Education and Science

Vol 39 (2014) No 175 41-57

Investigating Factors That Affect Turkish Students' Academic Success with Canonical Commonality Analysis According to PISA 2009 Results

Burhanettin Özdemir [1], Selahattin Gelbal [2]

Abstract

The aim of this study was to investigate relations between the factors related to Turkish students' reading abilities and the factors related to facilities that both students and families had, according to PISA 2009 results. The main reason that researchers do not employ canonical correlation analysis to determine relationship between the variables is that interpreting the results are difficult and complex. By calculating unique and common variance associated with variables in each variable sets, canonical commonality analysis helps to determine accurately degree of multicollinearity between the variables, suppressor variable (if there is) and related importance of variables in a canonical model. Thus, it helps researcher make more accurate and reliable interpretation. In this study, predictor variable set consists of factors related to facilities that students and family had and criterion variable set consists of factors related to students' reading abilities. Relationship between variable sets were investigated with canonical commonality analysis. As a result, predictor and criterion variable sets explained 31.7% of variance in students' academic success. In addition, according to commonality analysis, *utilizing information technologies while preparing homework* variable was a suppressor and there was a great multicollinearity between *facilities that students had at home* and socio-*economic status of families* variables.

Keywords

PISA

Factors affecting students' success

Canonical correlation

Commonality analysis

Article Info

Received: 01.13.2014

Accepted: 10.01.2014

Online Published: 11.10.2014

DOI: 10.15390/EB.2014.3025

Introduction

Education has a crucial role in the development of human resources, along with the individual's well-being and a better quality of life related to these facilities (Battle and Lewis, 2002). For educators, students' quality of performance has first priority for educators. This means to make changes at local, regional, national and international levels. International large-scale tests like PISA and TIMMS are used to reveal differences at the local, regional, national and international levels.

PISA Projects focus on 15 year-old students' reading, math and science abilities in OECD countries. İn addition, these projects aim to determine to what extend these students benefit from the skills that they obtained at the end of compulsory education. PISA not only tries to determine whether students use the skills learned in school environment, but also tries to determine whether students apply

[1] Hacettepe University, Faculty of Education, Department of Educational Sciences, Turkey, b.ozdemir@hacettepe.edu.tr
[2] Hacettepe University, Faculty of Education, Department of Educational Sciences, Turkey, s.gelbal@gmail.com

this skills within and outside of the school, and make prediction about unknown condition based on knowledge obtained in the school (MOE, 2007).

For each term (three-year intervals), PISA projects emphasize on only one of the main areas of reading, math and science literacy fields, as well as the other two fields are included in the scope of the evaluations. These three fields have been main concern once in a nine-year cycle. The project was based on reading literacy in 2000, while it was based on mathematical literacy in 2003 and science literacy in 2006. A new nine-year cycle began in 2009 and was focused on reading skills (PISA 2009 National Preliminary Report, 2010).

Educators and researchers were interested in exploring the factors that affect students' performance and the quality of education for a long time. Unfortunately, the quality of education is not a simple task to define and measure. The variability of quality criteria related to different perspectives of the stakeholders in this area increases the complexity of this process (Parra, 2006; Blevins, 2009).

Variables that affect student achievement have been discussed under the different titles. These variables may be classified as variables within school and outside of school. These factors also can be grouped as *student factors*, *family factors*, *school factors*, and *peer factors* (Crosno, Johnson and Elder, 2004).

Perelman and Santa (2011) examined factors such as students' opportunities, peer groups, personal characteristics and school characteristics that affect educational outcomes using data from PISA 2003. As a result, factors mentioned above and the student's school choice explained majority of differences between the achievements of students. Bender, Steel and Seferoğlu (2011) compared the results of PISA 2003 and 2009; and indicated that the main reason behind observed small increment was related to changes made in the educational system, the adoption of the constructivist approach and execution of the project relied on this approach.

According to literature review, studies that aimed to investigate the factors influencing students' reading achievement based on the PISA results varied in terms of methods adopted and the factors examined. Indeed, Shelley and Lightning (2013) investigated the factors that affected transformation of knowledge in the field of reading, science and mathematics based on PISA 2009 data set. It was found that students' mathematics and science achievement was a significant predictor of reading achievement. In addition, socio-economic and regional variables had a significant impact on the reading skills of students.

Perry and McConney (2013) studied the impact of school's socio-economic status on the student's math and reading achievement based on PISA 2006 data set and compared the results of Australia and Canada with respect to this factors. According to the results, the school's average socio-economic status had greater impact on math and reading skills compared to the impact of socio-economic status of students on math and reading skills. Indeed, different studies showed that beyond the students' socio-economic status, the schools' socio-economic status were positively related to educational outcomes (Palardy, 2008; Perry and McConney, 2010a, 2010b; Rumberg and Palardy, 2005; Southworth, 2010; Sui-Chu and Willms, 1996).

Chiu and McBride-Chang (2006) conducted a study to determine the factors that influenced students' reading skills and compared 43 different countries based on reading performance and the factors that affected reading performance of each country. Results of this study indicated that the factors such as family's socio-economic status, socio-economic status of the families of peers, number of books in a school and enjoying reading books variables showed a positive correlation with reading achievement.

Lee and Wu (2012) studied the relationship between reading printed texts and online electronic texts from the perspective of individual differences using information and communication technology (ICT) in a partial mediation model. For this purpose, they used the PISA 2009 data with 297,295 fifteen-year-old students across 42 countries. They found that availability of ICT at home was negatively associated with students' reading abilities.

Gürsakal (2012) conducted a study on determining the factors affecting students' math and science literacy and reading skills based on the results of PISA 2009. As a result, students 'reading achievement varied in terms of variables such as gender, school starting age, parents' education levels.

Apart from the studies mentioned above, different studies examined the factors affecting students' academic achievement and evaluated the performance of Turkey based on PISA results. (Alacacı & Erbas, 2010; Anil, 2009; Aypay, 2010; Iron & Sword, 2010; iron, Sword, & Unal, 2010, 2010b; Dincer & Uysal, 2010; Grisay & Monseur, 2007; Silver & Atalmıs, 2011; Beautiful & Berberoğlu, 2005; Beauties & Akin, 2011; Ovayol the & Kutlu, 2011; Unal & Anchor, 2009; Ziya Dogan, & Kelecioglu , 2010; Gürsakal, 2012).

In this study, in the light of information given above, factors affecting students' achievement were classified into two dimensions based on PISA 2009 results as "*variables related to students' reading skills*" and "*variables related to facilities of students and families*". Among the factors that assumed to affect students' achievement, the predictor variable set (variables related to facilities of students and families) consisted of *facilities that students had at home* (FACILITIES), *socio-economic status' of the families* (SES) and *utilizing information technologies while preparing homework* (INFO-TECH) variables, while the criterion variable set (variables related to students' reading skills) consisted of *students' reading scores* (READING SCORE) and *self-confidence of students while doing the senior task* (SELF-CONFIDENCE) variables. The relationship between these variable sets was examined by *canonical correlation* and *commonality analysis*.

Earlier studies that examined the factors affecting student achievement and reading skills did not take into account the interaction effect of these variables. However, ignoring the interaction of variables affects the reliability and validity of results. Therefore, in this study, the relationship between the variables related to students' reading skills and the variables that affect the reading skills; and unique and common effects of each factor in the model were examined by commonality analysis.

Purpose of Study

The aim of this study was to investigate relations between the factors related to Turkish students' reading abilities and the factors related to facilities that both students and families had based on PISA 2009 results by means of canonical commonality analysis. Since there has been few studies which aimed to discover the relationship between variable sets that are related to students' academic achievement in literature, it is assumed that this study would contribute the area of study.

This study was restricted with examining the factors that affect fifteen-year-old Turkish students' reading abilities according to PISA 2009 results. Definitely, there are many factors affecting students' reading abilities. Some of these factors were determined based on questionnaires in PISA 2009. Therefore, in this study, predictor variables were restricted with facilities *that students had at home, socio-economic status of the families* and *utilizing information technologies while preparing homework.* On the other hand, criterion variables were restricted with *students' reading scores* and *self-confidence of students while doing the senior task* variables. The students' math and science achievements and the factors that assumed to have effect on these factors in PISA 2009 were not included in the analysis.

In this study, it was assumed that all variables identified in PISA 2009 reflected the students' real time condition accurately. In addition, it was assumed that the tests and questionnaires applied in PISA 2009 managed to reveal students' reading abilities and self-confidence accurately.

Research Questions

In this study, the main goal was to determine the relationship between the predictor variable set which consisted of factors related to families and students' facilities and the criterion variable set which consisted of factors related to students' reading skills. Thus, the research questions were as follows:

1. What is the canonical correlation results of the model which was constructed with factors related to students' academic achievement?
2. What is the unique and common effect of factors related to families and students' facilities at home on factors related to students' reading abilities?
3. How do unique and common variance of factors affecting students' academic achievement differ based on commonality analysis?
4. What is the relative importance of factors within the predictor and criterion variable sets that consisted of variables related to students' academic achievement?
5. What is the most suitable model that can explain the relationship between the factors related to families and students' facilities at home and the factors related to students' reading abilities?

Method

Research Model

The model of this study was casual comparative research; since it aimed to investigate the relationship between the factors related to Turkish students' reading abilities which constituted the criterion variable set, and the factors related to facilities that both students and families had at home which constituted the predictor variable set based on PISA 2009 results by means of canonical commonality analysis.

Study Group

In this study, PISA 2009 Turkey data set was used in which 4,496 Turkish students from 15 year-old age group participated. Assumptions of canonical correlation analysis were checked before conducting the analysis. At the stage of checking the assumptions, two outliers related to SES variable, and seven outliers related to READING SCORE were excluded from the data set. In addition, mahallonobis distances related to variables in the model were calculated in order to determine multivariate outliers, and 85 multivariate outliers were excluded from the data set. As a result, the analysis was conducted on 4,902 participants of PISA 2009.

Data Analysis

In this study, SPSS 22 statistical software was used to calculate canonical correlation and commonality analysis (canonical commonality analysis). First of all, an SPSS syntax relevant to the model constructed was written in order to conduct the analysis. In addition, a macro file developed by Nimon (2010) was used to conduct commonality analysis. The syntax of canonical commonality analysis related to the model constructed was given as Appendix 1. At that point, it is assumed to be beneficial to give a more detailed information about canonical correlation and canonical commonality analysis.

Canonical correlation analysis is a novel technique used to examine theoretical and empirical relationships between two variable sets that are assumed to be related to each other (Capraro & Capraro, 2001). Allowing to analyze all variables simultaneously, canonical correlation analysis increases the validity of social science research and prevents Type I errors in univariate analysis (Thompson, 2000).

Canonical Commonality Analysis

Commonality analysis, which contributes to theory development, interpretation of research findings and interpretation of regression effect, has been applied in many disciplines such as social science research, education (e.g., Zientek & Thompson, 2006), counseling (e.g., Gill, Barrio Minton, & Myers, 2010), human resource development (c.f., Nimon, Gavrilova, & Roberts, 2010), behavioral science (e.g., Sorice& Conner, 2010; Nimon, 2010).

Commonality analysis helps researchers find out the contribution of each variable in a model by partitioning the explained variance into constituents (Zientek & Thompson, 2009). The contribution of each variable is determined by calculating how much of variance explained by variable itself (unique effect) and how much of variance explained by combination of other variables (common effect). Calculation of unique and common variance associated with each variable in commonality analysis was illustrated with a heuristic example in Figure 1. Let X_1 and X_2 be predictor variables, while X_0 be canonic variable of criterion variable set.

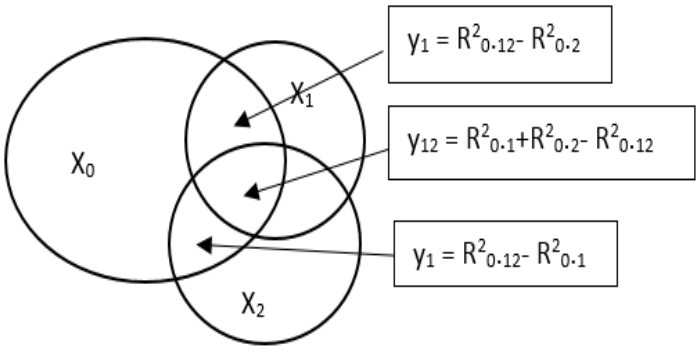

Figure 1. Venn Diagram of Unique (Y_1, Y_2) and Common Variance (Y_{12}) Components.

Figure 1 displays the variance of X_0 ($R^2_{0.12}$) which is explained by X_1 and X_2. This explained variance ($R^2_{0.12}$) consists of three components. These components are:

y_1 = unique variance related to X_1

y_2 = unique variance related to X_2

y_{12} = common variance related to X_1 and X_2 in $R^2_{0.12.}$

Where variance components are calculated by the formulas shown below;

$y_1 = R^2_{0.12} - R^2_{0.2}$

$y_2 = R^2_{0.12} - R^2_{0.1}$

$y_{12} = R^2_{0.1} + R^2_{0.2} - R^2_{0.12}$

$R^2_{0.1}$ *and* $R^2_{0.2}$, represent the sum of unique and common variance related to X_1 and X_2 respectively.

As illustrated in Figure 1, when the number of observed variables is equal to 2, commonality analysis (CA) yields two unique variances (y_1, y_2) and one common variance (y_{12}). As the number of variables increases, so does the number of the variance components calculated by CA. When the number of observed variables is equal to 3 and 4, then CA partitions the explained variance into 7 and 15 components respectively.

Results

This study consisted of two stages. At the first stage, canonical correlation analysis (CCA) was conducted in order to examine the relationship between the criterion variable set consisting of factors related to Turkish students' reading abilities and the predictor variable set consisting of factors related to facilities that both students and families had at home, and the results of analyses were interpreted. Then, significant canonical functions were selected based on CCA results and canonical variable sets obtained by means of these functions. At the second stage, commonality analysis was conducted in order to determine unique and common variance associated with each variable in both predictor and criterion variable sets, degree of multicollinearity between variables and suppressor variables which inflate the variance of other variables. Moreover, CA was conducted to select the best model that could be constructed with variables affecting students 'academic achievement.

Results of Canonical Correlation Analyses

Before conducting CCA, correlation coefficients related to variables within and between the predictor and criterion variable sets can be examined in order to get insight into the multicollinearity between the variables.

Table 1. Correlation Coefficients Related to Factors Affecting Students' Academic Achievement

		FACILITIES	SES
Correlation Coefficients Related to Predictor variables	FACILITIES	1.00	
	SES	.674	
	INFO-TECH	.384	.305
		READING SCORE	SELF-CONFIDENCE
Correlation Coefficients Related to criterion variables	READING SCORE	1.00	
	SELF-CONFIDENCE	.394	1.00
		READING SCORE	SELF-CONFIDENCE
Correlation Coefficients Related to variables between predictor and criterion variable set	FACILITIES	.397	.347
	SES	.454	.308
	INFO-TECH	-.034	.267

Table 1 displays correlation coefficients related to variables within and between the predictor and criterion variable sets. According to Table 1, FACILITIES and SES variables yielded the highest correlation coefficients (0.674). It is more likely to have collinearity between variables when the observed correlation between variables are equal or higher than 0.90 (Tabachnick & Fidell, 2007). The correlation coefficient shown in Table 1 indicates that there was no multicollinearity between the variables. However, these results do not provide any information about the degree of multicollinearity between the variables.

The number of canonical variable pairs and canonical correlations that can be calculated with CCA depends on the minimum number of variables in either the predictor or criterion variable set. In this study, the predictor variable set with variables related to students and families' facilities includes 3 variables, while the criterion variable set with variables related to students 'reading skills includes two variables. Thus, 2 canonical functions and 2 canonical correlations were calculated as a result of CCA. When canonical correlations were examined, the first canonical correlation (R_{c1}) related to the first canonical variable set was 0.563, and variance explained by the first canonical variable set was equal to 31.7 % (R_{c1}^2). On the other hand, the canonical correlation (R_{c2}) related to the other canonical variable set was 0.27 and the variance explained by the second canonical variable set was equal to 7.3 % (R_{c2}^2).

Table 2 presents Wilk's lambda and X^2 test statistics and also significance values of these statistics related to each canonical correlation pairs obtained from each variable sets.

Table 2. Significance Test Results of Canonical Functions.

	Wilk's	Chi-SQ	DF	Sig.
1. the first canonical function	.633	2241.608	6.000	.000
2. the second canonical function	.927	372.106	2.000	.000

According to the results shown in Table 2, the chi square value of the first canonical correlation pair was $X^2_{sd=6}$= 2241.608, and it was statistically significant (p= 0.000 <0.01). Likewise, the chi square value of the second canonical correlation pair was $X^2_{sd=6}$= 372.106, and it was also statistically significant (p= 0.000 <0.01). However, canonical correlation less than .30 is considered to be too small to interpret. Since it implies that the canonical variable set explains less than 10% of variance, it is not meaningful to interpret it (Capraro & Capraro, 2001). Thus, the findings related to the second canonical variable set were not interpreted. As a result, it is assumed that the first canonical variable set explained the relationship between the predictor and criterion variable sets.

Standardized canonical coefficients are used as coefficients of each observed variables in canonical functions that gives the canonical variable sets. Table 3 displays the standardized canonical coefficients and canonical loadings related to each variable in the predictor and criterion variable sets.

Table 3. Standardized Canonical Coefficients and Canonical Loadings Related to Each Variable

		Standardized canonical coefficients	Canonical loadings
Predictor variables	FACILITIES	-.525	-.863
	SES	-.636	-.918
	INFO-TECH	.236	-.160
Criterion variables	READING SCORE	-.836	-.898
	SELF-CONFIDENCE	-.444	-.561

The equations of the canonical variables, the significance of which were tested above, constructed by means of standardized canonical coefficients. The first canonical variable pair named as U_1 and V_1 were calculated with standardized canonical coefficients given in Table 3 as follows:

U_1= -0.525*FACILITIES -0.636*SES+ 0.236*INFO-TECH

V_1= -0.836*READING SCORE – 0.444*SELF-CONFIDENCE,

The other statistics that have to be interpreted are canonical loadings related to each observed variable within the predictor and criterion variable sets. Canonical loadings between canonical variables and observed variables are presented in Table 3. Tabachnick and Fidell (2007) suggest that the correlation coefficients and canonical loadings greater than 0.30 can be interpreted, and scale direction of the loading matrix should be taken into account while interpreting the canonical variables. Therefore, these conditions were taken into account while interpreting the canonical loadings.

The square of the canonical loadings of each variable gives the percentage of explained variance of the canonical variable belonging to the other variable set. Among the variables related to students' and families' facilities at home, FACILITIES (-.863) and SES (-.918) yielded large canonical loadings, while the INFO-TECH (-.160) yielded relatively small canonical loading. FACILITIES variable in predictor variable set explained 74.4% ($-.863^2$) of variance in V_1, while SES explained 84.2 % ($-.918^2$) of variance in V_1 which is the canonical variable associated with criterion variable set. The sum of the percentage of variance explained by these variables was larger than 100% which is an indicator of multicollinearity. However, this result does not provide information about the unique and common variance related to each variable.

Table 3 shows that the standardized canonical coefficient related to INFO-TECH was not equal to zero, and its canonical loading was relatively small. According to Pedhazur (1997), a variable serves as a suppressor variable, when it has a standardized canonical coefficient other than zero and a small canonical loading. However, this finding does not provide information about the degree of suppressor effect.

Table 3 also displays the canonical loadings associated with each variable in the criterion variable set. Among the variables related to students' reading abilities, READING *SCORE* (**-.898**) and SELF-CONFIDENCE (**-.561**) yielded large canonical loadings with the first canonical variable set. READING SCORE variable in the predictor variable set explained *80.6 %* (.8982) of variance in U_1, while SELF-CONFIDENCE explained 31.4 % ($-.561^2$) of variance in U_1 which is the canonical variable associated with the predictor variable set. Once again, the sum of the percentage of variance explained by these variables was larger than 100% which is an indicator of multicollinearity.

According to canonical correlation findings, the most important factor within the variable set related to the facilities that students and families had at home was the socio-economic status of families, and it was followed by students' facilities at home and using information technologies while doing homework. Likewise, the most important factor within the variable set related to students' reading abilities was the students' reading score, followed by students' self-confidence while taking part in senior tasks. Negative canonical loadings related to each factor within the predictor and criterion variable sets indicates that a decrease in a factor is highly related to the decrease in another factor. Thus, these negative values can be treated and interpreted as positive values.

To conclude, students' facilities at home and socio-economic status of families were highly correlated with students' reading scores and students' self-confidence while taking part in senior tasks. Therefore, as the facilities that students had at home and socio-economic status of families increased, so did the students' reading score and self-confidence. In addition, it was found that utilizing information technologies while doing homework did not have significant effect on students' reading skills.

Results of Canonical Commonality Analysis

Traditional CCA results are restricted with the interpretation of standardized canonical coefficients and canonical loadings. However, these findings does not provide enough information to uncover complex relationship between the variables in the model. Therefore, in this study, canonical commonality analysis was applied to the canonical variable set calculated by canonical functions. Findings of commonality analyses were displayed in Table 4, 5, 6 and 7, respectively.

Table 4 displays unique and common variance components of the variables within the criterion variable set. Total variance in Table 4 presents the variance explained in canonical variable (U_1) associated with students and families' facilities variable set.

Table 4. Variance Components Associated with Canonical Variable (U_1) of Predictor Variable Set

	Variables	Commonality coefficients (R^2)	Percentage (R^2 %)
Unique variance	READING SCORE	.217	68.53
	SELF-CONFIDENCE	.061	19.33
Common variance	READING SCORE & SELF-CONFIDENCE	.039	12.14
	Total	.317	100.00

Commonality coefficients represent the unique variance of each variable and common variance explained with combination of other variables. Moreover, each unique and common variance are divided by total variance explained by the model so that variance component can be expressed with percentages ($\frac{Ri2}{Rc2}\ x100$). Negative commonality coefficients related to variables implies that there is at least one predictor variable serving as suppressor variable, and this situation indicates that variables effect each other in a negative way (Pedhazur,1997).

All the unique and common variances associated with each variable in Table 4 were positive. Since there appeared to be no negative commonality coefficients, it can be concluded that there were no suppressor variable in the criterion variable set. According to the unique and common variance percentages presented in Table 4, the unique variance related to READING SCORE had the largest percentage(68.53%); and it was followed by the unique variance related to SELF-CONFIDENCE (19,33%) and common variance related to READING SCORE & SELF-CONFIDENCE (19.33%). Sum of commonality coefficients related to variables within criterion variable set was 0,317 which was equal to total variance explained by the model (R_c^2).

Table 5 displays the unique and common variance components of the variables within the criterion variable set. Total variance in Table 5 presents the total variance explained in canonical variable (V_1) associated with students' reading skills variable set.

Table 5. Variance Components Associated with Canonical Variable (V_1) of Criterion Variable Set

	Variables	Commonality coefficients (R^2)	Percentage (R^2 %)
Unique variance	FACILITIES	0.045	14.017
	SES	0.070	22.121
	INFO-TECH	0.016	4.933
Common variance	**FACILITIES & SES**	**0.195**	**61.465**
	FACILITIES & INFO-TECH	-0.010	-3.205
	SES & INFO-TECH	-0.004	-1.296
	FACILITIES &SES & INFO-TECH	0.006	1.965
	Total	**0.317**	**100.00**

According to the unique and common variance percentages presented in Table 5, the common variance related to SES and FACILITIES had the largest explained variance percentage (61.5%); and it was followed by the unique variance related to SES (22.12%) and the unique variance related to FACILITIES (14.01%). Unlike the large common variance related to SES and FACILITIES, small unique variance related to each FACILITIES and SES variable indicated that there appeared to be a large multicollinearity between students' facilities at home and families' socio-economic status. In other words, 65% of variance related to students' facilities at home and families' socio-economic status were in common. Therefore, it can be concluded that using only one of these variables in a model that aims to explain students' academic achievement, will provide more reliable results.

When the unique and common variance of INFO-TECH variable was examined, the unique variance of INFO-TECH was small and its common variances with FACILITIES and SES were negative. Normally, a variable cannot have a negative variance. As explained before, having negative common variances associated with variables indicates that there is at least one or more suppressor variables. In addition, in the presence of suppressor variable, unique variance of a variable can be greater than the total variance explained by this variable. The main reason behind this is negative variance components in the model. These findings imply that INFO-TECH served as suppressor variable in the model.

When the INFO-TECH variable was excluded from the model, unique variance associated with FACILITIES decreased from 14.01% to 10.8%. Likewise, unique variance associated with SES decreased from 22.12 % to 20.82%. These results imply that the role of *utilizing information technologies while doing homework* in *students' facilities at home* and *socio-economic status of families* had no significant effect on student reading abilities. In fact, the unique effect of students' facilities at home and socio-economic status of families on students' reading skills appeared to be larger than it really was, when INFO-TECH variable was added to the model with negative coefficients. Therefore, excluding the utilizing information technologies while doing homework variable will provide more accurate results.

Table 6 shows the unique variance related to each variable, the common variance explained with other variables and the sum of unique and common variance related to each variables that gives total variance associated with each variable. It is another way of displaying the commonality analysis results.

Table 6. Variance Components Related to Each Variable within Predictor and Criterion Variable Set

V_1 canonical variable				U_1 canonical variable			
Predictor variables	Unique variance	Common variance	Total variance	Criterion variables	Unique variance	Common variance	Total variance
FACILITIES	0.045	0.191	0.236	READING SCORE	.217	.039	.256
SES	0.070	0.197	0.267	SELF-CONFIDENCE	.061	.039	.100
INFO-TECH	0.016	-0.008	0.008				

Table 6 shows that the unique variance associated with FACILITIES was .045 and common variance explained by FACILITIES along with other variables was .191. The last column of Table 6 represents the sum of unique and common variance related to each variables which is equal to total variance associated with each variable ($R_i^2 = R_u^2 + R_c^2$). Therefore, total variance related to FACILITIES was equal to R_i^2= .236, and it was equal to sum of unique and common variance of FACILITIES (0.236=0.045+0.191). Likewise, the unique variance associated with READING SCORE variable in the criterion variable set was equal to .217 and common variance associated with READING SCORE and the other variables within the criterion variable set was .039. Therefore, total variance related to READING SCORE was equal to R_i^2= .256, which was equal to sum of unique and common variance related to it (0.256=0.217+0.039).

One of the most important aspect of commonality analysis which differs this method from the other methods is that it provides R^2 values of all possible sub-models that can be constructed with the variables within the predictor and criterion variable set which aims to predict canonical variables calculated by CCA. The main goal is to decide on the best variables that predict the canonical variable associated with the other variable set. Therefore, these findings help researcher decide on the variables and the number of variables given the variable set by examining R^2 values related to all possible models. Moreover, obtaining R^2 values related to all possible sub-models with one analysis make this method more advantageous over the other methods.

Table 7 provides explained variance values (R^2) of all possible sub-models that can be constructed with the variable within the predictor variable set which aims to predict canonical variable (V_1) calculated by CCA.

Table 7. R^2 Values of All Possible Sub-Models Related to Predictor Variable Set

Predictor variables	K	R^2
FACILITIES	1	0.235
SES	1	0.267
INFO-TECH	1	0.007
FACILITIES & SES	**2**	**0.301**
FACILITIES & INFO-TECH	2	0.247
SES & INFO-TECH	2	0.273
FACILITIES & SES & INFO-TECH	3	0.317

Table 7 presents the R^2 value of the model constructed with only SES variable, and it was equal to 0.267 which could be considered quiet large. When the predictor variable set consisted of FACILITIES and SES, explained variance associated with this model increased to 0.302 which was substantially close to the explained variance of the model including all three variables in the model (for K=3, R^2=0.317). When INFO-TECH was excluded from the model, the reduction in the explained variance was equal to 0.016 and relatively small (0.317 – 0.301=0.016). These findings imply that *utilizing information technologies while preparing homework* did not make significant contribution to the prediction of students' reading scores, and therefore could be excluded from predictor variable set. In addition, the small unique variance related to *utilizing information technologies while preparing homework variable* also supported this finding.

Table 8 provides explained variance values (R^2) of all possible sub-models that can be constructed with the variable within the criterion variable set which aims to predict canonical variable (U_1) calculated by CCA.

Table 8. R^2 Values of All Possible Sub-Models Related to Criterion Variable Set

Predictor variables	K	R^2
READING SCORE	1	.256
SELF-CONFIDENCE	1	.099
READING SCORE & SELF-CONFIDENCE	2	.317

Table 8 shows that the R^2 value (explained variance) of the model constructed with only READING SCORE variable, and it was equal to 0.256 which could be considered quiet large. The explained variance associated with this model increased to 0.317, when the predictor variable set consisted of READING SCORE and SELF-CONFIDENCE. Thus, one can infer from these findings that students' reading score and students' self-confidence while performing senior tasks factors increased the prediction power of the canonical variable (U_1) that represents the predictor variable set consisted of factors related to students' and families facilities.

Tablo 9. The Results of Canonical Commonality Analysis Associated with Model Constructed.

variable	β	r_s	r_s^2	Unique variance	Common variance	Total variance
READING SCORE	-.836	-.898		.2174	.0385	.2560
SELF-CONFIDENCE	-.444	-.561		.0613	.0385	.0998
R_c^2			0.317			
FACILITIES	-.525	-.863		0.045	0.191	0.236
SES	-.636	-.918		0.070	0.197	0.267
INFO-TECH	.236	-.160		0.016	-0.008	0.008

(Not: β= standardized canonical coefficients, r_s =canonical loadings, r_s^2 = explained variance)

Table 9 presents the results of both canonical correlation and commonality analysis together. According to some research, tabulation of findings in this format assumed to be more informative (Nimon, 2010). In addition to canonical correlation analysis findings, Table 9 also presents unique variance, common variance and total variance that is sum of unique and common variance associated with each variable. Moreover, it gives the percentage of each variable's variance in the total variance explained by the model.

Conclusion and Discussion

In this study, among the factors affecting students' academic achievement, relationship between the predictor variable set that consisted of factors related to students' and families' facilities at home and the criterion variable set that consisted of factors related to students' reading abilities was examined by canonical commonality analysis.

According to CCA results, the most important factor within the predictor variable set was socio-economic status of families; and it was followed by students' facilities at home and utilizing information technologies while performing senior tasks. Likewise, the most important factor within the criterion variable set appeared to be students' reading score, and it was followed by self-confidence of students while performing senior tasks. Standardized canonical coefficients related to utilizing information technologies while preparing for homework variable had values other than zero and canonical loading of it was relatively small. According to Pedhazur (1997), a variable serves a suppressor variable, when it has a small canonical loading and a standardized canonical coefficient other than zero.

According to CCA results, the students' facilities at home and socio-economic status of families factors were highly correlated with students' reading score and students' self-confidence while performing senior tasks factors. Therefore, one can infer from the results that as the students' facilities at home and socio-economic status of families increase, so do the students' reading score and self-confidence. Moreover, according to the commonality analyses results, the *utilizing information technologies while preparing for homework variable* served as suppressor variable and caused the predictive power of other predictor variables to appear higher than they really were. This results imply that the role of utilizing information technologies while preparing for homework within students' facilities at home and socio-economic status of families did not have significant effect on students' reading score. Moreover, the effect utilizing information technologies while preparing for homework variable should be controlled, while investigating relationship between the variables affection students' reading abilities.

According to CCA results, there appeared to be no collinearity between the variables, while commonality analysis indicated that there was a quite large amount of multicollinearity between the students' facilities at home and socio-economic status of families variables (65%). In this case, selecting only one of these two variables in a model that aims to explain students' academic achievement would yield more reliable and accurate results. In other words, in the presence of socio-*economic status of families* variable, *facilities that students had at home* variable could be excluded from the model. Chiu and McBride-Chang (2006) also found that that socio-economic status of families had a positive correlation with students' reading achievement along with other factors such as socio-economic status of the peers' families, number of books in a school and enjoying reading books.

Commonality analyses, conducted with variables affecting student reading performance, also gave R^2 value of all possible sub-models that could be constructed with the variables in the predictor and criterion variable set in order to predict canonical variables (U_1 and V_1) obtained from CCA. When the explained variance associated with each sub-models constructed with variables within the predictor variable set was examined, excluding *the utilizing information technologies while preparing for homework variable* caused a small reduction in the explained variance. Thus, *the utilizing information technologies while preparing for homework variable* can be excluded from the model, since it did not make a significant contribution to prediction of students' reading performance. Lee and Wu (2012) also found that using

information technologies was negatively correlated with students' reading skills. The main reason behind this result could be the risk that students might browse inappropriate content (Wolak, Finkelhor, Mitchell, &Ybarra, 2008; Wolak, Mitchell, & Finkelhor, 2007) or addiction to the online games and gambling (Grüsser, Thalemann, & Griffiths, 2006; Wan & Chiou, 2006), when they had free access to ICT.

As a result, only examining the standardized canonical coefficient in order to determine relative importance of variables in a model can lead to misinterpretation. To avoid this type of misinterpretation, CCA results should be reinforced with commonality analysis and the unique and common variance related to each variables should be taken into consideration.

In this study, relationship between variables affecting students' reading performance was investigated with canonical commonality analysis. Different factors and analysis methods can be used in order to uncover relationship between variables affecting students' reading performance and students' academic achievements. Moreover, different multivariate methods such as regression analysis, canonical correlation and variance analysis can be used together and results of this methods can be supported with commonality analysis results in the field of educational science.

The relationship between online reading and reading skills, which is beyond the scope of this study, can be examined with different factors based on PISA data, since the usage of information and communication technologies in education increases year by year. In addition, it is suggested that the reason behind the negative and positive effects of information technologies on different fields in education should be studied in more detail.

References

Adams, A. (1996). Even basic needs of young are not met. Retrieved from http://tc.education.pitt.edu/library/SelfEsteem

Alacaci, C., & Erbas, A.K. (2010). Unpacking the inequality among Turkish schools: Findings from PISA 2006.*International Journal of Educational Development, 30*(2), 182-192.

Anil, D. (2009). Factors affecting science achievement of science students in Program for International Students' Achievement (PISA) in Turkey. *Education and Science, 34* (152), 87-100.

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

Battle, J., & Lewis, M. (2002). The increasing significance of class: The relative effects of race and socioeconomic status on academic achievement. *Journal of Poverty*, 6(2), 21-35.

Blevins, B. M. (2009). Effects of socioeconomic status on academic performance in Missouri public schools. Retrieved from http://gradworks.umi.com/3372318.pdf

Capraro, R. M., & Capraro, M. M. (2001). Commonality analysis: Understanding variance contributions to overall canonical correlation effects of attitude toward mathematics on geometry achievement. *Multiple Linear Regression Viewpoints,* 27(2), 16-23.

Crosnoe, R., Johnson, M. K., & Elder, G. H. (2004). School size and the interpersonal side of education: An examination of race/ethnicity and organizational context. *Social Science Quarterly,* 85(5), 1259-1774

Chiu, M. M. & McBride-Chang, C. (2006) Gender, context, and reading: a comparison of students in 43 countries, *Scientific Studies of Reading*, 10:4, 331-362, DOI: 10.1207/s1532799xssr1004_1

Çelen, F. K. , Çelik, A. & Seferoğlu, S. S. (2011). Türk Eğitim Sistemi ve PISA Sonuçları. *Akademik Bilişim 2011, 2-4 Şubat 2011 / İnönü Üniversitesi, Malatya.*

Demir, I., & Kılıç, S. (2010). Usıng PISA 2003, examının the factors affectıng students' mathematıcs achıevement *H. Ü. Eğitim Fakültesi Dergisi* (*H. U. Journal of Education*), *38*, 44-54.

Demir, I., Kılıç, S., & Unal, H. (2010a). Effects of students' and schools' characteristics on mathematics achievement: Findings from PISA 2006. *Procedia Social and Behavioral Sciences, 2*(2), 3099–3103.

Demir, I., Kılıç, S., & Unal, H. (2010b). The effect of quality of educational resources on mathematics achievement: Turkish case from PISA-2006.*Procedia Social and Behavioral Sciences, 2,* 1855-1859.

Dincer, M.A., & Uysal, G. (2010). The determinants of student achievement in Turkey. *International Journal of Educational Development, 30,* 592–598.

Frederick, B. N. (1999). Partitioning variance in the multivariate case: A step-by-step guide to canonical commonality analysis. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 5, pp. 305-318). Stamford, CT: JAI Press.

Furstenberg, F. F., & Hughes, M. E., (1995). Social capital and successful development among at-risk youth. *Journal of Marriage and the Family*, 57, 580-592.

Gill, C. S., Barrio Minton, C. A., & Myers, j. E. (2010). Spirituality and religiosity: Factors affecting wellness among low-income, rural women. *Journal of Counseling & Development, 77*, 293-303.

Goddard, R. D. (2003). Relational networks, social trust, and norms: A social capital perspective on students' chances of academic success. *Educational Evaluations & Policy Analysis*, 25, 59-74.

Grisay, A., & Monseur, C. (2007). Measurıng the equıvalence of ıtem dıfficulty ın the varıous versıons of an ınternatıonal test. *Studies in Educational Evaluation, 33,* 69–86.

Grüsser, S. M., Thalemann, R., & Griffiths, M. D. (2006). Excessive computer game playing: Evidence for addiction and aggression? *Cyberpsychology & Behavior*, 10(2), 290–292.

Gumus, S., & Atalmıs, E.H. (2011). Explorıng the relatıonshıp between purpose of computer usage and readıng skılls of Turkısh students: Evıdence from PISA 2006.*The Turkish Online Journal of*

Educational Technology—TOJET, 10(3), 129-140.

Gunel, M. (2009). Writing as a cognitive process and learning tool in elementary science education. *Elementary Education Online, 8*(1), 201-211.

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

Güzel, I.Ç., & Berberoğlu, G. (2005). An analysis of the programme for international student assessment 2000 (PISA 2000) mathematical literacy data for Brazilian, Japanese, and Norwegian students. *Studies In Educational Evaluation, 31*(4), 283-314.

Güzeller, C.O., & Akın, A. (2011). An examination of the programme for international student assessment (PISA) 2003 Turkish database with the aim of exploring the relationship between homework variables and mathematics achievement. *Educational Research and Reviews, 6*(13), 793-803.

Henson, R. K. (2002, April). "The logic and interpretation of structure coefficients in multivariate general linear model analyses." *Paper presented at the annual meeting of the American Educational Research Association, New Orleans. (ERIC Document Reproduction Service No. ED467381)*

Lee, Y. H. & Wu, J. Y. (2012). The effect of individual differences in the inner and outer states of ICT on engagement in online reading activities and PISA 2009 reading literacy: Exploring the relationship between the old and new reading literacy, *Learning and Individual Differences* 22 (2012) 336–342

Mood, A. M. (1969). Macro-analysis of the American educational system. *Operations Research*, 17, 770–784.

Mood, A. M. (1971). Partitioning variance in multiple regression analyses as a tool for developing learning models. *American Educational Research Journal, 8, 191–202.*

Newton, R. G., & Spurrell, D. J. (1967). A development of multiple regressions for the analysis of routine data. *Applied Statistics,* 16, 51-64.

Nimon, K., & Gavrilova, M. (2010, February)."Commonality analysis: Demonstration of an SPSS solution for regression analysis". *Poster presented at the 2010 iConference, University of Illinois at Urbana-Champaign.*

Ovayolu, Ö., & Kutlu, Ö.(2011) The range of scores in competency clusters of Turkish students in mathematics sub-test according to PISA2006. *Procedia Social and Behavioral Sciences, 15,* 17–26.

Özenç, B., & Arslanhan, S. (2010, December). *An evaluation of the PISA 2009 results.* Ankara: Economic Policy Research Foundation of Turkey (Türkiye Ekonomi Politikaları Araştırma Vakfı).

Parri, J. (2006). Quality in higher education. *Vadyba/Management*, 2(11), 107-111.

Palardy, G. J. (2008). Differential school effects among low, middle, and high social class composition schools: A multiple group, multilevel latent growth curve analysis. *School Effectiveness and School Improvement,* 19, 21–49.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Ft Worth, TX: Harcourt Brace.

Perelman, D. & Santin, D. (2011). Measuring educational efficiency at student level with parametric stochastic distance functions: an application to Spanish PISA results. *Education Economics,* 19:1, 29-49, DOI: 10.1080/09645290802470475

Perry, L. B. & McConney, A. (2013). School socioeconomic status and student outcomes in reading and mathematics: A comparison of Australia and Canada. *Australian Journal of Education* 2013 57: 124. DOI: 10.1177/0004944113485836

Perry, L. B., & McConney, A. (2010a). Does the SES of the school matter? An examination of socioeconomic status and student achievement using PISA 2003. Teachers College Record, 112(4), 1137–1162.

Perry, L. B., & McConney, A. (2010b). School socio-economic composition and student outcomes in Australia: Implications for education policy. Australian Journal of Education, 54(1), 72–85.

Rumberger, R. W., & Palardy, G. J. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. Teachers College Record, 107(9), 1999–2045.

Seibold, D. R., & McPhee, R. D. (1979). Commonality analysis: A method for decomposing explained variance in multiple regression analysis. *Human Communication Research,* 363-355

Shelley, M. & Yildirim, A. (2013). Transfer of Learning in Mathematics, Science, and Reading among Students in Turkey: A Study Using 2009 PISA Data. *International Journal of Education in Mathematics, Science and Technology Volume 1, Number 2*, April 2013, 83-95 ISSN: 2147-611X

Sorice, M. G. & J. R. Conner. (2010). Predicting landowner intentions to enroll in an incentive program to protect endangered species. *Human Dimensions of Wildlife* 15(2):77-89.

Southworth, S. (2010). Examining the effects of school composition on North Carolina student achievement over time. *Education Policy Analysis Archives*, 18(29), 50–90.

Sui-Chu, E. H., & Willms, J. D. (1996). Effects of parental involvement on eighth-grade achievement. *Sociology of Education*, 69(2), 126–141.

Unal, H., & Demir, I. (2009). Divergent thinking and mathematics achievement in Turkey: Findings from the programme for international student achievement (PISA-2003). *Procedia Social and Behavioral Sciences,1*(1), 1767-1770

Thompson, B. (2000). Canonical correlation analysis. *In L. G. Grimm & P. R. Yarnold (Eds.), Reading and understanding more multivariate statistics (Vol. 1, pp. 192–196). Washington, DC: American Psychological Association.*

Zientek, L. R., & Thomspon, B. (2006). Commonality analysis: Partitioning variance to facilitate better understanding of data. *Journal of Early Intervention,* 28(4), 299-307.

Zientek, L. R., & Thompson, B. (2009). Matrix summaries improve research reports: Secondary analyses using published literature. *Educational Researcher, 38,* 343-352.

Ziya, E., Dogan, N., & Kelecioglu, H. (2010). What is the predicted level of which computer using skills measured in PISA for achievement in mathematics. *The Turkish Online Journal of Educational Technology-TOJET, 9*(4), 185-191

Wolak, J., Finkelhor, D., Mitchell, K. J., & Ybarra, M. L. (2008). Online "predators" and their victims: Myths, realities, and implications for prevention and treatment. American Psychologist, 63(2), 111.

Wolak, J., Mitchell, K., & Finkelhor, D. (2007). Unwanted and wanted exposure to online pornography in a national sample of youth Internet users. *Pediatrics*, 119(2), 247.

Appendix-1

SPSS Syntax of canonical correlation and canonical commonality analysis

```
*  Canonical correlation.sps (canonical correlation macro)
*commonality coefficients macro file http://profnimon.com/CommonalityCoefficients.sps Syntax
* Copy data file to working directory.
* Set to working directory.
CD "C:\Users\Burhan\Desktop\son".
* Get canonical correlation macro.
INCLUDE "Canonical correlation.sps".
* Get commonality coefficient macro.
INCLUDE "CommonalityCoefficients.sps".
* Get data file.
GET FILE="PISAOZ.sav".
* Run Canonical Correlation and create canonical variate scores.
CANCORR SET1= FACILITIES, ses, INFO-TECH/
SET2=basari, ozguven / .
* Save updated data file with canonical variate scores.
SAVE OUTFILE= PISAOZa.sav/COMPRESSED.
!cc dep=S1_CV1
db = PISAOZ a.sav
set= Predictor
ind= basari ozguven.
!cc dep=S2_CV1
db = PISAOZ a.sav
set= Criterion
ind= FACILITIES ses INFO-TECH.
* Commonality Data for Predictor Canonical Variate written to
* PredictorCommonalityMatrix.sav and PredictorCCByVariable.sav.
```