Using Structural Equation Modeling to Investigate Students’ Reading Comprehension Skills

Aylin ALBAYRAK SARI

ABSTRACT. In this study, Structural Equation Modeling was used to determine the factors that affect reading skills. To determine the factors data of PISA 2009 was used. It assessed students’ capacities to apply knowledge and skills in reading, mathematical and scientific literacy. It describes some wider findings about what lies behind results. Structural Equation Modeling is examined and four independent latent variables as “reading attitudes,” “study habits,” “stimulate,” and “strategies the teacher used” are determined. It was observed that the most important variable was the “strategies the teacher used” ($\gamma = 0.33$). The second important latent variable that affected the students’ reading comprehension skill was “teacher stimulating students” ($\gamma = 0.26$). Another latent variable affecting the students’ reading comprehension level was observed as “the students’ study habits” ($\gamma = 0.22$). The final latent variable was “attitude towards reading” ($\gamma = 0.16$).

Key Words: Reading Literacy, Reading Comprehension Skills, PISA 2009, Reading Strategies, Motivation

INTRODUCTION

Training individuals complying with the requirements of the rapidly developing age is among the basic objectives of all education systems. The knowledge, skills and proficiency of individuals forming society constitute the foundation of effective citizenship and social harmony. In this context, determination of the characteristics of students and monitoring of student performance on a national and international basis are important in terms of shaping the future of individuals and guiding them in accordance with desired characteristics.

The assessment and evaluation activities determining the levels of students for the purpose of determining the outputs of the education system of our country and directing education policies are executed both in national and international level. Determination of student performances in various areas also provides the opportunity to make interpretations regarding the performances of education systems.

One of the activities conducted in an international level is Programme for International Student Assessment (PISA), which Turkey has being participating since 2000. PISA is a program conducted by the Organization of Economic Cooperation and Development (OECD) in intervals of three years. It aims to evaluate the ability of students in the 15-year-old age group proceeding to formal education after compulsory education for using their knowledge and skills in circumstances they may encounter in the information society of today. The PISA exam is the most extensive and detailed international program evaluating student performance and collecting data on the factors of student, family and school in order to explain differences of student performance (Aşkar & Olkun, 2005). The most recent application of PISA was in 2009 and there was an emphasis on reading comprehension in the application.

The actual purpose of reading is to upskill correct and rapid comprehension to students (Akçamete, 1989). Reading comprehension skills are interpretations covering answering questions regarding information and details in the text and extracting the meaning of a word from its context, finding the main idea, making inferences, reasoning, generalizing, synthesizing, establishing cause and effect relations, and finding similarities (Boulware-Goode, Carreker, Thornhill & Joshi, 2007). Reading comprehension is generally considered one of the most important cognitive skills young students acquire during their school career and constitutes an important prerequisite for lifelong learning in adulthood (Alvermann & Earle, 2003; Mason, 2004).

1 Part of this work was presented at the 5th World Conference on Educational Sciences.

2 PhD., Hacettepe University, Faculty of Education, Department of Elementary Education, e-mail: aylinalb@hacettepe.edu.tr
The purpose of this study is to investigate the reading comprehension skills of students and the type of relation between the variables determined in the student questionnaire for Turkish students in the 15-year-old age group participating in the PISA 2009.

METHOD

Sample of Study
This study is a descriptive study conducted to reveal the existing status. The PISA 2009 application was conducted with the participation of 500,000 students in 74 participant countries in the world representing 28 million students in the 15-year-old age group. All students in the 15-year-old age group studying in the 8th grade are universe. The PISA 2009 sample selection was conducted randomly by utilizing the two-stage stratified sampling method by taking geographic regions and school types into account. 4996 students from 170 schools participated PISA 2009 in Turkey (OECD, 2012). The sample of this study consists of 4996 students.

Data Collection Tools
The data collection tools were the reading comprehension skill tests and the student questionnaires that were used in the PISA 2009. The data was taken from the official PISA website and OECD reports (OECD, 2012). There are different types of items in the PISA. These items consist of simple multiple choice (students mark one of the four or five options) or complex multiple-choice (students evaluated by selecting one of the two possible answers as "yes / no" or "agree / disagree"). The remaining items are short or long open-ended items asking to create students’ own answers (MEB, 2010).

Data Analysis
The study was conducted based on the analysis of PISA 2009 data through structural equation modeling (SEM). The reason for the selection of SEM in the study is to pair the available data with the propositions of the conceptual world in the mind of the researchers and the determination of the extent to which they are compatible with each other. The purpose of SEM is to investigate the set of relations between one or more independent variables and one or more dependent variables. SEM is an extensive statistics approach utilized to test models, in which there are causal and correlational relations together between observable and latent variables (Hoyle, 1995).

In this study, while independent variables were being determined, answers of students to the PISA 2009 student questionnaire were used. As a dependent variable, the achievement scores calculated from the answers to open ended and multiple-choice questions are prepared for the purpose of determining reading skills.

In the process of estimating the parameters and determining the significance of the model under the scope of the study; (1) a model was developed theoretically in line with the conducted literature review, (2) the dimensions of suitable questions taken from the student questionnaire were determined, (3) question groups those form the latent variables to be used in the SEM were selected, (4) the latent variables were assessed through the confirmatory factor analysis and (5) the fit indexes of the proposed model were assessed respectively. The principal components factor analysis was conducted in order to determine the dimensions of suitable questions taken from the student questionnaire and factors were determined according to factor loads (core values). The accuracy of the model constructed with the determined factors was checked with the SEM.

FINDINGS

The factor analysis was tested with the Kaiser-Meyer-Olkin (KMO) coefficient and the Barlett Sphericity test whether the data fit prior to conduct the factor analysis. The Barlett test is significant
if the determined value of KMO is higher than 0.60 (Büyüköztürk, 2011). In this study, KMO coefficient value was determined as 0.93 and it means that the data were fit for factor analysis. The internal consistency coefficients of factor loads, core values of factors and questions accumulated in the factors were examined. All questions in the questionnaire under the scope of the study were not taken into account and the questions with the most factor loads were included in the study. The most important criteria is the use of at least three questions in each dimension (Schumacker and Lomax, 1996). As a result of the factor analysis, it was determined that the core values of questions accumulated under 4 factors with a core value higher than 1, and these obtained results consist with the results described in the international report PISA 2009 (OECD, 2012). These factors and items of students’ questionnaire are presented in Table 1.

As shown in Table 1 third factor has negative factor loads. The reason of this result is these items have negative meanings. So these items were recoded before SEM analyses. The structural equation model constructed with variables determined with data obtained from the student questionnaire of students in 15-year-old age group in the PISA 2009 sample of Turkey has been provided in Figure 1.

As shown in Figure 1, four independent latent variables such as “reading attitudes,” “study habits,” “stimulate,” and “strategies the teacher used” are determined. The determined model was tested by employing maximum likelihood estimation method at a significance level of 0.05. Many Goodness-of-Fit indexes are used in order to assess model data fit. The most frequent used ones among those indexes may be classified under three groups such as Chi Square Goodness-of-Fit Tests (chi-square/sd), Descriptive Measures of Overall Model Fit (RMSEA) and Descriptive Measures Based on Model Comparisons (NFI, NNFI, CFI, GFI, AGFI) (Schermelleh-Engel and Moosbrugger, 2003).

The fit indexes obtained for model (chi-square/sd, RMSEA, NFI, NNFI, CFI, GFI, AGFI) were reviewed and it was observed that the model was significant at a significance level of 0.05. The “$\chi^2$” is determined as 3132.86 and “sd” is 291. So “$\chi^2 / sd$” coefficient is determined as 10.76. It was expected that rate of “$\chi^2 / sd$” was smaller than 5 but as the chi square value was relatively sensitive to the sample size, it was observed that this value generally appeared to be significant if the sample size is more than 200 so this coefficient is acceptable and the significance of chi square value indicates that the model fits data (Şimşek, 2007). As the sample size increased, the chi square value will be significant (Kline, 2005). The RMSEA and GFI values may be interpreted as the model indicates an acceptable fit and the NFI, NNFI, CFI, and AGFI values may be interpreted as the model indicates a perfect fit.

When the variables regarding the students’ reading comprehension skills were examined, it was observed that the most important variable was “strategies the teacher used (STRATEGIES).” The correlation coefficient was $\gamma=0.33$. The variable with the highest factor load among the components of “strategies the teacher used” latent variable was “the state of teacher motivating the students to ask questions.” This variable was followed by the observed variables such as “teacher’s explanation of what s/he expects of students” and “teacher’s discussion on students’ homework assignments after they are completed.”
### Table 1. Factor Loads Obtained According to the Factor Analysis

<table>
<thead>
<tr>
<th>Factors</th>
<th>Items in the Student Questionnaire</th>
<th>Factor Loads</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Strategies</td>
<td>Check Concentrating (The teacher checks that students are concentrating while working on the &lt;reading assignment&gt;)</td>
<td>0.712</td>
</tr>
<tr>
<td></td>
<td>Motivating questions (The teacher poses questions that motivate students to participate actively)</td>
<td>0.699</td>
</tr>
<tr>
<td></td>
<td>Discuss work (The teacher discusses students’ work, after they have finished the &lt;reading assignment&gt;)</td>
<td>0.691</td>
</tr>
<tr>
<td></td>
<td>Student questions (The teacher gives students the chance to ask questions about the &lt;reading assignment&gt;)</td>
<td>0.684</td>
</tr>
<tr>
<td></td>
<td>Explain judgments (The teacher tells students in advance how their work is going to be judged)</td>
<td>0.678</td>
</tr>
<tr>
<td></td>
<td>Immediate feedback (The teacher tells students how well they did on the &lt;reading assignment&gt; immediately after)</td>
<td>0.668</td>
</tr>
<tr>
<td></td>
<td>Ask if understood (The teacher asks whether every student has understood how to complete the &lt;reading assignment&gt;)</td>
<td>0.649</td>
</tr>
<tr>
<td></td>
<td>Explain Expectations (The teacher explains beforehand what is expected of the students)</td>
<td>0.564</td>
</tr>
<tr>
<td></td>
<td>Mark work (The teacher marks students’ work)</td>
<td>0.546</td>
</tr>
<tr>
<td>Stimulate</td>
<td>Express opinion (The teacher encourages students to express their opinion about a text)</td>
<td>0.686</td>
</tr>
<tr>
<td></td>
<td>Relate to lives (The teacher helps students relate the stories they read to their lives)</td>
<td>0.646</td>
</tr>
<tr>
<td></td>
<td>Time to think (The teacher gives students enough time to think about their answers)</td>
<td>0.641</td>
</tr>
<tr>
<td></td>
<td>Better understanding (The teacher asks questions that challenge students to get a better understanding of a text)</td>
<td>0.625</td>
</tr>
<tr>
<td></td>
<td>Recommend books (The teacher recommends a book or author to read)</td>
<td>0.615</td>
</tr>
<tr>
<td>Read Attitude</td>
<td>Hard to finish (I find it hard to finish books)</td>
<td>-0.586</td>
</tr>
<tr>
<td></td>
<td>Need information (I read only to get information that I need)</td>
<td>-0.585</td>
</tr>
<tr>
<td></td>
<td>Waste of time (For me, reading is a waste of time)</td>
<td>-0.575</td>
</tr>
<tr>
<td></td>
<td>Only if I have to (I read only if I have to)</td>
<td>-0.528</td>
</tr>
<tr>
<td>Study</td>
<td>Relate to Experience (When I study, I try to understand the material better by relating it to my own experiences.)</td>
<td>0.571</td>
</tr>
<tr>
<td></td>
<td>Figure Out (When I study, I start by figuring out what exactly I need to learn.)</td>
<td>0.568</td>
</tr>
<tr>
<td></td>
<td>Useful Outside School (When I study, I figure out how the information might be useful outside school.)</td>
<td>0.483</td>
</tr>
<tr>
<td></td>
<td>Memorize Details (When I study, I try to memorize as many details as possible.)</td>
<td>0.444</td>
</tr>
<tr>
<td></td>
<td>Important Points (When I study, I make sure that I remember the most important points in the text.)</td>
<td>0.443</td>
</tr>
<tr>
<td></td>
<td>Read Text Repeatedly (When I study, I read the text so many times that I can recite it.)</td>
<td>0.438</td>
</tr>
<tr>
<td></td>
<td>Memorize Everything (When I study, I try to memorize everything that is covered in the text.)</td>
<td>0.429</td>
</tr>
<tr>
<td></td>
<td>Additional Information (When I study and I don’t understand something, I look for additional information to clarify this.)</td>
<td>0.421</td>
</tr>
</tbody>
</table>
Figure 1 Structural Equation Model Regarding the Reading Comprehension Skill (Standardized Solution)

Table 2. Fit Criteria

<table>
<thead>
<tr>
<th>Fit Measures</th>
<th>Perfect fit</th>
<th>Acceptable fit</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\chi^2$/sd</td>
<td>$0 &lt; \chi^2$/sd $\leq 2$</td>
<td>$2 &lt; \chi^2$/sd $\leq 5$</td>
<td>10.76</td>
</tr>
<tr>
<td>RMSEA</td>
<td>$0 &lt; \text{RMSEA} &lt; 0.05$</td>
<td>$0.05 \leq \text{RMSEA} \leq 0.10$</td>
<td>0.044</td>
</tr>
<tr>
<td>NFI</td>
<td>$0.95 \leq \text{NFI} \leq 1$</td>
<td>$0.90 \leq \text{NFI} \leq 0.95$</td>
<td>0.97</td>
</tr>
<tr>
<td>NNFI</td>
<td>$0.97 \leq \text{NNFI} \leq 1$</td>
<td>$0.95 \leq \text{NNFI} \leq 0.97$</td>
<td>0.97</td>
</tr>
<tr>
<td>CFI</td>
<td>$0.97 \leq \text{CFI} \leq 1$</td>
<td>$0.95 \leq \text{CFI} \leq 0.97$</td>
<td>0.97</td>
</tr>
<tr>
<td>GFI</td>
<td>$0.95 \leq \text{GFI} \leq 1$</td>
<td>$0.90 \leq \text{GFI} \leq 0.95$</td>
<td>0.93</td>
</tr>
<tr>
<td>AGFI</td>
<td>$0.90 \leq \text{AGFI} \leq 1$</td>
<td>$0.85 \leq \text{AGFI} \leq 0.90$</td>
<td>0.92</td>
</tr>
</tbody>
</table>

(Source: Schermelleh-Engel & Moosbrugger, 2003)

It was observed that the second most important latent variable that affected the students’ reading comprehension skill was “the teacher stimulating the students (STIMULATE),” which was ranked after the latent variable of “strategies the teacher used” ($\gamma=0.26$). The variable with the highest factor load among the components of “the teacher motivating the students” latent variable was “the teacher motivating the students to explain their views.” This variable was followed by the observed variables as “the teacher providing sufficient time for students to think over their answers” and “the teacher asking questions to the students having difficulties to comprehend the text better.

It is observed that the strategies used by the teacher in the classroom and the teacher’s stimulation have a significant influence on the students’ achievement. This result is supported by the results of the studies performed on the subject (Aysan, Tannroğan and Tannroğan, 1996; Özler, 1998; Derman, 2002; Akdağ and Güneş, 2003, Anderman and Leake, 2005; Margolis and McCabe, 2006). Besides, Hattie (2009) performed a meta-analysis for the purpose of the elements affecting student achievement and reviewed 50,000 studies and 800 meta-analysis. The study stated that students’ achievement varies depending on genetics at the rate of 50%, teachers at the rate of 30%, and other factors at the rate of 20%.

Another latent variable affecting the students’ reading comprehension level was observed as “the students’ study habits (STUDY)” ($\gamma=0.22$). The variable with the highest factor load among the components of “students’ study habits” latent variable was “making sure to remember the most important points in the text while studying”. This variable was followed by “while studying, searching for additional information for clarification when the student did not understand a point” variable. The result of the study is also supported by the results of the studies (Geller and Ronald, 1986; Sübaşi, 2000; Brigman and Campbell, 2003) presenting that there was a positive relation between the students’ study habits and their academic achievements.

The final significant relation in the model is between “Attitude towards reading (READ)” latent variable and reading comprehension level (PV1READ) ($\gamma=0.16$). It was observed that “reading something only if the student is obliged to do” variable with the highest factor load ranked after “considering reading as a waste of time” variable. Many studies performed on the relation between attitude and achievement (Cooper, Lindsay and Nye, 1998; Saracalolu, 2000; Cecezolu, 2002; Gelbal, 2008; Karasakaloğlu and Saracalolu, 2009) argued that a positive attitude towards a course increases the achievement.

DISCUSSION and CONCLUSIONS

In this study an attempt was made to determine the relation between “teachers’ stimulation methods for students”, “students’ study habits”, “students’ attitudes towards reading” and “strategies the teacher used” variables that were obtained from the PISA student questionnaire and students’ reading skill. Therefore, a structural equation model was formed and the following recommendations are presented according to the findings:

516
The latent variables that explain the students’ reading comprehension skills are “the strategies the teacher used”, “teacher’s motivation for students”, “students’ study habits” and “their attitudes towards reading”, according to the order of importance.

It was observed in the study that strategies used by teachers in class and their encouragements to stimulate the students had a positive impact on students’ achievement. Therefore, teacher training program and their communications with students are subjects that need to be paid attention and activities should be performed in order to increase the teachers’ qualifications in terms of in-service trainings.

A variable that improves students’ reading comprehension scores is the frequency of constructive and supportive “strategies that teacher used”. The strategies for improving student comprehension may include providing students with background knowledge or experience, providing practical homework and class work based on reading. Scores of reading comprehension skills are increased by homeworks based on reading given by the teacher (Lei, et.al. 2010). Previous studies indicated that the strategies the teacher used that improves students' reading comprehension skills (Birch & Ladd, 1997; Taylor, Pearson, Peterson & Rodriguez, 2003; Lei, Rhinehart, Howard & Choi, 2010; Maslowski, Scheerens & Luyten, 2007; Yıldırım, 2012; Taşdelen Teker, Boztunç Öztürk ve Eroğlu 2014).

Another variable that affect students’ achievement is the “teacher’s motivation for students”. The teacher's behaviors, teaching specific methods and techniques used in lessons, positive interactions with students, encouraging them have a major effect on students’ motivation. Tuckerman (2003) offers that students, who are not motivated well, can not learn effectively and using specific motivation strategies and the structure of the course affect students’ achievement. These methods can strengthen students’ academic abilities.

The other variable is the students’ study habits. It can be said that students’ efficient study habits improve positive academic achievement. Credé and Kuncel (2008) had examined the effect of students’ study habits on academic performance and found the relationship to be positively significant and offer that study habit and skill measures improve prediction of academic performance more than any other noncognitive individual difference variable examined to date. Trainings should be provided to gain students both effective and efficient study skills and increase their achievement levels and self-confidence. Thus, the learning period will be shortened and the skills of memorizing and utilizing information will be developed.

Bloom (1979) states that student’s attitudes towards the course and his/her self-perception have important impacts on the achievement. It is not impossible to change the attitudes even though it is difficult. Teachers and families have great responsibilities on students’ acquisition of positive attitudes towards courses. It should be recommended that the guidance services at primary education schools conduct activities to acquire positive attitudes towards the reading course.
REFERENCES


Öğrencilerin Okuduğunu Anlama Becerilerinin Yapsal Eşitlik Modellemesi İle Kestirilmesi

Aylin ALBAYRAK SARI3

ÖZET


Bulgular: Faktör analizi yapılmadan önce, verilerin faktör analizine uygun olup olmadığını Kaiser-Meyer-Olkin (KMO) katsayısı ve Bartlett Sphericity testi ile kontrol edilmişdir (Büyüköztürk, 2006). Bu sonucu göre KMO katsayı değeri 0.93 bulunmaktadır. Bununla bu değer 0.60 değerinden yüksek ve Bartlett testi anlamlandırdığında, verilerin faktör analizine uygun olduğunu belirlemiştir. 

3 PhD., Hacettepe University, Faculty of Education, Department of Elementary Education, e-mail: aylinalb@hacettepe.edu.tr
SRMR (Standarized Root Mean Squared Residual) ve RMSEA (Root Mean Square Error of Approximation) uyum indeksleri değerlendirilmiştir. Bu uyum indekslerinden GFI ve AGFI için 0.90 ve üstündeki değerler verilen modele çok iyi uyduğunu; SRMR ve RMSEA uyum indeksleri için 0.10 ve altındaki değerler verilen modele iyi uyduğunu, 0.05 ve altındaki değerler verilen modele çok iyi uyduğunu göstermektedir (Steiger, 1990). Analiz sonucunda öğrencilerin okuma becerilerini açıklayan en önemli değişkenin “stratejiler” (öğretmenlerin öğretimde kullandığı stratejiler) olduğu belirlenmiştir ($\gamma=0.33$). İkinci en önemli değişken “motivasyon” (öğretmenlerin öğrencileri motive etmesi) olanak bulundu ($\gamma=0.26$). Üçüncü örtük değişken “çalışma” (öğrencilerin çalışma alışkanlıkları) ($\gamma=0.22$) ve okuma becerilerini en az açıklayan dördüncü örtük değişken ise “okuma” (okumaya karşı tutum) ($\gamma=0.16$) olarak hesaplanmıştır.