

PAPER DETAILS

TITLE: Investigating Factors Affecting Scientific Literacy with Structural Equation Modeling and Multilevel Structural Equation Modeling: Case of PISA 2015

AUTHORS: Eda AKDOGDU YILDIZ,Mehmet Can DEMIR,Selahattin GELBAL

PAGES: 795-824

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/1752543>



Investigating Factors Affecting Scientific Literacy with Structural Equation Modeling and Multilevel Structural Equation Modeling: Case of PISA 2015

Eda AKDOĞDU YILDIZ^{a*} (ORCID ID - 0000-0003-4374-4379)

Mehmet Can DEMİR^a (ORCID ID - 0000-0001-7849-7078)

Selahattin GELBAL^b (ORCID ID - 0000-0001-5181-7262)

^aBartın University, Faculty of Education, Bartın/Türkiye

^bHacettepe University, Faculty of Education, Ankara/Türkiye



Article Info

DOI: 10.14812/cufej.933101

Article history:

Received 05.05.2021

Revised 11.04.2022

Accepted 18.07.2022

Keywords:

Scientific Literacy,
Single-level Analysis,
Hierarchical Data,
Multi-level Analysis.

Research Article

Abstract

There is no empirical evidence in the literature regarding the problems encountered in the use of single-level analyzes on hierarchical data and the implementation of a single-multilevel structural equation model. In this study, the models were created by using Structural Equation Modeling and Multilevel Structural Equation Modeling for the effects of factors such as enjoyment in learning science, instrumental motivation, scientific self-efficacy, hinderances in education, and hinderance to learning which are claimed to predict Turkish students' science performance who participated PISA 2015. The effects of the predictive variables were estimated with two different single-level models constructed by aggregating and disaggregating the data. Then, single-level models are compared with the two-level model in terms of model fit and standardized parameters. As a result, since it was observed that standard error in regression coefficients decreased for the model which disregarded group levels, and variance-within-groups was not included in the model which disregarded individual levels which caused a data loss, the results were biased, and the effectiveness of the statistical test was weakened. In light of the results of this study, some recommendations were suggested for future studies which may consider dealing with analyzing hierarchical data.

Fen Okuryazarlığını Etkileyen Faktörlerin Tek ve Çok Düzeyli Yapısal Eşitlik Modeli ile İncelenmesi: PISA 2015 Örneği

Makale Bilgisi

DOI: 10.14812/cufej.933101

Makale Geçmişi:

Geliş 05.05.2021

Düzeltilme 11.04.2022

Kabul 18.07.2022

Anahtar Kelimeler:

Fen Okuryazarlığı,
Tek Düzeyli Analizler,
Hiyerarşik Veri,
Çok Düzeyli Yapısal Eşitlik Modeli.

Araştırma Makalesi

Öz

Hiyerarşik veriler üzerinde tek düzeyli analizlerin kullanımı ile tek ve çok düzeyli yapısal eşitlik modelinin uygulanmasında karşılaşılan sorunlara ilişkin literatürde ampirik bir kanıt bulunmamaktadır. Bu çalışmada, Türkiye’de PISA 2015 uygulamasına katılmış bireylerin fen başarısını yordadığı düşünülen fen öğrenmekten zevk alma, fen öğreniminde araçsal güdülenme, fen öz yeterliği, eğitim sürecindeki engeller, öğrenme engeli değişkenlerinin etkisi tek düzeyli ve çok düzeyli yapısal eşitlik ile modellenmiştir. Yordayıcı değişkenlerin etkileri, verilerin toplanması ve ayrıştırılması ile oluşturulan iki tek düzeyli model ile kestirilmiş ve model uyumu ile standartlaştırılmış parametreler açısından iki düzeyli model ile karşılaştırılmıştır. Sonuç olarak grup düzeyi göz ardı edilen modelde regresyon katsayılarına ait standart hataların azalmasından, birey düzeyi göz ardı edilen modelde ise grup içi varyans analize dâhil edilmediğinden ve veri kaybı yaşanmasından dolayı yanlış sonuçlar elde edilmiş ve istatistiksel testin gücünü azaltmıştır. Bu sonuçların, gelecekte hiyerarşik verilerde yapılacak çalışmalarda kullanılacak analizler için araştırmacılara fikir sunması beklenmektedir.

Introduction

Science is an organized human activity in which knowledge about the world is sought by the use of methods such as describing, observing, and experimental investigation. Moreover, it is especially a manner of discourse which expresses information in writing (Fang & Wei, 2010; Yore et al., 2004). This definition emphasizes the importance of correctly interpreting science, in other words; scientific literacy. One of the main objectives of scientific literacy is to enable individuals to make communications together with other people on their understandings of science, procedures, and research, and thus making conscious decisions and applying them (Yore et al., 2004). Accordingly, scientifically literate individuals are expected to solve problems and take decisions by using scientific processes and skills, acknowledge the importance of science in human life and thus be willing to explore the world and the universe while also being aware of the conditions that may interfere these explorations (Anagün, 2011; Organisation for Economic Co-operation and Development [OECD], 2015). OECD (2017) emphasizes the importance of scientific literacy by stating that scientific literacy plays a role in solving environmental, medical, economic, etc. problems. OECD (2017) also stated that the bringing-up of students who have high scientific literacy can affect the future high-tech sector and overall international competitiveness of the country of which they are a citizen.

In order to raise scientifically literate individuals, studies are conducted to identify scientific literacy rates after the necessary educational steps are taken. By measuring to what extend individuals are scientifically literate, contributions are made to planning education. For this reason, one of the internationally realized assessments is Programme for International Student Assessment (PISA hereafter) by Organisation for Economic Co-operation and Development (OECD hereafter). Scientific literacy of the students participating in PISA is defined as “being an active citizen by tackling with ideals and endeavours related to science” (OECD, 2017). In addition to individuals’ sociodemographic features, affective factors also affect this skill (Gardner, 1975). Accordingly, many studies were conducted on the variables affecting scientific literacy of individuals (Acosta & Hsu, 2014; Hanrahan, 1999; Schroeder et al., 2007). Many studies which aimed to investigate the factors affecting scientific literacy in Turkey were conducted by using data from PISA (Bilican Demir & Yıldırım, 2021; Döş & Atalmış, 2016; Kaya & Doğan, 2017; Taş et al., 2016). Among the variables affecting scientific literacy, which are the subject of this study, there are enjoyment of learning science, instrumental motivation for learning science, scientific self-efficacy, hinderance in teaching, and hinderance in learning.

According to the results of the previous studies, the effects of enjoyment of learning science (Anıl, 2009; Bussie're et al., 2007; Woods-McConney et al., 2013), instrumental motivation for learning science (Woods-McConney, et al., 2013; Yu, 2012), scientific self-efficacy (Acar & Öğretmen 2012; Bussie're et al., 2007; Woods-McConney et al., 2013), hinderance in teaching (Beese & Liang, 2010; Greenwald et al., 1996; Özkan et al., 2018) and hinderance in learning (Barutçu Yıldırım & Demir, 2020; Török et al., 2018) on scientific literacy were found to be significant. To clarify these variables, enjoyment of learning science can be stated as the interest or expressed liking for learning science; instrumental motivation for learning science can be stated as the attitude towards the perception that the course will contribute to future education and career; scientific self-efficacy can be stated as the belief in students' ability to use their science skills effectively in real life and to cope with difficulties; hinderance in teaching can be stated as the conditions that hinder the school's capacity to provide instruction (e.g., lack of teaching stuff, lack of educational staff, or lack of physical infrastructure), and hinderance in learning can be stated as the hinderance of the learning of students due to various behaviors related to teachers and students (e.g., student truancy, lack of respect for teachers, or teacher absenteeism). These variables, which are also discussed in the literature, are not always obtained from a single level. To put it more clearly, these variables obtained from PISA are obtained from students, school administrators, teachers, and parents. In this case, due to the use of some data sets together, the data set takes on a gradual feature. For example, in studies investigating the variables affecting student achievement, a hierarchical structure is formed because the variables obtained from school administrators and students have effects on an individual-level dependent variable. Data structure in a study has an effect on the data analysis method and the results. When the studies which

investigate the variables affecting scientific literacy are examined, it can be observed that although the data obtained from PISA tests were hierarchical, some studies used single-level analyses (Özer & Anıl, 2011; Usta & Çıkrıkçı Demirtaşlı, 2014). Because single-level analyses of variance, regression, factor, and structural equation modeling (SEM hereafter) depend on the independence of observations one another, using them may cause difficulties when the data is hierarchical (Çoker, 2009). When using single-level analyses for hierarchical data problems such as obtaining biased estimations or losing power in tests of significance by disregarding within group variation may occur due to making obligatory decisions in choosing unit of analysis, change in the sensitivity of measurement, and methodological constraints in estimations of parameters (Draper, 1995; Kaplan & Elliott, 1997; Raudenbush, 1995). Consequently, it is suggested that multi-level analyses are used when conducting a study with hierarchical data (Raudenbush & Bryk, 2002).

Before the discovery of multi-level analyses, researchers needed to disaggregate and aggregate the dataset to conduct a single-level analysis (Heck, 2001). For example, a researcher who aims to examine individuals' satisfaction with life in each state could only conduct analyses on a group level for aggregation regarding states, and on an individual level for aggregation in regard to the number of individuals. The results for both approaches are only valid when either the effects do not differ between groups or the affiliations to the group (states) are independent in all measurements. However, these conditions are mostly never met, which causes difficulties (Raudenbush & Bryk, 2002). The statistical and conceptual difficulties faced while approaching data hierarchic data with SEM led the way to the emergence and development of multilevel modelling methods (Heck & Thomas, 2015).

Multilevel modelling provided some advantages in comparison to single-level modelling methods. These advantages can be listed as a) imbalanced or loss data can be easily addressed, b) structural relations between different levels of the hierarchy can be determined and measured, c) hypotheses in different levels of the analysis can be tested, d) the validity of the suggested hierarchic structure can be assessed through likelihood or Bayes tests, e) standard errors are estimated correctly, not lower than the actual value, f) many modelling approaches can be used (Raudenbush & Bryk, 2002). Considering the advantages it brings, multi-level models are expected to be preferred in hierarchical data.

Multilevel models are beneficial and necessary to the extent of which the analyzed data show sufficient variation on all levels. Dividing variance into between and within group components provides evidence for making the decision to use a multilevel analysis or not (Heck & Thomas, 2015). Determining homogeneity between units in the same group emphasizes the dependence of the observations to the group. Increase in the degree of homogeneity, in other words the intraclass correlation, makes it necessary to conduct multilevel analyses (Can et al., 2011) and it is calculated by the following Equation 1 (Heck & Thomas, 2015):

$$\rho = \sigma_b^2 / (\sigma_b^2 + \sigma_w^2) \quad (1)$$

As seen in Equation 1, intraclass correlation (ρ) is equal to between group variance (Farmer, 2000). Generally, in the case that intraclass correlation is larger than .05, it is recommended that multileveled models are used (Dyer et al., 2005). However, in the case that it is larger than .70, it can be said that the groups are homogeneous; thus, the analyses are needed to be made on group level by using aggregation (Noyan, 2009).

Just as any other analysis, there are some limitations to multilevel modelling. One disadvantage of this modelling is that it renders examining indirect relationships between variables insufficient. In other words, analyses started to be made using structural equation modelling which has more than one level and enable conducting simultaneous tests on each level (individual and group levels) of hierarchic data (Uzun & Çokluk-Bökeoğlu, 2019). Multilevel structural equation modelling (MLSEM hereafter) provides the advantages of latent variables which serves as a direct mediator to aggregated data (Mehta & Neale, 2005).

MLSEM is also known as multilevel latent variable modeling (Kaplan et al., 2009), multilevel covariance structure analysis (Muthén, 1994; Yuan & Bentler, 2007), and multilevel covariance component models (Goldstein, 1987) in the literature. Because the variables are identified in between or within group levels in MLSEM which is composed of structural modelling components and measuring model, covariance matrix for data obtained from individuals are also disaggregated into levels. Model regarding the covariance structure of MLSEM is given in Equation 2 (Muthén, 1994):

$$\begin{aligned}\Sigma_T &= \Sigma_w + \Sigma_B \\ \Sigma_B &= \lambda \Psi_B \lambda' + \Theta_B \\ \Sigma_w &= \lambda \Psi_w \lambda' + \Theta_w\end{aligned}\quad (2)$$

As seen in Equation 2, covariance structures (Σ_T) in MLSEM are divided into within group (Σ_w) and between groups (Σ_B) variance components. The equation shows the λ factor loadings which identify between groups components, Ψ covariance matrix, and Θ margin of error covariance matrix (Depaoli & Clifton, 2015). For further information and related notations regarding estimation of covariance structures in multilevel structural equation models, please see Muthén (1994), and Heck and Thomas (2015).

Based on this information, and as the researchers have stated before, multilevel datasets continue to be handled with single level analyses, and this may lead to misleading results. Observation of these practices in studies conducted with the data of the country of origin of the researchers created a need for researchers to provide evidence for the application of analyzes that should be used in the analysis of multilevel data. For this reason, model comparison in order to present empirical evidence in the PISA 2015 Turkey data set constituted one dimension of this study. In the other dimension of the study, the aim was to reveal the effects of the variables that are thought to affect scientific literacy with the results obtained from the appropriate data analysis method. In line with these dimensions, this study aimed to create models by using SEM and MLSEM to investigate the relationship between enjoyment of learning science, instrumental motivation, scientific self-efficacy, hinderances in education, hinderance to learning variables, and scientific literacy scores of 15-year-old Turkish students who took PISA 2015. In line with these aims, this study sought to answer the following research questions:

1. To what extent do students' enjoyment of learning science, instrumental motivation, scientific self-efficacy, hinderances in education, and hinderance to learning variables account for their scientific literacy scores?
2. What is the level of structural equation relationship between enjoyment in learning science, instrumental motivation, scientific self-efficacy, hinderances in education, and hinderance to learning variables when compared to multilevel structural relationship?

Method

This chapter of the study presents information about the design of the study, demographic features of the sampling and population, data collection tools, and analysis of the data.

In this study, the relationship between enjoyment in learning science, instrumental motivation, scientific self-efficacy, hinderances in education, and hinderance to learning variables, and scientific literacy scores of 15-year-old Turkish students who took PISA 2015 was examined using structural equation modelling. Since examining the relationship between several variables was the aim, this study employed a correlational design (Büyüköztürk et al., 2014; Fraenkel et al., 2012). Concurrently, since the aim was also to describe the current situation by conducting SEM and MLSEM with hierarchic data, it is a descriptive study (Büyüköztürk et al., 2014).

Population and Sampling

The population of this study is the group of 1.324.089 students in Turkey who were between the age of 15 years, 3 months, and 16 years, 2 months. In PISA 2015 tests, the individuals to be taken into

sampling were identified in two steps. In the first step, the students were selected in consideration with various groupings (education type, school type, location, and administrative manner of the school) through random stratified sampling, and in the second step, students to participate in the test were selected through random sampling. All 5.895 students selected by PISA from 187 different schools in Turkey constitute the sample of this study. Distribution of these students regarding many variables are given in Table 1.

Table 1.
Distribution of the Sample for PISA 2015 in Turkey

	Class Grade												Age	
	7 th		8 th		9 th		10 th		11 th		12 th		16	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Female	6	0.2	42	1.4	494	16.8	2272	77.3	120	4.1	4	0.1	2938	100
Male	10	0.3	63	2.1	779	26.3	2036	68.9	66	2.2	3	0.1	2957	100
Total	16	0.3	105	1.8	1273	21.6	4308	73.1	186	3.2	7	0.1	5895	100

Data Collection Tools

Mathematics, science and reading comprehension of the students are assessed by PISA triennially. In order to explain the variance in students' performance in the test, data were also collected through questionnaire administered to students, family, and schools. In addition to demographic features of the individuals, data regarding the psychological nature were also collected through the questionnaires.

The data collection tools of this study were OECD's cognitive domain test for scientific literacy, and questionnaires which were administered to the students and schools. The results of the cognitive domain test were used as students' scientific literacy scores. In the questionnaire administered to students; 'enjoyment in learning science', 'instrumental motivation towards learning science', 'scientific self-efficacy' scales were used, and in the questionnaire administered to schools; 'hinderances in education', and 'hinderance to learning' scales were used. An increase in the result of the scales means that the specified characteristics also increases in the group.

Cognitive domain test: To measure the students' science performance, sub-tests of the cognitive domain test were administered rather than the test as a whole; thus, the scientific literacy scores of the students were obtained by estimating the likely answers they may give to all the questions. In PISA 2015, 10 plausible values (PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE, PV6SCIE, PV7SCIE, PV8SCIE, PV9SCIE, PV10SCIE) were determined in relation to scientific literacy by OECD. The reliability of the test for Turkish context was determined as Cronbach's $\alpha = .97$.

Scale for enjoyment of learning science: In PISA 2015, five items (ST094Q01NA, ST094Q02NA, ST094Q03NA, ST094Q04NA, ST094Q05NA) were used to determine the students' level of enjoyment of learning science. The internal consistency reliability for the scale for Turkish context was determined as Cronbach's $\alpha = .95$.

Scale for instrumental motivation for learning science: This scale was used to determine the perception of students towards the benefit of science courses they take in school for their career plans and education. The scale consists of 4 items (ST113Q01TA, ST113Q02TA, ST113Q03TA, ST113Q04TA). The reliability of the scale for Turkish context was determined as Cronbach's $\alpha = .90$.

Scale for scientific self-efficacy: This scale was used to determine the students' level of perception for their own knowledge and skills related to science in real world conditions. The scale includes 8 items (ST129Q01TA, ST129Q02TA, ST129Q03TA, ST129Q04TA, ST129Q05TA, ST129Q06TA, ST129Q07TA, ST129Q08TA). The reliability of the scale for Turkish context was determined as Cronbach's $\alpha = .89$.

Scale for hinderance in teaching: Teaching staff, assisting staff, educational materials, deficiency in physical infrastructure determine the hinderance in teaching, and to 8 items in the scale determine this variable (SC017Q01NA, SC017Q02NA, SC017Q03NA, SC017Q04NA, SC017Q05NA, SC017Q06NA,

SC017Q07NA, SC017Q08NA). The reliability of the scale for Turkish context was determined as Cronbach's $\alpha = .89$.

Scale for hinderance in learning: This scale identifies the hinderances in learning such as absence, negligence, etc. of teachers or students. The scale consists of 10 items (SC061Q01TA, SC061Q02TA, SC061Q03TA, SC061Q04TA, SC061Q05TA, SC061Q06TA, SC061Q07TA, SC061Q08TA, SC061Q09TA, SC061Q10TA). The reliability of the scale for Turkish context was determined as Cronbach's $\alpha = .83$.

Data Analysis

Before conducting SEM and MLSEM in line with the purpose of the study, data were initially cleansed and organized. After combining scientific literacy scores and questionnaires for students and schools, a Missing Completely at Random (MCAR; Little, 1988) test was conducted. The results of the test showed that missing data were not completely lost randomly ($\chi^2=1582.327$, $df=980$, $p<.05$). Expectation-maximization algorithm was chosen as the method to deal with data loss. After the missing data analysis, the group sizes were determined. Due to group sizes for sample of Turkey in PISA 2015 tests being larger than 100, secondary level groupings (schools) with sampling fewer than 10 for each group were subtracted from the dataset (Hox et al., 2017, p. 216). As a result of the data cleansing, data from 5800 individuals in 161 schools were analyzed. After the cleansing process, the assumptions were tested through linearity, homogeneity of variances, normality of residual tests. To establish the measurement model, it was determined that the variances for residuals at the individual level were homogenous ($p=.548$), there was a random pattern in the figure drawn between model residuals and dependent variables, and the residuals were normally distributed.

Before establishing the structural models, some arrangements were made to dataset to test SEM. Two separate datasets in which disregarded group level by disaggregation and individual level with aggregation were arranged. These steps were then followed by the generating process of structural equation models.

Arrangement of the data and the analyses regarding the reliability of measurement instruments were conducted on SPSS 23.0 software. The assumptions of the study were tested on RStudio, with 'lme4' package (Bates et al., 2015), and the model analyses regarding the research questions were conducted on RStudio with 'MplusAutomation' package (Hallquist & Wiley, 2018). In model analyses, MLF method in which the estimation of standard error is based on normal chi-square statistics and first derivatives (Heck & Thomas, 2015). The R functions used in data analyses are presented in Appendix 1.

Models

In line with the aim and research questions of the study, SEM and MLSEM were created. The indicators for the variables in the models are given in Table 2 below.

Table 2.
Indicators of the Variables in the Model

Variable	Indicator	Variable	Indicator
PV1SCIE	fb1	PV9SCIE	fb9
PV2SCIE	fb2	PV10SCIE	fb10
PV3SCIE	fb3	Science performance (Scientific literacy)	fb
PV4SCIE	fb4	Enjoyment in learning science	mza
PV5SCIE	fb5	Instrumental motivation for learning science	mag
PV6SCIE	fb6	Self-efficacy in science	moy
PV7SCIE	fb7	Hinderance in education	mee
PV8SCIE	fb8	Hinderance in learning	moe

Two-level structural equation model (MLSEM)

Two-level structural equation model which was created in line with the aim and research questions of the study is given in Figure 1. Continuous variables from fb1 to fb10 in the within group level of the structural equation model are explained by fb latent variable. Since they represent random intercepts, the continuous variables are shown in circles. In both levels, the variables are connected to the latent variable; however, fb latent variable is explained by various variables in within and between group. For within-groups level, mza, mag, moy and for between-groups level, mee, moe are the predicting variables which explain the variation in latent variable. The model tests were named as ‘Model 1’.

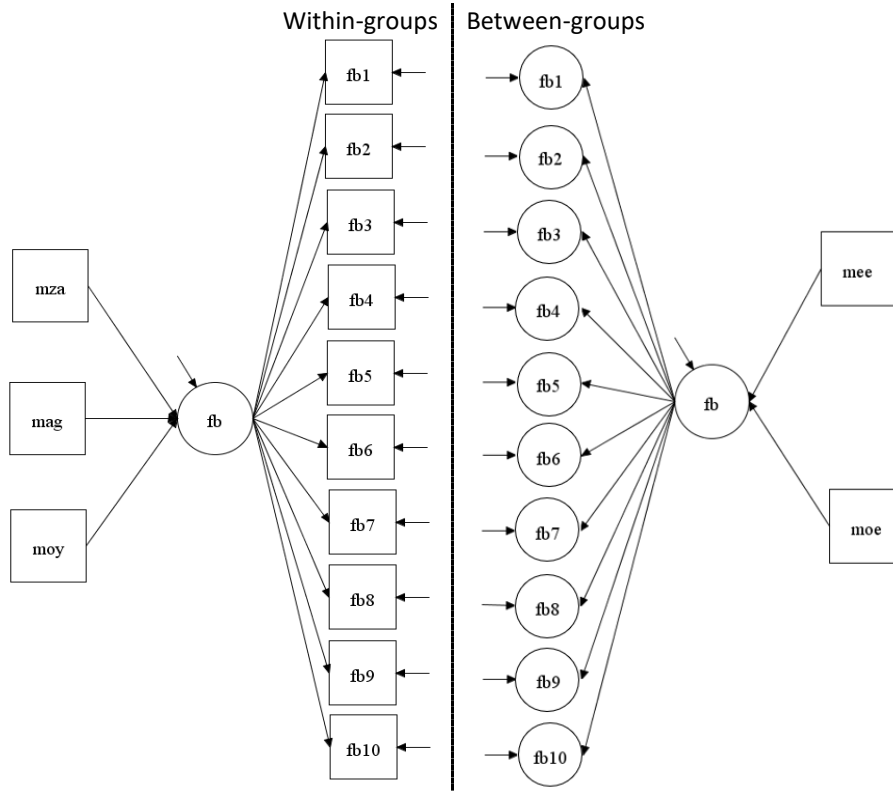


Figure 1. Two-levelled Structural Equation Model that was Tested (Model 1)

Structural equation model (SEM)

SEM which was created in line with the aim and second research question of the study is given in Figure 2. The variables mza, mag, moy, mee, and moe which explain fb latent variable in the two-levelled model explain science performance in SEM. The model which disregarded the group level by disaggregation was named ‘Model 2’, and the model which disregarded individual level by aggregation was named ‘Model 3’.

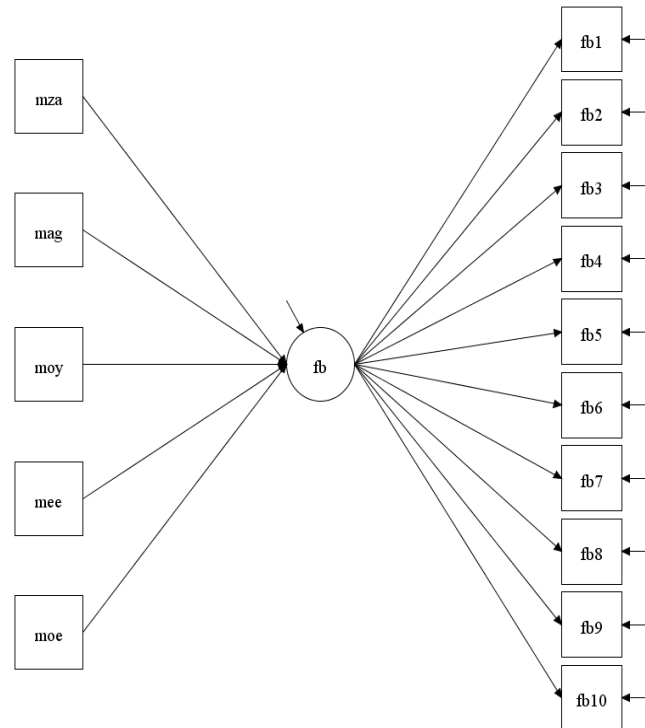


Figure 2. Structural Equation Model that was Tested (Model 2 – Model 3)

Findings

Intraclass correlation coefficients, which need to be higher than .05 to conduct multi-level model analyses, are given in Table 3 below.

Table 3.
Intraclass Correlation Coefficients

	Variables observed									
	fb1	fb2	fb3	fb4	fb5	fb6	fb7	fb8	fb9	fb10
Intraclass correlation coefficients	.475	.478	.483	.489	.495	.488	.476	.485	.488	.491

It can be seen in Table 3 that correlation coefficients of the 10 observed variables regarding science performance latent variable are between .475 and .495. Coefficients being larger than .05 and smaller than .70 indicates that the individuals show homogeneity in the group; however, it also indicates that it does not cause group homogeneity. According to these correlation coefficients, 48-49% of the total variance observed in variables which represent science performance are caused by difference in schools. In other words, while 48-49% of variance in students' science performance (fb1-fb10) can be explained by differences between schools, 51-52% can be explained by differences in students being taught within the schools. In the light of these findings, it was determined that conducting SEM for this specific dataset may cause biased parameter estimations, and thus, there was a need to create a multi-levelled structural equation model. Since this study aims to provide empirical evidence for the effect of the use of SEMs on model-fit indexes and parameter estimations when multi-levelled analyses are needed, both SEM and MLSEM were used. Model 1 is a two-level model, Model 2 which was created by using data made by disaggregation, and Model 3 which was created by using data made by aggregation were structural equation models.

Fit indices for two-levelled (Model 1) and single-levelled (Model 2 & 3) models which were created in line with the aim of this study are given in Table 4.

Table 4.
Model Fit Indices

	χ^2	χ^2/sd	RMSEA	CFI	TLI	SRMR
Model 1	155.701	1.353	.008	.999	.999	.004 ⁱ .006 ^a
Model 2	130.025	1.625	.010	.999	.999	.004
Model 3	139.928	1.749	.068	.990	.988	.006

i: within groups a: between groups

As seen in Table 4, SRMR index is .004 and .006 for between groups and within groups, respectively, and other indices are calculated for both levels. The χ^2/sd ratio which rates data-model fit being smaller than 2, RMSEA value being equal or smaller than .06, and CFI and TLI values being equal or larger than .95 indicate that the data shows a good fit to the model (Hu & Bentler, 1999; Tabachnick & Fidell, 2012). When fit indices for Model 1, 2, and 3 are examined, it can be stated that the data shows perfect fit to the model. However, when the fit indices of the models are compared to each other, Model 1 relatively shows a better fit. Factor loadings, regression coefficients, and standardized estimates for error variances for SEMs and MLSEM are given in Table 5.

Table 5.
Standardized Parameter Estimates

Results for two-levelled structural equation model (Model 1)					Results for structural equation models (Model 2 & 3)				
Level/ Variables	β	Std. Err.	t	p	Model/ Variables	β	St. Error	t	p
Within group					Model 2				
fb BY					fb BY				
fb1	.868	.004	194.573	.000	fb1	.934	.002	487.255	.000
fb2	.881	.004	240.905	.000	fb2	.941	.002	529.818	.000
fb3	.872	.004	238.275	.000	fb3	.937	.002	514.149	.000
fb4	.875	.004	233.119	.000	fb4	.939	.002	519.848	.000
fb5	.872	.004	230.641	.000	fb5	.938	.002	523.595	.000
fb6	.873	.004	237.037	.000	fb6	.937	.002	509.248	.000
fb7	.876	.004	239.565	.000	fb7	.938	.002	521.884	.000
fb8	.872	.004	248.895	.000	fb8	.935	.002	496.687	.000
fb9	.875	.004	217.448	.000	fb9	.939	.002	512.162	.000
fb10	.875	.003	252.746	.000	fb10	.939	.002	522.679	.000
fb ON					fb ON				
mza	.161	.016	10.235	.000	mza	.165	.013	13.164	.000
mag	.000	.019	0.024	.981	mag	.002	.013	0.128	.898
moy	-.077	.016	-4.811	.000	moy	-.101	.013	-7.791	.000
					mee	-.250	.013	-18.558	.000
					moe	-.111	.014	-7.871	.000
Error Variance					Error Variance				
fb1	.246	.008	31.749	.000	fb1	.127	.004	35.586	.000
fb2	.223	.006	34.596	.000	fb2	.114	.003	34.232	.000
fb3	.240	.006	37.546	.000	fb3	.123	.003	35.924	.000
fb4	.234	.007	35.700	.000	fb4	.119	.003	34.950	.000

fb5	.240	.007	36.422	.000	fb5	.119	.003	35.518	.000
fb6	.238	.006	37.075	.000	fb6	.121	.003	35.168	.000
fb7	.233	.006	36.354	.000	fb7	.120	.003	35.713	.000
fb8	.240	.006	39.305	.000	fb8	.125	.004	35.519	.000
fb9	.234	.007	33.304	.000	fb9	.119	.003	34.647	.000
fb10	.234	.006	38.638	.000	fb10	.118	.003	35.055	.000
fb	.964	.006	168.566	.000	fb	.848	.009	93.581	.000
Between Groups					Model 3				
fb BY					fb BY				
fb1	.998	.001	767.421	.000	fb1	.993	.002	600.224	.000
fb2	.998	.001	1069.992	.000	fb2	.994	.001	726.220	.000
fb3	.997	.001	896.949	.000	fb3	.993	.001	732.420	.000
fb4	.997	.001	883.131	.000	fb4	.994	.001	750.668	.000
fb5	.997	.001	919.439	.000	fb5	.994	.001	841.999	.000
fb6	.996	.001	855.469	.000	fb6	.992	.002	622.269	.000
fb7	.997	.001	1102.504	.000	fb7	.994	.001	834.062	.000
fb8	.994	.002	540.576	.000	fb8	.990	.002	492.825	.000
fb9	.996	.001	729.800	.000	fb9	.991	.002	535.639	.000
fb10	.997	.001	1004.562	.000	fb10	.993	.001	765.790	.000
fb ON					fb ON				
mee	-.339	.106	-3.215	.001	mza	.318	.082	3.899	.000
moe	-.159	.118	-1.344	.179	mag	.039	.081	0.486	.627
					moy	-.282	.077	-3.677	.000
					mee	-.227	.084	-2.686	.007
					moe	-.062	.088	-0.701	.483
Error Variance					Error Variance				
fb1	.005	.003	1.906	.057	fb1	.014	.003	4.404	.000
fb2	.004	.002	2.262	.024	fb2	.012	.003	4.362	.000
fb3	.007	.002	3.041	.002	fb3	.014	.003	5.341	.000
fb4	.006	.002	2.688	.007	fb4	.013	.003	4.890	.000
fb5	.005	.002	2.310	.021	fb5	.011	.002	4.785	.000
fb6	.007	.002	3.189	.001	fb6	.016	.003	5.009	.000
fb7	.005	.002	2.860	.004	fb7	.013	.002	5.365	.000
fb8	.013	.004	3.421	.001	fb8	.019	.004	4.896	.000
fb9	.007	.003	2.704	.007	fb9	.017	.004	4.709	.000
fb10	.007	.002	3.416	.001	fb10	.013	.003	5.214	.000
fb	.813	.072	11.348	.000	fb	.601	.071	8.428	.000

When the factor loadings for Model 1 is examined, it is seen that within group factors are between .878 and .881, and between group factor loadings are between .994 and .998. For Model 2 which was established with data obtained from individuals, and Model 3 established with data from groups, the loadings were between .934 and .941, and between .990 and .994, respectively. Considering this information, for both two-levelled model and the SEMs, it was determined that factor loadings for individual levels were smaller than group level factor loadings. Since error variance and factor loadings have an inverse relationship, error variance for individual level in all groups were observed to be larger than group levels. It was determined that all error variances except f1 variable, and all factor loading values in all models were significant. Thus, although there is variance for error and factor loadings for both SEM and MLSEM, these variances were not significant.

When the regression coefficients which predict science performance are examined in Model 1, it can be determined that among enjoyment in learning science, instrumental motivation for learning science,

and scientific self-efficacy variables, instrumental motivation for learning science was not determined as a significant predictor within groups. Moreover, among the predicting variables, one unit of increase in standard deviation in enjoyment in learning science will cause .161 increase for standard deviation in predicted science performance of students, and one unit of increase in standard deviation in scientific self-efficacy will cause .077 decrease. Predicting variables within groups explain 4% of the variance in latent variable. Hindrance in learning were not significant predictors between groups. It was also determined that one unit of increase for standard deviation in hinderances in education will cause .339 decrease in standard deviation. Predicting variables between groups explain 19% of the variance in latent variable.

When the regression coefficients which predict science performance are examined in Model 2, it can be determined that among enjoyment in learning science, instrumental motivation for learning science, scientific self-efficacy, hindrance in education and hindrance in learning variables, instrumental motivation for learning science was not determined as a significant predictor, either; however, in contrast to the two-levelled model, hindrance in learning is significant along other predictors. Furthermore, one unit of increase in standard deviation in enjoyment in learning science will cause .165 increase for standard deviation in predicted science performance of students, one unit of increase in standard deviation in scientific self-efficacy will lead .101 decrease, one unit of increase in standard deviation for hinderances in education will produce .250 decrease, and one unit of increase in standard deviation for hinderances in learning will cause .111 decrease for standard deviation. Predicting variables in Model 2 explain 15% of the variance in latent variable.

When Model 3 is examined, it can be seen that the significance level of regression coefficients level of enjoyment in learning science, instrumental motivation for learning science, and scientific self-efficacy variables are the same as two-levelled model. One unit of increase in standard deviation in enjoyment in learning science will cause .318 increase in predicted science performance, one unit of increase in standard deviation in scientific self-efficacy will cause .282 decrease, one unit of increase in standard deviation for hinderances in education will cause .227. Predicting variables in Model 3 explain 40% of the variance in latent variable.

Discussion, Conclusion & Recommendations

This study examined variables which are estimated to predict science performance in two-levelled structural equation model between schools and among students. In addition, this study offers empirical evidence regarding the use of structural equation models with hierarchic data.

It was determined through intraclass correlation coefficient that variables observed to affect science performance significantly vary between groups. It was determined that 48-49% of the variance observed in variables which represent science performance (fb1-fb10) were caused by difference in schools, and 51-52% were caused by differences in students at the same school. The two-levelled structural equation model, which was created after establishing the fact that the difference between schools were larger than 5%, was validated in accordance with fit indices.

When the variables in the model were examined, it was determined that the variables of instrumental motivation and hindrance in learning science learning did not significantly predict scientific literacy. Yetişir et al. (2018) found similar results with the results of this research in their study, as they examined the effect of instrumental motivation on scientific literacy. However, in the study conducted by OECD (2016), it was concluded that instrumental motivation was a significant predictor of scientific literacy. İlhan (2015) stated that if a student believes that the information he/she learns in the lesson will be beneficial in his/her life and career, the student's effort towards this lesson can increase (as cited in Güngör, 2019). However, the effect of instrumental motivation on achievement varies according to culture (Kjærnsli, & Lie, 2011; Yu, 2012). For this reason, there may be differences between the results of studies conducted with different countries and studies in Turkish culture. As stated, the effect of instrumental motivation depends on the perceived effect of learned information on career plans. The fact that individuals in Turkish culture have difficulty in making career planning due to the

changes in the country's education system may have changed the effect of instrumental motivation on scientific literacy. However, there is a limited number of studies examining the effect of instrumental motivation on scientific literacy in Turkish culture. Therefore, more studies are needed to be able to use strong statements about the prediction level of this variable. Moreover, no study was found in the literature examining the variable 'hinderance in learning' and this study found out that it is not a predictor of science performance the results indicate a significantly negative relationship between this variable and science performance. Many studies determined the negative effects of absence of teacher or students, drug use among students, and bullying among students which also constitute hinderance in learning variable in this study (Finlayson, 2009; Hanson et al., 2003; Konishi et al., 2010). Accordingly, it is recommended that future studies investigate the effects of instrumental motivation and hinderances in learning variables on science performance. However, in this study, hinderances in learning were analyzed using an index variable. The index variable defined by PISA includes many factors related to hinderances in learning. These factors include barriers originating from both students and teachers. Therefore, since data from both students and teachers of a particular sample will be under the same variable, obstacles or supports arising from the teacher or student group may be neutralized by the other group. With this point of view, it is recommended that researchers examine the effects of instrumental motivation and hinderances in learning on science performance.

It was determined that enjoyment in learning science had a low positive effect on science performance. One unit of increase in students' level of enjoyment in learning science would result in .161 standard deviation increase. In the report published by OECD (2016), it was stated that this variable had an indirect effect on science performance. Moreover, Yetişir et al. (2018) also concluded that the variable had a positive effect. Therefore, it is recommended that students are provided with learning environments in which they have the opportunity to enjoy, be content, and actively participate in learning science.

The effect of scientific self-efficacy on science performance was low and negative; meaning that one unit of increase in students' self-efficacy would cause .077 standard deviation decrease. Nevertheless, in many studies, it was reported that scientific self-efficacy had a positive effect on science performance (Aktamış et al., 2016; Kirbulut, & Uzuntiryaki-Kondakci, 2019; Uğraş, 2018) and even, the best predicting variable for science performance (Bircan, 2015; Uzun et al., 2010). Moreover, Bandura (1994) stated that self-efficacy is an important factor that affects an individuals' performance level. This finding of the study could not be supported by the related literature. However, in some studies, it is stated that there is no relationship between science self-efficacy and scientific literacy (Jamil & Mahmud, 2019). It is also stated that there may be mediating variables involved in the effect of self-efficacy on literacy (Doménech-Betoret et al., 2017). Based on these studies, it can be said that more studies are needed to examine the relationship between science self-efficacy and scientific literacy.

Another variable in the models which was thought to predict science performance was hinderances in education. Hinderances in education, which determined the level of adequacy in number of teachers at school, physical conditions, and resources had moderate level negative effect on science performance. One unit of increase in hinderances in education would cause .339 standard deviation increase for science performance. This result showed a general alignment with the literature which puts forward that hinderances in education is systematically related to students' performance, and this relationship carries significance (Greenwald et al., 1996; Hedges et al., 1994). Conversely, however, Hanushek (1997) emphasized that there was no relationship between students' performance and school resources. Similar to the studies in the literature, the research in this study found out that removing hinderances in education was proven to increase performance for students, and providing opportunities will support students in education. Thus, accommodating students' education-learning environments will yield better performance for them (Al Şensoy & Sağsöz, 2015). In line with this finding, lack of teachers, information technology materials, library, laboratory, and physical infrastructure should be eliminated, and the qualifications of the teachers should be increased.

In addition to examining the predicting effect of variables on science performance, this study also sought to answer the question ‘To what extent do the structural equation models and multi-level structural equation model which included variables differ for individual and school groups?’. Before comparing the parameter and fit indices for the models, intraclass correlation coefficients of the observed variables were examined, and it was determined that multi-levelled analyses were needed for the data. After examining the studies which made use of hierarchic data in the literature, and conducting aggregation and disaggregation, structural equation models were compared to multi-level structural equation model. Although difference in factor loadings in standardized parameters, error variances, and standard errors of these parameters was insignificant and at a low level for SEM and MLSEM, it was seen that there were significant differences in regression coefficients and standard errors. While hinderances in learning did not significantly predict science performance in regression coefficient obtained in Model 1 (two-levelled), it predicted science performance in Model 2 (disregarded group levels). This was caused by the high level of homogeneity among the students in schools, and thus; standard errors in regression coefficients created biased estimations in predictions (Heck, 2001; Hox et al., 2017; Muthén & Satorra, 1995). In Model 1 and Model 3 (disregarded individual groups), it was seen that although similarities occurred in significance for regression coefficients, the effects of within group predictors on latent variable were higher than they actually were. This was caused by not including within group variances into the analysis (Heck, 2001; Kaplan & Elliot, 1997). Moreover, a significant increase was observed for within group predictors. The reason for this finding was that the decrease in the efficiency of statistical test due to the data loss occurred during transforming individual level data into group level (Heck, 2001; Kaplan & Elliot, 1997). The results obtained from the three models indicated that use of two-levelled models were more appropriate for hierarchic data in terms of model fit indices. Results and conclusions obtained from comparison of analyses in the study presented ideas and suggestions for future research. In summary, using single level analyzes in hierarchical data yields inaccurate results regarding the relationships between the variables and this may lead to inaccurate interpretations and inferences.

Comparisons in this study were limited to the data obtained from Turkey for the science cognitive subdomain and student and school administrator questionnaires. In future research, more extensive comparisons can be made by including reading and/or mathematics cognitive subdomains, different countries and teacher questionnaires in the analysis.

Author Contribution Rates

The authors contributed equally to the study.

Ethical Declaration

The research has no unethical problems and the research and publication ethics have been fully observed. The ethics committee approval for present research was given by Bartın University Social and Humanities Ethics Committee with the protocol number 2020-17 and authors declare that the principals of research and publication ethics were followed.

All rules included in the “Directive for Scientific Research and Publication Ethics in Higher Education Institutions” have been adhered to, and none of the “Actions Contrary to Scientific Research and Publication Ethics” included in the second section of the Directive have been implemented.

Conflict of Interest

The authors declare no conflict of interest.

Türkçe Sürümü

Giriş

Fen, doğal dünya hakkında bilgiyi tanımlama, gözlem yapma, deneysel inceleme gibi bilimsel yöntemleri kullanarak arayan örgütlü bir insan etkinliğidir. Ek olarak, elde edilen bilgiyi özellikle yazılı dili kullanarak ifade eden bir söylem biçimidir (Yore vd., 2004; Fang & Wei, 2010). Bu tanımlama ile doğada var olan bilimi doğru şekilde okumanın ve yazı diline aktarmanın yani fen okuryazarlığının da önemi vurgulanmaktadır. Fen okuryazarlığının temel amaçlarından biri, bireylerin diğer insanlarla araştırmalar, prosedürler ve bilim anlayışları hakkında iletişim kurabilmesini ve böylece bilinçli kararlar alarak uygulamalarını sağlamaktır (Yore vd., 2004). Dolayısıyla fen okuryazarı olan bireylerin günlük yaşamında fene dair kavram, yasa ve ilkeleri kullanabilmesi, bilimsel süreç ve becerileri kullanarak problem çözmesi ve karar alması, fenin insan yaşamındaki önemini fark etmesi ve bu sebeple dünyayı ve evreni keşfetmeye istekli aynı zamanda keşiflerine engel olabilecek durumların farkında olması beklenmektedir (Anagün, 2011; Ekonomik İşbirliği ve Kalkınma Örgütü [OECD], 2015). OECD (2017) fen okuryazarlığının önemini çevresel, tıbbi, ekonomik vb. sorunların çözümünde fen okuryazarlığının rol aldığını belirterek vurgulamaktadır. OECD (2017) ayrıca, fen okuryazarlığı yüksek bireylerin yetişmesinin, öğrencilerin vatandaşları oldukları ülkenin gelecekteki ileri teknoloji sektörünü ve genel uluslararası rekabet gücünü etkileyebileceğini belirtmiştir.

Fen okuryazarı bireyler yetiştirebilmek için gerekli eğitim süreçlerinin ardından okuryazarlık düzeylerini belirlemeye yönelik çalışmalar yapılmaktadır. Bireylerin fen okuryazarı özelliklerini ne düzeyde taşıdıkları ölçülerek eğitim süreçlerinin planlamasına katkı sağlanmaktadır. Bu amaçla uluslararası düzeyde gerçekleştirilen değerlendirme çalışmalarından biri Ekonomik İşbirliği ve Kalkınma Örgütü'nün (OECD) düzenlediği Uluslararası Öğrenci Değerlendirme Programı (Programme for International Student Assessment-PISA)'dır. PISA'ya katılan öğrencilerin fen okuryazarlığı "etkin bir vatandaş olarak fenle ilgili fikirlerle ve fenle alakalı meselelerle uğraşabilme becerisi" olarak tanımlanmaktadır (OECD, 2017). Bireylerin bu becerilerini etkileyen sosyodemografik özelliklerin yanı sıra duyuşsal faktörler de bulunmaktadır (Gardner, 1975). Bu sebeple bireylerin fen okuryazarlığını etkileyen değişkenlere yönelik birçok çalışma yapılmıştır (Acosta & Hsu, 2013; Hanrahan, 1999; Schroeder vd., 2007). Türkiye'de fen okuryazarlığını etkileyen faktörleri incelemek için yapılan çalışmaların birçoğu PISA verileri kullanılarak gerçekleştirilmiştir (Bilican Demir & Yıldırım, 2021; Döş & Atalmış, 2016; Kaya & Doğan, 2017; Taş vd., 2016). Fen okuryazarlığını etkileyen değişkenler arasında bu çalışmaya da konu olan fen öğrenmekten zevk alma, fen öğreniminde araçsal güdülenme, fen öz yeterliği, eğitim süreci engeli ve öğrenme engeli yer almaktadır.

Daha önceki araştırmaların sonuçlarına göre fen öğrenmekten zevk alma (Anıl, 2009; Bussie`re vd., 2007; Woods-McConney vd., 2013), fen öğreniminde araçsal güdülenme (Woods-McConney, vd., 2013; Yu, 2012), fen öz yeterliği (Acar & Öğretmen 2012; Bussie`re vd., 2007; Woods-McConney vd., 2013), eğitim süreci engeli (Beese & Liang, 2010; Greenwald vd., 1996; Özkan vd., 2018) ve öğrenme engeli (Barutçu Yıldırım & Demir, 2020; Török vd., 2018) değişkenlerinin fen okuryazarlığı üzerindeki etkisi anlamlı bulunmuştur. Fen öğrenmekten zevk alma, fen öğrenmeye karşı ilgi veya dışavurulan beğeni; araçsal güdülenme, dersin gelecekteki eğitim ve kariyere katkısı olacağı algısına yönelik tutum olarak şeklinde ifade edilebilir. Ayrıca fen öz yeterliği, öğrencilerin fene ilişkin becerilerini gerçek yaşamda etkin bir şekilde kullanabilme ve karşılaşılan zorluklarla baş edebilme düzeylerine olan inanç; eğitim süreci engeli, okulun eğitim verme kapasitesini engelleyen koşullar/durumlar (örneğin, öğretmen eksikliği, öğretim materyali eksikliği veya fiziksel altyapı eksikliği) ve öğrenme engeli ise öğrenmelerin öğretmen ve öğrencilere bağlı çeşitli davranışlardan dolayı engellenmesi (örneğin, öğrencilerin okuldan kaçması, öğretmenlere saygısızlık veya öğretmen devamsızlığı) olarak açıklanabilir. Literatürde de ele alınan bu değişkenler daima tek bir düzeyden elde edilmemektedir. Daha açık bir ifadeyle, PISA'dan elde edilen bu değişkenler öğrenciler, okul yöneticileri, eğitimciler ve ebeveynlerden elde edilmektedir. Bu durumda bazı

veri setlerinin birlikte kullanımından kaynaklı olarak veri seti aşamalı bir özelliğe bürünmektedir. Örneğin, bireyin başarısını etkileyen değişkenlerin araştırıldığı çalışmalarda, birey ve okul yöneticilerinden elde edilen değişkenlerin birey düzeyindeki tek bir bağımlı değişkene yönelik etkileri olduğundan hiyerarşik bir yapı oluşturmaktadır. Veri yapısı araştırmada kullanılacak veri analizi yöntemine ve araştırmanın sonuçlarına etki etmektedir. PISA uygulamalarından elde edilen veriler aşamalı bir özellik taşımasına rağmen fen okuryazarlığını etkileyen faktörleri incelemek için bazı çalışmalarda tek düzeyli analiz yöntemlerinin kullanıldığı görülmektedir (Özer & Anıl, 2011; Usta & Çıkrıkçı Demirtaşlı, 2014). Varyans analizi, regresyon analizi, faktör analizi, yapısal eşitlik modellemesi gibi tek düzeyli analizler, gözlemlerin birbirlerinden bağımsız olması varsayımına dayandığından aşamalı veri yapısına sahip iken kullanılması bazı sorunların çıkmasına yol açmaktadır (Çoker, 2009). Tek düzeyli modellerin hiyerarşik verilere uygulanmasında, analiz birimini belirlemede zorunlu seçim yapma, ölçme hassasiyetinin değişimi, parametre tahminleri için yöntem sınırlaması, yanlış kestirimlerin elde edilmesi, grup içi varyasyon göz ardı edilmesinden dolayı anlamlılık testlerinin gücünün azalması gibi sorunlarla karşılaşmaktadır (Draper, 1995; Kaplan & Elliott, 1997; Raudenbush, 1995). Bu sebeple hiyerarşik veri yapısı üzerinde çalışma yürütülüyorsa çok düzeyli analizlerin kullanılması önerilmektedir (Raudenbush ve Bryk, 2002).

Çok düzeyli analizler bulunmadan önce, araştırmacılar tek bir düzeyde analiz edilebilecek bir veri kümesi oluşturmak için değişkenleri hiyerarşik bir yapıda ayırması (disaggregation) ya da birleştirmesi (aggregation) gerekiyordu (Heck, 2001). Örneğin, ABD’de tüm eyaletlerde bireylerin yaşam doyumlarını incelemek isteyen bir araştırmacı veri setinde eyaletlere göre yaptığı birleştirme yaklaşımında sadece grup düzeyinde, bireylerin sayısına göre yaptığı ayırma yaklaşımında ise sadece birey düzeyinde analizleri gerçekleştirebilmekteydi. Her iki yaklaşımda da elde edilen sonuçlar bütün ölçümlerdeki grup (eyalet) üyelikleri bağımsızsa veya etkiler gruplarda farklılık göstermezse geçerlidir. Ancak bu koşullar çoğunlukla sağlanmaz ve bazı sorunlara yol açar (Raudenbush & Bryk, 2002). Birleştirme yaklaşımında bilgi, güç kaybına ve katsayıların yanlış yorumlanmasına, ayırma yaklaşımında gruplar arası değişkenlik göz ardı edilmesine sebep olur (Raudenbush & Bryk, 2002). Hiyerarşik yapıdaki verilerle tek düzeyli modelleme ile yürütülen çalışmalarda karşılaşılan istatistiksel ve kavramsal sorunlar çok düzeyli modelleme tekniklerinin gelişimini sağlamıştır (Heck & Thomas, 2015).

Çok düzeyli modelleme tek düzeyli analiz yöntemlerine göre bazı avantajlar sağlamıştır. Bu avantajları şu şekilde ifade edilebilir: a) dengeli olmayan ve kayıp veriler kolaylıkla ele alınabilir, b) hiyerarşinin farklı düzeylerindeki yapısal ilişkiler belirlenebilir ve ölçülebilir, c) analizin farklı düzeylerindeki hipotezler test edilebilir, d) önerilen hiyerarşik yapının geçerliliği olabilirlik veya Bayes testleriyle değerlendirilebilir, e) standart hatalar olduğundan daha düşük değil doğru olarak kestirilir, f) çok sayıda modelleme yaklaşımı kullanılabilir (Raudenbush & Bryk, 2002). Getirdiği avantajlar göz önüne alındığında hiyerarşik verilerde çok düzeyli modellerin tercih edilmesi beklenmektedir.

Çok düzeyli modeller, analiz edilen verilerin her seviyede yeterli varyasyon sağladığı ölçüde faydalıdır ve gereklidir. Varyansı grup içi ve gruplar arası bileşenlerine bölmek, çok düzeyli bir analizin gerekçelendirilip gerekçelendirilmemesine karar vermek için araştırmacılara kanıt sağlar (Heck & Thomas, 2015). Aynı grupta bulunan birimler arasındaki homojenlik derecesini belirlemek gözlemlerin gruba bağımlılığını vurgulayacaktır. Homojenlik derecesinin, yani sınıf içi (intra-class) korelasyonunun artması çok düzeyli analizleri gerekli kılmaktadır (Can, Somer, Korkmaz, Dural ve Öğretmen, 2011) ve

$$\rho = \sigma_b^2 / (\sigma_b^2 + \sigma_w^2) \quad (1)$$

eşitliği ile hesaplanmaktadır (Heck & Thomas, 2015). Eşitlik 1’de görüldüğü üzere sınıf içi korelasyon (ρ), gruplar arası varyansın toplam varyansa oranıdır (Farmer, 2000). Genel olarak sınıf içi korelasyon 0.05’ten büyük olması durumunda çok düzeyli modellerin kullanılması önerilmektedir (Dyer vd., 2005). Ancak 0.70’ten büyük olması durumunda ise grupların homojen olduğu söylenebilir. Bu sebeple birleştirme yaklaşımı kullanılarak grup düzeyinde analizlerin yürütülmesi gerekmektedir (Noyan, 2009).

Her analizde olduğu gibi çok düzeyli modellerin de kısıtlamaları bulunmaktadır. Dezavantajlı yönlerinden birisi değişkenler arasındaki dolaylı ilişkilerin araştırılmasını yetersiz kılmasıdır. Yani hiyerarşik verinin her bir düzeyinde (birey ve grup düzeyi) eşzamanlı testlere imkân veren birden çok düzeye sahip yapısal eşitlik modelleri ile çözümlemeler yapılmaya başlanmıştır (Uzun & Çokluk Bökeoğlu, 2019). Çok düzeyli yapısal eşitlik modelleri (ÇDYEM), kümelenmiş verilere doğrudan ve aracılık etkisi olan gizil değişkenlerin ve daha karmaşık yol modellerinin avantajlarını sağlar (Mehta & Neale, 2005).

ÇDYEM alanyazında çok düzeyli örtük değişken modelleri (multilevel latent variable modeling) (Kaplan vd., 2009), çok düzeyli kovaryans yapı analizi (multilevel covariance structure analysis) (Muthén, 1994; Yuan & Bentler, 2007), çok düzeyli kovaryans bileşeni modelleri (multilevel covariance component models) (Goldstein, 1987) olarak da bilinmektedir. Ölçme modeli ve yapısal model bileşenlerinden oluşan ÇDYEM’de değişkenler grup içi ve gruplar arası seviyelerde tanımlandığından dolayı bireylerden elde edilen verinin kovaryans matrisi de düzeylere ayrılmaktadır. ÇDYEM’in kovaryans yapısına ilişkin model Eşitlik 2’de verilmiştir (Muthén, 1994):

$$\begin{aligned}\Sigma_T &= \Sigma_w + \Sigma_B \\ \Sigma_B &= \lambda\Psi\lambda' + \Theta_B \\ \Sigma_w &= \lambda\Psi_w\lambda' + \Theta_w\end{aligned}\quad (2)$$

Eşitlik 2’de görüldüğü üzere ÇDYEM’de toplam kovaryans yapıları (Σ_T) grup içinde (Σ_w) ve gruplar arasında (Σ_B) varyans bileşenlerine ayrılmıştır. Grup içi ve gruplar arası bileşenleri tanımlayan λ faktör yüklerini, Ψ kovaryans matrisini ve Θ hata kovaryans matrisini göstermektedir (Depaoli & Clifton, 2015). Çok düzeyli yapısal eşitlik modellerinde kovaryans yapılarının kestirimine ilişkin notasyon ve derinlemesine bilgi için Muthén (1994) ile Heck ve Thomas (2015) yayınları incelenebilir.

Bu bilgilerden hareketle ve araştırmacıların daha önce ifade ettiği üzere, çok düzeyli veri setleri tek düzeyli analizlerle ele alınmaya devam edilmektedir ve bu durum, yanıltıcı sonuçlara neden olabilmektedir. Bu uygulamaların araştırmacıların vatandaşı olduğu ülkenin verileri ile yapılan çalışmalarda gözlenmesi, araştırmacılara çok düzeyli verilerin analizinde kullanılması gereken analizlere yönelik uygulama için bir kanıt sunma ihtiyacı oluşturmuştur. Bu sebeple PISA 2015 Türkiye veri setinde ampirik kanıtlar sunmak amacıyla model karşılaştırması, bu çalışmanın bir boyutunu oluşturmuştur. Çalışmanın diğer boyutunda ise uygun veri analiz yönteminden elde edilen sonuçlarla, fen okuryazarlığını etkilediği düşünülen değişkenlerin etkisini ortaya koymak hedeflenmiştir. Bu boyutlar doğrultusunda araştırmada Türkiye’de PISA 2015 uygulamasına katılmış 16 yaşındaki bireylerin fen okuryazarlığı puanları ile fen öğrenmekten zevk alma, fen öğreniminde araçsal güdülenme, fen öz yeterliği, eğitim sürecindeki engeller, öğrenme engeli değişkenleri arasındaki ilişkinin çok düzeyli ve tek düzeyli yapısal eşitlik ile modellenmesi amaçlanmıştır. Bu amaçlara göre aşağıda yer alan problemlere yanıt aranmaktadır:

1. Öğrenci düzeyinde fen okuryazarlığı puanı; fen öğrenmekten zevk alma, fen öğreniminde araçsal güdülenme, fen öz yeterliği, okul düzeyinde; eğitim sürecindeki engeller, öğrenci öğrenmelerine engeller değişkenleri ile ne düzeyde açıklanmaktadır?
2. Öğrencilerin fen öğrenmekten zevk alma, fen öğreniminde araçsal güdülenme, fen öz yeterliği, eğitim sürecindeki engeller, öğrenme engeli değişkenlerinin fen okuryazarlığı puanı ile tek düzeyli yapısal ilişkisi, çok düzeyli yapısal ilişkisine göre ne düzeydedir?

Yöntem

Araştırmanın bu bölümde araştırma modeline, evren-örneklem özelliklerine, veri toplanma araçlarına ve verilerin analizine ilişkin bilgilere yer verilmiştir.

Araştırma Modeli

Bu çalışmada Türkiye’de PISA 2015 uygulamasına katılmış 16 yaşındaki bireylerin fen okuryazarlığı puanları ile fen öğrenmekten zevk alma, fen öğreniminde araçsal güdülenme, fen öz yeterliği, eğitim

sürecindeki engeller, öğrenme engeli değişkenleri arasındaki ilişki yapısal modelleme yoluyla incelenmiştir. Çalışmada iki ya da daha fazla değişken arasındaki ilişkileri belirlemek esas olduğu için çalışmanın türü korelasyonel araştırmadır (Büyüköztürk vd., 2014; Fraenkel vd., 2012). Aynı zamanda hiyerarşik yapı verilerde iki düzeyli ve tek düzeyli yapısal eşitlik modellemesi gerçekleştirilerek var olan durumun ortaya çıkarılması amaçlandığı için betimsel bir çalışmadır (Büyüköztürk vd., 2014).

Evren-Örneklem

Bu araştırmanın evrenini 2015 yılında Türkiye’de öğrenim gören 15 yaş 3 ay ve 16 yaş 2 ay aralığındaki 1.324.089 bireyden oluşturmaktadır. PISA 2015 uygulamasında örnekleme alınacak bireyler iki aşamada belirlenmektedir. İlk aşamada okullar, çeşitli tabakalar (eğitim türü, okul türü, okulların bulundukları yer ve okulların idari biçimleri) dikkate alınarak tabakalı seçkisiz örnekleme, ikinci aşamada ise bu okullarda uygulamaya katılacak olan öğrenciler seçkisiz yöntemiyle belirlenmektedir. PISA tarafından Türkiye’de 187 okuldan rastgele seçilen 5895 öğrencinin tamamı bu çalışmanın örneklemini oluşturmaktadır. Örneklemi oluşturan öğrencilerin çeşitli değişkenler bakımından dağılımlarına ilişkin bilgiler Tablo 1’de sunulmuştur.

Tablo 1.

PISA 2015 Türkiye Uygulamasına İlişkin Örneklem Dağılımı

	Sınıf düzeyi												Yaş	
	7. sınıf		8. sınıf		9. sınıf		10. sınıf		11. sınıf		12. sınıf		16	
	n	%	n	%	n	%	n	%	n	%	n	%	n	%
Kadın	6	0.2	42	1.4	494	16.8	2272	77.3	120	4.1	4	0.1	2938	100
Erkek	10	0.3	63	2.1	779	26.3	2036	68.9	66	2.2	3	0.1	2957	100
Toplam	16	0.3	105	1.8	1273	21.6	4308	73.1	186	3.2	7	0.1	5895	100

Veri Toplama Araçları

Matematik, fen ve okuduğunu anlama alanlarında her üç yılda bir öğrencilerin performansı PISA tarafından değerlendirilmektedir. Bu değerlendirme sürecinde öğrenci performanslarındaki farklılıkları açıklamak için öğrenci, aile ve okul anketleri ile de veri toplanmaktadır. Bu anketlerde bireylerin demografik özelliklerinin yanı sıra psikolojik yapılarla ilişkin de veriler elde edilmektedir.

Bu araştırmada ölçme araçları olarak OECD tarafından hazırlanan fen okuryazarlığı bilişsel alan testi, öğrenci ve okul anketleri kullanılmıştır. Bilişsel alan testinden aldıkları puanlar öğrencilerin fen okuryazarlığı puanları olarak kullanılmıştır. Öğrenci anketinde; “fen öğrenmekten zevk alma”, “fen öğreniminde araçsal güdülenme”, “fen öz yeterliği” ölçekleri, okul anketinde; “eğitim süreci engeli”, “öğrenme engeli” ölçekleri kullanılmıştır. Ölçeklerden alınan puanların yükselmesi belirtilen özelliğin grupta arttığı anlamına gelmektedir.

Bilişsel alan testi: PISA uygulamasında bireylerin fen başarılarını ölçmek amacıyla bireylere bilişsel alan testinin tamamı uygulanmamakta ve testte yer alan alt testleri yanıtlamaktadırlar. Bu sebeple öğrencilerin bütün sorulara verebileceği olası yanıtların kestiriminden yararlanılarak fen okuryazarlığı puanları elde edilmektedir. PISA 2015 uygulamasında OECD tarafından fen okuryazarlığına ilişkin on olası değer (PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE, PV6SCIE, PV7SCIE, PV8SCIE, PV9SCIE, PV10SCIE) kestirilmiştir. Testin güvenilirliği Türkiye için 0,97 olarak belirlenmiştir.

Fen öğrenmekten zevk alma ölçeği: PISA 2015’te öğrencinin okumaktan zevk alma düzeyini belirlemek amacıyla 5 madde (ST094Q01NA, ST094Q02NA, ST094Q03NA, ST094Q04NA, ST094Q05NA) kullanılmıştır. Ölçeğin iç tutarlılık anlamındaki güvenilirliği Türkiye için 0,95 olarak hesaplanmıştır.

Fen öğreniminde araçsal güdülenme ölçeği: Öğrencilerin okulda gördükleri fen dersinin eğitim ve kariyer planlarına ne kadar faydalı olduğuna dair algılarını belirlemek için kullanılmıştır. Ölçek 4 maddeden (ST113Q01TA, ST113Q02TA, ST113Q03TA, ST113Q04TA) oluşmaktadır. Ölçeğin Cronbach alpha güvenilirliği Türkiye için 0,90’dır.

Fen öz yeterliği ölçeği: Gerçek dünya şartlarında öğrencilerin fen bilgi ve becerilerini algılama düzeylerini belirlemek için kullanılmıştır. Fen öz yeterliğini belirleyen ölçek 8 maddeden (ST129Q01TA, ST129Q02TA, ST129Q03TA, ST129Q04TA, ST129Q05TA, ST129Q06TA, ST129Q07TA, ST129Q08TA) oluşmaktadır. Ölçeğin Cronbach alpha güvenirliği Türkiye için 0,89 olarak belirlenmiştir.

Eğitim süreci engeli ölçeği: Öğretim kadrosu, yardımcı eleman, eğitim materyali, fiziksel alt yapıdaki eksiklik ve yetersizlik eğitim sürecindeki engeller olarak belirtilmektedir ve değişkenini belirlemek için ölçekte 8 madde (SC017Q01NA, SC017Q02NA, SC017Q03NA, SC017Q04NA, SC017Q05NA, SC017Q06NA, SC017Q07NA, SC017Q08NA) yer almaktadır. Ölçeğin Cronbach alpha güvenirliği Türkiye için 0,89 olarak belirlenmiştir.

Öğrenme engeli ölçeği: Öğrenci-öğretmen devamsızlığı, özensizliği vb. durumlar öğrenme engeli olarak belirtilmektedir. Öğrenme engelini belirleyen ölçek 10 maddeden (SC061Q01TA, SC061Q02TA, SC061Q03TA, SC061Q04TA, SC061Q05TA, SC061Q06TA, SC061Q07TA, SC061Q08TA, SC061Q09TA, SC061Q010TA) oluşmaktadır. Ölçeğin Cronbach alpha güvenirliği Türkiye için 0,83 olarak hesaplanmıştır.

Verilerin Analizi

Araştırmanın amacı doğrultusunda yapılacak olan tek ve çok düzeyli yapısal eşitlik modellemesi için öncelikle veriler düzenlenmiş ve temizlenmiştir. Fen okuryazarlığı puanları, öğrenci ve okul anketleri ile birleştirilerek kayıp veri analizi için MCAR (Missing Completely at Random/Tamamen Rastlantısal Kayıp; Little, 1988) testi gerçekleştirilmiştir. Sonuç olarak eksik verilerin tamamen tesadüfi olarak kayıp olmadığı belirlenmiştir ($\chi^2=1582.327$, $sd=980$, $p<0.05$). Kayıp verilerle baş etme yöntemi olarak beklenti maksimizasyonu yöntemi tercih edilmiştir. Kayıp veri analizinin ardından grup sayısına göre grup büyüklükleri belirlenmiştir. 2015 PISA uygulamasında Türkiye örnekleminde ikinci düzey grup sayısının 100'den fazla olmasından dolayı her grup için örneklem büyüklüğü 10'dan daha az bireye sahip ikinci düzey (okul) kümeler veri setinden çıkarılmıştır (Hox vd., 2017, s. 216). Temizleme işlemi sonucunda 161 okul altında 5800 bireyin verileri analiz için kullanılmıştır. Verilerin temizlenmesinin ardından doğrusallık, varyansların homojenliği, artıkların normal dağılımına ilişkin analizler gerçekleştirilerek varsayımlar test edilmiştir. Ölçme modelini oluşturabilmek için bireysel düzeydeki artıkların varyansının homojen olduğu ($p=0.548$), model artıkları ve bağımlı değişken arasında çizilen grafikte rastgele bir desen oluştuğu ve artıkların normal dağılım gösterdiği belirlenmiştir.

Yapısal modeller oluşturulmadan önce tek düzeyli modeli test etmek için veri seti üzerinde tekrar düzenleme yapılmıştır. Ayırma (disaggregate) çalışmasıyla grup düzeyinin ve birleştirme (aggregate) çalışmasıyla birey düzeyinin göz ardı edildiği iki farklı veri seti düzenlenmiştir. Ardından yapısal modeller oluşturulmuştur.

Verilerin düzenlenmesi ve ölçme araçlarının güvenirliğine ilişkin analizler SPSS 23.0 programında yapılmıştır. Çalışmanın varsayımları RStudio "lme4" paketi (Bates vd., 2015) ve alt problemlerine ilişkin model analizleri ise RStudio "MplusAutomation" paketi (Hallquist & Wiley, 2018) ile gerçekleştirilmiştir. Modellerin analiz edilmesinde parametre ve standart hataların tahmini birinci dereceden türevlere ve normal ki-kare istatistiğine dayanan MLF yöntemi kullanılmıştır (Heck & Thomas, 2015). Verilerin analizinde kullanılan kodlar Ek 1'de sunulmuştur.

Modeller

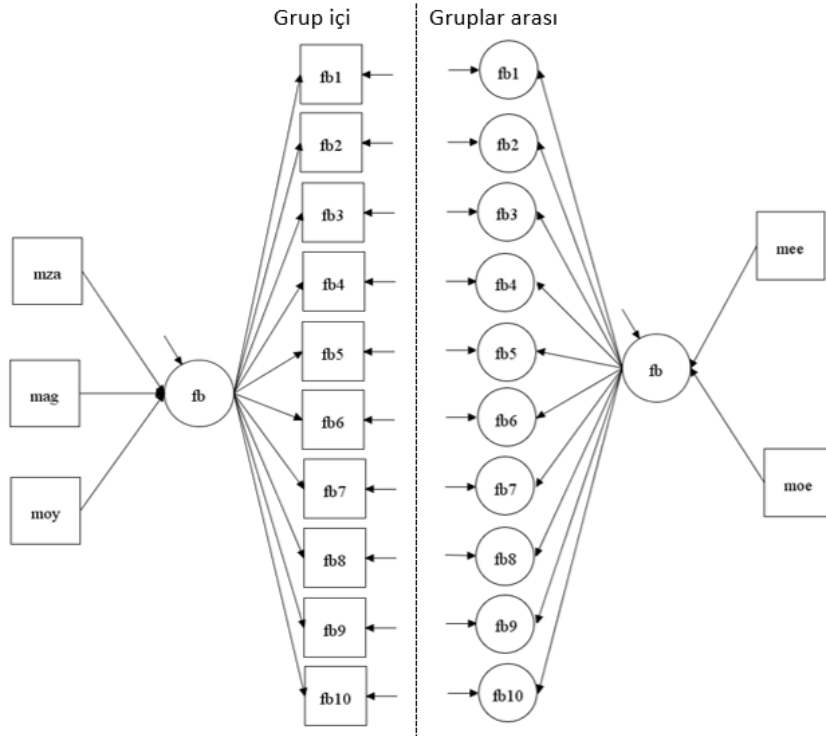
Araştırmanın amacı ve alt problemleri doğrultusunda fen başarısını açıklayan değişkenleri öngören iki düzeyli ve tek düzeyli yapısal eşitlik modelleri kurulmuştur. Bu modellerde yer alan değişkenlerin gösterimine ilişkin bilgiler Tablo 2'de sunulmuştur.

Tablo 2.
Modellerin Değişkenlerinin Gösterimi

Değişken	Gösterim	Değişken	Gösterim
PV1SCIE	fb1	PV9SCIE	fb9
PV2SCIE	fb2	PV10SCIE	fb10
PV3SCIE	fb3	Fen Başarısı (fen okuryazarlığı)	Fb
PV4SCIE	fb4	Fen öğrenmekten zevk alma	Mza
PV5SCIE	fb5	Fen öğreniminde araçsal güdülenme	Mag
PV6SCIE	fb6	Fen öz yeterliği	Moy
PV7SCIE	fb7	Eğitim süreci engeli	Mee
PV8SCIE	fb8	Öğrenme engeli	Moe

İki düzeyli yapısal eşitlik modeli

Çalışmanın amacı ve birinci alt problemi doğrultusunda oluşturulan iki düzeyli yapısal eşitlik modeli Şekil 1’de verilmiştir. Yapısal modelin grup içi düzeyinde fb1’den fb10’a kadar olan sürekli gözlenen değişkenler fb gizil değişkeni tarafından açıklanmaktadır. Belirtilen sürekli gözlenen değişkenler, gruplar arası düzeyde rastgele kesişimleri (random intercepts) temsil ettiği için daire şeklinde gösterilmektedir. Her iki düzeyde de gözlenen değişkenler gizil değişkene bağlanmaktadır. Ancak fb gizil değişkeni gruplar arasında ve grup içinde farklı değişkenler tarafından açıklanmaktadır. Grup içi düzeyde mza, mag, moy, gruplar arası düzeyde ise mee, moe gizil değişkendeki varyasyonu açıklayan yordayıcı değişken olarak yer almaktadır. Test edilen bu model “Model 1” olarak adlandırılmıştır.

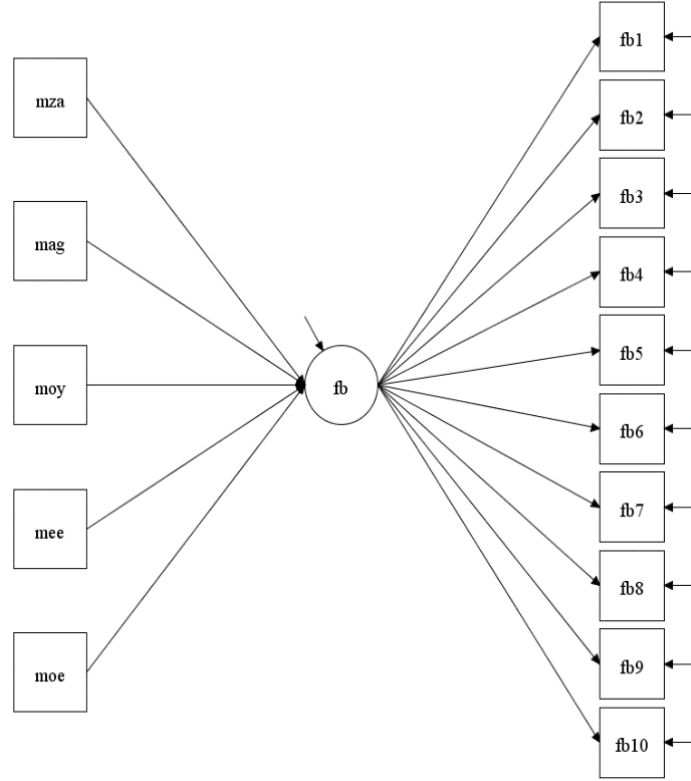


Şekil 1. Test Edilen İki Düzeyli Yapısal Eşitlik Modeli (Model 1)

Tek düzeyli yapısal eşitlik modeli

Çalışmanın amacı ve ikinci alt problemi doğrultusunda oluşturulan çok düzeyli yapısal eşitlik modeli Şekil 2’de verilmiştir. İki düzeyli modelde fb gizil değişkenini gruplar arası ve grup içi düzeyde açıklayan mza, mag, moy, mee ve moe değişkenleri, bu modelde tek düzeyde fen başarısındaki değişimi açıklamaktadır. Ayırma çalışmasıyla grup düzeyinin göz ardı edildiği veri setinin kullanıldığı yapısal model

“Model 2”, birleştirme çalışmasıyla birey düzeyinin göz ardı edildiği veri setinin kullanıldığı model ise “Model 3” olarak adlandırılmıştır.



Şekil 2. Test Edilen Tek Düzeyli Yapısal Eşitlik Modeli (Model 2 - Model 3)

Bulgular

Çok düzeyli analizlerin gerçekleştirilebilmesi için 0.05’ten az olmaması gereken gözlenen değişkenlere ilişkin sınıf içi korelasyon katsayıları Tablo 3’te verilmiştir.

Tablo 3.

Sınıf İçi Korelasyon Katsayıları

	Gözlenen değişkenler									
	fb1	fb2	fb3	fb4	fb5	fb6	fb7	fb8	fb9	fb10
Sınıf içi korelasyon katsayıları	.475	.478	.483	.489	.495	.488	.476	.485	.488	.491

Tablo 3 incelendiğinde fen başarısı gizli değişkenine ait 10 gözlenen değişkene ilişkin sınıf içi korelasyon katsayılarının 0.475 ve 0.495 arasında değiştiği görülmektedir. Gözlenen değişkenlere ilişkin katsayıların 0.05’ten büyük ve 0.70’ten küçük olması bireylerin grup içinde homojenlik gösterdiğinin ancak grup homojenliğine sebep olmadığını göstermektedir. Bu korelasyon değerlerine göre fen başarısını ifade eden gözlenen değişkenlerin puanlarındaki toplam varyansın yaklaşık %48-49’unun okullar arasında olduğu tespit edilmiştir. Yani öğrencilerin fen başarısı göstergelerindeki (fb1-fb10) değişimin %48-49’u okullar arası farklılıklardan, %51-52’si ise aynı okulda öğrenim gören öğrenciler arasındaki farklılıklardan kaynaklanmaktadır. Bu bilgiler doğrultusunda belirtilen veri setinde tek düzeyli yapısal eşitlik modeli kurmanın yanı sıra parametre tahminlerine sebep olacağı ve bu sebeple çok düzeyli yapısal eşitlik modeline ihtiyaç duyulduğu belirlenmiştir. Ancak bu çalışmada çok düzeyli analizler

uygulanması gerektiğinde tek düzeyli analizlerin kullanılmasının model uyum indeksleri ve parametre tahminleri üzerindeki etkisine ampirik kanıtlar sunulması amaçlandığından hem çok düzeyli hem tek düzeyli yapısal eşitlik modelleri kurulmuştur. Model 1 iki düzeyli, ayırma yaklaşımıyla oluşturulan veriler ile kurulan Model 2 ve birleştirme yaklaşımıyla oluşturulan veriler ile kurulan Model 3 tek düzeyli yapısal eşitlik modelleridir.

Çalışmanın amacı doğrultusunda oluşturulan iki düzeyli (Model 1) ve tek düzeyli (Model 2-Model 3) modellere ilişkin uyum indeksleri Tablo 4'te sunulmuştur.

Tablo 4.

Model Uyum İndeksleri

	χ^2	χ^2/sd	RMSEA	CFI	TLI	SRMR
Model 1	155.701	1.353	.008	.999	.999	.004 ⁱ .006 ^a
Model 2	130.025	1.625	.010	.999	.999	.004
Model 3	139.928	1.749	.068	.990	.988	.006

i: grup içi a: gruplar arası

Tablo 4 incelendiğinde Model 1 için sadece SRMR indeksi değerinin hem grup içi hem gruplar arası düzey için sırasıyla .004 ve .006 olarak, SRMR indeksi dışında tüm indekslerin her iki düzey için hesaplandığı görülmektedir. Veri-model uyumunu değerlendiren uyum indekslerinden χ^2/sd oranının 2'den küçük olması, RMSEA değerinin .06'ya eşit ya da bu değerden küçük olması, CFI ve TLI değerlerinin .95'e eşit ya da bu değerden büyük olması verinin modele iyi uyum gösterdiğini belirtir (Hu ve Bentler, 1999; Tabachnick ve Fidell, 2012). Model 1, 2 ve 3'e ait uyum indeksleri incelendiğinde ise tüm modeller için verinin modele mükemmel uyum gösterdiğini söyleyebiliriz. Ancak modellerin uyum indeksleri birbirleriyle kıyaslandığında Model 1'in veriye görece daha iyi uyum sağladığı görülmektedir.

İki düzeyli ve tek düzeyli yapısal eşitlik modelleri için elde edilen faktör yükleri, regresyon katsayıları ve hata varyansları için standartlaştırılmış kestirimler Tablo 5'te sunulmuştur.

Tablo 5.

Model Uyum İndeksleri

İki Düzeyli Yapısal Eşitlik Modeli Sonuçları (Model 1)					Tek Düzeyli Modellerin Sonuçları (Model 2-3)				
Düzye/ Değişkenler	β	St. Hata	t	p	Düzye/ Değişkenler	β	St. Hata	t	p
Grup içi					Model 2				
fb BY					fb BY				
fb1	.868	.004	194.573	.000	fb1	.934	.002	487.255	.000
fb2	.881	.004	240.905	.000	fb2	.941	.002	529.818	.000
fb3	.872	.004	238.275	.000	fb3	.937	.002	514.149	.000
fb4	.875	.004	233.119	.000	fb4	.939	.002	519.848	.000
fb5	.872	.004	230.641	.000	fb5	.938	.002	523.595	.000
fb6	.873	.004	237.037	.000	fb6	.937	.002	509.248	.000
fb7	.876	.004	239.565	.000	fb7	.938	.002	521.884	.000
fb8	.872	.004	248.895	.000	fb8	.935	.002	496.687	.000
fb9	.875	.004	217.448	.000	fb9	.939	.002	512.162	.000
fb10	.875	.003	252.746	.000	fb10	.939	.002	522.679	.000
fb ON					fb ON				
mza	.161	.016	10.235	.000	mza	.165	.013	13.164	.000
mag	.000	.019	0.024	.981	mag	.002	.013	0.128	.898
					moy	-.101	.013	-7.791	.000

					mee	-.250	.013	-18.558	.000
moy	-.077	.016	-4.811	.000	moe	-.111	.014	-7.871	.000
Hata varyansı					Hata varyansı				
fb1	.246	.008	31.749	.000	fb1	.127	.004	35.586	.000
fb2	.223	.006	34.596	.000	fb2	.114	.003	34.232	.000
fb3	.240	.006	37.546	.000	fb3	.123	.003	35.924	.000
fb4	.234	.007	35.700	.000	fb4	.119	.003	34.950	.000
fb5	.240	.007	36.422	.000	fb5	.119	.003	35.518	.000
fb6	.238	.006	37.075	.000	fb6	.121	.003	35.168	.000
fb7	.233	.006	36.354	.000	fb7	.120	.003	35.713	.000
fb8	.240	.006	39.305	.000	fb8	.125	.004	35.519	.000
fb9	.234	.007	33.304	.000	fb9	.119	.003	34.647	.000
fb10	.234	.006	38.638	.000	fb10	.118	.003	35.055	.000
fb	.964	.006	168.566	.000	fb	.848	.009	93.581	.000
Gruplararası					Model 3				
fb BY					fb BY				
fb1	.998	.001	767.421	.000	fb1	.993	.002	600.224	.000
fb2	.998	.001	1069.992	.000	fb2	.994	.001	726.220	.000
fb3	.997	.001	896.949	.000	fb3	.993	.001	732.420	.000
fb4	.997	.001	883.131	.000	fb4	.994	.001	750.668	.000
fb5	.997	.001	919.439	.000	fb5	.994	.001	841.999	.000
fb6	.996	.001	855.469	.000	fb6	.992	.002	622.269	.000
fb7	.997	.001	1102.504	.000	fb7	.994	.001	834.062	.000
fb8	.994	.002	540.576	.000	fb8	.990	.002	492.825	.000
fb9	.996	.001	729.800	.000	fb9	.991	.002	535.639	.000
fb10	.997	.001	1004.562	.000	fb10	.993	.001	765.790	.000
fb ON					fb ON				
					mza	.318	.082	3.899	.000
mee	-.339	.106	-3.215	.001	mag	.039	.081	0.486	.627
					moy	-.282	.077	-3.677	.000
moe	-.159	.118	-1.344	.179	mee	-.227	.084	-2.686	.007
					moe	-.062	.088	-0.701	.483
Hata varyansı					Hata varyansı				
fb1	.005	.003	1.906	.057	fb1	.014	.003	4.404	.000
fb2	.004	.002	2.262	.024	fb2	.012	.003	4.362	.000
fb3	.007	.002	3.041	.002	fb3	.014	.003	5.341	.000
fb4	.006	.002	2.688	.007	fb4	.013	.003	4.890	.000
fb5	.005	.002	2.310	.021	fb5	.011	.002	4.785	.000
fb6	.007	.002	3.189	.001	fb6	.016	.003	5.009	.000
fb7	.005	.002	2.860	.004	fb7	.013	.002	5.365	.000
fb8	.013	.004	3.421	.001	fb8	.019	.004	4.896	.000
fb9	.007	.003	2.704	.007	fb9	.017	.004	4.709	.000
fb10	.007	.002	3.416	.001	fb10	.013	.003	5.214	.000
fb	.813	.072	11.348	.000	fb	.601	.071	8.428	.000

Model 1'e ilişkin faktör yükleri incelendiğinde grup içi faktör yüklerinin .868 ile .881 aralığında ve gruplar arası faktör yüklerinin .994 ile .998 aralığında değiştiği görülmektedir. Bireylerden elde edilen veriler ile kurulan Model 2 ve gruplardan elde edilen veriler ile kurulan Model 3'te ise faktör yüklerinin sırasıyla .934 ile .941 ve .990 ile .994 aralığında değiştiği belirlenmiştir. Bu bilgilere göre hem iki düzeyli modelde hem tek düzeyli modellerde birey düzeyi için faktör yüklerinin grup düzeyi için faktör yüklerinden daha küçük olduğu tespit edilmiştir. Gözlenen değişkenlere ilişkin hata varyansları ise faktör yükleri ile ters yönlü olarak işlediğinden dolayı tüm modellerde birey düzeyi için hata varyanslarının grup

düzeyi için hata varyanslarından daha büyük olduğu görülmüştür. Belirtilen faktör yük değerleri tüm modellerde ve iki düzeyli modelde gruplar arası düzeyde yer alan f1 değişkeni dışında tüm hata varyanslarının anlamlı olduğu görülmektedir. Bu sebeple tek düzeyli ve iki düzeyli modellerin faktör yükleri ve hata varyansları bakımından farklılaşma olsa da bu farklılaşmanın manidar olmadığı söylenebilir.

Fen başarısını yordayıcı değişkenlerin regresyon katsayıları incelendiğinde Model 1’de grup içinde fen öğrenmekten zevk alma, fen öğreniminde araçsal güdülenme ve fen öz yeterliği yordayıcı değişkenlerinden fen öğreniminde araçsal güdülenmenin anlamlı bir yordayıcı değişken olmadığı tespit edilmiştir. Ayrıca yordayıcı değişkenlerden fen öğrenmekten zevk alma düzeyi bir standart sapma arttığında öğrencilerin yordanan fen başarısı puanının .161 standart sapma artacağını, fen öz yeterliği bir standart sapma arttığında ise .077 standart sapma azalacağını önermektedir. Grup içinde yer alan yordayıcı değişkenler gizil değişkendeki varyansın %4’ünü açıklamaktadır. Gruplar arası eğitim süreci engeli ve öğrenme engeli yordayıcı değişkenlerinden ise öğrenme engelinin anlamlı bir yordayıcı olmadığı görülmüştür. Eğitim süreci engelleri bir standart sapma arttığında ise .339 standart sapma azalacağı belirlenmiştir. Gruplar arasında yer alan yordayıcı değişkenler gizil değişkendeki varyansın %19’unu açıklamaktadır.

Model 2 için regresyon katsayıları incelendiğinde ise fen öğrenmekten zevk alma, fen öğreniminde araçsal güdülenme, fen öz yeterliği, eğitim süreci engeli ve öğrenme engeli yordayıcı değişkenlerinden fen öğreniminde araçsal güdülenmenin yine anlamlı bir yordayıcı olmadığı, ancak iki düzeyli modelin tersine “öğrenme engeli”nin diğer yordayıcılarla beraber anlamlı olduğu görülmektedir. Fen öğrenmekten zevk alma düzeyi bir standart sapma arttığında öğrencilerin yordanan fen başarısı puanının .165 standart sapma artacağı, fen öz yeterliği bir standart sapma arttığında ise .101 standart sapma azalacağı, eğitim süreci engelleri bir standart sapma arttığında .250 standart sapma azalacağı ve öğrenme engeli bir standart sapma arttığında ise .111 standart sapma azalacağı belirlenmiştir. Model 2 yer alan yordayıcı değişkenler gizil değişkendeki varyansın %15’ini açıklamaktadır.

Model 3 incelendiğinde ise fen öğrenmekten zevk alma, fen öğreniminde araçsal güdülenme, fen öz yeterliği, eğitim süreci engeli ve öğrenme engeli yordayıcı değişkenlerinin regresyon katsayılarının anlamlılık bakımından iki düzeyli model ile aynı olduğu görülmektedir. Fen öğrenmekten zevk alma düzeyi bir standart sapma arttığında öğrencilerin yordanan fen başarısı puanının .318 standart sapma artacağı, fen öz yeterliği bir standart sapma arttığında ise .282 standart sapma azalacağı, eğitim süreci engelleri bir standart sapma arttığında .227 standart sapma azalacağı tespit edilmiştir. Model 3 yer alan yordayıcı değişkenler gizil değişkendeki varyansın %40’ını açıklamaktadır.

Tartışma, Sonuç ve Öneriler

Bu çalışmada okullar arasında ve öğrenciler içinde fen başarısını açıkladığı öngörülen değişkenler iki düzeyli yapısal eşitlik modellemesi ile incelenmiştir. Ayrıca hiyerarşik verilerde tek düzeyli analizlerin kullanılmasının model uyum indeksleri ve parametre tahminleri üzerindeki etkisine ampirik kanıtlar sunulmuştur.

Fen başarısına ait gözlenen değişkenlerin gruplar arasında anlamlı olarak farklılaştığı sınıf içi korelasyon katsayı ile belirlenmiştir. Öğrencilerin fen başarısı göstergelerindeki (fb1-fb10) değişimin %48-49’u okullar arası farklılıklardan, %51-52’si ise aynı okulda öğrenim gören öğrenciler arasındaki farklılıklardan kaynaklandığı belirlenmiştir. Okullar arası farklılığın %5’ten fazla olması sebebiyle gerçekleştirilen iki düzeyli yapısal eşitlik modellemesi sonucunda araştırmacılar tarafından oluşturulan yapısal modelin doğrulandığı uyum indekslerine göre tespit edilmiştir.

Modelde yer alan değişkenler incelendiğinde, fen öğreniminde araçsal güdülenme ve öğrenme engeli değişkenlerinin fen başarısını anlamlı düzeyde yordamadığı belirlenmiştir. Yetişir ve diğ. (2018) fen okuryazarlık performansı üzerinde araçsal güdülenmenin de etkisini incelediği çalışmada bu araştırma ile benzer sonuçlara ulaşmıştır. Ancak OECD (2016)’nin yaptığı çalışmada fen başarısı üzerindeki etkisi incelenen araçsal güdülenmenin yordayıcı etkisi olduğu sonucuna ulaşılmıştır. İlhan (2015), öğrenci derste öğrendiği bilgilerin hayatında ve kariyerinde faydalı olacağına inanırsa, öğrencinin bu derse

yönelik çabasının artabileceğini ifade etmiştir (Akt: Güngör, 2019). Ancak, araçsal güdülenmenin başarı üzerindeki etkisinin kültüre göre değişiklik gösterdiği bilinmektedir (Kjærnsli & Lie, 2011; Yu, 2012). Bu sebeple, farklı ülkeler ile gerçekleştirilen çalışmaların ve Türk kültürüne ait çalışmaların sonuçları arasında farklılıklar olabilir. Belirtildiği üzere, araçsal güdülenmenin etkisi öğrenilen bilgilerin kariyer planları üzerindeki etkisine inanmaktan geçmektedir. Ülkenin eğitim sisteminde yaşanan değişikliklerden kaynaklı olarak Türk kültüründeki bireylerin kariyer planlaması yapmakta güçlük çekmesi, araçsal güdülenmenin fen başarısı üzerindeki etkisini değiştirmiş olabilir. Ancak, Türk kültüründe araçsal güdülenmenin fen başarısı üzerindeki etkisini inceleyen yeterli sayıda çalışma bulunmamaktadır. Dolayısıyla, bu değişkenin tahmin düzeyi hakkında güçlü ifadeler kullanabilmek için daha fazla çalışmaya ihtiyaç duyulmaktadır. Ayrıca fen başarısını yordamadığı tespit edilen ve araştırmacılar tarafından fen başarısıyla negatif yönlü bir anlamlı bir ilişkiye sahip olması beklenen öğrenme engeli değişkeni için de literatürde araştırmaya rastlanmamıştır. Öğrenme engeli değişkeninin içinde barındırdığı öğretmen devamsızlığı, öğrencilerin madde kullanımı, öğrenci zorbalığı gibi durumlar çalışmalarda (Finlayson, 2009; Hanson vd., 2003; Konishi vd., 2010) ayrı ayrı olarak incelemiş ve başarı üzerinde olumsuz etkileri olduğunu belirtmişlerdir. Ancak bu çalışmada öğrenme engelleri indeks değişkeni kullanılarak analiz edilmiştir. PISA tarafından tanımlanan indeks değişkeni, öğrenme engellerine ilişkin birçok faktörü içermektedir. Bu faktörler hem öğrencilerden hem de öğretmenlerden kaynaklı engelleri bünyesinde barındırmaktadır. Dolayısıyla belli bir örnekleme ait hem öğrenci hem öğretmen verileri aynı değişken altında yer alacağından, öğretmen ya da öğrenci grubundan kaynaklanacak engeller veya destekler diğer grup tarafından nötrlenmiş olabilir. Bundan hareketle araştırmacılara araçsal güdülenmenin ve öğrenme engeli değişkeninin fen başarısı üzerindeki etkisini inceleyen çalışmalar yürütmesi önerilmektedir.

Fen öğrenmekten zevk alma değişkeninin fen başarısı üzerinde düşük pozitif bir etkisinin olduğu görülmektedir. Öğrencilerin fen öğrenmekten zevk alma düzeyindeki bir standart sapmalı pozitif yönlü değişim .161 standart sapmalı artışa denk gelecektir. OECD (2016) yayınladığı raporda, bu değişkenin fen başarısı üzerinde dolaylı bir etkisi olduğunu belirtmiştir. Aynı zamanda Yetişir ve diğ. (2018) de çalışmalarında pozitif yönde etkisi olduğu sonucuna ulaşmışlardır. Bu sebeple öğrencilere fen öğretiminde onların çalışmaktan mutlu olacağı, zevk alacağı ve aktif katılım sağlayabilecekleri öğrenme ortamları sağlanması tavsiye edilmektedir.

Fen öz yeterliği değişkeninin fen başarısı üzerindeki etkisi ise düşük ve negatif yönlüdür. Yani öğrencilerin fen öz yeterliği düzeyindeki bir standart sapmalı artış .077 standart sapmalı azalmaya sebep olacağı sonucuna ulaşılmıştır. Ancak gerçekleştirilen birçok çalışmada fen öz yeterliği değişkeninin fen başarısı üzerinde pozitif yönde anlamlı (Uğraş, 2018) ve hatta fen başarısını en iyi yordayan değişken olduğu belirtilmiştir (Bircan, 2015; Uzun vd., 2010). Ayrıca Bandura (1994) öz yeterlik inancının bireylerin ulaşacağı başarı düzeyini etkileyen önemli bir unsur olduğunu belirtmiştir. Bu bakımdan çalışma bulguları literatür tarafından desteklenmemektedir. Ancak yapılan bazı çalışmalarda da fen öz yeterliği ve fen performansı arasında ilişki olmadığı da belirtilmektedir (Jamil & Mahmud, 2019). Ayrıca öz yeterliğin başarı üzerindeki etkisinde rol alan aracı değişkenlerin olabileceği de ifade edilmektedir (Doménech-Betoret vd., 2017). Bu çalışmalardan yola çıkarak fen öz yeterliği ve fen performansı arasındaki ilişkiyi incelemeye yönelik daha fazla çalışma yapılmasına ihtiyaç duyulduğu söylenebilir.

Fen başarısını etkilediği düşünülerek modele eklenen bir diğer değişken ise eğitim engeli değişkenidir. Okulun öğretmen yetersizliğini, fiziksel imkân ve kaynaklarını belirten eğitim engeli değişkeninin fen başarısı üzerinde orta düzeyde negatif yönlü bir etkisi olduğu tespit edilmiştir. Öğrencilerin eğitim engelleri bir standart sapma arttığında fen başarısında .339 standart sapma azalmaya sebep olacaktır. Elde edilen bu sonuç literatürle büyük oranda benzerlik göstermektedir. Eğitim engellerinin sistematik olarak öğrenci başarısıyla ilişkisi olduğu ve bu ilişkilerin eğitim açısından oldukça önem taşıdığı alanyazınla desteklenmektedir (Hedges vd., 1994; Greenwald vd., 1996). Ancak Hanushek (1997) belirtilen sonucun tersine öğrenci performansı ile okul kaynakları arasında güçlü ve tutarlı bir ilişki olmadığını çalışmada vurgulamıştır. Hem bu çalışmada hem de yapılmış diğer araştırmaların birçoğunda eğitim engellerinin kaldırılmasının başarıyı arttırdığına yönelik elde edilen bulgular öğrencilere sağlanacak okul imkânlarının da onlar için eğitim elemanı olacağını göstermektedir. Bu sebeple öğrencilerin eğitim-öğretim gördükleri ortamların uygun hâle getirilmesi, öğrenci

performansına olumlu katkılar sağlayacaktır (Al Şensoy & Sağsöz, 2015). Bunun için ise öğretmen, bilgi teknolojisi materyalleri, kütüphane, laboratuvar ve fiziksel alt yapı eksiklikleri giderilmeli ve nitelikleri artırılmalıdır.

Bu çalışmada değişkenlerin fen başarısına yordayıcı etkisinin incelenmesine ek olarak bu etkinin, okullar düzeyinde ve birey düzeyinde fen başarısını etkilediği düşünülen ve fen başarısına ait gözlenen değişkenlerin yer aldığı iki düzeyli ve tek düzeyli modellemede sonuçların ne derece farklılaştığı da incelenmiştir. Modellerden elde edilen parametre ve uyum indeksleri karşılaştırılmadan önce gözlenen değişkenlere ilişkin sınıf içi korelasyon katsayısı incelenmiş çok düzeyli analizlere ihtiyaç duyulduğu tespit edilmiştir. Literatürde hiyerarşik verilerde gerçekleştirilen tek düzeyli analizler göz önüne alınarak bu araştırmada da ayırma veya birleştirme yaklaşımı yürütülerek tek düzeyli modellemeler ve iki düzeyli modelleme karşılaştırılmıştır. Standartlaştırılmış parametreler bakımından faktör yükleri, hata varyansları ve bu parametrelerin standart hataları iki ve tek düzeyli modellerde anlamlı olmadan az düzeyde farklılaşsa da regresyon katsayıları ve standart hataları bakımından anlamlı farklılıklar ortaya çıktığı görülmüştür. Model 1 (iki düzeyli) ve Model 2 (grup düzeyi göz ardı edilen)'de elde edilen regresyon katsayılarında öğrenme engeli değişkeni fen başarısını Model 1'de anlamı olarak tahmin etmiyorken Model 2'de anlamlı tahmin ettiği sonucuna ulaşılmıştır. Bunun sebebi ise okullar içindeki öğrenciler arasındaki homojenliğin yüksek olması ve dolayısıyla regresyon katsayılarına ait standart hataların azalarak parametre tahmininde yanlı kestirimlerin oluşmasıdır (Heck, 2001; Hox vd., 2017; Muthén & Satorra, 1995). Model 1 ve Model 3 (birey düzeyi göz ardı edilen)'te ise regresyon katsayılarının anlamlılığı bakımından benzerlik gösterse de grup içi yordayıcı değişkenlerin gizil değişken üzerindeki etkisinin olduğundan daha yüksek olduğu görülmektedir. Bunun nedeni ise grup içi varyansın analize dâhil edilmemesidir (Heck, 2001; Kaplan & Elliot, 1997). Ayrıca gruplar arası yordayıcı değişkenlerin anlamlılık değerlerinde artış gözlenmiştir. Çünkü birey düzeyindeki verilerin grup düzeyine çekilmesinde veri kaybı yaşanması gerçekleştirilen istatistiksel testin gücünü azaltmıştır (Heck, 2001; Kaplan & Elliot, 1997). Model uyum indeksleri bakımından da üç modelden elde edilen sonuçlar hiyerarşik verilerde iki düzeyli modelin kullanılmasının daha iyi uyuma işaret ettiğini göstermektedir. Çalışmanın analiz karşılaştırmasından elde edilen bulgu ve sonuçlar ışığında gelecekte hiyerarşik verilerde yapılacak çalışmalarda kullanılacak analizler için araştırmacılara fikir sunacağı düşünülmektedir. Özet olarak, hiyerarşik yapıya sahip verilerde tek düzeyli analizlerin kullanılması, söz konusu değişkenler arasındaki ilişkilerle ilgili farklı/hatalı sonuçlar ortaya çıkarmaktadır. Bu durum da hatalı yorum ve çıkarımlara neden olabilmektedir.

Bu çalışmadaki karşılaştırmalar; fen bilişsel alt alanı, öğrenci ve okul yöneticisi anketleri için Türkiye'den elde edilen verilerle sınırlıdır. İleride yapılacak araştırmalarda, analize okuduğunu anlama ve/veya matematik bilişsel alt alanları, farklı ülkeler ve öğretmen anketleri dâhil edilerek daha kapsamlı karşılaştırmalar yapılabilir.

Yazar Katkı Oranı

Yazarlar, çalışmaya eşit oranda katkı sunmuşlardır.

Etik Beyan

Araştırmanın etik olmayan bir sorunu yoktur ve araştırma ve yayın etiği tam olarak gözetilmiştir. Mevcut araştırma için etik kurul onayı Bartın Üniversitesi Sosyal ve Beşerî Bilimler Etik Kurulu tarafından 2020-17 protokol numarası ile verilmiş ve yazarlar araştırma ve yayın etiğine uyulduğunu beyan etmişlerdir.

"Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesinde" yer alan tüm kurallara uyulmuş ve yönergenin ikinci bölümünde yer alan "Bilimsel Araştırma ve Yayın Etiğine Aykırı Eylemlerden" hiçbirisi gerçekleştirilmemiştir.

Çıkar Çatışması

Yazarlar herhangi bir çıkar çatışması beyan etmemektedir.

References

- Acar, T., & Öğretmen, T. (2012). Çok düzeyli istatistiksel yöntemler ile 2006 PISA fen bilimleri performansının incelenmesi. *Eğitim ve Bilim*, 37(163), 178-189.
- Acosta, S. T. & Hsu, H. Y. (2014). Negotiating diversity: An empirical investigation into family, school and student factors influencing New Zealand adolescents' science literacy. *Educational Studies*, 40(1), 98-115. <https://doi.org/10.1080/03055698.2013.830243>
- Aktamış, H., Kiremit, H. Ö., & Kubilay, M. (2016). Öğrencilerin öz-yeterlik inançlarının fen başarılarına ve demografik özelliklerine göre incelenmesi. *Adnan Menderes Üniversitesi Eğitim Fakültesi Eğitim Bilimleri Dergisi*, 7(2), 1-10.
- Al Şensoy, S., & Sağsöz, A. (2015). Öğrenci başarısının sınıfların fiziksel koşulları ile ilişkisi. *Kırşehir Eğitim Fakültesi Dergisi*, 16(3), 87-104.
- Anagün, Ş. S. (2011). PISA 2006 sonuçlarına göre öğretme-öğrenme süreci değişkenlerinin öğrencilerin fen okuryazarlıklarına etkisi. *Eğitim ve Bilim*, 36(162), 84-102.
- Anıl, D. (2009). Uluslararası Öğrenci Başarılarını Değerlendirme Programı'nda (PISA) Türkiye'deki öğrencilerin fen bilimleri başarılarını etkileyen faktörler. *Eğitim ve Bilim*, 34(152), 87-100.
- Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4) Is (pp. 71–81). Academic Press. <https://doi.org/10.1002/9780470479216.corpsy0836>
- Barutçu Yıldırım, F., & Demir, A. (2020). Self-handicapping among university students: The role of procrastination, test anxiety, self-esteem, and self-compassion. *Psychological Reports*, 123(3), 825-843. <https://doi.org/10.1177%2F0033294118825099>
- Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1-48. <https://doi.org/10.18637/jss.v067.i01>
- Beese, J., & Liang, X. (2010). Do resources matter? PISA science achievement comparisons between students in the United States, Canada and Finland. *Improving Schools*, 13(3), 266-279. <https://doi.org/10.1177%2F1365480210390554>
- Bilican Demir, S., & Yıldırım, O. (2021). Indirect effect of economic, social, and cultural status on immigrant students' science performance through science dispositions: A multilevel analysis. *Education and Urban Society*, 53(3), 336–356. <https://doi.org/10.1177/0013124520928602>
- Bircan, H. (2015). *Motivasyon ve bilişsel katılımın fen başarısındaki rolü* [Yayımlanmamış yüksek lisans tezi]. Orta Doğu Teknik Üniversitesi.
- Bussie're, P., Knighton, T., & Pennock, D. (2007). *Measuring up: Canadian results of the OECD PISA study—the performance of Canada's youth in science, reading and mathematics: 2006 first results for Canadians aged 15* (Report No: 590-593). Canadian Ministry of Industry. <http://www5.statcan.gc.ca/bsolc/olc-cel/olc-cel?catno=81-590-X&chprog=1&lang=eng>
- Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2014). *Bilimsel araştırma yöntemleri*. Pegem Akademi.
- Can, S., Somer, O., Korkmaz, M., Dural, S., & Öğretmen, T. (2011). Çok düzeyli yapısal eşitlik modelleri. *Türk Psikoloji Dergisi*, 26(67), 14-21.
- Çoker, E. (2009). *Çok-düzeyle regresyon modelleri ile çok-düzeyle yapısal eşitlik modellerinin uygulamalı karşılaştırılması* [Yayımlanmamış doktora tezi]. Mimar Sinan Güzel Sanatlar Üniversitesi.
- Depaoli, S., & Clifton, J. P. (2015) A Bayesian approach to multilevel structural equation modelling with continuous and dichotomous outcomes. *Structural Equation Modelling: A Multidisciplinary Journal*, 22(3), 327-351. <https://doi.org/10.1080/10705511.2014.937849>
- Doménech-Betoret, F., Abellán-Roselló, L., & Gómez-Artiga, A. (2017). Self-efficacy, satisfaction, and academic achievement: The mediator role of students' expectancy-value beliefs. *Frontiers in Psychology*, 8, 1193. <https://doi.org/10.3389/fpsyg.2017.01193>
- Draper, D. (1995). Inference and hierarchical modelling in the social sciences. *Journal of Educational Statistics*, 20(2), 115-148. <https://doi.org/10.3102%2F10769986020002115>
- Dyer, N. G., Hanges, P. J., & Hall, R. J. (2005). Applying multilevel confirmatory factor analysis techniques to the study of leadership. *The Leadership Quarterly*, 16(1), 149–167. <https://doi.org/10.1016/j.leaqua.2004.09.009>

- Döş, İ., & Atalmış, E. H. (2016). OECD verilerine göre PISA sınav sonuçlarının değerlendirilmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 16(2), 432-450. <https://doi.org/10.17240/aibuefd.2016.16.2-5000194936>
- Fang, Z., & Wei, Y. (2010) Improving middle school students' science literacy through reading infusion. *The Journal of Educational Research*, 103(4), 262-273. <https://doi.org/10.1080/00220670903383051>
- Farmer, G. L. (2000). Use of multilevel covariance structure analysis to evaluate the multilevel nature of theoretical constructs. *Social Work Research*, 24(3), 180-191. <https://doi.org/10.1093/swr/24.3.180>
- Finlayson, M. (2009). *The impact of teacher absenteeism on student performance: The case of the Cobb County School District* [Unpublished master's thesis]. Kennesaw State University.
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education*. McGraw Hill.
- Gardner, P. L. (1975) Attitudes to science: A review. *Studies in Science Education*, 2(1), 1-41. <https://doi.org/10.1080/03057267508559818>
- Greenwald, R., Hedges, L. V., & Laine, R. D. (1996). The effect of school resources on student achievement. *Review of Educational Research*, 66(3), 361-396. <https://doi.org/10.3102%2F00346543066003361>
- Goldstein, H. (1987). *Multilevel models in education and social research*. Oxford University Press.
- Güngör, M. (2019). *Fen motivasyonu ve özyeterliliği modeli'nin ölçme değişmezliğinin incelenmesi: PISA 2015 Türkiye örneği* [Yayınlanmamış yüksek lisans tezi]. Hacettepe Üniversitesi
- Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: A R package for facilitating large-scale latent variable analyses in M plus. *Structural Equation Modelling: A Multidisciplinary Journal*, 25(4), 621-638. <https://doi.org/10.1080/10705511.2017.1402334>
- Hanrahan, M. (1999). Rethinking science literacy: Enhancing communication and participation in school science through affirmational dialogue journal writing. *Journal of Research in Science Teaching*, 36(6), 699-717. [https://doi.org/10.1002/\(SICI\)1098-2736\(199908\)36:6%3C699::AID-TEA7%3E3.0.CO;2-P](https://doi.org/10.1002/(SICI)1098-2736(199908)36:6%3C699::AID-TEA7%3E3.0.CO;2-P)
- Hanson, T. L., Austin, G., & Lee-Bayha, J. (2003). *Student health risks, resilience, and academic performance*. WestEd.
- Hanushek, E. A. (1997). Assessing the effects of school resources on student performance: An update. *Educational Evaluation and Policy Analysis*, 19(2), 141-164. <https://doi.org/10.3102%2F01623737019002141>
- Heck, R. H. (2001). Multilevel modelling with SEM. In J. A. Marcoulides, & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modelling* (pp. 89-127). Lawrence Erlbaum Associates.
- Heck, R. H., & Thomas, S. L. (2015). *An introduction to multilevel modelling techniques: MLM and SEM approaches using Mplus*. Routledge.
- Hedges, L. V., Laine, R. D., & Greenwald, R. (1994). An exchange: Part I: Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher*, 23(3), 5-14. <https://doi.org/10.3102%2F0013189X023003005>
- Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge.
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modelling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Jamil, N. L., & Mahmud, S. N. D. (2019). Self-efficacy relationship on science achievement amongst national secondary school students. *Creative Education*, 10(11), 2509. <https://doi.org/10.4236/ce.2019.1011179>
- Kaplan, D., & Elliot, P. R. (1997). A didactic example of multilevel structural equation modelling applicable to the study of organizations. *Structural Equation Modelling*, 4(1), 1-24. <https://doi.org/10.1080/10705519709540056>

- Kaplan, D., Kim, J. S., & Kim, S. Y. (2009). Multilevel latent variable modelling: Current research and recent developments. In R. E. Millsap, & A. Maydeu-Olivares (Eds.), *The Sage handbook of quantitative methods in psychology* (pp. 592-612). SAGE Publications
- Kaya, V. H., & Doğan, A. (2017). Determination and comparison of Turkish student characteristics affecting science literacy in Turkey according to PISA 2012. *Research Journal of Business and Management (RJBM)*, 4(1), 34-51. <https://doi.org/10.17261/Pressacademia.2017.369>
- Kirbulut, Z. D., & Uzuntiryaki-Kondakci, E. (2019). Examining the mediating effect of science self-efficacy on the relationship between metavariabiles and science achievement. *International Journal of Science Education*, 41(8), 995-1014. <https://doi.org/10.1080/09500693.2019.1585594>
- Kjærnsli, M., & Lie, S. (2011). Students' preference for science careers: International comparisons based on PISA 2006. *International Journal of Science Education*, 33(1), 121-144. <https://doi.org/10.1080/09500693.2010.518642>
- Konishi, C., Hymel, S., Zumbo, B. D., & Zhen Li. (2010). Do school bullying and student—teacher relationships matter for academic achievement? A multilevel analysis. *Canadian Journal of School Psychology*, 25(1), 19–39. <https://doi.org/10.1177/0829573509357550>
- 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(404), 1198–1202. <https://doi.org/10.1080/01621459.1988.10478722>
- Mehta, P. D., & Neale, M. C. (2005). People are variables too: Multilevel structural equations modelling. *Psychological Methods*, 10(3), 259-284. <https://psycnet.apa.org/doi/10.1037/1082-989X.10.3.259>
- Muthén, B. O. (1994). Multilevel covariance structure analysis. *Sociological Methods & Research*, 22(3), 376-398. <https://doi.org/10.1177%2F0049124194022003006>
- Muthén, B., & Satorra, A. (1995). Complex sample data in structural equation modelling. *Sociological Methodology*, 25, 267-316. <https://doi.org/10.2307/271070>
- Noyan, F. (2009). Çok aşamalı yapısal eşitlik modellerinin iş tatmini ile örgütsel bağlılık arasındaki ilişki üzerine bir uygulaması [Yayınlanmamış doktora tezi]. Marmara Üniversitesi
- Organisation for Economic Co-operation and Development. (2015). Summary description of the seven levels of proficiency in science in PISA 2015. <https://www.oecd.org/pisa/test/summary-description-seven-levels-of-proficiency-science-pisa-2015.htm>
- Organisation for Economic Co-operation and Development. (2016). *PISA 2015 results (volume I): Excellence and equity in education*. PISA, OECD Publishing.
- Organisation for Economic Co-operation and Development. (2017). PISA 2015 assessment and analytical framework: Science, reading, mathematics, financial literacy and collaborative problem solving. OECD Publishing. <https://doi.org/10.1787/9789264281820-en>
- 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.
- Özkan, M., Balci, S., Kayan, S., & Is, E. (2018). Quality of educational resources: A comparative evaluation of schools that joined PISA 2015 from Turkey and Singapore. *International Education Studies*, 11(4), 132-143. <https://doi.org/10.5539/ies.v11n4p132>
- Raudenbush, S. W. (1995). Reexamining, reaffirming, and improving application of hierarchical models. *Journal of Educational Statistics*, 20(2), 210-220. <https://doi.org/10.3102%2F10769986020002210>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods (2nd ed.)*. Sage Publications.
- Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T. Y., & Lee, Y. H. (2007). A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching*, 44(10), 1436-1460. <https://doi.org/10.1002/tea.20212>
- Tabachnick, B., & Fidell, L. S. (2012). *Using multivariate statistics*. Pearson.
- Taş, U. E., Arıcı, Ö., Ozarkan, H. B., & Özgürlük, B. (2016). *Uluslararası öğrenci değerlendirme programı 2015 ulusal rapor*. http://odsgm.meb.gov.tr/test/analizler/docs/PISA/PISA2015_Ulusal_Rapor.pdf

- Török, L., Szabó, Z. P., & Tóth, L. (2018). A critical review of the literature on academic self-handicapping: Theory, manifestations, prevention and measurement. *Social Psychology of Education: An International Journal*, 21(5), 1175–1202. <https://doi.org/10.1007/s11218-018-9460-z>
- Uğraş, M. (2018). Yedinci sınıf öğrencilerinin motivasyon ile öz yeterlik inançlarının fen bilimleri dersindeki başarılarıyla ilişkisinin incelenmesi. *Bingöl Üniversitesi Sosyal Bilimler Enstitüsü Dergisi (BUSBED)*, 8(16), 495-508. <https://doi.org/10.29029/busbed.453792>
- Usta, G. H., & Çıkırcı Demirtaşlı, N. R. (2014). PISA 2006 sınavı sonuçlarına göre Türkiye'deki öğrencilerin fen bilimleri okuryazarlığını etkileyen duyuşsal faktörler. *Eğitim Bilimleri Araştırmaları Dergisi*, 4(2), 93-107. <https://doi.org/10.12973/jesr.2014.42.6>
- Uzun, G., & Çokluk Bökeoğlu, Ö. (2019). Akademik başarının okul, aile ve öğrenci özellikleri ile ilişkisinin çok düzeyli yapısal eşitlik modellemesi ile incelenmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 52(3), 655-685. <https://doi.org/10.30964/aeubfd.525770>
- Uzun, N. B., Gelbal, S., & Öğretmen, T. (2010). TIMSS-R fen başarısı ve duyuşsal özellikler arasındaki ilişkinin modellenmesi ve modelin cinsiyetler bakımından karşılaştırılması. *Kastamonu Eğitim Dergisi*, 18(2), 531-544.
- Woods-McConney, A., Oliver, M. C., McConney, A., Schibeci, R., & Maor, D. (2013). Science engagement and literacy: A retrospective analysis for indigenous and non-indigenous students in Aotearoa New Zealand and Australia. *Research in Science Education*, 43(1), 233–252. <https://doi.org/10.1007/s11165-011-9265-y>
- Yetişir, M. İ., Batı, K., Kahyaoglu, M., & Birel, F. K. (2018). Dezavantajlı öğrencilerin fen okuryazarlık performanslarının duyuşsal özellikleriyle ilişkisinin incelenmesi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 51(1), 143-158. <https://doi.org/10.30964/aeubfd.405014>
- Yore, L., Hand, B., Goldman, S., Hildebrand, G., Osborne, J., Treagust, D., & Wallace, C.S. (2004). New directions in language and science education research. *Reading Research Quarterly*, 39(3), 347-352. <https://doi.org/10.1598/RRQ.39.3.8>
- Yu, C. H. (2012). Examining the relationships among academic self-concept, instrumental motivation, and TIMSS 2007 science scores: A cross-cultural comparison of five East Asian countries/regions and the United States. *Educational Research and Evaluation*, 18(8), 713-731. <https://doi.org/10.1080/13803611.2012.718511>
- Yuan, K. H., & Bentler, P. M. (2007). 3. Multilevel covariance structure analysis by fitting multiple single-level models. *Sociological methodology*, 37(1), 53-82. <https://doi.org/10.1111%2Fj.1467-9531.2007.00182.x>

Appendix**Appendix 1**

```

syntax <- mplusObject(
  TITLE = 'Cok Duzeyli YEM1;',
  VARIABLE = 'NAMES ARE SCH STU EE1-EE8 OE1-OE10 ZA1-ZA5 AG1-AG4 OY1-OY8 FB1-FB10
    MEE MOE MZA MAG MOY MFB;
  USEVARIABLES ARE SCH FB1-FB10 MEE MOE MZA MAG MOY;
  WITHIN= MZA MAG MOY;
  BETWEEN= MEE MOE;
  CLUSTER IS SCH;';
  ANALYSIS = 'TYPE IS TWOLEVEL;
    ESTIMATOR IS MLF;';
  MODEL = '%WITHIN%
    LFB BY FB1-FB10;
    LFB ON MZA MAG MOY;
    %BETWEEN%
    GLFB BY FB1-FB10;
    GLFB ON MEE MOE;';
  OUTPUT = 'standardized modindices;';
  rdata=FEN1
)
FEN_fit1 <- mplusModeler(syntax, modelout = 'MODEL1.inp', run=1L)
When the data which disregards group levels by disaggregation is used;
Syntax2 <- mplusObject(
  TITLE = 'Tek Duzeyli YEM2;',
  VARIABLE = 'NAMES ARE SCH STU EE1-EE8 OE1-OE10 ZA1-ZA5 AG1-AG4 OY1-OY8 FB1-FB10
    MEE MOE MZA MAG MOY MFB;
  USEVARIABLES ARE FB1-FB10 MEE MOE MZA MAG MOY;';
  ANALYSIS = 'ESTIMATOR IS MLF;';
  MODEL = '
    LFB BY FB1-FB10;
    LFB ON MZA MAG MOY MEE MOE;';
  OUTPUT = 'standardized modindices;';
  rdata=FEN1
)
FEN_fit2 <- mplusModeler(syntax2, modelout = 'MODEL2.inp', run=1L)
When the data which disregards individual level by aggregation is used;
Syntax3 <- mplusObject(
  TITLE = ' Tek Duzeyli YEM3;',
  VARIABLE = 'NAMES ARE SCH STU FB1-FB10 MEE MOE MZA MAG MOY MFB;
  USEVARIABLES ARE FB1-FB10 MEE MOE MZA MAG MOY;';
  ANALYSIS = 'ESTIMATOR IS MLF;';
  MODEL = '
    LFB BY FB1-FB10;
    LFB ON MZA MAG MOY MEE MOE;';
  OUTPUT = 'standardized modindices;';
  rdata=FEN3
)
FEN_fit3 <- mplusModeler(syntax3, modelout = 'MODEL3.inp', run=1L)

```