How Classroom Assessment Affects Science and Mathematics Achievement?: Findings from TIMSS 2015

Melek Gülşah Şahin[a,*], Nagihan Boztunç Öztürk[b]

Abstract

In this study, it is aimed to examine the effect of classroom assessment on science and mathematics achievements. For this purpose, hierarchical linear modeling (HLM) is performed using variables of like learning science/maths, engage teaching in science/maths, confidence in science/maths, and home resources for learning variables at the student level, and experience, education level, homework, and assessment at the teacher level. The sample of the study consists of 4th grade students who participated in TIMSS 2015 in Turkey. According to the findings; 36% of variance in science achievement, and 40% of variance in mathematics achievement are due to variability between classes. In a random coefficient model, all student variables were found to be statistically significant predictors of science and mathematics achievement. Among these variables, the greatest effect size is self-confidence variability. Only the teacher variables are added according to the Means as the outcome model; the teacher's experience and emphasis to national achievement tests of monitoring students' progress had a statistically significant effect on science and mathematics achievement. Finally, according to the intercept and slopes of the outcomes model, the most important variable is the emphasis to national achievement tests of monitoring students' progress in both science and mathematics.

Keywords: Classroom assessment, student characteristics, science achievement, mathematics achievement, hierarchical linear modeling

Introduction

With today's increasing need for qualified human capital, priority is given to increasing the quality of education. Considering the system approach in education, classroom assessment plays an important role in determining the quality of the input, process and output elements of education. Classroom assessment is an instructor's way of gathering information about what students have learnt. They can then use them to reach important decisions about the students' grades, the content of future courses, and the revision of the structure or content of a course or curriculum (Brookhart, 1999). In addition, they might also be used to gain information about students' specific weaknesses and special instructional needs and to identify any concepts or procedures which may need to be re-taught or reviewed (Yu, 2012). To date, teachers have used assessments for different purposes and this has influenced teaching and learning differently. These are identified as summative assessments and formative assessments. Summative assessments are usually used to meet the requirements of graduation and academic placement, while formative assessments are mainly administered in order to garner advice on pedagogical decision-making processes and curriculum change. For instance, high-stakes standardised tests can be given as examples of summative assessment, while the tests developed by teachers or assigned homework and projects are examples of formative assessment (Liang, 2010).

It has become a matter of course that classroom assessment also plays an important role in student learning, rather than just as a means of showing their performance, by adopting a constructivist approach that deeply affects many educational applications in educational systems. In this context, the use of the formative assessments in education has increased. Both formative and summative assessments affect students' perceptions of classroom assessments. The perceptions of students against classroom assessments are related to the notion of the classroom assessment atmosphere. Classroom assessment is divided into learning and performance-focused environments (Buldur, 2014; Brookhart, 1997; İlhan, 2017; McMillan & Workman, 1998). While learning-focused assessments are dominated by an evaluation understanding aimed for students' learning; ultimately in performance-based assessment, their role is to understand and assess student learning (Buldur & Doğan, 2014; İlhan, 2017).

In general, assessment practices used by teachers play a decisive role in students' perceptions of the classroom assessment environment (İlhan, 2017). In addition, teachers communicate with students during the classroom assessment process. Teachers are in contact with students when determining the assessment criteria, for the giving of feedback, explaining results, and so on. It is inevitable, therefore, that the students' assessment perceptions, teacher's assessment tools in the process of formative assessment, assessment criteria, and teacher's communication with students during assessment impacts student motivation towards learning and of their self-efficacy, which is also positively related to motivation (Shunk, 1996). Studies in the literature suggest that classroom assessment and learning motivation are interrelated (Brookhart, 1997; Brookhart & Durkin, 2003; Crooks, 1988; Gronlund, 2006; Harlen & Crick, 2003; Rodríguez, 2004). The concept of classroom assessment is also related to success (Black & William, 1998; Brookhart & DeVoge, 1999; Brookhart & Durkin, 2003). As a result, it can be stated that the concept of classroom assessment is related to attitude, motivation, and teacher-student communication. The study of Brookhart and Durkin (2003) found that student performance was related to how teacher assessment was practised, student self-efficacy, motivation and effort.

In the current study, the impact of the teacher's classroom assessment practices on students' mathematics and science performances was investigated with hierarchical linear modeling (HLM) by using the TIMSS (Trends in International Mathe-

[a]Corresponding Author: Melek Gülşah Şahin, Gazi University, Faculty of Education, Department of Educational Sciences, Ankara, Turkey.
E-mail: melekgulsah@gmail.com
[b] Nagihan Boztunç Öztürk, Hacettepe University, Faculty of Education, Department of Educational Sciences, Ankara, Turkey,
E-mail: nagihanboztunc@gmail.com

matics and Science Study) 2015 Turkish sample. The TIMSS is a survey study implemented since 1995 in the Netherlands, which allows within-country and between-country comparisons conducted by the International Association for Evaluation of Educational Assessment [IEA]. The TIMSS study is conducted every four years in order to determine the trends of mathematics and science achievements of 4th and 8th grade students. In TIMSS, the educational and emotional factors effective in determining students' success tendencies have also been being investigated. While Turkey participated in this practice only at the 8th grade level in 1999 and 2007, its participation was extended to both 8th and 4th grade levels in 2011 and 2015. In the literature, it can be widely observed that by using the TIMSS data for Turkey, multi-level analyses of the teacher characteristics associated with mathematics and science achievements have been investigated (Abazaoğlu & Taşar, 2016; Akyüz, 2006, 2014; Akyüz & Berberoğlu, 2010; Atar, 2014; İpekçioğlu-Önal, 2015; Tavşancıl & Yalçın, 2015; Yavuz, Demirtaşlı, Yalçın, & İlgün Dibek, 2017). However, there have only been limited studies that have used the TIMSS Turkish sample and investigated classroom assessment variables which play an important role in student learning and in structuring course content (Nitko, 1989), and which has a relationship with student achievement (Akyüz, 2006; Arıkan, 2017; Yıldırım, Demirtaşlı, & Akbaş, 2012). In the international literature, Liang (2010) and Rodriguez (2004) studied Brookhart and Durkin's (2003) classroom assessment concept and related student variables with HLM analysis in different datasets.

In the current study, teacher's experience, education level, and classroom assessment were variables selected at the teacher level. The teachers' self-perceived assessment skills which were closely related to the teachers' assessment skills according to Gullikson (1984), were influenced mainly by teaching experience and professional training. Frey and Schmitt (2010) found that some teachers' preferences for classroom assessment differed according to their teaching experiences and their time to practice differed according to gender. The experience, and education level variables which relate to both teacher assessment practices and TIMSS success (Abazaoğlu & Taşar, 2016; Akyüz, 2006; Atar, 2014; Atar & Atar, 2012; Bietenbeck, 2011; İpekçioğlu-Önal, 2015; Kaya, 2008) were also selected within the context of the current study.

The literature is based on the determination of variables at the student level. As mentioned, studies have shown that classroom assessment practices are related to motivation. This criterion was taken into account especially while determining the variables. The variable of students' self-confidence is also included in the study, which relates to the motivation of pupils at the student level and strengthens pupils' participation in lessons. It can be correlated with the success in the research conducted on international exams; and it expresses the confidence of the students and their self-efficacy (Akyüz, 2014; Arıkan, van de Vijyer & Yağmur, 2016; Atar & Atar, 2012; Butakor, 2015; Chen, 2013; Demir & Kılıç, 2010; Kaya, 2008; Ker, 2015; Woo & Henfield, 2016; Yavuz et al., 2017). The variable of liking learning mathematics and science, which was related to the students' attitude and achievement, was also included (Erdinç-Akan, 2016; İpekçioğlu-Önal, 2015; Mo, Singh, & Chang, 2012; Yavuz et al., 2017). In addition, the variable of engaging teaching in science/maths, in which the expressions related to the teacher-student communication took place, and which was related to the classroom assessment variable, was also included. Finally, the variable of home resources (Akyüz, 2006; İpekçioğlu-Önal, 2015; Kaya, 2008; Tavşancıl & Yalçın, 2015), which is related to the students' success was included in the study. Other variables that

might be associated with the teacher's classroom assessment practices at the class level were also included in the study.

The concept of classroom assessment affects both the quality of in-class instruction and the motivational beliefs of the students (Brookhart, 1999). Given the importance of classroom assessment practices, it is important to increase the quality of teaching, determine the relevant variables, provide motivation for the students, and to increase student achievement. For the assessment of student learning, the classroom assessment variables which form the basis for many decisions concerning the structuring of courses and curriculum by using these assessments' results and which influence student learning; and examining the effects of the relevant variables in the TIMSS 4th Grade Turkish sample on student achievement constituted the significance of the study. In Rodriguez's (2004) research, the effect of different teacher variables on the mathematical achievement of students was examined. In the current study, both mathematical and science achievement are examined as dependent variables. For this reason, two separate multilevel models have been analysed for the two dependent variables.

In addition, determining the effect sizes of the variables closely related to success and presenting suggestions about the related variables is also important to this study. The main research question of the study is, 'What are the interrelationships of teacher assessment practices, student characteristics, and student achievement?'

In this context, answers were sought to the following sub-questions within the scope of the study.

> 1. Does the maths and science achievement of 4th graders participating in TIMSS 2015 differ significantly between classes?

> 2. If there is a significant difference between mathematics and science achievement of 4th grade students participating in TIMSS 2015, what are the student characteristics that explain this difference?

> 3. If there is a significant difference between mathematics and science achievement of 4th grade students participating in TIMSS 2015, what are the teacher characteristics and classroom assessment variables that explain this difference?

> 4. If there is a significant difference between mathematics and science achievement of 4th grade students participating in TIMSS 2015, what are the student characteristics associated with the teacher characteristics and classroom assessment variables explaining this difference?

Method

This study examined the effects of student-level and teacher-level factors on 4th graders' science and mathematics achievement. Due to the nested data structure, multilevel analysis was used with a largescale dataset. The effects of the teacher characteristics, and the classroom-assessment variables, along with the student characteristics on science and mathematics achievement at the elementary level were explored for the case of Turkey. Hierarchical Linear Modeling (HLM) was used in the analyses of the data from TIMSS 2015 which were obtained from the official TIMSS website.

Sample

The sample of the study consists of 4th grade Turkish students participating in TIMSS 2015. A total of 260 schools and 6,456 students participated within Turkey (Milli Eğitim Bakanlığı [Ministry of National Education], 2016). TIMSS uses a two-phase, random sampling design. In the first stage, some schools are taken as a sample, while determining one or more of the entire student classes from each of the sampled schools forms the second stage (IEA, 2016). It should be noted that the teachers in each country's sample are those selected within the student samples. The teachers were not selected in the sample and did not form a representative sample from each country. In other words, the teachers were not explicitly selected data and therefore no representative teacher sample was created from each country (Kaya, 2008).

Variables

Questionnaires are administered at the student, teacher, and school levels in TIMSS research. Depending on the purpose of the study; some variables are included in the student and/or teacher questionnaires. The dependent variables of the current study were the science and mathematics achievements, reported as five plausible values (PV) (ASSSCI01-ASSSCI05 for science, and ASMMAT01-ASMMAT05 for mathematics) that were estimated for each student in the sample. The plausible values are the assigned not individual scores of the students, based on similar students from whom the properties of the population can be accurately predicted – such as students with similar response patterns and experiences in the sample population. Final reliability values of .81 and .87 were found for the TIMSS 2015 science and mathematics field (IEA, 2016), respectively.

Student-Level Variables

In the student level, the student-level variables of liking learning science/maths (ASBGSLS, ASBGSLM; nine items [e.g., *I enjoy learning science/maths*] with Cronbach α_{SCI}= .76; Cronbach α_{MATH}= .84), engaging teaching in science/maths (ASBGESL, ASBGEML; ten items [e.g., *I know what my teacher expects me to do*] with Cronbach α_{SCI}= .77; Cronbach α_{MATH}= .73), confidence in science/maths (ASBGSCG, ASBGSCM; nine items [e.g., *I usually do well in science/maths*] with Cronbach α_{SCI}= .77; Cronbach α_{MATH}= .82), and the variable of home resources for learning (ASBGHRL; five items with Cronbach α= .74) were determined as the independent variables. The index values of the variables at the student level were used in the analysis. The TIMSS 2015 context questionnaire scaling was conducted by using ConQuest 2.0 software (Wu, Adams, Wilson, & Haldane, 2007 as cited in IEA, 2016).

The index of home resources for learning consists of the number of books in the home (students), number of home study supports (students), number of children's books in the home (parents), highest level of education of either parent (parents), and highest level of occupation of either parent (parents) variables.

Teacher-Level Variables

For the teachers level, the variables of experience (ATBG01), education level (ATBG04), frequency of science/maths homework assignments (ATBS06A, ATBM07A), time spent on homework (ATBS06B, ATBM07B), homework (ATBS06CA-ATBS06CC, ATBM07CA-ATBM07CC; *How often do you do the following with the science/mathematics homework assignments for this class?*), and the variable of as-sessment (ATBS07A-ATBS07C, ATBM08A-ATBM08C; How much emphasis do you place on the following sources to monitor students' progress in science/mathematics?) were adopted as the independent variables. The variables of educational level, homework, and assessment weren't used directly in the analysis. In the data, the educational level variable had a seven-point scale ranging from (1) 'Did not complete upper secondary' to (7) 'Doctor or equivalent'. In the current study, this variable was recoded as 1= Has a Master's or Doctoral degree, and 0= Does not have a Master's or Doctoral degree. In the data, the homework variable had a three-point scale ranging from (1) 'Always or almost always' to (3) 'Not always or almost always'. In the current study, this variable was recoded as 1= Always or almost always, and 0= Not always or almost always. In the data, the assessment variable had a three-point scale ranging from (1) 'Major emphasis' to (3) 'Little or no emphasis'. In the current study, this variable was recoded as 1= Major emphasis, and 0= Not major emphasis.

The student-level and teacher-level variables used in the current study and the codes representing the variables are displayed in Table 1.

Table 1. *Variables used at student and teacher level*

Variables	Codes
Student-Level	
Like learning science/math	LIKELEARN
Engaging teaching in science/math	ENGTEACH
Student confident in science/math	SELFCON
Home resources for learning	HOMERES
Teacher-level	
Experience	EXP
Education level	EDULEV
How often science/math home work assignment	FRQ-HWORK
Time spent on homework	TIMEHWORK
Homework-Correct assignments and give feedback to students	H-FEEDBACK
Homework-Discuss the homework in class	H-DISCUSS
Homework-Monitor whether or not the homework was completed	H-MONITOR
Assessment-Assessment of students' ongoing work	A-ONWORK
Assessment-Classroom tests (for exp., teacher-made or textbook tests)	A-CLASSTEST
Assessment-National or regional achievement tests	A-ACHTEST

Analysis of Data:

The correlation between teacher and student-related variables were investigated in the study. HLM was used as an inferential statistical procedure. The data were analysed with respect to the missing values, outliers, and univariate and multivariate normality in the preliminary analysis. The data were updated by assigning the mean values of the missing data at the student level. Additionally, missing data at the teacher level was removed from the data. Mahalanobis distance was calculated for the determination of the outliers. The sample size was determined using the 50/20 rule, with 50 groups and 20 individuals per group (Hox, 2010). After the analysis of the missing data, outliers,

and the sample sizes, the final sampling consisted of 5,163 students and 177 teachers in the science achievement data, and 4,817 students, and 166 teachers in the mathematics achievement data. Before starting the HLM analysis, it was examined as to whether or not the assumptions were violated. For this reason, homogeneity assumptions of normality and variances were tested. As a result, it was observed that the assumptions were not violated for both science or mathematics success models.

This study made use of the HLM as a more appropriate method for the analysis of the TIMSS data in which students are nested within classes (Raudenbush & Bryk, 2002). The HLM includes some of the deviations among students in the same class and looks at the differences in the students' science/mathematics achievements as well as the classes. The data were analysed using four models in order to answer the sub-problems of the study. These models are; Random Effects ANOVA Model, Random Co-efficient Model, Means as an Outcomes Model, and Intercept and Slopes as an Outcomes Model. In the analysis of the data, the HLM 7.03 program was used.

1. *Random Effects ANOVA Model*: A random ANOVA model was run with the HLM in order to determine within-classrooms and between-classrooms segments of the variance of the 4th grade science and mathematics achievements. None of the predictor variables were used at the student level and/or at the teacher level. Therefore, this model, which is the simplest among the other models, is named as the Null model. Fundamentally, this model divides the entire variance into student-level and teacher-level variance.

2. *Random Coefficient Model*: The second sub-research question examined the effects of the selected student-level variables on the science and mathematics scores of 4th grade students. In order to answer this, a random coefficient model was run with the HLM. The student-level variables of liking learning, engaging teaching in science/maths lessons, self-confidence, and home resources were introduced at the student level (level-1). None of the predictor variables were used at the teacher level.

3. *Means as an Outcome Model*: The third sub-research question investigated the impacts of the selected teacher-level variables on the science and mathematics scores of 4th grade students. In order to answer this, a means as outcomes model was run with the HLM. The teacher-level variables of experience, gender, age, education level, frequency of science/maths homework assignment, time spent on homework, homework, and assessment were introduced at the teacher level (level-2). None of the predictor variables were used at the student level.

4. *Intercept and Slopes as an Outcomes Model*: The fourth sub-research question examined the effects of the selected teacher-level variables on the science and mathematics scores of 4th grade students. In order to answer this, a fully conditional model with all student and teacher-level variables was run with the HLM.

In order to reduce the high correlation between the level 1 and level 2 variables and cross-level interactions in the analyses, while the variables in level 1 were centred around the group mean, the variables in level 2 were centred around the grand mean. Categorical variables were left uncentred.

Information about the direct and indirect associations of student characteristics and classroom assessment variables with student achievement is displayed in Tables 3, 4, and 5. In these tables, the results of the conventional statistical testing for determining the significance of the effects were reported. Yet, the variables might gain statistical significance due to the large sample size even if their substantive significance level is minimal. Therefore, the HLM results are represented in terms of effect sizes (*ES*) (Von Secker & Lissitz, 1999). Effect sizes are standard deviation (*SD*) units which permit the comparison of results with different measurements. These can be defined as the *SD* change in the independent variable linked to 1 *SD* change in an independent variable. The size interpretation of the effect sizes are as follows: 0.5 *SD* effect size or greater are regarded as large; moderate for 0.3 or greater but less than 0.5 *SD*; small for 0.1 or greater but less than 0.3 *SD*; and lastly < 0.1 *SD* effect sizes are considered as trivial (Rosenthal & Rosnow, 1984).

Findings

The study investigates the effects of four variables at the student level and ten variables at the teacher level on 4th grade science and mathematics achievement on the TIMSS 2015. First random effects ANOVA model was analysed with HLM, and the findings are presented in Table 2.

As can be seen from Table 2, the average class mean science and mathematics achievement were statistically different from zero (Y_{00}= 481.059, *p*< .001; Y_{00}= 480.774, *p*< .001). For science achievement, the intra-class correlation (ICC) was calculated as .36, showing that the variability of science achievement between classes is 36%. The ICC was calculated as .40 for mathematics achievement, indicating that the variability of mathematics achievement between classes is 40%. To put it in another way, the differences among 4th grade students in the same classroom caused 64% of the variance in science achievement and 60% of the variance in mathematics achievement. In order to explain this difference, four student variables were added to Model-1. The results of the Random-Coefficient Model are given in Table 3.

Table 2. *Results of random effects ANOVA model for science and mathematics achievement*

	Science			Mathematics		
	Coeff.	**SE**	***p*-Value**	**Coeff.**	**SE**	***p*-Value**
Intercept, Y_{00}	481.059	4.56	<.001	480.774	4.90	<.001
Between-class variability (τ)	3068.049			3720.504		
Within-class variability across all students (σ^2)	5524.650			5619.557		
Intra-class correlation (ICC)	.36			.40		

Table 3. *Results random coefficient model for estimating the effects of student-level variables on science and mathematics achievement*

Fixed effect	Science				Mathematics			
	Coeff.	SE	*p*-Value	Effect Size	Coeff.	SE	*p*-Value	Effect Size
Intercept, Y_{00}	481.043	4.56	<.001	---	480.752	4.90	<.001	---
LIKELEARN, Y_{10}	2.678	1.09	.022	.036	3.761	0.99	<.001	.050
ENGTEACH, Y_{20}	2.813	0.84	.001	.038	1.997	0.69	.004	.027
SELFCON, Y_{30}	11.292	0.70	<.001	.152	15.501	0.68	<.001	.207
HOMERES, Y_{40}	10.601	0.80	<.001	.143	10.185	0.82	<.001	.136

Random effect	Science		Mathematics	
	Variance	*p*-Value	Variance	*p*-Value
Mean achievement (u_0)	3109.093	<.001	3779.799	<.001
Level-1 effect, (r)	4329.558		3916.143	
LIKELEARN	18.243	.295	48.972	.003
ENGTEACH	17.403	.020	8.362	.137
SELFCON	10.613	.284	5.571	.470
HOMERES	30.903	<.001	26.773	.001

The results of random coefficient model showed that overall mean science and mathematics achievement across classrooms were statistically different from zero (Y_{00}= 481.043, *p*< .001; Y_{00}= 480.752, *p*< .001). The average effect of all student-level variables on both science and mathematics achievement was statistically significantly positive. When random effect in Table 3 is considered, the variance (u_0) of the differences between the corrected classroom averages and the general average was found to be statistically meaningful (*p*< .001). This result can be interpreted as meaningful differences exist among the corrected classroom averages. The effect of the engaging teaching variable in science achievement, liking learning variable in mathematics achievement, and the home resources for learning variable in both science achievement and mathematics achievement differ across classrooms.

According to Table 3, it seems that the greatest effect size in both science and mathematics achievement is the var-

iable of self-confidence. The effect size for the variable of self-confidence is likely to have an increase of roughly .152 standard deviation in students' mean science achievement which results from an increase of 1 standard deviation in the self-confidence. The effect size of the variable of self-confidence is likely to have an increase of roughly .207 standard deviation in students' mean mathematics achievement which results from an increase of 1 standard deviation in the self-confidence.

The within-classroom variability of science achievement decreased from 5524.650 to 4329.558 when these variables are added to the model as level-1 variables for TIMSS 2015, which indicates that these variables explain 21.64% of variance in student science achievement performance. With certain variables included in the model as level-1 variables for TIMSS 2015, a decline for the within-classroom variability of mathematics achievement was seen from 5619.557 to 3916.143, which indicates that 30.31% of with-

Table 4. *Results means as outcome model for estimating the effects of teacher-level variables on science and mathematics achievement*

Fixed effect	Science				Mathematics			
	Coeff.	SE	*p*-Value	Effect Size	Coeff.	SE	*p*-Value	Effect Size
Intercept, Y_{00}	457.141	29.30	<.001	---	425.974	33.46	<.001	---
EXP, Y_{01}	**2.774**	**0.38**	**<.001**	.050	**3.096**	**0.43**	**<.001**	.050
EDULEV, Y_{02}	0.603	20.84	.977	---	-0.506	22.73	.982	---
FRQ-HWORK, Y_{03}	5.799	7.99	.469	---	4.440	6.75	.511	---
TIMEHWORK, Y_{04}	-6.256	5.75	.278	---	-6.043	6.54	.357	---
H-FEEDBACK, Y_{05}	-13.250	11.12	.235	---	-18.138	13.13	.169	---
H-DISCUSS, Y_{06}	5.171	8.39	.538	---	6.471	8.96	.471	---
H-MONITOR, Y_{07}	8.445	12.08	.486	---	10.090	14.18	.478	---
A-ONWORK, Y_{08}	21.724	14.55	.137	---	34.046	18.06	.061	---
A-CLASSTEST, Y_{09}	-10.852	9.89	.274	---	12.479	11.55	.282	---
A-ACHTEST, Y_{010}	**16.304**	**8.27**	**.050**	.293	**21.178**	**9.10**	**.021**	.345

Random effect	Science		Mathematics	
	Variance	*p*-Value	Variance	*p*-Value
Mean achievement (u_0)	2213.939	<.001	2589.588	<.001
Level-1 effect, (r_{ij})	5524.702		5619.640	

Table 5. *Results of intercept and slopes as outcomes model for estimating direct and indirect effects of teacher-level variables on science and mathematics achievement*

Fixed effect	Science				Mathematics			
	Coeff.	SE	p-Value	Effect Size	Coeff.	SE	p-Value	Effect Size
Mean achievement								
Intercept, Y_{00}	472.630	5.48	<.001	---	467.702	6.20	<.001	---
EXP, Y_{01}	*2.909*	*0.37*	*<.001*	*.052*	*3.274*	*0.41*	*<.001*	*.053*
A-ACHTEST, Y_{02}	*16.346*	*7.45*	*.030*	*.293*	*23.500*	*8.33*	*.005*	*.382*
LIKELEARN, Y_{10}	*2.602*	*1.24*	*.039*	*.035*	*3.43*	*1.49*	*.023*	*.046*
LIKELEARN EXP Y_{11}*	*-0.182*	*0.09*	*.046*	*.013*	*-0.221*	*0.11*	*.047*	*-.017*
LIKELEARN*A-ACHTEST Y_{12}	0.401	1.59	.802	---	0.568	2.12	.789	---
ENGTEACH, Y_{20}	*2.710*	*1.12*	*.016*	*.036*	*1.440*	*0.95*	*.129*	*---*
ENGTEACH * EXP Y_{21}	-0.144	0.08	.068		-0.120	0.07	.067	---
ENGTEACH * A-ACHTEST Y_{22}	0.105	1.61	.948	---	0.830	1.34	.536	---
SELFCON, Y_{30}	*11.668*	*0.97*	*<.001*	*.157*	*16.259*	*1.02*	*<.001*	*.217*
SELFCON * EXP Y_{31}	-0.070	0.07	.284		*-0.128*	*0.06*	*.040*	*-.015*
SELFCON*A-ACHTEST Y_{32}	-0.699	1.34	.603	---	-1.375	1.22	.260	---
HOMERES, Y_{40}	*9.262*	*1.12*	*<.001*	*.125*	*9.457*	*1.31*	*<.001*	*.126*
*HOMERES * EXP Y_{41}*	*0.324*	*0.09*	*.001*	*.030*	*0.349*	*0.08*	*<.001*	*.033*
HOMERES*A-ACHTEST Y_{42}	2.703	1.67	.108	---	1.604	1.77	.376	---

Random effect	Science		Mathematics	
	Variance	p-Value	Variance	p-Value
Mean achievement (u_0)	2232.494	<.001	2641.458	<.001
Level-1 effect, (r_{ij})	4398.593		3950.289	
LIKELEARN slope, u_1	███	█	32.606	<.001
ENGTEACH slope, u_2	15.686	.028	███	█
HOMERES slope, u_4	18.661	.005	14.119	.026

in-classroom variability is justified by these variables in the performance of student mathematics achievement.

In Model-3, as seen in Table 4, only the teacher-level variables were added to Model-1.

The results of means as outcome model showed that overall mean science and mathematics achievement across classrooms were statistically different from zero (Y_{00}= 457.141, *p*< .001; Y_{00}= 425.974, *p*< .001). The averages effects of the teacher experience and the assessment of national or regional achievement tests have significant positive effects on both science and mathematics achievement. For this reason, other variables, which didn't have a statistically significant effect on achievement, were excluded from the upcoming model.

In this model, the effect size was calculated only on variables with statistically significant effect on achievement. As can be seen in Table 4, the greatest effect size in both science and mathematics achievement success is for the national or regional achievement variable. The effect size increased by around .293 standard deviation in students' mean science achievement, resulting from an increase of 1 *SD* in self-confidence; and increased by around .345 standard deviation in students' mean mathematics achievement, resulting from an increase of 1 *SD*.

The between-classroom variability of science achievement decreased from 3068.049 to 2213.939 when these varia-

bles are added to the model as level-2 variables for TIMSS 2015, which indicates that these variables explain 27.84% of variance in student science achievement performance. With certain variables included in the model as level-2 variables for TIMSS 2015, a decline for the between-classroom variability of mathematics achievement was seen from 3720.504 to 2589.588, which indicates that 30.40% of between-classroom variability is justified by these variables in the performance of students' mathematics achievement.

The Chi-square values obtained from the analysis shows that these explanatory variables at the classroom level do not explain the whole variability of the constants of science and mathematics achievement (*p*< .05). According to these results, other variables can be added to the model.

In Model 4, intercept and slopes as outcomes model, all significant student-level variables and teacher-level variables are included in the model. The findings are presented in Table 5.

According to Table 5, the results of intercept and slopes as outcomes model showed that overall mean science and mathematics achievement across classrooms were statistically different from zero (Y_{00}= 472.630, *p*< .001; Y_{00}= 467.702, *p*< .001).

Science achievement was significantly affected by all 4th grade student-level and teacher-level variables. Based on

the cross-level interactions, there was no significant interaction in TIMSS 2015, except for between the variables of like learning and teacher's experience, and between the variables of students' home resources for learning and teacher's experience in science achievement (Υ_{11}= -0.182, p< .05; Υ_{41}= 0.324, p< .05). The direction of those interactions differed in the examined model. According to the model, a statistically negative relationship was seen between the variables of teacher's experience and like learning (Υ_{11}= -0.182, p< .05). This finding shows that as the teachers' experience increases, the students' level of liking science decreases. When the variables of home resources and teacher experience are considered, a statistically positive relationship was found (Υ_{41}= 0.324, p< .05). In this case, it is seen that the students who have generally more books at home and whose parents have high levels of educational and professional status are taught by teachers with more experience.

Mathematics achievement at the 4th grade was significantly affected by all student-level variables, except for the engaging learning variable, and all teacher-level variables. Based on the cross-level interactions, there are three significant interactions in mathematics achievement. Considering among which variables there are cross-level interactions, it can be seen that there are statistically meaningful interactions between the variables of liking learning, self-confidence and home resources at the student level, and the variable of teacher experience at the teacher level. Accordingly, a statistically negative relationship was seen between the teacher experience variable and the variables of like learning (Υ_{11}= -0.221, p< .05) and self-confidence (Υ_{31}= -0.128, p< .05). This finding shows that as teachers' experience increases, students' level of liking mathematics and their self-confidence decreases. When the variables of home resources and teacher experience are considered, a statistically positive relationship was seen between these two variables (Υ_{41}= 0.349, p< .05).

The variable with the highest impact value in both science and mathematics is the frequency of emphasis of a teacher place on the national or regional achievement tests to monitor students' progress in science/mathematics variable. Accordingly, it is expected that a standard deviation increase in the frequency of the A-ACHTEST variable will increase the student's average science achievement by about .29 *SD*, and the mathematical achievement by .38 *SD* when the practical significance of the relevant variable is assessed.

Considering the stated variance values, it is determined that when the teacher experience variable and the other variables in the national or regional test model are fixed, it explains 27.23% of the variance in science achievement, and 29.00% of the variance in mathematics achievement.

Discussion

The current study aimed to investigate the relationship between classroom assessment practices, student characteristics, and achievement, and certain variables included in the student and teacher questionnaire at the 4th grade level of TIMSS 2015. Determination of the variables included in the research was made by examining the literature concerning the effect of classroom assessment on mathematical success. As a result of examining two levels of HLM analysis, four models (ANOVA model, random coefficient regression model, means as outcome model, and model with output of constant and slope coefficients) were analysed.

The randomised one-way ANOVA model examined whether or not the science and mathematical achievements differed between classrooms. In TIMSS 2015, the between-classroom differences within the scope of the current study were 36% for science achievement and 40% for mathematics achievement. In other words, 64% of the science achievement and 60% of the mathematics achievement seemed to result from students in the same class. For the within-classroom student differences, all the predictors of liking learning, engaging teaching, self-confidence and home educational resources variables were found to significantly affect both mathematics and science achievement.

Student-Level Variables:

Model 2 was formed by adding the student variables to the null model, and was found to be significant both for science and mathematics achievement. All the predictors addressed at this level for both science and mathematics had a statistically significant and positive affect on student success. When the effect sizes are considered, the variable that most affected both science and mathematics success was self-confidence. The effect size on science achievement was .152, while the effect size on mathematics achievement was .207. Although the highest effect size in the model belonged to the self-confidence variable, it appears that this variable had a low effect on both science and mathematics success (Rosenthal & Rosnow, 1984). Looking at the amount of in-class explained variance of student characteristics, it was seen that science achievement was 21.64% and mathematics achievement was 30.31%.

It was determined that the students' liking of learning mathematics or science variable was positively related to the students' achievement. In the literature, liking mathematics or science (Neale, 1969; Osborne, Simon & Collins, 2003) seems to be an effective component in determining the attitude towards mathematics and science. Liking learning of mathematics and science could be related to positive attitude towards mathematics and science. Students who develop a positive attitude towards a course will be motivated towards lessons and will make more effort to succeed. While Yavuz et al.'s (2017) study was conducted with 8th graders indicated that in TIMSS 2007, students who liked learning mathematics gained higher mathematical success; in TIMSS 2011, the opposite results were seen. However, Yavuz et al. (2017) noted that this could be ignored when considering the effect sizes of the mathematics success in TIMSS 2011. Erdinç-Akan (2016), in his study conducted with 8th graders on the TIMSS 2011 Turkish sample, obtained results which showed that students who liked learning science had higher science achievement, as well. In their courses, teachers may be recommended to employ appropriate activities according to the level of each student, which are aimed at positively changing the students' attitudes towards science courses, in order to improve their motivation and raise their achievement motive.

In the current study, the self-confidence variable was found to be positively correlated with student success. The self-confidence variable was also an indicator of attitudes towards the course; in other words, if it is high, it could lead to a positive attitude. In addition, the highly confident students' motivation towards lessons and their participation in classroom activities would also increase. In their study with the TIMSS 2007 Turkish sample, Atar and Atar (2012) found a meaningful and positive relationship between students' self-confidence and their success. Besides, among the variables addressed at the student level, the effect size of the self-confidence variable was found to

be the highest for achievement. Yavuz et al. (2017) pointed out that self-confidence as a variable had the highest effect size on mathematical success in TIMSS 2011. However, in TIMSS 2007, the self-confidence variable was not found to be meaningful, and the suggested reason for this was that the student-centred approach was not fully adopted at that time.

The current study found a meaningful and positive relationship between the home resources variable and student achievement at the student level. The home resources variable consisted of educational level of the family, as well as economic indicators. In general, higher levels of education can lead to careers in higher-wage professions, higher socioeconomic status, and greater home resources. Family income has also been shown to have a powerful influence on students' achievement in reading and mathematics (Baker, Goesling, & Letendre, 2002; Dahl & Lochner, 2005; Mullis, Martin, Foy, & Arora, 2012). The home background of the learner is not only related to student success, but is also related to attitudes towards learning. Children who have greater home educational resources can be considered more advantageous than others. Teachers should therefore be aware of the educational resources of their students and their learning environment to ensure that students with low educational home resources are not disadvantaged (İpekçioğlu-Önal, 2015). In his study of TIMSS 2011, Akyüz (2014) found that 8th graders' maths self-confidence and home resources variables had a positive and significant effect on the success of each country.

It can be said that the greater the number of books in the household, the existence of Internet connection and/or the availability of private rooms to study, which are included among the home resources variable, enhance the resources that students' can obtain academically and thereby the provision of an appropriate environment. Therefore, these sub-branches of the home resources variable play an important role in student success. Also, the importance given to a child's education can increase in line with the educational level of the parents. Consequently, the resources and opportunities for children can vary in proportion. In addition, as the parents' educational level increases, the stimuli that they are able to present to their children can provide enrichment in a parallel way.

Another variable in which a significant and positive correlation with the success was obtained at the student level, was the engaging in teaching variable. In this part, the variables related to teacher-student communication were included. Lay (2017) explored the predictive effects of student engagement in science lessons and their attitudes toward science in science achievement among Southeast Asian 8th graders in TIMSS 2015, and found that the engaging in teaching variable was a predictor of success. Teachers should be especially encouraged to communicate effectively with students both inside and outside the classroom, to help students with clear statements when they experience difficulty, and to encourage their students to succeed.

Teacher-Level Variables:

In Model-3, which was formed by adding only the teacher characteristics to the original null model, the current study examined whether or not the teacher experience, graduation status, homework and evaluation variables predicted science and mathematics achievement. As a result of the analysis, it was determined that the variables of experience and assessment in national or regional achievement tests were statistically significant. When the effect sizes were examined, it was seen that the assessment in nation-

al or regional achievement tests variable had a low-level effect size on science achievement (ES= .295) and a moderate level of effect size on mathematics achievement (ES= .345). When looking at the percentages of the explained variances of Model 3, it was seen that it explained 20.26% of the variance of science achievement, and 29.66% of the variance of mathematics achievement.

The experience variable addressed at the teacher level was found to have a meaningful and positive relationship with success. Atar and Atar (2012) found that as teachers' experience increased, the science achievement of students increased in their study based on TIMSS 2007 data. As the teachers' experiences increased, they will likely be more practical and knowledgeable in getting information about the recognition of their students. In addition, when experienced, teachers may become more successful in predicting the mistakes of students and guessing where they may have been challenged. Teachers can easily provide the necessary conditions for a healthier establishment of teacher-student communication.

Emphasis on national or regional achievement tests to monitor students' progress in science and mathematics, which were among the variables studied at the teacher level, seemed to have had a positive impact on the students' success. This might be as a consequence of TIMSS and the importance given to national examinations in Turkey. As known, while TIMSS aimed to determine the curricula success, the Turkish national examination at the end of secondary school (ÖBBS) included questions parallel to the curricula. In this case; an inference could be made about how effective the national examinations are on classroom practices.

The variable of time spent on homework at the teacher level was not associated with student success. Data on the length of time the teacher gave students for their assigned homework was obtained from the time spent on homework variable. The relationship strength between time spent on homework and success was dependent on the students' age; and was found to be stronger in post-compulsory secondary education than for compulsory education, while it was almost zero in primary education (Cooper, 1989; Cooper, Robinson, & Patall, 2006). In the study of Rodriguez (2004) using TIMSS data for 8th graders, the students' time spent on homework and success were not found to be significantly correlated. In the study of Akyüz (2013) using TIMSS 2007 data for 8th graders, he found no significant relationship between students' mathematical success and their time spent on homework.

The current study found no significant relationship between the frequency of variables under the title of 'homework' at the teacher level and success. These variables were correct assignments and give feedback to students, discuss the homework in class, and monitor whether or not the homework was completed. In the interpretation of the 'homework' and success relationship, consideration should be given to the content of the homework assigned by the teachers as well as the suitability of the subject's content to the level of the students, and what kinds of work it included. It should also be remembered that associating homework with its instruction will have an impact on the students' learning and success. Arıkan (2017) did not find any significant relationship between the practice frequency and success in his study of TIMSS 2011 8th grade data using the variables under the homework title from the teacher's questionnaire. The authors of the current study recommend that teachers assess these variables more extensively by considering the nature of their assignments' quality rather than considering frequency

of practice. Moreover, in TIMSS 1999 and TIMSS 2007, Yıldırım et al. (2012) found that the usage frequency of the correct assignments and give feedback to students variable did not change significantly over time. Teachers can provide feedback on their students' homework outside of class hours. By regularly evaluating the process, the development of students can be more reliably monitored.

Finally, there was no significant relationship found between student success and giving importance to the variables of assessment of students' ongoing work, and classroom tests (e.g., teacher-made or textbook tests) presented under the 'assessment' title at the teacher level. In their study conducted with 8th grade data from TIMSS 1999 and TIMSS 2007, Yıldırım et al. (2012) found that teachers paid more attention to tests made at the national level and to their own classroom examinations when assessing student achievement. It is important for teachers to attach significance to the assessment process in determining student success in terms of their progress by identifying their deficiencies and errors and by correcting them. However, suitability of this assessment to students is important for the realisation of learning. Again, the content of tests and what kind of test items included are also important when considering the teacher gives importance to the classroom tests variable. The assignments and tests in which achievements of lower cognitive skills such as knowledge and understanding are tested will be insufficient for learners to acquire high-level skills. Yıldırım et al. (2012) stated that teachers generally preferred to use test books or teacher-made tests with multiple-choice questions in their classroom practice.

Another variable to consider is the variety of the instruments that are preferred over assessment. Using multiple-choice tests predominantly in the assessment of students prevents students from being recognised as versatile. McMillan (2000) emphasised that teachers should use several assessment tools and approaches with different priorities instead of one single measurement tool. For this reason, it is advisable to use different assessment tools in the classroom. In addition, the level of reliability and validity of measurement tools also plays a role in determining their relationship with success. When considering the assessments by teachers with tests that were inappropriate or that did not serve their purpose, it is important for teachers to question the decisions made with regard to these tools rather than simply questioning their application frequency.

In conclusion, the findings of the current study were as expected for the student-level variables, whilst not as expected for the teacher-level variables. Especially on the emphasis of some sources to monitor students' progress in science and mathematics, the researchers were surprised that only one variable was found to be significant in predicting student achievement, and that this variable was national or regional achievement tests. For this reason, similar variables should be examined in experimental design studies.

References

Abazaoğlu, İ., & Taşar, M. F. (2016). Fen bilgisi öğretmen özelliklerinin öğrenci fen başarısı ile ilişkisi: TIMSS 2011 verilerine göre bir durum analizi [Relations of characteristics of science teachers and students achievement in science: A case analysis according to TIMSS 2011 data]. *İlköğretim online, 15*(3), 922-945.

Akyüz, G. (2006). Türkiye ve Avrupa Birliği ülkelerinde öğretmen ve sınıf niteliklerinin matematik başarısına etkisinin incelenmesi [Investigation of the effect of teacher and class characteristics on mathematics achievement in Turkey and European Union Countries]. *İlköğretim Online Dergisi, 5*(2), 61-74.

Akyüz, G. (2013). Öğrencilerin okul dışı etkinliklere ayırdıkları süreler ve matematik başarısı arasındaki ilişkinin incelenmesi [Evaluation of the relationship between the time spent on the leisure activities of students' and their achievements in mathematics]. *Elektronik Sosyal Bilimler Dergisi, 12*(46), 112-130.

Akyüz, G. (2014). The effects of student and school factor on mathematics achievement in TIMSS 2011. *Education and Science, 39*(172), 150-162.

Akyüz G., & Berberoğlu, G. (2010). Teacher and classroom characteristics and their relations to mathematics achievement of the students in the TIMSS. *New Horizons in Education, 580*(1), 77-95.

Altınok, H. (2004). Öğretmenlerinin fen öğretimine yönelik tutumlarına ilişkin öğrenci algıları ve öğrencilerin fen bilgisi dersine yönelik tutum ve güdüleri [Teacher candidates' evaluation of their teaching competencies]. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 26*, 1-8.

Arıkan, S. (2017). TIMSS 2011 verilerine göre Türkiye'deki ev ödevi ve matematik başarısı arasındaki ilişki [The relationship between homework and mahematics achievement in Turkey according to TIMSS 2011]. *International Journal of Eurasia Social Sciences, 8*(26), 256-276.

Arıkan, S., van de Vijver, F. J. R., & Yağmur, K. (2016). Factors contributing to mathematics achievement differences of Turkish and Australian students in TIMSS 2007 and 2011. *EURASIA Journal of Mathematics, Science and Technology Education, 12*(8), 2039-2059.

Atar, H. Y. (2014). Öğretmen niteliklerinin TIMSS 2011 fen başarısına çok düzeyli etkileri [Multilevel effects of teacher characteristics on TIMSS 2011 science achievement] [Özel Sayı]. *Eğitim ve Bilim Dergisi, 39*(172), 121-137.

Atar, H. Y., & Atar, B. (2012). Türk eğitim reformunun öğrencilerin TIMSS 2007 fen başarılarına etkisinin incelenmesi [Examining of the effects Turkish education reform on students' TIMSS 2007 sicence achivments]. *Kuram ve Uygulamada Eğitim Bilimleri Dergisi, 12*(4), 2621-2636.

Baker, D. P., Goesling, B., & Letendre, G. K. (2002). Socioeconomic status, school quality, and national economic development: A cross-national analysis of the "Heyneman-Loxley effect" on mathematics and science achievement. *Comparative Education Review, 46*(3), 291-312.

Bietenbeck, J. C. (2011). *Teaching practices and student achievement: Evidence from TIMSS* (Unpublished Master's thesis). Centro de Estudios Monetarios y Financieros (CEMFI) University, Madrid, Spain.

Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(12), 7-74.

Brookhart, S. M. (1997). A theoretical framework for the

role of classroom assessment in motivating student effort and achievement. *Applied Measurement in Education, 10*(2), 161-180.

Brookhart, S. M. (1999). *The art and science of classroom assessment: The missing part of pedagogy. ASHE-ERIC Higher Education Report (Vol: 27, No: 1).* Washington, DC: The George Washington University, Graduate School of Education and Human Development

Brookhart, S. M., & DeVoge, J. G. (1999). Testing a theory about the role of classroom assessment in student motivation and achievement. *Applied Measurement in Education, 12*(4), 409-425.

Brookhart, S. M., & Durkin, T. D. (2003). Classroom assessment, student motivation, and achievement in high school social study classes. *Applied Measurement in Education, 16*(1), 27-54.

Buldur, S. (2014). The investigation of the relationship between the students' perceptions about the classroom assessment environment and their achievement-goal orientations: Gender Perspective. *Education and Science, 39*(4), 213-225.

Buldur, S., & Doğan, A. (2014). Adaptation of the students' perceptions of the science and technology course classroom assessment environment scale (SP-CAES) into Turkish. *Education and Science, 39*(176), 199-211.

Butakor, P. K. (2015). *Multilevel modeling of factors that influence mathematics achievement in Ghana: A secondary analysis of TIMSS 2007 and 2011* (Unpublished Doctoral dissertation). University of Alberta, Edmonton, Alberta, Canada.

Chen, Q. (2013). Using TIMSS 2007 data to build mathematics achievement model of fourth graders in Hong Kong and Singapore. *International Journal of Science and Mathematics Education, 12*(6), 1519-1545.

Cooper, H. (1989). *Homework.* White Plains. NY: Longman

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A synthesis of research, 1987–2003. *Review of Educational Research, 76*(1), 1-62.

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research, 58*(4), 438-481.

Dahl, G. B. & Lochner, L. (2005). The impact of family income on child achievement, NBER Working Papers 11279, National Bureau of Economic Research, Inc.

Demir, İ., & Kılıç, S. (2010). Using Pisa 2003, examining the factors affecting students' mathematics achievement. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 38*(4), 44-54.

Erdinç-Akan, O. (2016). *TIMSS 2011 8. sınıf öğrencilerinin fen başarısı ile ilişkili öğrenci ve öğretmen niteliklerinin bilişsel alanlara göre incelenmesi: İki düzeyli hiyerarşik lineer model analizi [The investigation of the relationship between the characteristics of students and teachers and TIMSS 2011 8th grade students' achievement in respect to the cognitive domains: A two level hierarchical linear model analysis]* (Unpub-

lished Master's thesis). Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara, Turkey.

Frey, B. B., & Schmitt, V. L. (2010). Teachers' classroom assessment practices. *Middle Grades Research Journal, 5*(3),107-117.

Gronlund, N. E. (2006). *Assessment of student achievement* (8th ed.). Boston: Pearson.

Gullikson, A. R. (1984). Teacher perspectives of their instructional use of tests. *Journal of Educational Research, 77*(4), 244–248.

Harlen, W., & Crick, R. D. (2003). Testing and motivation for learning. *Assessment in Education: Principles, Policy & Practice, 10*(2), 169-207.

Hox, J. (2010). *Multilevel analysis techniques and applications.* Netherlands: Lawrence Erlbaum Associates Publishers.

IEA. (2016). *Methods and procedures in TIMSS 2015.* Chestnut Hill: International Association for the Evaluation of Educational Achievement (IEA).

İlhan, M. (2017). Öğrencilerin sınıf değerlendirme atmosferine ilişkin algılarının okula yönelik tutumları üzerindeki yordayıcı rolü [The predictive role of students' perceptions of classroom assessment environment on their attitudes towards school]. *Kastamonu Eğitim Dergisi, 25*(1), 111-128.

İpekçioğlu-Önal, S. (2015). *TIMSS 2011 cross country comparisons: Relationship between student- and teacher-level factors and 8th grade students' science achievement and attitude toward science* (Unpublished Doctoral dissertation). Middle East Technical University, Ankara, Turkey.

Kaya, S. (2008). *The effects of student-level and classroom-level factors on elementary students' science achievement in five countries* (Unpublished Doctoral thesis). College Education, Florida State University, Florida, USA.

Ker, H. W. (2015). The impacts of student-teacher-and school-level factors on mathematics achievement: An exploratory comparative investigation of Singaporean students and the USA students. *Educational Psychology, 36*(2), 254-276.

Lay, Y. F. (2017). The predictive effects if engagement in science lessons and attitudes toward science on southeast Asian grade 8 students' science achievement in TIMSS 2015. *The Eurasia Proceedings of Educational & Social Sciences (EPESS), 6*, 142-152.

Liang, X. (2010). Assessment use, self-efficacy and mathematics achievement: comparative analysis of PISA 2003 data of Finland, Canada and the USA. *Evaluation & Research in Education, 23*(3), 213-229.

McMillan, J. H. (2000). Fundamental assessment principles for teachers and school administrators. *Practical Assessment, Research &Evaluation, 7*(8).

McMillan, J. H., & Workman, D. J. (1998). *Classroom assessment and grading practices: A review of the literature.* Metropolitan Educational Research Consortium, Richmond, VA. ERIC Document Reproduction Service No: ED453263.

Milli Eğitim Bakanlığı. (2016). *TIMSS 2015 Ulusal matematik ve fen Bilimleri ön raporu 4. ve 8. sınıflar.* Ankara: Milli Eğitim Bakanlığı Ölçme, Değerlendirme ve Sınav Hizmetleri Genel Müdürlüğü.

Mo, Y., Singh, K., & Chang, M. (2013). Opportunity to learn and student engagement: a Hlm study on eight grade science achievement. *Education Research Policy & Practise, 12*(1), 3-19.

Mullis, I. V. S., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 International results in mathematics.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Neale, D. C. (1969). The role of attitudes in learning mathematics. *The Arithmetic Teacher, 16*(8), 631-640.

Nitko, A. J. (1989). Designing tests that are integrated with instruction. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed., pp. 447-474). New York: American Council on Education & Macmillan.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: a review of the literature and its implications. *International Journal of Science Education, 25*(9), 1049-1079.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods.* London: Sage.

Rodriguez, C. M. (2004). The role of classroom assessment in student performance on TIMSS. *Applied Measurement in Education, 17*(1), 1-24.

Rosenthal, R., & Rosnow, R. (1984). *Essentials of behavioral research: Methods and data analysis.* New York: McGraw-Hill.

Schunk, D. H. (1996, April). *Self efficacy for learning and performance.* Paper presented at the Annual Conference of the American Educational Research Association, New York.

Tavşancıl, E., & Yalçın, S. (2015). A determination of Turkish students' achievement using hierarchical linear models in trends in international mathematics-science study (TIMSS) 2011. *Anthropologist, 22*(2), 390-396.

von Secker, C. E., & Lissitz, R. W. (1999). Estimating the impact of instructional practices on student achievement in science. *Journal of Research in Science Teaching, 36*(10), 1110-1126.

Woo, H., & Henfield, M. S. (2016). Student and teacher factors' impact on fourth grade students' mathematics achievement: An HLM analysis of TIMSS 2007. *Journal of Mathematics Education, 9*(1),69-87.

Yavuz, H. Ç., Demirtaşlı, R. N., Yalçın, S., & İlgün Dibek, M. (2017). Türk öğrencilerinin TIMSS 2007 ve 2011 matematik başarısında öğretmen ve öğrenci özelliklerinin etkileri. *Eğitim ve Bilim, 42*(189), 27-47.

Yıldırım, Ö., Demirtaşlı, N. Ç., & Akbaş, U. (2012). The opinions of mathematics teachers on homework and in-class assessment: TIMSS 1999 and TIMSS 2007 periods. *Education & Science, 37*(163), 126-142.

Yu, V. (2012). *Six ways continuous assessment can help students to learn*. New York. Knewton.