International Journal of Progressive Education, Volume 15 Number 2, 2019

Investigation of the Mediator Variable Effect Using BK, Sobel and Bootstrap Methods (Mathematical Literacy Case)

Selda Örs Özdil [i]
Ankara University

Ömer Kutlu [ii]
Ankara University

Abstract

This study aimed to compare different mediation analysis methods (BK, Sobel, and bootstrapping) based on single mediation models for groups of different sizes. For this purpose, the PISA 2012 data for Turkey were used. In order to compare the mediation analysis methods, 4,848 students from Turkey that participated in PISA 2012 were divided into sample groups of 100, 200, 500 and 1,000 individuals. Among the mediation analysis methods discussed within the scope of the research, the BK method was implemented assisted by a regression analysis while for the remaining two methods, SPSS macros were utilized. For the analysis, syntax files were created to be run on SPSS. The results of the analysis of single mediation models revealed that the mathematics anxiety variable mediated the relationship between classroom climate and mathematical literacy. According to the analyses based on all three methods, it was observed that the standard error value increased as the sample group became smaller. Although the standard errors of the Sobel test and bootstrap method were close to each other in large study groups, the former produced less erroneous results in large samples whereas the latter yielded more reliable results in smaller samples.

Keywords: Mediator variable, mediation effect, Sobel test, bootstrap, BK method, PISA

DOI: 10.29329/ijpe.2019. 189.3

----

[i] **Selda Örs Özdil**, Assist. Prof. Dr., Ankara University, Measurement and Evaluation

Correspondence: seldaors85@gmail.com

[ii] **Ömer Kutlu,** Instructor Dr., Ankara University, Measurement and Evaluation

INTRODUCTION

The concept of mediation is used to indicate that the effect of one or more independent variables (X) is transferred by a third variable(s) to a dependent variable (Y). Numerous studies in the literature have examined not only direct effects but also other relationships considered to have indirect effects. In cases where there are indirect effects, there is a third variable called the mediator variable, which facilitates the relationship between two variables (MacKinnon, Fairchild, & Fritz, 2007). The mediator variable is very useful in providing an understanding of the mechanism by which a cause (independent variable) has an effect on a result (dependent variable) (Fairchild, & MacKinnon, 2009). Therefore, a mediator analysis tries to define the mediation process in which the effect is moved from an independent variable to a dependent variable (Muller, Judd, & Yzerbyt, 2005). Mediation hypotheses seek answers to how an independent variable (X) affects a dependent variable (Y) through one or more interacting variable(s) or mediator variable(s) (M) and the direction of this effect. In this process, models with one mediator variable are defined as simple/single mediation models (Baron & Kenny, 1986; MacKinnon et al., 2007; Preacher, & Hayes, 2008). Figure 1 presents a diagram of the single mediation model (Baron, & Kenny, 1986; Frazier, Tix, & Barron, 2004; Kenny, Kashy, & Bolger, 1998; MacKinnon et al., 2007; Preacher, & Hayes, 2008; Wu, & Zumbo, 2007).

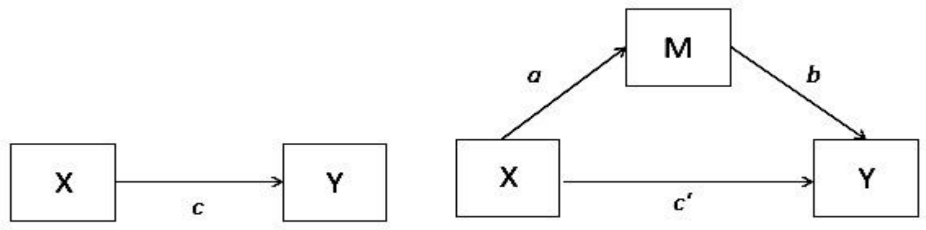

Figure 1. Single Mediation Model

In Figure 1, a causal relationship between the independent variable X and the dependent variable Y is defined and the total effect of X on Y is shown by the coefficient c. In this figure, coefficient a refers to the effect of X on the mediator variable M; coefficient b indicates the effect of M on Y except for the partial effect of X; and coefficient c is the effect of X on Y under the mediation of M (Hayes, 2013; MacKinnon et al., 2007; Preacher, & Hayes, 2008).

In order to estimate the coefficients in the defined model, basic regression equations (1), (2) and (3) are used (Hayes, 2013; MacKinnon et al., 2007; MacKinnon, Lockwood, Hoffman, West, & Sheets, 2002; Preacher, & Hayes, 2004).

$$Y = i_1 + c X + e_1 \quad (1)$$

$$Y = i_2 + c' X + b M + e_2 \quad (2)$$

$$M = i_3 + a X + e_3 \quad (3)$$

The coefficient c in this equation and in Figure 1 shows the total effect, c 'coefficient shows the direct effect and ab coefficient indicates the indirect effect. In this case, the total effect of X on Y will be equal to the sum of the direct and indirect effects. This is represented by the following mathematical equation:

$$c = c' + ab \quad (4)$$

Following this step, the mediation effect can be calculated using one of the following two equations (Hayes, 2009; MacKinnon et al., 2007; MacKinnon & Dwyer, 1993).

$$ab = c - c' \quad (5)$$

$$c' = c - ab \quad (6)$$

Studies by Judd, & Kenny (1981) and Baron, & Kenny (1986) in the field of social psychology prompted many other researchers to utilize mediation models in later studies (as cited in Burmaoglu, Polat, & Meydan, 2013). It is noteworthy that in the 1990s, there was a remarkable increase in the attempts to compare existing mediation analysis methods and develop alternative methods to determine the mediator variable effect (Cheung, & Lau, 2008; Frazier et al., 2004; Hayes, 2009; Hayes, & Preacher, 2014; MacKinnon et al., 2002; MacKinnon, & Dwyer, 1993; MacKinnon, Warsi, & Dwyer, 1995; Shourt, & Bolger, 2002; Taylor, MacKinnon, & Tein, 2008). The most comprehensive research on this subject belongs to MacKinnon et al. (2002), who investigated 14 different mediation analysis methods used in studies in the literature conducted in various disciplines. This methodological diversity in the literature also indicates that there is no clear consensus between disciplines concerning how to determine the mediator variable effect.

Mediation hypotheses are generally tested according to the Baron and Kenny (BK) method, and a partial or full mediation decision is made according to the result of this test. In the BK method, mediation relationships are established in four steps (with three regression equations). Baron, & Kenny (1986) explained these steps as follows:

1. Variable X significantly predicts variable Y (path c).

2. Variable X significantly predicts M (path a).

3. When the effect of variable X is controlled, variable(s) M significantly predicts Y (H0: b=0).

4. When the effect of variable M is controlled, there is a significant decrease in the relationship between X and Y or the relationship between these two variables is no longer significant (H0: c'=0).

According to this method, the greater the reduction in coefficient c, the greater the degree of mediation. Coefficient c' being zero or too close to zero indicates the presence of a mediator variable, and a smaller decrease in coefficient c’ (without approaching zero) suggests that there may be more than one mediator variable. As a result, Baron and Kenny’s approach makes a distinction between a full/excellent mediation (all effect of X on Y is through M) and partial mediation (only part of the effect of X on Y is through M). When the effect of M is controlled, if the relationship between X and Y completely disappears, then the data confirm the full mediation hypothesis, and the relationship is still present but significantly reduced, this supports the partial mediation hypothesis (Pardo & Moran, 2013).

Kenny et al. (1998) reconsidered the causal step approach and suggested that this method does not directly predict the size of the indirect effect (ab) or provide standard errors for the confidence interval values generated for the interpretation of the significance of the indirect effect; it was rather the process of testing each of the a, b and c coefficients individually. Zhao, Lynch, & Chen (2010) stated that the magnitude of mediation should be evaluated starting with the size of the indirect effect (ab), not the lack of a direct effect (c'), and it is not sufficient to know the statistical significance of coefficients c and c’ to determine whether they are actually different; instead, a comparison should be made between these coefficients. However, studies adopting the causal step approach generally do not test the significance of indirect effects in the mediation model. In addition, some of the disadvantages of this method have been previously reported. For example, in their simulation study including different sample sizes, MacKinnon et al. (2002) found that the BK method caused a type I error and the statistical power of the test was low in all conditions. Other approaches to testing mediation hypotheses focus on the product term ab value (this value is logically equal to the difference between the total effect and the direct effect), rather than individual paths in the mediation models. The Sobel test (Sobel, 1982), which is based on the product of coefficients a and b and also known as the multiplication of coefficients, is another method that is most commonly used in the literature (MacKinnon et al., 2002).

The Sobel test involves the multiplication of a and b coefficient estimates and determining the ratio of the resulting value to standard error. Numerous formulas have been proposed to estimate this standard error; however, the differences between them do not often have a significant effect on the test results (MacKinnon et al., 2002; Preacher & Hayes, 2004, 2008). Sobel (1982) proposed the use of the following formula:

$$z = \frac{ab}{\sqrt{b^2 s_a^2 + a^2 s_b^2}} \quad (7)$$

where coefficient a refers to the path between the independent variable and mediator variable, S_a is the standard error of this path (coefficient), b represents the path between the mediator variable and the dependent variable, and S_b is the standard error of path b. The result of this equation is the Z-score of the mediation effect. This score is used to determine whether the mediation effect is statistically significant through the use of probabilities corresponding to a standard normal distribution. If z-score is greater than 1.96, the mediation effect is interpreted to be statistically significant at the .05 level (MacKinnon et al., 2002; Mallinckrodt, Abraham, Wei, & Russell, 2006).

Studies investigating mediation analysis suggest that a multiplication result of two normally distributed variables is not normally distributed, and that the sampling distribution of ab multiplication can only be normal in large samples. Therefore, researchers have criticized the use of standard normal distribution to determine the probability value of the indirect effect and showed that the distribution of the ab product tends to be asymmetric. As a result of this asymmetry, the statistical power of the Sobel test in small samples is lower compared to the methods that attempt to correct this asymmetry (MacKinnon et al., 2002; MacKinnon et al., 1995; Mallinckrodt et al., 2006; Kenny et al., 1998). In order to overcome this problem, some authors (Preacher, & Hayes, 2004, 2008; Shrout, & Bolger, 2002) suggested using the bootstrap method.

Bootstrapping is a non-parametric resampling method and differs from other mediation methods in that it does not require the normality assumption of sampling distribution to test mediation. Bootstrapping is a computationally intensive method, which involves multiple data resampling processes and estimation of the indirect effect in each resampled data set. By repeating this process thousands of times, an empirical approach to ab sampling distribution is created and then used to estimate the confidence intervals of the indirect effect. Shrout, & Bolger (2002) explained the steps of the bootstrap percentile method in examining the mediation effect as follows:

1. In an original data set consisting of N observations, a desired number of bootstrap samples are created by randomly replacing observations.

2. For each bootstrap sample, a, b and ab are calculated and the results are saved.

3. Steps 1 and 2 are repeated j times.

4. The distribution of the estimates is examined, and if α=0.5, ab values and confidence intervals for the 2.5 and 97.5 percentiles of the distribution are determined.

Shrout, & Bolger (2002) determined that the bootstrap method was strong when the sample distribution of the mediation effects was non-zero or skewed. Cheung, & Lau (2008) expanded the simulation study of MacKinnon et al. (2002) and reported that bootstrapping could produce better results than the Sobel test. The authors also suggested that the bootstrap method was particularly useful when there was no information on the distribution or when the assumptions of distribution were violated. Hayes (2009) stated that bootstrapping had the highest power and provided the best type I error control in small samples. In a sample size of 60, Mallinckrodt et al. (2006) did not observe a statistically significant mediating effect using the BK method, but this effect was clearly revealed by the bootstrap method.

Although in recent years different methods have been developed for the identification of mediation effect in mediation models and examined in simulation studies, there is no definite agreement on the conditions in which these methods can be used or the limitations and advantages of each method. It is also noteworthy that the comparison of the methods used to determine the mediating effects is usually performed based on artificial (simulative) data. Furthermore, to the best of our knowledge, no study has been undertaken in Turkey to examine the use of different methods for mediation analysis. Therefore, it is considered important to investigate the mediator variable effect in an established single mediation model using the BK, Sobel and bootstrap methods and compare the efficiency of these methods in different sample sizes. By comparing the methods designed to determine the mediation effect in different situations, this study is expected to contribute to the accumulation of theoretical knowledge. Furthermore, the current research differs from most related previous studies in that it used real data sets, rather than artificial data to examine the mediation analysis methods and compare the results, which is considered to be another significant contribution to the literature regarding mediation tests. It is hoped that the results of the research will guide researchers in selecting the appropriate method to test the mediation effect in different group sizes.

The aim of this research was to compare the BK, Sobel and bootstrap mediation analysis methods in sample groups of different sizes using single mediation models based on the PISA 2012 mathematical literacy data for Turkey. In line with this purpose, the following research questions were constructed:

1. In the single mediation model for the classroom climate, mathematics anxiety, and mathematical literacy variables, does mathematics anxiety have a mediating effect on the whole group and sample groups of different sizes according to the BK method?

2. In the single mediation model for the classroom climate, mathematics anxiety, and mathematical literacy variables, does mathematics anxiety have a mediating effect on the whole group and sample groups of different sizes according to the Sobel test?

3. In the single mediation model for the classroom climate, mathematics anxiety, and mathematical literacy variables, does mathematics anxiety have a mediating effect on the whole group and sample groups of different sizes according to the bootstrap method?

METHOD

Research Model

This study had a basic (theoretical) research design to compare different methods for determining the effect of the mediator variable in mediation models using different sample sizes and contribute new data to the literature. The main purpose of basic research is to add new insights to the existing information (Karasar, 2008).

Study Group

In line with the general purpose of the research, the population of the study comprised Turkish students that participated in PISA 2012, and the sample consisted of 4,848 students selected from 965,736 students in the 15-year-old age group enrolled in grades 7 or higher in Turkey (Ministry of National Education, 2013). Within the scope of the study, study groups of 100, 200, 500 and 1,000 students were created to seeks answers to the research questions. In the selection of the study groups, a proportionate stratified selection was undertaken by taking into account the students' mathematical proficiency levels. In PISA, students with a proficiency level of 5 or 6 are considered to be in the upper performance group. The students were evaluated according to the three performance groups of upper (levels 5 and above), middle (levels 3 and 4) and lower (levels 2 and below) in proportionate stratified sampling. Before the selection of the sample, the possible missing data and extreme values of the variables of mathematical literacy, classroom climate and mathematics anxiety were examined. As a result of the analysis, no missing data was observed in the mathematical literacy variable, while the rate of missing data was 34% for the classroom climate and mathematics anxiety variables. Van Buuren (2011) stated that if the rate of missing data was less than 30%, data assignment could be made, but if this value is 30% or greater, then the missing data should be removed. As a result of examining the data set, it was determined that the majority of the missing data in both variables were related by the common students who did not respond to the items in the variables. Considering that exclusion of missing data from analysis would still leave a sufficient sample size, the missing data belonging to the variables were removed from the data set. Following the procedures related to missing data and extreme values, the final size of the sample was 3,133 students. Table 1 presents the distribution of these students according to the performance groups.

Table 1. Distribution of Students by Performance Group

Proficiency Level	Performance Group	Study Group									
		n=100		n=200		n=500		n=1000		n=3133	
		n	%	N	%	n	%	n	%	n	%
6 5	Upper	7	7	13	7	32	7	64	7	198	7
4 3	Middle	26	26	53	26	131	26	262	26	824	26
2 1 below 1	Lower	67	67	134	67	337	67	674	67	211 1	67

As shown in Table 1, 67% of the students selected for this research were in the lower performance group in terms of mathematical literacy scores, and 33% were in the middle and upper performance groups.

Data Collection Tools

This research utilized the responses of the selected sample to the items in the mathematical literacy test and student questionnaire in PISA 2012. The entire PISA 2012 data were obtained from the official website of OECD and the data belonging to Turkey were transferred to the SPSS program. The students' mathematics literacy scores were estimated according to the one-parameter logistic model of Matter Response Theory, and five different possible values were determined (OECD, 2013). In this research, the average of these five possible mathematics literacy scores (PV1MATH-

PV5MATH) was taken into consideration. In addition to cognitive tests, PISA includes a student questionnaire that takes approximately 30 minutes to complete. This questionnaire collects data on many dimensions, such as individual characteristics, socio-economic background, educational background, attitudes, learning strategies, learning motives, effectiveness of teaching, and classroom and school climate. In this research, analyses were conducted based on scale indexes of students' mathematics anxiety and classroom climate variables.

Data Analysis

The main purpose of a mediation analysis is to reveal how a relationship between two variables is connected with the presence of another variable. From the mediation analysis methods, the BK method was undertaken with the help of regression analysis while the Sobel test and bootstrapping were performed utilizing the SPSS macros developed by Preacher & Hayes (2004) and accessed from the website of Andrew F. Hayes. Prior to the data analysis, the data set was examined in terms of missing and extreme values.

Before proceeding to the mediation analysis, the assumptions of each analysis method must be tested. Regression equations are used in mediation analyses and each of these equations requires the assumptions of regression analysis to be met (Cohen, Cohen, West, & Aiken, 2003). Thus, in this study, for each data set, it was first determined whether the assumption of univariate normality was satisfied by examining the skewness and kurtosis coefficients of the variables. Table 2 presents these coefficients obtained from the study groups of different sizes.

Table 2. Skewness and Kurtosis Coefficients of the Variables

Variable	N=100		N=200		N=500		N=1,000		N=3,133	
	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis	Skewness	Kurtosis
Classroom Climate	0.163	0.315	-0.101	0.450	-0.001	0.304	0.120	0.230	0.025	0.227
Mathematics Anxiety	-0.278	0.739	0.103	0.217	0.011	0.476	-0.076	0.375	-0.119	0.458
Mathematical Literacy	0.644	-0.028	0.560	-0.257	0.668	0.032	0.559	-0.092	0.538	-0.142

The skewness and kurtosis coefficients of the variables were in the ± 1 range in the study groups (Table 2). This was evaluated as the variable scores not showing an extreme deviation from the normal distribution (Mertler, & Vannatta, 2005). However, the skewness coefficient of mathematical literacy scores being over 0.5 in all groups can be interpreted as showing a slightly skewed distribution. In each group, a scattering matrix was used to determine whether the variables met the assumptions of linearity and multivariate normality. The elliptical distributions in the matrix are evaluated as multivariate normality and linearity. Ensuring multivariate normality also requires satisfying the conditions for univariate normality (Mertler, & Vannata, 2005). A high correlation between the variables ($r > 0.80$) indicates the presence of a multicollinearity problem. Therefore, the correlation coefficients between the variables were calculated to determine whether there was multicollinearity between the variables. Table 3 shows the correlation coefficients between the variables for each study group.

Table 3. Correlation Coefficients Between the Variables

Study Group (N)	Variable	Classroom Climate	Mathematics Anxiety	Mathematical Literacy
3,133	Classroom Climate	1	-0.221**	0.259**
	Mathematics Anxiety		1	-0.316**
	Mathematical Literacy			1
1,000	Classroom Climate	1	-0.216**	0.289**
	Mathematics Anxiety		1	-0.282**
	Mathematical Literacy			1
500	Classroom Climate	1	-0.295**	0.256**
	Mathematics Anxiety		1	-0.329**
	Mathematical Literacy			1
200	Classroom Climate	1	-0.156*	0.237**
	Mathematics Anxiety		1	-0.384**
	Mathematical Literacy			1
100	Classroom Climate	1	-0.270**	0.262**
	Mathematics Anxiety		1	-0.311**
	Mathematical Literacy			1

**significant correlation at 0.01* significant correlation at 0.05

When the correlation coefficients between the variables were examined (Table 3), it was determined that all values were below 0.80, indicating that there was no multicollinearity problem. As a result of the analysis of the assumptions, the data were found to be suitable for analysis. After analysis of missing data and extreme values of the PISA 2012 Turkish sample, the mediation coefficients obtained from the final data set of 3,133 students (the entire group) were used as reference values in the comparisons between different study groups.

The SPSS program was used to analyze the data and examine the assumptions. In order to perform single and multiple mediation analyses, syntax files were created as described by Hayes (2013) and these files were used in SPSS. In addition, MedGraph program and the SPSS output files downloaded from http://pavlov.psyc.vuw.ac.nz/paul-jose/medgraph/Downloads.php were used to conduct single mediation analyses. For these analyses, the level of significance was accepted as .05.

RESULTS

Table 4 presents the results of the BK method concerning the mediating effect of mathematics anxiety in the relationship between classroom climate and mathematical literacy for each study group. These results were obtained from the three regression analyses undertaken for each group.

Table 4. Results of BK Mediation Analysis in Study Groups of Different Sizes

Study Group (N)	Coefficient	B	SHB	β	t	p
3,133	c	25.641	1.708	0.259	15.009	.000
	a	-0.247	0.019	-0.221	-12.677	.000
	b	-24.119	1.507	-0.272	-16.009	.000
	c'	19.683	1.684	0.199	11.685	.000
1,000	c	30.298	3.181	0.289	9.524	.000
	a	-0.248	0.036	-0.216	-6.973	.000
	b	-21.027	2.756	-0.230	-7.631	.000
	c'	25.087	3.168	0.239	7.918	.000
500	c	25.975	4.393	0.256	5.913	.000
	a	-0.319	0.046	-0.295	-6.883	.000
	b	-26.029	4.085	-0.278	-6.371	.000
	c'	17.666	4.425	0.174	3.993	.000
200	c	24.016	6.993	0.237	3.434	.001

	a	-0.177	0.080	-0.156	-2.220	.028
	b	-31.659	5.820	-0.355	-5.440	.000
	c'	18.407	6.617	0.182	2.782	.006
100	c	23.582	8.788	0.262	2.684	.009
	a	-0.323	0.116	-0.270	-2.779	.007
	b	-19.516	7.408	-0.259	-2.634	.010
	c'	17.273	8.863	0.192	1.949	.054

According to the results of the BK method, the classroom climate variable significantly predicted mathematical literacy (coefficient c) in the first step and mathematics anxiety (coefficient a) in the second step. In the third step, the mathematics anxiety variable significantly predicted mathematical literacy (coefficient b). The statistical significance of the coefficient values in the first three steps shows that the conditions of the BK method were met. In the reference group, when coefficient c representing the total effect on the relationship between classroom climate and mathematical literacy ($B = 25.64$, $\beta = 0.26$) was compared to coefficient c' that refers to the direct effect ($B = 19.68$, $\beta = 0.20$), it was found that there was a decrease in the predictive ability of classroom climate for mathematical literacy. According to Baron & Kenny's (1986) most widely used definition of mediation, in order for a variable to be a mediator, coefficient c' obtained from the regression equation when the mediator variable is added should be lower than coefficient c representing the value before the addition of the mediator. When the effect of a mediator variable is controlled, if the independent variable is no longer a significant predictor of the dependent variable, this indicates the presence of a full mediation, and if both the independent and mediator variables significantly predict the dependent variable, then this supports partial mediation. The values in Table 4 show that the mathematics anxiety variable was a partial mediator variable between classroom climate and mathematical literacy according to the BK method.

When the results obtained from different study groups (Table 4) are analyzed, it was observed that in all study groups, the classroom climate variable significantly predicted mathematical literacy (coefficient c) and mathematics anxiety (coefficient a), and the mathematics anxiety variable was a significant predictor of mathematical literacy (coefficient b). These results indicate that the conditions of the BK method for the first three steps were fulfilled; i.e., coefficients c, a and b were statistically significant. However, in the fourth step the method, coefficient c' values differed between the study groups. While the direct effect of classroom climate on mathematical literacy (coefficient c') was significant for all the study groups containing 1,000, 500 and 200 students, this coefficient was not significant in the group of 100 students. This suggests that the mathematics anxiety variable in the 1,000, 500, and 200 student groups was a partial mediator in the relationship between classroom climate and mathematical literacy according to the BK method. This result is interpreted as classroom climate not only directly affected mathematical literacy but also had an indirect effect on this variable through the mathematics anxiety mediator. In the group of 100 students, it was determined that the relationship between classroom climate and mathematical literacy was solely maintained by the mathematics anxiety mediator; i.e., there was a full mediation. In other words, for this sample size, classroom climate did not have a direct effect and only had an indirect effect on mathematical literacy through the mediation of mathematics anxiety.

In the second sub-problem of the research, it was examined whether the mathematics anxiety variable had a mediating effect on the relationship between classroom climate and mathematical literacy in the single mediation model according to the Sobel test. The results were examined first in the reference group, and then in the study groups of different sizes. Table 5 presents the results of the Sobel test on mediation for each study group.

Table 5. Results of the Sobel Test on the Mediator Effect in the Study Groups of Different Sizes

N=3,133								
Coefficient	B	z score	SH	P	Symmetric Confidence Interval		Asymmetric Confidence Interval	
					Lower		Upper	
A	-0.247	9.926	0.590	.000	4.794	7.106	4.994	7.276
B	-24.119							
Sa	0.019							
Sb	1.507							
N=1,000								
Coefficient	B	z score	SH	P	Symmetric Confidence Interval		Asymmetric Confidence Interval	
					Lower	Upper	Lower	Upper
A	-0.248	5.124	1.017	.000	3.218	7.204	3.563	7.499
B	-21.027							
Sa	0.035							
Sb	3.168							
N=500								
Coefficient	B	z score	SH	P	Symmetric Confidence Interval		Asymmetric Confidence Interval	
					Lower	Upper	Lower	Upper
A	-0.319	4.649	1.787	.000	4.806	11.810	5.414	12.329
B	-26.029							
Sa	0.046							
Sb	4.085							
N= 200								
Coefficient	B	z score	SH	P	Symmetric Confidence Interval		Asymmetric Confidence Interval	
					Lower	Upper	Lower	Upper
A	-0.177	2.049	2.768	.042	0.179	11.029	1.120	11.832
B	-31.659							
Sa	0.080							
Sb	5.820							
N= 100								
Coefficient	B	z score	SH	P	Symmetric Confidence Interval		Asymmetric Confidence Interval	
					Lower	Upper	Lower	Upper
A	-0.323	1.850	3.411	.064	-0.375	12.995	0.784	13.984
B	-19.516							
Sa	0.116							
Sb	7.408							

The Sobel test results for the reference group (Table 5) revealed that z score was statistically significant ($p < .05$) and the mathematics anxiety variable mediated the relationship between classroom climate and mathematical literacy. Another effective method for the determination of the significance of the indirect effect is calculation of the confidence interval. The range of confidence interval not including a zero indicates that the indirect effect is significant. MacKinnon (2008) suggested that since the indirect effect (ab) would not be normally distributed, it would be more accurate to evaluate the indirect effect based on an asymmetric confidence interval. In this study, both symmetric and asymmetric confidence intervals did not contain a zero value at the 95% level, which supports the significant mediating effect of mathematics anxiety.

When the Sobel z values obtained from the different study groups were examined, it was found that these values were significant for the study groups of 1,000, 500 and 200 students ($p < .05$). In addition, the symmetric and asymmetric confidence interval values in the same three groups did not contain a zero value. Therefore, according to the Sobel test, in the 1,000-, 500- and 200-student groups, the mathematics anxiety variable was a mediator variable in the relationship between classroom climate and mathematical literacy. In the group of 100 students, the z-score not being significant ($p> .05$) and the symmetric confidence intervals containing a zero value suggested that mathematics anxiety had no mediating effect; however, the asymmetric confidence interval did not include a zero value, which indicates that mathematics anxiety was actually a mediator variable. This finding supports the idea of MacKinnon (2008) that since the multiplication of ab does not have a normal distribution, it is better to evaluate the indirect effect based on the asymmetric confidence interval. In addition, it is noteworthy that as the size of the sample became smaller, the standard error of z-score increased; e.g., 0.590 in the reference group of 3,133 students but 3.411 in the group of 100 students.

In relation to the third research question, the mediation effect of the mathematics anxiety variable on the relationship between classroom climate and mathematical literacy was investigated in a single mediation model according to the bootstrap method first in the reference and then in the different-size study groups. Table 6 shows the results of mediation for each study group according to the bootstrap method.

Table 6. Bootstrapping Results on the Mediator Effect in Study Groups of Different Sizes

Study Group	Bootstrap	M	SH	Bootstrap Confidence Interval	
				Lower	Upper
3,133	ab	5.958	0.620	4.767	7.218
1,000	ab	5.211	0.986	3.379	7.221
500	ab	8.309	1.963	4.828	12.497
200	ab	5.609	2.902	0.026	11.626
100	ab	6.310	3.034	1.241	13.000

Note: Bootstrap resampling = 10,000

The bootstrap confidence intervals (Table 6) obtained at the 95% level from the reference group and the study groups of 1,000, 500, 200 and 100 students did not contain a zero value. Therefore, in all groups, the mathematics anxiety variable mediated the relationship between classroom climate and mathematics literacy according to the bootstrap method.

DISCUSSION, CONCLUSION AND RECOMMENDATIONS

The single mediation model analysis of mediation of mathematics anxiety in the relationship between classroom climate and mathematical literacy revealed the presence of a mediating effect in the reference group according to the BK, Sobel and bootstrap methods. This indicates that part of the students’ positive perception of classroom environment was affected by their reduced mathematics anxiety. The mediating effect of the mathematics anxiety variable was shown by all three analysis methods for the study groups of 1,000, 500 and 200 students, but the results were different for the 100-student group. In this group, although the BK and bootstrap methods found a mediating effect, the Sobel test did not show a significant mediation. In the group of 100 students, a significant mediation effect was only achieved by the finding that asymmetric confidence interval did not include a zero value, which was previously suggested by MacKinnon (2008). It was concluded that for smaller sample sizes, the multiplication of ab in the Sobel test tends to have an asymmetric distribution, which reduces its power to reveal mediating effects. Therefore, in such cases, the indirect effect should be assessed using the asymmetric confidence interval. This confirms the research results of MacKinnon et al. (1995), MacKinnon et al. (2002), and Mallinckrodt et al. (2006), who all reported that the Sobel test had lower statistical power in small sample sizes compared to the methods that involved the correction of this asymmetry.

Cheung, & Lau (2008) suggested that bootstrapping was particularly useful in small samples when there was no information on distribution or when the assumptions of normality were violated, and similarly, Shrout, & Bolger (2002) reported that bootstrapping was strong when the sample distribution of the mediation effects was non-zero or skewed. MacKinnon et al. (2002) stated that the bootstrap method was stronger in revealing indirect effects than the Sobel test in small samples. In the 100-student sample of the current study, the mediator effect was not significant according to the Sobel test, but significant according to the bootstrap method, which supports the findings of all three studies mentioned above. However, Mallinckrodt et al. (2006) suggested that it is not correct to make a generalization based on a small sample of real data and the bootstrap method may not always provide valid results concerning the mediation effect in small samples.

When the standard error values of the coefficients obtained according to different analysis methods were examined, it was found that this value increased as the study group size became smaller. Although the Sobel test and the bootstrap method produced similar standard errors for larger study groups, the Sobel standard error values were lower for the reference group while bootstrapping resulted in lower standard error for the 100-student group. Thus, it was determined that the tests generally produced less erroneous results in large samples, and bootstrapping provided more reliable results in small samples. In other words, when the sample size is increased or when the bootstrap method is used in small samples, estimation of the indirect effect can be performed with less errors.

In the large study groups, the standard errors of the Sobel test and bootstrap method were close to each other, but both were lower than the standard errors of the coefficients obtained by the BK method. In addition, since the BK method does not directly focus on the multiplication of ab, the Sobel test and bootstrapping should be preferred. Since the distribution of indirect effect size (multiplication of ab) tends to be asymmetric, it is recommended to use asymmetric confidence interval instead of symmetric confidence interval in determination of the mediation by the Sobel test in smaller samples. Due to its lower standard error value, the bootstrap method is preferable particularly for small study groups.

This study included PISA mathematical literacy, classroom climate and mathematics anxiety in the single mediation model to examine the effect of the mediator variable. Future research can investigate different mediator variables affecting science literacy and reading skills. In addition, in this study, the BK, Sobel and bootstrap methods were used. Other researchers can explore the strengths and weaknesses of different mediation analysis methods or undertake comparative studies on these methods by defining different simulation conditions.

REFERENCES

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology,* 51, 1173-1182.

Burmaoğlu, S., Polat, M., & Meydan, C. H. (2013). Örgütsel davranış alanında ilişkisel analiz yöntemleri ve Türkçe yazında aracılık modeli kullanımı üzerine bir inceleme. *Anadolu Üniversitesi Sosyal Bilimler Dergisi,* 13(1), 13-26.

Cheung, G. W., & Lau, R. S. (2008). Testing mediation and suppression effects of latent variables: Bootstrapping with structural equation models. *Organizational Research Methods*, 11(2), 296-325.

Cohen, J. P., Cohen, S. G., West, L. S., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Fairchild, A. J., & MacKinnon, D. P. (2009). A general model for testing mediation and moderation effects. *Prevention Science*, 10, 87-99.

Frazier, P. A., Tix, A. P., & Barron, K. E. (2004). Testing moderator and mediator effects in counseling psychology research. *Journal of Counseling Psychology*, 51, 115-134.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millenium. *Communication Monographs*, 76, 408-420.

Hayes, A. F. (2013). *An introduction to mediation, moderation, and conditional process analysis*. New York: The Guilford Press.

Hayes, A. F., & Preacher, K. J. (2014). Statistical mediation analysis with a multicategorical independent variable. *British Journal of Mathematical and Statistical Psychology*, 67, 451-470.

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

Kenny, D. A., Kashy, D., & Bolger, N. (1998). *Data analysis in socialpsychology*. In D. Gilbert, S. T. Fiske, and G. Lindzey (Eds.), Handbook of socialpsychology (4th ed., pp. 233-265). New York: McGraw-Hill.

MacKinnon, D. P. (2008). *Introduction to statistical mediation analysis*. Mahwah, NJ: Erlbaum.

MacKinnon, D. P., & Dwyer, J. H. (1993). Estimating mediated effects in prevention studies. *Evaluation Review*, 17(2), 144-158.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation analysis. *Annual Review of Psychology*, 58, 593-614.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, 7, 83-104.

MacKinnon, D. P., Warsi, G., & Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research*, 30, 41-62.

Mallinckrodt, B., Abraham, W. T., Wei, M., & Russell, D. W. (2006). Advances in testing the statistical significance of mediation effects. *Journal of Counseling Psychology*, 53, 372-378.

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

Mertler, C. A., & Vannatta, R. A. (2005). *Advanced and multivariate statistical methods: Practical application and interpretation* (3th ed.). CA: Pyrczak Publishing.

Muller, D., Judd, C. M., & Yzerbyt, V. Y. (2005). When moderation is mediated and mediation is moderated. *Journal of Personality and Social Psychology*, 89, 852-863.

OECD (2013). *PISA 2012 Results: Ready to Learn: Students' engagement, driveand self-beliefs* (Volume III). OECD Publishing. http://dx.doi.org/10.1787/9789264201170-en

Pardo, A., & Román, M. (2013). Reflections on the Baron and Kenny model of statistical mediation. *Anales de Psicología*, 29(2), 614-623.

Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, and Computers*, 36, 717-731.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40, 879-891.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods*, 7, 422-445.

Sobel, M. E. (1982). *Asymptotic confidence intervals for indirect effects in structural equation models*. In S. Leinhardt (Ed.), Sociological methodology, 1982 (pp. 290-312). Washington, DC: American Sociological Association.

Taylor, A. B., MacKinnon, D. P., & Tein, J. Y. (2008). Tests of three-path mediated effect. *Organizational Research Methods*, 11(2), 241-269.

Van Buuren, S. (2011). *Multiple imputation of multilevel data*. In J. J Hoxand J. K. Roberts (Eds.). The handbook of advanced multilevel analysis (pp. 173-196). New York: Routledge.

Wu, A. D., & Zumbo, B. D. (2007). Understanding and using mediators and moderators. *Social Indicators Research*, 87, 367-392.

Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37, 197-210.