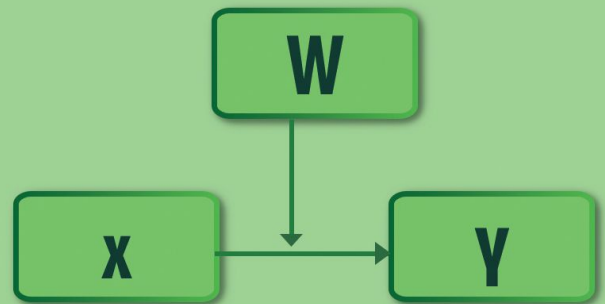
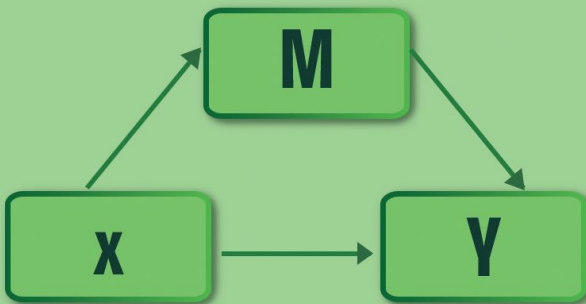


Regression, Mediation / Moderation, and Structural Equation Modeling with SPSS, AMOS, and PROCESS Macro



Lütfi SÜRÜCÜ | Harun ŞEŞEN |
Ahmet MASLAKÇI



LIVRE DE LYON

2023

Regression, Mediation/ Moderation, and Structural Equation Modeling with SPSS, AMOS, and PROCESS Macro

Dr. Lütü SÜRÜCÜ

Prof. Dr. Harun ŞEŞEN

Assoc. Prof.Dr. Ahmet MASLAKÇI



LIVRE DE LYON

Lyon 2023

Regression, Mediation/ Moderation, and Structural Equation Modeling with SPSS, AMOS, and PROCESS Macro

Dr. Lütflü SÜRÜCÜ

Prof. Dr. Harun ŞEŞEN

Assoc. Prof.Dr. Ahmet MASLAKÇI



LIVRE DE LYON

Lyon 2023

**Regression, Mediation/Moderation, and Structural Equation Modeling
with SPSS, AMOS, and PROCESS Macro**

Authors • Dr. Lütfi Sürücü Orcid • 0000-0002-6286-4184

Prof. Dr. Harun Şeşen Orcid • 0000-0002-7068-2487

Assoc. Prof.Dr. Ahmet Maslakçı Orcid • 0000-0001-6820-4673

Co ver Design • Motion Graphics

Book Layout • Motion Graphics

First Published • January 2023, Lyon

ISBN: 978-2-38236-515-1

copyright © 2023 by Livre de Lyon

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior written permission from the Publisher.

Publisher • Livre de Lyon

Address • 37 rue marietton, 69009, Lyon France

website • <http://www.livredelyon.com>

e-mail • livredelyon@gmail.com



LIVRE DE LYON

PREFACE

This book has been written by researchers who share the common trait of admiration for the research process. The aim of the book is to support the academicians and students who do research in social sciences, for gaining competence, skills, and knowledge on statistical analysis, interpretation, and reporting. For this purpose, attention was paid to using plain, simple, and understandable language as much as possible in the writing of the book and to supporting the examples with real research data. The book is short and concise in its design. An effort has been made to focus only on the basic concepts in the book.

The book aims to introduce undergraduate and graduate students to the scientific research process and to be a companion in the analysis process. Readers can do the analysis they need and report the results by simply selecting the topic they are interested in, without examining the entire book. In addition, the examples given about the topics in the book were first explained theoretically and then solved with SPSS and AMOS package programs. Unlike textbooks, this book aims to present different methodological applications about not only “research methods” (experimental data collection and analysis), but the whole “research process” from start to finish. In this respect, it shows that there may be more than one alternative on the way to the goal.

We would like to express that the success of the effort to produce this book has come from the continued support and encouragement we receive from many academics. You can obtain the SPSS data required to follow the analyzes in the book and to do it simultaneously with the book from **www.indataanalysis.com**

Your opinions are valuable to us.

lsurucu@yahoo.com

CONTENTS

Preface.....	I
CHAPTER 1	
CALCULATING SCALE AVERAGE	
Calculating Scale Average	1
Calculating The Scale Average with SPSS	2
Error Screens in SPSS	5
CHAPTER 2	
NORMAL DISTRIBUTION	
Normal Distribution	9
Normal Distribution Analysis With SPSS.....	11
Interpreting Statistical Methods in SPSS Output	13
Interpreting Graphical Methods in SPSS Output	16
Reporting Findings	20
Introduction to AMOS Program.....	21
Implementation of The AMOS Program	25
References	28
CHAPTER 3	
VALIDITY AND RELIABILITY	
Introduction.....	29
VALIDITY	30
Factor Analysis	33
Explanatory Factor Analysis (SPSS).....	34
Factor Analysis with SPSS	36
Reporting Findings	49
Confirmatory Factor Analysis	51
Confirmatory Factor Analysis and AMOS.....	51
Confirmatory Factor Analysis with AMOS	55
Modification of The Model.....	69
RELIABILITY.....	73
Reliability Analysis with SPSS.....	75
Interpretation of The Findings.....	78
Increasing Reliability	80
References	88

CHAPTER 4

CORRELATION AND REGRESSION ANALYSES

Introduction 91

CORRELATION ANALYSIS..... 93

 Correlation Analysis with SPSS 95

 Interpretation of Research Findings 98

 Correlation Analysis with AMOS 99

REGRESSION ANALYSIS..... 101

 Regression Analysis with SPSS 105

 Interpretation of Research Findings 114

 Regression Analysis With AMOS 115

References 120

CHAPTER 5

ANALYSIS OF THE INTERMEDIATE VARIABLE EFFEC

Differences between the Mediator and Moderator Variable 121

Analysis of The Mediator Variable Effect 124

 Baron & Kenny Causal Steps Approach 128

 Mediation Analysis With SPSS 130

 Reporting Research Findings..... 139

 Criticisms of Baron and Kenny’s Causal Steps Approach..... 140

 Structural EquatIon ModelIng (SEM) with AMOS 142

 Structural Equation Modeling with AMOS 145

 Process Macro 155

 Downloading the Process Macro 160

 Mediation Analysis With SPSS 162

 Reporting Research Findings..... 169

References 170

CHAPTER 6

ANALYSIS OF THE MODERATOR VARIABLE EFFECT

Introduction..... 173

Moderation Effect Analysis with Regression..... 177

 Moderation Analysis with SPSS 178

 Reporting Research Findings 187

Structural Equation Modeling (SEM) with AMOS 168

 Moderation Effect Analysis with AMOS 189

Process Macro 193

 Moderation Analysis with SPSS 195

 Reporting Research Findings 206

References 208

CHAPTER 1

CALCULATING SCALE AVERAGE

Almost all research in the social sciences is conducted by detecting latent variables through observed variables. For example, to measure (determine) job satisfaction in employees, certain statements are presented to the participants and they are expected to respond to these statements in varying degrees ranging from “Strongly Disagree” to “Strongly Agree”. For example, let’s assume that the researcher is studying job satisfaction, and there are 5 statements on the scale aimed at determining job satisfaction. The participant’s answers to these 5 statements determine their job satisfaction level. Calculating the scale average means dividing the sum of the scores obtained from the participants’ responses to the 5 statements in the scale by 5. Thus, the scores given by the participants to the expressions in the job satisfaction scale are represented by a single value and this data are used in all analyses. Therefore, before proceeding with the analysis, the scale average of all the variables included in the research should be obtained.

CALCULATING THE SCALE AVERAGE WITH SPSS

To perform the operations to calculate scale averages, the “Data-0.sav” file must be opened (You can access this file from the address www.indataanalysis.com).

The following steps should be followed:

1. Transform ---> Compute Variable....

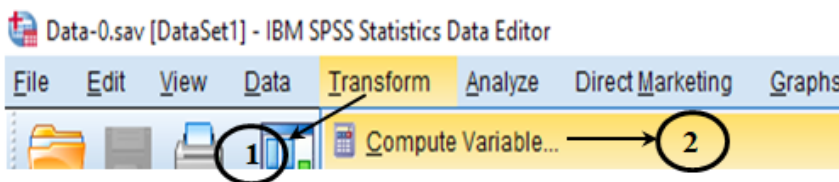


Figure 1.1. SPSS “Tranform” Login Screen

2. On the screen that opens, the relevant tabs are selected.

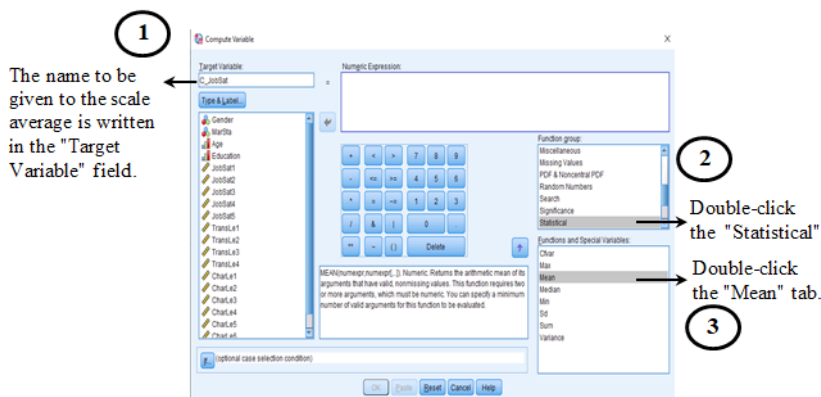


Figure 1.2. Formula Screen-1

After having completed these processes correctly, the following screen is displayed:

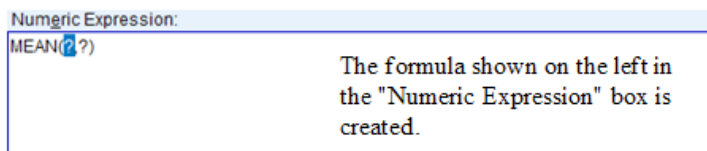


Figure 1.3. Numeric Expressions Screen-1

3. After creating the appropriate formula, each expression of the relevant variable is double-clicked and a comma (,) is added to the formula after each expression. Then, by double-clicking on the other expressions, it is ensured that all expressions in the scale are included in the formula.

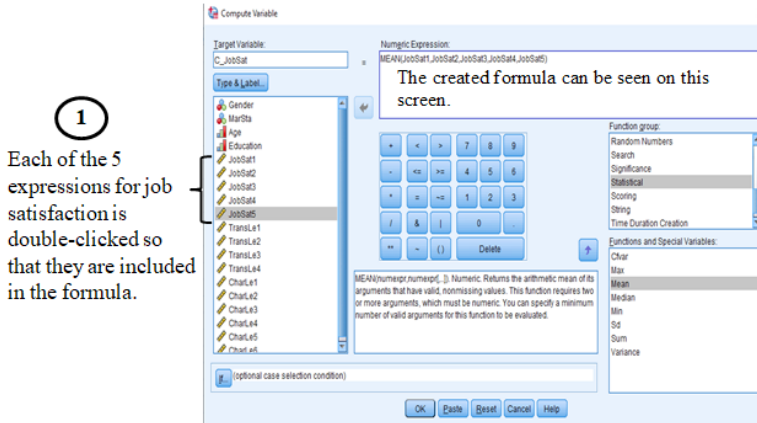


Figure 1.4. Formula Screen-2

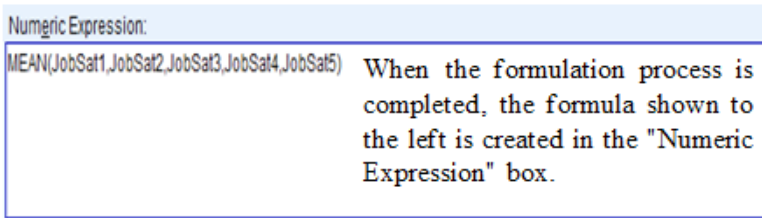


Figure 1.5. Numeric Expressions Screen-2

4. After the scale average has been created, the following output is seen in the SPSS output (Figure 1.6). This output indicates that the operation has been completed successfully. The researcher must verify from this output that the correct statements are included in the scale average. A mistake that is commonly made is that not all of the expressions in the scale are included in the formula. For example, although there are 5 expressions in the scale used for job satisfaction, a common mistake is for the scale average to only have 3 or 4 expressions. Another mistake is where the expression of another scale is included in the formula. For example, in addition to 5 statements about job satisfaction, another scale's statement may be included in the formula and the scale average of the job satisfaction scale is then evaluated with 6 statements.

```
COMPUTE C_JobSat=MEAN(JobSat1,JobSat2,JobSat3,JobSat4,JobSat5).
EXECUTE.
```

Figure 1.6. SPSS Output

5. On the “Variable View” screen of SPSS, the newly created “C_JobSat” scale average can be seen on the 20th line.

16	CharLe3	Numeric	8	0		[1, kesinikli... None	5	Right	Scale	Input	
17	CharLe4	Numeric	8	0		[1, kesinikli... None	5	Right	Scale	Input	
18	CharLe5	Numeric	8	0		[1, kesinikli... None	5	Right	Scale	Input	
19	CharLe6	Numeric	8	0		[1, kesinikli... None	5	Right	Scale	Input	
20	C_JobSat	Numeric	8	2		None	None	10	Right	Scale	Input

Figure 1.7. SPSS Variable View

6. In addition, the newly created “C_JobSat” scale average can also be seen on the “Data View” screen of SPSS. All subsequent analyses will be performed with the newly created “C_JobSat” scale average.

	Gender	MarSta	Age	Education	JobSat1	JobSat2	JobSat3	JobSat4	JobSat5	TransL...	TransL...	TransL...	CharLe1	CharLe2	CharLe3	CharLe4	CharLe5	CharLe6	C_JobSat	
1	2	1	4	5	5	5	5	5	5	3	4	3	3	5	4	4	4	4	3	5,00
2	1	1	1	4	4	4	4	5	4	4	4	3	5	5	5	4	3	4	4	4,40
3	1	1	4	4	3	3	3	3	4	5	3	4	4	3	3	3	3	3	3	3,20
4	1	1	3	4	5	5	5	5	5	3	4	4	4	4	5	4	3	4	3	5,00
5	2	1	2	3	1	3	3	5	2	4	3	5	5	4	5	5	5	5	4	2,80
6	1	1	5	4	4	3	5	3	5	4	4	3	5	4	3	4	5	4	3	4,00

Figure 1.8. SPSS Data View

ERROR SCREENS IN SPSS

In our example, the correct formula is as follows:

MEAN(JobSat1,JobSat2,JobSat3,JobSat4,JobSat5)

However, some mistakes could be made while creating a formula in SPSS. Possible error screens and their causes are summarized below.

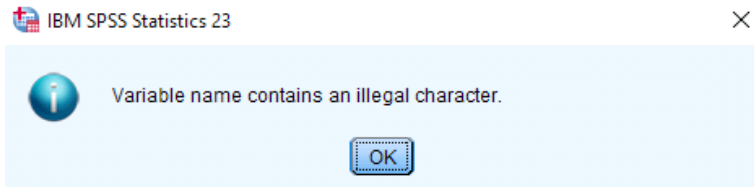


Figure 1.9. SPSS Error Screen-1

If an invalid character is used while specifying the variable name, the above warning will be displayed. In this case, it is likely that the researcher created the scale denomination in the “Target Variable” section with an inappropriate character.

For example;

(a) C JobSat ---> No spaces should be left in the denomination. It is actually C_JobSat (the underscore (_) symbol could also be used).

(b) C_JobŞat ---> Turkish characters should not be used (like the letter ş). It should be C_JobSat. However, this is not a problem in the latest versions of SPSS.

(c) Characters such as -, *, ?, \, /, !, ^, & should not be used. An underscore (_) or full stop (.) could be used.

(d) The creation of a scale denomination containing more than 64 characters in the “Target Variable” is another reason for this error.

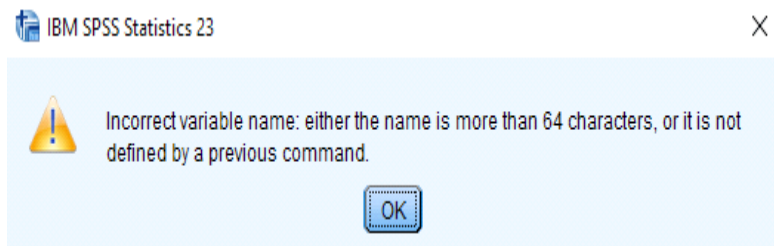


Figure 1.10. SPSS Error Screen-2

This is an error that may occur if there are extra characters in the formula (Figure 1.10).

MEAN(JobSat1,**JobSat2,,JobSat3**,JobSat4,JobSat5)

MEAN(JobSat1,JobSat2,JobSat3,JobSat4,**JobSat5**),

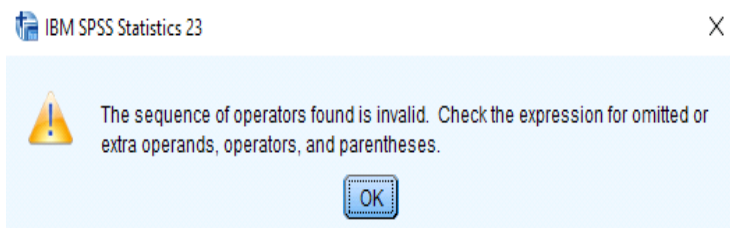


Figure 1.11. SPSS Error Screen-3

If the formula is created incorrectly, the above error is displayed (Figure 1.11).

As seen in this example

(MEAN(**JobSat1 JobSat2**,JobSat3,JobSat4,JobSat5)), a comma has not been added between the two expressions

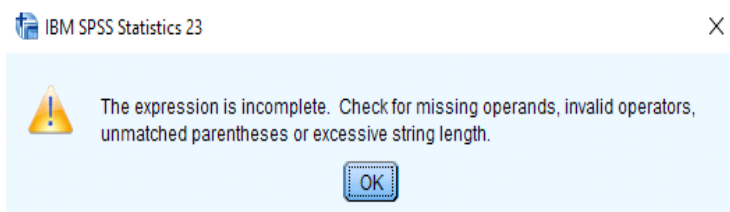


Figure 1.12. SPSS Error Screen-4

If the formula is not completed, this error is displayed (Figure 1.12). In the example below, the parenthesis “)” at the end of the formula has been omitted.

MEAN(JobSat1,JobSat2,JobSat3,JobSat4,JobSat5

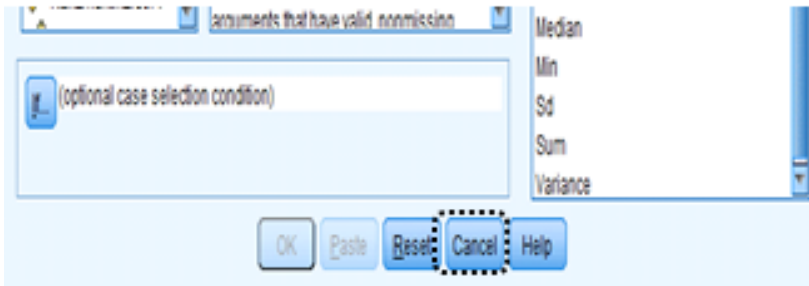


Figure 1.13. SPSS Error Screen-5

Another common situation is where the “OK” button is not active to complete the scale merging process. In this case, it is likely that the researcher used an invalid character in the formula. For example, there is a “Question mark” in the formula shown below.

`MEAN(JobSat1,JobSat2,JobSat3,JobSat4,JobSat5,?)`

CHAPTER 2

NORMAL DISTRIBUTION

Normal distribution tests are some of the most commonly used tests in statistical analysis. These tests are performed to determine how the data are distributed, which is also called “*the Gaussian distribution*” in the literature. After the studies of Carl Friedrich Gauss, the normal distribution formed the basis of parametric statistical analyses, and many statistical analysis methods in SPSS are based on the assumption that the data has a normal distribution. Therefore, normal distribution is a prerequisite for many analyses.

If analyses are conducted without checking the data distribution, this can negatively affect the validity and reliability of the statistical data. In addition, the normality test determines which analyses will be performed on the existing data set. Statistical tests in SPSS are divided into “*parametric tests*” and “*non-parametric tests*”. While “*parametric tests*” are applied for normally distributed data, “*non-parametric tests*” are applied for data that do not have normal distribution.

In *parametric tests*, the normal distribution of the data requires equal variances, while the data should be measured as intervals or ratios. Although parametric tests provide strong statistical values, according to the Central Limit Theorem (CLT), parametric tests should not be preferred for data with a sample size of less than 30. In general, it could be said that parametric tests give more robust results than non-parametric tests.

Non-parametric tests are performed when the data do not have a normal distribution, their variances are not equal, and the data are measured with a nominal or ordinal scale (categorical data). They are also preferred in cases where the number of participants (sample) is relatively small ($n < 30$).

Examples of parametric and non-parametric tests are presented below.

Table 2.1. Tests in accordance with the distribution of data

Parametric Tests	Non-parametric tests
Pearson Correlation Test (<i>r</i>)	Spearman Correlation Test (<i>rho</i>)
Dependent sample t-test	Wilcoxon Signed-Rank Test
Independent sample t-test	Mann-Whitney U Test
One-way ANOVA for independent samples	Kruskal-Wallis H Test
Repeated Measures Anovas	Friedman Test
----	Chi-Square Test

There are two main methods for evaluating data distribution: statistical tests and graphical methods (based on visual inspection).

Statistical tests provide more objective results about the distribution of data. However, they have significant disadvantages as they are less sensitive to low-sampled data and excessively (strictly) sensitive to large-sampled data. Therefore, when performing statistical testing, interpretation according to graphical methods may be preferred for evaluating the data distribution in cases where the statistical analysis may be overly sensitive. However, it should not be forgotten that interpretation according to graphical methods is used to make a subjective judgment and therefore lacks objectivity.

It is a more accurate approach for researchers who do not have sufficient experience in interpreting the data distribution according to graphical methods to rely on statistical tests and act accordingly.

Statistical and graphical tests performed to determine whether the data have a normal distribution are presented below.

Statistical tests:

1. *Test of Normality (Kolmogorov-Smirnov and Shapiro-Wilk tests)*
2. *Skewness and Kurtosis Values*

Graphical tests:

1. *Histogram Diagram*
2. *Q-Q Plot Charts*
3. *Stem and Leaf Plot*
4. *Box Plot Chart*

Determination of the distribution of the data can be performed with the AMOS program (with multiple normal distribution tests) as well as SPSS. At all stages of our book, we will firstly analyze with the SPSS program and then with the AMOS program. The selection of the program to use is at the individual researcher's discretion.

NORMAL DISTRIBUTION ANALYSIS WITH SPSS

Open the “Data-1.sav” file to perform analyses for the determination of data distribution (You can reach this file at www.indataanalysis.com).

The following steps should be followed:

1. Analyze ---> Descriptive Statistics ----> Explore....

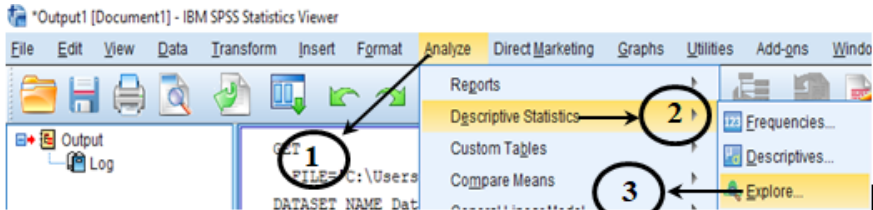


Figure 2.1. SPSS “Analyze” Screen

2. On the screen that opens, the scale averages of the variables included in the research are moved to the “*Dependent List*” section and the “*Statistics*” button is clicked.

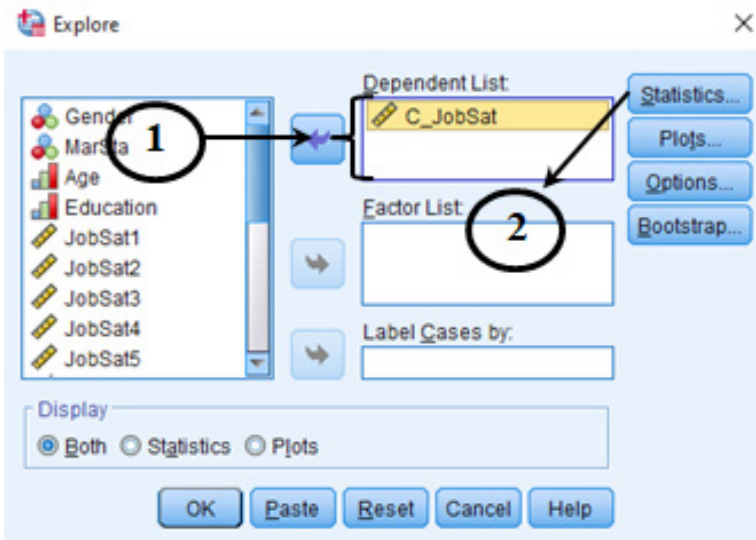


Figure 2.2. SPSS “Explore” Screen

If attention is paid here, the expressions about job satisfaction are transferred to the “*Dependent List*” section of the scale’s average (Compute), not one by one. The scale average topic is explained in detail in the first chapter of the book.

3. “On the *Statistics*” screen, the relevant sections are selected and the “Continue” button is clicked.

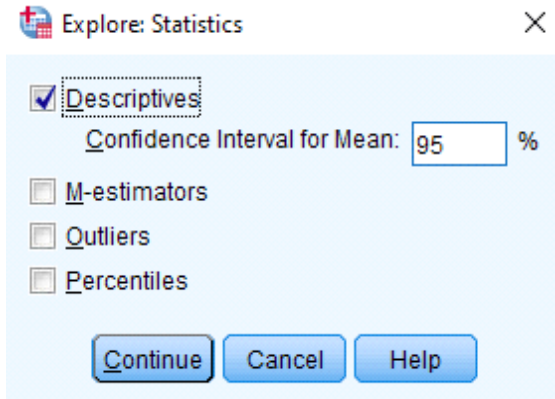


Figure 2.3. SPSS “Explore: Statistic” Screen

4. On the “Plots” screen, the relevant sections are selected and the “Continue” button is clicked.

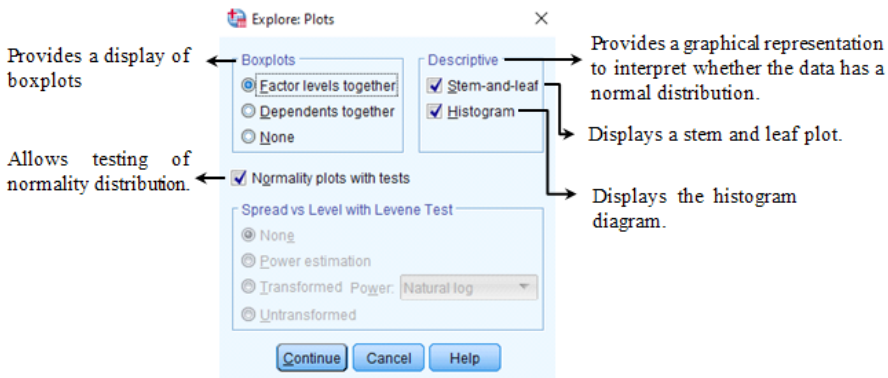


Figure 2.4. SPSS “Explore: Plots” Screen

5. On the “Options” screen, the relevant sections are selected and the “Continue” button is clicked.

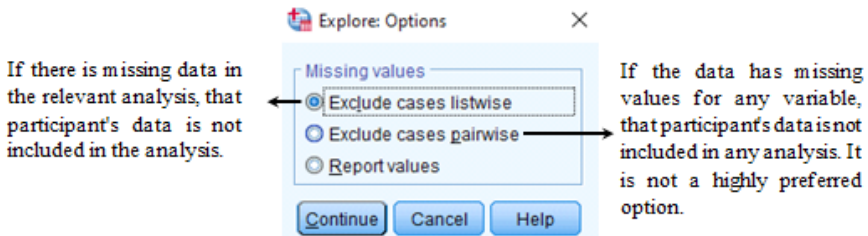


Figure 2.5. SPSS “Explore: Options” Screen

Interpreting Statistical Methods in SPSS Output

SPSS displays the output of several tables and graphs related to statistical data. However, not all of these outputs are actually for normal distribution. One reason for this is that the “Explore” command is not only used for testing normality, but also for describing data in many different ways. In our study, only the tables and graphs given for normality will be interpreted.

It was previously stated that there are two types of methods for assessing normality: statistical tests and graphical methods. In the outputs of SPSS, we will first examine the findings for statistical tests and then evaluate normality according to graphical methods.

Statistical tests:

Table 2.2. Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
C_JobSat	156	100,0%	0	0,0%	156	100,0%

Sample Size ← Percentage Missing Value Numbers Missing Value Percentage

This table contains information on the overall sample size, missing data, and percentages.

Table 2.3. Tests of Normality

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
C_istatmini	,082	156	,013	,969	156	,001

a. Lilliefors Significance Correction

This table contains the results of the normality analysis. The test of normality in SPSS is evaluated according to the results of the “Kolmogorov-Smirnov” and “Shapiro-Wilk” tests. These tests are generally accepted and well-known in the literature. The “Kolmogorov-Smirnov” test is more suitable for large-sampled data (N>50), while the “Shapiro-Wilk” test is more suitable for small-sampled data (N<50); however, the “Shapiro-Wilk” test can also be

preferred by researchers using data with large sample sizes ($N < 2000$). In this sense, the commonly preferred test in the literature is the “Shapiro-Wilk” test. The “Shapiro-Wilk” test is a regression type test that uses correlation to detect data distribution.

The Sig (p) value should be examined to determine whether the data follow a normal distribution. If the P-value is ≤ 0.05 (i.e., $\text{Sig} < 0.05$), the data are not normally distributed. In fact, since the results of our analysis (Kolmogorov-Smirnov = 0.082, $df = 156$, $p < 0.05$) and (Shapiro-Wilk = 0.969, $df = 156$, $p < 0.05$) are statistically significant, it could be stated that the data do not have a normal distribution according to the Kolmogorov-Smirnov and Shapiro-Wilk tests. This conclusion is based on the significant $p < 0.05$ value in the Kolmogorov-Smirnov test ($\text{sig} = 0.013$) and the Shapiro-Wilk test ($\text{sig}.001$). However, these results are obtained in the majority of studies conducted in the social sciences (i.e., $p < 0.05$). These tests do not provide robust results in Likert-type scales and particularly with regard to large-sampled data. For this reason, these tests are generally not taken into account in research conducted in the social sciences and the normal distribution is evaluated according to the “Skewness and Kurtosis” values. Here, as a small tip, it is useful to mention the following. If you plan to use a Likert-type scale in your research (5, 7, 9), using a 9-point Likert-type scale ensures that the data are normally distributed and the reliability coefficient is higher than other Likert-type scales.

The descriptive statistics table presents the “Skewness and Kurtosis” values for the determination of the normal distribution (Table 2.4).

Table 2.4. Descriptive Statistics

		Statistic	Std. Error		
C_JobSat	Mean	3,44	,077	→ Mean	
	95% Confidence Interval for Mean	Lower Bound	3,29		→ Average Confidence Interval
		Upper Bound	3,59		
	5% Trimmed Mean	3,46		→ Trimmed Mean Value (%5)	
	Median	3,40		→ Median	
	Variance	,920		→ Variance	
	Std. Deviation	,959		→ Standard Deviation	
	Minimum	1		→ Minimum Value	
	Maximum	5		→ Maximum Value	
	Range	4		→ Range	
	Interquartile Range	1		→ Interquartile Range	
	Skewness	-,306	,194	→ Skewness	
	Kurtosis	-,686	,386	→ Kurtosis	

The Descriptive statistics table presents a significant amount of information to the researcher, including the *median, variance, and standard deviation value* of the data. However, the most important values in the “Descriptives” table are the **Skewness and Kurtosis values**. *Skewness* in statistics is a measure of the asymmetry of the probability distribution about the mean of a random variable. In other words, the Skewness value informs us about the size and direction of the skewness (away from horizontal symmetry). The skewness value could be positive, negative or even undefined. If the skewness value is 0, the data are truly symmetrical, but this is unlikely to be the case for social science research. The Kurtosis value presents the height and sharpness of the central peak according to a standard bell curve.

If we return to our analysis to determine the data distribution, the Skewness value of job satisfaction is **-0.306** and the Kurtosis value is **-0.686**. These values show that the data have a normal distribution because Skewness and Kurtosis values between **-1.5 and +1.5** indicate that the data are normally distributed.

As a general approach:

For Skewness and Kurtosis values:

- A value between -1.5 and +1.5 indicates that the data have a normal distribution.

- If it is between -1 and +1, the data distribution is moderately skewed,

- A value between -0.5 and +0.5 suggests the data distribution is approximately symmetrical.

There are different interpretations in the literature regarding these values.

Hair et al. [1] stated that if the Skewness and Kurtosis values are in the range of **-1 to +1**, the data have a normal distribution.

George and Mallery [2] stated that a Skewness and Kurtosis value between **-1 and +1** is considered excellent for most psychometric tests, but a value between **-2 and +2** is also acceptable.

Tabachnick and Fidell [3] stated that a Skewness and Kurtosis value between **-1.5 and +1.5** is sufficient for normality.

Kim [4] stated that the value obtained by dividing the Skewness and Kurtosis values by the standard error should be below 1.96. The formula proposed by Kim [4] is presented below:

$$Z_{\text{Kurtois, Skewness}} = \frac{\text{Statistics}}{\text{Std.Error}} < |1.96|$$

Interpreting Graphical Methods in SPSS Output

The graphical methods used are as follows: histogram diagram, Q-Q plot, Stem and Leaf, and box plots. These methods give robust results if the researcher does not decide on the normal distribution based on a single graph but confirms the normality with several graphs.

Histogram Diagram

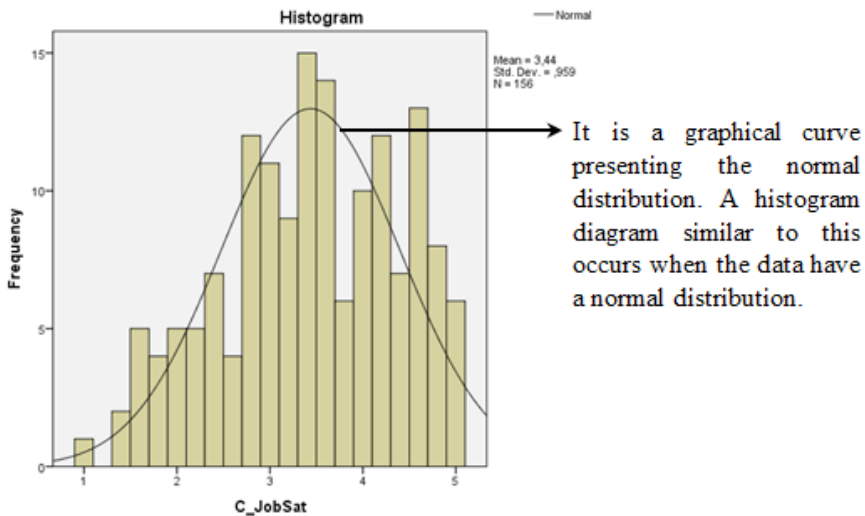
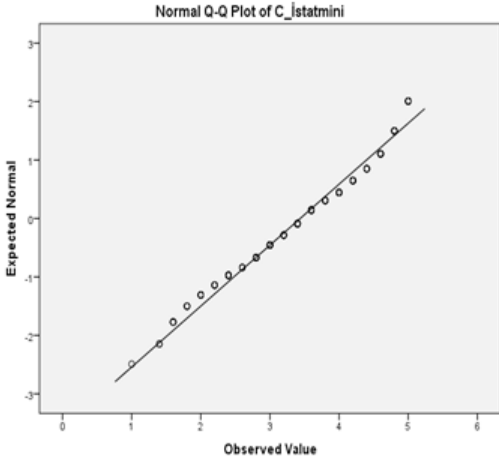


Figure 2.6. Histogram Diagram

Among the different graphical methods, interpreting a Histogram diagram is relatively more complex. In general, the frequency values of the bars in the above graph are expected to resemble a graphical curve with normal distribution. In other words, the bars in the graph should have a symmetrical distribution like a bell curve. Although there is a partially negative skewness in our study, it could be said that it is very close to normal distribution.

Accurately assessing the normality of data in interpretations based on graphs requires extensive experience. Therefore, mistakes are often made this process. If you are unsure about your ability to interpret the graph correctly, it is recommended that you rely on numerical methods instead. However, data that appear to have a normal distribution graphically may not be normally distributed in the theoretical sense. For this reason, the best approach is to decide by checking all methods (statistical and graphical methods) instead of making a decision by examining a single graph.

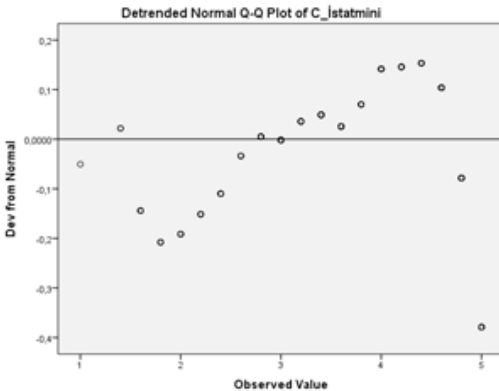
Q-Q Plot Charts



The Q-Q Plot Chart is another graphical method used to graphically determine normality. The small circles in the graph represent the data, and its collection on the 45° lines indicates that the data have a normal distribution.

If the data have a normal distribution, the data points will be close to the diagonal line. If the data points deviate non-linearly from the line, the data are not normally distributed. When the graph on the right is examined, we can say that the data have a normal distribution.

Figure 2.7. Normal Q-Q Plot Chart



In the other Q-Q Plot Chart, small circles represent data and horizontal linear lines represent normality.

If the data have a normal distribution, the data points will be close to the horizontal linear line. This graph also shows the degree of deviation of the data from the normal distribution (Dev from Normal). When the graph on the right is examined, it is seen that the data deviates to an insignificant degree (between -0.4 and 0.2) and are very close to normality.

Figure 2.8. Detrended Q-Q Plot Chart

Stem and Leaf Plots

Frequency	Stem & Leaf
3,00	1 . 044
9,00	1 . 66668888
17,00	2 . 0000222224444444
16,00	2 . 6666888888888888
35,00	3 . 00000000002222222244444444444444
20,00	3 . 66666666666668888888
29,00	4 . 000000000022222222222222444444
21,00	4 . 6666666666666888888888
6,00	5 . 000000

It is a graphical curve presenting normal distribution. When the data have a normal distribution, the stem-leaf plot will be similar to this graphical curve.

The graphical curve showing the normal distribution includes these values. In other words, the stem-leaf plots show that the data are partially skewed to the left (negatively skewed).

Figure 2.9. Stem and Leaf Plots

Another graphical method is the stem-leaf plot. This graph shows that the data for the job satisfaction scale used in our study are partly skewed to the left, but very close to normal distribution.

If we review all of the graphs again, we found that the data in the Histogram diagram were very close to the curve showing the normal distribution, but the data on the right of the diagram (shown as bars) extended beyond the normal distribution curve; in other words, the data are partially skewed to the left (negative skewness). In the Detrended Q-Q Plot, we determined that the deviation values of the data from the normal distribution were very low; however, the deviation in the negative value (approximately -0.4 on the chart) was higher than the positive value (approximately 0.15 on the chart); in other words the data were partially skewed to the left. Next, we checked the stem-leaf plot and we found that the curve showing the normal distribution in this graph did not include the values on the right; in other words, it was partially skewed to the left. Therefore, all the graphical methods examined thus far show that the data are very close to the normal distribution.

One of the main reasons why data do not have a normal distribution is that there are extreme values in the data. These values are caused by researchers making mistakes while transferring the data into the SPSS program (for example, writing 255 instead of 25 when entering the age of a participant) or the careless responses of a participant to the statements in the questionnaire. Another possibility is that one participant has contradictory thoughts compared to the other sample group; however, this situation is rarely encountered. Controlling and removing these extreme values that distort the normal distribution from the analysis ensures that the data, albeit partially, return to a normal distribution. Therefore, it should be verified whether such values exist. The boxplot, which shows the extreme values, can guide researchers in this sense.

To facilitate the understanding of extreme values in the boxplot, it will be useful to give an example. When the chart below is examined, it is seen that the 1st data is the over extreme value, while the 86th and 171st data are the extreme values. It is useful to exclude these data from the analysis. However, this decision is at the discretion of the researcher. If the data show a normal distribution despite the existence of extreme values, these data may not be excluded from the analysis.

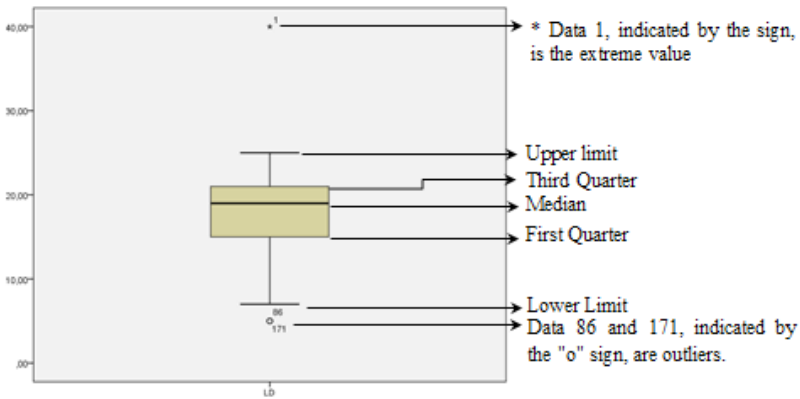


Figure 2.10. Example Box Plot Chart

After the example given to better understand the boxplot showing the extreme values, let's return to our study.

Box Plot Chart

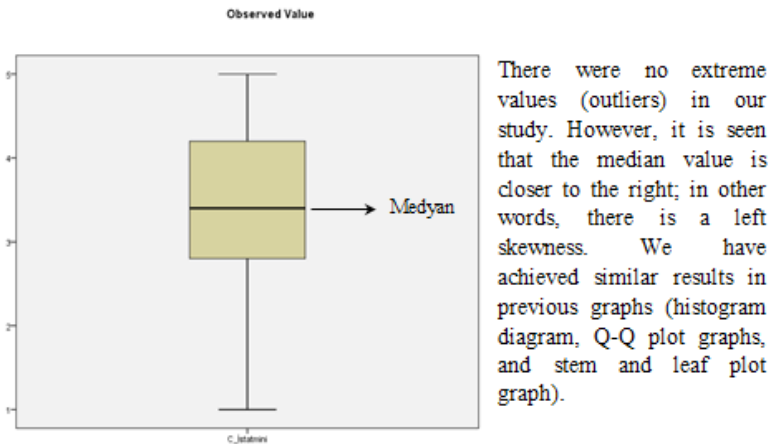


Figure 2.11. Box Plot Chart

All statistical and graphical methods performed show that the data in our study have a normal distribution. If the data does not have a normal distribution, the most important action is to remove outliers or transform the data (log or square root of the data). However, if normality cannot be achieved despite the omission of the extreme values, it is a more proper approach to perform non-parametric tests. Before performing non-parametric tests, the requirements of the data for non-parametric tests need to be checked. Remember that non-parametric tests also have certain assumptions and prerequisites.

REPORTING FINDINGS

Several methods can be used in the interpretation of analyses for normal distribution in studies. A sample report is presented below for researchers who want to provide detailed analysis results.

Table 2.5. Test of Normality

	Mean	Standard Deviation	Skewness	Kurtosis
Job Satisfaction	3.44	.959	-.306	-.686
Transactional Leader	3.83	.773	-.214	-.834
Charismatic Leader	3.33	1.128	-.245	-.858

An examination of Table 2.5 reveals that the skewness and kurtosis values are as follows: job satisfaction (-0.306/-0.686), transactional leader (-0.214/-0.834), and charismatic leader (-0.245/-0.858). The fact that the skewness and kurtosis values of the variables included in the research are between -1.5 and +1.5 reveals that the data have a normal distribution. For this reason, parametric tests were performed in the study.

INTRODUCTION TO AMOS PROGRAM

We consider that it would be beneficial to introduce the AMOS program before performing the analysis. Therefore, in this part of the book, the basic structure and usage features of the AMOS program will be briefly explained. More detailed information on the subject can be found in the book titled “Yapısal Eşitlik Modellemesi AMOS Uygulamaları (Applications of Structural Equation Modeling AMOS)” by Meydan and Şeşen [27].

When the AMOS Graphics program is opened, the main window shown in Figure 2.12 will be displayed.

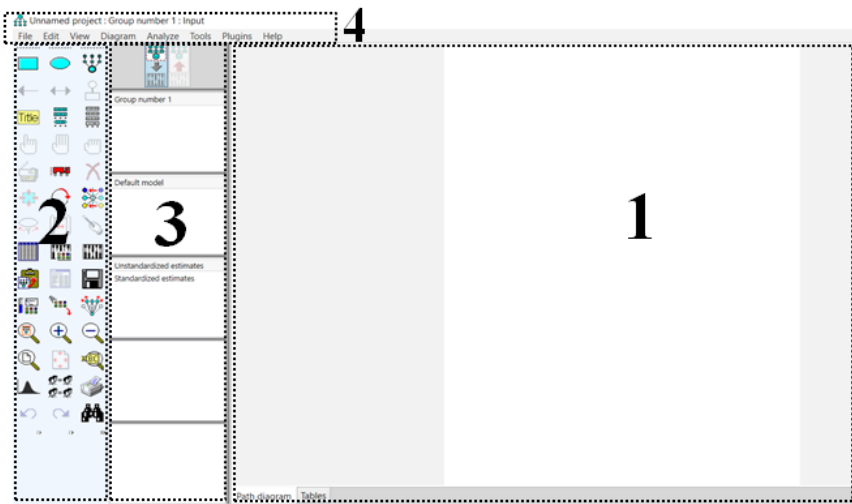
































Figure 2.12. Main Screen of AMOS Graphics Program

The main screen of the program consists of four parts. The first of these is a blank page that forms the majority of the screen on which the drawings will be made. The second part is the part on the left of the window that contains various quick function (shortcut) buttons. The third part is the darker part between these two. In this section, two structural model shapes will be drawn at the top, one of which is colored and the other is colorless. The colored figure is used to show the drawing model, and the colorless figure is used to show the resulting (completed research) model. The fourth part of the program screen contains the menus arranged on the top of the screen, which has the structure of classic Windows programs.

Table 2.6. Buttons in the Main Menu (adapted from Meydan and Şeşen [27])

Nu.	Button	Name of Button	Function
1		Rectangle	Draws an observed variable
2		Ellipse	Draws an unobserved (latent) variable
3		Indicator	Draws a latent variable or adds an indicator to a latent variable
4		Path (single-headed arrow)	Draws a regression path
5		Double-headed arrow	Draws a covariance (mutual relation)
6		Error	Adds an error term to an observed variable
7		Title	Figure captions
8		Variable list-1	Lists the variables in the model
9		Variable list-2	Lists the variables in the data set
10		Single selection	Selects an object at one time
11		Multiple selection	Selects many objects at one time
12		Deselect multiple selection	Deselects multiple selection
13		Copy	Creates copies of the selected object(s)
14		Move	Moves the selected object(s) to the desired location
15		Delete	Deletes the selected object(s)
16		Change the shape	Changes the shapes of the object(s)

17		Rotate	Rotates the indicators
18		Reflect	Reflects the indicators
19		Move parameter	Moves the parameter to the desired location
20		Move screen	Moves the main screen
21		Touch	Arranges the touched variable's paths
22		Data file	Selects the data file and reads it
23		Analysis properties	Used to design the analysis properties
24		Calculate the estimates	Runs the model
25		Copy to clipboard	Copies the model to the clipboard
26		View text	Shows the results in the text
27		Save	Saves the diagram
28		Object properties	Defines the object properties
29		Drag	Drags the properties between objects
30		Preserve symmetry	Equals the distance between objects
31		Zoom	Zooms a selected area
32		Zoom in	Views a smaller area of the path diagram

33		Zoom out	Views a larger area of the path diagram
34		Zoom page	Shows the entire page on the screen
35		Resize	Resizes the page to fit to the page
36		Magnify	Examines a specific part of the diagram
37		Bayesian	Calculates Bayesian statistics
38		Multigroup	Conducts a multi-group analysis
39		Print	Prints the diagram
40		Undo-1	Undoes the last action
41		Undo-2	Undoes the last undo
42		Specification search	Opens a screen for a specific search

The AMOS program offers many tools with which users can draw models. These tools are represented by a button that performs its specific functions. These tools are listed vertically on the left side of the main page of the program. There are 42 buttons in total. The buttons in the program, their names, and their functions are explained in Table 2.6. When the mouse is moved over the buttons when drawing a model, the operation will be automatically seen. However, those that are most frequently used in model drawing are the indicator and error buttons. The AMOS program also includes many user-friendly shortcut applications. For example, using the “Name Unobserved Variables” command under the “Plugins” tab, all variables in an extremely complex model can be named in a single operation. As users use the program in different models and analyses, their capability will also improve.

IMPLEMENTATION OF THE AMOS PROGRAM

A multivariate normal distribution test can be performed with the AMOS program. The data in the “Data-1.sav” file, for which a single normal distribution test was performed with SPSS above, will be tested with AMOS. Before performing this analysis, the model is first drawn as seen in Figure 2.13. For details on drawing the model, see the section on model drawing with AMOS. The model drawing process will not be explained here. However, according to the model above, charismatic and transactional leadership are the independent variables, and job satisfaction is the dependent variable. Here, drawing the model as a correlational model will not produce a different result in the context of normality testing. There are 6, 4, and 5 observed variables for each variable in the data set, respectively. After drawing the model, the “Data-1.sav” file is opened as the data set and the related observed variables are assigned. The first level latent variables are named “F” factors.

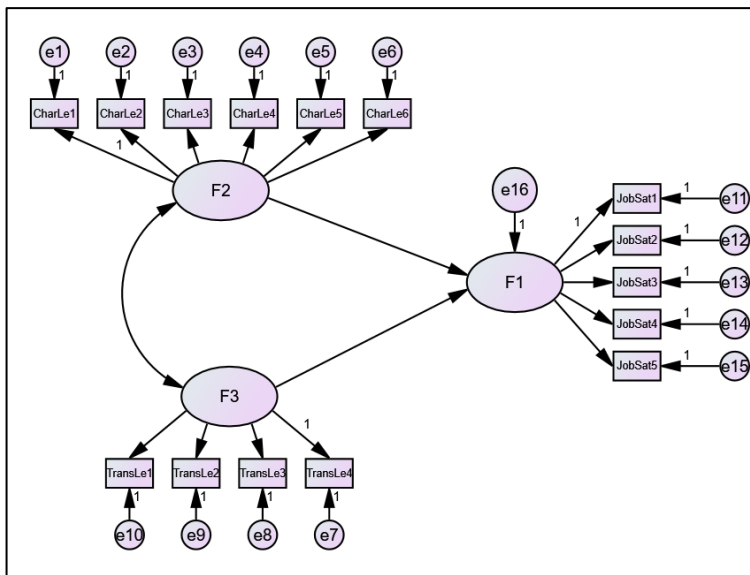


Figure 2.13. Multiple normal distribution test model

After the data set is attained, the “Analysis Properties” tab is selected under the “View” menu before proceeding to the analysis. In the Analysis Properties tab, the “Tests for normality and outliers” tab is selected, as shown in Figure 2.14. Thus, it will be ensured that the normality test is performed after the analysis.

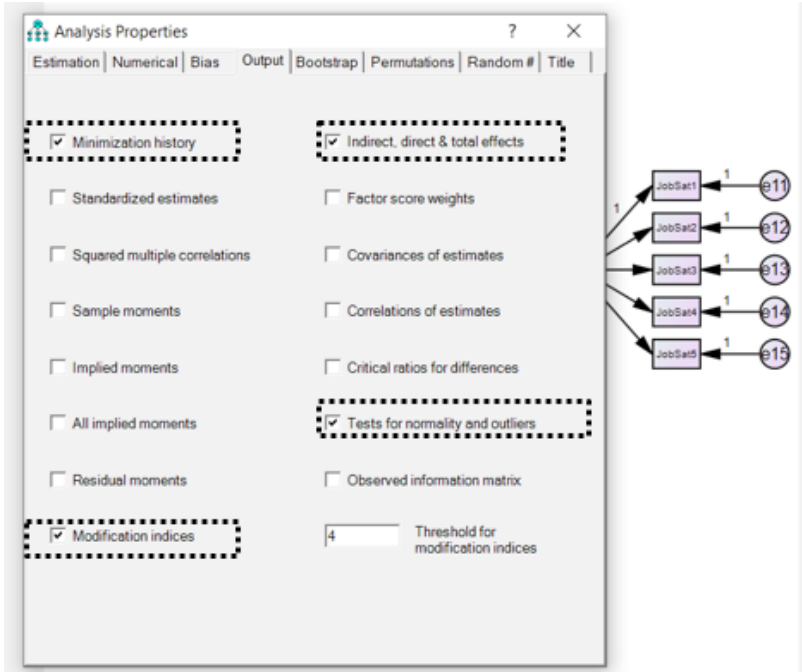


Figure 2.14. “Tests for normality and outliers” tab under “Analysis Properties”

Then, the analysis is performed by clicking on “Calculate Estimate”. The results of the analysis are displayed by clicking the “View Text” button. As seen in Figure 2.15, when the “Assessment of Normality” option is selected, the normality test results will be displayed.

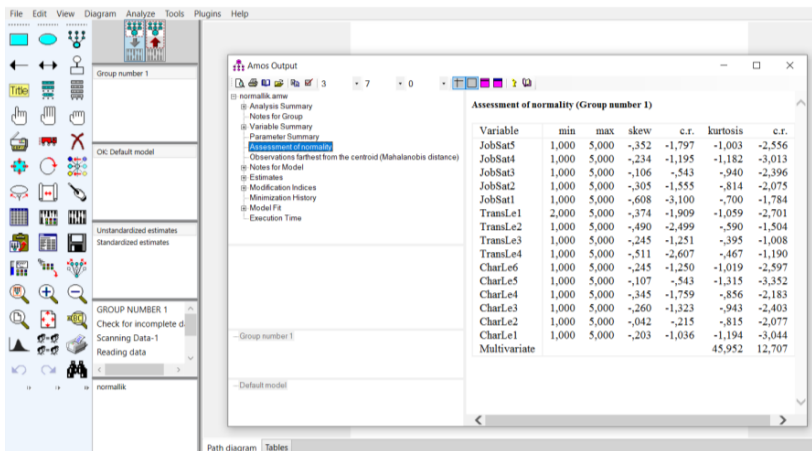


Figure 2.15. Results for the “Assessment of Normality” option

The results are shown in Table 2.7 for better visibility. Copy/paste functions can be carried out by using the right mouse button, which allows the results to be taken as a table and copied to the Microsoft Office Word program.

Table 2.7. The Results of the Tests of Normality

Variable	min	max	skew	c.r.	kurtosis	c.r.
JobSat5	1,000	5,000	-,352	-1,797	-1,003	-2,556
JobSat4	1,000	5,000	-,234	-1,195	-1,182	-3,013
JobSat3	1,000	5,000	-,106	-,543	-,940	-2,396
JobSat2	1,000	5,000	-,305	-1,555	-,814	-2,075
JobSat1	1,000	5,000	-,608	-3,100	-,700	-1,784
TransLe1	2,000	5,000	-,374	-1,909	-1,059	-2,701
TransLe2	1,000	5,000	-,490	-2,499	-,590	-1,504
TransLe3	1,000	5,000	-,245	-1,251	-,395	-1,008
TransLe4	1,000	5,000	-,511	-2,607	-,467	-1,190
CharLe6	1,000	5,000	-,245	-1,250	-1,019	-2,597
CharLe5	1,000	5,000	-,107	-,543	-1,315	-3,352
CharLe4	1,000	5,000	-,345	-1,759	-,856	-2,183
CharLe3	1,000	5,000	-,260	-1,323	-,943	-2,403
CharLe2	1,000	5,000	-,042	-,215	-,815	-2,077
CharLe1	1,000	5,000	-,203	-1,036	-1,194	-3,044
Multivariate					45,952	12,707

As can be seen in the table, the Skewness and Kurtosis values of each observed variable as a result of the normality test are displayed along with their critical ratios. Additionally, the Multivariate normal distribution value in the context of all variables can be seen at the bottom. The interpretation of these Skewness, Kurtosis, and Critical Ratio values is the same as in the SPSS program. According to these values, it can be accepted that the data are normally distributed. However, when the Kurtosis and Critical Ratio values of Multivariate normality are examined, it is seen that they are significantly above the acceptable limits. Therefore, it can be said that there is no multivariate normal distribution in this data set.

In this situation, the data set could be formed normally distributed by examining which observation disrupts the normality. To do this, the “Observations farthest from the centroid (Mahalanobis distance)” tab should be selected, as shown in Figure 2.16, and thus, the distance of each observation from the center is seen. At this point, it will be seen that multivariate normality improves as the observations farthest from the center are omitted from the data set.

Observation number	Mahalanobis d-squared	p1	p2
55	53.181	.000	.001
18	43.664	.000	.000
43	37.021	.001	.001
9	35.827	.002	.000
68	33.459	.004	.000
61	33.447	.004	.000
47	33.377	.004	.000
104	32.228	.006	.000
52	32.116	.006	.000
56	31.933	.007	.000
5	30.760	.009	.000
92	30.051	.012	.000
53	27.927	.022	.000
148	27.776	.023	.000
59	26.717	.031	.000
49	26.230	.036	.000
130	26.218	.036	.000
44	26.087	.037	.000
96	25.390	.045	.000
87	25.080	.049	.000
124	24.547	.056	.000
40	24.477	.057	.000
137	24.292	.060	.000
90	24.130	.063	.000
154	23.998	.065	.000
64	23.973	.066	.000
57	23.956	.066	.000
63	23.284	.078	.000
79	22.207	.102	.001
35	21.044	.135	.029
16	20.956	.138	.023
147	20.818	.143	.021
69	20.665	.148	.020
36	20.452	.155	.024

Figure 2.16. “Observations farthest from the centroid (Mahalanobis distance)” tab

References

- [1] Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2013). *Multivariate Data Analysis*: Pearson Education Limited.
- [2] George, D., & Mallery, M. (2010). *SPSS for Windows Step by Step: A Simple Guide and Reference, 17.0 update (10a ed.)* Boston: Pearson
- [3] Tabachnick and Fidell, 2013 B.G. Tabachnick, L.S. Fidell *Using Multivariate Statistics (sixth ed.)* Pearson, Boston (2013)
- [4] Kim, H. Y. (2013). Statistical Notes for Clinical Researchers: Assessing Normal Distribution (2) Using Skewness and Kurtosis. *Restorative Dentistry & Endodontics*, 38(1), 52-54.

CHAPTER 3

VALIDITY AND RELIABILITY

Although the concepts of validity and reliability are closely related to each other, these terms express different features of the measurement tool. In general, a measurement tool could be reliable without being valid, but if the measurement tool is valid, it is likely that it is also reliable. However, reliability alone is not sufficient to ensure validity. Even if a scale is reliable, it may not accurately reflect the behavior or quality that is intended to be measured (Figure 3.1). For this reason, researchers must test both the validity and reliability of the measurement tool they are using. The measurement tool must meet these two conditions. Otherwise, the interpretation of the research findings will not be correctly.

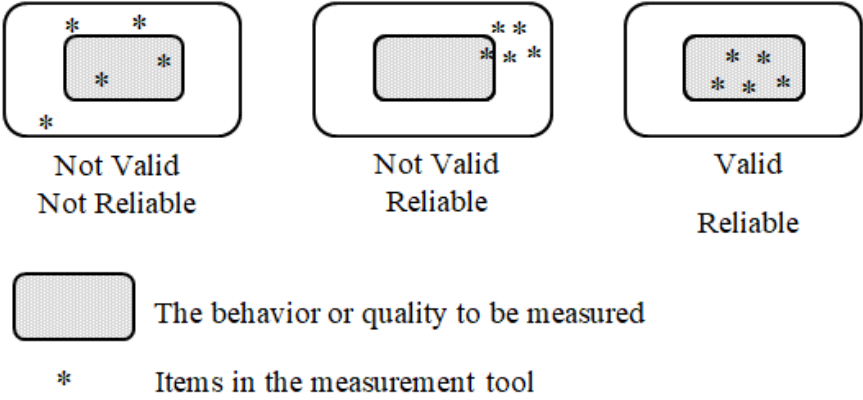


Figure 3.1. Validity and Reliability

VALIDITY

Validity is concerned with whether the measurement tool measures the intended behavior or quality, and it is a measure of how well the measurement tool performs its function [1]. Validity is determined by the meaningful and appropriate interpretation of the data from the measurement tool as a result of the analysis. Whiston [2] defines validity as the acquisition of appropriate data for the intended use of the measuring instrument. In this case, it is important what the measurement tool intends to measure, and whether the items in the measurement tool measure correctly according to the purposes of the research.

Although evaluating the validity of the measurement tool is more difficult than reliability, it is more important. For the research to yield healthy results, the measurement tool must measure what it claims. Using a validity-tested measurement tool will ensure that the findings produced by the analysis are valid.

To determine the validity of the measurement tool, different types of validity are described in the literature. These can be listed as “content validity, criterion-related validity, concurrent validity, internal validity, external validity, constructive validity, face validity, systemic validity, theoretical validity, jury validity, consequential validity, cultural validity, predictive validity, interpretive validity, descriptive validity, evaluative validity, statistical conclusion validity, and translation validity”. Four types of validity are important in the social sciences, although the list that has been created can be extended further. These are *internal validity*, *content validity*, *external validity*, and *construct validity*.

If the researcher does not develop a new scale and uses a previously developed scale in his/her research that has been tested for validity and reliability in the local language of the country where the research was conducted. It is sufficient to test the validity of the content, or in other words, the content validity.

Content validity: Bollen [3] defined *content validity* as a type of qualitative validity in which the domain of the measurement tool is clarified and that evaluates whether the items in the measurement tool fully represent the domain. In line with this definition, it can be said that content validity is a validity study that reveals the extent to which the measurement tool serves the purpose as a whole as well as each item in the measurement tool.

Content validity, which is used in scale development or adaptation studies for the culture and language of the developed scale, ensures the determination of the most appropriate items to increase the quality of the items in the

measurement tool and to serve the purpose of the scale. Thus, it is ensured that the measurement tool prepared to measure any behavior or quality is a useful scale with scope adequacy serving its purpose. The content area of many concepts that are included in the field of social sciences is unclear. For this reason, there is no consensus on the definitions and content of most concepts in the literature. Researchers who carry out content validity studies must make a theoretical definition of the related concept and determine the content (dimensions) of the concept.

Several methods are used for determining content validity in the literature. Among these, obtaining expert opinions and statistical methods are the two most frequently applied methods.

The first method, which involves the evaluation of more than one referee, is also known as “obtain expert opinions”. This method is a process that transforms qualitative studies based on expert opinions into statistical quantitative studies. In this method, the researcher asks experts to rate each item in the developed measurement tool in terms of fitting with its content, and estimations are made for each item in line with expert opinions. The experience and number of experts are of great importance in order to obtain objective results in the estimations to determine the content validity [4]. The experience level of the experts is very important for the results to be consistent and unbiased. For this reason, care should be taken in the selection of experts and academicians or practitioners with extensive knowledge about the measurement tool to be developed should be preferred.

Estimation of content validity with expert opinion is a statistical analysis based on content validity regarding whether the items in the measurement tool are on the scale or not, and it is evaluated according to the following formula:

$$KGD = \frac{N_U}{N} - 1 \quad \text{or} \quad KGD = \frac{N_U - \frac{N}{2}}{\frac{N}{2}}$$

In the formula:

N = Total number of experts who evaluated the items in the measurement tool

NU = Number of experts who evaluated the relevant statement as appropriate

According to Lawshe [5], each statement in the item pool is submitted to expert opinion. Experts evaluate these statements as “Appropriate”,

“Appropriate But Should Be Revised” and “Should be Removed”. If half of the experts express their opinion as “Suitable” about the item in the measurement tool, it will be $KGD=0$, if more than half of the experts have given the opinion “Suitable”, $KGD >0$, and if less than half of the experts have given their opinion as “Suitable”, it will be $KGD <0$. If the KGD rate is 0 (zero) or a negative value, that item should be removed from the measurement tool.

The second method involves testing the content validity with statistical methods. The most widely-used method among statistical methods is factor analysis. Factor analysis is performed to determine the structure under which the items in the measurement tool are gathered and whether they are suitable for the concept. Measurement tools used in social science research generally have a construct with more than one-subdimension. For example, the measurement tool developed by Allen and Meyer [6], which determines the level of organizational commitment in employees, consists of three sub-dimensions: *affective, continuance and normative commitment*. Likewise, the leader-member exchange scale developed by Liden and Maslyn [7] to measure the interaction between leaders and members is examined in four dimensions, namely contribution, commitment, influence and professional respect.

FACTOR ANALYSIS

Factor analysis emerged with the study of Charles Spearman in the early 1900s and is used in many fields such as social sciences, medicine, economics and geography. Factor analysis uses mathematical procedures to simplify interrelated measures to discover patterns in a set of variables. Basically, it summarizes the data so that the relationships and patterns of the observed variables in the measurement tool can be easily interpreted and understood.

Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) are the two most frequently-used main factor analysis techniques in the social sciences. *Exploratory Factor Analysis* is conducted according to the experience of the researcher, so EFA is intuitive in general. The EFA is an exploratory type of analysis that the researcher performs to obtain information about the number or nature of the variables. This analysis allows the researcher to discover the main dimensions of a relatively large hidden structure represented by a set of elements [8,9]. The main purpose is to examine a large number of variables related to the structure as well as to identify fewer variables that explain the structure and to increase the explanatory power of the structure. It is generally performed in order to reduce the number of variables observed in scale development studies and to determine which factors constitute it. In *Confirmatory Factor Analysis*, the main purpose is to test the accuracy of a previously validated scale or model.

In studies with previously tested scales, confirmatory factor analysis is first performed to test the validity of the scale or model. Exploratory factor analysis is performed if sufficient threshold values (fit indices) are not provided in the confirmatory factor analysis or if the structure of the measurement tool could not be verified. Thus, the relationship pattern between the items and the factors in the measurement tool should be discovered, the necessary revisions should be made and the confirmatory factor analysis should be repeated.

EXPLANATORY FACTOR ANALYSIS

Exploratory factor analysis functions based on the concept that measurable and observable variables can be reduced to less unmeasurable latent variables that share a common variance. [10]. Although these unobservable factors are not measured directly, they are hypothetical constructs used to represent the variables.

Researchers used exploratory factor analysis when they want to discover the number of factors affecting variables and analyze which statements are grouped together. The main hypothesis of EFA is to find common “hidden” factors in the dataset and to identify the least number of common factors that explain the structure. Thus, researchers use EFA to focus on fewer statements that explain the structure and place these statements into meaningful categories (sub-dimensions) instead of considering too many statements that may be unimportant. As a result, it is possible to reduce many expressions in the scale to a smaller data set, making the measurement tool easier to interpret.

In summary, exploratory factor analysis is used to reduce the number of statements in the measurement tool, to examine the structure or relationship between statements, to identify and evaluate a theoretical structure, to develop theoretical structures, to test the construct validity of the developed measurement tool, and to prove or reject a proposed theory.

Prerequisites for Exploratory Factor Analysis:

- There must be at least 3 variables for exploratory factor analysis to be performed, but this depends on the design of the study [11]. Factors with two or fewer variables should be interpreted with caution. Bivariate factor analysis is only considered reliable when the variables are highly correlated with each other ($r > 0.70$), but not highly correlated with other variables.
- The data should have univariate and multivariate normality. The absence of univariate and multivariate outliers is also important [12].
- The calculation of the correlation in exploratory factor analysis is made based on the assumption that there is a linear relationship between the factors and variables.
- Measurements of variables should be made with equal intervals (e.g., 5 Likert or 7 Likert).
- The sample size should be sufficient. As the sample size increases, the error rate in the data decreases. For this reason, EFA generally performs better with larger sample sizes.

Sample Size: Although the sample size is important for exploratory factor analysis, there are different opinions in the literature. While Tabachnick and Fidell [11] claimed that a sample size of 300 is sufficient for factor analysis, Hair et al. [13] stated that 100 is sufficient for factor analysis. On the other hand, Guadagnoli et al. [14] stated that if the factor load of the expressions in the measurement tool is $> .80$, smaller sample groups are sufficient for factor analysis. Accordingly, Sapnas and Zeller [15] stated that even a sample size of 50 may be sufficient for factor analysis (where the factor load is >0.80). As can be seen, the answers given by the participants to the statements greatly influence the sample size recommended for factor analysis. When the studies on the determination of the sample size in factor analysis are evaluated, it can be said that a sample size of 200 is sufficient for social science research.

There are also different opinions in the literature as to whether a certain number of times the number of expressions in the measurement tool is sufficient to determine adequate sample sizes. These rates are 3, 6, 10, 15, 20 times the number of expressions in the measurement tool [16, 17, 18]. However, researchers such as Hogarty et al. [19] and MacCallum et al. [20] conducted a series of studies to test the validity of these ratios in the literature, and the results showed that it is not appropriate to specify a minimum ratio for factor analysis.

Researchers who want to determine whether the sample size is sufficient for factor analysis should also check the Kaiser-Meyer-Olkin (KMO) value. If the KMO value, which can range between 1 and 0, is 0.5 and above, this indicates that the sample size is sufficient for factor analysis [11].

FACTOR ANALYSIS WITH SPSS

To perform factor analysis, we open the “Data-1.sav” file (You can access this file from the address www.indataanalysis.com).

The steps below should be followed:

1. Analyze ----> Dimension Reduction ----> Factor....

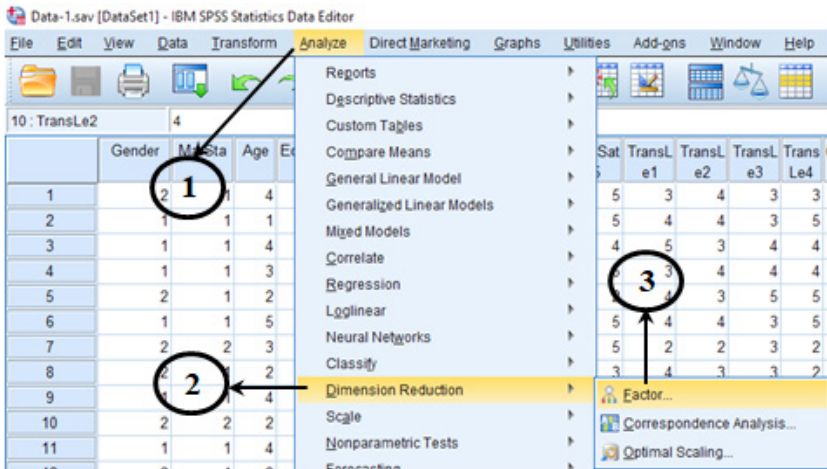


Figure 3.2. SPSS “Analyze” Screen

2. On the screen that opens, all the items of the variables included in the research are moved to the “Variables” section and the “Descriptives” button is clicked.

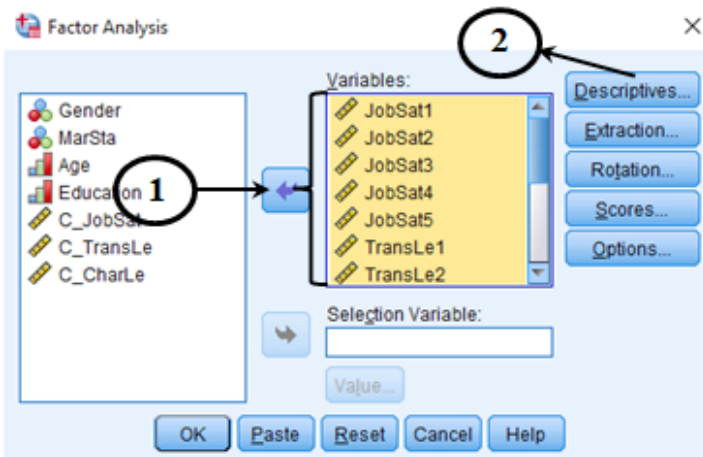


Figure 3.3. SPSS “Analyze” Screen

As can be observed, a total of 15 statements used to measure job satisfaction (5 statements), transactional leadership (4 statements), and charismatic leadership (6 statements) were transferred to the “Variables” section.

3. Relevant sections are selected on the “*Descriptives*” screen.

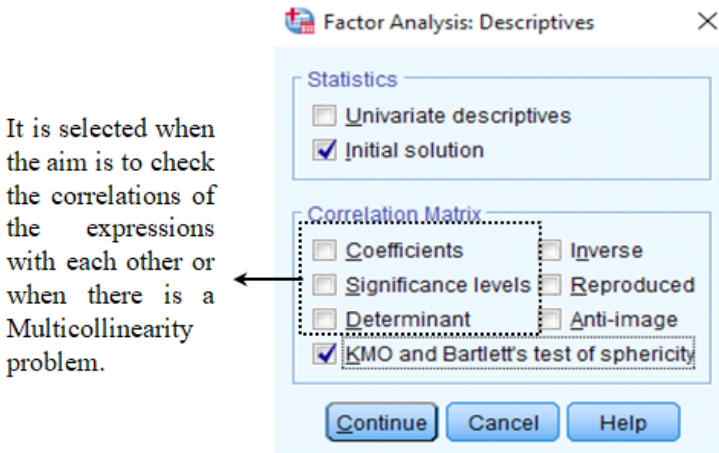


Figure 3.4. SPSS “Descriptives” Screen

4. Relevant sections are marked on the “*Extraction*” screen.

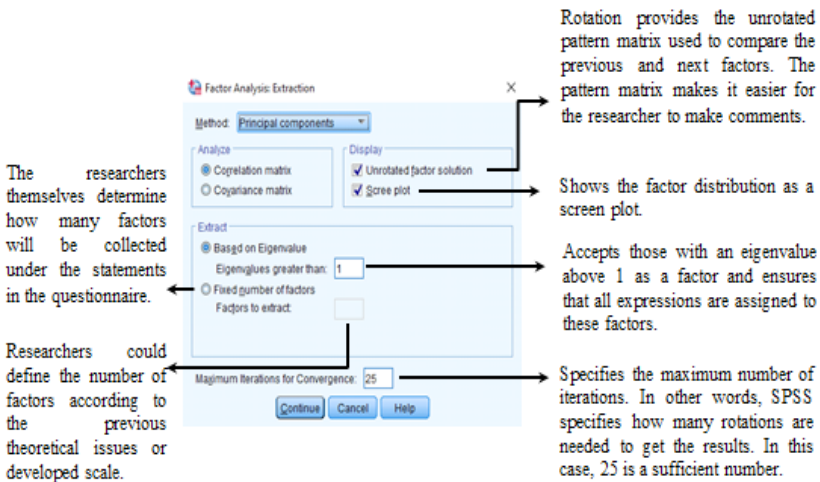


Figure 3.5. “*Extraction*” Screen

Factor analysis is based on a theoretical model called the “*common factor model*”. This model assumes that the observed variables are affected by common factors and unique factors, and their correlations need to be determined [21]. There are a number of “*Extraction*” methods available in SPSS. For this reason,

it is useful to briefly mention several commonly used techniques (Method tab in Figure 3.5).

Principal Component: Principal component analysis is used to extract the maximum variance from the dataset with each component. Thus, it reduces a large number of variables to a smaller number of components [11]. Principal component analysis is an expression reduction technique, and researchers can use principal component analysis as a first step to reduce expressions at scale. Although it varies depending on the research hypothesis and ease of interpretation, researchers mostly prefer this method.

Maximum Likelihood; The maximum likelihood method tries to analyze the maximum sampling probability of the observed correlation matrix [11]. When performing confirmatory factor analysis in research, it is a more appropriate approach to choose the Maximum Likelihood method from the “Method” section.

The Principal Axis Factoring; The Principal Axis Factoring method is based on the idea that all expressions belong to a group and a matrix is estimated when the factor is subtracted. The factors are then subtracted repeatedly until a sufficiently large variance is calculated in the correlation matrix. Principal Axis Factoring is recommended when the data violate the assumption of multivariate normality [22].

5. “Relevant sections are marked on the “Rotation” screen.

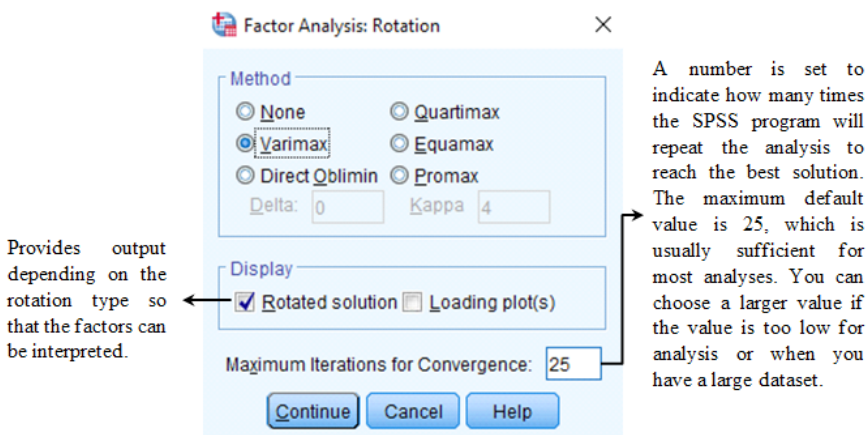


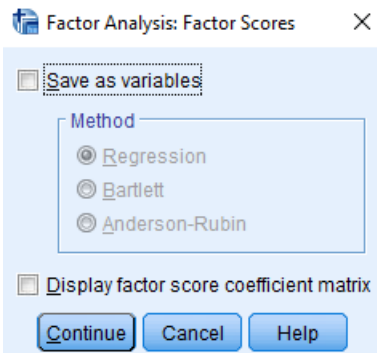
Figure 3.6. “Rotation” Screen

Factors are rotated for better interpretation because non-rotational factors are ambiguous. The purpose of the rotation is to try to load each variable on as few factors as possible, while still achieving an optimally simple structure that

maximizes the high number of loads on each variable. Ultimately, the simple structure seeks to ensure that each factor defines a separate set of interrelated variables. This will make interpretation easier.

In the “Method” part, there are two types of rotation: Orthogonal (varimax, quartimax and equamax) and Oblique (Direct Oblimin, Promax). These are related to the direction of the rotation. Orthogonal rotation is where the factors are rotated 90 degrees from each other [23]. Orthogonal rotation is generally preferred in social sciences and the most widely-used techniques are *Quartimax and Varimax* rotation. Varimax outputs as many factors as possible and loads as few expressions as possible on this factor. Thus, it creates a structure with more factors. In Quartimax, the situation is the opposite. In other words, it creates a structure with few factors by gathering the expressions under one factor as much as possible.

6. At this stage, there is no need to select any section on the “Scores” screen.



It is a necessary screen for generating factor scores. The image factor score coefficient matrix shows the correlation between the coefficients used to generate the factor scores by multiplying the factors. At this stage, we do not mark on this screen.

Figure 3.7. “Factor Scores” Screen

7. Relevant sections are marked on the “Options” screen.

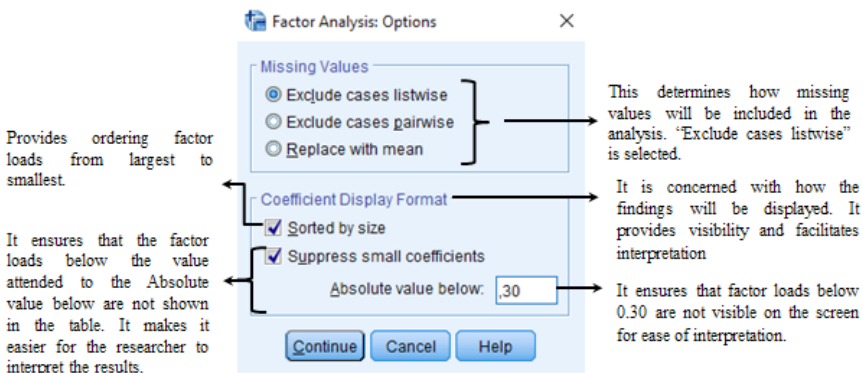


Figure 3.8. “Options” Screen

After all the selections are completed, the “OK” button is pressed and the OUTPUT of SPSS is interpreted.

Table 3.1. “KMO and Bartlett’s Test” Tables

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,869
Bartlett's Test of Sphericity	Approx. Chi-Square	1057,504
	df	105
	Sig.	,000

The “*KMO and Bartlett’s Test*” table shows the test results for evaluating the suitability of the data for factor analysis. These tests are the Sample Adequacy Measurement (Kaiser-Meyer-Olkin (KMO)) and Bartlett’s Test of Sphericity.

The adequacy of the sample size for factor analysis is determined according to the Kaiser-Meyer-Olkin Measure (KMO). The KMO value, which takes a value between 0 and 1, indicates that that the sample size is large enough for factor analysis when it is greater than 0.70. In the literature, there also other recommendations that the KMO value should be 0.5, 0.6, or 0.7 and above [11,12].

The Significant value for the suitability of our data set for factor analysis should be statistically meaningful, i.e., $p < .05$ (in our example it is Sig=0.000). This means that if this prerequisite condition is not met, reliable factors cannot be provided. In this case, it is necessary to increase the number of samples or to remove the expressions causing the scattered correlation models from the analysis and to perform the factor analysis again.

Table 3.2. “Communalities” Table

Communalities

	Initial	Extraction
JobSat1	1,000	,717
JobSat2	1,000	,705
JobSat3	1,000	,532
JobSat4	1,000	,621
JobSat5	1,000	,522
TransLe1	1,000	,624
TransLe2	1,000	,720
TransLe3	1,000	,520
TransLe4	1,000	,595
CharLe1	1,000	,607
CharLe2	1,000	,569
CharLe3	1,000	,697
CharLe4	1,000	,575
CharLe5	1,000	,618
CharLe6	1,000	,579

The “Communalities” table is used to determine the eigenvalue deduction based on certain criteria. However, to preserve the simplicity of our example, we will adhere to the Kaiser criteria. Both methods show that our dataset has three significant factors.

Extraction Method: Principal Component Analysis.

Table 3.3. “Total Variance Explained” Table

Component	Total Variance Explained								
	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6,288	41,917	41,917	6,288	41,917	41,917	3,709	24,725	24,725
2	1,624	10,826	52,743	1,624	10,826	52,743	2,907	19,380	44,105
3	1,287	8,583	61,326	1,287	8,583	61,326	2,583	17,221	61,326
4	,785	5,256	66,582						
5	,708	4,721	71,283						
6	,693	4,623	75,906						
7	,604	4,024	79,930						
8	,525	3,497	83,426						
9	,485	3,232	86,659						
10	,434	2,896	89,555						
11	,425	2,835	92,390						
12	,380	2,534	94,924						
13	,326	2,172	97,096						
14	,242	1,615	98,711						
15	,193	1,289	100,000						

The column to be interpreted.

Extraction Method: Principal Component Analysis.

The “Total Variance Explained” table is important and shows how many significant factors are formed from the expressions included in the analysis. In the table, all factors are arranged in descending order according to the most explained variance.

The “Initial Eigenvalues” column shows that all the expressions included in the research are gathered under 15 factors and the variance of each factor. In the “Extraction Sums of Squared Loadings” and “The Rotation Sums of Squared

Loadings” columns, factors with eigenvalues less than 1 are not displayed. As previously mentioned, the value “1” was selected in the “Eigenvalues greater than” box on the “Extraction” screen while performing factor analysis. Therefore, factors with eigenvalues less than 1 are not seen in these columns.

The “Extraction Sums of Squared Loadings” and “Initial Eigenvalues” columns show the eigenvalues and variances before rotation, while the “Rotation Sums of Squared Loadings” column shows the eigenvalues and variance after rotation.

In our study, we make all our comments according to the values in the “Rotation Sums of Squared Loadings” column. Here, the eigenvalues should be considered; in other words, the factors whose Eigenvalues are greater than 1. When the region shown in the dashed rectangle is examined (Table 3.3), all the statements in the study are gathered under three (3) factors. The eigenvalue of the first (1) factor is 3.709 and the variance explained is 24.725%, the eigenvalue of the second factor is 2.907 and the variance explained is 19.380%, and the eigenvalue of the third factor is 2.583 and the variance explained is 17.221%. The total explained variance is 61.326%. The total explained variance must be 50% or more.

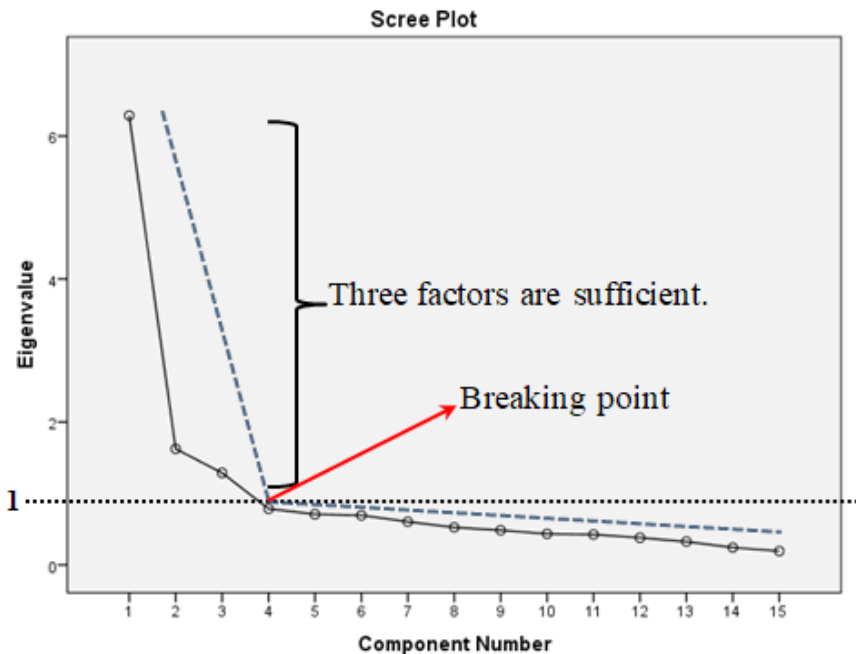


Figure 3.9. “Scree Plot”

The Scree Plot test is used to determine how many factors occur in factor analysis. It works based on the logic that factors with an eigenvalue over 1 are preserved, which is Kaiser’s basic criterion [24]. As can be seen, the Scree Plot consists of eigenvalues and factors. The Scree Plot test gives reliable results on data with at least 200 samples.

The Scree Plot (Figure 3.9) shows how many factors the expressions included in the study are gathered under. It is a graphical representation of the data shown in the previous “Total Variance Explained” table (Table 3.3). The “Component Number” indicates how many factors the expressions are gathered in (15 factors). However, the eigenvalues of all these factors are not greater than 1. Factors below the breakpoint (eigenvalue=1) should be excluded from the analysis. Three factors above the breaking point provide sufficient qualifications. If you recall, in the previous “Total Variance Explained” table (Table 3.3), we determined that the expressions in the study were gathered under 3 factors. The validity of these factors was confirmed by the Scree Plot chart.

Table 3.4. “Component Matrix” Table

Component Matrix^a

	Component		
	1	2	3
CharLe3	,772		-,313
CharLe5	,763		
JobSat5	,710		
CharLe6	,696		-,304
JobSat4	,679		,366
JobSat2	,669	-,339	,378
CharLe4	,668		-,341
CharLe1	,668		-,362
CharLe2	,658		-,348
TransLe2	,638	,530	
JobSat1	,613	-,308	,496
JobSat3	,578	-,317	,310
TransLe4	,545	,541	
TransLe3	,409	,567	
TransLe1	,554	,561	

Extraction Method: Principal Component Analysis.

Component Matrix: The Factor Matrix table shows the factor loads before the rotation and we do not have any operations with this table at this stage. The important thing is the “Rotated Component Matrix” table, which shows the factors after rotation. We make our comments according to this “Rotated Component Matrix” table.

Table 3.5. “Rotated Component Matrix” Table
Rotated Component Matrix^a

	Component		
	1	2	3
CharLe3	,764		
CharLe1	,748		
CharLe4	,722		
CharLe2	,721		
CharLe6	,703		
CharLe5	,666	,321 ← ?	
JobSat1		,826	
JobSat2		,790	
JobSat4		,706	
JobSat3		,683	
JobSat5	,464	,498 ← ?	
TransLe2			,787
TransLe1			,749
TransLe4			,732
TransLe3			,714

Extraction Method: Principal Component Analysis.

When interpreting factors, it is necessary to control factor loadings to determine the strength of relationships. The signs of the loads indicate the direction of the correlation and do not affect the interpretation of the magnitude of the factor loading or the expressions to be preserved [25]. The square of the factor load of an expression indicates the extent to which the expression explains the factor (variance). For example, the expression CharismaticLeader3 explains 58.3% ($(0.764)^2 = 0.583$) of the variance of the charismatic leadership factor.

As can be seen, there are no factor loads of 0.30 and below in the “Rotated Component Matrix” table, which facilitates interpretation. If you recall, we assigned this by typing 0.30 into the “Absolute value below” box on the “Options” screen. If this had not been done, factor loads would be seen in places that are now empty and interpretation would have been difficult due to the presence of many numbers. It is worth remembering here that the loads of all the factors in the blanks are values of 0.30 or below, and it is therefore not necessary to take them into account.

Researchers can use one of two approaches when deciding which expressions to delete: (1) deletion of expressions below a certain threshold by examining their factor loadings and (2) cross-loading of expressions.

(1) *Factor Loading*: A general rule of thumb used to determine factor reliability is to examine the relationship between factor loading and sample size. In other words, the larger the sample size, the more a factor is allowed to have a low factor loading and it is not excluded from the research [26]. Also, as a rule of thumb, at a significance level of 0.01, a factor load after rotation for sample sizes of at least 300 must be at least 0.32 to be considered statistically significant [11]. A factor load of 0.32 indicates that it explains about 10% of the overlapping variance. However, the threshold value of the factor load can also be determined by the researcher. For example, if the researcher wants the factor to consist of strong expressions, they can set the threshold value as 0.5. However, it would be an appropriate approach for the researchers not to prefer a factor load below 0.32 as a threshold value. In our factor analysis, the expression with the lowest factor load is JobSat5 (.498). This expression can be removed from the analysis. However, the decision to exclude the expression of satisfaction5 from the analysis should not be made immediately. Exclusion from the analysis should firstly be made according to the cross-loading condition, and factor loadings should be checked after the cross-loading problem is solved.

(2) *Cross-Loading*: Cross-loading occurs when an expression is loaded on two or more factors. In general, researchers expect the factor loadings of the expressions to be gathered under the relevant factor in the “Rotated Component Matrix” table. In our analysis, we expect to find 6 statements of the charismatic leadership in the factor related to charismatic leadership (column 1), 5 statements related to job satisfaction in the factor of job satisfaction (column 2) and 4 statements related to transactional leadership in the factor related to transactional leadership (column 3). However, the CharLe5 and JobSat5 statements are cross-loaded. Therefore, these statements should be removed from the analysis in order, and factor analysis should be performed again after each removal. Before this removal process, it is also useful to mention the following. Depending on the design of the study, an expression that is cross-loaded may be retained or omitted, assuming the latent nature of the variable. If the decision is made not to remove the statement, the researchers should explain this situation theoretically in their study.

Returning to our analysis, we found that two expressions were cross-loaded (JobSat5 and CharLe5). Firstly, the process of extracting expressions should be

performed one-by-one, and factor analysis should be performed again after each expression is removed. To decide which expression should be subtracted first, the difference between the cross-loading is examined and the expression with the closest value is subtracted first. The cross-loadings of JobSat5 are 0.464 and 0.498, while those of CharLe5 are 0.666 and 0.321. The differences between the cross-loadings are 0.034 (0.498-0.464= 0.034) for JobSat5 and 0.345 (0.666-0.321=0.345) for CharLe5. For this reason, we first remove the JobSat5 expression from the analysis and perform the factor analysis again.

When the processes described above for factor analysis are repeated (with JobSat5 excluded from the factor analysis), the following tables will be displayed.

Table 3.6. “KMO and Bartlett’s Test” Table

KMO and Bartlett’s Test			
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.			,862
Bartlett's Test of Sphericity	Approx. Chi-Square		965,459
	df		91
	Sig.		,000

The KMO value is greater than 0.70 and the Significant value indicating its suitability for factor analysis is significant (p <0.05), indicating that we can continue with the factor analysis.

Table 3.7. “Total Variance Explained” Table

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5,827	41,623	41,623	5,827	41,623	41,623	3,532	25,229	25,229
2	1,615	11,533	53,157	1,615	11,533	53,157	2,650	18,926	44,155
3	1,284	9,171	62,327	1,284	9,171	62,327	2,544	18,173	62,327
4	,781	5,582	67,909						
5	,702	5,011	72,921						
6	,692	4,938	77,860						
7	,574	4,099	81,959						
8	,497	3,548	85,506						
9	,438	3,131	88,637						
10	,427	3,051	91,689						
11	,382	2,728	94,416						
12	,328	2,341	96,758						
13	,243	1,737	98,494						
14	,211	1,506	100,000						

The column to be interpreted.

Extraction Method: Principal Component Analysis.

As a result of the second-factor analysis performed after removing the JobSat5 expression, it is seen that all expressions are still gathered under 3 factors. However, the variances explained by the factors have changed compared to the previous analysis. As a result of the second analysis, the eigenvalue of the

first (1) factor is 3.532 and the variance explained is 25.229%, the eigenvalue of the second factor is 2.650 and the variance explained is 18.926%, and the eigenvalue of the third factor is 2.544 and the variance explained is 18.173%. The total explained variance is 62.327%. It is important to remember that the total explained variance should be 50% or more.

Table 3.8. “Rotated Component Matrix” Table

Rotated Component Matrix^a

	Component		
	1	2	3
CharLe3	,766		
CharLe1	,753		
CharLe4	,724		
CharLe2	,724		
CharLe6	,704		
CharLe5	,669	,312 ← ?	
JobSat1		,828	
JobSat2		,804	
JobSat4		,702	
JobSat3		,676	
TransLe2			,789
TransLe1			,750
TransLe4			,733
TransLe3			,715

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

When the factor analysis is performed again, it is seen that the expression Charismatic_leader5 is still cross-loaded. Although this statement seems to be problematic, it would be appropriate not to exclude this statement from the analysis because the high factor load (0.669) is gathered under the right factor and the value between two-factor loads is greater than 0.1 ($0.669 - 0.321 = 0.348$).

Table 3.9. “Component Transformation Matrix” Table**Component Transformation Matrix**

Component	1	2	3
1	,701	,534	,473
2	-,227	-,462	,857
3	-,676	,708	,203

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

Our last table is the “*Component Transformation Matrix*” table. The values in this table are used to determine whether the rotation technique is appropriate. If you are not familiar with Field’s [12] factor analysis techniques, the “Component Transformation Matrix” can be ignored when interpreting Orthogonal (Varimax, Quartimax, and Equamax) rotation methods.

REPORTING FINDINGS

Exploratory factor analysis was performed on the data using the Varimax technique, one of the principal component analysis vertical rotation techniques. As a result of the analysis, the Kaiser Meyer Olkin (KMO) test result (0.862) indicates that the sample size is sufficient, while the Bartlett sphericity test result being significant ($\chi^2 (91) = 1057.504$, $p < 0.001$) shows that the correlational relationship between the expressions is suitable for factor analysis (*For readers: The number 91 is the df value in the KMO and Bartlett's Test table, 1057.504 is the Approx. Chi-Square value in the same table, and 862 is the KMO value*).

In the first factor analysis, one of the job satisfaction expressions was excluded from the analysis and was not included in the study because the factor load was low and had a tendency to cross-load. As a result of the second EFA analysis, taking into account the variances explained by the factors, it was ensured that those with an eigenvalue greater than 1 were formed a factor, and thus, the three-factor structure was deemed to be appropriate. In this context, the first factor explains 25.229% of the total variance, the second factor explains 18.926%, and the third factor explains 18.173%. The total explained variance is 62.327% and the factor loads of the items are 0.669 and above. The results are presented in Table 3.10.

Table 3.10. Factor Analysis Results

Factors and Statements	Factor Loadings		
	1	2	3
Charismatic Leader			
Charismatic Leader3	.766		
Charismatic Leader1	.753		
Charismatic Leader4	.724		
Charismatic Leader6	.704		
Charismatic Leader5	.669		
Job Satisfaction			
Job Satisfaction1		.828	
Job Satisfaction2		.804	
Job Satisfaction4		.702	
Job Satisfaction3		.676	
Transactional Leader			
Transactional Leader2			.789
Transactional Leader1			.750
Transactional Leader4			.733
Transactional Leader3			.715
Eigenvalues	3.532	2.650	2.544
% Variance	25.229	18.926	18.173
% Total Variance		62.327	

CONFIRMATORY FACTOR ANALYSIS AND AMOS

In confirmatory factor analysis, it is possible to test a previously determined model or hypothesis about the relationship between variables [27,28]. In terms of the operations performed, exploratory factor analysis is a method that is largely used to test the construct validity of newly-created scales and aims to reach fewer unobserved factors based on the observed variables in the scale. Confirmatory factor analysis, on the other hand, is performed to test whether scales that have been previously discovered and gathered under fewer factors are similar in the sample in which the research was conducted. In this context, it can be said that four different models can be tested in confirmatory factor analysis [29,30]. These models are called the *single-factor model*, *first-level multifactor model*, *second-level multifactor model*, and *unrelated model*.

Single Factor Model

This is a model in which all observable variables are gathered under a single factor. An example model is shown in Figure 3.11. The essence of the model is that all observable variables are gathered under a larger and more inclusive upper variable. An example of this is the collection of all observed variables under a single factor in a six-item self-efficacy scale

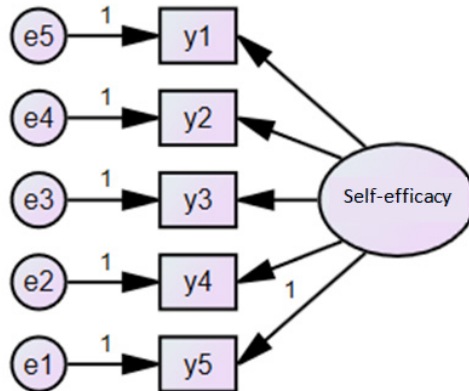


Figure 3.11. Single Factor Model

First-Level Multi-Factor Model

This is a model in which more than one observed variable is collected under a single factor that is independent of each other but related. An example model is presented in Figure 3.12. The basis of the model is that the observable variables are collected under more than one independent dimension. It could be that this type of model is determined completely theoretically by the researcher and tested with confirmatory factor analysis, or it could be a model obtained as a result of exploratory factor analysis. An example is the transformational leadership scale presented in Figure 3.12. The scale has four sub-dimensions: idealized influence, inspirational motivation, intellectual stimulation, and individual consideration. As an example, each dimension is shown here with only three items. In this example, the researcher performs confirmatory factor analysis on twelve items. If the goodness-of-fit indices are within acceptable limits, as will be shown in the next chapter, the first-order multi-factor model is validated.

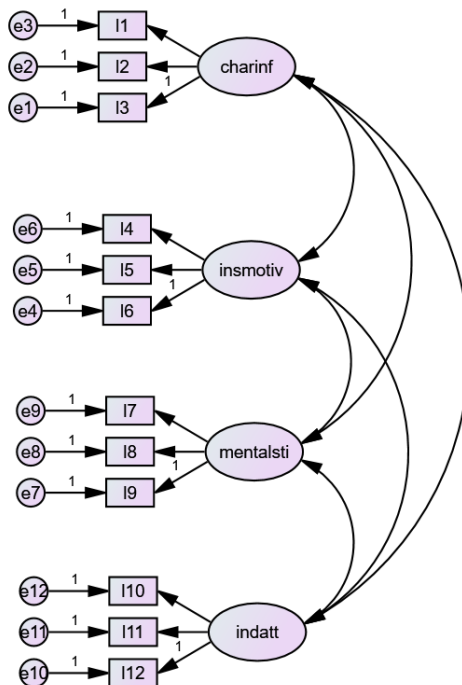


Figure 3.12. First-Level Multi-Factor Model

Second Level Multi-Factor Model

This can be defined as a model in which the observed variables are gathered under more than one unrelated factor, and these factors are then gathered under a larger and more inclusive factor. In Figure 3.13, the transformational leadership scale is shown as an example. The basis of the model is that the observable variables are gathered under four independent sub-dimensions, and these factors then come together under the transformative leadership variable, which is a more inclusive factor.

When the example in Figure 3.13 is examined, the question “Where is transformational leadership?” comes to mind. Although latent sub-dimensions are seen in the first-level multi-factor model, the variable that is the main research topic does not appear. Here, the model in which transformational leadership is included in the model as a single overarching latent variable is the second-level multi-factor model.

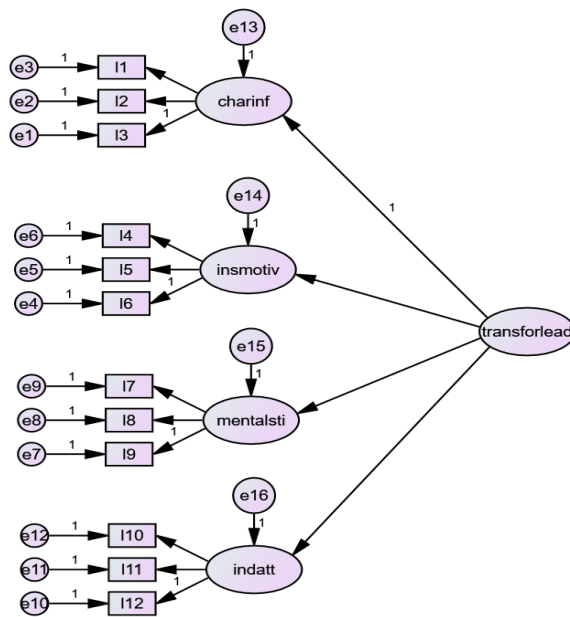


Figure 3.13. Second Level Multi-Factor Transformational Leadership Model

While some scales are first-level multifactorial due to their structure, some scales can be second-level multifactorial. For this reason, when applying confirmatory factor analysis, second-level multi-factor models of multidimensional scales should also be tested. Some scales can be used as both multi-factor and single-factor scales. Researchers should make this decision

according to the model and hypotheses in their study. For example, positive psychological capital, which is a current issue, can be used as both a first-level multi-factor with four factors and as a second-level multi-factor with a single inclusive dimension. At this point, the important factor is that the researcher's purpose and the data set are compatible with the structure being used.

Unrelated Model

The last model is the unrelated model, which can be defined as the model in which the observed variables are gathered under more than one unrelated factor. The unrelated model of the transformational leadership scale, which is given as an example of a first-level multi-factor model, is presented in Figure 3.14.

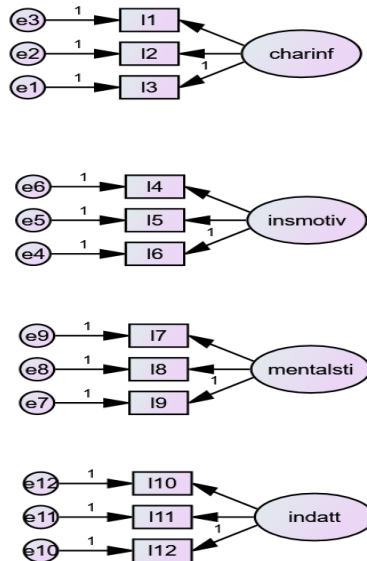


Figure 3.14. Unrelated Model

The process in confirmatory factor analysis involves testing different models and identifying which model is the most suitable. Particularly in multi-factor scales, it is necessary to test different models and decide which model is the most suitable by examining the goodness of fit and indices values. In this regard, the important factors are that the structure of the scale can be defended theoretically and that the structure to be used in accordance with the purpose (multi-dimensional or one-dimensional) can be demonstrated to be compatible with the data set.

CONFIRMATORY FACTOR ANALYSIS WITH AMOS

When performing confirmatory factor analysis with AMOS, the model of the scale is firstly drawn based on the theory and previous research findings. DataSeti-dfa.sav will be used for this sample application. The Self-Efficacy Scale in the sample data set has a two-dimensional structure comprised of job begin and job continue. The scale has been used in two dimensions in many previous studies. The *Job begin* sub-dimension consists of items 3-4-5-6-8 and the *Job continue* sub-dimension consists of items 1-2-7. Since this is the first time, the process of drawing a model with AMOS will be explained step by step, and then confirmatory factor analysis will be performed. The following AMOS applications in the book will continue with the drawn version of the model.

Before drawing a model with AMOS, particularly in complex models, drawing the model manually on paper and then editing it on AMOS makes the process much easier. To draw a model with AMOS, first click on “AMOS Graphics” and open the AMOS program (Figure 3.15).

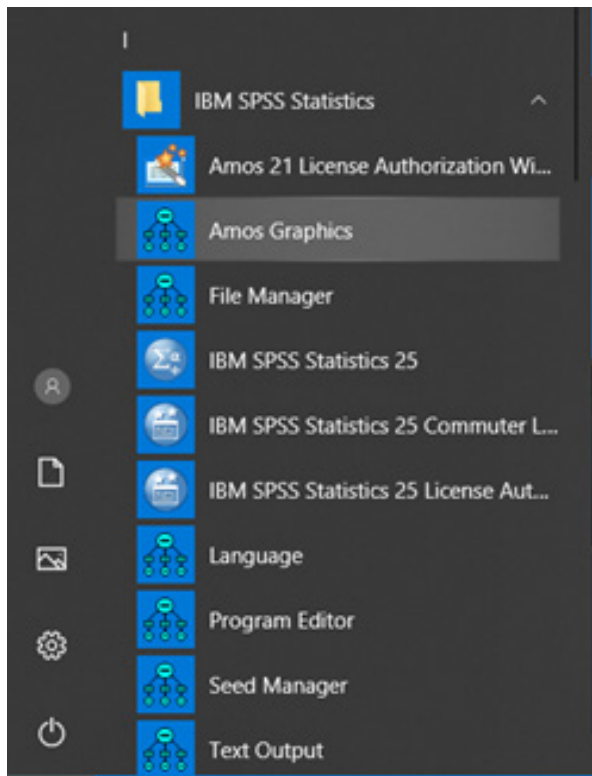


Figure 3.15. Running AMOS Program

Figure 3.16 shows the initial screen that opens. As can be seen in the figure, when the program starts, a blank worksheet will be displayed. The first thing to do here is to draw a model that we have built in our research, or better, a sketch on a piece of paper, step by step, on AMOS. In our example, there are two latent dimensions (*Job begin and Job continue*). There are five items for *Job begin* and three for *Job continue* on the scale. In the AMOS program, latent variables are symbolized by ellipses and observed items by rectangles. Therefore, the first task is to draw two ellipses and eight rectangles. In order to draw these shapes, the shape to be drawn is first selected with the left mouse button, and then the blank worksheet is clicked. This draws that figure on the page. Practicing this several times will allow users to increase the visibility of the drawn models.

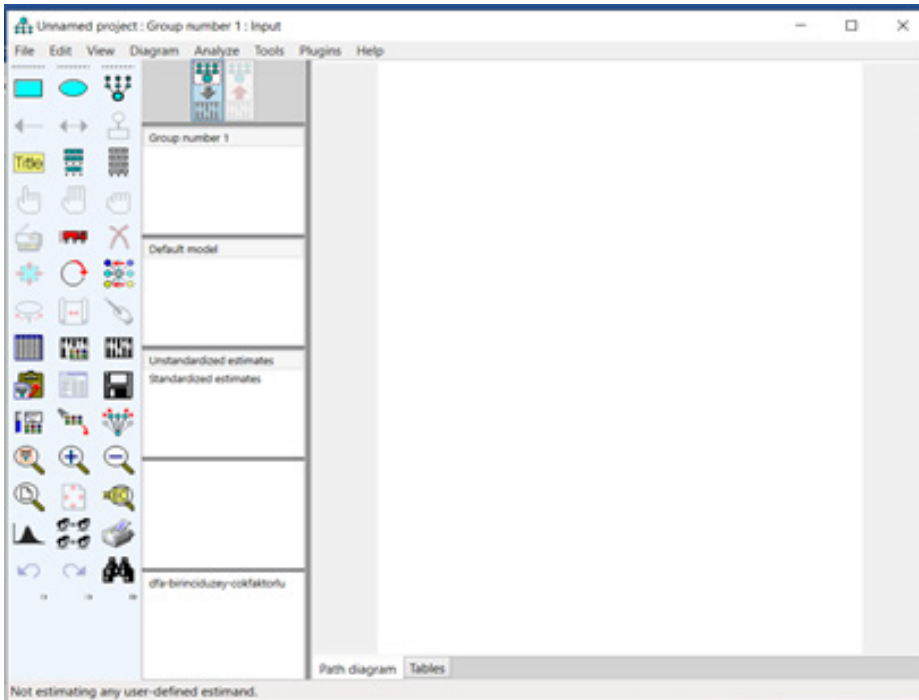


Figure 3.16. AMOS Opening Screen

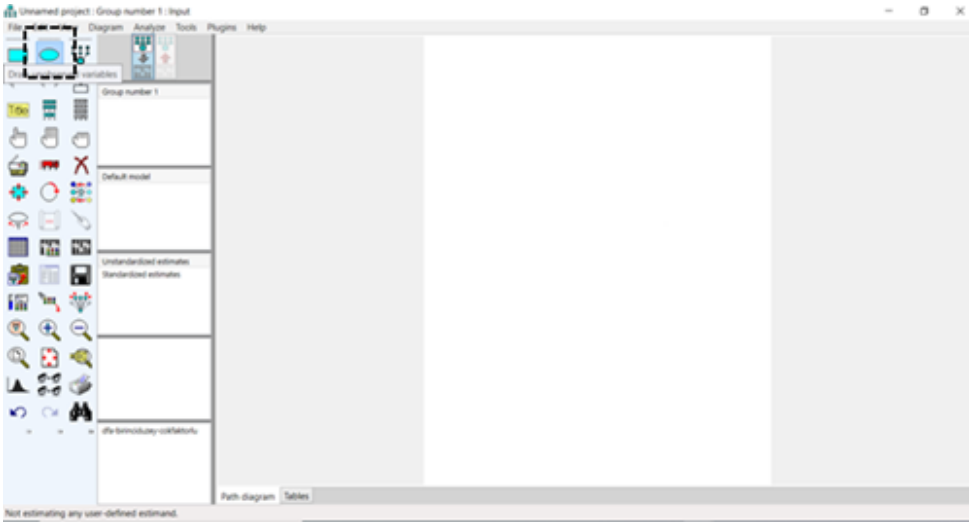


Figure 3.17. An Unobservable (Latent) Variable Icon

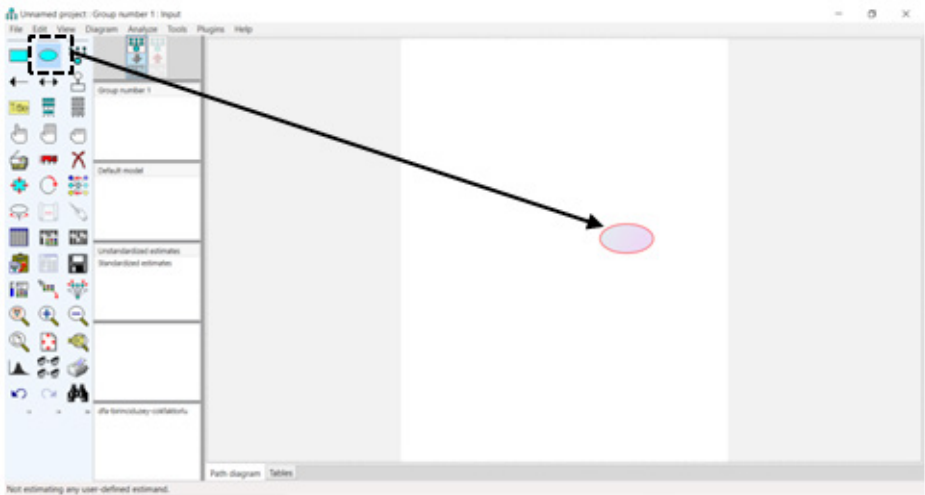


Figure 3.18. Process of Drawing an Unobservable (Latent) Variable

The first unobservable (Latent) variable, “ Job begin “, is selected by touching the ellipse icon, as shown in Figure 3.17, and it is drawn on the blank worksheet until it reaches the desired size by holding down the left mouse button, as shown in Figure 3.18. The same process is also repeated for the second unobservable (latent) variable “Job continue”. Subsequently, five and three observed variables of these variables are respectively drawn by selecting the rectangle icon. Thus, Figure 3.19 is formed.

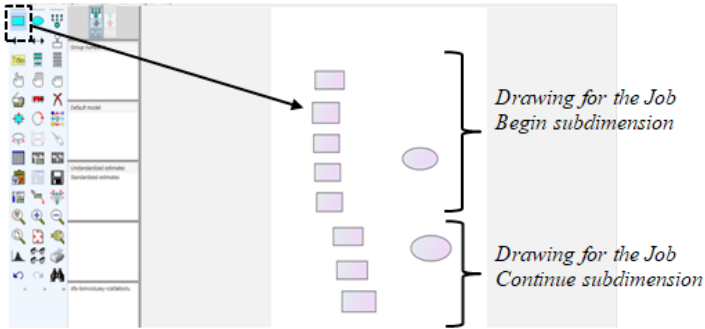


Figure 3.19. Process of Drawing the Unobservable (Latent) and Observed Variables

Subsequently, a path is drawn for each of the observed variables from the first and second dimensions, respectively. According to the main logic of AMOS, one of the paths should be selected as the priority path. It could be any of the paths. As seen in Figure 3.20, the first path is selected as the priority path (the value 1 is seen on them). Although the model may seem highly complex when the paths are first drawn, a much smoother shape will be obtained when the “Touch” icon is selected and the unobservable variables are left-clicked.

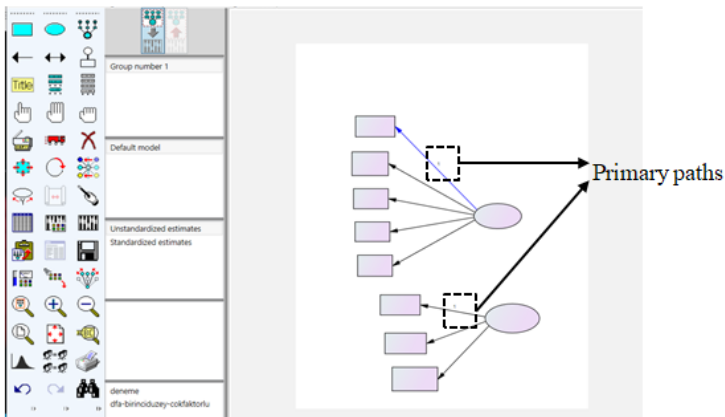


Figure 3.20. Drawn Model with Regression Paths

In addition, to assign a value of 1 to the first path, it is necessary to double-click on the path. Thus, the “Object Properties” screen shown in Figure 3.21 will be opened and the desired information can be entered on the information screen. Here, information such as the name of the object, parameter values, size, color, and format can be edited via the drop-down box. These revisions can be made for all objects in the model. In particular, it is necessary to use this function to assign names to objects.

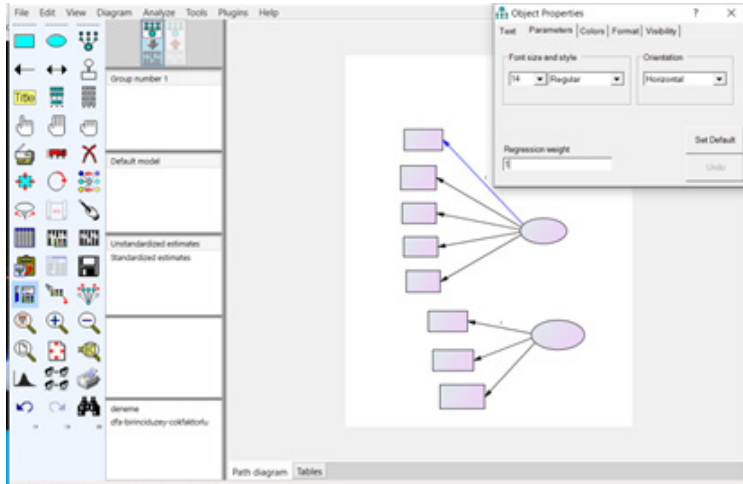


Figure 3.21. “Object Properties” Menu

To continue drawing the model, an error term should be added to all observed and unobserved variables on which regression is applied, and the paths from the error terms to the variables must be drawn and the value “1” must be assigned to them. Drawing and assigning values to them can be performed as previously described. Note that the direction of the arrows will be from the error term to the variable. Drawing the error terms and arrows one-by-one is a long and difficult task. However, an error term can be automatically added to each observed variable instead by selecting the “Add error” term button. This will result in the model displayed in Figure 3.22.

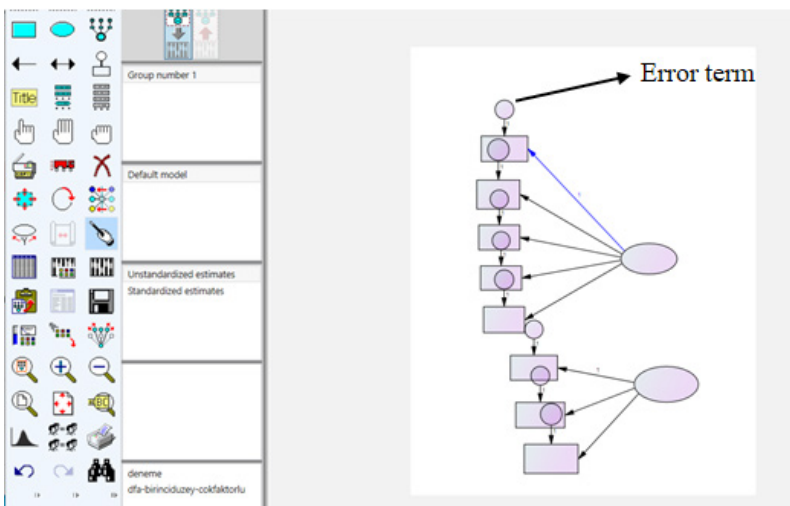


Figure 3.22. Adding Error Terms

The model created is visually quite complex. Certainly, there are ways in which such models can be drawn in a more aesthetic and visually simple manner. It will firstly show the long and arduous path, and then the short and much more visually correct model drawing will also be shown. However, to understand the functions of the different buttons, a long way model is drawn first.

The next step in the process is to name the variables and assign the observed variables in the data set to the relevant places. In the figure, two sub-dimensions are latent variables and the error terms are also latent variables. Therefore, they can be given any name. Although objects can be named one-by-one, this process can be quite time consuming. Instead, the “Name Unobserved Variables” command under the “Plugins” menu is run. Thus, as can be seen in Figure 3.23, all unobservable variables are assigned names automatically. As can be seen, the program gives sequential names starting with “e” to the error terms, and names starting with “F” to the sub-dimensions.

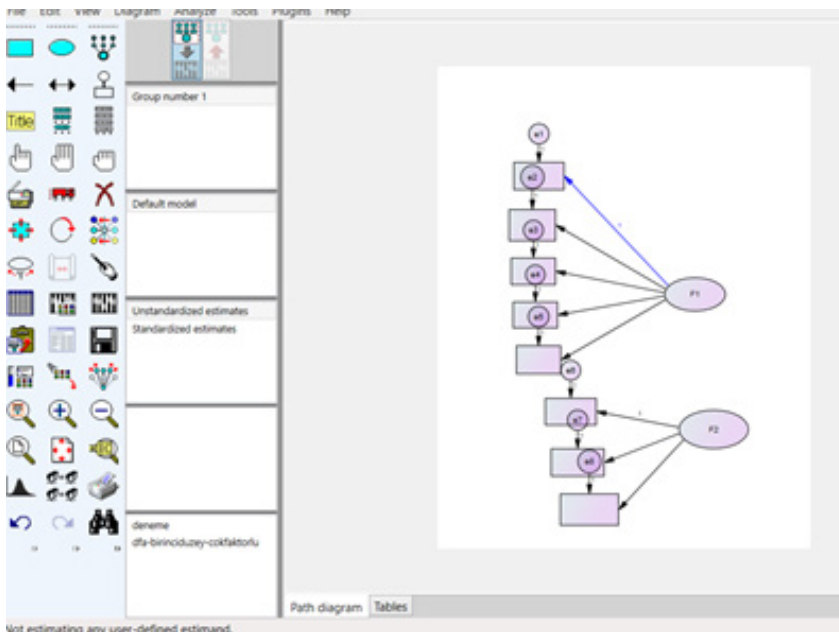


Figure 3.23. Naming Error Terms

The last step at this stage is to assign variables from the data set to the observed variables. In this process, the relevant data file is firstly opened by selecting the “File Name” tab in the “Data Files” command under the “File” menu. For the purposes of this study, a data file named VeriSeti-dfa will be opened. The actions to be taken are shown in Figure 3.24.

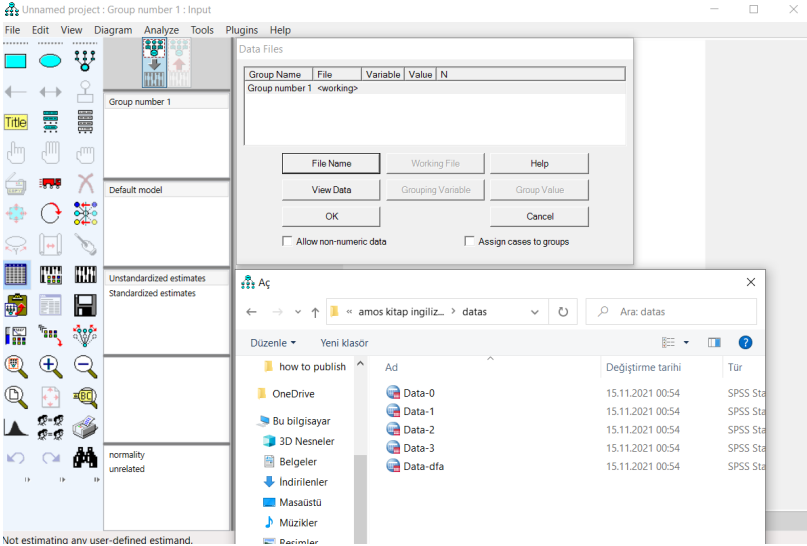


Figure 3.24. Opening the Data Set

After the data set is opened, the variables in the data set can be listed by clicking the “List Variables in Data Set” button, as shown in Figure 3.25.

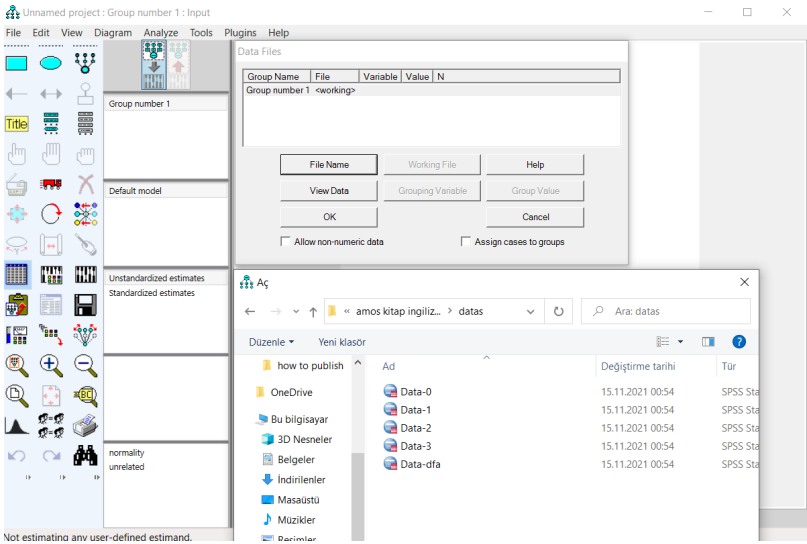


Figure 3.25. List Variables in Dataset Command Screen

It should be remembered that the Self-Efficacy Scale, on which confirmatory factor analysis is performed in this model, has a two-dimensional structure comprised of Job begin and Job continue. In the currently drawn model, these

are shown as two separate factors, F1 and F2. The Job begin sub-dimension consists of items 3-4-5-6-8 and the Job continue sub-dimension consists of items 1-2-7. The mentioned items are moved from the drop-down data setlist into the related observed variable by left clicking with mouse. Thus, the model seen in Figure 3.26 is created. In the model, the variables F1 and F2 were double left clicked, and their names were changed. In AMOS, *Turkish characters, special signs or spaces* should not be used when naming objects, variables, and even files and data in some versions

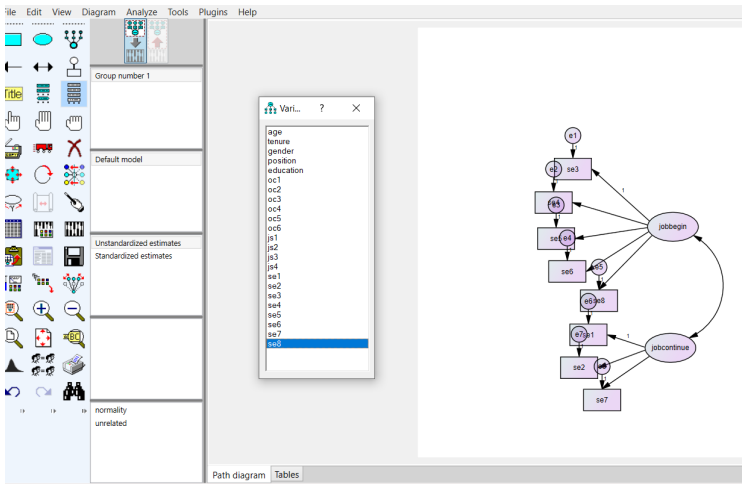




Figure 3.26. Data Assigned Complete Model

The model drawing process described thus far involves the use of different buttons, which takes a little longer and creates a visually more irregular appearance. However, the model creation process can be performed much faster, easier, and more visually. For this purpose, first of all, a blank page is opened by selecting the “New” command under the “File” menu. In the model we will draw, there are two latent sub-variables, one with five observed variables and the other with three observed variables. The main button used for quick and easy model drawing is the “Display”  button. By selecting this button, by left clicking with the mouse on the blank page, a latent variable is created as seen in Figure 3.27, and each time this latent variable is left clicked with the mouse, the observed variable with an error term is added to the latent variable. After repeating this process five times, a latent variable with five observed variables is obtained that is visually very smooth. The same process is repeated slightly below the first figure, creating another latent variable with three observed variables.

In the last stage, the observed variables are moved to the desired location with the “Rotate”  button, and the model shown in Figure 3.28 is formed. As can be seen, in the model, both the numbers 1 on the error terms regressions and observed variable are automatically assigned to the model.

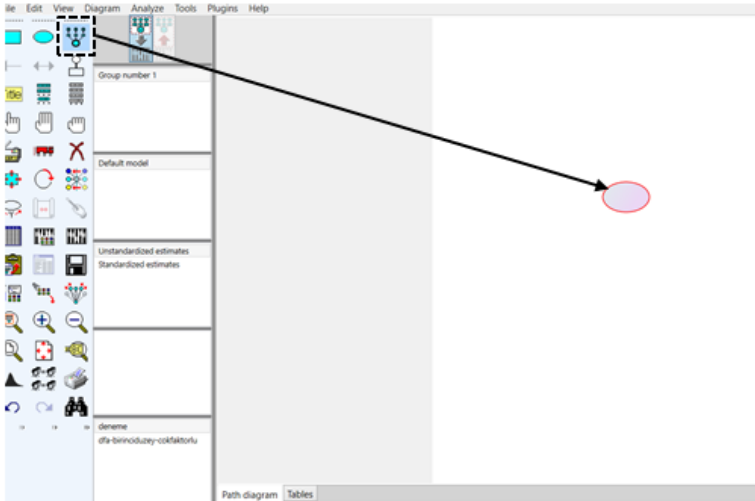


Figure 3.27. Using Add an Indicator Button

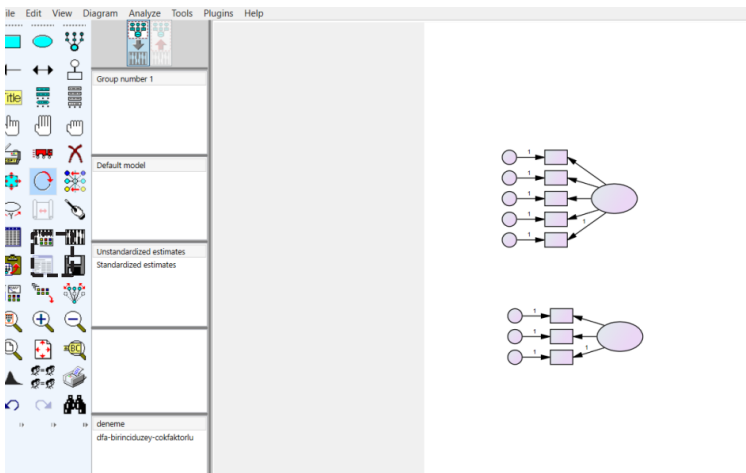


Figure 3.28. Model Formed Using Add an Indicator and Rotate Button

After this stage, by selecting the “Name Unobserved Variables” command from the Plugins menu as previously explained, the latent variables are automatically named, the sub-dimensions are renamed, and the data file is opened and the variables in the data file are moved over the observed variables

one-by-one, as described above. Finally, the model ready to be analyzed shown in Figure 3.29 is obtained by drawing the “Covariance”, which is a double-sided arrow, between the two sub-dimensions.

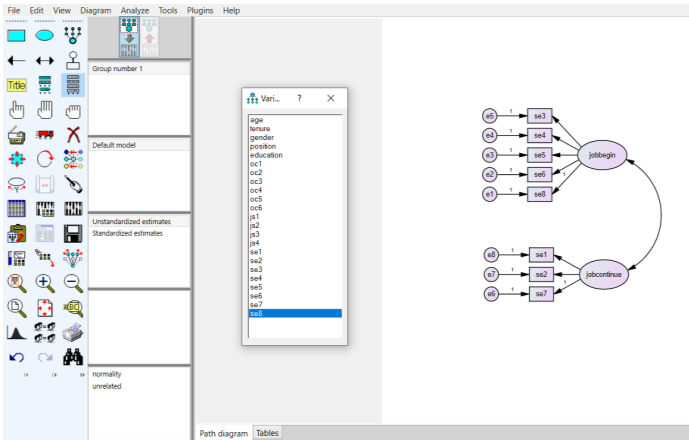


Figure 3.29. Drawn Model Ready for Analysis

Before performing the analysis, the first action that must be taken is to save the model by clicking the “Save” button that looks like a floppy disk, and then select the “Analysis Properties” option under the “View” menu and select the Minimization History, Standardized estimates and Modification indices options from the “Output” tab. This tab is shown in Figure 3.30.

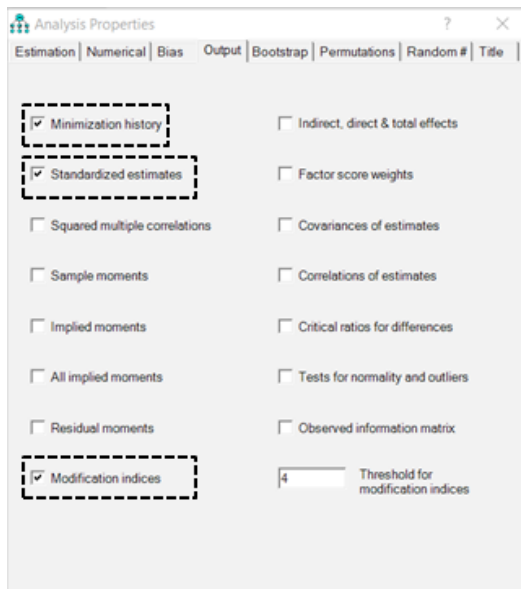



Figure 3.30. Analysis Properties Tab

Click on the “Calculate estimates”  button to start testing the model. Afterward, the test results can be accessed by clicking on the “View text” button.



These findings will be briefly explained step-by-step below.

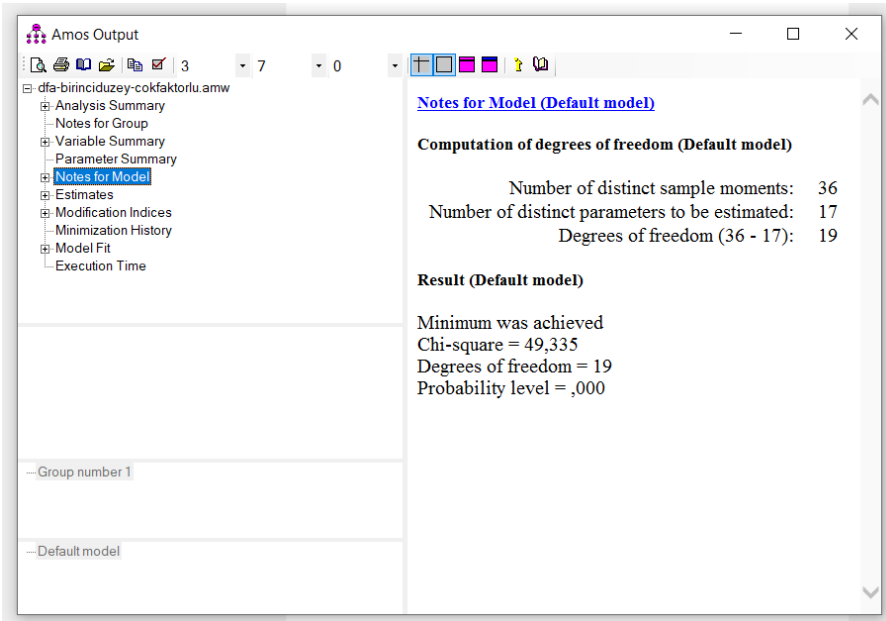


Figure 3.31. AMOS Output Screen

The results of testing the model are presented in Figure 3.31. When the results are opened, the “Notes for Model” tab will be displayed. As seen in the results here, the Chi-square value of the model is 49.335, while the degrees of freedom is found to be 19. For the model to be accepted, the significance level is expected to be statistically not significant. However, since this value is significant in many models, this time the value obtained by dividing the Chi-square value by the degrees of freedom is checked in order for the model to be accepted. If this value is less than 3 for good fit and less than 5 for acceptable fit, the overall fit of the model is accepted. These fit indices values are expressed in different ways in different sources. However, sources such as Meydan and Şeşen [27], Tabachnick and Fidell [31], Schumaker and Lomax [32], and Şimşek [33] could be accepted as references for acceptable indices values. Here, the most commonly-used model goodness-of-fit indices values reported in research will be briefly explained and their acceptable limits will be shared.

Structural equation model tests provide evaluation criteria, or fit indices, of the extent to which the model being tested is suitable for the data collected for that model [27,34]. The compatibility or incompatibility of a model with the data is determined by evaluating the various fit indices revealed as a result of the test. Although the package programs used in SEM calculate some similar goodness-of-fit indices, they also consider different indices. Since this book is purely focused on the AMOS program, some of the important fit indices that will be encountered in AMOS will be explained.

Chi-Square Goodness of Fit: It is the most widely-used statistic to test the overall model fit and can be considered as an initial fit index in a sense. The result of the chi-square test is the test of the fit between the data and the model. In this context, the Chi-Square test tests the hypothesis of whether the developed model and the model that emerges in the covariance structure of the observed variables are different. Provided that the estimated Chi-square statistical value is small, it is decided that the fit is good. In addition, since this is a difference value, if the value is significant, it means that the two models differ significantly from each other. Therefore, the insignificance of the chi-square value indicates the fit of the model.

Degrees of freedom: It is important in the chi-square test. The chi-square value can be significant in most model tests. For this reason, for the significance of the model, the ratio of the chi-square to the degrees of freedom (SD) is considered rather than whether the chi-square alone is significant. In this case, the chi-square/SD ratio being less than 5 shows that the overall fit of the model is acceptable, even if the chi-square is significant. In our example, Chi-square/SD was found to be $49.335/19=2.59$. In this case, although the chi-square was found to be significant, it can be said that the general fit of the model is sufficient.

The “Model Fit” tab is selected on the AMOS Output screen to view other model goodness-of-fit indices. In this case, the values shown in Figure 3.32 are obtained.

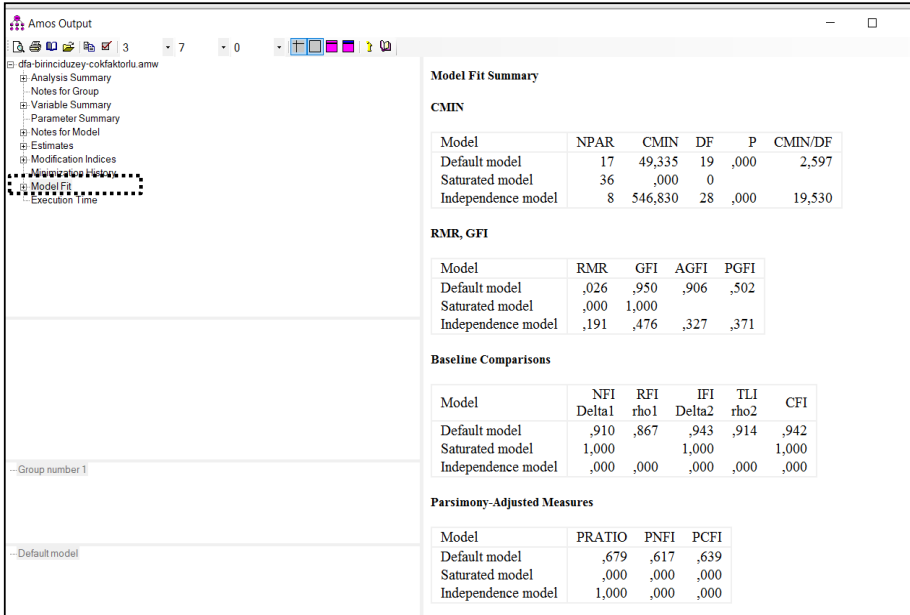


Figure 3.32. “Model Fit” Tab Outputs

Root Mean Square Residual, RMR: The residual-based fit index, RMR, is defined as the square root of the arithmetic mean of the squares of the differences between the obtained and included correlations. Its values can range between 0 and 1, and it is accepted that the fit of the model increases as the value approaches zero. An RMR of less than 0.08 indicates that the model is within acceptable limits. In this case, the RMR value obtained in the model we tested is 0.026, indicating a strong fit.

Goodness of Fit Index-GFI: GFI, one of the absolute fit indices, is a measure of the relative amount of variance and covariance that can be explained by the model. It has a value between 0 and 1, and values above 0.85 are considered acceptable. In the model we tested, the GFI was 0.95, which indicates a good fit.

Adjustment Goodness of Fit Index-AGFI: It is based on a correction to the figure showing the degrees of freedom in a less constrained model by freeing more parameters [33]. It is the GFI value that takes into account the sample size and has a value between 0 and 1, where 0.90 and above is considered a good fit. In our test model, the AGFI was found to be 0.906, indicating a good fit.

Normed Fit Index, NFI: NFI, one of the comparative fit indices, is calculated by dividing the chi-square value of the tested model by the chi-square value of the independent model [35]. The index takes a value between 0 and 1

and can give unreliable results in small sample values. Values above 0.90 in the index are considered acceptable. In our model, the NFI has a value of 0.910, which is within acceptable limits.

Incremental Fit Index, IFI: Another comparative fit index, IFI, was introduced to find solutions to the problems created by the variability in NFI. Index values of 0.90 and above indicate a good fit. In our test model, the IFI value of 0.943 indicates a relatively good fit.

Comparative Fit Index, CFI: An index close to 1 indicates the goodness of fit. Like other comparison indices, CFI is sensitive to sample size. A CFI value of 0.95 and above indicates a good fit. The CFI value of 0.942 in our test model indicates an acceptable fit.

Root Mean Square Error of Approximation, RMSEA: Its value ranges from 0 to 1. Values close to zero (0) are required (minimum error between the observed and produced matrices). Values up to 0.08 indicate acceptable fit. Index values of 0.10 and above indicate poor fit. RMSEA, like other comparative fit indices, is very sensitive to the number of samples. Hu and Bentler [36] found that in small-sample models (sample size less than 250), RMSEA rejects a model that should in fact be accepted. In the model we tested, the RMSEA was found to be 0.086, indicating an acceptable fit.

There is no absolute consensus on which fit indices should be reported for SEM models. This mostly depends on the purpose of the research. However, chi-square, DF, and goodness fit index values are always given in each study, and the results of one or two of the other indices are expressed in accordance with the purpose [27]. The results of our test model are shown in Table 3.12. When the obtained values are examined, it is seen that the index values of the model are within acceptable limits. Therefore, the results of the first level confirmatory factor analysis we tested show that the scale used has two dimensions, with 5 items in the first dimension and 3 items in the second dimension. This type of presentation will suffice when providing test results for most models.

Table 3.12. Confirmatory Factor Analysis Results of the Self-Efficacy Scale

Model	$\Delta\chi^2$	sd	$\Delta\chi^2/sd$	RMSEA	CFI	GFI	AGFI
First-level multi-factor	49.335*	19	2.597	0.086	0.94	0.95	0.90

RMSEA= Root Mean Square Error of Approximation; CFI= Comparative Fit Index; GFI= Goodness of Fit Index; AGFI= Adjusted of Goodness Fit Index, Df= Degree of freedom
* $p > .05$

MODIFICATION OF THE MODEL

In structural equation modeling, after computing the fit indices, the next step is to calculate the modification indices (MI), which give information about the arrangement of the model. MI suggests modifications that should be made to the model by evaluating the covariance between observed and latent variables. These modifications are created based on error terms and represent the amount of chi-square that was not originally predicted in the model but will be gained in the model by making the corresponding revisions. In this context, modifications include proposed new connections between observed or latent variables. The suggested modifications for the model we tested are shown in Figure 3.33 by selecting the “Modification Indices” tab in the output file.

In the data in Figure 3.33:

Column 1 indicates error terms that can be modified (covariates) in AMOS.

Column 2 shows the modification indices coefficients. These numbers represent the change in CMIN/DF when two error terms covariate. Therefore, researchers who want to bring the CMIN/DF value to the desired level should first examine column 2 (Figure 3.33) and make the covariation between the error terms with the highest value. One of the most important questions that arises is what the minimum value of M.I should be. There is no recommended lower threshold for the M.I value to be able to conduct the covariate process in the extant literature. However, in line with the experience and knowledge of the authors of this book, it is recommended not to covariate error terms with M.I values below 15.000.

The values in the 3rd column express the change in CMIN/DF as a percentage when two error terms covariate.

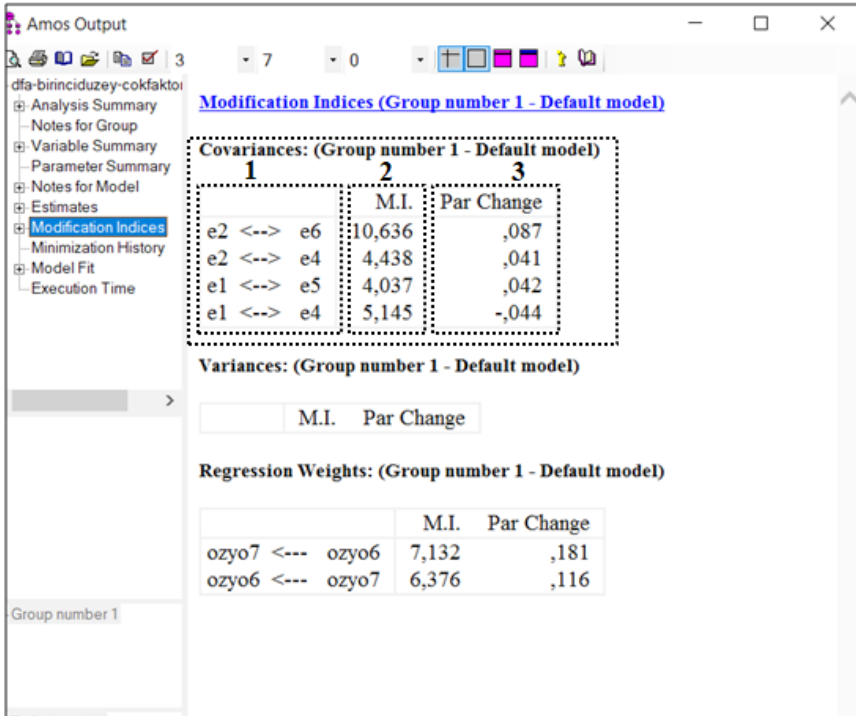


Figure 3.33. Modification Indices

Care must be taken when making modifications because the modification process could contradict the basic principles of SEM [27]. where every covariance could distort the originally designed model. Therefore, every modification in the model made to revise the fit indices must be based on a theoretical background. Otherwise, associating the variables with each other just to correct the indices may mean that a model is tested that will never exist. Particularly care should be taken when making a revision if a modification suggested by MI causes a very large decrease in the chi-square value of the model. The tested model in our cases does not require any modifications. However, if such a modification were to be made as an example, since covariating the error terms e2 and e6 would give the highest revision, the modification would have to be made here first. One such modification model is presented in Figure 3.34.

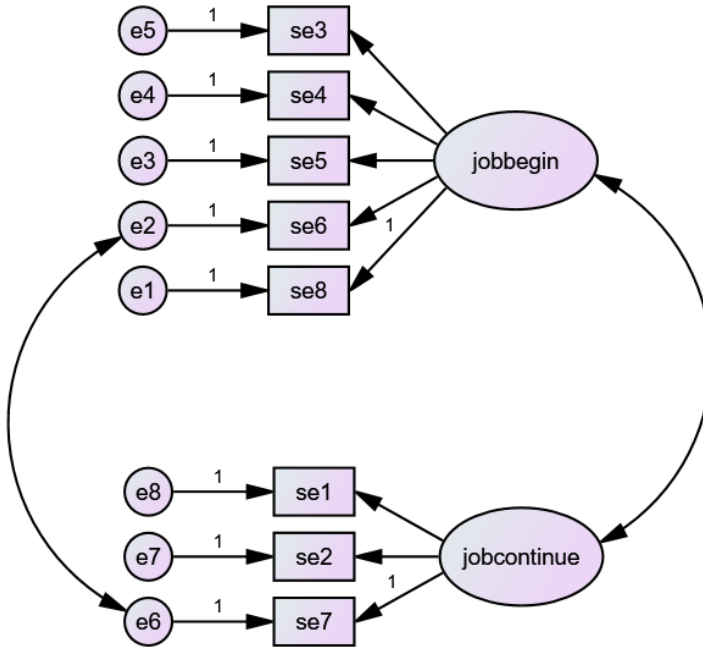


Figure 3.34. Modified Model

The fit indices after modification are also shown in Figure 3.35. As can be seen from the figure, all values improved after modification. While the chi-square/SD value decreased to 2.114, the other fit index values also increased. However, a point to be noted here is that the covariation of e2 and e6 has moved away from the original scale, the original structure has been distorted, and the items *ozyo6* and *ozyo7* have become equal to each other. In other words, these two items are reduced to a single item. The fact that both items are now under different dimensions shows that such a modification would not be very accurate. This modification, made to be purely exemplary, is not a highly preferred situation in real applications. If a modification is to be made, covariation of two items under the same dimension should be preferred.

The screenshot shows the Amos Output window with the following sections and data:

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	18	38,055	18	,004	2,114
Saturated model	36	,000	0		
Independence model	8	546,830	28	,000	19,530

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	,022	,960	,919	,480
Saturated model	,000	1,000		
Independence model	,191	,476	,327	,371

Baseline Comparisons

Model	NFI	RFI	IFI	TLI	CFI
	Delta1	rho1	Delta2	rho2	
Default model	,930	,892	,962	,940	,961
Saturated model	1,000		1,000		1,000
Independence model	,000	,000	,000	,000	,000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI

Figure 3.35. Fit Indices of the Modified Model

After the modifications have been made, the next step is to test the model again. If appropriate fit indices are obtained as a result of the retest, the model will be accepted. Otherwise, a modification could be made again if it is possible. This process is repeated until it is not possible to make a new modification and the model is accepted or rejected according to the resulting fit indices.

RELIABILITY

Reliability of measurement tools is a necessary prerequisite for a healthy interpretation of research results [40]. Therefore, researchers should ensure that the measurement tool they are using is reliable. Reliability refers to the stability of the measurement tool and its consistency over time [40]. In other words, reliability is related to the fact that the measurement tool gives similar results when applied at different times. Of course, it is unlikely that the measurement tool will give the same results every time due to differences in the time of application, as well as changes in the population and sample. However, a strong positive correlation between the results of the measurement tool indicates the reliability of the measurement tool.

Different methods for determining the reliability of measurement tools are available for research in social sciences. These are *test-retest reliability*, *alternative forms*, *split-halves*, *inter-rater reliability* and *internal consistency tests*. The most widely used in social sciences is the internal consistency test [40].

Internal consistency is related to the reliability of the items in the measurement tool. It measures the consistency of the items in the measurement tool and examines how well the measurement tool measures a particular behavior or attribute. Internal consistency of the measurement tool can be determined using item-total correlations, Kuder-Richardson-20 & 21, and Cronbach's alpha. However, the most widely-used method used to determine internal consistency in social science research is to calculate the Cronbach's alpha value [40].

The most popular method for testing internal consistency is the Cronbach's alpha internal consistency value, which was developed by and named after Cronbach [37]. It has been generally accepted in the literature and still maintains its validity. As the Cronbach's alpha value, which takes a value between 0 and 1, approaches +1, the internal consistency of the measurement tool becomes higher

The following formula is used to measure the Cronbach's alpha value:

$$\alpha = \left[\frac{N}{N-1} \right] \left[\frac{S_x^2 - \sum S_j^2}{S_x^2} \right]$$

N = Items in the measurement tool

S_j^2 = Variance of each item

S_x^2 = The sum of variance scores of each item in the measurement tool

Although there are different interpretations in the literature regarding the Cronbach's alpha value, the generally accepted approach is presented in the table below.

Table 3.13. Confidence Coefficients

Cronbach Alpha Value	Interpretation of the Cronbach's Alpha Value
$\alpha \geq 0.9$	The internal consistency of the scale is high,
$0.7 \leq \alpha < 0.9$	The scale has internal consistency,
$0.6 \leq \alpha < 0.7$	The internal consistency of the scale is acceptable,
$0.5 \leq \alpha < 0.6$	The internal consistency of the scale is weak,
$\alpha \leq 0.5$	The scale has no internal consistency.

If researchers calculate the Cronbach's alpha value to be 0.95 or higher, which indicates the internal consistency of the scale, they believe that the measurement tool is very reliable. However, this is not the correct approach. The fact that the reliability of the measurement tool has a very high value indicates that some expressions in the measurement tool are the same as each other and they do not have distinctive features [40]. In other words, it states that there are more expressions in the measurement tool than there should be, and that this behavior or quality could be measured with fewer expressions. When you submit an article for publication in an SSCI/SCI indexed journal, it is likely that the referees will criticize this issue. Likewise, a Cronbach's alpha value between 0.6 and 0.7 may not be considered sufficient for SSCI/SCI indexed journals. Therefore, a Cronbach's alpha value of 0.7 and above indicates that the scale has internal consistency.

RELIABILITY ANALYSIS WITH SPSS

Previously, we conducted analysis to determine the validity of the scales. Now, we will determine the reliability of the scales we are using by conducting a reliability analysis of each scale that we have included in the research. To perform the analysis, we first open the “Data-1.sav” file. You can access this file at www.indataanalysis.com

We follow the steps below:

1. Analyze ---> Scale ----> Reliability Analysis...

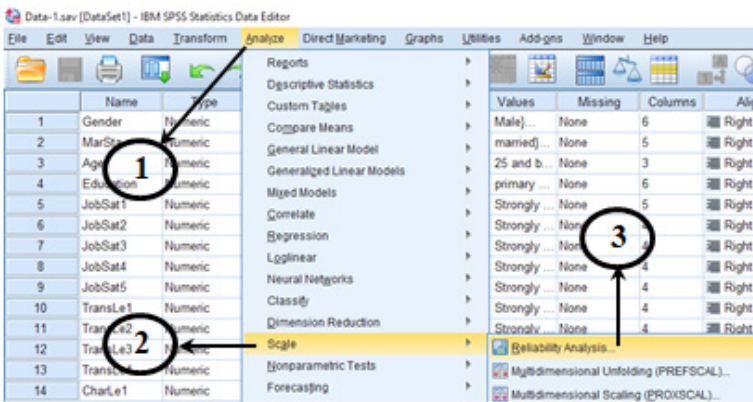


Figure 3.36. SPSS “Analyze” Screen

2. The items of the relevant variable are moved to the opened screen and the “Statistics” button is clicked.

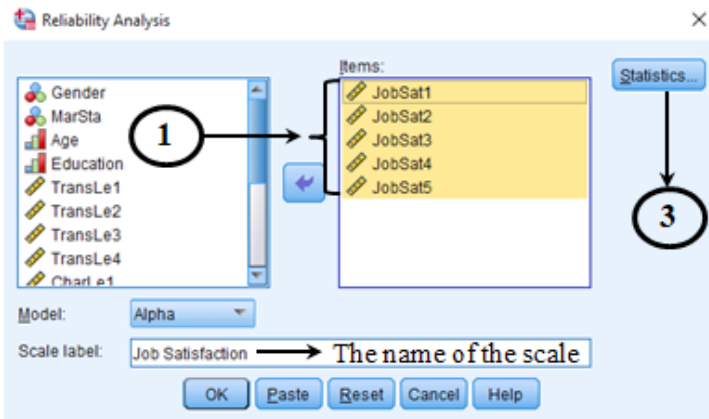


Figure 3.37. “Reliability Analysis” Screen

3. Analyses to be made are selected on the Statistics screen.

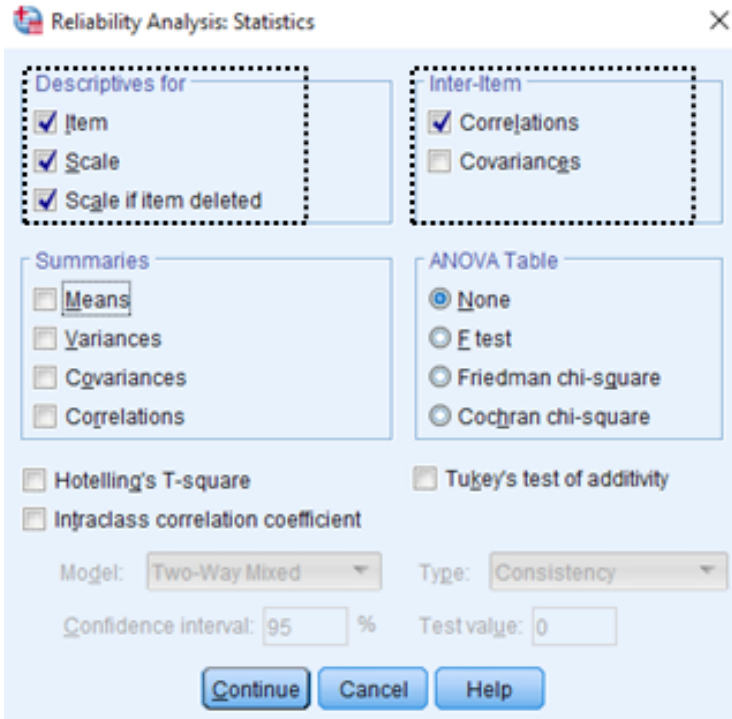


Figure 3.38. “Statistics” Screen

Table 3.14. Case Process Summary Table

Scale: Job Satisfaction

Case Processing Summary

		N	%
Cases	Valid	156	100,0
	Excluded ^a	0	,0
	Total	156	100,0

a. Listwise deletion based on all variables in the procedure.

Table 3.15. Reliability Value Table

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,830	,830	5

It is the standardized Cronbach's alpha estimated on the basis of the correlation. This value is used in research.

Standardized Cronbach's alpha estimated on the basis of covariance

Total number of items in the scale

Table 3.16. Inter-Item Correlation Table

	JobSat1	JobSat2	JobSat3	JobSat4	JobSat5
JobSat1	1,000	,693	,398	,548	,455
JobSat2	,693	1,000	,517	,495	,410
JobSat3	,398	,517	1,000	,489	,456
JobSat4	,548	,495	,489	1,000	,487
JobSat5	,455	,410	,456	,487	1,000

Shows the correlations of the items with each other.

In our example, it is seen that all statements in the scales have a positive correlation with each other.

Table 3.17. Item Total Statistics Table

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
JobSat1	13,60	14,641	,670	,551	,783
JobSat2	13,71	15,394	,681	,552	,782
JobSat3	13,95	15,959	,581	,377	,808
JobSat4	13,75	15,027	,644	,422	,791
JobSat5	13,73	15,682	,565	,333	,814

This table shows the changes in the Cronbach's alpha value when each statement in the scale is deleted. The most important column in the table is the "Cronbach's Alpha if Deleted" column.

As mentioned above, the Cronbach's alpha value is 0.830. It is clear that regardless of what expression we delete, the Cronbach's alpha value will not increase.

INTERPRETATION OF THE FINDINGS

Researchers prefer different methods in their research when calculating the Cronbach's alpha value. All of these approaches are correct. The choice of the examples presented below is at the discretion of the researcher.

In the first approach, the researcher indicates the Cronbach's alpha value on the table. In this method, the researcher determines the Cronbach's alpha value of all the variables included in the research and presents them in a table in the findings section.

Example 1:

Cronbach's alpha values showing the internal consistency of the variables included in the study were estimated and presented in Table 3.18.

Table 3.18. Cronbach's alpha values

	Number of Items	Cronbach's Alpha Values
Transactional Leadership	4	.768
Charismatic Leadership	6	.868
Job Satisfaction	5	.830

Having examined Table 1.18, it is seen that the Cronbach's alpha values of variables are 0.70 and above, which is the acceptable limit. These values indicate that the scales have internal consistency [40].

In the second approach, the researcher states the Cronbach's alpha values of the scales in parentheses in the table where the mean, standard deviation and correlations are shown. The mean and standard deviation values are explained in this analysis, and the correlation analysis will be explained in the following sections.

Example 2:

Mean, Standard Deviation, Correlation and Cronbach's Alpha values of the variables included in the study are presented in Table 3.19.

Table 3.19. Mean, Standard Deviation, Correlation and Cronbach's Alpha Values

Variables	Mean	Sd	1	2	3
1. Job Satisfaction	3.44	.959	(.830)		
2. Transactional Leadership	3.33	.773	.417**	(.768)	
3. Charismatic Leadership	3.33	1.128	.591**	.451**	(.868)

* $p < 0.10$ ** $p < 0.05$ (N=156), Cronbach's alpha values showing the internal consistency of the scale are presented in parentheses.

An examination of Table 3.19 reveals that the variables have a positive correlation with each other and the Cronbach's alpha values, which show the internal consistency of the variables, are 0.768 and above.

Example 3:

In the third approach, the researcher states the Cronbach's alpha value in the "scales" section, which is the subsection of the "Method" section of the article.

Scales:

Transactional Leader: Extant literature on the transactional leadership perceptions of employees was examined and the scale developed by Avolio and Bass [38] was used. All statements of the scale were graded between 1 ("strongly disagree") and 5 ("strongly agree") in a 5-point Likert type. Sample questions of the scale include "He/she sets an example for the employees with his/her behaviors.", "He/she deals with the personal and career development of the employees and guides them." and "Encourages employees to devise creative ideas." The Cronbach's alpha value, which indicates the internal consistency of the scale, was measured as 0.868.

Based on the example in the data set, the same format is applied for the "Charismatic Leader" and "Job Satisfaction" scales. In the last sentence, the Cronbach's alpha values of the scale are stated as in transactional leadership above.

NOTE: Researchers often use a combination of approaches 2 and 3.

INCREASING RELIABILITY

The main reason for the reliability coefficient of the scales used in the research being low is generally that the researcher used a scale in his/her research whose validity and reliability had not previously been tested. In order to avoid this problem, scales whose validity and reliability have been tested and used in previous research should be preferred in planned studies.

The literature emphasizes that the scales developed and applied vary according to the population and culture in which the research is conducted [40]. For this reason, it is important that the scales developed in different cultures are adapted to the local language and culture of the country where the research is conducted, and their validity and reliability are tested [40]. Another mistake frequently made by researchers is that a scale developed in another language is translated into the local language and used in the research. This issue usually produces problems such as low scale reliability and the inability to reach healthy results. However, the population in which the research is conducted is also important. It would be beneficial for the researchers to use previously employed scales whose validity and reliability have been tested for the population in which they intend to conduct the research. Generally, a scale developed for the manufacturing sector may not yield the same results when applied to the service sector.

Various different methods can be applied when it is determined that the measurement tool does not have sufficient reliability:

1. If the Cronbach's alpha value is low, the best method is to increase the number of samples. Increasing the sample size will probably increase the Cronbach's alpha value, which indicates the internal consistency of the scale. However, it is not true that increasing the number of samples will excessively increase the Cronbach's alpha value.

2. One of the methods used to increase the reliability of the measurement tool in scale development studies is to increase the number of items in the measurement tool. As the number of statements increases, the Cronbach's alpha value of the scale will also increase. This is related to the formula used to calculate the Cronbach's alpha value. Therefore, researchers now present the Cronbach's Alpha value as well as the Composite reliability value in their research.

3. If the Cronbach's alpha value is very low (0.40 and below), there are two possible reasons. Firstly, it could mean that the majority of the participants in the sample group answered without reading the statements in the measurement

tool or gave random answers. In this case, the researcher should examine each survey exclude those that have been completed in a certain systematic way. In cases where the researcher does not personally administer the questionnaires to the participants (such as when a questionnaire is administered by the human resources manager or interviewers, or questionnaires are administered in electronic environments, etc.), he/she will frequently encounter this problem.

The second possibility is that the scale used by the researcher is a formative scale. Formative scales may have negative or zero correlations between statements. Scale statements in formative measurement tools determine the latent variable and show the causes of the latent variable, not its effects [39]. In this case, a very low Cronbach’s alpha value indicates that the measurement tool actually works very well. As a result, researchers should know whether their measurement tool in their research is ‘formative’ or ‘reflective’ and they should use a scale appropriate to the research model and theoretical background.

4. One of the main reasons why the Cronbach’s alpha value is so low is that the researchers do not have sufficient information about the measurement tool. There may be reverse coded items in the scale. The researcher’s ignorance of this will cause a low Cronbach’s alpha value. In this case, it is useful to examine the correlations between the items of the scale in the research. It would be more useful to explain this with an example: First, let’s open the “Data-2.sav” file (You can find this file at www.indataanalysis.com) and follow the steps below to perform a reliability analysis.

1. Analyze ---> Scale ----> Reliability Analysis...

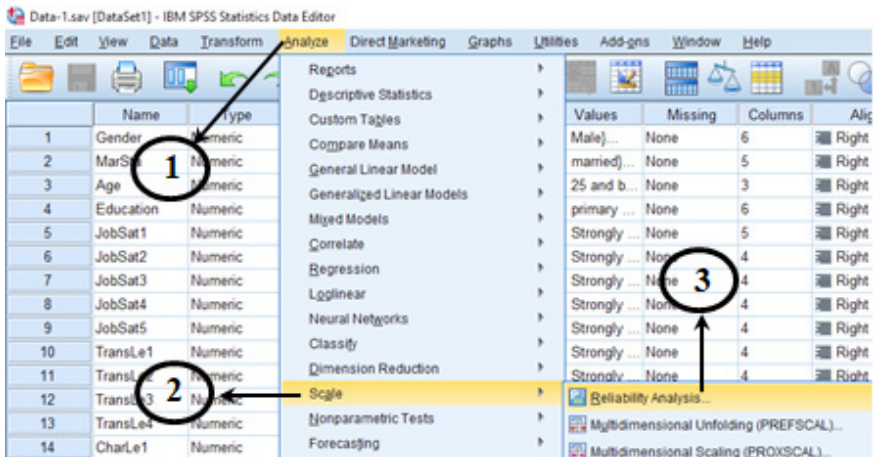


Figure 3.38. SPSS “Analyze” Screen

2. All items of the relevant variable are moved to the “Items” section of the screen that opens and the “Statistics” button is clicked.

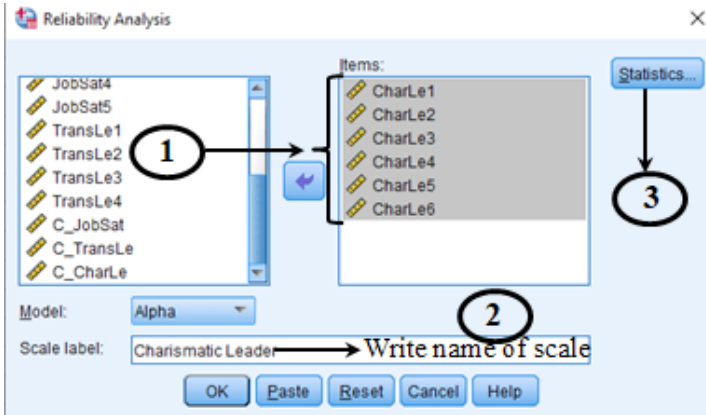


Figure 3.39. “Reliability Analysis” Screen

3. Necessary analyses are selected on the “Statistics” screen and the “OK” button is pressed.

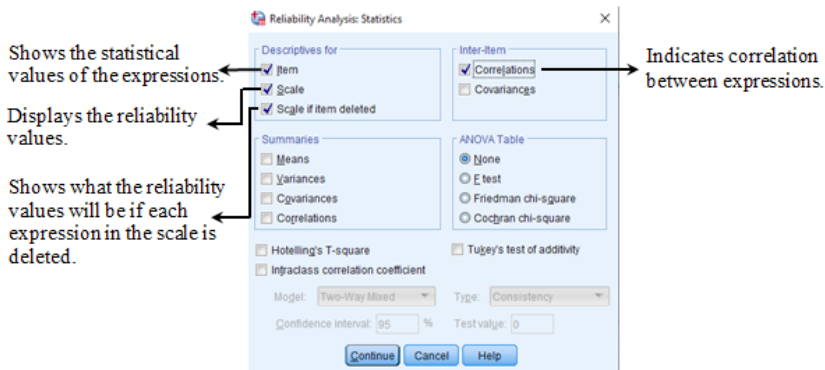


Figure 3.40. “Statistics” Screen

Having examined the tables in the “Output” file of SPSS:

Table 3.20. Reliability Statistics Table

Reliability Statistics		
Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,565	,574	6

The Cronbach’s alpha value of the Charismatic Leadership scale was measured as 0.565.

Table 3.21. Item Total Correlation Table

Item-Total Statistics					
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
CharLe1	16,0128	10,968	,615	,474	,355
CharLe2	15,9679	12,160	,597	,426	,392
CharLe3	16,0064	11,219	,648	,554	,350
CharLe4	15,7500	12,318	,537	,444	,415
CharLe5	15,9744	11,148	,585	,475	,372
CharLe6	16,5064	24,961	-.648	,435	,849

In the “Item-Total Statistics” table, the “Cronbach’s Alpha if Deleted” column shows the researcher the Cronbach’s alpha value that will be obtained by deleting each statement. When this column is examined, it is seen that the Cronbach alpha value of the scale will increase from 0.565 to 0.849 by removing the expression “CharLe6” from the analysis. However, it would be inappropriate to decide to exclude the item “CharLe6” from the analysis. This is because when we look at the column “Corrected Item-Total Correlation”, it is seen that the correlation value of the expression “CharLe6” is negative ($r = -0.648$). Before deleting this statement, the “Inter-Item Correlation Matrix” table should be checked.

Table 3.22. Inter-Item Correlation Table

Inter-Item Correlation Matrix						
	CharLe1	CharLe2	CharLe3	CharLe4	CharLe5	CharLe6
CharLe1	1,000	,570	,617	,440	,488	-.454
CharLe2	,570	1,000	,515	,457	,519	-.455
CharLe3	,617	,515	1,000	,571	,582	-.564
CharLe4	,440	,457	,571	1,000	,553	-.545
CharLe5	,488	,519	,582	,553	1,000	-.542
CharLe6	-.454	-.455	-.564	-.545	-.542	1,000

The “Inter-Item Correlation Matrix” table shows us the correlations between the expressions in the scale. It is seen that the expression “CharLe6” has a negative (-) correlation with all other expressions. This indicates that the expression “CharLe6” is reverse coded compared to other expressions. It is useful to give an example to facilitate the understanding of this subject:

Let's assume that there are 5 statements below in the scale developed to determine charismatic leadership perception.

1. Takes a high degree of personal risk for the sake of the team. (+)
2. He/she creates a collective sense of duty. (+)
3. Makes significant personal sacrifices to achieve goals. (+)
4. Constantly seeks new opportunities to achieve goals. (+)
5. Excites team members with their ideas. (+)
6. Does not seek different ways of doing things. (-)

While most of these items express a positive situation, item 6 expresses a negative situation. Therefore, it is natural for this expression to be negative in the correlation analysis. This situation shows that the 6th statement is reverse coded. The next step is to reverse-code this expression before conducting the analysis.

Reverse-coding process in SPSS:

We follow the procedures described below:

1. Transform ---> Recode into Different Variables...

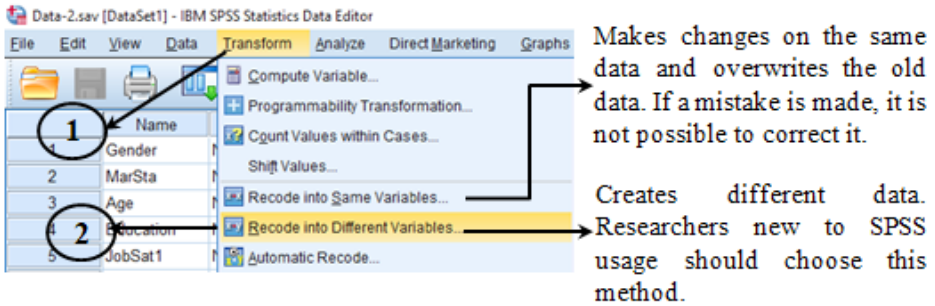


Figure 3.41. “Transform” Screen

Operations on the screen that opens:

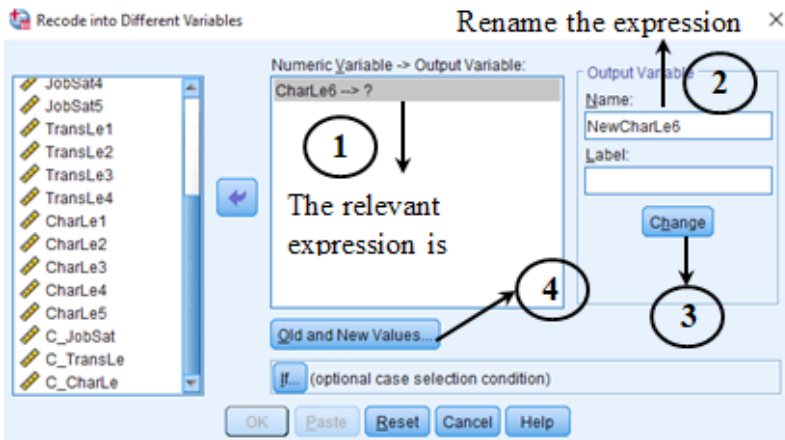


Figure 3.42. “Reverse Code” Screen

Actions to be taken on the screen:

The scale in our research is a 5-point Likert scale. Therefore, the actions to be taken are the assignment of “1 instead of 5”, “2 instead of 4”, “3 instead of 3”, “4 instead of 2”, and “5 instead of 1”. On the screen below, it can be seen that the operations in the first 4 stages have been made. In order to understand the subject, the process to be applied in the last step is explained in Figure 3.43. All steps are the same.

Figure 3.43 shows how to assign “5 instead of 1”.

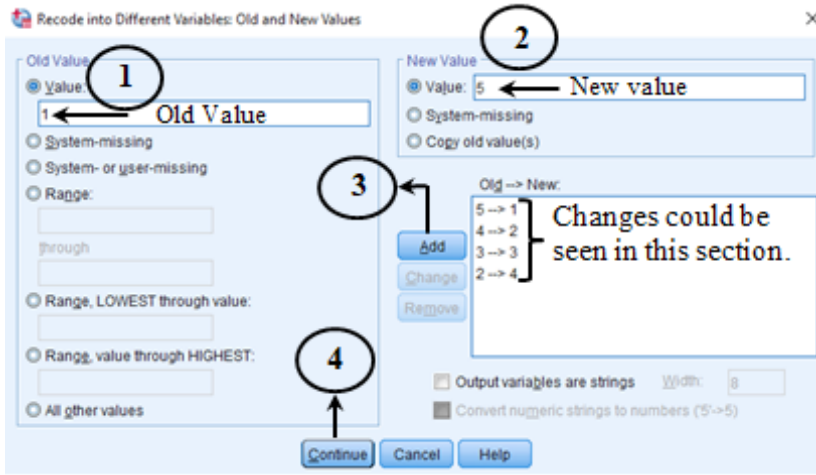


Figure 3.43. “Reverse Coding Operation” Screen

After performing all the operations, the new variable is seen on the 23rd line of the “Variable” screen of SPSS.

19	CharLe6	Numeric	8	2
20	C_JobSat	Numeric	8	0
21	C_TransLe	Numeric	8	0
22	C_CharLe	Numeric	8	0
23	NewCharLe6	Numeric	8	2
24				
25				

The variable “NewCharLe6” has been added to the 23rd row as a new variable.

Figure 3.44. SPSS “Variable View” Screen

In all subsequent operations we will perform, the expression “NewCharLe6” should be used. In other words, the term “CharLe6” will no longer be used at any stage of our analysis, and the term “NewCharLe6” will be used instead. If the researcher is sure of the correctness of his/her actions, it would be appropriate to delete the phrase “CharLe6” in order not to confuse it in future analyses.

After the reverse coding process, the results of the reanalysis are as follows:

Table 3.22. Reliability Statistics Value

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
,868	,869	6

With the reverse coding of the 6th statement (NewCharLe6), the Cronbach’s alpha value of our scale was measured as 0.868. If you remember, in our previous analysis, if this expression was removed, the Cronbach’s alpha value of the scale would have been measured as 0.849. Analyses should be interpreted well before deciding to extract statements from the scale. “Inter-Item Correlation Matrix” and “Item-Total Statistics” tables should be rechecked.

Table 3.23. Item Total Statistics Table

Item-Total Statistics

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
CharLe1	16,0128	10,968	,615	,474	,355
CharLe2	15,9679	12,160	,597	,426	,392
CharLe3	16,0064	11,219	,648	,554	,350
CharLe4	15,7500	12,318	,537	,444	,415
CharLe5	15,9744	11,148	,585	,475	,372
CharLe6	16,5064	24,961	-.648	,435	,849

When the “Cronbach’s Alpha if Deleted” column is examined in the “Item-Total Statistics” table, it is seen that the Cronbach’s alpha value will not exceed 0.868, irrespective of which expression is removed. In addition, when the “Corrected Item-Total Correlation” column is checked, it is seen that all expressions are positively (+) correlated. This indicates that our reliability analysis was completed correctly.

A mistake that researchers frequently encounter and make is that they directly omit some expressions in order to increase the Cronbach’s alpha value by considering the “Cronbach’s Alpha if Deleted” column in the “Item-Total Statistics” table. Omitting a statement is an important decision. Before doing this, factor analysis should be performed and the reason for the omission of the statement should be explained theoretically. It is not a correct approach to delete expressions, particularly in scales with few expressions (3-4 expressions).

References

- [1] Anastasi, A., & Urbina, S. (1997). *Psychological testing*. Prentice Hall/Pearson Education.
- [2] Whiston, S. C. (2012). *Principles and applications of assessment in counseling*. Cengage Learning.
- [3] Bollen, K. A. (1989). *The consequences of measurement error. Structural Equations with Latent Variables*, 151-178.
- [4] Ayre, C., & Scally, A. J. (2014). Critical values for Lawshe's content validity ratio: Revisiting the original methods of calculation. *Measurement and Evaluation in Counseling and Development*, 47(1), 79-86.
- [5] Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563-575.
- [6] Allen, N. J., & Meyer, J. P. (1990). The measurement and antecedents of affective, continuance and normative commitment to the organization. *Journal of Occupational Psychology*, 63(1), 1-18.
- [7] Liden, R.C., & Maslyn, J. M. (1998). Multidimensionality of leader-member exchange: An empirical assessment through scale development. *Journal of Management*, 24(1), 43-72.
- [8] Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, 66(3), 393-416.
- [9] Swisher, L. L., Beckstead, J. W., & Bebeau, M. J. (2004). Factor analysis as a tool for survey analysis using a professional role orientation inventory as an example. *Physical Therapy*, 84(9), 784-799.
- [10] Bartholomew, D. J., Knott, M., & Moustaki, I. (2011). *Latent variable models and factor analysis: A unified approach* (Vol. 904). John Wiley & Sons.
- [11] Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5, pp. 481-498). Boston, MA: Pearson.
- [12] Field, A. (2009). *Discovering statistics using SPSS: Introducing statistical Method* (3rd ed.). Thousand Oaks, CA: Sage Publications.
- [13] Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (1998). *Multivariate data analysis* (Vol. 5, No. 3, pp. 207-219). Upper Saddle River, NJ: Prentice hall.
- [14] Guadagnoli E., & Velicer W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin*, 103(2), 265-75.

- [15] Sapnas, K. G., & Zeller, R. A. (2002). Minimizing sample size when using exploratory factor analysis for measurement. *Journal of Nursing Measurement, 10*(2), 135-154.
- [16] Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Sage.
- [17] Gorsuch, R. L. (1983). *Factor Analysis*. Hillsdale, NJ: Erlbaum.
- [18] Everitt, B. S. (1975). Multivariate analysis: The need for data, and other problems. *The British Journal of Psychiatry, 126*(3), 237-240.
- [19] Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford, K. R. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communalities, and overdetermination. *Educational and Psychological Measurement, 65*(2), 202-226.
- [20] MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84.
- [21] Yong, A. G., & Pearce, S. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *Tutorials in quantitative methods for psychology, 9*(2), 79-94.
- [22] Costello, A. B., & Osborne, J. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research, and Evaluation, 10*(1), 7.
- [23] Rummel, R. J. (1970). *Applied factor analysis*, Evanston, IL: Northwestern Univer.
- [24] Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement, 20*(1), 141-151.
- [25] Kline, P. (2014). *An easy guide to factor analysis*. Routledge.
- [26] Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences*. Routledge.
- [27] Meydan, C. H., & Şeşen, H. (2011). *Yapısal eşitlik modellemesi AMOS uygulamaları*. Detay Yayıncılık.
- [28] Büyüköztürk, Ş. (2004). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem A Yayıncılık.
- [29] Byrne, B. M. (1998). Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming.
- [30] Sümer, N. (2000). Yapısal eşitlik modelleri: Temel kavramlar ve örnek uygulamalar. *Türk Psikoloji Yazıları, 3*(6), 49-74.

[31] Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics*. Boston: Allyn and Bacon.

[32] Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling*. Psychology Press.

[33] Şimşek, Ö. F. (2007). *Yapısal eşitlik modellemesine giriş:(Temel ilkeler ve LISREL uygulamaları)*. Ekinoks.

[34] Raykov, T., & Marcoulides, G. A. (2006). *A First Course in Structural Equation Modeling*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

[35] Ullman, J. B. (2001). *Structural equation modeling*. (Edi.: B. G. Tabachnick ve L. S. Fidell). Using Multivariate Statistics. Needham Heights, MA: Allyn & Bacon.

[36] Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.

[37] Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.

[38] Avolio, B. J., & Bass, B. M. (1995). Individual consideration viewed at multiple levels of analysis: A multi-level framework for examining the diffusion of transformational leadership. *The Leadership Quarterly*, 6(2), 199-218.

[39] Aksay, B., & Ünal, A. Y. (2016). Yapısal eşitlik modellemesi kapsamında formatif ve reflektif ölçüm. *Çağ Üniversitesi Sosyal Bilimler Dergisi*, 13(2), 1-21.

[40] Sürücü, L., & Maslakçı, A. (2020). Validity and reliability in quantitative research. *Business & Management Studies: An International Journal*, 8(3), 2694-2726.

CHAPTER 4

CORRELATION AND REGRESSION ANALYSES

While correlation analysis is performed to determine whether two variables change together and to determine the strength of the relationship between the variables, the purpose of regression analysis is to express the effect of the independent variable on the dependent variable as an equation [1].

For example, correlation analysis could be used to test whether an employee with high job satisfaction also has high performance. In regression analysis, the effect of job satisfaction on performance is determined. As a result of the regression analysis, the extent to which an increase in job satisfaction corresponds to an increase in performance can be estimated. In other words, regression analysis estimates the value of the dependent variable based on the current value of the independent variable, assuming that there is an average mathematical relationship between two or more variables [1].

Correlation and regression analyses have three main purposes:

- To determine the direction/strength of the relationship between two variables without inferring a cause and effect relationship.
- To test hypotheses about the cause and effect relationship.
- To determine the effect of a one-unit increase in the independent variable on the dependent variable.

Table 4.1. Differences in correlation and regression analyses

	Correlation Analysis	Regression Analysis
Content	Correlation analysis is a statistical measure that determines the direction and strength of the linear relationship between two variables.	The statistical relationship between an independent variable and a dependent variable is explained by regression analysis.
Purpose	Determines a statistical value about the direction and strength of the relationship between the variables.	Expresses the effect of the independent variable on the dependent variable in the form of an equation.
Usage	Shows the linear relationship between two variables.	Estimates the independent variable on the basis of the dependent variable.
Dependent and Independent Variables	No difference between variables	Both variables are different.
Meaning of Achieved Value	The correlation coefficient shows the extent to which the two variables move together.	The regression coefficient shows the effect of a unit change in the independent variable (x) on the dependent variable (y).
Adapted from the study of Sürücü and Maslakçı [1].		

Correlation and regression analysis are interrelated in the sense that they both deal with the relationships between variables. However, correlation and regression analyses should not be interpreted as establishing a cause-effect relationship. These analyses show how and the extent to which the variables are related to each other. Any inference about the cause-effect relationship should be based on the researcher's decision according to the theoretical framework. If you say "two variables are related to each other" you should perform correlation analysis, whereas if you say "one affects the other", you should perform regression analysis.

CORRELATION ANALYSIS

Correlation analysis is performed to determine the direction and strength of the linear relationship between two variables [1]. The correlation coefficient is denoted by “ r ” and has a value ranging from -1 to $+1$. As the correlation value approaches $+1$, it shows a positive linear relationship. In other words, it means that as the value of one variable increases, the value of the other variable will also increase. As the correlation value approaches -1 , it means that there is a negative linear relationship; in other words, as the value of one variable increases, the value of the other variable decreases. However, a correlation value of “ 0 ” indicates that there is no linear relationship between the variables. In short, correlation is a similar quantity trend and shows how much a certain increase in one variable will increase the other variable [1].

The sign of the correlation coefficient indicates the direction of the relationship (negative or positive), and the obtained value indicates the strength of the relationship.

For example, a correlation of $r=0.8$ indicates a high and positive relationship between two variables, while a correlation of $r = -0.3$ indicates a very weak and negative relationship between two variables. If $r = 0$, this indicates that there is no linear relationship between the two variables.

The table below can be used to evaluate the correlation coefficient.

Table 4.2. Evaluation of the correlation coefficient

Strength of Correlation Coefficient				
Very weak	Weak	Moderate	Strong	Very Strong
$0 < r \leq 0.3$	$0.3 < r \leq 0.5$	$0.5 < r \leq 0.7$	$0.7 < r \leq 0.9$	$0.9 < r \leq 1$

The figures below show three hypothetical scenarios where the independent variable is plotted along the horizontal axis (X-axis) and the dependent variable along the vertical axis (Y-axis).

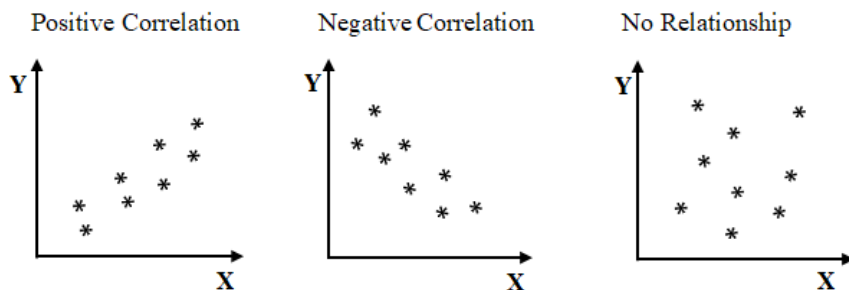


Figure 4.1. Direction of Correlation

Calculation of the correlation coefficient is performed with the following formula:

$$r = \frac{Cov(x, y)}{\sqrt{s_x^2 \times s_y^2}}$$

In the formula:

$Cov(x, y)$ = the covariance of x and y.

s_y^2 ve s_x^2 = the variances of x and y

Pearson product-moment and Spearman rank-order Correlation analyses are performed to determine the correlation value. The suitable analysis method is determined by the distribution of the variables. Pearson correlation analysis should be preferred if the variables are obtained with a ratio or interval scale and conform to a normal distribution, and if it does not comply with normal distribution (it is sufficient that just one variable does not show normal distribution), Spearman correlation analysis should be preferred [1]. If the variables are obtained with an ordinal scale, then Spearman correlation analysis can be performed.

CORRELATION ANALYSIS WITH SPSS

Research question: *'Is there a relationship between leadership styles and job satisfaction?'*

If you remember, transactional and charismatic leadership styles are identified as leadership styles in our research. Therefore, to answer the research question, we first need to determine whether there is a relationship between these two leadership styles and job satisfaction, and if a relationship exists, the direction and strength of this relationship. Correlation analysis will be performed.

As mentioned in a previous chapter, we conducted analysis to determine the normal distribution of the data. Since it was confirmed that the data have a normal distribution, it is not necessary to perform an analysis to determine the distribution of the data again. Since the data have a normal distribution, the Pearson Product Moments Correlation Analysis will be performed. First of all, we open the "Data-1.sav" file (You can access this file from the address www.indataanalysis.com).

We follow the steps listed below:

1. Analyze ---> Correlate ----> Bivariate options are clicked.

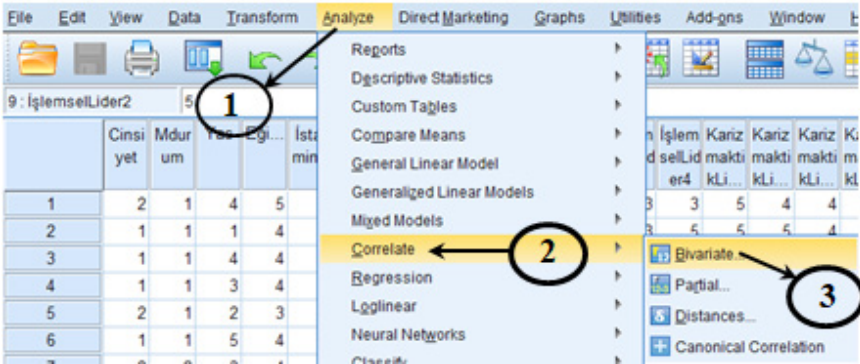


Figure 4.2. SPSS "Analyze" Ekranı

2. The items of the relevant variable are moved to the opened screen and the “Options” button is clicked.

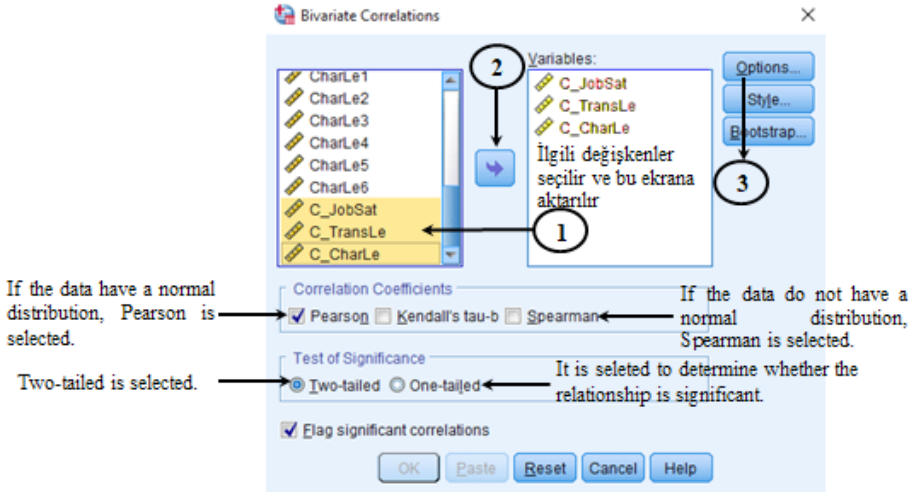


Figure 4.3. “Bivariate Correlations” Screen

3. “Exclude cases pairwise” and “Means and standard deviations” are selected from the “Options” menu.

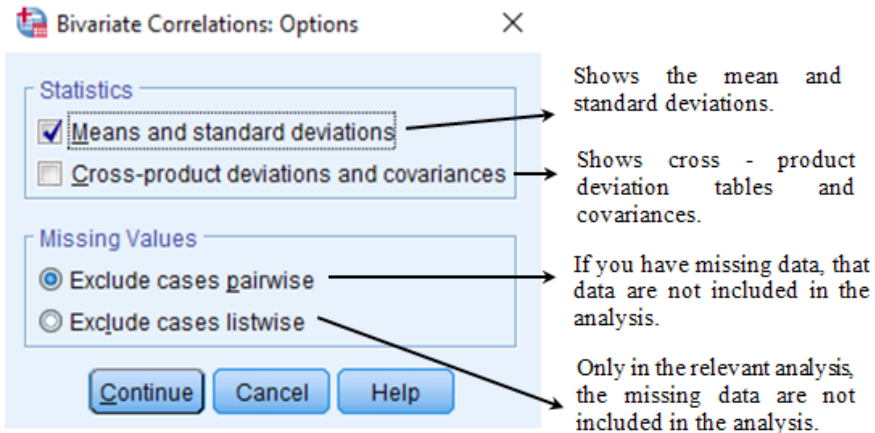


Figure 4.4. “Options” Screen

Having examined the below tables in the “Output” file of SPSS in detail:

Table 4.3. SPSS Output

Descriptive Statistics

	Mean	Std. Deviation	N
C_JobSat	3,44	,959	156
C_TransLe	3,83	,773	156
C_CharLe	3,33	1,128	156

Means

Standard
Deviations

Sample
Size

Table 4.4. Correlation Table

Correlations

		C_JobSat	C_TransLe	C_CharLe
C_JobSat	Pearson Correlation	1	,417**	,591**
	Sig. (2-tailed)		,000	,000
	N	156	156	156
C_TransLe	Pearson Correlation	,417**	1	,451**
	Sig. (2-tailed)	,000		,000
	N	156	156	156
C_CharLe	Pearson Correlation	,591**	,451**	1
	Sig. (2-tailed)	,000	,000	
	N	156	156	156

** . Correlation is significant at the .01 level (2-tailed).

Correlation coefficient
between transactional
leadership and job
satisfaction

Correlation coefficient
between charismatic
leadership and job
satisfaction

Shows that there is a relationship between
two variables at a significance level of 0.01.

INTERPRETATION OF RESEARCH FINDINGS

Although researchers prefer different methods when specifying the correlation values in their studies, the most preferred approach is to show the mean, standard deviation and correlations between variables in a single table.

Example:

The mean, standard deviation, and correlation values of the variables included in the study are presented in Table 4.5.

Table 4.5. Mean, Standard Deviation and Correlation

Variables	Mean	Sd	1	2	3
1. Job satisfaction	3.44	.959	1		
2. Transactional Leadership	3.83	.773	.417**	1	
3. Charismatic Leadership	3.33	1.128	.591**	.451**	1

** $p < 0.05$ (N=156)

As a result of the correlation analysis, it was determined that there is a weak and positive relationship between transactional leadership and job satisfaction at the 0.05 significance level ($r = 0.417$, $p < 0.05$), and a moderate positive relationship between charismatic leadership and job satisfaction ($r = 0.591$, $p < 0.05$).

Note: The mean and standard deviation values in the table can be determined by selecting “Mean and Standard Deviations” from the “Options” menu while performing the correlation analysis (described in the 3rd item above).

CORRELATION ANALYSIS WITH AMOS

Before proceeding to the next sections, where more complex models will be examined with structural equation models using AMOS, we would like to state that correlation analysis can also be performed with AMOS, although it does not fully comply with the capabilities of the program. At this stage, we will test the correlation analysis with AMOS, which was described in the previous parts of the book and tested with SPSS on Data-1.sav. The AMOS model plot of the correlation data previously tested with SPSS is presented in Figure 4.5. As can be seen, three observed variables covariate with each other (a bidirectional correlation relationship exists between them). The stages of drawing the model will not be repeated here. However, the model can be drawn using the shortcuts described above.

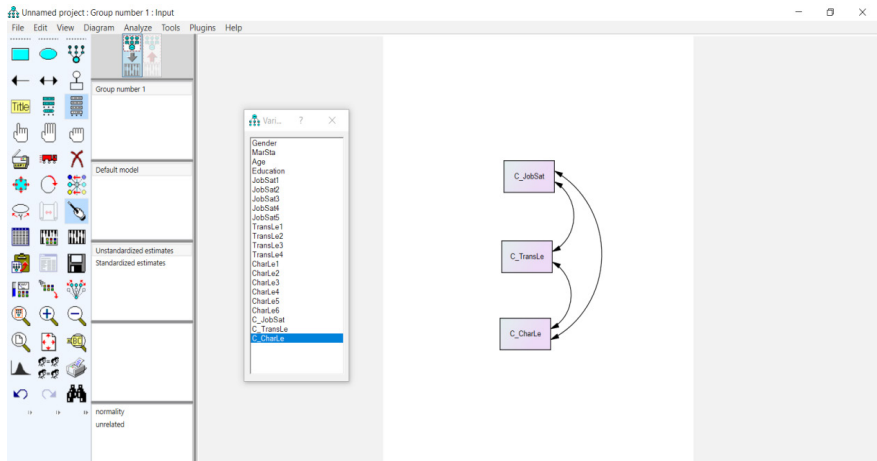


Figure 4.5. Correlational Model

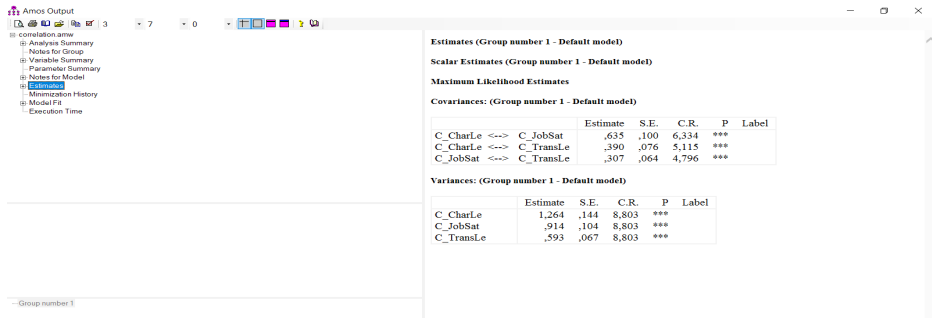


Figure 4.6. Test Statistics of the Correlational Model

The correlation test statistics resulting from the analysis of the drawn model are shown in Figure 4.6. Accordingly, there is a positive and significant relationship between job satisfaction and transactional leadership ($r = .417$), job satisfaction and charismatic leadership ($r = .591$), and transactional leadership and charismatic leadership ($r = .451$). Values in parentheses here are standardized Pearson correlation coefficients. If this correlation analysis is repeated with SPSS, the same values will be generated.

REGRESSION ANALYSIS

Regression analysis is used to determine the strength and direction of the relationship between two or more variables [1]. Regression analysis is performed using the least-squares method to estimate the value of the dependent variable based on the known values of the independent variable or variables [1]. The least-squares method is the most common procedure used in regression analysis to develop estimates of model parameters.

The variable on which you will make predictions is expressed as the “dependent variable” and is denoted by “y”. The variable whose value is given to estimate the value of the dependent variable (y) is expressed as the “independent variable” and is denoted by “x”.

When regression analysis is used to measure the strength of the relationship between a dependent (y) and an independent (x) variable, it is called “simple regression analysis” and is calculated with the following formula:

$$y = \beta_0 + \beta_1 x_1 + e$$

y = Dependent variable

x_1 = Independent variable

β_0 = The y-intercept for the population.

β_1 = The regression coefficients of the independent (x_1) and dependent variables (y).

e = Residual (error), a probabilistic value that cannot be explained by the linear relationship between y and x .

The following three assumptions are required for regression analysis:

- i) Normality of error
- ii) Homoscedasticity
- iii) Independence of errors

i) Normality of error: This assumption requires that errors around the regression line are normally distributed for each value of X (independent variables). Inferences about the regression line and the regression coefficients will not be seriously affected provided the errors around the regression line for each value of the independent variables are not very different from the normal distribution.

ii) Homoscedasticity: This assumption means that the variation around the regression line should be constant for all values of the independent variables (X). Thus, it indicates that errors when X has a low-value change by the same

amount as when X has a high value. An important criterion when using the least square method is that the homoscedasticity assumption fits the regression line. If there are serious deviations from this assumption, data transformations or applying the weighted least square method is the correct approach.

iii) Independence of errors: This assumption requires that the errors around the regression line are independent for each value of the explanatory variables. The independence of errors is particularly important when collecting longitudinal data over a certain period (spread over time). In this situation, errors in a given period are usually associated with errors in the previous period. The independence of errors is a necessary assumption for regression analysis, especially in analyses that examine the change of organizational behavior and that are examined with data collected in certain periods. For example, if a researcher collects data with the same measurement tool 3 times during a period of 1 year in the 1st, 6th, and 12th months, and then examines the change in organizational behavior with the data, the errors must be independent.

Significance test for the regression coefficient (β_1):

To determine the existence of a significant linear relationship between the dependent variable (y) and the independent variable (x), a hypothesis test is performed by determining zero (null) and alternative hypotheses regarding the determined regression coefficient (β_1), as stated below.

Null hypothesis (H_0): [$\beta_1 = 0$]. It means that there is no linear relationship between the dependent and independent variables.

Alternative hypothesis (H_1): [$\beta_1 \neq 0$]. It means that there is a significant linear relationship between the dependent and independent variables. In general, alternative hypotheses are tested in research. E.g., the hypothesis “Charismatic leadership affects job satisfaction significantly and positively” is an alternative hypothesis.

If the null hypothesis is supported, it can be concluded that there is no linear relationship between the dependent and independent variables. However, if the alternative hypothesis is supported, it should be concluded that there is a significant linear relationship between the dependent and independent variables.

The researcher uses the following formula to decide on the hypothesis result:

$$t_{cal} = \frac{b_1}{S_{b_1}}$$

Decision: If the calculated value of the test statistics (t_{cal}) is lower than the tabulated value (t_{tab}), the null hypothesis is supported. Otherwise, the null hypothesis is rejected and the alternative hypothesis is tested.

B_1 = The regression coefficients of Y over X1

S_{b_1} = The standard error of the regression coefficient (b_1)

Researchers must decide which regression analysis to use based on the model they have developed.

Types of Regression Analysis:

1. Multiple Regression
 - a. Standard Multiple Regression
 - b. Stepwise Multiple Regression
2. Hierarchical Regression

1. Simple Linear Regression: Regression analysis used to describe the relationship between a single dependent variable y and a single independent variable x .

2. Multiple Regression: Regression analysis conducted to determine the effect of more than one independent variable on a dependent variable.

a. Standard Multiple Regression: Multiple independent variables affect the dependent variable, and it is a type of regression analysis in which these independent variables are put into the model at the same time.

b. Stepwise Multiple Regression: This is the inclusion of independent variables in the model with a certain level of significance or a statistical approach. Three types of Stepwise multiple regression analysis could be performed:

1) In forward selection analysis, independent variables are included in the model one-by-one in terms of their effects and their effects are examined. As a result of this examination, the relevant independent variable is either included or eliminated from the model.

2) In backward elimination analysis, all independent variables are included in the model at the same time. Variables that do not contribute significantly are excluded from the model, respectively.

3) In bidirectional elimination (stepwise selection) analysis, the independent variables are included in the model one-by-one according to the preference of the researcher who performed the analysis. The researcher makes the decision as to whether the variable should stay in the model or not according to its effect.

In stepwise multiple regression analysis, the variables to be added or subtracted are selected according to the test statistics of the estimated coefficients. While the technique has many benefits, such as its ability to solve the multicollinearity problem, it requires a certain level of expertise. Therefore, it should be preferred by researchers who have a strong familiarity with statistical tests.

3. Hierarchical Regression: The researcher determines the order in which the independent variables enter the regression equation. The inclusion of independent variables in the model is at the discretion of the researcher. The researcher uses this initiative in line with the literature review on the subject. Independent variables are included in the model first, in order of importance. It would be helpful to give an example to facilitate the understanding of the subject.

In the next study, the hypothesis of “The effect of leadership styles on job satisfaction” will be tested. In the literature review on the subject, there are findings that the leadership perceptions of the employees vary according to demographic variables [2] as well as that job satisfaction is affected by demographic variables such as age, marital status and educational status [3]. For this reason, demographic variables (age, marital status and educational status) should be controlled in order to obtain more accurate results for the determination of the effect of the leadership styles specified in the research on job satisfaction. To do this, hierarchical regression analysis should be performed.

REGRESSION ANALYSIS WITH SPSS

Before proceeding to the SPSS application, it is useful to show the research model. The research model is presented in Figure 4.7.

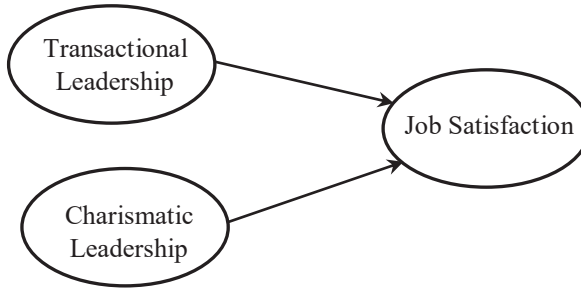


Figure 4.7. Research Model

Research hypotheses:

Hypothesis 0 (H_0): There is no linear relationship between leadership styles (transactional and charismatic leadership) and job satisfaction.

Hypothesis 1 (H_1): Transactional leadership significantly and positively affects job satisfaction.

Hypothesis 2 (H_2): Charismatic leadership has a statistically significant and positive effect on job satisfaction.

Regression analysis will be performed to test these two hypotheses that have been determined within the scope of the research. With this regression analysis, we can determine the effect of transactional and charismatic leadership on job satisfaction.

For regression analysis, we first open the “Data-1.sav” file (You can access this file from www.indataanalysis.com).

We follow the process steps described below:

1. Analyze ---> Regression ----> Linear... option is clicked.

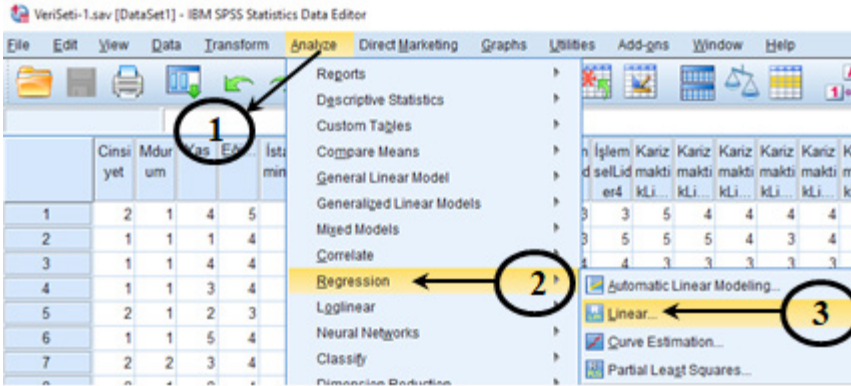


Figure 4.8. SPSS “Analyze” Screen

2. On the screen that opens, “Job satisfaction” is moved to the “Dependent” section, the demographic variables to be controlled are moved to the “Independent” section and the “Next” button is then clicked.

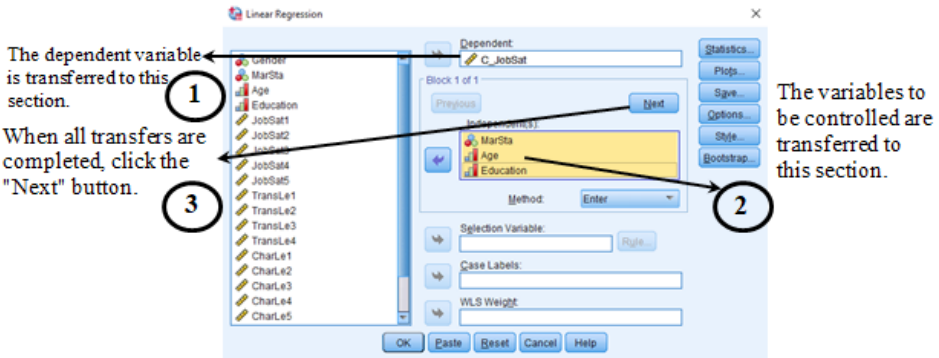


Figure 4.9. “Linear Regression” Screen

3. After clicking the “Next” button, the independent variables are moved to the “Independent” screen. As you can see here, “Block 2 of 2” appears on the screen (just above the Previous button). The next step is step 2 of the hierarchical regression analysis.

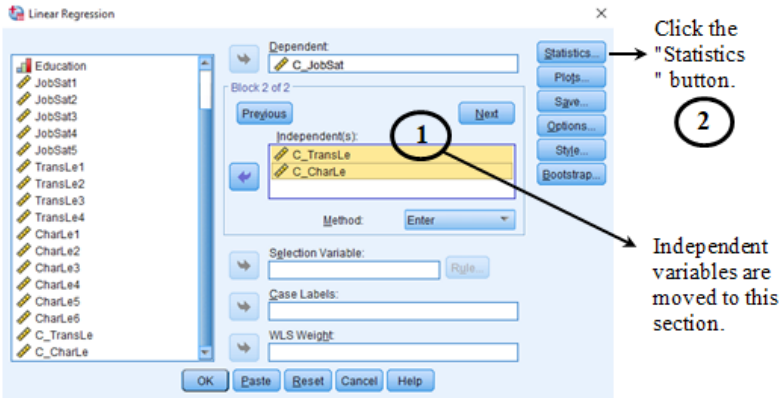


Figure 4.10. “Linear Regression” Screen

4. When the “Statistics” tab is clicked, the tabs shown on the screen that opens are selected.

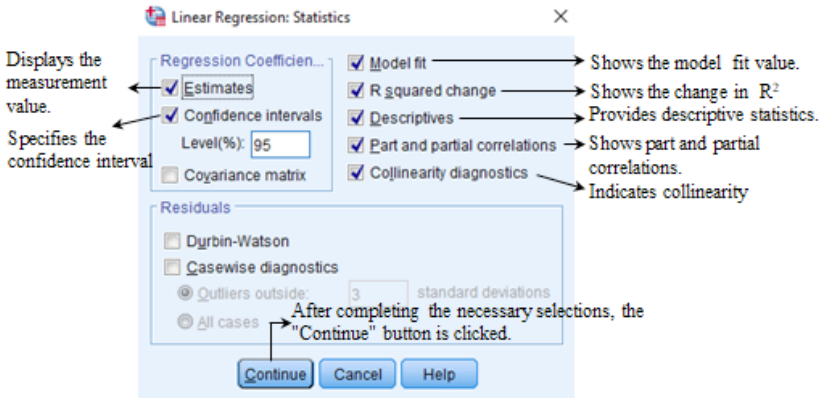


Figure 4.11. “Statistics” Screen

Multicollinearity: The problem is that the dependent variables included in the research act together because the correlation coefficient among themselves is high and significant. To understand whether such a problem is encountered in the research, the “Part and partial correlations” and “Collinearity Diagnostics” tabs should be selected from the “Statistics” menu.

After all the processes are completed, the “OK” button is clicked and the outputs of the SPSS are interpreted.

5. The “Output” file of SPSS is presented below.

Before interpreting SPSS tables, it should be noted that many tables have two models in the “Model” column. As previously mentioned, when performing hierarchical regression analysis, we first included demographic variables (age, marital status, and educational status) into the model, and then the independent variables (transactional and charismatic leadership). In other words, we created two different models. Therefore, there are two models and values for these two models in each table.

Table 4.6. Descriptive Statistics

Descriptive Statistics

	Mean	Std. Deviation	N
C_JobSat	3,44	,959	156
MarSta	1,71	,457	156
Age	2,05	1,168	156
Education	3,74	,708	156
C_TransLe	3,83	,773	156
C_CharLe	3,33	1,128	156

→ Sample Size

↓ Mean
↓ Standart Deviation

Table 4.7. Correlation Table

Correlations

		C_JobSat	MarSta	Age	Education	C_TransLe	C_CharLe
Pearson Correlation	C_JobSat	1,000	,166	-,169	,041	,417	,591
	MarSta	,166	1,000	-,611	-,056	,232	,246
	Age	-,169	-,611	1,000	,039	-,341	-,368
	Education	,041	-,056	,039	1,000	-,015	-,006
	C_TransLe	,417	,232	-,341	-,015	1,000	,451
	C_CharLe	,591	,246	-,368	-,006	,451	1,000
Sig. (1-tailed)	C_JobSat		,019	,018	,307	,000	,000
	MarSta	,019		,000	,245	,002	,001
	Age	,018	,000		,313	,000	,000
	Education	,307	,245	,313		,425	,468
	C_TransLe	,000	,002	,000	,425		,000
	C_CharLe	,000	,001	,000	,468	,000	
N	C_JobSat	156	156	156	156	156	156
	MarSta	156	156	156	156	156	156
	Age	156	156	156	156	156	156
	Education	156	156	156	156	156	156
	C_TransLe	156	156	156	156	156	156
	C_CharLe	156	156	156	156	156	156

The “*Correlations*” table shows the correlations between the variables. The values in triangle number 1 are important. These values show the correlations of the independent variables with each other. If this value is greater than 0.800, it indicates that the multicollinearity problem has occurred. In other words, independent variables with a correlation of 0.800 and above will act together as a single variable. In this case, it would be problematic to measure the effects of the relevant independent variables separately.

One more value can be used to detect the multicollinearity problem. In the “Coefficients” table below (Table 4.11), the “VIF” column should be examined. In the literature, different values have been proposed for the VIF value. Hair et al [4] stated that the VIF value should be 4 or below, whereas Allison [5] stated that if the VIF value is 2.5 or below, the multicollinearity problem will not occur.

The values in triangle number 2 show the significance of the relationship between the variables. If this value is above 0.05, the relationship is not significant. For example, when the education line is examined, it can be seen that the coefficients of significance are 0.307 (job satisfaction), 0.245 (marital status), and 0.313 (age). These values show that education is not associated with job satisfaction (0.307), marital status (0.245), or age (0.313).

Table 4.8. Table of Variables

Variables Entered/Removed ^a			
Model	Variables Entered	Variables Removed	Method
1	Education, Age, MarSta ^b	.	Enter
2	C_TransLe, C_CharLe ^b	.	Enter

a. Dependent Variable: C_JobSat
 b. All requested variables entered.

Shows the independent variables included in the analysis in the first stage (Education, Age, and Marital Status).
 Shows the independent variables included in the second stage of analysis (Transactional and Charismatic leadership).
 Shows the dependent variable.

Table 4.8 reveals that firstly, the effect of demographic variables such as education, age, and marital status on job satisfaction was analyzed (Model 1), while in the second stage (Model 2), the effect of transactional and charismatic leadership on job satisfaction was examined.

Table 4.9. Model Summary Table

Model Summary									
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	,193 ^a	,037	,018	,950	,037	1,969	3	152	,121
2	,626 ^b	,392	,371	,760	,354	43,660	2	150	,000

a. Predictors: (Constant), Education, Age, MarSta

b. Predictors: (Constant), Education, Age, MarSta, C_TransLe, C_CharLe

This is the table where the values for the results of the model are given (Table 4.9). In the “Model Summary” table, the regression coefficient is presented as R (remember that the correlation coefficient is represented by “r”). The R-Squared value in the second column also shows the variance explained as a percentage. It should be noted here that if simple linear regression had been applied, in other words, if a dependent and independent variable were included in the model, our interpretations would have been shaped according to the R-Squared value. Since we have two independent variables in our analysis, it is more appropriate to interpret the Adjusted R-Squared value. In the light of this information, it is seen that the variables in Model 1 (education, age, and marital status) do not have a statistically meaningful effect on job satisfaction (this can be understood from the sig value of 0.121 in the sig. F change column). In Model 2, transactional and charismatic leadership were included in the analysis along with the demographic variables (education, age, and marital status). It shows that at least one these variables included affects job satisfaction. This interpretation can be made because the sig value is less than 0.05 (it can be seen in the table that the Sig value of Model 2 is 0.000). After determining that at least one variable is effective on the dependent variable, the data in the “Coefficients” table below should be examined to determine which independent variable or variables are effective. The Adjusted R-Squared value of Model 2 is 0.371, which indicates that the variables in Model 2 (education, age and marital status, transactional leader, and charismatic leadership) can explain 37.1% of job satisfaction.

Table 4.10. ANOVA Summary Table

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	5,332	3	1,777	1,969	,121 ^b
	Residual	137,213	152	,903		
	Total	142,544	155			
2	Regression	55,818	5	11,164	19,308	,000 ^c
	Residual	86,727	150	,578		
	Total	142,544	155			

a. Dependent Variable: C_JobSat

b. Predictors: (Constant), Education, Age, MarSta

c. Predictors: (Constant), Education, Age, MarSta, C_TransLe, C_CharLe

Table 4.10 is the ANOVA summary table testing the null hypothesis. The null hypothesis is that the model is not significant. In this case, we reject the null hypothesis because the p-value is less than 0.05 (Model 2, sig=0.000). Therefore, our model is significant ($p < 0.05$). However, Model 1 is not significant. In other words, the demographic variables (age, education, marital status) do not affect job satisfaction ($p > 0.05$).

Table 4.11. Coefficients Table

Coefficients ^a														
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	95.0% Confidence Interval for B			Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF	
1	(Constant)	2,991	,642		4,659	,000	1,723	4,259						
	MarSta	,217	,211	,103	1,028	,305	-,200	,634	,166	,083	,082	,625	1,599	
	Age	-,088	,083	-,107	-1,068	,287	-,251	,075	-,169	-,086	-,085	,626	1,597	
	Education	-,069	,108	,051	,636	,526	-,145	,282	,041	,052	,051	,997	1,003	
2	(Constant)	,186	,629		,296	,768	-1,058	1,430						
	MarSta	,162	,169	,077	,960	,339	-,172	,496	-,166	,078	,061	,624	1,601	
	Age	,118	,070	,143	1,686	,094	-,020	,256	-,169	,136	,107	,561	1,782	
	Education	,062	,086	,046	,721	,472	-,108	,233	,041	,059	,046	,997	1,003	
	C_TransLe	,261	,091	,210	2,878	,005	,082	,440	,417	,229	,183	,761	1,314	
	C_CharLe	,451	,063	,530	7,184	,000	,327	,575	,591	,506	,458	,745	1,343	

a. Dependent Variable: C_JobSat

The “Coefficients” table shows the regression coefficients and is the table where we test our hypotheses (Table 4.11). The Unstandardized Coefficients column shows the B constant coefficients. The B value in the constant line is called the intersection (=2.991 for Model 1, 0.186 for Model 2). This means that if there were no demographic variables for Model 1 (education, age, and marital status), job satisfaction among employees would be 2.991 units, and the effect of the demographic variables on job satisfaction would start from this constant coefficient of 2.991. Although this statement might seem meaningless

at first glance, we found that the Model 1 column in the “Model Summary” table explained 1.8% of the change in job satisfaction. Therefore, there is a 98.2% value that we cannot explain in the formation of job satisfaction ($100 - 1.8 = 98.2$). The constant-coefficient (constant value) is high due to the large unexplained variance. When we examine Model 2, we see that the constant coefficient value (constant) is lower (0.186). When we look at Model 2 in the “Model Summary” table (Table 4.9), we see that the demographic variables + transactional leadership + charismatic leadership explain 37.1% of the change in job satisfaction. Unexplained variance is less than in Model 1. Therefore, the constant coefficient value (constant) is correspondingly lower (0.186). The B values in the prediction line (0.261 for transactional leadership and 0.451 for charismatic leadership) are called the slope. Since the slope has a positive value, the correlation will also be positive. B values indicate how many units an increase in the independent variable causes the dependent variable to increase by. The following comment can be made regarding our example. A one-unit increase in transactional leadership causes a 0.261-unit increase in job satisfaction, while a one-unit increase in charismatic leadership causes an increase of 0.451 units in job satisfaction. In line with these coefficients:

We can say that the effect of transactional leadership on job satisfaction is significant and positive ($\beta=0.261$, $p<0.05$), and the effect of charismatic leadership on job satisfaction is significant and positive ($\beta=0.451$, $p<0.05$).

As previously stated, our hypotheses are:

Hypothesis 0 (H_0): There is no linear relationship between leadership styles (transactional and charismatic leadership) and job satisfaction.

Hypothesis 1 (H_1): Transactional leadership significantly and positively affects job satisfaction.

Hypothesis 2 (H_2): Charismatic leadership has a statistically significant and positive effect on job satisfaction.

H_0 : By examining the sig value in the ANOVA table (Sig = .000, $p<0.05$), it can be stated that the model is statistically meaningful and H_0 hypothesis is not supported.

H_1 : Based on an examination of the regression coefficients in the “Coefficients” table ($\beta=0.261$, $p<0.05$), it can be stated that the effect of transactional leadership on job satisfaction is statistically significant and positive, and therefore, H_1 hypothesis is supported.

H_2 : According to the regression coefficients in the “Coefficients” table ($\beta=0.451$, $p<0.05$), it can be stated that the effect of charismatic leadership

on job satisfaction is statistically significant and positive, and therefore, H1 hypothesis is supported.

Note: “Collinearity Diagnostics” and “Excluded Variables” tables are provided in SPSS’s Output file. These tables include data related to the calculation of the Collinearity Statistics values in the far-right column of the “Coefficients” table. In this part of the research, the data in these two tables will not be interpreted.

INTERPRETATION OF RESEARCH FINDINGS

As will be recalled from a previous section, the mean and standard deviation values of the variables are given in the correlation table. To avoid duplication in studies, these values are not shown in the regression table. Researchers usually state the regression values in a separate table in their studies.

Hierarchical regression analysis was conducted to determine the effect of leadership styles on job satisfaction. In this context, the regression results are presented in Table 4.12.

Table 4.12. Regression Results

Variables	Job Satisfaction	
	Model 1	Model 2
Age	-.088	.118*
Marital status	.217	.162
Education	.069	.062
Transactional leadership		.261**
Charismatic leadership		.451***
F	1.969	19.308
R ²	.018	.760
ΔR ²	.037	.354

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$

In the first stage of the hierarchical regression analysis, the demographic variables were controlled (Model 1), while the independent variables were included in the model in the second stage (Model 2). While the demographic variables were under control, it was determined that transactional leadership ($\beta = 0.261$, $p < 0.05$) and charismatic leadership ($\beta = 0.451$, $p < 0.05$) both affected job satisfaction significantly and positively. In the light of these findings, while the H0 hypothesis was not supported, the H1 and H2 hypotheses were supported.

NOTE: Today, the algorithm used by SPSS for regression analysis is gradually losing its validity, and there are increasing opinions that the strength of the statistical analysis provided by SPSS is not sufficient. Therefore, it is useful to use structural equation modeling, particularly for regression analysis. For structural equation modeling, AMOS, R, or Lisrel programs can be used.

REGRESSION ANALYSIS WITH AMOS

While performing regression analysis with AMOS, the model to be tested must first be drawn with AMOS Graphics. The process of drawing the model will not be described at this stage. The methods previously described in this book should be used for model drawing.

The Data-1.sav file will be used for regression analysis with AMOS. As will be recalled, transactional leadership, charismatic leadership and job satisfaction variables were measured in this file. Additionally, gender, marital status, age and education were also included as demographic variables. In the section on regression analysis with SPSS, the aim, model and hypotheses of the research were explained in detail. Here, testing will be performed with already drawn models.

Observed values or average values calculated on the SPSS dataset can be used when performing regression analysis with AMOS. Here, the first step will be to evaluate the mean values calculated in SPSS as the observed variable, and then we will perform a regression analysis by drawing the latent variable and observed variable model, which is the main difference with AMOS.

The first regression model to be tested is the model that treats the mean variables calculated as a repetition of SPSS as directly observed variables. The model is shown in Figure 4.12.

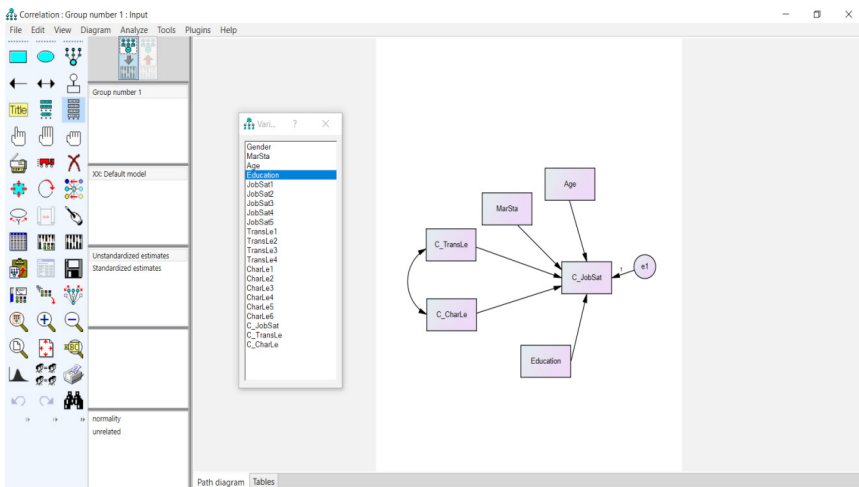


Figure 4.12. Regression Model with Mean Values as Observed Variable

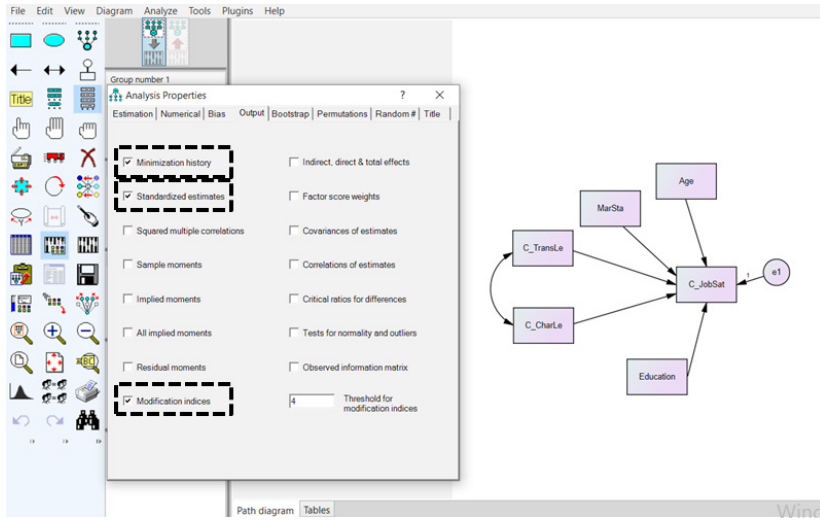


Figure 4.13. Selection of Analysis Properties

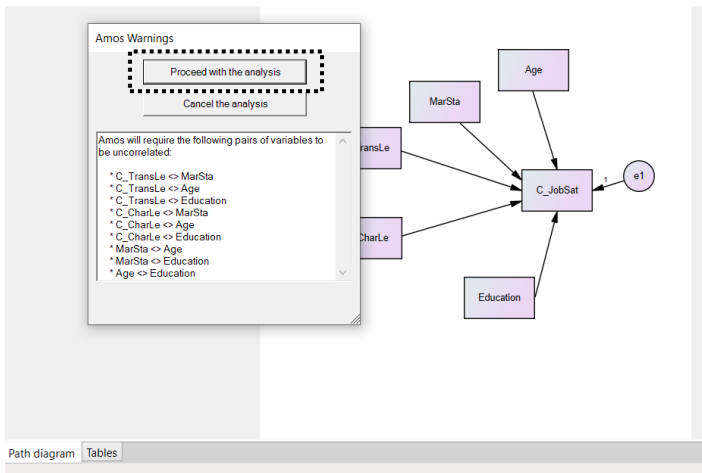


Figure 4.14. Pre-Analysis Warning

As can be seen in the model, transactional and charismatic leadership predict job satisfaction, and marital status, age and gender are also included in the regression as demographic variables. At this stage, before proceeding with the analysis, as shown in Figure 4.13, the “Standardized Estimates” and “Modification indices” options should be checked in the “Output” tab under the “Analysis Properties” menu and both standardized beta values and model fit indices should be calculated at the end of the analysis. When the “Calculate Estimates” button is clicked, an error message will appear, as shown in

Figure 4.14. This error message indicates that the process will continue without a correlational relationship between some observed variables. Here, we continue by selecting the “Proceed with the analysis” option.

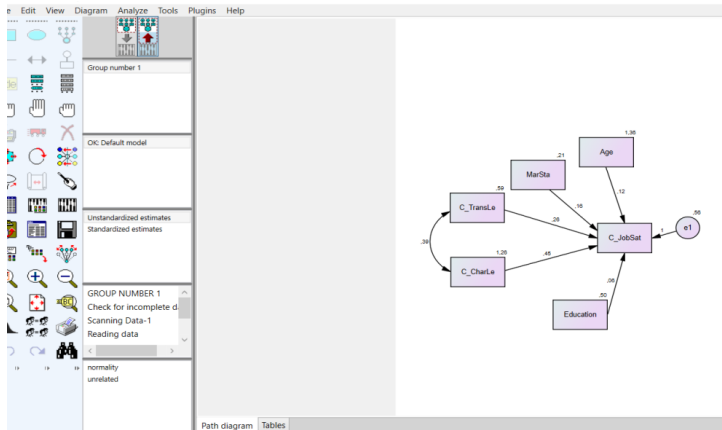



Figure 4.15. Standardized Beta Coefficients on the Default Model Tab

The standardized beta coefficients obtained as a result of the analysis are shown in Figure 4.15. This shape can be obtained by clicking the “Default Model” tab in the AMOS menu. Click the “View Text”  button to see the values obtained as a result of the regression analysis. When the “Estimates” tab is selected in the window that appears, the values in Figure 4.16 will be displayed.

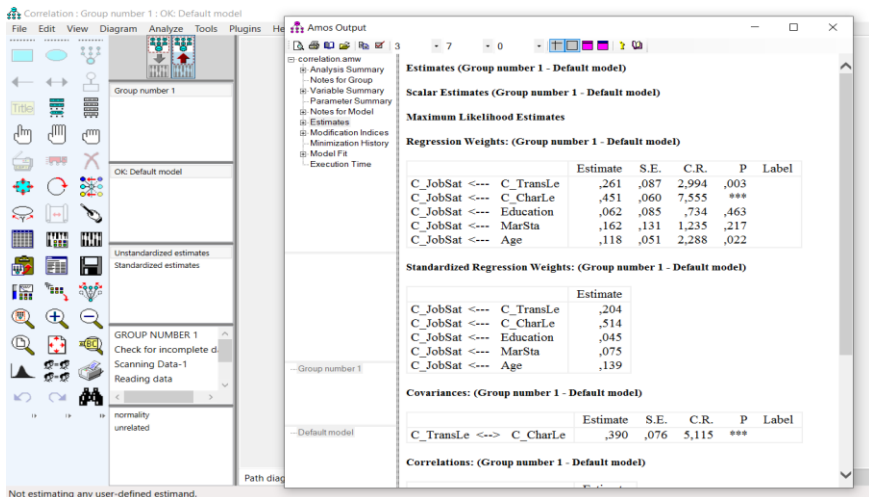


Figure 4.16. “Estimates” Tab Screen

When the obtained values are examined, it is seen that the variables of transactional leadership, charismatic leadership and age have a significant effect on job satisfaction. The beta and standardized beta values are the same as those with SPSS above. It is important to remember that the same model test could also be tested with different prediction models and different analysis parameters could be selected. Since the focus of this book is not to introduce the AMOS program, details on this topic will not be provided. However, for example, different estimation models can be tried by selecting “Generalized least squares” instead of “Maximum likelihood” in the “Estimation” tab, and “Indirect, direct & total effects”, “Tests for normality and outliers” in the “Output” tab. Additionally, direct, indirect and total effects on the dependent variable can be calculated and normality tests of the data can be performed. The reason why we mention the “Indirect, direct & total effects” tab in particular at this stage is that it will be used in the mediation test with AMOS, which will be presented in the following sections. Researchers who use the program could make different calculations. The reporting and interpretation of the findings will be as described in the SPSS analysis section above.

The second way of performing regression analysis with AMOS is to design a structural regression model. While doing this, unobservable and observed variables are plotted on the model, instead of taking the calculated averages directly as the observed variable. The same example above is drawn as a structural regression model shown in Figure 4.17.

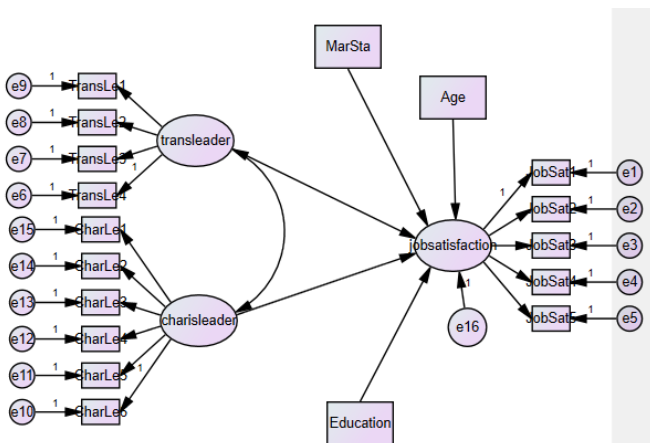


Figure 4.17. Structural Regression Model

As shown in Figure 4.17, latent variables and observed variables are included in the structural regression model. In fact, this is perhaps the different

and most meaningful aspect of testing regression analysis with AMOS. The results are presented in Figure 4.18.

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

		Estimate	S.E.	C.R.	P	Label
jobsatisfaction <---	transleader	,234	,148	1,583	,113	
jobsatisfaction <---	charisleader	,698	,126	5,532	***	
jobsatisfaction <---	Education	,041	,092	,447	,655	
jobsatisfaction <---	Age	,100	,056	1,796	,073	
jobsatisfaction <---	MarSta	,163	,142	1,144	,252	
JobSat1 <---	jobsatisfaction	1,000				
JobSat2 <---	jobsatisfaction	,924	,101	9,180	***	
JobSat3 <---	jobsatisfaction	,790	,104	7,618	***	
JobSat4 <---	jobsatisfaction	,950	,109	8,678	***	
JobSat5 <---	jobsatisfaction	,875	,110	7,993	***	
TransLe4 <---	transleader	1,000				
TransLe3 <---	transleader	,806	,140	5,764	***	
TransLe2 <---	transleader	1,314	,166	7,898	***	
TransLe1 <---	transleader	1,092	,148	7,385	***	
CharLe6 <---	charisleader	1,000				
CharLe5 <---	charisleader	1,130	,130	8,714	***	
CharLe4 <---	charisleader	,921	,114	8,070	***	
CharLe3 <---	charisleader	1,110	,121	9,161	***	
CharLe2 <---	charisleader	,854	,109	7,811	***	
CharLe1 <---	charisleader	1,033	,128	8,050	***	

Covariances: (Group number 1 - Default model)

Figure 4.18. Structural Regression Model Test Results “Estimates” Tab

As can be seen in Figure 4.18, while the variables charismatic leadership and age have a significant effect on job satisfaction, the effects of other variables are insignificant. As can be immediately observed, these results are quite different from the above regression analyses despite the use of exactly the same data set. The most important difference is that the effect of transactional leadership is meaningless. As can be understood, analyses with SPSS or directly observed variables with AMOS yield different results since latent variables are calculated by averaging. On the other hand, AMOS first calculates three latent variables (transactional leadership, charismatic leadership and job satisfaction) in the confirmatory factor analysis model and then performs the regression. In the standard regression calculation conducted by taking the average, the basic assumption is that all observed variables that constitute the latent variable have an equal effect on the latent variable. However, the results reveal that this is

not true because this structural regression analysis with AMOS indicates that not all observed variables are equally effective in calculating latent variables. Therefore, this error is inevitable in regression analyses based on means (it is useful to remember the Type-1 and Type-2 error issues at this point). Therefore, as previously mentioned, structural regression models are increasingly preferred today because they can work with small sample sizes, minimize measurement errors, and do not require normal distribution assumptions.

References

- [1] Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (2013). *Multivariate data analysis*. Pearson Education Limited.
- [2] George, D., & Mallery, M. (2010). *SPSS for Windows Step by Step: A simple guide and Reference*, (10 Ed.) Boston: Pearson
- [3] Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (sixth ed.). Pearson, Boston
- [4] Kim, H. Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution using Skewness and Kurtosis. *Restorative Dentistry & Endodontics*, 38(1), 52-54.

CHAPTER 5

ANALYSIS OF THE INTERMEDIATE VARIABLE EFFECT

Differences between the Mediator and Moderator Variable

When researchers detect an empirical relationship between two variables, they want to include a third variable in the model as the next step to clarify the nature of this simple relationship. This third variable, called the mediator or moderator, is frequently preferred by researchers to explain the relationship between the independent and the dependent variable in depth and to examine the effects in the relationship between them that do not appear at first glance.

Researchers are mostly undecided about whether the third variable is the “mediator” or the “moderator” variable while constructing a model. The basis of this indecision lies in the fact that the difference between the two variables is not well known. Although moderating and mediating variables are different concepts, they are often confused by researchers. For this reason, it is useful to know the meaning of the two concepts and what purpose they serve. These two terms are related to better understanding the relationship between an independent and a dependent variable. As a third variable, the mediating and moderating variables are concerned with controlling how they fit into the relationship between the independent variable and the dependent variable, and the similarities between the mediator/moderator variables are limited to this. It is easy to confuse these variables, but it should be noted that the mediator and moderator variables are two different terms that require different statistical approaches. The differences between them are briefly summarized below:

1. The moderator variable affects the strength of the relationship between two variables (dependent and independent variable), whereas the mediator variable explains the relationship between two variables. In other words, the moderator variable changes the relationship between the independent and

dependent variables. The mediating variable, on the other hand, functions as a channel that affects the dependent variable.

2. Analyses for the moderator variable are a way of determining whether it affects the strength or direction of the relationship between the dependent and independent variable as a third variable. When the intensity/strength of the moderator variable is changed, the relationship between the dependent and independent variables could change statistically. The mediator variable, on the other hand, could reduce or eliminate the strength of the relationship between the dependent and independent variables. As can be seen, the moderator variable is effective in its value/strength in the relationship between the dependent and independent variables. This is not the case with the mediator variable.

3. In models for the mediator variable, there is a third (intermediate) variable with a constant strength/effect in the relationship between the independent variable and the dependent variable, while the strength/intensity (low, medium, and strong) of the third (moderator) variable is not constant in the models for the moderator variable. Therefore, while the effect of the independent variable on the dependent variable is shaped by the strength of the moderator variable, this relationship changes with the constant presence of the mediator variable in models for the mediating variable.

4. The mediator variable is a more explicit concept than the moderator variable. The mediator variable mediates the relationship between the dependent and independent variables and explains the reason for the existence of this relationship. The main purpose of mediator variable analysis is to determine whether the effect of the mediating variable is stronger than the direct effect of the independent variable.

5. While models for the mediating variable make theoretical claims about the causal relationships between variables, this is not the case for models for the moderator variable.

6. Models for the mediating variable try to determine how and why the relationship between two variables occurs, while models for the moderator variable try to determine in which situations or when the relationship between two variables increases/decreases.

7. In models for the mediating variable, the mediating variable is the consequence of the independent variable and the antecedent of the dependent variable. In other words, there is a causal relationship between all variables. On the other hand, in models for the moderator variable, the moderator variable

may not be the consequence of the independent variable, and nor may it be the antecedent of the dependent variable.

8. Moderation refers to the conditions under which an impact changes in size, while mediation refers to the underlying mechanisms and processes that unify antecedents and consequences.

As can be seen, there are fundamental differences between the mediator and the moderator variables. Whether a variable is hypothesized as a mediator or moderator is primarily determined by your research hypothesis, which you have formulated in line with the theoretical framework. In determining your research hypothesis, you should review the literature on your research topic and analyze the research in the context of your research question. If the literature for your research question is relatively limited, it would be appropriate to make use of theories Lipsey [1] argued that theory should be used to shed light on the black box (mediator, moderator) of how the independent variable affects the dependent variable. Many researchers in the literature constantly discuss their research questions as being “theory-based”. In particular, the theory of social change, the theory of reasoned action, the theory of equality, the theory of self-determination, the theory of planned behavior, and the theory of expectation are among the frequently used theories.

ANALYSIS OF THE MEDIATOR VARIABLE EFFECT

The mediator variable works as a “mechanism” between the dependent and independent variables and helps to explain the relationship between the two variables ([2,3,4]). The mediator variable tries to explain how or why the relationship between two variables occurs and defines the psychological process between the variables. The existence of the mediator variable occurs when the relationship between the dependent and independent variable could be partially or completely explained by a third variable. In other words, if the effect of the independent variable on the dependent variable occurs partially or completely through a third variable, it is possible to discuss the existence of a mediator variable.

Statistical tests for the mediator variable aim to directly evaluate and test this proposition. In tests for the mediator variable, when the mediator variable is kept constant or controlled, it is determined to what extent the relationship between the independent variable and the dependent variable changes. The mediator variable is sometimes denoted as the intervening variable. An intervening variable in statistics is generally considered a subtype of a mediator variable. However, the lines between the two terms are somewhat blurred and they are often used interchangeably. In our study, we will consistently use the term “mediator variable” without having a concept discussion.

In simple mediator models, the effect of an independent variable (X) on a dependent variable (Y) is transmitted via a third variable (mediating variable, M). In other words, X affects M and M affects Y . Figure 5.1 below shows the path diagram for a simple mediation model.

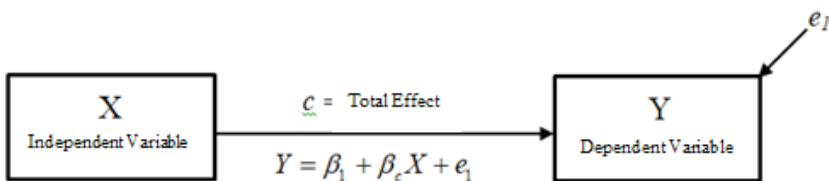


Figure 5.1. Simple Regression Model

The example above shows a model that associates an independent variable (X) with a dependent variable (Y), and it is a model that simply shows the relationship between one variable and another variable. Note that the path from

X to Y is assigned the symbol c . The coefficient e_1 represents the unexplained part in the relationship between X and Y.

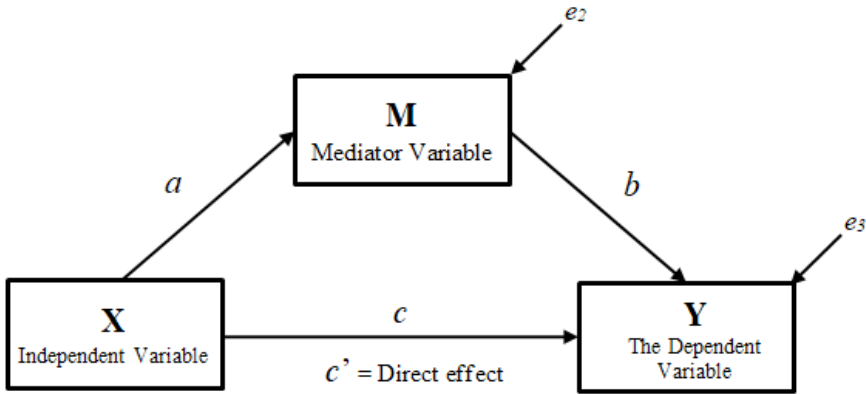


Figure 5.2. Simple Mediation Model (Extensive demonstration)

$$Y = \beta_3 + \beta_c X + \beta_b M + e_3$$

$$M = \beta_2 + \beta_a X + e_2$$

The above model represents a simple mediation model, which will be mentioned throughout most of the book. Arrows show the direction of the relationship between variables (X to M, M to Y, and X to Y). Also, symbols are displayed above each arrow, which correspond to the relationships between variables (a , b , and c'). The parameter e_3 represents the unexplained part of the relationship of X and M with Y. The parameter e_2 represents the unexplained part of M in its relation to X.

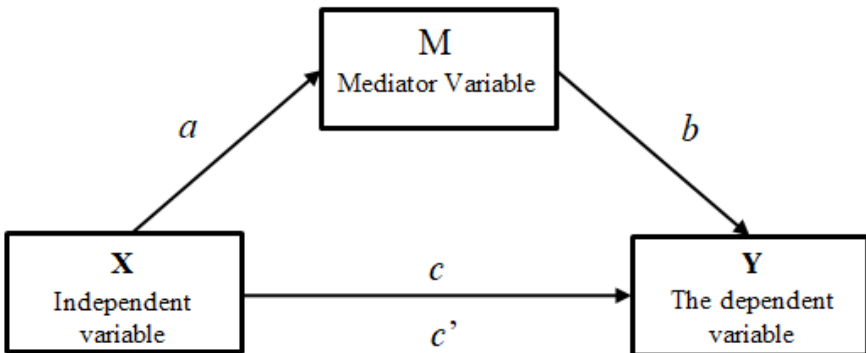


Figure 5.3. Simple Mediation Model

In the model, the independent variable (X) affects the dependent variable (Y) through a third variable, the mediator variable (M). Thus, the mediator variable (M) helps to explain the effect of X on Y and also explains how and why the relationship between the two variables occurs. For example, organizational commitment could be an important mediator variable in the effect of job satisfaction on job performance. Based on this example, the effect of job satisfaction on job performance is possible through organizational commitment. Therefore, if job satisfaction affects job performance and the relationship between the two variables weakens when the organizational commitment variable is included in this relationship, partial mediation can be mentioned, whereas if the relationship between the two variables disappears completely, then full mediation is mentioned.

As can be seen, there are two mediating effects: partial and full mediation.

Partial mediation occurs when the mediator variable is only responsible for part of the relationship between the dependent and independent variables. In other words, even when there is no mediator variable, there is a relationship between the dependent and independent variables, but the relationship is not that strong.

Full mediation is the situation in which the entire relationship between the dependent and independent variables occurs through the mediator variable. In the case of full mediation, the independent variable cannot directly affect the dependent variable and rather does so through a third variable, namely a “mediator variable”. If the mediator variable is removed from the model, the relationship between the dependent and independent variables disappears. In other words, the relationship between the independent and the dependent variables becomes statistically insignificant.

There are four different methods for determining the mediator variable.

1. *The four-step method proposed by Baron and Kenny [5]*

It is a four-stage regression (Multiple regression) method introduced to the literature by Baron and Kenny [5]. It is the oldest method in the field of mediation modeling. Although Baron and Kelly's method is one of the original methods used for mediation testing, it has low statistical power. This method will be described in our book because it provides a very explicit approach to relating variables and is occasionally requested by reviewers (journal reviewers or editors).

2. *Mediation testing with AMOS Structural Equation Modeling*

Many researchers argue that structural equation modeling (SEM) is better than Baron and Kenny's [5] "causal steps" approach [6, 7]. The main advantage of SEM is that it simultaneously estimates the statistical values calculated in each step in Baron and Kenny's [5] method, instead of assuming that the analyses conducted in the four steps specified in Baron and Kenny's [5] method are independent. However, due to the technical complexity of SEM, it is not preferred by researchers. In our book, SEM will be explained in its most basic form and the technical complexity of SEM will be avoided.

3. *Mediation test with SPSS Process Macro*

This is the contemporary bootstrap method developed by Hayes [8] which is frequently used by today's researchers to detect the mediating effect. This method provides gives stronger statistical results than Barron and Kenny's [5] method and is easier to perform than SEM. It evaluates according to the confidence interval, not the p (significance) coefficient applied by Baron and Kenny [5].

4. *Least Squares Method (Smart Pls)*: It is a kind of structural equation modeling. However, its validity has not yet been fully proven, and for this reason, this method will not be included in our book.

ANALYSIS OF THE MEDIATION EFFECT WITH THE BARON & KENNY CAUSAL STEPS APPROACH [5]

In 1986, Baron and Kenny published a paper in which they described how statistical tests of mediation hypotheses are performed. Since the publication of Baron and Kenny's [5] article describing a method for evaluating mediation, the use of mediation models in the social sciences has increased dramatically. Until the bootstrap method developed by Hayes was recently introduced, Baron and Kenny's [5] causal step approach remained the most widely used mediation test.

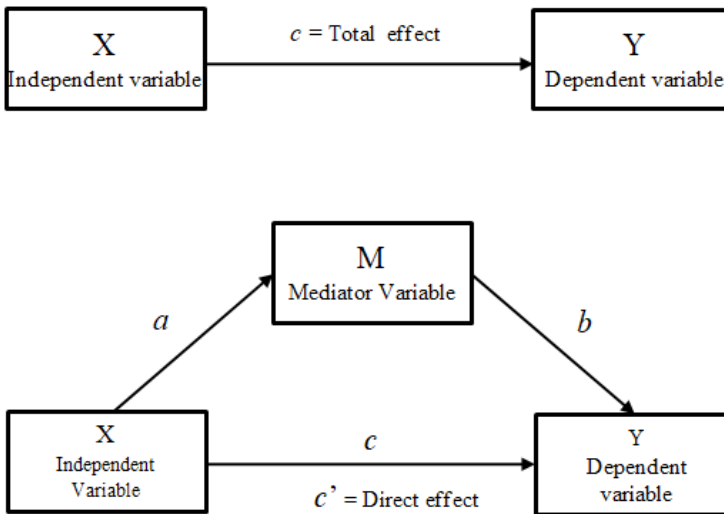


Figure 5.4. Simple Mediation Model

The above model represents the simple mediation model, and in the following chapters of the book, the analysis will be explained with this model.

Baron and Kenny's [5] causal steps approach is a 4-step method: Steps 1 and 2 use basic linear regression, while steps 3 and 4 use multiple regression. For mediation to be tested, each of the four steps must be implemented.

1. The total effect of the independent variable (X) on the dependent variable (Y) must be significant (path c).
2. The effect of the independent variable (X) on the mediator variable (M) should be significant (path a).

3. The mediator variable (M) must have a significant effect on the dependent variable (Y) (path b).

4. When the independent variable and the mediator variable are included in the regression analysis, the effect of the independent variable on the dependent variable (c') should be less than the total effect of X on Y (c) (partially mediated), or the statistically significant effect should disappear (full mediation). As a result of the analysis performed in the fourth step, if $c' = 0$, full mediation can be mentioned, and if $c' < c$, partial mediation can be mentioned.

Baron and Kenny [5] stated that partial mediation is a more realistic expectation than full mediation in studies conducted in the social sciences. Studies have found that most articles result in partial mediation when all the tests suggested by Baron and Kenny [5] are conducted and reported appropriately [9].

Finally, in addition to these four causal steps, Baron and Kenny [5] also recommended that the Sobel test be performed to test the statistical significance of the indirect effect.

MEDIATION ANALYSIS WITH SPSS

Research question: Does self-efficacy play a mediator role in the effect of transformational leadership on the perception of social support?

Let’s recall Baron and Kenny’s [5] four-stage causal steps method according to our research question. Our research model is presented in Figure 5.5.

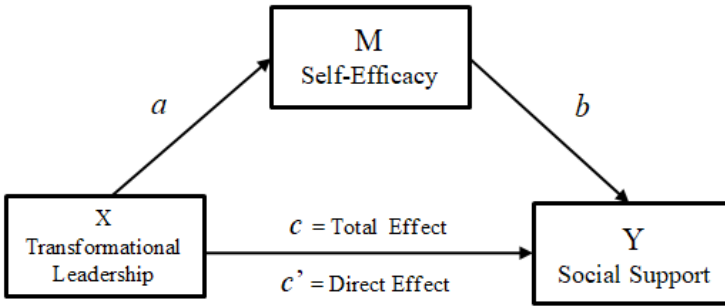


Figure 5.5. Research Model

We will now test mediation using the four-step causal step method of Baron and Kenny [5]. For this, the following steps will be followed.

Step 1: The effect of X on Y will be tested by simple regression analysis and path c will be determined.

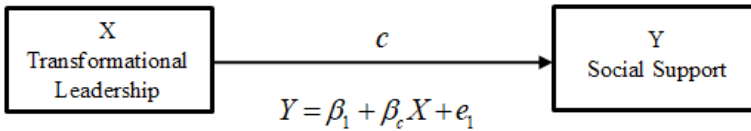


Figure 5.6. Effect of X on Y

Step 2: The effect of X on M will be tested by simple regression analysis and path a will be determined.

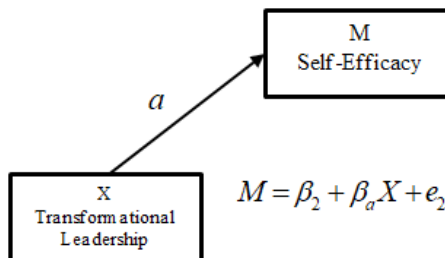


Figure 5.7. Effect of X on M (path a)

Step 3: The effect of M on Y (with X under control) will be tested by multiple regression analysis and path b will be determined.

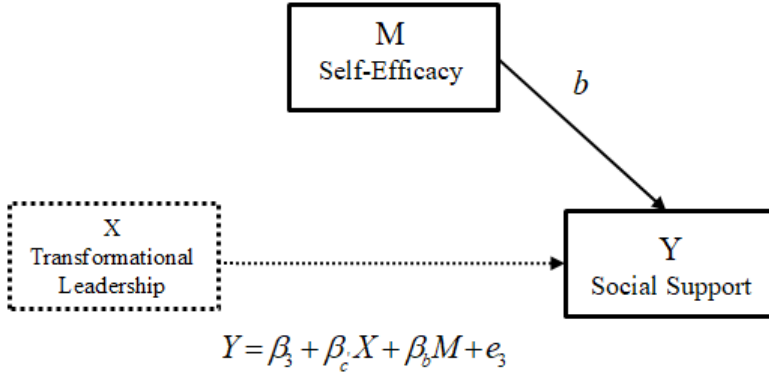


Figure 5.8. Multiple Regression Analysis (path b)

Step 4: The effect of X and M on Y will be tested by multiple regression analysis and the path c' , in other words its direct effect (the effect of X on Y), will be determined.

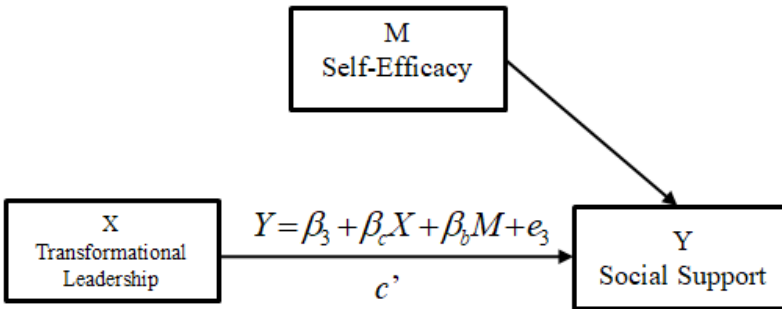


Figure 5.9. Multiple Regression Analysis (path c')

In this method proposed by Baron and Kenny [5], if the c' effect differs from the c effect (decreases or loses its meaning), this effect is interpreted as being shared by the mediator variable, and the presence of the mediator variable (partial or full mediation) can be mentioned

In the causal steps method of Baron and Kenny [5]:

Total effect (c) = direct effect (c') + indirect effect ($a \times b$)

- Full Mediation: $c' = 0$

- Partial Mediation: $c' < c$

The four-stage 'Causal Steps' method proposed by Baron and Kenny [5] is frequently used in social science research, and the decision regarding partial

or full mediation is made according to the research result. An important issue to consider at this stage is that the Sobel test checks the statistical significance of the last step of the causal steps method. Therefore, researchers should check the statistical significance of the regression analysis performed as the last step with the Sobel test [10].

For regression analysis, we first open the “Data-3.sav” file (You can access this file from www.indataanalysis.com).

We follow the steps below:

For Step 1:

Click on Analyze ----> Regression ----> Linear...

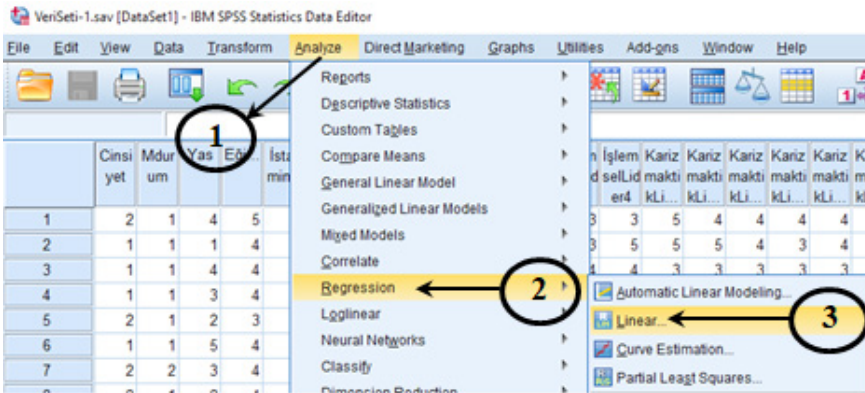
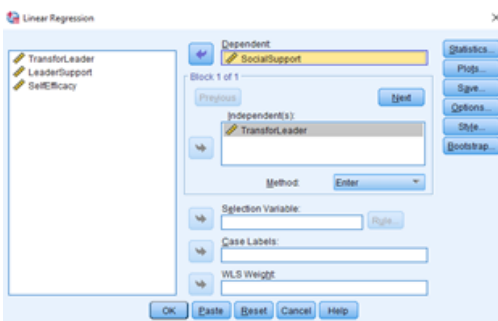


Figure 5.10. SPSS “Analyze” Screen



In the window that opens, assign the “Perception of Social Support” variable to the dependent variable and the “Transformational Leadership” variable to the Independent variable columns, and finally, click the “OK” button.

Figure 5.11. “Linear Regression” Screen

As a result of the performed analysis, the following tables will appear in SPSS output.

Table 5.1. ANOVA Table

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	229,455	1	229,455	25,714	,000 ^b
	Residual	1909,605	214	8,923		
	Total	2139,060	215			

a. Dependent Variable: SocialSupport
 b. Predictors: (Constant), TransforLeader

The model is statistically significant ($p < 0.05$)

Table 5.2. Regression Coefficients Table

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	10,129	,884		11,455	,000
	TransforLeader	,298	,059	,328	5,071	,000

a. Dependent Variable: SocialSupport

path c $\beta = 0.328$

Effect is significant at the 0.001 level level

When the related tables are examined, it is seen that transformational leadership has a positive and significant effect on the perception of social support ($\beta = 0.328, p < 0.05$).

After completing Step 1, we can proceed to Step 2.

For Step 2

Click on Analyze ---> Regression ----> Linear....

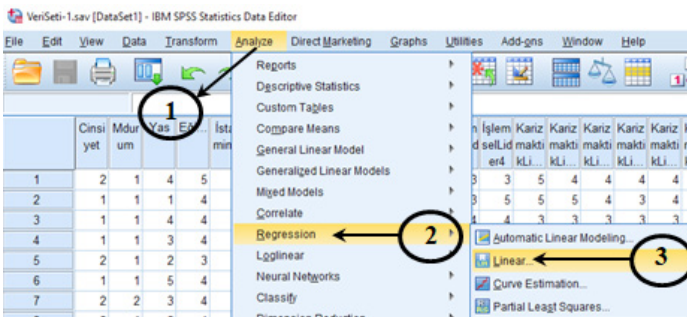


Figure 5.12. SPSS “Analyze” Screen

In the window that opens, we assign the “Self-Efficacy” variable to the dependent variable and the “Transformational Leadership” to the Independent variable part.

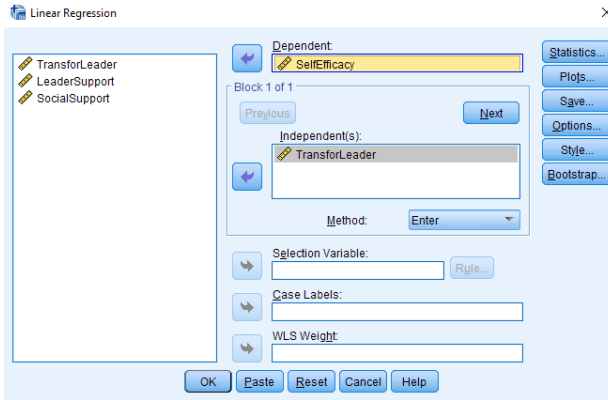


Figure 5.13. “Linear Regression” Screen

As a result of the analysis, the following tables will appear in SPSS.

Table 5.3. ANOVA Table

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3256,839	1	3256,839	72,793	,000 ^b
	Residual	9574,582	214	44,741		
	Total	12831,421	215			

a. Dependent Variable: SelfEfficacy

b. Predictors: (Constant), TransforLeader

The model is significant
($p < 0.05$)

Table 5.4. Regression Coefficients Table

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	22,114	1,980		11,169	,000
	TransforLeader	1,121	,131	,504	8,532	,000

a. Dependent Variable: SelfEfficacy

path a $\beta = 0.504$

The effect is significant at the 0.001 level

When the related tables are examined, it is seen that transformational leadership has a positive and significant effect on the self-efficacy of the employees ($\beta = 0.504, p < 0.05$).

After completing the second step, we can proceed to Step 3.

For Steps 3 and 4

In this step, multiple regressions will be performed where the independent variable (Transformational Leadership) and the mediator variable (Self-Efficacy) will be included in the analysis at the same time, and their effect on the perception of social support will be determined.

Click on Analyze ---> Regression ----> Linear....

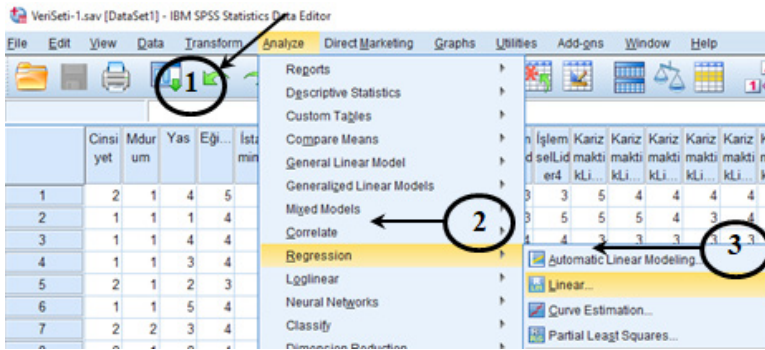


Figure 5.14. SPSS “Analyze” Screen

In the window that opens, we assign the “Perception of Social Support” variable as the dependent variable, and the “Transformational Leader” and “Self-Efficacy” variables as independent variables.

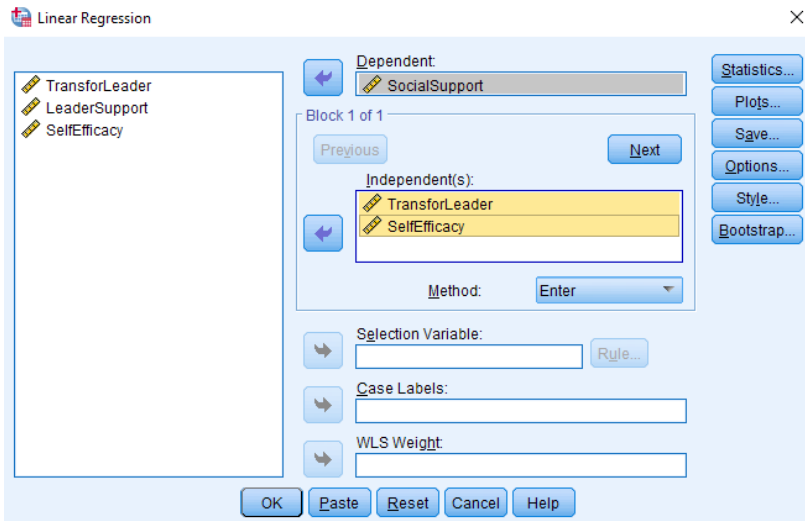


Figure 5.15. “Linear Regression” Screen

Table 5.5. ANOVA Table

ANOVA^a

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	296,726	2	148,363	17,153	,000 ^b
	Residual	1842,334	213	8,649		
	Total	2139,060	215			

a. Dependent Variable: SocialSupport

b. Predictors: (Constant), SelfEfficacy, TransforLeader

The model is significant (p < 0.05).

Table 5.6. Regression Coefficients Table

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	8,275	1,095		7,555	,000
	TransforLeader	,204	,067	,224	3,044	,003
	SelfEfficacy	,084	,030	,205	2,789	,006

a. Dependent Variable: SocialSupport

Path c' $\beta = 0.224$, The effect is significant at the 0.05 level.

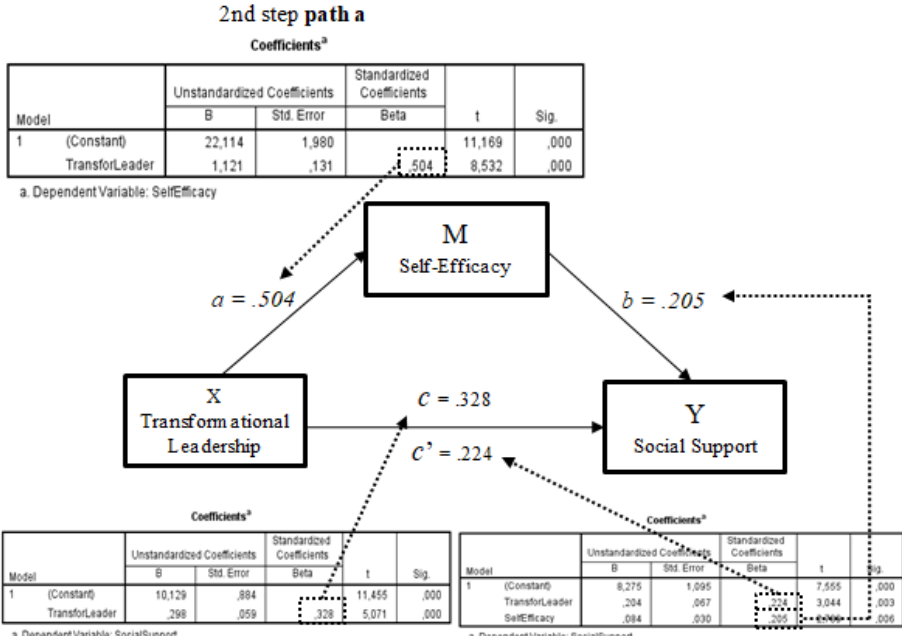
Path b $\beta = 0.205$, The effect is significant at the 0.05 level.

Therefore, self-efficacy has a positive and significant effect on the perception of social support ($\beta = 0.205$, $p < 0.05$). With this finding, the third step of the causal steps suggested by Baron and Kenny [5] was also provided.

As the fourth step, path c' should be checked. The indirect effect of transformational leadership on the employees' perception of social support is positive and significant ($\beta = 0.224$, $p < 0.05$). This finding reveals that the fourth step is achieved.

If we present all the analyses conducted in the four steps of the model:

2nd step **path a**



Total effect (c) = direct effect (c') + indirect effect (a x b)
 $0.328 = 0.224 + (0.504 \times 0.205)$

Note:

Full mediation: $c' = 0$ Partial mediation: $c' < c$

In the process performed in the first step (detection of the c path), it was found that there was a $\beta = 0.328$ effect at the 0.001 level of significance on the independent variable “Transformational Leadership” and the dependent variable “Social Support”; on the other hand, in the regression analysis performed with the inclusion of the mediator variable in the model, it was significant at the 0.05 level. It has been determined that there is an effect at the level of $\beta = 0.224$. In other words, both the effect and significance of transformational leadership on the perception of social support decreased ($0.328 - 0.224 = 0.104$). In this case, it can be said that partial mediation is supported. However, to make this decision,

the significance of partial mediation should be tested using the Sobel test. There are multiple sites on the Internet on which the Sobel test can be performed. To find a suitable website, simply type “Sobel Test” into any search engine. The Sobel test at <http://quantpsy.org/sobel/sobel.htm> was used in this research.

The significance of the effect of the mediator variable is checked filling in the relevant results on the fields displayed on the web page.

Table 5.7. Sobel Test Table

Input:		Test statistic:	Std. Error:	p-value:
a	<i>a</i>	Sobel test:	If this value is $p < 0.05$, it is	?
b	SH	Aroian test:	concluded that the mediating effect is	
s_a	<i>b</i>	Goodman test:	significant.	
s_b	SH	Reset all	Calculate	

Perform the Sobel Test, the research findings must be entered in the relevant sections of the Sobel Test table and the “Calculate” button must be clicked. For the convenience of the readers, all the procedures are summarized below.

Coefficients^a 2nd step path a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	22,114	1,980		11,169	,000
	TransforLeader	1,121	.131	.504	8,532	,000

a. Dependent Variable: SelfEfficacy

Input:		Test statistic:	Std. Error:	p-value:	
a	0.504	Sobel test:	3.35248678	0.03081891	0.00080089
b	0.205	Aroian test:	3.32555718	0.03106848	0.00088242
s_a	0.131	Goodman test:	3.38008138	0.03056731	0.00072464
s_b	0.030	Reset all	Calculate		

If this value is $p < 0.05$, it is concluded that the mediating effect is significant.

Coefficients^a 3rd step path b

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	8,275	1,095		7,555	,000
	TransforLeader	.204	.067	.224	3,044	,003
	SelfEfficacy	.084	.030	.205	2,789	,006

a. Dependent Variable: SocialSupport

After determining that the effect of the mediator variable is significant, the data are reported in the findings section of the research.

REPORTING RESEARCH FINDINGS

The four-stage regression analysis proposed by Baron and Kenny [5] was performed to determine the mediating role of self-efficacy in the effect of perceived transformational leadership in an organization on employees' perceptions of social support. The results of the analysis are presented in Table 5.8.

Table 5.8. Regression Analysis Results

Causal Steps	B	SH	β	Model Statistics
<i>Step 1</i> IV: Transformational Leadership DV: Social Support	.298	.059	.328***	R ² =.107 F(5.027)=25.714 p<.001
<i>Step 2</i> IV: Transformational Leadership DV: Self-Efficacy	1.121	.131	.504***	R ² =.254 F(8.532)=72.793 p<.001
<i>Step 3 and 4</i> IV: Transformational Leadership IV2: Self-Efficacy DV: Social Support	.204 .084	.067 .030	.224** .205**	R ² =.139 F(3.044,2.789)=17.153 p<.05

*p < 0.1, **p < 0.05, ***p < 0.001

In the first stage of the regression analysis, which comprised four stages in total, it was found that transformational leadership had a positive and significant effect on employees' perception of social support ($\beta = 0.328$, $p < 0.001$), and in the second stage, transformational leadership had a significant positive effect on self-efficacy ($\beta = 0.504$, $p < 0.001$). In the third stage, self-efficacy had a positive and significant effect on the perception of social support ($\beta = 0.205$, $p < 0.05$), while in the fourth stage, multiple regression analysis was conducted to determine the indirect effect of transformational leadership on the perception of social support and it was determined that this effect was significant and positive ($\beta = 0.224$, $p < 0.05$). When self-efficacy was included in the model, the effect ($\beta = 0.328$, $p < 0.001$) and significance of transformational leadership on the perception of social support decreased ($\beta = 0.224$, $p < 0.05$). In the light of these findings, it has been determined that self-efficacy plays a partial mediating role in the effect of transformational leadership on the perception of social support. To test the significance of the effect of the mediating variable, the Sobel test was performed and the test results revealed that the mediation effect was statistically significant ($z = 3.352$, $p < 0.05$).

Criticisms of Baron and Kenny's [5] Causal Steps Approach

Baron and Kenny's procedure for determining whether an independent variable influences a dependent variable through a mediator has been adopted by researchers and used by many authors for many years, to the extent that it was requested almost reflexively by the editors who evaluated the article in the journals. However, over time, concerns about the statistical power of the Baron and Kenny [5] method have emerged. This reached the point that as a result of these criticisms, Baron and Kenny [5] were ultimately forced to change the statistical method used to test mediation [11].

The most important criticism directed towards the method proposed by Baron and Kenny [5], which is recognized as the most commonly used method, is in relation to path c . Although Baron and Kenny [5] stated that the total effect (c -path) of the independent variable on the dependent variable must be significant for the mediator effect to occur, subsequent studies have shown that this effect does not need to be significant [12]. When determining whether variable M mediates the effect of X on Y , the path from X to Y need not be considered because this path is not part of the mediating effect [13]. Indeed, the total effect specified as path c does not determine the presence or magnitude of the indirect effect [7]. In other words, it can be said that there is an indirect effect even in the absence of the total effect. With the imposition of this condition, in some cases, even though the necessary ways (paths a , b , and c') to test mediation are significant, mediation tests are not performed and mediation is ignored. This has led researchers to abandon promising projects and journals to reject articles that may be worthy of publication. Today, there is a consensus among researchers that the significance of path c is not important in detecting mediation [8, 14]. Reputable journals such as SSCI/SCI include studies in which mediation has been determined despite the insignificance of path c . Therefore, in future studies, researchers who prefer Baron and Kenny's [5] causal steps approach should continue their mediation tests regardless of the presence or absence of the direct effect when the indirect effect is supported.

Baron and Kenny [5] stated that when there is an indirect effect (ab) but no direct effect (c'), a full mediation effect occurs, but when there are both indirect and direct effects, a partial mediation effect occurs. The strength of the mediation should not be measured by the absence of a direct effect, but by the magnitude of the indirect effect.

Baron and Kenny [5] also predicted that research in the social sciences will result in partial mediation rather than full mediation. Iacobucci et al. [9] stated in a study that “most of the articles result in partial mediation when all tests are conducted and reported appropriately”. This is due to the constructional defect in the Baron and Kenny [5] method. These concepts are very sensitive to the sample size of the research. In addition, in Baron and Kenny’s method, partial and full mediation are determined according to the total effect. However, the existence of the total effect is not important for the determination of mediation.

Although Baron and Kenny discussed procedures for testing the indirect effect (ab), this test is not among the conditions specified in the causal steps approach itself. However, it is important for assessing the extent of the mediating effect and comparing effects identifying alternative mediation mechanisms.

Finally, researchers using the method of Baron and Kenny [5] predominantly rely on the Sobel [10] test to control the statistical significance of the mediating effect and also determine the significance of the mediating effect with this test. However, the Sobel test is largely considered to be an outdated method, and it is emphasized that it is also inadequate and unreliable [13]. The Sobel test examines whether the difference between the total effect and the direct effect, in other words, the indirect effect, is statistically significant. The Sobel test is based on the assumption that the sample distribution for this indirect effect is normally distributed. However, since the indirect effect is the outcome of two parameters, in the case of a positive indirect effect, the distribution is positively skewed to the left with a shorter and fatter tail (the near-zero end of the distribution). Therefore, the sample of the indirect effect does not usually show a normal distribution. Therefore, tests based on the non-normal distribution of the indirect effect produce stronger statistical results than the Sobel test [8]. The bootstrap test developed by Preacher and Hayes [15,16] solves this problem associated with the Sobel test by creating an empirical sampling distribution. In this respect, the bootstrap test, popularized by Preacher and Hayes [15], gives stronger results compared to the Sobel test. Therefore, it is recommended that the bootstrapping method is used for determining the significance of the mediation.

STRUCTURAL EQUATION MODELING (SEM) WITH AMOS

Criticisms of Baron and Kenny's [5] causal steps approach have caused researchers to move away from this method. To detect the mediating effect, researchers now often prefer the "Process Macro" method developed by Hayes. This method both provides stronger statistical results than Baron and Kenny's method and is easier to apply. However, many researchers argue that structural equation modeling (SME) gives stronger statistical results than the approach of Baron and Kenny [5] [6,7].

The main advantage of SEM is that each equation is estimated at the same time, instead of assuming that the four analysis steps specified in Baron and Kenny's [5] method are independent. This prevents the Type 1 error from occurring. In addition, it is recommended that structural equation models are used in complex models where many independent variables and more than one mediator variable are included in the model [17]. Therefore, researchers can abandon the regression approach of Baron and Kenny [5] and follow the structural equation approach. Despite all these positive aspects, the significant technical complexity of SEM means that researchers are reluctant to use this approach. However, as explained in a previous section of the book, SEM is a method that can be easily learned and applied.

Structural Equation Modeling (SEM) is a multivariate statistical method used in many disciplines, which is based on a specific theory according to the definition of observable and unobservable variables in a causal and relational model [18]. SEM can also be considered as more than one regression analysis performed at the same time. For this reason, it is observed that some authors have also used various different concepts when referring to SEM such as causal modeling, causal analysis, simultaneous structural modeling, covariance structure analysis, path analysis or confirmatory factor analysis [19].

Two important features in structural equation modeling are that the studied process includes a series of structural equations (i.e., regression equations) and that these structural equations can be visually presented with drawings such that the hypotheses can be understood more easily. In this context, SEM analysis testing all latent and/or observed variables of the created model to reveal the extent to which the result is compatible with the data at hand. If the fit indices resulting from testing the model show that there is a fit between the model

and the data, the hypotheses formed structurally are accepted, otherwise the proposed hypotheses are rejected.

SEM differs from common multivariate statistical methods in terms of certain features [18]. The main differences of SEM are presented below.

(1) Unlike other statistical methods, SEM adopts a confirmatory approach rather than an exploratory approach. Therefore, while many other statistical methods try to discover the relationships in the data set, SEM theoretically confirms the compatibility of the established relationships with the data.

(2) While traditional multivariate methods do not have the ability to calculate or correct any measurement errors, SEM clearly takes measurement errors into account in all analyses and provides very clear results in its calculations.

(3) While traditional methods can only operate on observable variables, SEM can perform tests on both observable and latent variables in the same model.

(4) In most statistical methods, analyses are carried out on individual observations, and models are established. For example, in methods such as multiple regression or Analysis of Variance (ANOVA), regression coefficients or error variance estimates are calculated by minimizing the sum of squares of the difference between the observed and estimated values of each observation. However, in structural model tests, covariances are taken into account rather than observations. Instead of minimizing the observed and expected individual values, the difference between the sample covariance matrix and the covariance matrix estimated by the model is minimized

(5) A structural model, in its most basic sense, consists of observed or latent variables and arrows showing the cause-effect relationship between them. In SEM, unlike classical statistical methods, exogenous and endogenous variable definitions are used instead of dependent and independent variables. The main reason for this is that in SEM models, while a variable is an independent variable for some variables, it can also be a dependent variable for some variables too [20]. Therefore, it is not a suitable approach to use the dependent and independent variable definitions used in traditional regression equations when discussing SEM models. In this sense, mediated models, which are the primary focus of this book, are structural equation models and SEM can be used effectively in testing such models. A simple mediation model shown in Figure 5.16 will be analyzed as a structural equation model. Notice that our research model is the same as the one we used when describing the causal steps method of Baron and Kenny [5]. The main research question of the model is “Does self-

efficacy have a mediating role in the effect of transformational leadership on the perception of social support?”. In the model, the perception of transformational leadership affects both the perception of social support and self-efficacy of the individuals, and at the same time, self-efficacy affects the perception of social support. In other words, some (perhaps all) of the effect of transformational leadership on the perception of social support is through self-efficacy. Therefore, the perception of social support is included in the model as both an affected (dependent) and an influencing (independent) variable. As previously stated, when the model is evaluated in general, self-efficacy in the model is both a dependent and independent variable. Therefore, it would be more accurate to define self-efficacy as an internal (affected) or external (affecting) variable. In this section, mediation and moderation models will be evaluated using SEM and tested with the AMOS program.

STRUCTURAL EQUATION MODELING WITH AMOS

The model that we will consider first is the simplest test of mediation, which will be performed with directly observed variables. Our research question is: Does self-efficacy have a mediating role in the effect of transformational leadership on the perception of social support? The model is presented in Figure 5.16.

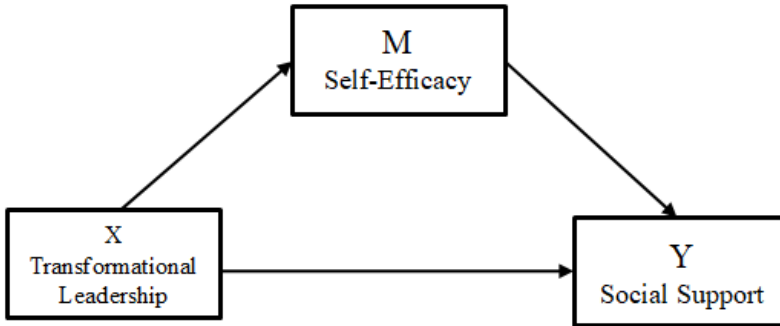


Figure 5.16. Research Model

Notice that our research model is the same as the one we used when describing the causal steps method of Baron and Kenny. Before proceeding to the SEM application, we open the “Data -3.sav” file (You can access this file at www.indataanalysis.com). The AMOS drawing of this model is presented in Figure 5.17 (the drawing in AMOS will not be explained again here, readers can refer to the section on drawing a model with AMOS in Chapter 3).

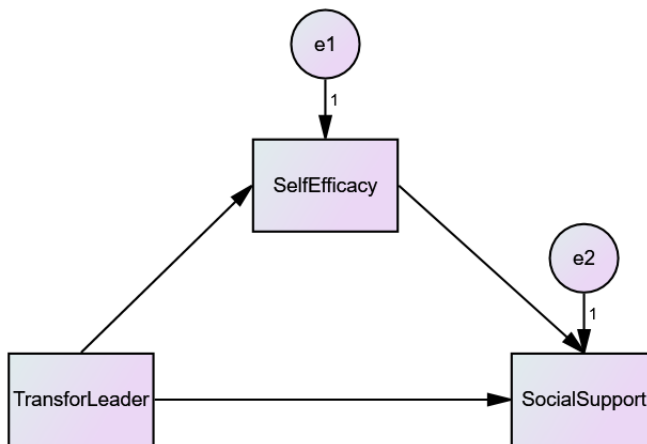


Figure 5.17. AMOS drawing of the research model

Two important steps must be taken before analyzing the model. The first of these is to select the “Standardized estimates” and “Indirect, direct & total effects” options from the “Analysis Properties” menu from the “Output” tab, and then secondly, to select the “Perform bootstrap” from the “Bootstrap” tab then enter the value of 2000, for example, and tick the “Bias-corrected confidence intervals” option. By selecting the Bootstrap option, p values are calculated at the end of the analysis, and the amount of virtual data to be produced is determined by the number 2000. By selecting the “Indirect, direct & total effects” tab, direct, indirect, and total effects will be calculated. The processes are shown in Figure 5.18 and Figure 5.19, respectively.

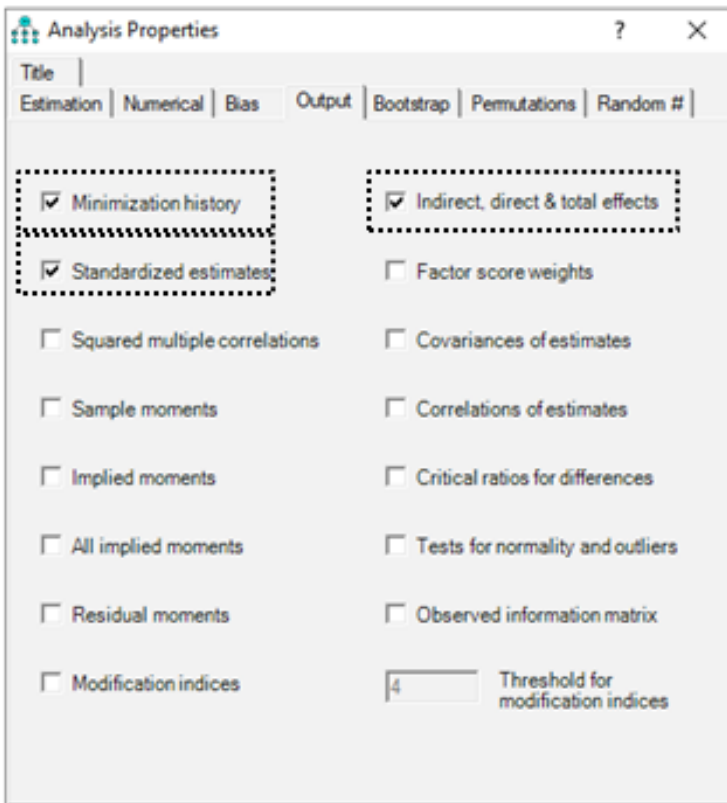


Figure 5.18. Output Tab

Analysis Properties ? X

Title |

Estimation | Numerical | Bias | Output | **Bootstrap** | Permutations | Random #

Perform bootstrap 200 Number of bootstrap samples

Percentile confidence intervals 90 PC confidence level

Bias-corrected confidence intervals 90 BC confidence level

Bootstrap ADF Monte Carlo (parametric bootstrap)

Bootstrap ML Report details of each bootstrap sample

Bootstrap GLS Bollen-Stine bootstrap

Bootstrap SLS 1 Bootfactor

Bootstrap ULS

Figure 5.19. Bootstrap Tab

The “Estimates” values generated by testing the model are presented in Table 5.9.

Table 5.9. Estimates Values

Estimates (Group number 1 - Default model)
Scalar Estimates (Group number 1 - Default model)
Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
SelfEfficacy	<---	TransforLeader	1.121	.131	8.552	***	
SocialSupport	<---	SelfEfficacy	.084	.030	2.802	.005	
SocialSupport	<---	TransforLeader	.204	.067	3.058	.002	

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
SelfEfficacy	<---	TransforLeader	.504
SocialSupport	<---	SelfEfficacy	.205
SocialSupport	<---	TransforLeader	.224

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
TransforLeader	11.992	1.157	10.368	***	
e1	44.327	4.275	10.368	***	
e2	8.529	.823	10.368	***	

Matrices (Group number 1 - Default model)

Total Effects (Group number 1 - Default model)

	TransforLeader	SelfEfficacy
SelfEfficacy	1.121	.000
SocialSupport	.298	.084

Standardized Total Effects (Group number 1 - Default model)

	TransforLeader	SelfEfficacy
SelfEfficacy	.504	.000
SocialSupport	.328	.205

Direct Effects (Group number 1 - Default model)

	TransforLeader	SelfEfficacy
SelfEfficacy	1.121	.000
SocialSupport	.204	.084

Standardized Direct Effects (Group number 1 - Default model)

	TransforLeader	SelfEfficacy
SelfEfficacy	.504	.000
SocialSupport	.224	.205

Indirect Effects (Group number 1 - Default model)

	TransforLeader	SelfEfficacy
SelfEfficacy	.000	.000
SocialSupport	.094	.000

Standardized Indirect Effects (Group number 1 - Default model)

	TransforLeader	SelfEfficacy
SelfEfficacy	.000	.000
SocialSupport	.103	.000

When Table 5.9 is examined, the values in the “Regression Weights: (Group number 1 - Default model)” table reveals that all regression coefficients are significant (p values are less than 0.05). Standardized coefficients of the same values are given in the next table, “Standardized Regression Weights: (Group number 1 - Default model)”. The values in these tables are the results of the model that includes the independent variable and the mediator variable together. However, the values in the “Total Effects (Group number 1 - Default model)”, “Direct Effects (Group number 1 - Default model)” and “Indirect Effects (Group number 1 - Default model)” tables should be examined to determine whether the mediation effect exists.

First, the total effects are examined from the “Total Effects (Group number 1 - Default model)” table or the “Standardized Total Effects (Group number 1 - Default model)” table. Accordingly, the total effect of transformational leadership on the perception of social support is 0.328 as a standardized beta coefficient. By examining the subsequent “Standardized Direct Effect” and “Standardized Indirect Effect” tables, it can be understood that 0.224 of this effect is a direct effect and 0.103 is an indirect effect. The presence of an indirect effect indicates that there is a mediation relationship. However, the main question here is whether this mediating effect is significant.

To understand whether the mediation effect is significant, the “Bias-corrected percentile method” is firstly selected from the “Estimates/Bootstrap” Tab, as shown in Figure 5.20, and “Two-Tailed Significance (BC)” is selected from the “Bias-corrected percentile method” tab from the new submenu that opens, as shown in Figure 5.21. Thus, the degree of significance of the beta

coefficients will be seen in the table that is displayed. The significance of the beta value of 0.103, which is the mediating effect value obtained, is 0.012, as can be seen in Figure 5.21. This value reveals that the mediating effect is significant.

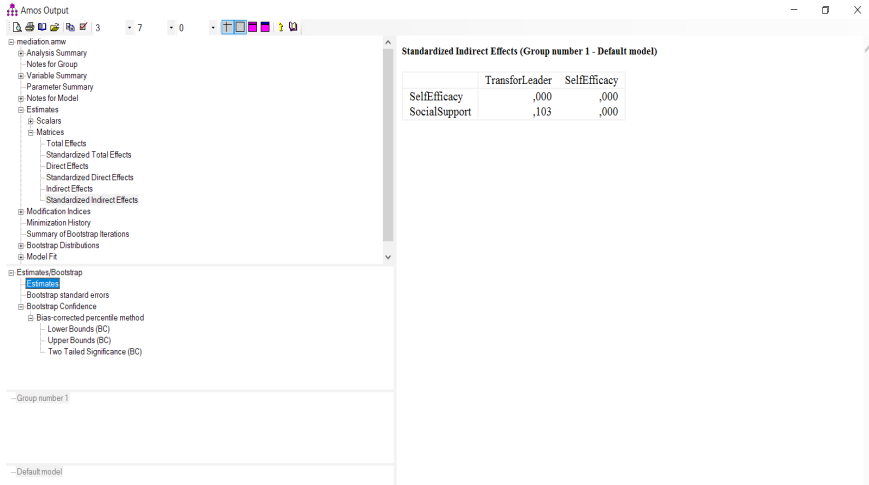


Figure 5.20. Bias-corrected percentile method selection from the Estimates/Bootstrap tab

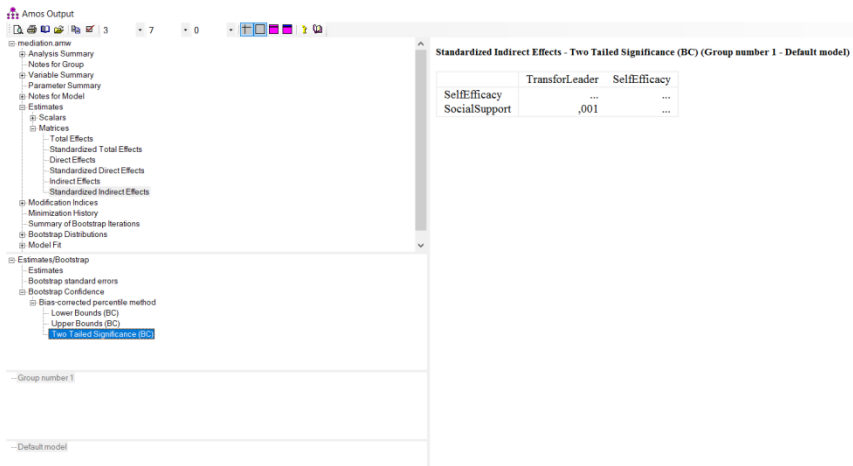


Figure 5.21. Two-Tailed Significance (BC) Selection from the Bias-corrected percentile method tab

The processes described thus far can also be performed using the common Baron and Kenny method or PROCESS macro with SPSS. However, the main advantage of performing a mediation test with the structural equation model is that the relationships between more than one independent, dependent, and

mediator variable could be tested in the same model at the same time. As explained in the relevant section, the effect of more than one mediator variable can be examined simultaneously in the SPSS program with PROCESS Macro. However, dependent and independent variables are one each. In the case of more than one, the test should be performed separately and the obtained values should be evaluated together. However, there is no such requirement in the structural equation model and mediator test with AMOS. Theoretically, the number of dependent, independent, and mediator variables can be as many as desired and they can be tested simultaneously. Certainly, it should be taken into account that as the number of variables included in the model increases, its complexity will increase and its theoretical basis will weaken.

We want to show what we mean on the model in Figure 5.22, which has increased complexity. While the mediator role of self-efficacy in the effect of transformational leadership on the perception of social support was examined in the first model, in this model, in addition to transformational leadership, “leader support” was also considered as an independent variable. The model is presented in Figure 5.22. To test this model with the SPSS PROCESS macro, two separate tests will be required. The drawing of the model and the selection of the relevant options should be performed as described in the sections above. Here, the direct test results will be discussed.

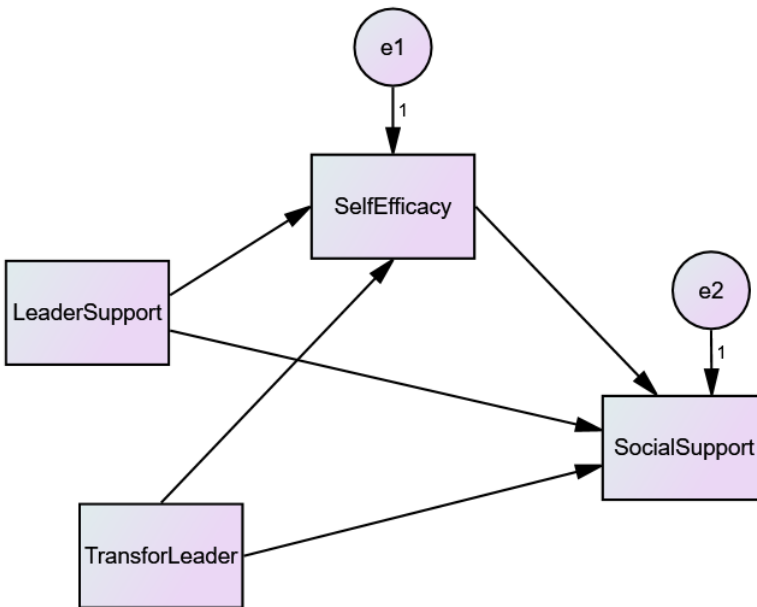


Figure 5.22. AMOS Drawing of the Two Independent Variable Research model

The results obtained by testing the model are presented in Table 5.10. As can be seen from the table, the beta coefficients in all paths are significant. Examining the “Standardized Indirect Effects (Group number 1 - Default model)” table, self-efficacy in the effect of leader support on social support perception was 0.028; on the other hand, there is a mediating effect of 0.079 in the effect of transformational leadership on the perception of social support.

Table 5.10. Test Results of the Two Independent Variable Research Model

Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
SelfEfficacy	<---	TransforLeader	.968	.129	7.484	***	
SelfEfficacy	<---	LeaderSupport	.291	.108	2.692	.007	
SocialSupport	<---	SelfEfficacy	.073	.030	2.428	.015	
SocialSupport	<---	TransforLeader	.152	.064	2.382	.017	
SocialSupport	<---	LeaderSupport	.122	.048	2.522	.012	

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
SelfEfficacy	<---	TransforLeader	.449
SelfEfficacy	<---	LeaderSupport	.161
SocialSupport	<---	SelfEfficacy	.176
SocialSupport	<---	TransforLeader	.171
SocialSupport	<---	LeaderSupport	.164

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
TransforLeader	11.992	1.157	10.368	***	
LeaderSupport	17.125	1.652	10.368	***	
e1	43.153	4.162	10.368	***	
e2	8.329	.803	10.368	***	

Matrices (Group number 1 - Default model)

Total Effects (Group number 1 - Default model)

	LeaderSupport	TransforLeader	SelfEfficacy
SelfEfficacy	.291	.968	.000
SocialSupport	.143	.222	.073

Standardized Total Effects (Group number 1 - Default model)

	LeaderSupport	TransforLeader	SelfEfficacy
SelfEfficacy	.161	.449	.000
SocialSupport	.192	.250	.176

Direct Effects (Group number 1 - Default model)

	LeaderSupport	TransforLeader	SelfEfficacy
SelfEfficacy	.291	.968	.000
SocialSupport	.122	.152	.073

Standardized Direct Effects (Group number 1 - Default model)

	LeaderSupport	TransforLeader	SelfEfficacy
SelfEfficacy	.161	.449	.000
SocialSupport	.164	.171	.176

Indirect Effects (Group number 1 - Default model)

	LeaderSupport	TransforLeader	SelfEfficacy
SelfEfficacy	.000	.000	.000
SocialSupport	.021	.070	.000

Standardized Indirect Effects (Group number 1 - Default model)

	LeaderSupport	TransforLeader	SelfEfficacy
SelfEfficacy	.000	.000	.000
SocialSupport	.028	.079	.000

Again, when “Two-Tailed Significance (BC)” is selected from the “Bias-corrected percentile method” tab, it is understood that these mediator effect values obtained are significant, as can be seen in Figure 5.23. This is because the beta coefficient of the mediating role of self-efficacy in the effect of leader support on the perception of social support is 0.041, and it is 0.013 in the effect of transformative leadership.

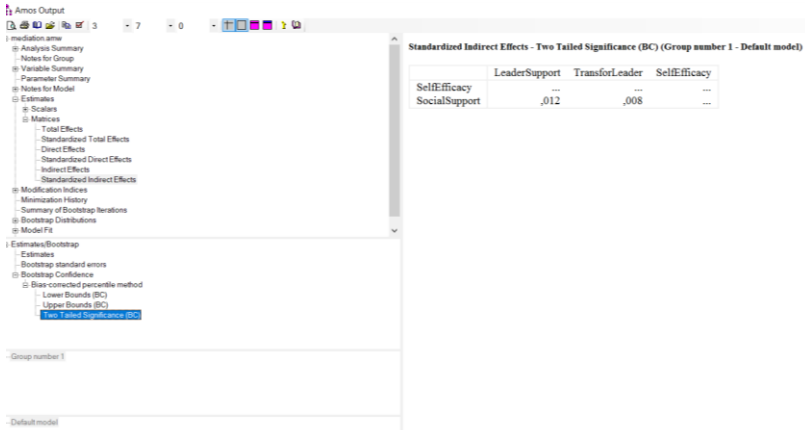


Figure 5.23. Two-Tailed Significance (BC) Selection from the Bias-corrected percentile method tab in the model with two independent variables

Tabulation of the obtained values is performed as described in the PROCESS macro method, which will be explained in the next section. This is because what is performed in PROCESS macro is actually “Bootstrap” analysis and the values are similar. As a result, tables are created using the same values.

ANALYSIS OF THE MEDIATOR VARIABLE EFFECT WITH PROCESS MACRO

Mediation analysis is a popular statistical technique that is often used by researchers to test hypotheses that examine causal effects. It offers researchers a richer understanding of the relationship between the mediating variable and the independent and dependent variable. For this reason, researchers often form hypotheses for the mediator model. Although many methods can be used to test mediator model hypotheses, Baron and Kenny's [5] causal steps method and structural equation modeling (SEM) are the most preferred by researchers.

In the Baron and Kenny [5] method, the coefficients in the mediator model are typically determined using a series of multiple regressions, while in structural equation modeling (SEM), the coefficients in the model are determined by simultaneously creating a structural equation. The causal steps approach popularized by Baron and Kenny [5], and then the Sobel test to check the significance of the mediation, do not yield strong statistical results. SEM analysis, on the other hand, requires a certain level of expertise. Considering the criticisms of the approach advocated by Baron and Kenny [5] [14,15,16,21] and the complexity of SEM, the "Process Macro" method proposed by Hayes, which has gained significant popularity in the statistical mediation literature, is considered more attractive by researchers. In addition to its ease of application, this method is frequently preferred, as it gives strong statistical results. The "Process Macro" method proposed by Hayes is accepted as a contemporary approach in today's literature. In the resampling approach applied in this method, Type I error rates are lower than predicted [22], so it is considered a method that can provide more accurate confidence intervals.

Confidence intervals include the error in the analysis and determine confidence intervals for the analysis result rather than a p coefficient (significance). These limits denoted as the Lower confidence limit (LCL) and Upper confidence limit (UCL) are asymmetrical. Asymmetric Confidence intervals are more accurate because the mediation effect is not always normally distributed. The mediation effect is tested on the basis of whether it is within the confidence interval of zero (0). If it is outside the zero confidence interval, the mediation effect is statistically significant. To understand this subject, it will be useful to examine the table below.

Confidence intervals for the indirect effect from four different studies are given below.

Table 5.11. Confidence intervals

Examples	LLCI	LCLI	Explanation
1	0.512	0.250	Statistically significant
2	-0.154	-0.021	Statistically significant
3	-0.471	0.122	Not statistically significant
4	0.842	-0.022	Not statistically significant

When the table above is examined, it can be seen that the confidence intervals for the indirect effect in examples 1 and 2 do not include zero (0). For this reason, it can be stated that there is a mediating effect in the 1st and 2nd examples. However, it is not possible to discuss the existence of a mediating effect in the 3rd and 4th examples because the confidence intervals contain zero (0).

Another important contribution of the “Process Macro” method is that it tests the strength and magnitude of the indirect effect. The indirect effect refers to the causal hypothesis that an independent variable (X) affects the mediating variable (M) and, accordingly, causes the independent variable (Y) [23]. It is important to note that the indirect effect is determined by multiplying the effect of the independent variable on the mediating variable (path a) and the effect of the mediating variable on the dependent variable (path b) (ab).

As emphasized by many researchers, the indirect effect plays a very important role in detecting mediation [15]. The “Process Macro” method provides a bootstrapping estimate for the indirect effect (ab), an estimated standard error, and 95% confidence intervals for the effective value of ab. The Bootstrapping procedure in the “Process Macro” method allows the estimation of the model coefficient and ab effect for each equation by taking a large number of n-dimensional random samples from the data that defines the model (the data obtained within the scope of the research). This method is a non-parametric approach to effect size estimation and hypothesis testing that makes no assumptions about the sample distribution regarding the shape or statistics of the distributions of the variables [15,21]. This approach has been suggested by researchers as a way of dealing with the power problem posed by asymmetries in the ab sampling distribution and other non-normal forms. If you recall, the Sobel test analyzes based on the assumption that the data are normally distributed, and for this reason, it is not recommended by many researchers [24]. In this context,

bootstrapping gives more reliable results as it makes no assumptions about the sample distribution.

There is no mention of partial or full mediation in the “Process Macro” method. Instead, there is effect size. In the literature, three methods are used to measure the effect size of the mediator variable. What these methods have in common is that they focus on the indirect effect (ab) to measure the mediating effect.

Three methods are generally accepted in the literature:

- (a) Ratio Effect Measures
- (b) R^2 Effect Measures
- (c) Standardized Effect Measures

The above-mentioned effect size measurements may differ from each other in terms of meaning and statistical properties.

(a) Ratio Effect Measure: In this method, proposed by Alwin and Hauser [25] as a way of measuring the size of the mediating effect, the total effect has to be considered the ratio ab/c , which is algebraically equal in the normal least squares regression. The greatest problem encountered in this criterion is where the proportionality value is greater than 1 or has a negative result, and in this case, the total effect is very close to zero or the mediating effect value gives a larger value than it actually is. Generally, this method requires sample groups of 500-1000 to give accurate values, but smaller sample sizes are also sufficient if the direct effect is bigger [26]. In this respect, this method is not generally preferred by researchers in determining the effect size [17].

(b) R^2 Effect Measures: To determine the effect size of the mediating variable, three different values can be taken, namely the part explained by the direct effect, the part that is unexplained and the part explained by the mediation effect. However, a common problem of all values is that effect size measurements primarily work with a single mediator model with consistent effects. The level of accuracy of results in more complex models is a matter of debate [27].

(c) Standardized Effect Measures: Although it is a new approach, the standardized effect measure is frequently preferred by researchers. The standardized indirect effect of the independent variable on the dependent variable is considered as a measure of the mediating effect.

Preacher and Kelley [28] stated that the fully standardized effect value of the mediator variable is as follows:

High effect if it is close to 0.25,

Moderate effect if close to 0.9

Low effect if close to 0.1

Although alternative effect size measures have been proposed for mediation effects, the value of the mediating effect will be evaluated with a fully standardized effect size in the next part of this book.

Let us briefly summarize what has been described thus far:

The “Process Macro” method makes inferences for the indirect effect (ab). It does not test the significance of the indirect effect with the Sobel test, but uses the bootstrap technique, which gives stronger and more valid statistical results.

The bootstrap technique determines upper and lower bound confidence intervals and accepts that mediation is significant if there is no zero within this confidence interval.

There are no concepts of partial or full mediation as in the Baron and Kenny [5] method. Instead, the mediator is in the reliable range and the mediator has an effect size.

As in Baron and Kenny’s [5] method, the coefficients in the whole model are determined in a single analysis instead of an ordered regression analysis. This prevents the occurrence of Type 1 errors in the analyses.

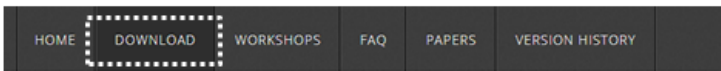
Finally, as in Baron and Kenny’s [5] method, it does not offer prerequisites for talking about mediation. If you remember, in their method, the mediation effect can only be mentioned if the three prerequisites are met (the total effect is statistically significant (path c), the independent and dependent variable (path a), and the effect of the mediator variable on the independent variable is statistically significant when the dependent variable is under control (path b)).

Even if these prerequisites are not met in the contemporary approach, Process Macro mediation can be mentioned.

Downloading the Process Macro

Analyzing the mediation effect is easy with the Process Macro proposed by Andrew F. Hayes. PROCESS is an easy-to-use plugin to SPSS or SAS that can be used to predict mediation models by multiple regression or logistic regression. Before proceeding to the application, readers must download the Process Macro application as an add-on to the SPSS program. This program, which can be downloaded free of charge from <http://processmacro.org>, is constantly being developed and there are minor differences in each version. At the time of publishing this book, the latest version is Process v3.4. All explanations in the book will be made according to this version. After the plug-in is downloaded to the SPSS program, it can be switched to the SPSS application.

Downloading the plugin:



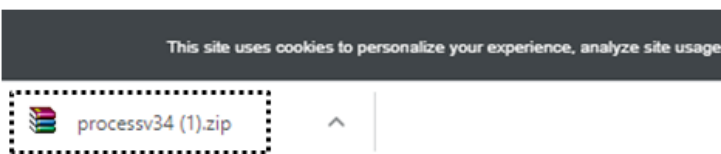
The PROCESS macro for SPSS, SAS, and R

The <http://processmacro.org> address is accessed and the "DOWNLOAD" tab is clicked

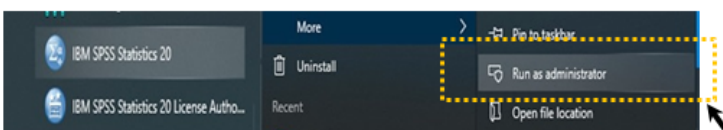
ATTENTION: MacOS "Catalina" users: There is a bug in the latest release of MacOS related to file access permissions that also affects SPSS. It has nothing to do with PROCESS or its operation on the Mac or SPSS. Consult your local tech support person for advice. [Here](#) is a video that might be helpful in solving your problem.



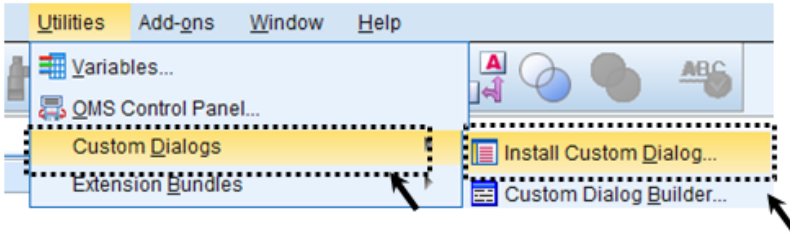
Click on the "Download PROCESS v3.4" tab in the middle of the opened page.




When the PROCESS v3.4 version is installed, the Zip file is displayed at the bottom left of the window. The downloaded zip file is extracted to the folder.

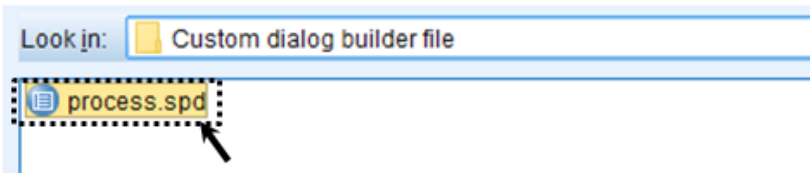


SPSS program is run as administrator.

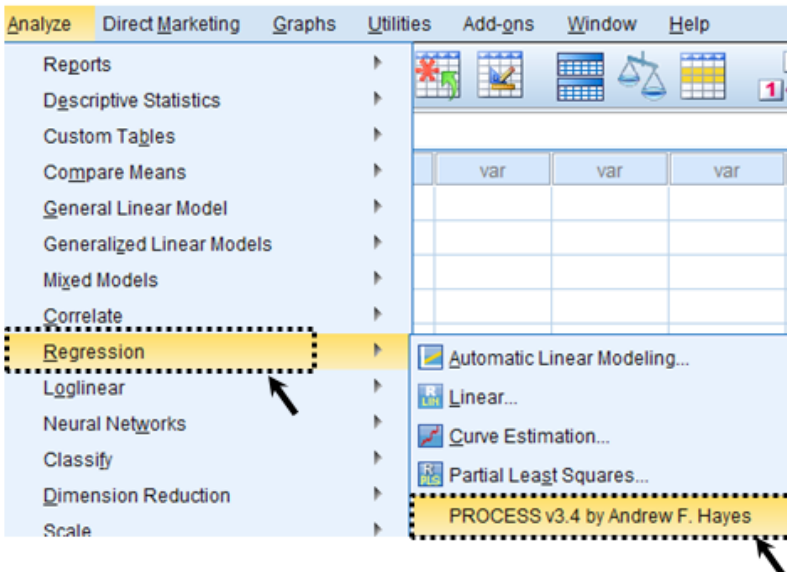


In the SPSS program "Utilities/CustomDialogs/
The "Install Custom Dialog" option is checked.

 Open a Dialog Specification



In the window that opens, select the "process.spd" file in the Zip file that was extracted to the folder and press the "OK" button.



When the installation process is completed successfully, the "Analyze/Regression/PROCESS v3.4 by Andrew F. Hayes" tab will appear on the screen.

Figure 5.12. Images of installing Process Macro

MEDIATION ANALYSIS WITH SPSS

As the readers will remember, for both the causal steps approach proposed by Baron and Kenny [5] and the structural equation modeling for the effect of the mediator variable, the research question was: “Does self-efficacy have a mediating role in the effect of the transformational leadership on the perception of social support?” We will continue our analysis using the same research question.

First of all, we open the “Data -3.sav” file (You can access this file at www.indataanalysis.com).

Let’s recall the model for our research question:

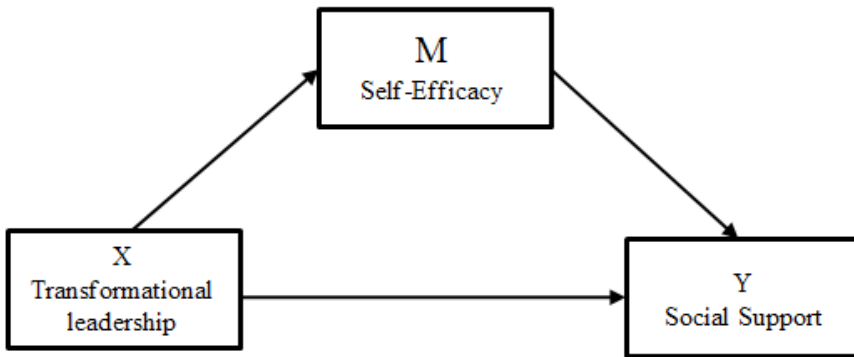


Figure 5.13. Research Model

We will analyze the mediation effect with the Process Macro proposed by Andrew F. Hayes. For this, the following steps will be followed.

“Analyze ---> Regression ----> PROCESS v3.4 by Andrew F. Hayes” option is clicked.

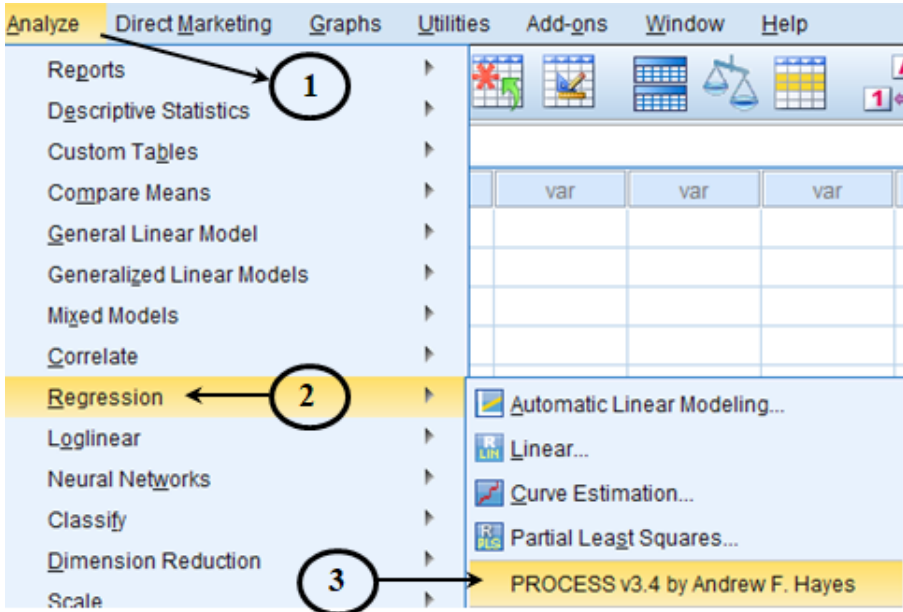


Figure 5.14. SPSS “Analyze” Screen

In the window that opens, we assign the relevant variables.

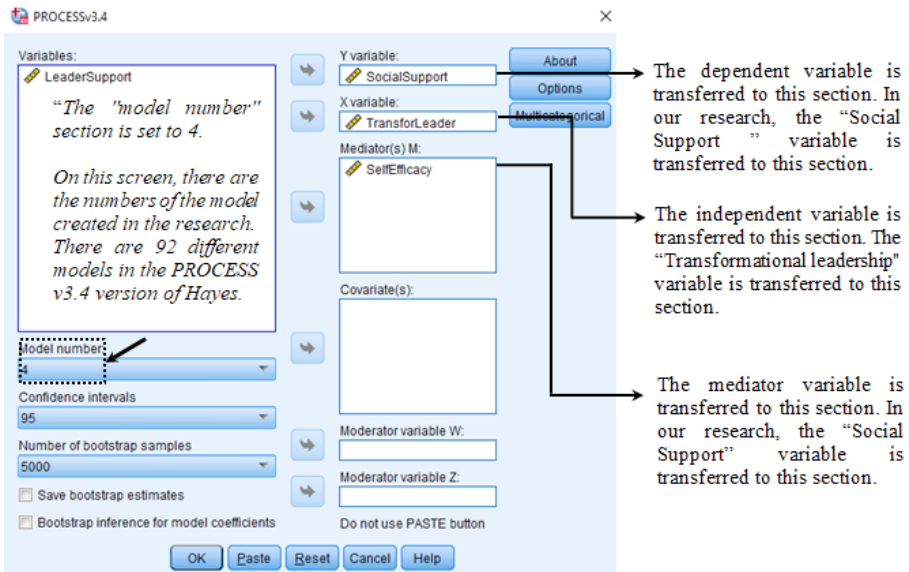


Figure 5.15. SPSS “Process Macro” Screen

After all processes are completed, the “Options” tab is checked.

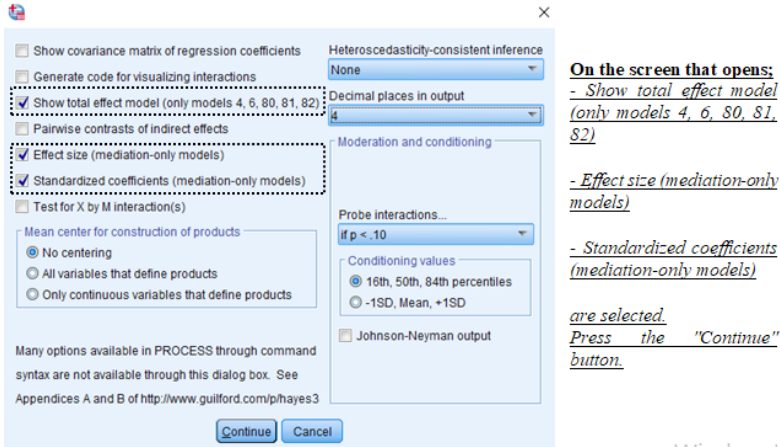


Figure 5.16. SPSS “Process Macro” Screen

Because the bootstrap method is used to calculate standard errors and confidence intervals, it may take some time for the outputs to occur. The interpretation and reporting of the outputs after they appear on the screen are explained separately in each table below.

When the above outputs are examined, it is seen that the independent variable, transformational leadership, affects the mediator variable, namely self-efficacy, significantly and positively ($\beta = 1.1213$, 95% CI [LLCI= 0.8622, ULCI=1.3803], $t = 8.5319$, $p < 0.001$). It is concluded that the effect is significant because the confidence interval for the effect (LLCI= .8622, ULCI=1.3803) does not include zero (0). Transformational leadership explains 25.38% of the variance in the change in self-efficacy in employees ($R^2 = 0.2583$).

Table 5.12. Results showing the effect of the independent variable (Transformational Leader) on the mediating variable (Self-Efficacy)

```

*****
Model : 4 -----> Model number of the study
  Y : SocialSu
  X : Transfor
  M : SelfEffi
  }
  -----> Variables included in the research

Sample
Size: 216 -----> Sample Size

*****
OUTCOME VARIABLE:
SelfEffi

Model Summary
  R          R-sq.      MSE      F      df1      df2      p
,5038      ,2538      44,7410  72,7931  1,0000  214,0000  ,0000

  Explained variance
  Significance of the model

Model
  Non-standardized beta coefficient
  Standardized beta coefficient

  coeff      se      t      p      LLCI      ULCI
constant  22,1136  1,9800  11,1687  ,0000  18,2109  26,0164
Transfor  1,1213   ,1314   8,5319  ,0000  ,8622   1,3803
  Path a

Standardized coefficients
  coeff
Transfor  ,5038 -----> Standardized beta coefficient

  The effect of X on M is significant
  95% CI, confidence interval
  
```

When the above outputs are examined, it is seen that the independent variable, transformational leadership, has a significant and positive effect on the dependent variable, namely the perception of social support ($\beta = 0.2036$, 95% CI [LLCI= 0.0718, ULCI=0.3355], $t = 3.0441$, $p < .05$). It is concluded that the effect is significant because the confidence interval for the effect (LLCI=0.0718, ULCI=0.3355) does not include zero (0). In addition, the results on the effect of self-efficacy, which is the mediating variable, on the perception of social support, which is the independent variable, are also seen in the same output. The findings reveal that the effect of self-efficacy on the perception of social support is significant and positive ($\beta = 0.0838$, 95% CI [LLCI= 0.0246, ULCI=0.1431], $t = 2.788$, $p < .05$). Transformational leadership and self-efficacy explain 13.87% of the variance of the change in perception of social support in employees ($R^2 = 0.1387$).

Table 5.13. Results showing the effect of the independent variable (Transformational Leadership) and mediator variable (Self-Efficacy) on the dependent variable (Perception of Social Support) (path b and c’):

OUTCOME VARIABLE:
SocialSu

Model Summary		Explained variance		MSE	F	df1	df2	p
R	R-sq							
,3724	,1387			8,6495	17,1529	2,0000	213,0000	,0000

Model	coeff	se	t	p	LLCI	ULCI
constant	8,2750	1,0953	7,5552	,0000	6,1160	10,4340
Transfor	,2036	,0669	3,0441	,0026	,0718	,3355
SelfEffi	,0838	,0301	2,7888	,0058	,0246	,1431

Standardized coefficients

Model	coeff
Transfor	,2241
SelfEffi	,2053

Annotations:
 - Path b: points to the coefficient for Transfor in the non-standardized table.
 - Path c: points to the coefficient for SelfEffi in the non-standardized table.
 - Non-standardized beta coefficient: points to the 'coeff' column in the non-standardized table.
 - Standardized beta coefficient: points to the 'coeff' column in the standardized table.
 - Indicates the significance of the effect: points to the 'p' column in the non-standardized table.
 - 95% CI, confidence interval: points to the LLCI and ULCI columns in the non-standardized table.

Table 5.14. Results showing the Total effect of the independent variable (Transformative Leader) on the dependent variable (Perception of Social Support) (path c):

```
***** TOTAL EFFECT MODEL *****
OUTCOME VARIABLE:
SocialSu

Model Summary
      R      R-sq      MSE      F      df1      df2      p
,3275  ,1073  8,9234  25,7139  1,0000  214,0000  ,0000

Model
      coeff      se      t      p      LLCI      ULCI
constant  10,1286  ,8842  11,4546  ,0000  8,3857  11,8716
Transfor  ,2976  ,0587  5,0709  ,0000  ,1819  ,4133

Standardized coefficients
      coeff
Transfor  ,3275

Annotations:
- R-sq (,1073) is circled and labeled 'Explained variance'.
- Path c (,2976) is circled and labeled 'Path c'.
- Non-standardized beta coefficient (,2976) is circled and labeled 'Non-standardized beta coefficient'.
- Standardized beta coefficient (,3275) is circled and labeled 'Standardized beta coefficient'.
- p-value (,0000) is circled and labeled 'Indicates the significance of the effect'.
- LLCI (,1819) and ULCI (,4133) are circled and labeled '95% CI, confidence interval'.
```

When the above outputs are examined, it is seen that the total effect of the independent variable, transformational leadership, on the dependent variable, that is, the perception of social support, is significant and positive ($\beta = 0.2976$, 95% CI [LLCI= 0.1819, ULCI=0.4133], $t = 5.0709$, $p < 0.001$). It is concluded that the effect is significant since the confidence interval for the effect (LLCI= 0.1819, ULCI=0.4133) does not include zero (0). Transformational leadership explains 10.73% of the variance of the change in the perception of social support in employees ($R^2 = 0.1073$).

Table 5.15. Results showing the direct effect, indirect effect and total effect of the independent variable (Transformational Leadership) on the dependent variable (Perception of Social Support) (paths c', ab and c):

```

***** TOTAL, DIRECT, AND INDIRECT EFFECTS OF X ON Y *****

Total effect of X on Y
  Effect      se      t      p      LLCI      ULCI      c_ps      c_cs
  -----
  ,2976      ,0587      5,0709      ,0000      ,1819      ,4133      ,0944      ,3275
  -----
  Path c

Direct effect of X on Y
  Effect      se      t      p      LLCI      ULCI      c'_ps      c'_cs
  -----
  ,2036      ,0669      3,0441      ,0026      ,0718      ,3355      ,0646      ,2241
  -----
  Path c'

Indirect effect(s) of X on Y:
SelfEffi      Effect      BootSE      BootLLCI      BootULCI      Confidence interval for indirect effect
  -----
  ,0940      ,0373      ,0267      ,1737

Partially standardized indirect effect(s) of X on Y:
SelfEffi      Effect      BootSE      BootLLCI      BootULCI
  -----
  ,0298      ,0117      ,0086      ,0544

Completely standardized indirect effect(s) of X on Y:
SelfEffi      Effect      BootSE      BootLLCI      BootULCI
  -----
  ,1034      ,0404      ,0297      ,1901
  -----
  Effect Size

***** ANALYSIS NOTES AND ERRORS *****

Level of confidence for all confidence intervals in output:
  95,0000
  -----
  Indicates the confidence interval

Number of bootstrap samples for percentile bootstrap confidence intervals:
  5000
  -----
  Number of resamples (Bootstrap)
    
```

The outputs above show the data with which the research hypothesis was tested as well as the results. The decision about whether self-efficacy mediates the effect of transformational leadership on the perception of social support is made according to the results of the indirect effect. Accordingly, the indirect effect of transformational leadership on the perception of social support is $\beta = 0.094$ (Indirect effect) and the confidence interval does not include zero (0) (BootLLCI= 0.0255, BootULCI=0.1733), indicating that self-efficacy plays a mediating role in the effect of transformational leadership on the perception of social support. The mediation effect is close to the low effect size ($R^2 = 0.1034$). In the light of these findings, the research hypothesis was supported.

Note: Preacher and Kelley [28] stated that if the standardized effect size of the mediator variable is close to 0.25, there is a high effect, if it is close to 0.9, there is a medium effect, and if it is close to 0.1, there is a low effect.

REPORTING RESEARCH FINDINGS

The causal steps approach proposed by Baron and Kenny [5] is commonly used in testing mediating hypotheses. However, the extant empirical literature indicates that this approach does not give strong enough statistical values and has many potential shortcomings [7,8,12,14]. For this reason, “Process Macro” software, which is a contemporary approach developed by Hayes [8] as an add-on to SPSS, was used to test the research hypotheses. In the Process Macro method, 5000 resampling options were preferred with the bootstrap technique. The regression results are presented in Table 5.16.

Table 5.16. Regression Results

Effects	Standardized			
	Beta	SE	LLCI - ULLC	
TL → Social Support	.2036**	.0669	.0718,	.3355
TL → Self-Efficacy	1.1213***	.1314	.8622,	1.3803
Self-Efficacy → Social Support	.0838**	.0301	.0246,	.1431
Indirect Effect (TL → SE → SS)	.0940***	.0370	.0255,	.1733

Note: TL; Transformational Leadership, SE; Self-Efficacy, SS; Perception of Social Support

When Table 5.16 is examined, it can be seen that transformational leadership ($\beta = 0.2036^{**}$, 95% CI= [0.0718, 0.3355], $t=3.0441$, $p < 0.05$) and self-efficacy ($\beta = 0.0838^{**}$, 95% CI= [0.0246, 0.1431]), $t=2.7888$, $p < 0.05$) seem to affect the perception of social support in a significant and positive way. In addition to these findings, it is determined that the indirect effect of transformational leadership on the perception of social support is significant, and therefore, self-efficacy plays a mediation role in the relationship between transformational leadership and social support ($\beta = 0.0940^{***}$, SE = 0.0370, $p < 0.05$, 95%). BCA CI = [0.0255, 0.1733]. The standardized effect size of the mediation is 0.1034. Accordingly, the mediation effect in the tested model is close to the low effect value [28], which supports the research hypothesis.

References

- [1] Lipsey, M. W. (1993). *Theory as method: Small theories of treatments*. In L. B. Sechrest & A. G. Scott (Eds.), *Understanding causes and generalizing about them: New directions for program evaluation*, San Francisco: Jossey-Bass.
- [2] MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, 55, 95-107.
- [3] Mathieu, J. E., DeShon, R. P., & Bergh, D. D. (2008). Mediation inferences in organizational research; Then, now, and beyond. *Organizational Research Methods*, 11, 203-223.
- [4] Ndofor, H. A., Sirmon, D. G., & He, X. (2011). Firm resources, competitive actions and performance: Investigating a mediated model with evidence from in-vitro diagnostics industry. *Strategic Management Journal*, 32, 640-657.
- [5] 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(6), 1173.
- [6] Iacobucci, D. (2008), *Mediation Analysis*, Thousand Oaks, CA: Sage.
- [7] Zhao, X., Lynch Jr, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of consumer research*, 37(2), 197-206.
- [8] Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.
- [9] Iacobucci, D., Saldanha, N., & Deng, X. (2007). A meditation on mediation: Evidence that structural equations models perform better than regressions. *Journal of Consumer Psychology*, 17(2), 139-153.
- [10] Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, 13, 290-312.
- [11] Krause, M. R., Serlin, R. C., Ward, S. E., Rony, R. Y. Z., Ezenwa, M. O., & Naab, F. (2010). Testing mediation in nursing research: Beyond Baron and Kenny. *Nursing Research*, 59(4), 288.
- [12] Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. *The handbook of Social Psychology*, 1(4), 233-265.
- [13] Aguinis, H., Edwards, J. R., & Bradley, K. J. (2017). Improving our understanding of moderation and mediation in strategic management research. *Organizational Research Methods*, 20(4), 665-685.

- [14] 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(1), 83.
- [15] Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, 36(4), 717-731.
- [16] 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(3), 879-891.
- [17] MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention science*, 1(4), 173-181.
- [18] Byrne, B. M., (2016). *Structural Equation Modelling with AMOS: Basic Concepts, Applications, and Programming*, 3rd ed. Routledge, New York.
- [19] Tabachnick, B.G. and Fidell, L.S. (2013), *Using multivariate statistics* (6th ed.), Boston: Pearson Education, Inc.
- [20] Kelloway, E. K. (1998). *Using LISREL for structural equation modeling: A researcher's guide*. Sage.
- [21] Hayes, A. F., & Preacher, K. J. (2010). Quantifying and testing indirect effects in simple mediation models when the constituent paths are nonlinear. *Multivariate Behavioral Research*, 45(4), 627-660.
- [22] MacKinnon, D. P., Lockwood, C. M., & Williams, J. (2004). Confidence limits for the indirect effect: Distribution of the product and resampling methods. *Multivariate Behavioral Research*, 39(1), 99-128.
- [23] Sobel, M. E. (1990). Effect analysis and causation in linear structural equation models. *Psychometrika*, 55(3), 495-515.
- [24] Coutts, J. J., Hayes, A. F., & Jiang, T. (2019). Easy statistical mediation analysis with distinguishable dyadic data. *Journal of Communication*, 69(6), 612-649.
- [25] Alwin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review*, 37-47.
- [26] MacKinnon, D. P., Warsi, G., & Dwyer, J. H. (1995). A simulation study of mediated effect measures. *Multivariate Behavioral Research*, 30(1), 41-62.
- [27] MacKinnon, D. (2012). *Introduction to statistical mediation analysis*. Routledge.
- [28] Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: quantitative strategies for communicating indirect effects. *Psychological Methods*, 16(2), 93.

CHAPTER 6

ANALYSIS OF THE MODERATOR VARIABLE EFFECT

Introduction

Statistically, it can be said that the moderator variable is synonymous with “interaction” because the interaction between the independent and the dependent variable changes through the moderator variable. Therefore, the effect of the independent variable on the dependent variable depends on the power/magnitude of the moderator variable [1]. Based on this information, it can be said that the independent variable has an effect on the dependent variable, although this effect varies according to the size of the moderator variable. For example, consider the relationship between workload (X), job stress (Y), and leader support (W). According to one hypothesis for the moderator variable, the workload of the employees causes work stress and there is a causal relationship between the two variables. On the other hand, it is expected that an employee with high leader support will have low job stress, whereas an employee with low leader support will have high job stress. In other words, the effect of workload on work stress varies according to the level of leader support. As seen in this example, leader support has a moderating effect on the relationship between workload and work stress.

You may remember that the hypotheses for the mediator variable were based on the assumption that the independent variable affected the mediator variable, and as a result, the dependent variable was affected. In the example given above, we do not assume that workload affects leader support, and as a result, job stress is affected. Instead, we hypothesize that there is a direct causal relationship between workload and stress, and that leader support modifies the strength of the workload-stress relationship. Therefore, leader support is not a mediator but a moderator of the relationship between workload and work stress. As can be seen, this moderator variable and the mediator variable are different concepts. The purpose of this introduction is to state as clearly as possible the conceptual distinction between the mediator and moderator variables. Thus,

researchers will be able to easily determine whether the third variable is a mediator or moderator variable.

An example of a model created for the moderator variable is presented below:

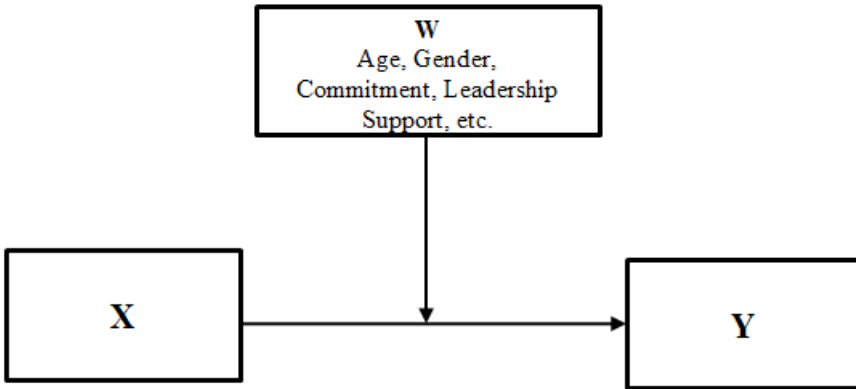


Figure 6.1. Simple Moderation Model

Although the model established here states that the relationship between two variables differs according to the level of a third variable, it is also possible to construct higher-level interactive models that include more than one moderator variable (models that include two or more moderator variables). However, to preserve simplicity in our book, this will be explained through simple models.

As seen in the model above, the moderator variable is a qualitative (e.g., gender, race, class, income level) or quantitative (e.g., commitment, leader support) variable that affects the direction and/or strength of the relationship between the dependent and independent variables. The moderator variable can strengthen, weaken, or change the direction of this relationship. In addition, the moderator variable specifies when or under what conditions a particular effect may change. With the inclusion of the moderator variable in the model, researchers have the opportunity to further investigate the relationship between the independent and dependent variables.

The moderator variable could be continuous (age, income, height, etc.) or categorical (gender, marital status, education, etc.). Demographic variables (gender, marital status, education level, etc.) are examined as moderator variables in social science research. Before proceeding to the example, it would be useful to give information about the statistical measurements of the moderator effect.

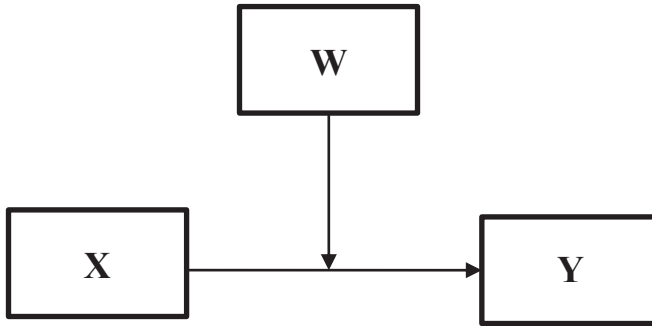


Figure 6.2. Simple Moderation Model

As will be remembered, the theoretical model for the moderator variable is presented above. The statistical representation of this theoretical model is presented below.

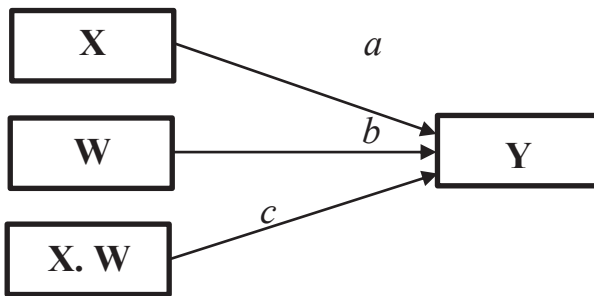


Figure 6.3. Statistical Representation

When the statistical representation is examined, it is seen that three variables explain the dependent variable (Y). These are the independent variable (X), the moderator variable (W), and the interaction variables ($X.W$) formed by the multiplicative effect of the dependent variable and the moderator variable. The regression equation describing the independent variable is as follows:

$$Y = i + \beta_a X + \beta_b W + \beta_c X.W + e$$

The interaction variable ($X.W$) is created by the multiplicative effect of the dependent variable (X) and the moderator variable (W). The interaction variable is important in detecting the existence of the moderator variable. If the effect of the interaction variable on the dependent variable is significant, it can be said to have a moderation effect. In this case, the effect of the independent variable on the dependent variable will be shaped in the presence of the moderator variable. In other words, as the moderator variable (W) takes different values (low, medium,

high), the effect of the independent variable on the dependent variable will also change according to these values. The important thing here is that the values of the dependent variable (X) and the mediator variable (W) should be centralized or standardized (z score) to reduce multicollinearity and facilitate interpretation [1]. Centralization can be performed using the scale function, which subtracts the mean of a variable from each value of that variable. Standardization is performed by subtracting the mean from the value of the variable and dividing the obtained value by the standard error. In the absence of standardization or centralization, the results will not be valid. It is important to note that categorical data does not need to be centralized or standardized (z score).

The test stages of the model created for the moderator variable consist of the following four basic steps:

- The values of the independent variable (X) and the moderator variable (W) are centralized or standardized (z score).
- The interactional variable ($X.W$) is created by the multiplicative effect of the independent variable (X) and the moderator variable (W).
- The effect of the interactional variable on the dependent variable is controlled by performing analyses. If the effect is significant, it can be mentioned that there is a moderator variable.
- If the moderation effect is significant, slope analyses are performed and the results are reported graphically.

Several programs such as SPSS, AMOS, and LISREL can be used to test models for the moderator variable. In our book, the moderation effect analyses performed in SPSS (linear regression and PROCESS Macro) and AMOS programs will be explained. Researchers can choose any of these three methods.

MODERATION EFFECT ANALYSIS WITH REGRESSION

To confirm the existence of a third variable that has a moderator effect on the relationship between the two variables, we should test whether the nature of this relationship changes as the values of the moderator variable change. This is performed by adding an interaction effect to the model and checking whether such an interaction is indeed significant. Linear regression is the most common way to test moderation. Many researchers prefer this method because it is simple and understandable.

The procedures are carried out in the order specified below and the findings are reported:

1. First, the values of the independent variable (X) and moderator variable (W) need to be centralized or standardized (z score) to facilitate interpretation and reduce the problem of multicollinearity. Categorical data do not need to be centralized or standardized (z-score), but categorical data should be puppeted.

2. The interactional variable (X.W) is formed by the multiplicative effect of the independent variable (X) and the moderator variable (W).

3. The effect of the interactional variable on the dependent variable is checked by performing analyses. If the effect is significant, the presence of a moderator variable could be mentioned.

4. If the moderation effect is significant, slope analyses are performed and the results are reported graphically.

If the independent and moderator variable are not significant with the addition of the interaction term, full modulation has occurred. If the independent and moderator variable are significant with the added interaction term, then the modulation effect has occurred, but the main effects are also significant.

MODERATION ANALYSIS WITH SPSS

Research hypothesis: Does leader support have a moderator role in the effect of self-efficacy on the perception of social support?

Our research model is presented below:

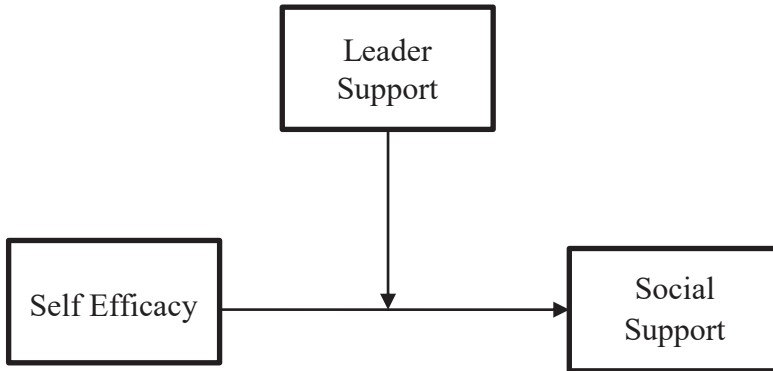


Figure 6.4. Research Model

The statistical representation of our research hypothesis is presented below:

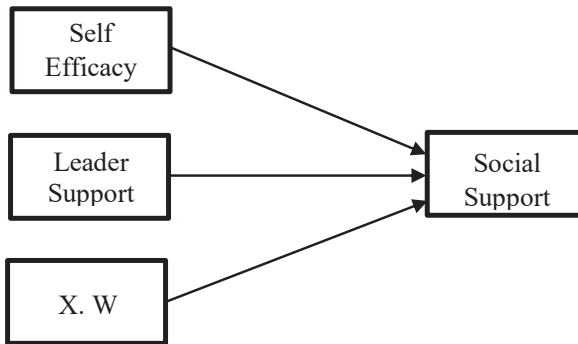


Figure 6.5. Research Model Statistical Representation

Let's briefly recall the procedures that need to be performed so that we can our research hypothesis:

- First, we will centralize or standardize (z score) the values of the dependent variable (Self-Efficacy) and moderator variable (Leader Support) to facilitate interpretation and reduce the problem of multicollinearity.

- In the second stage, we will create an interaction variable (X.W) with the multiplicative effect of the dependent variable (Self-Efficacy) and the moderator variable (Leader Support).

- We will control the effect of the interactional variable on the dependent variable (Perception of Social Support) by performing the analyses. If the effect is significant, we will validate our hypothesis and report the findings.

For regression analysis, we first open the “Data-3.sav” file (You can access this file from www.indataanalysis.com).

1- Before proceeding to hierarchical regression analysis, we will first standardize (z score) the values of the dependent variable (Self-Efficacy) and the moderator variable (Leader Support).

We follow the steps below in the “Data-3.sav” file that we have opened.:
Analyze ---> Descriptive Statistics ----> Descriptives....

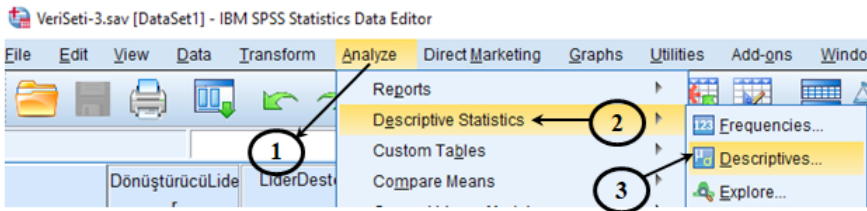


Figure 6.6. SPSS “Analyze” screen

On the screen that opens, the dependent variable and the moderator variable are transferred to the “Variable(s)” box, the “Save standardized values as variables” box is checked, and finally, the “OK” button is pressed.

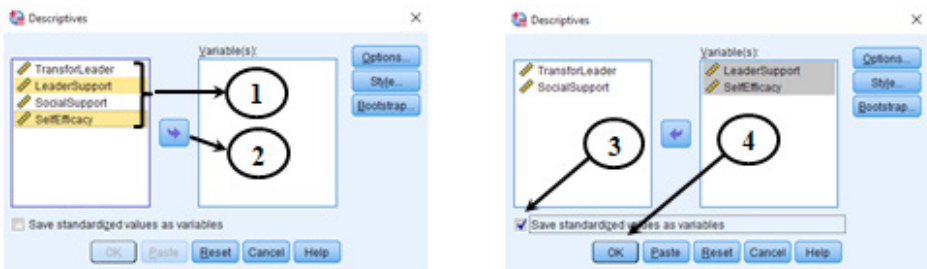


Figure 6.7. “Descriptive” screen

After completing this process, the following screen will be displayed in the “data view” window of SPSS.

	TransforLeader	LeaderSupport	SocialSupport	SelfEfficacy	ZLeaderSupport	ZSelfEfficacy
1	17	20	16	4	,45960	,57549
2	17	20	15	4	,45960	,57549
3	17	14	16	4	-,98691	,31660
4	20	11	12	5	-1,71017	1,48160
5	8	10	13	2	-1,95126	-1,75450
6	19	13	7	4	-1,22800	,83438
7	8	19	12	5	,21851	1,48160
8	16	17	15	4	-,26366	,18716
9	20	10	10	5	-1,95126	1,48160
10	16	15	13	4	-,74583	1,35215

Figure 6.8. “Data View” screen

As can be observed here, the values of the dependent variable (Self-Efficacy) and the moderator variable (Leader Support) are standardized (z score). We will perform our subsequent analysis with these data.

2- In the second step, we will create an interaction variable (X.W) with the multiplicative effect of the independent variable (Self-Efficacy) and the moderator variable (Leader Support).

Transform ---> Compute Variable....

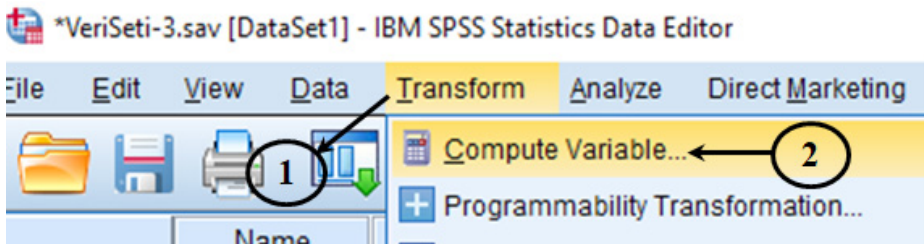


Figure 6.9. SPSS “Transform” screen

On the screen that opens, the name to be given to the interaction variable is denoted in the “Target Variable” section, the dependent variable and the moderator variable are transferred to the “Numeric Expression” box (with a * cross between the two variables), and finally, the “OK” button is clicked.

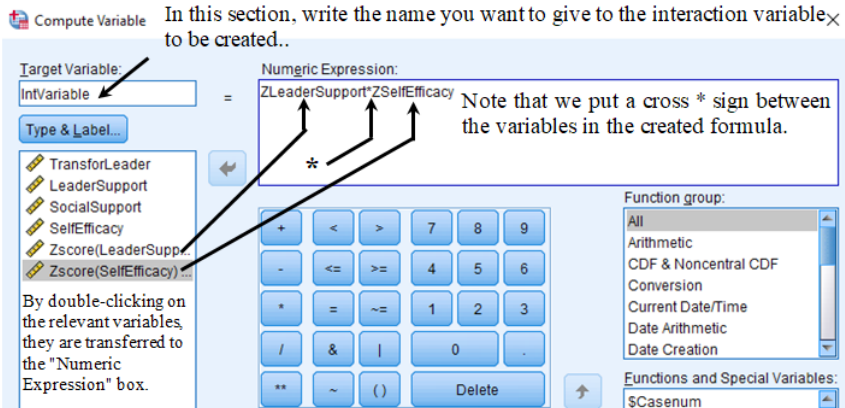


Figure 6.10. "Compute Variable" screen

After completing this process, the following screen will appear in the "Data View" window of SPSS.

	TransforLeader	LeaderSupport	SocialSupport	SelfEfficacy	ZLeaderSupport	ZSelfEfficacy	IntVariable
1	17	20	16	43	.45960	.57549	.26
2	17	20	15	43	.45960	.57549	.26
3	17	14	16	41	-.98691	.31660	-.31
4	20	11	12	50	-1.71017	1.48160	-2.53
5	8	10	13	25	-1.95126	-1.75450	3.42
6	19	13	7	45	-1.22800	.83430	-1.02
7	8	19	12	50	.21851	1.48160	.32
8	16	17	15	40	-.26366	.18716	-.05
9	20	10	10	50	-1.95126	1.48160	-2.89
10	16	15	13	49	-.74583	1.35215	-1.01
11	17	22	20	40	.94177	.18716	.18
12	19	15	16	43	-.74583	.57549	-.43

Figure 6.11. "Data View" screen

Note that a new variable has been created here. We will perform our next analysis with these data.

3- By performing the analysis, we will control the effect of the interaction variable on the independent variable. If the effect is significant, our hypothesis will be validated.

Analyze ---> Regression ---> Linear . . .

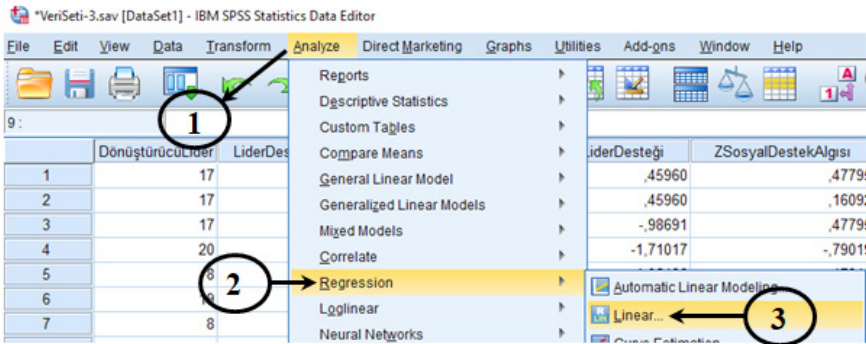


Figure 6.12. SPSS “Analyze” screen

On the screen that opens, we perform our process consisting of three stages:

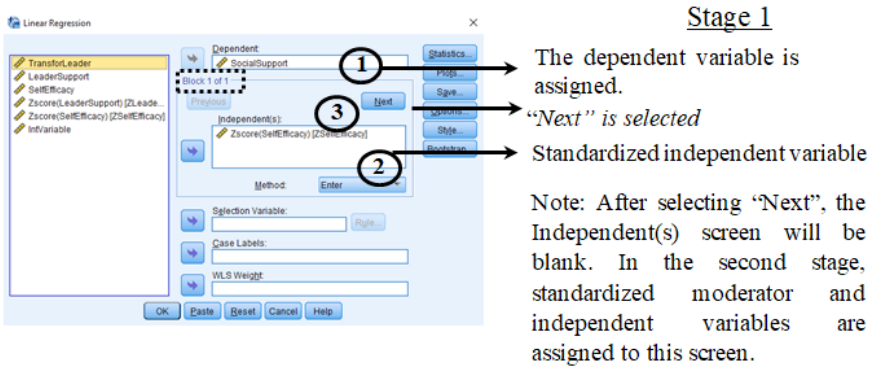


Figure 6.13. “Linear Regression” screen

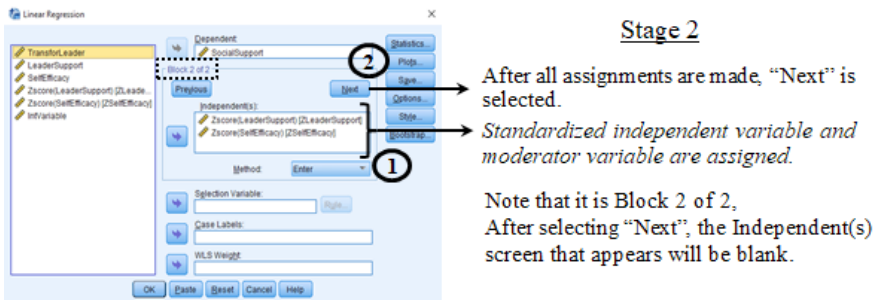
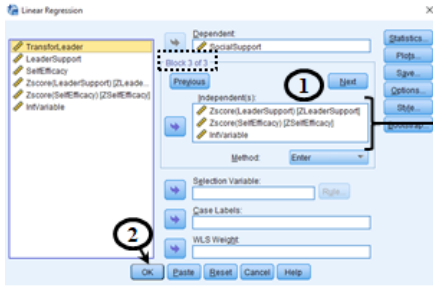


Figure 6.14. “Linear Regression” screen



Step 3

In the third stage, all variables, including the interaction variable, are included in the model and the assignment of the variables is made.

Note that it is Block 3 of 3,

“OK” is selected and SPSS outputs are interpreted.

Figure 6.15. “Linear Regression” screen

SPSS outputs are as follows.

Table 6.6. Model Summary Table

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,318 ^a	,101	,097	2,997
2	,375 ^b	,140	,132	2,938
3	,439 ^c	,193	,181	2,854

a. Predictors: (Constant), Zscore(SelfEfficacy)

b. Predictors: (Constant), Zscore(SelfEfficacy), Zscore (LeaderSupport)

c. Predictors: (Constant), Zscore(SelfEfficacy), Zscore (LeaderSupport), IntVariable

This table provides summary information about hierarchical regression analysis:

Model 1: If you remember, in the first step of the hierarchical regression analysis, we checked for the effect of self-efficacy on the perception of social support. The results of Model 1 are for this analysis. Model 1 shows that the effect of self-efficacy on the dependent variable (perception of social support) is $\beta=0.097$ and explains 10.1% of the change in the dependent variable (R^2 value is 0.101).

Model 2: In the second step of the hierarchical regression analysis, the effect of self-efficacy and leader support on the perception of social support was examined. The results of Model 2 are for this analysis. Model 2 shows that the effect of self-efficacy and leader support on the dependent variable (perception of social support) is $\beta=0.132$ and explains 14% of the change in the dependent

variable (R^2 value is 0.140). Compared to the previous model (Model 1), Leader support seems to contribute an additional 3.9% in explaining the change in the dependent variable.

Model 3: In the third step of the hierarchical regression analysis, we included an interactional variable in the analysis. Thus, in Model 3, the effects of the independent variable (self-efficacy), the regulatory variable (leader support), and the interaction variable [$Zscore(Leader\ Support)*Zscore(OzEfficacy)$] on the dependent variable (perception of social support) were examined. The analysis results show that this effect is $\beta=0.181$ and explains 19.3% of the change in the dependent variable (R^2 value is 0.193). Compared to the previous model (Model 2), it is seen that the interaction variable contributes an additional 3.9% to the explanation of the change in the dependent variable. Model 3 is the most important part of this table, as it shows how much of the variance in the dependent variable is explained by the interaction variable. Thus, in our analysis, the first condition for mentioning the moderation effect is that the interaction variable explains the variance in the dependent variable. Next, the significance of the effect of the moderator variable should be examined. The next table (Table 6.7) provides information to researchers on this subject.

Table 6.7. ANOVA Table

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	216,574	1	216,574	24,108	,000 ^b
	Residual	1922,487	214	8,984		
	Total	2139,060	215			
2	Regression	300,034	2	150,017	17,375	,000 ^c
	Residual	1839,026	213	8,634		
	Total	2139,060	215			
3	Regression	411,896	3	137,299	16,853	,000 ^d
	Residual	1727,164	212	8,147		
	Total	2139,060	215			

a. Dependent Variable: SocialSupport

b. Predictors: (Constant), Zscore(SelfEfficacy)

c. Predictors: (Constant), Zscore(SelfEfficacy), Zscore(LeaderSupport)

d. Predictors: (Constant), Zscore(SelfEfficacy), Zscore(LeaderSupport), IntVariable

The table above is the ANOVA table showing the significant results of the models included in the hierarchical regression analysis. In the ANOVA table, if the Sig. values are less than 0.05, it is considered that the variables included in the regression model (recall that different variables were included in the analysis

in all three models) explain the variance in the dependent variable very well. Sig. values are $p = 0.001$ (rightmost column). Accordingly, all three models are statistically significant.

Table 6.8. Coefficients Table

Coefficients ^a						
Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	14,492	,204		71,063	,000
	Zscore(SelfEfficacy)	1,004	,204	,318	4,910	,000
2	(Constant)	14,492	,200		72,488	,000
	Zscore(SelfEfficacy)	,773	,214	,245	3,615	,000
	Zscore(LeaderSupport)	,664	,214	,211	3,109	,002
3	(Constant)	14,260	,204		69,874	,000
	Zscore(SelfEfficacy)	,918	,211	,291	4,345	,000
	Zscore(LeaderSupport)	,691	,208	,219	3,324	,001
	IntVariable	,671	,181	,234	3,705	,000

a. Dependent Variable: SocialSupport

This is the table with the data we will use in reporting. Regression coefficients are located in column “B”, which is the sub-column of the “Unstandardized Coefficients” column. The rightmost “Sig.” column shows the significance of the effect. Accordingly, when Model 3, which includes the interaction variable and the last stage of hierarchical regression analysis, is examined, it can be seen that the effects of self-efficacy ($\beta=0.918$, $p<0.001$), leader support ($\beta=0.691$, $p<0.05$), and the interaction variable ($\beta=0.671$, $p<0.001$) on the perception of social support are significant and positive.

In addition to these findings, researchers are expected to graphically show the effect of the moderator variable in their articles or theses. To do this, a suitable approach would be to create the Slope graph proposed by Aiken and West [2]. However, there is no macro for graphical display in SPSS (The moderation effect will be explained with the Hayes method in the following sections. There is a macro for graphical display in this method). Therefore, the graphical representation will be made with “Microsoft Excel”. Although many files have been prepared for this purpose and are available online, the Excel file at <http://www.jeremydawson.co.uk/slopes.htm> will be used in this part of our book. When the relevant web page is accessed, the following screen will be displayed at the bottom of the web page. On this screen, there are many “Excel” files but the “2-way_standardised.xls” file is downloaded.

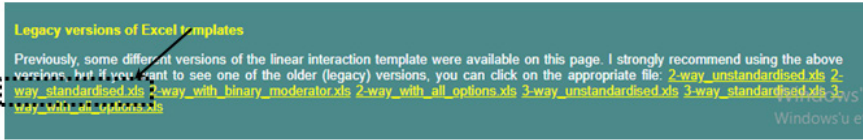
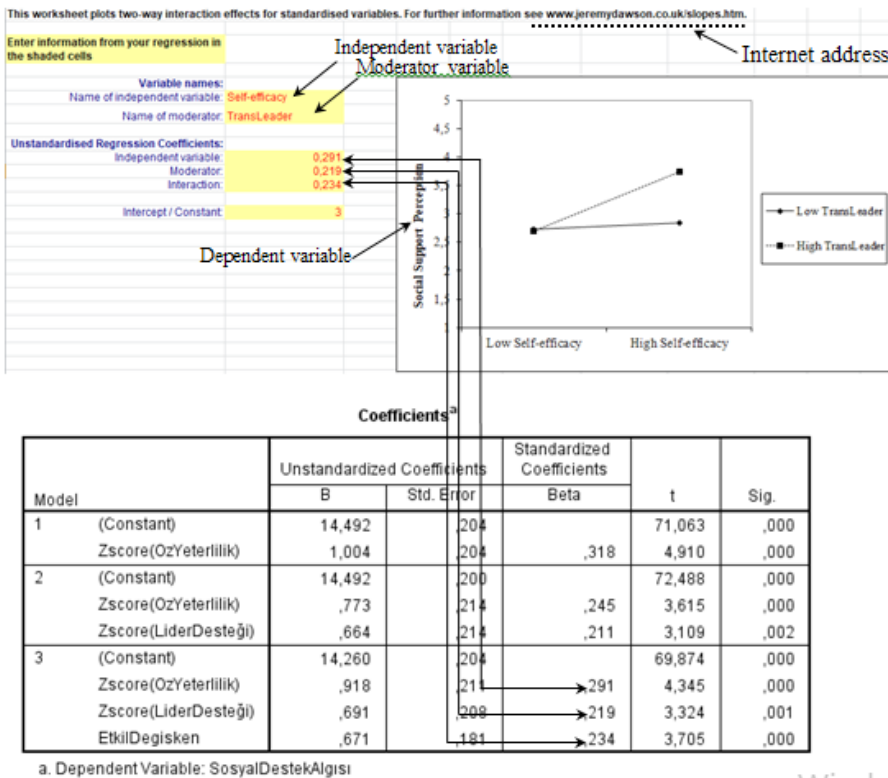


Figure 6.16. File download image

When file is opened, the Beta values obtained are assigned to the relevant sections

Table 6.9. Slope Chart Creation Table



Note here that the Intercept/Constant value is set to 3. Our Constant value is 14.260. However, when we assign this value in the relevant place, the graph goes off the screen. Therefore, a value of 3 is assigned. The main reason for this situation is that although the Excel file was prepared for 5-point Likert-type scales, the variables in our study were measured with a 7-point Likert-type scale. Setting the Intercept/Constant value to 3 in this sense does not cause errors in the interpretation of the findings visually. The main purpose is to provide readers with a graphical representation of the moderation effect.

REPORTING RESEARCH FINDINGS

Hierarchical regression analysis was conducted to determine the moderator role of leader support in the effect of self-efficacy on the perception of social support. To test our hypotheses, the independent variable and moderator variable were firstly standardized and then the interaction variable was created using ± 1 standard deviation [2]. The three-step regression analysis is as follows:

In the first step, the primary effect of self-efficacy on the perception of social support was examined. In the second step, the leader support variable was included in the analysis, and in the third step, the interaction variable was also included in the model and the moderation effect of the leader support was examined. The reason why the interaction variable was included in the model last was to determine how much of the variance the interaction variable explains the dependent variable, in addition to the independent and moderation variable included in the analysis in the previous step. If the interaction variable included in the model is statistically significant and contributes to explaining the variance in the dependent variable, it could be said to have a moderation effect. The results are presented in Table 6.10.

Table 6.10. Regression Results

	Perception of Social Support		
	Model 1	Model 2	Model 3
Self-Efficacy	1.004***	0.773***	0.918***
Leader Support		0.664**	0.691**
Self-Efficacy * Leader Support			0.671***
R ²	0.097	0.132	0.181
F	24.108***	17.375***	16.853***

In the third stage of the regression analysis to determine the moderation effect, the interaction variable (Self-Efficacy*Leader Support) was included in the regression model (Model 3), as suggested by Baron and Kenny [3]. The interactional variable has a positive and significant effect on the perception of social support ($\beta = 0.671$, $p < 0.001$).

To better explain the interaction patterns, a simple slope plot of the high (1 SD) and low (-1 SD) levels of the moderation variable, leader support, was drawn

using the standardized regression coefficients (β) in the regression equation [2,4]. This procedure aims to provide less biased regression coefficients for measuring the moderator variable effect. The simple slope graph of the moderator effect of leader support on the effect of self-efficacy on the perception of social support is presented in Figure 6.17.

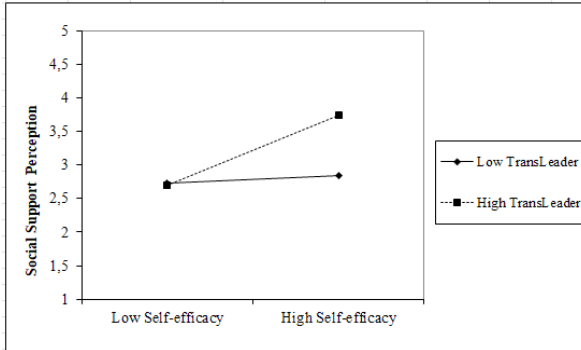


Figure 6.17. Simple Slope Graph

Simple slope analysis reveals that when there is low leader support (below the mean 1 SD), self-efficacy has a positive effect on the perception of social support, whereas in the existence of high leader support (above the mean 1 SD), self-efficacy does not have a significant effect on the perception of social support. Simple slope analysis reveals that leader support has a moderator role in the effect of self-efficacy on perceived social support. In the light of the findings, the research hypothesis was supported.

MODERATION EFFECT ANALYSIS WITH AMOS

As the readers will recall, in the hierarchical regression analysis regarding the effect of the moderator variable, the research question was determined as “Does leader support have a moderator role in the effect of self-efficacy on the perception of social support?”. We will conduct our analysis on the same research question.

First of all, we open the “Data-3.sav” file (You can access this file at www.indataanalysis.com).

Let’s recall the model for our research question:

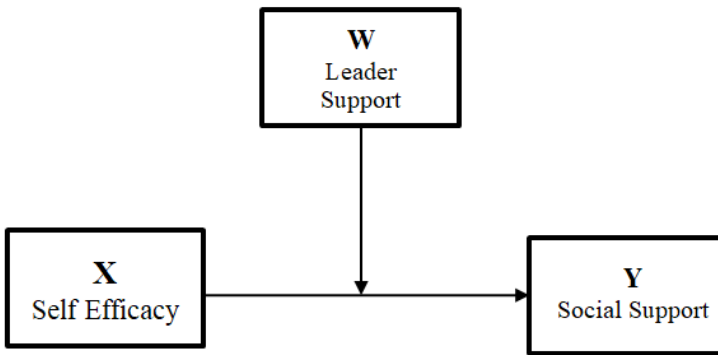


Figure 6.18. Research Model

The statistical representation of our research question is presented below.

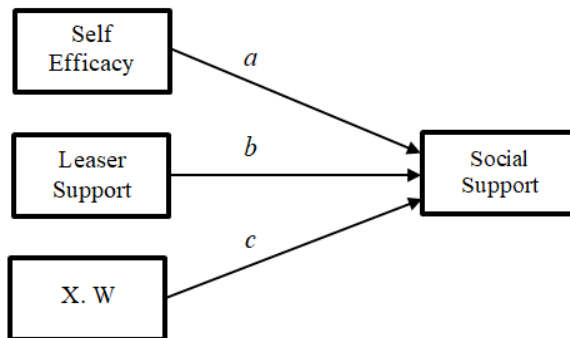


Figure 6.19. Statistical Representation of Moderator Variable Model

To test the relevant model, it is firstly necessary to draw it on AMOS. The model drawing process will not be described again here. However, readers who wish can refer to the previous model drawing section (Chapter 3). The model to be tested for the moderation effect is presented in Figure 6.20.

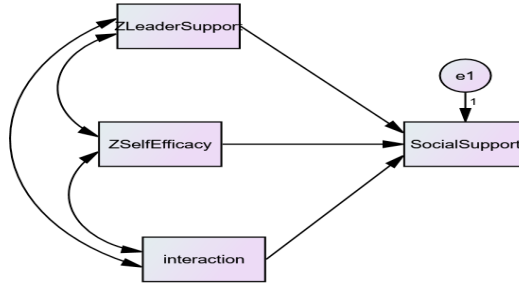


Figure 6.20. AMOS Drawing of the Moderation Effect Model

As can be seen in the figure, the above model has basically been drawn on AMOS. However, before drawing the model, the interaction term must first be created on SPSS. The process of creating the interaction term was explained before in the section on Moderation analysis with SPSS.

As can be observed, while drawing the model, the independent variable self-efficacy, the regulatory variable leader support, and the interaction term obtained by multiplying these variables were drawn as exogenous variables. The values of the independent variable, moderator variable, and interaction variable used in the analyses were normalized. Since the normalization of variables was described earlier section in the analysis with SPSS, it will not be repeated here. All these exogenous variables were also covaried by connecting with a double arrow. The coefficients after loading the relevant data and entering the analysis settings are presented in Figure 6.21. The same values are presented in tabular form in Table 6.11

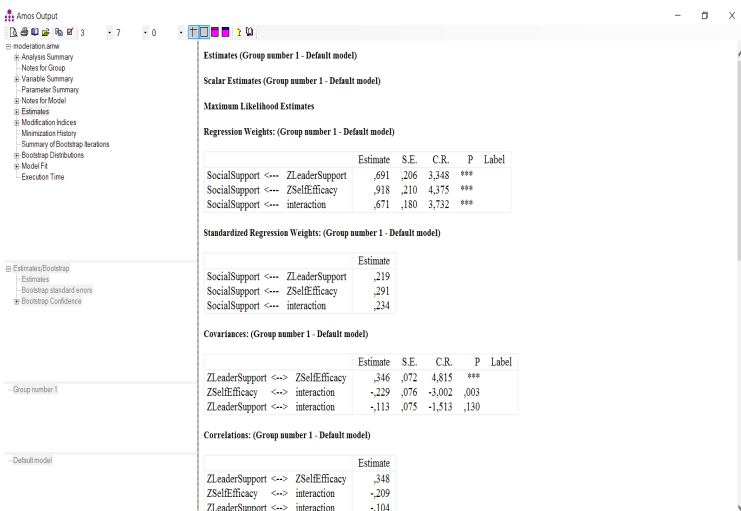


Figure 6.21. “Estimates” tab after analysis

Table 6.11. Results for “Estimates”**Estimates (Group number 1 - Default model)****Scalar Estimates (Group number 1 - Default model)****Maximum Likelihood Estimates****Regression Weights: (Group number 1 - Default model)**

			Estimate	S.E.	C.R.	P	Label
SocialSupport	<---	ZLeaderSupport	.691	.206	3.348	***	
SocialSupport	<---	ZSelfEfficacy	.918	.210	4.375	***	
SocialSupport	<---	interaction	.671	.180	3.732	***	

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
SocialSupport	<---	ZLeaderSupport	.219
SocialSupport	<---	ZSelfEfficacy	.291
SocialSupport	<---	interaction	.234

Covariances: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
ZLeaderSupport	<-->	ZSelfEfficacy	.346	.072	4.815	***	
ZSelfEfficacy	<-->	interaction	-.229	.076	-3.002	.003	
ZLeaderSupport	<-->	interaction	-.113	.075	-1.513	.130	

Correlations: (Group number 1 - Default model)

			Estimate
ZLeaderSupport	<-->	ZSelfEfficacy	.348
ZSelfEfficacy	<-->	interaction	-.209
ZLeaderSupport	<-->	interaction	-.104

Variances: (Group number 1 - Default model)

	Estimate	S.E.	C.R.	P	Label
ZLeaderSupport	.995	.096	10.368	***	
ZSelfEfficacy	.995	.096	10.368	***	
Interaction	1.202	.116	10.368	***	
e1	7.996	.771	10.368	***	

As can be seen from the values in the table, the standardized effect of self-efficacy on social support is 0.291. The effect of leader support is found to be 0.219 and the effect of the interactional term is 0.234. The “p” values in the

regression table show that all these coefficients are significant. As a result, it can be said that leader support has a moderator role in the effect of self-efficacy on the perception of social support. A simple slope graph can be drawn to understand how this role is achieved. The estimation and standard error variables in the regression table are used while plotting the slope graph. The process by which this graph is drawn has been explained in detail in the previous example of moderation effect analysis with SPSS. Again, the results here are assigned in the relevant Excel table, and the slope graph is drawn and interpreted.

ANALYSIS OF THE MODERATOR VARIABLE EFFECT WITH PROCESS MACRO

The moderator variable affects the strength of the relationship between two variables (dependent and independent variable). The effect of the independent variable on the dependent variable is formed by the strength of the moderator variable. In other words, moderation is used for trying to show that a variable changes the relationship between the independent and dependent variables. Therefore, analysis for the moderator variable is a way of determining whether it affects the strength or direction of the relationship between the independent and dependent variables as a third variable.

In the figure illustrated below, a model was established stating that the relationship between two variables differs according to the level of a third variable. In models for the moderator variable, as the strength/intensity of the moderator variable (low, medium, and strong) changes, the effect of the independent variable on the dependent variable will also change in line with these values.

As readers will recall, in the hierarchical regression analysis regarding the effect of the moderator variable, the research question was determined as “Does leader support have a moderator role in the effect of self-efficacy on the perception of social support?” We will conduct our analysis on the same research question.

First, we open the “Data-3.sav” file (You can access this file at www.indataanalysis.com).

Let's recall the model for our research question:

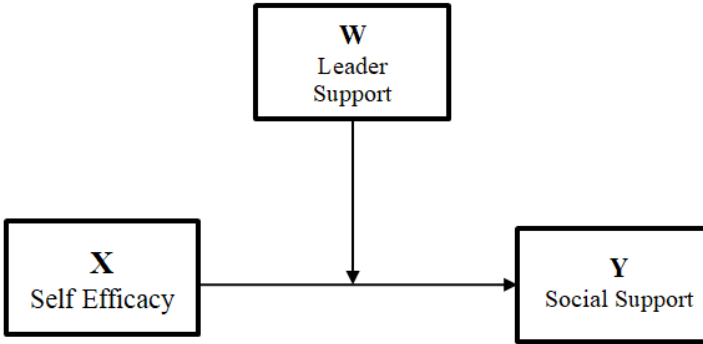


Figure 6.22. Research Model

The statistical representation of the research problem is presented below.

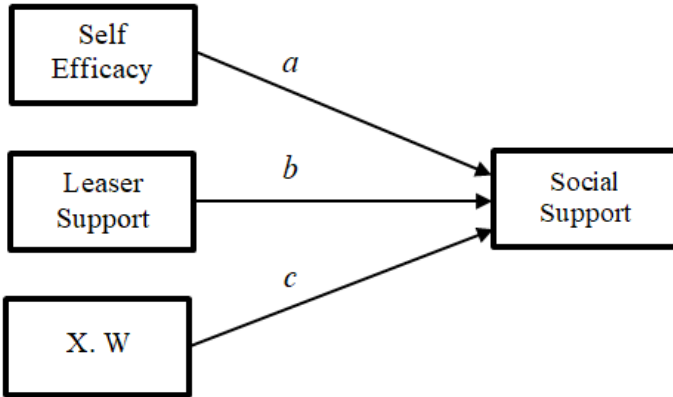


Figure 6.23. Statistical Representation of the Research Model

MODERATION ANALYSIS WITH SPSS

The following steps will be followed to perform analysis for the determination of the moderation effect with the Process Macro proposed by Andrew F. Hayes.

Click that option: the Analyze ---> Regression ----> PROCESS v3.4 by Andrew F. Hayes...

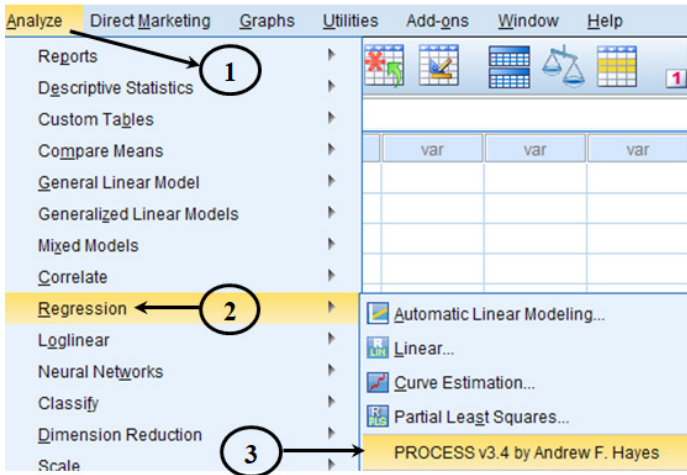


Figure 6.24. SPSS “Analyze” Screen

In the window that opens, we assign the relevant variables.

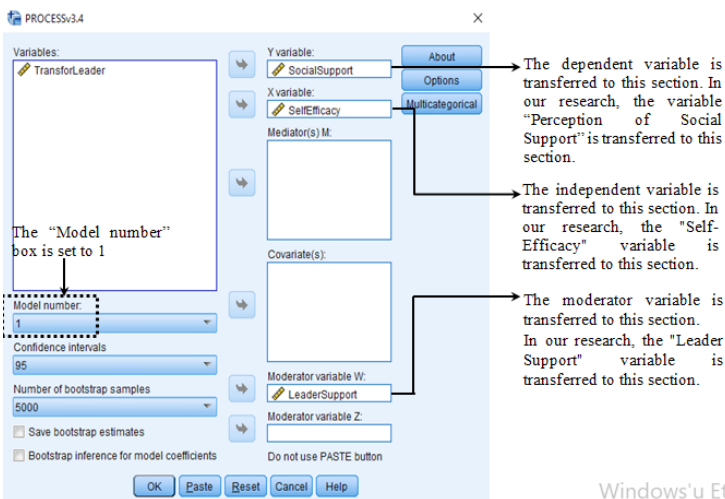


Figure 6.25. “PROCESS Macro” Screen

After the above processes are completed, the “Options” tab is selected. The actions to be taken are presented below.

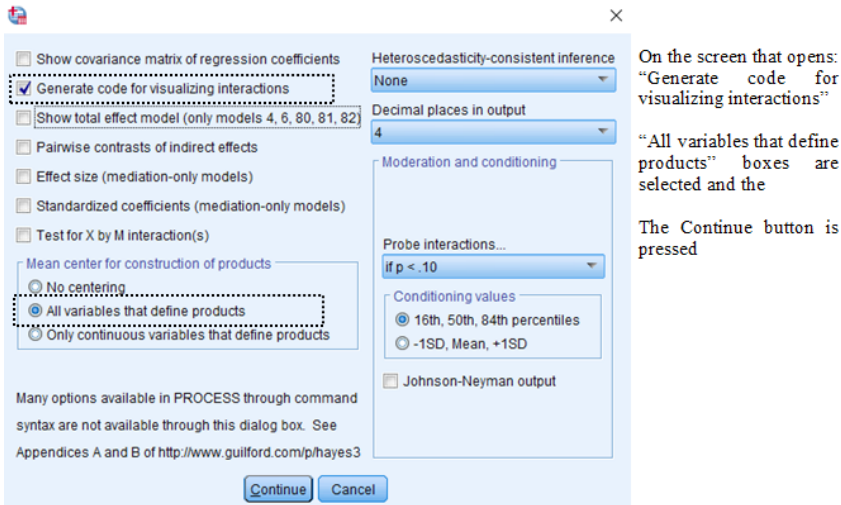


Figure 6.26. “Options” Screen

It may take some time for the outputs to be generated, as the bootstrap technique is used to calculate standard errors and confidence intervals. After the outputs are displayed on the screen, the findings are interpreted and reported.

In 5-10 seconds, the screen in Figure 6.27 will appear as the SPSS output and then the outputs will appear on the screen.

```

DATASET ACTIVATE DataSet1.

* Encoding: UTF-8.

/* PROCESS version 3.4 */.
/* Written by Andrew F. Hayes */.
/* www.afhayes.com */.
/* www.processmacro.org */.
/* Copyright 2019 by Andrew F. Hayes */.
/* Documented in http://www.guilford.com/p/hayes3 */.
/* PROCESS workshop schedule at http://www.processmacro.org/workshops.html */.

/* Distribution of this code in any form, except through processmacro.org, is prohibited */.
/* without the permission of the copyright holder */.

/* THIS SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND */.
/* EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF */.
/* MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT */.
/* IN NO EVENT SHALL THE COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, */.
/* DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT */.
/* OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE */.
/* SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE */.
/* USE OF THIS SOFTWARE IMPLIES AGREEMENT WITH THESE TERMS */.

set printback=off.

```

Figure 6.27. SPSS Output Screen

Table 6.12. Results showing the role of the moderator variable (Leader Support) in the effect of the independent variable (Self-Efficacy) on the dependent variable (Perception of Social Support):

```

*****
Model : 1 -----> Model number of the study
  Y : SocialSu
  X : SelfEffi
  W : LeaderSu } Variables included in the research

Sample
Size: 216 -----> Sample Size

*****
OUTCOME VARIABLE:
  SocialSu

Model Summary
      R      R-sq      MSE      F      df1      df2      p
      ,4388      ,1926      8,1470      16,8527      3,0000      212,0000      ,0000
                                     95% CI, Confidence Interval

Model
      coeff      se      t      p      LLCI      ULCI
constant      14,2601      ,2041      69,8739      ,0000      13,8578      14,6624
SelfEffi Path a      ,1188      ,0273      4,3447      ,0000      ,0649      ,1727
LeaderSu Path b      ,1665      ,0501      3,3243      ,0010      ,0678      ,2652
Int_1 Path c      ,0210      ,0057      3,7055      ,0003      ,0098      ,0321
-----> Interaction Variable

Product terms key:
  Int_1 :      SelfEffi x      LeaderSu
                                     Moderation Effect is Significant
                                     (Confidence interval does not include Zero (0))

Test(s) of highest order unconditional interaction(s):
      R2-chng      F      df1      df2      P
X*W      ,0523      13,7305      1,0000      212,0000      ,0003
-----> Additional variance explained by the
           interaction variable
    
```

In the table above, findings are shown in terms of whether leader support has a moderator effect on the effect of self-efficacy on the perception of social support. The significance and effect of the moderation effect are determined by interpreting the findings in the “Int_I” line (path c). In this example, the effect of the interaction variable (Int_I) is significant and is contained within the suggested confidence interval ($\beta = 0.0210$, 95%, CI [LLCI=0.0098, ULCI=0.0321], $t=3.7055$, $p<0.05$). This finding reveals that leader support has a moderator effect.

The significance of the effect is understood from the fact that the p-value is less than 0.05 ($p=0.0003$) and the effect is in the suggested confidence interval [(between the lower limit value (LLCI=0.0098) and the upper limit value (LLCI=0.0098) and zero (0) is not included)].

Table 6.13. Results showing the effect of the independent variable (Self-Efficacy) on the dependent variable (Perception of Social Support) when the interaction variable has Low, Medium, and High values:

```

-----
Focal predict: SelfEffi (X)
Mod var: LeaderSu (W)

Conditional effects of the focal predictor at values of the moderator(s):

```

	LeaderSu	Effect	se	t	p	LLCI	ULCI
Low	-4,0936	,0330	,0324	1,0202	,3088	-,0308	,0969
Medium	,9064	,1378	,0287	4,7944	,0000	,0812	,1945
High	3,9064	,2007	,0382	5,2516	,0000	,1254	,2760

The effect is not significant when the interaction value is low

Effects of the interaction variable at different levels (low, medium, high)

Significance Value

95% CI, Confidence Interval

As readers will remember, we stated that in models for the moderator variable, as the strength/intensity of the moderator variable (low, medium, and strong) changes, the effect of the independent variable on the dependent variable also changes according to these values. The table above shows the changes according to these values. The first line shows the effect of self-efficacy on the perception of social support when the leader support is low, the second line is moderate leader support, and the third line is high leader support.

Table 6.14. Data required for graphical representation

```

DATA LIST FREE/
  SelfEffi  LeaderSu  SocialSu  .
BEGIN DATA.
  -8,5541  -4,0936  13,2959
   1,4459  -4,0936  13,6263
   9,4459  -4,0936  13,8907
  -8,5541   ,9064  13,2321
   1,4459   ,9064  14,6102
   9,4459   ,9064  15,7128
  -8,5541  3,9064  13,1937
   1,4459  3,9064  15,2006
   9,4459  3,9064  16,8060
END DATA.
GRAPH/SCATTERPLOT=
  SelfEffi WITH      SocialSu BY      LeaderSu .

```

Data required for graphical representation

The table above shows the data required to create a graphical representation (simple slope graph). These data will be used in the reporting the findings.

To see the moderation effect graphically, a simple slope graph should be drawn. This graph is used in the reporting of research findings. While a simple slope graph could be drawn in “Excel” with regression analysis, this graph can also be drawn in SPSS with the plugin in Process Macro. The following steps will be followed for drawing the graph via SPSS:

First, the data in the SPSS output are copied.

Table 6.15. Data required for graphical representation

SelfEffi	LeaderSu	SocialSu
-8,5541	-4,0936	13,2959
1,4459	-4,0936	13,6263
9,4459	-4,0936	13,8907
-8,5541	,9064	13,2321
1,4459	,9064	14,6102
9,4459	,9064	15,7128
-8,5541	3,9064	13,1937
1,4459	3,9064	15,2006
9,4459	3,9064	16,8060

DATA LIST FREE/
SelfEffi LeaderSu SocialSu .
BEGIN DATA.
-8,5541 -4,0936 13,2959
1,4459 -4,0936 13,6263
9,4459 -4,0936 13,8907
-8,5541 ,9064 13,2321
1,4459 ,9064 14,6102
9,4459 ,9064 15,7128
-8,5541 3,9064 13,1937
1,4459 3,9064 15,2006
9,4459 3,9064 16,8060
END DATA.
GRAPH/SCATTERPLOT=
SelfEffi WITH SocialSu BY LeaderSu .

Cut
Copy
Paste
Delete

Data are selected and copied by double-clicking on the screen.

First, a new Syntax table opens. To open the syntax table; File ---> New ----> Syntax... option is selected.

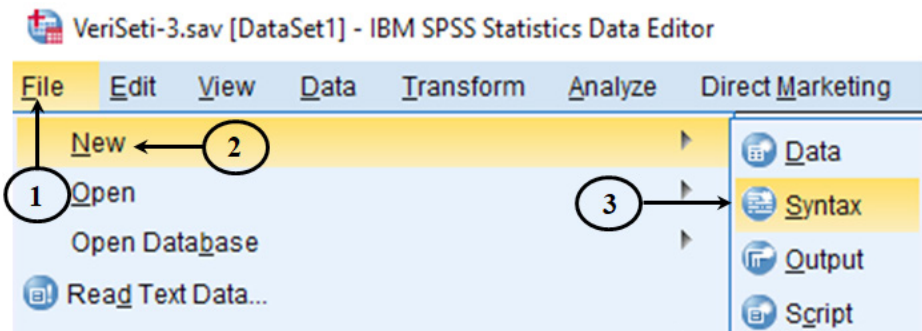


Figure 6.28. Opening the Syntax Screen

The data we copied is pasted into the Syntax screen that opens.

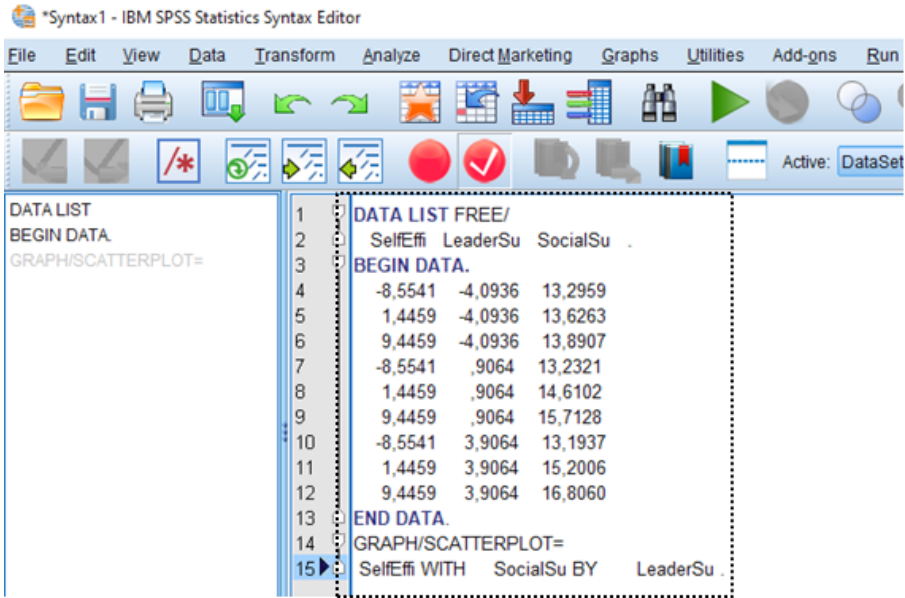


Figure 6.29. Syntax Screen

After the copy process is complete, Run ---> All.... option is selected.

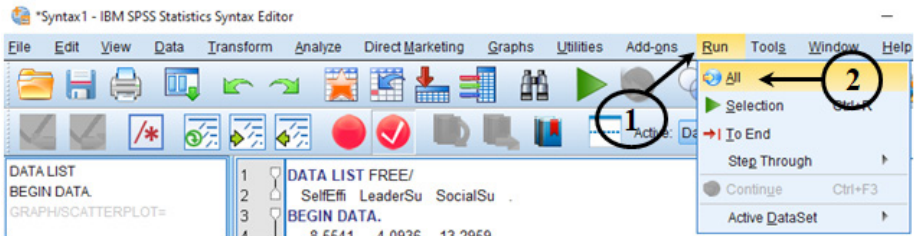


Figure 6.30. SPSS “Run” Screen

When this process is completed, the graphical representation in Figure 6.31 will be created in SPSS's Output file (SPSS output).

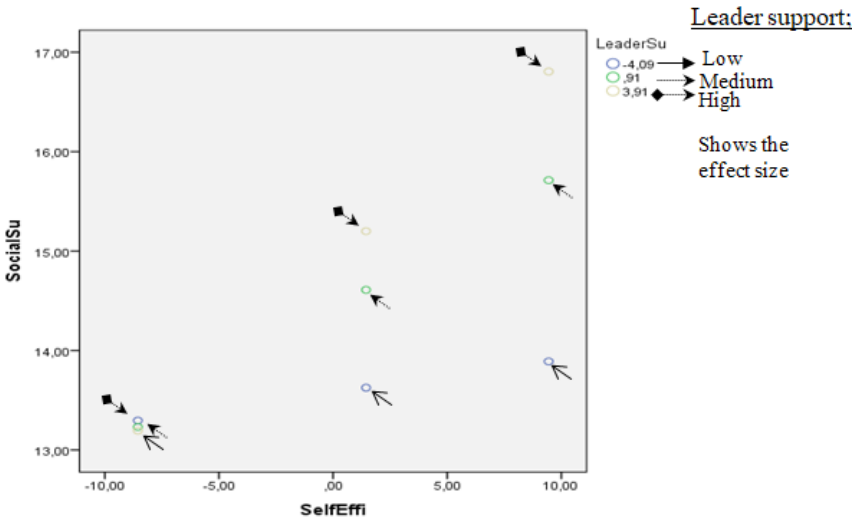


Figure 6.31. Graphical Display

Although the table above provides researchers with a graphical representation of the effect of the moderation variable, it is not sufficient for reporting. For this reason, actions will be taken to increase the visuality.

First of all, the SPSS output is opened and the graphical display is prepared for operation by double-clicking.

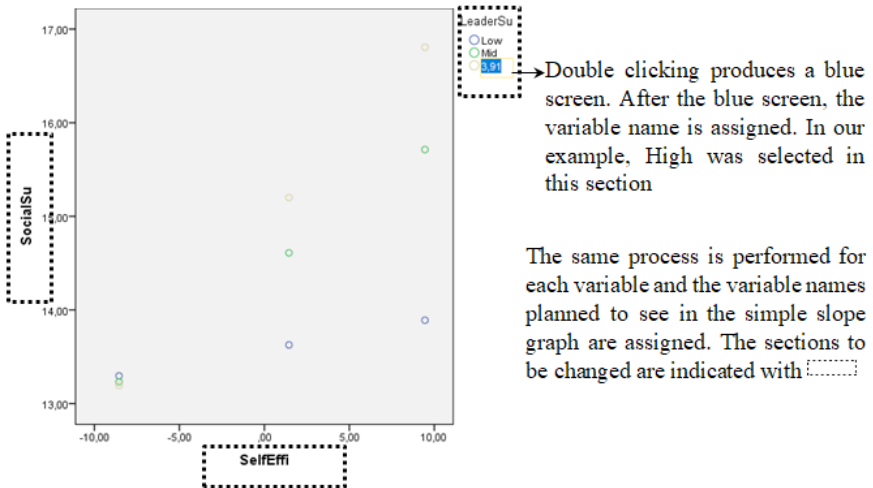


Figure 6.32. Graphical Display, Editing-1

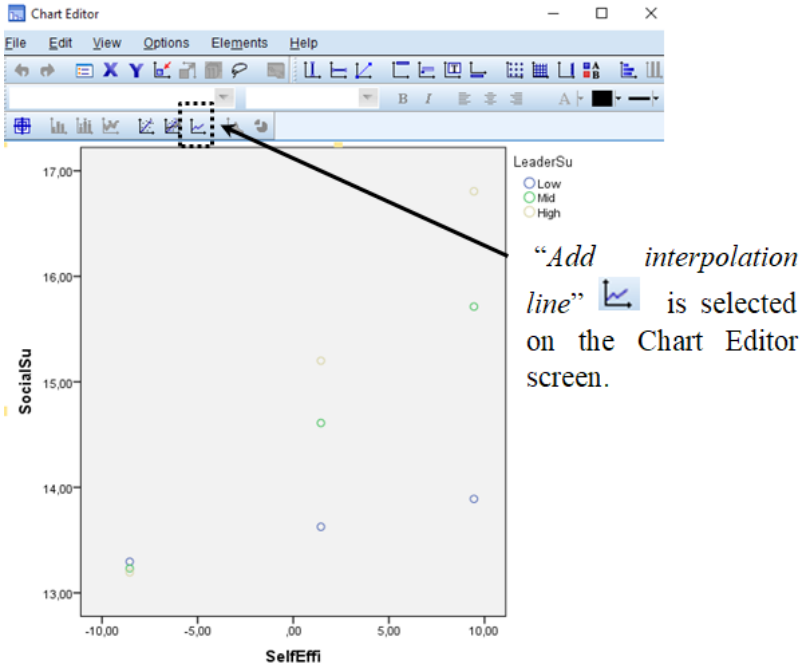


Figure 6.33. Graphical Display, Editing-2

The following procedures are implemented according to the preference of the researcher. The following screen shows a preferable example.

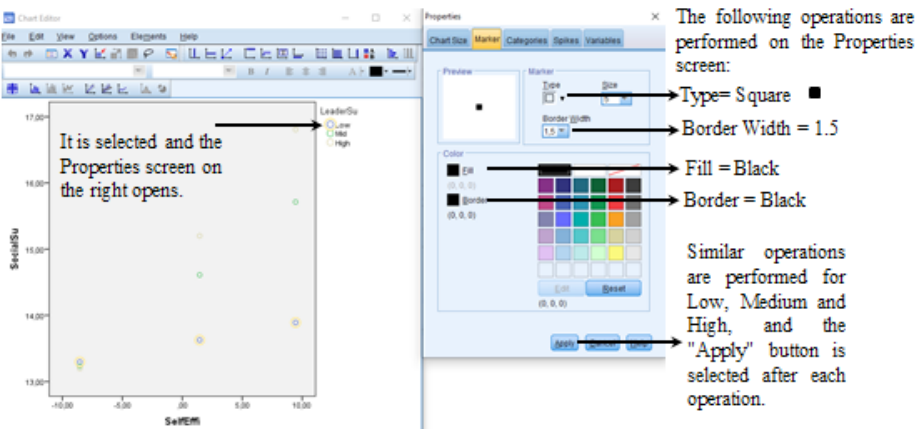


Figure 6.34. Graphical Display, Editing-3

After the processes are completed, we close the screens by clicking on x in the upper right corner of all screens.

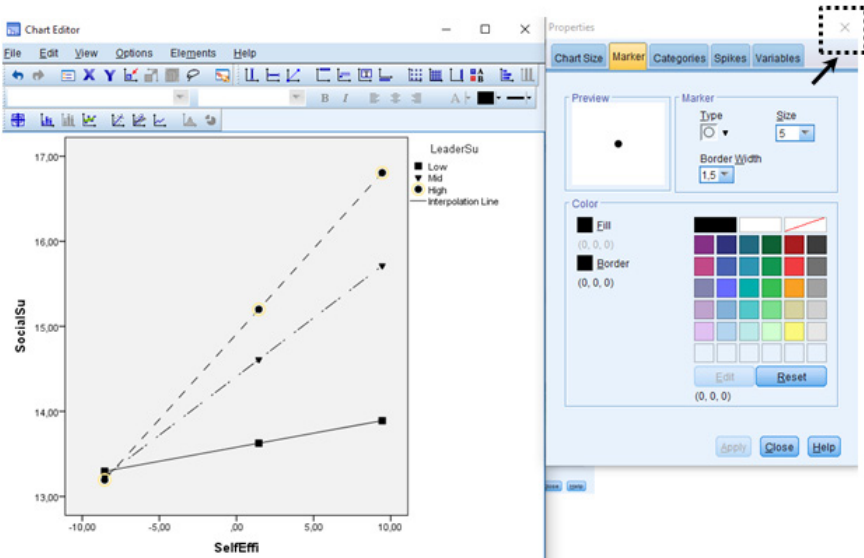


Figure 6.35. Graphical Display, Editing-3

When all operations are completed successfully, the following simple slope graph will be seen in the output of SPSS.

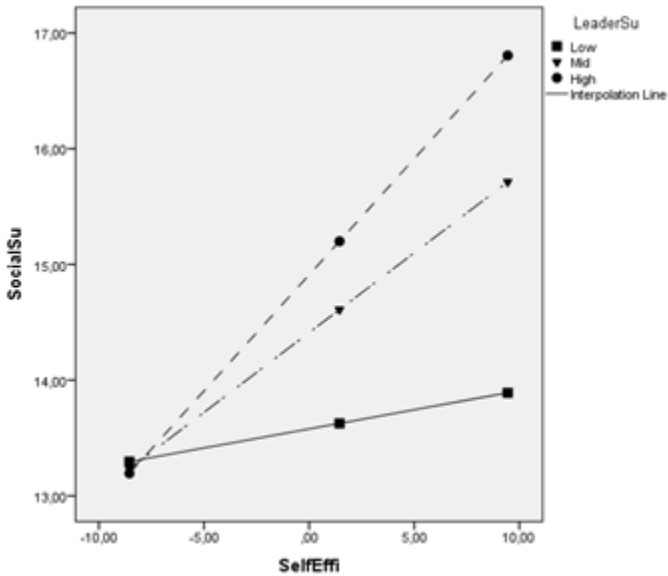


Figure 6.36. Simple Slope Graph

Interpreting a simple slope graph

As the readers will remember, there is a table in the outputs of SPSS that shows the effect of the moderation effect in low, medium, and high situations. The table below shows this effect.

Table 6.16. Effect of Moderation on Low, Medium, and High Situations

```

-----
Focal predict: SelfEffi (X)
Mod var: LeaderSu (W)

Conditional effects of the focal predictor at values of the moderator(s):
The effect is not significant when the interaction value is

```

LeaderSu	Effect	se	t	p	LLCI	ULCI
<u>Low</u>	-4,0936	,0330	1,0202	,3088	-,0308	,0969
<u>Medium</u>	,9064	,1378	4,7944	,0000	,0812	,1945
<u>High</u>	3,9064	,2007	5,2516	,0000	,1254	,2760

Effects of the interaction variable at different levels (low, medium, high)
Significance Value
95% CI, Confidence Interval

This data on is shown on a simple slope graph.

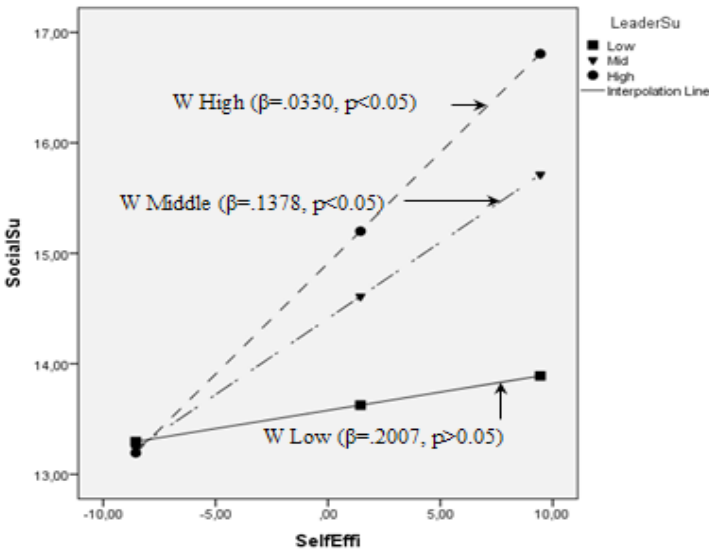


Figure 6.37. Effects of Moderator Variable

Please pay attention to the simple slope graph in Figure 6.37; when the interaction effect is low ($\beta=0.2007, p<0.05$), the slope is close to a straight line.

This line indicates that there is no significant effect when the interaction effect is low. In cases where the interaction effect is high ($\beta=0.0330$, $p<0.05$) or moderate ($\beta=0.1378$, $p<0.05$), the slope has a vertical appearance. In this case, the effect is significant when the interaction effect is high and medium. The greater the slope, the greater the effect. The graph in the simple slope graph corresponds to the findings in the slope analysis.

REPORTING RESEARCH FINDINGS

After the simple slope graph has been created, the process of reporting the findings can begin.

The regression method proposed by Baron and Kenny [3] is commonly used in the testing of moderation hypotheses. However, the extant empirical literature indicates that this approach does not give strong enough statistical values and has many potential shortcomings [5,6,7,8]. Therefore, the contemporary approach, “Process Macro” software developed by Hayes [5] as an add-on to SPSS, was used to test the research hypotheses. In the Process Macro method, 5000 resampling options were preferred with the bootstrap technique.

Regression results for determining the moderator effect of leader support on the effect of self-efficacy on the perception of social support are presented in Table 6.17.

Table 6.17. Regression Results

Variables	B	SH	P	t	LLCI	ULCI
(Constant)	14.260	.2041	.000	69.874	13.857	14.662
Self Efficacy (X)	.1188	.0273	.000	4.3447	.0649	.1727
Leader Support (W)	.1665	.0501	.001	3.3243	.0678	.2652
Interaction Variable (X*W)	.0210	.0057	.000	3.7055	.0098	.0321
R				.4388		
R ²				.1926		
Adjusted R ²				.0523		

Table 6.17 shows that all the variables included in the study explain approximately 19% of the change in the perception of social support. Furthermore, the additional variance explained by the interaction variable is approximately 5%. In addition to these findings, self-efficacy ($\beta = 0.1188$, 95% CI= [0.0649, .1727], $t=4.3447$, $p<0.05$) and leader support ($\beta = 0.1665$, 95% CI= [0.0678, 0.2652], $t=3.3243$, $p<0.05$) were found to have a significant and positive effect on the perception of social support. In addition, the findings reveal that the interaction effect of the moderator term (Self-efficacy * leader support) on the perception of social support is significant ($\beta = 0.0210$, 95% CI= [0.0098, 0.0321], $t=3.7055$, $p<0.05$). According to these results, leader support has a moderating effect on the effect of self-efficacy on the perception of social support. Aiken et al. [2] suggested that a graphical representation should be

created to visually consider the possible results of moderation effect tests. Therefore, the simple slope graph is presented in Figure 6.38.

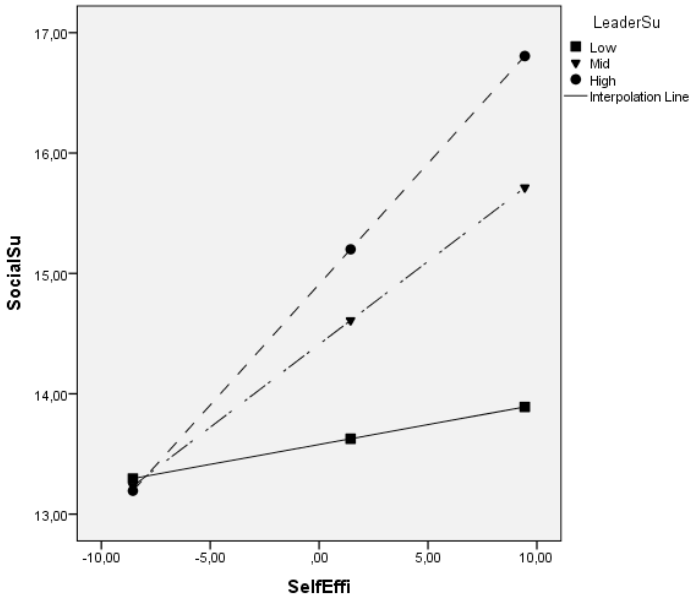


Figure 6.38. Simple Slope Graph

In cases where leader support is high ($\beta=0.0330$, $p<0.05$) or moderate ($\beta=0.1378$, $p<0.05$), the effect of self-efficacy on the perception of social support is significant. However, the effect of self-efficacy on the perception of social support is not significant in cases where leader support is low ($\beta=0.2007$, $p>0.05$). These results reveal that as leader support increases, the effect of self-efficacy on the perception of social support will also increase. Leader support has a moderator effect on the effect of self-efficacy on the perception of social support.

References

- [1] MacKinnon, D. P. (2011). Integrating mediators and moderators in research design. *Research on Social Work Practice, 21*(6), 675-681.
- [2] Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.
- [3] 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*(6), 1173.
- [4] Hayes, A. F., & Matthes, J. (2009). Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods, 41*(3), 924–936.
- [5] Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.
- [6] 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*(1), 83.
- [7] Kenny, D. A., Kashy, D. A., & Bolger, N. (1998). Data analysis in social psychology. *The handbook of Social Psychology, 1*(4), 233-265.
- [8] Zhao, X., Lynch Jr, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research, 37*(2), 197-206.