African Educational Research Journal
Vol. 10(3), pp. 277-290, August 2022
DOI: 10.30918/AERJ.103.22.035
ISSN: 2354-2160
Full Length Research Paper

A comparison of fixed and random effect models by the number of research in the meta-analysis studies with and without an outlier

Seda Demir[1]* and Mehmet Fatih Doğuyurt[2]

[1]Department of Measurement and Evaluation in Education, Tokat Gaziosmanpasa University, Tokat, Turkey.
[2]Ministry of National Education, Ankara, Turkey.

Accepted 8 June, 2022

ABSTRACT

The purpose of this research was to compare the performances of the Fixed Effect Model (FEM) and the Random Effects Model (REM) in the meta-analysis studies conducted through 5, 10, 20 and 40 studies with an outlier and 4, 9, 19 and 39 studies without an outlier in terms of estimated common effect size, confidence interval coverage rate and heterogeneity measures. In this descriptive study, real data set consisting of different studies examining teachers' emotional burnout in terms of gender were used and a total of 72 meta-analyses were performed with R program. The results indicated that REM was more advantageous when compared to FEM for the meta-analysis of data sets with an outlier. On the other hand, without an outlier, it was determined that the common effect size was generally estimated to be similar for all methods. Moreover, the increase in the number of studies included in the meta-analysis reduced the effect of the outlier on the effect size estimation and decreased the heterogeneity. When the examination of the confidence interval coverage accuracy rates of the meta-analysis methods was examined, it was concluded that the confidence intervals included the estimated effect sizes in all data sets and all methods. The findings of the current study showed that the methods used in meta-analysis studies with 20 or more studies were less affected by the outlier runs in the estimated common effect size.

Keywords: Meta-analysis, outlier, fixed effect model, random effects model, heterogeneity measures.

*Corresponding author. E-mail: seddadmr@gmail.com.

[#]This study was presented as an oral presentation at the 7th International Congress on Measurement and Evaluation in Education and Psychology (CMEEP), held on September 1-4, 2021.

INTRODUCTION

Since technology has been developing day to day, there occur differentiating and increasing needs, which leads to an increase in the number of scientific research. In this sense, it is likely to come across numerous studies on the same or similar research themes in the literature. Although the common research questions are included in these studies, it becomes very difficult to draw a general conclusion from all of these studies due to such factors as the change in the effect sizes of the studies in positive and negative intervals, the differentiation of research methodologies and designs, population and samples (Demir and Başol, 2014). At this point, there needs a thorough review of studies, namely a meta-analysis of studies in which different studies on the same theme are brought together to combine their results so that a common conclusion can be drawn. Glass (1976), who first used the term meta-analysis, defines the statistical analyses towards data in research as the *primary analysis*, and those towards re-analysis of primary analysis results via more advanced statistical analyses as *secondary analysis*. Based on this, it can be noted that samples in a meta-analysis study consist of previous research, and data is drawn from the results of previous research. Meta-analysis is a statistical analysis technique that combines different research results on a subject through one or more statistical methods, standardizes them in a common metric, summarizes the calculated statistical results together with the research characteristics, and provides more information than the primary results of the previous research (Glass, 1976; Hedges and Olkin, 1985). Comparisons can be made and common (overall) effect size can be calculated thorough combining the results of multiple studies (Rudy, 2001). In this respect, meta-analysis can be considered as combining a group of studies on the same subject to reach some conclusions and generalizations on the related subject (Lipsey and Wilson, 2001). Since the results of a wide range of studies are combined in the meta-analysis, it is possible to examine over a larger sample than primary studies and to increase the statistical power of the study (Borenstein et al., 2009; Normand, 1999; Schmid et al., 1991).

Just as it is in the primary studies, the distribution of the data set is of great importance in terms of summarizing the results in meta-analyses studies, as well (Riley et al., 2011). Furthermore, the variance of the effect sizes obtained from the primary studies included in the meta-analysis should be taken into account in the selection of the model to be used in the meta-analysis (Borenstein et al., 2009). The most frequent models used in the literature to calculate the common effect size by combining each effect size of the studies via meta-analyses are the Fixed Effect Model (FEM) and Random Effects Model (REM) (Borenstein et al., 2009; Hedges and Vevea, 1998). In case the heterogeneity between effect sizes of the studies included in a meta-analysis is high, REM is preferred instead of the FEM (Borenstein et al., 2009; Schwarzer et al., 2015). When the effect sizes of the studies are heterogeneously distributed, there are several estimation methods used under REM (Schwarzer et al., 2015). In addition to the most widely used method, the moments estimator DerSimonian-Laird (DL), other methods included under the REM are as follows: maximum likelihood estimator (ML), restricted maximum likelihood estimator (REML), empirical Bayes estimator (EB), Sidik-Jonkman estimator (SJ), Paule-Mandel estimator (PM), Hunter-Schmidt estimator (HS) and Hedges estimator (HE). There have been different perspectives toward these estimators. Langan et al. (2015), for example, suggest other methods instead of DL in a meta-analysis, whereas Thorlund et al. (2011) suggest that the REML method often produces better results than the DL and that the HE-based and the SJ-based estimates are, on average, larger estimates when compared to the DL-based estimates. Petropoulou and Mavridi (2017), on the other hand, note that SJ presents poor performance when compared to other estimators. Accordingly, it is possible to claim that there is no clarity or agreement in the literature about which method is more convenient to use than others under what conditions. However, it should be taken into account that because the weights of the studies in the meta-analysis change based on the method used, the method to be used is of great importance in that it has a direct impact on determining the common effect size (Borenstein et al., 2009; Hedges and Vevea, 1998).

In meta-analyses studies, the reason why there is a lack of the between-study homogeneity, namely the heterogeneous distribution of the effect sizes may be due to several issues, including sampling errors, the problem of outliers, differentiating research characteristics, etc. (Borenstein et al., 2010; Schwarzer et al., 2015; Viechtbauer and Cheung, 2010). Outlier detection is needed when the effect size of one or more studies deviates or falls outside most of the others, by taking high or low values in a distribution or pattern. In the presence of outliers, there will be some bias in the estimation of the common effect size, no matter which model and method are used for meta-analysis (Lin et al., 2017).

In a meta-analysis study, if an outlier is detected, the study with an outlier may be removed from the meta-analysis (Gumedze and Jackson, 2011). However, although the necessity of detection of outliers in meta-analysis studies has been emphasized in many studies in the literature, it is not recommended to remove the outlier from the data set as it may lead to statistical bias (Viechtbauer and Cheung, 2010).

Applying the procedures toward the identification and elimination of outliers by considering the detection and removing the outliers as a complicated and problematic process is unnecessary. In addition, the possible extreme values detected when sample sizes are small or moderate may have been caused by sampling errors, and so they may not be true outliers. Therefore, the elimination of these primary studies with outlier extreme values may lead to an overcorrection of sampling error or a smaller estimate of the variance of the population (Hunter and Schmidt, 2004). Similarly, Schmidt (2008) suggests that it is almost impossible to distinguish between large sampling errors and true outliers. Viechtbauer and Cheung (2010), on the other hand, argues that if the exclusion of one or more studies determined to be outliers from the meta-analysis significantly affects the results, careful attention must be paid to these results. Accordingly, it can be noted that instead of removing the primary studies with an outlier from the data set, it would be better to use the most convenient model covering all the data.

Based on this, in the current study, the main purpose was to compare the performances of a FEM method with eight different REM methods by the number of the research included in the meta-analysis studies with and without an outlier in terms of estimated common effect size, confidence interval coverage rate and heterogeneity measures. In this respect, the research problem of the study was as follows:

How are the performances of DL, ML, REML, EB, SJ, PM, HS and HE methods under REM and a method of FEM in the meta-analyses research conducted through 5, 10, 20 and 40 studies with an outlier and 4, 9, 19 and 39 studies without an outlier in terms of estimated common effect size, confidence interval coverage rate and heterogeneity measures (Cochran's *Q* Test, *Higgins' and Thompson's* I^2 test, and the parameter τ^2, often called *between-studies variance* component)?

This study is thought to contribute to the literature by giving an idea about which meta-analysis methods under FEM or REM would be relatively more functional than others in meta-analysis studies with an outlier.

Effect size

The term effect size was first used in the literature in 1977 through Cohen's *d*. (Cohen, 1977). It is used for calculating sampling size and is a parameter used in meta-analyses studies for different purposes. The common effect size obtained as a result of the meta-analysis is equal to the arithmetic mean of the effect sizes of the studies included in the research. In this sense, it is of great importance to select the appropriate effect size statistic for research findings and to make correct calculations. The effect size is not affected by the number of observations and could represent the size of the difference between the groups (e.g., the difference in test scores for males versus females) (Borenstein et al., 2009).

The sample size, mean response and standard deviation of the groups are basics to calculate the effect size over continuous variables (Schwarzer et al., 2015). While calculating the effect size over continuous variables, the difference between the averages and standardized mean differences are mostly employed. If the studies included in the meta-analysis use the same scale, the difference between the averages is used as an effect size measure. The standardized mean difference, on the other hand, is employed if different studies use different scales (e.g. aiming at combining the outcomes of different psychological tests) (Borenstein et al., 2009). In this regard, standardized mean difference as the effect size measure was used in the current study.

Fixed effect model (FEM)

If the studies included in the meta-analysis have homogenous distributions, in other words, if there is such an assumption that all studies are from the same population with the same effect, it is accepted that the differences in the outcomes of the studies are due to sampling errors, and the use of the fixed effect model (FEM) is recommended (Borenstein et al., 2010; Cooper, 2010; Lipsey and Wilson, 2001). The difference between the effect sizes of the studies included in the meta-analysis in the FEM is only due to the within-study variances (Borenstein et al., 2009). FEM estimates a lower variance than REM. Therefore, the confidence interval in FEM is narrower than in REM. Accordingly, the confidence interval coverage rate of FEM is lower than that of REM (Hedges and Vevea, 1998).

Random effects model

If the studies included in the meta-analysis have heterogeneous distributions, in other words, if it is not possible to include all studies from the same population with the same effect, the use of the random effects model (REM) is recommended on the grounds that the difference between the study results may be due to the sampling error as well as the difference between the effect sizes of the studies included in the sample (Borenstein et al., 2010; Cooper, 2010; Lipsey and Wilson, 2001). In REM, the differences among the effect sizes of studies stem from both within-study variances and between-study variances (Borenstein et al., 2010). Several methods (DL, ML, REML, EB, SJ, PM, HS, HE) have been developed in REM to estimate the true effect size, between-estimated effect size variance and estimation of the between-study variance (Schwarzer et al., 2015; Veroniki et al., 2016; Viechtbauer, 2010). It is seen that the most established method in the literature is the moments estimator DL. However, in the current study, all eight estimators mentioned above (DL, ML, REML, EB, SJ, PM, HS, HE) were considered.

Heterogeneity measures

Heterogeneity measures are those that quantify the percentage of the observed variance in the primary studies (Borenstein et al., 2009). It is crucial to determine the heterogeneity correctly while selecting the method to be used in meta-analysis studies. Several methods have been developed to measure heterogeneity among studies included in the meta-analysis. The most common measures of heterogeneity used in the literature are as follows: Cochran's *Q* statistic, Higgins and Thompson's I^2 statistic, and τ^2 statistic (Higgins et al., 2003). Langan et al. (2015) argue that the selection of heterogeneity estimation method can affect the results of a meta-analysis, and so the use of a single estimate of heterogeneity may cause inappropriate conclusions in meta-analyses studies. In the current study, all of these three measures of heterogeneity, Cochran's *Q* statistic, Higgins' and Thompson's I^2 statistic and τ^2 statistic, were used to make comparisons among methods.

Cochran's Q statistic is sensitive to the rate of the observed variance to the within-study error and can be calculated with the following formula (Borenstein et al., 2009):

$$Q = \sum_{i=1}^{k} W_i Y_i^2 - \frac{(\sum_{i=1}^{k} W_i Y_i)^2}{\sum_{i=1}^{k} W_i}$$

In this formulation, W_i represents the study weight, Y_i is for the study effect size and k is the number of studies.

Tau-squared (τ^2) parameter is only used in REM and is defined as the variance of the true effect sizes. In other words, in an infinite number of studies, the variance of estimations that represent true effect sizes for each study, each has an infinite sample size calculated via τ^2. However, true effect sizes cannot be observed, so they cannot be directly calculated. Instead, they are estimated through observed effects (Borenstein et al., 2009).

Higgins' and Thompson's I^2 statistic derived from Cochran's Q statistic represents the between-study heterogeneity as percentages and takes values ranging from 0 (0%) to 1 (100%) (Borenstein et al., 2009). I^2 is not an absolute measure of heterogeneity, but the percentage of variability in the estimates. Its estimate is not systematically affected by the sampling sizes of the studies included in the meta-analysis (Schwarzer et al., 2015). As the sampling sizes are increased, then sampling error (1- I^2) is closer to zero, and therefore I^2 is closer to 1 (Higgins and Thompson, 2002).

Determining outliers

The data used in meta-analysis studies are the effect sizes of primary studies, and these effect sizes can differ from each other depending on random error and various factors (Borenstein et al., 2009). It is possible to test this portion of variability through measures of heterogeneity. The source of observed variability (heterogeneity) may be outliers. An outlier refers to the effect size that denotes the effect size different from the others in the positive or negative direction among included studies in the meta-analysis. Outliers can considerably affect meta-analysis results. In this respect, it is necessary to determine the outliers (Viechtbauer 2010; Viechtbauer and Cheung, 2010). In the current study, graphical values (e.g., funnel plots and radial graphs) and statistical values (e.g., studentized residuals obtained by the DL method) were considered in order to determine the outliers The fact that the calculated studentized residual is outside the parameter range of (-2.58, 2.58) is regarded as evidence that the study to which the residual value belongs is an outlier.

METHODOLOGY

This study, designed as a descriptive one in its nature, aimed to compare the results obtained from meta-analysis studies performed under one FEM method and eight REM methods by the number of the research included in the meta-analysis studies with and without an outlier. In this sense, descriptive research designs are studies in which a given state of affairs is investigated as detailed as possible and provide information about the distribution of the phenomenon (Fraenkel et al., 2012).

Data collection

The phenomenon addressed through meta-analysis in this study is the investigation into teachers' emotional burnout in terms of gender. For this purpose, the keywords "burnout", "occupational burnout", "teacher burnout" were searched in the databases, including Web of Science, SAGE Journals, ERIC, EBSCOhost, Springer, Taylor and Francis, Proquest Digital Dissertations, Ulakbim, and 6248 studies were detected in total. However, a thorough examination of the studies has revealed that only 106 studies met the inclusion criteria of meta-analysis in that they include necessary data to be able to estimate the effect size as well as other variables for estimations. Finally, the sample size with one study with an outlier was formed and 39 studies were randomly selected from 105 studies. Therefore, the meta-analyses in the current study were performed over 40 studies in total.

Table 1 shows the effect sizes, variances of the effect sizes, and studentized residuals obtained by the DL method of the data set with 40 studies.

As shown in Table 1, the residual value (-8.15) of the study by Çil (2016) is less than -2.58. Accordingly, this study is an outlier for the meta-analysis. The data sets with an outlier are as follows: The first data set was formed with the first 5 studies, the second data set with the first 10 studies, the third data set with the first 20 studies, and the fourth data set with all 40 studies. After removing Çil's (2016) study, to which the outlier belongs, from these data sets, data sets without an outlier were formed through 4, 9, 19, and 39 studies. The effect sizes of the remaining 39 studies, the variances of the effect sizes, and the studentized residuals calculated according to the DL method are shown in Table 2.

As shown in Table 2, there is no outlier. In other words, while there is an outlier in the meta-analysis of 40 studies, it is seen that the residual values of the remaining 39 studies are at the acceptable level as a result of iterated analyzes after the removal of the detected outlier from the data set. Further, when Tables 1 and 2 are compared, it is seen that all the residual values of all studies show differences.

Figures 1 and 2 display the funnel plots and the radial plots were drawn to examine the heterogeneity and outliers by graphically evaluating the outcomes of the meta-analysis in Tables 1 and 2.

As displayed in Figure 1, one study diverges from the other studies by points away from the triangular region in the meta-analysis of 40 studies. In the funnel plot generated for the remaining 39 studies after the study with an outlier was removed from the data set, it is seen that all points remain in the triangular region.

As shown in Figure 2, as a result of the meta-analyses

Table 1. Effect size, variance and residual values in the data set with an outlier.

Study number	Study name	Effect size*	Variance	Residual**
1	Çil	-1.3190	0.0253	-8.15
2	Adiloğulları	0.1647	0.0367	0.84
3	Alkan	0.1105	0.0157	0.76
4	Arıkan	0.1784	0.0321	0.93
5	Acun	0.2469	0.0361	1.18
6	Ataç	-0.1101	0.0183	-0.33
7	Ayvaz	-0.0238	0.0186	0.09
8	Başören	-0.2140	0.0434	-0.66
9	Cemaloğlu & Erdemoğlu	-0.1114	0.0083	-0.38
10	Cinay	-0.0751	0.0163	-0.17
11	Belgi	0.1211	0.0751	0.52
12	Biçen	0.0000	0.0089	0.22
13	Boyraz	0.0274	0.0103	0.37
14	Bümen	0.1492	0.0058	1.11
15	Cihan	-0.1888	0.0101	-0.79
16	Çelebi	0.0047	0.0178	0.22
17	Çelik	0.0000	0.0112	0.22
18	Çağlar & Demirbaş	0.1698	0.0189	1.02
19	Çoban & Demir	-0.1631	0.0178	-0.59
20	Diri	0.1172	0.0247	0.72
21	Kale	-0.2271	0.0194	-0.89
22	Karacan	-0.0853	0.0247	-0.20
23	Karahan	0.0299	0.0172	0.35
24	Kaya	-0.1228	0.0306	-0.35
25	Korayay	-0.0134	0.0150	0.14
26	Korkulu	-0.0511	0.0077	-0.05
27	Kuvan	0.2033	0.0158	1.23
28	Mede	-0.2800	0.0649	-0.80
29	Mumcu	0.1328	0.0264	0.77
30	Nane	-0.1120	0.0086	-0.38
31	Oruç	-0.0225	0.0601	0.07
32	Öktem	0.1280	0.0284	0.74
33	Özcan	-0.1108	0.0101	-0.37
34	Özdemir	0.1083	0.0099	0.81
35	Özdoğan	-0.2988	0.0380	-1.04
36	Özipek Karabıyık	-0.0341	0.0192	0.04
37	Özkan	-0.1541	0.0778	-0.35
38	Sürgen	0.0144	0.0084	0.30
39	Şahinkaya Güven	-0.0803	0.0074	-0.21
40	Tümkaya & Türker	-0.0093	0.0121	0.17

Note. *The difference between the averages was used as effect size measure. **Studentized residuals obtained by the DL method.

Table 2. Effect size, variance and residual values in the data set without an outlier.

Study number	Study name	Effect Size*	Variance	Residual**
1	Adiloğulları	0.1647	0.0367	0.93
2	Alkan	0.1105	0.0157	1.00
3	Arıkan	0.1784	0.0321	1.07
4	Acun	0.2469	0.0361	1.37

Table 2. Countinues.

5	Ataç	-0.1101	0.0183	-0.73
6	Ayvaz	-0.0238	0.0186	-0.08
7	Başören	-0.2140	0.0434	-0.97
8	Cemaloğlu & Erdemoğlu	-0.1114	0.0083	-1.11
9	Cinay	-0.0751	0.0163	-0.49
10	Belgi	0.1211	0.0751	0.49
11	Biçen	0.0000	0.0089	0.14
12	Boyraz	0.0274	0.0103	0.40
13	Bümen	0.1492	0.0058	2.20
14	Cihan	-0.1888	0.0101	-1.79
15	Çelebi	0.0047	0.0178	0.13
16	Çelik	0.0000	0.0112	0.12
17	Çağlar & Demirbaş	0.1698	0.0189	1.34
18	Çoban & Demir	-0.1631	0.0178	-1.14
19	Diri	0.1172	0.0247	0.83
20	Kale	-0.2271	0.0194	-1.56
21	Karacan	-0.0853	0.0247	-0.47
22	Karahan	0.0299	0.0172	0.33
23	Kaya	-0.1228	0.0306	-0.63
24	Korayay	-0.0134	0.0150	-0.01
25	Korkulu	-0.0511	0.0077	-0.45
26	Kuvan	0.2033	0.0158	1.74
27	Mede	-0.2800	0.0649	-1.05
28	Mumcu	0.1328	0.0264	0.90
29	Nane	-0.1120	0.0086	-1.10
30	Oruç	-0.0225	0.0601	-0.04
31	Öktem	0.1280	0.0284	0.84
32	Özcan	-0.1108	0.0101	-0.99
33	Özdemir	0.1083	0.0099	1.24
34	Özdoğan	-0.2988	0.0380	-1.48
35	Özipek Karabıyık	-0.0341	0.0192	-0.16
36	Özkan	-0.1541	0.0778	-0.51
37	Sürgen	0.0144	0.0084	0.30
38	Şahinkaya Güven	-0.0803	0.0074	-0.81
39	Tümkaya & Türker	-0.0093	0.0121	0.03

Note. *The difference between the averages was used as effect size measure. **Studentized residuals obtained by the DL method.

performed under the FEM and the REM, the point under the shaded area for both FEM and REM in the analyzes with 40 studies can be the study with an outlier in the meta-analysis. After the removal of the study with an outlier, it is seen that the values of all studies are in the shaded area of the graph, and there is no outlier.

Data analysis

In the current study, 72 meta-analyses were carried out to answer the research problem. Accordingly, nine meta-analyses methods were compared (one FEM method, eight REM methods) by examining teachers' emotional burnout in terms of gender variable in the data sets which met the inclusion criteria of the meta-analysis. The number of studies included in the meta-analysis was 5, 10, 20, and 40 for meta-analyses with an outlier, and was 4, 9, 19, and 39 for meta-analyses without an outlier. Data sets with an outlier were designed in such a way that each data set contains one outlier. The primary study to which this outlier belongs is common to all data sets with an outlier. Data sets without an outlier were designed by testing the graphical (funnel plots and radial graphs) and statistical values (studentized residuals obtained by the DL method) for the determination of outliers and removing the outlier whose presence was determined based on the results obtained from the data

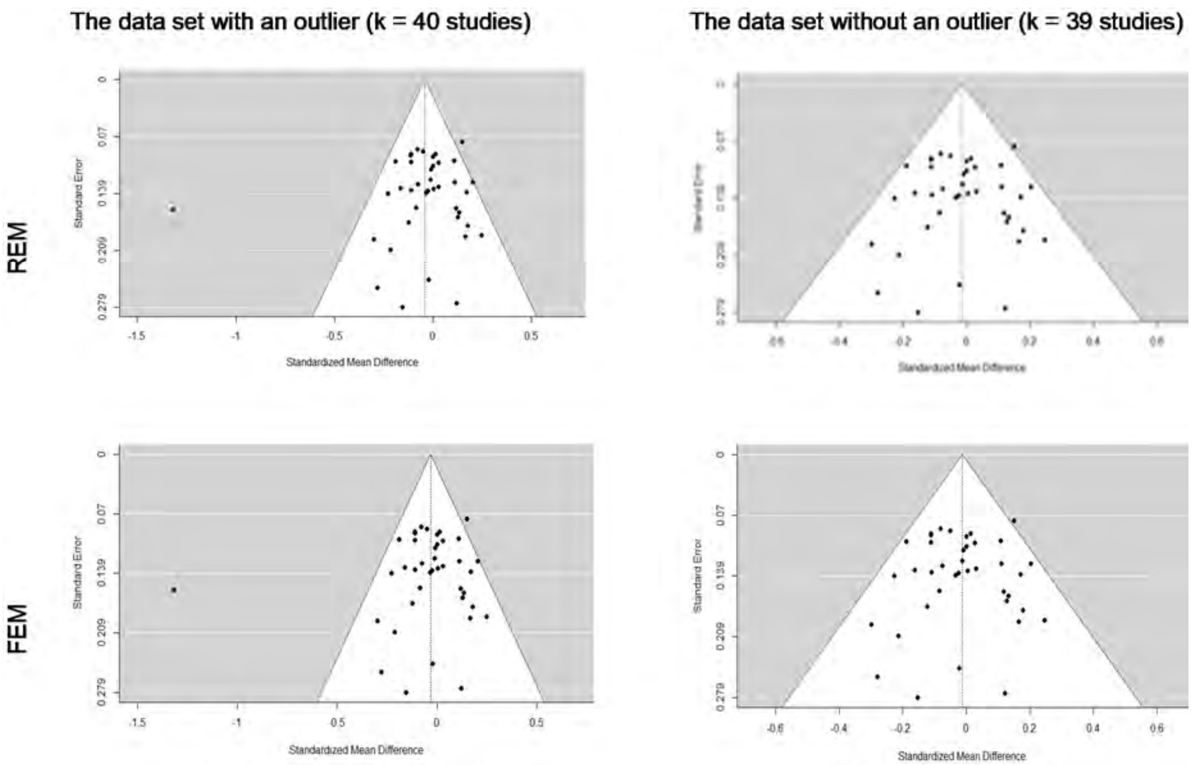

Figure 1. Funnel plots generated for the data set with and without an outlier.

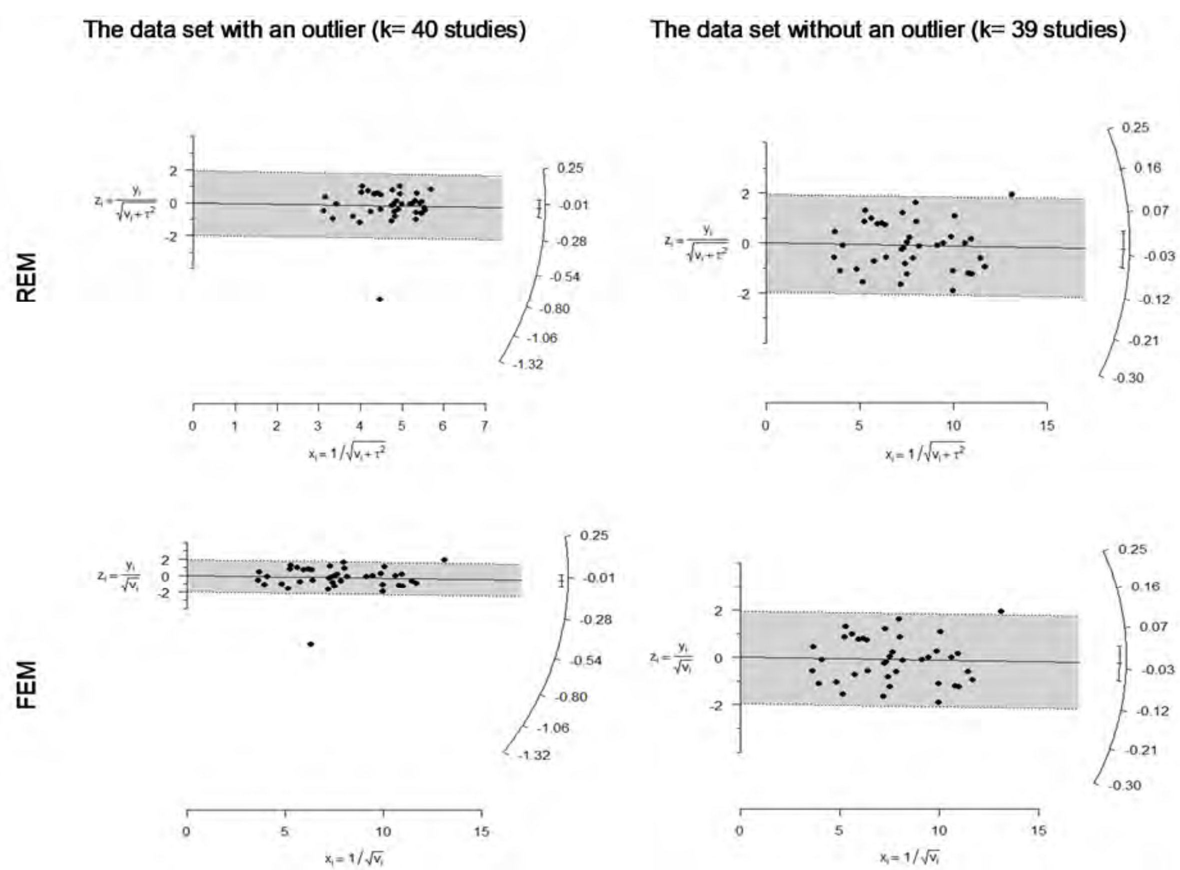

Figure 2. Radial plots generated for the data set with and without an outlier.

set. A series of meta-analyses were performed on the data sets consisting of studies with and without an outlier through the FEM method and such methods as DL, ML, REML, EB, SJ, PM, HS and HE under the REM. All the outcomes of the meta-analyses were compared in terms of estimated common effect size, confidence interval coverage rate, and measures of heterogeneity including Cochran's *Q* Test, *Higgins' and Thompson's* I^2 test, and the parameter τ^2. Their performances were evaluated, as well. However, since the meta-analysis under the FEM does not include the τ^2 statistical value, the comparisons made over this value only include the methods under REM. Analysis of the research data was conducted through the R program. In the meta-analyses with FEM and REM, R with the *metafor* package by Viechtbauer (2010) was used.

FINDINGS

In this section, findings obtained from the analyses are presented. Table 3 displays the results obtained from the meta-analysis of 5 studies, one of which includes an outlier to compare the FEM and eight estimation methods such as DL, ML, REML, EB, SJ, PM, HS and HE under the REM. Table 4, on the other hand, shows the results obtained from the meta-analysis of 4 studies after the removal of the study with an outlier.

Table 3. Meta-analysis results of 5 studies with an outlier.

Method	ES	%95 lower bound	%95 upper bound	*p*	*Q*	I^2	τ^2
FEM	-.1494	-.2919	-.0070	.0398	68.8147	94.19	-
REML	-.1268	-.7172	.4636	.6739	68.8147	93.96	.4246
DL	-.1267	-.7285	.4752	.6800	68.8147	94.19	.4424
ML	-.1275	-.6562	.4012	.6364	68.8147	92.46	.3348
EB	-.1268	-.7156	.4620	.6730	68.8147	93.93	.4222
HE	-.1268	-.7139	.4603	.6721	68.8147	93.89	.4196
HS	-.1275	-.6579	.4029	.6375	68.8147	92.51	.3371
SJ	-.1268	-.7128	.4591	.6714	68.8147	93.87	.4178
PM	-.1268	-.7156	.4620	.6730	68.8147	93.93	.4222

Note. ES = Effect Size, *p* = 0.05 significance level, *Q* = Cochran's *Q* statistics, I^2 = Higgins' and Thompson's I^2 Heterogeneity statistics, τ^2 = between-study variance.

Table 4. Meta-analysis results of 4 studies without an outlier.

Method	ES	%95 lower bound	%95 upper bound	*p*	*Q*	I^2	τ^2
FEM	.1597	-.0005	.3199	.0507	.3762	0.00	-
REML	.1597	-.0005	.3199	.0507	.3762	0.00	0.00
DL	.1597	-.0005	.3199	.0507	.3762	0.00	0.00
ML	.1597	-.0005	.3199	.0507	.3762	0.00	0.00
EB	.1597	-.0005	.3199	.0507	.3762	0.00	0.00
HE	.1597	-.0005	.3199	.0507	.3762	0.00	0.00
HS	.1597	-.0005	.3199	.0507	.3762	0.00	0.00
SJ	.1599	-.0012	.3210	.0518	.3762	0.94	.0003
PM	.1597	-.0005	.3199	.0507	.3762	0.00	0.00

Note. ES = Effect Size, *p* = 0.05 significance level, *Q* = Cochran's *Q* statistics, I^2= Higgins' and Thompson's I^2 Heterogeneity statistics, τ^2= between-study variance.

Based on the meta-analysis results with an outlier in Table 3, the common effect size was estimated lower in the FEM compared to those methods under the REM. The estimations methods under the REM, however, provided very close estimations. On the other hand, according to the results of the meta-analysis without an outlier in Table 4, it was found that the common effect size was similar for all methods and was estimated to be approximately 0.160. Accordingly, it can be noted that the meta-analysis results under the FEM were much more affected by the outlier when compared to the estimations under the REM. When the true effect size was considered as 0.160 in Table 4, it can be noted that confidence intervals include this value under all methods.

When it comes to the measures of heterogeneity in Table 3 and Table 4, there is a high level of heterogeneity in the data set with an outlier based on the *Q*, I^2 and between-study variance (τ^2) as it had been expected. However, when the outlier was removed from the data set, heterogeneity disappeared in all methods except the SJ method, which gives positive τ^2 values according to the formula.

Table 5 displays the results obtained from the meta-analysis of 10 studies with an outlier. Table 6, on the other hand, shows the results obtained from the meta-analysis of 9 studies after the removal of the study with an outlier.

Table 5. Meta-analysis results of 10 studies with an outlier.

Method	ES	%95 lower bound	%95 upper bound	*p*	*Q*	I^2	τ^2
FEM	.1165	.2040	-.0291	.0090	69.81	87.11	-
REML	-.1171	-.3927	.1585	.4049	69.81	89.37	.1732
DL	-.1170	-.3681	.1332	.3584	69.81	87.11	.1391
ML	-.1173	-.3783	.1437	.3784	69.81	88.13	.1529
EB	-.1171	-.3941	.1598	.4072	69.81	89.48	.1751
HE	-.1171	-.3951	.1609	.4090	69.81	89.56	.1766
HS	-.1177	-.3524	.1170	.3256	69.81	85.26	.1191
SJ	-.1171	-.3946	.1604	.4082	69.81	89.52	.1760
PM	-.1171	-.3941	.1598	.4072	69.81	89.48	.1751

Note. ES = Effect Size, *p* = 0.05 significance level, *Q* = Cochran's *Q* statistics, I^2 = Higgins' and Thompson's I^2 Heterogeneity statistics, τ^2= between-study variance.

Table 6. Meta-analysis results of 9 studies without an outlier.

Method	ES	%95 lower bound	%95 upper bound	*p*	*Q*	I^2	τ^2
FEM	-.0137	-.1048	.0775	.7688	7.6959	0.00	-
REML	-.0137	-.1048	.0775	.7688	7.6959	0.01	.0000
DL	-.0137	-.1048	.0775	.7688	7.6959	0.00	.0000
ML	-.0137	-.1048	.0775	.7689	7.6959	0.02	.0000
EB	-.0137	-.1048	.0775	.7689	7.6959	0.02	.0000
HE	-.0123	-.1053	.0807	.7951	7.6959	2.99	.0006
HS	-.0137	-.1048	.0775	.7688	7.6959	0.00	.0000
SJ	.0000	-.1181	.1180	.9997	7.6959	35.62	.0112
PM	-.0137	-.1048	.0775	.7689	7.6959	0.02	.0000

Note. ES = Effect Size, *p* = 0.05 significance level, *Q* = Cochran's *Q* statistics, I^2 = Higgins' and Thompson's I^2 Heterogeneity statistics, τ^2= between-study variance.

According to the outcomes of the meta-analysis with an outlier displayed in Table 5, the common effect size was estimated higher in the FEM compared to the REM. The estimates under the REM were very close to each other. On the other hand, the outcomes of meta-analysis without an outlier displayed in Table 6 show that the effect size was similar and estimated as -0.014 for all methods except SJ and HE methods. Accordingly, it can be noted that the meta-analysis results under the FEM and the meta-analysis results using SJ and HE, respectively, under the REM were more affected by the outlier than the other methods. Furthermore, when the outcomes displayed in Tables 5 and 6 were compared to the ones in Tables 3 and 4, it can be concluded that the increase in the number of studies included in the meta-analysis may have reduced the effect of the outlier on the estimation of the effect size. When the true effect size is accepted as -0.014 for Table 6, it is seen that the confidence intervals included the estimated effect sizes in all data sets and all methods. On the other hand, when Tables 5 and 6 are examined in terms of measures of heterogeneity, there is a high level of heterogeneity as it has been expected. However, when the outlier is removed from the data set, it is seen that heterogeneity disappears in all methods except SJ and HE methods.

Table 7 displays the results obtained from the meta-analysis of 20 studies with an outlier. Table 8, on the other hand, shows the results obtained from the meta-analysis of 19 studies after the removal of the study with an outlier.

According to the outcomes of the meta-analysis with an outlier displayed in Table 7, the common effect size was

Table 7. Meta-analysis results of 20 studies with an outlier.

Method	ES	%95 lower bound	%95 upper bound	p	Q	I^2	τ^2
FEM	-.0330	-.0875	.0216	.2365	86.4168	78.01	-
REML	-.0475	-.1848	.0898	.4976	86.4168	83.09	.0775
DL	-.0475	-.1684	.0735	.4419	86.4168	78.01	.0560
ML	-.0475	-.1808	.0858	.4847	86.4168	82.00	.0720
EB	-.0475	-.1881	.0931	.5079	86.4168	83.88	.0822
HE	-.0475	-.1901	.0952	.5142	86.4168	84.37	.0853
HS	-.0474	-1646	.0698	.4277	86.4168	76.52	.0515
SJ	-.0475	-.1908	.0958	.5161	86.4168	84.52	.0862
PM	-.0475	-.1881	.0931	.5079	86.4168	83.88	.0822

Note. ES = Effect Size, p = 0.05 significance level, Q = Cochran's Q statistics, I^2 = Higgins' and Thompson's I^2 Heterogeneity statistics, τ^2= between-study variance.

Table 8. Meta-analysis results of 19 studies without an outlier.

Method	ES	%95 lower bound	%95 upper bound	p	Q	I^2	τ^2
FEM	.0077	-.0477	.0632	.7844	18.8961	4.74	-
REML	.0073	-.0544	.0689	.8167	18.8961	15.30	.0028
DL	.0074	-.0498	.0647	.7990	18.8961	4.74	.0008
ML	.0073	-.0529	.0674	.8127	18.8961	11,80	.0021
EB	.0075	-.0494	.0644	.7963	18.8961	3.73	.0006
HE	.0077	-.0477	.0632	.7844	18.8961	0.00	.0000
HS	.0077	-.0477	.0632	.7844	18.8961	0.00	.0000
SJ	.0084	-.0636	.0804	.8190	18.8961	35.51	.0085
PM	.0075	-.0493	.0643	.7959	18.8961	3.60	.0006

Note. ES = Effect Size, p = 0.05 significance level, Q = Cochran's Q statistics, I^2= Higgins' and Thompson's I^2 Heterogeneity statistics, τ^2= between-study variance.

estimated higher in the FEM compared to the REM. The estimates under the REM were very close to each other. On the other hand, the outcomes of meta-analysis without an outlier displayed in Table 8 show that the common effect size is estimated to be approximately 0.007 by REML, DL and ML methods, while it is estimated to be approximately 0.008 by other methods. In addition, when the outcomes displayed in Tables 7 and 8 were compared to the ones in Tables 3 and 4, it can be concluded that the increase in the number of studies included in the meta-analysis may have reduced the effect of the outlier on the estimation of the effect size. There is evidence in Table 8 to suggest that the confidence intervals included the estimated effect sizes in all data sets and all methods. On the other hand, when the outcomes in Table 7 are compared to Tables 3 and 5 in terms of measures of heterogeneity, it can be noted that the increase in the number of studies included in the meta-analysis reduces the heterogeneity. After the removal of the outlier, it is seen that heterogeneity disappeared in HE and HS methods.

Table 9 displays the results obtained from the meta-analysis of 40 studies with an outlier. Table 10, on the other hand, shows the results obtained from the meta-analysis of 39 studies after the removal of the study with an outlier.

According to the outcomes of the meta-analysis with an outlier displayed in Table 9, the common effect size was estimated higher in the FEM compared to the REM. The estimates under the REM were very close to each other. On the other hand, the outcomes of meta-analysis without an outlier displayed in Table 10 show that the common effect size is similar for all methods and estimated as -0.013. Furthermore, it can be noted that the FEM was less affected by the outlier as the number of studies included in the meta-analysis increased compared to other data sets with fewer studies. Besides, when the outcomes displayed in Tables 9 and 10 were compared to the ones in Tables 3 and 4, it can be concluded that the increase in the number of studies included in the meta-analysis may have reduced the effect of the outlier on the estimation of the effect size. According to Table 10, it is suggested that the confidence intervals included the estimated effect sizes in all data sets and all methods. Tables 9 and 10 show that the data set with an outlier indicates heterogeneity as it has been

Table 9. Meta-analysis results of 40 studies with an outlier.

Method	ES	%95 lower bound	%95 upper bound	p	Q	I^2	τ^2
FEM	-.0327	-.0712	.0059	.0968	101.4617	61.56	-
REML	-.0421	-.1116	.0274	.2354	101.4617	66.24	.0307
DL	-.0414	-.1067	.0238	.2130	101.4617	61.56	.0250
ML	-.0419	-.1102	.0264	.2290	101.4617	65.00	.0290
EB	-.0425	-.1156	.0305	.2539	101.4617	69.54	.0357
HE	-.0427	-.1170	.0316	.2602	101.4617	70.56	.0374
HS	-.0413	-.1055	.0230	.2078	101.4617	60.36	.0238
SJ	-.0431	-.1208	.0347	.2779	101.4617	73.22	.0427
PM	-.0425	-.1156	.0305	.2539	101.4617	69.54	.0357

Note. ES = Effect Size, p = 0.05 significance level, Q = Cochran's Q statistics, I^2= Higgins' and Thompson's I^2 Heterogeneity statistics, τ^2= between-study variance.

Table 10. Meta-analysis results of 39 studies without an outlier.

Method	ES	%95 lower bound	%95 upper bound	p	Q	I^2	τ^2
FEM	-.0127	-.0515	.0262	.5232	34.9645	0.00	-
REML	-.0128	-.0526	.0271	.5298	34.9645	3.56	.0006
DL	-.0127	-.0515	.0262	.5232	34.9645	0.00	.0000
ML	-.0127	-.0521	.0266	.5263	34.9645	1.82	.0003
EB	-.0127	-.0515	.0262	.5232	34.9645	0.00	.0000
HE	-.0127	-.0515	.0262	.5232	34.9645	0.00	.0000
HS	-.0127	-.0515	.0262	.5232	34.9645	0.00	.0000
SJ	-.0129	-.0629	.0372	.6143	34.9645	34.45	.0081
PM	-.0127	-.0515	.0262	.5232	34.9645	0.00	.0000

Note. ES = Effect Size, p = 0.05 significance level, Q = Cochran's Q statistics, I^2= Higgins' and Thompson's I^2 Heterogeneity statistics, τ^2= between-study variance.

expected. On the other hand, when the outcomes in Table 9 are compared to Tables 3, 5 and 7 in terms of measures of heterogeneity, it can be noted that the increase in the number of studies included in the meta-analysis reduces the heterogeneity. After the removal of the outlier, it is seen that heterogeneity disappeared in all methods except for REML, ML and SJ methods.

DISCUSSION AND CONCLUSION

In the current study, it was aimed to compare the performances of the FEM and such methods as DL, ML, REML, EB, SJ, PM, HS and HE under the REM in the meta-analysis studies conducted through 5, 10, 20 and 40 studies with an outlier and 4, 9, 19 and 39 studies without an outlier. Based on the findings, it was concluded that when outliers in the studies included in the meta-analysis, regardless of the number of studies in the data set, the common effect size estimations were quite close to each other in the methods under REM, but were relatively higher or lower estimations under FEM. In this sense, it can be concluded that REM is more functional compared to FEM when *meta-analysis* data sets with an outlier. In line with this finding, Borenstein et al. (2009) recommended using REM in the studies collected from the literature, regardless of the heterogeneity test. FEM, on the other hand, is recommended when the effect sizes of the studies in the meta-analysis show homogenous distributions (Borenstein et al., 2010; Cooper, 2010; Lipsey and Wilson, 2001). The evidence for this conclusion can be attributed to the fact that outliers existing in *meta-analysis* data sets may increase heterogeneity between studies. Accordingly, it can be noted that this heterogeneity increased by outliers can be considered a great disadvantage for the estimation of common effect sizes under the FEM.

There is also evidence to suggest that when there is no outlier in the studies included in the meta-analysis, the common effect size is similarly estimated under all methods. This finding is supported by Langan et al. (2015) who concluded that although the measures of heterogeneity estimated by different methods used in the meta-analysis may differ, no significant difference was found in the estimations of the common effect size. In other words, when the data sets with and without an outlier are compared, FEM was found to be the most differentiating method for estimating the common effect size. SJ and HE methods, on the other hand, differed more than other REM methods in the dataset consisting of only 10 studies. Accordingly, it can be noted that FEM as well as SJ and HE methods for meta-analysis with fewer studies are more nonfunctional in the common effect size estimation than other methods. Petropoulou and Mavridi (2017) supported this result, by concluding that the SJ method performed worse than other estimators.

Findings showed that such measures of heterogeneity as *Q*, I^2 and between-study variance (τ^2) obtained high values in all meta-analyses with outliers as had been expected. In addition, in data set consisting of 4 studies without an outlier, heterogeneity disappears in all methods except for the SJ method, in data set consisting of 9 studies without an outlier, heterogeneity disappears in all methods except SJ and HE methods, in the data set consisting of 19 studies without an outlier, heterogeneity disappears only in HE and HS methods, in the data set consisting of 39 studies without an outlier, heterogeneity disappeared in all methods except for REML, ML and SJ methods. Accordingly, it was concluded that heterogeneity was eliminated when only the HS method was used for all data sets where the outlier was removed from the data set and the meta-analysis was repeated. In this regard, it can be noted that the HS method is more appropriate than other methods. Petropoulou and Mavridi (2017) supported this finding, by concluding that the performance of the HS method was high in meta-analysis studies in the absence of outliers. However, there have been different findings in the previous literature. Bowden et al. (2011), DerSimonian and Kacker (2007), as well as Novianti et al. (2014) recommended the use of the PM method, whereas Veroniki et al. (2016) and Viechtbauer (2005) were in favor of the REML method rather than others. Moreover, in their studies conducted by Langan et al. (2015), it was concluded that the DL and the PM methods were more consistent with each other in terms of indicating heterogeneity, and the estimates of heterogeneity under these methods were similar. This finding was in line with the current study about the meta-analysis outcomes without an outlier.

Another important finding obtained from the current study was that the methods used in the estimation of the common effect size were slightly affected by the outlier, especially when the number of studies included in the meta-analysis was 20 or more. In addition, it was found that the increase in the number of studies included in the meta-analysis reduced the effect of the outlier on the effect size estimation and decreased the heterogeneity. Umaroğlu (2020), partly in line with the result of the current study, noted that when the measure of heterogeneity was τ^2, the HS and ML methods performed better if the number of studies included in the meta-analysis was few, but as the number of studies in the meta-analysis increased, the difference between the methods disappeared and the methods made close estimations to each other.

When the confidence interval coverage rates of the meta-analysis methods were examined, it was determined that the confidence intervals included the estimated effect sizes in all data sets through all methods.

Based on the findings in the current study, the use of REM in the data set was recommended in the meta-analysis data sets with an outlier. This study should, therefore, be of value to practitioners wishing to conduct a meta-analysis. Furthermore, it can be noted that meta-analysis studies can be carried out with at least 20 studies in order to ensure that the estimated common effect size is affected by the outlier as little as possible.

In the data sets used in the current study, there was one outlier. Future studies, especially simulation-based, are needed to fully understand the performances of the meta-analysis methods when there are outliers in more than one data point. In addition, methods under the *Robust* variance estimation and different measures of heterogeneity can be included in future studies. Besides, such parameters as data set size, number and direction of outliers in data can be diversified through simulations studies to be able to make more detailed comparisons.

REFERENCES

References marked with an asterisk (*) indicate studies included in the meta-analysis.

*Acun, M. (2010). The investigation of the biology teachers occupational burnout levels according to some variables [Unpublished Master's Dissertation]. Dicle University.

*Adiloğulları, G. E. (2013). Examination of relation between emotional intelligence levels and professional burnout levels of physical education teachers [Unpublished Master's Dissertation]. Kahramanmaras Sutcu Imam University.

*Alkan, M. F. (2014). Burnout level and causes of middle school teachers [Unpublished Master's Dissertation]. Istanbul Aydin University.

*Arıkan, S. (2007). The investigation of the burnout levels and reasons of the classroom teachers who are working at the first stage of Muğla centrum primary schools [Unpublished Master's Dissertation]. Mugla University.

*Ataç, H. D. (2015). The investigation of burnout level according to the audit focus in primary school teachers [Unpublished Master's Dissertation]. Halic University.

*Ayvaz, U. (2015). Examination of physical education teachers' levels of burnout in terms of specific variables [Unpublished Master's Dissertation]. Yeditepe University.

*Başören, M. (2005). Examining the level of the exhaustion of the guide teachers according to some variables (Zonguldak example) [Unpublished Master's Dissertation]. Zonguldak Karaelmas University.

*Belgi, S. (2016). To researching the self-sufficiency of high school counselors and the burnout of high school counselors [Unpublished Master's Dissertation]. Nisantası University.

*Biçen, H. (2014). Burnout syndrome in the newly appointed teachers [Unpublished Master's Dissertation]. Dicle University.

*Boyraz, S. (2015). Investigation of the relationship between the levels of conflict management strategies with teacher burnout [Unpublished Master's Dissertation]. Istanbul Sabahattin Zaim University.

*Bümen, N. T. (2010). The relationship between demographics, self

efficacy, and burnout among teachers. Eurasian Journal of Educational Research, 40: 16-35.
*Cemaloğlu, N., and Erdemoğlu-Şahin, D. (2007). A study of the teacher's burnout level according to various variables. Kastamonu Education Journal, 15(2): 465-484.
*Cihan, B. B. (2011). Examination and comparison of occupational burnout levels of physical education teachers working at primary schools in different provinces [Unpublished Master's Dissertation]. Gazi University.
*Cinay, F. (2015). The relationship between primary school teachers' professional burnout and organizational citizenship behaviors [Unpublished Master's Dissertation]. Okan University.
*Çağlar, Ç., and Demirtaş, H. (2011). Private course teachers' burnout and job satisfaction. E-İnternational Journal of Educational Research, 2(2): 30-49.
*Çelebi, E. (2013). The burnout levels of the teachers working at special educational institutions in Elazig and Malatya city centres and related factors [Unpublished Doctoral Dissertation]. Firat University.
*Çelik, M. (2015). The relationship between teachers' occupational professionalism and burnout [Unpublished Master's Dissertation]. Dumlupınar University.
***Çil**, F. (**2016**). Researching the level of burnout of the computer educators in terms of some variables [Unpublished Master's Dissertation]. Okan University.
*Çoban, A. E., and Demir, A. (2004). Examination of the burnout levels of psychological counselors working in the Southeast Anatolia Region and the relationship between some demographic variables and burnout. Cukurova University Faculty of Education Journal, 2(28): 20-28.
*Diri, M. S. (2015). The effect to secondary school teachers' job satisfaction to the occupational burnout level [Unpublished Master's Dissertation]. Adnan Menderes University.
*Kale, F. (2007). To examine the physical education teacher's vocational satisfaction and the level of exhaustion in aspect of some kinds of variables [Unpublished Master's Dissertation]. Nigde University.
*Karacan, A. (2012). An investigation about the level of burnout of the teachers who work with individuals can be trained and taught at the private educational institutes: The case of Anatolian side of Istanbul [Unpublished Master's Dissertation]. Maltepe University.
*Karahan, Ş. (2008). The analysis of self-efficacy perception and burnout level of special education school educators [Unpublished Master's Dissertation]. Marmara University.
*Kaya, A. (2010). Burnout level of occupational groups which serve disabled people for the private and public institutions in the province of Isparta [Unpublished Master's Dissertation]. Suleyman Demirel University.
*Korayay, F. D. (2014). Professional burnout levels of the teachers in primary schools [Unpublished Master's Dissertation]. Dokuz Eylul University.
*Korkulu, N., Feyzioğlu, B., Özenoğlu-Kiremit, H., and Aladağ, E. (2012). Teachers' burnout levels in terms of some variables. Educational Sciences: Theory and Practice, 12(3): 1813-1831.
*Kuvan, Ö. (2009). Informatic technology teacher's problems and their level of being exhausted [Unpublished Master's Dissertation]. Sakarya University.
*Mede, E. (2009). An analysis of relations among personal variables, perceived self-efficacy and social support on burnout among Turkish EFL teachers. Inonu University Journal of the Faculty of Education, 10(2): 39-52.
*Mumcu, L. (2014). Physical education teachers' job satisfaction and burnout levels of investigation [Unpublished Master's Dissertation]. Karadeniz Teknik University.
*Nane, A. (2013). The relation level between the burnout of teachers and their perception of leadership behaviors of their administrators: Sample of Mersin province [Unpublished Master's Dissertation]. C+ag University.
*Oruç, S. (2007). Examining the burnout levels of teachers employed in the special education schools with some variables (The sample of Adana) [Unpublished Master's Dissertation]. Cukurova University.
*Öktem, E. (2009). An investigation of the burnout degrees of primary school teachers in terms of some specific variables (Afyonkarahisar-Sandıklı sample) [Unpublished Master's Dissertation]. Ege University.
*Özcan, T. (2008). A study of burnout levels of teachers, and attendants in Pendik district, according to some demographic variables [Unpublished Master's Dissertation]. Yeditepe University.
*Özdemir, T. (2009). The relationship between the undesired student behavior in the classroom and teachers' level of burnout [Unpublished Master's Dissertation]. Abant İzzet Baysal University.
*Özdoğan, H. (2008). Job enthusiasm of physical education teachers [Unpublished Master's Dissertation]. Cumhuriyet University.
*Özipek Karabıyık, A. (2006). The reasons and the level of professional burnout in teachers, work in secondary schools [Unpublished Master's Dissertation]. Trakya University.
*Özkan, Ş. Y. (2007). To research professional exhausted level of physical education teachers who work in the province of Niğde [Unpublished Master's Dissertation]. Nigde University.
*Sürgen, S. (2014). *Correlation between class teacher's focus of control and level of burnout* [Unpublished Master's Dissertation]. Balıkesir University.
*Şahinkaya Güven, S. (2013). Primary school teachers school principals occupational burnout educational leader relationship between the detection levels [Unpublished Master's Dissertation]. Yuzuncu Yıl University.
*Tümkaya, S., and Türker, P. (2010). The analysis of primary school teachers' perceptions of problem behavior levels according to the socio-demographic variants and burnout. Çukurova University Journal of the Faculty of Education, 3(38): 92-109.
Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (**2009**). Introduction to meta-analysis. Wiley.
Borenstein, M., Hedges, L. V., Higgins, J. P. T., and Rothstein, H. R. (**2010**). A basic introduction to fixed effect and random effect models for meta-analysis. Research Synthesis Methods, 1(1): 97-111.
Bowden, J., Tierney, J. F., Copas, A. J., and Burdett, S. (**2011**). Quantifying, displaying and accounting for heterogeneity in the meta-analysis of RCTs using standard and generalised Q statistics. BMC Medical Research Methodology, 11(1): 1-12.
Cohen, J. (**1977**). Statistical power analysis for the behavioral sciences. Lawrence Erlbaum Associates.
Cooper, H. (**2010**). Research synthesis and meta-analysis: A step-by-step approach. Sage.
Demir, S., and **Başol**, G. (**2014**). Effectiveness of computer-assisted mathematics education (CAME) over academic achievement: A meta-analysis study. Educational Sciences: Theory & Practice, 14(5): 2013-2035.
DerSimonian, R., and **Kacker**, R. (**2007**). Random-effects model for meta-analysis of clinical trials: An update. Contemporary Clinical Trials, 28(2): 105-114.
Fraenkel, J. R., Wallen, N. E., and Hyun, H. H. (**2012**). How to design and evaluate research in education (8th ed.). McGraw-Hill.
Glass, G. V. (**1976**). Primary, secondary, and meta-analysis of research. Educational Researcher, *5*(10): 3-8.
Gumedze, F. N., and **Jackson**, D. (**2011**). A random effects variance shift model for detecting and accommodating outliers in meta-analysis. BMC Medical Research Methodology, 11(1): 11-19.
Hedges, L. V., and **Olkin**, I. (**1985**). Statistical methods for meta-analysis. Academic Press.
Hedges, L. V., and **Vevea**, J. L. (**1998**). Fixed- and random-effects models in meta-analysis. Psychological Methods, 3(4): 486-504.
Higgins, J. P. T., and **Thompson** S. G. (**2002**). Quantifying heterogeneity in a meta-analysis. Statistics in Medicine, 21(11): 1539-1558.
Higgins, J., Tompson, S., Deeks, J., and Altman, D. (**2003**). A meta-analysis on the effectiveness of smart-learning. BMJ, 327(1): 557-560.
Hunter, J. E., and **Schmidt** F. L. (**2004**). Methods of meta-analysis: Correcting error and bias in research findings (2nd ed.). Sage.
Langan, D., Higgins, J. P. T., and Simmonds, M. (**2015**). An empirical comparison of heterogeneity variance estimators in 12894 meta-analyses. Research Synthesis Methods, 6(2): 195-205.
Lin, L, Chu, H., and Hodges J. S. (**2017**). Alternative measures of between-study heterogeneity in meta-analysis reducing the impact of outlying studies. Biometrics, 73(1): 156-166.
Lipsey, M. W., and **Wilson**, D. B. (**2001**). Practical meta-analysis.

Sage.
Normand, S. L. T. (**1999**). Tutorial in biostatistics meta-analysis: Formulating, evaluating, combining, and reporting. Statistics in Medicine, 18: 321-359.
Novianti, P. W., Roes, K. C., and van der Tweel, I. (**2014**). Estimation of between-trial variance in sequential meta-analyses: A simulation study. Contemporary Clinical Trials, 37(1): 129-138.
Petropoulou, M., and Mavridis, D. (2017). A comparison of 20 heterogeneity variance estimators in statistical synthesis of results from studies: A simulation study. Statistics in Medicine, 36(27): 4266-4280.
Riley, R. D., Higgins, J. P., and Deeks, J. J. (**2011**). Interpretation of random effects meta-analyses. BMJ, 342: 964-967.
Rudy, A. C. (**2001**). A meta-analysis of the treatment of anorexia nervosa: A proposal. New York: Ithaca College.
Schmid, E. J., Koch, G. G, and LaVange L. M. (**1991**). An overview of statistical issues and methods of meta-analysis. Journal of Biopharmaceutical Statistics, 1(1): 103-120.
Schmidt, F. (**2008**). Meta-analysis: A constantly evolving research integration tool. Organizational Research Methods, 11: 96-113.
Schwarzer, G., Carpenter, J. R., and Rücker, G. (**2015**). *Meta-Analysis with R*. Springer International Publishing. https://doi.org/10.1007/978-3-319-21416-0
Thorlund, K., Wetterslev, J., Awad, T., Thabane, L., and Gluud, C. (**2011**). Comparison of statistical inferences from the DerSimonian–Laird and alternative random-effects model meta-analyses–an empirical assessment of 920 Cochrane primary outcome meta-analyses. Research Synthesis Methods, 2(4): 238-253.
Umaroğlu, M. M. (**2020**). Examination of outliers in meta-analysis and comparison of methods used in outliers [Unpublished Doctoral Dissertation]. Hacettepe University.
Veroniki, A. A., Jackson, D., Viechtbauer, W., Bender, R., Bowden, J., Knapp, G., Kuss, O., Higgins, J. P., Langan, D., and Salanti, G. (**2016**). Methods to estimate the between-study variance and its uncertainty in meta-analysis. Research Synthesis Methods, 7(1): 55-79.
Viechtbauer, W. (**2005**). Bias and efficiency of meta-analytic variance estimators in the random-effects model. Journal of Educational and Behavioral Statistics, 30(3): 261-293.
Viechtbauer, W. (**2010**). Conducting meta-analyses in R with the metafor package. Journal of Statistical Software, 36(3): 1-48.
Viechtbauer, W., and **Cheung**, M. W. L. (**2010**). Outlier and influence diagnostics for meta-analysis. Research Synthesis Methods, 1(2): 112-125.

Citation: Demir, S., and Doğuyurt, M. F. (2022). A comparison of fixed and random effect models by the number of research in the meta-analysis studies with and without an outlier. African Educational Research Journal, 10(3): 277-290.