Article History
Received: 19.11.2020
Received in revised form: 05.03.2021
Accepted: 04.04.2021
Article Type: Research Article

International Journal of Contemporary Educational Research (IJCER)

www.ijcer.net

The Effect of Professional Development Programs on Teachers' Self-Efficacy Beliefs in Science Teaching: A Meta-Analysis Study

Yunus Emre Baysal[1], Fatma Mutlu[2]
[1]Inonu University, 0000-0001-6125-3165
[2]Inonu University, 0000-0002-8643-1236

To cite this article:

Baysal, Y. E. & Mutlu, F. (2021). The Effect of Professional Development Programs on Teachers' Self-Efficacy Beliefs in Science Teaching: A Meta-Analysis Study. *International Journal of Contemporary Educational Research,* *8*(4), 1-15. https://doi.org/10.33200/ijcer.828528

International Journal of Contemporary Educational Research
Volume 8, Number 4, December 2021, Page 1-15 ISSN: 2148-3868

The Effect of Professional Development Programs on Teachers' Self-Efficacy Beliefs in Science Teaching: A Meta-Analysis Study

Yunus Emre BAYSAL[1*], Fatma MUTLU[1]
[1]Inonu University

Abstract

This study examined the effect size of professional development programs on teachers' science teaching by meta-analysis method. For this purpose, literature review was made on existing national and international studies. In this context, master's thesis, doctoral dissertation, and articles conducted between 1990 and 2019 were appropriate for the research problem and had statistical data to be included in the meta-analysis study were reviewed and investigated in Turkish and English from national and international databases. As a result of the literature review, 1072 national and international related studies were reached. Among these studies, it was determined that 14 studies (N=969) met the inclusion criteria. The effect sizes and the combined effect size of the studies were calculated using "Comprehensive Meta-Analysis v2.0 (CMA) Statistical Package Program".

In the studies combined with random effects in the model, the effect size of professional development programs on teachers' self-efficacy beliefs in science teaching was found to be "moderate". As a result of the heterogeneity test, the study was found to have a high level of heterogeneity.Moderator analysis was performed in order to determine the sources of high levels of heterogeneity between studies. As a result of the ANOVA similarity analysis conducted for the publication type and branch categorical moderators, it was determined that the effect sizes of teachers' science teaching self-efficacy beliefs differed significantly according to the branch variable. As a result of the meta-regression analysis conducted for publication year, sample size, and application period, which were assessed as continuous variables, it was found that publication year and sample size as moderators caused significant differences in the effect sizes of science education teachers' self-efficacy beliefs. Teacher professional development programs were found to have a positive influence on science education teachers' self-efficacy beliefs, and teachers were suggested to be encouraged to participate in such programs.

Keywords: Professional development programs, Science teaching self-efficacy Beliefs, Self-efficacy, Effect size, Meta-analysis

Introduction

The teaching profession, one of the most important components of the education system, is carried out by people who manage the learning and teaching process at different levels of education and have high qualifications in terms of general culture, professional knowledge and skills, and expertise (Ada & Unal, 2009).The competencies of a teacher enhance the quality and qualification of the education. Professional development is seen as an important component of education reforms in recent years in increasing teacher quality (Borko, 2004; Hiebert, Gallimore & Stigler, 2002; Higgins &Parkons, 2009; Loucks-Horsley, Love, Stiles, Mundry, Love & Hewson, 2003). Professional development of teachers is accepted as a dynamic process that occurs throughout the teacher's life and covers different learning types (Menezes, 2011; Özer, 2005; Villegas-Reimers, 2003). According to Blandford (2000), professional development provides significant contributions to increasing professional performance, correcting inefficient practices, facilitating the implementation of educational policies, and ensuring change. This is supported by studies showing that the lack of professional development is an important reason why primary school teachers cannot teach science course effectively (Madden, 2016; Martin, 2016; Prentiss Bennett, 2016; Trimmell 2015).

* Corresponding Author: *Yunus Emre baysal, y.emrebaysal@gmail.com*

In this context, some programs are prepared and applied for teachers to ensure their professional development. Studies have shown that teachers' teaching skills and performances would be improved with long-term professional development programs applied practically in teaching environments (Banilower, Boyd, Pasley, & Weiss, 2005; Garet, Porter, Desimone, Birman &Yoon, 2001; Guskey, 2003). According to Avalos (2011), professional development programs for teachers should teach new things to teachers and teach new things to teachers and give them the chance to practice what they have learned. Desimone (2009) sees the professional development programs as many activities and interaction experiences where teachers can increase their knowledge and skills and improve their teaching practices. A change that took place in teachers' knowledge, beliefs, attitudes and self-efficacies was that teachers showed a strong relationship with the classroom practices (Fishman, Marx, Best & Tal, 2003; Haney, Lumpe & Czerniak, 2002; Luft & Hewson, 2014). Professional development and teaching practices have been shown as a way in studies conducted to improve self-efficacy (Posnanski, 2002, Cakiroglu, Capa-Aydin & Hoy 2012).

Self-efficacy is based on Bandura's social cognitive behavior and motivation theory and defined as the belief that individuals can overcome a difficult situation or task (Bandura, 1997). Self-efficacy, which is associated with many aspects of the individuals such as cognitive, affective, motivational, and decision-making, affects how individuals cope with difficulties, motivate themselves, and make important decisions about their lives (Bandura, 2002). On the other hand, self-efficacy is also evaluated as a measure of individuals' trust on abilities to perform complex tasks successfully (Mintzes, Marcum, Messerschmidt-Yates & Mark, 2013). Therefore, high self-efficacy level of teachers affects their efficient teaching performance as well as the success levels of students (Bleicher, 2006).

Self-efficacy concept defined by Bandura (1977) has been defined specifically for many areas. When this concept is applied to science teaching, it is called as "science teaching self-efficacy belief". Science teaching self-efficacy belief is defined as the teacher's belief that he/she will positively affect students' behaviors and success and can teach science effectively (Dembo & Gibson, 1985). Science education self-efficacy belief is closely related with the teachers' beliefs about science teaching and in-class activities. Teachers with a high level of science teaching self-efficacy have characteristics such as using student-centered teaching methods and techniques, sparing more time on teaching science, and showing a tendency to conduct research-based teaching.

On the other hand, teachers with low level of science teaching self-efficacy belief prefer to use teacher-centred approaches such as reading information from a book and explaining them verbally (Schriver & Czerniak, 1999). In other words, there is a positive correlation between science teaching self-efficacy beliefs and science teaching practices (Bhattacharyya, Volk & Lumpe, 2009; Czerniak & Shriver, 1994). For this reason, it is necessary to provide opportunities to increase teachers' self-efficacy beliefs (Ramakrishnan & Salleh, 2018). In this context, professional development programs involve extensive activities, including designing teaching practices (Avidav, 2000) and providing teachers opportunities to share their experiences with their colleagues and increase their professional knowledge (Vescio, Ross & Adems, 2008).

In the literature, there are studies that state that professional development programs are effective on teachers' self-efficacy beliefs in science teaching (Atia, 2012; Deniz & Akerson, 2013; Eshach, 2003; Ewing-Taylor, 2012) or they have no effect (Luera & Murray, 2016; Peters-Burton, Merz, Ramirez & Saroughi, 2015). It is evident that there are inconsistencies regarding the effect of professional development programs on science classroom teachers' self-efficacy beliefs. Therefore, there is a need to gather the studies under one umbrella to reanalyze them and make new decisions (Sağlam & Yuksel, 2007). When examining the relevant literature, there are national and international studies with independent and different results that investigate the effect of professional development programs on teachers' self-efficacy beliefs in science education; however, a meta-analysis on this topic was not found.

Aim of the Study

The purpose of the present study is to examine national and international studies that examine the effects of professional development programs on science teachers' self-efficacy beliefs through a meta-analysis. This meta-analysis is intended to provide information about the effect of in-service training programs on science teachers' self-efficacy beliefs and to serve to clarify inconsistencies in studies of teachers' self-efficacy beliefs.
Based on these points, answers to the following questions were sought within the scope of the study:

1. What is the overall effect level of the professional development programs on teachers' science teaching self-efficacy beliefs?

2. Does the general effect of the professional development programs on teachers' science teaching self-efficacy beliefs differ significantly according to the study moderators (publication type, branch, publication year, sample size and application time)?

Method

Research Design

In this study, a meta-analysis method from quantitative research synthesis methods was used. Meta-analysis is a statistical method used for combining experimental results obtained from individual studies, synthesizing and interpreting them in the form of effect size (Card, 2012; Wolf, 1986).

Data Collection and Selection Criteria

In this meta-analysis study investigating the effect of professional development programs on science teaching self-efficacy beliefs, "Eric, Science Direct, Web of Science, Taylor & Francis Online, Scopus, JSTOR, Ulakbim, Google Academic, Proquest and YÖK National Thesis Center" databases were used in the literature review to reach the individuals' studies. In conducting the literature review, "professional development and science teaching self-efficacy beliefs", "professional development and self-efficacy", "scale/tool for science teaching self-efficacy beliefs", "self-efficacy beliefs for professional development and science teaching", "self-efficacy beliefs for professional development and science teaching", "professional development and self-efficacy", "scale/tool for science teaching self-efficacy beliefs", and "self-efficacy (beliefs) (related to/for/of) science teaching" were used as keywords.. As a result of the literature review and examinations, 1072 studies were reached. Among these studies, 14 studies (N=969) met the inclusion criteria and contained appropriate data.

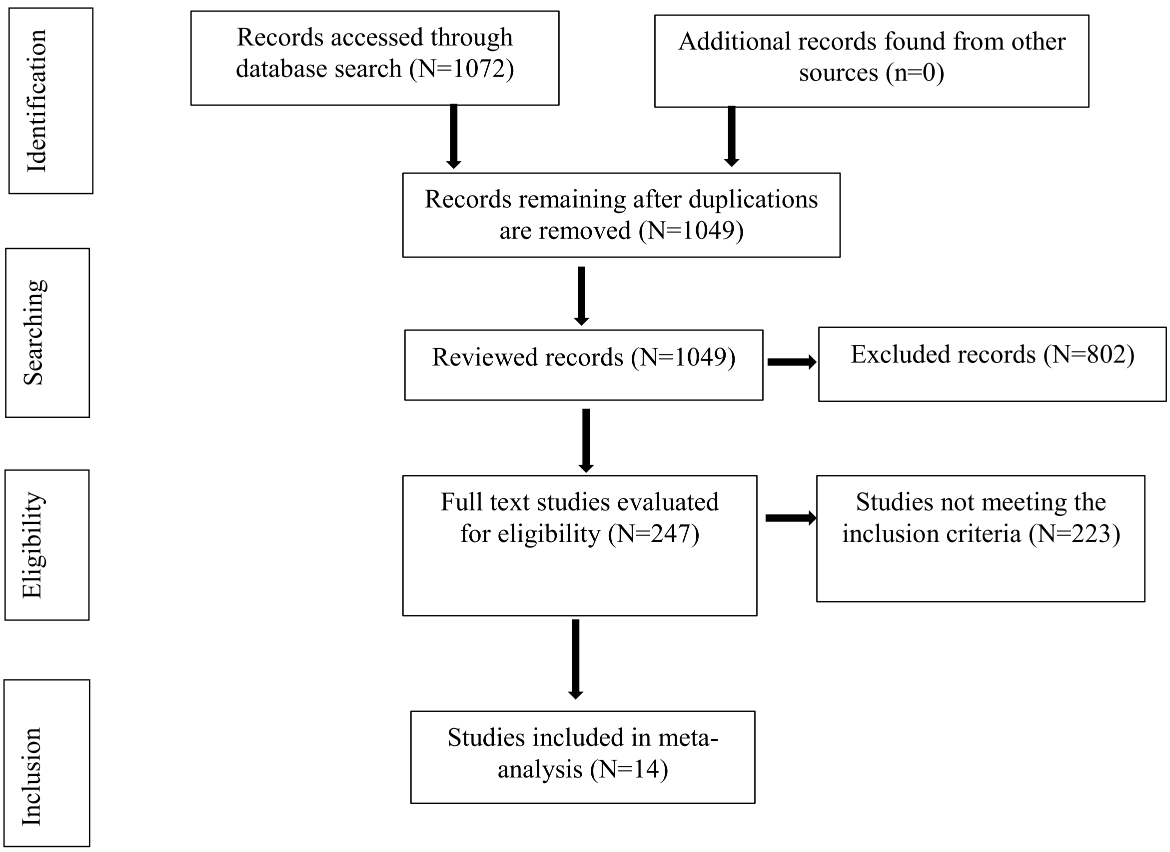


Figure 1. Flow Chart for Selection of Studies

While selecting the studies to be included in the meta-analysis, the following criteria were taken into consideration:

1) Having been published between 1990-2019 for studies in the meta-analysis. Self-efficacy concept defined by Bandura (1977) has been defined specifically for many areas. When this concept is applied to science teaching,

it is called “science teaching self-efficacy belief”. Science teaching self-efficacy belief was introduced into the literature by Enoch and Riggs (1990). For this reason, 1990 was chosen as the starting year.
2) Being published or unpublished master’s theses, doctoral dissertations, articles in electronic academic journals, and papers presented in congresses and symposiums.
3) Being conducted with teachers for studies in the meta-analysis.
4) Investigating the effect of professional development programs (conferences, workshop, in-service programs) on teachers’ science teaching self-efficacy beliefs by the studies included in the meta-analysis and using *“Science Teaching Self-Efficacy Beliefs Scale”* to collect data in these studies
5) Being studies with pretest and posttest single-group gives arithmetic mean, sample size, standard deviation, t or p values to calculate the effect size of the studies.
6) Being published in Turkish or English for studies to be included in the analysis.

The exclusion criteria of this meta-analysis study were determined as studies not examining the effect of the professional development programs on teachers’ science teaching self-efficacy beliefs, qualitative studies, and descriptive survey studies.

Data Coding

One of the critical steps for combining or comparing the results of the study to be included in the meta-analysis study is to encode the data. The first task for this purpose is to develop a coding form to classify studies meeting the inclusion criteria (Lipsey & Wilson, 2001).

In this study, a coding form consisting of three parts as “study identity”, “study content” and “data in the study” was developed. The first part of the form contains the title of the study, the name of the study, the name of the author or authors, the year of publication, the country, the type of publication, the publication status and the study pattern to record the studies that have emerged from the literature review. The second part includes information such as sample group, branch, the scale used, sample size, and application time. The third part consists of the sections for recording numerical data such as arithmetic mean, sample size, standard deviation, t value, and p-value in the individual studies.

The detailed explanation of the literature review in meta-analysis studies, recording of the studies reached as a result of the review using coding form, and obtaining similar results by other researchers using the same steps affect the reliability (Card, 2012). With the inclusion and exclusion criteria determined in this study, literature review was performed and the reliability was tried to be ensured by recording the studies reached as a result of the review into the coding form. It is also recommended to ensure the reliability of the coding form in meta-analysis studies (Card, 2012; Petitti, 2000). Agreement rate and Cohen’s Kappa statistics are widely used while performing reliability analysis between the coders in synthesizing the studies. In this study, the agreement between the coders was calculated as 93.3%. It is stated that when the variables are categorical, the agreement rate can be affected by the chance factor and a rate higher than expected can be obtained (Hartmann, 1977).
For this reason, using Cohen’s Kappa statistics providing more reliable results against the chance factor is recommended (Card, 2012). The kappa reliability value between coders was calculated as 0.90. This value indicates that there is a “very good level of agreement” between the coders according to the interpretation classification proposed by Landis & Koch (1977).

Data Analysis and Interpretation

Effect sizes constitute the basis of meta-analysis. Effect size indicates the sensitivity of an experimental procedure and the size of the experimental effect (Thalheimer & Cook, 2002). Combining the effect sizes obtained from the individual studies to be included in meta-analysis studies is done using statistical models. Two models including “fixed effects model” and “random effects model” are preferred in the literature. The fixed effects model assumes that there is only one effect size in all studies, and the sample variations' deviations in the effect sizes are caused by the sample variations (Card, 2012). The random-effects model does not include the assumption indicating that there is only one average effect size in the studies included in the analysis.

Conversely, effect sizes are thought to vary across studies and this variation is due to central tendency and study variance (Card, 2012). Researchers should decide which of these models to use prior to analysis (Borenstein, Hedges, Higgins & Rothstein, 2010). All analyses were conducted using the random effects model, taking into account that factors such as the individual studies included, conducted with teachers from different industries, in different countries, and with different sample sizes, may lead to variations in effect sizes.

While interpreting the effect sizes obtained as a result of the analysis performed using statistical models in the meta-analysis studies, some classifications are used to interpret the obtained results' level. There is more than one classification in the literature. One of the most used of these is the classification of Cohen et al. (2007). According to Cohen et al., effect size classification is as follows (2007):

- 0 ≤ Effect size value ≤ 0.20: Effect level is Poor,
- 0.21 ≤ Effect size value ≤ 0.50: Effect level is modest,
- 0.51 ≤ Effect size value ≤ 1.00: Effect level is moderate,
- 1.01 ≤ Effect size value: Effect level is strong.

One of the issues to consider in meta-analyses is publication bias. Publication bias is a condition that occurs because studies with statistically significant and positive results tend not to be published compared to studies with negative and statistically insignificant results. Due to this situation, it is very likely that the average effect size value is high (Borenstein, Cooper, Hedges & Valentine, 2009). In this study, the funnel plot, Orwin's fail-safe N, and Duval and Tweedie's trim and fill methods were used to assess publication bias.
In the study, Q (Cochran's Q) and I^2 statistics were benefited to evaluate heterogeneity. Cochran's Q can be used to measure heterogeneity and is calculated as the sum of the differences of the weighted squares between the effects of combined studies and the individual study with the weights used in the combining method (Borenstein et al., 2009). I^2 statistics, on the other hand, is a statistic that includes heterogeneity against chance factor and shows the percentage of variance in studies included in the analysis (Higgins & Thompson 2002; Higgins, Thompson, Deeks & Altman, 2002).

In the present study, analogue ANOVA similarity analysis was performed to analyze the categorical moderators. Analog ANOVA is a technique that shows whether effect sizes differ significantly between subgroups of categorical variables (Lipsey and Wilson, 2001). This analysis examines whether the difference between subgroups [between group (Qb)] is statistically significant or not. Meta-regression analysis was conducted to analyze the continuous moderators.

In this meta-analysis study, the Comprehensive Meta Analysis Version 2 (CMA Ver. 2.0) statistical package (Borenstein, Hedges, Higgins, & Rothstein, 2005) was used for effect sizes, heterogeneity tests, moderator, meta-regression, and publication bias analyzes. The SPSS 22.0 package program was used to calculate the inter-coder agreement rate and Cohen's kappa statistic. The value of 0.05 was accepted as a reference for statistical significance.

Results

Findings of Descriptive Analysis

Table 1 shows the distribution of the studies, included in the meta-analysis, in terms of the variables of their publication year, publication type, country, branch and the scale type used.

Table 1. Descriptive results of studies included in the meta-analysis

Variables	Frequency (f)	Percentage (%)
Publication Year		
1998	1	7.14
2000	1	7.14
2002	1	7.14
2003	1	7.14
2010	1	7.14
2012	3	21.42
2013	1	7.14
2014	2	14.28
2015	2	14.28
2016	1	7.14
Publication Type		
Article	10	71.42

Master's Thesis	1	7.14
Doctoral Dissertation	3	21.44
Country		
USA	13	92.86
Israel	1	7.14
Branch		
Science Fields (Science, Biology, Chemistry)	6	42.86
Other Fields (Primary + Preschool)	8	57.14
Scale used		
Ready Scale	13	92.86
Adapted Scale	1	7.14

When Table 1 was examined, 2012 in terms of year of publication, article in terms of type of publication, ABD in terms of countries, other areas (primary + preschool) in terms of branches, and finished scale in terms of scale used were the variables according to which most studies were conducted.

Publication Bias Results

"Funnel Plot, Orwin's fail-safe N, Duval and Tweedie's trim and fill" methods were used to evaluate the publication bias in this study.

Figure 2 shows the funnel plot results of the studies included in the meta-analysis.

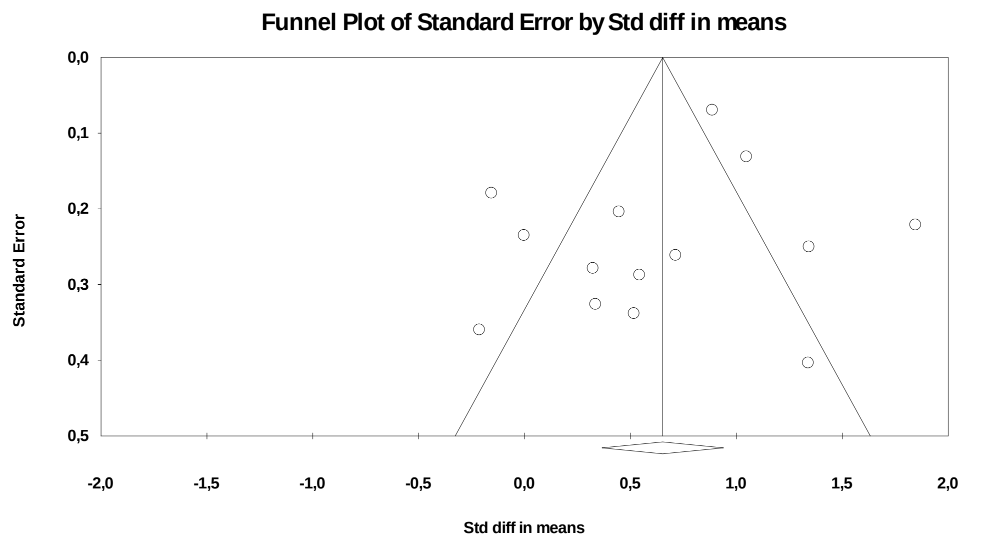


Figure 2. Funnel plot of the studies included in the meta-analysis

When Figure 2 was examined, it was seen that eight studies were situated in the left of the average effect size and 6 studies were situated in the right of the average effect size. The studies are expected to be distributed symmetrically on both sides of the average effect size in the funnel plot. However, it cannot be asserted to be fully symmetrical when the current shape was examined. On the other hand, the publication bias results of the studies included in the meta-analysis for the "Orwin's Fail-Safe N and Duval and Tweedie's trim and fill" methods were also given (Table 2).

Table 2. Testing results of publication bias

Number of Studies Included	Number of studies required for Orwin's Fail-Safe N "Poor" SMD	Duval and Tweedie's trim and fill method	
		Number of trimmed studies	Observed (Filled) for SMD
14	SOF 1070 for 0.01	1	0.652 (0.706)

SMD: Standardized Mean Difference

When Table 2 was examined, the number of studies that may reduce the effect size to a Poor level according to the Orwin's Fail-Safe N was seen to be 1070. This number is approximately 77 times of the number of studies included in the present study. 14 studies used in the study were all of studies conducted in Turkey and abroad for the research question, and reaching further 1070 studies is out of possibility. In literature, when the Orwin's Fail-Safe N is more than 5-10 times the number of studies included, it is interpreted that there is no publication bias problem for meta-analysis (Borenstein et al., 2009).

Duval and Tweedie's trim and fill method is another test used in publication bias. In this test, points causing the deterioration of the symmetry in the Funnel Plot are determined, and these points are temporarily filled in the second stage, and the general effect size is calculated again. The increase in the difference between the two general effects is interpreted as the likelihood of publication bias (Card, 2012). According to Table 2, there was no difference between the effect size value (0.652) observed and the virtual effect size (0.706) formed to correct the effect caused by publication bias.

Findings of General Effect Size

Fourteen studies investigating the effect of professional development programs on teachers' science teaching self-efficacy beliefs were combined under the random-effects model, and Table 3 shows their results.

Table 3. Combined results of teachers' science teaching self-efficacy beliefs

Model	k	ES	SE	Confidence interval of 95%		Z	P
				Lower Limit	Upper Limit		
Random Effects Model	14	0.652	0.146	0.365	0.939	4.455	0.000

k: Number of studies ES: Effect size SE: Standard Error

When Table 3 was examined, it was observed that the average effect size value of 14 studies included in the meta-analysis was calculated as 0.652 (confidence interval of 95%, lower limit of 0.365 and upper limit of 0.939) according to random-effects model. This average effect size value is a moderate effect size range according to Cohen et al. (2007). In other words, the effect of the professional development programs on teachers' science teaching self-efficacy beliefs was moderate in favor of the posttest.

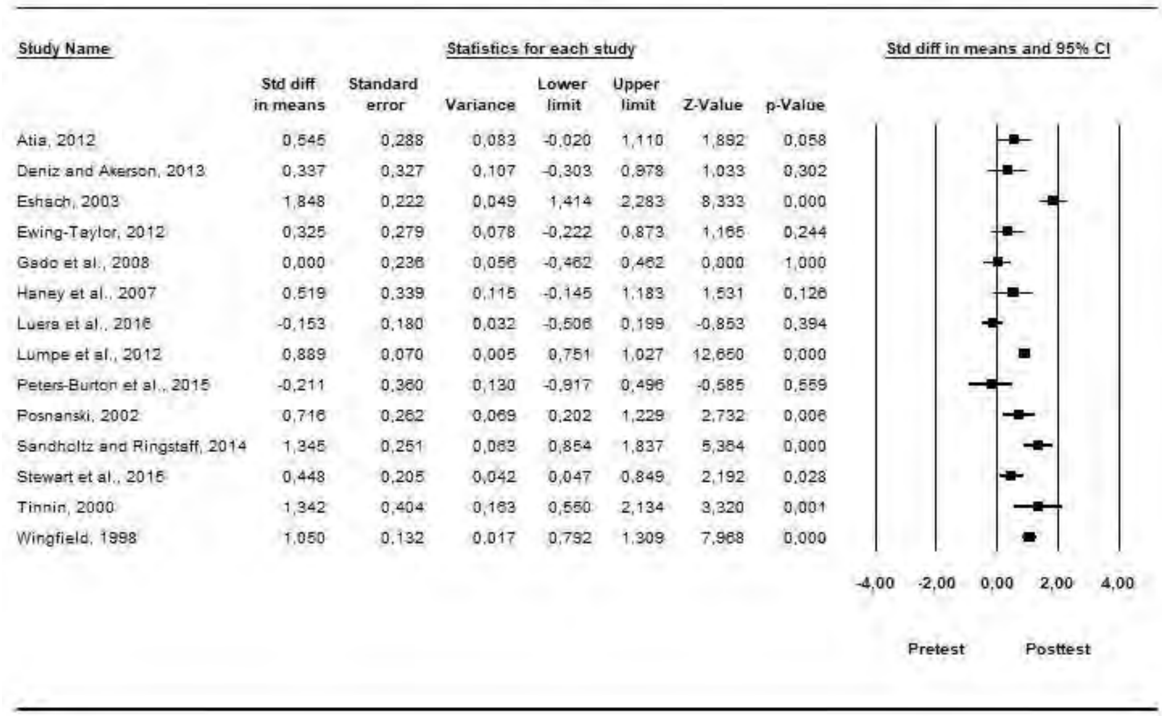
Table 4 shows the heterogeneity test results of 14 studies included in the study.

Table 4. Heterogeneity test results

Q value	df	Chi-square (χ^2)	P value	I^2
90.679	13	22.362	0.000	85.664

When Table 4 was examined, it was determined that Q value (90.679) was higher than the 13 degrees of freedom chi-square table value (22.362). This result means that distribution of the effect sizes was heterogeneous. When I^2 value showing the heterogeneity value was examined, it was seen to be 85.664. This value showed that there was a high level of heterogeneity among the studies included in the meta-analysis.

Figure 3 shows the forest plot showing the statistics of each study included in the meta-analysis. When Figure 3 was examined, it was seen that the highest effect size value was 1.848 (Eshach, 2003) and the lowest effect size value was 0.000 (Gado, Verma & Simonis, 2008). It was calculated that the effect size value was in favor of the pretest in 2 studies (Luera & Murray, 2016; Peters-Burton et al., 2015) and in favor of the posttest in 12 studies. It was determined that p-value was statistically significant in 7 studies ($p<0.05$) and statistically insignificant in 7 studies ($p>0.05$).



Study Name	Std diff in means	Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Atia, 2012	0,545	0,288	0,083	-0,020	1,110	1,892	0,058
Deniz and Akerson, 2013	0,337	0,327	0,107	-0,303	0,978	1,033	0,302
Eshach, 2003	1,848	0,222	0,049	1,414	2,283	8,333	0,000
Ewing-Taylor, 2012	0,325	0,279	0,078	-0,222	0,873	1,166	0,244
Gado et al., 2008	0,000	0,236	0,056	-0,462	0,462	0,000	1,000
Haney et al., 2007	0,519	0,339	0,115	-0,145	1,183	1,531	0,126
Luera et al., 2016	-0,153	0,180	0,032	-0,506	0,199	-0,853	0,394
Lumpe et al., 2012	0,889	0,070	0,005	0,751	1,027	12,660	0,000
Peters-Burton et al., 2015	-0,211	0,360	0,130	-0,917	0,496	-0,585	0,559
Posnanski, 2002	0,716	0,262	0,069	0,202	1,229	2,732	0,006
Sandholtz and Ringstaff, 2014	1,345	0,251	0,063	0,854	1,837	5,364	0,000
Stewart et al., 2015	0,448	0,205	0,042	0,047	0,849	2,192	0,028
Tinnin, 2000	1,342	0,404	0,163	0,550	2,134	3,320	0,001
Wingfield, 1998	1,050	0,132	0,017	0,792	1,309	7,968	0,000

Figure 3. Forest plot of the studies included in the meta-analysis

Findings of Moderator Analysis

As a result of the heterogeneity test, the study was found to have a high degree of heterogeneity. In such cases, the possible reasons for the heterogeneity of effect sizes should be investigated. Each characteristic of each study included in the current study was recorded and classified on the coding form. As a result of this classification, it is hypothesized that publication year, publication type, sample size, industry, scale used, and country variables may be the moderators. However, due to the insufficient number of subsets of the scale and country variables used, they could not be included in the moderator analysis. Moreover, when reviewing the literature, it was found that many studies included similar variables in the moderator analysis (Chesnut and Burley, 2015; Huang, 2016; Turhan, 2020; Yıldırım et al., 2019).

To explain the high level of heterogeneity between the studies in this research, the variables of publication year, publication type, sample size, branch and application time were determined as potential moderator variables. Analog ANOVA similarity analysis was conducted by considering the publication type and branch variables as categorical variables. The meta-regression analysis was performed by considering the variables of publication year, sample size and application time as continuous variables.

Table 5 shows analog ANOVA similarity results for the publication type and branch moderators.

Table 5. Categorical Moderator Analysis Results

				95%	Confidence		
Moderators	**K**	**ES**	**SE**	**Lower**	**Upper**	**Qb**	**p**
Publication type						0.546	0.460
Article	10	0.591	0.195	0.209	0.972		
Thesis	4	0.804	0.214	0.385	1.224		
Branch						21.281	0.000*
Other fields	8	1.022	0.140	0.747	1.296		
Science Fields	6	0.143	0.129	-0.111	0.396		

*p<0.05

When the moderators were examined by type of publication in Table 5, it was found that the effect size of articles was 0.591 (95% confidence interval, lower bound of 0.209, upper bound of 0.972), the value of effect size for dissertation studies was 0.804 (95% confidence interval, lower bound of 0.385, upper bound of 0.1.224) and these values of effect size do not represent a significant difference by type of publication (Qb=0.546, p=0.460). When examined by branch, the average effect size value for the studies from the other fields was 1.022 (95% confidence interval, lower bound 0.747, upper bound 1.296), the effect size value for the studies from the field of science was 0.143 (95% confidence interval, lower bound -0.111, upper bound 0.396) and these effect size values represent a significant difference in terms of branch (Qb=21.281, p=0.000).

Figure 4 shows meta-regression results for the publication year moderator.

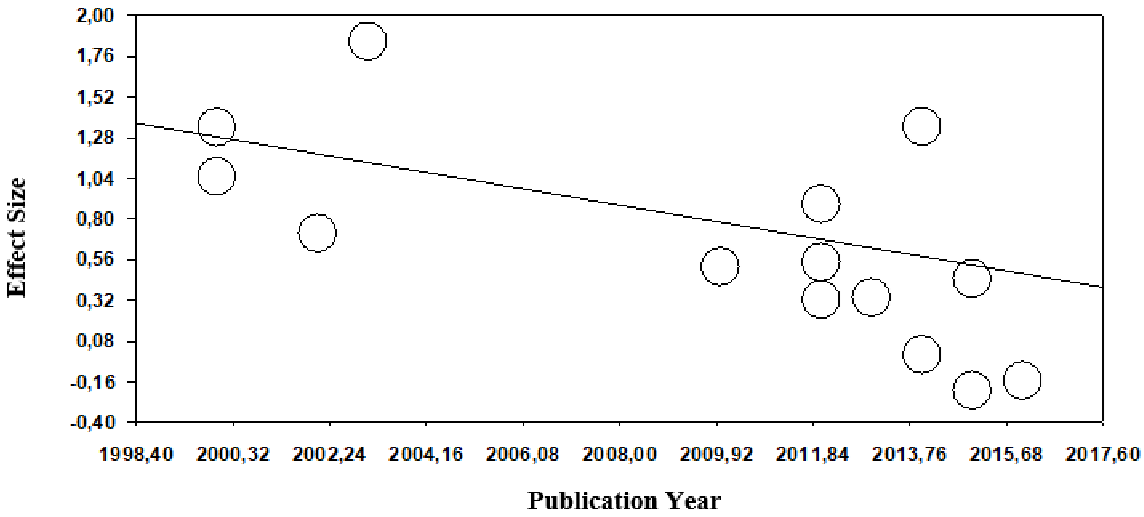


Figure 4. Correlation between publication years and effect sizes

Examining Figure 4, it was found that the slope of the line decreases as the year of publication progresses from the past to the present. Table 6 shows the results of the statistical significance of this decrease.

Table 6. Statistical results for the publication year and the effect sizes

	Point Estimation	Standard error	Lower Limit	Upper Limit	Z value	p-value
Slope	-0.05039	0.00923	-0.06849	-0.03229	-5.45604	0.00000*
Intercept	102.05758	18.56367	65.67345	138.44171		

*p<0.05

When Table 6 was examined, the publication year progressing from past to the present caused a decrease of 0.050 in the effect size. This decrease was seen to be statistically significant (p<0.05).

Figure 5 shows meta- regression results for the sample size moderator.

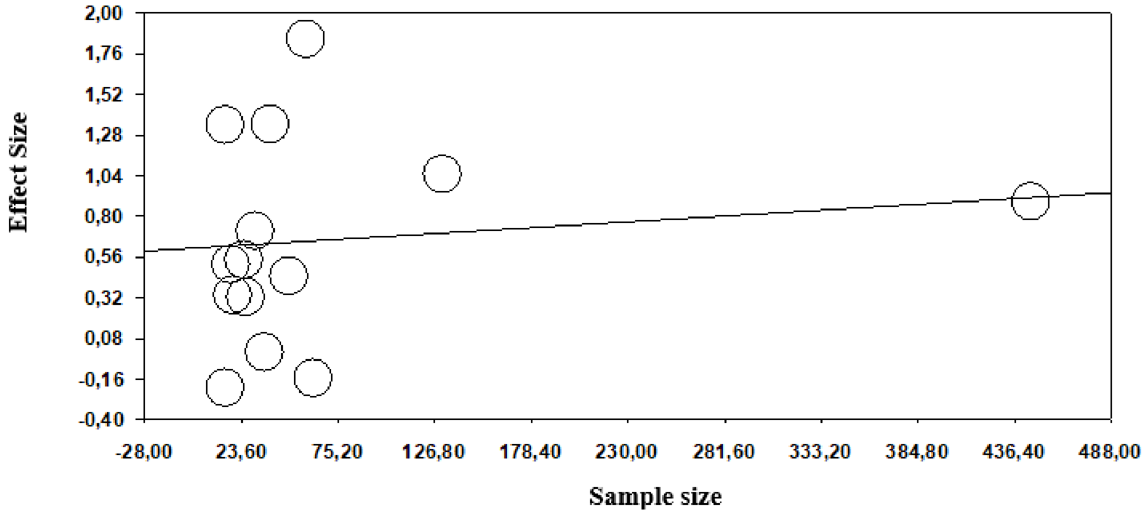


Figure 5. Correlation between the sample size and the effect sizes

When Figure 5 was examined, an increase was observed in the line slope with the increasing sample size. Table 7 shows statistical significance results regarding this increase.

Table 7. Statistical results regarding the sample size and effect sizes

	Point Estimation	Standard error	Lower Limit	Upper Limit	Z value	P value
Slope	0.00067	0.00025	0.00018	0.00115	2.70837	0.00676*
Intercept	0.61506	0.07555	0.46698	0.76314		

*p<0.05

When Table 7 was examined, it was seen that there was an increase of 0.00067 in the effect size with the increasing sample size and this increase was statistically significant ($p<0.05$).

It was determined that the application time in the studies included in the meta-analysis was given in hours. However, the application time was not specified in 3 studies (Peters-Burton et al., 2015; Tinnin, 2000; Wingfield, 1998). Therefore, 11 studies were included in the analysis. Figure 6 shows meta-regression results for the application time (hour) moderator.

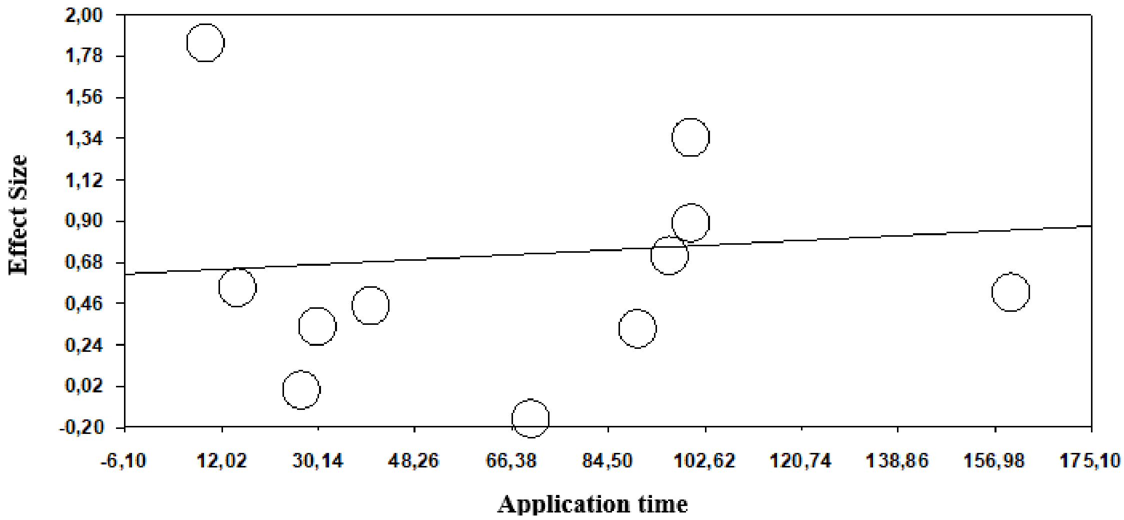


Figure 6. Correlation between the application time and the effect sizes

When Figure 6 was examined, it was observed that there was an increase in the line slope with the increasing application time. Table 8 shows statistical significance results for this increase.

Table 8. Statistical results of the application time and effect sizes.

	Point Estimation	Standard error	Lower Limit	Upper Limit	Z value	P value
Slope	0.00141	0.00155	-0.00163	0.00444	0.90679	0.36452
Intercept	0.62744	0.13649	0.35992	0.89497		

When Table 8 was examined, it was found that increasing the application time resulted in an increase in effect size of 0.00141. However, this increase did not result in a statistically significant difference in effect size ($p > 0.05$).

Conclusion and Discussion

The study included 14 studies that examined teachers' self-efficacy beliefs in science classrooms in the meta-analysis. It was found that the majority of these studies were conducted in 2012 according to the categorical descriptive characteristics. "Articles" according to the type of publication, "USA" according to the countries, "other fields" according to the industries, "finished scale" according to the scale used were the most studied categories. Another finding was that there was no study in Turkey that met the criteria of the present study.
It was also concluded that the publication bias was low in 14 studies determined for the meta-analysis and it was not enough to change the effect size classification. As a result of the combination made under the random-effects model, the average effect size value was determined as 0.652. This effect size value found under the random-effects model was in a moderate level according to the classification of Cohen et al. (2007). Deehan (2017) suggested that professional development practices such as the nature of science teaching can positively improve teachers' science teaching efficacy beliefs.

When the effect sizes of the individual studies included in the meta-analysis were examined, the result indicated that teachers' self-efficacy beliefs were moderate in the studies with moderate effect sizes (Atia, 2012; Haney, Jing, Keil, & Zoffel, 2007; Lumpe, Czerniak, Haney, & Beltyukova, 2012; Posnanski, 2002), as well as the studies that examined teachers' self-efficacy beliefs using the meta-analysis method (Chesnut and Burley, 2015; Kalkan, 2020; Shoji, Cieslak, Smoktunowiez, Rogala, Benight, & Luszczynska, 2016), supports the current study. However, among the studies included in the meta-analysis, those with a small effect size (Deniz and Akerson, 2013; Ewing-Taylor, 2012; Peters-Burton et al., 2015; Stewart et al., 2015) and the meta-analysis study that found that teachers' self-efficacy levels are low (Klassen and Tze, 2014) differ from the findings of the current study. The reason for this difference may be that the other meta-analysis studies in the literature examined teachers' self-efficacy beliefs, whereas this study examined teachers' self-efficacy beliefs in science teaching.

A moderator analysis was conducted to explain the high degree of heterogeneity between the studies. The results of the moderator analysis suggest that the moderators industry, publication year, and sample size have a statistically significant effect on the effect sizes of teachers' self-efficacy beliefs in science education. When the results for the branch moderator were examined, it was found that science teachers (science, biology, and chemistry teachers) had lower self-efficacy beliefs about teaching science than teachers of other subjects (classroom teachers and preschool teachers). There are many studies that conclude that science teachers' self-efficacy beliefs should be higher because they have a science degree (Hechter, 2008; Jarett, 1999). In these studies, the number of science classes completed and a science-focused school experience are found to have a positive impact on teachers' self-efficacy beliefs. (Cripe, 2009; Hechter, 2008; Jarrett, 1999).

When analyzing the results of the publication year moderator, it was found that the effect sizes of teachers' self-efficacy beliefs in science education decreased from 1990 to 2019 (Eshach, 2003; Luera et. al., 2016). This could be related to the inhomogeneous distribution of the studies included in the meta-analysis in terms of publication years.

When the results were evaluated for the moderator of sample size, it was found that increasing sample size had a positive effect on teachers' self-efficacy beliefs in science teaching. According to Bandura (2003), social modeling or indirect experiences are among the resources that improve self-efficacy. Individuals can learn from observing others or from the experiences they have with others. Social models are important aids to learning, especially when one's own competence is limited (Pajares, 2002). This situation suggests that high participation professional development programs enriched teachers' professional experiences and increased their self-efficacy beliefs in science teaching by providing opportunities to examine practices.

Recommendations

Recommendations for Researchers

- When the studies included in the meta-analysis were examined by publication year, it was found that teachers' self-efficacy beliefs in science teaching have decreased from the past to the present. Researchers may investigate the reasons for this.This meta-analysis study covered the studies conducted between 1990 – 2019. In a new meta-analysis study, the study can be repeated by extending this time interval.
- When the studies in the research were examined in terms of countries, it was determined that there was no study in Turkey investigating the effect of professional development programs on teachers’ science teaching self-efficacy beliefs. It can be recommended for researchers to design studies on this subject.
- In this present meta-analysis study, the effect sizes on teachers’ science teaching self-efficacy beliefs were examined. In a new meta-analysis study, the correlation between the science teaching self-efficacy beliefs and different variables such as attitude, motivation can be investigated.
- This study investigated the effect of professional development programs on teachers’ science teaching self-efficacy beliefs with meta-analysis method. A new study can be planned with other systematic synthesis methods on this subject.

Recommendations for Implementers

Increasing the sample size in the study resulted in a statistical increase in teachers' self-efficacy beliefs in science teaching. Teachers' participation in professional development programs can be encouraged.

Acknowledgement

This study was funded by Inonu University Scientific Research Projects Unit (SDK-2019-1577).

References

(* shows the studies included in the meta-analysis.)

Ada, S. and Ünal, S. (2009). *Eğitim bilimine giriş [Introduction to education science].* Ankara: Nobel Yayınları

*Atia, K. M. (2012). *Exploring the influence of a summer professional development program on teachers' self-efficacy belief outcomes and the possible effect on science CST scores* (Publication No. 923853630) (Master thesis, California State University). ProQuest Dissertations and Theses Global.

Avalos, B. (2011). Teacher professional development in teaching and teacher education over ten years. *Teaching and Teacher Education*, *27*(1), 10-20.

Avidav, A. (2000). *Professional career empowerment amongst teachers in primary school: The concept and its components* [Unpublished doctoral thesis]. Hebrew University.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191-215.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.

Bandura, A. (2002). Social cognitive theory in cultural context. *Journal of Applied Psychology, 51*(2), 269-290.

Bandura, A. (2003). On the psychosocial impact and mechanisms of spiritual modelling. *International Journal for the Psychology of Religion, 13,* 167–173.

Banilower, E. R., Boyd, S. E., Pasley, J. & Weiss, I. R. (2005). *Lessons from a decade of mathematics and science reform.* Chapel Hill, NC: Horizon, Research,

Bennett, J. M. P. (2016). *An investigation of elementary teachers' self-efficacy for teaching integrated science, technology, engineering, and mathematics (STEM) education* (Publication No. 10137835) [Doctoral dissertation, Regent University]. ProQuest Dissertations and Theses Global.

Bhattacharyya, S.,Volk, T., & Lumpe, A. (2009). The Influence of an extensive inquiry-base field experience on pre-service elementary student teachers’ science teaching beliefs. *Journal of Science Teacher Education, 20*(3), 199-218.

Blanford, P. E. (2008). *Teachers' participation in formal induction programs and their subsequent level of professional growth in wisconsin public schools* (Publication No. 3326690) [Doctoral dissertation, The University of Southern Mississippi]. ProQuest Dissertations and Theses Global.

Blandford, S. (2000). *Managing professional development in schools*. London: Routledge.

Bleicher, R. E. (2006). Nurturing confidence in preservice elementary science teachers. *Journal of Science Teacher Education, 18(*6), 841-860.

Borenstein, M., Cooper, H., Hedges, L., & Valentine, J. (2009). Effect sizes for continuous data. *The Handbook of Research synthesis and Meta-Analysis*, *2*, 221-235.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive Meta-analysis Version 2.* Englewood, N. J: Biostat. Available from: http://www.meta-analysis.com/downloads/Meta-Analysis-Manual.pdf

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed effect and random effects models for meta-analysis. *Research Synthesis Methods, 1*(2), 97-111.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3-15.

Cakiroglu, J., Capa-Aydin, Y., & Hoy, A. W. (2012). Science teaching efficacy beliefs. In *Second international handbook of science education* (pp. 449-461). Dordrecht: Springer.

Card, N. A. (2012). *Applied meta-analysis for social science research*. New York: The Guilford Press.

Chesnut, S. R., & Burley, H. (2015). Self-efficacy as a predictor of commitment to the teaching profession: A meta-analysis. *Educational Research Review, 15*, 1-16. doi:10.1016/j.edurev.2015.02.001.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th Edition). New York: Routledge

Czerniak, C. M., & Schriver, M. (1994). An examination of preservice science teachers' beliefs and behaviors as related to self-efficacy. *Journal of Science Teacher Education, 5*(3),77-86.

Deehan, J. (2017). *The Science Teaching Efficacy Belief Instruments (STEBI A and B): A comprehensive review of methods and findings from 25 years of science education research.* Cham, Switzerland: Springer International Publishing.

Dembo, M. H., & Gibson, S. (1985). Teachers' sense of efficacy: An important factor in school improvement. *The Elementary School Journal, 86*, 173-184.

*Deniz, H., & Akerson, V. (2013). Examining the impact of a professional development program on elementary teachers' views of nature of science and nature of scientific inquiry, and science teaching efficacy beliefs. *Electronic Journal of Science Education*, *17*(3), 1-19.

Desimone, L.M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher, 38*(3), 181-199. doi: 10.3102/0013189X08331140.

*Eshach, H. (2003). Inquiry-events as a tool for changing science teaching efficacy belief of kindergarten and elementary school teachers. *Journal of Science Education and Technology, 12*(4), 495-501.

*Ewing-Taylor, J. M. (2012). *A longitudinal evaluation study of a science Professional development program for K-12 teachers*: *NERDS* (Publication No.3539187) [Doctoral dissertation, University of Nevada, Reno]. ProQuest Dissertations and Theses Global.

Fishman, B. J., Marx, R. W., Best, S., & Tal, R. T. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and Teacher Education*, *19*(6), 643-658.

*Gado, I.,Verma, G., & Simonis, D. (2008). Middle grade teachers' perceptions of their chemistry teaching efficacy: Findings of a one year long professional development program. *Georgia Educational Researcher, 6*(1), 46-57.

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S., (2001). What makes professional development effective: Results from a national sample of teachers. *American Educational Research Journal, 38*(4), 915-945.

Guskey, T. R. (2003). What makes professional development effective?. *Phi Delta Kappan*, *84*(10), 748-750.

Haney, J., Lumpe, A., & Czerniak, C. (2002). From beliefs to actions: The beliefs and actions of teachers implementing change. *Journal of Science Teacher Education*, *13*, 171-187.

*Haney, J. J., Jing, W., Keil, C., & Zoffel, J. (2007). Enhancing teachers' beliefs and practices through problem-based learning focused on pertinent issues of environmental health science. *Journal of Environmental Education*, *38*(4), 25-33.

Hartmann, D. P. (1977). Considerations in the choice of interobserver reliability estimates. *Journal of Applied Behavior Analysis, 10*(1), 103-116.

Hechter, R. P. (2008). *Changes in preservice elementary teachers' personal science teaching efficacy and science teaching outcome expectancies: The influence of context*, (Publication No. 3324268) [Doctoral dissertation, University of North Dakota]. ProQuest Dissertations and Theses Global.

Hiebert, J., Gallimore, R., & Stigler, J., (2002). A knowledge base for the teaching profession: What would it look like and how can we get one? *Educational Researcher, 31*(5), 3-15.

Higgins, J., & Parkons, R., (2009). A successful professional development model in mathematics: A system-wide new zealand case. *Journal of Teacher Education, 60*(3), 231-24.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539-1558.

Higgins, J.,Thompson, S., Deeks, J., & Altman, D. (2002). Statistical heterogeneity in systematic reviews of clinical trials: A critical appraisal of guideline sandpractice. *Journal of Health Services Research & Policy*, *7*(1), 51-61.

Huang, C. (2016). Achievement goals and self-efficacy: A meta-analysis. *Educational Research Review, 19,* 119-137. http://dx.doi.org/10.1016/j.edurev.2016.07.002

Jarrett, O. S. (1999). Science interest and confidence among preservice elementary teachers. *Journal of Elementary Science Education, 11*, 49-59.

Kalkan, F. (2020). The relationship between teachers' self-efficacy beliefs and job satisfaction levels: A meta-analysis study. *Education and Science, 45*(204), 317-343.

Klassen, R. M.,& Tze, V. M. (2014). Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis. *Educational Research Review, 12*, 59-76.

Landis, J. R. & Koch, G. G. (1977) The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174.

Lipsey, M. W., & Wilson, D. B. (2001). *Applied social research methods series; Practical Meta-analysis.* Thousand Oaks, CA, US: Sage Publications, Inc.

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2009). *Designing professional development for teachers of science and mathematics*. Corwin press.

*Luera, G., & Murray, K. (2016). A mixed methods approach to determining the impact of a geoscience field research program upon science teachers' knowledge, beliefs, and instructional practices. *Journal of Geoscience Education*, *64*(4), 303-313. https://doi.org/10.5408/14-064.1

Luft, J. A., & Hewson, P. W. (2014). Research on teacher professional development programs in science. *Handbook of research on science education*, *2*, 889-909.

*Lumpe, A., Czerniak, C., Haney, J., & Beltyukova, S. (2012). Beliefs about teaching science: The relationship between elementary teachers' participation in professional development and student achievement. *International Journal of Science Education*, *34*(2), 153-166.

Madden, M. R. (2016). *Systemic changes occurring in elementary schools that pursue a STEM focus,* (Publication No. 10240305) [Doctoral dissertation, University of West Georgia]. ProQuest Dissertations and Theses Global.

Martin, B. J. (2016). *A case study investigation of practices and beliefs of teachers at a STEM-focused elementary school* (Publication No. 10183244) [Doctoral dissertation, San Francisco State University]. ProQuest Dissertations and Theses Global.

Menezes, L. (2011). Collaborative research as a strategy of professional development of teachers. In Ubuz, B. (Ed), *Proceedings of the 35th Int. Conference on the Psychology of Mathematics Education*, (Vol. 3, pp. 225-232). Ankara, Turkey: PME

Mintzes, J. J., Marcum, B., Messerschmidt-Yates, C., & Mark, A. (2013). Enhancing self efficacy in elementary science teaching with professional learning communities. *Journal of Science Teacher Education*, *24*(7), 1201-1218.

Özer, B. (2005). Ortaöğretim öğretmenlerinin mesleki gelişime ilgisi [Secondary school teacher's interest in professional development] . *Eğitim Bilimleri ve Uygulama*, *4*(8), 209-219.

Pajares, F. (2002). *Overview of social cognitive theory and of self-efficacy.* Retrieved from http://people.wku.edu/richard. miller/banduratheory.pdf

*Peters-Burton, E. E., Merz, S. A., Ramirez, E. M., & Saroughi, M. (2015). The effect of cognitive apprenticeship-based professional development on teacher self-efficacy of science teaching, motivation, knowledge calibration, and perceptions of inquiry-based teaching. *Journal of Science Teacher Education*, *26*(6), 525-548. https://doi.org/10.1007/ s10972-015-9436-1

Petitti D. B. (2000). *Meta-analysis, Decision Analysis and Cost-effectiveness Analysis: Methods for Quantitative Synthesis in Medicine* (2nd ed). New York: Oxford University Press.

*Posnanski, T. J. (2002). Professional development programs for elementary science teachers: An analysis of teacher self-efficacy beliefs and a professional development model. *Journal of Science Teacher Education*, *13*(3), 189-220.

Ramakrishnan, R., & Salleh, N. M. (2018). Teacher's self-efficacy: A systematic review. *International Journal of Academic Research in Business and Social Sciences*, *8*(12), 2379-2402.

Sağlam, M., ve Yüksel, İ. (2007). Program değerlendirmede meta-analiz ve meta değerlendirme yöntemleri [Meta-analysis and meta-evaluation in program evaluation]. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, 18*, 175-188.

*Sandholtz, J. H., & Ringstaff, C. (2014). Inspiring instructional change in elementary school science: The relationship between enhanced self-efficacy and teacher practices. *Journal of Science Teacher Education*, *25*(6), 729-751. doi:10.1007/s10972-014-9393-0

Schriver, M., & Czerniak, C. M. (1999). A comparison of middle and junior high science teachers' levels of efficacy and knowledge of developmentally appropriate curriculum and instruction. *Journal of Science Teacher Education, 10*(1), 21-42.

Shoji, K., Cieslak, R., Smoktunowicz, E., Rogala, A., Benight, C. C., & Luszczynska, A. (2016). Associations between job burnout and self-efficacy: A meta-analysis. *Anxiety, Stress, and Coping*, *29*(4), 367-386.

*Stewart, A. E., Knox, J. A., & Schneider, P. (2015). Piloting and evaluating a workshop to teach Georgia teachers about weather science and safety. *Journal of Geoscience Education*, *63*(4), 271-284.

Thalheimer, W., & Cook, S. (2002). How to calculate effect sizes from published research: A simplified methodology. *Work-Learning Research, 1*, 1-9.

*Tinnin, R. K. (2000). *The effectiveness of a long-term professional development program on teachers' selfefficacy, attitudes, skills, and knowledge using a thematic learning approach* (Publication No. 9992926) [Doctoral dissertation, University of Texas at Austin]. ProQuest Dissertations and Theses.

Trimmell, M. D. (2015). *The effects of stem-rich clinical professional development on elementary teachers' sense of self-efficacy in teaching science* (Publication No. 3704976) [Doctoral dissertation, California State Univetsity]. ProQuest Dissertations and Theses.

Turhan, N. S. (2020). Gender Differences in Academic Motivation: A Meta-Analysis. *International Journal of Psychology and Educational Studies, 7(2),* 211-224. http://dx.doi.org/10.17220/ijpes.2020.02.019

Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and Teacher Education*, *24*(1), 80-91.

Villegas-Reimers, E. (2003). *Teacher professional development: An international review of the literature*. Paris: International Institute for Educational Planning, UNESCO.

Yıldırım, İ., Çırak-Kurt, S., & Şen, S. (2019). The effect of teaching "learning strategies" on academic achievement: A meta-analysis study. *Eurasian Journal of Educational Research, 79,* 87-114. DOI: 10.14689/ejer.2019.79.5

*Wingfield, M. E. (1998). *The effect of site based preservice experiences on elementary science teaching self-efficacy beliefs*. (Publication No. 9828336) [Doctoral dissertation, University of Houston]. ProQuest Dissertations and Theses.

Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis*. Sage, Newbury Park, CA.