ANKARA ÜNİVERSİTESİ SOSYAL BİLİMLER DERGİSİ
ANKARA UNIVERSITY JOURNAL OF SOCIAL SCIENCES

Ankara Üniversitesi Sosyal Bilimler Dergisi | 13(1) | 2022 | 15 - 27

10.33537/sobild.2022.13.1.2

THE EFFECT OF COMPLEMENTARY MEASUREMENT AND EVALUATION TECHNIQUES ON ACADEMIC ACHIEVEMENT: A META-ANALYSIS STUDY

TAMAMLAYICI ÖLÇME VE DEĞERLENDİRME TEKNİKLERİNİN AKADEMİK BAŞARIYA ETKİSİ : BİR META-ANALİZ ÇALIŞMASI

Ceren MUTLUER
BAİBÜ, EFi Eğitim Bilimleri Bölümü,
cmutluer@yandex.com

Makale Bilgisi

Gönderildiği tarih:	*15.10.2021*
Kabul edildiği tarih:	*01.01.2022*
Yayınlanma tarihi:	*31.01.2022*

Article Info

Date submitted:	*15.10.2021*
Date accepted:	*01.01.2022*
Date published:	*31.01.2022*

Keywords

Meta-Analysis; Complementary Measurement and Evaluation Techniques; Academic Achievement; Random-Effects Model.

Anahtar sözcükler

Meta-Analiz; Tamamlayıcı Ölçme ve Değerlendirme Teknikleri; Akademik başarı; Rastgele Etkiler Modeli.

Abstract

This research aims to investigate the effect of complementary measurement and evaluation techniques on academic achievement. This research is a meta-analysis study and consists of 40 research studies (articles, thesis, and full-text presentations). A random-effects model was chosen for the meta-analysis process. In three studies discussed in the study, complementary measurement and evaluation techniques have a negative effect on academic achievement, while 40 studies have a positive effect of complementary measurement and evaluation techniques on academic achievement. The overall effect calculated using the random-effects model was +1,080. As a result of this research, complementary measurement and evaluation techniques have a wide impact on academic achievement.

Öz

Bu araştırma, tamamlayıcı ölçme ve değerlendirme tekniklerinin akademik başarıya etkisini araştırmayı amaçlamaktadır. Bu araştırma bir meta-analiz çalışması olup 40 araştırma çalışmasından (makale, tez ve tam metin sunumlar) oluşmaktadır. Meta-analiz süreci için rastgele etkiler modeli seçilmiştir. Çalışmada ele alınan üç çalışmada tamamlayıcı ölçme ve değerlendirme tekniklerinin akademik başarı üzerinde olumsuz etkisi bulunurken, 40 çalışmada tamamlayıcı ölçme ve değerlendirme tekniklerinin akademik başarı üzerinde olumlu etkisi bulunmaktadır. Rastgele etkiler modeli kullanılarak hesaplanan genel etki +1.080'dir. Bu araştırma sonucunda tamamlayıcı ölçme ve değerlendirme teknikleri akademik başarı üzerinde geniş bir etkiye sahiptir.

1. Introduction

The change in the behavior of the individual in the education process is an indicator used to evaluate the educational outcomes. Whether the individual acquires the desired behaviors in the education process is the focal point of the and evaluation field. The measurement and evaluation process for the purpose of providing feedback plays an active role in examining the existing program inputs, processes, and outputs of the specified program.

Traditional measurement and evaluation techniques used to evaluate individual behaviors in the education process were insufficient in measuring all behavioral (cognitive, affective and psychomotor) changes. With traditional measurement and evaluation, the objectives of the curriculum focus on whether it causes a change in cognitive behavior at the end of the semester. Open-ended written exams, short-answer written exams, oral exams, multiple-choice tests, true-false type tests and matching type tests are used in the traditional measurement and evaluation process. With these tests, it is not possible to go beyond remembering the knowledge before the exam in the cognitive structure of the individuals and marking the correct answer. According to Berenson (1995), traditional measurement approaches contribute to tracking students' grades rather than learning. Unfortunately, the individual's affective behavior regarding the concept, the process of organizing knowledge in a cognitive structure, her interests, and psychomotor skills cannot be accurately measured with traditional measurement and evaluation tools.

Since the second half of the 20th Century, researchers such as Piaget, Vygotsky, Ausubel, Bruner, and Von Glasersfeld have emphasized the constructivist approach depending on their research (Açıkgöz, 2003). Von Glasersfeld (1989) explains constructivism according to the concept of knowledge and thinks that the knowledge is actively received through the senses or communication and the knowledge is structured by cognition.

With the changing learning styles and approaches, individual behaviors have been analyzed in more detail, and education programs have been updated by emphasizing the constructivist approach for the current teaching process. As an important element of the education process, the measurement and evaluation process has also been reconsidered in this process. The measurement and evaluation process, which is handled in the constructivist approach, comes to the front that it is necessary to consider what the change in the individual is and how he/she learns. In current education programs, it is impossible to carry out measurement-evaluation activities only with product-oriented traditional measurement-evaluation methods. Traditional measurement and evaluation are insufficient in terms of the fact that the cognitive behaviors of the individual are limited to the exams held at the end of the subject and the end of the term, inability to evaluate the cognitive process at the beginning and in the process and focusing on cognitive behaviors for only one concept.

In the process, it is necessary to evaluate not only cognitive learning but also other learning areas, considering individual differences. The realization of this is possible using complementary measurement and evaluation methods throughout the education process with the participation of teachers and students (MEB, 2017). Performance tasks, portfolio, e-portfolio, self-peer-group assessment, holistic and analytical rubric (graded rubric), rating scale, tally tool, word association, structured grid, diagnostic branched tree, Vee diagram supplementary measurement used to evaluate performance and assessment tools.

Complementary measurement and evaluation techniques are superior to traditional measurement and evaluation techniques in terms of the following facts:

1. being a concrete product and allowing students to measure their success in creating original ideas and their creativity,
2. making scoring by examining the process,
3. enabling the measurement of high-level cognitive skills,
4. preventing memorization for a single correct answer,
5. having the opportunity to give feedback at every stage while the student is performing the requested task or
6. answering a question, the task that the students are asked to fulfill, or the question he/she is asked to answer is related to situations he/she will encounter in real life,
7. solving real-life problems,
8. using tools such as rubrics, rubrics, etc. (Baker, O'neil and Linn, 1993; Gipps and Stobart, 2003; Oosterhof, 2003; Kutlu, Doğan and Karakaya, 2014).

On the other hand, traditional measurement and evaluation methods are more practical because the scoring of complementary measurement and evaluation techniques are more subjective and time-consuming (Nitko, 1996; Goodrich, 2000).

Many studies have been conducted in our country focusing on complementary measurement and evaluation techniques (Goodrich, 1996; Corcoran, Dershimer and Tichenor, 2004; Doğan, 2005; Karakaya, 2007; McMillan, 2007; Kutlu, Karakaya and Doğan, 2008; Yurdabakan, 2011; Karalok, 2014; Mutluer and Okyay, 2017;). Various studies have been carried out to determine the effect of complementary measurement and evaluation techniques on the academic achievement of students (Allen and Flippo, 2002; Şeker ve Sert, 2015; Mutluer and Nartgün, 2017; Oluk and Ekmekçi, 2017; Buldur and Doğan, 2017; Tünkler and Güven, 2019). Apart from the indicated studies, many studies in the literature have

been carried out to investigate the effects of complementary measurement and evaluation and the relationship and effect of these techniques with academic achievement. The fact that the researches are conducted for different courses for students at different education levels and the resulting data are different from each other necessitates a meta-analysis for these studies. In this way, it is thought that it will be possible to make a holistic interpretation of the effect of complementary measurement and evaluation techniques on academic achievement by examining all the studies conducted in Turkey that reveal the effects of complementary measurement and evaluation techniques on students' success.

1.1. Purpose of Research

This study aimed to integrate and interpret the findings of similar studies conducted to determine the effect of complementary measurement and evaluation techniques on students' academic achievement in Turkey by combining them with meta-analysis method. In line with this purpose, the question "What is the effect of complementary measurement and evaluation techniques on the academic achievement of students?" is posed in this study.

A meta-analysis was felt on this subject to consider the research results holistically and form a common opinion. This study is deemed to be important as it enables a synthesis by integrating the findings of similar studies conducted to determine the effect of complementary measurement and evaluation techniques on students' academic achievement. Bringing together the findings of similar studies on a subject and making a synthesis of them will provide an opportunity to make meaningful comments about the effect of complementary measurement and evaluation techniques on the academic achievement of students. Thus, it is thought that it will contribute to both the applications to be made on this subject and the research to be done in the future by providing a broader and more comprehensive view of the effect on the students' academic achievement.

2. Methodology

2.1. Research Design

In this study, a meta-analysis design was chosen to integrate the results of studies investigating the effect of complementary measurement and evaluation techniques on academic achievement and draw a general conclusion. Glass first bandied about the term meta-analysis in 1977. According to Borenstein, Hedge, Higgins and Rothstein (2009), meta-analysis studies are studies that can be more generalized as a result of integrating the results of studies with the same or related purpose and reach results that many studies have confirmed. Field (2001) describes meta-analysis as -a statistical technique by which information from independent studies is assimilated.

2.1.1. Meta-analysis steps

There are eleven sequential steps to conducting a quality meta-analysis (Borenstein, Hedge, Higginsand Rothstein, 2009; Field and Gillett, 2010; Şen and Yıldırım, 2020):

1. Choose the subject
2. Reviewing the literature
3. Decide on inclusion criteria
4. Decide on theme
5. Decide on the research questions
6. Coding
7. Calculate the effect sizes
8. Do the heterogeneity test
9. Choose the model
10. Calculate the overall effect
11. Interpret

The overall effect in the meta-analysis should be interpreted in line with the specified stages. In the study, the following classification was made regarding the effect size values calculated for individual studies and the overall effect (Thalheimer and Cook, 2002).

Table1: Classification According to Effect Sizes

Effect Size Coefficient Ranges	Interpretation of Effect Size
-0,15≤Effect Size<0,15	Insignificant
0,15≤Effect Size<0,40	Small
0,40≤Effect Size<0,75	Medium
0,75≤Effect Size<1,10	Wide
1,10≤Effect Size<1,45	Very Wide
1,45≤Effect Size	Enormous/Excellent

Under the fixed-effect model we assume that there is one true effect size that underlies all the studies in the analysis, and that all differences in observed effects are due to sampling error (Borenstein, Hedges, Higgins and Rothsteind, 2009). Therefore, the true effect size is the same in all studies.

In a random-effects meta-analysis model, the effect sizes in the studies that were performed are assumed to represent a random sample from a particular distribution of these effect sizes (DeCoster, 2004; Borenstein, Hedges, Higgins and Rothsteind, 2010).

While deciding on the model, it is necessary to answer some questions theoretically. These questions are: 'Is the overall effect one?', 'Is the overall effect one in the universe?', 'Do the results of the study vary in different subgroups?', 'Is there a single source of error? In the

answers to these questions, the model should be determined by considering whether the widespread effect is shared by the subgroups and the theoretical framework is drawn theoretically. Although the heterogeneity test in model selection provides a statistical finding, it should first be determined whether the reseach resuls vary in the subgroups with a detailed literature review. Determining the theoretical framework is a priority here. The choice of a random-effects model would be appropriate for education and social sciences.

2.1.2. Data sources and search strategies

During the research process, YÖK (Councıl Of Higher Education) National Thesis Center, Google Scholar, and Web of Science databases were used. Since the word "alternative" instead of "complementary" is a common misconception, the word term "alternative" is also handled as thesis, article, and declaration in databases. Considering this situation, among the studies conducted between 2000-2021, 'Complementary measurement and evaluation', 'alternative measurement and evaluation', studies that are based on these measurement and evaluation techniques and the concept of 'academic achievement' are examined. Alternative measurement and evaluation techniques were used in 122 theses, and complementary measurement and evaluation techniques were used in 18 theses. Apart from these mentioned concepts, 592 theses were examined one by one. A single complementary measurement and evaluation technique is used, or more than two complementary measurement and evaluation techniques are used, and the academic success in the course can be examined, one by one. Of these theses, 32 were found in which the effects of the specified complementary measurement and evaluation techniques on academic achievement were examined. As a result of a detailed examination, 24 theses were selected by the research criteria. Apart from these, it was observed that complementary measurement and evaluation technique was applied in 345 articles, and alternative measurement and evaluation technique was applied in 1510 articles. There were 43 articles on the effect of complementary measurement and evaluation on academic achievement. After detailed analysis, it was found that 16 articles were suitable to be examined in the meta-analysis process.

The following criteria were used to select the studies included in the meta-analysis study:

- It is research that examines the effect of complementary measurement and evaluation techniques on students' academic achievement,
- Master's, doctoral thesis, article, or statement published in Turkey,
- Involvement of control and experimental groups to examine the concept of effect in the research,
- Include sufficient information (sample size, mean, standard deviation) in studies to calculate the effect size

2.1.3 Coding of studies

A coding page was created to examine in detail the features of 40 studies that met the criteria specified in the study. With this coding method, which gives the chance to examine according to their various characteristics, studies can be examined according to common and different groups. In this study, an Excel table was created with the name of the study, the year of the study, the names and surnames of the author(s), the design of the study, the parameters of the control and experimental groups to calculate the effect sizes, and 40 study information was coded to be able to code comfortably.

2.2. Analysis of Data

Effect sizes were calculated and weighted with the findings of the studies considered as research data. Comprehensive Meta-Analysis (version 3) program was used to calculate both effect sizes and overall effects in the process. The mean, standard deviation, and sample size data of the experimental and control groups for 40 studies were used to calculate the overall effect. In the study, the random-effects model was chosen in line with the knowledge that theoretically, there is not a single effect size from different studies and that there are different effects in the study. The suitability of this model selection was also determined using the heterogeneity test. The effect size Cohen's d coefficient was calculated using independent groups, and the J coefficient was used to transform this d coefficient into Hedge's g coefficient.

The effect size, Cohen's d coefficient, was calculated using independent groups and the J coefficient was used to transform this d coefficient into Hedge's g coefficient. Using the standardized mean difference (δ) obtained from the use of two different groups in the study, the effect size is calculated using the following equation (Cohen, 1969):

$$d=\frac{\overline{X_1}-\overline{X_2}}{S_{overall}} \quad (1)$$

To calculate the standard deviation using sample sizes of two independent groups (Cohen, 1969);

$$S_{Within}=\sqrt{\frac{(n_1-1)S_1^2+(n_2-1)S_2^2}{n_1+n_2-2}} \quad (2)$$

used.

The third and fourth equations were applied to obtain the variance and standard error of the calculated effect size (Hedges, 1981).

$$V_d=\frac{n_1+n_2}{n_1n_2}+\frac{d^2}{2(n_1+n_2)} \quad (3)$$

$$SE_D=\sqrt{V_d} \quad (4)$$

In estimating the mean difference (δ) between two independent groups over the Cohen d coefficient, slightly biased results can be obtained. To correct this and to make a more unbiased estimation, the J transform is performed for the conversion to Hedge's g coefficient. The statistical solution for transforming Cohen's d coefficient to Hedge's g coefficient statistically is presented in the fifth equation.

$$J=1-\frac{3}{4df-1} \quad \text{ve} \quad g=J*d \quad (5)$$

3. Findings

In this part of the research, first, the type of publication, year, percentage and frequency of the calculated effects are presented in the coding method. Then, publication bias was examined in the meta-analysis and CMA results were included for the overall effect.

3.1. Descriptive Statistics of Studies

The publication type of the 40 studies included in the study and the descriptive analyzes indicating the direction of the calculated effect are given in Table 2, Table 3 and Table 4, respectively

Table 2: Frequency And Percentage Values for The Type of Publication of The Studies İncluded in The Meta-Analysis Study

Publication Type	Frequency (f)	Percentage (%)
Article	17	42,5
Master Thesis	21	52,5
Doctoral thesis	2	5

Table 2 presents the percentage and frequency values showing the distribution according to the type of publication. Of the 40 studies that met the criteria for meta-analysis, 17 (42.5%) were written as articles, 21 studies (52.5%) were master's thesis and the remaining 2 (5%) were doctoral thesis. Among the 40 studies, the most suitable type of study consists of the theses written in the type of master's degree.

Table 3: Frequency And Percentage Values Of The Studies Included In The Meta-Analysis Study By Years

Year	Frequency (f)	Percentage (%)
2002	1	2,5
2005	1	2,5
2006	7	17,5
2007	4	10
2008	6	15
2009	2	5
2012	4	10
2014	2	5
2015	2	5
2016	1	2,5
2017	3	7,5
2018	3	7,5
2019	4	10

When Table 3 is examined, studies included in the meta-analysis according to years are seen. The study was planned to be composed of studies that met the criteria for meta-analysis between the years 2000-2021. Unfortunately, when the individual years are examined, there are also years in which there is no study that meets the desired criteria. Among the 40 studies, 2006 studies draw attention with maximum of 7 studies (17.5%). No studies are meeting the criteria in 2000, 2001, 2003, 2004, 2010, 2011, 2013, 2016, 2020 and 2021.

Table 4: Frequency And Percentage Values of The Calculated Effect Size Direction of The Studies Included ın The Meta-Analysis Study

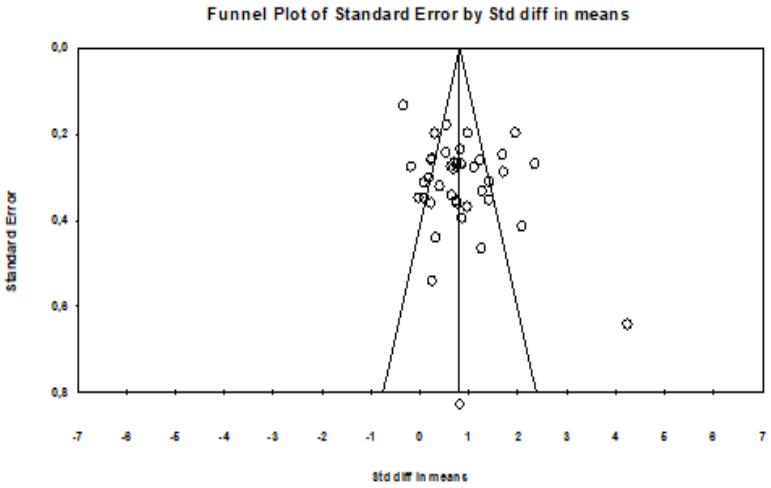
Direction of Calculated Effect	Frequency(f)	Percentage (%)
Negative Effect	3	7,5
Pozitive Effect	37	92,5

In the meta-analysis, the effect direction that has a significant effect, especially in publication bias, is discussed in Table 4. When the table is examined, it is seen that 3 of 40 studies (7.5%) have a negative effect of complementary measurement and evaluation techniques on academic achievement, while the remaining 37 studies (92.5%) have a positive effect.

3.2. Publication Bias in the Meta-Analysis Process

In the process of integrating the results of the studies selected in line with the criteria determined in the meta-analysis process, only studies that support certain hypotheses are selected, only studies with significant and favorable treatment effects are included, and some studies are excluded in the context of sample size, which causes publication bias. Various methods have been proposed to identify publication bias, remove and add studies. In this study, Funnel-Plot, Classic Fail-Safe N analysis, Egger regression estimation coefficient were calculated to determine publication bias. With Funnel Plot, the distribution of the effect sizes on the funnel plot is presented. The symmetric distribution of the sizes of the funnel plot indicates the absence of publication bias (Rothstein, Sutton, Borenstein, 2005; Sterne, Sutton, Ioannidis, Terrin, Jones, Lau & Higgins, 2011). The effect size distribution obtained in this study is given in Figure-1.

Figure 1: Funnel Plot for Effect Coefficients



When Figure 1 is examined, it is seen that the distribution has a symmetrical distribution except for one effect size. The study, which has an extreme effect size, has a very large sample size and therefore an extreme effect size was obtained.

It is a statistic that seeks to answer the question of how many more studies should be added so that the effect size average is 0 to reject the H0 hypothesis for the affected groups in the Classic Fail-Safe N analysis (Borenstein, Hedges, Higgins, & Rothstein, 2013).

The Classic fail-safe N was calculated as 4662($p < .01$). According to these results, it was seen that the number of studies to be added was quite high so that not the summary effect was significant. Egger's reggression test was used as the final analysis for publication bias. Egger's regression test, the regression intercept was not significant (intercept = 5,903, $p = .06$). The hypothesis was accepted to show that the regression constant didn't deviate from zero significantly.

3.3. Findings from Meta-Analysis

The effect sizes of 40 studies that met the criteria in the study were calculated. The effect sizes of 40 studies included in the meta-analysis are given in Table 5.

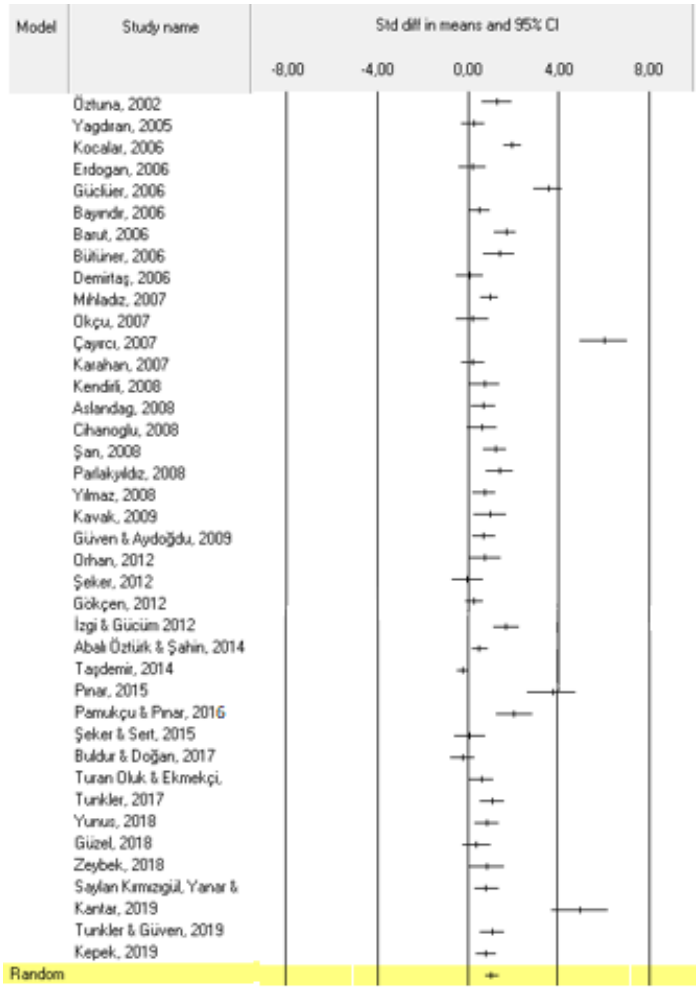
Table 5. Effect Sizes of Studies Included in the Meta-Analysis

Author(s) of the Study	Year	Effect Size	Standart Error	Variance
Öztuna	2002	1,275	0,332	0,11
Yağdıran	2005	0,241	0,257	0,066
Kocalar	2006	1,953	0,198	0,039
Erdoğan	2006	0,186	0,302	0,091
Güçlüler	2006	3,571	0,329	0,108
Bayındır	2006	0,533	0,243	0,059
Barut	2006	1,686	0,248	0,062
Bütüner	2006	1,409	0,353	0,125
Demirtaş	2006	0,076	0,313	0,098
Mıhladız	2007	0,975	0,198	0,039
Okçu	2007	0,224	0,361	0,13
Çayırcı	2007	6,024	0,534	0,285
Karahan	2007	0,228	0,259	0,067
Kendirli	2008	0,747	0,355	0,126
Aslandağ	2008	0,686	0,28	0,078
Cihanoğlu	2008	0,645	0,342	0,117
Şan	2008	1,224	0,26	0,068
Parlakyıldız	2008	1,416	0,31	0,096
Yılmaz	2008	0,759	0,267	0,071
Kavak	2009	0,969	0,368	0,136
Güven & Aydoğdu	2009	0,707	0,266	0,071
Orhan	2012	0,757	0,361	0,13
Şeker	2012	-0,03	0,348	0,121
Gökçen	2012	0,3	0,198	0,039
İzgi & Gücüm	2012	1,707	0,288	0,083
Abalı-Öztürk & Şahin	2014	0,539	0,179	0,032
Taşdemir	2014	-0,208	0,129	0,017
Pınar	2015	3,786	0,557	0,31
Şeker & Sert	2015	0,082	0,348	0,121
Pamukçu & Pınar	2016	2,081	0,414	0,171
Buldur & Doğan	2017	-0,183	0,275	0,076
Turan Oluk & Ekmekçi	2017	0,632	0,277	0,075
Tunkler	2017	1,107	0,27	0,077
Yunus	2018	0,846	0,321	0,103
Güzel	2018	0,403	0,395	0,156
Zeybek	2018	0,857	0,269	0,072
Saylan Kırmızıgül, Yanar & Kaya	2019	0,836	0,645	0,416
Kantar	2019	5,031	0,277	0,077
Tunkler & Güler	2019	1,107	0,236	0,056
Kepek	2019	0,828	0,147	0,022

When Table 5 is examined, it is seen that the effect sizes, standard error values and variance values of 40 studies are reported. Taşdemir's (2004) study has the lowest effect size with -0.208. The highest effect size is the study by Kantar (2019) with 5,031.

The overall effect values calculated according to the forest plots and random-effects model of the 40 studies included in the study are presented in Figure 2.

Figure 2. Florest Plot for Academic Achievement



In Figure 2, the effect sizes of all studies and the overall effect calculated according to the random effect model are visualized. As a result of the analysis, the overall effect is 1,080. The classification of effect sizes in line with the effect size comments of Thalheimer and Cook (2002) is presented in Table 6.

Table 6. Classification of Effect Sizes of Studies

Classification of Effect Size	f	%
Insignificant	5	12,5
Small	5	12,5
Medium	8	20
Wide	8	20
Very Wide	6	15
Enormous	8	20

Table 6 shows the distribution of the classification of effect sizes within 40 studies. 5 studies with insignificant effect size (12.5%), 5 studies with small effect size (12.5%), 8 studies with medium effect size (20%), 8 studies with large effect size (20%), very large effect size There are 6 studies (15%) and 8 studies (20%) of enormous / excellent effect size. The overall effect value obtained as a result of the meta-analysis of 40 studies in the study was found to be +1,080. The overall effect was found to have a very large effect size.

In the research, although it was determined by the literature review that the results of the study changed in the context of subgroups and the effect was not the only one in the universe and the random effect model was chosen. Using the random effect model, the results obtained are presented in Table 7.

Table 7. Overall Effect Size and Confidence Intervals for Heterogeneity by Random Effect Model

				τ^2
Model	N	Hedges' g	Lower Limit	Upper Limit
Rastgele Etkiler Modeli	40	1,08	0,792	1,369

According to the random effects model, it is seen that the lower limit of τ^2 in the 95% confidence interval of the effect sizes is 0.792, the upper limit is 1.369 and the overall effect size value is 1.080. The fact that the overall effect size value is (+), that is, in a positive direction, shows that the effect of TÖD techniques on the academic achievement of the students is positive. According to Thalheimer and Cook (2002), the overall effect size value obtained from the analysis is at a very large level. It can be said that the use of complementary measurement and evaluation techniques has a very high effect on increasing the academic achievement of students.

4. Discussion

In the study, it was aimed to investigate the effect of complementary measurement and evaluation techniques which made between 2000-2021 on academic achievement. For the research, meta-analysis was carried out with 40 studies in line with the criteria specified in the method section of the databases. Random effects model was used in the study. Of the 40 studies included in the research, 17 (42.5%) were articles and 23 (57.5%) were thesis studies. When these theses are examined, 2 of them are doctoral theses and the remaining 21 studies are master's theses. When the distribution of studies according to years is examined, it is seen that in 2002 and 2005, the least with one study (2.5%) and the highest number of studies with 7 studies (17.5%) in 2006. In the last 21 years, no studies that met the criteria were found in 2000, 2001, 2003, 2004, 2010, 2011, 2013, 2016, 2020 and 2021. When the effect size direction is examined, 3 studies (7.5%) have a negative effect, while 37 studies (92.5%) have a positive effect. When the individual effect sizes are examined, the lowest effect size is -0.208 in Taşdemir (2004). The highest effect size is the study by Kantar (2019) with 5,031. In the context of the effect size classification of Thalheimer and Cook (2002), when the effect sizes are examined individually, 5 studies with insignificant effect size (12.5%), 5 studies with small effect size (12.5%), 8 studies with medium effect size There are 8 studies with large effect size (20%), 6 studies with very large effect size (15%), and 8 studies with enormous/perfect effect size (20%). The overall effect value obtained as a result of the meta-analysis of 40 studies in the study was found to be +1.080. The overall effect appears to have a very large effect size. As a result of the intended meta-analysis study, the use of complementary measurement and evaluation techniques for academic achievement has a great effect on academic achievement.

The positive effect of the complementary measurement and evaluation tools used in the research on academic achievement is in parallel with the study of Özeren (2013). Özeren (2013) found in his study that complementary measurement and evaluation tools meet the expectations of students and teachers. As a result of his study, Özeren (2013) determined that these techniques make the student more active in the lesson. Başol and Erbay (2017) conducted a meta-analysis study using only portfolio, one of the complementary measurement and evaluation tools. The results of this research are consistent with the study of Başol and Erbay (2017). Başol and Erbay (2017) found that portfolio use has a positive and large-scale effect on academic achievement.

As a result of the research, it is suggested that complementary measurement and evaluation techniques used in the measurement of cognitive, affective and psychomotor skills of the individual should be used more in the education process.

5. References

Abalı Öztürk, Y., & Şahin, Ç. (2014). The Effects of Alternative Assessment and Evaluation Methods on Academic Achievement, Persistence of Learning, Self-Efficacy Perception And Attitudes. *Journal of Theory and Practice in Education 2014, 10(4),* 1022-1046.

Acar, A., Nemutlu, E., Gürhan, G., & Liman, V. (2004). Hacettepe Üniversitesi Eczacılık Fakültesi Araştırma Görevlilerinin İş Memnuniyeti ve Bunu Etkileyen Faktörler. *Hacettepe Üniversitesi Eczacılık Fakültesi Dergisi, 24*(2), 95-106.

Açıkgöz, K. (2003). *Aktif Öğrenme.* İzmir: Kanyılmaz.

Allen, D. D., & Flippo, R. F. (2002). Alternative Assessment in The Preparation of Literacy Educators: Responses from Students. *Reading Psychology, 23*(1), 15-26.

Aslandağ Soylu, B. (2008). *İngilizce Öğretiminde İşbirlikli Öğrenme Yönteminin İlköğretim 6. Sınıf Öğrencilerinin Akademik Başarılarına Etkisi.* (Yayınlanmamış Yüksek Lisans Tezi) Niğde Üniversitesi, Sosyal Bilimler Enstitüsü, Niğde.

Baker, E. L., O'Neil, H. F., & Linn, R. L. (1993). Policy And Validity Prospects for Performance-Based Assessment. *American Psychologist, 48*(12), 1210.

Barut, Ö. (2006). *İlköğretim 7. Sınıf Fen Bilgisi Konularının Kavram Haritaları ile Öğretilmesi.* (Yayınlanmamış Yüksek Lisans Tezi) Yüzüncü Yıl Üniversitesi, Fen Bilimleri Enstitüsü, Van.

Başol, G., & Erbay, Ş. (2017). Portfolyo Kullanımının Akademik Başarıya Etkisi: Bir Meta Analiz Çalışması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 32*(2), 396-412.

Bayındır, P. (2006). *İlköğretim Altıncı Sınıf Sosyal Bilgiler Dersi Coğrafya Konularında Kavram Haritalarının Öğrenci Başarısına Etkisi (Erzurum İli Ömer Nasuhi Bilmen İlköğretim Okulu Örneği).* (Yayınlanmamış Yüksek Lisans Tezi) Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü, Erzurum.

Berenson,S. (1995). Changing Assessment Practices in Science and Mathematics. *School Science Mathematics, 95*(4), 182–187.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2009). *Introduction To Meta-Analysis*. New York: John Wiley & Sons.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A Basic Introduction to Fixed-Effect and Random-Effects Models for Meta-Analysis. *Research Synthesis Methods, 1*(2), 97-111.

Buldur, S., & Doğan, A. (2017). Performansa Dayalı Tekniklerle Yürütülen Biçimlendirmeye Yönelik Değerlendirme Sürecinin Öğrencilerin Hedef Yönelimlerine Etkisi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 32*(1), 143-167.

Buldur, S., & Doğan, A. (2017). Performansa Dayalı Tekniklerle Yürütülen Biçimlendirmeye Yönelik

Bütüner, S. Ö. (2006). *Açılar ve Üçgenler Konusunun İlköğretim 7. Sınıf Öğrencilerine Vee Diyagramları ve Zihin Haritaları Kullanılarak Öğretimi.* (Yayınlanmamış Yüksek Lisans Tezi) Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü, Balıkesir.

Çayırcı, Ç. (2007). *İlköğretim 7. Sınıfta Web Tabanlı Portfolyo Uygulaması: Fen Bilgisi ve Sosyal Bilgiler Örnekleri.* (Yayınlanmamış Yüksek Lisans Tezi) Dokuz Eylül Üniversitesi, Eğitim Bilimleri Enstitüsü, İzmir.

Cihanoğlu, M. O. (2008). *Alternatif Değerlendirme Yaklaşımlarından Öz ve Akran Değerlendirmenin İşbirlikli Öğrenme Ortamlarında Akademik Başarı, Tutum ve Kalıcılığa Etkileri.* (Yayınlanmamış Doktora Tezi) Dokuz Eylül Üniversitesi, Eğitim Bilimleri Enstitüsü, İzmir.

Cohen, J. (1969). *Statistical Power Analysis for The Behavioral Sciences.* New York, NY: Academic.

Corcoran, C. A., Dershimer, E. L., & Tichenor, M. S. (2004). A Teacher's Guide to Alternative Assessment: Taking the First Steps. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 77*(5), 213-218.

Decoster, J. (2004). *Meta-Analysis Notes.* Retrieved from http://www.stat-help. com/notes.html.

Değerlendirme Sürecinin Öğrencilerin Hedef Yönelimlerine Etkisi. *Journal of Education, 32*(1), 143-167.

Demirtaş, B. (2006). *Kimya Deneylerinde "V" Diyagramları ile Öğretim Etkinliğinin İncelenmesi.* (Yayınlanmamış Yüksek Lisans Tezi) Dokuz Eylül Üniversitesi, Eğitim Bilimleri Enstitüsü, İzmir.

Doğan, D. (2005). *İlköğretim Öğretmen, Öğrenci ve Velilerin Dosya oluşturma (Portfolyo) Uygulamaları Hakkındaki Görüşleri Üzerine Bir Çalışma.* Yayımlanmamış Yüksek Lisans Tezi. Ankara: Ankara Üniversitesi Eğitim Bilimleri Enstitüsü.

Erdoğan, T. (2006). *Yabancı Dil Öğretiminde Portfolyoya Dayalı Değerlendirmenin Öğrenci Başarısı ve Derse Yönelik Tutumlarına Etkisi.* (Yayınlanmamış Yüksek Lisans Tezi) Dokuz Eylül Üniversitesi, Eğitim Bilimleri Enstitüsü, İzmir.

Erten, P (2015). *Çevrimiçi İşbirlikçi Öğrenme Ortamında E-Portfolyo Uygulamasının Akademik Başarıya, Tutumlara, Motivasyona ve Kalıcılığa Etkisi.* (Yayınlanmamış Doktora Tezi) Fırat Üniversitesi, Eğitim Bilimleri Enstitüsü, Elazığ.

Field, . P., & Gillett, R. (2010). How To Do a Meta-Analysis. *British Journal of Mathematical and Statistical Psychology, 63*(3), 665-694.

Field, A. P. (2001). Meta-Analysis of Correlation Coefficients: A Monte Carlo Comparison Of Fixed-And Random-Effects Methods. *Psychological Methods, 6*(2), 161.

Gipps, C., & Stobart, G. (2003). *Alternative Assessment. In International Handbook of Educational Evaluation* (pp. 549-575). Springer, Dordrecht.

Gökçen, B.B. (2012). Kavram Haritalarının Genel Biyoloji Dersine Yönelik Tutum ve Akademik Başarı Üzerine Etkileri. (Yayınlanmamış Yüksek Lisans Tezi) Çanakkale Onsekiz Mart Üniversitesi, Eğitim Bilimleri Enstitüsü, Çanakkale.

Goodrich, A. H. (2000). Using Rubrics to Promote Thinking and Learning. *Educational Leadership, 57*(5), 13-18.

Goodrich, H. (1996). *Student Self-Assessment: At the Intersection of Metacognition and Authentic Assessment. Doctoral dissertation.* Cambridge, MA: Harward University.

Güçlüer, E. (2006). *İlköğretim Fen Bilgisi Eğitiminde Kavram Haritaları ile Verilen Bilişsel Desteğin Başarıya, Hatırda Tutmaya ve Fen Bilgisi Dersine İlişkin Tutuma Etkisi.* (Yayınlanmamış Yüksek Lisans Tezi) Dokuz Eylül Üniversitesi, Eğitim Bilimleri Enstitüsü, İzmir.

Güven, E., & Aydoğdu, M. (2009). Portfolyonun 6. Sınıf Fen ve Teknoloji Dersi Vücudumuzda Sistemler Ünitesi'nde Başarı ve Kalıcılığa Etkisi. *Journal of Turkish Science Education, 6*(2), 115-128.

Güzel, Z. (2018). *Fen Bilimleri Öğretiminde Öz ve Akran Değerlendirme Uygulamalarının Yer Aldığı Probleme Dayalı Öğrenme Yaklaşımının Öğrencilerin Başarı ve Tutumlarına Etkisi.* (Yayınlanmamış Yüksek Lisans Tezi) Necmettin Erbakan Üniversitesi, Eğitim Bilimleri Enstitüsü, Konya.

Hedge, L. (1981). Distribution Theory for Glass's Estimator of Effect Size and Related Estimators. *Journal Of Educational Statistics, 6,* 107-128

İzgi, Ü. & Gücüm, B. (2012). Fen Eğitiminde Portfolyo Değerlendirme Kullanımının Sınav Kaygısı ve Öğrenmenin Kalıcılığı Üzerine Etkisi. *Education and Science, 37*(164), 71-80.

Kantar, N. (2019). *Alternatif Ölçme ve Değerlendirme Etkinliklerinin 6. Sınıf Öğrencilerinin Fen Bilimleri Dersindeki Başarısına ve Tutumuna Etkisi.* (Yayınlanmamış Yüksek Lisans Tezi) Atatürk Üniversitesi, Eğitim Bilimleri Enstitüsü, Erzurum.

Karahan, U. (2007). *Alternatif Ölçme ve Değerlendirme Metodlarından Grid, Tanılayıcı Dallanmış Ağaç ve Kavram Haritaları'nın Biyoloji Öğretiminde Uygulanması.* (Yayınlanmamış Yüksek Lisans Tezi) Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Karakaya, İ. (2007). Öğrenci Algısının Geliştirilmesi Öz Değerlendirme Akran Değerlendirme ve Grup Değerlendirme Formları. *İlköğretim Eğitimci Dergisi, 8*, 35-38.

Karalök, S. (2014). *Ortaokul Matematik Öğretmenlerinin Matematik Dersi Tamamlayıcı Ölçme Değerlendirme Tekniklerine İlişkin Profilleri.* Yayımlanmamış Yüksek Lisans Tezi. Denizli: Pamukkale Üniversitesi Eğitim Bilimleri Enstitüsü.

Kavak, S. (2009). *İlköğretim 8. Sınıf Fen ve Teknoloji Dersi "Maddenin Halleri ve Isı" Ünitesinde Kavram Haritası Tekniği Kullanımının Öğrencilerin Başarısına, Bilgilerin Kalıcılığına ve Fene Karşı Tutumlarına Etkisi.* (Yayınlanmamış Yüksek Lisans Tezi) Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Kendirli, B. (2008). *Fen ve Teknoloji Dersinde Kavram Haritası Kullanımının Öğrenci Tutumu, Başarısı ve Bilgi Kalıcılığına Etkisi.* (Yayınlanmamış Yüksek Lisans Tezi) Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Kepek, T. (2019). *Alternatif Değerlendirme Yöntemlerinin Öğrenci Başarısına Etkisinin ve Bu Yöntemlerle İlgili Öğretmen ve Öğrenci Görüşlerinin İncelenmesi.* (Yayınlanmamış Yüksek Lisans Tezi) Necmettin Erbakan Üniversitesi, Eğitim Bilimleri Enstitüsü, Konya.

Kocalar, A. O (2006). *Ortaöğretim Konularında Kavram Haritalarının Coğrafya Öğretiminde Kullanımı.* (Yayınlanmamış Yüksek Lisans Tezi) Marmara Üniversitesi, Eğitim Bilimleri Enstitüsü, İstanbul.

Kutlu, Ö., Doğan, C., & Karakaya, İ. (2014). *Ölçme ve Değerlendirme: Performansa ve Portfolyoya Dayalı Durum Belirleme.* Ankara: Pegem.

Kutlu, Ö., Karakaya, İ. & Doğan, C.D. (2008). Üst Düzey Zihinsel Becerilerin Belirlenmesi Performans Görevi Yazma, *İlköğretmen Eğitimci Dergisi, 18,* 10-15.

McMillan, J. H. (2007). *Classroom Assessment: Principles and Practice for Effective Standarts-Based Instruction. (4th Edition).* USA: Pearson Education.

Mıhladız, G. (2007). *İlköğretim Fen Bilgisi Öğretiminde Portfolyo Uygulamasının Öğrencilerin Akademik Başarılarına ve Derse Yönelik Tutumlarına Etkisi.* (Yayınlanmamış Yüksek Lisans Tezi) Muğla Üniversitesi, Fen Bilimleri Enstitüsü, Muğla.

Milli Eğitim Bakanlığı (MEB) (2017). *Fen Bilimleri Dersi Öğretim Programı ve Kılavuzu.* Ankara.

Mutluer, C & Okyay, Ö. (2017). *Öğretmen Adaylarının Portfolyo Uygulamasına Karşı Tutumlarının İncelenmesi. Sözlü Bildiri,* 26th International Congress Educational Sciences (ICES/UEBK-2017), Antalya.

Mutluer, C. &. Nartgün, Z. (2017). *Öğretmen Adaylarına Uygulanan Portfolyo Uygulamasının Akademik Öz düzenlemeStratejileri ve Akademik Başarı İçin Arabulucu Etkisi. Sözlü Bildiri,* 26. Uluslararası Eğitim Bilimleri Kongresi, Antalya.

Nitko, A. J. (1996). *Educational Assessment of Students.* Prentice-Hall Order Processing Center, PO Box 11071, Des Moines, IA 50336-1071.

Okçu, Y. (2007). *Matematik Eğitiminde Portfolyo Değerlendirme.* (Yayınlanmamış Yüksek Lisans Tezi) Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü, Balıkesir.

Oluk, N. T., & Ekmekci, G. (2017). Alternatif Değerlendirme Teknikleri İle Klasik Değerlendirme Tekniklerinin Öğrenci Başarısını Ölçme Açısından Karşılaştırılması. *Eğitim ve Toplum Araştırmaları Dergisi, 4*(2), 172-199.

Oosterhof, A. (2003). *Developing Using Classroom Assessment.* USA: Pearson Education.

Orhan, A. (2012). *Etkinliklerinin 6. Sınıf Fen ve Teknoloji Dersi Yaşamımızdaki Elektrik Ünitesindeki Öğrenci Başarısına Etkisi.* (Yayınlanmamış Yüksek Lisans Tezi) Ahi Evran Üniversitesi, Fen Bilimleri Enstitüsü, Kırşehir.

Özeren, E. (2013). *Alternatif Ölçme ve Değerlendirme Araçları Üzerine Bir Meta Analiz Çalışması.* Yayımlanmamış Yüksek Lisans Tezi. Elazığ: Fırat Üniversitesi Eğitim Bilimleri Enstitüsü.

Öztuna, A. (2002). *Kavram Haritalarının Grup Döngüsünde Yapılandırılmasının Başarıya ve Kavram Gelişimine Etkisi.* (Yayınlanmamış Doktora Tezi) Marmara Üniversitesi, Eğitim Bilimleri Enstitüsü, İstanbul.

Pamukçu, C., & Pınar,A. (2016). Tamamlayıcı Ölçme ve Değerlendirme Gelişim Programının Coğrafya Öğretmen Adaylarının Bilgi Düzeyi Üzerine Etkisi. *Selçuk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 35*, 225-235.

Parlakyıldız, B. (2008). *Üniveriste Öğrencilerinin Portfolyoya Dayalı Değerlendirmenin Üniversite Öğrencilerinin Akademik Başarılarına ve Bişilsel Yaşam Becerilerine Etkisi.* (Yayınlanmamış Doktora Tezi) Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.

Pınar, A. (2015). *Tamamlayıcı Ölçme ve Değerlendirme Gelişim Programının Coğrafya Öğretmen Adaylarının Yeterlik Algısı Üzerine Etkisi.* (Yayınlanmamış Doktora Tezi) Necmettin Erbakan Üniversitesi, Eğitim Bilimleri Enstitüsü, Konya.

Rothstein, H. R., Sutton, A. J., & Borenstein, M. (2005). *Publication Bias in Meta-Analysis: Prevention, Assessment and Adjustments.* England: John Wiley & Sons.

Şan, İ. (2008). *Lise II. Sınıf Biyoloji Dersinde Okutulan Bitkilerde Taşıma Sistemi Konusunun Kavram Haritalarıyla Öğretilmesinin Başarıya Etkisi.* (Yayınlanmamış Yüksek Lisans Tezi) Selçuk Üniversitesi, Fen Bilimleri Enstitüsü, Konya.

Saylan Kırmızıgül, A., Yanar, S. & Kaya, H. (2019). Fen Bilimleri Dersinde Tamamlayıcı Ölçme ve Değerlendirme Yaklaşımlarından Portfolyo Kullanımının Akademik Başarı ve Kalıcılığa Etkisinin Araştırılması. *Journal of Social Sciences of Mus Alparslan University,7*(2), 21-30.

Şeker, F. & Sert,H. (2015). Fen Bilimleri Dersinde Tamamlayıcı Ölçme ve Değerlendirme Yaklaşımlarının Tutum ve Başarıya Etkisi. *Mediterranean Journal of Humanities, 2*, 351-363.

Şeker, F. (2012). *İlköğretim Fen ve Teknoloji Dersinde Tamamlayıcı Ölçme ve Değerlendirme Yaklaşımlarının Tutum ve Başarıya Etkisi.* (Yayınlanmamış Yüksek Lisans Tezi) Akdeniz Üniversitesi, Sosyal Bilimler Enstitüsü, Antalya.

Seker, F., & Sert, H. (2015). Fen Bilimleri Dersinde Tamamlayıcı Ölçme ve Değerlendirme Yaklaşımlarının Tutum ve Başarıya Etkisi. *Akdeniz İnsani Bilimler Dergisi, (2)*, 351-363.

Sterne, J. A., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., … Higgins, J. P. T. (2011). Recommendations For Examining And İnterpreting Funnel Plot Asymmetry in Meta-Analyses of Randomised Controlled Trials. *BMJ, 343*, 1-8.

Taşdemir,M. (2014).Kendini Değerlendirme, Akran Değerlendirme ve Öğretmen Değerlendirmenin Yazılı Sınav Sonuçlarına Etkisi Ve Başarı Yordayıcılığı. *International Periodical for The Languages, Literature and History of Turkish or Turkic*, *9*(5), 1911-1929.

Thalheimer, W., & Cook, S. (2002). How To Calculate Effect Sizes from Published Research: A Simplified Methodology. *Work-Learning Research, 1,* 1-9.

Tünkler, V. & Güven, C. (2019). Mikroöğretim Uygulamasının Öğretmen Adaylarının Tamamlayıcı Ölçme-Değerlendirme Tekniklerine Yönelik Okuryazarlık Düzeylerine Etkisi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 34*(2), 541-564.

Tünkler, V. (2017). *Sosyal Bilgiler Öğretmen Adaylarının Tamamlayıcı Ölçme-Değerlendirme Tekniklerine Yönelik Okuryazarlık Düzeylerinin Mikroöğretim Aracılığıyla Geliştirilmesi.* (Yayınlanmamış Doktora Tezi) Necmettin Erbakan Üniversitesi, Eğitim Bilimleri Enstitüsü, Konya.

Tünkler, V., & Güven, C. (2018). Mikroöğretim Uygulamasının Öğretmen Adaylarının Tamamlayıcı Ölçme-Değerlendirme Tekniklerine Yönelik Okuryazarlık Düzeylerine Etkisi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, 34*(2), 541-564.

Turan Oluk, N. & Ekmekçi, G. (2017). Alternatif Değerlendirme Teknikleri ile Geleneksel Değerlendirme Tekniklerinin Öğrenci Başarısını Ölçme Açısından Karşılaştıtılması. *Eğitim ve Toplum Araştırmaları Dergisi, 4*(2),172-199.

Von Glasersfeld, E. (1989). Cognition, Construction of Knowledge, And Teaching. *Synthese, 80*(1), 121-140.

Yağdıran, E. (2005). *Ortaöğretim 9. Sınıf Fonksiyonlar Ünitesinin Çalışma Yaprakları, Vee Diyagramları ve Kavram Haritası Kullanılarak Öğretilmesi.* (Yayınlanmamış Yüksek Lisans Tezi) Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü, Balıkesir.

Yılmaz, H. (2008). *İlköğretim Birinci Kademe 5. Sınıf Sosyal Bilgiler Dersinde Kavram Haritalarının Kullanılmasının Başarıya Olan Etkisi* (Yayınlanmamış Yüksek Lisans Tezi)

Yunus, Ö. (2018). *Alternatif Ölçme ve Değerlendirme Tekniklerinin 6. Sınıf "Bitki ve Hayvanlarda Üreme, Büyüme ve Gelişme" Ünitesinde Kullanılmasının Öğrencilerin Akademik Başarılarına Etkisinin İncelenmesi.* (Yayınlanmamış Yüksek Lisans Tezi) Mustafa Kemal Üniversitesi, Fen Bilimleri Enstitüsü, Hatay.

Yurdabakan, İ. (2011). The View of Constructivist Theory on Assessment: Alternative Assessment Methods in Education. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi, 44*(1), 51-78.

Zeybek, G. (2019). Veritabanı Organizasyonu Dersinde Elektronik Portfolyo Uygulamalarının Akademik Başarıya Etkisi. *Afyon Kocatepe Üniversitesi Sosyal Bilimler Dergisi, 21*(4), 1045-1058.