Research Article

The Effect of Multiple Intelligences Theory-Based Education on Academic Achievement: A Meta-Analytic Review

Gökhan Baş
Niğde University

Abstract

The main purpose of this study is to determine the effect of multiple intelligences theory (MIT)-based education on students’ academic achievement. In this research, the meta-analytic method was adopted to determine this effect, and studies related to this subject carried out in Turkey were compiled. The effect sizes of the studies included in the meta-analysis were compared in terms of the variables of educational level, course type, being a postgraduate thesis or doctoral dissertation, geographical region, sample size, and duration of experimental procedure. Some specific criteria were taken into consideration for including studies in the meta-analysis. In accordance with these criteria, experimental postgraduate theses/dissertations (N = 75) were included in the research. At the end of the research, MIT-based education was determined to have a large effect size (d = 1.077) on students’ academic achievement levels. Additionally, while the effect sizes obtained in the research as a result of the meta-analysis were compared in terms of the variables of educational level, course type, postgraduate thesis/dissertation, geographical region, sample size, and duration of experimental procedure, no significant difference was seen in regard to the effect sizes of studies with respect to any variable apart from educational level.

Keywords
Theory of multiple intelligences-based instruction • Academic achievement • Meta-analysis • Postgraduate theses/dissertations
The questions of what intelligence is and how it is defined are issues that have occupied mankind for centuries (Selçuk, 2005). As studies on human intelligence had gained great momentum by the end of the 19th century, this led to the emergence of various theories on the issue (Denig, 2004). The first studies on intelligence can be said to be Darwin’s studies on animal intelligence, and his nephew Galton’s studies on human intelligence (Boring, 1950). At the beginning of the 20th century, Binet and Simon in particular made significant contributions to the understanding of human intelligence (Armstrong, 2000; Gardner, 1993a). Also, the studies of Spearman and Thurstone brought great innovations and expansions to the understanding of intelligence itself (Bümen, 2005). Traditionally, intelligence has been widely accepted by psychologists as a general intelligence, and it has been defined as a general capacity for conceptualization and problem solving that can be measured through IQ tests (Visser, Ashton, & Vernon, 2006). The concept of intelligence took its first step with Binet’s implementation of IQ tests (Zazzo, 1993). Since then, many intelligence tests have appeared mostly measuring verbal memory, verbal reasoning, numerical reasoning, understanding of logical sequences, and ability to state how one would solve every-day problems (Gardner, 1999). However, the concepts of intelligence and IQ gained a different dimension with the publication of Howard Gardner’s (1993a) inspirational book Frames of Mind: The Theory of Multiple Intelligences in 1983. Rather than defining intelligence in terms of mental test results or IQ scores, Gardner defined intelligence as a psychobiological information processing capacity to solve problems or fashion products that are valued in at least one community and culture (Kornhaber, Fierros, & Veenema, 2004). After years of study and observation on gifted and handicapped children, Gardner (1999) concluded that the mental potential of a human does not consist of a single intelligence but of multiple intelligences processed autonomously and localized in different parts of the brain. In his theory of multiple intelligences (MIT), Gardner (1993a) argues for a pluralistic view of the brain rather than a single part responsible for many functions. Also, according to MIT, the human brain has multiple sections, each functioning independently to modify our life. However, when we also attend to a complex task, these units engage their power and work in great harmony (Gardner, 1993a).

Meanwhile, after Gardner’s (1993b) Frames of Mind: The Theory into Practice was published, the understanding that was brought to intelligence was transformed into applications. MIT emerged to become popular in the field of educational sciences after 1998, with various studies conducted on the issue (e.g., Başbay, 2000; Bümen, 2001; Coşkungönüllü, 1998; Demirel et al., 1998; Talu, 1999; Tarman, 1999). Furthermore, aside from all the literature and results supporting this theory, some criticisms are also seen to have been made on the theory itself. One of the main points of these criticisms is whether the eight potentials, which are accepted as modalities of intelligence in the theory, form abilities or mental domains (White & Breen, 1998). It has also been stated
that, while MIT is consistent with much empirical evidence, it has not been subjected to
strong experimental tests (Denig, 2004). MIT is considered to be too broad to be useful
for planning curriculum, as well as that it presents a static view of student competence
(Klein, 1997). However, although there are some criticisms about MIT, very few
theories in the scope of education have been said to create such a significant impact as
MIT has done on teaching and learning (Saban, 2009; Shearer, 2004).

Multiple Intelligences Theory

As a neuropsychology and development specialist, Gardner started to analyze the
cognitive capacity of individuals in the 1970s and 80s after analyzing the traditional
concept of intelligence. When he was working in Boston University, he tried to understand
the talent patterns and determine the effects of cognitive or affective accidents. At the same
time, he conducted a study on normal and gifted children to analyze the developments of
cognitive talents in a project called “Project Zero” at Harvard University (Bümen, 2005).
Gardner (1993a) conducted the research on geniuses, intelligent people, brain-damaged
patients, savants, normal children, and normal adults.

Gardner, the developer of MIT, defines intelligence beyond its classical definitions
as “... a talent of an individual’s effort on revealing a product which is considered
valuable in one or more than one culture, or [sic] a talent or individual’s ability to
solve daily or professional life problems effectively and efficiently;” he also adds
that there are at least eight areas of intelligence (Oral, 2004). Gardner (1993a) called
people’s talents, which come in different forms, potentials, and abilities, modalities
of intelligence, thereby giving a broader perspective on intelligence. When Gardner
(1993a) revealed his multiple intelligences theory in 1983, he first mentioned seven
areas of intelligence; however, he later mentioned the existence of an eighth area of
intelligence in an interview with Checkley (1997). Thus, Gardner (1999) added the
eighth intelligence modality to his original list of seven in his book Intelligences
Reframed. In that case, the eight modalities of intelligence as revealed by Gardner
(1993a; 1993b; 1999) can be listed as: (a) verbal-linguistic, (b) logical-mathematical,
(c) visual-spatial, (d) bodily-kinesthetic, (e) musical-rhythmic, (f) social-individual,
(g) personal-intrapersonal, and (h) natural-naturalist.

While the classical IQ condemns intelligence into a single dimension by digitizing
it, MIT handles intelligence in multiple dimensions. Gardner also mentions that
intelligence is seen in every action, result, and problem-solving procedure, so it
cannot be calculated numerically (Saban, 2004). While the supporters of the IQ
test defend that intelligence is inherited and cannot be changed, MIT defends that
intelligence can be affected by many factors and cannot only be inherited (Campbell
& Campbell, 1999). MIT emphasizes that undeveloped areas of intelligence can be
developed and, subsequently, change (Silver, Strong, & Perini, 2000).
Gardner (1993a) claimed that all people have at least eight, maybe even nine, modalities of intelligence. In this sense, MIT suggests that intelligence cannot be in one dimension and that each individual has a variety of intelligences at different levels. Thus, MIT provides a faculty such as curriculum with the ability to strengthen and value students’ individual differences through teachers; creative use of the basic principles of this theory, emphasizing individual learning styles, interests, talents, and dispositions (Selçuk, Kayılı, & Okut, 2000). Moreover, MIT poses a template for the task of improving students’ academic achievements (Armstrong, 2000; Hoerr, 2000). Based on the research about MIT, it has become very popular among families and educators as a result of being student-centered, providing depth of learning through various educational processes, and providing different educational dimensions in the educational environment (Iyer, 2006).

MIT is not just an intelligence theory (Campbell, Campbell, & Dickinson, 1996). MIT is more widely accepted as a method of instruction, beyond just improving students’ modalities of intelligence (Campbell et al., 1996; Fasko, 2001; Lazear, 1998). Thus, many activities and models can be administrated in the process of teaching and learning that are based on this theory (Demirel, Başbay, & Erdem, 2006). Regardless of which model is chosen, even the planning stages of these activities are full of fun, colorful, and creative works. Certainly this theory brings a new point of view to educators. The significant difference brought by this theory emphasizes individual differences instead of defining one as smart or stupid (Bümen, 2005). Providing education that is appropriate to students’ intelligence types increases students’ academic achievements. Conversely, when students cannot get an education suited to their intelligence modalities, then they have difficulty learning (Bacanlı, 2006). At this point, the teacher has to continuously shift from one intelligence modality to another in MIT-based classrooms (Armstrong, 2000). According to MIT, everyone has different styles and speeds of learning. Even teachers’ consideration of this concept in the teaching and learning process affects many variables. Students can gain self-confidence, self-knowledge, creative thinking, respect for individual differences, and also easily learn and start thinking about professions for themselves with the help of educational applications based on this theory (Bümen 2005).

MIT emphasizes that students’ unimproved areas of intelligence can be improved at a later point (Silver et al., 2000). It can be accepted as a method of instruction, as well as an educational approach for improving students’ modalities of intelligence (Fasko, 2001), such that “it makes its greatest contribution to education by suggesting that teachers need to expand their repertoire of techniques, tools, and strategies beyond the typical linguistic and logical ones predominantly used” (Armstrong, 2009, p. 54).
Significance of the Study

Nowadays an increasing number of studies from all areas emerge with conflicting results on a number of topics. This situation leads to problems in attaining the objectives of research. For this reason, studies need to be gathered under one roof and re-analyzed (Demir & Başol, 2014). However, the increasing number of these studies inhibits data analysis from reaching the desired information. In this case, a wide analytical approach is needed that leads to new studies and interprets the findings from studies (Kablan, Topan, & Erkan, 2013). In this context, combining the results obtained from studies made in recent years is necessary as a basis for all related studies. This meta-analysis intends to combine results related to the effect of MIT-based instruction, which has grown popular in Turkey, on their students’ academic achievement levels. In this regard, combining the findings of these studies through meta-analysis will combine the findings obtained from small-scale individual studies that have been made at different times and places (Yıldız, 2002).

Meta-analysis means the quantitative combining of individual studies that have been performed for the same or similar issues under a common measure to discover common results. With individual studies, results can be provided regarding the effects of a dependent variable (i.e., a specific method of instruction) on an independent variable (i.e., an attitude towards a course); however, this prevents one from revealing the big picture about the effects of an independent variable on a dependent variable. The meta-analytic method, although it has some critics in the related literature, is thought to be important in revealing the overall picture of studies that have been carried out in an area with solid theoretical foundations using statistics. The main purpose of this meta-analysis is to determine the effect of MIT-based education on students’ academic achievement levels. Although meta-analyses are very well-known abroad, this method has recently begun to be used in Turkey. Thus the number of meta-analytic studies in the literature of Turkey is few (e.g., Aydin, Kaşarcı, & Yurt, 2012; Batdı, 2014; Camnalbur & Erdoğan, 2008; Çelik, 2013; Demir & Başol, 2014; Gözüyeşil & Dikici, 2013; Kablan et al., 2013; Kaşarcı, 2013; Kaşarcı, Aydin, Yurt, & Sünbül, 2012; Okursoy, 2009; Özdemirli, 2013; Semerci & Batdı, 2015; Şahin, 2005; Şen & Yılmaz, 2013; Tomakin & Yeşilyurt, 2013; Toraman & Demir, 2016; Yeşilyurt, 2012). A meta-analysis takes studies obtained from the relevant literature as a whole, and identifies the common aspects of these studies in terms of the effectiveness of certain variables (i.e., a particular method of instruction, academic achievement, attitudes towards a course, etc.). In the studies of this meta-analysis, the effectiveness of different instruction methods (i.e., problem-based learning, project-based learning, cooperative learning, computer-assisted learning, etc.) based on certain variables inspired by the curriculum is determined in the study. Reflections on MIT over curriculum without any meta-analytic study having been performed are considered to be a significant deficiency. Indeed, in 2005, through MIT’s implementation into
the curriculum of Turkey, this theory became an increasingly popular method with a significant increase in the number of studies based on this issue (see Table 1). Therefore, in addressing this deficiency in the relevant literature, studies about the effectiveness of MIT-based instruction on students’ academic achievement will be conducted using the meta-analytic model. Through this model, it is believed that Turkey can reach important evidence regarding the effectiveness of MIT-based instruction on students’ academic achievement levels. Findings obtained from the research regarding its effectiveness can be presented to education politicians and curriculum-development experts, as well as to teachers and school principals.

**Purpose of the Study**

The main purpose of this study is to determine the effect of MIT-based instruction on students’ academic achievement. In this research, the meta-analytic model was adopted to determine this effect and combine studies related to this subject that have been carried out in Turkey. In order to find an answer to the study’s purpose, the following sub-problems attempt to be answered in this meta-analysis:

1. What kind of effect does MIT-based instruction have on students’ academic achievement?

2. Is there a significant difference between the effect size of MIT-based instruction in regard to educational level (elementary school, high school, university)?

3. Is there a significant difference between the effect size of MIT-based instruction in regard to the type of course (social science, science, foreign language, vocational/technical)?

4. Is there a significant difference between the effect size of MIT-based instruction in regard to being either a postgraduate thesis or doctoral dissertation?

5. Is there a significant difference between the effect size of MIT-based instruction in regard to geographical region (central Anatolia, eastern Anatolia, Aegean, Marmara, Mediterranean, Black Sea, southeastern Anatolia) where the studies were conducted?

6. Is there a significant difference between the effect size of MIT-based instruction in regard to sample size (1-20, 21-40, 41+)?

7. Is there a significant difference between the effect size of MIT-based instruction in regard to the duration of experimental procedure (1-5 weeks, 6-10 weeks, 11+ weeks)?
Method

Research Model

In this study, the meta-analytic model proposed by Glass, McGaw, and Smith (1981) is used in order to calculate the effect size of MIT-based instruction on students’ academic achievement. A meta-analysis can be defined as a model statistically analyzing quantitative data, which has been obtained from a number of studies that are independent from each other, and reaching a general conclusion in terms of these studies’ results (Hunter & Schmidt, 2004; Lipsey & Wilson, 2001). Meta-analytic procedures require a number of steps: (a) locating all possible studies, (b) coding the studies for salient features and calculating effect sizes, and (c) carrying out statistical analyses of effect sizes and interpreting the acquired data (Höffler & Leutner, 2007). By considering these steps, analyses related to the meta-analytic method were performed in the current research.

Scope of Research

In order to determine the effect of MIT-based instruction on students’ academic achievement, quantitative studies (master’s theses and doctoral dissertations) carried out on MIT between 1998 and 2014 were examined.

The biggest problem with meta-analyses is that studies with strong effects and significant findings are viewed as convenient for publishing. However, if studies do not have any strong effects or statistical significances, these studies can be viewed as unworthy of publishing. In this regard, rarely are studies included in meta-analysis that have opposite or null effects (Rust, 1990). Hence, all relevant master’s theses and doctoral dissertations that could be accessed were included in this meta-analysis; studies were considered appropriate for their research content and for having the required data structures. Therefore, post-graduate theses and doctoral dissertations that took the effect of MIT-based instruction into account were considered in the content of this research. When reviewing the related literature in this context, the meta-analysis is seen to include studies carried out (as per the coding criteria) between 1998, when the earliest studies were conducted on the effect of MIT-based instruction on students’ academic achievement, and 2013 when the latest related study was conducted. Although the research reviewed studies between 1998 and 2014, no study was found conducted in 2014 that met the search criteria.

In the research, the studies included in the meta-analysis were determined based on certain criteria. First, studies published in Turkey between 1998 and 2014 and conducted to determine the effect of MIT-based instruction on students’ academic achievement were analyzed. In addition, in order to arrive at a standard effect size in the meta-analysis, included studies should have both experimental and control
groups. Single-group experimental designs or those without a control group were not included. Studies that adopted MIT-based instruction in the experimental group but not in the control group were also included. In studies involving more than one experimental group, data belonging only to the experimental group that was taught through MIT-based education were included in the analysis. Additionally, only one experimental group taught simultaneously based on MIT was selected randomly and included in the analysis. Furthermore, any study not reporting statistical data (i.e., mean scores, standard deviations, sample sizes), which are needed to calculate effect size, was excluded from the research. Also in this study, all postgraduate theses and doctoral dissertations examining the effect of MIT-based instruction in Turkey were included in the meta-analysis; however, studies conducted abroad were not.

Based on the above-mentioned criteria and in line with the research goals, postgraduate theses and doctoral dissertations stored in a PDF format in the Dissertations Center of the Higher Education Council (HEC) of Turkey were examined. For this, keywords such as instruction based on multiple intelligence theory, multiple intelligence theory, multiple intelligences and academic achievement, multiple intelligences instruction, and multiple intelligences achievement/success were searched in HEC’s Dissertations Center; as a result, a total of 231 studies were found and 179 of these studies were considered appropriate. Among these studies, some of them were not included in the analysis because they had no control group, lacked the necessary statistical data to calculate effect size, or because they could not be accessed from the Dissertations Center. Thus, a total of 75 studies (see Appendix 1) that had been conducted to examine the effect of MIT-based instruction on students’ academic achievement and which met the determined criteria constituted the sample of the current research.

Data Codification

In the study, 75 studies as research samples were included in the analysis, and the pretest and posttest mean scores and standard deviation values for the experimental and control groups were gathered in a data pool formed in Microsoft Excel 2007. Here, all studies were descriptively listed according to authors’ names and date of the thesis/dissertation; the number of samples in each study, their pre-test and post-test mean scores, and the standard deviations were gathered into the data pool. In the study, certain links were established between the authors’ surnames and the PDF files. Thanks to this, it was possible to access and find a study more easily. In addition, columns were prepared in the Microsoft Excel 2007 file for whether the studies included in the analysis were a master’s thesis or doctoral dissertation and its educational level (elementary school, high school, university), course type (social sciences courses, science courses, foreign languages courses, ability
courses), geographical region (central Anatolia, eastern Anatolia, Aegean, Marmara, Mediterranean, Black Sea, southeastern Anatolia), sample size (1-20, 21-40, 41+), and duration of experimental procedure (1-5 weeks, 6-10 weeks, 11+ weeks).

Data Analysis

In this study, the meta-analytic method of procedure effectiveness was used to analyze the data. The purpose of this method is to compare the effect sizes of independent-variable data used in more than one study by transforming the data into a common unit of measurement (Yıldız, 2002). In this method, the basic purpose is to calculate the differences between the mean scores of experimental and control groups in experimental studies with the formula $d = (X_e - X_c) / SD$ (see Hunter & Schmidt, 2004). In this formula, $X_e$ refers to the experimental group’s mean score, whereas $X_c$ refers to the control group’s mean score, and $SD$ refers to the pooled standard deviation value (see Borenstein, Hedges, Higgins, & Rothstein, 2009; Hedges & Olkin, 1985). The data collection tools in independent studies conducted by different researchers, the analyses of the data collected, and the results of related measurements were different from one another (Yıldırım, 2014). For this reason, in order to gather statistical data in different studies, the data should be transformed into an effect size (Şahin, 2005). In the meta-analysis of procedure effectiveness, standardized effect sizes, referred to as Cohen’s $d$ or Hedges’ $g$, are used (Başol-Göçmen, 2004; Borenstein et al., 2009; Cooper, 1989; Hartung, Knapp, & Sinha, 2008; Hedges & Olkin, 1985). The common point of these formulas is that all of them were developed for studies designed so as to involve a group (Yıldırım, 2014). The effect sizes of $d$ and $g$ are calculated by dividing the total standard deviation by the difference between the experimental and control groups’ mean scores (Borenstein et al., 2009; Şahin, 2005). In meta-analysis studies, calculation of the effect size is fairly important for obtaining accurate findings with standard deviations and for interpreting these findings (Lipsey & Wilson, 2001; Topçu, 2009). In this study, to calculate the effect size, Hedges’s $g$ was used and the significance level for the statistical analyses was taken as 95% (Borenstein et al., 2009; Hedges & Olkin, 1985). Furthermore, to interpret the effect sizes obtained as a result of meta-analysis, coefficient classification has been taken into account (Hartung, Knapp, & Sinha, 2008). While interpreting effect sizes in the current study, the effect size classification put forward by Cohen (1992) was adopted. According to Cohen’s (1992) coefficient for classifying effect size, effect size is small for values between 0.20 and 0.50; medium for values between 0.50 and 0.80, and large for values of 0.80 or higher. Also, if the effect size of a study is zero, it means that there is no effect. In this case, it is right to make an interpretation regarding whether or not there is an effect; however, if the effect size is negative, then it also means there is no effect (Çelik, 2013).
In the study, the most common approach, the chi-square heterogeneity test (Cochran’s $Q$), was used in order to see whether there was real heterogeneity between the included studies (Borenstein et al., 2009). According to the results of the heterogeneity test, a $p$-value lower than the accepted significance level demonstrates that the research results should be considered heterogeneous in line with the hypothesis put forth. Heterogeneous effect sizes demonstrate that studies are gathered from more than one distribution. The statistically significant chi-square value shows that the research results are homogeneous; in this case, the effect size can be used for all studies (Yıldırım, 2014).

The meta-analytic model to be used is fairly important for gathering studies which have different effect sizes, and thus for combining these studies (Borenstein et al., 2009). As the statistical model to be chosen will be influential on combining research results, the model should be chosen according to the research results (DeCoster, 2004; Dinçer, 2014). In a meta-analysis, there are two types of statistical research models for combining research results: the fixed effects model and the random effects model (see Borenstein et al., 2009; Rosenthal, 1991). Of these two models, the fixed effects model is based on the assumption that all the gathered studies should predict exactly the same effect when the data is homogeneous (Yıldız, 2002). In the fixed effects model, variance is thought to be a result of interrelated data (Shelby & Vaske, 2008). According to the fixed effects model, there is an effect size unit that measures the same effect size for all studies and is shared by these studies (Yıldırım, 2014). Using the random effects model seems more appropriate when the included studies are heterogeneous, as the fixed effects model is then inappropriate (DeCoster, 2004; Card, 2012; Durlak, 2005; Hedges, 1983). In the random effects model, analysis is conducted by considering the variance within studies and between studies (Okursoy, 2009). In other words, in meta-analyses formed using that model, both the differences between studies and within studies are included in the analysis (Shelby & Vaske, 2008). Thus, whether the effect sizes demonstrate homogeneous distribution or not should be tested when deciding on which statistical model to choose or use during meta-analysis. Based on the result of this test, the fixed effects model should be used if the effect sizes demonstrate a homogeneous distribution; if not, the random effects model should be adopted (Borenstein et al., 2009; Ellis, 2010). In this study, Comprehensive Meta-Analysis (CMA) Software (version 2.0; Borenstein et al., 2009) was performed to calculate the effect size and variances for each study; MS Excel 2007 was used to code and record the data.

Validity and Reliability of the Data
In the study, validity and reliability studies for each of the studies included in meta-analysis were seen to have been conducted. According to DeCoster (2004), validity and reliability in a meta-analysis depends on the individual validities and
reliabilities of all studies included in the meta-analysis. Additionally, including all the studies in a meta-analysis that meet the criteria for inclusion regarding a certain subject in the related literature influences the validity of the study (Borenstein et al., 2009; DeCoster, 2004). In this respect, depending on the purpose of the study, all studies conducted in Turkey were accessed. All accessed studies included in the meta-analysis were separately coded by two different researchers, and inter-rater agreement and reliability was calculated. Both researchers coded the data in their respective columns on the MS Excel file that had been created. Following this, the codings that had been done separately by the two researchers were compared. To calculate the inter-rater agreement and reliability for the two researchers, the formula \[ \text{Reliability} = \frac{\text{consensus}}{\left(\text{consensus} + \text{disagreement}\right) \times 100} \] was used, as suggested by Miles and Huberman (1994). In the study, the inter-rater value was calculated as 100%. Thus, the reliability for the two coders was found to be 100%, and data were thought to be reliable. For this reason, the two researchers were concluded to be in full consensus regarding the coding of the data.

Findings

In this part of the research, the meta-analytic findings of the studies in relation with the effect of MIT-based instruction on students’ academic achievement were given. In this study, descriptive data regarding the included studies are first presented; then the frequency and percentage values of these studies are calculated. Next, effect-size analyses of these studies are done, and the findings are presented through the help of tables. Additionally, comparisons are made through the fixed effects or random effects model, depending on the variable, to determine the effect of MIT-based instruction on students’ academic achievement levels.

Findings Regarding the Descriptive Features of the Studies

In this research, which examined the effectiveness of MIT-based instruction on students’ academic achievement, descriptive data and the effect sizes were determined for the studies \((N = 75)\) that met the inclusion criteria using sample sizes, pre-test and post-test mean scores, and standard deviation values. Descriptive data in relation to the included studies is given in Table 1.

According to Table 1, 85.33\% \((n = 64)\) of the studies included in the current meta-analysis were seen to be master’s theses; 14.66\% \((n = 11)\), doctoral dissertations; 69.33\% \((n = 52)\) were carried out in elementary schools; 28.0\% \((n = 21)\), high schools; and 2.66\% \((n = 2)\), universities. While 30.6\% \((n = 23)\) of the studies were seen to be carried out in central Anatolia, 14.6\% \((n = 11)\) were conducted in eastern Anatolia; 18.6\% \((n = 14)\), the Aegean; 17.33\% \((n = 13)\), the Marmara; 6.66\% \((n = 5)\), the Mediterranean; 9.33\% \((n = 7)\), the Black Sea; and 2.66\% \((n = 2)\), southeastern
Anatolia in Turkey. Also, 18.66% \( (n = 14) \) of the studies were seen carried out in social science courses; 69.33% \( (n = 52) \), science courses; 5.33% \( (n = 4) \), foreign languages courses; and 6.66% \( (n = 5) \), vocational/technical courses. In addition, 21.33% \( (n = 16) \) of the studies included in the research were understood to be carried out between 2000 and 2005; 73.33% \( (n = 55) \), between 2006 and 2010; and 5.33% \( (n = 4) \), after 2011. Of these studies, 12.0% \( (n = 9) \) were understood to have a sample size of between 1 and 20 subjects; 77.33% \( (n = 58) \), a sample size of 21-40; and 10.66% \( (n = 8) \), a sample size of 41+. Lastly, 56.0% \( (n = 42) \) of the studies included in the research were seen to have adopted a duration of experimental procedure of 1-5 weeks; 40.0% \( (n = 30) \), of 6-10 weeks; and 4.0% \( (n = 3) \), 11+ weeks.

**Findings Regarding the Entire Distribution of Effect Sizes**

The studies were combined into effect sizes with standard error and variance in the current meta-analysis. Therefore, a forest plot diagram was created in order to demonstrate confidence intervals, effect sizes, and their weight on the total effect sizes in the meta-analysis (see Figure 1).
**Figure 1.** Forest plot of the meta-analysis.
According to Figure 1, 73 studies of the 75 studies in the meta-analysis were seen to have positive effect sizes, whereas only two studies were seen to have a negative effect size. Thus, 97.33% of the studies (except the two indicated) can be said to suggest the effectiveness of MIT-based instruction on students’ academic achievement was positive. According to Wolf (1986), estimated effect sizes that are found to be positive mean that the performance is in favor of the experimental group, whereas estimated effect sizes that are found to be negative indicate that the performance is in favor of the control group. In this sense, performance can be claimed in favor of the experimental groups who had received MIT-based instruction.

Also, while the largest confidence interval in the meta-analysis was understood to belong to Başli (2006), the smallest one was seen to belong to Hepyaşar (2006; see Figure 1). The effect size of 43 studies was also large in the research; 19 studies, medium; nine studies, small; and four studies, unimportant, according to the classifications suggested by Cohen (1992).

Findings in Regard to the Effect Size of the Model

In order to put together the statistical data of different studies, the acquired data must first be converted to a common unit of measure, an effect size (Şahin, 2005). However, there is importance in the meta-analytic model used for combining the studies, which have different effect sizes from each other (Borenstein et al., 2009). As the chosen statistical model will affect how the study results are combined, the model must be selected in accordance with the results of these studies (DeCoster, 2004). There are two types of statistical models that combine study results in meta-analyses: the fixed effects and the random effects models (see Borenstein et al., 2009; Hartung, Knapp, & Sinha, 2008). In the research, Cochran’s $Q$ (chi-square heterogeneity test) was adopted, as it is the most common approach for evaluating heterogeneity of acquired data to determine which model to use. The result of the heterogeneity test yielded a $Q$-statistic of 458.793, while the chi-square table indicated a 95% significance level, and at 74 degrees of freedom the value was found to be 95.081. This $Q$ value indicates that the distribution of effect sizes in this collection of studies was heterogeneous, $Q (74) = 458.793, p < .001$ (see Table 2). In other words, the variance of the effect sizes of the studies included in the research is larger than can be explained by simple sampling error; therefore, a random effects model was adopted (see Borenstein et al., 2009; Hedges & Olkin, 1985; Yıldız, 2002; Dinçer, 2014; Özcan, 2008). The heterogeneity distribution value, average effect sizes, and confidence intervals of the studies included in the meta-analysis are given in Table 2.
Table 2
Studies’ Overall Effect Sizes, Heterogeneity Distribution Value According to Effects Model, and Confidence Intervals

<table>
<thead>
<tr>
<th>Model</th>
<th>k</th>
<th>ES</th>
<th>SE</th>
<th>Variance</th>
<th>Z</th>
<th>Q</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed Effects</td>
<td>75</td>
<td>0.946</td>
<td>0.031</td>
<td>0.001</td>
<td>30.138</td>
<td>458.793</td>
<td>0.885 – 1.008</td>
</tr>
<tr>
<td>Random Effects</td>
<td>75</td>
<td>1.077</td>
<td>0.079</td>
<td>0.006</td>
<td>13.582</td>
<td>458.793</td>
<td>0.921 – 1.232</td>
</tr>
</tbody>
</table>

Note. k = number of effect sizes; ES = effect size; SE = standard error; CI = confidence interval for the average value of ES.

When the total number of 75 studies included in the meta-analysis were compared according to the random effects model, the standard error was found to be 0.079 and the upper limit in the 95% confidence interval was 1.232, while the lower limit was 0.921 with an effect size of $d = 1.077$ (see Table 2). As a result of the acquired values, MIT-based instruction was revealed to be more successful on students’ academic achievement levels than other methods of instruction. According to the classification suggested by Cohen (1992), this value can be considered as a larger effect size. Also in the research, the result of the Z-test indicated that the effect size was statistically significant, $Z = 13.582, p < .000$.

Findings Regarding the Effect Sizes in Terms of Educational Level

The studies included in the meta-analysis were grouped under three main characteristics in terms of educational level: elementary school, high school, and university. The results of the analysis according to these groups are given in Table 3.

Table 3
Studies’ Effect-Size Values in Terms of Educational Level

<table>
<thead>
<tr>
<th>Educational Level</th>
<th>k</th>
<th>ES</th>
<th>SE</th>
<th>Variance</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower  Upper</td>
</tr>
<tr>
<td>Elementary School</td>
<td>52</td>
<td>0.925</td>
<td>0.081</td>
<td>0.007</td>
<td>0.766 – 1.084</td>
</tr>
<tr>
<td>High School</td>
<td>21</td>
<td>1.516</td>
<td>0.205</td>
<td>0.042</td>
<td>1.115 – 1.917</td>
</tr>
<tr>
<td>University</td>
<td>2</td>
<td>0.940</td>
<td>0.187</td>
<td>0.035</td>
<td>0.573 – 1.307</td>
</tr>
</tbody>
</table>

Note. k = number of effect size; ES = effect size; SE = standard error; CI = confidence interval for the average value of ES.

According to the results of analysis, the largest effect size was concluded to belong to the high-school level with $d = 1.516$ (95% CI = 1.115 – 1.917), while the smallest effect size belonged to the elementary-school level with $d = 0.925$ (95% CI = 0.766 – 1.084), as shown in Table 3. In addition, the result of the heterogeneity test yielded a $Q$-statistic of 7.317, while the chi-square table indicated a 95% significance level and the value at 2 degrees of freedom was found to be 5.991. The $Q$ value in the research indicates that the distribution of effect sizes in this collection of studies is heterogeneous, $Q_B$ (2) = 7.317, $p < .001$. Thus, the variance of the effect sizes of the studies included in the research is larger than can be explained by simple sampling error, so the random effects model was adopted. As a result, the effect of MIT-based instruction based on students’ academic achievement levels can be said to be significant in terms of educational level.
Findings Regarding Effect Sizes in Terms of Course Type

The studies included in the meta-analysis were grouped under four main characteristics in terms of course type: social science, science, foreign language, and tech courses. The results of the analysis according to these groups are given in Table 4.

<table>
<thead>
<tr>
<th>Course Type</th>
<th>k</th>
<th>ES</th>
<th>SE</th>
<th>Variance</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Science</td>
<td>52</td>
<td>0.863</td>
<td>0.038</td>
<td>0.001</td>
<td>0.788</td>
<td>0.938</td>
</tr>
<tr>
<td>Social Science</td>
<td>14</td>
<td>0.968</td>
<td>0.070</td>
<td>0.005</td>
<td>0.830</td>
<td>1.106</td>
</tr>
<tr>
<td>Foreign Language</td>
<td>4</td>
<td>1.958</td>
<td>0.169</td>
<td>0.029</td>
<td>1.627</td>
<td>2.290</td>
</tr>
<tr>
<td>Ability</td>
<td>5</td>
<td>1.139</td>
<td>0.105</td>
<td>0.011</td>
<td>0.932</td>
<td>1.346</td>
</tr>
</tbody>
</table>

Note. k = number of effect sizes; ES = effect size; SE = standard error; CI = confidence of interval for the average value of ES.

According to the results of analysis, the largest effect size was concluded to belong to foreign language courses with $d = 1.958$ (95% CI = 1.627 – 2.290), while the smallest effect size belonged to science courses with $d = 0.863$ (95% CI = 0.788 – 0.938), as shown in Table 4. The result of the heterogeneity test also yielded a $Q$-statistic of 7.085, while the chi-square table indicated 95% significance level and the value at 3 degrees of freedom was found out to be 7.815. The $Q$ value in the research indicates that the distribution of effect sizes in this collection of studies is homogeneous, $Q_B (3) = 7.085$. Hence, the fixed-effects model was adopted in the research. As a result, the studies included in the meta-analysis were seen to be homogeneous, thus there was not a significant difference in terms of the effect sizes when grouped according to course type. Therefore, the effect of MIT-based instruction on students’ academic achievement levels may be stated as insignificant in terms of course type in the research.

Findings Regarding Effect Sizes in Terms of Being a Postgraduate Thesis or Doctoral Dissertation

The studies included in the meta-analysis were grouped under two main characteristics: master’s thesis or doctoral dissertation. The results of the analysis according to these groups are given in Table 5.

<table>
<thead>
<tr>
<th>Type</th>
<th>K</th>
<th>ES</th>
<th>SE</th>
<th>Variance</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master’s Thesis</td>
<td>64</td>
<td>0.937</td>
<td>0.035</td>
<td>0.001</td>
<td>0.869</td>
<td>1.006</td>
</tr>
<tr>
<td>Doctoral Dissertation</td>
<td>11</td>
<td>0.984</td>
<td>0.072</td>
<td>0.005</td>
<td>0.842</td>
<td>1.125</td>
</tr>
</tbody>
</table>

Note. k = number of effect sizes; ES = effect size; SE = standard error; CI = confidence of interval for the average value of ES.
According to the results of the analysis, the largest effect size was concluded to belong to doctoral dissertations with $d = 0.984$ (95% CI = 0.842 – 1.125), while the smallest effect size belonged to master’s theses with $d = 0.937$ (95% CI = 0.869 – 1.006), as shown in Table 5. The result of the heterogeneity test also yielded a $Q$-statistic of 0.001, while the chi-square table indicated a 95% significance level with the value at 1 degree of freedom being 3.841. The $Q$-value in the research indicates that the distribution of effect sizes in this collection of studies is homogeneous, $Q_B(1) = 0.001$. Hence, a fixed-effects model was adopted in the research. As a result, the studies included in the meta-analysis were seen to be homogeneous; thus, there was not a significant difference in terms of the effect sizes grouped according to master’s thesis or doctoral dissertation. Therefore, the effect in this research of MIT-based instruction on students’ academic achievement levels is not significant in terms of being a Master’s thesis or doctoral dissertation.

### Findings Regarding Effect Size in Terms of Geographical Region

The studies included in this meta-analysis were grouped under seven main characteristics by geographical region in Turkey: central Anatolia, eastern Anatolia, Aegean, Marmara, Mediterranean, Black Sea, and southeastern Anatolia. The results of the analysis made according to these groups are given in Table 6.

<table>
<thead>
<tr>
<th>Geographical Region</th>
<th>k</th>
<th>ES</th>
<th>SE</th>
<th>Variance</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
</tr>
<tr>
<td>Central Anatolia</td>
<td>23</td>
<td>0.889</td>
<td>0.060</td>
<td>0.004</td>
<td>0.773 – 1.006</td>
</tr>
<tr>
<td>Eastern Anatolia</td>
<td>11</td>
<td>1.047</td>
<td>0.074</td>
<td>0.005</td>
<td>0.903 – 1.192</td>
</tr>
<tr>
<td>Aegean</td>
<td>14</td>
<td>0.989</td>
<td>0.075</td>
<td>0.006</td>
<td>0.852 – 1.149</td>
</tr>
<tr>
<td>Marmara</td>
<td>13</td>
<td>0.774</td>
<td>0.075</td>
<td>0.006</td>
<td>0.634 – 0.932</td>
</tr>
<tr>
<td>Mediterranean</td>
<td>5</td>
<td>1.107</td>
<td>0.117</td>
<td>0.014</td>
<td>0.888 – 1.352</td>
</tr>
<tr>
<td>Black Sea</td>
<td>7</td>
<td>0.938</td>
<td>0.100</td>
<td>0.010</td>
<td>0.750 – 1.147</td>
</tr>
<tr>
<td>Southeastern Anatolia</td>
<td>2</td>
<td>1.423</td>
<td>0.227</td>
<td>0.051</td>
<td>0.993 – 1.893</td>
</tr>
</tbody>
</table>

Note. k = number of effect sizes; ES = effect size; SE = standard error; CI = confidence of interval for the average value of ES.

According to the results of the analysis, the largest effect size was concluded to belong to southeastern Anatolia with $d = 1.423$ (95% CI = 0.993 – 1.893), while the smallest effect size belonged to the Marmara region with $d = 0.774$ (95% CI = 0.634 – 0.932), as shown in Table 6. The result of the heterogeneity test also yielded a $Q$-statistic of 2.570; the chi-square table indicated a 95% significance level and the value at 6 degrees of freedom was found to be 12.592. The $Q$ value indicates that the distribution of effect sizes in this collection of studies is homogeneous, $Q_B(6) = 2.570$. Hence, a fixed effects model was adopted in the research. As a result, the studies included in the meta-analysis were seen to be homogeneous, thus there was no significant difference in terms of the effect sizes grouped according to geographical
regions. Thus in the research, the effect of MIT-based instruction on students’ academic achievement levels is not significant in terms of Turkey’s geographical regions.

**Findings Regarding Effect Size in Terms of Sample Size**

The studies included in the meta-analysis were grouped under three main characteristics in terms of sample size: 1-20, 21-40 and 41+ participants. The results of the analysis according to these groups are given in Table 7.

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>K</th>
<th>ES</th>
<th>SE</th>
<th>Variance</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-20</td>
<td>9</td>
<td>1.100</td>
<td>0.120</td>
<td>0.014</td>
<td>0.864</td>
<td>1.336</td>
</tr>
<tr>
<td>21-40</td>
<td>58</td>
<td>0.984</td>
<td>0.036</td>
<td>0.001</td>
<td>0.912</td>
<td>1.055</td>
</tr>
<tr>
<td>41+</td>
<td>8</td>
<td>0.738</td>
<td>0.073</td>
<td>0.005</td>
<td>0.595</td>
<td>0.881</td>
</tr>
</tbody>
</table>

*Note.* k = number of effect sizes; ES = effect size; SE = standard error; CI = confidence of interval for the average value of ES.

According to the results of the analysis, the largest effect size was concluded to belong to sample sizes of 1-20 with $d = 1.100$ (95% CI = 0.864 – 1.336), while the smallest effect size belonged to sample sizes of 41 or more with $d = 0.738$ (95% CI = 0.595 – 0.881), as shown in Table 7. The result of the heterogeneity test also yielded a $Q$-statistic of 2.800; the chi-square table indicated a 95% significance level and the value at 2 degrees of freedom was found to be 5.991. The $Q$ value indicates that the distribution of effect sizes in this collection of studies is homogeneous, $Q_B (2) = 2.800$. Hence, a fixed effects model was adopted in the research. As a result, the studies included in the meta-analysis were seen to be homogeneous, thus there was no significant difference in terms of effect size grouped according to sample size. Therefore, the effect of MIT-based instruction on students’ academic achievement levels is not significant in terms of sample sizes in the research.

**Findings Regarding Effect Size in Terms of Duration of Experimental Procedure**

The studies included in the meta-analysis were grouped under three main characteristics in terms of duration of experimental procedure: 1-5 weeks, 6-10 weeks, and 11+ weeks. The results of the analysis made according to these groups are given in Table 8.
According to the results of analysis, the largest effect size was concluded to belong to procedural durations of 6-10 weeks with $d = 0.994$ (95% CI = 0.899 – 1.089), while the smallest effect size belonged to procedural durations of 11+ weeks with $d = 0.850$ (95% CI = 0.585 – 1.114), as shown in Table 8. The result of the heterogeneity test also yielded a $Q$-statistic of 3.122; the chi-square table indicated a 95% significance level and the value at 2 degrees of freedom was found to be 5.991. The $Q$ value indicates that the distribution of effect sizes in this collection of studies is homogeneous, $Q_B(2) = 3.122$. Hence, a fixed effects model was adopted in the research. As a result, the studies included in the meta-analysis were seen to be homogeneous, thus there was no significant difference in terms of effect size grouped according to duration of experimental procedure. In conclusion, the effect of MIT-based instruction on students’ academic achievement levels can be said to be insignificant in terms of duration of experimental procedure.

### Discussion

In this research, the theory of multiple intelligences (MIT), which has been a popular method of instruction particularly since its implementation in the 2005 elementary curriculum, was taken into consideration and studies that examined the effect of MIT-based instruction on students’ academic achievement levels were combined through a meta-analysis. The research tried to determine the effect sizes and the direction of MIT-based instruction on students’ academic achievement levels; also, the effect of MIT-based instruction on students’ academic achievement was compared in regard to certain variables: educational level, course type, master’s thesis or doctoral dissertation, geographical region, sample size, and duration of experimental procedure.

In the current research, the data of a total of 75 studies that met the inclusion criteria were examined under the meta-analysis; the overall effect size obtained in accordance with the analysis made using the fixed effects model was found to be $d = 0.946$. However, the emergence of heterogeneous studies after conducting the homogeneity test showed that the obtained data could not be generalized over the entire population. Therefore, the calculations made by the fixed effects model were re-performed according to the
random effects model. As a result of the analysis made by the random effects model, the overall effect size in the research was found to be $d = 1.077$ in total variance with a 0.079 standard error at 95% confidence interval (upper limit = 1.232 and lower limit = 0.921). Thus, the values obtained showed that MIT-based instruction is more successful than other educational methods. Hence, the successful results of MIT-based instruction can be concluded as an indication of its effectiveness in improving students’ academic achievement levels. In addition, the acquired effect size was determined to be large according to the classification suggested by Cohen (1992) in the research. According to the data obtained from the studies that were included in the meta-analysis, MIT-based instruction can be said to have positive effects on students’ academic achievement, taking place on a wide range of effect sizes at the largest level. Also, as the effect’s direction distribution of the studies included in the meta-analysis were looked at, no study was found with a zero value, whereas 93.33% ($n = 73$) of the studies had positive values, and 2.66% ($n = 2$) had negative values. Accordingly, almost all studies included in the meta-analysis can be said to have positive values. A review of the related literature shows that MIT-based instruction is more successful in improving students’ academic achievement levels than other methods of instruction. The results of research conducted nationally support many findings of research in the international literature. So, both nationally (e.g., Başlı, 2006; Çirakoğlu & Saracaloğlu, 2009; Karadeniz, 2006; Tunç, 2005; Uçak, Bağ, & Uşak, 2006; Yağcı, 2006) and internationally (e.g., Al-Balhan, 2006; Douglas, Burton, & Reese-Durham, 2008; Johnson, 2007; Nyugen, 2000), research literature shows that MIT-based instruction has a positive effect on students’ academic achievement levels. Therefore, the findings obtained from the current research are seen to be consistent with individual studies conducted in the national and international literature.

Although the overall effect size of MIT-based instruction on students’ academic achievement was examined in general, these effect sizes in regard to the variables of educational level, course type, master’s thesis or doctoral dissertation, geographical region, sample size, and duration of experimental procedure were analyzed in the research. In this sense, according to the analysis made in terms of educational level, the largest effect size was concluded to belong to high schools ($d = 1.516$), while the smallest effect size belonged to elementary schools ($d = 0.925$). According to the classification made by Cohen (1992), effect sizes were seen to be large in regard to educational level. When looking at the effect sizes of studies in the meta-analysis grouped according to educational level, the effect of MIT-based instruction on students’ academic achievement was seen to differ, $Q_B (2) = 7.317, p < .001$. According to this result, MIT-based instruction was understood to have a larger effect size in high schools than in elementary schools or universities. Thus, this result is considered significantly important. Additionally, even though a significant difference was seen in relation to effect size grouped according to educational level, the meta-analysis literature mostly shows that no significant difference between educational level in relation to the effect
sizes of applying various other methods of instruction in classrooms (e.g., Batdı, 2014; Camnalbur & Erdoğan, 2008; Gözüyeşil & Dikici, 2013; Kablan et al., 2013; Kaşarcı, 2013; Okursoy, 2009; Özdemirli, 2011; Şahin, 2005). Very few studies are also seen to report a significant difference between educational levels in the related literature (e.g., Çapar & Tarım, 2015; Nunnery, Chappell, & Arnold, 2013). Therefore, the reasons underlying the difference between educational levels in the current meta-analysis are recommended for deeper examination so as to reach a better understanding regarding the effect of MIT-based education on students’ academic achievement in terms of this variable. On the other hand, no study included in the meta-analysis was understood to have occurred at the preschool level. Hence, research conducted to test the effectiveness of MIT-based instruction in preschool education is recommended.

Another variable that was compared in terms of effect size was course type. In this research, the overall effect size was analyzed in the context of a total of four distinct course types: social science, science, foreign language, and tech courses. As a result of analysis, the largest effect size in the research was determined to belong to foreign language courses ($d = 1.958$), while the smallest effect size belonged to science courses ($d = 0.863$). According to the classification made by Cohen (1992), the overall effect size related to all course types was found to be large in the research. When looking at the effect sizes according to different course types, the effect of MIT-based instruction on students’ academic achievement was not seen to differ, $Q_B (3) = 7.085$. According to this result, the finding obtained in the research is suggested as generalizable to all course types. At this point, instruction based on multiple intelligences theory can be used in different course types at school. However, as the largest effect size in terms of course types is understood to belong to foreign language courses and the smallest effect size to science courses, a high-rise of effect size in foreign language courses may be due to the variety of rich learning methods used in these courses that adopt MIT-based instruction whereas a more abstract quality of science courses may be the reason for its lower level of effect size compared to foreign language courses. In a study conducted by Baş (2014), MIT-based instruction was concluded to offer various methods and techniques for student learning, so that students are constantly using different modalities of intelligence during the teaching and learning process in class.

According to the analysis conducted in terms of being a Master’s thesis or doctoral dissertation, the largest effect size was seen to belong to doctoral dissertations ($d = 0.984$), while the smallest effect size was seen to belong to master’s theses ($d = 0.937$). According to the classification made by Cohen (1992), the overall effect size regarding type of research was found to be large in the study. When looking at the effect sizes of studies included in the meta-analysis grouped according to being a master’s thesis or doctoral dissertation, the effect of MIT-based instruction on students’
academic achievement was not seen to differ, \( Q_a (1) = 0.001 \). Accordingly, the finding obtained in the research can be suggested as generalizable to all types of theses and dissertations. On the other hand, most of the studies that met the inclusion criteria for the meta-analysis were seen to be master’s theses \((n = 64)\), whereas few of them were doctoral dissertations \((n = 11)\). Although no significant difference between master’s theses and doctoral dissertations was detected, it was considered quite significant that the highest effect size belonged to doctoral dissertations. In fact, doctoral dissertations are generally known to be conducted throughout considerably larger areas, whereas master’s theses are carried out in relatively smaller circles. Of course, although there is no general judgment that doctoral dissertations are applied longer than master’s theses, the duration of the experimental procedures in doctoral dissertations are prepared more professionally. Also, it is quite thought-provoking that MIT-based instruction, which has inspired the curriculum development process in Turkey and then was reflected into the elementary curriculum especially, has a limited presence in master’s theses and even less so in doctoral dissertations. This may be due to the acceptance rate of students to doctoral and master’s programs in education and social sciences graduate schools. While master’s programs are known to accept more students than doctorate programs in Turkey, this result may affect the quantity of master’s theses and doctoral dissertations completed at the end of the programs in education and social sciences graduate schools. Though many studies have focused on MIT-based instruction at the master’s level, more research needs to be conducted at the doctorate level regarding this method of instruction. By studying MIT-based instruction at the doctorate level more, it is believed that the differences between postgraduate theses and doctoral dissertations can be discussed more healthily.

Another variable that was also compared in regard to effect sizes was the geographical region where studies were conducted. The purpose of such a comparison was to see whether significant differences in terms of geographical region had occurred or not to better analyze the reasons for the obtained results. Therefore, the overall effect size was analyzed in the context of seven geographical regions in Turkey: central Anatolia, eastern Anatolia, Aegean, Marmara, Mediterranean, Black Sea, and southeastern Anatolia. As a result of the analysis, the largest effect size was concluded to belong to southeastern Anatolia \((d = 1.423)\), while the smallest effect size belonged to the Marmara region \((d = 0.774)\). As there were only two studies conducted in southeastern Anatolia on the effect of MIT-based instruction on students’ academic achievement, generalizing the effect size in this region is probably not appropriate. According to the classification suggested by Cohen (1992), the effect sizes for all geographical regions except Marmara were found to be large, whereas the effect size for Marmara was seen to be at a medium level. When looking at the effect sizes of the studies grouped according to geographical region, the effect of MIT-based instruction on students’ academic achievement was seen to not differ in terms of geographical
region, $Q_B (6) = 2.570$. According to this result, the findings are suggested to be generalizable to all geographical regions. At the same time, this finding indicates that MIT-based instruction has similar results in different geographical regions of Turkey. When looking at the numerical distribution of the experimental studies using MIT-based instruction in terms of geographical region, many studies were seen to be conducted in all geographical regions of Turkey except for southeastern Anatolia ($n = 2$); this is considered too few compared to all other regions of the country. There may be various factors affecting this result. However, as Hedge’s $g$ value, which is used in calculating effect size, can give healthy and reliable results when at least five studies are taken into account (Rosenberg, Adams, & Gurevitch, 2000), the limited number of the studies conducted in the context of the southeastern Anatolia restricts further discussions and interpretations made in relation to the effectiveness of MIT-based instruction. Conducting more research in the context of southeastern Anatolia is considered necessary to better understand the underlying reason for the result obtained in the current meta-analysis.

As a result of the analysis conducted regarding the variable of sample size, the largest effect size was concluded to belong to the sample sizes of 1-20 ($d = 1.100$), whereas the smallest effect size belonged to sample sizes of 41 or more ($d = 0.738$). Moreover, according to the classification made by Cohen (1992), the effect size regarding sample sizes of 1-20 was found to be large, while the effect size in terms of the sample sizes of 41 or more was found to be at a medium level. When looking at the effect sizes of studies included in the meta-analysis grouped according to sample size, the effect of MIT-based instruction on students’ academic achievement was seen to not differ, $Q_B (2) = 2.800$. According to this result, this research finding is suggested to be generalizable to all sample sizes. In other words, the effect of MIT-based instruction on students’ academic achievement cannot be claimed to change in terms of sample size. At the same time, the finding obtained in the research indicates that instruction based on multiple intelligences theory has similar results for different sample sizes. Although no significant difference between different sample sizes was seen in terms of the effect of MIT-based instruction on students’ academic achievement, sample sizes of 1-20 were determined to have a larger effect size compared to sample sizes of 41 or more, which have a medium effect size in the study. According to this result, MIT-based instruction is seen to have a significantly larger effect size in a sample size of 1-20 compared to other sample sizes. Therefore, this result may indicate that MIT-based instruction is more effective in classes where there are fewer students, or rather small populations. As the relevant literature was reviewed, a positive correlation was seen between classrooms with small student populations and students’ academic achievement levels (e.g., Ferguson, 1991; Finn, Gerber, Achilles, & Boyd-Zahiras, 2001; Öğülmüş & Özdemir, 1995). Thus, by taking the sample sizes of classrooms into account, it seems possible to think that the
academic achievement levels of students is higher in classrooms with smaller sample sizes than classrooms with larger ones. Furthermore, additional research can also be performed as needed to better discuss and interpret this result more deeply.

Lastly, according to the results obtained from the analysis carried out in terms of the duration of experimental procedures, the largest effect size was concluded to belong to experimental durations of 6-10 weeks ($d = 0.994$), whereas the smallest effect size belonged to experimental durations of 11+ weeks ($d = 0.850$). According to the classification suggested by Cohen (1992), all of the studies experimental durations included in this meta-analysis were found to be large. When looking at the studies’ effect sizes in the meta-analysis grouped according to experimental duration, the effect of MIT-based instruction on students’ academic achievement did not differ in terms of duration, $Q_p (2) = 3.122$. According to this result, this finding may be suggested as generalizable to all durations of experimental procedure. In other words, the effect of MIT-based instruction on students’ academic achievement can be said to not change in terms of the experimental duration. At the same time, the finding obtained in the research indicates that MIT-based instruction has similar results in different durations of experimental procedure. However, even though no significant difference was seen among the different durations of experimental procedure in terms of the effect of MIT-based instruction on students’ academic achievement, the experimental duration of 6-10 weeks was observed to have a larger effect size compared to the others (i.e., 1-5 weeks of time, 11+ weeks of time). In this regard, many factors are thought to have a possible impact on this result. In the research, this result can be interpreted in many ways. Therefore, further research carried out on different durations of experimental procedure simultaneously should be conducted to discuss and interpret this result more deeply.

In this research, studies in terms of being a master’s thesis or doctoral dissertation that had considered the effect of MIT-based instruction on students’ academic achievement in Turkey were combined through the meta-analysis model. The results obtained from the studies included in the meta-analysis should be interpreted by taking some of the limitations inherent in experimental research into account (Kablan et al., 2013). Many variables in experimental studies can affect the dependent variables (Büyüköztürk, 2011). Although it is possible to prevent or control extraneous variables apart from the independent variables that aim to assess their effect on the dependent variable, many factors that the research cannot prevent or control directly may affect the dependent variable, so it cannot be known whether the obtained result is actually caused by the independent variable or not. Also, many factors can affect the dependent variable, such as whether the subjects or participants are familiar with the conditions of the experimental procedure or whether they know or have learned that they are part of an experiment, as well as reasons deriving from the pre-tests or post-tests, the conditions
in which the experimental procedure is carried out, the conditions of the distribution of subjects or participants into experimental and control groups, the behaviors or attitudes of the implementer of the experimental procedure, and more (Büyüköztürk, Çakmak-Kılıç, Akgün, Karadeniz, & Demirel, 2012; Fraenkel & Wallen, 2009). Thus, the expressed issues, which may affect students in both experimental and control groups depending on the circumstances, can result in a different situation that is not caused by the independent variable. This can therefore lead to inaccurate data obtained from experimental studies. The results of experimental research without having the necessary validity and reliability studies conducted can negatively reflect on the findings obtained from a meta-analysis. Accordingly, one should be cautious about the interpretation of the results of a meta-analysis, and the issues arising from the nature of experimental studies should be taken into full consideration. Indeed, as Gözüyeşil and Dikici (2013) have stated, some doubts may occur about the relevance of data arising from differences in research setting, time, and practices in the social sciences. Hence, it is quite important to make more flexible interpretations rather than precise and accepted interpretations based on the findings obtained from social sciences research. In this sense, it can be implied that the obtained results of research should be interpreted by taking the limitations and features of the actual implementation conditions of experimental studies into account (Kablan et al., 2013).

In this meta-analytic research, the effect of MIT-based instruction on students’ academic achievement was examined, and other variables on which MIT-based instruction had an effect were excluded from the study. Therefore, further meta-analytic studies can be conducted regarding the effect of MIT-based instruction on various variables such as motivation, attitudes towards course, locus of control, and so on. In this research, while the studies that had examined the effect of MIT-based instruction on students’ academic achievement in terms of postgraduate theses and dissertations carried out only in Turkey were combined through the meta-analysis, studies in terms of postgraduate theses and dissertations conducted on this issue outside of Turkey were not included. In this sense, it may be important to compare postgraduate theses and dissertations that examine the effectiveness of MIT-based instruction on students’ academic achievement both from Turkey and the world in a future meta-analytic study. In addition, while collecting the data to be used in this meta-analysis study from the Dissertations Center of the HEC, some studies were not accessible, so these could not be included in the current meta-analysis. In this context, that HEC should start the relevant work in relation with inaccessible postgraduate theses and dissertations and make these studies available for all researchers is seen as quite important. Such work is believed to increase future meta-analytic studies quickly, and more concrete and then reliable data obtained as a result of meta-analytic studies can occur. On the other hand, both the central examination system held by the Ministry of National Education (MoNE) and the examination system held in schools by teachers should be questioned critically.
Although the MoNE is stated to have been applying constructivist approaches in the teaching and learning process, the reflection of these approaches cannot be seen in the examination system itself. When both the central examination system held by MoNE and the examination system held in schools by teachers are examined carefully, the examination systems are seen to be far beyond the adopted constructivist approaches in the teaching and learning process. While MoNE is said to support constructivist approaches, such as the MIT-based instruction, it still continues to assess students through multiple-choice tests based mainly on verbal-linguistic and scientific-arithmetic courses without taking students’ abilities (such as spatial, kinesthetic, musical, etc.) into account. Also, while MoNE makes teachers try to evaluate students in alternative ways, it still assesses students through two specific intelligence types, such as IQ. Therefore, MoNE should take consistent steps in evaluating students through central and school-based examinations, as well as consider students’ different intelligence modalities in evaluating them on schooling. It was also understood that academic achievement tests of the studies included in the current meta-analysis were developed by the researchers themselves, thus these tests were found to have no standard. According to Gözüyeşil and Dikici (2013), standard academic achievement tests are generally used in research conducted abroad. In this context, it is believed that unity amongst researchers can be sustained by making standard academic achievement tests available nationwide; therefore, it is thought that more reliable results from meta-analytic studies can be obtained through the use of standardized tests by researchers. In this research, while only studies assessing the effectiveness of MIT-based instruction on students’ academic achievement were combined through the meta-analytic model, future meta-analytic studies can be conducted to assess the effectiveness of various methods of instruction (i.e., cooperative learning, project-based learning, active learning, differentiated instruction, constructivist learning, etc.) on students’ academic achievement levels.

References

Studies marked with asterisk (*) were included in the meta-analysis.


---

**Appendix 1**

*Studies Included in the Meta-Analysis*

Retrieved from http://tez2.yok.gov.tr, the website of The Dissertations Centre of the Higher Education Council (HEC) of Turkey.