

## Effect Size Reporting in Educational Research \*

Sibel ÖZSOY\*\*

Gökhan ÖZSOY\*\*\*

**ABSTRACT.** The purpose of this study is to investigate effect size reporting and interpreting practices in Turkish academic journals in the field of education which are indexed in Social Sciences Citation Index. For this purpose, 990 articles published in four journals between the years 2007-2011 were selected for investigation. Among these articles, 480 of them were included in the study in accordance with reportable effect sizes. The results of the study showed that only in 35 of the articles (i. e. 7.2%) researchers reported the effect size values obtained. Besides, researchers interpreted the effect size only in 21 of them. The results of the study also revealed that there are no significant differences among journals with respect to effect size reporting and interpreting  $\chi^2(3, n = 483) = 7.30, p = .06$ , Cramer's  $V = .12$ .

**Key Words:** Education, research, effect size, practical significance, statistical significance.

### SUMMARY

Effect size calculations have become important criteria for reporting of research, supporting null hypothesis significance testing and obtaining higher quality statistical results. According to APA (American Association of Psychology) (2001), the effect sizes calculated (Cohen's  $d$ , Hedge's  $g$ ,  $\eta^2$ , vb.) should always be reported together with  $p$  significance values. Reporting the effect sizes calculated in a study in comparison with the effect sizes previously obtained is the primary prerequisite of a good study report. It is generally observed in research published in Turkey that the effect size has not been reported, which makes conducting comparative studies like meta-analysis difficult. The aim of the current study is to investigate the state of the effect size reporting in articles published in journals issued in Turkey and indexed by SSCI (Social Science Citation Index®).

**Method:** A total of 990 articles published in four Turkish journals indexed in SSCI during the period of 2007 and 2011 were examined within the context of the present study. In order to examine the articles and to carry out a systematic coding procedure, the items in the checklist developed by Alhija and Levy (2007) were used after adapting them into the present study. Articles in each journal were evaluated in relation to each item in the checklist by considering the major statistical methods.

**Results:** Out of the 990 articles examined in the present study, 783 of them are quantitative (79.4%), 148 of them are qualitative (14.9%) and 56 of them are theoretical (5.7%). 480 quantitative articles where effect sizes should be reported were analyzed in the present study. When the effect size reporting was analyzed, it was found that only in 35 of these studies (7.2%) proper effect size reporting was observed, and only in 21 of them reported effect size was interpreted. When the distribution of effect size reporting according to the journals, it was found that throughout the five-year period, effect size was reported in 14 articles published in A, 11 articles published in B, 7 articles published in C and 3 articles published in D. Moreover, chi-square results revealed that there is no statistically significant relationship among chi-square reporting frequencies among the journals  $\chi^2(3, n = 483) = 7.30, p = .06$ , Cramer's  $V = .12$ . When the effect size reporting was investigated according to years, it was found that there are three articles reporting effect size for years 2007 and 2008 each, there are six in 2009, and 14 for 2010 and 9 for 2011. Chi-square analysis was carried out to determine whether there is a statistically significant difference in effect size reporting according to years and no significant difference was found  $\chi^2(4, n = 483) = 6.77, p = .15$ , Cramer's  $V = .12$ .

**Discussion:** In the former studies conducted on the reporting of effect size, reporting frequency rate ranging from %1 to 90% was found (Meline & Schmitt, 1997; Thompson, 1999b). In the present study, effect size reporting frequency rate was found to be 7.2%; hence, it may be concluded that in studies conducted in our country, effect size reporting rate is quite low. When the fact that the studies analyzed in the present study are indexed in SSCI is considered, it is obvious that the effect size reporting rate is highly inadequate. Moreover, it is estimated that in articles published in the journals not indexed in field indexes, the effect size reporting rate is much lower. A researcher needs to understand the findings of his/her study and report them in the most appropriate and correct manner. Hence, a researcher should be able to report and interpret both the results of statistical significance tests and results of effect size calculations.

\* This study is presented as an oral presentation at 5th World Conference on Educational Sciences (Rome, Italy, 2013).

\*\* Assist. Prof. Dr., Aksaray University Faculty of Education, Aksaray, Turkey. E-mail: [sozsoy@aksaray.edu.tr](mailto:sozsoy@aksaray.edu.tr)

\*\*\* Assist. Prof. Dr., Aksaray University Faculty of Education, Aksaray, Turkey. E-mail: [gozsoy@gmail.com](mailto:gozsoy@gmail.com)

# Eğitim Araştırmalarında Etki Büyüklüğü Raporlanması\*

Sibel ÖZSOY\*\*

Gökhan ÖZSOY\*\*\*

**ÖZ.** Bu araştırmanın amacı, eğitim alanında Türkiye kaynaklı yayımlanan ve SSCI’da taranan akademik bilimsel dergilerde yayımlanan makalelerde etki büyüklüğü raporlanması durumunu araştırmak ve okuyuculara etki büyüklüğü, hesaplanması ve raporlanması konularında bilgi sunmaktır. Araştırma kapsamına alınan dört dergide 2007-2011 yılları arasında yayımlanan toplam 990 makale incelenmiş; bunlardan etki büyüklüğü raporlanmaya uygun olan 480 tanesi araştırma kapsamına alınmıştır. Araştırma sonunda 480 makalenin sadece 35’inde (%7.2) etki büyüklüğü raporlandığı; bunların da 21’inde raporlanan etki büyüklüklerinin aynı zamanda yorumlandığı görülmüştür. Çalışmadan elde edilen sonuçlar incelendiğinde dergiler arasında etki büyüklüğü raporlama ve yorumlama açısından anlamlı bir farklılık olmadığı görülmüştür  $\chi^2(3, n = 483) = 7.30, p = .06$ , Cramer’s V = .12.

**Anahtar kelimeler:** Eğitim, araştırma, etki büyüklüğü, pratik anlamlılık, istatistiksel anlamlılık.

## GİRİŞ

Varoluşundan beri insanoğlu, yaşadığı ortamda olan biteni anlama ve açıklama çabası içerisinde olmuş, bu çaba sonucu elde ettiği bilgi birikimiyle yaşadığı çevreyi denetlemeye ve kendine uygun olarak değiştirmeye çalışmıştır. İnsanoğlunun, “bir alandaki varlıkları ve olayları inceleme, açıklama, onlara ilişkin genelleme ve ilkeler bulma, bu ilkeler yardımıyla gelecekteki olayları kestirme gayreti” bilim olarak (İşman, vd. 2002, s.41), bu süreçte izlediği yol da bilimsel araştırma yöntemi olarak tanımlanabilir. Bilimsel araştırma ile bilim insanları bilimsel yöntemleri kullanarak belli bir sorunun çözümüne yanıt ararlar. Bu sürecin temelini merak ve gözlem oluşturmaktadır. Bilimsel araştırma süreci, sosyal bilimler ve fen bilimleri arasında uygulamada bazı farklılıklar olsa da, temelde, problemin tanımlanması ile başlar, çeşitli yöntemleri kullanarak veri toplama ve verilerin analizi ile devam eder, elde edilen sonuçların raporlanması ile sonuçlanır.

Bilimsel araştırma sonucunda elde edilen bilgilerin doğru ve geçerli olması çok önemlidir. Doğru ve geçerli bilgi elde edebilmek için, bilimsel araştırmanın planlanmasından, uygulanmasına ve raporlanmasına kadar gerçekleştirilen tüm aşamalar araştırmacılar tarafından büyük bir titizlikle yürütülmelidir. Bu aşamaların herhangi birinde yapılacak olan hata, araştırmada elde edilen bilgilerin doğruluğunu ve geçerliliğini tehlikeye atacaktır. Bir araştırmaya başlarken ilk temel soru araştırma konusunu belirlemektir. Araştırmacı, araştırma konusunu belirlerken konunun ilgi çekici olmasına, araştırılabilir olmasına, bilime katkısı olmasına dikkat etmelidir (Polonsky ve Waller, 2005; Brewerten ve Millward, 2001; Daymon ve Holloway, 2002). Ayrıca araştırmacının, konu hakkında yeterli ön bilgisinin ve yöntem bilgisinin olması sağlıklı bir araştırma süreci için önem taşımaktadır.

Araştırmacı, araştırma konusunu belirledikten sonra çalışmanın odak noktasını belirleyecek olan araştırma problemini tanımlamalıdır. Araştırma probleminin belirlenmesi, araştırmanın sınırlarını doğru belirleyebilmek, ele alınacak hipotezleri doğru kurabilmek, uygun evren ve örnekleme seçebilmek, uygun ölçme araçları ile doğru analizleri yapabilmek için önemlidir. Bu nedenle, araştırma probleminin, araştırmacının kendisi için de yol gösterici bir öge olduğu unutulmamalı, araştırma problemi açık bir şekilde ifade edilmeli ve çalışmada ele alınan değişkenler arasındaki ilişkiyi yansıtmalıdır (Ghauri ve Grønhaug, 2003). Araştırma problemi oluşturulurken, problemin açık olmasına, çalışmaya değer ve

\* Bu çalışma, 2013 yılında Roma, İtalya’da düzenlenen “World Conference on Educational Sciences”ta sözlü bildiri olarak sunulmuştur.

\*\* Yrd. Doç. Dr., Aksaray Üniversitesi Eğitim Fakültesi, Aksaray Türkiye. E-posta: [sozsoy@aksaray.edu.tr](mailto:sozsoy@aksaray.edu.tr)

\*\*\* Yrd. Doç. Dr., Aksaray Üniversitesi Eğitim Fakültesi, Aksaray Türkiye. E-posta: [gozsoy@gmail.com](mailto:gozsoy@gmail.com)

önemli olmasına, zaman, çaba ve maliyet açısından araştırılabilir olmasına ve etik olmasına dikkat edilmelidir (Fraenkel ve Wallen, 2003).

Araştırma problemi de belirlendikten sonra yapılacak iş, probleme uygun hipotezlerin yazılması ve bu hipotezlerin uygun ölçme araçları ile test edilmesidir. Hipotez, bilimsel araştırmanın sonucunda elde edilmesi muhtemel sonuçlara yönelik bir tahmin ifadesi olarak tanımlanabilir (Frankel ve Wallen, 2003). Bir başka deyişle, hipotezler, doğruluğu bir araştırma ya da deney ile test edilen denenceler ya da öngörülerdir. Doğru yapılandırılmış bir hipotez, çalışmada kullanılacak istatistiksel analiz yöntemi, çalışmada ele alınan bağımlı ve bağımsız değişkenler konusunda yol göstericidir. Araştırmalarda, hipotezler doğrultusunda evrene ait parametrelere ulaşmak amaçlanır. Ancak araştırmacının her zaman evren üzerinde araştırma yapması mümkün olmayabilir. Bu nedenle araştırmalar, evrenden yansız olarak çekilen örneklem üzerinde yürütülür ve bu örneklemde elde edilen istatistiksel değerlerle evrene ait parametreler kestirilir. Bunu yapmadaki amaç, ulaşımı ve üzerinde çalışması daha kolay olan örneklemden yararlanarak, evren hakkında bir takım yorumlarda bulunmaktır. Yapılan bu tahmin ve kestirme işlemlerine hipotez testi ya da anlamlılık testi denmektedir. Hipotez testleri, örneklem ortalaması ile örneklemin alınmış olduğu evrenin ortalamaları arasında bir fark olmayacağı varsayımına dayanır. Bu anlamda oluşturulan bir yokluk (null, sıfır) hipotezinin bir anlamlılık/manidarlık testi ile yanlışlanıp yanlışlanmadığı incelenir (Balcı, 2009). Hipotez testi gözlemlenen örneklerin beklenen sonuçlardan farklı olup olmadığının belirlenmesine imkân verir ve böylelikle hipotezin reddine/kabulüne karar verilmesine yardım eder (Spiegel ve Stephans, 1999).

Bir hipotezin doğruluğunu test etmek için bir yokluk hipotezi ile bir alternatif hipotez oluşturulur. Yokluk hipotezi çalışmada ele alınan gruplar arasında istatistiksel olarak anlamlı bir farklılık olmadığını belirtir. Alternatif hipotez ise yokluk hipotezinin içermediği tüm durumları içerir. Örneğin, ilköğretim birinci kademe öğrencilerinin matematik dersine karşı tutumunu incelediğimiz bir çalışma düşünelim. Bu çalışmada ele alacağımız cinsiyet değişkenine göre kız ve erkek öğrenciler arasında matematik dersine karşı tutumları arasında anlamlı bir farklılık olup olmadığını incelemek için kuracağımız yokluk ve alternatif hipotezi şu şekilde olacaktır:

**H<sub>0</sub>:** İlköğretim birinci kademe öğrenim gören kız ve erkek öğrencilerin matematik dersine karşı tutum ortalamaları arasında anlamlı bir farklılık yoktur.

**H<sub>A</sub>:** İlköğretim birinci kademe öğrenim gören kız ve erkek öğrencilerin matematik dersine karşı tutum ortalamaları arasında anlamlı bir farklılık vardır.

H<sub>0</sub> ve H<sub>A</sub> 'yı sembolik olarak göstermek istersek aşağıdaki gösterimi kullanabiliriz:

$$\mathbf{H_0: \mu_k - \mu_e = 0}$$

$$\mathbf{H_A: \mu_k - \mu_e \neq 0}$$

Bundan sonra, yapılacak olan istatistiksel testlerden elde edilen sonuca göre yokluk hipotezini reddeder ya da kabul ederiz. İşte bu karar verme aşamasında, elde edilen veriler örneklemden alındığı için araştırmacının hatalı karar verme ihtimali vardır. Bir hipotez testi sonucunda karşımıza dört olası sonuç çıkar; yanlış bir hipotezi reddetme, doğru bir hipotezi kabul etme, doğru bir hipotezi reddetme ya da yanlış bir hipotezi kabul etme. İlk iki durum, araştırmacı için istendik durumlardır. Ancak son iki durumda yani doğru bir hipotezi reddetme ya da yanlış bir hipotezi kabul etme durumlarında, yanlış bir karar veya hatalı bir hüküm ortaya çıkmaktadır. Hipotez testi sonucunda kabul edilmesi gereken bir hipotez reddedildiğinde Tip I hata, reddedilmesi gereken bir hipotez kabul edildiğinde ise Tip II hata yapılmış olur. Bir araştırmada yapılabilecek Tip I hatanın maksimum olasılığı testin anlamlılık düzeyi olarak adlandırılır ve  $\alpha$  (alfa) ile gösterilir. Bu olasılık değeri araştırmacı tarafından araştırma öncesinde belirlenir. Sosyal bilimler araştırmalarında genelde .05 veya .01 anlamlılık düzeyi tercih edilir. Tercih edilen değerle, Tip I hata yapma olasılığı de belirlenmiş olur. Yani bir araştırmacı .05 değerini anlamlılık düzeyi olarak belirlemişse doğru olan bir hipotezi reddetme olasılığı %5 iken, .01 değerini anlamlılık düzeyi olarak belirlediyse bu olasılık %1'e düşmektedir. Araştırmada gerçekte elde edilen anlamlılık düzeyine olasılık değeri adı verilir ve  $p$  sembolü ile gösterilir (Balcı, 2009, Spiegel ve Stephans, 1999).

Elde edilen olasılık değerinin yani  $p$ 'nin, araştırma öncesinde belirlenen anlamlılık düzeyinden yani  $\alpha$ 'dan büyük ya da küçük olmasına göre istatistiksel çıkarım yapılır. Ancak araştırmacının hipotez testi konusunda duyması gereken kaygılar bu noktada bitmez. Araştırmacı Tip I hatanın olasılığını belirleyerek ya da bu olasılığı minimuma çekerek tüm sorunlardan kurtulamaz. Araştırmacının araştırmada belirlenen reddedilmesi gereken bir sıfır hipotezinin kabul edilmesi durumunda yapılacak olan Tip II hata da çözmesi gereken sorunlardan bir tanesidir. Araştırmacının bu türden bir hata yapma olasılığı  $\beta$  (beta) ile gösterilir. Örneklem hacmi sabit tutulduğunda, Tip I hata ne kadar küçültülürse Tip II hata da o kadar artacaktır.

Yokluk hipotezi anlamlılık testleri, araştırma problemleri ile ilgili istatistiksel çıkarım yapmak için 1900'lü yıllardan beri araştırmacılar tarafından kullanılan bir yaklaşımdır (Hubert ve Pike; 1999). Ancak ilgili literatürde bu yaklaşımın yeterliği ve kullanılabilirliği sorgulanır hale gelmiştir (Cohen, 1990; 1994; Kirk, 1996; Thompson, 1996; Yates, 1951; Yıldırım ve Yıldırım, 2011). Etki büyüklüğü (effect size) hesaplamaları, pratik anlamlılık için uzun zamandır bilimsel araştırmaların raporlanmasında yokluk hipotezi testlerinin desteklenmesi ve daha nitelikli istatistiksel sonuçlar elde edilmesi için önemli bir kriter durumuna gelmiştir (American Psychology Association [APA], 2001; American Educational Research Association [AERA], 2006; Sun, Pan, ve Wang, 2010). APA (Amerikan Psikoloji Derneği)'ne (2001) göre hesaplanan etki büyüklükleri (Cohen's  $d$ , Hedge's  $g$ ,  $\eta^2$ , vb.),  $p$  anlamlılık değerleri ile birlikte her zaman mutlaka raporlanması gerekir ve daha önceden elde edilen etki büyüklükleri ile mevcut araştırmadan elde edilen etki büyüklüklerini karşılaştırarak raporlamak, iyi bir araştırma raporunun öncelikli gerekliliğidir.

### **Etki Büyüklüğü Nedir?**

Etki büyüklüğü, örneklemde elde edilen sonuçların yokluk hipotezinde tanımlanan beklentilerden sapma düzeyini gösteren istatistiksel değerdir (Cohen, 1994; Vacha-Haasse ve Thompson, 2004). Etki büyüklüğü, genel olarak, yokluk hipotezleri ile alternatif hipotezler arasındaki farkın büyüklüğü olarak tanımlanmaktadır. Bu da, araştırma sonuçlarının pratikteki anlamlılığının bir göstergesi niteliğindedir.

Tüm araştırmalarda olduğu gibi eğitim alanında yapılan araştırmalarda da, elde edilen istatistiksel bulguların anlamlılığı en önemli hususlardan birisidir. Araştırmalarda anlamlılık konusunda iki temel yaklaşımdan istatistiksel anlamlılık, geniş bir şekilde bilinir ve uygulanır durumda olmasına rağmen pratik anlamlılık konusunda aynı şeyi söylemek mümkün değildir. Araştırmacılar, istatistiksel anlamlılık ölçütü olarak kullanılan  $p$  değerinin ne kadar küçük olursa, uygulamanın etkisinin ya da gücünün de aynı oranda büyük olacağı yönünde bir yanlış inanç vardır (Nickerson, 2000). Ancak istatistiksel anlamlılık testleri, örneklemde elde edilen sonucun şans faktörü ile elde edilme ihtimalini değerlendirirken; etki büyüklüğü pratik anlamlılığın bir göstergesidir. İstatistiksel anlamlılık, örneklem sayısından etkilenirken (Fan, 2001), etki büyüklüğü değeri, bu örneklem sayısından kaynaklanan sonuçları ortadan kaldırarak elde edilen sonuçlar hakkında daha doğru bir karar verilmesine yardımcı olur.

Çeşitli etki büyüklüğü tanımlarının ortak noktaları olarak “*ortalamlar arasındaki farkın standartlaştırılması*” ya da “*ilişkinin standartlaştırılmış ölçümü*” ifadeleri belirlemektedir. Pratikte, araştırmacıların iki kategoride etki büyüklüğü hesapladığı görülmektedir: Etki büyüklüğü ölçümleri (*grup ortalamaları farkına göre*) ve ilişki gücü ölçümleri (*hesaplanan varyansa göre*) (Kortlik ve Williams, 2003). Grup ortalamaları farkına göre hesaplanan etki büyüklüğü ölçümleri Cohen's  $d$  (Cohen, 1988), Glass's  $g$  (Glass, 1976) ve Hedge's  $d$  (Hedges, 1981) ile gösterilmektedir. Varyansa göre hesaplanan ilişki gücü ölçümleri ise bağımsız değişken ile bağımlı değişken arasındaki korelasyona göre hesaplanmakta, R-kare ( $R^2$ ) ve eta-kare ( $\eta^2$ ) ölçümleriyle ifade edilmektedir (Maxwell ve Delaney, 1990; Synder ve Lawson, 1993).

### **Etki Büyüklüğünün Hesaplanması, Raporlanması ve Yorumlanması**

Çalışmanın bu kısmında, çeşitli araştırma desenleri için kullanılacak etki büyüklüğü hesaplama yöntemleri açıklanmaktadır. Bu kısımda yer alan formüller, araştırmacılar için hesaplamalarının ve kullanımlarının kolaylığı dikkate alınarak seçilmiştir. Çoğu durumda verilen formüller, birçok istatistik

yazılımlarından (SPSS, vb.) elde edilen verilerin basitçe yerine koyulmasıyla etki büyüklüğü hesaplamalarını sağlayabilmektedir.

#### *Grup ortalamaları farkına göre hesaplamalar*

İki grup ortalaması arasındaki farkın hesaplandığı istatistiksel yöntemler (tek grup t-test, ilişkili örneklem için t-testi, ilişkisiz örneklem için t-test, vb.) için etki büyüklüğü hesaplanmasında Cohen's *d* formülü (Cohen, 1988) yaygın biçimde tercih edilmektedir. Cohen's *d* formülü ile hesaplama yapabilmek için grupların ortalamalarına ve harmanlanmış standart sapma (*pooled standart deviation*) bilgilerine ihtiyaç vardır. Harmanlanmış standart sapma ve Cohen's *d* hesaplanması aşağıda gösterilmiştir:

$$\text{Harmanlanmış standart sapma} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2}}$$

$$\text{Cohen's } d = \frac{\text{Örneklem ortalamaları farkı}}{\text{Harmanlanmış standart sapma}}$$

Hesaplamalar sonucunda elde edilen *d* değeri şu şekilde yorumlanır: .20- küçük (*small*) etki büyüklüğü; .50- orta (*medium*); .80 ise büyük (*large*) etki büyüklüğü (Cohen, 1988). Bununla birlikte Cohen's *d* değeri aynı zamanda örneklem büyüklüğünün kestirilmesi amacıyla da kullanılmaktadır. Cohen's *d* formülü ile elde edilen küçük etki büyüklükleri, daha büyük örneklem gerekliliğini gösterir.

Grup ortalamaları farkını temel alan istatistiksel yöntemlerde, etki büyüklüğü hesaplamaları için Cohen's *d* formülünün yanı sıra Hedges's *g* ve Glass's  $\Delta$  hesaplamaları da yapılabilmektedir. Hedges's *g* hesaplanması *d* ile oldukça benzer olmakla birlikte sadece harmanlanmış standart sapma hesaplamasında farklılık vardır. *g* hesaplanırken:

$$\text{Hedge's } g = \frac{\text{Grup ortalamaları farkı}}{\text{Harmanlanmış standart sapma}}$$

$$\text{Harmanlanmış standart sapma} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}}$$

#### *Varyansa göre etki büyüklüğü hesaplanması*

Varyans analizlerinde etki büyüklüğü hesaplanması için birçok yöntem kullanılmaktadır. Cohen's *f* ve Omega kare ( $\Omega^2$ ) varyans analizinde kullanılan etki büyüklüğü hesaplamaları arasında yer almaktadır. Her iki ölçüm de kategorik değişken tarafından açıklanan varyans oranı hakkında bir tahmin değeri verir. Cohen's *f* değeri, örneklem tarafından hesaplanan varyans oranını tahmin ederken, Omega kare ise evren tarafından oları tahmin eder.

Cohen's *f* hesaplamak için öncelikle eta kare ( $\eta^2$ ) hesaplanması aşağıdaki formüle göre yapılmalıdır:

$$\eta^2 = \frac{\text{Kareler toplamı (gruplar arası)}}{\text{Kareler toplamı (toplam)}}$$

Daha sonra ise Cohen's *f* değeri:

$$\text{Cohen's } f = \sqrt{\frac{\eta^2}{1 - \eta^2}}$$

formülü ile hesaplanır.

Evren bakımından varyans oranı tahmini yapan Omega kare ( $\Omega^2$ ) değeri ise ( $k$  = grup sayısı olmak üzere) şöyle hesaplanır:

$$\Omega^2 = \frac{\text{Kareler toplamı (gruplar arası)} - (k - 1)\text{Kareler ortalaması (grup içi)}}{\text{Kareler toplamı (toplam)} + \text{Kareler ortalaması (grup içi)}}$$

Cohen's  $f$  ya da Omega kare hesaplandıktan sonra yorumlanırken, Cohen's  $f$  için .10 = küçük , .25 orta, .40 geniş (Cohen, 1988) ; Omega kare için .010 küçük, .059 orta, .138 geniş kabul edilir (Kirk, 1996).

#### *Korelasyon ve regresyon hesaplamalarında etki büyüklüğü*

Etki büyüklüğü ölçümleri göz önünde bulundurulduğunda, en basit hesaplama, korelasyon istatistikleri için yapılmaktadır. Esasında korelasyon katsayısının kendisi, bir etki büyüklüğü değeridir (Kotrlık ve Williams, 2003). Korelasyon katsayıları ( $r$ ) için etki büyüklüklerinin yorumlanmasında, Davis'e (1971) göre .01 ile .09 arası ihmal edilebilir ilişki; .10 ile .29 arası düşük ilişki; .30 ile .49 arası orta; .50 ile .69 arası güçlü; .70 ve sonrası ise çok güçlü ilişki olarak yorumlanmakta; Hinkle, Wiersman ve Jurs'a göre (1979) .00 ile .30 arası çok düşük; .30 ile .50 arası düşük; .50 ile .70 arası orta; .70 ile .90 arası yüksek; .90 ile 1.00 arası ise çok yüksek olarak yorumlanmaktadır. Hopkins'e (1997) göre ise .00-.10 arası göz ardı edilebilir; .10-.30 arası küçük; .30-.50 arası orta; .50-.70 arası yüksek; .70-.90 arası çok yüksek; .90-1.00 arası ise mükemmel ilişki olarak yorumlanmaktadır.

Etki büyüklüğü hesaplamaları için bir başka basit ölçüm ise çoklu regresyon katsayısı olan  $R^2$  dir. Bütün temel istatistik yazılımları,  $R^2$  değerini otomatik olarak hesaplamaktadır. Cohen'e (1988) göre elde edilen bu etki büyüklüğü sonuçları ( $R^2$ ): .0196 küçük; .1300 orta; .2600 ise büyük etki değeri olarak yorumlanabilmektedir.

Araştırmacı, çalışmasından elde ettiği verileri kullanarak etki büyüklüğünü hesaplayabileceği gibi, etki büyüklüğünü hesaplamak için tasarlanmış hesap makinelerini de kullanabilir. İnternet'te konu ile ilgili yapılacak basit bir tarama ile birçok hesap makinesi içeren web sayfası bulmak mümkün olacaktır (örn. [www.uccs.edu/~faculty/lbecker/](http://www.uccs.edu/~faculty/lbecker/) ).

Etki büyüklüklerinin yorumlanmasında genellikle kullanılan yöntem etki büyüklüklerini “küçük”, “orta” ve “büyük” olmak üzere ayırmak yönündedir (Cohen, 1988). Ancak Cohen'in oluşturduğu bu göstergelerin genelde tek başına faydalı olmadığı belirtilmektedir (Thompson, 2008). Etki büyüklüğünün raporlanmasında araştırmacıların hem elde ettikleri etki büyüklüğünü yorumlamaları hem de bu değeri daha önceden aynı konuda yapılmış çalışmalarda elde edilen değerlerle karşılaştırmaları (Henson, 2006; Kline, 2004; Thompson, 2008) gerekmektedir. Etki büyüklüğü raporlama ve yorumlama, akademik çalışmalarda henüz bir alışkanlık haline gelmediği için, araştırmacının elde ettiği etki büyüklüğünü diğer çalışmalardan elde edilenlerle karşılaştırması zor olabilir. Ancak bu durumda bile araştırmacı, önceki çalışmalarda etki büyüklüğü hesaplaması için gereken verileri çıkararak, kendisi o çalışmalar için bir hesaplama yapmakla ve çalışmasından elde ettiği etki büyüklüğü değerlerini bu hesaplamalarıyla karşılaştırmakla yükümlüdür.

Etki büyüklüğü değerleri, araştırmalarda genel olarak üç amaçla kullanılabilir. Bunlardan birincisi, çalışmanın uygulaması öncesinde beklenen (umulan) etki büyüklüklerinin tahmin edilmesi örneklem sayısını belirlemede yararlı olacaktır. Önceden belirlenecek etki büyüklüğü, alfa değeri ve kuvvet (power) değeri ile hesaplanacak olan minimum örneklem sayısı, örneklem sayısının yetersizliğinden kaynaklanacak istatistiksel açıdan anlamlı olmayan değerlerin elde edilmesini engelleyecektir (Olejnik, 1984; Plucker, 1997). İkinci olarak etki büyüklüğü hesaplamaları diğer araştırmacıların uygulamanın gerçek değeri/etkisi hakkında fikir sahibi olmalarını sağlayacaktır. Üçüncü olarak, etki büyüklükleri, ortalamalardan ya da ilişkilerden etkilenmediği için, çalışmaların sonuçlarını nicel olarak karşılaştırabilmek için gereklidir. Çalışmadan elde edilen bir  $p$  değeri, uygulamanın etkililiği ile ilgili yeterli ve güvenilir bir bilgi veremez (Falk ve Greenbaum, 1995; Thompson, 1993; Thompson, 1999a).

Bu nedenlerden dolayı, yapılan bir hipotez testini etki büyüklüğü hesaplamaları takip etmelidir (Plucker, 1997; Robinson ve Levin 1997; Thompson ve Snyder, 1997). Ancak bu durumda elde edilen sonuçların pratikteki anlamlılığı konusunda bir şey söylemek mümkün olacaktır. Etki büyüklüğü, geleneksel anlamlılık testlerinden farklı bir ölçümdür. Çünkü etki büyüklüğü, sonuçların anlamlılığından bahseder ve çalışmalar arasında kolaylıkla karşılaştırma yapabilmemizi sağlar. Ayrıca araştırmacıların elde ettikleri sonuçların pratik anlamlılığını da sorgulamalarını sağlar. Bu yüzden yapılan çalışmaların tümünde etki büyüklüğünü gösteren herhangi bir yöntem raporlanmalıdır ki okuyucular, çalışmadan elde edilen sonuçların önemini daha iyi kavrayabilsinler ve çalışmalar arasında karşılaştırma yapabilsinler (Cohen, 1990; 1994; Thompson, 2002).

Bununla birlikte, APA tarafından, araştırmalarda *p* anlamlılık değeri ile birlikte mutlaka etki büyüklüğü değerinin de hesaplanarak raporlanması gerektiği belirtilmektedir; etki büyüklüğü raporlamamak, bir çalışma için kusur olarak görülmektedir. 2006 yılında AERA (Amerikan Eğitim Araştırmaları Birliği) tarafından yayımlanan araştırma standartları raporunda sosyal bilimlerdeki araştırmalar için standart hata, etki büyüklüğü, güven aralığı raporlanmasının önemli gerekliliklerden birisi olduğu belirtilmiştir (AERA, 2006). APA ve AERA'nın bilimsel çalışmalarda etki büyüklüğünün raporlanmasına yönelik yapmış olduğu düzenlemelerle bu derneklerin kendi yayınladığı akademik dergiler ile oluşturmuş oldukları yazım kurallarını şart koşan diğer dergiler de yayımlanacak bilimsel çalışmalarda etki büyüklüğünün raporlanmasını zorunlu hale getirmişlerdir.

Uluslararası literatürde, 1960'lı yıllardan beri akademik dergilerde yayınlanan makalelerde etki büyüklüğü raporlama sıklığını belirleme üzerine çalışmalar yapılmaktadır (Alhija ve Levy, 2007; Dunleavy, Barr, Glenn ve Miller, 2006; Ives, 2003; Kirk, 1996; Meline ve Schmitt, 1997; Meline ve Wang, 2004; Ottenbacher ve Barrett, 1989; Paul ve Plucker, 2003; Plucker, 1997; Snyder ve Thompson, 1998; Thompson ve Snyder, 1997; Vacha-Haase ve Ness, 1999). Türkiye'de yayınlanan araştırmalarda genellikle etki büyüklüklerinin raporlanmadığı gözlenmektedir, bu durum da meta-analiz gibi karşılaştırmalı çalışmaları imkânsız hale getirmektedir. Gözlenen bu eksiklik üzerine planlanan bu araştırmanın öncelikli amacı Türkiye'de yayınlanan ve SSCI (Social Science Citation Index®) tarafından indekslenen dergilerde yer alan makalelerdeki etki büyüklüklerinin raporlanma durumlarını incelemektir.

## YÖNTEM

Çalışmada incelenen dergiler amaçlı örnekleme yoluyla seçilmiştir. Amaçlı örneklemede çalışmanın yürütülmesi için gereken bilgi açısından zengin durumlar daha derin bir araştırma yapılabilmesi için seçilir, araştırma kapsamında yer alacak örneklem sayısı ve özel durumlar çalışmanın amacına uygun olarak belirlenir (Patton, 1999). Çalışmada yer alacak dergileri belirlemek için öncelikle Türkiye'de eğitim alanında yayınlanan dergilerin bir listesi çıkarılmış, daha sonra bu dergilerden SSCI'da tarananlar belirlenmiştir. Bunun sonucunda Türkiye merkezli olarak yayımlanan toplam 6 adet derginin SSCI'da tarandığı görülmüştür. Ayrıca dergilerin tam metinlerine ulaşılabilir olması da gözetilmiş, bu dergilerden birinde yayımlanan makalelerin tam metinlerine elektronik ya da basılı olarak ulaşamadığı için bu dergi çalışma kapsamı dışında bırakılmıştır. Öte yandan bu dergilerden bir diğeri de eğitim bilimlerinin özel bir alt alanındaki çalışmaları kapsadığı için çalışma dışında tutulmuş, böylelikle toplam dört dergi ve bunlarda yayımlanan toplam 990 makale araştırma kapsamına alınmıştır. Dergilerde 2007-2011 yılları arasında yayımlanan makaleler incelenmiştir.

### *İnceleme ve Kodlama İşlemleri*

Araştırma kapsamına alınan makaleleri incelemek ve kodlama işlemlerini sistematik olarak yürütebilmek amacıyla, Alhija ve Levy (2007) tarafından geliştirilen kontrol listesinde yer alan maddeler, çalışmaya adapte edildikten sonra kullanılmıştır. Buna göre, makaleleri incelemek için aşağıda verilen maddeler kullanılmıştır:

1. Çalışmanın Başlığı
2. Çalışmanın Yayınladığı Yıl
3. Çalışmanın Basıldığı Dergi
4. Çalışmanın temel alanı nedir?
5. Araştırmanın temel problemi nitel mi nicel mi?

6. Çalışmada yer alan temel analiz nedir?
7. Çalışmadan elde edilen sonuç nasıl raporlanmıştır? İstatistiksel olarak anlamlı/istatistiksel olarak anlamlı değil/karışık
8. Elde edilen istatistiksel sonuç yorumlanmış mı?
9. Etki büyüklüğü raporlanmış mı?
10. Hangi etki büyüklüğü raporlanmış?
11. Raporlanan etki büyüklüğü yorumlanmış mı?

Her bir dergide bulunan makaleler kullanılan temel istatistiksel yöntemleri göz önünde bulundurularak kontrol listesinde bulunan her bir madde açısından değerlendirilmiştir. Çalışmanın temel alanı maddesi, etki büyüklüğü raporlamasının herhangi bir alanda yoğunlaşmış ya da yoğunlaşmadığını gözlemlemek için kontrol listesine alınmıştır. Araştırma kapsamında yer alan makaleler incelendiğinde bazı makalelerin birden çok araştırma problemine cevap aradığı gözlenmiştir. Araştırma kapsamında makalelerde yer alan ana problem incelenmiş, makalede yer alan temel istatistiksel yöntem, araştırma problemini cevaplamak için seçilen istatistiksel yöntem olarak belirlenmiştir. Araştırma kapsamında yer alan makalelerde kullanılan istatistiksel yöntemler temelde üçe ayrılmıştır:

1. Basit Testler (*simple tests*): *t*-testi, basit korelasyon, ki-kare, Mann-Whitney U testi
2. Genel doğrusal modeller (*general linear models*): ANOVA, kovaryans analizleri, MANOVA, regresyon
3. Karmaşık modeller (*complex models*): Yapısal eşitlik modelleri (*structural equation models*) ve hiyerarşik doğrusal modeller (*hierarchical linear models*).

Çalışmalarda yer alan etki değerleri (a) grup ortalamaları farkına göre hesaplamalar; Cohen's *d*, *f*; (b) Varyansa göre etki büyüklüğü hesaplamaları ve (c) Korelasyon ve regresyon hesaplamaları olarak gruplandırılmıştır.

Çalışmada yer alacak makaleler seçildikten ve kontrol listesine göre incelendikten sonra elde edilen ham veriler araştırmacılar tarafından kategoriler altında kodlanmıştır. Çalışma kapsamında yer alan makalelerin kontrol listesinde bulunan maddelere göre değerlendirilmesi öznal bir yargıya bağlı olmadığından değerlendirmeciler arası güvenilirlik hesaplaması yapılmamıştır.

### **Veri Analizi**

Çalışmadan elde edilen tüm veriler nominal skalada olduğundan ve herhangi bir sayısal değer üretmediğinden verilerin analizi için parametrik olmayan testlerden ki-kare testi kullanılmıştır. Tüm ölçümler .05 anlamlılık düzeyine göre yapılmış etki büyüklüğü ölçümü olarak da Cramer's *V* hesaplanmıştır. Cohen'e göre (1988), serbestlik derecesi 2 olan ki-kare testlerinde elde edilen Cramer's *V* değeri .07 ile .21 arasında değişiyorsa küçük, .21 ile .35 arasında değişiyorsa orta ve .35'den büyük ise geniş etki büyüklüğü elde edilmiştir.

## **BULGULAR VE YORUM**

### ***1. Çalışmada yer alan dergilerde yayınlanan makalelerde etki büyüklüğünün raporlanma sıklığı nedir?***

İncelenen 990 makaleden 783 tanesi nicel (%79.4) 148 tanesi nitel (%14.9) ve 56 tanesinin (%5.7) kuramsal olduğu görülmüştür. İnceleme kapsamına 783 nicel makale dâhil edilmiş ancak bu makalelerin araştırma problemleri dikkatle incelendiğinde 303 tanesinde etki büyüklüğü raporlamaya gerek olmayan istatistiksel işlemlerin (frekans, yüzde hesabı) yapıldığı görülmüştür. Etki büyüklüğü raporlanması gereken 480 nicel makale çalışmada incelemeye alınmıştır. Etki büyüklüğü raporlanmaları incelendiğinde toplam 35 çalışmada yani toplam makalelerin %7,2'sinde uygun etki büyüklüğü raporlamasının yapıldığı görülürken; bunlardan 21 tanesinde raporlanan etki büyüklüğünün yorumladığı görülmüştür. Etki büyüklüğü raporlamasının dergilere göre dağılımı göz önüne alındığında toplam 5 yıl içerisinde A dergisinde 14, B dergisinde 11, C dergisinde 7, D dergisinde ise 3 makalede etki büyüklüğü değerinin raporlandığı görülmektedir. Dergilerde yıllara göre yayınlanan makale sayısı ve etki büyüklüğü raporlanan makale sayısının dergilere ve yıllara göre dağılımı Tablo 1.de gösterilmiştir.



Etki büyüklüğünün raporlandığı çalışmaların alanları incelendiğinde 10 tanesinin fen eğitimi alanında yapıldığı, 6 tanesinin psikolojik danışma ve rehberlik alanında yapıldığı, 5 tanesinin eğitim bilimlerinde, 3 tanesinin bilgisayar ve öğretim teknolojileri eğitiminde, 2 tanesinin okul öncesi eğitiminde yapıldığı bulunmuştur.

**Tablo 1.** Dergilerde Yayımlanan Toplam Makale Sayısı ve Etki Büyüklüğü Raporlanan Makale Sayısının Yıllara ve Dergilere Göre Dağılımı

Dergi	Yıllar	Toplam Makale Sayısı	Etki Büyüklüğü Raporlanan Makale Sayısı
A Dergisi	2007	53	2
	2008	55	2
	2009	50	2
	2010	59	7
	2011	37	1
	Toplam	254	14
B Dergisi	2007	31	0
	2008	38	1
	2009	53	2
	2010	57	2
	2011	70	6
	Toplam	249	11
C Dergisi	2007	40	1
	2008	31	0
	2009	54	1
	2010	62	4
	2011	54	1
	Toplam	241	7
D Dergisi	2007	61	0
	2008	43	0
	2009	47	1
	2010	51	1
	2011	44	1
	Toplam	246	3

### 2. Etki büyüklüğü raporlanan makaleler dergiye göre değişmekte midir?

Etki büyüklüğü raporlama sıklıkları dergilere göre incelendiğinde A dergisinde toplam 14, B dergisinde toplam 11, C dergisinde toplam 7 ve D dergisinde toplam 3 etki büyüklüğü raporlandığı görülmektedir. Ayrıca bağımsız ki-kare sonuçları dergilere göre raporlanma sıklıkları arasında istatistiksel olarak anlamlı bir ilişki bulunmadığını göstermektedir  $\chi^2(3, n = 483) = 7.30, p = .06$ , Cramer's V = .12.

### 3. Yıllara göre etki büyüklüğü raporlamaları bir farklılık göstermekte midir?

Yıllara göre etki büyüklüğü raporlanmaları incelendiğinde 2007 ve 2008 yıllarının her birinde üçer makalede etki büyüklüğü raporlanması bulunurken 2009 yılında 6; 2010 yılında 14 ve 2011 yılında 9 makalede etki değerlerinin raporlandığı bulunmuştur. Yıllara göre etki büyüklüğü raporlanmalarında istatistiksel farklılık olup olmadığını incelemek için ki-kare analizleri yapılmış analiz sonucunda yıllara göre etki büyüklüğü raporlanmalarında istatistiksel bir farklılık bulunamamıştır  $\chi^2(4, n = 483) = 6.77, p = .15$ , Cramer's V = .12.

## TARTIŞMA

Etki büyüklüğü (*effect size*), son yıllarda eğitim alanında yapılan araştırmalarda önem verilen bir kavram olarak karşımıza çıkmaktadır. Aynı zamanda APA tarafından, araştırmalarda *p* anlamlılık değeri ile birlikte mutlaka etki büyüklüğü değerinin de hesaplanarak raporlanması gerektiği belirtilmekte; APA tarafından yayımlanan akademik dergilerde, etki büyüklüğü raporlamayan araştırmalara yer

verilmemektedir. APA dergilerinde bir çalışmanın yayımlanabilmesi için etki büyüklüğü raporlanması bir ön koşul olarak kabul edilmeye başlanmıştır.

Etki büyüklüğü raporlaması üzerine yapılan önceki çalışmalarda raporlama sıklıklarının %1'den %90'a kadar değişiklik gösterdiği bulunmuştur (Meline ve Schmitt, 1997; Thompson, 1999b). Bu çalışmada elde edilen etki büyüklüğü raporlama sıklığının %7.2 olduğu göz önünde bulundurulduğunda ülkemizde yapılan çalışmalarda etki büyüklüğü raporlamanın hala çok düşük bir seviyede olduğu sonucuna varılabilir. Çalışmadan elde edilen veriler, ayrıca, etki büyüklüğü raporlanan çalışmaların %60'ında elde edilen değer yorumlandığını göstermektedir. Literatürde yer alan diğer bazı çalışmalarda etki büyüklüğü yorumlaması üzerine yakın yüzdeler raporlanmıştır (Alhija ve Levy, 2008; Hutchins ve Henson, 2002; Meline ve Wang, 2004; Sun ve Wang, 2010). Çalışmada incelenen dergilerin SSCI'da taranan dergiler olduğu göz önüne alındığında elde edilen sonucun ne kadar yetersiz olduğu açıktır. Ayrıca, elde edilen bu sonucun ülkemizde yayınlanan ve alan indekslerinde taranmayan diğer dergilerde yayınlanan makalelerde daha da düşük olacağı tahmin edilmektedir. Araştırmacının elde ettiği bulguları anlaması ve en doğru ve iyi bir şekilde uygun bir şekilde raporlaması gerekmektedir. Bunun için bir araştırmacının hem istatistiksel anlamlılık testlerinin sonuçlarını hem de etki büyüklüğü hesaplama sonuçlarını raporlaması ve yorumlaması gerekmektedir.

Etki büyüklüğü raporlayan çalışmalar incelendiğinde raporlanan etki büyüklüğünün çoğunlukla ortalama ölçümlerine ait olduğu ve eta kare ve Cohen's *d* sonuçlarının raporlandığı görülmektedir. Raporlanan çalışmalarda yapılan temel analizler dikkate alındığında, çalışmaların 24 tanesinde ANOVA; 7 tanesinde MANOVA, 4 tanesinde ise t-testi yapıldığı gözlenmektedir. Elde edilen bu sonuçlar literatürden farklılık göstermektedir (Alhija ve Levy, 2007; Hutchins ve Henson, 2002; Kirk, 1996). Konuyla ilgili literatürde çalışmalarda temel analiz olarak genelde genel doğrusal modeller kullanıldığı ve bu yüzden de ilişki gücü ölçümlerinin raporlandığı belirtilmiştir.

Çalışmada yer alan 35 makaleden 21'inde hesaplanan etki değerleri yorumlanmıştır. Etki değerlerinin yorumlandığı bu 21 makalede Cohen'in belirttiği gibi etki değerleri küçük, orta ve büyük olarak yorumlanmış, önceki araştırmalardan elde edilen etki büyüklüklerine değinilmemiştir. Makalelerde etki büyüklüğü ile ilgili daha ayrıntılı bir yorumlamanın olmayışı dikkat çekicidir.

Bu çalışmada, amaçlı örnekleme kullanıldığı için çalışmanın sonuçlarını yayınlanan tüm makalelere genellemek mümkün değildir. Ancak, ele alınan dergiler, eğitim alanında en önemli indekslerden biri olan, SSCI'da taranan Türk dergileri olduğu için elde edilen veriler dikkate değer görülmektedir. Bu dergilerde bile etki büyüklüğü raporlamalarının %10'u geçmemesi, diğer dergilerde bu oranın daha düşük olacağını düşündürmesine rağmen, bu konuda daha detaylı araştırmalara ihtiyaç duyulmaktadır. Etki büyüklüğü raporlaması ve yorumlanmasını konu alan çalışmaların artması, araştırmacıların ve dergi editörlerinin bu konuya olan ilgisini arttıracak, bu durum etki büyüklüklerinin raporlanma sıklığını arttıracaktır.

Bir araştırmacının, araştırması sırasında etki büyüklüğünü hesaplaması ve yayınında bu etki büyüklüğünü raporlaması ve yorumlaması, o konuyla ilgilenen diğer tüm araştırmacılara oldukça kıymetli bilgiler verecektir. Bu sayede, ortak bir dil konuşmak mümkün olacak, farklı çalışmaların sonuçları rahatlıkla karşılaştırılabilir, elde edilen istatistiksel anlamlılığın pratikteki büyüklüğü de belirlenmiş olacaktır. Yapılan her çalışma kendi başına bir değer olduğu gibi, bu çalışmalardan elde edilen sonuçların toplanması ve çalışılan konunun etkiliği hakkında bir değerlendirme yapılması da gerekmektedir. Bu anlamda meta-analiz çalışmalarının önemi kaçınılmazdır. Meta-analiz çalışmaları ile bir alanda yapılan bireysel çalışmalardan elde edilen bulgular birleştirilir, birlikte analiz edilir ve yorumlanır. Bu sayede, ele alınan bir yöntemin, etkinliğin veya materyalin etkililiği sınanmış olur. Meta-analiz, birçok araştırma sonucunun ortak bir ölçü birimine çevrilerek karşılaştırılmasını ve istatistiksel işlemlerle etki büyüklüklerinin hesaplanmasını sağlar. Etki büyüklüklerinin hesaplanması bu anlamda da büyük önem taşımaktadır. Bazı durumlarda, araştırmacılar en basit istatistiksel sonuçları bile çalışmalarında raporlamayarak, etki büyüklüğünün hesaplanmasını imkânsız hale getirmektedirler.

Bu çalışmanın kapsamında yer almasa da ilgili literatürde yayınlanan makale ve tezlerde birçok istatistiksel ve metodolojik yanlışlıklar yapıldığı çeşitli araştırmacılar tarafından raporlanmıştır (Kabaca

ve Erdoğan, 2007; Özsoy, Keleş ve Uzun, 2011; Sayın, 2008; Sönmez, 2005; Tonta, 1999; Toy ve Tosunoğlu, 2007). Araştırma modelinin doğru raporlanmaması, araştırmacının belirtilen modele uygun olarak tasarlanmamış olması, evren ve örneklemin uygun şekilde belirlenmemiş olması, evren ve örneklemin seçiminde kullanılan ölçütlerin raporlanmaması, istatistiksel analiz yöntemlerinin seçiminde testlerin gerektirdiği ön koşulların sağlanıp sağlanmadığına bakılmaması, istatistiksel verilerin eksik ve hatalı raporlanması daha önceki çalışmalarda raporlanan en temel sorunlar arasında yer almaktadır. Durum böyleyken, araştırmacıların etki büyüklüğünün raporlanmasını bir alışkanlık haline getirmeleri zaman alacak gibi gözükmektedir.

Ülkemizde sosyal bilimlerin gelişmesi ve bu alanda yapılan çalışmaların dünyada daha çok kabul görmesi için doğru, güvenilir ve hatasız yapılmış olan çalışmalara ihtiyaç duyulmaktadır. Bu nedenle, araştırmacıların, araştırmanın planlanmasından raporlanmasına kadar geçen her aşamada titizlikle çalışması, olası hatalara karşı hazırlıklı olması gerekmektedir. Bu hataların en aza indirilebilmesi için, araştırmacıların araştırma sırasında yaşayacakları sorunların farkında olmaları ve olası çözüm yolları konusunda bilgilerinin olması gerekmektedir. Araştırmacıların bu donanımla yetişebilmesi için lisansüstü eğitimde araştırma teknikleri ve istatistik derslerinin içeriği geliştirilmeli bu derslerin sayısının artırılması gerekmektedir. Bunun yanı sıra dergilerde yayınlanan makalelerin bir inceleme sürecinden geçtiği düşünüldüğünde, bu sürecin daha sağlıklı yürütülmesi bu hataların azaltılması için alınabilecek önlemlerden birisidir. Bu nedenle yazarlara olduğu kadar dergi editörleri ve hakemlerine de önemli bir sorumluluk düşmektedir.

## KAYNAKÇA

- Alhija, F. N. & Levy, A. (2007, April). *Effect size reporting practices in published articles*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.
- American Educational Research Association. (2006). Standards on reporting on empirical social science research in AERA publications. *Educational Researcher*, 35, 33-40.
- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.
- Balcı, A. (2009). Sosyal Bilimlerde Araştırma: Yöntem, Teknik ve İlkeler. Ankara: Pegem Akademi Yay.
- Brewerton, P. & Millward, L. (2001). *Organizational Research Methods: A Guide for Students and Researchers*, Sage Publications, London, GBR.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cohen, J. (1990). Things I have learned so far. *American Psychologist*, 45, 1304-1312.
- Cohen, J. (1994). The earth is round ( $p < .05$ ). *American Psychologist*, 49, 997-1003.
- Davis, J. A. (1971). *Elementary survey analysis*. Englewood Cliffs, NJ: Prentice-Hall.
- Daymon, C. & Holloway, I. (2002). *Qualitative Research Methods in Public Relations and Marketing Communications*, Routledge, Florence, KY, USA.
- Dunleavy, E. M., Barr, C. D., Glenn, D. M., & Miller, K. R. (2006). Effect size reporting in applied psychology: How are we doing? *The Industrial-Organizational Psychologist*, 43, 29-37.
- Falk, R. & Greenbaum, C. W. (1995). Significance tests die hard: The amazing persistence of a probabilistic misconception. *Theory & Psychology*, 5, 75-98.
- Fan, X. (2001). Statistical significance and effect size in education research: Two sides of a coin. *Journal of Educational Research*, 94, 275-283.
- Fraenkel, J. R. & Wallen, N. E. (2003). *How to Design and Evaluate Research in Education*, McGraw-Hill, New York.
- Ghauri, P. & Gronhaug, K. (2002). *Research Methods in Business Studies*, Prentice Hall, London.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5, 3-8.
- Hedges, L. V. (1981). Distributional theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6, 107-128.
- Henson, R. K. (2006). Effect size measures and meta-analytic thinking in counseling psychology research. *The Counseling Psychologist*, 34, 601-629.
- Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1979). *Applied statistics for the behavioral sciences*. Chicago: Rand McNally College Publishing.

- Hopkins, W. G: (1997). New view of statistics. Retrieved, September 12, 2011 from <http://www.sportsci.org/resource/stats/effectmag.html>
- Huberty, C. J., & Pike, C. J. (1999). On some history regarding statistical testing. In B. Thompson (Ed.), *Advances in social science methodology* (Vol. 5, pp.1-23). Stamford, CT: JAI Press.
- Hutchins, H. M., & Henson, R. K. (2002). *In search of OZ: Effect size reporting and interpretation in communication research*. Paper presented at the annual meeting of the Southwest Educational Research Association, Austin, TX.
- Ives, B. (2003). Effect size use in studies of learning disabilities. *Journal of Learning Disabilities*, 36,490-504.
- İşman, A. Baytekin, Ç., Balkan, F., Horzum, B. & Kıyıcı, M. (2002). Fen Bilgisi Eğitimi ve Yapısalçı Yaklaşım. *The Turkish Online Journal of Educational Technology*, 1(1), 41-47.
- Kabaca, T. ve Erdoğan, Y. (2007). Fen Bilimleri, Bilgisayar ve Matematik Eğitimi Alanlarındaki Tez Çalışmalarının İstatistiksel Açından İncelenmesi. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 22(2).
- Kirk, R. E. (1996). Practical significance: A concept whose time has come. *Educational and Psychological Measurement*, 56, 746-759.
- Kline, R. B. (2004). *Beyond significance testing: Reforming data analysis methods in behavioral research*. Washington, DC: American Psychological Association.
- Kotrlík, J. W. & Williams, H. A. (2003). The incorporation of effect size in information technology, learning, and performance research. *Information Technology, Learning, and Performance Journal*, 21(1), 1-7.
- Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analyzing data: A model comparison perspective*. Belmont, CA: Wadsworth.
- Meline, T., & Schmitt, J. F. (1997). Case studies for evaluating statistical significance in group designs. *American Journal of Speech-Language Pathology*, 6, 33-41.
- Meline, T., & Wang, B. (2004). Effect-size reporting practices in AJSLP and other ASHA journals, 1999-2003. *American Journal of Speech Language Pathology*, 13,202-207.
- Nickerson, R. S. (2000). Null hypothesis significance testing: A review of an old and continuing controversy. *Psychological Methods*, 5, 241-301.
- Olejnik, S. F. (1984). Planning educational research: Determining the necessary sample size. *Journal of Experimental Education*, 53, 40-48.
- Ottenbacher, K. J., & Barrett, K. A. (1989). Measures of effect size in the reporting of rehabilitation research. *American Journal of Physical Medicine & Rehabilitation*, 68, 52-58.
- Özsoy, S., Keleş Ö. & Uzun, N. (2011). Methodological and Statistical Errors Found in Science Education Master's Theses. *The International Journal of Educational Researchers*, 2(5), 35-46.
- Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research*, 34, 1189-1208.
- Paul, K. M., & Plucker, J. A. (2003). Two steps forward, one step back: Effect size reporting in gifted education research from 1995-2000. *Roeper Review*, 26, 68-72.
- Plucker, J. A. (1997). Debunking the myth of the "highly significant" result: Effect sizes in gifted education research. *Roeper Review*, 2, 122-126.
- Polonsky, M. J. & Waller, D. S. (2005). *Designing and Managing a Research Project*, Sage Publications, U.S.A.
- Robinson, D. H., & Levin, J. R. (1997). Reflections on statistical and substantive significance, with a slice of replication. *Educational Researcher*, 26, 21-27.
- Sayın, S. (2008). Bilimsel araştırmalarda yapılan bazı istatistiksel ve yöntembilimsel Hatalar-III: Güvenirlik Kestirimlerine Yönelik Hatalar . *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, 15.
- Snyder, P. A. & Lawson, S. (1993). Evaluating results using corrected and uncorrected effect size estimates. *Journal of Experimental Education*, 61, 331-349.
- Snyder, P. A., & Thompson, B. (1998). Use of tests of statistical significance and other analytic choices in a school psychology journal. *School Psychology Quarterly*, 13, 335-348.
- Spiegel, M. R. & Stephens, L. J. (1999). *Schaum's outline of theory and problems of statistics*. 3rd edition. *Schaum's outline series*, McGraw-Hill, New York.
- Sönmez, V. (2005). Bilimsel Araştırmalarda Yapılan Yanlışlıklar. *Eurasian Journal of Educational Research*, 18, 150-170.

- Sun, S., Pan W., & Wang, L. L. (2010). A comprehensive review of effect size reporting and interpreting practices in academic journals in education and psychology. *Journal of Educational Psychology*. Advance online publication. doi: 10.1037/a0019507
- Thompson, B. (2008). Computing and interpreting effect sizes, confidence intervals, and confidence intervals for effect sizes. In J. W. Osborne (Ed.), *Best practices in quantitative methods* (pp. 246-262). Thousand Oaks, CA: Sage.
- Thompson, B. (2002). "Statistical", "practical", and "clinical": How many kinds of significance do counselors need to consider? *Journal of Counseling & Development*, 80, 64-71.
- Thompson, B. (1999). If statistical significance tests are broken/misused, what practices should supplement or replace them? *Theory and Psychology*, 82, 165-181.
- Thompson, B. (1996). AERA editorial policies regarding statistical significance testing: Three suggested reforms. *Educational Researcher*, 25, 26-30.
- Thompson, B. (1993). The use of statistical significance tests in research: Bootstrap and other alternatives. *Journal of Experimental Education*, 61, 361-377.
- Thompson, B., & Snyder, P. A. (1997). Statistical significance testing practices in the Journal of Experimental Education. *Journal of Experimental Education*, 66, 75-83.
- Tonta, Y. (1999). Bilimsel Arařtırmalarda İstatistik Tekniklerin Kullanımı ve Bulguların Sunumu Üzerine. *Türk Kütüphaneciliđi*, 13 (2), 112-124.
- Toy, Y. B., ve Tosunođlu, G. N. (2007). Sosyal Bilimler Alanındaki Arařtırmalarda Bilimsel Arařtırma Süreci, İstatistiksel Teknikler ve Yapılan Hatalar, *Ticaret ve Turizm Eđitim Fakóltesi Dergisi*, Sayı: 1.
- Vache-Haase, T., & Ness, C. M. (1999). Statistical significance testing as it relates to practice: Use within professional psychology: Research and practice. *Professional Psychology: Research and Practice*, 30, 104-105.
- Vacha-Haase, T. & Thompson, B. (2004). How to estimate and interpret various effect sizes. *Journal of Counseling Psychology*, 51, 473-481.
- Yates, F. (1951). The influence of "statistical methods for research workers" on the development of science of statistics. *Journal of the American Statistical Association*, 46, 19-34.
- Yıldırım, H. H. & Yıldırım, S. (2011). Hipotez testi, güven aralığı, etki büyüklüğü ve merkezi olmayan olasılık dağılımları üzerine. *İlköđretim Online*, 10(3), 1112-1123.