



KENDİNİ DEĞERLENDİRME, AKRAN DEĞERLENDİRME VE ÖĞRETMEN DEĞERLENDİRMEİNİN YAZILI SINAV SONUÇLARINA ETKİSİ VE BAŞARI YORDAYICILIĞI*

*Mehmet TAŞDEMİR***

ÖZET

Kendini değerlendirme, akran değerlendirme ve öğretmen değerlendirmenin yazılı sınav sonuçlarına etkisi ve başarı yordayıcılığını ortaya çıkarmayı amaçlayan bu araştırma, nicel türde tek faktörlü yordayıcı korelasyon araştırması desenindedir.

Bu araştırma ile öğrenci başarısını ortaya çıkarmada kullanılan yazılı yoklama sınavlarının puanlamasında işe koşulan öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarının not vermeye olan etkisi ve bu yaklaşımlarla yapılan ölçüm sonuçlarının öğrenci başarısını olan yordayıcılık gücünün ne olduğunun ortaya çıkarılması amaçlanmıştır? Bu amaç doğrultusunda aşağıdaki sorulara cevap aranmıştır:

1. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçları değişmekte midir?
2. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımları göre ölçme sonuçları arasında anlamlı bir ilişki var mıdır?
3. Cinsiyete göre öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçları değişmekte midir?
4. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçlarında cinsiyetin etkisi nedir?
5. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçlarının dönem sonu başarıyı yordama gücü nedir?

Araştırma 121 öğrenci ile yürütülmüştür. Araştırma verileri geliştirilen yazılı yoklama ara sınav testi ve çoktan seçmeli türde dönem sonu sınavı testi ile elde edilmiştir. Ara sınav testinin hesaplanan Cronbach Alpha iç tutarlık katsayısı 0,97, dönem sonu sınavının hesaplanan iç tutarlık katsayısı ise $K-R_{21} = 0,78$ 'dir. Verilerin analizinde

*Bu makale Crosscheck sistemi tarafından taranmış ve bu sistem sonuçlarına göre orijinal bir makale olduğu tespit edilmiştir.

** Doç. Dr. Ahi Evran Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Eğitim Programları ve Öğretim ABD, El-mek: mtasdemir@ahievran.edu.tr & mttasdemir1963@gmail.com



t-testi, Pearson Correlation, Kendall's tau_b ve Spearman's rho, Cohen's d ve eta-kare(η^2), R, R Square ve F analiz teknikleri kullanılmıştır.

Araştırma bulgularına göre belirlenen alt problemlere ilişkin olarak aşağıdaki sonuçlara ulaşılmıştır:

1. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçları arasındaki fark anlamlıdır.

2. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçları arasında anlamlı yüksek bir ilişki vardır.

3. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçlarında kızlar lehine anlamlı bir fark vardır.

4. Akran değerlendirme puanlarında gözlenen varyansın yaklaşık 0,008'nincinsiyete bağlı olduğu ifade edilebilir. Kız ve erkek öğrencilerin akran değerlendirme puanları arasındaki fark 0,60 standart sapma kadardır.

5. Öğretmen değerlendirme puanlarında gözlenen varyansın yaklaşık 0,06'nın cinsiyete bağlı olduğu ifade edilebilir. Kız ve erkek öğrencilerin öğretmen değerlendirme puanları arasındaki fark 0,54 standart sapma kadardır.

6. Kendi kendini değerlendirme puanlarında gözlenen varyansın yaklaşık 0,06'nın cinsiyete bağlı olduğu ifade edilebilir. Kız ve erkek öğrencilerin kendi kendini değerlendirme puanları arasındaki fark 0,55 standart sapma kadardır.

7. Ara sınav sonuçları ile final başarısı arasında yüksek oranda doğrusal bir ilişki olduğunu söylenemez.

8. Araştırmada üç farklı puanlama yaklaşımına dayalı olarak yapılan ara sınav sonuçlarının final sınavına ilişkin yordayıcılığı akran değerlendirmeyle %23, kendi kendine değerlendirmeyle %24 ve öğretmen değerlendirmesi ile %30 oranında açıklamaktadır.

Araştırma bulgularına göre farklı puanlama tekniklerine dayalı elde edilen ölçüm sonuçları arasında yüksek ilişki olmasına karşın fark anlamlıdır. Ara sınav sonuçları ile final başarısı arasında yüksek oranda doğrusal bir ilişki olduğunu söylenemez.

Bu sonuçlara göre kullanılan puanlama yöntemleri arası ilişkisinin yüksek olması her bir farklı puanlama tekniğinde de öğrencilerin ara sınav sonuçları arasındaki tutarlılığın yüksek olduğunu göstermektedir. Ortalamalar arasındaki farkın anlamlı olması işe koşulan tekniklerle yapılan puanlamaların farklılaştığını göstermektedir.

İşe koşulan değerlendirme teknikleri içerisinde final sınavına ilişkin yordayıcılık değerlerinin en yüksek oranda öğretmen değerlendirme yaklaşımında olduğu görülmektedir. Bununla birlikte üç farklı puanlama yaklaşımı ile elde edilen ara sınav sonuçları ile final başarısı arasında yüksek oranda doğrusal bir ilişki olduğunu söylenemez. Bu sonuca ara sınav ve dönem sonu sınavı kapsamlarının

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



ara sınav lehine daha sınırlı olması ara sınav başarı ortalamasının daha yüksek çıkmasında önemli bir etken olabilir.

Anahtar Kelimeler: Akran Değerlendirme, Kendi Kendini Değerlendirme, Öğretmen Değerlendirme, Korelasyon, Yordama

THE EFFECT ON ESSAY ASSESSMENT SCORES AND REGRESSION OF SELF EVALUATION, PEER EVALUATION AND TEACHER EVALUATION TECHNIQUES

ABSTRACT

This study, which aims to discover the effect of self evaluation, peer evaluation and teacher evaluation techniques on written examination results and the success regression is one-factor predictor correlation research pattern in quantitative style.

This study has aimed to reveal the effect of teacher evaluation, peer evaluation and self evaluation approaches on grading which is used in scoring of written exams in determining student's success and how powerful the Regression of student success made by these approaches is. With this aim, it is looked for answers for the questions below.

1. Do the assessment results differ according to teacher evaluation, peer evaluation and self evaluation approaches?
2. Is there a significant relationship among the assessment results according to teacher evaluation, peer evaluation and self evaluation approaches?
3. Do the assessment results differ according to teacher evaluation in terms of gender, peer evaluation and self evaluation approaches?
4. What's the effect of gender on assessment results according to teacher evaluation, peer evaluation and self evaluation approaches?
5. How do the assessment results have an impact on the prediction of final success according to teacher evaluation, peer evaluation and self evaluation approaches?

The study has conducted with 121 students. Research data has gathered from midterm written examination test which has been developed and final test which is consisted of multiple choices. The midterm exam's calculated Cronbach's alpha internal consistency coefficient is 0,97 and the final exam's calculated internal consistency coefficient is $K-R_{21} = 0,78$. In analyzing data, T test, Pearson Correlation, Kendall's tau-b, Spearman's rho, Cohen's d, Eta-squared (η^2), R, R Square and F analysis techniques have been used

According to findings, these results have been found about the determined sub-problems.

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



1. The variation among the assessment results is significant in terms of teacher evaluation, peer evaluation and self evaluation approaches.

2. There is a high, significant relationship between the assessment results according to teacher evaluation, peer evaluation and self evaluation approaches.

3. There is a significant variation in assessment results in favor of girls according to teacher evaluation, peer evaluation and self evaluation approaches.

4. It can be stated that approximately 0,008 of the variance observed in peer evaluation points depends on gender. The variation between the female and male students' peer evaluation points is about 60 standardized deviations.

5. It can be stated that approximately 0,06 of the variance observed in teacher evaluation points depends on gender. The variation between the female and male students' teacher evaluation points is about 54 standardized deviations.

6. It can be stated that approximately 0,06 of the variance observed in self-evaluation points depends on gender. The variation between the female and male students' self evaluation points is about 55 standardized deviations.

7. It cannot be stated that there is a high linear relationship between the mid-term exam results and final success.

8. In the research it's seen that the mid-term exam results occurred in terms of three different evaluation approaches explain the prediction of final exam as peer evaluation 23% ,self evaluation 24% and teacher evaluation at the rate of 30%.

Although there are high correlations among evaluation results which have been obtained from different scoring techniques according to the research findings, the variation is significant. But it is unspeakable that there is a high linear correlation between the results of midterm exam and the success rates of final exam.

According to these results, high relationship between the used grading techniques points out that in each different grading technique, students' coherence between mid-term exam results is high. The variation's being significant among averages shows that used techniques and scoring differ.

Among the used techniques, the most highly prediction values related with final exam have been seen in teacher evaluation approach. However, it cannot be said that there is a high linear relationship between the final success and mid-term results, which are gained by using three different techniques. The scopes of mid-term and final exam's being limited in favor of the mid-term exam may be an important factor for mid-term exam's average coming out higher.

Key Words: Peer Evaluation, Self Evaluation, Teacher Evaluation, Corelation, Regression

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



doi



GİRİŞ

Geleneksel ölçme ve değerlendirme yaklaşımlarına öğrenenlerin başarıları, gelişimleri, neler öğrenip/öğrenemediklerini yeterli derecede yansıtamadığı şeklinde eleştiriler yapılmaktadır (Burke, 1999; Stiggins, 1999). Yeşilyurt'a (2012)göre, öğretmenler ölçme tekniklerinden en fazla geleneksel yöntemlerden yazılı yoklama, kısa cevap testleri, seçmeli sınavlar, sınıflama sınavları ve eşleştirmeli sınavlar ile alternatif ölçme tekniklerinden performans dayalı işlemler, projeler ve ürün seçki dosyasını kullanmaları yanında objektif değerlendirme yapma konusunda güçlük yaşamaktadırlar. Dönder, Elaldı ve Özkaya (2012) araştırma bulgularına göre ise tamamlayıcı ölçme ve değerlendirme tekniklerinin öğretim elemanları tarafından büyük ölçüde önemli ve faydalı görülmesine karşın en fazla kullanılan ölçme ve değerlendirme tekniği olarak ise kâğıt-kalem tekniğinin (%81) öne çıktığı görülmektedir. Gökyer ve Özer (2014)'e göre öğretmenler genel ölçme ve değerlendirme yeterlikleri bakımından kendilerini yeterli görmektedirler. Buna karşın, Abazaoğlu, Yıldırım ve Yıldızhan (2014) ve Kazu ve Aslan (2012) bulgularında ise öğretmenlerin yaklaşık yarısı alternatif ölçme değerlendirme eğitimi ihtiyacında olmalarına rağmen hemen hepsi (%89)'u klasik ölçme değerlendirme konusunda eğitim ihtiyacı içerisinde değildirlir. Yaygın kullanımına karşın yazılı yoklama puanlarına olan güven genelde az olmaktadır. Aynı şekilde yazılı yoklamaların puanlamasında da birçok farklı teknik kullanılmaktadır. Bunlardan en yaygın kullanılanı öğretmenin kendi sınavlarını yine kendisinin puanlamasıdır. Bu puanlama sürecinde *genel izlenim, sınıflama, sıralama ve analitik puanlama* yaklaşımlarından biri işe koşulmaktadır (Turgut, 1993; Tekin, 2007; Taşdemir, 2011).

Yazılı yoklamalar üzerine odaklaşan eleştiriler genel olarak kapsam geçerliği ve puanlama güvenilirliği üzerine olmaktadır. Özellikle bu iki nitelik yazılı yoklama puanlarının yordayıcılık gücü hakkında da önemli tartışmaların odağı olmaktadır.

Kendi kendini değerlendirme/öz değerlendirme

Öz değerlendirme, bireyin bir konuda öğrenme sürecindeki başarı düzeyini ve sonuçlarını yargılaması, bireyin kendisini değerlendirmesidir (MEB-URL, 2013; Çepni, ark., 2007). Bu değerlendirme yaklaşımının birçok farklı uygulaması bulunmakla birlikte yazılı yoklamaların ya da sınavların okunması sürecinde öğrencinin bizzat kendi sınav evrakını kendisinin puanlaması şeklinde gerçekleşmektedir. Literatürde bu uygulamanın üstünlüklerine dönük birçok araştırma bulunmaktadır. Chung (2000) bulgularına göre ilköğretim öğrencilerinin öz yeterlilikleri alternatif değerlendirme araçlarından biri olan öz değerlendirmenin kullanımından sonra anlamlı bir şekilde yükselmektedir. Benzer şekilde Zimbicki (2007)'nin alternatif değerlendirme yaklaşımlarının öğrencilerin motivasyon ve öz yeterlilik düzeyleri üzerine etkisini araştırdığı çalışmasına göre, alternatif değerlendirme yaklaşımları üzerine temellendirilmiş ders gören öğrencilerin öz yeterlilik düzeyleri yükselmektedir. Özan (2008) bulgularında öz ve akran değerlendirmenin temel iletişim becerileri üzerinde olumlu etkiye sahip olduğu saptanmıştır.

Akran değerlendirme

Akran değerlendirme bireyin yapmış olduğu çalışmanın kendi arkadaşları tarafından belirli ölçüler çerçevesinde kendi bakış açılarına göre değerlendirildiği bir süreçtir (Topping, 1998; Çepni ve ark., 2007).

Yazılı yoklama ve akran değerlendirme üzerine yapılmış araştırmalarda, değerlendirme ölçütlerini arkadaşlık ilişkileri üzerine kurma, akran değerlendirmeye katılmama ve bu konuda kendilerine güven duymama, akranını değerlendirmeye kaygılı yaklaşma, baskılı hissetme gibi bulgulara ulaşılmıştır (Sluijsmans ve ark., 1999; Sluijsmans ve ark., 2003; Sluijsmans, 2002; Cartney, 2010; Koç, 2011). Bunlara karşın birçok araştırmada akran değerlendirmenin öğrencileri gelecekteki meslek yaşamlarında karşılaşacakları değerlendirilme süreçlerine hazırladığı, eleştirel

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



düşünmeyi, işbirliği becerilerini ve etkileşimi geliştirdiği, öğrenciler ile öğretmenler arasındaki sosyal etkileşimin kalitesini artırdığı, sözlü anlatım becerileri, eleştirme ve eleştirilere açık olma gibi pek çok sosyal ve iletişim becerilerini, öğrenmeyi artırdığı, öğretmenlik yeterliklerini geliştirdiği, karşılaştırma yapma imkânı verdiği yönünde bulgulara ulaşılmıştır (Sluijsmans ve ark., 2003; Sluijsmans ve Prins, 2006; Topping, 2009; Vickerman, 2009; Gielen ve ark., 2010; Wiley ve Gardner, 2010; Koç, 2011; Baskonuş ve Taşdemir, 2012).

Yordama

Ölçme sonuçlarının taşınması gereken önemli niteliklerden birisinin de yordayıcılık gücünün yüksek olmasıdır. Sınav sonuçları işe koşulduğu amaca uygun olarak bireyin olası gelişiminin bir kestirimcisi de olabilmelidir. Sınavların yordayıcılık değerlerine ilişkin araştırmalar genelde orta öğretim ve yüksek öğretime geçişte işe koşulan merkezi birimlerce hazırlanıp uygulanan sınavlar üzerine odaklanmıştır. Araştırma bulgularında ÖSS ve benzer amaçlı yapılan sınavların belli ölçüde (-0,28 ila 0,30) yükseköğretimdeki başarıyı yordadığını ortaya koymuştur (Büyükoztürk, 2004; Çetin ve Mahir, 2006; Deaton ve Schutz, 2001; Gülleroğlu, 2005; Kutlu ve Karakaya, 2003; Karakaya ve Tavşancıl, 2008; Noble ve Sawyer, 2002; Önen ve Demirtaşlı, 2004; Qualls ve Ansley, 1995; Zwick ve Sklar, 2005). Buna karşın Deniz ve Kelecioğlu (2005) araştırmalarında Vatandaşlık dersinin OKÖSYS testindeki başarıyı yordamadığı ve onun dışındaki derslerden bazılarının birinci dönem notları ile ilgili alt test arasında; bazılarının ise ikinci dönem notları ile ilgili alt test arasında manidar ilişkiler bulunmuştur. Benzer şekilde Çelikkaleli ve Akbaş (2007) araştırmasında “sonuç beklentisi puanları”nın öğretmen adaylarının fen bilgisi dersine yönelik tutum varyansını yordamadığı bulgusuna ulaşılmıştır.

Cronbach’a, (1984) göre yordama çalışmalarının en zor tarafı test puanlarının yordanmasında uygun bir ölçütün bulunmamasıdır. Genelde yordama araştırmalarında ölçüt olarak “akademik başarı”, “yaşamdaki başarı” ve “mesleki başarı” alınmaktadır. Bu araştırmada “dönem sonu başarı durumu” ölçüt olarak alınmıştır.

Araştırmanın amacı

Yazılı yoklamaların puanlamasında akran değerlendirme, kendi kendini değerlendirme yaklaşımı ve öğretmen değerlendirme yaklaşımlarından hangisinin diğerine göre daha uygun olduğu ve bu yaklaşımlarla yapılan puanlamaların genel akademik başarıyı yordama düzeyine ilişkin araştırmalara olan ihtiyaç vardır. Bu araştırma ile bu amaca hizmet edeceği düşünülen bulgulara ulaşmak hedeflenmiştir.

Araştırmanın önemi

Okullarda yazılı yoklama sınavlarının puanlamasında genel olarak öğretmen değerlendirme yaklaşımlarının işe koşulduğu göz önüne alındığında özellikle akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarının başarı puanlarının ortaya çıkarılmasında işe koşulabilecek bir puanlama yaklaşımı olarak kullanılıp kullanılmayacağına ilişkin araştırma bulgularına ihtiyaç vardır.

Problem cümlesi

Başarı değerlendirme sürecinde işe koşulan yazılı yoklamaların puanlama yöntemlerinden “öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarının not vermeye olan etkisi ve bu puanlama sonuçlarının öğrenci başarısını yordayıcılık gücü farklılaşmakta mıdır?

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



Alt problemler

1. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçları değişmekte midir?
2. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçları arasında anlamlı bir ilişki var mıdır?
3. Cinsiyete göre öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçları değişmekte midir?
4. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçlarında cinsiyetin etkisi nedir?
5. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçlarının dönem sonu başarıyı yordama gücü nedir?
 - 5.1. Ara sınavda işe koşulan kendi kendini değerlendirme sonuçları final sınavı sonuçları ile doğrusal bir özellik göstermekte midir?
 - 5.2. Ara sınavda işe koşulan akran değerlendirme sonuçları final sınavı sonuçları ile doğrusal bir özellik göstermekte midir?
 - 5.3. Ara sınavda işe koşulan öğretmen değerlendirme sonuçları final sınavı sonuçları ile doğrusal bir özellik göstermekte midir?

YÖNTEM

Araştırma nicel araştırma türlerinden korelasyonel bir araştırmadır. Korelasyonel araştırma iki ya da daha çok değişken arasındaki ilişkinin herhangi bir şekilde bu değişkenlere müdahale edilmeden incelendiği araştırmalardır (Büyüköztürk ve ark., 2012:186). Frankel & Wallen (2006)'e göre korelasyonel araştırmalar keşfedici ve yordayıcı korelasyonel araştırmalar olarak ikiye ayrılır. Keşfedici korelasyonel araştırmaları iki değişken arası ilişkileri çözümleyerek önemli bir olayı anlamaya çalışmak için kullanılırken, yordayıcı korelasyon araştırmalarında değişkenler arası ilişkiler incelenerek değişkenlerin birinden yola çıkarak diğeri yordanmaya çalışılır (Akt. Büyüköztürk ve ark., 2012:186). İki değişken arasındaki ilişki ne kadar yüksekse bu belirlemenin o kadar doğru yapılabileceğinden söz edilebilir (Büyüköztürk ve ark., 2012:186).

Bu araştırmada yazılı yoklamaların puanlama yöntemlerinden akran değerlendirme, kendi kendini değerlendirme ve öğretmen değerlendirme teknikleri ile not verme ile öğrenci başarı puanlarının değişip değişmediği ve bu farklı teknikler ile verilen notlar arası ilişki ile ara sınav notlarının final sınav sonuçlarını yordama düzeyi belirlemeye çalışmıştır. Bu yönü ile bu araştırma korelasyonel araştırma desenlerinden tek faktörlü yordayıcı korelasyon deseniindedir.

Çalışma grubu

Araştırma 2012–2013 öğretim yılı güz döneminde Ahi Evran Üniversitesi Eğitim Fakültesi İlköğretim sınıf öğretmenliği ABD üçüncü sınıf öğrencileri ile yürütülmüştür. Araştırmada toplam 135 öğrenciden 35'i erkek, 86'sı kız, toplamda 121 öğrenci çalışma grubunu oluşturmuştur. Çalışma grubunun oluşturulmasında puanlama çalışmasının yapıldığı anda sınıfta mevcut bulunan tüm öğrencilerin gönüllü katılımı sağlanmıştır.

Veri toplama teknikleri

Araştırmada veri toplama aracı olarak araştırmacı tarafından geliştirilen sekiz yazılı yoklama türü sorudan oluşan ara sınav testi ile elli çoktan seçmeli sorudan oluşan dönem sonu sınavı testi kullanılmıştır.

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



Ölçme aracının hazırlanması ve puanlamasında işe koşulan süreç şöyledir:

1. Ölçme amacı belirlenmiştir. Bu amaçla araştırmacı tarafından yürütülen dersin ara sınavı araştırmacının ilk verilerinin toplama aşamasını, dönem sonu sınavı da ikinci veri toplama aşamasını oluşturmuştur. Dolayısı ile katılımcıların sınav evrakına maksimum performanslarını yansıtmaları sağlanmıştır.
2. Ölçme hedefleri belirlenmiştir. Dersin kapsam geçerliğini sağlamak üzere hazırlanmış olan belirtke tablosundaki oranlara göre kapsamı örnekleyecek sayı ve özellikte sorular oluşturulmuştur. Dersi yürüten araştırmacının aynı dersin önceki yıllarda uygulanmış olan sorularından da yararlanılarak geçerlik ve güvenilirliği artırılmıştır.
3. Soruların niteliği ve kapsamı örnekleyiciliği ile ilgili olarak beş uzman görüşüne yer verilmiştir. Bu süreçte Miles ve Huberman (1994) tarafından önerilen *görüş birliği* ve *görüş ayrılığı sayıları* hesaplanarak uygulamaya sunulmuştur. Nitel araştırmalarda bu uyum katsayısının 0,70'in üzerinde olması kabul edilebilir güvenilirlik değeri olarak alınır. Bu araştırmada uzman görüşlerinin uyum yüzdesi: 1.soru için 0,80; 2.soru için 0,60; 3.soru için 0,80; 4.soru için 0,80; 5.soru için 1,00 ; 6.soru için 0,80 ;7.soru için 1,00 ; 8.soru için 0,80 ve genel olarak ise 0,83 olarak hesaplanmıştır.
4. Sınavlar yasal bir sınav sürecinde öğretim elemanları gözetmenliğinde gerçekleştirilmiştir. Bu açıdan sınav güvenliği sağlanmıştır.
5. Arasınav kağıtlarının puanlanması için; İlk olarak araştırmacı (aynı zamanda dersin yürütücüsü) tarafından ayrıntılı bir analitik puanlama anahtarı hazırlanmıştır. İkinci olarak her bir öğrencinin yazılı yoklama kağıtları ve analitik puanlama anahtarı orijinalinden kopya edilerek yeterli miktarda çoğaltılmıştır. Üçüncü olarak her bir öğrenciye sınıfındaki bir akranının cevap kağıdı random yolla verilerek analitik puanlama anahtarı yardımı ile puanlaması yaptırılmıştır. Dördüncü olarak öğrencilere kendi cevap kağıdı verilerek puanlaması yaptırılmıştır. Bu süreçte araştırmacı gerekli açıklama ve yönlendirmelerle öğrencilerin gerekli duyarlılıkta yaklaşım göstermelerini sağlamıştır. Son olarak başka bir ortamda öğretmen aynı puanlama anahtarı yardımı ile her bir öğrenci kağıdını puanlamıştır.
6. Katılımcı öğrencilerin dönem sonu başarı durumları araştırmacı tarafından geliştirilen elli sorudan oluşan çoktan seçmeli türündeki bir ölçme aracı ile elde edilmiştir.
7. Katılımcı her bir öğrencinin puanlanmış sınav kâğıtları kodlanarak SPSS ortamında analiz edilmiştir.

Geçerlik- güvenilirlik çalışmaları

Frankel & Wallen (2006:317)'e göre korelasyonel araştırmalarda olabildiğince seçkisiz yöntemleri kullanarak örneklem seçilmelidir. Araştırma dış geçerliğinin sağlanması açısından ve korelasyon katsayısının anlamlılığının örneklem büyüklüğünden etkilendiği göz önüne alınarak bu araştırmalarda her bir gözenekteki örneklem büyüklüğünün 30'un altına düşmemesine özen gösterilmelidir Bu araştırmada örneklem alma yoluna gidilmemiş olup, dersi alan tüm öğrenciler araştırmaya dâhil edilmiştir. Ancak verilerin elde edildiği hafta derse katılım yapan toplam 135 öğrenciden 121 öğrenci çalışmaya dâhil edilmiştir.

Frankel & Wallen (2006:319) araştırmalarda iç geçerliğin sağlanması açısından tüm ölçümlerin mümkünse deneklerin benzer şekilde etkilenecekleri aynı ortamda yapılmasını önermektedir. Bu araştırmada veriler aynı ortamda araştırmacının kontrolü altında ve aynı anda elde edilmiştir. Her iki testin geçerlik ve güvenilirlik çalışması Cohen, Manion & Morrison (2007:506)'da formüle edilen Cronbach Alpha iç tutarlık katsayısı tekniği ile hesaplanmıştır.

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



Dönem sonu sınavı sonuçlarına dayalı olarak ise $K-R_{21}$ hesaplanmıştır. Yapılan üç farklı yöntemle elde edilmiş puanlar arası ilişkiye dayalı olarak hesaplanan Cronbach Alpha katsayıları sırası ile Akran-Öğretmen: 0,95; Akran-Kendi: 0,98; Öğretmen-Kendi: 0,95; Genelde 0,97 ve final değerlendirmesine dayalı hesaplanan iç tutarlık katsayısı ise $K-R_{21} = 0,78$. olarak hesaplanmıştır.

Analiz teknikleri

İşe koşulan değerlendirme teknikleri ile elde edilen puanlar arasındaki farkın manidarlığı ilişkisiz örneklem t-testi tekniği analiz edilmiştir. Aynı şekilde değerlendirme teknikleri ile elde edilen puanlar arasındaki ilişki Pearson Correlation, Kendall's tau_b ve Spearman's rho teknikleri ile analiz edilmiştir. İlişkisiz örneklem için t testi, iki ilişkisiz örneklem ortalamaları arasındaki farkın manidar olup olmadığını test etmek için kullanılır (Büyüköztürk, Çokluk ve Köklü, 2012: 39). Heiman (1996)'a göre bir ilişkinin gücü, iki değişken arasındaki ilişkinin miktar ya da derecesi anlamına gelir. Kuvvetli ilişkilerde, bir X değerine bağlı olarak daha tutarlı bir Y değeri veya bir birine yakın Y değerine rastlanmaktadır (Akt. Büyüköztürk, Çokluk ve Köklü, 2012: 86).

Richardson (1996), Capraro & Capraro'a (2002) göre etki büyüklüğünün hesaplanmasında r^2 , adjusted R^2 , η^2 , ω^2 , Cramer's V , Kendall's W , Cohen's d , and Eta gibi birkaç farklı teknik vardır (Akt. Cohen, Manion & Morrison, 2007). Bu çalışmada işe koşulan puanlama yaklaşımlarına göre ortalama puanların etki büyüklüklerinin karşılaştırılması için sık kullanılan bir teknik olarak Cohen's d ve eta-kare(η^2) korelasyon katsayıları hesaplanmıştır. Böylece cinsiyetin değerlendirme yaklaşımları üzerinde ne düzeyde etkili olduğuna bakılmıştır. Etki büyüklüğü (genişliği) indeksi olan eta-kare(η^2), test puanlarındaki varyansın ne kadarının bağımsız değişkene ya da grup değişkenine bağlı olduğuna ilişkin yorum yapma imkânı verir. Aynı şekilde standardize edilmiş etki büyüklüğü indeksi olan d değeri karşılaştırılan ortalamaların bir birinden kaç standart sapma uzaklaştığına dair yorumlama imkânı sağlar. 0 ile 1 arasında değişen etki değerinde (bazı yaklaşımlarda 1 den büyük olabilir) işaretine bakılmaksızın Cohen's d değeri anlamlandırılır. Buna göre; 0,00 - 0,20 zayıf etki; 0,21- 0,50 küçük etki; 0,51-,1,00 orta etki; > 1,00 kuvvetli etkidir (Cohen, Manion & Morrison, 2007:521). Eta-kare değerinin ise, 0,01; 0,06 ve 0,14 olması sırası ile küçük (small), orta (medium) ve geniş (large) etki büyüklüğü olarak yorumlanır. Cohen's d değeri potansiyel olarak “-∞”ile “+∞” ; η^2 ise 0 ile 1 arasında değerler alabilir (Cohen, Manion & Morrison, 2007; Büyüköztürk, Çokluk ve Köklü, 2012).

İşe koşulan ölçme teknikleri ile elde edilen ölçüm sonuçlarının dönem sonu başarıyı ne düzeyde yordadığına ilişkin olarak regresyon analizi yapılmıştır. Cohen, Manion & Morrison (2007) ve Howell (1987)'e göre regresyon, bir değişkene ilişkin ölçümlerin grup ortalamasına doğru çekilmesidir. Regresyon analizi ise, aralarında ilişki olan iki ya da daha fazla değişkenin birinin bağımlı değişken, diğerlerinin bağımsız değişkenler olarak ayrımı ile aralarındaki ilişkinin bir matematiksel eşitlik ile açıklanması sürecini anlatır (Akt. Büyüköztürk, Çokluk ve Köklü, 2012: 122).

BULGULAR

Bu bölümde araştırma problemi alt problemler altında analiz edilerek açıklanmıştır.

Alt Problem 1. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçları değişmekte midir?

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



Tablo 1. Puanlama Yaklaşımlarına Göre Ölçüm Sonuçları Ortalamalarının Karşılaştırılması

	N	\bar{X}	S	sd	t	p
Akran Değ.	121	58,2149	16,31217	120		
Öğretmen Değ.	121	62,0992	20,69074	120	-4,230	,000
Akran Değ.	121	58,2149	16,31217	120		
Kendi Kendini Değ.	121	61,1983	16,09898	120	-5,017	,000
Öğretmen Değ.	121	62,0992	20,69074	120		
Kendi Kendini Değ.	121	61,1983	16,09898	120	1,104	,272

Tablo 1 bulgularına göre akran değerlendirme-öğretmen değerlendirme $t(120) = -4,230$, $p < .05$ ve akran değerlendirme-kendi kendini değerlendirme $t(120) = -5,017$, $p < .05$ ortalamaları arasında ikinciler lehine anlamlı bir farklılık bulunmaktadır. Farkın kaynağı olarak akran değerlendirme puanları ortalamasının ($\bar{X} : 58,2149$) öğretmen değerlendirme ($\bar{X} : 62,0992$) ve kendi kendini değerlendirme ($\bar{X} : 61,1983$) ortalamalarından düşük olması oluşturmaktadır. Buna karşın öğretmen ve kendi kendini değerlendirme puanları arasında öğretmen değerlendirmesi lehine bir fark görülmesine karşın bu fark iki değerlendirme türü arasında puanların ortalamasına anlamlı bir farklılık yaratmamaktadır $t(120) = 1,104$, $p > .05$.

Alt Problem 2. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımları göre ölçme sonuçları arasında anlamlı bir ilişki var mıdır?

Tablo 2. Puanlama Yaklaşımlarına Göre Ölçüm Sonuçları Arası İlişki

	N	Spearman's rho	Kendall's tau_b	Sig.
Akran Değ. & Öğretmen Değ.	121	,853	,685	,000
Akran Değ. & Kendi Kendini Değ.	121	,897	,746	,000
Öğretmen Değ. & Kendi Kendini Değ.	121	,880	,732	,000

Tablo 2 bulgularına göre işe koşulan puanlama yaklaşımları arası hesaplanan korelasyon katsayıları yüksektir. Buna göre her üç tür puanlamalar arasında anlamlı ($p < ,01$) yüksek bir ilişki vardır (tau_b ve $\text{rho} \geq ,70$).

Alt Problem 3. Cinsiyete göre öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçları değişmekte midir?

Cinsiyet değişkenine göre ölme tekniklerinden alınan notların ortalama, standart sapma, F testi ve t testi sonuçları tablo 3 de verilmiştir.

Tablo.3. Cinsiyet Değişkenine Göre Ölçme Tekniklerinden Alınan Notların Ortalama, Standart Sapma, T Testi Sonuçları

	Cinsiyet	N	\bar{X}	S	sd	t	p
Akran Değ.	Erkek	35	51,4571	15,85661	119	-3,003	,003
	Kız	86	60,9651	15,76776			
Öğretmen Değ.	Erkek	35	54,2857	22,62686	119	-2,720	,008
	Kız	86	65,2791	19,08473			
Kendi Kendini Değ.	Erkek	35	55,0571	16,25521	119	-2,749	,007
	Kız	86	63,6977	15,43874			

Tablo 3 bulgularına göre, cinsiyet değişkeni açısından üç farklı not verme tekniğine göre verilen puanlar arasında kızlar lehine olan fark istatistiksel olarak anlamlıdır. [(Akran

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



değerlendirme $t(119) = -3,003$, $p < .05$; öğretmen değerlendirme $t(119) = -2,720$, $p < .05$; kendi kendini değerlendirme $t(119) = -2,749$, $p < .05$). Tablo 3 bulgularında üç farklı not verme tekniğinde de kızların başarı ortalamaları erkeklerin başarı ortalamalarından daha yüksektir.

Alt Problem 4. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçlarında cinsiyetin etkisi nedir?

Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçlarında cinsiyetin etkisi Etki Büyüklüğü Değerleri Cohen's d ve eta-kare(η^2) katsayıları ile analiz edilip açıklanmıştır. Değerlendirme yaklaşımlarına göre ortalama puanların etki büyüklüklerinin karşılaştırılması için sık kullanılan bir teknik olarak eta-kare(η^2) korelasyon katsayıları hesaplanmıştır. Böylece cinsiyetin değerlendirme yaklaşımları üzerinde ne düzeyde etkili olduğuna bakılmıştır. Aynı şekilde standardize edilmiş etki büyüklüğü indeksi olan Cohen's d değeri karşılaştırılan ortalamaların bir birinden kaç standart sapma uzaklaştığına imkân sağlar. Hesaplanan etki büyüklüklerine dayalı analizler aşağıda açıklanmıştır.

- Tablo 3 bulgularına göre akran değerlendirme puanları cinsiyete göre anlamlı bir farklılık göstermektedir. Kızların akran değerlendirme puanları (\bar{X} :60,9651), erkeklerin puanlarından (\bar{X} :51,4571) daha yüksektir. Bu bulgu, kızların akran değerlendirmeye dayalı başarı puanlarının erkeklere göre daha yüksek olduğunu göstermektedir. Hesaplanan η^2 değeri 0,0084 "**küçük**" tür. Buna göre akran değerlendirme puanlarında gözlenen varyansın yaklaşık 0,008'incinsiyete bağlı olduğu ifade edilebilir. Diğer yandan hesaplanan Cohen's d değeri 0,6020 ile "**orta**" dır. Bu sonuç kız ve erkek öğrencilerin akran değerlendirme puanları arasındaki farkın 0,60 standart sapma kadar olduğunu gösterir.
- Tablo 3 bulgularına göre öğretmen değerlendirme puanları cinsiyete göre anlamlı bir farklılık göstermektedir. Kızların öğretmen değerlendirme puanları (\bar{X} : 65,2791), erkeklerin puanlarından (\bar{X} : 54,2857) daha yüksektir. Bu bulgu, kızların öğretmen değerlendirmeye dayalı başarı puanlarının erkeklere göre daha başarılı olduğunu göstermektedir. Hesaplanan η^2 değeri 0,0585 "**orta**" dır. Buna göre öğretmen değerlendirme puanlarında gözlenen varyansın yaklaşık 0,06'incinsiyete bağlı olduğu ifade edilebilir. Diğer yandan hesaplanan Cohen's d değeri 0,5453 "**orta**" dır. Bu sonuç kız ve erkek öğrencilerin öğretmen değerlendirme puanları arasındaki farkın 0,54 standart sapma kadar olduğunu gösterir.
- Tablo 3 bulgularına göre kendi kendini değerlendirme puanları cinsiyete göre anlamlı bir farklılık göstermektedir. Kızların kendi kendini değerlendirme puanları (\bar{X} : 63,6977), erkeklerin puanlarından (\bar{X} : 55,0571) daha yüksektir. Bu bulgu, kızların kendi kendini değerlendirmeye dayalı başarı puanlarının erkeklere göre daha başarılı olduğunu göstermektedir. Hesaplanan η^2 değeri 0,0597 "**orta**" dır. Buna göre kendi kendini değerlendirme puanlarında gözlenen varyansın yaklaşık 0,06'nın cinsiyete bağlı olduğu ifade edilebilir. Diğer yandan hesaplanan Cohen's d değeri 0,5511 "**orta**" dır. Bu sonuç kız ve erkek öğrencilerin kendi kendini değerlendirme puanları arasındaki farkın 0,55 standart sapma kadar olduğunu gösterir.

Ara sınav sonuçların final sınav sonuçlarını yordayıcılığı

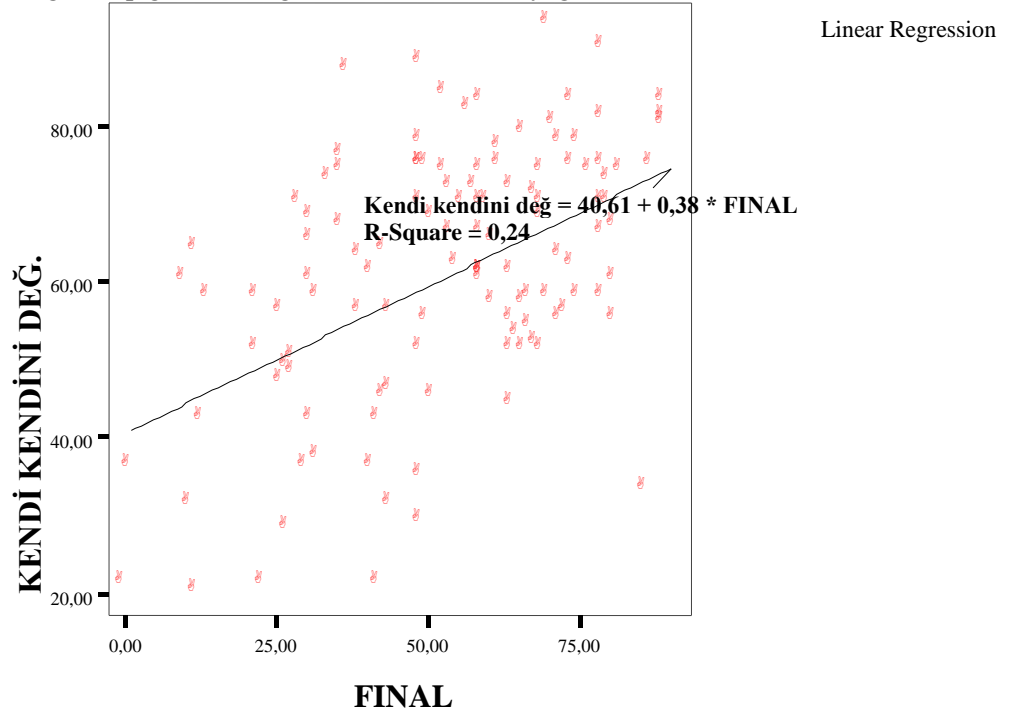
Alt Problem 5. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçlarının dönem sonu başarıyı yordama gücü nedir?

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



5.1. Ara sınavda işe koşulan kendi kendini değerlendirme sonuçlarının final sınavı sonuçları ile doğrusal bir özellik gösterip göstermediğine ilişkin veriler Diyagram 1’de verilmiştir.



Diyagram 1. Kendi Kendini Değerlendirme Ara Sınav Sonuçları ve Final Sınav Sonuçları Saçılma Diyagramı

Diyagram 1 incelemesinde kendi kendine değerlendirmeye dayalı ara sınav sonuçları ile final başarısı arasında yüksek oranda doğrusal bir ilişki olduğunu söylenemez. Deneklerin iki değişkene ilişkin değerlerini gösteren noktalar, bir doğru etrafında toplanmamaktadır.

Tablo 4. Kendi Kendine Değerlendirme Sonuçları ile Final Sınavı Regrasyon Analizi Sonuçları

Değişken	B	Standart Hata _B	β	T	p
Sabit	14,973	6,599		2,269	,025
Kendi Kendini Değ.	,647	,104	,494	6,199	,000
R=,494		R ² =,244		F _(1,119) = 38,425, p=.000	

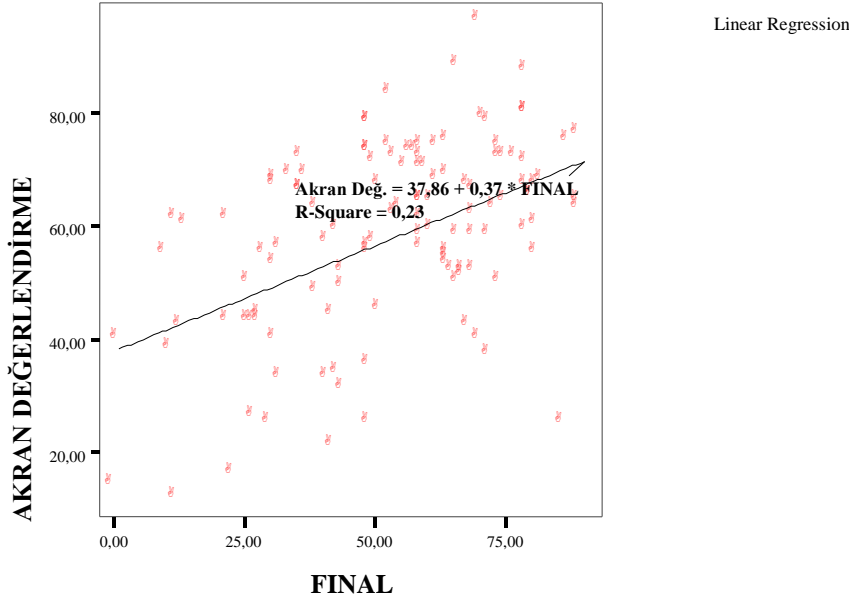
Tablo 4’de verilen analiz sonuçlarına göre yordayıcı değişken (kendi kendine değerlendirme) ile yordanan değişken (final başarısı) arasındaki ilişkinin miktarının, R ,494(a), R Square ,244; açıklanan varyansın, F(1,119) = 38,425, ve anlamlılık düzeyinin p<,000(a) olduğu görülmektedir. Bu sonuçlara göre öğrencilerin final sınav sonuçlarına ilişkin varyansın %24’nün kendi kendini değerlendirme sonuçları ile açıklandığı ifade edilebilir.

5.2. Ara sınavda işe koşulan akran değerlendirme sonuçlarının final sınavı sonuçları ile doğrusal bir özellik gösterip göstermediğine ilişkin veriler Diyagram 2’de verilmiştir.

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014





Diyagram 2. Akran Değerlendirme Ara Sınav Sonuçları ve Final Sınav Sonuçları Saçılma Diyagramı

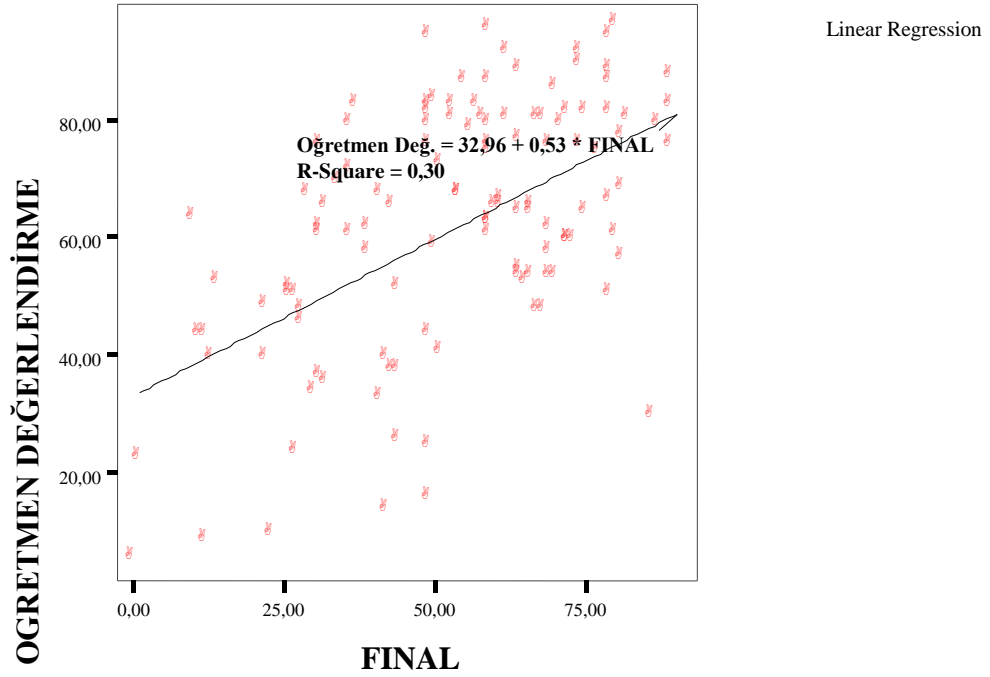
Diyagram 2 incelemesinde akran değerlendirmeye dayalı ara sınav sonuçları ile final başarısı arasında yüksek oranda doğrusal bir ilişki olduğunu söylenemez. Deneklerin iki değişkene ilişkin değerlerini gösteren noktalar, bir doğru etrafında toplanmamaktadır.

Tablo 5. Akran Değerlendirme Sonuçları ile Final Sınavı Regrasyon Analizi Sonuçları

Değişken	B	Standart Hata _B	β	T	p
Sabit	18,290	6,270		2,917	,004
Akran	,623	,104	,482	6,003	,000
R=,482		R ² =,232		F _(1,119) = 36,039, p=.000	

Tablo 5’de verilen analiz sonuçlarına göre yordayıcı değişken (akran değerlendirmeye dayalı ara sınav sonuçları) yordanan değişken (final başarısı) arasındaki ilişkinin miktarının, R ,482(a), R Square ,232; açıklanan varyansın, F(1,119) = 36,039, ve anlamlılık düzeyinin p<,000(a) olduğu görülmektedir. Bu sonuçlara göre öğrencilerin final sınav sonuçlarına ilişkin varyansın %23’nün akran değerlendirme sonuçları ile açıklandığı ifade edilebilir.

5.3. Ara sınavda işe koşulan öğretmen değerlendirme sonuçlarının final sınavı sonuçlarının doğrusal bir özellik gösterip göstermediğine ilişkin veriler Diyagram 3’de verilmiştir.



Diyagram 3. Öğretmen Değerlendirme Ara Sınav Sonuçları ve Final Sınav Sonuçları Saçılma Diyagramı

Diyagram 3 incelemesinde öğretmen değerlendirmeye dayalı ara sınav sonuçları ile final başarısı arasında yüksek oranda doğrusal bir ilişki olduğunu söylenemez. Deneklerin iki değişkene ilişkin değerlerini gösteren noktalar, bir doğru etrafında toplanmamaktadır.

Tablo 6. Öğretmen Değerlendirme Sonuçları ile Final Sınavı Regrasyon Analizi Sonuçları

Değişken	B	Standart Hata _B	β	T	p
Sabit	20,139	5,125		3,930	,000
Öğretmen	,554	,078	,544	7,074	,000
R=,544		R ² =,296		F _(1,119) = 50,034, p=.000	

Tablo 6'de verilen analiz sonuçlarına göre yordayıcı değişken (öğretmen değerlendirme) ile yordanan değişken (final başarısı) arasındaki ilişkinin miktarının, R ,544(a), R Square ,296; açıklanan varyansın, F(1,119) = 50,034, ve anlamlılık düzeyinin p<,000(a) olduğu görülmektedir. Bu sonuçlara göre öğrencilerin final sınav sonuçlarına ilişkin varyansın %30'nun öğretmen değerlendirme sonuçları ile açıklandığı ifade edilebilir.

Sonuç olarak, üç farklı yaklaşıma dayalı olarak yapılan ara sınav sonuçlarının final sınavını ilişkin yordayıcılığı üç farklı puanlama yaklaşımlarında sırası ile akran değerlendirmenin %23, kendi kendini değerlendirmenin %24 ve öğretmen değerlendirmesinin %30 oranında açıkladığı görülmüştür. Bunlar içerisinde final sınavına ilişkin yordayıcılık değerlerinin en yüksek oranda öğretmen değerlendirme yaklaşımında olduğu görülmektedir.

SONUÇ VE TARTIŞMA

Araştırma bulgularına göre belirlenen alt problemlere ilişkin olarak aşağıdaki sonuçlara ulaşılmıştır:

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



1. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçları arasındaki fark anlamlıdır. Bu durum Tablo 1 değerlerine göre, öğretmen değerlendirme-kendi kendini değerlendirme puanı ortalamaları arasında istatistiksel olarak anlamlı bir fark görülmemesine karşın, akran değerlendirme-öğretmen değerlendirme puanları ile akran değerlendirme-kendi kendini değerlendirme puanları ortalamaları arasında ikinciler lehine olan farktan kaynaklanmaktadır.
2. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımları göre ölçme sonuçları arasında anlamlı yüksek bir ilişki vardır.
3. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçlarında kızlar lehine anlamlı bir fark vardır.
4. Akran değerlendirme puanlarında gözlenen varyansın yaklaşık 0,008'nincinsiyete bağlı olduğu ifade edilebilir. Kız ve erkek öğrencilerin akran değerlendirme puanları arasındaki fark ,60 standart sapma kadardır.
5. Öğretmen değerlendirme puanlarında gözlenen varyansın yaklaşık 0,06'nın cinsiyete bağlı olduğu ifade edilebilir. Kız ve erkek öğrencilerin öğretmen değerlendirme puanları arasındaki fark ,54 standart sapma kadardır.
6. Kendi kendini değerlendirme puanlarında gözlenen varyansın yaklaşık 0,06'nın cinsiyete bağlı olduğu ifade edilebilir. kız ve erkek öğrencilerin kendi kendini değerlendirme puanları arasındaki fark ,55 standart sapma kadardır.
7. Öğretmen değerlendirme, akran değerlendirme ve kendi kendini değerlendirme yaklaşımlarına göre ölçme sonuçlarının dönem sonu başarıyı yordama gücüne ilişkin olarak şu sonuçlara ulaşılmıştır;

Kendi kendini değerlendirmeye dayalı ara sınav sonuçları ile final başarısı arasında yüksek oranda doğrusal bir ilişki olduğunu söylenemez. Tablo 4'de verilen analiz sonuçları yordayıcı değişken (kendi kendini değerlendirme) ile yordanan değişken (final başarısı) arasındaki ilişkinin miktarının, $R = ,494(a)$, $R \text{ Square} = ,244$; açıklanan varyansın, $F(1,119) = 38,425$, ve anlamlılık düzeyinin $p < ,000(a)$ olduğu görülmektedir. Bu sonuçlara göre öğrencilerin final sınav sonuçlarına ilişkin varyansın %24'nün kendi kendini değerlendirme sonuçları ile açıklandığı ifade edilebilir.

Akran değerlendirmeye dayalı ara sınav sonuçları ile final başarısı arasında yüksek oranda doğrusal bir ilişki olduğunu söylenemez. Tablo 5'de verilen analiz sonuçları yordayıcı değişken (akran değerlendirmeye dayalı ara sınav sonuçları) yordanan değişken (final başarısı) arasındaki ilişkinin miktarının, $R = ,482(a)$, $R \text{ Square} = ,232$; açıklanan varyansın, $F(1,119) = 36,039$, ve anlamlılık düzeyinin $p < ,000(a)$ olduğu görülmektedir. Bu sonuçlara göre öğrencilerin final sınav sonuçlarına ilişkin varyansın %23'nün akran değerlendirme sonuçları ile açıklandığı ifade edilebilir.

Öğretmen değerlendirmeye dayalı ara sınav sonuçları ile final başarısı arasında yüksek oranda doğrusal bir ilişki olduğunu söylenemez. Tablo 6'de verilen analiz sonuçları yordayıcı değişken (öğretmen değerlendirme) ile yordanan değişken (final başarısı) arasındaki ilişkinin miktarının, $R = ,544(a)$, $R \text{ Square} = ,296$; açıklanan varyansın, $F(1,119) = 50,034$, ve anlamlılık düzeyinin $p < ,000(a)$ olduğu görülmektedir. Bu sonuçlara göre öğrencilerin final sınav sonuçlarına ilişkin varyansın %30'nun öğretmen değerlendirme sonuçları ile açıklandığı ifade edilebilir.

Bu sonuçlara göre kullanılan puanlama yöntemleri arası ilişkinin yüksek olması her bir farklı puanlama tekniğinde de öğrencilerin ara sınav sonuçları arasındaki tutarlığın yüksek

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



olduğunu göstermektedir. Ortalamalar arasındaki farkın anlamlı olması işe koşulan tekniklerle yapılan puanlamaların farklılaştığını göstermektedir. Üç farklı puanlama sonuçlarında kızlar lehine anlamlı fark olması testin ölçtüğü özellikleri aynı duyarlılıkla, tutarlılıkla ve kararlılıkla ölçtüğünün önemli bir göstergesidir. Araştırmada öğretmen değerlendirme puanlarının diğer tekniklerle yapılan puanlamaya göre daha yüksek çıkması öğrencilerin akran ve kendi sınav kâğıtlarını puanlamalarına göre öğretmenlerin öğrencilerine toleranslı yaklaşımda bulunmuş olmalarının ya da mesleki eğitim ve deneyimine dayalı olarak daha nesnel puanlama yapmış olmasının bir göstergesi olabilir. Bu bulgu Sluijsmans ve diğerleri (1999) ve Sluijsmans ve diğerleri,(2003) bulgularının “Öğrenciler değerlendirme kriterlerini akranlarının çalışmasındaki kaliteden ziyade aralarındaki arkadaşlık ilişkisine göre belirleyebilirler, akranlarını değerlendirme sorumluluğunu almak istemeyebilirler ve bunu yapmak için kendilerine güvenmeyebilirler” görüşlerini destekler şekildedir. Aynı şekilde “öğrenciler akranlarını değerlendirirken ve değerlendirilirken kaygı duydukları” (Cartney, 2010) ve Koç (2011) bulgularında “akran değerlendirme sürecinde, öğretmen adaylarının kendilerini baskı altında hissettiklerini ve değerlendirirken nesnel değerlendirememeye kaygısı yaşadıkları” bulgularını destekleyen bir durumdan da kaynaklanmış olabilir. Aynı sonuç genelde akran değerlendirme ve kendi kendini değerlendirmeye olan nesnel olmayan yüzeysel bir yaklaşımla daha yüksek puanların verileceği eleştirilerini desteklememesi bakımından anlamlıdır.

Ölçme sonuçlarının taşınması gereken önemli niteliklerden birisinin yordayıcılık gücünün yüksek olmasıdır. Bu araştırmada üç farklı puanlama yaklaşıma dayalı olarak yapılan ara sınav sonuçlarının final sınavına ilişkin yordayıcılığı akran değerlendirmeye %23, kendi kendine değerlendirmeye %24 ve öğretmen değerlendirmesi ile %30 oranında açıkladığı görülmüştür. Bu sonuçlar Büyüköztürk (2004), Çetin ve Mahir (2006), Gülleroğlu (2005), Deaton ve Schutz (2001), Karakaya ve Tavşancıl (2008), Kutlu ve Karakaya (2003), Önen (2003), Önen ve Demirtaşlı (2004), Qualls ve Ansley (1995), Zwick ve Sklar (2005) bulguları ile paralellik göstermektedir. Araştırmada final sınavına ilişkin yordayıcılık değerlerinin en yüksek oranda öğretmen değerlendirme yaklaşımında olduğu görülmektedir. Bununla birlikte üç farklı puanlama yaklaşımı ile elde edilen ara sınav sonuçları ile final başarısı arasında yüksek oranda doğrusal bir ilişki olduğu söylenemez. Bu bulgu Zırhlı ve Atlı (2011) bulguları ile paralellik göstermektedir. Bu sonuca ara sınav ve dönem sonu sınavı kapsamlarının ara sınav lehine daha sınırlı olması ara sınav başarı ortalamasının daha yüksek çıkmasında önemli bir etken olabilir.

ÖNERİLER

Ölçmenin genel olarak tanımlanan sorunları *temsil, teklik ve anlamlılık* olarak ifade edilmektedir (Turgut, 1993; Taşdemir, 2011). Akran değerlendirme, kendi kendini değerlendirme, öğretmen değerlendirme ve diğer ölçme ve değerlendirme tekniklerinin dayadığı temel felsefeye uygun olarak ilkelerine uygun olarak işe koşulması şüphesiz olarak *temsil* sorunun üstesinden gelecektir. Alan yazın incelenmesi bu konuda uygulamalı araştırmalara ihtiyaç olduğunu göstermektedir. Literatürde yordama araştırmaları daha çok genel seçme sınavları, farklı disiplin alanları ve sınıflar arasındaki ilişki üzerine odaklanmış araştırmalar özelliğindedir. Özellikle ara sınav dönem sonu sınavları ve sınav teknikleri ve yordayıcılık üzerine araştırmalara ihtiyaç vardır. Araştırmacıların bu tür araştırmalara yönelmesi alana önemli katkı sağlayacaktır.

KAYNAKÇA

ABAZAOĞLU, İlkay, YILDIRIM, Osman ve Yılmaz YILDIZHAN (2014). “Türkiyenin Öğretmen Profili” *TURKISH STUDIES - International Periodical For The Languages, Literature and History of Turkish or Turkic* Volume 9/2 Winter 2014, www.turkishstudies.net, <http://dx.doi.org/10.7827/TurkishStudies.6491>, p. 1-20

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014



- BASKONUŞ, T. ve TAŞDEMİR, M. (2012). “İlköğretim Birinci Kademe Sosyal Bilgiler Programında (4. ve 5. sınıf) Yer Alan Alternatif Ölçme Değerlendirme Araç ve Yöntemlerinin Kullanılma ve Kullanılmama Durumları” *11. Ulusal Sınıf Öğretmenliği Eğitimi Sempozyumu*, 24-26 Mayıs, Rize
- BEST, WJ.,&KAHN,V.J. (2006). *Research in education* (11th ed.). Pearson Education Inc.
- BURKE, K. (1999). *How to assess authentic learning (3rd Ed)*, Arlington Heights, Illinois: Skyligh Professional Development.
- BÜYÜKÖZTÜRK, Ş. (2004). Predictors of academic achievement for elementary teacher education students in Turkey. *International Journal of Educational Reform*, 13(4), 388-402.
- BÜYÜKÖZTÜRK, Ş., KILIÇ-ÇAKMAK, E., AKKÜN, Ö.E., KARADENİZ, Ş. ve DEMİREL, F. ((2008). *Bilimsel araştırma yöntemleri*.(Geliştirilmiş 2.baskı), Pegem Akademi.
- CARTNEY, P. (2010). Exploring the use of peer assessment as a vehicle for closing the gap between feedback given and feedback used. *Assessment & Evaluation in Higher Education*, 35 (5), 551-564.
- CHEN, L., MANION, L. & MORRISON, K. (2007). *Research methods in education* (Sixth Edition).USA, Routledge
- CHUNG, M. (2000). The Development of Self-Regulated Learning. *Asian Pacific Review*, 1 (1), 55-66.
- CRONBACH, L. J. (1984). *Essentials of Psychological Testing (4th ed.)*. New York: Harper & Row Publishers.
- ÇELİKKALELİ, Ö. ve AKBAŞ, A. (2007). Sınıf Öğretmeni Adaylarının Fen Bilgisi Dersine Yönelik Tutumlarını Yordamada Fen Bilgisi Öğretimi Öz-Yeterlik İnançları. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 3(1), 21-34
- ÇEPNİ, S., BAYRAKÇEKEN, S., YILMAZ, A., YÜCEL, C., SEMERCİ, Ç., KÖSE, E., SEZGİN, F., DEMİRCİOĞLU, G. & GÜNDOĞDU, G. (2007). *Ölçme ve Değerlendirme*. Pegem A Yayıncılık, Ankara.
- ÇETİN, N., ve MAHİR, N. (2006). Genel Matematik Dersindeki Öğrenci Başarısı İle ÖSS Başarısı Arasındaki İlişki. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*,7(11), 37-46.
- DEATON, R. & SCHUTZ, G.J. (2001). Examining the predictive power of the ACT and high school GPA. Tennessee Higher Education Commission Southern Association of Institutional Research, Panama City Beach, Florida.
- DENİZ, Z.K. ve KELECİOĞLU, H. (2005). İlköğretim Başarı Ölçüleri İle Ortaöğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavı Arasındaki İlişkiler. *Ankara University, Journal of Faculty of Educational Sciences*, 38(2), 127-143
- DÖNDER, Ayşenur, ELALDI, Şenel ve Özlem Miraç ÖZKAYA (2012). “Öğretim Elemanlarının Üniversite Düzeyindeki Temel İngilizce Öğretiminde Kullanılabilecek Tamamlayıcı Ölçme Ve Değerlendirme Yöntemlerine İlişkin Görüşleri” *TURKISH STUDIES - International Periodical For The Languages, Literature and History of Turkish or Turkic* Volume 7/1 Winter 2012, www.turkishstudies.net,http://dx.doi.org/10.7827/TurkishStudies.2902,p.953-968

- FRANKEL, J.R. & WALLEN, N. E. (2006). *How to design and evaluate research in education*. (6.Baskı) New York: McGraw-Hill Book Company.
- GELBAL, S. ve KELECİOĞLU, H. (2007). Öğretmenlerin Ölçme ve Değerlendirme Yöntemleri Hakkındaki Yeterlik Algıları ve Karşılaştıkları Sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33, 135-145.
- GIELEN, S., PEETERS, E., DOCHY, F., ONGHENA, P., & STRUYVEN, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and Instruction*, 20, 304-315.
- GÖKYER, Necmi ve Fulya ÖZER (2014). “Aday Öğretmenlerin Sınıf Yönetimi Yeterliliklerine İlişkin Algıları” *TURKISH STUDIES - International Periodical For The Languages, Literature and History of Turkish or Turkic* Volume 9/2 Winter 2014, www.turkishstudies.net, <http://dx.doi.org/10.7827/TurkishStudies.6163>.p. 691-712
- GÜLLEROĞLU, H. (2005). “Üniversite Öğrencilerinin Akademik Başarılarının Yordanmasına İlişkin Karşılaştırmalı Bir Araştırma.” *Yayınlanmamış doktora tezi*, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- KARAKAYA, İ., ve TAVŞANCIL, E. (2008). Yükseköğretime Öğrenci Seçme Sınavı (ÖSS)'nin Yordama Geçerliği. *Kuram ve Uygulamada Eğitim Bilimleri*.8(3), 987-1019.
- KAZU, Hilal ve Serkan ASLAN (2012) “Birleştirilmiş Sınıf Öğretmenlerinin Hayat Bilgisi Dersi Öğretim Programına Yönelik Görüşlerinin İncelenmesi (Nitel Bir Araştırma)” *TURKISH STUDIES - International Periodical For The Languages, Literature and History of Turkish or Turkic* Volume 7/2 Spring 2012, www.turkishstudies.net, <http://dx.doi.org/10.7827/TurkishStudies.3364>,p.693-706
- KOÇ, C. (2011). Sınıf Öğretmeni Adaylarının Öğretmenlik Uygulamasında Akran Değerlendirmeye İlişkin Görüşleri. *Kuram ve Uygulamada Eğitim Bilimleri*, 11(4), 1965-1989
- KUTLU Ö. ve KARAKAYA, İ. (2003). Ortaöğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavının Faktör Yapısına ve Yordama Gücüne İlişkin Bir Araştırma. *Eğitim Bilimleri ve Uygulama Dergisi*, 2 (4),209-223.
- MEB-URL- http://iogm.meb.gov.tr/files/size_ozel/olcme_ve_degerlendirme.pdf, adresinden 25.05.2013 tarihinde indirilmiştir.
- NOBLE, J. & SAWYER, R. (2002). *Predicting different levels of academic success in college using high school GPA and ACT Composite Score*. (ACT Research Report Series 2002-4), Iowa City, IA: ACT, Inc.
- ÖNEN, E., ve DEMİRTAŞLI, N. (2004). OKÖSYS Başarısı ve Fen Lisesi 1. Sınıftaki Akademik Başarıya İlişkin Bir Yordama Geçerliği Çalışması. *Kuram ve Uygulamada Eğitim Bilimleri*, 4(1): 126-145.
- ÖZAN, S. (2008). “Öz ve Akran Değerlendirmenin Temel İletişim Becerileri Üzerindeki Etkileri.” *Yayımlanmamış doktora tezi*, Dokuz Eylül Üniversitesi, İzmir.
- QUALLS, A.L. & ANSLEY, T.N. (1995). The Predictive relationship of ITBS and ITED to measures of academic success. *Educational and Psychological Measurement* 55: 485 - 498.
- SLUIJSMANS, D., & PRINS, F. (2006). A conceptual framework for intergrating peer assessment in teacher education. *Studies in Educational Evaluation*, 32, 6-22.

- SLUIJSMANS, D., DOCHY, F. & MOERKERKE, G. (1999). Creating a learning environment by using self-peer and co-assessment. *Learning Environment Research 1*, 293–319.
- SLUIJSMANS, M. A. D., GRUWEL, S. B., VAN MERRIENBOER, J. J. G., & BASTIAENS, T. J. (2003). The training of peer assessment skills to promote the development of reflection skills in teacher education. *Studies in Educational Evaluation*, 29, 23-42.
- TAŞDEMİR, M. (2011). *Eğitimde ölçme ve değerlendirme*. Sohbet Kitabevi Yayınları. Kırşehir.
- TEKİN, H. (2007). *Eğitimde ölçme değerlendirme* (18.basım). Ankara: Yargı Yayınevi
- TOPPING, K. (1998). Peer assessment between students in colleges and universities. *Review Of Educational Research*, 68, 249–276.
- TURGUT, F. (1993). *Eğitimde, Ölçme Değerlendirme* (9.Basım). Ankara: Saydam Matbaacılık.
- VICKERMAN, P. (2009). Student perspectives on formative peer assessment: An attempt to deepen learning? *Assessment & Evaluation in Higher Education*, 34 (2), 221-230.
- WILLEY, K., & GARDNER, A. (2010). Investigating the capacity of self and peer assessment activities to engage students and promote learning. *European Journal of Engineering Education*, 35 (4), 429-443.
- YEŞİLYURT, Etem (2012) “Fen ve Teknoloji Dersinde Kullanılan Ölçmedeğerlendirme Yöntemleri ve Karşılaşılan Güçlükler” *TURKISH STUDIES - International Periodical For The Languages, Literature and History of Turkish or Turkic* Volume 7/2 Spring 2012, www.turkishstudies.net, <http://dx.doi.org/10.7827/TurkishStudies.3142>, p.1183-1205
- ZIRHLIOĞLU, G. ve ATLI, M. (2011). Beden Eğitimi Bölümü Özel Yetenek Sınavı Puanlarının Akademik Başarı Üzerinde Yordama Geçerliği. *Eğitim ve Bilim*, 36(161),176-185.
- ZIMBICKI, D. (2007). *Examining the effects of alternative assessment on student motivation and self efficiency*. Unpublished doctoral dissertation, The Walden University, U.S.A.
- ZWICK, R. & SKLAR J. (2005). Predicting college grades and degree completion using high school grades and SAT scores: The Role of student ethnicity and first language. *American Educational Research Journal*, 42 (3), 439- 464.

Turkish Studies

International Periodical For the Languages, Literature and History of Turkish or Turkic
Volume 9/5 Spring 2014

