

PERFORMANS GÖREVİNİN DEĞERLENDİRİLMESİNE İLİŞKİN BİR ÜÇLEME: ÖZ, AKRAN VE ÖĞRETİM ÜYESİ DEĞERLENDİRMESİ

Yrd. Doç. Dr. Şeyma UYAR

Mehmet Akif Ersoy Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü.

seymayukselster@gmail.com

Yrd. Doç. Dr. Kenan DEMİR

Mehmet Akif Ersoy Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü.

kenandemirkfe@gmail.com

Arş. Gör. Burcu AKSEKİOĞLU

Mehmet Akif Ersoy Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü,

ÖZ

Bu çalışmada, ölçme ve değerlendirme dersinin vize sınavı kapsamında yapılan öğrenci performans görevlerinin öz, akran ve öğretim üyesi tarafından puanlandığı durumlar arasındaki farklılıklar incelenmiştir. Araştırma, Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Eğitim Fakültesi Matematik, Okul Öncesi ve Türkçe Öğretmenliği bölümünde öğrenim gören toplam 157 öğrenci ile gerçekleştirilmiştir. Çalışmada öğrencilerden ders sürecinde işlenen kavramların gerçek yaşamdaki karşılıklarını şiir, karikatür, bulmaca, müzik, dans, şarkı sözü, drama oyunu, film, hikâye gibi farklı ve yaratıcı yollarla ifade etmeleri ve performans görevi hazırlayarak sunmaları istenmiştir. Puanlama sürecinde öğrencilerden kendilerinin ve arkadaşlarının performans görevlerini önce genel olarak, sonra dereceli puanlama anahtarı (rubric) kullanarak puanlamaları istenmiştir. Daha sonra da dersi yürüten öğretim üyesi aynı puanlama anahtarına göre performans görevlerini puanlamıştır. Kısmen, öğrencilerle birlikte geliştirilen dereceli puanlama anahtarı “yaratıcılık”, “öğrenme”, “anlatım” ve “emek-çaba” ölçütleri ile “usta”, “kalfa”, “çırak” ve “acemi” derecelerinden oluşturulmuştur. Bu çalışma sonunda, üç ayrı puanlayıcıdan elde edilen beş ayrı puan arasındaki ilişkiler incelenmiştir. Yapılan bağımlı gruplar t testi sonuçlarına göre akranların verdiği genel ve rubric puanlar arasında anlamlı bir fark olduğu, akranların rubric kullanarak verdikleri puanların daha düşük olduğu bulunmuştur. Verilen puanların cinsiyete göre anlamlı bir farklılık gösterip göstermediğine ilişkin bağımsız gruplar t testi sonuçlarına göre kız öğrencilerin kendilerine ve akranlarına daha yüksek puan verdikleri görülmüştür. Öz, akran ve öğretim üyesinin verdiği puanlar arasındaki ilişkiler ise tek faktörlü varyans analizi ile incelenmiş ve puanlar arasında anlamlı bir farklılık olduğu tespit edilmiştir.

Anahtar Sözcükler: Öz değerlendirme, akran değerlendirme, öğretim üyesi değerlendirmesi, dereceli puanlama anahtarı (rubric)

A TRIPLET ON PERFORMANCE-BASED ASSESSMENT: SELF, PEER AND RATER ASSESSMENTS

ABSTRACT

In this study, the differences among the processes of scoring of performance assignments which are done for the mid-term exam of the lesson, Measurement and Assessment, and by students, peers and raters are examined. The sample group of this study consists of 150 students from the division of Elementary Mathematics, Preschool Education and the department of Turkish language Teaching at Mehmet Akif Ersoy University. During the study, it is asked that the students use the equivalents of the concepts, which have been mentioned in the lesson process, by various creative ways such as poem, caricature, puzzle, music, dance, lyrics, drama, film, and story and present after preparing a performance assignment. It is asked that the students first score by regarding the whole and then by use of rubrics. Afterwards, the raters, who are the instructors at the departments of the students in the sample group, score performance assignments by the way the students have used. The rubric, which have partly been formed with the students, is consisted of the criterion of, ‘‘creativity’’, ‘‘learning’’, ‘‘expression’’, ‘‘effort-extersion’’ and levels of ‘‘master’’, ‘‘master-builder’’, ‘‘apprentice’’, ‘‘inexperienced’’. At the end of this study, the relationships among five different scorers formed from three different scorers are examined. It is reached that peers give lower scores by use of rubric and there is a remarkable difference between scoring the whole and scoring by means of rubric by the peers in respect of the results of applied paired sample t-test. It is seen that female students give higher scores to themselves and their peers in respect of the results of the Independent Samples t test on whether given scores differ remarkably according to gender. The relationships among the scores given by the students, the peers and the raters are examined by the one way analysis of variance and it is determined that there is a significant difference among these scores.

Keywords: *Self-assessment, peer assessment, rater assessment, rubric;*

1.GİRİŞ

Öğrenciler girdikleri sınavlarda hatalarının ya da eksikliklerinin nereden kaynaklandığına ilişkin geribildirim ihtiyacı duyarlar. Onların başarılarına, öğrenmelerin eve bilgi düzeylerine ilişkin geribildirimde bulunmak için değerlendirme yöntemlerine başvururlar.

Değerlendirme, geleneksel yöntemlerde öğretmenin tek başına değerlendirici olarak rol aldığı biçimde karşımıza çıkmaktadır. Tek doğru cevabı olan çoktan seçmeli testlerde değerlendirmeler öğretmen tarafından yapılır ve puanlamanın tutarlı olduğu düşünülür (Matsuno, 2007). Ancak, okuduğunu anlama, problem çözme, eleştirel düşünme, araştırma yapmavb. üst düzey zihinsel becerileri belirleyebilmek amacıyla kullanılan performans görevi, drama, sunum, proje gibi alternatif ölçme araçlarını puanlamak ve değerlendirmek çok kolay değildir. Hatta öğretmenin yanlı davrandığı kanısını bile oluşturabilmektedir. Bu nedenle farklı değerlendirme yaklaşımlarına duyulan ihtiyaç artmaktadır (Orsmond, Merry ve Reiling, 1997, 2000; Parlak, 2010).

Son yıllarda ölçme ve değerlendirme sürecine öğrencilerin de dâhil edildiği öğrenci merkezli yaklaşım benimsenmeye başlanmıştır (Lewkowicz ve Moon, 1985). Boud ve Falchikov (1989), öğretmenlerin öğrenciler hakkında edindiği izlenimin sınırlı olabileceğini ve öğrencilerin kendi başarılarına ilişkin daha geniş bir bakış açısına sahip olduğunu belirtmektedir. Rogers (2003) ise demokratik eğitim anlayışlarında öğrencinin birey olarak tanımlanması gerektiğini savunmaktadır. Ona göre bireye doğrudan öğretmek mümkün değildir. Birey kendi sorumluluklarını aldığı anda, başkaları ile karşılaştırma yaptığı anda ve öğretimin merkezinde yer aldığı anda yani özgür olduğunda öğrenme gerçekleşebilir. Özgür (otantik) öğrenmede öğretim programı, öğretmen, eğitim ortamı ve öğrenen etkileşim içerisindedir Bu dört öge birlikte çalıştığında sistemin başarılı olacağı düşünülmektedir. Ayrıca böyle bir sistem öğrencileri eğitimin her aşamasında aktif kılabilirdir (Higgs, 1988).

Black ve Williams (1998)'e göre öğrencilerin sorumluluğunu arttırmak ölçme ve değerlendirme süreçlerine katılmasıyla mümkündür. Ona göre değerlendirme yöntemleri ve öğrencilerin sınıf içi birbirleri ile ilişkilidir. Değerlendirme öğrenmenin ne kadar geliştiğini belirlerken bir yandan öğrenmeyi gerçekleştirir. Bu nedenle öğrenmenin standartlarını yükseltmek ve güçlü öğrenmelerin gerçekleşmesini sağlamak için değerlendirmenin önem verildiği programların yapılması gerektiği düşünülmektedir (Cartney, 2010; Orsmond, Merry ve Reiling, 1997).

Öğrenciler ölçme değerlendirme sürecine kendilerini ve akranlarını puanladıkları öz ve akran değerlendirme yöntemleriyle katılabilmektedirler (Oscarson, 1989). Öz değerlendirme, öğrencilerin kendi bilgi ve becerilerini değerlendirmeye ve yargılamaya yönelik uygulama olarak tanımlanabilmektedir (Bailey, 1998). Brown ve Hudson (1998), öz değerlendirme yapıldığında öğrencinin motivasyonunun arttığını belirtmektedir. Yapılan bazı çalışmalar öz değerlendirmenin, öz düzenleme becerilerini ve öz yeterliği arttırdığını göstermiştir (Panadero, Tapia ve Huertas, 2012; Chung, 2000; Zimbicki 2007). Özan (2008) öz ve akran değerlendirmenin iletişim becerileri üzerinde olumlu etkiye sahip olduğunu ifade etmiştir.

Akran değerlendirme, öğrencilerin diğer grup üyelerinin çalışmalarını değerlendirdiği yaklaşımdır. Bu yöntem ile öğrenciler kendi çalışmalarını başka öğrencilerle karşılaştırma fırsatı bulabilmekte ve puanlamanın adil olup olmadığını görebilmektedir. Aynı zamanda öğrenme gerçekleşmektedir (Hanrahan ve Isaacs2001, ;Kennedy, 2006, Temizkan, 2009). Keaten ve Richardson (1993) öğrencilerin akranlarını değerlendirirken adil ve tutarlı davrandıkları koşulda, sorumluluk alma bakımından geliştiklerini ifade etmektedir. Ellis (2001) öz ve akran değerlendirmenin yansıtıcı öğrenmeye destek olduğunu belirtmektedir. Yurdabakan (2011) ise işbirlikli öğrenme etkinlikleri ve akran değerlendirme etkinliklerinin birlikte yapılması gerektiğini belirtmektedir.

Öğrenciler öz ve akran değerlendirmede puanlamayı genel izlenimle (ölçme aracı kullanmadan) ya da ölçütlerin kullanıldığı bir araçla (rubrik) yapabilmektedir (Panadero, Tapia ve Huertas (2012). Rubrik, bir ölçme işlemine ilişkin gözlemleri iyiden kötüye doğru derecelendirmeye yarayan değerlendirme ve puanlama aracıdır (Goodrich, 1997). Puanlama anahtarı, ölçütlerden ve performans düzeylerinden oluşan yapısı sayesinde puanlamadaki tutarsızlıkları azaltabilmektedir (Parlak, 2010). Popham (1997) dereceli puanlama anahtarının öğrencilerin yazılı kompozisyonlarının ve sözel sunumlarının değerlendirilmesinde rehber olabileceğini belirtmektedir.

Öğrenmede katkı sağlamasına rağmen öz ve akran değerlendirmede geçerlik ve güvenilirlik sorunları yaşanabilmektedir. Öğrenciler kendilerini ya da akranlarını değerlendirirken yanlı davranabilmektedir. Bu nedenle iki yaklaşım, değerlendirmede pek tercih edilmemektedir (Falchinov ve Goldfinch, 2000; Magin 2001). Akranların kendi aralarında verdiği puanların tutarlı olması yöntemin güvenilirliğine; öz, akran değerlendirme ve öğretmen puanlarının tutarlı olması ise geçerliğe hizmet etmektedir. Yapılan birçok çalışma genellikle kızların kendilerine ve kız akranlarına daha yüksek puan verdiklerini göstermektedir (Kaufman, Felder ve Fuller, 2000; Sherrard ve Raafat, 1994; Baker, 2008). Bunun yanında Mattheos ve ark. (2004) öz değerlendirmede cinsiyet farklılığının etkili olmadığını belirtmiştir. Matsuno (2009), öğretmen puanlarının öz ve akran değerlendirme puanlarına göre ne çok katı ne de yüksek olduğunu ifade etmiştir.

Yapılan çalışmalarda genellikle yazılı sınavların değerlendirildiği belirtilmiştir. Bir performans görevinin öz, akran ve öğretmen tarafından puanlandığı duruma ilişkin araştırmaya rastlanmamıştır. Bu nedenle ilgili çalışmada performans görevinin öz, akran ve öğretim üyesi tarafından puanlandığı durumda puanlar arasındaki ilişkiyi belirlemek amaçlanmıştır.

2.YÖNTEM

Araştırmanın Türü

Bu çalışmada, ölçme ve değerlendirme dersinin vize sınavı kapsamında yapılan öğrenci performans görevlerinin öz, akran ve öğretim üyesi tarafından puanlandığı durumlar arasındaki farklılıklar incelenmiştir. Bu yönüyle çalışma betimsel bir araştırma niteliğindedir.

Çalışma Grubu

Bu araştırma Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Matematik, Okul Öncesi ve Türkçe Öğretmenliği bölümünde öğrenim gören ve puanlama süreçlerinin hepsine dahil olan toplam 157 öğrenci ile yürütülmüştür. Verilerin normal dağılıma uygunluğunun sağlanması açısından uç değerler çıkartılarak

çalışma 150 öğrenci ile sınırlandırılmıştır. Öğrencilerin cinsiyete ve bölüme göre dağılımları Tablo 1’de verilmektedir.

Tablo 1: Öğrencilerin bölüme ve cinsiyete göre dağılımları

Cinsiyet	Bölüm			
	Matematik	Okul Öncesi	Türkçe	Toplam
Kız	16	59	38	113
Erkek	5	8	24	37
Toplam	21	67	62	150

Tablo 1 incelendiğinde araştırmaya 113 kız ve 37 erkek öğrencinin katıldığı görülmektedir. Matematik öğretmenliğinden 21, Okulöncesi öğretmenliği bölümünden 67 ve Türkçe öğretmenliği bölümünden ise toplamda 62 öğrenci çalışmaya dahil edilmiştir.

Veri Toplama Araçları

Çalışmada öğrencilerden ölçme ve değerlendirme ders sürecinde işlenen kavramların gerçek yaşamdaki karşılıklarını şiir, karikatür, bulmaca, müzik, dans, şarkı sözü, drama oyunu, film, hikâye gibi farklı ve yaratıcı yollarla ifade etmeleri ve performans görevi hazırlayarak sunmaları istenmiştir. Performans görevinin tamamlanması için öğrencilere yaklaşık 8 hafta süre verilmiştir. Performans görevinin puanlanması belli bir sıraya uygun olarak gerçekleştirilmiştir. Öncelikle her öğrenci kendi görevini bir ölçüt kullanmadan genel olarak puanlamıştır. Daha sonra her öğrenci, başka bölümlerde öğrenim görmesi ve birbirini tanımaması dikkate alınarak, akranlarını ölçüt kullanmadan puanlamıştır.

İkinci aşamada öğrenciye rubrik verilerek akranını puanlaması ve daha sonra aynı puanlama anahtarına göre kendi görevini puanlaması istenmiştir. Buna göre öz genel- akran genel, akran rubrik ve öz rubrik olmak üzere 4 ayrı puan elde edilmiştir. Puanlama sürecine en son dersin öğretim üyesi katılmıştır. Bu aşamada her öğrencinin puanlanmasında yalnızca rubrikten yararlanılmıştır. Sonuçta beş ayrı puan türü elde edilmiştir. Dereceli puanlama anahtarı için örnek bir gösterim Tablo 2’de yer almaktadır

Tablo 2: Dereceli puanlama anahtarı

Ölçütler	Dereceler			
	Usta(4)	Kalfa(3)	Çırac (2)	Acemi(1)
Yaratıcılık				
Öğrenme				
Anlatım				
Emek-Çaba				

Tablo 2’de yer alan puanlama anahtarı kısmen öğrencilerle birlikte geliştirilmiştir. Dereceli puanlama anahtarı “yaratıcılık”, “öğrenme”, “anlatım” ve “emek-çaba” ölçütlerinden ve “usta”, “kalfa”, “çırak” ve “acemi” olmak üzere dört dereceden oluşmaktadır. Öğrencilerden kendilerinin ve akranlarının görevini yaratıcılık, öğrenme ve anlatım bakımından derecelendirmeleri istenmiştir. Emek-çaba ölçütünde ise öğrenciden puanlamaya daha fazla katkı sağlaması açısından görüş alınmıştır. Bu kısımda derecelendirme yapılmamıştır.

Verilerin Analizi

Bu çalışma sonunda, üç ayrı puanlayıcıdan elde edilen beş ayrı puan arasındaki ilişkiler incelenmiştir. Öncelikle veriler analizlere uygunluk açısından değerlendirilmiştir. Verilerin normalliği basıklık ve çarpıklık katsayıları ve histogram grafikleri ile incelenmiştir. Bu aşamada verilerin normal dağıldığı kanısına varılmıştır.

Akranların verdiği genel ve rubrik puanları arasındaki ilişkileri belirlemek amacıyla parametrik yöntemlerden bağımlı gruplar t testi kullanılmıştır. Puanların cinsiyete göre değişiklik gösterip göstermediği durum bağımsız gruplar t testi ile incelenmiştir. Öz, akran ve öğretim üyesinin verdiği puanlar arasındaki ilişkiler ise tek faktörlü varyans analizi ile incelenmiştir. Analizler SPSS 20.0 paket programında gerçekleştirilmiştir.

3.BULGULAR

Bu bölümde, araştırmanın amacı doğrultusunda cevap aranan sorulara ilişkin ulaşılan bulgulara ve yorumlara yer verilmiştir.

Tablo 3 bağımlı gruplar t testi sonuçlarına göre öğrencilerin genel izlenimle verdikleri öz ve akran değerlendirme puanları ile rubriğe dayalı olarak verilen puanları arasında fark olup olmadığını göstermektedir.

Tablo 3: Öz ve akran değerlendirme puanlarına ilişkin bağımlı gruplar t testi sonuçları

	sd	N	Ortalama	t	p	r
Öz Değerlendirme (Genel-Rubrik)			74,173			
			74,440	-,808	.420	.604
Akran Değerlendirme (Genel-Rubrik)	149	150	75,05			
			70,96	8,901	.00	.663

Tablo 3'e göre rubrik kullanılmadığında ve kullanıldığında verilen öz değerlendirme puanları arasında fark olmadığı görülmektedir ($t(149)=.0808>.05$). Akranlarını, değerlendirme aracı kullanmadan ve rubriğe dayalı olarak değerlendirdikleri puanlar arasında anlamlı bir fark olduğu söylenebilir ($t(149)=8,901<.05$). Akranların rubrik kullanarak verdikleri puanların daha düşük olduğu görülmektedir. (Akran genel ort=75,05; akran rubrik=75,05).

Öz ve akran değerlendirme puanlarının cinsiyete göre değişimi Tablo 4'te verilmektedir.

Tablo 4: Öz ve akran değerlendirme puanlarının cinsiyete göre değişimine ilişkin bağımsız gruplar t testi sonuçları

		Cinsiyet	sd	N	Ort.	t	p
Öz Değerlendirme (Genel)	Kız	14	11	74,55	1,743		.083
	Erkek	8	3	73,03			
Öz Değerlendirme (Rubrik)	Kız	14	11	74.90	2.261		.025
	Erkek	8	3	73.03			
Akran Değerlendirme (Genel)	Kız	14	11	75,44	1,757		.081
	Erkek	8	3	73,86			
Akran Değerlendirme (Rubrik)	Kız	14	11	71,87	2,634		.009
	Erkek	8	3	68,19			

Tablo 4'e göre genel izlenimle yapılan öz değerlendirme puanlarının cinsiyet bakımından farklılaşmadığı söylenebilir ($t(148)=.083>.05$). Akran değerlendirme genel izlenimle yapıldığında cinsiyetin anlamlı bir fark yaratmadığı görülebilmektedir ($t(148)=.081<.05$). Rubrik kullanılarak yapılan öz ve akran değerlendirme puanları cinsiyete göre anlamlı bir farklılık göstermektedir. Kız öğrenciler hem kendilerine hem akranlarına daha yüksek puan verme eğilimindedir.

Öğrencilerin kendilerine ve akranlarına verdikleri puanlar ile öğretim üyesi puanları arasında yapılan ANOVA testine ilişkin sonuçlar Tablo 5'te verilmektedir.

Tablo 5. Öz, akran ve öğretim üyesi puanlarına ilişkin ANOVA sonuçları

		Kareler Toplamı	sd	F	p	Anlamlı Fark
Değerlendirme Puanları	Gruplar	909,888	2	13,863	.00	*Öz-Akran değ.
	Arası Gruplar içi	14669,70 8	447			*Öz değ.-Öğretim üyesi

Tablo 5'e göre öz, akran ve öğretim üyesi değerlendirme puanları arasında anlamlı bir fark olduğu söylenebilir ($F(2,447) = 13,863, p < .001$). Çoklu karşılaştırma testi sonuçlarına göre öz değerlendirme puanları ile akran değerlendirme ve öğretim üyesi puanları anlamlı şekilde farklıdır. Öz değerlendirme puanlarının ortalaması akran değerlendirme puanı ile öğretim üyesi puan ortalamalarından yüksektir. (Öz ortalama= 74.44, akran ort.=70.96). Öğretim üyesinin verdiği puanlara ilişkin ortalamayla akranların verdiği puanlar arasında çoklu karşılaştırma testi sonuçlarına göre anlamlı bir fark bulunmamaktadır.

4.TARTIŞMA, SONUÇ VE ÖNERİLER

Bu çalışma öğrencilerin öz, akran değerlendirmelerine ilişkin puanlar ile öğretim üyesi puanları arasındaki ilişkileri belirlemeye yönelik tasarlanmıştır. Aynı zamanda değerlendirmede ölçme aracının kullanılıp kullanılması ile elde edilen puanlar da karşılaştırılmıştır. Sonuçlar öz değerlendirmede araç kullanılması ve kullanılmaması durumunda elde edilen puanların tutarlı olduğunu göstermiştir. Akran değerlendirmede rubrik kullanılması durumunda daha düşük puanlar elde edilmiştir. Çalışmalar akran değerlendirmede rubrik kullanmanın avantajlı olduğunu ve bu puanlara güvenilebileceğini göstermiştir (Sadler ve Good, 2006; Panadero ve Romero, 2014).

Genel olarak puanladıklarında kızlar ve erkekler kendilerine ve akranlarına benzer puanlar vermektedir. Ancak rubrik kullanıldığında ölçüte göre verilen puanlar daha düşük düzeylerde kalmaktadır. Öğrenciler neyi, nasıl değerlendirdiklerini bildiklerinde daha geçerli puanlar verebilmektedir. JonssonveSvingby, (2007); Sadler ve Good(2006), rubrik kullanmasının akran değerlendirmeye ilişkin yapı geçerliğini arttırdığını ve öğrencilerin daha doğru puanlama yaptıklarını belirtmiştir. Öz ve akran değerlendirmede kız öğrencilerin verdiği puan ortalaması erkek öğrencilerden daha yüksek bulunmuştur. Sherrard ve Raafat (1994), Kaufman, Felder ve Fuller (2000), May and Gueldenzoph (2006) ve Baker (2008)'in bulguları da akran değerlendirmede erkeklerin daha düşük puan aldıkları yönündedir.

Çalışmasının bir başka sonucu öz ve akran değerlendirme puanları arasında anlamlı bir fark bulunduğu yönündedir. Öğrenciler kendilerine daha yüksek puan verme eğilimindedir. Matsuno (2009)'un bulguları bazı öğrencilerin kendilerini değerlendirirken objektif olmadıklarını, ancak yine de birçok öğrencinin kendilerine daha düşük puan verdiklerini göstermiştir. Bunu kültürün etkisi ile açıklamış ve bazı kültürlerde öğrencilerin kendilerini değerlendirmelerinin zor olduğunu, kendilerine diğerlerinden daha yüksek puan vermenin kültürlerine aykırı olduğunu bu nedenle öz değerlendirmenin nadiren kullanıldığını belirtmiştir.

Evans, Aghabeigi, Leeson, O'Sullivan ve Eliahoo (2002) ise öz değerlendirme ile asıl puanlayıcılar arasında elde edilen uyumun zayıf olduğunu ve öğrencilerin kendilerine yüksek puan verdiğini belirtmişlerdir. Bu çalışmada öz değerlendirme ve öğretim üyesi puanlarına ilişkin elde edilen sonuçlar ilgili çalışma ile tutarlılık sergilemiştir.

Panadero, Tapia ve Huertas (2012) öğretmenin verdiği puanların, öz ve akran değerlendirme puanları ile tutarlı olduğunu belirtmiştir. Bu çalışmada akran puanlarının öz değerlendirme puanlarından düşük olduğu bulgusu ilgili çalışma ile çelişmektedir. Ancak, akran puanları öğretim üyesi puanları ile tutarlılık göstermiştir. Matsuno (2009), akran ve öğretmen puanlarının birbirine benzer olduğunu, yanlılık göstermediğini ifade etmiştir. Falchinov ve Goldfinch (2000) akran ve öğretmen puanlarının karşılaştırılmasına ilişkin yaptıkları meta analiz sonucunda, ölçütler iyi belirlendiğinde ve yeterince yargılayıcı olduğunda iki puan türünün birbirine oldukça benzer olduğunu belirtmişlerdir.

Öğrencilerin değerlendirmenin içerisinde olması farklı puanların ortaya çıkmasını sağladığından daha yansız ve kararlı sonuçlar elde etmeye olanak vermektedir. Bu nedenle öğretmenlerin bu yöntemleri kullanması önerilebilir. Özellikle akran değerlendirmenin öğretmen puanına benzer sonuçlar vermesinden dolayı bu puanlar güvenle kullanılabilir. Bunun yanında daha nesnel puanlama yapabilmeleri için puanların karar almada (sınıf geçme, başarısız olma vb.) etkili olamayacağı konusunda öğrenciler bilgilendirilebilir. Bu durum öğrencilerin sadece öğrenmek için değerlendirme yapmalarını sağlayabilir. Cinsiyetin puanlamada yarattığı farklılıklardan dolayı kız ve erkek öğrencilere puanlamayı daha nesnel yapmalarını sağlayacak eğitimler verilebilir. Akran değerlendirme tutarlılık sağlaması açısından aynı öğrenciyi farklı sayıda akran puanlayabilir. Bu sayede değerlendirmeye ilişkin güvenilirlik sağlanmış olur. Benzer şekilde birden fazla öğretmenin puanlaması da puanlara ilişkin güvenilirliği arttırabilir.

5.KAYNAKÇA

- BAILEY, K. M. (1998). Learning about Language Assessment: Dilemmas, Decisions, and Directions. Heinle&Heinle: US.
- BAKER, D. F. 2008. "Peer Assessment in Small Groups: A Comparison of Methods." *Journal of Management Education* 32 (2): 183–209.
- BLACK, P. & WILLIAM, D. (1998). Assessment and Classroom Learning. *Assessment in Education*, Vol. 5, No.1
- CARTNEY, P. (2010). Exploring The Use Of Peer Assessment As a Vehicle For Closing The Gap Between Feedback Given and Feedback Used. *Assessment And Evaluation in Higher Education*, 35 (5), 551-564.
- ELLIS, A. K. (2001). Teaching, Learning, & Assessment Together. Larchmont, NY: Eye on Education, Inc.
- EVANS, A. W., Aghabeigi, B., Leeson, R., O'Sullivan, C., & Eliahoo, J. (2002). Are We Really As Good As We Think We Are? *Annals of the Royal College of Surgeons of England*, 84(1), 54-56.
- FALCHIKOV, N. & BOUD, D. (1989). Student Self-Assessment in Higher Education: A Meta-Analysis. *Review of Educational Research* Winter 1989 vol. 59 no. 4395-430. doi: 10.3102/00346543059004395
- FALCHIKOV, N., & GOLDFINCH, J. (2000). Student Peer Assessment in Higher Education: A Meta- Analysis Comparing Peer And Teacher Marks. *Review of Educational Research*, 70 (3), 287-322.
- GOODRICH, H. (1997). Understanding Rubrics. *Educational Leadership* 54 (4): 14–17.
- HANRAHAN, S. & ISAACS, G. (2001) Assessing Self- and Peer-Assessment: The Students' Views. *Higher Education Research & Development*, 20(1), 53 70.
- HIGGS, J. (1988). Planning Learning Experiences to Promote Autonomous Learning. In *Developing Student Autonomy in Learning* (ed) Boud, D., pp. 40-58. London, UK: Kogan Page
- KAUFMAN, D. B., FELDER, R. M. & FULLER, H. (2000). "Accounting for Individual Effort in Cooperative Learning Teams." *Journal of Engineering Education-Washington* 89 (2): 133–140.
- KEATEN, J. A. & RICHARDSON, M., E. (1993). A Field Investigation of Peer Assessment as Part of the Student Group Grading Process. Paper presented at the Western Speech Communication Association Convention, Albuquerque, NM.
- LEWKOWICZ, J. A., & MOON, J. (1985). Evaluation: A Way of Involving the Learner. In J. C. Alderson (Ed.), *Evaluation* (pp. 45-80). Oxford: Pergamon Press.
- MAGIN, D. J. (2001). A Novel Technique for Comparing the Reliability of Multiple Peer Assessments with That of Single Teacher Assessments of Group Work. *Assessment and Evaluation in Higher Education*, 26, 139-152.

- MATTHEOS, N., NATTESTAD, A, CHRISTERSSON, C., JANSSON, H . & ATTSTROM. R. (2004). "The Effects of an Interactive Software Application on the Self Assessment Ability of Dental Students." *European Journal of Dental Education* 8 (3): 97–104.
- MATSUNO, S. (2009). Self-, Peer-, and Teacher-Assessments in Japanese University EFL writing classrooms. *Language Testing*, 26 (1) 075-100.
- MAY, G. L. & GUELLENZOPH, L. E. (2006). The Effect of Social Style on Peer Evaluation Ratings in Project Teams. *Journal of Business Communication*, 43, (1): 4-20
- ORSMOND, P., MERRY, S. & REILING, K. (1997) 'A Study in Self-Assessment: Tutor and Students' Perceptions of Performance Criteria.' *Assessment & Evaluation in Higher Education* 22.4pp. 357- 369
- ORSMOND, P., MERRY, S., & REILING, K. (2000). The Use of Student Derived Marking Criteria in Peer and Self-Assessment. *Assessment and Evaluation in Higher Education*, 25, pp. 23–38.
- OSCARSON, M. (1989). Self Assessment of Language Proficiency: Rationale and Applications *Language Testing*.
- ÖZAN, S. (2008). Öz ve Akran Değerlendirmenin Temel İletişim Becerileri Üzerindeki Etkileri. Yayınlanmamış doktora tezi, Dokuz Eylül Üniversitesi, İzmir.
- PANADERO, E., ALONSO-TAPIA, J., & HUERTAS, J. A. (2012). Rubrics and Self-Assessment Scripts Effects on Self-Regulation, Learning and Self-Efficacy in Secondary Education. *Learning and Individual Differences*, 22(6), 806-813.
- PARLAK, B. (2010). Öğrenci Performansının Belirlenmesinde Puanlama Anahtarı ve Dereceli Puanlama Anahtarının Karşılaştırılması. Hacettepe Üniversitesi, Yüksek lisans tezi
- POPHAM, J. W. (2000). *Modern Education Measurement*. Needham: Allyn & Bacon.
- ROGERS, C. R. (2003). *Client-Centered Therapy*. London, UK: Constable and Robinson.
- SHERRARD, W. R., & RAAFAT, F. (1994). "An Empirical Study of Peer Bias in Evaluations: Students rating students." *Journal of Education for Business* 70 (1): 43.
- TEMİZKAN, M. (2009). Akran Değerlendirmenin Konuşma Becerisinin Geliştirilmesi Üzerindeki Etkisi. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 6(12), 90-112.
- YURDABAKAN, İ. & OLGUN, M. (2011). Öz ve Akran Değerlendirmenin Öğrenme ve Bilişüstü Bilgi Üzerindeki Etkisi: Sonuçsal Geçerlik. 2nd International Conference on New Trends in Education and Their Implications'da sunulan bildiri, Antalya

ZIMBICKI, D. (2007), “ExaminingThe Effects of Alternative Assessment on Student Motivation and Self- Efficacy”, Doktora Tezi, WaldenUniversity.