



Dereceli Puanlama Anahtarı Kullanan Öğrencilerin Başarıları Üzerine Bir Araştırma¹

Ayşegül BAYRAKTAR²

Canan ASLAN³

Öz

Bu çalışmanın genel amacı, performans ödevlerini hazırlarken Derecelendirilmiş Puanlama Anahtarı (DPA) kullanan ortaokul öğrencileriyle kullanmayan öğrencilerin başarıları arasında fark olup olmadığını belirlemektir. Çalışma grubu, aynı Türkçe öğretmeninden ders alan ve Türkçe dersindeki akademik başarıları birbirine denk iki 6. sınıf öğrencisinden (n=44) oluşmaktadır. Bu gruplardan birisi rastlantısal atama yoluyla deney grubu (n=22), diğeri ise kontrol grubu (n=22) olarak belirlenmiştir. Araştırmanın verileri öğrencilerin performans ödevlerinden aldıkları puanlardan toplanmıştır. Çalışmanın ilk aşamasında öğrencilerin düzeyine ve Türkçe Öğretimi Programı'nın 6. sınıf kazanımlarına uygun bir performans ödevi hazırlanmıştır. Daha sonra bu ödevin nasıl yapılacağını açıklayan ve her iki gruptaki öğrencilere de verilen ayrıntılı bir yönerge ile alanyazın taraması, uzmanların görüş ve önerileri doğrultusunda dokuz boyutu olan çözümleyici (analitik) bir DPA hazırlanmıştır. Ayrıntılı yönerge her iki gruba, DPA ise yalnızca deney grubuna verilmiştir. Öğrencilerin hazırladıkları ödevler bu DPA'nın belirlenen dokuz boyutuna göre deneyimli iki Türkçe öğretmeni tarafından değerlendirilmiştir. Shapiro-Wilk testine göre dağılımların normal olduğu belirlenmiş, bu nedenle Pearson korelasyon katsayısı hesaplanmış ve değerlendiriciler arasındaki tutarlılık .77 olarak belirlenmiştir. Bu değer .05 düzeyinde anlamlı olup, aynı zamanda iki puanlayıcı uyumu için yeterli düzeydedir. İki grubun performans ödevlerindeki başarı puanları arasında fark olup olmadığına Mann-Whitney U testi ile bakılmış; sonuçta yalnızca dil ve anlatım boyutlarında deney grubu öğrencilerinin lehine anlamlı bir fark olduğu görülmüştür.

Anahtar sözcükler: Dereceli puanlama anahtarı, Türkçe dersi, performans ödevi, değerlendirme.

A Study on the Achievement of Students who Use Scoring Rubrics

Abstract

The general aim of this study is to determine achievement differences between students who use rubrics while preparing their performance homework assignments and those who don't. The study group consisted of forty-four 6th grade students who were taking the Turkish course from

¹ Bu çalışmanın bir bölümü 24-26 Nisan 2014 tarihinde İstanbul Üniversitesi Kongre Merkezinde düzenlenen 1. Avrasya Eğitim Araştırmaları Kongresi (1st Eurasian Educational Research Congress)'nde sözlü bildiri olarak sunulmuştur.

² Yrd. Doç. Dr., Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, İlköğretim Bölümü. Ankara. E-posta: abaraktar@ankara.edu.tr

³ Doç. Dr., Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, Güzel Sanatlar Eğitimi Bölümü. Ankara. E-posta: caslan@ankara.edu.tr

the same teacher and had equal academic success rates in two separate sections of the same school. Both groups had the same teacher and demonstrated an equal level of academic achievement. One group was randomly selected as the experiment group ($n=22$) and the other was used as the control group ($n=22$). The data were collected from the scores students obtained on their performance homework assignments. Firstly, a performance homework assignment which was appropriate to the level of the students and the learning outcomes specified in the 6th grade Turkish curriculum was prepared. Afterwards, detailed written instructions that explained how the assignment were to be done and an analytic scoring rubric with nine dimensions which was based on related literature, expert opinions, and recommendations was prepared by the researchers. Detailed instructions were given to both groups, but the rubric was given to the experiment group. The students' written assignments were graded by two experienced teachers, using the nine dimensions specified in the rubric. The Shapiro-Wilk test results showed normal distribution and Pearson correlation co-efficiency was calculated and the consistency between the raters was found to be 0.77. This value was meaningful at .05 level and sufficient regarding the consensus level between the scorers. Mann-Whitney U test was conducted to see whether there were differences in the achievement scores of the two groups, and a statistically significant difference was found only in the *language and expression* dimension between the experiment and control groups.

Key Words: Scoring rubrics, Turkish course, performance homework, assessment.

Giriş

Eğitimde performans, öğrenme ürünlerini kazanmadaki başarı; performansın ölçülmesi de o öğrenme ürünlerini kazanırken işe koşulan sürecin ya da başvuru yöntemin ve elde edilen ürünün ölçülmesidir. Performans ödevi, “programda öngörülen eleştirel düşünme, problem çözme, okuduğunu anlama, yaratıcılığını kullanma, araştırma yapma gibi öğrencinin bilişsel, duyuşsal ve devinişsel alandaki becerilerini kullanmasını, geliştirmesini ve bir ürün ortaya koymasını gerektiren çalışmaları kapsayan ve öğretmen rehberliğinde yaptırılan görevler”dir (MEB Talim ve Terbiye Kurulu Başkanlığı, 2007: 614). Tüm yazılı çalışmaların olduğu gibi performansa dayalı çalışmaların da nesnel, tutarlı ve yansız puanlanması çok önemlidir. Bu nedenle öğretmenlerin başarıyı uygun değerlendirme araçlarıyla puanlamaları gerekmektedir. Performansın ölçülmesinin en önemli yararı, okul içindeki ve dışındaki olgu ve yapılarla ilgili karmaşık becerilerin ölçülmesine olanak sağlamasıdır (Turgut ve Baykul, 2010).

Derecelendirilmiş Puanlama Anahtarı (DPA), öğrencilerin yanıtlarını ve performanslarını belirlenen boyutlara göre puanlamada kullanılacak bir kılavuzdur, öğrencilerin performanslarının hangi yetkinlikte olduğunu belirtir (Jonsson ve Svingby, 2007; Turgut ve Baykul, 2010). Kutlu, Doğan ve Karakaya'ya (2014) göre DPA, öğrencilere yaptıkları çalışmaların hangi ölçütlere göre değerlendirileceğini ve performanslarının hangi düzeydeki puana denk geldiğini gösteren araçlardır. DPA'lar, öğretmenler tarafından genellikle öğrencilerin ürünlerini puanlamak için hazırlanır; birer kontrol listeleri değildir (Mertler, 2001; Petkov ve Petkova, 2006). Bir görevin niteliğinin

değerlendirilmesi için gerekli olan boyutları barındırır (Rezaei ve Lovorn, 2010). Belirli öğrenme hedeflerine yönelik ölçütleri ve standartları belirtir; her bir kategoriye sayısal bir değer atar (Cyr, Smith, Broyles ve Holt, 2014).

Öğretimin değerlendirilmesinde DPA kullanılması hem öğretmen hem öğrenci hem de aileler için önemli ve gereklidir. Öğrenci açısından bakıldığında DPA'lar öğrencilerin ödevlerin içeriğini anlamalarına yardımcı olur (Kan, 2007; Panadero ve Jonsson, 2013; van Helvoort, 2012), onların ilerlemelerini düzenler (Andrade ve Du, 2005; van Helvoort, 2012), aldıkları puanların nedenlerini anlamalarını sağlar (Moskal, 2000; Oakleaf, 2008; Scharf, Elliot, Huey, Briller ve Joshi, 2007; van Helvoort, 2012) ve nitelikli ödevlerin özelliklerini anlatır (Bolton, 2006; Cyr, Smith, Broyles ve Holt, 2014; DelleBovi, 2012; Goodrich, 2000; Kan, 2007; Moskal, 2000). Öğrenciler, DPA'lardan aldıkları dönütlerle güçlü ve zayıf yönleri hakkında bilgi sahibi olurlar (Cyr, Smith, Broyles ve Holt, 2014; Fehring, 2005; Goodrich, 2000; Jonsson ve Svingby, 2007; Panadero ve Jonsson, 2013; van Helvoort, 2012). Başka bir deyişle, DPA'lar öğrencilere dönüt vermesi bakımından da öğretimin ayrılmaz birer parçasıdır. Durum belirleme öncesinde öğrencilere görevde kendilerinden neler beklendiğini ve görevlerin nasıl yapılacağını gösterirler. Durum belirleme sürecinde, ödevlerin yanlılıktan uzak değerlendirilmesini; sonrasında ise, öğrenciye açıklayıcı ve doğru bir geribildirim verilmesini sağlarlar. DPA, değerlendirme için ulaşılması istenen nitelikleri tanımlar ve bunların ne büyüklükte incelendiğine dayanak oluşturur. Öğrenciler, hem kendilerinin hem de arkadaşlarının çalışmalarındaki sorunları ve eksiklikleri görebilir; böylece sorun çözme becerilerini geliştirip eksikliklerini giderebilirler (Goodrich, 2000; Kutlu vd., 2014; Moskal, 2000). DPA kullanmak öğrencinin kendi kendisini değerlendirmesine (özdeğerlendirme), bireysel olarak birçok öğrenmeyi gerçekleştirmesine ve yaşam boyu öğrenme konusunda sorumluluk almasına katkı sağlar (Erman Aslanoğlu ve Kutlu, 2003; Fehring, 2005; Jonsson ve Svingby, 2007; Mason ve Steedly, 2006; Rhodes, 2010; van Helvoort, 2012). Bu yolla öğrencilerin üstbilişleri gelişir. Nitekim DPA kullanan yedinci sınıf öğrencilerinin DPA kullanmayanlara göre kendilerini daha iyi değerlendirdikleri ve konu alanını daha iyi öğrendikleri görülmüştür (Goodrich, 1996). Öğrencilerin kaygı düzeylerini düşürme (Panadero ve Jonsson, 2013) ve öğrencilere ödevlerinden yüksek puan almak için neyin önemli olduğunu gösterme bakımlarından da DPA kullanımının çok önemli olduğu düşünülmektedir (Wiggins, 1989).

Öğretmenler ise DPA'lar yardımıyla öğrenci çalışmalarını daha nesnel, hızlı (Aslanoğlu, 2003; Goodrich, 2000, 2005; Kutlu, Doğan ve Karakaya, 2014; Tompkins, 2008), etkili (DelleBovi, 2012) ve tutarlı değerlendirebilir (Cyr, Smith, Broyles ve Holt, 2014), derslerini öğrencilerin yetenek ve gereksinimlerine uygun biçimde planlayabilir (Mason ve Steedly, 2006) ve öğrencilere dönüt verilmesi gerektiğinde uzun uzun açıklamalar yapmak yerine, değerlendirmede kullanılan DPA'nın boyutlarına dikkat çekebilirler (Goodrich, 2000). Kısaca bu anahtarlar, öğrenci ve öğretmenlerin öğrenme

görevlerinin en önemli ögesine odaklanmalarını sağlar. Hem öğretmenler hem de öğrenciler; öğrenmenin niteliğine, öğrencinin güçlü ve zayıf yönlerine ilişkin ayrıntılı dönütler alırlar (Boston, 2002; DelleBovi, 2012). Eğitimde DPA'ların kullanılması çocuklarına ödevlerini tamamlarken süreç içerisinde daha fazla yardımcı olmalarını sağladığından aileler için de rahatlatıcıdır (Goodrich, 2000).

Yapısal özellikleri bakımından bütünsel (holistic rubric) ve çözümleyici (analytical rubric) olmak üzere iki tür DPA bulunmaktadır. Çözümleyici DPA'larda herhangi bir özel amaca yönelik birden fazla dereceli ölçüt bulunur. Bütünsel puanlama anahtarında ise, çoğunlukla tek ve genel bir dereceli ölçüt bulunur, belli bir konudaki öğrenmenin ve yetkinliğin (performansın) farklı boyutlarını yüzeysel olarak ölçerken özel boyutlara ilişkin herhangi bir dönüt vermez. Örneğin, öğrencinin hazırladığı ödev, içerik olarak bir bütünlük gösteriyorsa bu DPA'nın kullanımıyla dilbilgisel yanlışlar göz ardı edilebilir. Bütünsel DPA'lar daha hızlı bir değerlendirme yapmaya olanak sağlarken çözümleyici DPA'lar her bir boyutun tek tek puanlanması nedeniyle puanlayıcının daha fazla zaman ayırmasını gerektirir. Buna karşın çözümleyici anahtarlar hem öğretmene hem de öğrenciye öğrenmelerine ilişkin ayrıntılı dönüt verdiğinden anlamlı ve önemli kabul edilmektedir (Chase, 1999; Elbow, 2000; Kutlu, Doğan ve Karakaya, 2014; Mertler, 2001).

Hem bütünsel hem de çözümleyici DPA'lar genel amaçlı olabileceği gibi belirlenmiş bir görev için geliştirilmiş; başka bir deyişle göreve özel de (task specific) olabilir. Tüm görevler aynı DPA kullanılarak değerlendirilebiliyorsa bu DPA geneldir (generic rubric). Özel bir durumla ilgili görevlerin verildiği durumlarda genel bir DPA'nın kullanılması, gözlenmesi gereken kimi ölçütlerin göz ardı edilmesine, kimi ölçütlerinse yeterince duyarlı ölçülmemesine neden olur. Böyle durumda öğretmenlerin ilgili göreve ilişkin boyut ve tanımlamalarını içeren özel bir DPA kullanmaları gerekir (task specific rubric). Hiçbir yetenek yalnızca bir boyutuyla değerlendirilecek kadar basit yapıda olmadığından genel DPA'ların kullanılması, ölçmelerin güvenilirliğini düşürür. Hazırlanması daha karmaşık ve zaman alıcı olmakla birlikte belirlenen boyutlara ve ayrıntılı tanımlamalara yer verdiğinden göreve özel DPA'lar, öğrencilerin öğrenmelerine ilişkin daha fazla bilgi verirler (Kutlu, Doğan ve Karakaya, 2014; Moscal, 2000).

Öğrenci başarısında DPA'ların etkisi birçok araştırmanın konusu olmuştur. DPA'nın etkisi yalnızca dil ve edebiyat derslerinde değil (Andrade vd., 2008, 2009, 2010; Çetin, 2011; Goodrich Andrade, 2001; Goodrich Andrade and Boulay, 2003; Reynolds-Keefer, 2010), sanat (Mason ve Steedly, 2006), fen (Huerta, Lara-Alecio, Tong ve Irby, 2014; Reitmeier ve Vrchota, 2009), teknoloji ve bilim (Petkov ve Petkova, 2006), bilgi yönetimi (van Helvoort, 2012) ve eğitim psikolojisi (Andrade ve Du, 2005) gibi başka alanlarda da kullanılmıştır.

Goodrich ve Boulay (2003), birini deney diğerini kontrol grubu olarak belirlediği iki gruba da bir DPA vermiş ve birer öykü yazmalarını istemiştir. Deney grubundaki öğrencilere ayrıca

araştırmacılar tarafından 40'ar dakikalık iki oturumluk özdeğerlendirme dersleri verilmiştir. Çalışmanın sonucu, öğrencilerin öykü yazma performansları bakımından iki grup arasında anlamlı bir fark çıkmadığını göstermiştir. Araştırmacılara göre bunun nedenlerinden biri hem deney hem de kontrol gruplarındaki tüm öğrencilere DPA verilmiş olmasıdır; çünkü DPA ve içerdiği boyutlar yoluyla öğrencilere nitelikli metinlerin özellikleri konusunda bilgiler verilmiştir. Araştırmacılar, özdeğerlendirmenin nasıl yapılacağını öğrencilere anlatmanın, bu bilgilerin öğrencilerce hemen benimseneceği ve yazılarına aktarılacağı anlamına gelemeyeceği sonucuna ulaşmışlardır.

Goodrich (2001), DPA kullanımının sekizinci sınıf öğrencilerinin (n=242) kompozisyon yazma başarılarına etkisinin olup olmadığını ve katılımcıların nitelikli kompozisyonun özellikleri konusundaki görüşlerini incelemiştir. Deney grubundaki öğrencilere DPA vermiş, kontrol grubundakilere vermemiştir. Kontrol grubundaki öğrenciler nitelikli bir kompozisyonun özelliklerine ilişkin “yazım yanlışları”, “noktalama işaretleri” ve “güzel yazı” gibi özellikleri vurgularken deney grubundaki öğrencilerin bu özelliklere ek olarak kendilerine verilen DPA'daki boyutlara da değindikleri görülmüştür. Bu sonuçtan yola çıkarak araştırmacılar, DPA'ların öğrencilere nitelikli yazılı metinlerin özelliklerini öğretmede etkili olduklarını belirtmiştir.

Koutsoftas ve Gray (2012), öğrenme güçlüğü çeken öğrencilerle normal gelişim gösteren öğrencilerin kurmaca ve bilgilendirici türde metin yazma performanslarını karşılaştırdığı çalışmada, DPA kullanmanın öğrencilerin yazılı anlatım becerilerini geliştirdiği sonucuna varmıştır. Goodrich (1996), DPA kullanmanın ve özdeğerlendirme yapmanın yedinci sınıf öğrencilerinin öğrenmelerine ve üstbiliş gelişimlerine etkisini araştırdığı deneysel çalışmada DPA kullanan öğrencilerin daha başarılı olduğu bulgusuna ulaşmıştır. Mason ve Steedly'nin (2006) bilimsel konuların tarihinin öğretilmesinde sanattan ve DPA'dan yararlanmanın önemine vurgu yapan çalışmasının sonucunda DPA kullanmanın; diğer derslerdeki görsel sanatları değerlendirmede, öğretmenin etkinliklerini planlamasında, sanatçılar ve eğitimciler arasında işbirliğini artırmada, sanatsal etkinlikleri yapılandırmada ve öğrencilere özdeğerlendirme yaptırmada etkili olduğunu belirlemişlerdir.

Huerta, Lara-Alecio, Tong ve Irby (2014), DPA kullanmanın beşinci sınıf öğrencilerinin (n=210) bilgiyi yapılandırmalarına olan etkisini araştırdıkları çalışmalarında DPA kullanımının öğrencilerin öğrenmelerini ve deneyimlerini yazılarına aktarmalarında daha etkili olduğu sonucuna varmıştır. Benzer bir çalışmada, Petkov ve Petkova (2006), DPA kullanmanın öğrencilerin fen derslerindeki projelerini hazırlamadaki etkisini araştırmış; sonuçta üç dönem boyunca DPA kullanan öğrencilerin performanslarının olumlu yönde değiştiği ve geliştiği sonucuna varmıştır. van Helvoort (2012), öğrencilerin DPA'ları derslerde ve günlük yaşamlarında nasıl kullandıklarını araştırmış; sonuç olarak öğrencilerin bu araçları diğer derslerde de kullandıklarını ve kullandıkları bu DPA'lar yardımıyla hem kendi yazılarına hem de başkalarının yazılarına daha eleştirel bakabildiklerini bulgulamıştır.

Dereceli Puanlama Anahtarı Kullanan Öğrencilerin Başarıları Üzerine Bir Araştırma

Anılan çalışmalardan da görüldüğü gibi öğretimde DPA kullanmanın öğrenci başarısına, öğrenmesine ve özdeğerlendirme yapmasına olan etkileri birçok farklı alanda araştırılmıştır. Bazı çalışmalar da DPA'ların güvenilir ölçme araçları olduğunu ve gene birçok alanda kullanıldığını göstermektedir. Örneğin, East'in (2009) yaptığı çalışmada, Almanca'yı yabancı dil olarak öğrenen öğrencilerin yazdığı metinler ayrıntılı bir DPA eşliğinde iki farklı okuyucu tarafından değerlendirilmiştir. Çalışmanın bulguları amaca uygun ayrıntılı DPA'ların öğrencilerin yazma performanslarının ölçülmesinde güvenilir sonuçlar verdiğini göstermiştir.

Cyr, Smith, Broyles ve Holt'un (2014) çalışmasında beş öğretim üyesi "Aile Hekimliği" dersinde staj yapan öğrencilerinin (n=83) hazırladıkları durum belirleme raporlarını DPA kullanarak değerlendirmiş ve öğrencilerin DPA kullanımına ilişkin görüşlerini almışlardır. Araştırma sonuçları okuyucular arasında yüksek oranda tutarlılık bulunduğunu ve öğrenci çalışmalarının değerlendirilmesinde güvenilir ölçme araçlarının kullanılmasının önemli olduğunu göstermiştir. Erman Aslanoğlu ve Kutlu'nun (2003) yaptığı çalışmada, 14 fen bilgisi öğretmeni, öğrencilerin (n=24) sunu başarılarını DPA kullanarak değerlendirmiştir. Sonuç, DPA kullanmanın farklı okuyucular arasında tutarlı bir değerlendirme yapmaya olanak sağladığını göstermiştir. Beyreli ve Arı'nın (2009) DPA kullanmanın farklı değerlendiriciler arasındaki tutarlılığa olan etkisine vurgu yaptığı çalışmada öğrenciler (n= 200) tarafından yazılan kurgusal metinler altı farklı okuyucu tarafından değerlendirilmiş; puanlayıcılar arasında yüksek oranda tutarlılık olduğu görülmüştür.

Puanlama sonucunda öğrencilerin belirlenmiş boyutlar ve bu boyutlara ilişkin tanımlamalar doğrultusunda, başarılı ve eksik oldukları yanlarını görebilmeleri açısından öğretim sürecinin gelişmesine katkı sağlayan ve öğrenme hedeflerini vurgulayan göreve özel-çözümleyici dereceli puanlama anahtarlarının kullanılması büyük önem taşımaktadır. Türkçe derslerinde çoğunlukla kompozisyon biçiminde hazırlanması istenen performans ödevlerinin nesnel bir biçimde değerlendirilmesi öğretmenler için güç bir iştir. Bu nedenle öğrencilere verilen görevlerin geçerliği ve güvenilirliği olan "göreve özel-çözümleyici DPA"larla puanlanması önemlidir. Türkiye'de 2006 yılında eğitim-öğretim programlarının yenilenmesiyle birlikte başarının belirlenmesinde Türkçe derslerinde de DPA kullanımı ön plana çıkmıştır. Bu durumda Türkçe dersleri için göreve özel çözümleyici DPA'ların etkililiğini gösteren çalışmaların yapılması gerekmektedir.

Amaç

Bu çalışmanın genel amacı, performans ödevlerini hazırlarken DPA kullanan öğrencilerle kullanmayan öğrencilerin başarıları arasında fark olup olmadığını belirlemektir. Dolayısıyla bu araştırmanın sorusu "Göreve Özel Çözümleyici DPA kullanan öğrencilerin performans ödevlerinden

aldıkları başarı notları DPA kullanmayan öğrencilerin başarı notlarından anlamlı bir biçimde farklılaşmakta mıdır?" biçiminde oluşturulmuştur.

Yöntem

Bu çalışma deneysel bir araştırmadır ve yarı deneysel desenlerden son test kontrol gruplu eşleştirilmiş gruplu desende yürütülmüştür. Bu desen gerçek deneysel desenlere göre neden-sonuç ilişkisini ortaya koymada göreceli olarak daha zayıftır. Bu da araştırmanın bir sınırlılığı olarak düşünülebilir. Araştırmada DPA kullanan öğrencilerle kullanmayan öğrencilerin başarıları karşılaştırılmıştır. Bütünsel DPA'ların nesnel ve güvenilir değerlendirme için uygun olmadığı gerekçesiyle (Elbow, 2000; Gunning, 2006; East, 2009; Koutsoftas ve Gray, 2012) bu çalışmada göreve özel çözümleyici puanlama anahtarının kullanılması yeğlenmiştir.

Çalışma Grubu

Çalışma, 2014-2015 eğitim-öğretim yılının ikinci (bahar) döneminde Milli Eğitim Bakanlığına bağlı bir devlet okulunda gerçekleştirilmiştir. Araştırmanın çalışma grubu, aynı okuldaki iki altıncı sınıfta öğrenim gören 44 öğrencidir. Sınıflardan biri deney, öteki ise kontrol grubu olarak belirlenmiştir. Deney grubunda 10'u kız, 12'si erkek 22; kontrol grubunda 8'i kız, 14'ü erkek 22 öğrenci bulunmaktadır. Çalışmada öğretmen farklılığından; başka bir deyişle denkleştirme sorunundan kaynaklanabilecek sorunların oluşmaması için her iki grubun Türkçe derslerini de aynı öğretmenin yürütmesi sağlanmış; öğretmene deney grubuna verilen DPA bilgilerini kontrol grubuna hiçbir şekilde yansıtmaması gerektiği söylenmiştir. Ayrıca, deney grubundaki öğrencilerin DPA'dan etkili yararlanabilmeleri için öğretmene neler yapması gerektiği adım adım ve ayrıntılı biçimde açıklanmıştır.

Çalışmanın başında okul idaresinden her iki gruptaki öğrencilerin akademik başarılarının birbirine denk oluşturulduğu bilgisinin alınmasına karşın araştırmacılar öğrencilerin bir dönem önceki Türkçe ders notlarını temel alarak gruplar arasında anlamlı bir fark olup olmadığını t-testi ile belirlemişlerdir. Deney ve kontrol gruplarının bir dönem önceki Türkçe ders notlarının normal dağılım gösterip göstermediği Shapiro-Wilk testine testi ile sınınanmış, her iki grup için de dağılımın normal olduğu belirlenmiştir. Sonuçta kontrol grubu öğrencilerinin Türkçe dersi ortalama puanlarının ($X=79,93$) deney grubu öğrencilerinin ortalama puanlarından ($X=75,24$) yüksek olduğu; ancak ortalama puanlar arasındaki bu farkın anlamlı olmadığı belirlenmiştir ($t(52)=-1,463, p>0,01$). Başka bir deyişle, çalışmanın başında deney ve kontrol gruplarının Türkçe dersi akademik başarı puanları arasında anlamlı bir farklılık olmadığı görülmüştür. Bu bulgu grupların ölçülen özellik bakımından denk olduğu biçiminde yorumlanabilir.

Performans Ödevlerinin, Yönergenin ve Göreve Özel Çözümleyici DPA'nın Hazırlanma Süreci

Bu çalışmada öncelikle, ortaokul 6. sınıf öğrencilerinin düzeylerine ve Türkçe dersinin amaçlarına uygun bir performans ödevi hazırlanmıştır. 6. sınıf Türkçe Ders Kitabında “Okuma Kültürü”, “Atatürk”, “Sevgi”, “Duygular”, “Zaman ve Mekân” ile “Doğa ve Evren” olmak üzere toplam altı tema bulunmaktadır. Hazırlanan performans ödevi, “Sevgi” temasının içeriğine uygun olarak seçilmiştir. Ödev kapsamında öğrencilerden ders kitaplarında, basın yayın organlarında (gazete, radyo, televizyon, vb.) ve günlük yaşantılarında (mağaza ve dükkân adlarında, konuşma dilinde) Türkçe karşılıkları olmasına karşın kullanılan yabancı sözcükleri belirlemeleri ve bunların Türkçe karşılıklarını üç hafta içinde sözlük (ürün) biçiminde hazırlamaları istenmiştir. Daha sonra Türkçe Öğretimi Programından bu performans ödeviyle öğrencilere kazandırılacağı düşünülen 13 kazanım belirlenmiş ve bunlar da performans ödevi kâğıdına yazılmıştır. Ödev kâğıtlarında ürünlerini tamamlayabilmeleri için gereken tüm bilgiler (ilgili kaynaklardan konuya ilişkin makaleler okuma, çalışma planı hazırlama ve bunu öğretmene gösterme, plana uygun çalışma, dosya oluşturma ve çalışmayla ilgili verileri bu dosyada toplama, ödevin değerlendirileceği boyutlar, amaca uygun görseller, kullanılan kaynaklar) ile öğrencilerin bu ödevi yaparken hangi becerileri kazanacaklarını (Yazılı Anlatım Becerisi, Araştırma Becerisi, Eleştirel Düşünme Becerisi) aşamalı ve dizgesel bir biçimde anlatan ayrıntılı bir yönerge eklenmiştir.

İkinci aşamada deney grubundaki öğrencilerin ödevlerini hazırlarken yararlanmaları amacıyla “göreve özel çözümleyici bir DPA” hazırlanmıştır. Bu amaçla yerli ve yabancı alanyazındaki çeşitli kaynaklardan birçok göreve özel çözümleyici DPA örneği incelenmiş; nelerin göz önünde bulundurulması, boyutların ve tanımlarının hazırlanmasında nelere dikkat edilmesi gerektiği yine birçok yerli ve yabancı kaynağın tanıklığında not edilmiştir. Hazırlanan çözümleyici DPA’da, 9 boyutun olmasına karar verilmiş (Verilerin Toplanması, Görsellerle Destekleme, Sayfa Düzeni, Plan, İşleyiş, Dil ve Anlatım, Sözcük Seçimi, Yazım Noktalama ve Sözlük Düzeni); boyutların uygun ve açık biçimde tanımlanarak önemli bir teknik gereklilik olan tutarlığın sağlanmasına ve DPA’nın tasarımı ile amaçlarının birbirine uygun düşmesine önem verilmiştir. “Verilerin Toplanması” boyutu, öğrencinin seçtiği konuya uygun ayrıntılı araştırma yapıp yapmadığını açıklamaktadır. “Görsellerle Destekleme” boyutunda ödevin içeriğine uygun görsellerin olup olmadığı, varsa sayısının yeterliği dikkate alınmıştır. “Sayfa Düzeni” boyutunda, öğrencinin ödevini uygun kenar boşlukları bırakarak hazırlayıp hazırlamadığına, okunaklı ve bitişik eğik yazı ile yazıp yazmadığına bakılmıştır. “Plan” boyutunda öğrencinin ödevini giriş, gelişme, sonuç bölümleri biçiminde düzenleyip düzenlemediği ve bu bölümlerde yönergede istenen bilgilere ulaşip ulaşmadığı göz önünde bulundurulmuştur. “İşleyiş”te, öğrencinin listelediği sözcük sayısı temel alınmıştır. “Dil ve Anlatım” boyutunda, dilbilgisi kurallarına uyulup uyulmadığına bakılmıştır. “Sözcük Seçimi”nde sözcüklerin Türkçe karşılıklarıyla yazılıp

yazılmadığı göz önünde bulundurulmuştur. “Yazım ve Noktalama” boyutunda, yazım kurallarına uygunluk ile sözcüklerin doğru yazılıp yazılmadığına dikkat edilmiştir. “Sözlük Düzeni” boyutunda ise, öğrencinin performans ödevini bir sözlük biçimine getirip getirmediği, ödevde gereken bilgileri eksiksiz yazıp yazmadığı dikkate alınmıştır.

Boyutlar ve tanımlar hazırlanırken öğretim programındaki kazanımlar da göz önünde bulundurulmuştur. DPA’ların işlevsel ve verimli olabilmesi için kullanılan dilin okuyucular açısından yalın, anlaşılır ve tutarlı olması gerekir (Tierney ve Simon, 2004). Bu çalışmada kullanılan DPA’da açık ve anlaşılır bir dilin kullanılmasına ve amaca uygun sözcüklerin seçilmesine özen gösterilmiştir. Hem performans ödevi ve ödevin nasıl yapılması gerektiğine ilişkin yönerge hem de geliştirilen çözümleyici DPA, alana ve konuya hâkim 10 Türkçe öğretmene gönderilerek uzman görüşü alınmış; onların görüş ve önerileri doğrultusunda gerekli düzeltmeler yapılarak ödev, yönergeye ve DPA’ya son biçimleri verilmiştir. Uzman görüşü alınan öğretmenlerin ikisi dil ve edebiyat öğretimi alanında doktorasını, geriye kalan sekizi de yüksek lisanslarını tamamlamak üzere olan öğretmenlerden seçilmiştir.

Verilen ödevi nasıl hazırlamaları gerektiğine ilişkin ayrıntılı yönerge, her iki gruba da verilmiş; oluşturulan göreve özel çözümleyici DPA ise, yalnızca deney grubundaki öğrencilere verilmiştir. Bu gruba DPA’nın her bir alt boyutunun ne anlama geldiği ve nasıl puanlanacağı da açıklanmıştır. Ödev verildikten sonra öğretmenle belirli zaman aralıklarında görüşülmüş, sürecin sağlıklı ve başarılı yürümesi açısından gereken önlemler alınmıştır. Performans ödevleri önceden belirlenen tarihte toplanmış, çalışmaya katılan tüm öğrenciler ödevlerini zamanında teslim etmiştir. Toplanan her ödev bir numara verilmiş ve bu numaralar Microsoft Excel dosyasında öğrencilerin adlarıyla eşleştirilmiştir.

DPA kullanan ve kullanmayan öğrencilerin çalışmaları; özel okullarda çalışan, yüksek lisanslarını Türkçe öğretimi alanında yapan ve aynı zamanda 6. sınıf öğrencilerinin Türkçe derslerini yürüten iki Türkçe öğretmeni tarafından ödevlerin kime ve hangi gruba ait olduğu bilinmeden puanlanmıştır. Puanlama başlamadan önce gelişigüzel seçilen beş ödev, geliştirilen DPA’ya göre ortak bir anlayış geliştirilene ve kodlamada birlik sağlanana dek DPA boyutlarına göre ortaklaşa kodlanmıştır. Değerlendiriciler arasındaki tutarlık ise Pearson korelasyon katsayısı ile hesaplanmış ve sonuç, .77 bulunmuştur. Buna göre ödevlerin güvenilir bir biçimde puanlandığı söylenebilir.

Veri Kaynağı ve Verilerin Analizi

Bu araştırmanın veri kaynağı, öğrencilerin göreve özel çözümleyici DPA ile yönergeyi kullanarak hazırladıkları performans ödevlerinden aldıkları puanlardır. DPA kullanan ve kullanmayan öğrencilerin puanları arasındaki farkın anlamlılığı ise, ilişkisiz Ölçümler İçin Mann-Whitney U testi kullanılarak test edilmiştir.

Bulgular

Performans ödevlerini DPA kullanarak ve kullanmadan hazırlayan öğrencilerin başarı puanları arasındaki farkı belirlemek amacıyla gerçekleştirilen bu araştırma sonucunda, yalnızca bir boyutta (Dil ve Anlatım) deney grubunun lehine istatistiksel olarak anlamlı bir farkın olduğu, diğer boyutlardaki farkın anlamlı olmadığı görülmüştür. Çizelge 1, deney ve kontrol gruplarındaki öğrencilerin DPA'nın dokuz farklı boyutundan aldığı puanları göstermektedir.

Çizelge 1. Deney ve Kontrol Gruplarının DPA'nın Ölçütlerine İlişkin Puanlarına Ait Betimsel İstatistikler ve Mann Whitney U Testi Sonuçları

Ölçütler	Grup	n	Ort.	Ortn.	SS	Min.	Maks.	Çarpıklık	Sivrilik	Shapiro- Wilk	Sıra ort.	U, p
Verilerin toplanması	Deney	22	2,95	3	,84	1	4	-,43	-,24	,86 p=.005*	20,73	203,00
	Kontrol	22	3,18	3	,85	1	4	-,88	,42	,82 p=.001*	24,27	p=.328
Görsellerle destekleme	Deney	22	2,81	3	1,05	1	4	-,41	-,97	,86 p=.005*	25,32	179,00
	Kontrol	22	2,27	2	1,24	1	4	,41	-1,48	,80 p=.001*	19,64	p=.126
Sayfa düzeni	Deney	22	2,95	3	,38	2	4	-,59	5,66	,52 p=.000*	22,00	231,00
	Kontrol	22	3	3	,00	3	3	-	-	-	23,00	p=.555
Plan	Deney	22	2,13	2	,89	1	4	,16	-,84	,87 p=.007*	22,98	213,50
	Kontrol	22	2,09	2	,97	1	4	,49	-,64	,86 p=.006*	22,02	p=.796
İşleyiş	Deney	22	3,59	4	1,05	1	4	-2,28	3,50	,41 p=.000*	21,43	218,50
	Kontrol	22	3,90	4	,43	2	4	-4,69	22,00	,22 p=.000*	23,57	p=.268
Dil ve anlatım	Deney	22	2,72	3	1,07	1	4	-,66	-,75	,80 p=.000*	26,23	160,00
	Kontrol	22	2	1,5	1,15	1	4	,61	-1,19	,78 p=.000*	18,77	p=.042*
Sözcük seçimi	Deney	22	3,5	4	1,10	1	4	-1,88	1,82	,49 p=.000*	21,73	225,00
	Kontrol	22	3,86	4	,35	3	4	-2,28	3,49	,41 p=.000*	23,57	p=.531
Yazım ve noktalama	Deney	22	2,73	3	,88	1	4	-,77	,22	,81 p=.001*	25,73	171,00
	Kontrol	22	2,18	3	,95	1	3	-,40	-1,90	,68 p=.000*	19,27	p=.062
Sözcük düzeni	Deney	22	3,73	4	,63	2	4	-2,23	3,90	,49 p=.000*	23,82	213,00
	Kontrol	22	3,59	4	,73	1	4	-2,33	6,61	,60 p=.000*	21,18	p=.368
Toplam puan	Deney	22	27,14	29	5,35	15	34	-,88	-,25	,90 p=.026*	24,23	204,00
	Kontrol	22	26,09	27	4,59	18	33	-,12	-1,09	,94 p=.185*	20,77	p=.370

*p<.05

Çizelge 1 incelendiğinde DPA'nın altıncı boyutunda (Dil ve Anlatım) deney grubunun lehine anlamlı bir farklılık olmasına karşın diğer boyutlarda iki grubun sıra ortalamaları arasında anlamlı bir fark olmadığı söylenebilir. Bu bulgu, araştırma için hazırlanan DPA'nın performans ödevlerinin değerlendirilme sürecinde öğrenci başarıları üzerinde önemli bir etki göstermediği biçiminde yorumlanabilir. Başka bir söyleyişle, performans ödevinin değerlendirilmesi sürecinde öğrencilere DPA verilmesiyle onların performans ödevini hazırlama başarıları arasında tam bir ilişki bulunamamıştır.

Tartışma ve Sonuç

Performans ödevlerini hazırlarken çözümleyici DPA kullanan öğrencilerle kullanmayan öğrencilerin başarıları arasında fark olup olmadığını belirlemek amacıyla yürütülen bu çalışmada yapılan analizler, DPA'nın yalnızca altıncı boyutunda (Dil ve Anlatım) deney grubu lehine istatistiksel olarak anlamlı bir farkın olduğunu göstermiştir. Sonucun bu şekilde çıkması birçok nedene bağlanabilir. Bunlardan biri öğrencilerin DPA'dan işlevsel biçimde yararlanamaması olabilir. Başka bir deyişle, sonucun deney grubu lehine yalnızca bir boyutta anlamlı çıkması, öğrencilerin DPA kullanımına alışkın olmamasına ve her iki gruptaki öğrencilerin de ödevlerini daha çok onlara verilen ayrıntılı yönergeye bağlı kalarak hazırlamalarına bağlanabilir. Dolayısıyla öğrenciler DPA'yla çalışmaya ve özdeğerlendirme yapmaya alışkın olmadıkları için ayrıntılı yönergeyi dikkate almış olabilirler ya da DPA'da açıklanan ifadeleri görevlerine aktaramamış olabilirler.

Sonucun bu şekilde çıkmasının bir başka nedeni, öğretmenin deneysel uygulamayı eksik ya da yanlış yapması olabilir. Uygulama öncesi ilgili bilgilerin verilmesine ve uyarıların yapılmasına; başka bir deyişle gereken tüm önlemlerin alınmasına karşın öğretmen, deney grubundaki uygulamayı ve DPA'nın boyutlarını farkında olmadan kontrol grubu öğrencilerine de yansıtmış olabilir. Yine, DPA kullanma konusundaki deneyimsizliği nedeniyle öğretmen, araştırmacıların deney grubunda kullanmak üzere kendisine verdiği yönergeler (DPA'yı doğru, işlevsel ve etkili kullanmaları konusunda) bütünüyle uymamış olabilir.

Her iki gruptaki öğrencilere de ödevi nasıl hazırlayacaklarına ilişkin verilen yönergenin ayrıntılı olması da sonucun bu şekilde çıkmasının bir başka önemli nedeni olabilir. Başka bir söyleyişle ödev hazırlarken DPA kullanmaya alışık olmayan öğrenciler, ödevlerini daha çok alışkın oldukları yönergeye bağlı kalarak hazırlamış olabilirler. Ayrıca, kullanılan DPA, dört aşamalı başarı puanı esas alınarak hazırlanmıştır. Puan aşamasının fazla olmaması da sonucun bu şekilde çıkmasına neden olmuş olabilir.

DPA kullanımının öğrenci başarısında anlamlı bir fark yaratmadığı sonucu Goodrich ve Boulay'ın (2003) çalışmasında da görülmüştür. Araştırmacılar sonucun bu şekilde çıkmasını DPA'nın her iki gruba birden verilmesine bağlamaktadır. DPA ve ölçütlerini gören öğrenciler elbette nitelikli

yazıya ilişkin bilgi sahibi olmuşlardır. Araştırmacılara göre sonucun böyle çıkmasının bir diğer nedeni de öğrencilerin öğrendiklerini kendilerine verilen benzer görevlere aktaramamasıdır. Yine Reitmeier ve Vrchota'nın (2009) 35 katılımcı ile yürüttüğü çalışmada da DPA kullanan ve kullanmayan öğrenciler arasında anlamlı bir fark bulunmamıştır. Onların çalışmasında öğrenciler DPA'nın ölçütlerine ilişkin bilgi sahibi olmuşlar; ancak bu çalışmada da olduğu gibi kendilerine verilen yönergeye daha fazla bağlı kalmışlardır. DPA'lar hazırlanırken kullanılan dilin öğrenci düzeyinin üstünde olmaması ve görevlerini tamamlamaları için onlara yeterli sürenin verilmesi gerekir (Goodrich Andrade, 2001). Bu çalışmada kullanılan DPA hazırlanırken çok sayıda öğretmenden görüş alındığı için DPA'daki ifadeler çocukların anlayacağı şekilde yazılmış ve görevlerin tamamlanması için öğrencilere yeterli süre (üç hafta) verilmiştir. Dolayısıyla çalışma sonuçlarının bu etkenlerden etkilenmediği düşünülmektedir.

Hem bu çalışmanın hem de Goodrich ve Boulay'ın (2003) çalışmasının sonuçlarına bakıldığında öğretmenlerin DPA kullanımı konusunda öğrencileri güdülemelerinin, ölçütleri daha iyi anlayabilmeleri ve önemsemeleri için öğrencilerle birlikte oluşturmalarının daha etkili olabileceği söylenebilir. Bu nedenlerle öğretmenler performans ödevlerini puanlarken kullandıkları DPA'ları ve bunları kullanmaya dönük yönergeleri öğrencilere de vermeli, ödevlerini yaparken bu anahtarlardaki ölçütleri kullanmalarının yararlarını söyleyerek onları DPA kullanma konusunda güdülemelidir. Hatta göreve özel DPA'ların boyutlarını öğrencilerle birlikte de hazırlayabilir, onları da işin içine sokarak düşüncelerinden etkin bir biçimde yararlanabilirler. Aynı şekilde DPA kullanmak öğrencilere ödevlerinde olması gereken ölçütlere ilişkin bilgi verebilir; ancak öğrencilerin bu bilgileri kullanmaları için DPA kullanımının süreklilik göstermesi gerekir (Goodrich Andrade, 2001).

İlgili alanyazın ve bu çalışmanın sonuçları doğrultusunda aşağıdaki öneriler geliştirilebilir. Wiggins'in (1989) de belirttiği gibi standardize edilmiş testler öğrencilerin entelektüel gelişimlerini değerlendirmede sınırlı kalabilmektedir. Dolayısıyla öğrenci başarısını değerlendirmede DPA gibi daha etkili ve dönüt vermeye de olanak sağlayan (Bangert-Drowns, Kulik, Kulik ve Morgan, 1991; Butler ve Winne, 1995) araçlardan yararlanılabilir. Buna ek olarak öğretmenlerin değerlendirme yaparken DPA kullanma konusunda daha bilgili ve deneyimli olmaları gerekmektedir (Özdemir, 2010).

Alanyazında ve bilgisayar (internet) ortamında eğitimcilerin kullanması amacıyla hazırlanan birçok DPA örneği bulunmaktadır. Ancak Rezaei ve Lovorn (2010) ile Popham'ın (1997) da vurguladığı gibi tüm bu örnekler doğru kabul edilerek olduğu biçimiyle kullanılmamalıdır. Araştırmacıların çalışmalarında eğitim fakültelerindeki bir grup profesörün hazırladığı DPA'ların bile güvenilir olmadığı bulgulanmıştır. Bu çalışmada olduğu gibi, eğitimciler hem konunun içeriğine hem de çalıştıkları grubun düzeyine uygun DPA'lar hazırlamalıdır. Performans ödevlerinin değerlendirilmesinde, DPA'ların kullanılmasının önemini vurgulayan, bu konuda farkındalık yaratmayı sağlayan

araştırmaların yapılması ve bu araştırmalar doğrultusunda önerilerin geliştirilerek uygulamaların yaşama geçirilmesi bu çalışma kapsamında getirilen bir başka öneridir.

Kaynaklar

- Andrade, H., Buff, C., Terry, J., Erano, M. ve Paolino, S. (2009). Assessment-driven improvements in middle school students' writing. *Middle School Journal*, 40(4), 4–12.
- Andrade, H. ve Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment, Research & Evaluation*, 10(3), 1-11.
- Andrade, H., Du, Y. ve Mycek, K. (2010). Rubric-referenced self-assessment and middle school students' writing. *Assessment in Education: Principles, Policy & Practice*, 17(2), 199–214.
- Andrade, H., Du, Y. ve Wang, X. (2008). Putting rubrics to the test: The effect of a model, criteria generation, and rubric-referenced self-assessment on elementary school students' writing. *Educational Measurement: Issues and Practices*, 27(2).
- Aslanoğlu, A. E. (2003). *Eğitimde sunu becerilerinin değerlendirilmesinde dereceli puanlama anahtarı (rubric) kullanımına ilişkin bir araştırma*. Yayınlanmamış Yüksek Lisans Tezi, Ankara: Ankara Üniversitesi Eğitim Bilimleri Enstitüsü.
- Bangert-Drowns, R., Kulik, C., Kulik, J. ve Morgan, M. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research*, 61, 213-238.
- Beyreli, L. ve Arı, G. (2009). Yazma performansını değerlendirmede çözümleyici puanlama yönergesi kullanımı-değerlendirmeciler arası uyum araştırması. *Kuram ve Uygulamada Eğitim Bilimleri*, 9(1), 85-125.
- Bolton, F. (2006). Rubrics and adult learners: Andragogy and assessment. *Assessment Update*, 18(3), 5-6.
- Boston, C. (2002). An introduction to performance assessment scoring rubrics. İçinde *Understanding Scoring Rubrics: A Guide for Teachers*. ERIC Clearinghouse on Assessment and Evaluation. ERIC Publications, 071.
- Butler, D. ve Winne, P. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*. 65(3), 245-281.
- Chase, C. I. (1999). *Contemporary assessment for educators*. New York: Longman.
- Çetin, Y. (2011). Kompozisyon değerlendirmesinde değerlendiricilerin güvenilirliği: Analitik-holistik, analitik-analitik, holistik-holistik. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 8(16), 471-486.
- Cyr, P. R., Smith, K. A., Broyles, I. L. ve Holt, C. T. (2014). Developing, evaluating and validating a scoring rubric for written case reports. *International Journal of Medical Education*, 5, 18-23.
- DelleBovi, B. M. (2012). Literacy instruction: From assignment to assessment. *Assessing Writing*, 17, 271–292.
- East, M. (2009). Evaluating the reliability of a detailed analytic scoring rubric for foreign language writing. *Assessing Writing*, 14, 88–115.
- Elbow, P. (2000). *Everyone can write essays toward a hopeful theory of writing and teaching writing*. New York: Oxford University Press.
- Erman Aslanoğlu, A. ve Kutlu, Ö. (2003) Öğretimde sunu becerilerinin değerlendirilmesinde dereceli puanlama anahtarı (rubrik) kullanılmasına ilişkin bir araştırma. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 36,(1-2), 25-36.
- Fehring, H. (2005). Critical, analytical and reflective literacy assessment: Reconstructing practice. *Australian Journal of Language and Literacy*, 28(2), 95–113.
- Goodrich, A. H. (2005). Teaching with rubrics: The good, the bad, and the ugly. *Teaching*, 53(1), 27-31.
- Goodrich Andrade, H. (2001). The effects of instructional rubrics on learning to write. *Current Issues in Education*, 4 (4). Available: <http://cie.ed.asu.edu/volume4/number4/>
- Goodrich A. H. (2000). What do we mean by using rubrics to promote thinking and learning? *Educational Leadership*, 57(5), 13-18.
- Goodrich, H. (1996). *Student self assessment: At the intersection of metacognition and authentic assessment*. Yayınlanmamış Doktora Tezi, Cambridge, MA: Harvard University.
- Goodrich Andrade, H. ve Boulay, B. A. (2003) Role of rubric-referenced self-assessment in learning to write. *The Journal of Educational Research*, 97(1), 21-30.
- Gunning, T. G. (2006). *Assessing and correcting reading and writing difficulties* (3th ed.). Boston: Pearson Education Inc.

Dereceli Puanlama Anahtarı Kullanan Öğrencilerin Başarıları Üzerine Bir Araştırma

- Huerta, M., Lara-Alecio, R., Tong, F. ve Irby, B. J. (2014). Developing and validating a science notebook rubric for fifth-grade non-mainstream students. *International Journal of Science Education*, <http://dx.doi.org/10.1080/09500693.2013.879623>
- Jonsson, A. ve Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130-144.
- Kan, A. (2007). An alternative method in the new educational program from the point of performance-based assessment: Rubric scoring scales. *Educational Sciences: Theory & Practice*, 7(1), 144-152.
- Koutsoftas, A. D. ve Gray, S. (2012). Comparison of narrative and expository writing in students with and without language-learning disabilities. *Language, Speech, and Hearing Services in Schools*, 43, 395-409.
- Kutlu, Ö., Doğan, C. D. ve Karakaya, İ. (2014). *Ölçme ve değerlendirme performans ve portfolyoya dayalı durum belirleme*. Ankara: PegemAkademi.
- Mason, C. Y. ve Steedly, K. S. (2006). Lessons and rubrics for arts integration. *TEACHING Exceptional Children Plus*, 3(1) Article 1. Retrieved [date] from <http://escholarship.bc.edu/education/tecplus/vol3/iss1/art1>
- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25). Retrieved March 3, 2004 from <http://PAREonline.net/getvn.asp?v=7&n=25>
- Moskal, B. M. (2000). Scoring rubrics: what, when, how? *Practical Assessment, Research and Evaluation*, 7 (3). Retrieved from August 20, 2004 <http://PAREonline.net/getvn.asp?v=7&n=3>.
- Oakleaf, M. (2008). Dangers and opportunities: A conceptual map of information literacy assessment approaches. *Portal: Libraries and the Academy* 8(3), 233-253.
- Özdemir, S. M. (2010). İlköğretim öğretmenlerinin alternatif ölçme ve değerlendirme araçlarına ilişkin yeterlikleri ve hizmetiçi eğitim ihtiyaçları. *Türk Eğitim Bilimleri Dergisi*, 8(4), 787-816. metin içinde yok
- Panadero, E. ve Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review*, 9, 129-144.
- Petkov, D. ve Petkova, O. (2006). Development of scoring rubrics for IS projects as an assessment tool. *Issues in Informing Science and Information Technology*, 3, 499-510.
- Popham, W. J. (1997). What's wrong -and what's right -with rubrics. *Educational Leadership*, 72-75.
- Reitmeier, C. A. ve Vrchota, D. A. (2009). Self-assessment of oral communication presentations in food science and nutrition. *Journal of Food Science Education*, 8(4), 88-92.
- Reynolds-Keefer, L. (2010). Rubric-referenced assessment in teacher preparation: An opportunity to learn by using. *Practical Assessment Research & Evaluation*, 15(8). Retrieved from <http://pareonline.net/getvn.asp?v=15&n=8>.
- Rezaei, A. R. ve Lovorn, M. (2010). Reliability and validity of rubrics for assessment through writing. *Assessing Writing*, 15, 18-39.
- Rhodes, T. (2010). Since we seem to agree, why are the outcomes so difficult to achieve? *New Directions for teaching and Learning*, 121, 13-21.
- Scharf, D. Elliot, N., Huey, H. A., Briller, V. ve Joshi, K. (2007). Direct assessment of information literacy using writing portfolios. *Journal of Academic Librarianship*, 33(4), 462-477.
- Tierney, R. ve Simon, M. (2004). What's still wrong with rubrics: Focusing on the consistency of performance criteria across scale levels. *Practical Assessment, Research and Evaluation*, 9 (2). Retrieved from December 12, 2005 from <http://PAREonline.net/getvn.asp?v=9&n=2>
- Tompkins, G. E. (2008). *Teaching writing: Balancing process and product*. Upper Saddle River, NJ: Merrill Prentice Hall.
- Turgut, M. F. ve Baykul, Y. (2010). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem AkademiYayını.
- van Helvoort, A. A. J. (2012). Perspectives on how adult students in information studies use a scoring rubric for the development of their information literacy skills. *The Journal of Academic Librarianship*, 38(3), 165-171.
- Wiggins, G. (1989). A true test toward more authentic and equitable assessment. *Phi Delta Kapan*, 70, 703-713.