

PERFORMANS DEĞERLENDİRME: PERFORMANS GÖREVLERİNİN HAZIRLANMASI VE PUANLANMASI

Arş. Gör. Sevilay KAPLAN*

ÖZET

Yenilenen ilköğretim okulları I.-V sınıf öğretim programında klasik değerlendirme yaklaşımlarının yerine öğrenci merkezli değerlendirme yaklaşımlarının kullanılması önerilmektedir. Fakat yeni programın işleyişine ilişkin yapılan araştırmalarda öğretmenlerin yenilenen ilköğretim müfredatının öngördüğü öğrenci merkezli değerlendirme yaklaşımları hakkında bilgileri olmadığı, verilen hizmet içi eğitimin de yeterli olmadığı saptanmıştır. Bu araştırma, öğretmenlere ve araştırmacılara, üst düzey becerilerin ölçülmesinde kullanılan ve öğrenci merkezli değerlendirme yaklaşımlarından biri olan performans değerlendirmelerine ilişkin bilgi sunmak amacıyla gerçekleştirilmiştir. Problem çözme, eleştirel düşünme, iletişim becerileri gibi becerilerin klasik değerlendirme yaklaşımlarıyla ölçülmesi zor hatta belli durumlarda imkânsızdır. Performans değerlendirmeleri, temel olarak, performans görevleri ve puanlama olmak üzere iki bölümden oluşmaktadır. Performans görevleri, eğitim hedefleri, süre, maliyet, gerekli materyaller göz önünde bulundurularak hazırlanmalıdır. Performans değerlendirmede kontrol listeleri ve dereceleme ölçekleri kullanılmaktadır. Dereceleme ölçekleri sayısal, grafiksel ve betimsel olmak üzere üçe ayrılırlar. Betimsel dereceleme ölçeklerine, dereceli puanlama anahtarı (rubric) adı verilmektedir. Dereceli puanlama anahtarları çeşitli aşamalardan geçerek hazırlanmaktadır. Dereceli puanlama anahtarının puanlanmasında, holistik ya da analitik puanlama kullanılır. Bu makalede performans değerlendirme yaklaşımının hangi amaçlarla yapıldığı, eğitimdeki yeri, performans ödevi oluşturma aşamaları ve bu ödevlerin puanlanması ele alınmış ve ilgili literatür çerçevesinde incelenmiştir.

Anahtar kelimeler: Performans değerlendirme, dereceli puanlama anahtarı, performans ölçütleri.

Abstract

In the revised primary school education curriculum, the use of alternative assessment approaches has been recommended alongside the traditional ones. However, in a recent research which has been conducted about the implementation of new curriculum, it has been found out that teachers do not know anything about the alternative assessment approaches and that the provided in-service training is not adequate. Skills such as problem solving, critical thinking, communicative skills are

* Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Eğitim Bilimleri Bölümü

difficult to measure through traditional methods of assessment, which is why the alternative assessment methods need to be applied in the assessment of such skills. The aim of this article is to provide information to teachers and researchers about performance assessment which is used to measure higher level thinking skills. It is one of the methods of alternative assessment. It includes performance tasks and scoring rubrics. Performance tasks need to be well-constructed and scored with rubrics in order to obtain accurate and reliable measures. In this article, the aims of performance assessment, the significance of it in education, the steps of constructing and scoring performance tasks have been studied in the light of related literature review.

Keywords: Performance assessment, rubric, performance criteria.

1. GİRİŞ

Son yıllarda dünyada eğitim ve öğretimin kalitesini arttırmak için çeşitli yenileşme hareketleri başlamıştır. Ülkemiz de bu yenileme hareketlerinden etkilenmiş, davranışçı yaklaşımdan yapılandırmacı yaklaşıma geçişi içeren bir sürece girmiş ve 1-5 ilköğretim programı yapılandırmacı yaklaşıma göre yeniden düzenlenmiştir. Yenilenen programda üst düzey becerileri geliştiren bir eğitim öğretim süreci önerilmektedir. Bu becerilerin izlenmesi ve değerlendirilmesinde, geleneksel yöntemler yetersiz kalmaktadır. Bu yetersizlik nedeniyle, değerlendirme uygulamalarında geleneksel yöntemlerin yanı sıra öğrenci merkezli ölçme ve değerlendirme yaklaşımlarının kullanılması da öngörülmektedir.

Yenilenen ilköğretim programına göre, programda sadece ürün değil, öğrencilerin öğrenme süreçleri değerlendirilir. Değerlendirme, öğrencilerin ne bilmediklerini değil, neyi bildiklerini görmek ve sahip oldukları becerileri, günlük yaşamda kullanma ve uygulayabilmelerine katkıda bulunan bir araçtır. Program, her öğrencinin kendini farklı yansıtabileceği düşüncesiyle değişik değerlendirme araç ve yöntemlerini kullanmayı önermektedir. Bu amaçla yeni öğretim programı değerlendirmede, öğretmenlerin halen kullandıkları klasik ölçme araçları (çoktan seçmeli, doğru - yanlış, eşleştirmeli testler, yazılı yoklamalar vb.) yanında, süreci değerlendirmek için, performans değerlendirmeleri yapılmasını da öngörmektedir (MEB, 2005).

Yenilenen ilköğretim programının uygulanmasına dönük çok sayıda araştırma yapılmıştır. Bu araştırmalarda, öğretmenlerin öğrenci merkezli ölçme ve değerlendirmeyi tam olarak ifade edemedikleri, fakat açıklamalarında bu konu ile ilgili çabaları olduğu gözlenmiştir (Aydın, 2005). Öğretmenlerin yenilenen programın ölçme ve değerlendirme boyutuna ilişkin eğitim gereksinimi duydukları; gözlem, çalışma dosyası, tartışma, deneyler, projeler, araştırma kâğıtları, öğrenci

ürün dosyası ve performans değerlendirme gibi ölçme araçlar ve yaklaşımlarına yer verildiği, ancak şu anda Milli Eğitim sistemi içerisinde yer alan öğretmenlerimizin çoğunluğunun bu ölçme araçlarını kullanmayı bilmedikleri ve bu nedenle öğretmenlerin hizmet içi eğitimden geçmesi gerektiği saptanmıştır (Gözütok ve diğerleri, 2005; Merter, 2005; Yaşar ve diğerleri, 2005).

Yapılan araştırmalara göre, yenilenen ilköğretim müfredatının öngördüğü öğrenci merkezli değerlendirme yaklaşımları hakkında bilgileri olmadığı, verilen hizmet içi eğitimin de yeterli olmadığı saptanmıştır (Aydın, 2005; Gözütok ve diğerleri, 2005; Merter, 2005; Sever, 2005; Yaşar ve diğerleri, 2005). Bu nedenle araştırma, öğretmenlere ve araştırmacılara, üst düzey becerilerin ölçülmesinde kullanılan ve öğrenci merkezli değerlendirme yaklaşımlarından biri olan performans değerlendirmelerine ilişkin bilgi sunmak amacıyla gerçekleştirilmiştir.

NEDEN PERFORMANS DEĞERLENDİRME?

Arter ve McTighe (2000)'e göre değerlendirme yöntemleri iki kategoride incelenebilir. Bunlar: (1) seçmeli cevap/kısa cevap gerektiren ve (2) yapılandırılmış cevap gerektiren yöntemlerdir. Öğrenciler için önemli olan eğitim amaçlarından birisi anlamlı gerçek yaşam durumlarına, bilgisini ve akıl yürütme becerisini uygulayabilmesidir. Gerçek yaşam uygulamalarını görmek istediğimizde seçmeli cevap/kısa cevap yöntemleri yetersiz kalmaktadır. Bu becerilerin ölçülmesinde yapılandırılmış cevap yöntemi kullanılmaktadır. Yapılandırılmış cevaplar öğrencinin bir soruya yazılı olarak uzun bir cevap hazırlaması istendiği yazılı sınavlarının kullanılması, aynı zamanda somut bir ürünün üretildiği (yazılı, görsel ve üç boyutlu) performans değerlendirmelerini, öğrencinin ne anladığını ve yapabildiklerini gösteren uygulamaları (konuşma, grupla çalışma, bir enstrüman çalma) içermektedir. Çünkü yapılandırılmış cevap gerektiren değerlendirmeler, bir soruya cevap verirken veya bir görevi tamamlarken, sahip olunan bilgi ve becerilerin geri çağırılması veya farkına varılmasından ziyade, bunların kullanılmasını ve organize edilmesini gerektirmektedir.

Performans değerlendirmeleri, öğrencilerin, bilgi ve becerilerini göstermelerini sağlayan ve bir cevap veya bir ürün yarattıkları değerlendirmelerdir (Airasian, 1994). Öğrencinin yapabilecekleri ve bildiklerini tanımlayan değerlendirme olarak bilinmektedir. (Marzona, Pickering, McTighe, 1993). Performans değerlendirme, öğrencilerin edindikleri bilgileri ve becerileri kullanabilecekleri, zihinsel ve duyuşsal yeteneklere alışkanlıklarını uygulayabilecekleri, yaşantılara ilişkin anlayışlarını, yaklaşımlarını gösterebilecekleri durumlarda yapılan değerlendirmelerdir (Kutlu, 2002).

Mehrens (1992), performans değerlendirmenin önem kazanma nedenini, çoktan seçmeli testlerin bazı becerileri yansıtmada yetersiz kalması, bilişsel psikolojinin etkisi ve geleneksel testlerin olumsuz etkisi olmak üzere üç temel

nedene dayandırarak açıklamaktadır. Performansa dayalı durum belirlemeden yana olanlar, kağıt kaleme dayalı çoktan seçmeli ve doğru yanlış gibi sınıflama gerektiren testlerin sadece belleği çalıştırdığı, öğrencinin problem çözme, sentez yapma ve bağımsız düşünebilme becerilerini kullanmayı gerektirmediğini savunmaktadırlar. Bilişsel psikologlar ise, öğrencinin hem içeriksel bilgi, hem de işlemsel bilgiye sahip olması gerektiğine ve kavramaya dayalı ödevlerin her iki bilgi türünü gerekli kıldığını ileri sürmektedirler. Öğrencinin sahip olduğu süreç bilgileri, seçmeli testlerle ölçülemeyeceğinden dolayı, performansa yönelik değerlendirmelere gerek duyulduğu üzerinde durmaktadırlar. Eğitimcilerin birçoğu ise, büyük oranda şans başarısına bağlı sonuçlar içeren çoktan seçmeli testlerin öğretmenin öğrettiklerine sürekli bir biçimde etkide bulunacağı konusu üzerinde durduğundan, performansa dayalı durum belirlemenin geleneksel yaklaşıma göre çok daha değerli eğitimsel hedeflere yönelik olduğu kanısı taşımaktadırlar (akt. Popham, 2000).

Geleneksel değerlendirme yöntemleri; bilgi, kavrama ve uygulamanın bazı türleri için geçerli ve güvenilir olan etkili araçlardır. Fakat, karmaşık düşünce becerisi, zihin alışkanlıkları (habit of mind) ve sosyal becerilerin değerlendirilmesinde, performans testleri –eğer iyi yapılandırılmışsa- daha iyi işlemektedir (Kubizyn ve Borich, 1996). Öğretmenler, öğrenci başarıları ile ilgili bilgi toplarken, geleneksel değerlendirme yöntemlerini kullanamadıkları şarkı söyleme, kompozisyon yazma, dans etme, tiyatro yapma gibi bazı durumlarda, performans ürünlerini gözlemlemek ve değerlendirmek durumundadır. Performansa dayalı değerlendirme, öğrencilerin, gerçek hayatta ne yapabileceklerini göstermelerine izin vermektedir. Kağıt – kalem testlerinin büyük çoğunluğu öğrencinin cevaba ulaşırken kullandığı zihinsel yöntemin çok azını görmeyi sağlar (Airasian, 1994). Öğretmenler performans testlerini, üst düzey becerileri ölçmenin yanında kendini yönetme, başkalarıyla çalışma, sosyal farkındalık gibi bilişsel olmayan hedefler için de kullanmaktadırlar (Kubizyn ve Borich, 1996).

Cevaplar arasından birini seçmesini gerektiren doğru-yanlış testleri veya çoktan seçmeli testlere hiç benzemeyen performans değerlendirmeleri, öğrencinin doğru cevabı üretmesini gerektirmektedir. Örneğin kompozisyon yazmayla ilgili performans değerlendirmeleri, noktalama veya dilbilgisi ile ilgili çoktan seçmeli sorulara kolayca yanıt vermesinden ziyade, öğrencinin yazma işini eyleme dönüştürmesini gerektirir. Performans ödevleri bir ürün, performans veya yazılı çalışmalar olabilir. Sözlü sunumlar, fen projeleri, araştırma projeleri, müzik performansı, açık uçlu matematik problemleri performans değerlendirme örneklerinden bazılarıdır (Perlman, 2003).

Çoktan seçmeli, doğru yanlış testleri gibi testler sınıf içinde bir kez kullanılabilmesine rağmen, performans değerlendirme materyalleri tekrar tekrar kullanılabilirler. Performans değerlendirme materyalleri tanılayıcıdır ve

performansın farklı kısımlarına odaklanır. Böylece öğretmenin, öğrenciyi performansın her kademesini gözlemlemesini, değerlendirmesini ve düzeltmesini gerektirir. Öğretmenler performans değerlendirmeleri ile, öğrencinin zayıf ve güçlü yönlerini, özel davranışlar bakımından tanımlayabilir (Airasian, 1994).

Kubizyn ve Borich (1996)'e göre performans değerlendirmeleri dört çeşit öğrenci başarısının değerlendirilmesinde kullanılır:

1. Ürünler (şirler, düz yazılar, grafikler, gösteriler, resimler, haritalar vb.)
2. Karmaşık bilişsel süreçler (bilgiyi edinme, organize etme ve kullanma)
3. Gözlenebilir performans (dans, jimnastik gibi fiziksel hareketler, sözlü sunumlar, yönerge uygulamaları, kurbağanın incelenmesi gibi işlem basamaklarının takip edildiği durumlar, mikroskopla inceleme yapmak gibi uzmanlık gerektiren malzemelerin kullanıldığı durumlar vb.)
4. Sosyal beceriler ve zihin alışkanlıkları (kararlı olmak, işbirliği yapabilmek gibi zihinsel ve davranışsal alışkanlıklar ve sunum becerileri).

PERFORMANS GÖREVLERİNİN HAZIRLANMASI

Stiggins (1991)'e göre performans değerlendirmelerinin kaliteli bir şekilde yapılması için iki anahtar unsur vardır. Bunlar, belirlenen davranış ve ürün örneklemleri için uygun yöntemin seçilmesi ve değerlendirme kararları için esas alınan performans ölçütlerinin açık ve anlaşılır bir şekilde belirlenmesidir.

Performans değerlendirmelerinin ilk aşaması, eğitime odaklanmış bilgi, beceri, zihin alışkanlıkları ve öğrenme göstergelerini açıkça belirten bir amaç listesi hazırlamaktır (Kubizyn ve Borich, 1996). Performans değerlendirmelerinde amaç belirlendikten sonra, bu amaca yönelik performans ölçütleri hazırlanmalıdır. Performans ölçütü, bir performansı sergilerken, bireyin göstermesi gereken belirli davranışlar veya bir ürün oluşturulacaksa ürünün sahip olması gereken özelliklerdir. Bu ölçütler öğrencilerle birlikte sınıf içinde tartışılmalı ve belirlenmelidir. Bu tür değerlendirmelerin, klasik değerlendirme yöntemlerden ayırıcı özelliği, öğrencinin nasıl değerlendirileceğini bilmesidir. Değerlendirme ölçütlerinin bilinmesi öğrencilerin değerlendirilmeye motive edilmesini sağlamaktadır.

Airasian (1994)'a göre performans ölçütleri oluşturulurken aşağıdaki öneriler dikkate alınmalıdır:

1. Değerlendirilecek olan ödevi ya da performansı bütünüyle tanımlamalı ve öğretmenin, önce performansı kendisi gerçekleştirmeli veya gerçekleştirdiğini hayal etmeli, kendine aşağıdaki soruları sormalıdır:

- Bu ödevi tamamlamak için ne yapmak zorundayım?
- İzlemem gereken basamaklar nelerdir?
- Ürünün ya da performansın en önemli boyutu hangisidir?

2. Performans veya ürün ölçütlerinden önemli olanları listelenmelidir.

- Görevi doğru bir şekilde tamamlamak için hangi davranışlar ya da özellikler daha önemlidir.
- Eğitimde hangi davranışlara vurgu yapılmalıdır?

3. Performans ölçütlerinin tümünün, öğrencinin, performansı gerçekleştirmesi sırasında gözlenebilmesi için bu ölçütlerin sınırlı tutulması gerekmektedir. Bir grup öğretmen ile birlikte, belirlenen performans görev/ödevlerindeki önemli davranışlar tartışılmalıdır.

4. Performans ölçütleri, gözlenebilen öğrenci davranışları şeklinde ifade edilmelidir.

5. Performans ölçütünün anlamını gölgeleyen 'doğru bir şekilde', 'uygun bir şekilde' gibi belirsiz kelimeler kullanılmamalıdır. Çünkü, bu belirsiz ölçütlerin yorumlanması kişiden kişiye değişir.

Performans değerlendirmelerinde öğrenciye verilen sözlü ve yazılı açıklamalar öğrencinin anlayabileceği şekilde ifade edilmelidir. Verilen ödevde başlamadan önce açıklamalarla ilgili soru sorma fırsatı verilmelidir (Moscals, 2003). Performans değerlendirmelerinde, gerçek yaşama benzer ortam ve durumlar karşısında, öğrencilerin sahip oldukları bilgi, beceri ve yetenekleri sergilemelerine izin verici ödev/görev yaratılmalıdır. Bu ödevler kitaplardan, röportajlardan, gazetelerden esinlenerek yazılabilir. Performans ödevi, konu kapsamıyla ilgili önemli problemler, kavramlar ve ürünler üzerine odaklanmalıdır (Kubizyn ve Borich, 1996).

- **Zaman sınırı**

Bazı performans test maddelerinin yanıtları, sadece birkaç dakikada oluşturulabilir. Bazıları ise, yoğun çaba gerektirebilir ve haftalarca hatta aylarca zaman alabilir. Bu nedenle ödevin en geç sonuçlanacağı tarih açıkça belirtilmelidir. Tüm öğrenciler değerlendirilmeden önce süre sınırlamasından haberdar olmalıdırlar (Haladyna, 1997). Öğrencinin yaptığı yanlışları düzeltmesine izin verilip verilmeyeceği, diğer öğrencilerle yardımlaşıp yardımlaşmayacağı önceden belirlenmelidir. Okul dışındaki yaşamda, nasıl ki, bir matematikçi hatalar yapıp, daha sonra o hatasını düzeltiyorsa; bir gazeteci taslak yazısını yazdıktan sonra, o yazıyı tekrar gözden geçiriyorsa, gerçek hayata benzer durumlar karşısında, bir öğrencinin performansı değerlendirilirken, gerekli düzeltmeleri yapmasına izin verilmelidir. (Kubizyn ve Borich, 1996).

- **Dayanışma:**

Dayanışma, öğrencilerin birbirlerine yardım etmelerini veya yol göstermelerini gerektirir. Dayanışmanın ilgi çekici veya istek uyandırıcı olup olmadığına karar verilmelidir. Performans değerlendirmelerinde dayanışmanın olması doğal bir

performansın farklı kısımlarına odaklanır. Böylece öğretmenin, öğrenciyi performansın her kademesini gözlemlemesini, değerlendirmesini ve düzeltmesini gerektirir. Öğretmenler performans değerlendirmeleri ile, öğrencinin zayıf ve güçlü yönlerini, özel davranışlar bakımından tanımlayabilir (Airasian, 1994).

Kubizyn ve Borich (1996)'e göre performans değerlendirmeleri dört çeşit öğrenci başarısının değerlendirilmesinde kullanılır:

1. Ürünler (şiiirler, düz yazılar, grafikler, gösteriler, resimler, haritalar vb.)
2. Karmaşık bilişsel süreçler (bilgiyi edinme, organize etme ve kullanma)
3. Gözlenebilir performans (dans, jimnastik gibi fiziksel hareketler, sözlü sunumlar, yönerge uygulamaları, kurbağanın incelenmesi gibi işlem basamaklarının takip edildiği durumlar, mikroskopla inceleme yapmak gibi uzmanlık gerektiren malzemelerin kullanıldığı durumlar vb.)
4. Sosyal beceriler ve zihin alışkanlıkları (kararlı olmak, işbirliği yapabilmek gibi zihinsel ve davranışsal alışkanlıklar ve sunum becerileri).

PERFORMANS GÖREVLERİNİN HAZIRLANMASI

Stiggins (1991)'e göre performans değerlendirmelerinin kaliteli bir şekilde yapılması için iki anahtar unsur vardır. Bunlar, belirlenen davranış ve ürün örneklemleri için uygun yöntemin seçilmesi ve değerlendirme kararları için esas alınan performans ölçütlerinin açık ve anlaşılır bir şekilde belirlenmesidir.

Performans değerlendirmelerinin ilk aşaması, eğitime odaklanmış bilgi, beceri, zihin alışkanlıkları ve öğrenme göstergelerini açıkça belirten bir amaç listesi hazırlamaktır (Kubizyn ve Borich, 1996). Performans değerlendirmelerinde amaç belirlendikten sonra, bu amaca yönelik performans ölçütleri hazırlanmalıdır. Performans ölçütü, bir performansı sergilerken, bireyin göstermesi gereken belirli davranışlar veya bir ürün oluşturulacaksa ürünün sahip olması gereken özelliklerdir. Bu ölçütler öğrencilerle birlikte sınıf içinde tartışılmalı ve belirlenmelidir. Bu tür değerlendirmelerin, klasik değerlendirme yöntemlerden ayırıcı özelliği, öğrencinin nasıl değerlendirileceğini bilmesidir. Değerlendirme ölçütlerinin bilinmesi öğrencilerin değerlendirilmeye motive edilmesini sağlamaktadır.

Airasian (1994)'a göre performans ölçütleri oluşturulurken aşağıdaki öneriler dikkate alınmalıdır:

1. Değerlendirilecek olan ödevi ya da performansı bütünüyle tanımlamalı ve öğretmenin, önce performansı kendisi gerçekleştirmeli veya gerçekleştirdiğini hayal etmeli, kendine aşağıdaki soruları sormalıdır:

- Bu ödevi tamamlamak için ne yapmak zorundayım?
- İzlemem gereken basamaklar nelerdir?
- Ürünün ya da performansın en önemli boyutu hangisidir?

2. Performans veya ürün ölçütlerinden önemli olanları listelenmelidir.

- Görevi doğru bir şekilde tamamlamak için hangi davranışlar ya da özellikler daha önemlidir.
- Eğitimde hangi davranışlara vurgu yapılmalıdır?

3. Performans ölçütlerinin tümünün, öğrencinin, performansı gerçekleştirmesi sırasında gözlenebilmesi için bu ölçütlerin sınırlı tutulması gerekmektedir. Bir grup öğretmen ile birlikte, belirlenen performans görev/ödevlerindeki önemli davranışlar tartışılmalıdır.

4. Performans ölçütleri, gözlenebilen öğrenci davranışları şeklinde ifade edilmelidir.

5. Performans ölçütünün anlamını gölgeleyen 'doğru bir şekilde', 'uygun bir şekilde' gibi belirsiz kelimeler kullanılmamalıdır. Çünkü, bu belirsiz ölçütlerin yorumlanması kişiden kişiye değişir.

Performans değerlendirmelerinde öğrenciye verilen sözlü ve yazılı açıklamalar öğrencinin anlayabileceği şekilde ifade edilmelidir. Verilen ödevde başlamadan önce açıklamalarla ilgili soru sorma fırsatı verilmelidir (Moscal, 2003). Performans değerlendirmelerinde, gerçek yaşama benzer ortam ve durumlar karşısında, öğrencilerin sahip oldukları bilgi, beceri ve yetenekleri sergilemelerine izin verici ödev/görev yaratılmalıdır. Bu ödevler kitaplardan, röportajlardan, gazetelerden esinlenerek yazılabilir. Performans ödevi, konu kapsamıyla ilgili önemli problemler, kavramlar ve ürünler üzerine odaklanmalıdır (Kubizyn ve Borich, 1996).

- **Zaman sınırı**

Bazı performans test maddelerinin yanıtları, sadece birkaç dakikada oluşturulabilir. Bazıları ise, yoğun çaba gerektirebilir ve haftalarca hatta aylarca zaman alabilir. Bu nedenle ödevin en geç sonuçlanacağı tarih açıkça belirtilmelidir. Tüm öğrenciler değerlendirilmeden önce süre sınırlamasından haberdar olmalıdırlar (Haladyna, 1997). Öğrencinin yaptığı yanlışları düzeltmesine izin verilip verilmeyeceği, diğer öğrencilerle yardımlaşıp yardımlaşmayacağı önceden belirlenmelidir. Okul dışındaki yaşamda, nasıl ki, bir matematikçi hatalar yapıp, daha sonra o hatasını düzeltiyorsa; bir gazeteci taslak yazısını yazdıktan sonra, o yazıyı tekrar gözden geçiriyorsa, gerçek hayata benzer durumlar karşısında, bir öğrencinin performansı değerlendirilirken, gerekli düzeltmeleri yapmasına izin verilmelidir. (Kubizyn ve Borich, 1996).

- **Dayanışma:**

Dayanışma, öğrencilerin birbirlerine yardım etmelerini veya yol göstermelerini gerektirir. Dayanışmanın ilgi çekici veya istek uyandırıcı olup olmadığına karar verilmelidir. Performans değerlendirmelerinde dayanışmanın olması doğal bir

durumdur. Öğretmenler dayanışmayı bazı durumlarda engellememelidirler. Burada önemli olan, öğrenciye, kendi cevaplarını/ürünlerini oluşturması gerektiği ve her öğrencinin kendi final performans ya da ürününden sorumlu olacağı belirtilmesidir (Haladyna, 1997).

- **İşbirliği:**

Gerçek yaşam performans ödevleriyle ve işbirlikli öğrenmenin giderek artan kullanımıyla, öğrenciler, bir proje üzerinde başkalarıyla birlikte çalışmalarını için cesaretlendirilmelidir. Öğretmen performans değerlendirmelerinde işbirliği yapıp yapılmayacağına karar vermelidir. Eğer yapılacaksa, puanlamanın nasıl yapılacağı belirlenmelidir. İşbirliği ile yapılan çalışmalarda yaşanan ortak sorun herkesin eşit derecede iş yapmamasıdır. Bir veya iki öğrenci genellikle çok çalışmaktadır ve bu öğrencilerin, diğerlerine göre ürüne katkısı daha fazladır. Fakat işbirliği ekibine verilen notlar herkes için eşit olmaktadır (Haladyna, 1997).

- **Araç kullanma:**

Performans değerlendirirken kullanılacak araçlar önceden belirlenmelidir. Araçlar değerlendirme uygulamalarını tamamlamaya uygun olmalıdır. Bu uygulamalarda öğrencilerin kütüphane kaynakları, bilgisayar, hesap makinesini veya başka araçları kullanmaları gerekebilir. Bu nedenle öğretmen, öğrencilerin hangi araçlara gereksinim duyacaklarını açıklamalı ve uygun araçları temin etmelidir.

PERFORMANS GÖREVLERİNİN PUANLANMASI:

Mertler (2001)'e göre performans değerlendirmelerinde iki tür araç kullanılmaktadır. Bunlar kontrol listeleri ve derecelendirme ölçekleridir.

- **Kontrol listesi:**

Performans ölçütleri bazen kontrol listesi formunda yazılabilir. Puanlayıcı, her bir performans ölçütünü, yaratılan ürünlerin veya gözlenen performansın karşılayıp karşılamadığını belirler. Kontrol listeleri öğrenci performanslarının detaylı kayıtlarını içerir. Öğrencilerin, gelişimlerinin hangi noktasında olduklarını görmelerini sağlar. Çünkü, kontrol listesi, bir performansın özel tanımlanmış davranışları üzerine odaklanır. Aynı kontrol listesi, aynı öğrencilere farklı zamanlarda veya farklı öğrencilere uygulanabilir. Kontrol listesinin bazı dezavantajları da vardır. Önemli bir sınırlılığı, öğretmene her bir performans ölçütü için iki seçenek vermesidir. Kontrol listesi, öğrencinin performansı her iki uç noktanın arasında olsa bile, öğretmeni her bir performans ölçütü için kesin kararlar almaya zorlamaktadır. Kontrol listelerinin kullanımındaki ikinci bir sınırlılık ise öğrenci performansını tek bir puanla ifade etme güçlüğüdür. Bunun için öğretmenin bir kuralı veya yöntemi olmalıdır (Airasian,1994).

Kontrol listeleri, öğrencinin çalışmasını sunmadan önce, ürününü tekrar gözden geçirmesini sağlayıcı bir araç olarak kullanılabilir. Her madde genellikle bir soru şeklinde ifade edilir. Kontrol listeleri, öğrencinin, verilen ödevi yerine getirmesini sağlayan ve öğrencinin, ödevin dışına çıkmasını, sapmasını ve yüzeysel yapmasını engellemenin bir yoludur (Haladyna, 1997).

Kontrol listesi, evet-hayır; var-yok veya 1-0 şeklinde puanlanabilir. Bazı kontrol listeleri aynı zamanda öğrenene, performans ödevini yerine getirirken yaptığı hataların miktarını gösterir. Böyle durumlarda olumlu davranışlara +1; her hataya ise -1; performans ödevi gözlenmemişse 0 puan verilebilir (Kubizyn ve Borich, 1996).

• **Dereceleme ölçeği:**

Hem kontrol listesi hem de dereceleme ölçeği performans ölçüt takımı üzerine odaklanır. Dereceleme ölçeği kontrol listesine benzese de, aralarında önemli farklılıklar vardır. Bir kontrol listesi gözlemciye, yargıda bulunması için iki seçenek verirken, bir dereceleme ölçeği ikiden fazla seçenek verir. Dereceleme ölçeği, öğretmenin öğrenci performansına ilişkin daha hassas ayrımlar yapmasına izin verir. Dereceleme ölçeklerinin birçok türü vardır. En çok kullanılanları ise sayısal (numerical), betimsel (descriptive) ve grafiksel (graphic) dereceleme ölçekleridir (Airasian, 1994).

Tablo 1: Sayısal dereceleme ölçeği

Yönerge: Sözlü sunum esnasında, öğrencinin, aşağıda verilen davranışları hangi sıklıkta sergilediğini belirleyiniz. Öğrenci her zaman belirtilen davranışı gösteriyorsa 1'i; genellikle gösteriyorsa 2'yi; ara sıra gösteriyorsa 3'ü; hiç göstermiyorsa 4'ü işaretleyiniz.				
Fiziksel ifadeler:				
a)	Dik durdu ve yüzünü öğrencilere döndü.			
	1	2	3	4
b)	Ses tonuna göre, yüz ifadelerini değiştirdi.			
	1	2	3	4
			

Kaynak: Airasian, P. W. (1994). Classroom Assessment. New York: McGraw-Hill, s: 258

Tablo 2: Grafiksels dereceleme ölçeđi

Yönerge: Sözlü sunum yapılırken, öğrencinin, aşağıda verilen davranışları hangi sıklıkta sergilediğini çizgi üzerinde (X) işareti koyarak belirleyiniz.

1) Dik durdu ve yüzünü öğrencilere döndü.

her zaman	genellikle	ara sıra	hiçbir zaman
-----------	------------	----------	--------------

2) Ses tonuna göre yüz ifadelerini deđiştirdi.

her zaman	genellikle	ara sıra	hiçbir zaman
-----------	------------	----------	--------------

.....

Kaynak: Airasian, P. W. (1994). Classroom Assessment. New York: McGraw-Hill, s:258.

Betimsel dereceleme ölçeklerine dereceli puanlama anahtarı (rubric) adı verilmektedir. Dereceli puanlama anahtarı öğrenciden beklenenlerin listesini içermektedir (Airasian, 1994, Goodrich, 1997).

Dereceli puanlama anahtarlarının hazırlanması için çeşitli aşamalardan geçmesi gerekmektedir. Dereceli puanlama anahtarının hazırlanma aşamaları aşağıdaki gibidir:

1) **Amacın belirlenmesi:** Derecelenme anahtarı hazırlanırken ilk önce kullanım amacı belirlenmelidir.

2) **Ölçütlerin belirlenmesi:** Öğrencilerin ürün, süreç ve performansları ile sergilemeleri istenen spesifik gözlenebilir özellikler yani değerlendirme kriterleri, öğrencilerle birlikte tanımlanmalıdır. İyi yapılandırılmış dereceli puanlama anahtarı, öğrencilere, değerlendirme ölçütlerinin nasıl karşılanacağını gösterir. Dereceli puanlama anahtarları, daha iyi bir final ürünü üretebilmeleri için öğrencilere birer rehberdir. Böylece, öğrenciler, öğretmenlerinin ve akranlarının ön değerlendirmelerinden sonra ödevini yeniden gözden geçirip eksiklerini tamamlayabilirler (Arter & McTighe, 2000; Goodrich, 1997; Popham, 2000).

3) **Puanlama stratejisinin belirlenmesi:** Dereceli puanlama anahtarları, analitik dereceli puanlama anahtarı ve holistik dereceli puanlama anahtarı olmak üzere ikiye ayrılırlar. Hipotezler, yöntem, ürünler ve sonuçlar gibi bölümlere ayrılan

fen konularında kullanılan iki veya daha fazla ölçekli dereceli puanlama anahtarı, analitik dereceli puanlama anahtarlarına bir örnektir. Sadece bir ölçek kullanımını gerektiren dereceli puanlama anahtarlarına holistik adı verilmektedir. Holistik puanlama sisteminde öğrenci cevaplarının toplam niteliği, mükemmel, yeterli, sınırdan veya yetersiz şeklinde yargılanır (Perlman, 2003). Eğer genel bir karar verilecekse (notlandırma, seçme gibi), holistik puanlama kullanılması daha uygundur. Holistik puanlamada performansın tümüne ilişkin genel tek bir puan elde edilir. Ancak, amaç öğrencinin performanstaki güçlüklerini tanımamak ya da her bir performans ölçütünde yeterliliğini belirlemek ise, analitik puanlama tercih edilmelidir. Analitik puanlamada her bir performans ölçütü ayrı ayrı puanlanır ya da derecelendirilir (Airasian, 1994).

4 a) Holistik dereceli puanlama anahtarı kullanılacaksa, her bir özelliği kapsayan en iyi ve en kötü çalışmaları tam olarak açıklayan tanımlamaların yazılması.

4 b) Analitik dereceli puanlama anahtarı kullanılacaksa, her bir özellik için ayrı ayrı, en iyi ve en kötü çalışmaları tam olarak açıklayan tanımlamaların yazılması.

5 a) Holistik dereceli puanlama anahtarı kullanılacaksa, ortak özellikler için, en iyi ve en kötü çalışmalar arasında kalan diğer dereceleri tam olarak açıklayan tanımlamaların yazılması.

5 b) Analitik dereceli puanlama anahtarı kullanılacaksa, her bir özellik için ayrı ayrı, en iyi ve en kötü çalışmalar arasında kalan diğer dereceleri tam olarak açıklayan tanımlamaların yazılması.

6) Dereceli puanlama anahtarının yeniden gözden geçirilmesi: Dereceli puanlama anahtarı hazırlandıktan sonra tekrar gözden geçirilerek, hata yapıp yapılmadığına dair kontrol yapılmalıdır.

Tablo 3: Holistik dereceli puanlama anahtarı örneği.

Dereceler	Yazma becerileri dereceli puanlama anahtarı
6 Olağanüstü başarı	Bütünleşik ve iyi odaklanmış bir kompozisyon. Konu ve ana fikir açık. Konu dışına çıkılmamış. Giriş, gelişme ve sonuç bölümleri oldukça açık. Paragraflar arası geçişler farklı ve etkili. Yardımcı fikirler temel mantığı ve ana fikri destekliyor. Amaç kapsamlı olarak sık sık belirtilmiş (8-10 cümle). Teknik hatalar çok az ve seyrek.
5 Övgüye değer bir başarı	Genel olarak bir plana göre hareket edilmiş ve iyi organize edilmiş. Konu ve ana fikir genel olarak açık. Giriş ve gelişme bölümleri açık olarak verilmiş. Yardımcı fikirler genellikle farklı ve etkili. Kompozisyonun her bölümünde en azından bir amaç tam olarak açıklanmış (6-8 cümle). Teknik hatalar okumayı engellemese de birkaç küçük ve bir-iki büyük hata var.
4 Yeterli başarı	Konuya, ana fikre ve tüm plana hakim görünmekte fakat yazıya odaklanılmamış. Sonuç bölümünün bazı yerlerinde başarısız ve/veya birbirini tutmayan ifadeler kullanılmış. Geçişler tipik olarak mantıksal fakat ara sıra doğrudan ilişkiler ve açıklamalar yetersiz.
3 Başarıya ilişkin bazı kanıtlar var	Konu ve kompozisyon planı açıkça ifade edilememiş. Konu dışına taşmalar ve ayrıntılar okuyucunu kafasını karıştırıyor. Bazen geçişler kullanılmış. Sonuç bölümünde başarısız ve/veya birbirini tutmayan ifadeler kullanılmış. Anahtar öğeler beceriksizce kullanılmış veya değinilmemiş. Yardımcı fikirler arasında tutarsızlık var. Ana fikrin ayrıntıları sınırlı sayıda verilmiş (2-4 cümle). Bazı yerlerde okumayı engelleyen birtakım küçük ve büyük hatalar var.
2 Sınırlı sayıda başarı göstergesi var	Konu ve ana fikir açıkça ifade edilememiş. Ana fikri desteklemek için kullanılan yardımcı fikirler mantıklı değil. Okuyucunun anlamasını ciddi anlamda engelleyecek şekilde konu dışına çıkmış. Giriş ve sonuç bölümünde ana fikirden az miktarda söz edilmiş. Birkaç geçiş cümlesi kullanılmış. Destekleyici detaylar birkaç cümle ile ifade edilmiş fakat bu fikirler ana fikirle bağlantısız ve okuyucunun aklını karıştırıyor. Okumayı engelleyen teknik hataları içeriyor.
1 Asgari sayıda başarı göstergesi var	Ana fikri ve örgütsel bir planı yok. Çok sayıda gereksiz ayrıntı ve konu dışına taşmalar var. Giriş ve sonuç bölümlerinde ana fikre çok az rastlanıyor. Okumayı engelleyen birçok teknik hata var. Tamamlanmamış cümleler var.

Kaynak: Herman, J. L., Gearhart, M., Baker, E. (1994). Assessing Writing Portfolio: Issues in the Validity and Meaning of Scores. *Educational Measurement*, 1 (3), 201-224.

taşımaktadır. Performans değerlendirmeleri, temel olarak, performans görevleri ve puanlama olmak üzere iki bölümden oluşmaktadır. Performans görevleri, eğitim hedefleri, süre, maliyet, gerekli materyaller göz önünde bulundurularak hazırlanmalıdır. Performans değerlendirme kontrol listeleri ve derecelendirme ölçekleri kullanılmaktadır. Derecelendirme ölçekleri sayısal, grafiksel ve betimsel olmak üzere üçe ayrılır. Betimsel derecelendirme ölçeklerine, dereceli puanlama anahtarları (rubric) adı verilmektedir. Dereceli puanlama anahtarları çeşitli aşamalardan geçerek hazırlanmaktadır. Dereceli puanlama anahtarının puanlanmasında, holistik ya da analitik puanlama kullanılır. Eğer genel bir karar verilecekse (notlandırma, seçme gibi), holistik puanlama daha uygundur. Holistik puanlamada performansın tümüne ilişkin genel tek bir puan elde edilir. Ancak, amaç öğrencinin performansındaki güçlüklerini tanımlamak ya da her bir performans ölçütünde öğrencinin yeterliğini belirlemek ise, analitik puanlama tercih edilmelidir. Analitik puanlamada her bir performans ölçütü ayrı ayrı puanlanır ya da derecelendirilir. Sonuç olarak, performans değerlendirmeleri, çeşitli aşamalardan geçmeyi gerektirmektedir. Bu aşamaların yerine getirilmesi geçerli ve güvenilir ölçümler yapmayı sağlamaktadır.

KAYNAKÇA

- Airasian, P. W. (1994). **Classroom Assessment**. New York: McGraw-Hill.
- Arter, J. ve McTighe, J. (2000). **Scoring Rubrics in The Classroom**. California: Corwin Press.
- Aydın, F. (2005). **Öğretmenlerin Alternatif Ölçme Değerlendirme Konusundaki Düşünceleri ve Uygulamaları**. 14. Ulusal Eğitim Bilimleri Kongresi'nde sunulan bildiri, Denizli, Türkiye.
- Goodrich, H. (1997). Understanding Rubrics. **Educational Leadership**, 54 (4), 14-17.
- Gözütok, D., Ergün, E.Ö. ve Karacaoğlu, C. Ö., (2005). **İlköğretim Programlarının Öğretmen Yeterlilikleri Açısından Değerlendirilmesi**. Yeni İlköğretim Programlarını Değerlendirme Sempozyumu'nda sunulan bildiri, Kayseri, Türkiye.
- Haladyna, T. M. (1997). **Writing Test Items to Evaluate Higher Order Thinking**. USA: Allyn and Bacon
- Herman, J. L., Gearhart, M. ve Baker, E. (1994). Assessing Writing Portfolio: Issues In The Validity And Meaning Of Scores. **Educational Measurement**, 1 (3), 201-224.

- Kubizyn, T. ve Borich, G. (1996). **Educational Testing and Measurement: Classroom Application and Practice** (5th ed.). New York: Harper Collins.
- Kutlu, Ö. (2002). **Öğretmen Yetiştirme Programlarının Yeni Ölme ve Değerlendirme Yaklaşımlarını İçerme Düzeyi**. Uluslararası Katılımlı 2000'li Yıllarda I. Öğrenme ve Öğretme Sempozyumu'nda sunulan bildiri, İstanbul, Türkiye.
- Marzona, R. J., Pickering, D. Ve McTighe, J. (1993). **Assessing Student Outcomes**. Alexandria: Association for Supervision and Curriculum Development.
- MEB, TTKB (2005). **İlköğretim 1-5 sınıf Programları Tanıtım Kitapçığı**. Ankara.
- Mehrens, W. A. (1992). Using Performance Assessment For Accountability Purposes. **Educational Measurement: Issues and Practice**, 11 (1), 3-9.
- Merter, F. (2005). **Hayat Bilgisi ve Sosyal Bilgiler Dersi Programının Eğitim Sosyolojisi Bakış Açısından Olumlu ve Olumsuz Eleştirisi**. Yeni İlköğretim Programlarını Değerlendirme Sempozyumu'nda sunulan bildiri, Kayseri, Türkiye.
- Mertler, Craig A. (2001). Designing Scoring Rubrics For Your Classroom. **Practical Assessment, Research & Evaluation**, 7 (25), <http://PAREonline.net/getvn.asp?v=7&n=25> adresinden 3 Mart 2006 tarihinde ulaşılmıştır.
- Perlman, C. (2003). **Performance Assessment: Desining Appropriate Performance Task And Scoring Rubrics**. <http://www.eric.ed.gov/ERICDocs/data/ericdocs2.pdf> adresinden 05 Mayıs 2006 tarihinde ulaşılmıştır.
- Popham, J. W. (2000). **Educational Measurement**. Needham: Allyn & Bacon.
- Sever, S. (2005). **2004 Öğretim Programında Türkçe Öğretimi Anlayışı**. Yeni İlköğretim Programlarını Değerlendirme Sempozyumu'nda sunulan bildiri, Kayseri, Türkiye.
- Stiggins, R. J. (1991). Facing The Challenges Of A New Era Of Educational Assessment. **Applied Measurement in Education**, 4 (4), 263-273.
- Stiggins, R. ve Chappuis, J. (2005). Using Student-Involved Classroom Assessment to Close Achievement Gaps. **Theory Into Practice**, 44 (1), 11-18.
- Yaşar, Ş., Gültekin, M., Türkan, B., Yıldız, N. ve Girmen, P. (2005). **Yeni İlköğretim Programlarının Uygulanmasına İlişkin Sınıf Öğretmenlerinin Hazırbulunuşluk Düzeylerinin ve Eğitim Gereksinimlerinin Belirlenmesi**. Yeni İlköğretim Programlarını Değerlendirme Sempozyumu'nda sunulan bildiri, Kayseri, Türkiye.