



Performansa Dayalı Ölçümler: Teori ve Uygulama

Feral OGAN BEKİROĞLU¹

¹ Yrd. Doç. Dr., Marmara Üniversitesi, Atatürk Eğitim Fakültesi, OFMA Eğitimi Bölümü, İstanbul

Alındı: 16.08.2006

Düzeltildi: 05.10.2007

Kabul Edildi: 11.11.2007

The original language of article is Turkish (v5, n1, April 2008, pp.113-131)

ÖZET

Gelişen teknoloji ve toplumun ihtiyaçlarının da etkisiyle değişen bilimsel bilginin edinilmesi anlayışı müfredata entegre edildiğinde, kullanılan ölçme yöntemleri de bu anlayışı yansıtmalıdır. Bu bağlamda, öğrencilerin yüksek seviyede düşünme, akıl yürütme ve problem çözme becerilerinin ve kavramları anlamalarının ölçülmesi gündeme gelmiştir. Bu çalışmada, bütünlendirici yaklaşımın desteklediği ve öğrencilerin öğrenme süreci içinde ölçülmesine imkân veren performansa dayalı ölçümlere yer verilmiştir. Belgesel tarama yöntemi kullanılarak yapılan çalışmanın sonunda, performansa dayalı ölçümlerin kullanımının önemi geniş bir literatürle desteklenmiştir. Performansa dayalı ölçme yöntemlerinden verilen örnekler de bu makalenin kapsamı içindedir. Makalenin bitiminde reform sürecinde olan eğitim sistemimizde performansa dayalı ölçme yöntemlerinin kullanılması ile ilgili bazı önerilerde bulunulmuştur.

Anahtar Kelimeler: Performans Ölçümü; Bireysel Gelişim Dosyası; Dereceli Puanlama Anahtarı; Ölçme ve Değerlendirme.

GİRİŞ

Türk Eğitim Sistemi'nde son yıllarda devam eden reform hareketleriyle birlikte, öğrencilerin fen alanında yüksek seviyede düşünme, muhakeme yapma ve problem çözme becerilerinin ve kavramları anlamalarının ölçümünün önemi gündeme gelmiştir. Gelişen teknoloji ve toplumun ihtiyaçlarının da etkisiyle değişen bilimsel bilginin edinilmesi anlayışı müfredata entegre edildiğinde, kullanılan ölçme yöntemleri de bu anlayışı yansıtmalıdır (Zollman & Jones, 1994). Bu bağlamda, bütünlendirici yaklaşımın desteklediği öğrenme teorileriyle uyuşan ve bilginin aktif olarak yapılandırılmasını gerektiren performansa dayalı ölçümlerin de kullanılması önerilmektedir (Maeroff, 1991). Performansa dayalı ölçme yöntemlerini kullanabilmek için, öncelikle bu tür ölçümlere neden ihtiyaç duyulduğunu, bütünlendirici yaklaşım ile uyumunu ve performansa dayalı ölçümlerin yapısını anlamak gerekir. Bu nedenle performansa dayalı ölçümler ile ilgili yapılan araştırmaları derleyerek, bir organizasyon bütünlüğü içinde sunan çalışmalara ihtiyaç vardır.

AMAÇ

Bu çalışmada, performansa dayalı ölçümlerin temelindeki mantığın yansıtılması ve mevcut öğrenme teorilerine göre gelişen yeni eğilimler ile örtüştüğü noktaların ortaya konulması amaçlanmıştır. Ayrıca, performansa dayalı ölçme yöntemlerinden çeşitli örnekler verilerek, bahsedilen yöntemlerin öğretim sürecine entegre edilmesinde bu çalışmanın bir rehber olması hedeflenmiştir.

YÖNTEM

Çalışmada belgesel tarama (döküman metodu) yöntemi kullanılmıştır. “Belgesel tarama tekniğinde kaynak olarak kullanılan belgelerde bir değişme söz konusu olmadığından, veri toplama işleminde güvenilirlik en üst düzeylere çıkabilir” (Karasar, 2003, s.197).

Belgesel tarama yaparken Zwicky (1969)’nin kullandığı aşağıdaki adımlar takip edilmiştir:

1. Performansa dayalı ölçümler ile ilgili olan ve ulaşılabilen çalışmalardan çıkarımlar yapılması,
2. Elde edilen çıkarımlar arasında ortak temalar oluşturulması,
3. Oluşturulan temaların kategoriler altında toplanması,
4. Kategoriler arası ilişkilerin incelenmesi.

Çalışmada performansa dayalı ölçümlerin temelindeki mantık incelenmek istendiğinden, kullanılan literatürün güncelliğinden çok, bu konuda temel teşkil eden çalışmaların kaynak olarak kullanılması amaçlanmıştır. Belgesel tarama sonucunda “öğrencinin ölçülmesi kavramı”, “öğrencinin ölçülmesinde ve değerlendirilmesinde yeni eğilimler”, “bütünleştirici yaklaşım ve ölçme-değerlendirme”, “performansa dayalı ölçümler” ve “performansa dayalı ölçme yöntemleri” olmak üzere beş kategori belirlenmiştir. Her bir kategori içinde yer alan ancak birbirinden farklı olan temalar da kendi aralarında gruplandırılarak, alt kategoriler oluşturulmuştur. Örneğin, “bütünleştirici yaklaşım ve ölçme-değerlendirme” ile ilgili kategorinin altında “öğretim ve ölçme”, “dönüt verme” ve “öz değerlendirme ve akran değerlendirme” olmak üzere üç alt kategori yer almaktadır. Kategorilerin sunumu aralarındaki ilişkiler göz önüne alınarak yapılmıştır.

a) Öğrencinin Ölçülmesi Kavramı

“Öğrencinin ölçülmesi” kavramı farklı bağlamlar içinde farklı amaçlar için tanımlanmıştır. Öğrencinin ölçülmesi, onun gösterdiği performansın kapsamlı ve çok yönlü bir analizidir ve yargıya dayalıdır (Wiggins, 1993). Gainen ve Locatelli (1995)’e göre ölçme, öğrencinin öğrenmesini ve memnuniyetini artırmak için onun hakkında sistematik olarak bilgi toplamak, yorumlamak ve kullanmaktır.

b) Öğrencinin Ölçülmesinde ve Değerlendirilmesinde Yeni Eğilimler

“Geleneksel ölçme yöntemlerinin öğrencilerin sınav becerisine odaklanması ve ön bilgiyi ve düşünme becerilerini göz ardı etmesi nedeniyle” (Zollman & Jones, 1994, s.3), öğrencinin ölçülmesinde farklı eğilimler ortaya çıkmıştır. Objektif testler kullanarak, öğretimin sonunda yapılan ölçümler yerine; performansa dayalı ölçme yöntemleri kullanarak, öğretmenin anlık kararlar alabilmesine imkân veren ve öğretim sırasında yapılan ölçümler önem kazanmıştır (McMillan, 1997). Literatüre bakıldığında terminolojide bir tutarsızlık olduğu görülmektedir. Bazı yazarlar alternatif ölçme yöntemleri, bazı yazarlar ise performansa dayalı ölçme yöntemleri ifadesini kullanmaktadırlar. Her iki ifadenin de temel mantığı, bilginin ve becerilerin farklı

bağlamlarda uygulanmasının gösterilmesidir. Bu makalede performansa dayalı ölçme yöntemleri ifadesi kullanılmıştır.

Performansa dayalı ölçümler, öğrenciden cevabı seçmesinin istendiği ölçümlerin tersine, öğrenciden cevabı kendisinin oluşturması istenen ölçümlerdir (Stiggins, 1995). Bu tür ölçümler, soyutlanmış gerçeklerin tekrarını değil, bilginin aktif olarak yapılandırılmasını gerektirir. Performans ölçümleri (performance assessments), otantik ölçümler (authentic assessments) ve bireysel gelişim dosyası (portfolio) performansa dayalı ölçme yöntemlerinden bazılarıdır (Smith, 2003). Bu yöntemler üst düzey becerilerin kullanılmasını gerektirdiği ve öğrenciyi öğrenmeye yönlendirdiği için, bilişsel öğrenme ve motivasyon teorileri ile uyumludur; ayrıca, öğrencinin bilimsel süreç becerilerini geliştirir ve onu günlük hayat içinde yaşanan daha karmaşık problemlere hazırlar (Maeroff, 1991; Briscoe & Wells, 2002).

Ölçme ve değerlendirme alanında yaşanan değişimle beraber öğretim ve ölçmenin birbirini tamamlayan iki süreç olması, öğrenciye farklı bağlamlarda farklı görevler verilmesi, zamanında ve yeterli dönüt (feedback) verilmesi ve öğrencinin kendisini değerlendirmesi önem kazanmıştır. Öğrencinin ölçülmesinde ve değerlendirilmesinde yeni eğilimler Tablo 1’de verilmiştir.

Tablo 1. Ölçme ve Değerlendirme Alanında Yaşanan Değişim.

Odaklanılan Durumlar	Zamanla Terk Edilen Durumlar
Öğrencinin ne bildiğinin ölçülmesi	Öğrencinin ne bilmediğinin ölçülmesi
Ölçmenin öğretimin bir parçası olması ve öğretim sırasında da ölçüm yapılması	Ölçmenin yalnızca öğretim sonunda yapılması
Daha geniş bağlamlarda farklı görevler verilmesi ve bütünsel bir bakış sağlanmaya çalışılması	Belli konularda spesifik görevler verilmesi
Pek çok becerinin bir arada uygulanmasını gerektiren problemlerin verilmesi	Yalnızca birkaç becerinin kullanılmasını gerektiren soruların sorulması
Çeşitli ölçme yöntemlerinin kullanılması	Yalnızca yazılı soruların kullanılması
Sürecin de sonuçla beraber ölçülmesi	Yalnızca sonucun ölçülmesi
Zamanında ve yeterli dönüt verilmesi	Çok az dönüt verilmesi
Öğrencinin öz değerlendirme yapması ve yalnız öğretmen değil, akranları tarafından da değerlendirilmesi	Öğrencinin yalnızca öğretmen tarafından değerlendirilmesi

Tablodaki bazı maddeler 1992 yılında yayınlanan “*Alternative Assessments in Math And Science: Moving Toward A Moving Target*” adlı çalışmadan, bazı maddeler ise, McMillan (1997)’den derlenmiştir.

c) Bütünleştirici Yaklaşım ve Ölçme-Değerlendirme

Eğer öğrenme gerçeklerin, kuralların ve becerilerin edinilmesi olarak görülüyorsa; yapılan ölçüm daha çok hatırlama becerileri ve bilginin doğrulanması üzerine odaklı olur ve öğrenci “öğrendi” ya da “öğrenmedi” diye sınıflandırılır. Ancak, öğrenme tecrübelerin yapılandırılmasıyla sürekli gelişen bir süreç olarak düşünülüyorsa; ölçme öğrenmenin bir parçası olarak görülür ve dönüt verme olarak algılanır (Delandshere & Jones, 1999).

Klasik ölçme modelinde amaç belli özelliklere göre öğrencileri sınıflandırmak iken, bütünleştirici yaklaşımda amaç öğrencinin öğrenme sürecinin neresinde olduğunu belirlemektir (Wilson, 1994). Bütünleştirici yaklaşımda öğrencinin zihin yapısının doğru bir modeli çıkarılmaya çalışılır. Öğrencinin ne bildiği ya da anladığı hakkında geçerli çıkarımlar yapabilmek için, öğretmen farklı kaynaklardan bilgi edinmelidir, yani farklı ölçme yöntemleri kullanılmalıdır. Birçok becerinin ölçülmeye çalışıldığı tek bir ölçme yöntemi kullanmak, bütünleştirici yaklaşımla uyumsuzdur (Wilson, 1994). Öğretmenin öğrencinin kapasitesi (sahip olduğu bilgi ve beceriler) hakkında doğru karar

verebilmesi için farklı zamanlarda, farklı yöntemlerle edinilmiş, farklı bilgilere ihtiyacı vardır (Stiggins, 1999). Ayrıca, bütünleştirici yaklaşım, öğrencinin kapasitesini daha iyi gösterebileceği öz değerlendirmeyi ve akran değerlendirmesini desteklemektedir (Wilson, 1994).

1. Öğretim ve Ölçme: Kullanılan ölçme yöntemleri, öğrencilerin neyin önemli olduğunu anlamalarına ve kendilerini o yönde geliştirmeye çalışmalarına neden olur. Performansa dayalı ölçme yöntemlerinin kullanılması öğrencilerin öğrenmesini artıracığından (Novak, Mintzes & Wandersee, 2000), öğretim ve ölçme birbirinden ayrılan değil, birbirini tamamlayan iki süreç olur. Bütünleştirici yaklaşım öğretim ile ölçmenin birbiriyle entegrasyonunu öngörmektedir (Pilcher, 2001; Sluijsmans, Brand-Gruwel, van Merriënboore & Bastiaens, 2003; Wilson, 1994).

Öğretim ve ölçmenin birbiri ile ilişkilendirilerek, öğrencinin sürekli olarak ölçülmesi aşağıda verilen koşullar göz önüne alınarak sağlanabilir (Palmer, 1994):

1. Ölçüm sırasında kullanılan materyallerle, öğretim sırasında kullanılan materyallerin uyumlu olması,
2. Öğrencinin gelişimini ölçebilmek için çeşitli yöntemler kullanılması,
3. Öğrencilerin gözlemlenmesi ve kaydedilmesi,
4. Öğrencilerin verdikleri cevapların arkasındaki sebepleri ortaya çıkartacak yöntemler kullanılması.

Öğretmen, öğrencileri ölçmeye başlamadan önce neyin öğrenilmesinin önemli olduğunu ve buna göre ne tür bilgi ve becerilerin ölçüleceğini belirleyerek, belirlenen bilgi ve becerilere göre en uygun ölçme yöntemine karar vermelidir.

2. Dönüt (Geri Bildirim) Verme: Dönüt ya da geri bildirim (feedback), varolan düzey ile olması istenen düzey arasındaki boşluğun tamamlanması için gerekli olan bilgidir (Ramaprasad, 1983). Öğrencinin öğrenmesi ve kendisini geliştirmesi için, ölçme ve değerlendirme süreci içinde zamanında ve yeterli miktarda verilen dönütün çok önemli bir yeri vardır. Doğru cevabın verildiği dönütler öğrenci için faydalı değildir. Dönütün zamanında verilmesi için öğrencinin yalnız öğretimin sonunda değil, öğretim süreci içinde de ölçülmesi uygun olur. Burada performansa dayalı ölçümün önemi ortaya çıkmaktadır. Ancak, dönütün faydalı olması ve rehberlik edebilmesi için öğrencinin verilen dönütü anlaması ve kullanması gerekir. Bunun için de dönüt verirken aşağıda verilen hususlara dikkat edilmesi tavsiye edilmektedir (Sadler, 1989):

1. Öğrencinin hangi bilgi ve becerilere sahip olması gerektiğinin açıklanması,
2. Öğrencinin sahip olması gereken bilgi ve beceriye göre yaptığı çalışmanın karşılaştırılması,
3. Aradaki boşluğu kapatması için öğrencinin neler yapabileceğinin önerilmesi.

3. Öz Değerlendirme ve Akran Değerlendirme: Öğrencilerin kendilerini değerlendirmeleri, bilgi ve beceri olarak ne düzeyde olduklarını anlamalarına imkân vererek, kendilerini geliştirmek için neler yapmaları gerektiğini belirlemelerini sağlar. Böylece öğrencilerin pasif birer alıcı olmaları engellenir. Öz değerlendirme ve akran değerlendirme herhangi bir not verilmeden ve şekillendirici değerlendirme sürecinde yapılırsa amacına daha çok ulaşabilir (Taras, 2002). Öğretmenin öğrencileri değerlendirirken kullandığı kriterlerin aynısını, öğrenciler öz değerlendirme yaparken ve akranlarını değerlendirirken kullanabilirler.

d) Performansa Dayalı Ölçümler

Öğrenciden bilgisini bir cevap oluşturarak veya bir ürün yaparak göstermesinin istendiği ölçümler performansa dayalı ölçümler olarak ifade edilir (Century, 2002). Öğrenciden yapması istenilen kısa veya uzun dönemli etkinliklere ise performans görevleri denir. Sözel sunum yapma, deney yapma, problem çözme, proje hazırlama, model oluşturma, bir müzik aleti çalma ve bilimsel yazı veya kompozisyon (essay) yazma gibi görevler performans görevleri olabilir. Performansa dayalı ölçümlerde öğrenci daha bağımsız düşünür ve bir cevabı doğrulamaya çalışmaz (Shepard, 1991). Bu tür ölçümlerde öğrenci verilen problemi çözebilme ya da görevi yapabilmek için ön bilgisini kullanmaya, araştırma yapmaya ve sorgulamaya ihtiyaç duyar. Bu durumda ölçme, aynı zamanda öğrenmeyi de artıran bir unsurdur.

Öğrenmenin bağlamsal olduğu, diğer bir deyişle öğrencilerin performansının dinamik olduğu ve farklı durumlarda zamanla değiştiği göz önünde bulundurulduğunda, bir günde yapılan bir sınavın öğrencinin ne yapabileceğini ölçmede yetersiz kaldığı anlaşılmış ve alternatif yollar aranmıştır (Johnsen, 1996). Öğrencinin gelişimini değerlendirmek için çeşitli ölçme yöntemlerini kullanmak daha uygundur. Ayrıca öğrenciye ne öğrendiğini gösterme fırsatı da sunulmalıdır ki, bu da yine performansa dayalı ölçme yöntemleri kullanarak olabilir (Henessey, 1995). Ölçmenin amacının öğretimi geliştirmek ve öğrenimi artırmak olduğu düşünüldüğünde, yapılan ölçümün hem öğrencinin iyi olduğu hem de kendisini geliştirmesi gerektiği alanları belirlemesi gerekir. Performansa dayalı ölçümler bunu sağlar.

Performansa dayalı ölçümler öğrencinin öğrendiklerine büyüteç tutmak gibidir (Korkmaz & Kaptan, 2005) ve bu nedenle ne öğretilmesi gerektiği hakkında da bilgi verir (Kamen, 1991). Ayrıca, yapılan araştırmalar öğrencilerin performansa dayalı ölçümleri meydan okuyucu ve heyecan verici bulduklarını göstermiştir (Century, 2002)

Performans görevleri öğrencilerin günlük yaşamlarında karşılaştıkları durumlardan oluşabilir. Bu tür görevlere otantik görevler (authentic tasks) denir (Wiggins, 1989). İlköğretim seviyesinde verilebilecek bir matematik otantik görevi şöyle olabilir (Zollman & Jones, 1994): “45 YTL paranız var ve 10 arkadaşınızın katılacağı bir doğum günü kutlaması yapmak istiyorsunuz. Nasıl yaparsınız?”. Bu problemi çözmek için öğrenci ilk başta neleri yapması gerektiğini (yiyecek ve içecek almak, kutlamanın yapılacağı yeri ayarlamak, tabak, çatal ve bardak almak, davetiye göndermek ya da telefon etmek gibi) planlamalıdır ve buna göre bazı fiyat tahminlerinde bulunup hesaplamalar yapmalıdır.

Öğretmenlerin performansa dayalı bir ölçüm hazırlarken aşağıda verilen aşamaları takip etmesi tavsiye edilmektedir (Reichel, 1994; Moskal, 2003):

1. Ölçülmek istenilen öğrenim hedeflerinin belirlenmesi,
2. Bu hedefler içinde kavramsal anlama, kritik düşünme ve problem çözme gibi üst düzey becerilere odaklanması,
3. Hedeflenen becerileri ölçmeye yönelik, öğrencinin kendisinin yapabilmesine veya üretebilmesine olanak tanıyan bir performans görevi hazırlanması,
4. Performans görevi altında öğrenciden istenecek kriterlerin belirlenmesi,
5. Verilen performans görevinin içeriğinin bazı öğrenciler için dezavantaj oluşturmamasına ve meyilli olmamasına dikkat edilmesi,
6. Beklentilerin öğrencilere açıkça anlatılması,
7. Öğrencilerin kendi yaptıklarını değerlendirmelerine imkân verilmesi.

1. Performansa Dayalı Ölçümlerin Geçerliliği ve Güvenirliği: Bir ölçümün geçerliliği yalnızca kullanılan ölçme yönteminin veya elde edilen puanların geçerliliği değil, ölçümden elde edilen çıkarımların doğruluğudur (Maclellan, 2004). Düşük geçerliliği olan bir ölçüm, öğrencinin sahip olduğu kapasite hakkında doğru olmayan çıkarımlar

yapılmasına neden olur. Geçerlik şartlarından birisi olan yapı geçerliği amaca uygun ölçme yöntemini kullanmakla ilgilidir. Örneğin, bir öğrencinin motivasyonunu ölçmek için en uygun yöntem ona motivasyon hakkında bir yazı yazdırmak değil, onu ders sırasında gözlemlemektir. Performansa dayalı ölçümlerin, ölçülecek öğrenim hedefine uygun olması, kriterlerin çok dar kapsamlı veya çok geniş kapsamlı olmaması geçerlik için aranan bazı özelliklerdir.

Ölçümün geçerliği, psikometrik analizler yerine kullanılan ölçme yönteminin, ölçülmek istenen hedefle uyumluluğuna bakılarak irdelenmelidir (Pilcher, 2001). Dolayısıyla, geçerlik şartının sağlanması, ölçümün öğretimin iyileştirilmesi için yapılmasına bağlıdır (Cole, 1988). Bütünleştirici yaklaşıma göre geçerli ölçümler öğretmen ve öğrenci için faydalı olan ölçümlerdir, bu durumda ölçümün geçerliği öğrencilerin zayıf ve güçlü yönlerini göz önünde bulundurup, farklı ölçme yöntemleri kullanarak sağlanır (Graue, 1994).

Performansa dayalı bir ölçümün geçerliğini irdelerken aşağıdaki sorular dikkate alınabilir (Wiggins, 1996):

1. Ölçülmek istenen öğrenim hedefleri (bilgi ve beceriler) açıkça belirlenmiş mi?
2. Ölçülmek istenilen bilgi ve beceriler öğretim ile uyumlu mu?
3. Performans görevi ölçülmek istenen bilgi ve beceriler için uygun bir görev mi?
4. Performans görevi içinde belirlenen alt kriterler, öğrenciye verilen yönergeler ve sorulan sorular ölçülmek istenen bilgi ve becerileri ortaya çıkarmaya yönelik mi?
5. Kriterler, yönergeler ve sorular açık ve anlaşılır olarak ifade edilmiş mi?
6. Performansa dayalı ölçüm önemsiz bilgi ve becerileri değil, yalnızca öğrencinin öğrenmesi için gerekli bilgi ve becerileri kapsıyor mu?
7. Puanlama kriterleri ölçülmek istenilen bilgi ve becerileri ölçmeye yönelik mi?
8. Puanlama kriterleri açık ve anlaşılır olarak ifade edilmiş mi?
9. Puan ağırlıkları ve dağılım, ölçülecek bilgi ve becerilerin seviyesi ile uyumlu mu?

Ölçümün geçerliği irdelenirken “ölçümün kapsamının” göz önünde bulundurulması gerekir. Performansa dayalı ölçümün kapsamı, öğretimin içeriğiyle uyumlu olmalı ve problem çözme, muhakeme yapma, kritik düşünme, bilginin ve becerilerin farklı koşullarda uygulanması gibi performansa dayalı yöntemlerle ölçülebilecek hedefler içermelidir. Ölçümün kapsam olarak uygun olup olmadığını belirlerken yukarıda verilen sorulardan 2., 4. ve 6. sorular göz önünde bulundurulabilir (Carter, 1994).

Öğrenciye aynı bilgi ve becerileri ölçmeye yönelik farklı performans görevleri verilmesi, öğrencinin kapasitesi hakkında verilen kararın genellenebilmesini sağlar. Yukarıda verilen sorulardan 5., 8. ve 9. sorular ölçümün ve puanlandırmanın güvenilirliği ile ilgilidir. İyi tanımlanmış ve anlaşılır maddelerden oluşan bir dereceli puanlama anahtarının (rubric) kullanılması, puanlamadan kaynaklanabilecek güvenilirlik problemlerini ortadan kaldırır (Smith, 2003). Bu durumda farklı kişilerin değerlendirmesiyle elde edilen puanlar zaten birbiriyle tutarlı olacak ve kişiler arası güvenilirlik yüksek olacaktır (Shavelson & Ruiz-Primo, 2000).

2. Performansa Dayalı Ölçümlerin Değerlendirilmesinde Dereceli Puanlama Anahtarı Kullanımı: Dereceli puanlama anahtarı (rubric), kabul edilebilir ve kabul edilemez performans aralığını kriterlerle belirleyen bir ölçektir (Pate, Homestead & McGinnis, 1993). Öğrencinin gösterebileceği performansın her bir seviyesinin tanımlandığı ve bu seviyelere değerlerin atandığı kriterlerden oluşur (Herman, Aschbacher & Witner, 1992). Verilen performans görevi süresince öğrencinin gösterdiği performans ve/veya görev sonunda öğrencinin hazırladığı ürün, iyi hazırlanmış bir dereceli puanlama anahtarı kullanarak doğru ve adil bir şekilde değerlendirilebilir.

Dereceli puanlama anahtarının diğer bir amacı öğrencilerin hangi kriterlere göre ölçüleceğini bilmeleri ve görevi tamamlamak için kendilerine hedefler koymalarıdır (Jensen, 1995). Bu kriterler öğrencilerin katılımı ile de belirlenebilir. Dereceli puanlama anahtarının değerlendirilmesi sonunda öğrencilerin hangi beceriler için kendilerini ne kadar geliştirmeleri gerektiği belirlenerek, öğrencilere daha yapıcı dönütler verilebilir. Dereceli puanlama anahtarı hazırlarken belirlenen kriterlerin ölçülemeyecek karmaşıklıkta olmamasına ve kriterler için hazırlanan ölçeğin dört aşamadan fazla olmamasına dikkat etmek gerekir. Kullanılan ölçek nicel (nümerik) ve/veya nitel olabilir.

Dereceli puanlama anahtarı ile öğrencinin yalnızca bilişsel becerileri değil, derse katılımı, işbirlikçi yaklaşımı, akranlarıyla ilişkisi ve üstbiliş becerileri de ölçülebilir. Bu açıdan dereceli puanlama anahtarı kullanımı, öğrenmenin aktif bir süreç olduğunu belirten günümüz teorileriyle de uyumludur.

Bütünsel ve analitik olmak üzere iki tür dereceli puanlama anahtarı vardır. Bütünsel dereceli puanlama anahtarı, sürecin veya ürünün parçalarının ayrı ayrı değerlendirilmeden, bir bütün olarak değerlendirilmesini sağlar (Nitko, 2001). Analitik dereceli puanlama anahtarı kullanarak ise süreç veya ürün parçalarına ayrılarak, her bir parça kendisi içinde ayrı ayrı değerlendirilebilir, dolayısıyla öğrenciye daha spesifik konularda dönüt verilebilir (Nitko, 2001). Analitik dereceli puanlama anahtarında her bir kriterin kendi içindeki aşamaları için belirlenmiş bir ölçek olduğundan, yapılan ölçümün objektifliği için önemli bir gereklilik sağlanmış olur.

Örnek 1’de bir problem çözme etkinliğinde kullanılabilecek genel içerikli bütünsel bir dereceli puanlama anahtarı verilmiştir. Probleme göre dereceli puanlama anahtarının içeriğindeki kriterler daha spesifik olarak yazılabilir.

Örnek 1: *Bir Problem Çözme Etkinliğinde Kullanılabilecek Bütünsel Dereceli Puanlama Anahtarı*

Kategoriler	3: İleri Seviye	2: Orta Seviye	1: Düşük Seviye	0: Çok zayıf
Sorgulama	- Görev anlaşılmış ve analiz edilmiş. - İşe yarayan bir strateji geliştirilmiş. - Geliştirilen strateji doğru uygulanmış. - Strateji nitelikli bir çözümü destekliyor. - Kritik düşünme becerileri uygulanmış.	- Görev anlaşılmış. - İşe yarayan bir strateji geliştirilmiş. - Bir plan yapılmış ancak anlaşılır değil. - Strateji uygun bir çözümü destekliyor. - Kritik düşünme becerileri uygulanmış.	- Görev kısmen anlaşılmış. - Uygun bir strateji geliştirilmiş. - Bir plan yapılmış olabilir ancak anlaşılır değil. - Strateji uygun bir çözümü kısmen destekliyor. - Kritik düşünme becerileri kısmen uygulanmış.	- Görev kısmen anlaşılmış. - Uygun olmayan bir strateji geliştirilmiş. - Bir plan yapıldığına dair herhangi bir delil yok. - Çözüm ulaşılamamış. - Kritik düşünme becerileri uygulanmamış.
Sunum	- Bilgi veya veriler uygun ve düzenli bir şekilde sunulmuş. - Sunulan bilgi veya veriler kesin, doğru ve tam. - Yapılan çıkarımlar ve açıklamalar mantıklı ve açıkça ifade edilmiş.	- Bilgi veya veriler uygun ve düzenli bir şekilde sunulmuş. - Sunulan bilgi veya veriler kısmen kesin, doğru ve tam. - Yapılan çıkarımlar ve açıklamalar mantıklı ve oldukça açıkça ifade edilmiş.	- Bilgi veya veriler açıkça sunulmamış. - Sunulan bilgi veya veriler çok az kesin, doğru ve tam. - Yapılan çıkarımlar ve açıklamalar çok az açıklıkta ifade edilmiş.	- Bilgi veya veriler sunulmak üzere düzenlenmemiş. - Sunulan bilgi veya veriler kesin, doğru ve tam değil. - Yapılan çıkarımlar ve açıklamalar açıkça ifade edilmemiş.

Dereceli puanlama anahtarı hazırlamak zaman alır ve kriterlerin belirlenmesi kolay olmayabilir ancak, birçok ölçümün (performans ölçümleri ve bireysel gelişim dosyası gibi) değerlendirmesinde dereceli puanlama anahtarı kullanılabilir. Dereceli puanlama anahtarında yer alan kriterlerin öğrenci tarafından ne kadar sağlandığını belirlemek için gözlemler yapılabilir, sözel sorular sorulabilir, yazılı dökümanlar ya da ortaya çıkan ürün incelenebilir ve dereceli puanlama anahtarı bunlara göre doldurulabilir.

Dereceli puanlama anahtarı hazırlarken aşağıdaki aşamaların takip edilmesi önerilmektedir (Mertler, 2001):

1. Verilen performans görevi ile ölçülecek öğrenim hedeflerinin belirlenmesi,
2. Öğrencilerin performansı sırasında gözlemlenebilir olan ve öğrencilerin sahip olması istenen (ve istenmeyen) bilgi ve becerilerin belirlenmesi.
3. İkinci adımda belirlenen her bilgi veya beceri için ortalamanın üstü, ortalama ve ortalamanın altı olabilecek durumların belirlenmesi,
4.
 - a. Bütünsel bir dereceli puanlama anahtarı hazırlanıyorsa, her bilgi veya beceri için gözlemlenebilir en yüksek performansların, orta performansların ve en düşük performansların bir araya getirilerek, ayrı ayrı grupların oluşturulması.
 - b. Analitik bir dereceli puanlama anahtarı hazırlanıyorsa, her bilgi veya beceri için gözlemlenebilir en yüksek, orta ve en düşük performansların yazılarak bir grup oluşturulması.
5. Her bir seviye için öğrencilerin gözlemlenmesi ve dereceli puanlama anahtarının doldurulması,
6. Dereceli puanlama anahtarının gözden geçirilerek, gerekiyorsa bir kez daha kullanmadan önce üzerinde değişiklikler yapılması.

Dereceli puanlama anahtarı ile değerlendirme yaparken, verilen puanların tutarlılığını belirleyebilmek için, “puanlama güvenilirliği” incelenir. Moskal (2003)’a göre bunun için:

1. Dereceli puanlama anahtarı kriterlerini birbirlerinden bağımsız olarak puanlayan iki kişinin verdikleri puanlar birbiri ile tutarlı olmalıdır.
2. Öğrencinin performansı değişmedikçe, puanlandıran kişinin verdiği puanlar zaman içinde tutarlı olmalıdır.
3. Farklı performanslar (iyi, orta, zayıf gibi) gösteren öğrencilerin dereceli puanlama anahtarlarından seçilen örnekler, puanlandırma yapan kişinin öğrenci için verdiği puanların tutarlılığını test etmesi için kullanılabilir (örneğin zayıf performans gösterenlerin kendi aralarında karşılaştırılması).

e) Performansa Dayalı Ölçme Yöntemleri

Makalenin bundan sonraki bölümünde performansa dayalı ölçme yöntemlerinden performans ölçümüne, bireysel gelişim dosyasına ve yine bir çeşit performans ölçümü olan projelere örneklerle yer verilmiştir.

1. Performans Ölçümleri: Performansa dayalı ölçme yöntemlerinden birisi performans ölçümleridir (performance assessments). Örnek 2’de performans ölçümü örneği olarak laboratuvar ortamında yapılan bir deney verilmiştir (farklı performans ölçümü örnekleri için bkz. Ogan-Bekiroğlu, 2004).

Örnek 2: Laboratuarda Yapılan Bir Fizik Deneyi İçin Öğrencinin Performansının Performans Ölçümü İle Belirlenmesi.

Ölçülen öğrenim/performans hedefi	Bir cismin hareketini farklı değişkenlere göre açıklayabilme.												
Performansın tanımı	Bir balonun hareketinin incelenmesi ve hızının ve ivmesinin bulunması.												
Uygulama	Öğrenciler grupça bir balonun hızını ve ivmesini bulabilmek için, verilen materyalleri kullanarak bir deney dizayn edecekler ve balonun hareketini izleyeceklerdir.												
Kullanılacak Araç-Gereç	Her bir gruba aşağıdaki araç-gereçler verilecektir. • 8 m uzunluğunda naylon ip • 22 cm çapında bir balon • Cetvel • Kronometre • Kağıt ve kalem • Bant • 1 adet pipet • Grafik kağıdı												
Öğrenciye Yönelik Direktifler	• Balonun hızını ve ivmesini bulabileceğiniz bir deney dizayn ediniz. • Yaptığınız her denemede aldığınız ölçümleri kaydedebileceğiniz bir veri tablosu oluşturunuz. • Sonuçları gösterebileceğiniz bir grafik çiziniz. • Balonu hareket ettiren kuvvet ya da kuvvetleri belirleyiniz ve tanımlayınız. • Balona karşı olan kuvvet ya da kuvvetleri belirleyiniz ve tanımlayınız. • Balonun şu koşullarda nasıl hareket edeceğine dair hipotez oluşturunuz: a) vakum ortamında, b) uzayda, c) Dünya'nın iki katı kütleyle sahip bir gezegende d) Dünya'daki atmosferden iki kat daha yoğun olan atmosfere sahip bir gezegende. • Balonun hareketini Newton'un Birinci Kanunu'nu göz önünde bulundurarak açıklayınız. • Sürtünmenin bu deneyi nasıl etkilediğini açıklayınız. Not: Öğretmen eğer tercih ediyorsa deneyin nasıl yapılacağını öğrencilerine açıklayabilir ve doldurulacak tablo hakkında bilgi verebilir.												
Deneyin Yapılışı	İpi pipetten geçiriniz ve pipetin ucunun ipin ucuna yakın olmasına dikkat ediniz. İpin diğer ucunu sabit bir yere bağlayınız. Balonu şişiriniz ve açık ucunu sıkıca tutunuz. Balonu pipete tutturunuz. Balonun ucunu bıraktığınızda, balon ipe geçirilmiş pipetle birlikte hareket edecektir. Balonun hareket ettiği mesafeyi ve bu sürede geçen süreyi ölçünüz. En az 3 deneme yapınız ve elde ettiğiniz verileri tabloya kaydediniz. Daha sonra hız ve ivme değerlerini hesaplayınız. <i>Veri Tablosu</i> <table><tr><th>Deneme</th><th>Mesafe (metre)</th><th>Zaman (Saniye)</th></tr><tr><td>1</td><td></td><td></td></tr><tr><td>2</td><td></td><td></td></tr><tr><td>3</td><td></td><td></td></tr></table>	Deneme	Mesafe (metre)	Zaman (Saniye)	1			2			3		
Deneme	Mesafe (metre)	Zaman (Saniye)											
1													
2													
3													

Bu performans ölçümünün değerlendirilebilmesi için grubun çalışması gözlemlenerek ve hazırladıkları deney raporu incelenerek doldurulabilecek bir oranlama ölçeği hazırlanmıştır. Verilen ölçekteki rakamlar (1, 2, 3, 4) belli bir katsayı ile çarpılarak öğrencinin performansı notlandırılabilir. Değerlendirme için oranlama ölçeği yerine bir kontrol listesi veya dereceli puanlama anahtarı da hazırlanabilir.

4: Mükemmel 3: İyi 2: Orta 1: Zayıf G: Gözlenemedi					
Balonun hızının ve ivmesinin bulunabileceği bir deney dizayn edilmiş.	1	2	3	4	G
Alınan ölçümler/toplanan veriler doğru ve tam.	1	2	3	4	G
Her denemede alınan ölçümlerin kaydedildiği bir veri tablosu oluşturulmuş.	1	2	3	4	G
Sonuçların gösterilebileceği tam ve doğru bir grafik çizilmiş.	1	2	3	4	G
Balonu hareket ettiren kuvvet ya da kuvvetler belirlenerek, tanımlanmış.	1	2	3	4	G
Balona karşı olan kuvvet ya da kuvvetler belirlenerek, tanımlanmış.	1	2	3	4	G
Balonun vakum ortamında nasıl hareket edeceğine dair doğru bir hipotez oluşturulmuş.	1	2	3	4	G
Balonun uzayda nasıl hareket edeceğine dair doğru bir hipotez oluşturulmuş.	1	2	3	4	G
Balonun Dünya'nın iki katı kütleye sahip bir gezegende nasıl hareket edeceğine dair doğru bir hipotez oluşturulmuş.	1	2	3	4	G
Balonun Dünya'daki atmosferden iki kat daha yoğun olan atmosfere sahip bir gezegende nasıl hareket edeceğine dair doğru bir hipotez oluşturulmuş.	1	2	3	4	G
Balonun hareketi Newton'un Birinci Kanunu göz önünde bulundurularak açıklanmış.	1	2	3	4	G
Sürtünmenin bu deneyi nasıl etkilediği açıklanmış.	1	2	3	4	G

2. Bireysel Gelişim Dosyası (Öğrenci Ürün Dosyası) : Performansa dayalı ölçme yöntemlerinden birisi olan bireysel gelişim dosyası (portfolio) öğrencinin becerilerini, başarılarını ve gelişimini yansıtan işlerin toplamıdır (Century, 2002). Bireysel gelişim dosyası, öğrencinin belli bir süreç içinde yaptığı çalışmaların bir arada toplanmasıdır ve öğrencinin gelişimini, neler başardığını veya özel ilgilerini gösterir. Yazılı bir sınav belli bir anda ne öğrenildiğini gösteren bir fotoğraf ise, bireysel gelişim dosyası zaman içindeki değişimin ve gelişimin gösterildiği bir fotoğraf albümüne benzetilebilir (Wiggins, 1993). Bireysel gelişim dosyasında yer alan çalışmalar/ürünler, belirlenen öğrenim hedefleriyle ilişkili olmalı ve öğrencinin kendisini yansıtmaya ve öz değerlendirme yapabilmesine imkân vermelidir (Pierce & O'Malley, 1992). Bireysel gelişim dosyası sürecin değerlendirildiği şekillendirici değerlendirme (formative evaluation) için olabileceği gibi sonucun değerlendirildiği tamamlayıcı değerlendirme (summative evaluation) için de kullanılabilir (Ogan-Bekiroğlu, 2004).

Öğrencinin bireysel gelişim dosyasında yer alacak çalışmalar öğretmen tarafından belirlenebileceği gibi, öğrenci tarafından da seçilebilir. Eğer çalışmalar öğrenci tarafından belirlenecekse, öğrenciye kendisinden ulaşması beklenen öğrenim hedefleri verilerek, ondan bu hedeflere ulaştığını dosyasına koyduğu çalışmalarıyla göstermesi istenir. Bu durumda öğrenci çalışmalarını gözden geçirerek ve ilerlemesi için neler yapması gerektiğini belirleyerek, kendi öğreniminden sorumlu aktif bir öğrenen olur (Farr & Tone, 1994). Öğrencinin seçtiği her bir çalışma için neden o çalışmayı seçtiğini anlatan yansıtıcı bir açıklamayı da çalışmaya eklemesi istenir. Bireysel gelişim dosyası öğrencinin kendisini değerlendirmesine olanak sağladığından, üstbilis (metacognition) becerisinin gelişmesine katkıda bulunur (Hamilton, 1994). Bireysel gelişim dosyası, bir öğrencinin bireysel performansı hakkında bilgi sahibi olmak için kullanıldığında oldukça etkili bir ölçme yöntemidir (Zollman & Jones, 1994). Valdez (2001)'e göre bireysel gelişim dosyası öğrencilerin soyut kavramları somutlaştırmasına yardımcı olmaktadır. Öğrenciler bireysel gelişim dosyası için verilen görevle uğraşırken kuralları, kavramları ve kanunları daha iyi öğrendiklerini ifade etmekte, herşeyi hatırlamak zorunda olmadıklarını fark edip daha cesaretli olmaktadır (Slater, Ryan & Samson 1997). Bireysel gelişim dosyası uygulaması, geleneksel ölçme araçlarına göre öğrenciyi daha gerçekçi ve ayrıntılı olarak izleme ve hakkında daha isabetli kararlar alma fırsatı sunar

(Baki & Birgin, 2004). Bireysel gelişim dosyası öğrencinin neler yapabileceği hakkında somut örnekler içerdiğinden, velinin de öğrenciyi değerlendirme fırsatı oluşur.

Bir öğrenci bireysel gelişim dosyasında öğrencinin kendisini yansıtmaya olanak sağlayacak laboratuvar raporları, bir konu hakkında yazılan kompozisyon veya bilimsel yazılar, projeler, çizimler, tablolar, grafikler, alan gezileri raporları, fotoğraflar, araştırma raporları, kavram haritaları, bir yazının kritik edildiği raporlar, çalışma kağıtları, gözlem raporları, sertifikalar ve ödüller gibi pek çok ürün bulunabilir. Ayrıca, yalnızca yazılı dökümanlar değil, ses ve görüntü kasetleri, disketler veya modeller gibi materyaller de öğrencinin bireysel gelişim dosyasında yer alabilir. Bireysel gelişim dosyası bir bütün olarak değerlendirilebileceği gibi, dosya içindeki her bir ürün ayrı ayrı da değerlendirilebilir.

Bireysel gelişim dosyasının kullanımına başlamadan önce hangi tür çalışmaların öğrencinin dosyasında yer alacağı; bu çalışmaların sıradan çalışmalar mı, yoksa en iyi çalışmalar mı olacağı; çalışmaların konuya göre spesifik mi, yoksa müfredat içinde farklı konuların birleşmesinden oluşan bir ürün mü olacağı ve bu çalışmaları kimin seçeceği konularında karara varılmalıdır (Adams at all, 1996).

Bireysel gelişim dosyası kullanımının avantajları şu şekilde sıralanabilir (McMillan, 1997; Zollman & Jones, 1994):

1. Öğrenciye öz değerlendirme yapma fırsatı vermesi,
2. Öğretmen ve öğrencinin beraber değerlendirme yapmasına olanak sağlaması,
3. Karşılaştırma yapmaya değil, gelişime önem verilmesi,
4. Sistemli ve devam eden bir ölçüm sağlaması, bu nedenle öğrencinin farklı zamanlarda, farklı bağlamlarda gösterdiği performansın belirlenebilmesi,
5. Öğrencinin iyi olduğu becerilerine ve ne yapabileceğine odaklanılması, bu nedenle öğrencinin kendisine güvenmesini sağlaması,
6. Her öğrencinin bireysel olarak değerlendirilmesini sağlaması,
7. Öğrencinin kendisini ifade etmesine ve kendisine özgü başarılarını göstermesine imkân vermesi,
8. Öğrencinin kendi öğreniminden sorumlu olmasına olanak sağlaması ve onun üstbilis becerilerini artırması,
9. Öğrencinin derste gördüğü kavramların günlük hayatla ilişkisini kurmasına yardımcı olması,
10. Amaca göre farklı şekilde uygulanabilir olması ve öğretimle ilişkilendirilebilmesi,
11. Velilere, öğrencinin durumunun somut örneklerle açıklanabilmesini sağlaması.

Bireysel gelişim dosyası kullanımının bazı dezavantajları ise şunlardır (McMillan, 1997):

1. Değerlendirme sırasında karşılaşılan zorluklar güvenilirliğin düşük olmasına sebep olabilir.
2. Değerlendirme kriterlerinin belirlenmesi, puanlandırılması ve öğrencilerle ayrı ayrı buluşulması zaman alabilir,
3. Öğrenciler dosyalarına koydukları ürünleri/çalışmaları iyi seçemeyebilirler ve gerçek durumlarını yansıtamayabilirler,
4. Öğrenci dosyasında yer alan çalışmalar/ürünler öğrencinin bilgi ve becerisi hakkında genelleme yapılmasına olanak vermeyebilir, ancak aynı hedefi ölçen farklı çalışmaların dosyada yer alması bu durumu engelleyebilir.

Bir öğrencinin bireysel gelişim dosyasında şu dökümanlar yer alır:

- Ölçülecek hedefler,
- Hangi ürünlerin/çalışmaların hangi hedefler için hazırlandığını belirten bir açıklama,

- Değerlendirme kriterleri (tüm dosya için bütünsel olabilir veya her bir ürün için ayrı ayrı olabilir),
- Öğrencinin, dosyada yer alan ürünün/çalışmanın onun öğrenim hedeflerine ulaştığını nasıl gösterdiği ile ilgili olarak yazdığı yansıtıcı bir yazı (her bir ürün için),
- Öğrencinin öz değerlendirmesi,
- Öğretmenin değerlendirmesi .

Öğrencinin bireysel gelişim dosyasında yer alacak çalışmaların önceden mi belirleneceği, yoksa dosya oluşturulurken öğrenci tarafından mı seçileceği dosyanın kullanım amaçlarına göre belirlenir. Farklı kişiler tarafından farklı türde bireysel gelişim dosyaları tanımlanmıştır, fakat bu çeşitleri birbirinden kesin sınırlarla ayırmak zordur (Zollman & Jones, 1994). Örneğin Enger ve Yager (1998), pek çok çalışmanın konulduğu geniş kapsamlı “koleksiyon” gelişim dosyasından ve öğrencinin yaptığı en iyi çalışmaları gösteren “vitrin” gelişim dosyasından bahsetmektedir. Zollman ve Jones (1994) ise “en iyiler” (öğrencinin en iyi çalışmalarının dosyada yer alması) ve “betimsel” (verilen öğrenim hedeflerine ulaşıp ulaşılmadığının değerlendirilmesi için, dosyada yer alacak ürünlerin öğretmen tarafından veya öğretmen ve öğrenciler tarafından önceden belirlenmesi) gibi çeşitlerden bahsetmiştir. Slater (1996), ise öğrenci bireysel gelişim dosyasını 3 farklı kategori altında toplamıştır: “vitrin” (showcase) gelişim dosyası (dosyada öğrencilerin i) en iyi, ii) en ilgilerini çeken, iii) kendilerini en iyi ifade ettikleri, iv) kendilerini en çok hayal kırıklığına uğratan ve/veya en favori çalışmalarının yer alması), “kontrol listeli” (checklist) gelişim dosyası (belirlenen öğrenim hedeflerine göre dosyada yer alacak ürünlerin öğrencilerin önceden yaptığı çalışmalar arasından önceden belirlenmesi), “açık format” (open format) gelişim dosyası (dosyada öğrencinin öğrenim hedeflerine ulaştığını gösteren kendisinin seçtiği her türlü çalışmanın yer alması).

Örnek 3’de bir fizik dersinde dönem boyunca devam edebilecek kapsamlı bir öğrenci gelişim dosyası verilmiştir (farklı bireysel gelişim dosyası örnekleri için bkz. Ogan-Bekiroğlu, 2004).

Örnek 3: Fizik Dersinde Dönem Boyunca Devam Edebilecek Bir Öğrenci Gelişim Dosyası (Nickelson, 2004).

Öğrencinin bireysel gelişim dosyası ile ölçülecek öğrenim/performans hedefleri:
• Grafikleri yorumlayabilme ve grafik içinde tanımlanan değişkenler arasındaki matematiksel ilişkiyi çıkarabilme. • Fizik kavramları ile günlük yaşam arasındaki ilişkiyi anlayabilme. • Fiziğin ve teknolojinin toplum üzerindeki etkisini anlayabilme. • Bilimsel süreç becerilerini geliştirebilme. • Yüksek seviyede düşünebilme ve problem çözme becerilerini geliştirebilme. • Yaptıkları bilimsel araştırmaların raporlarını sözel ve yazılı olarak sunabilme.
Dosyalarına koydukları her çalışma/ürün için öğrencilerden aşağıdaki soruların cevaplarını içeren bir açıklama yazısı yazmaları istenir:
• Hedef için seçilen ürün nedir? • Neden bu ürün seçildi? • Seçilen ürünü hedefe ulaşıldığını nasıl gösteriyor?
Bu açıklama yazısı değerlendirilirken şu kriterler göz önünde bulundurulur:
• Cümlelerin açık ve kolay anlaşılır olması, • Yanlış kavramlar veya hatalar içermemesi, • Yapılan açıklama ile seçilen ürün arasında bağlantı kurulması, • Seçilen ürün ile, öğrencinin bilgi ve becerisi arasında bağlantı kurulması, • Öğrencinin bilgi ve becerisinde zaman içinde nasıl bir değişim olduğunu göstermesi, • Dosyada yer almayan bir ürünü anlatmaması.

Ölçülecek her hedef için ayrı ayrı belirlenir. Örneğin 1. hedef için aşağıdaki dereceli puanlama anahtarı oluşturulmuştur. Dereceli puanlama anahtarında yer alan her bir seviye bir katsayı ile çarpılarak öğrenciye puan verilebilir.

6. Seviye	Grafikleri fevkalade iyi yorumluyor. Dosyada yer alan çalışma/ürün laboratuvarında elde edilen örnekleri içeriyor. Aşağıdaki kriterlerden 3 tanesini sağlıyor: - Bağımlı ve bağımsız değişkenleri doğru belirleyebiliyor, - İki değişken arasındaki ilişkiyi açıklayabiliyor, - İlişkiyi doğru matematiksel formül ile ifade edebiliyor, - Eğimin altında kalan alanın neyi ifade ettiğini anlayabiliyor.
5. Seviye	Grafikleri fevkalade iyi yorumluyor. Dosyada yer alan ürün laboratuvarında elde edilen örnekleri içeriyor. Altıncı kategoride verilen kriterlerden 2 tanesini sağlıyor.
4. Seviye	Grafikleri iyi yorumluyor. Altıncı kategoride verilen kriterlerden 3 tanesini sağlıyor.
3. Seviye	Grafikleri iyi yorumluyor. Altıncı kategoride verilen kriterlerden 2 tanesini sağlıyor.
2. Seviye	Grafikleri iyi yorumluyor. Altıncı kategoride verilen kriterlerden 1 tanesini sağlıyor.
1. Seviye	Grafığın herhangi iki değişken arasındaki ilişkiyi verdiğini ifade edebiliyor ancak, ilişkiyi açıklayamıyor.
0. Seviye	Hedefe ulaşıldığını gösteren herhangi bir ürün mevcut değil.

Öğrencilerin dosyalarında yer almak üzere seçtikleri farklı çalışmaların değerlendirilmesinden kaynaklanabilecek güvenirlik problemini ortadan kaldırmak ve öğrencinin bilgisi ve becerileri hakkında elde edilen fikrin genellenebilmesini sağlamak amacıyla, tüm öğrencilerden belirlenen bir öğrenim hedefini ölçmek üzere açık uçlu görevlerden oluşan bir bireysel gelişim dosyası içeriği istenebilir (Vitale & Romance, 2000). Örnek 4’de yer alan proje bu tür bir dosya içindir.

Örnek 4: Isı Yalıtımı Projesi

Bu projede öğrenciler ısı iletkenliği kavramını ve çeşitli materyalleri kullanarak ev izolasyonunun nasıl yapıldığını anlayacaklardır. Bu projede kavramların anlamlı öğrenilmesi ölçülmektedir.
Projede yer alan aşamalar şunlardır: 1. Ev izolasyonu hakkında bir yazının okunması, 2. Bazı materyallerin (fiber cam, pamuk, toprak, hava, yün v.b.) ısı iletkenlik özelliklerinin incelenmesi, 3. Çeşitli sıcaklık dereceleri altında bir kutunun ısıtılması ve soğutulması durumları için ikinci şıkta incelenen farklı izolasyon materyallerinin kullanılmasının araştırılması, 4. İkinci ve üçüncü şıklar için yapılan çalışmaların, iletkenlik ve izolasyon değeri kavramları göz önünde bulundurularak özetlenmesi, 5. Sonuçların ısı iletkenliği ve radyasyon kavramlarına göre açıklanması.

Öğrencilerden sonuçların gösteriminde grafik kullanmaları istenir. Projede yer alan temel kavramlar şunlardır: ısı iletkenliği, iletken, yalıtkan, sıcaklık değişimi, radyasyon. Öğretmenin yapacağı değerlendirme, öğrencilerin gözlemlerin açıklanmasında temel kavramları nasıl kullandıklarını belirlemek üzerine olmalıdır. Öğrencilerin kavramlar arasındaki ilişkileri nasıl kurduklarını ölçmek için, yukarıda verilen aşamaların yanında onlardan dosyalarına koymak üzere ayrıca bir kavram haritası yapmaları istenebilir.

3. Projeler: Projeler, öğrencinin bir zaman sürecindeki performansını ve sonunda nasıl bir sonuç elde ettiğini ölçmek için çok uygundur. Proje grupça yapılıyorsa, her bir grup üyesinin katkısını değerlendirmek zor olabilir. Her bir üyenin gruba katkısını ayrı ayrı belirleyebilmek için, öğrencilerin görevlerinin ayrı ayrı belirlenmesi ve proje sürecince yapılan ölçümlerle grup üyelerinin birbirleriyle nasıl bir etkileşim içinde olduklarının gözlemlenmesi uygun olur. Bir başka seçenek ise öğrencilere öz-değerlendirme yaptırmaktır.

Öğrenciler yaptıkları projeleri bilim şenliklerinde sunabilirler. Örnek 5’de bilim şenliğinde sunulan bir proje için bir ilköğretim okulunun (Ogden Elementary School, Chicago) kullandığı analitik dereceli puanlama anahtarı verilmiştir.

Örnek 5: *İlköğretim Seviyesinde Düzenlenen Bir Bilim Şenliğinde Sunulan Projenin Değerlendirilmesinde Kullanılabilecek Analitik Dereceli Puanlama Anahtarı.*

Kategoriler	5 Puan	4 Puan	3 Puan	2 Puan	1 Puan
Kaynak Kullanma	- Farklı çeşitte çok sayıda (>6) kaynak kullanılmış. - Verilen bütün bilgi konu ile ilgili. - Yapılan araştırma mantıklı bir organizasyon içinde yazılmış.	- Bazıları farklı çeşitte olan oldukça fazla sayıda kaynak kullanılmış. - Bir-iki kısım dışında verilen bilgi konu ile ilgili. - Yapılan araştırma mantıklı bir yapıda yazılmış.	3-5 kaynak kullanılmış. - Verilen bilgi çoğunlukla konu ile ilgili. - Yapılan araştırma belli bir sırada yazılmış.	- Az sayıda (<3) ve çoğunlukla ansiklopedik kaynaklar kullanılmış. - Verilen bilginin büyük bir kısmı konu ile ilgisiz. - Yapılan araştırmanın belli bir sırada yazılması için çok az çaba sarf edilmiş.	- Araştırmada 1 veya 2 kaynak kullanılmış. - Verilen bilgi genellikle konu ile ilgisiz. - Yazıda yapılan araştırmadan düzensizce bahsedilmiş.
Bilimsel Araştırma	- Bilimsel araştırmanın temel prensipleri doğru uygulanmış. - Hipotez, değişken, kontrol gibi terimler doğru kullanılmış.	- Bilimsel araştırmanın temel prensipleri doğru uygulanmış. - Terminoloji büyük ölçüde doğru kullanılmış.	- Bilimsel araştırmanın temel prensipleri doğru uygulanmış. - Yalnızca bazı terimler doğru kullanılmış.	- Bilimsel araştırmanın temel prensipleri doğru uygulanmamış. - Terminoloji uygun kullanılmamış.	- Bilimsel araştırmanın temel prensipleri doğru uygulanmamış. - Temel terimler kullanılmamış veya anlaşılmamış.
Sunum	- Sonuçlar grafik kullanarak, açık bir şekilde sunulmuş. - Yapılan yorum elde edilen sonuçlarla uyumlu. - Sunum, projenin tüm aşamalarını (hipotez, araştırma, sonuçlar, yorum) gösteriyor.	- Sonuçlar grafik kullanarak, açık bir şekilde sunulmuş. - Yapılan yorum, bir-iki küçük fark dışında elde edilen sonuçlarla uyumlu. - Sunum, projenin tüm temel aşamalarını baştan sona gösteriyor.	- Sonuçlar bir-iki ufak hata ile grafik kullanarak sunulmuş. - Yapılan yorumun geçerli olması için küçük bazı düzeltmeler gerekli. - Sunum, projenin tüm temel aşamalarını gösteriyor.	- Sonuçların grafikte gösterimi yeterli değil. - Yapılan yorumun geçerli olması için büyük düzeltmeler gerekli. - Sunum, projenin bazı temel aşamalarını gösteriyor.	- Sonuçlar grafikte gösterilmemiş. - Yapılan yorum geçerli değil. - Sunum, projenin temel aşamalarını göstermiyor ve gelişigüzel hazırlanmış.

SONUÇ VE ÖNERİLER

Bilimsel bilginin öğreniminin nelere bağlı olduğu ve öğrenmenin nasıl gerçekleştiği konularındaki teoriler geliştikçe, öğrenci için belirlenen öğrenim hedefleri de zaman içinde değişmiştir. Artık öğrencinin yalnızca bilgiyi hatırlaması ve uygulayabilmesi değil, üst düzey becerilerini de kullanarak bilgiyi analiz edebilmesi, sentezleyebilmesi ve değerlendirebilmesi, pek çok beceriyi aynı anda kullanabilmesi ve günlük hayatta karşısına çıkabilecek problemleri çözebilmesi beklenmektedir. Öğrencinin bu hedeflerin neresinde olduğunu belirleyebilmek için var olan klasik yöntemlerin yanında, performansa dayalı ölçme yöntemlerine de ihtiyaç duyulmaktadır. Ayrıca, bütünleştirici yaklaşımda öğretim ve ölçmenin birbirinden ayrı süreçler değil, birbirini tamamlayan süreçler olduğu ve kullanılan ölçme yöntemlerinin öğrencinin öğrenmesini etkilediği vurgulanmaktadır. Bu açıdan bakıldığında da, bilginin aktif olarak yapılandırılmasını gerektiren performansa dayalı ölçme yöntemlerine duyulan ihtiyaç belirginleşmektedir. Bütünleştirici yaklaşımda öğrenciye zamanında ve yeterli dönüt verilmesi vurgulanmakta, dolayısıyla öğretim devam ederken öğrencinin ölçülüp, değerlendirilmesi gereği ortaya çıkmaktadır. Benzer şekilde bireyin ön plana çıktığı bütünleştirici yaklaşımda, öğrencinin gerçek kapasitesini göstermesine imkân verecek farklı ve çeşitli ölçme yöntemlerinin uygulanması önerilmektedir. Tüm bunlar performansa dayalı ölçme yöntemlerinin kullanımını gerektirmektedir. Öğrencinin kendi kendini değerlendirmesine olanak sağlayarak, üstbiliş becerilerinin gelişmesini sağlaması da performansa dayalı ölçme yöntemlerinin önemini artırmaktadır.

Türk Eğitim Sistemi'nde son yıllarda yaşanan reform hareketleriyle birlikte, performansa dayalı ölçme yöntemleri ilköğretim Fen ve Teknoloji müfredatına entegre edilmiştir. Benzer entegrasyonun ortaöğretim fen alanları dersleri müfredatı için de yapılması beklenmektedir. Öğretmenler yapılan reformların başarılı olmasında çok önemli bir role sahiptirler (Dori & Herscovitz, 2005). Bu nedenle öğretmenlerin, performansa dayalı ölçme yöntemlerini uygulayabilecek ve öğrencilerin gelişimini takip ederek, onları yönlendirebilecek yeterlikte olmaları gereklidir. Oysa yapılan çalışmaların önemli bir bölümü öğretmenlerin ölçme-değerlendirme bilgi ve becerilerinin istenen ve olması gereken düzeyin altında olduğunu göstermiştir (Çakan, 2004). Bu nedenle, görevine devam eden öğretmenlerin hizmet içi eğitimler ile bu yeterlikleri edinmeleri amaçlanmalıdır. Benzer şekilde öğretmen yetiştirme programlarının, öğretmen adaylarının alanlarında performansa dayalı ölçme yöntemlerini uygulayabilecek yeterliği edinmelerini sağlayacak şekilde düzenlemesi önerilmektedir. Öğretmenlerin performansa dayalı ölçme yöntemlerini uygulamaları için yalnızca yöntemi bilmeleri yeterli değildir, performansa dayalı ölçümlerin temelindeki mantığı ve bütünleştirici yaklaşımın ölçme ve değerlendirme ile ilgili felsefesini de anlamaları önemlidir. Öğretmenlere verilen hizmet içi eğitimde ve öğretmen adaylarına verilen hizmet öncesi eğitimde bu hususun dikkate alınması önerilmektedir. Yapılan bir çalışmada belirtilen husus dikkate alındığında öğretmen adaylarının ölçme ve değerlendirmeye karşı bütünleştirici bir yaklaşım benimsedikleri ve öğretim ve ölçmeyi birbirini besleyen iki süreç olarak gördükleri belirlenmiştir (Ogan-Bekiroğlu, basımda).

Araştırmalar ulusal düzeyde uygulanan çoktan seçmeli sınavların reformlar için bir engel teşkil ettiğini göstermektedir (Tamir, 2003). Tamir bunun nedenlerini çoktan seçmeli sınavların reform içeriği ile uyuşmaması, öğretmenlerin bu tür sınavlara göre öğretim yapması ve öğrencilerin de bu yönde başarılı olacak şekilde çalışmaları olarak vermektedir. Ortaöğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavı ve Öğrenci Seçme Sınavı gibi sınavlar nedeniyle benzer çelişkilerin eğitim sistemimizde yaşanması, yapılan reformların başarılı olmasında bir engel teşkil edebilir.

KAYNAKLAR

- Adams, N., Cooper, G., Johnson, L. & Wojtysiak, K. (1996). Improving student engagement in learning activities. *Master's thesis, Saint Xavier University, Lincolnshire, IL. (ERIC Document Reproduction Service No. ED40076).*
- Alternative Assessments in Math and Science: Moving Toward a Moving Target (1992). *(ERIC Document Reproduction Service No. ED 355 256).*
- Baki, A. & Birgin, O. (2004). Alternatif değerlendirme aracı olarak bilgisayar destekli bireysel gelişim dosyası uygulamasından yansımalar: bir özel durum çalışması. *The Turkish Online Journal of Education Technology – TOJET*, 3(3), Article 11, <http://www.tojet.net/articles/3311.htm>
- Briscoe, C. & Wells, E. (2002). Reforming primary science assessment practices: a case study of one teacher's professional development through action research. *Science Education*, 86, 417-435.
- Carter, C. (1994). Evaluating the quality and equity of alternative assessments. *Improving science and mathematics education: A toolkit for professional developers: Alternative assessment issues* (pp. 243-323). Regional Educational Laboratory Network Program on Science and Mathematics Alternative Assessment for Oregon.
- Century, D. N. (2002). *Alternative and traditional assessments: their comparative impact on students' attitudes and science learning outcomes: an exploratory study.* Unpublished Doctoral dissertation, University of Temple , USA.
- Cole, N. (1988). A realist's appraisal of prospects for unifying instruction and assessment. C. V. Bunderson (Ed.), *Assessment in the Service of Learning issues* (pp. 103-117). Princeton, NJ: Educational Testing Service.
- Çakan, M. (2004). Öğretmenlerin ölçme-değerlendirme uygulamaları ve yeterlik düzeyleri: ilk ve ortaöğretim. *Ankara Üniversitesi Eğitim Bilimleri Dergisi*, 37(2), 99-114.
- Delandshere, G. & Jones, J. H. (1999). Elementary teachers' beliefs about assessment in mathematics: a case of assessment paralysis. *Journal of Curriculum and Supervision*, 14(3), 216-240.
- Dori, Y. J. & Herscovitz, O. (2005). Case-based long-term professional development of science teachers. *International Journal of Science Education*, 27(12), 1413-1446.
- Enger, S. & Yager, R. E. (1998). *The Iowa assessment handbook*. Iowa City: Science Education Center of the University of Iowa.
- Farr, R. & Tone, B. (1994). *Portfolio and performance assessment: helping students evaluate their progress as readers and writers*. Fort Worth: Harcourt Brace College Publishers.
- Gainen, J. & Locatelli, P. (1995). *Assessment for the new curriculum: a guide for professional accounting programs*. Florida: American Accounting Association.
- Graue, M. (1994). Connecting visions of authentic assessment in the realities of educational practice. T. A. Romberg (Ed.), *Assessment in school mathematics issues*. Albany: State University of New York Press.

- Hamilton, L. S. (1994, April). *Validating hands-on science assessments through an investigation of response process*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
- Henessey, G. S. (1995, April). *Teaching with students' ideas in mind*. Paper presented at the annual meeting of the National Association for Research in Science Teaching, San Francisco.
- Herman, J. L., Aschbacher, P. R. & Witner, L. (1992). *A practical guide to alternative assessment*. Virginia: Association for Supervision and Curriculum Developments.
- Jensen, K. (1995). Effective rubric design. *The Science Teacher*, 62(5), 34-37.
- Johnsen, S. (1996). What are alternative assessments?, *Parenting the Gifted*, 19(4), 12-15.
- Karasar, N. (2003). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayın.
- Kamen, M. (1991). Use of drama to evaluate elementary school students. In G. Kulm & S. Malcolm (Eds.), *Science Assessment in the Service of Reform* issues. Washington, D.C.: American Association for Advancement of Science.
- Korkmaz, H. & Kaptan, F. (2005). Fen eğitiminde öğrencilerin gelişimini değerlendirmek için elektronik portfolyo kullanımı üzerine bir inceleme. *The Turkish Online Journal of Education Technology-TOJET*, 4(1), Article 13, <http://www.tojet.net/articles/4113.htm>
- Maclellan, E. (2004). How convincing is alternative assessment for use in higher education? *Assessment & Evaluation in Higher Education*, 29 (3), 311-321.
- Maeroff, G.I. (1991). Assessing alternative assessment. *Phi Delta Kappan*, 73(4), 272-281.
- McMillan, J. H. (1997). *Classroom assessment: principles and practice for effective instruction*. Needham, MA: Allyn & Bacon.
- Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation*, 7(25).
- Moskal, B. M. (2003). Recommendations for developing classroom performance assessments and scoring rubrics. *Practical Assessment, Research & Evaluation*, 8(14).
- Nickelson, D. (2004). Portfolios in physics. *The Science Teacher*, 52-55.
- Nitko, A. J. (2001). *Educational assessment of students*. Upper Saddle River, NJ: Merrill.
- Novak, J. D., Mintzes, J. J. & Wandersee, J. H. (2000). Learning, teaching, and assessment: a human constructivist perspective. In J. J. Mintzes, J. H. Wandersee & J. J. Novak (Eds.), *Assessing science understanding: A human constructivist view issues* (pp. 1-13). San Diego, CA: Academic Press.
- Ogan-Bekiroglu, F. (2004). *Ne kadar başarılı? Klasik ve alternatif ölçme-değerlendirme yöntemleri ve fizikte uygulamalar*, Ankara: Nobel Yayın.
- Ogan-Bekiroglu, F. (in press). Assessing assessment: examination of pre-service physics teachers' attitudes towards assessment and factors affecting their attitudes. *International Journal of Science Education*. doi: 10.1080/09500690701630448

- Palmer, J. (1994). Integrating assessment with curriculum and instruction. *Improving Science and Mathematics Education: A Toolkit for Professional Developers: Alternative Assessment issues* (pp. 112-182). Regional Educational Laboratory Network Program on Science and Mathematics Alternative Assessment for Oregon.
- Pate, E. P., Homestead, E. & McGinnis, K. (1993). Designing rubrics for authentic assessment. *Middle School Journal*, 25-27.
- Pierce, L. V. & O'Malley, J.M. (1992). *Performance and portfolio assessment for language minority students*. Washington, DC: National Clearinghouse for Bilingual Education.
- Pilcher, J. K. (2001, March). *The standards and integrating instructional and assessment practices*. Paper presented at the annual meeting of the American Association of Colleges for Teacher Education, Texas, IL. (ERIC Document Reproduction Service No. ED451190).
- Ramaprasad, A. (1983). On the definition of feedback. *Behavioural Science*, 28(1), 4-13.
- Reichel, A. (1992). Performance assessment: five practical approaches. *Science and Children*, 32(2), 21-25.
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 145-165.
- Shavelson, R. J. & Ruiz-Primo, M. A. (2000). On the psychometrics of assessing science understanding. J. J. Mintzes, J. H. Wandersee & J. D. Novak (Eds), *Assessing science understanding: A human constructivist view issues* (pp. 303-341). San Diego, CA: Academic Press.
- Shepard, L. (1991). Psychometricians' beliefs about learning. *Educational Researcher*, 20(7), 21-27.
- Slater, T. F. (1996). Portfolio assessment strategies for grading first-year university physics students in the USA. *Physics Education*, 31, 82-86.
- Slater, T. F., Ryan, J. M. & Samson, S. L. (1997). Impact and dynamics of portfolio assessment and traditional assessment in a college physics course. *Journal of Research in Science Teaching*, 34(3), 255-271.
- Sluijsmans, D. M. A., Brand-Gruwel, S., van Merriënboer, J. J. G. & Bastiaens, T. J. (2003). The training of peer assessment skills to promote the development of reflection skills in teacher education. *Studies in Educational Evaluation*, 29, 23-42.
- Smith, C. B. (2003). *Alternative forms of assessment*. (ERIC Document Reproduction Service No. ED482404).
- Stiggins, R. (1995). *Student-centered classroom assessment*. New York: Macmillan.
- Stiggins, R. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, 18(1), 23-27.
- Tamir, P. (2003). Assessment and evaluation in science education: opportunities to learn and outcomes. In B. J. Fraser & K. G. Tobin (Eds.), *International Handbook of Science Education issues* (pp. 761-789). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Taras, M. (2002). Using assessment for learning and learning from assessment. *Assessment & Evaluation in Higher Education*, 27(6), 501-510.

- Wiggins, G. (1989). A true test: toward more authentic and equitable assessment. *Phi Delta Kappan*, 70, 703-713.
- Wiggins, G. (1993). *Assessing student performance: exploring the purpose and limits of teaching*. San Francisco, CA: Jossey-Bass Publishers.
- Wiggins, G. (1996). Practicing what we preach in designing authentic. *Educational Leadership*, 54(4), 18-25.
- Wilson, L. D. (1994, April). *A theoretical framework linking beliefs with assessment practices in school mathematics: assessment reforms in search of a theory*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, IL. (ERIC Document Reproduction Service No. ED377215).
- Valdez, P.S. (2001). Alternative assessment: a monthly portfolio project improves student performance. *Science Teacher*, 68(8), 41-43.
- Vitale, M. R. & Romance, N. R. (2000). Portfolios in science assessment: a knowledge-based model for classroom practice. J. J. Mintzes, J. H. Wandersee & J. D. Novak (Eds), *Assessing Science Understanding: A Human Constructivist View* issues (pp. 167-196). San Diego, CA: Academic Press
- Zollman, A. & Jones, D. L. (1994, February). *Accommodating assessment and learning: utilizing portfolios in teacher education with preservice teachers*. Paper presented at the annual meeting of the Research Council on Diagnostic and Prescriptive Mathematics, Texas, IL. (ERIC Document Reproduction Service No. ED368551).
- Zwicky, F. (1969). *Discovery invention and research through the morphological approach*. New York: Macmillan.