

# Fen ve Matematik Öğretmenlerinin Değerlendirme Araçlarıyla İlgili Yaklaşımları ve Uygulamaları

Nergiz Nazlıççek ve Füsun Akarsu

## Özet

*Bu çalışmada matematik, fizik ve kimya öğretmenlerinin değerlendirme araçları/uygulamaları hakkındaki bilgi düzeyleri, yaklaşımları ve bu araçları kullanma alışkanlıkları incelenmektedir. Araştırmacılar tarafından geliştirilen ölçek 143 öğretmene verilmiştir. Verilerin analizi öğretmenlerin yazılı sınav ve soru-cevapla ilgili bilgi düzeylerinin yüksek; buna karşın portfolyo, kavram haritaları, günlük ve deney raporları gibi alternatif araçlar konusundaki bilgi düzeylerinin düşük olduğunu ortaya çıkarmıştır. Öğretmenlerin bu dört alternatif değerlendirme aracına verdikleri önem ile duydukları güvenin ve bunları kullanma alışkanlıklarının diğerlerine göre daha düşük olduğu belirlenmiştir. Ayrıca öğretmenlerin bu araçlarla ilgili yaklaşımları konusunda farklılaşma olduğu gözlenmiştir. En çok güvenilen ve en fazla önemsenen iki uygulamanın derse katılım ve gayret olduğu tespit edilmiştir. Tecrübeye ve cinsiyete göre kullanılan uygulamalarının farklılaşmadığı, sadece fizik ve matematik öğretmenlerinin geleneksel değerlendirme uygulamalarını kullanım düzeylerinin anlamlı bir şekilde farklı olduğu gözlenmiştir.*

*Anahtar sözcükler:* Fen öğretmenleri, matematik öğretmenleri, geleneksel ve alternatif değerlendirme uygulamaları

## Giriş

Öğrenme sürecinde sonuçtan ziyade sürecin ön plana çıkması hem kullanılacak öğrenme ve öğretme etkinliklerini hem de değerlendirme uygulamalarını değiştirmeyi zorunlu kılmaktadır. Ürünü ölçmeye çalışan geleneksel değerlendirme araçlarının yanı sıra öğrencinin öğrenme sürecindeki tüm performansını ve gücünü ortaya çıkarmayı hedefleyen alternatif değerlendirme araçlarının kullanılması son zamanlarda sıkça vurgulanmaya başlanmıştır.

Amerika'daki Ulusal Matematik Öğretmenleri Organizasyonu (NCTM, 1995) değerlendirmenin iki amacını öğretimle ilgili kararların verilmesi ve öğrencinin gelişiminin izlenmesi olarak tanımlamıştır. NCTM (1995) bu amaçlar bağlamında değerlendirme yaklaşımında nasıl bir değişim yaşandığını da çok detaylı bir şekilde açıklamaktadır. Öğretimle ilgili doğru kararları verebilmek için tek bir bilgi kaynağı yerine çok çeşitli değerlendirme uygulamalarından veri toplanması önerilmektedir. Ayrıca, değerlendirmenin belirli aralıklarla yapılması yerine öğretimle entegre edilmesi tavsiye edilmektedir. Diğer bir değişim de öğrencilerin belirli olguları ve becerileri ne kadar bildiklerini değerlendirmekten ziyade, öğrencinin alanla ilgili tüm gücünü değerlendirme yönündedir. Bunlara ek olarak öğrenciye sadece cevabının doğru olup olmadığını söylemek yerine performansı konusunda sürekli ve kapsamlı bir geri

---

*Nergiz Nazlıççek, Öğr. Gör., Boğaziçi Üniversitesi, Eğitim Fakültesi, Ortaöğretim Fen ve Matematik Alanları Eğitimi Bölümü, İstanbul.*

*Füsun Akarsu, Prof. Dr., Boğaziçi Üniversitesi, Eğitim Fakültesi, Ortaöğretim Fen ve Matematik Alanları Eğitimi Bölümü, İstanbul. [akarsuf@boun.edu.tr](mailto:akarsuf@boun.edu.tr)*

bildirim verilmesinin çok önemli olduğunun altı çizilmektedir. Ayrıca öğrencilerin değerlendirme sürecinde kendi gelişimlerini değerlendirerek daha aktif hale gelmeleri önerilmektedir (bulduğu eser: Reys, Suydam, Lindquist ve Smith, 1998, 53).

Önemli değişimlerden biri de geleneksel değerlendirme araçlarının yanı sıra alternatif değerlendirme araçlarının kullanılmasıdır. Geleneksel değerlendirme araçları eşleştirme, çoktan seçmeli, doğru/yanlış, boşluk doldurma ya da açık uçlu soruları içeren test, yazılı sınav ve küçük sınav gibi değerlendirme araçlarını kapsamaktadır. Otantik değerlendirme olarak da adlandırılan alternatif değerlendirme araçları ise öğrencinin performansını ölçmeye yönelik işler, projeler, portfolyo ve günlük gibi çeşitli uygulamaları içermektedir. Gözlem, öğrencinin öz değerlendirmesi ve öğrenciyle yapılan yüz yüze görüşmeler kullanılacak diğer alternatif değerlendirme araçlarından bazılarıdır (Bol ve diğerleri, 1998, 54).

Değerlendirme yaklaşımındaki değişime paralel olarak, öğretmenlerin söz konusu değişimleri ne ölçüde hayata geçirdiği, değerlendirmeyle ilgili fikirleri ve hangi değerlendirme uygulamalarını kullandığı araştırılmaya başlanmıştır. Örneğin Adams ve Yang-Hsu (1998) öğretmenlerin değerlendirmeyi nasıl kavradıkları ve değerlendirmeyle ilgili uygulamalarını araştırmıştır. Çalışmaya 1., 2., 3. ve 4. sınıfta ders veren toplam 269 matematik öğretmeni katılmıştır. Genel olarak bakıldığında öğretmenler değerlendirmeyle ilgili tüm uygulamaları önemli ya da çekimser olarak işaretlemişlerdir. Öğretmenlerin en çok önem verdiği ve kullandığı değerlendirme uygulaması öğretmen tarafından yapılan gözlem, ikincisi ise öğrenci performansdır. Verilen önem ve kullanım sıklığı bakımından ortalaması en düşük maddenin ise kompozisyon (essay) olduğu görülmüştür.

Bunlara ek olarak, öğretmenlerin değerlendirmeyi nasıl algıladıkları ve kullandıkları araçlar ile sınıf arasındaki ilişkiye bakıldığında; birinci ve dördüncü sınıfta ders veren öğretmenler için açık uçlu soruların daha önemli olduğu, üçüncü ve dördüncü sınıf öğretmenlere göre ödevin daha fazla önemli olduğu görülmektedir. Oysa ikinci ve üçüncü sınıf öğretmenlerinin de öğretmen tarafından hazırlanmış testlerin daha önemli olduğunu düşündükleri ve bunları daha fazla kullandıkları gözlenmiştir. Ayrıca, ikinci ve üçüncü sınıf öğretmenlerinin öğretmen tarafından hazırlanan testleri, dördüncü sınıf öğretmenlerinin ise standart testleri birinci sınıf öğretmenlerine göre daha sık kullandıkları belirlenmiştir.

Bol, Stephenson, O'Connel ve Nunnery (1998) sınıf düzeyine ek olarak öğretmenlik tecrübesinin ve branşın kullanılan değerlendirme uygulamaları üzerindeki etkisini araştırmıştır. Çalışmaya 34 okuldan toplam 893 öğretmen katılmıştır. Öğretmenler en sık kullandıkları ve geçerliliğine en fazla inandıkları ölçme metotlarının performans ve gözleme dayalı olduğunu belirtmiştir. Ayrıca, öğretmenlerin alternatif değerlendirme metotlarını kullanma sıklıkları artınca, bu tip değerlendirmeleri geliştirme yönünde kendilerine daha fazla güvenmeye başladıkları ve bunların geçerliliğine daha fazla inandıkları ortaya çıkmıştır. Benzer ilişkinin geleneksel değerlendirme yöntemleri için de geçerli olduğu bulunmuştur. Tecrübenin kullanılan değerlendirme araçları üzerindeki etkisi incelendiğinde, daha fazla tecrübeye sahip öğretmenlerin az tecrübelilere göre alternatif değerlendirme araçlarını daha sık kullandıkları belirlenmiştir. Bunlara ek olarak, ilkökul öğretmenlerinin lisedeki öğretmenlere kıyasla alternatif değerlendirme stratejilerini daha sık kullandıkları gözlenmiştir. Değerlendirme araçları branşlara göre incelendiğinde ise matematik

öğretmenlerinin geleneksel değerlendirme yöntemlerini fen, sosyal bilgiler ve İngilizce öğretmenlerine göre daha az kullandıkları ortaya çıkarılmıştır.

Street-Trepanier, McNair ve Donegan (2001) tarafından yapılan çalışmada ise anaokulu ile 3. 4. ve 5 sınıf öğretmenlerinin kullandıkları standart testler ve diğer değerlendirme uygulamaları ve bunlarla ilgili görüşleri kıyaslanmıştır. İki grup öğretmen de gözlem, portfolyo, yazılı ödev gibi çeşitli değerlendirme ölçülerini kullandıklarını ve güvendiklerini belirtirken, standart testleri kullanmadıklarını ve bunlara değer vermediklerini söylemişlerdir. Ancak anaokuluna ders veren öğretmenler belirli becerileri değerlendirme, gözlem, notlar yazma ve kontrol çizelgeleri (checklist) gibi etkinlikleri, üst sınıfları okutan öğretmenler kendi hazırladıkları ve kitaplardan aldıkları testleri daha sık kullanmakta ve bunlara daha fazla güvenmektedirler.

McMillan, Myran ve Workman (2002) öğretmenlerin kullandıkları değerlendirme uygulamalarına ek olarak nasıl not verdiğini de araştırmıştır. Araştırmaya katılan toplam 900 öğretmen not verme süreçlerini, kullandıklarını değerlendirme türünü ve değerlendirmenin bilişsel seviyesini belirlemede kullandıkları ölçütleri belirtmiştir. Öğretmenlerin not vermek için akademik performans, gayret ve gelişim, ödev, diğer öğrencilerle kıyas, diğer öğretmenlerin verdikleri notların dağılımı ve daha önceden belirlenmiş ölçütlerle kıyas olmak üzere altı temel etmeni dikkate aldıkları tespit edilmiştir. Ayrıca öğretmenlerin nesnel sınavları ve kendilerinin hazırladıkları sınavları daha fazla kullandıkları bulunmuştur. Matematik ve dil bilgisi olmak üzere iki derste kullanılan değerlendirme araçlarında fark olmadığı ve seviyelere göre kullanılan değerlendirme uygulamalarının değişmediği görülmüştür.

Schmidt ve Brosnan (1996) tarafından çalışmada öğretmenlerin %18'den azının alternatif değerlendirme araçlarını kullandıkları görülmüştür. Öğretmenlerin çoğu kendilerine hizmet içi eğitim verilmesinin, bu tür araçlarının nasıl uygulanacağını görmesinin ve zaman gibi etmenlerin söz konusu araçları kullanmalarını teşvik edeceğini belirtmiştir. Öğretmenler dönem notunu verirken çoğunlukla test, küçük sınav, ödev gibi geleneksel araçları kullanmaktadır. Alternatif değerlendirme araçlarının raporlamada çok az kullanıldığı gözlenmiştir. Bunun nedeni de, raporlama yöntemlerinin bu araçları kullanmayı gerektirmemesi olarak açıklanmaktadır.

Görüldüğü gibi, yapılan araştırmaların çoğu daha çok ilköğretim düzeyinde kullanılan değerlendirme etkinlikleri ve bunların seviyelere göre nasıl değiştiğini incelemeye yöneliktir. Ancak, ortaöğretim düzeyinde de değerlendirme yaklaşımında yaşanan değişimlerin ne kadar hayata geçirildiği, öğretmenlerin çeşitli değerlendirme uygulamaları ile ilgili yaklaşımları, bu uygulamaları hangi seviyede kullandıkları ve uygulamaların branşlar ve tecrübe düzeyine göre nasıl değiştiği konusunda da bilgiye ihtiyaç vardır. Eldeki çalışmada bu değişkenler incelenmiş ve buna bağlı olarak araştırmanın problemleri aşağıdaki gibi belirlenmiştir:

Ortaöğretim düzeyindeki fizik, kimya ve matematik öğretmenlerinin

1. Değerlendirme araçları/uygulamaları ile ilgili bilgi düzeyleri nedir?
2. Değerlendirme araçlarını/uygulamaları ile ilgili yaklaşımları nelerdir?
3. Değerlendirme araçlarını/uygulamalarını kullanım sıklıkları nedir?
4. Değerlendirme araçlarını/uygulamalarını kullanma sıklıkları branşa, tecrübeye, okul türüne ve cinsiyete göre nasıl değişmektedir?

5. Değerlendirme araçlarını/uygulamalarını kullanım sıklıklarına ilişkin kendi beyanları ile yapılan gözlemler arasında nasıl bir ilişki vardır?

### **Yöntem**

#### **Örneklem**

Çalışmaya özel okullardan 59 ve devlet okullarından 87 olmak üzere toplam 143 fizik, kimya ve matematik öğretmeni katılmıştır. Öğretmenlerden 11'i bayan, 15'i erkektir. Öğretmenlerin 12'si matematik öğretmeni, 8'i kimya ve 6'sı fizik öğretmenidir. Öğretmenler tecrübelerine göre iki gruba ayrılmıştır. Altı seneden az tecrübeye sahip olanlar "az tecrübeli", daha fazla tecrübeliler ise "çok tecrübeli" kategorisinde incelenmiştir. Buna göre öğretmenlerin 10'u az tecrübeli, 16'sı çok tecrübeli grubunda yer almaktadır.

Çalışmaya Boğaziçi Üniversitesi, Ortaöğretim Fen ve Matematik Alanları Eğitimi Bölümü öğrencilerden Okul Deneyimi I ve II dersini alan 143 3. sınıf öğrencisi de gözlemci olarak katılmıştır. Öğrenciler bu derste 8 hafta boyunca özel okullara giderek rehber öğretmenlerinin gözetiminde ders gözlemi, okulun sayısal olarak incelenmesi, sınıf dışı etkinliklerin ve değerlendirme uygulamalarının gözlenmesi gibi çeşitli etkinliklerle eğitim ve öğretim ortamını tanımaya çalışmaktadır. Öğrenciler değerlendirme uygulamaları ile ilgili verileri bu dersin gereklerinden biri olarak toplamıştır.

#### **Değişkenler ve Ölçme Araçları**

##### *Ölçme Aracı*

Araştırma kapsamında "Öğretmenlerin Değerlendirme Araçlarıyla İlgili Yaklaşımları ve Kullandıkları Değerlendirme Araçları" adlı anket geliştirilmiştir. Bu anketin geliştirilmesinde Adams ve Yang-Hsu (1998) ile Bol ve diğerlerinin (1998) öğretmenlerin yararlandıkları değerlendirme uygulamalarını ölçmek için kullandıkları ölçeklerden faydalanılmıştır.

Anket toplam üç bölümden oluşmaktadır. Her bölümde aynı 23 değerlendirme uygulaması/aracı bulunmaktadır. Maddeler geleneksel, alternatif ve diğer olarak üç gruba ayrılmaktadır. Geleneksel değerlendirme araçlarıyla ilgili boyutta test, yazılı sınav, ödev gibi araçlar bulunmaktadır. Buna karşın alternatif değerlendirme araçları boyutu ise portfolyo, günlük, kavram haritaları, öğrencinin kendini değerlendirmesi ve gözlem gibi araçları kapsamaktadır. Problem çözme alternatif değerlendirme araçlarından biri olarak görülmesine rağmen, bu çalışmada öğretmenler tarafından büyük bir ihtimalle fen ve matematik problemlerini çözmeye eş tutulacağı düşünüldüğünden, geleneksel değerlendirme araçları boyutunda yer almıştır. Üçüncü boyutta ise değerlendirmede dikkate alınan; derse katılım, gayret ve gelişim gibi unsurlar veya değerlendirme ölçütlerinin kullanımı vardır.

### *Anketin Güvenilirliği*

Her bölümün güvenilirlik analizi ayrı ayrı alfa katsayısı hesaplanarak yapılmıştır. Öğretmenlerin değerlendirme uygulamaları/araçlarıyla ilgili yaklaşımlarına yönelik ilk iki bölümden birincisinin alfa katsayısı 0,86, ikinci bölümünki ise 0,92 olarak bulunmuştur. Öğretmenlerin değerlendirme uygulamaları/araçlarını kullanma alışkanlıklarına ilişkin üçüncü bölümün alfa katsayısı 0,89 ve değerlendirme uygulamalarını kullanım sıklığını ölçmek için kullanılan gözlem formunun alfa katsayısı da 0,87 olarak bulunmuştur. Sonuçlar incelendiğinde tüm bölümlere ilişkin alfa katsayısının güvenilirlik için kabul edilebilir nitelikte olduğu düşünülmektedir.

### *Anketin Geçerliliği*

Anketin kapsam geçerliliği için iki uzmanın görüşüne başvurulmuştur. İki uzman da anketin kapsam geçerliliği konusunda olumlu görüşlerini bildirmişlerdir.

### *Değişkenler*

1) *Öğretmenlerin Değerlendirme Araçları/ Uygulamaları Hakkındaki Bilgi Düzeyleri:* Bu değişken “Öğretmenlerin Değerlendirme Araçlarıyla İlgili Fikirleri ve Kullandıkları Değerlendirme Araçları” anketinin I. bölümün sol tarafında sorulan maddelerle ölçülmüştür. Her madde için 3 cevap seçeneği sunularak, 1 ile 3 arasında “hiç bilmiyorum” dan “biliyorum” a doğru derecelendirilmiştir.

2) *Öğretmenlerin Değerlendirme Araçları/Uygulamalarıyla İlgili Yaklaşımları:* Anketin birinci ve ikinci bölümünde verilen maddelerle ölçülen bu değişken değerlendirme uygulamalarına/araçlarına verilen önem ile öğrenci performansını ve başarısını gösterme konusunda bu araçlara/uygulamalara duyulan güveni kapsamaktadır. İki bölümde de her madde için verilen cevap seçeneği 1 den 5'e kadar derecelendirilmiştir.

3) *Öğretmenlerin Değerlendirme Araçlarını/ Uygulamalarını Kullanma Alışkanları:* Bu değişken, anketin III.bölümünde yer alan 1'den 5'e doğru derecelendirilen maddelerle ölçülmüştür.

4) *Öğretmenlerin Değerlendirme Araçlarını/Uygulamalarını Kullanım Sıklığı:* Aday öğretmenler tarafından, öğretmenlerin derslerinde yapılan gözlemler sırasında doldurulan gözlem formuyla ölçülmüştür.

### *Veri Toplama Süreci*

Veri toplama aşamasına geçilmeden önce gözlem yapacak öğrencilere çalışmayla ilgili kısa bir bilgi verilmiştir. Öğrenciler daha sonra anketi rehber öğretmenlerine götürüp doldurmasını rica etmişler ve rehber öğretmenlerinin bir hafta boyunca beş dersini izleyerek gözlem formunu doldurmuştur.

### ***Verilerin Analizi***

Değerlendirme araçları/uygulamaları ile ilgili bilgi düzeyi, yaklaşım ve bu araçları/uygulamaları kullanma alışkanlığına ilişkin maddelerin ortalamaları ve standart sapmaları hesaplanmıştır. Cinsiyet ile tecrübenin öğretmenlerin değerlendirme araçlarını/uygulamalarını kullanma alışkanlıkları üzerindeki etkisini incelemek üzere bağımsız gruplar için t-testi ve branşın etkisini incelemek için ise, tek yönlü varyans analizi kullanılmıştır. Ayrıca değerlendirme araçlarını/ uygulamalarını kullanma sıklığı ve kullanma alışkanlığı değişkenleri arasındaki ilişkinin analizi için Pearson korelasyon testinden faydalanılmıştır.

### **Bulgular**

#### ***Öğretmenlerin Değerlendirme Uygulamalarıyla İlgili Bilgi Düzeyleri***

Öğretmenlerin ölçekte verilen değerlendirme uygulamaları ile ilgili bilgi düzeylerinin ortalamaları Tablo 1’de gösterilmiştir.

Ortalamalar incelendiğinde öğretmenlerin en çok bildikleri değerlendirme uygulamalarının/araçlarının yazılı sınav, soru cevap, problem çözme ve ödev olduğu gözlenmiştir. Ölçme aracıyla ilgili bölümde belirtildiği gibi problem çözme bu çalışmada geleneksel değerlendirme araçları boyutunda düşünüldüğünden, soru-cevap hariç ortalamaları yüksek çıkan değerlendirme uygulamaları/araçlarının geleneksel değerlendirme araçları olduğu söylenebilir. Ortalaması en düşük maddeler ise deney raporları, günlük, portfolyo ve kavram haritaları olarak bulunmuştur. Bu maddelerin hepsi de alternatif değerlendirme araçlarını içermektedir. Dolayısıyla, öğretmenlerin alternatif değerlendirme araçlarıyla ilgili bilgi düzeylerinin geleneksel araçlarla ilgili bilgi düzeylerine göre daha düşük olduğu söylenebilir.

**Tablo 1.** Öğretmenlerin değerlendirme araçları/uygulamalarıyla ilgili bilgi düzeylerinin ortalamaları

No	Değerlendirme Uygulaması/ Araçları	Ortalama (1-3)
1	Yazılı Sınav	3,00
2	Test	2,95
3	Küçük Sınav	2,96
4	Açık Uçlu Sorular	2,83
5	Soru-Cevap	2,98
6	Tartışma	2,93
7	Ödev	2,98
8	Gözlem	2,87
9	Portfolyo	2,12
10	Günlük	2,47
11	Öğrencinin Öz Değerlendirmesi	2,71
12	Yüz yüze görüşme	2,93
13	Gelişim	2,92
14	Gayret	2,95
15	Derse katılım	2,95
16	Derse karşı tutum	2,94
17	Öğrenme stili	2,81
18	Problem çözme	2,98
19	Projeler	2,69
20	Kavram haritaları	2,34
21	Deney raporları	2,59
22	Grupça yapılan ödevler	2,85
23	Öğrenci sunumları	2,82

### ***Öğretmenlerin Değerlendirme Uygulamalarıyla İlgili Yaklaşımları***

Öğretmenlerin değerlendirme uygulamaları/araçları ile ilgili yaklaşımları daha önce de belirtildiği gibi bu uygulamalara verilen önem ve duyulan güven olmak üzere iki bölümde incelenmiştir. Değerlendirme uygulamalarına/araçlarına verilen önemle ilgili maddelerin ortalamaları ve standart sapmaları Tablo 2’de verilmiştir.

Tablo 2’de görülebileceği gibi, madde ortalamaları 3,00 ve 4,78 arasında değişmektedir. Diğer bir deyişle, en düşük ortalamaya sahip değerlendirme uygulaması günlük iken, en yüksek ortalamaya sahip madde ise gayrettir. Aynı zamanda bu madde standart sapması en düşük olan üçüncü maddedir. Gayretten sonra en yüksek ortalamaya sahip diğer maddeler ise derse katılım, tutum ve gelişimdir. Ayrıca standart sapması

0,53 olan derse katılım maddesi en düşük standart sapmaya sahiptir. Tutum maddesi de standart sapması en düşük olan ikinci maddedir. Bunlara bağılı olarak hem ortalamaları yüksek olan hem de standart sapmaları çok düşük olan derse katılım, tutum ve gayret maddelerinin önemi konusunda öğretmenlerin fikir birliği içinde oldukları söylenebilir. Gayret, derse katılım, tutum ve gelişme gibi maddelerin ortalamalarının yüksek çıkması, öğretmenlerin geleneksel değerlendirme araçları dışındaki uygulamalara da önem verdiklerini göstermektedir.

**Tablo 2.** Öğretmenlerin değerlendirme araçları/ uygulamalarına verdikleri önemle ilgili betimsel istatistikler

No	Değerlendirme Uygulaması/Araçları	Ortalama	SD
		(1-5)	
1	Yazılı Sınav	4.40	0.88
2	Test	3.78	1.16
3	Küçük Sınav	3.88	1.12
4	Açık Uçlu Sorular	3.91	1.11
5	Soru-Cevap	4.54	0.72
6	Tartışma	4.36	0.89
7	Ödev	4.07	1.10
8	Gözlem	4.22	1.01
9	Portfolyo	3.09	1.46
10	Günlük	3.00	1.48
11	Öğrencinin Öz Değerlendirmesi	4.14	1.04
12	Yüz yüze görüşme	4.43	0.92
13	Gelişim	4.67	0.72
14	Gayret	4.78	0.63
15	Derse katılım	4.77	0.53
16	Derse karşı tutum	4.75	0.58
17	Öğrenme stili	4.39	0.94
18	Problem çözme	4.61	0.68
19	Projeler	3.99	1.11
20	Kavram haritaları	3.40	1.34
21	Deney raporları	3.50	1.61
22	Grupça yapılan ödevler	3.68	1.20
23	Öğrenci sunumları	3.91	1.16

Ortalaması en düşük olan dört madde sırasıyla günlük (3,00), portfolyo (3,09), kavram haritaları (3,40) ve deney raporlarıdır (3,50). Ancak bu maddeler standart sapması da en yüksek çıkan maddelerdir. Dolayısıyla bu maddeler üzerinde öğretmenlerin hemfikir olmadıkları düşünülmektedir.



Öğretmenlerin çeşitli değerlendirme araçlarına ne kadar güvendiklerine ilişkin maddelerin betimsel istatistikleri Tablo 3’de sunulmuştur.

**Tablo 3.** Öğretmenlerin değerlendirme araçları/uygulamalarına duydukları güvenle ilgili betimsel istatistikler

No	Değerlendirme Uygulaması/ Araçları	Ortalama	SD
1	Yazılı Sınav	4,16	0,89
2	Test	3,45	0,98
3	Küçük Sınav	3,78	0,96
4	Açık Uçlu Sorular	3,85	0,95
5	Soru-Cevap	4,31	0,77
6	Tartışma	4,12	1,11
7	Ödev	3,68	1,04
8	Gözlem	3,81	1,11
9	Portfolyo	2,95	1,31
10	Günlük	2,96	1,31
11	Öğrencinin Öz Değerlendirmesi	3,88	1,13
12	Yüz yüze görüşme	4,37	0,79
13	Gelişim	4,45	0,77
14	Gayret	4,55	0,76
15	Derse katılım	4,60	0,68
16	Derse karşı tutum	4,39	0,88
17	Öğrenme stili	4,12	0,95
18	Problem çözme	4,51	0,69
19	Projeler	3,76	1,27
20	Kavram haritaları	3,35	1,37
21	Deney raporları	3,26	1,53
22	Grupça yapılan ödevler	3,50	1,19
23	Öğrenci sunumları	3,74	1,25

Ortalaması en yüksek çıkan maddenin derse katılım (4,60) olduğu gözlenmiştir. En düşük ortalamaya sahip madde ise portfolyodur (2,95). Derse katılım maddesinin 0,68 olarak bulunan standart sapması da gözlenen en düşük standart sapmadır. Bu yüzden öğretmenlerin derse katılımın öğrencinin başarısını ve performansını göstermede çok doğru sonuçlar verdiği konusunda hem fikir oldukları söylenebilir. Derse katılım maddesinden sonra en yüksek ortalamaya sahip diğer üç madde sırasıyla gayret (4,55), problem çözme (4,51) ve gelişimdir (4,45). Gayret ve gelişim önem sıralamasında da ilk sıralarda yer alan maddelerdir.

Bu bölümde de en az güven duyulan araçların önemle ilgili bulgulara paralel şekilde alternatif değerlendirme araçları olduğu tespit edilmiştir. Önem ve güven

ortalaması sıralamasında en altta kalan maddeler farklı olmasına rağmen, en düşük güven ortalamalarına sahip dört maddenin bir öndeki bölümde olduğu gibi deney raporları, portfolyo, günlük ve kavram haritaları olduğu belirlenmiştir. Yine önemle ilgili bulgulara paralel şekilde bu maddelerin standart sapmaları diğer maddelere göre daha yüksek çıkmıştır.

### ***Öğretmenlerin Değerlendirme Uygulamalarını Kullanma Alışkanlıkları***

Tablo 4'de görüldüğü gibi en yüksek ortalama problem çözmede (4,38) görülürken, en düşük ortalama ise portfolyoda görülmektedir. Problem çözme maddesinden sonra en yüksek ortalama sahip maddeler sırasıyla soru-cevap (4,32), katılım (4,29) ve gayrettir. Problem çözme maddesi önem sıralamasında beşinci sırada, güven sıralamasında ise üçüncü sırada yer almıştır. Önem ve güven sıralamasında ilk ikide yer alan gayret ve derse katılım ise üçüncü ve dördüncü sırada yer almaktadır. Kullanım sıklığı açısından ikinci sırada yer alan soru-cevap maddesine ise ne önem ne de güven sıralamalarında üst sıralarda yer almıştır.

Dolayısıyla öğretmenlerin en çok önem verdikleri ve en çok kullandıklarını söyledikleri değerlendirme uygulamaları veya en çok güvendikleri ve en çok uyguladıklarını söyledikleri değerlendirme yollarının ne kadar örtüştüğü konusunda kesin bir sonuç çıkarmak zor görünmektedir. Ancak öğretmenlerin en az kullandıklarını söyledikleri değerlendirme uygulamalarının/araçlarının güven ve önem sıralamasındaki aynı dört alternatif değerlendirme aracı yani kavram haritaları, portfolyo, günlük ve deney raporları olduğu gözlenmiştir. Aslında bu maddelerin öğretmenler tarafından en az bilinen değerlendirme araçları olduğu bulgusu dikkate alındığında, öğretmenlerin bu araçları bilmedikleri için kullanmadıkları söylenebilir.

**Tablo 4.** Öğretmenlerin değerlendirme araçları/uygulamalarını kullanma alışkanlıkları ile ilgili betimsel istatistikler

No	Değerlendirme Uygulaması/Araçları	Ortalama	SD
1	Yazılı Sınav	4,09	0,83
2	Test	3,57	1,14
3	Küçük Sınav	3,25	1,36
4	Açık Uçlu Sorular	3,38	1,16
5	Soru-Cevap	4,32	0,83
6	Tartışma	3,64	1,13
7	Ödev	3,75	1,18
8	Gözlem	3,34	1,21
9	Portfolyo	2,35	1,26
10	Günlük	2,41	1,29
11	Öğrencinin Öz Değerlendirmesi	3,05	1,24
12	Yüz yüze görüşme	3,72	1,12
13	Gelişim	3,83	1,09
14	Gayret	4,20	0,89
15	Derse katılım	4,29	0,86
16	Derse karşı tutum	4,03	0,94
17	Öğrenme stili	3,54	1,16
18	Problem çözme	4,34	0,79
19	Projeler	2,80	1,22
20	Kavram haritaları	2,38	1,23
21	Deneysel raporları	2,60	1,40
22	Grupça yapılan ödevler	2,80	1,24
23	Öğrenci sunumları	2,78	1,34

#### ***Okul Türü ve Değerlendirme Uygulamaları***

Okul türünün öğretmenlerin değerlendirme araçlarını/uygulamalarını kullanma alışkanlıkları üzerindeki etkisi geleneksel, alternatif ve diğer olmak üzere üç boyutta incelenmiştir. Her öğretmenin üç boyutla ilgili maddelerden aldığı toplam puan hesaplanmıştır. Tablo 5’te sunulan sonuçlar incelendiğinde devlet ve özel okulda çalışan öğretmenler arasında sadece geleneksel değerlendirme araçlarının kullanımında anlamlı bir farklılık gözlemlenmiştir.

**Tablo 5.** Üç değerlendirme boyutunun kullanım düzeyinin okul türüne göre kıyaslanması

		N	Ortalama	SD	t	P
<b>Geleneksel</b>	Devlet	68	17,66	3,70	-5,06	,00
	Özel	49	20,78	2,58		
<b>Alternatif</b>	Devlet	50	37,08	9,41	-1,69	,94
	Özel	44	40,68	11,25		
<b>Diğer</b>	Devlet	65	19,42	3,91	-1,17	,246
	Özel	48	20,31	4,22		

**Branş ve Değerlendirme Uygulamaları**

Fizik, kimya ve matematik öğretmenlerinin geleneksel, alternatif ve diğer değerlendirme araçlarıyla ilgili üç boyuta ilişkin maddelerden aldıkları puanların ortalamaları arasındaki farka bakmak için tek faktörlü varyans analizi yapılmıştır. Tablo 6'da görüldüğü gibi sadece geleneksel değerlendirme araçlarıyla ilgili boyutta anlamlı bir farklılık çıkmıştır.

**Tablo 6.** Üç değerlendirme boyutunun kullanım düzeyinin branşa göre kıyaslanması

Boyutlar	Kareler Toplamı	df	Ortalama Kareler	F	p
<b>Geleneksel</b>	Gruplar arası	87.37	2	43.693.460.03	
	Gruplar içi	1425.41	113	12.61	
	Toplam	1512.78	115		
<b>Alternatif</b>	Gruplar arası	358.11	2	179.051.700.19	
	Gruplar içi	9456.37	90	105.07	
	Toplam	9814.47	92		
<b>Diğer</b>	Gruplar arası	64.05	2	32.032.000.14	
	Gruplar içi	1746.95	109	16.03	
	Toplam	1811.00	111		

Geleneksel değerlendirme araçlarıyla ilgili boyuta ilişkin olarak grup ortalamaları arasındaki farkın daha detaylı analizi için Tukey testi yapıldığında sadece matematik ve fizik öğretmenlerin ortalamaları arasında anlamlı bir fark bulunmuştur (Tablo 7).

**Tablo 7.** Geleneksel değerlendirme araçlarını kullanım düzeyinin branşlara göre çoklu karşılaştırılması

	<b>Branş</b>	<b>Branş</b>	<b>Ortalama Farkı</b>	<b>SD Hata</b>	<b>P</b>
<b>Geleneksel Matematik</b>	Kimya		1.27	0.87	0.32
	Fizik		1.93	0.75	0.03
Kimya	Matematik		-1.27	0.87	0.32
	Fizik		0.66	0.93	0.76
Fizik	Matematik		-1.93	0.75	0.03
	Kimya		-0.66	0.93	0.76

***Tecrübe, Cinsiyet ve Değerlendirme Uygulamaları***

Az ve çok tecrübeli öğretmenlerin geleneksel, alternatif ve diğer değerlendirme araçları/uygulamalarını kullanım düzeyleri arasında anlamlı bir fark gözlenmemiştir. Aynı şekilde bayan ve erkek öğretmenler arasında hiç bir boyutta anlamlı bir fark bulunmamıştır.

***Değerlendirme Uygulamalarını/Araçlarını Kullanım Sıklığı ve Kullanma Alışkanlığı Arasındaki İlişki***

Kullanım sıklığı ve alışkanlığı arasındaki ilişki geleneksel, alternatif ve diğer değerlendirme uygulamalarıyla ilgili yapılan gözlem ve bu uygulamaları kullanma alışkanlığı arasındaki korelasyon katsayısı hesaplanarak incelenmiştir (Tablo 8). Geleneksel ( $r=0,64$ ,  $p<0,01$ ), diğer ( $r=0,410$ ,  $p<0,01$ ) ve alternatif ( $r=0,410$ ,  $p<0,01$ ) değerlendirme uygulamalarıyla ilgili yapılan gözlemler ve öğretmenlerin beyanları arasında arasında manidar bir ilişki belirlenmiştir. Fakat, üç gruba ilişkin yapılan gözlem ve öğretmenlerin bu araçları kullanma alışkanlıklarının ortalamalarının anlamlı bir şekilde farklı olduğu gözlenmiştir. Diğer bir deyişle öğretmenlerin beyanlarının ortalamaları gözlem ortalamalarından daha yüksek çıkmıştır.

**Tablo 8.** Değerlendirme uygulamalarını/araçlarını kullanım sıklığı ve kullanma alışkanlığı arasındaki ilişki

	<b>N</b>	<b>Ortalama</b>	<b>SD</b>	<b>t</b>	<b>P</b>	<b>r</b>	<b>p</b>
<b>Geleneksel</b> Alışkanlık	87	19.03	3.91	7.37	,00	0,64	,00
	Gözlenen	87	16.33				
<b>Alternatif</b> Alışkanlık	82	19.54	4.33	12.48	,00	0,41	,00
	Gözlenen	82	13.44				
<b>Diğer</b> Alışkanlık	68	37.51	10.23	11.27	,00	0,41	,00
	Gözlenen	68	23.03				

### Tartışma

Öğretmenlerin değerlendirme uygulamaları/ araçları ile ilgili bilgi düzeylerine bakıldığında, yazılı sınav ve ödev gibi geleneksel değerlendirme uygulamalarıyla ilgili bilgi düzeylerinin günlük, portfolyo, deney raporları ve kavram haritaları gibi alternatif değerlendirme araçlarına göre daha yüksek olduğu belirlenmiştir.

Öğretmenlerin değerlendirme uygulamaları ile yaklaşımları güven ve önem olmak üzere iki boyutta incelenmiştir. Güven ve önem sıralamasında yerleri farklı olsa da ilk iki sırada derse katılım ve gayret yer almıştır. En çok güvenilen ve önem verilen iki değerlendirme uygulamasının derse katılım ve gayret olduğu söylenebilir. En az güvenilen ve önem verilen değerlendirme uygulamalarının portfolyo, günlük, kavram haritası ve deney raporları olduğu tespit edilmiştir. Ayrıca bu maddelere verilen önem ve duyulan güven puanlarının standart sapmalarının diğer maddelere göre daha yüksek çıkması, öğretmenlerin bu araçlarla ilgili yaklaşımları konusunda fikir birliğine varamadıklarına işaret etmektedir. Öğretmenlerin bu araçlarla ilgili bilgi düzeylerinin diğer değerlendirme uygulamalarıyla ilgili bilgi düzeylerinden düşük olduğu dikkate alınırsa bilgi eksikliğinin, alternatif değerlendirme araçlarına daha az güvenilmesinde, önem verilmesinde, bunların daha az kullanılmasında ve öğretmenlerin bu araçlarla ilgili yaklaşımlarının farklılaşmasında önemli bir etkiye sahip olduğu düşünülmektedir. Öğretmenlerin değerlendirme uygulamalarını/ araçlarını kullanma alışkanlıkları incelendiğinde, en sık kullanılan iki değerlendirme aracının problem çözme ve soru-cevap olduğu belirlenmiştir. En fazla önem verilen ve güven duyulan iki araç derse katılım ve gayret olmasına rağmen, bu uygulamanın kullanım düzeyi soru-cevap ve problem çözmeye göre düşük çıkmıştır. Öte yandan en az güvenilen ve önem verilen değerlendirme araçları olan portfolyo, günlük, kavram haritaları ve deney raporlarının kullanım düzeyleri diğer değerlendirme araçlarına göre çok düşük çıkmıştır. Dolayısıyla öğretmenlerin değerlendirme yaklaşımları ile değerlendirme uygulamalarını kullanma sıklıklarının ne kadar örtüştüğü konusunda kesin bir sonuç çıkarmak zordur. Ancak bu sonuç Adams ve Yang-Hsu (1998) tarafından yapılan araştırmanın sonuçlarıyla çelişmektedir. Adams ve Yang-Hsu (1998) öğretmenlerin en çok önem verdikleri ile en çok kullandıkları değerlendirme uygulamalarının paralel olduğunu bulmuştur.

Kullanım düzeyine ilişkin olarak öğretmenlerin beyanları ile gözlem sonuçları arasındaki ilişkiye bakıldığında ise tüm boyutlarda anlamlı bir ilişki bulunmuştur. Ancak üç boyuta ilişkin yapılan gözlem ve öğretmenlerin bu araçları kullanma alışkanlıklarının ortalamaları kıyaslandığında öğretmenlerin beyanlarının ortalamaları gözlem ortalamalarından anlamlı bir şekilde yüksek çıkmıştır.

Cinsiyete ve tecrübeye göre kullanılan değerlendirme araçlarının değişmediği bulunmuştur. Fakat Bol ve diğerleri (1998) tarafından yapılan çalışmada ise daha fazla tecrübeye sahip öğretmenlerin daha az tecrübeli öğretmenlere göre alternatif değerlendirme araçlarını daha sık kullandıkları belirlenmiştir.

Branşlara göre geleneksel, alternatif ve diğer değerlendirme uygulamalarının kullanım düzeyleri incelendiğinde sadece geleneksel uygulamalara ilişkin olarak matematik ve fizik öğretmenleri arasında anlamlı bir farklılık bulunmuştur. Fakat Bol ve diğerleri (1998) tarafından yapılan araştırmanın bulgularından biri matematik öğretmenlerinin geleneksel değerlendirme yöntemlerini fen öğretmenlerine göre daha az kullandıkları yönündedir. Lise düzeyinde olmasa bile McMillan, Myran ve Workman

(2002) da matematik ve dil bilgisi olmak üzere iki derste kullanılan değerlendirme araçları arasında fark olmadığını belirlemiştir.

Araştırmaya katılan öğretmen sayısının çok fazla olmaması ve gözlemcilerin öğretmenleri sadece 5 ders saati gibi kısa bir süre izlemeleri gibi sınırlılıklarına rağmen bu araştırma, ortaöğretim düzeyindeki matematik, fizik ve kimya öğretmenlerinin kullandıkları değerlendirme araçları/uygulamaları, bunlarla ilgili yaklaşımları ile tecrübenin, branşın ve cinsiyetin kullanılan araçlar üzerindeki etkisini anlama konusunda önemli ipuçları vermektedir.

### Sonuç

Çalışmada öğretmenlerin alternatif değerlendirme araçlarıyla ilgili bilgi düzeylerinin diğer araçlara/uygulamalara göre daha düşük olduğu bulunmuştur. Bu nedenle, öğretmen yetiştiren kurumların eğitim programlarında değerlendirmeye ilgili çeşitli araçlara ve uygulamalara geniş bir şekilde yer verilmelidir. Öğretmenlik yapanlara da çeşitli değerlendirme uygulamalarına yönelik hizmet içi eğitimler verilmelidir. Öğretmenler de hizmet içi eğitimlerin alternatif değerlendirme araçlarını kullanmalarını teşvik edeceğini belirtmiştir (Schmidt & Brosnan, 1996). Ayrıca öğretmenlerin bu araçların dayandığı yaklaşımları nasıl algıladıkları üzerinde durulmalıdır.

Daha sonra yapılacak araştırmalarda öğretmenlerle yüz yüze görüşmeler yapılarak onların çeşitli değerlendirme araçları/uygulamaları ile ilgili yaklaşımları ve kullanım düzeylerine ilişkin daha ayrıntılı veriler toplanabilir. Kullanılacak ölçeklerde daha az sayıda değerlendirme aracına odaklanarak öğretmenlerin bu araçlarla ilgili ne anladıkları daha net ortaya çıkarılabilir. Ayrıca kullanım düzeyini belirlemek için gözlem süresi mutlaka arttırılmalıdır. Bunların yanı sıra öğretmenlerin kullandıkları değerlendirme araçlarının örnekleri toplanarak bu araçların nitelikleri konusunda daha doğru tespitler yapılabilir.

### Kaynaklar

- Adams, T. L., & Yang-Hsu, J. W. (1998). Classroom assessment: Teachers' conceptions and practices in mathematics. *School Science and Mathematics*, 98(4), 174-180.
- Bol, L., Stephenson, P.L. O'Connell A. A., & Nunnery, J. A. (1998). Influence of experience, grade level, and subject area on teachers' assessment practices. *The Journal of Educational Research*, 91(6), 323-330.
- Mcmillan, J. H., Myran, S., & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research*, 95(4), 203-213.
- National Council of Teachers of Mathematics (NCTM). (1995). *Assessment standards for school mathematics*. Reston, VA.
- Reys, R. E., Suydam, M. N., Lindquist, M. M., & Smith, N. L. (1998). *Helping learn mathematics*. USA: Allyn & Bacon.
- Schmidt, M. E., & Brosnan, P. A. (1996). Mathematics assessment: practices and reporting methods. *School Science and Mathematics*, 96(1), 17-20.
- Street-Trepanier, M. L., McNair, S., & Donegan, M. M. (2001). The views of teachers on assessment: A comparison of lower and upper elementary teachers. *Journal of Research on Childhood Education*, 15(2), 234 – 242.

### Science and Mathematics Teachers' Approaches to Assessment Tools and Their Assessment Practices

#### Abstract

*Mathematics, physics and chemistry teachers' approaches and knowledge about assessment tools and their assessment practices are explored. The instrument developed by the researchers was given to 143 teachers. Data analyzed revealed that teachers' knowledge about alternative assessment tools such as portfolio, concept maps, journals and experiment reports was low, whereas their knowledge about achievement tests and questioning were high. It was identified that these four alternative assessment tools were ranked lowest in terms of frequency of usage, importance attached and reliance upon. In addition, it was observed that there is no agreement on teachers' approaches to these tools. The two most valued and important tool were found to be participation and effort. Reported assessment practices didn't differ with respect to gender and experience; however physics and mathematics teachers' use of traditional tools revealed to be significantly different.*

*Keywords:* Science teachers, mathematics teachers, traditional and alternative assessment practices