

BÖLÜM 7

SOLFEJ ÖĞRETİMİNDE DEŞİFRE BAŞARISININ DEĞERLENDİRİLMESİ

Özgür EROĞLU¹

GİRİŞ

Müziği okuma (müzik notasyonunu seslendirme) becerisi, müzik eğitiminin en öncelikli hedefleri arasındadır. Müziksel okuma, müzik öğrencisinin ilerlemesi ve nihayetinde tutarlı bir şekilde performans göstermesi için önemli bir araçtır. Müzisyenlerin bir eğitmeninden yardım almadan yeni bir repertuar öğrenebilmeleri için müziksel okuma becerisini kazanmaları gerekir (Sloboda, 1984: 222). Sloboda'ya (1978: 1) göre, nota okuma becerisi üst seviyede olan müzisyenler, müzik yaşamının hemen her alanında diğerlerine göre çok önemli bir avantaja sahiptirler. Gilman'a (2000) (akt. Jensen, 2016: 2) göre de, öğretme, eşlik etme veya performans gösterme gibi kariyer fırsatları için müziği hızlı ve etkin bir şekilde okumak esastır.

Müziksel okumanın alt becerilerinden birisi de vokal deşifredir. Müzisyenlerin, önceden çalışmadıkları ve aşına olmadıkları ezgilerin notalarını ilk görüşte seslendirmesi olarak tanımlanan bu beceri, müziksel okuma gelişimi için gereklidir (Jensen, 2016: 4). Bir performans becerisi olan vokal deşifrenin değerlendirilme süreci genellikle değerlendiricinin ruh hali ve kişisel önyargılarından belli oranlarda etkilenmektedir (Nelson, 1970: 14). Ayrıca, müzik eğitimi araştırmacıları arasında uzun süredir ilgi duyulan bir konu olan vokal deşifre üzerinde birçok çalışma yapılmış olmasına rağmen, tam olarak hangi becerilerin ne şekilde ölçülmesi gerektiği konusunda hala bir fikir birliğine varılamamıştır (Henry, 1999: 9).

Literatür derlemesi niteliğindeki bu bölümde, öğretimde değerlendirme ile ilgili genel bilgilerden sonra, müzikte performans değerlendirmeye ilişkin temel sorunlardan bahsedilmiş ve performans değerlendirme araçları geliştirmede kullanılacak geçerlilik ve güvenilirlik türleri açıklanmıştır. Ardından, performans değerlendirmenin belki de en fazla güçlük içeren boyutu olan vokal deşifre başarısının değerlendirilmesinde kullanılan ölçme yaklaşımları, geliştirilmiş ölçme araçlarının özellikleri ve bu alanda teknolojinin kullanımı konularına yer veril-

¹ Dr. Öğr. Üyesi, Atatürk Üniversitesi/Türk Musikisi Devlet Konservatuvarı/Türk Sanat Müziği AD.

miştir. Bölümün sonunda ise, elde edilen bilgilerden ulaşılan sonuçlar sunulmuş ve bu sonuçlara ilişkin önerilerde bulunulmuştur.

ÖĞRETİMDE DEĞERLENDİRMEYE GENEL BAKIŞ

Öğretim genellikle öğretmenlerin öğrettiği ve öğrencilerin öğrendiği basit bir süreç olarak kavramsallaştırılır. Bu bakış açısıyla öğretim, basit bir öğretme-öğrenme süreci olarak görülmektedir. Uygulamada ise, değerlendirmeyi öğretim sürecinin ayrılmaz bir parçası olarak görmek daha gerçekçidir. Aslında öğretmenlerin mesleki zamanlarının en az üçte birini değerlendirme ile ilgili etkinliklere ayırdıkları tahmin edilmektedir (Stiggins ve Conklin, 1992; akt. Reynolds, 2009: 2).

Değerlendirme, en temel anlamda, öğrencinin anladığına dair kanıt toplama sürecidir (Wiggins ve McTighe, 1998: 22). Bu kanıt, öğrenci ihtiyaçlarını teşhis etme, öğrencilere geri bildirim sağlama, öğrencileri gruplara ayırma, not verme ve öğretimi planlama gibi çeşitli amaçlar için kullanılabilir (Airasian, 2004: 32). Değerlendirme ayrıca, öğrenciler için güçlü bir motive edici olabilir. İlerlemenin somut kanıtı, önemli bir gurur kaynağı ve içsel motivasyonu teşvik etmek için etkili bir araçtır (Henry, 1999: 7).

Abeles (2010: 169) öğretimde değerlendirmenin iki genel işlevini şöyle tanımlar: Biçimlendirici değerlendirme, öğretimin değerlendirilmesini temsil eder ve öğrenmeyi ilerletmek için, devam eden öğrenci öğrenmesini izlemeyi amaçlar. Özetleyici değerlendirme ise, öğrenmenin bir değerlendirmesi olarak kullanılır. Kullanım amacı, öğrencinin ne öğrendiğini tespit etmek, genellikle not vermek ve yeni bir öğrenme düzeyine hazır olup olmadığını belirlemek gibi nihai yargılarda bulunmaktır.

Değerlendirmeyle ilgili literatürde, birbirinden farklı çok sayıda ilkedен bahsedilir. Pretorius (2007: 55) bu ilkeleri şu şekilde özetlemiştir:

- Değerlendirme, hevesli, açık fikirli, yetkin ve hesap verebilir öğretmenler tarafından uygulanır.
- Değerlendirme, açıkça tanımlanmış sonuçları ve değerlendirme kriterleri ile iyi yapılandırılmış bir müfredata dayalıdır.
- Değerlendirme, öğretimin sürekli ve ayrılmaz bir parçasını oluşturur.
- Değerlendirme geçerli, güvenilir ve adildir.
- Değerlendirme, olumlu bir öğrenme ortamı içinde yapılır.
- Değerlendirme öğrenci merkezlidir.
- Değerlendirme çeşitli ve esnekler.
- Değerlendirme öz değerlendirmeyi de içerir.
- Değerlendirme etkileşimi teşvik eder.
- Değerlendirme, akran değerlendirmesini de içerir.

- Değerlendirme orijinaldir.
- Değerlendirme, üst düzey bilişsel düşünmeyi destekler.
- Değerlendirme, kavramların entegre bir şekilde anlaşılmasını teşvik eder.
- Değerlendirme işlevseldir.
- Değerlendirme biçimlendirici olarak kullanılır.

Değerlendirme, hem öğretimi geliştiren hem de öğrenmeyi teşvik eden ilgili bilgileri sağlamalıdır. Diğer bir deyişle, öğretim, öğrenme ve değerlendirme arasında yakın bir karşılıklı ilişki olmalıdır. Öğretimin bu genişletilmiş kavramsallaştırması ile öğretim, öğrencilerin ne öğrendikleri, ne kadar iyi öğrendikleri, öğretimin ne kadar etkili olduğu ve hangi bilgi, kavram ve hedeflerin daha fazla dikkat gerektirdiği hakkında nesnel geri bildirim sağlayan değerlendirme ile bütünlük olarak ilişkilidir. Öğretme, bir öğretme-öğrenme süreci ile sınırlandırılmak yerine daha doğru bir şekilde öğretme-öğrenme-değerlendirme süreci olarak kavramsallaştırılmaktadır. Bu modelde değerlendirmenin amacı, öğretimde de olduğu gibi, öğrenci başarısını kolaylaştırmaktır. Eğitimin gerçek dünyasında, bir tür değerlendirme içermeyen etkili bir öğretimi hayal etmek zordur. Öğretmenler, öğrencilerin öğrenmesini değerlendirmede ne kadar iyi iş çıkarırsa, öğretimleri de o kadar iyi olacaktır (Reynolds, 2009: 2).

MÜZİKTE PERFORMANS DEĞERLENDİRME VE TEMEL SORUNLAR

Müziği yorumlamayı, yapılandırmayı ve fiziksel olarak yaratmayı içeren müzik performansı, psikolojik, fiziksel, akustik, sosyal ve sanatsal kararları içeren karmaşık bir dizi eylemi gerektirir (Palmer, 1997: 115-117). Müzik performanslarının değerlendirilmesi, kendine özgü bazı zorluklar içermektedir. Burada karşılaşılan önemli bir sorun, müzikal performans ölçümünün öznel doğasıdır (Abeles, 1973: 246). Değerlendirme için bazı yönergelerin sağlandığı durumlarda bile, değerlendiriciler, öznel görüş ve yargılardan etkilenmektedirler (Allen, 2005: 221).

Öğrenci performansının nüanslarını doğru bir şekilde tanımlayan kriterleri ifade etmek zordur. Performans değerlendirmede yaygın olarak kullanılan dereceli puanlama anahtarları genellikle öğrencilerin ve öğretmenlerin, performansın önemli özelliklerini tutarlı bir şekilde belirlemelerini sağlamak için fazla öznel ve belirsiz bir dil kullanır. Öğrenciler ve öğretmenler bu ifadeleri farklı yorumlayabilirler. Sonuç olarak, performans tutarsız bir şekilde değerlendirilecektir. Ek olarak, bu kriterler öğrencilere biçimlendirici değerlendirme yoluyla performanslarını nasıl geliştirecekleri konusunda çok az rehberlik sağlar. Bunun yanında, çoğu değerlendirme listesi, performansın ayrı bileşenlerinin kalitesini ölçer, ancak çeşitli müzik bileşenlerinin entegrasyonunu ve aralarındaki ilişkiyi dikkate almaz. Müzik performansı gibi karmaşık fenomenleri daha iyi anlamak için, temel unsurları

izole etmek yararlıdır. Bununla birlikte, karşılıklı ilişkiler ağlarından yapay olarak çıkarılmış ayrıık değişkenlerin incelenmesine dayalı sonuçlara götüren indirgemecilik tuzağından kaçınmak çok önemlidir (DeLuca ve Bolden, 2014: 71).

Dereceli puanlama anahtarı ölçütleri genellikle yalnızca teknik becerilere odaklanır. Çünkü bunlar daha kolay gözlemlenir ve ölçülür. Teknik beklentileri ölçmek önemli ve faydalı olsa da, yalnızca teknik beklentilere odaklanan değerlendirme listeleri, öğrencilerin müzisyenliğini ve performansın ifade edici yönlerine katılımını tam olarak teşvik etmez. Ayrıca, dereceli puanlama anahtarlarının biçimlendirici kullanımları öğrencilerin öz değerlendirme ve yansıtma kapasitesini desteklerken, çok az dereceli puanlama anahtarı bu becerileri kriterlere dâhil eder (DeLuca ve Bolden, 2014: 71).

Müzik dersinde performans değerlendirmeyi etkileyen müzik dışı faktörler şunlardır (McPherson & Thompson, 1998: 14):

- Performans gösterme sırası
- Sosyal yargılar
- Değerlendiricinin müzisyen hakkındaki ilk izlenimi
- Değerlendiricinin özellikleri

Müzik dışı faktörler, müzik performansının değerlendirilmesini olumsuz yönde etkiler ve bir performansın yargılanmasında öznel bakış açılarının rolünü vurgular. Bu faktörler, haksız önyargılar ürettiğinden, eğitimcilerin, değerlendiricilerin ve araştırmacıların bunları anlamak için çalışması önemlidir (McPherson ve Schubert, 2004: 73). Öğrencilerin müzikal performanslarını derecelendirmek, öğretmenlerin öznel yargılarına yol açabilse de, bu yargılar için koyulan kriterlerin açık bir şekilde tanımlanması, önyargıyı hafifletmeye yardımcı olur ve değerlendirmenin genel güvenilirliğini ve geçerliliğini artırır (Lopez, 2001: 33).

PERFORMANS DEĞERLENDİRMEDE GEÇERLİLİK VE GÜVENİLİRLİK

Değerlendiriciler tarafından yapılan performans değerlendirmeleri, puanlayıcıların yargılarının geçerliliği, güvenilirliği ve adaleti ile ilgili endişeler barındırır. Modern ölçme alanındaki gelişmeler, müzik alanında performans değerlendirmelerinin adillliğini ve hakkaniyetini geliştirmek için yeni yaklaşımlar sunabilir. Güvenilir ve geçerli derecelendirmeler üreten bir değerlendirme sistemi geliştirmek için bazı yöntemler kullanılabilir (Wesolowski, 2015: 149).

Performans değerlendirme amacıyla geliştirilen ölçekler için uygulanacak geçerlilik ve güvenilirlik yöntemleri, keşfedici faktör analizi, değerlendirici içi (intra-rater) güvenilirlik, değerlendiriciler arası (inter-rater) güvenilirlik, yüzey geçerliliği, içerik geçerliliği ve tahmin geçerliliği olarak sıralanabilir.

Keşfedici faktör analizinde ölçüm maddelerinin belirli faktörleri veya kavramsal yapıları temsil etme güvenilirliği araştırılır. Bir testteki maddelerin hangi temel bileşenlere işaret ettiğini belirlemek, test maddelerini etkileyen arka plandaki gizli yapıyı veya gizli değişkenleri ortaya çıkarmak veya faktörler ve değişkenler arasındaki ilişkilerin niteliğini saptamak için kullanılır. Değerlendirici içi güvenilirlik, değerlendirmeyi yapan araştırmacının, öğretmen veya gözlemcinin, aynı performans için farklı zamanlarda ne ölçüde tutarlı puanlama yaptığını belirlemeye yöneliktir. Değerlendiriciler arası güvenilirlik ise, öğrencinin performansının birden fazla değerlendirici tarafından kriterler dikkate alınarak değerlendirilmesi ve bu değerlendirme sonucunda verilen puanların tutarlı çıkmasıdır (Şencan, 2005).

Yüzey geçerliliği ilk aşamada araştırmacının kendisi tarafından yapılır. Testin veya ölçeğin maddeleri ile araştırılan kavramsal yapı arasında anlamlı bir bağ kurulabiliyorsa test yüzeysel geçerliliğe sahiptir. Sonraki aşamada pilot uygulama sırasında katılımcılardan gelen görüşler de değerlendirilir. İçerik geçerliliği, örneklem olarak belirlenen test veya ölçek maddelerinin belirli bir amaca yönelik olarak kavramsal ana kütle temsil etme derecesidir. İçerik geçerliliğinde ölçüm aracının ölçmek istediği yapıyı ölçüp ölçmediği, uzman görüşleriyle belirlenir. Tahmin geçerliliği, bir ölçekten/testten yüksek puan alan bir öğrencinin elde ettiği sonuçların, o öğrencinin daha sonraki ilgili başarı düzeylerinde de görülmesidir. Tahmin geçerliliği için test/ölçek öğrencilere uygulandıktan sonra elde edilen sonuçlar, uygun bir sürenin (6 ay ila 5 yıl) sonunda sorgulanır. Bu süre içinde kişilerin ilgili başarılarına ilişkin veriler toplanır. Bu verilerin ortalaması alınarak test puanlarıyla arasındaki korelasyon hesaplanır ve ölçeğin tahmin geçerliliği saptanır (Şencan, 2005).

VOKAL DEŞİFRE BAŞARISI DEĞERLENDİRME YAKLAŞIMLARI

Müzik okuma becerilerini değerlendirmenin yaygın bir yöntemi, vokal deşifre olarak adlandırılan, öğrencilerin önceden çalışmadıkları ve aşına olmadıkları ezgilerin notalarını ilk görüşte seslendirme becerilerini test etmektir. Vokal deşifre, hem işitsel algıyı hem de motor beceriyi kullanan bir süreçtir. Deşifre yapan kişi uygun perdeyi sadece algılamakla kalmamalı, aynı zamanda da seslendirebilme-lidir (Lucas, 1991: 2).

Vokal deşifre başarısının değerlendirilmesiyle ilgili literatürde yar alan farklı değerlendirme yaklaşımları şöyle özetlenebilir:

- Nota-nota değerlendirme – Ölçü-ölçü değerlendirme
- Perde ve ritim becerilerinin eş zamanlı değerlendirilmesi – Yalnızca perde becerisinin değerlendirilmesi

- Kısa ezgi parçalarının kullanılması – Uzun ezgilerin kullanılması
- Sadece temel unsurların değerlendirilmesi – Estetik özelliklerin de değerlendirilmesi

Vokal deşifre değerlendirmesinde puanlama yapılırken, tipik olarak iki değerlendirme stratejisinden biri kullanılır: her doğru perde veya ritim için puan verilen nota-nota değerlendirme veya hiçbir hatanın olmadığı her bir ölçü için puan verilen ölçü-ölçü değerlendirme (Henry, 2015: 58). Bu alanda geliştirilmiş olan birçok test arasındaki benzerliklerden birisi de, perde değerlendirmesinin ritim değerlendirmesinden ayrılmasıdır. Bazı testler, perde becerisini ritmik beceriyi dikkate almadan değerlendirirken, bazıları ise test materyallerinden tüm ritmik öğeleri çıkarmıştır (Henry, 2001: 22).

Geliştirilmiş olan testlerin bazılarında materyal olarak kısa ezgi parçaları kullanılırken, bir kısmında ise uzun ezgiler kullanılmaktadır. Nelson (1970), üniversite öğrencileriyle yaptığı bir deneyde, kısa maddeli ve uzun maddeli vokal deşifre testlerinin etkinliğini karşılaştırmıştır. Nelson, proje için her biri iki cümle uzunluğundaki sekiz melodiden oluşan uzun maddeli bir deşifre aracı hazırlamış, testin güvenilirliğini ve geçerliliğini hesaplamak için iki pilot çalışma yapmıştır. Sonuçta, uzun maddeli ve kısa maddeli vokal deşifre testlerinin güvenilirlik ve geçerliliğinin kayda değer ölçüde değişmediği anlaşılmıştır.

Birçok müzik eğitimcisi, deşifrede öğretim ve değerlendirme için tonal kalıpların kullanılmasını savunur. Edgar Gordon (1958), müziksel okuma sürecindeki önemli bir unsuru “zaten öğrenilmiş olan ezgi kalıplarının bir görüntüsünü elde etmek” olarak tanımlar (Henry, 1999: 14). Şema geliştirme ile bir deşifre testinin oluşturulması arasındaki kritik ilişki, müzikal bir “bütün” içinde yer alan tüm öğelerin birbiriyle ilişkili olduğunun anlaşılmasıdır. Krumhansl’a göre (1990; akt. Scott, 1996: 6) “Algısal deneyimimiz, tek tek müzik olaylarının duyusal kaydının ötesine geçer. Ses öğeleri bağlam içinde duyulur, perde ve ölçü birimi içinde düzenlenir ve bu bağlam içindeki işlevleri açısından anlaşılır.” Yine de, müzikal unsurlar arasındaki güçlü ilişkileri destekleyen çok sayıda kanıtla rağmen, çoğu deşifre değerlendirmeleri yalnızca melodik parçalar, yapay olarak düzenlenmiş yapılar ve kısa ritmik kalıplar kullanarak performans doğruluğunu ölçmüştür. Gabriëlsson’a (1979; akt. Scott, 1996: 7) göre, «Genel olarak ritmin melodi, armoni ve tını gibi müziğin diğer öğelerinden ayıramayacağını akılda tutmalıyız. Müzik deneyiminde herhangi bir tek yönü izole etmek her zaman tehlikelidir.» Müzik öğelerinin bir test aracı içinde sunulduğu bağlam belirlenirken «müziksel bütünü» önemi göz önünde bulundurulmalıdır.

Öğrencilerin performanslarını geliştirmelerine yardımcı olabilecek ek bir biçimlendirici değerlendirme yöntemi olarak öz değerlendirmeye de yer verilmesi

önerilir. Ancak veriler, öğrencilerin müzik bilgilerini kendi kendilerine değerlendirmede güvenilir olduklarını, buna karşın, müzik becerilerini değerlendirirken genellikle daha az güvenilir olduklarını göstermektedir (Darrow, Johnson, Miller ve Williamson, 2002; akt. Darrow ve Marsh 2006: 22). Lohmann ve Marsh (1998), deşifre şarkı söylemenin müzikal becerisine ilişkin öğrencilerin kendi beyanlarının güvenilirliğini incelemeye çalışmışlardır. Araştırmacılar yaptıkları çalışmada, deşifre becerisiyle ilgili en doğru tahminlere sahip öğrencilerin, en az müzik eğitimi almış ve en düşük deşifre testi puanlarına sahip öğrenciler olduğunu bulmuşlardır. Bu öğrenciler vokal deşifrede başarılı olamayacaklarının tamamen farkındaydılar. Deşifre testi puanları yüksek olan öğrencilerin genel eğilimi ise, tahminlerinde abartma yönünde olmuştur (akt. Darrow ve Marsh, 2006: 23).

Solfej sınıfında vokal deşifrenin bireysel olarak değerlendirilmesi kendine özgü bazı zorluklar içerir (Henry, 2001: 22):

- Sınıf ortamında bireysel değerlendirme için çok fırsat olmayabilir
- Son derece zaman alıcı olabilir
- Puanlama süreci öznel olabilir
- Vokal deşifrenin değerlendirilmesi için kaliteli değerlendirme araçları bulunmamaktadır.

Bu alanda yaygın olarak kabul edilen bir değerlendirme aracının olmaması nedeniyle, öğretmenler dersleri için özel olarak kendi değerlendirme araçlarını oluşturmak zorunda kalırlar (Henry, 2003: 24). Barnicle (2001) (akt. Goss, 2010: 21), değerlendirmelerin öğrencilerin deşifre becerilerinin doğru bir ölçümünü sağlaması için, öğretmen tarafından oluşturulan alıştırmaların performanslarının ses kayıtlarının alınmasını önermektedir. Ürünler daha sonra performans değerlendirme rubrikleri de dâhil olmak üzere çeşitli yöntemler kullanılarak puanlanabilir. Öğretmen, bu alıştırmalarla ilgili öğrenci performanslarının bir arşivini tutarak, öğrencinin deşifre becerilerinin belirli bir süre boyunca gelişiminin grafiğini çıkarabilir.

GELİŞTİRİLMİŞ ÖLÇME ARAÇLARININ ÖZELLİKLERİ

Enstrümantal deşifre becerilerini ölçümüyle ilgili çalışan araştırmacılar, Watkins-Farnum Performans Ölçeğini (1962) yıllardır kullanmaktadırlar. Vokal deşifre ile ilgilenen araştırmacılar tarafından ise, evrensel olarak kabul edilen veya kullanılan eşdeğer bir test yoktur (Henry, 1999: 9).

Bu alandaki ilk girişimlerden biri, 1924 yılında Earl Hillbrand tarafından yapılan bir çalışmadır. Çalışmasının amacı “vokal deşifre becerisinin ölçülmesi için bir ölçek geliştirmek” idi. Test, ilköğretim müzik kitaplarında bulunan repertuar

kullanılarak oluşturulmuştur. On iki alıştırmaya için hata birimi tek notadır ve hatalar şu şekilde tanımlanmıştır: tam hatalar, kısmi hatalar, atlanan notalar, fazladan eklenen notalar, tekrarlar ve tereddütler. Hillbrand, çalışması için denek olarak 1487 dördüncü, beşinci ve altıncı sınıf öğrencisini seçmiş, denekler araştırmacı tarafından bireysel olarak test edilmiş ve puanlama öğrencilerin deşifresi sırasında yapılmıştır. Temel istatistiksel hesaplama, bireysel olarak öğrenciler için ve tüm şarkılardaki çeşitli düzeyler için hata yüzdesidir (Scofield, 1980: 45).

1925'te ortaya çıkan Mosher Bireysel Söyleme Testi, öğrenci tarafından tekrarlanmak üzere zorluk sırasına göre düzenlenmiş on iki alıştırmadan oluşur. Skor, tonal ve ritmik olarak doğru seslendirilen ölçülerin toplam sayısıdır (Graves, 1980: 11).

Otterstein-Mosher Vokal Deşifre Testi (1932), öğrencinin ilk görüşte müzik okuma başarısını belirlemek amacıyla yapılan bireysel bir testtir. Test, hem melodik hem de ritmik zorluk düzeyi gittikçe artan yirmi sekiz alıştırmadan oluşur. İlk alıştırmalar temel olarak dizi seslerinde, yanaşık hareket içeren majör tonlarda melodilerden oluşurken, daha sonraki alıştırmalar ise genellikle tonikten farklı bir notayla başlayan, sık atlamalar ve dizi dışı sesler içeren minör melodilerdir. Test sadece perde ve ritmi ölçmek için tasarlanmıştır. Ton kalitesi, dinamik varyasyonlar ve artikülasyon değişiklikleri testin bir parçası değildir. Sınav görevlisi tonaliteyi ve başlangıç perdesini duyurur ve öğrenci egzersizi seslendirir. Hata birimi tam ölçüdür ve her bir ölçü başına, biri perde ve diğeri ritim için iki ayrı puan tanımlanmıştır. Toplam puanı 448 olan test, 224 ölçüden oluşmaktadır (Scofield, 1980: 46).

John Cooper (1965), üniversite müzik öğrencileri için aşamalı olarak düzenlenmiş bir başarı testi geliştirmiştir. Test için performanslar kaydedilmekte ve daha sonra bir derecelendirme ölçeğini doldurmak için sınav uygulayıcısı tarafından birkaç kez incelenmekteydi. Testte melodik (aralığa dayalı), ritmik ve birleşik puanlar verilmekteydi. Cooper, testi oluştururken deşifre testine dâhil edilmesi gereken faktörlere dikkat etmiştir. Geleneksel Batı müziğinin tarzında olduğunu belirttiği melodiler, test için özel olarak yazılmıştır. Testte, eksik beşli dışındaki geleneksel aralıkların yanı sıra artık ikili aralığı da kullanılmaktaydı. Ritimler, literatürde normalde karşılaşılabilecek olan ve giderek daha zorlaşan figürleri içeriyordu. Cooper, notayı hata birimi olarak kullanmayı seçmiştir. Test, birinci dinlemede ritmin, ikinci dinlemede ise sesin puanlanması amacıyla teybe kaydediliyordu. Araştırmacı, puanlama süresinin denek başına kırk beş dakikadan iki saate kadar olduğunu belirlemiştir. Cooper, elde ettiği verilerden, ritmin ve melodinin ayrı unsurlar olarak düşünülmemeyeceğine, bir bütün olarak düşünülmesi gerektiğine karar vermiştir. Ayrıca, ritmik zorlukların tonal zorluklardan daha az

sorun teşkil ettiğini belirlemiştir. Zorluk düzeyinin, deşifre başarısının bir faktörü olarak, mutlak bir ritim veya melodi niteliği olarak tanımlanamayacağı, büyük ölçüde bağlamla ilişkili olduğu sonuçlarına varmıştır. Cooper, puanlama için gereken zaman miktarı nedeniyle bu testi kullanışsız bulmuştur (Graves, 1980; Scofield, 1980).

Thostenson'ın CSS76 Criterion Sight-Singing Testi (1967) şu dört bölümü içermektedir: 1) on iki temel aralığı içeren bir aralık testi; 2) yirmi dört adet ritimsiz perde dizisi; 3) yirmi adet perdesiz ritim dizisi ve 4) yirmi adet iki ölçülük hem perde hem de ritim içeren melodi. Bu test yaklaşık otuz dakika içinde uygulanabiliyordu. Test bireysel olarak uygulanıyor ve sınav uygulayıcısı tarafından derecelendiriliyordu. Thostenson, testi lisans ve lisansüstü müzik öğrencileri üzerinde uygulamış ve 0.95'lik bir güvenilirlik bildirmiştir (Graves, 1980; Scofield, 1980).

Belwin-Mills Vokal Başarı Testi, Belwin Yayıncılık için 1971'de Bowles tarafından oluşturuldu. Artık ticari olarak mevcut olmamasına rağmen, yayınlanmış bir testtir. Test, ilkokuldan üniversiteye kadar olan öğrenciler için vokal deşifre başarısını ölçmek üzere tasarlanmıştır, ancak yazar enstrümantal deşifre için de kullanılabilirliğini önermektedir. Test için, deneklerden bir egzersizi belirli bir tempoda söylemeleri istenir. Puanlama birimi tam ölçüdür; bu nedenle bir ölçüde yapılan herhangi bir hata o ölçünün tamamını etkiler ve hatasız ölçülerden toplam puan hesaplanır (Jensen, 2016: 9).

Scofield (1980), öğrencilerin vokal deşifre başarısını ölçmek için güvenilir ve geçerli bir test geliştirmiştir. Testin pratikliği ve kullanılabilirliği özel bir önem taşıyordu. Çalışmanın bulgularından iki önemli sonuç çıkarılmıştır. İlk olarak, ritimsiz bir test kullanarak deşifre performansını ölçme yöntemi, çalışılan örneklem için geçerlidir. İkincisi, bu test, diğerleri gibi yalnızca güvenilir ve geçerli olmakla kalmayıp aynı zamanda oldukça kullanışlı olan bir deşifre testidir (akt. Farenga, 2013: 31).

Scott (1996), lise sopranoaları için Ulusal Sanat Eğitimi Standartlarına dayalı bütünsel bir deşifre testi geliştirmek için, öncelikle standart koro repertuarından dört ayrı zorluk seviyesini temsil eden eserleri araştırmıştır. Uzmanlardan oluşan bir komite, seçilen literatürün zorluk seviyeleri hakkında fikir birliğine vardıldıktan sonra, Scott, birden dörde kadar olan zorluk seviyelerini eşit olarak temsil eden 12 eserden 16 ölçülük alıntılar seçmiştir. Standart birim olarak ölçüyü kullanarak testi puanlamıştır. Her bir ölçüde, perde için iki puan ve ritim için iki puan verilmiş, ölçünün bir kısmı doğruysa bir puan verilmiştir. Pilot çalışma, testin çok uzun olduğunu ortaya koyduğundan bazı düzenlemeler yapılmış ve uygulama süresi ortalama 12 dakikaya, puanlama süresi de yaklaşık 7 dakikaya inmiştir. Belli bir seviyede performans göstermede açıkça başarısız olan deneklerden daha

yüksek bir zorluk seviyesine devam etmeleri istenmemiştir. Araştırmada, yalnızca bir öğrenci Dördüncü Düzey'i başarıyla tamamlayarak ileri düzey okuyucular için önerilen standardı elde etmiştir. Scott, ölçümünde elde edilen düşük puanların, koro sınıfındaki kaliteli eğitim ve öğrenim eksikliğinden kaynaklandığını öne sürmüştür.

Eğitsel yararlılık açısından, çoğu testin birincil eksikliği, deşifre becerisini genel olarak değerlendirmeleridir. Genel bir doğruluk yüzdesi sağlanarak bir kıyaslama yapılabilse de, çok az tanısal bilgi elde edilmiştir. Bu testlerin hiçbiri, öğrencilerin hangi becerileri geliştirdiği veya hangilerini geliştiremediği hakkında bilgi sağlamamıştır (Henry, 2001: 23).

Henry (2001: 23), bir bireysel deşifre değerlendirme aracı olan Vokal Deşifre Envanterini (VSRI) şu gereksinimleri karşılayacak şekilde geliştirmiştir:

- Sesli deşifrede başarılı olmak için gerekli olan tüm becerileri içermeli
- Zaman ve içerik açısından uygulanabilir olmalı
- Öğretmenler ve öğrenciler için biçimlendirici ve özetleyici değerlendirme verileri sağlamalı
- Güvenilir ve geçerli olmalı, standartlaştırılmış veriler sağlamalı

Henry (2001: 24) öncelikle vokal deşifre için gerekli olan, 28 perde becerisini tanımlamıştır. Perde becerileri, çıkıcı ve inici yanaşık hareket, akor sesleri arasında küçük ve büyük atlamalar, kadans kalıpları, modülasyon kalıpları ve kromatik kalıpları içermektedir. 28 becerinin her biri, ritmik bir çerçeve içeren melodiler içinde yer almaktadır. Ritmik figürler çoğunlukla dörtlük notalar ve az sayıda sekizlik notalardan oluşmaktadır, ancak ritim doğruluğu değerlendirilmemektedir. Veri analizinin sonuçları, 28 perde becerisi için zorluk seviyelerinin belirlenmesini sağlamıştır. Deneğin puanı, doğru şekilde gerçekleştirilen toplam perde becerilerini yansıtmaktadır. Ek olarak, testten alınan puan, deneğin başarılı bir şekilde sergilediği veya gerçekleştiremediği bireysel becerileri belirleyerek, testi tanısal bilgi sağlayan bir teşhis aracı olarak da faydalı hale getirmektedir.

DEŞİFRE BAŞARISINI DEĞERLENDİRMEDE TEKNOLOJİNİN KULLANIMI

Güvenilir bir değerlendirme sağlamak için sistematik ve iyi tasarlanmış bir ölçme çok önemlidir. Çok sayıda müzik eğitimcisi, öğrencileri değerlendirmelerine yardımcı olacak çeşitli teknolojiler kullanır. Russell ve Austin (2010: 46) ortaöğretim müzik eğitimcilerinin değerlendirme uygulamaları üzerine bir araştırma yapmışlardır. Veriler, ortaöğretim öğretmenlerinin %32'sinin öğrenci performansını değerlendirmek için ses kayıtları kullandığını göstermiştir. Benzer şekilde, La-

Cognata (2010: 82) öğretim elemanlarının %33'ünün bir değerlendirme stratejisi olarak öğrencilere performanslarını kaydetme görevi verdiklerini belirtmiştir. La-Cognata ayrıca bazı okul orkestrası yöneticilerinin SmartMusic programı (%13.1) gibi belirli müzik yazılımlarını kullandıklarına da dikkat çekmiştir

Müzik testleri ve ölçümleri alanındaki bilim adamları, uzun zamandır müzik performansını ölçmenin daha nesnel bir aracını geliştirmekle ilgilenmektedir. On dokuzuncu yüzyılın sonlarında Alman müzisyenler, seslendirilen perdeleri sabit bir org perdesiyle karşılaştırarak, aralarındaki farkı ölçmek için çaba harcamışlardır. Bununla birlikte, yirminci yüzyılın başlarında, müzikal testler ve ölçümlerde öncü olan Carl E. Seashore, perdeyi doğru bir şekilde ölçecek bir elektronik cihaz tasarlamak için çalışmalara başlamıştır. İlk olarak 1914'te tanımlanan Seashore tonoskopunun daha da geliştirilmesi, çalınan veya söylenen seslerin bir grafiğini ekranda üretme kabiliyetini sağlamıştır. Harold Seashore, Max Schoen ve Arnold Small gibi diğer müzik psikologları da, profesyonel şarkıcıların ve kemancıların perde reproduksiyonlarını incelemek için bu yöntemi kullanmışlar, özellikle ton kalitesi, vibrato, perde sapması ve atak gibi sanatsal nüanslarla ilgilenmişlerdir (Graves, 1980: 16).

Tını kalitesi, perde entonasyonu vb. iyileştirmenin bir aracı olarak elektronik geri bildirim kullanılması içeren bir çalışma, 1969'da Heller tarafından yapılmıştır. İlk girişimleri, Seashore'un yönteminin güncellenmiş bir versiyonuna yöneliktir. Bu yöntem için ihtiyaç duyulan ekipmanın erişilemezliği ve kullanımının pratik olmaması nedeniyle, daha modern elektronik cihazlar kullanma ihtiyacı duymuş, ancak bu yöntem de sahip olduğu dezavantajları nedeniyle yaygınlaşmamıştır (Graves, 1980: 17).

Müzik performansını analiz etmede bilgisayar kullanmak için de çalışmalar yapılmıştır. Enstrümantal müzikte bilgisayar destekli öğretimin pratikliği üzerine bir araştırma yapan Peters, bir arayüz tasarlamıştır. Bilgisayar, bir grup üniversite trompet öğrencisinin performanslarındaki perde ve ritmi doğru bir şekilde ölçebilmiştir. Ancak Peters, çalışmada kullanılan perde toleranslarının pratik amaçlar için fazla katı olduğunu, bu yüzden hiçbir deneğin egzersizleri hatasız gerçekleştirmediğini belirtmiştir (Graves, 1980: 18).

Enstrümantal öğretimde bilgisayarın kullanımı, vokal öğretimi geride bırakmıştır. Peters ilk olarak 1974'te çalgı öğrencilerine yönelik bilgisayar destekli öğretim olanaklarını araştırmıştır. Özellikle, web tabanlı SmartMusic programının evrimi enstrümantal değerlendirme için bilgisayarların kullanımında devrim yaratmıştır. Esasen 1994'te piyasaya sürülen Vivace adındaki bilgisayarlı bir eşlik takip programından geliştirilen sonraki formu, hem repertuar performansı hem

de deşifre çalma için değerlendirmeye izin vermektedir. En son sürümü artık vokal deşifre için de benzer yetenekler içermektedir. Bu teknolojik gelişmeler aynı zamanda sınıf eğitimi ve bireysel uygulama için yeni seçenekler de getirmektedir (Henry, 2015: 59).

Son yıllarda geliştirilen SmartMusic benzeri başka birçok yazılım, bir bilgisayar aracılığıyla perde doğruluğunu ve ritmik doğruluğu değerlendirebilme kabiliyetine sahiptir. Bilgisayara bağlı yerleşik bir mikrofon veya harici mikrofon, sesleri anında kaydedebilir. Bu işlev, yalnızca öğretmenlerin öğrencilere geri bildirim vermesine izin vermekle kalmaz, aynı zamanda öğrencilere otomatik geri bildirim ve düzeltmeler sağlar. Bu programlar ayrıca standart bir test veya dönem sınavı düzenlemek amacıyla kullanılabilir. Öğretmenlerin, büyük bir grup içinde her bir öğrenciyi değerlendirme göreviyle başa çıkmalarına da yardımcı olur.

Bununla birlikte, bilgisayar aracılığıyla öğretime veya değerlendirmeye alışkın olmayan öğrenciler, prosedürlerden rahatsız olabilir veya bilgisayar tabanlı değerlendirmenin doğruluğuna şüpheyle bakabilir. İnsan değerlendiriciler tarafından geleneksel değerlendirme uygulamalarının bilgisayar güdümlü süreçlere dönüştürülmesi, özellikle not verme veya seçmeler için kullanıldığında puanların güvenilirliğinin sorgulanmasına neden olabilir veya performans kaygısında artışa neden olabilir (Henry, 2015: 63).

Buna karşın, bilgisayar tabanlı değerlendirme, kişisel veya görsel önyargı ve fiziksel yorgunluk sorunlarının olmaması gibi kendine özgü güçlü yönleri nedeniyle insan hatalarının azaltılmasına yardımcı olabilir. Dowsing, Long ve Craven (2000; akt. Shih, 2018: 23), Bilgi Teknolojileri beceri sınavlarında insan denetçiler ile bilgisayarlı değerlendirme arasındaki farklara ilişkin bir çalışma yürütmüştür. Sonuçlar, bilgisayarlı değerlendirme sisteminin insan kontrolörlere kıyasla gelişmiş güvenilirliğe ve doğruluğa sahip olduğunu gösterdi. Teknolojinin kullanımı, yalnızca biçimlendirici ve özetleyici değerlendirmelerin tutarlılığını sağlamaya yardımcı olmakla kalmaz, aynı zamanda puanların tarafsız şekilde verilmesini sağlayarak öğrenciler, öğretmenler ve yöneticiler arasında hesap verebilirliğin artırılmasına yardımcı olur.

SONUÇ VE ÖNERİLER

Vokal deşifre başarısının etkili şekilde değerlendirilmesi için eğitimcilerin kullanabileceği geçerli, güvenilir ve kullanışlı ölçme araçları bulunmamaktadır. Bu özellikleri barındıran araçların geliştirilmesi için yapılacak çalışmalarda öncelikle değerlendirme kriterleri, ilgili literatür ve uzman görüşleri aracılığıyla, gerekli olan tüm becerileri içerecek şekilde dikkatle belirlenmeli, geçerlilik ve güveni-

lirlik için uygun yöntemler titizlikle uygulanmalıdır. Bunun yanında, testlerin özellikle uygulama ve puanlama süresi açısından kullanışlı olmasına ve özetleyici değerlendirmenin yanı sıra biçimlendirici değerlendirme verileri de sağlamasına dikkat edilmelidir. Ayrıca, hangi değerlendirme yaklaşımlarının gerekli özellikleri daha iyi sağladığının belirlenmesi açısından, farklı yaklaşımların karşılaştırıldığı çalışmalar alana üst düzeyde katkı sağlayabilir.

Değerlendiricinin öznel görüş ve yargıları ile öğrencinin performans gösterme sırası gibi müzik dışı faktörler, müzik performanslarının değerlendirilmesi sürecini olumsuz şekilde etkilemektedir. Bilgisayar tabanlı değerlendirme, müzik dışı faktörlerin etkisini ortadan kaldırmayı sağlayacak bir yöntem olabilir. Bu tür yazılımların sayılarının artması ve kullanımlarının yaygınlaşması, vokal deşifre için daha adil bir değerlendirme sağlayacaktır. Özellikle akıllı telefonlar için geliştirilecek uygulamalar, bu teknolojinin öğrenciler tarafından da etkili şekilde kullanılarak, solfej öğretiminde hem eğitimcilere hem de öğrencilere düzenli olarak geri bildirim sağlanmasına imkân verebilir. Solfej eğitimcileri ve araştırmacıları, diğer ilgili alanlarla ortak projeler yürüterek bu yazılımların daha da geliştirilmesi için çaba harcamalıdır.

KAYNAKLAR

- Abeles, H. F. (1973). Development and validation of a clarinet performance adjudication scale. *Journal of Research in Music Education*, 21(3), 246–255. Retrieved from <http://www.jstor.org/stable/3345094>
- Abeles, H. F. (2010). Assessing music learning. In H. Abeles & L. Custodero (Eds.), *Critical issues in music education: From theory to practice* (pp. 167-193). New York, NY: Oxford University Press.
- Airasian, P. W. (2004). *Classroom assessment: Concepts and applications*. New York, NY: McGraw-Hill.
- Allen, J. D. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 78(5), 218-223. Retrieved from <http://www.jstor.org/stable/30189912>
- Darrow, A.A., & Marsh, K. (2006). Examining the validity of self-report: middle-level singers' ability to predict and assess their sight-singing skills. *International Journal of Music Education*, 24(1), 21–29. <https://doi.org/10.1177/0255761406063101>
- DeLuca, C. ve Bolden, B. (2014). Music Performance Assessment: Exploring Three Approaches for Quality Rubric Construction. *Music Educators Journal*, 101(1), 70–76. <https://doi.org/10.1177/0027432114540336>
- Farenga, J. (2013). Arizona high school choral educators' attitudes toward the teaching of group sight singing and preferences for instructional practices (Order No. 3605207). Available from ProQuest Dissertations & Theses Global. (1477861721).
- Goss, D. A. (2010). Sight-singing assessment: A study of current beliefs and practices of Georgia middle and high school choral directors (Order No. 3398032). Available from ProQuest Dissertations & Theses Global. (305244889).
- Graves, D. L. (1980). The Development of an Objective Sight Singing Achievement Test Employing Electronic Measurement Apparatus (Order No. 8107914). Available from ProQuest Dissertations & Theses Global. (303008678).

- Henry, M. (1999). The development of an individual vocal sight-reading inventory (Doctoral dissertation, University of Minnesota, Minneapolis, 1999). Dissertation Abstracts International, 60, 683A.
- Henry, M. (2001). The Development of a Vocal Sight-Reading Inventory. *Bulletin of the Council for Research in Music Education*, 150, 21–35. <http://www.jstor.org/stable/40319097>
- Henry, M. (2003). A comparison of testing formats for vocal sight reading. *Texas Music Education Research*, 24-31.
- Henry, M. (2015). Vocal Sight-Reading Assessment: Technological Advances, Student Perceptions, and Instructional Implications. *Update: Applications of Research in Music Education*, 33(2), 58–64. <https://doi.org/10.1177/8755123314547908>
- Jensen, M. C. (2016). Measuring Music Reading: A Guide to Assessment Methods (Master's thesis). University of Ottawa.
- LaCognata, J. P. (2010). Current student assessment practices of high school band directors (Doctoral dissertation). Available from ProQuest Dissertations & Theses Global. (UMI No. 3436343)
- Lopez, C. (2001). Assessing elementary improvisation. In MENC (Ed.), *Spotlight on assessment in music education* (pp. 32-34). Reston, VA: Music Educators National Conference.
- Lucas, K. V. (1991). The effect of contextual condition on the sightsinging achievement of middle school choral music students. Dissertation Abstracts International.
- McPherson, G. E. ve Schubert, E. (2004). Measuring performance enhancement in music. Musical excellence: Strategies and techniques to enhance performance. In A. Williamon (Ed.), *Enhancing Musical Performance* (pp. 61-82). Oxford, UK: Oxford University Press.
- McPherson, G. E. ve Thompson, W. F. (1998). Assessing music performance: Issues and influences. *Research Studies in Music Education*, 10(1), 12-24.
- Nelson, J. (1970). A comparison of two methods of measuring achievement in sight singing. (Doctoral dissertation). Retrieved from Proquest Dissertations and Theses. (UMI 7015629).
- Palmer, C. (1997). Music performance. *Annual Review of Psychology*, 48, 115-138.
- Pretorius, J. (2007). A Critical Evaluation of Assessment Practices in Music Literacy Programmes for Young Adults (Master's thesis). North-West University.
- Reynolds, C. R. (2009) *Measurement and Assessment in Education*. N.J.: Pearson.
- Russell, J. A. ve Austin, J. R. (2010). Assessment Practices of Secondary Music Teachers. *Journal of Research in Music Education*, 58(1), 37-54.
- Scofield, W. R. (1980). The Construction and Validation of a Method for The Measurement of the Sight-singing Abilities of High School and College Students (Order No. 8020739). Available from ProQuest Dissertations & Theses Global. (288220319).
- Scott, T. B. (1996). The construction of a holistic, criterion-referenced sight-singing test for high school sopranos based on the voluntary national standards for music education (Order No. 9712431). Available from ProQuest Dissertations & Theses Global. (304302233).
- Shih, Y. J. (2018) Evaluation of Music Performance: Computerized Assessment Versus Human Judges (Doctoral dissertation). University of Hawaii.
- Sloboda, J. A. (1978). The psychology of music reading. *Psychology of Music*, 6, 3-20.
- Sloboda, J. A. (1984). Experimental studies of musical reading: A review. *Music Perception*, 2, 222–236.
- Şencan, H. (2005). Sosyal ve davranışsal ölçümlerde güvenilirlik ve geçerlilik. Ankara: Seçkin Yayıncılık.
- Wesolowski, B. C., Wind, A., & Engelhard, G. (2015). Rater fairness in music performance assessment: Evaluating model-data fit and differential rater functioning. *Musicae Scientiae*, 19(2), 147–170. <https://doi.org/10.1177/1029864915589014>
- Wiggins, G., & McTighe, J. (1998). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.