



Tanılayıcı Dallanmış Ağacın Psikometrik Özellikleri

Ümit Çelen ¹

Öz

Türkiye’de yapılandırmacı öğrenme-öğretme yaklaşımının benimsenmesiyle birlikte ölçme yöntem ve araçlarında bazı değişiklikler yaşanmaktadır. Yenilenen programlarda önerilen ölçme araçlarından birisi de tanılayıcı dallanmış ağaçtır. Bu aracın psikometrik özelliklerini belirlemek amacıyla yapılan bu araştırmada veriler üç farklı çalışma grubundan elde edilmiştir. Çalışma gruplarındaki toplam 525 öğrencinin ortak özellikleri öğretmen adayı olmaları ve 2012-2013 öğretim yılında ölçme ve değerlendirme dersini almalarıdır. Veriler ölçme ve değerlendirme dersindeki başarıyı ölçen ve araştırmacı tarafından geliştirilen tanılayıcı dallanmış ağaç, doğru-yanlış testi ve çoktan seçmeli başarı testleri kullanılarak toplanmıştır. Elde edilen bulgulardan hareketle, tanılayıcı dallanmış ağacın doğru-yanlış testlerinin barındırdığı dezavantajları barındırdığı, güvenilirlik ve geçerliliğinin yüksek olmadığı ve adında yer alan “tanılayıcılık” özelliğinden uzak olduğu sonuçlarına ulaşılmıştır.

Anahtar Kelimeler

Tanılayıcı dallanmış ağaç
Doğru-yanlış testleri
Geçerlilik
Güvenilirlik

Makale Hakkında

Gönderim Tarihi: 02.04.2013
Kabul Tarihi: 29.04.2014
Elektronik Yayın Tarihi: 06.08.2014

DOI: 10.15390/EB.2014.2690

Giriş

Eğitimde, öğrencileri belirli okul programlarına seçme, onların belirli özelliklerini tanıma, hazır bulunuşluklarını, öğrenme/yeterlilik seviyelerini ve dersin kazandırmayı hedeflediği özelliklere ulaşip ulaşamadıklarını belirleme gibi değişik amaçlarla ölçme yapılmaktadır. Eğitim ve öğretimin niteliği, değerlendirme sürecinin niteliğinden etkilenir; öğretim süreci ve değerlendirme arasında sıkı bir bağ var olduğu kabul edilmektedir. Öğretmen, öğretim ve değerlendirmenin içi içe olduğu gerçeğinden hareketle, sınıf içi değerlendirmeleri sadece not vermek için değil, öncelikle öğrencilerin akademik gelişimini izlemek ve öğretim sürecine dönüt vermek için yapar ve değerlendirmede farklı araç ve yöntemleri bir arada kullanmaya özen gösterir (Çıkrıkçı-Demirtaşlı, 2012).

Türkiye’de Milli Eğitim Bakanlığı (MEB), 2004 yılında geliştirilen ve 2005-2006 öğretim yılından itibaren uygulanan öğretim programlarında yapılandırmacılığı temel öğrenme-öğretme yaklaşımı olarak benimsenmiş ve bu yaklaşım ölçme ve değerlendirme süreçlerine de yansıtılmaya çalışılmıştır. Yenilenen programların ölçme ve değerlendirme yaklaşımı olarak, klasik ölçme araçlarının yanında süreci değerlendirmeye yönelik araçların da kullanılması önerilmiştir (MEB, 2005).

Yapılandırmacı yaklaşım, 1960’lı yılların ilk yarısında Bruner tarafından gündeme getirilmiş olsa da bu anlayışın daha eski felsefi kökenleri olduğunu söylemek mümkündür. Dubs (1993)’e göre davranışçı ve bilişçi yaklaşımlar anlam ve bilgiyi nesnel bir temele oturtmakta iken yapılandırmacılık

¹ Amasya Üniversitesi Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Ölçme ve Değerlendirme Anabilim Dalı, Türkiye, umitcelen@yahoo.com

bilginin içsel süreçlerle, bireysel olarak yapılandırıldığını varsaymaktadır. Bilgi ve doğrunun ne olduğu ve nerede olduğu konusunda davranışçı ve bilişçi gibi nesnel yaklaşımlarla yapılandırmacılık arasında önemli bir farklılaşma olduğu görülmektedir (Akt. Şimşek, 2004). Yapılandırmacılık, son dönem eğitim bilimleri literatüründe önemli bir yönelim ve tartışma konusu haline gelmiştir. “Nesnel bilgiyi tümüyle ya da büyük oranda reddeden bu yaklaşım; uzlaşmayı, işbirliğini, kültürü, bilginin değişkenlik, geçicilik ve durumsallığını temel almakta, öznellik ve göreliliği vazgeçilmez ilkeler olarak sunmaktadır. Bu yönleriyle benimsendiği kadar hafife de alınmaktadır” (Şimşek, 2004: 115).

Türkiye’de 2005–2006 öğretim yılında uygulanmaya başlanan ve yapılandırmacı yaklaşımı temel alan yeni öğretim programlarında, öğrenci merkezli öğretim yöntemi benimsenmektedir. Yenilenen programlarla birlikte ders içerikleri, öğretim yöntemleri, ders araç ve gereçleri ve ölçme ve değerlendirme yöntemleri de değişikliğe uğramıştır (Gelbal ve Kelecioğlu, 2007). Eğitim çevresinde yaygınlaşan yeni değerlendirme yaklaşımlarından biri performansa dayalı durum belirlemedir. Performansa dayalı durum belirlemede öğrencilerin basit, yalın, alt zihinsel düşünme gerektiren görevleri değil, daha üst düzey düşünme gerektiren karmaşık yapıdaki bazı görevleri yerine getirmeleri istenmektedir. Buradaki temel amaç öğrencinin yaratıcılık, problem çözme, eleştirel düşünme karar verme, empati kurma gibi yeteneklerini ortaya çıkarmak, geliştirmek, bu yeteneklerin ne düzeyde geliştiğini ve kullanılabildiğini belirlemektir (Kutlu, Doğan ve Karakaya, 2008). Geleneksel yöntemlerde öğrenci başarısı belirlenirken öğretim sürecinden ziyade ürüne ağırlık verilmekte; daha çok çoktan seçmeli, kısa yanıtı, yazılı ve sözlü sınavlar önemsenmektedir. Yapılandırmacı yaklaşımda ise ölçme ve değerlendirme, sürecin bir parçasıdır. Sürece ağırlık verilmesi geleneksel yaklaşıma göre daha sık ve farklı nitelikte ölçme araç veya yöntemlerinin kullanılmasını gerektirir (Gelbal ve Kelecioğlu, 2007).

Linn ve Gronlund (1995), klasik yöntemlerle ölçülemeyen öğrenme çıktılarının performansa dayalı durum belirleme ile ölçülebileceğini belirtmektedir. Yapılandırmacı yaklaşımda, kağıt-kalem testlerinin yanı sıra, öğrencinin davranışlarını ve süreçteki performansını gözlemleyerek, ilgilerini ve tutumunu ölçerek ve öğrenen bireyleri de değerlendirme sürecine katarak, öğrencilerin performansları her yönüyle değerlendirilebilmektedir (Gelbal ve Kelecioğlu, 2007).

Bachman (2002)’a göre, bazen alternatif veya otantik olarak da adlandırılan performans değerlendirme pek çok ülkede ilgi görmüş ve tartışma konusu olmuştur. Böylelikle araştırmacılar için özellikle geçerlilik gibi testin kalitesini gösteren özelliklere geleneksel bakışı yeniden gözden geçirmek için bir alan açmıştır. Literatürde performans değerlendirme süreçlerinde kullanılan pek çok ölçme ve değerlendirme yöntem ve araçları bulunmaktadır. Öğrenci ürün dosyaları (portfolyo), performans görevleri, projeler, öz değerlendirme formları, akran değerlendirme formları, grup değerlendirme formları, gözlemler, görüşmeler, posterler, sunum değerlendirme, dereceli puanlama anahtarları (rubrik), yapılandırılmış gridler, tanılayıcı dallanmış ağaçlar (TDA), kelime ilişkilendirmeler, kavram haritaları ve kontrol listeleri bu yöntem ve araçlar arasında yer alır (Özdemir, 2003). Performans değerlendirmede kullanılabilecek araçlar arasında sayılan TDA, MEB (2013) tarafından yayınlanan öğretim programlarında da yer almaktadır. Bu araç “bilişim teknolojileri” ve “fen ve teknoloji (4-5)” öğretim programlarında önerilmiş; “fen teknoloji (6-8) ve “biyoloji (9)” öğretim programlarında ise araç hakkında bilgi ve örnek verilerek açıklanmıştır. Biyoloji (9) dersi öğretim programında ölçme ve değerlendirme kapsamında kullanılacak araç ve yöntemlerin çoktan seçmeli, kısa cevaplı, eşleştirme, doğru yanlış ve essay tipi sorularından oluşan testler, performans değerlendirme, portfolyo değerlendirme, araştırma ve proje, kavram haritaları, yapılandırılmış grid, dallanmış ağaç, öz ve akran değerlendirme vb. ölçme teknikleri ve bu teknikler kapsamındaki ölçme araçları (dereceleme ve puanlama ölçekleri vb.) olduğu belirtilmiştir. Fen ve teknoloji (6-8) öğretim programında araçtan şöyle bahsedilmektedir (MEB, 2013):

“TDA belli bir konuda öğrencinin neleri öğrendiğini ve neleri öğrenemediğini belirlemek için kullanılabilir değerlendirme araçlarından biridir. Bu teknikte, temelden ayrıntıya giden bir sıraya göre doğru ve yanlış ifadeler seçilerek öğrenciden doğru seçimi yapması istenir. Böylece, 8 veya 16 seçimli bir ifadeler listesi ile sonlanan bir dallanmış ağaç oluşturulur.”

Johnstone ve arkadaşlarının 1986 yılında yayımlanan “Branching Trees and Diagnostic Testing” adlı makalesi, TDA isminin eğitim literatüründe ilk kez kullanıldığı yazıdır (Bahar, 2001). TDA tekniğinde bir ağaç grafiğinde yer alan birbiriyle ilişkili önerme cümleleriyle, yanıtlayıcıların doğru veya yanlış yanıtlarıyla bir çıkışa ulaştıkları ve bu sayede zihin yapılarındaki bilgi örüntülerini ve kavram yanılgılarını belirlemeyi amaçlayan bir araçtır (Kocaarslan, 2012). Bahar ve ark. (2009), TDA ile, yanıtlayıcının zihnindeki yanlış bağlantıların, hatalı stratejilerin ve yanlış bilgilerin açığa çıkarıldığını savunmaktadır. Kocaarslan (2012) ise yanıtlayıcının zihnindeki yapılarda yer alan bilgi örüntüleri ve anlamlı öğrenmelerini ölçmek için geliştirilen TDA’nın, öğrencilerin yeni bir şey öğrenirken önceki bilgi ve deneyimlerini kullanma ve bilgiler arasında ilişki kurma gibi amaçları gerçekleştirilmede uygun bir teknik olduğunu iddia etmektedir. Ayrıca, bu araç ile doğru yanlış madde türündeki %50 olan şansın azalarak %12.5’e düştüğünü belirtmektedir. TDA kullanmanın öğrenci başarısını artıracığına ilişkin görüşler (Şeyihoğlu ve Erbaş, 2010) ve kullanıldığında başarıyı artırdığına ilişkin bulgular (Karahana, 2007) vardır.

Aracın adında geçen “tanılayıcı” kavramının Türkçe sözlüklerdeki karşılığı “teşhis etmek”tir. Arapça kökenli teşhis kelimesi ise “kim ve ne olduğunu anlama, tanıma, seçme” anlamına gelmektedir (TDK, 2013). Tanılama yapması gibi bir görev yüklenen TDA’nın temelinde doğru-yanlış (DY) maddeleri yer almaktadır. DY maddeleri ise öğrencilerin yanlış öğrendikleri yerleri ortaya çıkarmada teşhis yönü bakımından çok zayıftır (Tekin, 1996). Bu zayıflığın iki sebebi vardır: Birincisi maddeyi şansa doğru yanıt olma olasılığı %50’dir. Bir birey o maddenin ölçtüğü özelliğe sahip olmadan da maddeyi %50 gibi yüksek bir ihtimalle doğru yanıtlayabilir. İkinci zayıf yön ise maddenin önermesinin yanlış olduğunu belirten birisinin, önermenin doğrusunu bilmeden de başka sebeplerle yanlıştır deme ihtimalinin bulunmasıdır. “Fransa’nın başkenti Roma’dır.” maddesine yanlış diyen öğrenci, doğru yanıtın Paris olduğunu bilmeden de yanlış demiş olabilir. 7 maddeli bir TDA’da yanıtlayıcı 3; 15 maddeli TDA’da ise 4 maddeyi yanıtlamak durumunda kalmaktadır. Yanıtlayıcının şansa doğru yanıt verme olasılığı 3 madde için $\left(\frac{1}{2}\right)^3 = \frac{1}{8}$; 4 madde için $\left(\frac{1}{2}\right)^4 = \frac{1}{16}$ ’dir. TDA’larda çıkış sayısı 8 ve 16’dır. Dolayısıyla bir bireyin hiçbir maddeyi okumadan bir çıkışı seçmesi durumunda en doğru çıkışı bulma olasılığı 3 ve 4 madde yanıtlatılan doğru yanlış testlerine göre değişmemekte, aynı kalmaktadır.

DY maddelerinin bir diğer zayıf yönü de, üst düzey zihinsel becerilerin ölçülmesindeki yetersizliğidir. Bir önerme okuyarak doğru ve yanlış olduğuna karar verirken kullanılan zihinsel süreç genellikle tanıma-hatırlama ve kavrama düzeyinde kalmaktadır. Maddelerin birbirine bağlantılı bir set oluşturularak sunulduğu TDA da DY maddelerinden oluştuğundan, üst düzey bir zihinsel beceriyi ölçmede yetersiz kalacaktır. Kutlu (2006), öğrenci davranışlarını değerlendirmek amacıyla kullanılan çoktan seçmeli (ÇS), kısa yanıt, DY, eşleştirmeli, boşluk doldurmalı gibi bazı klasik test yöntemlerinin problem çözme, okuduğunu anlama, eleştirel düşünme, analitik düşünme, empati kurma, araştırma yapma, karar verme, toplumsal tarihin önemini anlama, yaratıcılık gibi üst düzey zihinsel düzeyleri belirlemede yetersiz kaldığını belirtmektedir. Bu nedenle, öğrencilerin okul programlarında öngörülen kazanımlara ne derece ulaşıldığının belirlenmesinde performans, gerçek yaşam durumlarına ve portfolyoya dayalı durum belirleme yollarının kullanıldığını vurgulamaktadır. Eğer performansa dayalı durum belirlemede kullanılan yöntemler “klasik yöntemlerle ölçülemeyen öğrenme çıktıları ölçmede kullanılan yöntemler” olarak tanımlanırsa (Linn ve Gronlund, 1995), TDA’nın bu araçlar arasında yer alması tartışmaya açık bir konudur.

Yapılandırmacı yaklaşımla birlikte TDA üzerinde yapılan çalışmaların daha çok öğretmenlerin bu araç hakkındaki bilgi düzeyleri ve derslerinde bu aracı kullanıp kullanmadıklarını belirlemeye dönük araştırmalar olduğu görülmektedir (Aydoğmuş ve Coşkun Keskin, 2012; Karamustafaoğlu, Çağlak ve Meşeci, 2012; Şaşmaz Ören, Ormancı ve Evrekli, 2011; Çepni ve Şenel Çoruhlu, 2010; Özdemir, 2010; Şenel Çoruhlu, Er Nas ve Çepni, 2009; Okur, 2008; Pullu, 2008). Bu araştırmada ise geliştirilecek bir örnek üzerinde TDA’nın psikometrik özelliklerinin belirlenmesi amaçlanmıştır.

Yöntem

Araştırmanın modeli, Karasar (2003) tarafından ilişkiisel tarama modeli olarak adlandırılan modeldir. Araştırmanın amacı, kullanılan TDA'nın psikometrik özelliklerini belirlemek ve aracı bu özellikleri bakımından DY ve ÇS testlerle karşılaştırmaktır.

Araştırma Grubu

Araştırmada veriler Eskişehir Osmangazi Üniversitesi Eğitim Fakültesi ve İlahiyat Fakültesi'nde öğrenim gören 3 farklı çalışma grubundan elde edilmiştir. Birinci çalışma grubu (ÇG1), 2012-2013 öğretim yılı güz döneminde ölçme ve değerlendirme dersini alan ve çalışmaya katılmayı kabul eden lisans öğrencilerinden oluşmaktadır. İkinci çalışma grubundaki öğrenciler (ÇG2) aynı üniversitede formasyon sertifika programına devam eden ve aynı dönemde ölçme ve değerlendirme dersini alan öğrencilerdir. Üçüncü çalışma grubu (ÇG3) ise ölçme ve değerlendirme dersini 2012-2013 öğretim yılı bahar döneminde alan lisans öğrencilerden oluşturulmuştur. Her üç çalışma grubunun bölümlere göre dağılımına ilişkin bilgiler Tablo 1'de sunulmuştur.

Tablo 1. Çalışma Gruplarında Yer Alan Öğrencilerin Bölümlere Göre Dağılımı

Çalışma Grubu 1			Çalışma Grubu 2			Çalışma Grubu 3		
Bölüm	Sayı	%	Mezun Olunan Lisans Bölümü	Sayı	%	Bölüm	Sayı	%
RPD*	48	30.2	Türk Dili ve Edebiyatı	71	37.1	İlköğr. Matematik Ö.	66	37.9
Sınıf Ö.	91	57.2	Çağdaş Türk Leh. ve Ed.	5	2.6	İlköğr. Fen B. Ö.	65	37.4
DKAB**	20	12.6	İlahiyat	19	9.9	BÖTE***	29	16.7
Toplam	159	100.0	Tarih	16	8.3	Diğer****	14	8.0
			Hemşirelik	2	1.0	Toplam	174	100.0
			Matematik	49	25.5			
			Fizik	10	5.2			
			Kimya	10	5.2			
			Biyoloji	10	5.2			
			Toplam	192	100.0			

*Rehberlik ve Psikolojik Danışma Programı, **Din Kültürü ve Ahlak Bilgisi Öğretmenliği Programı, ***Bilgisayar ve Öğretim Teknolojileri Programı, ****Derse diğer bölümlerden kayıt yaptırmış öğrenciler.

Veri Toplama Araçları

TDA: Araştırmada kullanılan TDA araştırmacı tarafından geliştirilmiştir. Lisans programlarında 3; formasyon sertifika programlarında 2 kredilik bir ders olan "Ölçme ve Değerlendirme" dersi kapsamında hazırlanmıştır. Araç 15 adet DY maddesinden oluşmaktadır. Yanıtlayıcı bu 15 maddeden 4'üne yanıt vermek durumundadır. Maddelerden 8'inin anahtar yanıtı "yanlış", 7'sinin "doğru"dur. Ölçme değerlendirmede temel kavramlar konusundan hazırlanan araçta ölçme, değerlendirme, hata ve ölçek kavramlarına yönelik maddeler yer almıştır. Araç hazırlandıktan sonra, ölçme ve değerlendirme alanında çalışan araştırmacı dışında bir kişiden görünüş geçerliliği hakkında görüş alınmış ve öneriler doğrultusunda gerekli değişiklikler yapılmıştır.

Ara Sınav: ÇG1'in girmiş olduğu ölçme ve değerlendirme dersi ara sınavı 25 maddeli, 5 seçenekli maddelerden oluşan öğretmen yapımı bir testtir ve araştırmacı tarafından hazırlanmıştır. Ölçmenin önemi, ölçme ve değerlendirmeyle ilgili temel kavramlar ve ölçmede hata, ölçme araçlarında bulunması istenilen özellikler (güvenilirlik, geçerlilik ve kullanılabilirlik) konularını kapsayan testin kapsam geçerliliği sağlanmaya çalışılmıştır. Düzeltme formülü uygulanmayan testin ortalama güçlüğü'nün 0.70; nokta çift serili madde ayırıcılık güçleri ortalamasının 0.29, KR-20 iç tutarlılık katsayısının 0.54, aritmetik ortalamasının $\bar{x}=17.56$ ve standart sapmasının ise $S=3.04$ olduğu görülmüştür.

Dönem Sonu (Final) Sınavı: ÇG2'nin girmiş olduğu dönem sonu sınavı 40 maddeli, 5 seçenekli maddelerden oluşan öğretmen yapımı bir testtir ve araştırmacı tarafından hazırlanmıştır. Testin kapsamında yukarıda bahsedilen ara sınav konularının yanı sıra eğitimde kullanılan ölçme araçları ve özellikleri ve ölçme sonuçları üzerinde yapılan istatistiksel işlemler konuları da vardır. Uygulama sonucunda düzeltme formülü uygulanmayan testin ortalama gücünün 0.70, nokta çift serili madde ayırıcılık güçleri ortalamasının 0.29, KR-20 iç tutarlılık katsayısının 0.74, aritmetik ortalamasının $\bar{x}=27.93$, ve standart sapmasının ise $S=5.04$ olduğu görülmüştür.

Çalışmada ayrıca ÇG3'e doldurtulan 3 maddelik bir anket formu kullanılmıştır. Bu formdaki üç madde TDA'yı başarılarını belirlemede yeterli bulma durumları ve öğretmen olduklarında bu aracı kullanmayı düşünüp düşünmediklerini belirlemeye yöneliktir.

Verilerin Toplanması

ÇG1'den elde edilen veriler dönem ortasında yapılan ara sınavda ve bundan 3 hafta sonra yapılan TDA uygulamasıyla toplanmıştır. Ders saatinde yapılan TDA uygulaması sırasında öğrencilere önce testi yanıtlayarak bir çıkışa ulaşmaları gerektiği söylenmiştir. Uygulama tamamlandıktan sonra öğrencilere TDA'da yer alan 15 madde DY testi formatında verilmiş ve tüm maddeleri yanıtlamaları istenmiştir.

ÇG2'den elde edilen veriler dönem sonunda yapılan final sınavı esnasında toplanmıştır. A ve B grubu şeklinde maddelerin yerleri değiştirilerek oluşturulan iki test formundan birinin sonuna TDA, diğerine ise TDA'da yer alan maddelerden doğru yol üzerinde yer alan 4 madde eklenmiştir.

Öğrencilerin dersten geçme notları hesaplanırken, ara sınavın %30'u ve final sınavının %50'si alınmaktadır. Ders kapsamında verilen öğretmen yapımı bir test geliştirme ödevinin puanlarının da %20'si alınmakta ve geçme puanları bu şekilde oluşturulmaktadır.

ÇG3'ten elde edilen veriler, TDA kapsamında yer alan konuların işlenilmesinden 1 hafta sonra ders saatinde gerçekleştirilen uygulamada toplanmıştır. Öğrencilere TDA dağıtılarak yanıtlamaları istenmiş, uygulama bittikten sonra puanlama açıklanmıştır. Ulaşılan çıkışa göre her öğrencinin puanının belirlenmesinden sonra TDA'da yer alan tüm maddeler araştırmacı tarafından yanıtlanmıştır. Sonra öğrencilere TDA'yla ilgili görüşlerinin sorulduğu anket formu uygulanmıştır.

Verilerin Analizi

Araştırmada elde edilen tüm veriler Microsoft Excel programı kullanılarak bilgisayara aktarılmış ve analizler bu program ve SPSS 15.0 programı kullanılarak gerçekleştirilmiştir. Ayırıcılık gücü indeksi olarak hesaplanan istatistiklerde nokta çift serili korelasyon tekniğinden yararlanılmıştır. Test puanları arasındaki korelasyonlar incelenirken değişkenlerin normal dağılım varsayımını sağladığı görüldüğü için Pearson çarpım momentler korelasyon tekniği kullanılmıştır. TDA için madde gücü hesaplanırken birinci bölümü için oradaki tek maddenin gücü hesaplanmış, sonraki üç bölüm için o bölümden alınan puanların ortalamasından yararlanılmıştır. TDA'dan elde edilen puanlarla TDA'daki iki maddenin yeri değiştirilerek elde edilen tahmini puanlar arasındaki farkın incelenmesinde yine normallik varsayımı kontrol edilip sağlandığı gözlemlendiğinden, parametrik bir test olan ilişkili t-testi uygulanmıştır. ÇG1'e uygulanan TDA, DY testi ve tahminle bulunan TDA puanlarının ara sınav puanlarını yordayıp yordamadıklarını görmek için yapılan çoklu doğrusal regresyon analizinde stepwise yöntemi kullanılmıştır. Anket maddelerine verilen yanıtlar ise frekans ve yüzdelik dağılımlarına bakılarak incelenmiştir. Öğrencilerin TDA'yı yeterli bulma durumlarıyla aldıkları puanlar arasındaki ilişki incelenirken, normallik varsayımı sağlanmadığı için Kruskal Wallis varyans analizi uygulanmıştır.

Bulgular

TDA ve TDA'da yer alan 15 DY maddesinden oluşturulan testin ÇG1'e uygulanması sonucunda elde edilen puanlara ilişkin tanımlayıcı istatistikler Tablo 2'de sunulmuştur. Öğrenciler TDA'dan ortalama 2.59; DY testinden ise 10.45 puan almışlardır. Ortalama güçlükler incelendiğinde DY testinin öğrencilere daha kolay geldiği görülmektedir ancak aradaki fark sadece 0.05'tir. Öğrencilerin iki testten aldıkları puanlar arasındaki korelasyon ise 0.45 olarak hesaplanmıştır. TDA'nın dört bölümü için iç tutarlılık katsayısı 0.03'tür. Bu katsayı DY testinin 15 maddesi için 0.37'dir. DY testinde TDA'daki bölümlere karşılık gelen maddeler için bölüm puanı bulunup Cronbach alfa iç tutarlılık katsayısı hesaplandığında ise 0.29 olduğu görülmüştür. TDA puanlarında ölçmenin standart hatası 0.94; DY testinde ise 1.66 puandır.

Tablo 2. ÇG1'e Uygulanan TDA ve DY Testi Puanlarına İlişkin Test İstatistikleri (n=159)

İstatistik	TDA	DY Testi
\bar{x}	2.59	10.45
Medyan	3	11
Mod	3	11
Standart Sapma	0.90	4.36
En Düşük	0	5
En Yüksek	4	15
Ortalama Güçlük	.65	0.70
KR-20	0.03	0.37
Ölçmenin Standart Hatası	0.94	1.66

ÇG1'e uygulanan TDA ve DY Testine ilişkin madde istatistikleri Tablo 3'te verilmiştir. TDA madde güçlük indekslerinin 0.55 ile 0.79 arasında, DY testindeki maddelerin güçlük indekslerinin ise 0.32 ile 0.92 arasında değiştiği görülmektedir. TDA'da üçüncü bölümden itibaren maddelerin öğrencilere giderek daha zor geldiği görülürken, bu bölümlere denk gelen maddelerin DY testindeki ortalama güçlükleri incelendiğinde güçlüklerin benzer şekilde devam ettiği gözlenmektedir. TDA'nın bölümleri için hesaplanan ayırıcılık gücü indeksleri 0.29 ile 0.72 arasında değişmektedir ve güçlükte olduğu gibi üçüncü ve dördüncü bölümde bu değerler düşmektedir. DY maddelerinin ayırıcılık gücü indeksleri ise 0.12 ile 0.53 arasındadır. DY testinde, TDA'nın bölümlerine karşılık gelen maddelerden elde edilen bölüm toplam puanlarının toplam puanlarla korelasyonu incelendiğinde ise TDA'nın aksine üçüncü bölümden itibaren ayırıcılığın arttığı görülmektedir.

Öğrencilerin ara sınav puanlarıyla TDA ve DY testinden aldıkları puanlar bir geçerlilik kanıtı olarak incelenmiştir. Ara sınav, şans başarısının daha az olduğu 5 seçenekli ÇS maddelerden oluşmuş ve puanlar gerçek bir sınav uygulamasıyla elde edilmiştir. Bahsedilen bu korelasyonlar TDA için $r=0.17$ ve DY testi için 0.23 olarak bulunmuştur. Diğer puanın etki sabitlenerek hesaplanan kısmi korelasyonların ise TDA puanları için $r=0.09$; DY testi puanları için $r=0.17$ olmuştur. Ara sınavda kullanılan testin kapsamı TDA ve DY testine göre daha geniştir. Sadece TDA'nın kapsamında olan ölçme, değerlendirme, hata ve ölçek kavramlarıyla ilgili maddelerin yer aldığı 9 çoktan seçmeli maddenin puanlarıyla olan ilişki incelendiğinde ise bu katsayıların sırasıyla $r=.07$ ve $r=.08$ olduğu görülmüştür.

Tablo 3. ÇG1'e Uygulanan TDA ve DY Testi Madde İstatistikleri (n=159)

TDA			DY TESTİ				
Bölüm	Güçlük	Ayıcılık Gücü	Madde No	Güçlük	Ayırıcılık Gücü	Ortalama Güçlük	Bölüm Ayırıcılık Gücü
1	.62	.63	1	.60	.45	.60	.45
2	.79	.72	2	.67	.53	.77	.56
3	.64	.43	3	.87	.38	.71	.48
			4	.60	.35		
			5	.68	.28		
			6	.64	.12		
4	.55	.29	7	.92	.21	.69	.76
			8	.80	.36		
			9	.65	.34		
			10	.55	.39		
			11	.88	.16		
			12	.32	.27		
			13	.73	.38		
14	.79	.29					
15	.76	.24					

TDA'nın maddelerinin yerlerinin değiştirilmesi durumunda testten alınacak puanların nasıl değişeceğini belirlemek amacıyla öğrencilerin 15 DY maddesine verdiği yanıtlar kullanılarak elde edilen yeni yanıt örüntüsü incelenmiştir. İkinci bölümde yer alan ve verilen bir örnekte yapılan değerlendirmenin türünün belirleme hedefini ölçen iki maddenin yerleri değiştirilmiş, böylece her öğrenci için bu iki maddenin yeri değişik olması koşulunda hangi çıkıştan çıkacakları belirlenmiştir. Bu çıkışlar için yeni puan değerleri hesaplanmış ve tahmin edilen puanlarla mevcut TDA puanları arasındaki fark ilişkili t-testi ile incelenmiş ve sonuçlar Tablo 4'te sunulmuştur. TDA puanı ile tahmini TDA puanı arasındaki korelasyon 0.51'dir. İki maddenin yeri değiştirilerek hesaplanan puanlar TDA puanlarından istatistiksel olarak anlamlı şekilde yüksektir ($t=-3.01$, $p<0.01$).

Tablo 4. TDA puanları ve İki Maddenin Yeri Değiştirilerek Elde Edilen Tahmini TDA Puanlarının Karşılaştırılmasına İlişkin Test Sonuçları (n=159)

Puan Türü	\bar{x}	S	r	t	p
TDA'dan Elde Edilen	2.59	.95	.51	-3.013	.003
Tahmin Edilen	2.82	.97			

TDA puanları, DY testi puanları ve tahmin edilen TDA puanlarının ara sınav puanlarını yordayıp yordamadığını görmek üzere çoklu doğrusal regresyon analizi yapılmıştır. Analiz sonucunda sadece 15 maddeli DY testinden alınan puanların ara sınav puanlarını yordadığı görülmüştür ($F=8.35$, $p<0.01$).

Tablo 5. ÇG2'e Uygulanan TDA ve DY Testi İlişkin Test İstatistikleri

İstatistik	TDA (n=96)	DY Testi (n=96)
\bar{x}	2.47	2.13
Medyan	3	2
Mod	3	2
Standart Sapma	0.74	0.91
En Düşük	1	0
En Yüksek	4	4
Ortalama Güçlük	0.62	0.53

Final sınavında ÇG2'nin yarısına TDA ve diğer yarısına TDA'nın 4 puanlık çıkışına giden yoldaki 4 DY maddesi uygulanmış ve alınan puanlara ilişkin test istatistikleri Tablo 5'te sunulmuştur. Tablodan da görüleceği gibi, TDA'dan alınan puanların ortalaması 2.47'dir. DY testindeki maddeler ise öğrencilere daha zor gelmiştir ve ortalama 2.13 puan alınmıştır. DY testinden alınan puanların ranjı daha geniş, ortalaması daha düşük olmasına rağmen standart sapması daha yüksektir. Diğer bir deyişle TDA puanlarına göre daha heterojen bir dağılım göstermiştir.

ÇG2'ye uygulanan TDA ve DY testine ilişkin madde istatistikleri Tablo 6'da verilmiştir. TDA madde güçlük indekslerinin 0.45 ile 0.84 arasında, DY testindeki maddelerin güçlük indekslerinin ise 0.27 ile 0.91 arasında değiştiği görülmektedir. TDA'nın madde puanları ile TDA puanı arasındaki korelasyonlar ile hesaplanan ayırıcılık gücü indeksleri 0.17 ile 0.65 arasında değişmektedir. ÇG1'e uygulanan TDA'da olduğu gibi üçüncü ve dördüncü bölümde bu değerler düşmektedir. TDA madde puanlarının final puanlarıyla olan korelasyonları ise -0.04 ile 0.38 arasındadır. DY testinin güçlük indeksleri 0.27 ile 0.91 olarak hesaplanmıştır ve ortalama güçlüğü bakıldığında DY testinin öğrencilere daha zor geldiği görülmektedir. DY maddelerinin bu dört maddeden alınan puanla korelasyonu incelendiğinde 0.32 ile 0.65 arasında değişen değerler görülmektedir. Bu değerlerin ortalaması TDA maddelerinin ayırıcılık gücü ortalamasından 0.09 yüksektir ancak, final sınavından alınan puanlarla olan korelasyonlar incelendiğinde TDA maddelerinin korelasyonlarının ortalama 0.27 puan daha yüksek olduğu görülmektedir.

Tablo 6. ÇG2'ye Uygulanan TDA ve DY Testi Madde İstatistikleri

Bölüm	TDA (n=96)			Madde No	DY Testi (n=96)		
	Güçlük	Ayırıcılık*	Ayırıcılık**		Güçlük	Ayırıcılık*	Ayırıcılık**
1	.56	.65	.12	1	.55	.55	-.13
2	.84	.63	.38	2	.91	.32	.08
3	.45	.17	-.04	3	.27	.61	-.34
4	.62	.21	.24	4	.41	.55	.02
Ortalama	.62	.42	.18		.54	.51	-.09

* İlgili bölümden alınan puanla olan korelasyon.

** Final sınavından alınan puanlarla olan korelasyon.

ÇG2'deki öğrencilerin TDA'dan ve DY testinden aldıkları puanlarla ara sınav, final sınavı ve bu ikisine ek olarak bir ödev puanıyla birlikte hesaplanan geçme notu arasındaki ilişkiler incelenmiş ve bulgular Tablo 7'de sunulmuştur. TDA puanlarıyla her üç puan arasında istatistiksel olarak anlamlı pozitif korelasyonlar vardır. DY puanları ise üç puanla da oldukça düşük ve negatif korelasyonlar göstermiştir.

Tablo 7. ÇG2'ye Uygulanan TDA ve DY Testinin Geçerlilik Katsayıları

Test	n	Ara Sınav	Final	Geçme Notu
TDA	96	.31**	.40**	.44**
DY	96	-.03	-.20*	-.05

* $p < 0.05$, ** $p < 0.01$

ÇG3'te yer alan öğrencilere TDA uygulandıktan, puanları açıklandıktan ve 15 DY maddesinin de yanıtları tartışıldıktan sonra öğrencilere bu aracın başarılarını belirlemede ne derece yeterli olduğunu düşündükleri ve öğretmen olduklarında derslerinde bu aracı kullanmayı düşünüp düşünmedikleri sorulmuştur. Bu sorulara verdikleri yanıtlar Tablo 8'de sunulmuştur. Tablodan da görüleceği gibi uygulama yapılan tüm bölümlerde öğrencilerin yarısı başarılarını belirlemede TDA oldukça yeterli bulmuşlardır. Orta düzeyde yeterli bulan öğrenciler ise grupların yaklaşık 3'te birini oluşturmaktadır. Hiç yeterli değil diyen öğrenci olmamıştır. Öğretmen olduğunda TDA'yı kullanmayacağını belirten sadece 1 öğrenci vardır, %99.42ü TDA'yı kullanacağını bildirmiştir.

Tablo 8. ÇG3'te Yer Alan Öğrencilerin Kendilerine Uygulanan TDA Hakkındaki Görüşleri

TDA'yı Yeterli Bulma Durumu	BÖLÜM									
	İlköğ. Mat. Ö.		İlköğ. Fen B. Ö.		BÖTE		Diğer		Toplam	
	Sayı	%	Sayı	%	Sayı	%	Sayı	%	Sayı	%
Hiç Yeterli Değil	-	-	-	-	-	-	-	-	-	-
Çok Az Yeterli	2	3.0	2	3.1	5	17.2	2	14.3	11	6.3
Orta Düzeyde Yeterli	21	31.8	23	35.4	10	34.5	4	28.6	58	33.3
Oldukça Yeterli	39	59.1	34	52.3	14	48.3	5	35.7	92	52.9
Tamamen Yeterli	4	6.1	6	9.2	-	-	3	21.4	13	7.5
Öğretmen Olduğunda Derslerinde TDA Kullanmayı Düşünme Durumu										
Kullanacak	66	100.0	65	100.0	29	100.0	13	92.9	173	99.4
Kullanmayacak	-	-	-	-	-	-	1	7.1	1	0.6
Toplam	66	100.0	65	100.0	29	100.0	14	100.0	173	100.0

TDA'yı yeterli bulma durumunun, TDA'dan alınan puana göre değişiklik gösterip göstermediğini belirlemek amacıyla belirtilen her yeterlilik düzeyindeki öğrencilerin uygulanan TDA'dan almış oldukları puanların ortalaması incelenmiş ve bulgular Tablo 9'da sunulmuştur. Aracı tamamen yeterli bulanlar araçtan ortalama 3.50 puan alanlardır. Çok az yeterli bulduğunu bildiren öğrenciler ise en düşük ortalama (2.45) sahip gruptur. Yapılan Kruskal Wallis varyans analizi sonuçlarına göre TDA'dan alınan puanların belirtilen yeterlilik düzeyine göre farklılaştığı görülmüştür ($KW\chi^2= 10.41, p<0.05$).

Tablo 9. ÇG3'te Yer Alan Öğrencilerin TDA Puanları ve TDA'yı Yeterli Bulma Durumları Arasındaki İlişki

TDA'yı Yeterli Bulma Durumu	n	\bar{x}	S	Sıra Ortalaması	χ^2	p
Çok Az Yeterli	11	2.45	1.04	63.64	10.41	0.015
Orta Düzeyde Yeterli	53	2.70	0.97	73.33		
Oldukça Yeterli	83	2.89	0.84	81.62		
Tamamen Yeterli	12	3.50	0.80	113.25		
Toplam	159*	2.84	0.92			

*Bazı öğrencilerin puanları yönergeye uymamaları nedeniyle tespit edilememiştir.

Tartışma, Sonuç ve Öneriler

Bu çalışmada, bir grup öğrenciye TDA uygulanmış ve aynı öğrencilere daha sonra TDA'da yer alan 15 DY maddesi yanıtlatılmıştır. TDA uygulaması sonucunda, ortalama güçlük 0.65; 15 maddeli DY testi uygulamasında ise 0.70 olarak bulunmuştur. Bir başka ifadeyle, sadece 4 madde yanıtlanan TDA ile 15 madde yanıtlanan DY testinin güçlükleri birbirine çok yakındır. İki formdan alınan puanlar arasındaki korelasyon 0.45 olarak hesaplanmıştır. Aynı maddelerden oluşturulmuş ve aynı ders saati içinde uygulanmış, yanıtlarken aynı 4 maddenin her iki testte de mutlaka yanıtlandığı test puanları arasında bulunan bu orta düzeyli korelasyon, puanlar arasındaki paralelliğin çok yüksek olmadığını işaret etmektedir. Şeyihoğlu ve Erbaş (2010) da DY ve TDA kullanarak yaptıkları karşılaştırmada kullanılan tekniğe göre öğrenci başarısının değiştiği sonucuna varmışlardır. Korelasyonun daha yüksek çıkmamasının bir sebebi, kullanılan madde formatı olan DY madde türünün şans hatasını puanlara fazlasıyla karıştırması ve teşhis yönünün zayıf olmasıyla açıklanabilir. Diğer bir sebep ise iki test formatı arasındaki farklılık olabilir. İç tutarlılık güvenilirlikleri incelendiğinde 15 maddeli testten elde edilen puanların KR-20 katsayısı 0.37; TDA'dan elde edilen puanların ise 0.03 olarak bulunmuştur. Madde sayıları ve puanların varyanslarının farklı olması nedeniyle karşılaştırmayı daha sağlıklı yapabilmek için ölçmenin standart hatası hesaplanmıştır. TDA'da 0-4 puan aralığında elde edilen puanlardaki standart hata 1 puana yakındır. Puan ranjı TDA puanlarınınkinden 2 katından fazla olan ve teorik olarak 0-15 aralığında puanlar alınabilen DY testinde ise standart hata 1.67 puandır. Bu bulgulara bakılarak TDA puanlarına göre olarak daha fazla hata karıştığı söylenebilir.

Yine ÇG1'e uygulanan TDA ve DY testi madde istatistikleri incelendiğinde, TDA'da 3. sırada yanıtlanan maddelerden itibaren maddelerin zorlaştığı görülmektedir. 2. bölümdeki iki maddenin doğru yanıtlanma yüzdesi 0.79 iken, son yanıtlanan maddelerin yüzdesi 0.55'tir. DY uygulamasından elde edilen, aynı bölümlerde yer alan maddelerin güçlük ortalamaları incelendiğinde ise, güçlük düzeyinin yaklaşık olarak aynı kaldığı gözlenmiştir. Madde sayısının arttığı 3. bölümden itibaren TDA'da güçlük artarken DY testinde aynı kaldığından daha fazla maddenin yanıtlatıldığı DY testiyle önemli bir farklılaşma gerçekleşmiştir. Madde ayrıcalık gücü indeksleri incelendiğinde ise ilk üç madde ve aynı bölümdeki maddeler için bir paralellik söz konusu iken son bölümde DY lehine önemli bir farklılık vardır. TDA'nın son maddesinin ayrıcalık gücü 0.29 iken DY testinde bu bölümün ayrıcalık gücü 0.76'dır.

Uygulanan TDA ve DY testinin bir geçerlilik kanıtı olarak öğrencilerin aynı konuların yoklandığı ve çoktan seçmeli maddelerden oluşan ara sınav puanlarıyla ilişkisi incelenmiştir. Bu inceleme sonucunda DY testi için biraz daha yüksek olmakla birlikte $r=0.17$ ve $r=0.23$ gibi düşük geçerlilik katsayıları elde edilmiştir. Kısmi korelasyonlar incelendiğinde ise TDA puanlarının ara sınav puanlarıyla ilişkisi neredeyse hiç kalmamaktadır. Ara sınav puanları ölçüt kabul edildiğinde, her ikisinin de geçerliliği düşük olmakla birlikte DY testinin TDA'ya göre daha yüksek geçerliliğe sahip olduğu söylenebilir.

TDA'da maddelerin dizilişi çok önemlidir. Çünkü daha fazla sayıda madde olmasına rağmen öğrenci verdiği yanıtların gösterdiği yolu kullandığı için bazı maddeleri yanıtlanmamaktadır. Bazı maddelerin yerlerinin değişmesinin puanları nasıl etkileyeceğini görmek için, öğrencilerin 15 maddeye verdiği yanıtlar kullanılarak ikinci bölümdeki iki maddenin yeri değiştirilmiş ve böylece yeni puanlar tahmin edilmiştir. Bu şekilde elde edilen iki puan arasındaki korelasyon yine orta düzeyli bir değer olarak ($r=0.51$) bulunmuştur. Bu da iki maddenin yerinin değiştirilmesinin TDA'dan alınan puanları önemli ölçüde değiştirdiğinin bir göstergesidir. Tahminle elde edilen puanlarla uygulamadan elde edilen puanların farkı anlamlı mı diye bakıldığında ise tahmin puanlarının istatistiksel olarak anlamlı şekilde TDA puanlarından yüksek olduğu görülmektedir ($t=-3.01$, $p<0.01$).

TDA, tahmin edilen TDA ve DY testi puanlarından hangilerinin ara sınavda uygulanan test puanlarını yordadığına ilişkin çoklu doğrusal regresyon analiziyle yapılan incelemede TDA ve tahmin

edilen TDA puanlarının regresyon denkleminde giremediği görülmüştür. Bir başka ifadeyle her iki şekilde hesaplanan TDA puanları, 15 maddenin de uygulandığı DY testi kadar geçerli bulunmamıştır.

ÇG2'nin yarısına TDA, diğer yarısına da TDA'nın 4 puanlık çıkışına giden yoldaki 4 DY maddesi final sınavıyla birlikte uygulanmıştır. Benzer iki gruba uygulanan bu testlerin ortalama güçlükleri incelendiğinde TDA'nın öğrencilere daha zor geldiği görülmektedir. Kocaarslan'ın (2012) "TDA, DY testine göre şans başarısını azaltır" iddiasının doğru olması durumunda ortalama güçlük bakımından DY testinin daha kolay olması beklenebilir. Şeyihoğlu ve Erbaş'ın (2010) çalışmasında da TDA'da şans başarısının daha düşük olduğuna ilişkin bulgulara ulaşıldığı belirtilmektedir. Oysa bu çalışmadaki bulgular bu durumun tam tersini göstermektedir. DY testi maddelerinin bu maddelerden alınan toplam puanla ilişkisi hesaplanarak bulunan ayırıcılık gücü indekslerinin TDA bölümlerine göre daha yüksek olduğu görülmektedir. Final sınavından alınan puanlar ölçüt kabul edildiğinde ise tüm katsayıların önemli ölçüde düştüğü ve TDA bölümlerinin ayırıcılığının DY testi maddelerinden yüksek olduğu görülmüştür. TDA ve DY testinden alınan puanların öğrencilerin ara sınav, final ve geçme notlarıyla ilişkileri incelendiğinde ise TDA'nın daha yüksek geçerlilik katsayılarına sahip olduğu görülmektedir. Bu bulgu, TDA'nın 4 maddeli bir DY testinden daha geçerli olduğu şeklinde yorumlanabilir.

ÇG3'te yer alan öğrencilere TDA uygulandıktan ve puanları açıklandıktan sonra, TDA'yı bu konulardaki başarılarını belirlemede ne derece yeterli buldukları sorulmuştur. Öğretmen adayları öğrencilerin hiç birinin bu aracın yetersiz olduğunu belirtmemesi ilginç bir bulgudur. Çok az yeterli diyen öğrencilerin oranı ise % 6.3'tür. Öğrencilerin yarıdan fazlası bu aracı başarılarını belirlemede oldukça yeterli ve tamamen yeterli olduğunu düşünmektedir. TDA'yı yeterli bulma durumu ve TDA'dan alınan puan arasındaki ilişki incelendiğinde ise aslında başarılarını belirlemede TDA'yı yeterli bulanların TDA'dan yüksek puan alan öğrenciler olduğu, düşük puan alanların ise aracı yetersiz gördükleri dikkat çekmektedir.

TDA üzerinde yapılan araştırmalarda, öğretmenlerin TDA konusunda kendilerini yeterli görmedikleri ve hazırlanmasının çok zaman alması, bu konudaki bilgi yetersizlikleri ve sınıfların kalabalıklığı gibi sebeplerle derslerinde çoğunlukla TDA kullanmadıkları sonuçlarına ulaşmıştır (Aydoğmuş ve Coşkun Keskin, 2012; Çepni ve Şenel Çoruhlu, 2010;; Karamustafaoğlu, Çağlak ve Meşeci, 2012; Okur, 2008; Özdemir, 2010; Pullu, 2008; Şaşmaz Ören, Ormancı ve Evrekli, 2011; Şenel Çoruhlu, Er Nas ve Çepni, 2009). Bu çalışmada ise öğrencilere öğretmen olduklarında TDA kullanmayı düşünüp düşünmedikleri sorulduğunda neredeyse tamamının (%99.4) aracı kullanmayı düşündüğü görülmüştür.

Üç farklı gruba uygulanan TDA'ya ilişkin şu sonuçlara ulaşılmıştır:

- Doğru yanlış maddelerinden oluşan TDA, bu madde türünün dezavantajlarını içinde barındırmaktadır. Şans hatası yine çok yüksektir, teşhis gücü zayıftır. Aracın adının tanılayıcı olması şekilsellikten ileriye gitmemektedir.
- Rasgele hatalardan olan şans hatasının yüksek olması, araçla elde edilen puanların güvenilirliğini ve geçerliliğini olumsuz etkilemektedir. TDA uygulandığında, 4 maddeli DY testine göre daha geçerli puanlar bulunmuştur ancak, TDA'da yer alan 15 DY maddesinin hepsinin uygulanması durumunda puanlar daha geçerli ve güvenilir olacaktır.
- Öğretmen adayları TDA'yı geçerli bir araç olarak görmüşlerdir. Halen öğretmenlik yapanlar bu aracı çeşitli sebeplerle kullanmıyorlarken, TDA uygulanan öğretmen adayları derslerinde bu aracı kullanacaklarını belirtmektedir.
- TDA üzerinde yapılan çalışmalarda, aracı kullanmanın öğretimi desteklediğine ilişkin bulgulara ulaşılmıştır ancak bir aracın işe yaraması için geçerli, güvenilir ve kullanışlı olması şarttır. Bu çalışmada elde edilen bulgular TDA'nın bu özellikler bakımından yeterliliğinin kısıtlı olduğunu göstermektedir. Ayrıca, geleneksel ölçme araçlarıyla ölçülemeyen her hangi bir özelliği de ölçmediği açıktır. Bu nedenle aracın performansa dayalı değerlendirme yöntemleri arasında yer almasının da uygun olmadığı düşünülebilir.

Kaynakça

- Aydoğmuş, A. & Coşkun Keskin, S. (2012). Sosyal bilgiler öğretmenlerinin süreç odaklı ölçme ve değerlendirme araçlarını kullanma durumları: istanbul ili örneği. *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 8 (2), 110-123.
- Bachman, L. F. (2002). "Alternative interpretations of alternative assessments: some validity issues in educational performance assessments." *Educational Measurement: Issues and Practice*, 21 (3), 5-18.
- Çepni, S. & Şenel Çoruhlu, T. (2010). Alternatif ölçme ve değerlendirme tekniklerine yönelik hazırlanan hizmet içi eğitim kursundan öğretime yansımalar. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 28 (1), 117-128.
- Çıkrıkçı-Demirtaşlı, N. (2012). Öğrenme, öğretim ve değerlendirme arasındaki ilişkiler, N.Çıkrıkçı-Demirtaşlı (Ed.) *Eğitimde Ölçme ve Değerlendirme* (ss. 3-29). Ankara: Elhan Yayınları.
- Gelbal, S. & Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33, 135-145.
- Johnstone, A. H., McAlpine, E. & MacGuire, P.R.P. (1986). Branching trees and diagnostic testing. *A Journal for Further and Higher Education in Scotland*, 2, 4-7.
- Karahan, U. (2007). *Alternatif ölçme ve değerlendirme metodlarından grid, tanılayıcı dallanmış ağaç ve kavram haritaları'nın biyoloji öğretiminde uygulanması*. Yayınlanmamış yüksek lisans tezi, Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Karamustafaoğlu, S., Çağlak, A. & Meşeci, B. (2012). Alternatif ölçme değerlendirme araçlarına ilişkin sınıf öğretmenlerinin öz yeterlilikleri. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, 1 (2), 167-179.
- Karasar, N. (2000). *Bilimsel araştırma yöntemi (10. Baskı)*. Ankara: Nobel Yayın Dağıtım.
- Kocaarslan, M. (2012). Tanılayıcı dallanmış ağaç tekniği ve ilköğretim 5. sınıf fen ve teknoloji dersi maddenin değişimi ve tanınması adlı ünite de kullanımı. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9 (18), 269-279.
- Kutlu, Ö., Doğan, C.D. & Karakaya, İ. (2008). *Öğrenci başarısının belirlenmesi, performansa ve portfolyoya dayalı durum belirleme*. Ankara: Pegem Akademi.
- Linn, R. L. & Gronlund, N.E. (1995). *Measurement assesment in teaching*. New Jersey: Prentice Hall Inc.
- MEB (2013). *Öğretim programları*, Talim ve Terbiye Kurulu, <http://ttkb.meb.gov.tr/www/ogretim-programlari/icerik/72> adresinden 19.03.2013 tarihinde erişilmiştir.
- MEB (2005). *İlköğretim 1-5. sınıf programları tanıtım el kitabı*. MEB, TTK Başkanlığı, Eğitim Öğretim ve Program Dairesi Başkanlığı. Ankara: Devlet Kitapları Müdürlüğü Basımevi.
- Okur, M. (2008). *4. ve 5. Sınıf öğretmenlerinin fen ve teknoloji dersinde kullanılan alternatif ölçme ve değerlendirme tekniklerine ilişkin görüşlerinin belirlenmesi*. Yayınlanmamış yüksek lisans tezi, Bülent Ecevit Üniversitesi Sosyal Bilimler Enstitüsü, Zonguldak.
- Özdemir, S. M. (2010). İlköğretim öğretmenlerinin alternatif ölçme ve değerlendirme araçlarına ilişkin Yeterlilikleri ve Hizmet İçi Eğitim İhtiyaçları. *Türk Eğitim Bilimleri Dergisi*. 8(4), 787-816.
- Pullu, S. (2008). *Sınıf öğretmenlerinin ilköğretim programlarındaki ölçme ve değerlendirmeye yönelik görüşleri ve uygulamaları (Elazığ ili örneği)*. Yayınlanmamış yüksek lisans tezi, Fırat Üniversitesi Sosyal Bilimler Enstitüsü, Elazığ.
- Şaşmaz Ören, F., Ormancı, Ü. & Evrekli, E. (2011). Fen ve Teknoloji Öğretmen Adaylarının Alternatif Ölçme-Değerlendirme Yaklaşımlarına Yönelik Öz-Yeterlilik Düzeyleri ve Görüşleri. *Kuram ve Uygulamada Eğitim Bilimleri (KUYEB)*, 11(3), 1675-1698.
- Şenel Çoruhlu, T., Er Nas, S. & Çepni, S. (2009). Fen ve teknoloji öğretmenlerinin alternatif ölçme değerlendirme tekniklerini kullanmada karşılaştıkları problemler: trabzon örneği. *Yüzcüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 1(I), 122-141.

- Şeyihođlu, A. & Erbaş, A.A. (20-22 Mayıs 2010). *Hayat bilgisi dersinde tanılayıcı dallanmış ağaç tekniđiyle doğru-yanlıř test tekniđinin karşılaştırılması*. 9. Ulusal Sınıf Öğretmenliđi Eđitimi Sempozyumu, Elazığ.
- Şimşek, N. (2004). Yapılandırmacı öğrenme ve öğretime eleştirel bir yaklaşım. *Eđitim Bilimleri ve Uygulama*, 3 (5), 115-139.
- TDK (2013). *Büyük türkçe sözlük*. <http://tdkterim.gov.tr/bts/> adresinden 19.03.2013 tarihinde erişilmiştir.
- Tekin, H. (1996). *Eđitimde ölçme ve deđerlendirme* (9. Baskı), Ankara: Yargı Yayınevi.