



Okul Etkililiğini Belirlemede Kullanılan Katma-Değerli Değerlendirme Yaklaşımının Uygulanması *

Sedat Şen ¹, İbrahim Yıldırım ², Mehmet Fatih Karacabey ³

Öz

Çalışmanın amacı okul ve öğretmen değerlendirmesinde kullanılan yaklaşımlardan biri olan katma-değerli değerlendirme modelinin tanıtılması ve gerçek bir veri üzerinden uygulamasının gösterilmesidir. Bu çalışmada Şanlıurfa ilindeki 539 okula ait TEOG sınav puanları kullanılarak katma-değerli değerlendirme modeli vasıtasıyla okullar arasında başarı sıralaması yapılmıştır. Bunun yanında sadece ortalamalara bakılarak yapılan okul sıralamaları da sunulmuştur. Çalışmada elde edilen katma-değerli puanlarına dayalı sıralamalar ile sadece ortalamalara dayalı yapılan geleneksel sıralamalar arasında farklılıklar olduğu tespit edilmiştir. Sadece bir sınav üzerinden okul ortalamalarına bakarak okulların başarısının değerlendirilmesinin doğru olmadığı okulların bir sınavdan diğerine ne kadar değişim gösterdiğini yansıtan katma-değerli değerlendirme yaklaşımının daha objektif sonuçlar sunduğu görülmüştür. Katma-değerli değerlendirme yaklaşımının Türkiye’de kullanımına örnek teşkil eden bu çalışma ileride yapılması planlanan okul ve öğretmen değerlendirmelerine yardımcı olacak önerilerde bulunmaktadır.

Anahtar Kelimeler

Okul etkililiği
Katma-değerli değerlendirme
Katma-değerli değerlendirme modelleri

Makale Hakkında

Gönderim Tarihi: 05.07.2018
Kabul Tarihi: 13.02.2020
Elektronik Yayın Tarihi: 22.04.2020

DOI: 10.15390/EB.2020.8023

Giriş

Bireylerin eğitimiyle ve eğitim ortamlarıyla ilgili bazı radikal yaklaşımlar bulunmakla beraber, dünya genelinde bireylerin eğitimi okullarda gerçekleştirilmektedir. UNESCO tarafından 2019 yılında yayımlanan rapora göre dünya çapında ilköğretime kayıt yaptıran öğrenci oranı %92 iken mezun olan öğrenci oranı %84’tür (UNESCO, 2019). Okullaşma oranı açısından Türkiye incelendiğinde ise 2005-2006 eğitim öğretim yılında ilköğretimde okullaşma oranı %89.77, ortaöğretimde %56.63, yükseköğretimde ise % 18.85 iken bu oranlar 2018-2019 eğitim öğretim yılında ilköğretim için %91.92’ye, ortaöğretimde %93.28’e, yükseköğretimde ise %44.10’a çıkmıştır (Milli Eğitim Bakanlığı [MEB], 2019). Bu okullar birçok ülkede devlet tarafından finanse edilirken, bazı ülkelerde özel girişimin ağırlıkta olduğu da görülmektedir. İster devlet ister özel finans kaynakları ile desteklendiği fark etmeksizin, bu kurumlara ayrılan maddi kaynağın beklenen faydayı sağlayıp sağlayamadığının belirlenebilmesi bir problem durumu olarak tanımlanabilir. Bu problem durumunun açıklığa kavuşturulabilmesi okulun anlamını ve işlevini daha net ortaya koyabilme yolunda yani etkililik bağlamında önemli bir adım olacaktır.

* Bu makale 16’ncı ICTEL 2019 konferansında sözlü bildiri olarak sunulmuştur.

¹ Harran Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Türkiye, sedatsen06@gmail.com

² Gaziantep Üniversitesi, Gaziantep Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Türkiye, iyildirim84@gmail.com

³ Harran Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Türkiye, karacabeyakademi@gmail.com

Çok yaygın bir şekilde kullanılıyor olmasına rağmen, etkililik kavramı örgütün içinde bulunduğu çevre, ulaşmayı hedeflediği amaç ve paydaşlarına göre anlam kazanan ve bu değişkenlere bağlı olarak farklılık gösteren bir kavramdır. Etkililik, bir örgütün gerçekleştirdiği faaliyetler sonucunda, belirlediği amaçlara ne derece ulaştığına ilişkin bir performans göstergesi (Yükçü ve Atağan, 2009), örgütün amaçlarına ulaşma ve bu amaçlarını gerçekleştirme derecesidir.

Bir örgütün etkililiği bağlamında Barnard (1938) formal örgütlenmenin önemini vurgularken, formal yapılanma yoluyla bir örgütün etkililiği ve verimliliğini arttırmanın mümkün olduğunu iddia etmektedir. Bir örgütün varlığının devamlılığı için örgüt içinde denge kurmanın, örgütün dış çevreye uyum sağlamanın ve üniter bir sistem olarak örgüt içinde yöneticilerin performanslarının etkilerinin analiz edilmesinin önemli olduğunu vurgulamaktadır. Ona göre bir örgütün formal bir yapıya gereksinimi vardır çünkü örgüt paydaşlarının örgütün devamlılığı, etkililiği ve verimliliği için çalışması gereklidir. Bu doğrultuda yöneticiler etki ve kontrol yoluyla çalışanların eylemlerini ve amaçlarını değiştirebilirler, çevresel dış güçlere karşı örgütün iç dengesinin korunmasına yönelik eylem ve politika geliştirebilirler (Barnard, 1938).

Bir örgütün etkililiği genellikle amaç merkezli bir yaklaşım (örgütlerin belirledikleri amaçlara ulaşmadaki başarısı) ya da doğal sistem yaklaşımı (örgütün yaşamını devam ettirmesi) olmak üzere iki yaklaşım çerçevesinden incelenmektedir (Rowan, 1985). Eğitim örgütleri olan okulların etkililiği amaç-merkezli yaklaşıma göre incelendiğinde, onların önceden belirlenmiş öğrenim çıktıklarına ulaşip ulaşmadıkları ya da ne kadar ulaştıkları onların etkililiğinin göstergesidir. Daha geniş bir çerçeveden etkililiği inceleyen doğal sistem yaklaşımına göre ise bir okulun etkililiği, bireyin aldığı eğitim sonucu gelişen sosyal ve yaşam becerileri ile değer ve davranışlarını kapsayacak şekilde incelenmektedir (Karip ve Köksal, 1996). Okulların etkililiğinin değerlendirilmesinde karşılaşılan en büyük sorun, eğitim sisteminin kısa ve uzun dönem, gizil ve açık hedefleri olması nedeniyle bu hedeflere ne derece ulaşıldığının belirlenmesinin oldukça zor olmasıdır. Fakat ne kadar zor olsa da bir okulun etkililiğinin ölçülmesinde önceden belirlenen amaç ve hedeflere ne kadar ulaşıldığının bir ölçüt olarak belirlenmesi çok yanlış olmaz (Creemers ve Kyriakides, 2008). Sonuç olarak bir okulun etkililiği onun önceden belirlenen hedeflerine ulaşma derecesidir ve bu bağlamda etkili bir okul ise kurumsal, fiziksel, insan kaynakları gibi açılardan bu hedeflere ulaşmaya olumlu yönde etkisi olan özelliklere sahip olan okuldur (Botha, 2010). Kısaca; bir okulun etkililiğinin öğrenci öğrenmelerinin ne kadar gerçekleştiği ile ilişkili olduğu söylenebilir.

Okul etkililiğinin belirlenmesi bağlamında dünya genelinde birçok girişim ortaya çıkmıştır. Bir okulun etkililiğini okulun genel başarı puanı ile değerlendirmenin yanı sıra veri zarflama metodu ve benzeri yaklaşımlar aracılığıyla öğretmenlerin öğretim yeterlik düzeyleri, okuldaki eğitim girdileri, kullanılan materyallerin çeşidi/miktarı dikkate alınarak değerlendirme gibi yaklaşımlar sergilenmiştir. Veri zarflama metodu benzeri bir yaklaşımla EARGED de (2010) bir okulun başarısının değerlendirilmesinde "Okul Yönetimi", "Çalışanların Yeterliliği", "Eğitim Öğretim Süreçlerinin Niteliği" ve "Destek Hizmetler" ana boyutlarının dikkate alınması gerektiğini belirtmektedir. Adams (1993) 80'li ve 90'lı yıllarda kaliteli öğretim kurumlarına olan ihtiyaca dikkatin yöneldiğini, fakat her paydaşın bu kaliteyi farklı tanımlamasından ötürü ortak bir paydada buluşmanın pek mümkün olmadığını belirtirken, Hesapçioğlu (2006) ise kalite olgusunun, etkililik, verimlilik, performans, yeterlilik kavramları ile iç içe olduğunu ifade etmektedir. Bu fikirden hareketle, bir okulun etkililiğinin en önemli göstergesi nedir? sorusunun cevabı büyük oranda öğrenci başarısı anlamına geldiğini söylemek mümkündür. Bu durumda bu kadar farklı değişkenlerle ve kontrol edilemeyen durumlarla uğraşmak yerine direkt olarak öğrenci başarısına odaklanmak daha akla yakın gelmektedir. Bu durumda, okul kalitesinin ve öğrenme süreçleri çıktıklarına dair kalite göstergelerinin en önemli unsurlarından olan öğrenci başarısı tüm dikkati üzerinde toplamaktadır.

Eğitimde kalitenin artırılmasına yönelik ve öğrencilerin başarılarını dikkate alarak gerçekleştirilebilecek okul performans ölçütleri geliştirmek uzun yıllardır eğitimcilerin ilgisini çeken bir konu olmuştur. Hâlihazırda yapılan uygulamada MEB her bir okulun ne kadar başarılı olduğunu o okuldaki öğrencilerin ulusal ölçekli sınavlardan (TEOG ve YGS) aldıkları puanlara göre belirlemektedir. Okulların başarı durumlarının tek bir noktada (genelde yılsonunda) yapılan bir sınavdan elde edilen tek bir ölçme sonucuna göre değerlendirilmesi araştırmacılar ve uygulamacılar tarafından eleştirilen bir uygulamadır (Beardsley, 2008). Okul yöneticileri tarafından sıkça bahsedilen bir başka uygulama ise mevcut yılda elde edilen okul ortalamasının bir önceki yılda elde edilen okul ortalaması ile karşılaştırılıp okulun daha başarılı olup olmadığına karar verilmesidir. Öğrencilerin başarılarını etkileyen okul dışı birçok değişim faktörü bulunduğu düşünüldüğünde sadece ortalama puanlara bakarak okul başarısını değerlendirmenin doğru olmadığı söylenebilir. Ayrıca bu uygulamalarda öğrencilerin öğretim dönemi başındaki durumları göz ardı edilerek sadece öğretim dönemi sonundaki başarısı ele alınmaktadır. Öğrencilerin öğrenme yaşantıları, sosyo-ekonomik durumları ve bireysel farklılıkları dolayısıyla öğretim dönemine farklı seviyelerde başlayabilecekleri düşünüldüğünde bu farklılıkların kontrol altına alınıp her bir öğrencinin dönem başından dönem sonuna kadar ne kadar gelişim gösterdiğinin belirlenmesi okul başarısını belirlemek adına daha doğru bir yaklaşım olacaktır. Son yıllarda okul ve/veya öğretmen başarısını değerlendirmede öğrencilerin başarılarındaki değişim miktarını (gelişimini) temel alan birçok yöntem geliştirilmiştir. Bu yöntemlerin başında katma-değerli değerlendirme (value-added assessment) sistemleri gelmektedir. Katma-değerli değerlendirme (KDD) yaklaşımı öğrenci ve okullara özgü farklılıkları kontrol altına alıp her bir öğrencinin bir sınavdan diğerine ne kadar gelişim gösterdiğine bakarak bu gelişimi okullara ya da öğretmenlere atfedebilme imkânı sağlayan nesnel bir değerlendirme sistemidir.

Katma-Değerli Değerlendirme

Yirminci yüzyılın başlarından itibaren okul etkililiğine ilişkin çalışmalar bulunmakla beraber 20. yüzyılın son yıllarında ve 21. yüzyılın başlarında okul etkililiği ve okulların hesap verebilirliği konularına giderek artan bir ilgi oluşmuştur. Özellikle 2001 yılında A.B.D.'de kabul edilen 'No Child Left Behind Act' (Hiçbir Çocuk Geride Kalmaması Hareketi) ile 1900'lerin başından beri eyalet düzeyinde yapılan sınavların (Kelly, 1916) tüm eyaletler tarafından yapılarak eyaletlerin öğrencilerin akademik başarılarını ölçmesi hedeflenmiştir (Beardsley, 2008). Bu eyalet çapında değerlendirmenin ilk uygulamaları, öğrencilerin mevcut durumuna odaklanmıştır. Mevcut durum yaklaşımı, farklı sınıf düzeyindeki öğrencileri zaman içindeki tek bir noktada karşılaştırmaktaydı (Doran ve Izumi, 2004). Eğitimciler, bir noktada elde edilen test puanının okulun öğrenci performansı üzerindeki etkilerini tahmin etmede kullanışlı bir yol olmadığını kabul etmektedir. Mevcut durum yöntemleri okulların etkililiğini değerlendirirken sosyo-ekonomik faktörleri dikkate almamaktadır. Bu yöntemler okul hesap verebilirlik sisteminin kalbinde yer almasına rağmen, bu yöntemlerin geçersiz olması ve okul karşılaştırmaları amacıyla kullanılmamasının birçok gerekçesi vardır. Bunların başında öğrencilerin okullara farklı arka planlar ile gelmesi ve öğrencilerin okullara rastgele atanmamış olmasıdır (Doran ve Izumi, 2004). Dolayısıyla bu durum sosyo-ekonomik durum açısından avantajlı ve dezavantajlı okulların karşılaştırılmasında bir adaletsizlik doğurmaktadır. Başka bir gerekçe ise mevcut durum yöntemlerinin eklenerek artan bir başarıya bağlı olmasıdır. Yani öğrencilerin şu anki başarılarının sadece mevcut okullarına değil daha önceki okullarından elde edilen öğrenmelere de dayanıyor olmasıdır. Nitekim Kerbow (1996) okul değiştiren öğrencilerin devam ettikleri yeni okullarının sistemini ve başarısını etkilediklerini ifade etmektedir. Ayrıca Temple ve Reynolds (1999) okul değiştiren öğrencilerin başarılarının düştüğünü belirtmektedirler. Dolayısıyla öğrenim görülen son okul öğrencinin başarısından ya da başarısızlığından sorumlu tutulamaz. Mevcut duruma dayalı hesap verebilirlik sistemlerinin yanlış olduğu ve okul kalitesinin belirlenmesinde hatalı kararlara yol açtığı alanyazında sıklıkla vurgulanmaktadır (Drury ve Doran, 2003). Drury ve Doran'a (2003) göre bu yöntemin geçersiz ve yanıltıcı olmasının arkasında birçok sebep bulunmaktadır. Bunlar aile, toplum, arka plan ve önceki okulların birleştirilmiş etkisini yansıtmaması, öğrencinin daha önceki okullarının kümülatif başarısını yansıtmaması ve belirli sınıf ve ders düzeyindeki başarıları yansıtmada yetersiz olmasıdır. Bu yöntemin eksikliklerinin fazla olması, hesap verebilirlik sisteminde okul etkililiğini

değerlendirmenin alternatif bir yolunun ortaya çıkmasına zemin oluşturmuştur. Bu yeni yöntem öğrencinin okulda yıl boyunca gösterdiği “gelişme” üzerine odaklanmaktadır. Doran ve Izumi (2004) farklı sınıf düzeyindeki öğrenci gruplarının bilgi düzeyinin ne durumda olduğunu göstermek yerine, “değişim” anlamına gelen “öğrenme” açısından öğrencilerin zaman içindeki ilerlemelerini bir ölçmeden diğerine ölçmenin daha mantıklı olduğunu belirtmektedirler.

Araştırmacılar etkili okulları belirlemede öğrencilerin bireysel başarılarının boylamsal olarak kullanılmasına izin veren katma-değerli analiz (value-added analysis) yöntemlerini geliştirmişlerdir. Tekwe ve diğerlerinin (2004, s.12) tanımına göre “katma-değer bir yıldan sonrakine okul ve/veya öğretmenin yaptığı katkıyı ifade eden bir terimdir ve bu ölçüm performans değerlendirme sisteminin temelinde kullanılır.” Katma-değerli analiz (KDA) yönteminin geliştiricileri bu yöntemin tek bir noktadaki başarı yerine başarıdaki değişimi dikkate alması bağlamında mevcut durum yönteminden daha adil olduğunu ileri sürmektedir (Hanushek, 1972; Sanders, Saxton ve Horn, 1997; Tekwe vd., 2004). KDA yöntemlerinin temel amacı öğrencilerinin gelişimine dayanarak okulların etkililiğine karar vermektir. Öğrenci gelişimleri öğrencilerin bir yıl arayla aldıkları testlerin farkından hesaplanan kazanç puanları ile belirlenmektedir. Her öğrenci için elde edilen bu kazanç puanları katma-değerli değerlendirme (KDD) sistemlerinde kullanılmaktadır. Kısacası KDD sistemi iki temel fikre dayanmaktadır: (1) her öğrencinin okullar içerisinde kümelenmiş olmasına ve (2) bir yıldan sonraki yıla performans değişimine dayanmasıdır. KDD sisteminin mevcut-durum yaklaşımına olan başka bir üstün yanı da test puanlarını etkileyebilen okul ve öğrenci değişkenlerinin etkisini kontrol edebilmesidir. Bu sayede öğrencilerin sosyokültürel özelliklerinin ve önceki yaşantılarının etkisinin azaltılması sağlanmış olur.

Katma-Değerli Değerlendirme Modelleri

Meyer ve Dokumacı (2010) KDD sisteminin yüksek kalitede katma-değerli puanlar üretebilmesinin öğrenci başarısını ölçen çıktının uygun ve kaliteli olmasına, boylamsal verinin mevcut olmasına ve KDD modelinin desenine bağlı olduğunu belirtmektedir. KDD modelleri genel olarak, öğretmen ve okul gibi eğitim birimlerinin öğrencilerin başarılarına etkisini araştırmada kullanılan istatistiksel modelleri temsil etmektedir. Bu modeller öğrencilerin zaman içindeki gelişimlerini okul ve öğrenci özelliklerinin bir fonksiyonu olarak inceleme şansı verir. KDD modelleri son yıllarda araştırmacılar ve eğitim politikacıları tarafından çok ilgi görmektedir. Kazanç puanlarına dayanan bu modellerin bazıları rastgele etkilere bağlı olan hiyerarşik modeller olarak geliştirilmiştir. Okulların sabit dağılım gösterdiğini varsayan ve regresyon yöntemlerini içeren mevcut durum yaklaşımının aksine KDD analizleri için kullanılan hiyerarşik modeller okulların rastgele olduğunu varsaymaktadır. Mevcut durum analizlerinde olduğu gibi başka değişkenlerin öğrenci başarısı ile okul etkililiği arasındaki ilişkiye karışmasını önlemek için hiyerarşik modellerde istatistiksel kontrol uygulanabilmektedir (Sanders, 2000).

Hanushek (1972) KDD modellerini okul değerlendirme sistemlerinde kullanılmasını öneren ilk araştırmacıdır. Sanders ve diğerleri (1997) ise Tennessee KDD sistemini (TVAAS) oluşturarak bu KDD yaklaşımını eyalet bazında uygulayabilen ilk araştırmacılar olmuşlardır. McCaffrey, Lockwood, Koretz ve Hamilton'un (2003) yapmış olduğu geniş çaplı KDD araştırmasına göre ilk KDD modelleri sabit etkiler (Hanushek, 1972; Murnane, 1975) varsayımı yaparken daha sonraları geliştirilen KDD modelleri (hiyerarşik modeller ve TVAAS modeli) rastgele etkiler varsayımında bulunmaktadır. Başlıca KDD modelleri basit sabit etkiler modeli (simple fixed effects model), hiyerarşik modeller (hierarchical linear models) ve tabakalı karışık etkiler modelini (layered mixed effects model) içermektedir. McCaffrey, Lockwood, Koretz, Louis ve Hamilton (2004) KDD yaklaşımında öğretmen ya da okul etkilerini boylamsal veriler yoluyla analiz ederken tüm öğrenci sonuçlarının tam ortak dağılımını doğrudan modelleyen “çok değişkenli modeller” olarak adlandırdıkları genel bir modelin takip edildiğini ayrıca bu tür verilerin katma-değerli değerlendirme sisteminde analiz edilebilmesini sağlayan alternatif model yaklaşımlarının da olduğunu belirtmektedirler. Bu alternatif yaklaşımlar mevcut puanın önceki puanlar üzerine regresyonuna dayanan ortak değişken düzeltme modellerini (covariate adjustment models), yılsonu kazanımlarını sonuç olarak ele alan kazanç puanları (gain scores) modelini ve çapraz-

sınıflandırılmış modelleri (cross-classified models) içermektedir. Bu modellerin çoğu aynı öğrenciye ait eşitlenmiş ve standartlaştırılmış birden fazla test puanına sahip olmayı gerektirmektedir (Ballou, Sanders ve Wright, 2004; Doran ve Cohen, 2005). Eşitleme varsayımı yapan modellerin yanı sıra eşitleme varsayımı gerektirmeyen KDD modelleri de son yıllarda alanyazında yerini almıştır (Mariano, McCaffrey ve Lockwood, 2010). Eğitim kurumlarında uygulanan ardışık sınavların çoğunlukla eşitlenmemiş ya da aynı ölçek üzerinde standartlaştırılmamış olması sınavlar arası eşitleme varsayımı gerektiren KDD modellerinin uygulanmasını engellemektedir. Eşitlenmemiş sınavların bu modeller ile analiz edilmesi yanlış ve hatalı sonuçlar elde edilmesine yol açmaktadır (Briggs ve Domingue, 2013; Briggs ve Weeks, 2009). Bu modellerin sınırlılığını ortadan kaldıran ve eşitleme varsayımı gerektirmeyen Genelleştirilmiş Süreklilik modeli (Generalized Persistence [GP] model) Mariano ve diğerleri (2010) tarafından geliştirilmiştir.

Araştırmanın Amacı ve Önemi

Türkiye’de MEB tarafından 23 Ekim – 31 Ekim 2017 tarihleri arasında 12 ilde ve 132 okulda pilot uygulaması gerçekleştirilen “Öğretmen Performans Değerlendirme” çalışması yürütülmüş ve bu çalışmaya ülke geneli yetkili sendika karşı eylem kararı almıştır. Yani MEB tarafından öğretmen değerlendirme kapsamında yapılan bu çalışma yetkili sendikanın eylem kararı ile işlemez hale getirilmiştir. Yetkili sendikanın eylem kararına gerekçe olarak öğretmen performansının; öğrencilerin, idarecilerin, velilerin puanlarına ve öğretmenlerin gireceği bir sınavdaki başarı puanlarına dayandırılmasıdır. Bu bağlamda bu çalışma ile önerilen yapı, dünyada birçok ülkede uygulanmakta olan öğretmen veya okul değerlendirme modelinin Türkiye’de (özelde Şanlıurfa’da) uygulanma durumunu incelemekle kalmayıp, öğrenci başarı gelişimine göre öğretmen veya okulun değerlendirilebileceği bir modeli hayata geçirmeye çalışmaktadır.

Türkiye’de okul etkililiği üzerine yapılan çalışmalar incelendiğinde okul etkililiğinin daha çok ölçek, anket ve görüşme gibi veri toplama araçları üzerinden belirlenmeye çalışıldığı gözlenmiştir (Turhan, Şener ve Gündüzalp, 2017). Okul etkililiğinin bu tür ölçme araçları kullanılarak yönetici, öğretmen, öğrenci ya da veli gibi okul paydaşlarının beyanları üzerinden ölçülmesi bu fenomenin ölçülmesine öznellik katmaktadır. Bu doğrultuda daha nesnel ölçütlerin kullanılmasına ihtiyaç vardır. Nitekim mevcut eğitim sisteminde öğrencilerin başarılarından okulları sorumlu tutacak nesnel ve bilimsel bir yöntemin olmaması, katma-değerli değerlendirme (KDD) yaklaşımının birçok avantaj sunuyor olması ve bu yaklaşımın Türkiye’deki eğitim politikacıları tarafından daha önce hiç denenmemiş olması bu konuda çalışma yapılması ihtiyacı doğurmuştur. Bu ihtiyaçlar doğrultusunda gerçekleştirilen bu çalışmanın Şanlıurfa’daki öğrencilerin öğrenme çıktılarının izlenmesi ve bu çıktıların nesnel bir okul değerlendirmesinde kullanılması açısından önemli olduğu düşünülmektedir.

Bu bağlamda çalışmanın amacı gelişmiş ülkelerde etkin olarak kullanılan katma-değerli değerlendirme modelini; MEB’in kendine bağlı eğitim kurumlarını değerlendirmesinde yardımcı olacak bir formda tasarlayarak bu modelin prototipini Şanlıurfa ili verileri üzerinde denemektir. Bu çalışmanın diğer bir amacı da, eğitim politikaları ve/veya müdahale stratejileri geliştirilmesine yardımcı olmak ve eğitimde kalitenin artırılmasına yönelik bir dayanak noktası ortaya koyabilmektir. Bu çalışma ile okul başarısının değerlendirilmesinde geçerli, güvenilir ve kullanışlı bilgiler sunan katma-değerli değerlendirme sisteminin tanıtılması ve uygulanması amaçlanmaktadır. Bu bağlamda aşağıdaki sorulara cevap aranmaktadır:

1. Şanlıurfa ilinde bulunan ortaokul kurumlarının etkililik bağlamında sıralamaları nasıldır?
2. Şanlıurfa ilinde bulunan ortaokul kurumlarının tek sınava dayalı başarı sıralamaları ile katma-değer puan sıralamaları örtüşmekte midir?
3. Şanlıurfa ilinde bulunan ortaokul kurumlarından elde edilen puanlara göre katma-değerli değerlendirme modeli işlevsel midir?

Araştırma neticesinde elde edilen sonuçlar katma-değerli değerlendirmenin ön uygulamalarının Şanlıurfa bağlamında sonuçlarını kapsayacak ve bu yaklaşımın kullanılabilirliğine ilişkin ön bilgiler sunacaktır.

Yöntem

KDD analizleri ile okul etkililiğini belirlemek adına her öğrenciye ait en az iki ölçüm yapılması gerekmektedir. Bu çalışmada bir öğretim yılı içerisinde elde edilen iki standart test verisi kullanılarak bir KDM aracılığıyla verideki okullara ait katma-değer puanları hesaplanmıştır. Bu bağlamda 2015-2016 eğitim - öğretim yılında uygulanan Kasım ve Nisan aylarına ait TEOG (Temel Eğitimde Ortaöğretime geçiş) sınavı verileri kullanılmıştır. Analizler neticesinde elde edilen katma-değer puanlarına göre okul sıralamaları yapılmıştır. Şanlıurfa ilinde TEOG sınavına giren öğrencilere ait toplam puanlar kullanılmıştır. Örnekleme, Şanlıurfa ilinde her iki TEOG sınavına katılan tüm ortaokul ve öğrencileri kapsamakla beraber Şanlıurfa ilinde yer alan 64,852 öğrenci ve 539 ortaokuldan oluşmaktadır. Tüm okulların TEOG sınavı net ortalamaları incelendiğinde birinci sınava ait net ortalamaları 25.24 ile 69.52 arasında değişirken ikinci sınava ait net ortalamaları 27.70 ile 70.21 arasında değişim göstermektedir. Bu çalışmada kullanılan iki sınav verisi eşitlenmemiş olduğundan eşitleme varsayımı yapmayan KDM'lerden olan Genelleştirilmiş Süreklilik modeli (Generalized Persistence [GP] model) kullanılarak analizler gerçekleştirilmiştir. Raporlama sürecinde okul adları yerine okullara verilen kod numaraları kullanılmıştır.

Genelleştirilmiş Süreklilik Modeli. Mariano ve diğerlerinde (2010) gösterildiği üzere genelleştirilmiş süreklilik modeli çok değişkenli KDD modelleri (McCaffrey vd., 2004) grubunda yer almaktadır. Bu yaklaşım okul etkisini birden fazla yıl için aynı anda tahmin etmeye dayalıdır. Öğrencilerin yıl içerisinde okul ya da sınıf değiştirmeleri HLM'lerde istenen kümelenmiş yapıyı bozmaktadır. Bu tür daha karmaşık yapılarda kullanılacak çok değişkenli model aşağıda açıklanmaktadır.

Genelleştirilmiş süreklilik modeli, öğrencinin t yılına ait puanının tüm öğrenciler için genel t yılı ortalamasına, mevcut yılın ve geçmiş yıl öğretmenlerinin etkilerinin kümülatif toplamına ve testin bu yılındaki öğrenci için rastgele bir artık hata terimine bağlı olduğunu varsayar. y_{it} teriminin öğrenci i 'nin t yılındaki başarı puanını temsil ettiği düşünülürse bu puan için model:

$$y_{it} = \mu_t + \left(\sum_{g=1}^t \sum_{j=1}^{J_g} \phi_{igj} \theta_{g[jt]} \right) + \varepsilon_{it}$$

şeklinde verilir. Bu eşitlikte μ_t terimi o yıl için genel ortalamayı temsil ederken, ϕ_{igj} terimi öğrencinin g yılında öğretmen (ya da okul) j tarafından okutulduğu durumda 1 diğer durumlarda 0 olarak kodlanır böylelikle $\phi_{igj} \theta_{g[jt]}$ çarpımı mevcut yıl ve önceki yıllara ait öğretmen ya da okulun etkisini gösterir. Eşitlikteki ε_{it} de artık hata terimini temsil etmektedir. Bu modelde ε_{it} teriminin ortalaması sıfır ve yapılandırılmamış kovaryans matrisine sahip bir normal dağılıma sahip olduğu varsayılır ($\varepsilon_{it} \sim MVN(0, \Sigma)$). Mariano ve diğerleri (2010) katma-değerli değerlendirme sistemlerinde önceki yıllara ait okul ya da öğretmenlerin süreklilik etkisini hesaba katan tam süreklilik (complete persistence) ve değişken süreklilik (variable persistence) modellerinin bu model üzerinden elde edilebileceğini de göstermektedir. Bu çalışmada GP modeline ait model parametreleri ve okul etkileri R programında GPVam paketi (Karl, Yang ve Lohr, 2012) kullanılarak elde edilmiştir. Analizlerde maksimum olabilirlik kestirim yöntemi kullanılmıştır. Okul etkileri (katma-değer puanları) ampirik en iyi doğrusal yansız yordayıcı (the empirical best linear unbiased predictors [EBLUP]) olarak kestirilmiştir.

Bulgular

GPVam paketinde genel süreklilik (persistency) modelleri altında GP (generalized persistence), VP (variable persistence) ve CP (complete persistence) modelleri olmak üzere üç seçenek mevcuttur. Kullanılan testlerin eşitlenmemiş olması dolayısıyla çalışmada GP model kullanılmış ve bu modelden elde edilen bulgular raporlanmıştır. GP model kullanılarak elde edilen parametre tahminleri Tablo 1’de sunulmaktadır. Analiz edilen GP model sonucunda göreceli uyum değeri olan AIC değeri 500137 olarak elde edilmiştir. Ayrıca 539 okula ait katma-değer puanları da elde edilmiştir. Bu değerler I. ve II. sınav için ayrı ayrı hesaplanmaktadır. Bu değerler Şekil 1 ve 2’de grafiklerle gösterilmektedir.

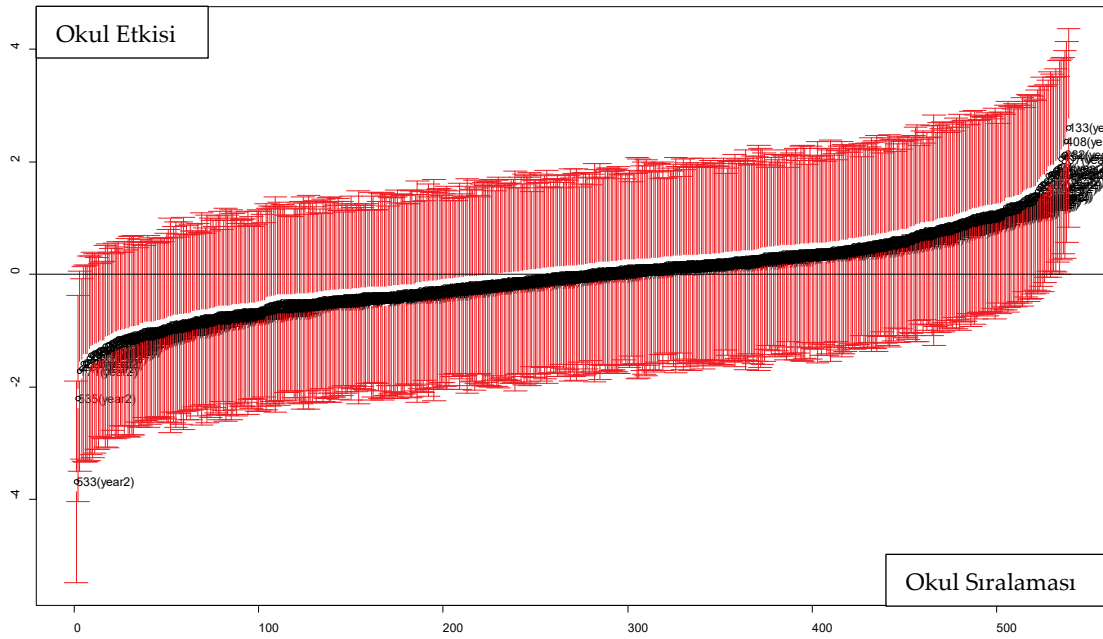
Tablo 1. GP Modelden Elde Edilen Parametre Tahminleri

	Katsayı	St. Hata
I. Yıl Okul etkisi	40.535	0.341
II. Yıl Okul etkisi	45.106	0.331

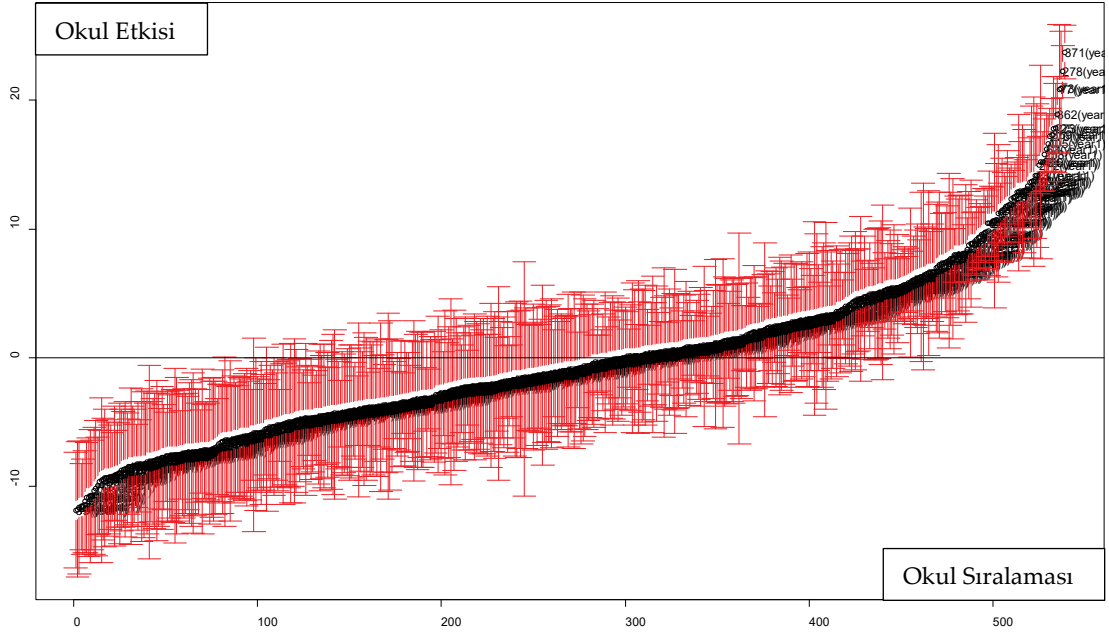
Tablo 2. GP Modelden Elde Edilen Kovaryans Değerleri

	I. Yıl	II. Yıl
I. Yıl	53.509	49.871
II. Yıl	49.871	48.126

Not. I. Yıl ile II. Yıla ait okul etkileri arası korelasyon değeri $r=.983$ olarak hesaplanmıştır.



Şekil 1. Okulların Birinci Yıla Ait Katma-Değer Puanları ve Güven Aralıkları



Şekil 2. Okulların İkinci Yıla Ait Katma-Değer Puanları ve Güven Aralıkları

Okul etkililiği değerlendirmesinde asıl bakılması gereken ilk sınavdan ikinci sınava öğrencilerin göstermiş oldukları değişimi yansıtan ikinci sınava ait katma-değer puanlarıdır. Tablo 3 ve 4'te öğrencilerin puanlarındaki artış ve azalışa göre elde edilmiş olan katma-değer puanlarını gösteren EBLUP değerleri sunulmaktadır. Bu tablolarda en yüksek puana sahip 20 okul ve en düşük puana sahip 20 okulun katma-değer puanları sunulmaktadır. Ayrıca bu okullara ait TEOG sınavı net ortalamaları ($\bar{X}_1 = \text{Kasım}$, $\bar{X}_2 = \text{Nisan}$) da bu tablolarda sunulmuştur.

En başarılı 20 okula ait değerlerin sunulduğu Tablo 3 incelendiğinde en yüksek katma-değer puanına sahip olan üç okulun sırasıyla 371, 278 ve 73 numaralı okullar olduğu görülmektedir. Bu okullara ait aritmetik ortalama değerleri incelendiğinde bu okulların genel olarak tüm okullar arasında yüksek aritmetik ortalamaya sahip okullar oldukları görülmektedir. Tablo 3 dikkatli incelendiğinde en yüksek ortalamaya sahip olan okulların (birinci sınavda 14. okul, ikinci sınavda 77. okul) en yüksek katma-değer puanına sahip olmadığı görülmektedir. Bu da hâlihazırda sadece okul ortalamalarına göre okul etkililiği değerlendirmenin katma-değerli değerlendirme yöntemine göre yapılan değerlendirmeden farklı sonuçlar sunacağını göstermektedir. Bir okula ait katma-değer puanı elde edilirken o okuldaki bütün öğrencilerin birinci sınavından ikinci sınavına olan değişimleri dikkate alınmaktadır. Katma-değer puanlarına göre yapılan sıralamaların yanında birinci sınav ve ikinci sınava ait sıralamalar da Tablo 3'te sunulmuştur. Bu sıralamalar genel olarak tutarlılık gösterse de farklı sıralamaya sahip okullar göze çarpmaktadır. Bu üç değere göre yapılan okul sıralamaları arasındaki korelasyon değerleri Spearman Rho formülü ile hesaplanmıştır. Katma-değer puanlarına göre yapılan sıralama ile birinci sınavın aritmetik ortalamasına göre yapılan sıralama arasındaki korelasyon .648 olarak bulunurken, katma-değer puanlarına göre yapılan sıralama ile ikinci sınavın aritmetik ortalamasına göre yapılan sıralama arasındaki korelasyon .418 olarak hesaplanmıştır.

Tablo 3. En Başarılı 20 Okula Ait Hesaplanan Katma-Değer Puanları ve Aritmetik Ortalamaları

Başarı sırası	Okul no	EBLUP*	St. Hata	\bar{X}_1	Başarı sırası (\bar{X}_1)	\bar{X}_2	Başarı sırası (\bar{X}_2)
1.	371	23.730	1.259	67.459	3.	68.071	4.
2.	278	22.210	1.239	65.465	5.	66.818	5.
3.	73	20.906	2.700	66.833	4.	68.646	3.
4.	77	20.847	3.041	67.708	2.	70.208	1.
5.	362	18.885	1.794	63.087	6.	63.556	14.
6.	273	17.825	2.108	61.679	7.	63.750	13.
7.	125	17.763	2.488	60.609	10.	66.707	6.
8.	359	17.319	1.709	60.941	9.	62.086	18.
9.	276	17.195	1.301	60.183	11.	61.687	19.
10.	405	16.617	1.331	59.121	14.	61.583	21.
11.	63	16.174	1.695	58.888	15.	61.673	20.
12.	258	15.759	1.681	57.501	20.	62.184	17.
13.	426	15.221	1.411	57.703	17.	60.198	24.
14.	484	15.218	4.567	69.524	1.	69.762	2.
15.	272	14.952	1.364	57.590	18.	59.658	28.
16.	237	14.143	3.376	59.305	13.	64.213	11.
17.	43	14.136	2.952	57.327	21.	63.794	12.
18.	199	13.650	3.991	61.666	8.	65.454	7.
19.	213	13.646	3.599	59.388	12.	64.477	9.
20.	351	13.605	2.120	57.561	19.	58.354	32.

*EBLUP = Empirical Best Linear Unbiased Predictors (ampirik en iyi doğrusal yordayıcı)

En başarısız 20 okula ait değerlerin sunulduğu Tablo 4 incelendiğinde en düşük katma-değer puanına sahip olan üç okulun sırasıyla 26, 18 ve 39 numaralı okullar olduğu görülmektedir. Tablodaki aritmetik ortalama değerleri incelendiğinde bu okullardan daha düşük aritmetik ortalamaya sahip okullar olduğu gözlenmektedir. En başarılı okullar listesindeki katma-değer puanları ile okul ortalamaları arası fark burada da gözlenmektedir. Bu da hâlihazırda sadece okul ortalamalarına göre okul etkililiği değerlendirmenin katma-değerli değerlendirme yöntemine göre yapılan değerlendirmeden farklı sonuçlar sunacağını göstermektedir. Katma-değer puanlarına göre yapılan sıralamaların yanında birinci sınav ve ikinci sınava ait sıralamalar da Tablo 4'te sunulmuştur. Bu sıralamalar genel olarak tutarlılık gösterse de farklı sıralamaya sahip okullar göze çarpmaktadır. Bu üç değere göre yapılan okul sıralamaları arasındaki korelasyon değerleri Spearman Rho formülü ile hesaplanmıştır. Katma-değer puanlarına göre yapılan sıralama ile birinci sınavın aritmetik ortalamasına göre yapılan sıralama arasındaki korelasyon .564 olarak bulunurken, katma-değer puanlarına göre yapılan sıralama ile ikinci sınavın aritmetik ortalamasına göre yapılan sıralama arasındaki korelasyon .397 olarak hesaplanmıştır.

Tablo 4. En Başarısız 20 Okula Ait Hesaplanan Katma-Değer Puanları ve Aritmetik Ortalamaları

Başarı sırası	Okul no	EBLUP*	St. Hata	\bar{X}_1	Başarı sırası (\bar{X}_1)	\bar{X}_2	Başarı sırası (\bar{X}_2)
520.	190	-9.257	2.465	30.450	498.	33.452	515.
521.	275	-9.321	2.641	31.238	478.	31.690	528.
522.	300	-9.329	2.731	29.271	509.	33.880	507.
523.	137	-9.345	3.251	28.083	524.	33.241	521.
524.	25	-9.437	2.872	28.214	522.	34.464	500.
525.	311	-9.518	3.880	27.361	533.	29.375	536.
526.	185	-9.645	2.444	29.783	503.	33.333	517.
527.	437	-9.949	2.292	28.810	515.	34.166	502.
528.	70	-10.146	3.194	27.706	531.	31.547	530.

Tablo 4. Devamı

Başarı sırası	Okul no	EBLUP*	St. Hata	\bar{X}_1	Başarı sırası (\bar{X}_1)	\bar{X}_2	Başarı sırası (\bar{X}_2)
529.	138	-10.327	2.996	27.766	529.	31.933	526.
530.	4	-10.675	2.586	27.716	530.	32.725	524.
531.	10	-10.731	3.140	27.386	532.	30.477	533.
532.	224	-10.849	2.799	28.500	517.	30.361	534.
533.	395	-10.864	2.613	26.569	536.	33.578	513.
534.	495	-11.181	2.423	28.022	526.	31.590	529.
535.	314	-11.565	2.260	28.116	523.	31.053	532.
536.	389	-11.586	2.030	28.502	516.	31.248	531.
537.	26	-11.641	3.140	25.947	537.	29.696	535.
538.	18	-11.787	3.194	26.785	534.	27.698	239.
539.	39	-11.838	2.731	28.000	527.	28.567	538.

* EBLUP = Empirical Best Linear Unbiased Predictors (ampirik en iyi doğrusal yansız yordayıcı)

Verilerden elde edilen bulgulara göre ise sistemin işlerliğinde herhangi bir problem rastlanılmamıştır. Yani merkezi bir sınav olan TEOG verileri ile katma-değerli değerlendirme sisteminin işler olduğu söylenebilir.

Sonuç ve Tartışma

Bu çalışmada okul etkililiğinin değerlendirmesinde sıklıkla kullanılan yaklaşımlardan biri olan katma-değerli değerlendirme sistemi tanıtılmış ve gerçek bir veri üzerinden bu yaklaşımın uygulaması gösterilmiştir. Katma-değerli değerlendirme sistemi, öğrencilerin gelişimlerini öğretmene ya da okula atfederek öğretmen/öğrenci ya da okul/öğrenci kümelenmiş veri yapısını kullanarak bir katma-değer puanı hesaplamaktadır. Bu çalışmada da Şanlıurfa'daki okullara o okullarda öğrenim görmekte iken iki farklı TEOG sınavına girmiş öğrencilerin puanlarına göre birer katma-değer puanı hesaplanarak okul başarı sıralaması yapılmıştır. Aynı zamanda okullardaki öğrencilerin sadece puan ortalamalarına bakılarak bir sıralama yapılmış ve katma-değerli değer puanlarına göre yapılan sıralamadan farklılıklar gösterdiği görülmüştür. Bulgularımız ışığında genel bir bakışla yüksek ortalamaya sahip okulların katma-değer puanlarının yüksek olduğu, düşük ortalamaya sahip okulların da genel olarak düşük katma-değer puanına sahip olduğu görülmüştür. Her ne kadar sonuçlar birbirleriyle tutarlı gibi görünse de katma-değerli değerlendirme yaklaşımının daha doğru sonuçlar ortaya koyduğu ifade edilmektedir. Bulgulara ve literatürdeki çalışmalara (Hanushek, 1972; McCaffrey vd., 2003; Sanders vd., 1997; Tekwe vd., 2004) dayanarak tek bir noktada elde edilen ortalama puanlara göre okulların performansları hakkında yargılarda bulunmanın yanlış değerlendirmelere yol açabileceği düşünülmektedir. Okul ortalamasına dayalı olarak okulların başarılı olup olmadığı yargısına varırken okula ve öğrenciye ait değişkenlerin hesaba katılmaması geleneksel yöntemin dezavantajları arasında yer alırken bu tür değişkenleri KDM'ler aracılığıyla kontrol ederek daha geçerli sonuçlar elde etmek mümkün görünmektedir. Bu çalışmada öğrencilere ya da okullara ait ek değişkenler olmadığı için model analizleri yordayıcı değişkenler olmadan analiz edilmiştir. Bu durum araştırmanın sınırlılığı olarak belirtilebilir.

KDD sistemi okul etkililiğinin değerlendirilmesine güçlü bir alternatif olma potansiyeline sahiptir. Ayrıca öğretmen hareketliliğinin minimuma indirildiği varsayıldığında, öğretmen etkililiğinin değerlendirilmesinde de güçlü bir alternatif olabilecektir. Bu değerlendirme sisteminin öğretmen değerlendirmede kullanılması, Türkiye gündemini meşgul eden 'öğretmenlere puan verme' problemini (Tedm Raporu, 2018) ortadan kaldıracak olduğu düşünülmektedir.

Okul yöneticilerinin öğretmenleri değerlendirirken somut araçlara sahip olmadıkları düşünülürse, KDD sistemi yöneticilere nesnel bir biçimde değerlendirme yapma olanağı sağlayabilir. Bazı öğretmenlerin yöneticileri tarafından sınıfta gözlemlenerek değerlendirilmelerini uygun bulmadıkları ve yöneticilerin de bu konuda eğitilmiş olmadıkları için öğretmen değerlendirmesini ihmal ettikleri belirtildiğinden (Bozkuş, 2018) KDD ile uygulamadaki bu olumsuzluklar giderilebilir.

2001 yılında Amerika'da çıkarılan "Hiçbir Çocuk Geride Kalmasın" (No Child Left Behind [NCLB]) yasasıyla birlikte birçok eyalet kendi sınırları içerisindeki okullarda öğrenim gören öğrencilere her yıl standartlaştırılmış sınavlar uygulayarak öğrencilerin başarılarını ölçmekle ve beklenen başarının altında kalan öğrencileri geliştirmekle yükümlüdür (McCaffrey vd., 2004). Bu eyaletlerin çođu kendi katma-deđerli deđerlendirme sistemlerini kurarak öğrencilere ait boylamsal veriler üzerinden okulları ve öğretmenleri deđerlendirmektedir. Yeterli başarıyı gösteren okullara teşvik verilirken beklenen gelişimi gösteremeyen okullara da yaptırımlarda bulunmaktadır. Buna benzer bir uygulamanın Milli Eğitim Bakanlığı yönetiminde Türkiye'deki illerde ve okullarda kullanılmasının ülke genelinde öğrenciler arası başarı farkını ortadan kaldırmak için dođru bir uygulama olacađı düşünölmektedir. Bu çalışma KDD sisteminin Türkiye'deki okullara uygulanabilirliğini göstermek adına iyi bir örnek teşkil etmektedir.

Katma-deđerli deđerlendirme sisteminde okulların ya da öğretmenlerin performanslarını deđerlendirmede kullanılan katma-deđer puanlarını elde etmek için alanyazında birçok istatistiksel model önerilmiştir. Bu modellerin okulların hesap verilebilirliği üzerinde kullanımı son yıllarda artmaya başladığı bunun neticesinde öğretmenler ve okulların sıralamaya tabi tutulduđu, bu sıralamalar esas alınarak ödöl ve yaptırımlarda bulunduğu Amerikan İstatistik Kurumu (American Statistical Association, 2014) tarafından altı çizilen bir nokta olmuştur. American Statistical Association (2014) ve birçok araştırmacı (Baker vd., 2010; Ballou ve Springer, 2015; Beardsley, 2008; Goldhaber, 2015; Johnson, 2015) sadece bu modellerden elde edilen sonuçların okullar ve öğretmenler hakkında önemli kararlar alınmasında kullanılmasını önermektedir. Araştırmacılar katma-deđerli deđerlendirme yaklaşımının okul ve öğretmen deđerlendirmesinde asıl karar verici olarak deđil de karar vermeye yardımcı olarak kullanılmasını önermektedir. Bu çalışma da bu bağlamda Mili eğitim Bakanlığı için bir öneri mahiyetinde deđerlendirilebilir.

Kaynakça

- Adams, E. (1993). School's out!: New initiatives for British school grounds. *Children's Environments*, 10(2), 180-191.
- American Statistical Association. (2014). *ASA statement on using value-added models for educational assessment*. Alexandria, VA.
- Baker, E. L., Barton, P. E., Darling-Hammond, L., Haertel, E., Ladd, H. F., Linn, R. L. ... ve Shepard, L. A. (2010). *Problems with the use of student test scores to evaluate teachers* (EPI Briefing Paper #278). Washington: Economic Policy Institute.
- Ballou, D. ve Springer, M. G. (2015). Using student test scores to measure teacher performance: Some problems in the design and implementation of evaluation systems. *Educational Researcher*, 44(2), 77-86.
- Ballou, D., Sanders, W. L. ve Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, 29(1), 37-66.
- Barnard, C. (1938). *The functions of the executive*. Cambridge, MA: Harvard University Press.
- Beardsley, A. (2008). Methodological concerns about the education value-added assessment system. *Educational Researcher*, 37(2), 65-75.
- Botha, R. (2010). School effectiveness: Conceptualising divergent assessment approaches. *South African Journal of Education*, 30(4), 605-620.
- Bozkuş, K. (2018). *Öğretmenlerin mesleki gelişiminde dinamik yaklaşımın uygulanması: Bir eylem araştırması* (Yayımlanmamış doktora tezi). Anadolu Üniversitesi, Eskişehir.
- Briggs, D. C. ve Domingue, B. (2013). The gains from vertical scaling. *Journal of Educational and Behavioral Statistics*, 38(6), 551-576.
- Briggs, D. C. ve Weeks, J. P. (2009). The sensitivity of value-added modeling to the creation of a vertical score scale. *Education Finance and Policy*, 4(4), 384-414.
- Creemers, B. P. ve Kyriakides, L. (2008). *The dynammics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. London: Routledge.
- Doran, H. ve Izumi, L. T. (2004). *Putting education to the test: A value-added model for California*. San Francisco, CA: Pacific Research Institute.
- Doran, H. C. ve Cohen, J. (2005). The confounding effect of linking bias on gains estimated from value-added models. R. Lissitz (Ed.), *Value-added models in education: Theory and application* içinde (s. 80-110). Maple Grove, MN: JAM Press.
- Drury, D. ve Doran, H. C. (2003). The value of value-added analysis. *Policy Research Brief for National School Board Association*, 3(1), 1-4.
- EARGED. (2010). *İlköğretim okulları başarı göstergeleri*. Ankara: MEB Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı.
- Goldhaber, D. (2015). Exploring the potential of value-added performance measures to affect the quality of the teacher workforce. *Educational Researcher*, 44(2), 87-95.
- Hanushek, E. A. (1972). *Education and race*. Lexington, MA: Lexington Books.
- Hesapçioğlu, M. (2006). Eğitim kurumlarında kalite olgusu ve kalite güvence sistemleri. *M.Ü. Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi*, 23(1), 143-160.
- Johnson, S. M. (2015). Will VAMS reinforce the walls of the egg-crate school?. *Educational Researcher*, 44(2), 117-126.
- Karip, E. ve Köksal, K. (1996). Etkili eğitim sistemlerinin geliştirilmesi. *Eğitim Yönetimi*, 2(2), 245-257.
- Karl, A. T., Yang, Y. ve Lohr, S. (2012). GPvam: Maximum likelihood estimation of multiple membership mixed models used in value-added modeling. R Package Version 2.0-0 [Computer software]. <https://cran.r-project.org/web/packages/GPvam/GPvam.pdf> adresinden erişildi.

- Kelly, F. J. (1916). The Kansas Silent Reading Tests. *Journal of Educational Psychology*, 7(2), 63-80.
- Kerbow, D. (1996). Patterns of urban student mobility and local school reform. *Journal of Education for Students Placed at Risk*, 1(2), 147-169.
- Mariano, L. T., McCaffrey, D. F. ve Lockwood, J. R. (2010). A model for teacher effects from longitudinal data without assuming vertical scaling. *Journal of Educational and Behavioral Statistics*, 35(3), 253-279.
- McCaffrey, D., Lockwood, J. R., Koretz, D. ve Hamilton, L. (2003). *Evaluating value-added models for teacher accountability*. Washington, DC: RAND.
- McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A. ve Hamilton, L. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, 29(1), 67-101.
- Meyer, R. H. ve Dokumaci, E. (2010). *Value-added models and the next generation of assessments*. Princeton, NJ: Center for K-12 Assessment & Performance Management. www.k12center.org/rsc/pdf/MeyerDokumaciPresenterSession4.pdf adresinden erişildi.
- Milli Eğitim Bakanlığı. (2019). *Milli eğitim istatistikleri örgün eğitim*. Ankara: Milli Eğitim Bakanlığı. http://sgb.meb.gov.tr/meb_iys_dosyalar/2019_09/30102730_meb_istatistikleri_organ_egitim_2018_2019.pdf adresinden erişildi.
- Murnane, R. J. (1975). *The impact of school resources on the learning of inner city children*. Cambridge, MA: Ballinger Publishing Co.
- Rowan, B. (1985). The assessment of school effectiveness. R. Kyle (Ed.), *Reaching for excellence: An effective schools sourcebook*. Washington, DC: National Institute of Education.
- Sanders, W. L. (2000). Value-added assessment from student achievement data: Opportunities and hurdles. *Journal of Personnel Evaluation in Education*, 14(4), 329-339.
- Sanders, W. L., Saxton, A. M. ve Horn, S. P. (1997). The Tennessee Value-Added Educational Assessment System (TVAAS): A quantitative, outcomes-based approach to educational assessment. J. Millman (Ed.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* içinde (s. 137-162), Thousand Oaks, CA: Corwin Press.
- Tedmem Raporu. (2018). *Öğretmen performans değerlendirme ve aday öğretmenlik iş ve işlemleri yönetmelik taslağı üzerine değerlendirmeler*. <https://tedmem.org/download/ogretmen-performans-degerlendirme-aday-ogretmenlik-is-islemleri-yonetmelik-taslagi-uzerine-degerlendirmeler?wpdmdl=2641> adresinden erişildi.
- Tekwe, C. D., Carter, R. L., Ma, C., Algina, J., Lucas, M. E., Roth, J. ve ... Resnick, M. B. (2004). An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics*, 29(1), 11-36.
- Temple, J. A. ve Reynolds, A. J. (1999). School mobility and achievement: Longitudinal findings from an urban cohort. *Journal of School Psychology*, 37(4), 355-377.
- Turhan, M., Şener, G. ve Gündüzalp, S. (2017). Türkiye'de okul etkililiği araştırmalarına genel bir bakış. *Turkish Journal of Educational Studies*, 4(2), 103-151.
- UNESCO. (2019). *Combining data on out-of-school children, completion and learning to offer a more comprehensive view on SDG 4* (Information Paper No. 61). Montreal, Quebec: United Nations Educational Scientific and Cultural Organization. <http://uis.unesco.org/sites/default/files/documents/ip61-combining-data-out-of-school-children-completion-learning-offer-more-comprehensive-view-sdg4.pdf> adresinden erişildi.
- Yükçü, S. ve Atağan, G. (2009). Etkinlik, etkililik ve verimlilik kavramlarının yarattığı karışıklık. *Atatürk Üniversitesi İktisadi ve İdari Bilimler Dergisi*, 23(4), 1-13.