



İngilizce Öğretmenliği Programında Verilen Ölçme Değerlendirme Eğitiminin Hedeflenen, Uygulanan, Öğrenilen ve Değerlendirilen Programları Arasındaki Uyum

Serpil Tekir ¹

Öz

Son zamanlarda yapılan araştırmalar, ulusal bağlamda İngilizce öğretmenliği lisans programlarında verilen ölçme değerlendirme eğitimi ile ilgili bir sorunu işaret etmektedir. Bu sorunun nedenlerinden biri, ölçme değerlendirme eğitimin farklı programları arasındaki uyumun düşük olması olabilir çünkü öğretmen eğitimi konusunda çalışmalar yapan çok sayıda araştırmacı, program uyumun, öğretmen adaylarının etkili öğrenme ve uygulamaları için gerekli bir koşul olduğunu ileri sürmektedir. Bu nedenle, araştırmacı bu çalışmada, İngilizce öğretmenliği lisans programında öğretmen adaylarına verilen ölçme değerlendirme eğitiminin hedeflenen, öğretilen, değerlendirilen ve öğrenilen programları arasındaki uyumu incelemiştir. Bu çalışma hem nitel hem de nicel veri toplama araçlarının kullanıldığı karma yöntem yaklaşımını benimseyen bir durum çalışmasıdır. Ayrıca, ölçme değerlendirme eğitiminin farklı programları arasındaki uyumu incelemek için Porter'ın Öğretilen Program Anketleri (ÖPA) uyum metodolojisinin bir uyarlamasını kullanılmıştır. Hedeflenen ve değerlendirilen program verileri, program belgeleri ve ölçme araçları vasıtasıyla toplanmış ve içerik analizleri yapılmıştır. Öğretilen ve öğrenilen programlara ilişkin veriler ise öğretmen adayları ve öğretim üyelerine uygulanan anketler aracılığıyla toplanmış ve betimsel yöntemle analiz edilmiştir. Analiz edilen nitel ve nicel veriler bir sonraki aşamada yapılacak uyum analizi için matrislere aktarılmış ve Porter'ın uyum indeksi formülü kullanılarak programlar arasındaki uyum hesaplanmıştır. Sonuçlar, İngilizce öğretmenliği lisans programında verilen ölçme değerlendirme eğitiminin farklı programları arasındaki uyumun 0, 44 ile 0, 78 arasında, orta ila yüksek indekslerde olduğunu göstermiştir. Bu çalışma, hedeflenen ve öğretilen programlar ile değerlendirilen ve öğrenilen programlar arasında güçlü bir uyum tespit ederken hedeflenen ve değerlendirilen; hedeflenen ve öğrenilen; öğretilen ve öğrenilen ve hedeflenen ve öğrenilen programlar arasında orta düzey bir uyum bulunmuştur. Bu çalışmanın sonuçları, öğretmen adaylarına verilen ölçme değerlendirme eğitiminin programları arasındaki

Anahtar Kelimeler

Ölçme değerlendirme eğitimi
Değerlendirilen program
Program uyumu
Öğretilen program
Öğretmen yetiştirme
Hedeflenen program
Öğrenilen program

Makale Hakkında

Gönderim Tarihi: 09.05.2020
Kabul Tarihi: 22.04.2021
Elektronik Yayın Tarihi: 18.05.2021

DOI: 10.15390/EB.2021.9750

¹ Orta Doğu Teknik Üniversitesi, Yabancı Diller Yüksek Okulu, Türkiye, stekir@metu.edu.tr

uyumsuzluğu tespit etmeye hizmet etmektedir. Çalışmada ayrıca, öğretmen yetiştirme programlarının uyumunu iyileştirmeye yönelik öneriler sunulmuş ve ileride hizmet öncesi ölçme değerlendirme eğitimi üzerine yapılacak araştırmalar için çeşitli önerilerde bulunulmuştur.

Giriş

Ölçme değerlendirme okuryazarlığı, öğretmenler için temel bir mesleki gerekliliktir ve son yirmi yılda öğretmen eğitiminin odak noktası olmuştur (DeLuca, LaPointe-McEwan ve Luhanga, 2016). Ölçme değerlendirme hem öğrencilerin akademik başarısı hem de öğretmenlerin mesleki gelişimi üzerinde önemli etkilere sahiptir. Bu nedenle eğitimde önemli bir rol oynar. Uygun ölçme değerlendirme tekniklerini ve notlandırma uygulamalarını kullanan öğretmenler, kendi öğretmenlik uygulamalarını değerlendirip geliştirebilir, öğrencilerin öğrenme motivasyonlarını ve başarılarını artırabilir ve öğrenci başarısı hakkında doğru değerlendirmeler yapabilir (Dinther, Dochy ve Segers, 2015; Mellati, Khademi ve Shirzadeh, 2015; Mertler, 2003).

Ölçme değerlendirme okuryazarlığının öğretmenlere sağladığı bu önemli katkılar göz önünde bulundurulduğunda, öğretmenlerin ölçme değerlendirme bilgilerini artırmak sistemsel bir öncelik haline gelmiştir. Ulusal düzeyde örneklendirmek gerekirse, Milli Eğitim Bakanlığı, *Öğretmenlik Mesleği Genel Yeterlilikleri* belgesinde ölçme değerlendirmeyi, öğretmenlerin sahip olması gereken yeterliklerden biri olarak tanımlamaktadır (Milli Eğitim Bakanlığı [MEB], 2017). Ayrıca, Yükseköğretim Kurulu (YÖK), öğretmen adaylarının yeterli ölçme değerlendirme okuryazarlık düzeyine erişmeleri ve gerekli ölçme değerlendirme yetkinliklerini kazanmalarını sağlamak için, Türkiye'de öğretmen yetiştirme programlarının öğretmen adaylarına ölçme değerlendirme eğitimi vermesini zorunlu kılmıştır (YÖK, 2016). Tüm bunların sonucu olarak, öğretmen yetiştirme programları, öğretmen adaylarına bu konuda okuryazarlık ve yetkinlik kazandırmak için, zorunlu ölçme değerlendirme dersleri vermektedir.

Fakat bu konuda yapılan araştırmalar, mesleğe yeni başlayan İngilizce öğretmenlerinin, lisans programlarında aldıkları zorunlu ölçme değerlendirme derslerine rağmen, öğrenci kazanımlarını ölçme konusunda güvensiz hissettiklerini ve ölçme değerlendirme okuryazarlık düzeylerinin düşük olduğunu göstermektedir (Hatipoğlu, 2015; Karaman ve Şahin, 2014; Mertler, 2003; Öz ve Atay, 2017). Ayrıca, pek çok öğretmenin, ölçme ve öğretimi etkili bir şekilde birleştirme konusunda yetersiz olduğu gözlemlenmiş ve özellikle mesleğe yeni başlayan öğretmenlerin bu konuda yetersiz hissettikleri tespit edilmiştir (Mertler, 2003). Farklı programlara devam eden öğretmen adaylarından toplanan veriler, aday öğretmenlerin öğrenci öğrenmesini etkili bir şekilde ölçmek için hazır olmadıklarını ve bu durumun sebebinin büyük ölçüde öğretmen yetiştirme programında aldıkları ölçme değerlendirme eğitiminin yetersizli olmamasından kaynaklandığını iddia etmişlerdir (Karaman ve Şahin, 2014). Benzer şekilde, Mede ve Atay (2017), İngilizce öğretmenlerinin ölçme değerlendirme okuryazarlıklarının oldukça düşük seviyelerde olduğunu ve bu öğretmenlerin neredeyse ölçme değerlendirmenin tüm alanlarında daha fazla eğitime ihtiyaç duyduklarını ifade etmiştir.

Türkiye'de tüm İngilizce öğretmenliği programlarının, öğretmen adaylarına ölçme değerlendirme okuryazarlığı kazandırmak için zorunlu ölçme değerlendirme dersleri vermesine rağmen, mesleğe yeni başlayan İngilizce öğretmenlerinin yeterli ölçme değerlendirme okuryazarlığına sahip olmaması ve kendilerini öğrencilerin öğrenmesini ölçmek için yeterli hissetmemeleri oldukça ilginçtir. Araştırmalar, öğretmenlerin düşük ölçme değerlendirme okuryazarlıklarının sebeplerinden birinin program uyumsuzluğu olduğunu göstermektedir (DeLuca ve Klinger, 2010; Graham, 2005; Stiggins, 2005). Ölçme değerlendirme alanında yapılan araştırmalar, mesleğe yeni başlayan öğretmenlerin ve öğretmen adaylarının ölçme değerlendirme okuryazarlık ve yeterlik düzeylerini incelemiş ve ele alınması gereken bir soruna işaret etmiştir. Fakat bugüne kadar, bu soruna yol açan etkenlerden biri olabilecek ölçme değerlendirme dersinin farklı programlarının uyumunu inceleyen bir

çalışma bulunmamaktadır. Bu çalışma, bir İngilizce öğretmenliği lisan programında öğretmen adaylarına verilen ölçme değerlendirme ders programının uyumunu analiz ederek literatürdeki bu boşluğu gidermeyi amaçlamaktadır.

Program Uyumu

Program uyumu terimi, Webb (2007) tarafından “hedeflenen kazanımlar ve ölçme değerlendirme arasındaki uyumun, öğrencilerin bu kazanımları edinmesi için ortak hareket etme derecesi” olarak tanımlanmıştır” (s. 1). Roach, Niebling ve Kurz (2008) program uyumunu “program hedefleri ve ölçme araçları arasındaki uyum ve bu uyumun öğretmenlerin öğrencilerin akademik hedeflere ulaşmalarını desteklemek için gösterdikleri çabaları yönlendirme derecesi” (s. 1) olarak tanımladılar. Squires (2012) ise uyumu, programın tasarlanmasında, ölçme değerlendirme ve öğrenme hedeflerinin, öğretim sürecinde ele alınmasını sağlayacak şekilde düzenlemesi şeklinde tanımlamıştır.

Program uyumu, öğrenme deneyimleri boyunca tutarlı fikirlerle karşılaşan öğrencilerin daha iyi öğrenecekleri varsayımına dayanmaktadır (Bransford, Brown ve Cocking, 2000; Bruner, 1990). Biggs (1999), tüm bileşenlerinin uyumlu olduğu bir eğitim sisteminde, öğrencilerin herhangi bir içeriği öğrenmemesinin mümkün olmadığını iddia etmiş ve uyum ve başarı arasında bir ilişki olduğunu vurgulamıştır. Cohen (1987), ölçme değerlendirme ve hedefler arasındaki uyumun öğrenmeyi iki standart sapma kadar arttırdığını öne sürerek program uyumu ve öğrenme arasındaki ilişkiye dikkat çekmiştir. Benzer şekilde, diğer bazı araştırmacılar, program uyumunun öğrenci başarısını arttırdığını belirtmiştir (EdSource, 2006; Kercheval, 2001; McFadden, 2009; Murphy, 2007; Schuenemann, Jones ve Brown, 2011; Zavadsky, 2006).

Bugüne kadar program uyumu alanında yapılan çalışmalar, program uyumunun öğrenci başarısını artırabileceğini göstermektedir. Program uyumunu ölçmek, farklı programlar arasındaki uyumu hesaplamayı gerektirir ve bunu ölçmek çok kolay değildir. Porter'a göre (2002, 2004), program uyumunu ölçmenin zor olmasının sebebi, her bir program türünün içeriğini ortak bir dile çevirmek gerekmesidir. Yani yazılı (hedeflenen) program, sınıfta uygulanan öğretim faaliyetleri (öğretilen program) ve ölçme değerlendirme araçlarının (değerlendirilen program) içeriklerinin ortak bir dile çevrilmesi gerekmektedir. Bu amaçla, Porter (2004) hedeflenen, öğretilen ve değerlendirilen programların içeriğini tanımlamıştır. Porter, hedeflenen programı, sınıfta uygulanması gereken öğretimin içeriği olarak tanımlamıştır. Öğretilen programı, öğrencilere öğrenme imkânı verilen içerik olarak tanımlayan Porter, değerlendirilen programı ise öğrencilerin test edildiği içerik olarak ifade etmiştir. Programlar arasındaki uyum, bu programların içerikleri arasındaki uyum derecesini ve yine bu programların, öğrenmeyi (öğrenilen programı) desteklemek için uyumlu çalışma derecelerini ifade eder.

ÖPA (Öğretilen Program Anketleri) Program Uyum Modeli

Program uyumunu ölçmek için, Webb ve Porter iki ana model geliştirmiş ve bu modeller için ölçme araçları tasarlamıştır. Bu modeller, Webb modeli ve Öğretilen Program Anketleri (ÖPA) modeli olarak bilinmektedir. Webb uyum modeli, içerik, hedefler ve değerlendirme araçları arasındaki uyumu ölçmek için oluşturulmuş bir çerçevedir. ÖPA modeli ise hedefler, değerlendirme araçları ve sınıf içi öğretim uygulamaları arasındaki uyumu ortak bir içerik matrisi aracılığıyla ölçmeyi sağlayan bir modeldir (Bhola, Impara ve Buckendahl, 2003; Council of Chief State School Officers [CCSSO], 2006; Roach vd., 2008). Bu çalışmada, öğretilen program uyumunu da ölçmeyi sağladığı için ÖPA modeli kullanılmıştır. Ayrıca, ÖPA modeli, Webb modeline göre iki önemli avantaj sunmaktadır. İlk olarak bu model, farklı program türlerini karşılaştırmayı ve aralarındaki uyumla ilgili bir sonuca varmayı sağlayan ortak bir dil çerçevesi kullanılmaktadır. İkinci olarak ise, bu modelde, programlar arasındaki uyumun derecesini ifade etmeyi kolaylaştıran 0 ile 1 aralığında bir ölçek kullanılmaktadır. Bu ölçekte “0” programlar arasında hiç uyum olmadığını “1” ise tam bir uyum olduğunu göstermektedir (Kurz, Elliott, Wehby ve Smithson, 2010).

Ayrıca bu model, program içeriklerinin iki farklı boyutta incelenmesini ve incelenen içeriklerin bir içerik matrisine aktarılmasını sağlar. İçeriğin incelendiği boyutlardan ilki içerik konusunu, ikinci ise

bilişsel düzeyi göstermektedir. Bloom'un (1949) taksonomisinde olduğu gibi, ÖPA de bilişsel görevleri beş bilişsel düzey kullanarak tanımlamaktadır. Bunlar 1) hatırlama, 2) prosedür becerileri, 3) uygulama, 4) analiz ve 5) sentezdir (Smithson, 2007).

Hedeflenen ve değerlendirilen programların içerikleri genellikle konunun uzmanları tarafından içerik matrisine dönüştürülür. Bu süreçte ilk olarak, uzmanlar ÖPA kodlama ilkeleri konusunda eğitilirler. Uzmanlar, daha sonra, değerlendirilen programı (yani ölçme maddelerini) ve yazılı/hedeflenen programı (yani ders programında belirtilen hedefleri, ders materyallerini vb.) konuya özgü içerik dilini ve bilişsel düzey kategorilerini kullanarak kodlarlar (Smithson, 2007). Bundan sonra, kodlar o programa özel içerik matrisine aktarılır. Öğretilen programın içeriğiyle ilgili bilgiler ise bizzat öğretmenlerden toplanır. Öğretmenler, ÖPA kodlama kuralları konusunda bilgilendirildikten sonra, ders içeriklerini, derse özgü içerik dili kullanarak bildirirler. Bunu yaparken, sınıf içi öğretim uygulamalarının her bir konuyu ve bilişsel düzeyi ne kadar kapsadığını sayısal olarak ifade ederler (Blank, 2002).

Daha sonra, her program türü için toplanan ve ilgili programın matrisine aktarılan içerik verileri matristeki tüm satırlar ve sütunların toplamı 1.00'e eşit olacak şekilde hücre düzeyine indirgenir. Bu işlem, analiz edilen her bir birimin (hücrenin) bir bütün içinde belli bir oranı olduğu anlamına gelmektedir (Roach vd., 2008). Buna Oransal Değer (OD) denir ve içerik matrisinin tüm hücrelerindeki oransal değerlerin toplamı 1.00 olmalıdır. Oransal değerleri hesaplamak için, her hücredeki içerik değeri, matristeki içeriklerin toplam değerine bölünür. Bu, belirli bir içerik kategorisinin belirli bir bilişsel düzeydeki oranının matristeki değerlerin toplamına oranını ölçmenin bir yoludur (DeLuca ve Bellara, 2013). Örnek vermek gerekirse, hedeflenen program matrisindeki değerlerin toplamı 43'tür. Bu matriste 3 değerine sahip bir hücre için oransal değer (OD) $3/43 = 0,07$ 'dir. Bu şekilde, hücrelerdeki veriler, tüm satır ve sütunlardaki değerlerin toplamı 1'e eşit olacak şekilde, oransal olarak küçültülür. Bu hesaplama yöntemi, tüm ÖPA içerik matris hesaplamaları için geçerlidir. Hedeflenen, öğretilen, değerlendirilen ve öğrenilen programların her biri için oluşturulan içerik matrisler 25 hücreden (5 sütun \times 5 konu) oluşturulmuştur ve bu hücrelerin içlerindeki değerlerin toplamı 1'e eşit olacak şekilde OD'le ifade edilmiştir. Bu nedenle bu matrisler eşleştirilebilir ve aralarındaki uyumu ölçmek için nicel analizler uygulanabilir.

Program Uyumu Araştırmaları

Program uyumu konusunda bu güne kadar yapılan araştırmaların çoğu, program uyumu ve öğrenci başarısı arasındaki ilişkiyi incelemiştir. Örneğin, Schmidt ve diğerleri (2001) 40'dan fazla ülkede matematik ve fen derslerinin hedeflenen programlarını, TIMMS sınav sorularını ve bu ülkelerdeki öğrencilerin sınav notlarını incelediler. Sonuç olarak hedeflenen program ve değerlendirilen program içerikleri arasındaki uyumun öğrenci başarısını artırdığını buldular. Bir başka deyişle, hedeflenen programda konuların ele alınma düzeyinin, öğrencilerin TIMSS sınavındaki başarıları üzerinde etkili olduğunu iddia ettiler (Schmidt vd., 2001). Benzer şekilde, Porter, Kirst, Osthoff, Smithson ve Schneider (1994), lise birinci sınıf matematik dersinin hedeflenen programı ve sınıfta öğretilen programı arasındaki uyumunu ve bu derste öğrenci başarısını incelediler. Çalışmanın bulguları, öğretilen programın hedeflenen programla uyumlu olduğunda, öğrenci başarısında olumlu ve anlamlı bir artışa yol açtığını göstermiştir (Porter vd., 1994). Ayrıca, Schmidt ve diğerleri (2001) ABD'de sekizinci sınıf öğrencilerinin TIMMS'deki matematik puanlarını ve öğretim süresini incelediler. Çalışmanın bulguları, öğretilen ve değerlendirilen programlar arasındaki uyumun öğrencilerin sınav başarısına katkıda bulunduğunu ortaya koymuştur. Genel olarak, bir konunun sınıfta öğretimine yaklaşık bir hafta daha fazla zaman ayırmak, öğrencilerin sınav puanlarında %3 ila 24'lük bir artışa neden olmaktadır. Bu nedenle, önemli konular için yapılacak ek öğretimin faaliyetlerinin öğrencilerin daha iyi öğrenmelerine ve sınavlarda daha yüksek puanlar almalarına katkı sağlayacağı düşünülmektedir (Schmidt vd., 2001). Cohen (1987) çalışmalarında benzer bulgular ortaya koymuştur. Araştırmacı, deneysel tasarımdaki çalışmasında, öğretim ve ölçme arasındaki ilişkiyi araştırmış ve öğretilen ve değerlendirilen programlar arasındaki uyumun, hem düşük hem de yüksek yetenekli öğrencilerin öğrenmesini olumlu etkilediğini

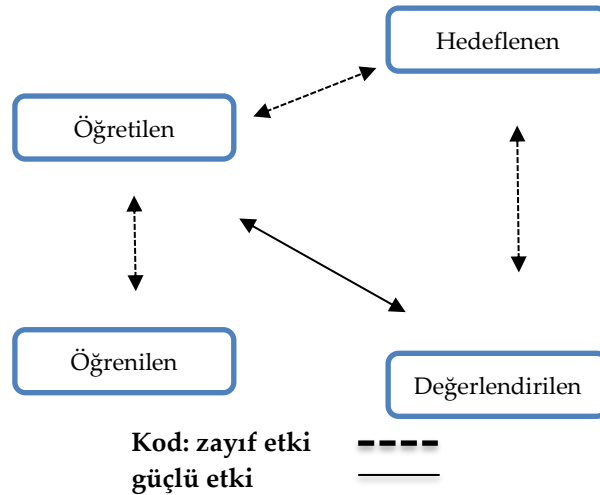
bulmuştur. İlginç bir şekilde, bu iki program arasındaki uyumun düşük yetenekli öğrencilerin öğrenmesini yüksek yetenekli öğrencilerden daha fazla artırdığını öne sürmüştür.

Son zamanlarda yapılan program araştırmalarının birinde Kurz ve meslektaşları, genel ve özel eğitimde sekizinci sınıf matematik dersinin hedeflenen, öğretilen ve değerlendirilen programları arasındaki uyumu Öğretilen Program Anketleri (ÖPA) modelini kullanarak araştırmıştır. Çalışmalarının sonunda, araştırmacılar, genel ve özel eğitimde hedeflenen ve öğretilen programlar arasındaki uyumun nispeten düşük olduğunu bulmuştur (1.00 üzerinden .12). Ayrıca, bu çalışmanın sonunda, araştırmacılar öğretilen ve değerlendirilen programlar arasındaki uyum ile öğrenci başarısı arasında önemli bir ilişki olduğunu ifade etmiştir (Kurz vd., 2010). Benzer bir şekilde, Seitz (2017), dokuzuncu sınıf matematik dersinin hedeflenen, öğretilen ve değerlendirilen programları arasındaki uyumu içerik ve bilişsel düzey boyutunda araştırmıştır ve çalışmasının sonunda hedeflenen, öğretilen ve değerlendirilen programlar arasında içerik açısından yüksek, bilişsel düzey açısından düşük uyum olduğunu tespit etmiştir. Bu konudaki literatür, program türleri arasındaki uyumun öğrenci başarısını olumlu etkilediğini göstermektedir ancak program türlerinin birbirlerini nasıl etkilediği henüz tam olarak bilinmemektedir. Bu konuda yapılan sınırlı çalışmalar, programlar arası etkileşimle ilgili aşağıdaki noktalara değinmektedir.

Mevcut araştırmalar, program türleri arasında farklı etkileşim örüntüleri olduğunu söylemektedir. Örneğin, Glatthorn (2000), Şekil 1'de gösterildiği gibi, öğretilen program ile değerlendirilen programlar arasında daha güçlü bir etkileşim olduğunu söylemiştir. Bu durumun eğitimde hesap verebilirlik (accountability) çabalarının bir sonucu olduğu söylenebilir. Gooding'in (1994) iddia ettiği gibi, hesap verebilirlik çabaları, öğretmenlerin öğrencilerinin sınav performansı ile daha fazla ilgilenmelerine yol açmış ve sınıfta geçirdikleri zamanın çoğunu öğrencilerin sınav başarılarını artırmak için bölge, eyalet veya ulusal sınavlardan örnek sorular çözmeye harcamalarına yol açmıştır (Gooding, 1994).

Araştırma bulguları, öğretmenlerin sınıftaki tüm çabalarına rağmen, öğretilen ve öğrenilen programlar arasındaki uyumun çok güçlü olmadığını göstermektedir. Başka bir deyişle, öğrenciler öğretmenlerinin sınıfta öğrettiklerini her zaman öğrenmemektedir Bunun nedenleri arasında, öğrenci motivasyonunun düşük olması, öğrencinin bilişsel becerisinin kısıtlı olması, öğrencilerin dikkat sürelerinin kısa olması ve öğretmenlerin, öğrencilerin öğrenmesini takip etme veya programı onlar için anlamlı ve zorlayıcı hale getirme konusunda yetersiz kalmaları gibi farklı nedenler sayılabilir (Glatthorn, Carr ve Harris, 2001).

Ayrıca, Glatthorn ve diğerleri (2001) hedeflenen programın hem öğretilen hem de değerlendirilen programlar üzerinde sadece kısmi bir etkiye sahip olduğunu iddia etmiştir. Araştırmacılar, hedeflenen ve öğretilen programlar arasındaki nispeten zayıf etkileşimin nedeninin, çoğu öğretmenin, özellikle de deneyimli olanların, program belgesine yılın başında şöyle bir bakıp yılın geri kalanında bu belgeyi pek kullanmaması olduğunu söylediler. Öğretmenlerin sınıfta neyi ve nasıl öğreteceklerine karar vermek için başka bazı faktörleri göz önünde bulundurdıklarını ifade ettiler. Örneğin, öğretmenlerin, öğretim faaliyetlerini planlarken öğrencilerin ilgisini, sınıfta neyin işe yarayıp neyin yaramayacağına ilişkin deneyimlerini ve ulusal sınavlarda hangi konuların sorulacağı ile ilgili düşünceleri gibi diğer faktörleri dikkate aldıklarını belirttiler.



Şekil 1. Program türleri arasındaki etkileşim (Glatthorn, 2000'den adapte edilmiştir)

Program uyumu ile ilgili yapılan araştırmalar, etkili öğrenme için hedeflenen, öğretilen ve değerlendirilen programların uyumlu olması gerektiğini göstermektedir (Elliott, Braden ve White, 2001; Webb, 1997a, 1997b). Benzer şekilde, öğretmen eğitimi üzerine yapılan araştırmalar, öğretmen adaylarına tutarlı öğrenme deneyimleri sağlanmadığı takdirde, düşük okuryazarlık ve yeterlik düzeyine sahip olabileceklerini iddia etmektedir (Darling-Hammond, 2006; Howey ve Zimpher, 1989; Russell ve Mcpherson, 2001). Öğretmen yetiştiren programlarının verdiği ölçme değerlendirme eğitimi üzerine yapılan araştırmalar, İngilizce öğretmenlerinin ölçme değerlendirme konusunda yeterliklerinin ve performanslarının düşük olduğunu göstermektedir (Hatipoğlu, 2015; Karaman ve Şahin, 2014; Mede ve Atay, 2017; Mertler, 2003; Öz ve Atay, 2017). İngilizce öğretmenlerinin düşük ölçme değerlendirme yeterlikleri ve okur yazarlıklarının nedeninin öğretmen eğitimi programında aldıkları ölçme değerlendirme eğitiminin programındaki uyumsuzluktan kaynaklanıp kaynaklanmadığı henüz araştırılmamıştır. Yani, öğretmenlerin yaşadıkları bu sorunun nedeninin, ölçme değerlendirme dersinin hedeflenen, öğretilen ve değerlendirilen programları arasında ki uyumsuzluk olup olmadığı henüz incelenmemiştir. Ayrıca, şimdiye kadar program uyumu konusunda yapılan çalışmaların çoğu hedefler ve değerlendirme araçları arasındaki uyumu araştırmıştır (örneğin, Resnick, Rothman, Slattery ve Vranek, 2003; Webb, 1997a, 1997b; Webb, Herman ve Webb, 2006). Öğretilen programla ilgili yapılan sınırlı sayıda ki çalışma ise K-12 bağlamında yapılmıştır. Öğretmen eğitiminde öğretilen program ve bu programın diğer programlarla uyumunu inceleyen bir çalışma bulunmamaktadır. Bu çalışma, daha önce araştırılmamış bir alan olan öğretmen eğitimi bağlamında öğretilen programı incelemeyi hedeflediği için program uyumu literatürüne katkı sağlayacaktır. Ayrıca, bu çalışmanın, öğretmen yetiştirme programında verilen ölçme değerlendirme eğitiminin program uyumu açısından eksiklerini tespit edip programın öğretmen adaylarının ihtiyaçlarını daha iyi karşılaması için öneriler sunmayı hedeflediği için de öğretmen eğitimi literatürüne katkı sağlaması beklenmektedir.

Bu amaçlarla, araştırmacı, Türkiye'nin İç Anadolu bölgesinde bir İngilizce öğretmenliği programı tarafından verilen ölçme değerlendirme eğitiminin hedeflenen, öğretilen, öğrenilen ve değerlendirilen programları arasındaki uyumu incelemeyi amaçlamıştır. Ayrıca, mevcut çalışmanın bir başka hedefi de program türleri arasında ki etkileşim örüntüsünü bulmaktır. Bu bağlamda, çalışma aşağıdaki araştırma soruları tarafından yönlendirilmiştir:

1. İngilizce öğretmenliği programında verilen ölçme değerlendirme eğitiminin hedeflenen, öğretilen, öğrenilen ve değerlendirilen programları ne ölçüde uyumludur?
2. İngilizce öğretmenliği programında verilen ölçme değerlendirme eğitiminin hedeflenen, öğretilen, öğrenilen ve değerlendirilen programları arasında nasıl bir etkileşim örüntüsü vardır?

Yöntem

Araştırma deseni

Çalışma, karma yöntemin kullanıldığı bir durum çalışmasıdır. Bu çalışma, bir durumun derinlemesine incelenmesi ve kanıt sunulması için nicel ve nitel veri toplama, sonuçlandırma ve birleştirme yöntemlerini kullanan bir karma yöntem çalışmasıdır (Creswell ve Plano Clark, 2018). Durum ve karma yöntem çalışmaları birbirinden tamamen farklı çalışmalar değildir. İkisi arasında oldukça ince bir çizgi vardır. Bu durum bunlardan herhangi birinin araştırmayı desteklemesine veya yönlendirmesine olanak sağlar (Carolan, Forbat ve Smith, 2016).

Mevcut çalışma, karma yöntemli durum çalışması yaklaşımını benimsemiştir ve belirli bir program içindeki bireysel unsurları ve ilişkileri hem nitel hem de nicel veri toplama araçlarıyla derinlemesine araştırmayı amaçladığı için literatürdeki karma yöntemli durum çalışması tanımıyla tutarlılık göstermektedir. Rowley'in (2002) iddia ettiği gibi, durum çalışması, anketler, röportajlar, belge analizi ve gözlem gibi nicel ve nitel yaklaşımların ve araçların bir harmanını kullanabilir. Bu çalışmanın amaçları doğrultusunda, veri toplamak için anket ve belge analizi araçları kullanılmıştır. Öğretilen ve öğrenilen programların içeriğine ilişkin veriler, öğretim üyelerine ve öğretmen adaylarına uygulanan anketlerle toplanmıştır. Hedeflenen ve değerlendirilen program için ise program belgeleri (ör. ders programı, ders materyalleri ve okumaları) ve ölçme değerlendirme araçları (ör. ara sınavlar ve final sınavı) içerik analizine tabi tutulmuştur.

Araştırma Alanı ve Katılımcılar

Araştırmacı, İngilizce öğretmenliği programlarının ortak özelliklerini belirleyip genellemeler yapmak yerine, araştırmayı daha doğru ve pragmatik hale getirmek için, belirli bir öğretmen eğitimi programı tarafından verilen ölçme değerlendirme dersinin programları arasındaki uyumu araştırmayı amaçlamıştır. Bu çalışmada, programın kendi bağlamı, uygulamaları, ve karmaşıklığı içindeki uyumunu incelemeyi amaçlamıştır (Stake, 1995). Ayrıca, uyum, programların ayırt edici bir özelliğidir. Her program, farklı paydaş profiliyle benzersizdir ve her programın farklı bir uyum örüntüsü olması doğaldır. Bu nedenle, Türkiye'deki tüm İngilizce öğretmenliği programlarının uyumunu araştırmak ve bu programların tamamının uyumu hakkında genel bir karar vermek mantıklı veya bilimsel olmayacağı için, bu çalışmada, hizmet öncesi ölçme değerlendirme programının uyumu hakkında kapsamlı ve kesin bir sonuca varabilmek için örnek bir araştırma alanı seçilmiştir.

Araştırmacı, Türkiye'nin İç Anadolu bölgesinde bir İngilizce öğretmenliği programı tarafından verilen İngilizce öğretiminde ölçme değerlendirme dersinin uyumunu incelemeyi seçmiştir. Araştırma yeri olarak, kolay ulaşılabilir olduğu ve araştırma için gerekli zengin ve ayrıntılı bilgiyi sağlama potansiyeline sahip olduğu için belirli bir öğretmen yetiştirme kurumu seçilmiştir. Pickart (2018), araştırmacılara çalışmanın başından itibaren zengin veri sağlayacak bir araştırma alanı seçmelerini önermektedir. Araştırma alanını seçerken, araştırmacıların, birden çok veri toplama tekniğini kullanabilecekleri, olay hakkında önemli bilgi sağlayacak kişilere ve belgelere ulaşabilecekleri sahaları seçmeleri önerilmektedir. Araştırmacıların ayrıntılı veri elde etmek için, araştırma alanına girme izninden daha fazlasına ihtiyacı vardır. Örneğin, tüm paydaşlarla güven ilişkisi geliştirmeleri gerekir (Pickart, 2018). Mevcut araştırma, ölçme değerlendirme araçları gibi bazı gizli belgelere erişmeyi gerektirdiğinden, araştırmacı, ilgili bilgi ve belgeleri sağlayabilme potansiyeli olan bu araştırma alanını seçmiştir.

Araştırmaya, seçilen İngilizce öğretmenliği lisans programı tarafından 2017-2018 eğitim öğretim yılında verilen İngilizce öğretiminde ölçme değerlendirme dersini alan öğretmen adayları anket yöntemiyle katılmıştır. Anketler, dönemin son dersinde, dersi üç farklı sınıfta alan toplam 90 öğretmen adayından sınıfta bulunan 57'sine (%51,3) uygulanmıştır. Anketin uygulandığı 57 öğretmen adayın 42'si kadın, 15'i erkektir. Diğer katılımcı grubunu ise, aynı eğitim öğretim yılında İngilizce öğretiminde ölçme değerlendirme dersini veren öğretim üyeleri oluşturmaktadır. Ders, üç farklı öğretim üyesi tarafından üç farklı grupta veriliyordu. Tablo 1'de gösterildiği gibi, çalışmaya katılan öğretim üyelerinden biri profesör, diğerleri ise doçentti. Katılımcı akademisyenlerin biri erkek, diğer

ikisi ise kadındı. Öğretim üyelerinin öğretmen yetiştirme deneyimleri ise 15 ila 30 yıl arasında değişmekteydi. Öğretim üyelerinin tamamı doktora derecelerini Türkiye'deki devlet üniversitelerinden almışlardı. İki öğretim üyesi doktorasını İngilizce öğretmenliği (ELT) alanında yapmış, diğer öğretim üyesi ise İngiliz dil bilimi alanında doktora yapmıştı. Araştırmacı, ders döneminin başında öğretim üyeleriyle görüşerek onlara araştırma ve kendilerinden beklenen katkı hakkında bilgi vermiştir. Araştırmacı, çalışmanın en başında, öğretim üyelerinin çalışmaya katılmayı kabul ettiklerini gösteren onayı almış ve bundan sonra araştırma sürecini başlatmıştır.

Tablo 1. Öğretim Üyelerinin Profilleri

Katılımcılar	Yaş	Cinsiyet	Unvan	Tecrübe (Yıl)	Lisansüstü Uzmanlık Alanı
Öğr. Gör. 1	62	Erkek	Profesör	30	İngilizce Öğretmenliği
Öğr. Gör. 2	43	Kadın	Doçent	15	Dil Bilim
Öğr. Gör. 3	45	Kadın	Doçent	18	İngilizce Öğretmenliği

Veri kaynakları

Tablo 2'de görüldüğü gibi, bu çalışmadaki veriler iki ana kaynaktan toplanmıştır. Bunlar program belgeleri ve anketlerdir. Hedeflenen ve değerlendirilen programa ilişkin veriler, program ve ölçme değerlendirme belgeleri aracılığıyla toplanırken, öğretilen ve öğrenilen program verileri, öğretim üyeleri ve öğretmen adaylarından anketler aracılığıyla toplanmıştır.

Tablo 2. Veri Toplama Araçları

Hedeflenen Prg.	Öğretilen Prg.	Öğrenilen Prg.	Değerlendirilen Prg.
Program Belgeleri	Öğretim üyesi anketi	Aday öğretmen anketi	Ölçme değerlendirme araçları/belgeleri

Program belgeleri. Hedeflenen program ile ilgili veriler ders programı (syllabus) ve ders materyallerinden (ders kitapları, okumalar ve basılı ders notları vs.) toplanmıştır. Araştırmacı, öğretim üyelerini araştırma ve veri toplama süreci hakkında bilgilendirmek için her biriyle yüz yüze görüşmeler yapmış ve katılımcılardan araştırmaya katılmayı kabul ettiklerini gösteren onayı aldıktan sonra, en güncel ders programlarının, ders kitaplarının, ders okumalarının ve ders notlarının birer kopyasını istemiştir. Tüm öğretim üyeleri iki hafta içinde e-posta yoluyla belgelerin dijital kopyalarını araştırmacıya iletmiştir. Ders programı (syllabus), çalışmaya katılan üç akademisyen tarafından ortaklaşa hazırlanmıştır. Bu nedenle, ders içeriği ve hedefleri üç sınıfın hepsinde aynıdır. Ders kapsamında, tüm sınıflarda iki ana ders kitabı kullanılmıştı. Ayrıca, dersi veren öğretim üyeleri tarafından dönem boyunca öğretmen adaylarıyla paylaşılmak üzere İngilizce öğretiminde ölçme değerlendirmenin farklı yönleri üzerine altı makale seçilmişti. Öğretim üyeleri, haftalık konuları sınıfta sunmak için slaytlar kullanıyorlardı. Bu slaytlar, içerik ve tasarım açısından küçük farklılıklar gösterebilir genel hatları aynıydı. Sonuç olarak, ölçme değerlendirme dersinin hedeflenen program içeriğini oluşturan program belgeleri, ders programı, iki ana ders kitabı, altı sınıf dışı okunması gereken makale ve öğretim üyelerinin sunum slaytlarından oluşmuştur.

Ölçme değerlendirme araçları. Dersi alan öğretmen adayları, dersi geçmek için, iki ara sınav ve bir final sınavında başarılı olmak, programda yer alan konulardan birini sınıfta başarılı bir şekilde sunmak ve sınıf içi etkinliklere aktif olarak katılmak zorundaydı. Araştırmacı, dersin değerlendirme içeriğini belirlemek için, resmi ve yazılı değerlendirme araçlarını yani ara sınavları ve final sınavını kullanmaya karar verdi. Araştırmacı, dönem sonunda, ölçme araçları sınıfta uygulandıktan sonra, dersi veren öğretim üyeleriyle yüz yüze görüşmeler yaparak ölçme araçlarının (ara sınavların ve final sınavlarının) birer kopyasını talep etti. Değerlendirme araçları gizli belgeler olduğundan, araştırmacı öğretim üyelerine, bu belgelerin yalnızca araştırma amacıyla kullanılacağını ve belgeleri kodlayacak uzmanlar dışında kimseyle paylaşılmayacağını açıklayan bir gizlilik mektubu verdi. Araştırmacı dönem sonunda ara sınavların ve final sınavının birer kopyalarını edindi. Sınavlar, üç öğretim üyesi tarafından ortak hazırlanmıştı. Ancak, uygulama biçimleri bazı yönlerden değişiklik gösteriyordu.

Örneğin, final sınavı, öğretmen adaylarının belirli bir durum için bir değerlendirme aracı tasarımlarını gerektiren bir proje çalışmasıydı. Öğretim üyelerinden ikisi projelerin çevrim içi gönderimine izin verirken, diğer öğretim üyesi öğretmen adaylarından ödevi basılı olarak göndermelerini istedi. Ayrıca, ikinci ara sınav evde yapılacak (take home) sınavdı ve birinci ve ikinci öğretim üyeleri bu sınavı bireysel bir çalışma olarak uygularken; üçüncü öğretim üyesi sınavı ikili gruplarla yapılacak bir değerlendirme aracı olarak kullanılmıştır. Kısacası, dersin değerlendirilen programının içeriğini, öğretim üyelerinden edinilen iki ara sınav ve bir final sınavı oluşturmuştur.

Öğretim üyesi anketi. Öğretilen programın içeriğine ilişkin veriler, öğretim üyelerinden anket yoluyla toplanmıştır. Öğretim üyeleri, her bir konuyla ilgili yaptıkları öğretim faaliyetlerinin içeriğini derse özel içerik dili kullanarak aktardılar. Sınıf içi öğretim uygulamalarının, hem konu içeriklerini kapsama düzeyini hem de her bir bilişsel düzeye yaptığı vurguyu sayısal olarak derecelendirdiler (Blank, 2002).

Katılımcı öğretim üyelerine, dönem sonunda, Öğretilen Program Anketleri (ÖPA) içerik anketi uygulanmıştır. ÖPA içerik anketleri, öğretmenlerin öğretim uygulamalarının içeriği hakkında güvenilir veriler elde etmek için kullanılan pratik ve tutarlı veri toplama araçlarıdır (CCSSO, 2006). Anket maddelerini, iki uzmanın, hedeflenen program içeriğini ilk kodlamalarının sonunda oluşturdukları konu içerik listesi ve bilişsel talep düzeyleri oluşturmuştur. İçerikte, birden fazla alt konuyu kapsayan 5 ana içerik alanı vardır: *İngilizce öğretiminde kullanılan ölçme yöntemleri ve ilkeleri* (6 konu), *dil becerilerini ölçme değerlendirme ve sınav geliştirme* (8 konu), *alternatif ölçme değerlendirme yöntemleri* (5 konu), *sınav sonuçlarının öğrenci değerlendirmesi için analizi* (4 konu), *kullanılan sınavların etkililiğinin değerlendirilmesi* (6 konu). Öğretim üyeleri, anketlerde, hedeflenen programda yer alan her bir konuyu sınıfta ne düzeyde ve hangi bilişsel alanda öğrettiklerini bildirdiler. Diğer bir deyişle, öğretim üyelerinden, öğrencilerinin her bir konuyu hedeflenen programda belirtilen bilişsel düzeyde öğrenmelerini sağlayıp sağlamadıklarını belirtmeleri istenmiştir. Öğretim üyelerinin öğretilen programın içeriğini belirtmeleri için dördümlü Likert ölçeği kullanılmıştır: "Hiç kapsamıyor" için "0", "Hafif Kapsıyor" için (bir ders ve daha az saat ayrılmışsa) "1", "Orta Derecede Kapsıyor" için (beş ders saati ayrılmışsa) "2" ve "Çok Kapsıyor" için (beşten fazla ders saati ayrılmışsa) "3". Anketler öğretim üyelerine e-posta ile gönderilmiştir ve bir hafta içinde tüm katılımcılar anketleri tamamlayarak araştırmacıya geri göndermiştir.

Öğretmen adayı anketi. İngilizce öğretiminde ölçme değerlendirme dersini alan öğretmen adaylarına uygulanan ankette, dersin hedeflediği konularda ve bilişsel düzeylerde kendilerini ne derece yetkin hissettiklerini belirtmeleri istenmiştir. Diğer bir deyişle, öğretmen adaylarından programda belirtilen öğrenme içeriklerini programda belirtilen bilişsel düzeylerde öğrenip öğrenmediklerini belirtmeleri istenmiştir. Hedeflenen program belgeleri iki uzman tarafından kodlanmış ve ortaya konu içerik ve bilişsel düzey listesi çıkartılmıştır. Anket maddeleri oluşturulurken bu liste kullanılmıştır. Katılımcılardan bu konularda ve bu bilişsel düzeylerde kendilerini ne kadar yeterli hissettiklerini dördümlü Likert ölçeğinde belirtmeleri istenmiştir. Ölçekte kullanılan kategoriler şu şekildedir: "Hiç Yetkin Değil" için "0", "Biraz Yetkin" için "1", "Yetkin" için "2" ve "Çok Yetkin" için "3".

Veri Analizi

Hedeflenen ve değerlendirilen program içerikleri için program belgeleri ve ölçme araçları aracılığıyla toplanan veriler içerik analizi ile analiz edilmiştir. Öğretilen ve öğrenilen program içerikleri için ise anketler yoluyla toplanan veriler betimsel istatistiklerle analiz edilmiştir. Her iki durumda da veriler daha sonraki analiz için matrislere aktarılmıştır.

Program belgeleri. Program belgelerinde yer alan hedefler, beklenen çıktılar ve ders içeriği, içerik analizi ile analiz edilmiştir. İki uzman, ders programı, ders kitapları, sunumlar ve okumalarını kodlayıp hedeflenen kazanımların bir listesini oluşturmuştur. Kodlamayı yapan uzmanların her ikisi de eğitim bilimlerinde doktora derecesi bulunan ve İngilizce öğretiminde ölçme değerlendirilme tecrübesine sahip akademisyenlerdi. Bu iki uzman, kodlamaya başlamadan önce, program uyumu metodolojisinin kodlama kuralları ve derse özgü içerik dilinin kodlamada nasıl kullanılacağı konusunda eğitilmiştir. İçerik analizi sürecinin güvenilirliğini sağlamak için kodlayıcılar arası

güvenirlilik analizi yapılmıştır. Bu süreçte ilk olarak iki kodlayıcı örnek olarak seçilen program belgelerini birlikte kodladılar. Belgelerde dikkat edilmesi gereken noktaları konuştuktan sonra, bir ön kod listesi oluşturdular. Daha sonra uzmanlar, belgeleri bağımsız olarak incelediler. Kodlama sürecinin sonunda, uzmanlar kodlarını ve önerdikleri temaları karşılaştırmak için bir araya geldiler. Uzmanların kodlamaları arasındaki korelasyon katsayısı .916 olarak hesaplanmıştır. Bu oran uzmanlar arasında güvenilirliğin yüksek olduğunu göstermektedir. Uzmanlar, sürecin sonunda, tüm belgeleri birlikte kodlayarak ve kodlamalardaki farklılıkları tek tek tartışarak ve ortak bir karara vararak, analiz için son kod listesini oluşturdular.

Program türleri arasında uyumu analiz etmek için uzmanlar, öncelikle hedeflenen program içeriğini belirlediler. Bunun için ders programı, ders materyalleri, sunumlar ve okumalar gibi program belgelerini incelediler. Başlangıçta toplam 29 kod belirlendi, daha sonra bu kodlar tematik olarak beş içerik temasına ayrıldı. Bu sürecin sonunda hedeflenen program içeriğinin konu başlıkları belirlendi. Alt konular, genel başlıklar altında konuya özgü bir içerik dili kullanılarak gruplandırıldı. İçerik beş ana başlıkta toplandı 1) İngilizce öğretiminde ölçme yöntemleri ve ilkeleri; 2) Dil becerilerinin değerlendirilmesi ve sınav geliştirilmesi; 3) Alternatif ölçme değerlendirme; 4) Sınav sonuçlarına dayalı öğrenci değerlendirmesi; 5) Kullanılan sınavların etkinliğinin değerlendirilmesi.

Uzmanlar, içerik alanlarını belirledikten sonra, tüm program belgelerini tekrar incelediler. Bu kez içerik konularını bilişsel düzeylerine göre kodladılar. Kodlama sisteminde hiç vurgu yok ise "0", hafif vurgu yani hedeflerin veya içeriğin %25'inden azında vurgu varsa "1", orta düzey vurgu yani hedeflerin veya içeriğin %33'ünde vurgu varsa "2" ve güçlü vurgu varsa yani hedeflerin veya içeriğin %33'ünden fazlasında vurgu varsa "3" kullanıldı. Bu sürecin sonunda oluşan kodlar hedeflenen program matrisine aktarıldı. Matrislerde bu değer "D" sütununda gösterilmiştir. Oluşturulan bu matrislerle hedeflenen programın diğer program türleriyle karşılaştırılabilmesi ve aralarındaki uyumla ilgili bir değerlendirme yapılabilmesi sağlanmıştır.

Değerlendirilen program. Araştırmacı tarafından edinilen ölçme değerlendirme araçları uzmanlar tarafından incelenmiştir. Uzmanlar, önce ölçme görevlerini performans adımlarına ayırdılar, daha sonra ise her adımı, her bir konu içeriğine ve bilişsel düzeye yaptığı vurguya göre kodladılar. Kodlama sisteminde hiç vurgu yok ise "0", hafif vurgu için yani ölçme aracının içeriğinin %25'inden azı konuya harcanmış ise "1", orta düzey vurgu için yani ölçme aracının içeriğinin %25-33'ü konuya harcanmış ise "2" ve güçlü vurgu için yani ölçme aracının içeriğinin %33'ünden fazlası konuya harcanmış ise "3" kodları kullanıldı. Oluşan kodlar, analizin güvenilirliği için öğretim üyelerine gönderildi ve yanlış yorumlama olasılığını ortadan kaldırmak için öğretim üyelerinin kodları gözden geçirmeleri istendi. Diğer bir deyişle, Lincoln ve Guba'nın (1985) önerdiği gibi, üye kontrol yöntemi ile çalışmanın güvenilirliği desteklendi. Üç öğretim üyesi de bu talebe olumlu dönüş yaptı ve uzmanların kodlarını onayladı. Ortaya çıkan sayısal kodlar, değerlendirilen programa ait matrisin "değer" yani "D" sütununa aktarılmıştır.

Öğretilen ve öğrenilen program anketleri. Öğretmen adayları ve öğretim üyelerinden anketlerle toplanan nicel veriler, betimsel istatistiklerle analiz edilmiştir. Anketlerdeki her bir maddeye verilen puanları hesaplamak için SPSS 18 programında ortalamalar hesaplanmıştır. Daha sonra, bu ortalamalar, ileri uyum analizleri için, öğretilen ve öğrenilen program matrislerinin "değer" (D) sütununa aktarılmıştır.

Uyum analizi

Uyum matrislerinin "değer" (D) sütununa aktarılan veriler, oransal değerlere (OD) dönüştürülmüştür (DeLuca ve Bellara, 2013). Bu adımda, her bir hücredeki veri sayısı, matristeki toplam veri sayısına bölünmüştür. Bu, belirli bir içerik kategorisine belirli bir bilişsel düzeyde yapılan vurgunun matristeki toplam veri sayısına oranını ölçmenin bir yoludur.

Her bir program türü için oluşturulan matrisler ikişerli eşleştirilerek aralarındaki uyumun nicel analizlerle ölçülebilmesi sağlanmıştır (Kurz vd., 2010). Bu aşamada, program eşleştirmeleri için toplam altı tablo oluşturulmuştur. Bu eşleştirmeler şöyledir: 1) hedeflenen ve öğretilen program, 2) hedeflenen ve değerlendirilen program, 3) hedeflenen ve öğrenilen program, 4) öğretilen ve öğrenilen program, 5) öğretilen ve değerlendirilen program, 6) öğrenilen ve değerlendirilen program. Programlar arası karşılaştırmayı kolaylaştırmak için “değerler” (D) ve “oransal değerler” (OD), matrislerde ayrı sütunlarda sunulmuştur. Bu tablolar bulgular bölümünde yer almaktadır.

Bir sonraki adımda, Porter’ın uyum indeksini hesaplamak için veriler arasındaki mutlak farkın toplamına ($\sum |x-y|$) ihtiyaç vardır. Bu nedenle, programların eşleştirildiği matrislerde, Ms. Excel kullanılarak hücre hücre karşılaştırma yapılmıştır ve farklar mutlak fark ($|x-y|$) sütununa yazılmıştır. Örneğin, hedeflenen program ve öğretilen programı karşılaştıran matrisin (Tablo 4) ilk hücresinde, “x” hedeflenen programda birinci konu içeriğine “hatırlama” bilişsel düzeyinde yapılan vurgunun oransal değerini gösterirken, “y” ise öğrenilen programda aynı konuya aynı bilişsel düzeyde yapılan vurguyu belirtmektedir. Bu iki programda bu konuya bu bilişsel düzeyde yapılan vurgunun oransal değerleri arasındaki mutlak fark $|0-0,08|=0,08$ olarak hesaplanmıştır ve matris üzerindeki mutlak fark sütununa ($|x-y|$) aktarılmıştır. Bu yöntemle, birbirine karşılık gelen her bir hücrenin oransal değerleri arasındaki fark elde edildikten sonra, her bir matris için bu farkların toplamı hesaplanmıştır (Fulmer, 2011). Örneğin, hedeflenen ve öğrenilen program eşleştirmesi için matristeki mutlak fark toplamı $0,08 + 0,2 + 0,11 + 0,12 + 0,16 = 0,67$ olarak hesaplanmıştır. Son adımda ise, program çiftleri arasındaki uyum, Porter’ın aşağıdaki uyum indeksi formülü kullanılarak hesaplanmıştır (Porter, 2002).

$$\text{Porter's Uyum İndeksi} = 1.0 - \frac{\sum |x-y|}{2}$$

$$\text{Hedeflenen ve Öğretilen Programlar Arası Uyum} = 1.0 - \frac{0,67}{2} = 0,67$$

Aynı formül, diğer programlar arasındaki uyum indekslerinin hesaplanmasında da kullanılmış ve sonuçlar bulgular bölümünde rapor edilmiştir.

Çalışmanın Sınırlılıkları

Çalışma önemli bulgular sunmasına rağmen, bazı sınırlılıkları da içermektedir. Öncelikle bu çalışma, sadece araştırmanın yapıldığı öğretmen yetiştirme programı ile sınırlandırılmıştır. Program türleri arasındaki uyumla ilgili yapılan bu çalışmanın bulguları, çalışmada incelenen programın ötesinde genellenemez. Diğer öğretmen yetiştirme programlarının farklı hedefleri, sınıf uygulamaları, ölçme değerlendirme uygulamaları ve sonuç olarak farklı öğrenci kazanımları olması muhtemeldir. Bu nedenle, farklı programların uyum indeksleri bu çalışmada bulunanlardan farklı olabilir.

Bununla birlikte, çalışmanın metodolojisi benzer bağlamlarda tekrar edilebilir. Çalışmanın kavramsal ve metodolojik çerçevesinin, ileride yapılacak program uyumu araştırmaları için model olması beklenmektedir. Ayrıca, mevcut çalışma, bir İngilizce öğretmenliği programı tarafından verilen İngilizce öğretiminde ölçme değerlendirme dersiyle sınırlıdır. Bu nedenle İngilizce öğretmenliği programının genel uyumunu incelemek için araştırmacı sonraki çalışmalarında programda yer alan her bir dersin ve okul uygulamalarının da dahil edildiği bir program uyumu çalışması planlamaktadır.

Çalışmanın bir başka sınırlılığı da anket verilerinin niteliğiyle ilgili endişelerdir. Öğretim üyeleri, sınıf içi uygulamalarının kapsamlarını rapor ederken hata yapmış olabilirler; ancak, mevcut literatür, sınıf öğretimiyle ilgili öğretmenlere uygulanan anket sonuçlarının sınıf gözlemleri ve öğretmen günlükleri ile yüksek bir korelasyona sahip olduğunu göstermektedir (Mullens ve Gayler, 1999; Shavelson, Webb ve Burstein, 1986; Smithson ve Porter, 1994). Ayrıca, Porter ve meslektaşlarının raporlarında ÖPA ölçümlerinin güvenilirliği ve geçerliliği ile ilgili ayrıntılı veriler bulunmaktadır (Porter vd., 1994).

Araştırma Etiği

Bu çalışma için etik onay ODTÜ İnsan Araştırmaları Etik Kurulu'ndan 28620816/158-300 onay numarası ile alınmıştır. Araştırmacı, öğretim üyelerine, e-posta aracılığıyla, araştırmanın amacı ve araştırma kapsamında incelenmesi gereken program belgeleri hakkında bilgi vermiştir. Araştırmacı ayrıca öğretim üyelerine, kendilerinden talep edilen program belgelerinin ifşa edilmeyeceğini, yetkisiz kişilerden korunacağını ve uygun şekilde muhafaza edileceğini bildiren yazılı imzalı bir sözleşme göndermiştir. Araştırmanın amacı ve gizlilik konuları hakkında bilgi verdikten sonra, araştırmacı öğretim üyeleri ve öğretmen adaylarından araştırmaya gönüllü olarak katıldıklarını gösteren imzalanmış gönüllü katılım formu almıştır.

Bulgular

Hedeflenen, Öğretilen, Öğrenilen ve Değerlendirilen Programlar Arasındaki Uyum Derecesi

Tablo 3'ün gösterdiği gibi hizmet öncesi İngilizce öğretiminde ölçme değerlendirme dersine ait hedeflenen, öğretilen, öğrenilen ve değerlendirilen programlar arasındaki uyum indeksleri 0,44 ila 0,78 arasında değişmektedir ve bu programlar arası orta ila yüksek uyum olduğunu göstermektedir. Porter ve Smithson (2000, 2001, 2002), program uyumuyla ilgili 1,00 indeks değerinin programlar arası her iki boyutta tam bir uyum olduğunu, 0,00 indeks değerinin ise programlar arası konu içeriği ve bilişsel düzey açısından hiçbir uyum olmadığını gösterdiğini belirtmiştir. Bu çalışmada programlar arası uyumla ilgili bulgular bu açıklamalara dayanmaktadır.

Tablo 3. Altı Farklı Program Eşleştirmesi için Uyum İndeksi

UI 1	UI 2	UI 3	UI 4	UI 5	UI 6
0,67	0,48	0,45	0,44	0,78	0,44

UI 1: Hedeflenen-Öğretilen, UI 2: Hedeflenen-Değerlendirilen, UI 3: Hedeflenen-Öğrenilen, UI 4: Öğretilen-Öğrenilen, UI 5: Değerlendirilen- Öğrenilen, UI 6: Öğretilen- Değerlendirilen

Hedeflenen program ile öğretilen program arasındaki uyumu ifade eden uyum indeksi (UI 1) 0,67 olarak hesaplanmıştır. Bu durum, öğretim üyelerinin hazırladıkları yazılı program belgesinde öğretmeyi amaçladıkları içeriğin %67'sini sınıfta öğrettiklerini göstermektedir. Hedeflenen programda içerik kategorileri %13 ile %23 arasında değişen oranlarda yer alırken, öğretilen programda bu konular %16 ila %24 arasında, benzer fakat daha az dengeli bir dağılıma sahiptir (Bkz. Tablo 4). Bilişsel düzey açısından ise her iki programdaki bilişsel seviyeler küçük farklılıklarla paralel bir dağılım göstermiştir. Hedeflenen ve öğretilen programlar arasındaki en önemli uyumsuzluk, "işlemsel beceriler" talep düzeyinin, hedeflenen programda (%26), öğretilen programdan (%16) daha yüksek bir oranla temsil edilmesidir.

Tablo 4. Hedeflenen ve Öğretilen Programlar Arası Uyum Matrisi

İçerik	Hatırlama			İşlemsel Beceriler			Uygulama			Analiz			Sentez			Toplam													
	Hedeflenen		Mutlak fark	Öğretilen		Mutlak fark	Hedeflenen		Mutlak fark	Öğretilen		Mutlak fark	Hedeflenen		Mutlak fark	Hedeflenen		Öğretilen											
	D	OD (x)		D	OD (y)		D	OD (x)		D	OD (y)		D	OD (x)		D	OD (y)		D	OD (x)	D	OD (y)							
1	0	0	3	,08	,08	3	,06	3	,08	,02	1	,02	0	0	,02	0	0	0	0	3	,06	0	0	,06	7	,13	6	,16	
2	0	0	0	0	0	3	,06	3	,08	,02	3	,06	3	,08	,01	3	,06	0	0	,06	2	,04	0	0	,04	11	,21	6	,16
3	0	0	0	0	0	3	,06	0	0	,06	3	,06	2	,05	,02	3	,06	3	,08	,02	3	,06	3	,08	,02	12	,23	8	,21
4	0	0	0	0	0	2	,04	0	0	,04	3	,06	3	,08	,02	3	,06	3	,08	,02	3	,06	3	,08	,02	11	,21	9	,24
5	0	0	0	0	0	3	,06	0	0	,06	3	,06	3	,08	,04	3	,06	3	,08	,02	3	,06	3	,08	,02	2	,22	9	,24
Toplam	0	0	3	,08	,08	14	,26	6	,16	,2	13	,25	11	,29	,11	12	,23	9	,24	,12	14	,26	9	,24	,16	43	1	38	1

1) İngilizce öğretiminde ölçme yöntemleri ve ilkeleri; 2) Dil becerilerinin değerlendirilmesi ve sınav geliştirilmesi; 3) Alternatif ölçme değerlendirme; 4) Sınav sonuçlarının analizine dayalı öğrenci ölçme değerlendirmesi; 5) Kullanılan sınavların etkinliğinin değerlendirilmesi.

D: Değer (Value), OD: Oransal Değer (Proportional Value)

Hedeflenen ve deęerlendirilen programlar arasındaki uyumu gsteren ikinci uyum analizi (UI 2) 0,48 olarak hesaplanmıřtır. Bu deęer, 0,50'den biraz dřk olduęu iin, đretim yelerinin đretmeyi planladıkları program ile lme aralarının test ettikleri arasında orta dzeyde bir uyum olduęu anlamına gelmektedir. Konu ierięi aısından iki program tr arasında kk farklılıklar vardır. Aralarındaki temel farklılıklar biliřsel dzey farklılıklarından kaynaklanmaktadır (Bkz. Tablo 5). Her iki program trnde ierik kategorileri %17 ile %23 arasında deęiřen oranlarda yeterli dzeyde yer almıřtır. Bu durumun dıřında kalan tek ierik konusu “İngilizce đretiminde lme yntemleri ve ilkeleri” olmuřtur. Bu ierik, hedeflenen programın %13’n, deęerlendirilen programın ise %19’unu oluřturmaktadır. Biliřsel dzeyler aısından ise, iki program arasında ok zayıf bir iliřki vardır. “Hatırlama” ve “iřlemsel beceriler” gibi dřk biliřsel dzeyler, deęerlendirilen programda sırasıyla %36 ve %42 gibi yksek oranlarda yer alırken, “uygulama” ve “analiz” gibi yksek biliřsel dzeyler %17 ve %6 gibi dřk oranlarda yer almıřtır. En yksek biliřsel dzey olan “sentez” dzeyinin deęerlendirilen programda hi yer almaması da ilgintir. Hedeflenen programda ise, biliřsel dzeylerin her biri, %23 ile %26 aralıęında yeterli oranda temsil edilmiřtir. Fakat, en dřk biliřsel dzey olan “hatırlama” hi yer almamıřtır.

Tablo 5. Hedeflenen ve Değerlendirilen Programlar Arası Uyum Matrisi

İçerik	Hatırlama			İşlemsel beceriler			Uygulama			Analiz			Sentez			Toplam													
	Hedeflenen		Değerlendirilen	Hedeflenen		Değerlendirilen	Hedeflenen		Değerlendirilen	Hedeflenen		Değerlendirilen	Hedeflenen		Değerlendirilen	Hedeflenen		Değerlendirilen											
	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)	x-y	D	OD	D	OD					
1	0	0	3	,08	,08	3	,06	3	,08	,02	1	,02	1	,03	,01	0	0	0	0	0	3	,06	0	0	,06	7	,13	7	,19
2	0	0	3	,08	,08	3	,06	3	,08	,02	3	,06	2	,06	0	3	,06	0	0	,06	2	,04	0	0	,04	11	,21	8	,22
3	0	0	2	,06	,06	3	,06	3	,08	,02	3	,06	2	,06	0	3	,06	0	0	,06	3	,06	0	0	,06	12	,23	7	,19
4	0	0	3	,08	,08	2	,04	3	,08	,04	3	,06	1	,03	,03	3	,06	1	,03	,03	3	,06	0	0	,06	11	,21	8	,22
5	0	0	2	,06	,06	3	,06	3	,08	,02	3	,06	0	0	,06	3	,06	1	,03	,03	3	,06	0	0	,06	2	,22	6	,17
Toplam	0	0	13	,36	,36	14	,26	15	,42	,12	13	,25	6	,17	,1	12	,23	2	,06	,18	14	,26	0	0	,28	43	1	36	1

1) İngilizce öğretiminde ölçme yöntemleri ve ilkeleri; 2) Dil becerilerinin değerlendirilmesi ve sınav geliştirilmesi; 3) Alternatif ölçme değerlendirme; 4) Sınav sonuçlarının analizine dayalı öğrenci ölçme değerlendirme; 5) Kullanılan sınavların etkinliğinin değerlendirilmesi.

D: Değer (Value), OD: Oransal Değer (Proportional Value)

Üçüncü uyum analizi (UI 3), öđretim üyelerinin hedefledikleri program ile öğrencilerin öğrendikleri program arasındaki uyumu göstermektedir. Bu programlar arasındaki uyum indeksi 0.45 olarak hesaplanmıştır. Bu durum, programlar arası orta düzey bir uyum olduğunu ve öđretim üyeleri tarafından hedeflenen içeriđin %45'inin öđretmen adayları tarafından öğrenildiđini göstermektedir. Her iki programın matris tabloları, hedeflenen ve öğrenilen programlarda içerik açısından farklılıklar olduğunu göstermektedir. Özellikle, "İngilizce öđretiminde ölçme yöntemleri ve ilkeleri", öğrenilen programda (%24), hedeflenen programdan (%13) daha fazla yer bulurken, "kullanılan sınavların etkinliđinin deđerlendirilmesi" öğrenilen programda (%12), hedeflenen programda olduğundan (%22) daha az yer almıştır (Bkz. Tablo 6).

Bu iki programda, bilişsel düzeyler açısından da uyumsuzluklar olduğđ görülmüştür. Örneđin, hedeflenen programda "hatırlama" düzeyinde içerik bulunmazken, öğrenilen programda kodların neredeyse yarısı "hatırlama" düzeyindedir (%44). Katılımcılar, kazanımlarının yarısının bu bilişsel düzeyde olduğunu bildirirken öđretim üyeleri program belgelerinde bu bilişsel düzeye hiç yer vermemişlerdir. "İşlemsel beceriler" bilişsel düzeyinde ise hem hedeflenen programında (%26) hem de öğrenilen programda (%28) benzer güçlü bir vurgu vardır. İlginçtir ki, öğrenilen programda, daha yüksek bilişsel seviyeler, "uygulama" (%6), "analiz" (%4) ve "sentez" (%8), hedeflenen programdan çok daha az yer almıştır. Hedeflenen programda bu yüksek bilişsel düzeylerin oranı %20'dir.

Tablo 6. Hedeflenen ve Öğrenilen Programlar Arası Uyum Matrisi

İçerik	Hatırlama			İşlemsel beceriler			Uygulama			Analiz			Sentez			Toplam													
	Hedeflenen		Mutlak fark	Hedeflenen		Mutlak fark	Hedeflenen		Mutlak fark	Hedeflenen		Mutlak fark	Hedeflenen		Mutlak fark	Hedeflenen		Mutlak fark											
	D	OD (x)		D	OD (y)		D	OD (x)		D	OD (y)		D	OD (x)		D	OD (y)		D	OD (x)	D	OD (y)							
1	0	0	3	,12	,12	3	,06	2	,08	,02	1	,02	0	0	0	,02	0	0	0	0	3	,06	1	,04	,02	7	,13	6	,24
2	0	0	2	,08	,08	3	,06	1	,04	,02	3	,06	1	,04	,02	3	,06	0	0	,06	2	,04	0	0	,04	11	,21	4	,16
3	0	0	2	,08	,08	3	,06	2	,08	,02	3	,06	1	,04	,02	3	,06	0	0	,06	3	,06	1	,04	,02	12	,23	6	,24
4	0	0	2	,08	,08	2	,04	2	,08	,04	3	,06	1	,04	,02	3	,06	1	,04	,02	3	,06	0	0	,06	11	,21	6	,24
5	0	0	2	,08	,08	3	,06	0	0	,06	3	,06	1	,04	,02	3	,06	0	0	,06	3	,06	0	0	,06	2	,22	3	,12
Toplam	0	0	11	,44	,44	14	,26	7	,28	,16	13	,25	4	,16	,1	12	,23	1	,04	,2	14	,26	2	,8	,2	43	1	25	1

1) İngilizce öğretiminde ölçme yöntemleri ve ilkeleri; 2) Dil becerilerinin değerlendirilmesi ve sınav geliştirilmesi; 3) Alternatif ölçme değerlendirme; 4) Sınav sonuçlarının analizine dayalı öğrenci ölçme değerlendirme; 5) Kullanılan sınavların etkinliğinin değerlendirilmesi.

D: Değer (Value), OD: Oransal Değer (Proportional Value)

Öđretilen program ile öđrenilen program arasındaki ierik uyumunu gsteren drdnc uyum analizi (UI 4) 0,45 olarak bulunmuřtur. Konuların kategorik uyumuna bakıldıđında, her iki programında, “İngilizce lme yntemleri ve ilkeleri” ve “kullanılan testlerin etkililiđinin deđerlendirilmesi” dıřındaki ierik kategorilerini benzer oranlarda ierdiđi grlebilir. “Kullanılan testlerin etkililiđinin deđerlendirilmesi” ieriđi, đretilen programda (%6), đrenilen programda olduđundan (%24) ok daha az yer almıřtır (Bkz. Tablo 7). Ancak “İngilizce lme yntem ve ilkeleri” teması đretilen programda (%23), đrenilen programda olduđundan (%12) daha fazla ele alınmıřtır. İeriklerdeki biliřsel dzey farklılıkları da nispeten dřk uyum indeksine yol amıř olabilir. “Hatırlama” biliřsel dzeyi đretilen programda ok sınırlı temsil edilirken (%8), đrenilen programda daha yksek oranlarda (%44) temsil edilmiřtir. Benzer řekilde “iřlemsel beceriler” biliřsel dzeyi đrenilen programda %28, đretilen programda ise ok daha az bir oranda (%16) yer almıřtır. İlgintir ki, diđer biliřsel dzeyler, zellikle en yksek iki dzey olan “analiz” ve “sentez”, đrenilen programda ok az vurgulanırken (sırasıyla %4 ve %8), đretilen programda her ikisi de %24 gibi ok daha fazla bir oranda yer almıřtır. Bu durum, đretim yelerinin sınıfta đrettikleri ile đretmen adaylarının đrendikleri arasında biliřsel dzey aısından farklar olduđunu iřaret etmektedir.

Tablo 7. Öğretilen ve Öğrenilen Programlar Arası Uyum Matrisi

İçerik	Hatırlama			İşlemsel beceriler			Uygulama			Analiz			Sentez			Toplam													
	Öğretilen	Öğrenilen	Mutlak fark	Öğretilen	Öğrenilen	Mutlak fark	Öğretilen	Öğrenilen	Mutlak fark	Öğretilen	Öğrenilen	Mutlak fark	Öğretilen	Öğrenilen	Mutlak fark	Öğretilen	Öğrenilen												
	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)	x-y	D	OD	D	OD					
1	3	,08	3	,12	,04	3	,08	2	,08	0	0	0	0	0	0	0	0	0	0	0	0	0	1	,04	,04	6	,16	6	,24
2	0	0	2	,08	,08	3	,08	1	,04	,04	3	,08	1	,04	,04	0	0	0	0	0	0	0	0	0	0	6	,16	4	,16
3	0	0	2	,08	,08	0	0	2	,08	,08	2	,05	1	,04	,01	3	,08	0	0	,08	3	,08	1	,04	,04	8	,21	6	,24
4	0	0	2	,08	,08	0	0	2	,08	,08	3	,08	1	,04	,04	3	,08	1	,04	,04	3	,08	0	0	,08	9	,23	6	,24
5	0	0	2	,08	,08	0	0	0	0	0	3	,08	1	,04	,04	3	,08	0	0	,08	3	,08	0	0	,08	9	,23	3	,12
Toplam	3	,08	11	,44	,36	6	,16	7	,28	,2	11	,29	4	,16	,13	9	,24	1	,04	,2	9	,24	2	,08	,24	38	1	25	1

1) İngilizce öğretiminde ölçme yöntemleri ve ilkeleri; 2) Dil becerilerinin değerlendirilmesi ve sınav geliştirilmesi; 3) Alternatif ölçme değerlendirme; 4) Sınav sonuçlarının analizine dayalı öğrenci ölçme değerlendirmesi; 5) Kullanılan sınavların etkinliğinin değerlendirilmesi.

D: Değer (Value), OD: Oransal Değer (Proportional Value)

Öđretim üyelerinin deđerlendirilen programı ile öđretmen adaylarının öđrenilen programı arasındaki uyumuna iliřkin beřinci uyum analizi (UI 5) 0,78 olarak hesaplanmıřtır. Sonular, her iki programın da tüm ierik kategorilerine olduka dengeli vurgu yaptığını, dolayısıyla aralarında anlamlı bir fark olmadığını gstermektedir (Bkz. Tablo 8). Tek fark, öđrenilen programda “dil becerilerinin deđerlendirilmesi ve test geliřtirme” temasının (%16), deđerlendirilen programa (%22) gre daha az vurgulanmasıdır. Ayrıca, öđrenilen program ile deđerlendirilen program arasında biliřsel seviyeler aısından da bazı kk farklılıklar vardır. Deđerlendirilen programdaki ieriđin, öđrenilen program ieriđine gre daha yksek biliřsel dzeylerde olduđu grlmektedir. rnek vermek gerekirse, deđerlendirilen programda, “iřlemsel beceriler” %42'lik bir orana sahipken, aynı biliřsel dzey öđrenilen programda yine yksek ancak deđerlendirilen programdaki orandan daha dřk bir temsile sahiptir (%28). Benzer řekilde, “uygulama” ve “analiz” biliřsel dzeyleri, deđerlendirilen programda (sırasıyla %18, %6) öđrenilen programdan (sırasıyla %16, %4) daha fazla yer almıřtır. Ancak, deđerlendirilen programda “sentez” dzeyinde hi ierik bulunmazken, öđrenilen programda ieriđin %8'i bu biliřsel dzeyde ifade edilmiřtir.

Tablo 8. Değerlendirilen ve Öğrenilen Programlar Arası Uyum Matrisi

İçerik	Hatırlama			İşlemsel beceriler			Uygulama			Analiz			Sentez			Toplam														
	Değerlendirilen	Öğrenilen		Değerlendirilen	Öğrenilen		Değerlendirilen	Öğrenilen		Değerlendirilen	Öğrenilen		Değerlendirilen	Öğrenilen		Değerlendirilen	Öğrenilen													
	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)											
1	3	,08	3	,12	,04	3	,08	2	,08	0	1	,03	0	0	,03	0	0	0	0	0	0	1	,04	,04	7	,19	6	,24		
2	3	,08	2	,08	0	3	,08	1	,04	,04	2	,06	1	,04	,02	0	0	0	0	0	0	0	0	0	0	0	8	,22	4	,16
3	2	,06	2	,08	,02	3	,08	2	,08	0	2	,06	1	,04	,02	0	0	0	0	0	0	1	,04	,04	7	,19	6	,24		
4	3	,08	2	,08	0	3	,08	2	,08	0	1	,03	1	,04	,01	1	,03	1	,04	,01	0	0	0	0	0	0	8	,22	6	,24
5	2	,06	2	,08	,02	3	,08	0	0	,08	0	0	1	,04	,04	1	,03	0	0	,03	0	0	0	0	0	0	6	,17	3	,12
Toplam	13	,36	11	,44	,08	15	,42	7	,28	,12	6	,17	4	,16	,12	2	,06	1	,04	,04	0	0	2	,08	,08	36	1	25	1	

1) İngilizce öğretiminde ölçme yöntemleri ve ilkeleri; 2) Dil becerilerinin değerlendirilmesi ve sınav geliştirilmesi; 3) Alternatif ölçme değerlendirme; 4) Sınav sonuçlarının analizine dayalı öğrenci ölçme değerlendirme; 5) Kullanılan sınavların etkinliğinin değerlendirilmesi.

D: Değer (Value), OD: Oransal Değer (Proportional Value)

Altıncı uyum analizi (UI 6), öđretim üyelerinin öđretilen programı ile deđerlendirilen programı arasındaki ierik uyumunu gstermektedir. Bu programlar arası uyum 0,44 olarak bulunmuřtur. Bu, öđretim üyelerinin sınıfta yaptıkları ile sınavlarda test ettikleri arasında orta düzeyde bir uyum olduđunu gstermektedir. Bunun nedeni hem ierik konuları hem de biliřsel düzeylerdeki farklılıklardır. Öđretilen programda, “İngilizce ölçme yöntem ve ilkeleri” ile “dil becerilerinin deđerlendirilmesi ve sınav geliştirilmesi” dışındaki ierik konuları deđerlendirilen programda olduđundan daha fazla yer almıřtır. Bu iki konu ise öđretilen programın %16’sında yer alırken, deđerlendirilen programda %19 ve %22 gibi biraz daha yüksek oranlarda temsil edilmiřtir (Bkz. Tablo 9). Biliřsel düzey aısından ise, deđerlendirilen programdaki ieriđin çođu daha düşük biliřsel düzeyde yer alırken, öđretilen programda ierik daha yüksek biliřsel düzeyde yer almıřtır. Deđerlendirilen programda “hatırlama” (%6) ve “işlemsel beceriler” (%42) oldukça yüksek oranlarda temsil edilirken, öđretilen programda bu biliřsel düzeyler sırasıyla %8 ve %16 oranlarında yer almıřtır. Bununla birlikte, “uygulama”, “analiz” ve “sentez” gibi daha yüksek biliřsel düzeyler öđretilen programda, oldukça yüksek oranlarda (%20’nin üzerinde) ele alınmıřtır. Fakat deđerlendirilen programda, %17’lik bir orana sahip “uygulama” düzeyi dışında, “analiz” düzeyi %6 gibi düşük bir oranda temsil edilmiř ve “sentez” düzeyi ise hi temsil edilmemiřtir.

Tablo 9. Öğretilen ve Değerlendirilen Programlar Arası Uyum Matrisi

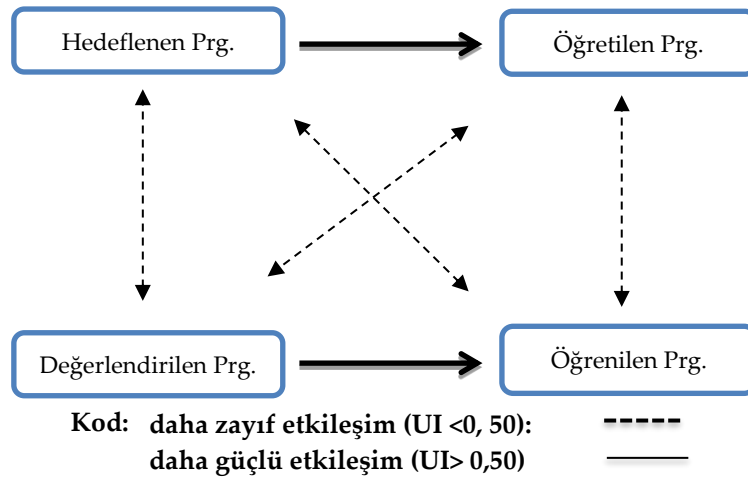
İçerik	Hatırlama			İşlemsel beceriler			Uygulama			Analiz			Sentez			Toplam														
	Öğretilen	Değerlendirilen	Mutlak fark	Öğretilen	Değerlendirilen	Mutlak fark	Öğretilen	Değerlendirilen	Mutlak fark	Öğretilen	Değerlendirilen	Mutlak fark	Öğretilen	Değerlendirilen	Mutlak fark	Öğretilen	Değerlendirilen													
	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)	x-y	D	OD (x)	D	OD (y)						
1	3	,08	3	,08	0	3	,08	3	,08	0	0	0	1	,03	,03	0	0	0	0	0	0	0	0	0	0	0	6	,16	7	,19
2	0	0	3	,08	,08	3	,08	3	,08	0	3	,08	2	,06	,02	0	0	0	0	0	0	0	0	0	0	6	,16	8	,22	
3	0	0	2	,06	,06	0	0	3	,08	,08	2	,05	2	,06	,01	3	,08	0	0	,08	3	,08	0	0	,08	8	,21	7	,19	
4	0	0	3	,08	,08	0	0	3	,08	,08	3	,08	1	,03	,05	3	,08	1	,03	,05	3	,08	0	0	,08	9	,24	8	,22	
5	0	0	2	,06	,06	0	0	3	,08	,08	3	,08	0	0	,08	3	,08	1	,03	,05	3	,08	0	0	,08	9	,24	6	,17	
Toplam	3	,08	13	,36	,28	6	,16	15	,42	,24	11	,29	6	,17	,19	9	,24	2	,06	,18	9	,24	0	0	,24	38	1	36	1	

1) İngilizce öğretiminde ölçme yöntemleri ve ilkeleri; 2) Dil becerilerinin değerlendirilmesi ve sınav geliştirilmesi; 3) Alternatif ölçme değerlendirme; 4) Sınav sonuçlarının analizine dayalı öğrenci ölçme değerlendirme; 5) Kullanılan sınavların etkinliğinin değerlendirilmesi.

D: Değer (Value), OD: Oransal Değer (Proportional Value)

Hedeflenen, Öğretilen, Öğrenilen ve Değerlendirilen Programlar Arasındaki Etkileşim

Çalışmanın bir diğer önemli bulgusu, aynı dersin farklı program türleri arasındaki etkileşim örüntüsüyle ilgidir. Program çiftleri arasındaki uyum indeksleri, Şekil 2'de gösterilen program türleri arasındaki etkileşim örüntüsünün ortaya çıkmasını sağlamıştır. Şekilde görüldüğü gibi, hedeflenen ve öğretilen programlar arasında güçlü bir etkileşim varken (UI 1 > 0,50), hedeflenen ve değerlendirilen ve hedeflenen ve öğrenilen programlar arasında orta düzey bir etkileşim vardır (UI 2 ve UI 3 < 0, 50). Öğretilen programın hem değerlendirilen hem de öğrenilen programlar ile orta düzeyde bir etkileşimi olduğu da görülmektedir (UI 6 ve UI 4 < 0, 50). Değerlendirilen programın ise hem hedeflenen hem de öğretilen programla orta düzeyde bir etkileşimi vardır (UI 2 ve UI 6 < 0, 50). Öğrenilen programın, değerlendirilen programla olan ilişkisi (UI 5 > 0, 50), öğrenilen programla hedeflenen veya öğretilen programlarla (UI 3 ve UI 4 < 0, 50) olan ilişkisinden daha güçlüdür.



Şekil 2. Program türleri arasındaki etkileşim

Tartışma

Uyum indeksleri, hizmet öncesi ölçme değerlendirme dersinin hedeflenen, öğretilen, öğrenilen ve değerlendirilen programları arasındaki uyum düzeyi hakkında bilgi vermenin yanı sıra bu program türleri arasındaki etkileşim hakkında da bilgi vermektedir. Araştırmanın bulguları, öğretim üyelerinin hedeflediği ve öğrettiği programlar arasındaki uyumun (0,67), hedefledikleri ve değerlendirdikleri programlar arasındaki uyumun (0,48), öğrettikleri ve değerlendirdikleri programlar arasındaki uyumdan (0,44) daha yüksek olduğunu göstermektedir. Öğretim üyelerinin planladığı ve öğrettiği programlar arasındaki uyum analizi, öğretim görevlilerinin kendi planlamış oldukları öğretim içeriğini ne kadar uyguladıkları hakkında bilgi sağlamıştır. Uyum indeksi, öğretim üyelerinin planladıkları hedef içeriklerin %67'sini öğrettiğini ve bunun da yüksek bir uyum olduğunu göstermiştir. Çalışmanın bu bulgusu, Kurz ve meslektaşlarının bulgularıyla çelişmektedir. Araştırmacılar, Öğretilen Program Anketleri (ÖPA) aracılığıyla sekizinci sınıf matematik dersinin genel ve özel eğitimde hedeflenen, öğretilen ve değerlendirilen programları arasındaki uyumu araştırmış ve araştırma sonunda, hedeflenen ve öğretilen programlar arasında düşük bir uyum bulmuştur. Araştırmacılar, çalışma sonunda, katılımcı öğretmenlerin yarısının planladıkları programın yarısından azını uyguladıklarını ifade etmiştir (Kurz vd., 2010).

Araştırmanın bulguları öğretim üyelerinin hedeflenen programlarındaki içeriğin yaklaşık yarısını (%48) ölçme araçlarıyla ölçtüklerini ve sınıfta öğrettiklerinin yarısından azını (%44) ölçtüklerini göstermiştir. Hizmet öncesi ölçme değerlendirme eğitiminin hedeflenen ve değerlendirilen programları ve öğretilen ve değerlendirilen programları arasındaki nispeten daha düşük uyum, öğretmen adaylarının sahip olduğu düşük ölçme değerlendirme okuryazarlık düzeyinin nedeni olabilir. Schmidt ve diğerlerinin (2001) iddia ettiği gibi, hedeflenen ve değerlendirilen programlar ile öğretilen ve değerlendirilen programlar arasındaki uyum, öğrenci başarısını artırmaktadır.

Uyum indeksleri, öğretim üyelerinin değerlendirilen programı ile öğretmen adaylarının öğrendikleri program arasındaki uyumun oldukça yüksek olduğunu (0, 78), fakat öğretim üyelerinin hedeflenen programıyla öğretmen adaylarının öğrenilen programı (0, 45) ve öğretim üyelerinin öğretilen programı ve öğretmen adaylarının öğrenilen programları arasındaki uyumun (0, 44) daha düşük olduğunu göstermiştir. Bu programlar arasındaki nispeten düşük uyum, esas olarak içerik temalarındaki farklılıklardan değil, yoğunlukla içerdikleri bilişsel seviyelerdeki farklılıklardan kaynaklanmaktadır. Öğretim üyelerinin hedefledikleri ve öğrettikleri programlarda içeriğin nispeten yüksek bilişsel seviyelerde yer aldığı görülürken, öğretmen adaylarının içeriği yoğunlukla daha düşük bilişsel düzeylerde öğrendikleri tespit edilmiştir. Diğer bir deyişle, öğretim üyeleri kazanımlarını belirlerken ve sınıfta kullanacakları öğretim etkinliklerini düzenlerken nispeten daha yüksek bilişsel düzeyde tasarlama eğilimindeyken öğretmen adaylarının öğrenmelerini ölçerken daha düşük bilişsel seviyelerde ölçtükleri tespit edilmiştir. Çalışmanın bu bulgusu önceki araştırmalarla uyumludur. Dokuzuncu sınıf matematik dersinde hedeflenen, öğretilen ve değerlendirilen programlar arasındaki tutarlılığı inceleyen Seitz (2017), programın içerik açısından uyumlu olduğunu ancak bilişsel düzey açısından uyumunun zayıf olduğunu bulmuştur. Benzer şekilde Fuhrman (2001), test maddelerinin, “değerlendirme” ve “sentez” gibi üst düzey bilişsel süreçler yerine, “hatırlama” ve “anlama” gibi daha düşük bilişsel düzeyleri ölçtüğünü öne süren benzer bulgular ortaya koymuştur.

Öğretmenlerin planladıkları, uyguladıkları ve değerlendirdikleri programlar ile öğrencilerin öğrendikleri programlar arasında farklılık olması normaldir. Ancak, belirli bir programın sürekli belirli bilişsel düzeyleri içermesi ve diğer bilişsel düzeyleri içermemesi endişe vericidir. Bu durum, bu çalışmada incelenen programın benzer bir sorunu olduğunu göstermektedir. Örnek vermek gerekirse, hedeflenen program daha yüksek bilişsel seviyeleri fazlaca içerirken, “hatırlama” seviyesini hiç içermemektedir. Öğretilen program, “uygulama”, “analiz” ve “sentez” gibi yüksek bilişsel seviyeleri çok fazla içerirken, “hatırlama” seviyesine yeteri kadar yer vermemiştir. Değerlendirilen program, “hatırlama” ve “işlemsel becerileri” fazlaca içerirken “analiz” ve “sentez” seviyelerine yeteri kadar yer vermemiştir. Bu durum, öğretim üyelerinin planlama, öğretme ve ölçme faaliyetlerinde bilişsel düzeylere yaptıkları vurgunun tutarsız olduğunu ve daha da kötüsü, öğretmen adaylarının öğrenmelerinin daha çok düşük bilişsel düzeylerde gerçekleştiğini ortaya koymaktadır. Hedeflenen ve değerlendirilen programlarda alt düzey bilişsel becerilere yapılan aşırı vurgu, öğretmen adaylarının öğrenmesini olumsuz etkilemiş olabilir. Bu nedenle, öğretmen adaylarının performanslarını iyileştirmek için, konuları sadece “hatırlama” veya “işlemsel beceri” seviyelerinde ele almak yeterli olmayabilir. İçerik, aynı zamanda üst düzey bilişsel seviyelerde öğretilmelidir ve sınavlarda da aynı düzeylerde test edilmelidir. Diğer bir deyişle, öğrenenlerin hem alt hem de üst bilişsel düzey becerilerini geliştirmeleri için, ders içeriğinin yalnızca düşük düzeylerde değil, yüksek bilişsel düzeylerde de öğretilmesi ve test edilmesi gerekir.

Öğrenilen programda, düşük bilişsel düzeyin fazla bulunması, öğretmen adaylarının yetersiz öğrenmelerini açıklayabilir. Bunun nedeni, ders programında hedeflenen, sınıfta öğretilen ve sınavlarda test edilen içeriğin farklı bilişsel düzeylerde olmasından kaynaklanan kafa karışıklığı olabilir. Ancak araştırmacı, öğretmen adaylarının gelecekteki mesleklerine iyi hazırlanabilmeleri için üst düzey bilişsel becerilere ihtiyaç duyduklarının altını çizmektedir. Diğer bir deyişle, İngilizce öğretiminde ölçme değerlendirme hakkında öğrendiklerini uygulayabilmeleri gerekir. Bu durum, öğrenme hedeflerinin, öğretme/öğrenme etkinliklerin ve performans değerlendirme görevlerinin daha yüksek bilişsel seviyelerde tasarlanmasını gerektirir. Bu çalışmada incelenen hizmet öncesi ölçme değerlendirme ders programının bu beklentiyi karşılaması için bazı değişikliklere ihtiyacı duyduğu görülmektedir. Örneğin, öğretim üyeleri, hedeflenen ve değerlendirilen programlarında, yalnızca “hatırlama” ve “işlemsel beceriler” gibi düşük bilişsel seviyeleri vurgulamakla kalmamalı, aynı zamanda “analiz” ve “sentez” gibi daha yüksek bilişsel seviyeleri de yeterince içermelidir. Daha yüksek bilişsel seviyelerde oluşturulmuş öğrenme hedeflerinin ve performans görevlerinin sayısının artırılması kesinlikle gereklidir. Bu amaçla, öğretim üyeleri, programlarını ve ölçme değerlendirme araçlarını gözden geçirmeli ve geliştirmelidir.

Araştırmanın bir diğer önemli bulgusu ise, hizmet öncesi ölçme değerlendirme eğitiminin farklı programları arasındaki etkileşimdir. Uyum indeksleri, hedeflenen ve öğretilen program ile değerlendirilen ve öğrenilen programlar arasında güçlü bir etkileşim olduğunu, ancak hedeflenen ve değerlendirilen, öğretilen ve değerlendirilen; öğretilen ve öğrenilen ve hedeflenen ve öğrenilen program çiftleri arasında daha düşük bir etkileşim olduğunu göstermektedir. Bu çalışmanın tespit ettiği hedeflenen ve öğretilen programlar arasındaki güçlü iletişim, Glatthorn ve diğerlerinin (2001) araştırma sonuçlarıyla çelişmektedir. Araştırmacılar çalışmalarında, hedeflenen programın öğretilen program üzerinde kısmi bir etkisi olduğunu iddia etmişlerdir. Araştırmacılara göre, çoğu öğretmenin, özellikle de deneyimli olanların, ne öğreteceklerine karar verirken sınıfla ilgili diğer faktörleri de dikkate almakta ve program belgesini sene başında kontrol edip sonra unutmaktadır. Bu nedenle, hedeflenen programın öğretmenlerin sınıfta yaptıkları üzerinde güçlü bir etkisi olmadığını söylemişlerdir. Benzer şekilde, UNICEF'in resmi bir belgesinde, her ikisi de aynı öğretmene ait hedeflenen ve öğretilen programlar arasında farklılık olmasının çok doğal olduğu belirtilmiştir. Bunun nedeni olarak ise hedeflenen programın yeterli olmadığı durumlarda, öğretmenlerin, öğrencilerinin belirli performans standartlarına ulaşması için farklı stratejiler geliştirdikleri söylenmiştir (UNICEF, 2000).

Araştırmacı, bu çalışmada, Glatthorn ve diğerlerinin (2001) çalışmasından farklı olarak hedeflenen ve öğretilen programlar arasında güçlü bir ilişki bulmuştur. Bunun nedeni, araştırmanın bir öğretmen eğitimi programında, yani bir yükseköğretim kurumunda yapılmış olması olabilir. Öğretim üyeleri, dışarıdan bir otoritenin baskısı olmaksızın hedeflenen veya yazılı programı kendileri hazırlamaktadır. Özellikle K-12 bağlamında, bölge ve okul yöneticileri, hedeflenen programı sınıfta neyin uygulanacağını kontrol etmek için bir araç olarak kullanabilir, bu nedenle K-12'de çalışan öğretmenler, sınıf içi öğretim uygulamalarını dışardan dayatılan yaptırımlardan çok, deneyimlerine ve sezgilerine dayandırmayı tercih edebilirler. Ancak bu çalışmada incelenen araştırma alanı bir yükseköğretim kurumu olduğu için hedeflenen/yazılı program, dersi veren öğretim elemanları tarafından tasarlanmıştır. Bu durum, bulgular arasındaki farklılıkları açıklayabilir. Sonuç olarak, araştırmanın bulguları, yazılı programın öğretmenler tarafından hazırlandığı durumlarda, yazılı program ile öğretilen (uygulanan) programlar arasında daha güçlü bir uyum olduğunu göstermektedir. Araştırmanın bu bulgusunun literatüre katkı sağlaması beklenmektedir.

Araştırmanın bir başka bulgusu, hedeflenen programın değerlendirilen program üzerinde zayıf bir etkiye sahip olmasıdır. Hedeflenen program, öğretim üyelerinin, öğrencilerin dersin sonunda başarımlarını bekledikleri içeriği ifade eder. Bu nedenle, hazırladıkları sınavlarda öğretim üyelerinin hedefledikleri kazanımlara yer vermeleri çok muhtemeldir. Değerlendirilen program, hedeflenen programın iyi bir temsilcisi olmalıdır, bu nedenle değerlendirilen programın içeriği, hedeflenen programdaki içerikle iyi eşleşmelidir. Aslında, bu çalışmada bulunan zayıf uyum, içerik kapsamındaki farklılıklardan değil, bilişsel seviyelerdeki farklılıklardan kaynaklanmaktadır. Bunun nedeni, öğrenme beklentileri yazılırken, bilişsel seviyelerinin doğru bir şekilde ifade edilmemesi olabilir. Öğretim üyeleri, ders programında ele alınacak içerik kategorilerini oluştururken derste amaçladıkları çıktıların hepsinin akıllarında olduğunu varsayıyor olabilirler. Bu durum, hedeflenen programın içerik konularının neden genelde "hatırlama" düzeyinde ele alındığını ve değerlendirilen programda ortaya çıkan diğer bilişsel talepleri yeterince yansıtmadığını açıklayabilir.

Çalışma aynı zamanda öğretilen programın değerlendirilen program üzerinde zayıf bir etkisi olduğunu işaret etmektedir. Normal sınıflarda, öğretmenler sınıfta öğrettiklerini değerlendirme eğilimindedir. Böylece en çok ders saatlerinin ayrıldığı içeriğin sınavlarda test edilmesi beklenmektedir. Ancak bu çalışma, hizmet öncesi ölçme değerlendirme dersinde durumun böyle olmadığını ortaya koymuştur. Bu durum yine araştırma sahasının özelliklerinden kaynaklanıyor olabilir. Araştırma sahası bir yükseköğretim kurumu, özellikle bir öğretmen eğitimi programı olduğu için, ölçme değerlendirme araçlarını tasarlayanlar dersi veren öğretim üyeleridir. Dersi veren öğretim üyeleri, farklı amaçlara yönelik farklı değerlendirme görevleri kullanıyor olabilirler. Diğer birçok öğretmen eğitimi programında olduğu gibi, incelenen programda da sınıf sunumları öğretmen adaylarının değerlendirilmelerinin bir parçasıdır. Öğretim üyeleri, öğretmen adaylarını gelecekteki mesleklerine

hazırlamak ve sözlü sunum becerilerini geliştirmek amacıyla farklı değerlendirme yöntemleri kullanılabilirler. Bu durum, öğretim üyelerinin sınıfta öğrettikleri ile ölçtükleri arasındaki uyumsuzluğu açıklayabilir.

Çalışma ayrıca, değerlendirilen program ve öğrenilen program arasında daha güçlü bir ilişki olduğunu bulmuştur. Bu durum önceki çalışma sonuçlarıyla eşleşmektedir. Glatthorn (2000), çalışmasında, iki program arasında güçlü bir uyum olduğunu iddia etmiştir ve sınavlarda değerlendirilen içeriğin öğrenciler tarafından öğrenilmesinin daha olası olduğunu eklemiştir. Ayrıca, Glatthorn'a göre, öğrenciler sadece sorumlu tutuldukları konulara önem vermektedir. Bu nedenle, öğretmenin belirlediği hedef ne olursa olsun veya öğretmen sınıfta neyi vurgularsa vurgulasın, öğrenciler yalnızca sınavlarda test edilen içeriği hatırlama eğilimindedir.

Öğretilen programın öğrenilen program üzerindeki zayıf etkisi ile ilgili olarak Marsh ve Willis (2003), öğrencilerin resmi programı tam olarak nasıl anladıklarının hem önceki sınıf içi öğrenimleriyle hem de medya veya başka kaynaklar aracılığıyla elde ettikleri sınıf dışı öğrenimleriyle ilgili olduğunu belirtmiştir. Farklı deneyimlere, sosyopolitik etkilere ve analitik tutumlara maruz kaldıkları için, öğrencilerin aynı dersten farklı kazanımlar edinmesi muhtemeldir. Benzer şekilde, Glatthorn ve diğerleri (2001) çalışmalarında öğretilen ve öğrenilen programlar arasındaki farklılıktan bahsetmiş ve zayıf motivasyon, sınırlı bilişsel yetenek veya öğrencilerin kısa dikkat süresi ve öğretmenlerin öğrenciyi izleme konusundaki eksiklikleri ve programı öğrencilerin öğrenmesi için anlamlı ve zorlayıcı hale getirmekteki yetersizlikleri gibi çeşitli faktörlerin bunun nedeni olabileceğini belirtmişlerdir.

Kısaca, bu çalışma, hedeflen ve öğretilen ile değerlendirilen ve öğrenilen programlar arasında güçlü bir uyum tespit ederken, hedeflenen ve değerlendirilen, hedeflenen ve öğrenilen ve öğretilen ve öğrenilen programlar arasında orta düzeyde bir uyum bulmuştur. Ancak programlar arasındaki farklılıklar içerikle ilgili değil, esas olarak bilişsel düzeylerle ilgilidir. Bu durum, öğretim üyelerinin, öğrenme beklentilerini planlarken, öğretirken ve ölçme değerlendirme görevlerini tasarlarken bilişsel düzeylere daha fazla dikkat etmeleri gerektiğini göstermektedir. Bununla birlikte, hedefler, sınıf etkinlikleri ve ölçme değerlendirme görevlerini yalnızca düşük bilişsel düzeyde değil, aynı zamanda öğretmen adaylarının daha iyi öğrenmesine ve performans göstermesini sağlayacak yüksek bilişsel düzeylerde oluşturmaları gerektiği anlaşılmaktadır. İhtiyaç duyulan şey hem düşük hem de yüksek seviyelerdeki bilişsel becerilerin dengeli bir şekilde öğretilmesi ve değerlendirilmesidir.

Sonuç

Program uyumu ile ilgili araştırmalar, öğretmen adaylarına sunulan hedeflerin, sınıf içi öğretim faaliyetlerinin ve ölçme değerlendirme araçlarının birbiri ile uyumlu olduğu ve tutarlı mesajlar verdiği durumlarda, öğretmen yetiştirmede daha yüksek başarı sağlayacağını ifade etmektedir (Elliott vd., 2001; Webb, 1997a, 1997b). Bunun tam tersi, öğretmen adaylarına uyumlu öğrenme deneyimleri sunulmaz ise, okuryazarlık ve yeterlik seviyeleri düşük olabilmektedir (Darling-Hammond, 2006; Howey & Zimpher, 1989; Russell & Mcpherson, 2001). Ulusal bağlamda, birçok araştırmacı İngilizce öğretmenlerinin ölçme değerlendirme yeterliği ve performansı ile ilgili bir soruna işaret etmiştir (Hatipoğlu, 2015; Karaman ve Şahin, 2014; Mede ve Atay, 2017; Mertler, 2003; Öz ve Atay, 2017). Ancak, İngilizce öğretmenlerinin ölçme değerlendirme okuryazarlık düzeylerinin düşük olmasının program uyumuyla ilgili bir problemten kaynaklanıp kaynaklanmadığını araştıran bir çalışma bulunmamaktadır. Bu açıdan, bu çalışmanın amacı, Türkiye'nin İç Anadolu bölgesinde bir İngilizce Öğretmenliği programı tarafından verilen hizmet öncesi ölçme değerlendirme eğitiminin hedeflenen, öğretilen, öğrenilen ve değerlendirilen programları arasındaki uyumu incelemektir. Bu çalışmada araştırmacı, öğrenme hedefleri, öğretme etkinlikleri ve ölçme değerlendirme gibi tüm program öğelerinin tutarlı olduğu ve nihayetinde öğretmen adaylarının öğrenmesini desteklediği uyumlu bir öğretmen eğitimi programının önemini vurgulamaya çalışmıştır.

Çalışma hem nitel hem de nicel veri toplama araçlarını kullanan karma yöntemli bir durum çalışması yaklaşımını benimsemiş ve Öğretilen Program Anketleri (ÖPA) uyum metodolojisinin (Porter ve Smithson, 2000, 2001, 2002) bir uyarlamasını kullanmıştır. Hedeflenen ve değerlendirilen program

içerikleri, yazılı program belgelerinin ve ölçme değerlendirme araçlarının, konu uzmanları tarafından her bir konunun içerik kapsamı ve bilişsel düzeyleri açısından kodlanmasıyla oluşturulmuştur. Öğretilen ve öğrenilen programın içeriğine ilişkin bilgiler, katılımcı öğretim üyelerinden ve öğretmen adaylarından toplanmıştır. Öğretim üyeleri anket yöntemiyle öğretim uygulamalarının içeriği hakkında bilgi vermişlerdir. Öğretmen adayları da anket aracıyla dersin sonunda hedeflenen kazanımları edinme düzeylerini rapor etmişlerdir. Bu sürecin sonunda ortaya çıkan tüm kodlar uyum matrislerine aktarılmıştır. Matrislerdeki her bir verinin matristeki tüm verilere oranını gösteren oransal değerler hesaplanmıştır (DeLuca ve Bellara, 2013). Bu yöntemle oluşturulan matrisler eşleştirilmiş ve program çiftleri arasındaki uyumu ölçmek için Porter'in uyum indeksi formülü kullanılmıştır (2002).

Çalışmanın sonunda, İngilizce öğretiminde ölçme değerlendirme dersinin hedeflenen ve öğretilen programları ile değerlendirilen ve öğrenilen programları arasındaki uyum indeksleri nispeten yüksek bulunmuştur. Ancak, hedeflenen ve değerlendirilen, öğretilen ve değerlendirilen, hedeflenen ve öğrenilen ve öğretilen ve öğrenilen programları arasındaki uyum orta düzeyde bulunmuştur. Bununla birlikte, programlar arası farklılıklar genelde içerik kapsamında değil, içeriğin bilişsel düzeyinden kaynaklanmaktadır. Bu durum, ders hedefleri, sınıfta uygulanan öğretim ve ölçme değerlendirme arasında sadece içerik kapsamı açısından değil, bilişsel düzeyler açısından da uyumun sağlanmasının öğretmen adaylarının öğrenmelerini destekleyeceğini göstermektedir. Araştırmanın bulguları, öğretim üyelerinin hedeflenen program ve ölçme değerlendirme araçlarını planlarken bilişsel düzeylere dikkat etmeleri gerektiğini ve bunları dersi işlerken de göz önünde bulundurmaları gerektiğini göstermektedir. Hedeflerin belirlenmesi, öğretim etkinlikleri ve ölçme değerlendirme araçlarının planlanması aşamasında düşük bilişsel düzeylerin yanı sıra yüksek bilişsel düzeyde içeriğe yer verilmesi öğretmen adaylarının daha üst düzeyde ölçme bilgisine sahip olmasını ve daha iyi öğrenmelerini sağlayacaktır.

Özet olarak, çalışmanın belirlediği uyum sorunlarının, özellikle de en kritik uyum türü olan öğretilen ve öğrenilen programlar arasındaki farklılıkların nedenlerinin ileriki araştırmalarda incelenmesi bu konuyla ilgili bildiklerimize katkı sağlayabilir. Bu çalışmada, hizmet öncesi ölçme değerlendirme eğitimi yalnızca zorunlu ölçme değerlendirme ders programı analiz edilerek incelenmiştir; ancak gelecekteki araştırmalar, öğretmen adaylarının ölçme değerlendirme öğrenebilecekleri diğer hizmet öncesi derslerin yanı sıra uygulama deneyimlerini de araştırabilir.

Kaynakça

- Bhola, D. S., Impara, J. C. ve Buckendahl, C. W. (2003). Aligning tests with states' content standards: Methods and issues. *Educational Measurement: Issues and Practices*, 22(3), 21-29.
- Biggs, J. B. (1999). *Teaching for quality learning at university*. Buckingham: Open University Press.
- Blank, R. K. (2002). Using surveys of enacted curriculum to advance evaluation of instruction in relation to standards. *Peabody Journal of Education*, 77(4), 86-121.
- Bloom, B. S. (1949, Kasım). *A taxonomy of educational objectives*. Examiner Konferansında sunulan bildiri, Monticello, IL.
- Bransford, J. D., Brown, A. L. ve Cocking, R. R. (2000). *How people learn: Brain, mind, experience, and school*. Washington DC: National Academy Press.
- Bruner, J. (1990). *Acts of meaning*. Cambridge, MA: Harvard University Press.
- Carolan, C. M., Forbat, L. ve Smith, A. (2016). Developing the DESCARTE model: The design of case study research in health care. *Qualitative Health Research*, 26(5), 626-639. doi:10.1177/1049732315602488
- Cohen, S. A. (1987). Instructional alignment: Searching for a magic bullet. *Educational Researcher*, 16(8), 16-20.
- Council of Chief State School Officers. (2006). *Alignment analysis*. http://www.ccsso.org/Projects/Alignment_Analysis/ adresinden erişildi.
- Creswell, J. W. ve Plano Clark, V. (2018). *Designing and conducting mixed methods research* (3. bs.). Thousand Oaks, CA: SAGE.
- Darling-Hammond, L. (2006). Constructing 21st-century teacher education. *Journal of Teacher Education*, 57(3), 1-15.
- DeLuca, C. ve Bellara, A. (2013). The current state of assessment education: Aligning policy, standards, and teacher education. *Journal of Teacher Education*, 64(4), 356-372.
- DeLuca, C. ve Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy and Practice*, 17(4), 419-438.
- DeLuca, C., LaPointe-McEwan, D. ve Luhanga, U. (2016). Approaches to classroom assessment inventory: A new instrument to support teacher assessment literacy. *Educational Assessment*, 21(4), 248-266. doi:10.1080/10627197.2016.1236677
- Dinther, M. V., Dochy, F. ve Segers, M. (2015). The contribution of assessment experiences to student teachers' self-efficacy in competence-based education. *Teaching and Teacher Education*, 49, 45-55. doi:10.1016/j.tate.2015.02.013
- EdSource. (2006). *Similar students, different results (SSDR): Why do some schools do better?*. http://www.edsource.org/pub_abs_simstu05.cfm adresinden erişildi.
- Elliott, S. N., Braden, J. P. ve White, J. (2001). *Assessing one and all: Educational accountability for students with disabilities*. Arlington, VA: Council for Exceptional Children.
- Fuhrman, S. (Ed.). (2001). *From the capitol to the classroom: Standards-based reform in the states: One-hundredth yearbook of the National Society for the Study of Education*. Chicago, IL: University of Chicago Press.
- Fulmer, G. W. (2011). Estimating critical values for strength of alignment among curriculum, assessments, and instruction. *Journal of Educational and Behavioural Statistics*, 36(3), 381-402.
- Glatthorn, A. A. (2000). *The principal as curriculum leader: Shaping what is taught and tested* (2. bs.). Thousand Oaks, Calif: Corwin Press.
- Glatthorn, A. A., Carr, J. F. ve Harris, D. E. (2001). *Planning and organizing for curriculum renewal*. <http://www.ascd.org/publications/curriculum-handbook/398.aspx> adresinden erişildi.

- Gooding, A. (1994). Recognition of effective group discussion. G. Bell, B. Wright, N. Leeson ve J. Geake (Ed.), *Challenges in mathematics education: Constraints on construction* içinde (s. 295-305). Lismore, NSW: MERGA.
- Graham, P. (2005). Classroom-based assessment: Changing knowledge and practice through pre-service teacher education. *Teaching and Teacher Education*, 21(6), 607-612.
- Hatipoğlu, Ç. (2015). English language testing and evaluation (ELTE) training in Turkey: Expectations and needs of pre-service English language teachers. *ELT Research Journal*, 4(2), 111-128. <https://dergipark.org.tr/tr/pub/eltrj/issue/28780/308006> adresinden erişildi.
- Howey, K. R. ve Zimpher, N. L. (1989). Pre-service teacher educators' role in programs for beginning teachers. *The Elementary School Journal*, 89(4), 450-470.
- Karaman, P. ve Şahin, Ç. (2014). Öğretmen adaylarının ölçme değerlendirme okuryazarlıklarının belirlenmesi. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 15(2), 175-189.
- Kercheval, A. (2001). *A case study of key effective practices in Ohio's improved school districts*. Bloomington, IN: Indiana Center for Evaluation. [http://www.schooleffects.com/gen/corp_generated_bin/documents/basic_module/Indiana_Effecti ve_Practices.pdf](http://www.schooleffects.com/gen/corp_generated_bin/documents/basic_module/Indiana_Effecti_ve_Practices.pdf) adresinden erişildi.
- Kurz, A., Elliott, S., Wehby, J. ve Smithson, J. (2010). Alignment of the intended, planned, and enacted curriculum in general and special education and its relation to student achievement. *Journal of Special Education*, 44(3), 131-145.
- Lincoln, Y. ve Guba, E. (1985). *Naturalistic inquiry*. Newbury Park, CA: Sage.
- Marsh, C. J. ve Willis, G. (2003). *Curriculum: Alternative approaches, ongoing issues* (3. bs.). Upper Saddle River, NJ: Prentice Hall.
- McFadden, L. (2009). District learning tied to student learning. *Phi Delta Kappan*, 90(8), 545-553.
- Mede, E. ve Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi*, 168(1), 1-5.
- Mellati, M., Khademi, M. ve Shirzadeh, A. (2015). The relationships among sources of teacher pedagogical beliefs, teaching experiences, and student outcomes. *International Journal of Applied Linguistics and English Literature*, 4(2), 177-184. doi:10.7575/aiac.ijalel.v.4n.2p.177
- Mertler, C. A. (2003). *Pre-service versus in-service teachers' assessment literacy: Does classroom experience make a difference?*. Orta Batı Eğitim Araştırmaları Derneği Yıllık Toplantısında sunulan bildiri, Columbus.
- Milli Eğitim Bakanlığı. (2017). *General competencies for teaching profession*. <http://oygm.meb.gov.tr> adresinden erişildi.
- Mullens, J. E. ve Gayler, K. (1999). *Measuring classroom instructional processes: Using survey and case study field test results to improve item construction* (NES 1990-08). Washington, DC: National Center for Education Statistics.
- Murphy, J. (2007). Restructuring through learning-focused leadership. H. J. Walberg (Ed.), *Handbook on Restructuring and Substantial School Improvement* içinde (s. 71-83). Charlotte, NC: Information Age Publishing. https://www.adi.org/pla/docs/Restructuring_Handbook.pdf adresinden erişildi.
- Öz, S. ve Atay, D. (2017). Turkish EFL instructors' in-class language assessment literacy: Perceptions and practices. *ELT Research Journal*, 6(1), 25-44.
- Pickart, A. J. (2018). *Research methods in information*. London: Facet Publishing.
- Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *Educational Researcher*, 31(7), 3-14.

- Porter, A. C. (2004). Curriculum assessment. J. Green, G. Camilli ve P. Elmore (Ed.), *Complementary methods for research in education*. Washington, DC: American Educational Research Association.
- Porter, A. C. ve Smithson, J. L. (2000, Nisan). *Alignment of state testing programs NAEP and reports of teacher practice in mathematics and science in Grades 4 and 8*. Amerikan Eğitim Araştırmaları Derneği yıllık toplantısında sunulan bildiri, New Orleans, LA.
- Porter, A. C. ve Smithson, J. L. (2001). Are content standards being implemented in the classroom? A methodology and some tentative answers. S. Fuhrman (Ed.), *From the capitol to the classroom: Standards-based reform in the states. One hundredth yearbook of the national society for the study of education* içinde (s. 60-80). Chicago: University of Chicago Press.
- Porter, A. C. ve Smithson, J. L. (2002, Nisan). *Alignment of assessments, standards, and instruction using curriculum indicator data*. Amerikan Eğitim Araştırmaları Derneği yıllık toplantısında sunulan bildiri, New Orleans, LA.
- Porter, A. C., Kirst, M. W., Osthoff, E., Smithson, J. L. ve Schneider, S. (1994). *Reform of high school mathematics and science and opportunity to learn*. Consortium for Policy Research in Education Policy Briefs. New Brunswick, NJ: Rutgers University, Consortium for Policy Research in Education.
- Resnick, L. B., Rothman, R., Slattery, J. B. ve Vranek, J. L. (2003). Benchmarking and alignment of standards and testing. *Educational Assessment*, 9(1), 1-27.
- Roach, A. T., Niebling, B. C. ve Kurz, A. (2008). Evaluating the alignment among curriculum, instruction, and assessments: Implications and applications for research and practice. *Psychology in the Schools*, 45(2), 158-176.
- Rowley, J. (2002) Using case studies in research. *Management Research News*, 25(1), 16-27. doi:10.1108/01409170210782990
- Russell, T. ve Mcpherson, S. (2001, May). *Indicators of success in teacher education: A review and analysis of recent research*. Pan-Kanada Eğitim Araştırmaları Gündem Sempozyumu Öğretmen Eğitimi/Eğitimci Eğitimi: Güncel Trendler ve Gelecek Yönergeler. Laval University, Quebec City, Canada.
- Schmidt, W. H., McKnight, C. C., Houang, R. T., Wang, H. C., Wiley, D. E., Cogan, L. S. ... Wolfe, R. G. (2001). *Why schools matter: A cross-national comparison of curriculum and learning*. San Francisco: Jossey-Bass.
- Schuenemann, D., Jones, D. ve Brown, M. (2011). The impact of a curriculum model on the mathematics and science achievement of economically disadvantaged students. *National Forum of Educational Administration and Supervision Journal*, 29(1), 63- 87.
- Seitz, P. (2017). Curriculum alignment among the intended, enacted and assessed curricula for grade 9 mathematics. *Journal of the Canadian Association for Curriculum Studies (JCACS)*, 15(1), 72-94.
- Shavelson, R. J., Webb, N. M. ve Burstein, L. (1986). Measurement of teaching. M. Wittrock (Ed.), *Handbook of research on teaching* (3. bs., s. 55-91). Washington, DC: American Educational Research Association.
- Smithson, J. L. (2007). *Report of alignment analyses of state content standards and assessment instruments* (Yayımlanmamış makale). University of Wisconsin-Madison, Madison.
- Smithson, J. L. ve Porter, A. C. (1994). *Measuring classroom practice: Lessons learned from efforts to describe the enacted curriculum. The Reform Up Close study* (CPRE Research Rep. Series No. 31). Madison: University of Wisconsin-Madison, Consortium for Policy Research in Education.
- Squires, D. A. (2012). Curriculum alignment research suggests that alignment can improve student achievement. *The Clearing House*, 85(4), 129-135.
- Stake, R. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.
- Stiggins, R. (2005). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, 18, 23-27.

- UNICEF. (2000). *Curriculum report card*. New York: Working Paper Series Education Section Programme Division.
- Webb, N., Herman, J. ve Webb, N. L. (2006). *Alignment of mathematics state-level standards and assessments: The role of reviewer agreement* (CSE Report 685). Los Angeles: University of California, National Center for Research on Evaluations, Standards, and Student Testing.
- Webb, N. L. (1997a). *Criteria for alignment of expectations and assessments in mathematics and science education* (Research Monograph No. 6). Madison: University of Wisconsin, National Institute for Science Education.
- Webb, N. L. (1997b). *Determining alignment of expectations and assessments in mathematics and science education* (NISE Brief, Vol. 1, No. 2). Madison: University of Wisconsin, National Institute for Science Education.
- Webb, N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied Measurement in Education*, 20(1), 7-25.
- Yükseköđretim Kurulu. (2016). *Eđitim fakóltesi öđretmen yetistirme lisans programları*. Ankara: Yüksek Öđretim Kurulu Başkanlıđı.
- Zavadsky, H. (2006). How NLCB drives success in urban schools. *Educational Leadership*, 64(3), 69-73.