

MADDE YAZMA EĞİTİMİNİN ÖZEL OKULDA ÇALIŞAN İLKÖĞRETİM İNGİLİZCE ÖĞRETMENLERİNİN MADDENİN KALİTESİNİ DEĞERLENDİRMEYE İLİŞKİN YETERLİKLERİNE ETKİSİ

ARAŞTIRMA MAKALESİ

Münibe KARTAL¹, Münevver İLGÜN DİBEK²

* Bu çalışma birinci yazarın tezsiz yüksek lisans dönem projesinden oluşturulmuştur.

1 Öğretmen, TED Kayseri Koleji, İngilizce, munibe.yazici@tedkayseri.k12.tr, ORCID: 0000-0002-6377-9977.

2 Dr. Öğretim Üyesi, TED Üniversitesi, Eğitim Bilimleri Bölümü, munevver.ilgun@tedu.edu.tr, ORCID: 0000-0002-7098-0118 .

Geliş Tarihi: 04.09.2021 Kabul Tarihi: 07.01.2022 DOI: 10.37669/milliegitim.991124

Öz: Kaliteli madde yazma eğitiminin özel okulda çalışan ilköğretim İngilizce öğretmenlerinin madde yazma yeterliliklerine olan etkisini belirlemek amacıyla yapılan bu çalışmada deneysel desenlerden tek grup ön-test son-test desen kullanılmıştır. Araştırmanın çalışma grubunu, kolay ulaşılabilir örnekleme yöntemi kullanılarak seçilmiş özel okulda ilköğretim kademesinde görev yapan 14 İngilizce öğretmeni oluşturmaktadır. Araştırmada veri toplama aracı olarak alan yazın taranarak oluşturulmuş olan kaliteli madde yazma kriterlerinin ölçüldüğü 19 maddeden oluşan bir test kullanılmıştır. Bu test, madde yazmayla ilgili hizmet içi eğitim verilmeden önce ön test ve hizmet içi eğitim verdikten sonra son test olarak İngilizce öğretmenlerine uygulanmıştır. Madde yazma eğitimi öncesi mezun olunan fakülte türüne göre öğretmenlerin ön test puanları açısından anlamlı bir fark olup olmadığını belirlemek amacıyla Mann Whitney U-testi, ön test puanlarının meslekte geçirdiği süreye göre de farklılaşp farklılaşmadığını belirlemek amacıyla Kruskal Wallis H-testi ve ön test ile son test puanları arasındaki farkın anlamlılığını test etmek amacıyla Wilcoxon işaretli sıralar testi kullanılmıştır. Öğretmenlerin madde yazma yeterliliklerinin mezun oldukları fakülte türüne ve meslekte geçirdiği süreye göre farklılaşmadığı bulunmuştur. Bunun yanı sıra, verilen eğitimin öğretmenlerin madde yazma yeterliliklerini geliştirmede önemli bir etkisinin olduğu sonucuna ulaşılmıştır.

Anahtar Kelimeler: ölçme ve değerlendirme, madde yazma yeterlilikleri, İngilizce öğretmeni

THE EFFECT OF ITEM CONSTRUCTION TRAINING ON COMPETENCIES REGARDING EVALUATING THE QUALITY OF ITEMS OF PRIMARY EDUCATION ENGLISH TEACHERS WORKING IN PRIVATE SCHOOL

Abstract:

In this study, which was conducted to determine the effect of qualified item writing training on the item writing proficiency of primary school English teachers working in private schools, one group pre-test post-test design, one of the experimental designs, was used. The study group of the research consists of 14 English teachers working in primary education in a private school selected by using convenience-sampling method. As a data collection tool in the research, a test consisting of 19 items was used to measure the criteria for writing qualified items, which was developed by utilizing from the literature. This test was administered to English teachers as a pre-test before in-service training on item writing and as a post-test after in-service training. Mann Whitney U-test to determine whether there is a significant difference in terms of pretest scores of teachers according to the type of faculty graduated before item writing training, Kruskal Wallis H-test to determine whether pretest scores differ according to the time spent in the profession, and Wilcoxon signed-rank test to determine whether there is a difference between pre- test and post test scores were used. It was found that teachers' item writing proficiency did not differ according to the type of faculty they graduated from and the time they spent in the profession. In addition, it was concluded that the in-service training provided had a significant effect on improving teachers' item writing competencies.

Keywords: measurement and evaluation, competency in item writing, English teacher

Giriş

Ölçme ve değerlendirme süreci, öğrencilerin gelişimleri hakkında dönüt sağlamak, öğrencilerin desteklenmesi gereken güçlü ve zayıf yönlerini, eksikliklerini belirlemek, öğretim sürecinin etkinliği hakkında fikir sahibi olmak ve süreç hakkında üst makamları da bilgilendirmek gibi pek çok süreci içerisinde barındırır. Hangi alanda olursa olsun, öğretimin hedefine ulaşp ulaşmadığına dair geçerli ve güvenilir saptamaların yapılabilmesi aynı zamanda geçerli ve güvenilir bir ölçme ve değerlendirme aracının

geliştirilmesine bağlıdır. Bu aşamada sınıf içi ölçme ve değerlendirmeleri gerçekleştiren öğretmenlere büyük rol düşmektedir. Test puanlarından elde edilen sonuçların geçerli ve güvenilir olabilmesi için bir testin sistematik ve iyi organize edilmiş bir yaklaşımla hazırlanmış olması gerekmektedir (Downing ve Haladyna, 2006). Kaliteli bir test geliştirmenin ilk şartı ise kaliteli test maddeleri yazabilmektir (Thorndike, 1967). Bir test maddesi, testteki temel gözlem birimidir ve katılımcının yanıtını ortaya çıkaran ifadeler içerir (Haladyna, 2004).

Her ne kadar günümüzde farklı ölçme ve değerlendirme yöntemleri kullanılmaya başlanmış olsa da öğretmenler, halen kendi hazırladıkları testleri diğer tüm test türlerinden çok daha sık kullanma eğilimindedir (Wenglinisky, 2000). Öğretmenlerin yüksek düzeyde ölçme ve değerlendirme okuryazarlığına sahip olması eğitim öğretim kalitesinin artırılması için vazgeçilmez bir unsurdur. Ölçme ve değerlendirme alanındaki yeterliklerin öğretmenlerin sahip olmaları gereken temel becerilerden biri olmasına rağmen pek çok araştırma, öğretmenlerin ölçme ve değerlendirme araç ve yaklaşımlarını yeterince iyi kullanamadıklarını ifade etmektedir (Calveric, 2010; Çakan, 2004; Erdal, 2007; Gelbal ve Kelecioğlu, 2007).

Özellikle dil öğretimi becerilerinin, dil yapısının ve kelime bilgisinin öğretimini kapsayan çok boyutlu bir yapı olduğu göz önünde bulundurulduğunda bu boyutların her birinin geçerli ve güvenilir bir şekilde ölçülmesi büyük bir önem arz etmektedir. Bu amacın gerçekleştirilebilmesi için farklı türlerde madde ve ölçme araçları dil öğretimi alanında sıklıkla kullanılmaktadır (Altmışdört, 2017). Daha da özele gidilecek olunursa İngilizce öğretimi bu alanlardan biridir. Öğretmenlerin gerek sürecin ölçülmesinde gerekse sonucun değerlendirilmesinde aktif test yazarları olduğu bu alanda, yapılan ölçme ve değerlendirmelerin geçerliliğinin ve güvenilirliğinin ve buna bağlı olarak kalitesinin artırılması İngilizce eğitim ve öğretiminde öğretmenlerin bu konuda yeterince bilgi ve beceri sahibi olmalarıyla yakından ilişkilidir (Jin, 2010). Bu durum ölçme ve değerlendirmede kullanılan ölçme araçlarının temel yapı taşı olan kaliteli test maddelerinin oluşturulmasına ilişkin eğitimin gerekliliğini doğurmaktadır. Alan yazında, öğretmenlerle yapılan araştırmalarda özellikle test maddesi yazma konusunda ihtiyacın olduğu, hizmet öncesi (Hatipoğlu, 2015; Lam, 2015) ve hizmet içi eğitimde (Firoozi, Razavipour ve Ahmadi, 2019; Zhang ve Yan, 2018) yetersizlikler olduğu gözlemlenmiştir. Ülkemizde İngilizce öğretmenlerinin ölçme ve değerlendirme okuryazarlıkları ile ilgili yapılan bir başka çalışma (Atay ve Mede, 2017) ise öğretmenler tarafından hazırlanan pek çok sınavın yeteri kadar kaliteli olmadığı ve İngilizce öğretmenlerinin dil alanında ölçme ve değerlendirme terimleri hakkında çok kısıtlı bilgiye sahip olduklarını ortaya koymaktadır. Bu araştırmanın bulgularına göre Türkiye'deki İngilizce öğretmenlerinin %22'si geleneksel test hazırlama konusunda hiç eğitim almadıklarını, %60'ı bu alanda çok az eğitim aldıklarını ifade etmişlerdir. Katılımcıların %18'i İngilizce ölçme ve değerlendirme alanında ileri derecede eğitim aldıklarını bildirmişlerdir. Bir başka çalışmada (Köksal, 2004) ise 2001-2002 ve 2002-

2003 akademik yıllarında ELT bölümü öğrencilerinin öğretmenlik uygulaması yaptığı devlet okullarından 56 farklı test örneği toplanmış, bu testlerde karşılaşılan hatalar değerlendirilmiş ve İngilizce öğretmenliği bölümündeki öğretmen adaylarının test oluşturma eğitimi almadıkları bu nedenle test hazırlayabilecek yeterlilikte olmadıkları ifade edilmiştir. Bunların yanı sıra, hazırlanan testlerle öğrencilere öğretilen içeriklerin örtüşmediği ve sınıfta öğretilen becerilerin sınavda sorulmadığı belirtilmiştir.

İngilizce öğretmenlerinin sınav hazırlama yeterlikleri üzerine uluslararası düzeyde yapılan araştırmalar, hazırlanan sınavların yeteri kadar kaliteli olmadığı (Alderson, 2005) ve İngilizce öğretmenlerinin dil alanında ölçme ve değerlendirme terimleri konusunda çok kısıtlı bilgiye sahip olduklarını göstermektedir (Gardner ve Rea-Dickins, 2001). İngilizce alanında sınıf içi değerlendirmeye ilgili yapılan bir başka çalışma da (Campbell, 2013), öğretmenlerin ölçme ve değerlendirme yeterlikleri üzerine verilen eğitimin yetersizliğine vurgu yapmıştır. Tüm bu çalışmalar göstermektedir ki, bu alanla ilgili eğitimin yetersizliği sadece ülkemizde değil, dünyanın başka ülkelerinde de karşılaşılan bir öğretmen eğitimi sorunudur. Tüm bu bilgiler ışığında bu çalışmanın amacı, test maddesi yazma eğitiminin özel okulda ilköğretim düzeyinde çalışan İngilizce öğretmenlerinin maddenin kalitesini değerlendirmeye ilişkin yeterliliklerine etkisini belirlemektir. Bu çalışma bağlamında cevap aranacak araştırma soruları şunlardır:

1) Etkili madde yazmayla ilgili hizmet içi eğitim verilmeden önce İngilizce öğretmenlerinin maddenin kalitesini değerlendirmeye ilişkin yeterlilikleri, mezun oldukları fakülte türüne (fen edebiyat ve eğitim fakültesi) göre anlamlı bir şekilde farklılaşmakta mıdır?

2) Etkili madde yazmayla ilgili hizmet içi eğitim verilmeden önce İngilizce öğretmenlerinin maddenin kalitesini değerlendirmeye ilişkin yeterlilikleri, meslekte geçirdikleri süreye (1-5, 6-10, 11-15, 16 ve üzeri yıl) göre anlamlı bir şekilde farklılaşmakta mıdır?

3) Etkili madde yazmayla ilgili hizmet içi eğitim verilmeden önce ve verildikten sonra İngilizce öğretmenlerinin maddenin kalitesini değerlendirmeye ilişkin yeterlilikleri anlamlı bir şekilde farklılaşmakta mıdır?

4) İngilizce öğretmenlerine etkili madde yazma eğitimi verilmeden önce ve sonra maddenin kalitesini değerlendirmeye ilişkin yeterlilikleri madde bazında nasıl değişmektedir?

Yöntem

Araştırmanın Modeli

Etkili madde yazma eğitiminin verilmesinin İngilizce öğretmenlerinin maddenin kalitesini değerlendirmeye ilişkin yeterliliklerine etkisinin incelendiği bu araştırmanın modeli, zayıf deneysel desenlerden tek grup ön test-son test desendir. Bu desende katı-

lımcıların bağımlı değişkene ilişkin ölçümleri uygulama öncesinde ön test, sonrasında son test olarak aynı denekler ve aynı ölçme araçları kullanılarak elde edilir (Creswell, 2014).

Çalışma Grubu

Araştırmanın çalışma grubuna dâhil edilecek öğretmenler 2020-2021 eğitim öğretim yılında Kayseri TED kolejinde çalışan öğretmenlerden seçilmiştir. Bu okulun seçiminde, bulunduğu ilde Türkiye’de uygulanan merkezi sınavlarda yüksek başarı elde etmesi, yapısı itibariyle öğretmen seçiminde belirli işlem adımlarını takip etmesi gibi faktörler rol oynamıştır. Araştırmada kolay ulaşılabılır örnekleme yöntemi kullanılarak 14 İngilizce öğretmeni seçilmiştir. Bu yöntemde araştırmacı, yakın olan ve erişilmesi kolay olan bir durumu seçmektedir (Yıldırım ve Şimşek, 2011). Araştırmacıardan birinin de bu kolejde çalışması zaman ve mekân açısından öğretmenlere ulaşmada avantaj sağlayacağından dolayı bu yöntem seçilmiştir. Çalışma grubuna dâhil edilen İngilizce öğretmenlerine ilişkin demografik bilgiler Tablo 1’de sunulmuştur:

Tablo 1. Çalışma Grubuna İlişkin Demografik Bilgiler

Değişken	Alt Boyut	f	
Mezun Olunan Fakülte Türü	1-5 yıl	1	
	6-10 Yıl	-	
	Eğitim Fakültesi	11-15 Yıl	1
	16 yıl ve üzeri	2	
	Toplam	4	
Mesleki Kıdem	1-5 yıl	2	
	Fen Edebiyat Fakültesi	6-10 Yıl	3
	11-15 Yıl	5	
	16 yıl ve üzeri	-	
	Toplam	10	

Tablo 1’de belirtildiği gibi öğretmenlerin çoğu fen-edebiyat fakültesinden mezun olmakla beraber mesleğinde 11 ile 15. yılları arasındadır.

Verilerin Toplanması ve Veri Toplama Aracı

Araştırmaya ilişkin veri toplanması amacıyla TED Üniversitesi İnsan Araştırmaları Etik Kurulu onayı (Sayı Numarası: E-27535802-100-4831) 10.06.2021 tarihinde alınmıştır. Araştırmada veri toplama aracı olarak 4-8. sınıf düzeylerinden farklı becerilere ilişkin kaliteli madde yazma ilkeleri açısından hatalı olan ve araştırmacı tarafından alan yazın taranarak hazırlanan 19 tane maddeden oluşan test kullanılmıştır. Hazırlanan testin ilk bölümünde katılımcılara gerekli bilgilendirmelerin yapıldığı ve katılım için

onaylarının alındığı bir bölüm yer almıştır. İkinci bölümünde ise İngilizce öğretmenlerine her bir maddenin hatasız veya hatalı olup olmadığını belirlemeleri ve hatalı ise sebebini açıklamaları istenmiştir. Hazırlanan testte buna ilişkin yer alan yönerge ve örnek bir madde Şekil 1’de verilmiştir:

Şekil 1. Madde Yazma İlkelerine İlişkin Hazırlanan Testteki Yönerge ve Örnek madde

Aşağıda verilen örnek soruları ve varsa cevap seçeneklerini, ölçme değerlendirme kriterleri açısından, belirtilen kazanım ve beceri türlerini dikkate alarak değerlendiriniz. Hatasız olduğunu düşündüğünüz sorularda yanıt için ayrılan alana "Soru Hatasızdır" yazınız. Hatalı bulduğunuz sorularda yanıt için ayrılan alana "Soru Hatalıdır" ifadesini yazarak sebebini açıklayınız. Çalışmaya katılımınızdan dolayı teşekkür ederiz.

Seviye	4. Sınıf
Kazanım:	Cümlelerin anlamına göre en doğru tamamlayan kelimeyi bulma.
Beceri:	Vocabulary

4. Fill in the blank with the correct word. (2 points.)

She wakes up late every morning, so she needs to buy an _____ clock.

A. time B. alarm C. watch D. bell

Yanıtınız

Şekil 1’de örnek olarak verilen madde, madde yazma ilkeleri açısından hatalıdır. Çünkü bu maddenin kökünde cevabı açık edecek şekilde dilbilgisi vb. ipucu ("an") verilmiştir. Oluşturulan taslak testin görünüş ve kapsam geçerliğine ilişkin kanıt elde etmek amacıyla bir İngilizce öğretmenine ve iki ölçme değerlendirme uzmanın görüşlerine başvurulmuş ve gelen dönütler doğrultusunda test uygulamaya hazır hale getirilmiştir.

Araştırmannın Uygulanma Süreci

Testin öğretmenlere uygulanmasından önce öğretmenlere katılımları için davet metni gönderilmiştir. Öğretmenlerin madde yazma yeterliliklerini belirlemek amacıyla geliştirilen test pandemi koşulları nedeniyle çevrimiçi platformda uygulanmıştır. Maddeler "Google forms" ortamında hazırlanarak araştırmacıyla ortak belirlenen bir saatte bütün katılımcılara uygulanmıştır. Böylelikle uygulama koşullarının bütün katılımcılar için eşit olması sağlanmıştır. Testin cevaplanması süresi olan bir saat içinde katılımcılardan gelebilecek soruları cevaplamak amacıyla çevrimiçi olarak gerçekleştirilen bu oturuma araştırmacılarından biri de katılmıştır. Bunun yanı sıra, aynı katılımcı-

nun ön ve son test puanlarını eşleştirebilmek amacıyla kendilerinden unutmayacakları bir kod yazmaları istenmiştir. Böylelikle katılımcı gizliliği de sağlanmıştır.

Ön-testin uygulanmasından sonraki hafta araştırmacı tarafından ilgili alan yazın taranarak ve Clay (2001) tarafından sunulan öneriler ışığında madde yazmayla ilgili hizmet içi eğitim hazırlanmış ve çevrimiçi platformda öğretmenlere verilmiştir. Bu eğitimde, (i) sınav ya da çalışma kâğıdı hazırlarken dikkat edilmesi gereken temel unsurlardan bahsedilmiş ve (ii) madde türlerine göre öğretmenlerin sınav ve çalışma kâğıdı hazırlarken sıklıkla düştükleri hatalar ve bunları önleme yolları üzerinde durulmuştur. Bu eğitimin teorik bilgiler sunmasının yanı sıra öğrenmelerin kalıcılığını artırmak amacıyla uygulamalı da olmasına dikkat edilmiştir. Öğretmenler gruplara ayrılarak hazırlanan çalışma kâğıdında yer alan maddelerdeki hataları tespit etme sürecini deneyimlemişlerdir. Bu hizmet içi eğitim yaklaşık iki saat sürmüştür. Ön test ve son test arasından geçen zamanın ön testte yer alan soruların hatırlanmasını engellemek amacıyla ne çok kısa ve arada başka öğrenmelerin veya faktörlerin karışmasının önüne geçmek amacıyla ne de çok uzun olmamasına dikkat edilmiştir. Bu amaçla, ön testin uygulanması, eğitimin verilmesi ve son testin uygulanması üç hafta içerisinde tamamlanmıştır.

Verilerin Analizi

Ön test ve son test sonuçları incelenerek frekans tablosu oluşturulmuştur, ön-test ve son-testte çeşitli hataların yapılma sıklığı belirlenmiştir. Katılımcıların sayısı 50'den az olduğu için Shapiro-Wilks testi yapılarak hazırlanan testten elde edilen puanların normallik varsayımını karşılayıp karşılamadığı test edilmiştir (Büyüköztürk, 2012). Hesaplanan p değerinin manidar çıkması nedeniyle normallik varsayımının karşılanmadığı tespit edilmiş ve parametrik olmayan yöntemler kullanılmıştır. Bu noktada madde yazma eğitimi öncesi mezun olunan fakülte türüne (eğitim fakültesi ve fen edebiyat fakültesi) göre ön test puanları açısından anlamlı bir fark olup olmadığını belirlemek amacıyla Mann Whitney U-testi kullanılmıştır. Bu test, iki ilişkisiz örneklemden/gruptan elde edilen puanların birbirinden anlamlı bir şekilde farklılaşp farklılaşmadığını gösterir (Büyüköztürk, 2012). Bunun yanı sıra, öğretmenlerin ön test puanlarının meslekte geçirdiği süreye (1-5, 6-10, 11-15, 16 ve üzeri yıl) göre de farklılaşp farklılaşmadığını belirlemek amacıyla Kruskal Wallis H-testi kullanılmıştır. Bu test, ilişkisiz en az iki grubun ortalamasının birbirinden anlamlı farklılık gösterip göstermediğini belirler (Büyüköztürk, 2012). Son olarak, öğretmenlerin ön test ve son test puanları arasındaki farkın anlamlılığını test etmek amacıyla Wilcoxon işaretli sıralar testi kullanılmıştır. Bu test ilişkili iki ölçüm setine ait fark puanlarını yönüne ve sıra miktarlarını dikkate almaktadır (Pallant, 2010). İstatistiksel olarak anlamlı farklılıklara ilişkin pratikteki anlamlılığı belirlemek amacıyla etki büyüklüğü değeri hesaplanmıştır. Bu değerin hesaplanmasında kullanılan formül şu şekildedir (Pallant, 2010):

$$Cohen\ d = \frac{z}{\sqrt{N}}$$

z, Wilcoxon işaretli sıralar testinden elde edilen z katsayısını, N ise gözlem sayısını ifade etmektedir.

Etki büyüklüğü değerinin düşük veya yüksek olarak yorumlanmasında Cohen (1988, s.22) tarafından belirlenen kriterler temel alınmıştır. Buna göre, "0.1" düşük etki büyüklüğünü, "0.3" orta düzeyde etki büyüklüğünü ve "0.5" yüksek düzeyde etki büyüklüğünü göstermektedir.

Bulgular

Öğretmenlere madde yazma eğitimi verilmeden önce mezun oldukları fakülte türüne (eğitim fakültesi ve fen edebiyat fakültesi) göre ön test puanları açısından anlamlı bir fark olup olmadığını belirlemek amacıyla yapılan Mann Whitney U-testi sonuçları Tablo 2’de verilmiştir:

Tablo 2. Mezun Olunan Fakülte Türüne göre Yapılan Mann Whitney U-testi Sonucu

Fakülte Türü	N	Sıra Ortalaması	Sıra Toplamı	U	p
Eğitim	4	7.00	28.00	18	.78
Fen-Edebiyat	10	7.70	77.00		

Tablo 2’de belirtildiği gibi öğretmenlerin madde yazma kriterleriyle ilgili hazırlanan maddelerden oluşan ön test puanlarının mezun oldukları fakülte türüne göre istatistiksel olarak farklılaşmadığı bulunmuştur, U=18, p=.78. Bu durum, eğitim fakültesi mezunlarının ön testten aldıkları puanlara ilişkin sıra ortalamalarının fen edebiyat fakültesinden mezun olan öğretmenlerin sıra ortalamasına yakın olmasıyla da desteklenmektedir.

İngilizce öğretmenlerinin ön test puanlarının meslekte geçirdiği süreye (1-5, 6-10, 11-15, 16 ve üzeri yıl) göre farklılaşp farklılaşmadığını belirlemek amacıyla yapılan Kruskal Wallis H-testi sonuçları Tablo 3’de verilmiştir:

Tablo 3. Meslekte Geçirilen Süreye göre Yapılan Kruskal Wallis H-testi Sonucu

Mesleki Yıl	N	Sıra Ortalaması	Serbestlik derecesi	χ^2	p
1-5	3	3.67	3	3.82	.28
6-10	3	9.17			
11-15	6	7.75			
16 yıl ve üzeri	2	10.00			

Tablo 3'te görüldüğü üzere her ne kadar meslekte 16 yıl üzeri geçiren İngilizce öğretmenlerinin ön testten aldıkları puanların ortalamasının diğer öğretmenlere göre yüksek olduğu bulunsu da Kruskal Wallis H-test sonucunda öğretmenlerin ön test puanlarının meslekte geçirdiği süreye göre istatistiksel olarak manidar farklılığın olmadığı bulunmuştur, χ^2 (sd=3, n=14)=3.82, p = .28. Diğer bir ifadeyle, meslekte geçirdiği süre bakımından farklı olan İngilizce öğretmenlerinin, eğitim verilmeden önce maddenin kalitesini değerlendirme açısından yeterliliklerinde farklılık olmadığı ifade edilebilir.

İngilizce öğretmenlerinin kaliteli madde yazma eğitimi öncesi ve sonrası maddenin kalitesini değerlendirme yeterliliklerinin anlamlı bir farklılık gösterip göstermediğine ilişkin yapılan Wilcoxon işaretli sıralar testi sonucu Tablo 4'te verilmiştir:

Tablo 4. Ön Test ve Son Test Puanlarına göre Yapılan Wilcoxon İşaretli Sıralar Testi Sonucu

Son test-Ön test	n	Sıra Ortalaması	Sıra Toplamı	z	p	Cohen d
Negatif Sıra	1	1.00	1.00	-3.24*	.001	.61
Pozitif Sıra	13	8.00	104.00			
Eşit	0	-	-			

*Negatif sıralar temeline dayalı

Tablo 4'e göre öğretmenlerin madde yazma eğitimi öncesi ve sonrasındaki test puanları arasında anlamlı fark olduğu bulunmuştur, z=-3.24, p<.001. Fark puanlarının sıra ortalaması ve toplamları dikkate alındığında, gözlenen bu farkın pozitif sıralar, yani son-test puanı lehinde olduğu görülmektedir. Bu sonuçlara göre, verilen bu eğitimin öğretmenlerin maddenin kalitesini değerlendirme açısından yeterliliklerini geliştirmede önemli bir etkisinin olduğu söylenebilir. Tablo 4'te istatistiksel olarak anlamlı çıkan bu farkın pratikteki anlamlılığını test etmek amacıyla hesaplanan etki büyüklüğü değerinin .61 olduğu görülmektedir. Bu değer .5'den büyük olması etki büyüklüğü değerinin yüksek olduğunu göstermektedir (Cohen,1998, s.22). Bir diğer ifade ile İngilizce öğretmenlerinin ön ve son testten aldıkları puanlar arasındaki fark yüksek düzeydedir.

Araştırma kapsamında katılımcı olarak dâhil edilen İngilizce öğretmenlerine kaliteli madde yazmayla ilgili hizmet içi eğitim verilmeden önce ve verildikten sonraki madde yazma ilkeleri açısından maddenin kalitesini değerlendirmeye ilgili durumlarını saptamak amacıyla verilen testte yer alan her bir maddeye ilişkin doğru ve yanlış yanıt frekansları belirlenmiştir. Bu frekans değerleri Tablo 5'te verilmiştir:

Tablo 5. Ön Test ve Son Testteki Maddelere Verilen Doğru ve Yanlış Yanıt Frekansları

Soru No	Ön Test		Son Test	
	Doğru Yanıt	Yanlış Yanıt	Doğru Yanıt	Yanlış Yanıt
1	5	9	8	6
2	8	6	9	5
3	10	4	14	0
4	4	10	13	1
5	8	6	13	1
6	8	6	11	3
7	5	9	11	3
8	12	2	13	1
9	4	10	11	3
10	5	9	10	4
11	11	3	10	4
12	6	8	9	5
13	6	8	10	4
14	4	10	10	4
15	9	5	13	1
16	5	9	13	1
17	4	10	11	3
18	9	5	13	1
19	12	2	14	0

Hizmet içi eğitim verilmeden önce öğretmenlerin bir kısmı, madde kalitesiz olmasına rağmen maddedeki hataları göz ardı ederek maddelerin hatasız olduğunu belirtmişlerdir. Özellikle çoğu öğretmenin (10 öğretmen) dikkat etmediği madde yazma kriterleri şunlardır: (i) cevabı açık edecek şekilde dilbilgisi vb. ipucu verilmiş olmaması, (ii) madde ile maddeyle ölçülmek istenen kazanımın eşleşmiş olması, (iii) doğru cevabın bariz bir şekilde belirgin olmaması ve (iv) maddenin kökünde yer alan olumsuz ifadesinin kalın, italik ya da büyük harflerle yazılmaması" olmasıdır. Göz ardı edilen bu kriterlerin yanı sıra dokuz öğretmen tarafından ihmal edilen kriterler ise şunlardır: (i)

Çeldiriciler maddenin köküyle ilgili ve anlamlı olmalı, (ii) doğru ve yanlış türü maddelerde birden fazla yargının yer almaması, (iii) eşleştirme türü maddelerde cevap seçeneklerinin homojen olması ve (iv) açık uçlu maddede cevaplayıcıdan istenen bilginin net bir biçimde belirtilmesidir. Öte yandan öğretmenlerin çoğunun ön testte doğru yanıtladığı maddenin 19. madde olduğu bulunmuştur. Bu madde, kapsam geçerliliğinin sağlanmasına yönelik belirte tablosu kullanılarak testin oluşturulmasıyla ilgilidir.

Tablo 5'te de belirtildiği gibi İngilizce öğretmenlerine kaliteli madde yazmayla ilgili hizmet içi eğitim verildikten sonra her bir maddeye ilişkin yanlış yanıt frekanslarında bir azalma olduğu bulunmuştur. Bir diğer ifade ile İngilizce öğretmenleri, her bir maddeyi madde yazma kriterleri açısından daha iyi analiz ederek maddelerde yer alan sorunları doğru bir biçimde belirleyebilmişlerdir. Örneğin, 3. madde ve 19. madde öğretmenlerin tamamı tarafından doğru bir biçimde kritik edilmiştir. Bunun yanı sıra hizmet içi eğitim verilmeden önce çoğu öğretmen tarafından doğru değerlendirilmeyen ve barındırdığı hataların doğru tespit edilmediği maddeleri (madde 1, madde 4, madde 7, madde 9, madde 14, madde 16 ve madde 17), eğitim aldıktan sonra öğretmenler doğru bir şekilde analiz edebilmiştir.

Tartışma, Sonuç ve Öneriler

Kaliteli madde yazma eğitiminin İngilizce öğretmenlerinin madde yazma/maddenin kalitesini değerlendirme yeterliliklerine etkisinin araştırıldığı bu çalışmada, tek grup ön test ve son test deneysel deseni kullanılmıştır. Bu amaçla öğretmenlerin kaliteli madde yazma kriterlerini ne düzeyde bildiklerini belirlemek amacıyla kaliteli madde olma açısından hatalar barındıran 19 maddelik geliştirilen bir test öğretmenlere uygulanmıştır. Daha sonrasında önemli olmasına rağmen alan yazında sıklıkla göz ardı edildiği belirtilen madde yazma kriterleri ve etkili bir test hazırlama ilkeleri hakkında İngilizce öğretmenlerine hizmet içi eğitim verilmiştir. Başlangıçta verilen testin eğitimden sonraki hafta son test olarak verilmesinden sonra her bir madde düzeyinde yapılan hatalara ilişkin frekans değerleri belirlenmiştir. Çeşitli açılardan öğretmenlerin madde yazma yeterliliklerinin farklılaşp farklılaşmadığını belirlemek amacıyla parametrik olmayan yöntemler kullanılmıştır.

Eğitim fakültesi veya fen edebiyat fakültesi mezunu olan İngilizce öğretmenlerinin bir test maddesinin kaliteli olma durumunu değerlendirme yeterlilikleri açısından farklılaşp farklılaşmadığını belirlemek amacıyla yapılan Mann Whitney-U testi sonucunda, öğretmenlerin başlangıç düzeylerinin birbirine yakın olduğu sonucuna ulaşılmıştır. Bu çalışmadan elde edilen bulguların aksine Aydın (2001) tarafından yürütülen çalışmada eğitim fakültesinden mezun olan öğretmenlerin ölçme ve değerlendirme yeterlilikleri fen edebiyat fakültesinden mezun olan öğretmenlerinkine kıyasla daha yüksek olduğu bulunmuştur. Araştırma sonuçlarındaki bu tutarsızlıkta, araştırmaya katılan İngilizce öğretmenlerinin mesleki kıdemleri ile mezun oldukları fakülte türü arasındaki etkileşim etkisi olabilir. Bir diğer anlatımla, mevcut araştırmaya katılan fen

edebiyat fakültesinden mezun olan İngilizce öğretmenlerinin çoğunun mesleki süresinin en az 6 yıl olması ve zamanla maddenin kalitesini değerlendirme veya kaliteli madde yazma konusuyla ilgili yeterince deneyim kazanmaları olabilir.

Hizmet içi eğitim verilmeden önce İngilizce öğretmenlerinin meslekteki hizmet yılına göre maddenin kalitesini değerlendirme açısından farklılaşıp farklılaşmadığı belirlemek amacıyla yapılan Kruskal Wallis H-testi sonucunda, mesleki süresi daha fazla olan öğretmenler ile mesleki süresi daha az olan öğretmenlerin maddenin kalitesini değerlendirmeyle ilgili yeterlilikleri açısından fark olmadığı ifade edilebilir. Bu bulgu, hizmet yılı değişkeni açısından öğretmenlerin ölçme ve değerlendirme yeterliliklerinin incelendiği araştırmanın (Aydın, 2001) bulgularıyla tutarlılık göstermektedir. Benzer şekilde, öğretmenlerin madde yazma, test geliştirme bilgilerinin mesleki tecrübeye göre farklılık gösterip göstermediğinin incelendiği bir başka araştırmanın (Inko-Tariah ve Okon, 2019) bulguları da bu sonucu destekler niteliktedir.

Kaliteli madde yazma veya maddenin kalitesini değerlendirmeyle ilgili eğitim verildikten sonra İngilizce öğretmenlerinin madde yazma yeterliliklerinin nasıl değiştiğini belirlemek amacıyla yapılan Wilcoxon İşaretli Sıralar testi sonucunda öğretmenlerin başlangıçtaki ve sonrasındaki yeterliliklerinde farklılık olduğu belirtilebilir. Bir diğer ifade ile öğretmenlerin bu konudaki yeterliliklerinde bir gelişme ve artış olduğu sonucuna ulaşılmıştır. Benzer şekilde bir başka çalışmada ders veren öğretim elemanlarına fakülte düzeyinde madde yazmayla ilgili eğitim verilmesinin öğretim elemanlarının yazdıkları maddelerin kalitesi üzerindeki etkisinin incelendiği çalışmada (Naeem ve diğ., 2012), eğitim sonrasında yazılan maddelerin kalitesinin arttığı belirtilmiştir.

Çoğu İngilizce öğretmenin çoktan seçmeli maddelerde kaliteli ve anlamlı olarak yazılmamış çeldiricileri, maddenin kökünde yer alan altı çizilmemiş olumsuz ifadeleri ve madde ile kazanımın eşleşmemiş olmasını fark etmede problem yaşadıkları sonucuna ulaşılmıştır. Benzer bulguya, Rush ve diğ.(2016)'nin üniversitede uygulanan testlerdeki maddelerin karmaşıklığının ve kalitesizliğinin madde ayırt ediciliği ve güçlüğüne olan etkisini belirlemek amacıyla yürütmüş oldukları çalışmada da rastlanılmıştır. Araştırmacılar, üniversite öğretim elemanlarının madde yazarken çoğunlukla çeldiricileri uygun bir şekilde yazmadığı, olumsuz madde kökü yazmada problem yaşadıklarını belirtmişlerdir. Bu durum, çoktan seçmeli maddelerin puanlamasının kolay olmasının verdiği avantajla çoğu uygulayıcılar tarafından sıklıkla kullanılması ve hatta sıklıkla yazılmasıyla açıklanabilir. Bu durumun sonucunda çok sayıda kalitesiz çoktan seçmeli maddede piyasada türemiştir. Oysaki her ne kadar çoktan seçmeli maddenin puanlanması kolay olsa da kaliteli çeldirici yazmak zor bir süreçtir (Shin ve diğ., 2019). Bu süreçte dikkat edilmeden yazılan çok sayıda madde ise öğretmenler ya da uygulayıcılar tarafından özüm senerek maddenin kalitesi göz ardı edilmeye başlanılmış olabilir.

Araştırma bulgularının analizi sonucunda madde yazma eğitimi öncesinde öğretmenlerin konuyla ilgili yetersizlikleri olduğu ve verilen eğitim neticesinde öğretmen-

lerin madde yazma yeterliliklerinde anlamlı bir artış olması sonucuna dayanılarak bu eğitimin diğer branşlarda da yaygınlaştırılması ve eğitimin veya çalıştayların belli aralıklarla tekrarlanarak sistematik hale getirilmesi önerilmektedir. Araştırmaya katılabilecek öğretmen sayısının az olmasından dolayı kontrol grubunun kullanılmaması çeşitli iç geçerlilik tehditlerine neden olmuş olabilir. Bu yüzden gelecekte yapılacak araştırmalar için daha büyük bir çalışma grubuyla kontrol grubu dâhil edilerek daha geniş çapta veri sağlayacak bir araştırma yürütülebilir.

Kaynakça

- ALDERSON, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. Continuum.
- ALTIŞDÖRT, G. (2017). The proper tests in language teaching: A study on instructors, students' and researchers' perspectives. *Science Journal of Turkish Military Academy*, 27(2), 167-186.
- AYDIN, A. (2001). *Eğitim fakültesi mezunu olan ve olmayan öğretmenlerin ölçme ve değerlendirme yeterliliklerinin karşılaştırılmasına yönelik bir çalışma* [Yayınlanmamış yüksek lisans tezi]. Ankara Üniversitesi.
- ATAY, D. and Mede, E., (2017). English teacher's assessment literacy: The Turkish context, *Dil Dergisi*, 1, 43-60.
- BÜYÜKÖZTÜRK, Ş.(2012). *Sosyal bilimler için veri analizi el kitabı*. Pegem Akademi.
- CALVERİC, S.B. (2010). *Elementary Teachers'assessment beliefs and practices* [Unpublished doctoral dissertation thesis]. Virginia Commonwealth University.
- CAMPBELL, C. (2013). Research on teacher competency in classroom assessment. In J. H. McMillan *SAGE handbook of research on classroom assessment* (pp. 71-84). SAGE Publications.
- CLAY, B. (2001, October). *Is this a trick question? A short guide to writing effective test questions*. Kansas Curriculum Center. Retrieved from <https://www.k-state.edu/ksde/alp/resources/Handout-Module6.pdf>
- COHEN, J.W.(1988). *Statistical power analysis of the behavioral sciences*. Lawrence Erlbaum Associates.
- CRESWELL, J. W. (2014). *Research design: Qualitative, quantitative and mixed Methods Approaches* (4th ed.). Sage.
- ÇAKAN, M. (2004). Öğretmenlerin ölçme-değerlendirme uygulamaları ve yeterlik düzeyleri: İlk ve ortaöğretim. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 37, 99-114.
- DOWNİNG, S.M. and Haladyna. T.M. (2006). *Handbook of Test Development*. Lawrence Erlbaum Associates Publishers.
- ERDAL, H. (2007). *2005 ilköğretim matematik programı ölçme değerlendirme kısmının incelenmesi (Afyonkarahisar ili örneği)* [Yayınlanmamış yüksek lisans Tezi]. Afyon Kocatepe Üniversitesi

- FİROOZİ, T., Razavipour, K. ve Ahmadi, A. (2019). The language assessment literacy needs of Iranian EFL teachers with a focus on reformed assessment policies. *Language Testing in Asia*, 9(2), 1-14. <https://doi.org/10.1186/s40468-019-0078-7>.
- GARDNER, S. and Rea-Dickins, P. (2001). Conglomeration or Chameleon? Teachers' representations of language in the assessment of learners with English as an additional language. *Language Awareness*, 10(2-3), 161-177.
- GELBAL, S. ve Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlikleri ve karşılaştıkları sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33, 135-145.
- HALADYNA, T.M. (2004). *Developing and validating multiple choice test items* (3rd Ed.). Lawrence Erlbaum Associates Publishers.
- 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.
- INKO-TARIAH, D. C. and Okon, E. J. (2019) Knowledge of test construction procedures among lecturers in Ignatius Ajuru University of Education, Port Harcourt, Nigeria. *Academic Research International*, 10(1), 130-138.
- JİN, Y. (2010). The place of language testing and assessment in the professional preparation of foreign language teachers in China. *Language Testing*, 27(4), 555-584.
- KÖKSAL, D.,(2004). Assessing teachers' testing skills in ELT and enhancing their professional development through distance learning on the net. *Turkish Online Journal of Distance Education*, 5(1), 1-11.
- LAM, R. (2015). Language assessment training in Hong Kong: Implications for language assessment literacy. *Language Testing*, 32(2), 169-197.
- NAEEM, N., Van der Vleuten C., and Alfaris, E. A. (2012). Faculty development on item writing substantially improves item quality. *Advances in Health Sciences Education*, 17, 369-376. <http://dx.doi.org/10.1007/s10459-011-9315-2>.
- RUSH, B. R., Rankin, D. C. and White, B. J. (2016). The impact of item-writing flaws and item complexity on examination item difficulty and discrimination value. *BMC Medical Education*, 16(250), 1-10. <https://doi.org/10.1186/s12909-016-0773-3>
- SHİN J, Guo Q and Gierl MJ. (2019) Multiple-choice item distractor development using topic modeling approaches. *Frontiers in Psychology*. 10(825), 1-14. <http://dx.doi.org/doi:10.3389/fpsyg.2019.00825>
- THORNDİKE, R. L. (1967). *The analysis and selection of test items*. In S. Messick ve D.Jackson (Eds.), *Problems in Human Assessment*. McGraw-Hill.
- WENGLİNSKY, H., (2000) *How teaching matters: Bringing the classroom back into discussions of teacher quality*, Educational Testing Service.
- ZHANG, C and Yan, X.(2018) Assessment literacy of secondary EFL teachers: Evidence from a regional EFL test. *Chinese Journal of Applied Linguistics*, 41(1), 25-46. <https://doi.org/10.1515/cjal-2018-0002>