

Yabancı/İkinci Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme-Değerlendirme Okuryazarlığı Öz Yeterlik Ölçeğinin Geliştirilmesi

Gizem Yıldırım*

Neslihan Karakuş**

MAKALE BİLGİSİ

Geliş: 20.02.2025
Düzeltilme: 18.08.2025
Kabul: 15.09.2025
Doi: 10.31464/jlere.1643713

Anahtar Sözcükler:

*yabancı dil olarak Türkçe
ölçek geliştirme
ölçme ve değerlendirme*

ÖZET

Bu çalışmanın amacı, yabancı/ikinci dil olarak Türkçe öğreten öğretmenlerin ölçme ve değerlendirme okuryazarlık öz yeterliklerinin belirlenmesinde kullanılabilecek geçerli ve güvenilir bir ölçme aracı geliştirmektir. Bu amaç doğrultusunda araştırmada, nicel araştırma yöntemlerinden tarama deseni kullanılmıştır. Alan yazını incelemesi neticesinde madde havuzu oluşturulmuş ve taslak maddeler, uzman görüşüne sunulmuştur. Kapsam geçerliğine kanıt olarak uzman görüşleri Lawshe tekniği ile analiz edilmiştir. Yapı geçerliği kapsamında açıklayıcı ve doğrulayıcı faktör analizi verileri, farklı zamanlarda 419 öğreticiden toplanmıştır. Analizler sonucunda 36 maddeden oluşan dört faktörlü bir yapı elde edilmiştir. Bu faktör yapıları; “Dil becerileri”, “Temel kavramlar”, “Teknoloji tabanlı ölçme ve değerlendirme” ve “Etik” olarak adlandırılmıştır. Bu faktörlerin Cronbach alfa katsayıları sırasıyla .903, .945, .930 ve .884 olarak hesaplanmıştır. Ayrıca ölçeğin tamamına ait Cronbach alfa katsayısı .912 ve toplam açıklanan varyans yüzdesi %61,32 olarak bulunmuştur. Bu değerler, araştırma kapsamında geliştirilen ölçeğin öğretmenlerin ölçme ve değerlendirme okuryazarlık öz yeterliklerini belirleyebilecek nitelikte geçerli ve güvenilir bir araç olduğunu göstermektedir.

Bilgilendirme

Bu araştırma, birinci yazarın ikinci yazar danışmanlığında hazırladığı yüksek lisans tezinden üretilmiştir.

Yayın Etiği Bilgilendirme

Bu çalışma için Etik Kurul onayı alınmıştır: YTÜ Sosyal ve Beşeri Bilimler Araştırmaları Etik Kurulu, tarih:31.10.2024 ve Toplantı no:2024.10.

Yazarların Katkı Oranı

Yazarların katkı oranı eşittir.

Çıkar Çatışması

Bu çalışmada bir çıkar çatışması bulunmamaktadır.

Gönderim

Yıldırım, G. ve Karakuş, N. (2025). Yabancı/ikinci dil olarak Türkçe öğretmenlerine yönelik ölçme-değerlendirme okuryazarlığı öz yeterlik ölçeğinin geliştirilmesi. *Dil Eğitimi ve Araştırmaları Dergisi, 11 (2)*, 511-540.

* Yüksek Lisans, ORCID ID: <https://orcid.org/0000-0002-9621-7612>, Yıldız Teknik Üniversitesi, Yabancılar Türkçe Öğretimi, yildirangizem15@gmail.com

** Prof, ORCID ID: <https://orcid.org/0000-0001-7808-1099>, Yıldız Teknik Üniversitesi, Türkçe ve Sosyal Bilimler Eğitimi, neslihanoksuz61@gmail.com

Giriş

Modern eğitim yaklaşımlarının benimsenmesiyle öğrenme sürecinin sistematik olarak izlenmesi, yapıcı geri bildirim sağlanması, dijital teknolojilerin değerlendirme sürecine dâhil edilmesi, ihtiyaç ve özel gereksinimlere göre çeşitlendirilmesi, uygun öğretim yöntemlerinin belirlenmesi için öğretmenlerin ölçme ve değerlendirme yetkinliğine sahip olması beklenmektedir. Öğretmenlik Mesleği Genel Yeterlilikleri (MEB, 2017), öğretmenlere yönelik beklentileri göstermesi açısından bir kaynak niteliği taşımaktadır. Bu yeterlilikler; mesleki bilgi, mesleki beceri, tutum ve değerler olmak üzere üç alan, on bir alt alan ve altmış beş göstergeden oluşmaktadır. Öğretmenlerden beklenen yeterlikler kapsamında mesleki beceri alanının alt boyutlarından biri olarak ölçme ve değerlendirme öne çıkmaktadır. Ölçme ve değerlendirme etkinlikleri, öğretmenlerin öğretim sürecinin etkinliğini anlaması, buna göre öğretim yöntem ve tekniklerini güncelleştirmesi, çeşitlendirmesi ve süreci yeniden yapılandırabilmesi için gerekli uygulamalardır. Bu gereklilik, ölçme ve değerlendirme okuryazarlığını ön plana çıkarmaktadır.

Ölçme ve değerlendirme okuryazarlığı kavramı Gürsoy'a (2015, s.9) göre "öğretmenin öğrenci başarısını, uyguladığı yöntemlerin geçerliğini ve öğretim programının uygunluğunu doğru bir şekilde değerlendirebilmesi için gerekli olan ölçme aracını etik kurallar ve ölçme prensipleri içinde seçme/geliştirme, uygulama, puanlama, yönetme, bilgilendirme, sonuçları yayımlama süreçlerini doğru yönetme bilgi ve becerisi"dir. Popham (2006) değerlendirmeye ilişkin temel ilke ve kavramlar hakkında yeterli bilgiye sahip olmak olarak tanımlarken McMillan (2000) bunun yeterli olmadığını, öğrenmeyi teşvik etmek için uygun olanı seçebilmeleri olarak tanımlanması gerektiğini savunmuştur. Bu yargıyı destekler nitelikte Rogler (2014) öğretmenlerin değerlendirmenin temel kavramları hakkında yeterli bilgiye sahip olmasının yanı sıra değerlendirme uygulamalarının planlanması ve değerlendirmeye bağlı olarak geri bildirim, not verme, sonuçları paylaşma gibi süreçlere etkilerinin farkında olmaları gerektiğini dile getirmiştir.

Ölçme ve değerlendirme okuryazarlığını, Stiggins (1995) öğretmenlerin nesnel değerlendirme yapabilme, McMillan (2001) öğretim süreci ve değerlendirme etkinliklerini bir bütün halinde planlayarak uygun yöntem teknikleri kullanabilme, Webb (2002) ölçme ve değerlendirme sonuçlarına göre süreci yeniden düzenleyebilme, Mertler(2004) ölçme ve değerlendirme etkinlikleriyle öğrencini başarısını belirlemenin yanı sıra öğrenmeyi de sağlayabilme, Gottheiner ve Siegel (2012) özel öğretim amacına uygun ölçme ve değerlendirme aracını seçebilme gibi yeterliliklere sahip olması bakımından eğitim süreci için vazgeçilmez bulmaktadır.

Tüm eğitim alanlarında olduğu gibi Türkçenin yabancı dil olarak öğretimi sürecinde de ölçme ve değerlendirmenin önemi yadsınamaz. Yabancı/ikinci dil olarak Türkçe öğreten öğrencilerin de ölçme ve değerlendirme okuryazarlık becerisi farkındalığında olmasının eğitim sürecine olumlu yansımaları olacaktır. Fakat Kolcuoğlu'nun (2022) çalışmasında yabancı dil olarak Türkçe öğretiminde ölçme ve değerlendirme alanında öğrencilerin okuryazarlık düzeylerinin istenen ve olması gereken düzeyin altında olduğu tespit edilmiştir. Boylu (2019) Türkçenin yabancı dil öğretim sürecinde öğrencilerin ölçme ve değerlendirme okuryazarlık yetkinliğinde olmamasından

kaynaklı olarak, geçerliliği ve güvenilirliği sağlanmamış ölçme araçları kullandıklarına, süreç yerine sonuç değerlendirmeye öncelik verdiklerine, anlatma becerilerine yönelik dereceli puanlama anahtarı kullanmadıklarına değinmiştir. Çakan'ın (2004) çalışmasında öğretmenler de genel olarak ölçme ve değerlendirme becerilerinin yetkinlik düzeyinde olmadığını dile getirmişlerdir. Işıkoğlu (2015) çalışmasında yabancı dil olarak Türkçe öğreten öğretmenlerin sınav hazırlama konusundaki yetersizlikleri sonucu, güvenilirliği ve geçerliği düşük sınavlar hazırlandığını ifade etmiştir. Bunun sonucunda uluslararası öğrencilerin dil seviyelerinin doğru tespit edilememesi, öğrencinin altında olduğu bir üst kura geçmesi ya da kuru tekrar etmesi, hatalı sertifikalandırma gibi olumsuz sonuçlara neden olmaktadır. Bu nedenle Türkçenin yabancı/ikinci dil olarak öğretiminde sorumluluk alan öğretmenlerin ölçme ve değerlendirme okuryazarı olmaları, eğitim sürecini doğru değerlendirebilmeleri ve yorum yapabilmeleri, geçerli ve güvenilir bir ölçme ve değerlendirme sürecinin gerçekleşebilmesi için ilk koşuldur.

Amerikan Öğretmen Federasyonu (AFT), Eğitimde Ölçme Ulusal Konseyi (NCME) ve Ulusal Eğitim Derneği (NEA) (1990) tarafından öğrencilerin eğitsel değerlendirilmesinde öğretmen yeterlik standartları belirlenmiştir. Bu standartlara göre ölçme ve değerlendirme okuryazarı öğretmenler; ölçme ve değerlendirme yöntemlerini seçme, geliştirme, uygulama, puanlama, sonuçları yorumlama, kullanma ve paylaşma, güvenilir ve geçerli notlandırma konusunda yeterli olmalı ve etik ilkelere özen göstermelidir.

Yabancı/ikinci dil olarak Türkçe öğretimi alanında görevli öğretmenlerin ölçme ve değerlendirme yeterlikleri ise şöyle sıralanmıştır (Barın ve diğerleri, 2017):

- Yabancı/ikinci dil olarak Türkçe öğretimiyle ilgili ölçme değerlendirme uygulamalarının amaçlarını belirleyebilme
- Yabancı/ikinci dil olarak Türkçe öğretiminde ölçme-değerlendirme araçlarını ve yöntemlerini kullanabilme
- Hedef kitlenin dil gelişimini belirlemek amacıyla ölçme sonuçlarını yorumlama ve geri bildirim verebilme
- Hedef kitlenin dil gelişimini belirlemek amacıyla ölçme değerlendirme sonuçlarını çalışmalarına yansıtabilme

Bu yeterlikler, öğretmenlere yabancı/ikinci dil olarak Türkçe öğretiminde ölçme ve değerlendirme hakkında yol göstericidir. Öğreticilerin, mesleki yeterliliklerini sorgulamaları, mevcut durumlarını belirlemeleri gelişim hedeflerini oluşturmalarını ve güncellemelerini sağlamak için öz değerlendirme bilincine ve yetkinliğine sahip olması gerekmektedir. Özdemir'e (2008) göre öğretmenlerin kendi yeteneklerine ve becerilerine ilişkin kişisel yargıları nitelikli öğretim yapmalarına ve öğretim sürecinde karşılaştıkları sorunları çözmelerine katkı sağlamaktadır.

Alan yazını incelendiğinde ölçme ve değerlendirme okuryazarlıkları üzerine çalışmalarda kullanılan ölçekler (Plake & Impara 1993; McMillan 2000; Mertler & Campell 2005; Nartgün, 2008; Kesici, 2022; Kırık, 2008; Bütün vd., 2010; Maden & Durukan, 2011; Mete, 2012; Ogan-Bekiroğlu & Suzuk, 2014; Gücüyeter & Karadoğan, 2019; Kılınç, 2011; Gülbahar & Büyüköztürk, 2008; Güneş, 2007; Gül, 2011; Davies, 2008) göz önünde bulundurulduğunda hedef kitleye yönelik yabancı/ikinci dil olarak

Türkçe öğretimi sürecinde dil becerilerini değerlendirilmesi, teknoloji tabanlı ölçme ve değerlendirme yöntemleri ve etik ilkeleri kapsayan alana özgü Türkçe ölçek geliştirme çalışmasının bulunmadığı tespit edilmiştir. Bu durum, alandaki mevcut eksikliği ortaya koymaktadır.

Bu bağlamda yabancı/ikinci dil olarak Türkçe öğreten öğretmenlere yönelik ölçme ve değerlendirme okuryazarlıklarının öz değerlendirilmesine olanak sağlayacak geçerli ve güvenilir bir ölçeğin tespit edilememesi, çalışmanın yapılma gerekliliğini ortaya koymaktadır. Geliştirilmesi amaçlanan ölçekle öğretmenlerin ölçme ve değerlendirme okuryazarlığı çerçevesinde geliştirilmesi gereken yönleri ile güçlü yönlerini değerlendirerek bu konudaki durumlarına farkındalık sağlaması hedeflenmektedir.

Araştırmanın Amaçları ve Araştırma Soruları

Bu araştırmanın amacı, yabancı/ikinci dil olarak Türkçe öğreten öğretmenlerin ölçme ve değerlendirme okuryazarlık öz yeterliklerinin belirlenmesinde kullanılabilecek geçerli ve güvenilir bir ölçme aracı geliştirmektir. Bu amaç doğrultusunda çalışmada yanıtlanması hedeflenen sorular sıralanmıştır:

1. Yabancı/İkinci Dil Olarak Türkçe Öğreten Öğreticilere Yönelik Ölçme ve Değerlendirme Okuryazarlığı Öz Yeterlik Ölçeği geçerli bir ölçme aracı mıdır?
2. Yabancı/İkinci Dil Olarak Türkçe Öğreten Öğreticilere Yönelik Ölçme ve Değerlendirme Okuryazarlığı Öz Yeterlik Ölçeği güvenilir bir ölçme aracı mıdır?

Yöntem

Araştırma Tasarımı/Modeli

Bu çalışmada nicel araştırma yöntemlerinden tarama (survey) deseni, araştırmanın modeli olarak belirlenmiştir. Tarama deseni araştırmalar, “geçmişte ya da hâlen var olan bir durumu var olduğu şekliyle tespit etmeyi amaçlayan” (Karasar, 2024 s.109), “bir konuya ya da olaya ilişkin katılımcıların görüşlerinin ya da ilgi, beceri, yetenek, tutum vb. özelliklerinin belirlendiği genellikle diğer araştırmalara görece daha büyük örneklemeler üzerinden yapılan araştırmalar” (Büyüköztürk ve diğerleri, 2023 s.231) olarak tanımlanmıştır. Tarama deseni, bir örneklem üzerinden evrene dair beceri, tutum veya görüşlerin sayısal verilere dayanarak nesnel bir betimleme yapmaya olanak sağlamaktadır (Bal, 2023). Ayrıca tarama deseninin; geniş bir evrene ait birtakım özellikleri tanımlama, evrenin genel özelliklerini ortaya koyabilen güvenilir sonuçlar elde etme, birden fazla değişken arasındaki ilişkiyi aynı anda tespit edebilme, araştırmacıya anket maddeleri ve analiz teknikleri bakımından esneklik sağlama, katılımcıların görüş, tutum ve özelliklerini standart bir ölçekle tespit ederek daha objektif sonuçlara ulaşma gibi diğer yöntemlere göre üstünlük sağlayan özellikleri bulunmaktadır (Gülbüz & Şahin, 2018).

Bu çalışmada, büyük gruplarla çalışılması, nesnel verilere dayalı güvenilir sonuçlar elde edilmek istenmesi ve yabancı/ikinci dil olarak Türkçe öğretmenlerin ölçme ve

değerlendirme okuryazarlığı öz yeterlik becerilerinin belirlenmesine yönelik bir ölçme aracı geliştirilmesi amacına uygun olarak tarama deseni tercih edilmiştir.

Yayın Etiği

Bu araştırma, bilimsel yayın etiğine uygun olarak hazırlanmıştır. Yıldız Teknik Üniversitesi Sosyal ve Beşeri Bilimler Araştırmaları Etik Kurulu tarafından 31.10.2024 tarihinde 2024.10 toplantı nolu etik kurul onayı verilmiştir. Araştırmaya katılım gönüllülük esasına dayalı olarak yürütülmüş ve katılımcılardan bilgilendirilmiş onam formu alınmıştır.

Katılımcılar

Yabancı/ikinci dil olarak Türkçe öğretiminde görev alan 419 öğreticiyle araştırma yürütülmüştür. Açıklayıcı faktör analizi (AFA) için ilk aşamada amaçlı örnekleme yoluyla seçilen 217 öğreticiden veri toplanmıştır. AFA sonucunda elde edilen faktör yapısının doğruluğunu test etmek amacıyla ikinci aşamada 202 öğreticiden veri toplanmıştır. Her iki veri grubu da farklı zamanlarda ve birbirinden bağımsız katılımcılardan elde edilmiştir. Ayrıca geliştirilmek istenen ölçeğin kapsam geçerliğini belirlemek amacıyla görüşlerine başvurulmuş 3 ölçme ve değerlendirme ve 20 alan uzmanı araştırmaya dahil edilmiştir.

Çizelge 1. Faktör Analizi Katılımcılarının Demografik Bilgileri

1. Çalışma Grubu	Değişken	Alt Değişkenler	f	%	2. Çalışma Grubu	Değişken	Alt Değişkenler	f	%				
	Cinsiyet	Kadın	140	64,5		Cinsiyet	Kadın	127	62,9				
		Erkek	77	35,5			Cinsiyet	Erkek	75	37,1			
	Eğitim Durumu	Lisans	23	10,6		Eğitim Durumu		Lisans	153	75,7			
		Yüksek Lisans	132	60,8			Eğitim Durumu	Yüksek Lisans	45	22,3			
		Doktora	62	28,6				Eğitim Durumu	Doktora	4	2		
	Deneyim Süresi	0-5 yıl	87	40,1		Deneyim Süresi			0-5 yıl	5	2,5		
		6-10 yıl	60	27,6			Deneyim Süresi		6-10 yıl	173	85,6		
		11-15 yıl	50	23,0				Deneyim Süresi	11-15 yıl	13	6,4		
		16-20 yıl	12	5,5					Deneyim Süresi	16-20 yıl	4	2	
		21-25 yıl	5	2,3						Deneyim Süresi	21-25 yıl	5	2,5
		26 yıl ve üzeri	3	1,4							Deneyim Süresi	26 yıl ve üzeri	2
	Sertifika	Evet	193	88,9		Sertifika						Evet	189
		Hayır	24	11,1			Sertifika					Hayır	13
Toplam		217	100	Toplam		202		100					

Çizelge 1. incelendiğinde araştırmanın AFA ve DFA grubu katılımcılarının cinsiyet, eğitim durumu, deneyim süresi, “Yabancılar Türkçe Öğretimi” Sertifikasına sahip olma durumu değişkenlerine yönelik frekans analizi sonuçları görülmektedir. Bu çizelgeye göre 1. Çalışma Grubunda yer alan 217 katılımcının %64,5’i kadın ve %35,5’i erkektir. Katılımcıların 23’ü (%10,6) lisans, 132’si (%60,8) yüksek lisans, 62’si (28,6) ise doktora mezunudur. Buna göre katılımcıların çoğunluğunun (%89,4) lisans üstü eğitime sahip olduğu söylenebilir. Bu alandaki deneyim süreleri değişkenine bakıldığında, 87 (%40,1) katılımcı 0-5 yıl, 60 (%27,6) katılımcı 6-10 yıl, 50 (%23) katılımcı 11-15 yıl, 12 (%5,5) katılımcı 16-20 yıl, 5 (%2,3) katılımcı 21-25 yıl, 3 (%1,4) katılımcı 26 yıl ve üzeri

deneyime sahip olduğunu belirtmiştir. Katılımcıların sadece 24'ü (%11,1)'ü “Yabancılar Türkçe Öğretimi” sertifikasına sahip değilken geri kalanı (%88,9) söz konusu sertifikaya sahiptir. 2. Çalışma Grubu bilgileri incelendiğinde %62,9 oranında kadın, %37,1 oranında erkek katılımcı bulunduğu gözlenmektedir. Katılımcıların 153 (75,7)'ü lisans, 45 (%22,3)'i yüksek lisans ve 4 (%2)'ü doktora mezunu olduklarını beyan etmişlerdir. Katılımcıların çoğunluğunun (%85,6) 6-10 yıl deneyim süresi bulunurken, bu süreyi 11-15 yıl (%6,4), 21-25 yıl (%2,5), 0-5 yıl (%2,5), 16-20 yıl (%2) ve son olarak 26 yıl ve üzeri (%1) deneyim süresi bulunan katılımcılar takip etmektedir. Katılımcıların çoğunluğunun (%93,6) “Yabancı Dil Olarak Türkçe Öğretimi” sertifikası vardır. Sadece 13 (%6,4) katılımcı bu sertifikaya sahip olmadığını belirtmiştir.

Veri Toplama ve Verilerin Çözümlemesi

Araştırma kapsamında Türkçenin yabancı/ikinci dil olarak öğretiminde görev alan öğretmenlerin çevrim içi veya yüz yüze geri dönüşleri ile veriler elde edilmiştir. Açıklayıcı faktör analizi verileri, 01.11.2024-29.12.2024 tarihleri arasında üniversitelerin Türkçe ve Yabancı Dil Uygulama ve Araştırma Merkezleri (TÖMER), Dil Öğretim Uygulama ve Araştırma Merkezi (DİLMER) veya Yabancı Diller Yüksek Okulunda Türkçenin yabancı dil olarak öğretiminde yer alan öğretim görevlileri ile yürütülmüştür. Doğrulamalı faktör analizi verileri ise 12.01.2024-01.02.2024 tarihleri arasında toplanmıştır.

Elde edilen veriler excel dosyasında birleştirilerek bütünlük sağlanmıştır. Ardından araştırmanın istatistiksel analizleri için IBM SPSS Statistics 30.0 ve IBM SPSS AMOS 24.0 programları kullanılmıştır. Keşfedici ve tanımlayıcı istatistikler SPSS kullanılarak hesaplanırken, yapısal eşitlik modellemesi (YEM) ve doğrulamalı faktör analizleri (DFA) için AMOS'tan yararlanılmıştır.

Bu çalışmada geliştirilen Yabancı/İkinci Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme ve Değerlendirme Okuryazarlık Öz Yeterliği Ölçeği'nin geçerliliği ve güvenilirliği, iki aşamalı bir yaklaşımla test edilmiştir. İlk aşamada, 217 katılımcıdan oluşan bir pilot çalışma veri seti üzerinde açıklayıcı faktör analizi (AFA) uygulanarak ölçeğin boyut yapısı belirlenmiştir. İkinci aşamada ise, 202 katılımcıdan oluşan farklı bir ana çalışma veri seti kullanılarak doğrulamalı faktör analizi (DFA) gerçekleştirilmiş ve AFA ile keşfedilen faktör yapısının mevcut veri ile uyumu sınanmıştır.

Ölçeğin güvenilirliğini değerlendirmek amacıyla Cronbach's Alpha (α) ve McDonald's Omega (ω) katsayıları hesaplanmıştır. Yapısal geçerliliğin bir parçası olan yakınsama geçerliliği (convergent validity) için birleşik güvenilirlik (CR) ve açıklanan ortalama varyans (AVE) değerleri kullanılmıştır. Ayrışan geçerliliğin (discriminant validity) yeterli düzeyde olduğunu doğrulamak amacıyla ise, yapısal eşitlik modellemesi öncesinde detaylı analizler yapılmıştır. Elde edilen bulgular, ölçeğin yapısal modelleme için gerekli olan istatistiksel koşulları karşıladığını ve geçerli bir ölçüm aracı olduğunu kanıtlamaktadır.

İşlem Süreci

Alan yazınındaki ölçek geliştirme aşamaları (Tezbaşaran 2008; Yurdabakan & Çüm 2017; Çarkçı, 2020; Erkuş, 2022; DeVellis, 2022) doğrultusunda izlenen adımlar sıralanmıştır:

Ölçülmek istenen özelliğin tanımlanması ve amacının belirlenmesi: Ölçek geliştirilmenin ilk aşaması ölçülmesi hedeflenen özellik, kavram veya davranışın açık, anlaşılır, ölçülebilir ve gözlemlenebilir bir şekilde çerçevesinin belirlenmesi, tanımlanması ve sınırlandırılarak araştırmanın amacını ortaya koyma aşamasıdır. Geliştirilmek istenen ölçeğe konu olan ölçme ve değerlendirme okuryazarlığı kuramsal olarak incelenmiş ve ölçeğin amacı belirlenmiştir.

Literatüre dayalı ölçek maddelerinin üretilmesi: Alan yazınındaki Türkçenin yabancı/ikinci dil olarak öğretiminde araştırma konusuna yönelik ulusal ve uluslararası araştırmalar incelenmiştir (Newfields 2006; Köse, 2008; Mete, 2012; Yıldız & Tepeli, 2014; Beyreli, 2017; Barın ve diğerleri, 2017; Boylu, 2019; Karadoğan & Uyumaz 2022; Bozkurt 2022; Dincel, 2021; Çiçek ve diğerleri, 2023; Bayrakdar, 2024). Bunun yanı sıra MEB tarafından belirlenen Öğretmenlik Mesleği Genel Yeterlikleri ve Diller İçin Avrupa Ortak Başvuru Metni'nden (D-AOBM) madde havuzu oluşturulma sürecinde yararlanılmıştır. Bu bağlamda konu ile ilgili 86 madde üretilmiştir.

Madde havuzunun uzman görüşüne sunulması: Madde havuzu kapsamında oluşturulan 86 madde içeren uzman görüş formu, kapsam geçerliğini belirleyebilmek, hedef kitleye uygunluğu ve maddelerin anlaşılabilirliğini tespit edebilmek amacıyla alan uzmanı ve ölçme ve değerlendirme uzmanlarına e-posta aracılığıyla görüşleri alınmak üzere ulaştırılmıştır. On uzmanın uzman görüş formuna verdikleri yanıtlar analiz edilerek madde havuzu 49 maddeye indirgenmiştir. Uzman görüşleri neticesinde madde havuzu değerlendirilmek üzere ikinci bir uzman görüşü için form yeniden düzenlenerek e-posta aracılığıyla farklı 13 uzmana ulaştırılmıştır. Bu bağlamda uzmanlardan ölçek kapsamında kalmasını istedikleri madde için “yeterli”, belli değişiklikler öngörerek kalmasını düşündükleri maddeler için “geliştirilmeli”, ölçekte bulunmasının uygun olmayacağını düşündükleri madde içinse “yetersiz” olarak görüş bildirmeleri istenmiştir. Ayrıca “geliştirilmeli” veya “yetersiz” gördükleri maddeleri neden böyle değerlendirdiklerine ilişkin nedenleri ve varsa önerilerini yazılı olarak açıklamaları beklenilmiştir.

Uzman görüşlerinin veri analizi: İkinci aşama uzman görüşlerinin nesnel olarak değerlendirilebilmesi amacıyla Lawshe (1975) tarafından geliştirilen maddelerin kapsam geçerlik oranı (KGO) ve kapsam geçerlik indeksi (KGİ) hesaplanmıştır. Bu hesaplama iki ölçme ve değerlendirme uzmanı ve 11 alan uzmanı olmak üzere 13 uzman dâhil edilmiştir. Ayre ve Scally (2014, s.85) tarafından kapsam geçerlik oranına dair kritik değer, 13 uzman için 0,538 olarak kabul edilmiştir.

Ön deneme uygulamasının gerçekleştirilmesi: Ölçek maddelerinin okunabilirliğinin ve anlaşılabilirliğinin kontrol edilebilmesinin yanı sıra ölçek uygulama süresinin ortalama olarak belirlenebilmesine yönelik 10 öğreticiyle görüşme gerçekleştirilmiştir. Öğreticilerin geri bildirimleri değerlendirilerek düzenlemeye gidilmiştir. Uygulama süresi ortalama 4 dakika 53 saniye olarak hesaplanmıştır. Ayrıca pilot uygulamaya katılan öğretmenler, deneme uygulaması örneklemini dışında bırakılmıştır.

Deneme maddelerinin belirlenmesi: Pilot uygulama sonrası yapılan düzenlemelerin ardından nihai ölçek maddelerine karar verilerek maddeler Google forma aktarılmıştır.

Deneme uygulamasının gerçekleştirilmesi: Yıldız Teknik Üniversitesi E-44513635-399-12732 sayılı yazısı kurum izinleri alınmak üzere Türkiye genelinde bulunan üniversitelerin ilgili bölümlerine ulaştırılmıştır. Alınan kurum izinleri sonrası örneklem büyüklüğü belirlenen hedef kitleye geliştirilmekte olan ölçek uygulanmıştır. Araştırmaya katılmayı kabul eden gönüllü katılımcılardan elektronik veya basılı olarak bilgilendirilmiş onam formu alınmıştır.

Ölçek verilerinin istatistik analizinin yapılması: Deneme uygulama sonrası elde edilen veriler doğrultusunda ölçeğin yapı geçerliğini değerlendirmek, güvenilirliğini belirlemek ve ölçülmek istenen özelliğin alt boyutlarını ortaya koyabilmek amacıyla faktör analizinden yararlanılmıştır.

Ölçeğe son hâlinin verilmesi: Ölçeğin son aşamasında, geçerlik ve güvenilirlik analizleri sonrası gerekli maddeler çıkarılarak geçerlik ve güvenilirliği kanıtlanan ölçeğe nihai şekli verilmiştir. Ölçeğin yönergesi, maddeleri ve seçenekleri Türkçe ifadelerle oluşturulmuştur. Likert ölçek dörtlü derecelendirilerek 1: “Hiç katılmıyorum”; 2: “Katılmıyorum”; 3: “Katılıyorum”; 4: “Tamamen katılıyorum” şeklinde puanlanmaktadır. Ölçekte olumsuz maddeye yer verilmemiştir. Ölçekten alınabilecek en düşük puan 36 en yüksek puan ise 144 olarak belirlenmiştir. Ölçekten alınan puanlar arttıkça öğrencilerin ölçme ve değerlendirme okuryazarlığı öz yeterlik düzeyleri olumlu yönde arttığı söylenebilir.

Bulgular

Kapsam Geçerliği

Araştırmaya katılan 13 uzmanın kapsam geçerlik değerlendirmeleri Çizelge 2.’de yer almaktadır:

Çizelge 2. Kapsam Geçerlik Oranı (KGO)

Madde	Yeterli görüş sayısı(n)	KGO	Madde	Yeterli görüş sayısı(n)	KGO	Madde	Yeterli görüş sayısı(n)	KGO
m1	11	0,692	m17	7	0,077	m33	10	0,538
m2	10	0,538	m18	10	0,538	m34	13	1,000
m3	10	0,538	m19	7	0,077	m35	13	1,000
m4	10	0,538	m20	12	0,846	m36	10	0,538
m5	12	0,846	m21	10	0,538	m37	10	0,538
m6	12	0,846	m22	12	0,846	m38	10	0,538
m7	12	0,846	m23	10	0,538	m39	10	0,538
m8	13	1,000	m24	10	0,538	m40	11	0,692
m9	10	0,538	m25	9	0,538	m41	13	1,000
m10	10	0,538	m26	10	0,538	m42	12	0,846
m11	11	0,692	m27	12	0,846	m43	5	-0,231
m12	8	0,231	m28	10	0,538	m44	6	-0,077
m13	12	0,846	m29	10	0,538	m45	10	0,538
m14	10	0,538	m30	12	0,846	m46	13	1,000
m15	10	0,538	m31	10	0,538	m47	6	-0,077
m16	11	0,692	m32	12	0,846	m48	13	1,000
						m49	12	0,846
Toplam Kapsam Geçerlik İndeksi (KGİ)								0,696

Ölçek maddeleri kritik değere göre (0,538) değerlendirilmiş, elde edilen KGO sonuçları doğrultusunda kritik değerin altında kalan maddeler (m12, m17, m19, m43, m44, m47) ölçekten çıkarılmıştır. Kapsam geçerlik oranı hesaplamaları doğrultusunda deneme uygulamasına alınacak ölçek maddeleri, 43 madde olarak belirlenmiştir. Ayrıca bu maddeler için elde edilen KGİ değeri (0,696), 0,538 olarak belirlenen kritik değerden büyük olarak tespit edildiğinden tüm ölçeğin kapsam geçerliğinin istatistiksel olarak anlamlı olduğu söylenebilir.

Yapı Geçerliği

Ölçeğin yapı geçerliğini sağlayabilmek için AFA ve DFA değerleri incelenmiştir.

Açımlayıcı Faktör Analizi (AFA)

Ölçeğin faktör yapısını keşfetmek için AFA yapılmıştır. Faktör çıkartma yöntemi olarak maksimum olabilirlik (maximum likelihood) tercih edilmiştir. Ayrıca bu süreçte elde edilen faktör yapısının yorumlanabilirliğini arttırmak amacıyla faktörler arası korelasyonun sıfır olduğu varsayımı doğrultusunda Varimax döndürme yöntemi uygulanmıştır. Bu yöntem sayesinde, maddelerin belirli bir faktörde yüksek yük almasını sağlayarak faktör yapısının daha açıklığı ve yorumlanabilirliği artırılmıştır. Veri setinin AFA için uygunluk durumunu belirlemek amacıyla Kaiser-Meyer-Olkin (KMO) katsayısı ve Bartlett testi incelenmiştir.

Değişkenler arasındaki korelasyon matrisinin faktör analizi için anlamlı olduğunu gösteren Bartlett's Küresellik Testi sonucunun istatistiksel olarak anlamlı olduğu ($p < .01$) bulunmuştur. Örneklem büyüklüğü ve veri yapısının faktör analizi için yeterliğini gösteren Kaiser-Meyer-Olkin (KMO) katsayısı ise .893 olarak hesaplanmıştır. Bu değer, alan yazınında kabul edilen .80 üzerindeki değerler göz önüne alındığında, veri yapısının faktör analizi için mükemmel düzeyde uyumlu olduğunu göstermektedir. Bu bulgular, veri setinin ölçek yapısını ortaya koymak için son derece elverişli olduğunu kanıtlamıştır.

Çizelge 3. Açımlayıcı Faktör Analizi Özet Çizelgesi

Boyutlar	Maddeler	Faktör yükü	Varyans %
Dil Becerileri (Dinleme, Konuşma, Okuma, Yazma)	Ölçme ve değerlendirmeye yönelik dilsel görevleri gerçek yaşamdaki dil olaylarına benzer seçebilirim.	,705	%17,99
	Okuma becerisi sınav metinlerini, öğrencilerin düzeyine uygun okunabilirlik özelliklerine göre seçebilirim.	,694	
	Dil yapılarını alımlama ve üretime yönelik bir bütün olarak değerlendirebilirim.	,671	
	Sınav metinlerini öğrencinin dil seviyelerine uygun olarak seçebilirim.	,664	
	Dinleme metinlerinin sesletim hızını dil seviyelerine uygun olarak düzenleyebilirim.	,651	
	Yazma becerisinin değerlendirilmesinde açık ve anlaşılır bir yönerge kullanabilirim.	,647	
	Yazma becerisini; dilsel bileşen, toplum dilbilimsel bileşen ile edimbilimsel bileşen boyutlarını göz önünde bulundurarak değerlendirebilirim.	,645	
	Kullandığım öğretim programları ile ölçme ve değerlendirme uygulamaları arasında tutarlılık sağlayabilirim.	,620	

	Sınav metinlerinde yer alan söz varlığını öğrencilerin dil düzeyine uygun seçebilirim.	,605	
	Konuşma becerisini; sözlü üretim (kendini ifade etme) ve sözlü etkileşim (karşılıklı konuşma) olmak üzere iki farklı durumda ölçüp değerlendirebilirim.	,597	
	Konuşma becerisini değerlendirmeye yönelik dereceli puanlama anahtarı kullanabilirim.	,577	
	Yazma becerisini değerlendirmeye yönelik dereceli puanlama anahtarı kullanabilirim.	,558	
	Hedef kitlenin ihtiyacına göre dinleme sayısını belirleyebilirim.	,556	
	Yazma becerisinin değerlendirilmesinde metin türlerinden (deneme, mektup, günlük vs.) uygun olanı seçebilirim.	,528	
	Konuşma becerisinin ölçümünü en az bir dış değerlendirici ile iş birliği içerisinde gerçekleştirebilirim.	,503	
Temel Kavramlar	Ölçme ve değerlendirme araçlarında geçerliğin anlamını açıklayabilirim.	,840	%16,66
	Ölçme türlerini açıklayabilirim.	,831	
	Ölçme ve değerlendirme araçlarında güvenilirliğin anlamını açıklayabilirim.	,826	
	Değerlendirme türlerini açıklayabilirim.	,815	
	Sonuç odaklı değerlendirmenin anlamını açıklayabilirim.	,795	
	Süreç odaklı değerlendirmenin anlamını açıklayabilirim.	,783	
	Ölçme ve değerlendirme araçlarında kullanışlılığın anlamını açıklayabilirim.	,783	
	Ölçme ve değerlendirme kavramları arasındaki farkları ayırt edebilirim.	,729	
Teknoloji Tabanlı Ölçme ve Değerlendirme	Geleneksel veya yenilikçi ölçme ve değerlendirme araçlarını teknolojiye uygun hâle getirebilirim.	,851	%13,27
	Teknoloji tabanlı programlarda ölçme ve değerlendirmeye yönelik karşılaştığım sorunları çözebilirim.	,844	
	Öğrenme çıktılarına uygun teknoloji tabanlı ölçme araçları geliştirebilirim.	,840	
	Sınav süresini teknoloji tabanlı ölçme araçlarına uygun biçimde düzenleyebilirim.	,824	
	Teknoloji tabanlı oyunları ölçme ve değerlendirme amaçlı kullanabilirim.	,816	
	Teknoloji tabanlı ölçme ve değerlendirme araçlarından amaca uygun olanı seçebilirim.	,766	
	Sınav sürecinde tüm öğrencilere eşitlik sağlayabilirim.	,863	%13,18
Etik	Sınav sonuçlarının hatalı olması durumunda gerekli düzeltmeyi yapabilirim.	,856	
	Sınav güvenliği (kopya çekilmemesi vb.) için önlem alabilirim.	,795	
	Sınav evraklarını ve sonuçlarını güvenli bir alanda saklayabilirim.	,785	
	Değerlendirme sürecinin yasal ve etik olması için almam gereken sorumlulukları bilirim.	,658	
	Öğrencilerin din ve milliyet yapısını gözetmeden ölçüp değerlendirebilirim.	,658	
	Sınav sonuçlarının gizliliğini sağlayabilirim.	,628	
Toplam Açıklama Oranı			%61,32

Yabancı/İkinci Dil Olarak Türkçe Öğreticilere Yönelik Ölçme ve Değerlendirme Okuryazarlığı Öz Yeterlik Ölçeği'nde uygulanan açımlayıcı faktör analizi sonucuna göre toplam varyans açıklayıcılık yüzdesi %61,32 olarak tespit edilmiştir. Bununla birlikte

standart faktör yükleri de 0.50 değerinden yüksek bulunduğundan açımlayıcı faktör analizinin yeterli seviyede olduğu belirlenmiştir. Ölçekten 7 madde düşük faktör yüklenimleri ($<0,50$) ya da binişik yük oluşturmaları nedeniyle elenmiştir. Bu maddeler analizden elenirken her seferinde yalnızca 1 madde elenmiş ve analiz yinelenmiştir. Kalan 36 madde ile açımlayıcı faktör analizinde faktör sayısı belirlenirken yalnızca özdeğer > 1 (Kaiser Kriteri) kuralı ile birlikte daha güvenilir sonuçlar elde etmek amacıyla Scree Plot (Yamaç Grafiği) ve alan yazınında en çok önerilen yöntemlerden biri olan Paralel Analiz yöntemleri de kullanılmıştır. Analiz sonucunda, her üç yöntemin de 4 faktörlü bir yapıya işaret ettiği görülmüştür. Bu boyutlar Dil Becerileri (DBC), Temel Kavramlar (TMK), Teknoloji Tabanlı Ölçme ve Değerlendirme (TÖD) ve Etik (ETK) olarak adlandırılmıştır. Analizden elenen maddeler ve elenme sebepleri Çizelge 4'te yer almaktadır.

Çizelge 4. Açımlayıcı Faktör Analizinde Elenen Maddeler ve Elenme Nedenleri

No	Maddeler	Elenme Sebebi
1	14. madde: Dil becerilerini ölçme ve değerlendirmede hem süreç hem de sonuç değerlendirme araçlarını kullanabilirim.	Binişik yük oluşumu
2	15. madde: Dil gelişimini izlemek ve desteklemek için akran ve öz değerlendirme araçları kullanabilirim.	Binişik yük oluşumu
3	16. madde: Dil becerilerini ölçmeye ve değerlendirmeye yönelik araçları düzenli olarak güncelleyebilirim.	Binişik yük oluşumu
4	17. madde: Dil becerilerinin değerlendirilmesinde çoklu ölçme araçlarından amaca uygun olanı seçebilirim.	Binişik yük oluşumu
5	24. madde: Dinleme ve okuma becerilerinin değerlendirilmesinde yazma gerektirmeyen sorulara yer verebilirim.	Düşük faktör yükü değerleri ($<0,50$)
6	27. Madde: Yazma becerisinin değerlendirilmesine yönelik yazma konularını birden çok seçenek arasından seçme hakkı tanırım.	Düşük faktör yükü değerleri ($<0,50$)
7	30. madde: Dil gelişimi hakkında öğrenciye geri bildirim verebilirim.	Düşük faktör yükü değerleri ($<0,50$)

Ölçeğin Alt Boyutlarında Ayrıntılı Güvenilirlik Analizi

Katılımcı öğretmenlerin Yabancı/İkinci Dil Olarak Türkçe Öğreten Öğreticilere Yönelik Ölçme ve Değerlendirme Okuryazarlık Öz Yeterlik Ölçeği alt boyutlarından elde edilen güvenilirlik ve iç tutarlılık değerleri Çizelge 5'te sunulmuştur.

Çizelge 5. Yabancı/İkinci Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme ve Değerlendirme Okuryazarlığı Öz Yeterlik Ölçeği Tutarlılık Güvenilirliği Değerleri

Boyutlar	Madde sayısı	Cronbach's Alpha	Omega (ω)
Dil Becerileri	15	,903	0,938

Temel Kavramlar	8	,945	0,955
Teknoloji Tabanlı Ölçme ve Değerlendirme	6	,930	0,934
Etik	7	,884	0,917
Ölçeğin tamamı	36	,912	0,945

Yabancı/İkinci Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme ve Değerlendirme Okuryazarlığı Öz Yeterlik ölçeğinin iç tutarlılık ve güvenilirliği, iki temel metrik olan Cronbach's Alpha (α) ve McDonald's Omega (ω) katsayıları ile değerlendirilmiştir.

Çizelge 5'te yer alan bulgular, ölçeğin tamamının Cronbach's Alpha değerinin .912 olduğunu göstermektedir. Bu değer, alan yazınında genel kabul gören ve yüksek güvenilirliği ifade eden .90 eşiğinin üzerindedir. Bu sonuç, ölçeğin 36 maddeden oluşan bütünüyle son derece tutarlı bir ölçüm aracı olduğunu kanıtlamaktadır. Alt boyutların güvenilirlik değerleri incelendiğinde ise şu sonuçlara ulaşılmıştır:

- Temel Kavramlar boyutu, .945'lik Cronbach's Alpha değeri ve .955'lik Omega değeri ile ölçeğin en yüksek güvenilirlik katsayısına sahip boyutudur. Bu durum, boyutun maddelerinin ölçtüğü kavramın son derece net ve tutarlı olduğunu gösterir.
- Dil Becerileri ($\alpha=.903$) ve Teknoloji Tabanlı Ölçme ve Değerlendirme ($\alpha=.930$) boyutları da mükemmel düzeyde güvenilirliğe sahiptir. Her iki boyut için de Omega değerlerinin Alpha değerlerinden biraz daha yüksek olması (.938 ve .934), bu boyutların güvenilirliğinin Alpha katsayısından daha güçlü bir şekilde doğrulandığını gösterir.
- Etik boyutu, .884'lük Cronbach's Alpha değeriyle diğer boyutlara göre biraz daha düşük olmasına rağmen, bu değer de yüksek güvenilirlik seviyesini teyit etmektedir. Bu boyut için hesaplanan .917'lik Omega değeri, Cronbach's Alpha'dan daha yüksek çıkarak bu boyutun güvenilirliğinin aslında daha da güçlü olduğunu ortaya koymaktadır.

Omega (ω) katsayısı, Cronbach's Alpha'ya kıyasla maddelerin faktör yüklerindeki farklılıkları da hesaba kattığı için daha doğru bir güvenilirlik tahminidir. Çizelgedeki tüm Omega değerlerinin .90'ın üzerinde olması, ölçeğin her bir boyutunun ve bütününe mükemmel düzeyde güvenilir olduğunu güçlü bir şekilde kanıtlamaktadır.

Bu kapsamlı analiz sonuçları, Yabancı/İkinci Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme ve Değerlendirme Okuryazarlığı Öz Yeterlik Ölçeği'nin geçerli ve bilimsel araştırmalarda güvenle kullanılabilecek, yüksek güvenilirliğe sahip bir ölçüm aracı olduğunu net bir şekilde ortaya koymaktadır.

Doğrulayıcı Faktör Analizi (DFA)

Doğrulayıcı faktör analizinde örnek büyüklüğü arttıkça, özellikle 200'den büyük örneklerde Ki Kare (χ^2) değeri de yüksek çıkmakta ve Ki-Kare (χ^2) testinin istatistiksel anlamlılık düzeyi düşük çıkmaktadır (Bollen, 1989; Fornell & Larcker, 1981; Bagozzi vd., 1999). Araştırma için kullanılan ölçeklerin doğrulayıcı faktör analizi değerlendirmesinde ve genel test edilen modellerin uygun olup olmadığına, serbestlik derecesi ile düzeltilmiş Ki- Kare (χ^2) değeri (Ki-Kare değeri/Serbestlik derecesi), diğer uyum iyiliği indeksleri ve standartlaştırılmış artık (residual) kovaryans matrisinde yer alan değerlerin incelenmesi sonucunda karar verilmiştir (Bayram, 2013).

Çizelge 6. Doğrulayıcı Faktör Analizinde Kullanılan Uyum İyiliği İndeksleri ve Uyum Değerleri (Meydan, 2011)

İndeksler	İyi Uyum	Kabul Edilebilir Uyum
χ^2 / df	$0 \leq \chi^2/df \leq 2$	$2 < \chi^2/df \leq 3$
CFI	$\geq 0,95$	$\geq 0,90$
IFI	$\geq 0,95$	$\geq 0,90$
SRMR	$\leq 0,05$	$,06 \leq SRMR \leq ,08$
RMSEA	$\leq 0,05$	$,06 \leq RMSEA \leq ,08$

Yapısal eşitlik modellemesi (YEM), parametre tahminlerinin doğruluğu ve model uyum indekslerinin güvenilirliği için çoklu normal dağılım (multivariate normality) varsayımının karşılanması gerektirir (Hair ve diğerleri, 2019). Tekil normallik (univariate normality) çoklu normalliğin sağlanması için gerekli bir koşul olsa da tek başına yeterli değildir (Tabachnick & Fidell, 2019). Bu nedenle, bu çalışmada çoklu normallik durumu, Mardia's çoklu basıklık katsayısı (Mardia's multivariate kurtosis) ve bu katsayıya ait kritik oran (CR) değeri ile değerlendirilmiştir.

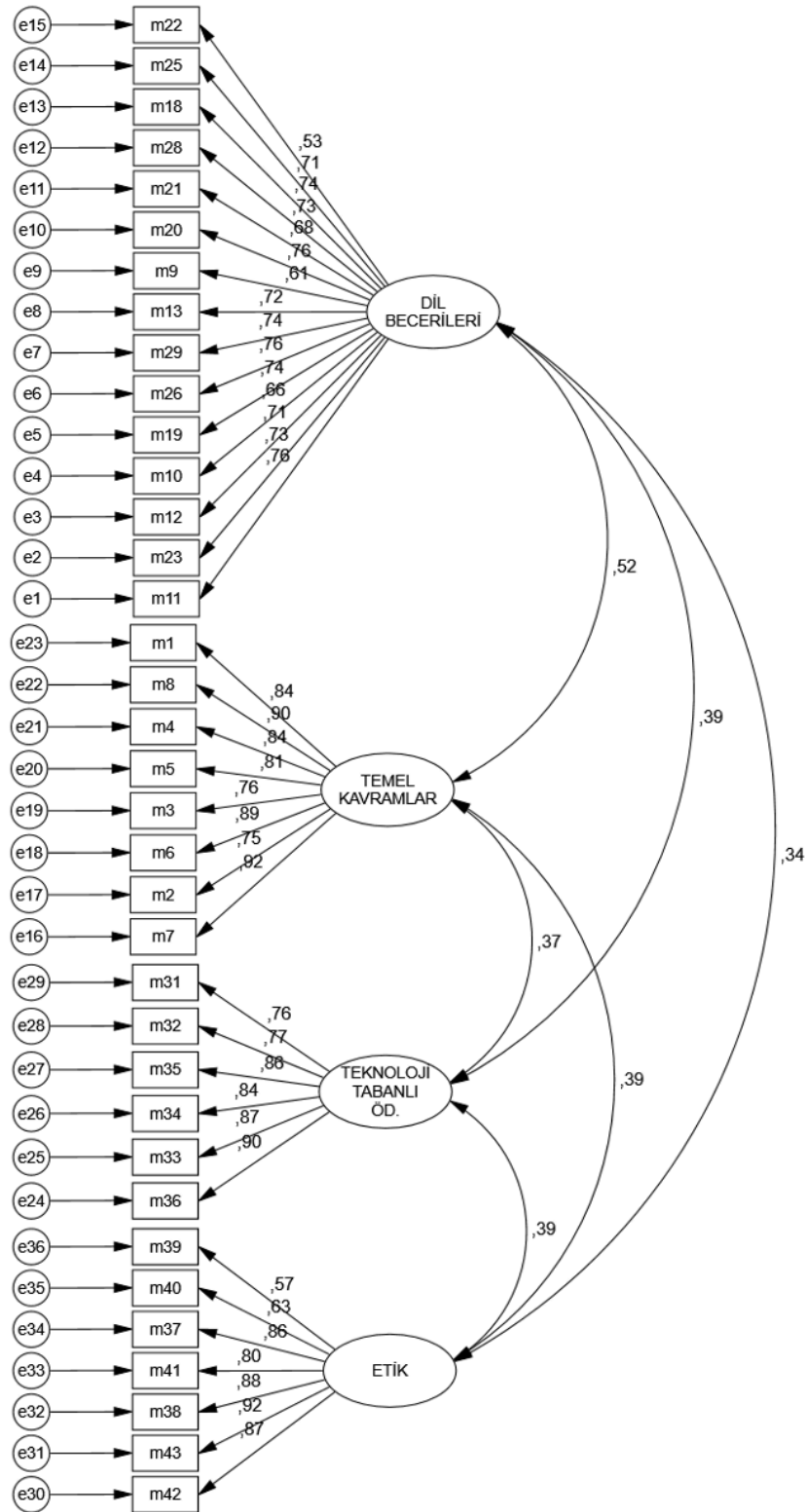
Alan yazınında, çoklu normalliğin önemli ölçüde ihlal edilmediğini gösteren çeşitli eşik değerler önerilmektedir. Tabachnick ve Fidell (2019), çoklu normallik için Mardia's basıklık katsayısının kritik oranının 5'in altında olmasını önermektedir. Doğrulayıcı faktör analizi (DFA) için en uygun tahmin yönteminin belirlenmesinde, veri setinin çoklu normallik varsayımı temel alınmıştır. Analiz için kullanılan 202 katılımcıdan oluşan veri setinde, çoklu normallik durumu Mardia's katsayısı (Mardia's multivariate kurtosis) ile değerlendirilmiştir.

Elde edilen Mardia's katsayısı 4,55 olarak bulunmuştur. Alan yazınında, bu değer için kritik oranı (critical ratio) 5'in altında olduğunda çoklu normallik varsayımının karşılandığı kabul edilmektedir (Tabachnick & Fidell, 2019). Bu durum, veri setinin DFA için en uygun ve en sık kullanılan tahmincilerden biri olan Maksimum Olabilirlik (Maximum Likelihood - ML) yöntemini kullanmaya elverişli olduğunu göstermektedir.

ML yöntemi, geniş örneklerde ve normal dağılım varsayımı sağlandığında en verimli ve sapmasız parametre tahminlerini sunmasıyla bilinir. Bu nedenle, çalışmanın analizlerinin güvenilirliğini en üst düzeye çıkarmak amacıyla DFA ve yapısal eşitlik modellemesi analizlerinde ML yöntemi tercih edilmiştir. Bu seçim, elde edilen model uyum indekslerinin ve parametre standart hatalarının güvenilir olduğunu garanti eder.

Doğrulamalı faktör analizi öncesinde, veri setinin güvenilirliğini ve modelin tahminlerinin geçerliliğini sağlamak amacıyla çok değişkenli aykırı değer (multivariate outlier) analizi yapılmıştır. Tek değişkenli aykırı değerler (univariate outliers) kolayca tespit edilebilse de YEM gibi çok değişkenli teknikler için bir araya geldiklerinde normallik varsayımını bozan gözlemlerin belirlenmesi kritik öneme sahiptir. Bu amaçla, her bir gözlemin çok değişkenli uzaydaki merkeze olan uzaklığını ölçen Mahalanobis Mesafesi (Mahalanobis Distance) kullanılmıştır (Tabachnick & Fidell, 2019, s. 147-151). Mahalanobis mesafesi, değişkenler arasındaki korelasyonu da dikkate alarak aykırı değerleri daha hassas bir şekilde belirleme olanağı sunar. Bu analizde, Mahalanobis mesafesi değeri için bir Ki-kare (χ^2) dağılımı referans alınmış ve anlamlılık düzeyi $p < .001$ olarak belirlenmiştir. Elde edilen Mahalanobis mesafesi değerleri ve ilgili p değerleri incelenerek, analizden önce DFA modelinin uyum ve tahmin parametrelerini olumsuz etkileme potansiyeli olan gözlemler (9 adet) belirlenmiştir. Bu gözlemlerin, veri giriş hatalarından ya da dikkatsiz yanıtlamadan kaynaklandığına karar verilmiş ve analiz dışı bırakılmıştır. Bu ön işlem, DFA'nın varsayımlarına daha uygun bir veri yapısı oluşturarak, model uyum indekslerinin ve parametre tahminlerinin güvenilirliğini artırmıştır (Hair ve diğerleri, 2019).

Yabancı/İkinci Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme ve Değerlendirme Okuryazarlığı Öz Yeterlik Ölçeği doğrulamalı faktör analizinde, (n=193) açımlayıcı faktör analizinden elde edilen 4 boyutlu yapının geçerliliği sınanmıştır. Analizde standart faktör yüklerinin (,52; ,92) değerleri arasında yer aldığı AMOS programında çizilen Faktör analizi grafik yapısı Çizim 1'de izlenmektedir.

Çizim 1. Doğrulayıcı Faktör Analizi Grafik Yapısı

Çizim 1 incelendiğinde doğrulayıcı faktör analizinden elde edilen model parametreleri; ($p < 0.05$) olmak üzere χ^2 (1220,618), χ^2/df (2,029) bulunduğundan modelin anlamlı olduğu görülmektedir. Modelin uyum indis değerleri IFI (.915), CFI (.909), SRMR (.0529), RMSEA (.0714) kabul edilebilir uyum sınırları içinde yer aldığından modelin geçerli olduğu anlaşılmaktadır.

Çizelge 7. Doğrulamalı Faktör Analizi Özet Çizelgesi

Alt boyut	Madde	Faktör yükü	Standart faktör yükü	Z	P
Dil Becerileri	m11	1,000	,760		
	m23	1,008	,732	10,765	***
	m12	1,050	,711	12,231	***
	m10	,887	,663	9,631	***
	m19	1,065	,740	10,906	***
	m26	1,037	,763	11,296	***
	m29	1,169	,740	10,912	***
	m13	1,039	,721	12,636	***
	m9	,865	,612	8,808	***
	m20	1,145	,759	11,228	***
	m21	1,198	,679	9,891	***
	m28	1,189	,733	10,785	***
	m18	1,129	,739	10,891	***
	m25	1,063	,709	10,382	***
	m22	1,090	,528	7,499	***
Temel Kavramlar	m7	1,000	,922		
	m2	,806	,750	14,122	***
	m6	,918	,885	20,342	***
	m3	,829	,761	14,508	***
	m5	,835	,809	16,384	***
	m4	,820	,836	17,647	***
	m8	,960	,904	21,591	***
	m1	,827	,841	17,881	***
Teknoloji Tabanlı Ölçme ve Değerlendirme	m36	1,000	,895		
	m33	,995	,866	17,409	***
	m34	1,059	,844	16,550	***
	m35	1,072	,857	17,075	***
	m32	,895	,769	13,823	***
	m31	,846	,755	13,502	***
Etik	m42	1,000	,874		
	m43	1,008	,920	19,274	***
	m38	1,046	,879	17,506	***
	m41	,968	,796	14,511	***
	m37	,967	,860	16,740	***
	m40	,914	,633	10,231	***
	m39	,851	,574	8,967	***

***p<0,001

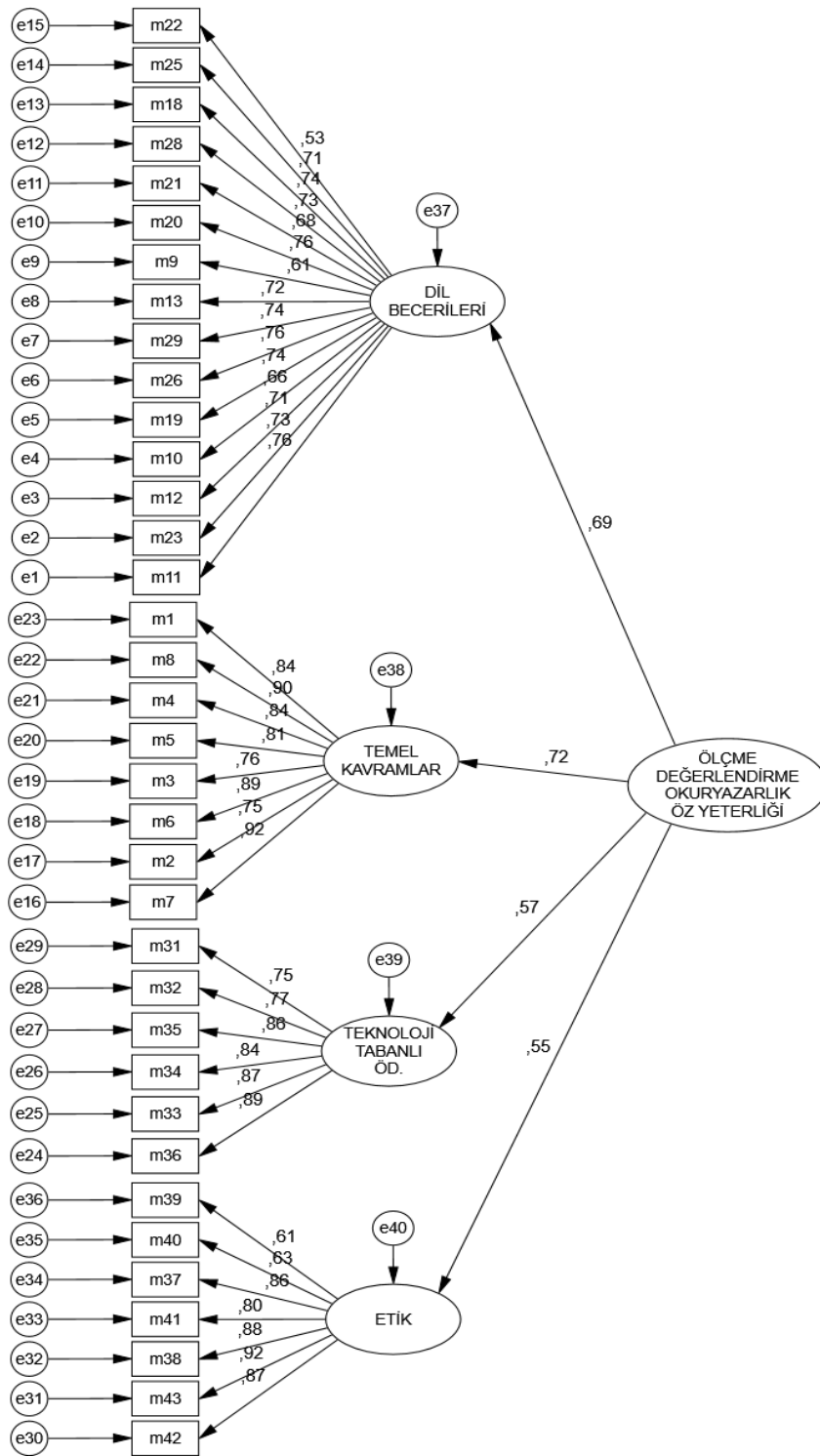
Yabancı/İkinci Dil Olarak Türkçe Öğreten Öğreticilere Yönelik Ölçme ve Değerlendirme Okuryazarlık Öz Yeterlik Ölçeği doğrulamalı faktör analizinden elde edilen

standart faktör yüklerinin tamamında bulunan standart yükler istatistiksel olarak anlamlı olduğu belirlenmiştir ($p<0,05$).

2. Düzey Doğrulayıcı Faktör Analizi

Doğrulayıcı faktör analizinin ikinci aşamasında, açımlayıcı faktör analizinden elde edilen 4 boyutlu yapının, tek bir üst kavram olan Ölçme Değerlendirme Okuryazarlık Öz Yeterliği'nin alt boyutları olarak değerlendirilip değerlendirilemeyeceği sınanmıştır. Bu amaçla, ikinci düzey doğrulayıcı faktör analizi (second-order confirmatory factor analysis) modeli uygulanmıştır (Byrne, 2016).

Bu model, her bir boyutun (Dil Becerileri, Temel Kavramlar, Teknoloji Tabanlı Ölçme ve Değerlendirme ve Etik) kendi maddelerini açıklamasının yanı sıra, bu dört boyut arasındaki korelasyonun nedenini de açıklayan tek bir genel üst faktör (second-order factor) tarafından açıklanıp açıklanmadığını değerlendirir (Kline, 2015). Bu yaklaşım, ölçeğin boyutlarının tek bir yapıyı temsil ettiğine dair güçlü bir kanıt sunarak, ölçeğin yapısal geçerliliğini ve kuramsal bütünlüğünü desteklemektedir (Hair ve diğerleri, 2019). Elde edilen sonuçlar, Yabancı/İkinci Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme ve Değerlendirme Okuryazarlığı Öz Yeterliği'nin çok boyutlu ancak tek bir üst yapıya sahip olduğunu doğrulamıştır.

Çizim 2. 2. Düzey Doğrulayıcı Faktör Analizi Grafik Yapısı

Doğrulayıcı faktör analizinden elde edilen model parametreleri; ($p < 0.05$) olmak üzere χ^2 (1254,09), χ^2/df (2,131) bulunduğundan modelin anlamlı olduğu görülmektedir. Modelin uyum indis değerleri IFI (.913), CFI (.922), SRMR (.0435) ve RMSEA (.0714) kabul edilebilir uyum sınırları içinde yer aldığından modelin geçerli olduğu anlaşılmaktadır.

Çizelge 8. 2. Düzey Doğrulayıcı Faktör Analizi Özet Çizelgesi

Alt boyut	Madde	Faktör yükü	Standart faktör yükü	Z	P
Ölçme ve Değerlendirme Okuryazarlık Öz Yeterliği	Dil Becerileri	1,000	,692		
	Temel Kavramlar	1,382	,716	5,866	***
	Teknoloji Tabanlı Ölçme ve Değerlendirme	1,229	,566	5,349	***
	Etik	1,105	,553	5,270	***
Dil Becerileri	m11	1,000	,760		
	m23	1,009	,733	10,787	***
	m12	1,050	,711	12,219	***
	m10	,887	,663	9,632	***
	m19	1,064	,740	10,900	***
	m26	1,037	,764	11,308	***
	m29	1,169	,740	10,903	***
	m13	1,038	,720	12,622	***
	m9	,866	,613	8,818	***
	m20	1,145	,759	11,230	***
	m21	1,197	,679	9,883	***
	m28	1,189	,733	10,788	***
	m18	1,129	,740	10,899	***
	m25	1,063	,708	10,376	***
	m22	1,090	,528	7,498	***
Temel Kavramlar	m7	1,000	,922		
	m2	,806	,750	14,133	***
	m6	,918	,885	20,356	***
	m3	,829	,761	14,507	***
	m5	,835	,808	16,380	***
	m4	,820	,836	17,653	***
	m8	,960	,904	21,597	***
	m1	,826	,841	17,874	***
Teknoloji Tabanlı Ölçme ve Değerlendirme	m36	1,000	,895		
	m33	,996	,867	17,431	***
	m34	1,059	,844	16,545	***
	m35	1,074	,857	17,088	***
	m32	,894	,768	13,783	***
	m31	,845	,754	13,455	***
Etik	m42	1,000	,872		
	m43	1,014	,924	19,376	***
	m38	1,050	,881	17,587	***

Alt boyut	Madde	Faktör yükü	Standart faktör yükü	Z	P
	m41	,970	,796	14,484	***
	m37	,967	,858	16,688	***
	m40	,905	,626	10,081	***
	m39	,903	,609	9,529	***

***p<0,001

Ölçeğin 2. düzey doğrulayıcı faktör analizinden elde edilen standart faktör yüklerinin tamamında bulunan standart yükler istatistiksel olarak anlamlı olduğu belirlenmiştir ($p<0,05$). Alt boyutların tek bir kavramın alt boyutu olduğu bu modelle ortaya konmuştur.

Yakınsama ve Ayrışma Geçerliliği

Yapı geçerliliği (construct validity), bir ölçüm aracının teorik bir yapıyı ne kadar doğru ölçtüğünü gösterir. Yapı geçerliliği, yakınsak ve diskriminant olmak üzere iki ana bileşenden oluşur. Diskriminant geçerliliği (discriminant validity), birbirinden teorik olarak farklı olması beklenen yapıların, ampirik olarak da yeterli düzeyde ayrıştığını doğrulamayı amaçlar. APA 7'ye göre, diskriminant geçerliliğini test etmek için en yaygın kullanılan iki yöntem, Fornell-Larcker Kriteri ve Heterotrait-Monotrait (HTMT) Oranı'dır. Fornell-Larcker kriterine göre diskriminant geçerliliği, her bir yapının ortalama varyans açıklayıcılığı (Average Variance Extracted-AVE) karekökünün, ilgili yapının diğer tüm yapılarla olan en yüksek korelasyon katsayısından daha büyük olmasıyla sağlanır (Fornell & Larcker, 1981). Bir ölçeğin güvenilirliği, aynı kavramı ölçen maddelerin tutarlılığını ifade eder. Bu çalışmada, Yabancı/İkinci Dil Olarak Türkçe Öğreten Öğreticilere Yönelik Ölçme ve Değerlendirme Okuryazarlık Öz Yeterlik Ölçeği'nin güvenilirliği iki ana metrikle değerlendirilmiştir: Cronbach's Alpha (CA) ve Birleşik Güvenilirlik (Composite Reliability-CR).

Cronbach's Alpha (CA): Ölçeğin iç tutarlılığını ölçen en yaygın katsayıdır. Bir ölçeğin maddelerinin ne kadar tutarlı bir şekilde aynı yapıyı ölçtüğünü gösterir. Elde edilen CA değerleri, tüm alt boyutlarda .914 ile .953 arasında bulunmuştur. Literatürde, .90'ın üzerindeki CA değerleri "mükemmel" iç tutarlılık olarak kabul edilir. Bu bulgular, ölçeğin maddelerinin son derece tutarlı olduğunu ve katılımcıların yanıtlarının güvenilir olduğunu göstermektedir.

Birleşik Güvenilirlik (CR): Bu metrik, özellikle doğrulayıcı faktör analizi (DFA) bağlamında, her bir boyutun latent değişkeni ne kadar iyi temsil ettiğini gösteren daha gelişmiş bir güvenilirlik ölçüsüdür. CR, maddelerin standart faktör yüklerini kullanarak hesaplanır ve klasik Cronbach's Alpha'ya göre daha hassas bir değerlendirme sunar (Raykov, 1997). Çalışmada, tüm boyutlar için hesaplanan CR değerleri .922 ile .947 arasında değişmektedir. Bu değerlerin, literatürde güvenilirlik için genel kabul gören .70 eşliğinin oldukça üzerinde olması, ölçeğin boyutlarının güvenilir bir şekilde ölçüldüğünü ve bulguların sağlam olduğunu kanıtlamaktadır. Bu iki güvenilirlik ölçütünün de yüksek değerler vermesi, Yabancı/İkinci Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme ve

Değerlendirme Okuryazarlığı Öz Yeterlik Ölçeği'nin son derece güvenilir bir ölçüm aracı olduğu sonucuna varmamızı sağlamıştır.

Çizelge 9. Yapıların Ortalama Varyans Açıklayıcılığı (AVE) Karekökleri ve Yapılar Arası Korelasyon Katsayıları

Boyut	DBC	TMK	TÖD	ETK
Dil Becerileri (DBC)	(,712)			
Temel Kavramlar (TMK)	,548**	(,830)		
Teknoloji Tabanlı Ölçme-değerlendirme (TÖD)	,404**	,369**	(,828)	
Etik (ETK)	,472**	,277**	,293**	(,800)
Cronbach's Alpha (CA)	,930	,945	,961	,921
Composite reliability (CR)	,941	,933	,931	,912
Açıklanan Ortalama Varyans (AVE)	,508	,689	,687	,640

*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$

Ayrışma geçerliliği, bir ölçeğin farklı boyutlarının birbirinden bağımsız ve özgün kavramları ölçtüğünü kanıtlama sürecidir. Bu geçerliliğin en yaygın göstergelerinden biri, Açıklanan Ortalama Varyans (AVE) değerinin karekökünün, ilgili boyutun diğer boyutlarla olan korelasyon katsayılarından daha yüksek olmasıdır. Bu kriter, Fornell ve Larcker (1981) tarafından geliştirilen ve alanda geniş kabul gören bir kuraldır.

Çizelge 9'da sunulan bulgular, bu kriterin tüm boyutlar için karşılandığını göstermektedir. Bu durum, her bir boyutun, diğer boyutlarla yüksek düzeyde ilişkili olsa bile, kendine özgü bir kavramı ölçtüğünü ve ölçeğin boyutlarının kuramsal olarak birbirinden ayrıştığını doğrulamaktadır. Bu güçlü ayrışma geçerliliği, ölçeğin boyutlarının tek bir kavramın farklı yönlerini değil, bağımsız ve farklı yapıları temsil ettiğini kanıtlamaktadır.

Tartışma ve Sonuç

Bu araştırmada yabancı/ikinci dil olarak Türkçe öğreten öğretmenlerin ölçme ve değerlendirme okuryazarlık öz yeterliklerini belirlemek amacıyla geliştirilen “Yabancı/İkinci Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme ve Değerlendirme Okuryazarlığı Öz Yeterlik Ölçeği”nin güvenilirlik ve geçerlik analizleri yapılmıştır. İlk aşamada alan yazını incelenerek 86 madde içeren madde havuzu oluşturulmuştur. Ölçeğin kapsam geçerliğini sağlamak için on uzman görüşü alınarak gerekli düzenlemeler yapılmış ve ardından maddeler on üç uzmanın görüşü alınmak üzere tekrar değerlendirmelerine sunulmuştur. Uzmanların geri bildirimleri doğrultusunda Lawshe (1975) tekniği kullanılarak maddelerin kapsam geçerlik oranları (CVR) hesaplanmış ve kritik değerin (.538) altında kalan maddeler ölçekten çıkarılmıştır. Benzer şekilde Lawshe tekniğini kapsam geçerliğine delil olarak kullanan çalışmalar (Akçay & Şahin, 2015; Sallabaş ve diğerleri, 2018; İnan Yıldız & Karadağ, 2023; Sevim & Varışoğlu, 2023) bulunmaktadır. Ayrıca maddelerin KGİ (.696) olarak bulunmuştur. Bu oranlar, kapsam geçerliğinin yüksek olduğunu göstermiştir.

Ölçeğin yapı geçerliğini sağlamak için açımlayıcı faktör analizi ve doğrulayıcı faktör analizi yapılmıştır. Bu yöntemi destekler nitelikte her iki analizin bir arada tercih

edildiği çalışmalar (Ülper ve diğerleri, 2013; Yaşar, 2014; Çetinkaya & Dikici 2023) mevcuttur. Açımlayıcı faktör analizinde Maximum likelihood metodu kullanılmış ve Varimax yöntemiyle de rotasyon uygulanmıştır. Yabancı/ikinci dil olarak Türkçe öğreten 217 öğreticinin katılmış olduğu veri setinin Bartlett's testi ($p<0,01$) ve Kaiser Meier Olkin (KMO) katsayısı (.893) olarak belirlenmiştir. Bu değerler sonucunda veri yapısının faktör analizi için iyi düzeyde uyumlu olduğunu göstermiştir (Çokluk ve diğerleri, 2023).

Henson ve Roberts (2006) toplam varyans yüzdesinin %52 ve üzeri olması gerektiğini belirtmiştir. Ölçeğin toplam varyans açıklayıcılık yüzdesi %61,32 ve standart faktör yükleri de 0.50 değerinden yüksek bulunduğundan açımlayıcı faktör analizinin yeterli düzeyde olduğu görülmüştür. AFA sonrası hesaplanan DFA'dan elde edilen standart faktör yüklerinin tamamında bulunan standart yükler istatistiksel olarak anlamlı olduğu belirlenmiştir ($p<0,05$). Modelin uyum indeks değerleri IFI (.913), CFI (.922), SRMR (.0435), RMSEA (.0714) kabul edilebilir uyum sınırları içinde yer aldığından modelin geçerli olduğu sonucuna varılmıştır (Meydan, 2015). Ayrıca tüm alt boyutların tek bir kavramın alt boyutu olduğu saptanmıştır.

Yabancı/İkinci Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme ve Değerlendirme Okuryazarlığı Öz Yeterlik Ölçeği'nin iç tutarlılığı, Cronbach's Alpha (α) ve McDonald's Omega (ω) katsayıları ile değerlendirilmiş; ölçeğin tamamı için $\alpha=.912$ bulunmuştur. Alt boyutlardan Temel Kavramlar en yüksek güvenilirlik değerine ($\alpha=.945$, $\omega=.955$) sahipken, Dil Becerileri ($\alpha=.903$) ve Teknoloji Tabanlı Ölçme ve Değerlendirme ($\alpha=.930$) boyutları da yüksek güvenilirlik göstermiştir. Etik boyutu $\alpha=.884$ olmasına rağmen $\omega=.917$ ile güçlü güvenilirlik sunmaktadır. Tüm Omega değerlerinin .90 üzerinde olması, ölçeğin hem boyutlar hem de bütün olarak mükemmel düzeyde güvenilir olduğunu göstermektedir.

Yapılan analizler sonucunda güvenilirliği ve geçerliği kanıtlanan "Yabancı/İkinci Dil Olarak Türkçe Öğreten Öğreticilere Yönelik Ölçme ve Değerlendirme Okuryazarlık Öz Yeterlik Ölçeği" otuz altı madde ve dört boyutlu bir yapıya sahiptir. 1. Boyut: "Dil Becerileri", 2. Boyut: "Temel Kavramlar", 3. Boyut: "Teknoloji Tabanlı Ölçme ve Değerlendirme" ve 4. Boyut: "Etik" olarak adlandırılmıştır.

Bu veriler doğrultusunda araştırma kapsamında geliştirilen "Yabancı Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme ve Değerlendirme Okuryazarlık Öz Yeterlik Ölçeği"nin öğreticilerin ölçme ve değerlendirme okuryazarlık öz yeterliklerini belirleyebilecek nitelikte geçerli ve güvenilir bir araç olduğu sonucuna varılmıştır.

Öneriler

- Ölçek yabancı/ikinci dil olarak Türkçe öğretimi ölçme ve değerlendirme okuryazarlığını kapsamaktadır. Farklı hedef dil öğretimi için uyarlaması yapılabilir.
- Bu çalışmada öğreticilerin ölçme ve değerlendirme okuryazarlık öz yeterliğine yönelik likert tipi ölçme aracı geliştirilmiştir. Sonraki çalışmalarda öğreticilerin ölçme ve değerlendirme okuryazarlık becerisini ölçümünde kullanılmak üzere senaryo odaklı görevler veya açık uçlu sorulardan oluşan ölçekler geliştirilebilir.

- Ölçeği yabancı/ikinci dil olarak Türkçe öğreten öğreticilere uygulanarak farklı değişkenlere göre öğreticilerin ölçme ve değerlendirme okuryazarlık öz yeterliklerini ortaya koyan çalışmalar yürütülebilir.
- Geliştirilen ölçek, öğreticilerin ölçme ve değerlendirme okuryazarlık becerilerini geliştirmeye yönelik yapılacak deneysel araştırmalarda ön test-son test olarak uygulanabilir.

Kaynakça

- Akçay, A., & Şahin, N. (2015). Türk dilinin eski yazılı örneklerinin ortaokul Türkçe ders kitaplarında metin olarak yer alabilirliğinin incelenmesi. *Uluslararası Eğitim Bilim ve Teknoloji Dergisi*, 1(1), 53-64. <https://dergipark.org.tr/tr/pub/uebt/issue/21609/232108>
- American Federation of Teachers, National Council on Measurement in Education, & National Education Association (1990). *The standards for Teacher competence in the educational assessment of students*. Retrieved from: <http://files.eric.ed.gov/fulltext/ED323186.pdf>
- American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.). Author.
- Ayre, C., & Scally, A. J. (2014). Critical Values for Lawshe's Content Validity Ratio: Revisiting the Original Methods of Calculation. *Measurement and Evaluation in Counseling and Development*, 47, 79-86. <https://doi.org/10.1177/0748175613513808>
- Bagozzi, R. P., Gopinath, M., & Nyer, P. U. (1999). The role of emotions in marketing. *Journal of the Academy of Marketing Science*, 27(2), 184-206. <https://doi.org/10.1177/0092070399272005>
- Bal, H. (2023). *Araştırma Yöntemleri*. Otorite Yayınları.
- Barın, E., Çangal, Ö., & Başar, U. (2017). Yabancı dil olarak Türkçe öğretimi alanında görev yapacak öğretmenlerin özel alan yeterliklerine ilişkin bir öneri. *International Journal of Language Academy*, 5(20), 81-98. <https://doi.org/10.18033/ijla.3779>
- Bayrakdar, E. (2024). Yabancı/ikinci dil olarak Türkçe öğretiminde konuşma becerisi. Ş. Dilidüzgün (Ed.), *Yabancılar Türkçe öğretimi: Eğiticinin eğitimi* (ss. 300-335). Asos Yayınları.
- Bayram, N. (2013). *Yapısal eşitlik modellemesine giriş: AMOS uygulamaları*. Ezgi Kitabevi Yayınları.
- Beyreli, L. (2017). Türkçenin yabancı dil olarak öğretiminde öğretmen nitelikleri. İçinde H. Develi, C. Yıldız, M. Balcı, İ. Gültekin & D. Melanlıoğlu (Ed.), *Uygulamalı Türkçenin yabancı dil olarak öğretimi el kitabı* (ss. 325-359). Kesit Yayınları.
- Bollen, K. A., & Hoyle, R. H. (2012). Latent variables in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 56-67). Guilford Press.
- Boylu, E. (2019). *Yabancılar Türkçe öğretiminde ölçme değerlendirme uygulamaları ve standart oluşturma* [Yayımlanmamış doktora tezi]. Çanakkale Onsekiz Mart Üniversitesi.
- Bozkurt, B. Ü. (2022). Konuşma eğitimi sürecinde ölçme ve değerlendirme. In G. Çetinkaya (Ed.), *Türkçe eğitimi sürecinde ölçme ve değerlendirme* (s. 145-180). Anı Yayıncılık.
- Bütünler, S. Ö., Yiğit, N., & Çimer, S. O. (2010). Ölçme değerlendirme okuryazarlığı envanterinin Türkçeye uyarlanması. *e-Journal of New World Sciences Academy*, 5(3), 1C0165. <https://dergipark.org.tr/tr/download/article-file/185783>
- Büyüköztürk Ş., Kılıç Çakmak E., Akgün Ö. E., Karadeniz Ş., & Demirel F. (2023). *Eğitimde Bilimsel Araştırma Yöntemleri*. Pegem Akademi Yayıncılık.

- Byrne, B. M. (2016). *Structural equation modeling with Mplus: Basic concepts, applications, and programming*. Routledge.
- Çakan, M. (2004). Öğretmenlerin ölçme-değerlendirme uygulamaları ve yeterlik düzeyleri: ilk ve ortaöğretim. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 37(2), 99-114
- Çarkçı, J. (2020). *Sosyal bilimlerde ölçek geliştirme kılavuzu*. Çizgi Kitabevi Yayınları.
- Çetinkaya, G. ve Dikici, A. (2023). Özetleme özyeterlik algısı ölçeğinin geliştirilmesi. *Diyaletolog-Uluslararası Sosyal Bilimler Dergisi*, 32, 13-33. <http://dx.doi.org/10.29228/diyalektolog.68896>
- Çiçek, S., Asar, A., & Mete, F. (2023). Yabancı dil olarak Türkçe öğreticilerinin özyeterlikleri üzerine bir araştırma. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, (33), 75-89. <https://doi.org/10.29000/rumelide.1279071>
- Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2023). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları*. Pegem Akademi Yayıncılık.
- Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing*, 25(3), 327-347. <https://doi.org/10.1177/0265532208090156>
- DeVellis, R. F. (2022). *Ölçek geliştirme: Kuram ve uygulamalar* (T. Totan, M. Uluman, A. Ş. Avşar, A. S. Sağkal, D. Ç. Eser, & Ö. Bıkmaz, Çev.). Nobel Akademik Yayıncılık.
- Dincel, Ö. (2021). *Yabancılar Türkçe öğretimi ders kitaplarının ölçme değerlendirme açısından incelenmesi ve konuya dair öğretici görüşleri* [Yayımlanmamış doktora tezi]. Trabzon Üniversitesi.
- Erkuş, A. (2022). Ölçek geliştirmeye hazırlık. Acar G. M. ve Özkan Ö. Y.(Ed.)*Tüm yönleriyle ölçek geliştirme süreci*.(s. 1-25). Pegem Akademi Yayıncılık.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.1177/002224378101800104>
- Gottheiner, D. M., & Siegel, M. A. (2012). Experienced middle school science teachers' assessment literacy: Investigating knowledge of students' conceptions in genetics and ways to shape instruction. *Journal of Science Teacher Education*, 23(5), 531-557. <https://doi.org/10.1007/s10972-012-9278-z>
- Güçüyeter, B., & Karadoğan, A. (2019). Türk dili ve edebiyatı öğretmenlerinin ölçme değerlendirme tutumlarını belirlemeye yönelik ölçek geliştirme çalışması: Geçerlik ve güvenirlik. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi*, 17, 47-60. <https://doi.org/10.29000/rumelide.656614>
- Gül, E. (2011). *İlköğretim öğretmen adaylarının ölçme- değerlendirme okuryazarlığı ve ölçme-değerlendirmeye ilişkin tutumlarının belirlenmesi* [Yayımlanmamış yüksek lisans tezi]. Erciyes Üniversitesi.
- Gülbahar, Y., & Büyüköztürk, Ş. (2008). Değerlendirme tercihleri ölçeğinin Türkçeye uyarlanması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 35(35), 148-161. <https://dergipark.org.tr/tr/pub/hunefd/issue/7803/102282>
- Güneş, A. (2007). *Sınıf öğretmenlerinin kendi algılarına göre ölçme ve değerlendirme yeterlikler* [Yayımlanmamış yüksek lisans tezi]. Marmara Üniversitesi.
- Gürbüz, S., & Şahin, F. (2018). *Sosyal bilimlerde araştırma yöntemleri: Felsefe – yöntem – analiz* (5. baskı). Seçkin Yayıncılık.
- Gürsoy, G. A. (2015). *Alan ile ilişkilendirilmiş uygulamalı ölçme ve değerlendirme dersinin öğretmen adaylarının ölçme ve değerlendirme okuryazarlık düzeylerine, tutumlarına ve alan bilgilerine etkisi* [Yayımlanmamış doktora tezi] Gazi Üniversitesi.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.

- Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), 115–135. <https://doi.org/10.1007/s11747-014-0403-8>
- Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological measurement*, 66(3), 393-416. <https://doi.org/10.1177/001316440528248>
- Işıkoğlu, M. (2015). *Yabancı dil olarak Türkçe öğretiminde kullanılan yeterlik sınavlarının madde yazımı bakımından incelenmesi (Mersin ve Sakarya Üniversiteleri örneği)* [Yayımlanmamış yüksek lisans tezi]. Mersin Üniversitesi.
- İnan Yıldız, F., & Karadağ, Ö. (2023). İlkokul düzeyinde yapılandırmacılık temelli öykü değerlendirmeye yönelik rubrik geliştirme çalışması. *Ana Dili Eğitimi Dergisi*, 11(4), 936-950. <https://doi.org/10.16916/aded.1321531>
- Karadoğan, A., & Uyumaz, F. (2022). Yabancı dil olarak Türkçe öğretiminde öğretici yeterliklerinin belirlenmesi. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi* (31), 174-196. <https://doi.org/10.29000/rumelide.1220510>
- Karagöz, Y., & Bardakçı, S. (2020). *Bilimsel araştırmalarda kullanılan ölçme araçları ve ölçek geliştirme*. Nobel Akademik Yayıncılık.
- Karasar, N. (2024). *Bilimsel araştırma yöntemi*. Nobel Akademik Yayıncılık.
- Kesici, S. (2022). Yabancılar Türkçe öğretiminde ölçme ve değerlendirme: öğretmen algıları üzerine bir inceleme. *Uluslararası Filoloji Bengü*, 2(2), 45-69. <https://doi.org/10.29228/filolojibengu.21>
- Kılınç, M. (2011). Öğretmen adaylarının eğitimde ölçme ve değerlendirmeye yönelik öz yeterlik algı ölçeği. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 12(4), 81-93.
- Kırık, M. Y. (2008). *Yabancı dil olarak İngilizce öğretmenlerinin ölçme değerlendirme bağlamında tutum ve yaklaşımları* [Yayımlanmamış doktora tezi]. İstanbul Üniversitesi.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Kolcuoğlu, B. (2022). *Yabancı dil olarak Türkçe öğretmenlerin ölçme ve değerlendirme okuryazarlıkları üzerine bir inceleme* [Yayımlanmamış yüksek lisans tezi]. İstanbul Aydın Üniversitesi.
- Köse, D. (2008). Türkçenin yabancı dil öğretiminde sınavların hazırlanması ve uygulanması, *Dil Dergisi*, 139, 36-47.
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563–575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Maden, S., & Durukan, E. (2011). Türkçe dersi öğretmenlerinin ölçme değerlendirmeye ilişkin algıları. *Milli Eğitim*, 40(190), 212-233. <https://search.trdizin.gov.tr/en/yayin/detay/114764>
- McMillan, J. H. (2000). Fundamental assessment principles for teachers and school administrators. *Practical Assessment, Research & Evaluation*, 7(8). <http://pareonline.net/getvn.asp?v=7&n=8>.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20(1), 20-32. <https://doi.org/10.1111/j.1745-3992.2001.tb00055.x>
- MEB (2017). *Öğretmenlik mesleği genel yeterlilikleri*, Öğretmen Yetiştirme ve Geliştirme Genel Müdürlüğü.
- Mertler, C. A. (2004) Secondary teachers' assessment literacy: Does classroom experience make a difference?. *American secondary education*, 49-64.
- Mertler, C. A. & Campbell, C. (2005, April). *Measuring teachers' knowledge and application of classroom assessment concepts: Development of the assessment literacy inventory*. Paper presented at the annual meeting of the American Educational Research Association,

- Mete, F. (2012). *Türkçeyi yabancı dil olarak öğreten öğretmenlerin özel alan yeterlikleri üzerine bir araştırma* [Yayınlanmamış doktora tezi]. Gazi Üniversitesi.
- Nartgün, Z. (2008). Öğretmen adayları için ölçme ve değerlendirme genel yeterlik algısı ölçeği: geçerlik ve güvenirlik çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 8(2), 85-94. <https://dergipark.org.tr/tr/pub/aibuefd/issue/1495/18086>
- Newfields, T. (2006). Teacher development and assessment literacy. In *Authentic Communication: Proceedings of the 5th Annual JALT Pan-SIG Conference* (pp. 48–73). Tokai University College of Marine Science. <https://pansig.org/publications/pansig/2006/PDF/Newfields.pdf>
- Ogan-Bekiroglu, F., & Suzuk, E. (2014). Pre-service teachers' assessment literacy and its implementation into practice. *The Curriculum Journal*, 25(3), 344–371. <https://doi.org/10.1080/09585176.2014.899916>
- Özdemir, S. M. (2008). Sınıf öğretmeni adaylarının öğretim sürecine ilişkin öz-yeterlik inançlarının çeşitli değişkenler açısından incelenmesi. *Kuram ve Uygulamada Eğitim Yönetimi*, 54(54), 277-306.
- Plake, B. S. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Mid-Western Educational Researcher*, 6(1), 21-24. <https://scholarworks.bgsu.edu/mwer/vol6/iss1/8>
- Popham, W. J. (2006). Needed: A dose of assessment literacy. *Educational Leadership*, 63(6), 84–85.
- Rogler, D. (2014). Assessment literacy: Building a base for better teaching and learning. *English Language Teaching Forum*, 3, 2-13.
- Sallabaş, M. E., Göktentürk, T., & Yazıcı, E. (2018). Türkçe metinleri okumaya karşı yabancı dil olarak Türkçe öğrenen öğrencilerin tutumunun belirlenmesi: bir geçerlilik ve güvenirlik çalışması. *Ana Dili Eğitimi Dergisi*, 6(2), 369-385. <https://doi.org/10.16916/aded.402001>
- Sevim, O., & Varışoğlu, B. (2023). Türkçe öğrenen yabancı öğrenciler için üretici dil becerisi öz yeterlilik ölçeği: geçerlik ve güvenirlik çalışması. *Firat University Journal of Social Sciences*, 33(1), 187-202. <https://doi.org/10.18069/firatsbed.1214107>
- Stiggins, R. J. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), 238-245.
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.
- Tezbaşaran, A. A. (2008). Likert tipi ölçek hazırlama kılavuzu (e-kitap). http://www.academia.edu/1288035/Likert_Tipi_Ölçek_Hazırlama_Kılavuzu.
- Ülper, H., Yaylı, D., & Karakaya, İ. (2013). Okur özyeterlik ölçeğinin geliştirilmesi. *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 14(1), 85-100. <https://dergipark.org.tr/tr/pub/kefad/issue/59473/854634>
- Webb, N. (2002). *Assessment literacy in a standards-based urban education setting*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
- Yaşar, M. (2014). İstatistiğe yönelik tutum ölçeği: geçerlilik ve güvenirlik çalışması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 36(36), 59-75. <https://dergipark.org.tr/tr/pub/pauefd/issue/32869/374831>
- Yıldız, Ü., & Tepeli, Y. (2014). A study on teacher qualifications in teaching Turkish as a foreign language [Yabancı dil olarak Türkçe öğretiminde öğretmen yeterlilikleri üzerine bir çalışma]. *International Journal of Language Academy*, 2(4), 564–578. <https://doi.org/10.18033/ijla.182>
- Yurdabakan İ, Çüm S. (2017) Scale development in behavioral sciences (Based on exploratory factor analysis). *TJFMPC*, 11(2), 108-26. <https://doi.org/10.21763/tjfmpe.317880>

Ekler

Ek: Yabancı/İkinci Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme ve Değerlendirme Okuryazarlığı Öz Yeterlik Ölçeği (YADOT-ÖDYÖ)

Yabancı/İkinci Dil Olarak Türkçe Öğreticilerine Yönelik Ölçme ve Değerlendirme Okuryazarlığı Öz Yeterlik Ölçek Maddeleri		Kesinlikle Katılmıyorum (1)	Katılmıyorum (2)	Katıyorum (3)	Kesinlikle Katılıyorum (4)
Temel Kavramlar					
1	Ölçme ve değerlendirme kavramları arasındaki farkları ayırt edebilirim.				
2	Ölçme türlerini açıklayabilirim.				
3	Değerlendirme türlerini açıklayabilirim.				
4	Süreç odaklı değerlendirmenin anlamını açıklayabilirim.				
5	Sonuç odaklı değerlendirmenin anlamını açıklayabilirim.				
6	Ölçme ve değerlendirme araçlarında güvenilirliğin anlamını açıklayabilirim.				
7	Ölçme ve değerlendirme araçlarında geçerliğin anlamını açıklayabilirim.				
8	Ölçme ve değerlendirme araçlarında kullanışlılığın anlamını açıklayabilirim.				
Dil Becerileri (Dinleme, Konuşma, Okuma, Yazma)					
9	Sınav metinlerini öğrencinin dil seviyelerine uygun olarak seçebilirim.				
10	Sınav metinlerinde yer alan söz varlığını öğrencilerin dil düzeyine uygun seçebilirim.				
11	Ölçme ve değerlendirmeye yönelik dilsel görevleri gerçek yaşamdaki dil olaylarına benzer seçebilirim.				
12	Dil yapılarını alımlama ve üretime yönelik bir bütün olarak değerlendirebilirim.				
13	Kullandığım öğretim programları ile ölçme ve değerlendirme uygulamaları arasında tutarlılık sağlayabilirim.				
14	Hedef kitlenin ihtiyacına göre dinleme sayısını belirleyebilirim.				
15	Dinleme metnlerinin sesletim hızını dil seviyelerine uygun olarak düzenleyebilirim.				
16	Konuşma becerisini; sözlü üretim (kendini ifade etme) ve sözlü etkileşim (karşılıklı konuşma) olmak üzere iki farklı durumda ölçüp değerlendirebilirim.				
17	Konuşma becerisini değerlendirmeye yönelik dereceli puanlama anahtarı kullanabilirim.				
18	Konuşma becerisinin ölçümünü en az bir dış değerlendirici ile iş birliği içerisinde gerçekleştirebilirim.				
19	Okuma becerisi sınav metinlerini, öğrencilerin düzeyine uygun okunabilirlik özelliklerine göre seçebilirim.				
20	Yazma becerisinin değerlendirilmesinde metin türlerinden (deneme, mektup, günlük vs.) uygun olanı seçebilirim.				
21	Yazma becerisinin değerlendirilmesinde açık ve anlaşılır bir yönerge kullanabilirim.				

22	Yazma becerisini değerlendirmeye yönelik dereceli puanlama anahtarı kullanabilirim.				
23	Yazma becerisini; dilsel bileşen (yazım, noktalama, dil bilgisi, sözcük bilgisi), toplum dilbilimsel bileşen (yerindelik, incelik kuralları vb.) ile edimbilimsel bileşen (bağdaşıklık, bağıntı, metinsellik vb.) boyutlarını göz önünde bulundurarak değerlendirebilirim.				
Teknoloji Tabanlı Ölçme ve Değerlendirme					
24	Teknoloji tabanlı ölçme ve değerlendirme araçlarından amaca uygun olanı seçebilirim.				
25	Teknoloji tabanlı oyunları ölçme ve değerlendirme amaçlı kullanabilirim.				
26	Teknoloji tabanlı programlarda ölçme ve değerlendirmeye yönelik karşılaştığım sorunları çözebilirim.				
27	Öğrenme çıktılarına uygun teknoloji tabanlı ölçme araçları geliştirebilirim.				
28	Sınav süresini teknoloji tabanlı ölçme araçlarına uygun biçimde düzenleyebilirim.				
29	Geleneksel veya yenilikçi ölçme ve değerlendirme araçlarını teknolojiye uygun hale getirebilirim.				
Etik					
30	Değerlendirme sürecinin yasal ve etik olması için almam gereken sorumlulukları bilirim.				
31	Sınav güvenliği (kopya çekilmemesi vb.) için önlem alabilirim.				
32	Sınav sonuçlarının gizliliğini sağlayabilirim.				
33	Öğrencilerin din ve milliyet yapısını gözetmeden ölçüp değerlendirebilirim.				
34	Sınav evraklarını ve sonuçlarını güvenli bir alanda saklayabilirim.				
35	Sınav sürecinde tüm öğrencilere eşitlik sağlayabilirim.				
36	Sınav sonuçlarının hatalı olması durumunda gerekli düzeltmeyi yapabilirim.				

EXTENDED ABSTRACT

DEVELOPMENT OF A LITERACY SELF-EFFICACY SCALE FOR TEACHERS OF TURKISH AS A FOREIGN/SECOND LANGUAGE

In this study, the reliability and validity analyses of the ‘Measurement and Evaluation Literacy Self-Efficacy Scale for Teachers of Turkish to Foreigners’, which was developed to determine the measurement and evaluation literacy self-efficacy of teachers who teach Turkish as a foreign/second language, were conducted. In the first stage, the literature was analysed and an item pool containing 86 items was formed. In order to ensure the content validity of the scale, the necessary arrangements were made by taking the opinions of ten experts and then the items were re-submitted to thirteen experts for their opinions. In line with the feedback of the experts, the CVR value of the items was calculated using Lawshe (1975) technique and the items below the critical value (.538) were removed from the scale. Similarly, there are studies using Lawshe technique as evidence for content validity (Akçay & Şahin, 2015; Sallabaş et al., 2018; İnan Yıldız & Karadağ, 2023; Sevim & Varışoğlu, 2023). In addition, the CVI of the items was found to be (.699). This ratio showed that the content validity was high.

Exploratory Factor Analysis and Confirmatory Factor Analysis were conducted to ensure the construct validity of the scale. There are studies (Ülper et al., 2013; Yaşar, 2014; Çetinkaya & Dikici 2023) in which both analyses were preferred together to support this method. Maximum likelihood method was used in Exploratory Factor Analysis and rotation was applied with Varimax method. Bartlett's test ($p < 0.01$) and Kaiser Meier Olkin (KMO) coefficient (.893) were determined for the data set in which 217 instructors teaching Turkish as a foreign/second language participated. These values showed that the data structure was (perfectly) compatible for factor analysis.

Henson and Roberts (2006) stated that the percentage of total variance should be 52% and above. Since the total variance explanatory percentage of the scale was found to be 61.32% and the standard factor loadings were higher than 0.50, the Exploratory Factor Analysis was found to be sufficient. The standard factor loadings obtained from the CFA calculated after EFA were found to be statistically significant ($p < 0.05$). Since the model's fit index values IFI (.913), CFI (.922), SRMR (.0435), RMSEA (.0714) were within acceptable fit limits, it was concluded that the model was valid (Meydan, 2015). In addition, all sub-dimensions were found to be sub-dimensions of a single concept. As evidence of the reliability and internal consistency of the scale, Cronbach's alpha coefficient (.912) was found in the Exploratory Factor Analysis. When the sub-dimensions of the scale were analysed, Cronbach's alpha coefficient was found (.903) for Dimension 1, (.945) for Dimension 2, (.930) for Dimension 3 and (.884) for Dimension 4. These values showed that all dimensions were ‘highly reliable’ (Karagöz & Bardakçı, 2020).

The ‘Assessment and Evaluation Literacy Self-Efficacy Scale for Teachers of Turkish as a Foreign/Second Language’, whose reliability and validity were proved as a result of the analyses, has thirty-six items and a four-dimensional structure. Dimension 1: ‘Language Skills’, Dimension 2: ‘Basic Concepts’, Dimension 3: ‘Technology-Based Measurement and Evaluation’ and Dimension 4: ‘Ethics’.

In line with these data, it was concluded that the ‘Assessment and Evaluation Literacy Self-Efficacy Scale for Teachers of Turkish to Foreigners’ developed within the scope of the research is a valid and reliable tool that can determine the assessment and evaluation literacy self-efficacy of the teachers.