



# Öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği'nin Türk Kültürüne Uyarlanması

## Adaptation of the Teachers' Formative Assessment Practices Scale to Turkish Culture

Ceyhun OZAN<sup>1</sup>, Eyüphan BAHADIR<sup>2</sup>

<sup>1</sup>Atatürk University, Erzurum, Türkiye  
· ozanceyhun@atauni.edu.tr · ORCID > 0000-0002-1415-7258

<sup>2</sup>Atatürk University, Erzurum, Türkiye  
· eyuphanbahadir@gmail.com · ORCID > 0000-0002-2381-1012

### Makale Bilgisi/Article Information

**Makale Türü/Article Types:** Araştırma Makalesi/Research Article

**Geliş Tarihi/Received:** 24 Mart/March 2024

**Kabul Tarihi/Accepted:** 08 Ekim/October 2024

**Yıl/Year:** 2024 | **Cilt-Volume:** 43 | **Sayı-Issue:** 2 | **Sayfa/Pages:** 1119-1158

**Atıf/Cite as:** Ozan, C. & Bahadır, E. "Öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği'nin Türk Kültürüne Uyarlanması - Adaptation of the Teachers' Formative Assessment Practices Scale to Turkish Culture"

Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi, Ondokuz Mayıs University Journal of Faculty of Education, 43(2), December 2024: 1119-1158.

**Sorumlu Yazar/Corresponding Author:** Ceyhun OZAN

**Etik Kurul Beyanı/Ethics Committee Approv:** "Araştırma için Atatürk Üniversitesi Bilimsel Araştırma ve Yayın Etiği Kurulu'ndan 31.01.2024 tarihli ve 01/59 karar sayısı ile etik kurul izni alınmıştır - Ethics committee approval was received for the research from the Scientific Research and Publication Ethics Committee of Atatürk University with the decision number 01/59 dated 31.01.2024."

## ÖĞRETMENLERİN BİÇİMLENDİRİCİ DEĞERLENDİRME UYGULAMALARI ÖLÇEĞİ'NİN TÜRK KÜLTÜRÜNE UYARLANMASI

### ÖZ

Biçimlendirici değerlendirme öğrenme sürecinin tüm aşamalarında aktif bir şekilde kullanılarak öğrencilerin öğrenme eksikliklerinin belirlenmesi ve giderilmesine dayanır. Biçimlendirici değerlendirme sürecinde öğrencinin gelişimi not ile ifade edilmez. Son yıllarda biçimlendirici değerlendirme uygulamalarının kullanımı yaygınlaşmıştır. Bu çalışma ile biçimlendirici değerlendirme uygulama ölçeği Türk kültürüne uyarlanarak öğrenme öğretme sürecinde kullanımı sağlanacaktır. Bu çalışmanın amacı, öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği'nin Türk kültürüne uyarlanarak geçerlik ve güvenilirlik analizlerinin yapılmasıdır. Çalışmanın araştırma deseni tarama modelidir. Çalışma grubunda iki grup yer almaktadır. İlk grupta Millî Eğitim Bakanlığına bağlı farklı kademelerde görev yapan toplam 177 öğretmen bulunmaktadır. İkinci grupta ise Millî Eğitim Bakanlığına bağlı farklı kademelerde görev yapan 151 öğretmen bulunmaktadır. Çalışmada veri toplama aracı olarak Yan ve Pastore tarafından geliştirilen “Öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği” kullanılmıştır. Çalışma grubundan elde edilen verilerin analizine yönelik açıklayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) teknikleri kullanılmıştır. Çalışma sonucunda yapılan açıklayıcı faktör analizine göre ölçeğin orijinalindeki yapı ile birebir uyumlu sonuçlar elde edilmiştir. Çalışmanın doğrulayıcı faktör analizine göre ölçeğin iki faktörlü yapısı doğrulanmıştır. Çalışmanın doğrulayıcı faktör analizine göre ölçeğin iki faktörlü yapısı doğrulanmıştır. Orijinal ölçek ile uyumlu şekilde, altı maddeden oluşan birinci faktör “Öğretmen Yönelimli Biçimlendirici Değerlendirme”, dört maddeden oluşan ikinci faktör “Öğrenci Yönelimli Biçimlendirici Değerlendirme” olarak isimlendirilmiştir. Güvenirlik için elde edilen veriler ölçeğin geneli ve alt boyutları için ölçümlerin güvenilir olduğunu göstermektedir. Sonuç olarak Türk kültüründe öğretmenlerin biçimlendirici uygulamalarını incelemek amacıyla kullanılacak geçerli ve güvenilir bir ölçek elde edilmiştir.

**Anahtar Sözcükler:** Biçimlendirici Değerlendirme, Ölçek Uyarlama, Ölçme ve Değerlendirme.



## ADAPTATION OF THE TEACHERS' FORMATIVE ASSESSMENT PRACTICES SCALE TO TURKISH CULTURE

### ABSTRACT

Formative assessment is based on identifying and eliminating students' learning deficiencies by actively using them at all stages of the learning process. In the formative assessment process, student progress is not expressed through grades. In recent years, the use of formative assessment practices has become widespread. With this study, the formative assessment application scale will be adapted to Turkish culture and used in the learning and teaching process. The purpose of this study is to analyze the validity and reliability of the teachers' formative assessment practices scale by adapting it to Turkish culture. The research design of the study is the survey model. There are two groups in the working group. In the first group, there are a total of 177 teachers working at different levels under the Ministry of National Education. In the second group, there are 151 teachers working at different levels under the Ministry of National Education. "Teachers' Formative Assessment Practices Scale" developed by Yan and Pastore was used as a data collection tool in the study. Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) techniques were used to analyze the data obtained from the study group. According to the exploratory factor analysis conducted as a result of the study, results were obtained that were exactly compatible with the structure of the original scale. According to the exploratory factor analysis conducted as a result of the study, results were obtained that were exactly compatible with the structure of the original scale. According to the confirmatory factor analysis of the study, the two-factor structure of the scale was confirmed. According to the confirmatory factor analysis of the study, the two-factor structure of the scale was confirmed. Consistent with the original scale, the first factor, consisting of six items, was named "Teacher-Oriented Formative Assessment", and the second factor, consisting of four items, was named "Student-Oriented Formative Assessment". The data obtained for reliability show that the measurements are reliable for the general scale and its sub-dimensions. As a result, a valid and reliable scale that can be used to examine the formative practices of teachers in Turkish culture has been obtained.

**Keywords:** Formative Assessment, Scale Adaptation, Measurement and Evaluation.



## GİRİŞ

Ölçme herhangi bir niteliğin gözlemlenerek, gözlem sonuçlarının sayısal veya sözel olarak ifade edilmesidir (Stanford & Reeves, 2005). Değerlendirme kavramı ise ölçme sonuçlarının belli kriterler kullanılarak bir yargıya varma süreci olarak tanımlanabilir. Ölçme işleminde var olan durumu gözlemleyip tanımlama varken değerlendirme eyleminde bir yargıya, sonuca varmak amaçlanır. Bu açıdan bakıldığında ölçme ve değerlendirme kavramları birbirine karıştırılmamalıdır (Küçükahmet, 2008).

Değerlendirme işlemleri öğrenme öğretme sürecinin farklı aşamalarında kullanılabilir. Öğrenme sürecinin başında, konuların ortasında ve sonunda değerlendirme faaliyetleri sürdürülebilir (Marzano, 2000). Değerlendirme etkinlikleri öğrencilerin başarılarının belirlenmesi, öğrenme sürecinde öğrenmeye karşı öğrencilerde var olan eksikliklerin belirlenmesi ve programın etkililiğinin belirlenmesi amacıyla yapılır (Karaman, 2014). Öğrenme sürecinde kullanılan değerlendirme teknikleri tanılayıcı değerlendirme, biçimlendirici değerlendirme ve düzey belirleyici değerlendirme olarak sınıflandırılabilir.

Tanılayıcı değerlendirme öğrenme sürecinin başında yapılır. Yapılan değerlendirme ile öğrencilerin bir programa, kursa veya etkinliğe yerleştirilmesi esas alınır. Biçimlendirici değerlendirme ise öğrenme sürecinde öğrencilerin öğrenme eksikliklerinin belirlenmesi ve giderilmesi amacıyla yapılan değerlendirmedir. Eğitim programının aksayan yönleri biçimlendirici değerlendirme etkinlikleri ile belirlenebilir (Fryer vd., 2012). Düzey belirleyici değerlendirme öğrenme sürecinin sonunda yapılır. Öğrenme sürecinin sonunda yapılan düzey belirleyici değerlendirme öğrenme hedeflerine öğrencilerin ne düzeyde ulaştığını sayısal olarak belirler. Amaç öğrencilerin belirlenen hedef kazanımlara ne kadarına ulaştığının belirlenmesidir.

Biçimlendirici değerlendirme ise öğrenme için kullanılan değerlendirme tekniği olarak tanımlanabilir. Öğrenme sürecinin tüm aşamalarında kullanılarak öğrencilere öğrenmeleri ile ilgili geri bildirimler verilerek öğrenme hedeflerine ulaşmaları sağlanır (Sadler, 1989). Öğrencilerin süreç içerisinde gelişimleri desteklenerek öğrenme eksikliklerinin belirlenmesi ve giderilmesi ile süreç devam ettirilir. Öğrenme sürecinde amaç öğrenciye not vermekten ziyade bütüncül bir anlayışla gelişimi desteklemektir (Zengin vd., 2017).

Öğrenme öğretme sürecinin başından sonuna kadar tüm aşamalarda biçimlendirici değerlendirme etkinlikleri kullanılabilir (MEB, 2020). Biçimlendirici değerlendirmede iki önemli nokta bulunmaktadır. Birincisi öğrencilerin öğrenme hedeflerine ulaşabilmek için kendilerinde var olan potansiyelin farkına varmalarıdır. İkincisi ise öğrenme hedeflerine ulaşabilmek için var olan öğrenme eksikliklerini

giderebilmek için çaba sarf etmeleridir. Bu iki eylem başarılı bir şekilde gerçekleştirildiğinde biçimlendirici değerlendirme amacına ulaşır (Black & William, 1998a).

Wiliam ve Thompson (2008) çeşitli biçimlendirici değerlendirme uygulamalarının beş temel strateji halinde kategorize edildiği bir çerçeve önermiştir. Bu beş strateji şunları içermektedir: (1) Öğrenme hedeflerini ve başarı kriterlerini açıklamak ve paylaşmak, (2) Etkili sınıf tartışmaları, sorular ve öğrenme görevleri tasarlamak, (3) Öğrencileri geliştirecek geri bildirim vermek, (4) Öğrencileri birbirleri için öğretim kaynağı olarak harekete geçirmek ve (5) Öğrencileri kendi öğrenmelerinin sahibi olarak harekete geçirmek (Wiliam & Thompson, 2008, s. 64).

Öğrenme hedefleri ve başarı kriterlerinin paylaşılması, öğretim sürecini daha şeffaf ve öğrenci odaklı hâle getirmenin önemli bir unsuru olarak kabul edilmektedir. Bu yaklaşım, öğrencilere öğrenme süreçleri hakkında net bir anlayış kazandırmak, öğrenci motivasyonunu artırmak ve öğretmenlerin öğretim stratejilerini iyileştirmek amacını taşımaktadır (Black & Wiliam, 1998b; Hattie, & Timperley, 2007; Wiggins, & McTighe, 2005).

Etkili sınıf tartışmaları, sorular ve öğrenme görevlerinin tasarlanması da biçimlendirici değerlendirmenin önemli ve temel unsurlarından biridir. Etkili sınıf tartışmaları, öğrencilerin fikirlerini paylaşmalarını, çeşitli bakış açılarını anlamalarını ve eleştirel düşünme becerilerini geliştirmelerini sağlar. Tartışma ortamının yönetimi, öğrencilerin katılımını artırmak ve saygılı bir atmosfer oluşturmak önemlidir. Öğrencilere açık uçlu sorular sorarak onların düşünce süreçlerini ortaya çıkarmak, tartışmaların kalitesini artırabilir (Brookfield & Preskill, 2005). Etkili sorular, öğrencilere düşünme süreçlerini açığa çıkarmaları ve derinlemesine anlamaları için fırsat tanımalıdır. Sorular, bilgi düzeyinden öteye geçerek eleştirel düşünme, problem çözme ve analitik becerileri hedef almalıdır. Bloom'un taksonomisine dayalı olarak özellikle analiz, değerlendirme ve yaratma düzeylerine hitap eden çeşitli sorular tasarlamak önemlidir (McTighe & Wiggins, 2013; Walsh & Sattes, 2016). Öğrenme görevleri, öğrencilerin bilgiyi pratiğe dökme yeteneklerini geliştirmelerini sağlar. Öğrencilere somut, ilgi çekici ve gerçek dünya ile bağlantılı görevler vermek, öğrenmenin anlamını artırabilir. Proje tabanlı öğrenme, grup çalışmaları, saha çalışmaları gibi çeşitli öğrenme görevleri, öğrencilerin etkileşimde bulunmalarını ve derinlemesine öğrenmelerini destekler (Conrad & Donaldson, 2016).

Öğrencilere nitelikli ve etkili geri bildirim vermek, biçimlendirici değerlendirmenin önemli bir unsuru olup, öğrenmeyi iyileştirmek ve öğrencilerin gelişimine katkıda bulunmak için kritik bir rol oynamaktadır. Geri bildirim öğrenciyi geliştirecek nitelikte olması için şu noktalara dikkat edilmesi gerekir (Black & Wiliam, 1998b; Hattie & Timperley, 2007; Higgins vd., 2002; Kluger & DeNisi, 1996): Geri bildirim, öğrencinin ihtiyaçlarına ve hedeflerine odaklanmalıdır. Geri bildirim, öğrencinin belirlenen öğrenme hedefleri ve standartlar doğrultusunda nasıl

performans gösterdiğini değerlendirmelidir. Biçimlendirici değerlendirme süreci içinde sürekli geri bildirim sağlamak, öğrencilere hızlı bir şekilde tepki verme ve eksiklerini düzeltme şansı tanır. Geri bildirim, öğrencinin güçlü yönlerini vurgulayarak ve olumlu dil kullanarak yapıcı olmalıdır. Geri bildirim, öğrencileri öğrenmeye aktif olarak katılmaya teşvik etmelidir. Düzenli ilerleme raporları ve öğrenci-öğretmen iletişimi, öğrencilere sürekli geri bildirim sağlamak adına önemlidir.

Biçimlendirici değerlendirmenin bir diğer ögesi, öğrencilerin birbirlerini öğretim kaynağı olarak kullanmalarıdır; bu da öğrenme süreçlerini zenginleştirir ve öğrenciler arasında iş birliği ile etkileşimi teşvik edebilir. Bu strateji, öğrencilerin sadece öğrenci-öğretmen etkileşiminden değil, aynı zamanda öğrenci-öğrenci etkileşiminden de öğrenmelerini sağlar. Akran değerlendirme de bu strateji kapsamında yer almaktadır (Slavin, 1996; Topping, 1996).

Biçimlendirici değerlendirmenin son ögesi öğrencileri kendi öğrenmelerinin sahibi olarak harekete geçirmek, öğrencilerin öğrenme süreçlerinde daha aktif ve bilinçli bir rol oynamalarını sağlamayı amaçlar. Bu strateji, öğrencilerin öğrenmeyi sadece bilgi almakla değil, aynı zamanda bilgi oluşturmak, bağlamak ve anlamakla ilişkilendirmelerini hedefler. Bu strateji öz değerlendirme uygulamalarını da kapsamaktadır (Deci vd., 1991; Pintrich, 2003; Zimmerman, 2002).

Eğitim sisteminin etkin ve sürdürülebilir bir şekilde işleyebilmesi, öğretmenlerin öğrencileri değerlendirme ve rehberlik etme yeteneklerine bağlıdır. Bu bağlamda, öğretmenlerin biçimlendirici değerlendirme uygulamaları, öğrenci başarısını artırmak, öğrenci gelişimine katkıda bulunmak ve eğitim süreçlerini iyileştirmek amacıyla önemli bir araç olarak öne çıkmaktadır (Akbaş & Ozan, 2024; Kıncal & Ozan; Ozan & Kıncal). Yan ve Pastore (2022) de biçimlendirici değerlendirmenin son yirmi yılda eğitim araştırmalarında popüler bir konu olmasına rağmen, yazarların bilgisi dahilinde, yaygın olarak tanınan ve uygulanan bir anket aracının olmadığına işaret etmektedir. Türkiye’de de biçimlendirici değerlendirme ile ilgili ölçme araçlarının (Karaman, 2017; Karaman & Şahin, 2017) oldukça sınırlı olduğu söylenebilir. Karaman (2017) tarafından yapılan çalışmada Yan ve Cheng (2015) tarafından öğretmenler için geliştirilmiş olan biçimlendirici değerlendirmeye ilişkin tutum, niyet ve uygulamalar ölçeğinin öğretmen adaylarına yönelik olarak uyarlaması yapılmıştır. Karaman ve Şahin (2017) ise Planlı Davranış Teorisi’ne göre öğretmenlerin biçimlendirici değerlendirme niyet ve uygulamalarını incelemeye yönelik geliştirilen bir ölçeğin Türkçe uyarlamasını yapmışlardır. Ölçekte, duyuşsal tutum, tutum, öznel norm, kontrol edilebilirlik, öz yeterlik, niyet ve davranış boyutları bulunmaktadır.

Ancak, bu araştırmada uyarlaması yapılacak ölçek, Wiliam ve Thompson (2008) tarafından geliştirilen çerçeveye dayanmaktadır. Bu ölçek, biçimlendirici değerlendirme uygulamalarını daha spesifik olarak beş temel strateji bağlamın-

da değerlendirmektedir. Diğer ölçeklerden farkı, biçimlendirici değerlendirmenin daha odaklanmış ve uygulamaya dönük bir biçimde ele alınmasıdır. Bu sayede, öğretmenlerin değerlendirme süreçlerindeki stratejik uygulamaları daha net bir şekilde ölçülmektedir.

Öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği'nin Türk kültürüne uyarlamasıyla yerel normlara ve değerlere uygun bir şekilde öğretmen biçimlendirici değerlendirme stratejileri değerlendirilebilecek ve bu stratejilerin Türk eğitim sistemi içindeki rolü belirlenebilecektir. Öğretmenlerin biçimlendirici değerlendirme stratejilerini etkili bir şekilde uygulamaları, öğrenci başarısını artırmada ve sınıf içi etkileşimleri geliştirmede kritik bir rol oynamaktadır. Uyarlanacak ölçekle birlikte öğretmenlerin biçimlendirici değerlendirme uygulamalarının değerlendirilip geliştirilmesi açısından katkı sağlayabilir. Biçimlendirici değerlendirme stratejilerinin uygulanması konusunda eksiklikler belirlenerek eğitim politikalarının bu alanda daha etkili ve yönlendirici olmasına katkı sağlayabilir. Ayrıca, ölçeğin Türk kültürüne uyumlu bir şekilde geliştirilmesi, uluslararası düzeyde gerçekleştirilecek eğitim araştırmalarında Türkiye'nin verilerinin daha anlamlı ve karşılaştırılabilir olmasına olanak tanıyacaktır.

Bu gerekçeler doğrultusunda araştırmanın amacı, Yan ve Pastore (2022) tarafından geliştirilen "Öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği'nin dilsel eşdeğerlik, geçerlik ve güvenilirlik analizlerinin yapılarak ölçeğin Türk kültürüne uyarlanmasıdır.

## YÖNTEM

### Araştırma Modeli

Araştırma tarama modelindedir. Tarama araştırmaları bir olayın özelliklerinin veya bir grubun görüşlerinin belirlenmeye çalışıldığı çalışmalardır (Büyüköztürk vd., 2020). Bir başka tanımlamaya göre tarama modeli geçmişte var olan veya günümüzdeki bir olay, durum, olgu veya kişileri kendi doğal ortamlarında hiçbir müdahale olmaksızın betimlemeye çalışan bir araştırma modelidir (Karasar, 2000). Ölçek uyarlaması olan bu çalışmada da ölçülmek istenen özelliklerle ilgili var olan durum betimlenmeye çalışılacağı için tarama modeli tercih edilmiştir.

### Çalışma Grubu

Araştırmada iki çalışma grubu bulunmaktadır. Birinci çalışma grubundan elde edilen verilerle açıklayıcı faktör analizi, ikinci çalışma grubundan elde edilen verilerle de doğrulayıcı faktör analizi yapılmıştır. Alan yazınında farklı analizlerin birbiriyle bağımsız çalışılmasının daha güvenilir sonuçlar vereceği belirtilmektedir (Brown, 2015; Tabachnick & Fidell, 2019). Bu yöntem, ölçüm aracının yapısını ilk

grupta açığa çıkardıktan sonra, ikinci grupta bu yapının doğrulanmasını sağlar. İki ayrı grup kullanmak, aynı örneklem üzerinde hem AFA hem de DFA yaparak model uyumunu aşırı uyarlama (overfitting) riskini azaltır (Worthington & Whittaker, 2006). Bu nedenle araştırmada iki farklı grup tercih edilmiştir. Birinci çalışma grubunu 2023-2024 eğitim öğretim yılı güz döneminde Erzurum ili bünyesinde farklı kademe ve düzeylerde görev yapan 177 öğretmen oluşturmaktadır. DeVellis (2016) ile Worthington ve Whittaker (2006) ölçek geliştirme çalışmalarında örneklem büyüklüğünün madde sayısının 5 ila 10 katı olmasının uygun olduğunu belirtirken, Nunnally ve Bernstein (1994) de madde başına en az 10 katı örneklem önerisinde bulunmaktadır. Bu çalışmada uyarlaması yapılan ölçeğin 10 maddesi bulunduğu için her iki çalışma grubunda yer alan örneklem sayısı yeterli kabul edilmiştir. Tablo 1'de birinci çalışma grubuna ait demografik bilgiler verilmiştir.

**Tablo 1.** Birinci Çalışma Grubuna Ait Demografik Bilgiler

Değişkenler	Kategoriler	n	%
Cinsiyet	Kadın	86	49
	Erkek	91	51
Görev Yaptığınız Kademe	İlkokul	33	19
	Ortaokul	85	48
	Lise	59	33
Görev Yılı	1-5	65	37
	6-10	53	30
	11 ve üzeri	59	33

Birinci çalışma grubundaki öğretmenlerin 86'sı kadın, 91'i erkek; 33'ü ilkokul, 85'i ortaokul ve 59'u lise kademesinde görev yapmaktadır. Öğretmenlerin 65'i 1-5 yıl arası, 53'ü 6-10 yıl arası, 59'u 11 ve üzeri yıl görev yapmışlardır. Tablo 2'de ikinci çalışma grubuna ait demografik bilgiler verilmiştir.

**Tablo 2.** İkinci Çalışma Grubuna Ait Demografik Bilgiler

Değişkenler	Kategoriler	n	%
Cinsiyet	Kadın	122	81
	Erkek	29	19
Görev yaptığınız kademe	İlkokul	38	25
	Ortaokul	74	49
	Lise	39	26
Görev yılı	1-5	69	46
	6-10	45	30
	11 ve üzeri	37	24



İkinci çalışma grubundaki öğretmenlerin 122'si kadın, 29'u erkek; 38'i ilkokul, 74'ü ortaokul ve 39'u lise kademesinde görev yapmaktadır. Öğretmenlerin 69'u 1-5 yıl arası, 45'i 6-10 yıl arası, 37'si 11 ve üzeri yıl görev yapmışlardır.

### Veri Toplama Araçları

Çalışmada veri toplama aracı olarak Yan ve Pastore (2022) tarafından geliştirilen Öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği kullanılmıştır. Ölçeğin geliştirilme süreci şu şekilde gerçekleştirilmiştir:

1. Alan Yazınının İncelenmesi: Ölçeğin geliştirilmesine başlanmadan önce ilgili alan yazını detaylı bir şekilde incelenmiştir.
2. Odak Grup Görüşmeleri: Hong Kong'da görev yapan ilkokul ve ortaokul öğretmenleri ile 15 maddelik bir ölçek üzerinde odak grup görüşmeleri yapılmıştır. Bu süreçte, öğretmenlerin görüşleri alınmıştır.
3. Uzman Görüşleri: Odak grup görüşmeleri ve uzmanların değerlendirmeleri sonucunda ölçekten 5 madde çıkarılmıştır. Geriye kalan maddeler üzerinde düzenlemeler yapılmıştır.
4. Likert Tipi Ölçek: Ölçekte yer alan maddeler, "Hiçbir Zaman", "Neredeyse Hiç", "Nadiren", "Bazen", "Sıklıkla" ve "Her Zaman" arasında değişen altı dereceli Likert tipi ölçek ile puanlanmıştır.

Örneklem gruplarının katılımı ve yapılan analizler şu şekildedir:

1. Örneklem Grupları: Ölçeğin uygulanacağı iki örneklem grubu seçilmiştir: Hong Kong'dan 12 okulda görev yapan toplam 449 öğretmen ve İtalya'dan 10 okulda görev yapan 309 öğretmen çalışmaya katılmıştır.
2. Açıklayıcı Faktör Analizi: Yapılan açıklayıcı faktör analizi sonucunda ölçeğin iki faktörlü bir yapıya sahip olduğu belirlenmiştir:
  - Öğretmen Yönelimli Biçimlendirici Değerlendirme (6 madde)
  - Öğrenci Yönelimli Biçimlendirici Değerlendirme (4 madde)

Her bir faktör, Wiliam ve Thompson (2008) çerçevesindeki biçimlendirici değerlendirme kategorilerine göre iki maddeden oluşmuştur. İki faktör arasındaki korelasyon 0,52 olarak hesaplanmıştır.

3. Doğrulayıcı Faktör Analizi: Doğrulayıcı faktör analizi sonuçları ise şu şekildedir: RMSEA = 0,056, SRMR = 0,062, CFI = 0,94. Ayrıca, Madde 1 ile Madde 2 ve Madde 9 ile Madde 10 arasında modifikasyon yapılmıştır.

4. Rasch Analizi: Rasch analizine göre, tüm maddeler Rasch modeline iyi uyum sağlamıştır.

Güvenirlilik analizi sonuçları ise şu şekildedir:

1. Hong Kong Grubu: Ölçeğin Hong Kong örnekleminde birinci faktör için Cronbach alfa katsayısı 0,77, ikinci faktör için ise 0,87 olarak hesaplanmıştır.
2. İtalya Grubu: İtalya örnekleminde birinci faktör için Cronbach alfa katsayısı 0,60, ikinci faktör için ise 0,66 olarak belirlenmiştir.
3. Birleşik Güvenirlilik: İki grubun birleştirilen verileri üzerinde yapılan analizde, birinci faktör için Cronbach alfa katsayısı 0,70, ikinci faktör için ise 0,75 olarak hesaplanmıştır.

### Dilsel Eşdeğerlik

Çalışmada dilsel eşdeğerliği sağlamak amacıyla, çeviri tekrar çeviri yöntemi kullanılmıştır. Ölçek maddeleri, bu süreçte üç farklı İngilizce öğretmeni tarafından bağımsız olarak Türkçeye çevrilmiştir. Çevirisi tamamlanan ölçek maddelerinin Türkçe karşılıkları, üç ayrı Türkçe öğretmeni tarafından incelenmiş ve daha sonra İngilizce öğretmenliği bölümünde görev yapan bir öğretim üyesi tarafından tekrar İngilizceye çevrilmiştir. Türkçeden İngilizceye çevrilen maddeler ile ölçeğin orijinal formundaki maddeler karşılaştırılmış ve maddelerin örtüşmesi göz önüne alınmıştır. Ölçek maddeleri, bu değerlendirmeler sonucunda gözden geçirilerek çeviri-tekrar çeviri süreci tamamlanmış ve ölçeğin Türkçe formu oluşturulmuştur.

Dilsel eşdeğerliği sağlamak amacıyla uygulanan çeviri-tekrar çeviri yöntemi tamamlandıktan sonra, İngilizce ve Türkçe ölçekler aynı örneklem grubuna uygulanmış ve iki formdan elde edilen puanlar arasında korelasyon analizi yapılmıştır. Bu aşamada, dilsel eşdeğerliğin sağlanıp sağlanmadığını test etmek için Pearson korelasyon katsayısı hesaplanmıştır. Korelasyon analizine göre, İngilizce ve Türkçe formdan elde edilen puanlar arasında yüksek düzeyde pozitif bir ilişki bulunmuştur ( $r = 0.84$ ,  $p < 0.001$ ). Bu sonuç, ölçeğin iki dilde de tutarlı olduğunu ve dilsel eşdeğerliğin sağlandığını göstermektedir. Bu bulgu, ölçeğin Türkçeye yapılan uyarlamasının orijinal İngilizce formuyla büyük ölçüde örtüştüğünü ve ölçeğin farklı dil versiyonları arasında anlamlı bir farklılık olmadığını ortaya koymaktadır.

### Verilerin Toplanması

Verilerin toplanmasına başlanmadan önce uygulamanın yapılabilmesi için gerekli yasal izinler alınmıştır. Öncelikli olarak ölçeği geliştiren iki yazardan biri olan Zi Yan'dan e-posta yoluyla uyarılama çalışması için izin alınmıştır. Ardından Eğitim

Bilimleri Birim Etik Kurulundan ve İl Millî Eğitim Müdürlüğünden uygulama izni alınmıştır. Veriler, Google Forms üzerinden anket oluşturularak toplanmıştır.

### Verilerin Analizi

Çalışma grubundan elde edilen verilerin analizine yönelik açılımlayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) teknikleri kullanılmıştır. Araştırmada Öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği'nin orijinal formunun Türk kültüründeki yapısını belirlemek amacıyla AFA kullanılmıştır. AFA, değişkenler arasındaki ilişkileri belirleme işlevini yerine getirirken doğrulayıcı faktör analizi değişkenler arasındaki belirlenmiş ilişkilerin beklenen düzeyde olup olmadığı ile ilgilenmektedir. Değişkenlerin hangi değişkenlerle ilişkili olduğu, faktörlerin birbirlerinden bağımsız olup olmadığı ve faktörlerin ilişkileri açıklama noktasında yeterli olup olmadığını sınamaktadır. AFA ile elde edilen verilerin yapı geçerliği doğrulayıcı faktör analizi ile test edilir (Güvendir & Özkan, 2015). Madde faktör yük değerinin genellikle 0.45 ve üzeri olması tercih edilir. Ancak az sayıda maddeden oluşan ölçeklerde faktör yük değerleri 0.30'a kadar kabul edilebilmektedir (Hair vd., 2010; Kline, 1994). Bu çalışmada faktör yüklerinin yeterli kabul edilmesi için belirlenen ölçüt .30'dur.

Analize başlamadan önce, tüm maddeler arasındaki korelasyon matrisi incelenerek, önemli düzeyde anlamlı korelasyonların olup olmadığı değerlendirilmiş ve faktör analizi için uygunluk gösteren anlamlı ilişkiler tespit edilmiştir. Daha sonra, örneklem uygunluğunu değerlendirmek için Kaiser-Meyer-Olkin (KMO) ve Bartlett Küresellik testleri yapılmıştır. Verilerin faktör analizine uygun olması için KMO değerinin .60'dan yüksek olması ve Bartlett testinin anlamlı çıkması gerekmektedir (Kaiser, 1974; Pallant, 2005).

### Etik Kurul İzin Bilgileri

Yapılan bu çalışmada "Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Etiği Yönergesi" kapsamında uyulması belirtilen tüm kurallara uyulmuştur.

Etik Değerlendirmeyi Yapan Kurul Adı: Atatürk Üniversitesi Eğitim Bilimleri Birim Etik Kurulu

Etik Değerlendirme Kararının Tarihi: 31.01.2024

Etik Değerlendirme Belgesi Sayı Numarası: 01/59

## BULGULAR

Ölçeğin Türk kültürüne uyarlama çabaları kapsamında geçerlik ve güvenilirlik çalışmaları gerçekleştirilmiştir.

### Yapı Geçerliği

**Açımlayıcı Faktör Analizi (AFA).** AFA sonuçlarına göre, KMO (Kaiser-Meyer-Olkin) Örneklem Uygunluk Katsayısı .84 olarak hesaplanmıştır. Aynı zamanda Bartlett Küresellik Testi'nin  $\chi^2$  değeri 614.962 ( $p < .01$ ) olarak belirlenmiştir. Bu değerler, verilerin faktör analizi için uygun olduğunu göstermektedir.

Analiz sonuçlarına göre, iki faktörün ortaya çıkmasıyla açıklanan toplam varyans oranı %53.82'dir. Ayrıca, tüm maddelerin faktör yükleri .30'dan yüksek olduğu için ölçekten herhangi bir madde çıkarılmamıştır. Faktör analizi sırasında, verilerin daha iyi bir şekilde yorumlanabilmesi amacıyla Varimax döndürme yöntemi kullanılmıştır. Bu döndürme yöntemi, faktörlerin birbirinden bağımsız olmasını sağlayarak, her bir faktörün daha açık ve net bir şekilde ayrışmasını amaçlamaktadır (Tabachnick & Fidell, 2019). Orijinal ölçek ile karşılaştırıldığında, kullanılan döndürme yönteminde herhangi bir farklılık bulunmamaktadır. Türkçe formun faktör yapısıyla ilgili değerler Tablo 4'te sunulmuştur.

**Tablo 3.** Maddeler ve faktör yükleri

Madde No	Faktör Yükleri	
	1	2
1	.501	
2	.794	
3	.690	
4	.688	
5	.772	
6	.769	
7		.858
8		.834
9		.556
10		.595
Açıklanan Varyans	%40.933	%12.885
Toplam=	%53.818	
OAV = 1. Faktör = 0.50; 2. Faktör = 0.52		

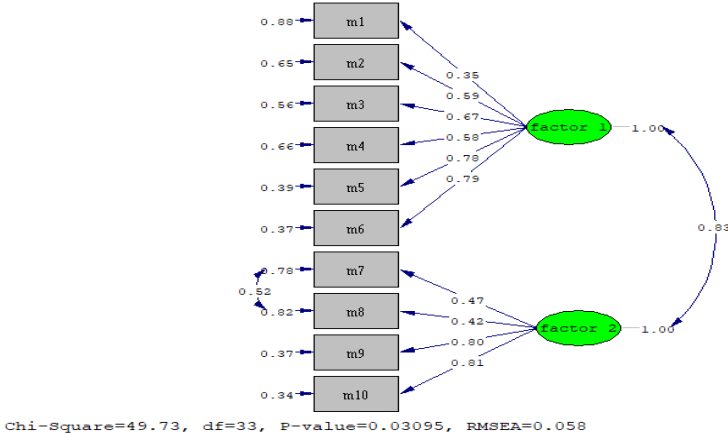
Tablo 3'te görüldüğü gibi Ölçeğin Türkçe formu iki boyuttan oluşmaktadır. Ölçeğin açıklanan toplam varyansı %53.82'dir. Maddelerin faktör yükleri .50 ile .86 arasında değişmektedir. Ölçeğin ortalama açıklanan varyans (OAV) değeri 1. Faktör için .050; 2. Faktör için .052 olarak hesaplanmıştır. OAV değerleri yakınsama ve ayırt edicilik geçerliğinin sağlandığını göstermektedir (Bagozzi & Yi, 1988; Nunnally & Bernstein, 1994).

Birinci düzey DFA sonuçlarına göre modelin uyum indekslerine ait değerler Tablo 4'te sunulmuştur.

**Tablo 4.** Birinci Düzey DFA Uyum İyiliği İndekslerine Yönelik Değerler

Ki-kare	sd	p	Ki-kare/sd	RMSEA	SRMR	GFI	AGFI	NFI	NNFI	CFI
49.73	33	p=0.03	1.51	0.058	0.042	0.94	0.90	0.95	0.98	0.98

Birinci düzey DFA sonuçlarına göre ilk etapta RMSEA değeri ve ki-kare/sd değerinin istenilen düzeyde olmadığı belirlenmiştir. Bu nedenle aynı faktör içerisinde yer alan M7 ile M8 arasındaki korelasyonlar serbest bırakılarak modifikasyon yapılmıştır. Modifikasyon sonrasında ki-kare değerinin serbestlik derecesine oranı iyi olarak değerlendirilebilecek şekilde 2'den küçük olarak hesaplanmıştır. Diğer uyum indekslerine yönelik değerlendirmeler yapıldığında, tüm indekslerin iyi veya kabul edilebilir sınırlar içinde olduğu görülmüştür (Schermelleh-Engel vd., 2003). Bu sonuçlara göre, model-veri uyumunun sağlandığı ve ölçeğin yapı geçerliğinin doğrulandığı belirlenmiştir. Doğrulayıcı faktör analizinin path diyagramı Şekil 1'de gösterilmiştir.



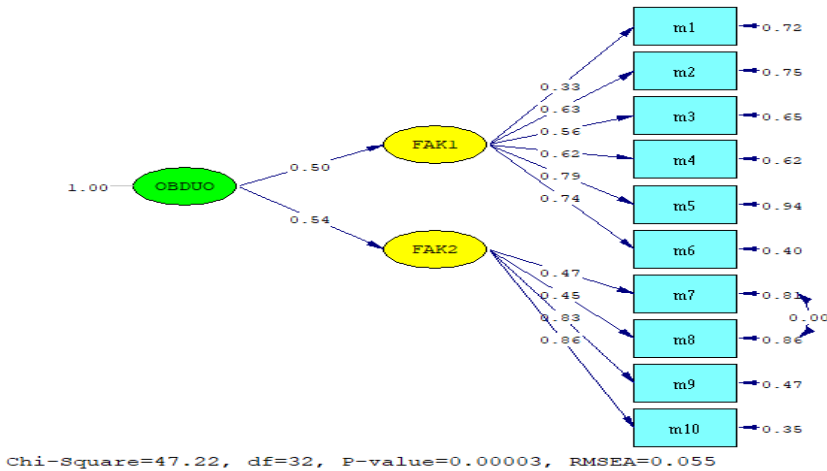
**Şekil 1.** Birinci Düzey DFA İçin Path Diyagramı

Birinci düzey DFA sonucunda ortaya çıkan iki faktörün uyumu ve bu faktörlerin tek bir üst faktörü temsil edip etmediğini belirlemek (Byrne, 2010; Kline, 2016) amacıyla ikinci düzey DFA yapılmıştır. İkinci düzey DFA sonuçlarına göre uyum iyiliği indeksleri Tablo 5'te gösterilmiştir.

**Tablo 5.** İkinci Düzey DFA Uyum İyiliği İndekslerine Yönelik Değerler

Ki-kare	sd	p	Ki-kare/sd	RMSEA	SRMR	GFI	AGFI	NFI	NNFI	CFI
47.22	32	p=0.03	1.48	0.055	0.040	0.92	0.91	0.96	0.97	0.97

İkinci düzey DFA sonuçlarına göre ilk etapta RMSEA değeri ve ki-kare/sd değerinin istenilen düzeyde olmadığı belirlenmiştir. Bu nedenle birinci düzey DFA'da olduğu gibi aynı faktör içerisinde yer alan M7 ile M8 arasındaki korelasyonlar serbest bırakılarak modifikasyon yapılmıştır. Modifikasyon sonrasında ikinci düzey DFA uyum iyiliği indekslerine yönelik sonuçlar, modelin iyi bir uyum gösterdiğini ortaya koymaktadır. Ki-kare/sd oranı 1.48 olup, 2'nin altında kaldığı için iyi bir uyuma işaret etmektedir (Kline, 2016). RMSEA değeri 0.055 ve SRMR değeri 0.040 olup, her iki değer de mükemmel uyum sınırları içindedir (Hu & Bentler, 1999). GFI ve AGFI değerleri sırasıyla 0.92 ve 0.91'dir; bu sonuçlar, modelin kabul edilebilir bir uyuma sahip olduğunu göstermektedir (Jöreskog & Sörbom, 1996). NFI, NNFI ve CFI değerleri sırasıyla 0.96, 0.97 ve 0.97 olup, bu sonuçlar güçlü ve mükemmel bir uyumun göstergesidir. İkinci düzey DFA'ya ilişkin path diyagramı Şekil 2'de gösterilmiştir.



**Şekil 2.** İkinci Düzey DFA İçin Path Diyagramı

Şekil 2'ye göre ÖBDUÖ'nün Faktör 1 ve Faktör 2 ile olan ilişkileri sırasıyla 0.50 ve 0.54 olarak verilmiştir. Bu değerler, ÖBDUÖ'nün her iki faktörü de anlamlı bir şekilde temsil ettiğini göstermektedir.

### Madde Analizi

Her bir ölçek maddesinin, öğretmenleri biçimlendirici değerlendirme uygulamaları açısından ne kadar iyi ayırt edebildiğini belirlemek için ilk olarak madde toplam korelasyonları hesaplanmıştır. Ardından, toplam puana göre belirlenen %27'lik üst ve %27'lik alt gruplar arasındaki madde puanlarındaki farkın anlamlılığını değerlendirmek için t-testi kullanılmıştır. Elde edilen sonuçlar Tablo 6'da sunulmuştur.

**Tablo 6.** Madde Toplam Korelasyonları ve Madde Ayırt Ediciliği İçin T Testi

Madde No	Düzeltilmiş Madde-Toplam Korelasyonları	t
1	.482	2.497*
2	.441	2.431*
3	.561	2.108*
4	.538	2.600*
5	.461	2.831*
6	.568	2.640*
7	.533	2.588*
8	.490	2.469*
9	.597	2.412*
10	.614	2.192*

\*p < .001

Ölçeğin düzeltilmiş madde-toplam korelasyonlarının .44 ile .61 arasında değiştiği belirlenmiştir. Ayrıca, toplam puanlara göre belirlenmiş %27'lik alt ve üst gruplar arasındaki madde puanlarındaki farklarla ilgili t değerleri 2.108 (p < .001) ile 2.831 (p < .001) arasında çeşitlilik göstermektedir (Nunnaly & Bernstein, 1994). Yapılan t-testi sonuçları, tüm maddeler için üst %27'lik grubun madde toplam puanının, alt %27'lik grubun aynı puanından anlamlı bir şekilde (p < 0.001) yüksek olduğunu ortaya koymuştur.

Öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği'nin Türkçe formunda iki faktörde toplanan 10 madde bulunmaktadır. Birinci faktör 6 maddeden oluşmakta olup "Öğretmen Yönelimli Biçimlendirici Değerlendirme

Uygulamaları”, ikinci faktör 4 maddeden oluşmakta olup “Öğrenci Yönelimli Biçimlendirici Değerlendirme Uygulamaları” şeklinde isimlendirilmiştir. Olumsuz ifadeye sahip herhangi bir madde bulunmamaktadır. Ölçek, orijinal ölçeğe uygun olarak tüm maddeler “Hiçbir zaman”, “Neredeyse hiç”, “Nadiren”, “Bazen”, “Sıklıkla” ve “Her zaman” arasında değişen şekilde altılı Likert tipinde derecelendirilmiştir. Ölçekten alınabilecek en yüksek puan 60, en düşük puan ise 10’dur. Ölçekten toplam puan elde edilebilmektedir. Ölçekten alınan toplam puan öğretmenlerin biçimlendirici değerlendirme uygulamalarına yer verme düzeylerini göstermektedir. Ölçeğin ilk boyutundan alınan puanlar öğretmenlerin öğretmen yönelimli biçimlendirici değerlendirme uygulamalarının düzeyini, ikinci boyutundan alınan puanlar ise öğretmenlerin öğrenci yönelimli biçimlendirici değerlendirme uygulamalarının düzeyini göstermektedir.

### Güvenirlilik

Ölçeğin güvenilirliğini belirlemek amacıyla uygulanan Cronbach alfa katsayısı ile eş yarılar yöntemine ilişkin veriler Tablo 6’da yer almaktadır.

**Tablo 6.** Ölçeğin Güvenirliğine İlişkin Veriler

Faktör	Madde Sayısı	Cronbach Alfa Katsayısı	Guttman Eş Yarılar Katsayısı
Genel	10	.84	.73
Öğretmen Yönelimli	6	.80	
Öğrenci Yönelimli	4	.72	

Ölçeğin geneli için Cronbach alfa katsayısı .84; Guttman eş yarılar katsayısı .73 olarak hesaplanmıştır. Öğretmen Yönelimli Biçimlendirici Değerlendirme alt boyutu için .80; Öğrenci Yönelimli Biçimlendirici Değerlendirme alt boyutu için .72 olarak hesaplanmıştır. Buna göre ölçekten elde edilen ölçümlerin güvenilir olduğu sonucuna ulaşılmıştır (Cronbach, 1951).

## TARTIŞMA, SONUÇ VE ÖNERİLER

Araştırmada Öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği’ni Türk kültürüne uyarlamak için dilsel eşdeğerlik, geçerlik ve güvenilirlik çalışmaları yapılmıştır. Ölçeğin orijinalinde iki alt boyuttan oluşan 10 madde bulunmaktadır. Bu araştırmada yapılan açımlayıcı faktör analizine göre ölçeğin orijinalindeki yapı ile birebir uyumlu sonuçlar elde edilmiştir. Yapılan doğrulayıcı faktör analizine göre de ölçeğin iki faktörlü yapısı doğrulanmıştır. Orijinal ölçek ile uyumlu şekilde, altı maddeden oluşan birinci faktör “Öğretmen Yönelimli Biçimlendirici Değerlendirme”, dört maddeden oluşan ikinci faktör “Öğrenci Yönelimli Biçimlendirici Değerlendirme” olarak isimlendirilmiştir. Ölçeğin birinci



boyutu genellikle öğretmenler tarafından başlatılan ve uygulanan biçimlendirici değerlendirme stratejileri (öğrenme hedefleri ve başarı kriterlerini açıklamak ve paylaşmak, etkili sınıf tartışmaları, sorular ve öğrenme görevleri tasarlamak ve öğrencileri geliştirecek geri bildirim vermek) ile ilgiliyken ikinci boyut, biçimlendirici değerlendirmenin temel unsurlarından olan, öğrenciyi merkeze alan öz ve akran değerlendirme stratejileri ile ilgilidir. Ölçekten toplam puan elde edilebilir. Puanların yükselmesi öğretmenlerin öğretim sürecinde biçimlendirici değerlendirme uygulamalarına daha çok yer verdikleri anlamına gelmektedir. Ayrıca, ölçekten elde edilen puanlar, alt boyutlar kapsamında da değerlendirilebilir. Alt boyutlardan alınan puanlara göre öğretmenlerin biçimlendirici değerlendirme uygulamalarının öğretmen yönelimli ya da öğrenci yönelimli olma düzeyi değerlendirilebilir.

Ölçeğin güvenilirlik analizleri de ölçekten elde edilen verilerin güvenilir olduğunu göstermektedir. Sonuç olarak Öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği'nin Türk kültüründe öğretmenlerin biçimlendirici değerlendirme uygulamalarının ölçülmesinde geçerli ve güvenilir bir ölçek olduğu belirlenmiştir.

Bu çalışma, biçimlendirici değerlendirme ile ilgili mevcut ölçeklerden bazı yönlerden farklılık göstermektedir. Örneğin, Karaman (2017) tarafından yapılan çalışmada, Yan ve Cheng (2015) tarafından geliştirilen Biçimlendirici Değerlendirme Tutum, Niyet ve Uygulama Ölçeği öğretmen adaylarına uyarlanmıştır. Karaman ve Şahin (2017) ise Planlı Davranış Teorisi'ne dayanan bir biçimlendirici değerlendirme ölçeğini Türkçeye uyarlamışlardır. Bu ölçeklerde biçimlendirici değerlendirme tutumları, niyetler ve kontrol edilebilirlik gibi çeşitli boyutlar ele alınmıştır. Bu araştırmanın ölçeği ise Wiliam ve Thompson (2008) çerçevesine dayanarak geliştirilmiş ve biçimlendirici değerlendirme uygulamalarını beş temel strateji kapsamında değerlendirmektedir. Karaman (2017) ve Karaman ve Şahin'in (2017) çalışmalarında tutum, niyet ve uygulamalara yönelik boyutlar bulunurken, bu çalışmada uyarlanan ölçek, doğrudan öğretmenlerin biçimlendirici değerlendirme stratejilerini ne ölçüde uyguladıklarına odaklanmaktadır. Karaman ve Şahin'in (2017) ölçeği, geniş kapsamlı bir yaklaşımla duyuşsal tutum ve davranışları incelemektedirken bu araştırmanın ölçeği daha çok uygulamaya yönelik bir ölçüm aracı sunmaktadır. Ayrıca, bu ölçek öğrenci merkezli stratejilere daha fazla vurgu yaparak, biçimlendirici değerlendirmenin hem öğretmen hem de öğrenci odaklı boyutlarını kapsamlı bir şekilde ele almaktadır.

Bugüne kadar öğretmenlerin biçimlendirici değerlendirmeyi etkili bir şekilde nasıl uyguladıkları, eğitim araştırmalarında ve eğitim reformlarında önemli bir gündem olmuştur. Ancak biçimlendirici değerlendirme uygulamalarını değerlendirmek için geçerli ve kullanımı kolay bir aracın bulunmaması, uzun süredir öğretmen geliştirme programlarının yanı sıra öğretmen değerlendirme uygulamalarını da olumsuz etkilemiştir (Yan & Pastore, 2002). Bu açıdan Türk kültürüne uyarla-

ması yapılan  leđin biimlendirici deđerlendirme uygulamalarının arařtırılması ve geliřtirilmesine katkı sađlama potansiyeli bulunmaktadır.

Sonu olarak,  đretmenlerin Biimlendirici Deđerlendirme Uygulamaları  leđi, T rk eđitim sistemi bađlamında  đretmenlerin biimlendirici deđerlendirme stratejilerini ne  l de uyguladıklarını deđerlendiren geerli ve g venilir bir ara olarak kullanılabilir. Bu  lek,  đretmenlerin uygulama d zeylerini hem genel olarak hem de alt boyutlar kapsamında deđerlendirmeye olanak tanıyarak,  đretmenlerin biimlendirici deđerlendirme uygulamalarını daha etkin h le getirmelerine katkı sunabilir.

Arařtırmadan elde edilen sonulara g re arařtırmacılara y nelik řu  nerilerde bulunulmuřtur:

-  lek kullanılarak  đretmenlerin biimlendirici deđerlendirme uygulamalarının eřitli deđiřkenlere g re farklılařıp farklılařmadıđı incelenebilir.
-  đretmenlere verilebilecek biimlendirici deđerlendirmeye y nelik hizmet ii eđitimlerin  ncesi ve sonrasında uygulanabilir.
- T rkiye ve farklı  lkelerde g rev yapan  đretmenlerin biimlendirici deđerlendirme uygulamalarını incelemeye y nelik karřılařtırmalı alıřmalar yapılabilir.
-  leđin farklı  rneklem grupları  zerinde geerlik ve g venirlik alıřmaları yapılabilir.

## TEŐEKK R VE AIKLAMALAR

alıřmamıza katkı sađlayan t m  đretmenlerimize teŐekk r ederiz.

## IKAR ATIŐMASI

Makalenin yazarları arasında, alıřma kapsamında herhangi bir kiřisel ve finansal ıkar atiŐması bulunmamaktadır.

## YAZAR KATKILARI

Arařtırma Y ntemi: CO(%70), EB(%30)

Veri Toplama: CO(%50), EB(%50),

Veri Analizi: CO(%100)

Makalenin Hazırlanması: CO(%50), EB(%50)

## KAYNAKLAR

- Akbaş, Ç., & Ozan, C. (2024). The effects of formative assessment on student achievement and attitudes in math. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, 18(2).
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16(1), 74-94.
- Bandalos, D. L., & Finney, S. J. (2010). Factor analysis: Exploratory and confirmatory. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 93-114). New York, NY: Routledge.
- Black, P., & William, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>
- Black, P., & William, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.
- Brookfield, S. D., & Preskill, S. (2005). *Discussion as a way of teaching: Tools and techniques for democratic classrooms* (2nd ed.). Jossey-Bass.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford.
- Büyükköztürk, Ş., Çakmak-Kılıç, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2020). *Bilimsel araştırma yöntemleri* (28. baskı). Pegem Akademi.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Routledge.
- Conrad, R. M., & Donaldson, J. A. (2016). *Engaging online learners: Activities and resources for creative instruction*. Jossey-Bass.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Çepni, S. (2010). *Araştırma ve proje çalışmalarına giriş* (5. baskı). Erol.
- Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist*, 26(3-4), 325-346. [https://doi.org/10.1207/s15326985sep2603e4\\_6](https://doi.org/10.1207/s15326985sep2603e4_6)
- DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). Sage.
- Fryer, R. G., Levitt, S. D., List, J., & Sadoff, S. (2012). *Enhancing the efficacy of teacher incentives through loss aversion: A field experiment* (No. w18237). National Bureau of Economic Research.
- Güvendir, M. A. & Özkan, Y. Ö. (2015). Türkiye'deki eğitim alanında yayımlanan bilimsel dergilerde ölçek geliştirme ve uyarılma konulu makalelerin incelenmesi. *Elektronik Sosyal Bilimler Dergisi*, 14(52), 23-33.
- Hair, J. F. Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson Prentice Hall.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27, 53-64. <https://doi.org/10.1080/03075070120099368>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Scientific Software International.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.
- Kaptan, F. (1998). *Fen bilgisi öğretimi*. İstanbul: MEB Öğretmen Kitapları Dizisi. Fen Bilgisi Öğretimi, Anı Yayıncılık.
- Karaman, P. (2014). Öğretmen adaylarının ölçme değerlendirme okuryazarlıklarının belirlenmesi ve mikro öğretim yolu ile geliştirilmesi. (Tez No. 10028977) [Doktora Tezi, On Sekiz Mart Üniversitesi, Çanakkale]. Yükseköğretim Kurulu Ulusal Tez Merkezi.
- Karaman, P. (2017). Öğretmen adayları için biçimlendirici değerlendirmeye yönelik tutum ve niyet ölçeği uyarılması ve bir yapısal eşitlik modeli. İnönü Üniversitesi Eğitim Fakültesi Dergisi, 18(3), 118-131. <https://doi.org/10.17679/inuefd.300688>
- Karaman, P., & Şahin, Ç. (2017). Adaptation of Teachers' Conceptions and Practices of Formative Assessment Scale into Turkish culture and a structural equation modeling. *International Electronic Journal of Elementary Education*, 10(2), 185-194. <https://doi.org/10.26822/iejee.2017236114>
- Karasar, N. (2000). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayınları.
- Kıncal, R. Y., & Ozan, C. (2018). Effects of formative assessment on prospective teachers' achievement, attitude and self-regulation skills. *International Journal of Progressive Education*, 14(2), 77-92. <https://doi.org/10.29329/ijpe.2018.1396>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford.
- Kline, P. (1994). *An easy guide to factor analysis*. Routledge.

- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284. <https://doi.org/10.1037/0033-2909.119.2.254>
- Küçükahmet, L. (2008). Öğretimde planlama ve değerlendirme. Ankara: Nobel Dağıtım.
- Marzano, R. J. (2000). *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.
- McTighe, J., & Wiggins, G. (2013). *Essential questions: Opening doors to student understanding*. Association for Supervision and Curriculum Development.
- Millî Eğitim Bakanlığı [MEB]. (2020). Okul ve Sınıf Tabanlı Değerlendirmeye Dayalı Öğretmen Kapasitesinin Güçlendirilmesi Matematik Dersi Öğretmen Rehber Kitapçığı.
- Nunnally, J., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
- Ozan, C., & Kınca, R. Y. (2018). The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills. *Educational Sciences-Theory & Practice*, 18, 85-118. <https://doi.org/10.12738/estp.2018.1.0216>
- Pallant, J. (2005). *SPSS survival manual: A step by step guide to data analysis using SPSS for Windows version 12* (2nd ed.). Allen & Unwin.
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95(4), 667-686. <https://doi.org/10.1037/0022-0663.95.4.667>
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144.
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research*, 8(2), 23-74.
- Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, 21(1), 43-69.
- Stanford, P., & Reeves, S. (2005). Assessment that drives instruction. *Teaching Exceptional Children*, Volume: 37 Number: 4, 18-22.
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.
- Topping, K. (1996). The effectiveness of peer tutoring in further and higher education: A typology and review of the literature. *Higher Education*, 32(3), 321-345.
- Walsh, S. A., & Sattes, D. E. (2016). *Quality questioning: Research-based practice to engage every learner* (2nd ed.). Corwin.
- Wiggins, G. P., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Association for Supervision and Curriculum Development.
- William, D., & Thompson, M. (2008). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53-82). Erlbaum.
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6), 806-838. <https://doi.org/10.1177/0011000006288127>
- Yan, Z., & Pastore, S. (2022). Assessing teachers' strategies in formative assessment: The teacher formative assessment practice scale. *Journal of Psychoeducational Assessment*, 40(5) 592-604.
- Zengin, Y., Bars, M., & Şimşek, Ö. (2017). Matematik öğretiminin biçimlendirici değerlendirme sürecinde Kahoot! ve Plickers uygulamalarının incelenmesi. *Ege Eğitim Dergisi*, 18(2), 602-626. <https://doi.org/10.12984/egedfd.318647>
- Zimmerman, B. J. (2002). *Becoming a self-regulated learner: An overview*. *Theory Into Practice*, 41(2), 64-70.



## ÖĞRETMENLERİN BİÇIMLENDİRİCİ DEĞERLENDİRME UYGULAMALARI ÖLÇEĞİ

Değerli katılımcı, bu çalışmanın amacı, öğretmenlerin bıçimlendirici değerlendirme uygulamalarına yönelik öğretmenlerin görüşlerinin belirlenmesidir. Ankette yer alan hiçbir sorunun doğru veya yanlış cevabı yoktur. Bu nedenle, vereceğiniz içten yanıtların araştırma sonuçlarının güvenilirliği açısından çok önemli olduğunu unutmayınız. Anketlerden elde edilecek kişisel bilgiler sadece bilimsel amaçlar için kullanılacak olup kesinlikle gizli tutulacaktır. Çalışmaya gönüllü olarak katıldığınızı ve verdiğiniz bilgilerin bilimsel amaçlar için kullanılmasını onayladığınızı belirtmiş olacaksınız. Çalışmaya yapacağınız değerli katkılar için şimdiden teşekkür ederiz.

1. Kesinlikle katılmıyorum
2. Nadiren katılıyorum
3. Bazen katılıyorum
4. Katılıyorum
5. Kesinlikle katılıyorum

	Original Version	Turkish Version					
1	I share the learning intention before students start working in class.	Derse başlamadan önce kazanımları öğrencilerle paylaşıyorum.	1	2	3	4	5
2	I clarify what is valued for each assessment task.	Verdiğim değerlendirme ödevlerinde nelere dikkat etmeleri gerektiğini öğrencilere açıklarım.	1	2	3	4	5
3	I use various assessment activities in the classroom to check students' mastery of course content.	Öğrencilerin kazanımlara ulaşım ulaşmadıklarını belirlemek için çeşitli ölçme değerlendirme araçları kullanırım.	1	2	3	4	5
4	I ensure homework can check students' learning progress.	Ev ödevlerinin öğrencilerin öğrenmelerinin niteliğini belirleyebilmesini sağlarım.	1	2	3	4	5
5	I point out students' strengths and weaknesses in my feedback.	Geri bildirimlerimde öğrencilerin güçlü ve zayıf yönlerini belirtirim.	1	2	3	4	5
6	I provide suggestions for students to improve their performance.	Öğrencilere performanslarını artıracak öneriler sunarım.	1	2	3	4	5
7	I ask students to evaluate peers' work.	Öğrencilerden arkadaşlarının çalışmalarını değerlendirmelerini isterim.	1	2	3	4	5
8	I ask students to provide feedback to help peers improve.	Öğrencilerden, arkadaşlarının gelişmesine yardımcı olmaları için geri bildirim vermelerini isterim.	1	2	3	4	5
9	I ask students to identify strengths and weaknesses in their own work.	Öğrencilerden kendi çalışmalarındaki güçlü ve zayıf yönleri belirlemelerini isterim.	1	2	3	4	5
10	I ask students to identify strategies that will improve their own work.	Öğrencilerden kendi çalışmalarını geliştirecek stratejiler belirlemelerini isterim.	1	2	3	4	5

## ADAPTATION OF THE TEACHERS' FORMATIVE ASSESSMENT PRACTICES SCALE TO TURKISH CULTURE

### ABSTRACT

Formative assessment is based on identifying and eliminating students' learning deficiencies by actively using them at all stages of the learning process. In the formative assessment process, student progress is not expressed through grades. In recent years, the use of formative assessment practices has become widespread. With this study, the formative assessment application scale will be adapted to Turkish culture and used in the learning and teaching process. The purpose of this study is to analyze the validity and reliability of the teachers' formative assessment practices scale by adapting it to Turkish culture. The research design of the study is the survey model. There are two groups in the working group. In the first group, there are a total of 177 teachers working at different levels under the Ministry of National Education. In the second group, there are 151 teachers working at different levels under the Ministry of National Education. "Teachers' Formative Assessment Practices Scale" developed by Yan and Pastore was used as a data collection tool in the study. Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) techniques were used to analyze the data obtained from the study group. According to the exploratory factor analysis conducted as a result of the study, results were obtained that were exactly compatible with the structure of the original scale. According to the exploratory factor analysis conducted as a result of the study, results were obtained that were exactly compatible with the structure of the original scale. According to the confirmatory factor analysis of the study, the two-factor structure of the scale was confirmed. According to the confirmatory factor analysis of the study, the two-factor structure of the scale was confirmed. Consistent with the original scale, the first factor, consisting of six items, was named "Teacher-Oriented Formative Assessment", and the second factor, consisting of four items, was named "Student-Oriented Formative Assessment". The data obtained for reliability show that the measurements are reliable for the general scale and its sub-dimensions. As a result, a valid and reliable scale that can be used to examine the formative practices of teachers in Turkish culture has been obtained.

**Keywords:** Formative Assessment, Scale Adaptation, Measurement and Evaluation.



## ÖĞRETMENLERİN BİÇİMLENDİRİCİ DEĞERLENDİRME UYGULAMALARI ÖLÇEĞİ'NİN TÜRK KÜLTÜRÜNE UYARLANMASI

### ÖZ

Biçimlendirici değerlendirme öğrenme sürecinin tüm aşamalarında aktif bir şekilde kullanılarak öğrencilerin öğrenme eksikliklerinin belirlenmesi ve giderilmesine dayanır. Biçimlendirici değerlendirme sürecinde öğrencinin gelişimi not ile ifade edilmez. Son yıllarda biçimlendirici değerlendirme uygulamalarının kullanımı yaygınlaşmıştır. Bu çalışma ile biçimlendirici değerlendirme uygulama ölçeği Türk kültürüne uyarlanarak öğrenme öğretme sürecinde kullanımı sağlanacaktır. Bu çalışmanın amacı, öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği'nin Türk kültürüne uyarlanarak geçerlik ve güvenilirlik analizlerinin yapılmasıdır. Çalışmanın araştırma deseni tarama modelidir. Çalışma grubunda iki grup yer almaktadır. İlk grupta Millî Eğitim Bakanlığına bağlı farklı kademelerde görev yapan toplam 177 öğretmen bulunmaktadır. İkinci grupta ise Millî Eğitim Bakanlığına bağlı farklı kademelerde görev yapan 151 öğretmen bulunmaktadır. Çalışmada veri toplama aracı olarak Yan ve Pastore tarafından geliştirilen "Öğretmenlerin Biçimlendirici Değerlendirme Uygulamaları Ölçeği" kullanılmıştır. Çalışma grubundan elde edilen verilerin analizine yönelik açımlayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) teknikleri kullanılmıştır. Çalışma sonucunda yapılan açımlayıcı faktör analizine göre ölçeğin orijinalindeki yapı ile birebir uyumlu sonuçlar elde edilmiştir. Çalışmanın doğrulayıcı faktör analizine göre ölçeğin iki faktörlü yapısı doğrulanmıştır. Çalışmanın doğrulayıcı faktör analizine göre ölçeğin iki faktörlü yapısı doğrulanmıştır. Orijinal ölçek ile uyumlu şekilde, altı maddeden oluşan birinci faktör "Öğretmen Yönelimli Biçimlendirici Değerlendirme", dört maddeden oluşan ikinci faktör "Öğrenci Yönelimli Biçimlendirici Değerlendirme" olarak isimlendirilmiştir. Güvenirlik için elde edilen veriler ölçeğin geneli ve alt boyutları için ölçümlerin güvenilir olduğunu göstermektedir. Sonuç olarak Türk kültüründe öğretmenlerin biçimlendirici uygulamalarını incelemek amacıyla kullanılabilir geçerli ve güvenilir bir ölçek elde edilmiştir.

**Anahtar Sözcükler:** Biçimlendirici Değerlendirme, Ölçek Uyarlama, Ölçme ve Değerlendirme.



## INTRODUCTION

Measurement refers to observing any attribute and expressing the results either numerically or verbally (Stanford & Reeves, 2005). On the other hand, the concept of assessment can be defined as the process of reaching a judgment based on measurement results using specific criteria. While measurement involves observing and describing the current state, the act of assessment aims to draw a conclusion or make a judgment. From this perspective, the concepts of measurement and assessment should not be confused (Küçükahmet, 2008).

Assessment processes can be used at different stages of the teaching and learning process. Assessment activities can be carried out at the beginning, middle, or end of the learning process (Marzano, 2000). These activities aim to determine students' achievements, identify deficiencies in their learning process, and assess the effectiveness of the curriculum (Karaman, 2014). Assessment techniques used in the learning process can be classified into diagnostic assessment, formative assessment, and summative assessment.

Diagnostic assessment is conducted at the beginning of the learning process, primarily focusing on placing students into a program, course, or activity. Formative assessment, however, is carried out during the learning process to identify and address students' learning deficiencies. It can also identify problematic aspects of an curriculum (Fryer et al., 2012). Summative assessment, conducted at the end of the learning process, quantitatively determines the extent to which students have achieved the learning objectives. The primary aim is to ascertain how many of the intended learning outcomes have been achieved by the students.

Formative assessment, often defined as an assessment technique used for learning, involves providing students with feedback throughout all stages of the learning process to help them achieve their learning objectives (Sadler, 1989). By supporting students' development during the process and identifying and addressing learning deficiencies, the process is maintained. The goal is to support development holistically rather than merely assigning grades to students (Zengin et al., 2017).

Formative assessment activities can be employed at every stage of the teaching and learning process (MEB, 2020). Two key points stand out in formative assessment. The first is enabling students to recognize their potential to achieve learning objectives. The second is encouraging them to address their learning deficiencies to reach these objectives. Formative assessment achieves its purpose when these two actions are successfully carried out (Black & William, 1998a).

William and Thompson (2008) proposed a framework categorizing various formative assessment practices into five key strategies: (1) Clarifying, sharing, and



understanding learning intentions and success criteria, (2) Designing effective classroom discussions, questions, and learning tasks, (3) Providing feedback that moves learners forward, (4) Activating students as instructional resources for one another, and (5) Activating students as owners of their learning (Wiliam & Thompson, 2008, p. 64).

Sharing learning objectives and success criteria is considered a crucial element for making the teaching process more transparent and student-centered. This approach aims to provide students with a clear understanding of their learning processes, enhance student motivation, and improve teaching strategies (Black & Wiliam, 1998b; Hattie & Timperley, 2007; Wiggins & McTighe, 2005).

Designing effective classroom discussions, questions, and learning tasks is another fundamental aspect of formative assessment. Effective classroom discussions enable students to share their ideas, understand different perspectives, and develop critical thinking skills. Managing the discussion environment to foster respectful participation is crucial. Open-ended questions can elicit students' thought processes and enhance the quality of discussions (Brookfield & Preskill, 2005). Questions should target critical thinking, problem-solving, and analytical skills beyond mere knowledge-level inquiries. Designing questions that align with Bloom's taxonomy—particularly those addressing analysis, assessment, and creation levels—is essential (McTighe & Wiggins, 2013; Walsh & Sattes, 2016). Learning tasks that connect to real-world contexts and engage students help deepen learning. Examples include project-based learning, group work, and field studies (Conrad & Donaldson, 2016).

Providing high-quality feedback is a vital component of formative assessment, playing a critical role in improving learning and supporting student development. Effective feedback should focus on the student's needs and goals, assess performance relative to learning objectives and standards, and be constructive by emphasizing strengths and using positive language (Black & Wiliam, 1998b; Hattie & Timperley, 2007). Continuous feedback during the formative assessment process enables students to make timely corrections and improvements. Regular progress reports and consistent communication between teachers and students are important for providing ongoing feedback.

Another element of formative assessment is encouraging students to serve as instructional resources for each other, fostering collaboration and interaction. Peer assessment is part of this strategy, allowing students to learn from one another, enhancing the overall learning experience (Slavin, 1996; Topping, 1996).

Finally, activating students as owners of their learning aims to involve them more actively and consciously in the learning process. This strategy encourages

students to associate learning not only with acquiring knowledge but also with creating, connecting, and understanding it. Self-assessment practices are also included in this strategy (Deci et al., 1991; Pintrich, 2003; Zimmerman, 2002).

The effective and sustainable operation of the education system relies on teachers' ability to assess and guide students. In this context, teachers' formative assessment practices emerge as a critical tool to enhance student achievement, contribute to student development, and improve educational processes (Akbaş & Ozan, 2024; Kıncal & Ozan; Ozan & Kıncal). Yan and Pastore (2022) point out that although formative assessment has been a popular topic in educational research over the past two decades, there is still a lack of widely recognized and applied survey instruments. In Türkiye, measurement tools related to formative assessment are also quite limited (Karaman, 2017; Karaman & Şahin, 2017). Karaman (2017) adapted the scale developed by Yan and Cheng (2015) for teachers on attitudes, intentions, and practices related to formative assessment to preservice teachers. Karaman and Şahin (2017) translated a scale into Turkish based on the Theory of Planned Behavior to examine teachers' formative assessment intentions and practices. The scale includes dimensions such as affective attitude, attitude, subjective norm, controllability, self-efficacy, intention, and behavior.

The scale to be adapted in this study is based on the framework developed by Wiliam and Thompson (2008). This scale evaluates formative assessment practices within the context of five key strategies more specifically. Its distinction from other scales lies in its more focused and practice-oriented approach to formative assessment. Consequently, it measures teachers' strategic practices in assessment processes more clearly.

Through the adaptation of the Teachers' Formative Assessment Practices Scale to Turkish culture, it will be possible to evaluate formative assessment strategies of teachers in alignment with local norms and values, as well as determine the role of these strategies within the Turkish education system. Teachers' effective implementation of formative assessment strategies plays a critical role in enhancing student achievement and improving classroom interactions. With the adapted scale, it will be possible to evaluate and improve teachers' formative assessment practices. Identifying deficiencies in the implementation of formative assessment strategies may contribute to the development of more effective and guiding education policies in this area. Additionally, the development of the scale in a manner aligned with Turkish culture will enable Türkiye's data to be more meaningful and comparable in international educational research.

In line with these justifications, the aim of this study is to adapt the Teachers' Formative Assessment Practices Scale developed by Yan and Pastore (2022) into Turkish culture by conducting linguistic equivalence, validity, and reliability analyses.

## METHOD

### Research Design

This study employs a survey design. Survey research aims to identify the characteristics of an event or the opinions of a group (Büyüköztürk et al., 2020). According to another definition, the survey model seeks to describe events, situations, phenomena, or individuals, whether from the past or present, in their natural settings without any intervention (Karasar, 2000). Given that this study, as a scale adaptation, aims to describe the existing state regarding the characteristics being measured, the survey design was deemed appropriate.

### Study Group

The research includes two study groups. Data from the first group were used for exploratory factor analysis (EFA), while data from the second group were utilized for confirmatory factor analysis (CFA). The literature indicates that conducting independent analyses for EFA and CFA yields more reliable results (Brown, 2015; Tabachnick & Fidell, 2019). This approach ensures that the structure of the measurement instrument is revealed in the first group and then validated in the second group. Using two separate groups reduces the risk of overfitting when conducting both EFA and CFA on the same sample (Worthington & Whittaker, 2006). Thus, two distinct groups were selected for this study.

The first study group comprised 177 teachers working at various levels and stages within Erzurum province during the fall semester of the 2023-2024 academic year. DeVellis (2016) and Worthington and Whittaker (2006) suggest that in scale development studies, the sample size should be 5 to 10 times the number of items, while Nunnally and Bernstein (1994) recommend a minimum of 10 times the number of items. Given that the scale being adapted in this study consists of 10 items, the sample size in both study groups was deemed sufficient. The demographic information for the first study group is presented in Table 1.

**Table 1.** *Demographic information of the first study group*

Variables	Categories	n	%
Gender	Female	86	49
	Male	91	51
Teaching level	Primary School	33	19
	Middle School	85	48
	High School	59	33
Years of Experience	1-5 years	65	37
	6-10 years	53	30
	11 years or more	59	33

Of the teachers in the first study group, 86 are female and 91 are male; 33 work at the primary school level, 85 at the middle school level, and 59 at the high school level. Among these teachers, 65 have 1–5 years of experience, 53 have 6–10 years of experience, and 59 have 11 years or more. Table 2 presents the demographic information of the second study group.

**Table 2.** *Demographic information of the second study group*

Variables	Categories	n	%
<b>Gender</b>	Female	122	81
	Male	29	19
Teaching level	Primary School	38	25
	Middle School	74	49
	High School	39	26
<b>Years of Experience</b>	1-5 years	69	46
	6-10 years	45	30
	11 years or more	37	24

Of the teachers in the second study group, 122 are female and 29 are male; 38 work at the primary school level, 74 at the middle school level, and 39 at the high school level. Among these teachers, 69 have 1–5 years of experience, 45 have 6–10 years of experience, and 37 have 11 years or more.

### Data Collection Tools

The Formative Assessment Practices Scale for Teachers, developed by Yan and Pastore (2022), was used as the data collection instrument in this study. The development process of the scale was carried out as follows:

1. **Review of the Literature:** Before developing the scale, the relevant literature was thoroughly reviewed.
2. **Focus Group Discussions:** Focus group discussions were conducted with primary and secondary school teachers in Hong Kong using a preliminary 15-item scale. During this process, teachers' opinions were gathered.
3. **Expert Opinions:** Based on the feedback from focus group discussions and experts, five items were removed from the scale, and adjustments were made to the remaining items.
4. **Likert-Type Scale:** The items on the scale were rated on a six-point Likert scale ranging from "Never," "Almost Never," "Rarely," "Sometimes," "Often," to "Always."

The participation of the sample groups and the conducted analyses are outlined below:

1. **Sample Groups:** Two sample groups were selected for the implementation of the scale: 449 teachers from 12 schools in Hong Kong and 309 teachers from 10 schools in Italy participated in the study.
2. **Exploratory Factor Analysis (EFA):** The EFA revealed that the scale had a two-factor structure:
  - Teacher-Oriented Formative Assessment (6 items)
  - Student-Oriented Formative Assessment (4 items)

Each factor included two items based on the formative assessment categories outlined by Wiliam and Thompson (2008). The correlation between the two factors was calculated as 0.52.

3. **Confirmatory Factor Analysis (CFA):** The results of the CFA were as follows: RMSEA = 0.056, SRMR = 0.062, CFI = 0.94. Additionally, modifications were made between Item 1 and Item 2, and Item 9 and Item 10.
4. **Rasch Analysis:** According to the Rasch analysis, all items demonstrated a good fit to the Rasch model.

The reliability analysis results were as follows:

1. **Hong Kong Sample:** Cronbach's alpha was calculated as 0.77 for the first factor and 0.87 for the second factor.
2. **Italy Sample:** Cronbach's alpha was calculated as 0.60 for the first factor and 0.66 for the second factor.
3. **Combined Reliability:** For the combined data of both groups, Cronbach's alpha was calculated as 0.70 for the first factor and 0.75 for the second factor.

### Linguistic Equivalence

To ensure linguistic equivalence, a translation and back-translation method was applied. The items of the scale were independently translated into Turkish by three English language teachers. The translated items were reviewed by three Turkish language teachers, and the revised items were then back-translated into English by a faculty member from the English Language Teaching Department. The back-translated items were compared with the original English items to ensure consistency. Following this evaluation, the translation-back-translation process was finalized, and the Turkish version of the scale was prepared.

After completing the translation-back-translation process, both the English and Turkish forms of the scale were administered to the same sample group. A correlation analysis was conducted to assess the linguistic equivalence of the scale. The Pearson correlation coefficient was calculated, revealing a strong positive relationship between the scores obtained from the English and Turkish forms ( $r = 0.84$ ,  $p < 0.001$ ). This result indicates that the scale is consistent across both languages and that linguistic equivalence has been achieved. These findings demonstrate that the adapted Turkish version of the scale closely aligns with the original English form, with no significant differences between the language versions.

### Data Collection

Before commencing data collection, the necessary legal permissions were obtained to conduct the study. Initially, adaptation approval was sought via email from Zi Yan, one of the two authors who developed the scale. Subsequently, permissions were secured from the Ethics Committee of the Educational Sciences Unit and the Provincial Directorate of National Education. Data were collected using a survey created on Google Forms.

### Data Analysis

Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) techniques were employed to analyze the data collected from the study group. EFA was used to determine the structural validity of the original form of the Teachers' Formative Assessment Practices Scale in Turkish culture. While EFA identifies the relationships among variables, CFA examines whether the predefined relationships among variables meet the expected level. CFA evaluates which variables are related to one another, whether the factors are independent, and whether the factors are adequate in explaining relationships. The construct validity of the data obtained through EFA is tested via CFA (Güvendir & Özkan, 2015). An item's factor loading value is typically preferred to be 0.45 or higher. However, for scales with fewer items, factor loading values as low as 0.30 may be acceptable (Hair et al., 2010; Kline, 1994). In this study, the threshold for acceptable factor loadings was set at 0.30.

Before beginning the analysis, the correlation matrix of all items was examined to assess the presence of significant correlations, identifying meaningful relationships appropriate for factor analysis. To evaluate sample adequacy, the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's Test of Sphericity were conducted. A KMO value above 0.60 and a significant Bartlett's test result are required for the data to be suitable for factor analysis (Kaiser, 1974; Pallant, 2005).

## Ethics Committee Approval

*(There is no requirement of Ethics Committee Approval for review articles)*

Ethics committee approval was received for this study from Atatürk University, Faculty of Kâzım Karabekir Education

The Title of The Ethics Committee: Atatürk University Educational Sciences Unit Ethics Committee

Approval Date: 31.01.2024

Ethics Document's Number: 01/59

## FINDINGS

In the efforts to adapt the scale to Turkish culture, validity and reliability studies were conducted.

### Construct Validity

**Exploratory Factor Analysis (EFA).** The EFA results revealed that the Kaiser-Meyer-Olkin (KMO) Sampling Adequacy Coefficient was calculated as .84. Additionally, the Bartlett's Test of Sphericity  $\chi^2$  value was found to be 614.962 ( $p < .01$ ). These values indicate that the data were suitable for factor analysis.

According to the analysis results, two factors emerged, explaining 53.82% of the total variance. Since all items had factor loadings greater than .30, no items were removed from the scale. During the factor analysis, the Varimax rotation method was employed to achieve better interpretability of the data. This rotation method aims to ensure that the factors are independent of each other, allowing each factor to be more distinct and interpretable (Tabachnick & Fidell, 2019). There was no difference in the rotation method used compared to the original scale. The values related to the factor structure of the Turkish version of the scale are presented in Table 4.

**Table 3.** *Items and Factor Loadings*

Item No	Factor Loadings	
	1	2
1	.501	
2	.794	
3	.690	
4	.688	

	5	.772	
	6	.769	
	7		.858
	8		.834
	9		.556
	10		.595
Explained Variance		%40.933	%12.885
Total=		%53.818	
AVE A= 1. Factor = 0.50; 2. Factor = 0.52			

As seen in Table 3, the Turkish version of the scale consists of two dimensions. The total explained variance of the scale is 53.82%. The factor loadings of the items range from .50 to .86. The average explained variance (AEV) values for the factors were calculated as .050 for Factor 1 and .052 for Factor 2. These AEV values indicate that both convergent and discriminant validity have been established (Bagozzi & Yi, 1988; Nunnally & Bernstein, 1994).

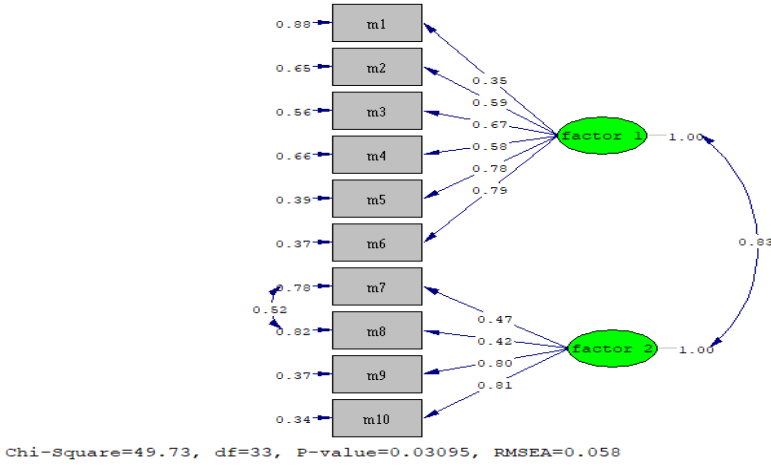
According to the first-level Confirmatory Factor Analysis (CFA) results, the fit indices of the model are presented in Table 4.

**Table 4.** Values for First-Level CFA Fit Indices

Chi-square	df	p	Chi-square/df	RMSEA	SRMR	GFI	AGFI	NFI	NNFI	CFI
49.73	33	p=0.03	1.51	0.058	0.042	0.94	0.90	0.95	0.98	0.98

According to the first-level CFA results, it was determined that the RMSEA value and the chi-square/df ratio were not at the desired level initially. Therefore, a modification was made by freeing the correlation between M7 and M8, which are in the same factor. After the modification, the ratio of the chi-square value to the degrees of freedom was calculated to be less than 2, which is considered good. Upon evaluating other fit indices, it was observed that all indices were within good or acceptable limits (Schermelleh-Engel et al., 2003). Based on these results, it was concluded that the model-data fit was achieved, and the construct validity of the scale was confirmed. The path diagram of the confirmatory factor analysis is shown in Figure 1.





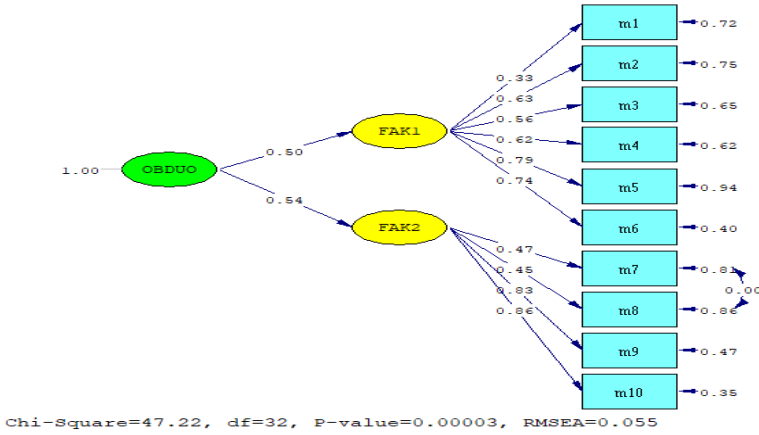
**Figure 1.** Path Diagram for First-Level CFA

A second-level CFA was conducted to determine the fit of the two factors emerging from the first-level CFA and whether these factors represent a single higher-order factor (Byrne, 2010; Kline, 2016). The fit indices for the second-level CFA results are presented in Table 5.

**Table 5.** Values for the Goodness of Fit Indices in the Second-Level CFA

Chi-square	df	p	Chi-square /df	RMSEA	SRMR	GFI	AGFI	NFI	NNFI	CFI
47.22	32	p=0.03	1.48	0.055	0.040	0.92	0.91	0.96	0.97	0.97

According to the results of the second-level CFA, the initial RMSEA value and chi-square/df ratio were found to be below the desired level. Therefore, similar to the first-level CFA, correlations between M7 and M8 within the same factor were freed, and a modification was made. After the modification, the results of the second-level CFA goodness-of-fit indices show that the model demonstrates good fit. The chi-square/df ratio is 1.48, which is below 2, indicating good fit (Kline, 2016). The RMSEA value is 0.055, and the SRMR value is 0.040, both of which are within the excellent fit limits (Hu & Bentler, 1999). The GFI and AGFI values are 0.92 and 0.91, respectively, indicating that the model has an acceptable fit (Jöreskog & Sörbom, 1996). The NFI, NNFI, and CFI values are 0.96, 0.97, and 0.97, respectively, suggesting strong and excellent fit. The path diagram for the second-level CFA is shown in Figure 2.



**Figure 2.** Path Diagram for the Second-Level CFA

According to Figure 2, the relationships between TFAPS and Factor 1 and Factor 2 are 0.50 and 0.54, respectively. These values indicate that TFAPS significantly represents both factors.

### Madde Analizi

To determine how well each scale item discriminates teachers in terms of formative assessment practices, item-total correlations were first calculated. Then, a t-test was used to assess the significance of the differences in item scores between the top 27% and bottom 27% groups, based on the total score. The results are presented in Table 6.

**Tablo 6.** Madde Toplam Korelasyonları ve Madde Ayırt Ediciliği İçin T Testi

Item No	Corrected Item-Total Correlations	t
1	.482	2.497*
2	.441	2.431*
3	.561	2.108*
4	.538	2.600*
5	.461	2.831*
6	.568	2.640*
7	.533	2.588*
8	.490	2.469*
9	.597	2.412*
10	.614	2.192*

\*p < .001

The corrected item-total correlations of the scale were found to range from .44 to .61. Additionally, t-values for the differences in item scores between the top 27% and bottom 27% groups, based on total scores, varied between 2.108 ( $p < .001$ ) and 2.831 ( $p < .001$ ) (Nunnally & Bernstein, 1994). The t-test results revealed that for all items, the item total scores of the top 27% group were significantly higher ( $p < 0.001$ ) than those of the bottom 27% group.

The Turkish version of the Teachers' Formative Assessment Practices Scale consists of 10 items distributed across two factors. The first factor, named "Teacher-Oriented Formative Assessment Practices," includes 6 items, while the second factor, named "Student-Oriented Formative Assessment Practices," contains 4 items. There are no negatively worded items. In accordance with the original scale, all items are rated on a six-point Likert scale ranging from "Never," "Almost Never," "Rarely," "Sometimes," "Frequently," to "Always." The highest possible score on the scale is 60, while the lowest is 10. A total score can be obtained from the scale. The total score reflects the extent to which teachers engage in formative assessment practices. Scores from the first dimension indicate the level of teacher-oriented formative assessment practices, while scores from the second dimension indicate the level of student-oriented formative assessment practices.

### Güvenirlilik

The data related to the Cronbach's alpha coefficient and the split-half method applied to determine the reliability of the scale are presented in Table 6.

**Table 6.** *Reliability of the Scale*

Factor	Item Number	Cronbach's Alfa Coefficient	Guttman Split-Half Coefficient
General	10	.84	.73
Teacher-Oriented	6	.80	
Student-Oriented	4	.72	

The Cronbach's alpha coefficient for the overall scale was calculated as .84, and the Guttman split-half coefficient was calculated as .73. For the Teacher-Oriented Formative Assessment subscale, the Cronbach's alpha was .80, and for the Student-Oriented Formative Assessment subscale, it was .72. Based on these results, it can be concluded that the measurements obtained from the scale are reliable (Cronbach, 1951).

## RESULT AND DISCUSSION

In the study, linguistic equivalence, validity, and reliability studies were conducted to adapt the Teacher Formative Assessment Practices Scale to Turkish culture. The original scale consists of 10 items across two subscales. The exploratory factor analysis conducted in this study yielded results that were perfectly consistent with the structure of the original scale. The confirmatory factor analysis also confirmed the two-factor structure of the scale. In accordance with the original scale, the first factor, consisting of six items, was named “Teacher-Oriented Formative Assessment,” and the second factor, consisting of four items, was named “Student-Oriented Formative Assessment.” The first factor is generally related to formative assessment strategies initiated and implemented by teachers (e.g., explaining and sharing learning goals and success criteria, conducting effective class discussions, designing questions and learning tasks, and providing feedback to enhance student development), while the second factor is related to student-centered formative assessment strategies, which are core components of formative assessment, such as self-assessment and peer assessment. A total score can be obtained from the scale. A higher score indicates that teachers incorporate more formative assessment practices into their teaching process. Additionally, the scores obtained from the scale can be evaluated within the subscales. Scores from the subscales allow for the evaluation of whether teachers’ formative assessment practices are teacher-oriented or student-oriented.

The reliability analyses of the scale also show that the data obtained from the scale are reliable. As a result, it has been determined that the Teacher Formative Assessment Practices Scale is a valid and reliable tool for measuring teachers’ formative assessment practices in the Turkish culture.

This study differs from some existing scales related to formative assessment in certain aspects. For example, in the study by Karaman (2017), the Formative Assessment Attitude, Intention, and Practice Scale developed by Yan and Cheng (2015) was adapted for teacher candidates. Karaman and Şahin (2017) adapted a formative assessment scale based on the Theory of Planned Behavior into Turkish. These scales address various dimensions such as attitudes, intentions, and controllability regarding formative assessment. The scale used in this research, however, is based on the framework of Wiliam and Thompson (2008) and evaluates formative assessment practices within five key strategies. While Karaman (2017) and Karaman and Şahin’s (2017) studies focus on attitudes, intentions, and practices, the scale adapted in this study directly focuses on how teachers implement formative assessment strategies. While Karaman and Şahin’s (2017) scale investigates affective attitudes and behaviors with a broader approach, the scale in this study provides a more application-oriented measurement tool. Furthermore, this scale places greater emphasis on student-centered strategies and comprehensively addresses both teacher- and student-oriented dimensions of formative assessment.

So far, how teachers effectively apply formative assessment has been an important agenda in educational research and educational reforms. However, the lack of a valid and easy-to-use tool for evaluating formative assessment practices has negatively impacted teacher development programs as well as teacher evaluation practices for a long time (Yan & Pastore, 2002). In this regard, the scale adapted for Turkish culture has the potential to contribute to the investigation and development of formative assessment practices.

In conclusion, the Teacher Formative Assessment Practices Scale can be used as a valid and reliable tool to assess the extent to which teachers apply formative assessment strategies in the context of the Turkish education system. This scale, by allowing an assessment of teachers' application levels both overall and within the subscales, can contribute to making teachers' formative assessment practices more effective.

Based on the results of the research, the following suggestions are made for researchers:

- The scale can be used to examine whether teachers' formative assessment practices differ according to various variables.
- In-service training on formative assessment can be applied before and after to teachers.
- Comparative studies can be conducted to examine the formative assessment practices of teachers in Turkey and other countries.
- Validity and reliability studies of the scale can be conducted with different sample groups.

## ACKNOWLEDGEMENTS

We would like to thank all the teachers who contributed to our study.

## CONFLICT OF INTEREST

There is no personal or financial conflict of interest between the authors of the article within the scope of the study.

## AUTHOR CONTRIBUTIONS

Research design: CO(%70), EB(%30)

Data collection: CO(%50), EB(%50),

Statistical analysis: CO(%100)

Preparation of the Article: CO(%50), EB(%50)

## REFERENCES

- Akbaş, Ç., & Ozan, C. (2024). The effects of formative assessment on student achievement and attitudes in math. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education*, 18(2).
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16(1), 74-94.
- Bandalos, D. L., & Finney, S. J. (2010). Factor analysis: Exploratory and confirmatory. In G. R. Hancock & R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 93-114). New York, NY: Routledge.
- Black, P., & Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. <https://doi.org/10.1080/0969595980050102>
- Black, P., & Wiliam, D. (1998b). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139-148.
- Brookfield, S. D., & Preskill, S. (2005). *Discussion as a way of teaching: Tools and techniques for democratic classrooms* (2nd ed.). Jossey-Bass.
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). Guilford.
- Büyüköztürk, Ş., Çakmak-Kılıç, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2020). *Bilimsel araştırma yöntemleri* (28. baskı). Pegem Akademi.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Routledge.
- Conrad, R. M., & Donaldson, J. A. (2016). *Engaging online learners: Activities and resources for creative instruction*. Jossey-Bass.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Çepni, S. (2010). *Araştırma ve proje çalışmalarına giriş* (5. baskı). Erol.
- Deci, E. L., Vallerand, R. J., Pelletier, L. G., & Ryan, R. M. (1991). Motivation and education: The self-determination perspective. *Educational Psychologist*, 26(3-4), 325-346. [https://doi.org/10.1207/s15326985sep2603&4\\_6](https://doi.org/10.1207/s15326985sep2603&4_6)
- DeVellis, R. F. (2016). *Scale development: Theory and applications* (4th ed.). Sage.
- Fryer, R. G., Levitt, S. D., List, J., & Sadoff, S. (2012). *Enhancing the efficacy of teacher incentives through loss aversion: A field experiment* (No. w18237). National Bureau of Economic Research.
- Güvendir, M. A. & Özkan, Y. Ö. (2015). Türkiye'deki eğitim alanında yayımlanan bilimsel dergilerde ölçek geliştirme ve uyarlama konulu makalelerin incelenmesi. *Elektronik Sosyal Bilimler Dergisi*, 14(52), 23- 33.
- Hair, J. F. Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson Prentice Hall.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112.
- Higgins, R., Hartley, P., & Skelton, A. (2002). The conscientious consumer: Reconsidering the role of assessment feedback in student learning. *Studies in Higher Education*, 27, 53-64. <https://doi.org/10.1080/03075070120099368>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Scientific Software International.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36.
- Kaptan, F. (1998). *Fen bilgisi öğretimi*, İstanbul: MEB Öğretmen Kitapları Dizisi. Fen Bilgisi Öğretimi, Anı Yayıncılık.
- Karaman, P. (2014). Öğretmen adaylarının ölçme değerlendirme okuryazarlıklarının belirlenmesi ve mikro öğretim yolu ile geliştirilmesi. (Tez No. 10028977) [Doktora Tezi, On Sekiz Mart Üniversitesi, Çanakkale]. Yükseköğretim Kurulu Ulusal Tez Merkezi.
- Karaman, P. (2017). Öğretmen adayları için biçimlendirici değerlendirmeye yönelik tutum ve niyet ölçeği uyarlanması ve bir yapısal eşitlik modeli. İnönü Üniversitesi Eğitim Fakültesi Dergisi, 18(3), 118-131. <https://doi.org/10.17679/inuefd.300688>
- Karaman, P., & Şahin, Ç. (2017). Adaptation of Teachers' Conceptions and Practices of Formative Assessment Scale into Turkish culture and a structural equation modeling. *International Electronic Journal of Elementary Education*, 10(2), 185-194. <https://doi.org/10.26822/iejee.2017236114>
- Karasar, N. (2000). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayınları.
- Kıncal, R. Y., & Ozan, C. (2018). Effects of formative assessment on prospective teachers' achievement, attitude and self-regulation skills. *International Journal of Progressive Education*, 14(2), 77-92. <https://doi.org/10.29329/ijpe.2018.139.6>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford.
- Kline, P. (1994). *An easy guide to factor analysis*. Routledge.

- Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, 119(2), 254-284. <https://doi.org/10.1037/0033-2909.119.2.254>
- Küçükahmet, L. (2008). Öğretimde *planlama ve değerlendirme*. Ankara: Nobel Dağıtım.
- Marzano, R. J. (2000). *Transforming classroom grading*. Alexandria, VA: Association for Supervision and Curriculum Development.
- McTighe, J., & Wiggins, G. (2013). *Essential questions: Opening doors to student understanding*. Association for Supervision and Curriculum Development.
- Millî Eğitim Bakanlığı [MEB]. (2020). Okul ve Sınıf Tabanlı Değerlendirmeye Dayalı Öğretmen Kapasitesinin Güçlendirilmesi Matematik Dersi Öğretmen Rehber Kitapçığı.
- Nunnally, J., & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
- Ozan, C., & Kincal, R. Y. (2018). The effects of formative assessment on academic achievement, attitudes toward the lesson, and self-regulation skills. *Educational Sciences-Theory & Practice*, 18, 85-118. <https://doi.org/10.12738/estp.2018.1.0216>
- Pallant, J. (2005). *SPSS survival manual: A step by step guide to data analysis using SPSS for Windows version 12* (2nd ed.). Allen & Unwin.
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95(4), 667-686. <https://doi.org/10.1037/0022-0663.95.4.667>
- Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119-144.
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research*, 8(2), 23-74.
- Slavin, R. E. (1996). Research on cooperative learning and achievement: What we know, what we need to know. *Contemporary Educational Psychology*, 21(1), 43-69.
- Stanford, P., & Reeves, S. (2005). Assessment that drives instruction. *Teaching Exceptional Children*, Volume: 37 Number: 4, 18-22.
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.
- Topping, K. (1996). The effectiveness of peer tutoring in further and higher education: A typology and review of the literature. *Higher Education*, 32(3), 321-345.
- Walsh, S. A., & Sattes, D. E. (2016). *Quality questioning: Research-based practice to engage every learner* (2nd ed.). Corwin.
- Wiggins, G. P., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Association for Supervision and Curriculum Development.
- William, D., & Thompson, M. (2008). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53-82). Erlbaum.
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6), 806-838. <https://doi.org/10.1177/0011000006288127>
- Yan, Z., & Pastore, S. (2022). Assessing teachers' strategies in formative assessment: The teacher formative assessment practice scale. *Journal of Psychoeducational Assessment*, 40(5) 592-604.
- Zengin, Y., Bars, M., & Şimşek, Ö. (2017). Matematik öğretiminin biçimlendirici değerlendirme sürecinde Kahoot! ve Plickers uygulamalarının incelenmesi. *Ege Eğitim Dergisi*, 18(2), 602-626. <https://doi.org/10.12984/eegefd.318647>
- Zimmerman, B. J. (2002). *Becoming a self-regulated learner: An overview*. *Theory Into Practice*, 41(2), 64-70.



## Teacher's Formative Assessment Practices Scale

Dear participant, the aim of this study is to determine teachers' views on their formative assessment practices. Please note that there are no right or wrong answers to any of the questions in this survey. Therefore, your sincere responses are crucial for the reliability of the research results. The personal information obtained from the surveys will be used solely for scientific purposes and will be kept strictly confidential. By participating in the study, you indicate that you consent to the use of the information for scientific purposes. Thank you in advance for your valuable contribution to the study.

1. Strongly disagree
2. Rarely agree
3. Sometimes agree
4. Agree
5. Strongly agree

	Original Version	Turkish Version						
1	I share the learning intention before students start working in class.	Derse başlamadan önce kazanımları öğrencilerle paylaşıyorum.	1	2	3	4	5	
2	I clarify what is valued for each assessment task.	Verdiğim değerlendirme ödevlerinde nelere dikkat etmeleri gerektiğini öğrencilere açıklarım.	1	2	3	4	5	
3	I use various assessment activities in the classroom to check students' mastery of course content.	Öğrencilerin kazanımlara ulaşip ulaşmadıklarını belirlemek için çeşitli ölçme değerlendirme araçları kullanırım.	1	2	3	4	5	
4	I ensure homework can check students' learning progress.	Ev ödevlerinin öğrencilerin öğrenmelerinin niteliğini belirleyebilmesini sağlarım.	1	2	3	4	5	
5	I point out students' strengths and weaknesses in my feedback.	Geri bildirimlerimde öğrencilerin güçlü ve zayıf yönlerini belirtirim.	1	2	3	4	5	
6	I provide suggestions for students to improve their performance.	Öğrencilere performanslarını artıracak öneriler sunarım.	1	2	3	4	5	
7	I ask students to evaluate peers' work.	Öğrencilerden arkadaşlarının çalışmalarını değerlendirmelerini isterim.	1	2	3	4	5	
8	I ask students to provide feedback to help peers improve.	Öğrencilerden, arkadaşlarının gelişmesine yardımcı olmaları için geri bildirim vermelerini isterim.	1	2	3	4	5	
9	I ask students to identify strengths and weaknesses in their own work.	Öğrencilerden kendi çalışmalarındaki güçlü ve zayıf yönleri belirlemelerini isterim.	1	2	3	4	5	
10	I ask students to identify strategies that will improve their own work.	Öğrencilerden kendi çalışmalarını geliştirecek stratejiler belirlemelerini isterim.	1	2	3	4	5	



