

Sınıf İçi Ölçme ve Değerlendirme Faaliyetlerini Değerlendirme Ölçeği: Geçerlik ve Güvenirlik Çalışması *

A Scale for Assessing Classroom Measurement and Evaluation Activities: A Validity and Reliability Study

Mehmet ŞATA¹, Fuat ELKONCA²

¹Ağrı İbrahim Çeçen Üniversitesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme
Anabilim Dalı. msata@agri.edu.tr.

²Muş Alparslan Üniversitesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme
Anabilim Dalı. f.elkonca@alparslan.edu.tr.

Makalenin Geliş Tarihi: 08.02.2021

Yayına Kabul Tarihi: 02.10.2021

ÖZ

Bu çalışmanın amacı, "sınıf içi ölçme ve değerlendirme faaliyetlerini değerlendirme" ölçeğini geliştirmek ve psikometrik özelliklerini belirlemektir. Araştırma var olan bir durumu ortaya çıkartmaya yönelik olduğundan nicel araştırma yaklaşımlarından tarama modeli ile yapılmıştır. Ölçeğin geliştirilmesi sürecinde alanyazın taraması yapılmış ve 56 maddeden oluşan bir taslak form hazırlanmıştır. Ölçme aracının kapsam geçerliği için uzman görüşleri alınmış ve altı maddenin yeterli kapsama sahip olmadığı belirlenip araştırma kapsamından çıkartılmıştır. Araştırmadaki katılımcılar çeşitli bölümlerde okuyan 400 öğretmen adayından oluşmaktadır. Geliştirilen ölçme aracından elde edilen ölçümlerin yapı geçerliğine kanıt sağlamak amacıyla açılımlı faktör analizi yapılmış olup üç faktör ve 40 maddeden oluşan bir ölçme aracı elde edilmiştir. Ölçme aracından elde edilen ölçümlerin güvenilirliğine ilişkin kanıt için Cronbach alfa tabakalı Cronbach alfa ve McDonald (ω) katsayıları hesaplanmış ve güvenirliliğin yüksek olduğu bulunmuştur. Ölçme aracının standart belirleme çalışması amacıyla iki aşamalı kümeleme analizi yapılmış düşük, orta ve yüksek olmak üzere üç küme belirlenmiştir. Sonuç olarak geliştirilen ölçme aracı kullanılarak elde edilen ölçme sonuçlarının güvenilir olduğu ve bu sonuçlara dayalı yapılan çıkarımların ise geçerli olduğu söylenebilir.

Anahtar Sözcükler: Sınıf içi ölçme ve değerlendirme, Uzaktan eğitim, Ölçek geliştirme, Güvenirlik, Geçerlik.

***Alıntılama:** Şata, M. ve Elkonca, F. (2021). Sınıf içi ölçme ve değerlendirme faaliyetlerini değerlendirme ölçeği: geçerlik ve güvenirlik çalışması. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi*, 41(3), 1409-1436.

ABSTRACT

The aim of this study is to develop the scale of "assessing classroom measurement and evaluation activities" and to determine its psychometric properties. Since the research is aimed at revealing an existing situation, it has been done by survey model, one of the quantitative research approaches. During the development of the scale, the literature was reviewed and a draft form consisting of 56 items was prepared. Expert opinions were taken for the content validity of the measuring tool and it was determined that six items did not have sufficient scope and were excluded from the study. The participants in the research consist of 400 teacher candidates studying in various departments. In order to provide evidence for the construct validity of the measurements obtained from the developed measurement tool, an exploratory factor analysis was performed and a measurement tool consisting of three factors and 40 items was obtained. Cronbach alpha stratified Cronbach alpha and McDonald (ω) coefficients were calculated for the proof of the reliability of the measurements obtained from the measurement tool and the reliability was found to be high. In order to determine the standard of the measurement tool, two-stage cluster analysis was performed and three clusters, low, medium and high, were determined. As a result, it can be said that the measurement results obtained using the developed measurement tool are reliable and the inferences made based on these results are valid.

Keywords: Classroom measurement and evaluation, Distance education, Scale development, Reliability, Validity.

GİRİŞ

Ölçme ve değerlendirme faaliyetlerinin eğitim alanındaki önemi son yıllarda gittikçe artmaya başlamıştır (Arastaman, Yıldırım ve Daşçı, 2015; Şata, 2016). Çünkü, eğitim programının hedeflerine ne ölçüde ulaşıldığının, öğrencilerin başarısının, öğrenci öğrenmelerinin zayıf ve güçlü yönlerinin ve öğretme-öğrenme sürecinin etkililiğinin belirlenmesinde sınıf içi ölçme ve değerlendirme faaliyetlerinden faydalanılmaktadır (Güler ve Gelbal, 2010). Sınıf içi ölçme ve değerlendirme faaliyetlerinin etkililiği de öğretmenlerin veya öğretim elemanlarının bu faaliyetleri uygulama düzeyleri ve mesleki yeterlikleri ile yakından ilişkilidir (Smith ve Southerland, 2007). Bununla beraber öğretmenlerin veya öğretim elemanlarının sınıf içi ölçme ve değerlendirme faaliyetlerini uygulama düzeyleri de öğrencilerin öğrenme durumlarını doğrudan etkilemektedir (Boud, 1995). Bu bağlamda öğretmenlerin veya öğretim üyelerinin sınıf içinde uyguladıkları ölçme ve değerlendirme faaliyetlerinin değerlendirilmesi eğitim ve

öğretim süreci açısından önem arz etmektedir. Nitekim, ölçme ve değerlendirme bilgi ve becerisi öğretmenlerin sahip olması gereken önemli mesleki yeterliklerinden biridir.

Öğretmenlerin sahip olduğu mesleki yeterlikler, öğrenci ile etkileşimlerinde önemli bir faktördür (Çakan, 2004; Gelbal ve Kelecioğlu, 2007; Kilmen ve Çıkrıkçı-Demirtaşlı, 2009). Uluslararası ve ulusal eğitim sistemindeki değişimler ve yeniliklerden dolayı Milli Eğitim Bakanlığı (MEB) 2006 yılında yürürlüğe koyduğu “öğretmenlik mesleği genel yeterliklerini” (MEB, 2006) 2017 yılında yeniden güncellemiştir. Öğretmenlik mesleğinin genel yeterliklerine yönelik yapılan her güncellemede ölçme ve değerlendirme yeterlik alanında da önemli değişiklikler yapılmıştır. Bunun temel nedeni; ölçme ve değerlendirme faaliyetlerinin eğitim öğretim sürecindeki değişimler ile paralellik göstermesidir (Shepard, 2005). Öğretmen yeterliklerini güncelleme çalışmalarında, her bir öğretmenlik alanı için alan yeterlikleri belirlemek yerine genel yeterliklere alan eğitimi ve alan bilgisi eklenerek bütünsel ve tek bir yeterlik metni oluşturulmuştur. Bu bağlamda güncellenmiş olan öğretmenlik mesleğinin yeterlikleri “mesleki bilgi”, “mesleki beceri” ve “tutum ve değerler” olmak üzere birbirini tamamlayan üç yeterlik alanı, bu alanlar altında yer alan 11 yeterlik ve bu yeterliklere ilişkin 65 performans göstergesinden oluşmaktadır. Güncellenen öğretmenlik mesleği yeterlikleri incelendiğinde, tanımlanan 11 yeterlikten biri de mesleki beceri genel yeterliğin altındaki ölçme ve değerlendirme yeterlik alanıdır. Bu yeterlikler içinde ölçme ve değerlendirme yeterlik alanına ilişkin beş performans göstergesinin olduğu görülmektedir (MEB, 2017).

Ölçme ve değerlendirme yeterlik alanına ilişkin belirlenen beş göstergenin etkili ve verimli bir biçimde gerçekleştirilebilmesi öğretmenlerin/eğitmenlerin bu yeterlik alanındaki donanımları ne kadar kazandığıyla birebir ilişkilidir (Aktaş ve Alicı, 2012; Çakan, 2004; Nartgün, 2008). Öğretmenlerin veya öğretim elemanlarının bu donanımlara sahip olup olmadığının belirlenmesi eğitim ve öğretim süreci içindeki öğrenciler için oldukça önemlidir (Abell ve Siegel, 2011; Xu ve Brown, 2016). Bu bağlamda eğitim öğretim sürecinde yapılan sınıf içi ölçme ve değerlendirme

faaliyetlerinin hangi düzeyde yapıldığı ve ölçme ve değerlendirme faaliyetlerinden hangisinin tercih edildiğinin belirlenmesi önemli görülmektedir.

Son bir yıllık eğitim öğretim sürecine bakıldığında, tüm dünyada olduğu gibi ülkemizde de eğitim ve öğretimin tüm kademeleri yüz yüze eğitimden uzaktan veya çevrim içi eğitime geçmiştir. Bu bağlamda mevcut çalışmada ayrıca sınıf içi ölçme ve değerlendirme faaliyetlerinin değerlendirilmesi sürecinde uzaktan eğitim boyutu da dâhil edilmiştir. Öğrenme-öğretme süreci yüz yüze ya da uzaktan olması ölçmenin temel odağını değiştirmemekle birlikte bu süreçte kullanılan farklı öğretim yöntem ve teknolojileri ölçme ve değerlendirme yaklaşımlarını da etkilemektedir. Nitekim, teknolojiye hızlı değişimden etkilenen alanlardan biri de ölçme ve değerlendirmedir (Kaya ve Tan, 2014). Uzaktan eğitim sürecinde kullanılan çevrim içi sistemler aracılığıyla öğrencilerin eğitime erişilebilirlik ve öğretim programında esneklik gibi temel avantajları ön plana çıkmakta iken, öğretmen ve öğrenci etkileşiminin azalması ve bireyselleşme gibi dezavantajlar da karşımıza çıkmaktadır. Özellikle uzaktan eğitim sürecinde ölçme ve değerlendirme faaliyetleri daha fazla önem kazanmıştır. Alanyazın incelendiğinde uzaktan eğitim sürecinde geleneksel ölçme ve değerlendirme yöntemlerinin yanı sıra tamamlayıcı ölçme ve değerlendirme faaliyetlerine de başvurulduğu görülmektedir. Geleneksel ölçme ve değerlendirme faaliyetlerinde sıklıkla çevrim içi sınavların kullanıldığı ve bu sınavların çoktan seçmeli, doğru-yanlış veya boşluk doldurma şeklinde olduğu görülmekte iken (Costa, Mullan, Kothe ve Butow, 2010; Gikandi, Morrow ve Davis, 2011; Marriott, 2009) tamamlayıcı ölçme ve değerlendirme faaliyetlerinde ise e-portfolio, kavram haritaları, projeler, iş birlikli çalışmalar, ödevler, performans görevleri, öz değerlendirme, akran değerlendirme ve çevrim içi tartışmaların yapıldığı görülmektedir (Ada, Suna, Elkonca ve Karakaya, 2016; Kear, Donelan ve Williams, 2012; Şata ve Karakaya, 2020; Terzis, Moridis ve Economides, 2012).

Sınıf içi ölçme ve değerlendirme faaliyetlerine ilişkin yapılan araştırmalar incelendiğinde; daha çok öğretmenlerin/öğretmen adaylarının tutumlarının (Aktaş ve Alıcı, 2012; Çalışkan ve Yazıcı, 2013; Erdoğan, 2010; Yaşar, 2014), yeterliklerinin

(Akdağ ve Ekmekçi, 2012; Çakan, 2004; Çelik ve Arslan, 2012; Erdoğan ve Kurt, 2012; Gelbal ve Kelecioğlu, 2007; Gök ve Şahin, 2009; Nartgün, 2008; Özdemir, 2010) ve uygulama yöntemlerinin (Çoruhlu, Nas ve Çepni, 2009; Özsevgeç ve Çepni, 2006; Yaşar, 2017) incelendiği görülmektedir. Bu araştırmaların genel olarak öğretmenlerin veya öğretmen adaylarının yeterlik düzeyleri ve geleneksel ya da tamamlayıcı ölçme ve değerlendirme türünden hangisinin kullanıldığına yönelik olduğu görülmektedir. Ancak öğretmenlerin veya öğretim elemanlarının sınıf içi ölçme ve değerlendirme faaliyetlerini nasıl veya hangi biçimde yürüttüğüne ilişkin bir araştırmanın olmadığı görülmektedir. Bu bağlamda mevcut çalışmada öğretim üyelerinin sınıf içi ölçme ve değerlendirme faaliyetlerinin değerlendirilmesine ilişkin bir ölçme aracının geliştirilmesi amaçlanmıştır. Alanyazın incelendiğinde öğretmen veya öğretim üyelerinin sınıf içi ölçme ve değerlendirme faaliyetlerinin öğrenci, eğitim programı, öğretim süreci ve ulaşılması beklenen hedefler için önemli olduğu görülmektedir. Bu bağlamda öğretmenlerin sınıf içi ölçme ve değerlendirme faaliyetlerinin durum belirlemesinin yapılması eğitim ve öğretim süreci için önem arz etmektedir. Buradan hareketle bu araştırmada öğretim üyelerinin sınıf içi ölçme ve değerlendirme faaliyetlerinin öğrenci/öğretmen adayı görüşlerine göre değerlendirilmesine ilişkin bir ölçme aracının geliştirilmesi önemli görülmektedir.

Araştırmanın Amacı

Bu çalışmanın amacı, “sınıf içi ölçme ve değerlendirme faaliyetlerini değerlendirme” ölçeğini geliştirmek ve ölçeğin psikometrik özelliklerini belirlemektir.

YÖNTEM

Araştırmanın Modeli

Bu araştırmada “Sınıf İçi Ölçme ve Değerlendirme Faaliyetlerini Değerlendirme” ölçeğinin geliştirilmesi ve psikometrik özelliklerinin belirlenmesi amaçlanmıştır. Araştırma bu yönüyle temelde bir tarama çalışmasıdır. Tarama modeli var olan durumu olduğu gibi betimlemeyi amaçlayan bir araştırma yaklaşımıdır (Karasar, 2012).

Ölçeğin Geliştirilme Süreci

Madde havuzunun oluşturulması

Araştırma kapsamında madde havuzu oluştururken, sınıf içi ölçme ve değerlendirme süreçlerinde uygulanan ya da uygulanması gereken standartlar belirlenmeye çalışılmıştır. Bu amaçla ilgili alanyazında yer alan çalışmalar detaylı şekilde incelenmiştir. Buna ek olarak farklı ulusal ve uluslararası değerlendirme kuruluşlarının sınıf içi ölçme ve değerlendirme standartları incelenmiş ve bu standartlardan yola çıkarak 56 maddeden oluşan madde havuzu oluşturulmuştur (Brookhart, 2011; Joint Committee for Standards on Educational Evaluation [JCSEE], 2015; MEB, 2017; Rundle, 2016; White, 2009). Madde havuzu oluşturulurken, ölçme ve değerlendirme temel kavramları, ölçme araçlarının taşınması gereken temel özellikler, değerlendirme yaklaşımları, gibi başlıklar dikkate alınarak maddeler yazılmıştır. Likert tipi dereceleme türünde geliştirilmesi amaçlanan ölçeğin kategori sayısı ölçülen özellik, katılımcıların yaş düzeyleri ve madde ifadeleri de dikkate alınarak “hiçbir zaman”, “nadiren”, “bazen”, “çoğunlukla” ve “her zaman” şeklinde 5 kategori olarak derecelendirilmiştir.

Uzman görüşünün alınması

Sınıf içi ölçme ve değerlendirme faaliyetlerini değerlendirme ölçeğinin (SİÖDFD) taslak formu için yazılan maddelerin, ölçülmesi amaçlanan yapıya uygunluğunun, ifadelerin açıklığının, anlaşılabilirliğinin ve ilgili kapsamı iyi düzeyde temsil edip etmediğinin belirlenebilmesi amacıyla uzman görüşü alınmıştır. Bu kapsamda Ölçme ve Değerlendirme alanında uzman altı ve Türkçe Eğitimi alanında uzman iki akademisyen olmak üzere toplamda sekiz akademisyenden uzman görüşü alınmıştır. Uzmanlara, ilgili ölçme aracındaki her bir madde için; (1) gerekli, (2) gerekli ama düzeltilmeli ve (3) gereksiz şeklinde üçlü derecelendirmeye sahip bir ölçme aracı kullanılarak ölçme aracındaki maddelerin değerlendirilmesi istenmiştir. Araştırmada uzman görüşleri kapsam geçerliği olarak ele alınmıştır. Maddelerin kapsam geçerliği belirlenirken, Lawshe (1975) yaklaşımı dikkate alınmıştır. Maddelerin yeterli kapsama sahip olup olmadığının belirlenmesi için kapsam geçerlik oranı (KGO) kullanılmış olup 8 uzman

sayısı dikkate alınarak minimum (kritik) değer 0,693 ($p = 0.05$ için) olarak belirlenmiştir (Wilson, Pan ve Schumsky, 2012). Her bir madde için hesaplanan KGO değerleri incelenmiş ve taslak formda bulunan maddelerden 6 tanesi 0,693 değerinden düşük olduğu için araştırmacılar tarafından tekrar incelenmiş ve uzmanların gerekçeleri uygun görüldüğünden ilgili maddeler ölçekten çıkartılmıştır.

Ön deneme uygulaması

Ölçeğin taslak formunda yer alan 50 maddenin öğretmen adayları tarafından anlaşılabilirliği ve ölçeğin uygulanma süresinin belirlenebilmesi amacıyla ön deneme uygulaması yapılmıştır. Bu amaçla araştırmanın hedef evrenine benzer özellikler taşıyan 30 kişilik bir grupta ön deneme uygulaması yapılmıştır. Yüz yüze eğitimin olmaması nedeniyle çevrim içi olarak uygulanan ölçekte yer alan maddelerin genel olarak öğrenciler tarafından anlaşıldığı ve ortalama 15 dakikada doldurulabildiği gözlenmiştir.

Evren ve Örneklem

Ölçeğin pilot uygulaması kapsamında iki aşamalı bir örnekleme yöntemi kullanılmıştır. Araştırmanın evrenini yükseköğretimdeki öğretmen adayları oluşturduğu için ilk aşamada kolay ulaşılabilir örnekleme yöntemi ile Doğu Anadolu Bölgesindeki bir üniversitede öğrenim gören öğretmen adayları seçilmiştir. İkinci aşamada örnekleme seçilecek öğretmen adayları ölçme ve değerlendirme temel kavramlarına daha aşina olması dolayısıyla eğitim fakültesinde öğrenim gören ve ölçme ve değerlendirme dersini almış olan öğrenciler ölçüt örnekleme kullanılarak seçilmiştir. Bu amaçla eğitim fakültesinin farklı bölümlerinde öğrenim gören üçüncü ve dördüncü sınıf öğretmen adayları ile pedagojik formasyon eğitiminde ölçme ve değerlendirme dersini alan öğretmen adaylarından olmak üzere toplam 462 kişi ile ölçeğin pilot uygulaması gerçekleştirilmiştir. Veri temizleme sürecinde maddelerin çoğuna cevap vermeyen 33 ve uç değer analizi sonucunda ± 3 aralığı dışında kalan 29 öğretmen adayı analizlerden çıkartılmış olup toplam 400 öğretmen adayı ile analizler gerçekleştirilmiştir.

Kolay ulaşılabilir/uygun örnekleme yöntemi ölçek geliştirme çalışmalarında sıklıkla tercih edilen bir yöntemdir. Buna ek olarak özellikle ölçek geliştirme çalışmalarında

ölçeğin yapısı ve analizde kullanılacak yöntemin özelliğinin dikkate alınmasının doğru bir yaklaşım olacağı belirtilmektedir (Erkuş, 2012). Pilot uygulama sürecinde temel amaçlardan biri ölçülen özelliği temsil eden bir örneklem üzerinde çalışmak olduğu için araştırmada kolay ulaşılabilir örnekleme yöntemi tercih edilerek olabildiğince çok kişiye daha ekonomik ve kolay şekilde ulaşılmaya çalışılmıştır. Ayrıca ölçme sonuçları üzerinde yapılacak geçerlik ve güvenirlik analizleri de dikkate alınarak katılımcı sayısının ölçekte yer alan madde sayısının en az 5 katı olmasına dikkat edilmiştir. Nitekim yapı geçerliğine kanıt sunmak amacıyla yapılan Açıklayıcı Faktör analizinin yapılabilmesi için araştırmaya alınacak kişi sayısının ölçekte yer alan madde sayısının en az 5 katı olması gerektiği alanyazında belirtilmiştir (Büyüköztürk, 2012). Araştırma kapsamında pilot uygulama yapılan grubun bazı demografik bilgilerine ait bulgular Tablo 1’de verilmiştir.

Tablo 1. Deneme/Pilot Uygulaması Çalışma Gurubunun Seçildiği Katılımcıların Demografik Bilgilerine ait Frekans ve Yüzde Dağılımları

Değişken	Kategori	f	%
Cinsiyet	Kadın	236	59.0
	Erkek	164	41.0
Bölüm	Türk Dili ve Edebiyatı Öğretmenliği	70	17.5
	Beden Eğitimi Öğretmenliği	65	16.3
	Okul Öncesi Öğretmenliği	61	15.2
	İlköğretim Matematik Öğretmenliği	43	10.8
	Türkçe Öğretmenliği	41	10.3
	Tarih Öğretmenliği	35	8.8
	Coğrafya Öğretmenliği	29	7.2
	Resim Öğretmenliği	21	5.2
	Mütercim ve Tercümanlık	21	5.2
	Fen Bilgisi Öğretmenliği	14	3.5
Sınıf	Üçüncü sınıf	356	89.0
	Dördüncü sınıf	44	11.0
Toplam		400	100.00

Tablo 1 incelendiğinde, çalışma grubunun büyük çoğunluğunun kadın ve 3. sınıfta olan öğretmen adaylarından oluştuğu görülmektedir. Öğretmen adaylarının okudukları

bölümler incelendiğinde ise en fazla Türk Dili ve Edebiyatı bölümünden en az ise Fen Bilgisi bölümünden olduğu görülmektedir.

Etik Kurallara Uygunluk

Bu çalışmanın hem pilot hem de asıl uygulama aşaması araştırma ve yayın etiği kurallarına uygun olarak gerçekleştirilmiştir. Çalışmanın etik ilkelere uygun olduğu araştırmacıların çalıştığı kurumdan da belgelenmiştir. Çalışmaya katılan öğrencilere sorulara cevap verirken içten ve dürüst bir şekilde cevaplandırmaları, uygulama sonunda not olarak değerlendirmeye tabi tutulmayacakları ve verecekleri cevapların herhangi başka kişi ya da şahıslarla çalışmanın amacı dışında paylaşılmayacağı açık ve net bir şekilde ifade edilmiştir. Ağrı İbrahim Çeçen Üniversitesi Bilimsel Araştırmalar Etik Kurulu'nun 29.12.2020 tarihli ve 170 sayılı toplantı kararı ile araştırmanın etik açıdan uygun olduğuna karar verilmiştir. Etik kurul onayına ilişkin belge Ek 1'de sunulmuştur.

Veri Analizi

Araştırmada 50 maddeden oluşturulan ölçeğin nihai hâli pilot uygulamada öğretmen adaylarına uygulanmış ve elde edilen ölçümler üzerinden ölçeğin faktör yapısını belirlemek amacıyla faktör çıkarma yöntemlerinden Temel Bileşenler Analizi kullanılmıştır. Verinin faktör analizine uygunluğunu değerlendirmek amacıyla Kaiser-Meyer-Olkin (KMO) ve Barlett küresellik testi kullanılmıştır. Barlett Küresellik Testi'nin istatistiksel olarak anlamlı çıkması ölçekteki maddelerin ilişkili ve veri matrisinin faktör analizi için uygun olduğu ve KMO değerinin 0.50 ve üzerinde olması ise verinin faktör analizi yapılabilmesi için uygun olduğu şeklinde yorumlanmaktadır (Field, 2009; Büyüköztürk, 2012). Faktör sayısına karar vermede, faktör özdeğerleri, yamaç birikinti grafiği (Scree Plot), uzman görüşü ve maddelerin içerikleri dikkate alınmıştır. Her bir maddenin ilgili faktörle ilişkisinin bir ölçüsü olarak ele alınan faktör yükü sınır değeri olarak 0.32'nin alınması uygun görülmüştür (Tabachnick ve Fidell, 2013). Bu değer altında yüke sahip olan maddeler ölçekten çıkarılmıştır. Faktör analizi yapılırken döndürme yöntemi olarak hem eğik hem de dik döndürme yöntemleri

denenmiş ölçekte yer alan maddelerin en anlamlı olarak kümeleştiği durum varimax döndürme yöntemi ile elde edilmiştir. Nitekim Erkuş (2012) genel olarak döndürme yöntemi seçilirken alt faktörler arası ilişkilerden yola çıkılsa da ölçek geliştirme sürecinde faktörler ile ilgili tamamen önsel bilgilerden hareket etmenin hatalı sonuçlara yol açabileceği ve dolayısıyla farklı yöntemlerin denenerek ortaya çıkan yapının kavramsal yapı ile örtüşmesinin önemli olduğunu belirtmiştir. Ölçekten elde edilen ölçme sonuçlarının güvenilirliğine kanıt sunmak amacıyla, Cronbach Alpha (α) tabakalı Cronbach alfa (strat α) ve McDonald's omega (ω) güvenilirlik katsayıları hesaplanmıştır. Buna ek olarak madde ayırt ediciliğinin bir ölçüsü olarak madde toplam korelasyon katsayıları raporlanmıştır. İç tutarlık katsayısı olarak ele alınan güvenilirlik katsayısı değerinin 0.70 ve üzerinde olması ölçme sonuçlarının güvenilir olduğunu şeklinde yorumlanmakta iken (Salvucci, Walter, Conley, Fink ve Saba, 1997), madde-toplam korelasyon katsayısı değerinin ise 0.20 ve üzerinde değer alması maddenin ölçeğin geneliyle uyumlu şekilde çalıştığına işaret etmektedir (Crocker ve Algina, 2006). Araştırma kapsamında veri analizi SPSS (versiyon 22) istatistik paket programı kullanılarak yapılmıştır.

BULGULAR

Pilot uygulama ile elde edilen verilerin faktör analizi için uygunluğunun belirlenmesi amacıyla ilk olarak Kaiser-Meyer Olkin ve Barlett küresellik testi yapılmıştır. Elde edilen bulgular Tablo 2'de verilmiştir.

Tablo 2. KMO ve Barlett Testi Sonuçları

Kaiser-Meyer-Olkin (KMO)		0,94
	χ^2	10093.57
Barlett Testi	sd	1225
	<i>p</i>	0.000

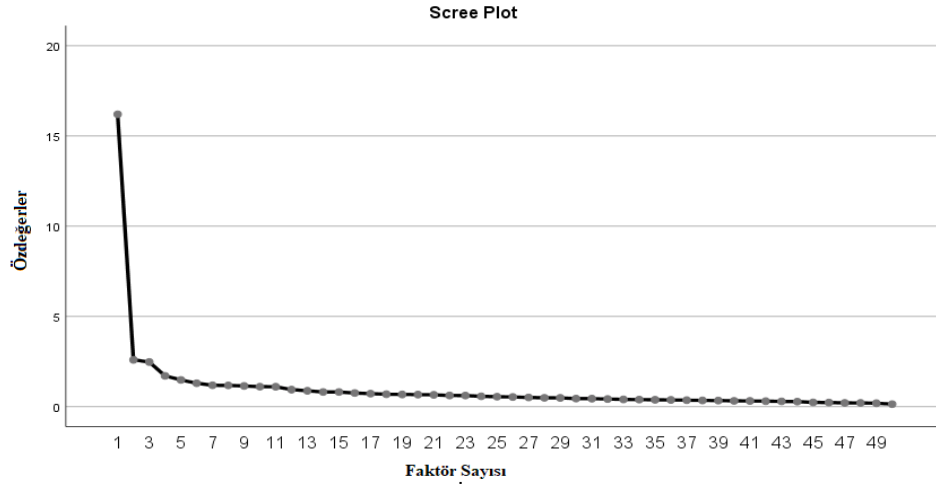
Tablo 2'de görüldüğü üzere, KMO değeri 0,94 olarak elde edilmiştir. Elde edilen bu değer verinin faktör analizi yapılabilmesi için uygun olduğunu göstermektedir (Field, 2009). Barlett küresellik testi sonuçları dikkate alındığında elde edilen değerlerin

istatistiksel olarak anlamlı olduğu görülmektedir ($\chi^2=10093.57$; $p < 0.05$). Elde edilen bu değer ölçekte yer alan maddelerin genel olarak ilişkili olduğunu göstermektedir. Buna ek olarak küresellik testinin anlamlı çıkması çok değişkenli normalliğin sağlandığı şeklinde de yorumlanmaktadır (Büyüköztürk, 2012). Pilot uygulama ile elde edilen verinin faktör yapısının ortaya çıkarılması için uygunluğu test edildikten sonra ilk olarak faktör sayısına herhangi bir sınırlama getirilmeden AFA yapılmıştır. Elde edilen bulgulara göre, ilk durumda özdeğeri 1'in üstünde olan 11 faktör oluşmuştur. Her bir faktöre ait özdeğer ve açıklanan varyans oranları Tablo 3'te verilmiştir.

Tablo 3. Özdeğer, Açıklanan Varyans ve Toplam Açıklanan Varyans Oranları

Faktörler	Öz Değer	Açıklanan Varyans (%)	Toplam Açıklanan Varyans (%)
1	16.20	32.40	32.40
2	2.59	5.18	37.58
3	2.47	4.95	42.52
4	1.70	3.40	45.93
5	1.48	2.96	48.89
6	1.30	2.59	51.48
7	1.18	2.36	53.84
8	1.18	2.35	56.19
9	1.14	2.29	58.48
10	1.11	2.21	60.69
11	1.10	2.20	62.89

Tablo 3'te görüldüğü üzere, ilk faktöre ait özdeğer 16.20 ve bu faktörün tek başına açıkladığı varyans oranı %32.40'tır. 11 faktörün açıkladığı toplam açıklanan varyans oranı yaklaşık olarak %63'tür. Faktörlere ait özdeğerler incelendiğinde özellikle üçüncü faktörden sonra özdeğerdeki düşüş miktarlarının azaldığı görülmektedir. Faktör sayısına karar vermek amacıyla ek olarak özdeğerlere dayalı olarak oluşturulan yamaç birikinti (Scree Plot) grafiği de incelenmiştir.



Şekil 1. Faktör Özdeğerlerine Ait Yamaç Birikinti Grafiği (Scree Plot)

Şekil 1'de görüldüğü üzere, özdeğerler arası en yüksek kırılmanın ilk faktör ile ikinci faktör arasındadır. Özellikle üçüncü faktörden sonra özdeğerler arası düşüş miktarının oldukça yakın olduğu görülmektedir. Buna ek olarak üç faktörlü yapıya kanıt sağlamak amacıyla paralel analiz yapılmıştır. Paralel analiz sonuçlarına göre %95 güven aralığında üç faktörlü yapının uygun olduğu görülmüştür. Modele ait uyum değerleri; CFI = 0.96, GFI = 0.98, AGFI = 0.97, RMSEA = 0.07 ve NNFI = 0.95 olarak elde edilmiş olup kurulan üç faktörlü modelin iyi düzeyde uyum gösterdiği bulunmuştur (Browne & Cudeck, 1992). AFA, paralel analiz sonuçları, uzman görüşü ve ilgili kuramsal yapı göz önüne alınarak faktör sayısının üç olmasına karar verilmiştir.

Faktör sayısına karar verildikten sonra faktör sayısı üçe sabitlenerek faktör analizi tekrarlanmıştır. Döndürme yöntemi olarak varimax dik döndürme yöntemi seçilmiştir. Elde edilen bulgulara göre, faktör yük değeri 0.32'nin altında olan bir madde (M47) ve farklı faktörlerde çok yakın faktör yükü değeri olarak binişiklik gösteren yedi madde (M15, M16, M18, M19, M20, M23, M25) ve ilgili faktörde çalıştığı hâlde madde ifadesi açısından kendi faktörü ile uyumlu olduğu düşünülmeyen iki (M39, M41) madde

de uzman görüşü alınarak ölçekten çıkarılmıştır. Ölçeğin son hâli üç faktörlü bir yapıda ve 40 maddeden oluşmaktadır.

Tablo 4. Ölçme ve Değerlendirme Faaliyetlerinin Değerlendirme Ölçeğinin Son Hâline Ait Faktör Yük Değerleri, Özdeğerler ve Açıklanan Varyans Oranları

Faktör 1		Faktör 2		Faktör 3	
Madde No	Faktör Yüğü	Madde No	Faktör Yüğü	Madde No	Faktör Yüğü
M1	0.64	M21	0.52	M45	0.47
M2	0.62	M22	0.49	M46	0.58
M3	0.63	M24	0.56	M48	0.76
M4	0.60	M26	0.53	M49	0.77
M5	0.63	M27	0.63	M50	0.49
M6	0.72	M28	0.55		
M7	0.60	M29	0.52		
M8	0.59	M30	0.60		
M9	0.59	M31	0.50		
M10	0.51	M32	0.57		
M11	0.50	M33	0.56		
M12	0.58	M34	0.47		
M13	0.56	M35	0.58		
M14	0.36	M36	0.60		
M17	0.52	M37	0.62		
		M38	0.52		
		M40	0.58		
		M42	0.55		
		M43	0.54		
		M44	0.52		
Özdeğer	6.44		7.17		3.79
Açıklanan Varyans (%)	16.10		17.92		9.49
Toplam Açıklanan Varyans (%)	43.50				

Tablo 4'te görüldüğü üzere birinci faktörün özdeğeri 6.44 ve açıkladığı varyans oranı %16'dır. Bu faktöre ait faktör yükü değerleri 0.32 ile 0.72 arasında değişmektedir. İkinci faktörün özdeğeri 7.17 ve açıkladığı varyans oranı %17.92'dir. Bu faktöre ait maddelerin faktör yükü değerleri 0.47 ile 0.63 arasında değişmektedir. Üçüncü faktörün ise özdeğeri 3.79 ve tek başına açıkladığı varyans oranı %9.49'dur. Bu faktöre ait maddelerin faktör yükü değerleri 0.47 ile 0.77 arasında değişmektedir. Üç faktörün açıkladıkları toplam varyans oranı yaklaşık olarak %44'dür. Her üç faktörde yer alan madde ifadeleri ve ilgili kuramsal yapı dikkate alınarak birinci faktör "Ölçmede Temel

Kavramlar”, ikinci faktör “Kalite (güvenirlilik, geçerlik, kullanılabilirlik) ve üçüncü faktör ise “Uzaktan Eğitim” olarak isimlendirilmiştir.

Araştırmada her bir faktörün diğer faktörlerle ve toplam puan ile ilişkileri Pearson Momentler Çarpımı Korelasyon katsayısı ile hesaplanmış ve bulgular Tablo 5’te verilmiştir.

Tablo 5. Faktörler Arası İlişkilere ait Pearson Momentler Çarpımı Korelasyon Katsayısı Değerleri

		Temel Kavramlar	Kalite	Uzaktan Eğitim	Toplam
Temel Kavramlar	r	1			
	p				
Kalite	r	0.66**	1		
	p	0.000			
Uzaktan Eğitim	r	0.44**	0.53**	1	
	p	0.000	0.000		
Toplam	r	0.86**	0.93**	0.67**	1
	p	0.000	0.000	0.000	

*p<0.05; **p<0.01

Tablo 5’te görüldüğü üzere ölçeğin alt boyutları genel olarak birbirleri ile orta düzeyde bir ilişki göstermişken ölçeğin toplam puanı ile yüksek düzeyde ilişki göstermektedirler. Temel kavramlar alt boyutu ölçeğin geneli ile arasındaki korelasyonu $r=0.86$ iken, Kalite boyutundan elde edilen puanların ölçeğin geneli için hesaplanan puanlarla korelasyonu $r=0.93$ ve Uzaktan Eğitim boyutundan elde edilen puanların ölçeğin geneli için elde edilen puanlarla korelasyonu $r=0.67$ ’dir. Tüm katsayılar 0.01 düzeyinde istatistiksel olarak anlamlıdır. Ölçeğin alt boyutlarının ölçeğin geneli ile orta düzeyde ilişkili olması alt faktörlerin aynı temel yapıyı ölçtükleri şeklinde yorumlanabilir.

Tablo 6. SİÖDFD Ölçeğinin Alt Faktörleri ve Genelinden Elde Edilen Ölçme Sonuçlarına ait Güvenirlik Katsayısı Değerleri

	N*	Cronbach Alpha (α)	McDonald's omega (ω)	Tabakalı Cronbach Alpha
Temel Kavramlar	15	0.89	0.89	--
Kalite	20	0.92	0.92	--
Uzaktan Eğitim	5	0.76	0.78	--
Toplam	40	--	--	0.94

*N: madde sayısı

Araştırmada Ölçme ve Değerlendirme Faaliyetlerini Değerlendirme ölçeğinden elde edilen ölçümlerin yapı geçerliğine kanıt sunmak amacıyla AFA yapıldıktan sonra bu ölçüm sonuçlarının güvenilirliğini belirlemek için de Cronbach Alpha (α), McDonald ω güvenilirlik katsayısı değerleri hesaplanmıştır. Elde edilen bulgular Tablo 6'da verilmiştir. Bu sonuçlara göre, alt faktörler için elde edilen güvenilirlik katsayısı değerleri hem Cronbach Alpha hem de McDonald's omega (ω) yöntemleri için oldukça benzerdir. Bu sonuçlara göre güvenilirlik katsayısı değerleri 0.76 ile 0.95 arasında değişmektedir. Ayrıca ölçeğin geneli için hesaplanan tabakalı Cronbach alfa katsayısının yüksek olduğu bulunmuştur. Elde edilen bu bulgu ölçeğin geneli ve alt faktörlerinden elde edilen ölçme sonuçlarının yüksek derecede güvenilir olduğuna işaret etmektedir (Salvucci vd., 1997).

SİÖDFD ölçeğinin son hâli için madde istatistikleri ve düzeltilmiş madde-toplam korelasyon değerleri hesaplanmış ve bulgular Tablo 7'de verilmiştir.

Tablo 7. SİÖDFD Ölçeğinin Son Hâline İlişkin Madde İstatistikleri ve Düzeltilmiş Madde-Toplam Korelasyon Değerleri

Madde No	Madde Ortalaması (\bar{X})	Düzeltilmiş Madde-Toplam Korelasyonu (r)	Madde No	Madde Ortalaması (\bar{X})	Düzeltilmiş Madde-Toplam Korelasyonu (r)
M1	3.66	0.56	M28	4.41	0.34
M2	3.77	0.46	M29	3.58	0.34
M3	3.54	0.57	M30	3.87	0.51
M4	3.37	0.52	M31	3.41	0.52
M5	3.50	0.60	M32	3.55	0.65
M6	3.16	0.59	M33	3.82	0.44
M7	3.57	0.55	M34	3.51	0.52
M8	3.47	0.51	M35	3.82	0.61
M9	3.49	0.51	M36	3.56	0.56
M10	3.49	0.49	M37	3.90	0.56
M11	3.30	0.64	M38	3.24	0.59
M12	2.44	0.47	M40	3.45	0.56
M13	2.25	0.43	M42	3.77	0.57
M14	2.24	0.34	M43	3.17	0.68
M17	2.99	0.55	M44	3.37	0.64
M21	3.56	0.66	M45	3.06	0.52
M22	3.58	0.59	M46	2.96	0.32
M24	4.04	0.53	M48	2.80	0.47
M26	3.58	0.66	M49	2.83	0.47
M27	4.04	0.61	M50	3.15	0.38
Ölçeğin Geneli İçin Madde Ortalaması		: 3.41			
Ölçeğin Geneli Madde-Toplam Korelasyon değerleri		: 0.53			

Tablo 7’de görüldüğü üzere, ölçekte yer alan maddelerin ortalamaları 2.24 ile 4.41 arasında değişmektedir. Ölçeğin geneli için elde edilen madde ortalaması değeri ise 3.41 olarak elde edilmiştir. Madde toplam korelasyon değerleri dikkate alındığında 0.34 ile 0.68 arasında değiştiği ve ölçeğin geneli için hesaplanan madde-toplam korelasyonları değerlerinin ortalamasının 0.53 olduğu görülmektedir. Genel olarak Madde-toplam korelasyon katsayısı değerlerinin 0.20 ve üzerinde değer alması maddenin testin geneliyle uyumlu çalıştığı şeklinde yorumlanmaktadır (Crocker ve Algina, 2006).

Araştırma kapsamında elde edilen tüm bulgular dikkate alındığında, üç faktör ve 40 maddeden oluşan SİÖDFD ölçeğinden elde edilen ölçme sonuçlarının geçerli ve güvenilir olduğu sonucuna ulaşılmıştır.

SIÖDFD ölçme aracından elde edilen ölçümlerin güvenirliliği ve geçerliği için kanıtlar sağlandıktan sonra, ilgili ölçme aracının standart belirleme çalışması yapılmıştır. Standart belirleme için iki aşamalı kümeleme analizi kullanılmıştır. Ölçme aracının alt faktörleri arasındaki korelasyonların yüksek olmasından dolayı toplam puan alınmış ve standart belirleme çalışması toplam puan üzerinden yapılmıştır. İki aşamalı kümeleme analizi sonucunda üç küme ortaya çıkmıştır (Kümeleme kalitesi = 0.650 ve ortalama Silhouette katsayısı = 0.543). Ortalama Silhouette katsayısı -1 ile +1 aralığında değer almakta olup bir değerine yakın olması yapılan kümelemenin tutarlı olduğunu göstermektedir. Ayrıca Kümeleme kalitesinin 0.50'den büyük olması verinin iyi bir şekilde kümelendiğine işaret etmektedir (Dinh, Fujinami & Huynh, 2019). Üç küme elde edildiğinden, iki tane kesme puanı elde edilmiş olup sırasıyla 126.50 ve 155.50 şeklindedir. Buna göre ölçekten 40.00 ile 126.50 arasında puan alınması, öğretmen adayı gözünden öğretim üyesinin sınıf içi ölçme ve değerlendirme faaliyetlerini düşük düzeyde gerçekleştirdiğine işaret etmekte iken, 126.50 ile 155.50 arası puan orta düzeyde ve 155.50 üstü puan ise yüksek düzeye işaret etmektedir.

TARTIŞMA ve SONUÇ

Sınıf içi ölçme ve değerlendirme uygulamaları öğrencilerin öğrenmelerine etki eden önemli bir faktördür. Her ne kadar yapılan bu uygulamalara ilişkin öğrenci algıları ya da görüşleri genel olarak göz ardı edilse de uygulanan değerlendirme yaklaşımlarının etkililiğinin ve niteliğinin öğrenci görüşlerine göre değerlendirilmesi önemlidir (Struyven, Dochy ve Janssens, 2005). Bu açıdan sınıf içi ölçme ve değerlendirme faaliyetlerinin öğrenci görüşlerine göre değerlendirilmesi ve bu amaca yönelik olarak bir ölçme aracının geliştirilmesi bu araştırmanın temel odağını oluşturmaktadır. Bu amaçla araştırma kapsamında geliştirilen SIÖDFD ölçeği toplam 40 maddeden ve üç faktörden oluşmaktadır. İlk faktörün açıkladığı toplam varyans oranı yaklaşık olarak %16, ikinci faktörün %18 ve üçüncü faktör %10 olmak üzere bu üç faktörlü yapının açıkladığı toplam varyans oranı %44 olarak elde edilmiştir. Ölçekte yer alan maddelere ait madde toplam korelasyon değerleri 0.34 ile 0.68 arasında değişmektedir. Elde edilen

bu sonuç ölçekte yer alan maddelerin genel olarak ölçeğin geneli ile uyumlu çalıştığını göstermektedir (Crocker ve Algina, 2006). Faktörler arası ilişkiler 0.44 ile 0.66 arasında değişmekte iken faktörlerin ölçeğin toplam puanı ile olan ilişkileri 0.67 ile 0.93 arasında değişmektedir. Elde edilen sonuçlar faktörler arası ilişkilerin orta düzeyde faktörlerin ölçeğin geneli ile olan ilişkisinin ise yüksek düzeyde olduğunu göstermektedir.

Ölçekten elde edilen ölçme sonuçlarının güvenilirlik katsayısı değerleri 0.76 ile 0.95 arasında değişmektedir. Elde edilen bu sonuç ölçeğin geneli ve alt faktörlerinden elde edilen ölçme sonuçlarının yüksek derecede güvenilir olduğunu göstermektedir (Salvucci vd., 1997). 40 maddeden oluşan SİÖDFD ölçeğinden alınabilecek en yüksek puan 200 iken, en düşük puan 40'dır. Araştırma kapsamında ölçekten elde edilen ölçümlerin norm çalışması sonuçlarına göre, ölçekten 40.00 ile 126.50 arasında puan alınması, öğretmen adaylarının görüşlerine göre öğretim üyelerinin sınıf içi ölçme ve değerlendirme faaliyetlerini düşük düzeyde gerçekleştirdiğine işaret etmekte iken, 126.50 ile 155.50 arası puan orta düzeyde ve 155.50 üstü puan ise sınıf içi ölçme ve değerlendirme faaliyetlerinin yüksek düzeyde gerçekleştirildiğine işaret etmektedir.

Sınıf içinde uygulanan ölçme ve değerlendirme faaliyetlerinin örnek bir standartlar listesi gibi ele alınabilecek bu ölçme aracının ilk faktörü "ölçmede temel kavramlar" olarak isimlendirilmiştir. Bu faktörde yer alan maddeler genel olarak, dönem boyunca yapılacak ölçme ve değerlendirme faaliyetleri, uygulanacak değerlendirme yaklaşımları, dersin kazanımları ve bu kazanımları ölçmek için uygulanan değerlendirme yaklaşımları, kazanımların açık ve anlaşılabilirliği ile ilgili dönem başında öğrenciler ile paylaşılması gereken standartlardır. Nitekim JCSEE (2015) tarafından geliştirilen sınıf içi ölçme ve değerlendirme standartlarında, değerlendirme yaklaşımlarının, amaçlarının öğrencilerle sürecin başında paylaşılmasının önemini ve kullanılacak ölçme ve değerlendirme yaklaşımının öğrencilerle paylaşılmasının gerekliliğini belirtmiştir. Alanyazında temel kavramlar ile ilgili benzer standartlar farklı çalışmalarda da (Rundle, 2016; White, 2009) raporlanmıştır.

Ölçekte yer alan "kalite" faktörü genel olarak güvenilirlik ve geçerlik ile ilgili kavramlardan oluşmaktadır. Bu kapsamda, ölçme aracının anlaşılabilirliği, taşınması gereken

nitelikler, uygulama sürecinde dikkat edilmesi gereken hususlar, ölçme aracının kapsamı, maddelerin amaca uygunluğu, adilliği, yansızlığı, süreç ve sonuç değerlendirme gibi değerlendirme yöntemlerinin çeşitliliği, değerlendirme sonuçları ile ilgili geri bildirimler vb. madde içeriklerine yer verilmiştir. Sınıf içi ölçme ve değerlendirmede çok önemli bir yere sahip olan güvenilirlik ve geçerlikle ilgili dikkat edilmesi gereken hususlar ve standartlar benzer şekilde birçok rapor ve çalışmada özetlenmiştir (bkz. The American Federation of Teachers [AFT], National Council on Measurement in Education [NCME], and National Education Association developed the Standards for Teacher Competence in Educational Assessment of Students [NEA], 1990; Airasian ve Russell, 2007; Brookhart, 2011; JCSEE, 2015).

Ölçeğin üçüncü faktörü ise “uzaktan eğitim” olarak isimlendirilmiştir. Bu faktörde sınıf içi ölçme ve değerlendirme standartlarına ek olarak özellikle uzaktan eğitim sürecinde ortaya çıkabilecek göstergeler dikkate alınmıştır. Bu kapsamda, uzaktan eğitimde yapılan ölçme ve değerlendirme faaliyetlerinin güvenilirlik ve geçerliğini tehdit edebilecek özellikle kopya, uzaktan eğitimde karşılaşılabilecek teknik problemler, teknoloji kullanımı ve yetkinliği açısından dezavantajlı öğrenciler ile ilgili madde içeriklerine yer verilmiştir. Uzaktan eğitimde ölçme ve değerlendirme için standartlar genel olarak yüz yüze eğitimde benzer olmakla beraber bu çalışmada özellikle uzaktan eğitimde karşılaşılabilecek farklı ölçme ve değerlendirme standartlarına dikkat çekilmiştir. Nitekim yapılan çalışmalar, uzaktan eğitim sürecinin kendi doğasında ortaya çıkabilecek teknik problemler, teknoloji kullanım kaygısı, öğretmen ve öğrenci iletişimde ortaya çıkabilecek sorunlar gibi farklı durumları içerdiğini göstermiştir (Karal, Çebi ve Pekşen, 2010).

Araştırma kapsamında, sınıf içi ölçme ve değerlendirme süreçlerinde yapılan ölçme ve değerlendirme faaliyetlerini öğretmen adayları görüşlerine göre değerlendirmek amacıyla geliştirilen SİÖDFD ölçeğinden elde edilen ölçme sonuçlarının güvenilirliği ve geçerliği için sunulan tüm kanıtlar dikkate alındığında, ölçüm sonuçlarının tutarlı olduğu, kapsam geçerliği ve yapı geçerliğinin iyi düzeyde olduğu diğer bir ifade ile SİÖDFD ölçeğinin amaca uygun şekilde çalıştığı belirlenmiştir. Ayrıca ilgili ölçme

aracı kullanılarak elde edilen ölçme sonuçlarının güvenilir olduğu ve bu sonuçlara dayalı yapılan çıkarımların ise geçerli olduğu söylenebilir.

Araştırma sonuçlarına dayalı olarak bazı öneriler sunulmuştur;

Bu araştırmada geliştirilen ölçme aracı ile yükseköğretimdeki sınıf içi ölçme ve değerlendirme faaliyetlerinin değerlendirilmesi amaçlanmıştır. Benzer çalışmalar eğitim sisteminin diğer kademelerinde de (ilköğretim, ortaöğretim gibi) yapılabilir.

Geliştirilen bu ölçek farklı üniversitede öğrenim gören öğretmen adaylarına uygulanarak sınıf içi ölçme ve değerlendirme faaliyetlerinde eksik olan alanlar belirlenebilir.

Sınıf içi ölçme ve değerlendirme standartlarının daha genel bir formunun belirlenmesi amacıyla birçok farklı fakülte ve bölümler dikkate alınarak yeni çalışmalar yapılabilir.

KAYNAKLAR

- Abell, S. K., & Siegel, M. A. (2011). *Assessment literacy: What science teachers need to know and be able to do?* In D. Corrigan, J. Dillon, & R. Gunstone (Eds.), *The professional knowledge base of science teaching* (pp. 205-221). The Netherlands: Springer. https://doi.org/10.1007/978-90-481-3927-9_12
- Ada, S., Suna, H. T., Elkonca, F., & Karakaya, I. (2016). Views of academicians, school administrators, and teachers regarding the use of e-portfolios in transition from elementary education to secondary education. *Educational Sciences: Theory & Practice*, 16 (2), 375-397. <https://doi.org/10.12738/estp.2016.2.0148>
- Airasian, P. W., & Russell, M. K. (2007). *Classroom assessment: Concepts and applications* (6th ed.). Boston: McGraw Hill.
- Akdağ, G. & Ekmekçi, S. (2012, Haziran). *Fen ve teknoloji öğretmenlerinin ölçme-değerlendirmeye ilişkin yeterlilik alguları ve görüşleri (Adıyaman ili örneği)*. X. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi. Niğde.
- Aktaş, M., & Alici, D. (2012). Eğitimde ölçme ve değerlendirme dersine yönelik tutum ölçeği'nin (EÖD-TÖ) geliştirilmesi. *Journal of Qafqaz University, Philology and Pedagogy*, 33, 66-73.
- American Federation of Teachers, National Council on Measurement in Education, & National Education Association (1990). *Standards for teacher competence in educational assessment of students*. Washington, DC: National Council on Measurement in Education.
- Arastaman, G., Yıldırım, K., & Daşçı, E. (2015). Ölçme ve değerlendirme ölçeğinin geliştirilmesi: Geçerlilik ve güvenilirlik çalışması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 38(38), 219-228. <https://doi.org/10.9779/PUJE728>
- Boud, D. (1995). *Ensuring that assessment contributes to learning*. Proceedings. International conference on problem-based learning in higher education. University of Linköping. Sweden.
- Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: issues and practice*, 30(1), 3-12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research*, 21(2), 230-258. <https://doi.org/10.1177/0049124192021002005>
- Büyüköztürk, Ş. (2012). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem Akademi.


- Costa, D. S., Mullan, B. A., Kothe, E. J., & Butow, P. (2010). A web-based formative assessment tool for Masters students: A pilot study. *Computers & Education*, 54(4), 1248-1253. <https://doi.org/10.1016/j.compedu.2009.11.011>
- Crocker, L., & Algina, J. (2006). Introduction to classical and modern test theory. Belmont: Wadsworth Pub Co.
- Çakan, M. (2004). Öğretmenlerin ölçme-değerlendirme uygulamaları ve yeterlik düzeyleri: İlk ve ortaöğretim. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 37(2), 99-114. https://doi.org/10.1501/Egifak_0000000101
- Çalışkan, H., & Yazıcı, K. (2013). Ölçme ve değerlendirmeye yönelik tutum ölçeğinin geliştirilmesi ve sosyal bilgiler öğretmenlerinin tutum düzeylerinin çeşitli değişkenlere göre incelenmesi. *International Journal of Human Sciences*, 10(1), 398-415.
- Çelik, Z. & Arslan, Y. (2012). Aday beden eğitimi öğretmenlerin ölçme ve değerlendirme genel yeterlilik algılarının belirlenmesi. *Selçuk Üniversitesi Beden Eğitimi ve Spor Bilim Dergisi*, 14(2), 223-232.
- Çoruhlu, T. S., Nas, S. E. & Çepni, S. (2009). Fen ve Teknoloji öğretmenlerinin alternatif ölçme-değerlendirme tekniklerini kullanmada karşılaştıkları problemler: Trabzon örneği. *Yüzüncü Yıl Üniversitesi Eğitim Fakültesi Dergisi*, 6(1), 122-141.
- Dinh DT., Fujinami T., & Huynh VN. (2019). *Estimating the Optimal Number of Clusters in Categorical Data Clustering by Silhouette Coefficient*. In: Chen J., Huynh V., Nguyen GN., Tang X. (eds) Knowledge and Systems Sciences. KSS 2019. Communications in Computer and Information Science, vol 1103. Springer, Singapore. https://doi.org/10.1007/978-981-15-1209-4_1
- Erdoğdu, M. Y. (2010, Ekim). *Öğretmenlerin ölçme ve değerlendirmeye yönelik tutumlarının bazı değişkenler açısından incelenmesi*. International Conference on New Trends in Education and Their Implications. Antalya.
- Erdoğdu, M. Y., & Kurt, F. (2012). Öğretmenlerin ölçme ve değerlendirme yeterlilik algılarının bazı değişkenler açısından incelenmesi. *Elektronik Eğitim Bilimleri Dergisi*, 1(2), 23-36.
- Erkuş, A. (2012). *Psikolojide ölçme ve ölçek geliştirme*. Ankara: Pegem Akademi Yayınları.
- Field, A. (2009). *Discovering Statistics Using SPSS*. London: Sage Publications Ltd.
- Gelbal, S., & Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33(33), 135-145.
- Gikandi, J. W., Morrow, D., & Davis, N. E. (2011). Online formative assessment in higher education: A review of the literature. *Computers & Education*, 57(4), 2333-2351. <https://doi.org/10.1016/j.compedu.2011.06.004>


- Gök, B., & Şahin, A.E. (2009). İlköğretim 4. Ve 5. Sınıf öğretmenlerinin değerlendirme araçlarını çoklu kullanım ve yeterlik düzeyleri. *Eğitim ve Bilim*, 34 (153), 127-143.
- Güler, N., & Gelbal, S. (2010). Açık uçlu matematik sorularının güvenilirliğinin klasik test kuramı ve genellenebilirlik kuramına göre incelenmesi. *Kuram ve Uygulamada Eğitim Bilimleri*, 10 (2), 991-1019.
- Joint Committee for Standards on Educational Evaluation [JCSEE]. (2015). Classroom assessment standards: Practices for PK-12 teachers. Retrieved from <http://www.jcsee.org/the-classroomassessment-standards-new-standards>.
- Karal, H., Çebi, A. & Pekşen, M. (2010). Student opinions about the period of measurement and evaluation in distance education: the difficulties. *Procedia-Social and Behavioral Sciences*, 9, 1597-1601. <https://doi.org/10.1016/j.sbspro.2010.12.371>
- Karasar, N. (2012). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayın Dağıtım.
- Kaya, Z., & Tan, Ş. (2014). New trends of measurement and assessment in distance education. *Turkish Online Journal of Distance Education*, 15(1), 206-217. <https://doi.org/10.17718/tojde.30398>
- Kear, K., Donelan, H., & Williams, J. (2014). Using wikis for online group projects: Student and tutor perspectives. *The International Review of Research in Open and Distributed Learning*, 15(4), 70-90. <https://doi.org/10.19173/irrodl.v15i4.1753>
- Kilmen, S., & Çıkrıkçı-Demirtaşlı, N. (2009). Sınıf öğretmenlerinin ölçme ve değerlendirme ilkelerini uygulama düzeylerine ilişkin görüşleri. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 42(2), 27-54. https://doi.org/10.1501/Egifak_0000001187
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel psychology*, 28(4), 563-575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Marriott, P. (2009). Students' evaluation of the use of online summative assessment on an undergraduate financial accounting ödüle. *British Journal of Educational Technology*, 40, 237-254. <https://doi.org/10.1111/j.1467-8535.2008.00924.x>
- MEB. (2006). Temel eğitime destek projesi "öğretmen eğitimi bileşeni" öğretmenlik mesleği genel yeterlikleri. Ankara: Öğretmen yetiştirme ve geliştirme genel müdürlüğü. Erişim adresi: http://web.deu.edu.tr/ilyas/ftp/ogretmenlik_meslegi_genel_yeterlikleri_2006.pdf
- MEB. (2017). *Öğretmenlik mesleği genel yeterlikleri*. Ankara: Öğretmen yetiştirme ve geliştirme genel müdürlüğü. Erişim adresi:

- http://oygm.meb.gov.tr/meb_iys_dosyalar/2017_12/11115355_YYRETMENTLYK_MESLEY_YENEL_YETERLYKLERI.pdf
- Nartgün, Z. (2008). Öğretmen adayları için ölçme ve değerlendirme genel yeterlilik algısı ölçeği: geçerlilik ve güvenilirlik çalışması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 8(2), 85-94.
- Özdemir, S. M. (2010). İlköğretim öğretmenlerinin alternatif ölçme ve değerlendirme araçlarına ilişkin yeterlikleri ve hizmet içi eğitim ihtiyaçları. *Türk Eğitim Bilimleri Dergisi*, 8(4), 787-816.
- Özsevgeç, T., & Çepni, S. (2006). Relation between science teachers' assessment tools and students' cognitive development, *Educational Research and Reviews*, 1(7), 222-226.
- Rundle, N. (2016). *Guidelines for good assessment practice* (3rd ed.). Hobart, Tasmania, Australia: Tasmanian Institute of Learning and Teaching, University of Tasmania. Retrieved from https://www.teaching-learning.utas.edu.au/__data/assets/pdf_file/0004/158674/Good-Assessment-Practice.pdf.
- Salvucci, S., Walter, E., Conley, V., Fink, S., & Saba, M. (1997). *Measurement Error Studies at the National Center for Education Statistics*. Retrieved from <https://files.eric.ed.gov/fulltext/ED410313.pdf>
- Shepard, L. A. (2005). Linking formative assessment to scaffolding. *Educational Leadership*, 63(3), 66-70.
- Smith, L., & Southerland, S. A. (2007). Reforming practice or modifying reforms? Elementary teachers' response to the tools of reform. *Journal of Research in Science Teaching*, 44(3), 396-423. <https://doi.org/10.1002/tea.20165>
- Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment & Evaluation in Higher Education*, 30(4), 325-341. <https://doi.org/10.1080/02602930500099102>
- Şata, M. (2016). Türk eğitim sistemi'nde sınıf içi ile geniş ölçekli ölçme ve değerlendirmeye genel bir bakış. *Current Research in Education*, 2(1), 53-60.
- Şata, M., & Karakaya, İ. (2020). Investigation of the use of electronic portfolios in the determination of student achievement in higher education using the many-facet Rasch measurement model. *Educational Policy Analysis and Strategic Research*, 15(1), 7-21. <https://doi.org/10.29329/epasr.2020.236.1>
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics*. Boston: Pearson.
- Terzis, V., Moridis, C. N., & Economides, A. A. (2012). The effect of emotional feedback on behavioral intention to use computer based assessment. *Computers & Education*, 59(2), 710-721. <https://doi.org/10.1016/j.compedu.2012.03.003>

- White, E. (2009). Are you assessment literate? Some fundamental questions regarding effective classroom-based assessment. *OnCUE journal*, 3(1), 3-25.
- Wilson, F. R., Pan, W., & Schumsky, D. A. (2012). Recalculation of the critical values for Lawshe's content validity ratio. *Measurement and Evaluation in Counseling and Development*, 45(3), 197-210.
<https://doi.org/10.1177/0748175612440286>
- Xu, Y., & Brown, G. T. (2016). Teacher assessment literacy in practice: A reconceptualization. *Teaching and Teacher Education*, 58, 149-162.
<https://doi.org/10.1016/j.tate.2016.05.010>
- Yaşar, M. (2014). Eğitimde ölçme ve değerlendirme dersine yönelik tutum ölçeğinin geliştirilmesi. *Eğitim Bilimleri Araştırmaları Dergisi*, 4(1), 259-279.
<https://doi.org/10.12973/jesr.2014.41.13>
- Yaşar, M. D. (2017). Prospective science teachers' perception related to formative assessment approaches in Turkey. *Journal of Education and Training Studies*, 5(4), 29-43. <https://doi.org/10.11114/jets.v5i4.2190>

ORCID

Mehmet ŞATA  <https://orcid.org/0000-0003-2683-4997>

Fuat ELKONCA  <https://orcid.org/0000-0002-2733-8891>

SUMMARY

Introduction

The importance of measurement and evaluation activities in the field of education has started to increase in recent years (Arastaman, Yıldırım, & Daşçı, 2015; Şata, 2016). Because, classroom assessment and evaluation activities are used to determine the extent to which the objectives of the curriculum have been achieved, the success of the students, the strengths and weaknesses of student learning, and the effectiveness of the teaching-learning process (Güler & Gelbal, 2010). Teachers' or lecturers' classroom measurement and evaluation activities directly affect students' learning (Boud, 1995).

Assessment and evaluation activities have gained more importance, especially in the distance education process. When the literature is examined, it is seen that in the distance education process, complementary measurement and evaluation activities are used as well as traditional measurement and evaluation methods. It is seen that in online exams, traditional assessment and evaluation activities frequently used and these exams are in the form of multiple choice, true-false or gap-filling (Costa, Mullan, Kothe, & Butow, 2010; Gikandi, Morrow, & Davis, 2011; Marriott, 2009). On the other hand, it is seen that complementary assessment and evaluation activities such as e-portfolio, concept maps, projects, collaborative studies, assignments, self-assessment, peer assessment, online discussions are being used (Ada, Suna, Elkonca, & Karakaya, 2016; Kear, Donelan, & Williams, 2012; Şata and Karakaya, 2020; Terzis, Moridis and Economides, 2012).

When the literature is examined, it is seen that the classroom assessment and evaluation activities of teachers or lecturers are important for the students, the curriculum, the teaching process and the expected goals. In this context, determining the assessment of teachers' classroom assessment and evaluation activities is important for the teaching-learning process. Therefore, in this study, it is important to develop a measurement tool to evaluate the classroom assessment and evaluation activities of faculty members according to teacher candidates' opinions.

Method

In this study, it was aimed to develop the "Assessing of Classroom Assessment and Evaluation Activities" scale and to determine its psychometric properties. In this respect, the research is basically a survey study. Survey model is a research approach that aims to describe the current situation as it is (Karasar, 2012).

While creating an item pool within the scope of the research, it was tried to determine the standards that were or should be applied classroom assessment and evaluation processes. For this purpose, the studies in the relevant literature were examined in detail. In addition, classroom assessment and evaluation standards of different national and international assessment institutions were examined and an item pool consisting of 56 items was created based on these standards. For the draft form of the scale for evaluating the classroom assessment and evaluation activities of faculty members, expert opinions were obtained from a total of eight academicians, six of whom are experts in the field of Assessment and Evaluation, and two academics who are experts in the field of Turkish Education.

Results

EFA (Exploratory Factor Analysis) was performed to provide evidence for the construct validity of the measurements obtained from the scale. Considering the EFA results, expert opinion and the relevant theoretical structure, it was decided that the number of factors was three. Considering the item expressions in all three factors and the related theoretical structure, the first factor was named as "Basic Concepts in Measurement", the second factor as "Quality (reliability, validity, usefulness)" and the third factor as "Distance Education".

Conclusions

As a result, considering all the evidence for the reliability and validity of the measurement results obtained from the SİÖDFD scale, which was developed in order to evaluate faculty members' assessment and evaluation activities carried out in the classroom assessment and evaluation processes according to the opinions of the teacher candidates, the measurement results are consistent, the content validity and construct validity are at a good level. In other words, it was determined that the SİÖDFD scale was working properly. In addition, it can be said that the measurement results obtained using the relevant measurement tool are reliable and the inferences made based on these results are valid.

Ek 1: Etik Kurul Onay Sayfası**AĞRI İBRAHİM ÇEÇEN ÜNİVERSİTESİ BİLİMSEL ARAŞTIRMALAR ETİK KURULU**

Tarih: 29.12.2020

Sayı: 170

KONU: Eğitim Fakültesi Dekanlığı 27.11.2020 tarih ve E.21703 sayılı yazısı

KISACA ÖZET:

Eğitim Fakültesi Dekanlığının ilgi yazısına istinaden Dr. Öğr. Üyesi Mehmet ŞATA ve Dr. Öğr. Üyesi Fuat ELKONCA'nın "**Ölçme ve Değerlendirme Faaliyetlerini Değerlendirme Ölçeği: Geçerlik ve Güvenirlilik Çalışması**" isimli bilimsel araştırması için kurulumuzdan izin istenmiş olup, araştırma dosyası belirtilen yazı ekinde kurulumuza gönderilmiştir.

KONU İLE İLGİLİ KİŞİLER:

Dr. Öğr. Üyesi Mehmet ŞATA ve Dr. Öğr. Üyesi Fuat ELKONCA

KONU İLE İLGİLİ YARARLANILAN VERİLER / KAYNAKLAR / DOKÜMANLAR

1. Eğitim Fakültesi Dekanlığı 27.11.2020 tarih ve E.21703 sayılı yazısı
2. Ağrı İbrahim Çeçen Üniversitesinin Bilimsel Araştırmalar Etik Kurulu Yönergesi
3. Ağrı İbrahim Çeçen Üniversitesi Etik İlkeleri
4. İnsan Hakları Evrensel Bildirgesi,
5. Helsinki Bildirgesi,
6. Dünya Hekimler Birliğinin ilke Bildirgeleri,
7. Amerikan Kimya Derneği (ACS) ilkeleri
8. Amerikan Psikologlar Derneği'nin (APA) Deontoloji ilkeleri,
9. TÜBİTAK Araştırma - Yayın Etiği,
10. T.C. Anayasası, Yasalar ve ilgili mevzuat

Yapılan etik kurul toplantısı sonuçları;

1. Dr. Öğr. Üyesi Mehmet ŞATA ve Dr. Öğr. Üyesi Fuat ELKONCA'nın araştırmacılığını yaptığı araştırma dosyası kurul üyelerine gönderilmiştir. Söz konusu dosya kurul üyeleri tarafından incelenerek, Ağrı İbrahim Çeçen Üniversitesi Bilimsel Araştırmalar Etik Kurulu 29.12.2020 tarihi saat 13:30'da araştırma ile ilgili başvuruyu görüşmek üzere Prof. Dr. Güray OKYAR başkanlığında toplanmıştır.
2. Dr. Öğr. Üyesi Mehmet ŞATA ve Dr. Öğr. Üyesi Fuat ELKONCA'nın araştırmacılığını yaptığı araştırma dosyası Ağrı İbrahim Çeçen Üniversitesi Bilimsel Araştırmalar Etik Kurulu Yönergesi madde 3, 9, 10, 11, 12 ve 13'e göre incelenmiş olup, yapılacak araştırmaya izin verilmesine mevcut oy birliğiyle karar verilmiştir.

