

Ölçme ve Değerlendirme Ölçeğinin Geliştirilmesi: Geçerlilik ve Güvenirlik Çalışması*

Gökhan ARASTAMAN** Kamil Yıldırım*** Elif DAŞCI****

Öz

Bu çalışmanın amacı, ilkokul ve ortaokullarda görev yapan öğretmenlerin ölçme ve değerlendirmeye yönelik tutumlarını belirleyen bir ölçek geliştirerek, geçerlik ve güvenirliliğini saptamaktır. 36 maddeden oluşan ölçek Aksaray il merkezindeki ilkokul ve ortaokullardan basit seçkisiz yöntemle belirlenen 166 branş öğretmenine uygulanmış, deneme uygulamasından elde edilen verilerle Açıklayıcı Faktör Analizi (AFA) yapılarak, yapı geçerliği istatistiki olarak test edilmiştir. Ölçeğin faktör analizine uygunluğunu belirlemek için öncelikle "Kaiser-Meyer-Olkin (KMO) katsayısı" ve "Barlett Sphericity Testi" yapılmış ve ölçeğin faktör analizi için uygun olduğu görülmüştür. Deneme uygulamasına ait analiz sonuçları dikkate alınarak 15 madde ölçekten çıkarılmış, kalan 21 madde ile ana uygulamaya geçilmiştir. Aksaray İli merkez ilkokul ve ortaokullarda görev yapan ve pilot uygulamaya katılmamış, tabakalı örnekleme yöntemiyle seçilen 301 branş öğretmeni ana uygulamaya katılmıştır. AFA ile elde edilen bulgular incelendiğinde, kalan 21 maddenin toplam varyansın % 67'sini açıkladığı ve maddelerin 4 faktör altında toplandığı görülmüştür. Ölçeğin faktör yükleri .82 ile .52, madde toplam korelasyonları ise .73 ile .44 arasında sıralanmaktadır. Faktör analizi sonucunda elde edilen veriler, ölçeğin geçerliğinin yüksek olduğunu göstermiştir. Ölçeğin güvenirliliğini tespit etmek amacıyla Cronbach Alfa güvenirlilik katsayısı hesaplanmış; ölçeğin tamamı için iç tutarlılık katsayısı .94 olarak bulunmuştur. Ölçeğin geçerlilik kanıtlarını artırmak amacıyla yapılan birinci ve ikinci düzey DFA sonucunda, ölçeğin iyi uyum değerleri ortaya koyduğu ve bir model olarak doğrulandığı tespit edilmiştir. (RMSEA=.06, NFI=.98, CFI=.99, GFI=.82, AGFI= .77). Yapılan analizler, geliştirilen Öğretmenlerin Ölçme ve Değerlendirmeye İlişkin Tutumu Ölçeğinin yüksek derecede geçerli ve güvenilir bir ölçme aracı olduğunu ve tutumla ilgili çalışmalarda kullanılabileceğini göstermektedir.

Anahtar Sözcükler: Ölçme ve değerlendirme, tutum, ölçek geliştirme.

Development of Measurement and Evaluation Scale: A Study of Reliability and Validity

Abstract

The aim of this study was to develop "Attitude toward Measurement and Assessment Scale" for primary and secondary school teachers. The scale consists of 36 items were conducted to randomly selected 166 subject teachers employed in the primary and secondary schools of Aksaray province. After obtaining pilot study data, exploratory factor analysis were performed and construct validity was tested accordingly. Sampling adequacy of the data for factor analysis was tested with the Kaiser-Meyer-Olkin (KMO) coefficient and Bartlett test of sphericity and it indicated that data group is coherent with factor analysis. By considering the pilot study result, 15 items were removed from the scale and study was applied with the remaining 21 items. A total of 301 primary and secondary subject teachers drawn via stratified sampling method and who were not involved in the pilot study were surveyed. Results of exploratory factor analysis showed that the 21 items loaded on four factors. The total variance explained was 67 % and the factor loadings ranged

*Bu çalışma 9-13 Haziran 2014 tarihlerinde Hacettepe Üniversitesi'nde düzenlenen IV. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongre'sinde sözlü bildiri olarak sunulmuştur.

**Yrd. Doç. Dr. , Hacettepe Üniversitesi, Üniversitesi Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Ankara. e-posta: gokhanarastaman@gmail.com

*** Yrd. Doç. Dr., Aksaray Üniversitesi Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Aksaray. e-posta: kamilyildirim1@gmail.com

**** Ar. Gör. Aksaray Üniversitesi Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Aksaray. e-posta: elifdasci89@gmail.com

between .52 to .82 and corrected item-total correlations ranged from .44 to .73. The internal consistency reliability coefficient (Cronbach's Alpha) was found as .94. On the other hand the confirmatory model for the scale indicated close model fit (RMSEA=.06, NFI=.98, CFI=.99, GFI=.82, AGFI=.77). These results revealed that the scale provide reliable and valid evidences and suggest it may be useful in research applications related with attitude.

Keywords: *Measurement and assessment, attitude, scale development.*

Giriş

Ölçme ve değerlendirmenin eğitim alanındaki önemi son yıllarda daha çok vurgulanmaya başlanmıştır. Bir eğitim programının işlevselliğinin saptanmasında, eğitim yöntem ve stratejilerinin etkililik düzeylerinin belirlenmesinde, öğrenci başarısının, öğrencilerin öğrenmedeki güçlü ve zayıf yönlerinin tespit edilmesinde ölçme ve değerlendirmeden yararlanılmaktadır. Dolayısı ile eğitim sisteminin girdi, süreç ve ürün boyutlarının tümünde ölçme ve değerlendirmeye ihtiyaç duyulduğu söylenebilir (Güler ve Gelbal, 2010).

Ölçme ve değerlendirme, öğrencilerin ne öğrendikleri hakkında bilgi ve kanıt toplama süreci olarak tanımlanmaktadır (Chen 2003; Wishon vd. 1998; Akt: Şahin ve Karaman, 2013). Ölçme, bir özelliğin gözlenerek gözlem sonuçlarının sayı ya da sembollerle ifade edilmesi olarak tanımlanırken, değerlendirme daha geniş kapsamlı bir süreç olup; ölçüm, ölçüt ve karar verme basamaklarından oluşur (Baykul, 2000). Öğretmenlerin öğrenci başarısı ve performansına ilişkin ölçme ve değerlendirme işlemlerini güvenilir ve geçerli bir şekilde yapabilmeleri için ölçme ve değerlendirme konusunda yeterli donanıma sahip olmaları gereklidir (Özbaşı, 2009).

Türkiye'de 2005 yılında gerçekleştirilen ve sonraki yıllarda güncellemeleri devam eden ders müfredatlarındaki değişimin önemli bir bölümü, ölçme-değerlendirme alanında gerçekleştirilmiştir. Değişim ve beraberinde getirdiği uygulamalar, öğretmenlerin ölçme-değerlendirme alanına ilişkin bilgi ve becerilerini geliştirmeye yöneltmiştir. 2008 yılında yayınlanan Öğretmenlik Mesleği Genel Yeterlikleri de öğretmenlerin ölçme-değerlendirmenin hangi boyutlarında yeterliklere sahip olmaları gerektiğini ortaya koymuş, yapılmakta olan hizmetiçi eğitim

faaliyetleri arasında ölçme-değerlendirmeyle ilgili olanların sayısı, süresi artmaya başlamıştır (MEB, 2008). Ayrıca Temel Eğitimden Ortaöğretime Geçiş (TEOG) sistemi gereği süreç içinde öğrenci başarısını belirlemek daha önemli hale gelmiş ve Ulusal Düzeyli Merkezi Sınav uygulamaları ölçme-değerlendirme çerçevesinde öğretmenlere ek bir sorumluluk yüklemiştir.

Eğitimde ölçme-değerlendirmeyle ilgili tüm bu gelişmeler, öğretmenlerin mesleki yaşantılarında daha çok yer alma eğilimindedir. Öğretmenlerin yeterli düzeyde ölçme ve değerlendirme bilgi ve becerileriyle donanık olması onların sahip olması gereken önemli bir özelliktir (Zhang ve Burry-Stock, 2003). Öğretmenlerin ölçme ve değerlendirme konusundaki bilgileri ve öğretmen adaylarına bu konuda ne öğretilmesi gerektiği konusu öteden beri tartışılmalıdır. Öğrenme ve öğretmenin ayrılmaz bir parçası olan değerlendirme, öğretim programı kapsamında öğrencinin başarısı hakkında öğretmenin vardığı yargıyı içeren süreçtir. Bu süreç, belirli amaçlar dahilinde hedeflere varılıp varılmadığı hakkında kararlar vermeyi, ilgili verinin nasıl toplanıp yorumlanacağını ve bunun ilgili kişilere nasıl açıklanacağını içerir (Harlen, 2005, 207).

Öğretmenler sınıf içi eğitim öğretim faaliyetlerinde etkili kararlar verebilmek için çeşitli ölçme ve değerlendirme yöntemleri kullanırlar. Öğretmenlerin ölçme ve değerlendirmeye ilişkin bilgi düzeyi ile onların bu konuya karşı tutumları orta düzeyde ilişkilidir (Quilter, 1998). Öğretmenlerin ölçme ve değerlendirmeye karşı tutumları, onların ölçme sonuçlarını kullanma ve yorumlama becerilerini etkilemektedir. Uygun ölçme yapmak, testin etkililiğine ilişkin tutumla pozitif bir ilişkiye sahiptir (Green ve Stager, 1986). Eğer

öğretmen adayları ölçme ve değerlendirme yöntem ve tekniklerine karşı olumsuz bir tutum geliştirirlerse, bu yöntem ve teknikleri anlamaları ve kullanmaları da o oranda düşük olmaktadır (Quilter, 1998). Öğretmenlerin ölçme ve değerlendirmeye karşı olumlu bir tutum geliştirmelerinde verilen eğitimin önemli bir yeri vardır ancak bu yeterli değildir. Bazı yazarlar sınıf içi ölçme ve değerlendirme konusunda öğretmen eğitiminin yetersiz olduğunu belirtmektedir (Stiggins ve Conklin, 1992; Wise, Lukin ve Roos, 1991). Öğretmen adaylarına verilen ölçme ve değerlendirme derslerinin yetersizliğine ilişkin olarak, verilen eğitimin sınıf içindeki uygulamalarla uyumlu olmadığına dikkat çekilmektedir. Bu konuda verilen eğitim öğretmenlerin ölçme ve değerlendirme konusundaki ihtiyaçların sadece üçte birini karşılamaktadır (Stiggins, 1993).

Türkiye’de, öğretmenlerin ölçme ve değerlendirme yeterliklerinin incelendiği çeşitli çalışmalarda öğretmenlerin ölçme-değerlendirmeye ilişkin kendilerini tam yeterli görmedikleri, eğitime ihtiyaç duyduklarını göstermektedir (Akçadağ, 2010; Bıçak ve Çakan, 2004; Çelikkaya, Karakuş ve Demirbaş, 2010). Örneğin Maden ve Durukan (2011) yaptıkları çalışmada, öğretmenlerin 2005 programıyla gelen ölçme değerlendirme araç ve yöntemlerini karmaşık bulduğu, programın önerdiği ölçme değerlendirme yaklaşımları ve araç gereçleri hakkında bilgilendirilmesi gerektiği sonuçlarına ulaşmışlardır. Gelbal ve Kelecioğlu (2007) ise ölçme-değerlendirme uygulamalarında ölçme araçlarının geliştirilmesi, nasıl kullanılacağı ve sonuçlarının nasıl değerlendirileceği konusunda iyi örneklerin yaygın olmamasını, öğretmenlerin bu alanda güçlük çekmelerinin sebepleri olarak bildirmişlerdir. Öğretmenlerin ölçme-değerlendirme alanındaki durumuna ilişkin olarak Türkiye dışındaki araştırmalarda da benzer bulgulara ulaşılmıştır (Schafer, 1991; Quilter, 1998; Leavit, 2012). Eğitim-öğretim faaliyetlerinin ayrılmaz bir parçası olan ölçme-değerlendirmenin, uzmanlık gerektiren yönü, öğretmenlerin bu alana yönelik tutumlarını etkileyebilir. Campbell (1998) sahip olunan yeterlik düzeyi ile tutum arasında bir ilişki olmadığı sonucuna ulaşırken; Schafer (1991), Stiggins (1997) ve Quilter (1998) ölçme-değerlendirme

yeterlik düzeyi ile bu alana ilişkin tutumları arasında orta düzeyde bir ilişki saptamışlardır. Öğretmenlerin ölçme-değerlendirmeye ilişkin deneyimleri ve yaşantıları onların bu konuda olumsuz bir tutum geliştirmiş olabileceklerini düşündürmektedir.

Tutumlar, yaşantılar aracılığıyla ve bir öğrenme süreci sonunda oluşan belli bir süre devamlılık gösteren olumlu ya da olumsuz davranışlara yol açabilen bilişsel, duyuşsal ve davranışsal öğelerdir (Tavşancıl, 2010: 71-72). Tutumlar, davranışlara yol gösterme işlevi gördüğünden tutumların ortaya konulması (ölçülmesi) çalışmaların daha verimli olması için önlem alma fırsatı sağlamaktadır. Öğretmenlerin ölçme-değerlendirmeye ilişkin tutumlarının öğretim uygulamalarına yansıma potansiyeli dikkate alındığında onların ölçme-değerlendirmeye ilişkin tutumlarının belirlenmesi gerektiği anlaşılmaktadır.

Araştırmanın Amacı

Bu çalışmanın amacı, ilkökul ve ortaokullarda görev yapan öğretmenlerin ölçme değerlendirme yönelik tutumlarını belirleyen bir ölçek geliştirerek, geçerlik ve güvenirliliğini saptamaktır.

Yöntem

Bu çalışma tarama modelinde gerçekleştirilmiştir. Tutum geliştirme çalışmalarında geniş katılım sağlanması amaçlandığında tarama yöntemine başvurulabilmektedir. Öğretmenlerin Ölçme ve Değerlendirmeye İlişkin Tutum Ölçeğinin geliştirilmesi sürecinde *i)* deneme formunun üretilmesi, *ii)* denemenin yapılması, *iii)* deneme sonuçlarının analizi *iv)* ana uygulama ve *v)* ana uygulama sonuçlarının analizi basamakları izlenmiştir.

Katılımcılar

Katılımcılar, 2013-2014 öğretim yılında Aksaray İli merkez, ilçe ve köylerinde resmi ilkökul ve ortaokul kademelerinde görev yapan farklı branşlardan öğretmenlerden oluşmuştur. Katılımcıların belirlenmesinde tabakalı örnekleme yöntemi benimsenmiştir. Bu amaçla önce Aksaray ilindeki beş eğitim bölgesinin her biri birer tabaka olarak ele alınmıştır. Daha sonra her bir tabakadan basit seçkisiz yöntemle seçilen okullardaki

toplam 301 öğretmen araştırmaya dâhil edilmiştir (Balcı, 2011). 2014 yılı Mart-Mayıs ayları arasındaki dönemde öğretmenlerin görev yaptıkları okullara gidilerek katılımcılara erişilmiş ve veri toplama araçları araştırmacılar tarafından uygulanmıştır. Katılımcılara ait demografik bilgiler Tablo 1’de verilmektedir.

Katılımcıların 147’si erkek, 152’si kadındır.

Öğretmenlerin branşlara göre dağılımında bir denge olduğu görülmektedir. Mesleki kıdeme göre öğretmenlerin dağılımına bakıldığında ilk beş yılında olanların % 13.3; altı ile on beş yıl arasında olanların oranı ise % 65’tir. Katılımcıların mesleki kıdem açısından özellikle 6-15 yıl aralığında yoğunlaştığı bu nedenle genç ve orta yaşlarda öğretmenlerin ağırlıklı olduğu bir katılımcı kitlesi söz konusudur.

Tablo 1. Deneme ve Ana Uygulamaya Katılan Öğretmenlerin Demografik Bilgileri

Katılımcı Özellikleri	f/%	Ön Uygulama							Toplam
		1	2	3	4	5	6	7	
Cinsiyet 1: Kadın; 2: Erkek	f	63	103	-	-	-	-	-	166
	%	38	62	-	-	-	-	-	100
Branş 1: Türkçe; 2:Mat; 3: Fen; 4: İng.; 5: Sosyal; 6:Din Kült.; 7:Diğer	f	39	34	36	26	24	7	-	166
	%	23.5	20.5	21.7	15.7	14.5	4.2	-	100
Mesleki Kıdem 1: 1-5 yıl; 2:6-10 yıl; 3: 11-15 yıl; 4: 16-20 yıl; 5: 21+	f	34	58	46	20	8	-	-	166
	%	20.5	34.9	27.7	12.0	4.8	-	-	100
Katılımcı Özellikleri	f/%	Ana Uygulama							
Cinsiyet 1: Kadın; 2: Erkek	f	152	147	-	-	-	-	-	301
	%	51	49	-	-	-	-	-	100
Branş 1: Türkçe; 2:Mat; 3: Fen; 4: İng.; 5: Sosyal; 6:Din Kült.; 7:Diğer	f	57	54	56	47	25	35	26	301
	%	18.9	17.9	18.3	15.6	8.30	11.6	8.0	100
Mesleki Kıdem 1: 1-5 yıl; 2:6-10 yıl; 3: 11-15 yıl; 4: 16-20 yıl; 5: 21+	f	40	71	125	65	-	-	-	301
	%	13.3	23.6	41.5	21.6	-	-	-	100

Ölçme Aracının Geliştirilmesi

Bu ölçeğin geliştirilmesinde öncelikle alan yazın taramasından yararlanılmıştır. Bu ölçekte tutum konusu ölçme ve değerlendirmedir. Ölçek, bireylerin kendi ifadelerine dayalı tasarlanmıştır. Tavşancıl (2010)’a göre tutumlar üç farklı yolla ölçülebilmektedir:

“Bireylerin davranışları gözlenerek, fizyolojik tepkilerinden çıkarsamalar yaparak ya da bireylerin bir dizi

ifade/madde ya da sığata verdikleri tepkilere dayanarak. Bu yöntemler ölçekleme teknikleri olarak tanımlanmakta ve ölçek olarak ifade edilmektedir.”

Bu çalışmada tutum konusu olan ölçme-değerlendirmeye ilişkin öğretmenlerin tutumları likert tipi tutum ölçeği kullanılarak ölçülmeye çalışılmıştır. Eşit aralıklı ölçekte en olumsuz (1) ile en olumlu (7) arasında bir tercihte bulunmaları istenmiştir. Böylece

her bir maddeden alınabilecek puana göre öğretmenlerin ölçme-değerlendirmeye ilişkin tutum puanları hesaplanmaktadır. Bu ölçekte en düşük 21 en yüksek ise 147 puan alınmaktadır. Ölçeğin işe vuruk kullanımında puan aralıkları bulgular kısmında verilmektedir.

Öğretmenlerin ölçme-değerlendirmeye ilişkin tutumları bilişsel, duyuşsal ve davranışsal olmak üzere üç boyutta tanımlanmıştır. Ayrıca ölçme-değerlendirme alanının kapsamı, alan yazın (Erkan ve Gömleksiz, 2008; Özçelik, 2011; Turgut ve Baykul, 2010) taranarak ve MEB Öğretmenlik Mesleği Genel Yeterlikleri (MEB, 2008) incelenerek saptanmış, kapsamın beş alt başlıktan oluştuğu belirlenmiştir. Bunlar: *i)* ölçmeyle ilgili temel kavramlar, *ii)* ölçme araçları (nitelikleri, geliştirilmesi ve uygulanması), *iii)* istatistik-matematiksel işlemler, *iv)* puanlama, *v)* ölçme sonuçlarının kullanımı (dönüt verme-geliştirme)dir. İçerik ve boyutlar kullanılarak bir matris üretilmiştir. Matristen yararlanarak 36 maddeden oluşan bir ölçme aracı tasarlanmıştır. Bu ölçme aracı yedili likert tipindedir (1: Beni hiç temsil etmiyor; 7: Beni tam temsil ediyor). Tasarlanan ölçme aracının kapsam ve içerik geçerliğini kontrol amacıyla iki alan uzmanının ve bir ölçme uzmanının görüşlerinden yararlanılmıştır. Gelen görüşler doğrultusunda bir maddenin ifadesi değiştirilmiştir ("Ölçme ile ilgili bilmediğim işlemleri öğrenmek için çaba gösteririm" ifadesine "*matematiksel-istatistiksel*" ifadesi eklenmiştir). Ölçme aracının yapı geçerliğini test etmek amacıyla deneme uygulaması yapılmıştır.

Verilerin Analizi

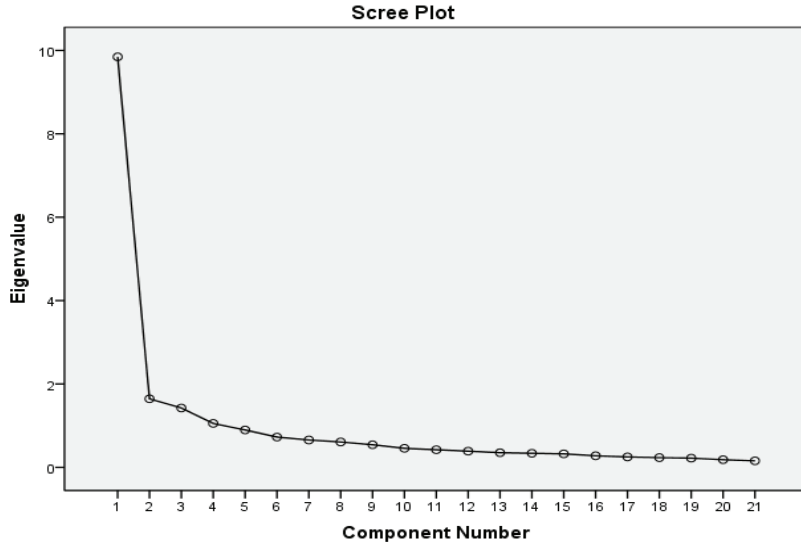
Deneme uygulamasına Aksaray il merkezindeki 4 ilkokul ve 10 ortaokuldan basit seçkisiz yöntemle belirlenen 166 branş öğretmeni katılmıştır. Ana uygulamada Aksaray İli merkez, ilçe ve köylerindeki 9 ilkokul ve 17 ortaokulda görev yapan ve pilot uygulamaya katılmamış 301 branş öğretmeni yer almıştır. Ölçekte boş bırakılan maddelere, verilerin ortalamalarına göre değerler atanmıştır. Veriler açıklayıcı (AFA) ve doğrulayıcı faktör analizi (DFA) teknikleriyle incelenmiştir. Analiz sürecinde SPSS 22.0 ve LISREL 8.80 istatistik paket programları kullanılmıştır.

Bulgular

Ölçeğin öncelikli olarak, faktör analizine uygun olup olmadığını anlamak amacıyla KMO ve Bartlett testi yapılmıştır. Bu kapsamda KMO testi ölçüm sonucunun .50 ve daha üstü, Bartlett küresellik testi sonucunun da istatistiksel olarak anlamlı olması gerekmektedir (Jeong, 2004). Ön uygulama sonucunda faktör analizi yapılmış ve yapılan KMO testi sonucu .89 olarak bulunmuştur. Bartlett Küresellik Testi sonuçları incelendiğinde ise, elde edilen ki-kare değerinin .01 düzeyinde manidar olduğu görülmüştür. Tüm maddeler ile yapılan ilk analizde, öz değeri 1'den büyük olan 6 faktör olduğu tespit edilmiştir. Bir maddenin faktördeki en yüksek yük değeri ile bu değerden sonra en yüksek olan yük değeri arasındaki farkın en az .10 olması önerilir (Büyüköztürk, 2008). Faktör yük değerleri birden fazla faktöre girip aralarında faktör yük değeri farkları 0,10'dan daha az olan 2, 13, 19, 21, 31 ve 35 numaralı maddeler (Temel kavramları birbirinden ayırt ederim. / Ölçme araçlarını hazırlarken sıkılırım. / Ölçme araçlarını uygularken sıkılırım. / Ölçme aracını uygularken zorlanırım. / Ölçme sonuçlarını kullanmak sıkıntı verir. / Ölçme ve değerlendirme sonuçlarını öğrencilerimle paylaşıyorum.) ölçekten çıkarılmıştır. Ayrıca, ölçekteki 32 numaralı maddenin (...ölçme sonuçlarını öğrenmenin geliştirilmesi için kullanırım.) bir boyutu tek başına ölçtüğü saptanarak, bu madde de ölçekten çıkarılmıştır. Daha sonra yapılan birinci ve ikinci düzey DFA sonucunda hata varyanslarına göre en fazla sorun teşkil eden 4, 22, 23, 24, 25, 26, 29 ve 36 numaralı maddeler (Temel kavramlarını öğrenmeye meraklıyım. / Ölçmedeki istatistiksel-matematiksel işlemleri bilirim. / Ölçmedeki matematiksel-istatistiksel işlemler ilgimi çeker. / Ölçmedeki matematiksel-istatistiksel işlemleri yapmakta başarısız olurum. / Ölçme sürecinde gerekli matematiksel-istatistiksel işlemleri kullanırım. / Ölçme ile ilgili bilmediğim işlemleri öğrenmek için çaba gösteririm. / Farklı puanlama türlerini kullanırım. / Ölçme ve değerlendirme sonuçlarını velilerle paylaşıyorum.) ölçme aracından çıkarılmıştır. Ön uygulama sonucunda ölçeğin toplam Cronbach Alpha güvenirlilik katsayısı .91 olarak bulunmuştur. Toplamda 15 madde ölçme aracından çıkarılarak, 21 madde ile ana uygulamaya geçilmiştir.

Ana uygulamada Aksaray İli merkez, ilçe ve köylerindeki 9 ilkokul ve 17 ortaokulda görev yapan ve pilot uygulamaya katılmamış 301 branş öğretmeni yer almıştır. Kalan 21 madde ile ana uygulamanın analizine geçilmiş,

sonucunda ölçeğin 4 faktör altında toplandığı görülmüştür. Analiz sonunda ölçme aracının Açıklayıcı Faktör Analizine ilişkin son yamaç eğri grafiği Şekil 1’de gösterilmektedir.



Şekil 1. Ana uygulamaya ilişkin faktörlerin öz değerlerine ait scree plot

Ana uygulama sonucunda yapılan KMO testi sonucu .92 olarak bulunmuştur. Bartlett Küresellik Testi sonuçları incelendiğinde ise, elde edilen ki-kare değerinin .01 düzeyinde manidar olduğu görülmüştür.

Ölçekte dört bileşen vardır. "Ölçme aracı hazırlama" boyutunu ölçen 6 (Örnek madde [ÖM]: Ölçme araçlarını hazırlamakta kendimi yeterli hissedirim), "uygulama" boyutunu ölçen 7 (ÖM: Ölçme sonuçlarını, öğrencilerin daha iyi öğrenmesi için kullanırım), "temel kavramları bilme" boyutunu ölçen 5 (ÖM: Ölçme ve değerlendirmenin temel kavramlarını kullanırım) ve "alternatif ölçme araçları" boyutunu ölçen 3 (ÖM: Alternatif ölçme araçları dikkatimi çeker) madde bulunmaktadır. Ana uygulamanın AFA'sına ilişkin faktör yük değerleri Tablo 2'de gösterilmektedir.

Öğretmenlerin Ölçme ve Değerlendirmeye İlişkin Tutum Ölçeği'nde yapılan AFA sonucunda 4 faktör elde edilmiştir. Ölçeğe ilişkin toplam varyansın; ölçme aracı hazırlama boyutu %20 'sini, uygulama boyutu %20 'sini, temel kavramları bilme boyutu %16'sını, alternatif ölçme araçları boyutu ise %11'ini açıklamaktadır. Bu dört faktörün varyansa

yaptığı katkının %67 olduğu görülmektedir. Yapılan faktör analizi sonucunda elde edilen veriler ölçeğin geçerliğinin iyi düzeyde olduğuna işaret etmektedir.

Ölçeğin güvenilirliğini tespit etmek amacıyla Cronbach Alfa güvenilirlik katsayısı belirlenmiştir. Yapılan istatistiklere göre ölçeğin Cronbach Alfa değerleri; ölçme aracı hazırlama boyutu için .90, uygulama boyutu için .89, temel kavramları bilme boyutu için .83, alternatif ölçme araçları boyutu için .82 olarak bulunmuştur. Ölçeğin tamamı için hesaplanan güvenilirlik katsayısının .94 olduğu tespit edilmiştir. Ölçeğin tüm boyutlarında madde-test (Item-Total) korelasyonu 0,44 ile 0,73 arasında değişmektedir. Bu durumda ölçeğin tamamının yüksek derecede güvenilir olduğu; ayrıca tüm boyutların da yüksek derecede güvenilirliğe sahip olduğu görülmektedir.

Ölçeğin geçerlik kanıtlarını artırmak amacıyla 301 üniversite öğrencisinin verileri üzerinde doğrulayıcı faktör analizi uygulanmıştır. Doğrulayıcı faktör analizi tüm maddelerin dört boyutu ölçtüğü bilgisine dayanarak gerçekleştirilmiştir. Doğrulayıcı faktör analizine ilişkin sonuçlar Tablo 3'de verilmiştir.

Tablo 2. Ana Uygulamaya Ait Faktör Yük Değerleri

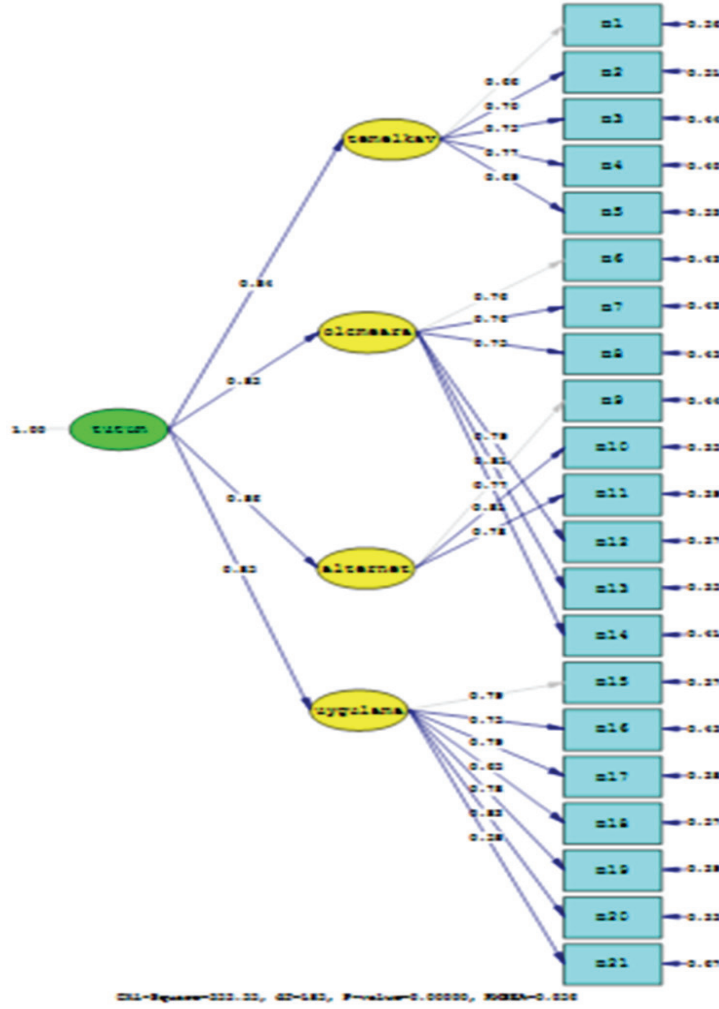
MADDELER	Döndürme Sonrası Yük Değerleri			
	F1	F2	F3	F4
1.BOYUT: Ölçme Aracı Hazırlama				
14. Ölçme araçlarını geliştirirken zorlanmam.	.788			
13. Ölçme araçlarını geliştirmekten hoşlanırım.	.761			
12. Ölçme araçlarını hazırlamakta kendimi yeterli hissederim.	.757			
6. Nitelikli bir ölçme aracı geliştirmekte yeterliyim.	.737			
8. İhtiyaç duyduğum ölçme aracını geliştirebilecek bilgiye sahibim.	.652			
7. Nitelikli bir ölçme aracının nasıl olması gerektiğini araştırırım.	.616			
2.BOYUT: Uygulama				
20. Ölçme sonuçlarını, öğrencilerin daha iyi öğrenmesi için kullanırım.		.755		
17. Puanlama bilgisine sahibim.		.742		
21. Ölçme ve değerlendirme sonuçlarının duyurulması gereklidir.		.737		
19. Ölçme sonuçlarının nasıl kullanılacağını bilirim.		.690		
16. Ölçme araçlarını planladığım gibi uygularım.		.662		
15. Ölçme aracının nasıl uygulanması gerektiğini bilirim.		.644		
18. Puanlama bende merak uyandırır.		.521		
3.BOYUT: Temel Kavramları Bilmek				
2. Ölçme ve değerlendirmenin temel kavramlarını araştırırım.			.731	
4. Ölçme ve değerlendirmenin temel kavramlarını kullanırım.			.722	
3. Ölçme ve değerlendirme temel kavramlarını bilirim.			.715	
1. Ölçme ve değerlendirmenin temel kavramları (ölçek, ölçüt vb...) dikkatimi çeker.			.699	
5. Ölçme araçlarının nitelikleri (geçerlik, güvenirlilik, kullanılabilirlik) hakkında yeterli bilgiye sahibim.			.579	
4.BOYUT: Alternatif Ölçme Araçları				
10. Alternatif ölçme araçları dikkatimi çeker.				.819
9. Aynı tip ölçme araçları yerine çeşitli ölçme araçlarını kullanırım.				.712
11. Alternatif ölçme araçları hakkında bilgi sahibiyim.				.602

Tablo 3. Ölçeğe Ait Doğrulayıcı Faktör Analizi Uyum Değerleri

Model	χ^2	$(\chi^2/sd)^*$	RMSEA	SRMR	NNFI	CFI	GFI	AGFI
Birinci Düzey	354,76	1,94	0,06	0,05	0,98	0,99	0,82	0,77
İkinci Düzey	355,55	1,92	0,06	0,06	0,98	0,99	0,82	0,77
* $p < 0,01$								

Elde edilen değerler incelendiğinde, ölçeğin 21 maddeden oluşan 4 faktörlü yapısının genel olarak iyi uyum değerleri ortaya koyduğu, bu

değerlerin kabul edilebilir olduğu ve bir model olarak doğrulandığı görülmektedir. Ölçeğe ait yol diyagramı Şekil 2'de verilmektedir.



Şekil 2. Öğretmenlerin Ölçme ve Değerlendirmeye İlişkin Tutum Ölçeğine ait yol diyagramı

Sonuç ve Tartışma

Ölçme ve Değerlendirmeye İlişkin Tutum Ölçeğinin geliştirilmesinde ilk aşama olarak, ilgili literatür taranmış, yurt dışında ve Türkiye’de kullanılmakta olan benzer ölçekler incelenmiştir. Ölçeğin yapı geçerliğini belirlemek amacıyla Açıklayıcı Faktör Analizi uygulanmış; madde geçerliğini saptamak için madde-toplam puan korelasyonları incelenmiştir. Ölçeğin geçerlik kanıtlarını artırmak amacıyla doğrulayıcı faktör analizi uygulanmıştır. Ölçeğin güvenilirliğini belirlemek amacıyla ise Cronbach Alfa iç tutarlılık katsayısı ve test-tekrar test güvenilirlik katsayısı incelenmiştir.

Öğretmenlerin ölçme ve değerlendirmeye ilişkin tutumlarını belirlemeye yönelik geliştirilen bu ölçekte elde edilen bulgular,

ölçeğin belirlenen 4 boyutta uygun nitelikte olduğunu göstermektedir. Geliştirilen ölçek, 7’li Likert tipi olup, 21 maddeden oluşmaktadır. Ölçeğe uygulanan Açıklayıcı Faktör Analizi (AFA) sonucunda öğretmenlerin ölçme ve değerlendirmeye yönelik tutumları belirlenmeye çalışılmış ve ölçeğin 4 boyutta toplandığı görülmüştür.

Boyutlar aşağıdaki şekilde isimlendirilmiştir:

1. Boyut: Ölçme Aracı Hazırlama
2. Boyut: Uygulama
3. Boyut: Temel Kavramları Bilme
4. Boyut: Alternatif Ölçme Araçları

Yapılan AFA sonucunda elde edilen boyutların toplamı, ölçeğin %67’sini açıklamaktadır.

AFA'ya ilişkin bulgular, ölçeğin geçerliğinin iyi düzeyde olduğunu göstermektedir. Ölçeğin tamamının Cronbach Alfa değeri .94 olarak hesaplanmıştır. Her bir faktör için hesaplanan Cronbach Alfa katsayı değeri .90, .89, .83 ve .82 olarak tespit edilmiştir. Bu durum ölçeğin yüksek güvenirlilik derecesine sahip olduğunu göstermektedir. Ölçeğe ait birinci ve ikinci düzey DFA uyum değerleri; $\chi^2/sd = 1,92$; RMSEA=0,06; SRMR=0,06; NNFI=0,98; CFI=0,99; GFI=0,82 ve AGFI=0,77 olarak bulunmuştur. Bu bulgular incelendiğinde, ölçeğin 21 maddeden oluşan 4 faktörlü yapısının genel olarak iyi uyum değerleri ortaya koyduğu, bu değerlerin kabul edilebilir olduğu ve bir model olarak doğrulandığı görülmektedir.

KAYNAKÇA

- Akçadağ, T. (2010). Öğretmenlerin ilköğretim programındaki yöntem, teknik, ölçme ve değerlendirme konularına ilişkin eğitim ihtiyacı. *Bilgi*, 53, 29-50.
- Balcı, A. (2011). *Sosyal bilimlerde araştırma yöntem, teknik ve ilkeler*. Ankara: PegemA
- Baykul, Y. (2000). Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması. Ankara: ÖSYM Yayınları.
- Bıçak, B. ve Çakan, M. (20-22 Aralık 2004). *Lise öğretmenlerinin sınıf içi ölçme ve değerlendirme uygulamalarına dönük görüşleri*. Milli Eğitim Bakanlığı, Orta Öğretimde Yeniden Yapılanma Sempozyumunda sunulmuş bildiri, Ankara.
- Büyükköztürk, Ş. (2008). *Sosyal Bilimler İçin Veri Analizi El Kitabı. İstatistik, Araştırma Deseni SPSS Uygulamaları ve Yorum*. Ankara: Pegem A.
- Campbell, C. S. (1998). *Preservice teachers' classroom assessment practices: An investigation into the factors that contribute to the discrepancy between measurement instruction and practical implementation* (Unpublished doctoral dissertation), Southern Dlinois University at Carbondale.
- Çakan, M. (2004). Öğretmenlerin ölçme-değerlendirme uygulamaları ve yeterlik düzeyleri: İlk ve ortaöğretim. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 37(2), 99-114.
- Çelikkaya, T., Karakuş, U. ve Demirbaş, Ç. Ö. (2010). *Sosyal bilgiler öğretmenlerinin ölçme-değerlendirme araçlarını kullanma düzeyleri ve karşılaştıkları sorunlar*. *Ahi Evran Üniversitesi Eğitim Fakültesi Dergisi*, 11, 57-76.
- Erkan, S. ve Gömleksiz, M. (2008). *Eğitimde ölçme ve değerlendirme*. Ankara: Nobel
- Gelbal, S. ve Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33, 135-145.
- Green, K. E., & Stager, S. F. (1986). Measuring attitudes of teachers toward testing. *Measurement and Evaluation in Counseling and Development* 19. 141-150.
- Güler, N. ve Gelbal, S. (2010). Açık uçlu matematik sorularının güvenirliliğinin klasik test kuramı ve genellenebilirlik kuramına göre incelenmesi. *Kuram ve Uygulamada Eğitim Bilimleri*, 10 (2), 991-1019.
- Harlen, W. (2005). Teachers' summative practices and assessment for learning – tensions and synergies. *The Curriculum Journal*, 16(2), 207 – 223
- Leavitt, L. H. (2012). *Preservice teachers' perceptions of assessment and teacher preparation in educational measurement* (Unpublished doctoral dissertation). University of South Dakota.
- Maden, S. ve Durukan, E. (2009). Türkçe dersi öğretmenlerinin ölçme-değerlendirmeye ilişkin algıları. *Milli Eğitim*, 190, 212-233.
- MEB (2008). *Öğretmen yeterlikleri, öğretmenlik mesleği genel ve özel alan yeterlikleri*. Ankara: Devlet Kitapları Müdürlüğü.
- Özbaşı, D. (2009). Sınıf Öğretmenleri İçin Öğrenci Başarısını Ölçme ve Değerlendirme İle İlgili Yeterlik Göstergelerinin ve Bunlara İlişkin Algılarının İncelenmesi. Yayımlanmamış Yüksek Lisans Tezi, Anlra Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara

- Özçelik, D. A. (2011). Ölçme ve değerlendirme. Ankara: PegemA
- Quilter, S. M. (1998). *Inservice teachers' assessment literacy and attitudes toward assessment* (Unpublished doctoral dissertation). University of South Carolina.
- Schafer, W. D. (1991). Essential assessment skills in professional education of teachers. *Educational Measurement: Issues and Practice*, 10(1), 3–6.
- Stiggins, R. J. (1997). *Student-centered classroom assessment*. Upper Saddle River, NJ: Prentice-Hall
- Stiggins, R. J. (1993). Teacher training in assessment: Overcoming the neglect. In S. L. Wise (Ed.), *Teacher training in measurement and assessment skills* (pp. 27-40).
- Stiggins, R. J., & Conklin, N. (1992). In teachers' hands: Investigating the practices of classroom assessment Albany, NY: SUNY Press.
- Şahin, Ç ve Karaman, P. (2013). Sınıf Öğretmen Adaylarının Ölçme ve Değerlendirmeye İlişkin İnançları. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*. 28(2). 394-407.
- Tavşancıl, E. (2010). *Tutumların ölçülmesi ve SPSS ile veri analizi*. Ankara: Nobel
- Turgut, M. F. ve Baykul, Y. (2010). *Eğitimde ölçme ve değerlendirme*. Ankara: PegemA.
- Wise, S. L., Lukin, L. E., & Roos, L. L. (1991). Teacher beliefs about training in testing and measurement. *Journal of Teacher Education*. 42(1). 37-42.
- Zhang, Z., & Burry-Stock, J. A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education*, 16(4), 323-342.

Summary

Introduction

Measurement and assessment is an integral component of learning and teaching. Assessment refers to all processes employed by academic staff to make judgments about the achievement of students in units of study and over a course of study. These processes include making decisions about what is relevant evidence for a particular purpose, how to collect and interpret the evidence and how to communicate it to intended users.

There is consensus indicating relationship between measurement practices, training, and attitudes toward assessment in the classroom. In this sense it is pointed out that teachers who had positive experiences in their personal past with evaluation tended to be more positive about assessment in general. This study focused on developing teachers' attitude scale regarding measurement and evaluation.

Methodology

The study is a descriptive research designed as a survey method. The sample of the study consists of 301 primary and secondary subject teachers employed in Aksaray province. The sample was drawn via stratified sampling

method from the population. 152 of the participants were female and 147 of them were male.

Findings

Results of exploratory factor analysis showed that the 21 items loaded on four factors. The total variance explained was 67 % and the factor loadings ranged between .52 to .82 and corrected item-total correlations ranged from .44 to .73. The internal consistency reliability coefficient (Cronbach's Alpha) was found as .94. On the other hand the confirmatory model for the scale indicated close model fit (RMSEA=.06, NFI=.98, CFI=.99, GFI=.82, AGFI=.77).

Discussion

These results revealed that the scale provide reliable and valid evidences and suggest it may be useful in research applications related with attitude. The results of this study have implications for researchers and teacher educators. It provides evidence for them for the value of university coursework in tests and measurement. Additionally policy makers of ministry of education can benefit the results in training in-service teachers.