

Öğretmenlerin Ölçme ve Değerlendirme Alanı Yeterliklerinin Yargıcı Kararlarına Dayalı Ölçeklenmesi¹

Recep Serkan ARIK² & Ömer KUTLU³

ÖZET

Bu çalışmanın amacı, okul eğitiminde gereksinim duyulan ölçme ve değerlendirme alanı öğretmen yeterliklerinin belirlenmesi, Delphi tekniği ile belirlenen yeterlik alanlarının yargıcı kararlarına dayalı olarak ölçeklenmesidir. Bu bağlamda üç oturum halinde uygulanan Delphi paneli uygulamalarından elde edilen nitel ve nicel veriler, ölçekleme yaklaşımlarından ikili karşılaştırma ve sıralama yargılarıyla, ölçeklemeden elde edilen bulgular ilişkilendirilerek yorumlanmıştır. Çalışma Ankara ilinde merkez ilçelerden dokuz ilköğretim okulunda ölçme ve değerlendirme konusunda tecrübe sahibi olan 300 öğretmen üzerinde gerçekleştirilmiştir. Çalışmada gerekli verilerin toplanmasında “iyi bir öğretmende olması gereken ölçme ve değerlendirme alan yeterlikleri nelerdir?” açık uçlu sorusuna verilen yanıtlar ve bunlarla oluşturulan yargıcı kararlarına dayalı ölçekleme formu kullanılmıştır. Elde edilen verilerin analizinde Thurstone V. Hal denklemi kullanılmıştır. En yüksek ölçek değerine sahip üç yeterlik ile en az ölçek değerine sahip üç ölçek değerlerinin kendi içlerinde birbirine yakın olduğu bulunmuştur. İkili karşılaştırma ile ölçeklemede ikinci ve üçüncü Delphi uygulamasında “Bilişsel, duyuşsal ve psikomotor alanda ölçme araçları hazırlar, teknolojik araç ve gereçleri ölçme ve değerlendirme yaparken kullanır ve Farklı soru formatları kullanma” yeterlikleri daha önemli olarak belirlenmiştir. Sıralama yargılarıyla ölçeklemede, “Performansa, projeye ve portfolyaya dayalı ölçmeler yapar, Öğrencinin öz değerlendirme ve akran değerlendirme yapmasına olanak tanır, Teknolojik araç ve gereçleri ölçme ve değerlendirme yaparken kullanma” yeterlik alanlarını öğretmenler önemli görmektedirler. Delphi uygulamaları arasında hem ikili karşılaştırma hem de sıralama ile ölçeklemede yüksek düzeyde ilişki ($r=0.99$, $p<.05$) olması dikkat çeken bulgular arasındadır. Bu yüksek ilişki Delphi tekniğinde uzlaşının olduğunun göstergesidir. İkili karşılaştırma ile sıralama yargılarıyla ölçekleme arasında ilişkinin olmaması önemli bir bulgu olarak bu araştırmada yer almıştır.

Anahtar Kelimeler: Öğretmen yeterlikleri, Ölçme ve değerlendirme, Ölçekleme, Delphi tekniği



DOI Number: <http://dx.doi.org/10.12973/jesr.2013.3210a>

¹ Bu makale “Öğretmenlerin Ölçme ve Değerlendirme Alanı Yeterliklerinin Yargıcı Kararlarına Dayalı Ölçeklenmesi” adlı doktora tezinden üretilmiştir”

² Arş. Grv. Dr. – Dumlupınar Üniversitesi Eğitim Fakültesi, Eğitim Bilimleri Bölümü - rserkan.arik@dpu.edu.tr

³ Yrd. Doç. Dr. – Ankara Üniversitesi Eğitim Bilimleri Fakültesi, Ölçme ve Değerlendirme Bölümü - omerkutlu@ankara.edu.tr

GİRİŞ

Toplum ve dünya değiştikçe öğretmenlere yüklenen görev ve sorumlulukların da değişip, çeşitlenmesi ve artması kaçınılmazdır. Gelişen eğitim anlayışında öğretmenlerin yerine getirmesi gereken temel görevin, ulusal ve evrensel değerleri benimseyen ve sorunlara çözüm üreten, milli eğitimin ve alanı ile ilgili ders programlarının amaçlarını davranışa dönüştüren, öğrenmeyi öğrenen bireyleri, her bireyin gereksinimini de dikkate alarak yetiştirmek olduğu bildirilmektedir (MEB, 2006). Bu görev ve sorumlulukları içinde barındıran öğretmenlik mesleğini yapabilecek yeterlikteki öğretmenlerin, eğitim sisteminde yerlerini almaları büyük önem taşımaktadır. Bu bağlamda, öğretmenlerin sahip olması gereken genel ve özel yeterlikler ayrıntılı olarak araştırılıp, tartışılmakta ve belirlenmektedir.

Alanyazında öğretmen yeterliği kavramı “qualification, competency, standard, quality, capacity, characteristics, effectiveness” gibi sözcüklerle tanımlanmaktadır. Bazı kaynaklarda ise “adequacy, efficiency, capability, sufficiency” gibi sözcüklerle, kimi zaman Türkçede yeterlik, yetkinlik, ehliyet, verimlilik gibi sözcüklerle karşılık bulmaktadır (Şişman, 2009). Yeterlik kavramı, bir işi ya da görevi etkili bir şekilde yerine getirebilmek için sahip olunması gereken özellikleri ifade eder. Yeterlik, bir görevi icra etmek ve görevin gerektirdiği sorumlulukları yerine getirmek için ihtiyaç duyulan yetenek, bilgi ve becerileri ifade eden bir kavramdır. Bu kavram, belirli bir görevi ya da rolü kabul edilebilir bir düzeyde yerine getirmek için sahip olunması gereken kapasiteyi vurgular (Şahin, 2004). Yeterli olmak, belli görevleri karşılayabilecek bilgi ya da hüner veya hem bilgi hem de beceri sahibi olmaktır. Yeterlik alanları, bir kişinin, belli bir makamın görevlerini yerine getirmesi için göstermesi gereken alanlardır (Bursalıoğlu, 1975).

Birçok Avrupa ülkesinde (İngiltere, Çekoslovakya, Estonya, Polonya, Slovenya, Hollanda, Almanya) öğretmenlerle ilgili yeterlik, standart, nitelik gibi kavramlaştırmalar altında bazı özelliklerin sıralandığı görülmektedir. Bunların, öğretmenlerle ilgili seçme, istihdam, ödüllendirme, hesap sorma, mesleki eğitim ve geliştirme gibi farklı amaçlar için kullanıldığına rastlanmaktadır.

Öğretmen yeterlikleri, son yıllarda üzerinde çok durulan ve tartışılan konulardan biridir. Özellikle yeterliğin öğretme ve öğrenmedeki rolü bu konuya araştırmacıların ve uygulayıcıların ilgisini artırmaktadır (Woolfolk-Hoy & Burke-Spero, 2005). Öğretmen yeterliği, öğretmenlerin belirli bir bağlamda öğretme görevini başarılı bir şekilde yerine getirebilmesi ve öğrenci davranışlarını değiştirebilmesi için gerekli olan davranışları gösterebilmesi konusundaki inanışları olarak tanımlanmaktadır (Tschannen-Moran & Woolfolk-Hoy, 2001; Yılmaz & Çokluk, 2008).

1999 yılında hazırlanan yeterlikler uygulamaya dönüştürülemeden, 2002 yılında uygulamaya konulan ve Avrupa Birliği (AB) tarafından desteklenen Temel Eğitime Destek Projesi kapsamında yeniden öğretmen yeterliklerinin belirlenmesi için bir çalışma başlatılmıştır. MEB tarafından yayınlanan “Öğretmen Yeterlikleri” kitabında “21. yüzyılda öğretimin niteliği nasıl olmalıdır?” ve “Hangi nitelikte öğrenci ve öğretmen istiyoruz?” sorularına cevap arandığı belirtilmiştir. Türk Eğitim Derneği’nin (TED) 2009 yılında yayınladığı öğretmen yeterlikleri özet raporunda Türkiye’de tanımlanan öğretmen yeterliklerinin birçok açıdan AB ülkelerinin belirlediği öğretmen niteliklerine benzediği belirtilmiştir (TED, 2009). Bu yeterlikler öğretmen niteliğini artırmanın yanında AB’ye uyum sağlama ile de bağlantılıdır.

Öğretmen yeterlikleri; uluslararası alanyazında ve uluslararası kuruluşların yaptığı çalışmalarda “öğretmen eğitimi”, “öğretmenlik ölçütleri”, “çağdaş öğretmende olması gereken özellikler”, “öğretmenin çağdaş rolleri”, “yeni arayışlar ışığında öğretmen”, “barış

kültürü oluşturmada öğretmenlerin rolü” başlıklarında araştırılmıştır (Navarro, 2003; NcNess, Broedfoot & Osborn, 2003; NBPTS, 2004; OECD, 2004).

Alanyazında, öğretmen yeterlikleri birbirleriyle ilişkili alt başlıklar halinde sınıflandırılmıştır. Demirel (2003), öğretmenin kişisel yeterliklerini; güdüleyicilik, başarıya odaklanmışlık ve profesyonellik başlıkları altında incelerken, mesleki yeterlikleri öğretim etkinliklerini planlama, öğretim yöntem ve tekniklerini kullanma, etkili iletişim kurma, sınıfı yönetme, zamanı etkili kullanma, öğrenmeleri değerlendirme ve rehberlik yapma başlıkları altında toplamıştır. Kavcar (2002), öğretmen yeterliğini, öğretmenlik mesleğine yatkınlık, örnek ve model olma olarak sınıflamıştır. Demirel (1999a), bir başka çalışmasında öğretmen yeterliklerinin coşku, içtenlik, güvenilirlik, yüksek başarı beklentisi içinde olma, destekleme, işbirliği, esneklik ve bilgililik olmak üzere sekiz aşamada incelenmesi gerektiğini belirtmektedir.

Öğretmen yeterlikleri konusunda yapılan araştırmalar incelendiğinde, birçok araştırmacının bu konuyu farklı ele aldığı görülmektedir. Araştırmaların çoğunda öğretmenlerin etkili bir öğretimi gerçekleştirmesi için gerekli meslek bilgisi, kişilerarası iletişim becerileri, alan bilgisi, genel kültür gibi öteden beri önemsenen özellikleri vurgulanmaktadır. Bu araştırmaların daha yakın zamanda yapılanlarında ise öğretmenin esprili olması, bilgi ve iletişim teknolojilerini kullanması, yaratıcı, yansıtıcı ve eleştirel düşünme gibi düşünme becerilerine sahip olması gerektiği belirtilmektedir (Senemoğlu & Özçelik, 1989; Alkan, 1998; Gündüz, 2001; Kavcar, 2002; Demirel, 2003; Senemoğlu, 2004; Ekşi, 2006).

Eğitim sürecinde gelişimi izlemek, eksiklik ve aksaklıkları gidermek için önemlidir. Çünkü giderilemeyen eksiklik ve aksaklıklar, artan bir etki ve ağırlıkla varlıklarını sürdürmekte, önlem alınmadığında ise en değerli kaynak olan insan gücünün boşa harcanmasına yol açabilmektedir (Özçelik, 1998). Eğitim sisteminin hatasız olabilmesi için nitelikli ölçme ve değerlendirme uygulamalarının yapılması, ilgili yöntemlerin öğretmenler tarafından çok iyi bilinmesi ve seçilmesi gerekir. Bir eğitim sisteminin başarısı, o sistemi hayata geçirip uygulayacak olan öğretmenin başarısından ayrı tutulamaz (Yılmaz, Köseoğlu, Erçek & Soran, 2004). Nitelikli bir öğretmenle ilgili değerlendirmelere bakıldığında, iyi bir öğretmen kendini mesleki ve kişisel açılardan sürekli olarak geliştiren kimse olarak tanımlanmaktadır (Seferoğlu, 2004). Öğretmenlerin, adaylıktan başlayarak, çok iyi eğitim almış olması gerekir. Daha sonra da öğretmenler, hizmet içi eğitimlerle, seminerlerle bu konudaki bilgilerini artırabilirler.

Öğretmen yeterlikleri dünyada ve Türkiye’de bir çok çalışmanın öncelikli konusunu ve eğitim bilimlerinin çeşitli alanlarının önemli bir çalışma alanını oluşturmaktadır. Çünkü etkili ve verimli bir eğitim, öğretmenlik mesleği yeterliklerine sahip öğretmenlerle gerçekleşmektedir. Öğretmen yeterlikleri çalışmalarına bakıldığında farklı yollar izlendiği ve bunlara ilişkin değişik yöntemlerin denendiği belirlenmiştir. Tüm bu araştırmalar eğitim sisteminin daha ileriye gitmesini nitelikli öğretmenlerin oluşturulmasını ve öğrencilerin daha iyi eğitilmesini sağlamaktadır.

Mansfield’e göre (1996) bir yeterlik modeli, işgörenin işinde etkili olabilmesi için gerek duyduğu becerilerin ve özelliklerin detaylı ve davranış boyutunda özel betimlemelerini içerir. Yeterliklerin belirlenmesine yönelik farklı yöntemler kullanılmaktadır. Bu yöntemler:

- Mesleğin ilgili alan literatürünün incelenmesi,
- İş analizi tekniklerinin kullanılması,
- Uzmanlardan oluşan çalışma gruplarının kullanılması şeklinde sıralanabilir.

Yukarıda sıralanan her bir yöntem, bazı üstünlükler yanında sınırlılığa da sahiptir. Bazı ülkelerde öğretmen yeterlikleri araştırma konusu yapılmış, yeterlik listeleri oluşturulmuştur. Ancak, ilgili alanyazın taramasıyla kolaylıkla belirlenecek bu yeterliklerin, ait olduğu ülkenin genel vizyonu, eğitim sisteminin vizyonu, felsefesi, temel öğrenme-öğretme yaklaşımları gözetilerek hazırlandığı unutulmamalıdır. Bu nedenle bir sistemde etkili sonuçlar veren öğretmen yeterlikleri bir başka sistemde aynen uygulandığında benzer sonuçları sağlayamayabilir. Hatta başlı başına bir problem kaynağı haline gelebilir. Bu yüzden ilgili alan yazın dikkatle izlenmeli ve kullanılabilir yönleri değerlendirilmelidir (Linstone & Turoff, 2002).

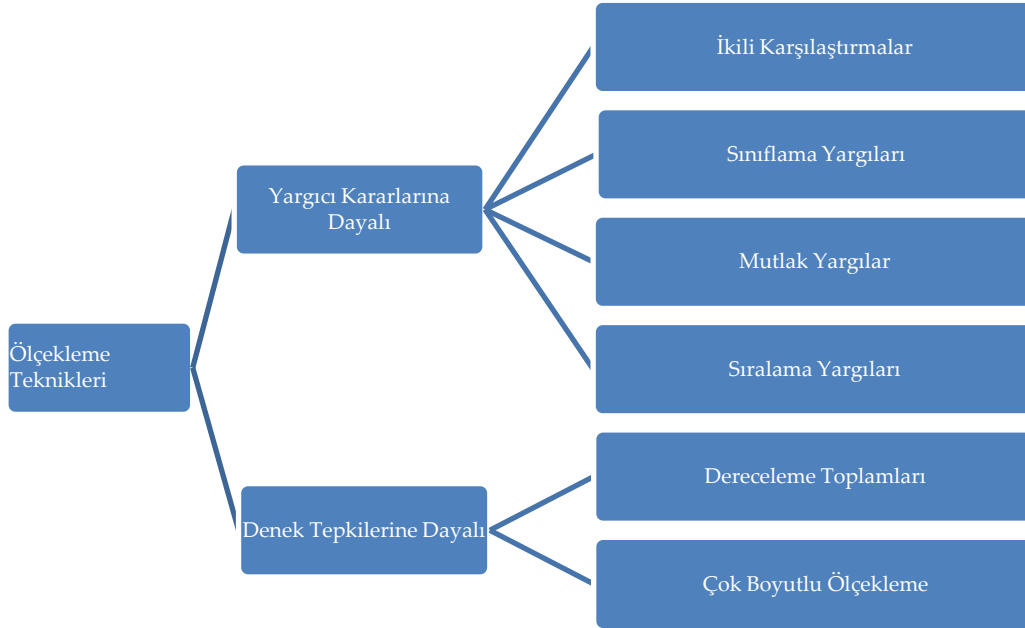
Öğretmen yeterlikleriyle ilgili çalışmalarda anket sorularıyla belirleme ve nitel bulgulardan yararlanma gibi farklı teknikler kullanılmış ve bunlardan yola çıkarak yorumlar yapılmıştır. Yeterliklerin belirlenmesi sürecinde etkili bir yol olan Delphi tekniği yapılan çalışmalarda sıklıkla karşımıza çıkmaktadır (Linston & Turoff, 1975; Sackman, 1975; Quinn, 1986; Paykoç & Ok, 1990; Woudenberg, 1991; Brady, 1994; Lang, 1994; Pollard & Tomlin, 1995; Gökçe, 1999; Şahin, 2000; Semerci & Semerci, 2001; Şahin, 2004).

Öğretmen yeterliklerini belirleme ve bir görüş birliği sağlama aracı olarak Delphi tekniği, bir problem durumuna ilişkin uzman görüşlerini sistematik bir şekilde elde etmeye yönelik olarak kullanılır. Bu teknik kullanılarak bir problem durumuna farklı açılardan bakan bireylerin ve grupların yüz yüze gelmeden ortak görüş oluşturmaları amaçlanmaktadır. Delphi tekniği ile katılımcıların farklı bakış açılarının yanında yaratıcılıklarından yararlanılması amaçlanır (Şahin, 2004). Uzlaşma sağlama aracı olarak ifade edilen Delphi, bir problem durumuna ilişkin uzman görüşlerini sistematik bir şekilde elde eden bir tekniktir (Sackman, 1975; Quinn, 1986; Semerci & Semerci, 2001). Delphi tekniği kullanılarak bir problem durumuna farklı açılardan bakan bireylerin ve grupların yüz yüze gelmeden uzlaşmaları amaçlanmaktadır (Linston & Turoff, 1975; Lang, 1994).

Yeterlikler konusunda uzmanların ortak bir görüş bildirmeleri (uzlaşmaları) önemlidir. Bu uzlaşma için aritmetik ortalama ve ortancaya göre belirlenmesi yerine daha belirgin ve ortaya koyucu istatistiksel tekniklere gereksinimi vardır. Bunun için özellikle araştırmacılar, duyuşsal alandaki pek çok davranışın ölçeklenmesinde ölçekleme tekniklerinin kullanılmasını önermişlerdir (Anıl & Güler, 2006; Kan, 2008a; Özer & Acar, 2011).

Davranış bilimleri alanına giren ölçme işlemleri, ölçülecek olan özellikler ve ölçme işlemi için kullanılan araçlar göz önüne alındığında, fiziksel ölçmelere nazaran daha titiz çalışmayı gerektirir (Kan, 2008a). Kişilik, ilgi, tutum, güdü, yetenek, özyeterlik (self-efficacy), benlik algısı (self-esteem) gibi birçok psikoloji ve eğitim biliminin alanında olan, doğrudan gözlenemeyen fiziksel büyüklükleri bilinmeyen değişkenler vardır. İnsan davranışlarına ve davranışlar arasındaki neden sonuç ilişkilerine ışık tutabilmek için psikoloji biliminin temel konuları olan bu yapıların ölçülebilir ya da gözlenebilir kılınması çok önemlidir. Psikolojik değişkenleri ölçme yollarının bulunması, ölçme araçlarının geliştirilmesi ve ölçeklenmesi, ölçme ve değerlendirme bilim dalının temel konuları içinde yer almaktadır. Ölçekleme, ölçek yapma işlemi ilk olarak psikolojinin bir alt dalı olan psikofizikte ortaya çıkmıştır. Psikofizik, ölçülmek istenen özelliğin ya da fiziksel uyarıcının bireyde oluşturduğu etkiyi araştıran bir bilim dalıdır. Burada üzerinde durulan nokta, uyarıcıların ölçülen özellikleri ile algılanan özellikleri arasında bir bağlantı kurmaktır (Guilford, 1954; Dunn-Runkin, 1982; Crocker & Algina, 1986; Brown & Peterson, 2009).

Ölçeklemede kullanılan yaklaşımlar iki grupta toplanabilir. Bunlar; 1) Denek tepkilerine dayalı (cevaplayıcı merkezli) yaklaşımlar ve 2) Yargıcı kararlarına dayalı yaklaşımlardır. Aşağıda bu yaklaşımların sınıflamasına ilişkin diyagram verilmiştir.



Şekil 1. Ölçeklemede kullanılan yaklaşımlar

Yargıcı kararlarına dayalı ölçekleme yaklaşımı, uyarıcıları uzman ya da bilir kişi yargılarına dayalı olarak belirli bir boyutta ölçeklemeyi içerir. Bu tür modeller gözlemcilerin her bir uyarıcının uyarıcılık derecesini belli bir yöntemle (sıralama, sınıflama, ikili karşılaştırma vb. gibi) belirlemelerine dayalıdır. Burada gözlemcilerden beklenen her bir uyarıcının ölçekleme boyutundaki yerini diğer uyarıcılarla karşılaştırarak belirlemesidir. Bunun sonucunda gözlemci yargılarının ortalama değeri uyarıcının ölçek değerini belirler. Yargıcı kararlarına dayalı ölçekleme işlemleri uyarıcı merkezlidir. Uyarıcı merkezli (stimulus centered) yaklaşımların temelinde her bir uyarıcının psikolojik ölçek üzerindeki yerini gözlemci yargılarına dayalı olarak belirlemek yer alır.

Denek tepkilerine dayalı yaklaşımlar ise uyarıcı değil, cevaplayıcı merkezlidir. Yani madde ya da uyarıcıların değil, cevapların ölçeklenmesi amacını güder (Torgerson, 1958; Tezbaşaran, 2004). Cevaplayıcı merkezli (subject centered) yaklaşım, her bir cevaplayıcının maddelere verdiği tepkilere dayalı olarak bireyleri ölçek üzerinde farklı bir yere yerleştirmeye odaklanır (Crocker & Algina, 1986). Uyarıcılara gösterilen bütün sistematik değişimler, cevaplayıcılar arasındaki farklılıklara atfedilir. Bu yaklaşım, K tane uyarıcıyı N kişilik denek grubuna uygulamayı ve onların tepkilerini toplamayı içerir. Bu yaklaşımda her bir kişi verilen uyarıcıya ilişkin kendi tepkisini belirten bir cevaplayıcı konumundadır. Bu yaklaşımda cevaplayıcının rolü uyarıcıların ölçekleme boyutundaki yerini tarafsız bir biçimde tayin etmek değil, her uyarıcının ölçekleme boyutundaki yerini kendi konumuna göre belirlemektir. Bu durum cevaplayıcı tepkilerine ve yargıcı kararlarına dayalı yaklaşımlar arasındaki farklılığı teşkil eder. Likert yöntemiyle ölçek geliştirme cevaplayıcı merkezli, denek tepkilerine dayalı yaklaşımın en bilinen örnekleridir. Bu ölçekler davranış

bilimlerinde birçok özelliğin, özellikle de tutumların ölçülmesinde en sık kullanılan ölçeklerdir (Crocker & Algina, 1986; Judd, Eliot & Kidder, 1991; Turgut & Baykul, 1993).

Ölçekleme, ölçme sürecinde nitel ayrımları gösteren gözlemlerden, nicel ayrımları gösteren ölçülere geçişte çok önemli bir halkayı içeren bir çalışma alanıdır (Anıl ve Güler, 2006). Türkiye’de ve dünyada ölçekleme alanına dönük çalışmalar son derece kısıtlıdır. Tezbaşaran (2004) yaptığı bir çalışmada konuyla ilgili çeşitli veri tabanlarını taramış ve bu konudaki çalışmaların çok sınırlı olduğunu rapor etmiştir. Ölçekleme alanında yapılan çalışmalar 2004’den günümüze incelendiğinde de çalışma sayısında önemli bir artış olmadığı dikkat çekmektedir (Anıl & Güler, 2006; Nartgün, 2006; Kan, 2008; Öğretmen, 2008; Ömür, 2009; Güler & Anıl, 2009; Bal, 2011; Özer & Acar, 2011; Öztürk, Özdemir & Gelbal, 2011; Ekinci, Bindak & Yıldırım, 2012; Albayrak & Gelbal, 2012; Bülbül & Acar, 2012).

Öğrenci başarısının belirlenmesinde bir öğretmenin ölçme ve değerlendirme yeterliklerine sahip olması beklenmektedir. Yeterlik alanları ve yeterlikler belirlenirken alanda uzmanlaşmış ve bu yeterlikleri sergileyen bireylerin görüşlerini almak yapılacak çalışmalarda büyük rol oynayacaktır. Sosyal bilimlerde çok fazla kullanılmayan yargıcı kararlarına dayalı ölçekleme tekniklerini uygulayıp, farklı bir açıdan incelemenin, öğretmen yeterlikleriyle ilgili yapılacak olan çalışmalara katkı getireceği düşünülmektedir.

Öğretmen yeterliklerinin belirlenmesi sürecinde, öğretmen görüşlerinin alınması ve bunların sistemli bir şekilde araştırmalara dahil edilmesi gereklidir. Bu durumda yeterlikleri öğretmenin belirlemesi ve bu konuda söz sahibi olması eğitimde kaliteyi artırmak adına önemlidir. Öğretmen yeterlikleri belirlenirken yalnızca ilgili alanyazının incelenmesi ve iş analizi tekniklerinin kullanılması bazı bilgilerin göz ardı edilmesine neden olacaktır. Yeterlikler bunlara ek olarak uyum ve uzlaşının olduğu farklı ölçme yaklaşımlarıyla da belirlenmelidir. Uzun yıllar öğretmen yeterlikleri belirlenmiş, bunlara ilişkin raporlar ve çalışmalar yayınlanmıştır. Okul uygulamalarında ölçme ve değerlendirmenin öğretmenler tarafından doğru ve yerinde kullanılması için yeterliklerin belirlenmesi bu çalışmanın önemini artırmaktadır.

Konuyla ilgili alanyazın incelendiğinde yeterlik algılarına dayalı doğrudan ölçekleme tekniklerinin kullanıldığı bir araştırmaya rastlanamamıştır. Buna rağmen az sayıda ölçekleme çalışmasından çıkan sonuçların ölçek geliştirme çalışmalarında, yeterlik boyutlarının belirlenmesinde kullanılması uzmanlar tarafından önerilmektedir. Bu çalışmada, yukarıda bahsedilen ölçekleme yaklaşımlarından ikili karşılaştırma ve sıralama yargılarına dayalı ölçekleme yöntemleri ve bu ölçekleme yöntemlerinden elde edilen sonuçların tutarlılığının karşılaştırılması üzerinde durulacaktır.

YÖNTEM

Araştırma Modeli

Bu araştırmada, Delphi tekniği ile belirlenen yeterliklerin yargıcı yargılarına dayalı olarak ölçeklenmesi amaçlanmıştır. Yargıcı kararlarına dayalı ölçekleme yöntemlerinin psikometrik özellikleri arasında bir fark olup olmadığının incelenmesine çalışılmıştır. Araştırma için çalışma grupları belirlenip, iki yönteme ilişkin özellikler üzerinde durumuştur. Bu nedenle araştırma “Betimsel araştırma” niteliğindedir. Geçmişte ya da halen var olan bir durumu varolduğu şekliyle betimlemeyi amaçlayan bir araştırma yaklaşımı olan tarama modelinde ayrıntı saptama düzeyi seçilmiştir. Bu düzeyde, değişkenler olabildiği ölçüde geçerli ve güvenilir bir şekilde sayısallaştırmaya çalışılır ve veri toplama işlemlerinin duyarlılığı yüksektir (Karasar, 2004). Katılımcıların bir konu ya da olaya ilişkin görüşlerinin, ilgilerinin, beceri, yetenek veya tutumlarının, özelliklerinin belirlendiği diğer araştırmalara

göre daha büyük örneklemeler ya da çalışma grupları üzerinde yapılan araştırmalara tarama araştırmaları denir (Fraenkel & Wallen, 2006). Yöntemsel olarak yeterliklerin belirlenmesi sürecinde ölçekleme yapıldığından çalışma temel araştırma özelliği de taşımaktadır.

Bu araştırma, Türkiye’de ilköğretimde görev yapan öğretmenlerin aldıkları eğitimlerle, edindikleri deneyimlerle oluşturdukları ölçme ve değerlendirme alanına ilişkin yeterlikleri belirlemek amacıyla yapılmıştır. Bu amaca hizmet etmek üzere, öğretmen görüşlerinin, ölçme ve değerlendirme alan yeterliklerinin belirlenmesi ve ölçeklenmesiyle araştırma gerçekleştirilmiştir.

Çalışma Grubu

Delphi araştırmalarının başarısı büyük ölçüde ilgili alanlardaki uzman kişilerin seçimine bağlıdır. Uygun uzmanların seçimi Delphi çalışmalarının güvenilirliğini artıran en önemli özelliktir. Bu nedenle, Lang’e göre (1994) katılımcıların seçkisiz (random) yöntemle seçilmesi kabul edilemez. Bunun yerine Delphi panel üyelerinin uzmanlıkları gözetilerek amaçlı seçimi gerekmektedir. Öğretmen çalışma grubunu, ölçme ve değerlendirme alanında eğitim almış deneyimli öğretmenler oluşturmaktadır.

Panel üyeleri uzman görüşlerini yansıtacak şekilde olmalıdır. Panel üyeleri deneyimleri ve nitelikleri neticesinde araştırma konusuna derin bir bakış sağlayabilmeli, konu hakkında önemli görüşlere sahip olmalıdırlar. Delphi tekniği kullanılarak büyük ya da küçük sayıda uzman grubuyla çalışmak mümkündür. En az 7 katılımcıdan oluşan bir grup olmalıdır. Grup genişliği 100 ya da daha fazla olabilir (Şahin, 2001).

Ankara ili merkez ilçeleri arasında öğretmen atamalarında dikkate alınan ilçenin ve okulların bulunduğu bölgeler incelenerek araştırma için ulaşım kolaylığı ve araştırma sürecine katılım göstereceğini belirten en çok öğretmenin çalıştığı üç ilçe seçilmiştir. Bu ilçeler Çankaya, Mamak ve Altındağ olarak belirlenmiştir. Çankaya’da yer alan okullar üst sosyoekonomik düzeyde yer alan okullar, orta sosyoekonomik düzeyde Mamak ilçesinde yer alan okullar, alt sosyoekonomik düzeyde ise Altındağ ilçesinde yer alan okullar araştırma grubu için seçilmiştir. Araştırma için seçilen okullar her ilçeden en çok öğretmen sayılarına sahip olma durumuna göre belirlenmiştir. Uygulama için planlanan öğretmen sayısı 468 olarak belirlenmiştir. Araştırma için seçilen 468 öğretmenin farklılıkları temsil edebilmesi adına okul sayısının büyük tutulmasına özen gösterilmiştir. Toplam 9 okulda görev yapan 468 öğretmen üzerinde araştırma yapılmıştır. Birinci Delphi anketi için okullara gidildiğinde öğretmenlerin bir kısmının tayininin çıkması, raporlu ve izinli olması ve araştırmaya gönüllü olarak katılmak istememeleri gibi nedenler araştırma grubu 334 kişi ile sınırlanmıştır. Birinci Delphi anketi ile öğretmenler uzmanlık alanları ve ölçme ve değerlendirme alanında bilgi ve tecrübeye sahip olup olmadıklarına ilişkin sorular sorulmuş ve bu kriterleri sağlayan 300 öğretmen ikinci ve üçüncü Delphi uygulamasına dahil edilmiştir. Sonuç olarak her üç Delphi paneline 300 öğretmenin verileri dahil edilmiştir.

Araştırma grubunda yer alan öğretmenlere ait demografik özellikler ve Delphi paneli için seçilen öğretmenlerin özellikleri Tablo 1’de verilmiştir.

Tablo 1. Araştırmaya katılan öğretmenlerin özellikleri

	Öğretmen Özellikleri	Sayı	Yüzde
Cinsiyet	Kadın	169	56,3
	Erkek	131	43,7
	Toplam	300	100,0
Yaş	26-35 arası	28	9,3
	36-45 arası	172	57,3
	46 ve üstü	100	33,3
	Toplam	300	100,0
Kıdem	1-10 yıl	78	26,0
	11-20 yıl	141	47,0
	21-30 yıl	65	21,7
	31-40 yıl	16	5,3
	Toplam	300	100,0
Mezun Olunan Fakülte	Eğitim Fakültesi	166	55,3
	Fen Edebiyat Fakültesi	41	13,7
	Eğitim Enstitüsü	49	16,3
	Diğer Fakülteler	44	14,7
	Toplam	300	100,0
Görev Yapılan Bölüm	Sınıf Öğretmeni	150	50,0
	Fen ve Teknoloji Öğretmeni	25	8,3
	Sosyal Bilgiler Öğretmeni	22	7,3
	Matematik Öğretmeni	13	4,3
	İngilizce Öğretmeni	22	7,3
	Teknoloji Tasarım Öğretmeni	31	10,3
	Türkçe Öğretmeni	19	6,3
	Tarih Öğretmeni	3	1,0
	Görsel Sanatlar Öğretmeni	6	2,0
	Din Kültürü Öğretmeni	6	2,0
	Rehberlik	3	1,0
Toplam	300	100,0	
Öğretmenlik Formasyonu Aldığı Yer	Lisans Eğitimi Sırasında	222	74,0
	Formasyon Programına Katılarak	78	26,0
	Toplam	300	100,0
Lisansüstü Eğitim Alma Durumu	Lisansüstü Eğitim Almadı	222	74,0
	Lisansüstü Eğitim Aldı	78	26,0
	Toplam	300	100,0
Sınıflarındaki Ortalama Öğrenci Sayısı	20'den Düşük	53	17,7
	21-40	183	61,0
	40'tan Fazla	64	21,3
	Toplam	300	100,0
Hizmet İçi Eğitime Katılma	Katılanlar	186	62,0
	Katılmayanlar	114	38,0
	Toplam	300	100,0

Öğretmenlerin mezun oldukları fakültelere ilişkin dağılım incelendiğinde 166 (% 55,3) öğretmenle en çok eğitim fakültesi mezunu, 49 (% 16,3) öğretmen eğitim enstitüsü mezunu, 44 (% 14,7) öğretmen mühendislik, iktisadi idari bilimler, ilahiyat gibi diğer fakültelerden ve 41 (% 13,7) öğretmen de fen edebiyat fakültelerinden mezun olmuşlardır. Görev yaptıkları branşlara ilişkin dağılım incelendiğinde en çok sınıf öğretmeni 150 (% 50) kişiyle araştırmaya dahil edilmiştir

Araştırmanın gönüllülük esasına dayanması ve öğretmenlerin iş yoğunluğunun fazla olması gibi etkenler göz önüne alındığında neredeyse tüm branşlardan öğretmenlerin delphi paneli üyesi olmayı kabul etmesi dikkat çekici bir bulgudur. Bu bulgu, öğretmenlerin

araştırmalara bakış açısının değiştiği, bilimsel çalışmalara ve öğretmenlik mesleği üzerine yapılan çalışmalara katılma konusunda istekli oldukları biçiminde yorumlanabilir.

Öğretmenlik formasyonunu nerede aldıklarına ilişkin 222 (% 74) öğretmen lisans eğitimleri sırasında aldıklarını belirtirken, 78 (% 26) öğretmen lisans sonrası formasyon eğitimi ile aldıklarını belirtmişlerdir. Öğretmen grubunun içerisinde 78'i (% 26) lisansüstü eğitim aldıklarını belirtirken, 222 (% 74) öğretmen lisansüstü eğitim almadıklarını belirtmişlerdir. Öğretmenlerin lisansüstü eğitim almaları, alanlarındaki gelişmeleri takip etmeleri göz önüne alındığında bu sayının yüksek olması dikkat çekicidir. Sınıflarındaki öğrenci sayılarına göre 21 ile 40 öğrencisi olan 183 (% 61) öğretmen, 40'tan fazla öğrenciye sahip 64 (% 21,3) öğretmen ve 20'den düşük öğrenci sayısına sahip 53 (% 17,7) öğretmen olduğu belirlenmiştir. Ankara'da farklı bölgelerden seçilen okulların sınıf mevcutlarının yüksek olması, uygulamalar sırasında öğretmenlerin bu konudaki sıkıntılarını ifade etmeleri, öğrenci başarısını değerlendirmede zorluklar yaşadıklarını belirtmeleri dikkat çekici bir bulgudur.

Öğretmenlerin ölçme ve değerlendirme alanında katıldıkları hizmet içi seminerlere ilişkin olarak 186 (% 62) öğretmen hizmetiçi eğitime katıldıklarını ifade ederken, 114 (% 38) öğretmen herhangi bir hizmetiçi eğitim programına katılmadıklarını ifade etmişlerdir. Çalışma grubunun içerisinde yer alan öğretmenlerin ölçme ve değerlendirme alanında hizmet içi eğitim almış olmaları bu alana ilişkin farkındalıklarının yüksek olduğunu göstermektedir.

Veri Toplama Araçları

Araştırmada veriler, Delphi tekniği ile belirlenmiş yeterlik alanlarında araştırmacı tarafından hazırlanan öğretmen yeterlikleri ölçeği, ikili karşılaştırma ve sıralama yargılarına dayalı olarak ölçeklenmesine olanak tanıyacak biçimde hazırlanmış ve uygulanmıştır. Ölçme ve değerlendirme alanı öğretmen yeterliklerinin belirlenmesi için oluşturulan anket formları Delphi panelleri için oluşturulmuş ve uygulanmıştır. Araştırma süresinin Delphi uygulaması ölçme ve değerlendirme alanı yeterliklerinin belirlenmesi 16 aylık bir sürede gerçekleşmiştir.

Birbirinden bağımsız ve habersiz uzman görüşlerinden ortak görüşler çıkarılmaya çalışılan, olayları ve eğilimleri kestirme, program planlama, politikalar geliştirme ve standartlar oluşturma amacıyla kullanılan Delphi tekniğine göre hazırlanan anketlerin, öğrencilere, öğretmenlere, velilere, yöneticilere ve mezunlara uygulanması önerilmektedir (Demirel, 1999b). Delphi tekniğinde uzmanlarla iletişim kurulurken, grup iletişim tekniklerinden en yaygın olanı, kâğıt-kalem (yazılı) versiyonu olarak bilinen "Geleneksel Delphi Tekniği" kullanılmıştır. Linstone ve Turoff'un (2002) önerdiği gibi bireylerin grup içerisindeki etkililiğine dayanan grup iletişiminin yapılandırıldığı Delphi tekniği, üç aşamada gerçekleşmiştir. Birinci aşamada, tartışılan konunun araştırılması, ikinci aşamada konu ile ilgili grup görüşlerinin alınmasına ilişkin işlemler, üçüncü aşamada uzlaşamayan görüşlerin ayrıştırılması, toplanan bilgilerin kontrol edilmesi, değerlendirilmesi ve geribildirim elde edilmesi yer almıştır.

Delphi Tekniğinin Uygulanması

Delphi tekniğinin uygulanması, odaklaşılan alanda uzman olan kişilerin ya da hedef kitlenin temsilcilerinin problem durumuna ilişkin yaklaşımlarını, bakış açılarını ortaya çıkarmaya, incelemeye ve bir uzlaşma sağlamaya yönelik bir dizi aşamadan oluşmuştur (Lang, 1994; Turof & Hiltz, 2001).

1. Problemin belirlenmesi

Araştırma problemi tüm katılımcılar tarafından aynı şekilde anlaşılacak biçimde bir cümle ile ifade edilmiştir. Araştırma için öğretmenlere görüşülmüş ve araştırmanın amacı ve nasıl bir süreç izleneceği aşamalarının neler olduğuna ilişkin bilgi verilmiştir.

2. Panel üyelerinin (katılımcıların) seçimi

Panel üyeleri uzman görüşlerini yansıtacak niteliktedir. Panel üyeleri deneyimleri ve nitelikleri neticesinde araştırma konusuna derin bir bakış sağlayabilmeli, konu hakkında önemli görüşlere sahiptir. Delphi tekniği kullanılarak büyük ya da küçük sayıda uzman grubuyla çalışmak mümkündür. En az 7 uzmandan oluşan bir grup olmalıdır. Grup genişliği 100 ya da daha fazla olabilir. İdeal grup büyüklüğü 120 uzmandan oluşur. Bu çalışmada 300 öğretmenle çalışılmıştır. Öğretmenlerin seçimi ve araştırmaya katılmaları için öncelikle öğretmenlere çalışma hakkında bilgi verilmiş ölçme ve değerlendirme alanında bilgi birikimine sahip oldukları, okul uygulamalarında tecrübeli ve araştırmaya katılmakta istekli oldukları bilgisi alındıktan sonra araştırma grubuna dâhil edilmişlerdir.

3. Araştırma problemine ilişkin açık uçlu olarak hazırlanan sorunun (birinci delphi anketi) panel üyelerine gönderilmesi

Anket katılımcılara elektronik ortamda gönderilebileceği gibi katılımcıların kağıt kalem kullanarak da yapmaları sağlanır. Bu araştırmada okullara tek tek gidilip her bir öğretmenle görüşerek araştırmanın amacı anlatılarak birinci Delphi anketini doldurmaları istenmiştir.

Birinci Delphi anketi: Birinci Delphi uygulamasında alan yazındaki belli başlı öğretmen yeterlikleri çalışmaları ve Delphi tekniği ile ilgili açıklamaların ardından açık uçlu bir soru sorulmuştur. Birinci Delphi uygulamasında "Her öğretilerde bulunması gereken ölçme ve değerlendirme öğretmen yeterlikleri sizce nelerdir?" sorusuna bütün katılımcıların belirttiği görüşler maddeler halinde sıralanmıştır.

Birinci Delphi anketi için öğretmenlere açık uçlu olarak hazırlanan ve öğretmen yeterliği ile ilgili bilgileri de içeren bir form verilmiştir. Öğretmenlerin diledikleri sayıda yeterlik ifadesi yazmaları sağlanmıştır.

"Lütfen, kişisel bir beyin fırtınası gerçekleştirerek konuya ilişkin mümkün olduğu kadar çok sayıda yeterliği belirleyip listeleyiniz. Belirlediğiniz bu yeterlik maddeleri diğer katılımcıların belirttikleriyle bir araya getirilip İkinci Delphi Anketi oluşturulacaktır."

Yeterlik 1:

Yeterlik 2:

Yeterlik 3:

.

.

.

Yeterlik 10:

Birinci delphi anketinin cevaplanması: Her bir katılımcı, sorulan soruya ilişkin düşüncelerini listelemesi istenmiş ve araştırmacı tarafından toplanmıştır. Katılımcılar düşüncelerini cümleler halinde ifade etmiştir. Anketlerin toplanmasıyla birinci Delphi uygulaması tamamlanmıştır. Birinci Delphi anketi sonrası toplanan veriler aynı anlama gelen birbirine yakın anlam taşıyan ifadeler bir araya getirilmiş ve Ek 3'te verilen yeterlik maddeleri halinde sunulmuştur.

İkinci Delphi Anketinin Düzenlenmesi ve Gönderilmesi: Birinci Delphi uygulamasında katılımcıların belirttiği görüşler maddeler halinde sıralanmış ve gerekirse alt başlıklar

halinde toplanmıştır. Bu şekilde oluşturulmuş anket formu niteliğinde olan İkinci Delphi anketi hazırlanmış ve katılımcılara sunulmuştur. Araştırmada İkinci Delphi uygulaması için ikili karşılaştırmalar ve sıralama yargularıyla ölçeklemeye uygun duruma getirilen form öğretmenlere verilmiştir.

İkinci delphi anketi: İlk turda toplanan görüşlere dayalı olarak hazırlanan maddeler ikili ve sıralı karşılaştırma ile ölçeklemeye uygun hale getirilmiştir. İkinci Delphi formunda, panel üyelerinin, maddeleri ölçeklemesi istenmiştir. Bu ölçekleme formu kolaylıkla yanıtlanabilecek ölçek maddeleriyle hazırlanmıştır. Ölçek maddeleri arasına işaret konulabilecek biçimde düzenlenmiştir. İkinci Delphi Anketi, hazırlanırken binişik görüşler ayıklanmış; gelen görüşler doğrultusunda gruplandırma yapılmıştır.

İkinci delphi anketinin cevaplanması: Katılımcılar her bir maddenin önem düzeyini ya da her bir maddeye katılma düzeylerini ikili karşılaştırma ve sıralama yargularıyla ölçekleyerek belirtmiştir. Öğretmen yeterlikleri ilgili eksik buldukları ve yer almasını istedikleri yeterlik alanını yorum ve açıklama ile yazmışlardır. Bu şekilde İkinci Delphi Uygulaması tamamlanmıştır.

İkinci delphi anketinin analizi: Yeterlik alanlarına ilişkin ölçekleme işlemi uygulama sonunda yapılır ve katılımcılara iletmek üzere her bir katılımcı için hazırlanan formlara eklenmiştir.

Üçüncü delphi anketinin düzenlenmesi ve gönderilmesi: Delphi tekniğinin üçüncü turunda, ikinci turda verilen yanıtlara göre her bir maddeye ilişkin sunulan ölçek değerleri belirlenerek maddelerin tekrar yanıtlanması sağlanmıştır. Üçüncü tur sonucunda öğretmenlerde olması gereken yeterlikler belirlenmiş, ikinci ve üçüncü Delphi uygulamaları arasındaki ilişkinin belirlenmesiyle sonuçlar Delphi üyelerinin tamamına duyurulmuştur.

Üçüncü delphi anketi: Üçüncü Delphi anketinde ikinci ankete verilen cevapların gözden geçirilmesi istenmiştir. Öğretmenlere ikinci Delphi uygulamasındaki formun aynısını üçüncü ankette de olduğunun bilgisi verilmiştir. Öğretmenlere sunulan bu ankete ikinci anketin istatistiksel analizleri de eklenmiştir. Her bir madde başında ikinci ankette yer alan öğretmene ait cevaplar yazılmıştır. Üçüncü Delphi formunda yanıtların yanında grubun cevaplarına ilişkin iki farklı ölçekleme sonucu da yer almıştır.

Üçüncü delphi anketinin cevaplanması: Üçüncü Delphi anketi istatistiksel sonuçlarla birlikte tekrar katılımcılara iletilmiştir. Katılımcılardan, ikinci ankete verdikleri cevapları gruba ilişkin istatistiklerle karşılaştırmaları, her bir maddeye ilişkin yapılan yorumları da değerlendirerek kararlarını yeniden gözden geçirmeleri istenmiştir. Eski kararında ısrar ediyorsa aynı ölçekleme işlemini yapmışlardır. Bu şekilde Üçüncü Delphi Uygulaması da tamamlanmıştır.

Üçüncü delphi anketinin analizi ve uygulamaların sonuçlandırılması: Üçüncü Delphi anketinin analizi için ikinci Delphi uygulamasında yapılan istatistikler kullanılmıştır. İkili karşılaştırma ve sıralama yargularına dayanan ölçekleme işleminde ölçek değerlerine bakılmış ve ikinci Delphi uygulamasından farklı olup olmadığı değerlendirilmiştir. Eğer ölçek değerlerindeki değişme sıralamayı etkiliyorsa bir kez daha form hazırlanmış ve yeniden bir ölçekleme yapılmaları istenmiştir.

Araştırma için öğretmenlerin ölçme ve değerlendirme alanı yeterliklerini belirlemeleri için Delphi tekniğine göre hazırlanan ve üç farklı oturumda gerçekleşen anketler kullanılmıştır. Ayrıca öğretmen özelliklerini ortaya koymayı amaçlayan bir anket de öğretmenlere uygulanmıştır.

Verilerin Çözülmesi

Ölçekleme ile ilgili yapılan çalışmalarda güvenilirlik için ölçek değerleri arasındaki tutarlık ve uyumun sağlanması gerekmektedir. Ölçek değerleri belirlenirken bazı varsayımlarda bulunulmuştur. Ayrıca elde edilen ölçek değerleri gözleyicilerin yargılarına dayandırılmıştır. Varsayımların doğru olup olmadığının ve yargıcıların dikkatli davranıp davranmadıklarının kontrol edilmesi gereklidir. Bu kontroller için ölçeklemenin iç tutarlılığı belirlenir. Ölçek değerlerinin iç tutarlılığının belirlenmesi, gözlenen p_{jk} oranlarının, ölçek değerlerinden elde edilen (ölçeklemeden beklenen) p_{jk} değerleriyle ne dereceye kadar bağdaştığının belirlenmesi yoluyla yapılır.

Tutarlık Ölçüsü

Ölçek değerinin iç tutarlığının hesaplanması, gözlenen p_{jk} değerleriyle ne kadar uyumlu olduğu hesaplanarak yapılmaktadır. Bu amaçla uygulama verilerinin analizinden elde edilen ölçek değerlerinden hareket edilerek bir Z' birim normal sapmalar matrisi ve bu matristen de teorik oranlar matrisi elde edilir. Teorik oranlar ile gözlenen oranlar arasındaki uygunluğa bakılır. Bu durumda öncelikle ölçek değerlerinin ortalama hatası hesaplanır. Elde edilecek küçük bir ortalama hata gözlemci yargılarının güvenilir olduğuna, büyük bir hata değeri ise gözlemci yargılarının güvenilir olmadığına işaret etmektedir. Büyük bir hata değeri ya modelin varsayımlarının sağlanmadığını ya gözlemci yargılarının güvenilir olmadığına ya da her iki hata kaynağının da etkili olduğunu göstermektedir. Ortalama hata değeri gözlenen değerlerle teorik değerler arasındaki uygunluğun ortalama değerini, yani uyumun bir ölçüsünü verir. Bu uygunluk derecesinin anlamlı olup olmadığı ise kay kare (chi-square) testi ile yoklanır. Bu test aşağıda verilen bağıntılar kullanılarak yapılır.

$$\chi^2 = \frac{\sum_{j>k} (q_{jk} - q'_{jk})^2}{\frac{821}{N}}$$

$$sd = \frac{K(K-1)}{2}$$

Yukarıdaki iki bağıntıda χ^2 test istatistiği, sd serbestlik derecesi, q_{jk} gözlenen oranlar matrisinin normalleştirilmiş değeri, q'_{jk} teorik oranlar matrisinin normalleştirilmiş değeri, N her bir uyarıcı için toplam yargı sayısı ve K uyarıcı sayısını ifade etmektedir. Elde edilen kay kare değeri ilgili serbestlik derecesinde belirlenen anlamlılık düzeyine göre tablo değeri ile karşılaştırılır. Bulunan kay kare değeri tablo değerinden küçük çıkması ölçeklemenin iç tutarlığının olduğunun göstergesidir.

Ölçek değerinin iç tutarlılığı

İç tutarlığı belirlemek için, verilerden elde edilen ölçek değerlerinden hareket edilerek bir Z birim normal sapmalar matrisi ve bu matristen teorik oranlar matrisi elde edilir. Teorik oranlar ile gözlenen oranlar arası uyuma bakılır. Bunun için ortalama hata formülü kullanılır.

$$OH = \frac{\sum_{j>k} |p_{jk} - p'_{jk}|}{K(K-1)}$$

OH: Gözlenen oranla teorik oran arasındaki farkın ortalama değeri

p_{jk} : Gözlenen frekanslardan elde edilen oran

p'_{jk} : Teorik oran

K: uyarıcı sayısı

Bu eşitlikten elde edilecek düşük ortalama hata ölçek değerlerinin güvenilir olduğuna, büyük bir hata gözlemci yargılarının güvenilir olmadığına, ya modeldeki varsayımların sağlanmadığına ya da bilinmeyen derecelerde olmak üzere her iki hata kaynağının da etkili olduğuna işaret eder.

Farkın manidar çıkması;

1. Varsayımlar işleminde olabilir,

2. Gözlemciler dikkatli gözlem yapmadılar,

3. Gözlenen değişken tek boyutlu değil ve normal dağılım göstermiyor,

şeklinde yorumlanabilir.

Yukarıdaki ortalama hata gözlenen değerlerle ampirik değerler arasındaki uyumun ortalama değerini, yani uyumun bir ölçüsünü verir. Fakat bunun anlamlı olup olmadığı hakkında bilgi vermez. Bunun için kay kare testi kullanılır.

$$\chi^2 = \frac{\sum_{j>k} (q_{jk} - q'_{jk})^2}{821/N}$$

sd= K(K-1)/2

χ^2 : Gözlenen kay kare değeri

sd: Serbestlik derecesi

q_{jk} : Gözlenen oranlar matrisinin j sütun ve k satırındaki p_{jk} oranına ait normalleştirilmiş değer

q'_{jk} : Teorik oranlar matrisinin j sütun ve k satırındaki p_{jk} oranına ait normalleştirilmiş değer

N: Her bir uyarıcı sayısı için toplam yargı sayısı

K: Uyarıcı sayısı

P oranlar matrisi için: $q = \arcsin \sqrt{p}$

q: p oranına ait normalleştirilmiş değer

p: Oran

elde edilen q değerinin dağılımı varyansı

$$\sigma_q^2 = 821/N$$

Açıklamalardan yola çıkarak teorik birim normal sapmalar matrisi oluşturulup ortalama hata miktarları Tablo 2'te verilmiştir.

Tablo 2. Ölçek değerlerinin içtutarlıklarına ait bulgular (n=300)

Ölçekleme	Hata	Ortalama Hata (OH)	Sd	χ^2
2. Dephi uygulaması ikili karşılaştırma ile ölçekleme	1,96	0,0178	55	35,3452
2. Delphi uygulaması sıralama ile ölçekleme	1,78	0,0161	55	33,4771
3. Dephi uygulaması ikili karşılaştırma ile ölçekleme	1,94	0,0176	55	33,1659
3. Dephi uygulaması sıralama ile ölçekleme	1.64	0,0149	55	33,1826

$$\chi^2_{(0.05;55)}=73,311$$

Tablo 4'te verilen ortalama hata değerlerine bakıldığında her dört ölçekleme işlemindeki ortalama hata oranlarının çok küçük olduğu görülmektedir.

Matematiksel modelin ampirik verilere uygunluğunun yoklanması

Gözlenen oranlar matrisinden $q = \arcsin \sqrt{p}$ dönüşümü yapılarak gözlenen ve beklenen teorik oranlar matrisleri çıkarılmıştır. Dönüştürülmüş Gözlenen ve Teorik Oranlar Farkları Matrisi $P(q_k - q_j k')$ 'nden elde edilen toplam kay kare değeri Tablo 3'te görüldüğü şekilde bulunmuştur. Tablo 3'te görüldüğü gibi ölçeklemelere ait ortalama hata değerleri oldukça düşüktür.

İkinci ve üçüncü Delphi uygulaması sonucu elde edilen ölçek değerlerine ait kay kare değerlerinin ise tablo değerinden ($\chi^2_{0.05;55} = 73,311$) küçük olması ölçek değerlerinin içtutarlığa sahip olduğunu göstermektedir. Belirlenen kay kare değerleri ve ortalama hatalar dikkate alındığında yargıcı kararlarında tutarlığın bir başka deyişle uyumun ya da yöntemin ölçeklemeye uygun olduğunu göstermektedir. İkinci ve üçüncü Delphi uygulamalarından elde edilen ölçekleme verileri ikili karşılaştırma ve sıralama ile ölçekleme yapılmasına istatistiksel olarak olanak sağlamaktadır.

BULGULAR

Araştırmaya katılan öğretmenlere açık uçlu olarak sorulan "İyi bir öğretmende olması gereken ölçme ve değerlendirme alanında gereksinim duydukları öğretmen yeterlikleri nelerdir?" sorusuna istedikleri kadar ifade yazmaları istenmiş ve bu ilk delphi paneli oturumu olarak kayıt altına alınmıştır.

Birinci Delphi paneli sonrası araştırma grubunda yer alan öğretmenlerin açık uçlu soruya verdikleri yanıtlar doğrultusunda nitel veriler elde edilmiştir. Toplamda 1120 farklı görüş 222 başlık halinde öğretmenler tarafından ölçme ve değerlendirme alanında birer yeterlik boyutu olarak ifade edilmiştir. İfadeler, ölçekleme teknikleri, ölçme ve değerlendirme alanında uzman üç akademisyen ile ölçme ve değerlendirme konusunda deneyimli iki öğretmen tarafından bir araya getirilerek ve 11 ana yeterlik boyutunda toplanarak ölçeklemeye uygun duruma getirilmiştir. Gruplandırma yapılırken ifadelerin benzerliği, ortak bir anlam çıkarımı ve uzmanların önerileri dikkate alınarak ortak bir görüşe varılmıştır. Uzmanların görüşleri birbiriyle bağlantılı olarak ele alınmış ve bir içerik analizi yerine ortak bir çözümleme (gruplandırma) yolu seçilmiştir. Yeterlik ifadelerinin kavramsallaştırılmasında ortak anlam içeren birbiriyle ilişkili maddeler aynı grupta toplanmış ve öğretmenlere sunulmuştur. Yeterlik maddeleri için uzlaşılan ve uzlaşılmayan ifadelerin de uygulamalarda yer alması ve bunun araştırma grubunda yer alan tüm öğretmenlerin görüşüne sunulması Delphi uygulamasının bir yararı olarak araştırmada yer almıştır.

Yeterlik alanları belirlenmesi sürecinde belirtilen görüşler, uzmanların önerileriyle grubu en iyi temsil edecek ifadeler seçilerek yazılmıştır. İkinci ve üçüncü Delphi

uygulamalarında yalnızca yeterlik başlığı değil alt başlıklar da katılımcılara verilmiştir. Ölçme ve değerlendirme alan yeterliklerinin bu biçimde gruplandırılmasının uygulamada getireceği kolaylık ve anlaşılabilirlik göz önünde bulundurularak böyle bir yol izlenmiştir. Belirlenen bu yeterlik boyutları temsil ettikleri gruplara göre 11 farklı yeterlik boyutunda toplanmış ve ikili karşılaştırma ve sıralama yargılarına dayalı ölçeklemeye hazır duruma getirilmiştir.

Delphi paneli katılımcıları tarafından belirlenen iyi bir öğretmende olması gereken ölçme ve değerlendirme alanında gereksinim duydukları öğretmen yeterlik boyutları aşağıdaki şekilde belirlenmiştir.

- Öğrencinin gelişim düzeylerini izleme ve buna uygun ölçme ve değerlendirme tekniklerini kullanma.
- Performans, proje ve portfolyaya dayalı ölçmeler yapma.
- Bireysel farklılıkları dikkate alan ölçme ve değerlendirme tekniklerini uygulama.
- Ölçme sonuçlarını geribildirim olarak kullanma.
- Bilişsel, duyuşsal ve psikomotor alanda ölçme araçları hazırlama.
- Objektif testler kullanma.
- Öğrencinin öz değerlendirme ve akran değerlendirme yapmasına olanak tanıma.
- Teknolojik araç ve gereçleri ölçme ve değerlendirme yaparken kullanma.
- Ölçütleri belirleme öğrenciyi bu ölçütlerden haberdar etme.
- Ölçme sonuçlarına göre eğitim programını değiştirme.
- Farklı soru formatları kullanma.

Birinci delphi uygulaması sonucu belirlenen yeterlikler gruplandırıldıktan sonra her bir grubu temsil edeceği düşünülen ifadeler ikili karşılaştırma ile sıralama yargılarına dayalı ölçeklemeye hazır duruma getirilmiştir. İkili karşılaştırmalar ve sıralama yargılarıyla ölçekleme formu öğretmenlere verilmiştir. Öğretmenler ilgili yeterlik boyutunu alt maddeleriyle görerek birinci ve ikinci delphi uygulamasını gerçekleştirmişlerdir.

2. Delphi paneli uygulaması için;

İkinci Delphi paneli uygulaması için: “İkili karşılaştırma yargılarına dayalı ölçekleme yönteminden elde edilen madde ölçek değerleri nasıldır?” sorusuna ilişkin bulgular.

Yukarıda belirtilen araştırma sorusuna yanıt bulmak için 11 yeterlik maddesine ait uyarıcılar ikişerli gruplar halinde eşlenerek, cevaplayıcıların her bir grupta tercih ettikleri alanı belirtmeleri istenmiştir. Cevaplayıcılara ait yargıların ($s_j > s_k$) frekansları Tablo 3'deki gibi, 11 satır ve 11 sütunlu bir ham puanlar matrisinde (F matrisi) toplanmıştır.

Her bir öğretmenin beş kompozisyonu ikili karşılaştırma yaparak sıralamaları sonucu oluşturulan frekans matrisi Tablo 3'de gösterildiği şekilde oluşturulmuştur. Burada yer alan her bir hücredeki eleman f_{jk} ile gösterilmek üzere $s_j > s_k$ değerini veren frekansları göstermektedir. Her yargıcı 11 yeterlik maddesini ölçeklediğinden F matrisinde 300 yargıcının yaptığı tüm değerlendirmelerin sayısı 16500'tür. Bu bize matrisin düzenlenmesinde bir hata yapılmadığını göstermektedir. Matris esas köşegene göre simetriktir ve simetrik elemanların toplamı toplam yargıcı sayısını verir.

Tablo 3. 11 yeterlik maddesine ait ham puanlar matrisi [$F(n(s_j > s_k))$]

U _k	U _j											Toplam
	A	B	C	D	E	F	G	H	I	J	K	
A		145	148	166	169	130	145	164	167	150	161	1545
B	155		152	155	164	140	160	173	155	156	177	1587
C	152	148		168	170	133	156	168	155	144	145	1539
D	134	145	132		166	144	165	164	155	140	155	1500
E	131	136	130	134		146	69	167	155	134	143	1345
F	170	160	167	156	154		144	140	155	166	133	1545
G	155	140	144	135	231	156		155	133	111	200	1560
H	136	127	132	136	133	160	145		123	144	122	1358
I	133	145	145	145	145	145	167	177		130	155	1487
J	150	144	156	160	166	134	189	156	170		120	1545
K	139	123	155	145	157	167	100	178	145	180		1489
Toplam	1455	1413	1461	1500	1655	1455	1440	1642	1513	1455		16500

Frekans (F) matrisinin her bir hücresindeki değer toplam kişi sayısına (N) bölünerek Tablo 4'deki (P) oranlar matrisi elde edilmiştir. Bu tablo frekans matrisinin elemanlarının 300 kişi sayısına bölünmesiyle oluşturulmuştur. P oranlar matrisinin elemanları esas köşegene göre simetrik olduğundan simetrik elemanların toplamı 1'e eşittir.

Tablo 4. Yeterlik maddesine ilişkin p oranlar matrisi (P_{jk})

U _k	U _j										
	A	B	C	D	E	F	G	H	I	J	K
A		0,483	0,493	0,553	0,563	0,433	0,483	0,547	0,557	0,500	0,537
B	0,517		0,507	0,517	0,547	0,467	0,533	0,577	0,517	0,520	0,590
C	0,507	0,493		0,560	0,567	0,443	0,520	0,560	0,517	0,480	0,483
D	0,447	0,483	0,440		0,553	0,480	0,550	0,547	0,517	0,467	0,517
E	0,437	0,453	0,433	0,447		0,487	0,230	0,557	0,517	0,447	0,477
F	0,567	0,533	0,557	0,520	0,513		0,480	0,467	0,517	0,553	0,443
G	0,517	0,467	0,480	0,450	0,770	0,520		0,517	0,443	0,370	0,667
H	0,453	0,423	0,440	0,453	0,443	0,533	0,483		0,410	0,480	0,407
I	0,443	0,483	0,483	0,483	0,483	0,483	0,557	0,590		0,433	0,517
J	0,500	0,480	0,520	0,533	0,553	0,447	0,630	0,520	0,567		0,400
K	0,463	0,410	0,517	0,483	0,523	0,557	0,333	0,593	0,483	0,600	

Oranlar matrisindeki hücre değerlerine (P) karşılık gelen (Z) standart değeri belirlenerek Tablo 5'de gösterilen birim normal sapmalar matrisi elde edilmiştir. P matrisinin elemanları esas köşegene göre simetrik olduğundan Z matrisinin esas köşegene göre simetrik olan elemanları mutlak değerce eşit, işaretçe zıttır. Bu nedenle Z matrisinin elemanları toplamı sıfıra eşittir. Matrisin sonunda her bir sütuna ait değerlerin toplamını gösteren bir satır oluşturulmuş ve bu satırdaki her bir Z hücre değerinin sütunlar boyunca ortalamaları alınarak yani 11'e bölünerek ölçek değeri hesaplanmıştır. Eksenin başlangıcı (0 noktası) bu satırdaki ortalama Z değerinden en küçük olanına kaydırılarak ölçek değerleri (S) sıralanmıştır.

Birim normal dağılım fonksiyonunun Z değerlerini P oranlarına bağlı olarak veren tablodan, P matrisinin elemanlarına karşılık gelen Z değerlerini bularak Tablo 5'de verilen birim normal sapmalar (Z) matrisi elde edilmiştir.

Tablo 5. Birim normal sapmalar matrisi (Z matrisi)

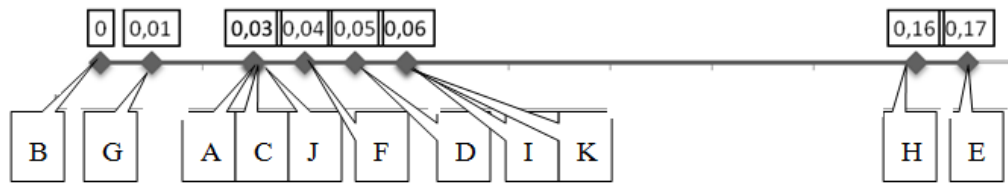
	U _j										
U _k	A	B	C	D	E	F	G	H	I	J	K
A		-0,042	-0,017	0,134	0,159	-0,168	-0,042	0,117	0,143	0,000	0,092
B	0,042		0,017	0,042	0,117	-0,084	0,084	0,193	0,042	0,050	0,228
C	0,017	-0,017		0,151	0,168	-0,143	0,050	0,151	0,042	-0,050	-0,042
D	-0,134	-0,042	-0,151		0,134	-0,050	0,126	0,117	0,042	-0,084	0,042
E	-0,159	-0,117	-0,168	-0,134		-0,033	-0,739	0,143	0,042	-0,134	-0,059
F	0,168	0,084	0,143	0,050	0,033		-0,050	-0,084	0,042	0,134	-0,143
G	0,042	-0,084	-0,050	-0,126	0,739	0,050		0,042	-0,143	-0,332	0,431
H	-0,117	-0,193	-0,151	-0,117	-0,143	0,084	-0,042		-0,228	-0,050	-0,236
I	-0,143	-0,042	-0,042	-0,042	-0,042	-0,042	0,143	0,228		-0,168	0,042
J	0,000	-0,050	0,050	0,084	0,134	-0,134	0,332	0,050	0,168		-0,253
K	-0,092	-0,228	0,042	-0,042	0,059	0,143	-0,431	0,236	-0,042	0,253	
ΣZ _{jk}	-0,034	-0,063	-0,028	-0,012	0,109	-0,019	-0,048	0,098	-0,003	-0,035	0,001
Σ _{jort.}	0,030	0,000	0,030	0,050	0,170	0,040	0,010	0,160	0,060	0,030	0,060
S _j	a	b	c	d	e	f	g	h	ı	j	k

Tablo 5'te görüldüğü gibi Z değerlerinin en küçüğü b olan -0,063'tür. Eksenin başlangıcı (0 noktası) doğrusal dönüştürme ile yani bulunan en küçük değer tüm uyarıcı değerlerine eklenmesi ile sıfır noktasına kaydırılarak S_j değerleri bulunmuş olur. S_j değerleri en küçükten en büyüğe doğru sıralanarak S_c ölçek değerleri Tablo 6'da sunulduğu şekliyle elde edilmiştir.

Tablo 6. Ölçek değerlerinin sıralanması

U _j	a	b	c	d	e	f	g	h	ı	J	k
S _c	0,03	0,00	0,03	0,05	0,17	0,04	0,01	0,16	0,06	0,03	0,06

İkili karşılaştırma yargıları ile ölçekleme değerlerinde yeterliklerin en küçükten en büyüğe sıralanmasının sayı doğrusunda gösterimi aşağıda verilmiştir.



Şekil 2. İkili karşılaştırma yargılarıyla ölçeklemenin sayı doğrusu üzerinde gösterimi

Ölçme ve değerlendirme alanı öğretmen yeterlikleri ikinci Delphi uygulamasından elde edilen verilerin 5. Durum denklemi ile ölçeklenmesi sonucu yeterlik maddelerinin tercih edilmeleri Tablo 6'da verilmiştir. Öğretmenler, "Bilişsel, duyuşsal ve psikomotor alanda ölçme araçları hazırlar" yeterliğini daha önemli görmektedir. En az önemli görülen yeterlik ise "Performans, proje ve portfolyaya dayalı ölçmeler yapar" olarak belirlenmiştir.

Tablo 3 ve 4'te verilen ve uygulama yapılan gruplara ait ölçekleme sonuçlarının 2. Delphi uygulaması için içtutarlılığının belirlenmesi amacıyla kay kare analizi yapılmıştır. Bu analizde elde edilen ölçek değerlerinin matematiksel modelin ampirik verilere uygunluk derecesine bakılması sonucunda ölçeklemenin iç tutarlılığı bulunmuştur. Analizleri yapılan

her bir ölçeklemeye ait ortalama hata ve χ^2 istatistik değerleri Tablo 2’te verilmiş ve ölçeklemeye uygun veriler olduğu gözlenmiştir. Ölçek değerlerinin içtutarlılığının sağlanması gözlenen p_{jk} oranlarının, ölçek değerinden elde edilen p_{jk} değerleriyle ne dereceye kadar bağdaştığının da kanıtıdır. Elde edilen ortalama hata değerinin düşük olması ölçek değerlerinin güvenilir olduğunu yüksek bir ortalama hata değeri ise gözlemci yargılarının güvenilir olmadığını veya modeldeki varsayımların sağlanmadığını ya da bilinmeyen derecelerde olmak üzere her iki hata kaynağının da etkili olduğunu gösterir. Araştırmada gözlenen değerler ile ampirik değerler arasında bir uyumun olduğu belirlenmiştir. Bu uyum değeri araştırmanın güvenilir sonuçlar verdiğinin göstergesidir.

2. Delphi paneli uygulaması için;

“İkinci Delphi panelinde sıralama yargılarına dayalı ölçekleme yönteminden elde edilen madde ölçek değerleri ne düzeydedir?” sorusuna ilişkin bulgular:

11 yeterlik maddesine ait uyarıcılar 300 gözlemcinin belirli bir boyutta sıralamalarıyla elde edilen veriler N satırlı ve K sütunlu bir matriste toplanmıştır. Tablo 7’de 300 gözlemcinin 11 uyarıcıyı büyükten küçüğe doğru sıralamasına ilişkin sıralamalar matrisi verilmiştir.

Tablo 7. Sıralamalar matrisi (R_i)

R_i	R_i	a	b	c	d	e	f	g	h	ı	j	k	Toplam
1	11	20	52	40	28	40	16	12	38	12	32	10	300
2	10	28	32	36	24	48	28	8	32	16	24	24	300
3	9	44	40	20	52	16	30	19	20	16	11	32	300
4	8	32	32	28	20	32	20	25	20	45	18	28	300
5	7	36	40	24	28	16	28	36	16	40	12	24	300
6	6	40	24	36	40	28	18	30	12	32	16	24	300
7	5	28	32	20	24	20	44	32	16	26	38	20	300
8	4	20	20	16	16	20	36	40	30	45	33	24	300
9	3	32	8	28	12	32	32	40	56	36	12	12	300
10	2	16	12	24	28	8	8	38	32	20	50	64	300
11	1	4	8	28	28	40	40	20	28	12	54	38	300
N		300	300	300	300	300	300	300	300	300	300	300	

Tablo 7’de verilen sıra frekanslarından $n(S_{ji}>S_{ki})=f_{ji}(f_{k<i}+1/2f_{ki})$ formülü ile her bir ölçekleme için $n(S_{ji}>S_{ki})$ frekansları Tablo 16’teki gibi bir matriste toplanmıştır. Sıralama yargılarına dayalı ölçekleme için oluşturulan Tablo 20’deki matris kullanılarak 300 gözlemcinin 11 uyarıcı için yapılan sıralamalardan her bir uyarıcının diğerlerinden kaç defa büyük olarak yargılandığına dair oranlar kullanılmıştır. Bu oranların kestirilmesi için, U_1, U_2, \dots, U_k K tane uyarıcı, r_1, r_2, \dots, r_k bu uyarıcılara her bir gözlemci tarafından verilen sıra sayıları ve $r_1>r_2>\dots>r_k$ olmak üzere U_j ve U_k gibi iki uyarıcı için yapılabilecek tüm ikili karşılaştırmalarının sayısı bulunmuştur.

Bu karşılaştırmaların içinde $S_j=S_k$ yargılarının da bulunabileceği dikkate alınarak bu eşit yargıların yarısının $S_j>S_k$ ve diğer yarısının $S_j<S_k$ olacağı $n(S_j>S_k)=n(S_j<S_k)$ sayılısı altında $S_j>S_k$ yargılarının bir r_i sırası için toplam sayısı, $n(S_{ji}>S_{ki})=f_{ji}(f_{k<i}+1/2f_{ki})$ formülü ile bulunmuştur. Burada;

$f_{k<i}$, U_k uyarıcısına r_i den küçük sıra değeri verenlerin sayısı olarak alındığında;

$$\sum_{i=1}^n n(S_{ji} > S_{ki}) = \sum_{i=1}^n f_{ji} (f_{k<i} + \frac{1}{2} f_{ki})$$

f_{ji} : U_j uyarıcısına r_i sıra değerinin verilme sayısı

f_{ki} : U_k uyarıcısına r_i sıra değerinin verilme sayısı

Bağıntısı ile U_j nin U_k dan büyük olduğunu gösteren frekanslar matrisi oluşturulmuştur. Sıralama yargıları ile ölçekleme yönteminde elde edilen frekanslardan sıra frekansları matrisi oluşturmak için yukarıda verilen formül uygulanarak tablolar oluşturulmuştur. Ölçekleme sıraları frekans matrisinin oluşturulması için her bir uyarıcıya ilişkin oluşturulan frekanslar matrisinin sütun toplamları alınmış, toplamlar Tablo 8’de gösterilmiştir.

Tablo 8. Toplam frekanslar matrisi

	A	B	C	D	E	F	G	H	I	J
B	37424									
C	46672	52992								
D	46592	53280	45216							
E	45632	52008	44432	44080						
F	53244	59688	50696	51084	51132					
G	58766	64768	54612	55480	54988	85867				
H	52368	57888	49600	49500	50012	45324	85872			
I	53198	60194	49886	50754	50638	44090	85413,5	86015		
J	58080	62778	54866	55070	54678	50103	85928,5	85835	86350,5	
K	56824	62228	53956	53920	54304	48720	85762	86078	86114	85314

Tablo 8’de verilen frekanslar toplamı N^2 ’ye bölünerek oranlara dönüştürülmüş ve aşağıdaki formül yardımıyla oranlar matrisi elde edilmiştir.

$$P_{j>k} = \frac{\sum_{i=1}^n f_{ji} (f_{k<i} + \frac{1}{2} f_{ki})}{N^2}$$

Bu formül kullanılarak sıralama yargılarıyla ölçekleme için gerekli olan oranlar matrisi oluşturulmuştur (Tablo 9).

Tablo 9. Sıralama yargılarıyla ölçekleme oranlar matrisi (p)

	A	B	C	D	E	F	G	H	I	J	K
A		0,584	0,481	0,482	0,493	0,408	0,347	0,418	0,409	0,355	0,369
B	0,416		0,411	0,408	0,422	0,337	0,280	0,357	0,331	0,302	0,309
C	0,519	0,589		0,498	0,506	0,437	0,393	0,449	0,446	0,390	0,400
D	0,518	0,592	0,502		0,510	0,432	0,384	0,450	0,436	0,388	0,401
E	0,507	0,578	0,494	0,490		0,432	0,389	0,444	0,437	0,392	0,397
F	0,592	0,663	0,563	0,568	0,568		0,046	0,496	0,510	0,443	0,459
G	0,653	0,720	0,607	0,616	0,611	0,954		0,046	0,051	0,045	0,047
H	0,582	0,643	0,551	0,550	0,556	0,504	0,954		0,044	0,046	0,044
I	0,591	0,669	0,554	0,564	0,563	0,490	0,949	0,956		0,041	0,043
J	0,645	0,698	0,610	0,612	0,608	0,557	0,955	0,954	0,959		0,052
K	0,631	0,691	0,600	0,599	0,603	0,541	0,953	0,956	0,957	0,948	

P oranlarının bulunmasının ardından birim normal dağılım fonksiyonunun Z değerini P oranlarına bağlı olarak veren tablodan P matrisinin elemanlarına karşı gelen Z değerleri bulunarak, birim normal sapmalar matrisi (Z) elde edilmiştir. Birim normal sapmalar matrisi Tablo 10'da verilmiştir.

Tablo 10. Birim normal sapmalar matrisi (z)

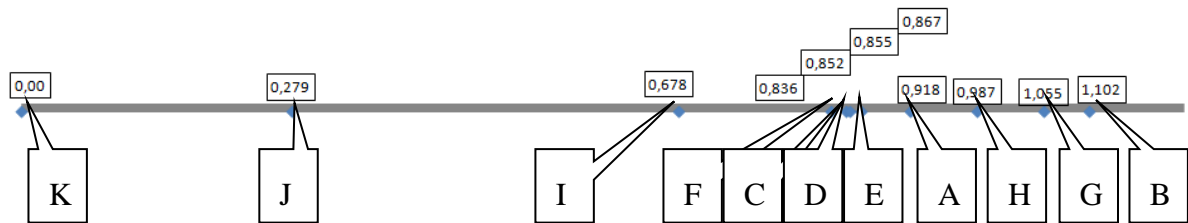
	A	B	C	D	E	F	G	H	I	J	K
A		0,213	-0,047	-0,044	-0,018	-0,232	-0,393	-0,207	-0,230	-0,373	-0,336
B	-0,213		-0,224	-0,233	-0,196	-0,421	-0,582	-0,367	-0,437	-0,517	-0,500
C	0,047	0,224		-0,006	0,016	-0,159	-0,271	-0,128	-0,137	-0,278	-0,252
D	0,044	0,233	0,006		0,026	-0,170	-0,296	-0,126	-0,161	-0,284	-0,251
E	0,018	0,196	-0,016	-0,026		-0,172	-0,282	-0,140	-0,158	-0,273	-0,262
F	0,232	0,421	0,159	0,170	0,172		-1,686	-0,009	0,025	-0,143	-0,104
G	0,393	0,582	0,271	0,296	0,282	1,686		-1,686	-1,636	-1,693	-1,674
H	0,207	0,367	0,128	0,126	0,140	0,009	1,686		-1,703	-1,682	-1,711
I	0,230	0,437	0,137	0,161	0,158	-0,025	1,636	1,703		-1,744	-1,715
J	0,373	0,517	0,278	0,284	0,273	0,143	1,693	1,682	1,744		-1,625
K	0,336	0,500	0,252	0,251	0,262	0,104	1,674	1,711	1,715	1,625	
Toplam	3,720	5,201	3,274	2,514	1,204	-2,977	-1,772	-2,581	-2,720	-2,667	-2,968
Uyarıcı Sayısı	11	11	11	11	11	11	11	11	11	11	11
S _j	1,666	3,690	0,945	0,980	1,114	0,762	3,179	2,433	-0,976	-5,362	-8,429
S _c	0,151	0,335	0,086	0,089	0,101	0,069	0,289	0,221	-0,089	-0,487	-0,766

Z matrisine göre ölçek değerlerinin bulunması amacıyla birim normal sapmalar matrisinde her bir uyarıcının yeri S_c ile belirtilmiştir. Ölçek değerlerinin en küçüğü sıfıra ötelenmiştir. Bu işlemde en küçük değer -0.766 ile K değeridir. Bu nedenle tüm ölçek değerlerine +0.766 eklenerek öteleme yapılır. Bu işlem sonucunda elde edilen ölçek değerleri Tablo 11'de gösterilmiştir.

Tablo 11. Sıralama yargılarıyla ölçeklemede ölçek değerlerinin sıralaması

A	B	C	D	E	F	G	H	I	J	K
0,918	1,102	0,852	0,855	0,867	0,836	1,055	0,987	0,678	0,279	0,000

Sıralama yargıları ile ölçekleme değerleri ile yeterliklerin en küçüğe en büyüğe sıralanmasının sayı doğrusunda gösterimi Şekil 3'te verilmiştir.



Şekil 3. Sıralama yargılarıyla ölçeklemenin sayı doğrusu üzerinde gösterimi

Sıralama yargılarıyla ölçekleme sonucu elde edilen ölçek değerlerine bakıldığında performans, projeye ve portfolyaya dayalı ölçmeler yapmayı en önemli yeterlik olarak belirtirken, en az önemli görülen yeterlik ise farklı soru formatları kullanma olarak belirlenmiştir.

İkinci Delphi uygulaması sonucu elde edilen ikili karşılaştırma ve sıralama yargılarıyla ölçekleme bulguları incelendiğinde Tablo 12’te belirtildiği gibi yeterlik ifadelerinin ölçek değerleri farklı ve sıralamanın da farklı olduğu bulunmuştur.

Tablo 12. İkinci delphi panelinde ikili karşılaştırma ve sıralama yargılarıyla ölçekleme sonucu ölçek değerlerinin karşılaştırılması

	Yeterlik İfadeleri	İkili Karşılaştırma		Sıralama	
		Ölçek Değeri	Sıra	Ölçek Değeri	Sıra
A	Öğrenci gelişim düzeylerini izleme buna uygun ölçme ve değerlendirme tekniklerini kullanma	0,03	7	0,918	4
B	Performans, proje ve portfolyaya dayalı ölçmeler yapma.	0,00	11	1,102	1
C	Bireysel farklılıkları dikkate alan ölçme ve değerlendirme tekniklerini uygulama.	0,03	8	0,852	7
D	Ölçme sonuçlarını geribildirim olarak kullanma	0,05	5	0,855	6
E	Bilişsel, duyuşsal ve psikomotor alanda ölçme araçları hazırlama	0,17	1	0,867	5
F	Objektif testler kullanma	0,04	6	0,836	8
G	Öğrencinin öz değerlendirme ve akran değerlendirme yapmasına olanak tanıma.	0,01	10	1,055	2
H	Teknolojik araç ve gereçleri ölçme ve değerlendirme yaparken kullanma	0,16	2	0,987	3
I	Değerlendirme ölçütlerini belirleme öğrenciyi bu ölçütlerden haberdar etme.	0,06	3	0,678	9
J	Ölçme sonuçlarına göre eğitim programını değiştirme.	0,03	9	0,279	10
K	Farklı soru formatları kullanma.	0,06	4	0,00	11

($r = -0.349$; $p > .05$)

Tablo 12’ye göre ikinci Delphi uygulamasında ikili karşılaştırmalar ve sıralama yargılarıyla ölçeklemede ölçek değerleri arasında manidar bir ilişki bulunmamıştır ($r = -0.349$; $p > .05$). İkili karşılaştırma ve sıralama yargılarıyla ölçekleme benzer formülleri kullanmasına karşın yargıcıların farklı yargılarda bulunması nedeniyle birbirine benzer sonuçlar üretmemektedir. Bu sonuca göre, yargıcı kararlarının elde edilmiş biçiminin uyarıcıların psikolojik boyuttaki yerinin fiziksel boyuttaki karşılığının olduğu söylenebilir. Bu durumda, eğer belli bir uyarıcı takımına ilişkin yargıcıların algılarına dayalı bilgi toplanmak istenirse, yargıcı kararlarının elde edilmiş biçimine dikkat edilmesi gerektiğini göstermektedir.

Turgut ve Baykul’a göre (1992) bu iki yaklaşıma ait değerlendirmeleri dikkate alınırsa, ikili karşılaştırmalar ile sıralama yargılarıyla ölçekleme sonuçlarının arasındaki ilişkinin düşük olmasına ve manidar çıkmamasına sıralama yargılarına ait ölçekleme yöntemlerinin neden olduğu düşünülebilir. Torgerson (1958), sıralama yargılarına dayalı yöntemlerin bazı sayıltılarının karşılanmasının zor olduğunu belirtmiştir. Sıra sayılarına ardışık aralıklar yöntemi uygulanması birinci ve sonuncu sıra sayılarının sınırları belirsiz olduğundan ve sıralamada aralık sınırları kararlılık gösteremeyeceğinden dolayı eleştirilmiştir (Guilford, 1954). Bundan dolayı sıralama yargılarına dayalı sürecin içinde en tutarlı ölçek değerlerini veren yöntemin ikili karşılaştırmalara dayalı yöntem olabileceğini öne sürmüştür. Diğer yandan mutlak yargılar ve sınıflama yargıları ile ölçekleme yöntemleriyle ikili karşılaştırma arasındaki uyumun yüksekliği, uyarıcı sayısının yüksek olduğu hallerde diğer yöntemlerin tercih edilmesiyle fazla hata yapılmayacağını göstermiştir (Turgut & Baykul, 1992). Bu bağlamda sıralama yargılarıyla ölçekleme ile ikili

karşılaştırmayla ölçekleme arasındaki ilişkinin manidar olmaması sıralama yargılarıyla ölçeklemenin sayılılarından kaynaklı olduğu söylenebilir. Bunun için Kan (2008)'ın yaptığı benzer bir araştırmada önerdiği gibi Guilford (1954) tarafından açıklanan birleşik standartla sıralama yargılarına dayalı ölçekleme yapılması ölçek değerleri arasında daha yüksek bir ilişkiyi gösterebilir.

3. Delphi paneli uygulaması için;

“Üçüncü Delphi panelinde ikili karşılaştırma yargılarına dayalı ölçekleme yönteminden elde edilen madde ölçek değerleri ne düzeydedir? Ve Üçüncü Delphi panelinde sıralama yargılarına dayalı ölçekleme yönteminden elde edilen madde ölçek değerleri ne düzeydedir?” Sorularına ilişkin bulgular aşağıda verilmiştir.

On bir yeterlik maddesi aynı ölçekleme formu 2. Delphi sonuçları ve ölçek değerleri öğretmenlere üç ay sonra yeniden verilmiştir. Bu ölçekleme formu ile birlikte 2. Delphi uygulamasında her bir öğretmenin yanıtlarının yanında ölçeklemeye ilişkin genel tablo ve her bir yeterlik alanının alt maddelerini içeren form öğretmenlere verilmiştir. Burada amaç Delphi uygulamalarında kontrollü geribildirim sağlanması, uzlaşılan ve uzlaşılmayan maddelerin tüm grup tarafından görülmesi ve gerektiğinde değiştirilmesi temeline dayanmaktadır. Öğretmenlerin 3. Delphi paneline verdikleri yanıtlar 2. Delphi panelinde izlenen ölçekleme yöntemiyle çözümlenmiştir. 3. Delphi uygulamasında öğretmenlerin verdikleri yanıtlar ölçeklenmiş, buna ilişkin ikili karşılaştırmalar ve sıralama ile ölçek değerleri Tablo 13’de sunulmuştur.

Tablo 13. Üçüncü delphi panelinde ikili karşılaştırma ve sıralama yargılarıyla ölçekleme sonucu ölçek değerlerinin karşılaştırılması

Yeterlik İfadeleri	İkili Karşılaştırma		Sıralama	
	Ölçek Değeri	Sıra	Ölçek Değeri	Sıra
Öğrenci gelişim düzeylerini izleme buna uygun ölçme ve değerlendirme tekniklerini kullanma	0,026	8	0,812	6
Performans, proje ve portfolyaya dayalı ölçmeler yapma	-0,005	11	1,013	2
Bireysel farklılıkları dikkate alan ölçme ve değerlendirme tekniklerini uygulama	0,033	7	0,77	8
Ölçme sonuçlarını geribildirim olarak kullanma	0,049	5	0,807	7
Bilişsel, duyuşsal ve psikomotor alanda ölçme araçları hazırlama	0,173	1	0,892	4
Objektif testler kullanma	0,042	6	0,827	5
Öğrencinin öz değerlendirme ve akran değerlendirme yapmasına olanak tanıma	0,017	10	1,024	1
Teknolojik araç ve gereçleri ölçme ve değerlendirme yaparken kullanma	0,16	2	0,932	3
Değerlendirme ölçütlerini belirleme öğrenciyi bu ölçütlerden haberdar etme	0,074	3	0,677	9
Ölçme sonuçlarına göre eğitim programını değiştirme	0,025	9	0,281	10
Farklı soru formatları kullanma	0,059	4	0,00	11

($r=-0.218$; $p>.05$)

Üçüncü Delphi uygulaması sonucu elde edilen ikili karşılaştırma ve sıralama yargılarıyla ölçekleme bulguları incelendiğinde Tablo 13’te belirtildiği gibi yeterlik ifadelerinin ölçek değerleri farklı ve sıralamanın da farklı olduğu bulunmuştur. Tablo 26’da görüldüğü gibi üçüncü Delphi uygulamasında ikili karşılaştırmalar ve sıralama yargılarıyla ölçeklemede ölçek değerleri arasında manidar bir ilişki bulunmamıştır ($r= -0.218$;

$p > .05$). İlişkinin negatif ve manidar olmaması ikinci Delphi uygulaması sonucu elde edilen ölçek değerleri arasındaki ilişkiyle benzerdir. Bu durumda ikinci Delphi uygulaması sonucu elde edilen ölçek değerleri arasındaki ilişkiye benzer bir yorum burada da yapılabilir.

Bu çalışmanın bulguları bu konuda az sayıda yapılmış çalışmaların sonuçlarıyla paralellik göstermektedir. Örneğin, O'Neil ve Chissom (1993) üç ölçekleme modelini ikili karşılaştırma, sıralama ve likert tipi ölçekleri birbiriyle karşılaştırmış ve yargıcı kararlarına dayalı olan ikili karşılaştırma ve sıralama yargılarına dayalı geliştirilen ölçek maddeleri arasında, denek tepkilerine dayalı geliştirilen ölçek (Likert tipi) maddeleri arasındakinden daha yüksek korelasyon olduğu sonucuna ulaşmışlardır. İkili karşılaştırma ve sıralama yargılarına dayalı ölçekleme arasındaki ilişkinin manidar olmaması sıralama yargılarına dayalı ölçeklemenin uyarıcı sayısının fazlalığından kaynaklı olduğu söylenebilir.

Delphi paneli uygulaması için;

“İkinci ve üçüncü delphi uygulamalarında ikili karşılaştırma yargılarına dayalı olarak elde edilen ölçek değerleri arasında uyum var mıdır? Ve “İkinci ve üçüncü delphi uygulamalarında sıralama yargılarına dayalı olarak elde edilen ölçek değerleri arasında uyum var mıdır?” Sorularına ilişkin bulgular aşağıda verilmiştir.

“İkili karşılaştırma yargılarına dayalı olarak elde edilen ölçek değerleri ile sıralama yargılarına dayalı olarak elde edilen ölçek değerleri arasında uyum var mıdır?” sorusuna yanıt aranırken; ikinci Delphi uygulaması ile üçüncü Delphi uygulaması arasındaki uyumun belirlenmesinde Pearson korelasyon katsayısı kullanılmıştır. Pearson r 'si, aralık ya da oran ölçeğinde ölçülen iki sürekli değişken arasındaki doğrusal ilişkiyi açıklamak üzere kullanılır (Büyüköztürk, Çokluk ve Köklü, 2011). Buna göre ikinci ve üçüncü Delphi uygulamalarında ikili karşılaştırma ile ölçeklemeden elde edilen ölçek değerleri arasında yüksek korelasyon ($r=0.993$; $p<.05$), sıralama ile ölçeklemeden elde edilen ölçek değerleri arasında da yüksek korelasyon ($r=0.991$; $p<.05$) bulunmuştur. Bu korelasyon değeri Delphi uygulamaları içine gerekli olan uzlaşının tam olarak sağlandığının da göstergesi kabul edilir.

Birinci ve ikinci Delphi uygulamaları arasında ikili karşılaştırma ile ölçekleme değerleri Tablo 14'te sunulmuştur.

Tablo 14. İkinci ve üçüncü delphi uygulamaları arasında ikili karşılaştırma ölçek değerleri

Yeterlik İfadeleri	İkinci Uygulama		Üçüncü Uygulama	
	Ölçek Değeri	Sıralama	Ölçek Değeri	Sıralama
Öğrenci gelişim düzeylerini izleme buna uygun ölçme ve değerlendirme tekniklerini kullanma	0,03	7	0,026	8
Performans, proje ve portfolyaya dayalı ölçmeler yapma	0,00	11	-0,005	11
Bireysel farklılıkları dikkate alan ölçme ve değerlendirme tekniklerini uygulama	0,03	8	0,033	7
Ölçme sonuçlarını geribildirim olarak kullanma	0,05	5	0,049	5
Bilişsel, duyuşsal ve psikomotor alanda ölçme araçları hazırlama	0,17	1	0,173	1
Objektif testler kullanma	0,04	6	0,042	6
Öğrencinin öz değerlendirme ve akran değerlendirme yapmasına olanak tanıma	0,01	10	0,017	10
Teknolojik araç ve gereçleri ölçme ve değerlendirme yaparken kullanma	0,16	2	0,16	2
Değerlendirme ölçütlerini belirleme öğrenciyi bu ölçütlerden haberdar etme	0,06	3	0,074	3
Ölçme sonuçlarına göre eğitim programını değiştirme	0,03	9	0,025	9
Farklı soru formatları kullanma	0,06	4	0,059	4

($r=0.993$; $p<.05$)

İkinci Delphi uygulaması ile üçüncü Delphi uygulamasında ikili karşılaştırma yargılarıyla ölçeklemeden elde edilen ölçek değerleri arasındaki korelasyon ($r=0.993$; $p<.05$) yüksek bulunmuştur. Her iki ölçekleme arasındaki korelasyonun yüksek olması, ölçek değerleri arasındaki uyumun yüksek olduğunu göstermektedir; bu da delphi uygulaması için gerekli olan uzlaşının sağlandığına işaret etmektedir. İkinci ve üçüncü Delphi uygulamasında öğretmenler ikili karşılaştırmalar ile ölçeklemede “Bilişsel, duyuşsal ve psikomotor alanda ölçme araçları hazırlar” yeterliğini diğer yeterlik maddelerinden daha önemli görürken “Performansa, projeye ve portfolyaya dayalı ölçmeler yapar.” yeterliğini en az önemli görmüşlerdir. Bu iki ölçeklemede yeterlik maddelerinin ölçek değerleri ve sıralamalarında az da olsa değişme olmasına rağmen sıralamada önemli bir değişiklik göstermemektedir. İkinci ve üçüncü Delphi uygulamaları arasında sıralama ile ölçekleme değerleri Tablo 15’te sunulmuştur.

Tablo 15. İkinci ve üçüncü delphi uygulamaları arasında sıralama ölçek değerleri

Yeterlik İfadeleri	İkinci Uygulama		Üçüncü Uygulama	
	Ölçek Değeri	Sıralama	Ölçek Değeri	Sıralama
Öğrenci gelişim düzeylerini izleme buna uygun ölçme ve değerlendirme tekniklerini kullanma	0,918	4	0,812	6
Performans, proje ve portfolyaya dayalı ölçmeler yapma	1,102	1	1,013	2
Bireysel farklılıkları dikkate alan ölçme ve değerlendirme tekniklerini uygulama	0,852	7	0,77	8
Ölçme sonuçlarını geribildirim olarak kullanma	0,855	6	0,807	7
Bilişsel, duyuşsal ve psikomotor alanda ölçme araçları hazırlama	0,867	5	0,892	4
Objektif testler kullanma	0,836	8	0,827	5
Öğrencinin öz değerlendirme ve akran değerlendirme yapmasına olanak tanıma	1,055	2	1,024	1
Teknolojik araç ve gereçleri ölçme ve değerlendirme yaparken kullanma	0,987	3	0,932	3
Değerlendirme ölçütlerini belirleme öğrenciyi bu ölçütlerden haberdar etme	0,678	9	0,677	9
Ölçme sonuçlarına göre eğitim programını değiştirme	0,279	10	0,281	10
Farklı soru formatları kullanma	0,00	11	0,00	11

($r=0.991$; $p<.05$)

İkinci Delphi uygulaması ile üçüncü Delphi uygulamasında sıralama yargılarıyla ölçeklemeden elde edilen ölçek değerleri arasındaki korelasyon ($r=0.991$; $p<.05$) yüksek bulunmuştur. Her iki ölçekleme arasındaki korelasyonun yüksek olması ölçek değerleri arasındaki uyumun yüksek olduğunu göstermektedir. Bu sonuç, delphi uygulaması için gerekli olan uzlaşının sağlandığına işaret etmektedir. İkinci ve üçüncü Delphi uygulamasında öğretmenler sıralama yargılarıyla ölçeklemede “Performansa, projeye ve portfolyaya dayalı ölçmeler yapar.” yeterliği ile “Öğrencinin öz değerlendirme ve akran değerlendirme yapmasına olanak tanıma.” yeterliğini diğer yeterlik maddelerinden daha önemli görürken “Farklı soru formatları kullanma.” yeterliğini en az önemli görmüşlerdir. Bu iki ölçeklemede yeterlik maddelerinin ölçek değerleri yalnızca performans, projeye ve portfolyaya dayalı ölçmeler yapar yeterliği ile öğrencinin öz değerlendirme ve akran değerlendirmesine olanak tanıma yeterliği arasında çok az ölçek değeri olmasına rağmen sıralaması değişmiştir. Diğer yeterlik maddelerinin sıralamalarında önemli bir değişiklik olmamıştır.

Sıralama ile ölçekleme yapıldığında öğretmenler, “Performansa, projeye ve portfolyaya dayalı ölçmeler yapar, Öğrencinin öz değerlendirme ve akran değerlendirme yapmasına olanak tanır, teknolojik araç ve gereçleri ölçme ve değerlendirme yaparken kullanır, ölçütleri belirler öğrenciyi bu ölçütlerden haberdar eder, ölçme sonuçlarına göre eğitim programını değiştirir ve Farklı soru formatları kullanır.” yeterliklerini en az önemli görmektedirler. Karacaoğlu’nun (2009) yaptığı çalışmada öğretmenlerin kendilerini en fazla yeterli gördükleri boyut “alanındaki temel kavram ve genellemeleri bilme” yeterliği olarak belirtmiş, en az yeterli gördükleri yeterlik boyutları ise verileri analiz ederek yorumlama, öğrenciye gelişimi ve öğrenmesi hakkında geri bildirim sağlama, değişik ölçme tekniklerini kullanarak öğrenci öğrenmelerini ölçme olarak belirlemiştir. Öğretmenlerin ölçme ve değerlendirme alanında kendilerini yeterince yeterli görmedikleri ve bu konuda eksik olduklarını belirtmişlerdir. Bu araştırmanın bulgusu da Karacaoğlu’nun (2009) araştırma bulgularıyla paralellik göstermektedir.

Ömür (2009) yaptığı bir çalışmada, ikili karşılaştırmayla, genel izlenimle ve dereceli puanlama anahtarıyla değerlendirmelerin sonucunda, hangi değerlendirme yapılırsa yapılsın kompozisyonların sıralamalarında bir değişikliğin olmadığını belirtmiştir. Kompozisyon değerlendirmelerinde sıralama söz konusu olduğu durumlarda, ikili karşılaştırmalar yönteminin kullanılmasının uygun olacağını önermektedir. Bu bulgu yeterlik belirleme sürecinde yer alan nitel verilerin analizinde yol gösterici olarak ölçeklemenin kullanılmasını önermektedir. Yeterlik belirleme sürecinde ölçek sıralamasındaki küçük değişiklikler olmasının yanında tekrarlı yapılan ölçekleme sonucunda da sıralamanın değişmediği bu araştırmanın bulgularıyla örtüşmektedir.

Yargıcı kararlarına ve denek tepkilerine dayalı ölçekleme yaklaşımları üzerine karşılaştırmalı bir çalışma yapan Kan (2008b), yargıcı kararlarına dayalı olarak geliştirilen ölçeğin, denek tepkilerine dayalı geliştirilen ölçeğe nazaran gruplar arası farklılığı ortaya çıkarmada daha objektif olduğunu belirtmiştir. Yargıcı kararlarına dayalı yapılan bu çalışmanın bulgularıyla benzer bir yorum yapılmıştır. Bülbül ve Acar (2012), Türkiye’de eğitim denetmenlerinin görevleri üzerine ikili karşılaştırma yöntemiyle bir ölçekleme çalışmasında, eğitim denetmenliği için seçilen beş genel görev Thurstone’un karşılaştırmalı yargılar kanununun 5. Durum denklemi ve tam veri matrisinde ölçeklenmiştir. İkili karşılaştırmalar ile ölçeklemede ölçek değerleri arasındaki uyumun yüksek olduğunu belirtmiştir.

Özbaşı (2009), öğretmen yeterliklerinin sıralanması araştırmasında öğretmenlerin birçok yeterlik alanında kendilerini yeterli gördüklerini bunların içinde en fazla sınavın/testin amacını belirlemek olarak belirlemiştir. Bu çalışmada Teknolojik araç ve gereçleri ölçme ve değerlendirme yaparken kullanma yeterliği her iki ölçekleme yönteminde de en önemli görülen yeterlik boyutu olarak yapılan diğer çalışmalarla paralellik göstermektedir. Bu çalışmada öğretmenlerin “Teknolojik araç ve gereçleri ölçme ve değerlendirme yaparken kullanır” yeterliğini diğer yeterlik boyutlarından daha önemli gördüğü bulunmuştur.

Özer ve Acar (2011), öğretmenlik mesleği genel yeterlikleri üzerine ikili karşılaştırma yöntemiyle ölçekleme çalışmasında eğitim fakültesi dördüncü sınıf öğrencilerinin Milli Eğitim Bakanlığı tarafından belirlenen öğretmenlik mesleği genel yeterliklerini ölçeklemelerini istemişlerdir. En önemli görülen öğretmen yeterliği “öğrenciyi tanıma” olarak belirlenirken en az önemli görülen yeterlik alanı “program ve içerik bilgisi” olarak ortaya çıkmıştır. Çalışmada yapılan ölçekleme işleminde ölçek değerlerinin büyükten küçüğe sıralanmasında madde toplam frekansları dikkate alınmadığı bu yüzden en büyük

ölçek değerinin en son tercih edildiği belirtilmiştir. Ölçekleme çalışmasında ölçek frekansları sütun toplamları en yüksek olan madde en çok tercih edilen madde olarak belirlenmelidir. Aynı ölçekleme yönteminin kullanıldığı araştırmanın bu bulgusu “öğretmenlerin ölçme sonuçlarına göre program ve içeriği değiştirme” bulgusuyla paralellik göstermektedir.

Güler ve Anıl (2009) “Lisansüstü Eğitime Başvuran Öğrencilerde Aranılan Özelliklerin İkili Karşılaştırma Yöntemiyle Ölçeklenmesi” adlı çalışmalarında ikili karşılaştırmayla ölçekleme yönteminde, akademik başarı puanının ilk sırada yer aldığı belirlenmiştir. Bu özelliği sırasıyla mülakat puanı, programa girme amacı, İngilizce yeterlik puanı, ALES puanı, bölüm öğrencisi olup olmaması ve bu nitelikleri en son sırada referansı olup olmaması özelliğinin izlediği belirlenmiştir. Ölçeklemede izlenen yollar uyum ve tutarlılık hesaplama yöntemi bu çalışmayla paralellik göstermektedir.

Türk Eğitim Derneği’nin 2009 yılında yaptığı çalışmada öğretmenlerin, öğrencilerin gelişim düzeyini ve bireysel farklılıklarını belirlemek için yaptıkları çalışmaların yetersiz kaldığını belirlenmiş, öğrencilerin bireysel gelişim dosyalarını inceleyen öğretmenlerin az sayıda olduğu, öğretmenlerin yenilenen programların gerektirdiği tamamlayıcı ölçme ve değerlendirme yöntem ve araçlarını kullanmada yetersiz kaldığı vurgulanmıştır. Bu bulgu, sıralama yargısıyla ölçekleme sonucu elde edilen ölçek değerlerine bakıldığında performans, proje ve portfolyaya dayalı ölçmeler yapmayı en önemli yeterlik olarak araştırmayı destekler niteliktedir. Öğretmenlerin ölçme ve değerlendirme alanında sahip olması gereken yeterlikleri sıralarken yeni ölçme ve değerlendirme uygulamalarını önemsedikleri ve bunların iyi bir öğretmende olması gereken yeterlikler olarak belirlenmiştir.

İkinci ve üçüncü Delphi uygulamasında öğretmenler ikili karşılaştırmalar ile ölçeklemede “Bilişsel, duyuşsal ve psikomotor alanda ölçme araçları hazırlar” yeterliğini diğer yeterlik maddelerinden daha önemli görürken “Performans, proje ve portfolyaya dayalı ölçmeler yapar” yeterliğini en az önemli görmüşlerdir. Bu iki ölçeklemede yeterlik maddelerinin ölçek değerleri ve sıralamalarında az da olsa değişme olmasına rağmen sıralamada önemli bir değişiklik göstermemektedir. Öğretmenlerin ölçme ve değerlendirme yeterlikleri içinde tamamlayıcı ölçme ve değerlendirme yöntemlerini tercih etmemelerinin uygulamalardaki zorluk ve klasik ölçme yaklaşımlarının daha kolay olduğunu ortaya çıkarmaktadır. Ancak yeterlikleri sıraladıklarında yeni ölçme ve değerlendirme uygulamalarını öncelikli olarak belirtmektedirler. Anıl ve Acar (2008) yaptıkları çalışmada, sınıf öğretmenlerinin özellikle alternatif ölçme araçlarının hazırlama ve uygulama aşamasında zaman açısından problem yaşadıkları sonucu ortaya çıkmıştır. Bu bulgu, ikili karşılaştırma ile ölçekleme sonucu “performans proje ve portfolyaya dayalı ölçmeler yapma” yeterliğini en az önemli olduğu bulgusuyla paralellik göstermektedir. Güneş (2007) yaptığı çalışmada, öğretmenlerin, verileri analiz etme, yorumlama, öğrenci gelişimi hakkında geri bildirim sağlama yeterliklerinin orta düzeyde olduğunu belirlemiştir. Bu çalışmada ise, ikili karşılaştırmalar ve sıralama yargısıyla ölçekleme sonucunda öğretmenler ölçme sonuçlarını geri bildirim olarak kullanır yeterliğini orta düzeyde önemli gördüklerini belirtmiş ve Güneş (2007)’in bulgusuyla paralellik göstermiştir. Yukarıda sunulan bulgular ve bulgularla ilişkili araştırmalar incelendiğinde; öğretmen yeterliklerinin yargıcı kararlarına dayalı ölçeklenmesinde, bu araştırma bulgularının diğer bazı araştırma bulgularıyla paralellik gösterdiği, bazı araştırma bulgularıyla ise ters düştüğü saptanmıştır.

SONUÇLAR VE ÖNERİLER

Bu bölümde, araştırmının bulguları doğrultusunda çıkarılan sonuçlara yer verilmiş ve bu sonuçlara dayalı önerilerde bulunulmuştur. Bu çalışmada, okul eğitiminde gereksinim duyulan ölçme ve değerlendirme alanı öğretmen yeterliklerinin belirlenmesi, yeterlik alanlarının yargıcı kararlarına dayalı olarak ölçeklenmesi amaçlanmıştır. Çalışma Ankara il sınırları içerisinde bulunan 9 okulda görev yapan 300 öğretmen üzerinde gerçekleştirilmiştir.

Bu çalışmada araştırma deseninin oluşturulmasında öğretmen yeterliklerinin belirlenmesinde Delphi tekniğinden yararlanılmıştır. Böylece nicel ve nitel verilerin bir arada kullanılarak aşamalı bir şekilde probleme ilişkin daha kapsamlı analizlerin yapılmasına ve ortak bir görüşün belirlenmesine olanak sağlanmıştır.

Alanyazında yapılan ölçekleme çalışmaları sınırlı sayıda olduğu bununla birlikte farklı ölçekleme yaklaşımlarının karşılaştırıldığı çalışmaların da az sayıda olduğu belirlenmiştir. Yapılan çalışmaların sonuçları bu araştırma ile bazı açılardan benzerlik bazı açılardan ise farklılıklar göstermektedir.

Bu çalışmada iki farklı ölçekleme tekniğinde ikili karşılaştırma ve sıralama yöntemiyle ölçeklemede farklı sıralamaların olduğu ortaya çıkmıştır. Ancak ölçek değerlerine bakıldığında çok küçük bir farklılıkla bu ölçek değerlerinin sıralamasında değişimler olmuştur. En yüksek ölçek değerine sahip üç ölçek değeri ile en düşük ölçek değerine sahip üç ölçek değerine bakıldığında ölçek değerlerinin birbirine yakın olduğu gözlenmiştir. İkili karşılaştırma ile ölçeklemede ikinci Delphi uygulamasıyla üçüncü Delphi uygulamasında “Bilişsel, duyuşsal ve psikomotor alanda ölçme araçları hazırlar, Teknolojik araç ve gereçleri ölçme ve değerlendirme yaparken kullanır ve Farklı soru formatları kullanma.” yeterlikleri öğretmenler tarafından daha önemli görülürken, “Öğrencinin öz değerlendirme ve akran değerlendirme yapmasına olanak tanır, Performansa, projeye ve portfolyaya dayalı ölçmeler yapar.” yeterlikleri en az önemli görülen yeterlikler olarak belirlenmiştir.

İkinci Delphi uygulaması ile üçüncü Delphi uygulaması arasındaki uyumun belirlenmesinde Pearson sıra farkları korelasyon katsayısı kullanılmıştır. Buna göre hem ikinci ve üçüncü Delphi uygulamalarında ikili karşılaştırma ile ölçeklemeden elde edilen ölçek değerleri arasında yüksek korelasyon ($r=0.993$; $p<.05$), hem de sıralama ile ölçeklemeden elde edilen ölçek değerleri arasında da yüksek korelasyon ($r=0.991$; $p<.05$) bulunmuştur. İkili karşılaştırma ve sıralama ölçeklemeleri için kendi aralarındaki bu korelasyon değerleri Delphi uygulamaları içinde gerekli olan uzlaşının tam olarak sağlandığının da göstergesi olarak kabul edilir. Delphi çalışmalarında uzlaşının olması, katılımdaki gizlilik ve kontrollü geribildirim bu çalışmada tam olarak sağlanmıştır. Ölçeklemeler arasındaki ilişkinin yüksek olması Delphi tekniği ile toplanan verilerin tutarlı sonuçlar vermesi araştırmının güvenilir sonuçlar vermesine olanak tanımıştır.

Araştırma bulgularına dayanarak hem ölçme ve değerlendirme alanında uygulamada çalışan uzmanlara hem de bundan sonra bu tür bir araştırma yapacak olanlara öneriler sunulmuştur. Ölçme ve değerlendirme alanında uygulamada çalışan uzmanlara yönelik olarak, öğrenciler için verimli bir eğitim hizmeti sunmaları için öğretmenlerin, onların ilgi ve gereksinimlerini tanımaları bunu yaparken de farklı ölçme yöntemlerini kullanmaları yerinde olacaktır. Çağdaş eğitim sistemlerinde çok yönlü ölçme yöntemlerinin kullanılması önerilmektedir. Araştırma sonuçlarına göre öğretmenlerin öz değerlendirme ve akran değerlendirme yapmayı çok önemli bir yeterlik alanı olarak görmedikleri belirlenmiştir. Bu yöntemlerin kullanılmasını artırmak için öğretmenlerin bilgilendirilmesi önerilebilir.

Öğretmenler, performansa, projeye ve portfolyaya dayalı ölçmeler yapmayı az önemli görmelerini, değişik ölçme yöntemlerinden yararlanabilmeleri için onlara hazır, geliştirilmiş ölçme araçlarını hazırlayıp kullanmalarını sağlamak yararlı olabilir. Okullarda yapılan hizmet içi eğitim seminerleri için örnek çalışmaların gönderilip öğretmenlerin bilgilendirilmelerini sağlamanın yararlı olacağı düşünülmektedir. Bu çalışmada belirlenen öğrenci başarısını ölçme ve değerlendirme ile ilgili yeterlikler, öğretmenlerin gereksinim duydukları konularda düzenlenecek hizmet içi eğitimlerde içerikler hazırlanırken rehberlik edebilir. Öğretmenlerin ölçme yöntemlerini kullanmada karşılaştıkları sorunları azaltabilmek için buldukları bölgede veya okullarda ölçme ve değerlendirme uzmanı buldurmak belirlenen yeterlikler ile ilgili bilgilendirme yapmak yararlı olabilir.

Bu çalışma sonucu ortaya çıkan öğrenci başarısını ölçme ve değerlendirme ile ilgili yeterlikler eğitim fakültelerinin programlarının hazırlanması aşamasında ölçme ve değerlendirme dersi içeriği için bir kaynak niteliğindedir. Bu programların oluşturulmasında dikkate alınması gerektiği önerilmektedir.

Bu çalışmadan sonra bu tür bir araştırma yapacak olanlara önerilerin başında çalışmada geliştirilen, öğretmenlere yönelik öğrenci başarısını ölçme ve değerlendirme ile ilgili yeterlikler, daha geniş ve farklı okul türündeki (devlet, özel, mesleki) öğretmenlere uygulanarak, öğretmenlerin sahip olmaları gereken, ulusal düzeyde öğrenci başarısını ölçme ve değerlendirme standartları geliştirilebilir.

Delphi uygulamasının kullanıldığı çalışmalar incelendiğinde farklı istatistiksel tekniklerin kullanılmadığı aritmetik ortalama ve ortanca gibi istatistiklerin kullanıldığı görülmektedir. Bundan sonra yapılacak bu çalışmaya benzer çalışmalarda, ölçekleme tekniklerinin de çalışmalara dahil edilmesi hem delphi uygulamaları hem de ölçekleme çalışmalarının gelişmesi bakımından fayda sağlayacaktır. Bu çalışmada ikili karşılaştırmalar ve sıralama yargılarıyla ölçekleme yöntemlerinden yararlanılmıştır. Bundan sonraki çalışmalarda sınıflama yargılarıyla, mutlak yargılarla ve derece toplamlarıyla ölçekleme tekniklerinin kullanılması ölçekleme tekniklerinin tanınması ve kullanımının yaygınlaşması adına araştırmacılara önerilmektedir. Çalışmada tek boyutlu ölçekleme tekniklerinden yararlanılmıştır. Çalışmanın amacı doğrultusunda öğretmenlerin özellikleri bakımından yeterlikleri karşılaştırma yoluna gidilmemiştir. Bundan sonra yapılacak çalışmalarda çok boyutlu ölçekleme tekniklerinden yararlanarak ölçekleme alanındaki bilgi birikimini artıracak çalışmaların yapılması önerilmektedir.

KAYNAKÇA

- Albayrak, A. & Gelbal, S. (2012). İkili karşılaştırmalar yargılarına ve sıralama yargılarına dayalı ölçekleme yaklaşımlarından elde edilen ölçek değerlerinin tutarlığının karşılaştırılması. *21. Ulusal Eğitim Bilimleri Kongresi*, İstanbul, 12-14 Eylül.
- Alkan, C. (1998). *Eğitim teknolojisi*. Ankara: Anı Yayıncılık.
- Anıl, D. & Güler, N. (2006). İkili karşılaştırma yöntemi ile ölçekleme çalışmasına bir örnek. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 30, 30-36.
- Bal, Ö. (2011). Seviye belirleme sınavı (sbs) başarısında etkili olduğu düşünülen faktörlerin sıralama yargıları kanunuyla ölçeklenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 2 (2), 200-209.
- Brady, D. (1994). *Development of vocational curricula handout*. UK: The University of Huddersfield.
- Brown, T. C. & Peterson, G. L. (2009). *An enquiry into the method of paired comparison: Reliability, scaling, and thurstone's law of comparative judgment*. Gen Tech.

- Bülbül, T. & Acar, M. (2012). A pair-wise scaling study on the missions of education supervisors in turkey. *International Journal of Human Sciences*, 9 (2), 623-640.
- Bursalıoğlu, Z. (1975). *Eğitim yöneticisinin yeterlikleri*. Ankara: Ankara Üniversitesi Eğitim Fakültesi Yayınları. No: 93.
- Büyüköztürk, Ş., Çokluk, Ö. & Köklü, N. (2011). *Sosyal bilimler için istatistik (7. baskı)*. Ankara: Pegem A Yayıncılık.
- Crocker, L. & Algina, J. (1986). *Introduction to classical & modern test theory*. Florida: Holt, Rinehart and Winston Inc.
- Demirel, Ö. (1999a). "Öğretmen nitelikleri" Cumhuriyetin 75. yılında öğretmen yetiştirme. Ankara: Milli Eğitim Basımevi. ss. 53-62
- Demirel, Ö. (1999b). *Eğitimde program geliştirme*. Ankara: Pegem A Yayıncılık.
- Demirel, Ö. (2003). *Planlamadan değerlendirmeye öğretme sanatı*. Ankara: Pegem A Yayıncılık.
- Dunn-Runkin, (1982). *Scaling methods*. New Jersey: Lawrence Erlbaum Associates Publishers.
- Ekinci, A., Bindak, R. & Yıldırım, C. (2012). İlköğretim okulu yöneticilerinin öğretmenlerin mesleki sorunlarına empatik yaklaşımlarının ikili karşılaştırmalar metodu ile incelenmesi. *Gaziantep Üniversitesi Sosyal Bilimler Dergisi*, 11 (3), 759-776.
- Ekşi, N. (2006). Bilgisayar teknolojisinde öğretmen yeterlikleri. *İlköğretim Eğitimci Dergisi*, 2, 8-10.
- Fraenkel, J. R. & Wallen, N.E. (2006). *How to design and evaluate research in education*. New York: McGraw-Hill.
- Gökçe, E. (1999). İlköğretim öğretmenlerinin yeterlikleri. *Yayımlanmamış Doktora Tezi*. Ankara Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Guilford, J. P. (1954). *Psychometrics methods*. New York: Mc Graw-Hill Book Co.
- Güler, N. & Anıl, D. (2009). Scaling through pair-wise comparison method in required characteristics of students applying for post graduate programs. *International Journal of Human Sciences*, 6 (1), 627-639.
- Gündüz, Y. (2001). Öğretmenlerin sınıf yönetimindeki yeterlikleri. *Yayımlanmamış Yüksek Lisans Tezi*. Yıldız Teknik Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Güneş, A. (2007). Sınıf öğretmenlerinin kendi algılarına göre ölçme değerlendirme yeterlikleri. *Yayımlanmamış Yüksek Lisans Tezi*. Marmara Üniversitesi Eğitim Bilimleri Enstitüsü, İstanbul.
- Judd, C. M., Eliot, E. R., & Kidder, L. H. (1991). *Research methods in social relations*. Newyork: Harcourt Brace Jovanovich College Publishers.
- Kan, A. (2008a). Psikolojik değişkenleri ölçmek için kullanılan ölçekleme yaklaşımları üzerine bir araştırma. *Eğitimde Kuram ve Uygulama*, 4 (1), 2-18. <http://eku.comu.edu.tr/index/4/1/akan.pdf>. İndirme Tarihi: 10 Ekim 2009.
- Kan, A. (2008b). Yargıcı kararlarına dayalı ölçekleme yöntemlerinin karşılaştırılması üzerine ampirik bir çalışma. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 35, 186-194.
- Karacaoğlu, Ö. C. (2008). Öğretmenlerin yeterlilik algıları. *Yüzüncü Yıl Üniversitesi, Eğitim Fakültesi Dergisi*, V (I), 70-97, <http://efdergi.yyu.edu.tr>. İndirme Tarihi: 30 Kasım 2009.
- Karasar, N. (2004). *Bilimsel araştırma yöntemi*. Ankara: Nobel Yayın Dağıtım.
- Kavcar, C. (2002). Cumhuriyet döneminde dal öğretmeni yetiştirme. *Eğitim Bilimleri Fakültesi Dergisi*, 35 (1-2), 1-14.
- Lang, T. (1994). *An overview of four futures methodologies*. Retrieved July 6, 2007. <http://www.futures.hawaii.edu/publications/half-fried-ideas/J7/LANG.pdf>. İndirme Tarihi: 14 Temmuz 2011.

- Linstone H. A. & Turoff M. (2002). *The delphi method: techniques and applications*. California: University of Southern California.
- Mansfield, R. S. (1996). Building competency models: approaches for HR professionals. *Human Resource Management*, 35 (1), 7-18.
- MEB (2006). *Temel eğitime destek programı öğretmenlik mesleği genel alan yeterlik çalışması ulusal raporu*. Ankara. <http://otmg.meb.gov.tr/belgeler/otmg/Yeterlikler.pdf>. İndirme Tarihi: 10 Aralık 2008.
- Nartgün, Z. (2006). Öğretmenlik meslek bilgisi derslerinin önem düzeyinin ikili karşılaştırmalarla ölçeklenmesi, *A.İ.B.Ü. Eğitim Fakültesi Dergisi*, 6 (2), 161- 176.
- Navarro, J.C. (2003). The teaching profession-beyond training. *Techknowlogia*. http://www.techknowlogia.org/TKL_active_pages2/CurrentArticles/main.aspFileType=HTML&ArticleID=206. İndirme Tarihi: 14 Temmuz 2003.
- NBPTS (2004). *The national board for professional teaching standards*. <http://www.nbpts.org/nbpts/standards/five-props.html>. İndirme Tarihi: 15 Şubat 2004.
- NcNess, E., Broedfoot, P. & Osborn, M. (2003). Is the effective ompromising the affective? *British Educational Research Journal*, 29 (2), 243-257.
- O'Neil, M. R. & Chissom, B. S. (1993). A comparison three methods for assessing attitudes. *Paper Presented at The Annual Meeting of The Educational Research Association*. New Orleans, USA.
- OECD (Organisation for Economic Cooperation and Development) (2004). *Learning for tomorrow's world first results from PISA 2003. The learning environment and the organisation of schooling*. <http://www.pisa.oecd.org/dataoecd/58/59/33918026.pdf>. İndirme Tarihi: 13 Aralık 2004.
- Öğretmen, T. (2008). Alan tercih envanteri: ölçeklenmesi, geçerliği ve güvenilirliği. *Türk Eğitim Bilimleri Dergisi*, 6 (3), 507-522.
- Ömür, S. (2009). Dereceli puanlama anahtarıyla, genel izlenimle ve ikili karşılaştırmalar yöntemiyle yapılan değerlendirmelerin karşılaştırılması. *Yayımlanmamış Yüksek Lisans Tezi*. Mersin Üniversitesi Sosyal Bilimler Enstitüsü, Mersin.
- Özbaşı, D. (2009). Sınıf öğretmenleri için öğrenci başarısını ölçme ve değerlendirme ile ilgili yeterlik göstergelerinin ve bunlara ilişkin algılarının incelenmesi. *Yayımlanmamış Yüksek Lisans Tezi*. Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Özçelik, D. A. (1998). *Ölçme ve değerlendirme*. Ankara: ÖSYM Yayınları.
- Özer, Y. & Acar, M. (2011). Öğretmenlik mesleği genel yeterlikleri üzerine ikili karşılaştırma yöntemiyle bir ölçekleme çalışması. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi*, 3 (40), 89-101.
- Öztürk, N., Özdemir, S. & Gelbal, S. (2011). İki farklı ölçekleme yaklaşımından elde edilen ölçek değerleri tutarlılığının incelenmesi. *20. Ulusal Eğitim Bilimleri Kurultayı*, 8-10 Eylül 2011. Burdur.
- Paykoç, F. & Ok, A. (1990). Delphi tekniği ile türk eğitim sistemindeki bazı problemlerin incelenmesi. *Eğitim ve Bilim*, 14 (75), 14-21.
- Pollard, R., & M. Tomlin. (1995). The use of expert teachers to improve education. *Education*, 116 (1), 3-9.
- Quinn, P. (1986). *Utilization-focused evaluation*. Newbury Park, CA: Sage Publications.
- Sackman, H. (1975). *Delphi critique: expert opinion*. Lexington, MA: Lexington Books.
- Şahin, A. E. (2000). İlköğretim okulu müdürlerinin yeterlikleri. *Kuram ve Uygulamada Eğitim Yönetimi*, 6 (22), 243-260.

- Şahin, A. E. (2001). Eğitim araştırmalarında delphi tekniği ve kullanımı. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 20, 215-220.
- Şahin, A. E. (2004). Öğretmen yeterliklerinin belirlenmesi. *Bilim ve Aklın Aydınlığında Eğitim Dergisi*, 5 (58). <http://yayim.meb.gov.tr/dergiler/sayi58/yilmaz-koseoglu.htm>. İndirme Tarihi: 10 Aralık 2008.
- Seferoğlu, S. S. (2004). Öğretmen yeterlikleri ve mesleki gelişim. *Bilim ve Aklın Aydınlığında Eğitim Dergisi*, 5 (58), 40-45.
- Semerci, Ç. & Semerci, N., (2001). Program geliştirmede delphi, dacum ve meslek analizi. *F.Ü. Sosyal Bilimler Enstitüsü Dergisi*, 11 (2), 241-250.
- Senemoğlu, N. & Özçelik D. A. (1989). Öğretmen adaylarına öğretmenlik bilgisi kazandırma bakımından fen-edebiyat ve eğitim fakültelerinin etkililiği. *Çağdaş Eğitim Dergisi*, 14, 142-144. http://bef.sdu.edu.tr/hocalar/dekanlik/nuray_senemoglu/Makaleler/ogretmenlik_bilgisi.htm. İndirme Tarihi: 10 Aralık 2008.
- Senemoğlu, N. (2004). Öğrenci görüşlerine göre öğretmen yeterlikleri. http://bef.sdu.edu.tr/hocalar/dekanlik/nuray_senemoglu/Makaleler/ogretmen_yeterli.htm. İndirme Tarihi: 10 Aralık 2008.
- Şişman, M. (2009). Öğretmen yeterliliklerini yeniden düşünmek. *Türk Yurdu Dergisi*, 29, 37-41.
- Tezbaşaran A. (2004). Likert tipi ölçeklere madde seçmede geleneksel madde analizi tekniklerinin karşılaştırılması. *Türk Psikoloji Dergisi*, 19 (54), 77 – 90.
- Thurstone, L. L. (1958). A law of comparative judgment. *Psychological Review*, 34, 273-286.
- Torgerson, W. S. (1958). *Theory and methods of scaling*. Newyork: John Wiley & Sons Inc.
- Tschannen-Moran, M. & Woolfolk-Hoy, A. (2001). Teacher efficacy: Capturing an exclusive construct. *Teaching and Teacher Education*, 17, 783-805.
- Turgut, M. F. & Baykul, Y. (1993). *Ölçekleme teknikleri*. Ankara: ÖSYM Yayınları.
- Türk Eğitim Derneği (2009). *Öğretmen yeterlikleri*. Ankara: Türk Eğitim Derneği Yayınları.
- Turoff, M. & Hiltz, S. R. (2001). *Computer based Delphi processes*. London: Kingsley.
- Woolfolk-Hoy, A. & Burke-Spero, R. (2005). Changes in teacher efficacy during the early years of teaching: a comparison of four measures. *Teaching and Teacher Education*, 21, 343-353.
- Woudenberg F. (1991). An evaluation of Delphi. *Technological Forecasting and Social Change*, 40, 131-150.
- Yılmaz, K. & Çokluk, Ö. (2008). Primary school teachers' belief of efficacy. *Ankara University, Journal of Faculty of Educational Sciences*, 41 (2), 143-167.
- Yılmaz, M., Köseoğlu, P., Erçek, C. & Soran, H. (2004). Öğretmen öz-yeterlik inancı. *Bilim ve Aklın Aydınlığında Eğitim Dergisi*, 5 (58), 50-54. <http://yayim.meb.gov.tr/dergiler/sayi58/yilmaz-koseoglu.htm>. İndirme Tarihi: 10 Aralık 2008.

Scaling Primary School Teachers' Competence Based On Judgmental Decisions in the Field Of Measurement and Evaluation⁴

Recep Serkan ARIK⁵ & Ömer KUTLU⁶

Introduction

Teacher competency is an issue that has been considered and discussed much in the last few years. In particular, the role of teaching and learning in this thread has increasingly interested researchers on the topic as well as practitioners (Woolfolk Hoy & Burke-Spero, 2005). Teacher competency is described as a set of beliefs that enable teachers to accomplish their teaching duties in a specific context and function in order to demonstrate necessary skills that are required to modify students' behaviors (Tschannen-Moran & Woolfolk Hoy, 2001; Yılmaz & Çokluk, 2008).

The common view of experts (reconciliation) concerning competencies is important. For that reconciliation, instead of determinations made with respect to arithmetic means or medians, the use of more distinct and established statistical techniques is needed. To accomplish this, researchers have proposed, in particular, the use of techniques of scaling for measuring many of the behaviors in the affective areas (Anıl & Güler, 2006; Kan, 2008a; Özer & Acar, 2011). In the process of identifying the competencies of teachers, the views of teachers themselves must be taken into account and included in the research in a systematic way. In this case, teachers' determinations of competencies and their own views in this regard are important in order to increase the quality of education. In order to determine teacher qualifications, examining only the related literature and using business-analysis techniques will result in ignoring some information. In addition, competencies should be determined by various types of evaluation and measurement approaches that include harmony and consensus as well.

No study based on competency perception has been found that has used direct scaling techniques. Nevertheless, using the results from the work of a small number of scaling studies is recommended in the development of scaling and to determine the dimensions of competencies by experts. In this study, the focus will be the scaling methods based on binary comparison and ranking judgment and the comparison of the consistency of results obtained from those scaling methods.

Method

The aim of this study is to determine the competencies in the field of measurement and evaluation of primary school teachers in Turkey formed by their experiences in regard to their previous teacher training. In order to accomplish this, research was carried out using the Delphi technique to determine measurement and evaluation competencies and scaling of teachers' views.

The success of Delphi research depends largely on the careful selection of experts in related fields. The most important feature of Delphi studies that will increase the reliability of a particular study is the selection of appropriate experts. Therefore, according to Lang

⁴ This study was produced from some of the data of R. Serkan Arık's doctorate thesis carried out in Ankara University under Ömer Kutlu's consultancy.

⁵ Dr. - Dumlupınar University, Faculty of Education - rserkan.arik@dpu.edu.tr

⁶ Asst. Prof. Dr. - Ankara University, Faculty of Educational Science - omerkutlu@ankara.edu.tr

(1994), a random selection of participants is not acceptable. Instead, considering the expertise of each member of the panel for the selection of a Delphi study is required. The study group of teachers is comprised of experienced teachers who are trained in the field of measurement and evaluation.

Members of the panel should be representative of the views of experts in the field. As a result of their experiences and qualities, panel members should be able to provide a deep overview and should have significant opinions on the topic. It is possible to work with large or small expert groups using the Delphi technique, but there must be at least seven participants in the group. Groups can have 100 or more members (Şahin, 2001).

Considering the counties and school districts with the largest number of teacher appointments in central Ankara, three counties were chosen that have better transportation facilities and more teachers willing to participate in the research. These counties are Çankaya, Mamak, and Altındağ. The schools chosen for the study group in Çankaya have a high socio-economic level; those in Mamak represent a middle socio-economic level; and those chosen in Altındağ have a low socio-economic level. The schools have been selected from each district according to the number of teachers they have; schools with the largest number of teachers were chosen. The target number of teachers for conducting the study is determined as 468. Attention has been given to keep the number of schools larger in order for the 468 selected teachers to be able to represent the differences. Research has been conducted with 468 teachers who served at nine different schools. When the schools were visited for the first part of the determination of the Delphi survey, the research group was limited to 334 teachers because of appointments, medical reports, and other reasons, such as unwillingness to participate voluntarily. At the first Delphi survey, teachers were asked about their expertise, knowledge, and experience in the fields of measurement and evaluation, and 300 teachers who met the criteria were included in the implementation of the second and third Delphi panels. As a result, the data of 300 teachers have been included in all three Delphi panels.

Results, Discussion, and Suggestions

In this study, using two different scaling techniques, the method of paired comparison and ranking, revealed that there are different rankings in scaling. However, from the values of the scale of this difference, there is a very small difference in the ranking of scale values. When the highest value of the three scales and the lowest value of the three scales were observed, it was found that the value of the scales is very close. From scaling with binary comparison, in the second and third Delphi applications, “to prepare cognitive, affective, and psychomotor measuring instruments, to use technological tools and materials to do measurement and evaluation, and to use various question formats,” competencies seemed more significant to teachers, whereas “to enable the student to make self-assessment and peer-assessment, to make performance measurements based on the project and portfolio” competencies appeared to be least important.

To determine the correlation between the second and the third Delphi implementations, Pearson correlation coefficient is used. Accordingly, the second and third Delphi applications and scaling values obtained high correlation by comparison ($r = 0.993$, $p < .05$), and the scale values obtained by ranking and scaling are found to have a high correlation ($r = 0.991$, $p < .05$). It is noted that the correlation coefficient of paired comparison and ranking scaling needed also for the Delphi applications is considered an exact indication of reconciliation. In this study, Delphi studies have consensus, participation of privacy, and

controlled feedback. Scaling is consistent with the relationship between the results of the data collected, which should be high with the Delphi technique if the study produced reliable results.

Based on the findings, suggestions are presented for assessment and evaluation in the field of research and practice, as well as for working professionals who then do the research. As for the specialists working in the field of measurement and evaluation in practice, teachers have to learn about their interests and have to represent students' needs when it would be appropriate to use different measurement methods for an efficient educational experience for students. Measurement methods are recommended for use by the most versatile contemporary educational systems. According to the results of the research, it appears that teachers do not consider self-assessment and peer assessment as important factors in determining competencies. Informing teachers about the need to increase the use of these methods may be recommended.

When the studies that use Delphi applications are examined, it is noted that there is a lack of use of different statistical techniques and an over-use of mean and median. After that, similar studies and inclusion of scaling techniques in the studies will provide an improvement of Delphi applications, as well as scaling studies. Pairwise comparisons and judgments by scaling methods have been used in this study. The researchers suggested that for future studies, the use of classification judgments, absolute judgments, and degree total scaling techniques can provide recognition and the widespread use of scaling techniques. The study benefited from one-dimensional scaling techniques. The purpose of this study comparing teachers' competencies in terms of features was not taken into account. After that, the scaling studies using multidimensional scaling techniques are recommended to increase knowledge in this field of research.

Key Words: Teachers' competency, Measurement and Evaluation, scaling, Delphi Technique

Atıf için / Please cite as:

Arık, R. S. & Kutlu, Ö. (2013). Öğretmenlerin ölçme ve değerlendirme alanı yeterliklerinin yargıcı kararlarına dayalı ölçeklenmesi [Scaling primary school teachers' competence based on judgmental decisions in the field of measurement and evaluation]. *Eğitim Bilimleri Araştırmaları Dergisi - Journal of Educational Sciences Research*, 3 (2), 163–196. <http://ebad-jesr.com/>