

İki Ölçekleme Yönteminin Karşılaştırılması: İkili Karşılaştırma ve Sıralama Yargıları

Meltem ACAR GÜVENDİR¹, Yeşim ÖZER ÖZKAN²

ÖZET

Çalışmanın amacı, yargıcı kararlarına dayalı ölçekleme yöntemlerinden ikili karşılaştırma ve sıralama yargılarına dayalı ölçekleme yöntemlerinden elde edilen ölçek değerleri arasındaki tutarlılığı belirlemektir. Bu amaç doğrultusunda, eğitim fakültesinde öğrenim görmekte olan 395 dördüncü sınıf öğrencilerinden ölçme ve değerlendirme dersini tekrar alma gibi bir şansları olsaydı bu dersin ağırlıklı olarak hangi öğretim yöntemine göre yürütülmesi gerektiğine ilişkin olarak görüşleri alınmıştır. İkili karşılaştırmaya dayalı ölçekleme sonucunda öğrenciler tarafından ölçme ve değerlendirme dersinin ağırlıklı olarak yürütülmesi gereken en önemli yöntem “Örnek Olay Yöntemi” ardından sırasıyla gösterip yaptırma, problem çözme, tartışma ve bireysel çalışma ve son sırada anlatma yöntemi; sıralama yargılarına dayalı olarak yapılan ölçekleme sonucunda ise benzer şekilde ilk sırada “Örnek Olay Yöntemi” ardından sırasıyla gösterip yaptırma, tartışma, problem çözme, anlatma ve son sırada bireysel çalışma yöntemi yer almaktadır. Her iki yöntemle elde edilen ölçek değerleri arasındaki tutarlılığı belirlemek amacıyla Spearman-Brown sıra farkları korelasyon katsayısı hesaplanmıştır. Elde edilen sonuca göre iki ölçekleme yöntemi benzer sonuçlar üretmektedir.

Anahtar Kelimeler: Ölçekleme, İkili karşılaştırma yöntemi, Sıralama yargıları yöntemi, Öğretim yöntemi

 DOI Number: <http://dx.doi.org/10.12973/jesr.2013.316a>

¹ Arş. Gör. Dr. - Trakya Üniversitesi, Eğitim Fakültesi - meltemacar@gmail.com

² Yrd. Doç. Dr. - Gaziantep Üniversitesi, Gaziantep Eğitim Fakültesi - yozer80@gmail.com

GİRİŞ

Ölçekleme, diğer bir deyişle, ölçek yapma işlemi ilk olarak psikolojinin bir alt dalı olan psikofizikte karşımıza çıkmıştır. Psikofizik, ölçülmek istenen özelliğin ya da fiziksel uyarıcının kişide oluşturduğu sinirsel 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)

Ölçeklemenin birinci anlamı ölçme aracı geliştirmek, ikinci anlamı, sınıflama düzeyinde elde edilen verileri eşit aralıklı ölçek düzeyine getirebilmektir; çünkü sınıflandırma ölçeğinde birim yoktur. Ölçekleme, bir ölçekle toplanan verilerin niteliğini artırdığı gibi aynı zamanda, sadece ölçme aracı geliştirmek değil, basit ölçme sonuçlarını işleyerek standart ölçme aracı geliştirmektir (Turgut & Baykul, 1992). Stevens'a göre (1966) ölçekleme, fiziksel objelerin bir kural dâhilinde sayılarla eşleştirilmesidir. Ona göre ölçeklemede yer alan bu fiziksel objeler, çoğunlukla tutumlar veya inançlardır. Ölçekleme temelde puanlama yapma amacıyla kullanıldığı gibi, bir kavramın tek veya çok boyutlu olduğunu belirlemek için de kullanılabilir.

Ölçekleme işlemleri yargı ve tepki kanununa göre yapılmaktadır. Yargı kanununda uzman kanıları dikkate alınır ve bu kanun uyarıcı merkezli bir yaklaşımdır, tepki kanununda deneklerin tepkilerine göre ölçekleme yapılır ve bu kanun tepki yaklaşımı olarak da bilinir. Uyarıcı merkezli yargı yaklaşımında amaç, uyarıcıları, gözlemci ya da uzman yargılarına dayalı olarak belirlenmiş bir boyutta ölçeklemektir. Bu yöntemde N tane gözlemciden K tane uyarıcıyı uyarıcılık derecelerini belli yöntemlerle belirtmeleri istenir. Bu yöntemler sıralama, sınıflama ve ikili karşılaştırma gibi yöntemlerdir. Burada önemli olan gözlemcilerin her bir uyarıcının ölçekleme boyutundaki yerini diğer uyarıcılarla karşılaştırarak belirlemesidir. Sonuç olarak uyarıcının ölçek değeri gözlemci yargılarının ortalama değeri ile elde edilir (Stevens, 1966; Turgut & Baykul, 1992). Tepki yaklaşımında ise K tane uyarıcıyı, N kişilik bir denekler grubuna uygulayarak onların tepkileri toplanır. Bu yaklaşımda tepkiyi veren kişiler, tarafsız bilirkişi rolünde değil, kendi tepkilerini belirten denek durumundadırlar. Dolayısıyla bu yöntemde bir denekğin görevi, uyarıcıların ölçekleme boyutundaki yerini tarafsızca tayin etmek değil, her uyarıcının ölçekleme boyutundaki yerini aynı boyuttaki kendi yerine göre belirtmektir (Turgut & Baykul, 1992).

Yargı yaklaşımından ikili karşılaştırma yöntemi, Thurstone tarafından geliştirilmiştir ve 10 yıldır, psikolojik araştırmalarda, özellikle bilişsel alanlarda kullanılmaktadır (Bradley, 1984; Nishisato, 1994; Martignon & Hoffrage, 2002). Bu yöntem basit bir ikili karşılaştırmadır. İkili karşılaştırma yaklaşımı ile uyarıcı kümesine veya maddelere karar verilir. Gözlemcilere bu uyarıcıların mümkün olan tüm ikili çifti verilir ve onlardan verilen bu çiftleri daha iyi belirlenmiş seçim ölçütüne göre karşılaştırmaları istenir. Yargıcı kararları için geliştirilmiş olan bu yöntemin bir avantajı basit ve karşılaştırmalı yargıların tümünü kapsamasıdır (Brown & Peterson, 2009). Titchener (1901), karşılaştırmalı yargılarla ilgili olarak, bireylerin belirli sayıda uyarıcı verildiğinde bu uyarıcıların hangilerinden daha çok hoşnut oldukları ile ilgili yargılarının kesin olmayacağını, ancak bu uyarıcıların ikili şekilde verildiğinde daha kesin yargılarda bulduklarını belirtmiştir. Benzer şekilde Nunnally (1976), insanların günlük hayatlarında kesin yargılarda bulunmaktan kaçındıklarını, uyarıcının mutlak önemi yargılandığında bilindik bir şekilde hatalı yargılamalar yapabildiklerine değinmiştir.

Karşılaştırmalı yargı kanununun uygulanmasında beş hal belirtilir (Guilford, 1954; Torgerson, 1958). I. Hal ölçeklemede karşılaştırmadan elde edilen p_{jk} ve z_{jk} değerlerini kullanarak doğrudan ölçeklemeye imkân vermediğinden, II. Hal ölçeklemenin de

çözümünün mümkün olmamasından dolayı diğer hallerden III. ve V. Hal ölçeklemede tercih edilir. III. ve V. Hal karşılaştırmalı yargı kanununun basitleştirilmiş halleridir. III. ve V. Hal denklemleri arasındaki fark ise ölçekleme yapılırken V. Hal denkleminde ayırt etme yargıları varyanslarının eşit kabul edilmesine karşın III. Hal denkleminde ayırt etme yargıları varyansları eşit kabul edilmeyip bu varyansların tek tek bulunmasıdır. Gerçekte varyanslar arası farklar büyük iken varyansları eşit sayarak V. Hal yöntemiyle ölçekleme yapmak işlemlerde biraz ekonomi sağlar (Turgut & Baykul, 1992).

Yargı yaklaşımından sıralama yargılarına dayalı ölçekleme yöntemi ise, bir sıra numarası verilebilecek tüm uyarıcılara uygulanabilir. Televizyon spikerlerinin beğenilme dereceleri, bir yarışmaya katılan resimler veya şarkılar gibi konulardaki ölçekleme problemleri bu yöntemin kullanılış alanına örnek olarak gösterilebilir (Turgut & Baykul, 1992). Bu yöntemde öncelikle gözlemciler/yargıcılara uyarıcıların tümü verilir. Yargıcılardan verilen her bir uyarıcıyı diğer uyarıcılarla karşılaştırarak bir sıra sayısı belirlemesi istenir. Böylelikle uyarıcıların tümü, her bir uyarıcının karşılaştırıldığı birleşik bir standarda dönüştürülmüş olur. Uyarıcılara verilen sıralama yargılarından elde edilen oranlar bu birleşik standartla karşılaştırılarak ölçek değerleri elde edilir (Guilford, 1954; Turgut & Baykul, 1992).

Ölçekleme teknikleri konusunda alanyazında sınırlı sayıda çalışmanın yapıldığı görülmektedir. Darcy, Lee ve Tracey (2004), çok kültürlü danışmanlık uygulamaları arasındaki farkları iki ayrı ölçekleme yöntemini kullanarak belirlemeye çalışmışlardır. Çalışmada üzerinde durulan konu çok kültürlü danışmanlık uygulamalarıdır. Kullanılan ölçekleme yöntemleri, ikili karşılaştırmaya dayalı ölçekleme yöntemi ile çok boyutlu ölçeklemedir.

Anıl ve Güler (2006) yargı yaklaşımına dayalı ikili karşılaştırma yöntemi ile nitelikli bir öğretilerde bulunması istenilen sekiz özelliği öğretmen adaylarının yargılarına göre ölçeklemişlerdir. Çalışmanın sonucunda nitelikli bir öğretilerde bulunması istenilen birinci özelliğin meslek sevgisi olduğu ortaya çıkmıştır.

Kan (2008), yargıcı kararlarına dayalı iki ölçekleme yönteminin benzer sonuçları üretip üretmediğini belirlemek amacıyla yapmış olduğu çalışmada tutum ölçeğini kullanarak elde etmiş olduğu veriler üzerinde sınıflama ve sıralama kararlarına dayalı ölçekleme yöntemlerini uygulamıştır. Bu iki ölçekleme yönteminden elde edilen ölçek değerleri arasındaki tutarlılığı incelemek için Sperman rho korelasyon katsayısını hesaplamıştır. Araştırma sonucunda her iki yöntemden elde edilen ölçek değerleri arasındaki korelasyonun pozitif ancak düşük düzeyde bir ilişki olduğu sonucuna ulaşmıştır.

Öğretmen (2008), alan tercihleri envanteri üzerinde ölçekleme, geçerlik ve güvenilirlik çalışmalarını yürütmüştür. Çalışmanın amacı öğretmen yetiştiren bölümlerde okuyan öğretmen adaylarının belirli alanlara karşı tercihlerini ölçmek ve öğretmeyi en çok tercih ettikleri konu alanlarına ait öncelikleri belirlemektir. Bunun için ikili karşılaştırmaya dayalı yargı kanununun V. Hal denklemi ile ölçekleme yapmıştır. Bunun yanında alan yazında yapılan çalışmalarda verilerin güvenilirlik ve geçerlik çalışmalarına yer verilmediğini belirterek, ölçek değerlerine ilişkin iç tutarlılık sonuçlarına da yer vermiştir.

Brown ve Peterson (2009), yayınlamış oldukları teknik raporda, ikili karşılaştırma yargılarına dayalı ölçekleme yöntemi konusunda kuramsal bilgilere yer vermişler, bunun yanında yöntemin geçerlik ve güvenilirlik çalışmaları hakkında bilgiler sunmuşlardır. Ayrıca raporda kuramsal bilgiler örnekler ile de desteklenmiştir.

Güler ve Anıl (2009) öğretim elemanlarının lisansüstü eğitime başvuran öğrencilerde bulunmasını istedikleri özelliklerin neler olması gerektiğini öğretim elemanlarının verdikleri

tepkilere göre ikili karşılaştırmayla ölçekleme yöntemi kullanarak belirlemişlerdir. Araştırma sonucunda, öğrencilerde bulunması gereken en öncelikli özelliğin akademik başarı puanı ardından mülakat puanı, programa girme amacı, İngilizce yeterlik puanı, ALES puanı, bölüm öğrencisi olup olmaması ve en son sırada referansı olup olmaması izlemiştir.

Özer ve Acar (2011), öğretmen yeterlikleri üzerinde yapmış oldukları çalışmalarında, öğretmen adaylarından bir öğretmende bulunması gereken öğretmenlik mesleği genel yeterliklerini daha önemliden en az önemliye göre ikili olarak karşılaştırmalarını istemişlerdir. Elde edilen veriler karşılaştırmalı yargı yaklaşımından ikili karşılaştırma yönteminin V. Hal denklemi ile tam veri matrisinden ölçeklenmiştir. Araştırma sonucunda öğrenciler açısından bir öğretmende bulunması gereken en öncül yeterlik alanı “öğrenciyi tanıma” olarak belirlenirken bunu sırasıyla “öğrenmeyi, gelişimi izleme ve değerlendirme”, “kişisel ve mesleki değerler-mesleki gelişim”, “öğrenme ve öğretme süreci”, “okul-aile ve toplum ilişkileri” ve son sırada “program ve içerik bilgisi” yeterliği izlemiştir.

Öztürk, Özdemir ve Gelbal (2011), sınıflama yargılarına dayalı olarak elde edilen ölçek değerleri ile denek tepkilerine dayalı ölçek değerleri arasındaki tutarlılığı belirlemeye çalışmışlardır. Araştırma sonucunda, iki farklı ölçekleme yaklaşımına göre elde edilen ölçek değerleri arasında orta düzeyde bir ilişki olduğu görülmüştür.

Bal (2011) yaptığı çalışmada lise birinci sınıf öğrencilerinin seviye belirleme sınavı (SBS) başarılarında etkili olduğu düşünülen faktörleri, sıralama yargıları kanunuyla ölçekleyerek belirlemiştir. Araştırmada, öğrenci özelliklerini belirlemeyi amaçlayan sorularla, ölçekleme yapılacak sorulardan oluşan bir veri toplama aracı kullanılmıştır. Çalışma sonucunda elde edilen bulgulara göre, kendine güven, başarıma isteği, dikkat ve güdü gibi bireysel özellikler, öğrencilerin SBS başarılarını etkileyen faktörlerin başında gelmektedir. Bunu sırasıyla, öğrencinin çalışma yöntemi ile SBS'ye hazırlanırken aile ve arkadaş gibi yakın çevreden aldığı destek izlemektedir. Dershaneye gitme, özel ders alma, soru çözme ve öğretmen desteği ise, SBS başarısında etkili olduğu düşünülen faktörler olarak alt sıralarda yer almıştır. Ayrıca öğrenciler, kitap okuma ve kendine zaman ayırma gibi sosyal etkinlikleri son sıraya koymuşlardır.

Albayrak ve Gelbal (2012), öğrencilerin internet kullanma amaçlarını iki farklı yöntem olan ikili karşılaştırma yargılarına dayalı ve sıralama yargılarına dayalı ölçekleme yöntemleri ile ölçeklendirmişlerdir. Çalışma sonucunda, kullanılan bu iki ölçekleme yaklaşımının benzer sonuçlar üretmediği, bu nedenle ölçekleme çalışmalarında seçilen ölçekleme yöntemine ve sayıtlılarına dikkat edilmesi gerektiği vurgulanmıştır.

Alanyazında ölçekleme ile ilgili olarak sosyal bilimlerin yanında işletme, ekonometri, sağlık gibi farklı alanlarda çok boyutlu ölçeklemenin kullanıldığı çalışmalar bulunmaktadır (Steyvers, 2001; Groenen & Velden, 2004; Şimşek, 2006; Yiğit, 2007; Alan, 2008). Ancak yukarıda değinildiği gibi tek boyutlu ölçekleme ile ilgili yapılan çalışmalar oldukça sınırlı sayıdadır.

Bu çalışmada ölçekleme yaklaşımlarından ikili karşılaştırma ve sıralama yargılarına dayalı ölçekleme yöntemleri ve her iki yöntemden elde edilen sonuçlar arasındaki ilişkinin belirlenmesi üzerinde durulmuştur. Türkiye’de ve yurt dışında ölçekleme alanında yapılan araştırmaların son derece sınırlı olmasından ve ayrıca ölçekleme yöntemlerinden elde edilen ölçek değerleri arasındaki tutarlılık konusunda bilgi edinileceğinden böyle bir çalışmanın yapılmasına gereklilik görülmüştür.

Çalışmanın Amacı ve Önemi

Bu çalışmada amaç, yargıcı kararlarına dayalı ölçekleme yöntemlerinden ikili karşılaştırma ve sıralama yargılarına dayalı ölçekleme yöntemlerinden elde edilen ölçek değerleri arasındaki tutarlılığı belirlemektir. Böylelikle farklı yöntemlere göre ölçekleme yapıldığında uyarıcıların ölçek değerlerinin farklılaşıp farklılaşmadığı belirlenmeye çalışılmıştır. Bu sayede uyarıcıların sahip olduğu sıra numaralarına ne kadar güvenileceği ortaya konulmuştur. Aynı zamanda alanyazında ölçekleme ile ilgili sınırlı sayıda çalışma yer almaktadır. Böylelikle, yapılan bu çalışmanın ölçekleme alanında sınırlı olan çalışmalara katkı sağlayacağı düşünülmektedir.

YÖNTEM

Araştırma Modeli

Araştırmada sıralama ve ikili karşılaştırma yargılarına dayalı ölçekleme yöntemleri ile elde edilen ölçek değerleri arasındaki tutarlılığın belirlenmesi amaçlandığı için bu çalışma, temel bir araştırma niteliğindedir.

Çalışma Grubu

Araştırmanın çalışma grubunu, Trakya Üniversitesi Eğitim Fakültesinden 254 ve Kilis 7 Aralık Üniversitesi Eğitim Fakültesinden 141 olmak üzere toplam 395 dördüncü sınıf öğrencisi oluşturmaktadır. Çalışma için dördüncü sınıf öğrencilerinin seçilmesinin başlıca iki nedeni vardır. Bunlardan ilki öğrencelere yöneltilen öğretim yöntemlerini ikinci sınıf düzeyindeyken aldıkları “Öğretim İlke ve Yöntemleri” dersinde öğrenmiş olmaları; ikinci nedeni ise bu öğrencilerin önceki yıl veya dönemde ölçme ve değerlendirme dersi almış olmaları ve bu sayede öğrencilerin, ölçme ve değerlendirme dersini tekrar alma gibi bir şansları olsaydı bu dersin ağırlıklı olarak hangi öğretim yöntemine göre yürütülmesi gerektiği ile ilgili daha doğru yargıda bulunabilecekleri düşüncesidir.

Veri Toplama Aracı

Öğretim yöntemleri ve bu yöntemlerin kullanılış biçimleri bilim adamları tarafından farklı şekillerde sınıflandırılmıştır. Bu çalışmada veri toplama aracının hazırlık aşamasında, alan yazında geçen altı öğretim yöntemi esas alınmıştır (Demirel, 2006). Bu çalışmada Demirel’in (2006) sınıflandırılmasının esas alınmasının nedeni, bu yöntemlerin değişik sınıf düzeylerinde sıkça kullanılan yöntemler olmasıdır. Ayrıca bu sınıflandırma öğrenme-öğretme süreçlerinde ve eğitim durumlarının düzenlenmesinde kullanılabilecek bazı genel öğretim yöntemlerini içermektedir. Veri toplama aracı olarak, “Öğretim Yöntemleri Değerlendirme Formu” hazırlanmıştır. Belirlenen altı öğretim yöntemi öğrencilerin sıralama ve ikili karşılaştırma yapabilecekleri şekilde düzenlenmiştir. Sıralama yargılarına dayalı ölçekleme yapabilmek için öğrencilerden bu altı öğretim yöntemini ölçme ve değerlendirme dersini düşünerek kendilerince öncelikli olarak tercih ettikleri şekilde her bir uyarıcıyı tüm uyarıcılarla karşılaştırarak bir sıra numarası istenmiştir. İkili karşılaştırmalara dayalı ölçekleme yapabilmek için ise aynı formda belirlenen öğretim yöntemleri ikili şekilde karşılaştırma yapabilecekleri şekilde verilmiştir. Bu şekilde öğretim yöntemleri değerlendirme formunda yer alan altı öğretim yöntemi hem sıralı şekilde hem de ikili şekilde öğrencilere sunulmuştur.

Verilerin Analizi

Çalışmada öğretim yöntemleri değerlendirme formu ile toplanan verilere sıralama kararlarına dayalı ve Thurstone'un karşılaştırmalı yargı kanununun V. Hal denklemi ile tam veri matrisinden ölçekleme işlemleri uygulanmıştır. İkili karşılaştırmalara dayalı ölçekleme için öncelikle öğrencilere altı öğretim yöntemi ikili şekilde sunulmuştur. İkili olarak yapmış oldukları karşılaştırmalara ait frekans değerleri belirlenmiş ve ardından bu değerler ile frekans matrisi oluşturulmuştur. Frekans matrisinden yola çıkarak oranlar matrisi bulunmuştur. Oranlar matrisinin elemanlarına karşılık gelen z değerleri belirlenerek birim normal sapmalar matrisinin oluşturulmasına geçilmiştir. Birim normal sapmalar matrisinin en alt satırına her bir sütuna ait değerlerin toplamı alınmış ve bu satırdaki her bir z değerinin sütunlar boyunca ortalamaları hesaplanmış ve böylelikle ölçek değerleri bulunmuştur. Eksenin başlangıç noktasını z ortalama değerlerinin en küçüğüne kaydırmak için, en küçük z ortalama değerinin mutlak değeri tüm değerlere eklenecek şekilde ölçek değerleri sıralanmıştır.

Sıralama yargılarına dayalı yargıların toplanması için öncelikle öğrencilere, altı tane öğretim yönteminin tümü verilmiştir. Öğrencilerin öğretim yöntemlerinin tümünü birden düşünmesi ve her bir yöntemi diğer yöntemlerle karşılaştırarak her bir öğretim yöntemine bir sıra sayısı tayin etmesi istenmiştir. Verilerin analizinde öncelikle, sıra frekansları matrisi oluşturulmuş ve sıralama yargılarından oranlar matrisi elde edilmiştir. Oranlar matrisinin elemanlarına karşılık gelen z değerleri belirlenerek birim normal sapmalar matrisinin oluşturulmasına geçilmiştir. Birim normal sapmalar matrisinin en alt satırına her bir sütuna ait değerlerin toplamı alınmış ve bu satırdaki her bir z değerinin sütunlar boyunca ortalamaları hesaplanmış ve böylelikle ölçek değerleri bulunmuştur.

Ölçek değerlerinin tutarlılığının belirlenmesi için ortalama hata ve bu değer anlamı olup olmadığını belirlemek için ise ki-kare istatistiği hesaplanmıştır.

Her iki yöntemden elde edilen ölçek değerleri arasındaki tutarlılığı incelemek için ise Spearman-Brown sıra farkları korelasyon katsayısı hesaplanmıştır.

BULGULAR

Araştırmada ölçekleme tekniklerinden ikili karşılaştırma ve sıralama yargılarına ölçekleme yöntemleri ile ölçme ve değerlendirme dersinin ağırlıklı olarak hangi öğretim yöntemi ile yürütülmesi gerektiği ölçeklenmiştir. Öncelikle öğretim yöntemlerinin ikili karşılaştırmalara dayalı olarak yapılan ölçekleme sonuçları verilmiştir.

Tablo 1. İkili karşılaştırmalara dayalı elde edilen öğretim yöntemleri ile ilgili frekans matrisi (F)

	A	B	C	D	E	F
A	-	112	43	65	69	201
B	283	-	132	142	145	273
C	352	264	-	200	213	306
D	330	253	195	-	221	284
E	326	250	182	174	-	293
F	194	122	89	111	103	-

Tablo 1'de öncelikle öğretim yöntemleri öğrenciler tarafından hangi öğretim yönteminin daha öncelikli olarak kullanılması gerektiğine ilişkin ikili olarak karşılaştırılmış ve karşılaştırma sonucunda her bir yönteme ait frekans değerlerini gösteren frekanslar

matrisi (F) oluşturulmuştur. Frekans matrisi, satırdaki uyarıcının sütundaki uyarıcıya göre tercih edilme durumuna göre oluşturulmuştur. Örneğin, A öğretim yöntemini B öğretim yöntemine tercih edenlerin sayısı 112'dir. Benzer şekilde B öğretim yöntemini A öğretim yöntemine tercih edenlerin sayısı 283'tür. Bu durumda F matrisinin sütun toplamı öğretim yöntemlerinin tercih edilmeme sayısını, satır toplamları ise tercih edilme sayısını göstermektedir. Ardından frekans matrisinin her bir elemanı ikili karşılaştırmayı yapan öğrenci sayısına bölünerek oranlar matrisi (P) bulunmuştur. Araştırmada ikili karşılaştırmayı yapan öğrenci sayısı 395 olduğundan frekans matrisinin her elemanı bu sayıya bölünmüş ve oranlar matrisi oluşturulmuştur. Tablo 2'de oranlar matrisine yer verilmiştir.

Tablo 2. Öğretim yöntemleri ile ilgili oranlar matrisi (P)

	A	B	C	D	E	F
A	-	0,284	0,109	0,165	0,175	0,509
B	0,716	-	0,334	0,359	0,367	0,691
C	0,891	0,668	-	0,506	0,539	0,775
D	0,835	0,641	0,494	-	0,559	0,719
E	0,825	0,633	0,461	0,441	-	0,742
F	0,491	0,309	0,225	0,281	0,261	-

Tablo 2'de oluşturulan oranlar matrisinde dikkat edilmesi gereken, matrisin esas köşegene göre simetrik olan elemanlarının toplamının bire eşit olması gerektiğidir. Bu kontrolün yapılmasının ardından, birim normal sapmalar matrisinin (Z) oluşturulmasına geçilmiştir. Z matrisi için, P matrisinin elemanlarına karşılık gelen z değerleri belirlenmiştir. Z matrisinin elemanları, esas köşegene göre işaret yönünden ters ve mutlak değer açısından eşittir. Tablo 3'teki matrisin her bir sütunun toplamları bulunarak $\sum Z_j$ satırına yazılmıştır. Ardından bulunan her bir sütunun toplamı uyarıcı sayısına bölünmüştür. Burada uyarıcı sayısı altı adet öğretim yöntemi olduğundan $\sum Z_j$ değerlerinin altıya bölünmesiyle Z değerleri bulunmuştur.

Tablo 3. Öğretim yöntemleri ile ilgili birim normal sapmalar matrisi (Z)

	A	B	C	D	E	F
A	-	-0,572	-1,233	-0,976	-0,936	0,022
B	0,572	-	-0,428	-0,360	-0,340	0,499
C	1,233	0,435	-	0,016	0,099	0,754
D	0,976	0,360	-0,016	-	0,150	0,580
E	0,936	0,340	-0,099	-0,150	-	0,649
F	-0,022	-0,499	-0,754	-0,580	-0,641	-
$\sum Z_j$	3,694	0,063	-2,530	-2,049	-1,668	2,504
\bar{z}_j	0,616	0,011	-0,422	-0,342	-0,278	0,417
S_j	1,038	0,433	0,000	0,080	0,144	0,839

Tablo 3'e bakıldığında \bar{z}_j değerlerinin en küçüğü C öğretim yöntemine ait olan -0,422 değeridir. Eksenin başlangıç noktası kaydırılarak S_j değerleri bulunur. Bunun için her bir Z_{jort} değerine -0,422 değerinin mutlak değeri olan 0,422 değeri eklenir. Bu eklemenin ardından elde edilen S_j değerleri A, B, C, D, E, F olarak adlandırılan altı adet öğretim yöntemine ait ölçek değerleridir. Bu ölçek değeri ve sıraları Tablo 4'te verilmiştir.

Tablo 4. İkili karşılaştırmalara dayalı öğretim yöntemleri ölçek değerleri ve sıraları

	Ölçek Değerleri (S _i)	Sıraları
C- Örnek Olay Yöntemi	0,000	1
D- Gösterip Yaptırma Yöntemi	0,080	2
E- Problem Çözme Yöntemi	0,144	3
B- Tartışma Yöntemi	0,433	4
F- Bireysel Çalışma Yöntemi	0,839	5
A- Anlatma Yöntemi	1,038	6

Tablo 4 incelendiğinde ikili karşılaştırma yöntemine göre ölçekleme yapıldığında öğrenciler tarafından ölçme ve değerlendirme dersinin ağırlıklı olarak yürütülmesi gereken en önemli yöntem “Örnek Olay Yöntemi” olarak belirlenmiştir. Ardından sırasıyla gösterip yaptırma, problem çözme, tartışma ve bireysel çalışma ve son sırada anlatma yöntemi yer almaktadır.

Öğrencilerin ikili şekilde kendilerine verilen altı öğretim yöntemini uygun şekilde karşılaştırma yapıp yapmadıklarını kontrol etmek için elde edilen ölçek değerlerinin iç tutarlılığı hesaplanmıştır. Bu amaç doğrultusunda birim normal sapmalar, oranlar ve en son olarak hatalar matrisleri oluşturulmuştur. Ayrıca elde edilen ortalama hata değerinin anlamlı olup olmadığı da ki-kare istatistiği ile belirlenmiştir. Elde edilen bulgular Tablo 5’te verilmiştir.

Tablo 5. Ölçek değerlerinin iç tutarlılığına ilişkin bulgular (ikili karşılaştırma)

Hatalar matrisi					Oran farklarının karesi matrisi				
Anlatma Yöntemi	-				-				
Tartışma Yöntemi	0,011	-			0,496	-			
Örnek Olay Yöntemi	0,041	0,001	-		12,227	0,004	-		
Gösterip Yaptırma Yöntemi	0,005	0,003	0,026	-	0,128	0,030	2,177	-	
Problem Çözme Yöntemi	0,011	0,019	0,018	0,034	-	0,677	1,310	1,074	3,861
Bireysel Çalışma Yöntemi	0,087	0,033	0,025	0,057	0,017	25,310	4,124	2,987	14,204
Σhata	0,387					χ ²	10,110		
Ortalama Hata	0,010					χ ² (0,05;10)	18,307	(tablo değeri)	

Tablo 5’te ikili karşılaştırmaya dayalı ölçekleme yönteminden elde edilen ölçek değerlerine ait ortalama hatanın düşük olduğu görülmektedir (ort.hata_{ikili}=0,010). Bu sonuçlara göre örneklemeye dâhil edilen öğrencilerin yargılarının güvenilir olduğu söylenebilir. Elde edilen bu değerlerin anlamlı olup olmadığının da kontrol edilmesi gerekir (Turgut & Baykul, 1992). Bunun için ki-kare istatistiği hesaplanmıştır. Hesaplanan ki-kare istatistiği tablo değerinden küçük olduğundan, bu sonuç öğrencilerin kararlarında tutarlı olduğunu ve aynı zamanda kullanılan ölçekleme yönteminin kullanılan verilere uygun olduğunu gösterir.

İkili karşılaştırma ile ölçekleme yönteminin ardından öğretim yöntemlerine sıralama yargılarına dayalı olarak ölçekleme yapılmıştır. İkili karşılaştırma yönteminde olduğu gibi burada da öncelikle 395 öğrencinin altı uyarıcıyı belirli boyutta sıralamalarıyla elde edilen sıra frekansları matrisi Tablo 6’da verilmiştir.

Tablo 6. Sıralama yargularına dayalı elde edilen öğretim yöntemleri ile ilgili frekans matrisi (F)

Ri	ri	Uyarıcılar						TOPLAM
		A	B	C	D	E	F	
1	6	187	15	7	16	19	151	395
2	5	78	72	16	45	54	130	395
3	4	24	87	53	73	105	53	395
4	3	25	90	97	77	82	24	395
5	2	24	71	142	71	70	17	395
6	1	57	60	80	113	65	20	395
		395	395	395	395	395	395	2370

Tablo 6’da yer alan frekans matrisinin oluşturulmasında ri esas alınmıştır. Örneğin A öğretim yöntemini altıncı sırada tercih edenlerin sayısı 187 iken birinci sırada tercih edenlerin sayısı 57’dir. Benzer şekilde C öğretim yöntemini altıncı sırada tercih edenlerin sayısı 7 iken, birinci sırada tercih edenlerin sayısı 80’dir. Frekans matrisinin sütun toplamaları $N^2 = (395)^2$ ’ye bölünerek oranlar matrisi elde edilmiştir. Oranlar matrisine ikili karşılaştırmalar yöntemi uygulanmıştır. Ardından, birim normal sapmalar matrisinin (Z) oluşturulmasına geçilmiştir. Tablo 7’deki Z matrisini elde etmek için, P matrisinin elemanlarına karşılık gelen z değerleri belirlenmiştir.

Tablo 7. Öğretim yöntemleri ile ilgili birim normal sapmalar matrisi (Z)

	A	B	C	D	E	F
A	-	-0,616	-0,819	0,180	0,196	0,842
B	0,616	-	-0,374	-0,215	-0,020	0,816
C	0,819	0,374	-	0,180	0,399	1,181
D	0,729	0,215	-0,180	-	0,196	0,968
E	0,623	0,020	-0,399	-0,196	-	0,842
F	0,004	-0,816	-1,181	-0,968	-0,842	-
$\sum z_j$	2,791	-0,823	-2,953	-1,019	-0,070	4,648
\bar{z}_j	0,469	-0,141	-0,496	-0,160	-0,012	0,764
s_j	0,965	0,355	0,000	0,336	0,484	1,260

Tablo 7’ye bakıldığında \bar{z}_j değerlerinin en küçüğü C öğretim yöntemine ait olan -0,496 değeridir. Eksenin başlangıç noktası kaydırılarak S_j değerleri bulunmuştur. Bunun için her bir \bar{z}_j değerine -0,496 değerinin mutlak değeri olan 0,496 değeri eklenmiştir. Bu eklemenin ardından elde edilen S_j değerleri Tablo 8’de verilmiştir.

Tablo 8. Sıralama yargularına dayalı öğretim yöntemleri ölçek değerleri ve sıraları

	Ölçek Değerleri (S_j)	Sıraları
C- Örnek Olay Yöntemi	0,000	1
D- Gösterip Yaptırma Yöntemi	0,336	2
B- Tartışma Yöntemi	0,355	3
E- Problem Çözme Yöntemi	0,484	4
A- Anlatma Yöntemi	0,965	5
F- Bireysel Çalışma Yöntemi	1,260	6

Tablo 8'e bakıldığında sıralama yargılarına dayalı ölçekleme yapıldığında öğrenciler tarafından bir ölçme ve değerlendirme dersinin ağırlıklı olarak yürütülmesi gereken öğretim yöntemi, "Örnek Olay Yöntemi" olarak belirlenmiştir. Ardından sırasıyla gösterip yaptırma, tartışma, problem çözme, anlatma ve son sırada bireysel çalışma yöntemi yer almaktadır.

İkili karşılaştırmalara dayalı ölçekleme yönteminde olduğu gibi sıralama yargılarına dayalı ölçekleme yönteminden elde edilen ölçek değerlerinin iç tutarlılığı ve ki-kare istatistiği belirlenmiştir. Elde edilen bulgular Tablo 9'da verilmiştir.

Tablo 9. Ölçek değerlerinin iç tutarlılığına ilişkin bulgular (sıralama yargıları)

Hatalar Matrisi					Oran Farklarının Karesi Matrisi				
Anlatma Yöntemi									
Tartışma Yöntemi	0,002				0,017				
Örnek Olay Yöntemi	0,039	0,007			8,321	0,171			
Gösterip Yaptırma Yöntemi	0,032	0,078	0,060		4,403	19,955	12,366		
Problem Çözme Yöntemi	0,049	0,059	0,031	0,019	9,556	11,622	3,610	1,169	
Bireysel Çalışma Yöntemi	0,118	0,024	0,015	0,011	0,019	46,378	3,106	1,866	0,710
Σhata	0,840					χ ²	18,127		
Ortalama Hata	0,030					χ ² (0,05;10)	18,307	(tablo değeri)	

Tablo 9'da sıralama yargılarına dayalı ölçeklemeden elde edilen ölçek değerlerine ait ortalama hatanın düşük olduğu görülmektedir [ort.hata_{sıralama}=0,030]. Bu sonuçlara göre örnekleme dâhil edilen öğrenci yargılarının güvenilir olduğu söylenebilir. Ayrıca ki-kare istatistiğinin tablo değerinden küçük olması öğrencilerin kararlarında tutarlı olduğunu ve aynı zamanda kullanılan ölçekleme yönteminin kullanılan verilere uygun olduğunu göstermektedir.

İkili karşılaştırma ve sıralama yargılarına dayalı ölçekleme yöntemlerinden elde edilen ölçek değerleri arasındaki tutarlılığı belirlemek amacıyla Spearman-Brown sıra farkları korelasyon katsayısı hesaplanmıştır (r=0.886, p<0.01). Bu sonuca göre, ikili karşılaştırma yöntemiyle elde edilen ölçek değerleri ile sıralama yargılarına dayalı ölçekleme yöntemlerinden elde edilen ölçek değerleri arasında yüksek düzeyde, pozitif ve manidar bir ilişki olduğu görülmektedir. Tablo 4 ve 8 incelendiğinde en yüksek ölçek değerine sahip yöntemin iki ölçekleme yönteminde de "Örnek Olay Yöntemi" olduğu görülmektedir. Ayrıca "Örnek Olay Yöntemi" ve "Gösterip Yaptırma Yöntemi"nin iki ölçekleme tekniğine göre de aynı sıralara sahip olduklarını diğer yöntemlerinin ise birbirine yakın sıralara sahip olduğu gözlenmiştir. Bu bulguya göre iki ölçekleme yönteminin benzer sonuçlar ürettiğini diğer bir ifadeyle ölçek değerlerinin büyük ölçüde benzerlik gösterdiği söylenebilir.

SONUÇ, TARTIŞMA VE ÖNERİLER

Bu çalışmada, yargıcı kararlarına dayalı ölçekleme yöntemlerinden ikili karşılaştırma ve sıralama yargılarına dayalı ölçekleme yöntemlerinden elde edilen ölçek değerleri arasındaki tutarlılığın belirlenmesi amaçlanmıştır. Öncelikle ikili karşılaştırmaya dayalı ölçekleme yapabilmek için, öğretim yöntemleri öğrenciler tarafından hangi öğretim yönteminin daha öncelikli olarak kullanılması gerektiğine ilişkin ikili olarak karşılaştırılmıştır. İkili karşılaştırma yöntemine göre ölçekleme yapıldığında öğrenciler tarafından ölçme ve değerlendirme dersinin ağırlıklı olarak yürütülmesi gereken en önemli

yöntem, “Örnek Olay Yöntemi” olarak belirlenmiştir. Ardından sırasıyla gösterip yaptırma, problem çözme, tartışma, bireysel çalışma ve son sırada anlatma yöntemi yer almaktadır.

İkili karşılaştırma ile ölçekleme yönteminin ardından öğretim yöntemleri sıralama yargılarına dayalı olarak ölçeklenmiştir. Elde edilen sonuca göre, öğrenciler tarafından bir ölçme ve değerlendirme dersinin ağırlıklı olarak yürütülmesi gereken öğretim yöntemi, “Örnek Olay Yöntemi” olarak belirlenmiştir. Ardından sırasıyla gösterip yaptırma, tartışma, problem çözme, anlatma ve son sırada bireysel çalışma yöntemi yer almaktadır.

En yüksek ölçek değerine sahip yöntem iki ölçekleme yönteminde de “Örnek Olay Yöntemi”dir. Ayrıca “Örnek Olay Yöntemi” ve “Gösterip Yaptırma Yöntemi”nin iki ölçekleme türüne göre de aynı sıralara sahip olduklarını diğer yöntemlerin ise birbirine yakın sıralara sahip olduğu gözlenmiştir. İkili karşılaştırma ve sıralama yargılarına dayalı ölçekleme yöntemlerinden elde edilen ölçek değerleri arasındaki tutarlılığı belirlemek amacıyla hesaplanan Spearman-Brown sıra farkları korelasyon katsayısı her iki yöntemden elde edilen ölçek değerlerinin olumlu yönde ilişkili olduğunu göstermektedir. Bu bulguya göre iki ölçekleme yönteminin benzer sonuçlar ürettiği söylenebilir. Elde edilen bu bulgu O’Neil ve Chissom’ın (1993) çalışmalarındaki bulguyla benzerlik göstermektedir. O’Neil ve Chissom (1993) öğrencilerin tutumlarını ölçmek için hazırladıkları beş maddeyi ikili karşılaştırma, sıralama yargıları ve Likert tipi yöntemlerini kullanarak ölçeklemişlerdir. Üç ölçekleme yöntemine göre elde ettikleri ölçek değerlerini karşılaştırarak aralarındaki korelasyonu hesaplamışlardır. En yüksek ilişkinin ($r=0,89$) ikili karşılaştırma ve sıralama yargılarına dayalı geliştirilen ölçek maddeleri arasında olduğunu vurgulamışlardır.

Bu bulgulara dayanarak iki ölçekleme yöntemi de benzer sonuçlar üretmiştir. Çünkü iki yaklaşımda da gözlemcinin görevi, her uyarıcının ölçekleme boyutundaki büyüklüğünü diğer uyarıcılara göre belirlemektir. Aralarındaki tek farklılık ise ikili karşılaştırmada uyarıcılar gözlemcilere ikili verilmekte, sıralamada ise bütün uyarıcılar birlikte verilmektedir. En son adımda ise iki yaklaşımda da ölçek değerlerinin sıralaması vardır. Bu sonuçlara dayanarak uyarıcıların verilmiş şeklinin elde edilen ölçek değerlerini yüksek düzeyde etkilemediğini söyleyebiliriz.

Alanyazında farklı ölçekleme yaklaşımlarını karşılaştıran çalışmalar (O’Neil & Chissom, 1993; Kan, 2008; Öztürk, Özdemir & Gelbal, 2011; Albayrak & Gelbal, 2012) oldukça azdır. Bu nedenle ölçekleme yöntemlerinden elde edilen verilerin tutarlılığını ortaya koymak açısından farklı ölçekleme yöntemlerini karşılaştıran çalışmalar yapılması önerilebilir.

KAYNAKÇA

- Alan, S. (2008). Çok boyutlu ölçekleme yöntemi ve bir uygulama. *Yayımlanmamış Yüksek Lisans Tezi*. Marmara Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Albayrak, A. & Gelbal, S. (2012). İkili karşılaştırmalar yargılarına ve sıralama yargılarına dayalı ölçekleme yaklaşımlarından elde edilen ölçek değerlerinin tutarlılığının karşılaştırılması. *21. Ulusal Eğitim Bilimleri Kongresi*. Marmara Üniversitesi Eğitim Fakültesi. 12-14 Eylül 2012. İstanbul.
- Anıl, D. & Güler, N. (2006). İkili karşılaştırma yöntemi ile ölçekleme çalışmasına bir örnek. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 30, 30–36.
- 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). <http://www.insanbilimleri.com/en>. İndirme Tarihi: 12.10.2012.

- 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.
- Bradley, R. A. (1984). Paired comparisons: some basic procedures and examples. *Handbook of Statistics 4*. (Edt: P. R. Kirshnajah & P. K. Sen). Amsterdam: Elsevier. pp. 299-326.
- 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. Rep. RMRS-GTR-216WWW. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 98 p.
- Crocker, L. & Algina, J. (1986). *Introduction to classical & modern test theory*. Florida: Holt, Rinehart and Winston Inc.
- Darcy, M. & Tracey, T. J. G. (2004). Individual differences using paired comparisons and multidimensional scaling: Applications to multicultural counseling competence. *Journal of Counseling Psychology*, 51 (2), 139-150.
- Demirel, Ö. (2006). *Eğitimde program geliştirme*. Ankara: Pegem Akademi.
- Dunn-Runkin, P. (1982). *Scaling methods*. New Jersey: Lawrence Erlbaum Assoc. Publishers.
- Groenen, P. J. F. & Velden, M. V. D. (2004). *Multidimensional scaling*. Econometric Institute Report EI 2004-15. <http://repub.eur.nl/res/pub/1274/ei200415.pdf>. İndirme Tarihi: 12.10.2012.
- Guilford, J. P. (1954). *Psychometrics methods*. New York: Mc Graw-Hill Book Co.
- Kan, A.(2008). Yargıcı kararlarına dayalı ölçekleme yöntemlerinin karşılaştırılması üzerine ampirik bir çalışma. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 35, 186-194.
- Martignon, L. & Hoffrage, U. (2002). Fast, frugal, and fit: Simple heuristics for paired comparison. *Theory and Decision*, 52 (1), 29-71.
- Nishisato, S. (1994). *Elements of dual scaling: An introduction to practical data analysis*. Hillsdale, NJ: Erlbaum.
- Nunnally, J. C. (1970). *Introduction to psychological measurement*. New York: Mcgraw-Hill.
- O'Neil, M. R. & Chissom, B. S. (1993). A comparison three methods for assessing attitudes. *Annual Meeting of the Educational Research Association*. 10-12 November, 1993. New Orleans, USA.
- Öğretmen, T. (2008). Alan tercih envanteri: Ölçeklenmesi, geçerliği ve güvenilirliği. *Türk Eğitim Bilimleri Dergisi*, 6 (3), 507-522.
- 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*. Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi. 8-10 Eylül 2011. Burdur.
- Stevens, S. S. (1966). *Handbook of experimental psychology*. New York: John Willey and Sons.
- Steyvers, M. (2001). Multidimensional scaling. *Encyclopedia of cognitive science*. (Edt: L. Nadel). London: Macmillan Reference. pp. 1-5.
- Şimşek, D. (2006). Kümeleme analizi, çok boyutlu ölçekleme, doğrulayıcı ve açıklayıcı faktör analizi ile elde edilen yapı geçerliği kanıtlarının karşılaştırılması. *Yayınlanmamış Yüksek Lisans Tezi*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Titchener, E. B. (1901). *Experimental psychology: a manual of laboratory practice. Volume 1: Qualitative Experiments*. New York: Macmillan.

- Torgerson, W. S. (1958). *Theory and methods of scaling*. New York: John Wiley & Sons.
- Turgut, M. F. & Baykul, Y. (1992). *Ölçekleme teknikleri*. Ankara: ÖSYM Yayınları.
- Yiğit, E. (2007). Çok boyutlu ölçekleme yöntemlerinin incelenmesi ve bir uygulama. *Yayımlanmamış Yüksek Lisans Tezi*. Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, Samsun.

A Comparison of Two Scaling Methods: Pair Wise Comparison and Rank-Order Judgments Scaling

Meltem ACAR GÜVENDİR³, Yeşim ÖZER ÖZKAN⁴

Introduction

Scaling was initially used in psychophysics, which is a subfield of psychology. Scaling can be considered as a process of developing a scaling tool and ranking the collected data at equal intervals according to the decided scale. The scaling operation is conducted according to the judgment and reaction rule. The judgment rule is a stimulus—oriented approach that takes an expert's opinions into consideration. For the reaction rule, the scaling operation is done according to the reactions of the subjects.

In the stimulus—oriented rule approach, stimulants are scaled by an observer's or expert's opinions. The paired comparison method, used within the judgment approach, was developed by Louis Leon Thurstone, and it has been used in psychology research (especially cognitive psychology) for 10 years. Simple paired comparison is a simple method. Within the paired comparison approach, the stimulant sets or items are decided by the researcher. Observers are provided with all of the possible pairs and required to compare these pairs according to a better particular criterion. An advantage of this approach is its inclusion of both simple and comparative judgments.

When applying the comparative judgment rule, five situations take place. Of these situations, the most preferred ones are the third and fifth, which are simplified situations of the comparative judgment rule. While in the fifth situation, equation discrimination judgments are accepted as equal, in the third situation, equation discrimination judgments are not accepted as equal, and the variances are found one by one, as well. The fifth situation's method is considered economical since it accepts the variances as equal (normally the difference between the variances are large).

Scaling operation based on rank order judgments can be applied to all stimuli that can provide a rank number. In this method, observers/judges are provided all of the stimuli. Judges are required to compare each stimulus with others to ascertain a rank number. Thus, all of the stimuli are converted into a unified standard against which they are compared. Values obtained from the ranking judgments of the stimuli are compared with that unified standard and scale values are acquired. This method can be applied to all stimuli that can be assigned a rank number.

Although, studies have been conducted on multidimensional scaling, in both Turkey and abroad, literature on scaling techniques is limited. Thus, this study focuses on paired comparison and rank-order judgments scaling techniques and examines the consistency between results gathered by using them. The purpose of this research is to determine the

³ PhD – Res. Asst. - Trakya University, Faculty of Education - meltemacar@gmail.com

⁴ PhD - Instructor - Kilis 7 Aralık University, Faculty of Education - yozzer80@gmail.com

relationship between scaling values obtained from ranking judgment and pairwise comparison.

Methods

This descriptive study determines the consistency between the results obtained from paired comparison and rank-order judgments scaling techniques. The research group comprised 395 students in their fourth year of study in the education programs at Trakya University and Kilis 7 Aralık University during the 2011-2012 fall semester. A Teaching Methods assessment form was used as the data collection instrument. The scaling procedure was conducted to obtain scaling values for pairwise comparison and ranking judgment. In order to determine the relationship between the scaling values obtained from these two different scaling procedures, the Spearman Brown rho correlation coefficient was calculated. A teaching methods evaluation form, -which includes six teaching methods, -was given to the participants. They were requested rank (from one to six,) the teaching methods to be used by their instructor if they had to take the measurement and evaluation course again. The teaching methods mentioned in the teaching methods evaluation form were 'the lecture method', 'the discussion method', 'the case method', 'the demonstration-performance method', 'the problem solving method', and 'the independent study method'. In order to enable scaling operation based on paired comparison, students were given the same form but in a different format in which paired comparison of teaching methods was possible.

Data gathered from the teaching methods evaluation form were scaled using rank-order judgments and the complete data matrix from Thurstone's law of comparative judgment case V. equation. Primarily, the frequency equation values were ascertained from students' pair-wise comparisons of the six teaching methods, and the frequency matrix was formed using these values. The proportion matrix was found by stemming from the frequency matrix. The z values that corresponded to ratio matrix elements were identified and the formation of the unit normal variance matrix was put into effect. Values that belonged to each column were added to the bottom line of the unit normal variance matrix and the mean of each z value in the column was calculated to find the scale values. In order to move the starting point of the axis to the lowest z mean values, the scale values were sequenced by adding the lowest z mean value's absolute value to all values.

In order to gather data based on rank-order judgments, the students were given the list of six teaching methods and required to rank them, considering their priority for the measurement and evaluation course. The first operation of data analysis was to form the rank frequency matrix. The proportion matrix was formed from the rank-order judgments. The z values that corresponded to the ratio matrix elements were identified and the formation of the unit normal variance matrix was put into effect. Values that belong to each column were added to the bottom line of the unit normal variance matrix, and the mean of each z value in this column was calculated to find the scale values. In addition, the internal consistency reliability of scale values was calculated. Following this operation, χ^2 was calculated to find the significance of the internal consistency coefficient. The Spearman-Brown rank differences correlation coefficient was used to calculate the consistency between the gathered scale values.

Findings

The paired comparison results revealed that students preferred 'the case method' for the measurement and evaluation course, followed by 'the demonstration-performance method', 'the problem solving method', 'the discussion method', 'the independent study method', and 'the lecture method'.

Both scaling techniques found that 'the case method' received the highest priority. Moreover, 'the case method' and 'the discussion method' were at the same rank, and the other methods were at similar ranks, as well. The Spearman-Brown rank differences correlation coefficient operation was conducted on both data collections. Using both scaling techniques displayed the positive and significant relationship between the scale values. This finding shows that both scaling techniques produce similar outcomes.

Conclusion and Recommendations

There are many unobservable variables in education. The purpose of this study was to compare two scaling techniques that are used for developing scaling stimuli that measures educational variables. Results revealed that the correlation coefficient between these two sets of scaling values was 0,89. This finding shows that paired comparison and rank-order judgments scaling operations produce similar results. Researcher interested in related research areas should bear this similarity in mind while developing a scale based on expert judgments.

Key Words: Scaling, Paired comparison, Rank-order judgments, Teaching methods

Atf için / Please cite as:

Acar Güvendir, M. & Özer Özkan, Y. (2013). İki ölçekleme yönteminin karşılaştırılması: İkili karşılaştırma ve sıralama yargıları [A comparison of two scaling methods: Pair wise comparison and rank-order judgments scaling]. *Eğitim Bilimleri Araştırmaları Dergisi - Journal of Educational Sciences Research*, 3 (1), 105–119. <http://ebad-jesr.com/>