



## Öğretmen Adaylarının Test Yönergesinde Aradıkları Bilgilerin Sıralama Yargıları Yöntemi ile Ölçeklenmesi<sup>1</sup>

MAKALE TÜRÜ	Başvuru Tarihi	Kabul Tarihi	Yayın Tarihi
Araştırma Makalesi	05.10.2018	30.04.2019	01.05.2019

**D. Bahar Şahin Sarkın** <sup>2</sup>  
İstanbul Okan Üniversitesi

**Fazilet Taşdemir** <sup>3</sup>  
Recep Tayyip Erdoğan Üniversitesi

### Öz

Bu çalışmada, eğitim fakültesinde öğrenim görmekte olan öğretmen adaylarının test yönergesindeki bilgilerden hangisini daha önemli gördüklerinin sıralama yargıları yöntemi ile ölçeklenmesi amaçlanmıştır. Araştırma tarama modelinde yürütülmüştür. Ölçeklendirme sürecinde araştırmacılarca geliştirilen Test Yönergesi Bilgilerini Değerlendirme Formu kullanılmıştır. Test yönergesinde sınav süresinin belirtilmesi ve sorulara uygun olarak ayarlanması, öğrencilerin yanıtlama davranışlarını etkilemektedir. Bu çalışma, bir devlet üniversitesinin farklı bölümlerinde [Sınıf Öğretmenliği (SÖ), Sosyal Bilgiler Öğretmenliği (SBÖ), Fen ve Teknoloji Öğretmenliği (FTÖ), Türkçe Öğretmenliği (TÖ)] eğitim gören üçüncü ve dördüncü sınıf düzeyindeki 252 öğretmen adayı ile yürütülmüştür. Öğretmen adayları test yönergesinde bulunan bilgileri önem sırasına göre en çoktan en aza doğru sıraladıklarında, sınav süresi hakkındaki bilginin ilk sırada yer aldığı görülmektedir. Bu bilgiyi puanlama, şans başarısına yönelik önlem alınıp alınmayacağı, soru sayısı ile ilgili bilgiler izlemektedir. Cinsiyete göre öğrencilerin test yönergelerinde bulunan bilgilerden en önemli gördükleri bilgilerin sıralaması incelendiğinde kadın ve erkek öğrenciler için de puanlama, sınav süresi, şans başarısına yönelik önlem alınıp alınmayacağı bilgileri daha çok önemsedikleri görülmektedir. Sınıf düzeyine göre bilgilerin sıralaması incelendiğinde üçüncü ve dördüncü sınıf öğrencileri için de puanlama, sınav süresi, şans başarısına yönelik önlem alınıp alınmayacağı bilgilerinin ilk sıralarda yer aldığı sonucuna ulaşılmıştır. Sonuç olarak, eğitimciler için test yönergelerinin hazırlanması ile ilgili seminerler düzenlenebilir, sınavların nasıl hazırlanacağı konusundaki eksikleri giderilebilir ve sorunları çözülebilir.

**Anahtar sözcükler:** Ölçekleme, sıralama yargıları, yönerge, test yönergesi, öğretmen adayı.

<sup>1</sup>Bu çalışma 11-14 Mayıs 2017 tarihleri arasında gerçekleşen IV. International Eurasian Educational Research Congress'de (EJER) sözlü bildiri olarak sunulmuştur.

<sup>2</sup>*Sorumlu Yazar:* Dr. Öğr. Üyesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Rehberlik ve Psikolojik Danışmanlık Anabilim Dalı, E-posta: bahar.sarkin@okan.edu.tr, <https://orcid.org/0000-0002-1155-6114>.

<sup>3</sup>Dr. Öğr. Üyesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Rehberlik ve Psikolojik Danışmanlık Anabilim Dalı, E-posta: fazilet.tasdemir@erdogan.edu.tr, <https://orcid.org/0000-0002-0430-9094>.

Ölçme aracının nitelikli olması; güvenilir, geçerli ve kullanışlı olmasını anlatır. Nitelikli ölçme araçları yardımıyla elde edilen ölçme sonuçlarına dayalı olarak yapılan değerlendirmeler de aynı oranda nitelikli olur (Crocker ve Algina, 1986). Ölçme aracının kullanışlılığı, ekonomiklik, uygulanabilirlik ve puanlanabilirlik başlıkları altında toplanmaktadır (Tekin, 2010). Ölçme aracının emek, zaman ve maliyet bakımından uygun düzeyde olması ekonomikliktir. Uygulanabilirlik, ölçme aracının ölçülen birey ve uygulayan kişi açısından uygulama kolaylığına sahip olmasıdır. Test yönergesinin açık ve anlaşılır olması, testin uygulanan yaş grubuna göre yazı puntosu, satır aralıkları, sayfa düzeni, yazı biçimi gibi ölçme aracının özelliklerine dikkat edilmesi, uygulanabilirliği sağlama açısından önemlidir. Puanlanabilirlik ise, yanıt kağıdının kolay algılanabilir olması, yanıtlayan için kolay, güvenilir, anlaşılır ve doğru hazırlanmış olmasını belirtmektedir (Atılğan, 2006). Bireylerin başarı düzeylerinin değerlendirilmesinde farklı ölçme ve değerlendirme araç ve yöntemleri kullanılmaktadır. Test puanlarının en önemli niteliklerinden geçerliği ve güvenilirliği zedeleyen şans başarısı, tahmin davranışı gibi puanlamaya etki edebilecek etmenlere karşı ne yapılabileceği test uygulayıcılarının sorunlarından biri olmuştur. Test yönergesi, sorulara geçilmeden önce sınav hakkında bilgilerin sunulduğu bölümü oluşturmaktadır. Sınavın cevaplama süresinin ne kadar olduğu, hangi amaçla yapıldığı, puanlamasının nasıl yapılacağı, sınav sırasında öğrencilerin dikkat edeceği kuralların neler olduğu konusunda bilgilerin sunulduğu bölümdür (Özçelik, 1989). Her ölçme aracının hangi amaçla uygulandığını açıklayan ve nasıl tepki verileceğini içeren bir yönergesi olmalıdır. Yönerge, cevaplayıcıya güven vermekle birlikte yanlış anlamlardan dolayı verilebilecek hatalı tepkileri önler. Alanyazında test yönergesinde bulunan açıklamaların örneğin düzeltme formülünün kullanılıp kullanılmayacağına belirtilmesinin bireyin yanıtlama davranışında farklılıklara neden olduğu görülmektedir. Ebel'e (1965) göre, düzeltme formülünün psikolojik yönü daha ağır basmaktadır. Eğer test alan bireyler düzeltme formülünün uygulanacağını bilirlerse, yanıtları daha bilinçli olmakta ve bu işlem test puanlarının psikometrik niteliklerini arttırmaktadır. Cronbach (1960), yanıtlama davranışlarının kişilik özelliği ile ilgili bir durum olduğunu ve cevaplarından kuşku duyan bireylerin yönergelerde açıklanan düzeltme formülünün uygulanma durumundan etkilenmediğini belirtmiştir.

Tutum, ilgi, başarı, yetenek, kişilik gibi psikolojik nitelikleri ortaya çıkaracak ölçme araçlarının geliştirilmesi, ölçeklenmesi psikometri bilim alanında önemli adımlardandır (Torgerson, 1958). Ölçme işlemi sonunda elde edilen ölçümlerin anlamlı nitelikler kazanması amacıyla yapılan işlemler bütünü, ölçekleme süreci olarak bilinmektedir. Ölçekleme, bir ölçme aracının derecelendirilmesinden farklı olarak gözlem yoluyla bilinen ilişkilerden, kurallara dayalı ilişkilere geçişin yöntemlerini belirlemeyi amaçlar (Brown ve Peterson, 2009; Guilford, 1954; Turgut ve Baykul, 1992). Ölçekleme yaparken tepki ve yargı olarak sınıflandırılan deneysel yaklaşımlara başvurulmaktadır (Turgut ve Baykul, 1992). Tepki yaklaşımında uyarıcılara tepki veren kişiler yansız olarak değil, kendi tepkilerini belirterek uyarıcıların ölçekleme boyutundaki yerini belirlemektedirler. Yargı yaklaşımında

ise bilirkişi veya uzmanın her bir uyarıcının ölçek değerini diğer uyarıcılara göre belirlemesi beklenmektedir. Bilirkişilerin yansız yargılarının ortalama değeri hesaplanarak uyarıcının son ölçek değeri elde edilir. Yargı yaklaşımında uyarıcıların uyarıcılık derecesi, sınıflama, sıralama, ikili karşılaştırma yargılarıyla ölçekleme yöntemlerinden biri tercih edilerek belirlenmektedir (Turgut ve Baykul, 1992). Uyarıcıların belirli nesneye yönelik tutum, görüş, gereklilik, inanış gibi nitelikler kazandırmak amacıyla ölçme aracının belli işlemlere tabi tutulmasını gerektiren yöntem sıralama yargıları kanununa dayanmaktadır. İşlem sürecinde objenin özelliklerinin büyüklüğüne göre sıralanarak her özelliğe sıra numarası verilmesi ve konumunun oluşturulması önemlidir. Ölçekleme konusu ile ilgili yapılan çalışmalar incelendiğinde farklı konularda pek çok araştırmanın yapıldığı (Arık ve Kutlu, 2013; Brown ve Peterson, 2009; Bülbül ve Acar, 2012; Groenen ve Velden, 2004; Güler ve Anıl, 2009; Kara ve Gelbal, 2013; Özer ve Acar, 2011; Öztürk, Özdemir ve Gelbal, 2011) görülmüştür. Ölçekleme çalışmalarının daha çok ikili karşılaştırmalar yöntemiyle yapıldığı belirlenmiştir. Sıralama yargıları kanuna dayalı araştırılan çalışmalara (Bal, 2011; Kan, 2008; Özdil ve Kınay, 2015; Şahin, Boztunç Öztürk ve Taşdelen Teker, 2015; Yalçın ve Şengül Avşar, 2014) ek olarak bu araştırmanın alana katkı sağlayacağı düşünülmektedir. Alanyazında test yönergesinin içeriğinin test alan bireyler tarafından ne derece önemli olduğunu belirlemeye yönelik doğrudan bir araştırma bulunmamasıyla birlikte yönerge bilgilerinin yanıtlama davranışlarında farklılığa neden olduğunu vurgulayan araştırmalar vardır (Çoban, 2015; Koçak, 2013; Özdemir-Tokat 2006). Bu bağlamda bu araştırma; öğretmen adaylarının yanıtlama davranışlarını etkilediği düşünülen etkenlerin önem derecesinin ve bu etkenlerin öğretmen adayları tarafından algılanan büyüklüğü ile gerçek büyüklüğü arasındaki ilişkinin belirlenmesi ile var olan diğer ölçekleme çalışmalarından farklılaşmaktadır.

Bu araştırma, test yönergesinin içeriğinin test alan bireyler tarafından ne derece önemli olduğunu belirlemek açısından gerekli görülmüştür. Ayrıca araştırmanın bir ölçekleme çalışması olması nedeniyle öğrencilerin yanıtlama davranışlarında neleri önemsediklerinin sıralanması bağlamında da geleceğin öğretmeni olacak ve test yönergesi hazırlayacak öğretmen adaylarına da katkıda bulunacağı düşünülmektedir. Bu bağlamda yapılan araştırma ile farklı branşlardaki öğretmen adaylarının test yönergesinde hangi bilgileri aradıklarına bilgilerin önem sıralarının nasıl olduğuna ve farklı demografik niteliklere (cinsiyet ve sınıf) göre önem sıralarının farklılık gösterip göstermeyeceğini belirlemek amaçlanmıştır. Genel amaca göre araştırma kapsamında şu sorulara cevap aranmıştır:

1. Öğretmen adaylarının test yönergesinde aradıkları bilgilerin ölçek değerlerine göre sıralaması nasıldır?
2. Yönerge bilgilerinin sıralaması cinsiyete göre değişmekte midir?
3. Yönerge bilgilerinin sıralaması sınıf düzeyine göre değişmekte midir?

## Yöntem

Araştırmanın modeli, çalışma grubu, veri toplama aracı, verilerin analizi ve bulgulara bu başlık altında verilmiştir.

### Araştırma Modeli

Bu araştırmada eğitim fakültesinde öğrenim görmekte olan öğretmen adaylarının test yönergesindeki bilgileri önem derecelerine göre sıralama yargıları yöntemi ile sıra numarası verilmesi, bir duruma müdahale edilmeden olduğu gibi betimlendiğinden tarama modelindedir (Karasar, 2007).

### Çalışma Grubu

Araştırma 2016-2017 öğretim yılında bir devlet üniversitesinin Eğitim Fakültesi'nin farklı bölümlerinde [Sınıf Öğretmenliği (SÖ), Sosyal Bilgiler Öğretmenliği (SBÖ), Fen ve Teknoloji Öğretmenliği (FTÖ), Türkçe Öğretmenliği (TÖ)] eğitim gören üçüncü ve dördüncü sınıf düzeyindeki toplam 252 öğrenci ile yürütülmüştür. Kadınların erkeklere oranla test yönergesini daha fazla dikkate aldığı ve yanıtlama davranışında farklılıklar olduğu sonucuna ulaşan araştırmalar (Çoban, 2015; Koçak, 2013) incelendiğinden bu çalışmada da cinsiyet bir değişken olarak kullanılmıştır. Araştırma grubunda yer alan öğretmen adaylarının tamamı 3. ve 4. sınıf olmaları nedeniyle öğrenim gördükleri programda "Ölçme ve Değerlendirme" dersini almıştır. Ölçme ve değerlendirme dersini almamış öğrencilerin yönergede bulunan bilgileri yeterince dikkatli okumadıkları, incelemedikleri ve farkındalıklarının yeterince oluşmadığı varsayıldığından araştırma grubuna dahil edilmemişlerdir. Çalışma grubuna yönelik betimsel istatistikler Tablo 1'de verilmiştir.

Tablo 1

#### Araştırma Grubuna Yönelik Frekans ve Yüzde Dağılımı

Sınıf Düzeyi	Bölüm	Cinsiyet				Toplam	
		Kadın		Erkek		f	%
		f	%	f	%		
3. sınıf	SBÖ	29	12	11	4	99	39
	TÖ	44	17	15	6		
4. sınıf	FTÖ	36	14	18	7	153	61
	SÖ	80	32	19	8		
Toplam		189	75	63	25	252	100

Tablo 1 incelendiğinde katılımcıların % 75'i kadın, % 25'i erkek olmak üzere çoğunluğunun kadınlardan oluştuğu görülmektedir. Farklı bölümlerde bulunan öğrencilerin % 39'u üçüncü sınıf, % 61'i dördüncü sınıftır.

### Veri Toplama Aracı

Veri toplama aracının geliştirilmesinde alanyazın taraması yapılmış, araştırma grubunun bulunduğu eğitim fakültesinde görev yapan 20 farklı öğretim üyesinin hazırladığı 20 sınav kağıdı incelenmiştir. Test geliştirme konusunda çalışmaları bulunan iki ölçme ve değerlendirme uzmanından test yönergesinde bulunması gereken bilgiler ile ilgiligörüş alınmıştır. Uzmanlardan alınan bilgiler doğrultusunda sınav kağıtlarında en çok tekrar eden bilgiler belirli başlıklar altında sınırlandırılmıştır. Yapılan alanyazın taraması ve alınan uzman görüşleri doğrultusunda ölçekleme işlemi için araştırmacılar tarafından oluşturulan “Test Yönergesi Bilgilerini Değerlendirme Formu” oluşturulmuştur. Bu formda öğrencilerin bir test yönergesinde aradıkları bilgilere ilişkin yargıların sıralanması için araştırmacılar tarafından belirlenmiş sekiz faktör yer almaktadır. Bu faktörler; sınav süresi (M1), soru sayısı (M2), puanlama (M3), soruların nasıl cevaplanacağı (M4), cevaplamanın nereye yapılacağı (M5), şans başarısına yönelik önlem alınıp alınmayacağı (M6), kopya durumunda ne yapılacağı (M7), sınav sırasında yasak davranışların neler olduğudur (M8). Öğretmen adaylarından uyarıcıları bir bütün olarak görmeleri ve her birini diğerleriyle kıyaslayarak sıra sayıları vermeleri talep edilmiştir. Test Yönergesi Bilgilerini Değerlendirme Formu cevaplandırılırken öğretmen adaylarından yönergedeki en önemli bilgiye 1 sıra sayısını vermeleri istenmiştir. Her bilginin önem sırasını belirten öğrenmen adayları 1-8 aralığında numaralandırma yapmışlardır. Ayrıca adaylarından her bir bilgiye sıra sayıları verilirken her sıra sayısının bir kez kullanılması konusunda dikkatli davranmaları istenmiştir.

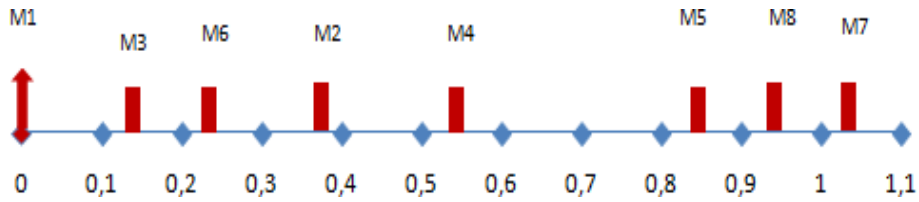
### Verilerin Analizi

Verilerin analizinde, ölçekleme yöntemlerinden yargıcı kararlarına ilişkin yaklaşımlardan sıralama yargıları kanundan yararlanılmıştır. Öğretmen adaylarına sınav yönergesinde aradıkları bilgilere yönelik yargıların elde edilmesi için sekiz uyarıcı verilmiştir. Veri toplama aracını oluşturan sekiz uyarıcının ölçek değerleri ve bağımsız değişkenlere göre sıralamaların değişip değişmediği tek tek Microsoft Excel programında analiz edilmiştir. Öğretmen adaylarından uyarıcıların hepsini birlikte düşünmesi ve her uyarıcıyı diğerleri ile kıyaslayarak numara vermeleri istenmiştir. Dolayısıyla uyarıcı kategorisi, tüm uyarıcıların kıyaslandığı bir bileşik ölçüme dönüştürülmüştür. Öncelikle uyarıcıların kaçınıcı sraya kaç defa getirildiğini belirten uyarıcılara ilişkin sıra frekansları matrisi ve ilgili matrise bağlı olarak oranlar matrisi oluşturulmuştur (Turgut ve Baykul, 1992). Oranlar matrisi kullanılarak birim standart normal sapmalar matrisi (zjk) hesaplanmış ve tek tek uyarıcılara dair ölçek değerleri (Sj) belirlenmiştir. En küçük ölçek değerinin mutlak değeri, bütün ölçek değerlerine eklenerek eksen başlangıcı (O noktası) en küçük ölçek değerine kaydırılmıştır. Böylece uyarıcıların ölçek değerleri (Sc) yeniden belirlenmiş ve sıralanmıştır. Bağımsız değişken kategorilerinin her biri için her bir uyarıcının ölçek değeri ayrı ayrı hesaplanmıştır.

Elde edilen ölçek değerleri öğretmen adaylarının yargılarına dayanmaktadır. Gözlemcilerin uyarıcılara tepkide bulunurken dikkatli davranıp davranmadıklarını kontrol etmek amacıyla ölçek değerlerinin iç tutarlılığı incelenmiştir. Turgut ve Baykul'a (1992) göre ölçek değerlerinin iç tutarlılığı; gözlenen frekans oranlarının ölçek değerlerinden hesaplanan frekans oranları ile ne derece benzer olduğunun hesaplanması ile belirlenmektedir. İç tutarlılığı hesaplamak için uyarıcıların ölçek değerlerinden yararlanılarak kuramsal birim sapmalar matrisi ve kuramsal oranlar matrisi hesaplanmıştır. Daha sonra da gözlenen oranlar ve teorik oranlar benzerliği incelenerek ölçek değerlerinin ortalama hatası elde edilmiştir. Ortalama hata değeri, uygunluk değerinin manidar olup olmadığı hakkında bilgi vermediği için uygunluk derecesinin manidarlığı Ki-kare testiyle incelenmiştir. Hesaplanan ki-kare bulgusu [ $\chi^2$ (sd=28, n=252)=182.427, p<.05] tablo değerini [ $\chi^2$ (sd=28, n=252)=41.337, p<.05] aştığından anlamlı bulunmuştur. Elde edilen Ki-kare değerinin tablo değerinden büyük olması iç tutarlığın düşük olduğunu göstermiştir. Bu bulgulara göre ölçeklemeye esas olan verilerin, V. halin varsayımlarını sağlamadığı sonucuna varılmıştır. Ayrıca ölçeklenen değişkenin tek boyutlu olmaması da Ki-kare değerinin anlamlı olmasına neden olduğu yorumu da yapılabilmektedir. V. hal denkleminde ayırıcılık yargıları varyansları eşit kabul edilirken III. hal denkleminde ayırıcılık yargıları varyansları ayrı ayrı hesaplanmakta yani eşit kabul edilmemektedir. Dolayısıyla varsayımları karşılandığı takdirde V. hal yöntemiyle ölçekleme yapmak işlemlerde ekonomiklik sağlamaktadır. Guilford (1954), uyum ölçüsünün anlamlı çıkması durumunda III. halin tercih edilmesini önermektedir (Akt. Turgut ve Baykul, 1992). Bu nedenle öğretmen adaylarının test yönergesinde aradıkları bilgilerin önem sırasını belirlemek için III. hal denklemleri ile ölçekleme yapılmıştır.

### Bulgular

Bu bölümde öğretmen adaylarının test yönergesinde aradıkları bilgilerin hangisini daha önemli gördüklerini belirlemek amacıyla yapılan ölçekleme çalışmasında sekiz uyarıcının ölçek değerleri hesaplanmıştır. Elde edilen en küçük ölçek değeri başlangıç noktası (0 noktası) olacak biçimde ötelenmiş ve hesaplanan değerlerin mutlak değeri büyüklüğünde diğer belirlenen ölçek değerlerinin büyüklüğü artırılmıştır. Uyarıcıların sonul (nihai) ölçek değerleri Şekil 1'de sayı doğrusunda verilmiştir.



Şekil 1. Uyarıcıların ölçek değerlerinin betimlenmesi

Şekil 1’de de belirtildiği gibi sınav süresi (M1); sıfır değerini taşıması nedeniyle öğrenciler tarafından en önemli uyarıcı olarak görülmektedir. Öğretmen adaylarının test yönergesinde aradıkları bilgilerin uyarıcı sıraları ve ölçek değerleri Tablo 2’de sunulmuştur.

Tablo 2

*Test Yönergesindeki Bilgilerin Ölçek Değerleri ve Uyarıcı Sıraları*

Test Yönergesinde Yer Alan Bilgiler	Ölçek Değerleri (Sj)	Uyarıcı Sıraları
M1. Sınav süresi	0.00	1
M2. Soru sayısı	0.39	4
M3. Puanlama	0.14	2
M4. Soruların nasıl cevaplanacağı	0.53	5
M5. Cevaplamanın nereye yapılacağı	0.85	6
M6. Şans başarısına yönelik önlem alınıp alınmayacağı	0.20	3
M7. Kopya durumunda ne yapılacağı	1.02	8
M8. Sınav sırasında yasak davranışların neler olduğu	0.91	7

Tablo 2’ye göre öğrenciler, test yönergelerinde bulunan bilgileri önem sırasına göre en çoktan en aza doğru sıraladığında, M1 maddesinde belirtilen sınav süresi ile ilgili bilginin ilk sırada yer aldığı görülmektedir. Bu bilgiyi puanlama (M3), şans başarısına yönelik önlem alınıp alınmayacağı (M6), soru sayısı (M2) ile ilgili bilgiler izlemektedir. Tablo 2 incelendiğinde öğrencilerin sınav sırasında yasak davranışların neler olduğu (M8) ve kopya durumunda ne yapılacağı (M7) bilgilerinin sınav yönergesinde daha az arandığı söylenebilir.

Öğretmen adaylarının test yönergesinde aradıkları bilgiler cinsiyet değişkenine ilişkin tek tek ölçeklendiğinde hesaplanan ölçek değerleri ve uyarıcı sıraları Tablo 3’te verilmiştir.

Tablo 3

*Cinsiyete Göre Öğrencilerin Test Yönergesinde En Önemli Gördükleri Bilgiler İçin Hesaplanan Ölçek Değerleri ve Uyarıcı Sıraları*

Test Yönergesinde Yer Alan Bilgiler	Kadın		Erkek	
	Ölçek Değerleri (Sj)	Uyarıcı Sıraları	Ölçek Değerleri (Sj)	Uyarıcı Sıraları
M1. Sınav süresi	0.00	1	0.16	2
M2. Soru sayısı	0.54	4	0.42	4
M3. Puanlama	0.21	2	0.00	1
M4. Soruların nasıl cevaplanacağı	0.73	6	0.90	6
M5. Cevaplamanın nereye yapılacağı	0.72	5	0.85	5

(devam ediyor)

Tablo 3 (devam)

Test Yönergesinde Yer Alan Bilgiler	Kadın		Erkek	
	Ölçek Değerleri (Sj)	Uyarıcı Sıraları	Ölçek Değerleri (Sj)	Uyarıcı Sıraları
M6. Şans başarısına yönelik önlem alınıp alınmayacağı	0.41	3	0.26	3
M7. Kopya durumunda ne yapılacağı	1.01	8	1.02	8
M8. Sınav sırasında yasak davranışların neler olduğu	0.96	7	1.00	7

Tablo 3'e göre kadın ve erkek öğrenciler, test yönergelerinde bulunan bilgileri önem sırasına göre en çoktan en aza doğru sıraladığında ilk iki sıradaki bilginin değişmesi dışında farklılık yoktur. Kadın öğrenciler için M1 maddesinde belirtilen sınav süresi ile ilgili bilginin ilk sırada yer aldığı görülmektedir. Erkek öğrenciler için ise M3 maddesinde belirtilen puanlama ile ilgili bilginin ilk sırada yer aldığı görülmektedir. Bu bilgileri tüm grup için elde edilen sıralamada olduğu gibi cinsiyet değişkeni için şans başarısına yönelik önlem alınıp alınmayacağı (M6), soru sayısı (M2) ile ilgili bilgiler izlenmektedir. Tablo 3'e göre kadın ve erkek öğrenciler sınav sırasında yasak davranışların neler olduğu (M8) ve kopya durumunda ne yapılacağı (M7) yönerge bilgilerini diğerlerine göre daha az önemli bulmuşlardır.

Öğrencilerin test yönergesinde aradıkları bilgiler sınıf düzeyi değişkenine göre 3. sınıf ve 4. sınıf öğrencileri için ayrı ayrı incelenmiş ve hesaplanan ölçek değerleri Tablo 4'te verilmiştir.

Tablo 4

*Sınıf Düzeyine Göre Öğrencilerin Test Yönergesinde En Önemli Gördükleri Bilgiler İçin Hesaplanan Ölçek Değerleri ve Uyarıcı Sıraları*

Test Yönergesinde Yer Alan Bilgiler	3. Sınıflar		4. Sınıflar	
	Ölçek Değerleri (Sj)	Uyarıcı Sıraları	Ölçek Değerleri (Sj)	Uyarıcı Sıraları
M1. Sınav süresi	0.00	1	0.33	3
M2. Soru sayısı	0.47	5	0.44	4
M3. Puanlama	0.11	2	0.00	1
M4. Soruların nasıl cevaplanacağı	0.34	4	0.86	6
M5. Cevaplamanın nereye yapılacağı	0.80	6	0.72	5
M6. Şans başarısına yönelik önlem alınıp alınmayacağı	0.20	3	0.19	2
M7. Kopya durumunda ne yapılacağı	1.08	8	1.02	8
M8. Sınav sırasında yasak davranışların neler olduğu	0.97	7	0.98	7

Tablo 4'e göre 3. sınıf öğrencileri, test yönergelerinde bulunan bilgileri önem sırasına göre en çoktan en aza doğru sıraladığında, M1 maddesinde belirtilen sınav



süresi ile ilgili bilginin ilk sırada yer aldığı görülmektedir. Bu bilgiyi puanlama (M3), şans başarısına yönelik önlem alınıp alınmayacağı (M6), soruların nasıl cevaplanacağı (M4) ile ilgili bilgiler izlemektedir. Tablo 4'e göre 3.sınıf öğrenciler, sınav sırasında yasak davranışların neler olduğu (M8) ve Kopya durumunda ne yapılacağı (M7) bilgilerini diğerlerine göre daha az önemli bulmuşlardır.

Tablo 4'e göre 4. sınıf öğrencileri, test yönergelerinde bulunan bilgileri önem sırasına göre en çoktan en aza doğru sıraladığında, M3 maddesinde belirtilen puanlama hakkındaki bilginin ilk sırada yer aldığı görülmektedir. Bu bilgiyi şans başarısına yönelik önlem alınıp alınmayacağı (M6), sınav süresi (M1), soru sayısı (M2) ile ilgili bilgiler izlemektedir. Tablo 4'e göre 4. sınıf öğrenciler de 3. sınıf öğrencileri gibi, sınav sırasında yasak davranışların neler olduğu (M8) ve kopya durumunda ne yapılacağı (M7) bilgilerini diğerlerinden düşük düzeyde önemli bulmuşlardır.

### **Tartışma, Sonuç ve Öneriler**

Bu araştırmada öğretmen adaylarının test yönergesinde aradıkları bilgilerin önem sırası sıralama yargılarına dayalı olarak belirlenmiştir. Ayrıca öğretmen adaylarının sıralama yargılarına dayalı olarak belirlenen ölçek değerlerinin cinsiyete ve sınıf düzeyine göre farklılaşıp farklılaşmadığı araştırılmıştır. Araştırma sonucunda sekiz faktörden en önemli bilginin sınav süresi olduğu, bu bilgiyi sırasıyla puanlama, şans başarısına yönelik önlem alınıp alınmayacağı, soru sayısı, cevaplamanın nereye yapılacağı, soruların nasıl cevaplanacağı, sınav sırasında yasak davranışların neler olduğu, kopya durumunda ne yapılacağı bilgilerinin izlediği bulunmuştur.

Çoktan seçmeli testlerde, öğrencinin zamanının büyük bir kısmı, soruların okunarak neyin sorulduğunun anlaşılmasına, sorunun yanıtının düşünülmesine ve bulunan yanıtın hangi seçenekte olduğunun bir işaretle belirtilmesine harcanmaktadır (Özçelik, 1989). Test yanıtlama süresi yanıtlayıcılar için önemli değişkenlerden biridir. Test yönergesinde sınav süresinin belirtilmesi ve sorulara uygun olarak ayarlanması öğrencilerin yanıtlama davranışlarını etkilemektedir. Frary'e (1980) göre çoktan seçmeli bir maddenin yanıtlama sürecinde, yanıtlayıcının sahip olduğu bilgiyi, seçeneklerle ilişkilendirmesi ve çeldiricilerden birinin doğru olduğuna inanması yaşanır. Bununla birlikte yanıtlayıcı doğru seçeneği tek ya da daha çok çeldiriciyle ilişkilendirebilir, geri kalan iki ya da daha çok çeldirici içerisinden kestirimde bulunabilir. Bu olasılıklar yanıtlayıcının düşüncesindeki değişkenlikleri, başka bir anlatımla tam olmayan bilginin değişik düzeylerdeki varlığını ortaya koyar. Ayrıca başarı testlerinde yeterlitest süresinin verilmemesi durumunda, yanıtlayıcı bazı maddeleri okumadan, rastgele yanıtlandırma yoluna gidebilmektedir (Turgut, 1971).

Test yönergesinde her bir sorunun puan değerinin ne kadar olduğunun ve düzeltme formülünün kullanılıp kullanılmayacağını açıklanması, yanıtlayıcıların aradıkları önemli bilgilerdendir. Yönergede yer alan puanlama bilgisinin öğrencilerin davranışlarında belirleyici olduğunu gösteren çalışmalar vardır (Çoban, 2015; Koçak, 2013; Özdemir-Tokat 2006). Dolayısıyla öğretmenler ve test geliştiricilerin, test

geliştirme sürecinde yanıtlayıcı davranışlarının neler olduğunu, hangi durumda ne tür yanıtlayıcı davranışları gösterildiğini göz önünde bulundurması gerekmektedir.

Test yönergesinde şans başarısına yönelik önlem alınıp alınmayacağı bilgisi, yanıtlayıcı davranışlarında farklılıklara neden olmaktadır (Çelen ve Demirtaşlı 2006; Espinosa ve Gardeazabal, 2010; Kadioğlu 2002; Umay, 1998). Çoktan seçmeli test maddelerini yanıtlayıcı hakkında deneyimli olan kişi, maddeleri yanıtlarken her bilgi parçasını kullanarak kestirimde bulunma eğilimindedir. Deneyimsiz ya da şansını denemekten kaçınan birey ise, düzeltme puanı uygulamasından etkilenir ve doğru yanıtlayıcı olasılığı (ihtimali) yüksek olan oldukça çok maddeyi de yanıtlayıcı bırakır. Ebel (1965), düzeltilmiş puan uygulamasının, test maddelerini şansla doğru yanıtlayıcı konusunda deneyimli olan bireylere özel bir kazanç sağladığını belirtmektedir. Koçak (2013) farklı yönergelerle verilen testlerde yanıtlayıcı davranışlarını incelemiş ve *Doğru yanıtı bulamadığım için rastgele işaretledim* davranışının, düzeltme formülü uygulanmayacağı bilgisi verildiğinde en çok gösterilen davranışlar arasında olduğu ancak herhangi bir uyarı yapılmadığı ya da düzeltme formülü uygulanacağı uyarısı yapıldığında bu davranıştan kaçınıldığı sonucuna ulaşmıştır.

Cinsiyet ve sınıf düzeyi değişkenleri için yapılan incelemede, test yönergesinde aranan bilgilerin önem sıralarının tüm grup için elde edilen sıralamaya benzer olduğu sonucuna ulaşılmıştır. Cinsiyete göre öğrencilerin test yönergelerinde bulunan bilgilerinden en önemli gördükleri bilgilerin sıralaması incelendiğinde kadın ve erkek öğrenciler için de puanlama, sınav süresi, şans başarısına yönelik önlem alınıp alınmayacağı bilgilerinin ilk sıralarda yer aldığı ve öğrencilerin bu bilgileri diğer bilgilerden daha çok önemsedikleri görülmektedir. Test yönergesinde bulunan kopya durumunda ne yapılacağı ve sınav sırasında yasak davranışların neler olduğu bilgilerinin cinsiyet değişkenine göre daha az önemsenenler arasında olduğu görülmüştür. Sınıf düzeyine göre öğrencilerin test yönergelerinde bulunan bilgilerinden en önemli gördükleri bilgilerin sıralaması incelendiğinde üçüncü ve dördüncü sınıf öğrencileri için de puanlama, sınav süresi, şans başarısına yönelik önlem alınıp alınmayacağı bilgilerinin ilk sıralarda yer aldığı ve öğrencilerin bu bilgileri diğer bilgilerden daha çok önemsedikleri görülmektedir. Sınav yönergesinde bulunan kopya durumunda ne yapılacağı ve sınav sırasında yasak davranışların neler olduğu bilgilerinin sınıf düzeyi değişkenine göre en az önemsenenler arasında olduğu görülmüştür.

Bu araştırma sonucunda sınav kağıdı hazırlarken bazı test geliştiriciler tarafından önemsenmeyen yönergenin, öğrencilerin yargılarından yola çıkarak sıralama yargıları yöntemi ile ölçeklenmiş ve gerekliliği belirtilmiştir. Araştırmanın bulguları ve sonuçları, öğretmen adaylarının yaptıkları sıralamalar, sınav kağıdı hazırlayan öğretim elemanları ile paylaşılabilir. Test yönergesi ile ilgili seminerler düzenlenerek yönergenin nasıl hazırlanması gerektiği konusunda öğretim elemanlarının eksiklikleri giderilebilir, sorunları çözülebilir. Bu araştırmanın sonuçları doğrultusunda temsil gücü yüksek farklı fakülteler, üniversiteler ve farklı değişkenler de dahil edilerek benzer çalışmalar yapılabilir.

### Kaynakça

- Atılgan, H. (Ed.) (2006). *Eğitimde ölçme ve değerlendirme*. Ankara: Anı Yayıncılık.
- Arık, R. S. ve 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.
- 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.
- Brown, T. C., and Peterson, G. L. (2009). An enquiry into the method of paired comparison: reliability, scaling, and Thurstone's law of comparative judgment. Retrieved from [https://www.fs.fed.us/rm/pubs/rmrs\\_gtr216.pdf](https://www.fs.fed.us/rm/pubs/rmrs_gtr216.pdf)
- Bülbül, T., and 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.
- Büyüköztürk, Ş., Çakmak, E., Akgün, Ö. E., Karadeniz, Ş. ve Demirel, F. (2010). *Bilimsel araştırma yöntemleri*. (7. Baskı). Ankara: Pegem A Yayıncılık.
- Crocker, L., and Algina, J. (1986). *Introduction to classical and modern test theory*. New York, NY: Holt.Rinehart and Winston.
- Cronbach, L. J. (1960). *Essentials of psychological testing*. (2nd Edition). New York, NY: Harper and Brothers, Publishers.
- Çelen, Ü. ve Demirtaşlı, N. (2006). Düzeltme yönergesinin testin psikometrik özelliklerine etkisi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 30(30), 82-91.
- Çoban, E. (2015). *Test yönergesi ve risk alma eğiliminin seçmeli testlerin psikometrik özelliklerine ve tahminle yanıtlama davranışına etkisi* (Yayımlanmış yüksek lisans tezi, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara, Türkiye). <https://tez.yok.gov.tr/UlusalTezMerkezi/>nden elde edilmiştir. (Tez No. 396161)
- Ebel, R. L. (1965). *Measuring educational achievement*. New Jersey, NJ: Prentice-Hall.
- Espinosa, M. P., and Gardeazabal, J. (2010). Optimal correction for guessing in multiple-choice tests. *Journal of Mathematical Psychology*, 54(5), 415-425.
- Frary, R. (1980). The effect of misinformation, partial information and guessing on expected multiple-choice test item scores. *Applied Psychological Measurement*, 4(1), 79-90.

- Groenen, P. J. F., and Velden, M. V. D. (2004). Multidimensional scaling. Retrieved from <http://repub.eur.nl/res/pub/1274/ei200415.pdf>.
- Guilford, J. P. (1954). *Psychometrics methods*. New York, NY: Mc Graw-Hill Book Co.
- Güler, N., and 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.
- Kadioğlu, B. (2002). *Çoktan seçmeli testlerde test maddelerini cevaplama davranışlarının madde ve öğrenci özelliklerine göre nasıl bir değişim gösterdiğinin incelenmesi* (Yayımlanmamış yüksek lisans tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara, Türkiye). <https://tez.yok.gov.tr/UlusalTezMerkezi>'nden elde edilmiştir. (Tez No: 113426)
- Kan, A. (2008). Yargıcı kararlarına dayalı ölçekleme yöntemlerinin karşılaştırılması üzerine ampirik bir çalışma. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 35, 186-194.
- Kara, Y. ve Gelbal, S., (2013). İlköğretim öğrencilerinin başarılarını etkileyen özelliklerin tam sıralama halinde ikili karşılaştırmalar yöntemiyle ölçeklenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 4(1), 33-51.
- Karasar, N. (2008). *Bilimsel araştırma yöntemi* (18. Baskı). Ankara: Nobel Yayın Dağıtım.
- Koçak, D. (2013). *Farklı yönergelerle verilen çoktan seçmeli testlerde yanıtlama davranışlarının incelenmesi*. (Yayımlanmış yüksek lisans tezi, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara, Türkiye). <https://tez.yok.gov.tr/UlusalTezMerkezi>'nden elde edilmiştir. (Tez No: 342461)
- Özçelik, D. A. (1989). *Test hazırlama kılavuzu*. Ankara: ÖSYM Yayınları.
- Özer, Y. ve Acar, M. (2011). Öğretmenlik mesleği genel yeterlikleri üzerine ikili karşılaştırma yöntemiyle bir ölçekleme çalışması. *Çukurova Üniversitesi Eğitim Fakültesi Dergisi*, 3(40), 89-101.
- Özdemir Tokat, Y. N. (2006). *Çoktan seçmeli testlerde yanıtlama davranışlarının belirlenmesi* (Yayımlanmamış yüksek lisans tezi, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara, Türkiye). <https://tez.yok.gov.tr/UlusalTezMerkezi>'nden elde edilmiştir. (Tez No: 204611)
- Özdil Örs, S. ve Kınay, E. (2015). 5.Sınıf seçmeli ders tercihlerinin sıralama yargıları kanunuyla ölçeklenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(2), 268-278.
- Öztürk, N., Özdemir, S. ve Gelbal, S. (2011, Eylül). *İki farklı ölçekleme yaklaşımından elde edilen ölçek değerleri tutarlılığının incelenmesi*. 20. Ulusal

Eğitim Bilimleri Kurultayı'nda sunulan bildiri. Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi, Burdur.

Şahin, G. M., Öztürk Boztunç, N. ve Taşdelen Teker, G. (2015). Öğretmen adaylarının başarılarının değerlendirilmesinde tercih ettikleri ölçme araçlarının belirlenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(1), 95-106.

Tekin, H. (2010). *Eğitimde ölçme ve değerlendirme*. Ankara: Yargı Yayınları.

Torgerson, W. S. (1958). *Theory and methods of scaling*. New York, NY: John Wiley and SonsInc.

Turgut, F. (1971). *Şans başarısının test puanlarına etkisi*. Ankara: ODTÜ Yayınları.

Turgut, M. F. ve Baykul, Y. (1993). *Ölçekleme teknikleri*. Ankara: ÖSYM Yayınları.

Umay, A. (1998). Seçmeli derslerde yanıtlayıcı davranışları ve şans başarısının elimine edilmesi işlemlerine ilişkin bazı öneriler. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 14(14), 54-61.

Yalçın, S. ve Şengül Avşar, A. (2014). Eğitim fakültesi meslek bilgisi derslerinin sıralama yargıları kanunuyla ölçeklenmesi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 5(2), 79-90.



## Scaling the Information Sought by the Pre-service Teachers in the Test Directive Through Method of Sorting Judgments<sup>1</sup>

ARTICLE TYPE	Received Date	Accepted Date	Published Date
Research Article	10.05.2018	04.30.2019	05.01.2019

**D. Bahar Şahin Sarkın** <sup>2</sup>  
İstanbul Okan University

**Fazilet Taşdemir** <sup>3</sup>  
Recep Tayyip Erdoğan University

### Abstract

In this study, it is aimed to scale the information sought by the students studying at the faculty of education in the test directive through the method of sorting judgments. The study has been conducted in screening model. The Test Directive Information Evaluation Form developed by the researchers has been used for the scaling process. The research was carried out in different departments of the Faculty of Education of a public university [Primary School Teaching (PST), Social Studies Teaching (SST), Science and Technology Teaching (STT), Turkish Language Teaching (TLT)] with 252 students who are attending the third and fourth grade. Specifying the exam duration and adjusting it according to the questions in the test directive affect the response behaviors of the students. The information in the test directives have been sorted by the students from most important to the least important as duration of the exam, scoring, whether the measures are taken against success by chance and the number of questions. When sorting based on the gender examined, it is seen that the information of scoring, duration of exam and whether the measures are taken against success by chance for both female and male students rank near the top and they pay more attention to this information than others. According to the grade level, the information of scoring, duration of exam and whether the measures are taken against success by chance rank near the top for the 3rd and 4th grade students. In conclusion, seminars related to the test directive could be organized and the deficiencies of the instructors about how to prepare the tests could be eliminated and their problems could be solved.

**Keywords:** Scaling, rank-order judgments, directive, test instruction, pre-service teachers.

<sup>1</sup>This article was presented at the IV. International Eurasian Educational Research Congress (EJER) held on 11-14 May 2017.

<sup>2</sup>*Corresponding Author:* Assist. Prof. Dr., Faculty of Education, Department of Educational Sciences, Guidance and Psychological Counselling Department, E-mail: bahar.sarkin@okan.edu.tr, <https://orcid.org/0000-0002-1155-6114>.

<sup>3</sup>Assist. Prof. Dr., Faculty of Education, Department of Educational Sciences, Guidance and Psychological Counselling Department, E-mail: fazilet.tasdemir@erdogan.edu.tr, <https://orcid.org/0000-0002-0430-9094>.

**Purpose and Significance**

Keeping the test guideline and the examples clear and understandable, and giving attention to the characteristics of the measurement tools such as font size, line spacing, page layout, type of writing according to the age group of the test are important in terms of ensuring the applicability of the tests. It was one of the problems of the test instructors that what could be done against the factors which could affect the scoring such as chance success, estimation behavior which damage the validity and reliability which are the important characteristics of the test scores. Test guideline forms the section where information about the exam is presented before proceeding to the questions. It is seen in the literature that the explanations in the test guideline, for example, whether to use the correction formula or not, cause differences in the response behavior of the individual. It is important to determine the importance level of the factors which is thought to affect the response behaviors of pre-service teachers and the relationship between the magnitude of these factors perceived by the pre-service teachers and their actual magnitude. Scaling presents the methods of transition from observational relations to rule-based associations. Scaling with sorting judgments means that the stimulants sort a specific characteristic from large to small or vice versa and that a sequence number is attached to each characteristic according to its position. The study is important to determine the extent to which the content of the directives that are considered unimportant by academics teaching at universities is required by the individuals taking the test. In addition, as the research is a scaling study, it is thought that ranking on what the students give importance in terms of responding behaviors will also contribute to the literature. In this study, it is aimed to scale the information in the test directive sought by the pre-service teachers studying in the faculty of education through the method of sorting judgments.

**Method**

The study has been conducted in screening model. The research was carried out in different departments of the Faculty of Education of a public university in the academic year of 2016-2017 [Primary School Teaching (PST), Social Studies Teaching (SST), Science and Technology Teaching (STT), Turkish Language Teaching (TLT)] with 252 students who are attending the third and fourth grade. The Test Directive Information Evaluation Form developed by the researchers has been used for the scaling process. In the analysis of the data, the method of scaling through sorting judgments has been used. In this form, there are eight factors determined by the researchers in order to sort the opinions related to the information which the students would like to see in a test guideline. These factors are; time, the number of questions, scoring, how to answer the questions, where to address the answers, whether to take measures for the success by chance, what to do in case of copy and what are the prohibited behaviors during the exam. The pre-service teachers were asked to think each stimulus as a whole and give a sequence number for each stimulant by comparing each of them with the others. While the Evaluation Form of Test Guideline Information was answered, the pre-service teachers were asked to rank the information in the guideline which they think the most important, by giving them the

sequence number of 1. The pre-serviceteachers made a gradation between 1 - 8 which states the order of importance of each information. Obtained scaling values are based on the judgments of the pre-serviceteachers. The internal consistency of the scale values was examined in order to check whether the observers reacted carefully to the stimuli. In order to calculate the internal consistency, theoretical unit deviation matrix and theoretical ratios matrix were obtained by using the scale values of stimuli. In the next stage, the average error of the scale values was calculated by considering the appropriateness of the ratios observed with the theoretical ratios. As the average error value did not provide information about whether the fidelity is appropriate or not, it is examined by Chi-square test to determine if the appropriateness degree is significant. The calculated chi-square value [ $\chi^2$ (sd = 28, n = 252) = 182.427, p < .05] exceeds the table value [ $\chi^2$ (sd = 28, n = 252) = 41.337, p < .05]. Having the calculated chi-square value higher than the table value showed that the internal consistency is low. According to these findings, it was concluded that the data based on scaling did not provide the assumptions of the V state. For this reason, in order to determine the importance sequence of the information that the pre-serviceteachers would like to see in the test guideline, the scaling was made by the equation of state III.

## **Results**

In the research findings, it is seen that the information of duration of exam, scoring and whether the measures are taken against success by chance on the test directives have been sorted by the students from the most important to the least important based on the order of importance. It could be stated that the information of what the prohibited behaviors during the exam are and what to do in case of cheating are not very important to the students. According to gender, it is seen that for both female and male students the information of scoring, duration of exam and whether the measures are taken against success by chance included in the test directives, rank near the top and the students pay more attention to this information than others. When sorting of the information that is regarded as the most important information in the test directives by the students according to the grade levels, it is seen that the information of scoring, duration of the exam, whether the measures are taken against success by chance rank near the top according to the third and fourth grade students and the students pay more attention to this information than others. It is seen that the information of what to do in case of cheating and what are the prohibited behaviors during the exam included in the exam directive are among the least important according to the grade level variable.

## **Discussion and Conclusions**

Specifying the exam duration and adjusting it according to the questions in the test directive affects the response behaviors of the students. The duration of answering the test is one of the important variables for responders. Indicating the duration of the exam in the test guideline and adjusting it according to the questions affects the students' response behaviors. According to Frary (1980), in responding to a multiple-choice item, the knowledge level of the respondent causes the process of relating the



correct option to distracters and believing that one of the distracters is true. Besides, the respondent can also associate the correct option with one or more distracters and make a prediction among the remaining two or more distracters. These conditions indicate the variabilities in the respondent's knowledge, in other words, the presence of partial information at different levels. In addition, the lack of appropriate test duration in the achievement tests may lead the responder to give random responses to these items as he/she could not find adequate time for reading some of the items (Turgut, 1971). Explaining the score value of each problem and whether the correction formula will be used or not in the test directive are important information that respondents look for. It is therefore necessary for teachers and test developers to take into consideration what are the response behaviors and what type of response behaviors are shown in which situations during the test development process.

The information of whether the measures are taken against success by chance causes differences in the test directive (Çelen and Demirtaşlı 2006, Espinosa and Gardezabal, 2010; Kadioğlu 2002; Umay, 1998). In the examination conducted for variables of gender and grade level, it is found that the order of importance for the information sought in the test directive is similar to ranking obtained for the whole group. As a result of this research, the guidelines which are not given enough attention by some of the test developers while preparing the exam paper were scaled by taking opinions of the students into consideration by the scaling method and their importance was emphasized. The findings and the results of the research, the scaling made by the pre-service teachers can be shared with the instructors who prepare the exam paper. In conclusion, seminars related to the test directive could be organized and the deficiencies of the instructors about how to prepare the tests could be eliminated and their problems could be solved.