

Çevrimiçi Sınav Görüş Anketi

Özkan YILMAZ

Matematik ve Fen Bilimleri Eğitimi Bölümü, Eğitim Fakültesi, Erzincan Üniversitesi, Erzincan,
ozkanyilmaz@erzincan.edu.tr

Özet

Bu çalışmanın amacı; eğitim öğretimde kullanılan kağıt-kalem sınavlarına alternatif olarak uygulanan çevrimiçi (online) sınavların değerlendirilmesine yönelik bir görüş anketi hazırlamaktır. Araştırma, Türkiye'nin doğusunda yer alan yakın zamanda kurulmuş orta büyüklükteki bir üniversitede yapılmıştır. Çevrimiçi sınav uygulaması Fen Bilgisi Öğretmenliği Bölümünde öğrenim gören Bilgisayar I dersini alan 80 öğrenci ile yapılmıştır. Uygulama sonucu, anket için cronbach alfa güvenilirlik katsayısı $\alpha=0.87$ olarak hesaplanmıştır. Anket üç alt faktörden oluşmaktadır. Alt faktörler için cronbach alfa güvenilirlik katsayıları; (1) pratiklik ve kullanılabilirlik $\alpha=0.89$ (8 Madde), (2) duyuşsal etki faktörü $\alpha=0.82$ (6 Madde), (3) güvenilirlik $\alpha=0.82$ (3 Madde) olarak hesaplanmıştır. Anket, çevrimiçi sınavlarının öğrenci görüşlerine göre belirtilen alt faktörler boyutunda değerlendirilmesinde kullanılabilir niteliktedir. Farklı çevrimiçi sınav uygulamalarında kullanılarak genel geçerlik ve güvenilirliğinin değerlendirilmesi önerilmektedir.

Anahtar kelimeler: Çevrimiçi sınav, Değerlendirme ölçeği, Fen öğretimi

Online Examination Assessment Survey

Abstract

The purpose of this study is to develop a new survey for evaluating online examination which is use as an alternative to paper-pencil exams. Study was conducted in medium-sized university which was located in east part of Turkey and it was established recently. The implementation was conducted with 80 students enrolled in Computer I course in Department of Science Education. After the online examination the cronbach alfa index was calculated as $\alpha=0.87$ for scale. Online examination scale has three sub-factors. cronbach alfa index for all three sub-factors were calculated (1) practicality-suitability $\alpha=0.89$ (8 Items), (2) affective factors $\alpha=0.82$ (6 Items), (3) reliability $\alpha=0.82$ (3 Items). The survey can be use, depend on students' views, to evaluate online examination with specific factors. Additionally, the survey needs to implement in different online examination to evaluate general validity and reliability.

Keywords: Online examination, Assessment scale, Science instruction

GİRİŞ

Yükseköğretimde öğrenci niteliklerinin geliştirilmesi önemli bir konudur. Niteliklerdeki gelişim düzeylerini belirlemek ise ölçme, değerlendirme ve dönütü sağlama ekseninde gerçekleşmektedir. Gerek öğretimin niteliğini gerek akademik başarıdaki ilerlemenin ve gerekse ihtiyaç duyulan dönüt sağlamak için yapılan sınavlar eğitim öğretimde önemli bir yer tutmaktadır.

Bir ölçüm aracı olan sınavlar, öğretmenin öğrencisini belirli bir alandaki performansını diğer öğrencilere göre değerlendirebilmesi veya belirlenmiş bir standarda göre değerlendirebilmesine imkân tanır (Woolfolk, Hughes ve Walkup, 2008). Formal olarak norm kaynaklı ve ölçüt kaynaklı olmak üzere iki tür sınav bulunmaktadır (Evers ve Walberg, 2004). Norm kaynaklı sınavlar bir kişinin önceden kabul edilen bir norma bağlı olarak değerlendirilmesinde kullanılmaktadır. Eğitimde en az üç tip norm grupları bulunmaktadır. Bunlar, sınıf içerisinde veya okul genelinde, il-bölge genelinde ve son olarak ülke genelinde olan norm gruplarıdır (Woolfolk vd., 2008). Örneğin, sınıf içerisinde yapılan bir final sınavında başarılı ilk üç öğrenci için norm o sınava giren öğrenci grubu iken, Temel Eğitimden Orta Öğretime Geçiş (TEOG) sınavında başarılı öğrencilerin ilgili liselere yerleştirilmesinde norm grubu Türkiye genelinde sınava giren tüm öğrencilerdir. Formal diğer bir sınav türü olan ölçüt kaynaklı sınavlarda ise belirlenmiş bir alanda öğrencilerin ne yapabildiği veya yapamadığının ortaya konmasını amaçlayan sınavlardır. Her hangi bir grup ile kıyaslama yerine belirlenen bir ölçüte göre değerlendirme yapılmaktadır (Woolfolk vd., 2008). Örnek olarak, bir öğrencinin bir dersi başarıyla tamamlayabilmesi için final sınavından yüz üzerinden en az yetmiş almasının gerektiği bir durumda, sınavda kullanılan bu ölçüt öğrencinin o ders için başarılı veya başarısız sayılmasında kullanılmaktadır.

Sınavlarda kullanılan soru tiplerine bakıldığında; doğru-yanlış, eşleştirme tipi sorular, çoktan seçmeli maddeler içeren sorular, tamamlama tipi sorular, açık uçlu sorular, kısıtlı cevaplı sorular türünde çeşitlilik göstermektedir (Borich, 2013). İlgili alan öğretmeni bu soruların birini veya birkaçını kullanarak öğrencilerin derste akademik başarısını belirli bir ölçüte bağlı olarak değerlendirmektedir. Sınıf içerisinde yüz-yüz yüze eğitim öğretim ortamında yapılan bu sınavlar teknolojiye bağlı olarak bilgisayar ortamında veya belirli bir fiziksel mekânda bulunmadan internet ortamında da yapılmaya başlanmıştır. Bu yeni sınav uygulaması

ile birlikte, sınavlar kağıt-kalem sınavları ve çevrimiçi (online) sınavlar olmak üzere iki kategoride incelenmiş ve bu alanda çeşitli çalışmalar yapılmıştır (Kulik ve Kulik, 1991; Ni, 2013; Vleeschouwer vd., 2014).

Ayberk, Şahin, Eriş, Şimşek ve Köse (2014) kâğıt-kalem ve bilgisayar formunda uygulanan testlerdeki öğrenci başarısının karşılaştırıldığı çalışmaların geneline yönelik yaptıkları meta-analiz çalışmasında, öğrencilerin performanslarında sınav türü bakımından manidar bir fark olmadığı sonucu ortaya konmuştur. Kamu Personeli Seçme Sınavı'na (KPSS) hazırlık amacı ile yapılan bir başka etkileşimli çevrimiçi sınav uygulamasında öğrencilerin web tabanlı platformda öğrencilerin sorulara her yerden ulaşabilmesi ve deneme sınavları ile tecrübe kazanmaları sağlanmıştır (Bozkurt ve Tekedere, 2013). Matematik dersinde uygulanan, isteğe bağlı çevrimiçi sınava yönelik öğrenci görüşlerinin değerlendirildiği bir çalışma sonucunda, öğrencilerin kısmen olumlu görüşe sahip olmalarına karşın genelde kararsızlık yaşadıkları görülmüştür (Çiğdem ve Tan, 2014). Benzer bir çalışmada öğrencilerin çevrim içi sınava yönelik olarak memnuniyet duydukları rapor edilmiştir (DeSouza ve Fleming, 2003). Gretes ve Green (2014) üniversite düzeyinde yaptıkları çalışmada, eğitim derslerinde kullanılan bilgisayar ortamındaki test etkinliklerinin öğrenci performanslarında artış sağladığını ifade etmişlerdir. Bununla birlikte farklı programlardaki farklı dersler için beklentilerin aynı olmadığını bu nedenle bilgisayar ortamında ölçme-değerlendirme etkinliklerinin tüm dersler için aynı oranda etkili olmadığı görülmüştür. Yine üniversite düzeyinde yapılan bir başka çalışmada, üniversite öğretmenlerinin bilgisayar temelli ve kâğıt-kalem temelli sınavlar hakkında görüşleri incelenmiştir. Üniversite öğretmenleri bilgisayar temelli sınav sistemine yönelik olarak olumlu tutuma sahip olmalarına karşın, bazı durumlarda kâğıt-kalem sınavlarını tercih ettikleri görülmüştür (Jamil, Tariq ve Shami, 2012). Psikometrik özellikleri bakımından incelendiğine, çevrimiçi ve kâğıt-kalem sınavlarının önemli düzeyde farklılıklarının olmadığı görülmüştür (Joubert ve Kriek, 2009).

Yapılan çalışmalar genel anlamda değerlendirildiğinde; daha çok çevrimiçi sınav ile kâğıt-kalem sınavlarının karşılaştırılmasının yapıldığı ve öğretmen görüşü, öğrenci görüşü, öğrenci performans düzeylerine katkısı, memnuniyet durumları gibi çeşitli faktörlerin ele alındığı görülmektedir. Kâğıt-kalem sınavlarına göre oldukça yakın bir zamanda geliştirilen ve kullanılmaya başlanan bilgisayar tabanlı, çevrimiçi sınavların etkin kullanılabilmesi için çeşitli araştırma faaliyetlerinin yapılması beklenen ve istenen bir durumdur.

Bu araştırma faaliyetlerinde çevrimiçi sınavın değerlendirilmesine yönelik standart bir ölçüm aracının olması yapılacak çalışmaların genel değerlendirilmesinde ve bu çalışmaların sonucunda çevrimiçi sınav uygulamalarında istenen başarının sağlanması için önemli olduğu düşünülmektedir. Çevrimiçi sınav uygulamalarının değerlendirilmesi için çeşitli faktörlerin olduğu yukarıda yapılan kısa bir literatür taramasında da görülmektedir. Bu faktörlerden biri olan öğrenci görüşleri, öğrenme ve öğretme etkinliklerinin odak noktasında yer alan öğrenci boyutu ile ayrı bir önem arz etmektedir. Bu çalışmada, öğretimde yapılan çevrimiçi sınavları öğrenci görüşlerine bağlı değerlendirmek için standart bir görüş anketi hazırlamak amaçlanmıştır.

YÖNTEM

Bu çalışma, çevrimiçi sınavları değerlendirmek için kullanılacak bir anket geliştirme çalışmasıdır. Bu kapsamda yapılan uygulama ve işlemler aşağıda detaylı olarak verilmiştir.

Çalışma grubu

Araştırmanın katılımcıları, Türkiye'nin doğu bölgesinde yer alan yakın zamanda kurulmuş, yeni bir üniversite, Eğitim Fakültesi Fen Bilgisi Öğretmenliği Bölümünde öğrenim gören ve "Bilgisayar I" dersini alan toplam 80 öğrenciden oluşmaktadır. İki farklı şube ile gerçekleştirilen çevrimiçi sınav uygulamasına katılan öğrencilerin 52'si bayan, 28'i erkek öğrencidir. Çevrimiçi sınav bir dönem içerisinde yapılan iki sınav (vize ve final sınavları) arasında uygulanmıştır.

Araştırma Süreci

Araştırmada öncelikle öğrencilere çevrimiçi bir sınav uygulaması yapılmıştır. Daha sonra hazırlanan görüş anketi uygulanarak bu sınav hakkında öğrenci görüşleri belirlenmeye çalışılmıştır.

Çevrimiçi sınav uygulaması

Çevrimiçi sınav uygulaması dönem içerisinde yapılan iki sınav (vize ve final) arasındaki zaman diliminde yapılmıştır. Bilgisayar laboratuvarında yapılan sınavda her öğrenci bilgisayarında bireysel olarak kendi sınavını başlatmış ve bitirmiştir. Olası problemlere karşı bir gözetmen sınav süresince sınıfta hazır bulunmuştur. Uygulama sınavı MOODLE isimli ders yönetim sistemi üzerinden uygulanmıştır. Her öğrenci kendi kullanıcı adı ve şifresi ile internet ortamında sisteme giriş yapmıştır. Sınav soruları 3-5 adetle

sınırlandırılmış süre 4-7 dakika arası gibi kısa tutulmuştur. Sınav sorularında çeşitliliğin sağlanması (boşluk doldurma, kısa cevap, çoktan seçmeli... vb.) ve şubeler arası soru alma-verme gibi durumları önlemek için uygulamada şubeler arası kısmi değişiklikler yapılmıştır. Sorular ekrana teker teker gelecek şekilde ve her bir öğrenciye aynı havuzdan farklı soru yönlendirilmek suretiyle planlanmıştır. Böylece yan yana oturan öğrencilerin bir birleri ile aynı soru üzerinde etkileşim kurması azaltılmaya çalışılmıştır. Öğrenciler sınavlarını başlatma ve bitirmede bireysel tercihlere sahip olmakla birlikte genel bir birliktelik sağlamak amacıyla, sınavın başladığı bilgisi ilgili gözetmen tarafından verilmiştir. Sınavın kısa süreli ve az sayıda soru ile yapılmasındaki amaç, öğrencilerin çevrimiçi sınava yönelik bir deneyim kazanmalarınıdır. Öğrencilerin çevrimiçi sınavı, vize ve final sınavlarında yapılan kâğıt-kalem sınavlarında olduğu gibi sınavı önemsemeleri için çevrimiçi kısa sınavdan alınan puanlar başarı notuna yansıtılmıştır. Çiğdem ve Tan (2014), çevrimiçi sınavdan alınan puanların başarı notuna yansıtılması gerektiğini yaptıkları çalışmada vurgulamışlardır.

Anket geliştirme süreci

Her bir anket amacına göre farklılık göstermektedir. Bu anket çalışmasında amaç, öğrenci deneyimlerine bağlı olarak çevrimiçi sınavın öğrenci görüşlerine göre değerlendirmesini yapacak bir standart anket geliştirmektir. Bu amaç doğrultusunda aşağıda belirtilen işlemler yapılmıştır.

Anket için öncelikle dikkat edilen konu cevaplayacakların sıkılmadan doldurabileceği bir formun oluşturulmasıdır. Bunun için ifadeler oldukça kısa tutulmaya çalışılmış ve madde sayısının az olmasına dikkat edilmiştir. Anket maddelerini hazırlamak için çeşitli kaynaklar incelenmiştir (Başol, Çiğdem ve Kocadağ, 2013; Cassidy, Budenz-Anders, Pavlechko ve Mock, 2001; Dermo, 2009). Bu kaynaklar içerisinde; Dermo (2009)'nun 'e-Assessment and the student learning experience: A survey of student perceptions of e-assessment' isimli çalışmasında kullanılan ölçek, altı faktörlü geniş kapsamlı bir ölçek olması ve üniversite düzeyinde uygulamalarda kullanılacak düzeyde olması sebebi ile temel alınmıştır. Diğer çalışmalardan ise soru maddelerinin yazımı konusunda fikir edinmek için faydalanılmıştır. Okulun, gerekli teknolojik sistem alt yapısının çevrimiçi sınav için ne derece uygun olduğunu belirleyebilmek için Bilgisayar ve Öğretim Teknolojileri (BÖTE) bölümünde görevli bir öğretim üyesinin görüşlerinden faydalanılmıştır. 7 ters, 18 düz madde tipi olmak üzere toplamda 25 maddelik anket, öğrenci değerlendirmelerine sunulmak üzere çevrimiçi form olarak MOODLE ders yönetim sistemine yüklenmiş ve çevrim içi sınav uygulamasından sonra öğrencilerin doldurmaları sağlanmıştır. Öğrencilerin, sisteme yüklenen formu doldurmaları için bir hafta süre tanınmıştır. 80 öğrenci görüşü alınan anket için sırası ile aşağıdaki istatistiksel analizler yapılmıştır. Analizlerde ters soru tipleri ilgili maddenin sonuna "T" yazılarak belirtilmiştir.

Hazırlanan anket maddelerinin ölçülmek istenen özelliği ölçüp ölçmediğini belirlemek için yapı geçerliği incelenmiştir. Yapı geçerliği, önceden kabul edilen ve gözlemlenebilir özelliklere bağlı olarak gözlemlenebilir olamayan değişkenlerin yorumlanmasına yönelik geçerlik durumudur (Pedhazur ve Schmelkin, 1991). Sosyal bilimlerde sıklıkla karşılaşılan bir durum ölçülmek istenen yapının, doğrudan ölçülememesidir. Buna örtük veya gizil değişkenler denir (Pedhazur ve Schmelkin, 1991). Bu değişkenlerin genel yapı ile olan ilişkisi veya değişkenlerin kendi içerisinde oluşturduğu grup yapıları belirleyebilmek için faktör analizi kullanılır (Field, 2009). Yapılan faktör analizi ile maddelerin ölçülmek istenen yapı ile olan ilişkisi belirlenmiştir. Anketin güvenilirliğini belirlemek için Cronbach alfa iç tutarlık katsayısı hesaplanmıştır. Ayrıca, elde edilen alt faktörlerin ilişki düzeylerini belirlemek için korelasyon matrisine bakılmıştır.

BULGULAR

Bu bölümde çevrimiçi sınav değerlendirme ölçeği için yapılan istatistiksel analizler ve elde edilen bulgulara yer verilmiştir.

Faktör analizi yapılabilmesi için KMO değeri 0,5 veya yukarı olması gerekmektedir. Tablo 1'deki indeks değerine bakıldığında oldukça yüksek bir değer (.82) olduğu görülmektedir. Ayrıca, küresellik testi anlamlı farklılık ($p \leq 0.05$) göstermiştir. Bu değerler maddeler arasında bir korelasyonel ilişkinin var olduğu ve belirgin bir kümeleşmenin olduğunu göstermektedir.

Tablo 1.

KMO ve Bartlett Testi

Kaiser-Meyer-Olkin Örneklem Yeterliği (KMO)	,820
Yaklaşık Ki-kare	1016,562
Bartlett Küresellik Testi	Serbestlik Derecesi (Sd) 231
	Anlamlılık (p) ,000

İkinci aşamada, her bir maddenin toplam yapı ile olan ilişkisine bakmak için madde-toplam korelasyon değerleri incelenmiştir. Tablo 2’de görüldüğü üzere 9, 17 ve 18 numaralı maddelerin madde-toplam korelasyon değerleri negatif değer aldığı görülmektedir. Bu maddeler ölçülmek istenen özellikle ilgili yapıyla ilişkisi olmadığını göstermektedir. Bu nedenle bu maddeler çıkarılarak tekrar madde-toplam istatistiğine bakılmıştır. 22 madde ile yapılan analiz sonucu her bir maddenin belirli bir değerde toplamla ilişkili olduğu görülmüştür.

Tablo 2.*Madde-Toplam İstatistiği*

Maddeler	Madde çıkarıldığında anket ortalaması	Madde çıkarıldığında anket varyansı	Düzeltilmiş madde-toplam korelasyon değeri	Madde çıkarıldığında anket alfa değeri
m1	112,21	432,524	,281	,843
m2	112,16	435,075	,249	,844
m3	112,75	427,911	,272	,844
m4	112,74	413,234	,511	,836
m5	113,60	401,737	,477	,836
M6T	113,68	423,893	,269	,844
M7T	114,63	408,946	,390	,840
m8	112,93	407,387	,580	,833
M9T	114,49	462,633	-,180	,860
m10	113,69	396,040	,624	,830
m11	113,05	395,491	,745	,827
m12	114,04	397,201	,628	,830
M13T	113,65	411,218	,471	,837
m14	113,34	405,340	,663	,831
M15T	113,85	424,711	,326	,842
M16T	114,58	403,615	,517	,835
m17	113,53	461,594	-,170	,860
m18	113,79	453,207	-,069	,855
m19	114,18	410,728	,485	,836
M20T	113,48	418,936	,364	,841
m21	112,78	421,923	,440	,838
m22	113,24	424,740	,310	,842
m23	112,74	429,285	,316	,842
m24	113,05	396,403	,709	,828
m25	113,18	398,020	,766	,827

Maddeler arası ilişkiye bağlı olarak ölçeğin kaç faktörden oluştuğunu belirlemek için açımlayıcı faktör analizi yapılmıştır. Çevirme metodu olarak “Varimax” kullanılmıştır. Yapılan analiz sonucunda Tablo 3’de gösterilen sonuçlar elde edilmiştir. Tablo 3 incelendiğinde analiz sonucu bileşenlerin altı tane olduğu görülmektedir. Ancak, bazı maddenin birden fazla bileşenle ilişkisi olduğu görülmektedir. Maddelerin sadece tek bir bileşen altında gruplanması gerekmektedir. Bu gruplamaları yapabilmek için birden fazla bileşenle ilişkili olan maddeler sırası ile çıkarılarak (24,19,1,2,3) yalnız bir bileşenle ilişkili maddeler belirlenmeye çalışılmıştır. Bunun için faktör analiz işlemi üç kez tekrarlanmıştır. Çıkarma işlemi birer madde azaltılarak yapılmıştır. Hangi maddenin çıkarılması gerektiğine öncelikle maddenin farklı faktörlerdeki yük değerleri arasındaki fark durumu incelenerek yapılmıştır. Farklı faktörlerdeki yük değerleri arasında 0,1’den küçük fark var ise, bu maddeler binişiklik eğilimi gösteren maddelerdir ve analizden çıkarılması gerekmektedir (Gürbüz ve Şahin, 2014). Çıkarma işlemi, yük değerlerine bağlı olarak sıralanmış listede sondan başa doğru 0,1 fark değerleri incelenerek teker teker yapılmıştır. Madde-toplam korelasyon katsayıları incelenirken 0,3’den küçük olan ve 0,9’dan büyük olan değere sahip maddelerin çıkarılması gerekmektedir (Field, 2009). Bu nedenle, tablo değeri 0,3 olarak sınırlandırılmıştır.

Tablo 3.
Madde-faktör Yükleri Dağılımı

Maddeler	Bileşenler					
	1	2	3	4	5	6
m10	,799					
m14	,722			,357	,327	
m5	,711					
m4	,703		,389			
m11	,559			,545		
m25	,557			,331		,485
m12	,542	,493		,475		
m8	,540			,405	,404	
M7T		,777		,314		
M16T		,772				
M15T		,726				
M6T		,683				
M13T	,423	,656				
m2			,839			
m1			,833			
m3			,575			
m19				,827		
m22					,807	
m23			,307		,762	
m21			,552		,654	
M20T		,454				,727
m24	,392			,473		,593

Açıklama Metodu: Temel Bileşenler Analizi

Çevirme Metodu: Varimax, Kaiser Normalizasyonu

Tablo 4’de bakıldığında geriye kalan maddelerle yapılan faktör analizinde çoğu maddenin ilgili bileşenle yüksek değerle ilişkisi olduğu görülmektedir. Bazı maddelerin ise birden fazla bileşenle korelasyonel ilişkili olduğu ortaya çıkmıştır. Ancak, bu katsayılar arasındaki fark 0,1’den fazla olduğu için ilgili maddeler yüksek katsayı değerine göre hangi bileşende değerlendirileceğine karar verilmiştir.

Tablo 4.
Düzeltilmiş Bileşenler Matrisi

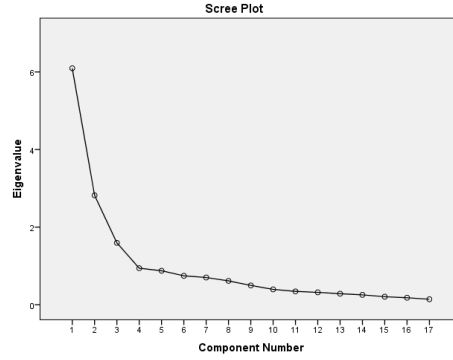
Maddeler	Bileşenler		
	1	2	3
m14	.828		
m10	.815		
m11	.764	.332	
m12	.716	.397	
m8	.682		.338
m25	.664		.374
m5	.658		
m4	.603		
M7T		.779	
M16T	.349	.735	
M6T		.733	
M15T		.665	
M20T		.656	
M13T	.404	.655	
m23			.857
m21			.837
m22			.783

Özdeğerler 6.09 2.82 1.59

Varyans yüzdeliği 35.85 16.58 9.38

Açıklama Metodu: Temel Bileşenler Analizi

Çevirme Metodu: Varimax, Kaiser Normalizasyonu



Grafik 1. Özdeğer ve bileşen çizgi grafiği

Tablo 4 incelendiğinde 17 anket maddesinin üç farklı bileşen altında toplandığı görülmektedir. Ayrıca Grafik 1’de özdeğer ve bileşen grafiği incelendiğinde eğim noktasının 3. ve 4. bileşenler arasında olduğu görülmektedir. Birinci bileşen varyansın yaklaşık %36’sını, ikinci bileşen yaklaşık %17 ve üçüncü bileşen yaklaşık %9’unu açıklamaktadır. Üç faktör kümülatif olarak toplam varyansın %61.8’ini açıklamaktadır. Bu bileşenlere ait iç tutarlılık güvenilirlik analizi sonrasında elde edilen güvenilirlik katsayıları Tablo 5’de verilmiştir. Görüldüğü üzere bileşenler iyi düzeyde güvenilirlik katsayılarına sahiptir.

Tablo 5.
Bileşenler İç Tutarlılık Alfa Güvenirlik Katsayıları

Bileşenler	Madde sayısı	Güvenirlik
1	8	0.88
2	6	0.82
3	3	0.81
Toplam	17	0.87

SONUÇ VE TARTIŞMA

Çevrimiçi sınavları değerlendirmek için hazırlanan anket için yapılan istatistiksel analizler sonucu hesaplanan cronbach alfa güvenilirlik katsayısı $\alpha=0.87$ ’dir. Bu değer, ölçeğin güvenilirlik düzeyinin yüksek olduğunu göstermektedir. Anket için alt bileşenlerin olup olmadığı değerlendirmek için yapılan faktör analizi sonucunda, üç faktörlü bir yapının olduğu ortaya çıkmıştır. Bu faktörlerin isimlendirmelerinde Dermo (2009)’nun çalışması önemli ölçüde referans alınarak; 1. faktör için “pratiklik ve kullanılabilirlik” 2. faktör için “duyuşsal etki faktörü” ve 3. faktör için “güvenirlik” isimlendirmesi yapılmıştır. Alt faktörler için güvenilirlik katsayıları (1) Pratiklik ve kullanılabilirlik: 0.88, (2) Etki faktörü: 0.82 ve (3) Güvenirlik: 0.81 olarak hesaplanmıştır.

Yapılan istatistiksel analizler sonucunda anketin belirlenen alt faktörler boyutunda ve öğrenci görüşlerine bağlı olarak çevrimiçi sınav değerlendirmelerinde kullanılabilirliği sonucu ortaya çıkmıştır. Bununla birlikte, çalışma fen bilgisi öğretmenliği öğrencileri ile sınırlıdır. Farklı bölümlerde öğrenim gören öğrencilere uygulanarak doğrulayıcı faktör analizi yapılmalı ve anketin genel geçerliğinin değerlendirilmesi gerekmektedir. Yine, çevrimiçi sınavında yer alan soru sayısı ve uygulama sayısı diğer bir sınırlılıktır. Yapılacak çalışmalarda, çevrimiçi sınav uygulamalarının kağıt-kalem uygulamalarının nitelik ve niceliklerine benzer şekilde yapılması önerilmektedir.

Çevrimiçi sınavları değerlendirmek için yapılan anket geliştirme çalışması sonucunda anketin belirlenen faktörler boyutunda kullanılabilir olduğu sonucu ortaya çıkmıştır. Hazırlanan öğrenci görüşlerine bağlı çevrimiçi sınav değerlendirme anket maddeleri Ek 1’de verilmiştir.

KAYNAKÇA

- Ayberk, E., Şahin, D., Eriş, H. M., Şimşek, A. S., & Köse, M. (2014). Kağıt kalem ve bilgisayar formunda uygulanan testlerde öğrenci başarısının karşılaştırıldığı çalışmaların meta-analizi. *Asya Öğretim Dergisi*, 2(2), 18-26.
- Başol, G., Çiğdem, H., & Kocadağ, T. (2013). Ölçme değerlendirme dersinde e-sınav uygulaması ve öğrenci görüşleri. Paper presented at the Uluslararası Eğitimde Değişim ve Yeni Yönelimler Sempozyumu, Konya.
- Borich, G. D. (2013). *Effective teaching methods* (8 ed.). England: Person Education.
- Bozkurt, A., & Tekedere, H. (2013). Etkileşimli çevrimiçi KPSS deneme sınavı yazılımı. *Gazi Üniversitesi Endüstriyel Sanatlar Eğitim Fakültesi Dergisi*, 31, 36-44.

- Cassady, J. C., Budenz-Anders, J., Pavlechko, G., & Mock, W. (2001). *The effects of internet-based formative and summative assessment on test anxiety, perception of threat, and Achievement*. Paper presented at the Annual Meeting of The American Educational Research Association, Seattle WA.
- Çiğdem, H., & Tan, Ş. (2014). Matematik dersinde isteğe bağlı çevrimiçi kısa sınav uygulanması hakkında öğrenci görüşleri. *Bilgisayar ve Eğitim Araştırmaları Dergisi*, 2(4), 51-73.
- Dermo, J. (2009). e-Assessment and the student learning experience: A survey of student perceptions of e- assessment. *British Journal of Educational Technology*, 40(2), 203-214.
- DeSouza, E., & Fleming, M. (2003). A comparison of In-class and online quizzes on student exam performance. *Journal of Computing in Higher Education*, 14(2), 121-134.
- Evers, W. M., & Walberg, H. J. (2004). *Testing student learning, evaluating teaching effectiveness*. Stanford, Calif.: Hoover Institution Press.
- Field, A. (2009). *Discovering Statistics Using SPSS*. 1 Oliver Yard 55 City Road London EC1Y 1SP: SAGE Publications Ltd. .
- Gretes, J. A., & Green, M. (2014). Improving undergraduate learning with computer-assisted assessment. *Journal of Research on Computing in Education*, 33(1), 46-54.
- Gürbüz, S., & Şahin, F. (2014). *Sosyal Bilimlerde Araştırma Yöntemleri*. Ankara: Seçkin yayıncılık.
- Jamil, M., Tariq, R. H., ve Shami, P. A. (2012). Computer-based vs paper-based examinations: Perceptions of university teachers. *The Turkish Online Journal of Educational Technology*, 11(4).
- Joubert, T., & Kriek, H. J. (2009). Psychometric comparison of paper-and-pencil and online personality assessments in a selection setting. *SA Journal of Industrial Psychology/SA Tydsrkif vir Bedryfsielkunde*, 35(1), #727.
- Kulik, C.-L. C., & Kulik, J. A. (1991). Effectiveness of computer-based instruction: An updated analysis. *Computers in Human Behavior*, 7(1-2), 75-94. doi: [http://dx.doi.org/10.1016/0747-5632\(91\)90030-5](http://dx.doi.org/10.1016/0747-5632(91)90030-5)
- Ni, A. Y. (2013). Comparing the effectiveness of classroom and online learning: Teaching research methods. *Journal of Public Affairs Education*, 19(2), 199-215.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, Design, and Analysis: An Integrated Approach*. 365 Broadway Hillsdale, New Jersey 07642: Lawrence Erlbaum Associates, Inc.
- Vleeschouwer, M., Schubart, C. D., Henquet, C., Myin-Germeys, I., Gastel, W. A. v., Hillegers, M. H. J., . . . Derks, E. M. (2014). Does Assessment Type Matter? A Measurement Invariance Analysis of Online and Paper and Pencil Assessment of the Community Assessment of Psychic Experiences (CAPE). *PLoS One*. doi: 10.1371/journal.pone.0084011
- Woolfolk, A., Hughes, M., & Walkup, V. (2008). *Psychology in Education*. Madrid: Pearson Education Limited.

EXTENDED SUMMARY

The purpose of this study is to develop a new scale for evaluating online examination which is using as an alternative to paper-pencil exams. Study was conducted in medium-sized university which was located in east part of Turkey and it was established in the more recently. The application made with 80 students enrolled in Computer I course in Department of Science Education. Application of online examination was used between two paper-pencil exams which were called mid-term exam and final exam. Examination was conducted in Computer Laboratory Classroom. All students were used own computer for examination starting and stopping. For possible technical problems, an observer was stands by in the classes. Students used their personal user name and password to access MOODLE which is a kind of learning management system. Questions showed on PC screen one by one and the different order in same questions pool. So that, every students answered same questions but different order. That application used for cheating prevention between students. Students could start their own examination when they ready. But for synchronizations, observer was announced the start of exam. Main point of online exam was students' experience. So that, exam was not taking a long time and no more than 5 questions.

Before the preparing scale items, aimed to prepare short and clear sentences. Therefore, review of the literature was used for preparing items pool. Also, one of academician's views used who is a staff in department of Computer and Instructional Technology. 7 reserve and 18 normal questions used in the items pool. Alfa index (α) was used to determine scale reliability. The internal consistency analysis was used to assess the construct validity of scale.

Factor analysis used to data reduction for scale. There are two types of factor analysis which are Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). In this study, used EFA to define sub-factors of scale because study does not have any basis of pre conceived special theories. KMO & Bartlett's Test of Sphercity used to check the sampling adequacy for factor analyses. The range of the KMO results falls between 0 to 1. And, the accepted index globally is 0.6 and up. The other important step for factor analyses, Bartlett's Test of Sphercity has to be less than 0.05. Analyze results showed that KMO index is 0.8 and Barlett's Test of Sphercity under the 0.05 ($p \leq 0.05$) for scale. Varimax rotation technique was used to define items and correlated factor. Factor analyses were done three times to find items and sub factor coherence.

The current study found that online examination assessment scale had three components. These are (1) practicality-suitability $\alpha=0.89$ (8 Items), (2) affective factor $\alpha=0.82$ (6 Items), (3) reliability $\alpha=0.82$ (3 Items). The scale can use in science instruction to evaluate online examination with specific factors. Additionally, it needs to use in other courses and departments to evaluate general validity and reliability of that scale.

Ek 1: Çevrimiçi sınav değerlendirme anket maddeleri ve faktörler

Faktör 1: *Pratiklik ve kullanılabilirlik*

14. İnternet üzerinden yapılan sınavdaki sorular daha anlaşılırdı.
10. Çevrimiçi sınavı kağıt-kalem sınavına tercih ederim.
11. Sınavın internet üzerinden olması hoşuma gitti.
12. Bütün sınavların internet üzerinden olmasını isterim.
8. Çevrimiçi sınav, kağıt-kalem sınavına göre daha düzenliydi.
25. Çevrimiçi sınavdaki sorular öğrenmemi daha iyi ölçüyordu.
4. İnternet üzerinden yapılan sınav etkili ve kullanışlıydı.
5. Kağıt-kalem sınavı yerine çevrimiçi sınavı tercih ederim.

Faktör 2: *Duyuşsal etki faktörü*

7. Çevrimiçi sınav, kağıt-kalem sınavına göre daha stresliydi.
16. Çevrimiçi sınavda sınav kaygım daha fazlaydı.
6. Çevrimiçi sınava uyum sağlamam zor oldu.
15. Çevrimiçi sınavdan sonra, kağıt-kalem sınavına göre daha çok yorulduğumu hissettim.
13. Bütün sınavların kağıt-kalem olmasını isterim.
20. Çevirim içi sınavda kağıt-kalem sınavına göre soruları okumak daha zordu.

Faktör 3: *Güvenirlilik*

23. Çevrimiçi sınavda soruların seçimi uygundu.
21. Çevrimiçi sınavda soruların zorluğu uygundu
22. Çevrimiçi sınavda sorular ayırt ediciydi.