

## SINAV TÜRLERİNİN MUTLAK YARGILAR YÖNTEMİ İLE ÖLÇEKLENMESİ VE ÖĞRENCİ GÖRÜŞLERİ ÜZERİNDEN DEĞERLENDİRİLMESİ

Esin TEZBAŞARAN\*

### ÖZ

*Bu araştırmanın amacı, öğrenci yargılarına göre sınav türlerini ölçeklemek ve ölçek değerlerine göre sınav türleri hakkındaki öğrenci görüşlerini incelemektir. Bu çalışma için 2014-2015 öğretim yılı İstanbul Üniversitesi Hasan Ali Yücel Eğitim Fakültesi Pedagojik Formasyon eğitimine kayıtlı 243 öğrenciden veri toplanmıştır. Öncelikle, mutlak yargılar yöntemi ile altı adet sınav türü ölçeklenmiştir. Ardından, sınav türlerinin ölçek değerlerine göre, öğrencilerin sınav türlerine vermiş oldukları derecelerin gerekçeleri, içerik analizi yöntemi ile incelenmiştir. Ölçek değerine göre birinci sırada tercih edilen sınav türü çoktan seçmeli, son sırada yer alan tür ise sözlü sınavlardır. Öğrencilerin sınav türleri hakkındaki olumlu ve olumsuz görüşleri on kategoride toplanmış ve alanyazındaki sınav türlerine ait özellikler ile örtüştüğü gözlenmiştir. Sonuç olarak, araştırmaya katılan öğrencilerin eğitim yaşantılarında, sınav türlerinin olumlu ve olumsuz yanlarını deneyimlemiş oldukları söylenebilir.*

**Anahtar Kelimeler:** ölçekleme, mutlak yargılar yöntemi, sınav türleri

### SCALING TYPES OF TEST WITH THE METHOD OF ABSOLUTE JUDGEMENT AND ASSESSING THEM WITH THE STUDENTS' VIEWS

#### ABSTRACT

*The aim of this study is to scale the types of tests with the method of absolute judgement and to investigate the reasons of the students' judgements on them. For this study, the data was gathered from 243 pedagogical formation students at Istanbul University Hasan Ali Yücel Education Faculty in 2014-2015 academic year. As a first step, six types of test have been scaled with the method of absolute judgement. As a second step, with regard to the scale values of test types, the reasons of the degrees given to the test types by the students have been investigated by using the content analysis. Considering their degrees, multiple choice tests are most favoured one among types of tests, oral exams, on the other hand, are least preferable. The reasons of the students' positive and negative judgements on types of tests were gathered under ten topics of category and it was observed that the arising judgements*

---

\* Yrd. Doç. Dr., İstanbul Üniversitesi Hasan Ali Yücel Eğitim Fakültesi, İstanbul-Türkiye, [tezbasaranesin@gmail.com](mailto:tezbasaranesin@gmail.com)

*correspond to the characteristics of the types of tests in literature. As a conclusion, it can be said that students who participated to this research had experienced the positive and negative sides of the types of tests in their previous educational lives.*

**Keywords:** *scaling, the method of absolute judgement, types of test*

## 1. GİRİŞ

### 1.1. Sınav Türleri

Eğitimde başarının ölçülmesi için kullanılan yöntemlerin seçimi, “başarı” adı altında neyin ölçüleceğinin tanımlanması ile başlar. Okul başarısı, genellikle bilişsel alana dayalı öğrenmeleri barındırdığından dolayı, kâğıt-kalem kullanılarak hazırlanan sınavlarla yapılmaktadır. Bunun yanı sıra öğretmenlerin sözlü sınavları da tercih edebildikleri görülmektedir. Tercih edilen sınav türü ne olursa olsun, her biri güvenilirlik, geçerlik ve kullanılabilirlik açısından bir ölçme aracı olarak incelenmesi gerekmektedir. Bu açılarından her bir sınav türünün kendi içinde avantajlı ya da dezavantajlı olduğu durumlar söz konusudur.

Sözlü sınavların bilişsel alan yeterliklerini ölçmede güvenilirlik ve geçerlikleri düşük sınavlar olmakla birlikte, konuşma becerileri veya sözlü anlatım gücünü ölçmek için diğerlerine göre en uygun seçenektir (Gelbal, 2013; Özçelik, 2009; Turgut ve Baykul, 2010).

Uzun cevaplı yazılı sınavlar ise yazılı anlatım gücünü ölçmede iyi bir yöntem olmakla birlikte, bilişsel alan yeterliklerini ölçmede güvenilirlik ve geçerlik açısından sıkıntıları vardır. Öğrenmeleri ölçmek amacıyla yazma becerileri ve yazılı ifade gücünün ölçülmesi dışında, zorunlu olmadıkça, uzun cevaplı yazılı sınavlara başvurulmaması gerekir (Özçelik, 2009). Ancak hazırlanması kısa zaman aldığından dolayı yaygın olarak kullanılmaktadır (Gelbal, 2013). Bu sınav türünün uygulaması ve puanlamasının uzun sürmesi, ayrıca puanlamanın öznel oluşu gibi özellikleri vardır (Doğan, 2006a; Gelbal, 2013; Turgut ve Baykul, 2010).

Kısa cevap gerektiren sınavlar sorularının yapısı bakımından çok çeşitli olabilir (Turgut ve Baykul, 2010). Ancak uygulamada boşluk doldurma ya da kısa cevap gerektiren sorulardan oluşan sınavlar ağırlıkta kullanılır. Kısa cevaplı yazılı sınavlar, uzun cevaplı yazılı sınavların olumsuz özelliklerini azaltma amacıyla geliştirilen sınavlardır (Özçelik, 2009). Bu sınav türünde cevaplar kısa, soru sayısı çok ve puanlaması görece daha objektiftir (Doğan, 2006b; Gelbal, 2013; Turgut ve Baykul, 2010).

Doğru-yanlış soru tipinden oluşan sınavlar üzerinden, bilgi düzeyinin yanı sıra üst düzey öğrenmeler de yoklanabilir (Özçelik, 2009). Bununla birlikte çok sayıda soru sorulabilmesi ile konu kapsamındaki davranışların ölçülmesine fırsat verir; uygulama süresi kısa ve puanlaması tamamen objektiftir (Gelbal, 2013). Bu açıdan geçerli bir sınav olmakla birlikte, şansla doğruyu yakalama olasılığının %50 olması, bu sınavın en belirgin olumsuz yanındır.

Çoktan seçmeli testler, özellikle büyük gruplarda uygulanmasının getirdiği avantajlardan dolayı, üzerinde çokça çalışılan bir ölçme araç ve yöntemidir (Gelbal, 2013). Çoktan seçmeli sınavlar, bugüne kadar bulunabilmiş en üstün ölçme araçları olmakla birlikte, iyi düzeyde seçmeli sorular belli bir hazırlık dönemi ve deneyimden sonra yazılabilmektedir (Özçelik, 2009). Aksi takdirde, yetkin olmadan çoktan seçmeli sınav hazırlamak ve böyle bir sınavın test ve madde istatistiklerini incelemeyi uygulamak, bu sınavı en kötü ölçme aracı haline getirebilir. Bu sınav türünde soru sayısı çok, uygulama ve puanlama kolay ve puanlama objektiftir (Doğan, 2006c; Gelbal, 2013; Turgut ve Baykul, 2010).

Sınav türlerinin alanyazında geçen ve altı çizilerek belirtilen üstün ve aksak yanları dikkate alınarak, bu sınavların öğretmenler tarafından kullanılması “eğitimde ölçme” açısından önemlidir. Sınavlar, öğrencilerin nasıl çalışacağına ve ne öğreneceğine doğrudan ve güçlü bir etki yapar. Bu nedenle öğretmenlerin hazırladığı sınavlar standart sınavlara göre eğitsel açıdan daha önemlidir (Ebel, 1965). Öğrenimleri boyunca sınavlarla karşılaşan öğrencilerin, hangi sınav türlerini, diğerlerine kıyasla, neden tercih ettiklerini belirlemek, öğretmenlerin sınavları, onların özelliklerini dikkate alarak, uygulayıp uygulamadıklarının bir göstergesi olabilir. Bundan dolayı bu araştırmada, öncelikle sınav türlerinin öğrenciler tarafından yargılanarak ölçme değerlerinin ortaya konması ve daha sonra nedenlerine yönelik düşüncelerin saptanması amaçlanmıştır.

## **1.2. Mutlak Yargılar Yöntemi ile Ölçme**

Psikolojik yapıların ölçümüne dair işlemler bazı kurallara göre yapılır. Sistemik kuralların ve deneysel gözlemleri nicelleştirmek için anlamlı ölçme biriminin gelişimi ölçme olarak adlandırılır (Crocker ve Algina, 1986). Psikolojik yapıların “ölçülebilirliği” hakkında temel iki denence bulunur. Bunlardan birincisi, psikolojik süreklilik de denilen, yapının tek boyutlu süreklilik içeren bir skala

üzerinde farklı değerler alacak biçimde nicelleştirilebilmesi; diğeri ise bu süreklilik üzerindeki değerler, gerçel sayılar sisteminin özelliklerine sahip olması gerekliliğidir (sıralılık, uzaklık ve orijin) (Crocker ve Algina, 1986).

Bu çerçevede ölçekleme için üç temel yaklaşım benimsenir. Birey odaklı yöntemler, uyarıcı odaklı ya da yargılama yöntemleri ve tepki odaklı yöntemler. Birey odaklı yöntemlerde temel amaç, bireyleri sürekli ölçek üzerinde farklı noktalara yerleştirmektir. Likert tipi ölçekleme bunun bir örneğidir. Uyarıcı odaklı yöntemlerde ise amaç, uyarıcıların/maddelerin psikolojik süreklilik içinde yerlerini belirlemektir. Karşılaştırmalı ve sınıflandırmalı yargılar yasası ile yapılan ölçeklendirme işlemleri, uyarıcı odaklı bir yaklaşımla geliştirilmiştir. Son olarak tepki odaklı ölçekleme yöntemlerinde amaç, maddenin gücüne göre bireyleri psikolojik süreklilik üzerinde ölçek değerini belirlemeye çalışırken; aynı zamanda, o maddelere tepki veren bireylerin sayısı ile maddeleri ölçeklemektir. Guttman'ın scolagram analizi bu yöntemlere bir örnektir (Torgerson, 1958).

Bu araştırmada kullanılan mutlak yargılar yöntemi, uyarıcıları ölçeklendirmek için kullanılan ölçekleme yöntemlerinden birisidir. Bu yöntem, psiko-fizikte yargılamanın açıkça belirlenmiş karşılaştırılabilir ölçütü olmadan, cevaplayıcıya uyarıcıları belli dereceler üzerinden derecelendirilmesinin istendiği bir işlem sürecidir (Colman, 2015). Psiko-fizikte temel problem olan psikolojik büyüklüklerin ölçülmesi, psiko-fizik dalının kendisi kadar eski bir problem olup, 1800'lerin ortalarında Fechner tarafından önerilen ve daha sonra Fechner Yasası olarak bilinen çalışma ile başlar. Ardından Stevens Güç Yasası ile psikolojik büyüklükler ile uyarıcının yoğunluğu arasındaki ilişkiyi açıklayan görgül bir yasa 1950'lerde Stevens tarafından önerilmiştir (Bolanowski ve Gescheider, 2013).

Bu araştırmada kullanılan mutlak yargılar yöntemi, sınıflama yargıları yasası altında geliştirilen, ardışık aralıklar yönteminin ilk uygulamalarından sayılır ve Guilford'un mutlak ölçekleme yöntemi olarak bilinir (Turgut ve Baykul, 1992). Bu yöntem eşit görünümlü aralıklar yönteminden elde edilen verilere, dereceli ölçeklerden elde edilen verilere ve tek uyarıcı yönteminden elde edilen verilere uygulanabileceğini belirtir (Guilford, 1938). Derecelendirme için iki türlü veri toplama formu hazırlanabilir. Birinde, uyarıcıları yargılamak üzere bireylere numara, rakam, kelime ya da kelime grubu üzerinden kategoriler sunulur; diğesinde, kategori sayısı hakkında

sınırlama olmadan, gram, desibel gibi fiziksel birimler kullanarak uyarıcı için uygun gördükleri değerler üzerinden dereceleme yaptırılabilir (Torgerson, 1958).

Önerilen yöntemlerin temel sıkıntısı, uyarıcı yoğunluğu ile psikolojik büyüklük değerleri arasındaki ilişkinin, bireylerin yargıları sonucu oluşmasıdır (Bolanowski ve Gescheider, 2013). Bu durum sonuçlarda yargı hatası oluşturmaktadır. Thurstone (1959) bu durumu, uyarıcı hatası problemi olarak ifade eder ve bu durumun, bireylerin uyarıcı ile ilgili geçmiş yaşantılarındaki yoğunluğa göre yargıda bulunmaları ile ilgili olduğunu belirtir (s. 186). Mutlak yargılar yönteminde, yargıda bulunacak katılımcılara somut başlangıç noktası ve birim verilmemesi, katılımcıların birer defa uyarıcılar hakkında yargıda bulunması durumu, her birey kendi zihnindeki başlangıç noktası ve birimi kullanacağından dolayı yargıcılar arasında farklılıkların doğmasına ve böylelikle hataların ortaya çıkmasına neden olur (Turgut ve Baykul, 1992).

Öztürk, Özdemir ve Gelbal (2011'den akt., Acar Güvendir ve Özer Özkan, 2013), 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. Yine Albayrak Sarı ve Gelbal (2015) tarafından yapılan çalışmada, yargıcı kararlarına göre iki ölçekleme yöntemi olan ikili karşılaştırma ve sıralama yargılarına dayalı ölçekleme sonuçları karşılaştırılmıştır. Sonuçta iki ölçekleme yaklaşımının benzer sonuçlar üretmediği; bu nedenle, ölçekleme çalışmalarında seçilecek olan ölçekleme yöntemine karar verilirken kullanılacak yöntemin sayıltularına dikkat edilmesi gerektiği ifade edilmiştir.

Bu çalışmada, öğrenci görüşlerine göre sınav türlerinin mutlak yargılar yöntemi ile ölçeklenmesi ve sınav türlerinin ölçek değerlerine göre öğrenci bakış açısıyla bir değerlendirme yapılması amaçlanmıştır. Böylelikle, sınav türleri hakkında teknik olarak bilinen özellikler yanında, öğrenciler açısından sınav türlerinin olumlu ve olumsuz yanlarının bilinmesi sağlanmış olacaktır. Bu durum, uygulamada sınav türlerinin yerini görebilmek ve uygun biçimde kullanılıp kullanılmadığına yönelik bir gösterge oluşturmak açısından önemli olabilir.

### **1.3. Problem Cümlesi**

Bu araştırmada birbirine bağlı iki problem cümlesi vardır:

1. Sınav türlerinin mutlak yargılar yöntemi ile ölçek değerleri nedir?
2. Sınav türlerinin ölçek değerleri üzerinden sınav türleri hakkında öğrenci görüşleri nelerdir?

### **1.4. Sınırlılıklar ve Sayılıtlar**

Sınav türleri hakkında hesaplanan ölçek değerleri ve sınav türlerine ait öğrenci görüşleri verinin elde edildiği araştırma grubu üzerinden yapılmıştır. Bu nedenle araştırmanın sınırlılığı, bulguların genellenememesidir. Araştırmanın sayılıtlarına bakılacak olursa, bunlardan ilki, katılımcıların tepkilerinin mutlak ölçek değerlerini verdiği kabulü; diğeri ise, verilen bir uyarıcının ölçekleme boyutundaki değeri, ardışık aralıklar yönteminde olduğu gibi, herhangi bir kategori derecesinin altında ise, uyarıcının o kategoride yer aldığı kabulüdür.

## **2. YÖNTEM**

### **2.1. Araştırmanın Türü**

Bu araştırmada nitel ve nicel araştırmalar birlikte kullanılmıştır. Verilerin sayısal formda olması durumunda araştırma nicel araştırma olarak adlandırılır (Trochim ve Donnelly, 2008). Bu nedenle sınav türlerinin ölçeklenmesi, nicel bir araştırma türü olarak adlandırılabilir. Aynı zamanda sınav türlerinin öğrenci bakış açısıyla değerlendirilmesi nitel bir çalışma olarak düşünülebilir. Nitel araştırmalar katılımcıların bakış açısından olayların ve durumların nasıl anlaşıldığı ya da algılandığı ile ilgilidir (Fraenkel ve Wallen, 2003). Tarama modeli var olan bir durumun, olduğu şekliyle, betimlenmesini amaçlayan bir araştırma yaklaşımıdır (Karasar, 2014). Var olan bilgiyi ortaya çıkarma açısından değerlendirildiğinde bu çalışma aynı zamanda bir tarama araştırmasıdır.

### **2.2. Çalışma Grubu**

Bu araştırmada verinin toplandığı grup 2014-2015 eğitim ve öğretim yılında İstanbul Üniversitesi Hasan Ali Yücel Eğitim Fakültesinde Pedagojik Formasyon eğitimine katılan ve farklı bölümlerde okuyan ya da mezun olmuş öğrencilerdir. Bu çalışmaya gönüllülük esasına dayalı olarak 243 öğrenci katılmıştır. 243 öğrenciden 4'ü bölümünü belirtmediğinden dolayı, verisi bulunan 239 öğrencinin bölümlerine

göre dağılımları Tablo 1’de görülmektedir. Katılımcıların %80’i kadın, %20’si ise erkektir.

Tablo 1  
*Öğrencilerin Bölümlere Göre Dağılımı*

<b>Bölümler</b>	<b>f</b>	<b>%</b>
<b>Türk Dili ve Edebiyatı</b>	67	28.03
<b>Fizik</b>	48	20.08
<b>Biyoloji</b>	47	19.67
<b>İlahiyat</b>	45	18.83
<b>Tarih</b>	25	10.46
<b>Arapça</b>	7	2.93

### 2.3. Veri Toplama Araçları

Bu araştırma için kullanılan veri toplama aracı üç kısımdan oluşmaktadır. Birinci kısımda demografik bilgiler bulunmaktadır. İkinci kısımda ise, altı adet sınav türünün bulunduğu ve bunları “çok kötü, kötü, orta, iyi, çok iyi” olacak şekilde sınıflandırmayı sağlayan bölüm bulunmaktadır. Diğer kısımda ise sınıflandırmalara dair, katılımcıların sınav türlerine verdikleri tepkilerin gerekçelerini yazabilecekleri açık uçlu alan bulunmaktadır.

Bu çalışmada, alanyazında sunulan sınav türlerinden sözlü sınav (SS), uzun cevaplı yazılı sınav (UCYS), kısa cevaplı yazılı sınav (KCYS), boşluk doldurmalı yazılı sınav (BDYS), doğru-yanlış sınıflandırmalı sınav (D-YS) ve çoktan seçmeli sınavlar (ÇSS) incelenmektedir.

Sözlü sınavlar uygulaması, İlköğretim Kurumları Yönetmeliği (Milli Eğitim Bakanlığı [MEB], 2006) ve Orta Öğretim Kurumları Yönetmeliği ile (MEB, 2013) Milli Eğitim Bakanlığı’nın okullarında kaldırılmıştır. Ancak, bu sınav türü katılımcıların yaş grubu düşünülerek, sözlü sınav ile karşılaştıkları varsayılarak çalışmaya eklenmiştir.

### 2.4. Verilerin Analizi

Birinci araştırma sorusunu cevaplamak üzere, mutlak yargılar yöntemi kullanılmıştır ve Excel üzerinde hesaplanmıştır. İlk olarak 243 öğrencinin sınav türlerini eşit aralıklı ardışık sınıflar üzerinden verdiği tepkiler, uyarıcı×sınıflar olmak üzere satır ve sütunlarından oluşan frekans matrisine dönüştürülür. Her uyarıcı için frekans değerleri, ardışık sınıflar için verilen dereceler ile çarpılır (çok kötü:1, kötü:2,

orta:3, iyi:4, çok iyi:5) ve satır toplamı katılımcı sayısına bölünerek, uyarıcıların ölçek değerleri hesaplanır. Ayrıca her bir uyarıcı için, yargıdaki değişkenliklerini saptamak üzere, standart sapmalar hesaplanır.

İkinci araştırma sorusunu cevaplamak üzere içerik analizi kullanılmıştır. İçerik analizi, yazılı metin ya da belgelerin hem içeriğine hem de bağlamına göre analiz edilmesidir (Berelson, 1952 ve Robson, 2002'den akt., Spencer, Ritchie ve O'connor, 2003). İlk olarak sınav türleri ölçek değerlerine göre hangi sınıflara ait olduğu saptanmıştır. Öğrencilerin her sınav türü için ortaya çıkan çok iyi, iyi, orta, kötü ya da çok kötü olarak tanımlanan sınıflardan, sınav türünü o sınıfta değerlendirmelerinin gerekçeleri kodlanmıştır. İçerik ve bağlamına göre yapılan kodlama üzerinden kategori başlıkları oluşturulmuş ve bu kategori başlıkları üzerinde frekanslar saptanarak bulgular elde edilmiştir.

### 3. BULGULAR

Bu çalışmada öğrenciler sınav türlerini, yargıcı olarak, çok kötüden çok iyiye doğru beş kategoride değerlendirmişlerdir. 243 öğrencinin altı sınav türüne verdikleri dereceler üzerinden mutlak yargılar yöntemi ile yapılan ölçekleme sonuçları Tablo 2'de görülmektedir.

Tablo 2

*Sınav Türlerinin Ölçek Değerleri ve Standart Sapmaları*

Sınav Türü	Ölçek Değeri	Standart Sapma
ÇSS	4.02	0.81
KCYS	3.69	1.03
D-YS	3.44	1.11
BDYS	3.34	1.34
UCYS	2.93	1.33
SS	2.74	1.3

Buna göre 4.02 değeri ile ÇSS türü diğerlerine göre yüksek derecede ölçeklenmiştir. Ayrıca diğerlerine göre standart sapması (0.81) düşük olduğu için ölçek değeri daha kararlı bir noktadadır. Yüksek dereceden başlamak üzere ölçek değerlerine göre ÇSS'nin ardından KCYS, D-YS, BDYS, UCYS ve son olarak SS türü gelmektedir.

Birinci araştırma sorusu üzerinden elde edilen ölçek değerlerine bakıldığında, ÇSS türü, araştırmaya konu olan diğer sınav türlerine göre 5 derece üzerinden 4.02 ile birinci sırada gelmektedir.



Öğrencilerin bu sınav türünü seçmelerinin gerekçeleri incelendiğinde en uygun kategori başlıkları belirlenmiştir (bkz. Tablo 3).

Tablo 3

*ÇSS Türü Tercihlerinde Öğrenci Gerekçeleri ve Frekansları*

<b>Kategori Başlığı</b>	<b>Gerekçeler</b>	<b>f</b>
<b>İpucu</b>	Bilginin hatırlanmasını sağlar.	17
	<b>Toplam</b>	<b>17</b>
<b>Öğrenme odaklılık</b>	Bilginin yorumlanmasına önem verir/yorumlama gücü kazandırır.	3
	Öğrencileri seçenekleri karşılaştırmaya ve düşünmeye sevk eder.	3
	Hatırlama ve akıl yürütme birlikte kullanılır.	3
	Eleme yöntemi ile doğru cevaba ulaşılabilir.	7
	<b>Toplam</b>	<b>16</b>
<b>Amaca uygunluk</b>	Soru ve cevap belirgindir.	1
	Öğrencinin öğrenmelerini ortaya çıkarabilir.	2
	Kapsamı geniş bir sınavdır.	5
	Bilmeyenin doğru yapma olasılığı düşüktür.	8
	<b>Toplam</b>	<b>16</b>
<b>Kullanışlılık</b>	Cevaplama daha kolaydır/pratik bir sınavdır.	7
	<b>Toplam</b>	<b>7</b>
<b>Ezber odaklılık</b>	Çalışırken ezber yapmayı gerektirmez.	5
	<b>Toplam</b>	<b>5</b>
<b>Şans başarısı</b>	Bilinmese de cevap tahmin edilebilir.	1
	Şansla doğru cevap bulunabilir.	3
	<b>Toplam</b>	<b>4</b>
<b>Puanlama güvenilirliği</b>	Daha adaletlidir/objektif puanlanabilir.	4
	<b>Toplam</b>	<b>4</b>
<b>Sınav kaygısı</b>	Daha az stres yaratan bir sınavdır.	2
	<b>Toplam</b>	<b>2</b>

En çok yorum alan kategori başlıkları, sırasıyla ipucu, öğrenme odaklılık, amaca uygunluk, kullanışlılık, ezber odaklılık, şans başarısı, puanlama güvenilirliği ve sınav kaygısı olduğu görülmektedir. Cevabın hatırlanmasını kolaylaştırması, gerekçe belirten 69 öğrenci arasında en çok tekrarlanan gerekçedir. Bunu sınavın öğrenmeyi sağlama ve düşünme odaklı olduğuna dair atıf yaptıkları, öğrenme odaklılık ve öğrenmeyi ölçebildiğine yapılan atıfla, amaca uygunluk kategorileri takip etmektedir. Ayrıca 7 öğrenci bu sınav türünün ezber

gerektirmediğini, 4 öğrenci şans başarısı, 4 öğrenci puanlama güvenilirliği açısından ÇSS için olumlu değerlendirme yapmış ve 2 öğrenci ise sınav kaygısının düşük olduğunu belirtmiştir.

KCYS türünün, ölçek değeri açısından ikinci sırada tercih edilen bir sınav türü olduğu görülmektedir (bkz. Tablo 2). Öğrenci bakış açısıyla bu sınav türünü 3.69 ile “iyi” olarak değerlendirenlerin gerekçeleri incelendiğinde, ana başlıkların geçerlik, kullanılışlık, puanlama güvenilirliği, ezber odaklılık, öğrenme odaklılık olarak adlandırılabilenleri görülmektedir (bkz. Tablo 4).

Tablo 4

*KCYS Türü Tercihlerinde Öğrenci Gerekçeleri ve Frekansları*

<b>Kategori Başlığı</b>	<b>Gerekçeler</b>	<b>f</b>
<b>Amaca uygunluk</b>	Her şey değil ama bazı sorular için kısa cevap yeterli olmayabilir.	6
	Bilgiyi ölçmek için gerekli/idealdir.	11
	Öğrenip öğrenmediğin ortaya çıkar, bilmiyorsan cevaplandıramazsın.	12
	İstenen cevapların kısa, öz ve net olması açısından iyidir.	16
	<b>Toplam</b>	<b>45</b>
<b>Kullanışlılık</b>	Zaman açısından kullanılışlıdır, kısa sürede cevaplanabilir.	4
	Cevaplama açısından öğrenciye kolaylık sağlar.	12
	<b>Toplam</b>	<b>16</b>
<b>Puanlama güvenilirliği</b>	Puanlamada objektiflik sağlar.	3
	<b>Toplam</b>	<b>3</b>
<b>Ezber odaklılık</b>	Ezber gerektirir.	2
	<b>Toplam</b>	<b>2</b>
<b>Öğrenme odaklılık</b>	Öğrenciyi ders çalışmaya sevk eder.	1
	<b>Toplam</b>	<b>1</b>

KCYS türüne 67 öğrenci gerekçe bildirmiştir. Buna göre 45 öğrenci geçerlik başlığı altında toplanabilecek gerekçeler belirtmiştir. Bunun içinde yalnızca 6 kişi kısa cevabın yeterli olmayabileceği çekincesi koymuş, diğer 39 öğrenci öğrenmeyi ölçebileceğini vurgulamıştır. Aynı zamanda 16 öğrenci zaman ve cevaplama açısından kullanılışlı bir

sınav olduğunu belirtmiştir. Bununla birlikte dezavantaj olarak bu sınav türünün ezber gerektirdiği 2 öğrenci tarafından belirtilmiştir. Cevapların doğru ve yanlış olarak iki kategoride belirlendiği sorulardan oluşan D-YS türünün, öğrenciler tarafından üçüncü sırada tercih edildiği görülmektedir (bkz. Tablo 2).

Tablo 5

*D-YS Türü Tercihlerinde Öğrenci Gerekçeleri ve Frekansları*

<b>Kategori Başlığı</b>	<b>Gerekçeler</b>	<b>f</b>
<b>Amaca uygunluk</b>	Bilgiyi ölçmede iyi bir sınavdır.	5
	Öğrenenleri kesin belirler/öğrenmeyi ölçmede iyi bir sınavdır.	10
	<b>Toplam</b>	<b>15</b>
<b>Kullanışlılık</b>	Kolay/pratik bir sınavdır.	12
	<b>Toplam</b>	<b>12</b>
<b>Şans başarısı</b>	%50 şans faktörü olduğu için iyidir.	6
	<b>Toplam</b>	<b>6</b>
<b>İpucu</b>	Bilginin hatırlanmasını sağlar.	6
	<b>Toplam</b>	<b>6</b>
<b>Sınav kaygısı</b>	İki seçenek arasında kalma sıkıntı yaratır.	5
	<b>Toplam</b>	<b>5</b>
<b>Öğrenme odaklılık</b>	Öğrenmeyi sağladığı için iyi bir sınavdır.	2
	<b>Toplam</b>	<b>2</b>
<b>Puanlama güvenilirliği</b>	Puanlaması objektiftir.	2
	<b>Toplam</b>	<b>2</b>

Bu sınav türü de KCYS gibi “iyi” kategorisinde yer almaktadır ve öğrencilerin bu kategoride değerlendirilmesinin gerekçeleri incelendiğinde, 44 öğrenciden 15’i isteneni ölçebilen, 12’si pratik, 12’si ise şans başarısının yüksek ve hatırlamayı sağlaması açısından iyi olduğunu belirtmiştir (bkz. Tablo 5).

Boşluk doldurma soru tipinden oluşan sınavlar, kısa cevaplı sınavlar kategorisine girse de bu sınavlar birbirinden farklı ölçek değerleri almıştır. BDYS türü sınavların ölçek değeri bu araştırmada 3,34 olarak belirlenmiştir (bkz. Tablo 2).

Tablo 6

*BDYS Türü Tercihlerinde Öğrenci Gerekçeleri ve Frekansları*

<b>Kategori Başlığı</b>	<b>Gerekçeler</b>	<b>f</b>
<b>İpucu</b>	Bilmeyen de ipuçlarından yararlanarak boşluğu doğru doldurabilir.	4

	Cevaplar hatırlanabilir olduğu için kolay bir sınavdır.	22
<b>Amaca uygunluk</b>	<b>Toplam</b>	<b>26</b>
	Kavramların öğrenilip öğrenilmediğini iyi ölçer.	3
	Anlaşılır sorulardır, cevaplar kısa ve nettir.	7
	Bilgi sorularında etkilidir.	16
<b>Ezber odaklılık</b>	<b>Toplam</b>	<b>26</b>
	Ezber gerektirir.	2
<b>Puanlama</b>	<b>Toplam</b>	<b>2</b>
<b>güvenilirliği</b>	Puanlama objektiftir.	2
<b>Sınav kaygısı</b>	<b>Toplam</b>	<b>2</b>
	Az stres yaratan bir sınavdır.	1
	<b>Toplam</b>	<b>1</b>

Buna göre öğrencilerin bu sınav türüne “iyi” düzeyinde değer vermelerinin gerekçeleri incelenmiştir. 57 öğrencinin gerekçelerine göre çoğunlukta olan, soru tipinin doğru cevaba ipucu vermesi (26) ve ölçmede etkili (26) bir sınav türü olduğunu ortaya çıkmaktadır (bkz. Tablo 6).

Uzun cevap gerektiren yazılı sınavların ölçek değeri 2.93’tür ve altı sınav türü arasında 5. sırada tercih edilen sınav olarak “orta” düzey kategorisinde yer almıştır (bkz. Tablo 2).

Tablo 7

*UCYS Türü Tercihlerinde Öğrenci Gerekçeleri ve Frekansları*

<b>Kategori Başlığı</b>	<b>Gerekçeler</b>	<b>f</b>
<b>Amaca uygunluk</b>	Konunun derinlemesine öğrenilip öğrenilmediğini ölçer.	1
	Herkes düşündüklerini uzunca yazıya dökemez.	2
	Bilgilerin bütünü ölçmek için iyi bir sınavdır.	2
	Soru sayısı az olunca az sayıda bilgi ölçer.	3
	Bilgi ölçer ise kötü, yorum ölçer ise iyi bir sınavdır.	8
	Cevapların uzun istenmesi iyi değil/uzatmak konuyu dağıtır/yanlış şeyler yazılabilir/yazarken tekrara düşülür.	11
	<b>Toplam</b>	<b>27</b>
<b>Sınav kaygısı</b>	Sınav kaygısı yüksek oluyor.	1

	Öğretmenin beklentisini karşılayamama kaygısı var.	1
	Süre yetiştirebilir.	11
	<b>Toplam</b>	<b>13</b>
<b>Ezber odaklılık</b>	Ezbercilik gerektirir.	6
	Öğrendiklerinin tümünü hatırlayamama sorun yaratır.	2
	<b>Toplam</b>	<b>8</b>
<b>Öğrenme odaklılık</b>	Öğrenciyi araştırmaya yönelir.	1
	<b>Toplam</b>	<b>1</b>
<b>Kullanışlılık</b>	Yorucu, pratik olmayan bir sınavdır.	1
	<b>Toplam</b>	<b>1</b>

Tepki veren 50 öğrenciden UCYS türünü orta düzeyde belirlemelerinin gerekçeleri incelendiğinde, amaca uygunluk kategori başlığında toplanabilecek 27 gerekçeden 24'ünde istenenin ölçülemeyeceği çekincesi görülmektedir. Bunlardan 8'i ancak yorum istenen durumlarda iyi bir sınav olabileceğini belirtmektedir (bkz. Tablo 7). Sınav kaygısının da bu sınav türünde belirginleştiği söylenebilir.

SS türü, en son tercih edilen sınav türüdür ve ölçek değeri 2.74'tür (bkz. Tablo 2). Buna göre sınav türünü "orta" düzeyde değerlendiren öğrencilerin gerekçeleri Tablo 8'de görülmektedir.

Tablo 8

*SS Türü Tercihlerinde Öğrenci Gerekçeleri ve Frekansları*

<b>Kategori Başlığı</b>	<b>Gerekçeler</b>	<b>f</b>
<b>Amaca uygunluk</b>	Her konuyu ölçemez.	1
	Kişinin kendini ifade edebilmesi açısından iyidir ancak heyecan yaratabilir.	3
	Heyecan ve kaygı nedeniyle unutma olur, kendini iyi ifade edemezsin.	56
	<b>Toplam</b>	<b>60</b>
<b>Öğrenme etkinliği</b>	Öğrencinin öğrenip öğrenmediği kontrol edilebilir.	3
	Eksik bilgilerin denetim ve takviyesi için ders içinde yapılabilir.	1
	<b>Toplam</b>	<b>4</b>
<b>Sınav kaygısı</b>	Verilen cevap yüzünden rencide olma kaygısı vardır.	1
	Öğrenciyi gereksiz yere strese sokar.	1

	<b>Toplam</b>	<b>2</b>
<b>İletişim</b>	Öğrenci öğretmen iletişimi açısından iyi olabilir.	2
	<b>Toplam</b>	<b>2</b>
<b>Puanlama</b>	Objektif puanlama yapılamaz.	1
<b>güvenirligi</b>	<b>Toplam</b>	<b>1</b>

SS türünde “amaca uygunluk” başlığı altında toplanan ve gerekçesini belirten 69 öğrenciden 60’ı sınav esnasında heyecanlanma nedeniyle istenenin ölçülemeyeceği vurgusu yapmıştır. Ayrıca bu türden canlı soru-cevap uygulamasının ders işlenirken kullanılmasının uygun olacağını belirten ifadeler, öğrenme etkinliği kategori başlığında toplanabilir.

Bir bütün olarak Tablo 9 üzerinden, öğrencilerin sınav türlerine yönelik değerlendirmeleri kategoriler bazında incelenmektedir.

Tablo 9

*Sınav Türlerine Göre Öğrenci Yargılarından Oluşan Kategori Başlıklarının Yüzdeleri*

	<b>ÇSS</b>	<b>KCYS</b>	<b>D-YS</b>	<b>BDYS</b>	<b>UCYS</b>	<b>SS</b>
<b>Amaca uygunluk</b>	23.2 (+)	67.2 (+)	34 (+)	45.6 (+)	6 (+)	87 (-)
					16 (+/-)	
					32 (-)	
<b>Öğrenme odaklılık</b>	23.2 (+)	1.5 (+)	4.5 (+)	-	2 (+)	
<b>İpucu</b>	24.6 (+)	-	13.6 (+)	45.6 (+)	-	
<b>Kullanışlılık</b>	10.1 (+)	23.9 (+)	27.3 (+)	-	2 (-)	
<b>Ezber odaklılık</b>	7.3 (-)	3 (+)	-	3.5 (+)	8 (+)	
<b>Şans başarısı</b>	5.8 (+)	-	13.6 (+)	-	-	
<b>Puanlama güvenirligi</b>	5.8 (+)	4.5 (+)	4.5 (+)	3. (+)	-	1.5 (-)
<b>Sınav kaygısı</b>	2.9 (-)	-	11.4 (+)	1.8 (-)	26 (+)	
<b>Öğrenme etkinliği</b>	-	-	-	-	-	5.8 (+)

<b>İletişim</b>	-	-	-	-	-	2.9 (+)
-----------------	---	---	---	---	---	------------

Ölçme amacına uygun olarak olumlu (+) gözlenen en iyi sınav türü KCYS'dir (%67.2). UCYS için amaca uygunluk %6'da kalmıştır. UCYS için %16'sı "Bilgi ölçer ise amaca uygun değil, yorum gücü ölçer ise amaca uygundur." ifadeleri ile koşul belirtmiştir. Öğrencilerin %32'si ise UCYS türünü amaca uygunluk açısından olumsuz (-) olarak değerlendirmiştir. Ancak bu kategoride en olumsuz görüş SS türü içindir (%87).

Öğrenmeyi sağlaması açısından en iyi ÇSS türü, öğrenciler tarafından değerlendirilmiştir. Cevaplama ipucu veren sınav olarak birinci sırada BDYS (%45.6), ikinci sırada ÇSS (%24.6) gelmektedir. Kullanışlılık açısından tercih edilen sınavlar ise, sırasıyla, D-YS (%27.3), KCYS (%33.9) ve ÇSS (%10.1) türleridir. Ezber yapmaya yönelen sınav ise UCYS (%8); tersine, çalışırken ezber yapmayı gerektirmeyen sınav olarak ÇSS (%7.3) olduğu görülmektedir. Şans başarısı ile puan elde etme en çok D-YS (%13.6) türünde ve ÇSS (%5.8) türündedir. Puanlamanın objektifliği sınav değerlendirmesinde çok fazla gerekçeler arasında yer almamakla birlikte SS türü için objektif puanlamanın olmadığı kanaati vurgulanmıştır (%1.5). Sınav kaygısı yaratan sınav türü UCYS türüdür (%26). SS türü için iki yeni kategori açılmıştır. Öğrenme etkinliği ve iletişim bu sınav türü için bir avantaj olarak öğrenciler tarafından görülmektedir (bkz. Tablo 9).

#### 4. TARTIŞMA VE SONUÇ

Bu araştırmada amaç, öğrencilerin bakış açısından sınav türlerinin derecelenmesi ve ortaya çıkan derecelere göre, olumlu ve olumsuz yönleriyle sınav türlerinin, yine öğrenci bakış açısıyla bir değerlendirmesini ortaya koymaktır.

İlk olarak öğrencilerin sınav türlerine beş derece üzerinden verdikleri yargılar veri olarak kullanılmış ve mutlak yargılar yöntemi ile sınav türlerinin ölçek değerleri elde edilmiştir. ÇSS türü ölçek değeri olarak aralarında en yüksek ölçek değerine sahip sınav türüdür ve arkasından KCYS gelmektedir. Sırasıyla, D-YS ve BDYS oldukça birbirine yakın ölçeklenmiştir. UCYS ve son olarak SS türü sınavlar yine birbirlerine yakın ve sıralamada arkalara yerleşen sınav türleridir. Tezbaşaran (2015) tarafından ikili karşılaştırma yöntemi ile yapılan araştırmada, benzer sonuçların elde edildiği görülmektedir. Özellikle ÇSS, UCYS ve SS türleri, sıralama açısından aynı sonuçları vermektedir.

Bu araştırmanın önemli bir diğer sonucu ise, sınav türlerine öğrenciler hangi açılardan yaklaşmakta ve onları ne tür gerekçelerle olumlu ya da olumsuz eleştirmekte olduklarının saptanmış olmasıdır. Bu çalışmada sınav türleri, kategori başlıklarında incelenmiştir. Kategorilerden oluşan üst başlıkların “Güvenirlilik”, “Geçerlik”, “Kullanışlılık” ve “Öğrenme” temaları olabileceği düşünülmektedir. Ancak öğrencilerin yaptığı değerlendirmelerin kategori başlıkları altında incelenmesinin, tartışma açısından daha anlamlı olduğu düşünülmüştür.

Sınav türlerinden ÇSS’nin, üst düzey öğrenmeleri ölçebilmesi, objektif puanlanabilmesi gibi özellikleri öğrenciler tarafından vurgulanmakla birlikte, cevabın ipucu ile yakalanabilmesi ve şansla doğru cevabın bulunabilmesi nedeni ile bu sınav türünün öğrenci tarafından avantaj olarak görülmesi, ÇSS türünün öğretmenler tarafından çok dikkatli olarak hazırlanması gerektiğini göstermektedir. Cevabın, sorunun altında, öğrencinin gözü önünde olması nedeniyle, iyi hazırlanmamış çeldiricilerden oluşan sorular, ÇSS’nin geçerliğini zedelemektedir. Cevabın bilmeyen tarafından yakalanabilmesi, sınavı hazırlayan tarafından çoktan seçmeli soru yazmada önemli kuralların yerine getirilmediği anlamına gelmektedir. Hatta iyi bir soru kökü, çalışmayan çeldiriciler yüzünden geçersiz bir soru haline dönüşmektedir. Hazırlanan ve uygulaması yapılan çoktan seçmeli bir testin maddeleri analiz edilerek, testin yeterince güvenilir ve geçerli olup olmadığı, testi oluşturan maddelerin iyi çalışıp çalışmadığı, bilenler ile bilmeyenleri birbirinden ayırıp ayırmadığı, güçlüğü ya da kolaylığı gibi özellikleri belirlenebilir (Gelbal, 2013). Çoktan seçmeli maddelerin analiz edilebilmesi, öğretmene doğru çalışmayan çeldiricileri düzelterek sorunun kalitesini iyileştirme fırsatı verir (Burton, Sudweeks, Merrill ve Wood, 1991).

KCYS türü özellikle amaca uygunluk ve kullanışlılık açısından, öğrenciler tarafından, iyi bir sınav türü olarak değerlendirilmiştir. Özellikle bilgiyi ölçmek için uygun olması, soruların anlaşılır ve istenen cevabın kısa ve öz olması bu sınav türünün olumlu yanı olarak öğrenciler tarafından vurgulanmıştır. Bunun yanı sıra, zaman açısından ve istenen cevabın kısa olması durumunun da sınavın getirdiği bir rahatlık olarak değerlendirilmiştir. Cevapların kısa olması, yazılı yoklamalarda olan yoruma açıklığı azaltır (Gelbal, 2013).



D-YS türü ise amaca uygunluk ve kullanılabilirlik açısından, öğrenciler tarafından beğenilen bir sınav türü olmakla birlikte şans başarısının yüksek olması da bir avantaj olarak görülmektedir. Bundan dolayı, büyük sınıflardaki öğrencilere bu sınav türünün uygulanması, geçerlik açısından sakıncalı olabilir. Şans başarısının yüksek olmasına rağmen, bu sınav türünün kaygı yarattığı ifadeleri, özellikle ikilemde kalmanın stres yarattığının belirtilmesi, bu araştırmanın ortaya çıkardığı sonuçlardan bir tanesidir.

BDYS türü, aslında KCYS altında yer almaktadır. Ancak görülmektedir ki, boşluk doldurma ve kısa cevaplı soru tipi öğrenciler tarafından farklı algılanmakta ve derecelendirilmektedir. KCYS’ndan farklı olarak bu sınav türünde cevaba götürecek ipuçlarının bulunması bu sınav türünü “iyi” kategorisinde değerlendirmelerinin bir gerekçesini oluşturmaktadır. Sınav sorularında beklenen özelliğe sahip olanın doğru cevaplama beklentisi vardır ki, buna madde ayırt edicilik gücü ya da madde geçerliği denir (Atılğan, 2006). O halde ipucu veren sorulardan oluşan bir sınavın geçerliği de düşük olur.

UCYS türü sınavlar amaca uygunluk açısından çekincelerin belirtildiği bir sınavdır. Özellikle yazılı anlatım gücü bu sınav türünde geçerliği zedeleyici bir etmendir (Özçelik, 2009; Turgut ve Baykul, 2010). Bu durum öğrenciler tarafından da ifade edilmiştir. Ayrıca bu sınavın kaygı yaratan bir sınav olduğu belirtilmiştir. Yine kapsam geçerliği açısından sıkıntılı bir sınavdır. Bununla birlikte, soruların açık uçlu olması nedeniyle öğretmenin istediği cevabın ne olduğunun anlaşılabilmesi ise düşük not alma kaygısı yaratabilir. Uzun cevabı olan soruların uygulanması, sınavın başta kullanılabilirliğini sonra da güvenilirlik ve geçerliğini düşürür. Az sayıda uzun cevaplı soru ile sınav kapsamı yeterince temsil edilemez. Bu nedenle uzun cevaplı sorular daha küçük alt sorulara ayrılarak kısa cevaplı hale dönüştürülebilir (Gelbal, 2013).

SS türü için amaca uygun olmayan bir sınav türü olarak değerlendirenlerin oranı oldukça fazladır. Bu kategorideki gerekçeler incelendiğinde, sınavın heyecan yarattığı vurgusu yapılarak amaca uygun olmadığı gerekçelendirilmiştir. Bu durumda sınav kaygısı ve amaca uygunluğun birlikte ele alındığı söylenebilir. Bu sınav türü ancak sözlü anlatım gücü, konuşma becerileri ölçülecek ise geçerli bir sınavdır (Özçelik, 2009).

Öğrencilerin ortaya koyduğu sınav türlerine yönelik açıklayıcı tepkiler, alanyazında bu sınavların olumlu ve olumsuz yönleri ile

örtüşmektedir. Özçelik (2009), kısa cevaplı yazılı sınavların en kullanışlı sınav olduğunu ve özellikle okullardaki derslerde kullanılmak üzere önceden hazırlanmış ve geliştirilmiş çoktan seçmeli testler yok ise, kısa cevaplı sınavlar hazırlanarak bunlardan yararlanılmasını önerir (s. 129).

Öğrenci tepkileri ile alanyazındaki bilgilerin bu denli uyuşması, öğrencilerin eğitim yaşantılarında sınav türlerinin olumlu ve olumsuz yanlarını deneyimlemiş olduklarını düşündürmektedir. Öğrencilerin hatırlamayı sağlayan sınav türlerini tercih etmeleri, geçmiş yaşantılarında daha çok hatırlama gerektiren bilgi basamağında hazırlanmış sınavlarla karşılaştıklarına dair bir çıkarım yapılabilir. Oysaki daha üst düzeyden soruların da sınavlarda yer alması gerekir. Öğrenme düzeyleri hiyerarşiktir; buna göre, bilgiyi kavramış biri onu aynı zamanda hatırlayabilir de (Bloom, Engelhart, Furst, Hill ve Krathwohl, 1956). Ayrıca şans başarısı gibi soru tipinden kaynaklı, bilmeden doğru cevaba ulaşabilme olasılığı bulunan sınavların, kendi avantajlarına olacağını düşündükleri için, o sınav türünü tercih ettikleri görülmektedir. Bu nedenle, eğitim veren tüm öğretmen ya da ölçme yapacak olan tüm eğitimcilerin, sınavların özelliklerini bilmeleri ve sınav hazırlamada kendilerini geliştirerek yetkin olmaları önemlidir. Sonuç olarak, eğitimde yapılan nitelikli ölçümler, “başarı” adına yapılan değerlendirmelerin sağlam olmasına katkıda bulunacağı gibi, öğrencilere sınav türlerinin olumsuzluklarının yaşatılmaması ve onların sınavlara ve ondan elde edilen puanlara karşı olan güvenlerinin zedelenmemesi açısından da önemlidir.

## 5. KAYNAKÇA

- Acar Güvendir, M. ve Özer Özkan, Y. (2013). İki ölçekleme yönteminin karşılaştırılması: İkili karşılaştırma ve sıralama yargıları. *Eğitim Bilimleri Araştırmaları Dergisi*, 3(1), 105-119.
- Albayrak Sarı, A. ve Gelbal, S. (2015). İkili karşılaştırmalar yargılarına ve sıralama yargılarına dayalı ölçekleme yaklaşımlarının karşılaştırılması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 6(1), 126-141.
- Atılğan, H. (2006). Madde ve test istatistikleri. H. Atılğan (Ed.), *Eğitimde ölçme ve değerlendirme* içinde (s. 353-375). Ankara: Anı Yayıncılık.
- Bloom, B. S, Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives*. London: Longman Group Ltd.
- Bolanowski, S. J., & Gescheider, G. A. (2013). *Ratio scaling of psychological magnitude*. Hove: Psychology Press.

- Burton, S. J., Sudweeks, R. R., Merrill, P. F., & Wood, B. (1991). *How to prepare better multiple-choice test items: Guidelines for university faculty*. Retrieved August 23, 2016, from <https://testing.byu.edu/handbooks/betteritems.pdf>
- Colman A. M. (2015). *A dictionary of psychology*. Oxford: OUP Oxford.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Florida: Holt, Rinehart and Winston, Inc.
- Doğan, N. (2006a). Yazılı yoklamalar. H. Atılğan (Ed.), *Eğitimde ölçme ve değerlendirme* içinde (s. 206-226). Ankara: Anı Yayıncılık.
- Doğan, N. (2006b). Kısa yanıtli sınavlar. H. Atılğan (Ed.), *Eğitimde ölçme ve değerlendirme* içinde (s. 248-255). Ankara: Anı Yayıncılık.
- Doğan, N. (2006c). Çoktan seçmeli testler. H. Atılğan (Ed.), *Eğitimde ölçme ve değerlendirme* içinde (s. 281-326). Ankara: Anı Yayıncılık.
- Ebel, R. L. (1965). *Measuring educational achievement*. New Jersey: Prentice Hall, Inc.
- Fraenkel, J. R., & Wallen, N. E. (2003). *How to design and evaluate research in education*. Boston: Mc Graw Hill.
- Gelbal, S. (2013). Eğitimde kullanılan ölçme araçları. S. Gelbal (Ed.), *Ölçme ve değerlendirme* içinde (s. 58-76). Eskişehir: Anadolu Üniversitesi.
- Guilford, J. P. (1938). The computation of psychological values from judgements in absolute categories. *Journal of Experimental Psychology*, 22(1), 32-42.
- Karasar, N. (2014). *Bilimsel araştırma yöntemi*. Ankara: Nobel Akademi.
- Milli Eğitim Bakanlığı (2006). Milli Eğitim Bakanlığı İlköğretim Kurumları Yönetmeliği. *Resmi Gazete*, 26156, 02.05.2006. [http://mebk12.meb.gov.tr/meb\\_iys\\_dosyalar/.../09043814\\_yonetmelik.doc](http://mebk12.meb.gov.tr/meb_iys_dosyalar/.../09043814_yonetmelik.doc) Erişim tarihi: 10.02.2015.
- Milli Eğitim Bakanlığı (2013). Milli Eğitim Bakanlığı Orta Öğretim Kurumları Yönetmeliği. *Resmi Gazete*, 28758, 07.09.2013. <http://www.resmigazete.gov.tr/eskiler/2013/09/20130907-4.htm> Erişim tarihi: 10.02.2015.
- Özçelik, D. A. (2009). *Ölçme ve değerlendirme*. Ankara: Pegem Akademi.
- Spencer, L., Ritchie, J., & O'conner, W. (2003). Analysis: Practices, principles and processes. In J. Ritchie, J. Lewis (Ed.) *Qualitative research practices* (pp. 199-218). London: Sage Publications.
- Tezbaşaran, E. (2015). Öğretmen adaylarının görüşlerine göre sınav türlerinin ikili karşılaştırma yöntemi ile ölçeklenmesi. *Eğitim ve Öğretim Araştırmaları Dergisi*, 4(3), s.283-288.
- Thurstone, L. L. (1959). *The Measurement of values*. Chicago: The University of Chicago Press.
- Torgerson, W. S. (1958). *Methods of scaling*. New York: John Wiley & Sons, Inc.

- Trochim, W. M. K., & Donnelly, J. P. (2008). *The research methods knowledge base*. USA: Cengage Learning.
- Turgut, M. F. ve Baykul, Y. (1992). *Ölçekleme teknikleri*. Ankara: ÖSYM Yayınları.
- Turgut, M. F. ve Baykul, Y. (2010). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem Akademi.