

## Yabancı Dil Eğitiminde Aynı Davranışları Yoklayan Çoktan Seçmeli ve Kısa Cevaplı İki Testin Madde ve Test Özelliklerinin Karşılaştırılması

### A Comparison of the Items and Specifications of Two English Tests in Multiple Choice and Short Answer Format Measuring the Same Behavior in Foreign Language Teaching

Adnan Kan ve Ulaş Kayapınar  
Mersin Üniversitesi

#### Öz

Bu araştırmanın amacı, aynı hedef davranışları ölçen çoktan seçmeli ve kısa cevaplı maddelerin ve bu maddelerden oluşan iki İngilizce testinin madde güçlük indeksleri, madde ayıt edicilik gücü indeksleri, test güvenilirlikleri, bir ölçüte dayalı geçerlik katsayıları ve test puanları ortalamaları arasında farklılıklar olup olmadığını incelemektir. 10 ayrı İngilizce hazırlık sınıfında öğrenimlerini sürdüren 139 öğrenciden elde edilen madde puanlarının yardımıyla hesaplanan madde ve test istatistikleri, bu iki testin madde güçlük indeksleri, madde ayırtıcılık gücü indeksleri ve test güvenilirlikleri arasında belirgin farklılıklar ortaya koyarken, test puanları ortalamaları arasında belirgin bir fark bulunmamıştır. Ayrıca, kısa cevaplı, test puanlarının, çoktan seçmeli test puanlarına nazaran final sınavı puanlarını daha iyi yordadığı saptanmıştır. *Anahtar Sözcükler:* Çoktan seçmeli test, kısa cevaplı test, madde ve test özellikleri, geçerlik, güvenilirlik.

#### Abstract

The purpose of the study is to determine the differences between two English tests constructed with multiple choice format and short answer format which measure the same behavior in accordance with their item difficulty indexes, item discrimination indexes, test reliability indexes, criterion related validity coefficients, and test score averages. Computed item and test statistics according to the item scores which were supplied from 139 students from 10 different English preparatory classes revealed that these two tests have clear differences related to their item difficulty indexes, item discrimination indexes, and test reliability indexes but there is not a distinct difference related to their test score averages. Moreover, it can be stated that the short answer format test scores predict the final exam scores better than the multiple choice format test scores.

*Key words:* Multiple choice test, short answer test, item and test properties, validity, and reliability.

#### Giriş

Bilim ve teknolojide meydana gelen sürekli ve hızlı gelişim toplumların da bu gelişime ayak uydurmasını zorunlu kılmaktadır. Gelişmiş toplumların oluşması için ise nitelikli bireyler gerekir ve nitelikli birey yetiştirme

çabalarının en önemli aracı eğitimidir. Var olan bilgilerin değişimi ve bilginin sürekli gelişimi, eğitim sisteminin de değişen koşullarla aynı hızda kendisini değiştirmesini ve yenilenmesini gerektirmektedir.

Eğitim genel anlamda, bireyin davranışlarında kendi yaşantıları yoluyla olumlu değişiklikler meydana getirme sürecidir (Ertürk, 1982: 12). Etkili olmak isteyen bir eğitim sistemi, sorunları zamanında ve sağlıklı tanımlar koyup çözmek zorundadır. Eğitimin hedefi, toplumların geleceğini sağlam temeller üzerine oturtma ışığında bireylere istedik davranışların kazandırılarak onların nitelikli birer insan haline getirilmesini sağla-

Yrd. Doç. Dr. Adnan Kan, Mersin Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Mersin Üniversitesi, Eğitim Fakültesi, İngilizce Öğretmenliği Mersin

Arş. Gör. Ulaş Kayapınar, Mersin Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Mersin Üniversitesi, Eğitim Fakültesi, İngilizce Öğretmenliği Mersin.

mak ise bu hizmetlerin etkin ve verimli bir şekilde artırılması öncelikli esastır. Bu durumun getirisi olarak eğitim sürecinin sonunda bireylere kazandırılması planlanmış olan davranış değişikliklerinin meydana gelmesi beklenir.

Bu sürecin en verimli şekilde sonuçlanmasını sağlamak için eğitimcilerin görevi, eğitim sürecinin temelini, işleyişini ve ürünlerini yakından ve sıkı bir şekilde izlemek ve denetim altına almaktır. Bu izleme ve denetim ise ölçme ve değerlendirmeyi gerektirir. Eğitim süreci içerisinde uygulanan programın sağlam olup olmadığı kanısına varma, öğretme ve öğrenme sürecinde kullanılan yöntemlerin ne derece etkili olduğunu saptama, öğrenme güçlüklerini ve eksikliklerini belirleme, öğrenci başarısını ve öğrenme düzeyini teşhis etme, istenilen davranışların kazanılıp kazanılmadığını ortaya koyma ve öğretme ve öğrenme sürecinin denetimi, ölçme ve değerlendirme hizmetinin iyi bir şekilde işlemesine dayanır (Baykul, 2000: 89).

Bu görevlerin yerine getirilmesi, istendik davranışların kazanılıp kazanılmadığının belirlenmesi ve kararların doğru bir şekilde alınabilmesi için ise yapılan ölçümlerin tutarlı olması beklenir ki bu da ölçme araçlarının kullanılması ile mümkün olabilir. Bir ölçme aracını yapılandırmak, temelde iki işlemi gerektirir; birincisi, ölçülecek davranışın ya da düzeyin belirlenmesi, diğeri ise ölçmenin yapılacağı aracın en iyi şekilde yapılandırılması ve ölçme sonuçlarının elde edilmesidir (Micheels & Karnes, 1950: 17).

Bu şekilde elde edilmiş bir ölçme aracının eğitim sürecinde rahatlıkla kullanılabilmesi için denenmiş ve yeterli geçerlik ve güvenilirlik ölçütlerini barındırıyor olması gerekir. Bu amaçla, ölçme aracı sadece ölçmek için düzenlendiği özelliği ölçmeli ve bu işi hatasız bir şekilde ya da hatalardan arınık olarak yapabilme yeterliğine sahip olmalıdır. Bu amaçla, kullanılacak ölçme aracının, ölçmek istenilen özelliği başka herhangi bir özelliklerle karıştırmadan, doğru olarak ölçüp ölçmediği tespit edilerek o ölçme aracının amacına hizmet etme niteliğine geçerlik adı verilir (Tekin, 2000: 42).

Ölçme aracının, belli zaman aralıklarıyla, aynı bireylere tekrar uygulanması durumunda, bu bireylerin o testten aldıkları puanların tutarlılığının ve birbirlerine yakınlığının göstergesi ise güvenilirlik olarak adlandırılır (Anastasi, 1982: 102).

Eğitim sürecinde kullanılan ölçme araçlarının bu niteliklere sahip olma gerekliliği ile birlikte ölçülecek özelliğin doğasına ve aracın kullanılacağı gruba göre çeşitlilik göstermesi de gerekir.

Öğrencilerin, eğitim ve öğretim süreci içerisinde öğrenme düzeylerini belirlemek ve bu düzeylere bağlı başarıyı ölçmek amacıyla geliştirilen başarı testleri; sözlü ve yazılı testler olarak, yazılı testler ise çoktan seçmeli ve kısa cevaplı testler olarak adlandırılabilir. Çoktan seçmeli testler genellikle üç, dört ya da beş alternatif cevap içeren sorulardan oluşur. Bu alternatiflerden bir tanesi doğru cevaptır, diğerleri ise çeldirici olarak adlandırılır (Keeves, 1988: 357). Çoktan seçmeli testler, başarı testleri arasında en yaygın kullanılanıdır. Kısa cevaplı testler ise öğrencinin bir sözcük, bir rakam, bir tarih ya da en çok bir cümle ile cevaplandırabileceği maddelerden oluşur (Tekin, 2000: 125). Kısa cevaplı testler de başarı testleri arasında oldukça yaygın olarak kullanılmaktadır.

Eğitim sürecinde bu soru tiplerinin nitelikli olduğu kanısıyla yaygın olarak kullanılması ve aynı öğrencilerin farklı tipteki aynı davranışları yoklayan sorulara verdikleri tepkilerin de farklılık göstermesi olasılığı düşünüülerek istendik davranışların ifade edilmesi ve gerçekleştirilmesinde soru tiplerinin önemi taşıyıp taşımadığı merak konusu olmuştur. Bu amaçla, aşağıdaki problem cümlesi ve alt problem cümlelerine cevap aranmaya çalışılmıştır.

### Problem

Aynı davranışları yoklayan çoktan seçmeli ve kısa cevaplı sorulardan oluşan iki testin madde ve test özellikleri arasında farklılıklar var mıdır?

### Alt Problemler

1. Aynı davranışları yoklayan çoktan seçmeli ve kısa cevaplı sorulardan oluşan iki testin güçlükleri arasında farklılık var mıdır?
2. Aynı davranışları yoklayan çoktan seçmeli ve kısa cevaplı sorulardan oluşan iki testin sorularının ayrımcılık güçleri nasıldır?
3. Aynı davranışları yoklayan çoktan seçmeli ve kısa cevaplı sorulardan oluşan iki testin test puanları ortalamaları arasında anlamlı bir farklılık var mıdır?

4. Aynı davranışları yoklayan çoktan seçmeli ve kısa cevaplı sorulardan oluşan iki testin test güvenilirlikleri arasında anlamlı bir farklılık var mıdır?
5. Aynı davranışları yoklayan çoktan seçmeli ve kısa cevaplı testten elde edilen puanlar, final sınavındaki başarıyı ne derece yordamaktadır?

#### İlgili Araştırmalar

Lancaster (1987), seçme gerektiren ve kısa cevaplı maddelerin madde güçlük ve madde ayırt edicilik gücü indekslerini ve bu tip maddelerden oluşturulan testlerin güvenilirliklerini karşılaştırmıştır. Kısa cevaplı maddelerin daha zor ve daha ayırt edici olmasının beklenmesine rağmen maddelerin güçlük ve ayırt edicilik gücü indekslerinde anlamlı bir fark bulunmamıştır.

Özalp (1992)'in çoktan seçmeli ve kısa cevaplı maddelerin psikometrik özelliklerinin karşılaştırmasını yaptığı çalışmasının analizleri sonucunda, kısa cevaplı ve çokta seçmeli dil yeteneği testlerinin KR-20 iç tutarlılık katsayıları, madde ayırt edicilik göstergelerinin ortanca değeri ve değişkenlikleri arasında anlamlı farklar bulunmazken, ortalama güçlük düzeyleri, davranışlar arası korelasyonları, faktör yapıları, uygulama süreleri ortalamaları anlamlı bulunmuştur.

Aydın (1993), seçme gerektiren ve kısa cevaplı test sorularının psikometrik özelliklerinin karşılaştırılmasını amaçlamış ve yapılan analizler sonucunda, kısa cevap gerektiren testin daha güç ve daha ayırt edici olduğunu görmüştür. Öte yandan, testlerin geçerlikleri arasında anlamlı bir fark bulunmazken, KR-20 güvenilirlikleri arasında kısa cevap gerektiren test lehine anlamlı bir fark bulunmuştur.

Uğurlu (1996), aynı davranışı ölçmeye yönelik kısa cevaplı ve çoktan seçmeli maddelerden oluşan testlerin psikometrik özelliklerinin incelenmesi konulu araştırmasında maddelerin ortalama güçlükleri arasındaki farkı z istatistiği yardımıyla test etmiş, bunun sonucunda kısa cevaplı test lehine anlamlı bir farklılık görülmüştür. Madde geçerlikleri arasındaki fark Fisher'in z dönüşümü ile test edilmiş ve kısa cevaplı test lehine anlamlı farklılık olduğu görülmüştür. Testlerin ortalama güçlük düzeyleri arasındaki fark ise t istatistiği ile test edilmiş anlamlı bir fark, bulunmamıştır. KR-20 güvenilirlikleri arasındaki fark Fisher'in z dönüşümü yardımıyla test

edilmiş, güvenilirlikler arasında anlamlı bir fark bulunmamıştır. Testlerin geçerlikleri arasındaki fark, Fisher'in z dönüşümü yardımıyla test edilmiş ve yine anlamlı bir fark bulunmamıştır.

Karaca (1996), seçme gerektiren, kısa cevaplı ve doğru-yanlış testlerinin madde ve test özelliklerinin karşılaştırılması konulu araştırmasında, testlerin madde güçlük indeksleri, madde ayırt edicilik gücü indeksleri ve madde güvenilirlikleri arasındaki farkı varyans analiziyle test etmiştir. Sonuç olarak, doğru-yanlış ve seçme gerektiren test maddelerinin güçlükleri daha yüksek bulunmuş, doğru-yanlış maddelerinin daha ayırt edici oldukları görülmüştür. Testlerin güvenilirlikleri KR-20 güvenilirlik katsayısı ile hesaplanmış ve aralarında anlamlı bir farklılık olmadığı görülmüştür.

#### Yöntem

Araştırmanın verileri, 2003-2004 eğitim-öğretim yılında Mersin Üniversitesi Yabancı Diller Yüksekokulu İngilizce hazırlık sınıflarında öğrenimlerini sürdüren 139 öğrenciye uygulanan ve orta öncesi (pre-intermediate) düzey sonunda İngilizce zaman ifadesinin bilgisini ölçme amacı ile öğrencilerin kendi öğretmenlerinin katkısıyla araştırmacılar tarafından geliştirilen, aynı davranışları yoklayan her biri 4 seçenekli 19 maddelik bir çoktan seçmeli test ve yine 19 maddelik bir kısa cevaplı test yardımıyla elde edilmiştir. Bu ölçme araçlarından kısa cevaplı testin geliştirilmesi, çoktan seçmeli test geliştirildikten sonra bu testin kısa cevaplı test formatına dönüştürülmesi ile oluşturulmuştur. Bu dönüştürme, uygulanan ölçme araçlarının sadece ölçülmek istenen aynı davranışları ölçmeleri için gerçekleştirilmiştir.

Araştırmada kullanılan testler (kısa cevaplı ve çoktan seçmeli), denenmiş ve madde ayırtıcılık güçleri yeterli düzeyde (0,30 ve 0,30'dan büyük) olduğu kanıtlanmış maddelerden oluşmaktadır. Araştırmada iki kategorili puanlama yönteminden yararlanılmıştır. İki kategorili puanlamada her bir maddeye verilen cevaplar, doğru cevap için "1"; yanlış, boş ya da birden fazla olan cevaplar için "0" puan verilmesi şeklinde puanlanmaktadır. Kısa cevaplı testin soruları üzerinde kısmi puanlama yapılmamış bunun yerine sorulara verilen cevaplar anahtara tamamen uygunsa 1, uygun değilse 0 puan vermek suretiyle puanlanmıştır.

Testin uygulanması iki aşamalı olarak yapılmıştır. Cevaplayıcılara öncelikle kısa cevaplı test, bu uygulamanın 21 gün sonrasında ise çoktan seçmeli test kendi dersliklerinde ve kendi öğretmenleri tarafından uygulandıktan sonra araştırmacılar tarafından puanlanmıştır.

#### Verilerin Analizi

Verilerin analizi ve kullanılan istatistiksel teknikler alt problemlere göre aşağıdaki gibidir.

1. alt probleme ilişkin olarak, madde güçlük indeksleri, aynı zamanda her bir maddeye ait başarı oranları olarak ele alınmış ve çoktan seçmeli test ve kısa cevaplı teste ait verilerden hesaplanan madde güçlük indeksi ( $P_j$ ) arasında fark olup olmadığı iki oran arasındaki farkı test etmeye yarayan z testi istatistiği ile belirlenmiştir.
2. alt probleme ilişkin olarak, her bir maddeye yönelik madde ayırıcılık güçleri Çift Serili Korelasyon Katsayısı kullanılarak hesaplanmıştır. Korelasyon katsayıları arasında anlamlı bir fark olup olmadığı z test istatistiği ile belirlenmiştir.
3. alt probleme ilişkin olarak, her iki teste ait güvenilirlik katsayıları (KR-20) hesaplanmış ve bu katsayılar Fischer'in Z dönüşümüne tabi tutulduktan sonra elde edilen Z değerleri arasında fark olup olmadığı Z testi ile test edilmiştir.
4. alt probleme ilişkin olarak, iki testin ortalamaları hesaplanarak aritmetik ortalamaları arasında fark olup olmadığı t testi ile test edilmiştir.
5. alt probleme ilişkin olarak, öğrencilerin final sınavı puanları ölçüt olarak alınmış ve her iki testten elde edilen puanların final sınavındaki başarıyı ne derece yordadığını belirlemek amacıyla Basit Doğrusal Regrasyon Analizi kullanılmıştır.

Her iki testin de gerekli betimsel istatistikleri ve test istatistiklerinin hesaplanmasında SPSS istatistik paket programından ve MS Excel Office programından yararlanılmıştır.

#### Bulgular

Madde güçlük indekslerinin kullanılan test ve madde tipine göre farklılık gösterip göstermediği z testi ile test edilmiş ve Tablo 1'de verilmiştir.

Tablo 1.

*Çoktan Seçmeli ve Kısa Cevaplı Testlerden Elde Edilen Madde Performanslarına ( $P_j$ ) İlişkin Z İstatistiği Tablosu*

Madde No	$P_{\text{çoktan seçmeli}}$	$P_{\text{kısa cevaplı}}$	Z değerleri
1	0,83	0,56	5,031**
2	0,24	0,04	4,995**
3	0,66	0,52	2,465*
4	0,60	0,47	2,184*
5	0,40	0,60	-3,429**
6	0,30	0,76	-8,671**
7	0,66	0,76	-1,866
8	0,44	0,61	-2,927**
9	0,60	0,71	-1,776
10	0,67	0,74	-1,319
11	0,10	0,01	3,527**
12	0,60	0,55	0,972
13	0,51	0,73	-3,941**
14	0,60	0,29	5,319**
15	0,18	0,01	4,852**
16	0,21	0,01	5,722**
17	0,51	0,26	4,466**
18	0,45	0,70	-4,387**
19	0,34	0,17	3,384**

Tablo 1 incelendiğinde, çoktan seçmeli teste ilişkin madde başarı oranlarının 0,10 ile 0,83 arasında, kısa cevaplı teste ilişkin madde başarı oranlarının ise 0,01 ile 0,76 arasında değerler aldığı gözlenmiştir. Ayrıca, kısa cevaplı teste ait sorulara ilişkin başarı oranları ile çoktan seçmeli teste ait sorulara ilişkin başarı oranları arasında 7, 9, 10 ve 12. sorular arasında anlamlı bir farklılık bulunmazken, diğer tüm sorularda 0,05 ve 0,01 düzeyinde anlamlı farklılıklar bulunmuştur. Bu durum, hazırlanan teste ilişkin soru tipinin, yani testin kısa cevaplı ya da çoktan seçmeli olmasının öğrencilerin sorulara ilişkin başarı oranlarını etkilediği şeklinde yorumlanabilir.

Hazırlanan testin madde tipinin madde ayırıcılık gücü indeksini nasıl etkilediğine ilişkin olarak iki farklı tipte testin (kısa cevaplı ve çoktan seçmeli) madde ayırıcılık gücü indeksleri Tablo 2'te verilmiş ve bu tablo üzerinden yorumlamalar yapılmıştır.

Tablo 2 incelendiğinde, kısa cevaplı testten hesaplanan madde ayırıcılık gücü indekslerinin çoktan seçmeli testten hesaplanan madde ayırıcılık gücü indekslerinden görece daha yüksek bir değere sahip olduğu söylenebilir. Ayrıca çoktan seçmeli teste ait madde ayırıcılık gücü indeksleriyle kısa cevaplı, teste ait madde

Tablo 2.

*Çoktan Seçmeli ve Kısa Cevaplı Testlerden Elde Edilen Madde Ayırıcılık Gücü İndeksleri Tablosu*

Madde No	rpb-çoktan seçmeli	rpb-kısa cevaplı	Z
1	0,25	0,51	2,54*
2	0,00	0,19	1,58
3	0,41	0,52	1,15
4	0,34	0,25	0,82
5	0,42	0,19	2,11*
6	0,36	0,55	1,99*
7	0,21	0,50	2,77**
8	0,28	0,43	1,42
9	0,32	0,51	1,90
10	0,28	0,60	3,34**
11	0,01	0,12	0,92
12	0,41	0,51	1,05
13	0,40	0,39	0,10
14	0,38	0,47	0,91
15	0,28	0,08	1,72
16	0,33	0,04	2,50*
17	0,46	0,46	0
18	0,36	0,64	3,14**
19	0,60	0,28	3,34**

\*p<0,05, \*\* p<0,01

ayırıcılık gücü indeksleri arasında 1, 5 ve 6. sorulara ait ayırıcılık gücü indeksleri arasında 1. ve 6. sorularda kısa cevaplı lehine 5. soruda ise çoktan seçmeli test lehine olmak üzere 0,05 düzeyinde anlamlı bir farklılık elde edilmiştir. 7, 10, 18 ve 19. soruların ayırt edicilik gücü indeksleri arasında ise 7, 10 ve 18. sorularda kısa cevaplı, 19. soruda ise çoktan seçmeli test lehine olmak üzere 0,01 düzeyinde istatistiksel olarak anlamlı bir fark vardır. Bu durum, madde ayırıcılık gücü baz alındığında, kısa cevaplı testlerin çoktan seçmeli testlere göre daha avantajlı olduğu söylenebilir. Bazı sorularda anlamlı farkın çoktan seçmeli test lehine çıkmasının çeldiricilerle ilgili olduğu düşünülmektedir. Bu düşünceden hareketle çeldiricilerin iyi düşünülerek hazırlanmasının, bu testin gücünü artıracığı söylenebilir. Tüm bu sonuçlar ışığında, testte kullanılacak madde tipinin, madde ayırıcılık gücü indeksini de etkilediği ve çoktan seçmeli testlerin eğer çeldirici seçenekler iyi oluşturulabilirse, diğer madde tipleriyle oluşturulan testlere göre avantajını koruyabileceği şeklinde yorum yapılabilir.

Çoktan seçmeli test ve kısa cevaplı testlerde öğrencilerin elde ettikleri başarı puanları ortalaması arasında fark olup olmadığı, her iki test tipinde uygulanan sınavlardan öğrencilerin elde ettikleri puanların ortalaması alındıktan sonra bağımlı gruplarda t testi ile test edilmiş ve sonuçlar Tablo 3'te verilmiştir.

Tablo 3.

*Çoktan Seçmeli ve Kısa Cevaplı Testlerin Aritmetik Ortalamalarının Karşılaştırılmasına İlişkin t-Testi Tablosu*

Ölçüm	N	$\bar{X}$	S	sd	t	p
K.C.T.	139	8,50	2,58	138	1,47	0,145
Ç.S.T.	139	8,91	3,29			

\*\* p > 0,01

Tablo 3'e göre, öğrencilerin çoktan seçmeli test ile kısa cevaplı testten elde ettikleri başarı puanları ortalaması arasında anlamlı bir farklılık gözlenmemiştir. Öğrencilerin, kısa cevaplı teste ilişkin başarı puanları ortalaması 8,50 iken, çoktan seçmeli teste ilişkin başarı puanları ortalaması 8,91 olarak elde edilmiştir. Çoktan seçmeli teste ait puan ortalaması biraz daha yüksek bir değere sahip olmasına rağmen anlamlı bir farklılık oluşturmamaktadır. Bu durum, testin madde tipinin kısa cevaplı ya da çoktan seçmeli olmasının, öğrencilerin ortalama başarılarını etkilemediği şeklinde yorumlanabilir.

Çoktan seçmeli teste ve kısa cevaplı teste ait KR-20 güvenilirlikleri ve bu iki güvenilirlik katsayısı arasında anlamlı bir farklılık olup olmadığına ilişkin z-testi sonuçları Tablo 4'te verilmiştir.

Tablo 4 incelendiğinde, çoktan seçmeli teste ait KR-20 güvenilirliğinin kısa cevaplı teste göre daha düşük bir

Tablo 4.

*Çoktan Seçmeli ve Kısa Cevaplı Testlerin KR-20 ile Hesaplanan Güvenirliklerine İlişkin z İstatistiği Tablosu*

TESTLER	N	K	KR-20	z
Ç.S.T.	139	19	0,42	
				4,12**
K.C.T.	139	19	0,74	

\*\* p < 0,01

değere sahip olduğu ve bu iki teste ilişkin KR-20 güvenilirlik katsayıları arasındaki farkın 0,01 düzeyinde kısa cevaplı test lehine istatistiksel olarak anlamlı olduğu gözlenmiştir. Bu durumu, güvenilirliğin ölçme sonuçlarına karışan hata miktarının derecesi olduğu düşünüldüğünde, hata faktörünün çoktan seçmeli testlerde daha etkili olabileceği şeklinde yorumlanabilir.

Tablo 5.  
KCT Puanlarının Final Sınavlarını Yordamasına İlişkin Regrasyon Analizi Sonuçları

Değişken	B	Standart hata	$\beta$	t	p
Sabit	48,615	2,951	0,578	16,473	0,00
K.C.T	2,196	0,305		7,187	0,00

$R^2 = 0,334$

KCT puanlarının final sınavı puanlarını yordamasına ilişkin regrasyon analizi sonuçları incelendiğinde, kısa cevaplı test puanlarının, final sınavı puanlarını anlamlı bir şekilde yordadığı tespit edilmiş ( $R = 0,578$ ,  $R^2 = 0,334$ ,  $F_{(1,138)} = 51,658$ ,  $p = 0,00$ ) ve kısa cevaplı test puanları ile final puanları arasında, 0,578 düzeyinde anlamlı bir ilişki bulunmuştur. Kısa cevaplı testten elde edilen puanlar, final puanlarına ilişkin toplam varyansın %33,4'ünü açıklamaktadır.

ÇST puanlarının final sınavlarını yordamasına ilişkin regrasyon analizi sonuçları incelendiğinde, çoktan seçmeli test puanlarının, final sınavı puanlarını 0,05 düzeyinde, anlamlı bir şekilde yordadığı ( $R = 0,231$ ,  $R^2 = 0,053$ ,  $F_{(1,138)} = 5,820$ ,  $p = 0,018$ ) fakat, 0,01 anlamlılık düzeyinde anlamlı yordamadığı tespit edilmiş ve çoktan seçmeli test puanları ile final puanları arasında, 0,23 düzeyinde anlamlı (0,05 anlamlılık

Tablo 6.  
ÇST Puanlarının Final Sınavlarını Yordamasına İlişkin Regrasyon Analizi Sonuçları

Değişken	B	Standart hata	$\beta$	t	p
Sabit	59,641	3,994	0,231	14,932	0,000
Ç.S.T.	1,012	0,419		2,413	0,018

$R^2 = 0,053$

düzeyinde) bir ilişki bulunmuştur. Çoktan seçmeli testten elde edilen puanlar, final puanlarına ilişkin toplam varyansın %5'ini açıklamaktadır.

Yukarıdaki sonuçlardan da anlaşılacağı gibi, kısa cevaplı test puanlarının, final sınavındaki başarıyı çoktan seçmeli testlere nazaran daha iyi yordadığı söylenebilir. Bu sonucun alınmasında, çoktan seçmeli testlerde etkili olan şans başarısının etkili olduğu düşünülmektedir.

### Sonuç ve Öneriler

Bu çalışmada, aynı davranışları yoklayan çoktan seçmeli ve kısa cevaplı testlerin madde ve test istatistikleri karşılaştırılmıştır. Böylece, testlerde kullanılan madde tipinin, madde ve test istatistiklerine etkisi olup olmadığı, öğrencilerin performansını nasıl etkilediğini ortaya çıkarmak amaçlanmıştır. Bu amaç doğrultusunda, testlerin madde güçlük indeksleri (aynı zamanda her bir maddeye ait başarı oranları olarak ele alınmıştır), ayırıcılık gücü indeksi, KR-20 güvenilirlik indeksi ve her iki teste ait ortalamalar hesaplanarak her iki testin madde güçlük indeksi, ayırıcılık gücü indeksi ve KR-20 güvenilirlik indeksleri arasında fark olup olmadığı z testi, test ortalamaları arasında fark olup olmadığı ise t testi ile test edilmiştir. Öğrencilerin final sınavı puanları ölçüt alınarak her iki testten elde edilen puanların final sınavındaki başarıyı ne derece yordadığı da Basit Doğrusal Regrasyon Analizi aracılığıyla test edilmiştir.

Analizler sonucunda, hazırlanan teste ilişkin madde tipinin, yani testin kısa cevaplı ya da çoktan seçmeli maddelerden oluşmasının, her bir hedef davranışın ne derece kazanıldığını gösteren performansın ölçüsü olan madde güçlük indeksi, madde ayırt edicilik gücü indeksi ve KR-20 güvenilirlik indeksini etkilediği gözlenmiştir. Fakat öğrencilerin her iki teste ait ortalama başarı düzeyleri açısından etkili olmadığı tespit edilmiştir. Bu durumda testin madde tipinin kısa cevaplı veya uzun cevaplı olmasının, maddelere ilişkin performans üzerinde etkili olduğu, fakat toplam test puanları üzerinde etkili olmadığı söylenebilir.

Madde ayırıcılık gücü baz alındığında, genel olarak kısa cevaplı testlerin çoktan seçmeli testlere göre daha avantajlı olduğu bulgusuna ulaşılmasına rağmen, çoktan seçmeli testlerde bazı soruların ayırıcılık gücü

indekslerinin (5 ve 19) kısa cevaplı teste göre istatistiksel olarak anlamlı düzeyde daha yüksek bir değere sahip olduğu ve yapılan incelemeler sonucunda, bu sorulara ait çeldirici seçeneklerin alt grup için cazip olduğu ve iyi işlediği gözlenmiştir. Sonuçların böyle çıkmasında, çoktan seçmeli testlerde şans faktörünün ve seçeneklerin hazırlanmasında çok büyük dikkat ve özen gösterilmesi gerekliliğinin etkili olduğu düşünülmektedir. Final sınavları ölçüt olarak alındığında, kısa cevaplı testten elde edilen puanların çoktan seçmeli teste nazaran final sınavlarındaki başarıyı daha iyi yordadığı söylenebilir. Bu durumun ortaya çıkmasında, kısa cevaplı testlerin şans başarısını ortadan kaldırmasının etkili olduğu düşünülmektedir.

Tüm bu sonuçlar ışığında, eğer çeldirici seçenekler çok iyi bir şekilde hazırlanırsa, şans faktörü ve hata faktörü en aza indirgenebilirse, çoktan seçmeli testlerin diğer test tiplerine göre üstünlüğünü ve avantajını koruyabileceği söylenebilir.

#### Kaynakça

- Anastasi, A. (1982). *Psychological Testing*. USA: Macmillan Publishing.
- Aydın, B. (1993). *Seçme Gerektiren Test Maddeleri ile Kısa Cevap Gerektiren Test Maddelerinin Psikometrik Özellikleri ve Öğrenci Başarısı Bakımından Karşılaştırılması* (Yayımlanmamış Yüksek Lisans Tezi) Ankara: Hacettepe Üniversitesi.
- Baykul, Y. (2000). *Eğitim ve Psikolojide Ölçme: Klasik Test Teorisi ve Uygulaması*. Ankara: ÖSYM Yayınları.

- Ertürk, S. (1982). *Eğitimde Program Geliştirme*. Ankara: Yelkenetepe Yayınevi.
- Karaca, E. (1996). *Seçme Gerektiren Testlerle Kısa Cevaplı ve Doğru Yanlış Testlerinin Madde Ve Test Özelliklerinin Karşılaştırılması* (Yayımlanmamış Yüksek Lisans Tezi) Ankara: Hacettepe Üniversitesi.
- Keeves, J. P. (1988). *Educational Research Methodology And Measurement: An International Handbook*. Oxford: Pergamon Press.
- Lancaster, D. M. (1987). *A Comparison of Item Type and Source on Difficulty and Discrimination Ability*. Annual Meeting of Midsouth Educational Research Association in Mobile Alabama, November 13.
- Micheels, W. J. & Karnes, M. R., (1950). *Eğitimde Başarının Ölçülmesi* (çev. İ. Yurt). Ankara: Ajans-Türk Matbaası.
- Özalp, D. (1992). *Kısa Cevaplı ve Çoktan Seçmeli Maddelerden Oluşan Testlerin Psikometrik Özelliklerinin Karşılıklı Olarak İncelenmesi* (Yayımlanmamış Doktora Tezi) Ankara: Hacettepe Üniversitesi.
- Tekin, H. (2000). *Eğitimde Ölçme ve Değerlendirme*. Ankara: Yargı Yayınevi.
- Uğurlu, N. (1996). *Aynı Davranışı Ölçmeye Yönelik Kısa Cevaplı, Çoktan Seçmeli ve Kompozisyon Tipi Maddelerden Oluşan Testlerin Özelliklerinin İncelenmesi* (Yayımlanmamış Yüksek Lisans Tezi) Ankara: Hacettepe Üniversitesi.

Makele Geliş	20 Temmuz 2004
İnceleme	18 Ocak 2005
1. Düzeltme	30 Mayıs 2005
2. Düzeltme	20 Eylül 2006
Kabul	20 Eylül 2006