

## YANITLAYICI DAVRANIŞLARIN ANALİZİ YOLU İLE MATEMATİKTE PROBLEM ÇÖZÜMLERİ İÇİN BİR GÜVENİRLİK VE GEÇERLİK ARAŞTIRMASI

Aysun Umay\*

**ÖZET :** Matematikte yanıtlayıcılara çoktan seçmeli test biçiminde sorulan problemler açık uçlu sorulduğunda ortaya çıkan değişikliklerin araştırıldığı bir çalışma yapılmıştır. Bu çalışma niteliksel bir araştırma özelliğindedir. 100 yanıtlayıcıya 26 maddelik bir çoktan seçmeli test ve çoktan seçmeli testte yer alan problemlerin açık uçlu biçimleri yöneltilmiş ve her yanıtlayıcının aynı sorulara vermiş oldukları yanıtlar (2600 adet) karşılaştırılmıştır. Araştırma sonuçları, matematikte çoktan seçmeli test biçiminde yöneltilen sorular açık uçlu sorulduğunda yanıtlayıcı davranışlarında yapısal bir değişiklik ortaya çıktığını göstermektedir. Ayrıca, iki sınav türünden elde edilen sonuçlar arasında geçerlik ve güvenilirlik açısından bir fark bulunmamış, ancak açık uçlu sınavdaki doğru olmayan yanıtlar türlerine göre gruplandırıldığında hata türlerinin yapısal bir karakteri olduğu saptanmıştır.

**Anahtar Sözcükler:** Çoktan seçmeli test; Açık uçlu sınav; Yazılı yoklama; Geçerlik; Güvenirlik; Matematikte ölçme.

**ABSTRACT:** In this study, the changes, which occurred when multiple-choice testing were asked as open-ended testing in mathematics, were examined. This research is a qualitative study. A multiple-choice test with 26 items and open-ended problems which placed in the multiple choice test were given to respondents. Answers (2600 items) were compared to each other. Results of this study showed that when multiple-choice testing was asked as open-ended testing in mathematics, structural changes have occurred in behaviours of the respondents. Besides, there were no differences between the results of these two types of tests on validity and reliability. However, when wrong answers in open-ended tests were grouped according to their types, error types had a structural characteristics.

**KEY WORDS:** Measurement in mathematics, Multiple-choice-test, Open-ended-test, Essay test, Validity and Reliability.

### 1. GİRİŞ

Eğitimde ölçme gereksinimi eğitim kavramının kendisi kadar eskidir. Eğitim sürecini ve bu süreci etkileyen faktörleri doğru bir şekilde değerlendirmek,

beklenen bir hedefe ulaşıldığına ya da bu hedefe ulaşabilmek için daha çalışmak gerektiğine karar vermek ancak geçerli ve güvenilir ölçme sonuçları ile olanaklıdır.

Eğitimde ölçme genellikle yazılı yoklama, sözlü sınavlar ve testler ile yapılır. Her sınav türünün kendisine özgü üstünlükleri ve sakıncaları vardır. Bu nedenle, hangi sınav türünün kullanılacağına karar verilirken grubun nitelikleri, birim sayısı, konu alanı gibi değişkenler göz önünde bulundurulur. Böyle yaklaşıldığında, sınav türünün belirlenmesinde eğitimciler arasında genellikle bir ortak anlayış bulunmasına karşın zaman zaman yazılı yoklama ve bir test türü olan çoktan seçmeli testler arasındaki seçimde bir ikircik yaşanmaktadır. Tartışma özellikle de bu sınav türlerinin geçerlik ve güvenilirlikleri üzerinde yoğunlaşmaktadır.

Ölçme sonuçlarının geçerliği, ölçülmek istenen bir niteliğin bir başka nitelikte karıştırılmadan ölçülebilme derecesi olarak tanımlandığına göre (1), yazılı yoklamaların puanlanmasında geçerliğin sağlanmasının ciddi sıkıntıları olacağını tahmin etmek zor olmaz. İfade gücü, yazı güzelliği gibi değişkenle ilgili olmayan bazı faktörlerin, soruların yazılarak yanıtlandığı bir sınavda puanlayıcıyı az ya da çok, ama büyük olasılıkla etkileyeceği ve puanlamanın varlığını arttıracığı söylenilebilir. Buna karşılık yanıt tek bir cümle ya da sayı ile ifade edilmediği için yanıtlayıcının düşünme biçimini daha doğrudan gözleme olanağı bulunacaktır. Bir başka deyişle, yazılı yoklamada yanıtlayıcının düşünme süreci rahatlıkla izlenebilir, tam olarak yerleşmemiş olan bilgileri saptanabilir, yeni, farklı, yaratıcı düşünme ürünleri yakalanabilir. Kuşkusuz bütün bunlar önemli bir avantajdır ve özellikle yazma becerilerinin ölçülmesinde yazılı yoklamanın geçerliğinin çoktan seçmeli testlerden daha yüksek bulunmasına neden olur (2).

Yazılı yoklamanın geçerliğini arttıran bir başka özelliği de özgün değerlendirmelere olanak vermesidir. Özgün değerlendirme, çocukların okul yaşantılarının yanı sıra okul dışı çevrelerindeki performansını da göz önüne alır ve standart başarı testlerine bir alternatif oluşturur (3). Buna karşılık, eğitim ortamı ve koşulları dikkate alınarak hazırlanan

\* Yard. Doç. Dr. Aysun Umay, Hacettepe Üniversitesi Fakültesi Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı Öğretim Üyesi.

özgün sorularla yapılan ölçmelere puanlayıcıların değer yargıları kaçınılmaz olarak daha çok yansıyacaktır. 1990 yılında yapılmış olan bir çalışmayı anlatırken Tucman, yazılı yoklamaların puanlanmasında ortak ölçütler geliştirmek üzere çağırılan uzmanların uyuşamadıkları noktaların uyuştuklarından fazla olduğunu belirtiyor (4).

Bridgeman ve Lewis, 32 değişik kolejin birinci sınıflarında okuyan 6243 öğrenci üzerinde yapmış oldukları bir çalışmada, bir üst sınıfa geçme sınavlarında kullanılan yazılı yoklama ve çoktan seçmeli testlerin sonuçları arasındaki korelasyonun çok yüksek olduğunu belirtiyorlar (5). Bu sonuç, iyi hazırlanmış çoktan seçmeli testlerin de yüksek bir geçerliğe sahip olabildiklerini açıkça göstermektedir. Çoktan seçmeli testlerde yazılı yoklamadan daha fazla soru sorabilme olanağına sahip olmanın, bu sınav türünün geçerliğini, özellikle kapsam geçerliğini arttırdığı ise herkesçe bilinmektedir. Yine de, öğrencileri daha az düşünmeye ve daha fazla ezbere yönelttiği iddia edilen (6) çoktan seçmeli testlerin öğrenimin ilk basamaklarında yaygın olarak kullanılması eğitimcilerce de pek fazla benimsenmemektedir. Bunun haklı nedenlerinden biri, küçük yaşlarda yazma ve anlatım becerilerinin gelişmesinin engelleneceği endişesi, bir diğeri öğretmenlerin kendilerini test hazırlama konusunda yeterli bulmamaları olabilir.

Ölçme sonuçlarının hatalardan arınlık derecesi onun güvenilirliğini gösterir (7). Şans eseri sınavda soruların bilinen (ya da bilinmeyen) konulardan çıkması, doğru yanıtın atıp tutturulması, puanlayıcının yanlılığı gibi bir çok kaynağı bulunabilen ya da bulunamayan etmen, kaçınılmaz olarak ölçme sonuçlarına hata karışmasına neden olur.

Yazılı yoklamalarda hata, sınav sonuçlarına değişik aşamalarda karışabilir. Uygulama öncesi alınabilecek önlemlerin başında soru sayısını artırarak şans faktörünün etkisini azaltmak gelir. Soru sayısını arttırmak ise, yazmanın da ayrıca zaman alması nedeniyle bu tür sınavlarda sınav süresinin daha da uzaması demektir. O halde soru sayısını arttırabilmek için soruların kapsamı daraltılarak bir anlamda kısa yanıt gerektiren testlere yaklaştırılması gerekir. Ancak, bu durumdaki bir sorunun yapısı değişmiştir ve artık yazılı yoklamanın geçerlikle ilgili olarak yukarıda belirtilen avantajlarını taşıması söz konusu değildir.

Hata, ölçme sonuçlarına uygulama sonrasında, puanlama aşamasında da karışabilir. Yazılı yoklamada puanlamanın objektifliğini sağlamanın zorluğu, bu sınav türünün en önemli sakıncalarından sayılır. Bu sakıncanın ortadan kaldırılabilmesi için alınabilecek olan önlemlerin başında ayrıntılı bir

yanıt anahtarı hazırlayarak puanlamayı anahtara uygunluğuna göre yapmak gelir. Ancak bu önlem, üst düzey bilişsel basamaklardaki sorulara verilen yanıtların puanlanmasında pek işe yaramaz. Özellikle analiz, sentez gibi, bireye özgü çözümlerin istendiği sorulara verilecek yanıtların önceden ayrıntılarıyla belirlenerek puan dağılımının yapılmasının olanaksızlığı açıktır. Oysa yazılı yoklama türü sınavların en büyük üstünlüğü olarak böyle yüksek düzey davranışları da ölçebilmesi gösterilmektedir (8). Verilmesi beklenen yanıtın kesin ve belli olduğu durumlarda ise, puanlamadaki kolaylığı da göz önüne alınarak genellikle çoktan seçmeli testler yeğlenir. Bir başka deyişle, yazılı yoklamayla yapılan ölçme işlemlerinde, güvenilirliği arttırmak için alınabilecek önlemler, eğitimcileri soru tipini değiştirmeye yöneltmektedir. Bu bir ikilemdir.

Çoktan seçmeli testin yanıtlarının puanlanmasında bir objektiflik problemi olmaması bu sınav türünün güvenilirlik açısından en önemli avantajıdır. Dahası, bu sınav türünde puanlamada mekanik yollar da kullanılabilir. Bu özelliği birim sayısının çok olduğu durumlarda çoktan seçmeli test uygulanmasını zorunlu hale getirir.

Çoktan seçmeli testlerin, yanıtlayıcıların düşünme sürecini ve yanıtlama davranışını izleme olanağını vermemesi ise en büyük eksikliğidir. Oysa unutulmamalıdır ki aynı özelliği nedeniyle puanlamada objektiflik sağlanabilmektedir. Bu da çoktan seçmeli testlerin ikilemidir.

Çoktan seçmeli testlerde doğru yanıtın seçenekler arasında bulunması şans başarısı olasılığını kaçınılmaz kılmaktadır. Şans başarısının elimine edilebilmesi için çeşitli yollar geliştirilmiştir (9). Çoktan seçmeli testlerde yanıtlayıcı davranışlarının incelendiği bir araştırmada, yanıtlayıcıların çoğunun doğru yanıtı bulduğunu düşünerek maddeyi işaretlediği, doğru yanıtı bulduğuna inanmadığında ise boş bırakmayı yeğlediği anlaşılmaktadır. Aynı araştırmada şans başarısını elimine etmek için önerilen yöntemlerin düzeltme işlemini gerçekten etkili ve doğru bir biçimde yaptıkları, yanıtlayıcılara "yanlışların belli bir oranda doğruları sileceği" söylendiği zaman kullanıldığında ise aşırı bir eliminasyona gittiği saptanmıştır (10). Şans başarısının büyük ölçüde elimine edilebiliyor olması önemli bir avantaj olarak görülmele birlikte eliminasyon formüllerinin puanlayıcılarca yaygın olarak kullanıldığını söylemek oldukça zordur. Bu nedenle şans başarısı çoktan seçmeli testlerin güvenilirliğini olumsuz yönde etkileyen önemli bir faktör olmayı sürdürmektedir.

Şans başarısı elimine edilmediğinde çoktan seçmeli testlerin başarı puanları ortalamasının yazılı yoklamalardan daha yüksek bulunduğunu gösteren pek

çok araştırma vardır. Beklenen form ve güçlüğün bir işlevi olarak test performansının incelendiği bir araştırma, bireylerin beklentileri için zor olduğu yönündeysen daha sıkı çalıştıklarını ve bu durumun her iki sınav türü için de geçerli olduğunu ortaya koymuştur (11). O halde performansın ölçülmesinde sınav türlerinin güçlüğü ya da kolaylığı değil yanıtlayıcıların beklentileri belirleyici olmaktadır.

Bütün bu tartışmalar acaba matematik alanına nasıl yansır? Matematikte yazılı yoklama ve çoktan seçmeli testle yapılan ölçmelerde geçerlik ve güvenilirlik hangi durumdadır?

Matematik sınavlarının diğer alanlardaki sınavlardan en belirgin farkı, çoğunlukla yanıtlayıcılardan problem çözmelerinin istenmesidir. Problemi çözerken, yanıtlayıcının düşünme sürecinin rahatlıkla izlenebilmesi, tam olarak yerleşmemiş olan bilgilerinin saptanabilmesi, yeni, farklı, yaratıcı çözüm yolları üretenlerin yakalanabilmesi matematiksel düşünmenin ölçülmesinde büyük önem taşır. Çözüm sırasında yapılan hataların bilgisizlikten mi, dikkatsizlikten mi yoksa yeni bir yaklaşım içinde olup doğru sonuca ulaşamamaktan mı kaynaklandığına bakılarak yanıtlayıcıların matematiksel yaklaşımları daha doğru değerlendirilebilir. Bu durum geçerliği olumlu yönde etkiler. Ancak burada da bir değerlendirme söz konusu olduğuna göre puanlayıcının yargısı devreye girecektir. Yazı güzelliği, ifade gücü gibi özellikle sözle ifadeye dayanan alanlarda çok etkili olan faktörler kadar varyansa sahip olmasa da bir yargının söz konusu olması geçerliği düşürür.

Çoktan seçmeli testin güvenilirliğini arttıran en önemli özelliklerinden biri, çok sayıda soru sorulabilmesidir. Aynı sınav süresi içinde daha fazla problem sorulabilmesi, matematikte de sınav yaparken çoğu zaman çoktan seçmeli testin yeğlenmesine neden olmaktadır. Ayrıca alanın özelliği gereği, problem çözümü sırasında düşülebilecek hataların kestirilmesinin genellikle güç olmaması, dolayısıyla güçlü çeldiriciler hazırlanabilmesi, puanlayıcıya süreci doğrudan gözlemeden de düşülen hataları izleyebilme olanağı vermektedir. Eğer, problem çözerken yapılan hatalar iyi analiz edilebilirse (ki bu çalışmada yapılmaya çalışılmıştır), çoktan seçmeli testin kısa sürede ve kolayca puanlanabilme gibi avantajlarından vaz geçmeden de problemin çözüm süreci konusunda fikir sahibi olunabilir.

### 1.1. Problem

Matematik eğitimindeki yeni yaklaşımın, problem çözen öğrenciyi özgürce düşünmeye ve problemlere alışılmış kalıplar dışında kendine uygun, yeni, yaratıcı çözümler üretmeye yöneltmek olduğu kabul edilirse, problem çözme sürecinin ölçülmesi sırasında gös-

terilen davranışların analizi daha büyük bir önem kazanmaktadır.

Yanıtlayıcılar matematikte problem çözerken nasıl davranmaktadır? Problemleri yazılı olarak yanıtlarken doğru yanıtı ulaşamayanlar aynı problem çoktan seçmeli test maddesi biçiminde karşısına çıktığında ne yapmaktadır? Yazılı yoklamada bulunduğu yanlış sonucu seçenekler arasında bulanların yüzde kaçını aynı seçeneği işaretlemektedir? Aslında problemi çözdüğü halde yalnızca dikkatsizlik (örneğin işlem hatası) nedeniyle yanlış bir sonuca ulaşanların yüzde kaçını testte doğru yanıtı ulaşmaktadır?

Bunlar daha bunlar gibi pek çok sorunun yanıtı matematikçilerce merak konusudur. Eğer bu sorulara tatmin edici yanıtlar alınabilirse, matematikte problem çözme becerisinin ölçülmesinde daha geçerli ve güvenilir sonuçlara ulaşılabilir.

Matematikte problem çözme becerisi yoklanırken kısa kısa, ama çok sayıda problem sorulması "şans" faktörünü büyük ölçüde azaltarak güvenirliliğin artmasına neden olur. Böyle soruları "açık uçlu" biçiminde tanımlamak daha doğru olacaktır. Yazma ve ifade becerilerinin etkili olmadığı bu soru tipindeki yanıtlar da 1 ve 0 olarak puanlanabilir. Böyle olduğunda problem çoktan seçmeli test biçiminde sorulduğunda verilen yanıtlar ile yanıtlayıcının kendisi çözmesi istendiğinde verilmiş olan yanıtları karşılaştırmak çok daha kolay olacaktır.

### 1.2. Problem Cümlesi

Matematikte problemlerin çoktan seçmeli test yerine açık uçlu sorular biçiminde sorulması yanıtlayıcıların davranışlarına nasıl yansımakta, bu durum güvenilirliği ve geçerliği nasıl etkilemektedir?

### 1.3. Alt Problemler

1. Matematikte problemlerin çoktan seçmeli test yerine açık uçlu sorular biçiminde sorulması yanıtlayıcıların davranışlarını nasıl etkilemektedir?

1.1. Matematik problemlerini çözenlerin çoktan seçmeli testteki başarılarıyla açık uçlu sınavdaki başarıları arasında fark var mıdır?

1.2. Çoktan seçmeli testte matematik problemi çözenlerin verdikleri yanıtlar, aynı problemler açık uçlu sorular biçiminde verildiğinde nasıl değişmektedir?

2. Matematikte problemlerin çoktan seçmeli test yerine açık uçlu sorular biçiminde sorulması güvenilirliği nasıl etkilemektedir?

2.1. Çoktan seçmeli test için hesaplanan güvenirlilik katsayısı ile aynı soruların açık uçlu sorulması ile hesaplanan güvenirlilik katsayısı arasında fark var mıdır?

2.2. Açık uçlu sınavdaki doğru olmayan yanıtların çoktan seçmeli testteki dağılımı nasıldır?

2.3. Açık uçlu sınavdaki doğru olmayan yanıtlar yapısal bir özellik göstermekte midir?

2.4. Açık uçlu sınavda yanıtlayıcıların verdikleri doğru olmayan yanıtlar yapılan hata türlerinin düzeyleri ile ne kadar açıklanabilmektedir?

3. Matematikte problemlerin çoktan seçmeli test yerine açık uçlu sorular biçiminde sorulması geçerliği nasıl etkilemektedir?

3.1. Çoktan seçmeli test ve açık uçlu sınava verilen yanıtlar arasındaki ilişki nasıldır?

3.2. Çoktan seçmeli testteki maddelerin geçerlik indeksleri, aynı maddeler açık uçlu sorulduğunda nasıl değişmektedir?

## 2. YÖNTEM

Yapılan çalışma, yanıtlayıcı davranışlarını analiz ederek bunların güvenilirlik ve geçerlik üzerindeki etkisini irdelemeye yönelik olduğundan niteliksel bir araştırma özelliği taşımaktadır.

Araştırma, Hacettepe Üniversitesi Fen Fakültesi'nde okumakta olan 100 öğrenci üzerinde yürütülmüştür.

Kullanılan ölçme araçlarından çoktan seçmeli test, daha önce araştırmacı tarafından geliştirilen ve geçerliği, güvenilirliği (sırasıyla 0,69 ve 0,96 olarak) hesaplanmış 50 maddelik bir çoktan seçmeli matematik muhakeme testi içinden, amaca uygun olarak seçilen 26 maddeden oluşmaktadır. Oluşturulan yeni testin geçerlik ve güvenilirlik çalışmaları aynı verilerle yeniden yapılarak sırasıyla 0,71 ve 0,93 olarak bulunmuştur. Açık uçlu sınavda ise çoktan seçmeli teste kullanılan problemler yanıtlayıcılara aynen ama bu kez açık uçlu olarak sorulmuştur.

Çoktan seçmeli test uygulanırken yanıtlayıcılara çözemedikleri soruları boş bırakmaları söylenmiştir. Aynı gruba üç hafta sonra uygulanan açık uçlu sınavda ise özellikle yanıtlayıcılardan çözümün açık olarak gösterilmesi istenmiştir.

## 3. BULGULAR VE YORUM

Bu bölümde, araştırmanın bulguları ve bu bulgulara ilişkin yorumlar, alt problemler teker teker ele alınarak tartışılmaktadır.

**1. Matematikte problemlerin çoktan seçmeli test yerine açık uçlu sorular biçiminde sorulması yanıtlayıcıların davranışlarını nasıl etkilemektedir?**

Yanıtlayıcı davranışlarına ilişkin bu alt problemle ilgili olarak şu iki soruya yanıt aranmıştır:

**1.1. Matematik problemlerini çözenlerin çok-**

**tan seçmeli testteki başarılarıyla açık uçlu sınavdaki başarıları arasında fark var mıdır?**

Sınavlara katılan 100 kişiden 74'ünün çoktan seçmeli teste daha yüksek puan aldığı, buna karşılık açık uçlu sınavda daha yüksek puan alanların sayısının yalnız 20 olduğu görülmektedir. İki sınavda da aynı sonucu elde edenlerin sayısı ise 6'dır. Çoğunluğun açık uçlu sınavda daha düşük puan almaları dışında yanıtlayıcıların çoktan seçmeli test ve açık uçlu sınavlardaki başarı puanları arasında da belirgin bir farklılık gözlenmektedir. İki sınavın başarı puanları ve puanlar arasındaki farklara ilişkin aritmetik ortalama ve standart sapmalar Tablo 1'de gösterilmiştir.

Tablo 1 : Çoktan Seçmeli Test ve Açık Uçlu Sınav Sonuçları

	Arit.Ort.	Stan. Sap.
Çoktan Seçmeli Test	16,15	4,59
Açık Uçlu Sınav	13,46	4,69
Fark	2,69	3,35

Tablo 1'de de görüleceği gibi açık uçlu sınav yanıtlayıcılarca daha zor bulunmuştur. Bunun nedeni sınavların yapılarının farklılığı olabileceği gibi çoktan seçmeli teste karışan şans başarısı da olabilir. Bilindiği gibi çoktan seçmeli teste seçeneklerden birinde doğru yanıtın verilmesi, kimi zaman yanıtlayıcılara hedeflenmediği durumlarda da seçeneklerden giderek, yani çeldiricileri teker teker eleyerek ya da tamamen şans eseri (1/seçenek sayısı olasılıkla) doğru yanıtı bulma olanağı vermektedir. Doğru yanıtın böyle, soru tipinden kaynaklanan avantajlar yardımıyla bulunması geçerlik ve güvenilirliği olumsuz yönde etkiler. Her iki sınav aracının da geçerlik ve güvenilirliklerinin yüksek bulunması (Bkz. 2. ve 3. altproblem) ve aradaki farkın büyüklüğü, böylesi bir sonucun elde edilmesinde soru tipinden kaynaklanan avantajlar dışında bazı yapısal özelliklerin daha etkili olduğu görüşünü güçlendirmektedir. Bu yapısal özelliklerin neler olabileceği ancak yanıtlayıcı davranışlarındaki değişikliklerin ayrıntılı olarak incelenmesiyle mümkün olabilir.

**1.2. Çoktan seçmeli testte matematik problemi çözenlerin verdikleri yanıtlar, aynı problemler açık uçlu sorular biçiminde verildiğinde nasıl değişmektedir?**

Çoktan seçmeli teste bulunan 26 maddeye 100 yanıtlayıcı tarafından verilen yanıtların dağılımı Tablo 2'de görülmektedir.

Tablo 2 : Çoktan Seçmeli Test Yanıtlarının Dağılımı

DOĞRU		YANLIŞ		BOŞ		TOPLAM	
f	%	f	%	f	%	f	%
1615	62	486	19	499	19	2600	100

Tablo 2'ye bakıldığında maddelerin % 62'sinin doğru olarak yanıtlandığı ve boş bırakılan madde sayısının yanlış yanıtlanan madde sayısından az bir farkla da olsa daha fazla olduğu görülmektedir. Bu durum yanıtlayıcıların bilmediklerini düşündükleri soruları yanıtlamamak eğiliminde oldukları izlenimini vermektedir. Bu ise "şans başarısı" faktörünün fazla etkili olmadığı ve Tablo 1'de görülen ortalamalar arasındaki farkın daha çok sınav türlerinin yapısından kaynaklandığı biçiminde yorumlanabilir.

Aynı şekilde açık uçlu sınavda bulunan 26 maddeye 100 yanıtlayıcı tarafından verilen yanıtların dağılımı Tablo 3'de gösterilmiştir.

Açık uçlu sınavda verilen "yanlış" yanıtlar ,

•**YD(yanlış/dikkatsiz):** Problemi çözdüğü, çözümlü bulduğu halde basit işlem hataları gibi dikkatsizlikle yanlış sonuç bulmuş,

•**YV(yanlış/var):** Bulduğu sonuç teste de çeldirici olarak verilmiş ve

•**YA(yanlış/aynı):** testi yanıtlarken bu seçeneği işaretlemiş,

•**YB(yanlış/başka):** testi yanıtlarken bu seçeneği değil, başkasını işaretlemiş,

•**YY(yanlış/yok):** Bulduğu sonuç teste çeldirici olarak verilmemiş,

olmak üzere üç ana grupta toplanmıştır. Açık uçlu sınavda "boş" olarak tanımlanan yanıtlar için ise,

•**TB(tamamen/boş):** Yanıtlayıcının hiç dokunmadığı, tamamen boş bıraktığı,

•**BÇ(boş/çalışmış):** Yanıtlayıcının çözmeye çalıştığı ama herhangi bir sonuca ulaşamayacağını anlayıp bıraktığı olmak üzere iki grup yapılmıştır.

Tablo 3 : Açık Uçlu Sınav Yanıtlarının Dağılımı

DOĞRU		YANLIŞ			BOŞ	
	YD	YV=YA+YB	YY	TB	BÇ	
	106	147=75+72	164	648	189	
f	%	f	%	f	%	
1346	52	417	16	837	32	

Tablo 3'de, açık uçlu sınavda yanıtları boş bırakılan soru sayısının (837) neredeyse yanlış yanıtların (417) iki katı olduğu görülmektedir. Burada gözden ka-

çırılmaması gereken nokta, "Boş" olarak kabul edilen yanıtların bir kısmı üzerinde uğraşmış ama bir sonuca ulaşamamış (BÇ=189) olmasıdır. Ancak yine de tamamen boş bırakılan soruların sayısı (TB=648) bile yanlış yanıtların sayısından çok fazladır. Bu durum yanıtlayıcıların çoktan seçmeli testlerde daha atak davrandıkları biçiminde yorumlanabilir.

Çoktan seçmeli testin başarı ortalamasının daha yüksek olması bu ataklığın olumlu sonuç verdiğini göstermektedir. Kaldı ki açık uçlu sınavda "yanlış" olarak tanımlanan yanıtlar arasında problemi çözdüğü halde yalnızca dikkatsizlik yüzünden bulunan yanlış sonuçlar (YD=106) da vardır. Öğretmenlerin çoğu kağıtları puanlarken bu tip hatalardan ya hiç puan kırmazlar ya da küçük indirimler yaparlar. Matematik sınavında amaç dikkati ölçmek olmadığına göre bu yaklaşım ölçme teknikleri açısından da doğrudur ve geçerliği arttırır. Böyle dikkatsizlik nedeniyle yanlış yanıtlanan soruların çoktan seçmeli teste (dikkat hataları üzerine çeldirici düzenlenmemişse) doğru yanıtının bulunması beklenir. Bu durum, açık uçlu sınavda verilen yanlış yanıtların çoktan seçmeli teste verilen yanıtlarla karşılaştırıldığı altproblemde ayrıntılı olarak incelenmektedir (Bkz. 2.2. altproblem).

Açık uçlu sınavda bulunan yanlış sonuçlardan bir kısmı çoktan seçmeli teste çeldiriciler arasında verilmiştir. Bazı sonuçları yanıtlayıcılar testi yanıtlarken de aynı biçimde bulmuş ve seçenekler arasında görüp işaretlemiştir (YA=75). Bu bulgu testin güvenirliliğinin bir işareti olarak kabul edilmelidir. (Güvenirlilik her zaman geçerliği arttırıcı bir özellik olarak kabul edildiğinden bu yanıtlar sınavların geçerliğine de (sınırlı olarak) bir işaret sayılabilir.) Ancak verilen yanlış yanıtlar arasında, açık uçlu sınavda bulunan yanlış sonuç testin seçenekleri arasında bulunduğu halde bu seçeneğin değil, bir başka seçeneğin (bu arada doğru yanıtın) işaretlenmesi ya da boş bırakılması sayısı da (YB=72) oldukça fazladır. Bu durum güvenirliliğin düşüklüğü ya da "yanlış yanıtlama" davranışlarının yapısı gereği fazla kararlı olmaması biçiminde açıklanabilir. Açık uçlu sınavın güvenirlilik katsayısı yeterince yüksek olduğuna göre (Bkz. 2.1. altproblem) daha çok yanlış yanıtlama davranışlarının yapısının etkisi söz konusudur. Nitekim açık uçlu sınavdaki yanlış sonuçlar arasında bulunan sonucun çoktan seçmeli teste olmadığı yanıtların sayısı (YY=164) en büyük grubu oluşturmaktadır. Bu gruptaki maddelerin çoktan seçmeli teste boş bırakılmış olması beklenir (Bkz. 2.2. altproblem). Oysa bilindiği gibi çoktan seçmeli teste boş bırakılan yanıtların sayısı açık uçlu sınavdakilerden çok daha azdır. Bu durumda açık uçlu sınavda yanlış yanıtlanan sorular için yanıtlayıcı davranışlarının kararlı olmadığını kabul etmek çıkarılabilecek sonuçların en mantıklısı olarak görülmektedir.

## 2. Matematikte problemlerin çoktan seçmeli test yerine açık uçlu sorular biçiminde sorulması güvenilirliği nasıl etkilemektedir?

Güvenirlige ilişkin bu alt problemle ilgili olarak şu dört soruya yanıt aranacaktır:

### 2.1. Çoktan seçmeli test için hesaplanan güvenilirlik katsayısı ile aynı soruların açık uçlu sorulması ile hesaplanan güvenilirlik katsayısı arasında fark var mıdır?

Ölçme araçlarının güvenilirlikleri KR-20 yöntemiyle hesaplanmış ve Tablo 4'de gösterilmiştir.

Tablo 4 : İki Ölçme Aracının Güvenirlik Katsayıları

Ölçme Araçları	KR-20
Çoktan Seçmeli Test	0,84
Açık Uçlu Sınav	0,82

Tablo 4'de görüldüğü gibi her iki ölçme aracının da güvenilirlik katsayıları oldukça yüksek ve birbirine yakındır. Bu durum, soruların çoktan seçmeli ya da açık uçlu biçimde sorulmasının güvenilirliği etkilemediği biçiminde yorumlanabilir. Her iki ölçme aracının da güvenilirlik katsayılarının yüksek olması bulguların güvenilirliği açısından da önemlidir.

### 2.2. Açık uçlu sınavdaki doğru olmayan yanıtların çoktan seçmeli testteki dağılımı nasıldır?

Açık uçlu sınavdaki doğru olmayan yanıtların çoktan seçmeli testte nasıl yanıtlandığı Tablo 5'de gösterilmiştir.

Tablo 5 : Açık Uçlu Sınavdaki Doğru Olmayan Yanıtların Çoktan Seçmeli Testteki Dağılımı

Açık Uçlu Yan.	Doğru	Yanlış	Boş	Toplam	
YD	81	17	8	106	
YV	YA	-	75	75	147
	YB	46	17	9	
YY	85	37	42	164	
TB	200	157	291	648	
BÇ	75	51	63	189	
TOP	488	353	413	1254	

Tablo 5'den, açık uçlu sorulara verilen doğru olmayan yanıtların (1254) üçte birinden çoğunun (% 39) çoktan seçmeli testte doğru olarak yanıtlandığı (488) anlaşılmaktadır. Bu oran göz ardı edilemeyecek kadar büyüktür. Dikkatsizlik sonucu yanlış sonuca ulaşılan sonuçlar (81) dikkate alınmasa bile, testin güvenilirlik katsayısı da oldukça yüksek olduğuna göre, bu durum "raslantı" ile açıklanamaz. Böyle ya-

nitlenen sorulara yakından bakıldığında en büyük bölümü açık uçlu sınavda tamamen boş bırakılmış olan soruların (TB=200) oluşturduğu dikkati çekmektedir. Bu bulgu bize çoktan seçmeli testlerle açık uçlu sınavlar arasındaki yapısal bir farkın açık bir işaretini verir. Açık uçlu sınavda çözmeye "çekinilen" ya da "üşenilen" soruların önemli bir kısmı çoktan seçmeli olarak sorulduğunda çözülmüş ve doğru yanıtla ulaşılmıştır. Çoktan seçmeli testte doğru yanıtla ulaşıldığı halde açık uçlu sınavda çözülmeye çalışılıp çözülemeyen (BÇ=75) ya da farklı sonuca ulaşılan (YB=46, YY=85) maddeler ise ancak "şans başarısı" ve "matematik problemlerinin çözümünde çözümün sezildiği anın her zaman aynı biçimde ortaya çıkmaması" ile açıklanabilir (12).

Açık uçlu sınavda doğru olarak yanıtlanmayan soruların (1254) yaklaşık üçte biri de (% 28) çoktan seçmeli testte yanlış (353) olarak işaretlenmiştir. Açık uçlu sınavda yanlış sonuca ulaşılanlar (417) ile karşılaştırıldığında bu oran % 85'i bulmaktadır. Çoktan seçmeli testte verilen yanlış yanıtlar içinde en büyük bölüm açık uçlu sınavda boş bırakılmıştır (TB=157). Bu bulgu, bazı yanıtlayıcıların çoktan seçmeli testte şanslarını denedikleri, bu arada testte doğru gibi görünen bazı yanıtların şans başarısı olduğu tezini de güçlendirmektedir (Bkz. 1.1. alt problem). Her iki sınavda da aynı sonuca ulaşılanlar (YA=75) ise çoktan seçmeli testte yanlış işaretlenenler içinde ikinci sırayı almaktadır. Buna karşılık açık uçlu sınavda ulaşılan yanlış sonuç seçenekler arasında bulunduğu halde bir başka çeldiricinin işaretlenmesi (YB=17) aynı yanlış seçeneği işaretleyenlerin ancak beşte biri (%18) kadardır. Bu bulgu, daha önce de belirtildiği gibi işlemin güvenilirliğine bir işaret olarak görülmelidir.

Açık uçlu sınavda doğru olarak yanıtlanmayan soruların (1254) üçte bir kadarını da (% 33) çoktan seçmeli testte boş bırakılmış olan (413) yanıtlar oluşturmaktadır. Burada dikkati çeken durum, açık uçlu sınavda doğru yanıtlanmayan ve çoktan seçmeli testte de boş bırakılanların büyük çoğunluğunun (% 70) açık uçlu sınavda tamamen boş bırakılmış (291) olmasıdır. Bu bulgu, yanıtlayıcıların çekindikleri, hatta hoşlanmadıkları konulardaki soruları çözmeye bile çalışmadıkları biçiminde yorumlanabilir. Böyle bir yorum ise "matematiğe karşı önyargının bir başka göstergesi" olarak kabul edilebilir.

Tablo 5'e, yanlışların türü esas alınarak bakıldığında da ilginç bulgularla karşılaşılmaktadır. Açık uçlu sınavda dikkatsizlik nedeniyle ulaşılan yanlış sonuçların (106) dörtte üçünden fazlasında (% 76) çoktan seçmeli testte doğru yanıt bulunabilmiştir. 1.2. alt-problemde de belirtildiği gibi bu sonuç istenilen yönde, geçerliliği arttırıcı bir özelliktir ve çoktan seçmeli testin avantajları hanesine yazılmalıdır.

Açık uçlu sınavda bulunan sonuca çoktan seçmeli testte çeldiriciler arasında yer verilmiş olan sorulara verilen yanıtların (147) yaklaşık yarısında (% 51) aynı seçenek işaretlenmekle birlikte (75) diğer yarının büyük ölçüde (% 64) doğru yanıtı yönelmesi (46) dikkat çekicidir. Bu bulgu çoktan seçmeli testte doğru yanıtı ulaşılan bir sorunun açık uçlu olarak sorulması durumunda, öngörülebilir bir bilgi hatasına düşülebildiğini göstermektedir. Bu durum ancak, böyle soruların çoktan seçmeli testte raslantıyla (şans başarısı) ya da altyapısı sağlam olmayan bir tür sezgiyle bulunmuş olması biçiminde açıklanabilir. Aynı biçimde açık uçlu sınavda ulaşılan sonuç testin çeldiricileri arasında bulunmayan sorulara verilen yanıtların (164) yarısında da (% 52) soru çoktan seçmeli iken doğru yanıtı ulaşılmıştır (85). Bu durum, belki de bulunduğu yanlış sonuç seçenekler arasında yer almadığından bazı deneklerin uyarılmalarına ve daha sonra büyük ölçüde doğru yanıtı yönelmelerine neden olmaktadır.

Bütün bu bulgular, çoktan seçmeli testte başarının daha fazla olduğunu göstermektedir. İstatistiksel olasılık hesapları ve boş bırakılmış yanıtların çokluğu göz önüne alınarak, içinde önemli bir payı olmakla birlikte bu farkı yalnızca "şans başarısı" olarak tanımlamak doğru olmaz. Belki sezgi, seçenekleri elemek gibi olanakları ya da başka özellikleri nedeniyle çoktan seçmeli testin yanıtlayıcılarca açık uçlu sınavdan daha kolay bulunduğunu söylemek yanlış olmaz.

### 2.3. Açık uçlu sınavdaki doğru olmayan yanıtlar yapısal bir özellik göstermekte midir?

Açık uçlu sınava verilen doğru olmayan yanıtların yapısını incelemek üzere yanlış yanıtlar (YD, YV, YY) ve boş bırakılan yanıtlar (TB, BÇ) üzerine faktör analizi uygulanmıştır.

Faktör analizi "Stat View II" paket programıyla yapılmış "Principal Components" yöntemi ve "Orthotran / Varimax Transformation" işlemi kullanılmıştır.

Yapılan analizde 3 faktör bulunmuştur. Bu faktörlere ilişkin eigen değerler Tablo 6'da gösterilmiştir.

Tablo 6 : Hata Değişkenlerinin Eigen Değerleri

	Büyüklik	Açıklanan Varyans
Faktör 1	1,47	,29
Faktör 2	1,23	,25
Faktör 3	1,03	,21

Tablo 6'da görüldüğü gibi üç faktörün büyüklükleri oldukça yakındır ve üçü birlikte toplam varyansın ,75'ini açıklamaktadır. Üç faktörün değişkenler içindeki yüklerini gösteren Varimax yöntemiyle döndürülmüş faktör matrisi Tablo 7'de gösterilmiştir.

Tablo 7 : Döndürülmüş Faktör Matrisi

	Faktör 1	Faktör 2	Faktör 3
YD	,87	-,04	-,22
YV	,09	,73	,16
YY	,01	,04	,96
BÇ	-,07	,80	-,10
TB	-,83	-,06	-,27

Tablo 7'de görüldüğü gibi, YD ve TB değişkenlerinin (ters yönde olmak üzere) Faktör 1'de yüklendiği görülmektedir. "Yalnızca dikkatsizlik yüzünden hata yapılması (YD= ,87)" bir bakıma problemin doğru biçimde çözüldüğünü göstereceği ve "çözmeye bile kalkışmadan boş bırakılan yanıtlar (TB= -,83)" ile ters yönde, aynı özelliği taşıdığı düşünülürse 1. faktörün "**Matematiksel Muhakeme Gücü**" olarak adlandırılabilir. "Çoktan seçmeli testin çeldiricileri arasında yer alan bir sonuca ulaşılan (YV= ,73)" ve "çözülmeye çalışılan ancak çözülemeyen boş bırakılan (BÇ= ,80)" maddelerin ortak olarak yüklendiği 2. faktörün adı "**Alışılmış Yollardan Çözüm Yaklaşımı**" olabilir. 3. faktör, yaygın biçimde düşülebilecek hatalar göz önüne alınarak hazırlanmış olan "çeldiriciler dışında sonuç bulunan yanıtlar (YY= ,96)"ın karakteristiği olarak "**Farklı Çözüm Yollarına Yönelme**" olarak adlandırılabilir.

Açık uçlu sınavların verdiği "doğru olmayan yanıtları ayrıntılarıyla inceleme olanağı" yardımıyla bu yanıtların yapısına ilişkin olarak ulaşılan sonuç, yapılan hataların yapısal farklılıklar gösterdiğini açıkça ortaya koymaktadır. Bu durum, yapılan ölçmenin güvenirliliği kadar geçerliliğini de yakından ilgilendirir ve geçerlik-güvenirlilik ilişkisine yeni bir boyut katabilir.

### 2.4. Açık uçlu sınavda yanıtlayıcıların verdikleri doğru olmayan yanıtlar yapılan hata türlerinin düzeyleri ile ne kadar açıklanabilmektedir?

Her bir hata türünün, doğru olmayan yanıtlardaki varyansın yüzde kaçını açıkladığını görmek için, hata türlerine ilişkin determinasyon katsayıları hesaplanmış ve elde edilen bilgiler Tablo 8'de gösterilmiştir.

Tablo 8 : Doğru Olmayan Yanıtlar İçin Determinasyon Katsayıları

Hata Türü	R	Deter. Kat (R <sup>2</sup> )	Düzeltilmiş R <sup>2</sup>	Hata Payı
YD	0,31	0,10	0,08	4,52
YA	0,23	0,05	0,03	4,63
YB	0,24	0,06	0,04	4,61
YY	0,11	0,01	-0,01	4,73
TB	0,81	0,66	0,65	2,79
BÇ	0,41	0,17	0,15	4,33

Tablo 8 incelendiğinde, doğru olmayan yanıtlardaki varyansın % 65'inin "tamamen boş bırakılmış yanıtlar"dan kaynaklandığı anlaşılmaktadır. Bu bulgu, açık uçlu sınavda soruların neden boş bırakıldıklarının daha ayrıntılı olarak incelenmesi gerektiğini ortaya koymaktadır. Bu durum, 2.2. alt-problemdede de belirtildiği gibi açık uçlu sınavda "çekinilen" ya da "çözmeye üşenilen" soruların önemli bir kısmının çoktan seçmeli testte yanıtladığı, hatta çoğunda doğru yanıt ulaşıldığı da göz önüne alınarak bir başka araştırmanın konusu yapılabilir.

Yanlış yanıt türlerinin kendi içinde, yanlış yanıtlardaki varyansın yüzde kaçını açıkladığını görmek için hesaplanan determinasyon katsayıları ise Tablo 9'da gösterilmiştir.

Tablo 9 : Yanlış Yanıtlar İçin Determinasyon Katsayıları

Hata Türü	R	Deter. Kat (R <sup>2</sup> )	Düzeltilmiş R <sup>2</sup>	Hata Payı
YD	0,49	0,24	0,22	1,87
YA	0,47	0,23	0,21	1,89
YB	0,46	0,21	0,19	1,91
YY	0,63	0,40	0,39	1,66

Tablo 9'dan, yanlış yanıtlanan soruların varyansının en büyük bölümünün "ulaşılacak yanlış sonuç çoktan seçmeli testte olası hatalar içinde öngörülmemiş olan sorular"dan kaynaklandığı (% 39); dikkatsizlik (% 22), aynı seçeneği (% 21) ya da bir başka seçeneği (% 19) işaretlemenin ise birbirine oldukça yakın paylara sahip oldukları anlaşılmaktadır.

Açık uçlu sınavda soruyu yanlış olarak yanıtlama davranışlarından, "öngörülmeyen bir hataya düşme"nin neden hataların diğerlerinden fazla bölümünü açıkladığı ancak bu hataları yakından inceleyerek doğru biçimde yorumlanabilir. Yine de, problemi açık uçlu olarak çözerken tahmin edilebilecek bilgi ve mantık hataları dışındaki nedenlerin daha etkili olduğunu bilmenin, bazı önlemler olarak hataların düzeltilebileceği umudunu da azaltacağı söylenebilir. Ayrıca bu bulgu, belki de bulunduğu yanlış sonuç seçenekler arasında bulunmadığından uyarılan bazı deneklerin daha sonra büyük ölçüde doğru yanıt ulaştıkları ve çoktan seçmeli testin bu nedenle daha kolay bulunduğu görüşünü de destekler niteliktedir.

### 3. Matematikte problemlerin çoktan seçmeli test yerine açık uçlu sorular biçiminde sorulması geçerliği nasıl etkilemektedir?

Açık uçlu sınavın geçerliğinin çoktan seçmeli testin geçerliği ile karşılaştırılmasına ilişkin bu alt problemle ilgili olarak şu iki soruya yanıt aranmıştır:

### 3.1. Çoktan Seçmeli Test Ve Açık Uçlu Sınava Verilen Yanıtlar Arasındaki İlişki nasıldır?

Çoktan seçmeli test ve açık uçlu sınava verilen yanıtlar arasındaki ilişki Tablo 10'da gösterilmiştir.

Tablo 10 : Sınavlardaki Yanıtların Dağılımı

İkisinde de doğru		İkisinde de yanlış		Biri doğru, biri yanlış	
f	%	f	%	f	%
1127	43,35	766	29,46	707	27,19

Tablo 10'da görüleceği gibi her iki sınava da verilen yanıtlar arasında her ikisine de doğru yanıt verilenlerin sayısı tüm yanıtların yaklaşık yarısını (% 43) oluşturmaktadır. İkisinde de yanlış olanlarla (%29) birlikte her iki sınavda sorulara verilen yanıtların % 72'si aynı biçimdedir. Bu durum iki sınav arasında oldukça yüksek bir korelasyon olduğunu gösterir.

İki sınav arasında, tüm yanıtları dikkate alması ve bunların ilişkilerini en doğru biçimde ortaya koyması açısından Tetrakorik Korelasyon Katsayısı hesaplanmış ve 0,69 olarak bulunmuştur. Bu sayı, daha önce geçerliği ve güvenilirliği hesaplanmış ve yüksek bulunmuş bir ölçme aracı olan çoktan seçmeli test "ölçüt" alındığında açık uçlu sınavın geçerlik katsayısı olarak kabul edilebilir.

### 3.2. Çoktan seçmeli testteki maddelerin geçerlik indeksleri, aynı maddeler açık uçlu sorulduğunda nasıl değişmektedir?

Testten alınan puanlar ölçüt kabul edildiğinde, her madde için "madde-test korelasyonu" (madde ayırıcılık gücü indeksi) maddenin geçerlik katsayısı olarak kabul edilmektedir (13). Tablo 11'de her iki sınav türü için soruların geçerlik katsayıları ve karşılıklı olarak korelasyonları gösterilmiştir.

Tablo 11 : Sınav Türlerine İlişkin Soruların Geçerlik Katsayıları

Soru No	Ç.S.T. Mad. Geç. Kat.	A.U.S. Sor. Geç. Kat.	Ara. Kore. Kat.
1	0,54	0,84	0,75
2	0,45	0,56	0,49
3	0,49	0,57	0,48
4	0,53	0,40	0,79
5	0,70	0,65	0,58
6	0,58	0,48	0,77
7	0,52	0,41	0,65
8	0,65	0,74	0,66
9	0,60	0,44	0,49
10	0,63	0,59	0,82
11	0,44	0,55	0,27
12	0,46	0,62	0,52
13	0,59	0,61	0,71
14	0,75	0,64	0,47
15	0,40	0,47	0,42
16	0,46	0,64	0,64
17	0,59	0,61	0,65
18	0,55	0,35	0,18



Soru No	C.S.T. Mad. Geç. Kat.	A.U.S. Sor. Geç. Kat.	Ara. Kore. Kat.
19	0,51	0,50	0,66
20	0,63	0,67	0,84
21	0,71	0,56	0,17
22	0,71	0,63	0,70
23	0,68	0,67	0,53
24	0,40	0,42	0,40
25	0,66	0,71	0,64
26	0,55	0,45	0,57
Medyan	0,57	0,58	0,61

Tablo 11'in ilk sütunda, yöneltilen sorular çoktan seçmeli test maddesi biçimindeyken verilen yanıtların madde geçerlik indeksleri bulunmaktadır. Buradan da görülebileceği gibi çoktan seçmeli testte yer alan maddelerin geçerlik indeksleri oldukça yüksektir.

Tablo 11'in ikinci sütununda da açık uçlu soruların geçerlik indeksleri yer almaktadır. Tablodan da görülebileceği gibi, aynı sorular açık uçlu sorulduğunda geçerlik indekslerinde önemli bir değişiklik gözlenmemektedir. Bir karşılaştırma yapmak için merkezi eğilimlerine bakılırsa, çoktan seçmeli test maddelerinin geçerlik indekslerinin medyanı 0,57 iken açık uçlu sınavdaki soruların geçerlik indekslerinin medyanının 0,58 olduğu görülür. Katsayılar arasında bir fark görülmediği gibi, 26 sorunun 13'ünde soru açık uçlu sorulduğunda, diğer 13'ünde ise çoktan seçmeli olduğunda daha yüksek bir geçerlik katsayısı bulunmuştur. Bu sonuç, soru tipinin geçerliği etkilemediğini açıkça göstermektedir. Daha önceki altproblemlerde elde edilen sonuçların sınavların yapısal olarak farklı olduklarını çeşitli biçimlerde gösterdikleri göz önüne alındığında bu bulgu, aradaki farkın geçerliği etkilemediği biçiminde yorumlanabilir. Çoktan seçmeli testte ve açık uçlu sınavda daha yüksek (ya da daha düşük) geçerlik indeksine sahip olan sorular üzerinde yapılan incelemede de ortak bir özellik gözlemlenmektedir.

İki sınav türünün madde geçerlikleri arasında önemli farklar bulunmadığına göre iki sınav türüne verilen yanıtlar arasındaki farklılık nasıl açıklanabilir? Yapısal olarak farklı olan bu sınav türlerinden elde edilen bilgileri birbirleri için geçerlik ölçütü olarak almak sakıncalı olabilir mi? Bu ve bunun gibi sorulara yanıt aramak için her maddeye verilen yanıtlar ayrı ayrı karşılaştırılarak korelasyon katsayıları hesaplanmış ve Tablo 11'in son sütununda gösterilmiştir. Tablodan bu katsayıların da oldukça yüksek olduğu görülmektedir. Aralarındaki korelasyon 0,40'ın altına düşen madde sayısı yalnızca 3'dür. Ayrıca bu katsayıların medyanı, sınav türlerinin kendi içlerinde hesaplanan geçerlik katsayılarından da yüksek ve 0,61 dir. O halde sınavlarda verilen yanıtların türünde (doğru, yanlış, boş) sayısal olarak bir fark görülse de ölçülen nitelikte bir fark ortaya çıkmamaktadır.

Sonuç olarak, iki sınav türüne verilen yanıtlar çeşitli açılardan farklılıklar gösterebilir de bu farkların geçerliği etkilemediği anlaşılmaktadır.

#### 4. SONUÇ

Geçerlik ve güvenirlik açısından çoktan seçmeli test ve yazılı yoklamaların karşılaştırılması eğitimciler tarafından yıllardır yapılmaktadır. Bu araştırmaların bir kısmında (yazılı yoklamada geçerlik lehine ve güvenirlik aleyhine) farklar bulunurken bazılarında önemli farklar olmadığına karar verilmiştir. Bu çalışmanın amacı bütün bu araştırmalara bir yenisini eklemek değil, bu güne kadar pek üzerinde durulmamış olan bir durumu, yanıtlayıcıların bu soru tipleri karşısında nasıl davrandıklarını incelemektir. Eğer bu davranışlar iyi analiz edilebilirse sınav türlerini avantaj ve dezavantajlarıyla daha iyi tanıyarak hangi koşulda hangisini kullanmak gerektiğine daha doğru karar vermek olanağı olur.

Yapılan çalışmada ulaşılan sonuçlar şöyle sıralanabilir:

1. Kullanılan ölçme araçlarının her ikisinin de geçerlik ve güvenirlikleri yüksek bulunmuştur. Böyle ölçeklerden elde edilen ölçme sonuçlarının da geçerli ve güvenilir olması beklenir.

2. Açık uçlu sınav, çoktan seçmeli testten daha zor bulunmuştur ki bu sonuç literatürle de uyumludur.

3. Araştırmaya başlarken öngörüldüğü gibi, çoktan seçmeli testteki problemler açık uçlu olarak sorulduğunda yanıtlayıcı davranışlarında bir farklılaşma görülmektedir. Çoktan seçmeli testte boş bırakılan ve yanlış işaretlenen yanıtların oranı aynı iken (% 19), bu oran problemler açık uçlu sorulduğunda boş bırakılanlar lehine (% 32) değişmektedir.

4. Açık uçlu sınavda daha çok yanıtın boş bırakılmasının nedenlerinden biri "motivasyon azlığı" olarak düşünülebilir. Motivasyon azlığının etkili olması durumunda güvenirliğin düşük olması beklenir. Oysa araştırmada kullanılan ölçme araçlarının ikisi de oldukça yüksek güvenirlik katsayılarına sahiptir. Yine de bu durum bir başka araştırmanın konusu yapılabilir.

5. Açık uçlu sınavda doğru olmayan yanıtlar için, deyim yerindeyse "hata analizi" yapılmış, hatalar gruplandırılarak çoktan seçmeli testteki sonuçlarla karşılaştırılmış ve ilginç sonuçlara ulaşılmıştır. Örneğin,

i, Açık uçlu sınavda boş bırakılan soruların üçte biri testte doğru olarak yanıtlanmıştır.

ii, Açık uçlu sınavda yanlış olarak yanıtlanan soruların yarısının ise çoktan seçmeli testte doğru yanıtının bulunduğu saptanmıştır. Bu tür yanıtların içinde "dikkatsizlik" nedeniyle yapılan yanlışlar (% 38) kadar ulaşılan yanlış sonuç seçenekler arasında bulunmayanlar (% 40) vardır.

Bütün bu bulgular ancak yanıtlayıcı davranışlarının sınav türüne göre değişmesiyle açıklanabilir.

6. Yapılan çalışmada ulaşılan en önemli sonuçlardan biri de açık uçlu sınavdaki doğru olmayan yanıtların, "matematiksel muhakeme gücü", "alışılmış yollardan çözüm yaklaşımı" ve "farklı çözüm yollarına yönelme" olarak adlandırılabilir üç boyut içerdiğinin saptanmasıdır. Bu durum, bir ölçüde yanıtlayıcıların açık uçlu sınavlardaki davranış değişikliklerini de açıklamaktadır.

7. Araştırmada ulaşılan bir başka önemli sonuç da, açık uçlu sınavda yanıtlayıcıların verdikleri doğru olmayan yanıtlardaki varyansın asıl kaynağının "tamamen boş bırakılan yanıtlar" olduğunun anlaşılmasıdır. Açık uçlu sınavda tamamen boş bırakılan soruların yarısının doğru yanıtının çoktan seçmeli testte bulunduğu da göz önüne alınarak bu durumun ayrı bir araştırmada incelenmesi gerekebilir.

8. Çoktan seçmeli testte yer alan maddelerin geçerlik indekslerinde, problemler yanıtlayıcıya açık uçlu olarak verildiğinde önemli bir değişme gözlenmemiştir. Bu sonuç yapısal olarak farklı da olsalar sınav türlerinin geçerlikleri arasında önemli bir fark olmadığını bir kez daha göstermektedir.

#### KAYNAKÇA

- (1) Turgut, M.F. (1977): Eğitimde Ölçme ve Değerlendirme Metotları. Ankara . Nüve Matbaası : 27
- (2) Cooper,P.L. (1984): The Assessment of Writing Ability: A Review of Research. Information Analyses General (070). New Jersey US.
- (3) Lockledge,A. and Hayn, J.A. (1992): Preparing Teachers for Open Response Testing. Reports Descriptive (141). North Carolina US.
- (4) Tuckman,B.W. (1991): Evaluating the Alternative to Multiple Choice Testing for Teachers. Contemporary Education. v62 n4 p299-300 Sum 1991
- (5) Bridgeman,B. ve Lewis, C. (1974): The Relationship of Essay and Multiple-Choice Scores with Grades in College Courses. Journal of Educational Measurement. v31 n1 p37-50 Spr 1994
- (6) Tuckman,B.W. (1991): Evaluating the Alternative to Multiple Choice Testing for Teachers. Contemporary Education. v62 n4 p299-300 Sum 1991
- (7) Turgut, M.F. (1977): Eğitimde Ölçme ve Değerlendirme Metotları. Ankara . Nüve Matbaası : 27
- (8) Tuckman,B.W. (1993): The Essay Test: A Look at the Advantages and Disadvantages. NASSP Bulletin. v77 n555 p20-26 Oct 1993.
- (9) Turgut, Fuat. (1971): Şans Başarısının Test Puanlarına Etkisi. Ankara. Orta Doğu Teknik Üniversitesi Yay.
- (10) Umay, A. (1995): Seçmeli Testlerde Yanıtlayıcı Davranışları ve Şans Başarısının Elimine Edilmesi İşlemlerine İlişkin Bazı Öneriler. II. Eğitim Bilimleri Kongresi. Ankara.
- (11) Foos,P.W. (1992): Test Performance as a Function of Expected Form and Difficulty. Journal of Experimental Education. v60 n3 p205-11 Spr 1992
- (12) Umay, A. (1997): Matematik Eğitimi ve Ölçülmesi. Eğitim Fakültesi Dergisi 12, s145
- (13) Turgut M.F. (1978): Madde ve Test İstatistikleri. Ders Notu 1. (Yayımlanmamış)