

**Yabancı Öğrencilerin Türkiye’de Üniversiteye Giriş Sınavı Başarılarının  
Karşılaştırılmalı Analizi**

Erhan ÇATAL-Alper SİNAN- Ömer Faruk KINA- Alper TOSUN



**Yabancı Öğrencilerin Türkiye’de Üniversiteye Giriş Sınavı Başarılarının**

**Karşılaştırılmalı Analizi**

 <https://doi.org/>

 10.55107/turksosbilder.1223854

**Erhan ÇATAL**

ecatal07@gmail.com

**Alper SİNAN**

Akdeniz Üniversitesi,

alpsin25@gmail.com

0000-0001-6632-5500

**Ömer Faruk KINA**

Akdeniz Üniversitesi,

0000-0001-6427-3700

omerfkina@gmail.com

**Alper TOSUN**

Akdeniz Üniversitesi,

0000-0001-9715-5209

alpertosun.003@gmail.com

**Öz**

Bu çalışmanın amacı Türkiye’de düzenlenen uluslararası öğrencilerin üniversiteye giriş sınavının madde ve test istatistiklerinin incelenerek analizlerinin yapılması ve bireylerden elde edilen puanların gruplar arasında farklılaşıp farklılaşmadığını saptamaktır. Araştırma Antalya ili örnekleminde yapılmış ve Akdeniz Üniversitesine devam eden yabancı uyruklu öğrencilerin katılımıyla oluşturulmuştur. Çalışmada öncelikle 2019 yılında yapılan Yabancı Öğrenci Sınavının madde analizleri incelenmiş, bu

kapsamda da madde güçlük değerleri, madde ayırt edicilik değerleri, maddelerin dağılımları ve madde güvenilirlik katsayıları çalışmada değerlendirilmiştir. Çalışmanın devamında ise, Yabancı Öğrenci Sınavının test istatistiklerine ilişkin sonuçlar incelemiştir. Bu sonuçlar kapsamında maddelerin çeldirici analiz sonuçlarına yer verilmiş, sınava katılan bireylerin dil değişkenine yönelik karşılaştırma yapılmıştır. Genel olarak araştırma kapsamında maddelerin analiz sonucuna göre, niteliklerine göre önerilerde bulunulmuştur.

**Anahtar kelimeler:** Başarı testi, Madde Analizi, Yabancı Öğrenci Seçme Sınavı  
**Abstract**

### **The Success of Foreign Students in the University Entrance Exam in Turkey Comparative Analysis**

The aim of this study is to analyze the item and test statistics of the university entrance exam of international students held in Turkey and to determine whether the scores obtained from the individuals differ between the groups. The research was conducted in the sample of Antalya province and was formed with the participation of foreign students attending Akdeniz University. In the study, first of all, the item analyzes of the Foreign Student Exam held in 2019 were examined, and in this context, item difficulty values, item discrimination values, distribution of items and item reliability coefficients were evaluated in the study. In the continuation of the study, the results of the test statistics of the Foreign Student Examination were examined. Within the scope of these results, the results of the distracting analyze of the items were included, and a comparison was made for the language variable of the individuals who took the exam. In general, within the scope of the research, suggestions were made according to the results of the analysis of the items and their qualities.

**Keywords:** Achievement test, Item Analysis, Foreign Student Selection Exam

### **Giriş**

Dijitalleşen ve küreselleşen dünyada bilgiye kolay ulaşım sağlanması, bilginin kolay yayılması uluslararası sınavların uygulanmasını kolaylaştırmıştır. Uluslararası sınavlar da eğitim sisteminin önemli bir boyutu olarak yerini almaya devam etmektedir (Mullins & Roberts, 1996). Diğer taraftan dünyada dil ve kültür çeşitliliğinin fazla olması uluslararası sınavların düzenlenmesi zorlaştırmakta, karşılaştırma yapılmasında güçlük çekilmesine yol açmaktadır. Dünya genelinde uluslararası sınavların düzenlenmesinde yapılan çalışmalar günden güne artmakta ve yaşanan bu zorlukların azaltılması için farklı araştırmalar yapılmaktadır. Eğitim seviyelerinin karşılaştırılmasında da uluslararası sınavlar önemli bir konumda yer almaktadır. Sınav sistemlerinin incelenmesi için yapılan çalışmalara bakıldığında değerlendirme ve uygulama biçimlerine odaklanılmış durumdadır (Knight, 2002).

Ölçme ve değerlendirmenin ilkeleri olduğu kadar ulusal ve uluslararası sınavların da kendi bağlamında ilkeleri mevcuttur. Eğitim seviyelerini kıyaslamak, eğitim standartlarını belirlemek ve norm çalışmalarında bulunmak bu ilkelerden birkaç tanesidir. Amaçlarına bakıldığında ise, sınav sisteminin ardından eğitimi ve eğitim programlarını başlatma sürecini belirlemesidir. Ölçme ve değerlendirme dijitalleşmenin arttığı bu dönemde özellikle geleneksel bir yapıdan modern bir yapıya geçiş aşamasında farklı metotların denendiği bir alandır. Bu metotların denendiği dönemde, kuramların güncellenmesi, yeni kuramların ortaya çıkması ve yöntemlerin değişmesi modern ilkelere geçişin ilk aşamalarını oluşturmaktadır. Çalışmaları yapılan yeni yöntem ve tekniklere ilişkin geleneksel yöntem olan sınav sistemi test tekniklerinin de artık bir geçiş sürecinde olduğundan bahsedebiliriz. Bu süreçte ölçme araçlarının denenmesi, ölçme araçlarının istenilen hedefe yönelik sonuçlar vermesi de en büyük amaçlardan birisidir. Ulusal ve Uluslararası sınav sistemlerinin de günden güne değişmesi, ülkelerin eğitim standartlarını belirlemede önemli bir yer edinmektedir. Aynı zamanda bu standartların yanında çağı önden yakalamanın da farklı bir boyutudur.

### 1.1 Ölçme Temel Kavramlar

Eğitim ve öğretim bireydeki var olan istedik davranışların değiştirilmesini gaye edinmektedir. İstenilen davranışların oluşması ise birçok ilkenin yerine getirilmesi, etkinliklerin yapılması ve buna yönelik programların yapılmasıyla oluşmaktadır (Baykal, 2015). Eğitim sistemlerinde de değerlendirme kavramı buradan yola çıkmaktadır. İstenilen davranışların ve eğitim programlarına uygun bireylerin yetiştirilmesinin en önemli ilkelerinden birisi de değerlendirme boyutunun sağlıklı ve doğru bir şekilde yapılmasıdır. Değerlendirme de tam anlamıyla bir karar verme sürecidir. Bu süreçte bireylerin başarı durumunun gerçekçi sonuçlar vermesinde dikkatli çalışmak gerekmektedir. (Baykul ve Turgut, 2015).

Değerlendirme sürecine ilişkin karar verme boyutunun yanı sıra ölçme işlemlerinde de hata olup olmamasının da tespit edilmesi gerekmektedir. Ölçme sonuçlarına ilişkin hata türlerine baktığımızda, hatalar, sabit, sistematik ve tesadüfi olmak üzere üçe ayrılmaktadır. Sabit hata, hatanın miktarı, kaynağı ve yönünün belli olduğu hata türüdür. Sistematik hata, sistematik bir şekilde yönü ve miktarı belirli olan hata türüdür. Tesadüfi hata ise kaynağı, yönü ve miktarı belirli olmayan hatadır (Baykul, 1999). Tesadüfi hatalar ayrıca güvenilirlik durumunu etkilemekte iken sistematik hata ise geçerliliği etkilemektedir. Özellikle sistematik hatalarda ortaya çıkan yanlışlık, yapı geçerliliği olumsuz etkilemektedir. Buna yönelik de ölçme sonuçlarının tesadüfi hatalardan ne kadar arınık ise ölçme sonucu o derecede güvenilirdir ve ölçme sonucunun güvenilirliği aynı zamanda duyarlılığın da olduğunu göstermektedir (Baykal,2015).

Ancak güvenilirliği de belirlemenin yolları bulunmaktadır. Ölçme sonuçlarının tutarlılık göstermesi için örneğin art arda ölçümler yapılması tutarlılığı olumlu yönde etkiler. Güvenirliği belirlemede kullanılan yöntemlerden birisi olan test tekrar test yöntemi, belirli bir gruba aynı testin daha

sonra belirlenmiş zamanlarında da uygulanarak alınan sonuçlardaki ilişkiye bakılmasıyla oluşur. Alınan sonuçlardaki ilişki eğer tutarlı ise o teste güvenilir diyebiliriz.

Bir diğer tutarlılık belirleme yöntemi ise eş değer yani paralel formlardır. Paralel formlarda testlerin madde sayısı, maddelerin gücüğü, maddelerin türü benzer olmalıdır bu yüzden her zaman paralel formların hazırlanması kolay değildir. Burada ölçme aracı hata kaynağı olarak söylenebilir. Paralel formlar her zaman istenilen varsayımı sağlamayabilir bu yüzden güvenilirlik belirleme yöntemlerinden biri olan eş değer yarılar (iki yarı güvenliği) formunu uygulamak daha iyi olabilir. İki ölçme sonucu arasındaki korelasyon katsayısı eş değer yarılar güvenilirliğini verir (DeVellis, 2016).

Güvenilirliği etkileyen başka bir durum ise puanlayıcı güvenirligidir. Bu yöntemde bireyin başarısının birden fazla puanlayıcı tarafından puanlanması ile elde edilen sonuçlar arasındaki ilişkiye bakılır. Burada ölçme sonucunun tutarlılığı önemli olduğu gibi, puanlayıcı hata kaynağı olarak görülür (Atılğan, Aydın ve Kan, 2017).

## 1.2. Ölçmede Kuramlar

Eğitim sistemlerini ve hedeflerini incelediğimizde önde gelen en önemli amaçlardan birisi de bireyin başarısının gerçekçi ve doğru olmasının yerine getirilmesi olgusudur. Neden olarak ise öğrencinin yeteneğinin net olarak ölçülmesi çok da mümkün olmamaktadır ve gizil durumdadır. Öğrencinin yeteneğinin net olarak kestirilememesi ve gizil bir halde olması ölçme aracını uygulayan kişinin, ölçülen psikometrik özelliğın ve bir çok farklı hatalar nedeniyle gerçek ve doğru değere ulaştırılmasını zorlaştırmaktadır (Çelen ve Aybek, 2013). Ortaya çıkan bu güçlükler nedeniyle de öğrenci ölçme araçlarında yer alan maddelerde hata paylarını da göze almak gerekmektedir (Erkuş, 2003). Sınav kavramının ortaya çıktığı günden bu yana günümüze kadar yapılan ölçme ve değerlendirme uygulamaları geleneksel sınav yöntemiyle uygulanmaktadır. Ancak sınavla ölçmede de soruların aşağıdaki parametreleri ne oranda bünyesinde barındırdığı ve bu durumun nasıl ölçülebileceği konusu gündeme gelmektedir;

- Ayırt edicilik,
- Yetenek (Zorluk)
- Şansla doğru cevaplanma ihtimali,
- Sorunun tam doğru yapılma ihtimali,

Çünkü tüm bu özellikleri bünyesinde barındırmayan bir soru yapısı öğrenci başarısını doğru ve adaletli bir şekilde ölçemeyecektir. Bu kapsamda özellikle sınavla ölçüm yaparken; “Sorulan sorular öğrenci başarısını ne oranda doğru ölçmektedir?”, sorusu öncelikle yanıtı bulunması gereken konulardandır. Bu nedenden dolayı da araştırma kapsamında açık ve uzaktan öğrenmedeki durumu görebilmek adına ölçme ve değerlendirmede iki temel kuram olan “Klasik Test Kuramı=KTK” ve “Madde Tepki Kuramı=MTK/Item Response Theory=IRT” üzerinden analizlerle soruların öğrenci başarısını temsil yeteneği kestirilmeye çalışılacaktır. Geçmiş eskilere dayanan KTK'nın alana faydası

daha çoktur. Ancak ölçme problemlerini çözmede yararı daha çok olduğu için MTK günümüzde daha çok tercih edilir. Kolay uygulanabilmesi açısından özellikle tercih edilen klasik test kuramının da avantajlı olduğu alanlar da elbette ki bulunmaktadır. (Hambleton ve Jones, 1993, Baykul 2000). Crocker ve Algina (2008) yaptıkları araştırmaya göre ölçme aracını geliştiren kişilerin maddelere yönelik farkların belirlemede dikkatli olmaları gerektiğini ve bu nedenle de farklı yetenek boyutlarına sahip öğrencilerin performans boyutu hakkında bilgi vermediğini ifade etmişlerdir. Ortaya çıkan bu sınırlılıklara da bakarak, belirlenen bir maddenin bütün eylemlerinin görüntüsünü görmek için Madde Tepki Kuramından tercih edilmesinin daha uygun olduğu vurgulanmıştır. Bu sayede de hangi yöntemle hesaplanan puanın daha doğru bir kestirim yapabildiği, yani KTK'ya göre hesaplamayla MTK 2PL, MTK 3PL modeline göre hesaplama arasında anlamlı bir fark olup olmadığı ortaya konulmuş olacaktır.

### 1.3 Başarı Testi

Eğitim programları ve uygulamaları, bireyin kendini aldığı eğitimlerle geliştirmesinde, gelişimindeki gereksinimlerini ve eksiklerini fark etmesinde, değerler eğitimine önem vermesinde, bireyde bilişsel farkındalıklar oluşmasında ve bireyin ön yargıların olmadığı bir düşünce sistemiyle karşılaştığı zaman ve en önemlisi de problemleri davranışlarla da karşılaştığında bu tip durumlara karşı olumlu davranışlar edinmesini sağlar (Atılğan, 2014).

Ölçme ve değerlendirmede kullanılan testler, belirli bir hedef, özellik doğrultusunda ve bu özellikler bağlamında gözlemlerde işlenen sistematik ölçme araçlarıdır. Ölçülmesi hedeflenen psikometrik özelliğin, hedefinin tanımlanması, test geliştirme ve ölçme sürecinin ilk aşamalarını oluşturmaktadır. Bireyin belirlenen bu eğitim sürecinde ya da sosyal şartlar altında ne kadar öğrendiği ölçen ölçme araçları da bir nevi başarı testleridir. Başarı testleri özellikle öğrencilerin ne öğrendikleri, öğrendikleri ile ilgili eksikleri ve geçmişteki bu öğrenmelerin geleceğe yol belirlemesi adına ortaya atılmıştır. Örnek vermek gerekirse, ortaokul öğretmenin bireylerin öğrendiği konular üzerine performans değerlendirmelerine ilişkin oluşturulan testler, başarıyı ölçmeye yönelik testlerdir. Akademik başarı, bireyin ders başarısı, öğrenme başarısı, yıllık, dönemlik başarı, ünite sonu başarıları gibi farklı biçimlerde ve kavramlarla da tanımlanabilmektedir (Brookhart, 2009).

Başarının tanımını incelediğimizde, bilişsel açıdan baskın olan gizil bir özelliktir. Bunun yanı sıra da başarı testleri, maksimum performans testleri olarak da literatürde yerini almaktadır. Başlıca başarı testlerini incelediğimizde;

- Yazılı Yoklama,
- Sözlü Yoklama,
- Çoktan Seçmeli Test,
- Kısa Cevaplı Test,

- Boşluk Doldurmalı Test,
- Eşleştirmeli Test,
- Doğru-Yanlış Testleri,
- Kompozisyon Testleri,
- Performans Değerlendirme,

araç ve yöntemleridir. Başarı testlerini sınıf içerisinde de değerlendirecek olursak; çoktan seçmeli test, yapılandırılmış maddelerle oluşan testler; doğru-yanlış eşleştirme testleri; kısa cevaplı, boşluk doldurmalı testler yarı yapılandırılmış testler; kompozisyon testleri, sözlü ve yazılı yoklamaları ise yapılandırılmamış maddelerden oluşan test grubunda değerlendirilir.

#### **1.4 Test Geliştirme**

Ölçme araçlarının en önemli araçlarından birisi de nitekim testlerdir. Testler, kapsam ve amaca göre yoklaması yapılan öğrenme ürünlerinin kararının verildiği alandır. Buna göre yoklanacak davranışlara uygun olması için kullanılacak olan soru türü de önem arz etmektedir. Psikolojide ve eğitimde ölçme araçları kullanıldığı zaman uyarıcı sorular kullanılır, bu soruların kullanılmasındaki amaç gözlenemeyen davranışları ortaya çıkarmaktır. Uyarıcı soruların da hazırlanma aşamasında hedeflenen ve gözlenen davranışların ortaya çıkarılacağı şekilde yapılması gerekmektedir. Aynı zamanda bu davranışların ortaya çıkmasında en iyi rolü oynayan sorular oluşturulmalıdır (Özçelik, 2013). Hazırlanacak olan maddelerin her birisinin, ölçülmek istenen psikometrik özelliğe sahip olup olmadığına, bireyin ölçülmesi istenen gelişim düzeyine uygun olup olmadığına, dil açısından anlaşılır olup olmadığına da bakılmalıdır. Özellikle dil bilgisi ve soruların anlaşılabilirliği büyük önem arz etmektedir. Sorunun anlaşılmaması ve dil bilgisine uygun olmaması ölçülmesi istenen özelliğin güvenilirliğini düşürmekte ve maddenin ölçülmesi istenen hedefe aykırı şekilde oluşmasına yol açmaktadır. Genel olarak yapılan bu düzenleme literatürde madde düzenlemesi olarak yerini almıştır. Madde düzenlemesinde, düzenlenen maddelerin incelemelerinin yapılarak daha nitelikli hale getirilmesi sağlanmaktadır (Baykul, 2015). Madde düzenlemesi, testin geliştirildiği grupta maddelerin kalitesi, madde analizleri ve madde parametreleri ile ilgili kanıtların ve yorumların elde edildiği süreçtir. Ele alınan bu süreçte maddelerin niteliği ile alakalı birçok gösterge elde edilebilir ancak bu göstergelerin en çok kullanılan terimi, madde ayırt edicilik indeksi ve madde güçlük indeksidir.

Belirli bir hedef doğrultusunda gerçek testte kullanılacak maddeler, gerçek testin uygulanacağı örnekleme benzeyen ve benzer koşullar altında uygulanacak olan çok sayıda sorunun yer aldığı deneme formundan madde ve test analizleri sonucu seçilir. Madde seçimlerinde ilk olarak geçerlilik ve güvenilirlik durumlarına bakılır. Her ikisinin de yeterli düzeyde olduğu maddeler önceliklidir. Bunun yanında testin amacı doğrultusunda diğer özellikler de önem sırasına bakılarak madde seçimleri yapılır. (Özçelik,2014).

#### **1.5 Yabancı Öğrenci Sınavı**

ÖSYM tarafından düzenlenen uluslararası öğrenci seçme sınavı, alınan kararlar eşliğinde üniversiteler tarafından yapılmaktadır. Üniversiteler kendi bünyesinde oluşturduğu, ölçme, değerlendirme ve belgelendirme merkezleri tarafından bu sınavları oluşturur ve uygularlar. Özellikle son zamanlarda üniversitelerdeki yabancı uyruklu öğrenci sayılarının artması da ÖSYM tarafından bu kararın alınmasında etkili olmuştur. Yapılan sınav da Yabancı Öğrenci Sınavı olarak adlandırılmaktadır. Araştırmada bu sınavın maddelerinin incelediği birim Antalya ili Akdeniz Üniversitesi seçilmiştir. Akdeniz Üniversitesi de kendi bünyesine yabancı öğrenci almak için uluslararası sınavlar düzenlemekte ve bu sınavda Temel Öğrenme Becerileri testini kullanmaktadır. Uygulanan bu test, çoktan seçmeli 80 sorudan oluşmaktadır. Testin ilk 40 sorudan oluşan kısmı Genel Yetenek, diğer 40 soruluk kısım ise Matematik Testi olarak geçmektedir. Genel Yetenek Testinde şekle dayalı akıl yürütme ve sayısal alanda akıl yürütme soruları yer almaktadır. Bu bölümde daha çok şekil dizileri, şekil benzeşimleri, şekil matrisleri, simetri, dönme, katlama, kesme, blok sayma, küp ve zar soruları, katı cisimlerin özelliklerine yönelik sorular sorulmaktadır. Sayısal alanda akıl yürütme soruları daha çok grafik ve çizgiler ile ilgili sorulardır. Bunlar genellikle ikili işlemler, sayı dizisi tamamlama, sayı-şekil ilişkileri gibi sorulardan oluşmaktadır. Testin diğer bölümü olan 40 soruluk Matematik Testi ise temel işlemler, trigonometri, logaritma, karmaşık sayılar, fonksiyonlar, türev, integral, limit, süreklilik, analitik geometri gibi sorulardan oluşmaktadır. Sınav soruları öncelikle alan uzmanları tarafından Türkçe olarak hazırlanmaktadır. Uzmanlar tarafından gerekli düzenlemeler yapıldıktan sonra sınav son hâlini alır. Sınava en son halini ise çeviri yapılacak dillerin uzmanlarının yaptığı çeviriler verir. Farklı dillerde olan test maddeleri, Türkçe testine ait maddelerin farklı dillere çevrilmesinden elde edilen maddelerdir.

Temel Öğrenme Becerileri Testinden alınan sonuçlar 3 farklı puan türüyle ayrı ayrı hesaplanır. Bunlar; Genel Yetenek Puanı, Matematik Puanı, Eşit Ağırlık Puanlarıdır. Her adayın üç farklı puan türünde t ve z puanları hesaplanır ve sonuçları üç farklı puan türüyle açıklanır. Öğrencilerin aldıkları sınav puanları bir yıl geçerli olmaktadır.

## **Yöntem**

Bu çalışma ölçme ve değerlendirmede sık kullanılan nicel araştırma yöntemlerinden tarama modeli ile yapılmıştır. Tarama araştırması ya da modeli bir konu hakkında çalışmaya katılan bireyin ilgi, tutum, yetenek ve beceri gibi psikometrik özelliklerin ortaya çıkarılmasının hedeflendiği araştırmalardır (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel, 2017). Karasar (2018) tarama araştırmasını farklı bir şekilde tanımlamakta ve günümüzde ya da geçmişten var olan bir durumu olduğu haliyle sunma biçimi olarak yorumlamaktadır. Frankel ve Wallen (2006) ise tarama araştırmalarını, genel olarak bir topluluğun psikometrik özelliklerini yansıtabilecek örneklem bölümü, bireylere yöneltilen soru ve

ifadelerden elde edilen cevaplara dayalı yaklaşım, her bireyden değil onu yansıtacak bir gruptan veri toplama olmak üzere üç önemli özelliğe sahip olduğunu belirtmiştir.

## Verilerin Analizi

Araştırmada elde edilen veriler ÖSYM tarafından Akdeniz Üniversitesi bünyesine verilen 2019 yılında düzenlenen Yabancı Öğrenci Sınavı, soru kitapçıkları cevap anahtarlarıdır. Sınavda iki farklı kitapçık uygulanmış ve sınava katılan bireylerin %50.5'ini A Kitapçığı (n=2720) ve % 49.5'inin ise B Kitapçığını (n=2663) almıştır. Sınav 80 sorudan oluşmaktadır. İlk 40 soru mantık sorularından, son 40 soru matematiksel işlem becerilerini ölçmeye yönelik sorulardan oluşmaktadır. Çalışmada soruların test istatistikleri, güvenilirlik değerleri, Ham ve Z puanları son olarak ise madde güçlük ve ayırt edicilik analizleri yapılmıştır. Elde edilen bulgular Tablo1.'de sunulmuştur.

### Bulgular

Tablo.1 Akdeniz Üniversitesi Yabancı Öğrenci Sınavına Ait Test İstatistikleri

<b>Katılımcı Sayısı</b>	5383
<b>Minimum Puan</b>	3
<b>Maksimum Puan</b>	80
<b>Medyan</b>	51.0
<b>Ortalama</b>	50.0
<b>Standart Sapma</b>	17.88
<b>Varyans</b>	319.60
<b>Basıklık Katsayısı</b>	-0.179
<b>Çarpıklık Katsayısı</b>	-1.744
<b>Ortalama Madde Güçlüğü</b>	0.625
<b>Ortalama Ayırt Edicilik</b>	0.555
<b>Ortalama Çift Serili Korelasyon Katsayısı</b>	0.49
<b>Minimum Madde Güçlüğü</b>	0.23
<b>Maksimum Madde Güçlüğü</b>	0.94
<b>Minimum Ayırt Edicilik İndeksi</b>	0.15
<b>Maksimum Ayırt Edicilik İndeksi</b>	0.84
<b>Minimum Nokta Çift Serili Korelasyon</b>	0.16
<b>Maksimum Nokta Çift Serili Korelasyon</b>	0.68

Yabancı Öğrenci Sınavının test istatistiklerini incelediğimizde, sınava katılan öğrencilerin puanlarının 3-80 arasında değiştiği görülmektedir. Diğer sonuçlara baktığımızda basıklık çarpıklık katsayısının ise beklenen aralık olan -1,+1 aralığında olduğunu yani puanların normal dağılım gösterdiğini ilişkin kanıtların olduğu söylenebilir. Testin ortalama güçlük değerinin 0.625, testteki minimum madde güçlüğü'nün 0.23, maksimum madde güçlüğü'nün ise 0.94 olduğu ve nispeten kolay bir sınav olduğu söylenebilir. Testin ortalama ayırt edicilik değerinin 0.555 ve ortalama çift



serili korelasyon katsayısının ise 0.49 olduğu dolayısıyla başarılı ve başarısız öğrenci ayırt etmede güçlü bir sınav olduğu söylenebilir. Sınava ait aritmetik ortalamanın 50, medyan değerinin 51, standart sapmanın 17.88 ve varyansın ise 319.60 olduğu görülmektedir. Bu test istatistikleri sonuçlarına göre sınava giren öğrencilerin başarılı olduğu ve puanlardaki değişkenliğin fazla olduğuna yönelik istatistiksel çıkarım yapılabilir.

**Tablo. 2** *Yabancı Öğrenci Sınavı Güvenirlilik Değerleri*

<b>KR20</b>	0.96
<b>KR21</b>	0.95
<b>İki Yarı Güvenirliliği (Birinci-İkinci)</b>	0.77
<b>İki Yarı Güvenirliliği (Tek-Çift)</b>	0.93
<b>Ölçmenin Standart Hatası</b>	3.56

YÖS Öğrenci sınavı güvenirlilik değerleri incelendiğinde ise, KR20 değerinin 0.96, KR21 değerinin 0.95, testin iki yarıya bölünerek hesaplanan yarılar arası güvenirliliğin 0.77, tek ve çift maddeler arası güvenirliliğin ise 0.93 olduğu dolayısıyla testin iç tutarlılık anlamında güvenilir bir sınav olduğu söylenebilir. Diğer yandan ölçmenin standart hata değerinin ise 3.56 olduğu görülmektedir.

**Tablo.3** *Yabancı Öğrenci Sınavına ait Ham ve Z Puan Dağılımları*

<b>Ham Puan</b>	<b>Frekans</b>	<b>Z Puanı</b>	<b>Ham Puan</b>	<b>Frekans</b>	<b>Z Puanı</b>
<b>3</b>	1	-2,63	<b>44</b>	96	-0,34
<b>6</b>	1	-2,46	<b>45</b>	73	-0,28
<b>8</b>	1	-2,35	<b>46</b>	91	-0,22
<b>9</b>	2	-2,29	<b>47</b>	91	-0,17
<b>10</b>	3	-2,24	<b>48</b>	80	-0,11
<b>11</b>	5	-2,18	<b>49</b>	76	-0,06
<b>12</b>	11	-2,13	<b>50</b>	93	0,00
<b>13</b>	10	-2,07	<b>51</b>	77	0,06
<b>14</b>	15	-2,01	<b>52</b>	63	0,11
<b>15</b>	18	-1,96	<b>53</b>	95	0,17
<b>16</b>	27	-1,90	<b>54</b>	79	0,22
<b>17</b>	29	-1,85	<b>55</b>	72	0,28
<b>18</b>	33	-1,79	<b>56</b>	69	0,34
<b>19</b>	50	-1,73	<b>57</b>	75	0,39
<b>20</b>	36	-1,68	<b>58</b>	90	0,45
<b>21</b>	37	-1,62	<b>59</b>	92	0,50
<b>22</b>	75	-1,57	<b>60</b>	86	0,56
<b>23</b>	67	-1,51	<b>61</b>	94	0,62
<b>24</b>	60	-1,45	<b>62</b>	88	0,67
<b>25</b>	92	-1,40	<b>63</b>	101	0,73
<b>26</b>	85	-1,34	<b>64</b>	103	0,78

27	86	-1,29	65	98	0,84
28	70	-1,23	66	107	0,89
29	78	-1,17	67	122	0,95
30	82	-1,12	68	124	1,01
31	79	-1,06	69	118	1,06
32	85	-1,01	70	108	1,12
33	90	-0,95	71	121	1,17
34	71	-0,90	72	113	1,23
35	87	-0,84	73	125	1,29
36	81	-0,78	74	120	1,34
37	97	-0,73	75	102	1,40
38	83	-0,67	76	89	1,45
39	94	-0,62	77	70	1,51
40	92	-0,56	78	60	1,57
41	89	-0,50	79	33	1,62
42	75	-0,45	80	14	1,68
43	78	-0,39			

Yapılan sınavın ham ve z puan dağılımlarına ait değerler incelendiğinde öğrencilerin sınavdan 3-80 arasında puanlar aldıkları görülmektedir. Alınan ham puanların Z puanları dönüştürülmesi sonucu elde edilen değerlerin -2.63 ile 1,68 arasında değiştiği görülmektedir. Sınavdan düşük puan kişi sayısının az olduğu puan arttıkça kişi sayısında da gözle görülür bir artışın olduğu söylenebilir. Ancak en çok yığılmanın 63-75 puan arasında olduğu gözlemlenmiştir.

**Tablo.4** *Yabancı Öğrenci Sınavı Puanlarının Dil Değişkenine Ait Normallik Değerleri*

Değişken	Alt Grup	Basıklık	Çarpıklık	Basıklık Z	Çarpıklık Z	K-S/S-W
Dil	Almanca	0.27	0.39	0.30	0.87	0.734*
	Arapça	-1.11	0.15	3.86	1.00	0.000
	Fransızca	0.82	0.59	0.95	1.34	0.200*
	İngilizce	-0.87	-0.48	4.14	4.80	0.000
	Rusça	-0.40	0.59	1.21	3.47	0.000
	Türkçe	-1.21	-0.23	1.50	5.89	0.000

Uluslararası öğrenci sınavına yönelik dil değişkeni test istatistiklerini incelediğimizde karşımıza çıkan sonuçlarda basıklık ve çarpıklık katsayılarının tüm alt gruplarda -2, +2 arasında yer almaktadır. Diğer yandan ise basıklık Z ve çarpıklık Z katsayılarının Almanca ve Fransızca alt gruplarında beklenen değer olan -1.96, +1.96 arasında yer aldığı fakat Arapça basıklık Z, İngilizce basıklık Z ve çarpıklık Z, Türkçe çarpıklık Z katsayıların ise bu kriterleri karşılamadığı istatistiksel olarak hesaplanmıştır. Sonuçları incelemeye devam ettiğimizde ise K-S veya S-W testi sonuçları Almanca ve Fransızca alt gruplarının test sonuçlarının  $p > 0.05$  ile manidar sonuç vermediği, diğer alt gruplarda ise  $p < 0.05$  olarak manidar oldukları görülmektedir.

**Tablo.5** *Yabancı Öğrenci Sınavı Puanlarının Dil Değişkenine Ait Kruskal Wallis H Testi Sonuçları*

Alt Grup	N	Ortalama	Sıra Ortalaması	$\chi^2$	SD	p	Farklılık*
Almanca	26	28.92	896.15				4-1,4-2,4-3,
Arapça	284	45.07	2248.55				4-5,6-1,6-2,
Fransızca	28	15.25	133.88	368.47	5	0.000	6-3,6-5,2-1,
İngilizce	558	52.54	2899.71				2-3,2-5,5-3,
Rusça	218	32.40	1183.31				3-1
Türkçe	4258	51.23	2792.45				

\*1=Almanca, 2=Arapça, 3=Fransızca, 4=İngilizce, 5=Rusça, 6=Türkçe

Yabancı öğrenci sınavlarının Kruskal Wallis H testi sonuçları incelendiğinde alt grup sıra ortalamaları arasında manidar bir fark olduğu görülmektedir ( $\chi^2=368.47$ ,  $p<0.05$ ). Ortaya çıkan bu farklılığın hangi alt gruplar arasında olduğunu test etmek adına gerçekleştirilen Mann Whitney U testi sonuçlarına göre, İngilizce ve Türkçe sınava giren bireylerin sıra ortalamalarının Almanca, Arapça, Fransızca ve Rusça sınava giren kişilerin sıra ortalamalarından daha fazla olduğu görülmektedir. Benzer şekilde de Arap dilinden sınava giren öğrencilerin sıra ortalamalarının Almanca, Fransızca ve Rusça sıra ortalamalarından, Rusçadan sınava girenlerin sıra ortalamalarının Fransızca sınava girenlerin sıra ortalamalarından ve son olarak Fransızca sınava girenlerin sıra ortalamalarının Almanca sınava girenlerin sıra ortalamalarından daha fazla olduğu görülmektedir.

### Sonuç

ÖSYM tarafından üniversitelerin bünyesine dahil edilen Yabancı Öğrenci Sınavı Akdeniz Üniversitesi birimindeki 2019 yılında düzenlenen sınavın sonuçlarını incelediğimizde, madde güçlük değerlerinden çıkarılan sonuç, testte yer alan mantık sorularının genellikle kolay maddelerden, diğer yandan matematik sorularının ise orta güçlükte ve zor sorulardan oluştuğuna ilişkin istatistiksel olarak çıkarımda bulunabiliriz. Yapılan sınavdaki testlerde genel olarak zor maddenin az sayıda olduğu, sınavdaki maddelerin orta güçlükte olduğunu söyleyebiliriz. Uluslararası öğrencilerin eğitim düzeylerinin standartlarını belirlemek ve bu standartlara göre sorular hazırlamak, maddelerin uygunluğunu oluşturmak ulusal sınavlara kıyasla daha zorlu bir süreç gerektirmektedir. Sınavlarda uygulanan maddeleri incelediğimizde, testte yer alan matematik becerilerinin olduğu soruların hepsinin istatistiksel olarak da iyi derecede ayırt edici maddelerden oluştuğu ve diğer yandan da genel olarak mantık sorularının düşük düzeyde ayırt ediciliğe sahip olduğu istatistiksel olarak yapılan çıkarımlardan birisidir. Maddelerin varyans ve standart sapma değerlerine bakıldığında ise, çok zor ve çok kolay maddelerin değerlerinin diğer alanlara göre daha düşük olduğu saptanmıştır. Madde çarpıklık ve basıklık değerlerine baktığımızda ise dağılımların genel olarak simetrik bir yapıda olduğu görülmüştür. Fakat birçok madde de çarpıklık ve basıklık değerlerinin kolay bir şekilde cevaplanabilmesi adına yığılma olduğu da görülmektedir. Bu nedenle de belirtilen maddelerden ötürü başarının arttığı gözlemlenmiştir.

Madde güvenilirlik katsayıları incelendiğinde, matematik becerilerini ölçen soruların daha

güvenilir olduğu, buna kıyasla da mantık sorularındaki güvenilirlik değerlerinin düşük olduğu gözlemlenmiştir. Uluslararası öğrencilerin katılım gösterdiği yabancı öğrenci sınavının betimsel istatistiklerini incelediğimizde ise, bireylerin sınavdan 3-80 arasında puanlar aldıkları görülmektedir. Sonuçlarına baktığımız ham puanların Z puanları dönüştürülmesi sonucu elde edilen değerlerin ise -2.63 ile 1.68 arasında değiştiği görülmektedir. Sınavdan düşük puan kişi sayısının az olduğu puan arttıkça kişi sayısında da gözle görülür bir artışın olduğu söylenebilir. Fakat en çok yığılmanın 63-75 puan arasında olduğu istatistiksel olarak söylenebilir. Orta ve yüksek puana sahip öğrenci sayılarının epey fazla olduğu dolayısıyla dağılımın sola çarpık yani negatif çarpık bir dağılım gösterdiği söylenebilir. Puanların basıklığı ilişkin olarak yapılan incelemelerde kısmen de olsa basık bir dağılım olduğu söylenebilir. Sonuçların diğer test istatistiklerini incelediğimizde aritmetik ortalamanın 50, medyan değerinin 51, standart sapmanın 17.88 ve varyansın ise 319.60 olduğu görülmektedir. Buna göre sınava giren grubun başarılı ve puanlardaki değişkenliğin fazla olduğu söylenebilir. Teste ait basıklık çarpıklık katsayısının ise beklenen aralık olan -1,+1 aralığında olduğu dolayısıyla puanların normal dağılım gösterdiğini ilişkin kanıtların olduğu söylenebilir. Testin ortalama güçlük değerinin 0.625, testteki minimum madde güçlüğü'nün 0.23, maksimum madde güçlüğü'nün 0.94 olduğu ve nispeten kolay bir sınav olduğu söylenebilir. Testin ortalama ayırt edicilik değerinin 0.555 ve ortalama çift serili korelasyon katsayısının ise 0.49 olduğu dolayısıyla başarılı ve başarısız öğrenci ayırt etmede güçlü bir sınav olduğunu söylenebilir. Teste ait hesaplanan güvenilirlik değerleri incelendiğinde KR20 değerinin 0.96, KR21 değerinin 0.95, testin iki yarıya bölünerek hesaplanan yarılar arası güvenilirliğin 0.77'i tek ve çift maddeler arası güvenilirliğin ise 0.93 olduğu dolayısıyla testin iç tutarlılık anlamında güvenilir bir sınav olduğu söylenebilir.

Öğrencilerinde karşılaştırmalı analizinin incelendiği bu araştırmada, İngilizce ve Türkçe sınavına giren öğrencilerin ortalama puanlarının diğer dil sınavları Rusça, Arapça, Almanca ve Fransızca ortalamalarından düşük olduğu görülmüştür. Bu istatistiksel hesaplamaya göre de sınava giren bireylerin evrensel dil olarak kabul edilen İngilizce 'de ve sınava girdikleri ülke olan Türkçe 'de başarı gösterdikleri görülmüştür. Öğrencilerin, İngilizce' den yüksek puan almasının sebebi bu dilin evrenselliği Türkçe 'den yüksek puan almasının sebebi ise ülke diline göre hazır bulunuşluklarının olduğu görülebilir.

## Kaynaklar

Akdeniz YÖS-2019 Sınavı'nı Kabul Eden Üniversiteler. (2019, 28 Mayıs).

Baykul, Y. (1999). *İstatistik: Metodlar ve uygulamalar*. Anı Yayıncılık.

Brookhart, S.M. (2009). Educational assessment knowled geand skillsfor teachers. *Educational Measurement: issues and practice*, 30(1), 3-12.

Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2017). Bilimsel araştırma

yöntemleri. *Pegem Atıf İndeksi*, 1-360.

Çatal, E. Akdeniz Üniversitesi Yüksek Lisans Tezi, *Akdeniz Üniversitesi 2019 Yabancı Öğrenci Seçme Sınavının (YÖS) başarı puanları açısından değerlendirilmesi*

Frankel, J. R., & Wallen, N. E. (2006). *How to Design and Evaluate Research in Education*. (6. Baskı) McGraw Hill.

Knight, P. (2002). A systemic approach to professional development: learning as practice. *Teaching and teacher education*, 18(3), 229-241.

Lord, F. M., & Novick, M. R. (2008). *Statistical theories of mental test scores*. IAP.

Mullins, L., & Roberts, M. (1996). Assessment strategies: some comparisons between the UK and the US systems of higher education. *International Journal of Educational Management*.

Tekin, C., Gunes, G., & Colak, C. (2014). Adaptation of Problematic Mobile Phone Use Scale to Turkish: A Validity and Reliability Study [Cep Telefonu Problemlü Kullanım (Pu) Ölçeğinin Türkçe'ye Uyarlanması: Geçerlik Ve Güvenilirlik Çalışması]. *Med Sci Int Med J*, 3(3), 1361-1381.

Turgut, M. F., & Baykul, Y. (2015). Eğitimde ölçme ve değerlendirme (Yedinci baskı). *Pegem A yayıncılık, Ankara*

### **Extended Summary**

When we examine the results of the Exam for Foreign Students, which was included in the body of universities by OSYM in 2019, we can statistically infer that the result derived from the item difficulty values is that the logic questions in the test are generally easy items, while the math questions are composed of medium difficulty and difficult questions. . We can say that there are few difficult items in the tests in the exam, and that the items in the exam are of medium difficulty. Setting the standards for the education level of international students, preparing questions according to these standards and establishing the appropriateness of the items require a more challenging process compared to national exams. When we examine the items applied in the exams, it is one of the statistical inferences that all of the questions with mathematical skills in the test consist of statistically good discriminating items, and on the other hand, logic questions in general have a low level of discrimination. When the variance and standard deviation values of the items were examined, it was determined that the values of very difficult and very easy items were lower than the other areas. When we look at the item skewness and kurtosis values, it is seen that the distributions are generally symmetrical. However, in many items, it is seen that there is aggregation in order to easily answer the skewness and kurtosis values. For this reason, it was observed that the success increased due to the mentioned items. When the item reliability coefficients were examined, it was observed that the questions measuring mathematics skills were more reliable, while the reliability values of the logic questions were lower. When we examine the descriptive statistics of the foreign student exam, in which international students participate, it is seen that individuals score between 3-80 in the exam. It is seen that the values obtained as a result of the conversion of the raw scores, the results of which we look at, into Z scores, vary between -2.63 and 1.68. It can be said that there is a noticeable increase in the number of people with a low score in the exam and as the score increases. However, it can be said statistically that the highest accumulation is between 63-75 points.

It can be said that the number of students with medium and high scores is quite high, so the distribution shows a left skewed, that is, a negatively skewed distribution. It can be said that there is a partially flat distribution in the studies on the kurtosis of the scores. When we examine the other test statistics of the results, it is seen that the arithmetic mean is 50, the median value is 51, the standard deviation is 17.88 and the variance is 319.60. Accordingly, it can be said that the group taking the exam is successful and the variability in scores is high. It can be said that the kurtosis skewness coefficient of the test is in the expected range of -1,+1, so there is evidence that the scores show a normal distribution. It can be said that the average difficulty value of the test is 0.625, the minimum item difficulty in the test is 0.23, the maximum item difficulty is 0.94 and it is a relatively easy exam. It can be said that the average discrimination value of the test is 0.555 and the average biserial correlation coefficient is 0.49, so it is a powerful test in distinguishing successful and unsuccessful students. When the calculated reliability values of the test are examined, it can be said that the KR20 value is 0.96, the KR21 value is 0.95, the inter-half reliability calculated by dividing the test into two halves is 0.77, and the reliability between odd and even items is 0.93, so it can be said that the test is a reliable test in terms of internal consistency.

In this study, in which the comparative analysis of the students was examined, it was seen that the average scores of the students who took the English and Turkish exams were lower than the averages of the other language exams in Russian, Arabic, German and French. According to this statistical calculation, it was seen that the individuals who took the exam were successful in English, which is accepted as the universal language, and in Turkish, the country where they took the exam. It can be seen that the reason why students get high scores from English is the universality of this language and the reason why they get higher scores from Turkish is their readiness for the language of the country.