



## Açık ve uzaktan öğrenmede şans başarısı: Klasik Test Kuramı (KTK) ve Madde Tepki Kurama (MTK) temelinde karşılaştırmalı bir analiz

Gülğün BULUT<sup>a</sup>

<sup>a</sup> Anadolu Üniversitesi Sosyal Bilimler Enstitüsü

### Özet

Bu çalışmada açık öğretim sınavlarında 4 yanlış 1 doğru uygulamasının şans başarısını etkileme düzeyi; Klasik Test Kuramı (KTK) ve Madde Tepki Kuramı (MTK) çerçevesinde karşılaştırılarak tespit edilmeye çalışılmıştır. Bu sebeple aradaki ilişkiyi görebilmek adına; 2013-2017 yılları arasında toplam 8 dönemi kapsayan ve öğrenci sayısı dönemler itibarıyla 1800-4000 aralığında değişen bir dersin datası şu çerçevede analiz edilmiştir: “Sınavlar aracılığıyla kestirilmeye çalışılan yetenek ölçütleri hangi madde tepki kuramı modeline daha uyumludur?”, “KTK’na göre kestirilen yetenek ile MTK’ya göre kestirilen yetenek ölçüleri arasında anlamlı bir ilişki var mıdır?”, “4 yanlışın 1 doğruyu iptal etmesi uygulamasıyla hesaplanan puanlarla, MTK 2PL ve MTK 3PL modellerine göre hesaplanan puanlar arasında anlamlı bir ilişki var mıdır?” Ayrıca Açık ve Uzaktan öğrenmede şans başarısının öğrenciler tarafından kullanılma oranları ne kadardır sorusuna da yanıt aranmıştır. Elde edilen bulgularda KTK’ya göre hesaplanan puanlarla 4 yanlış 1 doğru yöntemiyle hesaplanan puanlar arasındaki korelasyon ara sınavda 0.909 finalde 0.883 çıkarken, Madde tepki kuramının şans başarısı düzeltilmesi yapan 3 parametrelili modeliyle ise korelasyon ara sınavda 0.688 dönem sonu sınavında 0.641 çıkmıştır. Sonuç olarak Açıköğretim sınavlarında öğrencilerin soruları sistematik bir şansa dayalı olarak cevaplamadığı ve MTK 3 parametrelili modelin şans başarısını daha iyi düzelttiği görülmüştür.

**Anahtar Sözcükler:** Klasik Test Kuramı, Madde Tepki Kuramı, Şans Başarısı

### Abstract

In this study, we have tried to identify the extent to which 4 wrong 1 correct technic affects the guessing in open education exams with a view based on Classical Test Theory and Item Response Theory. In order to see the relationship between these two, students registered between 2013 and 2017 including 8 semesters and counted approximately as 1800-4000 students were analyzed in the present frame: Which IRT model is best suitable for identifying talent criteria by examinations? Is there a significant relationship between the talent scores identified by CTT and IRT? Is there a significant relationship between scores calculated by 4 wrong 1 correct technic, the IRT 2PL and IRT 3PL model? We have also tried to see the usage level of guessing among Open and Distance Education students. Findings indicate that the correlation between the scores calculated by CTT and 4 wrong 1 correct technic is 0,909 in midterm and 0,883 in final exams while . the three parameters model of IRT indicates a correlation of 0,688 in midterm and 0,641 in final exams. In short, it is found in this study that Open Education students do not answer the questions in a guessing basis and three parameters model of IRT work better in refining guessing.

**Keywords:** Classical Test Theory, Item Response Theory, Guessing

### Kaynak Gösterme

Bulut, G. (2018). Açık ve uzaktan öğrenmede şans başarısı: Klasik Test Kuramı (KTK) ve Madde Tepki Kurama (MTK) temelinde karşılaştırmalı bir analiz. *AUAd*, 4(1), 78-93.

## Giriş

Eğitim sisteminin önde gelen hedeflerinden birisi de öğrenci başarısının gerçeğe yakın ve adaletli bir şekilde nasıl kestirilebileceği olgusudur. Çünkü bireyin yeteneği gizlidir ve net olarak ölçülmesi pek mümkün değildir. Ayrıca bireyin yeteneğinin gizil olmasının yanı sıra, ölçmeyi yapan kişiden, ölçülen özellikten, kullanılan ölçme aracından vb. kaynaklanan çeşitli hatalar yüzünden öğrenci başarısının kestiriminde gerçek değerine ulaşmak çoğu zaman mümkün olmamaktadır (Çelen ve Aybek, 2013). Söz konusu ölçüm güçlükleri sebebiyle öğrencinin testte yer alan maddelere verdiği yanıtlara bakılarak gerçek başarısının kestirilmeye çalışılması bir miktar hata payını göze almayı gerektirmektedir (Erkuş, 2003). Geçmişten günümüze öğrenci başarısı en geleneksel sistem olan sınavla ölçülmeye çalışılmaktadır. 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?”, bir diğer ifadeyle “Soruların alt yapısı öğrenci başarısını tam ve doğru ölçmede yeterli mi?” gibi sorular öncelikle yanıtını bulması gereken konulardandır. Bu noktada da soruların niteliği yani öğrenci başarısını ölçme kabiliyeti devreye girmektedir. Bu sebepten ötürü 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. 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. Ayrıca 2016-2017 güz dönemi ara ve final sınavları için 4 yanlış 1 doğru uygulamasının da sürece dahil edildiği bir model ile puanlar tekrar hesaplanarak; sürecin öncesi ve sonrası arasındaki farka bakılıp, MTK 2PL ve MTK 3PL modelleri ile ilişkisi analiz edilecektir.

Bu çalışmanın amacı, açık ve uzaktan öğrenme yöntemiyle eğitim vermekte olan, 2013-2017 yılları arasında toplam 8 dönemi kapsayan ve öğrenci sayısı dönemler itibariyle 1800-

4000 aralığında deęişen bir dersin ara ve final sınavlarındaki şans başarısını test etmektir. Bu çalışma, bir testin KTK ve MTK'ya göre puanlanmasına baęlı olarak öğrencilerin başarı sıralamalarının nasıl deęiştirdiğini ve bir değerlendirme söz konusu olduğunda sonucu nasıl etkilediğini tespit etmeyi amaçlamaktadır.

### Araştırma Sorunsalı

Klasik Test Kuramı (KTK) ve Madde Tepki Kuramı (MTK) dięer adıyla Item Response Theory (IRT), yeteneęi gizil olan bireyin gerçek puanını (gerçek yeteneęini) ölçmede kullanılan alanyazındaki iki temel kuramdır. Bu sebeple araştırmada açık ve uzaktan öğrenme yöntemiyle eğitim vermekte olan bir dersin 2013-2017 yılları arasında güz döneminde yapılan ara ve final sınav puanları KTK, MTK 2PL ve MTK 3PL modeline göre ayrı ayrı hesaplanarak analize tabi tutulacaktır. Ayrıca 2016-2017 güz dönemi ara ve final sınavları için 4 yanlış 1 doğru uygulamasının sürece dahil olmasının öğrenci başarısına etkisi test edilmeye çalışılacaktır. Yapılan analizler sonucu hesaplanan puanların madde tepki kuramı modellerinden (2PL, 3PL) hangisine daha çok uyum gösterdiği ve söz konusu puanların Klasik Test Kuramına (KTK) ve Madde Test Kuramına (MTK) göre kestirilmesinin anlamlı bir farklılık gösterip göstermedięi ortaya çıkarılmaya çalışılacaktır. Bu kapsamda öğrenci puanlarını hesaplamada ilk olarak 4 yanlış 1 doğru uygulamasının öncesi ve sonrası şeklinde bir kıyaslamaya gidilecek, sonrasında ise KTK, MTK 2PL ve MTK 3PL ile hesaplanan puanlar çapraz bir değerlendirme yöntemi ile karşılaştırılacaktır.

Yukarıda bahsi geçen amaçlar doğrultusunda çalışma kapsamında cevap aranan sorular sıralanacak olursa;

1. Sınavlar aracılığıyla kestirilmeye çalışılan yetenek ölçütleri hangi madde tepki kuramı modeline daha uyumludur?
2. KTK'na göre kestirilen yetenek ile MTK'ya göre kestirilen yetenek ölçüleri arasında anlamlı bir ilişki var mıdır?
3. 4 yanlış 1 doğru uygulamasıyla hesaplanan puanla, MTK 2PL ve MTK 3PL modeline göre hesaplanan puan arasında anlamlı bir ilişki var mıdır?
4. Açık ve Uzaktan öğrenmede şans başarısının öğrenciler tarafından kullanılma oranları ne kadardır?

## İlgili Alanyazın

Alan yazında öğrenci başarısını tam ve doğru olarak ölçmede “Klasik Test Kuramı” ve “Madde Tepki Kuramı” olmak üzere iki tür kuramdan bahsedilmektedir.

### Klasik Test Kuramı

1905’li yıllarda Spearman tarafından temeli atılan “Klasik Test Kuramı” en yalın ifadeyle bireylerin gerçek puanlarının  $X=T+E$  eşitliğiyle bulunabileceğini ifade etmektedir (Turgut ve Baykul, 2010). Bu eşitlikte X gözlenen puanı, T gerçek puanı ve E ise tesadüfi hata miktarını göstermektedir. Bu kapsamda klasik test kuramının tesadüfi hataya odaklandığından (Akyıldız ve Şahin, 2017, s. 144) bahsetmek mümkündür. Çünkü hata ne kadar iyi tanımlanır ve tespit edilirse, ölçüm sonucu da bir o kadar gerçeği yansıtabilecek ve güvenilir olabilecektir. Klasik Test Kuramına göre;

- Testin uygulandığı gruptan elde edilen sonuçlar bir bütün olarak değerlendirilmektedir.
- Maddelerin güçlük ve ayırt edicilik indeksleri ancak grubun tüm verisi elde edildikten sonra kestirilebilmektedir.
- Grubun değişmesi durumunda elde edilen verilerinde değişmesi olası bir durumdur. Çünkü KTK’da madde istatistikleri gruba bağımlıdır. Uygulanılmış maddelerin parametrelerinden yararlanarak sonraki uygulamaların değerlerini tahmin etmek mümkün görünmemektedir.
- Bir ölçme ya da değerlendirme yapılırken her bir madde bazında değil bir bütün olarak toplam puana bakılmaktadır.

KTK’da iki tür madde istatistiği/madde parametresi bulunmaktadır (Akyıldız, 2015, s. 13). Bunlar:

- *Madde Güçlüğü* ( $p_j$ ), bir maddenin uygulandığı grupta doğru cevaplanma yüzdesidir.
- *Madde Ayırt Ediciliği* ( $r_{jx}$ ), maddenin ölçmeye çalışılan beceriye sahip olanlarla olmayanları ayırt etme gücüdür.

Yukarıda bahsi geçen madde istatistikleri testin uygulandığı gruba göre farklılık gösterebilmektedir. Örneğin; Bir gruba uygulanan testin sonucunda elde edilen “*Madde Güçlüğü*” ve “*Madde Ayırt Ediciliği*” istatistik değerleri ile aynı testin başka bir gruba uygulanması sonucu elde edilen değerler farklılık gösterebilmektedir. Bu sebeple KTK’da madde istatistikleri gruba bağımlıdır ve grubun değişmesi durumunda güvenilirliğin de

değişmesi söz konusu olmaktadır (Nartgün, 2002). KTK'ya göre başarı puanı öğrencinin maddelerden almış olduğu toplam puan üzerinden hesaplanmaktadır.

### **Madde Tepki Kuramı=MTK (IRT= Item Response Theory)**

Örtük Özellikler Kuramı olarak da bilinen Madde Tepki Kuramının (Crocker ve Algina, 1986) kökeni her ne kadar 1920'li yıllara uzansa da (Bock, 1997), ilk olarak 1930'lu yıllarda ortaya çıkmış ve kuramla ilgili asıl çalışmalar ise 1950'li yıllarda Lord tarafından uygulanmaya başlanmıştır. Madde Tepki Kuramı (IRT), özellikle 1970'lerden sonra yapılan test geliştirme çalışmalarında, büyük oranda KTK'nın yöntemlerine alternatif olarak uygulama alanı bulmuştur (Akyıldız ve Şahin, 2017).

Madde Tepki Kuramı (IRT); Klasik Test Kuramının eksiklik ve sınırlılıklarını giderebileceği iddia edilen, bireylerin yeteneklerinin matematiksel yöntemlerle kestirilmesinde kullanılan bir ölçme aracıdır. Kurama göre, bireylerin belli bir alandaki doğrudan gözlenemeyen yetenekleri ile söz konusu alanı yoklayan sorulardan oluşan test maddelerine verdikleri yanıtlar arasında bir ilişki vardır ve bu ilişki matematiksel olarak ifade edilebilmektedir. Bu kapsamda Madde Tepki Kuramı (IRT)'nin temel varsayımı; testi alan bireyin o teste ilişkin performansının, bireyin özellikleri (traits) olarak adlandırılan yeteneklerinin tanımlanması ve bu özelliklerle ilişkili olarak bireylerin yetenek puanlarının tahmin edilerek, madde ve test performanslarının açıklanıp kestirilebileceği esasına dayanmaktadır. Diğer bir deyişle, Madde Tepki Kuramı bir ya da birden fazla özellik sayesinde bir kişinin bir testte gösterebileceği performansın kestirilebileceğini varsayar. Çünkü Madde Tepki Kuramı (MTK), soru ve birey bazlı puanlama yöntemi kullanmaktadır. Bir sorunun her bir birey için puanı farklıdır. MTK yöntemiyle hesaplanan puanlara her bir maddenin katkısı her bireyde aynı değildir; kullanılan modele göre maddenin güçlüğünden, ayırt edicilik gücünden ve şansla doğru yanıtlanma olasılığından etkilenen madde puanları söz konusudur.

Madde Tepki Kuramına göre;

1. Bir madde, o maddeyi cevaplandırabilecek herhangi bir cevaplayıcı grubundan bağımsızdır.
2. Bir bireyin yeteneği, o bireye uygulanan herhangi bir madde grubundan bağımsızdır.
3. Bir testin özelliklerini testi uygulamadan önce kestirmek mümkündür.

Yukarda belirtilen ilk iki özellik, Klasik Test Teorisinin sınırlılıkları olarak belirtilen özelliklerin tam tersidir. Söz konusu özelliklerin KTK'nın sınırlılıklarını ortadan kaldırdığı varsayılmaktadır. Çünkü IRT gruba ve bireye uygulanan maddelere bağımlı değildir ve bu durumda Madde Tepki Kuramını daha kullanışlı hale getirmektedir.

Madde Tepki Kuramının diğer bir önemli özelliği de testi uygulamadan önce yeteneğin kestirilebileceğini iddia etmesidir. Madde Tepki Kuramına göre; gruptan, uygulanan maddeden bağımsız ve hatta testi uygulamadan dahi matematiksel yöntemlerle kestirilebilecek mutlak değişmez bir yetenek düzeyi vardır. Bu yetenek düzeyi her şeyden bağımsızdır ve farklı testlerde farklı gruplarda bir birey için aynı sonuçları vermektedir. Madde Tepki Kuramı gerek testlerin farklı yetenek gruplarına uyarlanabilmesinin kolaylığı nedeniyle gerekse bireylerin yetenek ölçülerini teste bağımlı kalmadan, maddeler aracılığıyla kestirebildiği için her geçen gün daha popüler ve tercih edilebilir olmaktadır (Reise, Ainsworth ve Haviland, 2005).

### **Madde Tepki Kuramının modelleri**

*Tek Parametrelili Lojistik Model (1 PL) (Rasch Model):* Rasch tarafından ilk kez kullanılan ve çoğu kez Rasch model olarak da anılan bu model, madde güçlüğünü esas almaktadır. Modelde madde güçlüğüne bağlı olarak belirli bir yetenek düzeyindeki bir kişinin o maddeye doğru yanıt verme olasılığının fonksiyonu belirlenir (Akyıldız, 2015).

*İki Parametrelili Lojistik Model (2 PL):* Bu model Lord tarafından geliştirilmiş, Birnbaum tarafından iki parametrelili lojistik modele ait madde karakteristik eğrisi fonksiyonu tanımlanmıştır. Madde güçlük parametresine (b) madde ayırtediciliğinin (a) eklendiği bu modelde madde ayırtediciliği, maddenin ölçmeye çalıştığı yeteneğe sahip olan bireyleri o yeteneğe düşük düzeyde sahip olanlardan maddeye doğru yanıt verme olasılığı üzerinden ayırt etme derecesini belirten bir parametredir (Akyıldız, 2015).

*Üç Parametrelili Lojistik Model (3 PL):* Bu modelde madde ayırt ediciliği ve madde güçlüğüne ek olarak doğru yanıtlanma olasılığı bir de şans başarısı göz önünde bulundurularak kestirilir. Şans başarısı daha çok alt yetenek gruplarındaki kişilerin test maddelerine şansla cevap verme olasılığını veren bir parametredir (Baykul, 1979; Crocker & Algina, 1986; Embretson & Hershberger, 1999; Akyıldız, 2015).

Her iki kuramda da test maddelerine verilen tepkilerin doğruluğu veya yanlışlığı üzerine odaklanılması KTK ve MTK'nın ortak yönüdür. Ayrıca normallik varsayımı her iki kuram için de söz konusudur. Çünkü normal dağılımın sağlanamaması yetenek kestiriminde sorunlara yol açabilmektedir.

Alanyazında KTK ve MTK'yı karşılatıran alıřmalara genel olarak bakıldıđında, testlerin yetenek kestiriminde gvenilirlik ve geerliliklerinin (Gzen ıtak, 2007; etin, 2009; řahin ve Akyıldız, 2017), psikometrik zelliklerinin (elen, 2008), đrencilerin bařarı ltlerinin (zer zkan, 2012) ve aralarındaki uyumun (Selvi, 2013; Sayın, 2014) karřılařtırıldıđı grlmektedir.

### **řans Bařarısı**

řans bařarısı, oktan semeli soru yapısının kullanıldıđı bir testte dođru yanıtların tahmine giderek dođru bulunma olasılıđıdır. Madde ile llen yeteneđe (bilgi dzeyine) sahip olmayan ya da kısmen sahip olan yanıtlayıcıların, dođru yanıtı tahmine giderek bulmaları durumuna řans bařarısı karıřmaktadır (Baykul, 2000). Bu ařamada da rastgele tahminin testin geerliliđine ve gvenilirliđine olumsuz katkısı devreye girmektedir. Bu alandaki arařtırmaların birođu gstermektedir ki, testlerdeki řans bařarisının geerlilik (Lord, 1962) ve gvenilirlik zerindeki olumsuz etkisini azaltmanın yolu dzeltmiř test puanlarıdır. Gnmzde en sık kullanılan test puanını dzeltme yntemi 4 yanlıřın 1 dođruyu gtrdđ sistemdir. Fakat bu yntemin test puanını řans hatasından arındırdıđı ve yanıtlayıcılar arasındaki řans bařarisından kaynaklanan farkları giderdiđini savunmak olanaksızdır (Umay, 1998; elen ve ıkırıkcı Demirtařlı, 2006). rneđin; birok đrenci dođru cevabı seme ihtimalini 5 seenekten belirgin yanlıřları eleme yntemiyle 2 seeneđe kadar indirmekte ve řansını 2 seenek zerinde denemektedir. Yani đrenci soru hakkında tam olarak bilgisiz deđildir. Bu durumdaki bir đrenciye tam puan vermenin dođru olmayacađı gibi hi puan vermemek de dođru bir yaklařım olmayacaktır. nk bazı tahminler kısmi bilgiye dayanmaktadır. Dolayısıyla dzeltme formlleri her zaman adil bir puanlama yntemi olmamaktadır.

## **Yntem**

### **Arařtırma Modeli**

Arařtırma nicel arařtırma yntemlerinden betimsel model kullanacak řekilde tasarlanmıřtır. Nicel arařtırmalar “bir fenomeni, istatistik gibi matematiksel yntemler kullanarak analiz edilmiř miktarsal (sayısal) veri toplayarak aıklamak” (Aliaga & Gunderson; 2002) olarak tanımlanmaktadır. Betimsel alıřmalar varolan durumu olduđu haliyle ortaya koyan alıřmalardır.

## Araştırma Grubu

Bu çalışmada kullanılan veriler açık ve uzaktan öğrenme yöntemiyle eğitim veren ve öğrenci sayısı dönemler itibariyle 1800-4000 aralığında değişen bir dersin 2013-2017 yılları arasında güz döneminde yapılan ara ve final sınav puanlarından oluşmaktadır.

## Veri Toplama Araçları

Veriler sınav sonuçlarının tutulduğu veri tabanından İstatistik Bilgi Edinme ve Değerlendirme biriminden alınan izinle çekilmiştir.

## Veri Toplama Süreci

Veri tabanından ham haliyle elde edilen datalar işlenerek analize uygun formata getirildikten sonra “Jmetrik 4.1.0” programında 1-0 matris olacak şekilde yeniden kodlanmıştır. Ayrıca KTK’da 4 yanlış 1 doğru uygulamasının etkisini görebilmek için yanlışların doğruları götürebildiği bir kodlama yoluyla datalar yeniden işlenmiştir. Analiz için elde edilen datalarda sadece “A” kitapçığı kullanılmıştır.

## Veri Analizi

Analiz için 2013-2017 yılları arasındaki güz eğitim öğretim dönemini kapsayan ara ve final sınavlarının yer aldığı toplam 8 ayrı veri seti kullanılmıştır. Veriler Jmetrik 4.1.0.’da KTK, IRT 2PL ve IRT 3PL modele göre analiz edilmiş ve sınav puanları her bir veri seti için ayrı ayrı hesaplanmıştır. Ayrıca 2016-2017 ara ve final sınavları için 4 yanlış 1 doğru uygulaması puanlamaya dahil edildiği bir kodlama yapılarak puanlar yeniden hesaplanmıştır.

Veriler SPSS 25 programına aktarılarak *Kendall’in tau-b* ( $\tau_b$ ) üzerinden korelasyonlarına bakılmıştır. *Kendall’in tau-b* ( $\tau_b$ )’nin temel alınmasının sebebi ise, bağıntı analizlerinde incelenen değişkenin tipine ve dağılım özelliğine (sıralı, sürekli, normal dağılıma uyuyor, uymuyor gibi) göre hangi testin uygulanması gerektiği durumu değişebilmektedir. Bu kapsamda yapılan analizlerde yaygın olarak kullanılan testler; *Pearson*, *Spearman* ve *Kendall’in tau-b* olmak üzere üç tür testten oluşmaktadır. Bunlardan ilki parametrik bir yöntem olan ve sürekli değişkenlerin birbirleriyle doğrusal ilişkisinin derecesi her iki değişkende de normal dağılım gösteren *Pearson* bağıntı katsayısıdır. İkincisi normal dağılıma uymayan sürekli değişkenler arası doğrusal ilişki incelemesinde kullanılan *Spearman* bağıntı katsayısıdır. Üçüncüsü ise incelenen değişkenlerin biri veya her ikisinin de sıralı değişken olması durumunda kullanılan *Kendall’in tau-b* katsayısıdır. Bu çalışmada da öğrenciler



puanları bakımından sıralandığından ve geçme durumlarını sıraları belirlediğinden (Kılıç, 2012) *Kendall'in tau-b* katsayısının kullanılmasına karar verilmiştir. Çünkü *Kendall'in tau-b* ( $\tau_b$ ) korelasyon katsayısı, en az bir sıra ölçeğinde ölçülen iki değişken arasında var olan ilişkinin kuvveti ve yönünün parametrik olmayan bir ölçümüdür. Bu sebeple *Spearman*'ın korelasyon katsayısına alternatif bir yöntem olarak düşünülmektedir (Leard Statistic, 2017).

Yukarıdaki açıklamalardan da anlaşılacağı üzere gerek diğer testlere üstünlüğü gerekse çalışmamızın yöntemine uygunluğu sebebiyle sonuçlar *Kendall'in tau-b* ( $\tau_b$ ) korelasyon katsayısı üzerinden yorumlanmıştır.

### Bulgular ve Yorumlar

Veri analizi başlığında da belirtildiği üzere *Kendall'in tau-b* katsayısı, incelenen değişkenlerin biri veya her ikisi de sıralı değişken olduğunda tercih edilen bir testtir. Çalışmamızda da büyük oranda sıralı değişkenlerin ön planda olması dolayısıyla bulguları yorumlamada *Kendall'in tau-b* katsayısı kullanılmıştır.

Çalışmanın birinci sorusu olan sınavlar aracılığıyla kestirilmeye çalışılan yetenek ölçütleri hangi madde tepki kuramı modeline daha uyumludur sorusunun cevabı için puanlar IRTPRO programında MTK 2 parametrelili ve MTK 3 parametrelili modellerine göre analiz edilmiş ve her bir model için uyum değerlerine bakılmıştır. Bu kapsamda elde edilen uyum değerleri Tablo 1'de gösterilmiştir.

Tablo1						
<i>MTK 2 PL ve MTK 3 PL Modele Göre Uyum Değerleri</i>						
	MTK 2PL			MTK 3PL		
Yıl	Log Likelihood	M2	RMSEA	Log Likelihood	M2	RMSEA
2013-2014 ara	96745.90	427,64	0,02	96752.08	376,99	0,02
2013-2014 ds	95870.30	1474,09	0,04	95788.60	821,30	0,03
2014-2015 ara	75829.26	637,79	0,03	75757.79	456,85	0,03
2014-2015 ds	73395.20	611,73	0,03	73445.68	446,03	0,03
2015-2016 ara	63040.64	568,69	0,03	62839.13	343,99	0,02
2015-2016 ds	62782.65	839,29	0,04	62943.81	654,78	0,04
2016-2017 ara	46448.77	323,34	0,02	46488.12	212,25	0,01
2016-2017 ds	43437.22	526,36	0,03	43464.24	365,72	0,03

Tablodan da görüleceği üzere uyum istatistikleri “Log Likelihood”, “M2” ve “RMSEA” değerleri üzerinden verilmiştir. Uyum istatistiklerinde en iyi model-veri uyumu en küçük değerle ifade edilmektedir.

Jmetrik 4.1.0. ve SPSS 25 programlarında yapılan analizler sonucu klasik test kuramının 4 yanlış 1 doğru (4y1d), madde tepki kuramı 2 parametrelili ve 3 parametrelili modellerine göre korelasyonları dönemler itibariyle Tablo 2’de gösterilmiştir.

Tablo 2					
<i>KTK'nın MTK 2PL, MTK 3PL ve 4y1d ile Korelasyonu</i>					
	Sınav Dönemleri	KTK-MTK2PL Korelasyon Katsayısı	KTK-MTK3PL Korelasyon Katsayısı	KTK-4y1d Korelasyon Katsayısı	KTK(4y1d)-MTK3PL Korelasyon Katsayısı
Kendall's tau_b	2013-2014_ara	.815**	.791**	-	-
	2013-2014_ds	.805**	.718**	-	-
	2014-2015_ara	.926**	.902**	-	-
	2014-2015_ds	.774**	.754**	-	-
	2015-2016_ara	.858**	.880**	-	-
	2015-2016_ds	.823**	.806**	-	-
	2016-2017_ara	.712**	.688**	.909**	.725**
	2016-2017_ds	.682**	.641**	.883**	.718**

\*\* 0.01 düzeyinde anlamlı.

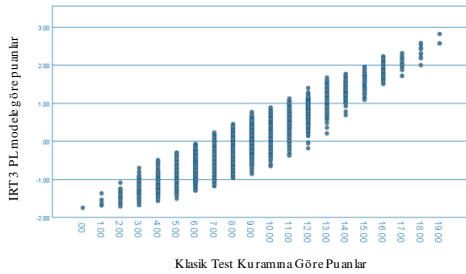
Jmetrik analizi sonucu 2013-2014 güz dönemi ara sınavından 2016-2017 güz dönemi final sınavına kadar klasik kurama göre hesaplanan toplam puanlar ile MTK 2PL modele göre hesaplanan toplam puanlar arasındaki korelasyonlar en yüksek 0.858-0.926 arasındayken en düşük 0.682-0.712 arasındadır. Aralık geniş değildir, genel olarak puanların yüksek sıra benzerliği gösterdiği görülmüştür.

Klasik kurama göre hesaplanan toplam puanlar ile MTK 3PL modele göre hesaplanan toplam puanlar arasındaki korelasyonlar ise en yüksek 0.880-0.902 arasındayken en düşük 0.641-0.688 arasındadır. Her ne kadar aralık geniş olmayıp, puanlar yüksek sıra benzerliği gösterse de sonuçlar MTK 3PL’ye daha çok uyum göstermektedir. MTK 3 PL’nin şans başarısı üzerinde diğer yöntemlere oranla daha çok düzeltme yaptığının bilindiği durumda KTK ve MTK 3PL öğrenci başarı sıralamaları farklılık göstermektedir. Ayrıca KTK’nın 4y1d uygulaması sonrasında korelasyonlar düşmekte ve öğrencileri doğru olmayan bir şekilde sıralamaya yol açmaktadır.

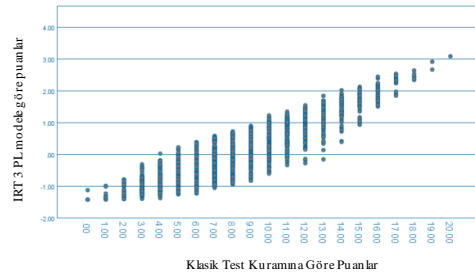
KTK ve 4y1d arasındaki korelasyon da ara sınav için 0.909, final sınavı için 0.883 çıkmıştır. Bu durum açık öğretim öğrencilerinin soruları şansa bağlı olarak cevaplamadığını bu sebeple 4y1d şans başarısı düzeltmesi geldiğinde de korelasyonda dikkate değer bir değişiklik meydana gelmediğini göstermektedir. Açık öğretim sınavlarında başarı sırası

önemlidir ancak 4yıld şans düzeltmesi sıraların aynı kalma derecesine herhangi bir etkiye bulunmamaktadır.

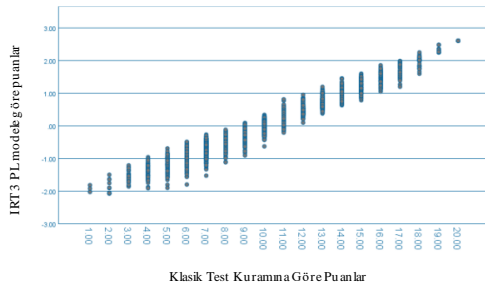
Yapılan analizlerde bu koşulu MTK 3PL modeli sağladığı için tüm dataların korelasyonlarında MTK 3PL modeli kullanılmıştır.



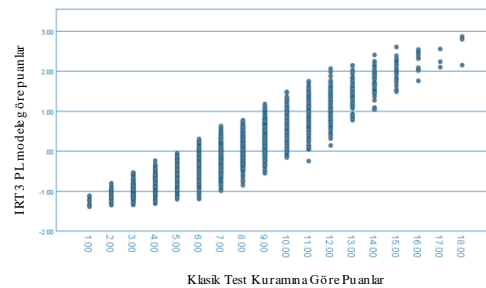
*Grafik 1.* 2013-2014 ara sınav puanlarının KTK ve IRT 3 PL'ye göre korelasyonu



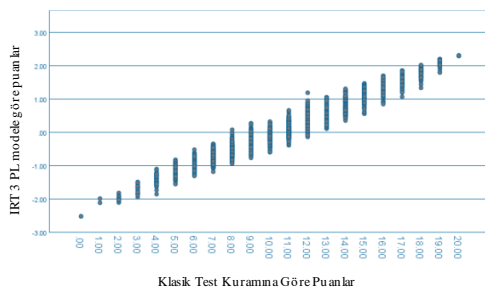
*Grafik 2.* 2013-2014 final sınav puanlarının KTK ve IRT 3 PL'ye göre korelasyonu



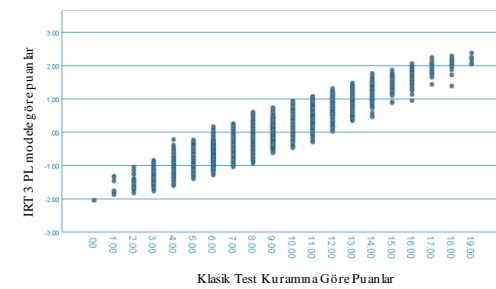
*Grafik 3.* 2014-2015 ara sınav puanlarının KTK ve IRT 3 PL'ye göre korelasyonu



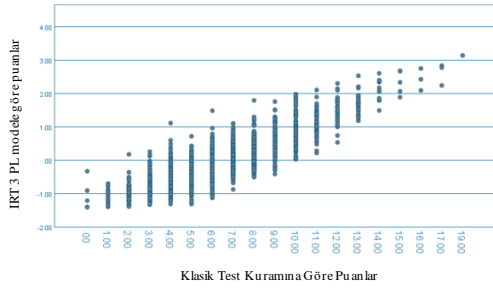
*Grafik 4.* 2014-2015 final sınav puanlarının KTK ve IRT 3 PL'ye göre korelasyonu



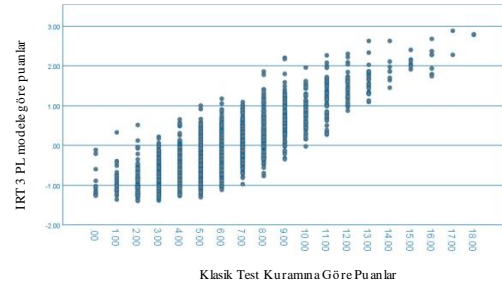
*Grafik 5.* 2015-2016 ara sınav puanlarının KTK ve IRT 3 PL'ye göre korelasyonu



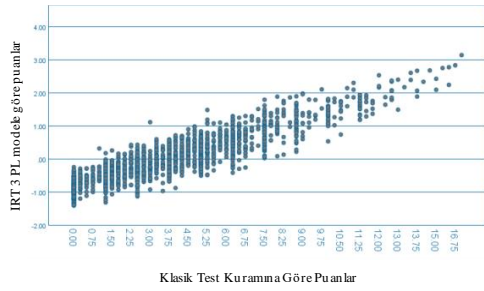
*Grafik 6.* 2015-2016 final sınav puanlarının KTK ve IRT 3 PL'ye göre korelasyonu



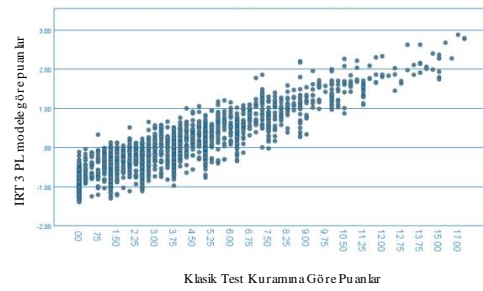
Grafik 7. 2016-2017 ara sınav puanlarının KTK ve IRT 3 PL'ye göre korelasyonu



Grafik 8. 2016-2017 final sınav puanlarının KTK ve IRT 3 PL'ye göre korelasyonu



Grafik 9. 2016-2017 ara sınav puanlarının KTK 4y1d ve IRT 3 PL'ye göre korelasyonu



Grafik 10. 2016-2017 final sınav puanlarının KTK 4y1d ve IRT 3 PL'ye göre korelasyonu

Araştırmanın temel sorunsallarından olan “KTK’na göre kestirilen yetenek ile MTK’ya göre kestirilen yetenek ölçüleri arasında anlamlı bir ilişki var mıdır?” sorusundan yola çıkarak puanlar her bir dönem için Jmetrik’te KTK ve MTK’ya göre ayrı ayrı analiz edilmiş ve karşılıklı korelasyonlarına bakılmıştır.

“4 yanlış 1 doğru uygulamasıyla hesaplanan puanla, MTK 2PL ve MTK 3PL modeline göre hesaplanan puan arasında anlamlı bir ilişki var mıdır?” sorusuna cevaben; 2016-2017 ara ve final sınav dönemi için 4 yanlış 1 doğru uygulaması sürece dahil edilerek hem öncesi ve sonrası arasındaki farka hem de diğer modellerle ilişkisine bakılmıştır. Elde edilen bulgular sonucu değer aralıkları yüksek olmamakla beraber MTK 3PL modeli şans başarısını diğer modellere oranla daha çok düzeltmektedir.

Araştırmamızın son sorusu “Açık ve Uzaktan öğrenmede şans başarısının öğrenciler tarafından kullanılma oranları ne kadardır?” a cevaben, Jmetrik 4.1.0.’da madde tepki kuramının şans başarısını da dikkate alan MTK 3PL modeline göre analiz yapılmıştır. Klasik kurama göre hesaplanan toplam puanlar ile şans düzeltmesini “c” parametresine göre yapan MTK 3PL modele göre hesaplanan toplam puanlar arasındaki ilişki en yüksek 0.902 en düşük 0.641 çıkmıştır.

## Sonuçlar

Yapılan analizler sonucu elde edilen bulgular bir bütün olarak incelendiğinde KTK ile hesaplanan puanlar ile 4y1d ile hesaplanan puanlar arasındaki korelasyonların 0.883-0.909 değerleri arasında değiştiği ve dikkate değer bir fark olmadığı görülmüştür. Bu kapsamda 4y1d uygulamasının açık öğretim sınavlarında şans düzeltmesini gerçekleştirmediği sonucuna ulaşılmıştır. Çünkü açık öğretim öğrencileri soruları şansa dayalı cevaplamamakta dolayısıyla sistematik bir şans başarısı kullanmamaktadır.

Madde tepki kuramı 2 parametrelili ve 3 parametrelili modele göre kestirilen yetenek ölçülerinde ise aralık geniş olmayıp, puanlar yüksek sıra benzerliği gösterse de sonuçlar MTK 3PL'ye daha çok uyum göstermektedir. Diğer bir ifadeyle MTK 3 PL'nin şans başarısını diğer yöntemlere oranla daha çok düzeltmektedir.

## Öneriler

Gelecekte yapılacak olan çalışmalarda Klasik Test Kuramına dayalı şans başarısından arındırma yöntemleri kullanılarak yetenek kestirimleri arasındaki ilişkiler daha detaylı olarak incelenmelidir. Madde Tepki kuramına dayalı 3 parametrelili kestirimlerin harf notlarının dağılımını ne kadar etkilediği belirlenerek şans başarısı düzeltmesinin harf notları üzerindeki etkisi ortaya çıkarılabilir.

### Kaynakça

- Akyıldız, M. & Şahin, M. D. (2017). Açıköğretimde kullanılan sınavlardan Klasik Test Kuramına ve Madde Tepki Kuramına göre elde edilen yetenek ölçülerinin karşılaştırılması, *Açıköğretim Uygulamaları ve Araştırmaları Dergisi (AUAd)*, 3(4), 141-159.
- Akyıldız, M. (2015). Açıköğretim Fakültesi 2014/2015 Öğretim Yılı Sınavlarından, Madde Tepki Kuramı ve Klasik Test Kuramına Göre Kestirilen Yetenek Ölçülerinin Karşılaştırılması, *Anadolu Üniversitesi Bilimsel Araştırma Projesi*, Proje no: 1503E094.
- Aliaga, M., & Gunderson, B. (2002). *Interactive statistics*. New Jersey: Prentice Hall.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. Ankara: ÖSYM Yayınları
- Bock, R. D., Thissen, D. and Zimowski, M. F. (1997). IRT Estimation of Domain Scores, *Journal of Educational Measurement*, 37(3), 197–211.
- Crocker, L. & Algina, J. (1986). *Introduction Classical and Modern Test Theory*. USA: CBS College Publishing Company.
- Çelen, Ü. & Aybek, E. C. (Kış 2013). Öğrenci Başarısının Öğretmen Yapımı Bir Testle Klasik Test Kuramı ve Madde Tepki Kuramı Yöntemleriyle Elde Edilen Puanlara Göre Karşılaştırılması, *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 4(2), 64-75.
- Çelen, Ü. (2008). Klasik Test Kuramı ve Madde Tepki Kuramı Yöntemleriyle Geliştirilen İki Testin Geçerlilik ve Güvenilirliğinin Karşılaştırılması, *İlköğretim Online*, 7(3), 758-768.
- Çelen, Ü. & Çıkrıkçı Demirtaşlı, N. (2006). Düzeltme yönergesinin testin psikometrik özelliklerine etkisi, *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 30, 82-91.
- Çetin, E. (2009). Dikey ölçekteleme klasik test ve madde tepki kuramına dayalı yöntemlerin karşılaştırılması, *Yayımlanmamış doktora tezi*, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Gözen Çıtak, G. (2007). Klasik Test ve Madde-Tepki Kuramlarına Göre Çoktan Seçmeli Testlerde Farklı Puanlama Yöntemlerinin Karşılaştırılması, *Yayımlanmamış doktora tezi*, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Hambelton, R. K. (1994). Item Response theory: a broad psychometric framework for measurement advances, *Psicothema*, 6(3), 535-556.
- İlhan, M. (2016). Açık Uçlu Sorularla Yapılan Ölçmelerde Klasik Test Kuramı ve Çok Yüzeysel Rasch Modeline Göre Hesaplanan Yetenek Kestirimlerinin Karşılaştırılması, *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi (H. U. Journal of Education)*, 31(2), 346-368.

- Kılıç, S. (2012). Bağlantı analizi Sonuçlarının Yorumlanması. *Journal of Mood Disorder*, 2(4), 191-3.
- Lord, F. M. (1962). Formula scoring and validity, *Research Bulletin (May)* <http://onlinelibrary.wiley.com/store/10.1002/j.23338504.1962.tb00290.x/asset/ets200290.pdf?v=1&t=jcqaltgl&s=e20f6ab10873be2f957ca17e9db7538a71efee59> (Erişim Tarihi: 25.11.2017)
- Nartgün, Z. (2002). Aynı Tutumu Ölçmeye Yönelik Likert Tipi Ölçek ile Metrik Ölçeğin Madde ve Ölçek Özelliklerinin Klasik Test Kuramı ve Örtük Özellikler Kuramına Göre İncelenmesi. Yayınlanmamış doktora tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Özer Özkan, Y. (2014). Öğrenci Başarılarının Belirlenmesi Sınavından Klasik Test Kuramı, Tek ve Çok Boyutlu Madde Tepki Kuramı Modelleri ile Kestirilen Başarı Puanlarının Karşılaştırılması, *International Journal of Human Sciences*, 11(1), 20-44.
- Reise, S. P., Ainsworth, A.T., and Haviland, M.G. (2005). Item Response theory. Fundamentals, applications, and promise in psychological research, *Current Directions in Psychological Science*, 14(2), 95-101.
- Sayın, A. (2014). Klasik test kuramı ve madde tepki kuramına göre kestirilen parametrelerle sınırlandırılan yapısal eşitlik modellerinin uyum indekslerinin karşılaştırılması, Yayınlanmamış doktora tezi, Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Selvi, H. (2013). Klasik test ve madde tepki kuramlarına dayalı değişen madde fonksiyonu belirleme tekniklerinin farklı puanlama durumlarında incelenmesi, Yayınlanmamış doktora tezi, Mersin Üniversitesi Eğitim Bilimleri Enstitüsü, Mersin.
- Leard Statistic (2017). <https://statistics.laerd.com/spss-tutorials/kendalls-tau-b-using-spss-statistics.php> (erişim tarihi: 15.12.2017).
- Turgut, M. F. & Baykul, Y. (2010). *Eğitimde Ölçme ve Değerlendirme*, Ankara: Pegem Akademi.
- Umay, A. (1998). Seçmeli testlerde yanıtlayıcı davranışları ve şans başarısının elimine edilmesi işlemlerine ilişkin bazı öneriler, *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 14, 54-61.

## Yazar Hakkında

### Glgn BULUT



Glgn Bulut, İktisat alanında lisans eęitimini aldıktan sonra, yksek lisansını Maliye alanında tamamlamıřtır. Bulut, halen doktora eęitimine Anadolu niversitesi Uzaktan Eęitim Anabilim Dalı'nda devam etmektedir. 2008 yılından beri Anadolu niversitesi Sosyal Bilimler Dergisinde grev yapmaktadır. Yazarın ilgi alanları arasında nicel arařtırma yntemleri, yapay sinir aęları, istatistiksel analiz, yetiřkin ęrenme ve yařam boyu ęrenme, e-ęrenme malzemelerinin tasarlanması, kitap tasarımı ve bilgisayar programları yer almaktadır.

Tel (İř) : +90 222 3350580

Eposta : gbulut@anadolu.edu.tr