

## Birey Açıklayıcı Madde Tepki Kuramı Analizi: Örtük Regresyon İki Parametrelili Lojistik Modeli

### Person Explanatory Item Response Theory Analysis: Latent Regression Two Parameter Logistic Model

Burcu ATAR\* Derya ÇOBANOĞLU AKTAN\*\*

Hacettepe Üniversitesi

#### Öz

Bu çalışmada, açıklayıcı madde tepki kuramı modeli çerçevesinde (AMTK) örtük regresyon iki parametrelili lojistik (2-PL) modelinin TIMSS 2007 Türkiye fen bilimleri verisinin bir bölümü üzerindeki uygulaması gösterilmektedir. Bunun için ilk olarak, 2-PL MTK ve örtük regresyon 2-PL modelleri ile madde parametreleri hesaplanmıştır. İkinci olarak, örtük regresyon 2-PL modelinde, cinsiyet, fene karşı olumlu tutum, fene verilen önem, fene yönelik özgüven ve ebeveynlerin eğitim düzeyleri bireysel özelliklerinin öğrenci başarıları üzerindeki etkileri analiz edilmiştir. Bu özelliklerden sadece fene yönelik özgüven ile ebeveynlerin eğitim düzeylerinin, öğrenci başarıları arasındaki farklılıkları açıklamada anlamlı istatistiksel etkilerinin olduğu gözlenmiştir. Uyum indekslerine bakıldığında, örtük regresyon 2-PL modelinin veriye 2-PL MTK modelinden daha iyi uyum gösterdiği bulunmuştur.

*Anahtar Sözcükler:* Açıklayıcı madde tepki kuramı modeli, örtük regresyon, TIMSS 2007 Türkiye.

#### Abstract

In this paper, an application of latent regression two-parameter logistic (2-PL) model as an explanatory item response model (EIRM) was illustrated using a part of TIMSS 2007 Science data for Turkey. For this purpose, initially, item parameters were calculated via 2-PL IRT and latent regression 2-PL models. Then, in the latent regression 2-PL model, the effects of gender, positive affect toward science, valuing science, self-confidence in learning science, and education level of parents as person properties on the student achievement were examined. It was seen that among those properties, only self confidence in learning science and education level of parents had a statistically significant effects on explaining the differences in students' achievements. Based on fit indices, it was found that latent regression 2-PL model had a better model fit to data than 2-PL IRT model.

*Keywords:* Explanatory item response models, latent regression, TIMSS 2007 Turkey

#### Summary

##### Purpose

The purpose of this study is to illustrate the application of latent regression two-parameter logistic (2-PL) model as an explanatory item response model (EIRM) on TIMSS 2007 Science data

\* Yrd. Doç. Dr. Burcu ATAR, Hacettepe Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, burcua@hacettepe.edu.tr

\*\* Dr. Derya ÇOBANOĞLU AKTAN, Hacettepe Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, dcaktan@hacettepe.edu.tr

for Turkish students. For this purpose, 2-PL IRT model was formulated under the generalized linear and nonlinear mixed models (GLMM and NLMM) framework. With this framework, it is possible to add item and/or person predictors into the model to explain the effects of those predictors on item responses, and then the model is called Explanatory Item Response Model (EIRM) (De Boeck & Wilson, 2004). If only person predictors are included to the EIRM it is called person explanatory item response model, also known as latent regression model. To illustrate the application of person explanatory IRT, gender, positive affect toward science, valuing science, self-confidence in learning science, and education level of parents were used as person properties in the latent regression 2-PL model. Through this analysis the effects of those person properties on the student achievement were examined by SAS PROC NLMIXED.

### *Results*

The estimates of the item difficulty and item discrimination parameters were calculated by PROC NLMIXED and MULTILOG programs for 2-PL IRT model. The results showed that item parameter estimates obtained from both programs were very close to each other. To investigate the effects of the person properties on students' achievement, data were analyzed under latent regression 2-PL model by PROC NLMIXED program. Gender, positive affect toward science, and valuing science person properties were not found to have statistically significant effects. On the other hand, self-confidence in learning science and education level of parents had statistically significant effects on students' science achievement. Moreover, to examine the model fit of 2-PL IRT and latent regression 2-PL models, deviance (-2 log likelihood), AIC, and BIC indices were compared. As a result, it was seen that latent regression 2-PL model had a better model fit to data than 2-PL IRT model.

### *Discussion*

With the traditional IRT models only item and person parameters are estimated. These models do not allow us to include item and/or person predictors to explain the differences among items and/or the differences among persons. On the other hand, Explanatory Item Response Models (EIRM) has this advantage over traditional IRT models. In this study, this advantage was used to articulate the variations among eight-grade Turkish students on Science achievement. For the further studies item characteristics such as cognitive domain of the items (e.g. reasoning) can be included into analyses in addition to the person characteristics. Moreover, EIRM can be employed to analyze multidimensional and polytomous data.

### *Conclusion*

It was demonstrated that 2-PL IRT model can be utilized for both descriptive and explanatory purposes to analyze test data. Our findings show that this approach provides similar results with the previous research.

## Giriş

Eğitimde madde tepki kuramı (MTK) modelleri, genellikle bireylerin örtük yeteneklerinin ölçülmesi ve bireylerin örtük yeteneklerinin ölçülmesinde yararlanılan ölçme aracındaki maddelerin analizi amacıyla ölçme modelleri olarak kullanılırlar (Lord, 1980; Embretson & Reise, 2000). Gerek çoktan seçmeli maddeler gibi iki kategorili puanlanan maddeler, gerekse uzun veya kısa yanıtli maddeler gibi çok kategorili puanlanan maddeler için kullanılan MTK modelleri olsun, birey ve maddelere özgü etkileri tanımlamak amacıyla kullanıldıkları için "tanımlayıcı" modeller olarak ele alınırlar (De Boeck ve Wilson, 2004).

Bir araştırmanın amacı eğer bireylerin aldıkları puanları veya madde parametrelerini tanımlamak ise, MTK modelleri yeterli olmakla birlikte, araştırmanın amacının bireyler ve/veya maddeler arasındaki farklılıkları açıklamak olduğu durumlarda, MTK modelleri tek başına

yeterli olmamaktadır. Bireylerin puanları arasındaki farklılıkları açıklama veya modelleme işlemi genellikle iki adımlı olarak yapılmaktadır. İlk adımda bireylerin testten elde ettiği puanlar hesaplanarak tanımlama işlemi yapılmakta, ikinci adımda ise puanlar arasındaki farklılıkları açıklayan ya da yordayan bireye ait değişkenlerin etkileri çeşitli yöntemlerle incelenmektedir. Uzun ve Öğretmen (2010), Aypay, Erdoğan ve Sözer'in (2007) çalışmaları bu konudaki örnekler olarak gösterilebilir. Açıklayıcı madde tepki kuramı modelleri (AMTK) ise bu tanımlayıcı ve açıklayıcı işlemlerin eş-zamanlı olarak yapılmasına olanak sağlayan modellerdir (De Boeck ve Wilson, 2004).

*Açıklayıcı madde tepki modelleri*, MTK modellerine bireylerin ve maddelerin özelliklerini eklemeye ve bunların etkilerini tek adımda incelemeye imkân sağlayan *açıklama ve tanımlama* özelliklerini birleştiren modelleme yaklaşımıdır. De Boeck ve Wilson (2004) modele dahil edilen yordayıcılara göre dört model tarif etmişlerdir (Tablo 1). Bunlardan hiçbir yordayıcının modele dahil edilmediği model çifte *tanımlayıcı* olarak adlandırılmıştır. Bu modeller bilinen MTK modelleridir. Birey yordayıcılarının dahil edildiği model, birey açıklayıcı; madde yordayıcılarının eklendiği ise madde açıklayıcı modeldir. Her iki özelliğin eklenmesi ile çifte açıklayıcı model elde edilir. Bu çalışmada ise birey yordayıcıları 2-PL MTK modeline eklenerek TIMSS 2007 Türkiye fen bilimleri verisindeki bireyler arasındaki farklar incelenmektedir.

Tablo 1.  
*Çeşitli AMTK Modelleri*

	Birey Yordayıcıları	
Madde Yordayıcıları	Özelliklerin dahil edilmemesi	Özelliklerin dahil edilmesi (Birey)
Özelliklerin dahil edilmemesi	Çifte tanımlayıcı	Birey açıklayıcı
Özelliklerin dahil edilmesi (Madde)	Madde açıklayıcı	Çifte açıklayıcı

(DeBoeck ve Wilson, 2004, s.47)

Birey ve/veya madde yordayıcılarının modele eklenmesi, madde tepki kuramı modellerinin genelleştirilmiş doğrusal ve doğrusal olmayan karma modeller (GDKM ve DOKM) çerçevesinde ele alınmasıyla mümkün olmaktadır (McCulloch ve Searle, 2001). Aşağıdaki bölümlerde bu çalışmada kullanılan 2-PL MTK modelinin ve bu modele birey yordayıcılarının eklenmesiyle ortaya çıkan birey açıklayıcı, *örtük regresyon 2-PL modelinin* GDKM ve DOKM çerçevesinde formülasyonu ayrıntılarıyla açıklanmaktadır.

#### *GDKM ve DOKM Çerçevesinde 2-PL MTK Modelleri*

Bilinen MTK modelleri GDKM ve DOKM'nin özel durumlarıdır. Bu modeller MTK modellerinin çok seviyeli modeller olarak formüle edilmesine imkân tanurlar. MTK çerçevesinde, her bir birey testteki birçok sayıda maddeye yanıt verir ve bu sebeple testte her bir bireyin tekrar tekrar gözlemlendiği düşünülür (De Boeck ve Wilson, 2004). Bu sebeple çok seviyeli yaklaşımda her bir birey için madde gözlemleri gruplanabilir. Bu durumda, en düşük seviyede (seviye 1) analiz birimi madde ve ikinci seviyede ise (seviye 2) analiz birimi bireydir. Bu açıdan, 2-PL MTK modeli iki seviyeli genelleştirilmiş doğrusal olmayan karma model (DOKM) olarak ele alınır. MTK modellerinin çok seviyeli formülasyonunda, verilen bir örtük yetenek için bir maddeye belli bir cevap verme olasılığı, doğrusal yordayıcı, bağlantı fonksiyonu (link function) ve dağılım ailesi yoluyla belirlenir (Rabe-Hesketh, Pickles, & Skrondal, 2004; Skrondal, & Rabe-Hesketh, 2004).

$J$  bireyinin  $i$  maddesine verdiği yanıt,  $Y_{ij}$ , iki kategorili bir değişken ve  $j$  bireyinin yeteneği,  $\theta_j$  sürekli bir değişken olduğunda 2-PL modelde,  $i$  maddesine doğru yanıt verme olasılığı, bireyin

örtük yeteneği (birey parametresi) ile madde güçlüğü ve ayırıcılığının (madde parametreleri) bir fonksiyonu olarak modellenir. Genelleştirilmiş doğrusal olamayan karma modelde ise 2-PL modeli olarak aşağıdaki biçimde modellenir.

$i$  madde ( $i=1,2,\dots, I$ ) ve  $j$  bireyi için ( $j=1,2,\dots, J$ ) 2-PL modelinde doğrusal yordayıcı formülü Eşitlik 1'de gösterilmektedir.

$$\eta_{ij} = x_{ij}' \beta + \theta_{1j}^{(2)} z_{ij}^{(2)} \alpha_i^{(2)} = \beta_i + \theta_{1j}^{(2)} \alpha_i^{(2)} \quad \text{Eş. 1}$$

Burada  $x_{ij}$  seviye 1'deki  $i$  sabit etki için  $i$  boyutlu madde yordayıcıları (madde göstergeleri) vektörüdür.  $x_{ij}$  vektörünün  $j$  bireyi için  $i$  ninci elemanı 1'e diğer elemanlar ise 0'a eşittirler.  $\beta$  ise  $x_{ij}$  yordayıcı değişkenler vektörü ile ilişkili  $i$  boyutlu katsayılar veya sabit etkiler vektörüdür.  $z_{ij}^{(2)}$  seviye 2'deki  $j$  rastlantısal etki için  $j$  boyutlu birey yordayıcıları (birey göstergeleri) vektörüdür.  $i$  maddesi için  $z_{ij}^{(2)}$  vektörünün  $j$  inci elemanı 1'e diğer elemanlar ise 0'a eşittirler.  $\theta_{1j}^{(2)}$ ,  $z_{ij}^{(2)}$  yordayıcı değişkenler matrisi ile ilişkili  $j$  bireyi için  $j$  boyutlu rastlantısal etkiler vektörüdür.  $\theta_{1j}^{(2)}$   $j$  bireyinin yeteneğidir.  $\left[-\frac{\beta}{\alpha_i^{(2)}}\right]$   $i$  maddesinin madde güçlük parametresi,  $\alpha_i^{(2)}$   $i$  maddesinin madde ayırıcılık parametresidir.  $\theta_{1j}^{(2)}$  örtük değişkeninin sıfır ortalama ve  $\sigma^2$  varyansı ile normal olarak dağıldığı varsayılır. ( $\theta_{1j}^{(2)} \sim N(0, \sigma^2)$ ).

Yukarıdaki Eşitlik 1 matris formatında aşağıdaki biçimde yazılır,

$$\begin{bmatrix} \eta_{1j} \\ \eta_{2j} \\ \vdots \\ \eta_{ij} \\ \vdots \\ \eta_{Ij} \end{bmatrix} = \begin{bmatrix} 1 & 0 & L & 0 \\ 0 & 1 & L & 0 \\ \vdots & \vdots & \vdots & \vdots \\ M & M & M & M \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & L & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_j \\ \vdots \\ \beta_I \end{bmatrix} + \theta_{1j}^{(2)} \begin{bmatrix} 1 & 0 & L & 0 \\ 0 & 1 & L & 0 \\ \vdots & \vdots & \vdots & \vdots \\ M & M & M & M \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & L & 1 \end{bmatrix} \begin{bmatrix} \alpha_1^{(2)} \\ \alpha_2^{(2)} \\ \vdots \\ \alpha_i^{(2)} \\ \vdots \\ \alpha_I^{(2)} \end{bmatrix} = \begin{bmatrix} \beta_1 + \theta_{1j}^{(2)} \alpha_1^{(2)} \\ \beta_2 + \theta_{1j}^{(2)} \alpha_2^{(2)} \\ \vdots \\ M \\ \beta_I + \theta_{1j}^{(2)} \alpha_I^{(2)} \end{bmatrix}$$

Genellikle verilen belli sabit ve rasgele etkiler ile belirli yordayıcıların yanıt dağılımları için iki parametrelili MTK modellerinde Bernoulli dağılımı kullanılır.

$i$  maddesine  $j$  bireyinin verdiği  $Y_{ij}$  yanıtının beklenen değeri, verilen belirli yordayıcı, sabit ve rastlantısal etkiler için,  $\eta_{ij}$  doğrusal yordayıcısına logit fonksiyonu ile bağlanır.

$$\text{logit}(E[Y_{ij} | x_{ij}, z_{ij}, \beta_i, \alpha_i, \theta_{1j}]) = \eta_{ij} \quad \text{Eş. 2}$$

$$E[Y_{ij} | x_{ij}, z_{ij}, \beta_i, \alpha_i, \theta_{1j}] = \Pr(Y_{ij} = 1 | x_{ij}, z_{ij}, \beta_i, \alpha_i, \theta_{1j}) = \pi_{ij} \quad \text{Eş. 3}$$

$$\text{logit}(\pi_{ij}) = \ln\left(\frac{\pi_{ij}}{1 - \pi_{ij}}\right) = \beta_i + \theta_{1j}^{(2)} \alpha_i^{(2)} \quad \text{Eş. 4}$$

Bu durumda logit bağlantı fonksiyonu aşağıdaki biçimde tekrar yazılabilir. Her iki tarafın eksponansiyeli alındığında, 2-PL modelinin eksponansiyel formunu elde ederiz.

$$\frac{\pi_{ij}}{1 - \pi_{ij}} = \exp(\beta_i + \theta_{1j}^{(2)} \alpha_i^{(2)}) \quad \text{Eş. 5}$$

$$\pi_{ij} = \frac{\exp(\beta_i + \theta_{1j}^{(2)} \alpha_i^{(2)})}{1 + \exp(\beta_i + \theta_{1j}^{(2)} \alpha_i^{(2)})} \quad \text{Eş. 6}$$

*Birey Açıklayıcı Model: İki Seviyeli Örtük Regresyon İki Parametrelili Lojistik MTK Modeli*

Birey yordayıcılarının 2-PL MTK modeline eklendiği durumda elde edilen model, örtük regresyon 2-PL MTK modeli olarak isimlendirilir. 2-PL MTK modeline birey yordayıcılarını eklemek için modelde  $\theta_{1j}^{(2)}$ , aşağıdaki regresyon eşitliğiyle yer değiştirir.

$$\theta_{1j}^{(2)} = \sum_{p=1}^P \theta_p Z_{pj} + \varepsilon_j \quad \text{Eş. 7}$$

Eşitlik 7'de  $Z_{pj}$  j bireyinin p'ninci birey özelliği ( $p = 1, 2, \dots, P$ ),  $\theta_p$  p'ninci birey özelliği için birey regresyon ağırlığı ve  $\varepsilon_j$  ise birey özelliklerinin etkileri göz önünde bulundurulduktan sonra kalan birey etkisidir.  $\varepsilon_j$ 'nin sıfır ortalama ve  $\sigma_\varepsilon^2$  varyansı ile normal olarak dağıldığı varsayılır ( $\varepsilon_j \sim N(0, \sigma^2)$ ).

Açıklayıcı madde tepki kuramı modellerinin nispeten yeni bir yaklaşım olması nedeni ile literatürde uygulamalarına az rastlanmaktadır. Bu konudaki başlıca kaynak olan De Boeck ve Wilson'ın (2004) editörlüğünü yaptığı *Explanatory Item Response Models*'ta bu yaklaşımın çerçevesi ve bazı uygulamaları açıklanmaktadır. Braeken ve Tuerlinckx (2009) çalışmasında, madde açıklayıcı ve birey açıklayıcı modellerinin MATLAB programıyla uygulanmasına imkân tanıyan paket tanıtılmıştır. Briggs (2008) çalışmasında ise fen başarısındaki etnik farklılıkları açıklamak ve tanımlamak için *örtük regresyon Rasch modelini* kullanmıştır. Bu modeli kullanarak elde ettiği sonuçları Briggs (2008) *iki adımlı* (tanımlamayı takiben açıklama) yaklaşımına karşılaştırmıştır. Araştırması sonunda Briggs (2008), test puanlarının güvenilirliğinin yüksek olduğu durumlarda iki yaklaşımın sonuçlarının benzer olacağını, aksi durumda ise, iki adımlı yaklaşımın birinci adımında yetenek kestirimlerinin evren ortalamasına doğru daralma göstereceğini ve bununla bulguların yorumlanmasında farklılıklar oluşturabilecek küçülmüş (attenuated) regresyon katsayılarına neden olacağı sonucuna varmıştır. Briggs'in (2008) çalışmasında, birey özelliklerinin öğrenci başarısındaki farklılıklara etkileri örtük regresyon Rasch modeli ile incelenmiştir. Bu çalışmada ise TIMSS 2007 Türkiye fen verisi örtük regresyon 2-PL modellemesi kullanılarak incelenmektedir. Bu incelemenin ayrıntıları ve bulguları aşağıdaki bölümlerde verilmiştir.

## Yöntem

### Veriler

Bu çalışmada, 2007 yılı Uluslararası Matematik ve Fen Eğilimleri Çalışması (**Trends in International Mathematics and Science Study** (TIMSS)) Türkiye 8. sınıf fen testi yedi numaralı kitapçığından elde edilen test verileri ve öğrencilerin bazı demografik bilgileri incelenmiştir. TIMSS, Eğitim Başarısını Değerlendirme Uluslararası Birliği (International Association for the Evaluation of Educational Achievement (IEA)) tarafından 1995'ten bu yana dört yıllık aralıklarla (1999, 2003, 2007) uluslararası ölçekte farklı ülkelerin 4. sınıf ve 8. sınıf öğrencilerinin matematik ve fen alanındaki başarılarını ölçmek ve karşılaştırmak için yapılan bir çalışmadır. Akademik başarının yanı sıra TIMSS çalışmasında, öğrenci, öğretmen ve okul yöneticisi anketleri ile fen ve matematik öğreniminin gerçekleştiği çerçeveyi betimleyen veriler de toplanmaktadır. Öğrencilerin demografik özellikleri, anne ve babalarının eğitim düzeyleri, matematik ve fen ile ilgili tutumları ve görüşleri, matematik ve fen derslerindeki sınıf içi etkinlikleri ile ilgili bilgiler bunların arasında yer almaktadırlar. Sekizinci sınıf TIMSS fen testi fizik, kimya, biyoloji ve yer bilimi olmak üzere dört ana bölümden oluşmaktadır. Bu bölümlerde toplam 227 adet soru bulunmaktadır. Bu soruların 120 tanesi açık uçlu ve 107 tanesi ise çoktan seçmeli sorulardır. Bunlar 14 farklı kitapçık içerisine matris örneklem yaklaşımıyla dağıtılarak öğrencilere sorulmuştur. Testin bu yapısından dolayı eşitleme konusuna girmemek ve çalışmada sadece açıklayıcı madde tepki modelleri yaklaşımına odaklanmak amacıyla bu çalışmada, 7 numaralı

kitapçığı yanıtlayan 327 adet öğrencinin verileri kullanılmıştır. Bu öğrencilerin 159'ü bayan ve 168'i erkektir. Kitapçık seçimi rasgele bir şekilde yapılmıştır. Bu kitapçıkta toplam 31 adet fen sorusu bulunmaktadır. Ancak çalışmaya sadece cevabı iki kategorili olan sorular dahil edildiğinden en yüksek puanı 2 olan iki adet soru çıkartılıp toplam 29 madde ile analizlere başlanmış, yakınsama problemi nedeni ile bir soru daha analizlerden çıkartılmıştır.

Test verilerinin yanı sıra, birey özelliklerini karakterize etmek için, çalışmada öğrenci anketleri verilerinden 14 tanesi kullanılarak dört adet indeks oluşturulmuştur. Bu indeksler, fene verilen önem, fene karşı olumlu tutum, fene yönelik özgüven ve ebeveynlerin eğitim düzeyleridir. Bu indekslerden ilk üçü, TIMSS 2007 (Olson ve ark., 2008) teknik raporunun 12. bölümünde yer alan bireysel özellik indekslerinden ve sonuncusu ise Aypay, Erdoğan ve Sözer (2007) çalışmasındaki sosyoekonomik düzey indeksinden türetilmiştir. Bu indekslerin detayları Tablo 2'de görülmektedir. Bu indekslerin haricinde cinsiyet değişkeni de analizlere bireysel özellik olarak eklenmiştir.

Tablo 2.

*Bireysel Özellik İndeksleri.*

İndeks adı	
Fene verilen önem (ÖNEM)	Fen bana günlük hayatımda yardımcı olacaktır Diğer konuları öğrenmek için fene ihtiyacım var Üniversiteye girmek için fene ihtiyacım var İstediğim işe girmek için fene ihtiyacım var
Fene karşı olumlu tutum (TUTUM)	Feni seviyorum Fen öğrenmeyi sevmiyorum Fen sıkıcıdır Başka fen dersi almak isterim
Fene yönelik özgüven (GÜVEN)	Genellikle fende iyiyimdir Fen benim için daha zordur Fen kuvvetli olduğum bir konu değildir Fen konularını çabuk öğrenirim
Ebeveynlerin eğitim düzeyleri (DÜZEY)	Annenin eğitim düzeyi Babanın eğitim düzeyi

*Verilerin analizi*

Tanımlayıcı MTK modellerinin analizinde genellikle BILOG-MG, PARSCALE, MULTILOG ve benzeri gibi paket programlar kullanılmakla beraber bu programlar açıklayıcı MTK modellerinin analizinde yetersiz kalmaktadır. De Boeck ve Wilson'ın (2004) önerdiği yazılımlardan olan SAS PROC NLMIXED (SAS Institute Inc., 1999) ise hem tanımlayıcı hem de açıklayıcı MTK modellemesi yapılabilmektedir. Bu çalışmada öncelikle 2-PL MTK modellemesi MULTILOG ve PROC NLMIXED ile yapılmış ve analiz sonuçları karşılaştırılmıştır. Daha sonra örtük regresyon 2-PL modellemesi PROC NLMIXED ile yapılmıştır. Sonuçlar bulgular kısmında ayrıntılarıyla verilmektedir.

**Bulgular**

*İki Parametrelili Lojistik Madde Tepki Kuramı Modeli (2-PL MTK Modeli)*

SAS PROC NLMIXED ile yapılan tanımlayıcı 2-PL MTK analizi sonucunda madde güçlük parametre tahminleri doğrudan verilmemektedir.  $i$  maddesi için madde güçlük parametre tahmini,  $\beta_i$  tahmininin  $\alpha_i$  tahminine bölünmesi sonucu elde edilir (Tablo 3). Madde ayrıricılık parametre tahminleri ( $\alpha_i$ ) ise doğrudan verilmektedir. Aynı analiz MULTILOG programı ile yapıldığında, benzer madde güçlük ve ayrıricılık parametre değerleri elde edilmiştir.

2-PL MTK modelinin PROC NLMIXED analizi sonucunda, madde güçlük parametreleri yaklaşık -5.89 logit (madde 6) ile 4.35 logit (madde 15) değerleri arasında tahmin edilmiş olup, ortalama madde güçlüğü yaklaşık 0.45 olarak bulunmuştur (Tablo 3). Verinin MULTILOG programı

ile analizi sonucunda, tahmin edilen madde güçlük parametreleri yaklaşık -6.09 logit (madde 6) ile 4.38 logit (madde 15) değerleri arasında değişmekte olup, ortalama madde güçlüğü yaklaşık 0.43'tür (Tablo 3). Madde güçlüğü 0 olan bir maddeyi, ortalama bir öğrencinin doğru cevaplama olasılığının %50 olduğu göz önüne alınırsa, ortalama bir öğrencinin madde güçlük değeri 0.43 olan bir maddeyi doğru cevaplama olasılığının %50'den düşük olduğu yorumuna ulaşılabilir.

PROC NLMIXED analizi sonucunda, madde ayırıcılık parametreleri yaklaşık 0.09 logit (madde 15) ile 2.84 logit (madde 22) değerleri arasında değişmekte olup, ortalama madde ayırıcılığı yaklaşık 1.34'tür (Tablo 3). Verinin MULTILOG analizi sonucunda tahmin edilen madde ayırıcılık parametreleri yaklaşık 0.09 logit (madde 15) ile 2.85 logit (madde 22) değerleri arasında değişmekte olup ortalama madde ayırıcılığı yaklaşık 1.35'tir (Tablo 3). İki programla tahmin edilen madde ayırıcılıklarının değerleri birbirlerine oldukça yakındır.

Tablo 3.

*Madde Güçlük ve Ayırıcılık Parametre Tahminleri*

	Madde Güçlüğü			Madde Ayırıcılığı		
	MULTILOG	PROC NLMIXED	PROC NLMIXED Örtük	MULTILOG	PROC NLMIXED	PROC NLMIXED Örtük
	2-PL Modeli	2-PL Modeli	regresyon 2- PL	2-PL Modeli	2-PL Modeli	regresyon 2- PL
Madde 1	-1.52 (0.29)	-1.51 (0.24)	-1.85 (0.45)	1.14 (0.29)	1.13 (0.22)	0.91 (0.18)
Madde 2	0.22 (0.11)	0.24 (0.10)	0.34 (0.39)	1.82 (0.29)	1.82 (0.25)	1.44 (0.20)
Madde 3	-0.09 (0.09)	-0.07 (0.09)	-0.05 (0.39)	2.16 (0.33)	2.15 (0.29)	1.65 (0.23)
Madde 4	0.00 (0.09)	0.02 (0.09)	0.05 (0.39)	2.33 (0.33)	2.33 (0.32)	1.83 (0.25)
Madde 5	0.36 (0.36)	0.38 (0.27)	0.53 (0.51)	0.48 (0.13)	0.47 (0.13)	0.37 (0.10)
Madde 6	-6.09 (4.46)	-5.89 (3.14)	-7.95 (4.61)	0.38 (0.20)	0.39 (0.22)	0.28 (0.17)
Madde 7	0.17 (0.17)	0.18 (0.13)	0.27 (0.41)	1.02 (0.16)	1.03 (0.17)	0.80 (0.13)
Madde 8	0.81 (0.29)	0.84 (0.22)	1.10 (0.48)	0.76 (0.15)	0.75 (0.15)	0.59 (0.12)
Madde 9	3.57 (2.19)	3.63 (1.73)	4.09 (1.73)	0.29 (0.13)	0.28 (0.14)	0.26 (0.11)
Madde 10	0.70 (0.13)	0.72 (0.11)	0.95 (0.41)	1.93 (0.34)	1.92 (0.27)	1.51 (0.21)
Madde 11	1.68 (0.34)	1.70 (0.27)	2.22 (0.53)	1.07 (0.21)	1.07 (0.19)	0.83 (0.15)
Madde 12	-0.24 (0.12)	-0.22 (0.10)	-0.25 (0.39)	1.58 (0.29)	1.55 (0.22)	1.25 (0.18)
Madde 13	0.23 (0.14)	0.25 (0.12)	0.35 (0.40)	1.33 (0.24)	1.31 (0.19)	1.05 (0.15)
Madde 14	-0.03 (0.10)	-0.02 (0.09)	0.01 (0.39)	2.00 (0.31)	1.97 (0.27)	1.57 (0.21)
Madde 15	4.38 (7.73)	4.35 (5.81)	7.32 (12.62)	0.09 (0.12)	0.09 (0.12)	0.06 (0.09)
Madde 16	-0.44 (0.10)	-0.43 (0.10)	-0.50 (0.38)	1.98 (0.32)	1.96 (0.28)	1.57 (0.22)
Madde 17	0.18 (0.11)	0.19 (0.10)	0.28 (0.39)	1.75 (0.26)	1.74 (0.24)	1.37 (0.19)
Madde 18	-0.38 (0.17)	-0.37 (0.14)	-0.42 (0.40)	1.05 (0.22)	1.03 (0.17)	0.85 (0.14)
Madde 19	0.98 (0.13)	1.00 (0.12)	1.30 (0.42)	2.23 (0.39)	2.22 (0.33)	1.76 (0.26)
Madde 20	1.18 (0.14)	1.21 (0.13)	1.60 (0.43)	2.56 (0.47)	2.55 (0.41)	1.89 (0.30)
Madde 21	0.57 (0.12)	0.59 (0.11)	0.78 (0.41)	1.93 (0.31)	1.92 (0.27)	1.51 (0.21)
Madde 22	1.25 (0.13)	1.27 (0.13)	1.66 (0.43)	2.85 (0.55)	2.84 (0.49)	2.18 (0.37)
Madde 23	0.09 (0.29)	0.11 (0.22)	0.18 (0.46)	0.56 (0.14)	0.55 (0.13)	0.46 (0.11)
Madde 24	2.23 (1.54)	2.25 (1.20)	1.72 (1.40)	0.25 (0.12)	0.25 (0.13)	0.21 (0.10)
Madde 25	1.26 (0.22)	1.29 (0.19)	1.69 (0.47)	1.25 (0.22)	1.23 (0.20)	0.96 (0.16)
Madde 26	0.41 (0.19)	0.43 (0.15)	0.59 (0.42)	1.00 (0.17)	1.00 (0.16)	0.79 (0.13)
Madde 27	-0.07 (0.14)	-0.05 (0.11)	-0.03 (0.39)	1.30 (0.24)	1.30 (0.19)	1.04 (0.15)
Madde 28	0.60 (0.33)	0.63 (0.24)	0.83 (0.49)	0.60 (0.14)	0.59 (0.14)	0.47 (0.11)

*Örtük Regresyon İki Parametrelili Lojistik Modeli*

2- PL MTK modeline cinsiyet, fene karşı olumlu tutum (TUTUM), fene verilen önem (ÖNEM), fene yönelik özgüven (ÖZGÜVEN) ve ebeveynlerin eğitim düzeyi (DÜZEY) birey yordayıcılarının eklenmesi ile örtük regresyon 2-PL modeli oluşturulmuştur. Bu yordayıcıların öğrencilerin fen başarıları üzerindeki etkilerinin tahminleri ve bu tahminlerin standart hataları Tablo 4'te verilmiştir. Her tahminin  $p$ -değerine bakıldığında, öğrencilerin fen başarılarının cinsiyet, fene karşı olumlu tutum veya fene önem vermeye göre 0.05 alfa düzeyinde istatistiksel olarak anlamlı bir farklılık göstermediği görülmektedir. Halbuki öğrencilerin fene yönelik özgüvenleri ile ebeveynlerinin eğitim düzeylerinin fen başarılarında 0.05 alfa düzeyinde istatistiksel olarak anlamlı olduğu bulunmuştur.

Fene yönelik özgüvenin tahmin edilen etkisi 0.15 logit olup, bu bulgu özgüven puanındaki bir birimlik artışın, fen başarı puanında 0.15 logitlik bir artışa neden olacağı şeklinde yorumlanabilir. Logit ölçeğindeki bu değer, ekponansiyeli alındığında odds oranı olarak ifade edilebilir. Odds oranı, herhangi bir maddeyi doğru cevaplama olasılığının yanlış cevaplama olasılığına oranı olarak tanımlanır; doğru cevaplama olasılığının 0.50 olması durumunda odds oranı 1'e eşittir (Eşitlik 5). Logit ölçeğindeki çalışmada elde edilen 0.15 değerinin odds oranı ise 1.16'dır. Bu da doğru cevaplama olasılığını 0.50'den (1 odds oranı) 0.54'e (1.16 odds oranı) yükseltecektir. Ebeveynlerin eğitim düzeyinin tahmin edilen etkisi ise 0.28 logit olup, ebeveynlerin eğitim düzeyindeki bir birimlik artışın, fen başarı puanında 0.28 logitlik bir artışa neden olacağı, diğer bir ifadeyle olasılığı 0.50'den 0.57'ye yükselteceği şeklinde yorumlanabilir.

Tablo 4.

*Örtük Regresyon 2-PL Modelinde Öğrenci Özelliklerinin Tahmini Etkileri*

Öğrenci Özellikleri	Tahmin	Standard Hata	Serbestlik Derecesi	$t$ -değeri	$p$ -değeri
CİNSİYET	-0.16	0.13	326	-1.24	0.22
TUTUM	-0.01	0.03	326	-0.44	0.66
ÖNEM	0.02	0.02	326	1.01	0.31
GÜVEN	0.15	0.03	326	5.38	< 0.001
DÜZEY	0.28	0.03	326	8.35	< 0.001

2-PL MTK modellemesinde olduğu gibi örtük regresyon 2-PL modellemesinde de madde parametreleri elde edilir. Ancak PROC NLMIXED ile yapılan analiz sonucunda madde güçlük parametreleri doğrudan verilmemektedir. Bu değerlerin hesaplanması 2-PL modelinde yapılan hesaplamalardan farklıdır.  $i$  maddesi için madde güçlük parametre tahmini,  $\left(\frac{\beta}{\alpha_i}\right)$  değerinden ortalama öğrenci etkisinin değeri çıkarıldığında elde edilen değerdir (Tablo 3). Ortalama öğrenci etkisi ise Eşitlik 7'ye göre aşağıdaki şekilde hesaplanır:

CİNSİYET Etkisi \* Ort. CİNSİYET + TUTUM \* Ort. TUTUM + ... + DÜZEY Etkisi \* Ort. DÜZEY + Ort. Hata

PROC NLMIXED yazılımı ile örtük regresyon 2-PL modeli analizi sonucunda elde edilen  $\alpha_i$  tahminleri,  $i$  maddesi için madde ayırıcılık parametre tahminini verir. Analiz sonucunda, madde ayırıcılık parametreleri yaklaşık 0.06 logit ile 2.18 logit değerleri arasında değişmekte olup ortalama madde ayırıcılığı yaklaşık 1.05'tir (Tablo 3). Örtük regresyon 2-PL modeli ile tahmin edilen madde ayırıcılık parametre değerleri, 2-PL MTK modeli ile tahmin edilen madde ayırıcılık parametre değerlerine yakındır.

2-PL MTK ile örtük regresyon 2-PL modellerinin uyum istatistikleri -2 log likelihood, Akaike Bilgi Ölçütü (Akaike Information Criterion, AIC) (Akaike, 1974) ve Bayesian Bilgi Ölçütü (Bayesian Information Criterion, BIC) (Schwarz, 1978) indeksleri kullanılarak karşılaştırılmıştır (Tablo 5). İndeksler ne kadar küçükse model uyumunun o kadar iyi olduğu söylenir. İndeksle-



göre örtük regresyon 2-PL modelinin uyumu 2-PL MTK modeline göre daha iyidir. İki model arasındaki -2 log likelihood farkları istatistiksel olarak anlamlı bulunmuştur ( $\chi^2 = 11.07 (0.05) < 328$ ).

Tablo 5.  
*Uyum İstatistikleri*

	2-PL MTK Modeli	Örtük Regresyon 2-PL Modeli
-2 Log Likelihood	10148	9820
AIC (smaller is better)	10260	9946
BIC (smaller is better)	10472	10185

### Tartışma ve Sonuç

Bu çalışmada, TIMSS 2007 Türkiye sekizinci sınıf fen bilimleri verisi ilk olarak 2-PL MTK modeli ile PROC NL MIXED ve MULTILOG programları kullanılarak analiz edilmiştir. Analizler sonucunda tahmin edilen madde güçlük ve madde ayırıcılık parametre değerleri, birbirlerine çok yakın bulunmuştur. Daha sonra, bireylerin fen başarıları arasında gözlenen farklılıkları açıklamak üzere örtük regresyon 2-PL modeli çerçevesinde birey özelliklerinden cinsiyet, fene karşı olumlu tutum, fene verilen önem, fene yönelik özgüven ve ebeveynlerin eğitim düzeyi 2-PL MTK modeline eklenmiştir. Modelde yer alan birey özelliklerinden sadece bireylerin fene yönelik özgüvenleri ile ebeveynlerinin eğitim düzeylerinin fen başarılarında istatistiksel olarak anlamlı etkileri olduğu bulunmuştur. Bu sonuçlar ilgili literatürdeki sonuçlara benzerdir (Uzun ve Öğretmen, 2010; Aypay, Erdoğan ve Sözer, 2007). Bu çerçevede elde edilen madde güçlük parametreleri 2-PL modelden elde edilen madde güçlük parametrelerine yakın olmakla birlikte biraz fark göstermektedir. Bunun nedeni ise yukarıda belirtildiği üzere ortalama bireysel etkilerin hesaplamalara katılmasıdır. Son olarak da örtük regresyon ile 2-PL modellerinin uyum indeksleri karşılaştırılmış, örtük regresyon modelinin veriye daha iyi uyum gösterdiği bulunmuştur. Literatürde bunun bir sebebinin modelde tahmin edilen parametre sayısının artması olarak gösterilmektedir (De Boeck ve Wilson, 2004).

Madde tepki kuramı modelleri, geleneksel olarak formüle edildiğinde yalnızca madde güçlüğü, madde ayırıcılığı gibi madde parametrelerinin ve bireylerin örtük puanları gibi birey parametrelerinin tahmin edilmesi ile sınırlı olup sadece ölçme modeli olarak kullanılırlar. Bu çalışmada olduğu gibi, MTK modelleri genelleştirilmiş doğrusal ve doğrusal olmayan karma modeller çerçevesinde ele alındığında, birey özellikleri modele eklenebilmekte ve böylece bireyler arasında gözlenen farklılıkları eşzamanlı istatistiksel analiz ile açıklamak mümkün olmaktadır. Bu yaklaşım ile yapılan formülasyonlarda birey özelliklerinin yanı sıra madde özellikleri de modele eklenerek istatistiksel analizler gerçekleştirilebilir. Rasch modeli ve 1-PL modelinin literatürde bu konudaki örnekleri bulunmakla birlikte, 2-PL ve diğer MTK modelleri için uygulamalarına pek rastlanmamaktadır.

Bu çalışmanın literatüre olan katkısı Rasch ve 1-PL MTK modeli uygulamalarına 2-PL uygulamasını eklemesi ve örneklemesidir. Bu çalışmada, yalnızca bireyler arasındaki farklılıklar açıklanmaya çalışılmış, maddeler arasındaki farklılıkları açıklamak üzere bir model oluşturulmamıştır. Birey özelliklerinin de modele eklenmesiyle maddeler arasındaki farklılıklar açıklanabilir. Bunların yanı sıra, açıklayıcı MTK modelleri çok boyutlu, çok kategorili puanlanan testleri de analiz etmede kullanılabilir. Bu da Türkiye’de uygulanan testlerin daha değişik yöntemlerle analizine olanak sağlar.

### Kaynakça

Akaike, M. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716-723.

- Aypay, A., Erdoğan, M., & Sözer, M. A. (2007). Variation among schools on classroom practices in science based on TIMSS-1999 in Turkey. *Journal Of Research In Science Teaching*, 44(10), 1417-1435.
- Braeken, J., & Tuerlinckx, F., (2009). Investigating latent constructs with item response models: A MATLAB IRTm toolbox. *Behaviour Research Methods*, 41, (4), 1127-1137.
- Briggs, D. C. (2008). Using explanatory item response models to analyze group differences in science achievement. *Applied Measurement in Education*, 21, 89-118.
- De Boeck, P., & Wilson, M. (2004). *Explanatory item response models: A generalized linear and nonlinear approach*. New York: Springer
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- McCulloch, C. E. and Searle, S. R. (2001). *Generalized, linear, and mixed models*. New York: Wiley.
- Olson, J.F., Martin, M.O., & Mullis, I.V.S. (Eds.). (2008). *TIMSS 2007 Technical Report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2004). *GLAMM manual*. U. C. Berkeley Division of Biostatistics Working Paper Series. Working Paper 160.
- SAS Institute Inc. (1999). *SAS 8 Help and Documentation*, Cary, NC: SAS Institute Inc.
- Schwarz, G. (1978). Estimating a dimension of a model. *Annals of Statistics*, 6, 461- 464.
- Skrondal, A. & Rabe-Hesketh, S. (2004). *Generalized latent variable modeling: Multilevel, longitudinal and structural equation models*. Boca Raton: Chapman & Hall/CRC.
- Uzun, B. ve Öğretmen, T. (2010). Fen Başarısı ile İlgili Bazı Değişkenlerin TIMSS-R Türkiye Örneğinde Cinsiyete Göre Ölçme Değişmezliğinin Değerlendirilmesi. *Eğitim ve Bilim*, 35, 26-35.