

ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

EĞİTİM BİLİMLERİ ANABİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME PROGRAMI

SÜREÇ VERİLERİ İLE BİREY VE MADDE
ÖZELLİKLERİNİN MADDELERİN DOĞRU
YANITLANMA OLASILIĞINA ETKİSİ

DOKTORA TEZİ

FATMA KIZILTOPRAK

ANKARA
ŞUBAT, 2025



**ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

**EĞİTİM BİLİMLERİ ANABİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME PROGRAMI**

**SÜREÇ VERİLERİ İLE BİREY VE MADDE
ÖZELLİKLERİNİN MADDELERİN DOĞRU
YANITLANMA OLASILIĞINA ETKİSİ**

DOKTORA TEZİ

FATMA KIZILTOPRAK

DANIŞMAN: PROF. DR. CELAL DEHA DOĞAN

**ANKARA
ŞUBAT, 2025**

Ankara Üniversitesi Eğitim Bilimleri Enstitüsü Müdürlüğüne,

Fatma KIZILTOPRAK adlı öğrencinin hazırladığı “Süreç Verileri ile Bazı Birey ve Madde Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığına Etkisi” başlıklı bu çalışma Eğitim Bilimleri Anabilim Dalı / Eğitimde Ölçme ve Değerlendirme Programı’nda jüri üyelerince oy birliği / oy çokluğu ile **Doktora Tezi** olarak kabul edilmiştir.

	<u>Jüri Üyeleri</u>	<u>İmza</u>
Başkan	Prof. Dr. Şeref TAN (Başkan)	
Üye	Prof. Dr. Celal Deha DOĞAN (Danışman)	
Üye	Prof. Dr. Nuri DOĞAN	
Üye	Doç. Dr. Ergül DEMİR	
Üye	Doç. Dr. Seher YALÇIN	

ONAY

Bu tez Ankara Üniversitesi Lisansüstü Eğitim Öğretim Yönetmeliği’nin ilgili maddeleri uyarınca, jüri üyeleri tarafından .../.../20... tarihinde, Enstitü Yönetim Kurulu tarafından ise .../.../20... tarihinde kabul edilmiştir.

.....

Prof. Dr. Mehmet İkbal YETİŞİR
Eğitim Bilimleri Enstitüsü Müdürü

ETİK İLKELERE UYGUNLUK BİLDİRİMİ

Tez içindeki bütün bilgileri akademik yazım kurallarına uygun biçimde raporlaştırdığımı ve bunları etik ilkelere (atıfta bulunulan tüm yapıtlara kaynaklarda yer verilmesi, tezde kullanılan bilgi ve belgelere resmi yollarla ulaşılması ve bunların aslı bozulmadan kullanılması vb.) uygun olarak elde ettiğimi ve sunduğumu bildiririm.

Fatma KIZILTOPRAK

ÖZET

SÜREÇ VERİLERİ İLE BİREY VE MADDE ÖZELLİKLERİNİN MADDELERİN DOĞRU YANITLANMA OLASILIĞINA ETKİSİ

KIZILTOPRAK, Fatma

Doktora Tezi, Eğitim Bilimleri Anabilim Dalı

Tez Danışmanı: Prof. Dr. Celal Deha DOĞAN

Şubat, 2025, xii + 90 sayfa

Bilgisayar tabanlı uygulamaların yaygınlaşması ile eğitimde durum belirleme çalışmalarında farklı madde türlerinin kullanımı, otomatik puanlama ve bireye uyarlanabilir testlere geçilmesi gibi çeşitli yenilikler ortaya çıkmaktadır. Bu yeniliklerden son zamanlarda üzerinde en çok çalışılanlardan biri de süreç verilerinin toplanması ve bu veriler yardımıyla bireylerin test alma süreçlerinin ve yapılan değerlendirmelerin geçerliğinin artırılmasıdır. Araştırmanın amacı, Açıklayıcı Madde Tepki Modellemesi (AMTM) ile bireylerin maddeleri doğru yanıtlama olasılıklarını etkileyen birey, madde ve süreç verilerini ve bunlar arasındaki etkileşimleri belirlemektir. Bu amaç doğrultusunda PISA 2015 Türkiye örnekleme ilişkin veriler kullanılmıştır. Birey ve madde özelliklerinden cinsiyet, evdeki kitap sayısı, madde türü, metin biçimi, bağlam, boyut ve madde zorluk düzeyi, süreç verilerinden ise tepki süresi, eylem sayısı ve görüntüleme sayısı değişkenleri çalışma kapsamına dahil edilmiştir. Verilerin analizinde Rasch model, Dinamik Etkileşim Modeli, Doğrusal Lojistik Test Modeli, Örtük Regresyon Rasch Modeli ve Örtük Regresyon Doğrusal Lojistik Test Modeli (ÖRDLTMTM) ile rastgele ve sabit etkiler kestirilmiş ve model veri uyumu incelenmiştir. Araştırma sonuçları veriye en iyi uyum gösteren modelin süreç verilerinin, madde ve birey değişkenlerinin birlikte ele alındığı ÖRDLTMTM3 olduğunu göstermiştir. Süreç verileri ile okuma maddelerinin doğru yanıtlanma olasılıkları arasındaki manidar ilişkiler olduğu ve okuma becerileri performansının açıklanmasında süreç verilerinin önemli bir katkı sağladığı görülmüştür. Eylem sayısı ve tepki süresindeki artış maddelerin doğru yanıtlanma olasılığını arttırmakta, görüntüleme sayısındaki artış ise azaltmaktadır.

Anahtar Sözcükler: Süreç Verileri, AMTM, Madde Bazında Birey Özellikleri.

ABSTRACT

THE IMPACT OF PROCESS DATA WITH PERSON AND ITEM COVARIATES ON ESTIMATION OF THE PROBABILITY OF CORRECT RESPONSE

KIZILTOPRAK, Fatma

Dissertation, Department of Educational Sciences

Supervisor: Prof. Dr. Celal Deha DOĞAN

February 2025, xii + 90 pages

With the widespread use of computer-based applications, various innovations such as the use of different item types, automatic scoring and adaptive tests have emerged in educational assessment studies. Among these innovations, one of the most studied ones in recent times is the collection of process data and increasing the validity of individuals' test-taking processes and assessments with the help of these data. The aim of this study is to determine the item and person characteristics and process data and the interactions between them that affect the probability of the correct response with Explanatory Item Response Modelling (EIRM). For this purpose, PISA 2015 Turkish data were used. Within the scope of the study, gender, the number of books in the house, item type, text type, context, aspect, and item difficulty level as person and item characteristics and response time, number of actions and number of visits as process data were included. In the analysis of the data, the data were modelled with Rasch Model, Dynamic Interaction Model, Linear Logistic Test Model, Latent Regression Rasch Model and Latent Regression Linear Logistic Test Model (LRLTM), random and fixed effects were estimated, and model-data fit was examined. The results of the study showed that the LRLTM in which process data, item and person variables were considered together fits better. It was found that there are significant relationships between process data and the probability of correct response of reading literacy items. Briefly, the increase in response time and number of actions increased the probability of correct response, while the increase in number of visits decreased the probability of correct response.

Key Words: Process Data, EIRM, person-by-item covariates.

ÖNSÖZ

Bu çalışmayı yapmamda emeği geçen başta değerli danışmanım Prof. Dr. Celal Deha DOĞAN'a süreçteki desteği, ilgisi, anlayışı ve katkılarından dolayı en içten teşekkürlerimi sunarım. Tez izleme kurulunda yer alan ve değerli görüşleri, destekleri, önerileri ve eleştirileri ile çalışmamı şekillendirmemi, daha iyi hale getirmemi sağlayan değerli hocalarım Prof. Dr. Şeref TAN, Prof. Dr. Hamide Deniz GÜLLEROĞLU ve Doç. Dr. Seher YALÇIN'a; tez jürimde yer alan, görüş ve önerileri ile tezime katkı sağlayan değerli hocalarım Prof. Dr. Nuri DOĞAN ve Doç. Dr. Ergül DEMİR'e teşekkür ederim. Doktora eğitimim süresince Ankara Üniversitesi Eğitimde Ölçme ve Değerlendirme Anabilim dalında görev yapan tüm değerli hocalarıma tek tek gerek akademik gerek manevi olarak sağladıkları tüm katkılar ve emeklerinden ötürü teşekkürlerimi sunarım. Doktora eğitimim süresince yanımda olan, beni benden çok düşünen ve her türlü desteği sunan değerli ailem ve bana en büyük gücü veren biricik oğlum başta olmak üzere, çalışmamı tamamlayabilmem için beni motive eden, bana destek ve cesaret veren, umutsuzluğa kapıldığımda yanımda ve yardımcı olan, bitmek bilmeyen sorularıma sabırla cevap veren tüm arkadaşlarıma ve hocalarıma en içten teşekkürlerimi sunarım.

Canım Oğlum Kuzey'e ...



İÇİNDEKİLER

	Sayfa
ETİK İLKELERE UYGUNLUK BİLDİRİMİ.....	iii
ÖZET	iv
ABSTRACT	v
ÖNSÖZ.....	vi
İÇİNDEKİLER.....	viii
TABLolar DİZİNİ.....	x
ŞEKİLLER DİZİNİ	xi
KISALTMALAR	xii
BÖLÜM 1	1
GİRİŞ.....	1
Problem.....	1
Amaç.....	5
Önem.....	6
Sınırlılıklar	7
Sayıtlar.....	7
Tanımlar.....	7
BÖLÜM 2.....	9
KAVRAMSAL ÇERÇEVE VE İLGİLİ ARAŞTIRMALAR.....	9
Bilgisayar Tabanlı Testlerde Yanıtlama Süreci	9
Süreç Verileri	10
Süreç Verilerinin Analizi ve Kullanım Alanları	12
Tepki Süresi	13
Eylem Sayısı	16
İlk Eyleme Kadar Geçen Süre.....	17
Görüntüleme Sayısı	18
Açıklayıcı Madde Tepki Kuramı Modelleri	19
Genelleştirilmiş Doğrusal Karma Modeller	20
Rasch Model	22
Doğrusal Lojistik Test Modeli	23
Örtük Regresyon Rasch Model	24
Örtük Regresyon Doğrusal Lojistik Test Modeli.....	25
Madde Bazında Birey Değişkenleri İçeren Genişletilmiş Rasch Model.....	25
BÖLÜM 3.....	27
YÖNTEM	27
Araştırmanın Modeli.....	27
Çalışma Grubu	27
Verilerin Elde Edilmesi	28
Araştırmada Ele Alınan Değişkenler	29
Verilerin Çözümlemesi	33

Verilerin Analize Hazırlanması.....	34
Varsayımların İncelenmesi	35
Veri Çözümleme Teknikleri.....	37
BÖLÜM 4.....	41
BULGULAR VE YORUMLAR	41
Boş Modele (Rasch Model) Yönelik Bulgular	41
Süreç Verilerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular.....	43
Tepki Süresinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular	45
Eylem Sayısının Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular	46
Görüntüleme Sayısının Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular.....	48
Tepki Süresinin, Eylem Sayısının ve Görüntüleme Sayısının Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular.....	49
Madde Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular.....	52
Birey Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular.....	58
Süreç Verilerinin, Madde ve Birey Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular	61
BÖLÜM 4.....	72
SONUÇLAR VE ÖNERİLER	72
Sonuçlar	72
Süreç Verilerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Sonuçlar.....	72
Madde Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Sonuçlar.....	73
Birey Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Sonuçlar.....	74
Süreç Verilerinin, Madde ve Birey Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Sonuçlar	74
Öneriler	75
Uygulamaya Yönelik Öneriler	75
Araştırmacılara Yönelik Öneriler.....	76
KAYNAKLAR.....	78
EKLER	85
EK 1. Etik Kurul Onayı	86
EK 2. Araştırmada Ele Alınan Değişkenlere İlişkin Betimsel İstatistikler.....	87
ÖZGEÇMİŞ.....	90

TABLÖLAR DİZİNİ

Tablo	Sayfa
Tablo 1. Açıklayıcı Madde Tepki Modelleri	22
Tablo 2. Araştırmada Ele Alınan Değişkenler.....	29
Tablo 3. Araştırmada Ele Alınan Maddeler ve Özellikleri.....	30
Tablo 4. Rasch Model İle Modellenen Maddelerin Infit ve Outfit Değerleri.....	37
Tablo 5. Rasch model ile Kestirilen Madde Güçlük Parametreleri ve Kestirimlerin Standart Hataları.....	42
Tablo 6. Süreç Verileri ile Kurulan Modellerin Sonuçları	44
Tablo 7. Madde Özellikleri ile Kurulan Modellerin Sonuçları.....	53
Tablo 8. Birey Özellikleri ile Kurulan Modellerin Sonuçları.....	58
Tablo 9. Etkileşim Modellerinin Sonuçları	62

ŞEKİLLER DİZİNİ

Şekil	Sayfa
Şekil 1. Süreç Verilerinin Üç Temel Kullanımı (Ercikan Vd., 2023, 212)	11
Şekil 2. GDKM Ait Grafikselsel Gösterim	21
Şekil 3. Madde Bazında Birey Değişkenleri İçeren Genişletilmiş Rasch Modelin Grafikselsel Temsili	26
Şekil 4. Paralel Analiz Yamaç Birikinti Grafiğı	36
Şekil 5. Rasch Model İle Elde Edilen Kestirimlere Dayalı Birey-Madde Haritası	42
Şekil 6. M42 Modelinde Manidar Bulunan Açıklayıcı Değişkenler İçin Marjinal Etki Grafiğı	51
Şekil 7. DLTM İle Manidar Bulunan Açıklayıcı Değişkenler İçin Marjinal Etki Grafiğı	56
Şekil 8. DLMT_E İle Manidar Bulunan Açıklayıcı Değişkenler İçin Marjinal Etki Grafiğı	57
Şekil 9 ÖRRM_E Modelinde Cinsiyet ve Kitap Sayısı İçin Marjinal Etki Grafiğı	60
Şekil 10. ÖRDLTM3 Modelinde Tepki Süresi, Görüntüleme Sayısı ve Eylem Sayısı Değişkenleri İçin Marjinal Etki Grafiğı	65
Şekil 11. ÖRDLTM3 İle Modellenen Madde Zorluk Düzeyi, Evdeki Kitap Sayısı ve Eylem Sayısı Değişkenleri İçin Marjinal Etki Grafiğı	66

KISALTMALAR

PISA	Uluslararası Öğrenci Değerlendirme Programı
TIMSS	Uluslararası Matematik ve Fen Eğilimleri Araştırması
PIAAC	Uluslararası Yetişkin Becerilerinin Ölçülmesi Programı
AMTM	Açıklayıcı Madde Tepki Modellemesi
OECD	Ekonomik İşbirliği ve Kalkınma Örgütü
MTK	Madde Tepki Kuramı
GDKM	Genelleştirilmiş Doğrusal Karma Modeller
DOKM	Doğrusal Olmayan Karma Modeller
DLTM	Doğrusal Lojistik Test Modeli
ÖRRM	Örtük Regresyon Rasch Modeli
ÖRDLTM	Örtük Regresyon Doğrusal Lojistik Test Modeli

BÖLÜM 1

GİRİŞ

Araştırmanın bu bölümünde problem durumuna, amaç ve araştırma sorularına, araştırmanın önemine, sınırlılıklarına ve araştırmada ele alınan tanımlara yer verilmiştir.

Problem

Dijital çağ olarak adlandırılan günümüz dünyasında eğitimde ölçme ve değerlendirme çalışmalarında kayda değer bir değişim dönüşüm gerçekleştiği görülmektedir. Bu değişim ve dönüşümün geleneksel kâğıt-kalem testlerinden bilgisayar tabanlı uygulamalara geçilmesiyle birlikte büyük bir hız kazandığı açıktır.

Bilgisayarların yaygınlaşması ile, test geliştiriciler, geleneksel kâğıt-kalem testlerinden, bilgisayar tabanlı uygulama yöntemlerine doğru geçiş yapmaya başlamışlardır (Bartram ve Bayliss, 1984). Amerika'da birçok eyalet ve okul bölgesi ölçme uygulamalarını bilgisayar destekli uygulamalara geçirmeye çalışmaktadır (Bennett, 2006). Quellmalz ve Pellegrino (2009), 2009 yılına kadar 27'den fazla ülkenin çevrimiçi değerlendirme uygulamalarının pilot çalışmalarını ya da asıl uygulamalarını hayata geçirdiklerini belirtmektedir. Benzer şekilde geniş ölçekli durum belirleme çalışmalarında da geleneksel kâğıt-kalem testlerinden bilgisayar tabanlı testlere doğru büyüyen bir hareket söz konusudur (Wise, 2019). Uluslararası Geniş Ölçekli Durum Belirleme çalışmalarında bilgisayar tabanlı ölçme uygulamaları bir standart olarak kabul edilmeye başlamıştır. Örneğin, küresel düzeyde yürütülen Uluslararası Öğrenci Değerlendirme Programı (Programme for International Student Assessment, [PISA]) 2015 ve Uluslararası Matematik ve Fen Eğilimleri Araştırması (Trends in International Mathematics and Science Study, [TIMSS]) 2019 yılında bilgisayar tabanlı olarak uygulanmaya başlanmıştır.

Bilgisayar tabanlı testlerin uygulanmasında ilk motivasyon ölçmelerin hızlı ve etkili (efficiency) yapılabilmesindeki kazanımlardır (Ramalingam ve Adams, 2018). Ölçmelerin uygulanmasının ve puanlanmasının yanı sıra test geliştirmede ve puanların

raporlaştırılmasında da etkili olmuştur. Bununla birlikte dinamik görseller, ses ve kullanıcı etkileşiminin yanı sıra yanıtlayıcıların her birine uyarlanabilmesi ve neredeyse eş zamanlı puanlanabilmesi sayesinde test olanaklarını geleneksel kâğıt kalem testlerinin sınırlılıklarının ötesine büyük ölçüde genişletmiştir (Scalise ve Gifford, 2006). Bu gelişmelere paralel olarak iş birlikli problem çözme ve yaratıcı düşünme gibi yeni yapıları ölçmenin ve çeşitli yapıları test etmenin yeni yollarını keşfetmeye odaklanılmıştır. Bunun bir sonucu olarak toplanabilen verilerin çeşitliliğinin artırılması yalnızca yeni yapıların keşfedilmesine değil aynı zamanda yerleşik yapıların etkili bir şekilde ölçülebilmesine de katkı sağlamıştır. Böylelikle teknoloji öğrenci performansının ölçülme ve değerlendirilmesinde en büyük yeniliklerden biri olarak kabul edilmiş, son yıllarda eğitimde ölçme ve değerlendirme araştırmalarının önemli bir odağı haline gelmiş (Ramalingam ve Adams, 2018) ve bilgisayar tabanlı ölçmelerden dijital değerlendirme ortamları olarak adlandırılan (Ercikan, Guo, ve Por, 2023; Provasnik, 2021) farklı bir boyuta geçilmiştir.

Karmaşık becerilerin değerlendirilmesine giderek daha fazla odaklanılması, nihai çözümlerin yanı sıra bireylerin yanıt süreçlerinin de incelenmesinin yolunu açmıştır. Yanıtlama sürecine ilişkin veriler, bir öğrencinin bir görevi yalnızca doğru bir şekilde gerçekleştirip gerçekleştirmediğini değil, aynı zamanda nasıl gerçekleştirdiğine ilişkin daha önceden ulaşılamayan iç görüler sağladığı için oldukça umut vaat etmektedir (Ramalingam ve Adams, 2018). Veri yakalama teknolojilerindeki gelişmeler, göz izleme, tıklama akışları ve çeşitli eylemlerle ilişkili zaman damgaları gibi teknikler aracılığıyla yanıtlama süreci verilerinin toplanabilmesine olanak sağlamıştır (Ercikan vd., 2023). Bu bağlamda bilgisayar tabanlı ölçmelere geçişin önemli bir diğer avantajı, testi alan kişinin dijital ortam ile olan her türlü etkileşimini içeren süreç verilerini kaynak olarak kullanabilme olanağıdır (Provasnik, 2021; Teig, Scherer, ve Kjærnsli, 2020).

Süreç verilerinin toplanmasının çeşitli avantajlarından bahsetmek mümkündür. Test ve madde tasarımına ve elde edilen puanların yorumlanabilirliğine katkısı (Embretson, 2023; Ercikan vd., 2023), geçerliğin ve güvenirliğin artırılabilmesini sağlaması (Bergner ve von Davier 2019; Provasnik, 2021), okuma ve matematiksel problem çözme gibi yerleşik yapılarla ilgili olarak incelikli ve potansiyel olarak teşhis edici bir rol oynaması (Bergner ve von Davier 2019) ve ölçme tasarımı, madde geliştirme, puan geçişleme ve adillik çalışmalarını önemli ölçüde geliştirebilecek olması (Ercikan vd., 2023) bunlardan bazılarıdır.

Süreç verileri testi alan kişinin madde ile karşılaştığında maddeyi ne kadar sürede yanıtladığı, yanıtını nasıl yapılandığı, maddeye geri dönüp dönmediği, yanıtını değiştirip değiştirmediği gibi pek çok değişkeni içerebilmektedir. Yanıtlayıcıların eylemleri (üniteyi başlatma, bir tuşa tıklama ya da bir yanıtı onaylamak için harcanan süre ve benzeri) ve tüm davranışları (klavye, fare ve gözleriyle neler yaptıkları) log dosyası (günlük dosyası) olarak adlandırılan yeni bir veri seti olarak kaydedilebilmektedir (Reis Costa ve Chen, 2023).

Embretson (2023) eğitimcilerin ve test geliştiricilerin öğrenci performansını daha kapsamlı bir şekilde değerlendirilmesine olanak tanıdığı için yanıt doğruluğu ve tepki sürelerinin birlikte analiz edilmesini önermektedir. Bergner ve von Davier (2019), süreç verilerinin güvenilirlik ve geçerlilik gibi temel psikometrik konuların değerlendirilmesinde nasıl bir rol oynadığını anlamak için daha fazla çalışmaya ihtiyaç olduğunu dile getirmiştir. Provasnik, (2021), süreç verilerinin şu anda en faydalı görünen kullanımlarından birinin maddelerin yanıtlanma sürecinin ve hangi özelliklerin ya da değişkenlerin maddeleri daha zor ya da daha güvenilir hale getirdiğinin daha iyi anlaşılması olduğunu belirtmektedir.

Anghel, Khorramdel ve von Davier (2024), süreç veri üzerine yaptıkları literatür taramasında en fazla çalışılan değişkenin tepki süresi olduğunu ve bu çalışmaların büyük çoğunluğunun metodolojik bir başka deyişle model geliştirme ağırlıklı olduğunu belirlemiş, Embretson (2023) geliştirilen test modellerinin daha fazla ampirik çalışma ile desteklenmesi gerektiğini önermiştir. Diğer yandan, süreç verileri ile bilişsel beceriler ve yetenek düzeyleri arasındaki ilişkiler test katılımcıları ve madde özellikleri göz önünde tutularak incelenmekte ve yürütülen araştırma sonuçları, veriler arasındaki ilişkilerin kültüre (Lee ve Habermann, 2016; Reis Costa ve Chen, 2023) ve ele alınan değişkenlere, uygulanan testte yer alan maddelerin özelliklerine göre farklılık gösterdiğini ortaya koymaktadır.

Birey özellikleri ile tepki süreleri arasındaki ilişkileri inceleyen araştırmaların bireyin yeteneği, yanıt doğruluğu ve tepki süresi arasındaki ilişkilere yönelik çelişkili sonuçlar elde ettiği göze çarpmaktadır. Bunun yanı sıra maddeyi görüntüleme sayısı üzerine yapılan az sayıda çalışma bulunmakta ve yanıtlama süreciyle ilgili bu değişkenin doğasını açıklamaya yönelik daha fazla çalışmanın yapılmasına ihtiyaç duyulmaktadır.

Alanyazın incelendiğinde süreç verilerinin öğrencilerin yalnızca çıktı değişkenlerinin değil aynı zamanda yanıtlama süreçlerini incelemeyi sağlaması (Embretson, 2023) ve süreç verileri ile performans ve diğer bağlamsal değişkenler

arasındaki ilişkilerin henüz tutarlı olarak açıklanamamış olması (Anghel vd., 2024; Provasnik, 2021), süreç verileri üzerine yapılan çalışmaların büyük çoğunun metodolojik ağırlıklı (Anghel vd., 2024) ya da test alma davranışlarını belirlemeye yönelik olması (Sahin ve Covin, 2020; Wise, 2019) ve çeşitli modeller geliştirilmiş (De Boeck ve Scaliase, 2019; Reis Costa ve Chen, 2023; Rebouças-Ju ve Cheng, 2021; van der Linden, 2006) olmasına karşın, bu modellerin kullanıldığı az sayıda çalışmanın bulunması (Embretson, 2023) ve bu verilerin doğasının ve rolünün anlaşılmasına yönelik daha fazla çalışma yapılmasının önerilmesi (Bergner ve von Davier, 2019) göze çarpmaktadır. Bunlara ek olarak Türkiye örneklemi üzerinde süreç verileri kullanılarak yapılan az sayıda çalışmaya (Yalçın, 2022) rastlanmıştır. Alanyazında Türkiye örnekleme özelinde süreç verilerinin çeşitli madde özelliklerinin, birey özelliklerinin ve yanıt doğruluğunun birlikte ele alındığı çalışmalara rastlanmamıştır.

Türkiye örnekleme için süreç verilerinin toplandığı üç farklı ölçme uygulaması (TIMSS, PISA ve PIAAC) bulunmakla birlikte, analizlerin daha sağlıklı yürütülebilmesi adına en çok maddenin en fazla öğrenci tarafından yanıtlandığı ve görüntüleme sayılarını içeren PISA 2015 uygulaması verilerinden yararlanılmasının uygun olacağı düşünülmüştür. Bu bağlamda Türkiye örnekleme özelinde, temel bir beceri olan okuma becerileri kapsamında, süreç verilerinin okuma becerileriyle ilişkili olduğu önceki araştırma sonuçları ile ortaya konulan birey ve madde özellikleriyle ele alınarak incelenmesinin alanyazına katkı sağlayacağı düşünülmüştür.

Okuma becerileriyle ilişkili olan madde özelliklerinden bazılarının madde türü (Bulut, Bulut ve Arıkan, 2022; Schulz-Heidorf ve Støle, 2018; Yousefpoori-Naeim, Bulut ve Tan, 2023), maddenin gerektirdiği bilişsel boyut (Bulut vd., 2022; OECD, 2017a; Yousefpoori-Naeim vd., 2023), maddenin içerdiği metnin biçimi (Hódi, Németh ve Tóth, 2017; Hódi, Tóth ve Nikolov, 2025), maddenin içerdiği metnin bağlamı (OECD, 2017a) olduğu görülmektedir. Bunun yanı sıra OECD tarafından maddelerin kullanıldığı ilk uygulama verisi üzerinden kestirilen ve birey yeteneği ile aynı ölçek üzerinde ölçeklenen madde güçlük düzeyinin okuma becerilerinin açıklanmasına katkı sağlayacağı düşünülmüştür. Ayrıca PISA 2015 okuma becerileri çerçevesinde bu özelliklere yer verilmiş olması ve uygulanan maddelere ait bu özelliklerin açıklanmış olması nedeniyle çalışmaya edilmesine karar verilmiştir.

Okuma becerileri maddelerinin doğru yanıtlanma olasılığıyla ilişkili olan çok çeşitli birey özelliklerinden bahsetmek mümkündür. Bunlardan bazıları cinsiyet, öğrenciye ait kitap sayısı, evde bulunan kitap sayısı, babanın eğitim durumu, evde internet

kullanımı, evde konuşulan dil, sahip olunan e kitap okuyucu sayısı, evdeki kitap çeşitliliği ve sayısı bunlardan bazılarıdır (Aydoğan ve Gelbal, 2022; Kutlu, Yıldırım, Bilican ve Kumandaş, 2011). Aydoğan ve Gelbal (2022) öğrencilerin okuma becerilerini yordayan en etkili değişkenlerin evdeki kitap sayısı, tepki süresi, evde konuşulan dil ve cinsiyet olduğu sonucuna ulaşmıştır. Bu bağlamda bu özelliklerden PISA 2015 veri setinde yer alan ve öğrenciler arasında daha dengeli bir dağılımı olan evdeki kitap sayısının ve cinsiyetin çalışmaya edilmesine karar verilmiştir. Bu doğrultuda bu araştırmanın problem cümlesi “PISA 2015 okuma becerileri kapsamında tepki süresi, eylem sayısı, görüntüleme sayısı, madde türü, madde bağlamı, bilişsel boyut, metin biçimi ve madde güçlük düzeyi, evdeki kitap sayısı ve cinsiyet değişkenleri öğrencilerin maddeleri doğru yanıtlama olasılığını nasıl etkilemektedir?” olarak belirlenmiştir.

Amaç

Bu araştırmanın amacı, PISA 2015 okuma becerileri kapsamında süreç verilerinin (tepki süresi, eylem sayısı ve görüntüleme sayısı), birey (cinsiyet ve evdeki kitap sayısı) ve madde (madde türü, bağlam, metin biçimi, bilişsel boyut ve madde zorluk düzeyi) özelliklerinin maddelerin doğru yanıtlanma olasılığı üzerindeki etkisini incelemektir. Bu genel amaç doğrultusunda aşağıdaki araştırma sorularına yanıt aranmıştır:

1. Ayrı ayrı ve eş zamanlı olarak ele alındığında süreç verilerinin maddelerin doğru yanıtlanma olasılığı üzerindeki etkisi nasıldır?
 - 1.1. Tepki süresi tek yordayıcı olarak,
 - 1.2. Eylem sayısı tek yordayıcı olarak,
 - 1.3. Görüntüleme sayısı tek yordayıcı olarak,
 - 1.4. Tepki süresi, eylem sayısı ve görüntüleme sayısı birlikte ele alındığında, maddelerin doğru yanıtlanma olasılığı üzerindeki etkisi nasıldır?
2. Madde özelliklerinin (madde türünün, madde bağlamının, maddenin içerdiği metin biçiminin, madde bilişsel boyutunun ve madde zorluk düzeylerinin) maddelerin doğru yanıtlanma olasılığı üzerindeki etkisi nasıldır?
3. Birey değişkenlerinin (cinsiyetin ve evdeki kitap sayısının) maddelerin doğru yanıtlanma olasılığı üzerindeki etkisi nasıldır?
4. Süreç verisi, birey ve madde özelliklerinin etkileşimlerinin maddelerin doğru yanıtlanma olasılığı üzerindeki etkisi nasıldır?
 - 4.1. Birey ve madde özellikleri etkileşimlerinin,

4.2. Süreç verileri, birey ve madde özellikleri etkileşimlerinin maddelerin doğru yanıtlanma olasılığı üzerindeki etkisi nasıldır?

Önem

Bu araştırmada ortaya çıkarılmak istenen etkiler gerek kuramsal alana gerek uygulamaya katkı sağlaması açısından önem arz etmektedir. Süreç verileri olarak adlandırılan ve son yıllarda popülerliği giderek artan verilerin ayıklanması ve incelenmesi, ölçülmek istenilen yapılarla olan ilişkilerinin belirlenmesinin ölçümlerin geçerliğine katkısı tartışmasız kabul edilmektedir. Bu verilerin doğasına ışık tutulması öğrencilerin sadece öğrenme çıktılarını değil aynı zamanda bilişsel süreçlerini de yansıtmaya potansiyeli barındırdığından kuramsal alana bazı katkılar sağlayabilir. Bu anlamda bu çalışmanın okuma becerileri maddeleri bağlamında maddelerin doğru yanıtlanma olasılığını etkileyen süreç verilerinin neler olduğunu ve nasıl bir etkisinin olduğunu, bunların birey ve madde yordayıcıları ile eş zamanlı etkileşimlerini ortaya koyması bakımından gelecekteki çalışmalara ışık tutmaya ve alandaki boşluğu doldurmaya katkı sağlayacağı düşünülmektedir.

Öğrencilerin test alma davranışlarından özellikle ilgisiz test alma davranışının (disengagement) belirlenmesinin ölçümlerin geçerliğine ve testlerin tasarımına katkısı alanyazında tartışmasız kabul edilmektedir. Bu çalışma ile incelenen süreç verilerinden hangisinin maddelerin doğru yanıtlanma olasılığına nasıl bir etkisinin olduğunun ortaya konması ilgisiz test alma davranışlarının belirlenmesinde yeni yaklaşımların ileri sürülmesinde göz önünde tutulabilir. Bu açıdan bu çalışmanın kuramsal alana katkı sağlayabileceği düşünülmektedir. Ayrıca çalışmanın düşük riskli geniş ölçekli durum belirleme çalışmalarından PISA uygulaması bağlamında yürütülmüş olması, spesifik olarak PISA Türkiye sonuçlarının daha farklı bir bakış açısıyla değerlendirilmesini sağlayabilir.

Süreç verileriyle ilgili yapılan çalışmaların genellikle yeni modellerin geliştirilmesine odaklanması ancak çoğu modelin yeterince çalışmadığı (Embretson, 2023) fikrinden yola çıkılarak, bu çalışmanın süreç verilerinin incelenmesinde AMTM'nin bir uygulaması olarak alinyasına katkı sağlayacağı düşünülmektedir. Diğer yandan süreç verileri ile madde özellikleri arasındaki ilişkilerin incelendiği çalışmalarda tutarlı sonuçlar elde edilmiş olmakla birlikte süreç verisi ve birey özellikleri arasındaki ilişkilerin incelendiği çalışmalarda farklı sonuçlara ulaşılmış olması (Anghel vd., 2024)

nedeniyle bu çalışmadan elde edilen sonuçlar bu değişkenlere ilişkin bilgilerin desteklenmesini ya da farklı bir bakış açısı sunulmasını sağlayabilir.

Son olarak süreç verilerinden maddeyi görüntüleme sayısını ya da Türkiye örneklemindeki öğrencilerin süreç verilerini dikkate alan çok az sayıda çalışmanın bulunması, yürütülen araştırmalarda farklı kültürler için farklı sonuçlara ulaşılmış olması ve bu çalışmada alanyazındaki birçok çalışmaya kıyasla daha geniş bir madde kümesi üzerinde çalışılmış olması yönüyle bu çalışmanın sonuçlarının Türk kültüründeki öğrencilerin yanıtlama süreçlerine kapsamlı bir bakış açısı sunmasının önemli olduğu düşünülmektedir.

Sınırlılıklar

1. Bu çalışma PISA 2015 uygulaması okuma becerileri alanından bir madde kümesinin Türkiye örneklemine ait verilerle sınırlıdır.
2. Bu çalışmada hazır veriler kullanıldığı ve bu verilerin ait olduğu maddeler yayınlanmadığından, sunulan maddelerin içerik ve bağlamının kontrol edilememesi çalışmanın sınırlılığıdır.

Sayıtlar

1. Kullanılan R programa dilinin ve paket programlarının hatasız kestirim yaptığı varsayılmıştır.
2. Öğrencilerin, süreç verilerini etkileyebilecek teknik aksaklıklar yaşamadığı varsayılmıştır.

Tanımlar

1. Tepki Süresi (response time): Madde üzerinde harcanan toplam süredir.
2. Eylem Sayısı (number of actions): Tıklama, tuşa basma, çift tıklama, tutma ve sürükleme gibi tüm eylemlerin sayısıdır.
3. Görüntüleme sayısı (total visit): Maddenin kaç kez görüntülediğini, öğrencilerin maddelere geri dönme sıklığının bir ölçüsüdür.
4. İlgisiz test alma davranışı (disengagement): Yanıtlayıcıların bir ya da daha fazla test maddesini dikkatlice okumadan, anlamaya çalışmadan ya da yeterli

aba gstermeden yanıtlamasıdır. Bu tr davranıřlar genellikle dřk motivasyon, ilgisizlik, yorgunluk veya testin sonucuna nem vermeme gibi nedenlerle ortaya ıkmaktadır (Wise ve DeMars, 2005).



BÖLÜM 2

KAVRAMSAL ÇERÇEVE VE İLGİLİ ARAŞTIRMALAR

Araştırmanın bu bölümünde süreç verileri ve Açıklayıcı Madde Tepki Modelleri ile ilgili ayrıntılı bilgilere yer verilmiştir. Bu bölümde ayrıca ilgili alan yazın ve kavramsal çerçeve sunulmuştur.

Bilgisayar Tabanlı Testlerde Yanıtlama Süreci

Bilgisayar tabanlı değerlendirme yapıldığında, testi alan kişinin dijital ortamdaki her türlü etkileşimi süreç verileri olarak kaydedilebilmektedir. Süreç verileri testi alan kişinin madde ile karşılaştığında maddeyi ne kadar sürede yanıtladığı, yanıtını nasıl yapılandığı, maddeye geri dönüp dönmediği, yanıtını değiştirip değiştirmediği gibi pek çok değişkeni içerebilmektedir. Süreç verisinin kullanımı, bir öğrencinin yalnızca bir görevi çözüp çözemediğini değil aynı zamanda nasıl çözdüğüne ilişkin daha önceden ulaşılamayan iç görüşler sağladığı için oldukça umut vaat etmektedir (Ramalingam ve Adams, 2018). Dahası, öğretme ve öğrenme bağlamında doğrudan hedeflenmeye uygun olması olası yirmi birinci yüzyıl becerilerinin değerlendirilmesine katkı sağlayan bu süreç verileridir.

Bilgisayar tabanlı ölçme uygulamalarında süreç verisinin toplanması sayesinde geçerlik ve güvenirlik arttırılabilmektedir. Bilgisayar tabanlı değerlendirmelere geçişin önemli bir avantajı, süreç verilerini kaynak olarak kullanabilme olanağıdır (Provasnik, 2021; Teig vd., 2020). Bu verilerin test ve madde tasarımına ve elde edilen puanların yorumlanabilirliğine katkısı olduğu gösterilmiştir (Embretson, 2023; Ercikan vd., 2023). Diğer yandan süreç verilerinin okuma ve matematiksel problem çözme gibi yerleşik yapılarla ilgili olarak incelikli ve potansiyel olarak teşhis edici bir rol oynayabileceği belirtilmiştir (Bergner ve von Davier 2019). Multimedya madde formatlarının öğrencilerin katılım düzeyini artırabildiği ancak hedeflenmemiş bilişsel gereklilikler (yapıyla ilişkili olmayan varyans) içerebilmesi nedeniyle geçerlilik sorunlarına yol açabildiği bilinmektedir (Ercikan vd., 2023). Bu potansiyel dezavantaja rağmen dijital değerlendirme ortamları, ölçme tasarımı, madde geliştirme, puan geçişleme ve adillik

alışmalarını önemli ölçüde geliştirebilecek süreç verilerinin toplanması için fırsatlar sunmaktadır.

Süreç Verileri

Süreç verileri bir test maddesinin yanıtlanma sürecini yansıtan bilişsel ve bilişsel olmayan özellikle de psikolojik yapıları yansıtan ampirik veriler şeklinde tanımlanmıştır (Provasnik 2021; Reis Costa ve Leoncio Netto, 2022). Bilgisayar tabanlı testlerde hem yanıtlayıcıların eylemleri (üniteyi başlatma bir tuşa tıklama girdiği için ya da bir yanıtı onaylamak için harcanan süre ve benzeri) hem de yanıtlayıcıların tüm davranışları (klavye fare ve gözleriyle neler yaptıkları) süreç verisi olarak adlandırılan yeni bir veri seti olarak kaydedilebilmektedir (Resi Costa ve Chen, 2023).

Süreç verileri bilgisayar tarafından otomatik olarak kaydedilen log dosyaları kullanılarak elde edilmektedir. Bilgisayar tabanlı ölçme uygulamaları ile yanıtlayıcıların çözüm süreçlerine ait tüm veriler günlük dosyaları olarak adlandırılan log dosyalarına kaydedilmektedir. Log dosyaları, öğrencilerin bir ölçme uygulaması sırasında gerçekleştirdikleri tüm eylemlerin kayıtlarını ve bunlara karşılık gelen zaman damgalarını içeren dosyalardır (Teig vd., 2020). Provasnik (2021) log dosyalarını bilgisayar tabanlı ölçmelerde girdilerin sırası ve hızından (örneğin, tıklamalar ve tuş vuruşları) ölçmeyi yapmak için kullanılan cihazın VPN'sine kadar yakalanan her şey olarak tanımlamaktadır. Bu günlük dosyalarında yer alan olaylar ve tarihleri kullanarak çeşitli süreç verisi değişkenleri elde edilebilmektedir. Diğer bir deyişle, bir bireyin test alma sürecindeki tüm eylemleri tarih ve saatleri ile bir log dosyasına kaydedilmektedir. Örneğin, bir log dosyasından bir bireyin bir maddeyi ilk görüntülediği an ile maddeyi yanıtladığı ana kadar geçen süre hesaplanarak maddeyi yanıtlama süresi değişkeni, ya da bireyin maddeyi yeniden ekranda görüntülediği olayların sayısı görüntüleme sayısı değişkeni olarak elde edilebilmektedir. Dolayısıyla, süreç verileri bireylere ait log dosyalarından çıkarılmaktadır. Diğer yandan süreç verilerinin günlük dosyalarından gelenlerle sınırlı olması gerekmez; örneğin, göz izleme bilgilerini veya manyetik rezonans görüntüleme ve bilgisayarlı tomografi taramaları gibi genellikle dijital temelli durum belirleme çalışmalarının bir parçası olarak henüz yakalanmayan beyin görüntülemelerini ya da dışarıdan gelen bilgileri (örneğin, görüşmecilerin ya da test personelinin belirli bir maddeye katılım derecesine ilişkin gözlemleri) içerebilir (Provasnik, 2021).

Açıkça görülebileceği gibi log dosyaları kullanılarak çok çeşitli süreç verileri elde edilebilir. Bilgisayar tarafından oluşturulan log dosyaları yanıtlama sürecine ilişkin herhangi bir bilgi parçasını (örneğin, tepki süresi ya da eylem sayısı) çıkarmak ve yanıt doğruluğunun ötesinde öğrencilerin yanıtlama süreçleriyle ilgili bilgileri toplamak için kullanılabilir. Örneğin, öğrencilerin her bir maddeyi yanıtlamak için harcadığı süre ve mevcut araçlarla etkileşimlerine (örneğin, çevrimiçi bir hesap makinesi kullanımı, sürükleme, vs.) ilişkin bilgiler çıkarılabilir. Bununla birlikte, zaman ve depolama alanı gibi maliyetlerinden dolayı bir log dosyasından çıkarılabilecek tüm bilgiler süreç verisi değişkenleri olarak yayınlanmamaktadır. Örneğin, PISA 2015 uygulamasında açıklanan süreç verileri tepki süreleri, eylem sayıları ve görüntüleme sayısı değişkenlerinden oluşmaktadır.

Süreç verileri yanıtlayıcıların test alma sürecinde nasıl bir yol izlediğini ve hangi stratejileri kullandığını anlamamıza yardımcı olur (Lee ve Haberman, 2016). Bergner ve von Davier (2019), süreç verilerinin birçoğunun tanısal nitelikte olduğunu, sadece yanıtların doğruluğunu tespit etmenin ötesinde öğrencilerin okuma, yazma veya problem çözme ile ilgili yaşadığı güçlüklerin doğasını açıklamaya yönelik olduğunu ifade etmiştir. Bu veriler içinde yanıtlayıcıların bilişsel süreçlerine ışık tutacak çok sayıda faydalı iç görü bulunabilir. Test puanlarının yorumlanmasında geçerliği artırma potansiyelinin olması bu tür verilerin kullanımını desteklemektedir. Eğitimde ve psikolojide ölçme standartlarının son versiyonunda (2014) geçerlik kaynaklarından biri de yanıt süreçleri ile elde edilen kanıtlardır (Embretson, 2023; Zumbo, Maddox ve Care, 2023). Ercikan ve diğerleri (2023) süreç verilerini kullanarak yanıtlama süreci göstergeleri oluşturmaya yönelik farklı yaklaşımları tartışan bir dizi konudan bahsetmektedir. Bu bağlamda süreç verilerinin Şekil 1’de görüldüğü gibi üç temel kullanımı olduğu belirtilmiştir.

Ölçme tasarımının, kalitesinin ve geçerliğinin geliştirilmesi	• Ölçme Tasarımı • Puanların Anlamının Geçerlenmesi
Yapıların Kanıtı Olarak	• Puanların Üretilmesine Katkı Sağlama • Formatif Amaçlar İçin Performansın Tanımlanması
Gruplar Arası Farkların ve Adilliğin İncelenmesi	• Farklılaşan Katılımın İncelenmesi • Adilliğin Değerlendirilmesi

Şekil 1. Süreç Verilerinin Üç Temel Kullanımı (Ercikan vd., 2023, 212)

Süreç verilerinin farklı kullanımları üç farklı amaç altında birleştirilmiştir. Bu amaçlardan ilki, ölçme değerlendirme tasarımı, kalitesini ve geçerliliğini iyileştirmektir. İkinci amaç, hedeflenen yapıyla ilgili kanıtlar sağlamaktır. Üçüncü amaç ise, farklı insan gruplarının değerlendirmeye katılımını incelemek, puanları karşılaştırmak ve sonuçları etkileyebilecek yapıyla ilgisi olmayan faktörleri belirlemektir.

Süreç Verilerinin Analizi ve Kullanım Alanları

Katılımcıların test maddelerini nasıl çözdüklerini gözlemlemenin etkili ve zorlayıcı olmayan bir yolunu sunması, süreç verilerinin analizine olan ilginin giderek artmasına neden olmuştur. Süreç verilerinin analizi, tıklama, zamanlama ve diğer eylem verilerinin bilişsel düşünme süreçleri bağlamında yorumlanmasını kapsamaktadır. Embretson'a (2023) göre, yanıt doğruluğu ile yanıtlama sürelerinin bir arada analiz edilmesi, eğitimcilerin ve test geliştiricilerin öğrenci performansını daha kapsamlı bir şekilde değerlendirebilmesine olanak tanımaktadır. Bu tür bir değerlendirme, yalnızca öğrencilerin ne bildiklerini değil, aynı zamanda bu bilgiyi nasıl ve ne kadar verimli kullandıklarını da ortaya koymaktadır.

Provasnik (2021), süreç verilerinin analizinin özellikle madde geliştirme ve test tasarımında tanısallık ya da adil uygulamaları iyileştirmek amacıyla faydalı olduğunu belirtmektedir. Bu bağlamda, süreç verileri ve log dosyalarının veri kalitesini artırmak için kullanılabileceği vurgulanmıştır. Süreç verilerinin kullanımına yönelik çeşitli uygulamalar şunlardır:

- Maddelerin işleyişinin daha iyi anlaşılması ve belirli özelliklerin veya değişkenlerin maddeleri daha zor ya da güvenilir hale getirmedeki etkilerinin incelenmesi,
- "Ulaşılamayan" ya da "uygulanmayan" (hiç görülmeyen), "atlanması gereken" (görülen, üzerinde zaman harcanan ancak sonuçta atlanan) ve "denenmeyen" (görülen ancak atlanmadan önce üzerinde zaman harcanmayan) eksik cevap türleri arasında ayırım yapılması,
- Olası kopya çekme durumlarına veya programlama hatalarına işaret edebilecek öğrenci tahminlerinin ve aykırı durumların belirlenmesi.

Anghel ve diğerleri (2024), geniş ölçekli testlerde süreç verilerini ele alan deneysel çalışmaları literatür taraması yöntemiyle incelemiş ve bu çalışmaları konu alanlarına göre sınıflandırmıştır. Çalışmaların büyük bir çoğunluğunun, tepki süresi

değişkenleri, test alma davranışlarının belirlenmesi, karmaşık problem çözme ve dijital yazma gibi konular üzerinde yoğunlaştığı görülmüştür. Ayrıca, bu çalışmaların önemli bir kısmı metodolojik yaklaşımlara odaklanmış ve bireylerin yetenek kestiriminde kullanılabilecek çeşitli yöntem ve teknikler geliştirilmiştir.

Süreç verilerinin analizinde kullanılan modeller, bireylerin test sırasında sergiledikleri yanıt örüntülerindeki farklılıkları yansıtmayı amaçlamaktadır. Bu bağlamda, farklı maddelerin işlenmesine bağlı olarak birden fazla özelliği ölçmek için çok bileşenli modeller (Embretson ve Yang, 2013), test sırasındaki değişiklikleri yansıtmak için dinamik etkileşim modelleri (Meulders ve Xie, 2004), madde güçlüğüne yönelik bireysel farklılıkları değerlendirmek için rastgele ağırlıklar LLTM (Rijmen ve De Boeck, 2002) ve madde yanıt stratejilerindeki bireysel farklılıkları belirlemek amacıyla karma dağılım modelleri (Rost ve von Davier, 1995) sıklıkla kullanılmaktadır (Embretson, 2023).

Buna ek olarak, süreç verilerinin analizinde örtük değişken modellerinin (De Boeck ve Scaliase, 2019; Reis Costa ve Chen, 2023) de kullanıldığı belirtilmiştir. Örneğin, Pokropek ve Khorramdel (2024), hızlı tahmin davranışlarını belirlemek amacıyla Tahmin Davranışı İçin AMTM adlı bir yöntem geliştirmişlerdir. Süreç verisi değişkenleri arasında tepki süresi ve eylem sayısı, alanyazında en sık analiz edilen ve modellenen unsurlar olarak öne çıkmaktadır.

Tepki Süresi

Alanyazında en sık çalışılan süreç verisi tepki süresidir. Tepki süresi, test alma süresinde her bir maddenin ne kadar sürede yanıtlandığını göstermektedir (Lee ve Habermann, 2016). Tepki süresinin öğrencilerin test alma davranışlarını belirlemede, özellikle ilgisiz test alma davranışı kapsamında sıklıkla ele alındığı görülmektedir. Tepki sürelerinin analizi motivasyonun filtrelenmesinde başvurulan yöntemlerden biridir (Bergner ve von Davier, 2019). Özellikle düşük riskli ölçmeler için bir geçerlik tehdidi olarak görülen ilgisiz test alma davranışlarından hızlı tahminin belirlenmesinde ve eşik değerlerin belirlenmesinde tepki süresinden yararlanılmaktadır (örn; Wise, 2019). Bu çalışmalarda ilgisiz test alma davranışları genellikle farklı tepki süresi eşikleri belirlenerek karşılaştırılmıştır. Belirlenen eşik değerlerin altında kalan süreler hızlı tahmin olarak adlandırılarak ilgisiz test alma davranışı saptanmaya çalışılmıştır. Hızlı tahminlere ve boş bırakılan maddelere dayalı istatistiksel modellerin geliştirilmesi

(Ulitzsch, von Davier ve Pohl, 2020), yetenek kestiriminde ilgisiz test alma davranışlarının istatistiksel etkilerinin birey ya da okul düzeyinde incelenmesi (Wise vd., 2020) de ilgisiz test alma davranışının saptanması kapsamında ele alınan konulardan bazılarıdır (Anghel vd., 2024).

Örneğin, Goldhammer, Martens ve Lüdtke (2017), ilgisiz test alma davranışını tepki süresine bağlı olarak saptayarak, düşük riskli testlerde, ilgisiz test alma davranışını koşullandıran faktörleri incelemiştir. AMTM kullanılarak birey ve madde değişkenlerinin etkilerinin incelendiği araştırma sonuçları ilgisiz test alma davranışının alana göre örtük bir boyutu olduğunu, düşük eğitim düzeyi, düşük bilişsel beceriler ve anadil ile test dilinin farklı olması durumunda arttığını göstermiştir. Cinsiyete göre belirgin bir farklılık bulunmamakla birlikte yaşın, madde gücü ve madde konumunun ilgisiz test alma davranışını etkilediği görülmüştür. Güçlük düzeyi yüksek maddeler ve maddenin testin ilerleyen bölümlerinde yer alması ilgisiz test alma davranışının ortaya çıkma olasılığını arttırmaktadır.

Embretson (2019), tepki sürelerinin geçerlik yönünü test geliştirme süreci bağlamında ele alarak madde tasarımı ve tepki sürelerini bir zekâ testi üzerinde iki farklı AMTM ile incelemiştir. Araştırma sonuçları, madde performansının bilişsel karmaşıklık değişkenlerinden güçlü bir şekilde tahmin edilebildiğini göstermiştir.

Anghel ve diğerleri (2024), süreç veriyle ilgili yaptıkları literatür taramasında tepki süresinin ele alındığı çalışmaların büyük çoğunluğu metodolojik bir başka deyişle model geliştirme ağırlıklı olup, bu çalışmalarda kuramsal boyuta fazla değinilmediğini belirlemiştir. Bu kapsamda tepki süresinin ve yanıtın doğruluğunun modellenmesi için geliştirilen Tepki Süresi Modellerine rastlanmaktadır (Örneğin, van der Linden, 2006; Rebouças-Ju ve Cheng, 2021). Metodolojik çalışmalara örnek olarak Klein Entink, Khun, Hornke, ve Fox, (2009) belli bir kurala dayalı olarak tasarlanmış bilişsel testlerin yanıt doğruluğu ve tepki sürelerinin eşzamanlı analizine olanak tanıyan yeni bir model geliştirmiştir. Bu modelde madde tasarım matrislerini madde tarafına entegre ederek, madde tasarım parametrelerinin (örneğin, bilişsel süreçler) hem madde gücü hem de tepki süresi üzerindeki etkilerini değerlendirmek mümkün hale gelmiş ve madde yapısına ilişkin daha derin bilgiler sunmuştur. Molenaar, Terlinckx ve van der Mass (2015), yanıt doğruluğu ve tepki süresinin ortak etkisini kullanarak birey özellik farklılıklarını modellemek için genelleştirilmiş bir doğrusal çerçeve ileri sürmüştür. İki değişkenli genelleştirilmiş doğrusal madde tepki kuramı (B-GLIRT) olarak adlandırılan bu çerçevede, yanıtlar ve yanıt süreleri için ayrı genelleştirilmiş doğrusal ölçme modelleri

belirlenir ve bunlar daha sonra çaprazlanmış ilişkilerle birbirine bağlanır. Molenaar ve De Boeck (2018) yanıtlayıcıların farklı maddeleri farklı yanıtlama süreçleri kullanarak çözdüklerinden ve bunun bireyler için heterojenlik oluşturduğundan yola çıkarak bu heterojenliğin dikkate alındığı bir karma model geliştirmiştir. Bir diğer çalışmada Klotzke ve Fox (2019) log dosyalarından elde edilen yuvalanmış ve çaprazlanmış değişkenler için Bayesçi çerçevede, kovaryans yapısı modellemesi önermiştir.

Tepki süresi ile madde özellikleri arasındaki ilişkileri inceleyen çalışmalar, madde türü, madde güçlüğü, madde ayıricılığı, madde uzunluğu, madde konumu ve maddenin bilişsel karmaşıklığı gibi madde özellikleriyle uzun tepki süresi arasında bağlantı olduğunu göstermiştir (Arenada vd., 2022; Chang, Plake ve Ferdous, 2005; Yalçın, 2022).

Diğer yandan birey özellikleri ile tepki süreleri arasındaki ilişkileri inceleyen araştırmaların sonuçlarının daha az tutarlı olduğu dikkati çekmektedir. Bireyin yeteneği, yanıt doğruluğu ve tepki süresi arasındaki ilişkilere yönelik çelişkili sonuçlar göze çarpmaktadır. Uzun tepki süresinin yetenek düzeyine bağlı olduğu sonucuna ulaşan araştırmalar (örneğin, Chang vd., 2005; Goldhammer, Naumann ve Greiff, 2015; Yalçın, 2022) olduğu gibi aralarında bir ilişki bulunamadığını ya da düşük bir ilişki bulunduğunu ortaya koyan araştırma sonuçları (Arenada, 2022) mevcuttur. Chang ve diğerleri (2005), farklı yetenek düzeylerindeki bireylerin maddeleri yanıtlarken harcadıkları tepki sürelerini karşılaştırdıklarında, yüksek yetenekli öğrencilerin düşük yeteneklilere göre yanıt doğruluğundan ve madde güçlüklerinden bağımsız olarak maddeler üzerinde daha fazla süre harcadıklarını ortaya çıkarmıştır. Bu sonuç farklı cinsiyet ve kültürler arasında tutarlıdır. Bunun yanı sıra öğrencilerin özgüven, kaygı gibi çeşitli psikolojik özellikleri de harcadıkları tepki sürelerinin azalmasına ya da artmasına neden olabilir. Örneğin, kaygı düzeyi yüksek bir öğrenci bir test maddesini yanıtladıktan sonra verdiği yanıtın doğruluğunu kontrol etmek için daha fazla süre harcayabilir.

Lee ve Habermann (2016) ise tüm test için ayrılan süreyi kullanırken farklı kültürlerden gelen katılımcıların oldukça farklı stratejiler izlediğini ve farklı stratejilerin, demografik özelliklerinden bağımsız olarak yüksek performans gösteren katılımcılar üzerinde genellikle çok az etkisi olduğunu belirtmişlerdir. Bu tutarsızlığın ikili işleme kuramında belirtilen, her görevin kendine özgü bir özelliği olan hızlı işleme ya da yavaş işleme gerektirmesi gibi bir özellikten kaynaklanabileceği düşünülmektedir (Anghel vd., 2024). Bu durumun bir diğer açıklaması ise, farklı kültürler için bireyin özellikleri ile süreç verileri arasında farklı ilişkiler bulunmasından kaynaklanabilir. Örneğin, Reis Costa ve Chen (2023) Danimarka, İsveç ve Norveç gibi benzer kültürel özelliklere sahip

lkelerdeki ğrencilerin performansları ve yanıt sreleri ile baėlamsal deėiřkenler arasındaki iliřkiyi inceledikleri alıřmalarında, her  lke iin ortak olarak yalnızca sosyoekonomik dzeyin manidar bir iliřkisi olduėu sonucuna ulařmıřlardır. Farklı kltrler ve kořullarda yařayan, hatta farklı okul trlerinde ėrenim gren ğrencilerin yanıtlama srelerinin ve bu yanıtlama sreleriyle iliřkili bireysel zelliklerin farklılık gstermesi beklenebilecek bir durumdur. Bu durumun nedenlerinin daha fazla alıřma ile aıklanması gereklidir. Bu nedenle alanyazın elde edilen sonuların farklı kltrler iin de benzerlik gsterip gstermediėinin incelenmesini ve lme deėiřmezliėi alıřmalarını nermektedir (rneėin De Boeck ve Scalise, 2019; Reis Costa ve Chen, 2023).

Eylem Sayısı

Alanyazında eylem sayısının hızlı tahmin davranıřının belirlenmesinde bir kořul olarak ele alındıėı (Sahin ve Covin, 2020), tepki sresi ve yanıt doėruluėu ile birlikte ele alınarak rtk zelliklerin incelendiėi (De Boeck ve Scalise, 2019; Reis Costa ve Chen, 2023), baėlamsal deėiřkenlerle iliřkisinin incelendiėi (Reis Costa ve Chen, 2023) ve tahmin davranıřının modellenmesi iin geliřtirilen bir modele diėer sre verileriyle birlikte entegre edilebildiėi (Pokropek ve Khorramdel, 2024) eřitli alıřmalar bulunmaktadır.

řahin ve Covin (2020) tepki sresi ve yanıtlama davranıřlarının (eylem sayısı ve trleri) ilgisiz test alma davranıřının belirlenmesine katkısını incelediklerinde, yanıtlama davranıřlarının eřik deėerin belirlenmesinde kullanımının anlamlı girdiler saėlayabileceėini ve bazı durumlarda ilgisiz test alma davranıřını belirlemede yalnızca yanıt srelerini kullanmaktan daha iyi performans gsterdiėini ortaya koymuřtur. İlgisiz test alma davranıřı gsteren ğrencilerin eylem sayılarının daha az ve eylem trlerinin ok eřitli olduėu grlmřtr. Yavuz (2019), tepki sresi, eylem sayısı ve yanıtlayıcıların performansları arasındaki iliřkileri ve okuma becerileri arasındaki etkileřimleri incelediėi alıřmada, daha ok sre harcayan ve daha fazla tepki davranıřı gsteren ğrencilerin maddeleri doėru yanıtlama olasılıėının daha yksek olduėunu ve bu etkinin maddeler ve ğrenciler arasında deėiřmediėini gstermiřtir. Arenada vd. (2022) ise nispeten dřk ve yksek yeterliliėe sahip ğrencilerin orta yeterliliėe sahip ğrencilere gre daha az sayıda eylem sergilediklerini belirlemiřtir.

Lee ve Haberman (2016) bireylerin maddeleri yanıtlarken sergiledikleri eylemleri istendik dzenle gerekleřtirip gerekleřtirmediklerini, tepki sresi, demografik

değişkenler ve test uzunluğu arasındaki ilişkiler üzerinden incelemişlerdir. Bir diğer çalışmada Qiao, Jiao ve He (2023) yanıt doğruluğu, tepki süresi ve eylem zincirleri değişkenlerini cinsiyete dayalı grup değişmezliğine odaklanarak çok gruplu ortak model kullanarak modellemişlerdir. Bunların yanı sıra eylem sayısı, eylem zincirleri, tepki süresi, yanıt doğruluğu ve arka plan değişkenleri gibi daha karmaşık değişkenlerin de dahil edilebildiği ortak modeller de kullanılabilmektedir (Ulitzsch vd., 2021). Ayrıca, senaryo tabanlı etkileşimli görevlerdeki eylemlerin sıra analizi, rastgele ya da görev dışı eylem dizileriyle meşgul görünen öğrencileri filtreleyerek puan kalibrasyonunu iyileştirmek için kullanılabilen bir başka yaklaşımdır (Bergner ve von Davier 2019).

Teig ve diğerleri (2020), öğrencilerin tepki süresi, eylem sayısı ve yanıtın doğruluğu, fen performansı ve bağlamsal değişkenleri analiz ederek, stratejik, ortaya çıkan ve ilgisiz olmak üzere üç farklı öğrenci profili belirlemiştir. Öğrencilerin bu profillere atanması cinsiyet, sosyoekonomik düzey, evde konuşulan dil gibi demografik özelliklerine, fene yönelik tutumlarına ve fen başarılarına bağlı olarak değişiklik göstermiştir. Reis Costa ve Chen (2023), tepki süresi ve eylem sayısı ile bağlamsal değişkenler arasındaki ilişkilerin kültüre göre farklılık gösterdiğini ortaya koymuştur.

İlk Eyleme Kadar Geçen Süre

İlk eyleme kadar geçen süre, madde güçlüğü, madde formatı ve öğrencilerin yetenek düzeyleri gibi çeşitli özelliklerle ilişkilendirilmektedir. Araneda vd. (2022) ilk eyleme kadar geçen süre ile öğrenci yeteneği arasında pozitif ilişki olduğu sonucuna ulaşmıştır. Albert ve Steinberg (2011) bu sürenin uzun olmasının maddenin çözümü için planlama yapmayı yansıttığını belirtmiştir. Ancak yine de ilk eyleme kadar geçen süre sadece planlama aşamasını değil, problemi anlama aşamalarını da yansıtabileceği göz önünde tutulmalıdır. Arenada vd. (2022) farklı madde türleri bağlamında ilk eyleme kadar geçen süre ile açıklanabilecek bilişsel süreçlerin incelenmesini önermektedir. Bu sürenin madde formatına bağlı olarak farklı bilişsel süreçleri yakalayabileceğini ve öğrenciler tarafından kullanılan bilişsel süreçlerin sınıf seviyeleri arasında farklılık gösterebileceğini belirtmiştir.

Özellikle teknolojiyle zenginleştirilmiş interaktif maddelerde olduğu gibi kritik bilgilerin keşfedilmesi için tıklama, sürükleme, medyayı oynatma gibi çeşitli eylemlerin gerçekleştirilmesi bu sürenin daha kısa olmasına neden olabilir. Bazı madde türleri için ilk eyleme kadar geçen süre çözüm süreçleri için harcanan zamanı yansıtabilecekken,

diğer madde formatlarında planlama ve çözüm süreçleri için harcanan zaman ilk eyleme kadar geçen süresinin çok ötesine uzanabilmektedir. Örneğin, yanıt seçeneklerinin başlangıçta görünmediği bir madde türünde ilk eyleme kadar geçen süre nispeten daha kısa olabilir çünkü öğrencilerin ilk eylemi yanıt seçeneklerini görüntülemek için bir menüye tıklamak olacaktır. Diğer yandan maddenin tamamen ekranda görüldüğü bir matematik maddesinde öğrencinin maddeyi anlama ve çözme sürecini planlaması daha uzun sürecektir. Bu nedenle ilk eyleme kadar geçen süre maddenin tasarımıyla ilgili olabileceğinden bu değişkenle ilgili sonuçların dikkatli yorumlanması gereklidir.

Görüntüleme Sayısı

Süreç verileri arasında en az incelenen değişkenin görüntüleme sayısı olduğundan görüntüleme sayısı ile ilgili olarak sınırlı sayıda çalışmaya ulaşabilmektedir. Bunun nedenlerinden biri, süreç verilerinin genellikle geniş ölçekli düşük riskli ölçme uygulamalarıyla toplanmış olması olabilir. Bu uygulamalardan elde edilen görüntüleme sayılarına ilişkin verilerin öğrenciler arasındaki varyansı çok düşüktür. Öğrencilerin çok büyük bir çoğunluğunun maddeyi yalnızca bir kez görüntülediği ve birden çok görüntüleyen öğrencilerin az sayıda olduğu ortaya çıkmaktadır. PISA 2018 teknik raporunun K ekinde, PISA 2018 uygulamasında uygulanan tüm maddelerin görüntüleme sayısına ilişkin yapılan analizler, maddelerin tüm öğrencilerin %93'ü tarafından yalnızca bir kez görüntülendiğini ortaya çıkarmıştır (OECD, 2024, s.5). Diğer bir deyişle öğrencilerin yalnızca %7'si maddeyi birden çok kez görüntülemiştir. Dahası, görüntüleme sayılarına logaritmik dönüşüm uygulandıktan sonra bile, artan görüntüleme sayısı ile logaritmik frekanslar arasında üstel bir düşüş göze çarpmaktadır.

Maddeyi yeniden görüntüleme davranışları ile ilgili öğrenci profillerinin belirlenmesine (Guo, Johnson, Saldivia, Worthington ve Ercikan, 2024; Zhang, Li ve Wang, 2023), görüntüleme sayıları ile maddelerin güçlük ve ayırt edicilik parametreleri arasındaki ilişkilerin (Lahza, Smith ve Khosravi, 2022), ve madde türü, öğrenci performansı ve madde güçlüğü arasındaki ilişkilerin (Zhang, vd., 2023) incelenmesine yönelik çalışmalara ulaşılmıştır.

Guo ve diğerleri (2024) NAEP matematik maddeleri üzerinden test puanı, tepki süreleri ve görüntüleme sayılarını temel alarak öğrenci profilleri oluşturmuştur. Görüntüleme sayısı fazla olan öğrencilerin, tepki sürelerinin düzensiz ve / veya kısa, test puanlarının düşük ya da çok düşük olduğu görülmüştür. Yüksek görüntüleme sayıları,

düzensiz veya stratejik olmayan test yaklaşımlarına işaret edebilir. Örneğin, düşük performans gösteren öğrenciler arasında, soruları defalarca görüntüleyen ancak etkili yanıt veremeyen profiller tespit edilmiştir. Yüksek başarı gösteren öğrenciler genellikle daha düzenli görüntüleme davranışları sergilemiş, soruları sıralı bir şekilde çözmüş ve toplamda daha düşük bir görüntüleme sayısına sahip olmuştur.

Lahza ve diğerleri (2022), öğrencilerin maddeleri yanıtlarken gerçekleştirdikleri görüntüleme sayıları ile maddelerin güçlük ve ayırt edicilik parametreleri arasındaki ilişkileri incelemişlerdir. Görüntüleme sayıları nihai yanıt kadar görüntüleme sayısı ve toplam görüntüleme sayısı olmak üzere iki farklı şekilde incelenmiştir. Öğrencilerin nihai yanıt kadar bir maddeyi kaç kez görüntülediğinin, bir maddenin zorluğu ve ayırt ediciliğinin güçlü bir göstergesi olduğunu ileri sürmüşlerdir. Toplam görüntüleme sayısı ise madde güçlüğü ile yüksek korelasyon göstermiş olup, düşük performanslı öğrencilerin maddeleri daha fazla görüntüledikleri saptanmıştır. Maddelerin daha fazla görüntülenmesinin öğrencilerin verdikleri yanıtlara güvenmemelerinden kaynaklanabileceği ileri sürülmüştür.

Zhang ve diğerleri (2023), bilgisayar tabanlı testlerde öğrencilerin verdikleri yanıtları geri dönüp değiştirme ve maddeleri görüntüleme sayılarının altında yatan gizil yapıları incelediklerinde bu değişkenlerin test koşullarına ve madde türüne göre değişiklik gösterdiğini ortaya koymuştur. Çoktan seçmeli maddeler doğru yanlış türündeki maddelere göre daha fazla görüntülenmiş ve yanıt üzerinde daha fazla değişiklik yapılmıştır. Zor sorular, öğrenciler tarafından daha az gözden geçirilmiş ve daha az değişiklik yapılmıştır. Cevaplarını sık sık değiştiren öğrencilerin ortalama başarı oranlarının daha yüksek olduğu gözlemlenmiştir. Araştırmanın sonuçlarına göre, yanıt değiştirme eğilimleri, öğrencinin genel performansı ile ilişkili olup görüntüleme sayısı öğrencilerin puanını doğrudan etkilemektedir.

Açıklayıcı Madde Tepki Kuramı Modelleri

Geleneksel Madde Tepki Kuramı (MTK) modelleri, bireylerin örtük özellik düzeyleri ve seçilen modele bağlı olarak maddelerin güçlük, ayırt edicilik ve şans parametreleri ile ilgili her bir madde düzeyinde doğrudan bilgi vermektedir (Desjardins ve Bulut, 2018). De Boeck ve Wilson (2004) tarafından ileri sürülen Açıklayıcı Madde Tepki Modellemesi (AMTM) ise, madde kümeleri, birey grupları ve bunların etkileşimleri arasındaki ortak değişkenliğin ölçülebilmesini sağlamaktadır. Geleneksel

MTK modellerinden farklı olarak AMTM maddeler ve bireyler ile ilgili açıklayıcı değişkenleri kullanarak madde yanıtlarındaki ortak değişkenliği açıklamayı amaçlamaktadır.

AMTM'nin temel avantajı, eş zamanlı olarak maddeler ve bireyler arasındaki ortak değişkenliği ayrıştırabilirken, maddeler ve bireyleri analiz etmedeki esnekliğidir. (Briggs, 2008). AMTM'nin genel özellikleri şu şekilde özetlenebilir:

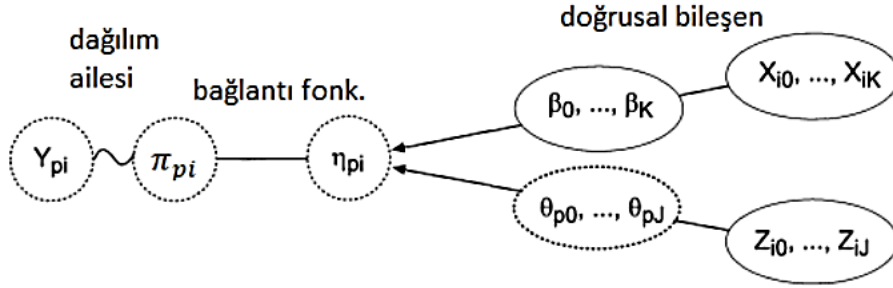
- Hem ölçme hem de açıklama amacıyla kullanılabilir (De Boeck ve Wilson, 2004; Desjardins ve Bulut, 2018).
- Açıklayıcı değişkenler kullanılarak bağımlı değişken daha tutumlu bir yolla kestirilebilir. Açıklayıcı değişkenlerin maddeler üzerindeki genel etkisinin açıklanabilmesini sağlar.
- Madde yanıtlarını çok düzeyli bir çerçevede inceleyerek, açıklayıcı değişkenlere ait etkinin düzeyler arasında sabit ya da rastgele etki olarak ele alınmasına olanak tanır (De Boeck ve Wilson, 2004).
- Maddelerin birinci, bireylerin ikinci düzeyde yer aldığı çok düzeyli bir yapı içinde, maddeleri tekrarlı ölçümler, bireyleri kümeler ve maddelere verilen yanıtları bağımlı değişken olarak ele almaktadır. Modellere dahil edilen açıklayıcı değişkenler ise, madde yanıtlarını yordayan bağımsız değişkenlerdir (Desjardins ve Bulut, 2018).

AMTM doğru yanıt olasılığını kişi ve madde özelliklerinin bir fonksiyonu olarak birlikte modeller. Ortak modelleme, kişi ve madde özelliklerinin etkilerini eş zamanlı olarak tahmin ettiği için örtük bir özellikteki bireysel farklılıkların daha doğru bir şekilde tahmin edilmesini sağlar (De Boeck ve Wilson, 2004). Yani, ortak modelleme, verideki tüm bilgileri (birey veya madde özellikleri) kullanarak doğru yanıt olasılığını kestirilir.

Genelleştirilmiş Doğrusal Karma Modeller

AMTM örtük değişken modellerinin bir alt sınıfı olarak, MTK modellerinin, içsel ve dışsal değişkenler dahil edilerek genelleştirilmiş doğrusal (GDKM) ve doğrusal olmayan karma modeller (DOKM) formunda yeniden modellenerek tasarlandığı esnek psikometrik bir yaklaşımdır (De Boeck ve Wilson, 2004). GDKM ve DOKM tekrarlı ölçümler gibi kümelenmiş verilerin analizinde kullanılabilirler. Bu modeller ile rastgele

etkiler modellenebilir ve normal dağılmayan verilerle analiz yapılmasına olanak sağlar. Şekil 2’de GDKM ait grafiksel gösterim sunulmuştur.



Şekil 2. GDKM Ait grafiksel gösterim (Wilson, DeBoeck ve Carstensen, 2008, 92)

Şekil 2’de GDKM’lerin, gözlem sonuçlarını olasılıklara bağlayan bir dağılım ailesi, olasılıklar arasındaki dönüşümü ifade eden bir bağlantı fonksiyonu ve birey ve madde özelliklerini olasılıklara bağlayan doğrusal bileşen olmak üzere üç bileşenden oluştuğu görülmektedir. GDKM’lerin, doğrusal bileşenine ilişkin matematiksel gösterim ve bu gösterimlerdeki bileşenler aşağıda sunulmuştur.

$$\eta_{pi} = \sum_{k=0}^K \beta_k X_{ik} + \sum_{j=0}^J \theta_{pj} Z_{ij}$$

GDKM’lerin doğrusal bileşenine ilişkin matematiksel gösterimde k maddeye ilişkin değişkenleri, j bireye ilişkin değişkenleri temsil etmektedir. Sabit etkiler, bireyler arasında değişmeyen β_k , rastgele etkiler bireyler arasında değişen θ_{pj} ile ifade edilmektedir. Eşitlikte β_k k yordayıcı değişkeninin etkisini, X_{ik} k değişkeni için i maddesinin aldığı değeri, θ_{pj} p bireyi için j yordayıcı değişkenin etkisini, Z_{ij} j değişkeni için i maddesinin aldığı değeri, η_{pi} p bireyinin i maddesini doğru yanıtlama olasılığının lojit ölçeğindeki değerini temsil etmektedir. Lojit ölçeğindeki değerlerin eksponansiyeli ise odds oranına karşılık gelmektedir.

AMTM sabit ve değişen etkiler olmak üzere iki tür etkiyi içermektedir (De Boeck, Cho ve Wilson, 2016). Sabit etkiler tüm bireyler için aynı değerlere sahipken değişen etkiler bireylere göre değişkenlik gösterir ve rastgele etkiler olarak adlandırılırlar. Rastgele etkiler, sabit bir ortalama ve bu ortalamadan bir sapma olmak üzere iki bileşene

sahiptir. AMTK modellemesinde birey parametreleri rastgele etkiler olarak ele alınmakta olup, ortalaması 0 olan normal dağılım göstermektedir ($\theta p \sim N(0, \theta_\varepsilon^2)$).

AMTM ile iki aşamalı bir süreç kullanılarak birinci aşamada madde ve yetenek parametreleri kestirilirken ikinci aşamada ölçümlerle ilişkili olabilecek dışsal faktörlerin etkisi incelenir. Modele dahil edilen değişkenlere göre Tablo 1’de görüldüğü gibi dört farklı model tanımlanabilmektedir.

Tablo 1

Açıklayıcı Madde Tepki Modelleri

Madde yordayıcıları	Birey yordayıcıları	
	Yok	Var
Yok	Çifte tanımlayıcı	Birey açıklayıcı
Var	Madde açıklayıcı	Çifte açıklayıcı

Tablo 1’de farklı türde yordayıcıları içeren dört tür model görülmektedir. Her bireyin birey özellikleri ve her maddenin de madde özellikleri ile açıklanamayan kendine özgü bir etkisi söz konusu olduğunda bu çifte tanımlayıcı model olarak adlandırılmaktadır. Bu model birey ve maddeye özgü etkileri açıklamaz, madde ve birey yordayıcılarını içermez ve genellikle ölçme amaçları için uygundur (Wilson, De Boeck ve Carstensen, 2008). Yalnızca birey yordayıcılarını içeren modeller birey açıklayıcı, yalnızca madde yordayıcılarını içeren modeller madde açıklayıcı, birey ve madde yordayıcılarının her ikisini birden içeren modeller ise çifte açıklayıcı modeller olarak adlandırılmaktadır (De Boeck ve Wilson, 2004).

Rasch Model

AMTM’nin temelini oluşturan Rasch model boş model ya da çifte tanımlayıcı model olarak adlandırılmaktadır. Herhangi bir açıklayıcı değişken içermeyen bu modelde yalnızca her bir madde ve bireye ait etkiler kestirilir. Birey yetenek puanlarının evrenden rastgele çekildikleri, yetenek puanlarının normal dağıldığı, ortalamalarının 0’a eşit olduğu ($\theta p \sim N(0, \theta_\varepsilon^2)$) varsayılır (De Boeck ve Wilson, 2004). Rasch modelin doğrusal bileşenine ilişkin matematiksel gösterim ve bu gösterimlerdeki bileşenler aşağıda sunulmuştur.

$$\eta_{pi} = \theta_p X_{i0} + \sum_{k=1}^K \beta_i X_{ik}$$

η_{pi} , p bireyinin i maddesini doğru yanıtlama olasılığı

θ_p , p bireyinin yetenek parametresi,

X_{i0} , tüm maddelerin ortak özelliği (bire eşittir),

β_i , i maddesinin kolaylık parametresi (madde güçlük parametresinin tersi),

X_{ik} , k değişkeni için i maddesinin aldığı değer (i=k ise 1, diğer durumlarda 0'a eşittir)

Doğrusal Lojistik Test Modeli

Yalnızca madde düzeyinde açıklayıcı değişkenler içeren modellerdir. Madde açıklayıcı madde tepki modeli olarak da adlandırılmakta olup, yalnızca madde açıklayıcı değişkenlerin eklendiği modeldir. Madde özelliklerinin modellenmesi Doğrusal Lojistik Test Modeli (DLTM) veya çok düzeyli MTK uygulamaları ile yapılabilir. Fischer (1973) tarafından geliştirilen doğrusal lojistik test modeli madde güçlük parametresinin doğrudan maddeye ilişkin faktörlerin bir fonksiyonu şeklinde tanımlanmasını sağlayan bir yöntemdir (Hartig ve diğerleri, 2012).

DLTM'in doğrusal bileşenine ilişkin matematiksel gösterim ve bu gösterimlerdeki bileşenler aşağıda sunulmuştur.

$$\eta_{pi} = \theta_p + \sum_{k=1}^K \beta_k X_{(p,i)k}$$

η_{pi} , p bireyinin i maddesini doğru yanıtlama olasılığı

θ_p , p bireyinin yetenek parametresi,

β_k , k yordayıcı değişkeninin etkisi,

$X_{(p,i)k}$, i maddesi k özelliğine dair değeri taşıyorsa 1, taşımıyorsa 0'a eşitlenir.

DLTM ile aynı kategorideki tüm maddeler için eşit ve hatasız madde kolaylığı (madde güçlüğü'nün negatifi) kestirilir. Bu temel haliyle tüm aynı kategorideki maddelerin eşdeğer olduğu varsayılır ki bu durum çoğu zaman gerçekçi olmayan bir

varsayımdır. Modele her bir madde için madde düzeyindeki artık değer gibi işlev gören rastgele hata terimi eklenerek DLTM_E modeli elde edilir (Desjardins ve Bulut, 2018).

$$\eta_{pi} = \theta_p + \sum_{k=1}^K \beta_k X_{(p,i)k} + \varepsilon_i$$

Yukarıda DLTM_E modelinin doğrusal bileşenine ilişkin matematiksel gösterim görülmektedir. DLTM_E modelinin yardımı ile her bir maddenin kendine özgü madde kolaylık parametresi hesaplanabilir. Yüksek hata varyansı madde değişkenlerinin madde güçlüklerini açıklamada daha az başarılı olduğu anlamına gelir (De Boeck ve Wilson, 2004; De Boeck, Cho ve Wilson, 2011).

Örtük Regresyon Rasch Model

Yalnızca birey düzeyinde açıklayıcı değişkenler içeren modellerdir. Birey açıklayıcı madde tepki modeli ya da Örtük Regresyon Rasch Modeli (ÖRRM) olarak adlandırılmaktadır. Örtük regresyon modelinde yetenek puanlarının normal dağılması varsayımı vardır. Bundan dolayı bu model, örneklem evren ortalamasından farklılaşmadığı sürece faydalı olabilmektedir (De Boeck, Bakker ve diğerleri, 2011).

ÖRRM'in doğrusal bileşenine ilişkin matematiksel gösterim ve bu gösterimlerdeki bileşenler aşağıda sunulmuştur.

$$\theta_{pi} = \sum_{j=1}^J \vartheta_j Z_{(p)j} + \varepsilon_p$$

$$\eta_{pi} = \sum_{j=1}^J \vartheta_j Z_{(p)j} + \varepsilon_p + \beta_i X_{pi}$$

$Z_{(p)j}$, p bireyinin J (j=1, ..., J) özelliğine ilişkin değeri (birey değişkeni),

ϑ_j , j özelliğinin sabit etkisi,

ε_p , birey değişkenleri kontrol altında tutulduğunda kalan birey etkisi ($\varepsilon_p \sim N(0, \theta_\varepsilon^2)$)

Örtük Regresyon Doğrusal Lojistik Test Modeli

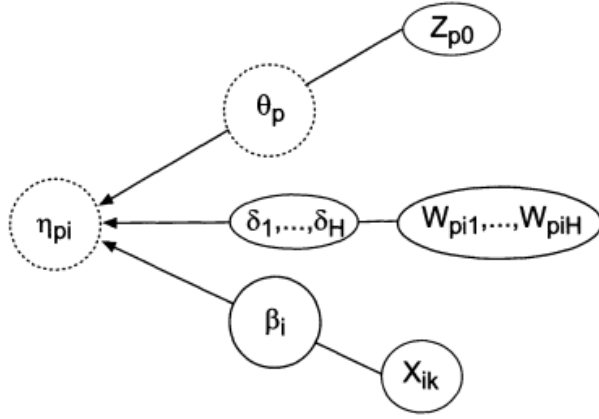
Örtük Regresyon Doğrusal Lojistik Test Modeli (ÖRDLTM) ya da diğer adıyla çifte açıklayıcı madde tepki modeli, birey ve madde düzeyinde açıklayıcı değişkenlerin eklendiği ve en az sınırlamanın olduğu modeldir. ÖRDLTM'in doğrusal bileşenine ilişkin matematiksel gösterim ve bu gösterimlerdeki bileşenler aşağıda sunulmuştur.

$$\eta_{pi} = \sum_{j=1}^J \vartheta_j Z_{pj} + \varepsilon_p + \sum_{k=1}^K \beta_k X_{ik}$$

$Z_{(p)j}$, p bireyinin J (j=1, ..., J) özelliğine ilişkin değeri (birey değişkeni), p=j şartı sağlanıyorsa bire, sağlanmıyorsa 0'a eşittir (De Boeck ve Wilson, 2004). Matematiksel gösterim incelendiğinde bu modelin DLTM ve ÖRRM'nin bileşiminden oluştuğu görülmektedir. AMTM kapsamında temel olarak yer alan bu modeller dışında özel durumlar için geliştirilmiş çeşitli modellere de ulaşmak mümkündür. Birey ve madde özelliklerinin çaprazlandığı madde bazında birey değişkenlerin dahil edilebildiği modellerden biri de Madde Bazında Birey Değişkenleri İçeren Genişletilmiş Rasch Modeldir.

Madde Bazında Birey Değişkenleri İçeren Genişletilmiş Rasch Model

Madde bazında birey yordayıcıları, birey ve madde yordayıcılarından farklı olarak hem bireyler içinde hem bireyler arasında değişmektedir. Bu değişkenlerin modellenmesi değişen madde fonksiyonu ve madde yerel bağımsızlığı gibi çeşitli olguların çalışılması için kullanılmaktadır. Şekil 3'te madde bazında birey yordayıcıları ile genişletilmiş bir Rasch modelin grafiksel temsili görülmektedir.



Şekil 3. Madde Bazında Birey Değişkenleri İçeren Genişletilmiş Rasch Modelin Grafikselsel Temsili (Meulders ve Xie, 2004 s.215)

Şekil 3'te görüldüğü gibi bu modelde rastgele bileşen ve lojit bağlantı çıkarılmıştır. Bu model, p bireyinin i maddesini doğru yanıtlama olasılığının lojitini (η_{pi}) açıklamaktadır. Bu açıklama bireyler üzerinde rastgele etkisi olan θ_p ağırlığı ile bir sabit Z_0 birey yordayıcısına; ($K=I$ olmak üzere) X_{i1} ile X_{iK} arasında değişen ve bunlara karşılık gelen β_1 ile β_I arasında sabit ağırlıkları olan I madde yordayıcılarına ve W_{pi1} ile W_{piH} arasında değişen δ_1 ile δ_H arasında sabit ağırlıkları olan H madde bazında birey yordayıcılarına bağlı olarak yapılmaktadır. W yordayıcıları kişi ve maddeye bağlı olan değerlere sahip özellikler ya da bir başka maddeye verilen yanıtlar olabilir.

Statik etkileşim modelleri, modellenen yanıtlara doğrudan dayalı olmayan ancak bağımsız bir şekilde destekleyen ya da birey ve madde yordayıcıları temelinde yapılandırılan maddeye göre birey yordayıcılarını içermektedir. Bu modellere örnek olarak DIF çalışmaları gösterilebilir. Örneğin, madde konum etkisi, birey ve madde değişkenin çaprazlanmasıyla üretilen ve grup üyeliğini işaret eden bir maddeye göre birey yordayıcısı olarak ele alınarak modele dahil edilebilir.

Dinamik etkileşim modellerinde madde yanıtları diğer maddelere verilen yanıtlar temel alınarak modellenir. Koşullu modeller olarak adlandırılan bu modeller bir yanıtın olasılığını, yordayıcılar ve diğer yanıtlar göz önüne alınarak belirlemektedir. Bilişsel ve eğitsel testler bağlamında, dinamik modeller genellikle test alma sırasında öğrenmeyi değerlendirmek için kullanılır. Bir başka uygulama ise bir maddeye verilen yanıtları, aynı uyarıcı içeriğini paylaşan (belli bir noktaya kadar) başka bir maddeye verilen yanıtlar temelinde modellemektir.

BÖLÜM 3

YÖNTEM

Bu bölümde yürütülen araştırmanın modeli, ilgili modelin seçilme gerekçesi, katılımcılar, kullanılacak veriler, verilerin çözümlenmesi aşamalarına yer verilmiştir.

Araştırmanın Modeli

Bu araştırmada süreç verilerinin, madde ve birey değişkenlerinin, maddelerin doğru yanıtlanma olasılığı üzerindeki etkileri incelendiğinden araştırma modeli korelasyonel araştırma türüne girmektedir. Korelasyonel araştırma, iki ya da daha fazla sayıdaki değişken arasında birlikte değişimin varlığını ve/veya derecesini belirlemeyi amaçlamaktadır (Fraenkel, Wallen ve Hyun, 2012). Araştırmada ele alınan değişkenlerden maddelerin doğru yanıtlanma olasılığı bağımlı değişken, açıklayıcı değişken olarak modele dahil edilen süreç verilerinin, birey ve madde özelliklerinin her biri bağımsız değişken olarak ele alınarak AMTM kapsamında modellenmiştir. Bu modeller GDKM çerçevesinde çaprazlanmış çok düzeyli regresyon modelleridir. Diğer bir deyişle, tepki süresi, eylem sayısı, görüntüleme sayısı, cinsiyet, evdeki kitap sayısı, madde türü, bağlam, metin biçimi, bilişsel boyut ve madde zorluk düzeyi bağımsız değişkenlerdir.

Çalışma Grubu

Bu araştırmada toplam tepki süresi, eylem sayısı ve görüntüleme sayısı verilerine ulaşılabilen ve maksimum sayıda öğrencinin aynı madde kümesini yanıtladığı PISA 2015 uygulamasından elde edilen veriler kullanıldığı için evren ve örneklem PISA’da tanımlandığı şekliyle ele alınmıştır. PISA uygulamalarında evren 15 yaş gurubunda ve 7. sınıf ya da üstü eğitime devam eden öğrenciler olarak tanımlanmaktadır. PISA uygulamalarında örneklem tabakalı seçkisiz örnekleme yöntemiyle iki aşamalı olarak seçilmektedir. Türkiye örneklemini için ilk olarak İstatistikî Bölge Birimleri Sınıflaması (İBBS) Düzey 1, eğitim türü, okul türü, okulların bulundukları yer ve idari biçimleri

tabakaları kullanılarak tabakalı seçkisiz örnekleme yöntemiyle okullar belirlenmiştir (MEB, 2016). Bu doğrultuda 12 bölgeyi temsil eden 61 ilden 187 okul örnekleme alınmıştır. İkinci aşamada ise bu okullarda öğrenim gören öğrenciler arasından seçkisiz örnekleme ile uygulamaya katılacak 5895 öğrenci seçilmiştir. Öğrenciler cinsiyete göre eşit dağılım göstermektedir başka bir ifadeyle örneklemin %50'si kız ve %50'si erkek öğrencilerden oluşmaktadır. Bu çalışmanın katılımcılarını PISA 2015 uygulamasına Türkiye'den katılan ve okuma becerileri alanında en fazla ikili kodlanan maddenin yer aldığı 16 maddelik 6 numaralı madde kümesini yanıtlayan 737 öğrenci oluşturmaktadır.

Verilerin Elde Edilmesi

Araştırma verileri PISA 2015 uygulamasından elde edilen okuma becerisine ilişkin madde puanları, süreç verileri (log data) ve Öğrenci Anketi ile toplanan verilerden oluşmaktadır. Bu çalışmada ele alınan değişkenler Öğrenci Anketi (Student Questionnaire) ve Öğrenci Bilişsel Verileri (Student Cognitive) ve Öğrenci Bilişsel Süre (Student Cognitive Timing) veri setlerinden çekilmiştir. Bu çalışmada PISA verisinin kullanılmasının başlıca nedenleri Türkiye'nin 2015 uygulamasından itibaren bilgisayar tabanlı ölçme uygulamasına katılmış olması, örnekleminin temsil ediciliğinin yüksek olması ve PISA verisinin psikometrik olarak güçlü ölçümler sağlaması ve benzer çalışmalarda etkili bir şekilde kullanılmış olmasıdır. PISA uygulamalarına ilişkin veriler OECD tarafından internet ortamında yayınlanmaktadır. Bahsedilen verilere internet ortamından (<https://www.oecd.org/en/data/datasets/pisa-2015-database.html>) erişim sağlanmıştır. Araştırma sürecinde etik kurul onayı alınmış ve ilgili izin belgesi EK-1 de sunulmuştur.

Okuma becerilerinin ölçülmesi için hazırlanan 7 farklı madde kümesi içerdiği ikili kodlanan madde sayısı ve maddeleri alan maksimum öğrenci sayısı temel alınarak incelendiğinde 6 numaralı madde kümesinin maksimum sayıda (16) madde içerdiği ve en çok öğrenci tarafından yanıtladığı görüldüğünden çalışma bu madde kümesi üzerinde yürütülmüştür (OECD, 2017b). Bu doğrultuda, 6 numaralı okuma becerileri madde kümesine ve bu madde kümesini alan öğrencilere ilişkin veriler bu çalışmanın verilerini oluşturmaktadır.

Araştırmada Ele Alınan Değişkenler

Araştırma kapsamında ele alınan değişkenler, türleri ve düzeyleri Tablo 2’de verilmiş ve kullanılan veri çözümleme yönteminin terminolojisine uygun olarak birey yordayıcıları, madde yordayıcıları ve madde bazında birey yordayıcıları olmak üzere üç kategoride açıklanmıştır.

Tablo 2

Araştırmada Ele Alınan Değişkenler

Kategori	Değişken	Türü	Ölçüm
Birey Yordayıcıları	Cinsiyet	Kategorik	Kız (0) Erkek (1)
	Evdeki Kitap Sayısı	Kategorik	0-10 arasında (1), 11-25 arasında (2), 26-100 arasında (3), 101-200 arasında (4), 201-500 arasında (5), 500’den çok (6)
Madde Yordayıcıları	Madde türü	Kategorik	Çoktan Seçmeli (1) Açık Uçlu (2)
	Bağlam	Kategorik	Eğitimsel (1), Kamusal (2), Kişisel (3)
	Metin biçimi	Kategorik	Karma (1), Sürekli (2), Sürekli Olmayan (3), Çoklu (4)
	Bilişsel boyut	Kategorik	Bilgiye ulaşma ve hatırlama (1), Bir araya getirme ve yorumlama (2), Yansıtma ve değerlendirme (3)
	Madde zorluk düzeyi	Kategorik	En düşük zorluk (1), ..., En yüksek zorluk (6)
Madde Bazında Birey Yordayıcıları	Toplam Tepki Süresi	Sürekli	
	Eylem Sayısı	Sürekli	
	Görüntüleme Sayısı	Sıralı Sınıflama	Bir kez (1) Birden çok (2)

Birey yordayıcıları, bireyler arasında değişen ancak maddeler arasında değişmeyen birey özelliklerini temsil etmektedir. Öğrencilerin cinsiyeti ve evdeki kitap sayısı birey yordayıcı değişkenler olarak AMTM çerçevesinde modellere dahil edilmiştir.

Cinsiyet iki kategorili bir değişkendir. Evdeki kitap sayısı ise altı kategorili bir değişken olarak tanımlanmıştır.

Madde yordayıcıları, maddeler arasında değişen ancak bireyler arasında değişmeyen madde özelliklerini ifade etmektedir. Madde özelliklerden madde türü, bağlam, metin biçimi, bilişsel boyut ve madde zorluk düzeyi değişkenleri madde yordayıcıları olarak AMTM çerçevesinde modellere dahil edilmiştir. Araştırmada kullanılan madde kümesinde yer alan maddelerin özellikleri PISA 2015 teknik raporundan (OECD, 2017b) özetlenerek Tablo 3'te sunulmuştur.

Tablo 3

Araştırmada Ele Alınan Maddeler ve Özellikleri

Madde Konumu	Madde Türü	Bağlam	Metin Biçimi	Bilişsel Boyut	Madde Zorluk Düzeyi
1	Çoktan Seçmeli	Eğitimsel	Karma	Bir araya getirme ve yorumlama	3
2	Açık Uçlu	Eğitimsel	Sürekli Olmayan	Bilgiye ulaşma ve hatırlama	2
3	Çoktan Seçmeli	Eğitimsel	Sürekli Olmayan	Bilgiye ulaşma ve hatırlama	3
4	Çoktan Seçmeli	Eğitimsel	Sürekli Olmayan	Bir araya getirme ve yorumlama	3
5	Açık Uçlu	Kişisel	Sürekli	Bilgiye ulaşma ve hatırlama	3
6	Açık Uçlu	Kişisel	Sürekli	Bir araya getirme ve yorumlama	3
7	Açık Uçlu	Kişisel	Sürekli	Yansıtma ve Değerlendirme	5
8	Açık Uçlu	Kişisel	Sürekli	Yansıtma ve Değerlendirme	5
9	Çoktan seçmeli	Kişisel	Sürekli	Bir araya getirme ve yorumlama	4
10	Karmaşık Çoktan Seçmeli	Kişisel	Çoklu	Bilgiye ulaşma ve hatırlama	3
11	Karmaşık Çoktan Seçmeli	Kişisel	Çoklu	Bir araya getirme ve yorumlama	3
12	Çoktan Seçmeli	Kamusal	Sürekli	Bir araya getirme ve yorumlama	4
13	Çoktan Seçmeli	Kamusal	Sürekli	Bir araya getirme ve yorumlama	2
14	Çoktan Seçmeli	Kamusal	Sürekli	Yansıtma ve Değerlendirme	3
15	Çoktan Seçmeli	Kamusal	Sürekli	Bir araya getirme ve yorumlama	2
16	Çoktan Seçmeli	Kamusal	Sürekli	Bir araya getirme ve yorumlama	5

Tablo 3'te madde türünün, çoktan seçmeli ve açık uçlu ve karmaşık çoktan seçmeli olmak üzere üç kategorisi olduğu görülmektedir. Karmaşık çoktan seçmeli maddeler diğer madde türlerine kıyasla az sayıda olduğundan, çoktan seçmeli madde olarak yeniden kodlanmış ve madde türü iki kategorili bir değişken olarak ele alınmıştır. Madde bağlamı (OECD, 2017a) eğitimsel, kamusal ve kişisel olmak üzere üç kategorili bir değişkendir. Eğitimsel bağlam, öğretim amaçları için özel olarak tasarlanan metinlerdir (örn; ders kitapları ya da etkileşimli öğrenme yazılımları gibi). Kişisel bağlam, bir bireyin hem pratik hem de entelektüel kişisel ilgilerini karşılaması amaçlanan metinlerle ilgilidir (örn; mektup, kurgu, biyografi gibi). Kamusal kategori, daha geniş bir toplumun faaliyetleri ve kaygılarıyla ilgili metinlerin tanımlar (örn; resmi dokümanlar, kamusal etkinliklerle ilgili bilgiler gibi).

Metin biçimi metnin nasıl organize edildiğini ifade etmektedir. Metin biçimlerinin çeşitliliği, “yalnızca belirli bir metin parçasını anlamayı değil, aynı zamanda hem okula yönelik hem de okul dışı amaçlarla okumayla ilişkili tüm metinlerden bilgi bulma, seçme, yorumlama ve değerlendirme becerisini de içeren” geniş bir okuma anlayışını yansıtmaktadır (OECD, 2017a). PISA 2015 okuma becerileri çerçevesinde metin biçimleri sürekli, sürekli olmayan, karma ve çoklu olmak üzere dört kategoride ele alınmaktadır. Sürekli metinler geleneksel yazılı formlardır ve daha kolay tanımlanabilen ve anlaşılabilen temel bir organizasyonel ve dilsel yapıya sahiptir. Sürekli olmayan metinler, metin organizasyonunun parçalı olduğu, yani olgusal, nicel, teknik veya matematiksel bilgilerin haritalar, çizelgeler, grafikler, zaman çizelgeleri, tablolar ve diyagramlarda sunulduğu bilgiye ulaşmak için gerekli okumaların yapılmasını içeren metinlerdir. Karma metinler, sürekli ve sürekli olmayan metinlerin özelliklerini birleştiren yazılı formlar olarak "sözel ve resimsel bilginin birlikte sunulması" ile karakterize edilir (Rouet vd., 2008). Gazeteler, dergiler, ya da web sayfaları karma metinlere örnek verilebilir. Çoklu metinler ise, bağımsız olarak oluşturulmuş ve bağımsız olarak anlam ifade eden metinler olarak tanımlanır; belirli bir durum için yan yana getirilirler veya değerlendirme amacıyla zayıf bir şekilde birbirine bağlanmış olabilirler (OECD, 2017a). Metinler arasındaki ilişki açık olmayabilir; birbirlerini tamamlayabilir veya birbirleriyle çelişebilirler. Örneğin, seyahat tavsiyesi veren farklı şirketlere ait bir dizi web sitesi turistlere benzer yönlendirmeler sağlayabilir veya sağlamayabilir. Çoklu metinler tek bir formata sahip olabilir (örneğin sürekli) veya hem sürekli hem de sürekli olmayan metinler içerebilir. İncelenen veri setinde her bir kategoriye temsil eden madde sayısı dengeli olmadığı için bu kategorilerin yeniden düzenlenmesi yoluna gidilmiştir. Bu

doğrultuda Karma, Sürekli Olmayan ve Çoklu metin biçimindeki maddeler “Diğer” şeklinde kodlanarak, maddeler “Sürekli” ve “Diğer” metin biçimi olmak üzere iki kategoride incelenmiştir.

Madde bilişsel boyutu ise “Bilgiye ulaşma ve hatırlama”, “Bir araya getirme ve yorumlama” ve “Yansıtma ve değerlendirme” olmak üzere üç kategorili bir değişkendir (OECD, 2017a). Bilgiye ulaşma ve hatırlama, bir veya birden fazla farklı bilgi parçasını bulma ve geri alma sürecini ifade etmektedir. Bu süreç, belirli bir bilgi alanına erişmeyi ve bu alan içinde etkili bir şekilde gezinmeyi gerektirir. Erişim ve geri getirme görevleri, geniş bir yelpazeye yayılmaktadır. Örneğin, bir işverenin iş ilanında talep ettiği ayrıntıları belirlemek, belirli bir alan koduna sahip bir telefon numarasını bulmak ya da bir iddiayı desteklemek veya çürütmek için ilgili bir bilgiyi tespit etmek gibi farklı amaçlarla gerçekleştirilebilir.

Bir araya getirme ve yorumlama, okunan metni anlamlandırma sürecini kapsar. Bir araya getirme, metnin tutarlılığını kavrayarak farklı bilgi parçalarını bir araya getirmeyi ve anlamlı ilişkiler kurmayı içerir. Bu süreç, benzerlik ve farklılıkları belirleme, karşılaştırmalar yapma ve neden-sonuç ilişkilerini anlama gibi becerileri gerektirir. Yorumlama ise, metinde açıkça belirtilmeyen anlamları çıkarma sürecidir. Okuyucu, metnin altında yatan varsayımları ve ima edilen anlamları analiz eder. Geniş bir anlayış oluşturabilmek için hem bütünleştirme hem de yorumlama gereklidir. Okuyucunun metni genel çerçevede değerlendirmesi, ana fikri, mesajı veya metnin amacını belirleyerek temel anlamı kavraması beklenir.

Yansıtma ve değerlendirme, metindeki bilgileri okuyucunun kendi bilgi birikimi, deneyimleri ve düşünsel çerçevesiyle ilişkilendirmesini gerektirir. Yansıtma, okuyucunun kendi deneyim veya bilgilerini kullanarak metindeki bilgileri karşılaştırmasını, farklılıkları ve benzerlikleri analiz etmesini ya da olasılıklar üzerine düşünmesini içerir. Değerlendirme ise, okuyucunun metinde sunulan bilgileri bağımsız ölçütlere göre yargılamasını ve değerlendirmesini gerektirir. Bu süreçte, okuyucu metni nesnel veya öznel ölçütlere göre ele alarak anlamını ve değerini belirler.

Madde zorluk düzeyi PISA uygulamalarında toplanan veriler üzerinden OECD tarafından kestirilmektedir. Bu süreçte maddelerin geliştirildiği ve ilk kez uygulandığı yıl toplanan veriler üzerinden hem öğrencilerin yeterlik düzeyleri hem de maddelerin güçlük düzeyleri aynı ölçek (contunium) üzerinde yerleştirilmektedir. Öğrenci yeterliklerinin puanlanmasında ve gruplanmasında olduğu gibi maddelerin güçlükleri de benzer şekilde puanlanmakta ve düzeylerine göre gruplanmaktadır. Buna göre madde güçlük düzeyi 1

en kolay ve 6 en zor olmak üzere sıralama ölçeğine yerleştirilmiş altı düzeyli bir değişkendir. Bu çalışmada incelenen madde kümesinde yer alan maddelerin zorluk düzeyleri ise 2 ile 5 arasında değişmektedir. Bu nedenle madde zorluk düzeyi dört kategorili bir değişken olarak incelenmiştir. Tablo 3 incelendiğinde maddelerin büyük çoğunluğunun çoktan seçmeli türünde, kişisel bağlamda, sürekli metin biçiminde, bir araya getirme ve yorumlama boyutunda ve üçüncü düzey zorlukta olduğu görülmektedir.

Madde bazında birey yordayıcıları, maddeden maddeye değişen birey özelliklerini temsil etmektedir. Örneğin bir bireyin harcadığı toplam süre, gerçekleştirdiği eylem sayısı ve maddeye geri dönüp dönmediği her bir madde için farklılık göstermektedir. Bu çalışma kapsamında süreç verilerinden tepki süresi, eylem sayısı ve görüntüleme sayısı madde bazında birey yordayıcıları olarak AMTM modellerine dahil edilmiştir. Araştırmada ele alınan süreç verileri şu şekilde tanımlanmaktadır (OECD, 2024):

- Eylem sayısı: Öğrencinin maddeyle etkileşimi süresince gerçekleştirdiği tıklama, çift tıklama, tuşa basma ve sürükleme eylemlerinin sayısıdır.
- Tepki süresi: Öğrencinin madde üzerinde harcadığı toplam süre, bir maddeyi görüntülemeye başladığı andan nihai yanıtını verdiği ana kadar harcadığı toplam süredir.
- Görüntüleme Sayısı: Öğrencilerin bir maddeyi kaç kez yeniden görüntülediği başka bir ifadeyle maddeye kaç kez geri dönerek gözden geçirdiğini göstermektedir. Görüntüleme sayısına ilişkin bilgiler, PISA 2018 uygulaması teknik raporunun K ekinde detaylı olarak açıklanmıştır (OECD, 2024).

Verilerin Çözümlemesi

Araştırmanın verileri AMTM (De Boeck ve Wilson, 2004) çerçevesinde yer alan modeller kullanılarak çözümlenmiştir. Bu modeller Rasch Model, Doğrusal Lojistik Test Modeli, Örtük Regresyon Rasch Modeli, Örtük Regresyon Doğrusal Lojistik Test Modeli ve Madde Bazında Birey Değişkenleri İçeren Genişletilmiş Rasch modellerdir (Meulders ve Xie, 2004). Örtük değişken modellerinin bir alt sınıfı olan AMTM, birey ve madde yordayıcılarını ayrı ayrı ve bir arada ele alarak bu değişkenlerin birey ve madde parametrelerinin kestirimine etkisini eş zamanlı olarak açıklayabilmekte (De Boeck ve Wilson, 2004) ve açıklanan varyansı ortaya koyabilmektedir. Bu modellerle bir maddenin doğru yanıtlanma olasılığı birey, madde ya da her ikisinin etkileşimleri temel alınarak modellenabilmektedir. Bu modeller çok düzeyli bir yapıya sahiptir. Hatanın dahil edildiği

modeller doğru yanıtlama olasılığının bireyler ve maddeler arasında değiştiği çaprazlanmış karma etkili çok düzeyli regresyon modelleridir. Hatanın dahil edilmediği modellerde ise doğru yanıtlama olasılığı bireyler arasında değişkenlik göstermektedir. Hatanın dahil edildiği ve edilmediği boş modelin matematiksel ifadesi aşağıda sunulmuştur.

$$\text{Logit}(\pi_{pi}) = \eta_{pi} = \beta_0 + \theta_p + \beta_i$$

$$\theta_p \sim N(0, \sigma_\theta^2) \text{ ve } \beta_i \sim N(0, \sigma_\beta^2)$$

$$\text{Logit}(\pi_{pi}) = \eta_{pi} = \beta_0 + \theta_p$$

$$\theta_p \sim N(0, \sigma_\theta^2)$$

Eşitliklerde, π_{pi} , p kişinin i maddesini doğru yanıtlama olasılığıdır. β_0 , genel kesişim noktası (sabit etki) olup, ortalama beceriye sahip bir öğrencinin ortalama kolaylıktaki bir maddeye doğru yanıt verme olasılığı için kestirilen logit değerine karşılık gelmektedir. θ_p ; bireylerin yetenek düzeyi ya da bireye özgü sapmayı, β_i ise maddeye özgü sapmayı ya da madde kolaylığını temsil eden rastgele etkilerdir. Bireye ve maddeye özgü sapmaların ortalaması 0 ve varyansı sırasıyla σ_θ^2 ve σ_β^2 olan normal dağılım gösterdiği varsayılmaktadır.

AMTM birey ve madde parametreleri kestirilirken aralarındaki ortak değişkenliği eş zamanlı olarak analiz etme esnekliği sağlar (Briggs, 2008) ve hem ölçme hem açıklama amacıyla kullanılabilir (De Boeck ve Wilson, 2004). AMTM ile iki aşamalı bir süreç kullanılarak birinci aşamada madde ve yetenek parametreleri kestirilirken ikinci aşamada ölçümlerle ilişkili olabilecek dışsal faktörlerin etkisi incelenir. Diğer yandan birden çok birey bazında madde değişkeninin, bu çalışma kapsamında süreç verilerinin, çaprazlanmış değişkenler olarak modellenebilmesi için dinamik etkileşim modelleri kullanılmıştır. Verilerin analizi R programlama dilinde ilgili paketler kullanılarak gerçekleştirilmiştir.

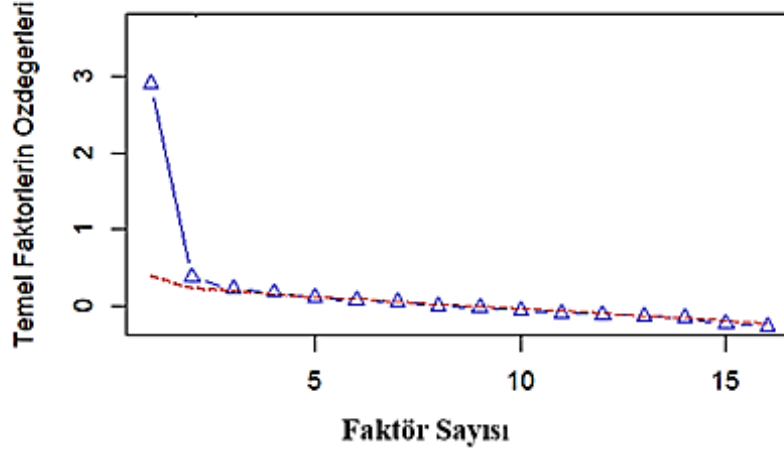
Verilerin Analize Hazırlanması

Veri analizine geçilmeden önce indirgenen veriler kayıp veriler açısından incelenmiştir. Maddenin görüntülendiği ancak yanıtlanmadığı kayıp veriler alinyazınla

tutarlı olarak yanlış kabul edilerek 0 olarak kodlanmıştır. Bunun nedeni bir kişinin bir maddeyi yanıtlama fırsatı olduğu halde yanıtlamamasının muhtemelen yanıtı ulaşamamasına dayandırılmıştır. Ardından veri seti kayıp veriler silinerek analize hazırlanmıştır. 737 öğrencinden 32 öğrenciye ait kayıp veriler silinerek 705 öğrenciye ait veriyle analizlere devam edilmiştir. Verilerin analize hazırlanmasında ikinci aşama olarak ele alınan değişkenlere ilişkin betimsel istatistikler hesaplanarak verilerin yapısı ve dağılımları incelenmiştir. Süreç verilerin dağılımları histogram grafikleri ile incelendiğinde sağa çarpık bir dağılım gösterdiği görülmüştür. Süreç verilerinin aşırı çarpık dağılımlar göstermesi nedeniyle, alanyazında sıklıkla kullanıldığı ve önerildiği şekilde (van der Linden, 2009) daha iyi bir model uyumu sağlanabilmesi için tepki süreleri ve eylem sayılarına log dönüşümü uygulanmıştır. Daha sonra standartlaştırılmış z puanlarına dönüştürülerek uç değerler ($|z| > 3.5$) belirlenmiş ve tüm maddeleri doğru yanıtlayan 3 öğrenciye ait verilerle birlikte veri setinden çıkarılmıştır. Analizlere 625 öğrenci verisi ile devam edilmiştir. Ayrıca maddeyi görüntüleme sayısı bir ve birden çok görüntüleme olmak üzere iki kategorili kategorik değişkene dönüştürülmüştür.

Varsayımların İncelenmesi

Bir sonraki aşamada veriler MTK modellerinin varsayımları açısından incelenmiştir. Tek boyutluluk varsayımının sağlanıp sağlanmadığı sirt paketi (Robitzsch, 2024) ile DETECT (Zhang, 2007) kullanılarak açıklayıcı ve doğrulayıcı faktör analizi yoluyla incelenmiştir. DETECT kullanılarak elde edilen açıklayıcı ve doğrulayıcı faktör analizi sonuçlarına göre sırasıyla üretilen DETECT, ASSI ve RATIO indisleri (DETECT1 = -0.18 , ASSI1 = -0.15, RATIO1 = -0.13 ; DETECT2 = -0.13, ASSI2 = -0.17, RATIO2 = -0.10) ölçüt değerlerin (DETECT<.20; ASSI<.25; RATIO<.36) altında kaldığından temel tek boyutluluğun sağlandığı görülmüştür. Buna ek olarak psych paketi (Revelle, 2024) kullanılarak yürütülen paralel analiz sonuçları Şekil 4'te verilmiş olup tek boyutluluğun sağlandığını desteklemektedir.



Şekil 4. Paralel Analiz Yamaç Birikinti Grafiği

Yerel bağımsızlık varsayımının sağlanıp sağlanmadığını incelemek için, Yen'in Q3 istatistiği hesaplanmıştır. Christensen, Makransky ve Horton (2017) yerel bağımsızlık için sabit bir kritik değer yerine artık korelasyonların ortalamalarına dayalı bir ölçüt kullanılmasını önermektedir. Buna göre, artık korelasyonların ortalamalarının 0.2 üzerindeki korelasyon değerleri yerel bağımlılığa işaret edebilir ve bağımsız maddeler arasındaki artık korelasyonların ortalamasının 0.3 üzerinde olması olası görünmemektedir. Bu çalışmada artık korelasyonların ortalamaları 0.002 bulunmuş olup, tüm madde çiftleri için artık korelasyonların 0.20'den küçük olduğu görülmüştür.

Ardından verilerin, AMTM'nin dayalı olduğu Rasch modelin varsayımlarını karşılayıp karşılamadıklarının incelenmesinde mirt (Chalmers, 2012) paketinden yararlanılmıştır. Rasch model ile maddeler için elde edilen veri uyumu indislerinden Infit ve Outfit değerleri incelenmiştir. Bu değerlerin 0.5 ile 1.5 arasında olması madde uyumunun sağlandığına işaret etmektedir (De Ayala, 2009). Tablo 4'te görüldüğü gibi bu değerlerin 0.69 ile 1.21 arasında olması madde uyumunun sağlandığını göstermektedir.

Tablo 4

Rasch Model İle Modellenen Maddelerin Infit ve Outfit Değerleri

Madde	Infit	Outfit
1	0.97	0.96
2	0.89	0.86
3	0.83	0.78
4	1.10	1.09
5	0.90	0.87
6	0.82	0.77
7	0.84	0.71
8	0.91	0.88
9	0.75	0.69
10	0.90	0.88
11	0.98	0.97
12	1.00	1.02
13	0.94	0.89
14	0.87	0.86
15	0.87	0.80
16	1.12	1.21

Ardından bireyler için elde edilen veri uyumu indislerinden Infit ve Outfit değerleri incelenmiştir. 17 bireye ait uyum istatistiklerinin kritik değer olan $[-2, 2]$ aralığı (Wright ve Masters, 1982) dışında kaldığı ve manidar olduğu görülmektedir. Bu bireylere ait veriler Rasch modele yeterince uyum sağlayamadığı için analiz dışında bırakılmış ve 608 öğrenci verisi ile analizlere devam edilmiştir. Modelin genel uyumunun değerlendirilmesinde incelenen RMSEA değeri 0.07 olup (%95 güven aralığı değerleri 0.063-0.076 arasındadır), kabul edilebilir uyuma (Browne ve Cudeck, 1993) işaret etmektedir.

Veri Çözümleme Teknikleri

AMTM ile yapılan analizlerde verilerin uzun formatta düzenlenmesine diğer bir deyişle her gözlem (maddeye verilen yanıt) bir satıra gelecek şekilde verinin yeniden düzenlenmesine ihtiyaç vardır. Bu nedenle veriler R programlama dilinde reshape2 (Wickham, 2007) paketi kullanılarak uzun formata dönüştürülmüştür. Verilerin uzun formata dönüştürülmesinin ardından bir sonraki aşamada bu modellerin test edilmesine geçilmiştir. AMTM kapsamındaki analizler R programlama dilinde “eirm” (Bulut, 2021) ve “lme4” (Bates vd., 2015) paketi kullanılarak gerçekleştirilmiştir. Çalışmada kurulan

modellerin tamamında Laplace yaklaşımı kullanılmıştır. Modellerin karşılaştırılmasında Akaike Bilgi Kriteri (AIC), Bayesian Bilgi Kriteri (BIC) ve Sapma (-2Log (L)) değerleri incelenmiştir. Düşük AIC, BIC ve Sapma değerleri daha iyi uyuma işaret etmektedir (Wilson, De Boeck ve Carstensen, 2008).

Verilerin analizine AMTM arasından çifte tanımlayıcı model olarak adlandırılan Rasch model ile başlanarak birey ve madde parametreleri kestirilmiştir. Süreç verilerinin maddelerin doğru yanıtlanma olasılığına etkilerinin tek tek ve bir arada incelendiği birinci araştırma sorusuna yönelik olarak hatanın dahil edilmediği ve hatanın dahil edildiği sekiz farklı model kurulmuştur. Süreç verileri kapsamında tepki süresi, eylem sayısı ve görüntüleme sayısı ele alınmıştır. Birinci araştırma sorusuna yönelik olarak kurulan modellerin genel matematiksel ifadesi aşağıda sunulmuştur.

$$\begin{aligned}\eta_{pi} &= \theta_p Z_{p0} + \beta_i X_{ik} + \sum_{h=1}^H \delta_h W_{pih} + \varepsilon_i \\ &= \theta_p Z_{p0} + \beta_i X_{ik} + \delta_1 W_{pi1} + \delta_2 W_{pi2} + \delta_3 W_{pi3} + \varepsilon_i\end{aligned}$$

Madde özelliklerinin maddelerin doğru yanıtlanma olasılığına etkilerinin incelendiği ikinci araştırma sorusuna yönelik olarak DLTM ve hatanın da dahil edildiği DLTM_E modelleri kurulmuştur. Madde özellikleri kapsamında madde türü, metin biçimi, bağlam, boyut ve zorluk düzeyi ele alınmıştır. Madde güçlük düzeyinin modele dahil edilmesinin tekillik sorunu oluşturup oluşturmadığı “performance” paketi kullanılarak tolerans değeri 0.00001 alınarak test edilmiş ve tekillik sorunu olmadığı görülmüştür. İkinci araştırma sorusuna yönelik olarak kurulan DLTM ve DLTM_E modellerinin genel matematiksel ifadesi aşağıda sunulmuştur.

$$\begin{aligned}\eta_{pi} &= \theta_p + \sum_{k=1}^K \beta_k X_{ik} = \theta_p + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} \\ \eta_{pi} &= \theta_p + \sum_{k=1}^K \beta_k X_{(p,i)k} + \varepsilon_i = \theta_p + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5} + \varepsilon_i\end{aligned}$$

Birey özelliklerinin maddelerin doğru yanıtlanma olasılığına etkilerinin incelendiği üçüncü araştırma sorusuna yönelik olarak ÖRRM ve hatanın da dahil edildiği

ÖRRM_E modelleri kurulmuştur. Birey özellikleri kapsamında cinsiyet ve evdeki kitap sayısı ele alınmıştır. Üçüncü araştırma sorusuna yönelik olarak kurulan ÖRRM ve ÖRRM_E modellerinin genel matematiksel ifadesi aşağıda sunulmuştur.

$$\eta_{pi} = \sum_{j=1}^J \vartheta_j Z_{pj} + \theta_p + \beta_i = \theta_p + \vartheta_1 Z_{p1} + \vartheta_2 Z_{p2} + \beta_i$$

$$\eta_{pi} = \sum_{j=1}^J \vartheta_j Z_{pj} + \theta_p + \beta_i + \varepsilon_i = \theta_p + \vartheta_1 Z_{p1} + \vartheta_2 Z_{p2} + \beta_i + \varepsilon_i$$

Süreç verilerinin, birey ve madde özelliklerinin bir arada ele alınarak maddelerin doğru yanıtlanma olasılığına etkilerinin incelendiği dördüncü araştırma sorusuna yönelik olarak üç farklı model kurulmuştur. İlk model birey ve madde özelliklerinin açıklayıcı değişken olarak ele alındığı ve hatanın dahil edildiği ÖRDLTm1 modelidir. İkinci model olan ÖRDLTm2 modeline süreç verileri de açıklayıcı değişken olarak dahil edilmiştir. ÖRDLTm2 modeli ile etkileri istatistiksel olarak manidar bulunmayan değişkenler modelden çıkarılarak üçüncü model olan ÖRDLTm3 modeli kurulmuştur. Madde güçlük düzeyinin modele dahil edilmesinin tekillik sorunu oluşturup oluşturmadığı “performance” paketi kullanılarak tolerans değeri 0.00001 alınarak test edilmiş ve tekillik sorunu olmadığı görülmüştür. Dördüncü araştırma sorusuna yönelik olarak kurulan ÖRDLTm_E modellerinin genel matematiksel ifadesi aşağıda sunulmuştur.

$$\eta_{pi} = \left(\sum_{j=1}^J \vartheta_j Z_{pj} + \theta_p \right) + \sum_{k=1}^K \beta_k X_{(p,i)k} + \sum_{h=1}^H \delta_h W_{pih} + \varepsilon_i$$

$$\eta_{pi} = (\theta_p + \vartheta_1 Z_{p1} + \vartheta_2 Z_{p2}) + (\beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i3} + \beta_4 X_{i4} + \beta_5 X_{i5}) + (\delta_1 W_{pi1} + \delta_2 W_{pi2} + \delta_3 W_{pi3}) + \varepsilon_i$$

Araştırma soruları kapsamında kurulan modeller kendi aralarında karşılaştırılarak en iyi uyum gösteren model belirlenmiştir. Ardından açıklayıcı değişkenlerin etkileri manidarlık düzeyleri göz önünde bulundurularak yorumlanmıştır. Manidarlık düzeyi $p < 0.05$ olarak ele alınmıştır. Maddelerin doğru yanıtlanma olasılığını manidar şekilde

etkileyen deęiřkenler belirlenmiřtir. Bu deęiřkenler ve model ile aıklanan varyans hesaplanmıřtır. GDKM’de aıklanan varyansın hesaplanmasında Nakagawa ve Schielzeth (2013) tarafından nerilen marjinal ve kořullu R^2 istatistięi kullanılmıřtır. Bu yaklařımda marjinal R^2 sabit etkiler ile aıklanan varyansı, kořullu R^2 ise model tarafından aıklanan toplam varyansı gstermekte olup, bu istatistikler “performance” paketi (Ldecke vd., 2021) ile kestirilmıřtir. Aıklayıcı deęiřkenler arasındaki etkileřim etkileri marjinal etki grafikleri ile incelenmiřtir. Birey ve madde varyansları yorumlanmıřtır.



BÖLÜM 4

BULGULAR VE YORUMLAR

.....
Bu bölümde araştırmanın bulgularına ve bu bulgulara dayalı olarak yapılan yorumlara yer verilmiştir. Bulguların sunumunda önce boş model olarak ele alınan Rasch modele ilişkin kestirimler yapılmış sonrasında sırasıyla süreç verilerinin, madde özelliklerinin, birey özelliklerinin ve bunların etkileşimlerinin etkilerine ilişkin kurulan modeller araştırma sorularına yönelik başlıklar altında sunulmuştur. Araştırmada ele alınan değişkenlere ait betimsel istatistikler EK-2’de sunulmuştur.

Boş Modele (Rasch Model) Yönelik Bulgular

Öncelikle madde parametrelerinin kestirildiği boş model olarak adlandırılan Rasch model (M0) ile analizlere başlanmıştır. Rasch model ikili kodlanan bağımlı değişkenin sabit (madde) ve rastgele (bireylerin örtük özellik düzeyleri) etkiler tarafından kestirildiği bir GDKM olarak tanımlanabilmektedir (Desjardins ve Bulut, 2018). Birey ve madde değişkenlerini içermeyen bu model ile sabit etki olarak madde güçlükleri, rastgele etki olarak ise bireylerin yetenek düzeyleri kestirilmiştir.

M0 modeli ile rastgele etkiler olarak kestirilen öğrencilerin okuma becerilerinin lojit ölçeğindeki varyansı 0.96 ve standart hatası 0.98’dir. Lojit ölçeğindeki varyans 2.61 ($\exp(0.96)$) odds oranına karşılık gelmekte olup, yetenek düzeyleri arasında 1 standart sapma fark olan iki öğrencinden yetenek düzeyi 1 standart sapma fazla olan öğrencinin bir maddeyi doğru yanıtlama olasılığı 2.66 kat daha fazladır. Diğer bir deyişle, testteki herhangi bir maddeyi doğru yanıtlama olasılığı 0.50 olan bir öğrenciye göre yetenek düzeyi 1 standart sapma yüksek olan öğrencinin aynı maddeyi doğru yanıtlama olasılığı 0.73’tür. Rasch model ile sabit etkiler olarak kestirilen madde güçlük parametreleri ve kestirimlerin standart hataları Tablo 5’te verilmiştir.

Tablo 5 incelendiğinde, madde güçlüklerinin lojit ölçeğinde -1.42 ile 1.68 arasında değiştiği, en kolay maddenin Madde 13, en zor maddenin ise Madde 7 olduğu görülmektedir. Madde güçlük parametrelerinin ortalaması 0.04 olup testin orta güçlükte bir test olduğu görülmektedir. Bu sonuç ortalama yetenek düzeyindeki bir öğrencinin

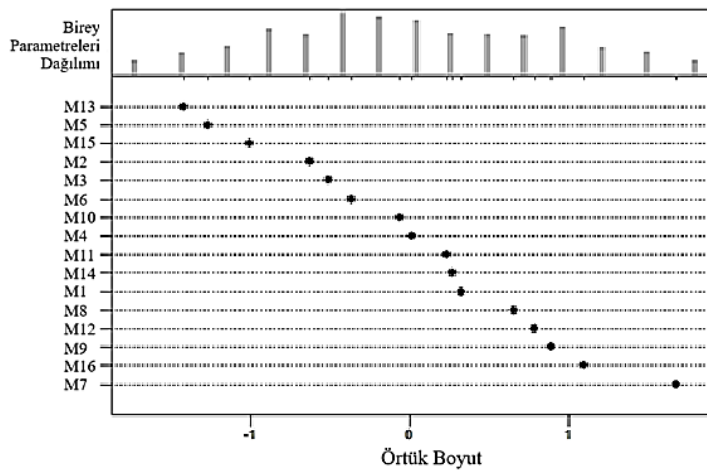
ortalama güçlükteki bir maddeyi doğru yanıtlama olasılığının yaklaşık olarak 0.5 düzeyinde olduğuna işaret etmektedir.

Tablo 5

Rasch model ile Kestirilen Madde Güçlük Parametreleri ve Kestirimlerin Standart Hataları

Maddeler	Tahmin	Standart Hata
Madde 1	0.32	0.10
Madde 2	-0.63	0.10
Madde 3	-0.51	0.10
Madde 4	0.01	0.10
Madde 5	-1.26	0.11
Madde 6	-0.36	0.10
Madde 7	1.68	0.12
Madde 8	0.66	0.10
Madde 9	0.89	0.10
Madde 10	-0.06	0.10
Madde 11	0.24	0.10
Madde 12	0.79	0.10
Madde 13	-1.42	0.11
Madde 14	0.27	0.10
Madde 15	-1.01	0.10
Madde 16	1.10	0.11

Rasch model ile elde edilen kestirimlere dayalı oluşturulan madde güçlük parametrelerinin ve bireylerin örtük özellik düzeylerinin yayılımını gösteren Birey-Madde Haritası aşağıda Şekil 5’te görülmektedir.



Şekil 5. Rasch Model ile Elde Edilen Kestirimlere Dayalı Birey-Madde Haritası

Birey madde haritası, özellikle katılımcıların örtük özellik aralığı göz önüne alındığında, bir testin güçlük aralığının yeterli olup olmadığını gözden geçirmek için yararlıdır (Desjardins ve Bulut, 2018). Şekil 5 incelendiğinde, maddelerin güçlük düzeyleri ile bireylerin yetenek düzeylerinin örtüştüğü, maddelerin ortalama bir güçlük düzeyinde yayılım gösterdiği ve madde güçlüklerinin yayılımının birey yeteneklerinin dağılımına göre kısmen daha dar bir aralıkta yayıldığı görülmektedir. Aşağıda yer alan eşitlik kullanılarak ortalama yetenek düzeyindeki bir öğrenci için ($\theta = 0$), ortalama güçlükte bir maddeyi ($\beta_i = 0.04$) doğru yanıtlama olasılığı hesaplanabilir.

$$\pi_{pi} = \frac{\exp(\beta_i + \theta_p)}{1 + \exp(\beta_i + \theta_p)}$$

Eşitlikten elde edilen 0.51 değeri okuma becerisi yeteneği ortalama düzeydeki bir öğrencinin ortalama madde güçlüğüne sahip bir maddeyi doğru yanıtlama olasılığının 0.51 olduğunu göstermektedir. Bu sonuç testin orta güçlükte bir test olduğunu desteklemektedir.

Süreç Verilerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular

Rasch model ile yapılan incelemelerin ardından, tepki süresi, eylem sayısı ve görüntüleme sayısının maddelerin doğru yanıtlanma olasılığına etkisinin araştırıldığı birinci araştırma problemine ilişkin olarak maddelerin rastgele etkiler olarak modellere dahil edildiği ve edilmediği her bir değişkenin etkisinin tek tek ve hep birlikte incelendiği sekiz farklı model kurulmuştur. Bu modeller uyum indisleri referans alınarak boş model ile karşılaştırılmıştır.

Her bir süreç verisinin bireylerin okuma becerileri üzerindeki etkisi hatanın dahil edilmediği ve maddelerin rastgele etkiler olarak ele alınarak hatanın modellere dahil edildiği iki farklı model ile incelenmiştir. Hatanın dahil edildiği modeller bireylere ek olarak maddeler arasında da rastgele etkilerin tanımlandığı modellerdir (Desjardins ve Bulut, 2018; Isbell ve Son, 2023; Janssen, Shepers ve Peres, 2004). Süreç verileri ile kurulan modeller ve boş model ile elde edilen kestirimler ile standart hataları ve model uyum indisleri Tablo 6’da verilmiştir.

Tablo 6

Süreç Verileri ile Kurulan Modellerin Sonuçları

	M0	M11	M12	M21	M22	M31	M32	M41	M42
Sabit Etkiler									
Tepki Süresi		0.00 (0.00)	0.18*** (0.05)					0.12*** (0.03)	0.07 (0.03)
Eylem Sayısı				0.05*** (0.01)	0.56*** (0.04)			0.07** (0.02)	0.57*** (0.04)
Görüntüleme Sayısı						-0.03 (0.04)	-0.03 (0.22)	-1.49*** (0.33)	-2.07*** (0.50)
						-0.08 (0.07)	-0.16 (0.23)	-1.60*** (0.34)	-2.39*** (0.51)
Rastgele Etkiler									
Birey σ^2	0.96 (0.98)	0.68 (0.83)	0.90 (0.95)	0.69 (0.83)	0.86 (0.93)	0.68 (0.83)	0.96 (0.98)	0.64 (0.80)	0.85 (0.92)
Madde σ^2			4.70 (2.17)		2.83 (1.68)		0.71 (0.84)		1.25 (1.12)
Model Karşılaştırması									
AIC	11638.6	12865.6	11677.3	12850.3	11460.1	12866.2	11697.4	12814.5	11432.7
BIC	11760.7	12879.9	11698.8	12864.6	11481.6	12887.7	11726.1	12850.4	11475.7
Sapma	11604.6	12861.6	11671.3	12846.3	11454.1	12860.2	11689.4	12804.5	11420.7
R^2_{marjinal}	.144	.000	.003	.001	.068	.000	.000	.008	.098
$R^2_{\text{koşullu}}$	.338	.171	.631	.175	.561	.171	.337	.168	.449

* p< .05, ** p< .01, *** p< .001 (Not: Parantez içindeki değerler standart hatayı göstermektedir.)

Tablo 6’da verilen tüm modellere ait kestirimler, ilgili modellere ilişkin bulguların sunumunda alt başlıklar halinde aşağıda yorumlanmıştır.

Tepki Süresinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular

Tepki süresi tek yordayıcı olarak ele alınarak, hatanın dahil edilmediği yalnızca bireyler için rastgele etkilerin ele alındığı M11 modeli ve hatanın dahil edildiği bireyler ve maddeler için rastgele etkilerin ele alındığı M12 modeli kurulmuştur. M11 modeli, yalnızca bireyler arasındaki rastgele etkileri ve tepki süresinin sabit etkisini içermektedir. M12 modeline ise birey varyansına ek olarak maddeler arası rastgele etkiler de dahil edilmiştir.

Model uyumunu değerlendirmek için AIC, BIC ve Sapma değerleri kullanılmıştır; daha düşük değerler, daha iyi uyumu işaret etmektedir. M11 ve M12 modelleri karşılaştırıldığında, M12 modeli AIC, BIC ve Sapma değerleri açısından daha iyi bir uyum sağlamaktadır. Ayrıca, maddeler arası rastgele etkilerin dahil edilmesi tepki süresi ile açıklanan varyansı ve modelin toplam açıklayıcı gücünü arttırmıştır ($M11$, $R^2_{marjinal} = .000$ ve $R^2_{koşullu} = .171$; $M12$, $R^2_{marjinal} = .003$ ve $R^2_{koşullu} = .631$). Ancak, her iki modelin de boş modele kıyasla uyum açısından iyileşme göstermediği anlaşılmaktadır.

M11 modelinde tepki süresinin sabit etkisi lojit ölçeğinde 0.00 olarak tahmin edilmiş ve standart hatası 0.00 bulunmuştur. Bu sonuca göre, tepki süresinin maddelerin doğru yanıtlanma olasılığı üzerinde anlamlı bir etkisinin olmadığı görülmüştür ($p = .988$). M12 modeli ile yapılan kestirimler sonucunda ise tepki süresinin sabit etkisi lojit ölçeğinde 0.18 ve standart hatası 0.05 olarak tahmin edilmiş ve bu etki istatistiksel olarak anlamlı bulunmuştur ($p < 0.001$). Tepki süresinin doğru yanıt verme olasılığı üzerinde anlamlı ve pozitif bir etkisi vardır. Lojit ölçeğindeki 0.18 katsayısı, 1.20 ($\exp(0.18)$) odds oranına karşılık gelmektedir. Bu, tepki süresi bir standart sapma yüksek olan bir öğrencinin bir maddeyi doğru yanıtlama olasılığının, aynı maddeyi doğru yanıtlama olasılığı 0.50 olan bir öğrenciye göre 1.20 kat daha fazla olduğunu göstermektedir. Bu durumda, bu öğrencinin aynı maddeyi doğru yanıtlama olasılığı yaklaşık 0.55’tir. Bu sonuç, maddeleri yanıtlamak için daha fazla süre harcayan öğrencilerin maddeleri doğru yanıtlama olasılığının daha yüksek olduğunu ortaya koymaktadır.

M12 modeli, bireyler ve maddeler arasındaki rastgele varyansı tahmin etme açısından daha karmaşık bir yapıya sahiptir. Maddeler rastgele etkiler olarak modele dahil edildiğinde, açıklayıcı değişken olarak tepki süresinin etkisi kontrol edilirken, madde güçlük parametrelerinin ve örtük özelliklerin rastgele etkiler olarak değişmesine izin verilmiştir. M12 modelinin daha iyi uyum sağlamış olması, maddeler arasındaki varyansın modelde sabit bir parametre yerine bir dağılım üzerinden tahmin edilmesi gerektiğini göstermektedir.

M11 modelinde birey varyansı lojit ölçeğinde 0.68 olarak tahmin edilmiş ve standart hatası 0.83'tür. Bu varyans, tepki süresi değişkeninin modele dahil edilmesinden sonra bireyler arasında açıklanamayan kalan varyansı temsil etmektedir. Bu sonucun boş model ile kestirilen birey varyansından daha küçük olması, tepki süresinin etkisi çıkarıldıktan sonra kalan varyansı göstermesinden kaynaklanmaktadır.

M12 modelinde birey varyansı lojit ölçeğinde 0.90 olarak tahmin edilmiş ve standart hatası 0.95'tir. Maddeler arası rastgele etkilerin modele dahil edilmesi, birey varyansını geniş bir ölçekte tahmin etmeyi sağlamış ve birey varyansı M11 modeline kıyasla artış göstermiştir. Bu artış, bireyler arası farklılıkların daha belirgin hale geldiğini göstermektedir. Diğer yandan maddeler arası varyans lojit ölçeğinde 4.70 olarak tahmin edilmiş ve standart hatası 2.17'dir. Bu, maddeler arasında tepki süresiyle açıklanamayan önemli farklılıklar olduğunu göstermektedir. Maddeler arasındaki bu yüksek varyans, maddelerin zorluk seviyesi, yapısal özellikleri veya başka gizli faktörlerden kaynaklanabilir. Maddelerin zorluk seviyeleri ve yapısal özelliklerinin doğru yanıt olasılığı üzerinde önemli bir etkisi olabileceği düşünülmektedir.

Eylem Sayısının Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular

Eylem sayısı tek yordayıcı olarak ele alınarak, hatanın dahil edilmediği yalnızca bireyler için rastgele etkilerin ele alındığı M21 modeli ve hatanın dahil edildiği bireyler ve maddeler için rastgele etkilerin ele alındığı M22 modeli kurulmuştur. M21 modeli, yalnızca bireyler arasındaki rastgele etkileri ve eylem sayısının sabit etkisini içermektedir. M21 modeline ise birey varyansına ek olarak maddeler arası rastgele etkiler de dahil edilmiştir.

M21 ve M22 modelleri karşılaştırıldığında, M22 modeli ile elde edilen AIC, BIC ve Sapma değerleri daha düşüktür ve daha iyi bir uyum sağlamaktadır. Ayrıca, maddeler

arası rastgele etkilerin dahil edilmesi eylem sayısı ile açıklanan varyansı ve modelin toplam açıklayıcı gücünü arttırmıştır ($M21, R^2_{\text{marjinal}} = .001$ ve $R^2_{\text{koşullu}} = .175$; $M22, R^2_{\text{marjinal}} = .068$ ve $R^2_{\text{koşullu}} = .561$). Bunun yansısı, M22 modeli boş modele kıyasla da her üç uyum iyiliği ölçütü açısından daha iyi uyum göstermektedir.

M21 modelinde eylem sayısının sabit etkisi lojit ölçeğinde 0.05 olarak tahmin edilmiş ve standart hatası 0.01 bulunmuştur. Bu sonuca göre, eylem sayısının maddelerin doğru yanıtlanma olasılığı üzerinde anlamlı bir etkisi vardır ($p < .001$). Benzer şekilde M22 modeli ile yapılan kestirimler sonucunda da eylem sayısının sabit etkisi lojit ölçeğinde 0.56 ve standart hatası 0.04 olarak tahmin edilmiş ve bu etki istatistiksel olarak anlamlı bulunmuştur ($p < .001$). Her iki modele göre eylem sayısının doğru yanıt verme olasılığı üzerinde anlamlı ve pozitif bir etkisi vardır. Lojit ölçeğindeki 0.56 katsayısı, 1.75 ($\exp(0.18)$) odds oranına karşılık gelmektedir. Bu, eylem sayısı bir standart sapma yüksek olan bir öğrencinin bir maddeyi doğru yanıtlanma olasılığının, aynı maddeyi doğru yanıtlanma olasılığı 0.50 olan bir öğrenciye göre 1.75 kat daha fazla olduğunu göstermektedir. Bu durumda, bu öğrencinin aynı maddeyi doğru yanıtlanma olasılığı yaklaşık 0.64'tür. Bu sonuç, maddeleri yanıtlamak için daha fazla eylemde bulunan öğrencilerin maddeleri doğru yanıtlanma olasılığının daha yüksek olduğunu ortaya koymaktadır.

M22 modeli, bireyler ve maddeler arasındaki rastgele varyansı tahmin etme açısından daha karmaşık bir yapıya sahiptir. Maddeler rastgele etkiler olarak modele dahil edildiğinde, açıklayıcı değişken olarak eylem sayısının etkisi kontrol edilirken, madde güçlük parametrelerinin ve örtük özelliklerin rastgele etkiler olarak değişmesine izin verilmiştir. M22 modelinin daha iyi uyum sağlamış olması, maddeler arasındaki varyansın modelde sabit bir parametre yerine bir dağılım üzerinden tahmin edilmesi gerektiğini göstermektedir.

M21 modelinde birey varyansı lojit ölçeğinde 0.69 olarak tahmin edilmiş ve standart hatası 0.83'tür. Bu varyans, eylem sayısının modele dahil edilmesinden sonra bireyler arasında açıklanamayan kalan varyansı temsil etmektedir. Bu sonucun boş model ile kestirilen birey varyansından daha küçük olması, eylem sayısının etkisi çıkarıldıktan sonra kalan varyansı göstermesinden kaynaklanmaktadır.

M22 modelinde birey varyansı lojit ölçeğinde 0.86 olarak tahmin edilmiş ve standart hatası 0.93'tür. Maddeler arası rastgele etkilerin modele dahil edilmesi, birey varyansını geniş bir ölçekte tahmin etmeyi sağlamış ve birey varyansı M21 modeline

kıyasla artış göstermiştir. Bu artış, bireyler arası farklılıkların daha belirgin hale geldiğini göstermektedir. Diğer yandan maddeler arası varyans lojit ölçeğinde 2.83 olarak tahmin edilmiş ve standart hatası 1.68'dir. Bu, maddeler arasında eylem sayısı ile açıklanamayan önemli farklılıklar olduğunu göstermektedir. Maddeler arasındaki bu yüksek varyans, maddelerin zorluk seviyesi, yapısal özellikleri veya başka gizil özelliklerden kaynaklanabilir. Maddelerin zorluk düzeyleri ve yapısal özelliklerinin doğru yanıt olasılığı üzerinde önemli bir etkisi olabileceği düşünülmektedir.

Görüntüleme Sayısının Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular

Görüntüleme sayısı tek yordayıcı olarak ele alınarak, hatanın dahil edilmediği yalnızca bireyler için rastgele etkilerin ele alındığı M31 modeli ve hatanın dahil edildiği bireyler ve maddeler için rastgele etkilerin ele alındığı M32 modeli kurulmuştur. M31 modeli, yalnızca bireyler arasındaki rastgele etkileri ve görüntüleme sayısının sabit etkisini içermektedir. M32 modeline ise birey varyansına ek olarak maddeler arası rastgele etkiler de dahil edilmiştir.

Model uyumu açısından M31 ve M32 modelleri karşılaştırıldığında, M32 modeli AIC, BIC ve Sapma değerleri açısından daha iyi bir uyum sağlamaktadır. Ayrıca, maddeler arası rastgele etkilerin dahil edilmesi modelin toplam açıklayıcı gücünü arttırmıştır $(M31, R^2_{\text{marjinal}} = .000 \text{ ve } R^2_{\text{koşullu}} = .171; M32, R^2_{\text{marjinal}} = .000 \text{ ve } R^2_{\text{koşullu}} = .337)$. Ancak, her iki modelin de boş modele kıyasla uyum açısından iyileşme göstermediği anlaşılmaktadır.

Görüntüleme sayısının sabit etkisi her iki modelde bir kez ve birden çok görüntüleme kategorilerinin her ikisi için kestirilmiştir. Bir kez görüntülemenin sabit etkisi M31 modelinde lojit ölçeğinde -0.03 olarak tahmin edilmiş ve standart hatası 0.04; M32 modelinde lojit ölçeğinde -0.03 olarak tahmin edilmiş ve standart hatası 0.22'dir. Bu sonuca göre, bir kez görüntülemenin maddelerin doğru yanıtlanma olasılığı üzerinde anlamlı bir etkisinin olmadığı görülmüştür ($p = .428, p = .908$). Birden çok görüntülemenin sabit etkisi M31 modelinde lojit ölçeğinde -0.08 olarak tahmin edilmiş ve standart hatası 0.07; M32 modelinde lojit ölçeğinde -0.16 olarak tahmin edilmiş ve standart hatası 0.23'tür. Bu sonuca göre, birden çok görüntülemenin maddelerin doğru yanıtlanma olasılığı üzerinde anlamlı bir etkisinin olmadığı görülmüştür ($p = .281, p = .481$). Bir kez

ve birden çok görüntüleme kategorilerinin hiçbirinin doğru yanıt olasılığı üzerinde istatistiksel olarak anlamlı bir etkisi bulunmamıştır. Bu durum, görüntüleme sayısının doğru yanıt olasılığını açıklamak için yeterli bir değişken olmadığını göstermektedir.

M31 modelinde birey varyansı lojit ölçeğinde 0.68 olarak tahmin edilmiş ve standart hatası 0.83'tür. Bu varyans, görüntüleme sayısının modele dahil edilmesinden sonra bireyler arasında açıklanamayan kalan varyansı temsil etmektedir. Bu sonucun boş model ile kestirilen birey varyansından daha küçük olması, görüntüleme sayısının etkisi çıkarıldıktan sonra kalan varyansı göstermesinden kaynaklanmaktadır.

M32 modelinde birey varyansı lojit ölçeğinde 0.96 olarak tahmin edilmiş ve standart hatası 0.98'dir. Maddeler arası rastgele etkilerin modele dahil edilmesi, birey varyansını geniş bir ölçekte tahmin etmeyi sağlamış ve birey varyansı M31 modeline kıyasla artış göstermiştir. Bu artış, bireyler arası farklılıkların daha belirgin hale geldiğini göstermektedir. Diğer yandan maddeler arası varyans lojit ölçeğinde 0.71 olarak tahmin edilmiş ve standart hatası 0.84'tür. Bu sonuç görüntüleme sayısının etkisinin manidar olmadığı bulgusuyla birlikte ele alındığında daha anlamlı bir şekilde yorumlanabilir. Bu durumda maddelerin görüntüleme sayısı açısından birbirine benzer davrandığına işaret etmektedir.

Tepki Süresinin, Eylem Sayısının ve Görüntüleme Sayısının Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular

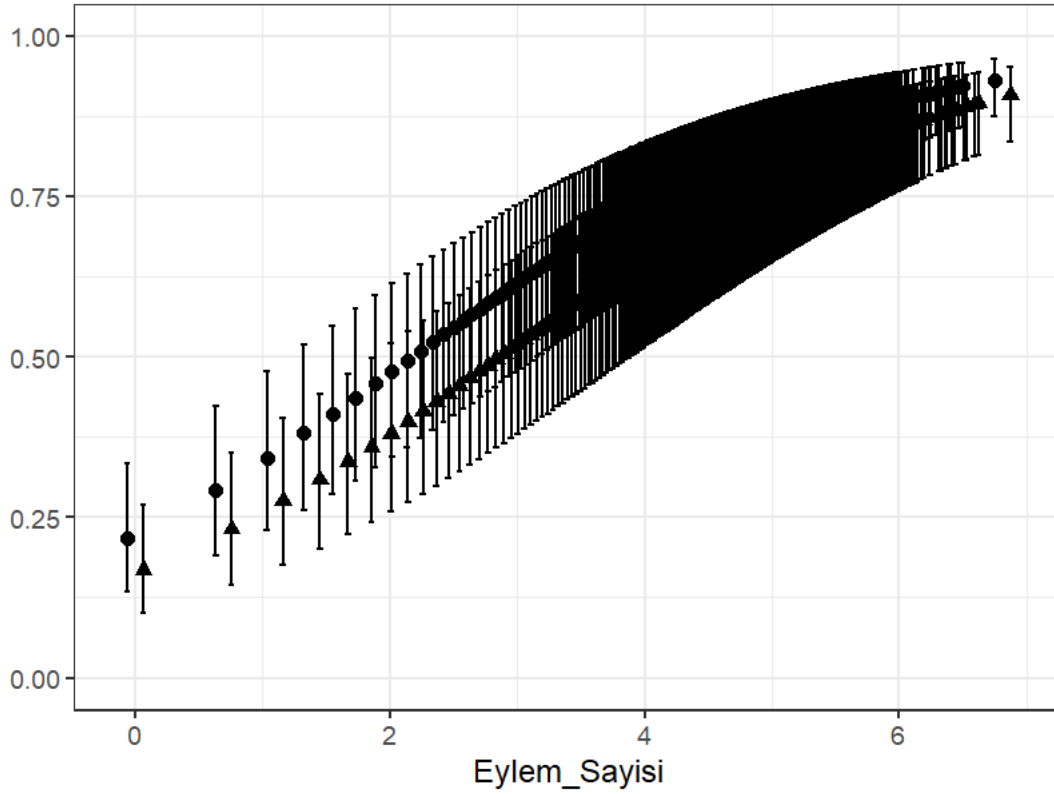
Tepki süresi, eylem sayısı ve görüntüleme sayısı yordayıcı olarak ele alınarak, hatanın dahil edilmediği yalnızca bireyler için rastgele etkilerin ele alındığı M41 modeli ve hatanın dahil edildiği bireyler ve maddeler için rastgele etkilerin ele alındığı M42 modeli kurulmuştur. M41 modeli, yalnızca bireyler arasındaki rastgele etkileri ve açıklayıcı değişkenlerin sabit etkisini içermektedir. M42 modeline ise birey varyansına ek olarak maddeler arası rastgele etkiler de dahil edilmiştir.

M41 ve M42 modelleri karşılaştırıldığında, M42 modeli ile elde edilen AIC, BIC ve Sapma değerleri daha düşüktür ve daha iyi bir uyum sağlamaktadır. Ayrıca, maddeler arası rastgele etkilerin dahil edilmesi, tepki süresi, eylem ve görüntüleme sayısı ile açıklanan varyansı (sabit etkilerin açıkladığı varyansı) ve modelin toplam açıklayıcı gücünü arttırmıştır ($M41, R^2_{\text{marjinal}} = .008$ ve $R^2_{\text{koşullu}} = .168$; $M42, R^2_{\text{marjinal}} = .098$ ve $R^2_{\text{koşullu}} = .449$). Bunun yansıya, M42 modeli boş modele kıyasla da her üç uyum

iyiliği ölçütü açısından daha iyi uyum göstermektedir. Buna ek olarak, M42 modeli kurulan 9 model arasında uyum iyiliği değerleri en düşük olan, en iyi uyum sağlayan modeldir.

M41 modelinde tüm açıklayıcı değişkenlerin etkisi istatistiksel olarak manidardır ($p < 0.001$). Lojit ölçeğinde sırasıyla tepki süresinin, eylem sayısının, bir kez görüntüleme sayısının ve birden çok görüntüleme sayısının etkisi ve standart hataları 0.12 (0.03), 0.07 (0.02), -1.49 (0.33); -1.60 (0.34) olarak tahmin edilmiştir. Bu sonuca göre, tepki süresi ve eylem sayısı arttıkça maddelerin doğru yanıtlanma olasılığı artmakta, görüntüleme sayısı arttıkça maddelerin doğru yanıtlanma olasılığı azalmaktadır. Tepki sayısının etkisi 1.13, eylem sayısının etkisi 1.07, bir kez görüntülemenin etkisi 0.23, birden çok görüntülemenin etkisi 0.20 odds oranına karşılık gelmektedir.

M42 modelinde tepki süresinin etkisi istatistiksel olarak manidar değildir ($p = 0.065$). Lojit ölçeğinde sırasıyla tepki süresinin, eylem sayısının, bir kez görüntüleme sayısının ve birden çok görüntüleme sayısının etkisi ve standart hataları 0.07 (0.03), 0.57 (0.04), -2.07 (0.50); -2.39 (0.51) olarak tahmin edilmiştir. Bu sonuca göre, eylem sayısı arttıkça maddelerin doğru yanıtlanma olasılığı artmakta, görüntüleme sayısı arttıkça maddelerin doğru yanıtlanma olasılığı azalmaktadır. Tepki sayısının etkisi 1.07, eylem sayısının etkisi 1.77, bir kez görüntüleme etkisi 0.13, birden çok görüntüleme etkisi 0.09 odds oranına karşılık gelmektedir. M42 modeline göre manidar bulunan açıklayıcı değişkenler için çizilen marjinal etki grafiği aşağıda Şekil 6'da verilmiştir.



Şekil 6. M42 Modelinde Manidar Bulunan Açıklayıcı Değişkenler İçin Marjinal Etki Grafiği (●= Bir Kez Görüntülenen Maddeler, ▲= Birden Çok Görüntülenen Maddeler)

Şekil 6’da eylem sayısı X ekseninde, kestirilen doğru yanıtlanma olasılığı Y ekseninde verilmiştir. Grafikte maddelerin doğru yanıtlanma olasılığının eylem sayısına bağlı olarak görüntüleme sayısına göre nasıl değiştiği görülmektedir. Bar çubukları üzerindeki noktalar farklı görüntüleme sayıları için kestirilen parametrelere ve çubukların uzunlukları %95 güven aralığına karşılık gelmektedir. Marjinal etki grafiği, eylem sayısı ve görüntüleme sayısı arasındaki etkileşim nedeniyle maddelerin kestirilen doğru yanıtlanma olasılığını göstermektedir. Grafiğe göre eylem sayısı arttıkça görüntüleme sayısının her iki kategorisi içinde doğru yanıtlanma olasılığının arttığı, kestirilen parametrelerin güven aralığının azaldığı ve birden fazla görüntüleme sayısı için doğru yanıtlanma olasılığının daha düşük olduğu görülmektedir.

M42 modeli, bireyler ve maddeler arasındaki rastgele varyansı tahmin etme açısından daha karmaşık bir yapıya sahiptir. Maddeler rastgele etkiler olarak modele dahil edildiğinde, açıklayıcı değişkenlerin etkisi kontrol edilirken, madde güçlük parametrelerinin ve örtük özelliklerin rastgele etkiler olarak değişmesine izin verilmiştir. M42 modelinin daha iyi uyum sağlamış olması, maddeler arasındaki varyansın modelde

sabit bir parametre yerine bir dağılım üzerinden tahmin edilmesi gerektiğini göstermektedir.

M41 modelinde birey varyansı lojit ölçeğinde 0.64 olarak tahmin edilmiş ve standart hatası 0.80'dir. Bu varyans, açıklayıcı değişkenlerin modele dahil edilmesinden sonra bireyler arasında açıklanamayan kalan varyansı temsil etmektedir. Bu sonucun boş model ile kestirilen birey varyansından daha küçük olması, açıklayıcı değişkenlerin etkisi çıkarıldıktan sonra kalan varyansı göstermesinden kaynaklanmaktadır.

M42 modelinde birey varyansı lojit ölçeğinde 0.85 olarak tahmin edilmiş ve standart hatası 0.92'tir. Maddeler arası rastgele etkilerin modele dahil edilmesi, birey varyansını geniş bir ölçekte tahmin etmeyi sağlamış ve birey varyansı M41 modeline kıyasla artış göstermiştir. Bu artış, bireyler arası farklılıkların daha belirgin hale geldiğini göstermektedir. Diğer yandan maddeler arası varyans lojit ölçeğinde 1.25 olarak tahmin edilmiş ve standart hatası 1.12'dir. Bu, maddeler arasında açıklayıcı değişkenlerle açıklanamayan önemli farklılıklar olduğunu göstermektedir. Maddeler arasındaki bu yüksek varyans, maddelerin zorluk seviyesi, yapısal özellikleri veya başka gizil özelliklerden kaynaklanabilir. Maddelerin zorluk seviyeleri ve yapısal özelliklerinin doğru yanıtlanma olasılığı üzerinde önemli bir etkisi olabileceği düşünülmektedir.

Madde Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular

Madde özelliklerinin maddelerin doğru yanıtlanma olasılıkları üzerindeki etkilerinin kestirilmesi amacıyla madde özellikleri boş modele eklenerek Doğrusal Lojistik Test Modeli (DLTM) kurulmuştur. Ardından maddelerin kestirimlerde rastgele etkiler olarak ele alınmasına olanak sağlayan hatanın dahil edildiği Doğrusal Lojistik Test Modeli (DLTM_E) kurularak maddelerden kaynaklanan varyans açıklanmaya çalışılmıştır (Desjardins ve Bulut, 2018). DLTM_E, normal DLTM ile regresyona dayalı mükemmel bir tahminin aksine kusurlu bir tahmine izin vermektedir (De Boeck ve Wilson, 2004). Hatanın dahil edildiği DLTM_E modelinde maddeler rastgele etkiler olarak modele dahil edilmektedir. Bu modeller ile her bir maddenin güçlük parametresi ayrı ayrı kestirilmek yerine madde gruplarının güçlük parametreleri kestirilmektedir. DLTM_E, açıklayıcı değişkenlerin etkisini kontrol ederken madde güçlük parametrelerinin ve örtük özellik tahminlerinin rastgele etkiler olarak değişmesine izin vermektedir. DLTM_E, maddeler arası değişkenliğini hesaba kattığından ve bu yolla

daha doğru parametre kestirimlerine izin verdiğiinden madde yordayıcılarının etkilerini incelemek için daha iyi bir seçenek gibi görünmektedir (Desjardins ve Bulut, 2018). Madde özelliklerinin maddelerin doğru yanıtlanma olasılığına etkilerini ortaya çıkarmak amacıyla kurulan modellere ilişkin kestirimler ile standart hataları ve model uyum indisleri aşağıda Tablo 7’de verilmiştir.

Tablo 7

Madde Özellikleriyle Kurulan Modellerin Sonuçları

	M0	DLTM	DLTM_E
Sabit Etkiler			
Madde Zorluk Düzeyi			
Düzy 3		-0.68*** (0.08)	-0.68** (0.22)
Düzy 4		-2.08*** (0.12)	-2.10*** (0.30)
Düzy 5		-1.96*** (0.10)	-1.99*** (0.27)
Madde Türü (Çoktan Seçmeli)		0.20 (0.11)	0.20 (0.28)
Bağlam			
Kamusal		-0.49*** (0.14)	-0.48 (0.39)
Kişisel		-0.17* (0.08)	-0.17 (0.23)
Bilişsel Boyut			
Bütünleştirme ve Yorumlama		-0.54*** (0.07)	-0.55** (0.19)
Yansıtma ve Değerlendirme		-1.02*** (0.10)	-1.03*** (0.26)
Metin Biçimi (Sürekli)		1.08*** (0.12)	1.10** (0.33)
Rastgele Etkiler			
Birey σ^2	0.96 (0.98)	0.93 (0.97)	0.96 (0.98)
Madde σ^2			0.05 (0.23)
Model Karşılaştırması			
AIC	11638.6	11738.3	11677.8
BIC	11760.7	11817.3	11764.0
Sapma	11604.6	11716.3	11653.8
$R^2_{marjinal}$	.144	.129	.131
$R^2_{koşullu}$	.338	.322	.335

* $p < .05$, ** $p < .01$, *** $p < .001$ (Not: Parantez içindeki değerler standart hatayı göstermektedir.)

DLTM ve DLTM_E modelleri karşılaştırıldığında, DLTM_E modeli ile elde edilen AIC, BIC ve Sapma değerleri daha düşüktür ve bu model daha iyi uyum

sağlamaktadır. Ayrıca, maddeler arası rastgele etkilerin dahil edilmesi, madde özellikleri ile açıklanan varyansı ve modelin toplam açıklayıcı gücünü arttırmıştır ($DLTM$, $R^2_{marjinal} = .129$ ve $R^2_{koşullu} = .322$; $DLTM_E$, $R^2_{marjinal} = .131$ ve $R^2_{koşullu} = .335$). Ancak, her iki modelin de boş modele kıyasla uyum ve açıklanan varyans oranları açısından bir iyileşme göstermediği anlaşılmaktadır.

Bu modeller ile kestirilen sabit etki değerleri birbirine yakın olmakla birlikte istatistiksel manidarlıkları modeller arasında farklılık göstermektedir. $DLTM$ ile madde türü dışındaki tüm değişkenlerin, $DLTM_E$ ile bağlam ve madde türü dışındaki diğer değişkenlerin etkilerinin manidar olduğu görülmektedir. $DLTM_E$ modeli daha iyi uyum gösterdiğinden madde özelliklerinin etkileri bu model sonuçlarına dayalı olarak yorumlanmıştır.

Madde zorluk düzeyi değişkeninin maddelerin doğru yanıtlanması üzerindeki etkisi üçüncü, dördüncü ve beşinci zorluk düzeyleri için tahmin edilmiştir. Buna göre zorluk düzeyi Düzey 3 ($\beta_i = -0.68(0.22)$, $p < .01$) Düzey 4 ($\beta_i = -2.10(0.30)$, $p < .001$) ve Düzey 5 ($\beta_i = -1.99(0.27)$, $p < .001$) olan maddelerin doğru yanıtlanma olasılığı, referans kategori olarak ele alınan Düzey 2 maddelerine göre daha düşüktür. Maddelerin güçlük düzeyleri arttıkça doğru yanıtlanma olasılıklarının düştüğü söylenebilir. Bununla birlikte beklenin aksine zorluk düzeyi 5 olan maddelerin doğru yanıtlanma olasılığının, zorluk düzeyi 4 olan maddelerden daha yüksek olması dikkat çekici bir bulgudur.

Madde türü değişkeninin maddelerin doğru yanıtlanması üzerindeki etkisi çoktan seçmeli maddeler için tahmin edilmiştir. Buna göre çoktan seçmeli maddelerin ($\beta_i = 0.20(0.28)$, $p = .486$) doğru yanıtlanma olasılığı, referans kategori olarak ele alınan açık uçlu maddelerine göre daha yüksek gibi görünse de bu sonuç istatistiksel olarak manidar değildir.

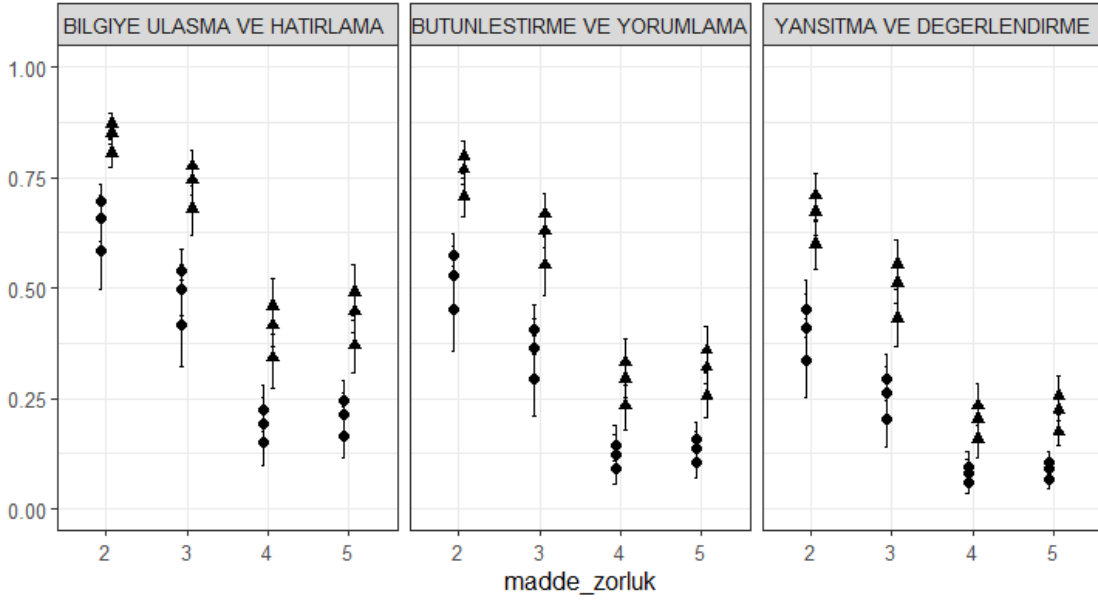
Madde bağlamı değişkeni için kestirilen etkiler her iki modelle de birbirine yakın sonuçlar vermekle birlikte bu etkilerin istatistiksel olarak manidarlığı açısından modeller arasında farklılık söz konusudur. $DLTM_E$ modelinin sonuçları referans olarak ele alınan eğitimsel bağlam içeren maddelere göre, kamusal bağlam içeren maddeler ($\beta_i = -0.48(0.39)$, $p = .222$) ile kişisel bağlam içeren maddelerin ($\beta_i = -0.17(0.23)$, $p = .456$) doğru yanıtlanma olasılığının daha düşük görünse de bu sonuç istatistiksel olarak manidar değildir. Diğer bir deyişle maddeler rastgele etkiler olarak ele alındığında bağlamın maddeler arasındaki etkisinin manidar olmadığı görülmektedir. Eğitimsel bağlam içeren maddelere kıyasla kamusal ya da kişisel bağlam içeren maddelerin güçlükleri arasında

manidar bir fark yoktur. Baęlam deęiřkeni maddelerin doęru yanıtlanma olasılıęı üzerinde etkili deęildir.

Maddenin biliřsel boyutunun maddelerin doęru yanıtlanmasına etkisi, her iki model ile birbirine yakın olarak kestirilmiş olup bu etkiler istatistiksel olarak manidardır. DLTM_E modelinin sonuçları, referans olarak ele alınan bilgiye ulaşma boyutundaki maddelere göre bütünleřtirme ve yorumlama ($\beta_i = -0.55(0.17)$, $p < .01$) ile yansıtma ve deęerlendirme ($\beta_i = -1.03(0.26)$, $p < .001$) boyutlarındaki maddelerin doęru yanıtlanma olasılıęının daha düşük olduęuna iřaret etmektedir. Dięer bir deyiřle bilgiye ulaşma ve hatırlama boyutundaki maddelere göre, bütünleřtirme ve yorumlama boyutundaki maddelerin doęru yanıtlanma olasılıęı 0.58 ($\exp(-0.55)$) kat, yansıtma ve deęerlendirme boyutundaki maddelerin doęru yanıtlanma olasılıęı 0.36 ($\exp(-0.55)$) kat daha düşüktür.

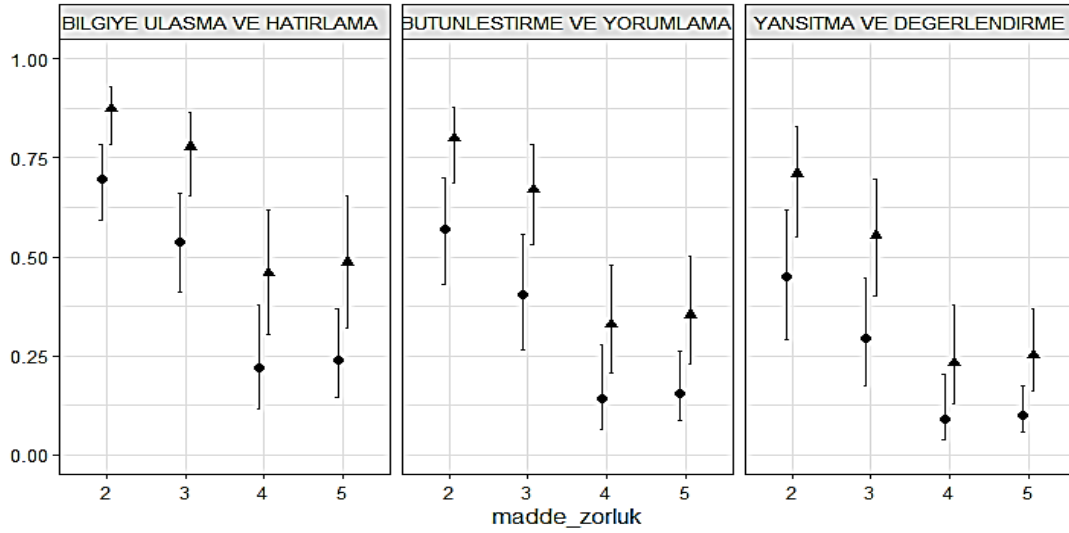
Metin biçiminin maddelerin doęru yanıtlanma olasılıęı üzerinde etkisi her iki model ile birbirine yakın olarak kestirilmiş olup bu etkiler istatistiksel olarak manidardır ($p < .01$). DLTM_E modelinin sonuçları, referans olarak ele alınan dięer metin biçimlerindeki (karma, çoklu, sürekli olmayan) maddelere göre sürekli metin biçimindeki ($\beta_i = 1.10(0.33)$, $p < .01$) maddelerin doęru yanıtlanma olasılıęının daha yüksek olduęuna iřaret etmektedir. Bu sonuç dięer metin biçimindeki maddelere göre sürekli metin biçimindeki maddelerin doęru yanıtlanma olasılıęının 3.00 ($\exp(1.10)$) kat fazla olduęunu göstermektedir.

DLTM ve DLTM_E ile modellenen ve etkileri istatistiksel olarak manidar çıkan madde yordayıcıları için çizilen marjinal etki grafikleri ařaęıda řekil 7 ve řekil 8’de verilmiřtir.



Şekil 7. DLTM İle Manidar Bulunan Açıklayıcı Değişkenler İçin Marjinal Etki Grafiği (▲= Sürekli Metin İçeren Maddeler, ●= Diğer Metin Biçimlerini İçeren Maddeler)

Şekil 7’de madde bilişsel boyutlarının her biri için 3 ayrı grafik yer almaktadır. Madde güçlük düzeyi X ekseninde, kestirilen doğru yanıtlanma olasılığı Y ekseninde verilmiştir. Grafikte maddelerin doğru yanıtlanma olasılığının madde özelliklerine bağlı olarak her bir metin biçimine göre nasıl değiştiği görülmektedir. Bar çubukları üzerindeki noktalar farklı bağlamlar için kestirilen parametrelere ve çubukların uzunlukları %95 güven aralığına karşılık gelmektedir. Marjinal etki grafiği, madde zorluk düzeyi, metin biçimi, bağlam ve bilişsel boyut arasındaki etkileşim nedeniyle maddelerin kestirilen doğru yanıtlanma olasılığını göstermektedir. Bağlam değişkeninin farklı kategorileri (sırasıyla eğitimsel, kişisel ve kamusal) aynı bar çubuğu üzerinde üç farklı nokta ile temsil edilmektedir. Grafiğe göre doğru yanıtlanma olasılığı en düşük (yaklaşık 0.025) olan maddeler zorluk düzeyi 4, metin biçimi diğer, bilişsel boyutu yansıtma ve değerlendirme ve bağlamı kamusal olan maddelerdir. Diğer yandan doğru yanıtlanma olasılığı en yüksek olan maddelerin (yaklaşık 0.88) zorluk düzeyi 2, metin biçimi sürekli, bilişsel boyutu bilgiye ulaşma ve hatırlama ve bağlamı eğitimsel olan maddeler olduğu görülmektedir.



Şekil 8. DLMT_E İle Manidar Bulunan Açıklayıcı Değişkenler İçin Marjinal Etki Grafiği (▲ = Sürekli Metin İçeren Maddeler, ● = Diğer Metin Biçimlerini İçeren Maddeler)

Şekil 8 incelendiğinde madde zorluk düzeyi, metin biçimi ve bilişsel boyut değişkenlerinin maddelerin doğru yanıtlanma olasılığına etkisi görülmektedir. Grafiğe göre doğru yanıtlanma olasılığı en düşük (yaklaşık 0.09) olan maddeler zorluk düzeyi 4, metin biçimi diğer, bilişsel boyutu yansıtma ve değerlendirme olan maddelerdir. Diğer yandan doğru yanıtlanma olasılığı en yüksek olan maddelerin (yaklaşık 0.86) zorluk düzeyi 2, metin biçimi sürekli, bilişsel boyutu bilgiye ulaşma ve hatırlama olan maddeler olduğu görülmektedir.

DLTM ve DLTM_E modelleri rastgele etkiler olan birey ve madde varyansı açısından incelendiğinde bireyler arasındaki varyans sırasıyla 0.97 hata ile 0.93 ve 0.98 hata ile 0.96 olarak tahmin edilmiştir. Okuma becerileri açısından bireyler arasındaki değişkenlik DLTM modeli ile görece daha düşük hesaplanmış olmakla birlikte bu varyansın boş model ile kestirilen birey varyansına çok yakın olduğu görülmektedir. Maddeye özgü özellikler modele dahil edildiğinde, bireyler arasında açıklanamayan varyansta kayda değer bir değişim olmadığı, bireylerin okuma becerileri arasındaki değişkenliği açıklamada madde özelliklerinin büyük bir etkisinin olmadığı söylenebilir. Diğer yandan DLTM_E ile maddeler arasındaki kestirilen varyans 0.05 olup, modele dahil edilen madde yordayıcılarının maddeler arasındaki varyansın büyük kısmını açıkladığı söylenebilir.

Birey Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular

Birey özelliklerinin maddelerin doğru yanıtlanma olasılıkları üzerindeki etkilerinin kestirilmesi amacıyla birey özellikleri boş modele eklenerek Örtük Regresyon Rasch Modeli (ÖRRM) kurulmuştur. Ardından maddelerin kestirimlerde rastgele etkiler olarak ele alınmasına olanak sağlayan hatanın dahil edildiği Örtük Regresyon Rasch Modeli (ÖRRM_E) kurularak maddelerden kaynaklanan varyans açıklanmıştır. ÖRRM modeli, yalnızca bireyler arasındaki rastgele etkileri ve açıklayıcı değişkenlerin sabit etkisini içermektedir. ÖRRM_E modeline ise birey varyansına ek olarak maddeler arası rastgele etkiler dahil edilmiştir. ÖRRM, okuma becerileriyle ilgili olarak bireyler arasında farklılıkların açıklanması için birey özelliklerini içermektedir (Wilson, De Boeck ve Carstensen, 2008). Birey özelliklerinden cinsiyet ve evdeki kitap sayısı değişkenleri modellere dahil edilerek maddelerin doğru yanıtlanma olasılığına etkileri incelenmiştir. Tablo 8’de kurulan modellere ilişkin kestirimler ile standart hataları ve model uyum indisleri sunulmuştur.

Tablo 8

Birey Özellikleriyle Kurulan Modellerin Sonuçları

	M0	ÖRRM	ÖRRM E
Sabit Etkiler			
Cinsiyet (Erkek)		-0.31*** (0.07)	-0.35*** (0.09)
Evdeki Kitap Sayısı			
11-25 arası		0.30** (0.10)	0.35** (0.12)
26-100 arası		0.70*** (0.10)	0.81*** (0.12)
101-200 arası		0.85*** (0.14)	0.98*** (0.16)
200-500 arası		1.20*** (0.17)	1.39*** (0.20)
500’den çok		1.18*** (0.21)	1.36*** (0.24)
Rastgele Etkiler			
Birey σ^2	0.96 (0.98)	0.51 (0.71)	0.73 (0.86)
Madde σ^2			0.70 (0.84)
Model Karşılaştırması			
AIC	11638.6	12758.5	11592.4
BIC	11760.7	12815.9	11657.1
Sapma	11604.6	12742.5	11574.4
$R^2_{marjinal}$	.144	.043	.046
$R^2_{koşullu}$	.338	.171	.336

* $p < .05$, ** $p < .01$, *** $p < .001$ (Not: Parantez içindeki değerler standart hatayı göstermektedir.)

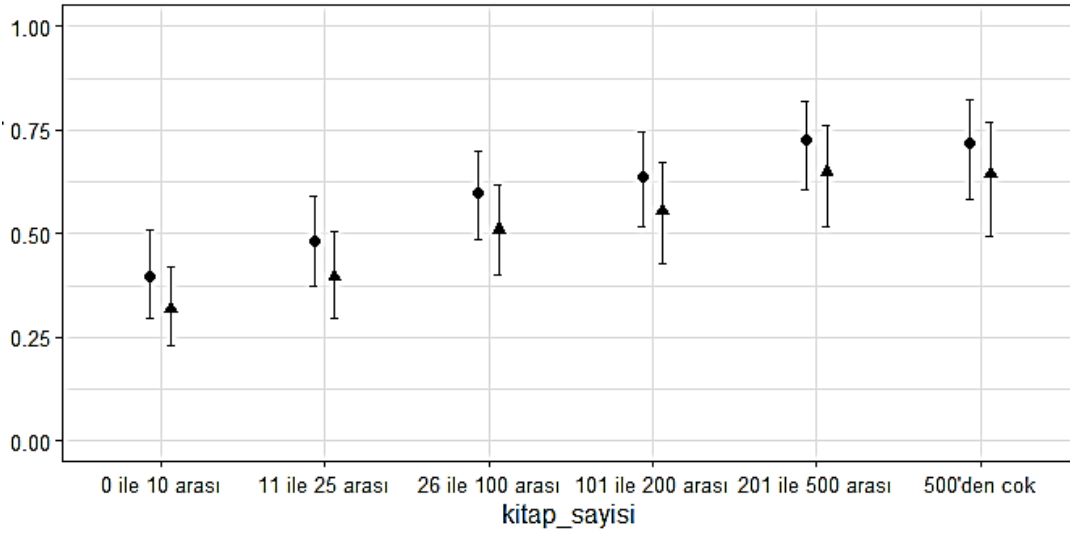
ÖRRM ve ÖRRM_E modelleri karşılaştırıldığında, ÖRRM_E modeli ile elde edilen AIC, BIC ve Sapma değerleri daha düşüktür ve bu model daha iyi uyum sağlamaktadır. Ayrıca, maddeler arası rastgele etkilerin dahil edilmesi, birey özellikleri tarafından açıklanan varyansı ve modelin toplam açıklayıcı gücünü arttırmıştır ($DLTM$, $R^2_{marjinal} = .043$ ve $R^2_{koşullu} = .171$; $DLTM_E$, $R^2_{marjinal} = .046$ ve $R^2_{koşullu} = .336$). Bunun yansıra, ÖRRM_E modeli boş modele kıyasla her üç uyum iyiliği ölçütü açısından da daha iyi uyum göstermektedir. Maddeye özgü açıklanmayan varyansın hesaba katılması, açıklayıcı değişkenlerin etkilerinin daha iyi bir tahminini sağlamaktadır. Bu sonuç cinsiyetin ve evdeki kitap sayısının etkilerinin maddeler arasında değişkenlik gösterdiğine işaret etmektedir.

Her iki model ile kestirilen sabit etki değerleri birbirine yakın olmakla birlikte istatistiksel manidarlık düzeyleri aynıdır. ÖRRM_E modeli daha iyi uyum gösterdiğinden ve bulguların gereksiz yere tekrarından kaçınmak için birey özelliklerinin etkileri bu model sonuçlarına dayalı olarak yorumlanmıştır.

Cinsiyetin maddelerin doğru yanıtlanması üzerindeki etkisi kızlar referans alınarak erkekler için tahmin edilmiştir. Buna göre erkek öğrencilerin maddeleri doğru yanıtlama olasılıkları ($\beta_i = -0.35(0.09)$, $p < .001$) kız öğrencilere göre daha düşüktür. Sonuçlar erkek öğrencilerin maddeleri doğru yanıtlama olasılığının kız öğrencilere göre 0.71 ($\exp(-0.35)$) kat daha düşük olduğuna işaret etmektedir. Buna göre, cinsiyetin erkek olmasının 0.50'lik bir olasılık üzerindeki etkisi, bu olasılığı 0.42'ye düşürmektedir. Diğer bir deyişle testteki herhangi bir maddeyi doğru yanıtlama olasılığı 0.50 olan bir kız öğrenciye göre bir erkek öğrencinin aynı maddeyi doğru yanıtlama olasılığı 0.42'dir.

Evdeki kitap sayısının maddelerin doğru yanıtlanması üzerindeki etkisi 0-10 arası kategorisi referans alınarak diğer kategoriler için tahmin edilmiştir. ÖRRM_E modelinin sonuçları, referans olarak ele alınan evindeki kitap sayısı 0 ile 10 arasında olan öğrencilere göre, 11-25 arasında ($\beta_i = 0.35(0.12)$, $p < .01$), 26- 100 arasında ($\beta_i = 0.81(0.12)$, $p < .001$), 101-200 arasında ($\beta_i = 0.98(0.16)$, $p < .01$) 201- 500 arasında ($\beta_i = 1.39(0.20)$, $p < .01$) ve 500'den çok ($\beta_i = 1.36(0.24)$, $p < .001$) olanların maddeleri doğru yanıtlanma olasılığının daha yüksek olduğuna işaret etmektedir. Evdeki kitap sayısı arttıkça maddelerin doğru yanıtlanma olasılığının arttığı söylenebilir. Evindeki kitap sayısı 0 ile 10 arasında olan öğrencilere kıyasla diğer kategorilerde kitap sayısına sahip olan öğrencilerin maddeleri doğru yanıtlama olasılığı istatistiksel olarak manidar artış göstermektedir. Maddeleri doğru yanıtlama olasılığı en yüksek olan öğrenciler evdeki kitap sayısı 200- 500 arasında olan öğrencilerdir ve maddeleri doğru yanıtlama olasılıkları

evdeki kitap sayısı 0-10 arasında olan öğrencilere göre 4.02 ($\exp(1.39)$) kat daha fazladır. Buna göre, evdeki kitap sayısının 200-500 arasında olmasının 0.50'lik bir olasılık üzerindeki etkisi, bu olasılığı 0.80'e çıkarmaktadır. Diğer bir deyişle evindeki kitap sayısı 0-10 arasında ve testteki herhangi bir maddeyi doğru yanıtlama olasılığı 0.50 olan bir öğrenciye göre evindeki kitap sayısı 200-500 arasında olan bir öğrencinin aynı maddeyi doğru yanıtlama olasılığı 0.80'dir. ÖRRM_E ile modellenen cinsiyet ve kitap sayısı için çizilen ve Şekil 9'da sunulan marjinal etki grafiğinde daha net bir şekilde görülebilir.



Şekil 9. ÖRRM_E Modelinde Cinsiyet ve Kitap Sayısı İçin Marjinal Etki Grafiği (▲ = Erkek, ● = Kız)

Şekil 9'da evdeki kitap sayısı X ekseninde, kestirilen bireylerin maddeleri doğru yanıtlama olasılığı Y ekseninde verilmiştir. Grafikte bireylerin maddeleri doğru yanıtlama olasılığının evdeki kitap sayısına bağlı olarak cinsiyete göre nasıl değiştiği görülmektedir. Bar çubukları üzerindeki noktalar kestirilen parametrelere ve çubukların uzunlukları %95 güven aralığına karşılık gelmektedir. Marjinal etki grafiği, evdeki kitap sayısı ve cinsiyet arasındaki etkileşim nedeniyle bireylerin maddeleri doğru yanıtlama olasılığını göstermektedir. Grafiğe göre maddeleri doğru yanıtlama olasılığı en yüksek olan bireyler, evdeki kitap sayısı 200'den fazla olan kız öğrencilerdir. Diğer yandan maddeleri doğru yanıtlama olasılığı en düşük olan bireylerin evdeki kitap sayısı 0 ile 10 arasında olan erkek öğrenciler olduğu ve evdeki kitap sayısı azaldıkça bireylerin maddeleri doğru yanıtlama olasılıklarının azaldığı görülmektedir.

ÖRRM_E modeli, bireyler ve maddeler arasındaki rastgele varyansı tahmin etme açısından daha karmaşık bir yapıya sahiptir. Maddeler rastgele etkiler olarak modele dahil edildiğinde, açıklayıcı değişkenlerin etkisi kontrol edilirken, madde güçlük parametrelerinin ve örtük özelliklerin rastgele etkiler olarak değişmesine izin verilmiştir. ÖRRM_E modelinin daha iyi uyum sağlamış olması, maddeler arasındaki varyansın modelde sabit bir parametre yerine bir dağılım üzerinden tahmin edilmesi gerektiğini göstermektedir.

ÖRRM modelinde birey varyansı lojit ölçeğinde 0.51 olarak tahmin edilmiş ve standart hatası 0.71'dir. ÖRRM_E modelinde birey varyansı lojit ölçeğinde 0.73 olarak tahmin edilmiş ve standart hatası 0.86'dır. Bu modellerle kestirilen birey varyansının Rasch model ile kestirilen birey varyansından daha düşük olması, cinsiyet ve evdeki kitap sayısı ile açıklanan varyansın çıkarılmasının bir sonucudur. Bireyler için kestirilen varyans cinsiyet ve evdeki kitap sayısı ile açıklanamayan bireysel farklılıklara işaret etmektedir ve boş modele göre daha düşük olması beklenir. Maddeler arası rastgele etkilerin modele dahil edilmesi, birey varyansını geniş bir ölçekte tahmin etmeyi sağlamış ve birey varyansı ÖRRM modeline kıyasla artış göstermiştir. Bu artış, bireyler arası farklılıkların daha belirgin hale geldiğini göstermektedir. Diğer yandan maddeler arası varyans lojit ölçeğinde 0.70 olarak tahmin edilmiş ve standart hatası 0.84'tür. Bu, maddeler arasında açıklayıcı değişkenlerle açıklanamayan önemli farklılıklar olduğunu göstermektedir.

Süreç Verilerinin, Madde ve Birey Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Bulgular

Madde, birey ve madde bazında birey özelliklerinin maddelerin doğru yanıtlanma olasılığına etkilerinin araştırıldığı dördüncü araştırma sorusuna ilişkin üç model kurulmuştur. Daha önce kurulan modellerde hatanın dahil edildiği modellerin daha iyi uyuma işaret ettiği görüldüğünden bu modellere maddelerin rastgele etkileri dahil edilmiştir. Birey ve madde özelliklerinin bir arada incelendiği modeller etkileşim modelleri ya da Örtük Regresyon Doğrusal Lojistik Test Modeli (ÖRDLTm) olarak adlandırılmaktadır.

Bu amaçla ilk olarak araştırmada ele alınan tüm birey ve madde değişkenleri boş modele eklenerek ÖRDLTm1 olarak adlandırılan ilk etkileşim modeli kurulmuştur. Diğer bir deyişle, cinsiyet, evdeki kitap sayısı, madde zorluk düzeyi, madde türü, bağlam,

bilişsel boyut ve metin biçimi değişkenleri modele eklenmiştir. Dördüncü araştırma sorusu kapsamında kurulan modellere ilişkin kestirimler ile standart hataları ve model uyum indisleri Tablo 9’da sunulmuştur.

Tablo 9

Etkileşim Modellerinin Sonuçları

	M0 Tahmin (SH)	ÖRDLT M1 Tahmin (SH)	ÖRDLT M2 Tahmin (SH)	ÖRDLT M3 Tahmin (SH)
Sabit Etkiler				
Madde Zorluk Düzeyi				
Düzy 3		-0.68** (0.22)	-1.05*** (0.21)	-1.12*** (0.18)
Düzy 4		-2.10*** (0.31)	-2.07*** (0.29)	-2.13*** (0.24)
Düzy 5		-1.99*** (0.27)	-2.45*** (0.26)	-2.52*** (0.25)
Madde Türü (Çoktan Seçmeli)		0.20 (0.28)	1.10*** (0.28)	1.11*** (0.19)
Bağlam				
Kamusal		-0.48 (0.39)	-0.18 (0.37)	
Kişisel		-0.17 (0.23)	-0.25 (0.22)	
Boyut				
Bütünleştirme ve Yorumlama		-0.55** (0.19)	-0.61*** (0.18)	-0.59** (0.19)
Yansıtma ve Değerlendirme		-1.03*** (0.26)	-0.98*** (0.25)	-0.95*** (0.26)
Metin Biçimi (Sürekli)		1.10** (0.33)	0.63* (0.32)	0.49** (0.18)
Cinsiyet (Erkek)		-0.35*** (0.09)	-0.36*** (0.08)	-0.36*** (0.08)
Evdeki Kitap Sayısı				
11-25 arası		0.35** (0.12)	0.33** (0.11)	0.33** (0.11)
26-100 arası		0.81*** (0.12)	0.73*** (0.11)	0.73*** (0.11)
101-200 arası		0.98*** (0.16)	0.93*** (0.16)	0.93*** (0.16)
200-500 arası		1.39*** (0.20)	1.29*** (0.19)	1.29*** (0.19)
500’den çok		1.36*** (0.24)	1.23*** (0.23)	1.23*** (0.23)
Tepki Süresi			0.09* (0.04)	0.09* (0.04)
Eylem Sayısı			0.57*** (0.04)	0.56*** (0.04)
Görüntüleme Sayısı			-0.23*** (0.05)	-0.23*** (0.05)
Rastgele Etkiler				
Birey σ_{0p}^2	0.96 (0.98)	0.73 (0.86)	0.65 (0.81)	0.65 (0.81)
Madde σ_{0i}^2		0.05 (0.23)	0.05 (0.22)	0.05 (0.23)
Model Karşılaştırması				
AIC	11638.6	11571.8	11302.3	11299.7
BIC	11760.7	11701.0	11453.1	11436.2
Sapma	11604.6	11535.8	11260.3	11261.7
$R^2_{marjinal}$	.144	.177	.226	.225
$R^2_{koşullu}$	.338	.335	.361	.362

* $p < .05$, ** $p < .01$, *** $p < .001$ (Not: Parantez içindeki değerler standart hatayı göstermektedir.)

İkinci aşamada ÖRDLTm1 modeline madde bazında birey değişkenleri olarak tepki süresi, eylem sayısı ve görüntüleme sayısı değişkenleri dahil edilerek ÖRDLTm2 modeli kurulmuştur. Bu modelle birey, madde ve madde bazında birey özellikleri bir arada ele alınarak, bu değişkenlerin bireylerin maddeleri doğru yanıtlama olasılığına etkisi incelenmektedir. ÖRDLTm2 modeli ile yapılan kestirimler sonucunda sabit etkisi istatistiksel olarak manidar bulunmayan bağlam değişkeni daha sade bir model elde etmek amacıyla çıkarılarak ÖRDLTm3 modeli oluşturulmuştur.

Tablo 9’da görülen uyum iyiliği indisleri temel alınarak modeller arası karşılaştırmalar yapıldığında, AIC ve BIC değerleri en düşük olan ÖRDLTm3 modelinin çalışmada kurulan tüm modeller arasında en iyi uyum gösteren model olduğu görülmektedir. Bu model ile elde edilen sapma değerinin ise ÖRDLTm2 ile elde edilen değerden az farkla daha büyük olduğu dikkat çekmektedir. Ayrıca, ÖRDLTm1 modeline süreç verilerinin eklenmesiyle kurulan ÖRDLTm2 modelinde yordayıcı değişkenler tarafından açıklanan varyansın (sabit etkilerin açıkladığı varyansı) ve modelin toplam açıklayıcı gücünün arttığı görülmektedir (ÖRDLTm1 , $R^2_{\text{marginal}} = .177$ ve $R^2_{\text{koşullu}} = .335$; ÖRDLTm2 , $R^2_{\text{marginal}} = .226$ ve $R^2_{\text{koşullu}} = .361$).

ÖRDLTm1 modeli ile kestirilen sabit etkiler, DLTM_E ve ÖRRM_E kestirilen sabit etkiler ve hataları ile tamamen aynı sonuçları vermektedir. Bu nedenle bu etkiler yeniden yorumlanmamıştır. Birey varyansı ÖRRM_E ile kestirilen birey varyansına, madde varyansı ise DLTM_E ile kestirilen madde varyansına eşittir.

ÖRDLTm2 modeli ile kestirilen sabit etkiler incelendiğinde bağlam değişkeni dışındaki tüm yordayıcıların etkilerinin istatistiksel olarak manidar olduğu görülmüştür. Modele eklenen tepki süresi değişkeninin etkisi 0.04 hata ile lojit ölçeğinde 0.09 olarak kestirilmiş olup, tepki süresi daha uzun olan bireylerin maddeleri doğru yanıtlama olasılığının daha yüksek olduğuna işaret etmektedir ($p < .05$). Benzer şekilde eylem sayısı etkisi 0.04 hata ile lojit ölçeğinde 0.57 olarak kestirilmiş olup, maddeleri yanıtlarken daha fazla eylemde bulunan bireylerin maddeleri doğru yanıtlama olasılığının daha yüksek olduğu görülmektedir ($p < .001$). Görüntüleme sayısının etkisi ise 0.05 hata ile lojit ölçeğinde -0.23 olarak kestirilmiştir ($p < .001$). Bu sonuç maddeleri birden çok görüntüleyen öğrencilerin maddeleri doğru yanıtlama olasılığının azaldığını ortaya koymaktadır.

Süreç verileri eklenmeden önce istatistiksel olarak manidar olmayan madde türü değişkeninin ($\beta_i = 1.10(0.28)$, $p < .001$) bu model ile hem manidar olarak kestirildiği hem

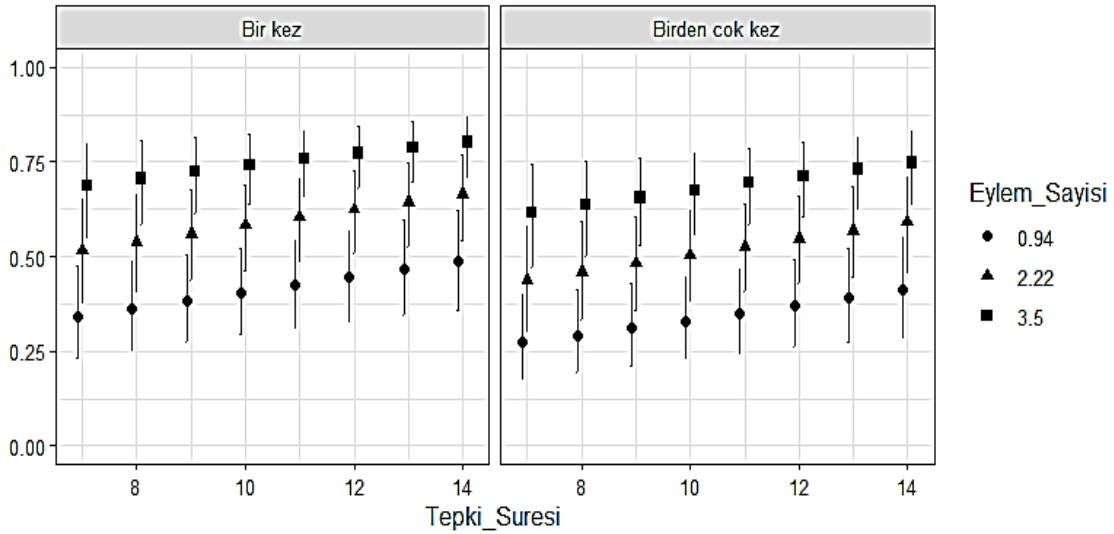
de etkinin büyüklüğünde bir artış olduğu dikkat çekmektedir. Bu sonuç, tepki süresi, eylem sayısı ve görüntüleme sayısı değişkenleri ile ele alındığında çoktan seçmeli maddelerin açık uçlu maddelere göre kolaylaştığına işaret etmektedir. Diğer bir deyişle tepki süresi, eylem sayısı, görüntüleme sayısı değişkenleri madde türünün etkisinin daha iyi ortaya çıkarılmasını sağlamıştır. Benzer bir iyileşmenin madde zorluk düzeylerinin etkileri için geçerli olduğu söylenebilir. Süreç verileri eklenmeden önce, zorluk düzeyi 4 olan maddelerin etkileri (-2.10), zorluk düzeyi 5 olan maddelerin etkilerinden (-1.99) daha düşük kestirilmiş olup doğru yanıtlanma olasılığının beklenenin aksine az da olsa daha düşük olduğuna işaret etmektedir. Süreç verilerinin eklendiği ÖRDLT2 modeli ise beklendiği gibi zorluk düzeyi 5 olan maddelerin etkilerinin (-2.40), zorluk düzeyi 4 olan maddelerin etkilerinden (-2.07) daha düşük olduğunu, dolayısıyla doğru yanıtlanma olasılığının daha az olduğunu ortaya koymaktadır. Bununla birlikte, kestirimlerin standart hataları incelendiğinde, ÖRDLT2 modeli ile yapılan kestirimlerin standart hatalarının daha düşük olduğu görülmektedir. Bu sonuçlar, birey ve madde özelliklerinin süreç verileri ile birlikte modellenmesinin modelin açıklama gücünü ve tutarlılığını artırma potansiyeline işaret etmektedir.

İstatistiksel olarak manidar etkisi olmayan bağlam değişkeninin atılmasıyla kurulan ÖRDLT3 modeli ile kestirilen etkiler incelendiğinde madde zorluk düzeyi, bilişsel boyut ve metin biçimi dışındaki değişkenlerin kestirimlerinde kayda değer bir farklılık olmadığı görülmektedir. Madde zorluk düzeylerinin etkileri artmış ve kestirimlerin hataları azalmıştır. Diğer yandan bilişsel boyutların etkileri azalmış ve kestirimlerin hataları artmıştır. Bu sonuçlar bağlamın etkisi çıkarıldığında bilişsel boyutlardaki etkinin bir kısmının madde zorluk düzeylerine kaymış olabileceğini düşündürmektedir. Ayrıca metin biçimi değişkeninin etkisi azalmış olmakla birlikte istatistiksel manidarlık düzeyi artmıştır ($\beta_i = 0.49(0.18)$, $p < .01$).

ÖRDLT2 ve ÖRDLT3 etkileşim modelleri ile birey varyansı 0.81 hata ile lojit ölçeğinde 0.65 olarak tahmin edilmiştir. Boş modele kıyasla birey varyansının daha düşük kestirildiği, bu modellere eklenen değişkenlerin bireylerin performansları arasındaki farklılıkların bir kısmını açıklayabildiği görülmektedir. Diğer yandan maddeler arası varyans ($\sigma^2_{0i} = 0.05$) çok daha az olduğu için maddelerin birbirine benzer bir şekilde işlev gördüğü söylenebilir.

ÖRDLT3 ile modellenen değişkenler arasındaki etkileşimler marjinal etki grafikleri yardımıyla incelenmiştir. Bu amaçla ilk olarak tepki süresi, görüntüleme sayısı

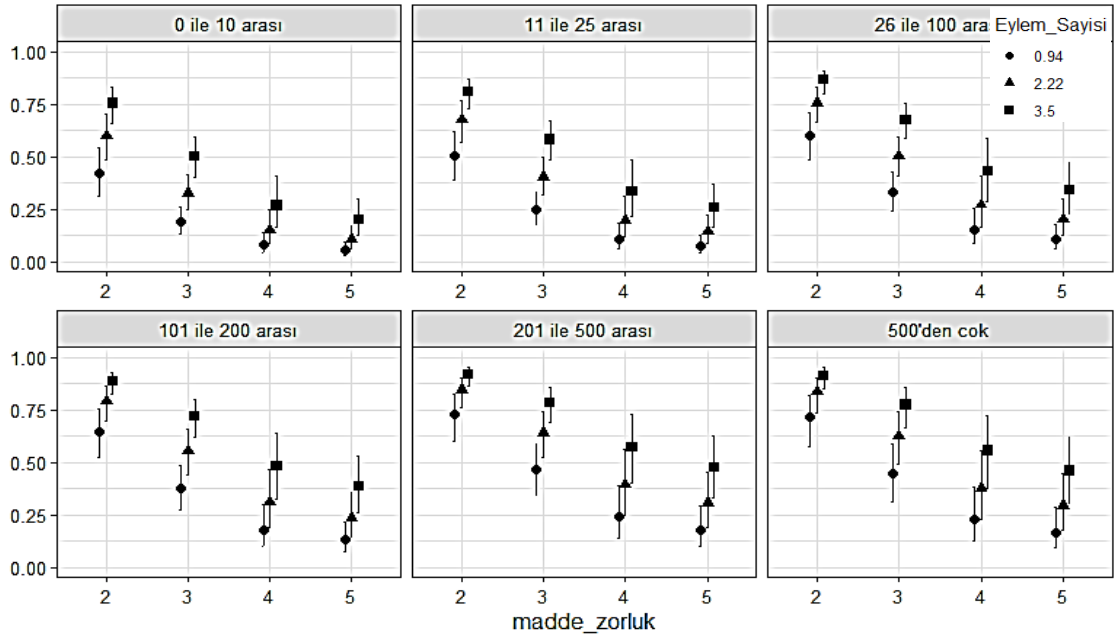
ve eylem sayısı değişkenleri ile doğru yanıtlanma olasılığı arasındaki ilişkiler için oluşturulan marjinal etki grafiği aşağıda Şekil 10’da sunulmuştur.



Şekil 10. ÖRDLTM3 Modelinde Tepki Süresi, Görüntüleme Sayısı ve Eylem Sayısı Değişkenleri için Marjinal Etki Grafiği

Şekil 10’da görüntüleme sayısının kategorileri için iki farklı grafik yer almaktadır. Tepki süresi X ekseninde, kestirilen doğru yanıtlanma olasılığı Y ekseninde verilmiştir. Grafikte maddelerin doğru yanıtlanma olasılığının tepki süresi ve görüntüleme sayısına bağlı olarak eylem sayısına göre nasıl değiştiği görülmektedir. Eylem sayısına ilişkin değerler, eylem sayısının ortalaması, ortalamanın bir standart sapma altı ve bir standart sapma üstüne eşittir. Bar çubukları üzerindeki noktalar bu eylem sayıları için kestirilen parametrelere ve çubukların uzunlukları %95 güven aralığına karşılık gelmektedir. Marjinal etki grafiği, tepki süresi, görüntüleme sayısı ve eylem sayısı arasındaki etkileşim nedeniyle maddelerin kestirilen doğru yanıtlanma olasılığını göstermektedir. Grafiğe göre, birden çok görüntülenene, az süre harcanan ve az sayıda eylemde bulunan maddelerin doğru yanıtlanma olasılıkları en düşüktür. Diğer yandan, çok süre harcanan, çok sayıda eylemde bulunan ve bir kez görüntülenene maddelerin doğru yanıtlanma olasılıkları en yüksektir.

Eylem sayısı ile güçlü etkilere sahip olan evdeki kitap sayısı ve madde güçlük düzeyi değişkenleri ve maddelerin doğru yanıtlanma olasılığı arasındaki ilişkiler için oluşturulan marjinal etki grafiği aşağıda Şekil 11’de verilmiştir.



Şekil 11. ÖRDLTM3 İle Modellenen Madde Zorluk Düzeyi, Evdeki Kitap Sayısı ve Eylem Sayısı Değişkenleri İçin Marjinal Etki Grafiği

Şekil 11’de evdeki kitap sayısının kategorileri için altı farklı grafik yer almaktadır. Madde zorluk düzeyi X ekseninde, kestirilen doğru yanıtlanma olasılığı Y ekseninde verilmiştir. Grafikte maddelerin doğru yanıtlanma olasılığının evdeki kitap sayısı ve madde zorluk düzeyine bağlı olarak eylem sayısına göre nasıl değiştiği görülmektedir. Marjinal etki grafiği, evdeki kitap sayısı, madde zorluk düzeyi ve eylem sayısı arasındaki etkileşim nedeniyle maddelerin kestirilen doğru yanıtlanma olasılığını göstermektedir. Grafığe göre, evdeki kitap sayısı fazla olan ve maddeler üzerinde daha çok eylemde bulunan bireylerin her güçlük düzeyi için maddeleri doğru yanıtlanma olasılığı daha yüksektir. Grafikte dikkat çekici bir bulgu olarak, evdeki kitap sayısı az olan bireylerin eylem sayılarındaki artış kolay maddelerin doğru yanıtlanma olasılığını daha geniş bir aralıkta etkilerken, evdeki kitap sayısı çok olan bireylerin eylem sayılarındaki artış zor maddelerin doğru yanıtlanma olasılığını daha geniş bir aralıkta etkilemektedir. Diğer bir ifadeyle, evdeki kitap sayısı az olan bireyler zorluk düzeyi yüksek maddeleri yanıtlarken ne kadar çok eylemde bulunsalar da doğru yanıtlayabilme olasılıkları düşük kalmaktadır. Diğer yandan evdeki kitap sayısı fazla olan bireylerin zorluk düzeyi yüksek olan maddeleri yanıtlarken eylem sayılarındaki artış doğru yanıtlanma olasılıklarını daha fazla arttırmaktadır. Maddelerin zorluk düzeyi arttıkça, evindeki kitap sayısı çok olanların ve maddelerin zorluk düzeyi azaldıkça evindeki kitap sayısı az olanların doğru yanıtlanma olasılığı eylem sayısından daha çok etkilenmektedir.

Araştırma soruları doğrultusunda elde edilen bulgular genel olarak değerlendirildiğinde, maddeler arası rastgele etkilerin modele dahil edilmesinin, model uyumunu ve modelin toplam açıklama gücünü arttırdığı görülmüştür. Rastgele hatanın dahil edildiği modeller, maddeye özgü açıklanamayan varyansı hesaba katarak ve sabit etkilerin daha iyi bir kestirimini sunmakta ve bu sabit etkilerin yeni maddelere genellenebilmesini sağlamaktadır (Isbell ve Son, 2023; Janssen vd., 2004).

Araştırma bulguları, süreç verilerinin etkilerinin maddeden maddeye farklılık gösterdiğini ortaya koymaktadır. Tepki süresinin etkisine yönelik bulgular, yalnızca maddeler arasındaki rastgele etkiler hesaba katıldığında, maddelerin doğru yanıtlanma olasılığı üzerinde pozitif bir etkisi olduğuna işaret etmektedir. Diğer bir deyişle tepki süresinin etkisinin maddeler arasında değişkenlik gösterdiği görülmüştür. Bu sonuç, tepki süresi ile yanıt doğruluğu arasındaki ilişkilerin maddelerin güçlük düzeylerine bağlı olarak değiştiğini ortaya koyan çalışmaların sonuçlarını desteklemektedir (Dodonova ve Dodonov, 2013; Goldhammer vd., 2014; Goldhammer, Naumann ve Greiff, 2015). Diğer yandan pozitif bir etkisi olduğuna yönelik bulgular, genellikle negatif bir ilişkiyi işaret eden çalışma sonuçları ile çelişmektedir (Dodonova ve Dodonov, 2013; Goldhammer vd., 2014, 2015; Su ve Davison, 2019). Bununla birlikte bu çalışmalardan bazıları (Dodonova ve Dodonov, 2013; Goldhammer vd., 2014, 2015), madde güçlüğü arttıkça negatif ilişkinin azaldığına ve hatta pozitifte dönebildiğine işaret etmektedir. Bu pozitif etkinin nedeni ikili işlem kuramına göre maddelerin kontrollü bir işlem gerektirmesinden (Goldhammer vd., 2014) ya da karmaşıklık düzeyi, içerdikleri metinlerin uzunluğu gibi farklı madde özelliklerinden kaynaklanıyor olabilir. Wu, Chen ve Stone (2018) öğrencilerin yanıtlama süreçlerinin, testle ölçülen özellikten çok bir reaksiyonu yansıttığını ve bu nedenle tepki süreleri ve diğer süreç verilerinin testin özelliklerine bağlı olduğunu belirtmektedir.

Araştırma bulguları, tepki süresinin etkisi çıkarıldıktan sonra bireyler arasında önemli ölçüde açıklanamayan varyans olduğuna işaret etmektedir. Bireyler arasında açıklanamayan varyans yetenek düzeylerinden kaynaklanıyor olabilir. Bu doğrultuda alanyazında tepki süresinin etkisinin bireylerin test ile ölçülmek istenen yetenek düzeylerine bağlı olarak değiştiğini ortaya koyan çalışmaların sonuçlarında tutarsızlık olduğu görülmektedir. Yetenek düzeyi yüksek öğrencilerin tepki sürelerinin daha kısa olduğuna işaret eden çalışmalar (Yalçın, 2022) olduğu gibi, yanıtın doğruluğundan ve madde güçlüklerinden bağımsız olarak maddeler üzerinde daha fazla süre harcadıklarını ve bu sonucun farklı cinsiyet ve kültürler arasında tutarlı olduğunu ileri süren (Chang vd.,

2005) alıřmalar da bulunmaktadır. Farklı kltrler ve kořullarda yařayan, hatta farklı okul trlerinde ğrenim gren ğrencilerin yanıtlama srelerinin ve bu yanıtlama sreleriyle iliřkili bireysel zelliklerin farklılık gstermesi beklenebilecek bir durumdur (Reis Costa ve Chen, 2023).

Arařtırmanın eylem sayısının etkisine ynelik bulguları, maddeler arası rastgele etkiler dahil edilerek modellendiğinde eylem sayısının modelin aıklama gcn arttırdığını, maddelerin doėru yanıtlanma olasılığı zerinde pozitif bir etkiye sahip olduğunu ve bu etkinin maddeler arasında deėiřkenlik gsterdiğini ortaya koymaktadır. Bu bulgular alanyazındaki alıřmaların bazı bulguları ile tutarlı iken bazı bulguları ile eliřkilidir. rneėin Yavuz (2019), daha ok sre harcayan ve daha fazla tepki davranıřı gsteren ğrencilerin maddeleri doėru yanıtlama olasılıėının daha yksek olduğunu ortaya koymuřtur. Ayrıca eylem sayısı ile okuma becerisi etkileřiminin maddelerin doėru yanıtlanması zerinde pozitif etkisi ve okuma becerisi yksek ğrencilerin daha fazla eylemde bulunma olasılığı olması bakımından paralellik gstermektedir. Diėer yandan bu etkinin maddeler ve bireyler arasında deėiřmediėi sonucu bu arařtırmanın bulguları ile eliřmektedir. řahin ve Covin (2020) ilgisiz test alma davranıřını belirlemede eylem sayılarının anlamlı katkılar saėlayabileceėini ve bu davranıřı gsteren ğrencilerin eylem sayısının daha az olduğunu belirtmektedir. İlgisiz test alma davranıřının yanıtın doėruluėu ile negatif iliřkisi gz nne alındığında, eylem sayısının az olmasının maddelerin doėru yanıtlanma olasılıėını da dolaylı olarak azalttığı sylenebilir. Bu bakımdan arařtırma sonuları paralellik gstermektedir. Diėer yandan Arenada vd. (2022), nispeten dřk ve yksek yeterliliėe sahip ğrencilerin orta yeterliliėe sahip ğrencilere gre daha az sayıda eylem sergilediklerine ynelik sonucu ile alıřma bulguları eliřmektedir. Ortaya ıkan bu eliřki, kullanılan modellerin doėrusal iliřkileri aıklayamamasından kaynaklanabileceėi gibi eylem sayısının ğrencilerin maddeleri yanıtlamaya ynelik biliřsel abalarının bir gstergesi olabileceėini dřndrmektedir.

Eylem sayısının etkisinin maddeler arasında deėiřkenlik gstermesi bulgusuyla uyumlu olarak, maddeleri doėru yanıtlama olasılığı en yksek olan stratejik profildeki ğrencilerin iki farklı grevden birinde ortalama dzeyde eylem sayısı ve tepki sresine sahip oldukları ancak biliřsel gereksinimi yksek olan diėer maddede daha uzun tepki sresi ve eylem sayısı sergiledikleri grlmřtr (Teig vd., 2020). Maddeleri doėru yanıtlama olasılığı en dřk olan ilgisiz profildeki ğrencilerin ise her iki madde iin dřk eylem sayısı ve tepki sresine sahip olduėu ortaya ıkmıřtır. Reis Costa ve Chen (2023), tepki sresi ve eylem sayısı ile baėlamsal deėiřkenler arasındaki iliřkilerin kltre

göre farklılık gösterdiğini ortaya koymuştur. Bu çalışmada eylem sayısının öğrencilerin bilişsel çabasının bir göstergesi ve tepki süresini de etkileyen ve açıklayan bir değişken olarak ele alınabileceği düşünülmektedir. Eylem sayısı arttıkça öğrencilerin tepki süreleri de doğal olarak artacağından, tepki süresindeki artışın pozitif etkisi eylem sayısı üzerinden açıklanabilir. Diğer yandan tepki süresi uzun olduğu halde, madde üzerinde bilişsel çaba harcamayan yeterince eylem sergilemeyen öğrencilerin maddeleri doğru yanıtlama olasılığının düşük olması beklenen bir sonuçtur. Bu öğrencilerin kopya çekmeye çalışması, ilgisiz test alma davranışı sergilemesi ya da dikkat ve odaklanma problemleri yaşaması gibi çeşitli açıklamalar muhtemeldir.

Bu çalışmanın bulgularına göre, görüntüleme sayısı açısından maddeler arası varyans düşüktür ve maddelerin doğru yanıtlanma olasılığını açıklamada tek başına yeterli bir değişken değildir. OECD (2024) raporunda da belirtildiği gibi öğrencilerin yalnızca %7'si maddeleri birden çok kez görüntülemiştir. Bu sonuç araştırmada incelenen verilerin düşük riskli ölçme uygulamalarından elde edilmiş olmasından kaynaklanabilir. Bu nedenle görüntüleme sayısına ilişkin sonuçların dikkatli bir şekilde yorumlanması gerekmektedir. Ortak etkiler modellendiğinde birden çok görüntülemenin maddenin doğru yanıtlama olasılığını düşürdüğü görülmektedir. Bu bulgu görüntüleme sayısının ele alındığı bazı çalışmaların sonuçları ile benzerlik göstermektedir. Örneğin, Lahza ve diğerleri (2022), toplam görüntüleme sayısının madde güçlüğü ile yüksek korelasyon gösterdiğini, düşük performanslı öğrencilerin maddeleri daha fazla görüntülediklerini belirtmiştir. Dolayısıyla maddelerin görüntülenme sayısındaki artışın doğru yanıtlanma olasılığını düşürdüğü ve maddeler arasında değişkenlik gösterdiği bulgusu ile benzerlik göstermektedir. Guo ve diğerleri (2024) görüntüleme sayısı fazla olan öğrencilerin, tepki sürelerinin düzensiz ve / veya kısa, test puanlarının düşük ya da çok düşük olduğu sonucuna ulaşmıştır. Örneğin, düşük performans gösteren öğrenciler arasında, soruları defalarca görüntüleyen ancak etkili yanıt veremeyen profiller tespit edilmiştir. Bu bakımdan birden çok görüntüleme sayısının doğru yanıt olasılığını düşürmesi bulgusu ile paralellik göstermektedir. Diğer yandan zor maddelerin, öğrenciler tarafından daha az gözden geçirildiğine ve daha az değişiklik yapıldığına ilişkin araştırma sonuçları (Zhang vd., 2023) bu araştırma sonucu ile çelişmektedir. Zhang ve diğerleri (2023), maddeleri görüntüleme sayılarının madde türüne göre değişiklik gösterdiğini, çoktan seçmeli maddelerin doğru yanlış türündeki maddelere göre daha fazla görüntülendiğini ortaya koymuştur. Görüntüleme sayılarının öğrencilerin test puanını doğrudan etkilediği sonucuna ulaşılmıştır. Ancak maddelerin görüntülenme sayıları ile testten alınan toplam

puan, ya da yetenek düzeyi arasındaki ilişkiler ile bir maddenin görüntülenme sayısının doğru yanıtlanma olasılığı üzerindeki etkisinin birbirine karıştırılmaması gerekir. Bilişsel yük kuramına göre, öğrencilerin doğru yanıtı kolayca ulaşamayacaklarını düşündükleri maddeleri atlayarak test sonunda geri dönmeleri ve yeniden görüntülemeleri test performanslarının artmasına katkı sağlayabilir. Bu süreçte öğrenciler öncelikle doğru yanıtı kolayca ulaşabilecekleri maddeleri yanıtlayıp daha sonra zorlandıkları maddeler üzerinde çalışarak bilişsel yükü yönetebilirler. Ancak bu durum öğrencilerin zorlandıkları maddelere geri döndüklerinde doğru yanıtlanma olasılıklarının arttığı anlamına gelmez. Öğrenci kendisini doğru yanıtı ulaştıracak bilgi ve yeteneğe sahip olmadığı süreçte maddeyi yeniden görüntülese de doğru yanıtı ulaşamayabilir. Bununla birlikte yüksek görüntüleme sayılarının, düzensiz veya stratejik olmayan test yaklaşımları ile zaman yönetimi ve dikkat sorunlarına işaret edebileceği (Guo vd., 2024) ya da öğrencilerin verdikleri yanıtlara güvenmemelerinden kaynaklanabileceği (Lahza vd., 2022) ileri sürülmüştür.

Madde özelliklerinin okuma becerileri maddelerinin doğru yanıtlanma olasılığı üzerindeki etkileri incelendiğinde, zorluk düzeyinin, madde türünün, bilişsel boyutun ve metin biçiminin manidar etkileri olduğu sonucuna ulaşılmıştır. Maddelerin zorluk düzeyi arttıkça doğru yanıtlanma olasılıkları beklenildiği gibi azalmaktadır. Bu değişkenin bir yordayıcı olarak ele alınmasının nedeni, araştırmada ele alınan veri setinden değil, farklı örneklemelere ait daha büyük veri setleri üzerinden kestirilmiş olmasıdır. Madde zorluk düzeyi bir nevi kontrol değişkeni olarak ele alınarak modellere dahil edilmiştir. Yalnızca madde özelliklerinin incelendiği DLTM_E modeli sonuçlarına göre madde zorluğunun doğru yanıtlanma olasılığına etkisi dördüncü düzeye kadar artmakta, dört ve beşinci düzeyler arasında etkinin beklenenin aksine azaldığı görülmektedir. Nitekim, süreç, madde ve birey parametrelerinin etkilerinin birlikte kestirildiği ÖRDLT3 modeliyle elde edilen sonuçlar, maddelerin güçlük düzeyleri arttıkça doğru yanıtlanma olasılıklarının azaldığını doğrulamaktadır.

Maddelerin doğru yanıtlanma olasılığı üzerinde (yalnızca süreç, madde ve birey özelliklerinin etkilerinin birlikte kestirildiği modellerde) etkili olan bir diğer değişken madde türüdür ve çoktan seçmeli maddelerin açık uçlu maddelere göre doğru yanıtlanma olasılığı daha yüksektir. Bu sonuç alanyazındaki çalışmalarla tutarlıdır (Bulut vd., 2022; Schulz-Heidorf ve Støle, 2018; Yousefpoori-Naeim vd., 2023).

Maddenin gerektirdiği bilişsel boyutun maddelerin doğru yanıtlanma olasılığı üzerinde anlamlı bir etkisi vardır. Maddelerin çözümü için gerekli olan bilişsel

süreçlerdeki karmaşıklık arttıkça, maddelerin doğru yanıtlanma olasılığının düştüğü görülmektedir. Genel olarak, öğrencilerin sadece metin içinde bilgi bulmalarını gerektiren maddeler en kolay, bilgiyi yansıtmayı ve değerlendirmeyi gerektiren maddeler en zor ve metni anlamaya yönelik maddeler orta zorluk seviyesinde yer almaktadır (Yousefpoori-Naeim vd., 2023). Bu bulgu, PISA 2015 değerlendirme çerçevesinde vurgulandığı üzere (OECD, 2017a) bilişsel karmaşıklık ve madde güçlüğü arasında beklenen ilişkiyle tutarlıdır. Ayrıca bu bulgu, daha önce yapılan araştırmalarda ileri sürülen üst düzey bilişsel süreçlere yönelik maddelerin daha zor olduğu bulgusunu desteklemektedir (Bulut vd., 2022).

Maddenin içerdiği metin biçiminin maddelerin doğru yanıtlanma olasılığı üzerinde manidar bir etkisi olduğu, sürekli metin biçimine sahip maddelerin doğru yanıtlanma olasılığının daha yüksek olduğu sonucuna ulaşılmıştır. Bu sonuç, önceki araştırmaların öğrencilerin karma ve sürekli olmayan metin biçimlerine göre sürekli metinlerde önemli ölçüde daha iyi performans gösterdiği bulguları (Hódi vd., 2017; Hódi vd., 2025) ile uyumludur. Bu sonucun muhtemel bir nedeni, öğrencilerin sürekli metin biçimindeki maddelere aşinalığı olabilir.

Evdeki kitap sayısı ve cinsiyetin maddelerin doğru yanıtlanma olasılığı üzerinde manidar etkileri olduğu görülmüştür. Bu sonuç alanyazındaki çalışmaların sonuçlarıyla örtüşmektedir (Aydoğan ve Gelbal, 2022; Gülleroğlu, Demir ve Demirtaşlı, 2014; Kutlu vd., 2011). Aydoğan ve Gelbal (2022) öğrencilerin okuma becerilerini yordayan en etkili değişkenin evdeki kitap sayısı olduğu ve cinsiyetin de manidar bir etkisi olduğu sonucuna ulaşmıştır. Kız öğrencilerin maddeleri doğru yanıtlanma olasılıkları erkek öğrencilere göre daha yüksektir (Bulut vd., 2022; Kutlu vd., 2011; Yousefpoori-Naeim vd., 2023). Evdeki kitap sayısı arttıkça öğrencilerin maddeleri doğru yanıtlanma olasılıklarının arttığı görülmüştür. Bu sonuç alanyazındaki çalışmaların sonuçlarıyla örtüşmektedir (Gülleroğlu, Demir ve Demirtaşlı, 2014; Kutlu vd., 2011).

Süreç verilerinin, yanıt doğruluğu, madde ve birey özellikleri ile modellenmesi boş modele göre veri uyumunu ve modelin açıklama gücünü arttırmış ve madde özelliklerinin etkilerinin daha tutarlı kestirimine anlamlı katkılar sağlamıştır. Bu bulgu daha önceki çalışmalarda belirtilen süreç verilerinin analizlere dahil edilmesinin öğrenci profillerinin belirlenmesine anlamlı katkılar sağladığı bulgularını desteklemektedir (Guo vd., 2024; Teig vd., 2020).

BÖLÜM 4

SONUÇLAR VE ÖNERİLER

Bu bölümde araştırmada bulgulara dayalı olarak elde edilen sonuçlar özetlenerek bu sonuçlar doğrultusunda gelecekteki araştırmalara ve uygulamaya dönük önerilere yer verilmiştir.

Sonuçlar

Bu çalışmada süreç verileri, birey ve madde özelliklerinin öğrencilerin okuma becerileri maddelerini doğru yanıtlama olasılıkları üzerindeki etkileri AMTM kapsamında ele alınarak modellenmiştir. Araştırma soruları doğrultusunda modellerin uyumu, modellere eklenen açıklayıcı değişkenlerin etkileri, kestirilen birey ve madde varyansları yorumlanarak maddelerin doğru yanıtlanma olasılığı üzerinde etkili olan değişkenler belirlenmiştir. Araştırma sonuçları süreç verileri ile okuma maddelerinin doğru yanıtlanma olasılıkları arasındaki manidar ilişkiler olduğunu ve okuma becerileri performansının açıklanmasında süreç verilerinin önemli bir katkı sağladığını göstermektedir. Bununla birlikte, tüm araştırma soruları genelindeki ortak sonuç, maddeler arası rastgele etkilerin modele dahil edilmesinin model uyumunu ve modelin toplam açıklayıcı gücünü arttırdığını göstermektedir. Bu sonuç, maddeler arasındaki varyansın modelde sabit bir parametre yerine bir dağılım üzerinden tahmin edilmesinin faydalı olacağına işaret etmektedir. Boş modele ilişkin sonuçlar ise, testin orta güçlükte bir test olduğu ve öğrencilerin örtük özellik aralığı göz önüne alındığında testin güçlük aralığının yeterli olduğunu göstermektedir. Diğer sonuçlar ise araştırma soruları bağlamında özetlenerek maddeler halinde sunulmuştur.

Süreç Verilerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Sonuçlar

1. Tepki süresi, eylem sayısı ve görüntüleme sayısı tek açıklayıcı değişken olarak ele alındığında,

- Maddeler arasında tepki süresi, eylem sayısı ya da görüntüleme sayısı ile açıklanamayan önemli farklılıklar vardır ve etkileri maddeden maddeye değişiklik göstermektedir.
 - Süreç verileri arasında okuma maddelerinin doğru yanıtlanma olasılığını en çok etkileyen değişken eylem sayısıdır, eylem sayısı arttıkça doğru yanıtlanma olasılığı artmaktadır.
 - Boş modele kıyasla uyum açısından anlamlı bir iyileştirme gösteren tek model eylem sayısının yordayıcı değişken olarak ele alındığı maddeler arası rastgele etkileri içeren modeldir.
 - Maddeler görüntüleme sayısı açısından birbirine benzerlik göstermekte olup, görüntüleme sayısı doğru yanıt olasılığını açıklamak için tek başına yeterli bir değişken değildir.
 - Maddeler arası rastgele etkiler modele dahil edildiğinde tepki süresinin maddelerin doğru yanıtlanma olasılığı üzerinde pozitif ve manidar bir etkisi olduğu görülmektedir.
2. Tepki süresi, eylem sayısı ve görüntüleme sayısı birlikte modellendiğinde,
- Açıklanamayan birey ve madde varyansı ile kestirim hataları azalmıştır.
 - Maddeler arası rastgele etkilerin dahil edildiği model, boş modele kıyasla uyum açısından anlamlı bir iyileştirme göstermektedir.
 - Görüntüleme sayısının maddelerin doğru yanıtlanma olasılığı üzerinde negatif ve manidar bir etkisi vardır. Maddelerin birden fazla görüntülenmesi durumunda doğru yanıtlanma olasılığının daha düşük olduğu görülmüştür.
 - Tepki süresinin maddelerin doğru yanıtlanma olasılığı üzerinde manidar bir etkisi yoktur.

Madde Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Sonuçlar

- Madde özelliklerinin modele dahil edilmesinin boş modele kıyasla uyum açısından anlamlı bir iyileştirme göstermediği, bireylerin okuma becerileri arasındaki değişkenliği açıklamada madde özelliklerinin büyük bir etkisinin olmadığı görülmüştür.

- Maddelerin doğru yanıtlanma olasılığı üzerinde zorluk düzeyi, bilişsel boyut ve metin biçimi değişkenleri manidar bir etkiye sahiptir.
- Maddelerin doğru yanıtlanma olasılığı madde zorluğunun dördüncü düzeyine kadar azalmaktadır.
- Bilgiye ulaşma boyutundaki maddelere göre bütünleştirme ve yorumlama ile yansıtma ve değerlendirme boyutlarındaki maddelerin doğru yanıtlanma olasılığı daha düşüktür.
- Diğer metin biçimlerindeki (karma, çoklu, sürekli olmayan) maddelere göre sürekli metin biçimindeki maddelerin doğru yanıtlanma olasılığı daha yüksektir.

Birey Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Sonuçlar

- Birey özelliklerinin modele dahil edilmesinin boş modele kıyasla uyum açısından iyileşme gösterdiği, bireylerin okuma becerileri arasındaki değişkenliği açıklamada birey özelliklerinin önemli bir etkisi olduğu görülmektedir.
- Erkek öğrencilerin maddeleri doğru yanıtlanma olasılıkları kız öğrencilere göre daha düşüktür.
- Evindeki kitap sayısı fazla olan öğrencilerin maddeleri doğru yanıtlanma olasılığının daha yüksek olduğu görülmüştür.

Süreç Verilerinin, Madde ve Birey Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığı Üzerindeki Etkisine Yönelik Sonuçlar

- Süreç verileri ile madde ve birey özelliklerinin birlikte modellenmesi boş modele kıyasla uyum açısından iyileşme göstermektedir.
- Madde bağlamının maddelerin doğru yanıtlanma olasılığı üzerinde etkisi yoktur ve modelden çıkarılması model uyumunu iyileştirmektedir.
- Tepki süresi daha uzun olan öğrencilerin maddeleri doğru yanıtlanma olasılığı daha yüksektir.
- Maddeleri yanıtlarken daha fazla eylemde bulunan bireylerin doğru yanıtlanma olasılığı daha fazladır.

- Maddeleri birden çok görüntüleyen öğrencilerin doğru yanıtlama olasılığı daha düşüktür.
- Tepki süresi, eylem sayısı ve görüntüleme sayısı değişkenlerinin modele dahil edilmesiyle madde türünün etkisinin daha iyi ortaya çıkarılabildiği, madde zorluk düzeylerindeki tutarsızlığın ortadan kalktığı ve kestirimlerin standart hatalarının azaldığı görülmüştür.
- Birey ve madde özelliklerinin süreç verileri ile birlikte modellenmesi, modelin açıklama gücünü ve tutarlılığını artırma potansiyeline sahiptir.
- Birden çok görüntülenen, az süre harcanan ve az sayıda eylemde bulunan maddelerin doğru yanıtlama olasılıklarının en düşük; çok süre harcanan, çok sayıda eylemde bulunan ve bir kez görüntülenen maddelerin doğru yanıtlanma olasılıklarının en yüksek olduğu sonucuna ulaşılmıştır.
- Maddelerin zorluk düzeyi arttıkça, evindeki kitap sayısı çok olanların ve maddelerin zorluk düzeyi azaldıkça evindeki kitap sayısı az olanların maddeleri doğru yanıtlama olasılığının eylem sayısından daha çok etkilendiği görülmektedir.

Öneriler

Araştırmanın sonuçlarına dayalı olarak geliştirilen öneriler gelecekteki araştırmalara ve uygulamaya yönelik öneriler olarak aşağıda sunulmuştur.

Uygulamaya Yönelik Öneriler

- Maddeler arasındaki rastgele etkilerin modele dahil edilmesinin model uyumunu iyileştirmesi sonucuna dayanarak, AMTM çerçevesinde kurulan modellere maddeler arasındaki rastgele etkilerin dahil edilmesi önerilmektedir.
- Eylem sayılarındaki ve tepki sürelerindeki artışın maddelerin doğru yanıtlanma olasılığını arttırdığı sonucuna dayanarak öğretmenlerin öğrencileri maddeler üzerinde daha fazla eylemde bulunmaya teşvik etmesi önerilmektedir.
- Madde özelliklerine ilişkin değişkenlerin (madde türü, madde zorluk düzeyi ve metin biçimlerinin) bireyler arasında farklılık göstermesi test tasarımında bu özelliklerin dengeli bir şekilde dağıtılmasının önemini göstermektedir. Madde

türü, metin biçimi, bilişsel boyutu ve madde zorluk düzeyi değişkenleri maddenin doğru yanıtlanma olasılığını manidar derecede etkilediği sonucuna dayalı olarak, madde kümelerinin bu özellikler bağlamında tasarlanması önerilmektedir.

- Okuma becerileri açısından kız öğrencilerin maddeleri doğru yanıtlanma olasılıklarının erkek öğrencilere göre daha yüksek olması sonucuna dayalı olarak eğitimcilerin erkek öğrencilerin okuma becerilerini geliştirmeye yönelik, erkek öğrencilerin de ilgi alanlarına hitap edebilecek yeni yaklaşımlar ve içerikler geliştirmesi önerilebilir.
- Evdeki kitap sayısının alanyazınla tutarlı olarak bu çalışmada da maddelerin doğru yanıtlanma olasılığını diğer bir deyişle öğrencilerin okuma performanslarını olumlu yönde etkilemesine dayalı olarak, ailelerin ve politika yapıcılarının evdeki kitap sayılarının arttırılmasına katkı sağlamaları önerilebilir. Yayınevleri, çocuklu ailelere yönelik kampanyalar düzenleyebilir ya da politika yapıcılar her dönem her öğrenciye belirli sayıda ücretsiz kitap desteği sağlayabilir.
- Düşük riskli bir durum belirleme çalışması olan PISA uygulamaları verilerine dayalı olarak elde edilen bu sonuçlar düşük riskli testlerde öğrencilerin test alma davranışlarının bir örneği olarak yorumlanmalıdır. Süreç verileri ile öğrenci performansı arasındaki ilişkilerin yüksek riskli testler için daha farklı sonuçlar verebileceği göz önünde tutulmalıdır. Öğrencilerin yaşamlarını önemli ölçüde etkileme potansiyeli taşıyan bu yüksek riskli ölçme uygulamalarında (Alan Yeterlilik Testi, Temel Yeterlilik Testi gibi) süreç verilerinin toplanması ölçme sonuçlarının ve değerlendirmelerin geçerliğine katkı sağlayacaktır. Bu nedenlerle politika yapıcılara bilgisayar tabanlı test uygulamalarında (Elektronik Yabancı Dil Sınavı, Elektronik Yükseköğretim Kurumları Yabancı Dil Sınavı gibi) süreç verilerinin toplanması ve ölçmelerin geçerliğini belirlemede göz önünde bulundurmaları önerilmektedir.

Araştırmacılara Yönelik Öneriler

- Bu araştırmada düşük riskli bir ölçme uygulaması olan PISA verileri kullanılmıştır. Buna paralel olarak görüntüleme sayısı açısından bireyler arası varyansın düşük olduğu görülmektedir. Yüksek riskli ölçme uygulamalarıyla elde

edilen veriler üzerinden çalışma tekrarlanarak süreç verilerine yönelik çıkarımlar karşılaştırılabilir.

- Araştırma sonuçları eylem sayısının tepki süresine göre daha etkili bir yordayıcı olduğunu ortaya koymaktadır. Eylem sayısı ve tepki süresi arasındaki ilişkilerin ortaya çıkarılmasına yönelik daha kapsamlı çalışmalar yapılabilir.
- Seçilen madde kümesinde maddelerin zorluk düzeyleri 2 ile 5 arasında değişmektedir ve çok kolay ya da çok zor maddeler bu madde kümesinde yer almamaktadır. Gelecekte yapılacak araştırmalarda güçlük düzeyi çok kolay ya da zor maddeler için birey parametrelerinin kestirimlerine ilişkin etkiler incelenebilir.
- Bu araştırmada birey özelliklerinden yalnızca cinsiyet ve evdeki kitap sayısı değişkenleri ele alınmıştır. Gelecekte yapılacak çalışmalarda lise türü gibi daha farklı birey özellikleri ele alınarak, farklı ölçme uygulamalarında ve farklı alt testlerde çalışma tekrarlanabilir.
- Bu çalışmada, PISA 2015 uygulaması kapsamında toplanan tepki süresi, eylem sayısı ve görüntüleme sayısı değişkenlerini içeren süreç verileri ele alınmıştır. Bilgisayar tabanlı ölçmelerle toplanabilen çok farklı türde süreç verisi mevcuttur. Gelecekteki araştırmalar farklı değişkenler içeren süreç verileri kullanarak bu çalışmada kullanılan yöntemleri izleyebilir.
- Süreç verilerinin bağımlı değişken olarak ele alınarak bu değişkenleri açıklayan birey ve madde özellikleri incelenebilir.

KAYNAKLAR

- Albert, D., & Steinberg, L. (2011). Age differences in strategic planning as indexed by the Tower of London. *Child Development*, 82(5), 1501-1517.
- Anghel, E., Khorramdel, L., & von Davier, M. (2024). The use of process data in large-scale assessments: a literature review. *Large-scale Assessments in Education*, 12(1), 13.
- Araneda S., Lee D., Lewis J., Sireci S. G., Moon J. A., Lehman B., ... Keehner, M. (2022). Exploring Relationships among Test Takers' Behaviors and Performance Using Response Process Data. *Education Sciences*, 12(2), 104. <https://doi.org/10.3390/educsci12020104>
- Aydoğan, İ., & Gelbal, S. (2022). Okuma Becerilerini Yordayan Özelliklerin Belirlenmesi: Genetik Algoritma Kestirimi. *Ana Dili Eğitimi Dergisi*, 10(1), 89-103. <https://doi.org/10.16916/aded.1030857>
- Bartram, D., & Bayliss, R. (1984). Automated testing: Past, present and future. *Journal of Occupational Psychology*, 57(3), 221-237. doi: 10.1111/j.2044-8325.1984.tb00164.x
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., ... & Bolker, M. B. (2015). Package 'lme4'. *Convergence*, 12(1), 2.
- Bennett, R. E. (2006) Inexorable and inevitable: The continuing story of technology and assessment. *Computer-based testing and the Internet*, 201–217.
- Bergner, Y., & von Davier, A. A. (2019). Process data in NAEP: Past, present, and future. *Journal of Educational and Behavioral Statistics*, 44(6), 706-732.
- Briggs, D. C. (2008). Using explanatory item response models to analyze group differences in science achievement. *Applied Measurement in Education*, 21(2), 89-118.
- Browne, M.W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In: Bollen, K.A., & Long, J.S. (Eds.), *Testing structural equation models* (pp. 136-162). Beverly Hills, CA: Sage
- Bulut, O. (2021). *eirm: Explanatory item response modeling for dichotomous and polytomous item responses*. doi:10.5281/zenodo.4556285, R package version 0.3.0, <https://CRAN.R-project.org/package=eirm>
- Bulut, H. C., Bulut, O., & Arikan, S. (2022). Evaluating group differences in online reading comprehension: The impact of item properties. *International Journal of Testing*, 23(1), 10-33.
- Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of statistical Software*, 48, 1-29.

- Chang, S. R., Plake, B. S., & Ferdous, A. A. (2005). *Response Times for Correct and Incorrect Item Responses on Computerized Adaptive Tests*. Paper presented at the 2005 annual meeting of the American Educational Research Association (AERA), Montréal, Canada (2005, April).
- Christensen, K. B., Makransky, G., & Horton, M. (2017). Critical values for Yen's Q3: Identification of local dependence in the Rasch model using residual correlations. *Applied psychological measurement*, 41(3), 178-194.
- De Ayala, R. J. (2009). *The theory and practice of Item Response Theory*. New York, NY: The Guilford Press.
- De Boeck, P. & Wilson, M. (Eds.) (2004). *Explanatory item response models*. New York, NY: Springer.
- De Boeck, P., Bakker, M., Zwitser, R., Nivard, M., Hofman, A., Tuerlinckx, F., & Partchev, I. (2011). The estimation of item response models with the lmer function from the lme4 package in R. *Journal of Statistical Software*, 39, 1-28.
- De Boeck, P., Cho, S. J., & Wilson, M. (2016). Explanatory item response models. In A. A. Rupp & J. P. Leighton (Eds.), *The Wiley handbook of cognition and assessment: Frameworks, methodologies, and applications*, (pp.247-266).
- De Boeck, P., & Scalise, K. (2019). Collaborative problem solving: Processing actions, time, and performance. *Frontiers in psychology*, 10, 1280.
- Desjardins, C. D., & Bulut, O. (2018). *Handbook of educational measurement and psychometrics using R*. CRC Press.
- Dodonova, Y. A., & Dodonov, Y. S. (2013). Faster on easy items, more accurate on difficult ones: Cognitive ability and performance on a task of varying difficulty. *Intelligence*, 41(1), 1-10.
- Embretson, S. (2019). Explanatory Item Response Theory Models: Impact on Validity and Test Development? In *Quantitative Psychology: 83rd Annual Meeting of the Psychometric Society*, (pp. 1-11). Springer International Publishing.
- Embretson, S. (2023). Understanding examinees' item responses through cognitive modeling of response accuracy and response times. *Large-scale Assess Educ*, 11(9). <https://doi.org/10.1186/s40536-023-00158-8>
- Ercikan, K., H. Guo, and H. Por (2023). Uses of process data in advancing the practice and science of technology-rich assessments. In Foster, N. and M. Piacentini (Eds.), *Innovating Assessments to Measure and Support Complex Skills*. OECD Publishing, Paris, <https://doi.org/10.1787/7b3123f1-en>.
- Fischer, G. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychologica*, 36, 359–374.

- Fraenkel, J., Wallen, N., & Hyun, H. (2012). *How to Design and Evaluate Research in Education*. Mc Grawall Hill.
- Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H., & Klieme, E. (2014). The time on task effect in reading and problem solving is moderated by task difficulty and skill: insights from a computer-based large-scale assessment. *Journal of Educational Psychology*, 106(3), 608.
- Goldhammer, F., Naumann, J., & Greiff, S. (2015). More is not always better: The relation between item response and item response time in Raven's matrices. *Journal of Intelligence*, 3(1), 21-40.
- Goldhammer, F., Martens, T., & Lüdtke, O. (2017). Conditioning factors of test-taking engagement in PIAAC: An exploratory IRT modelling approach considering person and item characteristics. *Large-Scale Assessments in Education*, 5, 1-25.
- Guo, H., Johnson, M., Saldivia, L., Worthington, M., & Ercikan, K. (2024). Human-Centered AI for Discovering Student Engagement Profiles on Large-Scale Educational Assessments. *Journal of Measurement and Evaluation in Education and Psychology*, 15(Special Issue), 282-301.
- Gülleroğlu, H. D., Demir, S. B., & Demirtaşlı, N. (2014). Türk Öğrencilerinin PISA 2003-2006-2009 Dönemlerindeki Okuma Becerilerini Yordayan Sosyoekonomik ve Kültürel Değişkenlerin Araştırılması. *Journal of Faculty of Educational Sciences*, 47(2).
- Hartig, J., Frey, A., Nold, G., & Klieme, E. (2012). An application of explanatory item response modeling for model-based proficiency scaling. *Educational and Psychological Measurement*, 72(4), 665-686.
- Hódi, Á., Németh, M. B., & Tóth, E. (2017). Text formats in the frameworks and reports of system-level reading assessment programs. *Docere*, 1(3-4), 26-60.
- Hódi, Á., Tóth, E., & Nikolov, M. (2025). Interactions Among Text Formats and Reading Processes in Young Learners' L1 Reading Literacy. *Reading Research Quarterly*, 60(1), e595. <https://doi.org/10.1002/rrq.595>
- Isbell, D. R., & Son, Y. A. (2023). Explanatory item response models for instrument validation: A tutorial based on an elicited imitation test. *Research Methods in Applied Linguistics*, 2(3), 100080.
- Janssen, R., Schepers, J., & Peres, D. (2004). Models with item and item group predictors. In P. De Boeck & M. Wilson (Eds.), *Explanatory item response models: A generalized linear and nonlinear approach* (pp. 189-212). New York, NY: Springer New York.
- Klein Entink, R. H., Kuhn, J.-T., Hornke, L. F., & Fox, J.-P. (2009). Evaluating cognitive theory: A joint modeling approach using responses and response times. *Psychological Methods*, 14(1), 54-75. <https://doi.org/10.1037/a0014877>

- Klotzke, K., & Fox, J. P. (2019). Bayesian covariance structure modeling of responses and process data. *Frontiers in psychology*, 10, 1675. <https://doi.org/10.3389/fpsyg.2019.01675>
- Kutlu, Ö., Yıldırım, Ö., Bilican, S., & Kumandaş, H. (2011). İlköğretim 5. sınıf öğrencilerinin okuduğunu anlamada başarılı olup-olmama durumlarının kestirilmesinde etkili olan değişkenlerin incelenmesi. *Journal of Measurement and Evaluation in Education and Psychology*, 2(1), 132-139.
- Lahza, H., Smith, T. G., & Khosravi, H. (2023). Beyond item analysis: Connecting student behaviour and performance using e-assessment logs. *British Journal of Educational Technology*, 54(1), 335-354.
- Lee, Y. H., & Haberman, S. J. (2016). Investigating test-taking behaviors using timing and process data. *International Journal of Testing*, 16, 240–267. <https://doi.org/10.1080/15305058.2015.1085385>
- Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). performance: An R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, 6(60), 3139. <https://doi.org/10.21105/joss.03139>
- Meulders, M., & Xie, Y. (2004). Person-by-item predictors. In P. De Boeck & M. Wilson (Eds.), *Explanatory item response models: A generalized linear and nonlinear approach* (pp. 213-240). New York, NY: Springer New York.
- Millî Eğitim Bakanlığı. (2016). *PISA 2015 Ulusal Rapor*. Erişim adresi: http://odsgm.meb.gov.tr/test/analizler/docs/PISA/PISA2015_Ulusal_Rapor.pdf
- Molenaar, D., & De Boeck, P. (2018). Response mixture modeling: Accounting for heterogeneity in item characteristics across response times. *Psychometrika*, 83, 279–297.
- Molenaar, D., Terlinckx, F., & van der Mass, H. L. (2015). A bivariate generalized linear item response theory modeling framework to the analysis of responses and response times. *Multivariate Behavioral Research*, 50, 56–74.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4(2), 133-142.
- OECD (2017a). *PISA 2015 Assessment and Analytical Framework: Science, Reading, Mathematic, Financial Literacy and Collaborative Problem Solving*. OECD Publishing, Paris. <https://doi.org/10.1787/9789264281820-en>.
- OECD (2017b). *PISA 2015 Technical Report*. OECD Publishing. https://www.oecd.org/content/dam/oecd/en/about/programmes/edu/pisa/publications/technical-report/PISA2015_TechRep_Final.pdf

- OECD (2024). *Annex K: Uses and Reporting of Process Data*. <https://www.oecd.org/content/dam/oecd/en/data/datasets/pisa/pisa-2018-datasets/ssas-sps-data-files/PISA2018%20Aneex%20K%20TechReport.pdf>
- Pokropek, A., and Khorramdel, L. (2024). Analyzing Test-Taking Behaviors Through Process Data Using the IRT Explanatory Model for Guessing. *PsyArXiv*. DOI: 10.31234/osf.io/3krct.
- Provasnik, S. (2021). Process data, the new frontier for assessment development: rich new soil or a quixotic quest?. *Large-scale Assessments in Education*, 9(1), 1.
- Qiao, X., Jiao, H., & He, Q. (2023). Multiple-group joint modeling of item responses, response times, and action counts with the conway-maxwell-poisson distribution. *Journal of Educational Measurement*, 60(2), 255-281. <https://doi.org/10.1111/jedm.12349>
- Quellmalz, E., & Pellegrino, J. W. (2009). Technology and testing. *Science*, 323, 75-79. doi: 10.1126/science.1168046
- Ramalingam, D., & Adams, R. (2018). How can use of data from computer-delivered assessments improve measurement? In E. Care, P. Griffin, & M. Wilson (Eds.). *Assessment and teaching of 21st century skills: Research and applications*, pp (225-238). Cham: Springer.
- Rebouças-Ju, D., & Cheng, Y. (2021). Explanatory Response Time Models. In Quantitative Psychology: *The 85th Annual Meeting of the Psychometric Society, Virtual* (pp. 35-46). Springer International Publishing.
- Reis Costa, D., & Chen, C. W. (2023). Exploring the relationship between process data and contextual variables among Scandinavian students on PISA 2012 mathematics tasks. *Large-Scale Assessments in Education*, 11(1), 5.
- Reis Costa, D., & Leoncio Netto, W. (2022). Process data analysis in ILSAs. In T. Nilsen, A. Stancel-Piątak, & J. E. Gustafsson (Eds.), *International handbook of comparative large-scale studies in education* (pp. 235-246). Springer International Handbooks of Education. https://doi.org/10.1007/978-3-030-38298-8_60-1
- Revelle, W. (2024). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois. R package version 2.4.12, <https://CRAN.R-project.org/package=psych>.
- Robitzsch, A. (2024). *sirt: Supplementary Item Response Theory Models*. R package version 4.1-15, <https://CRAN.R-project.org/package=sirt>.
- Rouet, J. F., Lowe, R., & Schnotz, W. (2008). Understanding multimedia documents: An introduction. In J. F. Rouet, R. Lowe, & W. Schnotz (Eds.), *Understanding multimedia documents* (pp. 1–16). Springer.
- Sahin, F., & Colvin, K. F. (2020). Enhancing response time thresholds with response behaviors for detecting disengaged examinees. *Large-Scale Assessments in Education*, 8, 1–24. <https://doi.org/10.1186/s40536-020-00082-1>

- Scalise, K., & Gifford, B. (2006). Computer-based assessment in e-learning: A framework for constructing "intermediate constraint" questions and tasks for technology platforms. *The Journal of Technology, Learning and Assessment*, 4(6). Retrieved from <https://ejournals.bc.edu/index.php/jtla/article/view/1653>.
- Schulz-Heidorf, K., & Støle, H. (2018). Gender differences in Norwegian PIRLS 2016 and ePIRLS 2016 results at test mode, text and item format level. *Nordic Journal of Literacy Research*, 4(1), 167–183. <https://doi.org/10.23865/njlr.v4.1270>
- Su, S., & Davison, M. L. (2019). Improving the predictive validity of reading comprehension using response times of correct item responses. *Applied Measurement in Education*, 32(2), 166-182.
- Teig, N., Scherer, R., & Kjærnsli, M. (2020). Identifying patterns of students' performance on simulated inquiry tasks using PISA 2015 log-file data. *Journal of Research in Science Teaching*, 57(9), 1400–1429.
- Ulitzsch, E., von Davier, M., & Pohl, S. (2020). A hierarchical latent response model for inferences about examinee engagement in terms of guessing and item-level non-response. *British Journal of Mathematical and Statistical Psychology*, 73, 83–112. <https://doi.org/10.1111/bmsp.12188>
- Ulitzsch, E., He, Q., Ulitzsch, V., Molter, H., Nichterlein, A., Niedermeier, R., & Pohl, S. (2021). Combining clickstream analyses and graph-modeled data clustering for identifying common response processes. *Psychometrika*, 86(1), 190–214.
- van der Linden, W. J. (2006). A lognormal model for response times on test items. *Journal of Educational and Behavioral Statistics*, 31(2), 181–204. <https://doi.org/10.3102/10769986031002181>
- van der Linden, W. J. (2009). Conceptual issues in response- time modeling. *Journal of Educational Measurement*, 46(3), 247-272. doi: 10.1111/j.1745-3984.2009.00080.x
- Wickham, H. (2007). "Reshaping data with the reshape package." *Journal of Statistical Software*, 21(12). <https://www.jstatsoft.org/v21/i12/>.
- Wilson, M., De Boeck, P., & Carstensen, C. H. (2008). Explanatory item response models: A brief introduction. In J. Hartig, E. Klieme, & D. Leutner (Eds.), *Assessment of competencies in educational contexts* (pp. 91–120). Hogrefe & Huber Publishers.
- Wise, S. L. (2019). Controlling construct-irrelevant factors through computer-based testing: disengagement, anxiety, & cheating. *Education Inquiry*, 10(1), 21-33. doi: 10.1080/20004508.2018.1490127
- Wise, S. L., & De Mars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment*, 10(1), 1–17.
- Wise, S. L., Soland, J., & Bo, Y. (2020). The (non) impact of differential test taker engagement on aggregated scores. *International Journal of Testing*, 20, 57–77. <https://doi.org/10.1080/15305058.2019.1605999>

- Wright, B. D. (1982). *Rating scale analysis*. MESA Press.
- Wu, Chen ve Stone, 2018 Wu, A. D., Chen, M. Y., & Stone, J. E. (2018). Investigating how test-takers change their strategies to handle difficulty in taking a reading comprehension test: Implications for score validation. *International Journal of Testing*, 18(3), 253-275.
- Yalçın, S. (2022). Examining Students' Item Response Times in eTIMSS According to their Proficiency Levels, Self-Confidence and Item Characteristics. *Journal of Measurement and Evaluation in Education and Psychology*, 13(1), 23-39.
- Yavuz, H. C. (2019). The effects of log data on students' performance. *Journal of Measurement and Evaluation in Education and Psychology*, 10(4), 378-390.
- Yousefpoori-Naeim, M., Bulut, O., & Tan, B. (2023). Predicting reading comprehension performance based on student characteristics and item properties. *Studies in Educational Evaluation*, 79, 101309.
- Zhang, J. (2007). Conditional covariance theory and detect for polytomous items. *Psychometrika*, 72(1), 69–91.
- Zhang, S., Li, A., & Wang, S. (2023). Exploration of Latent Structure in Test Revision and Review Log Data. *Educational Measurement: Issues and Practice*, 42(4), 53-65.
- Zumbo, B. D., Maddox, B., & Care, N. M. (2023). Process and product in computer-based assessments: Clearing the ground for a holistic validity framework. *European Journal of Psychological Assessment*, 39(4), 252.



EKLER

EK 1. Etik Kurul Onayı

ANKARA ÜNİVERSİTESİ SOSYAL BİLİMLER ALT ETİK KURULU KARAR ÖRNEĞİ

Karar Tarihi : 12/04/2021

Toplantı Sayısı : 6

Karar Sayısı : 127

127-Üniversitemiz Eğitim Bilimleri Enstitüsü Ölçme ve Değerlendirme Bölümü doktora öğrencisi **Fatma Kızıltoprak**'ın "Teknolojiyle Zenginleştirilmiş Maddelerle Yapılan Bilgisayar Tabanlı Değerlendirmelerin Geçerliliği" başlıklı tezi ile ilgili "İnsan Üzerinde Yapılan Klinik Dışı Araştırmalar Başvuru Formu" Etik Kurulumuzca incelendi.

Üniversitemiz Eğitim Bilimleri Enstitüsü Ölçme ve Değerlendirme Bölümü doktora öğrencisi **Fatma Kızıltoprak**'ın "Teknolojiyle Zenginleştirilmiş Maddelerle Yapılan Bilgisayar Tabanlı Değerlendirmelerin Geçerliliği" başlıklı tezinin, insan üzerinde yapılan bir çalışma olmaması nedeniyle etik kurul onayına ihtiyaç duymadığına oy birliği ile karar verildi.

ASLININ AYNIĞI
12/04/2021

EK 2. Araştırmada Ele Alınan Değişkenlere İlişkin Betimsel İstatistikler

Değişkenler	Kategori	N	Ortalama	Std. Sapma	Min.	Maks.	Açıklık	Çarpıklık	Basıklık
M1_ Tepki Süresi		705	63070.45	55327.88	1917	536203	534286	2.68	12.80
M2_ Tepki Süresi		705	67666.46	41536.81	1565	428491	426926	2.44	12.14
M3_ Tepki Süresi		705	47445.98	29347.22	1047	339250	338203	2.73	16.90
M4_ Tepki Süresi		705	61993.35	40187.22	3704	263330	259626	1.38	2.90
M5_ Tepki Süresi		705	178455.10	118779.37	1819	748053	746234	1.39	2.84
M6_ Tepki Süresi		705	125641.96	94236.10	1115	883355	882240	2.08	9.14
M7_ Tepki Süresi		705	97855.78	64720.76	681	548241	547560	1.87	6.71
M8_ Tepki Süresi		705	118760.33	95515.02	452	533596	533144	1.19	1.56
M9_ Tepki Süresi		705	31086.44	32271.97	2509	649974	647465	10.88	192.22
M10_ Tepki Süresi		705	100732.38	57223.51	2727	489106	486379	1.21	4.04
M11_ Tepki Süresi		705	53399.98	38776.15	3515	338066	334551	1.75	5.69
M12_ Tepki Süresi		705	58080.16	64094.33	1198	416442	415244	2.26	6.03
M13_ Tepki Süresi		705	36598.52	37118.90	1355	362491	361136	2.68	12.22
M14_ Tepki Süresi		705	38459.35	28331.34	929	268662	267733	2.36	10.29
M15_ Tepki Süresi		705	35296.67	30771.03	645	238044	237399	2.00	5.48
M16_ Tepki Süresi		705	47830.43	37855.32	3100	581201	578101	5.40	60.53
M1_ Görüntüleme Sayısı		705	1.56	1.32	1	12	11	3.58	16.33
M2_ Görüntüleme Sayısı		705	1.54	1.15	1	11	10	3.24	14.14
M3_ Görüntüleme Sayısı		705	1.36	0.79	1	7	6	2.76	9.31
M4_ Görüntüleme Sayısı		705	1.13	0.37	1	3	2	2.98	8.69
M5_ Görüntüleme Sayısı		705	1.11	0.38	1	5	4	4.39	26.33
M6_ Görüntüleme Sayısı		705	1.25	0.69	1	7	6	3.66	16.68
M7_ Görüntüleme Sayısı		705	1.28	0.77	1	8	7	4.01	22.10
M8_ Görüntüleme Sayısı		705	1.25	0.65	1	5	4	3.07	10.48
M9_ Görüntüleme Sayısı		705	1.11	0.35	1	4	3	3.67	15.30
M10_ Görüntüleme Sayısı		705	1.11	0.34	1	3	2	2.96	8.45
M11_ Görüntüleme Sayısı		705	1.11	0.34	1	3	2	2.96	8.45
M12_ Görüntüleme Sayısı		705	1.11	0.39	1	5	4	4.56	26.96
M13_ Görüntüleme Sayısı		705	1.18	0.51	1	5	4	3.24	11.54
M14_ Görüntüleme Sayısı		705	1.18	0.57	1	5	4	3.85	16.31
M15_ Görüntüleme Sayısı		705	1.13	0.51	1	5	4	4.38	20.60
M16_ Görüntüleme Sayısı		705	1.06	0.25	1	3	2	4.47	21.01
M1_ Eylem Sayısı		705	5.32	5.87	1	78	77	6.37	60.47
M2_ Eylem Sayısı		705	9.13	9.94	1	175	174	9.21	125.88
M3_ Eylem Sayısı		705	4.86	3.13	1	35	34	3.60	21.19
M4_ Eylem Sayısı		705	6.03	4.11	2	48	46	4.22	26.46
M5_ Eylem Sayısı		705	109.08	99.45	0	907	907	2.04	8.38
M6_ Eylem Sayısı		705	83.80	90.68	1	717	716	2.32	8.91
M7_ Eylem Sayısı		705	90.80	81.53	1	1222	1221	4.86	54.31
M8_ Eylem Sayısı		705	99.18	117.94	1	717	716	1.66	3.36

M9_Eylem Sayısı		705	4.69	3.45	2	51	49	8.06	94.44
M10_Eylem Sayısı		705	5.21	8.38	0	118	118	7.32	75.45
M11_Eylem Sayısı		705	6.43	6.20	2	82	80	5.32	42.91
M12_Eylem Sayısı		705	4.45	6.61	1	108	107	11.39	162.79
M13_Eylem Sayısı		705	4.07	5.01	1	69	68	8.47	88.85
M14_Eylem Sayısı		705	3.74	3.40	1	42	41	5.63	42.72
M15_Eylem Sayısı		705	3.70	3.37	1	50	49	6.99	69.39
M16_Eylem Sayısı		705	5.34	4.76	2	95	93	11.22	186.80
M1		705	0.43	0.49	0	1	1	0.29	-1.92
M2		705	0.61	0.49	0	1	1	-0.46	-1.79
M3		705	0.59	0.49	0	1	1	-0.38	-1.86
M4		705	0.49	0.50	0	1	1	0.04	-2.00
M5		705	0.71	0.45	0	1	1	-0.95	-1.10
M6		705	0.56	0.50	0	1	1	-0.25	-1.94
M7		705	0.19	0.39	0	1	1	1.58	0.49
M8		705	0.34	0.47	0	1	1	0.67	-1.56
M9		705	0.31	0.46	0	1	1	0.81	-1.35
M10		705	0.50	0.50	0	1	1	0.00	-2.00
M11		705	0.45	0.50	0	1	1	0.22	-1.95
M12		705	0.33	0.47	0	1	1	0.71	-1.49
M13		705	0.75	0.43	0	1	1	-1.15	-0.69
M14		705	0.45	0.50	0	1	1	0.21	-1.96
M15		705	0.68	0.47	0	1	1	-0.77	-1.41
M16		705	0.28	0.45	0	1	1	0.97	-1.05
Cinsiyet	Kız	346							
	Erkek	359	1.51	0.50	1	2	1	-0.04	-2.00
Kitap Sayısı	0 ile 10 arası	176							
	11 ile 25 arası	208							
	26 ile 100 arası	196							
	101 ile 200 arası	63							
	201 ile 500 arası	38							
	500'den çok	24	2.50	1.29	1	6	5	0.81	0.23

BENZERLİK BİLDİRİMİ

“Süreç Verileri İle Birey Ve Madde Özelliklerinin Maddelerin Doğru Yanıtlanma Olasılığına Etkisi” başlıklı tezimin ana bölümü (ön bölüm, kaynaklar ve ekler hariç) Turnitin İntihal Tespit Programı aracılığıyla incelenmiş ve ilgili rapor danışmanım tarafından da kontrol edilmiştir. Kontrol sırasında (1) “Yedi sözcükten daha az olan benzeşmeler” (2) “Kaynaklar” (3) “Doğrudan alıntılar” ve (4) “Lisansüstü tezim ile ilgili yapmış olduğum kendi yayınlarım ve yararlandığım mevzuat metinleri gibi kaynaklar” dışarıda tutulmuştur. Benzerlik kontrolüne ilişkin rapordan elde edilen sonuçlar aşağıda sunulmuştur.

Rapor Tarihi	: 14.03.2025
Gönderim Numarası	: 3182900499
Sayfa Sayısı	: 77
Sözcük Sayısı	: 20.580
Karakter Sayısı	: 138.912
Benzerlik Oranı	: %5
Savunma Tarihi	: 14.03.2025

Yukarıda belirtilen sonuçları gösteren Turnitin İntihal Tespit Programı’na ilişkin orijinal raporu, sonuçlarda herhangi bir değişiklik yapmaksızın bu beyanım ekinde Enstitüye teslim ettiğimi, tezimin %10’dan fazla benzerlik oranı içerdiğinin, tek bir kaynakla eşleşme oranının ise %2’den fazla olduğunun belirlenmesi durumunda, bundan doğabilecek tüm yasal sorumluluğu kabul ettiğimi bildirir, saygılarımı sunarım.

Öğrencinin Adı Soyadı: Fatma KIZILTOPRAK

Tarih:14.03.2025

İmza:

ÖZGEÇMİŞ

Kişisel Bilgiler

Adı ve Soyadı : Fatma KIZILTOPRAK

E-Posta Adresi :

İş Deneyimi :

Unvan	Görev Yeri	Yıl
Arş. Gör.	Muğla Sıtkı Koçman Üniversitesi EF	2021- halen
Arş. Gör.	Ankara Üniversitesi EBF	2017- 2021
Arş. Gör.	Gazi Üniversitesi EBF	2014- 2017

Akademik Bilgiler

Öğrenim Durumu:

Derece	Bölüm/Program	Üniversite	Yıl
Doktora	Matematik Eğitimi	Anadolu Üniversitesi	2017
Yüksek Lisans	Eğitimde Ölçme ve Değerlendirme	Gazi Üniversitesi	2016
Lisans	İlköğretim Matematik Öğretmenliği	Osmangazi Üniversitesi	2005

Yayınlar:

Özyeter, N. T., & Kızıltoprak, F. (2021). The Change of Students' Science Related Affective Characteristics over Years: The Comparison of PISA 2006-2015 Assessments. *Kastamonu Education Journal*, 29(5), 838-855.

Kızıltoprak, F. (2019). Matematik okuryazarlığının ölçülmesi ve değerlendirilmesi. T. Kabael. (Ed.), *Matematik okuryazarlığı ve PISA* (ss. 59-106). Ankara: Anı Yayıncılık.