

**ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

**ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME DOKTORA PROGRAMI**

**TIMSS 2015 UYGULAMASINDA TUTUM-BAŞARI PARADOKSUNUN UÇ
TEPKİ STİLİNE GÖRE ÇOK BOYUTLU MADDE TEPKİ KURAMI İLE
MODELENMESİ**

DOKTORA TEZ

MÜNEVVER İLGÜN DİBEK

Ankara, Ekim, 2018



**ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

**ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME DOKTORA PROGRAMI**

**TIMSS 2015 UYGULAMASINDA TUTUM-BAŞARI PARADOKSUNUN UÇ
TEPKİ STİLİNE GÖRE ÇOK BOYUTLU MADDE TEPKİ KURAMI İLE
MODELENMESİ**

DOKTORA TEZ

MÜNEVVER İLGÜN DİBEK

DANIŞMAN: R.NÜKHET ÇIKRIKÇI

Ankara, Ekim, 2018

ONAY

Eđitim Bilimleri Enstitüsü M¼d¼rl¼ę¼'ne

M¼nevver İLG¼N DİBEK'in hazırladıđı "TIMSS 2015 Uygulamasında Tutum-Başarı Paradoksunun Uç Tepki Stiline G¼re Çok Boyutlu Madde Tepki Kuramı İle Modellenmesi" başlıklı bu alıřma j¼rimiz tarafından ¼lme ve Deđerlendirme Anabilim Dalı/¼lme ve Deđerlendirme Programı'nda Doktora Tezi olarak kabul edilmiřtir.

İmza

Başkan Do.Dr. ¼may OKLUK B¼KEOĐLU

¼ye Prof. Dr. N¼khet IKRIKI

¼ye Prof.Dr. Nuri DOĐAN

¼ye Do.Dr. Sevilay KİLMEN

¼ye Do.Dr. C.Deha DOĐAN

ONAY

Bu tez Ankara ¼niversitesi Lisans¼st¼ Eđitim-¼đretim ve Sınav Y¼netmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri ¼yeleri tarafından ./.../20.. tarihinde uygun g¼r¼lm¼ř ve Enstit¼ Y¼netim Kurulunca .../.../20.. tarihinde kabul edilmiřtir.

Prof. Dr. Yasemin KEPENEKCI

Eđitim Bilimleri Enstit¼s¼ M¼d¼r V.

TEZ BİLDİRİMİ

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.



Münevver İLGÜN DİBEK

ÖZET

TIMSS 2015 UYGULAMASINDA TUTUM-BAŞARI PARADOKSUNUN UÇ TEPKİ STİLİNE GÖRE ÇOK BOYUTLU MADDE TEPKİ KURAMI İLE MODELLENMESİ

İlgün Dibek, Münevver

Doktora, Ölçme ve Değerlendirme Anabilim Dalı

Tez Danışmanı: Prof.Dr. Nükhet ÇIKRIKÇI

Ekim 2018, xi + 121 sayfa

Başarı-tutum paradoksu ülke düzeyinde tutum ile başarı arasındaki ilişkinin negatif, birey düzeyinde ise pozitif yönde olmasıdır. Bu bağlamda bu çalışmanın amacı, TIMSS 2015 uygulamasına katılan ülkeler arasında tutum-başarı paradoksuna neden olabilecek uç tepki stilinin (UTS) etkisini belirlemek ve bu etkiyi düzelterek tutum ve başarı arasındaki ilişkiyi hem ülkeler içinde hem de ülkeler arasında belirlemektir. Korelasyonel araştırma modelinde olan bu araştırmanın örneklemini TIMSS 2015 uygulamasına katılan ülkeler arasından belirlenen 15 ülkeden seçkisiz olarak seçilen 500'er 8.sınıf öğrencisi oluşturmaktadır. Ülkelerin UTS sergilemelerinde farklılıklar MANOVA ile belirlenmiştir. Bireylerin cevaplarında UTS'nin etkisini belirlemek amacıyla UTS'nin bir boyut olarak dahil edilmediği ve edildiği iki farklı Çok Boyutlu Madde Tepki Kuramı (ÇBMTK) modeli Latent GOLD 5.1 ve WinBUGS programlarıyla analiz edilmiştir. Öğrencilerin düzeltilmemiş ve düzeltilmiş tutum puanları arasındaki farkın manidarlığını değerlendirmede her bir ülke için ilişkili örneklem için t-testi analizi yapılmıştır. Tutumla ilgili değişkenlerin matematik başarısıyla ilişkisini belirlemek amacıyla gözlenen puanlara dayalı korelasyon değeri ile ÇBMTK modellerinden elde edilen korelasyon değerleri elde edilmiştir. Korelasyon değerleri arasında manidar farkın olup olmadığı Fisher'in r_z dönüşümü ile belirlenmiştir. Araştırmanın bulguları şu şekildedir: (a) Ülkelerin UTS gösterme eğilimleri farklılaşmaktadır; (b) UTS'nin bir boyut olarak dahil edildiği ÇBMTK modeli veriye daha iyi uyum sağlamaktadır; (c) düzeltilmemiş ve düzeltilmiş tutum puanları arasında manidar farklılaşmanın ülke sayısı açısından en fazla olduğu boyut matematiğe değer verme boyutudur;(d) gözlenen puanlara dayalı korelasyon değerleri negatif, ÇBMTK modellerine dayalı olarak kestirilen puanlara ilişkin korelasyon değerleri pozitifdir ve farklılaşmaktadır. Başarı- tutum paradoksuna neden olan faktörlerden biri UTS'dir. Fakat UTS bu paradoksu açıklamada tek başına yeterli değildir.

Anahtar Sözcükler: Çok boyutlu madde tepki kuramı, TIMSS, Tutum-başarı paradoksu, Uç tepki stili

SUMMARY

MODELLING OF ATTITUDE-ACHIEVEMENT PARADOX IN TIMSS 2015 WITH RESPECT TO EXTREME RESPONSE STYLE USING MULTIDIMENSIONAL ITEM RESPONSE THEORY

İlgün Dibek, Münevver

Dissertation, Department of Educational Measurement and Evaluation

Advisor: Prof. Dr. Nükhet ÇIKRIKÇI

October 2018, xi + 121 pages

Achievement-attitude paradox is that country-level correlation between attitude and achievement is negative while individual-level correlation between them is positive. In this regard, this study aimed to first investigate the effect of the extreme response style (ERS) which could lead to attitude-achievement paradox among the countries participating in TIMSS 2015 and then to determine the individual- and country-level relationship between attitude and achievement by adjusting the effect of ERS. For the sample of this correlational study, 500 students were randomly selected from each of the 15 countries that participated in TIMSS 2015. The differences in the ERS tendency of the countries were determined by performing MANOVA. To determine the effect of ERS, two different multidimensional item response theory (MIRT) models were used; the first not including the ERS trait as a dimension and the other including this trait as a dimension. The results were analyzed with Latent GOLD 5.1 and WinBUGS softwares. The differences between these scores were tested for each country by performing a separate paired samples t-test. To determine the relationship between attitudinal variables and achievement, the correlation values based on observed scores and MIRT models were obtained. Whether there was any significant difference between these correlation values was determined by Fisher's r_z transformation. The findings of this study were as follows: a) ERS tendency of the countries were different (b) The model in which the ERS trait was included as a dimension best fit the data; c) the value students placed on mathematics was the variable for which the highest significant difference was found between uncorrected and corrected attitude scores in terms of the number of countries; and d) the correlation values based on observed scores were negative and those based on the MIRT models were positive, with the two statistically differing from each other. ERS is one of the factors causing achievement-attitude paradox. However, it not sufficiently enough to explain achievement-attitude paradox.

Key Words: Multidimensional Item Response Theory, TIMSS, Attitude-achievement paradox, Extreme Response Style

ÖNSÖZ

Tutum ile başarı arasındaki ilişkinin birey düzeyinde pozitif grup düzeyinde ise negatif olmasının nedenlerinden biri ülkeler arasındaki olası tepki stilleri farklılıklarıdır. Tepki stilleri, bireylerin test puanlarına sistematik hata karıştırarak test puanlarının geçerliliğini düşürmektedir. Kültürlerarası karşılaştırmalarda sonuçların sistematik bir şekilde değişmesine yol açması nedeniyle bu tepki stillerinin etkilerinin düzeltilmesi gereklidir. Özellikle TIMSS gibi çeşitli ülkelerin katıldığı çalışmalarının sonuçlarının eğitim politikalarına yön verdiği ve buna göre gerekli uygulamaların gerçekleştirildiği düşünüldüğünde bu çalışmalarının sonuçlarının yansız bir şekilde sunulması önemlidir. Bu yüzden, bu çalışmada TIMSS 2015 uygulamasına katılan öğrencilerin cevaplarında UTS'nin etkisi belirlenmiştir. Bu amaçla, farklı ülkelerdeki sekizinci sınıf öğrencilerinin tutum ölçeklerine verdikleri cevaplar üzerinde UTS'nin etkisi düzeltilerek matematiğe yönelik tutum ve TIMSS matematik başarısı arasındaki ilişkinin derecesi hem birey hem de ülke düzeyinde belirlenmiştir.

Lisansüstü eğitim hayatım boyunca, bana desteğini esirgemeyen, gelişimde büyük katkısı olan değerli hocam ve danışmanım Prof. Dr. Nükhet ÇIKRIKÇI'ya, tez izleme komitesinde yer alan ve değerli bilgileriyle tezime katkıda bulunan saygıdeğer hocalarım Prof. Dr. Nuri DOĞAN'a, Doç. Dr. Deha DOĞAN'a; tez jürimde yer alan ve tezime katkı sağlayan değerli hocalarım Doç. Dr. Ömay Çokluk BÖKEOĞLU'na ve Doç. Dr. Sevilay KİLMEN'e; süreçte özellikle manevi anlamda da destek veren, ilgisini eksik etmeyen sevgili hocam Dr. Öğrt. Üyesi Ömer KUTLU'ya; tez sürecinde analizlerdeki sorunlarımı çözen, yardımını esirgemeyen saygı değer Prof. Dr. Daniel BOLT'a, çalışmamda görüşlerine başvurduğum, süreçte hızlanmamı sağlayan, tez sürecimde beni hep motive eden sevgili arkadaşlarım Arş. Gör. Hatice Çiğdem YAVUZ'a, Arş. Gör. Dr. Seher YALÇIN'a, Arş. Gör. Merve ŞAHİN KÜRŞAD'a ve Arş. Gör. Burcu TOPTAŞ'a; süreçte gösterdiği sabır için, zor günlerimde hep yanımda olup bana destek olan eşim Burak DİBEK'e; çalışmamı ileriye götürmemi sağlayan Samet DİBEK'e; bugünlere gelmemde büyük emeği olan, rehberlik eden anneme, babama ve abime; canımdan bir parça olan oğlumun emanet ettiğim, arka planda yatan iş yükümü azaltan kayınvalideme ve kayınbabama teşekkürlerimi sunarım.

Ayrıca doktora eğitimim süresince verdiği maddi destekten ötürü Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK)'na teşekkür ederim.

Münevver İLGÜN DİBEK



OĞLUM FURKAN'A

İÇİNDEKİLER

Sayfa

ONAY.....	ii
BİLDİRİM.....	iii
ÖZET.....	iv
SUMMARY.....	v
ÖNSÖZ.....	vi
İÇİNDEKİLER.....	viii
ÇİZELGELER DİZİNİ.....	x
ŞEKİLLER DİZİ.....	xi
BÖLÜM	
BÖLÜM 1.....	1
Problem Durumu.....	1
Amaç.....	14
Önem.....	15
Sınırlılıklar.....	17
Kısaltmalar.....	17
Tanımlar.....	17
BÖLÜM 2.....	19
Yöntem.....	19
Araştırma Modeli.....	19
Evren ve Örneklem.....	19
Veriler ve Elde Edilmesi.....	24
Verilerin Analizi.....	27
BÖLÜM 3.....	42
Bulgular ve Yorumlar.....	42
TIMSS 2015 Uygulamasına Katılan Ülkelerin Tutum Ölçeklerinden	
Elde Edilen UTS Puanlarına İlişkin Farklılıklara İlişkin Bulgular....	42

UTS'nin Bir Boyut Olarak Ele Alındığı ve Alınmadığı Modellerin Veri ile Uyumuna İlişkin Bulgular.....	42
Toplam Ölçek Puanlarına UTS'nin Karıştırdığı Yanlılığın Miktarının ve Yönünün Farklı UTS Değerlerine göre Değişimine İlişkin Bulgular	44
Öğrencilerin UTS'nin Etkisine göre Düzeltilmemiş ve Düzeltilmiş Ortalama Tutum Puanlarının Arasındaki Farklara İlişkin Bulgular...	47
UTS'nin Etkisine göre Düzeltilmemiş ve Düzeltilmiş Tutum Puanlarının Matematik Başarısıyla Olan Ülke İçi ve Ülkeler Arası İlişkilere İlişkin Bulgular.	54
UTS'nin Etkisine göre Düzeltilmemiş ve Düzeltilmiş Ortalama Tutum Puanlarının Matematik Başarısıyla Ülke İçi ve Ülkeler Arası İlişkilerin Arasındaki Farklara İlişkin Bulgular	57
BÖLÜM 4	61
Sonuç ve Öneriler	61
Sonuç.....	61
Öneriler	62
KAYNAKLAR.....	64
EKLER	71
EK A. Tutumla ilgili Boyutların Matematik Başarısıyla İlişkilerine Ait Saçılma Grafikleri	72
EK B. Tutum Değişkenlerine Ait Maddeler	75
EK C. Tutum Ölçeklerine Ait Cronbach Alfa Güvenirlilik Katsayısı Değerleri	77
EK D Örneklem-Alt Örneklem için Değişkenler Arası Korelasyon Değerleri	83
EK E.1 Parametrik Olmayan MANOVA için R Kodu	81
EK E.2 Parametrik Olmayan MANOVA Sonuçları	81
EK F Latent Gold Kodları	83
EK G WinBUGS Kodları	85
EK H. Üç Boyutlu Modele Ait Zaman Serisi Grafikleri	88
EK I.1 İlişkili Örneklem için t-testi Varsayımları	91
EK I.2 Herbir Ülke için Histogram Grafikleri	99

ÇİZELGELER DİZİNİ

Çizelge	Sayfa
1. Ülke İçi Korelasyon Değerleri.....	23
2. Ülkelerin Evren ve Örnekleme.....	25
3. Üç ve Dört Boyutlu Modellerle Elde Edilen Kategori Kesişim Parametreleri.....	33
4. MCMC Simülasyonlarına Ait Betimsel İstatistikler ve MC Hataları	38
5. Model Karşılaştırma Sonuçları.....	43
6. Bazı Bireylerin Düzeltilmemiş ve Düzeltilmiş Ortalama Tutum Puanları.....	53
7. Ülkelerin Düzeltilmemiş ve Düzeltilmiş Tutum Puanlarının Matematik Başarıyla Olan İlişkilerine Dair Ülke İçi Korelasyon Değerleri.....	54
8. Ülkelerin Düzeltilmemiş ve Düzeltilmiş Tutum Puanlarının Matematik Başarıyla Olan Ülkeler Arası İlişkileri.....	55
9. Ülkelerin Düzeltilmemiş ve Düzeltilmiş Ortalama Tutum Puanlarının Matematik Başarıyla Olan Ülke İçi İlişkilerin Arasındaki Farklar.....	57
10. Ülkelerin Düzeltilmemiş ve Düzeltilmiş Tutum Puanlarının Matematik Başarıyla Olan İlişkilerine Ait Ülkeler Arası Korelasyon Değerleri Arasındaki Farklar.....	59

ŞEKİLLER DİZİNİ

Şekil	Sayfa
1. Tepki Stilleri (Liu, 2015, s.3).....	6
2. Ülkelerin Matematik Başarısı ve Matematik Dersinde Kendine Güvenmemesine göre Dağılımları.....	21
3.Ülkelerin Matematik Başarısı ve Matematik Dersini Sevmemesine göre Dağılımları	22
4. Ülkelerin Matematik Başarısı ve Matematik Dersine Değer Vermemesine göre Dağılımları.....	23
5. Üç Boyutlu Modelde Başarı ve Sevme Örtük Değişkeni Arasındaki İlişkiye Ait Markov Zinciri Zaman Serisi Grafiği.....	37
6. Dört Boyutlu Modelde Başarı ve Sevme Örtük Değişkeni Arasındaki İlişkiye Ait Markov Zinciri Zaman Serisi Grafiği.....	37
7. Matematiği Öğrenmeyi Sevme Ölçeğinden Elden Edilen Toplam Puanlardaki Yanlılık.....	45
8. Matematikte Kendine Güven Ölçeğinden Elde Edilen Toplam Puanlardaki Yanlılık.....	45
9. Matematiğe Değer Verme Ölçeğinden Elde Edilen Toplam Puanlardaki Yanlılık	46
10. Ülkelerin Matematiği Öğrenmeyi Sevmeye İlişkin Düzeltilmemiş ve Düzeltilmiş Puanlarının Dağılımı.....	48
11. Ülkelerin Matematikte Kendine Güvene İlişkin Düzeltilmemiş ve Düzeltilmiş Puanlarının Dağılımı.....	49
12. Ülkelerin Matematiğe Değer Vermeye İlişkin Düzeltilmemiş ve Düzeltilmiş Puanlarının Dağılımı.....	50

BÖLÜM 1

GİRİŞ

Bu bölümde araştırmanın problemi, amacı, önemi, sınırlılıkları, kısaltmalar ve tanımlar açıklanmıştır.

1.1 Problem Durumu

Eğitim alanında, geniş ölçekli uluslararası karşılaştırma çalışmaları ülkelerin kendi değerlerini, eğitime ilişkin uygulamalarını, teorilerini incelemelerini ve daha gerçekçi standartlar ortaya koyarak eğitim sistemlerinin başarılarını kontrol etmeyi sağlamaktadır. Bu yüzden ülkeler karşılaştırma çalışmalarına katılmaktadır (Plomp, 1998). Bu çalışmalar, belirli bir yaş veya sınıf seviyesindeki bir grubun en az bir alandaki başarısının eğitim sistemleri boyunca karşılaştırıldığı ve bu sistemlerdeki, okullardaki, sınıflardaki ve başarıya ilişkin öğrenci düzeyindeki bağlamsal faktörlerin etkilerinin araştırıldığı çalışmalar olarak tanımlanmaktadır (Bos, 2002).

Ekonomik İşbirliği ve Kalkınma Örgütü (Organisation for Economic Co-operation and Development [OECD]) tarafından yürütülen Uluslararası Öğrenci Değerlendirme Programı (Program for International Student Assessment [PISA]) ve Uluslararası Eğitim Başarısını Değerlendirme Kuruluşu (International Association for the Evaluation of Educational Achievement [IEA]) tarafından yürütülen Uluslararası Matematik ve Fen Bilgisi Araştırma Projesi (Trends in International Mathematics and Science Study [TIMSS]), ülkelerarası araştırmalar yapılması imkanını sağlayan etkili uluslararası çalışmalardan bazılarıdır. Bu çalışmalardan, ilki 1995 yılında gerçekleşen TIMSS uygulaması, ülkelerin 4. ve 8. sınıf öğrencilerini dâhil ettiği dört yılda bir gerçekleşen tarama çalışmasıdır (Mullis, Martin ve Loveless, 2016). TIMSS, fen ve matematik başarısı ile ilgili bilişsel veriler sunmasının yanı sıra öğrencilere, öğretmenlere, velilere ve okul yöneticilerine sunulan farklı formatlardaki anketler aracılığıyla bilişsel olmayan değişkenlere ilişkin (tutum, algı, ilgi) veriler de sunmaktadır. Bilişsel olmayan özelliklerden biri olarak da tutumun ele alındığı TIMSS uygulamalarının sonuçlarından elde edilen uluslararası anket çalışmaları ve yayınların sayıca fazla olması, uluslararası karşılaştırmalarda testlerdeki başarı ile ilintili olarak

tutumlardaki farklılaşmaya olan ilgilinin göstergeleri olarak görülebilir. Bu çalışmalar farklı kültürlerdeki tutum ve değerler açısından farklılıklara ışık tutmaktadır.

TIMSS kapsamında ele alınan boyutlardan biri olan matematik başarısı üzerinde önemli bir etkiye sahip duyuşsal özellik matematięe yönelik tutumdur (Farooq ve Shah, 2008; Grootenboer ve Marshman, 2016; Kiwanuka, Van Damme, Van den Noortgate, Anumendem ve Namusisi, 2015). Matematięe yönelik tutum, bireylerin matematięi sevmesiyle, matematikle ilgili etkinliklere katılmasıyla, matematięin faydalı olmasıyla ve matematik başarısıyla ilgili düşüncelerinden oluşmaktadır (Neale, 1969). Öğrencinin matematięe yönelik tutumunu inceleyen araştırmalar (de Lourdes Mata, Monteiro ve Peixoto, 2012; Sundre, Barry, Gynnild ve Ostgard, 2012), öğrenmeye yönelik motivasyonunun ilgili alanın sevilmesi ve ona değer verilmesinden etkilenmesi nedeniyle motivasyon kavramına da odaklanmıştır. Bunun yanı sıra, öğrencilerin öğrenmeye yönelik motivasyonu dersi başarabileceklerine dair inançlarıyla da ilgili olduğu için derste kendilerine olan güvenlerinden de etkilenmektedir (Mullis, Martin, Foy ve Arora, 2012). Bu doğrultuda, TIMSS uygulamalarında matematięe yönelik motivasyonla ve dolayısıyla tutumla ilgili olan “matematięi öğrenmeyi sevmesi (students like learning mathematics)”, “öğrencinin matematikte kendine güvenmesi (students confident in mathematics)”, ve “matematięe değer vermesi (students value mathematics)” adlı üç psikolojik yapı ölçülmektedir (Kadijevich, 2006). Matematik başarısı ve tutum ile ilişkili olan matematikten hoşlanma, öğrencinin matematięe ilgisi ve içsel motivasyonu ile ilgilidir (Mullis ve diğerleri, 2012). Matematięe yönelik tutumla yakından ilgili olan matematikte kendine güven kavramı, öğrencinin matematik dersinde başarılı olduğuna inanması ve matematięi öğrenirken yapabileceğine inanması olarak tanımlanmaktadır (Demir ve Kılıç, 2010). TIMSS’de uygulanan öğrencinin matematikte kendine güvenmesi ölçeęi, öğrencilerin matematięi öğrenmeye ilişkin yeterliliklerine dair kendilerine güvenlerini veya benlik algılarını ölçmektedir. Matematięe değer verme ise, bireylerin dışsal motivasyonuna yönelik olmakla beraber matematięin hayatlarının çeşitli dönemlerindeki faydası ve önemiyle ilgili bireylerin tutumunu ifade etmektedir (Wigfield ve Eccles, 2000). Özetle, tutumla ilgili bu kavramların ortak noktasına dikkat edildiğinde, alanyazındaki araştırmaların da vurguladığı üzere bu kavramların matematik başarısı üzerinde önemli bir role sahip olduğu (Arıkan, van de Vijver ve Yağmur, 2016; Belbase, 2013; Khine, Al-Mutawah ve Afari, 2015; Kim, Park, Park ve Kim, 2013) görülmektedir. Bu nedenle, farklı

ülkelerdeki öğrencilerin matematik başarılarındaki farklılaşmanın anlaşılması açısından tutum boyutuna odaklanılması önemlidir.

Tutum ile başarı arasındaki ilişkiyi araştıran uluslararası karşılaştırma çalışmaları, başarısı yüksek olan öğrencilerin derse yönelik olumlu tutum sergileme eğiliminde olduğunu (Kadijevich, 2008; Marsh, Trautwein, Lüdtke, Köllerve Baumert, 2005) belirtirken bazı çalışmalar ise (Buckley, 2009; Bybee ve MaCrae, 2007; Van de Gaer ve Adams, 2010) ülkeler arasında öğrencinin başarısının yüksek olmasına karşın öğrencinin derse yönelik tutumun olumsuz olduğunu vurgulamaktadır. Tutum ile başarı arasındaki negatif ilişki TIMSS, PISA ve Uluslararası Okuma Becerileri Gelişim Projesi (Progress in International Reading and Literacy Study [PIRLS]) gibi uluslararası karşılaştırma çalışmalarında da gözlemlenmektedir. Tutum ile başarı arasında pozitif ilişki olduğunu vurgulayan beklenti değer kuramı (expectancy value theory) (Atkinson, 1957) gibi motivasyon teorilerinin aksine bu ilişkinin yönü, birey düzeyi (individual level) ya da grup düzeyinde (group level) yapılan belirlemelere göre değişmektedir. Bir diğer ifade ile ülke içinde öğrencilerin ilgili dersteki tutumu ile başarısı arasında pozitif ilişki söz konusu olabilirken ülkeler arasında öğrencilerin tutum ile başarısı arasında negatif ilişki söz konusu olabilir (Bofah ve Hannula, 2015; Van de gaer, Grisay, Schulz ve Gebhardt, 2012). Bu yüzden, birey ve grup düzeyinde elde edilen korelasyonların birbiri yerine kullanılması araştırmalardan elde edilen sonuçların geçerliliğini düşürmektedir (Robinson, 1950).

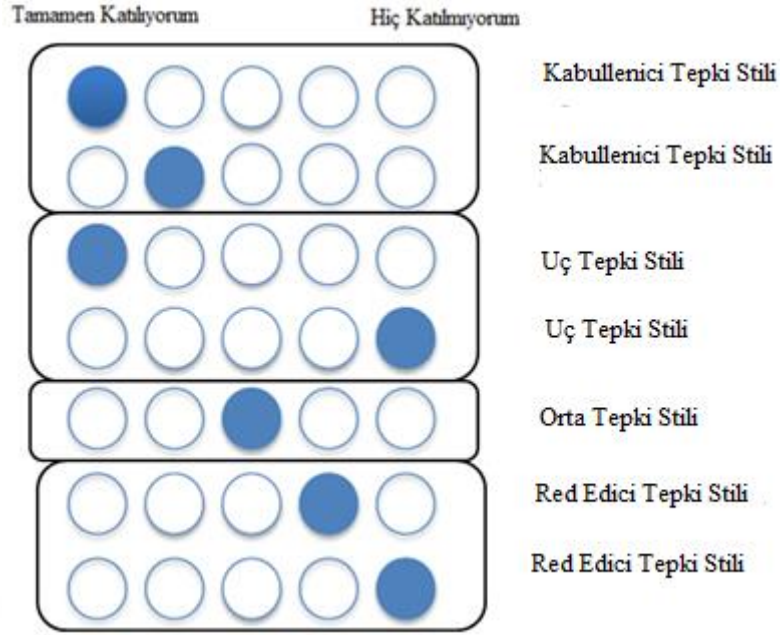
Alanyazında, tutum ile başarı arasındaki ilişkinin ülke içinde pozitif, ülkeler arasında ise negatif olması başarı-tutum paradoksu (achievement-attitude paradox) olarak tanımlanmaktadır (Van de ve diğerleri, 2012). Bu başarı-tutum paradoksunun ortaya çıkmasının başlıca nedeni, ölçeklerden elde edilen ortalama puanların kültürler arasında karşılaştırılmasını amaçlayan çalışmaların dayandığı bazı temel sayıtların geçerli olmamasıdır (Harzing, 2006). Bu varsayımlardan biri, bireylerin maddeleri cevaplandırırken maddelere yükledikleri anlamları dikkate alarak cevapladıklarıdır. Bir diğer ifade ile cevaplayıcıların tepki kategorilerini tercihlerinde madde içeriğinden bağımsız herhangi bir faktörün sistematik etkisinin olmadığı varsayımdır. Bir diğer varsayım, bir cevaplayıcı grubu içindeki bütün bireylerin, maddeleri cevaplarırken tepki kategorilerine yükledikleri anlamların da aynı olmasıdır (Cho, 2013). Fakat, bu varsayımlar her madde ve her yanıtlayıcı birey için doğru olmadığında bireylerin bu maddelere verdikleri cevaplar farklı faktörlerden etkilenmekte ve bu durum elde edilen sonuçların yanlı (biased) olmasına neden olmaktadır. Ölçek maddelerine verilen

tepkileri etkileyerek uluslar arası düzeyde karşılaştırma sonuçlarında yanlılığına neden olan faktörlerden biri; ortalama başarının düşük olduğu okula giden öğrencilerin daha yüksek akademik benlik algısına sahip olması veya ortalama başarının yüksek olduğu okula giden öğrencilerin daha düşük akademik benlik algısına sahip olması anlamına gelen "büyük balık küçük göl etkisi (big fish little pond effect)" olabilir (Marsh, Seaton, Trautwein, Ludthke, Han, O'Mara ve Craven, 2008). Bir diğer sebep ise ülkeler arasındaki olası tepki stilleri (response style) farklılıklarıdır (Buckley, 2009). Cronbach'ın 1941 yılında tepki stili kavramını geliştirmesiyle birlikte daha çok ilgi odağı olan tepki stili, "maddenin içeriğine bağlı kalmadan anket maddelerine sistematik bir biçimde tepki verme eğilimidir" (Paulhus, 1991, s.17). Bireylerin sergilemiş olduğu tepki stili verilerde çeşitli psikometrik sorunlara neden olmaktadır (Bolt ve Newton, 2011). Daha açıkça belirtmek gerekirse, aynı düzeyde bilgiye, tutuma veya benzer kişilik özelliklerine sahip bireylerin test puanlarına sistematik hata karıştırarak farklı test puanları elde etmesine yol açarak test puanlarının geçerliliğini düşürmektedir (Cronbach, 1946).

Tepki stilleri, araştırma sonuçlarının geçerliliğini iki türlü tehdit etmektedir (Baumgartner ve Steenkamp, 2001). Bunlardan birincisi, araştırmada ele alınan değişkene ilişkin tek değişkenli dağılımı (univariate distribution) etkilemesidir. Daha açıkça belirtmek gerekirse ölçek puanlarına ilişkin ortalama ve varyansa hata karışmasına neden olur. Bu yüzden, tepki stilinin etkisi kontrol edilmediğinde, t-testi veya F-testi gibi karşılaştırmalı testlerden yanlış sonuçlar elde edilebilir (Cheung ve Rensvold, 2000). Bu durum özellikle uluslararası karşılaştırma çalışmaları bağlamında düşünüldüğünde, ülkelerin kendi tepki stillerindeki farklılaşmalardan ötürü, ölçülmek istenen özellikteki (substantive trait) farklılaşmalar hakkında yanlış yorumlamalar yapılabilir. Bu yüzden, kültürlerarası çalışmalarda tepki stillerinden kaynaklı yanlılığı dikkate almayan çalışmalar, gruplar arasındaki tutum puanlarındaki ortalama farklılığa dair yanlış çıkarımlarda bulunabilir (Harumi, 2011). Bir diğer ifade ile, tutum puanlarında gözlenen farklılıklar tepki stillerinden kaynaklı farklılıklar olmasına rağmen ölçülen özellikteki farklılıklarımsı gibi yorumlanabilir. Tepki stillerinin, ölçek puanlarının geçerliliği tehdit ettiği durumlardan bir diğeri de çok değişkenli dağılımı (multivariate distribution) etkilemesi ile ilgilidir. Özellikle değişkenler arasındaki korelasyonun değişmesine neden olarak korelasyona dayanan Cronbach alfa, regresyon analizi, faktör analizi, yapısal eşitlik modellemesi gibi çeşitli istatistiksel tekniklerden elde edilen sonuçların yanlı olması olmasına neden olur (Reynolds ve Smith, 2010). Daha açıkça

belirtmek gerekirse, verilen tepkinin maddenin içeriğinden bağımsız olması gerçekte ilişkisiz olan özellikler arasındaki ilişkilerin var olduğundan yüksek çıkmasına ve korelasyon matrisine dayalı olan çok değişkenli analizlerde gizil faktörlerin yanlış yorumlanmasına neden olabilir (Austin, Deary, Gibson, McGregor ve Dent, 1998). Tepki stillerinin araştırma sonuçlarında yanlılığa neden olması maddelerde değişen madde fonksiyonuna (DMF) odaklanan çalışmalarda da gözlenmektedir. Özellikle kültürlerarası karşılaştırmalarda kullanılan ölçek maddelerinin DMF göstermesi farklı gruplardaki cevaplayıcıların tepki stillerindeki farklılaşmayla ilgili olabilir (Holland ve Wainer, 2012). Wetzel, Böhnke, Carstensen, Ziegler ve Ostendorf (2013)'e göre tepki stillerinin etkisi kontrol edildiğinde bir maddenin DMF gösterip göstermeme durumu ve DMF gösterme düzeyi değişmektedir. Bu sebeplerden dolayı kültürlerarası karşılaştırmalarda tepki stillerinden kaynaklı oluşabilecek yanlılığa dikkat edilmesi gerekir.

Alanyazında, test puanlarının geçerliliğini olumsuz etkileyen çeşitli tepki stillerinin var olduğu belirtilmektedir. Uç tepki stili (extreme response style-UTS), tepki ölçeğinin iki uç noktasından birinde tepki verme eğilimidir (Baumgartner ve Steenkamp, 2001). En yaygın olarak kullanılan tepki stillerinden, kabullenici tepki stili (acquiescence response style-KTS), diğer tepki kategorilerine veya seçeneklere kıyasla daha çok "katılıyorum" veya "evet" cevabını verme eğilimidir (Greenleaf, 1992; Van Herk, Poortinga ve Verhallen, 2004). Orta nokta tepki stili (midpoint response style-OTS), tepki ölçeğinin orta noktalarında (kararsızım vb.) tepki verme eğilimidir (Baumgartner ve Steenkamp, 2001; Weijters, Geuens ve Schillewaert, 2008). Red edici tepki stili (disacquiescence response style-RETS) ise, kabullenici tepki stiline tam tersi olmakla beraber, daha çok "katılmıyorum" şeklinde tepki verme eğilimidir (Baumgartner ve Steenkamp, 2001). Tepki stillerine ilişkin şematik gösterim Şekil 1'de verilmiştir:



Şekil 1. Tepki Stilleri (Liu, 2015, s.3)

Bireylerin farklı tepki stilleri sergileme eğiliminde olmasının nedenleri incelendiğinde, bireylerde olumlu bir etki yaratma isteğine neden olan kendini gösterme güduları (self-presentational motives) (Leary, 1996; Zickar ve Drasgow, 1996), durağan kişilik özellikleri (Eid ve Rauber, 2000) veya anketin cevaplanmasının karmaşıklığını azaltma isteği gibi bireye özgü faktörlerin etkili olduğu görülmektedir. Bunların yanı sıra, çeşitli araştırmalarda (Harzing, 2006; Johnson, Kulesa, Cho ve Shavitt, 2005), sosyo-kültürel özelliklerden bazılarının bir kültürdeki bireylerin belirli tepki stillerini sergilemede etkili olduğu belirtilmiştir. Hofstede (2001) tarafından tanımlanan kültürel boyutlardan toplulukçuluk (collectivism) ve bireycilik (individualism), belirsizlikten kaçınma (uncertainty avoidance), güç mesafesi (power distance) ve erkeklik (masculinity) ve kadınlık (femininity) kültürlerin birbirinden farklı tepki stillerine sahip olmasına neden olabilir. Daha açıkça belirtmek gerekirse, toplulukçu kültürlerde yaşayan bireyler için gruba bağımlılık, kişiler arası ilişkiler ve grup dayanışması önemli olmakla birlikte bu bireyler davranışlarını ve tutumlarını toplumdaki normlara veya taleplere göre belirlemektedirler (Hofstede, 2001). Öte yandan, bireyci toplumdaki bireyler ise eylemleri üzerinde kontrole sahip, eylemlerinin sorumluluğunu alabilen ve işbirliğinden çok rekabet eden kişilerdir. Bu bağlamda, bireyci toplumdaki bireyler daha uçlarda tepkiler vermeye eğilimliyken (Johnson ve diğerleri, 2005) toplulukçu kültürlerdeki bireyler ise toplumdaki uyumun sağlanması açısından daha çok orta noktalarda tepki verme eğilimindedirler (Smith, 2004).

Kültürlerin farklı tepki stillerine sahip olmalarını etkileyen faktörlerden bir tanesi de bu kültürlerin toplumdaki belirsizliği tolere edebilme ve kendilerini rahat hissedebilme dereceleri anlamına gelen “belirsizlikten kaçınma”dır (Hosftede, 2001). Belirsizlikten kaçınmaya eğilimli olan toplumlar, mutlak doğrulara inanmakta ve belirsizliği azaltmak için için katı kurallar ve kanunlar koymaktadır. Belirsizlikten kaçınma eğilimi düşük olan toplumlar ise daha fazla risk alabilir ve farklı fikirlere karşı daha hoşgörülüdür. Bu özelliklerinden dolayı, belirsizlik kaçınma eğilimi yüksek olan toplumlar, toplumda var olan belirsizliği azaltmak için daha çok kabullenici tepki verme eğilimindedirler (Harzing, 2006). Güç mesafesi ise, bir toplumdaki güçsüz olan bireylerin gücün eşit olarak dağıtılmadığını kabullenme dereceleridir (Hosftede, 2001). Güç mesafesi yüksek olan toplumlarda, eşitsizliğin toplumdaki düzenin sağlanabilmesi açısından gerekli olduğu düşüncesi hakim olmakla beraber yüksek güce sahip bireylerin görüşleriyle aynı görüşlere sahip olmak önemlidir. Düşük güç mesafesi olan toplumlarda ise her bir bireyin görüşlerine saygı duyulur. Dolayısıyla, güç mesafesi düşük olan toplumlardaki bireyler daha çok ortalarda tepki sergilerken yüksek güç mesafesinin görüldüğü toplumlar daha çok uçlarda tepki verme eğilimindedirler (Johnson ve diğerleri, 2005). Bir toplumun erkeklik veya kadınlık vasfının baskın olması da bireylerin farklı tepki stilleri sergilemesinde etkili olan durumlardan bir diğeridir. Erkeklik vasfı yüksek olan toplumda cinsiyetlere yüklenen roller farklılaşmaktadır. Bu toplumlarda, erkekler girişken, cesaretli ve maddi başarıya odaklanan ve gücü elinde tutan kişiler olarak tanımlanırken kadınlar ise duyarlı, ılımlı, hayatın kalitesine odaklanan cinsiyet olarak kabul edilir. Öte yandan, kadınlık vasfının egemen olduğu toplumlarda ise erkeğe ve kadına yüklenen roller arasında net bir ayrım yoktur. Erkeklik özelliğinin hakim olduğu kültürlerde güç ve istikrarın ön planda olması sebebiyle bireyler daha çok uçlarda tepki verme eğilimindedir (Johnson ve diğerleri, 2005). Özetle duyuşsal özelliklerin (kişilik, tutum, algı) ölçüldüğü ve bu kapsamda kültürlerarası karşılaştırmaların yapıldığı araştırmalarda, kültürlerin baskın özelliklerine göre cevaplayıcıların farklı tepki stillerini sergilediği ve farklı tepki stillerine odaklanıldığı görülmektedir

Bu çalışma kapsamında, (i)TIMSS 2015 uygulamasında öğrencilerin tutumunu ölçmek amacıyla kullanılan ölçeklerin 4’lü likert tipinde olmasından dolayı daha fazla sayıda tepki kategorisine sahip olmaması, (ii) diğer tepki stillerine kıyasla bireylerin cevaplarındaki varyansı daha fazla açıklaması (De Jong, Steenkamp, Fox ve Baumgartner, 2008), (iii) etkisinin çeşitli yöntemlerle daha güvenilir bir şekilde kontrol

altına alınabilmesi (Engle, 2016), (iv) kültürlerarası karşılaştırmalarda uç tepki stiline sergilenmesindeki farklılıkların sıklıkla dile getirilmesi ve (v) eğitimsel başarıyla yakından ilgili olması (Lu ve Bolt, 2015) gibi nedenlerden dolayı uç tepki stili üzerinde odaklanılmıştır.

Kültürlerarası karşılaştırmalarda bireylerin vermiş olduğu tepkilerin çeşitli yollarla tepki stillerinden etkilenmesi, sonuçların sistematik bir şekilde değişmesine neden olması, uluslararası değerlendirme sonuçlarına ve bunlara bağlı olarak ulusal düzeyde verilecek eğitim kararları ve politikalarına ilişkin hatalı çıkarsamalara ve yanlış uygulamalara yol açma olasılığı bu tepki stillerinin tutum gibi psikolojik yapılara ait puanlardaki etkilerinin düzeltilmesini önemli kılmaktadır. Bu gerekçelerle, alanyazında bu sorunu çözmeye yönelik çeşitli yöntemler sunulmuştur. Alanyazında, kültürlerarası karşılaştırmalarda yanlılığa neden olan ve bu çalışmada da ele alınacak olan tepki stillerinden UTS'nin etkisini belirlemek için sıklıkla önerilen yöntemler; Ad Hoc Uç Tepki Stili Kestirimlerine (Ad Hoc Extreme Response Style Measures) ve örtük değişken modellerine dayalıdır. Bu yöntemlerden ad-hoc UTS kestirimlerini kullanarak tepki stillerini ele alan araştırmalar sıradaki kısımda özetlenmiştir.

Ad-Hoc UTS Kestirimleriyle ilgili Yapılan Araştırmalar

Ad-Hoc UTS kestirimlerinde, bireylerin tepki ölçeğinin uç kategorilerindeki cevaplarının frekansı veya madde puanlarının standart hatası kullanılarak UTS indeksi oluşturulmaktadır. Alanyazında, UTS indeksi oluşturan ve bu indeksi kullanarak çeşitli değişkenler açısından farklılık olup olmadığını varyans analiziyle belirlemiş araştırmalar mevcuttur (Bachman ve O'Malley, 1984; Bachman, O'Malley ve Freedman-Doan, 2010; Hui & Triandis, 1989; İlgün Dibek, Yavuz ve Çokluk Bökeoğlu, 2018; Marín, Gamba, & Marín, 1992). Varyans analizi gerçekleştiren bu çalışmalardan bazıları bireylerin tepki stillerini çeşitli faktörler (ırk, cinsiyet, ailevi faktörler, kültürel yapı vb.) ve kişilik özellikleriyle olan ilişkileri bağlamında incelemiştir. Örneğin, ulusal alanyazında İlgün Dibek, Yavuz ve Çokluk Bökeoğlu (2018) tarafından yapılan araştırmada PISA 2012 uygulamasına katılan ülkelere her bir grupta üçer ülke olmak üzere ülkelerin mutluluk ve başarı durumlarına göre üç ülke grubu (*mutsuz ama başarılı ülkeler*: Kore, Estonya, Finlandiya, *ne mutlu ne de başarılı ülkeler*: Türkiye, Birleşik Arap Emirlikleri (Dubai) ve Sırbistan ve *mutlu ama başarısız ülkeler*: Kolombiya, Endonezya, Peru) belirlenerek ülke gruplarının çeşitli tepki stillerini (KTS, UTS, RETS) gösterme eğilimindeki farklılıklar varyans analizi ile incelenmiştir. Araştırmanın

sonunda UTS ve RETS gösterme eğilimleri açısından üç gruptaki ülkelerin farklılaştığı, KTS gösterme eğilimleri açısından mutsuz ama başarılı ülkeler ile ne mutlu ne de başarılı ülkelerin benzerlik gösterdiği bulunmuştur.

Bachman ve O'Malley (1984), farklı ırklardaki bireylerin tepki stillerini sergileme eğilimlerinin cinsiyete, ailevi faktörlere, zamana ve kişilik özelliklerine göre farklılaşıp farklılaşmadığını araştırmışlardır. Siyahi ve beyaz lise öğrencilerinin tepki stillerini sergileme eğilimlerinin araştırıldığı çalışmanın sonunda, siyahi öğrencilerin daha çok en olumlu uç olmak üzere uç kategorilerde tepki verdiği, öz saygılarının zaman içerisinde değişmemesine bağlı olarak uç tepki verme eğiliminin de değişmediği bulunmuştur. Bunun yanı sıra ailevi faktörler ile tepki stillerine sahip olma arasında manidar bir ilişki olmadığı ve cinsiyete göre tepki stillerini sergileme eğilimlerinin farklılaşmadığı belirtilmiştir.

Bachman, O'Malley ve Freedman-Doan (2010), farklı etnik kökenlerde, sınıf seviyelerinde ve cinsiyette bireylerin tepki stillerini sergileme eğilimlerinin nasıl değiştiğini belirlemeyi amaçlamıştır. Her bir sınıf seviyesinde ve etnik kökende erkeklerin kızlara göre uçlarda tepki gösterme eğilimlerinin daha fazla olduğu, 8.sınıftan 12.sınıfa doğru ilerledikçe uç tepki stili sergileme eğilimlerinin azaldığı, Afrika kökenli Amerikalıların, İspanyol kökenli ve Asya kökenli Amerikalılara göre uç kategorilerde tepki gösterme eğilimlerinin fazla olduğu bulunmuştur.

Model temelli olmayan Ad-Hoc UTS kestirimi yönetiminin en büyük sınırlılığı ölçülmek istenen özelliğin etkisi ile tepki stillerinin etkisini birbirinden ayıramamasıdır. Bir diğer anlatımla, tepki ölçeğinin sürekli olumlu uç noktasını seçen bir birey gerçekten de ölçülen özelliğe yüksek düzeyde sahip olabilir. Bu noktada, alan yazında bu karışıklığı önlemek amacıyla ölçme aracına tepki stilini ortaya koymayı amaçlayan yeni maddeler ekleyen araştırmalar yer almaktadır. Örneğin, Greanleaf (1992b) araştırmasında, bireylerin UTS eğilimlerini ölçmek amacıyla ölçülmek istenen özellik açısından düşük düzeyde madde-toplam test korelasyonu gösteren maddeler eklemiştir. Benzer şekilde başka çalışmalarda da (Arce-Ferrer, 2006; Clarke III, 2000) aynı yaklaşım benimsenmiştir. Greanleaf tarafından geliştirilen ölçeğin kullanıldığı üniversite öğrencileri üzerinde gerçekleştirilen bir başka çalışmada da farklı uluslar ve kültürler arasında UTS gösterme eğilimleri açısından farklılık olduğu bulunmuştur (Clarke III, 2000). Buna göre İspanyollar İspanyol olmayanlara, siyahi bireyler siyahi olmayanlara ve Fransız öğrenciler Avustralyalı öğrencilere kıyasla daha yüksek düzeyde uç tepki stili sergilemektedir.

ArcFerrer (2006), Meksika’da kırsal ve kentsel yerleşim yerlerinde yaşayan lise öğrencilerinin UTS tepki stilleri açısından farklılaşıp farklılaşmadığını birbiri ile düşük korelasyon gösteren maddelerden oluşturduğu bir ölçek ile incelediği çalışmasında farklı yerleşim yerinde yaşayan öğrencilerin UTS gösterme eğilimlerinin farklılaştığını, bazı maddelere kırsal kesimde yaşayanların daha uçlarda tepki gösterdiği bazılarında ise kentsel bölgelerde yaşayanların uç tepkiler sergilediğini bulmuştur.

Alanyazındaki bazı çalışmalarda ise regresyon modelleri (Harzing, 2006) veya hiyerarşik lineer modeller (Baumgartner & Steenkamp, 2001; Johnson, Kulesa, Llc, Cho, & Shavitt, 2005) kullanılarak birey düzeyinde tepki stilleri ile ülke düzeyindeki özellikler arasındaki ilişkiler incelenmiştir. Harzing (2006), kuzey, güney, doğu Avrupa, kuzey Amerika, Latin Amerika ve Asya ülkeleri dahil olmak üzere 26 ülkedeki üniversite öğrencilerinin çeşitli tutum ifadelerine verdikleri yanıtlardan yola çıkarak farklı tepki stillerini (UTS, KTS, RETS, OTS) yansıtan indeksler oluşturmuştur. Güç mesafesi, bireycilik/toplulukçuluk ve belirsizlikten kaçınma boyutlarının ülke düzeyinde bağımsız değişken olarak ele alındığı bu çalışmada, güç mesafesi fazla olan ülkelerin UTS'yi OTS'ye kıyasla daha fazla sergilediği, toplulukçuluğun bireylerin KTS'yi ve OTS'yi sergileme eğilimlerinde etkili olduğu, belirsizlikten kaçınma boyutunun KTS ile yüksek düzeyde ilişkili olduğu ve son olarak toplulukçuluğun diğer kültürel boyutlara kıyasla bireylerin tepki stilleri üzerinde en fazla etkiye sahip olduğu bulunmuştur.

Baumgartner ve Steenkamp (2001), gerçekleştirdiği hiyerarşik modelleme analizinde birey düzeyi ve ülke düzeyine ek olarak diğer çalışmalardan farklı olarak ölçeklerdeki olumlu-olumsuz madde sayılarının denk olma durumunu da ele alarak ölçek düzeyini de dahil etmiştir. Beş farklı tepki stilinin (KTS, RETS, UTS, OTS, noncontingent responding (NCR)) etkisinin ölçekten ölçeğe ve ülkeden ülkeye değişip değişmediğini belirlemek amacıyla 11 ülkedeki (Belçika, Danimarka, Fransa, Almanya, İngiltere, Yunanistan, İrlanda, İtalya, Hollanda, Portekiz ve İspanya) bireylerin beş farklı ölçeğe verdiği cevaplardan yola çıkarak üç düzeyli hiyerarşik lineer modelleme analizi gerçekleştirmişlerdir. Birinci düzeyin birey düzeyi, ikinci düzeyin ölçek düzeyi ve üçüncü düzeyin ülke düzeyi olduğu analize göre beş farklı tepki stilinin ölçek puanları üzerindeki etkisi ölçekten ölçeğe göre değişmektedir. Fakat ülkeden ülkeye göre tepki stillerinin ölçek puanları üzerindeki etkisindeki farklılık ihmal edilebilir. Aynı zamanda, ölçekteki olumlu ve olumsuz maddelerin sayısının birbirine denk olma

derecesi tepki stillerinin ölçek puanları üzerindeki olumlu ya da olumsuz yöndeki etkisini deęiřtirdięi sonucuna ulařılmıştır.

Johnson ve dięerleri (2005), ad-hoc UTS kestirimleriyle tepki stillerini inceleyen arařtırmalara benzer olarak bireylerin tepki stilleri üzerinde kültürel yönelimlerin etkisini incelemiřtir. Fakat bu çalıřmaların ötesine geçerek birey düzeyindeki tepki stilleri ile ülke düzeyindeki kültürel yönelimlerin olduęu 19 ülke (Avustralya, Belçika, Brezilya, Çek Cumhuriyeti, Almanya, Macaristan, Hindistan, Japonya, Malezya, Potekiz, Türkiye, İngiltere, Meksika, Filipinler, Polonya, Singapur, Hong Kong, Fransa ve İtalya) ile hiyerarřık lineer modelleme analizi gerçekleřtirmiřtir. Çalıřmanın sonucunda güç mesafesinin fazla ve erkeklik vasfının baskın olmasının uç tepki stili ile olumlu yönde iliřki gösterdięini bulmuşlardır. Öte yandan, bireycilik, belirsizlikten kaçınma, güç mesafesi ve erkeklik vasfının yüksek olması KTS ile negatif yönde iliřki göstermektedir.

Ad-hoc UTS kestirimlerini, regresyon veya hiyerarřık lineer modelleme analizini kullanan arařtırmalar, bireylerin ölçek puanlarındaki yanlılıęın üzerinde tepki stillerinin doęrusal bir etkisinin olduęunu varsaymıřtır. Aynı zamanda, bu yöntemler ölçülmek istenen özellięin farklı düzeylerinin tepki stillerine iliřkin oluřturulan indekslerle ne düzeyde bir iliřkisinin olduęunu ve uçlarda verilen tepkilerin gerçekten ölçülen özellięinin uç sınırlarıyla mı yoksa tepki stilleriyle mi iliřkili olduęunu belirlemede yetersiz kalmaktadır. Bu çalıřmaların yanı sıra örtük deęiřken modelleriyle ilgili yapılan arařtırmalar sıradaki kısımda özetlenmiřtir.

Örtük Deęiřken Modellerine dayalı UTS belirleme ile İlgili Yapılan Arařtırmalar

Örtük deęiřken modelleriyle ilgili yapılan arařtırmalarda UTS'yi belirlemede betimsel istatistiklere dayalı olarak UTS indeksi oluřturma yoluna gidilmemektedir. Bu arařtırmaların bir kısmının doęrulayıcı faktör analizi (DFA) bir kısmının ise gizil sınıf analizi (GSA) gerçekleřtirdięi görülmektedir. Bu arařtırmalardan DFA gerçekleřtiren arařtırmalardan bazıları çoklu grup DFA gerçekleřtirerek her bir grupta ayrı ayrı DFA modelini test etmekte, deęiřkenleri sürekli deęiřken olarak ele almakta ve sadece ölçülmek istenen özellięi ayrı bir faktör dahil etmektedir. Örneęin, farklı grupların tutumlarını karřılařtırmada tepki stillerinin etkisini belirlemek amacıyla yürütölen çalıřmada (Cheung ve Rensvold, 2000), faktör yükleri, kesiřimler, gizil deęiřken ortalamaları gruplar boyunca deęiřmez olduęunda gruplar arasında ölçülmek istenen

özelliğın denk olduđu ve grup karşılařtırmalarına tepki stillerinden kaynaklı yanlılığın karıřmadıđı vurgulanmıřtır.

Bazı arařtırmalar ise tepki stilini de ayrı bir faktör olarak ele almaktadır. Alanyazın incelendiğinde bu yöntemi kullanan arařtırmaların çoğunun KTS'nin etkisini belirlemeye odaklandıđı görölmektedir. Bu çalıřmalardan, Welkenhuysen-Gybels, Billiet, ve Cambre (2006) tarafından yürütölen arařtırmada, KTS'nin ayrı bir faktör olarak ele alınmadıđı durumda faktör yüklerinin gruplar boyunca deđiřmezliđine olan etkisi belirlenmiřtir. Bu amaçla dokuz batı Avrupa ölkesindeki bireylerin çeřitli tutum ifadelerine verdikleri cevaplar incelenmiřtir. KTS ayrı bir faktör olarak ele alınmadıđında faktör yüklerinin ve standart hataların yanlı olarak kestirilmesine neden olduđu vurgulanmıřtır. Benzer řekilde Billiet ve McClendon (2000) yürötmüş oldukları arařtırmada, KTS'nin ayrı bir faktör olarak ele alınmasının modelin veriyle daha iyi uyumu ve daha basit, teorik olarak daha anlamlı modellerin geliřtirilmesini sađladıđını vurgulamaktadır.

Alanyazında tepki stillerini gizil sınıf analizi ile inceleyen arařtırmalar incelendiğinde ise UTS'nin ölçme düzeyini kategorik olarak ele almaktadır. Bu arařtırmalardan bir kısmının (Moors, 2003, 2004; van Rosmalen, van Herk, & Groenen, 2010) gizil sınıf multinominal lojistik modellerle bir kısmının (Eid & Rauber, 2000) ise gizil sınıf karma modellerle analizlerini gerçekteřtirdikleri görölmektedir. Van Rosmalen, van Herk ve Groenen (2010) beř farklı ölkedeki (Fransa, Almanya, İtalya, İspanya ve İngiltere) bireylerin tepkileri üzerinde maddenin içeriğinin, bireye iliřkin özelliklerin (ölke, yař ve eđitim düzeyi) ve tepki stillerinin etkisini belirlemeyi amaçlamıřlardır. Arařtırmacılar, bađımlı ve bađımsız deđiřkenlerin kategorik olduđu gizil sınıf multinominal lojistik model geliřtirmişlerdir. Arařtırmalarında, bireylerin tepki stili davranıřlarındaki farklılıđı en fazla açıklayan faktörün maddenin içeriđi olduđu vurgulanmıřtır.

Moors (2004) Belçika'daki Türk ve Faslı bireylerin etnik ayrımcılıđa iliřkin tutum maddelerini cevaplar ken uç tepki stili sergileme eđilimlerini gizil sınıf faktör analizi yaklařımı ile belirlemiřtir. Arařtırmada, farklı ırklardaki bireylerin uç tepki stilini sergileme eđilimlerinin farklılařtıđı belirtilmiřtir. Maddelerin olumlu veya olumsuz bir yönde tutum belirtiđine bakmaksızın Türk bireylerin daha çok "katılıyorum" tepki kategorisini Faslı bireylerin ise daha çok "kesinlikle katılmıyorum" tepki kategorisini seçtiđi bulunmuřtur.

Eid ve Rauber (2000), alt grupların tepki verme eğilimlerinin başka değişkenler açısından farklılaşıp farklılaşmadığını ölçme değişmezliği çerçevesinde çoklu karma Rasch modeli (polytomous mixed Rasch model) kullanarak inceledikleri araştırmalarında iki sınıf olduğunu bulmuşlardır. Sınıflardan birinin tutum ölçeğinin farklı tepki kategorilerini seçtiği öte yandan diğer sınıfın ölçeğin uç tepki kategorilerinde tepki gösterme eğiliminde olduğu belirtilmiştir.

Alanyazında, örtük değişken modellerinden olan Çok Boyutlu Madde Tepki Kuramı (ÇBMTK) modeliyle tepki stillerini ele alan çalışmalar da mevcuttur. Bu çalışmalardan, Bolt ve Johnson (2009), bireylerin sigara içmeye yönelik motivasyonları ve maddelerin DMF gösterme durumları üzerinde UTS'nin etkisini belirlemek ve kontrol altına almak için ÇBMTK kullanmıştır. Araştırmacılar, sigara içmeye yönelik motivasyonun düzeyine göre uç tepki stiline de etkisinin değiştiğini, bireylerin tepki stiline göre düzeltilmiş puanlarının düzeltilmemiş puanlardan farklılaştığını belirtmişlerdir. Bunun yanı sıra, multinomial tepki modeline kıyasla nominal tepki modeli ile daha fazla sayıda maddenin DMF gösterdiği ve bunun nedeni olarak ise gruplarının tepki stillerindeki farklılık olduğu belirtilmiştir. Aynı araştırmacılar bir başka çalışmada (Johnson ve Bolt, 2010) ise bireylerin tepki stillerindeki farklılığı açıklamada çok boyutlu multinomial logit madde tepki modelinin (multidimensional multinomial logit item response model) nasıl kullanılacağını açıklamış ve tepki stillerine ilişkin faktör ile ölçülmek istenen asıl özelliğe ilişkin faktörün birbiriyle ilişkili olduğunda da bu modelin etkili bir şekilde kullanılabileceğini belirtmişlerdir.

Bolt ve Newton (2011), Bolt ve Johnson (2009)'ün yapmış olduğu çalışmayı genişleterek araştırmacıların geliştirdiği modellerle UTS'nin etkisini daha iyi kestirebilmek amacıyla PISA 2006 uygulamasına katılan Amerikalı öğrencilerin birden fazla ölçeğe verdikleri cevapları eş zamanlı analiz etmiştir. Araştırmanın sonunda birden fazla ölçekle UTS'nin etkisinin daha iyi kestirilebildiği bulunmuştur. Aynı zamanda, farklı ölçeklerle ölçülen özellikler arasında ilişki olmasının UTS'den kaynaklı yanlılığı düzeltmede bir etkisinin olmadığı bulunmuştur.

Lu (2012), araştırmasında ülke düzeyinde görülen tutum-başarı paradoksunda UTS'nin rolünü belirlemeyi amaçlamıştır. Bu doğrultuda PISA 2006 uygulamasına katılan ülkelerin tutum ölçeklerine verdikleri cevaplar üzerinde UTS'nin etkisini belirlemek ve düzeltmek amacıyla ÇBMTK modelini çok düzeyli bağlamda analiz etmiştir. ÇBMTK modeliyle kestirdiği parametrelerin doğruluğunu test etmek amacıyla ise simülasyon analizi gerçekleştirmiştir. Araştırmanın sonunda, ülke düzeyinde başarısı

yüksek olan öğrencilerin olumsuz yönde tutuma sahip olmasında UTS'nin etkisinin olduğu bulunmuştur. Bunun yanı sıra simülasyon analizi sonuçlarına göre ise çok düzeyli ÇBMTK modeli, ülke düzeyinde ortalamaların kestiriminde ve ülke için kovaryans yapılarının (within-covariance structures) belirlenmesinde etkili iken ülke düzeyinde başarı ile tutum arasındaki ilişkilerin ve ülkeler arası kovaryans yapılarının (between-covariance structures) belirlenmesinde nispeten daha az etkilidir.

Özetle, uluslararası alanyazında, psikoloji, işletme, pazarlama gibi alanlarda çeşitli amaçlarla duyuşsal özelliklerin ölçüldüğü durumlarda verilen tepkilerdeki uç tepki stilinden kaynaklı yanlılığa ilişkin kültürlerarası karşılaştırma yapan, çoğunlukla toplum yapısına, farklı ırklardaki değişime odaklanan çeşitli araştırmaların olduğu görülmektedir. Öte yandan eğitim alanında tutum ile başarı arasındaki ilişkiyi açıklamada etkisinin belirlenmesine odaklanan yeterince çalışma olmadığı ifade edilebilir. Aynı zamanda bu çalışmaların çoğunun kullandıkları yöntemlerden kaynaklı olarak UTS'nin neden olduğu yanlılığı belirlemede ve bu hatayı düzeltmede bazı sınırlılıklarının olduğu belirtilebilir. ÇBMTK modelini kullanan araştırmaların ise UTS'nin etkisinin kestirilmesi açısından diğer yöntemlerin eksik kalan yönlerini büyük ölçüde telafi ettiği görülse de tepki stillerinin etkisini çok düzeyli bağlamda incelemeye bu çalışmalar yetersiz kalmaktadır. Dolayısıyla eğitim alanındaki bu eksikliğin giderilmesine katkı getirmek ve daha geçerli çıkarımların yapılmasını sağlamak amacıyla bu çalışmada uç tepki stilinin etkisini belirleme ve düzeltme olanağı sağlaması nedeniyle çok düzeyli çok boyutlu madde tepki kuramı modeli kullanılmıştır.

1.2 Amaç

Bu çalışmanın amacı, TIMSS 2015 uygulamasına katılan ülkelerin matematik başarıları ile matematiğe yönelik tutumlarının ölçülmesinde tutum-başarı paradoksuna neden olabilecek tepki stillerinden uç tepki stilinin etkisinin etkisini ve bu etkiyi farklı ülkelerdeki sekizinci sınıf öğrencilerinin tutum ölçeklerine verdikleri cevaplar üzerinde düzelterek matematiğe yönelik tutum ve matematik başarıları arasındaki ilişkiyi belirlemektir. Bu amaç doğrultusunda bu ilişki hem birey hem de ülke düzeyinde araştırılmıştır. Bu doğrultuda araştırmada yanıt aranan sorular şunlardır:

1. TIMSS 2015 uygulamasına katılan tutum-başarı paradoksunun görüldüğü ülkeler arasında tutumla ilgili farklı değişkenlerin araştırıldığı ölçeklerden (“matematiği öğrenmeyi sevmesi”, “öğrencinin matematikte kendine güvenmesi”, ve “matematiğe değer vermesi”) ad-hoc kestirimleriyle elde edilen UTS puanları açısından manidar farklılık var mıdır?

2. TIMSS 2015 uygulamasına katılan tutum-başarı paradoksunun görüldüğü ülkelerdeki öğrencilerin

- a) Tutumla ilgili ölçeklere verdikleri cevaplara ilişkin oluşturulan UTS'nin bir boyutolarak ele alındığı ve alınmadığı modellerde, model- veri uyumu nasıldır?
- b) Tutumla ilgili farklı ölçeklerden elde edilen toplam ölçek puanlarına UTS'ye bağlı olarak karışan yanlılığın miktarı ve yönü farklı UTS değerlerine göre nasıl değişmektedir?
- c) UTS'nin etkisine göre düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasındaki farklar manidar mıdır?
- d) UTS'nin etkisine göre düzeltilmemiş ve düzeltilmiş tutum puanlarının matematik başarısıyla ülke içi ve ülkeler arası ilişkileri nasıl değişmektedir?
- e) UTS'nin etkisine göre düzeltilmemiş ve düzeltilmiş ortalama tutum puanlarının matematik başarısıyla ülke içi ve ülkeler arası ilişkileri arasındaki farklar manidar mıdır?

1.3 Önem

Uluslararası alanyazınında, psikoloji, işletme gibi alanlarda tepki stillerinden kaynaklı yanlılığa ilişkin kültürlerarası karşılaştırma yapan çok sayıda araştırma (Khorramdel ve von Davier, 2014; van Rosmalen, van Herk ve Groenen, 2010; Weijters, 2006) olmasına karşın eğitim alanında bu konu üzerinde yeterince odaklanılmadığı ve hatta ulusal alanyazında tepki stillerini ele alan uluslararası karşılaştırma yapan bir çalışmanın olduğu görülmektedir. Diğer alanlarda tepki stilinin neden olduğu problemlerin farkında olunması öte yandan eğitimde ise bu problemlerin göz ardı edilmesi eğitim sisteminde yanlış uygulamalara neden olabilir. Bu bakımdan bu çalışma eğitim alanında özellikle ulusal alanyazında özgün bir çalışma olması ve tepki stili tehditine dikkat çekmesinden dolayı önemlidir.

Alanyazında var olan çalışmalarının çoğu ise tepki stillerinin etkisini belirlemeye odaklanmasına karşın etkisini düzeltmeye odaklanmamıştır. Halbuki eğitim alanında, tepki stillerinin ölçme konusu olan özellik üzerindeki etkileri dikkate alınmadığında uluslar arası değerlendirmeler doğrultusunda yapılan çıkarımların geçerliliği tehlikeye düşmektedir. TIMSS gibi geniş ölçekli, çeşitli ülkelerin katıldığı çalışmaların sonuçlarının, ülkelerin eğitim sistemlerinin güçlü ve zayıf yönlerinin belirlenmesini sağlayarak eğitim programlarının düzeltilmesi gereken yönleri açısından eğitim politikalarına yön verdiği ve buna göre gerekli uygulamaların gerçekleştirildiği (Bos, 2002) göz önünde bulundurulduğunda bu çalışmaların sonuçlarının yansız bir şekilde sunulması doğru uygulamaların gerçekleştirilmesi açısından önemlidir. Bu bağlamda, eğitim araştırmalarından elde edilen çıkarsamaların daha geçerli olmasını sağlamak amacıyla, bu çalışma, öğrencilerin ülkeler içinde yuvalandığı TIMSS 2015 uygulamasında, ülkelerin sergilediği tepki stillerinin etkisini çok düzeyli bağlamda kontrol ederek eğitim alanyazınındaki bu eksikliğin giderilmesine katkı sağlaması açısından değerlidir. Aynı zamanda, mevcut çalışma betimsel istatistikleri veya tepki stilleri farklılıklarına ilişkin post hoc karşılaştırmaları sunan çalışmaların aksine model temelli bir yaklaşım sunmaktadır. Bu bakımdan bu çalışma ölçülmek istenen özelliğin dışında araştırma sonuçlarına karışarak ölçme kesinliğini etkileyen tepki stillerinin miktarını belirleme ve düzeltme imkanı sağlaması, öğrencilerin tutum gibi psikolojik özelliklerinin doğru olarak belirlenmesine imkan sunması, bu belirlemede kullanılacak yöntem önermesi açısından önemlidir. Bununla birlikte bu çalışma UTS'nin etkisine ilişkin geleceğe yönelik bilgilendirici öneriler vermesi bakımından değerlidir.

Öğrencinin derse yönelik tutumu ile dersteki başarısı arasında pozitif yönde bir ilişki olduğunu belirten kültürlerarası karşılaştırma araştırmaları ve teoriler (Bandura, 1994; Marsh, Trautwein, Lüdtke, Köller, ve Baumert, 2005), her bir okul sisteminin kendine özgü sosyal, ekonomik ve kültürel anlamda farklılaşmasını göz önünde bulundurmamaktadır. Bu yüzden, bu araştırmalar ve teoriler başarı ile tutum arasındaki ilişkiyi ve bu ilişkinin yönünün ülkeden ülkeye göre değişmesini açıklamada yetersiz kalmaktadır (Shen ve Tam, 2008). Bu bağlamda, öğrencilerin başarıları ile kendilerine yönelik algıları ve tutumları arasındaki ilişkilerin çok düzeyli analizlerle ülke düzeyinde (ülkeler arası) incelenerek karşılaştırılmasının, bireysel düzeyde bilgi sağlayan teorilerin de (örn. motivasyon teorileri ve öz-yeterlilik teorileri) ötesine geçmesi nedeniyle bu çalışmadan elde edilecek sonuçlar tutumun başarı ile olan ülke düzeyinde ilişkilerin anlaşılması ve ilişkilere dair geçerli yorumlamalar yapılabilmesine katkı sağlaması

açısından değerlidir. Bu açıdan, araştırmadan elde edilen sonuçların ülke düzeyinde tutum ile başarı arasındaki negatif yöndeki ilişkinin açıklanmasına yardımcı olacağı düşünülmektedir.

1.4 Sınırlılıklar

1. TIMSS 2015'de uygulanan matematiğe yönelik tutumu ölçen maddelerin dörtten fazla sayıda tepki kategorisine sahip olmaması ve maddelerde orta tepki kategorisi bulunmaması nedeniyle bu çalışma sadece uç tepki stilinin etkisinin modellenmesiyle sınırlıdır.
2. Kullanılan değişken sayısının ve buna bağlı olarak madde sayısının fazla olması, dahil edilen değişkenlerin eş zamanlı modellenmesi nedeniyle modellerin karmaşıklığını azaltmak amacıyla bu çalışma öğrencilerin matematik başarılarını temsil eden beş olası değerlerin ortalama değeriyle sınırlıdır.
3. Çalışmada kayıp veri sayısının fazla olması nedeniyle çoklu atama gibi yöntemlerin verinin yapısını bozma ihtimaline karşın kayıp veriler silindikten sonra seçkisiz olarak belirlenen alt örneklemelerle sınırlıdır.
4. UTS'nin bir boyut olarak dahil edilmediği ve dahil edildiği ÇBMTK modellerinin karmaşık olması nedeniyle analizlerin gerçekleştirildiği bilgisayarın kapasitesine bağlı olarak iterasyon sayısı ulaşılan maksimum iterasyon sayısı olan 8000 ile sınırlandırılmıştır.

1.5 Kısaltmalar

ÇBMTK : Çok Boyutlu Madde Tepki Kuramı

DFA : Doğrulayıcı Faktör Analizi

DMF : Değişen Madde Fonksiyonu

GSA : Gizil Sınıf Analizi

KTS : Kabullenici Tepki Stili

OTS : Orta Nokta Tepki Stili

PISA : Uluslararası Öğrenci Başarılarını Değerlendirme Programı (Programmes for International Students' Achievement Assessment)

RETS : Red Edici Tepki Stili

TIMSS : Uluslararası Matematik ve Fen Bilgisi Araştırma Projesi (Trends in International Mathematics and Science Study)

UTS : Uç Tepki Stili

1.6 Tanımlar

Paradoks: Çelişkidir. Bu çalışmada ise öğrencilerin derse yönelik tutumları ile dersteki başarıları arasında beklentilerin aksine çelişkili bir ilişkinin olması anlamına gelmektedir.

Başarı-tutum paradoksu: Öğrencilerin derse yönelik tutumları ile dersteki başarıları arasında birey düzeyinde pozitif grup düzeyinde ise negatif yönde bir ilişkinin olmasıdır (Van de ve diğerleri, 2012). Bu çalışmada ise bu paradoks söz konusu ilişkilerin ülke içinde pozitif ülkeler arasında negatif yönde olması anlamına gelmektedir.

Uç tepki stili: Tepki ölçeğinin iki uç tepki kategorisinden birinde tepki verme eğilimidir (Baumgartner ve Steenkamp, 2001). Bu çalışmada ise öğrencilerin tepki ölçeğinin uç noktaları olan “1” veya “4” tepki kategorisini seçme eğilimidir.

BÖLÜM 2

2.1 Yöntem

2.1.1 Araştırma Modeli

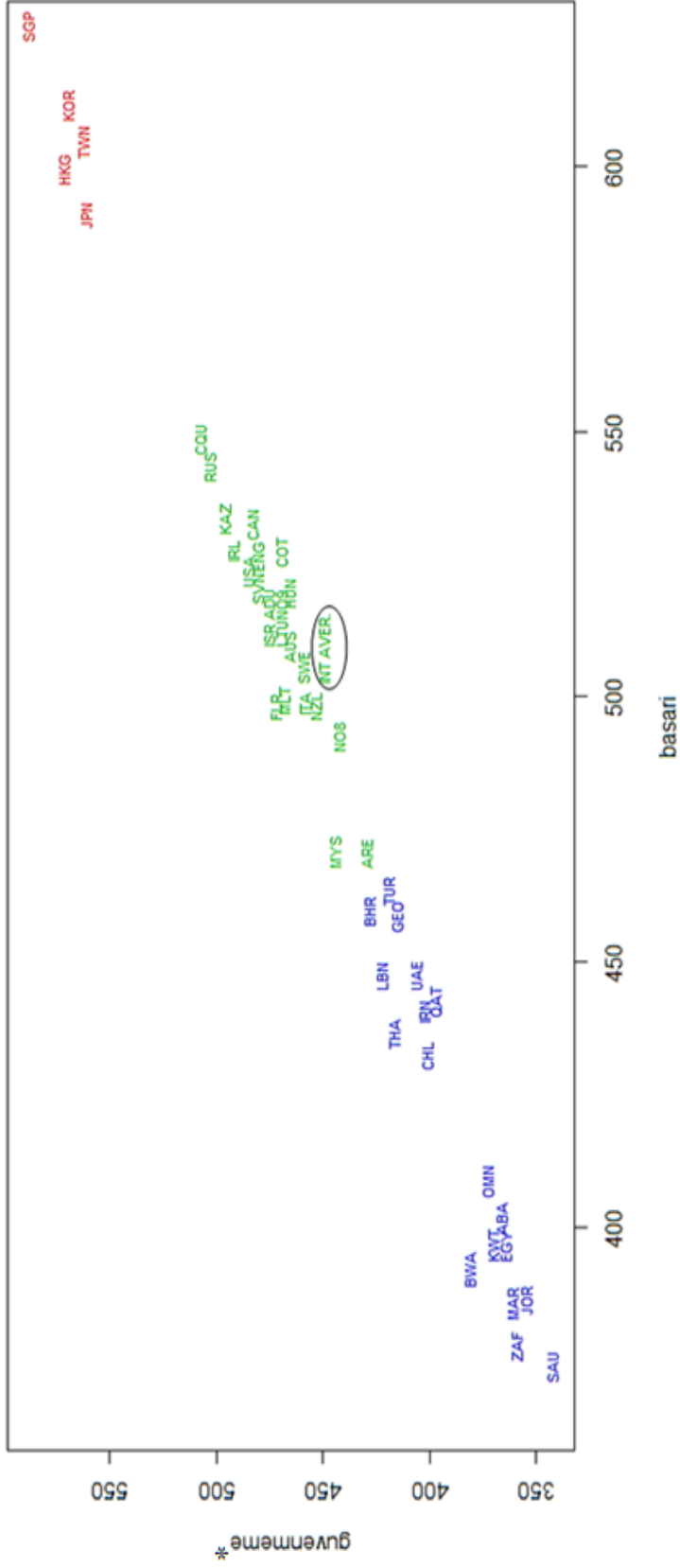
Bu çalışmada, TIMSS 2015 uygulamasına katılan ülkeler arasında tutum-başarı paradoksuna neden olabilecek tepki stillerinden UTS açısından farklılaşım farklılaşmadığının ve UTS'nin etkisinin belirlenmesi, farklı ülkelerdeki sekizinci sınıf öğrencilerin tutum puanlarında UTS'nin etkisinin düzeltilerek tutum ve matematik başarısı arasındaki ilişkinin belirlenmesi amaçlanmıştır. Bu bağlamda, bu çalışma korelasyonel bir çalışmadır. Korelasyonel araştırmalarda, değişkenler üzerinde herhangi bir değişilmeme yapılmaksızın aralarındaki ilişkiler belirlenmeye çalışılır. Bu yüzden korelasyonel araştırmalar daha çok betimsel niteliktedir (Fraenkel ve Wallen, 2006).

2.1.2 Evren ve Örneklem

Mevcut araştırmanın örneklemini, TIMSS 2015 uygulamasına katılan ve tutum-başarı paradoksunun görüldüğü ülkelerin sekizinci sınıf öğrencileri oluşturmaktadır. Ülkelerdeki sekizinci sınıf öğrencilerinin seçiminde iki aşamalı tabakalı örnekleme yöntemi kullanılmıştır. İlk aşamada, okullar, evrendeki oranları doğrultusunda seçkisiz olarak seçilmiştir. İkinci aşamada ise bu okullar içinden en az bir sınıf seçkisiz olarak belirlenmiştir. Belirlenen bu sınıflardaki sekizinci sınıf düzeyindeki öğrencilerin tamamı uygulamaya katılmıştır (LaRoche, Joncas ve Foy, 2016). TIMSS 2015 uygulamasına katılan sekizinci sınıf öğrencilerinin seçilmesinin nedenleri; (i) bu sınıf düzeyindeki öğrencilerin, matematik öğretim programında kazandırılması gereken bilgi ve becerilerin büyük ölçüde farklılaştığı (Rodriguez, 2004) kademeler olan ortaokuldan liseye geçiş döneminde olması (Reynolds, 1991) (ii) daha küçük bir yaş grubunda olan dördüncü sınıf öğrencilerinin kendi yeterliliklerinin ve tutumlarının farkında olmaması ve kendilerini bu konularda etkili değerlendirememeleridir (Harter, 1999).

TIMSS 2015 uygulamasına katılan ülkelerdeki öğrencilerin matematiği öğrenmeyi sevme, matematikte kendine güven ve matematiğe değer verme değişkenlerine ilişkin maddelere verdikleri cevaplar doğrultusunda, matematik başarısı ile matematiğe yönelik tutumla ilgili değişkenler arasında ülke düzeyinde negatif ilişki

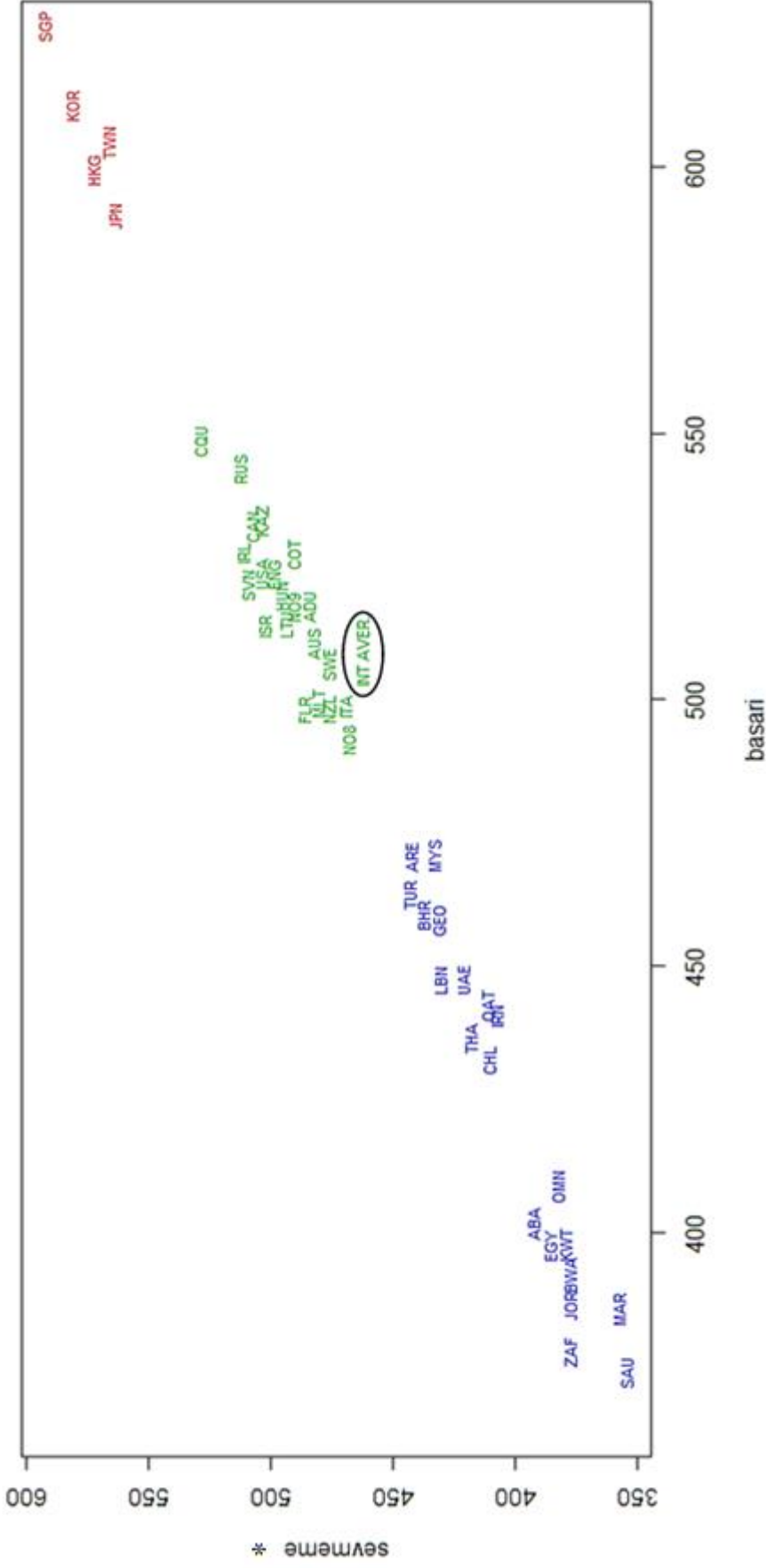
olduđu belirlenmiřtir. TIMSS 2015 uygulamasına katılan tm lkeler arasından lke seiminde tutum ile bařarı arasındaki iliřkiye dair rntnn daha iyi temsil edilmesini sađlamak amacıyla matematiđi sevmeyen, matematikte kendine gvenmeyen ve matematiđe deđer vermeyen đrencilerin matematik bařarı puanları kullanılarak kmeleme analizi yapılmıř ve oluřan her bir gruptan lkeler seilmiřtir. Kmeleme analizi, ok deđiřkenli istatistiksel tekniklerden biri olup verilerin benzerliklerine gre homojen alt gruplara ayrılması iin kullanılmaktadır (okluk, řekerciođlu ve Bykztrk, 2010). lkelerin matematik bařarılarına ve matematiđe ynelik tutumlarına gre ka kmeye ayrıldıđı nceden bilinmemesinden ve farklı evrenler olan lkelerden gelen rneklemeler seilmesinden dolayı tek bađlantı kmeleme yaklařımı ile birleřtirici hiyerarřik kmeleme yntemi kullanılmıřtır. Bu yntem, ayırık olan kmeleri belirlemede etkili bir yntemdir. Kme sayısı belirlendikten sonra ise daha sađlam teorik temellere dayanan hiyerarřik olmayan kmeleme yntemleri (Tatldil, 2002) arasından gvenilir bir yntem olması nedeniyle K-ortalamlar yntemi kullanılmıřtır (okluk ve diđ., 2010). Sonu olarak kmeleme analizi sonucunda ortaya ıkan  kmeden (I.kme “matematiđe ynelik tutumu olumsuz fakat bařarılı, II.kme “matematiđe ynelik tutumu ne olumlu ne de olumsuz ve bařarısı orta dzeyde” ve III. kme“ matematiđe ynelik tutumu olumlu fakat bařarısız”) beřer lke olmak zere toplam 15 lke seilmiřtir. Kmeleme analizi sonucunda ortaya ıkan kmeler ve iliřkilerin yn řekil 2., 3. ve 4’de verilmiřtir.



Şekil 2. Ülkelerin Matematik Başarısı ve Matematik Dersinde Kendine Güvenmesine göre Dağılımları ($r_1 = -.59$, $p < .05$)

Not 1. AUS= Avustralya, BHR= Bahreyn , BWA= Botswana, CAN = Kanada, CHL= Şile , TWN= Tayvan, EGY = Mısır, ENG= İngiltere, GEO =Gürcistan, HKG = Hong Kong, HUN= Macaristan, IRN= İran, IRL= İrlanda, ISR= İsrail, ITA= İtalya, JPN= Japonya, JOR= Ürdün, KAZ= Kazakistan, KOR= Kore, KWT= Kuveyt, LBN =Lübnan, LTU = Litvanya, MYS = Malezya, MLT =Malta, MAR = Fas, NZL= Yeni Zelanda, NOR = Norveç, OMN= Umman, QAT = Katar, RUS = Rusya, SAU= Suudi Arabistan, SGP = Singapur, SVN=Slovenya, ZAF = Güney Afrika, SWE= İsveç, THA= Tayland, TUR = Türkiye, ARE =Birleşik Arap Emirlikleri, USA = Amerika Birleşik Devletleri , ABA =Boyenes Ayres, COT = Ontario, CQU= Kubek, NO8= Norveç 8. sınıf, AAD= Abu Dabi, ADU = Dubai

*Grafığın y eksenini matematikte kendine **güvenmeyen** öğrencilerin matematik başarı puanlarını göstermektedir.



Şekil 3. Ülkelerin Matematik Başarısı ve Matematik Dersini Sevmemesine göre Dağılımları ($r_2 = -.74, p < .05$)

*Grafığın y eksenini matematiği öğrenmeyi **sevmeyen** öğrencilerin matematik başarı puanlarını göstermektedir.



Şekil 4. Ülkelerin Matematik Başarısı ve Matematik Dersine Değer Vermemesine göre Dağılımları ($r_3 = -.65, p < .05$)

*Grafikğin y eksenini **matematiğe değer vermeyen** öğrencilerin matematik başarı puanlarını göstermektedir.

Şekil 2, Şekil 3 ve Şekil 4 'te görüldüğü üzere matematik başarısı ile matematiği öğrenmeyi sevme, matematikte kendine güven ve matematiğe değer verme boyutları arasındaki ülkeler arası korelasyon değerlerinin negatif yönde olduğu görülmektedir ($r_1=-.59$, $r_2=-.74$, $r_3=-.65$, $p<.05$) Seçilen bu ülkelerde tutum-başarı paradoksunun gözlenme durumunu belirlemek amacıyla matematik başarısı ile matematiğe yönelik tutum arasında birey düzeyindeki ilişkileri gösteren ülke içi korelasyon değerleri incelenmiştir. Seçilen ülkelerin tutumla ilgili boyutlara ilişkin ülke içi korelasyonları Çizelge 1'de verilmiştir:

Çizelge 1
Ülke İçi Korelasyon Değerleri

Ülkeler	Değişken Çifleri Arasındaki İlişkiler		
	Başarı-Sevme	Başarı-Güven	Başarı-Değer
Singapur	.28	.39	.11
Kore Cumhuriyeti	.41	.50	.39
Çin- Tayvan	.44	.51	.35
Hong-Kong	.32	.37	.19
Japonya	.35	.44	.22
İsveç	.42	.63	.19
İtalya	.36	.52	.19
Malta	.29	.41	.11
Avustralya	.29	.50	.19
Norveç	.37	.61	.20
Türkiye	.19	.52	.11
Şili	.19	.45	.13
Kuveyt	.20	.33	.18
Mısır	.24	.37	.14
Sudi Arabistan	.28	.39	.11

Çizelge 1 'den görüldüğü üzere bu ilişkilerinin pozitif yönde olduğu belirlenmiştir. Bir diğer anlatımla, seçilen bu ülkelerde tutum-başarı paradoksu görülmektedir. Ülke içi korelasyonlara ilişkin saçılma grafikleri Ek A'da verilmiştir.

Ülkeler seçildikten sonra her bir ülkenin veri setindeki kayıp değerler silinmiş ve her bir ülkeden 500'er kişilik alt örneklem seçkisiz olarak seçilmiştir. Seçilen bu ülkelerin evren ve örnekleminde yer alan öğrenci sayıları Çizelge 2'de verilmiştir: (LaRoche ve Foy, 2016):

Çizelge 2.
Ülkelerin Evren ve Örneklemi

	Hedef Evren		Örneklem		Alt Örneklem
	Okul	Öğrenci	Okul	Öğrenci	Öğrenci
Singapur	167	47626	167	6116	500
Kore Cumhuriyeti	3007	587190	150	5309	500
Çin- Tayvan	991	285714	190	5711	500
Hong-Kong	158	4363	133	4155	500
Japonya	150	5037	147	4745	500
İsveç	1616	95438	150	4090	500
İtalya	5718	554401	161	4481	500
Malta	48	4063	48	3891	500
Avustralya	290	11968	285	10338	500
Norveç	1000	61174	142	4795	500
Türkiye	15583	1298955	218	6079	500
Şili	184	5285	171	4849	500
Kuveyt	327	39997	168	4503	500
Mısır	9900	1300305	211	7822	500
Sudi Arabistan	154	3962	143	3759	500

2.1.3 Veriler ve Elde Edilmesi

Bu araştırmada veri toplama aracı olarak TIMSS 2015'de uygulanan matematik başarı testinden ve öğrenci anketlerinden elde edilen 15 ülkeye ait veriler kullanılmıştır. Öğrenci anketinden tutumla ilgili değişkenlerin ölçüldüğü maddelere ait cevaplar ve öğrencilerin matematik başarı puanları seçilmiştir. TIMSS 2015 uygulamasına ait bu veri dosyaları TIMSS web sayfasından (<http://timssandpirls.bc.edu/timss2015/>) elde edilmiştir.

2.1.3.1 Veri Toplama Araçları

Bu çalışmada kullanılan verilerinde elde edilmesinde kullanılan ölçme araçları, TIMSS 2015 uygulamasına katılan sekizinci sınıf öğrencilerinin almış olduğu "Öğrenci Anketi" ve "Matematik Başarı Testi"dir.

2.1.3.1.1 Öğrenci anketi. TIMSS 2015 uygulamasına katılan her ülkeye öğrenci anketi uygulanmıştır. TIMSS uygulamalarında kullanılan anket ve başarı testleri, uzmanların madde geliştirme ve redaksiyonunu gerçekleştirmesi, pilot uygulamasının yapılması, pilot uygulamadan elde edilen sonuçlara göre maddelerin analizlerinin yapılması ve maddelerin düzeltilmesi, uzmanlarla sürekli iş birliği içinde olunması gibi standart süreçlerden geçerek IEA tarafından geliştirilmektedir (Martin ve Mullis, 2012).

Öğrenci anketinde, öğrencilerin demografik bilgilerine, evdeki ortamlarına, öğrenmeye, okul ortamlarına, kendi algılarına ve tutumlarına yönelik maddeler yer almaktadır (Hooper, Mullis ve Martin, 2013). Bu araştırmada ise öğrenci anketinde tutumla ilgili öğrencinin "matematiği öğrenmeyi sevmesi", "matematikte kendine güvenmesi" ve "matematiğe değer vermesi" boyutlarında başarısı yüksek olan ülkelerde olumsuz yönde tutuma sahip öğrencilerin yüzdesinin diğer ülkelere kıyasla fazla olması ya da başarısı düşük olan ülkelerde olumlu yönde tutuma sahip öğrencilerin diğer ülkelere kıyasla fazla olması nedeniyle mevcut çalışmada bu değişkenlerine ait ölçekler kullanılmıştır. Bu ölçeklerde yer alan maddelerin her biri dört tepki kategorisine sahiptir. Bu tepki kategorileri 1 "kesinlikle katılıyorum"u "4" ise "hiç katılmıyorum"u ifade edecek şekilde sıralanmıştır. Bu yüzden TIMSS 2015 uygulamasında yer alan bu ölçeklerden elde edilen toplam puanların yüksek olması olumsuz yönde tutumu, düşük olması ise olumlu yönde tutumu göstermektedir. Bu ölçeklere ait maddeler Ek B' de verilmiştir. Ölçeklerin seçilen ülkelere elde edilen Cronbach alfa güvenilirlik katsayılarının .70 ile .96 arasında değiştiği bulunmuştur (Martin, Mullis, Hooper, Yin, Foy ve Palazzo, 2016). Güvenirlik katsayılarının .60'dan büyük olması ölçeklerden elde edilen puanların güvenilir olduğunu göstermektedir (Büyüköztürk, 2005). Araştırmaya dahil edilen ülkelerin farklı tutum ölçeklerinden elde edilen puanlarına ilişkin güvenilirlik katsayılarının tamamı Ek C'de verilmiştir.

2.1.3.1.2 Matematik başarı testi. TIMSS uygulamalarında kullanılan başarı testleri, önceden belirlenen kazanımlar çerçevesinde hazırlanmakta ve başarı testlerine dahil edilecek maddelerin geliştirilme sürecini uzmanlar koordine etmektedir. Maddeler tekrar gözden geçirilmekte ve maddelere verilecek cevapların puanlanması için açık puanlama anahtarları hazırlanmaktadır. Bu süreçlerden geçen maddelerin katılımcı ülkelerde pilot uygulaması yapılmaktadır. Psikometrik özellikleri yeterli olan maddeler asıl uygulamalara dahil edilmektedir (Büyüköztürk, Çakan, Tan ve Atar, 2014).

Araştırmada sekizinci sınıf öğrencilerinin matematik başarısına ilişkin puanları, TIMSS 2015 uygulamasında uygulanan matematik başarı testlerinden elde edilmiştir. TIMSS 2015' deki matematik başarı testleri, 115'i çoktan seçmeli ve 97'si yanıtı yapılandırılmış olmak üzere toplamda 212 maddeden oluşmaktadır. Matematik başarı testinde yer alan maddelerin konu kapsamlarına göre dağılımı %30'u sayılar, %30'u cebir, %20'si geometri, %20'si veri ve olasılık konu alanlarına aittir. Bununla birlikte, başarı testinde yer alan maddelerin bilişsel beceri düzeylerine göre dağılımları; %35'i bilme, %40'ı uygulama ve %25'i ise akıl yürütme şeklindedir. Bilme, öğrencilerin matematiksel gerçekler, kavramlar ve işlemlerle ilgili bilgi tabanını ifade eder. Uygulama, öğrencinin bilgisini ve kavramsal anlayışını sorulan soruya uygulama yeteneği anlamına gelmektedir. Akıl yürütme ise alışılmış soru çözümünün daha ötesinde bir bilişsel düzeydir ve daha zor çok aşamalı soruların çözümünü kapsar (Mullis, Cotter, Fishbein ve Centurino, 2016).

TIMSS uygulamalarında başarı testinde çok sayıda madde yer almaktadır. Bu maddeler, sekizinci sınıflarda her bir öğrenciye bir kitapçık verilerek ve toplamda 14 kitapçığa dağıtılarak uygulanmaktadır. Her bir öğrenci farklı test kitapçığını aldığından öğrencilerin matematik başarısını temsil eden beş farklı makul (plausible) değer hesaplanmaktadır. Bu çalışmada ise, benzer çalışmalarda olduğu gibi (örn. Buckley, 2009) öğrencilerin matematik başarısını temsil etmek amacıyla makul beş değer ortalaması kullanılmıştır.

2.1.3.2 Verilerin Analizi

Her bir ülkeye ait veri setindeki kayıp değerler, örneklemelerde yer alan bireylerin sayıca fazlalığı ve çoklu atamanın yapılmasının bireylerin uç tepki stili puanlarını etkileyeceği gibi durumlar göz önünde bulundurularak veri setinden silinmiştir. Ölçeklerde yer alan maddelere ilişkin tepki kategorilerinin yüksek puan

olumsuz yönde tutumu temsil edecek şekilde sıralanması sebebiyle olumlu yönde tutumu bildiren maddeler için “1=4, 2=3, 3=2 ve 4=1” şeklinde ters kodlama yapılmıştır. Örneğin "matematiği öğrenmeyi severim" maddesine ait tepki kategorilerinin ters kodlama işlemi gerçekleştirilerek bu maddede "4" tepki kategorisinin seçilmesinin matematiği sevdiği anlamına gelmesi sağlanmıştır. Bu durumda ölçekten alınan yüksek puan, ilgili tutum nesnesine yönelik olumlu tutumu yansıtmaktadır.

Ters kodlama işlemi gerçekleştirildikten sonra analizlerin uzun sürmesi ve analizlerin gerçekleştirildiği bilgisayarın işlemci hızı ve hafızası göz önünde bulundurulduğundan analizlerin karmaşıklığını önlemek ve örneklerde yer alan öğrenci sayılarını denk tutmak amacıyla her bir ülkeye ait veri setinden 500'er kişilik alt örneklem seçkisiz olarak seçilmiştir. 500 kişilik alt örneklemin, örneklemin tamamını yeterli bir şekilde yansıtmadığını belirlemek amacıyla alt örneklemin ve örneklemin tamamının başarı puanı ve alt ölçek ortalama puanları arasındaki korelasyon incelenmiştir. Alt örnekleme ilgili değişkenler arasındaki korelasyon değerleri ile örneklemin tamamında değişkenler arasındaki korelasyon değerlerinin birbirine yakın olması alt örneklemin temsiliyetinin bir göstergesi olarak düşünülebilir. Örneklemere ve alt örneklemere ait korelasyon değerleri, Ek D’de verilmiştir.

Alt amaçlara göre verilerin analizinde yapılan işlemlerden önce UTS’yi belirlemede kullanılan yöntemler hakkında bilgiler verilmiştir:

UTS’yi Belirlemede Kullanılan Yöntemler

Alanyazında UTS’yi belirlemeye yönelik sunulan yöntemler; ad hoc uç tepki stili ölçümlerine ve örtük değişken modellerine dayalıdır. Bu yöntemlerden ilki olan ad hoc uç tepki stili ölçümlerine dayalı UTS’nin belirlenmesi için ilgili ölçek maddelerine verilen uç yanıtların sayısı belirlenerek UTS indeksi oluşturulmaktadır (Greenleaf, 1992b). Fakat, bu yöntemle indeks hesaplamak, ölçülmek istenen özelliğin asıl etkisinin gözlenmesinde yetersiz kalmaktadır. UTS’yi ölçülmek istenen özellikten ayırmak için, cevaplayıcıların UTS eğilimlerini ölçmeyi amaçlayan ve ölçülmek istenen özellik ile düşük korelasyon gösteren maddeler eklenebilir (Greenleaf, 1992b). Fakat yeni maddelerin eklenmesinin de bir takım sınırlılıkları vardır. İlk olarak, fazladan maddelerin eklenmesi cevaplayıcıların yorgunluk etkisiyle geçerliliğin düşmesine neden olabilir. Bir diğer sınırlılık ise eklenen maddelerin farklı ülkelerde, gruplarda aynı

yapıyı ölçmesini gerektirmesidir. Aksi bir durum geçerliliği düşürür. Bu nedenlerden dolayı UTS'yi belirlemede Ad Hoc yöntemleri örtük değişken modeli yöntemlerine kıyasla daha az kullanışlıdır (Bolt ve Newton, 2011).

UTS'yi belirlemek için kullanılan örtük değişken modelleri ise DFA, GSA ve madde tepki kuramına dayalı modellerden oluşmaktadır. DFA, tepki stillerini incelemek için iki yöntemle gerçekleştirilebilir. Bu yöntemlerden bir tanesi çoklu grup DFA'dır (Cheung ve Rensvold, 2000). İkincisi ise maddelere cevap vermeyi doğrudan etkileyen tepki stillerini ayrı bir faktör olarak ele alan çok faktörlü DFA (multifactor CFA approach) olarak da bilinen yöntemdir (Billiet ve McClendon, 2000).

Çoklu grup DFA çalışmalarında, her bir gruba ilişkin oluşturulan DFA modelleri sadece ölçülmek istenen özelliği içermektedir. Bu yüzden, UTS'yi ölçmek amacıyla oluşturulmaz. Faktör örüntüsü, faktör yükleri, kesişimler ve örtük ortalamalar gruplar boyunca değişmediğinde, ölçülen özelliğin gruplar arasında değişmediği ve grup karşılaştırmalarının tepki stilleri farklılıklarından etkilenmediği sonucuna ulaşılır (Cheung ve Rensvold, 2000). Fakat, tepki stillerinin var olduğu durumlarda da ölçme değişmezliği sağlanabilir (Little, 2000). Tepki stili, kültürleri tanımlayan bir özellik olduğunda ölçülebilir yapı olarak ele alınmalıdır. Bu bağlamda, çoklu grup DFA, tepki stillerinin ve ölçülen özelliğinin etkisini birbirinden ayırmada yeterli olmayabilir.

Çok faktörlü DFA'da tepki stilleri hem ayrı bir faktör olarak ele alınır hem de maddelerin ölçülen özelliğe ilişkin faktörde ve tepki stilleri faktöründe yük vermesine izin verilir. Fakat bu yöntem, maddelere verilen tepkiler üzerinde tepki stillerinin etkilerini doğrusal olarak modellediği için doğrusal olmayan etkilerini belirlemede yetersizdir. Bu yüzden, çok faktörlü DFA, kabullenici tepki stillerinin incelenmesinde kullanılabilirken (Billiet ve McClendon, 2000) UTS'yi belirlemede tercih edilen bir etkili yöntem değildir.

UTS'yi belirlemek için örtük değişken modellerine dayalı olarak gerçekleştirilen yöntemlerden biri olan gizil sınıf analizinde ise faktörleri sürekli değişken olarak ele alan DFA'nın aksine, faktörlerin ve gözlenen değişkenlerin ölçme düzeyleri kategoriktir (Moors, 2003). Aynı zamanda, GSA'da faktör analizinde olduğu gibi doğrusallık ve örtük faktörlerin çoklu normallik varsayımlarının sağlanmasına gerek yoktur. DFA, korelasyon ve kovaryans yapılarıyla sınırlı iken gizil sınıf analizi, cevaplayıcıların uç kategorilerde tepki verme eğilimine ilişkin bilgi veren tam frekans tablosunu (full frequency table) analiz eder. Fakat frekans tablosu seyrek (sparse) ise, model uyumunu değerlendirmek güçleşir (Hagenaars ve McCutcheon, 2002).

GSA'nın tepki stillerini belirlemede kullanılmasının çeşitli sınırlılıkları vardır. Birincisi, tepki stili süreksiz bir değişken olarak ele alınmaktadır. Fakat, psikoloji alanyazınında, genellikle tepki stilleri sürekli bir değişken olarak görülmektedir (Greenleaf, 1992a; Prediger, 1999). Bir diğer sınırlılık ise, puanların tepki stilinin etkilerinden arındırılmasına imkan sağlamamasıdır. Bu yüzden, ölçme konusu olan asıl özellik üzerinde UTS yanlılığını düzeltme konusunda kullanışlı bir yöntem değildir (Bolt ve Johnson, 2009).

Tepki stillerini belirlemede daha önceden önerilen yöntemlerin ölçülmek istenen asıl özellik ile tepki stillerini birbirinden ayıramaması sınırlılığına karşın ÇBMTK modeli kullanılmaktadır. Örtük değişken modellerinden bir diğeri olan ÇBMTK modeli, iki ya da daha fazla gizil değişken ile bir maddeye verilen cevap arasındaki ilişkiyi modellemek için kullanılır (Reckase, 2009). Bu bağlamda, ÇBMTK, tepki stilini, bireyin tepkisi üzerinde etkili olan psikolojik başka bir boyut olarak istatistiksel açıdan ele almakta ve ölçme konusu özelliğin etkisi ile beraber analiz etmektedir. ÇBMTK'da, $\theta_1, \theta_2, \dots, \theta_n$ yetenek düzeyinde bir bireyin j maddesinin k .tepki kategorisini seçme olasılığı şu şekilde belirlenmektedir (Bolt ve Johnson, 2009):

$$P(Y_j = k | \theta_1, \theta_{UTS}) = \frac{\exp(\alpha_{jk1}\theta_1 + \alpha_{jk2}\theta_{UTS} + c_{jk})}{\sum_{h=1}^k \exp(\alpha_{jh1}\theta_1 + \alpha_{jh2}\theta_{UTS} + c_{jh})} \quad (1)$$

a: j . maddenin kategori eğimini

c: kesişim parametresini

θ_1 : ölçülmek istenen özelliği;

θ_{UTS} : uç tepki stiline ilişkin özelliği göstermektedir.

Tepki stilinin modellenebilmesi için maddeler boyunca kategori eğim parametrelerine sabit değer kısıtlamaları uygulanmaktadır. Örneğin, 7'li likert türünde bir maddede, θ_1 'in ölçülmek istenen özellik ve θ_{UTS} 'nin uç tepki stiline ilişkin özellik olarak yorumlanabilmesi için, θ_1 için maddelere ait tepki kategorilerinin a parametresi -3, -2, -1, 0, 1, 2, 3 değerlerine sabitlenirken θ_{UTS} için 3, -1.2, -1.2, -1.2, -1.2, -1.2, 3 değerlerine sabitlenir. Burada dikkat edilmesi gereken nokta birbirinin simetriği olan tepki kategorilerinin mutlak değerce aynı değere sabitlenmesi ve her bir j maddesi için $\sum_k c_{jk} = 0$ olmasıdır (Lu ve Bolt, 2015). Böylelikle UTS'nin etkisi doğru bir şekilde kestirilebilmektedir.

ÇBMTK'nın diğer yöntemlerden en önemli farkı, sadece tepki stillerinin varlığını belirten diğer yöntemlerin aksine ölçme konusu özelliğın kestiriminde UTS'den kaynaklı yanlılığın etkisini düzeltme imkanı sağlamasıdır. Örneğın, θ_1 'in ölçülmek istenen asıl özellik ve θ_2 'nin uç tepki stili olduđu durumda θ_1 ve θ_2 'nin herhangi bir düzeyi için madde parametrelerinin kestirimlerine dayanan beklenen toplam puan (BTP-expected sum score) şu şekilde hesaplanmaktadır:

$$BTP(\theta_1, \theta_2) = \sum_{j=1}^m \sum_{k=1}^K k * P(Y_j = k | \theta_1, \theta_2) \quad (2)$$

k: kategori sayısı

m: madde sayısı

Bu eşitlikte $P(Y_j = k | \theta_1, \theta_2)$ 1. denkleme göre hesaplanmaktadır. UTS'den kaynaklı yanlılığı hesaplamak için $\theta_2 = 0$ referans noktası alınmaktadır. Bu durumda bu yanlılık gözlenen ve beklenen arasındaki farktır ve şu şekilde hesaplanmaktadır:

$$YANLILIK(\theta_1, \theta_2) = BTP(\theta_1, \theta_2) - BTP(\theta_1, 0) \quad (3)$$

UTS'ye ilişkin $\theta_2 = \theta_{UTS} = 0$ noktası keyfi referans noktasıdır (arbitrary reference point). Bu yüzden 3. eşitlikte elde edilen değer nispi/göreceli yanlılık (relative bias) değeri olarak adlandırılmaktadır. θ_{UTS} 'nin farklı düzeylerinde toplam puanlar da o doğrultuda değışeceğinden, aradaki fark hep aynı kalacaktır.

Özetle, ÇBMTK'nın tepki stillerinin varlığını belirlemesinin yanı sıra ölçme konusu özelliğın kestiriminde tepki stillerinden kaynaklı yanlılığının etkisini düzeltme imkanı sağlaması (Bolt ve Johnson, 2009) ve UTS belirlemeye yönelik alanyazında sunulan diğer yöntemlerin çeşitli sınırlılıkları bulunduđu için bu çalışmada bireylerin UTS'nin etkisine göre düzeltilmiş tutum puanları ÇBMTK kullanılarak kestirilmiştir.

Alt amaçlara göre gerçekleştirilen analizler hakkında ayrıntılı bilgiler aşağıda sunulmuştur:

Araştırmanın birinci alt amacı doğrultusunda ülkelerin UTS sergileme eğilimleri açısından manidar bir farkın olup olmadığını belirlemek amacıyla öncelikle ad hoc UTS kestirimi yöntemlerinden biri kullanılarak ülkelerin UTS sergileme eğilimlerine ilişkin UTS indeksi oluşturulmuştur. Bu indeksin hesaplanmasında Bachman ve O'Malley'in (1984) yöntemi kullanılmıştır. UTS indeksi için tepki kategorilerinin uç noktaları olan "hiç katılmıyorum" ile "kesinlikle katılıyorum" yanıtları 1, diğer yanıtlar 0 olarak kodlanmıştır. Ardından bireylerin her bir alt ölçekten aldıkları puanlar toplanıp ilgili

ölçekteki madde sayısına bölünerek ortalama UTS indeksi belirlenmiştir. Veri setlerindeki uç değerlerin çıkarılmasının uç tepki stiline sahip bireylerin de çıkarılmasına neden olabilmektedir. Aynı zamanda değişkenlere ilişkin normallik ve varyansların homojenliği varsayımı sağlanmamıştır. Bu yüzden bu varsayımların sağlanmasını gerektiren bağımlı değişken sayısının ikiden fazla olduğu durumlarda kullanılan parametrik analizlerden biri olan MANOVA (Mertler ve Vannatta, 2005) yerine parametrik olmayan analiz gerçekleştirilmiştir. Bu bağlamda, çok değişkenli veri setlerinin parametrik olmayan yorumlamalarını sağlayan R yazılım programında “nrmv” paketi (Woodrow ve diğerleri, 2017) kullanılarak birinci araştırma sorusuna yanıt aranmıştır. Bu amaçla “1” Singapur’u, “2” Kore’yi, “3” Tayvan’ı, “4” Hong Kong’u, “5” Japonya’yı, “6” İsveç’i, “7” İtalya’yı, “8” Malta’, “9” Avustralya’yı, “10” Norveç’i, “11” Türkiye’yi, “12” Şili’yi, “13” Kuveyt’i, “14” Mısır’ı ve “15” Sudi Arabistan’ı temsil edecek şekilde her bir ülkeye ait veri seti tek bir veri setinde birleştirilmiştir. Bu analize ilişkin kodlar Ek E1’de verilmiştir.

Araştırmanın ikinci alt amacına yönelik olarak belirli sınırlamaların yapıldığı iki farklı model geliştirilmiştir. İlk (temel) modelde, ölçülmek istenen tutumla ilgili boyutlar olan matematiği sevme (θ_S), matematikte kendine güven (θ_G), ve matematiğe değer verme (θ_D) olmak üzere üç ayrı boyuttan oluşmaktadır. İlk modelde, tepki kategori eğri parametreleri (a_{jkm}) her bir boyut için eşit aralıklı değerlere (-3, -1, 1, 3) sabitlenmiştir. İkinci modelde ise dördüncü boyut olan UTS yanlılığına dair boyut (θ_{UTS})da dahil edilmiştir. Dördüncü boyuta ilişkin kategori eğri parametreleri ölçeklerde yer alan bütün maddeler için sırasıyla “1, -1, -1 ve 1”e sabitlenmiştir. Daha sonrasında veriye daha iyi uyum sağlayan modeli belirlemek amacıyla modellere ilişkin model uyum istatistikleri Latent Gold programı kullanılarak elde edilmiştir. Model-veri uyumları çeşitli model uyum indekslerine göre belirlenmiştir. Bu indekslerin değerleri log olabilirlik fonksiyonuna bağlı olarak elde edilmektedir. Bayesçi yaklaşımlarda model seçilirken alternatif modellerin log sonsal olasılığı (log posterior probability) bilgi kriterlerinin (information criteria) hesaplanmasında kullanılmaktadır. Log sonsal, log olabilirlik ve önsel olasılığın bir fonksiyonudur. Bu bilgi kriterleri; Bayesçi bilgi kriteri (the Bayesian information criterion-BIC), Akaike bilgi kriteri (the Akaike information criterion- AIC) ve AIC indeksinin çeşitleri: Akaike bilgi kriteri 3 (the Akaike information criterion 3-AIC3) ve tutarlı Akaike bilgi kriteridir (the consistent Akaike information criterion- CAIC). Bu bilgi kriterlerinin düşük değerleri daha iyi uyum olduğunu göstermektedir (Vermunt & Magidson, 2005).

Latent Gold programında madde parametrelerinin kestiriminde marjinal en çok olabilirlik tahmin yöntemini kullanılmaktadır. Bu süreçte ilk önce EM algoritması sonrasında ise Newton Raphson iterasyonu gerçekleştirilir. Newton Raphson yöntemi parametrelerin kestiriminde birinci ve ikinci türevlerini almaktadır. Her bir boyut için 10 tane quadrant noktası kullanılarak Gauss-Hermite sayısal yakınlaştırma uygulanmaktadır. Belirlenen maksimum iterasyon sayısına ulaşıldığında Newton Raphson iterasyonları sonlanmaktadır (Magidson ve Vermunt, 2008).

Araştırmanın üçüncü amacına yönelik olarak öğrencilerin toplam ölçek puanlarında etkili olan UTS'den kaynaklı yanlılığı belirlemek ve araştırmanın dördüncü amacı için öğrencilerin düzeltilmemiş ve düzeltilmiş tutum puanlarını belirlemek için öncelikle ülke düzeyinde tutuma ilgili olan boyutlara ve tepki stiline dair madde ve birey parametreleri kestirilmiştir. Bu amaçla çok düzeyli ÇBMTK modeli kullanılarak iki aşamalı kestirim süreci gerçekleştirilmiştir. Birinci aşamada, ölçeklerde yer alan maddelerin kesişim parametreleri üç boyutlu ve dört boyutlu modellerle kestirilmiştir. Madde parametrelerinin kestiriminde Latent Gold programı kullanılmıştır (Vermunt ve Magidson, 2008). Madde parametrelerinin kestirimi için kullanılan kodlar Eke'de verilmiştir.

Maddelere ait kesişim parametreleri Çizelge 3'de sunulmuştur:

Çizelge 3.

Üç ve Dört Boyutlu Modelle Elde Edilen Kategori Kesişim Parametreleri

Boyut	Madde	Üç Boyutlu Model				Dört Boyutlu Model			
		Tepki Kategorileri				Tepki Kategorileri			
		1	2	3	4	1	2	3	4
Sevme	1	-1.92	.38	1.40	.14	-1.93	.42	1.42	.90
	2	-1.02	.56	.78	-.32	-.96	.56	.75	-.35
	3	-.93	.94	.81	-.83	-.91	.99	.83	-.91
	4	-2.39	.48	1.56	.36	-2.43	.55	1.58	.30
	5	-1.34	.46	1.07	-.19	-1.31	.48	1.7	-.24
	6	-.82	1.12	.89	-1.19	-.83	1.19	.94	-1.30
	7	-.90	.72	.90	-.72	-.86	.75	.91	-.80
	8	-.29	1.16	.74	-1.61	-.26	1.2	.77	-1.71
	9	-.11	.74	.40	-1.03	.10	.71	.36	-1.70
Güven	10	-1.60	.24	1.09	.26	-1.66	.34	1.2	.12
	11	-.91	.56	.67	-.32	-.96	.67	.76	-.47
	12	-.21	.57	.29	-.66	-.19	.64	.33	-.78
	13	-1.36	.55	.95	-.14	-1.45	.69	1.80	-.32
	14	-.90	.49	.62	-.21	-.94	.59	.70	-.35
	15	-.38	.79	.70	-1.11	-.44	.93	.83	-1.32
	16	-.76	.59	.72	-.55	-.81	.71	.83	-.72

	17	-.45	.34	.43	-.31	-.43	.39	.46	-.41
	18	-.48	.52	.42	-.45	-.49	.59	.47	-.58
	19	-2.68	-.32	1.57	1.43	-2.76	-.25	1.57	1.43
	20	-2.44	.12	1.58	.75	-2.58	.24	1.64	.70
	21	-2.94	-.27	1.50	1.71	-3.10	-.22	1.49	1.74
	22	-2.66	-.08	1.41	1.34	-2.75	.00	1.41	1.34
	23	-.34	.81	.54	-1.01	-.33	.84	.54	-1.60
	24	-2.70	-.01	1.51	1.20	-2.81	.90	1.54	1.19
	25	-3.05	-.28	1.64	1.69	-3.15	-.20	1.65	1.70
	26	-3.48	-.46	1.76	2.18	-3.55	-.41	1.73	2.23
	27	-3.94	-.66	2.02	2.58	-4.00	-.62	1.98	2.63

Çizelge 3’de UTS’nin etkisinin dikkate alınmadığı ve alındığı durumlar için geliştirilen sırasıyla üç boyutlu ve dört boyutlu modellerin analiz edilmesiyle kestirilen kategori kesişim parametreleri verilmiştir. Kategori kesişim parametresi, örtük özellik düzeyinin ortalaması “0” olduğunda o kategorinin seçilme eğilimini yansıtmaktadır (Bolt ve Newton, 2011). Örneğin, araştırma kapsamına alınan ülkelerdeki öğrencilerin her üç ölçüğe verdiği yanıtlar doğrultusunda maddelerin çoğunlukla 2. ve 3. kategorilerine ait kesişim parametrelerinin pozitif en yüksek değere sahip olduğu bulunmuştur.

İkinci aşamada, UTS’nin etkisinin dikkate alınmadığı ve alındığı durumlar için kestirilen kesişim parametreleri kullanılarak öğrencilerin tutumla ilgili üç değişkene ait theta değerleri kestirilmiştir. Bu aşamada, üç boyutlu model ile bireylerin düzeltilmemiş tutum puanları kestirilirken UTS’ye ilişkin theta değeri (θ_{UTS}) birey ve ülke düzeyine dahil edilmemiştir. Dört boyutlu modelde ise θ_{UTS} hem birey hem de ülke düzeyine dahil edilerek birey parametreleri üzerindeki etkisi iki düzeyde de belirlenmiş ve düzeltilmiş tutum puanlarının yanı sıra θ_{UTS} değerleri elde edilmiştir. Aynı zamanda, ülke düzeyinde ortalama başarı puanları kullanılarak ülkelerin tutumları ile başarıları arasındaki ilişkiler UTS’nin kontrol edilmediği ve edildiği durumlar için belirlenmiştir. Bu aşamada ise Markov Zinciri Monte Carlo (Markov Chain Monte Carlo-MCMC) yöntemi kullanılarak WinBUGS (Spiegelhalter, Thomas ve Best, 2004) programı ile analizler gerçekleştirilmiştir. Analize ilişkin kodlar Ek G’de verilmiştir.

MCMC, Bayes yaklaşımına dayalıdır. Bayesçi analizler, veri toplamadan önce modeldeki her bir parametre hakkında önsel bilgilerin kullanılmasıyla başlar. MTK modellerinde Bayesçi yaklaşımın kullanılması, istatistiksel yorumlamalarda güvenilirliği artırmaktadır (Fox, 2010). Bunun yanı sıra, Bayesçi yaklaşım (i) parametre

kestirimlerinin en çok olabilirlik tahmin yöntemlerine göre daha doğru, net (precise) olmasını, (ii) yetenek parametrelerinin madde parametreleriyle eş zamanlı olarak kestirilmesini sağlamakla birlikte (iii) Bayeşçi yaklaşımlarda önsel dağılımların kullanılması daha durağan (stable) parametre kestirimlerinin elde edilmesini mümkün kılmaktadır (Lord, 1986).

MCMC, Markov Zinciri ile Monte Carlo özelliklerinin birleşimidir. Monte Carlo, dağılımın özelliklerinin, dağılımdan çekilen seçkisiz örneklemeler üzerinden kestirilmesidir. Örneğin, Monte Carlo yaklaşımı, normal dağılımın ortalamasının dağılıma ait denklemler yoluyla hesaplanmasının yerine, normal dağılımdan yeterince fazla seçkisiz örneklemeler çekerek bunların ortalamasını bulmaya dayanmaktadır. Markov zinciri özelliği ise seçkisiz örneklemelerin ardışık bir süreç dahilinde oluşturulduğu fikrine dayanmaktadır. Her bir seçkisiz örneklem bir sonraki örneklemin oluşturulması için kullanılmaktadır. Bu zincirin temel özelliği, her bir yeni örneklemin sadece kendinden bir önceki örnekleme bağlı olmasıdır (van Ravenzwaaj, Cassey ve Brown, 2016). Özetle, MCMC, koşullu dağılımlardan simülasyon yolu ile örneklem çekerek, parametre tahminlerinin elde edilmesini sağlayan simülasyon tekniklerinin tamamıdır. MCMC’de dağılımdan örneklemeler çekilir ve daha sonra bu örneklemelerin ortalamaları beklenen değeri bulmak için kullanılır (Gilks, Richardson ve Spiegelhalter, 1996).

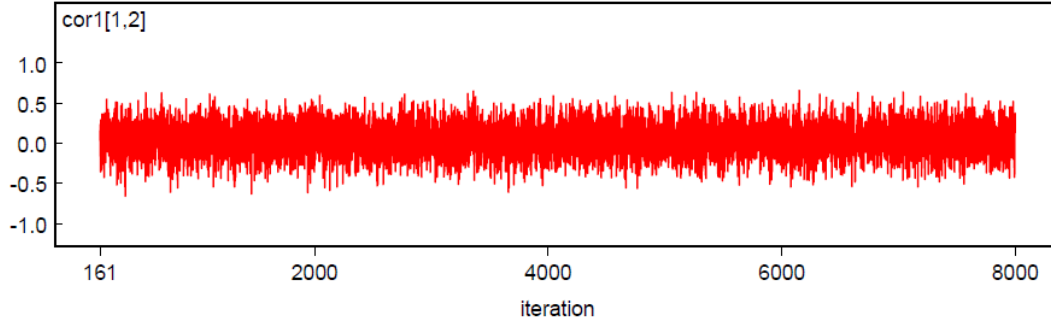
Markov zincirlerinde başlangıç değerleri önsellerden çalıştırılmaktadır. Analiz sonucunda kestirilecek parametreler hakkında kesin olmayan bilgilerin oluşturduğu olasılık dağılımı önsel dağılımdır. Önsel dağılımlar, tam Bayes tahmin yaklaşımlarının temelini oluşturmaktadır. Bu yüzden önsel dağılım olmadan Bayeşçi tahmin gerçekleştirilememektedir. MCMC’de önselleri belirlemede çeşitli yöntemler kullanılmaktadır. Bu yöntemler; eşlenik önseller (conjugate priors), uygun ve uygun olmayan önseller (proper and improper priors), aşırı önseller (hyperpriors), bilgi verici ve bilgi verici olmayan önseller (informative ve noninformative priors), zayıf bilgi verici önsellerdir (weakly informative priors)(Qian, 2011). Bu önsellerden, bilgi verici olmayan önseller, araştırma problemi ile ilgili önsel bilgi olmadığı veya araştırmacıların verinin parametre tahminine daha fazla ağırlıkvermesi gerektiği durumlarda kullanılmaktadır (Gelman ve Hill, 2007). Bu araştırmada da parametreler hakkında önsel bir bilginin bulunmaması nedeniyle bu çalışmada bilgi verici olmayan önseller kullanılmıştır.

MCMC’de başlangıç değerleri, simülasyonun ilk adımlarını etkilese de ilerleyen adımlarda oluşturulan zincirin üzerindeki etkisi kaybolmaktadır (Gill, 2002). Teorik olarak sonsuz sayıda iterasyon yapıldığında başlangıç değerinin etkisi söz konusu olmaz. Uygulamada ise sonsuz sayıda iterasyon yapmak mümkün olmayacağı için belli bir tekrardan sonra zincirin yakınsaması beklenir. Başlangıç değerinden etkilenen bölüm zincirden çıkarılır. Yakma periyodu (burn-in) olarak adlandırılan bu işlemle başlangıç değerlerinin sonsal değerler üzerindeki etkisi azaltılmaktadır. Bu yüzden parametre kestiriminde, bu işlemden sonra geriye kalan iterasyon sayısı kullanılması önemlidir (Congdon, 2003; Gill, 2002). Yakma periyodunda zincirden çıkarılacak iterasyon sayısının, toplam iterasyon sayısının %1 veya %2’si kadar olması yeterlidir (Geyer, 1992)

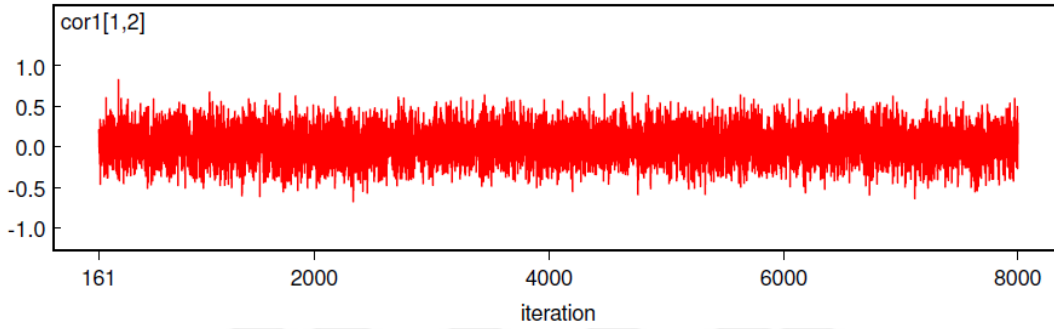
Parametre kestirimlerinin başlangıç değerlerinden etkilenmemesi ve yakınsamanın gerçekleşmesi için gerekli olan iterasyon sayısının belirlenmesi konusunda ortak bir görüş birliğine varılamamıştır. Yakınsamanın olması için bazı araştırmacılar (Gelman ve Rubin, 1996), çok sayıda kısa zincir yerine, uzun tek bir zincir çalıştırılmasını bazı araştırmacılar (Fırat, 2001a; Lee, 2007) ise farklı başlangıç değerleri ile birden çok zincir çalıştırılmasını önermişlerdir. Bu araştırmada ise Gelman ve Rubin (1996)’nin de önerdiği gibi tek bir zincir kullanılmış ve benzer çalışmalarda (Lu, 2012; Lu ve Bolt, 2015) yapılan iterasyon sayısının 2000 ile 10000 arasında değişmesi sebebiyle analizlerin gerçekleştirildiği bilgisayarın kapasitesine bağlı olarak her iki model için 8000 iterasyon yapılmıştır. Geyer (1992) tarafından önerildiği üzere bu iterasyonlardan 160’ı yakma periyodunda (burn-in period) kullanılmıştır.

Yapılan iterasyon sayısının yeterliliğine ve bunun sonucunda parametre kestirimlerinin doğruluğuna karar vermede Markov zincirinin yakınsayıp yakınsamadığına bakılmaktadır. MCMC’de yakınsamayı değerlendirmede birkaç yöntem vardır. Tek bir yöntem yakınsamayı değerlendirmede yeterli değildir (Qian, 2011). Bu yöntemler arasında sıklıkla kullanılanlar zaman serileri grafikleri (time series) (Sinharay, 2004) ve MC hataları (Ntzoufras, 2009) yer almaktadır.

Zaman serileri grafiklerinde, örüntü ya da düzensizlik gözlenmezse modelin yakınsadığı varsayılır (Ekici, 2005; Sinharay, 2004). Bu araştırma kapsamında yapılan MCMC simülasyonlarına dair bazı zaman serileri grafikleri Şekil 5 ve Şekil 6’da verilmiştir:



Şekil 5. Üç Boyutlu Modelde Başarı ve Sevme Örtük Değişkeni Arasındaki İlişkiye Ait Markov Zinciri Zaman Serisi Grafiği



Şekil 6. Dört Boyutlu Modelde Başarı ve Sevme Örtük Değişkeni Arasındaki İlişkiye Ait Markov Zinciri Zaman Serisi Grafiği

Şekil 5 ve Şekil 6'da TIMSS 2015 uygulamasına katılan ülkeler için yapılan MCMC simülasyonlarında düzensizlik görülmediği, yakınlığı görülmektedir. Diğer değişkenlere ilişkin zaman serisi grafikleri Ek H'da verilmiştir.

Yakınsamayı değerlendirme kullanılabilecek bir diğer yöntem olan MC hataları, simülasyon sonucunda ilgili parametrenin ortalama değişimini ölçer. İterasyon sayısı arttıkça MC hatası azalmaktadır. MC hatalarının .05'ten küçük olması yakınsamanın sağlandığının gösterir (Ntzoufras, 2009). Bu araştırma kapsamında gerçekleştirilen MCMC simülasyonları sonucunda elde edilen MC hatalarına örnek Çizelge 4'de verilmiştir:

Çizelge 4.

MCMC Simülasyonlarına Ait Betimsel İstatistikler ve MC Hataları

Düğüm	Ortalama		Standart Sapma		MC Hatası		Güven Aralığı Alt Sınır(0.25)		Medyan		Güven Aralığı Üst Sınır(0.75)		Toplam İterasyon	
	1	2	1	2	1	2	1	2	1	2	1	2	Başlangıç	Örnekleme
cor[1,1]	1.000	1.000	0.000	0.000	0.000	0.000	1.000	1.000	1.000	1.000	1.000	1.000	161	7840
cor[1,2]	0.048	0.036	0.207	0.213	0.002	0.003	-0.361	-0.382	0.048	0.037	0.445	0.452	161	7840
cor[1,3]	0.109	0.070	0.207	0.210	0.002	0.003	-0.307	-0.346	0.116	0.072	0.494	0.468	161	7840
cor[1,4]	0.084	0.071	0.209	0.211	0.002	0.002	-0.329	-0.347	0.086	0.073	0.482	0.465	161	7840
cor[1,5]	-	0.085	-	0.214	-	0.002	-	-0.339	-	0.088	-	0.492	161	7840
cor[2,1]	0.048	0.036	0.207	0.213	0.002	0.003	-0.361	-0.382	0.0483	0.037	0.445	0.452	161	7840
cor[2,2]	1.000	1.000	0.000	0.000	0.000	0.000	1.000	1.000	1	1.000	1.000	1.000	161	7840
cor[2,3]	0.069	0.062	0.208	0.213	0.002	0.002	-0.338	-0.358	0.0702	0.066	0.473	0.464	161	7840
cor[2,4]	0.040	0.041	0.211	0.215	0.002	0.003	-0.375	-0.389	0.0389	0.043	0.448	0.450	161	7840
cor[2,5]	-	0.050	-	0.215	-	0.002	-	-0.376	-	0.055	-	0.463	161	7840
cor[3,1]	0.109	0.070	0.207	0.210	0.002	0.003	-0.307	-0.346	0.1158	0.072	0.4937	0.468	161	7840
cor[3,2]	0.069	0.062	0.208	0.213	0.002	0.002	-0.338	-0.358	0.0702	0.066	0.4726	0.464	161	7840
cor[3,3]	1.000	1.000	0.000	0.000	0.000	0.000	1.000	1.000	1	1.000	1	1.000	161	7840
cor[3,4]	0.012	0.077	0.211	0.213	0.003	0.003	-0.395	-0.350	0.0122	0.081	0.4219	0.479	161	7840
cor[3,5]	-	0.031	-	0.216	-	0.002	-	-0.384	-	0.031	-	0.447	161	7840
cor[4,1]	0.084	0.071	0.209	0.211	0.002	0.002	-0.329	-0.347	0.0861	0.073	0.482	0.465	161	7840
cor[4,2]	0.040	0.041	0.211	0.215	0.002	0.003	-0.375	-0.389	0.0389	0.043	0.448	0.450	161	7840
cor[4,3]	0.012	0.077	0.211	0.213	0.003	0.003	-0.395	-0.350	0.0122	0.081	0.422	0.479	161	7840
cor[4,4]	1.000	1.000	0.000	0.000	0.000	0.000	1.000	1.000	1	1.000	1.000	1.000	161	7840
cor[4,5]	-	-0.087	-	0.211	-	0.002	-	-0.484	-	-0.092	-	0.334	161	7840
cor[5,1]	-	0.085	-	0.214	-	0.002	-	-0.339	-	0.088	-	0.492	161	7840
cor[5,2]	-	0.050	-	0.215	-	0.002	-	-0.376	-	0.055	-	0.463	161	7840
cor[5,3]	-	0.031	-	0.216	-	0.002	-	-0.384	-	0.031	-	0.447	161	7840
cor[5,4]	-	-0.087	-	0.211	-	0.002	-	-0.484	-	-0.092	-	0.334	161	7840
cor[5,5]	-	1.000	-	0.000	-	0.000	-	1.000	-	1.000	-	1.000	161	7840

1=Üç Boyutlu Model, 2= Dört Boyutlu Model

Çizelge 4’de belirtildiği üzere araştırma kapsamında geliştirilen modellere ilişkin gerçekleştirilen simülasyonlara ait MC hata değerlerinin .05’ten küçük olduğu görülmektedir. Buna göre, Markov zincirleri yakınsamaktadır.

İterasyon sayısının yeterli olduğu ve bunun sonunca Markov zincirlerinin yakınsadığı belirlendikten sonra her bir ölçek için toplam ölçek puanları eşitlik (2) ile hesaplanmış ve eşitlik (3) ile bireylerin toplam ölçek puanları üzerindeki yanlılık TIMSS 2015 uygulamasına katılan öğrencilerin farklı tutum düzeyleri için belirlenmiştir.

Araştırmanın dördüncü alt amacı doğrultusunda TIMSS 2015 uygulamasına katılan ve araştırmaya dahil edilen ülkelerdeki öğrencilerin UTS’nin etkisine göre düzeltilmemiş tutum puanları ile düzeltilmiş tutum puanları belirlendikten sonra bu puanlar arasındaki farkın manidarlığını değerlendirmede herbir ülke için ayrı ayrı ilişkili örneklem için t-testi analizi yapılmıştır. İlişkili örneklem için t-testi analizi, ilişkili iki örneklemden elde edilen iki ortalama arasındaki farkın manidarlığını test etmede kullanılan parametrik bir tekniktir (Büyüköztürk, Çokluk ve Köklü, 2012). Bu analiz gerçekleştirilmeden önce analize ilişkin varsayımların sağlanıp sağlanmadığı test edilmiştir. Bu amaçla, bağımlı değişkenler olan matematiği öğrenmeyi sevme boyutuna ilişkin UTS’nin etkisine göre düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasındaki fark puanlarının, matematikte kendine güvenme boyutuna ilişkin UTS’nin etkisine göre düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasındaki fark puanlarının ve matematiğe değer verme boyutuna ilişkin UTS’nin etkisine göre düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasındaki fark puanlarının aralık düzeyinde olup olmadığı ve bu fark puanlarının normal dağılıp dağılmadığı herbir ülke için ayrı ayrı ve ülkelerin tamamı için kontrol edilmiştir. Değişkenlerin çarpıklık ve basıklık değerlerinin -2 ve +2 arasında olması tek değişkenli normalliğin sağlandığını göstermektedir (Çokluk ve diğerleri, 2012). Buna göre bu değişkenlere ait fark puanlarının çarpıklık ve basıklık değerlerinin -2 ile 2 değerleri arasında olması ve histogram grafikleri, fark puanlarının normal dağıldığını göstermektedir. Fark puanlarının normalliğini gösteren çarpıklık, basıklık değerleri ve histogram grafikleri Ek I.1 ve Ek I.2’de verilmiştir.

Örneklem büyüklüğü arttıkça örneklem ortalamaları arasındaki farkın manidar bulunması gerçekte örneklem arasında fark olduğunu garanti etmemektedir. Bu nedenle farkların pratikteki manidarlığını test etmek amacıyla ilişkili örneklem için t-testi analizi sonucunda elde edilen manidar farklara ilişkin etki büyüklüğü değerleri

hesaplanmıştır. Etki büyüklüğü indeksi olan d değeri, ortalamaları standart sapma bağlamında karşılaştırır. Bu değer $.2$, $.5$ ve $.8$ civarında olması sırasıyla küçük, orta ve geniş etki büyüklüğü şeklinde yorumlanmaktadır (Büyüköztürk ve diğerleri, 2012). Etki büyüklüğü şu şekilde hesaplanmaktadır:

$$d = \frac{t}{\sqrt{N}} \quad (4)$$

Eşitlikte “ t ” ilişkili örneklem için t testi sonucunda elde edilen t değerini ve N ise örnekleme yer alan birey sayısını ifade etmektedir.

Araştırmanın beşinci amacı doğrultusunda gözlenen puanlara dayalı olarak matematik başarısı ile matematiğe yönelik tutumla ilgili değişkenler arasındaki ülke düzeyindeki ilişkileri belirlemek amacıyla ilk önce her bir ülkedeki bireylerin farklı tutum ölçeklerine ait ölçek puanlarının ortalaması hesaplanmış ve her bir ülkeye ait dört farklı değer (matematik başarı puanı, matematiği öğrenmeyi sevme boyutuna ilişkin ortalama ölçek puanı, matematikte kendine güven boyutuna ilişkin ortalama ölçek puanı ve matematiğe değer verme boyutuna ilişkin ortalama ölçek puanı) belirlenmiştir. Ardından, 15 ülkeye ait bu değerlerin normal dağılıp dağılmadığı kontrol edilmiştir. Örneklem büyüklüğünün 35’den küçük olması nedeniyle bu gibi durumlarda kullanılan normallik testlerinden Shapiro Wilk normallik testinin (Shapiro ve Wilk, 1965) sonuçları dikkate alınmıştır. Buna göre seçilen ülkelerin ülke düzeyindeki ortalama matematik başarı puanlarının normal dağılmadığı belirlenmiştir ($p=.03$). Bu yüzden, parametrik olmayan korelasyon katsayılarından Spearman Sıra Farkları Korelasyon Katsayısı (r_s) tutumla ilgili değişkenlerin gözlenen puanları ile matematik başarısı arasındaki ilişkiyi belirlemek amacıyla hesaplanmıştır. Bu noktada, ham verilerden yararlanılmıştır. Bunun yanı sıra, matematiğe yönelik tutum ile matematik başarısı arasındaki ülke düzeyindeki ilişkilerde UTS'nin kontrol edilmesinin etkisini belirlemek amacıyla WinBUGS programı ile üç ve dört boyutlu modeller analiz edilerek korelasyon değerleri elde edilmiştir. Bu noktada, üç boyutlu model ile dört boyutlu modelin ortak örtük özellik metriğinde (common latent trait metric) olmasından dolayı (Lu ve Bolt, 2015) üç boyutlu modelden elde edilen ülke düzeyindeki korelasyon değerleri dört boyutlu modelle elde edilen korelasyon değerleriyle karşılaştırma yapabilmek adına temel (baseline) değerler olarak ele alınmıştır.

Araştırmanın altıncı amacı kapsamında tutumla ilgili değişkenlerin matematik başarıyla ilişkilerini gösteren gözlenen puanlara dayalı korelasyon değeri, üç boyutlu modelle elde edilen korelasyon değeri ve dört boyutlu modelle elde edilen korelasyon değeri arasında manidar bir farkın olup olmadığını belirlemek amacıyla MedCalc 18.2.1 yazılım programı (MedCalc Software bvba,2018) kullanılmıştır. Bu program temelde korelasyon katsayılarını (r) z puanına dönüştürmektedir. Bu dönüşüm, Fisher'in r_z dönüşümü (Fisher's r to z transformation) olarak adlandırılmaktadır. Bu analizde sonraki aşamada $Z_{gözlenen}$ ($Z_{observed}$) puanı belirlenerek bu puanın .05 düzeyinde kritik değerden yüksek veya düşük olmasına göre korelasyon katsayılarının birbirinden manidar bir şekilde farklılaşıp farklılaşmadığı belirlenmektedir.

Araştırmada görsellerin oluşturulmasında R, SPSS, WinBUGS yazılım programı kullanılmıştır. Ülkelerin matematik başarıları ve tutumla ilgili değişkenlere göre dağılımlarına ait grafiklerin oluşturulmasında "cluster" (Maechler ve diğ., 2016)" paketi, ülkelerin matematik başarıları ile tutumla ilgili değişkenler arasındaki ilişkilere dair saçılma grafiklerinin oluşturulmasında "reshape2" (Wickham, 2007), "plotly" (Sievert, Parmer, Hocking, Chamberlain, Ram, Corvellec ve Despouy, 2017) ve "ggplot2" (Wickham, 2009) paketleri, ülkelerin tutumla ilgili değişkenlere ait düzeltilmemiş ve düzeltilmiş puanlarının dağılımına ait grafiklerin oluşturulmasında "ggplot2" (Wickham, 2009) paketleri kullanılmıştır. Bunların yanı sıra varsayımların test edilmesinde ve histogramların oluşturulmasında SPSS yazılım programı kullanılmıştır. Yakınsamaları değerlendirmek üzere oluşturulan zaman serisi grafikleri WinBUGS 1.4 (Spiegelhalter ve diğ., 2003) programından elde edilmiştir.

BÖLÜM 3

3.1 Bulgular ve Yorumlar

Bu bölümde, elde edilen bulgular alt amaçlara uygun olarak sırasıyla verilmiştir:

3.1.1 TIMSS 2015 Uygulamasına Katılan Ülkelerin Tutum Ölçeklerinden Elde Edilen UTS Puanları Farklılıklarına İlişkin Bulgular

Araştırmanın ilk sorusunda, analiz kapsamına alınan ülkelerin tutumla ilgili olan matematiği sevme, matematikte kendine güvenme ve matematiğe değer verme boyutlarında UTS gösterme eğilimleri arasında manidar bir fark olup olmadığı test edilmiştir. Bunu belirlemek için yapılan parametrik olmayan MANOVA sonuçlarına göre ülkeler UTS gösterme eğilimleri açısından manidar bir şekilde farklılaşmaktadır ($F_{(26.67,15864.54)}=69.87$, $p<.05$, Wilk's $\Lambda = 46.58$). Bunun yanı sıra UTS sergileme eğilimleri açısından bulunan farklılaşmanın hangi ülkeler arasındaki farklılaşmadan kaynaklandığını belirlemek amacıyla yapılan ülkelerin iki veya daha fazla olan çoklu karşılaştırmaları sonucunda her ülkenin birbirinden manidar bir şekilde farklılaştığı bulunmuştur. Analize ilişkin diğer sonuçlar Ek E.2'de verilmiştir. Bu bulgu Asya ülkelerinden (Çin, Endonezya, Malezya, Tai, Lübnan, Singapur ve Sri Lanka) gelen bireyler ile Avusturyalı bireylerin uç tepki stili sergileme eğilimindeki farklılığın varyans analizi ile incelendiği çalışmanın (Dolnicar ve Grun, 2007) bulguları ile benzerlik göstermektedir. Bireylerin hazır yiyecek (fast food) tüketimine ilişkin tutumları üzerinde UTS'nin etkisinin incelendiği bu çalışmanın sonunda Avustralyalı bireylerin Asya ülkelerinden gelen bireylere göre ölçeğin uç tepki kategorilerini daha fazla seçtikleri bulunmuştur. Mevcut araştırma kapsamına dâhil edilen ülkelerin tepki stillerindeki bu farklılaşma ülkelerin kültürel yönelimlerindeki farklılaşmayla açıklanabilir. Daha açıkça belirtmek gerekirse, mevcut araştırma kapsamında tepki stilleri incelenen ülkeler Avustralya haricinde genel olarak, Uzak Doğu (Singapur, Kore, Çin-Tayvan, Japonya ve Hong-Kong), Avrupa (İsveç, İtalya, Norveç ve Malta), Güney Amerika (Şili) ve Ortadoğu-Arap (Türkiye, Kuveyt, Mısır, Suudi Arabistan) ülkeleri olarak gruplandırıldığında Hofstede (2001), Uzak doğu ve Asya-Arap ülkelerinde güç mesafesinin fazla, toplulukçu bir yapının egemen olduğunu öte yandan Avrupa ülkelerinde ise güç mesafesinin az ve daha bireyci bir toplumun hâkim olduğunu belirtmiştir. Güç mesafesinin fazla olduğu toplumlarda, toplumdaki düzenin

sağlanabilmesi için eşitsizlik kabul edilebilmektedir. Bunun yanı sıra, toplulukçu kültürlerde grubun fikirlerindeki uyuşma önemlidir. Dolayısıyla bu toplumlarda uç noktalarda tepki verme eğilimleri daha azdır. Öte yandan, güç mesafesinin az olduğu toplumlarda ise her bir bireyin görüşlerine saygı duyulmakla beraber bireyci toplumlarda bireysel fikirler daha ön plandadır. Bu toplumdaki bireylerin tepki ölçeğinin orta kategorilerinde tepki sergileme eğilimleri daha azdır (Johnson ve diğ., 2005). Bu sebeplerden dolayı kültürel yönelimlerdeki farklılıkların ülkelerin UTS sergileme eğilimindeki farklılıklarla ilişkisinden söz etmek mümkündür. Bunun yanı sıra, öğrencilerin tepkileri üzerinde UTS gösterme eğilimindeki farklılıklar kendini gösterme stillerindeki (self-presentation style) farklılıklardan kaynaklı olabilir. Çünkü tepki stilleri agresif, aktif, pasif savunucu davranışların sergilenmesiyle ilgili olan kendini gösterme stillerinin göstergeleri olarak ele alınabilir ve iç içedir (He, 2015).

3.1.2 UTS'nin Bir Boyut Olarak Ele Alındığı ve Alınmadığı Modellerin Veri ile Uyumuna İlişkin Bulgular

UTS'nin bir boyut olarak ele alındığı ve alınmadığı modellerin veri ile uyumu belirlemek amacıyla iki farklı ÇBMK modeli analiz edilmiştir. Bu doğrultuda sadece tutumla ilgili olan matematiği öğrenmeyi sevme, matematikte kendine güven ve matematiğe değer verme değişkenlerinin birer boyut olarak dahil edildiği üç boyutlu model ile UTS'nin de ayrı bir boyut olarak dahil edildiği dört boyutlu model ayrı ayrı analiz edilerek modellere ilişkin çeşitli model-veri uyum istatistikleri elde edilmiştir. Bu model uyum istatistiklerinin değerleri Çizelge 5'de verilmiştir:

Çizelge 5.
Model Karşılaştırma Sonuçları

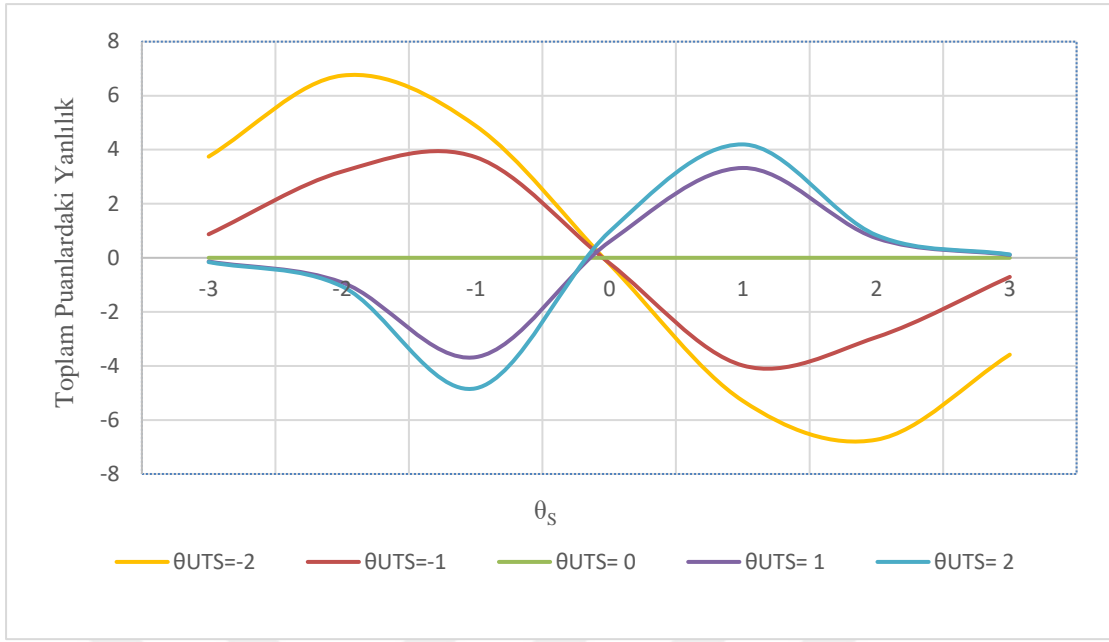
Modeller	Uyum İstatistikleri						
	LL	BIC	AIC	AIC3	CAIC	L ²	p
Üç Boyutlu Model	-209522.32	419820.91	419218.64	419305.64	419907.91	286174.19	7.7e-54656
Dört Boyutlu Model	-196641.21	394094.39	393464.43	393555.43	394185.39	260411.98	1.5e-49216

Çizelge 5'de görüldüğü üzere dört boyutlu model ile kestirilen model uyum indeks değerlerinin üç boyutlu model ile kestirilen model uyum indeks değerlerinden daha düşük olduğu bulunmuştur. Model uyum indekslerine ait daha düşük değerler daha iyi uyumun göstergesi olduğundan (Vermunt ve Magidson, 2005) dört boyutlu modelin veriye en iyi uyum gösteren model olduğu bulunmuştur. Bir diğer ifadeyle bireylerin

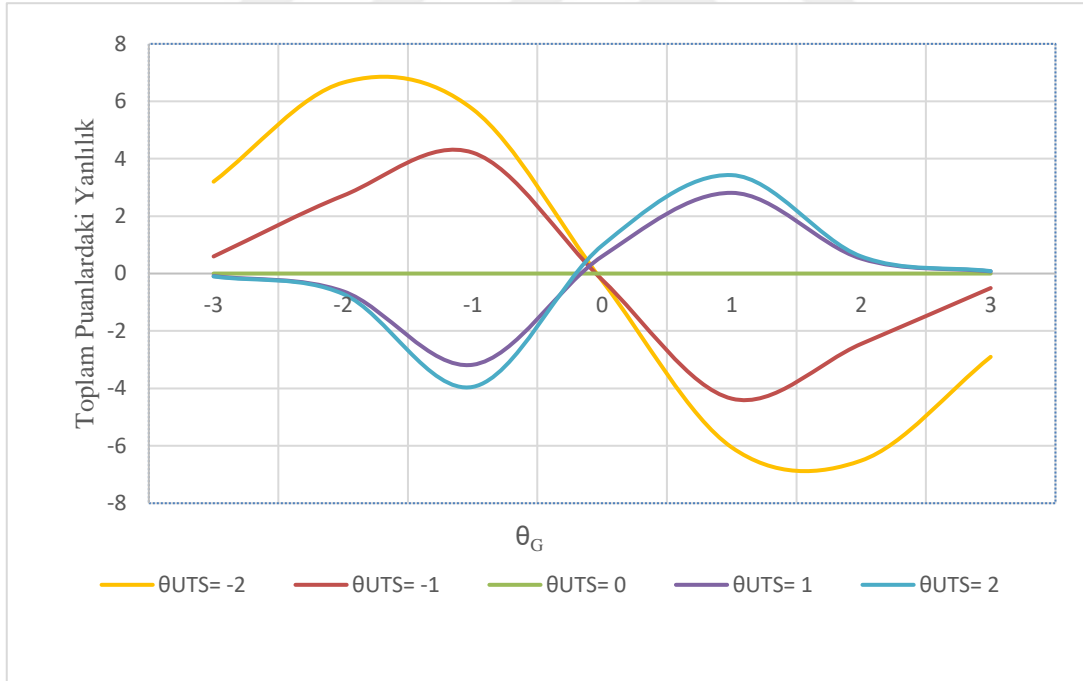
tutum maddelerine verdikleri tepkiler üzerinde UTS'nin etkisi söz konusudur. Bu durumda UTS'nin etkisi dikkate alındığında öğrencilerin tutumla ilgili değişkenlere ait örtük özellik düzeylerinin gerçek değerlerinin daha doğru ve net kestirildiği ifade edilebilir. Aynı zamanda, dört boyutlu modelde tutumla ilgili farklı ölçeklere ait madde tepki kategorilerine ve UTS'ye ilişkin yapılan sınırlamalar model-veri uyumunun artmasına katkı sağlamıştır. Benzer bulguya geniş ölçekli uluslararası çalışmalardan biri olan PISA 2006 uygulamasına katılan Amerika'daki öğrencilerin verdiği cevaplarda UTS'nin rolünün araştırıldığı çalışmada da (Bolt ve Newton, 2011) rastlanmıştır. Bu çalışmada, UTS'nin ayrı bir boyut olarak ele alındığı modelin veriyle daha iyi uyum sağladığı belirtilmiştir. Bu durum TIMSS, PISA gibi çeşitli kültürlerdeki öğrencilerin katıldığı geniş ölçekli uygulamalarda öğrencilerin tepkileri üzerinde uç tepki stilinin etkisinin olduğunu göstermektedir. Belirli standart test ve ölçek geliştirme adımlarının takip edildiği, geçerliliği ve güvenilirliği saptanmış ölçme araçlarının kullanıldığı TIMSS gibi geniş ölçekli uygulamadan elde edilen verilerin kullanıldığı bu çalışmada, UTS'nin etkisinin bulunmasında öğrencilerin farklı toplum yapılarından gelmiş olmasının (Hosftede, 2001) yanı sıra çalışmaya dâhil edilen öğrencilerin sosyo-ekonomik düzeylerinin de farklı olmasının rolü olabilir (Hooper ve diğerleri, 2013). Çünkü sosyo-ekonomik düzeyi düşük olan toplumların UTS sergileme eğilimleri yüksektir (Meisenberg ve Williams, 2008).

3.1.3 Toplam Ölçek Puanlarına UTS'nin Karıştırdığı Yanlılığın Miktarının ve Yönünün Farklı UTS Değerlerine göre Değişimine İlişkin Bulgular

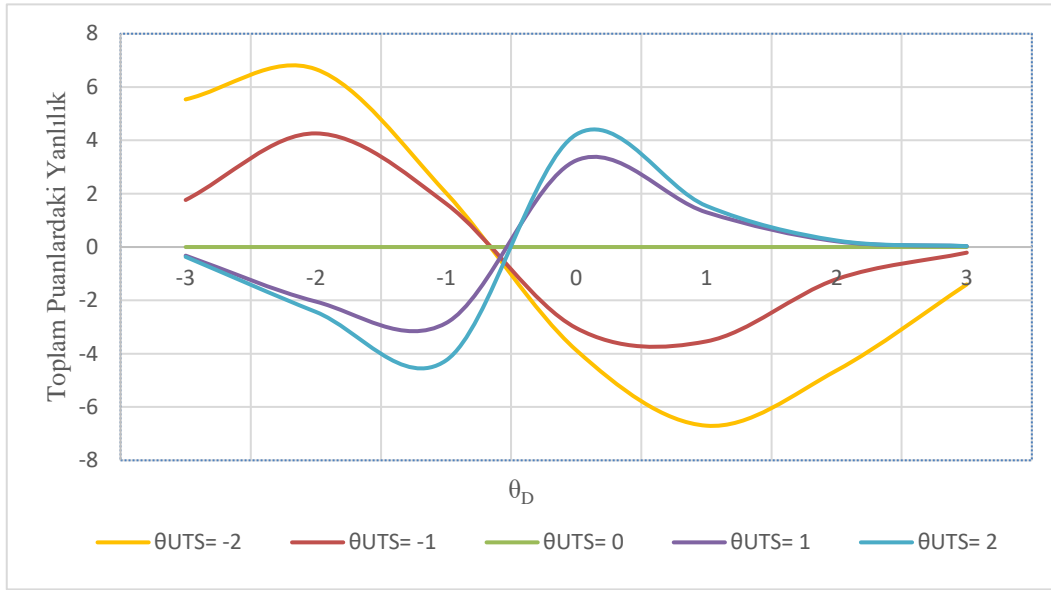
TIMSS 2015 uygulamasına katılıp çalışmaya dâhil edilen ülkelerdeki öğrencilerin UTS sergiledikleri belirlendikten sonra öğrencilerin matematiği öğrenmeyi sevmeye, matematikte kendine güvenme ve matematiğe değer verme değişkenlerinin ölçülmesinin amaçlandığı ölçeklerin her birinden elde edilen toplam puanlarındaki yanlılığın miktarı θ_{UTS} 'nin farklı değerlerine (örneğin $\theta_{UTS}=-2, -1,0,1,2$) göre belirlenmiştir. Bu doğrultuda eşitlik 2 ve eşitlik 3'den yararlanılmıştır. Buna göre tutumla ilgili her bir değişken için toplam puanlara karışan yanlılığın θ_{UTS} 'nin farklı değerlerine göre miktarı Şekil 7 , Şekil 8 ve Şekil 9'da gösterilmiştir:



Şekil 7. Matematiği Öğrenmeyi Sevme Ölçeğinden Elden Edilen Toplam Puanlardaki Yanlılık



Şekil 8. Matematikte Kendine Güven Ölçeğinden Elde Edilen Toplam Puanlardaki Yanlılık



Şekil 9. Matematığe Değer Verme Ölçeğinden Elde Edilen Toplam Puanlardaki Yanlılık

Şekil 7, Şekil 8 ve Şekil 9'da toplam puanlar üzerindeki yanlılığın miktarı ve etkisinin yönünün sırasıyla θ_S , θ_G ve θ_D ve θ_{UTS} 'nin bir fonksiyonu olarak değiştiği görülmektedir. Bunlara göre TIMSS 2015 uygulamasına katılan araştırma kapsamına dahil edilen öğrencilerin matematiği sevmeye, matematikte kendine güven ve matematiğe değer verme boyutlarına ilişkin tutum düzeylerinin sırasıyla yaklaşık olarak “-.2”, “-.2” ve “-.6” olduğu yerde UTS'nin farklı düzeylerini temsil eden bütün eğrilerin kesiştiği ve bu kesişim noktalarında yanlılığın miktarının yaklaşık olarak “0” olduğu bulunmuştur. Bir diğer anlatımla, bu öğrencilere ait θ_S , θ_G , θ_D 'nin değeri sırasıyla “-.2”, “-.2” ve “-.6”dan farklılaştıkça toplam puanlara yanlılığın karıştığı bulunmuştur. Daha açıkça belirtmek gerekirse, öğrencilerin maddelere ilişkin tepkileri üzerinde UTS'nin etkisinin mutlak değerce artması, öğrencilerin matematikte kendine güven, matematiği sevmeye ve matematiğe değer verme boyutlarına ilişkin tutum düzeylerinin sırasıyla “-.2”, “-.2” ve “-.6”dan yüksek olduğu durumda (olumlu yönde tutumu bildiren “katılıyorum” ve “kesinlikle katılıyorum” tepki kategorilerini daha çok tercih ettikçe) bu boyutlarla ilgili uygulanan ölçeklerden elde edilen toplam puanlara pozitif yönde yanlılığın karışmasına yol açtığı bulunmuştur. Öte yandan θ_{UTS} 'nin değerinin artması, öğrencilerin theta değerlerinin “-.2”, “-.2” ve “-.6”dan düşük olduğu durumda (olumsuz yönde tutumu bildiren “katılmıyorum” ve “hiç katılmıyorum” tepki kategorilerini daha çok tercih ettikçe) bu boyutlarla ilgili uygulanan ölçeklerden elde edilen toplam puanlara negatif yönde yanlılığın karışmasına neden olduğu bulunmuştur. Bu durum, UTS'nin ölçülmek

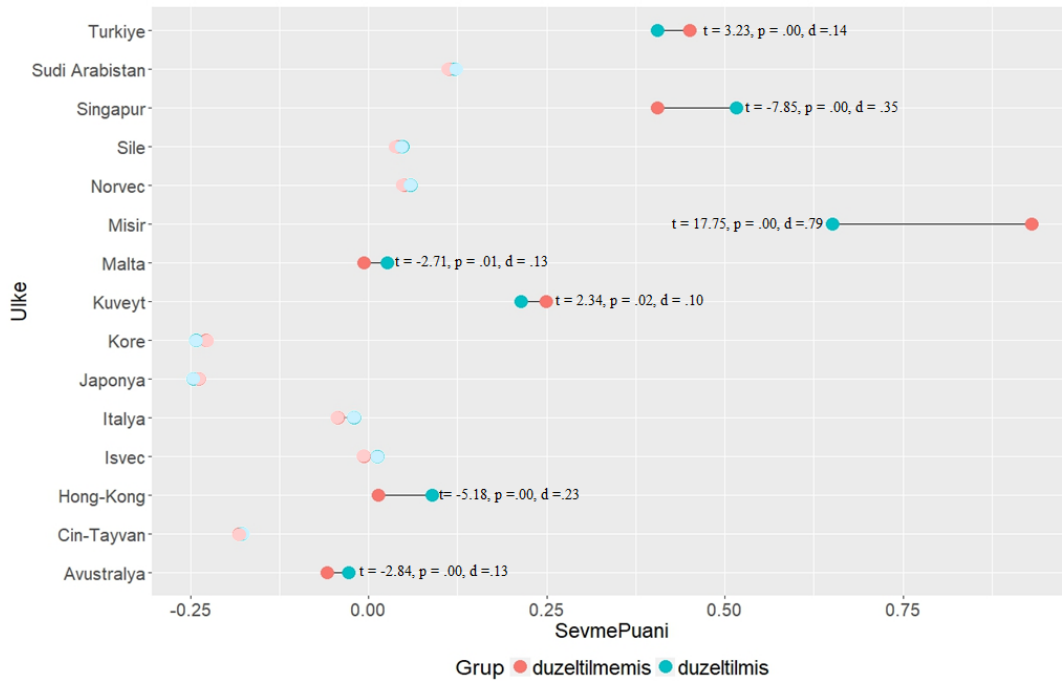
istenen özelliğin farklı düzeylerinde etkisinin değiştiğini ve bu etkinin doğrusal olmadığını göstermektedir. Aynı zamanda bu bulgudan yola çıkılarak örneğin matematiği öğrenmeyi sevme boyutuna ilişkin yetenek düzeyi -.2'den yüksek olan öğrencilerin 4 kategorisini 3 kategorisine göre daha çok tercih ettiği, yetenek düzeyi -.2'den düşük öğrencilerin ise 1 kategorisini 2 kategorisine kıyasla daha çok tercih ettiği ifade edilebilir.

Özetle, UTS'den kaynaklı yanlılık bazı durumlarda öğrencilerin tutum puanlarının artmasına bazı durumlarda ise azalmasına neden olmaktadır. Daha açıkça ifade etmek gerekirse bu çalışmada TIMSS 2015 uygulamasına katılan ülkelerdeki öğrencilerin yüksek düzeyde UTS sergilemeleri ve UTS'nin doğasına bağlı olarak sistematik bir biçimde uç noktayı seçmeleri tutumlarının daha olumlu olduğu durumlarda bu değişkenlerle ilgili ölçeklerden elde edilen toplam puanlarının olduğundan daha fazla olmasına yol açmaktadır. Öte yandan bu boyutlara ait tutumlarının daha olumsuz olduğu durumlarda fazla düzeyde UTS sergilemelerinin toplam puanlarının olduğundan daha az olmasına neden olduğu bulunmuştur. Bu bulgu, PISA uygulamasına katılan bazı ülkelerin fene yönelik tutumla ilgili farklı değişkenlerin ölçülmesinin amaçlandığı ölçeklerden elde edilen toplam puanlardaki UTS'nin etkisinin araştırıldığı çalışmanın (Lu ve Bolt, 2015) bulguları ile paralellik göstermektedir. Araştırmacılar, farklı yetenek düzeylerindeki bireylerin tepki kategorilerini seçme eğilimleri üzerinde UTS'nin etkisinin farklılaştığını bulmuşlardır. Öğrencilerin tutum puanlarının daha olumlu veya daha olumsuz olmasında tutumlarının gerçekten olumlu veya olumsuz olmasının değil UTS sergileme düzeylerinin etkili olduğu belirtilebilir. Bu durumda UTS'nin öğrencilerin ölçülmek istenen özelliklerinin doğru olarak ölçülmesinde sistematik bir hata olan yanlılığı karıştırdığı (Moors, 2004) ve ölçme sonuçlarının geçerliliğini tehdit ettiği ifade edilebilir (Messick, 1962).

3.1.4 Öğrencilerin UTS'nin Etkisine göre Düzeltilmemiş ve Düzeltilmiş Ortalama Tutum Puanlarının Arasındaki Farklara İlişkin Bulgular

TIMSS 2015 uygulamasına katılan ülkelerdeki öğrencilerin UTS'nin etkisine göre düzeltilmemiş ve düzeltilmiş tutum puanlarını belirlemek amacıyla geliştirilen üç ve dört boyutlu modellerin analizlerinin sonuçlarına göre öğrencilerin cevapları üzerinde UTS'nin etkisinin değiştiği bulunmuştur. Değişimin manidar olduğu ülkeler incelendiğinde, UTS bazı öğrencilerin tutumla ilgili yetenek düzeylerini olumlu yönde

etkilemekte bazı öğrencilerin ise tutumla ilgili yetenek düzeylerini olumsuz yönde etkilemektedir. Bu durum tutumla ilgili her üç değişken için ülke düzeyinde ele alındığında, bazı ülkelerin düzeltilmiş ortalama tutum puanlarının daha düşük olduğu öte yandan bazı ülkelerin ise düzeltilmiş ortalama tutum puanlarının daha yüksek olduğu bulunmuştur. Araştırmaya dahil edilen ülkelerin matematiği öğrenmeyi sevme, matematikte kendine güven ve matematiğe değer verme değişkenlerine ilişkin düzeltilmemiş ve düzeltilmiş ortalama tutum puanlarının dağılımı sırasıyla Şekil 10, Şekil 11 ve Şekil 12’de verilmiştir:



Şekil 10. Ülkelerin Matematik Öğrenmeyi Sevme Boyutuna ve İlişkin Düzeltilmemiş ve Düzeltilmiş Puanlarının Dağılımı

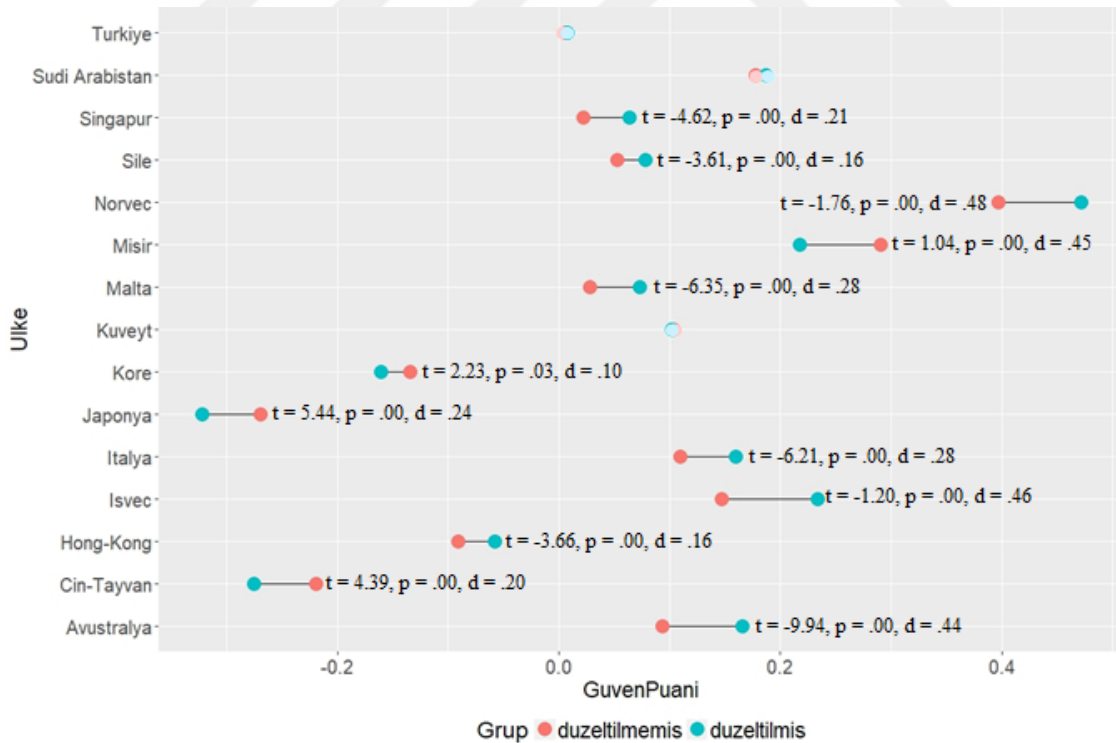
Not. Düzeltilmemiş ve düzeltilmiş tutum puanları arasında manidar fark olmayan ülkelerin puanları soluk olarak gösterilmiştir

Şekil 10’da görüldüğü üzere Singapur, Malta, Hong-Kong ve Avustralya’nın matematiği öğrenmeyi sevme boyutuna ilişkin UTS’nin etkisinin dikkate alınmadığı düzeltilmemiş ortalama tutum puanlarının düzeltilmiş tutum puanlarından daha düşük olduğu bulunmuştur. UTS’nin etkisi dikkate alındığında ise bu ülkelerin düzeltilmiş tutum puanlarının arttığı görülmektedir. Bu durumda, UTS’nin bu ülkelerin matematiği öğrenmeyi sevme boyutuna ilişkin ortalama tutum puanlarını azaltıcı yönde rol oynadığı ifade edilebilir. Öte yandan, Şekil 10’da Türkiye, Mısır ve Kuveyt’in UTS’nin etkisi dikkate alındığında matematiği öğrenmeyi sevme boyutuna ilişkin ortalama tutum puanlarının düştüğü görülmektedir. Bir başka anlatımla, UTS, bu ülkelerdeki

öğrencilerin tutum puanlarının olduğundan daha fazla olarak kestirilmesinde yol açan faktörlerden biri olabilir. Bunların yanı sıra, geriye kalan ülkelerin UTS'nin etkisine göre düzeltilmemiş ve düzeltilmiş ortalama tutum puanlarının birbirine yakın olduğu bulunmuştur.

Araştırmaya dahil edilen ülkelerdeki öğrencilerin matematiği öğrenmeyi sevme boyutuna ilişkin UTS'nin etkisine göre düzeltilmemiş ve düzeltilmiş ortalama tutum puanlarının arasında manidar farkın olup olmadığı incelendiğinde Şekil 10'de de görüldüğü üzere düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasında manidar fark bulunan ülkelerin Singapur, Hong-Kong, Malta, Avustralya, Türkiye, Kuveyt ve Mısır olduğu bulunmuştur ($p < .05$). Manidar olan farklılıklara ilişkin etki büyüklüğü değerleri (d) incelendiğinde ise etki büyüklüğü değerlerinin çoğunlukla küçük olduğu Mısır'a ait etki büyüklüğü değerinin orta olduğu görülmektedir. Diğer ülkelerdeki öğrencilere kıyasla Mısır'daki öğrencilerin cevaplarında UTS'nin etkisinin daha fazla olduğu ifade edilebilir.

Ülkelerin matematikte kendine güven değişkenine ilişkin düzeltilmemiş ve düzeltilmiş ortalama tutum puanlarının dağılımı Şekil 11'de verilmiştir:

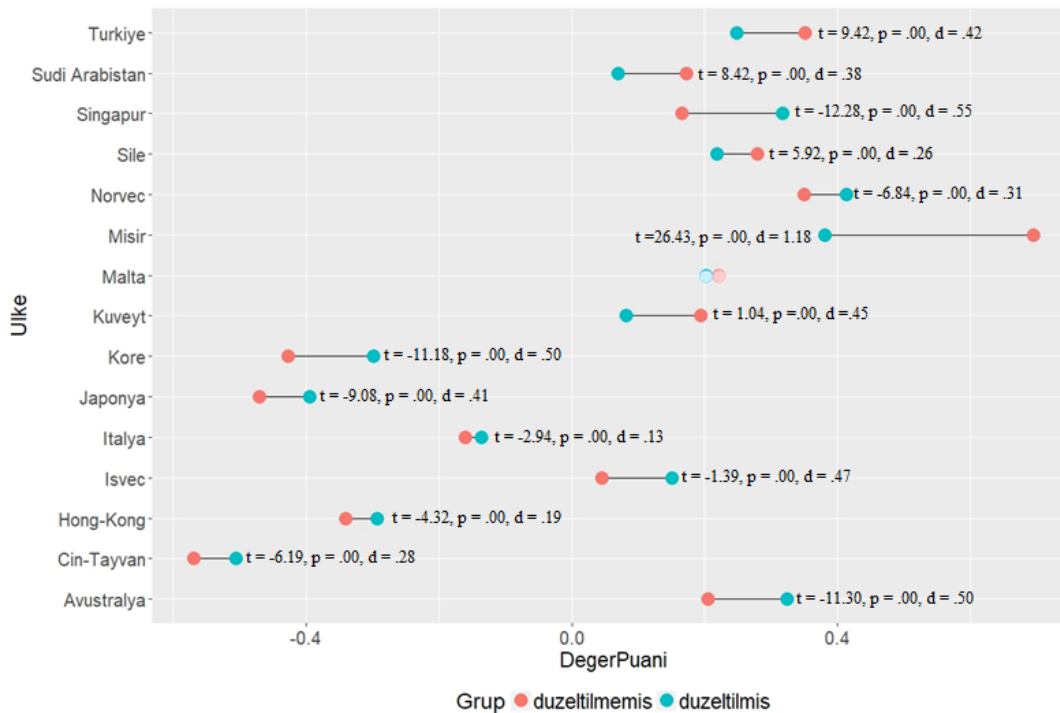


Şekil 11. Ülkelerin Matematikte Kendine Güvene İlişkin Düzeltilmemiş ve Düzeltilmiş Puanlarının Dağılımı

Şekil 11’de Singapur, Şili, Norveç, Malta, İtalya, İsveç, Hong-Kong ve Avustralya’nın matematikte kendine güven değişkenine ilişkin UTS’nin etkisinin dikkate alınmadığı durumda elde edilen düzeltilmemiş ortalama tutum puanlarının düzeltilmiş ortalama tutum puanlarına kıyasla daha düşük olduğu görülmektedir. Bir diğer ifadeyle, UTS bu ülkelerdeki öğrencilerin tutum puanlarının olduğundan daha düşük olarak kestirilmesine neden olmuş olabilir. Bu ülkelerin aksine Mısır, Kore, Japonya ve Çin-Tayvan’ın düzeltilmemiş ortalama tutum puanlarının daha yüksek olduğu bulunmuştur. Bu nedenle, UTS’nin öğrencilerin matematikte kendine güven değişkenine ilişkin ortalama tutum puanlarının olduğundan daha yüksek olarak kestirilmesine yol açtığı belirtilebilir. Geriye kalan ülkelerin ise düzeltilmemiş tutum puanları ile düzeltilmiş tutum puanlarının birbirine yakın olduğu bulunmuştur.

Şekil 11’de görüldüğü üzere matematikte kendine güven boyutuna ilişkin düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasında manidar fark bulunan ülkeler Singapur, Kore, Çin-Tayvan, Hong-Kong, Japonya, İsveç, İtalya, Malta, Avustralya, Norveç, Şili ve Mısır’dır ($p < .05$). Manidar olan farklılıklara ilişkin etki büyüklüğü değerlerinin çoğunlukla küçük olduğu bulunmuştur.

Matematiğe değer verme değişkenine ilişkin düzeltilmemiş ve düzeltilmiş ortalama tutum puanlarının dağılımı Şekil 12’de verilmiştir:



Şekil 12. Ülkelerin Matematiğe Değer Vermeye İlişkin Düzeltilmemiş ve Düzeltilmiş Puanlarının Dağılımı

Şekil 12'den anlaşıldığı üzere Türkiye, Sudi Arabistan, Şili, Mısır ve Kuveyt'in matematiğe değer verme değişkenine ait UTS'nin etkisine göre düzeltilmemiş ortalama tutum puanlarının düzeltilmiş ortalama tutum puanlarından daha yüksek olduğu bulunmuştur. Bu durumda, UTS'nin bireylerin tutum puanları üzerinde artırıcı yönde rol oynadığı ifade edilebilir. Bu ülkelerin aksine Singapur, Norveç, Kore, Japonya, İsveç, Hong-Kong, Çin-Tayvan ve Avustralya'nın düzeltilmemiş ortalama tutum puanlarının ise daha düşük olduğu bulunmuştur. Bu ülkelerde ise UTS, öğrencilerin tutum puanlarını düşürücü yönde rol oynamış olabilir. Bunların yanı sıra, Malta ve İtalya'nın düzeltilmemiş ve düzeltilmiş tutum puanlarının yakın olduğu bulunmuştur.

Matematiğe değer verme boyutuna ilişkin düzeltilmemiş ve düzeltilmiş tutum puanları arasında manidar fark bulunmayan ülkenin sadece Malta olduğu bulunmuştur. Bir diğer anlatımla araştırmaya geriye kalan ülkelerdeki öğrencilerin (Türkiye, Sudi Arabistan, Singapur, Şili, Norveç, Mısır, Kuveyt, Kore, Japonya, İtalya, İsveç, Hong-Kong Çin-Tayvan ve Avustralya) matematiğe değer verme boyutuna ilişkin düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasında manidar fark bulunmuştur ($p < .05$). Manidar farklılıkların pratikteki manidarlığı incelendiğinde ise etki büyüklüğü değerlerinin çoğunlukla küçük ve orta büyüklükte değiştiği, düzeltilmemiş ve düzeltilmiş tutum puanları arasında büyük fark olan ülkenin ise Mısır olduğu bulunmuştur.

Ülkelerin tutumla ilgili her üç değişkene ilişkin düzeltilmemiş ve düzeltilmiş ortalama tutum puanları bütüncül olarak değerlendirildiğinde, düzeltilmiş ortalama tutum puanlarının düzeltilmemiş ortalama tutum puanlarından manidar şekilde farklılaştığı ülkelerin sayıca fazla olduğu boyutun, matematiğe değer verme boyutu olduğu belirtilebilir. Bir diğer anlatımla, ülkelerin bu boyuta ilişkin tutum puanları uç tepki stilinden daha fazla etkilenmekte ve bu yüzden düzeltilmiş ve düzeltilmemiş puanlar birbirinden farklılaşmaktadır. Bu durum bir psikolojik yapı olarak değerler kültürdeki önemiyle açıklanabilir. Daha açıkça belirtmek gerekirse, kültürün temel çekirdeği olan değerler (Hofstede, Hofstede ve Minkov, 2010) sayesinde kültür kendi kendini beslemekte, eğitimle veya aileyle bireyden bireye aktarılmaktadır. Okul ortamında ise eğitim sisteminin bir parçası olan öğretmenler vasıtasıyla öğrencilere yeni değerler kazandırılmakta, ailelerinden gelen veya kendilerinin oluşturdukları değerler değiştirilebilmektedir (Österling, 2013). Bu noktada, eğitim sisteminde önemli bir yere sahip olan okul matematiğinin değer odaklı (value-laden) olduğu ifade edilebilir (Bishop, FitzSimons, Seahve Clarkson, 1999). Bu araştırmada incelenen matematiğe

değer verme boyutunda çoğunlukla öğrencilerin matematiğin gelecekteki faydası, buna dair beklentileriyle ilgili olması sebebiyle toplumun beklentilerinin, değer yargılarının, toplumun birer bireyi olan öğrencilerin derslerdeki beklentilerini ve derse verdikleri değeri şekillendirdiği belirtilebilir. Bu bağlamda, değerın kültürdeki kritik yeri dikkate alındığında farklı kültürlerin dahil edildiği bu araştırmada bireylerin matematiğe değer vermeleri üzerinde tutumla ilgili diğer boyutlara kıyasla çeşitli kültür yapılarının etkisinin daha baskın olduğunu söylemek mümkündür. Öğrencilerin cevapları üzerindeki bu kültürel etki kendini öğrencilerin farklı tepki stilleri sergilemesiyle göstermektedir (Harzing, 2006). Kültürün tepki stillerinin sergilenmesi üzerindeki etkisi uç tepki stilinin sergilenmesi açısından değerlendirildiğinde, çalışmaya dahil edilen ülkelerin UTS'nin etkisine göre düzeltilmemiş ve düzeltilmiş tutum puanları arasında manidar farkın olmasının ve UTS'nin etkisine göre düzeltilmiş tutum puanlarının daha yüksek bazılarının ise daha düşük olmasının altında yatan nedenlerden biri olduğu ifade edilebilir.

Araştırma kapsamına dahil edilen ülkelerin düzeltilmemiş ortalama tutum puanlarının düzeltilmiş ortalama tutum puanlarına kıyasla daha düşük olduğu ülkelerin genel olarak matematik başarısı yüksek ve orta düzeyde olan ülkeler olduğu ifade edilebilir. Bu ülkelerde UTS'nin, öğrencilerin matematiğe yönelik tutumlarının olduğundan daha olumsuz olarak kestirilmesinde rol oynadığı ifade edilebilir. Öte yandan, düzeltilmemiş ortalama tutum puanlarının düzeltilmiş ortalama tutum puanlarına kıyasla daha yüksek olduğu ülkelerin ise genel olarak matematik başarısı düşük olan ülkeler olduğu bulunmuştur. Bu durumda ise UTS'nin, bu ülkelerdeki öğrencilerin matematiğe yönelik tutumlarının olduğundan daha olumlu olarak kestirilmesine yol açan nedenlerden biri olduğu belirtilebilir. Benzer bulguya Van de Gaer ve Adams (2010)'un PISA 2003 uygulamasına katılan ülkelerin UTS'nin etkisinin dikkate alındığı ve alınmadığı durumlarda matematiğe yönelik ilgilerine ilişkin puanlarındaki değişimin incelendiği araştırmada da ulaşılmıştır. Araştırmacılar, matematik başarısı yüksek olan ülkelerin (Kore, Hong-Kong, Japonya ve Finlandiya) UTS'nin etkisine göre matematiğe yönelik ilgilerine ait düzeltilmemiş puanlarının düzeltilmiş puanlarına kıyasla daha düşük olduğunu öte yandan matematik başarısı düşük olan ülkelerin (Endonezya, Tunus ve Brezilya) ise matematiğe yönelik ilgilerine ait düzeltilmemiş puanlarının düzeltilmiş puanlarına kıyasla daha yüksek olduğunu belirtmişlerdir. Özetle, beklentilerin aksine başarılı olan ülkelerin tutumlarının olumsuz

yönde veya başarısız olan ülkelerin tutumlarının olumlu yönde olmasında UTS rol oynadığı belirtilebilir.

TIMSS 2015 uygulamasına katılan ve araştırmaya dahil edilen ülkelerdeki öğrencilerin birey düzeyinde düzeltilmemiş ve düzeltilmiş ortalama tutum puanlarının değişimi incelendiğinde ise ülke düzeyindeki duruma benzer olarak UTS'nin etkisinin ve yönünün değiştiği bulunmuştur. Bu bağlamda uç noktalardaki ve orta noktalardaki tepki kategorilerini seçen öğrencilere tipik olarak örnek gösterilebilecek bireylerin düzeltilmemiş ve düzeltilmiş ortalama tutum puanları Çizelge 6'da sunulmuştur:

Çizelge 6.
Bazı Bireylerin Düzeltilmemiş ve Düzeltilmiş Ortalama Tutum Puanları

Birey No	Tepki Örüntüsü	Üç Boyutlu Model			Dört Boyutlu Model			
		θ_S	θ_G	θ_D	θ_S	θ_G	θ_D	θ_{UTS}
471	4 4	2.84	1.68	1.63	2.26	1.50	1.31	.73
499	4 1 1 4 4 4 4 4 4 4 4 1 1 4 1 4 4 4 1 1 4 4 4 4 4 4 4 4 4 4 4	.78	-.17	1.18	.30	-.11	.52	1.78
1982	1 1	-2.90	-1.93	-2.42	-1.98	-1.5	-1.94	1.12
6158	3 3 3 2 2 2 2 2 2 2 3 3 2 2 3 2 2 3 2 2 3 3 2 2 3 3 3 2 3 3	-.24	-.02	-.58	-.33	-.02	-.52	-1.03

Çizelge 6'da 471 ve 1982 numaralı öğrencilerin tutumla ilgili maddeleri cevaplarken sırasıyla sürekli en olumlu ve en olumsuz tutumu betimleyen tepki kategorisini seçtiği görülmektedir. Bunun yanı sıra 499 numaralı bireyin ise maddelerin en olumlu ve en olumsuz tutumu betimleyen tepki kategorilerinin ikisini de seçtiği orta noktalardaki tepki kategorilerini ise seçmediği görülmektedir. Bu sebeple özellikle bu şekilde cevap örüntüsüne sahip öğrencilerin tutum düzeylerinin kestiriminde UTS'den kaynaklı yanlışlık söz konusu olabilir. Aynı zamanda, Çizelge 6'dan da görüldüğü üzere daha çok uç değerlerin yer aldığı 1., 2.ve 3. tepki örüntülerine ilişkin kestirilen θ_{UTS} değerlerinin UTS'nin doğasına/tanımına uygun olarak daha yüksek olduğu bulunmuştur. Ayrıca UTS'nin öğrencilerin ortalama tutum puanları üzerindeki ülke düzeyindeki etkisine benzer olarak UTS'nin birey düzeyindeki etkisinin de öğrencilerin tutum düzeylerine göre değiştiği bulunmuştur. Daha açıkça ifade etmek gerekirse UTS, matematiğe yönelik tutumu olumlu olan öğrencilerin ortalama tutum puanlarının gerçek tutum puanlarına göre yüksek olarak belirlenmesine neden olmaktadır. Öte yandan, UTS'nin matematiğe yönelik tutumu olumsuz olan öğrencilerin tutum puanlarının gerçekteki tutum puanlarından daha düşük olarak kestirilmesine neden olduğu bulunmuştur. Bu bulgu, hasta bireylerin nevrotikliğe (neuroticism), bilinçliliğe (conscientiousness) ve uygunluğa (agreeableness) ilişkin yetenek düzeylerinin UTS'nin etkisinin dikkate alınmadığı ve alındığı durumlar için belirlendiği çalışmanın (Dowling,

Bolt, Deng ve Li, 2016) bulguları ile benzerlik göstermektedir. Araştırmacılar, UTS'nin etkisinin dikkate alındığı modelde bireylerin bu değişkenlere ilişkin tutum düzeylerinin farklılaştığını, UTS'nin etkisinin yönünün bireylerin tutumlarının olumlu veya olumsuz yönde olmasına göre değiştiğini belirtmişlerdir.

3.1.5 UTS'nin Etkisine göre Düzeltilmemiş ve Düzeltilmiş Tutum Puanlarının

Matematik Başarısıyla olan Ülke İçi ve Ülkeler Arası İlişkilere İlişkin Bulgular

TIMSS 2015 uygulamasına katılan ve araştırmaya dahil edilen ülkelerdeki öğrencilerin UTS'nin etkisine göre düzeltilmemiş ve düzeltilmiş tutum puanlarının matematik başarısıyla olan ilişkilerine dair ülke içi korelasyon değerleri Çizelge 7'de verilmiştir:

Çizelge 7.

Ülkelerin Düzeltilmemiş ve Düzeltilmiş Tutum Puanlarının Matematik Başarısıyla Olan İlişkilerine Dair Ülke İçi Korelasyon Değerleri

Ülkeler	Matematiği Öğrenmeyi Sevme		Matematikte Kendine Güven		Matematiğe Değer Verme	
	a	b	a	b	a	b
Singapur	.28	.33	.39	.50	.11	.22
Kore	.41	.47	.50	.63	.39	.42
Çin- Tayvan	.44	.49	.51	.63	.35	.41
Hong-Kong	.32	.32	.37	.44	.19	.25
Japonya	.35	.39	.44	.54	.22	.28
İsveç	.42	.46	.63	.68	.19	.28
İtalya	.36	.38	.52	.57	.19	.25
Malta	.29	.32	.41	.50	.11	.22
Avustralya	.29	.35	.50	.60	.19	.27
Norveç	.37	.42	.61	.66	.20	.28
Türkiye	.19	.23	.52	.59	.11	.15
Şili	.19	.26	.45	.53	.13	.21
Kuveyt	.20	.23	.33	.44	.18	.27
Mısır	.24	.33	.37	.50	.14	.30
Sudi Arabistan	.28	.25	.39	.54	.11	.26

a = Düzeltilmemiş Tutum Puanlarının Matematik Başarısıyla İlişkilerine Dair Korelasyon Değeri, b = Düzeltilmiş Tutum Puanlarının Matematik Başarısıyla İlişkilerine Dair Korelasyon Değeri

Çizelge 7 de görüldüğü üzere araştırmaya dahil edilen ülkelerdeki öğrencilerin matematiği öğrenmeyi sevme, matematikte kendine güven ve matematiğe değer verme boyutlarına ilişkin düzeltilmemiş ve düzeltilmiş tutum puanlarının matematik başarısıyla ilişkilerine dair ülke içi korelasyon değerleri pozitif yönde olmasının yanı sıra birbirine benzerdir. Bir diğer ifade ile her iki durumda da (UTS'nin etkisinin kontrol

edilmediği ve edildiği) matematiğe yönelik tutumu olumlu olan öğrencilerin daha yüksek düzeyde matematik başarısına sahip olduğu ifade edilebilir. Bunun yanı sıra korelasyon değerlerinin her iki durumda da benzer çıkmasından dolayı UTS'nin öğrencilerin tutumlarının matematik başarısıyla ilişkileri üzerinde önemli bir etkiye sahip olmadığı belirtilebilir.

TIMSS 2015 uygulamasına katılan ülkelerdeki öğrencilerin UTS'nin etkisine göre düzeltilmemiş ve düzeltilmiş tutum puanlarının matematik başarısıyla ülkeler arası ilişkileri ise Çizelge 8'de verilmiştir:

Çizelge 8.

Ülkelerin Düzeltilmemiş ve Düzeltilmiş Tutum Puanlarının Matematik Başarısıyla Olan Ülkeler Arası İlişkileri

Değişken Çiftleri	Gözlenen Puanlara Dayalı Korelasyon Değeri (r_{s1})	Üç Boyutlu Modelle Kestirilen Puanlara İlişkin Korelasyon Değeri	Dört Boyutlu Modelle Kestirilen Puanlara İlişkin Korelasyon Değeri
Başarı-Sevme	-.59	.08	.15
Başarı-Güven	-.74	.04	.19
Başarı-Değer	-.65	.01	.13
Başarı-UTS	-	-	-.09

Çizelge 8' de görüldüğü üzere araştırmada, TIMSS 2015 uygulamasına katılan ülkelerdeki öğrencilerin matematik başarısı ile matematiği öğrenmeyi sevmeleri arasında ülke düzeyinde negatif yönde manidar ilişkisinin olduğu bulunmuştur ($r_{s1} = -.59$, $p < .01$). Bir diğer anlatımla, matematik başarıları yüksek olan öğrencilerin matematiği sevmeme eğiliminde olduğu bulunmuştur. Öte yandan, üç boyutlu MTK modeline göre kestirilen puanlara ilişkin ülke düzeyinde korelasyon değerine göre bu öğrencilerin matematik başarısı ile matematiği öğrenmeyi sevmeleri arasında pozitif yönde düşük düzeyde ilişki bulunmuştur. Benzer şekilde, UTS'nin etkisinin dikkate alındığı dört boyutlu MTK modeliyle kestirilen puanlara ilişkin ülke düzeyindeki korelasyon değerine göre de öğrencilerin matematik başarısı ile matematiği öğrenmeyi sevmeleri arasında pozitif yönde benzer düzeyde ilişki bulunmuştur.

Araştırmadan elde edilen bir diğer bulgu ise TIMSS 2015 uygulamasına katılan ülkelerdeki öğrencilerin matematik başarısı ile matematikte kendine güvenleri arasında ülke düzeyinde negatif yönde manidar ilişki olduğudur ($r_{s2} = -.74$, $p < .05$). Bir diğer anlatımla, öğrencilerin matematik başarısı yüksek olan öğrencilerin matematikte

kendine güvenmeme eğiliminde olduğu görülmektedir. Öte yandan, ÇBMTK modellerine göre kestirilen puanlara ilişkin ülke düzeyinde korelasyon katsayılarına göre bu öğrencilerin matematik başarısı ile matematikte kendine güvenmeleri arasında düşük düzeyde pozitif yönde ilişki bulunmuştur.

Çizelge 8’de belirtildiği gibi TIMSS 2015 uygulamasına katılan ülkelerdeki öğrencilerin matematik başarısı ile matematiğe değer vermeleri arasında ülke düzeyinde negatif yönde yüksek düzeyde manidar ilişkisinin olduğu bulunmuştur ($r_{s3} = -.65$, $p < .05$). Bir başka deyiş ile matematik başarısı yüksek olan öğrencilerin matematiğe daha az değer verme eğiliminde olduğu görülmektedir. Öte yandan, ÇBMTK modellerine göre kestirilen puanlara ilişkin ülke düzeyinde korelasyon değerlerine göre öğrencilerin matematik başarısı ile matematiğe değer vermeleri arasında pozitif yönde düşük düzeyde ilişki olduğu bulunmuştur.

Özetle, araştırmaya dahil edilen ülkelerdeki öğrencilerin gözlenen puanlara dayalı olarak tutumu ile matematik başarıları arasında ülke düzeyinde yüksek düzeyde negatif yönde ilişki olduğu, bu ilişkilerin örtük özellik puanlarına dayalı olarak üç boyutlu modele göre kestirildiğinde pozitif yönde değiştiği ve miktarının düştüğü, UTS’nin etkisi dikkate alındığında ise benzer miktarlarda pozitif yönde ilişkilerin olduğu bulunmuştur. Bir diğer anlatımla, matematik başarısı ile matematiğe yönelik tutum arasındaki ülke düzeyindeki ilişkilerin ÇBMTK modelleriyle kestirilen puanlara dayalı olarak belirlendiğinde birey düzeyindeki ilişkilere benzer olarak pozitif yönde olduğu ifade edilebilir. Bu bulgu TIMSS gibi geniş ölçekli uygulamalardan biri olan PISA 2006 uygulamasına katılan öğrencilerin fene yönelik tutumları ile fen başarıları arasında ilişkinin UTS’nin etkisinin dikkate alındığında nasıl değiştiğinin incelendiği ve ülke düzeyindeki negatif ilişkilerin UTS’nin etkisi dikkate alındığında pozitif yönde değiştiği ve ilişkilerin düzeyinin düştüğü araştırmanın bulguları ile desteklenmektedir (Lu ve Bolt, 2015). UTS'nin geniş ölçekli uygulamalarda da yaşanan bu duruma neden olan faktörlerden biri olduğu ifade edilebilir. Öte yandan, UTS ile matematik başarısı arasındaki ilişkinin düşük düzeyde bulunması ve ÇBMTK modellerinden elde edilen puanlara ilişkin olarak belirlenen ülke içi ve ülkeler arası korelasyon değerlerinin birbirine yakın olması, üç boyutlu modelden farklı olarak dört boyutlu modelde UTS’nin kontrol edilmesinin tutum-başarı paradoksunu tek başına açıklamada yeterli olmadığını göstermektedir. Bu duruma, bireylerin cevaplarında etkili olan ve korelasyonu etkileyen başka tepki stilleri (KTS, RETS, vb.) de neden olmuş olabilir (Reynolds ve Smith, 2010). Matematik başarısı ile matematiğe yönelik tutum arasındaki

ilişkiyi ve miktarını etkileyen bir diğer faktör ise büyük balık küçük göl etkisi olabilir (Marsh ve diğerleri, 2008). Çünkü öğrencilerin kendi beyanlarına (self-report) dayanan bu çalışmada, öğrenciler içinde bulunduğu öğrenci grubuna göre kendilerini değerlendirmiş ve tutumlarını olduğundan daha farklı olarak ifade etmiş olabilirler. Bu durumda matematik başarılarıyla olan ilişkilerinin yönü ve düzeyi etkilenmiş olabilir.

3.1.6 UTS'nin Etkisine göre Düzeltilmemiş ve Düzeltilmiş Ortalama Tutum Puanlarının Matematik Başarıyla Ülke İçi ve Ülkeler Arası İlişkilerin Arasındaki Farklara İlişkin Bulgular

TIMSS 2015 uygulamasına katılan ve araştırmaya dahil edilen ülkelerdeki öğrencilerin UTS'nin etkisine göre düzeltilmemiş ve düzeltilmiş ortalama tutum puanlarının matematik başarıyla olan ülke içi ilişkilerin arasında manidar bir fark olup olmadığını belirlemek için yapılan Fisher'in r_z dönüşümü sonuçları Çizelge 9'da sunulmuştur:

Çizelge 9.

Ülkelerin Düzeltilmemiş ve Düzeltilmiş Ortalama Tutum Puanlarının Matematik Başarıyla Olan Ülke İçi İlişkilerin Arasındaki Farklar

Ülkeler	Değişken	z	p
Singapur	Sevme	-.87	.38
	Güven	-2.17	.03*
	Değer	-1.78	.07
Kore Cumhuriyeti	Sevme	-1.17	.24
	Güven	-3.03	.00*
	Değer	-.57	.57
Çin-Tayvan	Sevme	-1.01	.31
	Güven	-2.82	.00*
	Değer	-1.11	.27
Hong Kong	Sevme	.00	1.00
	Güven	-1.32	.19
	Değer	-.99	.32
Japonya	Sevme	-.73	.46
	Güven	-2.08	.04*
	Değer	-1.01	.31
İsveç	Sevme	-.78	.43
	Güven	-1.38	.17
	Değer	-1.5	.13
İtalya	Sevme	-.37	.71
	Güven	-1.12	.26
	Değer	-.99	.32

(Devam ediyor)

Çizelge 9. (devam)

Ülkelerin Düzeltilmemiş ve Düzeltilmiş Ortalama Tutum Puanlarının Matematik Başarıyla Olan Ülke İçi İlişkilerin Arasındaki Farklar

Ülkeler	Değişken	z	p
Malta	Sevme	-.52	.60
	Güven	-1.79	.07
	Değer	-1.78	.07
Avustralya	Sevme	-1.05	.29
	Güven	-1.79	.07
	Değer	-1.78	.07
Norveç	Sevme	-1.05	.29
	Güven	-2.27	.02*
	Değer	-1.33	.18
Türkiye	Sevme	-.93	.35
	Güven	-1.32	.19
	Değer	-1.34	.18
Şili	Sevme	-.66	.51
	Güven	-1.6	.11
	Değer	-.64	.52
Kuveyt	Sevme	-.50	.61
	Güven	-2.04	.04*
	Değer	-1.5	.13
Mısır	Sevme	-1.55	.12
	Güven	-2.54	.01*
	Değer	-2.66	.00*
Sudi Arabistan	Sevme	.51	.61
	Güven	-3.03	.00*
	Değer	-2.45	.01*

*.05 düzeyinde manidar

Çizelge 9'dan görüldüğü üzere TIMSS 2015 uygulamasına katılan ülkelerdeki öğrencilerin genel olarak matematikte kendine güven boyutuna ilişkin düzeltilmemiş ve düzeltilmiş tutum puanlarının matematik başarıyla ilişkileri arasında manidar farklılık vardır ($p < .05$). Öte yandan, öğrencilerin matematiği öğrenmeyi sevmesi ve matematiğe değer vermesi boyutlarına ilişkin düzeltilmemiş ve düzeltilmiş tutum puanların matematik başarılarıyla ilişkileri arasında çoğunlukla manidar farklılık bulunmamıştır. Bu durum ülkelerin düzeltilmemiş ve düzeltilmiş tutum puanların matematik başarıyla ilişkilerini gösteren korelasyon değerlerinin birbirine benzer olmasıyla açıklanabilir.

Araştırmaya dahil edilen ülkelerin UTS'nin etkisine göre düzeltilmemiş ve düzeltilmiş ortalama tutum puanlarının matematik başarıyla olan ülkeler arası ilişkilerin arasında manidar bir fark olup olmadığını belirlemek için yapılan Fisher'in r_z dönüşümü sonuçları Çizelge 10'da sunulmuştur:

Çizelge 10.

Ülkelerin Düzeltilmemiş ve Düzeltilmiş Tutum Puanlarının Matematik Başarıyla Olan İlişkilerine Ait Ülkeler Arası Korelasyon Değerleri Arasındaki Farklar

Korelasyon Katsayısı Çiftleri	Değişken	z	p
a-b	Guven	-3.40	.01*
	Sevme	-2.63	.01*
	Değer	-2.75	.01*
a-c	Guven	-3.43	.01*
	Sevme	-2.64	.01*
	Değer	-2.81	.00*
b-c	Guven	-.61	.54
	Sevme	-.06	.95
	Değer	-1.16	.24

*.05 düzeyinde manidar, a = Gözlenen Puanlara Dayalı Korelasyon Değeri, b = Üç Boyutlu Modelle Kestirilen Puanlara İlişkin Korelasyon Değeri, c = Dört Boyutlu Modelle Kestirilen Puanlara İlişkin Korelasyon Değeri

Çizelge 10'dan görüldüğü üzere TIMSS 2015 uygulamasına katılan ülkelerdeki öğrencilerin matematiği öğrenmeyi sevmesi, matematikte güvenmesi ve matematiğe değer vermesi ile matematik başarıları arasındaki ülke düzeyindeki ilişkileri gösteren gözlenen puanlara dayalı korelasyon değerleri ile ÇBMTK modelleriyle kestirilen korelasyon değerleri arasında istatistiksel olarak manidar fark bulunmuştur ($p < .05$). Öte yandan, üç boyutlu MTK modeliyle kestirilen puanlara dayalı olarak belirlenen her bir değişkenin matematik başarıyla ilişkisini gösteren korelasyon değerleri ile dört boyutlu MTK modeliyle kestirilen puanlara ilişkin korelasyon değeri arasında manidar fark bulunmamıştır. Benzer bulguya Lu (2012) tarafından yapılan araştırmada da ulaşılmıştır. Buna göre PISA 2006 uygulamasına katılan öğrencilerin fene yönelik tutumları ile fen başarıları arasındaki ilişki, çeşitli ÇBMTK modelleriyle incelendiği ve UTS'nin etkisinin göz önüne alındığı ve alınmadığı durumlarda farklılaşmıştır.

ÇBMTK modelleriyle elde edilen puanlara dayalı olarak belirlenen ülke içi ve ülkeler arası korelasyon değerlerinin kendi içinde birbirine benzer çıkmasında UTS'nin kontrol edilmesinin, matematik başarıyla ile matematiğe yönelik tutum arasındaki ülke düzeyindeki ilişkileri açıklamada tek başına yeterli olmaması rol oynamış olabilir.

Bunun yanı sıra düzeltilmemiş tutum puanlarının kestirildiği üç boyutlu ve düzeltilmiş tutum puanlarının kestirildiği dört boyutlu modellerin aynı kuram olan MTK'ya dayanması, arka planda benzer algoritmaların yer alması, aynı örtük özellik metriğinde olması etkili olmuş olabilir. Öte yandan, gözlenen puanlara dayalı korelasyon değeri ile ÇBMTK modelleriyle elde edilen puanlara ilişkin ülkeler arası korelasyon değerleri arasında manidar farkın çıkmasında da farklı boyutlara ilişkin örtük özelliklerin eş zamanlı olarak kestirilmesini sağlayan ÇBMTK modellerinin aksine gözlenen puanlara dayalı korelasyon değerinin temelinde Klasik Test Kuramı'nın (KTK) yer alması ve bu kuramların farklı varsayımlara dayanması rol oynamış olabilir. Benzer durum Lu (2012) tarafından yapılan araştırmada da gözlenmiştir. Araştırmacı, ülke düzeyinde fene yönelik tutum ile fen başarısı arasındaki ilişkileri temelinde KTK'ya dayanan ad-hoc UTS kestirimleriyle ve ÇBMTK kuramı modelleriyle elde edilen korelasyon değerleriyle karşılaştırmış ve iki yaklaşıma göre hesaplanan korelasyon değerlerinin farklı olduğunu bulmuştur. Bunların yanı sıra, daha öncesinde Şekil 14, Şekil 15 ve Şekil 16'da da belirtildiği üzere dört boyutlu modelde öğrencilerin tutum puanları UTS'nin etkisinden arındırılmış ve bu puanlar öğrencilerin UTS'nin etkisine göre düzeltilmemiş gözlenen puanlara dayalı tutum puanlarından farklı kestirilmiştir. Bu durum öğrencilerin matematik başarısı ile matematiğe yönelik tutumları arasındaki ilişkiyi etkilemiştir. Bu yüzden, öğrencilerin matematik başarıları ile matematiğe yönelik tutumları arasındaki ilişkiyi belirlemek için gözlenen puanlara dayalı olarak hesaplanan korelasyon değerinden farklı değerlerin elde edilmesine neden olmuştur.

Öğrencilerin kendini beyanlarına dayalı likert tipi ölçeklerin kullanıldığı bu araştırmadan elde edilen bulgular bütün olarak değerlendirildiğinde, farklı ülkelerdeki öğrencilerin UTS gösterme eğilimlerinin ve düzeylerinin değiştiği sonucuna ulaşılmıştır. Bunun yanı sıra UTS'nin öğrencilerin tutumlarının üzerindeki etkisi öğrenciden öğrenciye, ülkeden ülkeye göre farklılaşmaktadır. Bu durum, öğrencilerin tutumlarının başarılarıyla olan ilişkilerinin yönünü ve düzeyini de etkilemektedir.

BÖLÜM 4

4.1 Sonuç ve Öneriler

Bu bölümde, araştırmadan elde edilen bulgulara dayalı sonuçlar ve öneriler sunulmuştur.

4.1.1 Sonuçlar

1. TIMSS 2015 uygulamasına katılan ülkelerden çalışmaya dahil edilen her ülkenin matematiğe yönelik tutum ölçeklerindeki maddelere verdikleri tepkilerde UTS gösterme eğilimleri birbirine göre değişmektedir.

2. TIMSS 2015 uygulamasına katılan öğrencilerin matematiğe yönelik tutum maddelerine verdikleri tepkilerde UTS'nin etkisi söz konusudur.

3. Farklı miktarlarda UTS'nin sergilenmesinin öğrencilerin matematiğe yönelik tutum ölçeklerinden elde edilen toplam puanlarının üzerindeki yanlılığa etkisi, öğrencilerin matematiğe yönelik tutum düzeylerine göre değişmektedir. Bir başka anlatımla, tutum maddelerine verilen tepkilerde yüksek düzeyde UTS'nin sergilenmesi; matematiğe yönelik tutumu olumlu olan bireylerde tutumlarını ölçmede kullanılan ölçeklerdeki toplam puanını artırıcı, matematiğe yönelik tutumu olumsuz olan bireylerin ölçeklerdeki toplam puanı azaltıcı yönde hareket etmektedir.

4. UTS'nin öğrencilerin matematiğe yönelik tutumları üzerindeki etkisinin olumlu ya da olumsuz yönde olması ülkeden ülkeye göre değişmektedir. UTS'nin sergilenmesi matematik başarısı yüksek ve orta düzeyde olan ülkelerde, öğrencilerin matematiğe yönelik tutumlarının olduğundan daha olumsuz olarak kestirilmesine yol açmaktadır. Öte yandan, matematik başarısı düşük olan ülkelerde ise matematiğe yönelik tutumlarının daha olumluymuş gibi belirlenmesine neden olmaktadır.

5. Ülkelerin tutumla ilgili farklı ölçeklerden elde edilen düzeltilmemiş ve düzeltilmiş tutum puanları arasındaki farklılık tutumla ilgili farklı boyutlara göre değişmektedir. En fazla matematiğe değer verme boyutunda, çoğu ülkenin düzeltilmemiş ve düzeltilmiş tutum puanlarının birbirinden farklılaştığı sonucuna ulaşılmıştır.

6. UTS, öğrencilerin matematiğe yönelik tutumla ilgili örtük özellik düzeylerini etkilemesinin yanı sıra tutumla ilgili değişkenlerin matematik başarılarıyla ilişkilerini de etkilemektedir. Bir başka anlatımla, araştırmaya katılan öğrencilerin tepkileri üzerinde UTS'nin etkisi dikkate alınmadığında öğrencilerin matematik başarıları ile matematiğe yönelik tutumları arasında ülke düzeyindeki ilişkiler negatif, UTS'nin etkisi dikkate alındığında ise bu ilişkiler pozitifdir.

7. Öğrencinin matematiği öğrenmeyi sevmesi, matematikte kendine güvenmesi ve matematiğe değer vermesi ile matematik başarısı arasındaki ilişkiyi belirlemek için gözlenen puanlara dayal korelasyon değerleri ile ÇBMTK modelleriyle elde edilen puanlara ilişkin korelasyon değerleri birbirinden farklılaşmaktadır.

8. UTS'nin etkisinin kontrol edilmediği ve edildiği ÇBMTK modellerinden elde edilen puanlara ilişkin korelasyon değerleri birbirinden farklılaşmamaktadır.

4.1.2 Öneriler

Araştırmanın bulguları doğrultusunda benzer konularda çalışma yapacak araştırmacılara yönelik öneriler aşağıda sunulmuştur

1. Araştırmada sadece uç tepki stilinin etkisi modellenmiştir. Kabullenici tepki stili, red edici tepki stili gibi diğer tepki stillerinin etkisi de araştırmada kurulan modellere dahil edilebilir. Bunun yanı sıra araştırmada kullanılan ölçekler dörtlü likert tipinde olmasından dolayı orta tepki kategorisi olmadığı için orta tepki kategorisinin seçilmesin tutum puanları üzerindeki etkisinin incelenmemiştir. Tepki kategorisinin sayısı artırılarak orta nokta tepki stilinin de etkisi incelenebilir.

2. Araştırmada ikiden fazla tepki kategorisine sahip likert tipi ölçeklerin kullanılması ve elde edilen bilgilerin bireylerin kendi hakkında bilgi vermesine dayalı olması öğrencilerin UTS sergilemelerinde etkili olabilir. Bu nedenle benzer amaçları taşıyan ilerideki araştırmalarda ikili puanlanan maddeler kullanılabilir. Bunun yanı sıra, başka psikolojik özelliğin ölçülmesinin amaçlandığı, bireylerin kendi hakkında bilgi vermesine dayalı ölçeklerin kullanıldığı başka araştırmalarda UTS'nin etkisi incelenebilir.

3. Araştırmada UTS'nin modellenmesinde örtük özellik modellerinden avantajlarından dolayı ÇBMTK kullanılmıştır. UTS'nin modellenmesinde örtük özelliklerin ölçülmesine dayalı başka modellerden yararlanılabilir ve modeller arası tutarlılık araştırılabilir.

4. Araştırmada UTS'nin etkisinin dikkate alınması, ülke düzeyindeki negatif ilişkilerin pozitif olmasını sağlamasının yanı sıra bu ilişkilerin düşük düzeyde olması öğrencilerin matematik başarıları ile tutumları arasındaki ilişkileri etkileyen başka faktörlerin de olabileceğini göstermektedir. Bu yüzden, ilerideki araştırmalarda bu faktörlerde modellere dahil edilerek analizler yapılabilir.

5. Araştırmada, ülkelerin UTS sergilemedeki farklılıklarının olası nedenlerinden birinin kültürel farklılar olabileceğinin belirtilmesi nedeniyle farklı ülkelerin, ırkların veya etnik kökenli bireylerin katıldığı, bireylerin kendi hakkında bilgi verdiği, likert tipi ölçeklerin kullanıldığı araştırmalarda, araştırmacılara bireylerin kültürel kökenlerinin de sorulduğu maddeleri eklemesi önerilmektedir.

6. Çalışmada TIMSS uygulamasına katılan öğrencilerin cevaplarında UTS'nin etkisinin olmasından dolayı TIMSS'i gerçekleştiren kuruluş olan IEA'ya bu yanlılığa dikkat etmeleri ve buna göre sonuçları raporlaştırmaları önerilmektedir.

7. Araştırmada, parametre kestirimlerindeki bütün belirsizliklerin aynı anda dikkate alınmasını, standart hataların daha kolay elde edilmesini sağlaması açısından Bayesçi yaklaşım benimsenerek MCMC simülasyonu gerçekleştirilmiştir. Fakat, Bayesçi yaklaşımlar belirli önseller gerektirmektedir. Örneğin bu araştırmada önsel olarak birim matris (identity matrix) kullanılmıştır. Fakat kullanılan önsellerin araştırma sonuçlarını etkileyebilmesi nedeniyle ileride aynı amaçla yapılacak çalışmalarda önsellerin kullanılmasını gerektirmeyen Marjinal Maksimum Olabilirlik yöntemi gibi alternatif yöntemler kullanılabilir.

8. Araştırmada kullanılan örneklem büyüklüğünün fazla olması ve tutumla ilgili değişkenlerin bir arada ele alınıp UTS ile eş zamanlı modellenmesi bir yandan kestirimin daha güvenilir olmasını sağlarken bir yandan da analizlerin çok uzun sürmesine neden olmaktadır. Kullanılan bilgisayarın kapasitesine de bağlı olarak her bir model için ortalama 120'şer saat süren analizler yapılmıştır. Bu yüzden tek seferde ayrı ayrı iki boyutlu model (tutumla ilgili bir değişken ve UTS'nin ayrı bir boyut olarak dahil edildiği) analiz edilebilir. Aynı nedenlerden dolayı, araştırmada öğrencilerin matematik başarılarını temsil etmek amacıyla sunulan beş farklı olası değerlerin ortalaması kullanılmıştır. İlerideki araştırmalarda öğrencilerin matematik başarılarını temsil etmek amacıyla beş farklı makul değer de modele dahil edilerek analizler gerçekleştirilebilir. Bu sınırlılıklar göz önünde bulundurularak temsiliyeti daha iyi sağlayabilmek adına ülke sayısı artırılabilir.

KAYNAKLAR

- Arce-Ferrer, A. J. (2006). An investigation into the factors influencing extreme-response style: Improving meaning of translated and culturally adapted rating scales. *Educational and Psychological Measurement*, 66, 374-392.
- Arikan, S., van de Vijver, F. J. R. ve Yağmur, K. (2016). Factors contributing to mathematics achievement differences of Turkish and Australian students in TIMSS 2007 and 2011. *EURASIA Journal of Mathematics, Science and Technology Education*, 12(8), 2039-2059.
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review*, 64,359-373.
- Austin, E. J., Deary, I. J., Gibson, G. J., McGregor, M. J., Dent, J. B. (1998). Individual response spread in self-report scales: personality correlations and consequences. *Personality and Individual Differences*, 24, 421-438.
- Bachman, J. G., O'Malley, P. M., & Freedman-Doan, P. (2010). *Response styles revisited:Racial/ethnic and gender differences in extreme responding* (Monitoring the Future Occasional Paper No. 72). Ann Arbor, MI: Institute for Social Research. Available:[TS](#)
- Bandura, A. (1994). Self-efficacy. *Encyclopedia of Human Behavior*, 4, 71–81.
- Baumgartner, H. and Steenkamp, J. E.M. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, 38, 143-156.
- Belbase, S. (2013). Images, anxieties, and attitudes toward mathematics. *International Journal of Education in Mathematics, Science and Technology*, 1(4), 230-237
- Billiet, J. B., and McClendon, M. J. (2000). Modeling acquiescence in measurement models for two balanced sets of items. *Structural Equation Modeling*, 7, 608-628.
- Bishop, A., FitzSimons, G., Seah, W. T. & Clarkson, P. (1999). *Values in mathematics education: making values teaching explicit in the mathematics classroom*. Paper Presented at the Combined Annual Meeting of the Australian Association for Research in Education and the New Zealand Association for Research in Education. Melbourne, Australia, November 29, December 2.
- Bofah, E. A. ve Hannula, M. S. (2015). TIMSS data in an African comparative perspective: Investigating the factors influencing achievement in mathematics and their psychometric properties. *Large-Scale Assessments in Education*, 3(1). doi:1.1186/s40536-015-0014-y
- Bolt, D. M., and Newton, J. (2011). Multiscale measurement of extreme response style. *Educational and Psychological Measurement*, 71, 814-833.
- Bolt, D. M., and Johnson, T. R. (2009). Addressing score bias and differential item

functioning due to individual differences in response style. *Applied Psychological Measurement*, 33,335-352.

- Bos, K. T. (2002). *Benefits and limitations of large-scale international comparative achievement studies: The case of IEA's TIMSS study*. (Unpublished doctoral thesis). University of Twente.
- Buckley, J. (2009). *Cross-national response styles in international educational assessment: Evidence from PISA 2006*. NCES Conference on the Program for International Student Assessment: What we can learn from. Retrieved from <https://edsurvey.rti.org/PISA/>
- Büyüköztürk, Ş. (2005). *Sosyal Bilimler için veri analizi el kitabı*. 5. baskı. Pegem A Yayıncılık.
- Büyüköztürk, Ş., Çakan, M., Tan, Ş. ve Atar, H. Y. (2014). *TIMSS 2011 ulusal matematik ve fen raporu 8. sınıflar*. Ankara: İşkur Matbaacılık.
- Büyüköztürk, Ş., Çokluk Bökeoğlu, Ö. ve Köklü, N. (2012). *Sosyal bilimler için istatistik*. Ankara: Pegem Akademi.
- Bybee, R., and McCrae, B. (2007). Scientific literacy and student attitudes: Perspectives from PISA 2006 science. *International Journal of Science Education*, 33, 7-26.
- Cheung, G. W., and Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross cultural research using structural equation modeling. *Journal of Cross- Cultural Psychology*, 31, 187-212.
- Cho, Y. (2013). *The mixture distribution polytomous rasch model used to account for response styles on rating scales: A simulation study of parameter recovery and classification accuracy* (Unpublished Doctoral Dissertation). University of Maryland, College Park.
- Clarke III, I. (2000). Extreme response style in cross-cultural research: An empirical investigation. *Journal of Social Behavior and Personality*, 15, 137-152.
- Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement*, 6, 475-494.
- Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2012). *Sosyal bilimler için çok değişkenli SPSS ve LISREL uygulamaları*. Ankara: PegemA Yayıncılık.
- Demir, İ. ve Kılıç, S. (2010). Öğrencilerin matematik başarısına etkileyen faktörlerin PISA 2003 kullanılarak incelenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 38, 44-54.
- De Jong M. G., Steenkamp J.-B. E. M., Fox J.-P., Baumgartner H. (2008). Using item response theory to measure extreme response style in marketing research: A global investigation. *Journal of Marketing Research*, 45(1), 104-115.

- de Lourdes Mata, M., Monteiro, V., and Peixoto, F. (2012). Attitudes towards mathematics: Effects of individual, motivational, and social support factors. *Child Development Research*. doi:1.1155/2012/876028).
- Dolnicar, S. & Grün, B. (2007) Cross-cultural differences in survey response patterns. *International Marketing Review*, 24(2), 127-143.
- Dowling, N. M., Bolt, D.M., Deng, S., & Li, C. (2016). Measurement and control of bias in patient reported outcomes using multidimensional item response theory. *BMC Medical Research Methodology*, 16(1), 1.
- Eid, M., & Rauber, M. (2000). Detecting measurement invariance in organizational surveys. *European Journal of Psychological Assessment*, 16, 20-3.
- Ekici, O. (2005). Bayesyen regresyon ve WinBUGS ile bir uygulama. Yayınlanmamış Yüksek Lisans Tezi, İstanbul Üniversitesi Sosyal Bilimler Enstitüsü, İstanbul.
- Engle, P. J. (2016). *Response Style in the Political Survey*. (Unpublished doctoral dissertation). University of Wisconsin Madison.
- Farooq, M. S. and Shah, S. Z. U. (2008). Students' attitude toward Mathematics. *Pakistan Economic and Social Review*, 48(1), 75-83.
- Fox, J. P. (2010). *Bayesian item response modeling: Theory and applications*. New York: Springer.
- Fraenkel, J.R., and Wallen, N.E. (2006). *How to design and evaluate research in education*. New York: McGraw-Hill.
- Gelman, A., and Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.
- Geyer, C. J. (1992). *On the convergence of Monte Carlo maximum likelihood calculations*. Technical Report 571, School of Statistics, Univ. Minnesota.
- Gilks, W.; Richardson, S.; and Spiegelhalter, D., eds. (1996). *Markov chain Monte Carlo in practice*. Chapman and Hall.
- Gill, J. (2002). *Bayesian methods: A social and behavioral sciences approach*. Boca Raton, FL: Chapman & Hall/CRC.
- Greenleaf, E. A. (1992a). Improving rating scale measures by detecting and correcting bias components in some response styles. *Journal of Marketing Research*, 29, 176- 188.
- Greenleaf, E. A. (1992b). Measuring extreme response style. *Public Opinion Quarterly*, 56, 328-351.
- Grootenboer, P., and Marshman, M. (2016). *Mathematics, affect and learning: middle school students' beliefs and attitudes about mathematics education*. Singapore: Springer.

- Hagenaars, J. A., and McCutcheon, A. L. (Eds.), (2002). *Applied Latent Class Analysis*. Cambridge: Cambridge University Press.
- Harter, S. (1999). *The construction of the self: A developmental perspective*. New York: Guilford Press.
- Harumi, C. A. (2011). *Cross-cultural differences in response styles* (Unpublished doctoral dissertation). Washington State University.
- Harzing, A. (2006). Response styles in cross-national survey research. *International Journal of Cross Cultural Management*, 6, 243-265.
- He, J. (2015). *The general response style from a cross-cultural perspective* (Unpublished doctoral dissertation). Tilburg University
- Hofstede, G. H. (2001). *Cultures consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Thousand Oaks, California: Sage Publications, Inc.
- Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations, software of the mind. Intercultural Cooperation and Its Importance for Survival*. 3rd ed. New York: McGraw Hill.
- Holland, P.W. and Wainer, H. (2012) *Differential Item Functioning*. Hillsdale: Routledge.
- Hooper, M, Mullis, I. V. S., and Martin, M.O. (2013). TIMSS 2015 Context Questionnaire Framework. Mullis, I.V.S. and Martin, M.O. (Eds.) TIMSS 2015 Assessment Frameworks. Retrieved from Boston College, TIMSS and PIRLS International Study Center website: <http://timssandpirls.bc.edu/timss2015/frameworks.html>
- Hui, C. H., & Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural psychology*, 20, 296-309.
- İlgün Dibek, M., Yavuz, H. & Çokluk Bökeoğlu, Ö. (2018). Tutum - başarı paradoksunda tepki stillerinin rolü: dokuz ülkenin karşılaştırılması, *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18(2), 932-952.
- Johnson, T.R., & Bolt, D. M. (2010). On the use of factor-analytic multinomial logit item response models to account for individual differences in response style. *Journal of Educational and Behavioral Statistics*, 35, 92-114.
- Johnson, T., Kulesa, P., Cho, Y. I., and Shavitt, S. (2005). The relation between culture and response styles: evidence from 19 countries. *Journal of Cross-Cultural Psychology*, 36, 264-277.
- Kadijevich, D.J. (2006). Developing trustworthy TIMSS background measures: A case study on mathematics attitude. *The Teaching of Mathematics*, 9 (2), 41-51.

- Kadijevich, D. (2008). TIMSS 2003: Relating dimensions of mathematics attitude to mathematics achievement. *Zbornik instituta za Pedagogical Research*, 40(2), 327–346. doi: 1.2298/ZIPI0802327K
- Khine, M. S., Al-Mutawah, M. ve Afari, E. (2015). Determinant of affective factors in mathematics achievement: Structural equation modeling approach. *Journal of Studies in Education*, 5(2), 199-211.
- Khorramdel L., von Davier M. (2014). Measuring response styles across the Big Five: A multiscale extension of an approach using multinomial processing trees. *Multivariate Behavioral Research*, 49, 161-177. doi:1.1080/00273171.2013.866536
- Kim, S. J., Park, J. H., Park, S. W. ve Kim, S. S. (2013). *The effects of school and students' educational contexts in Korea, Singapore, and Finland*. http://www.iea.nl/fileadmin/user_upload/IRC/IRC_2013/Papers/IRC-2013_Kim_etal.pdf adresinden erişildi.
- Kiwanuka, H. N., Van Damme, J., Van den Noortgate, W., Anumendem, D. N., and Namusisi, S. (2015). Factors affecting mathematics achievement of first-year secondary school students in Central Uganda. *South African Journal of Education*, 35(3), 1–16. doi:1.15700/saje.v35n3a1106
- LaRoche, S., and Foy, P. (2016). Sample Implementation in TIMSS 2015. In M. O. Martin, I. V. S. Mullis, and M. Hooper (Eds.), *Methods and Procedures in TIMSS 2015* (pp. 5.1-5.175). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timss.bc.edu/publications/timss/2015-methods/chapter-5.html>
- LaRoche, S., Joncas, M., and Foy, P. (2016). Sample Design in TIMSS 2015. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and Procedures in TIMSS 2015* (pp. 3.1-3.37). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timss.bc.edu/publications/timss/2015-methods/chapter-3.html>
- Lee, S. (2007). *Structural Equation Modeling: A Bayesian Approach*. John Wiley & Sons, London,UK,458.
- Little, T. D. (2000). On the comparability of constructs in cross-cultural research: A critique of Cheung and Rensvold. *Journal of Cross-Cultural Psychology*, 31, 213-219.
- Liu, M. (2015). *Response Style and Rating Scales: The Effects of Data Collection Mode, Scale Format, and Acculturation* (Unpublished doctoral dissertation). The University of Michigan.
- Lord, F. (1986). Maximum Likelihood and Bayesian Parameter Estimation in Item Response Theory. *Journal of Educational Measurement*, 23(2), 157-162. Retrieved from <http://www.jstor.org/stable/1434513>.

- Lu, Y. (2012). *A multilevel multidimensional item response theory model to address the role of response style on measurement of attitudes in PISA 2006*, (Dissertation). University of Wisconsin-Madison.
- Lu, Y. and Bolt, D.M. (2015). Examining the attitude-achievement paradox in PISA using a multilevel multidimensional IRT model for extreme response style. *Large-scale Assessments in Education*, 3(2), 1-18. doi:1.1186/s40536-015-0012-.
- Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., Hornik, K. (2018). cluster: Cluster Analysis Basics and Extensions. R package version 2.0.7-1.
- Marín, G., Gamba, R. J., & Marín, B.V. (1992). Extreme response style and acquiescence among Hispanics: The role of acculturation and education. *Journal of Cross-Cultural Psychology*, 23, 498-509.
- Marsh, H.W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K.T., O'Mara, A.J., and Craven, R.G. (2008). The big fish little pond effect stands up to critical scrutiny: Implications for theory, methodology, and future research. *Educational Psychology Review*, 20, 319–35.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O and Baumert, J. (2005). Academic self-concept, interest, grades and standardized test scores: Reciprocal effects models of causal ordering. *Child Development*, 76(2), 397-416.
- Martin, M. O. ve Mullis, I. V. S. (Ed.). (2012). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS and PIRLS International Study Center, Boston College.
- Martin, M. O., Mullis, I. V. S., Hooper, M., Yin, L., Foy, P., and Palazzo, L. (2016). Creating and Interpreting the TIMSS 2015 Context Questionnaire Scales. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and Procedures in TIMSS 2015* (pp. 15.1-15.312). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timss.bc.edu/publications/timss/2015-methods/chapter-15.html>
- MedCalc Software bvba (2018). MedCalc Statistical Software version 18.2.1, Ostend, Belgium. Retrieved from <https://www.medcalc.org>.
- Meisenberg, G. ve Williams, A. Are acquiescent and extreme response styles related to low intelligence and education?, *Personality and Individual Differences*, 44 (7), doi:10.1016/j.paid.2008.01.010.
- Mertler, C.A. and Vannatta, R.A. (2005) *Advanced and Multivariate Statistical Methods: Practical Application and Interpretation*. 3rd Edition, Pyrczak, Los Angeles.
- Messick S. (1962). Response style and content measures from personality inventories. *Educ. Psychol. Meas.* 22, 41–56.

- Moors G. (2003). Diagnosing response style behavior by means of a latent-class factor approach: Sociodemographic correlates of gender role attitudes and perceptions of ethnic discrimination reexamined. *Quality&Quantity*, 37, 277–302.
- Moors, G. (2004). Facts and artifacts in the comparison of attitudes among ethnic minorities. A multilevel latent class structure model with adjustment for response style behavior. *European Sociological Review*, 20, 303-32.
- Mullis, I. V. S., Cotter, K. E., Fishbein, B. G., and Centurino, V. A. S. (2016). Developing the TIMSS 2015 achievement items. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and Procedures in TIMSS 2015* (pp.1.1-1.22). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timss.bc.edu/publications/timss/2015-methods/chapter-1.html>
- Mullis, I. V. S., Martin, M. O., Foy, P. ve Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Mullis, I.V.S., Martin, M.O., and Loveless, T. (2016). 20 Years of TIMSS: International Trends in Mathematics and Science Achievement, Curriculum, and Instruction. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved from: <http://timss2015.org/timss2015/wp-content/uploads/2016/T15-20-years-of-TIMSS.pdf>
- Neale, D. (1969). The role of attitudes in learning mathematics. *The Arithmetic Teacher*, 16, 631-641.
- Ntzoufras, I. (2009). *Bayesian modeling using WinBUGS*. Wiley Series in Computational Statistics, Hoboken, USA.
- Österling, L. (2013). To Survey what Students Value in Mathematics Learning (Unpublished master's thesis). Sweden/Stockholm University.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver & L. S. Wrightman (Eds.), *Measures of Personality and Social Psychological Attitudes* (Vol. 1). San Diego, CA: Academic Press.
- Plomp, T. (1998). The potential of international comparative studies to monitor the quality of education. *Prospects*, 28(1), 45-59.
- Prediger, D. J. (1999). Basic structure of work-relevant abilities. *Journal of Counseling Psychology*, 46, 172-184.
- Qian. X. (2011). *A multi-level differential item functioning analysis of trends in international mathematics and science study: Potential sources of gender and minority difference among U.S. eighth graders' science achievement*. Unpublished doctoral dissertation, Faculty of the University of Delaware.

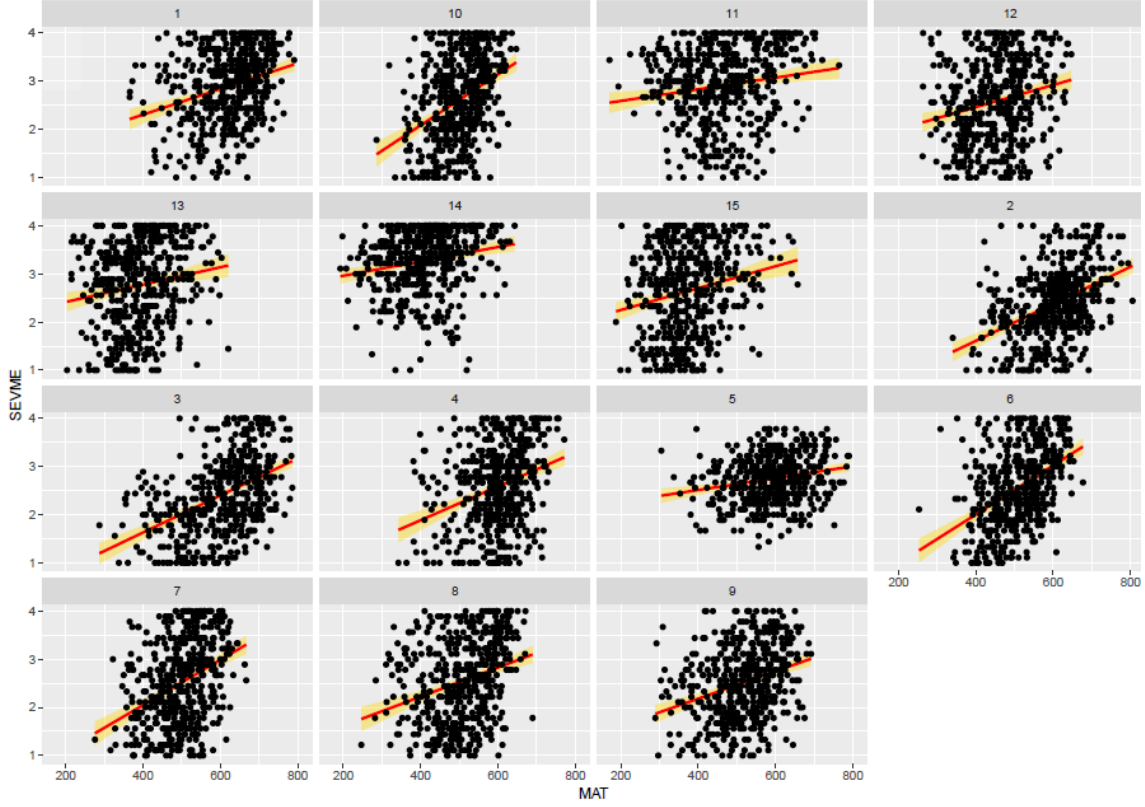
- Reckase, M. (2009). *Multidimensional item response theory*. Dordrecht: Springer.
- Reynolds, A. J. (1991). The middle schooling process: Influences on science and mathematics achievement from the longitudinal study of American youth. *Adolescence*, 26, 132-157.
- Reynolds, N., and Smith, A. (2010). Assessing the impact of response styles on cross-cultural service quality evaluation: A simplified approach to eliminating the problem. *Journal of Service Research*, 13, 230–243. doi: 1.1177/1094670509360408.
- Robinson, W. S. (1950). Ecological correlations and the behavior of individuals. *American Sociological Review*, 15, 351-357.
- Rodriguez, M. C. (2004). The Role of classroom assessment in student performance on TIMSS. *Applied Measurement in Education*, 17(1), 1-24. doi: 1.1207/s15324818ame1701_1
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (Complete samples). *Biometrika*, 52(3/4), 591-611.
- Shen, C. and Tam, H.P. (2008) The paradoxical relationship between student achievement and self-perception: a cross-national analysis based on three waves of TIMSS data, *Educational Research and Evaluation*, 14(1), 87-100, DOI: 1.1080/13803610801896653
- Sievert, C., Parmer, C., Hocking, T. Chamberlain, S., Ram, K., Corvellec, M. ve Despouy, P. (2017). plotly: Create Interactive Web Graphics via ‘plotly.js’. R package version 4.7.1. <https://CRAN.R-project.org/package=plotly>.
- Sinharay, S. (2004). Experiences with markov chain monte carlo convergence assessment in two psychometric examples. *Journal of Educational and Behavioral Statistics*, 29, 461– 488.
- Smith, P. B. (2004). Acquiescent response bias as an aspect of cultural communication style. *Journal of Cross-Cultural Psychology*, 35, 50-61.
- Spiegelhalter, D., Thomas, A., and Best, N. (2004). *WinBUGS version 1.4* [Computer program]. Cambridge, UK: MRC Biostatistics Unit, Institute of Public Health.
- Sundre, D., Barry, C., Gynnild, V., and Ostgard, E. T. (2012). Motivation for achievement and attitudes toward mathematics instruction in a required calculus course at the Norwegian University of Science and Technology. *Numeracy*, 5(1), 1–2. doi:1.5038/1936-466.5.1.4.
- van de Gaer, E. and Adams, R. (2010). *The Modeling of Response Style Bias: An Answer to the Attitude-Achievement Paradox?*, paper presented at the annual conference of the American Educational Research Association, Denver, Colorado, USA, 30 April – 4 May.

- Van de gaer, E., Grisay, A., Schulz, W. and Gebhardt, E. (2012). The reference group effect an explanation of the paradoxical relationship between academic achievement and self-confidence across countries. *Journal of Cross-Cultural Psychology*, 43(8), 1205-1228.
- van Herk, H., Poortinga, Y. H., and Verhallen, T. M. M. (2004). Response styles in rating scales: Evidence of method bias in data from six EU countries. *Journal of Cross-Cultural Psychology*, 35, 346-36.
- van Ravenzwaaij, D., Cassey, P., & Brown, S. D. (2016). A simple introduction to Markov Chain MonteCarlo sampling. *Psychonomic Bulletin & Review*. DOI: 1.3758/s13423-016-1015-8
- van Rosmalen J., van Herk H., Groenen P. J. F. (2010). Identifying response styles: A latent-class bilinear multinomial logit model. *Journal of Marketing Research*, 47, 157-172.
- Vermunt, J. K., and Magidson, J. (2008). *LG-Syntax User's Guide: Manual for Latent Gold 4.5 Syntax Module*. Belmont, MA: Statistical Innovations, Inc.
- Weijters, B. (2006). *Response Styles in Consumer Research* (Unpublished doctoral dissertation). Ghent University.
- Weijters, B., Geuens, M., Schillewaert, N.(2008). Assessing response styles across modes of data collection, *Journal of the Academy of Marketing Science* , 36, 409-422 doi: 1.1007/s11747-007-0077-6
- Welkenhuysen-Gybels, J., Billiet, J. B., & Cambre, B. (2006). Adjustment for acquiescence in the assessment of the construct equivalence of Likert-type score items. *Journal of Cross-Cultural psychology*, 34, 702-722.
- Wetzel, E., Böhnke, J., Carstensen, C.H., Ziegler, M.,and Ostendorf, F. (2013).Do individual response styles matter? assessing differential item functioning for men and women in the NEO-PI-R. *Journal of Individual Differences*, 34, 69–81.
- Wickham. H.(2007). Reshaping Data with the {reshape} Package. *Journal of Statistical Software*, 21(12),1-20.
- Wickham. H.(2009). *ggplot2: elegant graphics for data analysis*. Springer New York
- Wigfield, A. ve Eccles, J. S. (2000). Expectancy-value theory of motivation. *Contemporary Educational Psychology*, 25, 68-81.
- Zickar, M. J. and Drasgow, F. (1996). Detecting faking on a personality instrument using appropriateness measurement. *Applied Psychological Measurement*, 20, 71-87.

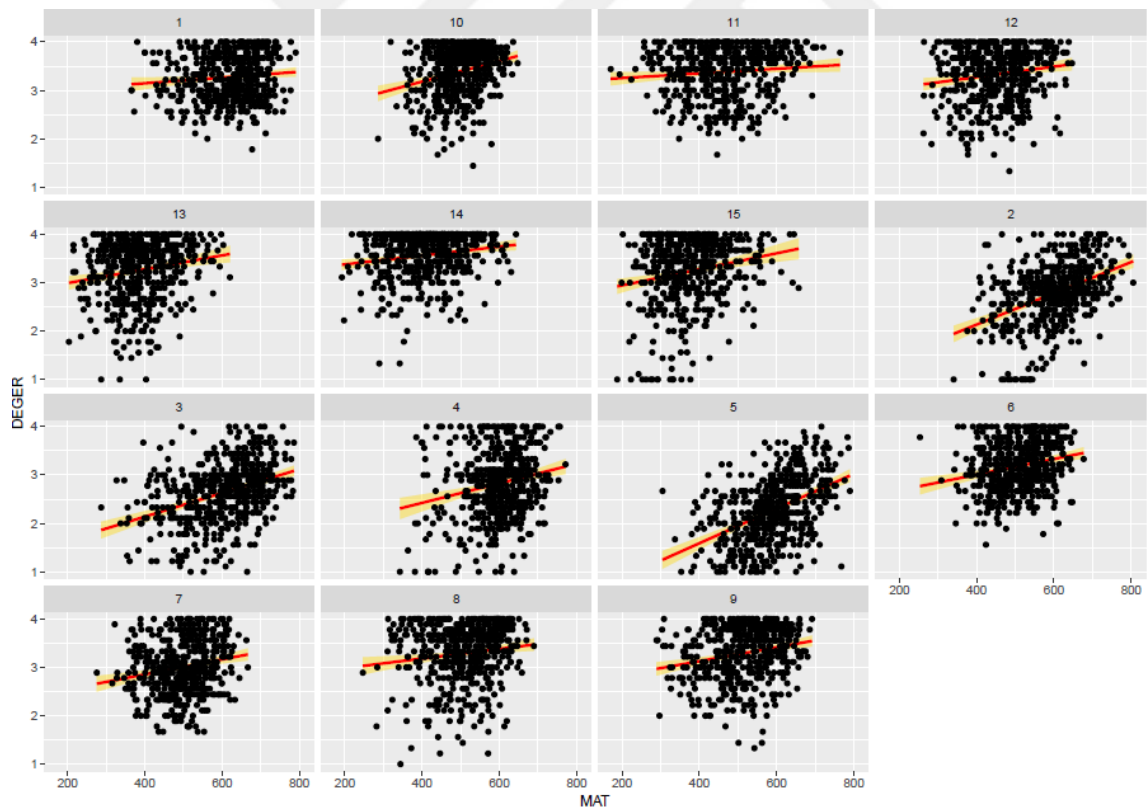
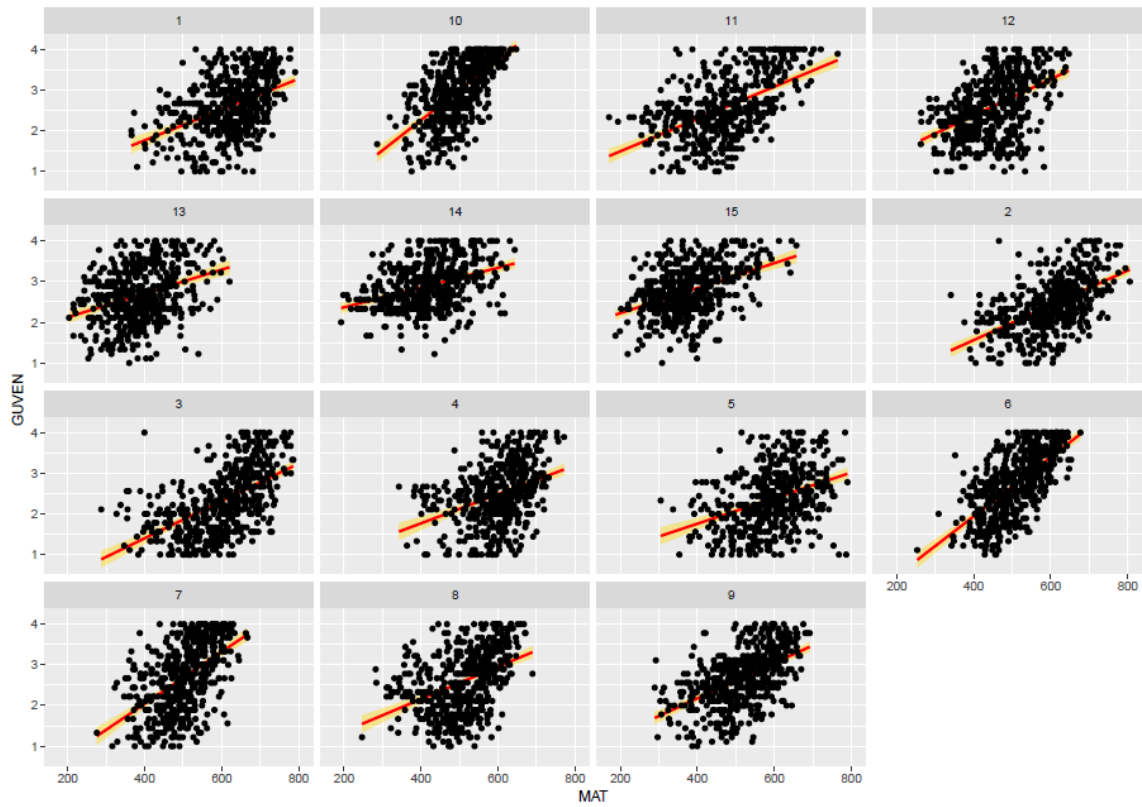


EKLER

EK A.Tutumla ilgili Boyutların Matematik Başarılarıyla İlişkilerine Ait Saçılma Grafikleri



Not. 1= Singapur, 2= Kore Cumhuriyeti ,3= Çin- Tayvan, 4= Hong-Kong, 5= Japonya ,
6= İsveç, 7 = İtalya, 8= Malta, 9 =Avustralya, 10 = Norveç, 11 = Türkiye, 12 =Şile, 13
= Kuveyt, 14 = Mısır, 15=Sudi Arabistan



EKB. Tutum Değişkenlerine Ait Maddeler

Maddeler	Tepki Kategorileri
Öğrencinin Matematik Öğrenmeyi Sevmesi İndeksi	
Matematik öğrenmeyi severim	
Keşke matematik çalışmam gerekmeseydi	
Matematik sıkıcıdır	
Matematikte pek çok ilginç şey öğrenirim	
Matematiği severim	Tamamen katılıyorum
İçinde sayı olan her türlü okul çalışmasını severim	Kısmen katılıyorum
Matematik problemlerini çözmekten hoşlanırım	Kısmen katılmıyorum
Matematik derslerini dört gözle bekliyorum	Hiç katılmıyorum
Matematik en sevdiğim derslerden biridir	
Öğrencinin Matematikte Kendine Güvenme İndeksi	
Matematikte genelde iyiyimdir	
Matematik birçok sınıf arkadaşşıma göre bana zor gelir	
Matematik başarılı olduğum alanlardan biri değildir	Tamamen katılıyorum
Matematik konularını hızlı öğrenirim	Kısmen katılıyorum
Matematik sinirlenmeme neden olur	Kısmen katılmıyorum
Zor matematik problemlerini çözmekte iyiyimdir	Hiç katılmıyorum
Öğretmenim matematikte iyi olduğumu söyler	
Matematik benim için diğer alanlardan daha zordur	
Matematik kafamın karışmasına neden olur	
Öğrencinin Matematiğe Değer Vermesi İndeksi	
Matematik öğrenmenin günlük yaşamıma yardımcı olacağını düşünüyorum	
Diğer okul derslerini öğrenmek için matematiğe ihtiyacım vardır	Tamamen katılıyorum
Seçtiğim üniversiteye gitmek için iyi bir matematiğe ihtiyacım var	Kısmen katılıyorum
İstedğim mesleği elde etmek için iyi bir matematiğe ihtiyacım var	Kısmen katılmıyorum
Matematik kullanacağım bir meslek isterim	Hiç katılmıyorum
Hayatta ilerlemek için matematiği öğrenmek önemlidir	
Matematiği öğrenmek yetişkin olduğumda daha fazla iş fırsatı sağlayacak	
Ailem matematikte iyi olmanın önemli olduğunu düşünür	
Matematiği iyi yapmak benim için önemlidir	

EKC. Tutum Ölçeklerine Ait Cronbach Alfa Güvenirlik Katsayısı Değerleri

Cronbach Alfa Güvenirlik Katsayıları			
Ülkeler	Öğrencinin Matematiği Öğrenmeyi Sevmesi Ölçeği	Öğrencinin Matematikte Kendine Güvenmesi Ölçeği	Öğrencinin Matematiğe Değer Vermesi Ölçeği
Singapur	.93	.91	.84
Kore	.95	.90	.90
Çin- Tayvan	.96	.94	.88
Hong Kong	.95	.89	.91
Japonya	.94	.91	.79
İsveç	.94	.92	.86
İtalya	.95	.93	.83
Malta	.93	.89	.87
Avustralya	.93	.89	.88
Norveç	.95	.92	.87
Türkiye	.91	.89	.81
Şili	.93	.88	.86
Kuveyt	.92	.81	.85
Mısır	.84	.70	.78
Sudi Arabistan	.92	.73	.90

EKD. Örneklem-Alt Örneklem için Değişkenler Arası Korelasyon Değerleri

Singapur

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.32**	.28**	.41**	.39**	.12**	.11*
	p			.00	.00	.00	.00	.00	.015
	N	5907	500	5907	500	5907	500	5907	500
SEVME	Korelasyon	.32**	.28**	1	1	.71**	.73**	.54**	.48**
	p	.00	.00			.00	.00	.00	.00
	N	5907	500	5907	500	5907	500	5907	500
GUVEN	Korelasyon	.41**	.39**	.71**	.73**	1	1	.37**	.34**
	p	.00	.00	.00	.00			.00	.00
	N	5907	500	5907	500	5907	500	5907	500
DEGER	Korelasyon	.12**	.11*	.54**	.48**	.37**	.34**	1	1
	p	.00	.01	.00	.00	.00	.00		
	N	5907	500	5907	500	5907	500	5907	500

** .01 düzeyinde manidar* .05 düzeyinde manidar

Kore

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.41**	.41**	.54**	.50**	.38**	.39**
	p			.00	.00	.00	.00	.00	.00
	N	5193	500	5193	500	5193	500	5193	500
SEVME	Korelasyon	.41**	.41**	1	1	.67**	.70**	.56**	.57**
	p	.00	.00			.00	.00	.00	.00
	N	5193	500	5193	500	5193	500	5193	500
GUVEN	Korelasyon	.54	.50**	.67**	.70**	1	1	.44	.41**
	p	.00	.00	.00	.00			.00	.00
	N	5193	500	5193	500	5193	500	5193	500
DEGER	Korelasyon	.38**	.39**	.56**	.57**	.44**	.41**	1	1
	p	.00	.00	.00	.00	.00	.00		
	N	5193	500	5193	500	5193	500	5193	500

** .01 düzeyinde manidar

Tayvan

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.43**	.44**	.53**	.51**	.36**	.35**
	p			.00	.00	.00	.00	.00	.00
	N	5466	500	5466	500	5466	500	5466	500
SEVME	Korelasyon	.43**	.44**	1	1	.73**	.76**	.61**	.62**
	p	.00	.00			.00	.00	.00	.00
	N	5466	500	5466	500	5466	500	5466	500
GUVEN	Korelasyon	.53**	.51**	.73**	.76**	1	1	.46**	.48**
	p	.00	.00	.00	.00			.00	.00
	N	5466	500	5466	500	5466	500	5466	500
DEGER	Korelasyon	.36**	.35**	.61**	.62**	.46**	.48**	1	1
	p	.00	.00	.00	.00	.00	.00		
	N	5466	500	5466	500	5466	500	5466	500

** .01 düzeyinde manidar

Hong Kong

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.34**	.32**	.37**	.37**	.24**	.19**
	p			.00	.00	.00	.00	.00	.00
	N	3952	500	3952	500	3952	500	3952	500
SEVME	Korelasyon	.34**	.32**	1	1	.70**	.71**	.57**	.56**
	p	.00	.00			.00	.00	.00	.00
	N	3952	500	3952	500	3952	500	3952	500
GUVEN	Korelasyon	.37**	.37**	.70**	.71**	1	1	.35**	.33**
	p	.00	.00	.00	.00			.00	.00
	N	3952	500	3952	500	3952	500	3952	500
DEGER	Korelasyon	.24**	.19**	.57**	.56**	.35**	.33**	1	1
	p	.00	.00	.00	.00	.00	.00		
	N	3952	500	3952	500	3952	500	3952	500

**.01 düzeyinde manidar

Japonya

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.37**	.35**	.46**	.44**	.25**	.22**
	p			.00	.00	.00	.00	.00	.00
	N	4583	500	4583	500	4583	500	4583	500
SEVME	Korelasyon	.37**	.35**	1	1	.72**	.70**	.43**	.35**
	p	.00	.00			.00	.00	.00	.00
	N	4583	500	4583	500	4583	500	4583	500
GUVEN	Korelasyon	.46**	.44**	.72**	.70**	1	1	.29**	.20**
	p	.00	.00	.00	.00			.00	.00
	N	4583	500	4583	500	4583	500	4583	500
DEGER	Korelasyon	.25	.22**	.43**	.35**	.29**	.20**	1	1
	p	.00	.00	.00	.00	.00	.00		
	N	4583	500	4583	500	4583	500	4583	500

**.01 düzeyinde manidar

İsveç

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.44**	.42**	.63**	.63**	.21**	.19**
	p			.00	.00	.00	.00	.00	.00
	N	3244	500	3244	500	3244	500	3244	500
SEVME	Korelasyon	.44**	.42**	1	1	.68**	.67**	.52**	.50**
	p	.00	.00			.00	.00	.00	.00
	N	3244	500	3244	500	3244	500	3244	500
GUVEN	Korelasyon	.63**	.63**	.68**	.67**	1	1	.35**	.38**
	p	.00	.00	.00	.00			.00	.00
	N	3244	500	3244	500	3244	500	3244	500
DEGER	Korelasyon	.21**	.19	.52**	.50**	.35**	.38**	1	1
	p	.00	.00	.00	.00	.00	.00		
	N	3244	500	3244	500	3244	500	3244	500

**.01 düzeyinde manidar

İtalya

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.37**	.36**	.53**	.52**	.22**	.19**
	p			.00	.00	.00	.00	.00	.00
	N	4115	500	4115	500	4115	500	4115	500
SEVME	Korelasyon	.37**	.36**	1	1	.78**	.79**	.59**	.57**
	p	.00	.00			.00	.00	.00	.00
	N	4115	500	4115	500	4115	500	4115	500
GUVEN	Korelasyon	.53**	.52**	.78**	.79**	1	1	.47**	.47**
	p	.00	.00	.00	.00			.00	.00
	N	4115	500	4115	500	4115	500	4115	500
DEGER	Korelasyon	.22**	.19**	.59**	.57**	.47**	.47**	1	1
	p	.00	.00	.00	.00	.00	.00		
	N	4115	500	4115	500	4115	500	4115	500

** .01 düzeyinde manidar

Malta

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.30**	.29**	.43**	.41**	.18**	.11*
	p			.00	.00	.00	.00	.00	.01
	N	3345	500	3345	500	3345	500	3345	500
SEVME	Korelasyon	.30**	.29**	1	1	.70**	.70**	.57**	.53**
	p	.00	.00			.00	.00	.00	.00
	N	3345	500	3345	500	3345	500	3345	500
GUVEN	Korelasyon	.43**	.41**	.70**	.70**	1	1	.39**	.38**
	p	.00	.00	.00	.00			.00	.00
	N	3345	500	3345	500	3345	500	3345	500
DEGER	Korelasyon	.18**	.11*	.57**	.53**	.39**	.38**	1	1
	p	.00	.01	.00	.00	.00	.00		
	N	3345	500	3345	500	3345	500	3345	500

** .01 düzeyinde manidar

* .05 düzeyinde manidar

Avustralya

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.35**	.29**	.53**	.50**	.24**	.19**
	p			.00	.00	.00	.00	.00	.00
	N	8635	500	8635	500	8635	500	8635	500
SEVME	Korelasyon	.35**	.29**	1	1	.66**	.66**	.57**	.52**
	p	.00	.00			.00	.00	.00	.00
	N	8635	500	8635	500	8635	500	8635	500
GUVEN	Korelasyon	.53**	.50**	.66**	.66**	1	1	.39**	.36**
	p	.00	.00	.00	.00			.00	.00
	N	8635	500	8635	500	8635	500	8635	500
DEGER	Korelasyon	.24**	.19**	.57**	.52**	.39**	.36**	1	1
	p	.00	.00	.00	.00	.00	.00		
	N	8635	500	8635	500	8635	500	8635	500

** .01 düzeyinde manidar

Norveç

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.38**	.37**	.61**	.61**	.19**	.20**
	p			.00	.00	.00	.00	.00	.00
	N	4011	500	4011	500	4011	500	4011	500
SEVME	Korelasyon	.38**	.37**	1	1	.64**	.61**	.51**	.51**
	p	.00	.00			.00	.00	.00	.00
	N	4011	500	4011	500	4011	500	4011	500
GUVEN	Korelasyon	.61**	.61**	.64**	.61**	1	1	.36**	.32**
	p	.00	.00	.00	.00			.00	.00
	N	4011	500	4011	500	4011	500	4011	500
DEGER	Korelasyon	.19**	.20**	.51**	.51**	.36**	.32**	1	1
	p	.00	.00	.00	.00	.00	.00		
	N	4011	500	4011	500	4011	500	4011	500

** .01 düzeyinde manidar

Türkiye

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.21**	.19**	.50**	.52**	.15**	.11*
	p			.00	.00	.00	.00	.00	.01
	N	5360	500	5360	500	5360	500	5360	500
SEVME	Korelasyon	.21**	.19**	1	1	.67**	.66**	.61**	.55**
	p	.00	.00			.00	.00	.00	.00
	N	5360	500	5360	500	5360	500	5360	500
GUVEN	Korelasyon	.50**	.52**	.67**	.66**	1	1	.44**	.40**
	p	.00	.00	.00	.00			.00	.00
	N	5360	500	5360	500	5360	500	5360	500
DEGER	Korelasyon	.15**	.11*	.61**	.55**	.44**	.40**	1	1
	p	.00	.01	.00	.00	.00	.00		
	N	5360	500	5360	500	5360	500	5360	500

** .01 düzeyinde manidar

Şile

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.26**	.19**	.46**	.45**	.14**	.13**
	p			.00	.00	.00	.00	.00	.00
	N	4159	500	4159	500	4159	500	4159	500
SEVME	Korelasyon	.26**	.19**	1	1	.65**	.63**	.53**	.45**
	p	.00	.00			.00	.00	.00	.00
	N	4159	500	4159	500	4159	500	4159	500
GUVEN	Korelasyon	.46**	.45**	.65**	.63**	1	1	.33**	.26**
	p	.00	.00	.00	.00			.00	.00
	N	4159	500	4159	500	4159	500	4159	500
DEGER	Korelasyon	.14**	.13**	.53**	.45**	.33**	.26**	1	1
	p	.00	.00	.00	.00	.00	.00		
	N	4159	500	4159	500	4159	500	4159	500

** .01 düzeyinde manidar

Kuveyt

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.20**	.20**	.34**	.33**	.16**	.18**
	p			.00	.00	.00	.00	.00	.00
	N	3488	500	3488	3488	3488	500	3488	500
SEVME	Korelasyon	.20**	.18**	1	1	.67**	.68**	.56**	.58**
	p	.00	.00	.00	.00	.00	.00	.00	.00
	N	3488	500	3488	500	3488	500	3488	500
GUVEN	Korelasyon	.34**	.33**	.67**	.68**	1	1	.44**	.46**
	p	.00	.00	.00	.00	.00	.00	.00	.00
	N	3488	500	3488	500	3488	500	3488	500
DEGER	Korelasyon	.16**	.18**	.56**	.58**	.44**	.46**	1	1
	p	.00	.00	.00	.00	.00	.00	.00	.00
	N	3488	500	3488	500	3488	500	3488	500

** .01 düzeyinde manidar

Mısır

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.22**	.24**	.39**	.37**	.14**	.14**
	p			.00	.00	.00	.00	.00	.00
	N	5771	500	5771	500	5771	500	5771	5771
SEVME	Korelasyon	.22**	.24**	1	1	.62**	.58**	.56**	.46**
	p	.00	.00	.00	.00	.00	.00	.00	.00
	N	5771	500	5771	500	5771	500	5771	500
GUVEN	Korelasyon	.39**	.37**	.62**	.58**	1	1	.39**	.30**
	p	.00	.00	.00	.00	.00	.00	.00	.00
	N	5771	500	5771	500	5771	500	5771	500
DEGER	Korelasyon	.14**	.14**	.56**	.46**	.39**	.30**	1	1
	p	.00	.00	.00	.00	.00	.00	.00	.00
	N	5771	500	5771	500	5771	500	5771	500

** .01 düzeyinde manidar

Sudi Arabistan

		MAT		SEVME		GUVEN		DEGER	
		Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem	Örneklem	Alt örneklem
MAT	Korelasyon	1	1	.32**	.28**	.41**	.39**	.12**	.11*
	p			.00	.00	.00	.00	.00	.01
	N	5907	500	5907	500	5907	500	5907	500
SEVME	Korelasyon	.32**	.28**	1	1	.71**	.73**	.54**	.48**
	p	.00	.00	.00	.00	.00	.00	.00	.00
	N	5907	500	5907	500	5907	500	5907	500
GUVEN	Korelasyon	.41**	.39**	.71**	.73**	1	1	.37**	.34**
	p	.00	.00	.00	.00	.00	.00	.00	.00
	N	5907	500	5907	500	5907	500	5907	500
DEGER	Korelasyon	.12**	.11*	.54**	.48**	.37**	.34**	1	1
	p	.00	.01	.00	.00	.00	.00	.00	.00
	N	5907	500	5907	500	5907	500	5907	500

** .01 düzeyinde manidar

EK E.1 Parametrik Olmayan MANOVA için R Kodu

```
library(npmv)

allcountriesUTS <- read.csv("C:/Users/can/Desktop/allcountriesUTS.csv")

nonpartest(SEVMEUTS | GUVENUTS | DEGERUTS ~ ulke, allcountriesUTS,
permreps = 1000)

ssnonpartest(SEVMEUTS | GUVENUTS | DEGERUTS ~ ulke, allcountriesUTS, test =
c(1, 0, 0, 0), alpha = .05, factors.and.variables = TRUE)
```

EKE.2 Parametrik Olmayan MANOVA Sonuçları

Test Statistic	df1	df2	p-value	Permutation Test p-value
ANOVA type test p-value			69.873	29.67 15864.54 0 0
McKeon approx. for the Lawley Hotelling Test			48.956	42.000 18282.470 0
Muller approx. for the Bartlett-Nanda-Pillai Test			44.266	42.024 22464.980 0
Wilks Lambda			46.583	42.000 22198.93 0

\$releffects

	sevmeuts	guvenuts	degeruts
1	0.47032	0.42701	0.47432
2	0.37078	0.34481	0.34987
3	0.40522	0.45039	0.30575
4	0.46950	0.44253	0.39167
5	0.37594	0.42334	0.30421
6	0.44570	0.48179	0.44148
7	0.51220	0.51755	0.42657
8	0.53532	0.51847	0.59447
9	0.44209	0.43013	0.51724
10	0.49076	0.54036	0.56815
11	0.58150	0.54496	0.61395
12	0.52571	0.53655	0.58154
13	0.58818	0.59184	0.61380
14	0.69406	0.66008	0.71078
15	0.59272	0.59020	0.60619

The Hypothesis of equality between factor levels 2 4 5 7 is rejected

....

...

...

The Hypothesis of equality between factor levels 1 2 is rejected

All appropriate subsets using factor levels have been checked using a closed multiple testing procedure, which controls the maximum overall type I error rate at $\alpha=0.05$

~Performing the Subset Algorithm based on Response Variables~

The Hypothesis of equality using response variables sevmeuts guvenuts degeruts is rejected

....

....

....

The Hypothesis of equality using response variables sevmeuts is rejected

All appropriate subsets using response variables have been checked using a multiple testing procedure, which controls the maximum overall type I error rate at $\alpha=0.05$



EK F Latent Gold Kodları

```

options
  maxthreads=all;
  algorithm
    tolerance=1e-008 emtolerance=0.01 emiterations=250 nriterations=50 ;
  startvalues
    seed=0 sets=16 tolerance=1e-005 iterations=50;
  bayes
    categorical=1 variances=1 latent=1 poisson=1;

  montecarlo
    seed=0 sets=0 replicates=500 tolerance=1e-008;
  quadrature nodes=10;
  missing excludeall;

output

  parameters=effect standarderrors probmeans=posterior profile frequencies
  bivariateresiduals classification classification=model;

variables

dependent BSBM17A nominal, BSBM17B nominal, BSBM17C nominal, BSBM17D
nominal, BSBM17E nominal, BSBM17F nominal, BSBM17G nominal, BSBM17H
nominal, BSBM17I nominal,

BSBM19A nominal, BSBM19B nominal, BSBM19C nominal, BSBM19D nominal,
BSBM19E nominal, BSBM19F nominal, BSBM19G nominal, BSBM19H nominal,
BSBM19I nominal,

BSBM20A nominal, BSBM20B nominal, BSBM20C nominal, BSBM20D nominal,
BSBM20E nominal, BSBM20F nominal, BSBM20G nominal, BSBM20H nominal,
BSBM20I nominal;

latent

  CFactor1 continuous,
  CFactor2 continuous,
  CFactor3 continuous,
  CFactor4 continuous,
  Cluster nominal 1;


equations

  CFactor1;
  CFactor2;
  CFactor3;
  CFactor4;
  Cluster <-1;
  CFactor1<-> CFactor2;

```

```
CFactor1<-> CFactor3;  
CFactor1<-> CFactor4;  
CFactor2<-> CFactor3;  
CFactor2<-> CFactor4;  
CFactor3<-> CFactor4;  
BSBM17A-BSBM17I<-1+ Cluster + (L1) CFactor1 +(L4) CFactor4;  
BSBM19A-BSBM19I<-1+ Cluster + (L2) CFactor2 +(L4) CFactor4;  
BSBM20A-BSBM20I<-1+ Cluster + (L3) CFactor3 +(L4) CFactor4;
```

```
L1[1,1]=-3;  
L1[1,2]=-1;  
L1[1,3]=1;  
L1[1,4]=3;  
L2[1,1]=-3;  
L2[1,2]=-1;  
L2[1,3]=1;  
L2[1,4]=3;  
L3[1,1]=-3;  
L3[1,2]=-1;  
L3[1,3]=1;  
L3[1,4]=3;  
L4[1,1]=1;  
L4[1,2]=-1;  
L4[1,3]=-1;  
L4[1,4]=1;
```



EKG WinBUGS Kodlari

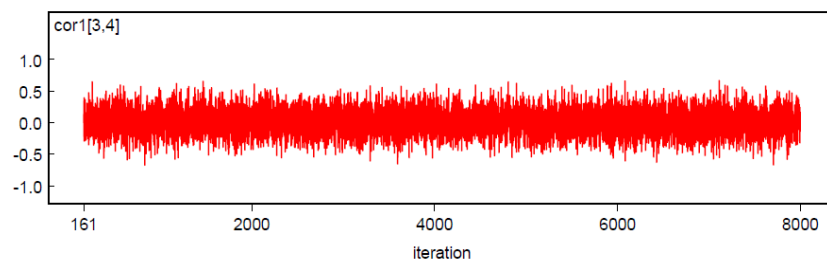
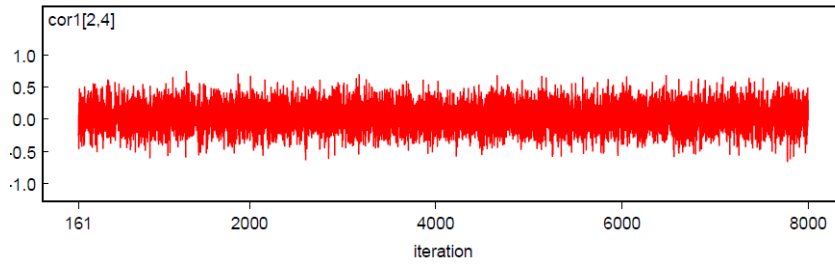
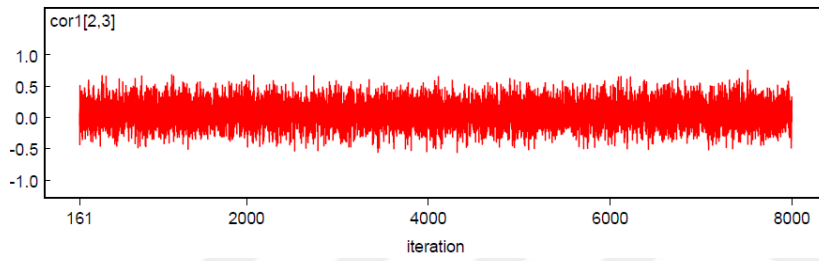
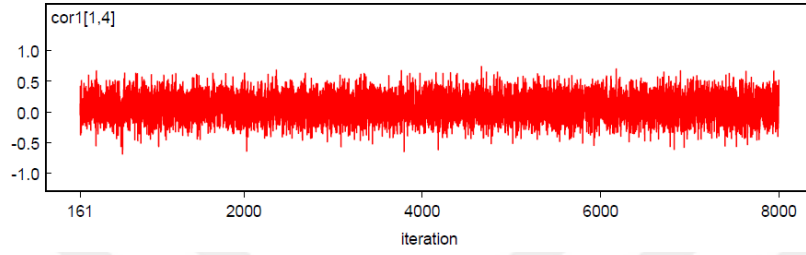
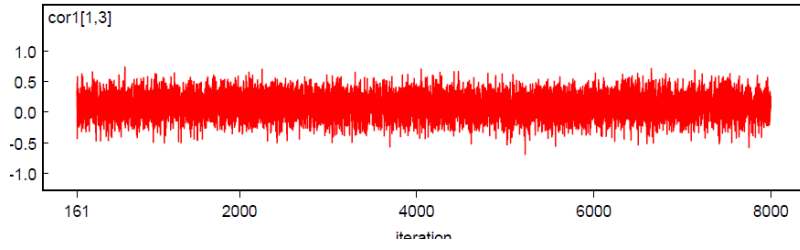
```

model
{
  ## read in individual response
  for (j in 1:N){
    for (i in 1:T+1){
      r[j,i]<-resp [j,i]
    }
  }
  ## nominal response mode
  for (j in 1 : N) {
    for (i in 1 : 9) {
      for (k in 1: 4) {
        prop [j, i, k] <- exp(a1[k]*theta[j,1] +a4[k]*theta[j,4]+ b[i,k])
        p[j,i,k]<-prop[j,i,k]/(sum(prop[j,i,]))
      }
      r[j,i]~dcat(p[j,i,])
    }
    .....
  }
  for (i in 19 : 27) {
    for (k in 1: 4) {
      prop [j, i, k] <- exp(a3[k]*theta[j,3] +a4[k]*theta[j,4]+ b[i,k])
      p[j,i,k]<-prop[j,i,k]/(sum(prop[j,i,]))
    }
    r[j,i]~dcat(p[j,i,])
  }
  theta[j,5]<-r[j,28]/100.
}
for (j in 1:N){
  theta[j, 1:5]~dmnorm (mu[coun[j],1:5], tau[1:5,1:5])
}
## specify priors for theta hyperparameters
for(g in 1:NC){
  mu[g, 1:5]~dmnorm (mupri[1:5],covpri[1:5, 1:5])
}
covpri[1:5,1:5]~dwish(covpri2[1:5,1:5],10)
tau[1:5, 1:5]~dwish(covpri2[1:5,1:5],10)
sigma1[1:5,1:5]<-inverse(covpri[, ])
sigma2[1:5,1:5]<-inverse(tau[, ])
for (g in 1:5){
  for (h in 1:5){
    cor1[g,h]<-sigma1[g,h]/sqrt(sigma1[g,g]*sigma1[h,h])
    cor2[g,h]<-sigma2[g,h]/sqrt(sigma2[g,g]*sigma2[h,h])
  }
}
}

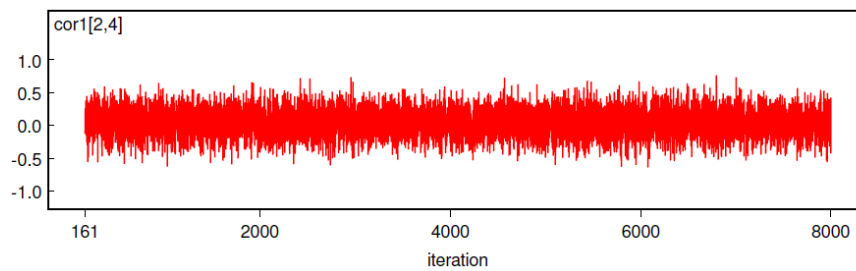
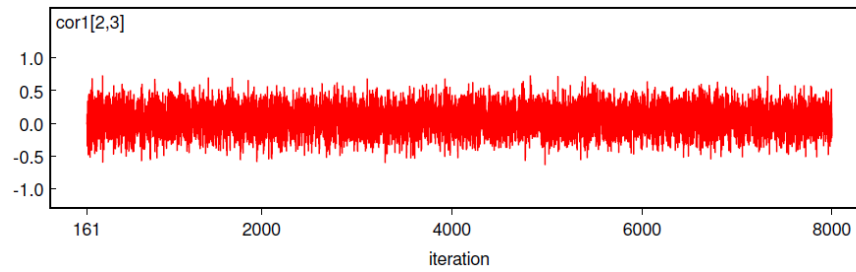
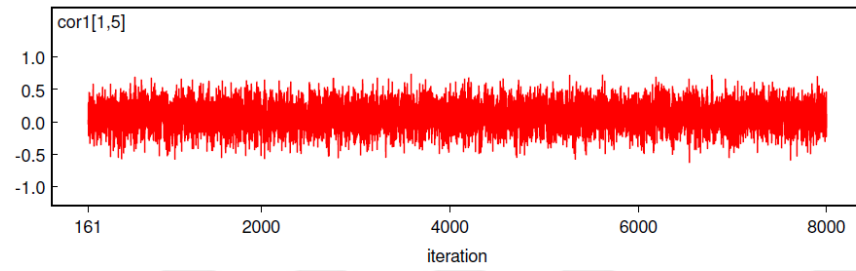
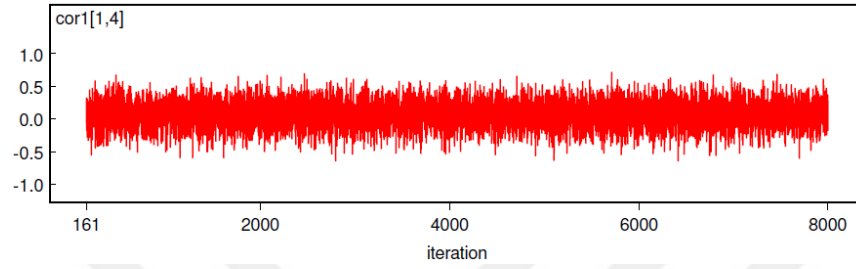
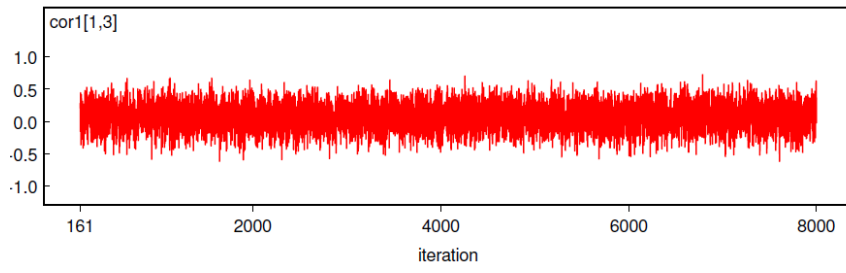
## Begining of data list
list(NC=15,mupri=c(0,0,0,0,0), covpri2=structure(.Data=c(
10,0,0,0,0,

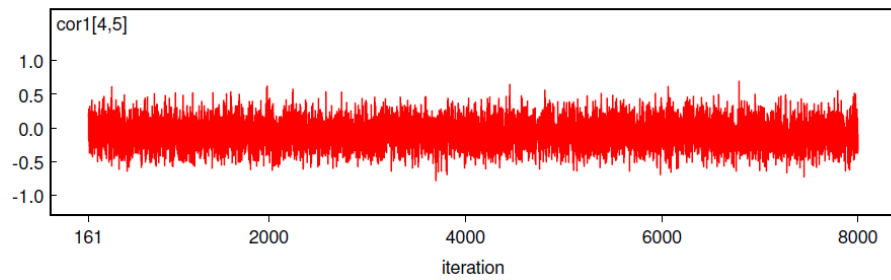
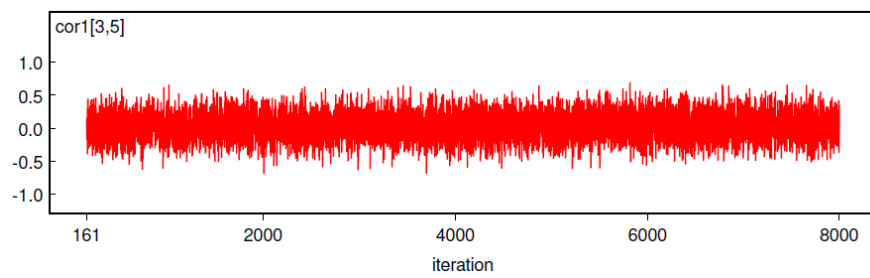
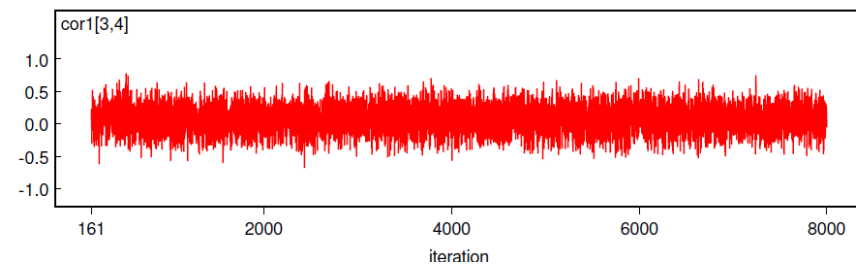
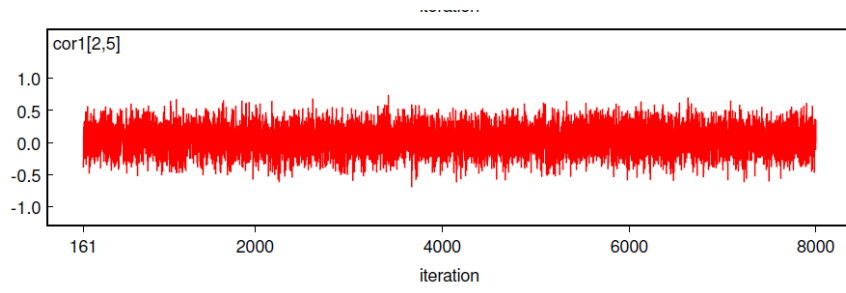
```


EKH. Üç Boyutlu Modele Ait Zaman Serisi Grafikleri



Dört Boyutlu Modele Ait Zaman Serisi Grafikleri





EK I.İlişkili Örneklemeler için t-testi Varsayımları

Çizelge 11
TIMSS 2015 Uygulamasına Katılan Herbir Ülke için Fark Puanlarına Ait Betimsel İstatistikler

Ülke		İstatistik değeri	Standart Hata	
Singapur	farksevme	Ortalama	-.11	
		Medyan	-.07	
		Mod	-.43	
		Std.Sapma	.32	
		Çarpıklık	-.19	.11
		Basıklık	.78	.22
	farkguven	Ortalama	-.04	
		Medyan	-.02	
		Mod	-.18	
		Std.Sapma	.20	
		Çarpıklık	-.16	.11
		Basıklık	.68	.22
Kore	farkdeger	Ortalama	-.15	0.01
		Medyan	-.14	
		Mod	-.43	
		Std.Sapma	.27	
		Çarpıklık	-.14	.11
		Basıklık	.05	.22
	farksevme	Ortalama	.01	
		Medyan	.01	
		Mod	-.31	
		Std.Sapma	.36	
		Çarpıklık	.07	.11
		Basıklık	.16	.22
farkguven	Ortalama	.03		
	Medyan	.02		
	Mod	-.36		
	Std.Sapma	.26		
	Çarpıklık	.17	.11	
	Basıklık	.06	.22	
farkdeger	Ortalama	-.13	0.01	
	Medyan	-.11		
	Mod	-.40		
	Std.Sapma	.26		
	Çarpıklık	.00	.11	
	Basıklık	.16	.22	

(devam ediyor)

Çizelge 11 (devam)

TIMSS 2015 Uygulamasına Katılan Herbir Ülke için Fark Puanlarına Ait Betimsel İstatistikler

Ülke			İstatistik değeri	Standart Hata
Çin-Tayvan	farksevme	Ortalama	.00	
		Medyan	-.01	
		Mod	-.70	
		Std.Sapma	.36	
		Çarpıklık	.04	.11
		Basıklık	.10	.22
	farkguven	Ortalama	.06	
		Medyan	.02	
		Mod	-.32	
		Std.Sapma	.28	
		Çarpıklık	.25	.11
		Basıklık	-.35	.22
Hong-Kong	farkdeger	Ortalama	-.06	
		Medyan	-.04	
		Mod	-.11	
		Std.Sapma	.23	
		Çarpıklık	-.27	.11
		Basıklık	.64	.22
	farksevme	Ortalama	-.08	
		Medyan	-.05	
		Mod	-.34	
		Std.Sapma	.33	
		Çarpıklık	-.10	.11
		Basıklık	.38	.22
farkguven	Ortalama	-.03		
	Medyan	-.02		
	Mod	-.19		
	Std.Sapma	.21		
	Çarpıklık	.38	.11	
	Basıklık	.22	.22	
farkdeger	Ortalama	-.05		
	Medyan	-.02		
	Mod	-.17		
	Std.Sapma	.25		
	Çarpıklık	-.19	.11	
	Basıklık	.58	.22	

(devam ediyor)

Çizelge 11 (devam)

TIMSS 2015 Uygulamasına Katılan Herbir Ülke için Fark Puanlarına Ait Betimsel İstatistikler

Ülke		İstatistik değeri	Standart Hata	
Japonya	farksevme	Ortalama	.01	
		Medyan	-.01	
		Mod	-.09	
		Std.Sapma	.27	
		Çarpıklık	-.09	.11
		Basıklık	.55	.22
	farkguven	Ortalama	.05	
		Medyan	.04	
		Mod	.10	
		Std.Sapma	.21	
		Çarpıklık	.23	.11
		Basıklık	.29	.22
	farkdeger	Ortalama	-.07	
		Medyan	-.04	
		Mod	.00	
		Std.Sapma	.18	
		Çarpıklık	-.36	.11
		Basıklık	.67	.22
İsveç	farksevme	Ortalama	-.02	
		Medyan	.00	
		Mod	-.16	
		Std.Sapma	.26	
		Çarpıklık	-.08	.11
		Basıklık	.53	.22
	farkguven	Ortalama	-.09	
		Medyan	-.06	
		Mod	-.21	
		Std.Sapma	.19	
		Çarpıklık	-.21	.11
		Basıklık	.19	.22
farkdeger	Ortalama	-.11		
	Medyan	-.07		
	Mod	-.17		
	Std.Sapma	.23		
	Çarpıklık	-.15	.11	
	Basıklık	.75	.22	

(devam ediyor)

Çizelge 11 (devam)

TIMSS 2015 Uygulamasına Katılan Herbir Ülke için Fark Puanlarına Ait Betimsel İstatistikler

Ülke			İstatistik değeri	Standart Hata
İtalya	farksevme	Ortalama	-.02	
		Medyan	-.02	
		Mod	-.11	
		Std.Sapma	.26	
		Çarpıklık	.15	.11
		Basıklık	.94	.22
	farkguven	Ortalama	-.05	
		Medyan	-.03	
		Mod	-.17	
		Std.Sapma	.18	
		Çarpıklık	-.22	.11
		Basıklık	.51	.22
farkdeger	Ortalama	-.02	0.01	
	Medyan	.00		
	Mod	-.05		
	Std.Sapma	.19		
	Çarpıklık	-.33	.11	
	Basıklık	1.21	.22	
Malta	farksevme	Ortalama	-.03	
		Medyan	-.03	
		Mod	-.26	
		Std.Sapma	.27	
		Çarpıklık	.33	.11
		Basıklık	1.18	.22
	farkguven	Ortalama	-.04	
		Medyan	-.03	
		Mod	-.29	
		Std.Sapma	.16	
		Çarpıklık	-.16	.11
		Basıklık	.44	.22
farkdeger	Ortalama	.02	0.01	
	Medyan	.02		
	Mod	-.14		
	Std.Sapma	.24		
	Çarpıklık	-.11	.11	
	Basıklık	.36	.22	

(devam ediyor)

Çizelge 11 (devam)

TIMSS 2015 Uygulamasına Katılan Herbir Ülke için Fark Puanlarına Ait Betimsel İstatistikler

Ülke			İstatistik değeri	Standart Hata
Avustralya	farksevme	Ortalama	-.03	
		Medyan	-.01	
		Mod	-.28	
		Std.Sapma	.23	
		Çarpıklık	.51	.11
		Basıklık	1.91	.22
	farkguven	Ortalama	-.07	
		Medyan	-.06	
		Mod	-.03	
		Std.Sapma	.16	
		Çarpıklık	-.11	.11
		Basıklık	1.06	.22
	farkdeger	Ortalama	-.12	
		Medyan	-.09	
		Mod	-.37	
		Std.Sapma	.24	
		Çarpıklık	-.04	.11
		Basıklık	.25	.22
Norveç	farksevme	Ortalama	-.01	
		Medyan	-.02	
		Mod	-.04	
		Std.Sapma	.23	
		Çarpıklık	.58	.11
		Basıklık	1.48	.22
	farkguven	Ortalama	-.07	
		Medyan	-.06	
		Mod	-.06	
		Std.Sapma	.15	
		Çarpıklık	-.18	.11
		Basıklık	.89	.22
farkdeger	Ortalama	-.06	0.01	
	Medyan	-.05		
	Mod	-.27		
	Std.Sapma	.21		
	Çarpıklık	-.13	.11	
	Basıklık	.65	.22	

(devam ediyor)

Çizelge 11 (devam)

TIMSS 2015 Uygulamasına Katılan Herbir Ülke için Fark Puanlarına Ait Betimsel İstatistikler

Ülke			İstatistik değeri	Standart Hata
Türkiye	farksevme	Ortalama	.04	
		Medyan	-.01	
		Mod	-.25	
		Std.Sapma	.31	
		Çarpıklık	.73	.11
		Basıklık	.93	.22
	farkguven	Ortalama	.00	
		Medyan	-.01	
		Mod	-.11	
		Std.Sapma	0.16	
		Çarpıklık	0.27	.11
		Basıklık	1.40	.22
	farkdeger	Ortalama	.10	
		Medyan	.08	
		Mod	-.03	
		Std.Sapma	.24	
		Çarpıklık	.22	.11
		Basıklık	.32	.22
Şili	farksevme	Ortalama	-.01	
		Medyan	-.02	
		Mod	.00	
		Std.Sapma	.27	
		Çarpıklık	.64	.11
		Basıklık	1.98	.22
	farkguven	Ortalama	-.02	
		Medyan	-.02	
		Mod	-.20	
		Std.Sapma	.16	
		Çarpıklık	-.08	.11
		Basıklık	1.23	.22
farkdeger	Ortalama	.06		
	Medyan	.04		
	Mod	-.09		
	Std.Sapma	.23		
	Çarpıklık	.14	.11	
	Basıklık	.41	.22	

(devam ediyor)

Çizelge 11 (devam)

TIMSS 2015 Uygulamasına Katılan Herbir Ülke için Fark Puanlarına Ait Betimsel İstatistikler

Ülke			İstatistik değeri	Standart Hata
Kuveyt	farksevme	Ortalama	.04	
		Medyan	.00	
		Mod	-.72	
		Std.Sapma	.34	
		Çarpıklık	.49	.11
		Basıklık	.67	.22
	farkguven	Ortalama	.00	
		Medyan	-.01	
		Mod	-.01	
		Std.Sapma	.17	
		Çarpıklık	.22	.11
		Basıklık	1.28	.22
	farkdeger	Ortalama	.11	
		Medyan	.08	
		Mod	.00	
Std.Sapma		.25		
Çarpıklık		.28	.11	
Basıklık		.31	.22	
Mısır	farksevme	Ortalama	.28	
		Medyan	.24	
		Mod	.17	
		Std.Sapma	.35	
		Çarpıklık	.36	.11
		Basıklık	-.35	.22
	farkguven	Ortalama	.07	
		Medyan	.02	
		Mod	-.09	
		Std.Sapma	.16	
		Çarpıklık	.83	.11
		Basıklık	.40	.22
	farkdeger	Ortalama	.31	
		Medyan	.31	
		Mod	.00	
Std.Sapma		.26		
Çarpıklık		-.06	.11	
Basıklık		-.66	.22	

(devam ediyor)

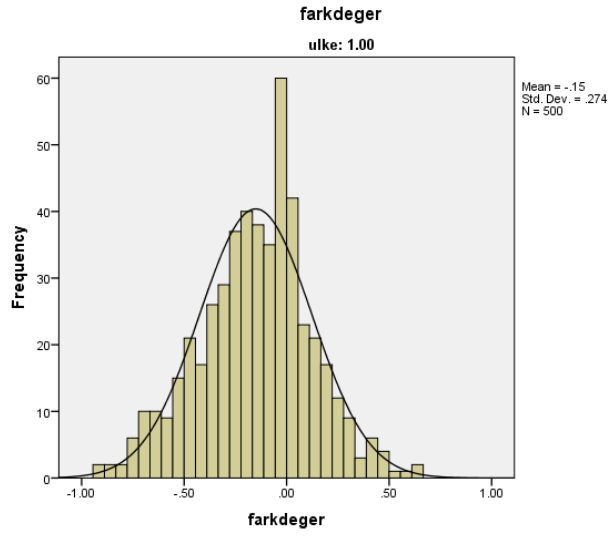
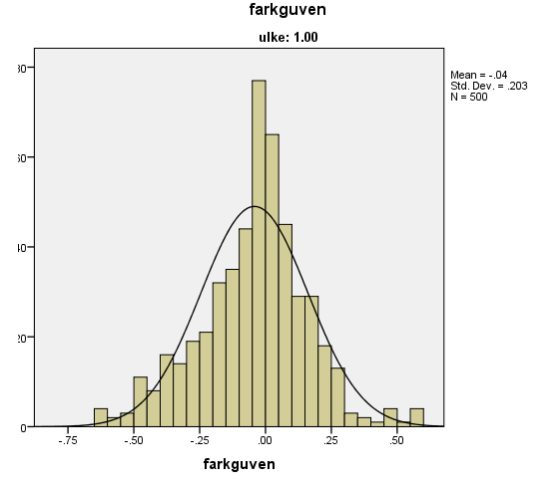
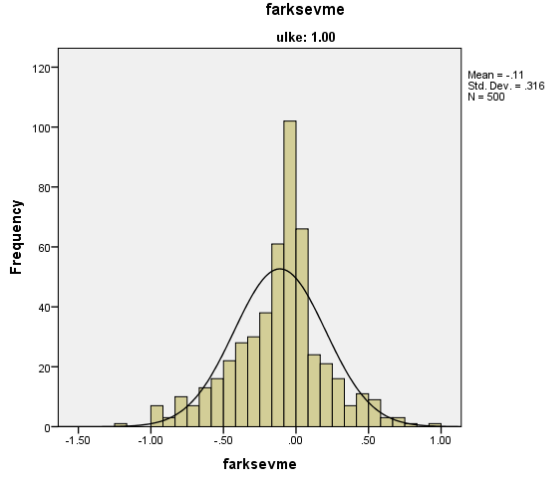
Çizelge 11 (devam)

TIMSS 2015 Uygulamasına Katılan Herbir Ülke için Fark Puanlarına Ait Betimsel İstatistikler

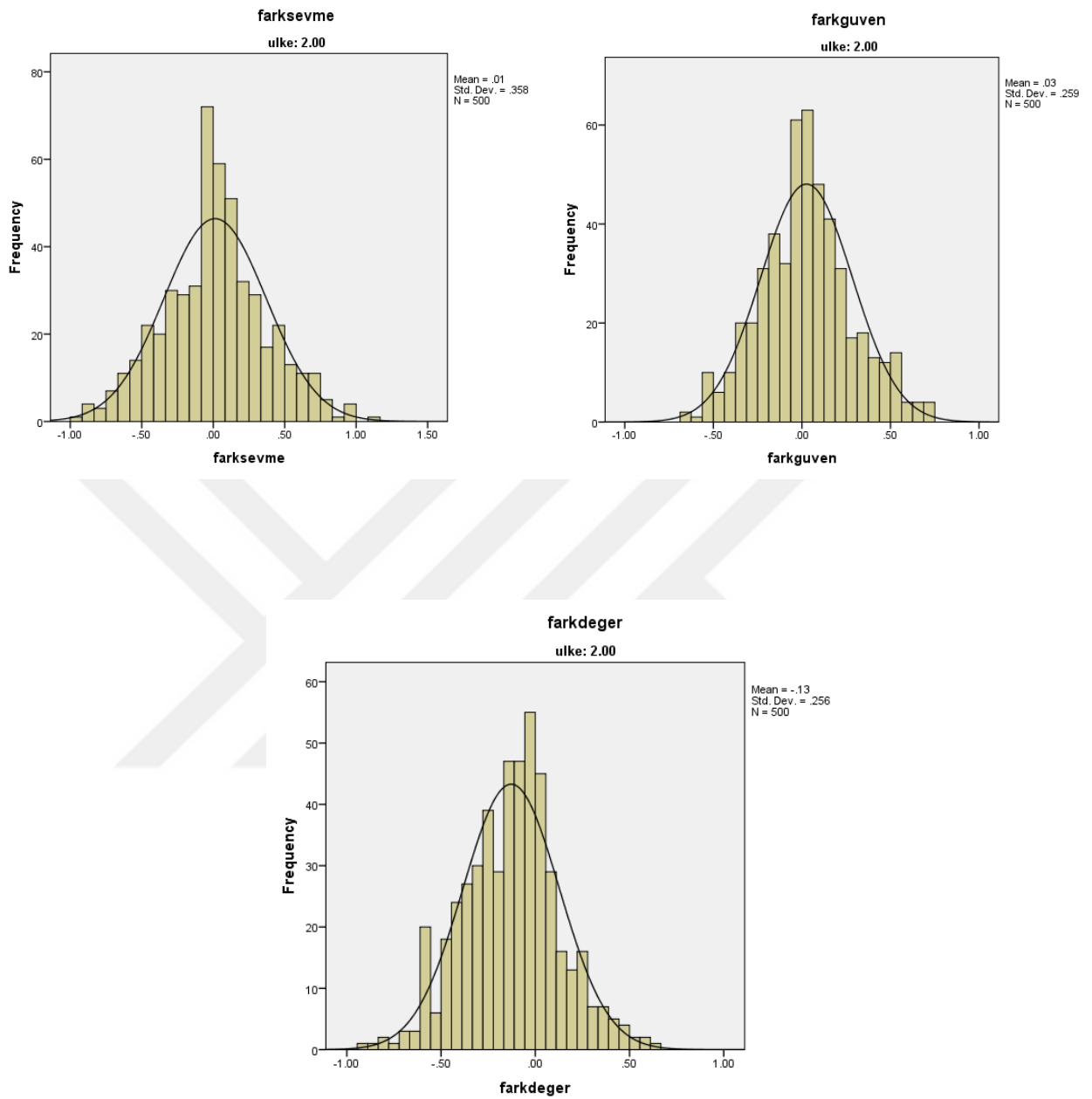
Ülke			İstatistik değeri	Standart Hata
Sudi Arabistan	farksevme	Ortalama	-.01	
		Medyan	-.03	
		Mod	-.09	
		Std.Sapma	.34	
		Çarpıklık	.40	.11
		Basıklık	.92	.22
	farkguven	Ortalama	-.01	
		Medyan	-.01	
		Mod	-.15	
		Std.Sapma	.16	
		Çarpıklık	.04	.11
		Basıklık	1.43	.22
	farkdeger	Ortalama	.10	
		Medyan	.07	
		Mod	.08	
Std.Sapma		.27		
Çarpıklık		.18	.11	
		Basıklık	.40	.22

EK I.2 Herbir Ülke için Histogram Grafikleri

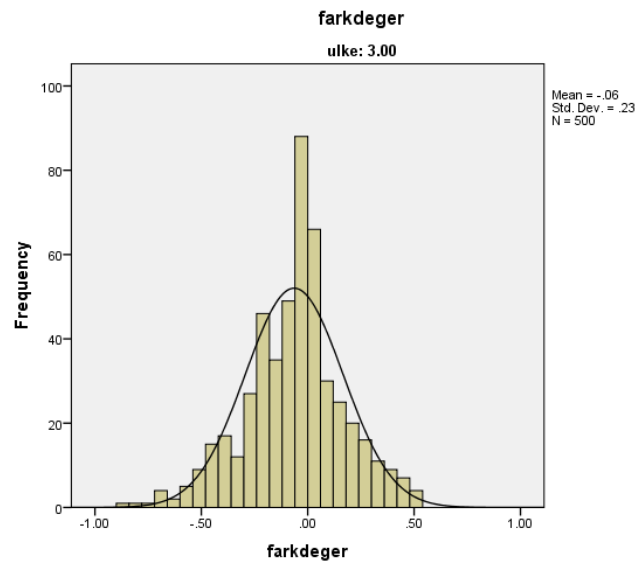
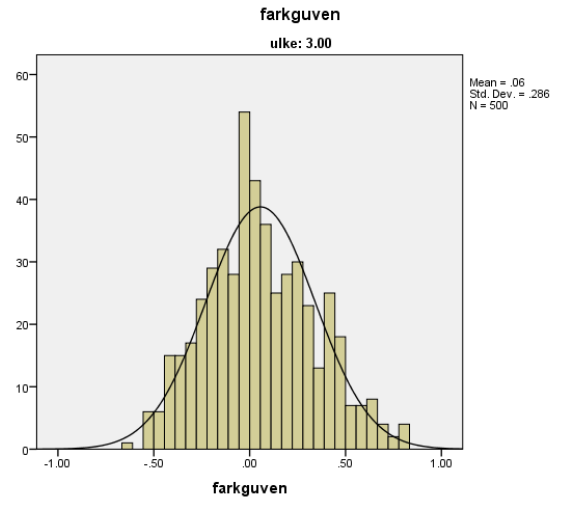
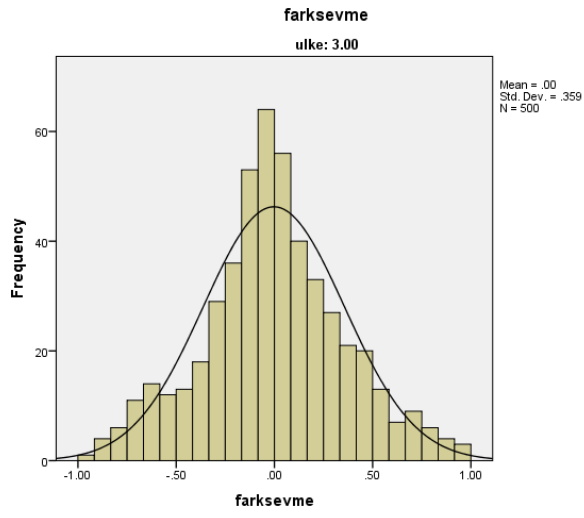
Singapur



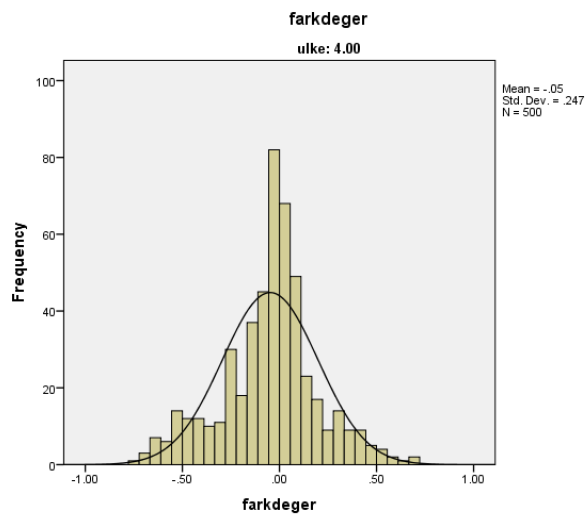
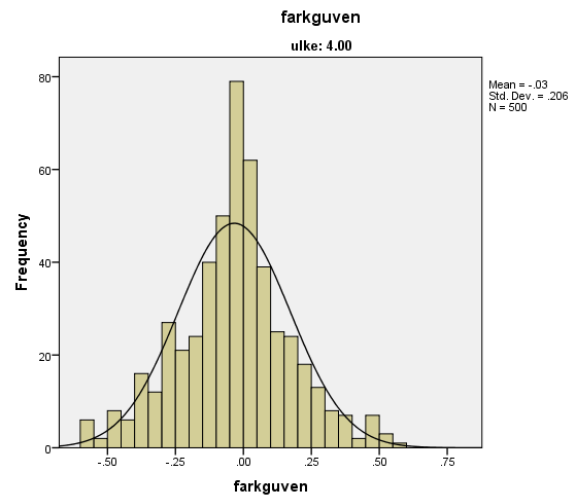
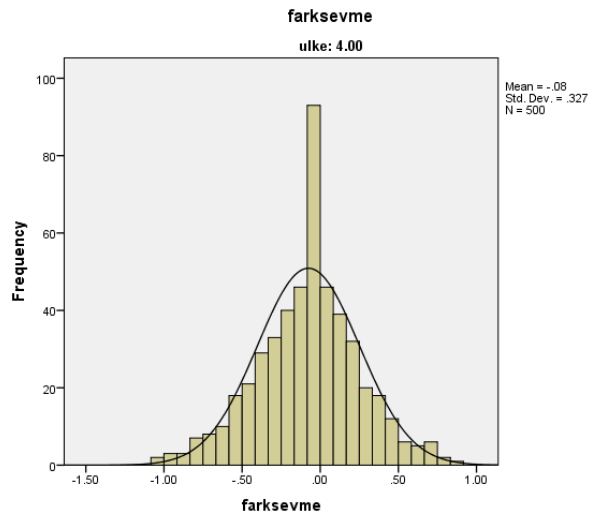
Kore



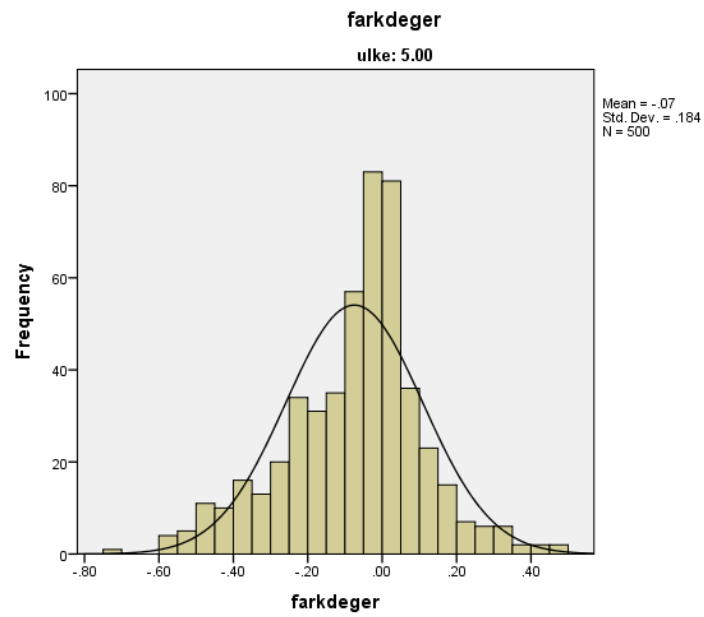
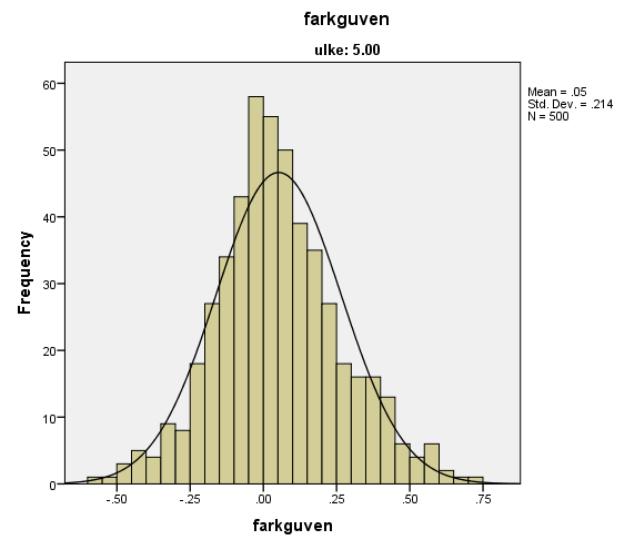
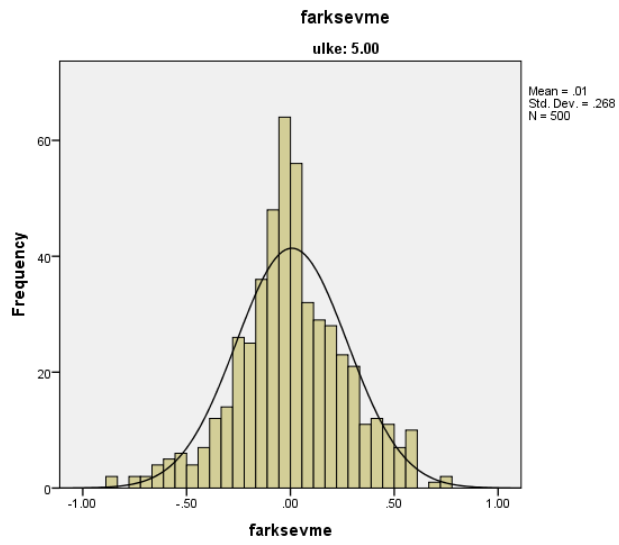
Çin-Tayvan



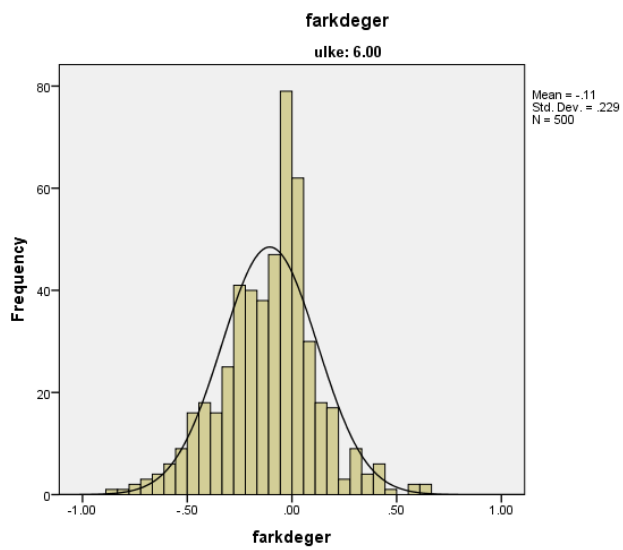
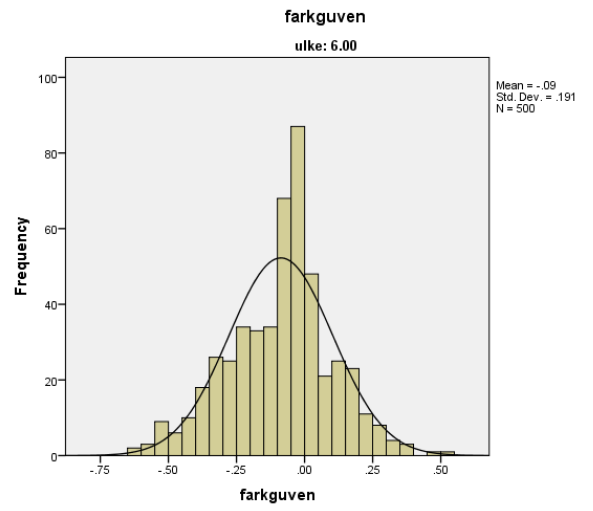
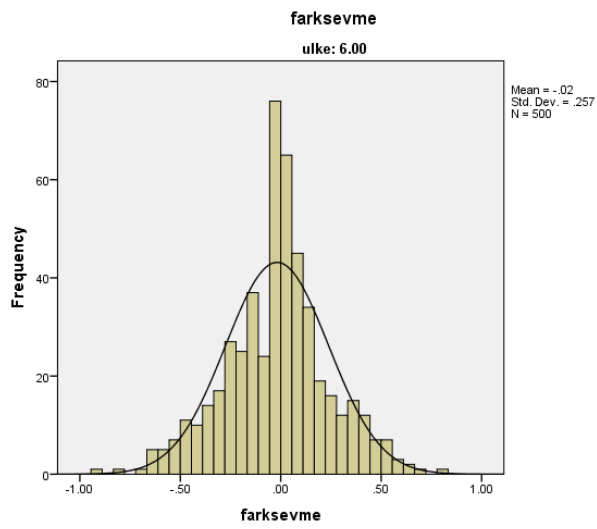
Hong-Kong



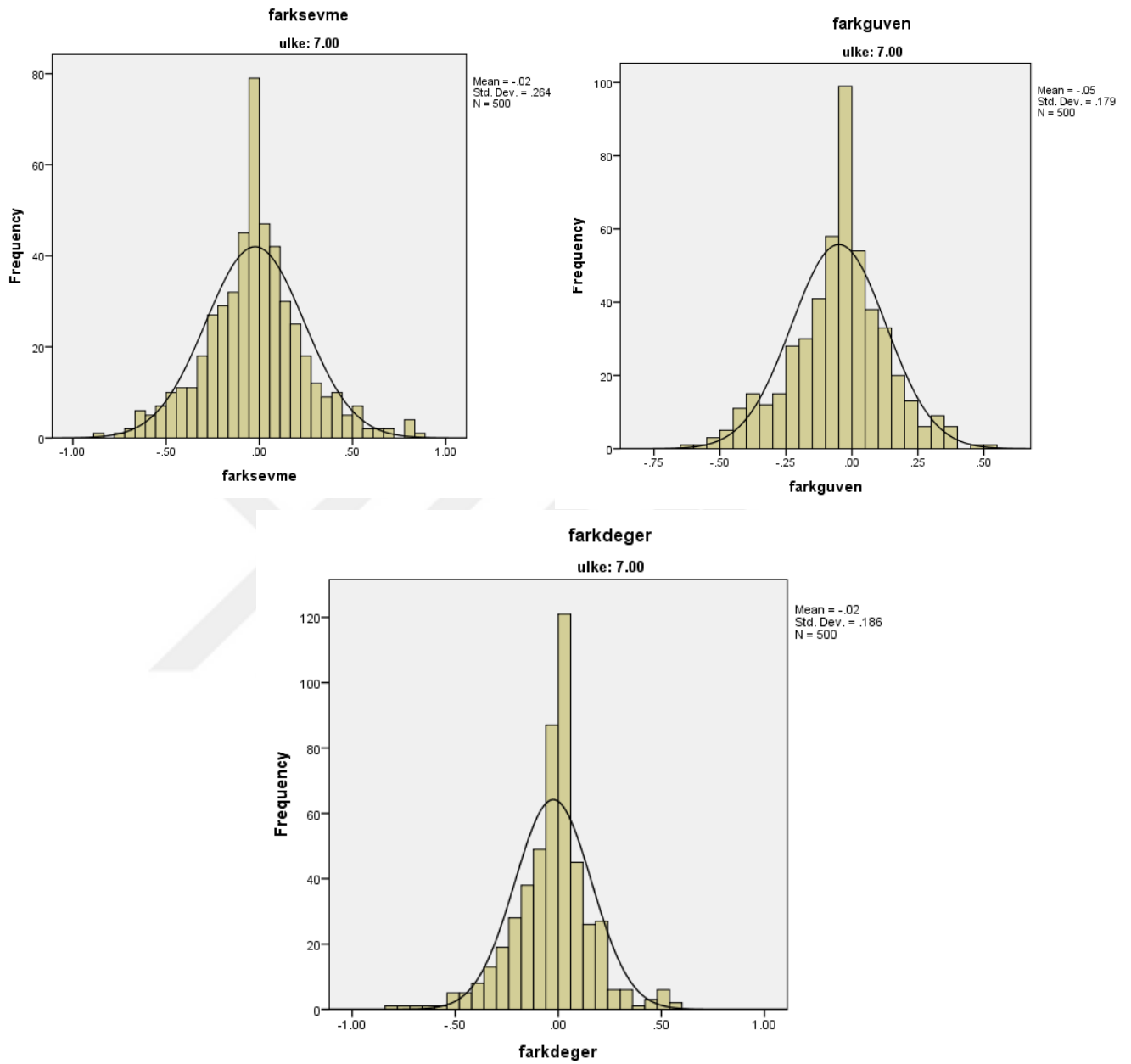
Japonya



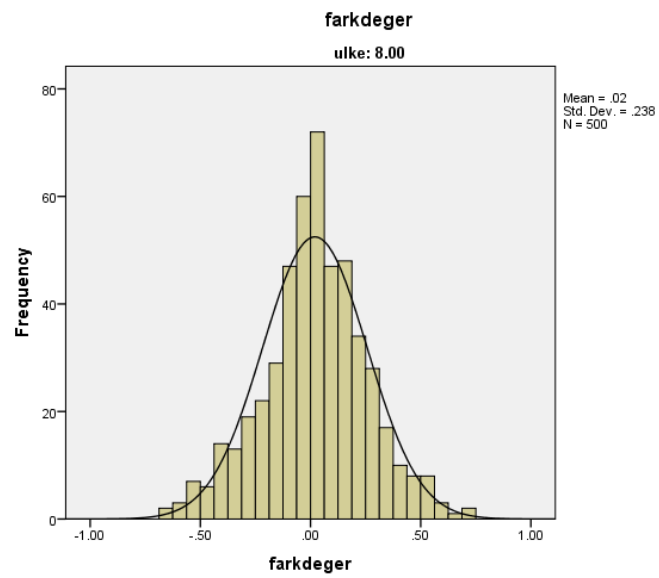
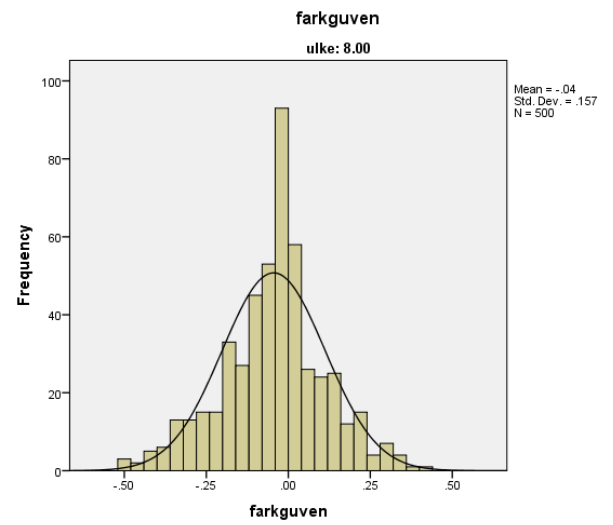
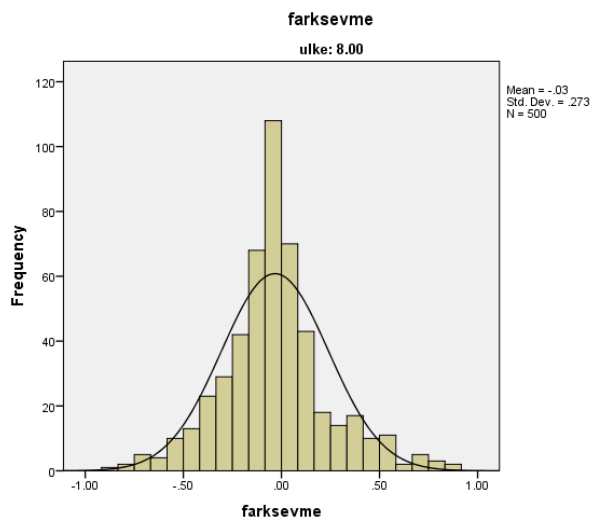
İsveç



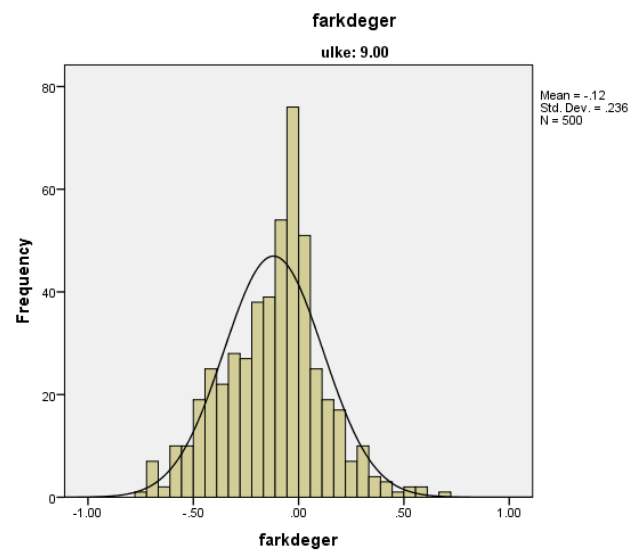
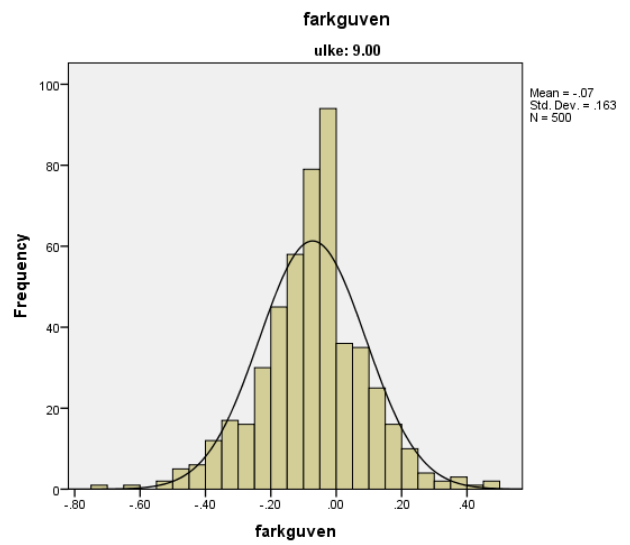
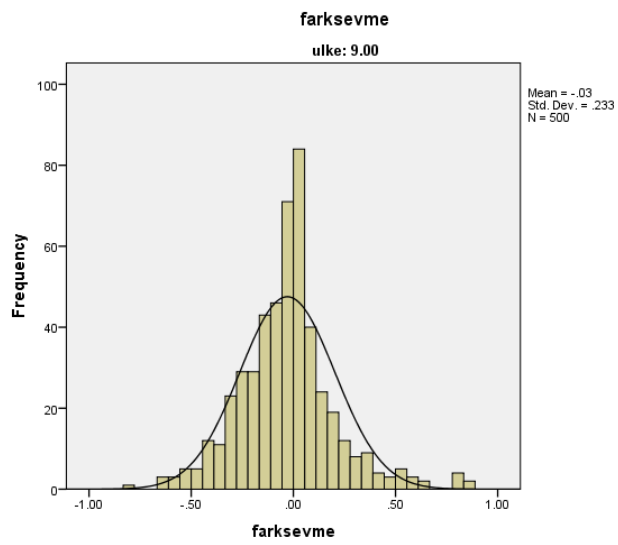
İtalya



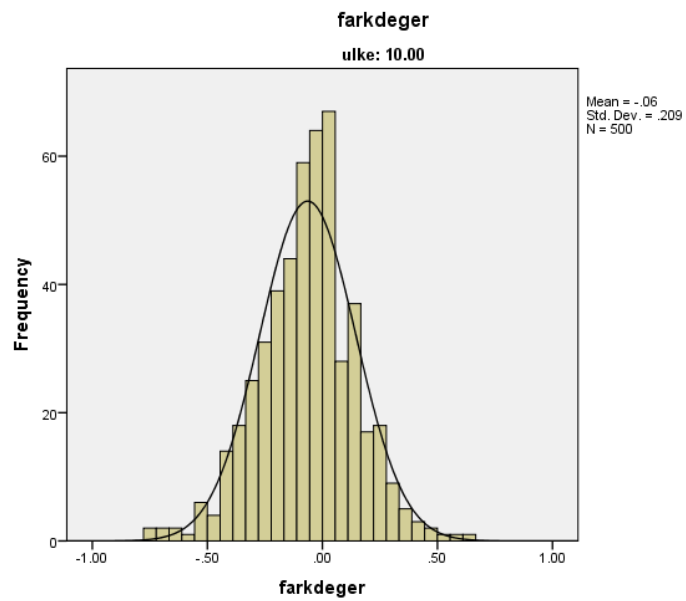
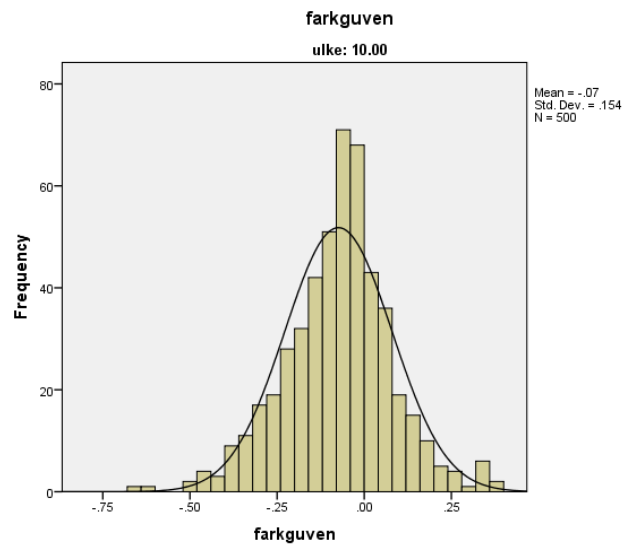
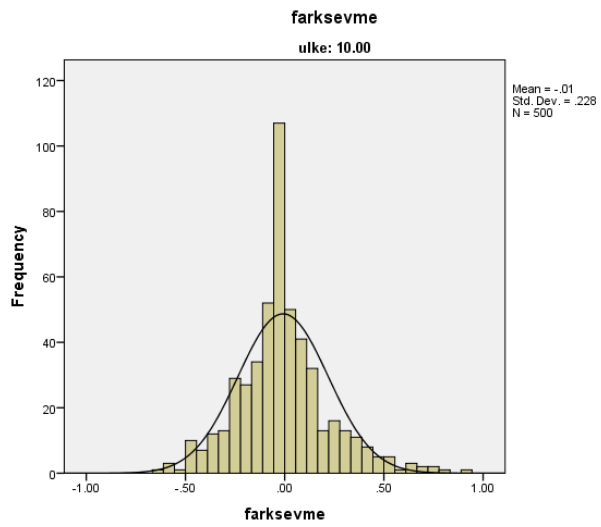
Malta



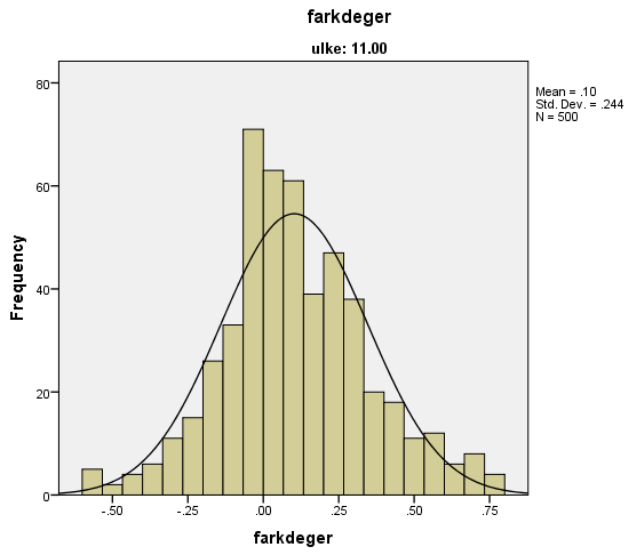
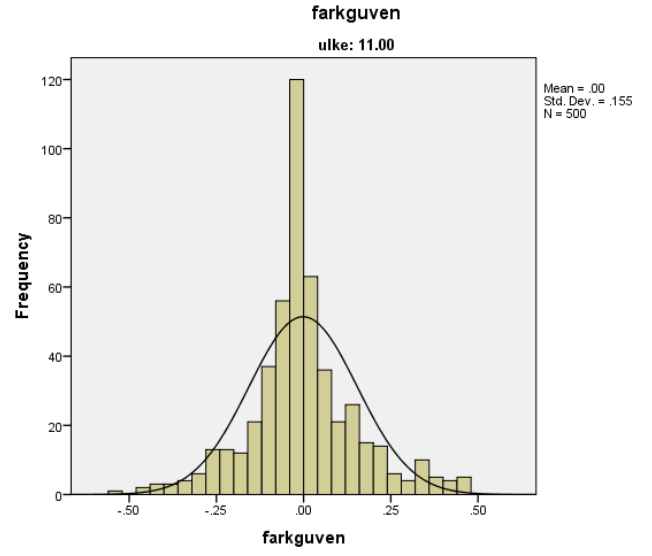
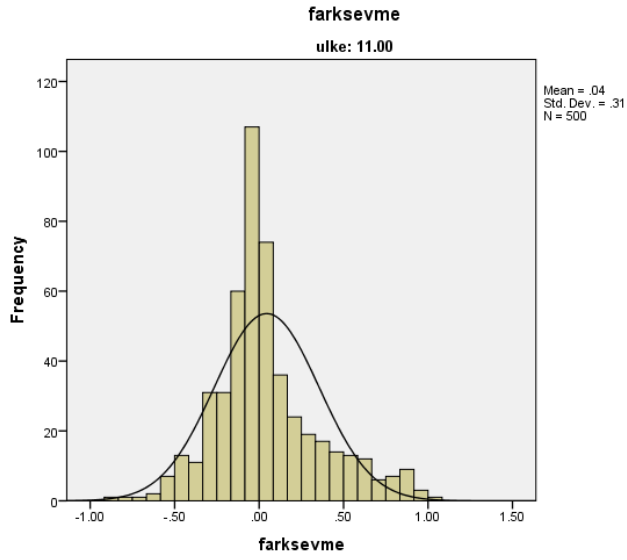
Avustralya



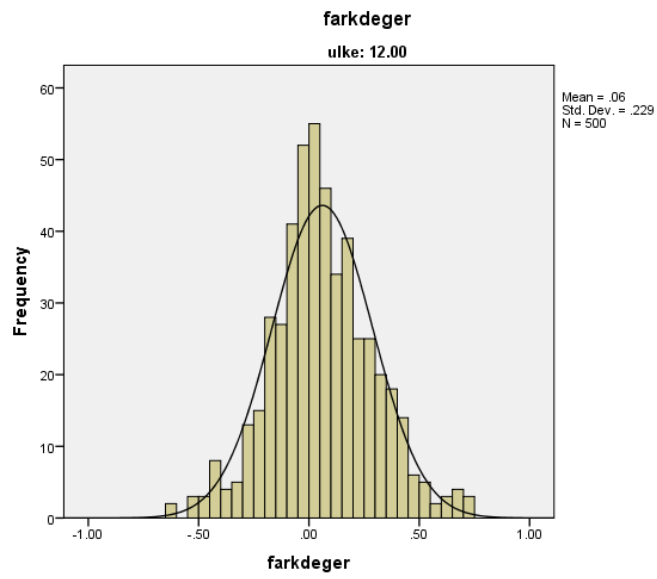
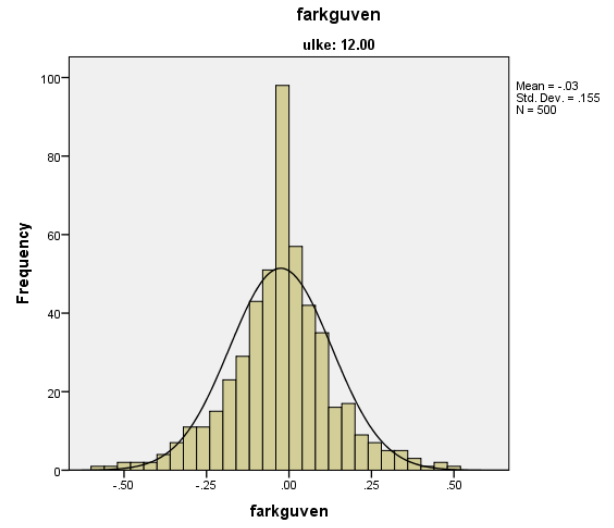
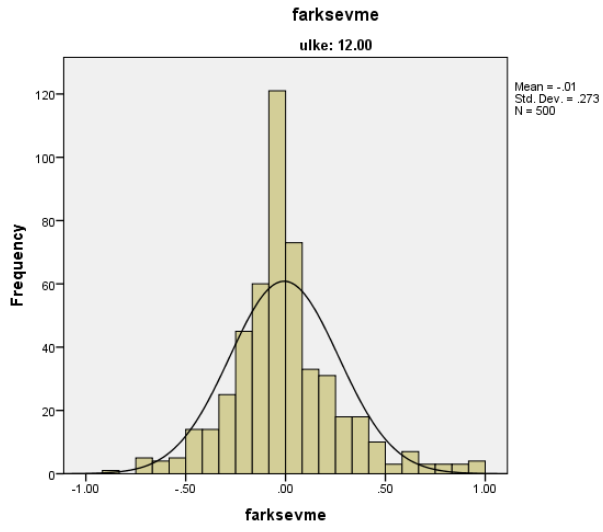
Norveç



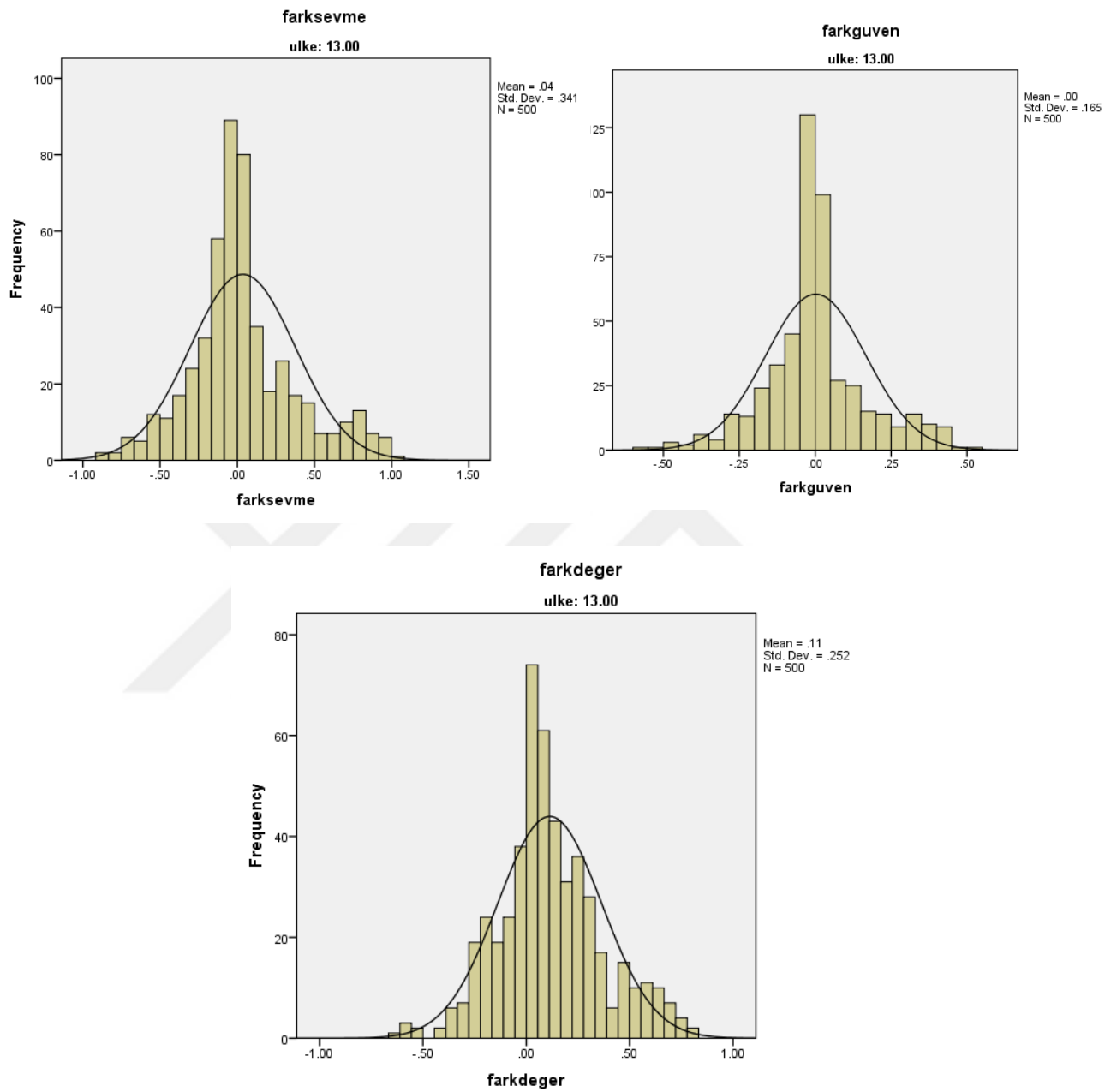
Türkiye



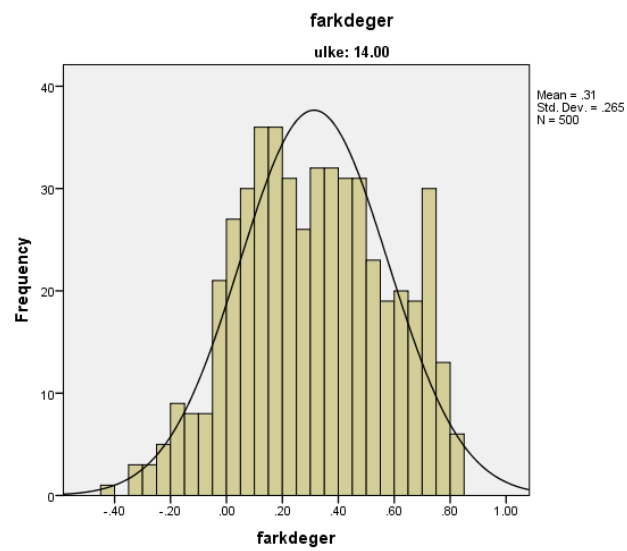
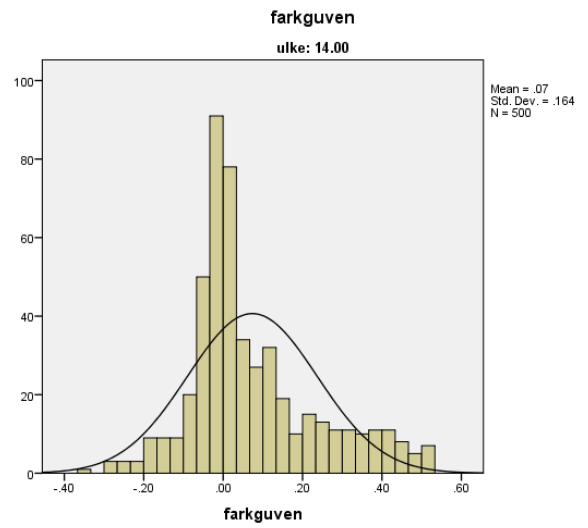
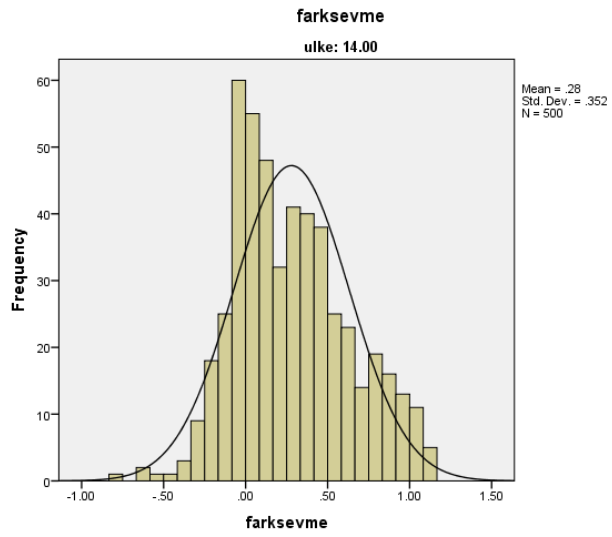
Şili



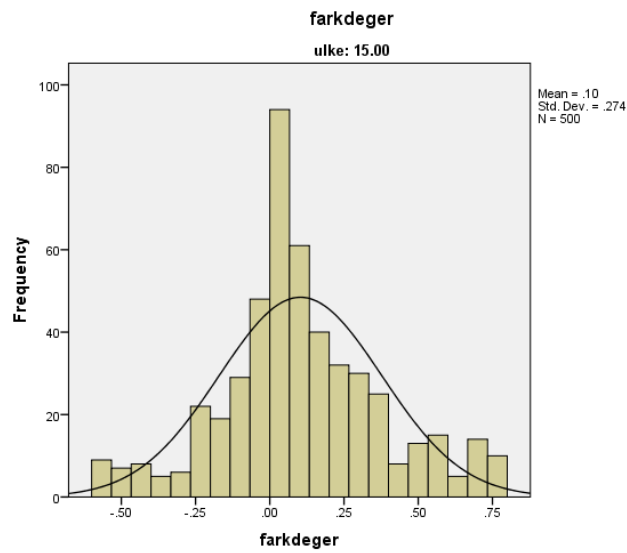
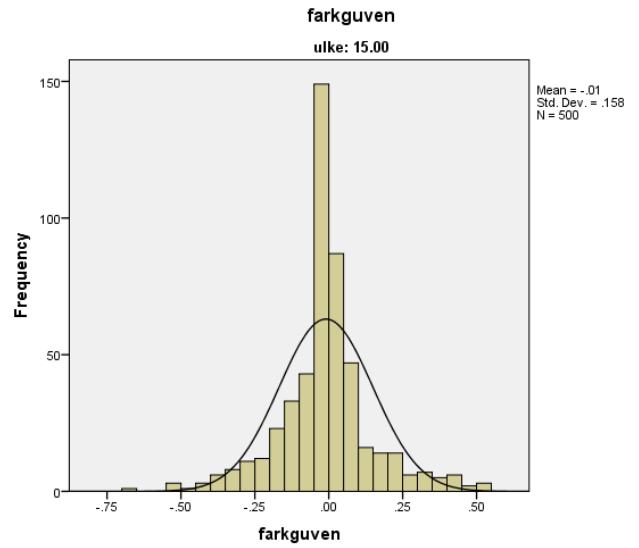
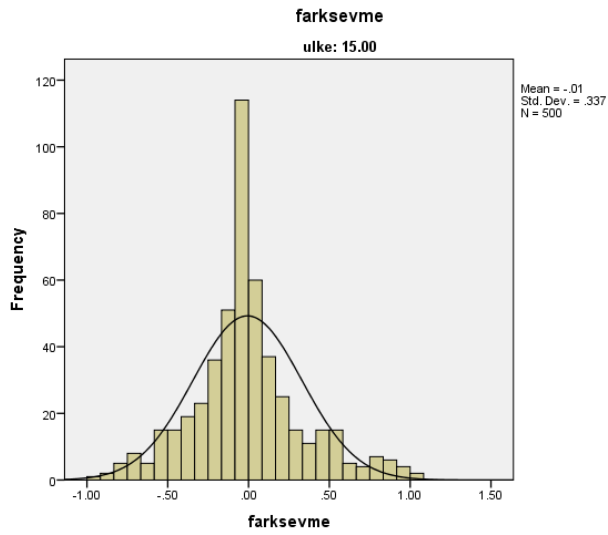
Kuveyt



Mısır



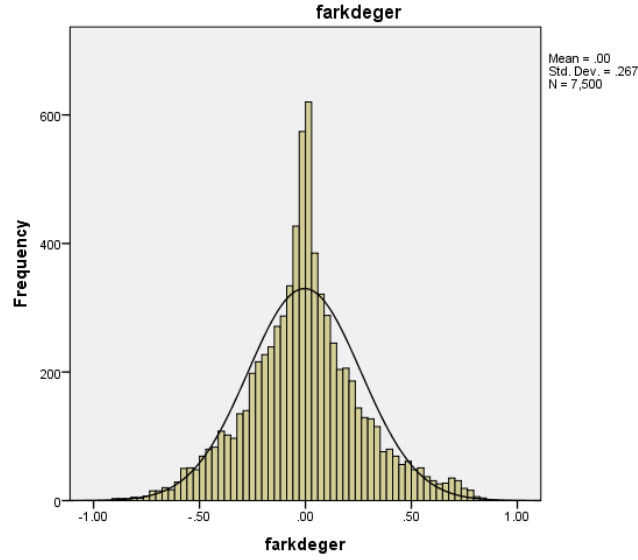
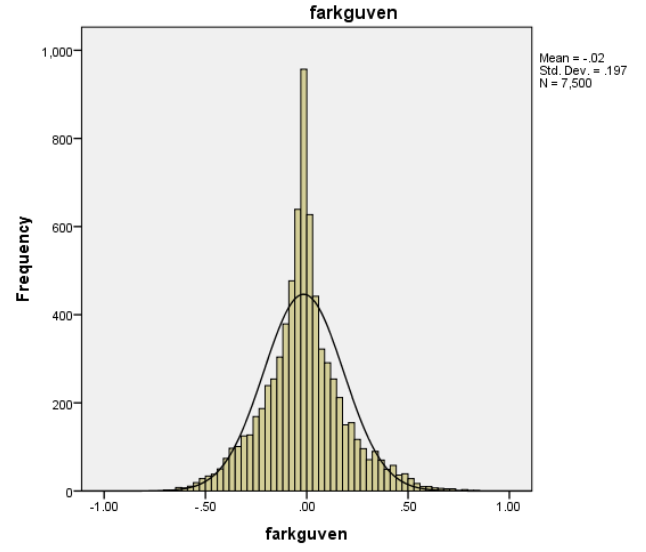
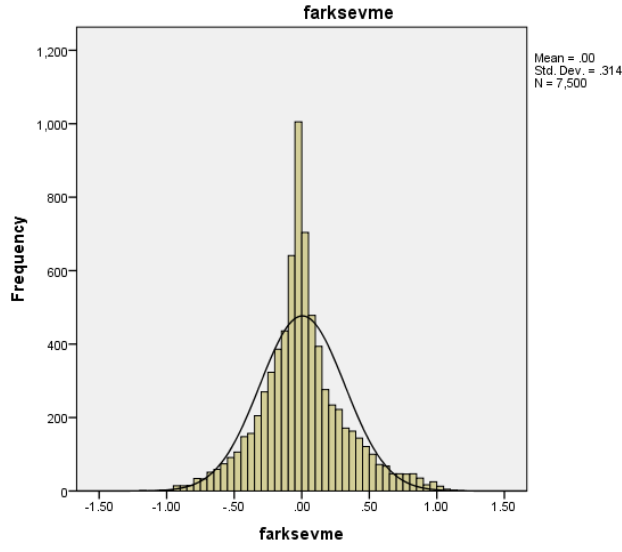
Sudi Arabistan



TIMSS 2015 Uygulamasına Katılan Ülkelerin Tamamı için Fark Puanlarına Ait Betimsel İstatistikler

		İstatistik değeri	Standart Hata
farksevme	Ortalama	.00	
	Medyan	-.01	
	Mod	-.01	
	Std.Sapma	.33	
	Çarpıklık	.25	.03
	Basıklık	1.02	.06
farkguven	Ortalama	-.01	
	Medyan	-.02	
	Mod	-.01	
	Std.Sapma	.20	
	Çarpıklık	.24	.03
	Basıklık	1.11	.06
farkdeger	Ortalama	.00	.01
	Medyan	.00	
	Mod	-.01	
	Std.Sapma	.27	
	Çarpıklık	.15	.03
	Basıklık	.59	.06

TIMSS 2015 Uygulamasına Katılan Ülkelerin Tamamı için Fark Puanlarına Ait Histogramlar



ÖZGEÇMİŞ

Adı ve Soyadı : Münevver İLGÜN DİBEK

Doğum Tarihi:27.09.1987

İletişim Bilgileri:Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, Ölçme ve Değerlendirme Bölümü

E-posta adresi: milgun@ankara.edu.tr

Öğrenim Durumu:

Derece	Bölüm/Program	Üniversitesi	Yıl
Lisans	İlköğretim Matematik Öğretmenliği	ODTÜ	2005-2010
Yüksek Lisans	İlköğretim Fen ve Matematik Eğitimi	ODTÜ	2010-2013
Yüksek Lisans	Ölçme ve Değerlendirme	Ankara Üniversitesi	2013-2015

İş Deneyimi:

Unvan	Görev Yeri	Yıl
Araştırma Görevlisi	Sakarya Üniversitesi	2010-2013
Araştırma Görevlisi	Ankara Üniversitesi	2013-devam ediyor