



**VERİLERİN KÜMELENMESİNDE TEPKİ STİLİ ETKİSİ: PISA 2015
ÖRNEĞİ**

Zafer Ertürk

DOKTORA TEZİ
EĞİTİM BİLİMLERİ ANABİLİM DALI
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME BİLİM DALI

GAZİ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ

AĞUSTOS, 2021

TELİF HAKKI VE TEZ FOTOKOPİ İZİN FORMU

Bu tezin tüm hakları saklıdır. Kaynak göstermek koşuluyla tezin teslim tarihi itibarıyla tezden fotokopi çekilebilir.

YAZARIN

Adı : Zafer

Soyadı : ERTÜRK

Bölümü : Eğitimde Ölçme ve Değerlendirme

İmza :

Teslim tarihi :

TEZİN

Türkçe ad: Verilerin Kümelenmesinde Tepki Stili Etkisi: PISA 2015 Örneği

İngilizce adı: The Effect of Response Style on Data Clustering: The Case of PISA 2015

ETİK İLKELERE UYGUNLUK BEYANI

Tez yazma sürecinde bilimsel ve etik ilkelere uyduđumu, yararlandıđım tüm kaynakları kaynak gösterme ilkelerine uygun olarak kaynakçada belirttiđimi ve bu bölümler dışındaki tüm ifadelerin şahsıma ait olduđunu beyan ederim.

Yazar Adı Soyadı: Zafer ERTÜRK

İmza:

JÜRİ ONAY SAYFASI

Zafer ERTÜRK tarafından hazırlanan “Verilerin Kümelenmesinde Tepki Stili Etkisi: PISA 2015 Örneği” adlı tez çalışması aşağıdaki jüri tarafından oy birliği ile Gazi Üniversitesi Eğitim Bilimleri Anabilim Dalı’nda Doktora Tezi olarak kabul edilmiştir.

Danışman: Prof. Dr. Şeref TAN

Eğitimde Ölçme ve Değerlendirme, Gazi Üniversitesi

Başkan: Prof. Dr. Selahattin GELBAL

Eğitimde Ölçme ve Değerlendirme, Hacettepe Üniversitesi

Üye: Prof. Dr. Mehtap ÇAKAN

Eğitimde Ölçme ve Değerlendirme, Gazi Üniversitesi

Üye: Prof. Dr. Nuri DOĞAN

Eğitimde Ölçme ve Değerlendirme, Hacettepe Üniversitesi

Üye: Prof. Dr. Hakan Yavuz ATAR

Eğitimde Ölçme ve Değerlendirme, Gazi Üniversitesi

Tez Savunma Tarihi: 17/06/2021

Bu tezin Eğitim Bilimleri Anabilim Dalı’nda doktora tezi olması için şartları yerine getirdiğini onaylıyorum.

Prof. Dr. Yücel GELİŞLİ

Eğitim Bilimleri Enstitüsü Müdürü

.....



Sevgili Aileme

TEŞEKKÜR

Doktora eğitimim boyunca birlikte çalışmaktan mutluluk duyduğum, akademik hayatımda yol gösterici olan ve gelişmemde çok büyük katkısı bulunan ve her konuda benden yardımlarını esirgemeyen değerli hocam ve danışmanım sayın Prof. Dr. Şeref TAN' a çok teşekkür ederim.

Bu çalışmanın başından sonuna kadar birçok katkı sağlayan, araştırma sürecinin şekillenmesinde katkılarını esirgemeyen ve bu süreçte her zaman yanımda olan tez izleme komitemde ve jürimde yer alan çok kıymetli hocalarım Prof. Dr. Hakan Yavuz ATAR ve Prof. Dr. Nuri DOĞAN'A çok teşekkür ederim. Tez savunmamda bulunarak değerli görüşler ve önerileri ile tezime değerli katkılar sağlayan Prof. Dr. Mehtap ÇAKAN ve Prof. Dr. Selahattin GELBAL'a çok teşekkür ederim.

Tezin analizlerinde yardımlarını esirgemeyen ve bütün sorularıma yanıt alabildiğim saygı değer Doç. Dr. Mariko Takagishi' ye teşekkür ederim.

Tezin hazırlanması sürecinde yardımını esirgemeyen arkadaşlarım Arş. Gör. Oya Erdiñç AKAN, Arş. Gör. Ayşenur ERDEMİR, Arş. Gör. Esra OYAR, Arş. Gör. Gül İNCE ARACI ve Arş. Gör. Vildan ÖZDEMİR'e teşekkür ediyorum. Anabilim dalı başkanımız Sayın Prof. Dr. Mehtap ÇAKAN başta olmak üzere Gazi Üniversitesi Eğitimde Ölçme ve Değerlendirme anabilim dalı öğretim üyeleri ve asistan arkadaşlarıma teşekkürlerimi sunarım.

Her daim yanımda olan, bu süreçte beni destekleyen ve motive eden kıymetli eşime çok teşekkür ederim. Bu günlere gelmemi sağlayan, benim için ellerinden gelen her şeyi yapan başta annem ve babam olmak üzere tüm aileme çok teşekkür ederim. Her zaman yanımda olan verdikleri destek ile beni mutlu eden kıymetli eşimin ailesine teşekkürlerimi sunarım. Her zaman yüzümü güldüren, ailemizin neşesi canım yeğenim Mustafa Ali TÜKENMEZ' e çok teşekkür ederim.

VERİLERİN KÜMELENMESİNDE TEPKİ STİLİ ETKİSİ: PISA 2015 ÖRNEĞİ

(Doktora Tezi)

Zafer Ertürk

GAZİ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Ağustos, 2021

ÖZ

Tutum-başarı paradoksu ülke içinde tutum ile başarı arasındaki ilişkinin pozitif, ülkeler arasında ise tutum ile başarı arasındaki ilişkinin negatif olmasıdır. Tepki stili yanlılığı da tutum-başarı paradoksunun sebeplerinden biri olarak kabul edilmekte ve istatistiksel analiz sonuçlarının geçerliğini olumsuz şekilde etkilemektedir. Bu doğrultuda bu çalışmanın amacı, PISA 2015 uygulamasına katılan ülkelerdeki öğrencilerin tepki stili yanlılıklarını düzelterek içeriğe dayalı kümelenmesini sağlamak ve tepki stili yanlılığının tutum-başarı paradoksu üzerindeki etkisini belirleyerek ülke içi ve ülkeler arasındaki tutum ile başarı arasındaki ilişkiyi açıklamaktır. Korelasyonel desende yürütülen araştırmanın örneklemini PISA 2015 uygulamasına katılan ve tutum-başarı paradoksunun görüldüğü ülkeler arasından seçilen 21 ülkedeki 15 yaş grubundaki öğrenciler oluşturmaktadır. Araştırmada öğrencilerin fene yönelik tutumlarını belirlemek amacıyla Fenden Hoşlanma, Başarı Motivasyonu ve Fen Öz Yeterliği ölçekleri kullanılmıştır. Tepki stili yanlılığının düzeltilmesi amacıyla verilerin analizinde Tepki Stili Yanlılığının Düzeltilmesi ve Kümelenmesi (TSYDK) yöntemi kullanılmıştır. TSYDK yöntemi ile tepki stili yanlılığından arındırılmış düzeltilmiş puanlar elde edilmekte ve bu puanlar üzerinden öğrencilerin kümelenmesi gerçekleştirilmektedir. Ayrıca, düzeltilmemiş ham puanlar üzerinden K-ortalamar kümeleme yöntemi de kullanılarak TSYDK yöntemi ile karşılaştırması yapılmıştır. Düzeltilmemiş ve düzeltilmiş

fene yönelik tutum puanları arasındaki farkın manidarlığını değerlendirmede her bir ülke için ilişkili örneklem t-testi analizi yapılmıştır. Fene yönelik tutum ile fen başarısı arasındaki ilişkide tepki stillerinin etkisini incelemek amacıyla düzeltilmemiş ve düzeltilmiş puanlara dayalı korelasyon değerleri elde edilmiştir. Çalışmadan elde edilen sonuçlara göre, TSYDK yöntemi ile araştırmada yer alan her üç fene yönelik tutum ölçeğindeki tepki stili yanlılığı düzeltilmiş ve içerik tabanlı kümeler elde edilmiştir. Ayrıca TSYDK yöntemi ile tepki stili tabanlı kümelerde elde edilmiştir. TSYDK yöntemi ile elde edilen tepki stili tabanlı ve içerik tabanlı kümelerin birbirinden farklı olduğu görülmüştür. Her üç ölçek için K-ortalamlar yöntemi ile elde edilen kümelemede ise ölçülen özellik açısından öğrencilerin kümelere ayrılmadığı, tepki stili tabanlı kümeleme sonuçlara benzer sonuçlar elde edilmiştir. TSYDK yönteminin bireyleri kümeleme performansı açısından K-ortalamlar kümeleme yöntemine göre daha başarılı olmuştur. Her üç ölçek içinde ülkelerin büyük çoğunluğunda düzeltilmemiş ve düzeltilmiş fene yönelik tutum puan ortalamaları arasındaki farklılığın manidar olduğu sonucuna ulaşılmıştır. Düzeltilmemiş ve düzeltilmiş puanlar aracılığıyla elde edilen ülkeler arası tutum-başarı korelasyon değerleri her üç ölçek içinde negatif bulunmuştur. Düzeltilmiş puanlar aracılığıyla elde edilen ülkeler arası tutum-başarı korelasyon değerleri, düzeltilmemiş puanlar aracılığıyla elde edilen ülkeler arası tutum-başarı korelasyon değerinden daha düşüktür. Elde edilen bütün sonuçlar birlikte değerlendirildiğinde; tutum-başarı paradoksunun sebeplerinden birinin tepki stili yanlılığının olduğu ancak tepki stili yanlılığının düzeltilmesinin bu paradoksu ortadan kaldırmada tek başına yeterli olmadığı ve TSYDK yöntemi ile tepki stili yanlılığının düzeltilebildiği, ölçülen özellik açısından öğrencilerin TSYDK yöntemi ile kümelenmesinin daha geçerli sonuçlar verdiği görülmüştür.

Anahtar Kelimeler : Tepki Stili, PISA, Kümeleme, Tutum-Başarı Paradoksu

Sayfa Adedi : xix + 182

Danışman : Prof. Dr. Şeref TAN

**THE EFFECT OF RESPONSE STYLES ON DATA CLUSTERING:
THE CASE OF PISA 2015**

(Ph. D Thesis)

Zafer Ertürk

GAZI UNIVERSITY

GRADUATE SCHOOL OF EDUCATIONAL SCIENCES

August, 2021

ABSTRACT

The attitude-achievement paradox is that the relationship between attitude and achievement in the country is positive, while it is negative among countries. Response style bias is considered one of the reasons for the attitude-achievement paradox and negatively affects the validity of statistical analysis results. In this regard, this study aims to ensure the content-based clustering of students from countries participating in PISA 2015 by correcting their response style bias and to explain the relationship between attitude and success within and between countries by determining the effect of response style bias on attitude-achievement paradox. The sample of the study, conducted in the correlational design, consists of 15-year-old students from 21 countries that participated in PISA 2015 survey and were selected among the countries with the attitude-achievement paradox. In this study, Enjoyment of Science, Achievement Motivation, and Science Self-Efficacy scales were used to determine the affective characteristics of students in science. In analyzing the data, the Correcting and Clustering Response Style Biased Data (CCRS) method was used to correct response style bias. Corrected scores, devoid of response style bias, are obtained using the CCRS method, and students are clustered based on these values. In addition, the K-means clustering method was also used over the uncorrected raw scores and compared with the CCRS method. To assess the significance of the difference between uncorrected and corrected science attitude scores, a related samples t-test analysis was performed for each country. The correlation values based on uncorrected and corrected scores were obtained in order to examine the

effect of response styles on the relationship between attitude towards science and science achievement. According to the results obtained from the study, response style bias in all three scales in the study was corrected with the CCRS method, and content-based clusters were obtained. Moreover, response style-based clusters were obtained with the CCRS method. It has been observed that the response style-based and content-based clusters obtained by the CCRS method are different from each other. In the clustering obtained by the K-means method for all three scales, the students were not divided into clusters in terms of the measured trait, and similar results to the response style-based clustering were obtained. CCRS method was more successful than K-means clustering method for clustering individuals. For all three scales, it was concluded that the difference between the uncorrected and corrected attitude towards science score averages in many of the countries was significant. Inter-country attitude-achievement correlation values obtained from uncorrected and corrected scores were found to be negative for all three scales. The attitude-achievement correlation values between countries obtained through corrected scores are lower than the correlation values obtained from uncorrected scores. When all of the findings are evaluated together; it has been seen that one of the reasons for the attitude-achievement paradox is the response style bias, but the correction of the response style bias solely is not sufficient to eliminate this paradox; the response style bias can be corrected with the CCRS method, and clustering the students with the CCRS method in terms of the measured trait gives more valid results.

Key Words: Response Style, PISA, Clustering, Attitude-Achievement Paradox

Page Number: xix + 182

Supervisor: Prof. Dr. Şeref TAN

İÇİNDEKİLER

TELİF HAKKI ve TEZ FOTOKOPİ İZİN FORMU.....	i
ETİK İLKELERE UYGUNLUK BEYANI	ii
JÜRİ ONAY SAYFASI	iii
TEŞEKKÜR.....	v
ÖZ.....	vi
ABSTRACT	viii
İÇİNDEKİLER.....	x
TABLolar LİSTESİ.....	xiv
ŞEKİLLER LİSTESİ	xvi
SİMGELER VE KISALTMALAR LİSTESİ.....	xviii
BÖLÜM I.....	1
GİRİŞ	1
Problem Durumu	1
Araştırmanın Amacı	9
Araştırmanın Önemi.....	9
Problem İfadesi	11
Araştırma Soruları.....	11
Sınırlılıklar.....	12

BÖLÜM II	13
KURAMSAL ÇERÇEVE	13
Tepki Stili Yanlılığının Düzeltilmesi ve Kümelenmesi (TSYDK)-Correcting and Clustering Response-Style-Biased Data-(CCRS)	13
Tepki Fonksiyonlarının Formüle Edilmesi	13
Tepki Fonksiyonları	15
İlgili Araştırmalar	25
Tepki Stillerini İncelemede Kullanılan Geleneksel Yaklaşımlar.....	25
Tepki Stillerini İncelemede Model Temelli Yaklaşımlar	31
DFA ile Yapılan Araştırmalar.....	31
GSA ile Yapılan Araştırmalar	32
MTK ile Yapılan Araştırmalar	36
BÖLÜM III	45
YÖNTEM	45
Araştırmanın Modeli	45
Evren ve Örneklem	45
Verilerin Elde Edilmesi	55
Veri Toplama Araçları	56
Fene Yönelik Tutum Anketleri.....	56
Verilerin Analizi.....	57
BÖLÜM IV	63
BULGULAR ve YORUMLAR	63
Birinci Araştırma Sorusuna İlişkin Elde Edilen Bulgular ve Yorumlar	63
Araştırma Sorusu 1: PISA 2015 uygulamasına katılan ve tutum-başarı paradoksunun görüldüğü ülkeler TSYDK yöntemi ile nasıl kümelenmektedir?	63
İkinci Araştırma Sorusuna İlişkin Elde Edilen Bulgular ve Yorumlar	79

Araştırma Sorusu 2: PISA 2015 uygulamasına katılan ve tutum-başarı paradoksunun görüldüğü ülkeler K-ortalamlar yöntemi ile nasıl kümelenebilir?	79
Üçüncü Araştırma Sorusuna İlişkin Elde Edilen Bulgular ve Yorumlar	91
Araştırma Sorusu 3: TSYDK ile K-ortalamlar yöntemlerinin ülkeleri kümeleme performanslarının karşılaştırması nasıl sonuçları vermektedir?	91
Dördüncü Araştırma Sorusuna İlişkin Elde Edilen Bulgular ve Yorumlar	94
a. Tepki stili yanlılığı bulunan veriler ile TSYDK yöntemi ile düzeltilmiş verilerde ortalama tutum puanları arasındaki farklar manidar mıdır?	94
b. Tepki stili yanlılığı bulunan veriler ile TSYDK yöntemi ile ortalama tutum puanlarının fen başarısı ile arasındaki ülke içi ve ülkeler arası ilişkiler nasıl değişmektedir?	102
BÖLÜM V	105
TARTIŞMA, SONUÇ ve ÖNERİLER	105
Tartışma ve Sonuç	105
Öneriler	122
Araştırmacılara Yönelik Öneriler	122
Uygulayıcılara Yönelik Öneriler	124
KAYNAKLAR	127
EKLER	145
Ek 1. Fene Yönelik Tutum Ölçekleri ve Maddeleri	146
Ek 2. Araştırmada Yer Alan Ülkelerin Öğrenci Sayıları ve Silinen Kayıp Veri Miktarları	147
Ek 3. Ülke Düzeyinde Örneklem ve Çalışma Örneklemi Fene Yönelik Tutum Ölçekleri Ortalama Puanları	148
Ek 4. Her Bir Ülke için Fene Yönelik Tutum Ölçekleri Cronbach Alfa Güvenirlik Katsayısı Değerleri	149

Ek 5. Fene Yönelik Tutum Ölçeklerinin Fark Puanlarına İlişkin Betimsel İstatistikler	150
Ek 6. Fene Yönelik Tutum Ölçeklerinin Fark Puanlarına İlişkin Histogram	159
Grafikleri	159
Ek 7. Bootstrap Örneklemelerinin Elde Edilmesinde Kullanılan R Kodları	180
Ek 8. TSYDK Yöntemi ve K-ortalamlar Tekniği ile Elde Edilen Kümeleme	182
Sonuçları için R Kodları	182



TABLÖLAR LİSTESİ

Tablo 1. <i>Fene Yönelik Tutum ile Fen Başarısı Arasındaki Ülke Ortalama Puanları Arasındaki (Ülkeler Arası) Korelasyon Değerleri</i>	48
Tablo 2. <i>Fene Yönelik Tutum ile Fen Başarısı Arasındaki Ülkeler Arası ve Ülke İçi Korelasyon Değerleri</i>	49
Tablo 3. <i>Belirlenen Ülkelerde Fene Yönelik Tutum Ölçekleri ile Fen Başarısı Arasındaki Ülkeler Arası ve Ülke İçi Korelasyon Değerleri</i>	52
Tablo 4. <i>Araştırmada Yer Alan Ülkelerin Örneklem ve Çalışma Örneklemeleri</i>	53
Tablo 5. <i>Ülkelerin Örneklem ve Çalışma Örneklemelerinin Fene Yönelik Tutum Ölçekleri Ortalama Puanlarına İlişkin İlişkisiz Örneklem t Testi Bulguları</i>	54
Tablo 6. <i>Ülkelerin Örneklemeleri ve Çalışma Örneklemelerinin Fen Başarısı ve Fene Yönelik Tutum Ölçekleri Arasındaki Ülke İçi Korelasyon Değerleri</i>	55
Tablo 7. <i>Fenden Hoşlanma Ölçeği için Ülkelerin TSYDK Yöntemi ile Elde Edilen Tepki Stili Tabanlı Kümeler (H=1,2,3) Üzerindeki Dağılımı</i>	67
Tablo 8. <i>Başarı Motivasyonu Ölçeği için Ülkelerin TSYDK Yöntemi ile Elde Edilen Tepki Stili Tabanlı Kümeler (H=1,2) Üzerindeki Dağılımı</i>	73
Tablo 9. <i>Fen Öz Yeterliği Ölçeği için Ülkelerin TSYDK Yöntemi ile Elde Edilen Tepki Stili Tabanlı Kümeler (H=1,2,3) Üzerindeki Dağılımı</i>	78
Tablo 10. <i>Fenden Hoşlanma Ölçeği için Ülkelerin K-ortalamalar Yöntemi ile Elde Edilen Kümelerdeki Dağılımı</i>	83
Tablo 11. <i>Başarı Motivasyonu Ölçeği için Ülkelerin K-ortalamalar Yöntemi ile Elde Edilen Kümelerdeki Dağılımı</i>	86

Tablo 12. <i>Fen Öz Yeterliği Ölçeği için Ülkelerin K-ortalamalar Yöntemi ile Elde Edilen Kümelerdeki Dağılımı</i>	90
Tablo 13. <i>Ülkelerin Fenden Hoşlanma Ölçeği için Düzeltilmemiş ve Düzeltilmiş Puanlarına İlişkin İlişkili Örneklem t Testi Bulguları</i>	95
Tablo 14. <i>Ülkelerin Başarı Motivasyonu Ölçeği için Düzeltilmemiş ve Düzeltilmiş Puanlarına İlişkin İlişkili Örneklem t Testi Bulguları</i>	97
Tablo 15. <i>Ülkelerin Fen Öz Yeterliği Ölçeği için Düzeltilmemiş ve Düzeltilmiş Puanlarına İlişkin İlişkili Örneklem t Testi Bulguları</i>	99
Tablo 16. <i>Ülkelerin Düzeltilmemiş ve Düzeltilmiş Fen Tutum Puanları ile Fen Başarıları Arasındaki Ülke İçi Korelasyon Değerleri</i>	102
Tablo 17. <i>Düzeltilmemiş ve Düzeltilmiş Fen Tutum Puanlarının Fen Başarısı ile Olan Ülkeler Arası Korelasyon Değerleri</i>	103

ŞEKİLLER LİSTESİ

Şekil 1. Tepki stili yanlılığı.	15
Şekil 2. Gözlemlenen değerin $x_{ij} = 6$ 'nın düzeltilmiş değere nasıl karşılık geldiğini gösteren örnek.	17
Şekil 3. Tepki fonksiyonları.	18
Şekil 4. (Sol) Sıralı sınır verilerinin bir örneği. (Sağ) Üç I-spline temel fonksiyonlar.	21
Şekil 5. Ülkeler arası fenden hoşlanma ile fen başarısı arasındaki saçılım diyagramı.	50
Şekil 6. Ülkeler arası başarı motivasyonu ile fen başarısı arasındaki saçılım diyagramı.	51
Şekil 7. Ülkeler arası fen öz yeterliği ile fen başarısı arasındaki saçılım diyagramı.	51
Şekil 8. Fenden hoşlanma ölçeği için TSYDK yöntemi ile elde edilen içeriğe ve tepki stiline dayalı bootstrap örneklemelerinin küme sayılarını gösteren histogram grafikleri.	64
Şekil 9. Fenden hoşlanma ölçeği için TSYDK yöntemi ile elde edilen içerik tabanlı kümelerde beş maddenin kutu grafikleri.	65
Şekil 10. (Sol) Fenden hoşlanma ölçeği için TSYDK yöntemi ile elde edilen tepki stili tabanlı kümelerin tahmini tepki fonksiyonları. (Sağ) Tepki stili tabanlı kümelerin β_h değeri için düşük boyutlu çizimi.	66
Şekil 11. Fenden hoşlanma ölçeği için TSYDK yöntemi ile elde edilen içerik tabanlı ve tepki stili tabanlı kümelerin mozaik grafiği.	68
Şekil 12. Başarı motivasyonu ölçeği için TSYDK yöntemi ile elde edilen içeriğe ve tepki stiline dayalı bootstrap örneklemelerinin küme sayılarını gösteren histogram grafikleri.	69
Şekil 13. Başarı motivasyonu ölçeği için TSYDK yöntemi ile elde edilen içerik tabanlı kümelerde beş maddenin kutu grafikleri.	70

<i>Şekil 14.</i> (Sol) Başarı motivasyonu ölçeği için TSYDK yöntemi tarafından elde edilen tepki stili tabanlı kümelerin tahmini tepki fonksiyonları. (Sağ) Tepki stili tabanlı kümelerin β_h değeri için düşük boyutlu çizimi.	72
<i>Şekil 15.</i> Başarı motivasyonu ölçeği için TSYDK yöntemi ile elde edilen içerik tabanlı ve tepki stili tabanlı kümelerin mozaik grafiği.	74
<i>Şekil 16.</i> Fen öz yeterliği ölçeği için TSYDK yöntemi ile elde edilen içeriğe ve tepki stiline dayalı bootstrap örneklemelerinin küme sayılarını gösteren histogram grafikleri	75
<i>Şekil 17.</i> Fen öz yeterliği ölçeği için TSYDK yöntemi ile elde edilen içerik tabanlı kümelerde sekiz maddenin kutu grafikleri.....	76
<i>Şekil 18.</i> (Sol) Fen öz yeterliği ölçeği için TSYDK yöntemi ile elde edilen tepki stili tabanlı kümelerin tahmini tepki fonksiyonları. (Sağ) Tepki - stili tabanlı kümelerin β_h değeri için düşük boyutlu çizimi.....	77
<i>Şekil 19.</i> Fen öz yeterliği ölçeği için TSYDK yöntemi ile elde edilen içerik tabanlı ve tepki stili tabanlı kümelerin mozaik grafiği.	79
<i>Şekil 20.</i> Fenden hoşlanma ölçeği için K-ortalamlar kümeleme yöntemine göre bootstrap örneklemelerinin küme sayılarını gösteren histogram grafiği.	80
<i>Şekil 21.</i> Fenden hoşlanma ölçeği için K-ortalamlar kümeleme yöntemi ile elde edilen kümelerde beş maddenin kutu grafikleri.....	81
<i>Şekil 22.</i> Başarı motivasyon ölçeği için K-ortalamlar kümeleme yöntemine göre bootstrap örneklemelerinin küme sayılarını gösteren histogram grafiği.	84
<i>Şekil 23.</i> Başarı motivasyonu ölçeği için K-ortalamlar kümeleme yöntemi ile elde edilen kümelerde beş maddenin kutu grafikleri.....	85
<i>Şekil 24.</i> Fen öz yeterliği ölçeği için K-ortalamlar kümeleme yöntemine göre bootstrap örneklemelerinin küme sayılarını gösteren histogram grafiği.	87
<i>Şekil 25.</i> Fen öz yeterliği ölçeği için K-ortalamlar kümeleme yöntemi ile elde edilen kümelerde sekiz maddenin kutu grafikleri.....	88

SİMGELER VE KISALTMALAR LİSTESİ

UTS	Uç Tepki Stili
KTS	Kabullenici Tepki Stili
ONTS	Orta Nokta Tepki Stili
RETS	Red Edici Tepki Stili
PISA	Uluslararası Öğrenci Değerlendirme Programı (Program for International Student Assessment)
TIMSS	Uluslararası Matematik ve Fen Bilgisi Araştırma Projesi (Trends in International Mathematics and Science Study)
PIRLS	Uluslararası Okuma ve Okuryazarlık Çalışmasında İlerleme (Progress in International Reading Literacy Study)
GSA	Gizil Sınıf Analizi
DFA	Doğrulayıcı Faktör Analizi
TSYDK	Tepki Stili Yanlılığının Düzeltilmesi ve Kümelenmesi
CCRS	Correcting and Clustering Response-Style- Biased Data
DASA	Doğu Asya Sosyal Anketi (East Asian Social Survey)
DMF	Değişen Madde Fonksiyonu

HLM	Hiyerarşik Lineer Modelleme
ÇBMTK	Çok Boyutlu Madde Tepki Kuramı
MTK	Madde Tepki Kuramı
ÇG-DFA	Çoklu Grup Doğrulayıcı Faktör Analizi
GSFA	Gizil Sınıf Faktör Analizi
KÇKMTK	Karma Çok Kategorili Madde Tepki Kuramı
IRTTree	Madde Tepki Kuramı Ağaç Modelleri
GLMM	Genelleştirilmiş Lineer Karma Model
SÇÖ	Sınırlandırılmış Çift Ölçeklendirme

BÖLÜM I

GİRİŞ

Bu bölümde araştırmanın problemine, amacına, önemine ve sınırlılıklarına ilişkin bilgilere yer verilmiştir.

Problem Durumu

Son yıllarda küreselleşmenin etkisiyle; ülkelerin kamusal, toplumsal bilimsel farkındalık değerleri ve tutumları ile ilgili uluslararası çalışmalarda büyük bir hız kazanmıştır. Uluslararası çalışmalar, kültürel olarak çeşitli ülkelerin eğitime yönelik tutumları ve değer yapılarındaki farklılıklarına dair önemli bilgiler sağlayabilir. Eğitim alanında, Uluslararası Öğrenci Değerlendirme Programı (PISA) ve Uluslararası Matematik ve Bilim Çalışmasında Eğilimler (TIMSS), ülkeler arasında karşılaştırma çalışmaları için fırsatlar sunan büyük ölçekli çalışmalardan ikisidir. PISA ve TIMSS gibi geniş ölçekli sınavlar aynı zamanda öğrenci başarısı ile öğrencinin geçmişi, tutumu, okul ve öğrencinin yaşadığı yere ilişkin özellikler arasındaki ilişkilerin incelenmesine de olanak vermektedir (Akyüz & Pala, 2010; Anderson, 2012; Aydın, Sarier, & Uysal, 2012). Bu geniş ölçekli sınavlar, belirli yaş ve sınıf seviyesindeki öğrencilerin akademik alanlardaki başarılarının sınıf, okul ve ülke düzeyindeki eğitim sistemleri perspektifinden incelendiği ve araştırıldığı çalışmalardır (Bos, 2002). Bu çalışmalardan elde edilen anketler ve yayınların sayısının artması, tutum ve değerler ile akademik başarı arasındaki ilişkide meydana gelen uluslararası farklılıklarda kanıt sunmaktadır (Lu, 2012).

PISA araştırması, her bir ülkedeki 15 yaş grubu öğrencilerin temsili bir örneğinin fen, matematik ve okuma becerileri alanlarındaki yeterliliklerini ölçmek için üç yılda bir düzenli

olarak uygulanmaktadır. İlk olarak 2000 yılında uygulanan PISA projesi, OECD üyesi ülkeler ve diğer katılımcı ülkelerdeki (dünya ekonomisinin yaklaşık olarak %90'ı) zorunlu eğitimi bitiren öğrencilerin çağdaş toplumda yerlerini alabilmeleri için gereken temel bilgi ve becerilere ne ölçüde sahip olduklarını ölçmeyi amaçlamaktadır (Demirez, 2018). PISA projesinin veri grubunu, 7. sınıf ve üzeri sınıf düzeylerinde örgün eğitime kayıtlı olan 15 yaş grubu öğrenciler oluşturmaktadır. PISA'da her üç yılda bir matematik, fen ve okuma becerileri alanlarından biri ana odak olarak seçilmektedir. PISA'nın altıncı döngüsü olan PISA 2015'de fen okuryazarlığı odak alan olarak belirlenmiştir. PISA projesinin temel amacı uluslararası bir referans olarak, her bir ülkenin eğitim sisteminin 15 yaş grubu gençleri ne kadar iyi eğittiğini ölçmektir. İlk olarak 2000 yılında OECD ülkelerinin katılımıyla gerçekleştirilen PISA'ya son yıllarda OECD ülkelerinin yanı sıra diğer pek çok ülke de katılmaktadır.

PISA'nın diğer büyük ölçekli sınavlardan farklı perspektifi, okul eğitimindeki takip ettiği çalışmalar açısından öğrencilerin bilgilerinin kalitesine bakmak değil, öğrencilerin günlük yaşam durumlarıyla özerk bir şekilde başa çıkma yeteneklerini değerlendirme fikridir. Bu perspektif, OECD uzmanları tarafından 'okuryazarlık' şeklinde ifade edilmektedir. PISA testleri tüm özel eğitim programlarından bağımsız olacak şekilde tasarlanmıştır (Bulle, 2011). PISA uygulamasında her dönem matematik okuryazarlığı, fen bilimleri okuryazarlığı ile okuma becerileri okuryazarlığı alanlarından birine odaklanılmaktadır. Yapılan PISA sınavı ile katılımcı ülkeler öğrencilerinin üst düzey bilişsel becerilere sahip olup olmadıklarını belirleyebilmekte ve elde edilen sonuçlarla hedefler arasındaki farklar incelenerek eğitim politikalarında değişiklik meydana getirecek göstergeler sağlanmaktadır (Aksu, Güzeller ve Eser, 2017). PISA sınavı ile ayrıca; öğrenciler, veliler ve okul yöneticilerinden anketler aracılığıyla veriler toplanmakta ve bu veriler vasıtasıyla öğrencilerin başarı ya da başarısızlık kaynakları incelenebilmektedir (Anıl, 2009; OECD, 2016).

PISA'nın altıncı döngüsü olan PISA 2015'de odak alan fen okuryazarlığıdır. Böylece PISA 2015 sonuçlarının matematik okuryazarlığı ve okuma becerilerinden daha çok ağırlıklı alan olan fen okuryazarlığına odaklanacağı anlaşılmaktadır. PISA'da her bir ülkedeki öğrenciler için hesaplanan fen okuryazarlık puanı ile öğrencilerin yalnızca fen alanındaki bilgileri değil, bu bilgileri nasıl kullandığı ve bu bilgileri gerçek hayatta ne derece uygulayabildiği de değerlendirilmektedir. Bunun yanında PISA'nın önemli bir hedefi, öğrenci, okul ve aile anketleri ile hem fen başarısını hem de fene karşı tutumlardaki kültürel farklılıkları daha iyi

anlamaktır (Timoçin, 2019). Diğer büyük ölçekli değerlendirmelerde olduğu gibi PISA 2015, öğrencilerin fen okuryazarlığı başarılarını değerlendirmek için olası değerleri (plausible values) kullanmaktadır. Geçmiş yıllarda yapılan PISA sınavlarında (PISA 2003, PISA 2006, PISA 2009, PISA 2012) beş olası değer kullanılmasına rağmen, PISA 2015’de on olası değer (plausible values) kullanılmıştır. Fen okuryazarlığı başarısına ek olarak, PISA 2015’de öğrencilerin fene yönelik duyuşsal özelliklerini ölçmek amacıyla fene yönelik öğrenci tutum anketleri de yer almaktadır.

Psikologlar uzun zamandır öğrencilerin olumlu tutumlarının ve kendi yeterliklerine ilişkin algılarının öğrenme motivasyonlarını olumlu yönde etkilediğine ve bu da başarılarının artmasına yol açtığına inanmaktadırlar (Bandura, 1994; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005). Bununla birlikte, büyük ölçekli uluslararası çalışmalardan elde edilen sonuçlar, sürekli olarak bu ilişki ile ilgili mevcut psikolojik modelleri zorlayan tuhaf bir paradoks göstermiştir (Min, Cortina, & Miller, 2016). Öğrencilerin akademik benlik kavramı, değer verme ve ilgi gibi tutumlarının başarıyı arttırdığı düşünülse de, matematik başarısında en üst sırada yer alan çoğu ülke akademik benlik düzeyi açısından en düşük puanlar arasında yer almıştır (Lee, 2009; Loveless, 2006). Bu durum PISA ve Uluslararası Okuma ve Okuryazarlık Çalışmasında İlerleme (PIRLS) gibi uluslararası karşılaştırmalı çalışmalarda da gözlenmiştir (Maaz, Trautwein, Lüdtke, & Baumert, 2008; Marsh & Hau, 2004, Van de gaer, Grisay, Schulz, & Gebhardt, 2012).

Fendeki başarı ile tutum arasındaki ilişkide, psikolojik modellerden farklı olarak kendine özgü bulgular elde edilmiştir. Özellikle, ülke içinde başarı ve tutum arasındaki ilişkiye bakıldığında öğrencilerin fene karşı olumlu tutumları arttıkça, fen başarılarının arttığı görülürken ülkeler arasında ise daha yüksek fen başarısı, fene karşı daha olumsuz tutum ile ilişkilidir. Yani ülke içinde fene yönelik tutum ile fen başarısı arasında doğru orantılı bir ilişki varken, ülkeler arasında fene yönelik tutum ile fen başarısı arasında ters orantılı bir ilişki bulunmaktadır. Bu durumda ülke içinde fene yönelik olumlu tutum artarken, fen başarısı da artmaktadır. Buna karşın ülkeler arasında ise fen yönelik olumlu tutum artarken, fen başarısı düşmektedir (Lu, 2012). PISA ve TIMSS gibi büyük ölçekli sınavlar üzerinden yapılan ikincil analizlerde de bur durum görülmüştür (Buckley, 2009; İlgün Dibek, 2018; Leventhal 2017; Lu,2012; Marsh & Hau, 2004; Shen & Tam, 2008).

PISA ve TIMSS gibi uluslararası düzeyde yapılan sınavlarda bireysel düzeyde öğrenci tutumları ve öğrenme ortamı (motivasyon ve öğretmen desteği) ile başarı arasında pozitif bir ilişki bulunurken ülke düzeyinde negatif bir ilişki bulunmaktadır. Yani ölçeklere verilen

puanlar bireysel bazda toplanıp tutum ve başarı arasındaki korelasyon hesaplandığında pozitif bir korelasyon katsayısı elde edilirken, ölçek puanları ülke düzeyinde toplandığında ve ülkelerin ortalama tutum ve başarı düzeyleri arasındaki korelasyon hesaplandığında negatif bir korelasyon katsayısı elde edilmektedir. Yapılan birçok çalışmada bu duruma benzer sonuçlar elde edilmiştir (Buckley, 2009; He ve Van de Vijver, 2015; İlgün Dibek, 2018; Kyllonen & Bertling, 2014; Loveless, 2006). Tutum ile başarı arasındaki ilişkinin ülke içinde pozitif, ülkeler arasında ise negatif olması durumu, tutum-başarı paradoksu olarak tanımlanmaktadır (Van de gaer vd., 2012).

Tutum-başarı paradoksunun oluşmasında birçok faktör rol oynamaktadır. Tutum-başarı paradoksunu meydana getiren etkenlerin başında, bireylerin maddeleri cevaplandırırken maddelere yükledikleri anlamları dikkate alarak cevaplandırmaları gelmektedir. Yani, bireyler madde kategorilerinde tercih yaparken madde içeriğinden bağımsız olarak herhangi bir faktörün sistematik etkisinin altında kalmaktadırlar. Bu nedenle bireylerin cevapları farklı faktörlerden etkilendiği için elde edilen sonuçlar yanlılık göstermektedir. Bireylerin madde kategori seçimlerini etkileyerek uluslararası düzeyde yanlı sonuçlar çıkmasına sebep olan bu faktörlerden biri “büyük balık küçük göl etkisi” dir (Marsh vd., 2008). Büyük balık-küçük göl etkisi, bir insanın kendini algılamasının mutlak olmaktan çok göreceli olması şeklindedir. Daha açıkça ifade etmek gerekirse; öğrenciler bir konudaki yetkinliklerini ortamlarındaki diğer öğrencilerle kıyaslama eğilimindedirler. Bu nedenle, ortalama düzeyde yetkinliği olan bir öğrenci başarısız bir sınıfa veya okula kayıtlıysa, kendisini “matematikte iyi” veya “fende iyi ”olarak kabul edecektir. Aynı öğrenci yüksek başarılı bir sınıfa veya okula kayıtlıysa, kendisini “matematikte kötü” veya “fende kötü” olarak değerlendirecektir. Öğrencinin kendisini diğer arkadaşlarından daha üstün olarak değerlendirmesi, kendisini olduğundan daha yüksek öz benliğe sahip biri olarak görmesine neden olacaktır. Ya d öğrencinin kendisini diğer arkadaşlarından daha düşük seviyede görmesi düşük öz benliğe sahip olmasına neden olacaktır. Öğrenciler üzerindeki bu baskılayıcı etki, muhtemelen geniş ölçekli sınavlara katılan katılımcı ülkelerin çoğunda benlik kavramı ölçeği puanlarında okullar arası veya sınıflar arası varyansın neden çok az gözlendiğini açıklayabilir (Kyllonen, Lietz, & Roberts, 2012).

Tutum-başarı paradoksunu meydana getiren etkenlerden biri de sosyal beğenirlik durumudur. Sosyal beğenirlik, bireysel özellikleri ölçen maddelere katılımcının doğru ve gerçekçi bir şekilde yanıt vermek yerine iyi görünmesini sağlayacak şekilde yanıt verme eğiliminde olmasıdır (Holtgraves, 2004, s. 161). Bu terim, bazı insanların (bilinçli ya da

bilinçsiz olarak) toplum içinde en “kabul edilebilir” olarak kabul edilen davranışları yansıtacak kategorileri seçme eğiliminde olduğu ve anlamlı derecede kabul edici davranışların olası yorumlarını içerir. Katılımcıların sosyal beğenirlik davranışı farklı kültürlerde farklı olabilir. Çünkü daha kolektivist kültürler (Japonya, Çin, Kore, Singapur) ara veya orta tepki kategorilerini seçme eğiliminde iken A.B.D gibi daha bireysel ülkelerde farklı olarak sosyal beğenirlik eğilimi bireylerin daha uç tepki kategorileri seçmelerini tetikleyebilir (Kyllonen, vd., 2012).

Tutum-başarı paradoksunun nedenlerinden bir diğeri ise ülkeler arasındaki tepki stili farklılıklarıdır (Buckley, 2009). Bu tez çalışmasında da fen başarısı ile tutum ölçekleri arasındaki beklenmedik ülke düzeyindeki ilişkinlerin ülkeler arasındaki tepki stili farklılıkları ile açıklanabileceği düşünülmektedir. Ayrıca kullanılacak yöntem ile tepki stilinin etkisinin ortadan kaldırılarak ülkelerin ölçülen içeriğe göre daha doğru bir sınıflandırmasının yapılması amaçlanmıştır.

Tepki stili, madde içeriğinden bağımsız olarak katılımcıların ölçek maddelerine sistematik bir şekilde tepki verme eğilimidir (Paulhus, 1991). Tepki stilleri, tepki stili yanlılığını meydana getirmektedir. Tepki stili yanlılığı, içerikten bağımsız olarak maddelerdeki belirli kategorileri seçme eğilimleri ile ilgilidir. İlk olarak Cronbach (1946) tarafından tartışılan tepki stilleri, ölçeklerin geçerliğini tehdit eden çeşitli faktörlerden biri olarak değerlendirilmektedir. Bu, farklı örneklemelerden alınan ölçek cevaplarının gerçek bir kültürel farklılığı yansıttığı ya da katılımcıların aynı içerik maddelerine aynı psikolojik yanıt vermesine rağmen ölçek maddelerini nasıl cevapladıkları arasındaki farkın açık olmadığı durumlarda göz önüne alınmaktadır. Sonuç olarak, tepki stilleri farklı kültürlerden gelen katılımcılar arasındaki önemli farklılığı gizleyebilir ya da gerçekte ilgili yapının özelliklerinde önemli farklılıklar olmadığında bu farklılıkları manipüle edebilir. Bu nedenle, kültürler arası yapılan karşılaştırma çalışmalarında, araştırmacılar tepki stili yanlılığı ile ilgilenmek zorundadır (Kemmelmeyer, 2016).

Çok kategorili ölçeklerin kullanıldığı araştırmalarda, tepki stillerinin başında ölçek maddelerindeki en uçlardaki kategorileri tercih etmek anlamına gelen Uç Tepki Stili (UTS) ve olumlu kategorileri tercih etme eğiliminde olmak anlamına gelen Kabullenici (Acquiescence) Tepki Stili (KTS) gelmektedir (Baumgartner & Steenkamp, 2001, Bolt & Johnson, 2009). UTS, adından da anlaşılacağı üzere tepki ölçeğinin her iki ucunda bulunan kategorileri tercih etme eğilimi iken, KTS ise diğer tepki kategorilerine ve seçeneklerine kıyasla daha çok ”katılıyorum” veya “evet” anlamındaki kategorileri tercih etme eğilimidir

(Greenleaf, 1992; Van Herk, Poortinga, & Verhallen, 2004). Bu iki tepki stilinin dışında Orta Nokta Tepki Stili (OTS) katılımcıların orta tepki kategorisini (kararsızım v.b) diğer kategorilerden daha sık seçme eğiliminde olması, Red Edici Tepki Stili (RTS) ise KTS'nin tam tersidir ve olumsuz ifadeleri temsil eden “kesinlikle katılmıyorum” ve “katılmıyorum” kategorilerini daha çok tercih etme eğilimi anlamına gelmektedir (Liu, 2015).

Tepki stili, bir tür ölçme hatasıdır ve rastgele hata olmaktan daha çok sistematik ölçme hatası olarak değerlendirilmektedir. Bir derecelendirme ölçeğinin ölçümünü iyileştirmek için ilk adım, tepki stili yanlılığının kaynağının iyi bir şekilde anlaşılmasıdır (Liu, 2015). Tepki stili eğilimleri bireysel düzeyde ele alınabileceği gibi kültürel düzeyde de ele alınabilmektedir. Bireysel düzeyde; kendini gösterme güduları, durağan kişilik özellikleri (Eid & Rauber, 2000), yaş (Meisenberg & Williams, 2008; Weijters, Geuens, & Schillewaert, 2010), cinsiyet (Weijters vd.,2010) ve eğitim düzeyi gibi özelliklerin tepki stilleri ile ilişkili olduğu bulunmuştur. Çeşitli çalışmaların sonuçları, eğitim düzeyi düşük olan kişiler arasında KTS'nin daha sık olduğunu göstermiştir (Narayan & Krosnick, 1996; Rammstedt, Goldberg, & Borg, 2010; Rammstedt & Kemper, 2011).

Çeşitli kültürel özellikler, ülkelere özgü tepki stillerini meydana getirebilmektedir (Buckley, 2009; Harzing, 2006; He & Van de Vijver, 2013; He & Van de Vijver, 2015; Rammstedt, Danner, & Bosnjak, 2017). Hofstede (2001)'nin boyutları olarak ifade edilen bazı kültürel özellikler, ülkelere özgü tepki stillerinin oluşmasına zemin hazırlamaktadır. Hofstede (2001)'e göre; kültürlere özgü toplulukçuluk, belirsizlikten kaçınma, kolektivizm, bireycilik ve güç mesafesi gibi özelliklerin bireylerin kategori tercihlerini etkilemekte ve kültürler arasında farklı tepki stillerin oluşmasına neden olmaktadır. Yapılan bazı çalışmalarda, tepki stilleri ile Hofstede (2001) boyutları arasındaki ilişkiler ele alınmıştır. Bireycilik, kendini ifade etme, bağımsızlık ve bireysel görüşlerle güçlü bir ilişki içindedir. Kolektivizm ise tevazu ve kişilerarası uyumla ilişkilidir. Bu nedenle, UTS'nin bireycilikle olumlu ilişki içinde olduğu (De Jong, Steenkamp, Fox, & Baumgartner, 2008; Smith & Fischer, 2008) ve KTS'nin ise bireycilik ile olumsuz ilişki içinde olduğu görülmüştür (Harzing, 2006; Johnson, Kulesa, Cho, & Shavitt, 2005).

Toplumlara özgü erkeklik ve kadınlık boyutlarından birinin baskın olması bireylerin tepki stillerini etkileyen başka bir özelliktir. Erkeklik boyutu yüksek olan toplumlar bir derecelendirme ölçeğinde en güçlü kategorilerin seçilmesini etkileyen atılganlığı ve kararlı davranışı vurgulamaktadır (Van Dijk, Datema, Welten, & Van de Vijver, 2009). Bundan dolayı toplumların erkeklik boyutu ile UTS pozitif korelasyon göstermektedir (De Jong vd.,

2008; Johnson vd., 2005). Kadınlık boyutu daha dominant olan toplumların ise nezaket, tevazu ve sosyal uyum özellikleri daha yüksektir. Bundan dolayı kadınlık boyutu daha yüksek olan toplumların KTS ile daha yüksek bir ilişkiye sahip olduğu düşünülmektedir (Van Dijk vd., 2009). Bir toplumdaki güçsüz olan bireylerin gücün eşit olarak dağıtılmadığını kabullenme derecesi olan güç mesafesi de, tepki stilleri ile ilişkili bir başka boyuttur. KTS'nin güç mesafesi ile pozitif korelasyon göstermesi beklenebilir (Harzing, 2006; Smith & Fischer, 2008). Kültürlere özgü “belirsizlikten kaçınma” boyutu kültürlerin toplumdaki belirsizliği tolere edebilme ve kendilerini rahat hissetme durumudur (Hosftede, 2001). Güçlü kültürlerin belirsizliğe ilişkin birçok kuralı ve hoşgörüsüzlüğü vardır. Bir derecelendirme ölçeğinin uç noktasının seçilmesi, açık ve net bir düşüncenin bir ifadesi olarak değerlendirildiği için UTS ve belirsizlikten kaçınma arasında pozitif bir ilişki bulunmaktadır (De Jong, Steenkamp, Fox, & Baumgartner, 2008).

Ülke düzeyindeki farklılıklar, daha büyük bölgelerin kültürel geleneklerini yansıtmaktadır ve diğer ulusal özelliklerle de bağlantılıdır. Örneğin, Kabullenici Tepki Stili (KTS), Latin Amerika ülkelerinde ABD ile karşılaştırıldığında daha yüksek görünmektedir (Harzing 2006; Ross & Mirowsky, 1984). Bazı araştırmalar KTS ve UTS' nin Akdeniz'de Kuzeybatı Avrupa'dan (Harzing, 2006; van Herk vd.,2004) daha yaygın olduğunu ve UTS'nin Kuzey Amerika'da Doğu Asya ülkelerinden (Chen, Lee, & Stevenson, 1995; Heine vd., 2001; Lee & Green, 1991; Takahashi, Ohara, Antonucci, & Akiyama, 2002) daha yaygın olduğunu göstermiştir. Tepki stilleri ile ilişkili ülke değişkenine başka bir örnekte uygunluktur. Uygunluk konusunda daha yüksek puan alan bireylerin ve ülkelerin anket sorularına verdikleri yanıtlarda tepki stillerine sahip olmaları daha olasıdır (Bond & Smith, 1996).

Bireylerin tepki stilleri, bazı tepki kategorilerinden kaçınarak veya sistematik olarak belirli kategorileri seçerek tepki ölçeğindeki sistematik bireysel farklılıklardan meydana gelmektedir. Tepki stilleri, tepki veren kişinin gerçek gizil yetenek düzeyinin üstünde veya altında madde yanıtları üzerinde bir etki oluşturursa, gizil yetenek tahminleri yanıltıcı olur. Sonuç olarak da bu yetenek tahminlerine dayalı olarak yapılan yorumlar yanıltıcı olur (Wetzel, Böhnke, & Rose, 2016). Bu durumda uluslararası sınavlardaki veriler kullanılarak yapılan çalışmalarda ülkelerin kendilerine özgü tepki stillerindeki farklılıklar dikkate alınmazsa ölçülen özellikteki farklılaşmalar hakkında yanlış yorumlamalar yapılabilir. Ayrıca, tutumlar ve kişilikler arasındaki kültürler arası çalışmalar genellikle likert ölçekli soruların yer aldığı anketlerdeki toplam ortalama puanların karşılaştırılmasına dayanmaktadır (Harzing 2006). Bu tür çalışmalarda, bireylerin vermiş oldukları yanıtların

sadece yanıtladıkları soruların içerik anlamına dayandığına dair temel bir varsayımda bulunulur. Bununla birlikte, bireylerin likert ölçeği maddelerine verdikleri cevaplar, çalışmalarda ortaya çıkan sonuçların yanlı olmasına neden olabilecek tepki stillerinden de etkilenmektedir.

Tepki stillerinin varlığı veri kümelerinin analizinde çeşitli psikometrik problemlere neden olabilmektedir. Bu psikometrik problemler test puan farklılıklarını geçersiz kılabilir, ilgi alanları arasındaki gerçek ilişkileri gizleyebilir, test güvenirliğini etkileyebilir ve grup düzeyinde karşılaştırma çalışmaların sonuçlarında yanlılığa sebep olabilir. Tepki stilleri, gözlenen madde puanlarını şişirerek veya düşürerek gerçek puanların değerlendirilmesini olumsuz şekilde etkileyebilir (Cho, 2013). Tepki stillerinin yanlılığından dolayı tutum puanlarına dayalı yapılan t-testi, F testi, kümeleme analizi, faktör analizi, regresyon analizi ve yapısal eşitlik modellemesi gibi çeşitli istatistiksel yöntemlerin sonuçları yanlı olabilmektedir (Reynolds & Smith, 2010). Ayrıca tepki stili yanlılığı güvenirliği de tehdit etmektedir. Liu, Wu ve Zumbo (2010) yaptıkları simülasyon çalışması ile tepki stilinin Cronbach alfa güvenirlik katsayısını ciddi şekilde etkilediğini ve aynı zamanda güvenirlik katsayısına ilişkin elde edilen standart hata değerini de yükselttiğini göstermişlerdir. Gözlemlerin tepki stillerinden etkilendiği verilere “tepki stili yanlı veriler” denmektedir. Tepki stili yanlı veriler üzerinden yapılan analizlerden elde edilen sonuçlar yanıltıcı olabilmektedir. Tepki stillerinden kaynaklanan yanlılıklar rastgele hata yerine “sistemik hata” olarak değerlendirilmektedir. Bu nedenle anlamlı bir veri analizi yapmak için bu tür sistemik hatalar dikkate alınmalıdır (Baumgartner & Steenkamp, 2001).

Kümeleme analizi, katılımcıların benzer gözlem gruplarına ayrıldığı analiz türüdür. Kümeleme analizinde katılımcılar ölçme aracında yer alan maddelere ilişkin tercihlerine göre kümelenmektedir. Bireylerin tercihlerini ölçmek için genellikle likert ölçeği gibi bir derecelendirme ölçeği ile önceden belirlenmiş bir dizi tercih kategorisinden oluşan ölçekler kullanılmaktadır. Bireylerin cevaplarına göre kümelenmesi, gizil yapıları tanımlamak için yararlı olmaktadır (Takagishi, van de Velden, & Yadohisha, 2019). Uygulamada veriler tepki stili yanlılığı içeriyorsa kümeleme analizi, ölçülen özellik açısından benzerlik gösteren (içerik tabanlı kümeler) kümeler yerine, benzer tepki stillerine (tepki stili tabanlı kümeler) sahip katılımcı kümeleri verebilir. Örneğin, bir ölçekte bir katılımcı grubu orta nokta kategorilerini seçme eğilimindeyken, başka bir grubun tercihlerine bakılmaksızın uç nokta kategorilerini tercih etme eğiliminde olduğunu varsayalım. Elde edilen verilere küme analizi uygulamak, orta nokta ve son nokta kategorilerini seçmiş katılımcıların kümelerini

çıkabilir. Ancak, bu kümeler yalnızca tepki stillerini yansıtır ve verilerdeki herhangi bir içerik tabanlı yapı tespit edilemez (Takagishi vd., 2019).

Özellikle çok sayıda ülkenin katıldığı geniş ölçekli sınavlarda, kültürel farklılıklar nedeniyle meydana gelen tepki stili yanlılığı yapılacak ikincil analizlerden elde edilen sonuçların geçerliğini olumsuz şekilde etkilemektedir. PISA ve TIMSS gibi sınavlarda yapılacak ikincil analizlerdeki sonuçların geçerliğine kanıt sağlanması ihtiyacı nedeni ile tepki stili yanlılığının düzeltilmesi önemli görülmektedir. Bu doğrultuda tutum-başarı paradoksuna neden olan tepki stili yanlılığının düzeltilmesinde ve verilerin içerik tabanlı kümelere ayrılmasında yeni bir analiz yöntemi olan Tepki Stili Yanlılığının Düzeltilmesi ve Kümelenmesi (TSYDK)-Correcting and Clustering Response Style Biased Data (CCRS)-modelinin eğitim verilerinde kullanılması, bu çalışmanın gerekçesini oluşturmaktadır.

Araştırmanın Amacı

Çalışmanın genel amacı, PISA 2015'e katılan ülkelerde tutum-başarı paradoksuna neden olan tepki stili etkisinin düzeltilmesi ve ülkelerin düzeltilmiş puanlara göre kümelenmesidir. Ayrıca çalışmada Tepki Stili Yanlılığının Düzeltilmesi ve Kümelenmesi (TSYDK) yönteminin etkililiğini incelemek amacıyla, TSYDK ile K-ortalamlar kümeleme yönteminin öğrencileri ölçülen içeriğe göre kümeleme performansları da karşılaştırılmıştır.

Araştırmanın Önemi

Tepki stilleri potansiyel olarak derecelendirme ölçeklerini kullanan ve sonuçları için alternatif açıklamaları olan tüm deneysel çalışmaları etkilemektedir. Ancak tepki stillerinin önemine rağmen birçok araştırmacı bu yanlılık kaynağını kontrol etmeden çalışmalarını yürütmektedirler (Vaerenbergh & Thomas, 2012). Tepki stilleri sadece ölçümlerin tek değişkenli dağılım özelliklerinde (Örn. ortalama, varyans, çarpıklık v.b) değil, aynı zamanda değişkenler arasındaki çok değişkenli ilişkiler üzerinde de olumsuz etkilere neden olmaktadır. Yapılan birçok deneysel araştırmalardan elde edilen sonuçlarda bu durum gözlenmiştir (Baumgartner & Steenkamp, 2001; Moors, 2008).

Tepki stilleri farklı bireysel özelliklerin sonucu olarak ortaya çıkmaktadır (Hamilton, 1968). Ayrıca tepki stili eğilimleri ülke bazında incelendiğinde; tepki stillerinin ülkeler arasındaki farklılaşmasından dolayı ham puanlar aracılığıyla doğrudan kültürler arası karşılaştırmalar

yapmak elde edilen sonuçların geçerliğini olumsuz şekilde etkileyebilir (Hui & Triandis, 1989). Van Herk vd. (2004) tarafından yapılan ve altı farklı Avrupa ülkesinin tepki stilini incelendikleri araştırmada; UTS eğilimi açısından ülkeler arasında önemli farklılıklar bulunmuş ve UTS'deki ulusal düzeydeki farklılıkları görmezden gelmenin kültürler arası araştırmalarda geçersiz sonuçlara yol açabileceği sonucuna varılmıştır. Eğitim alanında, PISA gibi geniş ölçekli ve farklı kültürlerden çok sayıda ülkenin katıldığı sınavlardan elde edilen sonuçların hem ulusal düzeyde hem de uluslararası düzeyde çok önemli olduğu ve ülkelerin eğitim politikalarına yön verdiği göz önüne alındığında, bu sonuçların geçerli ve doğru olması önemli görülmektedir. Bu doğrultuda tepki stili yanlılığının kontrol edilerek düzeltilmesini sağlayan bu araştırma, PISA 2015 verileri kullanılarak yapılacak ikincil analizlerin sonuçlarının geçerliğinin sağlanması açısından önemli görülmektedir.

UTS, sadece cevap verenin bir ölçek puanının, “gerçek” bir içerik farklılığını yansıtmaksızın, ölçeğin uç kategorilere doğru “yönlendirilmesi” anlamına gelir. Farklı kültürler farklı tepki stilleri sergiledikleri için, bu tarz faktörleri göz ardı etmek yanlış karşılaştırmalara yol açabilir. Dolayısıyla UTS, karşılaştırmalı çalışmalarda olası ölçme eşitsizliğinin nedenlerinden biri olarak karşımıza çıkar (Kankaraş & Moors, 2011). Bu doğrultuda PISA verisi kullanılarak yapılan ülkeler arası karşılaştırma çalışmalarının sonuçlarının geçerliğinin ön koşulu olarak ölçme değişmezliğinin sağlanması gerekmektedir. Dolayısıyla bu araştırmada UTS etkisinin düzeltilmesi ölçme değişmezliğinin sağlanmasında ve PISA verisini kullanacak araştırmacılar için sonuçlarının geçerliliklerini sağlamada faydalı olacaktır. Ayrıca bu araştırmada kullanılan TSYDK ile diğer veri analizi yöntemleri birlikte kullanılabilir.

Tepki stillerinin incelenmesinde betimsel istatistiksel yaklaşımlarının (Ör. Standart sapma, varyans) yanı sıra model temelli yöntemlerde (Ör. Doğrulayıcı Faktör Analizi, Örtük Sınıf Analizi, Madde Tepki Kuramı) bulunmaktadır. Ancak tepki stili yanlılığı bulunan verilerde madde puanlarının düzeltilmesinde birçok araştırmada sadece Çok Boyutlu Madde Tepki Kuramı yöntemi kullanılmaktadır (İlgün Dibek, 2018; Ju & Falk, 2019; Liu, 2015). Bu nedenle bu araştırmada, tepki stili yanlılığının düzeltilmesinde Çok Boyutlu Madde Tepki Kuramı'ndan farklı bir model olan TSYDK yönteminin kullanılması önemli görülmektedir.

Araştırma kapsamında kullanılan TSYDK yöntemi, tepki stili yanlılığı bulunan verileri düzeltmenin yanında düzeltilmiş verileri eş zamanlı olarak kümeleme olanağı da sağlamaktadır. TSYDK yöntemi ik olarak bireylere özgü tepki fonksiyonlarını kullanarak her bir birey için 0 ile 1 arasında yeniden oluşturulmuş madde puanları elde edilmesini

sağlamaktadır. Daha sonra elde edilen bu madde puanları üzerinden K-ortalamlar kümeleme analizi gerçekleştirilerek katılımcıların kümelenmesini sağlamaktadır. PISA verisi gibi büyük veri setlerinde genellikle kümeleme analizi gibi yöntemler uygulanarak benzer özelliklere sahip alt gruplar oluşturulmaktadır. Ayrıca kümeleme analizi, yeni bir gözlemin birçok kategori arasından doğru kategoriye yerleştirilmesini sağlamaktadır. Bu doğrultuda kümeleme analizinin sonuçlarının daha doğru şekilde elde edilmesi ve gerçek sonuçlara ulaşılması önemli hale gelmektedir. Bu nedenle tepki stili yanlılığı bulunan verilerin düzeltilerek eş zamanlı olarak bireylerin ölçülen özellik açısından kümelemesi, TSYDK yöntemini önemli kılmaktadır. TSYDK yöntemi ile tepki stili yanlılığı ortadan kaldırılarak içerik tabanlı bir sınıflandırma yapılabilmektedir. Yani bireyler ölçme aracı ile ölçülmek istenen özelliklerine göre gerçek veri kümelerine ayrılabilir. Bu açıdan bakıldığında PISA ve TIMSS gibi büyük veri setlerini kullanacak olan diğer araştırmacılar için TSYDK yöntemine yer vermeleri çalışmalarının geçerliliklerine kanıt sağlamaları açısından önemli görülmektedir.

Problem İfadesi

PISA 2015 uygulamasına katılan ve tutum-başarı paradoksunun (fene yönelik tutum ile fen başarısı arasında ülke içi pozitif, ülkeler arası negatif olduğu durum) görüldüğü ülkelerde tepki stili yanlılığının düzeltilmesi ve bireylerin ölçülen özellik açısından kümelenmesi.

Araştırma Soruları

Araştırmanın problem ifadesine göre aşağıdaki sorulara yanıt aranmıştır:

1. PISA 2015 uygulamasına katılan ve tutum-başarı paradoksunun görüldüğü ülkeler TSYDK yöntemi ile nasıl kümelenmektedir?
2. PISA 2015 uygulamasına katılan ve tutum-başarı paradoksunun görüldüğü ülkeler K-ortalamlar yöntemi ile nasıl kümelenmektedir?
3. TSYDK ile K-ortalamlar yöntemlerinin ülkeleri kümeleme performanslarının karşılaştırması nasıl sonuçlar vermektedir?
4. PISA 2015 uygulamasına katılan ve tutum-başarı paradoksunun görüldüğü ülkelerde
 - a) Tepki stili yanlılığı bulunan ve TSYDK yöntemi ile düzeltilmiş ortalama tutum puanları arasındaki farklar manidar mıdır?

- b) Tepki stili yanlılığı bulunan ve TSYDK yöntemi ile düzeltilmiş ortalama tutum puanlarının fen başarısı ile arasındaki ülke içi ve ülkeler arası ilişkiler nasıl değişmektedir?

Sınırlılıklar

1. PISA 2015 uygulamasında öğrenci anketlerinde yer alan maddelerde dört kategori bulunduğu için bu çalışma UTS'nin, KTS'nin ve RETS'nin modellenmesi ile sınırlıdır.
2. Araştırma, öğrencilerin fene yönelik tutumlarını ölçmek amacıyla PISA öğrenci anketlerindeki maddelerin ağırlıklandırılmış örneklem ortalamaları ile sınırlıdır.
3. Araştırma, öğrencilerin fen başarılarını belirlemek amacıyla PISA 2015 fen başarı testindeki on olası (plausible) değerin ağırlıklandırılmış örneklem ortalaması ile sınırlıdır.
4. Araştırmada, analizlerin yapıldığı bilgisayarın işlem yapabilme kapasitesi göz önüne alınarak veri sayısının fazla olması nedeniyle kayıp veriler silinmiştir. Bu nedenle araştırma, kayıp veriler silindikten sonra her bir ülkedeki örneklemde %10'luk rastgele örneklem seçme yöntemi yoluyla seçilmiş çalışma örneklemi ile sınırlıdır.

BÖLÜM II

KURAMSAL ÇERÇEVE

Bu bölümde araştırmanın kuramsal temellerine ve çalışma konusu ile ilgili yapılmış olan diğer araştırmalara yer verilmiştir.

Tepki Stili Yanlılığının Düzeltilmesi ve Kümelenmesi (TSYDK)-Correcting and Clustering Response-Style-Biased Data-(CCRS)

Araştırma kapsamında kullanılan ve tepki stili yanlılığı bulunan verilerin düzeltilip eş zamanlı olarak kümelenmesini sağlayan TSYDK yönteminin matematiksel çerçevesi aşağıda sunulmuştur. TSYDK yönteminde birinci adımda tepki stili yanlılığı bulunan veriler düzeltilmektedir. İkinci adımda tepki stili yanlılığının düzeltildiği verilerin içerik tabanlı kümelenmesi gerçekleştirilmektedir. TSYDK yönteminin avantajı bu adımları sırayla gerçekleştirmek yerine katılımcılara özgü tepki fonksiyonları ile tepki stillerinin aynı anda düzeltilmesi ve düzeltilmiş verilerin içerik tabanlı kümelenmesini sağlamasıdır.

Tepki Fonksiyonlarının Formüle Edilmesi

TSYDK yöntemini tanımlamak için, ilk olarak ‘tepki fonksiyonu’ kavramını formüle etmek gerekmektedir. Tepki stillerinin ve düzeltilmiş değerlerin formüle edilmesi van de Velden (2007), Schoonees, Velden ve Gronen (2015) ile Takagishi vd. (2019) tarafından yapılan tanımlama çerçevesine uygun şekilde bu çalışmada da yapılmıştır. Bu tanımlama, daha genel olarak tepki stili etkileri nedeniyle bozulmuş tercih kategorilerinden oluşan verilerin incelenmesinde kullanılmaktadır.

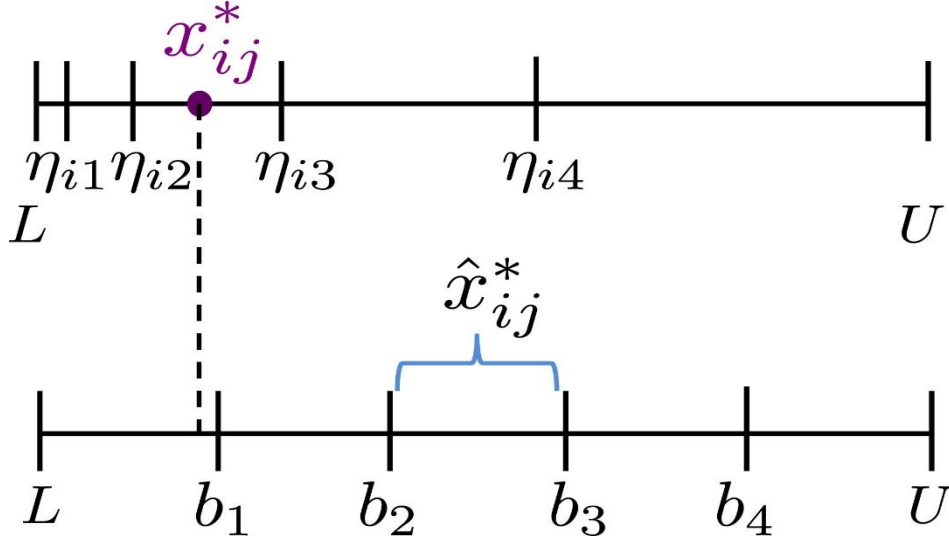
Tercih verilerindeki kategori sınırları

Tepki stili yanlılığı, tercih kategorilerinin yorumlanmasının farklı katılımcılar için farklı olduğu durumlarda ortaya çıkmaktadır. Örnek vermek gerekirse; *1-Kesinlikle Katılmıyorum..... 5-Kesinlikle Katılıyorum* şeklinde beş puanlı bir derecelendirme ölçeğinin kullanıldığı araştırmada bir katılımcının KTS eğilimi varsa, yani madde içeriğinden bağımsız olarak çoğunlukla olumlu olan en yüksek kategorileri seçme eğilimi varsa, üçüncü kategori-bu kategori ölçeğin orta noktası- olmasına rağmen yanıt verenin bu madde için düşük bir tercihi olduğunu gösterir.

Formül olarak yukarıda açıklanan durumu ifade etmek amacıyla q kategorili tercih ölçeğinde i. katılımcı tarafından j. madde ($i=1,2,\dots,n;j=1,\dots,m$) için tercih edilen kategori $x_{ij} \in \{1,2,\dots,q\}$ şeklinde gösterilsin. Gözlenen tercih verisi x_{ij} 'nin aşağıdaki gerçek tercih verisi $x_{ij}^* \in R$ ile ilgili olduğunu varsayalım:

$$x_{ij} = \sum_{l=1}^q l \cdot I\{\eta_{i(l-1)} < x_{ij}^* \leq \eta_{il}\} \quad (\text{Denklem 1})$$

Burada $I\{\cdot\}$ bir gösterge fonksiyonudur ve $\eta_{il} (l = 0, \dots, q)$ katılımcıya özgü sınırlardır. Tüm katılımcılar için eşit olan ve eşit aralıklarla yerleştirilmiş olan $b_{il} (l = 0, \dots, q)$ sınır kümesi referans sınırları olarak adlandırılmaktadır. Bu çalışmada sınırlı bir aralık, yani $\eta_{i0} = b_0 = L$ ve $\eta_{iq} = b_q = U$ olarak düşünülmektedir. Bu gösterimle, 'tepki stili eğilimli veriler', her bir katılımcı için katılımcıya özgü sınırlar η_{il} olmasına rağmen, x_{ij}^* gerçek tercihlerinin eşit aralıklı referans sınırlarının b 'ye göre sınıflandırıldığı verilerdir. Bu işlem Şekil 1'de gösterilmektedir.



Şekil 1. Tepki stili yanlılığı. Takagishi, M., van de Velden, M., & Yadohisa, H. (2019). Clustering preference data in the presence of response-style bias. *British Journal of Mathematical and Statistical Psychology*, 72(3), 401-425.

Şekil 1’de üst ölçekte $[l, u] \in r$, katılımcıya özgü sınırlar gösterilirken, alt ölçekte eşit aralıklı referans sınırları gösterilmekte ve x_{ij}^* gerçek tercih kategorisini göstermektedir. Bu durumda, $x_{ij}=3$ olduğu zaman $x_{ij}^* \in (b_2, b_3]$, referans sınırları b_l ile x_{ij}^* bir ölçek tahminidir. Üst ölçekteki $\eta_{il} (l = 1, \dots, q - 1)$ kümesi, dördüncü ve beşinci kategorilerin seçilme olasılığının daha yüksek olduğu bir tepki stilini temsil eder (Takagishi vd., 2019, s. 4).

Şekil 1’de, katılımcı i’nin gerçek tercihi x_{ij}^* ve sınırları $\eta_{il} (l = 0, \dots, q)$ üst ölçekte gösterildiği gibi olsun. Burada amaç, x_{ij} ’den x_{ij}^* tahmin etmektir. Bu örnekte, gözlenen tercih kategorisi $x_{ij}=3$ ’tür. Her katılımcının farklı sınırları olması olasılığını göz ardı edersek ve basitçe referans sınırlarının Şekil 1’deki alt ölçekte gösterildiği gibi kullanıldığını varsayarsak, x_{ij}^* ’nin kaba bir tahmini, $\hat{x}_{ij}^*$, gerçek x_{ij}^* olmaktan çok uzak olacaktır. Bu, gözlenemeyen katılımcıya özgü sınırlara bağlı olarak, gerçek x_{ij}^* ’den bir sapma elde ettiğimizi gösterir (Takagishi vd., 2019).

Tepki Fonksiyonları

Tepki stili yanlılığı içeren verileri düzeltmek için bir tepki fonksiyonu aşağıdaki gibi tanımlanabilir.

Tanım 1: (Tepki Fonksiyonu): Referans sınırları $b_l (l = 1, \dots, q - 1)$ ve katılımcıya özgü sınırlar $\eta_{il} (l = 1, \dots, q - 1)$ verildiğini varsayalım. Her iki sınırında l için monoton

olarak artmasına izin verilsin. Daha sonra, i. yanıtlayıcı için tepki fonksiyonu aşağıdaki gibi tanımlanır:

$$\phi_i: b_l \longmapsto \eta_{il} \ (l = 1, \dots, q - 1) \quad (\text{Denklem 2})$$

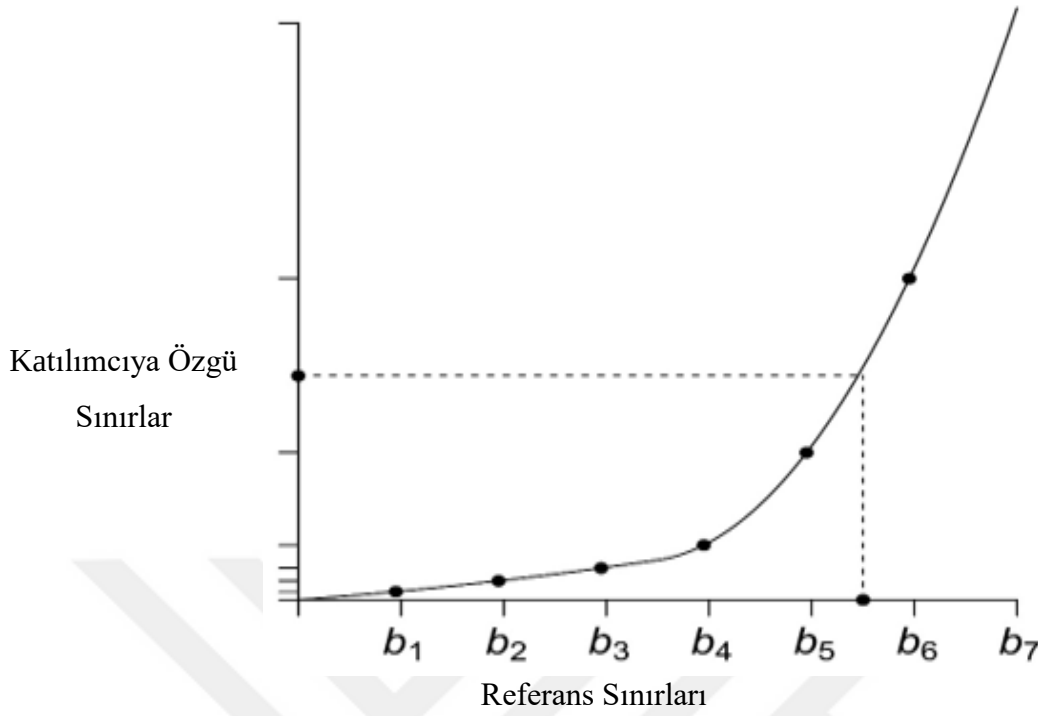
Tepki fonksiyonun tanımından yola çıkarak, ϕ_i 'nin monotonik olarak artan bir fonksiyon olduğu anlaşılmaktadır. Ayrıca, tepki fonksiyonun sürekli olduğu varsayılmaktadır. Çalışmanın daha sonraki amaçları için tepki stilinin bulunmaması durumunda tepki fonksiyonunu özel olarak tanımlamak yararlı olacaktır.

Tanım 2. (Tepki Stili Yok): Tepki stili bulunmayan tepki fonksiyonu, $\eta_{il} = b_l \ (l = 1, \dots, q - 1)$ eşitliğinin sağlanması durumudur ve i. yanıtlayıcının tepki stili eğilimi olmadığı şeklinde yorum yapılabilir. Eğer, ϕ_i tüm katılımcılar için biliniyorsa, tepki stili yanlılığı içeren verileri düzeltmek ve katılımcıların tepki stillerini yorumlamak için bu formülasyonu kullanabiliriz.

Tanım 3. Tepki Fonksiyonlarını Kullanarak Tercih Verilerini Düzeltmek: Referans sınırları $b_1, \dots, b_{q-1}$ olan q puanlı x_{ij} verisi ve ϕ_i tepki fonksiyonları verildiğinde; $x_{ij} = l$ olduğu zaman, x_{ij} 'nin düzeltilmiş değeri aşağıdaki gibi tanımlanmaktadır.

$$y_{ij} = \phi_i(\tau(l)), \ \tau(l) \in (b_{l-1}, b_l] \quad (\text{Denklem 3})$$

Bu tanım, x_{ij} 'nin düzeltilmiş değerinin, ϕ_i 'ye göre, $\tau(l)$, (b_l) ve (b_{l-1}) 'nin bazı değerleri arasında bir değer dönüştürülmesinin ürünü olarak tanımlandığını gösterir. Burada $\tau(l) = (b_l + b_{l-1})/2$ şeklinde düzeltilir. Tepki stili yanlılığını düzeltmek için, bir tepki stilinin nasıl kullanıldığı Şekil 2'de gösterilmektedir.



Şekil 2. Gözlenen değerin $x_{ij} = 6$ 'nın düzeltilmiş değere nasıl karşılık geldiğini gösteren örnek. Takagishi, M., van de Velden, M., & Yadohisa, H. (2019). Clustering preference data in the presence of response-style bias. *British Journal of Mathematical and Statistical Psychology*, 72(3), 401-425.

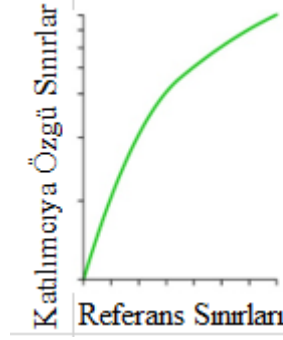
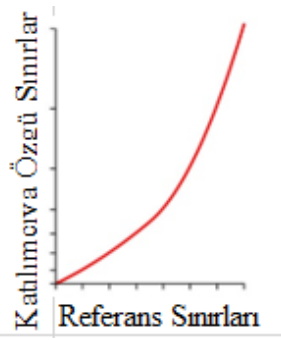
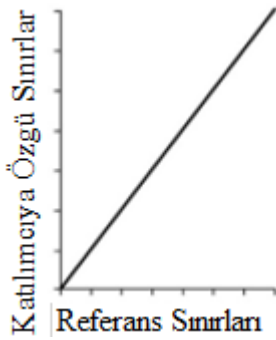
Şekil 2’de düz çizgi, tepki fonksiyonu ϕ_i ’yi göstermektedir. Yatay eksen referans sınırı (ölçeği) temsil ederken, dikey eksen katılımcıya özgü sınırı temsil etmektedir. (Takagishi vd., 2019, s. 5). Şekil 2’deki örneğe göre yedi puanlı ölçekte gözlenen değerin $x_{ij} = 6$ olduğunda x_{ij}^* ’ı bilmek istediğimizi varsayalım. Bu durumda, ϕ_i ’nin argümanı, $(b_5, b_6]$, aralığında herhangi bir değer olabilir. Denklem 3’e dayanarak, aralığın orta noktası kullanılır ve kategori 6’nın temsili değeri olarak adlandırılır. Eğer, $b_l = l (l = 1, \dots, q - 1)$ belirlenirse 5.5 (örneğin, Şekil 2’de yatay eksen üzerindeki puan) ϕ_i ’nin argümanı olacaktır. Gerçek tepki fonksiyonunun sürekli olduğu varsayılarak, kategorinin temsili değerine karşılık gelen tepki fonksiyonunun sonuç değeri (yani, Şekil 2’deki dikey eksen üzerindeki nokta), dikey eksenden okunabilir. Ortaya çıkan bu sonuç değeri (y_{ij}), bu durumda düzeltilmiş değerdir.

Tepki fonksiyonları, katılımcıların tepki stillerini yorumlamak için kullanılabilir. Şekil 3’te tepki stili olmayan, KTS, RETS, UTS ve OTS’ne sahip katılımcılara karşılık gelen tipik tepki fonksiyon örnekleri gösterilmektedir.

Tepki Stili Yok

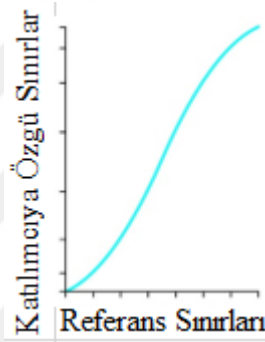
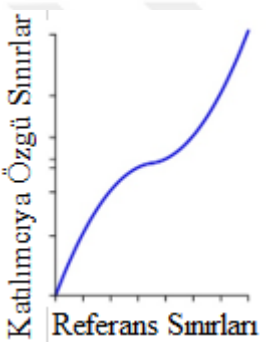
Kabullenici Tepki Stili (KTS)

Red Edici Tepki Stili (RETS)



Uç Tepki Stili (UTS)

Orta Nokta Tepki Stili (OTS)



Şekil 3. Tepki fonksiyonları. Takagishi, M., van de Velden, M., & Yadohisa, H. (2019). Clustering preference data in the presence of response-style bias. *British Journal of Mathematical and Statistical Psychology*, 72(3), 401-425.

Şekil 3'te yatay eksen referans sınırını (ölçeği) temsil ederken, dikey eksen yanıtlayana özgü sınırı temsil etmektedir (Takagishi vd., 2019, s. 6).

Tepki Stili Yanlılığın Varlığında Tercih Verilerinin Düzeltilmesi ve Kümelenmesi

Tepki stili yanlılığı bulunan verilerde yukarıda tanımlanan katılımcıya özgü tepki fonksiyonlarının tahminleri kullanılarak düzeltme yapılır. Ayrıca tahmini tepki fonksiyonları, tepki stili yanlılığını düzeltmek ve aynı zamanda katılımcıların kümelerini düzeltilmiş madde tercihlerine göre bulmak içinde kullanılabilmektedir (Takagishi vd., 2019).

Tepki Fonksiyonlarının Modellenmesi

Bir tepki fonksiyonunu tahmin etmek için katılımcıya özgü sınırları temsil eden veriler gereklidir. Daha sonra bu veriler TSYDK yönteminde 'sıralamalı sınır verileri' olarak kodlanır. Bu durum, belirtilen madde tercihlerinin ve referans sınırlarının her bir katılımcı için sıralama düzenine dönüştürüldüğü anlamına gelir. Elde edilen sınır sıralamaları, katılımcıların belirli derecelendirme kategorilerini seçme eğilimlerini yansıtır.

Örnek vererek açıklamak gerekirse, q-puanlı ölçek tercih verileri $X = (x_{ij})$ ($i = 1, \dots, n; j = 1, \dots, m$) şeklinde ve referans sınırları $b_1, b_2, \dots, b_{q-1}$ şeklinde olsun. Ardından sıralı sınır verisi $f_{il} = (l, \dots, q - 1)$ aşağıdaki gibi elde edilebilir:

$$f_{ij} = \sum_{l=1}^{m+q-1} (I\{c_{it} < b_l\} + \frac{1}{2}I\{c_{it} = b_l\}) - \frac{1}{2} \quad (\text{Denklem 4})$$

Burada, $c_{it} = \begin{cases} \frac{b_l + b_{l-1}}{2} & (t = 1, \dots, m, \dots, x_{it} = l) \\ b_{t-q+1} & (t = m + 1, \dots, m + q - 1) \end{cases}$ şeklinde olur.

Ayrıca $t = 1, \dots, m, \dots, c_{it}$ bir kategorinin temsili değerlerini gösterir. Bu durumda, bu değerler $(b_l + b_{l-1})/2$ olur. Bir diğer yandan, $t = m + 1, \dots, m + q - 1$, c_{it} referans sınırlarını gösterir.

Formülasyonun pratikte nasıl çalıştığını göstermek için, yedi kategorili ölçek tercih verilerinde, $x_i = (5, 6, 7)$ olarak verilsin. Denklem 4 kullanılarak, $c_i = (c_{it}) (t = 1, \dots, m + q - 1)$ olduğunda $c_i = (4, 5; 5, 5; 6, 5; 1, 2, 3, 4, 5, 6)$ elde edilir. Daha sonra, bunları sıralama düzenlerine dönüştürünce (0'dan başlayarak) aşağıdaki durum elde edilir.

$$c_i^{\text{sıralı}} = (1 \ 2 \ 3 \ 4 \ 4,5 \ 5 \ 5,5 \ 6 \ 6,5)$$

$$\text{Sıra} = (0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8)$$

$c_{i4} = 1, c_{i5} = 2, c_{i6} = 3, \dots, c_{i9} = 6$ 'dan beri sırasıyla 0, 1, 2, 3, 5, 7 seviyelerine karşılık gelir, $f_i = (0, 1, 2, 3, 5, 7)$ elde edilir. Şekil 4 (solda) bu referans sınırlarını $f_i = (f_{i1}) (l = 1, \dots, 6)$ karşı çizer. Bu dönüştürülmüş f_i 'yi kullanarak, i yanıtlayıcısının bir kabullenme tepki stili gösterdiğini görüyoruz. Örneğin, $f_{i1}, \dots, f_{i4}$, değerleri birer birer artar, bu da katılımcı i'nin ilk ve dördüncü referans sınırları arasındaki kategorileri sıklıkla seçmediğini gösterir (yani, katılımcı genellikle 4'ten küçük bir puan almaz). Öte yandan,

dördüncü ve altıncı referans sınırları arasındaki kategorilerin sıklıkla seçildiğini gösteren f_{i4} ve f_{i6} arasında büyük bir boşluk vardır. f_{il} 'yi ve I-spline temel fonksiyonlarını kullanarak tepki fonksiyonları için bir model oluşturulabilir. $\bar{f}_{il} = f_{il}/p$, $p = m + q - 1$ olsun. Böylece, $\bar{f}_{il} \in [0,1]$. Ayrıca, bundan sonra $b_l = l/q$ ($l = 1, \dots, q - 1$) şeklinde kullanılmaktadır. TSYDK' de, $\bar{f}_{il} = f_{il}$ yaklaşık olarak tahmin edilir.

$$\bar{f}_{il} \approx \phi_i^{TSDK}(l/q) \quad (i = 1, \dots, n; l = 1, \dots, q - 1)$$

$$\phi_i^{TSDK}(x) = \sum_{r=1}^3 \beta_{ir} M_r(x) \quad (\text{Denklem 5})$$

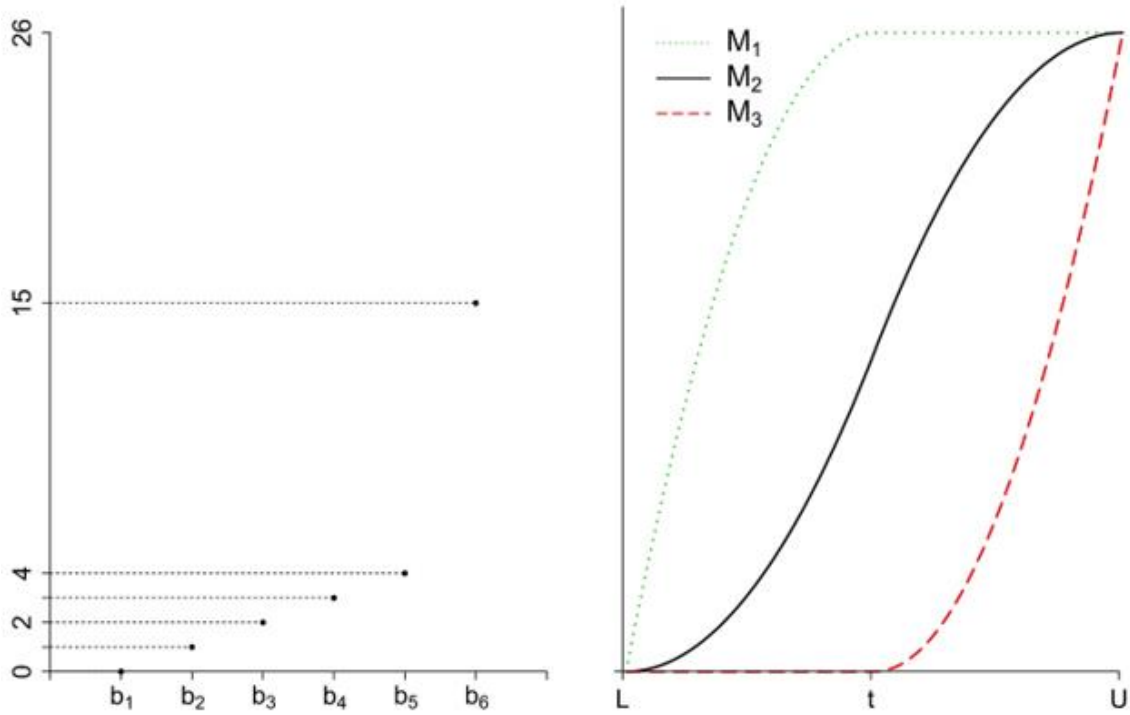
$$\text{s. t. } \sum_{r=1}^3 \beta_{ir} = 1, \beta_{ir} \geq 0 \quad (r=1,2,3).$$

$$M_1(x) = \begin{cases} \frac{2t(x-L) - (x^2 - L^2)}{(t-L)^2} & (L \leq x \leq t) \\ 1 & (t \leq x \leq U). \end{cases}$$

$$M_2(x) = \begin{cases} \frac{(x-L)^2}{(t-L)(U-L)} & (L \leq x \leq t) \\ \frac{t-L}{U-L} + \frac{2U(x-L) - (x^2 - t^2)}{(U-t)(U-L)} & (t \leq x \leq U) \end{cases}$$

$$M_3(x) = \begin{cases} 0 & (L \leq x \leq t) \\ \frac{(x-t)^2}{(U-t)^2} & (t \leq x \leq U) \end{cases}$$

$x \in [L, U]$, $t = L + 0.5(U - L)$. Burada M_r ($r=1,2,3$) I-spline temel fonksiyonlarıdır ve β_{i1} , β_{i2} , β_{i3} sırasıyla M_1, M_2, M_3 'ün katsayılarıdır (Ramsay, 1988). TSYDK'de $L = 0, U = 1$ kullanılır. ϕ_{ir} 'nin monoton arttırıcı bir fonksiyon olması için negatif olmayan koşullar, $\beta_{ir} \geq 0$ ($r=1,2,3$) gereklidir.



Şekil 4. (Sol) Sıralı sınır verilerinin bir örneği. (Sağ) Üç I-spline temel fonksiyonlar. Takagishi, M., van de Velden, M., & Yadohisa, H. (2019). Clustering preference data in the presence of response-style bias. *British Journal of Mathematical and Statistical Psychology*, 72(3), 401-425.

Şekil 4’de yatay eksen referans sınırlara karşılık gelirken, dikey eksen her sınıra karşılık gelen f_{il} değerlerini gösterir. Her nokta $f_{i1}, \dots, f_{i6}$ y1 temsil eder. (Sağ). M_1, M_2, M_3 sırasıyla düz, noktalı ve kesikli çizgiler olarak gösterilmiştir (Takagishi vd., 2019, s. 8).

Üç I-spline temel fonksiyonunu kullanarak (Şekil 4’te (sağda) gösterildiği gibi), Şekil 3’te gösterilen beş farklı tepki stilini elde edebiliriz. Ayrıca, bu modelde tepki stillerini yorumlamak için sadece $\beta_{i1}, \beta_{i2}, \beta_{i3}$ ’ün dikkate alınması gerekir. Örneğin, daha büyük bir β_{i3} değeri, yüksek kategorileri seçme konusunda daha güçlü bir eğilimi gösterir, çünkü M_3 üzerine daha fazla ağırlık yerleştirilmesine neden olur, bu da fonksiyonun şeklini, kabullenme tepki stiline karşılık gelen tepki fonksiyonuna şekline daha benzer olacak şekilde değiştirir (Şekil 3).

Denklem 5’de tanımlanan modeli kullanarak, tepki fonksiyonu sınırlandırılmış en küçük kareler yöntemi ile düzeltme (smoothing) tahmin edilebilir. Başka bir deyişle, bir $q \times 1$ vektörü $\bar{f}_i = (\bar{f}_{il})$ ve $(q - 1) \times 3$ matrisi $M = \left(M_r \left(\frac{l}{q} \right) \right) (l = 1, \dots, q - 1; r = 1, 2, 3), \beta_i$ minimize edilerek elde edilir.

$$\sum_{i=1}^n \| \bar{f}_i - M\beta_i \|^2, \quad \text{s.t.} \sum_{r=1}^3 \beta_{ir} = 1, \quad \beta_{ir} \geq 0, \quad (\text{Denklem 6})$$

Burada, $\beta_i = (\beta_{i1}, \beta_{i2}, \beta_{i3})'$ dir. Tahmin edilen $\hat{\beta}_i$ değerini kullanarak, "tahmini" tepki fonksiyonu oluşturulur, $\hat{\Phi}(x) = \sum_{r=1}^3 \hat{\beta}_{ir} M_r(x)$. $\hat{\Phi}(x)$ 'i kullanarak X tercih verilerindeki tüm yanıtları dönüştürerek, tepki stillerinin kaldırıldığı matris ($n \times m$) elde edilir. Burada ($n \times m$) matrisi “düzeltilmiş veri” olarak adlandırılır. Yanıtlayıcıları düzeltilmiş veri matrisindeki içeriğe dayalı olarak kümelemek için, düzeltilmiş verilere k-ortalama kümeleme analizi uygulanır.

Tepki stili verilerinin düzeltilip daha sonra kümelenmesi ayrı ayrı optimize edildiğinde, düzeltme ve içerik tabanlı kümeleme için en uygun sonuçları vermeyebilir (Arabie & Hubert, 1994). TSYDK yöntemi kullanılarak bu iki prosedür aynı anda yürütülebilmektedir. Böylece düzeltilmiş verilerin kümelenmesinde daha geçerli sonuçlar elde edilebilecektir.

TSYDK: Tepki Stili Yanlılığı Bulunan Verilerin Düzeltilmesi ve Kümelenmesi

Eşzamanlı düzeltme ve kümeleme, iki minimizasyon kriteri eklenerek elde edilebilir (Hwang, Dillon, & Takane, 2006). K, içerik tabanlı kümelerin sayısı olsun. TSYDK yönteminin temel fonksiyonu aşağıdaki gibi tanımlanır:

$$L(B, G, U | \bar{F}, Z, M_1, M_2, M_3) = \lambda \sum_{i=1}^n \| \bar{f}_i - M\beta_i \|^2 + (1-\lambda) \sum_{i=1}^n \sum_{k=1}^K u_{ik} \| Z_i \tilde{M}\beta_i - g_k \|^2$$

$$\text{s.t.} \sum_{r=1}^3 \beta_{ir} = 1, \quad \beta_{ir} \geq 0 \quad (r = 1, 2, 3; i = 1, \dots, n), \quad (\text{Denklem 7})$$

Burada, $B = (\beta_i)$, $G = (g_k)$, $U = (u_{ik})$, $F = (f_i)$ ($i = 1, \dots, n; k = 1, \dots, K$) ve $Z = (Z_i)$ ve $Z_i = (z_{ij})$ ($j = 1, \dots, m; l = 1, \dots, q$) şeklindedir. Denklem 7' de ilk terim düzeltme terimidir ve ikinci terim içeriğe dayalı kümeleme terimidir. $\lambda \in [0, 1]$ şeklindedir. λ terimi ağırlıklandırma terimidir ve analizden önce belirlenmesi gerekir.

İçeriğe dayalı kümelemede, k-ortalamalar kümeleme analizi düzeltilmiş veriler üzerinde gerçekleştirilir. Yani $Z_i M\beta_i = (\hat{y}_{ij})$ ($i = 1, \dots, n; j = 1, \dots, m$) şeklindedir. Eğer i yanıtlayıcısı j. maddenin l kategorisini seçerse, ($q \times 1$), $z_{ij} = (z_{ijl})$ ($l = 1, \dots, q$) bir dummy vektördür ve $z_{ijl}=1$ değerini alır, aksi takdirde $z_{ij} = 0$ olur. $q \times 3$ matrisi $M = (M_r(\tau(l)))$, ($l = 1, \dots, q; r = 1, 2, 3$) temel fonksiyon matrisidir; ancak M'den

farklı olarak, Tanım 3'te düzeltilmiş verileri oluşturmak için sınırların orta noktalarını argümanlar olarak alır. $K \times 1$ vektörü $u_i = (u_{ik})(k = 1, \dots, K)$ içerik tabanlı küme için bir gösterge vektörü olup, eğer yanıtlayıcı i k . içerik tabanlı kümeye ait ise $u_{ik} = 1$, aksi halde $u_{ik} = 0$ olur. $G, K \times m$ içerik tabanlı küme sentroid matrisidir.

TSYDK Parametre Tahminleri

$B; G; U$ parametrelerini elde etmek için, iki işlem - tepki fonksiyonlarının tahmini (B tahmini) ve içerik tabanlı kümeleme (G ve U tahmini) - sırayla gerçekleştirilir. B 'nin sabitlenmesi için Denklem 7' nin minimize edilmesi, $Z_i \tilde{M} \beta_i (i = 1, \dots, n)$ verisi (tepki stili düzeltilmiş) k -ortalamlar kümelenmesini azaltır. Diğer yandan, G ve U sabitlendiğinde, B için çözüm daha az önemsizdir, çünkü bu Denklem 7'de her iki açıdan da görülmektedir. Bununla birlikte, Denklem 7' nin B 'ye göre en aza indirilmesi, aşağıdaki gibi basit bir sınırlandırılmış en küçük kareler problemine indirgenebilir.

$$L(B, G, U | \bar{F}, Z, M_1, M_2, M_3) = \sum_{i=1}^n \left\| \left(\frac{\sqrt{\lambda} f_i}{\sqrt{(1-\lambda)}} G^T u_i \right) - \left(\frac{\sqrt{\lambda} M}{\sqrt{(1-\lambda)}} Z_i \tilde{M} \right) \beta_i \right\|^2 \quad (\text{Denklem 8})$$

TSYDK parametreleri aşağıdaki algoritma kullanılarak tahmin edilebilir.

Adım 1: Başlatma: B, G, U için rastgele bir başlangıç değeri ve ε yakınsama kriteri ve λ ayarlanır. Daha sonra yineleme sayısını $w=1$ 'den w 'ye ayarlanır.

Adım 2: Tepki fonksiyonu tahmini: G, U sabitleri için, B 'yi, Denklem (8) 'nin, Denklem (1)' deki sınırlandırma ile en aza indirilecek şekilde güncellenir (Haskell & Hanson, 1981).

Adım 3: İçeriğe dayalı kümeleme: Sabit B için G, U formülünü kullanarak güncellenir.

$$g_k = \frac{\sum_{i=1}^n u_{ik}(Z\tilde{M}\beta_i)}{\sum_{i=1}^n u_{ik}}$$

$$u_{ik} = \begin{cases} 1, & (k = \operatorname{argmin} \parallel \sqrt{1-\lambda}(g_s - Z_i\tilde{M}\beta_i) \parallel^2) \\ 0, & (\text{Diğer durumlar}) \end{cases} \quad (i = 1, \dots, n; k = 1, \dots, K).$$

Adım 4: Yakınsama testi: $L^{(w)}$, güncellenmiş parametreleri kullanarak temel fonksiyon (Denklem 7) değeri hesaplanır ve $w > 1$ için, eğer, $L^{(w)} - L^{(w-1)} < \varepsilon$ ise sonlandırılır; aksi halde $w = w + 1$ olur ve 2. adıma dönülür).

Sonraki adımlarda temel fonksiyon monoton olarak azaldığı için algoritmanın yakınsaması garanti edilir. Algoritmanın 1. adımında B, G, U için başlangıç değerlerinin seçilmesi gerekmektedir. Bu rastgele yapılır; örneğin, homojen bir dağılımdan rasgele değerler üreterek yapılabilir. Alternatif olarak, Denklem 7'deki birinci terim için yani sınırlı verilerin tepki fonksiyonlarının optimal uyarlaması için β_i ($i = 1, \dots, n$) çözülerek B, G, U için başlangıç değerleri ve düzeltilmiş verilere $Z_i\tilde{M}\beta_i$ ($i = 1, \dots, n$) K-ortalamlar uygulanmaktadır (Takagishi vd., 2019, s. 10).

İlgili Araştırmalar

İlgili araştırmalar bölümünde, tepki stillerinin model temelli olmayan geleneksel yaklaşımlar ile incelendiği araştırmalara ve tepki stillerinin model temelli yöntemlerle incelendiği araştırmalara yer verilmiştir.

Tepki Stillerini İncelemede Kullanılan Geleneksel Yaklaşımlar

Model temelli olmayan istatistiki yöntemler aracılığıyla tepki stilleri incelenirken tepki ölçeklerindeki madde puanlarının ortalaması, standart sapması veya madde kategorilerinin frekansları hesaplanarak tepki stili indeksleri oluşturulmaktadır. Oluşturulan tepki stili indeksleri varyans analizi, kümeleme analizi, regresyon veya Hiyerarşik Lineer Modelleme (HLM) yöntemlerinde kullanılarak tepki stillerinin çeşitli faktörler (etnik köken, cinsiyet, kişilik özellikleri, kültürler özellikler) ile ilişkileri incelenmektedir.

Van Dijk, Datema, Welten ve Van de Vijver (2009) tarafından yapılan araştırmada, Uluslararası Sosyal Araştırma Programı (International Social Survey Programme-ISSP) verileri kullanılarak, ülkelerdeki alan bağımlılıkları ile KTS ve UTS eğilimleri arasındaki farklılıklar incelenmiştir. Alan bağımlılıkları, “Sosyal Eşitsizlik”, “Din”, “Aile”, “Ulusal Kimlik” ve “Hükümetin Rolü” şeklindeki konuları içeren beş içerik alanını kapsamaktadır. Ayrıca araştırmada çeşitli kültürel, psikolojik ve ekonomik değişkenler ile tepki stilleri arasındaki ilişkilere de bakılmıştır. KTS indeksleri hesaplanırken, bireylerin olumlu kategori tercihleri pozitif olarak değerlendirilmiş ve 1 puan verilmiş, olumsuz kategori tercihleri negatif olarak değerlendirilmiş ve 0 puan verilmiştir. Daha sonra, ortalamalar alınarak her bir katılımcı için ağırlıklandırılmış KTS indeksi oluşturulmuştur. Benzer şekilde UTS içinde, ölçek maddelerinin her iki uç kategorisi 1 şeklinde puanlanmış ve diğer kategorilere 0 puan verilmiştir. Daha sonra KTS olduğu gib UTS indeksi oluşturularak analizler gerçekleştirilmiştir. Araştırmanın sonuçlarına göre, hem “sosyal eşitsizliğin” hem de “aile” içerik alanlarının diğer alanlara kıyasla daha çok bireysel düzeyde alakalı olduğu ortaya çıkmıştır. Anketteki bireysel düzeyle alakalı olduğu düşünülen sorularda, ülkeler içinde KTS’ nin daha homojen olduğu sonucuna varılmıştır. UTS’ de, ortalama veya standart sapma

puanları için içerik alanları arasında önemli bir fark bulunmamıştır. Tepki stillerinin psikolojik değişkenler ile ilişkisine bakıldığında; KTS' nin refah, bireycilik ve mutluluk ile negatif olarak ilişkili olduğu, UTS' nin ise sadece mutluluk ile negatif olarak ilişkili olduğu bulunmuştur. Belirsizlikten kaçınma ile hem KTS hem de UTS arasında pozitif ilişkiler bulunmuştur. Bunun yanında KTS ile ekonomi, eğitim ve iletişim faktörleri arasında tutarlı ve olumsuz bir ilişki bulunmuştur. Buna göre, bir ülkedeki yüksek ekonomik kalkınma, yüksek eğitim seviyesi ve kitle iletişiminin dah az KTS ile birlikte geldiği görülmüştür. Buna karşılık ülkelerin nüfusu ile KTS ile pozitif ve orta düzeyde bir ilişki bulunmuştur. Bu sonuç kalabalık ülkelerin daha fazla KTS gösterdiklerine işaret etmektedir.

Arce-Ferrer (2006) yaptığı çalışmada, çevrilmiş ve uyarlanmış derecelendirme ölçeklerinde UTS'yi etkileyen faktörleri incelemiştir. Derecelendirme ölçeklerinin çevirisini ve kültürel uyarlaması katılımcıların dilinden, kültüründen, oranlarından ve ölçüm süreçlerinin özelliklerinden etkilenmektedir. Araştırmada kırsal ve kentsel ortamlardan gelen ve İspanyolca konuşan katılımcıların UTS oranlarını etkileyen faktörler araştırılmıştır. Verilerin analizinde UTS indeksi oluşturulmuş ve bu indeks ANOVA'da değişken olarak kullanılmıştır. Katılımcıların derecelendirme ölçeklerine, öznel kategorilerine ve ölçek özelliklerine aşinalık sorunları araştırılmış ve bunların UTS ile ilişkili olduğu görülmüştür. Ayrıca UTS'nin, ölçek özellikleri ile yanıtlayanların öznel kategorileri, yanıtlayanların derecelendirme ölçeklerine aşinalıkları ve iletişim normları arasındaki uyumsuzluktan kaynaklandığı görülmüştür.

Beuthner, Friedrich, Herbes ve Ramme (2018) tarafından yapılan çalışmada, kültürlerarası pazarlamada Meksika ve Güney Kore katılımcıları arasındaki tepki stili farklılıkları incelenmiştir. Çalışmada UTS, OTS, KTS ve RETS için bireylerin cevaplarının frekansları kullanılarak tepki stili indeksleri oluşturulmuştur. Araştırmada veriler, 500 Meksikalı ve 525 Güney Koreli katılımcıya uygulanan çevresel tutumlar, bireysel sorumluluklar ve yeşil ambalaj özelliklerine odaklanan 28 soruluk bir anket aracılığı ile elde edilmiştir. Her iki ülkedeki katılımcılar soruları çoğunlukla "Katılıyorum" veya "Kesinlikle Katılıyorum" şeklinde yanıtlamayı tercih ederken, Meksika'daki katılımcıların "Katılmıyorum" veya "Kesinlikle Katılmıyorum" kategorilerini Güney Kore'dekilere göre iki kat daha fazla tercih ettikleri bulunmuştur. Araştırma sonuçlarına göre Meksika'daki katılımcılar UTS eğiliminde olurken, Güney Kore'dekiler OTS sergileme eğilimi göstermiştir.

Clarke (2001) tarafından yapılan çalışmada kültürler arası araştırmada UTS eğilimi incelenmiştir. Çalışmada Greenleaf'ın UTS ölçümünün dört ülkeye uygulamasıyla, UTS' nin

kültürler arasında ve madde formatları arasında farklılık gösterdiği tespit edilmiştir. Ayrıca, KTS'nin de kültürler ve madde formatları arasında değiştiğine dair kanıtlar da bulunmuştur. Çalışmada, ayrıca bir dizi ANOVA aracılığıyla, UTS ve KTS için gruplar arası farklılıkları en aza indirmek için post hoc tepki stili ayarlamasının nasıl kullanılabileceği gösterilmiştir. Araştırma sonuçlarına ek olarak, bu çalışma CETSCALE gibi pazarlama odaklı bir anket aracını kullanan kültürler arası pazar araştırmacılarının, UTS' deki gruplar arası farklılığa uyum sağlayamayarak hatalı sonuçlara nasıl ulaşabileceklerini göstermektedir.

Jakobs, Böldt, Selbach ve Steffens (2019) tarafından yapılan araştırmada, psikoakustik deneylerde dinleyicilerin sosyodemografik özelliklerinin UTS eğilimi üzerindeki etkisi incelenmiştir. Psikolojik literatüre göre UTS' nin kadınlarda erkeklere göre daha belirgin olduğu varsayılmıştır. Araştırmada Weinzierl, Lepa ve Ackermann (2018) tarafından elde edilen verileri yeniden analiz edilmiştir. Çalışmada 190 denek, Oda Akustik Kalite Envanteri (RAQI) aracılığıyla 35 ikili simüle edilmiş oda akustik izlenimleri derecelendirilmiştir. Doğrusal karma etki modellerinin (Linear Mixed-Effect) sonuçları, yaş değişkeninin UTS üzerinde önemli bir etkisi olmadığını ortaya koymuştur. Buna karşılık, elde edilen sonuçlar kadınların erkeklerden önemli ölçüde daha yüksek UTS sergilediklerini göstermektedir ve anketi cevaplama süresinin artması daha yüksek UTS'ye yol açmıştır. Bu sonuçlara göre, UTS'nin anketleri değerlendirmedeki rolü ve psikoakustik deneyleri yürütürken ve analiz ederken UTS faktörünün dikkate alınması gerektiği vurgulanmıştır.

Beuckelaer, Weijters ve Rutten (2010) tarafından yapılan çalışmada, KTS ve UTS' nin yaş, eğitim düzeyi ve ulusal kültür gibi değişkenler ile ilişkisi incelenmiştir. Ayrıca araştırmada frekansa dayalı geleneksel yöntemler ile tepki stilleri için temsili göstergeler (RIRS) yöntemi de karşılaştırılmıştır. Araştırmada Hollanda'dan alınan anket verileri kullanılmıştır. Geleneksel yöntem, ilgili maddelerin ad-hoc kümesine dayanır. Alternatif yöntem, maddelerin heterojenliğini ve temsiliyetini optimize etmek için rastgele örneklenmiş bir dizi madde kullanır. Elde edilen sonuçlara göre; geleneksel yöntemin, KTS ve UTS düzeylerini tahmin etmede RIRS yöntemine göre yetersiz kaldığı görülmüştür. Ayrıca geleneksel yöntemlerin kullanımı ile KTS ve UTS' nin diğer değişkenlerle (yaş, eğitim düzeyi ve ulusal kültür) ilişkilerine yol açabileceği bulunmuştur.

Harzing (2006) tarafından yapılan ve ülkeler arasındaki tepki stilleri farklılıklarının incelendiği araştırmada 26 farklı Avrupa ülkesindeki üniversite öğrencilerinden elde edilen veriler kullanılmıştır. Buna göre öğrencilerin çeşitli tutum ifadelerine verdikleri yanıtların frekansları alınarak UTS, KTS, RETS ve OTS indeksleri oluşturulmuştur. Araştırma

sonuçları daha önceki araştırmaları hem onaylayan hem de genişleten ülkeler arasında tepki stillerinde büyük farklılıklar olduğunu göstermektedir. Güç mesafesi, toplulukçuluk, belirsizlikten kaçınma ve dışadönüklük gibi ülke düzeyindeki özelliklerin tümü, UTS ve KTS gibi tepki stillerini önemli ölçüde etkilemiştir. Dahası, İngilizce anketlerinin öğrencilerin OTS eğilimlerini arttırdığı görülürken, ana dildeki anketlerde ise öğrencilerin daha çok UTS eğilimi gösterdikleri görülmüştür. Son olarak, İngilizce dil yeterliliği, UTS ile pozitif ilişkili bulunurken, OTS ile negatif ilişkilidir.

Tsekouras (2017) tarafından yapılan çalışmada, tüketici tarafından oluşturulan ürün derecelendirmelerinde meydana gelen UTS yanlılığı incelenmiştir. Verilerin analizinde regresyon analizi kullanılmıştır. Araştırmada iki ürün kategorisinde yapılan iki deneyden elde edilen bulgular, derecelendirme ölçeğinin orta noktasının açıkça etiketlenmesinin UTS eğilimini azalttığını, ancak bu etki uç noktalarda duygusal etiketler kullanıldığında zayıfladığını göstermiştir. Dahası duygusal etiketlerin kullanılması, derecelendirme ölçeğinin boyutu büyük olduğunda kullanıcıları uç tepki kategorilerinden caydırmıştır. Etkiler, farklı model özellikleri ve sağlamlık kontrollerinde tutarlı kalmıştır. Çalışma, tüketici ürün derecelendirmeleri bağlamında UTS yanlılığına ilişkin teorik çıkarımları genişletmiştir.

Benitez, He, Van de Vijver ve Padilla (2016) bir karma yöntem tasarımında nicel ve nitel kanıtları entegre ederek kültürler arası araştırmada UTS' nin daha iyi anlaşılmasına katkıda bulunmuşlardır. Araştırmanın nicel aşamasında, farklı uluslararası çalışmalardan alınan "Yaşam Kalitesi Ölçeği" verilerinden türetilen UTS indeksleri İspanya ve Hollanda arasında karşılaştırılmıştır. Çalışma sonuçlarına göre UTS yanlılığının İspanyollar arasında Hollandalılar' a kıyasla daha yaygın olduğu bulunmuştur. Nitel aşamada, her ülkeden 25 katılımcıyla bilişsel görüşmeler yapılmıştır. Nicel ve nitel bulguların birleşimine göre UTS' deki ülke farklılıklarının, Hollandalılar tarafından her bir maddenin daha bağımsız bir şekilde değerlendirilmesi, İspanyollar tarafından "Asla (Never)" kategorisi için daha güçlü çağrışımlar ve iş tatmini gibi belirli konular tarafından tetiklenen daha güçlü duygular dâhil olmak üzere çeşitli kaynaklardan kaynaklanabileceğini göstermiştir. Nicel ve nitel kanıtların birleşiminin kültürler arası benzerlikleri ve UTS' deki farklılıkları anlamaya yardımcı olabileceği sonucuna varılmıştır.

Bachman, O'Malley ve Freedman-Doan (2010) tarafından yapılan çalışmada, tepki stili indeksi oluşturularak etnik köken, sınıf seviyesi ve cinsiyetin UTS ile ilişkisi incelenmiştir. Araştırmadan elde edilen sonuçlara göre etnik köken, sınıf seviyesi ve cinsiyetin UTS ile

ilişkili olduğu bulunmuştur. Buna göre, her bir sınıf seviyesinde ve etnik kökeninde erkeklerin kız öğrencilere göre daha çok UTS eğilimi gösterdiği görülmüştür. Ayrıca 8. sınıftan 12. sınıfa gelindikçe öğrencilerin UTS eğilimlerinin azaldığı bulunmuştur. Etnik kökene göre ise, Afrika kökenli Amerikalıların, İspanyol kökenli ve Asya kökenli Amerikalılara göre UTS eğilimlerinin daha fazla olduğu saptanmıştır.

Baumgartner ve Steenkamp (2001), uluslararası pazarlama araştırmalarındaki tepki stillerini incelemişlerdir. Araştırmada 11 ülkedeki bireylerin beş farklı ölçeğe verdikleri yanıtlar göz alınarak beş farklı tepki stili (KTS, RETS, UTS, OTS, Koşulsuz Yanıt Verme (noncontingent responding)-NCR-) incelenmiş ve bunların ölçek puanları üzerindeki yanlılık etkilerini ve ölçekler arasındaki korelasyonları tartışılmıştır. Verilerin analizinde üç düzeyli Hiyerarşik Lineer Modelleme (HLM) kullanılmıştır. Üç düzeyli analizde birinci düzeyi bireyler, ikinci düzeyi olumlu ve olumsuz madde sayılarının denk olduğu ölçek ve üçüncü düzeyi ise ülkeler oluşturmuştur. Çalışma sonuçlarına göre tepki stillerinin ölçek puanları üzerindeki sistematik etkisinde ters puanlı maddelerin oranı, yanıt ölçeğinin orta noktasından sapma ve ölçekler arasındaki korelasyonların tepki stili bileşenleri arasındaki korelasyona bağlı olarak pozitif ve negatif yönde yanlılık olabileceğini göstermiştir. Beş farklı tepki stilinin ölçek puanları üzerindeki etkisi ölçekten ölçeğe göre değişmektedir. Ülkeler arasındaki tepki stili farklılıklarının ise ihmal edilecek kadar küçük olduğu bulunmuştur. Ayrıca olumlu ve olumsuz madde sayısının denkleğinin de tepki stili yanlılığı üzerinde etkili olduğu görülmüştür. Elde edilen tüm bu bulgular pazarlama araştırmacılarının ölçme araçlarını geliştirirken ve kullanırken tepki stili yanlılığına daha fazla dikkat etmeleri gerektiğini göstermektedir.

Johnson, Kulesa, Cho ve Shavitt (2005) tarafından HLM kullanılarak ülke düzeyinde Hofstede (2001) tarafından tanımlanan ve incelenen dört kültürel yönelimin, UTS ve KTS üzerindeki etkileri incelenmiştir. Veriler, beş kıtadaki 19 ülkedeki (Avustralya, Belçika, Brezilya, Çek Cumhuriyeti, Almanya, Macaristan, Hindistan, Japonya, Malezya, Portekiz, Türkiye, Birleşik Krallık, Meksika, Filipinler, Polonya, Singapur, Hong Kong, Fransa ve İtalya) çalışanlar tarafından doldurulan yaklaşık 18.000 anket verisinden elde edilmiştir. HLM, kişi düzeyinde tepki stilleri ile ülke düzeyinde kültürel yönelimler arasındaki ilişkileri incelemek için kullanılmıştır. Çalışma sonucunda, güç mesafesi ve erkekliğin UTS ile pozitif yönde ve bağımsız olarak ilişkili olduğunu göstermiştir. Bireycilik, belirsizlikten kaçınma, güç mesafesi ve erkeklik özelliklerinin KTS ile negatif yönde ilişkili olduğu bulunmuştur.

Ulusal alan yazına bakıldığında, İlgün Dibek, Yavuz ve Çokluk Bökeoğlu (2018) tarafından yapılan araştırmada PISA 2012 uygulamasına katılan öğrencilerin tepki stilleri incelenmiş ve ülkelerin tepki stillerine göre sınıflandırması yapılmıştır. Çalışmada UTS, KTS ve RETS ele alınmıştır. Ülkelerin seçiminde ölçüt olarak öğrencilerin okulda kendilerini mutlu hissetmesi ile PISA matematik okuryazarlığı performansı dikkate alınmıştır. Araştırmada veriler Okula Aidiyet ile Okula Karşı Tutum ölçeklerinden elde edilmiştir. Verilerin analizinde tek yönlü varyans analizi ile kümeleme analizi kullanılmıştır. Araştırmanın bulgularına göre ülkeler *mutluz ama başarılı* olan ülkeler: Kore, Estonya ve Finlandiya *ne mutlu ne de başarılı* olan ülkeler Türkiye, Birleşik Arap Emirliği (Dubai) şeklinde ve *mutlu ama başarısız* ülkeler: Kolomiya, Endonezya, Peru olarak üç kümeye ayrılmıştır. Araştırma sonuçlarına göre üç gruptaki ülkelerin UTS ve RETS gösterme eğilimleri açısından farklılaştığı görülmüştür.

Model temelli olmayan tepki stili kestirim yöntemlerinin sınırlılıklarının başında ölçülmek istenen özelliğin etkisi ile tepki stillerinin etkisini ayırmada yetersiz kalmaları gelmektedir. Diğer bir ifade ile tepki ölçeğinde sürekli olumlu uç kategorileri seçen bir birey geçekten de ölçülen özelliğe yüksek düzeyde sahip olabilir. Bu durumda model temelli olmayan yaklaşımlar bu durumun ölçülen özellikten mi yoksa tepki stili etkisinin varlığından mı kaynaklandığını ayırt edemez. Ayrıca, model temelli olmayan yaklaşımlarda ölçek puanlarındaki yanlışlığın üzerinde tepki stillerinin doğrusal bir etkisinin olduğu varsayılmaktadır (İlgün Dibek, 2018).

Geleneksel yöntemler ile tepki stillerinin incelendiği çalışmalar özetlenirse, bireylerin cevaplarının frekansı, madde puan ortalamaları ve madde puanlarının standart sapması kullanılarak UTS, KTS ve RETS indeksleri oluşturulmaktadır. Oluşturulan indeksler ANOVA, regresyon analizi ve HLM yöntemlerinde değişken olarak kullanılarak çeşitli değişkenler ile tepki stilleri arasındaki ilişkiler incelenmiştir. Ayrıca Tukey ve Bonferroni gibi post hoc karşılaştırma testleri kullanılarak ülkeler arasında ve kültürel arasındaki tepki stilleri karşılaştırılmıştır.

Tepki Stillerini İncelemede Model Temelli Yaklaşımlar

Model temelli olmayan yaklaşımlarda, bireylerin UTS, KTS ve RETS gib tepki stillerini belirlemek amacıyla maddelerin frekans, ortalama ve standart sapmaları kullanılarak tepki stili indeksi oluşturulmaktaydı. Model temelli yöntemlerde ise tepki stili indeksi oluşturulmadan tepki stilleri incelenmektedir. Model temelli yöntemler arasında Doğrulamalı Faktör Analizi (DFA), Gizil Sınıf Analizi (GSA), Madde Tepki Kuramı (MTK) yer almaktadır. DFA ile gerçekleştirilen çalışmalarda tepki stilleri iki farklı şekilde ele alınmaktadır. Bazı araştırmalarda tepki stillerini incelemek için çoklu grup DFA kullanılırken, bazı araştırmalarda ise tepki stilleri sürekli bir değişken olarak ele alınmakta ve analizler buna göre yapılmaktadır. GSA kullanılan araştırmalarda ise tepki stilleri kategorik değişken olarak kabul edilmekte ve analizler buna göre yapılmaktadır. Son yıllarda ise tepki stillerinin incelenmesinde daha çok MTK tabanlı modeller kullanılmaktadır. MTK tabanlı modellerde tepki stilleri ayrı bir parametre şeklinde ve sürekli bir değişken olarak ele alınmaktadır.

DFA ile Yapılan Araştırmalar

Cheung ve Rensvold (2000), tepki stillerinin farklı grupların tutumlarının karşılaştırılmasındaki etkisinin incelenmesinde Çoklu Grup Doğrulamalı Faktör Analizi (ÇG-DFA) yönteminden yararlanmışlardır. Araştırmadan elde edilen sonuçlara göre faktör yükleri, kesişimler, gizil değişken ortalamalarının gruplar boyunca değişmez olduğu bulunmuştur. Buna göre, ölçülmek istenen özelliğin gruplar arasında değişmez olduğu ve grup karşılaştırmalarında tepki stillerinden kaynaklı yanlılığın karışmadığı sonucuna varılmıştır.

Weijters, Schillewaert ve Geunes (2008) tarafından yapılan araştırmada ise kâğıt kalem anketleri, telefon görüşmeleri ve çevrimiçi anketler olmak üzere üç farklı veri toplama metodu arasında tepki stillerinin seviyeleri karşılaştırılmıştır. Metodlar KTS, RETS ve OTS açısından değerlendirilmiştir. Karşılaştırmaları yapmak için araştırmacılar temsili göstergeler olarak tepki stili ortalamaları ve kovaryans yapısı (RIRSMACS) yöntemini önermişlerdir. Bu yöntem birden çok tepki stilinin eşzamanlı olarak analiz edilmesini sağlamaktadır. Modelde, ortak madde setlerine bağlı olmaları nedeniyle tepki stili göstergeleri arasındaki bağımlılıklar hesaba katılır. Ayrıca yöntem tepki stili ölçülerindeki rastgele hatayı da açıklamaktadır. Yeni önerilen yöntem aracılığıyla, ölçme değişmezliği

testinin ve çok aşamalı çoklu yöntem yaklaşımlarının yetersiz olduğu durumlarda çapraz mod tepki stili farklılıklarını algılayabilir ve düzeltebilir.

Welkenhuysen-Gybels, Billiet, ve Cambre (2003) tarafından yapılan araştırmada tepki stili ayrı bir faktör ele alınarak incelenmiştir. Araştırmada KTS'nin ayrı bir faktör olarak ele alınmadığı durumda faktör yüklerinin gruplar boyunca ölçme değişmezliğine olan etkisine bakılmıştır. Araştırma sonuçlarına göre KTS ayrı bir faktör olarak ele alınmadığında faktör yüklerinin ve standart hatalarının yanlış olarak kestirildiği görülmüştür.

Jin ve Chen (2020) tarafından yapılan araştırmada UTS' nin varlığının yanlış puanlara yol açabileceğini ve dolayısıyla Diferansiyel Madde Fonksiyonu (DMF) tespitinin doğruluğunu etkileyebileceğini düşünmektedirler. Araştırmalarında, UTS'nin DMF tespitinde etkisini ortadan kaldırmanın bir yolu olarak çoklu göstergeli çoklu nedenler (MIMIC) çerçevesi altında yeni bir yöntem önerilmişlerdir. Bir dizi simülasyondan elde edilen bulgulara göre gruplar arasındaki UTS farkının, UTS hesaba katılmadığında DMF tespitinde yanlış pozitif oranların şişmesine ve gerçek pozitif oranların düşmesine neden olduğu görülmüştür. Geleneksel MIMIC, diskriminant analizi, sıralı lojistik regresyon ve bunların uzantılarına kıyasla modifiye edilmiş MIMIC modeli, durumlar genelinde yanlış pozitif oranları kontrol edebilir ve güvenilir gerçek pozitif oranlar sağlamıştır.

GSA ile Yapılan Araştırmalar

Gizil sınıf analizi aracılığıyla yapılan araştırmalarda; gizil sınıf faktör analizi, çoklu grup gizil sınıf analizi, gizil sınıf karma modelleri ve bazı araştırmalarda da gizil sınıf multinominal lojistik modelleri kullanılmıştır. Aşağıda GSA temelli yaklaşımların kullanıldığı çalışmalara yer verilmiştir.

Liu, Conrad ve Lee (2017) tarafından yapılan çalışmada yüz yüze uygulanan anketler ile internet üzerinden yapılan anketlerin KTS ve UTS üzerindeki etkileri karşılaştırılmıştır. Çalışmada, 2012 Amerikan Ulusal Seçim Çalışmalarından (ANES) elde edilen veriler GSA aracılığıyla analiz edilmiştir. Araştırma sonucunda hem yüz yüze hem de internet üzerinden yapılan anketlerde KTS ve UTS'nin mevcut olduğu görülmüştür. Ayrıca yüz yüze anket uygulanan katılımcıların internet üzerinden anketlerin uygulandığı katılımcılara kıyasla daha yüksek KTS ve UTS sergiledikleri görülmüştür.

Eid, Langeheine ve Diener (2003) tarafından yapılan araştırmada çoklu grup gizil sınıf analizinin psikoloji alanındaki kültürlerarası araştırmalar için avantajları tartışılmıştır. Araştırmada çok gruplu gizil sınıf analizi tanımlanmış, Çin ve Amerika kültürlerinin yaşam alanından memnuniyet profilleri karşılaştırmıştır. Araştırmada kültürler arasında çeşitli ölçme değişmezliği varsayımlarının gizil sınıf çerçevesinde istatistiksel olarak nasıl test edilebileceği gösterilmiştir. Araştırma sonuçlarına göre, Çin ve Amerika Birleşik Devletlerindeki kişiler iki gizil sınıfa ayrılmışlardır. Ayrıca, gizli sınıf analizi Çin ve Amerika Birleşik Devletleri'nin alan memnuniyeti sınıfları arasında yapısal farklılıklar olduğunu ortaya koymuştur. İki ülke arasındaki farklar, öncelikle uygun (satisfied) edilen kategorinin olasılıkları içindedir. Bu kategori için iki sınıftaki olasılıklar genellikle Çin'de daha düşüktür, bu da Çinlilerin yaşam memnuniyetlerini bildirmede daha mütevazı göründüklerini göstermektedir. Dahası, analizler iki sınıfın Çin'de neredeyse eşit derecede büyük olduğunu ortaya çıkarmıştır. Buna karşılık Amerika Birleşik Devletleri'nde, tüm alanlarda genel olarak büyük bir memnuniyet belirten ikinci sınıf, birinci sınıftan çok daha büyüktür. Bu nedenle, Amerikalılar Çinlilerden sadece daha az mütevazı tepki stillerinde değil, aynı zamanda genel olarak daha yüksek yaşam doyumlarında da farklılık göstermektedirler.

Morren, Gelissen ve Vermunt (2011), UTS'nin saptanması ve düzeltilmesi için istatistiksel yöntemlerle ilgilenmişlerdir. Araştırmada Moors (2003) tarafından tanıtılan Gizil Sınıf Faktör Analizi (GSFA) kullanılmış ve GSFA'nin diğer yöntemlere kıyasla birçok avantajı olduğu savunulmuştur. Moors'un önerdiği yöntem, ilgili önemli faktörlere ek olarak bir UTS faktörü içeren bir gizil değişken modelinin tanımlanmasını içermektedir. Moors'un yaklaşımı ve genişletilmiş GSFA yöntemi, kültürel olarak farklı dört grup arasındaki UTS'deki farklılıkların araştırılmasına ve düzeltilmesine olanak tanıyan Hollanda'daki "Etnik Azınlıkların Sosyal Konumu ve Hizmet Kullanımları" (SPVA) anketinden elde edilen veriler kullanılarak gösterilmiştir. Araştırmada kullanılan veriler, 1988'den 2002'ye kadar her dört yılda bir tekrarlanan SPVA adlı Hollanda anketinden elde edilmiştir. Araştırma kapsamında 2002 yılında Hollanda'daki en büyük dört etnik azınlık arasında toplanan verileri; yani Türkler, Faslılar, Surinamlılar ve Antiller'i içermektedir. Analizlerde sadece her haneden tek bir temsilcinin görüşleri alınmıştır. Yanıt oranları Surinamlılar ve Antiller için %44 ve Türkler için %52 arasındadır. Katılımcılardan, tamamen katılıyorum (1) ile tamamen katılmıyorum (5) arasında değişen, tarafsız bir orta nokta olarak ne katılıyorum ne de katılmıyorum olan 5 puanlı Likert ölçeğini cevaplandırmaları istenmiştir. Araştırma

sonuçlarına göre, UTS' nin varlığına dair güçlü kanıtlar elde edilmiş ve dahası grupların sadece tutumlarında değil aynı zamanda UTS'de de farklılık gösterdiği görülmüştür. Bu bulgular, kültürler arası araştırmalarda tutumları incelerken UTS'yi kontrol etmenin önemini göstermiştir.

Moors (2003) tarafından yapılan araştırmada tepki stilleri GSFA aracılığıyla incelenmiştir. Araştırmada, katılımcıların cinsiyet rolünün, tutumlarının ve etnik ayrımcılık algılarının sosyo-demografik ilişkileri de incelenmiştir. Araştırmacı tepki stili eğilimlerini teşhis etmede ve gizil değişkenlerle nedensel modellerden elde edilen bulguları ayarlamada gizil sınıf faktör analizinin esnekliğini göstermeyi amaçlamıştır. Araştırma verisini, etnik azınlıklar arasında entegrasyonla ilgili tutumlar üzerine Belçika MHSM araştırma projesinden elde edilen bir örneklem oluşturmaktadır. "Cinsiyet rolleri" ve "etnik ayrımcılık duyguları" na atıfta bulunan iki bağımsız likert tipi soru dizisinin analizinde UTS' nin tespit edildiği görülmüş ve 'geleneksel cinsiyet rolü' ve 'etnik ayrımcılık duygularının gizil tutum boyutlarının gerçek ilişkisinin, UTS yanlılığı hesaba katıldığında ortaya çıktığı gösterilmiştir. Tepki stili dikkate alınırsa, kovaryatların tutum boyutları üzerindeki etkisi daha yeterli bir şekilde tahmin edileceği gösterilmiştir.

Liu ve Keusch (2017) ölçek yönünün, sıralı derecelendirme ölçeklerindeki tepki stilleri üzerindeki etkisini inceldikleri araştırmalarında, 2012 Amerikan Ulusal Seçim Çalışmaları' ndan (ANES) alınmış veriler kullanılmıştır. ANES, Amerika Birleşik Devletleri'nde seçim yıllarında siyasi adaylar, partiler, genel olarak siyaset ve diğer ilgili konular üzerine yapılan ulusal bir zaman serisi anketidir. Çalışmada ölçek yön etkisinin KTS ve UTS örtük sınıf değişkenleri üzerindeki etkisi incelenmiştir. Çalışma sonuçlarına göre bir web anketi kullanarak pozitif sıfat kelimeleriyle başlayan ölçeklerde daha yüksek KTS eğilimi görülmüştür. Web tabanlı anketlerde ve yüz yüze uygulanan anketlerde, UTS üzerinde ölçek yönünün önemli bir etkisinin olmadığı tespit edilmiştir. Bu çalışma aynı zamanda modele tepki stili örtük sınıf değişkenleri dâhil edildiğinde ölçek yönünün önemli örtük sınıf değişkenlerini etkilemediğini göstermiştir.

Van Rosmalen, van Herk ve Groenen (2010) yaptıkları çalışmada, ölçek maddelerindeki tepki stillerinin ve içeriklerin derecelendirme ölçeği yanıtlarını nasıl etkilediğini incelemişlerdir. Buna göre yazarlar farklı türden tepki stillerini, maddenin içeriğini ve yanıtlayanların arka plan özelliklerini açıklayan yeni bir model geliştirmişlerdir. Buna göre, derecelendirme ölçeği yanıtlarına sahip veri kümelerinde yanıt davranışını açıklayarak içerik ve tepki stillerini ayırt etmek için yeni bir model olan Gizil Sınıf Multinomial Lojistik

Modeli kullanmışlardır. Araştırmada Fransa, Almanya, İtalya, İspanya ve İngiltere ülkelerindeki bireylerden elde edilen verilerle analizler gerçekleştirilmiştir. Araştırma sonuçları, katılımcıların fikirlerinde ve tepki stillerinde büyük farklılıklar olduğunu ortaya koymuştur. Araştırmaya katılan bireylerden %58,4' ünde KTS eğilimi, %20,6' sında uç kategorilerden kaçınma eğilimi, %9,3'ünde çok çeşitli yanıtları tercih etme eğilimi, %4,1' inde UTS eğilimi ve %2,5' inde ise OTS eğilimi gözlemlenmiştir. Bunun dışında araştırmadaki bireylerden %0,9 ise anketteki maddelere herhangi bir yanıt vermemişlerdir. Ayrıca, araştırmada yanıtlayanın özelliklerinin tepki stili davranışını ve bir maddenin içeriği üzerindeki etkilerini grafiksel olarak ayırt etmişlerdir. Bazı tepki stilleri geçerli iletişim stilleri gibi görünürken, diğer tepki stilleri genellikle maddelerin içerikleri ile tutarsızlık göstermekte ve tepki stili yanlılığı gibi görünmektedir.

Morren, Gellissen ve Vermunt (2013) karma yöntem araştırmalarını kullanarak bir tepki stiline sahip kültürel olarak farklı anket katılımcılarının tepki verme süreçlerini incelemişlerdir. Araştırmada, Hollanda'daki en büyük dört azınlığın - Türkler, Faslılar, Antilliler ve Surinamlılar' ın yanıt verme süreçlerini analiz etmek için nicel ve nitel yöntemleri birleştiren karma bir yöntem yaklaşımı sunmuşlardır. İlk olarak, bu azınlıkların maddelere farklı yanıt verip vermediğini değerlendirmek ve farklı tepki stillerine sahip katılımcıları ayırt etmek için geniş bir veri kümesi üzerinde gizli sınıf faktör analizleri gerçekleştirmişlerdir. Bu analizler, azınlıklar arasındaki grup farklılıklarının kısmen tepki stilleri ile açıklanabileceğini ve bazı katılımcıların uç kategorileri seçmeyi tercih ederken diğerlerinin bitişik kategorileri tercih ettiğini göstermiştir. Daha sonra, aynı kültürel gruplardan görüşülen kişiler seçilmiş ve nicel çalışmadan elde edilen yanıt profillerine göre sınıflandırılmışlardır. Nitel araştırma sonuçları UTS eğilimi olan kişilerin, anket sorularını yanıtlamak için komşu kategorileri seçme eğilimi olan kişilerden farklı stratejiler kullandığını göstermiştir.

Kieruj ve Moors (2013) tepki stili davranışının madde formatına ve kişisel özelliklere bağlı olup olmadığını inceledikleri araştırmada GSFA yaklaşımını kullanmışlardır. Araştırmada KTS ve UTS yanlılığının dış etkenler (Örn. test koşulları veya yorgunluk) tarafından etkilenip etkilendiği veya iç etkenler (Örn. Kişilik veya sosyal özellikler) sebebi ile meydana gelip gelmeyeceği incelenmiştir. Çalışmanın ilk bölümünde, ölçek uzunluğunun (manipüle edilmiş test koşulu) 5 ila 11 kategori arasında değiştirilmesinin UTS veya KTS oluşumunu etkileyip etkilemediği göz önüne alınmıştır. Araştırma sonuçları, UTS' nin ölçek uzunluğundan bağımsız olarak gerçekleştiğini göstermektedir. İkinci adımda, UTS' nin (a)

dış ölçümlerine bağlayarak ve (b) bir kişilik profili ve sosyo demografik özelliklerle ilişkilendirerek dahili bir kişisel stili yansıtip yansıtamayacağını kontrol etmişlerdir. Sonuçlar, UTS' nin oldukça istikrarlı olduğunu ve seçilen kişilik profili ve yaşla ilişkili olduğunu göstermiştir.

MTK ile Yapılan Araştırmalar

Klasik yöntemler, genellikle test maddelerinin psikometrik özelliklerini ve yanıt verenlerin yetenek düzeylerini hesaba katmadığı için araştırmacılar tepki stillerini ölçmek ve kontrol etmek için MTK modellerine dayalı yöntemlere yönelmişlerdir. MTK tabanlı yöntemlerde tepki stilleri sürekli değişken olarak ele alınmaktadır. MTK tabanlı çalışmalarda maddelerin DMF gösterme durumlarında tepki stillerinin etkisi incelenmiştir. Ayrıca tepki stilinin etkisini kontrol altına almak ve düzeltmek amacıyla yapılan çalışmalarda mevcuttur. Bu doğrultuda tepki stillerinin belirlenip düzeltilmesi amacıyla Çok Boyutlu Madde Tepki Kuramı (ÇBMTK) modellerine dayalı yöntemler geliştirilmiştir. ÇBMTK modelleri, tepki stilini etkileyen ek boyutlar olarak tepki stillerini birleştirir ve tepki kategorilerinin seçiminde gizil özellik ve tepki stilinin eşzamanlı etkisini açıklamaktadır (Bolt & Johnson, 2009). Kategori tepki olasılıklarının analizinde Karma Çok Kategorili Madde Tepki Kuramı (KÇKMTK) modelleri de kullanılmaktadır. KÇKMTK modellerinde MTK modelleri ile birlikte bireyleri gizil sınıflara ayıran yapılarda bulunmaktadır. Böylece farklı tepki stillerine sahip katılımcılar farklı gizil sınıflara ayrılabilirler. Belirlenen sınıflar içindeki her bir katılımcıya gizil bir özellik tahmini atanır ve böylece gizil özellik seviyeleri karşılaştırıldığında tepki stili etkileri kontrol edilebilir (Cho, 2013).

MTK tabanlı yaklaşımlardan bir diğeri ise Böckenholt (2012) tarafından geliştirilen Madde Tepki Kuramı Ağaç Modelleri (IRTree)' dir. Bu modeller ile tepki sürecinin yalnızca tepki stillerinden değil aynı zamanda madde özelliği etkilerinden veya ilgili nitelik ile ilgisi olmayan özelliklerden de etkilenebileceği durumlar göz önüne alınmaktadır. Bir ağacın dalları ve düğümlerine benzeyen yapısı ile IRTree modelleri tepki sürecini aşamalı hale getirerek her aşamada farklı bir MTK modelinin kullanılmasına izin vermektedir. Madde yanıtları bir dizi alt maddeden (sözde madde) oluşturularak MTK modelleri her bir aşamadaki maddelere uygulanmaktadır. Yapılan bazı araştırmalarda da çeşitli MTK modelleri karşılaştırılmıştır. MTK modelleri ile tepki stillerinin incelendiği çalışmalardan bazıları aşağıda sunulmuştur.

Çok Boyutlu Madde Tepki Kuramı Modelleri ile İlgili Araştırmalar

ÇBMTK modellerini diğer modellerden ayıran en önemli özelliği tepki stili yanlılığının belirlemek yanında düzeltilmesini de sağlamasıdır. Yapılan araştırmalar incelendiğinde ÇBMTK modelleri ile tıp, pazarlama ve eğitim başta olmak üzere birçok alanda çalışma yapıldığı görülmüştür.

Bolt ve Johnson (2009) ÇMTK modellerinden Bock'un (1972) Sınıflamalı Tepki Kuramı (Nominal Response Model) modelinin çok boyutlu uzantısını tanıttıkları araştırmada; yeni geliştirilen modelin sıralı derecelendirme ölçeği verilerindeki yanlılığı azaltmak için tepki stili etkilerinin incelenmesine ve kontrolüne izin verdiğini göstermişlerdir. Mevcut uygulamada model, Değişen Madde Fonksiyonu (DMF) altında yatan bir neden olarak tepki stilini araştırmak için de kullanılmıştır. Geliştirilen yeni yaklaşım, sigara içenlerin tütün bağımlılıklarındaki motivasyonlarının bir ölçüsü olan ve 68 maddeden oluşan Wisconsin Sigara Bağımlılığı Motifleri Anketine verilen madde yanıtları kullanılarak gösterilmiştir. Araştırma sonuçlarına göre sigara içme alışkanlığının düzeyine göre UTS' nin etkisinin de değiştiği ve bireylerin düzeltilmiş puanlarının düzeltilmemiş puanlarından anlamlı derecede farklı olduğu görülmüştür.

Falk ve Cai (2016) tarafından yapılan araştırmada birden çok özellik alanı yapısı boyunca birden çok kullanıcı tanımlı tepki stillerini modellemek için esnek bir tam bilgi yaklaşımı sunulmuştur. Model, her bir madde ve gizil özellik için genel madde eğimlerinin tahminini puanlama işlevlerinden (kategorilerin sırasını gösteren) ayıran çok boyutlu sınıflamalı tepki modelinin yeni bir parametreleştirilmesine dayanmaktadır. Bu özellik, tepki stillerinin tanımının hem yetenek hem de tepki stili boyutları için maddeler arasında değişen genel madde eğimlerinin yanı sıra maddeler arasında da değişmesine de olanak tanımaktadır. Araştırmada kurulan model Hasta Tarafından Bildirilen Sonuçlar Ölçüm Bilgi Sistemi (PROMIS) ölçeği maddeleri üzerinden değerlendirilmiştir. Ayrıca çalışmada simülasyon verisi de kullanılmıştır. Araştırmadan elde edilen sonuçlara göre geliştirilen modelin tahmin yaklaşımının model parametrelerini, faktör puanlarını ve standart hataların olası tahminlerini kurtarabildiğini göstermiştir. Ayrıca, simülasyon verisinden elde edilen sonuçlar tepki stili faktörlerini dahil etmeme durumunun, önemli yetenek faktörü puanının kestirilmesinde olumsuz sonuçlara yol açtığını göstermiştir. Böylece geliştirilen yeni model, birden çok tepki stilini modelleyebilmekte ve önceki yaklaşımları (Bolt & Johnson, 2009; Bolt & Newton, 2011; Johnson & Bolt, 2010) da kapsamaktadır.

Schneider (2018) tarafından yapılan araştırmada, yaşlanma araştırmalarında olumlu ve olumsuz etkinin değerlendirilmesinde tepki stili yanlılığının rolü araştırılmıştır. Araştırmada, ÇBMTK modelinden yararlanılarak Sağlık ve Emeklilik Çalışmasında (SEÇ) psikososyal ankete yanıt veren 50-100 yaş aralığındaki 6295 kişinin sağladığı duygu derecelendirmelerinden elde edilen veriler kullanılmıştır. Araştırma sonuçlarına göre, UTS olasılığının yaş ile birlikte önemli ölçüde arttığı ve bu etkiye SEÇ bilişsel test puanlarındaki yaşa bağlı düşüşlerin aracılık ettiği görülmüştür. Tepki stillerinin duygu derecelendirme ölçeğinden düzeltilmesinin yaş kalıplarını olumlu ve olumsuz duygu durumlarında değiştirmemiştir. Ayrıca duygu derecelendirme ölçeğinden elde edilen ölçümlerin yakınsak geçerliliğini (eşzamanlı depresyon ve akıl sağlığı sorunları ile ilişkiler) ve yordama geçerliliğini (hastane ziyaretleriyle ileriye dönük ilişkiler, fiziksel hastalık başlangıcı) sürekli olarak artırmıştır.

Dowling, Bolt, Deng ve Li (2016) ÇBMTK modelini kullanarak hasta tarafından bildirilen sonuç ölçümlerindeki verilerde UTS yanlılığını belirlemiş ve düzeltmişlerdir. Araştırmacılar ÇBMTK modelini kullanarak ölçülen özelliğin aynı düzeyde olduğu fakat farklı UTS seviyelerine sahip bireyler için UTS' nin toplam puanlar üzerindeki yanlı etkilerini incelemişlerdir. Araştırma sonuçlarına göre, UTS yanlılığını açıklayan dört boyutlu bir model, yanıt verilerine daha iyi uyum sağlamıştır. Artan UTS seviyeleri, özellik tahminlerinin bir fonksiyonu olarak toplam puanlarda yanlılığa neden olmuştur. UTS' nin etkisi, birden çok ölçeğin birlikte yer aldığı ve UTS' nin yer aldığı modelde daha büyüktür. Tahmini madde kategori kesişimleri (sabitleri), içerikten bağımsız kategori seçimlerin yapıldığının kanıtını sağlamıştır. Tahmin modellerinde açıklayıcı değişken olarak kullanılan düzeltilmemiş yetenek tahminleri aşağı yönlü yanlılık göstermiştir. Ayrıca, tepki stillerinin getirdiği yanlılık, bireysel özellik tahminlerini ciddi şekilde etkileyebilir ve bunların önemli olan diğer klinik sonuçlarla ilişkisini bozabilir.

Wetzel ve Carstensen (2015) tarafından yapılan araştırmada özelliklerin ve tepki stillerinin modellenmesinde ÇBMTK modeli kullanılmıştır. Araştırmada kullanılan ÇBMTK, eşzamanlı olarak özelliklerin ve tepki stillerinin modellenmesini sağlamaktadır. Farklı tepki stilleri ve kişilik özellikleri (Beş Büyük Kişilik Ölçeği) model uyumu açısından karşılaştırılmıştır. Kişilik özellikleri ve tepki stilleri arasındaki ilişkiler araştırılmış ve UTS' ni modellemeye yönelik farklı yaklaşımlar, özellik tahminleri üzerindeki etkileri açısından karşılaştırılmıştır. Araştırma sonuçlarına göre; tüm çok boyutlu modeller, tek boyutlu modellerden daha iyi bir uyum göstermiştir ve bu da tepki stillerinin madde yanıtlarını

etkilediğini ve UTS' nin bu etkideki en büyük varyansı açıkladığı görülmüştür. UTS ve OTS temelde kişilik özelliklerinden bağımsızdır. Buna karşılık KTS ve RETS birkaç kişilik özelliği ile güçlü şekilde ilişkili bulunmuştur. Bu yaklaşımdaki UTS boyutu genel bir UTS eğiliminden ziyade yalnızca ölçeğe özgü UTS' yi yansıtmaya rağmen, aynı maddeler UTS' yi ve kişilik özelliklerini modellemek için kullanıldığında özellik tahminlerinde küçük bir düzeltme meydana gelmiştir.

Lu (2012) yaptığı araştırmada, PISA 2006 uygulamasına katılan ülkelerin tutum-başarı paradoksunu açıklamada UTS' nin etkisini incelemiştir. Çalışmada ülkelerin UTS eğilimlerini belirlemek ve düzeltmek amacıyla ÇBMTK modelini çok düzeyli bağlamda kullanmıştır. ÇBMTK modeli ile kestirilen parametre kestirimlerinin doğruluğu simülasyon verisi kullanılarak test edilmiştir. Araştırmadan elde edilen sonuçlara göre ülkelerde görülen tutum-başarı paradoksuna neden olan faktörlerden biri de UTS olmuştur. Ancak UTS tek başına bu paradoksu açıklamada yeterli olmamıştır. Ayrıca UTS' nin bireylerin ölçek toplam puanlarını azaltıcı veya arttırıcı yönde rol oynadığı gösterilmiştir.

İlgün Dibek (2018) tarafından yapılan araştırmada ise TIMSS 2015 uygulamasına katılan ülkelerde görülen tutum-başarı paradoksu üzerinde etkisi olduğu düşünülen UTS' yi incelemiştir. Bu doğrultuda çalışmanın örneklemini, TIMSS 2015 uygulamasına katılan ülkeler arasından belirlenen 15 ülkenin her birinden basit seçkisiz olarak seçilen 500' er 8. sınıf öğrencisi oluşturmaktadır. Öğrencilerin UTS sergilemelerindeki farklılıklar non-parametrik MANOVA ile belirlenirken, UTS' nin etkisini belirlemek amacıyla iki farklı ÇBMTK modeli kullanılmıştır. Ayrıca öğrencilerin düzeltilmiş ve düzeltilmemiş tutum puanları arasındaki farkın ülke içinde test edilmesinde ilişkili örneklem t-testi kullanılmıştır. Araştırmadan elde edilen sonuçlara göre; UTS sergileme açısından ülkelerin farklılaştığı, UTS' nin bir boyut olarak dâhil edildiği ÇBMTK modelin veriye uyum sağlamıştır. Ayrıca, tepki stili etkisine göre düzeltilmemiş ve düzeltilmiş tutum puanları arasında anlamlı farklılığın olduğu görülmüştür. Böylece ülke düzeyinde görülen tutum-başarı paradoksuna neden olan faktörlerden birinin UTS olduğu belirlenmiştir.

Karma MTK Modelleri ile İlgili Araştırmalar

Karma MTK modelleri kullanılarak tepki stillerinin incelediği araştırmalara bakıldığında Karma Çok Kategorili Rasch Modelleri, Karma Kısmi Puanlama Modeli, Karma

Derecelendirilmiş Ölçek Modeli yaklaşımlarından yararlanıldığı görülmüştür. Aşağıda Karma MTK modelleri ile yapılan araştırmalardan bazılarına yer verilmiştir.

Wetzel, Böhnke ve Rose (2016) yaptıkları simülasyon çalışmasında, UTS'nin etkisinin düzeltilmesi için Karma Rasch modelleri, ÇBMTK modelleri ile regresyon yöntemlerini karşılaştırmışlardır. Üç yöntem gerçek gizil yetenek düzeylerinin geri kazanımları açısından karşılaştırılmıştır. Veriler, yalnızca bir niteliği ölçen tek boyutlu bir modele, iki UTS bulunan modele ve UTS bulunmayan örnekleme Karma Rasch modeli ve bir nitelik ve bir UTS boyutunu içeren iki boyutlu bir MTK modeline göre oluşturulmuştur. Elde edilen sonuçlara göre, Karma Rasch modelleri genel olarak diğer düzeltme yöntemlerinden daha düşük başarı göstermiştir. Simülasyon çalışmasının sonuçlarına göre ise, ortalama olarak UTS'yi göz ardı etmenin, eğer UTS ve gizil yetenek ilişkisiz ise veya zayıf bir şekilde ilişkili ise yetenek tahminlerini neredeyse hiç etkilemediğini göstermektedir (Wetzel ve Carstensen, 2015). Regresyon modellerinin en büyük dezavantajı, gizil yetenek ve UTS'nin ilişkisiz olduğu şeklindeki varsayımdır. Son olarak, incelenen yöntemlerden iki boyutlu madde tepki kuramı modeli en iyi performansı göstermiştir.

Wetzel, Carstensen ve Böhnke (2013) özellikler arasında uç tepki stili olma durumu ve uç tepki stili olmama durumların tutarlılığını Karma Rasch modelleri ile incelemiştir. Çalışmada ayrıca, ikinci dereceden gizil sınıf analizi de kullanılmıştır. Araştırma verisi olarak PISA 2006 tutum ölçeği ve Alman NEO-PI-R ölçeği kullanılmıştır. NEO-PI-R, beş büyük kişilik özelliği olan nevrotiklik, dışadönüklük, deneyime açıklık, anlaşılabilirlik ve vicdanlı olma alanlarını değerlendirmektedir. Bu üst düzey alanların her biri, sekiz madde ile değerlendirilen ve toplamda 240 maddeden oluşan altı özellikten oluşur. Ayrıca, ölçekteki maddeler kesinlikle katılmıyorum, katılmıyorum, kararsızım, katılıyorum ve kesinlikle katılıyorum yanıt seçenekleri şeklinde beşli likert tipindedir. İlk olarak, karma kısmi puan modeli ve sınırlandırılmış karma kısmi puan modelleri veri uyumu açısından karşılaştırılmıştır. Sınırlandırılmış karma kısmi puan modeli veriye iyi uyum göstermiş ve gizil sınıflar tepki stillerinde farklılık gösterirken, gizil değişkende farklılık göstermemiştir. Durumun böyle olduğu ölçekler için, katılımcıların her ölçekte uç tepki stili olmayan veya uç tepki stili olan üyelikleri bir gizil sınıf analizine girilmiştir. Her iki ölçme aracı için bu ikinci dereceden gizil sınıf analizi, tepki stilinin katılımcıların çoğu için gizil değişkenler arasında tutarlı olduğunu göstermiştir.

Eid ve Rauber'ın (2000) yaptığı çalışmada, örgütsel ölçeklerde ölçme değişmezliği sorunu tartışılmış ve örgütsel ölçeklerde tepki stillerini tespit etmek için karma dağılım modellerinin

nasıl kullanılabileceği gösterilmiştir. Bir liderlik performans ölçeğinin uygulandığı 4578 kişiden elde edilen veri üzerinden çok kategorili karma Rasch modeli analizinin sonuçları rapor edilmiştir. Sonuçlar, tepki stillerinde farklılık gösteren iki gizil sınıfın tespit edilebileceğini ortaya çıkarmıştır. Belirli kategoriler için güçlü bir tercih olmaksızın tüm yanıt ölçeğini kullanan bir sınıf (1. Boyut: % 71) ve uç tepki kategorilerini tercih eden ve orta nokta kategorilerinden kaçınma eğiliminde olanların bir sınıf (2. Boyut: % 29) olduğu ortaya çıkmıştır. Ayrıca, iki gizil sınıfın bireylere ilişkin diğer demografik değişkenler açısından farklılık gösterdiği görülmüştür.

Wetzel, Böhnke, Carstensen, Ziegler ve Ostendorf (2013) bireysel tepki stillerinin önemli olup olmadığını Karma Rasch modellerini kullanarak incelemişlerdir. Madde yanıtları, DMF etkilerinin yanı sıra tepki stillerinden de etkilenebilir. Çalışmada ölçme aracı olarak, Almanya’da uygulanan beş büyük kişilik alanını değerlendiren ölçme aracı (NEO-PI-R) kullanılmıştır. Araştırmanın örneklemini ise NEO-PI-R ölçeğinin uygulandığı klinik olmayan 11.724 katılımcının bir parçası oluşturmaktadır (Ostendorf & Angleitner, 2004). Araştırma kapsamında örneklem rastgele ikiye bölünmüştür. Örneklemin birinci yarısında analizler gerçekleştirilirken, ikinci yarısı ise analiz sonuçlarının doğrulanması amacıyla kullanılmıştır. Çalışmada kullanılan ölçek tepki stillerini dikkate alırken erkekler ve kadınlar için farklı madde işleyişine göre analiz etmektedir. Karma Rasch modelleri ile ilk olarak tepki stillerinde farklılık gösteren gizil sınıflar belirlenmeye çalışılmıştır. Buna göre uç tepki stili sergileyen ve uç tepki stili sergilemeyenler olmak üzere iki farklı gizil sınıf ortaya çıkmıştır. Daha sonra farklı tepki stilleri için ayrı ayrı DMF analizleri yapılmış ve tüm örneklem için DMF sonuçları ile de karşılaştırılmıştır. Özellikle Nevrotiklik, Uyumluluk ve Dürüstlük ölçeklerindeki birkaç madde cinsiyet açısından DMF göstermiş ve bu nedenle erkekler ve kadınlar arasında farklılıklar bulunmuştur. DMF sonuçları esas olarak tüm örneklem ve tepki stili alt örneklemeleri arasında boyut açısından farklılık göstermiştir; ancak DMF sınıflandırması uç tepki stili sergileyen alt örneklem, uç tepki stili sergilemeyen alt örneklem ve tüm örneklem için tutarlılık göstermiştir.

Madde Tepki Kuramı Ağacı (IRTTree) Modellerine Yönelik Araştırmalar

Madde Tepki Kuramı Ağacı (IRTTree) modelleri tek bir tepki stilinin ele alınmasında kullanılırken, bazı çalışmalarda ise birden fazla tepki stilleri de ele alınmıştır. IRTTree yöntemleri çok işlemli madde tepki kuramı modelleri olarak da tanımlanmaktadır. IRTTree

yöntemleri Genelleştirilmiş Doğrusal Karma Modeller'nin (GDKM) bir uzantısı olarak görülmektedir. IRTree modelleri daha esnek ve bilgilendirici bir yöntem olarak görülmektedir. Aşağıda IRTree yönteminin kullanıldığı araştırmalardan bazılarına yer verilmiştir.

Plieninger ve Meiser (2014) tepki stillerini ve ölçülen içeriğin birbirinden ayrılması konusunda IRTree yöntemini kullanmışlardır. IRTree yönteminin test edilmesi amacıyla iki büyük veri seti kullanılmıştır. İlk çalışmada yedi kategorili bir derecelendirme ölçeğine verilen yanıtlar çözümlenmiştir. Araştırmadan elde edilen sonuçlar, tepki stiliyle ilgili ve içerikle ilgili süreçlerin dış kriterlerle bağlantılı olduğunu göstermiştir. İkinci veri setinde dört kategorili bir derecelendirme ölçeği kullanılmıştır. Burada içerikle ilgili bir kritere odaklanılmış ve tepki stilinin önemli bir bastırma etkisi ortaya çıkarılmıştır. Araştırmadan elde edilen bulgular, tepki stillerini araştırmak ve kontrol etmek için IRTree modellerinin geçerliliği için güçlü kanıtlar sağlamıştır. Ayrıca, derecelendirme ölçeklerine verilen yanıtların altta yatan farklı boyutların çözülebileceğini ve ortaya çıkan boyutların psikolojik olarak anlamlı süreçleri yakaladığını göstermektedir.

Böckenholt ve Meiser (2017) yaptıkları çalışmada tepki stillerinin Karma Kısmi Puan Modeli ve IRTree yöntemleri ile nasıl analiz edileceğini göstermişlerdir. Tepki stilleri, sıralı veriler için Rasch modellerinde ve bu modellerin karma dağılım uzantılarında eşik parametreleri analiz edilerek değerlendirilebilir. Yanıt ölçeğinin ortaya çıkardığı temel özellikler ve tepki eğilimleriyle ilgili süreçleri çözmek için kullanılabilen çok süreçli madde tepki kuramı ağacı modelleri tarafından farklı bir çerçeve sağlanır. Çalışmada iki yaklaşım gözden geçirilmiş ve iki boyutlu “Yapı İçin Kişisel İhtiyaç (PNS)” ölçeğinden elde edilen veriler kullanılmış ve her iki modelde birden çok kriter açısından karşılaştırılmıştır. PNS ölçeğindeki maddeler; bir kişinin yapı, kesinlik ve kararlılık arzusunu ölçmek için tasarlanmıştır. Tipik olarak, yapıya yüksek ihtiyaç duyan katılımcılar netliği tercih ederler ve belirsiz durumları can sıkıcı ve sıkıntılı bulurlar. Karma Kısmi Puan Modeli ve IRTree modellerini resmetmek ve karşılaştırmak için, 12 PNS maddesine tam yanıt veren 2112 katılımcıdan oluşan bir örneklem kullanılmıştır. Veriler, Almanya'daki iki üniversitede 2003 ile 2011 yılları arasında gerçekleştirilen çeşitli deney ve anket çalışmalarında toplanmıştır. Her iki çerçevenin de tepki stillerini hesaba kattığı gösterilse de, model seçiminin nicel kriterleri, model tahmininin pratik yönleri ve psikolojik değerlendirmedeki bireysel farklılıkların sürekli ve çok boyutlu kaynakları olarak tepki stillerini temsil etmenin kavramsal sorunları bakımından farklılık göstermişlerdir. Karma Kısmi Puan Modeli ve

IRtree modelleri her iki perspektife de hizmet edebilir; yani, hedef özelliklerini tepki stillerinden ayırırlar ve kişinin tepki stillerini örtük sınıf veya örtük özellik değişkenleri açısından değerlendirilmesine izin verirler. Araştırmada belirlenen ve test edilen farklı model varyansları, her iki model ailesinin de tepki stili eğilimlerini ve tutum hakkındaki varsayımlarını araştırmak için nasıl kullanılabileceğini göstermektedir. Bu nedenle, Karma Kısmi Puan Modelleri ve IRTree modelleri, ilgili psikolojik fenomenler ve yargısal davranışlardaki bireysel farklılıkların kaynakları olarak tepki stilleri üzerine araştırmayı ilerletmek için istatistiksel araçlar olarak kullanılabilirler.

Park ve Wu (2019) tarafından yapılan araştırmada, likert tipi maddelerden oluşan derecelendirme ölçeklerinde KTS ve UTS' yi incelemek amacıyla IRTree modellerini kullanmışlardır. Araştırmada spesifik olarak, daha geniş bir modelleme çerçevesi altında algılanan, açıklayıcı madde yanıt modelleri adı verilen, De Boeck ve Wilson (2012) tarafından önerilen, tanımlayıcı ve açıklayıcı modeller olmak üzere iki tür IRTree modeli tanıtılmıştır. Araştırmada açık kaynaklı olan California'daki Boylamsal 2005 Nesil Araştırması (Silverstein & Bengtson, 2008) verisindeki 1596 kişiden elde edilen veriler kullanılmıştır. Araştırmada 10 maddeden oluşan Rosenberg'in Benlik Saygısı Ölçeği' ndeki maddelere verilen yanıtlar incelenmiştir. Araştırmadan elde edilen bulgular iki farklı UTS ve KTS' nin varlığına işaret etmektedir.

İlgün Dibek (2019) IRTree yöntemini kullanarak TIMSS 2015 uygulamasına katılan öğrencilerin matematiğe yönelik tutumları üzerinde UTS etkisini incelemiştir. Araştırmada ayrıca, Genelleştirilmiş Lineer Karma Model (GLMM) yaklaşımına (Rastgele kesişim, UTS' nin kesişim etkisine sahip rastgele model ve rastgele kesişim-eğim modeli) dayalı olarak hangi IRTree modellerinin TIMSS 2015 verilerine uyum gösterdiği belirlenmeye çalışılmıştır. Araştırmanın örneklemini TIMSS 2015'e katılan ülkeler arasından tutum-başarı paradoksunun görüldüğü 15 ülke oluşturmaktadır. Araştırma sonuçlarına göre model uyum indeksleri değerlendirildiğinde, rastgele kesişme eğimi modelinin verilere en uygun model olduğu görülmüştür. Bu modele göre; UTS, öğrencilerin matematiğe yönelik tutumlarında önemli miktarda farklılığı açıklamaktadır. Ayrıca, matematiğe karşı olumsuz bir tutuma sahip öğrencilerin daha fazla UTS' ye sahip oldukları anlaşılmıştır.

Jeon ve De Boeck (2019) IRTree yaklaşımı ile UTS yanlılığının incelenmesinde değişmezlik türleri üzerine bir değerlendirmede bulunmuşlardır. Uç tepki stillerini ölçmek için bir IRTree yaklaşımı benimsendiğinde, gizil değişken ve madde parametresi seviyelerinde yönsel (directional) ve içerik değişmezliği varsayılabilir. Bu çalışmada, bu değişmezlik

varsayımlarının deneysel geçerliliğini, genişletilmiş değişmezlik varsayımlarına sahip genel bir IRTree modeli kullanarak değerlendirilmiştir. Değerlendirme sonuçlarına göre, üç deneysel uygulamayla, bazı değişmezlik varsayımlarını gevşetmenin model uyumunu iyileştirdiği görülmüştür. Bu da tüm varsayılan değişmezliklerin deneysel olarak desteklenmediğini göstermektedir. Spesifik olarak, gizil değişken düzeyinde, yönsel değişmezlik için olası kanıtlar bulunmuştur. Ancak içerik değişmezliği için karışık kanıtlar elde edilmiştir. İçeriğe özgü uç tepki gizil değişkenleri arasındaki tahmini korelasyonların da yüksek olduğu ve genel bir uç tepki stili eğiliminin potansiyel varlığına işaret ettiği görülmüştür. Madde parametresi düzeyinde, eşikler için yönsel veya içerik değişmezliği elde edilemezken, eğimler içinde içerik değişmezliği bulunamamıştır. Üç deneysel uygulama ile bazı değişmezlik varsayımlarını gevşeten genel IRTree modellerinin verilere, değişmezlik varsayımlarına sahip standart IRTree modelinden daha iyi uyduğu görülmüştür.

Tepki stillerinin incelenmesinde kullanılan DFA, GSA ve MTK gibi model tabanlı yaklaşımların yanında bir diğer prosedür de bu araştırmada da kullanılan TSYDK yöntemidir. Takagishi vd. (2019) tarafından geliştirilen TSYDK yöntemi, verilerdeki tepki stili yanlılığının düzeltilmesini ve verilerin içerik tabanlı kümelenmesini sağlamaktadır. Takagishi vd. (2019) yaptıkları çalışmada simülasyon ve gerçek veri setlerini kullanarak geliştirdikleri TSYDK yönteminin avantajlarını açıklamışlardır. TSYDK yöntemi ile tepki stili yanlılığı bulunan verilerde tepki stili yanlılığı düzeltildikten sonra verilerin kümelenmesi yapılmaktadır. TSYDK yöntemini diğer tepki stili yanlılığının düzeltilmesinde kullanılan yöntemlerden ayıran en önemli özelliği tepki stili yanlılığın düzeltilmesi ve kümeleme işlemlerinin eş zamanlı olarak yapılabilmesine olanak sağlamasıdır. TSYDK yönteminin dışında Schoones vd. (2015) tarafından önerilen ve kategorik verilerdeki tepki stillerinin incelenmesinde kullanılan Sınırlandırılmış Çift Ölçeklendirme (SÇÖ) (Constrained Dual Scaling-CDS-) yöntemi de bulunmaktadır. Schoones vd. (2015) yaptıkları çalışmada, tepki stillerindeki farklılıkları tespit etmek için SÇÖ yönteminden yararlanılabileceğini göstermişlerdir. SÇÖ yöntemi, dört belirgin tepki stiline izin verecek şekilde genişletilebilen bir yöntemdir. SÇÖ’de parametreleri tahmin etmek için alternatif negatif olmayan en küçük kareler algoritması kullanılmaktadır. Yapılan çalışma ile SÇÖ yöntemi hem simülasyon hem de deneysel bir uygulama ile değerlendirilmiştir.

BÖLÜM III

YÖNTEM

Bu bölümde, araştırmanın modeline, evren ve örnekleme, veri toplama araçları ile verilerin analizine yönelik bilgilere ve açıklamalara yer verilmiştir.

Araştırmanın Modeli

Araştırmada PISA 2015 uygulamasına katılan ülkelerde tepki stili yanlılığı açısından farklılık olup olmadığı, tepki stili yanlılığının düzeltilmesi ile düzeltilmiş veriler aracılığıyla öğrencilerin kümelenmesi amaçlanmıştır. Bu amaç doğrultusunda araştırma korelasyonel desende yürütülmüştür. Korelasyonel desende, bağımsız değişkenin düzeyleri araştırmacı tarafından manipüle edilemez. Araştırmacı bağımsız değişkeni tanımlayabilir, ancak bu değişkenlerin düzeylerine ayrılması üzerinde hiçbir kontrole sahip değildir. Bu tür araştırmalarda, bağımsız değişken yordayıcı ve bağımlı değişken ölçüt olarak tercih edilir (Tabachnick & Fidell, 2013).

Evren ve Örneklem

Araştırmanın evrenini, PISA 2015 uygulamasına katılan ülkelerin 15 yaş grubu öğrencileri oluşturmaktadır. PISA uygulamasında veriler iki aşamalı tabakalı örnekleme yöntemi kullanılarak elde edilmektedir. Buna göre örneklem seçiminin birinci aşamasında okullar, ikinci aşamada ise bu okullarda yer alan sınıflar seçilmektedir. PISA'da örneklem belirleme işlemi yapılırken, ulusal merkezler aracılığıyla örneklemin ve okulların seçilme kriterleri (bölge, program türü, okul türü) düzenlenir. Düzenlenen tabakalara uygun olarak okul ve öğrenci sayılarının yer aldığı örneklem dosyası oluşturulur ve bu dosya ulusal merkezler

aracılığıyla uluslararası kuruluşlara gönderilir. Dosyada belirtilen tabakalara uygun olacak şekilde ve öğrenci oranlarına göre okullar seçilip dosya ülkelere tekrar geri gönderilir. Ülkeler belirlenen okullardan sınıf düzeyi 7. sınıf ve üzeri olan 15 yaş grubu öğrencilerden seçkisiz olacak şekilde 35 öğrenciyi ve öğrenci sayısı 35'ten az olan okullardan tüm 15 yaş grubundaki öğrencileri PISA uygulaması için seçer. Son aşamada oluşturulan bu listeler okullara gönderilip onların da onayı alındıktan sonra her bir ülke için PISA değerlendirmesinde yer alacak olan okullar ve öğrenciler belirlenmiş olur (MEB, 2016).

Araştırma kapsamında ülke örneklemeleri ve çalışma örneklemeleri belirlenirken takip edilen işlem adımları aşağıdaki gibidir.

1. PISA 2015 öğrenci verisi OECD internet sitesinden indirilmiş ve fen yönelik tutum ölçekleri belirlenmiştir. Daha sonra bu ölçeklerdeki maddeler tek tek kontrol edilerek kayıp veriler silinmiştir.
2. Araştırmada ülkelerin seçilme kriteri olarak tutum-başarı paradoksu segileme durumları göz önüne alınmıştır. Tutum-başarı paradoksunda ülke içi fene yönelik tutum ile fen başarısı arasındaki korelasyonların pozitif, ülkeler arasında ise fene yönelik tutum ile fen başarısı arasındaki korelasyonun negatif olması gerekmektedir. Bunun için ilk olarak, ülke düzeyinde IDB Analyzer programı aracılığıyla fene yönelik tutum ölçekleri ile fen başarısı için ağırlıklandırılmış ortalama puanlar hesaplanmıştır.
3. Hesaplanan ülke ortalama puanları ile fene yönelik tutum ölçeklerinin her biri için ülkeler arasında fene yönelik tutum ile fen başarı arasındaki korelasyon değerleri elde edilmiştir. Bu korelasyon değerlerinden pozitif olan ölçekler analiz dışı bırakılmıştır. Ayrıca, geriye kalan ölçekler üzerinden ülke içinde tutum ile başarı arasında korelasyonlarda hesaplanmıştır. Tutum ile başarı arasındaki korelasyon değerinin negatif olduğu ülkelerin bulunduğu ölçeklerde analiz dışı bırakılmıştır. Araştırma kapsamında; Fenden hoşlanma, Başarı motivasyonu ve Fen öz yeterliği ölçekleri ile kullanılmıştır.
4. Tutum-başarı paradoksunun olduğu ölçekler belirlendikten sonra araştırmaya dâhil edilecek ülkeleri seçmek amacıyla ülkelerin fene yönelik tutum ölçekleri ortalamaları ile fen başarı puan ortalamalarının saçılım diyagramları çizilmiştir.
5. Her üç ölçek içinde saçılım diyagramları oluşturulduktan sonra; birinci grupta, fene yönelik tutumu olumlu ve fen başarısı düşük olan ülkeler, ikinci grupta, fene yönelik tutumu olumsuz ve fen başarısı yüksek olan ülkeler ve üçüncü grupta ise fene yönelik

tutumu orta düzeyde ve fen başarısı orta düzeyde olan ülkelerin bulunduğu gruplar oluşturulmuştur. Oluşturulan her bir gruptan yedişer ülke seçilerek toplam 21 ülke araştırmaya dahil edilmiştir.

6. Araştırmaya dahil edilen 21 ülke içinde yeniden ülkeler arası ve ülke içi fene yönelik tutum ile fen başarısı arasındaki korelasyonlar hesaplanmıştır.
7. Analizlerin yapıldığı bilgisayarın işle yapabilme kapasitesi yeterli olmadığı için kayıp verilerinin silinmesinin ardından her bir ülke verisinden rastgele örneklem seçme yöntemi kullanılarak %10'luk çalışma örneklemeleri oluşturulmuştur.
8. Çalışma örneklemlerinin ülke örneklemelerini temsil edip etmediklerini belirlemek amacıyla ülkelerin örneklem ve çalışma örneklemlerinin fene yönelik tutum ölçekleri ortalama puanları ilişkisiz örneklem t testleri ile karşılaştırılmış ve fen başarısı ile fen yönelik tutum ölçekleri arasındaki ülke içi korelasyonlar hesaplanmıştır.

Ülkelerdeki öğrenci sayılarının fazla olması ve kayıp verilere çoklu atamanın yapılmasının uç tepki stilini etkileyeceği düşünülerek liste bazında veri silme işlemi yapılmıştır. Kayıp verilerin silinmesinin ardından her bir ülke için kalan veri sayısı Ek 2'de yer gösterilmektedir. Ülkeler seçilirken tutum-başarı paradoksu sergileme durumları göz önüne alınmıştır. Araştırmada bu paradoksu sergileyen ülkeler arasından 21 tanesi seçilmiştir. Tutum-Başarı paradoksu, bir konu alanına yönelik olumlu tutum sergileyen öğrencilerin beklenilenin aksine akademik başarıları düşük olurken, tutumları olumsuz olan öğrencilerin ise başarılarının yüksek olması durumudur. Buna göre tutum-başarı paradoksu, ülkeler arası tutum-başarı korelasyon sonuçlarının negatif, ülke içi tutum-başarı korelasyon değerlerinin pozitif olmasıdır.

PISA gibi geniş ölçekli test verilerinde dikkate alınması gereken iki temel konu vardır. Bunlar, örneklem seçme metoduna bağlı olarak çıkan örneklem ağırlıkları ve öğrencilerin başarılarını belirlemek için kullanılan olası değerlerdir (Rutkowski, Gonzalez, Joncas, & von Davier, 2010). Araştırmalarda kullanılan örneklem alındığı evrenden farklı olma olasılığı problemini çözmek amacıyla kullanılan yöntemlerden birisi örneklem ağırlıklarının kullanılmasıdır (Rust, 2013). Ulusal düzeyde Türkiye İstatistik Kurumu (TÜİK)'in yapmış olduğu olduğu çalışmalarda ve PISA, TIMSS ve PIRLS gibi geniş ölçekli testlerde istatistiksel analizler için örneklem ağırlıklandırması kullanılmaktadır. Ayrıca, yapılan birçok çalışmada da örneklem ağırlıklandırması kullanılmadan yapılan çalışmalarda istatistiksel olarak yanlı sonuçların elde edildiği görülmüştür (Arıkan, Özer, Şeker ve Ertaş, 2020; Tat, Koyuncu ve Gelbal, 2019; Rutkowski vd., 2010). Rastgele olmayan örneklem

seimlerinde yanlışlık olasılığı yüksek olduėu gibi, bazen rastgele rneklem seme ileminde yanlışlık olabilir ve her bir bireyin rneklem seilme ansı eit olmayabilir. Byle yanlışlık durumlarından kaınmak ve her bireyin evrendeki seilme olasılıklarını eitleyebilmek iin rneklem aėırlıkları kullanılması nerilmektedir (Arıkan vd., 2020). Grupların puan ortalamalarını hesaplamak iin rneklem aėırlıkları hesaba katıldığında daha doėru kestirimler yapılabilmekte ve evren temsiliyeti korunmu olur (Gonzales, 2012). Bu aratırmada da yanlışlıktan kaınmak ve evren temsiliyetini saėlamak iin leklerin ortalamalarını hesaplariken rneklem aėırlıklandırmasından yararlanılmı ve her bir ėrenci iin aėırlıklandırılmı rneklem ortalaması alınmıtır.

PISA gibi geni lekli sınavlarda bireylerin aldıkları puandan ziyade evrenin veya ierisindeki alt grupların performansı tanımlandığı iin (Monseur & Adams, 2009; Von Davier, Gonzalez, & Mislevy, 2009) bireyler iin tutarlı ve geerli baarı puanlarını hesaplamak bu testlerin amacı deėildir. Ayrıca, bireylere tm testler uygulanmak yerine belli sayıda test uygulanmaktadır. Bylece her bir ėrenci soruların belli bir kısmını cevaplandırmaktadır. Bu nedenle geni lekli testlerde, MTK yntemleri kullanılarak baarı test puanları kestirilmektedir. Her bir ėrenci iin birden fazla olası deėer (plausible values) elde edilmektedir. ėrencilerin fen baarılarını belirlemek amacıyla PISA 2015’de her bir ėrenci iin 10 olası (plausible) deėer hesaplanmıtır. Bu aratırmada hem bilgisayarın ilem yapabilme kapasitesi hem de sonuların yorumlarının anlaılabilirliėi aısından 10 olası deėerin aėırlandırılmı rneklem ortalaması alınarak tek bir ortalama puan kullanılmıtır. Aynı ekilde her bir lke iinde aėırlıklandırılmı rneklem ortalaması alınarak tek bir fen baarı puan ortalaması hesaplanmıtır. rneklem aėırlıklandırması hesaplanırken, ėrenci aėırlıklandırmaları gz nne alınmıtır. Tutum-Baarı paradoksunu belirlemek iin lke ii ve lkeler arası korelasyonlar hesaplanmıtır. lkeler arasında hesaplanan fene ynelik tutum ile fen baarısı arasındaki korelasyon deėerleri Tablo 1’de gsterilmektedir.

Tablo 1

Fene Ynelik Tutum ile Fen Baarısı Arasındaki lke Ortalama Puanları Arasındaki (lkeler Arası) Korelasyon Deėerleri

	Fene Ynelik Tutum lekleri					
	Fenden Holanma	Arasal Motivasyon	Test Kaygısı	Baarı Motivasyonu	Fen z Yeterliėi	Epistomolojik İnan
Fen Baarısı	-,491	-,573	,337	-,429	-,365	,451

Tablo 1 incelendiğinde, ülkeler arasında yapılan korelasyon analizi bulgularına göre; Fen Başarısı ile Fenden Hoşlanma, Araçsal Motivasyon, Başarı Motivasyonu ve Fen Öz Yeterliği ölçekleri arasında negatif korelasyon değerleri elde edilmiştir. Ayrıca Test Kaygısı ve Epistemolojik İnanç ölçeği ile Fen Başarısı arasında pozitif korelasyon bulunmuştur. Bu nedenle Test Kaygısı ve Epistemolojik İnanç ölçeklerinde Tutum-Başarı paradoksu görülmediği kabul edilerek analiz dışı bırakılmışlardır. Geriye kalan dört ölçek ile ülke içinde Fene yönelik tutum ile fen başarısı arasındaki korelasyonlar hesaplanmıştır. Yapılan korelasyon analizi sonucuna göre, bazı ülkelerde Araçsal Motivasyon ile Fen Başarısı arasındaki korelasyon değerleri negatif bulunmuştur (Bulgaristan, Kolombiya, Kosta-Rika, Dominik Cumhuriyeti, Hırvatistan, Macaristan, Karadağ, Peru, Rusya, Uruguay). Tutum-başarı paradoksunun görülmesi için ülke içi korelasyon değerlerinin pozitif olması gerekmektedir. Buna göre Araçsal Motivasyon ölçeği için yukarıdaki ülkelerde negatif korelasyonlar elde edildiğinden bu ölçekte analiz dışı bırakılmıştır. Böylece, araştırmada Tutum-Başarı paradoksunun görüldüğü ölçekler olarak Fenden Hoşlanma, Başarı Motivasyonu ve Fen Öz Yeterliği ölçekleri kalmıştır. Ülkeler arası ve ülke içinde Fen Başarısı ile Fenden Hoşlanma, Başarı Motivasyonu ve Fen Öz Yeterliği ölçekleri arasındaki korelasyon değerleri Tablo 2’de gösterilmektedir.

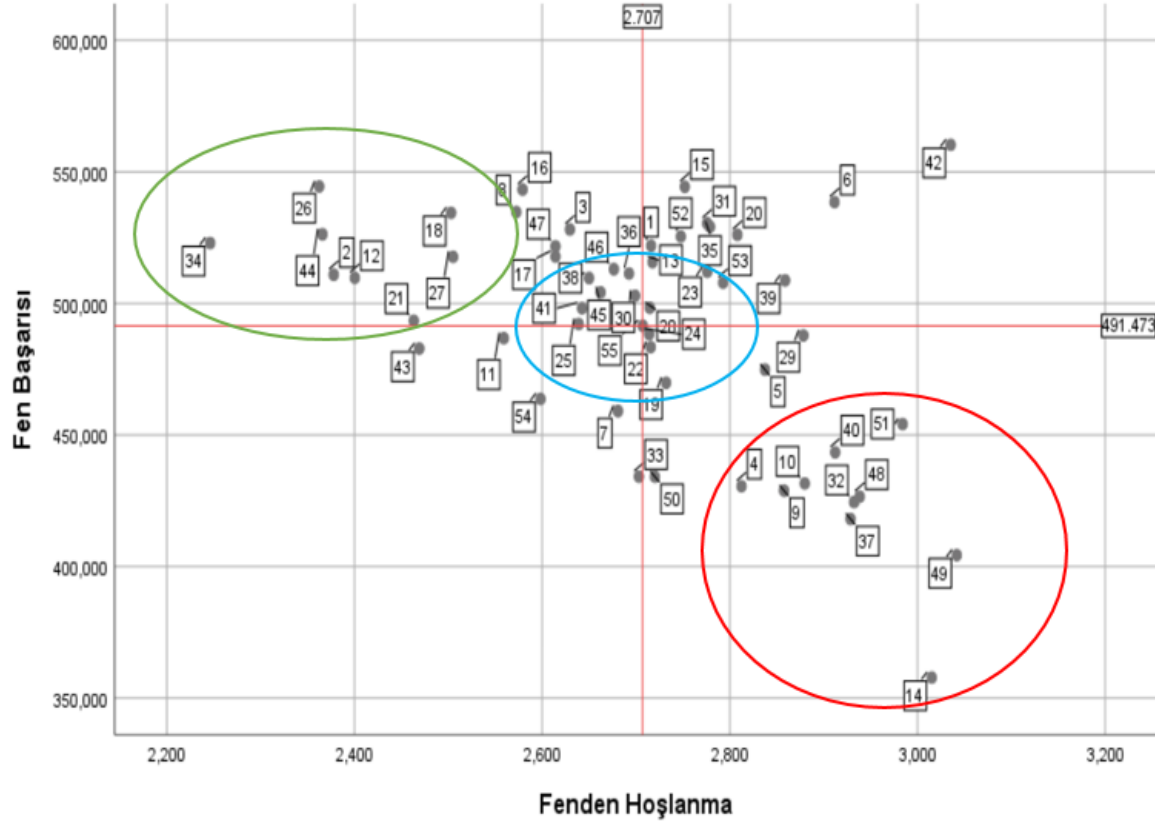
Tablo 2

Fene Yönelik Tutum ile Fen Başarısı Arasındaki Ülkeler Arası ve Ülke İçi Korelasyon Değerleri

Ölçekler	Fenden Hoşlanma		Başarı Motivasyonu		Fen Öz Yeterlik	
	Ülkeler Arası	Ülke İçi	Ülkeler Arası	Ülke İçi	Ülkeler Arası	Ülke İçi
Başarı	-,491	,266	-,429	,160	-,365	,235

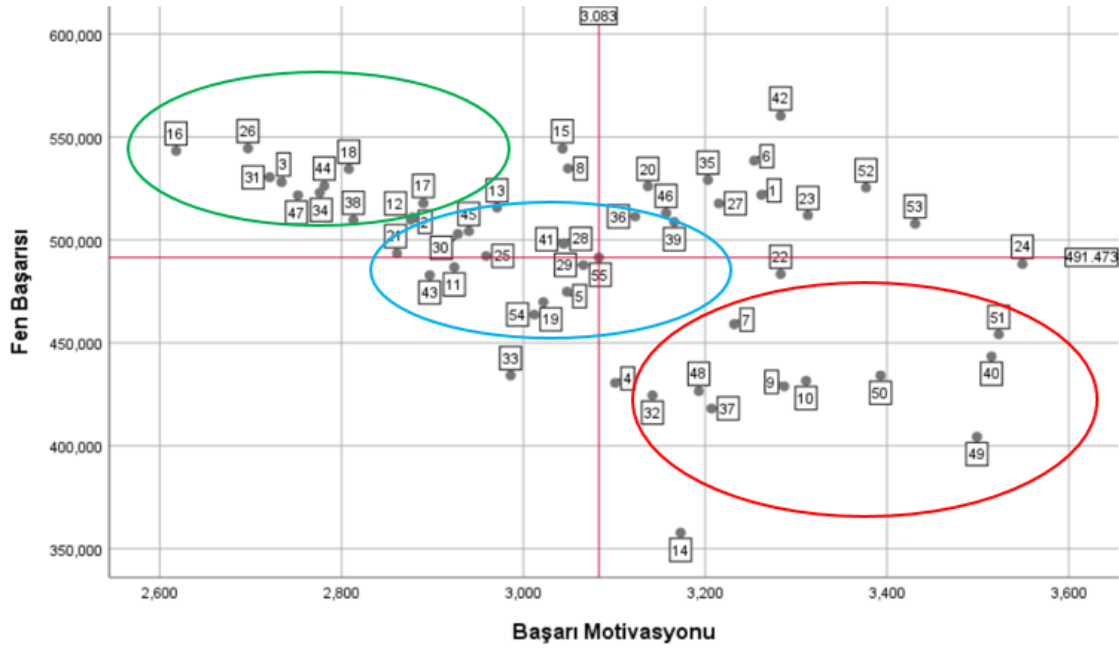
Tablo 2 incelendiğinde, ülkeler arası hesaplanan fen başarısı ile fenden hoşlanma, başarı motivasyonu ve fen öz yeterliği ölçekleri arasında negatif korelasyonlar elde edilirken, ülkeler içinde fen başarısı ile fenden hoşlanma, başarı motivasyonu ve fen öz yeterlik ölçekleri arasında pozitif korelasyonlar bulunmuştur. Buna göre fenden hoşlanma, başarı motivasyonu ve fen öz yeterlik ölçekleri üzerinden PISA 2015 uygulamasına katılan ülkelerde tutum-başarı paradoksunun görüldüğü söylenebilir. Araştırmaya dâhil edilecek ülkeleri belirlemek amacıyla ülkelerin fen başarı puanları ile fenden hoşlanma, başarı motivasyonu ve fen öz yeterliği ölçeklerinin ağırlıklandırılmış örneklem ortalamaları alınarak saçılım diyagramları elde edilmiştir. Fen başarısı ile fenden hoşlanma ölçeği için elde edilen saçılım diyagramı Şekil 5’de, fen başarısı ile başarı motivasyonu ölçeği için elde

edilen saçılım diyagramı Şekil 6’da ve fen başarısı ile fen öz yeterliği ölçeği için elde edilen saçılım diyagramı Şekil 7’de gösterilmektedir.

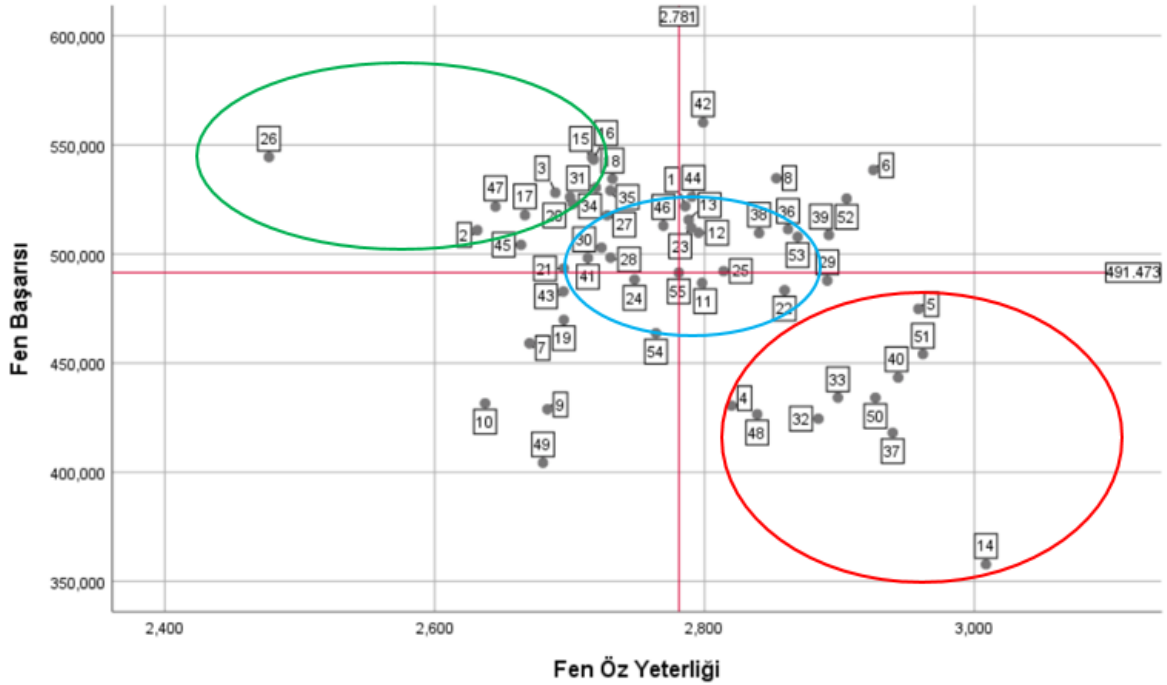


Şekil 5. Ülkeler arası fenden hoşlanma ile fen başarısı arasındaki saçılım diyagramı.

Şekil 5’deki grafikte y-ekseninde ülkelerin fen başarı puan ortalamaları gösterilirken, x-ekseninde ülkelerin fenden hoşlanma ölçeğinden aldıkları ortalama puanları gösterilmektedir.



Şekil 6. Ülkeler arası başarı motivasyonu ile fen başarısı arasındaki saçılım diyagramı. Şekilde 6'daki grafikte y-ekseninde ülkelerin fen başarı puan ortalamaları gösterilirken, x-ekseninde ülkelerin başarı motivasyonu ölçeğinden aldıkları ortalama puanları gösterilmektedir.



Şekil 7. Ülkeler arası fen öz yeterliği ile fen başarısı arasındaki saçılım diyagramı.

1-Avustralya 2-Avusturya 3-Belçika 4-Brezilya 5-Bulgaristan 6-Kanada 7-Şili 8-Çin 9-Kolombiya 10-Kosta Rika 11-Hırvatistan 12-Çek Cumhuriyeti 13-Danimarka 14-Dominik Cumhuriyeti 15-Estonya 16-Finlandiya 17-Fransa 18-Almanya 19-Yunanistan 20-Hong

Kong 21-Macaristan 22-İzlanda 23-İrlanda 24- İsrail 25-İtalya 26-Japonya 27-Kore 28-Letonya 29-Litvanya 30-Lüksemburg 31-Makao 32-Meksika 33-Karadağ 34-Hollanda 35-Yeni Zellanda 36-Norveç 37-Peru 38-Polonya 39-Portekiz 40-Katar 41-Rusya Federasyonu 42-Singapur 43-Slovak Cumhuriyeti 44-Slovenya 45-İspanya 46-İsveç 47-İsviçre 48-Tayland 49-Birleşik Arap Emirlikleri 50-Tunus 51-Türkiye 52-Birleşik Krallık 53-Amerika Birleşik Devletleri 54-Uruguay

Şekilde 7’deki grafikte y-ekseninde ülkelerin fen başarısı puan ortalamaları gösterilirken, x-ekseninde ülkelerin fen öz yeterliği ölçeğinden aldıkları ortalama puanları gösterilmektedir.

Her üç saçılma diyagramı da incelendiğinde kırmızı, mavi ve yeşil renkli olmak üzere üç dairenin yer aldığı görülmektedir. Bu daireler oluşturulurken ülkelerin fen başarı puan ortalamaları ile tutum ölçeklerinden aldıkları puan ortalamaları dikkate alınmıştır. Kırmızı renkli dairede yer alan ülkeler birinci grubu oluşturmakta ve bu ülkelerin fene yönelik tutumu olumlu ve fen başarısı düşüktür. Yeşil renkli dairede yer alan ülkeler ikinci grubu oluşturmakta ve bu ülkelerin fene yönelik tutumu olumsuz ve fen başarısı yüksektir. Son olarak mavi renkli daireyi oluşturan üçüncü gruptaki ülkeler ise fene yönelik tutumu orta düzeyde ve fen başarısı orta düzeyde olan ülkelerdir. Her üç saçılım diyagramı da incelenerek her grupta yedi ülke olmak üzere toplam 21 ülke araştırmaya dâhil edilmiştir. Bu doğrultuda her bir grup için seçilen ülkeler şu şekildedir:

1.Grup (*Fene Yönelik Tutumu Olumlu-Fen Başarısı Düşük*)= 14-Dominik Cumhuriyeti, 37-Peru, 32-Meksika, 49-Birleşik Arap Emirlikleri, 40-Katar, 4-Brezilya, 51-Türkiye.

2.Grup (*Fene Yönelik Tutumu Olumsuz-Fen Başarısı Yüksek*)= 26-Japonya, 16-Finlandiya, 3-Belçika, 34-Hollanda, 17-Fransa, 2-Avusturya, 18-Almanya.

3.Grup (*Fene Yönelik Tutumu Orta Düzeyde-Fen Başarısı Orta*)= 13-Danimarka, 28-Letonya, 41-Rusya Federasyonu, 24-İsrail, 25-İtalya, 11-Hırvatistan, 30- Lüksemburg.

Belirlenen 21 ülke için yeniden tutum-başarı paradoksunu göstermek amacıyla ülkeler arası ve ülke içi korelasyon değerleri hesaplanmış ve elde edilen bulgular Tablo 3’de gösterilmiştir.

Tablo 3

Belirlenen Ülkelerde Fene Yönelik Tutum Ölçekleri ile Fen Başarısı Arasındaki Ülkeler Arası ve Ülke İçi Korelasyon Değerleri

Ölçekler	Fenden Hoşlanma		Başarı Motivasyonu		Fen Öz Yeterlik	
	Ülkeler Arası	Ülke İçi	Ülkeler Arası	Ülke İçi	Ülkeler Arası	Ülke İçi
Başarı	-,855	,249	-,849	,147	-,854	,217

Tablo 3 incelendiğinde, ülkeler arası fen başarısı ile fene yönelik tutum ölçekleri arasındaki korelasyon değerlerinin negatif olduğu görülmektedir. Buna karşın, ülke içi hesaplanan fen başarısı ile fene yönelik tutum ölçekleri arasındaki korelasyonlar pozitif bulunmuştur. Buna göre fenden hoşlanma, başarı motivasyonu ve fen öz yeterlik ölçekleri üzerinden çalışmaya dahil edilen ülkelerde tutum-başarı paradoksunun görüldüğü söylenebilir. Analizlerin yapıldığı bilgisayarın işlem yapabilme kapasitesi de göz önüne alınarak her ülkedeki örneklemden %10'luk rastgele örneklem seçme yöntemi kullanılarak çalışma örneklemeleri elde edilmiştir. Her bir ülkenin örneklem ve çalışma örneklemeleri Tablo 4'de gösterilmektedir.

Tablo 4

Araştırmada Yer Alan Ülkelerin Örneklem ve Çalışma Örneklemeleri

Ülkeler	Örneklem	Çalışma Örneklemi
Avusturya (AUT)	4.937	470
Belçika (BEL)	6.705	651
Brezilya (BRA)	8.901	885
Hırvatistan (HRV)	4.513	425
Danimarka (DNK)	5.197	503
Dominik Cumhuriyeti (DOM)	2.336	219
Finlandiya (FIN)	4.718	445
Fransa (FRA)	4.322	409
Almanya (DEU)	2.590	243
İsrail (ISR)	4.742	448
İtalya (ITA)	8.919	888
Japonya (JPN)	5.834	574
Letonya (LVA)	3.924	377
Lüksemburg (LUX)	3.651	349
Meksika (MEX)	5.598	548
Hollanda (NLD)	4.356	413
Peru (PER)	4.639	440
Katar (QAT)	7.508	739
Rusya Federasyonu (RUS)	4.489	422
Birleşik Arap Emirlikleri (ARE)	10.165	1.006
Türkiye (TUR)	4.716	444
Toplam	112.760	10.898

Tablo 4 incelendiğinde, araştırmada yer alan ülke örneklemelerindeki toplam öğrenci sayısı 112.760 ve çalışma örneklemelerindeki toplam öğrenci sayısı 10.898' dir. Çalışma örneklemelerinin ülke örneklemelerini temsil edip etmediklerini belirlemek amacıyla ülke örneklemeleri ve çalışma örneklemelerinin fene yönelik tutum ölçeklerinin ağırlıklandırılmış örneklem ortalamaları $p=,01$ anlamlılık düzeyinde karşılaştırılmıştır. Ortalamaları karşılaştırmak için ilişkisiz örneklem t testi kullanılmıştır. Elde edilen bulgular Tablo 5'de

gösterilmektedir. Ayrıca her bir ülke örnekleme ve çalışma örnekleme ilişkin ölçek ortalamaları Ek 3’de yer almaktadır.

Tablo 5

Ülkelerin Örneklem ve Çalışma Örneklemelerinin Fene Yönelik Tutum Ölçekleri Ortalama Puanlarına İlişkin İlişkisiz Örneklem t Testi Bulguları

Ülkeler	Grup	Fenden Hoşlanma				Başarı Motivasyonu			Fen Öz Yeterliği		
		n	Ort.	t	p	Ort.	t	p	Ort.	t	p
Avusturya	Örneklem	4.937	2,38	-,495	,621	2,88	,036	,971	2,63	-1,142	,254
	Çalışma Örn.	470	2,36			2,88			2,65		
Belçika	Örneklem	6.705	2,63	-,173	,862	2,73	-,864	,382	2,69	-,290	,772
	Çalışma Örn.	651	2,67			2,76			2,71		
Brezilya	Örneklem	8.901	2,81	1,298	,194	3,10	,758	,448	2,82	,739	,460
	Çalışma Örn.	885	2,82			3,09			2,82		
Hırvatistan	Örneklem	4.513	2,56	-1,043	,297	2,92	-,990	,322	2,80	1,054	,292
	Çalışma Örn.	425	2,60			2,95			2,77		
Danimarka	Örneklem	5.197	2,72	-,147	,883	2,97	-,061	,951	2,79	-1,384	,167
	Çalışma Örn.	503	2,72			2,98			2,85		
Dominik Cum.	Örneklem	2.336	3,02	,273	,785	3,17	-,581	,561	3,01	,307	,759
	Çalışma Örn.	219	3,00			3,19			2,99		
Finlandiya	Örneklem	4.718	2,58	-,128	,898	2,62	,508	,611	2,72	-,654	,513
	Çalışma Örn.	445	2,58			2,60			2,74		
Fransa	Örneklem	4.322	2,61	,325	,745	2,89	-,802	,423	2,67	1,370	,171
	Çalışma Örn.	409	2,61			2,91			2,63		
Almanya	Örneklem	2.590	2,50	,517	,605	2,81	,014	,989	2,73	-,131	,896
	Çalışma Örn.	243	2,47			2,81			2,75		
İsrail	Örneklem	4.742	2,71	,723	,470	3,55	,070	,944	2,75	-,265	,791
	Çalışma Örn.	448	2,69			3,55			2,77		
İtalya	Örneklem	8.919	2,64	,515	,606	2,96	,978	,328	2,81	-,824	,410
	Çalışma Örn.	888	2,64			2,97			2,83		
Japonya	Örneklem	5.834	2,36	-1,189	,235	2,70	-,075	,940	2,48	-,319	,750
	Çalışma Örn.	574	2,40			2,71			2,49		
Letonya	Örneklem	3.924	2,71	-,689	,491	3,04	,129	,897	2,73	-1,071	,284
	Çalışma Örn.	377	2,74			3,06			2,77		
Lüksemburg	Örneklem	3.651	2,70	1,454	,146	2,93	,680	,497	2,72	-,399	,690
	Çalışma Örn.	349	2,63			2,90			2,74		
Meksika	Örneklem	5.598	2,93	-1,271	,204	3,14	-1,215	,224	2,88	-,450	,653
	Çalışma Örn.	548	2,96			3,18			2,88		
Hollanda	Örneklem	4.356	2,25	,073	,942	2,78	1,092	,275	2,70	-1,115	,265
	Çalışma Örn.	413	2,24			2,75			2,74		
Peru	Örneklem	4.639	2,93	-,792	,428	3,21	-1,000	,317	2,94	,370	,711
	Çalışma Örn.	440	2,95			3,22			2,92		
Katar	Örneklem	7.508	2,91	,484	,628	3,51	1,323	,186	2,94	,434	,664
	Çalışma Örn.	739	2,90			3,49			2,94		
Rusya Fed.	Örneklem	4.489	2,64	-,294	,768	3,04	,312	,755	2,71	-,926	,354
	Çalışma Örn.	422	2,65			3,01			2,76		
Birleşik Arap Emir.	Örneklem	10.165	2,98	1,344	,179	3,52	,063	,949	2,96	,487	,626
	Çalışma Örn.	1.006	2,99			3,55			2,97		
Türkiye	Örneklem	4.716	2,72	-,897	,370	3,39	,030	,976	2,93	,754	,451
	Çalışma Örn.	444	2,72			3,40			2,90		

Tablo 5 incelendiğinde, her üç ölçek içinde ülkelerin örneklem ve çalışma örneklemeleri arasındaki ilişkisiz örneklem t testi değerlerinin anlamlı olmadığı görülmektedir. Buna göre, her üç ölçek içinde ülkelerin örneklem ve çalışma örneklemelerin ortalamalarının birbirine yakın olduğu ve çalışma örneklemelerinin ülke örneklemelerini temsil ettiği söylenebilir. Çalışma örneklemelerinin ülke örneklemelerini temsil edip etmediklerini incelemek için ayrıca fen başarısı ile fene yönelik tutum ölçekleri arasındaki korelasyonlar da hesaplanmıştır. Elde edilen korelasyon değerleri Tablo 6’da gösterilmektedir.

Tablo 6

Ülkelerin Örneklemeleri ve Çalışma Örneklemelerinin Fen Başarısı ve Fene Yönelik Tutum Ölçekleri Arasındaki Ülke İçi Korelasyon Değerleri

Ülkeler	Fenden Hoşlanma ve Fen Başarısı		Başarı Motivasyonu ve Fen Başarısı		Fen Öz Yeterliği ve Fen Başarısı	
	Örneklem	Çalışma Örneklemi	Örneklem	Çalışma Örneklemi	Örneklem	Çalışma Örneklemi
Avusturya	,332*	,338*	,081*	,151*	,310*	,304*
Belçika	,327*	,337*	,036*	,047*	,319*	,301*
Brezilya	,201*	,195*	,189*	,079*	,184*	,200*
Hırvatistan	,268*	,264*	,154*	,112*	,265*	,258*
Danimarka	,337*	,291*	,240*	,230*	,347*	,325*
Dominik Cum.	,074*	,110*	,077*	,215*	,014*	,069*
Finlandiya	,333*	,293*	,249*	,217*	,337*	,338*
Fransa	,354*	,319*	,141*	,096*	,311*	,317*
Almanya	,351*	,359*	,112*	,167*	,338*	,360*
İsrail	,244*	,268*	,103*	,128*	,143*	,166*
İtalya	,251*	,222*	,057*	,093*	,164*	,135*
Japonya	,340*	,287*	,178*	,190*	,234*	,256*
Letonya	,185*	,225*	,199*	,214*	,226*	,190*
Lüksemburg	,340*	,329*	,125*	,170*	,285*	,265*
Meksika	,153*	,128*	,170*	,145*	,090*	,132*
Hollanda	,326*	,298*	,142*	,123*	,256*	,206*
Peru	,090*	,109*	,216*	,140*	,084*	,049*
Katar	,272*	,282*	,128*	,146*	,221*	,195*
Rusya Fed.	,190*	,164*	,159*	,177*	,176*	,206*
Birleşik Arap E.	,232*	,242*	,120*	,102*	,217*	,154*
Türkiye	,180*	,177*	,112*	,152*	,145*	,130*

* $p \leq ,01$

Tablo 6 incelendiğinde fene başarısı ile fene yönelik tutum ölçekleri arasında ülke örneklemeleri ve çalışma örneklemelerine ilişkin elde edilen korelasyon değerlerinin birbirine yakın olduğu görülmektedir. Ayrıca, hem ülke örneklemeleri hem de çalışma örneklemeleri için fene yönelik tutum ölçekleri ile fen başarısı arasındaki korelasyon değerleri ,01 düzeyinde anlamlı bulunmuştur ($p \leq ,01$). Elde edilen bu korelasyon değerlerine göre çalışma örneklemelerinin ülke örneklemelerini temsil ettiği söylenebilir.

Verilerin Elde Edilmesi

Araştırmada veriler PISA 2015'te uygulanan öğrenci anketlerinden ve fen başarı testlerinden elde edilmiştir. Veri dosyaları PISA web sayfasından (<https://www.oecd.org/pisa/data/2015database/>) elde edilmiştir.

Veri Toplama Araçları

Veriler PISA 2015 fene yönelik öğrenci tutum anketleri ve fen başarı testlerinden elde edilmiştir. Aşağıda PISA 2015’de yer alan fen yönelik tutum anketlerine ve fen başarı testine ilişkin bilgiler verilmiştir.

Fene Yönelik Tutum Anketleri

PISA 2015 uygulamasında odak alan fen okuryazarlığıdır. Bu araştırmada da fene yönelik tutum ölçekleri dikkate alınmıştır. PISA 2015 uygulamasında fen okuryazarlığına yönelik tutumu ölçmeyi amaçlayan 4’lü likert tipinde oluşturulmuş 6 alt ölçek bulunmaktadır. Araştırmada hangi alt ölçeklerin kullanılacağına karar vermek amacıyla tutum-başarı paradoksuna bakılmıştır. Buna göre, tutum-başarı paradoksunun görüldüğü alt ölçekler ST094 kodlu ve 5 maddeden oluşan “Fenden Hoşlanma” (Enjoyment of Science), ST119 kodlu ve 5 maddeden oluşan “Başarı Motivasyonu” (Achievement Motivation) ve ST129 kodlu ve 8 maddeden oluşan “Fen Öz Yeterliği” (Science Self-Efficacy) ölçekleridir. Bu alt ölçeklerde yer alan maddelerin her biri dört tepki kategorisine sahiptir. “Fenden Hoşlanma” ve “Başarı Motivasyonu” alt ölçeklerindeki maddeler 1=Kesinlikle Katılmıyorum, 2= Katılmıyorum, 3= Katılıyorum ve 4=Kesinlikle Katılıyorum şeklinde dörtlü likert tipinden oluşmaktadır. Fen Öz Yeterliği alt ölçeği ise 1= Bunu kolaylıkla yapabilirim, 2= “Bunu uğraşırsam yapabilirim”, 3= “Bunu yapabilmek için çok çaba göstermem gerekir” ve 4 = “Bunu yapamam” şeklindeki kategorilerden oluşmaktadır. Fen Öz Yeterliği alt ölçeğindeki maddelerde ters puanlama yapılmıştır. Böylece Fen Öz Yeterliği ölçeğinden elde edilen yüksek puan, öğrencilerin fendeki öz yeterliklerin yüksek olduğunu, elde edilen düşük puan ise öğrencilerin fendeki öz yeterlik düzeylerinin düşük olduğunu göstermektedir. Araştırmada kullanılan ölçek maddeleri Ek 1’de yer almaktadır. Ayrıca, her üç ölçek içinde Cronbach alfa (α) güvenilirlik katsayıları hesaplanmıştır. Her bir ülke için elde edilen Cronbach alfa (α) güvenilirlik katsayıları Ek 4’de yer almaktadır. Fenden Hoşlanma ölçeğinde her bir ülke için elde edilen güvenilirlik katsayı değerleri ,89 ile ,96 arasında değişmektedir. Başarı Motivasyonu ölçeğinde her bir ülke için elde edilen güvenilirlik katsayı değerleri ,68 ile ,87 arasında değişmektedir. Fen Öz Yeterliği ölçeğinde her bir ülke için elde edilen güvenilirlik katsayı değerleri ise ,82 ile ,91 arasında değişmektedir. Her üç ölçek içinde elde edilen Cronbach alfa güvenilirlik katsayılarının ,60’dan büyük olması ölçeklerin güvenilirliklerinin yüksek olduğunu göstermektedir (Büyüköztürk, 2005).

Fen Başarı Testi

PISA 2015 uygulaması, önceki PISA uygulamalarından farklı olarak bilgisayar tabanlı değerlendirme olarak yapılmıştır. Ancak bilgisayar yöntemini kullanmak istemeyen ülkeler, kâğıt-kalem tabanlı değerlendirmeleri seçerek uygulamayı gerçekleştirmişlerdir. Bu ülkeler için önceki döngülerde kullanılan sorular kullanılmıştır. PISA 2015 için geliştirilen başarı testi maddeleri sadece bilgisayar tabanlı değerlendirmeye uygun olarak yazılmıştır. PISA 2015'e katılan 72 ülkeden 57'si uygulamayı bilgisayar tabanlı değerlendirme, 15'i ise kâğıt-kalem tabanlı değerlendirme olarak gerçekleştirmiştir. Bilgisayar tabanlı değerlendirme için 66 farklı kitapçık bulunurken, kâğıt-kalem tabanlı değerlendirmeler için ise 30 farklı kitapçık bulunmaktadır (MEB, 2016, s.4). PISA uygulamasında öğrenciler bütün soruları cevaplamamakta ve yalnızca soruların belli kesimini cevaplamaktadırlar. Bu nedenle öğrenciler için tek bir başarı puanı yerine MTK kullanılarak 10 farklı olası değer (plausible values) (PV1SCIE-PV10SCIE) kestirilmiştir.

Verilerin Analizi

Analize başlamadan önce, ülkelerdeki öğrenci sayısının fazla olması ve kayıp değer atama yöntemlerinin tepki stillerini etkileyeceği göz önüne alınarak kayıp veriler liste bazında veri silme yöntemi kullanılarak silinmiştir. Kayıp verilerin silinmesinden sonra analizlerin yapıldığı bilgisayarın veri işleme kapasitesi ve işlemci hızı göz önüne alınarak her bir ülke örneklemelerinden %10'luk rastgele çalışma örneklemeleri seçilerek analizler bu çalışma örneklemeleri üzerinden gerçekleştirilmiştir.

Hem öğrencilerin hem de ülkelerin tutum ölçeklerine ilişkin ortalamalarını hem de fen başarılarına ilişkin ortalamalarını hesaplarken, örneklem ağırlıklandırmasından yararlanılmıştır. Bu doğrultuda, hem ortalamaları hesaplarken hem de korelasyon analizlerinde örneklem ağırlıklarını ve olası değerlerini dikkate alan International Association for Evaluation (IEA)'nın IDB Analyzer 4.0 programı kullanılmıştır. IDB Analyzer programı SPSS veri dosyalarını okuyabilen ve program için kod üretebilen yardımcı bir programdır. Araştırmada ölçek ortalamalarını ve olası fen başarı puan ortalamalarını hesaplamak için öğrenci ağırlıklandırması kullanılmıştır. PISA 2015 veri setinde öğrenci düzeyindeki örneklem ağırlığı olarak FINAL TRIMMED NONRESPONSE ADJUSTED STUDENT WEIGHT (W_FSTUWT) kullanılmıştır.

Araştırma sorularına göre gerçekleştirilen analizler hakkında bilgiler aşağıda sunulmuştur:

Birinci araştırma sorusu kapsamında tutum-başarı paradoksunun görüldüğü ülkelerdeki gözlenen puanlar TSYDK yöntemi ile analiz edilerek tepki stili yanlılığından arındırılmış düzeltilmiş puanlar elde edilmiştir. Düzeltilmiş bu puanlar aracılığıyla da ülkelerin nasıl kümelendiği yine TSYDK yöntemi ile incelenmiştir. Ayrıca TSYDK yönteminden yararlanılarak ülkelerin tepki stiline göre nasıl kümelendiği de gösterilmiştir.

İkinci araştırma sorusu kapsamında PISA 2015 uygulamasına katılan ve tutum-başarı paradoksunun görüldüğü ülkelerin gözlenen puanlarına K-ortalamlar kümeleme yöntemi uygulanarak nasıl kümelendikleri belirlenmiştir.

Üçüncü araştırma sorusu kapsamında verilerin TSYDK yöntemi aracılığıyla elde edilen kümeleme sonuçları ile düzeltilmemiş veriler üzerinden gerçekleştirilen K-ortalamlar kümeleme yöntemi sonuçları karşılaştırılmıştır.

Verileri analiz etmede temel olarak TSYDK yöntemi kullanılmıştır. TSYDK yöntemi de ÇBMTK yöntemi gibi ölçülmek istenen gerçek özellik ile tepki stiline birbirinden ayrılmasını sağlamaktadır. Böylece tepki stili yanlılığı ortadan kaldırılarak düzeltilmiş gerçek özellik puanları elde edilebilmektedir. TSYDK yöntemi hakkında detaylı bilgi araştırmanın giriş bölümünde verilmiştir. Bu bölümde TSYDK yönteminin hangi kriter ve kurallara göre uygulandığı açıklanmıştır. Ayrıca araştırmanın amaçları doğrultusunda TSYDK yönteminin elverişliliğini göstermek amacıyla K-ortalamlar kümeleme analizi ile de karşılaştırması yapılmıştır.

K-ortalamlar kümeleme analizi, Hartigan ve Wong (1979) tarafından geliştirilmiş olan bir bölümlendirme tekniğidir. K-ortalamlar kümeleme algoritması, çok büyük veri setlerinden az sayıda küme elde etmek amacıyla en çok kullanılan kümeleme yöntemlerinden biridir. K-ortalamlar yöntemi, veri setini meydana getiren nesnelerin, objelerin veya bireylerin bazı nitelik veya özelliklerine göre “k” sayıda gruba ayrılmasını sağlamaktadır. Buradaki “k” pozitif bir tam sayı olmakla beraber veri setindeki toplam kişi sayısından daha küçüktür. K-ortalamlar tekniği; kolay uygulanabilmesi, çıktılarının kolay yorumlanması ve hızlı bir şekilde kümeleme analizinin gerçekleşmesi gibi özellikleri nedeni ile en popüler kümeleme yöntemleri arasında gösterilmektedir. K-ortalamlar kümeleme yöntemi ayrıca normal dağılım varsayımının karşılanmadığı veri setlerinde bile diğer kümeleme yöntemlerine kıyasla oldukça dirençli bir kümeleme yöntemidir (Genolini & Falissard, 2010; Kaufman & Rousseeuw, 1990; Usami, 2014).

K-ortalamlar kümeleme analizinde önemli olan analiz öncesinde ideal küme sayısının belirlenmesidir. İdeal küme sayısı “k” ile ifade edilmektedir. Araştırmacı tarafından analiz öncesinde belirlenen “k” değeri analizin performansını etkileyebilmektedir. Veri setine ilişkin belirli özelliklerin yansıtılması hususunda olası “k” değerlerini kullanmak çok önemlidir (Eser, 2019). K-ortalamlar kümeleme analizinde uygun küme sayısına karar vermek için çok sayıda indeks bulunmaktadır. Bu çalışmada TSYDK ve K-ortalamlar yöntemlerinin ikisi içinde uygun küme sayısının belirlenmesinde KL indeksi kullanılmıştır. K-ortalamlar kümeleme analizinde bütün ülkelerin çalışma örneklemi birleştirilerek tek bir veri seti oluşturulmuştur. TSYDK ve K-ortalamlar kümeleme analizinde belirlenen küme sayısının tutarlılığını sağlamak için 200 bootstrap örnekleme oluşturulmuştur. Buna göre, bütün ülkelerin yer aldığı veri setinden 200 bootstrap örnekleme elde edilmiş ve her bootstrap örneği için minimum küme sayısı K=2 ile maksimum küme sayısı K=15 olacak şekilde analizler gerçekleştirilmiştir. Yapılan analizlerde KL indeksini maksimize eden K değeri uygun küme sayısı olarak seçilmiştir. KL indeksi Krzanowski ve Lai (1988) tarafından tanımlanmıştır. Krzanowski ve Lai (1988), küme içi kareler toplamı (WSS) değerinin azalışını kullanarak aşağıdaki hesaplamayı yapmışlardır.

$$DIFF(k) = (k - 1)^{2/m} WSS_{k-1} - k^{2/m} WSS_k \quad (\text{Denklem 9})$$

Denklem 9’da m değeri değişkenlerin sayısıdır ve bir kriter olarak kullanılmaktadır. Denklem 9’daki değeri maksimize eden k uygun küme sayısıdır. Denklem 9’daki istatistiğe bağlı olarak KL indeksi aşağıdaki gibi hesaplanmaktadır.

$$KL(k) = \left| \frac{DIFF(k)}{DIFF(k+1)} \right| \quad (\text{Denklem 10})$$

KL indeksi Denklem 10 aracılığıyla elde edilmektedir. Burada k=1 için KL(k) tanımlanmamaktadır. KL indeksinde uygun küme sayısına karar vermedeki temel düşünce WSS(k) değerinin uygun küme sayısına ulaşana kadar hızlı bir şekilde azaldığı ve uygun küme sayısına ulaşıldığında yavaş bir şekilde azalma gösterdiği yerdur. Bu durumda WSS(k) değerinin en yavaş azalma gösterdiği değer uygun küme sayısı olarak değerlendirilmektedir. Ayrıca 2’den büyük değerler için en yüksek Krzanowski-Lai indeks değerine sahip k sayısı uygun küme sayısı olarak kabul edilmektedir.

KL indeksini belirlemek için R programındaki NbClust paketi (Charrad, Ghazzali, Boiteau, & Niknafs, 2014) kullanılmıştır. Her bir bootstrap örnekleme için elde edilen uygun küme sayılarından en sık seçileni yani frekansı en yüksek olan değer temel analizlerde uygun küme

sayısı olarak yer almıştır. TSYDK yönteminde uygun küme sayısı ile birlikte λ değerinin belirlenmesi gerekmektedir. Takagishi vd. (2019)’un yaptıkları simülasyon çalışmasında $\lambda=0.8$ değerini en uygun değer olarak belirledikleri için bu araştırmada da bu değer kullanılmıştır. Ayrıca TSYDK yöntemi için 20 farklı başlangıç değeri seçilmiştir. TSYDK yönteminde K değeri içeriğe dayalı küme sayısını gösterirken, H değeri ise tepki stiline dayalı küme sayısını göstermektedir. K-ortalamar kümeleme analizi içinde 500 farklı başlangıç değeri kullanılmıştır. Araştırma kapsamında TSYDK yöntemi için R programındaki “ccrs” (Takagishi, 2019) paketi kullanılırken, K-ortalamar yöntemi için ise R programında kod yazılmıştır. Ayrıca elde edilen sonuçların geçerliliklerinin değerlendirilmesi aşamasında yararlanılan grafiklerinin oluşturulmasında R programındaki “reshape2” paketi ile “ggplot2” paketleri kullanılmıştır. Analizlerin gerçekleştirildiği R kodları Ek 7 ve Ek 8’ de yer almaktadır.

Dördüncü araştırma sorusunun a seçeneği kapsamında tepki stili yanlılığı bulunan veriler ile TSYDK yöntemi ile düzeltilmiş verilerdeki ortalama puanlar arasındaki fark puanlarının manidarlığı incelenmiştir. Fark puanların manidarlığını incelemeye her bir ülke için ayrı ayrı ilişkili örneklem t-testi yapılmıştır. Analizler için ilk olarak bireylerin araştırma kapsamında kullanılan fene yönelik tutumlarını ölçen üç ölçekten aldığı ham puanlar $0 \leq x \leq 1$ aralığında olacak şekilde yeni değerlere dönüştürülmüştür. Bunun sebebi TSYDK yöntemi ile tepki stili yanlılığı bulunan verilerin 0-1 aralığında düzeltilmiş puanlara dönüştürülmesidir. Düzeltilmemiş ve düzeltilmiş verilerden elde edilen puanların karşılaştırılması amaçlandığından ham puanlar için bu ölçek dönüştürmesinin yapılması gerekli görülmüştür. Madde puanlarını 0-1 aralığına dönüştürmek için min-max normalleştirme yapılmıştır. Min-Max normalleştirme ile orijinal veri üzerinde doğrusal bir dönüşüm yapılır. Bu yöntem aracılığıyla veriler 0-1 aralığına dönüştürülür. Min.-Max normalizasyon işlemi alan değerinin minimum değerden ne kadar büyük olduğuna bakar ve bu farkları sıralar (Larose & Larose, 2014). Bu işlem Denklem 11’deki gibi yapılmaktadır.

$$X^* = \frac{(X - \min(X))}{\text{aralık}(x)} = \frac{(X - \min(X))}{(\text{maks}(X) - \min(X))} \quad (\text{Denklem 11})$$

Denklem 11’de;

X^* =Normalize edilmiş veriyi,

x_i =Girdi değerini,

$x_{\min}$ =Girdi seti içerisinde yer alan en küçük sayıyı,

x_{max} =Girdi seti içerisinde yer alan en büyük sayıyı göstermektedirler.

Ölçeklerdeki ham puanlar Denklem 11'deki hesaplamalar yardımı ile 0-1 aralığına dönüştürülmüştür. Daha sonra dönüştürülen puanlar kullanılarak her bir değişken için ağırlıklandırılmış örneklem ortalama puanlar elde edilmiştir.

İlişkili örneklem için t-testi yapmadan önce varsayımlar her bir ülke için ve ülkelerin tamamı için test edilmiştir. Buna göre fenden hoşlanma, başarı motivasyonu ve fen öz yeterliği ölçekleri için düzeltilmemiş veriler ile TSYDK yöntemi ile düzeltilmiş verilerin ortalama puanları arasındaki fark puanlarının aralık ölçek düzeyinde olup olmadığı ve normal dağılım sergileyip sergilemediği test edilmiştir. Fark puanlarının aralık düzeyinde olduğu görülmüştür. Fark puanlarında tek değişkenli normalliğin sağlanıp sağlanmadığını göstermek amacıyla çarpıklık ve basıklık değerlerine bakılmış ve histogram grafikleri de incelenmiştir. Her bir ülke için elde edilen fark puanlarının normalliğini gösteren çarpıklık ve basıklık değerleri Ek 5'de gösterilmektedir. Ayrıca her bir ülke için elde edilen histogram grafikleri de Ek 6'da yer almaktadır. Fark puanlarını hesaplamak için düzeltilmiş değerlerden düzeltilmemiş değerler çıkartılmıştır. Çalışma örneklemelerinin ortalamaları arasındaki fark puanlarının manidarlığını incelemek için etki büyüklükleri de hesaplanmıştır. İlişkili örneklem için t testi analizi sonucunda elde edilen manidar farklara ilişkin Cohen d etki büyüklüğü değerinin nasıl hesaplandığı Denklem 12'de gösterilmektedir. Etki büyüklüğü indeksi (d) için ,20 değerine ulaşıldığında küçük ,50 değerine ulaşıldığında orta ve ,80 değerine ulaşıldığında ise geniş etki büyüklüğü şeklinde yorumlanmıştır.

$$d = \frac{t}{\sqrt{N}} \quad (\text{Denklem 12})$$

Denklem 12'de t değeri ilişkili örneklem için t testi sonucunda elde edilen değeri gösterirken, N değeri çalışma örneklemindeki kişi sayısını göstermektedir.

Dördüncü araştırma sorusunun b seçeneği kapsamında tepki stili yanlılığı bulunan verilerde ilk olarak her bir birey için fene yönelik tutum ölçeklerinden aldıkları puanların ağırlıklandırılmış örneklem ortalaması ve 10 olası fen başarı puanlarını ağırlıklandırılmış ortalaması elde edilmiştir. Daha sonra öğrencilerin ortalama puanlarından yararlanarak her bir ülke için fenden hoşlanma ölçeği puan ortalaması, başarı motivasyonu ölçeği puan ortalaması ve fen öz yeterlik ölçeği puan ortalaması ile ülkelerin fen başarısını temsil eden başarı puanı olmak üzere dört farklı puan elde edilmiştir. Aynı işlemler TSYDK yöntemi ile düzeltilmiş veriler üzerinden de gerçekleştirilmiştir. Daha sonra her bir ülke içinde öğrencilerin düzeltilmemiş ve düzeltilmiş ortalama puanlarından yararlanarak ülke içinde

fene yönelik tutum ile fen başarı puanları arasındaki Pearson Korelasyon Katsayısı (r_p) hesaplanmıştır. Ülkeler arasında fene yönelik tutum ile fen başarısı arasındaki ilişkiyi incelemek için her bir ülkenin düzeltilmemiş ve düzeltilmiş ortalama ölçek puanları ile ortalama fen başarı puanları elde edilmiştir. Daha sonra 21 ülkeden edilen bu puanlar için normallik testi yapılmıştır. Örneklem büyüklüğünün 30'dan küçük olması nedeni ile Shapiro Wilk normallik testinin (Shapiro & Wilk, 1965) kullanılması uygun görülmüştür. Ülkelerin ortalama ölçek puanları ile fen başarı puanlarının normal dağılım sergilemediği görülmüştür ($p<,05$). Buna göre ülkeler arası fene yönelik tutum ile fen başarısı arasındaki ilişkiyi incelemek için Sıra Farkları Korelasyon Katsayısı (r_s) hesaplanmıştır.



BÖLÜM IV

BULGULAR VE YORUMLAR

Bu bölümde PISA 2015 verileri üzerinden gerçekleştirilen analizlerden elde edilen bulgular her bir araştırma sorusuna dayalı olarak sırasıyla verilmiştir.

Birinci Araştırma Sorusuna İlişkin Elde Edilen Bulgular ve Yorumlar

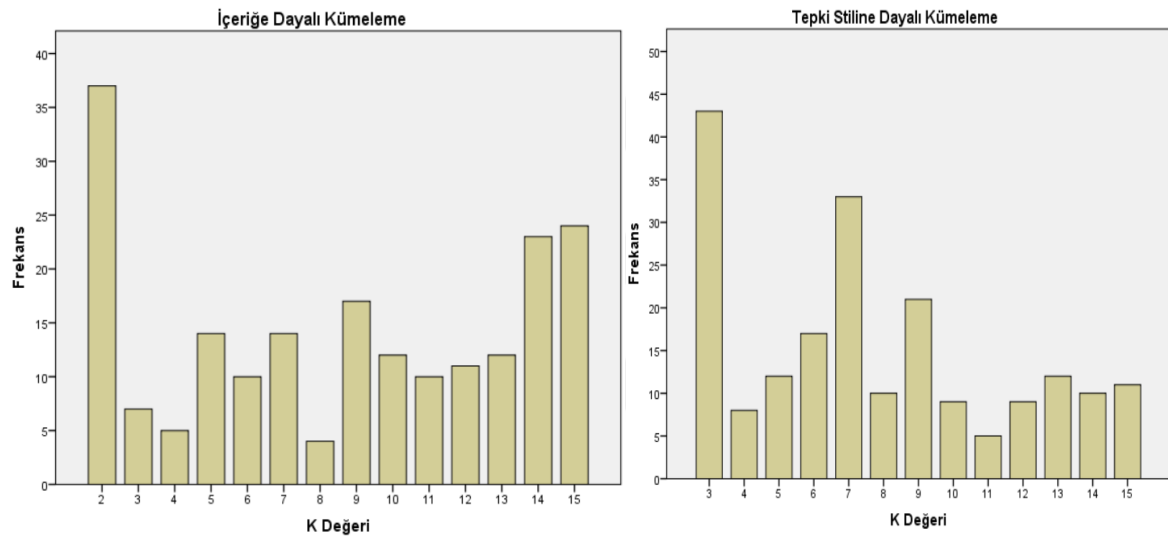
Araştırma Sorusu 1: PISA 2015 uygulamasına katılan ve tutum-başarı paradoksunun görüldüğü ülkeler TSYDK yöntemi ile nasıl kümelenmektedir?

Birinci araştırma sorusu kapsamında çalışmaya dâhil edilen 21 ülkenin fene yönelik tutumları ile ilgili olan Fenden Hoşlanma, Başarı Motivasyonu ve Fen Öz Yeterliği ölçeklerindeki tepki stillerinin TSYDK yöntemi ile düzeltiliş ve düzeltilen veriler üzerinden kümeleme analizi yapılmıştır. TSYDK yöntemi ile ülkelerin kümelenmesine geçmeden önce her bir ölçek için tüm ülkelerin verileri tek bir veri setinde birleştirilerek bütün öğrencilerin tepki stilleri düzeltilerek ölçeklerdeki her bir madde için düzeltilmiş değerler elde edilmiştir. Tepki stillerinin düzeltilmesi her bir ölçek için ayrı ayrı yapılmıştır. Buna göre, aşağıda ilk olarak Fenden Hoşlanma ölçeği için elde edilen bulgulara yer verilmiştir.

Fenden Hoşlanma Ölçeği için Elde Edilen Bulgular

Fenden Hoşlanma ölçeği için TSYDK yöntemi ile içeriğe dayalı kümeleme ve tepki stiline dayalı kümelemede ilk olarak uygun küme sayısı belirlenmiştir. Uygun küme sayısına karar vermek için araştırma verisi olarak kullanılan 10898 kişilik veri setinden 200 tane bootstrap

örneklem elde edilmiş ve her bir bootstrap örnekleminde minimum küme sayısı $K=2$ ile maksimum küme sayısı $K=15$ olacak şekilde her bir küme sayısı için KL indeksi hesaplanmıştır. Her bootstrap örneklemini için KL indeksini maksimize eden K değeri uygun küme sayısı olarak değerlendirilmektedir. Toplam 200 farklı örneklem için 200 tane küme sayısı belirlenmiştir. Temel analizler için içerik tabanlı ve tepki stiline dayalı kümeleme de uygun küme sayısı olarak 200 tane küme sayısı içinde en çok tekrar eden değerler belirlenmiştir. TSYDK yöntemi için ayrıca $\lambda=,8$ değeri analizlerde kullanılmıştır. İçeriğe dayalı kümeleme ve tepki stiline dayalı kümelemede uygun küme sayısına karar vermek için bootstrap örneklemelerinden elde edilen küme sayıları Şekil 8’de gösterilmektedir.



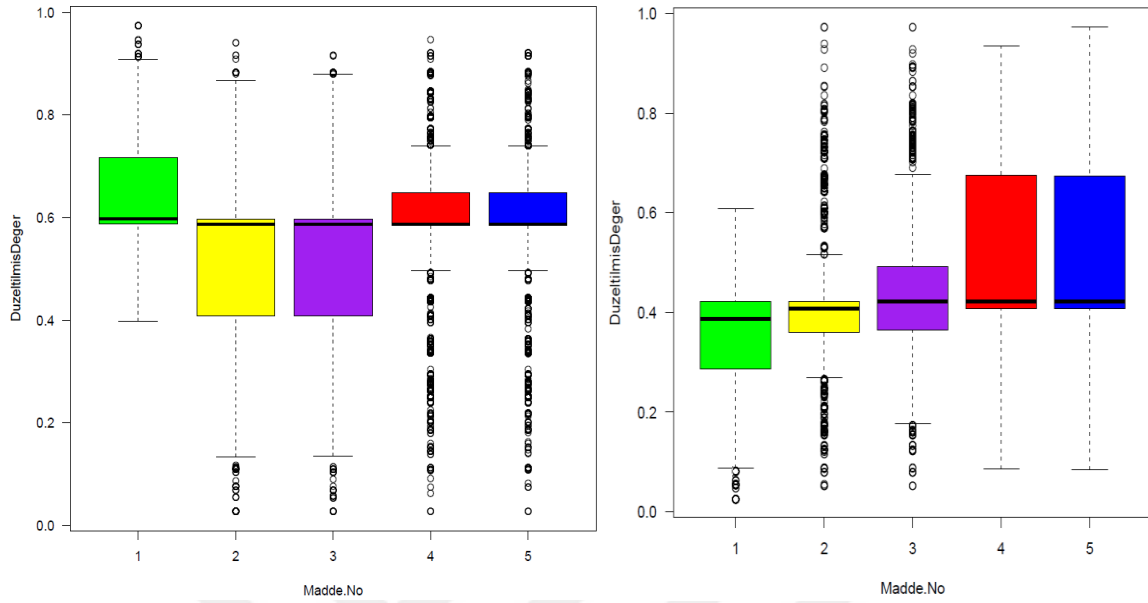
Şekil 8. Fenden hoşlanma ölçeği için TSYDK yöntemi ile elde edilen içeriğe ve tepki stiline dayalı bootstrap örneklemelerinin küme sayılarını gösteren histogram grafikleri.

Şekil 8’deki değerlere bakıldığında 200 bootstrap örnekleminde içeriğe dayalı kümelemede en fazla iki kümenin oluştuğu, tepki stiline dayalı kümelemede ise en fazla üç kümenin oluştuğu gözlenmektedir. Buna göre, TSYDK yöntemi ile yapılan temel analizlerde içeriğe dayalı küme sayısı $K=2$ ve tepki stiline dayalı küme sayısı ise $H=3$ olarak belirlenmiştir. Parametreler ayarlandıktan sonra TSYDK yöntemi 500 farklı başlangıç değeri için uygulanmıştır.

TSYDK yöntemi ile verilerin tepki stili yanlılığının düzeltilmesi ve kümelenmesinde ilk olarak düzeltilen değerler üzerinden kümelerin nasıl farklılaştığına bakılmıştır. Bu doğrultuda TSYDK yöntemi ile içeriğe dayalı iki küme oluştuğu için her iki küme içinde maddelerin düzeltilen değerlerinin kutu grafikleri Şekil 9’da gösterilmektedir.

1. Küme (%64,1)

2. Küme (%35,9)

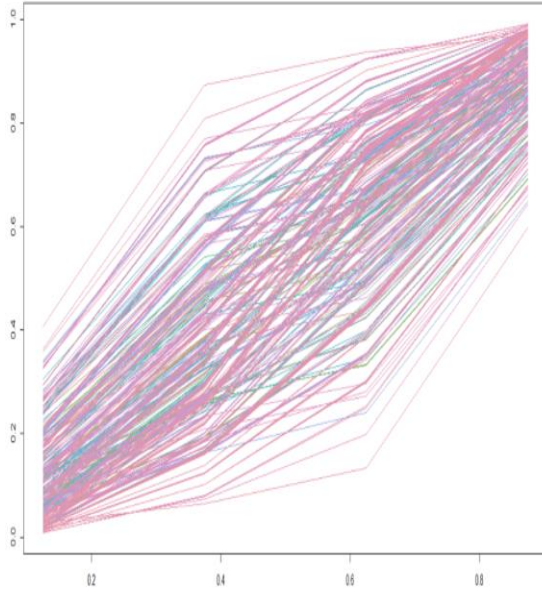


Şekil 9. Fenden hoşlanma ölçeği için TSYDK yöntemi ile elde edilen içerik tabanlı kümelerde beş maddenin kutu grafikleri.

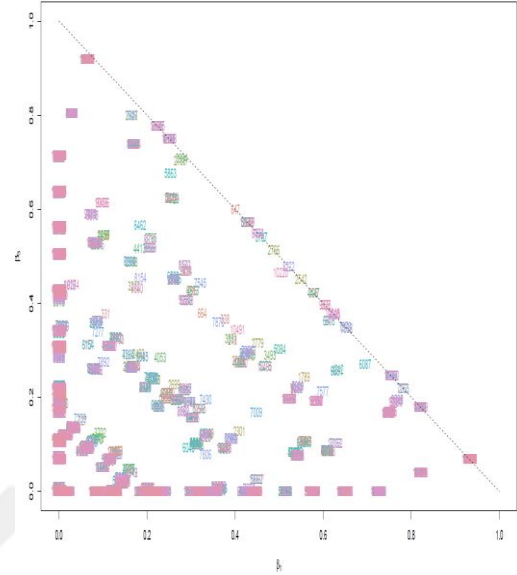
Şekil 9’da dikey eksen tahmini tepki fonksiyonları kullanılarak elde edilen düzeltilmiş verilerin ölçeğini gösterirken, yatay eksen madde numaralarını göstermektedir. Düzeltilmiş değerler kullanılarak elde edilen içerik tabanlı kümelemede öğrencilerin %64,1’i birinci kümede yer alırken, %35,9’u ikinci kümede bulunmaktadır. Şekil 9 incelendiğinde, kümelerin maddelerin değerlendirilmesinde nasıl farklılaştığı da görülmektedir. Buna göre, birinci ve ikinci kümedeki maddelerin kutu grafikleri incelendiğinde maddelerin değerlerinin ve grafiklerinin her iki küme içinde benzer olduğu görülmektedir. Maddelerin kutu grafiklerindeki farklılıkların sadece bir madde (Madde 1) için geçerli olduğu söylenebilir.

TSYDK yöntemi ile katılımcıya özgü tepki fonksiyonlarının kümelenmesi ile tepki stili tabanlı üç küme elde edilmişti. Araştırmada yer alan her bir öğrenciye karşılık gelen tepki fonksiyonları Şekil 10’da (solda) gösterilmektedir. Ayrıca, TSYDK yöntemi ile tahmin edilen tepki fonksiyonlarındaki katsayılar, tepki stillerinin özelliklerini görsel olarak göstermek için de kullanılabilir. Buna göre her bir öğrencinin β_1 ve β_3 eksenleri ile oluşturulan düzlemdeki yerleri Şekil 10’da (sağda) gösterilmektedir. “Kabullenme” tepki stili yüksek β_3 ve düşük β_1 değerine karşılık gelir. Benzer şekilde, ‘Orta nokta’ tepki stili ise düşük β_1 ve β_3 değerlerine karşılık gelmektedir.

Tahmini Tepki Fonksiyonları



Katsayı Grafiği



Şekil 10. (Sol) Fenden hoşlanma ölçeği için TSYDK yöntemi ile elde edilen tepki stili tabanlı kümelerin tahmini tepki fonksiyonları. (Sağ) Tepki stili tabanlı kümelerin β_h değeri için düşük boyutlu çizimi.

Şekil 10 incelendiğinde; sol taraftaki grafikte y eksenini düzeltilmiş değerler ölçeğini gösterirken, x eksenini gözlenen ölçek değerlerini göstermektedir. Ayrıca sağ taraftaki grafikte yer alan sayılar öğrenci numaralarını temsil etmektedir. Tepki stiline dayalı kümelerin ülkelerle ilgili olup olmadığını incelemek için ülkelerin tepki stili tabanlı kümeler üzerindeki dağılımları Tablo 7’de gösterilmektedir.

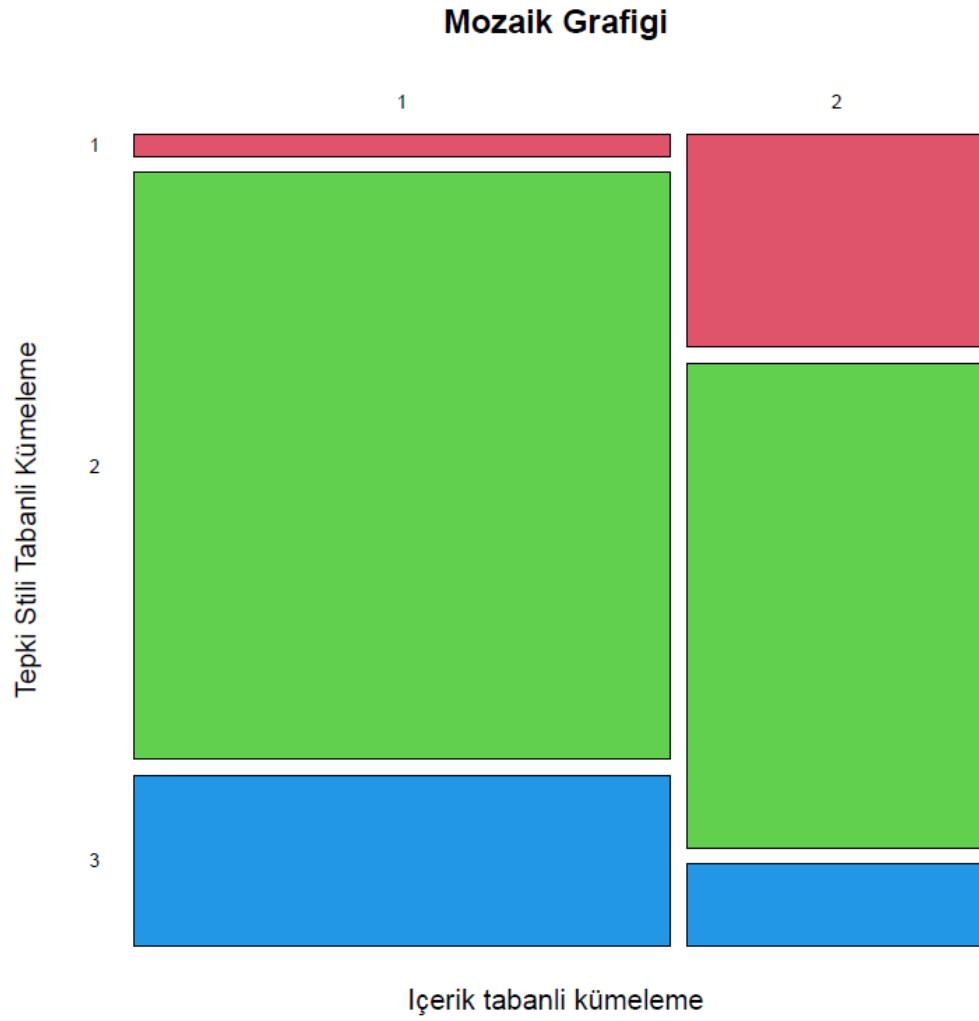
Tablo 7

Fenden Hoşlanma Ölçeği için Ülkelerin TSYDK Yöntemi ile Elde Edilen Tepki Stili Tabanlı Kümeler (H=1,2,3) Üzerindeki Dağılımı

Ülkeler	H=1		H=2		H=3	
	Frekans	Yüzde	Frekans	Yüzde	Frekans	Yüzde
Avusturya	106	22,6	291	61,9	73	15,5
Belçika	95	14,6	450	69,1	106	16,3
Brezilya	64	7,2	673	76,0	148	16,7
Hırvatistan	47	11,1	322	75,8	56	13,2
Danimarka	52	10,3	363	72,2	88	17,5
Dominik Cumhuriyeti	20	9,1	128	58,4	71	32,4
Finlandiya	41	9,2	362	81,3	42	9,4
Fransa	68	16,6	271	66,3	70	17,1
Almanya	49	20,2	155	63,8	39	16,0
İsrail	73	16,3	269	60,0	106	23,7
İtalya	98	11,0	694	78,2	96	10,8
Japonya	110	19,2	383	66,7	81	14,1
Letonya	28	7,4	296	78,5	53	14,1
Lüksemburg	54	15,5	228	65,3	67	19,2
Meksika	25	4,6	386	70,4	137	25,0
Hollanda	99	24,0	290	70,2	24	5,8
Peru	24	5,5	314	71,4	102	23,2
Katar	57	7,7	501	67,8	181	24,5
Rusya	32	7,6	354	83,9	36	8,5
Birleşik Arap Emir.	80	8,0	658	65,4	268	26,6
Türkiye	50	11,3	302	68,0	92	20,7
Toplam	1272	11,7	7690	70,6	1936	17,8

Tablo 7 incelendiğinde araştırmada yer alan öğrencilerin %11,7'si birinci kümede, %70,6'sı ikinci kümede ve %17,8' i ise üçüncü kümede yer almaktadır. Her bir kümedeki ülkelerin yüzde değerlerine bakıldığında, birinci kümede %24,0 değeri ile öğrencilerinin çoğunluğunun bulunduğu ülke Hollanda olurken, ikinci kümede %83,9 değeri ile öğrencilerinin çoğunluğunun bulunduğu ülke Rusya ve üçüncü kümede ise %32,4 değeri ile öğrencilerinin çoğunluğunun bulunduğu ülke Dominik Cumhuriyeti olmuştur.

TSYDK yönteminin etkililiğini araştırmak için içerik tabanlı kümeler ile tepki stili tabanlı kümelerin nasıl karşılık geldiği gösterilmiştir. Bunun için iki kümeleme çözümünün çapraz tablosunu görselleştiren mozaik çizim yapılmıştır. Buna göre Fenden Hoşlanma ölçeği için içerik tabanlı kümeleme ile tepki stili tabanlı kümeleme sonuçlarının mozaik grafiği Şekil 11'de gösterilmektedir.

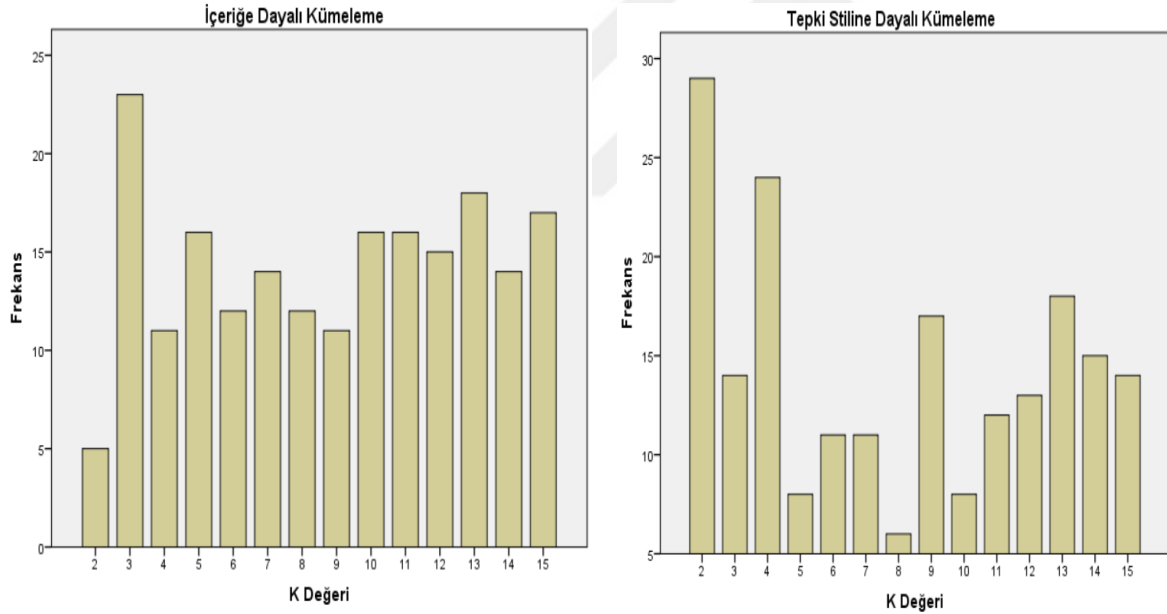


Şekil 11. Fenden hoşlanma ölçeği için TSYDK yöntemi ile elde edilen içerik tabanlı ve tepki stili tabanlı kümelerin mozaik grafiği.

Şekil 11 incelendiğinde, içerik tabanlı birinci ve ikinci kümeler ile tepki stili tabanlı üç kümenin çakışmadığı görülmektedir. Buna göre her iki içerik tabanlı kümede tüm tepki stili tabanlı kümelerden yanıt verenlerin bulunduğu anlaşılmaktadır. Diğer bir ifade ile içerik tabanlı kümeleme ile tepki stili tabanlı kümeleme sonuçları çakışmaz. Buna göre, TSYDK yönteminin öğrencileri ölçülen özelliğe göre sınıflandırmada başarılıdır.

Başarı Motivasyonu Ölçeği için Elde Edilen Bulgular

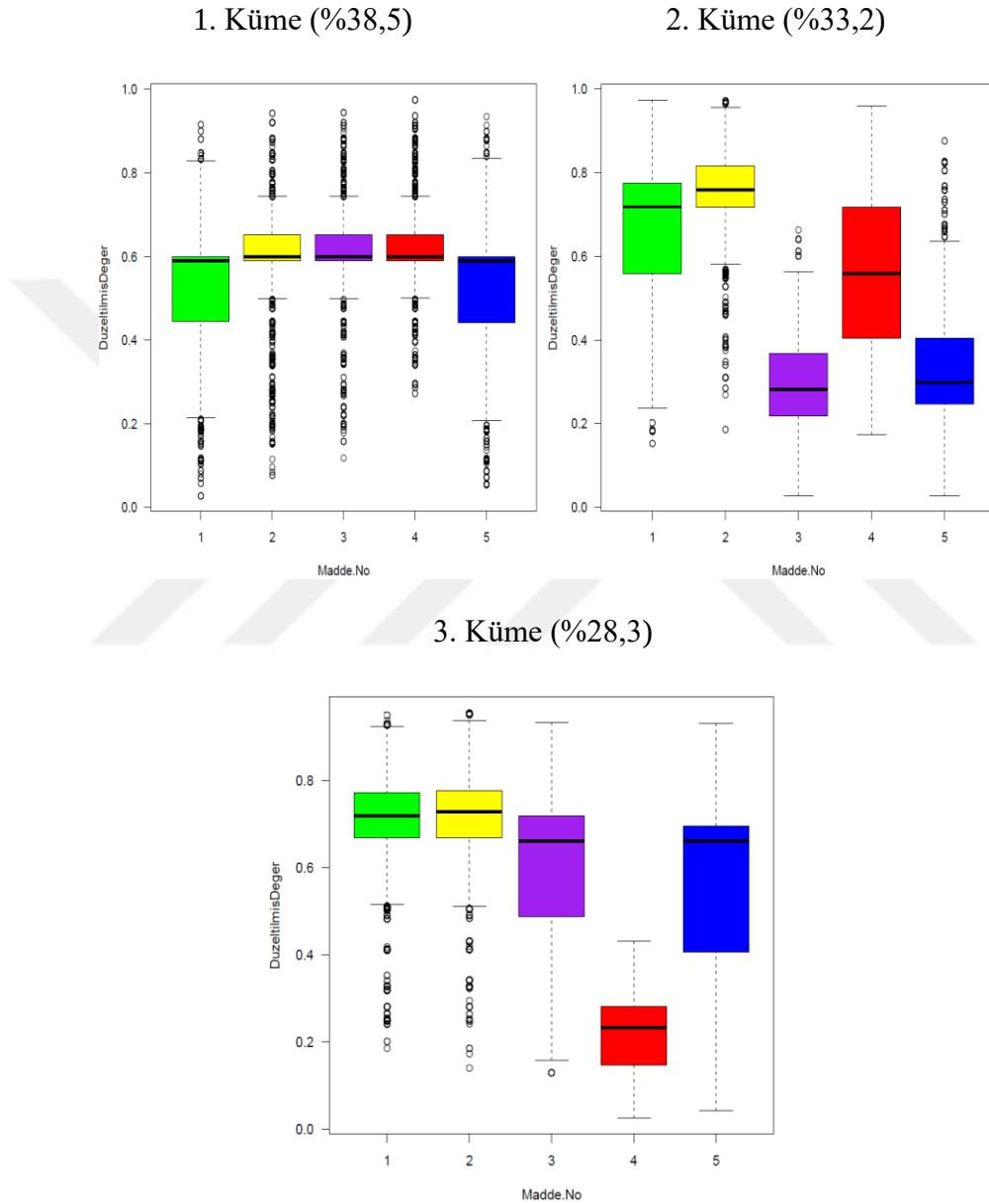
Başarı Motivasyonu ölçeği için TSYDK yöntemi ile içeriğe dayalı kümeleme ve tepki stiline dayalı kümelemede ilk olarak uygun küme sayısı belirlenmiştir. Uygun küme sayısına karar vermek için araştırma verisi olarak kullanılan 10898 kişilik veri setinden 200 bootstrap örneklem elde edilmiş ve her bir bootstrap örnekleminde minimum küme sayısı $K=2$ ile maksimum küme sayısı $K=15$ olacak şekilde her bir küme sayısı için KL indeksi hesaplanmıştır. Her bootstrap örneklemini için KL indeksini maksimize eden K değeri uygun küme sayısı olarak değerlendirilmektedir. Temel analizler için içerik tabanlı ve tepki stiline dayalı kümeleme de uygun küme sayısı olarak 200 küme sayısı içinde en çok tekrar eden değerler belirlenmiştir. TSYDK yöntemi için ayrıca $\lambda=,8$ değeri analizlerde kullanılmıştır. İçeriğe dayalı kümeleme ve tepki stiline dayalı kümelemede uygun küme sayısına karar vermek için bootstrap örneklemelerinden elde edilen küme sayıları Şekil 12’de gösterilmektedir.



Şekil 12. Başarı motivasyonu ölçeği için TSYDK yöntemi ile elde edilen içeriğe ve tepki stiline dayalı bootstrap örneklemelerinin küme sayılarını gösteren histogram grafikleri.

Şekil 12’deki değerlere bakıldığında 200 bootstrap örnekleminde içeriğe dayalı kümelemede en fazla üç kümenin oluştuğu, tepki stiline dayalı kümelemede ise en fazla iki kümenin oluştuğu görülmektedir. Buna göre, TSYDK yöntemi ile yapılan temel analizlerde içeriğe dayalı küme sayısı $K=3$ ve tepki stili küme sayısı ise $H=2$ olarak belirlenmiştir. Parametreler ayarlandıktan sonra TSYDK yöntemi 500 farklı başlangıç değeri için uygulanmıştır.

TSYDK yöntemi ile verilerin tepki stili yanlılığının düzeltilmesi ve kümelenmesinden elde edilen bulguların incelenmesinde ilk olarak düzeltilen değerler üzerinden kümelerin nasıl farklılaştığına bakılmıştır. Bu doğrultuda TSYDK yöntemi ile içeriğe dayalı üç küme olduğu için her üç küme içinde maddelerin düzeltilen değerlerinin kutu grafikleri Şekil 13’de gösterilmektedir.



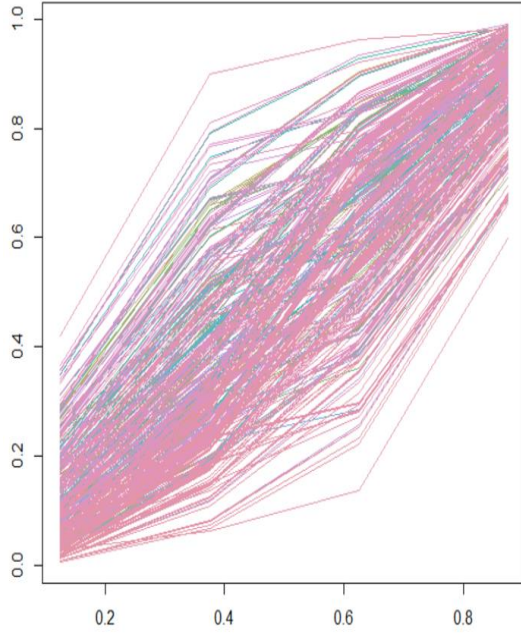
Şekil 13. Başarı motivasyonu ölçeği için TSYDK yöntemi ile elde edilen içerik tabanlı kümelerde beş maddenin kutu grafikleri.

Şekil 13’de dikey eksen tahmini tepki fonksiyonları kullanılarak elde edilen düzeltilmiş verilerin ölçeğini gösterirken, yatay eksen madde numaralarını göstermektedir. Düzeltilmiş değerler kullanılarak elde edilen içerik tabanlı kümelemede öğrencilerin %38,5’i birinci

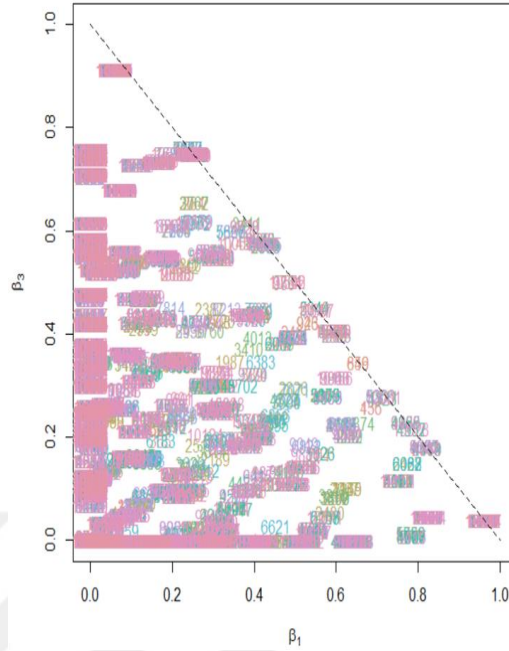
kümeye yer alırken, ikinci kümede öğrencilerin %33,2'si ve üçüncü kümede ise öğrencilerin %28,3'ü bulunmaktadır. Buna göre her üç kümedeki öğrencilerin yüzdelik değerlerinin birbirine yakın olduğu yorumu yapılabilir. Şekil 13 incelendiğinde, kümelerin maddelerin değerlendirilmesinde nasıl farklılaştığı görülmektedir. Buna göre, birinci ve üçüncü kümeler içinde maddelerin değerleri incelendiğinde farklılıkların bir madde (Madde 4) için geçerli olduğu görülmektedir. Ayrıca birinci ve üçüncü kümelerdeki diğer maddelerin kutu grafikleri incelendiğinde maddelerin değerlerinin ve grafiklerinin her iki küme içinde benzer olduğu görülmektedir.

TSYDK yöntemi ile katılımcıya özgü tepki fonksiyonlarının kümelenmesi ile iki küme çözümü elde edilmişti. Araştırmada yer alan her bir öğrenciye karşılık gelen tepki fonksiyonları Şekil 14'de (solda) gösterilmektedir. Ayrıca, TSYDK yöntemi ile tahmin edilen tepki fonksiyonlarındaki katsayılar, tepki stillerinin özelliklerini görsel olarak göstermek için de kullanılabilir. Buna göre her bir öğrencinin β_1 ve β_3 eksenleri ile oluşturulan düzlemdeki yerleri Şekil 14' de (sağda) gösterilmektedir.

Tahmini Tepki Fonksiyonları



Katsayı Grafiği



Şekil 14. (Sol) Başarı motivasyonu ölçeği için TSYDK yöntemi tarafından elde edilen tepki stili tabanlı kümelerin tahmini tepki fonksiyonları. (Sağ) Tepki stili tabanlı kümelerin β_h değeri için düşük boyutlu çizimi.

Şekil 14 incelendiğinde sol taraftaki grafikte y eksenini düzeltilmiş değerler ölçeğini gösterirken, x eksenini gözlenen ölçek değerlerini göstermektedir. Ayrıca sağ taraftaki grafikte yer alan sayılar öğrenci numaralarını temsil etmektedir. Tepki stiline dayalı kümelerin ülkelerle ilgili olup olmadığını incelemek için ülkelerin tepki stili tabanlı kümeler üzerindeki dağılımları Tablo 8’de gösterilmektedir.

Tablo 8

Başarı Motivasyonu Ölçeği için Ülkelerin TSYDK Yöntemi ile Elde Edilen Tepki Stili Tabanlı Kümeler (H=1,2) Üzerindeki Dağılımı

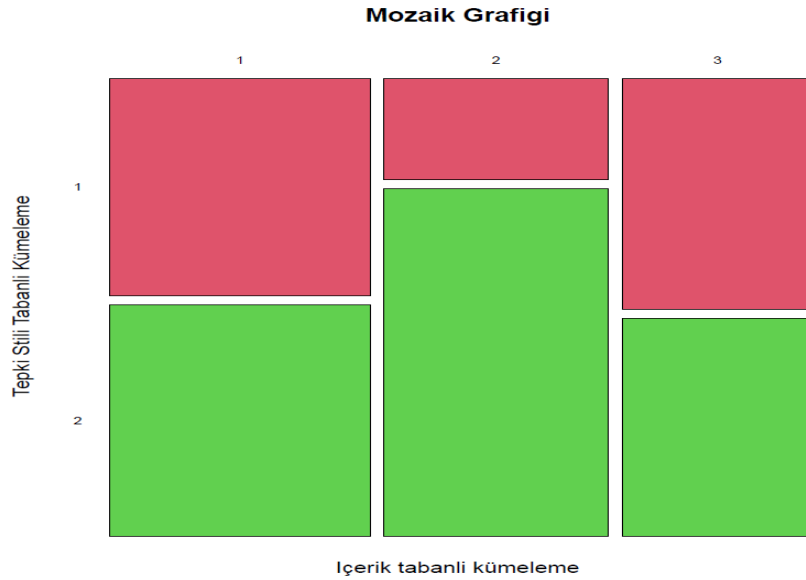
Ülkeler	H=1		H=2	
	Frekans	Yüzde	Frekans	Yüzde
Avusturya	160	34,0	310	66,0
Belçika	144	22,1	507	77,9
Brezilya	405	45,8	480	54,2
Hırvatistan	104	24,5	321	75,5
Danimarka	167	33,2	336	66,8
Dominik Cumhuriyeti	139	63,5	80	36,5
Finlandiya	73	16,4	372	83,6
Fransa	117	28,6	292	71,4
Almanya	70	28,8	173	71,2
İsrail	333	74,3	115	25,7
İtalya	202	22,7	686	77,3
Japonya	157	27,4	417	72,6
Letonya	134	35,5	243	64,5
Lüksemburg	115	33,0	234	67,0
Meksika	261	47,6	287	52,4
Hollanda	62	15,0	351	85,0
Peru	221	50,2	219	49,8
Katar	505	68,3	234	31,7
Rusya	125	29,6	297	70,4
Birleşik Arap Emir.	659	65,5	347	34,5
Türkiye	281	63,3	163	36,7
Toplam	4.434	40,7	6.464	59,3

Tablo 8 incelendiğinde araştırmada yer alan öğrencilerin %40,7'si birinci kümede ve %59,3'ü ise ikinci kümede yer almaktadır. Her bir kümedeki ülkelerin yüzde değerlerine bakıldığında, birinci kümede %74,3 değeri ile öğrencilerinin çoğunluğunun bulunduğu ülke İsrail olurken, ikinci kümede %85'lik değeri ile öğrencilerinin çoğunluğunun bulunduğu ülke Hollanda olmuştur.

Tablo 8 incelendiğinde, tutum-başarı paradoksuna göre üçe ayrılan ülkelerden birinci grupta yer alan ve Fene yönelik tutumu olumlu-Fen Başarısı düşük olan ülkelerden Dominik Cumhuriyeti, Katar, Birleşik Arap Emirlikleri ve Türkiye ülkelerinde öğrencilerin büyük çoğunluğunun birinci kümede yer aldığı görülmektedir. Tutum-başarı paradoksuna göre pozitif yönlü uç kategorileri seçme eğiliminin yüksek olduğu ülkelerin tepki stillerine göre kümelenmesinde de öğrencilerinin büyük çoğunluğunun aynı kümede yer aldıkları söylenebilir. Bu ülkelerin daha çok “4” kategorisini seçtiği düşünüldüğünde öğrencilerinin büyük çoğunluğunun aynı kümede yer almaları beklenen bir durumdur.

Tablo 8'e bakıldığında, tutum-başarı paradoksuna göre üçe ayrılan ülkelerden ikinci grupta yer alan ve Fene yönelik tutum ortalaması olumsuz-Fen Başarısı yüksek olan bütün ülkelerin (Avusturya, Belçika, Finlandiya, Fransa, Almanya, Japonya ve Hollanda) öğrencilerinin büyük çoğunluğunun ikinci kümede yer aldığı görülmektedir. Buna göre negatif yönlü uç kategorileri seçme eğiliminin yüksek olduğu ülkelerin tepki stiline göre kümelenmesinde de aynı kümede yer aldıkları görülmektedir. Bu ülkelerin daha çok "1" kategorisini seçtiği düşünüldüğünde, öğrencilerinin büyük çoğunluğunun aynı kümede yer almaları beklenen bir durumdur. Tablo 8'e göre, tutum-başarı paradoksuna göre üçe ayrılan ülkelerden üçüncü grupta yer alan ve Fene yönelik tutum ortalaması orta düzeyde-Fen Başarısı orta düzeyde olan ülkelere İsrail hariç diğer ülkelerin (Danimarka, Letonya, Rusya, İtalya, Hırvatistan, Lüksemburg) öğrencilerinin büyük çoğunluğunun ikinci kümede yer aldıkları görülmektedir.

İçerik tabanlı kümeler ile tepki stili tabanlı kümelerin nasıl karşılık geldiğini göstermek için tepki stili tabanlı kümeleme ile içerik tabanlı kümelemenin çapraz tablosunu görselleştiren mozaik grafik çizimi yapılmıştır. İçerik tabanlı kümeleme ile tepki stili tabanlı kümeleme sonuçlarının mozaik grafiği Şekil 15'de gösterilmektedir.



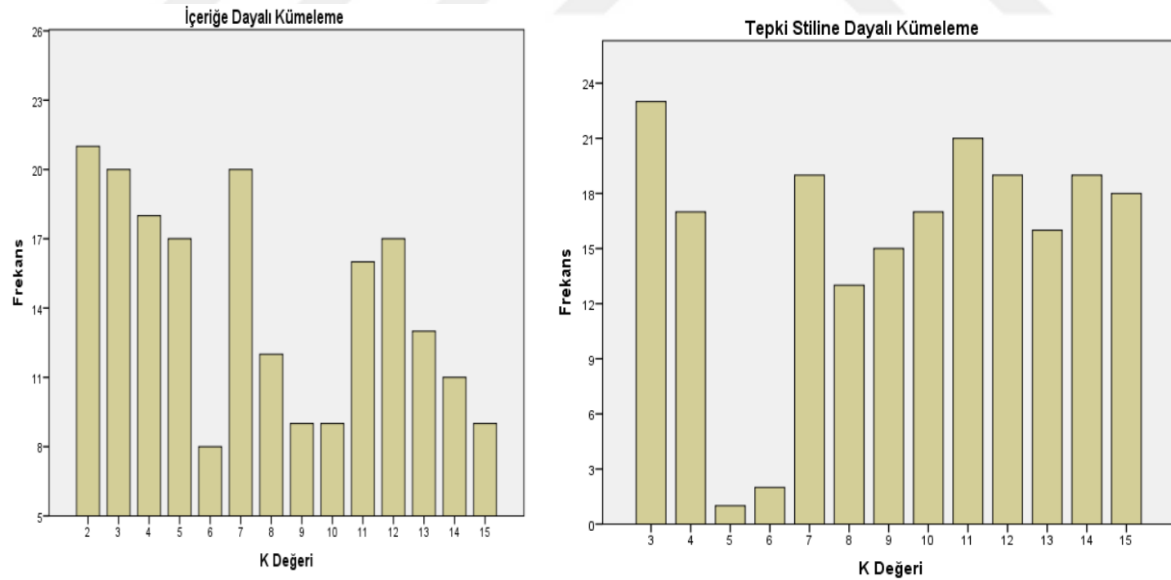
Şekil 15. Başarı motivasyonu ölçeği için TSYDK yöntemi ile elde edilen içerik tabanlı ve tepki stili tabanlı kümelerin mozaik grafiği.

Şekil 15 incelendiğinde, içerik tabanlı birinci, ikinci ve üçüncü kümeler ile tepki stili tabanlı iki kümenin çakışmadığı görülmektedir. Buna göre her üç içerik tabanlı kümede tüm tepki stili tabanlı kümelerden yanıt verenlerin bulunduğu anlaşılmaktadır. Diğer bir ifade ile içerik

tabanlı kümeleme ile tepki stili tabanlı kümeleme sonuçları çakışmaz. Buna göre, TSYDK yönteminin öğrencileri ölçülen özelliğe göre kümelemede başarılı olduğu söylenebilir.

Fen Öz Yeterliği Ölçeği için Elde Edilen Bulgular

Fen öz yeterliği ölçeği için TSYDK yöntemi ile içeriğe dayalı kümeleme ve tepki stiline dayalı kümelemede ilk olarak uygun küme sayısına karar verilmiştir. Uygun küme sayısına karar vermek için araştırma verisi olarak kullanılan 10898 kişilik veri setinden 200 bootstrap örneklem elde edilmiş ve her bir bootstrap örneklemi için minimum küme sayısı $K=2$ ile maksimum küme sayısı $K=15$ olacak şekilde her bir küme sayısı için KL indeksi hesaplanmıştır. Her bootstrap örneklemi için KL indeksini maksimize eden K değeri uygun küme sayısı olarak değerlendirilmektedir. Toplam 200 farklı örneklem için 200 farklı küme sayısı belirlenmiştir. Temel analizler için içerik tabanlı ve tepki stiline dayalı kümeleme de uygun küme sayısı olarak 200 tane küme sayısı içinde en çok tekrar eden değerler belirlenmiştir. TSYDK yöntemi için ayrıca $\lambda=,8$ değeri analizlerde kullanılmıştır. İçeriğe dayalı kümeleme ve tepki stiline dayalı kümelemede uygun küme sayısına karar vermek için bootstrap örneklemelerinden elde edilen küme sayıları Şekil 16'da gösterilmektedir.

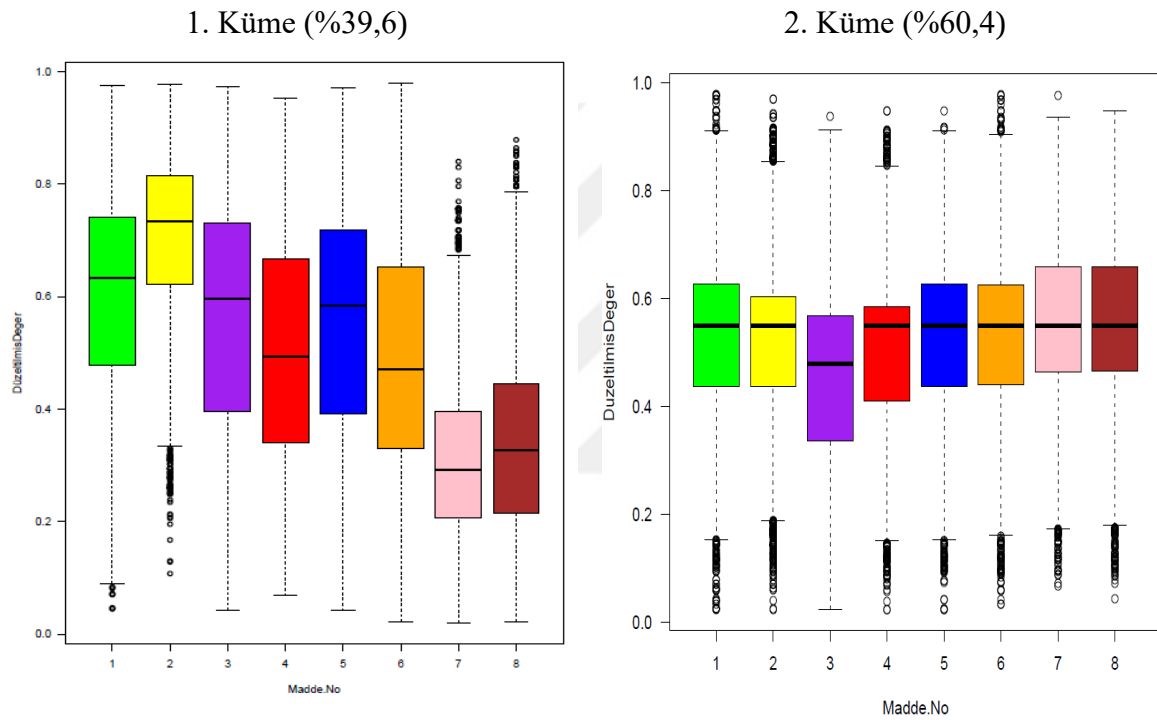


Şekil 16. Fen öz yeterliği ölçeği için TSYDK yöntemi ile elde edilen içeriğe ve tepki stiline dayalı bootstrap örneklemelerinin küme sayılarını gösteren histogram grafikleri.

Şekil 16'daki değerlere bakıldığında 200 bootstrap örneğinde içeriğe dayalı küme sayısında en fazla iki kümenin oluştuğu, tepki stiline dayalı kümelemede ise en fazla üç kümenin oluştuğu görülmektedir. Buna göre, TSYDK yöntemi ile yapılan temel analizlerde

içeriğe dayalı küme sayısı $K=2$ ve tepki stili küme sayısı ise $H=3$ olarak belirlenmiştir. Parametreler ayarlandıktan sonra TSYDK yöntemi 500 farklı başlangıç değeri için uygulanmıştır.

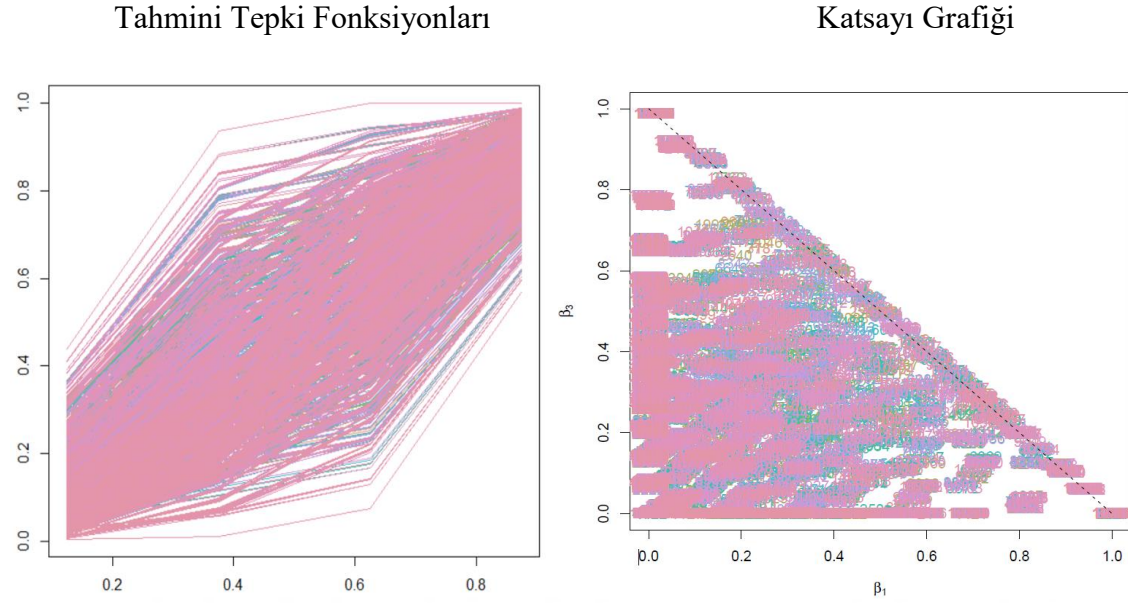
TSYDK yöntemi ile verilerin tepki stili yanlılığının düzeltilmesi ve kümelenmesinden elde edilen bulguların incelenmesinde, ilk olarak düzeltilen değerler üzerinden kümelerin nasıl farklılaştığına bakılmıştır. Bu doğrultuda TSYDK yöntemi ile içeriğe dayalı iki küme olduğu için her iki küme içinde maddelerin düzeltilen değerlerinin kutu grafikleri Şekil 17’de gösterilmektedir.



Şekil 17. Fen öz yeterliği ölçeği için TSYDK yöntemi ile elde edilen içerik tabanlı kümelerde sekiz maddenin kutu grafikleri.

Şekil 17’de dikey eksen tahmini tepki fonksiyonları kullanılarak elde edilen düzeltilmiş verilerin ölçeğini gösterirken, yatay eksen madde numaralarını göstermektedir. Düzeltilmiş değerler kullanılarak elde edilen içerik tabanlı kümelemede öğrencilerin %39,6’sı birinci kümede yer alırken, öğrencilerin %60,4’ü ise ikinci kümede bulunmaktadır. Şekil 17 incelendiğinde, kümelerin maddelerin değerlendirilmesinde nasıl farklılaştığı görülmektedir. Buna göre, birinci ve ikinci kümedeki maddelerin kutu grafikleri incelendiğinde maddelerin değerlerinin ve grafiklerinin her iki küme içinde benzer olduğu görülmektedir. Her iki küme içinde maddelerin değerleri incelendiğinde farklılıkların Madde2, Madde 7 ve Madde 8 için geçerli olduğu söylenebilir.

TSYDK yöntemi ile katılımcıya özgü tepki fonksiyonlarının kümelenmesi ile tepki stili tabanlı üç küme elde edilmişti. Araştırmada yer alan her bir öğrenci için karşılık gelen tepki fonksiyonları Şekil 18'de (solda) gösterilmektedir. Ayrıca, TSYDK yöntemi ile tahmin edilen tepki fonksiyonlarındaki katsayılar tepki stillerinin özelliklerini görsel olarak göstermek içinde kullanılabilmektedir. Buna göre her bir öğrencinin β_1 ve β_3 eksenleri ile oluşturulan düzlemdeki yerleri Şekil 18' de (sağda) gösterilmektedir.



Şekil 18. (Sol) Fen öz yeterliği ölçeği için TSYDK yöntemi ile elde edilen tepki stili tabanlı kümelerin tahmini tepki fonksiyonları. (Sağ) Tepki - stili tabanlı kümelerin β_h değeri için düşük boyutlu çizimi.

Şekil 18 incelendiğinde sol taraftaki grafikte y eksenini düzeltilmiş değerler ölçeğini gösterirken, x eksenini gözlenen ölçek değerlerini göstermektedir. Şekil 18'de sağ taraftaki grafikte yer alan sayılar öğrenci numaralarını temsil etmektedir. Tepki stiline dayalı kümelerin ülkelerle ilgili olup olmadığını görmek için ülkelerin tepki stili kümeleri üzerindeki dağılımları Tablo 9'de gösterilmektedir.

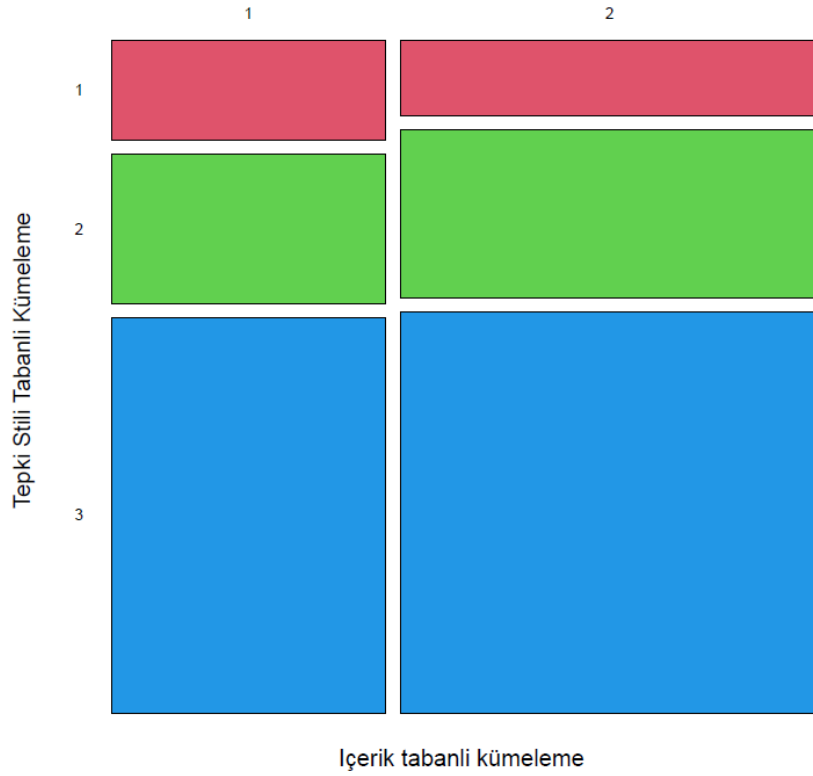
Tablo 9

Fen Öz Yeterliği Ölçeği için Ülkelerin TSYDK Yöntemi ile Elde Edilen Tepki Stili Tabanlı Kümeler (H=1,2,3) Üzerindeki Dağılımı

Ülkeler	H=1		H=2		H=3	
	Frekans	Yüzde	Frekans	Yüzde	Frekans	Yüzde
Avusturya	89	18,9	99	21,1	282	60,0
Belçika	114	17,5	157	24,1	380	58,4
Brezilya	125	14,1	249	28,1	511	57,7
Hırvatistan	59	13,9	98	23,1	268	63,1
Danimarka	70	13,9	133	26,4	300	59,6
Dominik Cumhuriyeti	22	10,0	81	37,0	116	53,0
Finlandiya	56	12,6	95	21,3	294	66,1
Fransa	90	22,0	84	20,5	235	57,5
Almanya	39	16,0	49	20,2	155	63,8
İsrail	79	17,6	127	28,3	242	54,0
İtalya	104	11,7	191	21,5	593	66,8
Japonya	90	15,7	43	7,5	441	76,8
Letonya	34	9,0	80	21,2	263	69,8
Lüksemburg	58	16,6	80	22,9	211	60,5
Meksika	35	6,4	133	24,3	380	69,3
Hollanda	68	16,5	92	22,3	253	61,3
Peru	28	6,4	124	28,2	288	65,5
Katar	80	10,8	231	31,3	428	57,9
Rusya	62	14,7	114	27,0	246	58,3
Birleşik Arap Emir.	98	9,7	339	33,7	569	56,6
Türkiye	47	10,6	117	26,4	280	63,1
Toplam	1.447	13,3	2.716	24,9	6.735	61,8

Tablo 9 incelendiğinde, araştırmada yer alan öğrencilerin %13,3'ü birinci kümede ve %24,9'u ikinci kümede ve %61,8'i ise üçüncü kümede yer almaktadır. Ayrıca tepki stili tabanlı kümeleme sonuçlarına göre, ülkelerin yüzde değerleri dikkate alındığında, Fransa'daki öğrencilerinin büyük çoğunluğunun (%22,0) birinci kümede yer aldığı görülmektedir. İkinci tepki stili kümede ise %37,0'lük değeri ile öğrencilerinin çoğunluğunun bulunduğu ülke Dominik Cumhuriyeti olmuştur. Üçüncü tepki stili kümede ise, %76,8'lik değeri ile öğrencilerinin çoğunluğunun bulunduğu ülke Japonya olmuştur.

Tepki stili tabanlı kümeler ile içerik tabanlı kümelerin nasıl karşılık geldiğini görmek için, iki kümeleme çözümünün çapraz tablosunu görselleştiren bir mozaik çizim yapılmıştır. Buna göre, Fen Öz Yeterliği için içerik tabanlı kümeleme ile tepki stili tabanlı kümeleme sonuçlarının mozaik grafiği Şekil 19'da gösterilmektedir.



Şekil 19. Fen öz yeterliği ölçeği için TSYDK yöntemi ile elde edilen içerik tabanlı ve tepki stili tabanlı kümelerin mozaik grafiği.

Şekil 19 incelendiğinde, içerik tabanlı kümelerden birinci ve ikinci kümeler ile tepki stili tabanlı üç kümenin çakışmadığı görülmektedir. Böylece, her içerik tabanlı kümede tüm tepki stili tabanlı kümelerden yanıt verenler bulunmaktadır. Bu durum TSYDK yönteminin öğrencileri ölçülen özelliğe göre kümelemede başarılı olduğunu göstermektedir.

İkinci Araştırma Sorusuna İlişkin Elde Edilen Bulgular ve Yorumlar

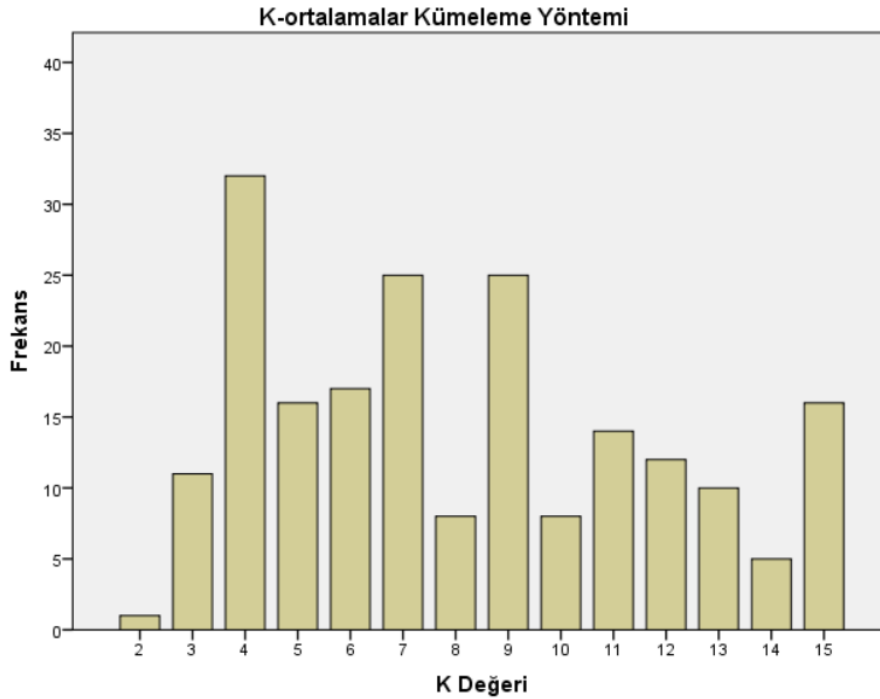
Araştırma Sorusu 2: PISA 2015 uygulamasına katılan ve tutum-başarı paradoksunun görüldüğü ülkeler K-ortalamlar yöntemi ile nasıl kümelenebilir?

İkinci araştırma sorusu kapsamında, çalışmaya dâhil edilen 21 ülkenin Fenden Hoşlanma, Başarı Motivasyonu ve Fen Öz Yeterliği ölçeklerinin gözlenen puanları üzerinden K-ortalamlar kümeleme analizi gerçekleştirilmiştir. Buna göre ilk olarak Fenden Hoşlanma

ölçeği için K-ortalamalar yöntemi ile elde edilen kümeleme analizi bulgularına yer verilmiştir.

Fenden Hoşlanma Ölçeği için Elde Edilen Bulgular

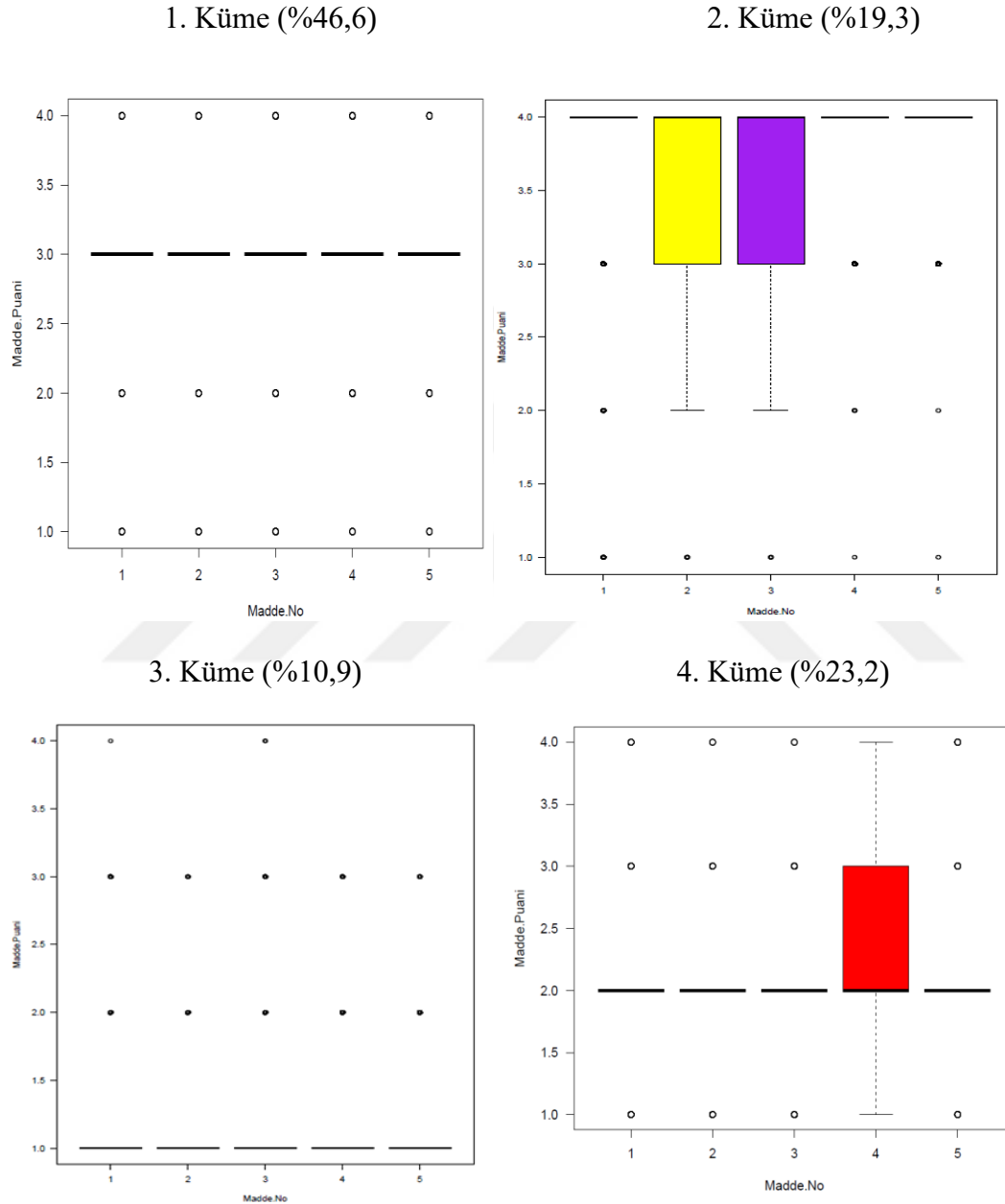
Fenden Hoşlanma ölçeği için K-ortalamalar kümeleme analizi uygulanmadan önce uygun küme sayısı belirlenmelidir. Bunun için 10898 kişilik veri setinden 200 bootstrap örneklem elde edilmiş ve her bir bootstrap örneğinde minimum küme sayısı $K=2$ ile maksimum küme sayısı $K=15$ olacak şekilde her bir küme sayısı için KL indeksi hesaplanmıştır. Her bootstrap örneklemini için KL indeksini maksimize eden K değeri uygun küme sayısı olarak değerlendirilmiştir. Toplam 200 K değeri içinde en çok tekrar eden değer uygun küme sayısı olarak belirlenmiştir. Her bir bootstrap örnekleminde elde edilen küme sayıları Şekil 20’de gösterilmektedir.



Şekil 20. Fenden hoşlanma ölçeği için K-ortalamalar kümeleme yöntemine göre bootstrap örneklemelerinin küme sayılarını gösteren histogram grafiği.

Şekil 20’deki değerlere bakıldığında toplam 200 bootstrap örneklemini üzerinden yapılan kümeleme analizlerinde 32 bootstrap örnekleminin küme sayısı 4 olarak bulunmuştur. Buna göre, K-ortalamalar kümeleme yöntemi ile yapılan temel analizlerde küme sayısı $K=4$ olarak belirlenmiştir.

K-ortalamlar kümeleme analizi ile ilk olarak madde puanları üzerinden kümelerin nasıl farklılaştığına bakılmıştır. Bu doğrultuda K-ortalamlar kümeleme yöntemi ile dört küme olduğu için her küme içinde maddelerin kutu grafikleri Şekil 21’de gösterilmektedir.



Şekil 21. Fenden hoşlanma ölçeği için K-ortalamlar kümeleme yöntemi ile elde edilen kümelerde beş maddenin kutu grafikleri.

Şekil 21’de dikey eksen orijinal tercih verilerinin ölçeğini gösterirken, yatay eksen madde numaralarını göstermektedir. K-ortalamlar kümeleme analizi sonuçlarına göre, öğrencilerin %46,6’sı birinci kümede yer alırken, %19,3’ü ikinci kümede, %10,9’u üçüncü kümede ve

%23,2'si dördüncü kümede yer almaktadır. Şekil 21'de kümelerin maddelerin değerlendirilmesinde nasıl farklılaştığı da görülmektedir. Birinci, ikinci, üçüncü ve dördüncü kümeler için maddelerin değerleri ve grafiklerinin dört küme içinde farklılık gösterdiği görülmektedir. Birinci kümedeki öğrencilerin çoğunlukla maddelerdeki 3. kategoriyi, ikinci kümedeki öğrencilerin ise çoğunlukla 4. tepki kategorisini tercih ettikleri görülmektedir. Buna karşılık, üçüncü kümede yer alan öğrencilerin çoğunlukla 1. tepki kategorisini tercih ederken, dördüncü kümede yer alan öğrenciler ise çoğunlukla 2. tepki kategorisini tercih etmişlerdir.

Birinci kümede öğrencilerin madde kategorilerinden daha çok 3. tepki kategorisini, “Katılıyorum” ifadesini tercih ettikleri anlaşılmaktadır. Maddelerdeki üçüncü kategoriler fene yönelik olumlu yönde tutumu belirttiği için bu kümede yer alan öğrencilerin fene yönelik tutumlarının olumlu yönde olduğu söylenebilir. İkinci kümede yer alan öğrenciler madde kategorilerinden daha çok 4. tepki kategorisini, yani “Kesinlikle Katılıyorum” ifadesini tercih etmektedirler. Dördüncü kategori fene yönelik olumlu yönde tutumu belirttiği için bu kümede yer alan öğrencilerin fene yönelik olumlu tutumlarının yüksek düzeyde olduğu söylenebilir. Üçüncü kümede yer alan öğrenciler madde kategorilerinden daha çok 1. tepki kategorisini, yani “Kesinlikle Katılmıyorum” ifadesini tercih etmektedirler. Birinci kategori fene yönelik olumsuz yönde tutumu belirttiği için bu kümede yer alan öğrencilerin fene yönelik olumsuz tutumlarının yüksek düzeyde olduğu söylenebilir. Dördüncü kümede yer alan öğrencilerin madde kategorilerinden daha çok 2. tepki kategorisini, yani “Katılmıyorum” ifadesini tercih ettikleri anlaşılmaktadır. Maddelerdeki ikinci kategoriler fene yönelik olumsuz yönde tutumları belirttiği için bu kümede yer alan öğrencilerin fene yönelik tutumlarının olumsuz yönde olduğu söylenebilir.

Elde edilen bu bulgulara göre K-ortalamlar kümeleme yöntemine göre kümelemenin bireyleri ölçülen içerik açısından değilde tepki stillerine göre kümelediği söylenebilir. K-ortalamlar kümeleme yöntemine göre elde edilen kümelerde ülkelerin dağılımları Tablo 10'da gösterilmektedir.

Tablo 10

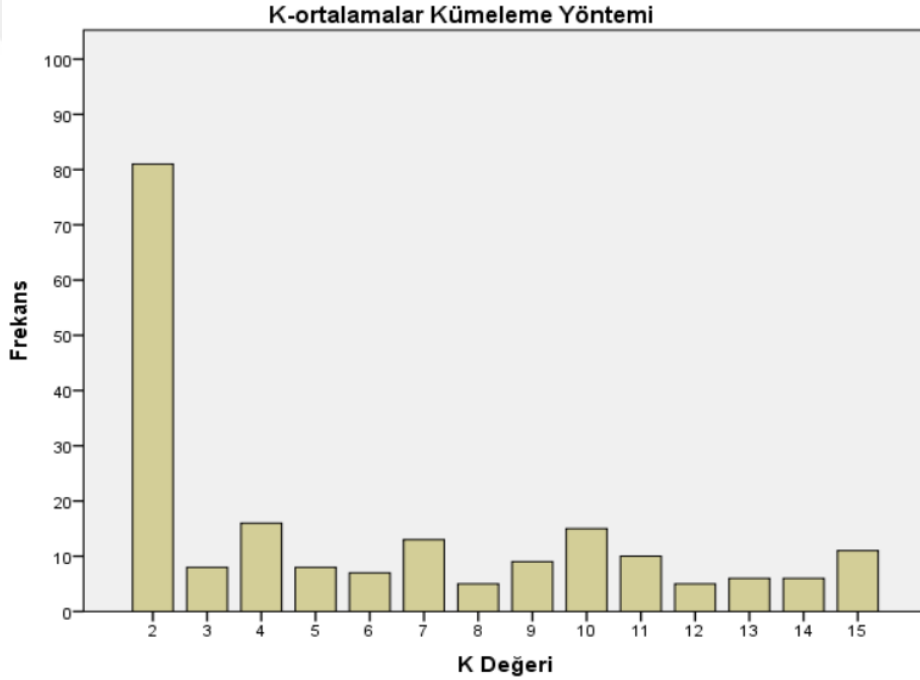
Fenden Hoşlanma Ölçeği için Ülkelerin K-ortalamlar Kümeleme Yöntemi ile Elde Edilen Kümelerdeki Dağılımı

	K=1		K=2		K=3		K=4	
Ülkeler	Frekans	Yüzde	Frekans	Yüzde	Frekans	Yüzde	Frekans	Yüzde
Avusturya	129	27,4	86	18,3	107	22,8	148	31,5
Belçika	298	45,8	111	17,1	88	13,5	154	23,7
Brezilya	488	55,1	153	17,3	55	6,2	189	21,4
Hırvatistan	197	46,4	57	13,4	42	9,9	129	30,4
Danimarka	244	48,5	96	19,1	52	10,3	111	22,1
Dominik Cumh.	106	48,4	76	34,7	18	8,2	19	8,7
Finlandiya	201	45,2	56	12,6	37	8,3	151	33,9
Fransa	198	48,4	68	16,6	58	14,2	85	20,8
Almanya	87	35,8	41	16,9	50	20,6	65	26,7
İsrail	172	38,4	109	24,3	68	15,2	99	22,1
İtalya	455	51,2	110	12,4	95	10,7	228	25,7
Japonya	178	31,0	84	14,6	103	17,9	209	36,4
Letonya	202	53,6	58	15,4	25	6,6	92	24,4
Lüksemburg	142	40,7	71	20,3	51	14,6	85	24,4
Meksika	290	52,9	160	29,2	23	4,2	75	13,7
Hollanda	145	35,1	26	6,3	92	22,3	150	36,3
Peru	235	53,4	120	27,3	24	5,5	61	13,9
Katar	370	50,1	192	26,0	52	7,0	125	16,9
Rusya	226	53,6	44	10,4	30	7,1	122	28,9
Birleşik Arap E.	500	49,7	290	28,8	70	7,0	146	14,5
Türkiye	214	48,2	96	21,6	50	11,3	84	18,9
Toplam	5.077	46,6	2.104	19,3	1190	10,9	2.527	23,2

Tablo 10 incelendiğinde, çalışmada yer alan öğrencilerin %46,6'sı birinci kümede, %19,3'ü ikinci kümede, %10,9'u üçüncü kümede ve %23,2'si ise dördüncü kümede yer almaktadır. Tablo 10'da her bir kümede ülkelerin yüzdelik değerlerine bakıldığında %55,1 değerleri ile öğrencilerinin çoğunluğunun birinci kümede olduğu ülkeler Brezilya olmuştur. İkinci kümede %34,7'lik değeri ile öğrencilerinin çoğunluğunun bulunduğu ülkeler Dominik Cumhuriyeti olmuştur. Öğrencilerinin çoğunluğunun üçüncü kümede olduğu ülke %22,8 değeri ile Avusturya olurken, dördüncü kümede ise %36,4 değeri ile öğrencilerinin çoğunluğunun bulunduğu ülke Japonya olmuştur. Birinci ve ikinci kümelerde daha çok fene yönelik olumlu yönde tutuma sahip öğrencilerin, üçüncü ve dördüncü kümelerde ise daha çok fene yönelik olumsuz yönde tutuma sahip olan öğrencilerin bulunduğu bilinmekteydi. Dominik Cumhuriyeti ve Brezilya gibi ülkelerin fene yönelik tutumu olumlu-fen başarısı düşük olan ülkelere göre ve öğrencilerinin büyük çoğunluğunun 1. ve 2. kümelerde yer alması buna karşılık, Avusturya ve Japonya gibi ülkelerin fene yönelik tutumu olumsuz-fen başarısı yüksek olan ülkelere göre ve öğrencilerinin büyük çoğunluğunun 3. ve 4. kümelerde yer alması, K-ortalamlar kümeleme yönteminin öğrencileri ölçülen özellik açısına göre değil, tepki stillerine göre kümelediği söylenebilir.

Başarı Motivasyonu Ölçeği için Elde Edilen Bulgular

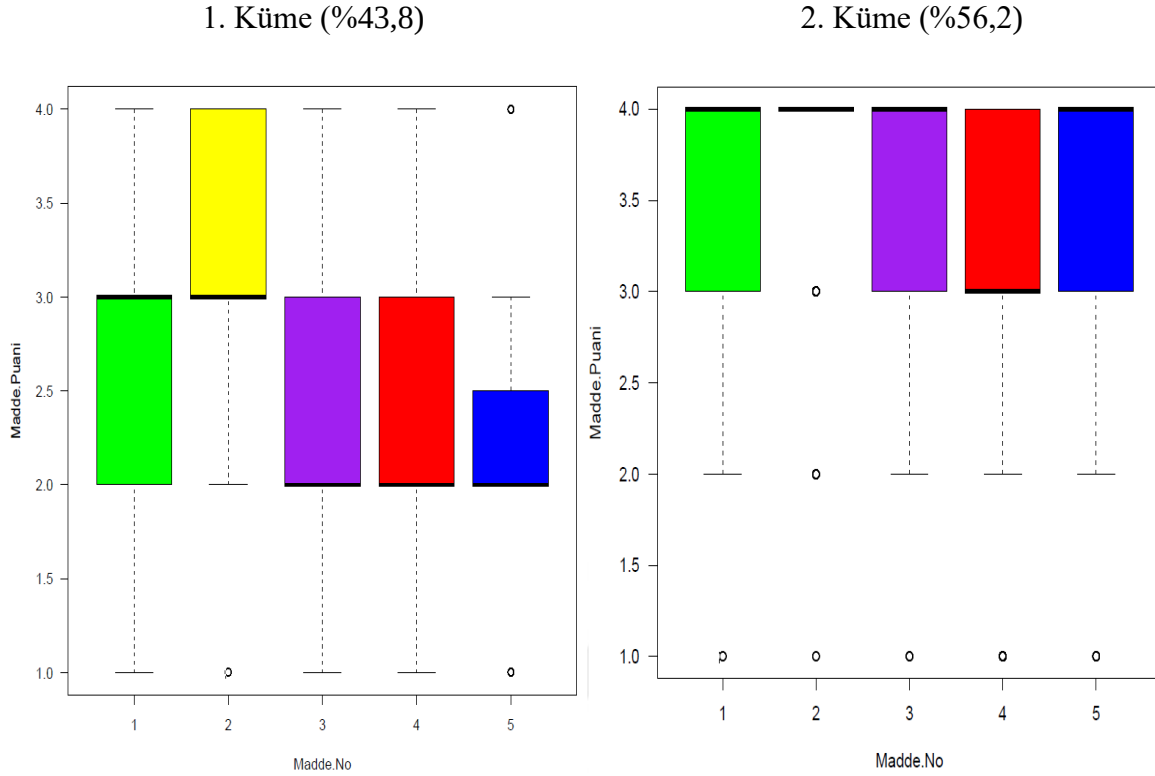
Başarı Motivasyonu ölçeği için K-ortalamlar kümeleme analizi uygulanmadan önce uygun küme sayısı belirlenmelidir. Bunun için 10898 kişilik veri setinden 200 tane bootstrap örneklem elde edilmiş ve her bir bootstrap örnekleminde minimum küme sayısı $K=2$ ile maksimum küme sayısı $K=15$ olacak şekilde her bir küme sayısı için KL indeksi hesaplanmıştır. Her bootstrap örneklemini için hesaplanan KL indeksini maksimize eden K değeri uygun küme sayısı olarak belirlenmiştir. Toplam 200 bootstrap örnekleminde KL indeksini maksimize eden küme sayılarının frekansı hesaplanmıştır. Toplam 200 K değeri içinde en çok tekrar eden değer uygun küme sayısı olarak belirlenmiştir. Her bir bootstrap örnekleminde elde edilen küme sayısı Şekil 22’de gösterilmektedir.



Şekil 22. Başarı motivasyon ölçeği için K-ortalamlar kümeleme yöntemine göre bootstrap örneklemelerinin küme sayılarını gösteren histogram grafiği.

Şekil 22’deki değerlere bakıldığında toplam 200 bootstrap örneklemini üzerinden yapılan kümeleme analizlerinde 81 bootstrap örnekleminin küme sayısı 2 olarak bulunmuştur. Buna göre, K-ortalamlar kümeleme yöntemi ile yapılan temel analizlerde küme sayısı $K=2$ olarak belirlenmiştir.

K-ortalamlar kümeleme analizi ile ilk olarak madde puanları üzerinden kümelerin nasıl farklılaştığına bakılmıştır. Bu doğrultuda K-ortalamlar kümeleme yöntemi ile iki küme olduğu için her iki küme içinde maddelerin kutu grafikleri Şekil 23’de gösterilmektedir.



Şekil 23. Başarı motivasyonu ölçeği için K-ortalamalar kümeleme yöntemi ile elde edilen kümelerde beş maddenin kutu grafikleri.

Şekil 23’de dikey eksen orijinal tercih verilerinin ölçeğini gösterirken, yatay eksen madde numaralarını göstermektedir. K-ortalamalar kümelemede öğrencilerin %43,8’i birinci kümede yer alırken, %56,2’si ikinci kümede yer almaktadır. Şekil 23’de kümelerin maddelerin değerlendirilmesinde nasıl farklılaştığı da görülmektedir. Birinci ve ikinci kümede maddelerin değerleri ve grafiklerinin iki küme içinde farklılık gösterdiği görülmektedir.

Birinci kümedeki öğrencilerin çoğunlukla maddelerdeki 2. ve 3. kategoriye, ikinci kümedeki öğrencilerin ise çoğunlukla 3. ve 4. tepki kategorilerini tercih ettikleri görülmektedir. Birinci kümede öğrencilerin madde kategorilerinden daha çok 2. tepki kategorisini “Katılmıyorum” ve 3. tepki kategorisi, “Katılıyorum” ifadesini tercih ettikleri anlaşılmaktadır. İkinci kümede öğrencilerin madde kategorilerinden daha çok 3. tepki kategorisini “Katılıyorum” ve 4. tepki kategorisi, “Kesinlikle Katılıyorum” ifadesini tercih ettikleri anlaşılmaktadır. Elde edilen bu bulgulara göre K-ortalamalar kümeleme yöntemine göre kümelemenin bireyleri ölçülen özelliğe göre değil, tepki stillerine göre kümelediği söylenebilir. K-ortalamalar kümeleme yöntemine göre elde edilen kümelerde ülkelerin dağılımları Tablo 11’de gösterilmektedir.

Tablo 11

Başarı Motivasyonu Ölçeği için Ülkelerin K-ortalamalar Yöntemi ile Elde Edilen Kümelerdeki Dağılımı

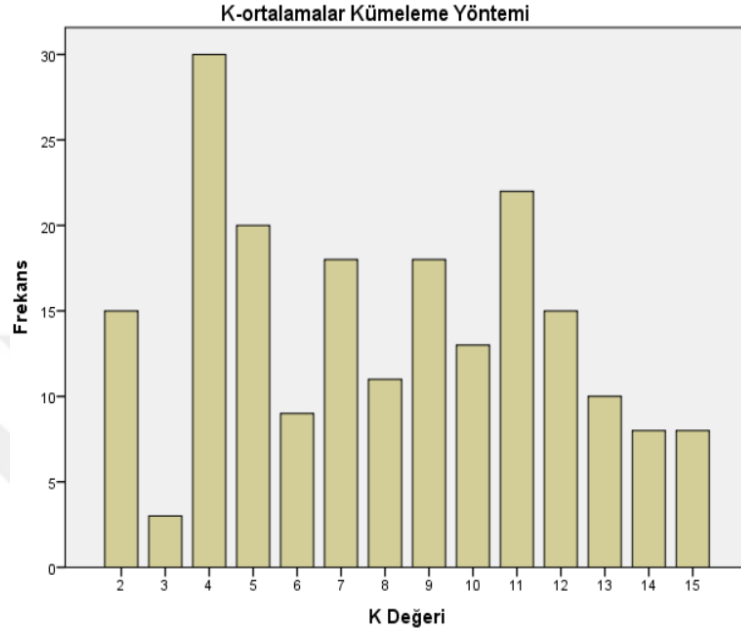
Ülkeler	K=1		K=2	
	Frekans	Yüzde	Frekans	Yüzde
Avusturya	286	60,9	184	39,1
Belçika	454	69,7	197	30,3
Brezilya	364	41,1	521	58,9
Hırvatistan	231	54,4	194	45,6
Danimarka	262	52,1	241	47,9
Dominik Cumhuriyeti	54	24,7	165	75,3
Finlandiya	327	73,5	118	26,5
Fransa	250	61,1	159	38,9
Almanya	157	64,6	86	35,4
İsrail	55	12,3	393	87,7
İtalya	530	59,7	358	40,3
Japonya	401	69,9	173	30,1
Letonya	171	45,4	206	54,6
Lüksemburg	200	57,3	149	42,7
Meksika	144	26,3	404	73,7
Hollanda	311	75,3	102	24,7
Peru	114	25,9	326	74,1
Katar	94	12,7	645	87,3
Rusya	190	45,0	232	55,0
Birleşik Arap E.	101	10,0	905	90,0
Türkiye	80	18,0	364	82,0
Toplam	4.776	43,8	6.122	56,2

Tablo 11 incelendiğinde, çalışmada yer alan öğrencilerin %43,8'i birinci kümede ve %56,2'si ise ikinci kümede yer almaktadır. Ayrıca, her bir kümedeki yüzde değerlerine bakıldığında, birinci kümede %75,3 değeri ile öğrencilerinin çoğunluğunun bulunduğu ülke Hollanda olmuştur. İkinci kümede %90'lık değeri ile öğrencilerinin çoğunluğunun bulunduğu ülke Birleşik Arap Emirlikleri olmuştur.

Fen Öz Yeterliği Ölçeği için Elde Edilen Bulgular

Fen Öz Yeterliği ölçeği için K-ortalamalar kümeleme analizi uygulanmadan önce uygun küme sayısı belirlenmelidir. Bunun için 10898 kişilik veri setinden 200 tane bootstrap örneklem elde edilmiş ve her bir bootstrap örneğinde minimum küme sayısı K=2 ile maksimum küme sayısı K=15 olacak şekilde her bir küme sayısı için KL indeksi hesaplanmıştır. Her bootstrap örnekleme için KL indeksini maksimize eden K değeri uygun

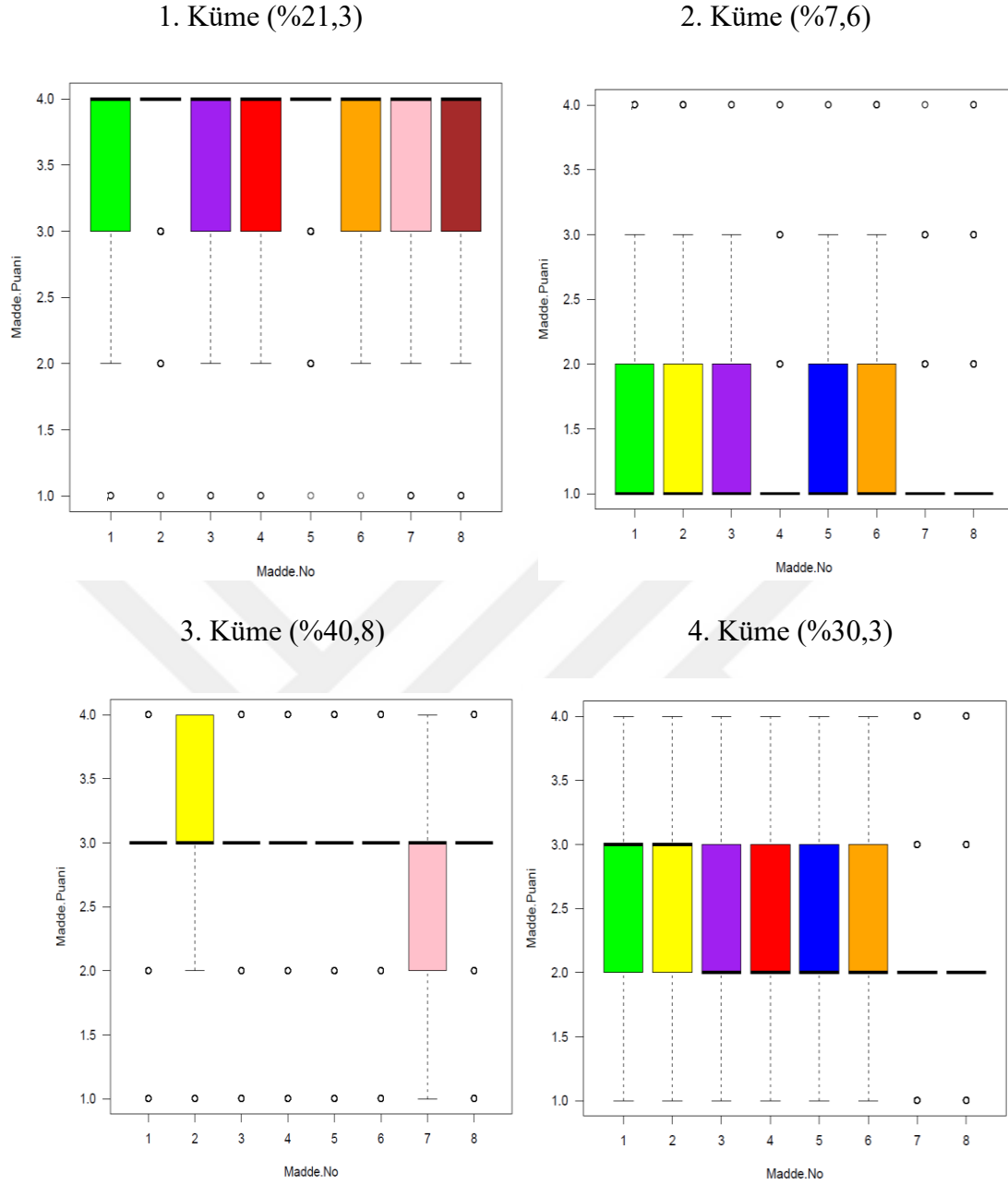
küme sayısı olarak değerlendirilmiştir. Toplam 200 bootstrap örnekleminde KL indeksini maksimize eden küme sayılarının frekansı elde edilmiştir. Toplam 200 K değeri içinde en çok tekrar eden değer, uygun küme sayısı olarak belirlenmiştir. Her bootstrap örnekleminde elde edilen küme sayısı Şekil 24’de gösterilmektedir.



Şekil 24. Fen öz yeterliği ölçeği için K-ortalamlar kümeleme yöntemine göre bootstrap örneklemlerinin küme sayılarını gösteren histogram grafiği.

Şekil 24’deki değerlere bakıldığında toplam 200 bootstrap örnekleminin küme sayısı 4 olarak bulunmuştur. Buna göre, K-ortalamlar kümeleme yöntemi ile yapılan temel analizlerde küme sayısı $K=4$ olarak belirlenmiştir.

K-ortalamlar kümeleme analizi ile ilk olarak madde puanları üzerinden kümelerin nasıl farklılaştığına bakılmıştır. Bu doğrultuda K-ortalamlar kümeleme yöntemi ile dört küme oluştuğu için her küme içinde maddelerin kutu grafikleri Şekil 25’de gösterilmektedir.



Şekil 25. Fen öz yeterliği ölçeği için K-ortalamalar kümeleme yöntemi ile elde edilen kümelerde sekiz maddenin kutu grafikleri.

Şekil 21’de dikey eksen orijinal tercih verilerinin ölçeğini gösterirken, yatay eksen madde numaralarını göstermektedir. K-ortalamalar kümeleme analizi sonuçlarına göre, öğrencilerin %21,3’ü birinci kümede yer alırken, %7,6’sı ikinci kümede, %40,8’i üçüncü kümede ve %30,3’ü dördüncü kümede yer almaktadır. Şekil 21’de kümelerin maddelerin değerlendirilmesinde nasıl farklılaştığı da görülmektedir. Birinci, ikinci, üçüncü ve dördüncü kümeler için maddelerin değerleri ve grafiklerinin dört küme içinde farklılık gösterdiği görülmektedir. Birinci kümedeki öğrencilerin çoğunlukla maddelerdeki 4.

kategoriyi, ikinci kümedeki öğrencilerin ise çoğunlukla 1. tepki kategorisini tercih ettikleri görülmektedir. Buna karşılık, üçüncü kümede yer alan öğrencilerin çoğunlukla 3. tepki kategorisini tercih ederken, dördüncü kümede yer alan öğrenciler ise çoğunlukla 2. tepki kategorisini tercih etmişlerdir.

Birinci kümede öğrencilerin madde kategorilerinden daha çok 4. tepki kategorisini, “Kesinlikle Katılıyorum” ifadesini tercih ettikleri anlaşılmaktadır. Maddelerdeki dördüncü kategoriler fene yönelik olumlu tutumu belirttiği için bu kümede yer alan öğrencilerin fene yönelik tutumlarının yüksek düzeyde olumlu olduğu söylenebilir. İkinci kümede yer alan öğrenciler madde kategorilerinden daha çok 1. tepki kategorisini, yani “Kesinlikle Katılmıyorum” ifadesini tercih etmektedirler. Birinci kategori fene yönelik olumsuz yönde tutumu belirttiği için bu kümede yer alan öğrencilerin fene yönelik olumsuz tutumlarının yüksek düzeyde olduğu söylenebilir. Üçüncü kümede yer alan öğrenciler madde kategorilerinden daha çok 3. tepki kategorisini, yani “Katılıyorum” ifadesini tercih etmektedirler. Üçüncü kategori fene yönelik olumlu yönde tutumu belirttiği için bu kümede yer alan öğrencilerin fene yönelik olumlu yönde tutumlarının olduğu söylenebilir. Dördüncü kümede yer alan öğrencilerin madde kategorilerinden daha çok 2. tepki kategorisini, yani “Katılmıyorum” ifadesini tercih ettikleri anlaşılmaktadır. Maddelerdeki ikinci kategoriler fene yönelik olumsuz yönde tutumları belirttiği için bu kümede yer alan öğrencilerin fene yönelik tutumlarının olumsuz yönde olduğu söylenebilir.

Elde edilen bu bulgulara göre K-ortalamlar kümeleme yöntemine göre kümelemenin bireyleri ölçülen içerik açısından değil de tepki stillerine göre kümelediği söylenebilir. K-ortalamlar kümeleme yöntemine göre elde edilen kümelerde ülkelerin dağılımları Tablo 12’de gösterilmektedir.

Tablo 12

Fen Öz Yeterliği Ölçeği için Ülkelerin K-ortalamalar Yöntemi ile Elde Edilen Kümelerdeki Dağılımı

	K=1		K=2		K=3		K=4	
Ülkeler	Frekans	Yüzde	Frekans	Yüzde	Frekans	Yüzde	Frekans	Yüzde
Avusturya	79	16,8	55	11,7	179	38,1	157	33,4
Belçika	121	18,6	70	10,8	235	36,1	225	34,6
Brezilya	214	24,2	78	8,8	349	39,4	244	27,6
Hırvatistan	72	16,9	32	7,5	194	45,6	127	29,9
Danimarka	116	23,1	31	6,2	191	38,0	165	32,8
Dominik Cumh.	78	35,6	11	5,0	81	37,0	49	22,4
Finlandiya	86	19,3	32	7,2	176	39,6	151	33,9
Fransa	68	16,6	58	14,2	151	36,9	132	32,3
Almanya	40	16,5	23	9,5	106	43,6	74	30,5
İsrail	95	21,2	38	8,5	178	39,7	137	30,6
İtalya	161	18,1	45	5,1	382	43,0	300	33,8
Japonya	38	6,6	68	11,8	207	36,1	261	45,5
Letonya	58	15,4	15	4,0	170	45,1	134	35,5
Lüksemburg	77	22,1	39	11,2	118	33,8	115	33,0
Meksika	128	23,4	15	2,7	246	44,9	159	29,0
Hollanda	86	20,8	41	9,9	162	39,2	124	30,0
Peru	119	27,0	12	2,7	192	43,6	117	26,6
Katar	200	27,1	39	5,3	324	43,8	176	23,8
Rusya	82	19,4	41	9,7	167	39,6	132	31,3
Birleşik Arap E.	299	29,7	56	5,6	430	42,7	221	22,0
Türkiye	109	24,5	26	5,9	203	45,7	106	23,9
Toplam	2.326	21,3	825	7,6	4.441	40,8	3.306	30,3

Tablo 12 incelendiğinde, çalışmada yer alan öğrencilerin %21,3'ü birinci kümede yer alırken, %7,6'sı ikinci kümede, %40,8'i üçüncü kümede ve %30,3'ü dördüncü kümede yer almaktadır. Tablo 12'de her bir kümede ülkelerin yüzdelik değerlerine bakıldığında %35,6 değeri ile öğrencilerinin çoğunluğunun birinci kümede olduğu ülke Dominik Cumhuriyeti olmuştur. İkinci kümede %14,2'lik değeri ile öğrencilerinin çoğunluğunun bulunduğu ülke Fransa olmuştur. Öğrencilerinin çoğunluğunun üçüncü kümede olduğu ülke %45,7'lik değeri ile Türkiye olurken dördüncü kümede ise %45,5'lik değeri ile öğrencilerinin çoğunluğunun 4. Kümede olduğu ülke Japonya olmuştur. Birinci ve üçüncü kümelerde daha çok fene yönelik olumlu yönde tutuma sahip öğrencilerin, ikinci ve dördüncü kümelerde ise daha çok fene yönelik olumsuz yönde tutuma sahip olan öğrencilerin bulunduğu bilinmekteydi. Dominik Cumhuriyeti ve Türkiye gibi ülkelerin fene yönelik tutumu olumlu-fen başarısı düşük olan ülkelere göre ve öğrencilerinin büyük çoğunluğunun 1. ve 3. kümelerde yer alması buna karşılık; Fransa ve Japonya gibi ülkelerin fene yönelik tutumu olumsuz-fen başarısı yüksek olan ülkelere göre ve öğrencilerinin büyük çoğunluğunun 2. ve 4. kümelerde yer alması görülmektedir. Buna göre, K-ortalamalar kümeleme

yönteminin öğrencileri ölçülen özelliğe göre değil de tepki stillerine göre kümelediği söylenebilir.

Üçüncü Araştırma Sorusuna İlişkin Elde Edilen Bulgular ve Yorumlar

Araştırma Sorusu 3: TSYDK ile K-ortalamlar yöntemlerinin ülkeleri kümeleme performanslarının karşılaştırması nasıl sonuçlar vermektedir?

Araştırmada TSYDK yönteminin etkinliğini değerlendirmek amacıyla, üçüncü araştırma sorusu kapsamında TSYDK yöntemi ile K-ortalamlar kümeleme yönteminden elde edilen bulgular karşılaştırılmıştır. Buna göre her üç ölçek için elde edilen bulgular aşağıda sırayla açıklanmıştır.

Fenden Hoşlanma Ölçeği için TSYDK ile K-ortalamlar Tekniği ile Elde Edilen Bulguların Karşılaştırılması

Fenden hoşlanma ölçeği için TSYDK yöntemi ile içerik tabanlı iki küme oluştuğu görülmüştür. Öğrencilerin bu kümelere dağılımında ise birinci kümede öğrencilerin %64,1'nin, ikinci kümede ise %35'9' nun yer aldığı anlaşılmıştır. Ayrıca TSYDK yöntemi ile tepki stili tabanlı kümeleme de gerçekleştirilebilmektedir. Fenden hoşlanma ölçeği için TSYDK yöntemi ile üç tepki stili küme elde edilmiştir. Öğrencilerin tepki stili tabanlı kümelere göre dağılımında; %11,7'sinin birinci kümede, %70,6'sı ikinci kümede ve %17,8' i ise üçüncü kümede yer almıştır. İçerik tabanlı kümeleme sonuçları ile tepki stili tabanlı kümeleme sonuçlarının mozaik grafiği incelendiğinde, içerik tabanlı kümeleme ile tepki stili kümelemenin çakışmadığı görülmüştür. Böylece, her iki içerik tabanlı kümede tüm tepki stili tabanlı kümelerden yanıt verenlerin olduğu söylenebilir.

Fenden hoşlanma ölçeği için K-ortalamlar tekniği ile elde edilen kümeleme analizine göre öğrenciler dört farklı kümeye ayrılmıştır. K-ortalamlar kümeleme analizine göre, öğrencilerin %46,6'sı birinci kümede yer alırken, %19,3'ü ikinci kümede, %10,9'u üçüncü kümede ve %23,2'si dördüncü kümede yer almıştır. Ayrıca birinci kümedeki öğrencilerin çoğunluğunun 3. kategoriyi, ikinci kümedeki öğrencilerin ise çoğunlukla 4. tepki kategorisini tercih ettikleri, üçüncü kümede yer alan öğrencilerin ise çoğunlukla 1. tepki

kategorisini tercih ettikleri ve dördüncü kümede yer alan öğrenciler ise çoğunlukla 2. tepki kategorisini tercih ettikleri görülmüştür.

TSYDK yöntemi ile elde edilen kümeleme analizi sonuçları ve K-ortalamalar yöntemi ile elde edilen kümeleme sonuçları birlikte değerlendirildiğinde; TSYDK yöntemi ile elde edilen kümelerin tepki stili yanlılığından arındırıldığı ve öğrencilerin ölçülen özelliğe göre kümelendikleri söylenebilir. Buna karşılık, K-ortalamalar tekniği ile elde edilen kümelemede ise öğrenciler belirli tepki kategorisini seçme özelliklerine göre kümelendiğidir. Maddelerin dört kategoriden oluşması ve K-ortalamalar kümeleme tekniği ile de öğrencilerin dört farklı kümeye ayrılması bu bulguyu destekler niteliktedir. Fenden Hoşlanma ölçeğinde, TSYDK yöntemi ile öğrencilerin ölçülen özelliğe sahip olma durumlarına göre kümelendikleri ve K-ortalamalar tekniğine göre ise öğrencilerin tepki stillerine göre kümelendikleri söylenebilir.

Başarı Motivasyonu Ölçeği için TSYDK ile K-ortalamalar Tekniği ile Elde Edilen Bulguların Karşılaştırılması

Başarı motivasyonu ölçeği için TSYDK yöntemi ile içerik tabanlı üç küme oluştuğu gözlemlenmiştir. Öğrencilerin bu kümelere dağılımında ise birinci kümede öğrencilerin %38,5'nin, ikinci kümede %33,2'sinin ve üçüncü kümede ise %28,3'nün yer aldığı anlaşılmıştır. Ayrıca TSYDK yöntemi ile tepki stili tabanlı kümeleme de gerçekleştirilebilmektedir. Başarı motivasyonu ölçeği için TSYDK yöntemi ile iki farklı tepki stili küme elde edilmiştir. Öğrencilerin tepki stili tabanlı kümelere göre dağılımında, %40,7'sinin birinci kümede ve %59,3'nün ise ikinci kümede yer aldıkları görülmüştür. İçerik tabanlı kümeleme sonuçları ile tepki stili tabanlı kümeleme sonuçlarının mozaik grafiği elde edildiğinde, içerik tabanlı kümeleme ile tepki stili tabanlı kümelemenin çakışmadığı görülmüştür. Böylece, her üç içerik tabanlı kümede tüm tepki stili tabanlı kümelerden yanıt verenlerin olduğu söylenebilir.

Başarı motivasyonu ölçeği için K-ortalamalar tekniği ile elde edilen kümeleme analizine göre öğrenciler iki farklı kümeye ayrılmıştır. K-ortalamalar kümeleme analizi ile elde edilen bulgulara göre; öğrencilerin %43,8'i birinci kümede yer alırken, %56,2'si ikinci kümede yer almıştır. Ayrıca birinci kümedeki öğrencilerin çoğunluğunun 2. ve 3. kategoriyi, ikinci kümedeki öğrencilerin ise çoğunlukla 3. ve 4. tepki kategorisini tercih ettikleri görülmüştür.

TSYDK yöntemi ile elde edilen kümeleme analizi sonuçları ile K-ortalamlar yöntemi ile elde edilen kümeleme sonuçları birlikte değerlendirildiğinde, TSYDK yöntemi ile elde edilen kümelerin tepki stili yanlılığından arındırıldığı ve öğrencilerin ölçülen özelliğe göre kümelendikleri söylenebilir. Buna karşılık, K-ortalamlar tekniği ile elde edilen kümelemede ise öğrenciler belirli tepki kategorisini seçme özelliklerine göre kümelendiğilerdir. TSYDK yöntemi ile elde edilen tepki stili tabanlı küme sayısının iki olması ve K-ortalamlar kümeleme tekniği ile de elde edilen küme sayısının iki olması bu bulguyu destekler niteliktedir. Ayrıca, her bir kümede yer alan öğrencilerin büyük çoğunluğunun belirli bir kategoriyi seçme yüzdeliklerinin yüksek olması (Örn. Birinci kümedeki öğrencilerin daha çok ikinci ve üçüncü kategoriyi seçmesi, ikinci kümedeki öğrencilerin daha çok üçüncü ve dördüncü kategoriyi seçmesi) K-ortalamlar kümeleme yönteminin öğrencileri tepki stillerine göre kümelediğine işaret etmektedir. Bu durumda “Başarı Motivasyonu” ölçeğinde, TSYDK yöntemi ile öğrencilerin ölçülen özelliğe göre kümeler ayrıldığı ve K-ortalamlar tekniğine göre ise öğrencilerin sahip oldukları tepki stillerine göre kümeler ayrıldığı söylenebilir.

Fen Öz Yeterliği Ölçeği için TSYDK ile K-ortalamlar Tekniği ile Elde Edilen Bulguların Karşılaştırılması

Fen Öz Yeterliği ölçeği için TSYDK yöntemi ile içerik tabanlı iki küme oluştuğu görülmüştür. Bu kümeler dağılımda ise, birinci kümede öğrencilerin %39,6’ sı ve ikinci kümede %60,4’ü yer almıştır. Ayrıca TSYDK yöntemi ile tepki stili tabanlı kümeleme de gerçekleştirilebilmektedir. Fen Öz Yeterliği ölçeği için TSYDK yöntemi ile üç farklı tepki stili küme elde edilmiştir. Ayrıca öğrencilerin tepki stili tabanlı kümeler göre dağılımına bakıldığında; öğrencilerin %13,3’ünün birinci kümede, %24,9’nun ikinci kümede ve %61,8’ nin ise üçüncü kümede yer aldıkları görülmüştür. İçerik tabanlı kümeleme sonuçları ile tepki stili tabanlı kümeleme sonuçlarının mozaik grafiği incelendiğinde, içerik tabanlı kümeleme ile tepki stili tabanlı kümelemenin çakışmadığı anlaşılmıştır. Böylece, her iki içerik tabanlı kümede tüm tepki stili tabanlı kümelerden yanıt verenlerin olduğu söylenebilir.

Fen Öz Yeterliği ölçeği için K-ortalamlar tekniği ile elde edilen kümeleme analizine göre öğrenciler dört farklı kümeye ayrılmıştır. K-ortalamlar kümeleme analizine göre öğrencilerin %21,3’ü birinci kümede yer alırken, %7,6’sı ikinci kümede, %40,8’ i üçüncü kümede ve %30,3’ ü ise dördüncü kümede yer almıştır. Ayrıca, birinci kümedeki

öğrencilerin çoğunlukla maddelerdeki 4. kategoriye, ikinci kümedeki öğrencilerin çoğunluğunun 1. tepki kategorisini tercih ettikleri görülmüştür. Buna karşılık, üçüncü kümede yer alan öğrenciler çoğunlukla 3. tepki kategorisini tercih ederken, dördüncü kümede yer alan öğrenciler ise çoğunlukla 2. tepki kategorisini tercih etmişlerdir.

TSYDK yöntemi ile elde edilen kümeleme analizi bulguları ve K-ortalamalar kümeleme yöntemi ile elde edilen kümeleme analizi bulguları birlikte değerlendirildiğinde, TSYDK yöntemi ile elde edilen kümelerin tepki stili yanlılığından arındırıldığı ve ölçülen özelliğe göre kümelendiklerini söyleyebiliriz. Buna karşılık, K-ortalamalar tekniği ile elde edilen kümelemede ise öğrenciler belirli tepki kategorisini seçme eğilimlerine göre kümelanmışlardır. Fen Öz Yeterliği ölçeği için TSYDK yöntemi ile öğrencilerin ölçülen özelliğe sahip olma durumlarına göre kümelendikleri ve K-ortalamalar tekniğine göre ise öğrencilerin sahip oldukları tepki stillerine göre kümelendikleri söylenebilir.

Dördüncü Araştırma Sorusuna İlişkin Elde Edilen Bulgular ve Yorumlar

a. Tepki stili yanlılığı bulunan veriler ile TSYDK yöntemi ile düzeltilmiş verilerde ortalama tutum puanları arasındaki farklar manidar mıdır?

Dördüncü araştırma sorusunun a seçeneği kapsamında PISA 2015’de tepki stili yanlılığının bulunduğu verilerin TSYDK yöntemi ile düzeltilmesinin, öğrencilerin cevapları üzerindeki tepki stillerinin etkisi araştırılmıştır. Bu doğrultuda fene yönelik tutum ile ilgili üç ölçek için ülke düzeyinde düzeltilmemiş ve düzeltilmiş ortalama puanları arasında ilişkili örneklem t testi yapılmıştır. Araştırmaya dâhil edilen toplam 21 ülkenin fenden hoşlanma, başarı motivasyonu ve fen öz yeterlik ölçeklerine ilişkin düzeltilmemiş ve düzeltilmiş ortalama tutum puanlarının t testi sonuçları aşağıda sırasıyla gösterilmiştir. Fenden hoşlanma ölçeği için ülkelerin düzeltilmemiş ve düzeltilmiş ortalama puanları arasındaki ilişkili örneklem t testi bulguları Tablo 13’ de gösterilmektedir.

Tablo 13

Ülkelerin Fenden Hoşlanma Ölçeği için Düzeltilmemiş ve Düzeltilmiş Puanlarına İlişkin İlişkili Örneklem t Testi Bulguları

Ülkeler	Ortalama Puanlar	N	$\bar{X}$	Std.S	Korelasyon	t	s.d	p	Etki Büyüklüğü
Avusturya	Düzeltilmemiş	470	,467	,303	,912	-2,938	469	,003*	,14
	Düzeltilmiş	470	,500	,072					
Belçika	Düzeltilmemiş	651	,539	,278	,903	2,086	650	,037**	,08
	Düzeltilmiş	651	,521	,070					
Brezilya	Düzeltilmemiş	885	,597	,230	,857	10,248	884	,000*	,34
	Düzeltilmiş	885	,535	,063					
Hırvatistan	Düzeltilmemiş	425	,534	,249	,877	1,913	424	,056	,09
	Düzeltilmiş	425	,516	,071					
Danimarka	Düzeltilmemiş	503	,578	,269	,884	4,942	502	,000*	,22
	Düzeltilmiş	503	,532	,071					
Dominik Cumhur.	Düzeltilmemiş	219	,666	,260	,872	8,212	218	,000*	,55
	Düzeltilmiş	219	,549	,059					
Finlandiya	Düzeltilmemiş	445	,529	,235	,883	1,245	444	,214	,06
	Düzeltilmiş	445	,518	,069					
Fransa	Düzeltilmemiş	409	,536	,273	,904	1,102	408	,271	,05
	Düzeltilmiş	409	,525	,067					
Almanya	Düzeltilmemiş	243	,492	,295	,909	-1,086	242	,279	,07
	Düzeltilmiş	243	,508	,071					
İsrail	Düzeltilmemiş	448	,562	,299	,904	3,539	447	,000*	,17
	Düzeltilmiş	448	,522	,072					
İtalya	Düzeltilmemiş	888	,536	,240	,883	2,449	887	,015**	,08
	Düzeltilmiş	888	,521	,067					
Japonya	Düzeltilmemiş	574	,469	,289	,901	-3,048	573	,002*	,13
	Düzeltilmiş	574	,498	,074					
Letonya	Düzeltilmemiş	377	,579	,231	,878	4,993	376	,000*	,26
	Düzeltilmiş	377	,533	,066					
Lüksemburg	Düzeltilmemiş	349	,543	,286	,904	1,809	348	,071	,10
	Düzeltilmiş	349	,521	,071					
Meksika	Düzeltilmemiş	548	,657	,215	,858	13,971	547	,000*	,60
	Düzeltilmiş	548	,552	,050					
Hollanda	Düzeltilmemiş	413	,415	,261	,886	-7,387	412	,000*	,36
	Düzeltilmiş	413	,488	,071					
Peru	Düzeltilmemiş	440	,651	,233	,847	11,211	439	,000*	,53
	Düzeltilmiş	440	,551	,058					
Katar	Düzeltilmemiş	739	,632	,251	,864	12,130	738	,000*	,45
	Düzeltilmiş	739	,543	,063					
Rusya Federasyonu	Düzeltilmemiş	422	,549	,223	,882	2,716	421	,001*	,13
	Düzeltilmiş	422	,527	,067					
Birleşik Arap Emir.	Düzeltilmemiş	1.006	,647	,260	,868	15,213	1.005	,000*	,48
	Düzeltilmiş	1.006	,547	,062					
Türkiye	Düzeltilmemiş	444	,580	,275	,880	4,942	443	,000*	,23
	Düzeltilmiş	444	,529	,070					

* $p \leq ,01$; ** $p \leq ,05$

Tablo 13 incelendiğinde araştırmaya katılan 14 ülkede alfa ,01 düzeyinde, 2 ülkede ise alfa ,05 düzeyinde fenden hoşlanma ölçeği düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasında anlamlı farklılıklar bulunmuştur. Ayrıca her bir ülke için elde edilen düzeltilmemiş ve düzeltilmiş puanlar arasındaki korelasyon değerlerinin ,80'den yüksek olduğu görülmektedir. Düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasında anlamlı farklılık bulunmayan ülkeler Hırvatistan, Finlandiya, Fransa, Almanya ve Lüksemburg'dur. Bu ülkelerin fenden hoşlanma ölçeği düzeltilmemiş ortalama puanları ile düzeltilmiş ortalama puanları evrende aynıdır yorumu yapılabilir.

Tablo 13 incelendiğinde, t değeri anlamlı bulunan ülkeler arasından Avusturya, Japonya ve Hollanda'nın fenden hoşlanma ölçeğine ilişkin tepki stili etkisinin dikkate alınmadığı düzeltilmemiş ortalama puanlarının düzeltilmiş ortalama puanlarından daha düşük olduğu görülmektedir. Tepki stili etkisi düzeltildikten sonra bu ülkelere ilişkin hesaplanan fenden hoşlanma ölçeği ortalama puanlarının arttığı görülmektedir. Dolayısıyla bu üç ülkede, tepki stilinin fenden hoşlanma ölçeği ortalama puanları üzerinde azaltıcı yönde rol oynadığı söylenebilir. Ayrıca bu üç ülkede araştırma kapsamında fene yönelik tutumu olumsuz, fen başarısı yüksek olan ülkeler arasında yer almaktaydı. Buna göre elde edilen bu bulgular tutum-başarı paradoksu üzerinde tepki stillerinin etkisinin olduğunu göstermektedir. Ayrıca, TSYDK yönteminin bu ülkelerde tepki stili yanlılığının düzeltilmesini sağladığı söylenebilir.

Tablo 13' e göre, t değeri anlamlı bulunan ülkeler arasından; Belçika, Brezilya, Danimarka, Dominik Cumhuriyeti, İsrail, İtalya, Letonya, Meksika, Peru, Katar, Rusya Federasyonu, Birleşik Arap Emirlikleri ve Türkiye' nin fenden hoşlanma ölçeğine ilişkin tepki stili etkisinin dikkate alınmadığı düzeltilmemiş ortalama puanlarının düzeltilmiş ortalama puanlarından daha yüksek olduğu görülmektedir. Tepki stili etkisi düzeltildikten sonra bu ülkelere ilişkin hesaplanan fenden hoşlanma ortalama puanlarının azaldığı görülmektedir. Dolayısıyla bu ülkeler için tepki stilinin, fenden hoşlanma ölçeği ortalama puanları üzerinde artırıcı yönde rol oynadığı söylenebilir.

Araştırmaya dâhil edilen ülkelerdeki öğrencilerin fenden hoşlanma ölçeğine ilişkin düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasındaki manidar farklılıklarına ilişkin etki büyüklükleri de hesaplanmıştır. Etki büyüklüğü değerleri (d) incelendiğinde, düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasındaki farkın manidar olduğu ve ülkeler içinde etki büyüklüğü değerlerinin küçük değerler altında ($d < .20$) veya küçük olduğu; Dominik Cumhuriyeti, Meksika ve Peru'ya ait etki büyüklüğü değerlerinin ise orta düzeyde ($.50 \leq d \leq .80$) olduğu görülmektedir.

Etki büyüklüğü değerinin en yüksek olduğu ülke Meksika'dır. Bu durumda diğer ülkelere kıyasla fenden hoşlanma ölçeği için Meksika'daki öğrencilerin cevaplarında tepki stili yanlılığının daha fazla olduğu ifade edilebilir. Başarı Motivasyonu ölçeği için ülkelerin düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasındaki ilişkili örneklem t testi bulguları Tablo 14' de gösterilmektedir.

Tablo 14

Ülkelerin Başarı Motivasyonu Ölçeği için Düzeltilmemiş ve Düzeltilmiş Puanlarına İlişkin İlişkili Örneklem t Testi Bulguları

Ülkeler	Ortalama Puanlar	N	$\bar{X}$	Std.S	t	Korelasyon	s.d	p	Etki Büyüklüğü
Avusturya	Düzeltilmemiş	470	,626	,217	10,971*	,920	469	,000	,51
	Düzeltilmiş	470	,535	,043					
Belçika	Düzeltilmemiş	651	,578	,202	8,146*	,886	650	,000	,32
	Düzeltilmiş	651	,526	,047					
Brezilya	Düzeltilmemiş	885	,694	,169	28,749*	,820	884	,000	,97
	Düzeltilmiş	885	,555	,033					
Hırvatistan	Düzeltilmemiş	425	,650	,174	14,819*	,846	424	,000	,72
	Düzeltilmiş	425	,548	,040					
Danimarka	Düzeltilmemiş	503	,667	,218	14,626*	,881	502	,000	,65
	Düzeltilmiş	503	,549	,044					
Dominik Cumhuri.	Düzeltilmemiş	219	,731	,188	15,504*	,860	218	,000	1,05
	Düzeltilmiş	219	,561	,031					
Finlandiya	Düzeltilmemiş	445	,533	,210	3,546*	,889	444	,000	,17
	Düzeltilmiş	445	,522	,049					
Fransa	Düzeltilmemiş	409	,638	,192	12,665*	,866	408	,000	,63
	Düzeltilmiş	409	,540	,042					
Almanya	Düzeltilmemiş	243	,601	,216	6,336*	,919	242	,000	,41
	Düzeltilmiş	243	,530	,045					
İsrail	Düzeltilmemiş	448	,850	,178	38,259*	,886	447	,000	1,81
	Düzeltilmiş	448	,577	,030					
İtalya	Düzeltilmemiş	888	,617	,198	14,826*	,882	887	,000	,50
	Düzeltilmiş	888	,538	,045					
Japonya	Düzeltilmemiş	574	,565	,235	5,287*	,914	573	,002	,22
	Düzeltilmiş	574	,523	,050					
Letonya	Düzeltilmemiş	377	,680	,199	15,052*	,855	376	,000	,78
	Düzeltilmiş	377	,551	,040					
Lüksemburg	Düzeltilmemiş	349	,634	,223	9,842*	,901	348	,000	,53
	Düzeltilmiş	349	,538	,047					
Meksika	Düzeltilmemiş	548	,725	,167	26,506*	,807	547	,000	1,13
	Düzeltilmiş	548	,562	,030					
Hollanda	Düzeltilmemiş	413	,582	,165	7,906*	,835	412	,000	,39
	Düzeltilmiş	413	,530	,042					
Peru	Düzeltilmemiş	440	,744	,158	26,999*	,761	439	,000	1,29
	Düzeltilmiş	440	,566	,028					
Katar	Düzeltilmemiş	739	,829	,193	41,192*	,850	738	,000	1,52
	Düzeltilmiş	739	,577	,032					
Rusya Federasyonu	Düzeltilmemiş	422	,678	,184	16,015*	,788	421	,000	,78
	Düzeltilmiş	422	,558	,040					
Birleşik Arap Emir.	Düzeltilmemiş	1.006	,841	,164	56,834*	,787	1.005	,000	1,79
	Düzeltilmiş	1.006	,580	,025					
Türkiye	Düzeltilmemiş	444	,796	,199	27,612*	,853	443	,000	1,31
	Düzeltilmiş	444	,572	,034					

* $p \leq ,01$

Tablo 14 incelendiğinde araştırmaya katılan tüm ülkelerde alfa ,01 düzeyinde başarı motivasyonu ölçeği için düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasında anlamlı farklılıklar bulunmuştur. Ayrıca her bir ülke için elde edilen düzeltilmemiş ve düzeltilmiş puanlar arasındaki korelasyon değerlerinin ,75’den yüksek olduğu görülmektedir. Tablo 14’e göre bütün ülkelerin başarı motivasyonu ölçeğine ilişkin tepki stili etkisinin dikkate alınmadığı düzeltilmemiş ortalama puanlarının düzeltilmiş ortalama puanlarından daha yüksek olduğu görülmektedir. Tepki stili yanlılığı düzeltildikten sonra bu ülkelere ilişkin hesaplanan başarı motivasyonu ortalama puanları azalmıştır. Dolayısıyla

bütün ülkeler için tepki stili yanlılığının başarı motivasyonu ölçeği ortalama puanlarını arttırıcı yönde rol oynadığı söylenebilir.

Araştırmaya dâhil edilen ülkelerdeki öğrencilerin başarı motivasyonu ölçeğine ilişkin düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasındaki manidar farklılıklarına ilişkin etki büyüklükleri de hesaplanmıştır. Etki büyüklüğü değerleri (d) incelendiğinde, düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasındaki farkın manidar olduğu ülkeler içinde etki büyüklüğü değerlerinin çoğunlukla orta ve geniş düzeyde olduğu görülmektedir. Etki büyüklüğü değerinin en yüksek olduğu ülke İsrail'dir. Bu durumda diğer ülkelere göre kıyasla başarı motivasyonu ölçeği için İsrail'deki öğrencilerin cevaplarında tepki stili yanlılığının daha fazla olduğu yorumu yapılabilir. Fen öz yeterliği ölçeği için ülkelerin düzeltilmemiş ve düzeltilmiş ortalama puanları arasındaki ilişkili örneklem t testi bulguları Tablo 15' te gösterilmektedir.

Tablo 15

Ülkelerin Fen Öz Yeterliği Ölçeği için Düzeltilmemiş ve Düzeltilmiş Puanlarına İlişkin İlişkili Örneklem t Testi Bulguları

Ülkeler	Ortalama Puanlar	N	$\bar{X}$	Std.S	Korelasyon	t	s.d	p	Etki Büyüklüğü
Avusturya	Düzeltilmemiş	470	,557	,225	,929	4,568*	469	,000	,21
	Düzeltilmiş	470	,515	,028					
Belçika	Düzeltilmemiş	651	,565	,236	,934	5,990*	650	,000	,23
	Düzeltilmiş	651	,516	,029					
Brezilya	Düzeltilmemiş	885	,601	,236	,910	11,569*	884	,000	,39
	Düzeltilmiş	885	,520	,031					
Hırvatistan	Düzeltilmemiş	425	,588	,210	,915	7,798*	424	,000	,38
	Düzeltilmiş	425	,518	,027					
Danimarka	Düzeltilmemiş	503	,608	,215	,925	10,441*	502	,000	,47
	Düzeltilmiş	503	,520	,027					
Dominik Cumhur.	Düzeltilmemiş	219	,665	,225	,892	10,148*	218	,000	,69
	Düzeltilmiş	219	,528	,028					
Finlandiya	Düzeltilmemiş	445	,580	,218	,929	6,908*	444	,000	,33
	Düzeltilmiş	445	,517	,028					
Fransa	Düzeltilmemiş	409	,541	,235	,929	2,798*	408	,005	,14
	Düzeltilmiş	409	,512	,030					
Almanya	Düzeltilmemiş	243	,583	,217	,932	5,240*	242	,000	,34
	Düzeltilmiş	243	,518	,027					
İsrail	Düzeltilmemiş	448	,590	,224	,913	7,631*	447	,000	,36
	Düzeltilmiş	448	,518	,028					
İtalya	Düzeltilmemiş	888	,595	,201	,911	12,618*	887	,000	,42
	Düzeltilmiş	888	,520	,026					
Japonya	Düzeltilmemiş	574	,495	,211	,934	-1,764	573	,078	-,07
	Düzeltilmiş	574	,509	,031					
Letonya	Düzeltilmemiş	377	,587	,185	,914	8,125*	376	,000	,42
	Düzeltilmiş	377	,519	,025					
Lüksemburg	Düzeltilmemiş	349	,579	,247	,949	5,306*	348	,000	,28
	Düzeltilmiş	349	,517	,031					
Meksika	Düzeltilmemiş	548	,633	,194	,901	14,588*	547	,000	,62
	Düzeltilmiş	548	,526	,025					
Hollanda	Düzeltilmemiş	413	,580	,231	,936	6,341*	412	,000	,31
	Düzeltilmiş	413	,517	,029					
Peru	Düzeltilmemiş	440	,644	,186	,887	14,946*	439	,000	,71
	Düzeltilmiş	440	,526	,024					
Katar	Düzeltilmemiş	739	,643	,222	,908	16,215*	738	,000	,60
	Düzeltilmiş	739	,526	,028					
Rusya Federasyonu	Düzeltilmemiş	422	,581	,236	,928	6,161*	421	,000	,30
	Düzeltilmiş	422	,518	,030					
Birleşik Arap Emir.	Düzeltilmemiş	1.006	,649	,222	,914	19,619*	1.005	,000	,62
	Düzeltilmiş	1.006	,527	,028					
Türkiye	Düzeltilmemiş	444	,632	,219	,903	11,454*	443	,000	,54
	Düzeltilmiş	444	,526	,028					

* $p \leq ,01$

Tablo 15 incelendiğinde, araştırmaya katılan bir ülke (Japonya) hariç geriye kalan 20 ülkede ,01 anlamlılık düzeyinde fen öz yeterliği ölçeği için düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasında anlamlı farklılıklar bulunmuştur. Japonya’da ise ,01 anlamlılık düzeyinde düzeltilmemiş ve düzeltilmiş fen öz yeterliği puanlarında anlamlı farklılık bulunmamıştır. Ayrıca her bir ülke için elde edilen düzeltilmemiş ve düzeltilmiş puanlar arasındaki korelasyon değerlerinin ,85’den yüksek olduğu görülmektedir. Japonya hariç diğer ülkelerin fen öz yeterliği ölçeğine ilişkin tepki stili etkisinin dikkate alınmadığı düzeltilmemiş ortalama puanlarının düzeltilmiş ortalama puanlarından daha yüksek olduğu

görülmektedir. Tepki stili etkisi düzeltildikten sonra bu ülkelere ilişkin hesaplanan fen öz yeterliği ortalama puanlarının azaldığı görülmektedir. Dolayısıyla bu ülkeler için tepki stilinin fen öz yeterliği ölçeği ortalama puanları üzerinde arttırıcı yönde rol oynadığı söylenebilir.

Tablo 15' e bakıldığında Japonya' nın fen öz yeterliği ölçeğine ilişkin tepki stili etkisinin dikkate alınmadığı düzeltilmemiş ortalama puanlarının düzeltilmiş ortalama puanlarından daha düşük olduğu görülmektedir. Tepki stili etkisi düzeltildikten sonra Japonya' ya ilişkin fen öz yeterliği ortalama puanlarının arttığı görülmektedir. Dolayısıyla Japonya için tepki stili, fen öz yeterliği ölçeği ortalama puanları üzerinde azaltıcı yönde rol oynamıştır. Ayrıca Japonya araştırma kapsamında fene yönelik tutumu olumsuz, fen başarısı yüksek olan ülkeler grubunda yer almaktaydı. Buna göre, Japonya'daki öğrencilerin cevaplarında tepki stili yanlılığı bulunduğu ve bu nedenle fendeki öz yeterliklerini olduğundan daha düşük düzeyde gösterme eğiliminde oldukları söylenebilir.

Araştırmaya dâhil edilen ülkelerdeki öğrencilerin fen öz yeterliği ölçeğine ilişkin düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasındaki manidar farklılıklarına ilişkin etki büyüklükleri de hesaplanmıştır. Etki büyüklüğü değeri en fazla olduğu ülke Peru olmuştur. Peru'daki etki büyüklüğü orta düzeydedir. Bu durumda diğer ülkelere kıyasla fen öz yeterliği ölçeği için Peru'daki öğrencilerin cevaplarında tepki stili yanlılığının daha fazla olduğu yorumu yapılabilir.

Her üç ölçek için ilişkili örneklem t testi bulguları birlikte değerlendirildiğinde, başarı motivasyonu ölçeğinde bütün ülkelerde düzeltilmemiş ve düzeltilmiş ortalama puanlarının manidar şekilde farklılaşmaktadır. Diğer bir ifade ile bütün ülkelerdeki başarı motivasyonu ölçeğine ilişkin tutum puanları tepki stilinden etkilenmektedir. Bu yüzden düzeltilmemiş ve düzeltilmiş ortalama puanları arasındaki farkın birbirinden anlamlı şekilde değiştiği söylenebilir. Buna karşılık Fenden Hoşlanma ölçeği için Hırvatistan, Finlandiya, Fransa, Almanya ve Lüksemburg ülkelerinde düzeltilmemiş ve düzeltilmiş puan ortalamaları arasında anlamlı farklılık bulunmamıştır. Buna göre, bu ülkelerdeki öğrencilerin tutum puanlarının tepki stili etkisinden fazla etkilenmediği söylenebilir.

Fenden hoşlanma ölçeği için elde edilen düzeltilmiş ortalama puanların düzeltilmemiş ortalama puanlarından yüksek olduğu ülkelerin genel olarak fenden hoşlanma düzeyleri olumsuz-fen başarısı yüksek olan ülkelerin olduğu görülmüştür. Bu ülkelerdeki öğrenciler için tepki stili yanlılığının fene yönelik tutumlarını olduğundan daha olumsuz olacak şekilde

kestirilmesine yol açtığı söylenebilir. Diğer yandan, fenden hoşlanma ölçeği için düzeltilmemiş ortalama tutum puanlarının düzeltilmiş ortalama puanlarından daha yüksek olduğu ülkelerin genellikle fenden hoşlanma düzeyleri olumlu ve fen başarısı yüksek olan ülkelerin olduğu görülmüştür. Bu ülkelerdeki öğrenciler için ise tepki stili yanlılığının fene yönelik tutumlarını olduğundan daha olumlu olarak kestirilmesine yol açtığı söylenebilir.

Başarı motivasyonu ölçeği için bütün ülkelerde düzeltilmemiş ortalama puanlarının düzeltilmiş ortalama puanlarda daha yüksek olduğu görülmüştür. Bu durumda bütün ülkeler için tepki stili yanlılığının ülkelerdeki öğrencilerin tutum puanlarını olduğundan daha olumlu olarak kestirilmesine yol açtığı söylenebilir. Başarı motivasyonu ölçeği için etki büyüklüğü değerleri incelendiğinde, başarı motivasyonu olumlu ve fen başarısı düşük olan ülkelerin geniş etki büyüklüğüne sahip oldukları görülmüştür. Buna göre bu ülkeler için tepki stili yanlılığının çok fazla olduğu ve bu ülkelerde tepki stili yanlılığının öğrencilerin başarı motivasyonlarını olduğundan çok daha yüksek kestirilmesine yol açtığı söylenebilir.

Fen öz yeterliği ölçeğinde Japonya' nın düzeltilmemiş ortalama puanlarının düzeltilmiş ortalama puanlarından daha düşük olduğu görülmüştür. Buna göre fen başarısı yüksek olan Japonya'da tepki stili etkisinin öğrencilerin fene yönelik öz yeterliklerinin olduğundan daha düşük kestirilmesine yol açmıştır. Buna karşılık Japonya hariç diğer ülkelerde fen öz yeterliklerine ilişkin düzeltilmemiş ortalama puanlarının düzeltilmiş ortalama puanlarından daha yüksek olduğu görülmüştür. Buna göre bu ülkeler için tepki stili yanlılığının öğrencilerin fen öz yeterliği puanlarını olduğundan çok daha yüksek kestirilmesine yol açtığı söylenebilir.

b. Tepki stili yanlılığı bulunan veriler ile TSYDK yöntemi ile düzeltilmiş ortalama tutum puanlarının fen başarısı ile arasındaki ülke içi ve ülkeler arası ilişkileri nasıl değişmektedir?

Dördüncü araştırma sorusunun b seçeneği kapsamında PISA 2015’de tepki stili yanlılığının bulunduğu verilerin TSYDK yöntemi ile düzeltilmesinin etkileri araştırılmıştır. Bu doğrultuda fene yönelik tutumla ilgili üç ölçek için düzeltilmemiş ve düzeltilmiş ortalama puanlarının fen başarısıyla olan ilişkilerine dair ülke içi korelasyonları hesaplanmıştır. Buna göre araştırmaya dahil edilen 21 ülkenin fenden hoşlanma, başarı motivasyonu ve fen öz yeterlik ölçeklerine ilişkin düzeltilmemiş ve düzeltilmiş ortalama puanlarının fen başarıları ile olan ülke içi korelasyon değerleri Tablo 16’ da gösterilmektedir.

Tablo 16

Ülkelerin Düzeltilmemiş ve Düzeltilmiş Fen Tutum Puanları ile Fen Başarıları Arasındaki Ülke İçi Korelasyon Değerleri

	Fenden Hoşlanma		Başarı Motivasyonu		Fen Öz Yeterliği	
Ülkeler	r_x	r_y	r_x	r_y	r_x	r_y
Avusturya	,338	,342	,151	,153	,304	,263
Belçika	,337	,307	,047	,043	,301	,248
Brezilya	,195	,109	,079	-,035	,200	,148
Hırvatistan	,264	,247	,112	,103	,258	,229
Danimarka	,291	,243	,230	,188	,325	,253
Dominik Cum.	,110	,104	,215	,140	,069	,079
Finlandiya	,293	,286	,217	,184	,338	,314
Fransa	,319	,306	,096	,078	,317	,256
Almanya	,359	,371	,167	,131	,360	,329
İsrail	,268	,241	,128	,071	,166	,109
İtalya	,222	,216	,093	,094	,135	,077
Japonya	,287	,302	,190	,180	,256	,188
Letonya	,225	,196	,214	,133	,190	,135
Lüksemburg	,329	,304	,170	,174	,265	,254
Meksika	,128	,128	,145	,081	,132	,088
Hollanda	,298	,269	,123	,105	,206	,140
Peru	,109	,095	,140	,084	,049	-,016
Katar	,282	,261	,146	,084	,195	,108
Rusya Fed.	,164	,160	,177	,071	,206	,138
Birleşik Arap E.	,242	,146	,102	,022	,154	,072
Türkiye	,177	,138	,152	,044	,130	,167

r_x = Düzeltilmemiş Tutum Puanları ile Fen Başarısı Arasındaki Korelasyon Değeri

r_y = Düzeltilmiş Tutum Puanları ile Fen Başarısı Arasındaki Korelasyon Değeri

Tablo 16 incelendiğinde araştırmaya dahil edilen ülkelerin fenden hoşlanma ölçeği için düzeltilmemiş ve düzeltilmiş puanları ile fen başarıları arasındaki korelasyon değerleri

pozitiftir. Başarı motivasyonu ölçeği için ise sadece Brezilya'nın düzeltilmiş puanları ile fen başarıları arasındaki korelasyon değerleri negatif olurken, diğer ülkelerin düzeltilmemiş ve düzeltilmiş puanları ile fen başarıları arasındaki korelasyon değerleri pozitif bulunmuştur. Fen öz yeterliği ölçeğinde, sadece Peru'nun düzeltilmiş puanları ile fen başarısı arasındaki korelasyon negatif, diğer bütün ülkelerin düzeltilmemiş ve düzeltilmiş tutum puanları ile fen başarısı arasındaki korelasyon değerleri pozitiftir.

PISA 2015 uygulamasına katılan ülkelerdeki öğrencilerin tepki stilinin etkisine göre düzeltilmemiş ve düzeltilmiş tutum puanları ile fen başarıları arasındaki ülkeler arası korelasyonları değerleri Tablo 17'de gösterilmektedir.

Tablo 17

Düzeltilmemiş ve Düzeltilmiş Fen Tutum Puanlarının Fen Başarısı ile Olan Ülkeler Arası Korelasyon Değerleri

Değişkenler	Düzeltilmemiş Puanlar	Düzeltilmiş Puanlar
Fenden Hoşlanma-Fen Başarısı	-,855	-,801
Başarı Motivasyonu-Fen Başarısı	-,849	-,849
Fen Öz Yeterliği-Fen Başarısı	-,854	-,849

Tablo 17 incelendiğinde araştırmaya dâhil edilen ülkelerdeki öğrencilerin düzeltilmemiş puanlara göre fenden hoşlanma düzeyleri ile fen başarıları arasında ülke düzeyinde negatif yönde manidar ilişki bulunmuştur ($r_{s1} = -,855; p < ,01$). Buna göre, fen başarıları yüksek olan öğrencilerin fenden hoşlanma düzeylerinin düşük olduğu veya fenden hoşlanmaları yüksek olan öğrencilerin fen başarılarının düşük olduğu söylenebilir. Ülke düzeyinde düzeltilmiş puanlar ile fenden hoşlanma ve fen başarısı arasında negatif yönde manidar ilişki bulunmuştur ($r_{s2} = -,801; p < ,01$). Düzeltilmiş puanlar aracılığıyla elde edilen fenden hoşlanma ile fen başarısı arasındaki korelasyon değerinin düzeltilmemiş puanlar aracılığıyla elde edilen korelasyon değerinden daha düşük olduğu görülmektedir. Buna göre, TSYDK yönteminin tepki stili yanlılığının düzeltilmesinde etkili olduğu söylenebilir.

Tablo 17 incelendiğinde, araştırmaya dahil edilen ülkelerde düzeltilmemiş puanlara göre başarı motivasyonu ile fen başarısı arasındaki ülkeler arası korelasyonun negatif yönde olduğu görülmektedir ($r_{s1} = -,849; p < ,01$). Buna göre, fen başarıları yüksek olan öğrencilerin başarı motivasyonlarının düşük olduğu veya başarı motivasyonları yüksek olan öğrencilerin fen başarılarının düşük olduğu söylenebilir. Ülke düzeyinde, düzeltilmiş veriler ile elde edilen başarı motivasyonu ile fen başarısı arasında negatif yönde manidar ilişki bulunmuştur ($r_{s2} = -,849, p < ,01$). Düzeltilmiş puanlar aracılığıyla elde edilen başarı motivasyonu ile fen

başarısı arasındaki korelasyon değeri, düzeltilmemiş puanlar aracılığıyla elde edilen korelasyon değeri ile aynıdır.

Tablo 17 incelendiğinde, düzeltilmemiş puanlara göre ülke düzeyinde fen öz yeterlikleri ile fen başarıları arasında negatif yönde manidar ilişki olduğu bulunmuştur ($r_{s1} = -,854$, $p < ,01$). Buna göre, fen başarıları yüksek olan öğrencilerin fen öz yeterlik düzeylerinin düşük olduğu veya fen öz yeterlikleri yüksek olan öğrencilerin fen başarılarının düşük olduğu söylenebilir. Ülke düzeyinde düzeltilmiş veriler ile elde edilen fen öz yeterlikleri ile fen başarısı arasında negatif yönde manidar ilişki bulunmuştur ($r_{s2} = -,849$, $p < ,01$). Düzeltilmiş puanlar aracılığıyla elde edilen fen öz yeterliği ile fen başarısı arasındaki korelasyon değerinin düzeltilmemiş puanlar aracılığıyla elde edilen korelasyon değerinden daha düşük olduğu görülmektedir. Buna göre, TSYDK yönteminin tepki stili yanlılığının düzeltilmesinde etkili olduğu söylenebilir. Ayrıca, düzeltilmemiş ve düzeltilmiş fen öz yeterliği ölçeği puanları ile fen başarısı arasındaki korelasyon değerlerinin birbirine yakın olduğu söylenebilir.

Her üç ölçek içinde fen başarısı ile elde edilen ilişkiler birlikte değerlendirildiğinde; düzeltilmiş ve düzeltilmemiş puanlar aracılığıyla elde edilen ülke içi ve ülkeler arası korelasyon değerlerinin birbirine yakın olması, tepki stil yanlılığının düzeltilmesinin tutum-başarı paradoksunu açıklamada tek başına yeterli olmadığını göstermektedir.

BÖLÜM V

TARTIŞMA, SONUÇ VE ÖNERİLER

Bu bölümde, araştırmanın bulguları doğrultusunda çıkarılan sonuçlar alan yazındaki diğer çalışma sonuçları ile tartışılmış ve elde edilen sonuçlar doğrultusunda önerilere yer verilmiştir.

Tartışma ve Sonuç

Yapılan araştırmanın genel amacı, PISA 2015'te görülen tutum-başarı paradoksunu açıklamada tepki stili yanlılığının etkisini belirlemek ve verilerin içerik tabanlı kümelenmesini sağlamaktır. Bu amaç doğrultusunda tepki stili yanlılığını ortadan kaldırmak ve içerik tabanlı kümeleme için TSYDK yöntemi kullanılmıştır. TSYDK yöntemi, eş zamanlı olarak tepki stili yanlılığı düzeltilmesini ve düzeltilmiş verilerin kümelenmesini sağlamaktadır. TSYDK yöntemi ile hem içerik tabanlı kümeleme hem de tepki stili tabanlı kümeleme sonuçları elde edilebilmektedir.

Birinci araştırma sorusu kapsamında PISA 2015'e katılan ve tutum-başarı paradoksunun görüldüğü ülkelerde TSYDK yöntemine göre; fenden hoşlanma ölçeği için 200 bootstrap örnekleminde içeriğe dayalı iki küme oluştuğu, tepki stiline dayalı ise üç kümenin oluştuğu görülmüştür. Başarı motivasyonu ölçeği için 200 bootstrap örnekleminde içeriğe dayalı üç kümenin oluştuğu, tepki stiline dayalı ise iki kümenin oluştuğu görülmüştür. Fen öz yeterliği ölçeği için ise 200 bootstrap örnekleminde içeriğe dayalı iki küme, tepki stiline dayalı üç küme elde edilmiştir. Her üç ölçek içinde, içerik tabanlı kümeleme ile tepki stili tabanlı kümelemenin mozaik grafiği incelenmiş ve içerik tabanlı kümeler ile tepki stili tabanlı kümelerin çakışmadığı görülmüştür. Diğer bir ifade ile her bir içerik tabanlı kümede tüm

tepki stili tabanlı kümelerden yanıt verenler bulunmaktadır. Buna göre TSYDK yönteminin ham verileri tepki stili yanlılığından arındırdığı ve öğrencileri ölçülen özellik açısından kümelere ayırdığı söylenebilir. TSYDK yönteminin her üç ölçek içinde verileri tepki stili yanlılığı açısından düzelttiği ve aynı zamanda ölçülen özelliğe göre öğrencileri kümelere ayırdığı görülmüştür. Böylece TSYDK yönteminin tepki stili yanlılığının düzeltilmesinde ve ölçülen özelliğe göre öğrencileri kümelemede başarılı olduğu sonucu ortaya çıkmıştır.

Mevcut araştırmadan elde edilen sonuçlara, Takagishi vd. (2019)' nin simülasyon ile gerçek veri setlerini kullanarak yaptıkları çalışmada da rastlanmıştır. Takagishi vd. (2019)' nin yaptığı çalışmada TSYDK yöntemi, Schoonees vd. (2015) tarafından geliştirilen ve ikiden fazla tepki stilini tespit etmek için tasarlanmış Sınırlandırılmış Çift Ölçeklendirme (SÇÖ) (Constrained Dual Scaling-CDS) yöntemi ve K-ortalamlar kümeleme yöntemi ile karşılaştırılmıştır. Takagishi vd.(2019) yaptıkları simülasyon çalışması sonucunda, TSYDK yönteminin tepki stili yanlılığının düzeltilmesinde ve verilerin içerik tabanlı kümeleneğinde SÇÖ'den daha başarılı olduğu sonucu elde edilmiştir. Çalışmada ayrıca TSYDK yöntemi, SÇÖ yöntemi ve K-ortalamlar kümeleme yöntemi sosyal değerlerle ilgili dört farklı ülkeden elde edilen verilere uygulanmıştır. Araştırma kapsamında analiz edilen veriler, 2008 yılında Çin, Japonya, Kore ve Tayvan'dan gelen 8745 katılımcıya uygulanan Doğu Asya Sosyal Anketi'nden (DASA) elde edilmiştir. Uygulamada ankette yer alan 7 kategorili maddelerden 10 tanesi ele alınmıştır. Araştırmada TSYDK yöntemine göre elde edilen içerik tabanlı ve tepki stili tabanlı küme sayıları birbirinden farklı bulunmuştur. Ayrıca çalışmada içerik tabanlı kümeleme ile tepki stili tabanlı kümelerin mozaik grafiği elde edilmiş ve her bir içerik tabanlı kümede, tepki stili tabanlı kümelerden öğrencilerin orantılı bir şekilde bulunduğu görülmüştür.

İkinci araştırma sorusu kapsamında, PISA 2015'e katılan ve tutum-başarı paradoksunun görüldüğü ülkelerde K-ortalamlar kümeleme yöntemi uygulanmış ve elde edilen sonuçlar incelenmiştir. K-ortalamlar kümeleme yöntemine göre, fenden hoşlanma ölçeği için uygun küme sayısı dört olarak belirlenmiştir. Elde edilen küme yapılarına bakıldığında; birinci kümede öğrencilerin çoğunlukla 3. tepki kategorisini, ikinci kümedeki öğrencilerin çoğunlukla 4. tepki kategorisini, üçüncü kümede yer alan öğrencilerin çoğunlukla 1. tepki kategorisini tercih ettikleri ve dördüncü kümede yer alan öğrencilerin ise çoğunlukla 2. tepki kategorisini tercih ettikleri görülmüştür. K-ortalamlar kümeleme yöntemine göre, başarı motivasyonu ölçeği için uygun küme sayısı iki olarak belirlenmiştir. Elde edilen küme yapılarında, birinci kümede öğrencilerin çoğunlukla 2. ve 3. tepki kategorisini, ikinci

kümedeki öğrencilerin ise çoğunlukla 3. ve 4. tepki kategorisini tercih ettikleri görülmüştür. K-ortalamlar yöntemine göre, fen öz yeterliği ölçeği için uygun küme sayısı dört olarak belirlenmiştir. Elde edilen küme yapılarına bakıldığında; birinci kümede öğrencilerin çoğunlukla 4. tepki kategorisini, ikinci kümedeki öğrencilerin çoğunlukla 1. tepki kategorisini, üçüncü kümede öğrencilerin çoğunlukla 3. tepki kategorisini ve dördüncü kümede ise öğrencilerin çoğunlukla 2. tepki kategorisini tercih ettikleri görülmüştür.

İkinci araştırma sorusu kapsamında elde edilen sonuçlar birlikte değerlendirildiğinde fenden hoşlanma ve fen öz yeterliği ölçekleri için kümelerdeki öğrencilerin çoğunlukla belirli bir kategoriyi tercih etmiş olmaları dikkat çekmektedir. Başarı motivasyonu ölçeğinde ise kümelerdeki öğrencilerin tek bir kategoriden ziyade çoğunlukla iki kategoriyi seçmiş oldukları görülmüştür. Buna göre K-ortalamlar kümeleme yönteminin fenden hoşlanma ve fen öz yeterliği ölçeklerinde, bireyleri ölçülen özelliğe göre değil tepki stillerine göre kümelediği söylenebilir. Aksu, Güzeller ve Eser, (2017) tarafından yapılan çalışmada PISA 2012 öğrenci anketinde yer alan öz yeterlik algısı, derslere karşı ilgi ve tutum duyuşsal özelliklerine göre OECD üyesi ve diğer katılımcı ülkelerin tamamının kümelenebilirliği yapılmıştır. Araştırmada kayıp verilerin silinmesinden sonra 43 ülke analize dâhil edilmiştir. Çalışmada kümelerin oluşmasında hiyerarşik kümeleme yöntemlerinden tam bağlantı kümeleme yöntemi kullanılmıştır. Analiz sonucunda elde edilen küme sayısının geçerliliğine kanıt sağlamak amacıyla K-ortalamlar kümeleme analizi ve ayırma analizi yöntemlerinden yararlanılmıştır. Çalışma sonucunda öz yeterlik ölçeğine göre 8, ilgi ölçeğine göre 7 ve tutum ölçeğine göre ise 6 farklı küme elde edilmiştir. Acar (2012) tarafından yapılan çalışmada ise PISA 2009'a katılan 65 ülkeden toplam 475.460 öğrencinin kümelenebilirliği gerçekleştirilmiştir. Kümeleme analizi sonuçlarına göre; 1. kümede 13 ülke yer alırken, 2. kümede 30 ülke, 3. kümede 10 ülke ve 4. kümede ise 12 ülkenin yer aldığı görülmüştür. Kümeleme analizinin sonuçlarının geçerliliğine kanıt sağlamak amacıyla ayırma analizi yapılmış ve %96,9 oranında doğru sınıflandırma yüzdesi elde edilmiştir.

PISA verisi üzerinden gerçekleştirilen kümeleme analizi çalışmalarından bir diğeri ise Linnakyla ve Malin (2008) tarafından yapılmıştır. Linnakyla ve Malin (2008) tarafından yapılan çalışmada, PISA 2003 uygulamasına katılan Finlandiyalı öğrencilerin profilleri belirlenmeye çalışılmıştır. Çalışma sonuçlarına göre Finlandiyalı öğrencilerin bilişsel başarıları üst düzeyde yer alırken, öğrencilerin okullarındaki sorumluluk profilleri bilişsel ve duyuşsal öğrenmeleri ile ilişkili bulunmuştur. Kjærnsli ve Lie (2011) tarafından yapılan çalışmada, PISA 2006 veri seti kullanılarak 15 yaşındaki öğrencilerin fen ile ilgili

gelecekteki bir kariyeri değerlendirme eğilimleri incelenmiştir. Araştırmada öğrencilerin fen ile ilgili alanlarda gelecekteki eğitim ve kariyerlerine katılma düşüncelerini ölçen dördümlük likert tipinde maddelerden oluşan duyuşsal bir yapı olan "gelecek fen yönelimi" adlı yapı ele alınmıştır. Kültürel yanlı tahminler sağlayan likert ölçeklerine ilişkin kanıtlar elde edildiği için çalışmada basit ülke ortalamalarının karşılaştırılmasının ötesine geçilmeye çalışılmıştır. Yapılan bir dizi regresyon ve korelasyon analizi ile katılımcı ülkelerin her birinde “gelecek fen yönelimi” yapısının varyansının diğer PISA öğrenci verileri tarafından ne kadar iyi açıklanabileceği araştırılmıştır. Fenle ilgili işleri, "ılıman" ve "zor" şeklinde ayırarak, her bir kategorideki ülke yüzdeleri hesaplanmış ve karşılaştırılmıştır. Ayrıca cinsiyet farklılıkları tartışılmış ve ilginç uluslararası modeller tespit edilmiştir. Araştırma sonuçları sadece tek tek ülkeler için değil, aynı zamanda ülke grupları için de rapor edilmiştir. Kümeleme analizi yapılarak ülkeler yapılarına göre sınıflandırılmıştır. Ülkelerin ortaya çıkan kümeleme yapısı, ülkeler arasındaki benzerlikler ve farklılıklar ile hem fen eğitiminde hem de daha genel olarak ülke kümelenmesinin arkasındaki faktörlere ilişkin literatüre katkı sağlamaktadır.

Olsen ve Lie (2011) tarafından yapılan araştırmada, PISA 2006’ya katılan ülkelerdeki öğrencilerin fene olan ilgisi kümeleme analizi ile incelenmiştir. Araştırmada likert tipi maddelerin kullanımıyla ülkeler arasında geçerli ölçümlerin elde edilememesi düşünülmüş ve bu nedenle öğrencilerin likert ölçeğinden oluşan maddelerdeki ifadelerle yanıt vermelerinin istendiği birkaç maddeden alınan bilgilerin birleştirilmesi yoluna gidilmiştir. Kültürel olarak yanlı ülke puanları sağlayan likert ölçeklerine kanıt sağlandığı için, çalışmada ilgi alanlarının göreceli profillerinin ülkeler arasında anlamlı bir şekilde analiz edilebileceği gösterilmiştir. Bu nedenle, madde düzeyinde ülkeye ve maddeye özgü puanlardan oluşturulmuş fene ilişkin ulusal ilgi profilleri geliştirilmiştir. Daha sonra, bu göreceli ilgi profilleri, madde bazında fene yönelik ilgi örüntülerine sahip farklı ülke gruplarının tanımlanmasını sağlayan bir kümeleme analizinde girdi değişkenleri olarak kullanılmıştır. Araştırmada, hiyerarşik kümeleme yöntemlerinden ortalama bağlantı kümeleme tekniği kullanılmıştır. Kümeleme analizi sonuçlarına göre iki büyük küme elde edilmiştir. Kümelerin birinde Avrupa/Batı ülkeleri bulunurken, diğer kümede ise sadece birkaç istisna dışında Avrupa dışındaki ülkeler yer almıştır. Kısmen dil ve kısmen yerleştirme ile tanımlanan bir dizi anlamlı ülke kümesi, iki ana kümenin her birinde tanımlanmıştır. Kümelerdeki ülkelerin karakteristik özelliklerinin daha ayrıntılı bir şekilde anlaşılmasını sağlamak amacıyla, maddeler hakkında açıklayıcı bilgilerde analize dâhil edilmiştir. Çalışmadan elde edilen en dikkat çekici sonuç, Avrupa/Batı ülkelerindeki

fiziksel/teknolojik sistemlerle ilgili maddelerin Avrupa dışındaki ülkelere kıyasla daha çok tercih edilmesidir. Buna karşın, Avrupa dışındaki ülkeler arasında yaşam ve sağlık ile ilgili maddelerin Avrupa ülkelerine kıyasla daha çok tercih edildiği görülmüştür. Qiao ve Jiao (2018), PISA 2012 Amerika Birleşik Devletleri örnekleminde elde ettikleri 426 kişilik çalışma grubu verilerini kullanarak K-ortalamlar kümeleme yöntemi ile Kohonen'in Öz Örgütlemeli Haritalama yöntemini karşılaştırmışlardır. Analizlerde 11 problem çözme maddesi kümeleme analizinde girdi değişkeni olarak kullanılmıştır. Çalışmadan elde edilen sonuçlarda ,84 kappa istatistiği ile K-Ortalamlar kümeleme yöntemi ile ideal küme sayısı beş olarak bulunurken ,96 kappa istatistiği ile Kohonen'in Öz Örgütlemeli Haritalama yöntemi ile elde edilen ideal küme sayısı dokuz olarak belirlenmiştir. Ayrıca her iki kümeleme yöntemine ilişkin doğru sınıflandırma yüzdelерinin tatmin edici düzeyde ve birbirine yakın olduğu bulunmuştur.

Üçüncü araştırma sorusu kapsamında TSYDK yöntemi ile elde edilen sonuçlar ile K-ortalamlar kümeleme yönteminden elde edilen sonuçlar karşılaştırılmıştır. Karşılaştırma sonuçlarına göre, ölçülen özellik açısından TSYDK yönteminin K-ortalamlar kümeleme yöntemine göre bireyleri kümeleme performansının daha başarılı olduğu ve geçerliği daha yüksek sonuçların elde edildiği görülmüştür. TSYDK yöntemi verileri tepki stili yanlılığından arındırarak içerik tabanlı kümelerin elde edilmesini sağlamıştır. Buna karşılık K-ortalamlar kümeleme yöntemi bireyleri daha çok belirli tepki kategorileri tercih etmelerine göre sınıflandırmış ve tepki stili tabanlı kümeler ayırmıştır. Bu durumda K-ortalamlar kümeleme yöntemine göre elde edilen kümeleme sonuçları yorumlanırken, içeriğe dayalı kümelemeden daha çok tepki stili temelli farklılıklar düşünülmelidir. Üçüncü araştırma sorusu kapsamında elde edilen sonuçlar Takagishi vd. (2019) tarafından yapılan araştırma sonuçları ile benzerlik göstermektedir.

Takagishi vd. (2019) tarafından yapılan çalışmada, 2008 yılında Doğu Asya Sosyal Anketi (East Asian Social Survey-EASS-) verileri kullanılarak TSYDK yönteminin kullanımı gösterilmiş ve TSYDK yönteminin etkililiği SÇÖ ve K-ortalamlar kümeleme yöntemi ile de karşılaştırılmıştır. Takagishi vd. (2019) tarafından yapılan çalışmada TSYDK yöntemi ile tepki stili yanlılığı düzeltilmiş ve dört ülkeden elde edilen veriler içerik tabanlı yedi kümeye ayrılırken, tepki stili tabanlı altı küme elde edilmiştir. TSYDK yöntemi ile elde edilen içerik tabanlı ve tepki stili tabanlı kümeleme çözümlerinin çapraz tablosuna bakıldığında; içerik tabanlı ve tepki stili tabanlı kümelerde önemli bir çakışma olmadığı görülmüştür. Her içerik tabanlı kümede tüm tepki stili tabanlı kümelerden yanıt verenler

bulunmuştur. SÇÖ yöntemine göre ise içerik tabanlı dört küme elde edilirken, tepki stili tabanlı üç küme elde edilmiştir. SÇÖ yönteminden elde edilen içerik tabanlı kümeleme ile tepki stili tabanlı kümeleme çözümlerinin mozaik grafiğine de bakılmıştır. Mozaik grafikten elde edilen sonuçlarda içerik tabanlı kümeleme ile tepki stili tabanlı kümelemenin çakıştığı görülmüştür. Buna göre SÇÖ yöntemine göre katılımcıların içerik tabanlı kümelenmesinde tepki stili yanlılığının bulunduğu söylenebilir. Böylece, SÇÖ yönteminin katılımcıları ölçülen özelliğe göre kümelemede TSYDK yöntemi kadar başarılı olmadığı sonucu ortaya çıkmıştır. K-ortalamalar kümeleme yöntemine göre ise üç küme elde edilmiş ve K-ortalamalar ile elde edilen sonuçlar tepki stili tabanlı kümelerden ayırt edilmesi zor içerik tabanlı kümelere yol açmıştır. Mevcut araştırmada da TSYDK yönteminin öğrencileri kümelemede K-ortalamalar kümeleme yönteminden daha başarılı olduğu sonucu ortaya çıkmıştı. Takagishi vd. (2019) tarafından yapılan araştırmada aynı coğrafi bölgede bulunan dört ülkedeki (Çin, Japonya, Kore ve Tayvan) katılımcılar kümelenirken, bu araştırmada ise farklı coğrafi bölgelerdeki toplam 21 ülkeden gelen öğrenciler kümelenmiştir. Ülke sayısının fazla olması ve farklı coğrafi bölgelerdeki ve kültürdeki ülkelerin bu çalışmada yer almasına rağmen elde edilen sonuçların, Takagishi vd. (2019) tarafından yapılan araştırma sonuçları ile benzerlik göstermesi TSYDK yönteminin tepki stili yanlılığının düzeltilmesi ve içerik tabanlı kümelemedeki etkililiğine ek kanıtlar sağlamıştır.

Araştırma sonuçlarına göre, tepki stili yanlılığının düzeltilmediği puanlar kullanılarak yapılan K-ortalamalar kümeleme analizi sonuçları, öğrencilerin tepki stili eğilimlerine göre kümelere ayrıldığını göstermiştir. Buna karşılık TSYDK yöntemi ile tepki stili yanlılığının düzeltilmesinin kümeleme analizinin sonuçlarının geçerliğini arttırdığı görülmüştür. Buna göre, tepki stili yanlılığının düzeltilmesinin kümeleme analizi gibi istatistiksel yöntemler üzerinde olumlu etkiler meydana getirdiği görülmüştür. Liu, Harbaugh, Harring ve Hancock (2017) tarafından DFA aracılığıyla UTS' nin ve UTS olmama durumunun ölçme değişmezliği üzerindeki etkileri, Monte Carlo simülasyon çalışmasıyla sistematik olarak incelenmiştir. UTS' ye sahip olma ve UTS' ye sahip olmama durumları için orijinal ve yanlı veriler oluşturulmuş ve farklı ölçme değişmezliği sınırlandırmaları ile ÇGDFA ile analiz edilmiştir. Araştırmadan elde edilen sonuçlar, tepki stilinin ölçme değişmezliği üzerinde yanlı sonuçlara neden olabileceğini göstermiştir. Daha spesifik olarak, UTS hem model uyumu hem de parametre tahminleri üzerinde önemli ölçüde olumsuz etkiye sahip olmuştur. Bu da gruplar arasında ölçme değişmezliği eksikliğinin ölçülen faktörlerden değil, tepki stillerinden kaynaklanmış olabileceğini düşündürmüştür. Bu nedenle hem mevcut

araştırmadan elde edilen sonuçlar hem de Liu vd. (2017) yapmış olduğu çalışma sonuçları tepki stillerinin istatistiksel yöntemlerden elde edilen sonuçların geçerliği üzerinde önemli bir tehdit olduğunu kanıtlamıştır.

Dördüncü araştırma sorusunun a seçeneği kapsamında TSYDK yöntemi ile düzeltilmiş ve düzeltilmemiş ortalama tutum puanları arasındaki farkların manidar olup olmadığı incelenmiştir. Bu doğrultuda araştırma kapsamında kullanılan fenden hoşlanma, başarı motivasyonu ve fen öz yeterlikleri ölçekleri için her bir ülkenin düzeltilmiş ve düzeltilmemiş ortalama puanları arasında ilişkili örneklem t testi yapılmıştır. Fenden hoşlanma, başarı motivasyonu ve fen öz yeterliği ölçekleri için düzeltilmiş ve düzeltilmemiş ortalama puanlar arasındaki ilişkili örneklem t testi sonuçları bütüncül bir şekilde değerlendirildiğinde araştırmaya dâhil edilen ülkelerin çoğunda düzeltilmemiş ve düzeltilmiş ortalama puanlar arasında anlamlı farklılık bulunmuştur. Bu durumda tepki stillerinin öğrencilerin fene yönelik tutumlarını etkilediği yorumu yapılabilir. Ayrıca araştırma kapsamına dâhil edilen ülkelerin düzeltilmemiş ortalama tutum puanları, düzeltilmiş ortalama tutum puanlarına kıyasla daha düşük olduğu ülkelerin genel olarak fen başarısı yüksek ve fen başarısı orta düzeyde olan ülkeler ve fene yönelik tutumları olumsuz olduğu görülmüştür. Buna karşılık, düzeltilmemiş ortalama tutum puanları düzeltilmiş ortalama tutum puanlarına kıyasla daha yüksek olduğu ülkeler ise genel olarak fen başarısı düşük ve fene yönelik tutumu olumlu olan ülkeler olduğu görülmüştür. Elde edilen bu sonuçlara göre tepki stillerinin fen başarısı yüksek olan ülkelerdeki öğrencilerin fene yönelik tutumlarını olduğundan daha olumsuz şekilde kestirilmesine yol açtığı söylenebilir. Ayrıca, tepki stillerinin fen başarısı düşük olan ülkelerdeki öğrencilerin fene yönelik tutumlarını olduğundan daha olumlu şekilde kestirilmesine yol açtığı söylenebilir.

Mevcut araştırmada dördüncü araştırma sorusunun a seçeneğinden elde edilen sonuçlar, Van de Gaer ve Adams (2010)'un PISA 2003 uygulamasına katılan ülkelerdeki öğrencilerin matematiğe yönelik ilgilerinde UTS' nin etkisini inceledikleri çalışma sonuçları ile benzerlik göstermektedir. Araştırmacıların yaptığı çalışmanın sonuçlarına göre, matematik başarısı yüksek olan Kore, Hong-Kong, Japonya ve Finlandiya gibi ülkelerin UTS' nin etkisine göre matematiğe yönelik ilgilerine ait düzeltilmemiş ortalama puanlarının düzeltilmiş ortalama puanlarından daha düşük olduğu ve matematik başarısı düşük olan Endonezya, Tunus ve Brezilya gibi ülkelerin ise matematiğe yönelik ilgilerine ilişkin düzeltilmemiş ortalama puanlarının düzeltilmiş ortalama puanlarından daha yüksek olduğu bulunmuştur.

Fenden hoşlanma ölçeği için bazı ülkelerde düzeltilmemiş ortalama puanların düzeltilmiş ortalama puanlardan düşük olduğu, bazı ülkelerde ise tam tersine düzeltilmemiş ortalama puanların düzeltilmiş ortalama puanlardan daha yüksek olduğu görülmüştür. Elde edilen bu sonuç, İlgün Dibek (2018) tarafından TIMSS 2015 verileri kullanılarak yapılan çalışma sonuçları ile benzerlik göstermektedir. İlgün Dibek (2018) yaptığı çalışmada, TIMSS 2015'deki matematik öğrenmeyi sevme değişkeninde UTS' nin etkisinin düzeltilmesi amacıyla ÇBMTK'ı kullanmıştır. ÇBMTK kullanılarak UTS yanlılığının düzeltilmesi ile düzeltilmiş puanlar elde edilmiştir. Ayrıca çalışmada ülkelerin düzeltilmemiş ve düzeltilmiş ortalama puanları arasında anlamlı farklılık olup olmadığı ilişkili örneklem t testi ile analiz edilmiştir. Yapılan analizler sonucunda bazı ülkelerin düzeltilmemiş ortalama tutum puanlarının düzeltilmiş ortalama puanlarından anlamlı derecede daha yüksek olduğu görülürken, bazı ülkelerde ise düzeltilmemiş ortalama tutum puanları düzeltilmiş ortalama tutum puanlarından anlamlı derecede daha düşük bulunmuştur. Çalışmadan elde edilen sonuçlara göre; Singapur, Malta, Hong-Kong ve Avustralya'nın matematiği öğrenmeyi sevme boyutuna ilişkin UTS yanlılığının bulunduğu düzeltilmemiş ortalama tutum puanlarının düzeltilmiş tutum puanlarından daha düşük olduğu bulunmuştur. Bu ülkeler için UTS yanlılığının etkisinin kaldırıldığı düzeltilmiş tutum puanları artmıştır. Diğer taraftan Türkiye, Mısır ve Kuveyt'in UTS' nin etkisi dikkate alındığında matematiği öğrenmeyi sevme boyutuna ilişkin ortalama tutum puanlarının azaldığı görülmüştür. Bu durumda, UTS' nin bu ülkelerdeki öğrencilerin matematik öğrenmeyi sevme puanlarını olduğundan daha yüksek kestirilmesinde yol açan faktörlerden biri olduğu söylenebilir. Düzeltilmemiş ve düzeltilmiş ortalama tutum puanları arasındaki manidar farklılıklara ilişkin etki büyüklüğü değerleri (d) incelendiğinde etki büyüklüğü değeri en yüksek olan ülkenin Mısır olduğu görülmüştür. Ayrıca, Mısır'a ait etki büyüklüğü değeri orta düzeyde bulunurken, diğer ülkelerin etki büyüklüğü değerleri küçük düzeyde bulunmuştur.

İlgün Dibek (2018) yaptığı çalışmada, TIMSS 2015'teki öğrenci anketinde yer alan matematikte kendine güven ve matematiğe değer verme değişkenleri için ÇBMTK' nı kullanarak UTS yanlılığını incelemiştir. Çalışmadan elde edilen sonuçlara göre; Singapur, Şili, Norveç, Malta, İtalya, İsveç, Hong-Kong ve Avustralya'nın matematikte kendine güven değişkenine ilişkin UTS yanlılığının dikkate alınmadığı durumda elde edilen düzeltilmemiş ortalama tutum puanlarının düzeltilmiş ortalama tutum puanlarına kıyasla daha düşük olduğu görülmüştür. Bu ülkeler için UTS yanlılığının kaldırıldığı düzeltilmiş ortalama tutum puanları artış göstermiştir. Bunun yanında Mısır, Kore, Japonya ve Çin-Tayvan'ın

düzeltilmemiş ortalama tutum puanları, düzeltilmiş ortalama puanlarından daha yüksek bulunmuştur. Bu ülkeler için UTS' nin öğrencilerin matematikte kendine güven değişkenine ilişkin ortalama tutum puanları olduğundan daha yüksek kestirilmesine yol açtığı söylenebilir. Matematiğe değer verme değişkeni için elde edilen sonuçlara göre; Türkiye, Sudi Arabistan, Şili, Mısır ve Kuveyt'in matematiğe değer verme değişkenine ilişkin UTS' nin etkisine göre düzeltilmemiş ortalama tutum puanlarının düzeltilmiş ortalama tutum puanlarından daha yüksek olduğu bulunmuştur. Bu doğrultuda bu ülkeler için UTS, matematiğe değer verme boyutundaki puanları artırıcı yönde rol oynamıştır. Bunun yanında Singapur, Norveç, Kore, Japonya, İsveç, Hong-Kong, Çin-Tayvan ve Avustralya'nın düzeltilmemiş ortalama tutum puanları düzeltilmiş ortalama tutum puanlarından daha düşük bulunmuştur. Bu ülkeler için UTS, öğrencilerin tutum puanlarını azaltıcı yönde rol oynamıştır. Düzeltilmemiş ortalama puanlar ile düzeltilmiş ortalama puanlar arasında manidar farklılık bulunan ülkeler için hesaplanan etki büyüklüğü değerlerinin çoğunlukla küçük ve orta büyüklükte olduğu, düzeltilmemiş ve düzeltilmiş tutum puanları arasında en büyük fark olan ülkenin ise Mısır olduğu görülmüştür.

Dördüncü araştırma sorusunun b seçeneği kapsamında düzeltilmemiş ortalama puanlar ile düzeltilmiş ortalama puanlar kullanılarak fenden hoşlanma, başarı motivasyonu ve fen öz yeterliği ölçekleri ile fen başarısı arasındaki ülke içi ve ülkeler arası korelasyonlar hesaplanmıştır. Buna göre ilk olarak ülke içi fene yönelik tutum ile fen başarısı arasındaki korelasyonlar hesaplanmıştır. Fenden hoşlanma ölçeği için araştırmada yer alan toplam 21 ülkede ülke içi korelasyon değerleri incelendiğinde; düzeltilmemiş ve düzeltilmiş fenden hoşlanma puanları ile fen başarıları arasındaki korelasyon değerleri pozitif bulunmuştur. Başarı motivasyonu ölçeğinde sadece Brezilya'nın düzeltilmiş puanları ile fen başarıları arasındaki korelasyon değerleri negatif olurken, diğer ülkelerin düzeltilmemiş ve düzeltilmiş başarı motivasyonu puanları ile fen başarıları arasındaki korelasyon değerleri pozitif bulunmuştur. Fen öz yeterliği ölçeğinde ise sadece Peru'nun düzeltilmiş puanları ile fen başarısındaki korelasyon negatif, diğer bütün ülkelerin düzeltilmemiş ve düzeltilmiş fen öz yeterliği puanları ile fen başarıları arasındaki korelasyon değerleri pozitif bulunmuştur.

Dördüncü araştırma sorusunun b seçeneği kapsamında; fenden hoşlanma, başarı motivasyonu ve fen öz yeterliği ölçekleri ile fen başarıları arasındaki düzeltilmemiş ortalama puanlar ile düzeltilmiş ortalama puanlar aracılığıyla ülkeler arası korelasyonlar hesaplanmıştır. Her üç ölçek içinde düzeltilmiş puanlar ile elde edilen ülkeler arası tutum-başarı korelasyon değerleri negatif ve düzeltilmemiş puanlar aracılığıyla elde edilen

korelasyon değerlerinden eşit ya da daha düşük bulunmuştur. Ancak düzeltilmiş puanlar aracılığıyla elde edilen korelasyon değerleri pozitif dönüşmemiştir. Bu durum, TSYDK yönteminin tepki stili yanlılığının düzeltilmesindeki etkiliğinin yeterli olmadığına işaret etmektedir. Ayrıca düzeltilmemiş ve düzeltilmiş puanlar aracılığıyla elde edilen tutum-başarı korelasyon değerleri birbirine yakındır. Araştırma kapsamında, her üç ölçek içinde düzeltilmemiş ortalama puanlar ve düzeltilmiş ortalama puanlar aracılığıyla elde edilen ülke içi ve ülkeler arası fene yönelik tutum ile fen başarısı arasındaki korelasyon değerlerinin birbirine yakın olması, tepki stil yanlılığının düzeltilmesinin tutum-başarı paradoksunu açıklamada yetersiz kaldığını göstermektedir.

Araştırmaya dahil edilen ülkelerde, fenden hoşlanma, başarı motivasyonu ve fen öz yeterliği ölçekleri için gözlenen puanlar ile hesaplanan ülke içi tutum-başarı korelasyon değerlerinin pozitif, ülkeler arası korelasyon değerlerinin ise negatif olması tutum-başarı paradoksuna işaret etmekteydi. Araştırma kapsamında bu durumu ortadan kaldırmak için, TSYDK yöntemi kullanılmıştır. TSYDK yöntemi ile tepki stili yanlılığı ortadan kaldırılarak düzeltilmiş puanlar elde edilmiştir. Düzeltilen puanlar üzerinden birinci araştırma sorusu kapsamında TSYDK yöntemi ile kümeleme analizi yapılmış ve öğrencilerin içerik tabanlı kümelenmesinin gerçekleştiği görülmüştür. Ayrıca TSYDK yöntemi ile K-ortalamlar kümeleme yöntemi karşılaştırılmış ve TSYDK yönteminin K-ortalamlar yöntemine göre daha geçerli sonuçlar verdiği ve kümeleme performansının daha başarılı olduğu gözlemlenmiştir. Diğer taraftan her üç ölçek içinde araştırmaya dâhil edilen ülkeler için düzeltilmemiş ve düzeltilmiş tutum puanları arasında ilişkili örneklem t testi yapılmıştır. Ülkelerin büyük çoğunluğunda düzeltilmemiş ve düzeltilmiş ortalama puanlar arasında anlamlı farklılık bulunmuştur. Elde edilen bu sonuçlar birlikte değerlendirildiğinde, TSYDK yönteminin tepki stili yanlılığının ortadan kaldırılmasında ve verilerin içerik tabanlı kümelenmesinde kullanılabilecek bir yöntem olduğu düşünülmektedir. Ayrıca, tutum-başarı paradoksuna yol açan sebeplerden birinin de tepki stili yanlılığın olduğu ortaya çıkmıştır.

Çalışmanın dördüncü araştırma sorusu kapsamında fenden hoşlanma, başarı motivasyonu ve fen öz yeterliği ölçeklerinin düzeltilmemiş ve düzeltilmiş tutum puanları ile fen başarısı arasındaki ülke içi ve ülkeler arası korelasyonlar hesaplanmıştır. Düzeltilmemiş ve düzeltilmiş puanlar aracılığıyla elde edilen ülke içi korelasyon değerlerinin birbirine yakın olduğu görülmüştür. Düzeltilmemiş puanlar aracılığıyla elde edilen ülkeler arası korelasyon değerleri negatif bulunurken, düzeltilmiş puanlar aracılığıyla elde edilen korelasyon değerleri daha düşük bulunmuş, ancak negatiftikten kurtulamamıştır. Elde edilen bu

sonuçlar, Lu ve Bolt (2015) tarafından yapılan çalışma sonuçları ile paralellik göstermektedir. Lu ve Bolt (2015) tarafından yapılan çalışmada, PISA 2006 uygulamasına katılan öğrencilerin fene yönelik tutumları ile fen başarıları arasındaki ilişkide UTS' nin etkisi incelenmiştir. Araştırmada UTS' nin etkisini incelemek amacıyla ÇBMTK kullanılmıştır. PISA 2006' da fene yönelik tutum anketlerindeki maddeler 1=Kesinlikle Katılıyorum, 2= Katılıyorum, 3= Katılmıyorum ve 4= Kesinlikle Katılmıyorum şeklinde puanlanmaktadır. Bu nedenle, anketlerden alınacak yüksek puan fene yönelik olumsuz tutumu göstermektedir. Ayrıca öğrencilerin fen başarısını temsil etmek amacıyla beş olası değerlerin ortalaması alınmıştır. Araştırmada tepki stili yanlılığının bulunduğu ve tepki stili yanlılığının düzeltildiği veriler üzerinden ülke içi ve ülkeler arası korelasyonlar hesaplanmıştır. Tepki stili yanlılığı bulunan veriler üzerinden yapılan ülke içi tutum-başarı korelasyon değerlerinin negatif olduğu ve ülkeler arası korelasyon değerlerinin ise pozitif olduğu görülmüştür. Diğer taraftan, ÇBMTK ile UTS yanlılığından arındırılmış düzeltilen veriler üzerinden hesaplanan ülke içi korelasyon değerlerinin de negatif ve ülkeler arası korelasyon değerlerinin ise pozitif olduğu görülmüştür. Bunun yanında, düzeltilmemiş veriler aracılığıyla elde edilen ülke içi ve ülkeler arası korelasyon değerleri ile ÇBMTK aracılığıyla düzeltilen veriler üzerinden hesaplanan ülke içi ve ülkeler arası korelasyon değerlerinin birbirine yakın olduğu görülmüştür. Elde edilen bu sonuçlarda UTS' nin tutum-başarı paradoksunu açıklama yetersiz kaldığını göstermiştir. Lu ve Bolt (2015) yaptıkları bu çalışma ile tepki stili yanlılığının düzeltilmesine rağmen, ülkeler arası korelasyon sonuçlarının büyük ölçüde neden değişmediğine dair bazı ipuçları elde etmişlerdir. Buna göre ilk olarak, UTS' deki ülke düzeyindeki farklılıklar tespit edilebilir olsa da, UTS' deki ülke içi değişkenliğe kıyasla nispeten küçük olmaktadır. Ayrıca, UTS ile ilgili sınıf içi korelasyon değeri yaklaşık olarak ,31 olarak bulunmuştur. Bu da her iki yönde de uç tepki eğiliminde olan ülkelerde bile makul miktarda ülke içi heterojenliğin kaldığını göstermektedir. Bunun yanında, ülke içinde başarı ile UTS arasında beklenen yönde (yani, daha yüksek başarı daha az UTS) bir ilişki bulunurken, ülke düzeyinde başarı ve UTS arasındaki korelasyon oldukça zayıf (-.07) bulunmuştur. Bu sonuçta, tutum ile başarı arasındaki korelasyonda UTS ile ilgili ülke etkilerinin kontrol edilmesinin anlamlı etkilere yol açma olasılığının düşük olduğunu göstermektedir.

Bolt, Lu ve Kim, (2014) tarafından yapılan ve ÇBMTK modelinin kullanıldığı kültürler arası bir çalışmada ise farklı ülkelerin geleneksel tepki stili türlerine uymayan şekillerde farklı derecelendirme ölçek kullanımı yapabildikleri gösterilmiştir. Bu nedenle, tutum-başarı

paradoksunun olası kaynağı olarak tepki stilindeki kültürler arası farklılıkların ek metodolojik çalışmalarla da incelenmesi gerektiği düşünülmektedir. Örneğin tepki stili eğilimlerinin daha iyi ölçülmesi için alternatif tasarım yöntemleri de kullanılabilir. Buna göre, sabitleme vinyetlerinin (anchoring vignettes) kullanımı (King, Murray, Salomon, & Tandon, 2004), farklı türlerdeki tepki stillerinin tanımlanmasında ve kontrol edilmesinde faydalı olabilir.

Buckley (2009) yaptığı çalışmada, PISA 2006 öğrenci anketindeki fene verilen değer ve fenden hoşlanma ölçekleri ile fen başarısı arasındaki ilişkide tepki stili veya ölçek kullanımındaki kültürler arası farklılıkların kapsamını, şeklini ve sonuçlarını araştırmıştır. Bu doğrultuda geleneksel yaklaşımlardan ad hoc kestirim yöntemi ile daha sistematik bir Bayes hiyerarşik yaklaşımı kullanılarak uluslararası tepki stili farklılıklarına kanıt gösterilmiştir. Yapılan çalışmada tepki stili yanlılığının düzeltilmesinden sonra, düzeltilmiş tutum ölçeklerine dayanan grafiklerde ülkelerin başarıları ve fene yönelik tutumları arasında doğrusal bir ilişkiden ziyade, daha karmaşık doğrusal olmayan ilişki olduğu görülmüştür. Her iki ölçek için de tepki stili, tutum ile başarı arasındaki doğrusal ilişkiyi değiştiriyor gibi görünmüştür. Doğrusal olmayan Bayes hiyerarşik modele göre; fene verilen değer ölçeğinde; hem olumlu hem de olumsuz tutuma sahip ülkelerin ortalama olarak benzer yüksek başarı düzeylerine sahip oldukları görülmüştür. Bu nedenle tepki stili yanlılığının dikkate alınmadığı durumda elde edilen doğrusal model sonuçlarındaki yanlışlığın, bireysel ve kültürel ilişkiler arasında gözlenen farklılıklar için yeterli bir açıklama olmayabileceği düşünülmüştür. Yani, bireysel düzeyde başarı ve tutumlar arasında gözlemlenen pozitif ilişkinin ülke düzeyinde tersine döndüğünü söylemek artık doğru olmadığı ifade edilmiştir. Her iki seviyede de daha karmaşık doğrusal olmayan bir ilişkiden söz edilmektedir. Bu nedenle doğrusal ilişkilerin elde edildiği korelasyon gibi istatistiksel yöntemlerle tepki stili yanlılığın düzeltilmesinin etkililiğini incelemenin zor olduğu söylenebilir.

Dördüncü araştırma sorusunun b seçeneği kapsamında düzeltilmiş puanlar aracılığıyla elde edilen fene yönelik tutum ile fen başarısı arasındaki ülkeler arası korelasyon değerlerinin negatif bulunmasının sebeplerinden biri ülkelerin heterojen bir yapı oluşturması gösterilebilir. Araştırmaya katılan ülkelerin farklı coğrafi bölgelerden gelmesi ve bu ülkelerin aynı veri seti içerisinde yer alması korelasyon sonuçlarının geçerliğini olumsuz şekilde etkilemiş olabilir. Aynı coğrafi bölgede yer alan ülkelerin benzer tarihe ve kültüre sahip olması tepki stili eğilimlerini de etkilemektedir. Literatür incelendiğinde bu durumu kanıtlar nitelikte çalışma sonuçlarının mevcut olduğu görülmektedir (Jia & van de Vijver,

2016). Ülkelerin geçmişten günümüze kadar sahip oldukları siyasi, tarihsel ve kültürel zenginliklerinde sayısız farklılıklar bulunmaktadır. Bu farklılıklar daha sonra ölçek kullanım tercihlerini etkiler ve likert ölçekli öz raporların karşılaştırılabilirliği ve geçerliği üzerinde daha fazla etkiye sahip olabilir. Tipik olarak PISA çalışmalarında ortalama başarının altında puanlara sahip olan Latin Amerika ülkeleri, öz bildirim ölçeklerinde ortalamanın üzerinde puanlara sahip olma eğilimindedir (Jia & van de Vijver, 2016). Chen vd. (1995) Orta ve Güney Amerikalı öğrencilerin Doğu Asyalılara göre UTS kullanmaya daha yatkın olduklarını bulmuşlardır. Hatta aynı ülkede yaşayan farklı etnik kökenden gelen bireyler arasında bile tepki stili farklılıkları bulunabilmektedir. Yapılan çalışmalarda A.B.D' de hem Hispanikler (Clarke, 2001; Hui & Triandis, 1989; Johnson vd, 1997; Marin, Gamba, & Marin, 1992; Ross & Mirowsky, 1984) hem de Afrika Amerikalıların (Bachman & O'Malley, 1984; Clarke, 2001; Johnson vd, 1997) Avrupalı Amerikalılara kıyasla özellikle tepki ölçeğinin olumlu ucuna doğru UTS ve KTS eğilimlerinin daha yüksek olduklarını göstermiştir.

Hofstede ve McCrae (2004) ve Van Hemert, van de vijver, Poortinga, ve Georgas (2002), faktör yapılarının yalnızca ülkeler içinde bireysel düzeyde değil, aynı zamanda ülke düzeyinde de kopyalandığını ve bu nedenle ülkeler arasında anlamlı karşılaştırmalar yapılabileceğini göstermişlerdir. Bu doğrultuda, beş büyük kişilik özelliklerinin (nevrotiklik, dışadönüklük, açıklık, anlaşılabilirlik ve vicdanlılık) kültürler arasındaki karşılaştırmasına dayanan çalışmalar bulunmaktadır. Smith (2004), kolektivist kültürlerdeki grup içi uyumun, katılımcıların kabul edici tepkiler vermesine yol açabileceğini iddia etmektedir. Ayrıca, Smith (2004) tepki stili yanlılığının yalnızca ulus düzeyindeki gruplararası ilişkilerin (güç mesafesi, bireysellik / toplulukçuluk ve belirsizlikten kaçınma gibi) tezahürü olarak değil, aynı zamanda bireysel iletişimin ulus düzeyinde bir yansıması olarak da görülebileceğini ifade etmiştir. Araştırmalar, belirsizlikten kaçınma düzeylerinin yüksek olduğu ülkelerdeki anket katılımcılarının belirsizliğe tahammülsüz olma eğiliminde olduğunu ve bu nedenle yanıtlarında orta kategorilerden daha çok uç kategorileri tercih etme eğiliminde olduklarını göstermiştir (Harzing, 2006; He, van de Vijver, Domínguez, & Mui, 2014).

TSYDK, SÇÖ ve ÇBMTK yöntemleri tepki stili yanlılığının düzeltilmesini sağladıkları için diğer istatistiksel yöntemlerden bir adım önde görülebilir. Bu yöntemlerden TSYDK ve SÇÖ Takagishi vd. (2019) tarafından yapılan çalışmada karşılaştırılmış ve TSYDK yönteminin SÇÖ yönteminden daha başarılı olduğu bulunmuştur. Bu yöntemlerin dışında tepki stili yanlılığının düzeltilmesi amacıyla daha önce var olan yöntemlerin karşılaştırıldığı

çalıřmalardan biri He ve Van de Vijver, (2016) tarafından yapılmıřtır. alıřmada, yedi Latin Amerika lkesi, Portekiz ve İspanya' nın PISA 2012 ğrenci anketindeki verileri kullanılmıřtır. Arařtırmaya dâhil edilen lkeler, gl fikirler ifade etme ve kendini geliřtirme konusunda gl bir eėilim gstermektedirler. Arařtırmada tepki stili yanlılıėını dzeltmek amacıyla sırasıyla UTS' nin istatistiksel olarak kontrol edilmesi, ařırı talep (overclaiming) ve sabitleme vinyetleri (anchoring vignettes) yntemleri kullanılarak dokuz lkedeki 39.045 ğrenci arasında ğretmen desteėi leėi kullanım tercihleri dzletilmiřtir. Dzeltme yntemlerinin etkililiėini arařtırmak iin her bir yntem ile elde edilen puan dzeltmelerinin lme deėiřmezlikleri incelenmiřtir. Her  yntemden elde edilen sonular birbirinden farklılık gstermiřtir. Buna gre, tm dzeltme yntemleri, lme deėiřmezliėini iyileřtirmeye yardımcı olmuřtur ancak sabitleme vinyetleri temelli dzeltme, UTS' nin istatistiksel olarak kontrol edilmesi ve ařırı talep dzeltme ile karřılařtırıldıėında skaler deėiřmezliėe ulařmada daha az etkili olmuřtur. UTS' nin istatistiksel olarak kontrol edilmesi ve ařırı talep, İspanya'nın ortalama puanını diėer lkelere gre daha byk lde deėiřtirmiřtir. Bu daha gereki bir model oluřtururken, diėer ltlerle korelasyonlardaki deėiřiklikler olduka sınırlıdır. Tepki stillerinin dzetilmesinde hangi yntemin tercih edileceėine dair kesin bir sonu elde edilememiřtir. nk bu lkelerde hangi yntemin puanların geerliliėini daha fazla artırdıėına dair yeterli kanıt bulunmamaktadır. Ancak, alıřmadan elde edilen nemli sonulardan biri bu dzeltme yntemleri farklı lek kullanım tercihlerini hedeflemektedir. UTS' nin istatistiksel olarak kontrol edilmesi ve ařırı talep, her trden z-bildirimi tek tip řekilde etkileyen daha genel lekli kullanım tercihleridir. Sabitleme vinyetleri bireysel dzeyde, madde ieriklerini ve tepki stillerini daha zel yollarla yorumlamadaki farklılıkları hedefler. Sabitleme vinyetlerine dayalı yeniden leklendirme, hem bireysel hem de kltrel dzeyde puanlarda daha fazla eřitlilik ortaya ıkarabilir. Bununla birlikte yeniden leklendirmenin, zellikle sabitleme vinyetlerinin katı varsayımlarının olası ihlali gz nne alındıėında, bařka trden yanlılıėa yol aıp amadıėı aık deėildir. Vinyetlerin sabitlenmesinden tretilenler gibi puan dzeltmelerinin tutum-bařarı paradoksunu Doėu Asya ve Latin Amerika gibi ok farklı blgeleri karřılařtırırken tersine evirebileceėi sonucuna varılabilir. Ancak bu prosedrler kltrel olarak daha homojen bir blgede uygulandıėında yorumlanması ok daha zor sonular elde edilebilmektedir.

PISA öğrenci anketindeki maddeler likert tipi madde kategorilerinden oluşmaktadır. Tepki stili yanlılığını düzeltip daha geçerli sonuçlar elde edebilmek amacıyla likert tipi maddelerin yanında farklı madde formatlarında sorular oluşturulabilir. Zorunlu seçim formatlı sorular (forced-choice format questions) ve durumsal yargı formatı (situational judgement format) soruları gibi başka madde tasarım yöntemlerinin kullanılması tepki stili yanlılığının ortadan kaldırmak amacıyla kullanılabilir. Farklı madde formatlarının dışında, karşılaştırılabilirlik ve geçerlilik endişeleri gidermeye yardımcı olabilecek iki faktörlü modeller gibi diğer istatistiksel yöntemlerde kullanılabilir (Brown & Maydeu-Olivares, 2011; Cheung & Rensvold, 2000; Rutkowski, von Davier, & Rutkowski, 2014).

Brown ve Maydeu-Olivares, (2011) yaptıkları çalışmada çok boyutlu zorunlu seçim formatlarının tipik olarak derecelendirme ölçekleriyle ilişkilendirilen çok sayıda tepki stili etkisini azaltıp azaltamayacağını incelemişlerdir. Çok boyutlu zorunlu seçim formatlı maddelerin analizinde MTK modellemesinin nasıl uygulanabileceği gösterilmiştir. Araştırmada, Thurstone'un karşılaştırmalı veri çerçevesine dayanan çok boyutlu bir MTK modeli tanıtılmıştır. Thurstone'ın MTK modelleri, yapılandırılmış faktör yükleri ve yapılandırılmış yerel bağımlılıkları olan normal ogive modellerdir. Çeşitli zorunlu seçim tasarımları altında gizil özelliklerin nasıl geri kazanıldığını araştırmak ve en uygun anket tasarımı için kılavuzlar sağlamak için bir dizi simülasyon çalışması yapılmıştır. Modelin pratikte nasıl uygulanabileceğini göstermek için deneysel bir uygulamada verilmiştir. Araştırmadan elde edilen sonuçlara göre, önerilen metodoloji ile zorunlu seçim ölçeklerinden tahmin edilen puanların tepki stili yanlılığından arındırıldığı ve gizil özellikleri iyi bir şekilde yeniden ürettiği görülmüştür. Mevcut araştırmada kullanılan TSYDK yönteminin farklı madde formatlarından elde edilen verilerin analizinde de kullanılması tepki stili yanlılığının düzeltilmesinde hangi yöntemlerin daha çok etkili olduğuna dair ipuçları verebilir.

Araştırmada tepki stili yanlılığı bulunan veriler üzerinden gerçekleştirilen korelasyon analizlerinde; ülke içinde düşük pozitif korelasyonlar elde edilirken, ülkeler arası tutum ile başarı arasında büyük negatif korelasyon değerleri elde edilmiştir. Buna göre, tepki stillerinin ülkeler arası korelasyonlar üzerinde kültürel yanlılığının kaynaklarından biri olduğu varsayılmaktadır. He, Chung ve van de Vijver, (2020) tarafından yapılan araştırmada PISA 2012'ye katılan 64 ülke verisi kullanılarak, sabitleme vinyetleri, aşırı talep ve post hoc istatistiksel prosedürlerin [yani, ipsatizasyon (ipsatization) ve uç tepki stili düzeltmesi] karşılaştırması yapılmıştır. Çalışmada ülke düzeyinde hesaplanan sınıf içi korelasyon

değerlerinin düşük olması, düzeltme etkilerindeki değişkenliğin büyük ölçüde ulusal bağlamlara ve kültürel özelliklere (örneğin, refah düzeyi, kültürel değerler ve yönelimler) bağlı olmadığını, aksine bireysel özelliklerin daha büyük ölçüde öne çıktığını göstermiştir. Yani bireysel düzeydeki kişilik özelliklerinin ülkelerin geneline yansıdığı ve böylece kültürlere özgü değişkenler olabileceği elde edilmiştir.

Araştırmadan elde edilen sonuçlara göre; büyük ölçekli değerlendirmelerde tepki stillerini test etmemek ve düzeltmemek, bilişsel olmayan değişkenler ile bilişsel yeterlilik puanları arasındaki ilişkiyi incelerken potansiyel olarak anlamsız grup karşılaştırmalarına ve geçersiz sonuçlara yol açabilir. (Khorramdel, von Davier, Bertling, Roberts, & Kyllonen, 2017). Ayrıca, tepki stili yanlılığı bulunan veriler üzerinden yapılan korelasyon, t testi, ANOVA ve kümeleme analizleri gibi istatistiksel yöntem sonuçlarının geçerliliğinin olumsuz etkilendiği anlaşılmıştır. TSYDK yöntemi ile tepki stili yanlılığının düzeltilebildiği ve dolayısıyla potansiyel olarak tutum ölçümlerinin daha iyi karşılaştırılmasının sağlandığı görülmüştür. TSYDK yönteminin kullanılması ile bilişsel olmayan yapıların kültürlerarası karşılaştırmaya sonuçlarının daha geçerli olduğu ve daha doğru gizil yetenek puanların elde edilebileceği söylenebilir.

Tepki stillerinin beklenmedik ülke düzeyindeki korelasyonlarda kısmen etkisi olduğu mevcut ve diğer araştırma (Buckley,2009; İlgün Dibek, 2018; Lu & Bolt, 2015) sonuçları ile kanıtlanmıştır. Ayrıca TSYDK ile bu etkinin azaltılabildiği ve ölçümler arasındaki ülke düzeyindeki ilişkilerin daha gerçekçi bir tahmininin sağlandığı görülmüştür. PISA gibi büyük ölçekli veri kümelerinde tepki stili farklılıklarını tespit etmede en uygun yöntemi belirlemek için daha fazla araştırma yapılması gerekmektedir birlikte, tepki stillerinin uluslararası değerlendirmelerin ikincil analizleri üzerindeki etkisinin göz ardı edilmemesi gerektiği açıktır.

Araştırmadan elde edilen sonuçlar aşağıda maddeler halinde kısaca sunulmuştur.

1. TSYDK yönteminin, PISA 2015 uygulamasına katılan ve araştırmaya dahil edilen ülkelerde fenden hoşlanma, başarı motivasyonu ve fen öz yeterliği ölçeklerinde tepki stili yanlılığının düzeltilmesinde etkili olmuştur. TSYDK yöntemi ile her üç ölçek maddeleri içinde düzeltilmiş puanlar elde edilmiştir.
2. TSYDK yöntemi, öğrencileri ölçülen özellik açısından içerik tabanlı kümeler ayırabilmektedir. Araştırmada kullanılan fenden hoşlanma, başarı motivasyonu ve fen öz yeterliği ölçeklerinde TSYDK yöntemi birbirinden farklı olan içerik tabanlı ve tepki stili tabanlı kümeler elde edilmesini sağlamıştır.
3. Fene yönelik tutum ölçeklerinde, TSYDK yöntemi ile elde edilen içerik tabanlı kümeler ile tepki stili tabanlı kümeler birbiri ile çakışmamaktadır. Böylece, her bir içerik tabanlı kümede bütün tepki stili tabanlı kümelerden öğrenciler yer almaktadır.
4. K-ortalamar kümeleme yönteminin araştırmaya dahil edilen ülkelere ölçülen özellik açısından kümeleme performansı düşük bulunmuştur. Diğer bir ifade ile K-ortalamar kümeleme yöntemi ile elde edilen kümelere öğrencilerin çoğunluğunun belirli bir tepki kategorisini seçme durumu ön plana çıkmaktadır. Bu durumda, K-ortalamar kümeleme yönteminin bireyleri daha çok sergiledikleri tepki stillerine göre kümelediğini göstermektedir.
5. TSYDK yöntemi ile K-ortalamar kümeleme yönteminin sonuçları karşılaştırıldığında; TSYDK yöntemi K-ortalamar kümeleme yöntemine göre bireyleri kümeleme performansı daha başarılı olmuştur. Araştırmada kullanılan fene yönelik tutum ölçeklerinde, TSYDK yöntemi öğrencileri tepki stili yanlılığından arındırarak içerik tabanlı kümelerin elde edilmesini sağlamıştır. Buna karşılık K-ortalamar kümeleme yöntemi öğrencileri daha çok belirli tepki kategorileri tercih etmelerine göre sınıflandırmış ve tepki stili tabanlı kümeler ayırmıştır.
6. Ülkelerin çoğunda fene yönelik tutum ölçeklerinden elde edilen düzeltilmemiş ve düzeltilmiş tutum puanları arasında manidar farklılık bulunmuştur. Düzeltilmemiş ve düzeltilmiş tutum puanları arasında manidar farklılaşmanın ülke sayısı açısından en fazla olduğu ölçek başarı motivasyonu ölçeği olmuştur.
7. Tepki stillerinin fene yönelik tutumlar üzerindeki etkisi ülkeden ülkeye olumlu veya olumsuz yönde değişmektedir. Tepki stilleri, fen başarısı yüksek olan ülkelere öğrencilerin fene yönelik tutumlarını olduğundan daha olumsuz şekilde

kestirilmesine, fen başarısı düşük olan ülkelerde ise öğrencilerin fene yönelik tutumlarını olduğundan daha olumlu şekilde kestirilmesine yol açmaktadır.

8. Öğrencilerin fenden hoşlanma, başarı motivasyonu ve fen öz yeterlikleri ile fen başarıları arasındaki ilişkide düzeltilmemiş ve düzeltilmiş puanlar aracılığıyla elde edilen ülke içi korelasyon değerleri pozitif ve birbirine yakın bulunmuştur. Ayrıca, her üç ölçek içinde düzeltilmemiş ve düzeltilmiş puanlar aracılığıyla elde edilen ülkeler arası tutum-başarı korelasyon değerleri negatiftir. Düzeltilmiş puanlar aracılığıyla elde edilen ülkeler arası tutum-başarı korelasyon değerleri, düzeltilmemiş puanlar aracılığıyla elde edilen ülkeler arası tutum-başarı korelasyon değerinden daha düşüktür.

Öneriler

Aşağıda araştırma kapsamında elde edilen sonuçlar doğrultusunda araştırmacılara ve uygulayıcılara yönelik bazı öneriler sunulmuştur.

Araştırmacılara Yönelik Öneriler

- PISA 2015’te fene yönelik tutum ölçeklerinde yalnızca dört derecelendirme seçeneği olduğu düşünüldüğünde, daha fazla tepki seçeneği olan derecelendirme ölçekleri kullanıldığında daha büyük farklılıklar bulunabilir. Araç geliştirmede yeterli bir yanıt aralığı önemlidir, çünkü bu olmadan maddeler arası korelasyonlar sınırlı ve zayıf olacaktır (Liu, 2015). Likert ölçekleri gibi sürekli ölçekler için, genel olarak önerilen derecelendirme seçeneklerinin sayısı beşten yediye kadardır (Kirchhoff, 1999, Pedhazur & Schmelkin, 1991, Sudman & Bradburn, 1982). Bu nedenle tepki kategorilerinin beşli, yedili veya dokuzlu olduğu maddelerden oluşan simülasyon verileri üretilerek kategori sayısına göre TSYDK yönteminin elverişliliği araştırılabilir.
- Farklı ölçek kullanım düzeltme prosedürlerinin eşzamanlı uygulamaları ve karşılaştırmaları nadirdir ve bu çalışmalar, her prosedürün göreceli yararlarına ve sınırlıklarına odaklanma eğilimindedir (He vd., 2017; Kyllonen & Bertling, 2014). Ancak uygulanan düzeltme yöntemlerinin hangi psikometrik özellikleri değiştirdiği ve bu düzeltmelerin bireysel ve ülke düzeyindeki tepkileri ne ölçüde farklı şekilde

etkileyebileceği belirsizliğini korumaktadır. Mevcut araştırmada düzeltme yöntemi olarak uygulanan TSYDK yönteminin aşırı talep ve sabitleme vinyetleri gibi farklı düzeltme yöntemleri ile karşılaştırma çalışmaları yapılabilir.

- Araştırmada TSYDK yöntemi aracılığıyla ülke düzeyindeki tepki stillerinin belirlenmesi ve düzeltilmesine odaklanılmıştır. Bireylerin farklı psikolojik özelliklerin ölçülmesinin amaçlandığı ileriki çalışmalarda bireysel düzeyde tepki stillerinin belirlenmesi ve düzeltilmesinde TSYDK yönteminin kullanımı araştırılabilir.
- Mevcut araştırmada analizlerin uzun sürmesi ve bilgisayarın işlem yapabilme kapasitesi gibi nedenler göz önüne alınarak toplam 21 ülke analize dâhil edilmiştir. Bu sınırlılıklar nedeni ile temsiliyeti daha iyi sağlayabilmek için ileriki araştırmalarda mevcut analiz ülke sayısı artırılarak gerçekleştirilebilir. Ayrıca, mevcut çalışmada analizlerin işlem süresinin çok uzun olması nedeni ile öğrencilerin fen başarı puanları olarak PISA 2015'deki 10 olası değerlerin ağırlıklandırılmış örneklem ortalaması kullanılmıştır. İlerideki araştırmalarda öğrencilerin fen başarılarını temsil etmek amacıyla 10 olası değer her biri için analizler gerçekleştirilebilir.
- Tepki stili yanlılığının ortadan kaldırılması veya hafifletilmesi amacıyla basit istatistiksel yöntemlerin yanında MTK ve GSA temelli yaklaşımlarında kullanıldığı birçok çalışma bulunmaktadır. Bu araştırmada ise bu yaklaşımlardan farklı olarak TSYDK yöntemi kullanılmıştır. TSYDK yönteminin en önemli avantajının eş zamanlı olarak tepki stili yanlılığının düzeltilmesi ve içeriğe dayalı kümelemenin yapılmasına olanak sağlamasıdır. Bu araştırmadan elde edilen sonuçlar neticesinde TSYDK yönteminin özellikle ölçme ve değerlendirme alanındaki araştırmacıların ilgisini çekeceği düşünülmektedir.
- Mevcut araştırmada tepki stili yanlılığının belirlenmesinde ve düzeltilmesinde sadece TSYDK yönteminden yararlanılmıştır. Tepki stili yanlılığının belirlenmesi ve düzeltilmesinde ÇBMTK ve SÇÖ gibi yöntemlerde mevcuttur. Hem TSYDK yönteminin etkililiğini hem de hangi düzeltme yönteminin gerçekten de yanlılığı azalttığı ve farklı ölçülerdeki geçerliliği arttırdığını belirlemek için ileriki araştırmalarda TSYDK yöntemi ile diğer model tabanlı yaklaşımların karşılaştırılması yapılabilir.

- Mevcut çalışmanın sonuçlarının derecelendirme veya likert tipi ölçeklerle ölçülen diğer bilişsel olmayan yapılara genellenip genellenemeyeceğine yönelik araştırmalar yapılabilir.
- Araştırmada tutum-başarı paradoksunu açıklamak için tepki stili yanlılığı göz önüne alınmıştır. Ancak tutum-başarı paradoksu üzerinde etkisi olduğu düşünülen büyük balık- küçük göl, referans grup ve sosyal kabul edilebilirlik gibi özelliklerinde ele alınıp incelendiği daha geniş kapsamlı araştırmalara gereksinim duyulmaktadır.
- Pek çok çalışma, ölçek kullanımındaki tepki stili yanlılığını düzeltme prosedürlerinin yapısal ilişkiler ve kültürler arası araştırmalarda ortalama karşılaştırmalar üzerinde önemli bir etkisi olduğunu bildirmiştir (He, Chung, & van de Vijver, 2020). Bu doğrultuda özellikle geniş ölçekli sınavlardaki ölçekler aracılığıyla yapılacak ikincil analizlerde, TSYDK yöntemi ile tepki stili yanlılığının düzeltilmesi ikincil analiz sonuçlarının geçerliği açısından önemli görülmektedir. Bu nedenle, geniş ölçekli sınavlar üzerinden yapılan kültürlerarası karşılaştırma çalışmalarının geçerliliklerini arttırmak amacıyla TSYDK yöntemi kullanılarak veriler tepki stili yanlılığından arındırılabilir.

Uygulayıcılara Yönelik Öneriler

- Mevcut analizin amaçlarından biri, PISA 2015 değerlendirmesinde ülke düzeyinde fene yönelik tutum ile fen başarısı arasındaki korelasyonlarda tepki stili farklılıklarının etkilerini değerlendirmektir. PISA' nın önemli bir amacı, hem fen başarısındaki hem de fene yönelik tutumdaki kültürler arası farklılıkları daha iyi anlamak olduğu için, mevcut uygulama yalnızca PISA' nın yukarıda söylenen amacına ilişkin sonuçlar elde etmekle kalmaz, aynı zamanda uluslararası değerlendirmelerde tepki stilini hesaba katılması gerektiğini belirtir. Bu nedenle, PISA ve TIMSS gibi uluslararası geniş ölçekli sınavların raporlarında, kültürlerin sahip olduğu tepki stillerin belirlenmesine yönelik çalışma sonuçlarına yer verilebilir.
- Uluslararası ve kültürler arası ölçeklerin dâhil olmak üzere tepki stillerinin ölçme yanlılığını gösteren geniş ve büyüyen literatürüne rağmen, araştırmacılar hem ölçme aracı tasarımında hem de ikincil analizlerde tepki stili yanlılığını dikkate

almamaktadırlar (Smith 2003, s. 69). Bu araştırmadan elde edilen sonuçlar doğrultusunda, PISA ve TIMSS gibi uluslararası sınavlarda özellikle bilişsel olmayan yapıların ölçülmesi amacıyla geliştirilen ölçme araçlarının tasarımında tepki stili yanlılığının göz önüne alınması gerektiği önemli görülmektedir. Bu nedenle, bu sınavların uygulayıcılarının tepki stili yanlılığını azalttığı birçok araştırmada kanıtlanmış olan sabitleme vinyetleri, aşırı talep ve çok boyutlu zorunlu seçim formatları gibi müdahalelerinin kapsamını değerlendirmek için ölçeğin pilot aşamasının bir parçası olarak bir dizi madde formatını kullanabilirler.

- PISA ve diğer uluslararası eğitim değerlendirmelerine katılan çok sayıda ülke göz önüne alındığında, eğitim alanında ülkeler arası farklılıkların ortaya çıkarılmasında bu tür anket deneylerinin derinlemesine etnografik araştırmalardan çok daha uygun olduğu düşünülmektedir. Bu deneylere paralel olarak, OECD veya üye ülkeler, ölçek kullanım heterojenliğini ölçmek ve düzeltmek için daha gelişmiş istatistiksel yöntemler hakkında ek araştırmalara sponsor olabilirler (Buckley, 2009).
- Araştırmada tutum-başarı paradoksunda tepki stillerinin açıklayıcı rolünün bulunması ile beraber, tepki stili yanlılığının düzeltilmesine rağmen ülkeler arası tutum-başarı korelasyon sonuçlarının negatifliğinin devam ettiği görülmüştür. Diğer taraftan eğitimin, alçakgönüllülüğün, özeleştirinin, çabaların ve sürekli kendini geliştirmenin değerini vurgulayan Konfüçyüsçü değer sistemi, paradoksal gibi görünen tutum ve başarı modeline kavramsal zemini sağlayabildiği düşünülmektedir (He & van de Vijver, 2016). Bu nedenle tutum-başarı paradoksunun daha iyi anlaşılması amacıyla PISA ve TIMSS gibi uluslararası sınavlarda ülkelerin değer yargılarını belirlemek amacıyla farklı anket ve ölçme araçları hazırlanabilir.
- Öğretme ve öğrenme kültürler etkinliklerdir ve PISA çok farklı kültürel, sosyal, tarihi ve politik geçmişe sahip ülkelerin katılım gösterdiği bir uygulamadır. PISA’ da yer alan bilişsel ve duyuşsal veriler aracılığıyla ülkelerin eğitim kalitesi, işleyişi, eğitim paydaşları (okul, öğretmen, öğrenci, aile) ve eğitim politikaları ile ilgili önemli bilgiler elde edilebilmektedir. PISA’ da yer alan duyuşsal veriler öğrencilerin kendi beyanlarına dayanmaktadır. Mevcut araştırmadan elde edilen sonuçlar da duyuşsal veriler üzerinde ülkelere özgü tepki stili yanlılığının bulunduğunu göstermiştir. Bu nedenle ülkelerdeki eğitim yöneticilerinin, PISA’ dan elde edilen öğrencilerin

duyuşsal  zellikleri ile başarıları arasındaki ilişkilerde  lkelere  zg  tepki stili eğilimlerini dikkate almaları gerektięi d ş n lmektedir.



KAYNAKLAR

- Acar, T. (2012). Türkiye'nin PISA 2009 sonuçlarına göre OECD'ye üye ve aday ülkeler arasındaki yeri, *Kuram ve Uygulamada Eğitim Bilimleri*, 12 (4), 2561-2572.
- Aksu, G., Güzeller, C. O., & Eser, M. T. (2017). Öğrencilerin matematik okuryazarlığı performanslarının aşamalı doğrusal model (HLM) ile incelenmesi: PISA 2012 Türkiye örneği. *Eğitim ve Bilim*, 42(191), 247-266.
<http://dx.doi.org/10.15390/EB.2017.6956>.
- Aksu, G., Güzeller, C. O., & Eser, M. T. (2017). PISA 2012 sonuçlarının duyuşsal özelliklere göre kümeleme çalışması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 32(4), 838-862.
- Akyüz, G., & Pala, N. M. (2010). PISA 2003 sonuçlarına göre öğrenci ve sınıf özelliklerinin matematik okuryazarlığına ve problem çözme becerilerine etkisi. *İlköğretim Online*, 9(2), 668-678.
- Anderson, D. (2012). *Hierarchical linear modeling (HLM): An introduction to key concepts within cross-sectional and growth modeling frameworks*. (Technical Report 1308). Eugene: Behavioral Research and Teaching.
- Anıl, D. (2009). Uluslararası öğrenci başarılarını değerlendirme programı (PISA)'nda Türkiye'deki öğrencilerin fen bilimleri başarılarını etkileyen faktörler. *Eğitim ve Bilim*, 34(152), 87-100.
- Arabie, P., & Hubert, L. (1994). Cluster analysis in marketing research. In R. P. Bagozzi (Ed.), *Advanced methods of marketing research* (pp. 160–189). Oxford, UK: Blackwell.
- 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(3), 374-392. <https://doi.org/10.1177/0013164405278575>
- Arıkan, S., Özer, F., Şeker, V., & Ertaş, G. (2020). Geniş ölçekli testlerde örneklem ağırlıklarının ve olası değerlerin önemi. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 11(1), 43-60. <https://doi.org/10.21031/epod.602765>
- Aydın, A., Sarier, Y., & Uysal, Ş. (2012). Sosyoekonomik ve sosyokültürel açılarından PISA matematik sonuçlarının karşılaştırılması. *Eğitim ve Bilim*, 37(164), 20-30.
- Bachman, J. G., & O'MALLEY, P. M. (1984). Yea-saying, nay-saying, and going to extremes: Black-white differences in response styles. *Public Opinion Quarterly*, 48(2), 491-509. <https://doi.org/10.1086/268845>
- 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. Retrieved from <http://www.monitoringthefuture.org/>
- Bandura, A. (1994). Self-efficacy. *Encyclopedia of Human Behavior*, 4, 71–81.
- Baumgartner, H., & Steenkamp, J. B. E. M. (2001). Response styles in marketing research: A cross-national investigation. *Journal of Marketing Research*, 38(2), 143-156. <https://doi.org/10.1509/jmkr.38.2.143.18840>
- Benítez, I., He, J., Van de Vijver, F. J., & Padilla, J. L. (2016). Linking extreme response style to response processes: A cross-cultural mixed methods approach. *International Journal of Psychology*, 51(6), 464-473. <https://doi.org/10.1002/ijop.12379>
- Beuthner, C., Friedrich, M., Herbes, C., & Ramme, I. (2018). Examining survey response styles in cross-cultural marketing research: A comparison between Mexican and South Korean respondents. *International Journal of Market Research*, 60(3), 257-267. <https://doi.org/10.1177/1470785318762015>
- Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37(1), 29-51. <https://link.springer.com/content/pdf/10.1007/BF02291411.pdf> sayfasından erişilmiştir.

- Bolt, D. M., & Johnson, T. R. (2009). Addressing score bias and differential item functioning due to individual differences in response style. *Applied Psychological Measurement*, 33(5), 335-352. <https://doi.org/10.1177/0146621608329891>
- Bolt, D. M., & Newton, J. R. (2011). Multiscale measurement of extreme response style. *Educational and Psychological Measurement*, 71(5), 814-833. <https://doi.org/10.1177/0013164410388411>
- Bolt, D.M., Lu, Y., & Kim, J-S. (2014). Measurement and control of response styles using anchoring vignettes: A model-based approach. *Psychological Methods*, 19, 528–541. <https://doi.org/10.1037/met0000016>
- Bond, M. H., & Smith, P. B. (1996). Cross-cultural social and organizational psychology. *Annual Review of Psychology*, 47(1), 205-235. <https://doi.org/10.1146/annurev.psych.47.1.205>
- Bos, K. (2002). Benefits and limitations of large-scale international comparative achievement studies (Doctoral dissertation). Retrieved from <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.472.359&rep=rep1&type=pdf>
- Böckenholt, U. (2012). Modeling multiple response processes in judgment and choice. *Psychological Methods*, 17, 665–678. <https://doi.org/10.1037/a0028111>
- Böckenholt, U., & Meiser, T. (2017). Response style analysis with threshold and multi-process IRT models: A review and tutorial. *British Journal of Mathematical and Statistical Psychology*, 70(1), 159-181. <https://doi.org/10.1111/bmsp.12086>
- Brown, A., & Maydeu-Olivares, A. (2011). Item response modeling of forced-choice questionnaires. *Educational and Psychological Measurement*, 71(3), 460-502. <https://doi.org/10.1177/0013164410375112>
- Buckley, J. (2009). *Cross-national response styles in international educational assessment: Evidence from PISA 2006*. Paper presented at the NCES Conference on the Program for International Student Assessment: What we can learn from PISA, Washington, DC, May 2010. Retrieved from <https://edsurvey.rti.org/PISA/>
- Bulle, N. (2011). Comparing OECD educational models through the prism of PISA. *Comparative Education*, 47(4), 503-521. <https://doi.org/10.1080/03050068.2011.555117>

- Büyüköztürk, Ş. (2018). *Sosyal bilimler için veri analizi el kitabı*. Ankara: Pegem.
- Charrad, M., Ghazzali, N., Boiteau, V., Niknafs, A., & Charrad, M. M. (2014). *Nbclust: An examination of indices for determining the number of clusters*. R package version 2.0.3. Retrieved from <https://www.jstatsoft.org/article/view/v061i06>
- Chen, C., Lee, S. Y., & Stevenson, H. W. (1995). Response style and cross-cultural comparisons of rating scales among East Asian and North American students. *Psychological Science*, 6(3), 170-175. <https://doi.org/10.1111/j.1467-9280.1995.tb00327.x>
- Cheung, G. W., & Rensvold, R. B. (2000). Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling. *Journal of Cross-Cultural Psychology*, 31(2), 187-212. <https://doi.org/10.1177/0022022100031002003>
- 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* (Doctoral dissertation). Retrieved from <https://drum.lib.umd.edu/handle/1903/14511>
- Clarke, I. (2001). Extreme response style in cross-cultural research. *International Marketing Review* 18(3), 301-324. <https://doi.org/10.1108/02651330110396488>
- Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement*, 6, 475-494. <https://journals.sagepub.com/doi/abs/10.1177/001316444600600405?journalCode=epma> sayfasından erişilmiştir.
- De Beuckelaer, A., Weijters, B., & Rutten, A. (2010). Using ad hoc measures for response styles: A cautionary note. *Quality & Quantity*, 44(4), 761-775.
- De Boeck, P., & Wilson, M. (2004). *Explanatory item response models: A generalized linear and nonlinear approach*. New York, NY: Springer.
- De Jong, M. G., Steenkamp, J. B. E., 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.

- Demirez, G. (2018). *Bazı değişkenlerin fen başarı puanına etkisi: PISA 2015 Türkiye, Singapur ve Almanya örneği*. (Yüksek Lisans Tezi). <https://tez.yok.gov.tr> sayfasından erişilmiştir.
- Dibek, M. I. (2019). Examination of the extreme response style of students using IRTree: The case of TIMSS 2015. *International Journal of Assessment Tools in Education*, 6(2), 300-313. <https://doi.org/10.21449/ijate.534118>
- 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), 1-12.
- Eid, M., & Rauber, M. (2000). Detecting measurement invariance in organizational surveys. *European Journal of Psychological Assessment*, 16(1), 20-30. <https://doi.org/10.1027/1015-5759.16.1.20>
- Eid, M., Langeheine, R., & Diener, E. (2003). Comparing typological structures across cultures by multigroup latent class analysis: A primer. *Journal of Cross-Cultural Psychology*, 34(2), 195-210. <https://doi.org/10.1177/0022022102250427>
- Eser, M. T. (2019). *Uluslararası öğrenci değerlendirme programı 2015 verilerinin verimadenciliğinde kümeleme yöntemleriyle incelenmesi*, (Doktora tezi). <https://tez.yok.gov.tr> sayfasından erişilmiştir.
- Falk, C. F., & Cai, L. (2016). A flexible full-information approach to the modeling of response styles. *Psychological Methods*, 21(3), 328-347. <https://doi.org/10.1037/met0000059>
- Genolini, C., & Falissard, B. (2010). KmL: k-means for longitudinal data. *Computational Statistics*, 25(2), 317-328.
- Greenleaf, E. A. (1992). Improving rating scale measures by detecting and correcting bias components in some response styles. *Journal of Marketing Research*, 29(2), 176-188. <https://doi.org/10.1177/002224379202900203>
- Gonzalez, E. J. (2012). Rescaling sampling weights and selecting mini-samples from large-scale assessment databases. *IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments*, 5, 117-134.

- Hamilton, D. L. (1968). Personality attributes associated with extreme response style. *Psychological Bulletin*, 69(3), 192-203. <https://doi.org/10.1037/h0025606>
- Hartigan, J. A., & Wong, M. A. (1979). AK-means clustering algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 28(1), 100-108. <https://doi.org/10.2307/2346830>
- Harzing, A. (2006). Cross cultural management response styles in a 26-country study. *International Journal of Cross Cultural Management*, 6(2), 243-266. <https://doi.org/10.1177/1470595806066332>
- Haskell, K. H., & Hanson, R. J. (1981). An algorithm for linear least squares problems with equality and nonnegativity constraints. *Mathematical Programming*, 21(1), 98-118. <https://doi.org/10.1007/BF01584232>
- He, J., & Van de Vijver, F. J. (2013). A general response style factor: Evidence from a multi-ethnic study in the Netherlands. *Personality and Individual Differences*, 55(7), 794-800. <https://doi.org/10.1016/j.paid.2013.06.017>
- He, J., & Van De Vijver, F. J. (2015). Effects of a general response style on cross-cultural comparisons: Evidence from the Teaching and Learning International Survey [Special issue]. *Public Opinion Quarterly*, 79(1), 267-290. <https://doi.org/10.1093/poq/nfv006>
- He, J., & Van de Vijver, F. J. (2016). The motivation-achievement paradox in international educational achievement tests: Toward a better understanding. *In the psychology of Asian learners* (pp. 253-268). Springer, Singapore.
- He, J., & Van de Vijver, F. J. R. (2015). The value of keeping an open eye for methodological issues in research on resilience and culture. In L. Theron, M. Ungar, & L. Liebenberg (Eds.), *Youth resilience and culture. Commonalities and complexities* (pp. 189-201). New York, NY: Springer.
- He, J., Chung, J. M., & Van De Vijver, F. J. (2020). An examination of different scale usage correction procedures to enhance cross-cultural data comparability. *Journal of Cross-Cultural Psychology*, 51(10), 794-813. <https://doi.org/10.1177/0022022120960815>
- He, J., van de Vijver, F. J. R., Fetvadjev, V. H., Espinosa, A. C. D., Adams, B.,.....Hapunda, G. (2017). On enhancing the crosscultural comparability of

- Likert-scale personality and value measures: A comparison of common procedures. *European Journal of Personality*, 31, 642–657. <https://doi.org/10.1002/per.2132>
- He, J., Van de Vijver, F. J., Espinosa, A. D., & Mui, P. H. (2014). Toward a unification of acquiescent, extreme, and midpoint response styles: A multilevel study. *International Journal of Cross Cultural Management*, 14(3), 306-322. <https://doi.org/10.1177/1470595814541424>
- He, Jia & Van de Vijver, Fons (2016). Correcting for scale usage differences among Latin American countries, Portugal, and Spain in PISA. *RELIEVE*, 22(1), M9. <http://dx.doi.org/10.7203/relieve.22.1.8282>
- Heine, S. J., Kitayama, S., Lehman, D. R., Takata, T., Ide, E., Leung, C., & Matsumoto, H. (2001). Divergent consequences of success and failure in japan and north america: an investigation of self-improving motivations and malleable selves. *Journal of Personality and Social Psychology*, 81(4), 599–615. <https://doi.org/10.1037/0022-3514.81.4.599>
- Hofstede, G. H. (2001). *Cultures consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Thousand Oaks, California: Sage Publications.
- Hofstede, G., & McCrae, R.R. (2004). Personality and culture revisited: Linking traits and dimensions of Culture. *Cross- Cultural Research*, 38(1), 52–88. <https://doi.org/10.1177/1069397103259443>
- Holtgraves, T. (2004). Social desirability and self-reports: Testing models of socially desirable responding. *Personality and Social Psychology Bulletin*, 30(2), 161-172. <https://doi.org/10.1177/0146167203259930>
- Hui, C. H., & Triandis, H. C. (1989). Effects of culture and response format on extreme response style. *Journal of Cross-Cultural Psychology*, 20(3), 296-309. <https://doi.org/10.1177/0022022189203004>
- Hwang, H., Dillon, W. R., & Takane, Y. (2006). An extension of multiple correspondence analysis for identifying heterogeneous subgroups of respondents. *Psychometrika*, 71(1), 161–171. <https://doi.org/10.1007/s11336-004-1173-x>

- İlgün Dibek M. (2018). *TIMSS 2015 uygulamasında tutum-başarı paradoksunun uç tepki stiline göre çok boyutlu madde tepki kuramı ile modellenmesi*. (Doktora Tezi). <https://tez.yok.gov.tr>. sayfasından erişilmiştir.
- İlgün Dibek, M., Yavuz, H. Ç., & Çokluk Bökeoğlu, Ö. (2018). Tutum-başarı paradoksunda tepki stillerinin rolü: Dokuz ülkenin karşılaştırılması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18 (2), 932-952.
- Jakobs, C., Selbach, D., Böldt, S., & Steffens, J. (2019). *The influence of extreme response style on sound evaluations*. Universitätsbibliothek der RWTH Aachen.
- Jeon, M., & De Boeck, P. (2019). Evaluation on types of invariance in studying extreme response bias with an IRTree approach. *British Journal of Mathematical and Statistical Psychology*, 72(3), 517-537. <https://doi.org/10.1111/bmsp.12182>
- Jin, K. Y., & Chen, H. F. (2020). MIMIC approach to assessing differential item functioning with control of extreme response style. *Behavior Research Methods*, 52(1), 23-35. Retrieved from <https://link.springer.com/article/10.3758/s13428-019-01198-1>
- 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(1), 92-114. <https://doi.org/10.3102/1076998609340529>
- Johnson, T., Kulesa, P., Cho, Y. I., & Shavitt, S. (2005). The relation between culture and response styles: Evidence from 19 countries. *Journal of Cross-Cultural Psychology*, 36(2), 264-277. <https://doi.org/10.1177/0022022104272905>
- Johnson, T., O'Rourke, D., Chavez, N., Sudman, S., Warnecke, R., Lacey, L., & Horm, J. (1997). Social cognition and responses to survey questions among culturally diverse populations. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 87–113). New York: John Wiley & Sons.
- Ju, U., & Falk, C. F. (2019). Modeling response styles in cross-country self-reports: An application of a multilevel multidimensional nominal response model. *Journal of Educational Measurement*, 56(1), 169-191.
- Kankaraš, M., Vermunt, J. K., & Moors, G. (2011). Measurement equivalence of ordinal items: a comparison of factor analytic, item response theory, and latent class

- approaches. *Sociological Methods & Research*, 40(2), 279-310.
<https://doi.org/10.1177/0049124111405301>
- Kaufman, L., & Rousseeuw, P. (1990). *Finding groups in data: An introduction to cluster analysis*. New York: Wiley.
- Kemmelmeier, M. (2016). Cultural differences in survey responding: Issues and insights in the study of response biases. *International Journal of Psychology*, 51(6), 439-444.
<https://doi.org/10.1002/ijop.12386>
- Khorramdel, L., von Davier, M., Bertling, J. P., Roberts, R. D., & Kyllonen, P. C. (2017). Recent IRT approaches to test and correct for response styles in PISA background questionnaire data: A feasibility study. *Psychological Test and Assessment Modeling*, 59(1), 71-92.
- Kieruj, N. D., & Moors, G. (2013). Response style behavior: question format dependent or personal style?. *Quality & Quantity*, 47(1), 193-211.
<https://link.springer.com/article/10.1007/s11135-011-9511-4> sayfasından erişilmiştir.
- King, G., Murray, C. J., Salomon, J. A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, 98(1), 191-207.
- Kirchhoff, K. T. (1999). Design of questionnaires and structured interviews. In M. A. Mateo & K. T. Kirchhoff (Eds.), *Using and conducting nursing research in the clinical setting* (pp. 229-239). Philadelphia: W. B. Saunders.
- Kjærnsli, M., & Lie, S. (2011). Students' preference for science careers: International comparisons based on PISA 2006. *International Journal of Science Education*, 33(1), 121-144. <https://doi.org/10.1080/09500693.2010.518642>
- Krzanowski, W. J., & Lai, Y. T. (1988). A criterion for determining the number of groups in a data set using sum-of-squares clustering. *Biometrics*, 44(1), 23-34.
<https://doi.org/10.2307/2531893>
- Kyllonen, P. C., & Bertling, J. J. (2014). Innovative questionnaire assessment methods to increase cross-country comparability. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (pp. 277-286). CRC.

- Kyllonen, P., Lietz, P., & Roberts, R. (2012). Cross-cultural comparative questionnaire issues. *Paper prepared at the PISA*. https://edsurveys.rti.org/PISA/documents/Buckley_PISAresponsestyle.pdf sayfasından erişilmiştir.
- Larose, D. T., & Larose, C. D. (2014). *Discovering knowledge in data: an introduction to data mining (Vol. 4)*. John Wiley & Sons.
- Lee, C., & Green, R. T. (1991). Cross-cultural examination of the Fishbein behavioral intentions model. *Journal of international business studies*, 22(2), 289-305. <https://doi.org/10.1057/palgrave.jibs.8490304>
- Lee, J. (2009). Universals and specifics of math self-concept, math self-efficacy, and math anxiety across 41 PISA 2003 participating countries. *Learning and Individual Differences*, 19(3), 355–365.
- Leventhal, B. (2017). *Extreme response style: Which model is best?* (Doctoral dissertation). Retrieved from http://d-scholarship.pitt.edu/31405/1/Leventhal_Dissertation.pdf.
- Linnakylä, P., & Malin, A. (2008). Finnish students' school engagement profiles in the light of PISA 2003. *Scandinavian Journal of Educational Research*, 52(6), 583-602. <https://doi.org/10.1080/00313830802497174>
- Liu, M. (2015). Response style of rating scales: The effects of data collection mode, scale format, and acculturation (Doctoral dissertation). Retrieved from https://deepblue.lib.umich.edu/bitstream/handle/2027.42/111591/liumn_1.pdf?sequence=1&isAllowed=y
- Liu, M., & Keusch, F. (2017). Effects of scale direction on response style of ordinal rating scales. *Journal of Official Statistics*, 33(1), 137-154. <https://doi.org/10.1515/jos-2017-0008>
- Liu, M., Conrad, F. G., & Lee, S. (2017). Comparing acquiescent and extreme response styles in face-to-face and web surveys. *Quality & Quantity*, 51(2), 941-958.
- Liu, M., Harbaugh, A. G., Harring, J. R., & Hancock, G. R. (2017). The effect of extreme response and non-extreme response styles on testing measurement invariance. *Frontiers in psychology*, 8, 726. <https://doi.org/10.3389/fpsyg.2017.00726>

- Liu, Y., Wu, A. D., & Zumbo, B. D. (2010). The impact of outliers on Cronbach's coefficient alpha estimate of reliability: Ordinal/rating scale item responses. *Educational and Psychological Measurement*, 70(1), 5-21. <https://doi.org/10.1177/0013164409344548>
- Loveless, T. (2006). *The 2006 Brown center report on American education: How well are American students learning?* Washington, D.C.:Brookings Institution. Retrieved from <http://www.brookings.edu/index/publications.htm>
- Lu, Y. (2012). *A multilevel multidimensional item response theory model to address the role of response style on measurement of attitudes in PISA 2006*. Doctoral Dissertation, University of Wisconsin-Madison.
- Lu, Y., & 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(1), 1-18. <https://doi.org/10.1186/s40536-015-0012-0>
- Maaz, K., Trautwein, U., Lüdtke, O., & Baumert, J. (2008). Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Development Perspectives*, 2(2), 99-106.
- Marin, G., Gamba, R.J., & Marin, B.V. (1992). Extreme response style and acquiescence among hispanics. The role of acculturation and education. *Journal of Cross-Cultural Psychology* 23(4), 498–509. <https://doi.org/10.1177/0022022192234006>
- Marsh, H. W., & Hau, K. T. (2004). Explaining paradoxical relations between academic self-concepts and achievements: Cross-cultural generalizability of the internal/external frame of reference predictions across 26 countries. *Journal of Educational Psychology*, 96(1), 56-67. <https://doi.org/10.1037/0022-0663.96.1.56>
- Marsh, H. W., Seaton, M., Trautwein, U., Lüdtke, O., Hau, K. T., O'Mara, A. J., & 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(3), 319-350.
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self concept, interest, and standardized test scores: Reciprocal effects models of causal

- ordering. *Child Development*, 76(2), 397–416. <https://doi.org/10.1111/j.1467-8624.2005.00853.x>
- MEB (2016). *PISA 2015 ulusal raporu*. ANKARA: Ölçme, değerlendirme ve sınav hizmetleri genel müdürlüğü, http://pisa.meb.gov.tr/wp-content/uploads/2016/12/PISA2015_Ulusal_Rapor1.pdf sayfasından erişilmiştir.
- MedCalc Software bvba (2018). MedCalc Statistical Software version 18.2.1, Ostend, Belgium. Retrieved from <https://www.medcalc.org>
- Meisenberg, G., & Williams, A. (2008). Are acquiescent and extreme response styles related to low intelligence and education?. *Personality and Individual Differences*, 44(7), 1539-1550. <https://doi.org/10.1016/j.paid.2008.01.010>
- Min, I., Cortina, K. S., & Miller, K. F. (2016). Modesty bias and the attitude-achievement paradox across nations: A reanalysis of TIMSS. *Learning and Individual Differences*, 51, 359-366. <https://doi.org/10.1016/j.lindif.2016.09.008>
- Monseur, C., & Adams, R. (2009). Plausible values: How to deal with their limitations. *Journal of Applied Measurement*, 10(3), 1-15.
- Moors, G. (2003). Diagnosing response style behavior by means of a latent-class factor approach. Socio-demographic correlates of gender role attitudes and perceptions of ethnic discrimination reexamined. *Quality and Quantity*, 37(3), 277-302. Retrieved from <https://link.springer.com/article/10.1023/A:1024472110002>
- Moors, G. (2008). Exploring the effect of a middle response category on response style in attitude measurement. *Quality & Quantity*, 42(6), 779-794. <https://doi.org/10.1007/s11135-006-9067-x>
- Morren, M., Gelissen, J. P., & Vermunt, J. K. (2011). Dealing with extreme response style in cross-cultural research: A restricted latent class factor analysis approach. *Sociological Methodology*, 41(1), 13-47. <https://doi.org/10.1111/j.1467-9531.2011.01238.x>
- Morren, M., Gelissen, J. P., & Vermunt, J. K. (2013). Exploring the response process of culturally differing survey respondents with a response style: A sequential mixed methods study. *Field Methods*, 25(2), 162-181. <https://doi.org/10.1177/1525822X12453535>

- Narayan, S., & Krosnick, J. A. (1996). Education moderates some response effects in attitude measurement. *Public Opinion Quarterly*, 60(1), 58-88. <https://doi.org/10.1086/297739>
- OECD, (2016). *PISA 2015 Results in Focus*. <http://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf> sayfasından erişilmiştir.
- Olsen, R. V., & Lie, S. (2011). Profiles of students' interest in science issues around the world: Analysis of data from PISA 2006. *International Journal of Science Education*, 33(1), 97-120. <https://doi.org/10.1080/09500693.2010.518638>
- Ostendorf, F., & Angleitner, A. (2004). *NEO-PI-R: NEO-Persönlichkeitsinventar nach Costa und McCrae* [NEO-PI-R: NEO Personality Inventory after Costa and McCrae]. Göttingen, Germany: Hogrefe
- Park, M., & Wu, A. D. (2019). Item response tree models to investigate acquiescence and extreme response styles in Likert-type rating scales. *Educational and Psychological Measurement*, 79(5), 911-930. <https://doi.org/10.1177/0013164419829855>
- 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* (pp. 17-59). San Diego, CA: Academic.
- Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum.
- Plieninger, H., & Meiser, T. (2014). Validity of multiprocess IRT models for separating content and response styles. *Educational and Psychological Measurement*, 74(5), 875-899. <https://doi.org/10.1177/0013164413514998>
- Qiao, X., & Jiao, H. (2018). Data mining techniques in analyzing process data: A didactic. *Frontiers in Psychology*, 9, 2231. <https://doi.org/10.3389/fpsyg.2018.02231>
- Rammstedt, B., & Kemper, C. J. (2011). Measurement equivalence of the Big Five: Shedding further light on potential causes of the educational bias. *Journal of Research in Personality*, 45(1), 121-125. <https://doi.org/10.1016/j.jrp.2010.11.006>
- Rammstedt, B., Danner, D., & Bosnjak, M. (2017). Acquiescence response styles: A multilevel model explaining individual-level and country-level differences.

Personality and Individual Differences, 107, 190-194.
<https://doi.org/10.1016/j.paid.2016.11.038>

- Rammstedt, B., Goldberg, L. R., & Borg, I. (2010). The measurement equivalence of Big-Five factor markers for persons with different levels of education. *Journal of Research in Personality*, 44(1), 53-61. <https://doi.org/10.1016/j.jrp.2009.10.005>
- Ramsay, J. O. (1988). Monotone regression splines in action. *Statistical Science*, 3(4), 425–441. <https://doi.org/10.1214/ss/1177012761>
- Reynolds, N., & 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(2), 230-243.
- Ross, C.E., & Mirowsky, J. (1984). Socially desirable response and acquiescence in a cross-cultural survey of mental health. *Journal of Health and Social Behavior*, 25(2), 189–97. <https://doi.org/10.2307/2136668>
- Rust, K. (2013). Sampling, weighting, and variance estimation in international large-scale assessments. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (1st ed., pp. 117–154). New York, NY: Chapman and Hall/CRC.
- Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39(2), 142-151.
- Rutkowski, L., von Davier, M., & Rutkowski, D. (Eds.). (2014). *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis*. Boca Raton, FL: CRC.
- Schneider, S. (2018). Extracting response style bias from measures of positive and negative affect in aging research. *The Journals of Gerontology: Series B*, 73(1), 64-74. <https://doi.org/10.1093/geronb/gbw103>
- Schoonees, P. C., van de Velden, M., & Groenen, P. J. (2015). Constrained dual scaling for detecting response styles in categorical data. *Psychometrika*, 80, 968–994. <https://doi.org/10.1007/s11336-015-9458-9>

- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4), 591-611. <https://doi.org/10.2307/2333709>
- Shen, C., & 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. <https://doi.org/10.1080/13803610801896653>
- Silverstein, M., & Bengtson, V. L. (2008). *Longitudinal study of generations, 1971, 1985, 1988, 1991, 1994, 1997, 2000, 2005 [California]*. Ann Arbor, MI: Inter-University Consortium for Political and Social Research.
- Smith, P. B., & Fischer, R. (2008). Acquiescence, extreme response bias and culture: A multilevel analysis. In F. J. R. van de Vijver, D. A. van Hemert, & Y. H. Poortinga (Eds.), *Individuals and cultures in multilevel analysis* (pp. 285-314). Mahwah, New Jersey: Erlbaum.
- Smith, P.B. (2004). Acquiescent response bias as an aspect of cultural communication style. *Journal of Cross-Cultural Psychology*, 35(1), 50–61. <https://doi.org/10.1177/0022022103260380>
- Smith, T. W. (2003). Developing comparable questions in cross-national surveys. In J. A. Harkness, F. J. R. V. de Vijver, & P. P. Mohler (Eds.), *Cross-cultural survey methods* (pp. 69-92). Hoboken, New Jersey: Wiley.
- Sudman, S., & Bradburn, N. M. (1982). *Asking questions: A practical guide to questionnaire design*. San Francisco: Jossey-Bass.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th Ed.) Pearson.
- Takagishi, M.(2019). *ccrs: Correct and cluster response style biased data*. (R package version 0.1.0). Retrived from <https://rdrr.io/cran/ccrs/man/correct.rs.html>
- Takagishi, M., van de Velden, M., & Yadohisa, H. (2019). Clustering preference data in the presence of response-style bias. *British Journal of Mathematical and Statistical Psychology*, 72(3), 401-425.
- Takahashi, K., Ohara, N., Antonucci, T. C., & Akiyama, H. (2002). Commonalities and differences in close relationships among the Americans and Japanese: A comparison

- by the individualism/collectivism concept. *International Journal of Behavioral Development*, 26(5), 453-465. <https://doi.org/10.1080/01650250143000418>
- Tat, O., Koyuncu, İ., & Gelbal, S. (2019). Makul Değer ve Ağırlıklandırma kullanımının doğrusal regresyon ve hiyerarşik doğrusal model parametrelerine etkisi. *Psychology*, 10(3), 235-248. <https://doi: 10.21031/epod.486999>
- Timoçin, G. (2019). *Uluslararası öğrenci değerlendirme programına (PISA) göre Türkiye'deki öğrencilerin fen okuryazarlık puanlarının çok düzeyli modeller ile incelenmesi*. (Yüksek Lisans Tezi). <https://tez.yok.gov.tr.sayfasından> erişilmiştir.
- Tsekouras, D. (2017). The effect of rating scale design on extreme response tendency in consumer product ratings. *International Journal of Electronic Commerce*, 21(2), 270-296. <https://doi.org/10.1080/10864415.2016.1234290>
- Usami, S. (2014). Constrained k-means on cluster proportion and distances among clusters for longitudinal data analysis. *Japanese Psychological Research*, 56(4), 361-372. <https://doi.org/10.1111/jpr.12060>
- Van de gaer, E., Grisay, A., Schulz, W., & 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. <https://doi.org/10.1177/0022022111428083>
- van de Gaer, E.,& Adams,R.(2010, April). *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.
- van de Velden, M. (2007). *Detecting response styles by using dual scaling of successive categories* (Research No. EI 2007-41). Report/Econometric Institute, Erasmus University Rotterdam. Retrieved from <http://hdl.handle.net/1765/10561>
- van Dijk, T. K., Datema, F., Welten, S. C. M., & van de Vijver, F. J. R. (2009). Acquiescence and extremity in cross-national surveys: Domain dependence and country-level correlates. In G. Aikaterini & K. Mylonas (Eds.), *Quod Erat Demonstrandum: From Herodotus' ethnographic journeys to cross-cultural research: Proceedings from the 18th International Congress of the International Association for Cross-Cultural Psychology*. Retrieved from https://scholarworks.gvsu.edu/iaccp_papers/51/

- Van Hemert, D. A., Van de Vijver, F. J., Poortinga, Y. H., & Georgas, J. (2002). Structural and functional equivalence of the Eysenck Personality Questionnaire within and between countries. *Personality and Individual Differences*, 33(8), 1229-1249. [https://doi.org/10.1016/S0191-8869\(02\)00007-7](https://doi.org/10.1016/S0191-8869(02)00007-7)
- Van Herk, H., Poortinga, Y. H., & Verhallen, T. M. (2004). Response styles in rating scales: Evidence of method bias in data from six EU countries. *Journal of Cross-Cultural Psychology*, 35(3), 346-360. <https://doi.org/10.1177/0022022104264126>
- Van Rosmalen, J., Van Herk, H., & Groenen, P. J. (2010). Identifying response styles: A latent-class bilinear multinomial logit model. *Journal of Marketing Research*, 47(1), 157-172. <https://doi.org/10.1509/jmkr.47.1.157>
- Van Vaerenbergh, Y., & Thomas, T. D. (2013). Response styles in survey research: A literature review of antecedents, consequences, and remedies. *International Journal of Public Opinion Research*, 25(2), 195-217. <https://doi.org/10.1093/ijpor/eds021>
- Von Davier, M., Gonzalez, E., & Mislevy, R. (2009). What are plausible values and why are they useful. *IERI Monograph Series*, 2(1), 9-36.
- Weijters, B., Geuens, M., & Schillewaert, N. (2010). The stability of individual response styles. *Psychological Methods*, 15 (1), 96–110. <https://doi.org/10.1037/a0018721>
- Weijters, B., Schillewaert, N., & Geuens, M. (2008). Assessing response styles across modes of data collection. *Journal of the Academy of Marketing Science*, 36(3), 409-422.
- Weinzierl, S., Lepa, S., & Ackermann, D. (2018). A measuring instrument for the auditory perception of rooms: The Room Acoustical Quality Inventory (RAQI). *The Journal of the Acoustical Society of America*, 144(3), 1245-1257. <https://doi.org/10.1121/1.5051453>
- Welkenhuysen-Gybels, J., Billiet, J., & Cambré, B. (2003). Adjustment for acquiescence in the assessment of the construct equivalence of Likert-type score items. *Journal of Cross-Cultural Psychology*, 34(6), 702-722. <https://doi.org/10.1177/0022022103257070>
- Wetzel, E., & Carstensen, C. H. (2015). Multidimensional modeling of traits and response styles. *European Journal of Psychological Assessment*.33(5) 352-364.

- Wetzel, E., Böhnke, J. R., & Rose, N. (2016). A simulation study on methods of correcting for the effects of extreme response style. *Educational and Psychological Measurement*, 76(2), 304-324. <https://doi.org/10.1177/0013164415591848>
- Wetzel, E., Böhnke, J. R., Carstensen, C. H., Ziegler, M., & Ostendorf, F. (2013). Do individual response styles matter?. *Journal of Individual Differences*, 34(2), 69-81.
- Wetzel, E., Carstensen, C. H., & Böhnke, J. R. (2013). Consistency of extreme response style and non-extreme response style across traits. *Journal of Research in Personality*, 47(2), 178-189. <https://doi.org/10.1016/j.jrp.2012.10.010>



EKLER



Ek 1. Fene Yönelik Tutum Ölçekleri ve Maddeleri

Fenden Hoşlanma (Enjoyment of Science); (ST094Q01NA - ST094Q05NA) 5 Madde

- | | |
|---|---------------------------|
| 1. Fen bilimleri ile ilgili konuları öğrenirken genellikle eğlenirim. | 1-Kesinlikle Katılmıyorum |
| 2. Fen bilimleri ile ilgili yazıları okumaktan zevk alırım. | 2-Katılmıyorum |
| 3. Fen bilimleri ile ilgili konularda çalışmak beni mutlu ediyor. | 3-Katılıyorum |
| 4. Fen bilimleri hakkında yeni bilgiler edinmekten zevk alırım. | 4-Kesinlikle Katılıyorum |
| 5. Fen bilimleri ile ilgili bir şeyler öğrenmeye ilgi duyarım. | |

Başarı Motivasyonu(Achievement Motivation);(ST119Q01NA-ST119Q05NA) 5 Madde

- | | |
|--|---------------------------|
| 1. Tüm derslerimden ya da derslerimin çoğundan yüksek notlar almak isterim. | 1-Kesinlikle Katılmıyorum |
| 2. Mezun olduğumda uygun olan en iyi fırsatlar arasından seçim yapabilmek isterim. | 2-Katılmıyorum |
| 3. Her ne yaparsam yapayım, en iyisi olmak isterim. | 3-Katılıyorum |
| 4. Kendimi hırslı biri olarak görüyorum. | 4-Kesinlikle Katılıyorum |
| 5. Sınıfımın en iyi öğrencilerinden biri olmak isterim. | |

***Fen Öz Yeterliği (Science Self-Efficacy) ; (ST129Q01TA-ST129Q08TA) 8 Madde**

- | | |
|--|--|
| 1. Bir sağlık sorununu ele alan gazete raporunun temelinde yatan, fen bilimleri ile ilgili soruyu tanıma. | |
| 2. Bazı alanlardaki depremlerin diğer alanlardan niçin daha sık meydana geldiğini açıklama. | 1-Bunu kolaylıkla yapabilirim. |
| 3. Antibiyotiklerin hastalıkların tedavisindeki rolünü belirtme. | 2-Bunu uğraşırsam yapabilirim. |
| 4. Fen bilimleri açısından, atıkların uygun şekilde toplanarak işlenmesi ile ilgili sorunu belirleme. | 3-Bunu yapabilmek için çok çaba göstermem gerekir. |
| 5. Çevredeki değişmelerin belli canlı türlerinin hayatta kalmasını nasıl etkileyebileceğini önceden tahmin etme. | 4-Bunu yapamam. |
| 6. Gıda maddelerinin üzerine konan etiketlerde verilen bilimsel bilgileri yorumlama. | |
| 7. Yeni kanıtların, Mars'ta hayat olup olmadığına ilişkin anlayışınızı nasıl değiştirebileceğini tartışma. | |
| 8. Asit yağmurlarının nasıl oluştuğunu açıklayan iki görüşten hangisinin daha iyi olduğunu belirleme. | |

* Ters Kodlama yapılmıştır.

Ek 2. Araştırmada Yer Alan Ülkelerin Öğrenci Sayıları ve Silinen Kayıp Veri Sayıları

Ülkeler	Örneklem	Silinen Kayıp Veri Miktarı	Kalan Veri
Avusturya (AUT)	7.007	2.070	4.937
Belçika (BEL)	9.651	2.946	6.705
Brezilya (BRA)	23.141	14.240	8.901
Hırvatistan (HRV)	5.809	1.296	4.513
Danimarka (DNK)	7.161	1.964	5.197
Dominik Cumhuriyeti (DOM)	4.740	2.404	2.336
Finlandiya (FIN)	5.882	1.164	4.718
Fransa (FRA)	6.108	1.786	4.322
Almanya (DEU)	6.504	3.914	2.590
İsrail (ISR)	6.598	1.856	4.742
İtalya (ITA)	11.583	2.664	8.919
Japonya (JPN)	6.647	813	5.834
Letonya (LVA)	4.869	945	3.924
Lüksemburg (LUX)	5.299	1.648	3.651
Meksika (MEX)	7.568	1.970	5.598
Hollanda (NLD)	5.385	1.029	4.356
Peru (PER)	6.971	2.332	4.639
Katar (QAT)	12.083	4.575	7.508
Rusya Federasyonu (RUS)	6.036	1.547	4.489
Birleşik Arap Emirlikleri (ARE)	14.167	4.002	10.165
Türkiye (TUR)	5.895	1.179	4.716
Toplam	112.760	56.344	112.760

Ek 3. Ülke Düzeyinde Örneklem ve Çalışma Örnekleme Fene Yönelik Tutum Ölçekleri Ortalama Puanları

Ülkeler	Fenden Hoşlanma		Başarı Motivasyonu		Fen Öz Yeterliği	
	Örneklem	Çalışma Örnekleme	Örneklem	Çalışma Örnekleme	Örneklem	Çalışma Örnekleme
Avusturya	2,38	2,36	2,88	2,88	2,63	2,65
Belçika	2,63	2,67	2,73	2,76	2,69	2,71
Brezilya	2,81	2,82	3,10	3,09	2,82	2,82
Hırvatistan	2,56	2,60	2,92	2,95	2,80	2,77
Danimarka	2,72	2,72	2,97	2,98	2,79	2,85
Dominik Cumh.	3,02	3,00	3,17	3,19	3,01	2,99
Finlandiya	2,58	2,58	2,62	2,60	2,72	2,74
Fransa	2,61	2,61	2,89	2,91	2,67	2,63
Almanya	2,50	2,47	2,81	2,81	2,73	2,75
İsrail	2,71	2,69	3,55	3,55	2,75	2,77
İtalya	2,64	2,64	2,96	2,97	2,81	2,83
Japonya	2,36	2,40	2,70	2,71	2,48	2,49
Letonya	2,71	2,74	3,04	3,06	2,73	2,77
Lüksemburg	2,70	2,63	2,93	2,90	2,72	2,74
Meksika	2,93	2,96	3,14	3,18	2,88	2,88
Hollanda	2,25	2,24	2,78	2,75	2,70	2,74
Peru	2,93	2,95	3,21	3,22	2,94	2,92
Katar	2,91	2,90	3,51	3,49	2,94	2,94
Rusya Federasyonu	2,64	2,65	3,04	3,01	2,71	2,76
Birleşik Arap Emir.	2,98	2,99	3,52	3,55	2,96	2,97
Türkiye	2,72	2,72	3,39	3,40	2,93	2,90

Ek 4. Her Bir Ülke için Fene Yönelik Tutum Ölçekleri Cronbach Alfa Güvenirlik Katsayısı Değerleri

Ülkeler	Cronbach Alfa Güvenirlik Katsayıları		
	Fenden Hoşlanma Ölçeği	Başarı Motivasyonu Ölçeği	Fen Öz Yeterliği Ölçeği
Avusturya	,94	,80	,87
Belçika	,94	,81	,88
Brezilya	,92	,68	,90
Hırvatistan	,94	,74	,86
Danimarka	,96	,87	,87
Dominik Cumhur.	,93	,76	,90
Finlandiya	,94	,83	,89
Fransa	,93	,78	,88
Almanya	,94	,81	,87
İsrail	,95	,83	,86
İtalya	,92	,79	,85
Japonya	,95	,86	,90
Letonya	,93	,81	,82
Lüksemburg	,94	,83	,91
Meksika	,89	,71	,88
Hollanda	,95	,77	,89
Peru	,93	,69	,85
Katar	,93	,87	,89
Rusya Federasyonu	,93	,82	,89
Birleşik Arap Emirlikleri	,94	,83	,89
Türkiye	,95	,85	,90

Ek 5. Fene Yönelik Tutum Ölçeklerinin Fark Puanlarına İlişkin Betimsel İstatistikler

PISA 2015'e Katılan Ülkelerin Fenden Hoşlanma Ölçeği için Fark Puanlarına İlişkin Betimsel İstatistikler

Ülkeler	İstatistik	Değeri	Standart Hata
Avusturya	Ortalama	,032	
	Std.Sapma	,239	
	Medyan	,032	
	Mod	,410	
	Çarpıklık Katsayısı	-,044	,113
	Basıklık Katsayısı	-,672	,225
Belçika	Ortalama	-,017	
	Std.Sapma	,217	
	Medyan	-,030	
	Mod	-,080	
	Çarpıklık Katsayısı	,275	,096
	Basıklık Katsayısı	-,085	,191
Brezilya	Ortalama	-,061	
	Std.Sapma	,179	
	Medyan	-,079	
	Mod	-,080	
	Çarpıklık Katsayısı	,301	,082
	Basıklık Katsayısı	1,107	,164
Hırvatistan	Ortalama	-,017	
	Std.Sapma	,189	
	Medyan	-,030	
	Mod	-,080	
	Çarpıklık Katsayısı	,229	,118
	Basıklık Katsayısı	,680	,236
Danimarka	Ortalama	-,046	
	Std.Sapma	,208	
	Medyan	-,079	
	Mod	-,080	
	Çarpıklık Katsayısı	,303	,109
	Basıklık Katsayısı	,270	,217
Dominik Cumh.	Ortalama	-,117	
	Std.Sapma	,210	
	Medyan	-,079	
	Mod	-,080	
	Çarpıklık Katsayısı	,656	,164
	Basıklık Katsayısı	,331	,327
Finlandiya	Ortalama	-,010	
	Std.Sapma	,177	
	Medyan	-,007	
	Mod	-,080	
	Çarpıklık Katsayısı	,129	,116
	Basıklık Katsayısı	1,052	,231

PISA 2015'e Katılan Ülkelerin Fenden Hoşlanma Ölçeği için Fark Puanlarına İlişkin
Betimsel İstatistikler Devamı

Ülkeler	İstatistik	Değeri	Standart Hata
Fransa	Ortalama	-,011	
	Std.Sapma	,214	
	Medyan	-,030	
	Mod	-,080	
	Çarpıklık Katsayısı	,318	,121
	Basıklık Katsayısı	-,160	,241
Almanya	Ortalama	,0163	
	Std.Sapma	,23328	
	Medyan	,0088	
	Mod	,090	
	Çarpıklık Katsayısı	,039	,156
	Basıklık Katsayısı	-,546	,311
İsrail	Ortalama	-,0395	
	Std.Sapma	,23616	
	Medyan	-,0757	
	Mod	-,080	
	Çarpıklık Katsayısı	,276	,115
	Basıklık Katsayısı	-,460	,230
İtalya	Ortalama	-,015	
	Std.Sapma	,183	
	Medyan	-,030	
	Mod	-,080	
	Çarpıklık Katsayısı	,423	,082
	Basıklık Katsayısı	,711	,164
Japonya	Ortalama	,028	
	Std.Sapma	,224	
	Medyan	,032	
	Mod	,090	
	Çarpıklık Katsayısı	-,055	,102
	Basıklık Katsayısı	-,305	,204
Letonya	Ortalama	-,045	
	Std.Sapma	,176	
	Medyan	-,071	
	Mod	-,080	
	Çarpıklık Katsayısı	,209	,126
	Basıklık Katsayısı	1,077	,251
Lüksemburg	Ortalama	-,021	
	Std.Sapma	,224	
	Medyan	-,030	
	Mod	-,080	
	Çarpıklık Katsayısı	,225	,131
	Basıklık Katsayısı	-,320	,260

PISA 2015'e Katılan Ülkelerin Fenden Hoşlanma Ölçeği için Fark Puanlarına İlişkin
Betimsel İstatistikler Devamı

Ülkeler	İstatistik	Değeri	Standart Hata
Meksika	Ortalama	-,103	
	Std.Sapma	,173	
	Medyan	-,079	
	Mod	-,080	
	Çarpıklık Katsayısı	,359	,104
	Basıklık Katsayısı	,578	,208
Hollanda	Ortalama	,072	
	Std.Sapma	,200	
	Medyan	,038	
	Mod	,090	
	Çarpıklık Katsayısı	,159	,120
	Basıklık Katsayısı	-,189	,240
Peru	Ortalama	-,099	
	Std.Sapma	,186	
	Medyan	-,079	
	Mod	-,080	
	Çarpıklık Katsayısı	,342	,116
	Basıklık Katsayısı	,613	,232
Katar	Ortalama	-,089	
	Std.Sapma	,199	
	Medyan	-,079	
	Mod	-,080	
	Çarpıklık Katsayısı	,361	,090
	Basıklık Katsayısı	,355	,180
Rusya Federasyonu	Ortalama	-,0220	
	Std.Sapma	,16634	
	Medyan	-,0309	
	Mod	-,080	
	Çarpıklık Katsayısı	,288	,119
	Basıklık Katsayısı	1,666	,237
Birleşik Arap Emir.	Ortalama	-,100	
	Std.Sapma	,208	
	Medyan	-,079	
	Mod	-,080	
	Çarpıklık Katsayısı	,431	,077
	Basıklık Katsayısı	,241	,154
Türkiye	Ortalama	-,050	
	Std.Sapma	,216	
	Medyan	-,079	
	Mod	-,080	
	Çarpıklık Katsayısı	,386	,116
	Basıklık Katsayısı	,089	,231

PISA 2015'e Katılan Ülkelerin Başarı Motivasyonu Ölçeği için Fark Puanlarına İlişkin Betimsel İstatistikler

Ülkeler	İstatistik	Değeri	Standart Hata
Avusturya	Ortalama	-,090	
	Std.Sapma	,178	
	Medyan	-,076	
	Mod	-,400	
	Çarpıklık Katsayısı	,106	,113
	Basıklık Katsayısı	-,391	,225
Belçika	Ortalama	-,051	
	Std.Sapma	,161	
	Medyan	-,024	
	Mod	,010	
	Çarpıklık Katsayısı	-,197	,096
	Basıklık Katsayısı	,286	,191
Brezilya	Ortalama	-,138	
	Std.Sapma	,142	
	Medyan	-,120	
	Mod	-,400	
	Çarpıklık Katsayısı	-,058	,082
	Basıklık Katsayısı	-,577	,164
Hırvatistan	Ortalama	-,102	
	Std.Sapma	,142	
	Medyan	-,076	
	Mod	-,147	
	Çarpıklık Katsayısı	-,241	,118
	Basıklık Katsayısı	-,300	,236
Danimarka	Ortalama	-,118	
	Std.Sapma	,181	
	Medyan	-,076	
	Mod	-,400	
	Çarpıklık Katsayısı	,190	,109
	Basıklık Katsayısı	-,115	,217
Dominik Cumh.	Ortalama	-,170	
	Std.Sapma	,162	
	Medyan	-,185	
	Mod	-,241	
	Çarpıklık Katsayısı	1,256	,164
	Basıklık Katsayısı	2,678	,327
Finlandiya	Ortalama	-,010	
	Std.Sapma	,168	
	Medyan	,007	
	Mod	-,076	
	Çarpıklık Katsayısı	-,170	,116
	Basıklık Katsayısı	,603	,231

PISA 2015'e Katılan Ülkelerin Başarı Motivasyonu Ölçeği için Fark Puanlarına İlişkin
Betimsel İstatistikler Devamı

Ülkeler	İstatistik	Değeri	Standart Hata
Fransa	Ortalama	-,098	
	Std.Sapma	,157	
	Medyan	-,062	
	Mod	-,400	
	Çarpıklık Katsayısı	-,352	,121
	Basıklık Katsayısı	-,562	,241
Almanya	Ortalama	-,071	
	Std.Sapma	,175	
	Medyan	-,062	
	Mod	-,247	
	Çarpıklık Katsayısı	,233	,156
	Basıklık Katsayısı	,058	,311
İsrail	Ortalama	-,273	
	Std.Sapma	,151	
	Medyan	-,309	
	Mod	-,400	
	Çarpıklık Katsayısı	1,717	,115
	Basıklık Katsayısı	4,076	,230
İtalya	Ortalama	-,079	
	Std.Sapma	,160	
	Medyan	-,062	
	Mod	-,400	
	Çarpıklık Katsayısı	,013	,082
	Basıklık Katsayısı	,180	,164
Japonya	Ortalama	-,0420928	
	Std.Sapma	,19074032	
	Medyan	-,0240339	
	Mod	-,40048	
	Çarpıklık Katsayısı	-,062	,102
	Basıklık Katsayısı	-,271	,204
Letonya	Ortalama	-,128	
	Std.Sapma	,166	
	Medyan	-,118	
	Mod	-,400	
	Çarpıklık Katsayısı	,307	,126
	Basıklık Katsayısı	,161	,251
Lüksemburg	Ortalama	-,095	
	Std.Sapma	,181	
	Medyan	-,076	
	Mod	-,400	
	Çarpıklık Katsayısı	,207	,131
	Basıklık Katsayısı	-,052	,260

PISA 2015'e Katılan Ülkelerin Başarı Motivasyonu Ölçeği için Fark Puanlarına İlişkin
Betimsel İstatistikler Devamı

Ülkeler	İstatistik	Değeri	Standart Hata
Meksika	Ortalama	-,162	
	Std.Sapma	,143	
	Medyan	-,163	
	Mod	-,163	
	Çarpıklık Katsayısı	,801	,104
	Basıklık Katsayısı	2,007	,208
Hollanda	Ortalama	-,051	
	Std.Sapma	,132	
	Medyan	,002	
	Mod	,010	
	Çarpıklık Katsayısı	-,580	,120
	Basıklık Katsayısı	,879	,240
Peru	Ortalama	-,178	
	Std.Sapma	,138	
	Medyan	-,185	
	Mod	-,400	
	Çarpıklık Katsayısı	,340	,116
	Basıklık Katsayısı	,171	,232
Katar	Ortalama	-,252	
	Std.Sapma	,166	
	Medyan	-,294	
	Mod	-,400	
	Çarpıklık Katsayısı	1,459	,090
	Basıklık Katsayısı	2,616	,180
Rusya Federasyonu	Ortalama	-,120	
	Std.Sapma	,154	
	Medyan	-,076	
	Mod	-,076	
	Çarpıklık Katsayısı	-,154	,119
	Basıklık Katsayısı	,163	,237
Birleşik Arap Emir.	Ortalama	-,260	
	Std.Sapma	,145	
	Medyan	-,294	
	Mod	-,400	
	Çarpıklık Katsayısı	1,059	,077
	Basıklık Katsayısı	1,092	,154
Türkiye	Ortalama	-,223	
	Std.Sapma	,170	
	Medyan	-,278	
	Mod	-,400	
	Çarpıklık Katsayısı	1,206	,116
	Basıklık Katsayısı	1,740	,231

PISA 2015’e Katılan Ülkelerin Fen Öz Yeterliği Ölçeği için Fark Puanlarına İlişkin Betimsel İstatistikler

Ülkeler	İstatistik	Değeri	Standart Hata
Avusturya	Ortalama	,042	
	Std.Sapma	,199	
	Medyan	,059	
	Mod	,117	
	Çarpıklık Katsayısı	-,423	,113
	Basıklık Katsayısı	,005	,225
Belçika	Ortalama	,049	
	Std.Sapma	,208	
	Medyan	,060	
	Mod	,117	
	Çarpıklık Katsayısı	-,374	,096
	Basıklık Katsayısı	-,013	,191
Brezilya	Ortalama	,081	
	Std.Sapma	,208	
	Medyan	,117	
	Mod	,117	
	Çarpıklık Katsayısı	-,389	,082
	Basıklık Katsayısı	-,144	,164
Hırvatistan	Ortalama	,070	
	Std.Sapma	,186	
	Medyan	,085	
	Mod	,117	
	Çarpıklık Katsayısı	-,308	,118
	Basıklık Katsayısı	,185	,236
Danimarka	Ortalama	,088	
	Std.Sapma	,190	
	Medyan	,095	
	Mod	,432	
	Çarpıklık Katsayısı	-,294	,109
	Basıklık Katsayısı	,078	,217
Dominik Cumh.	Ortalama	,137	
	Std.Sapma	,200	
	Medyan	,117	
	Mod	,117	
	Çarpıklık Katsayısı	-,456	,164
	Basıklık Katsayısı	,012	,327
Finlandiya	Ortalama	,063	
	Std.Sapma	,192	
	Medyan	,069	
	Mod	,117	
	Çarpıklık Katsayısı	-,341	,116
	Basıklık Katsayısı	,112	,231

PISA 2015’e Katılan Ülkelerin Fen Öz Yeterliği Ölçeği için Fark Puanlarına İlişkin Betimsel İstatistikler Devamı

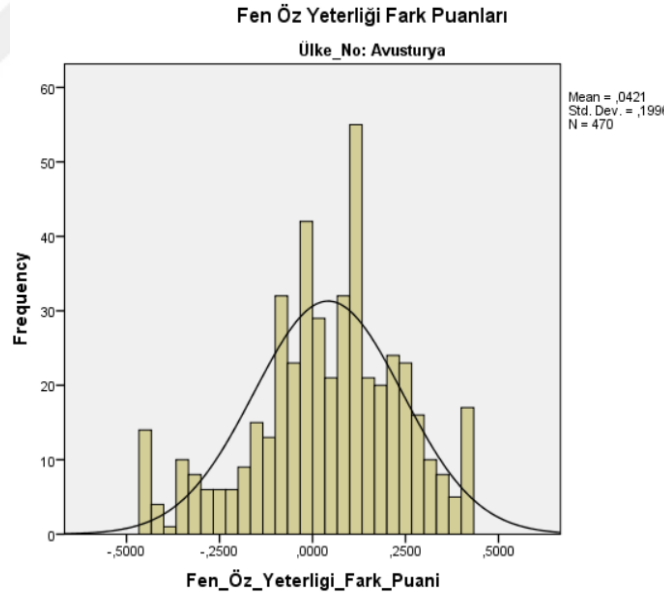
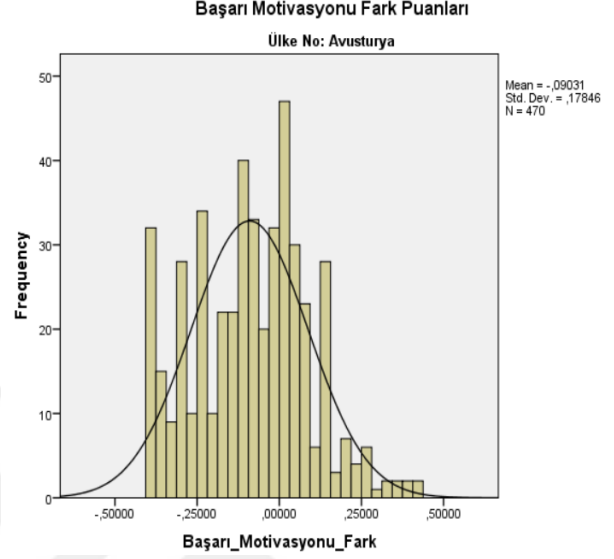
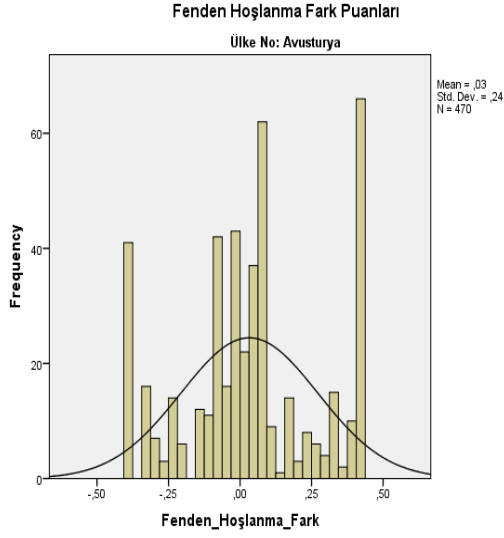
Ülkeler	İstatistik	Değeri	Standart Hata
Fransa	Ortalama	,028	
	Std.Sapma	,207	
	Medyan	,061	
	Mod	,117	
	Çarpıklık Katsayısı	-,398	,121
	Basıklık Katsayısı	-,160	,241
Almanya	Ortalama	,064	
	Std.Sapma	,191	
	Medyan	,068	
	Mod	,117	
	Çarpıklık Katsayısı	-,379	,156
	Basıklık Katsayısı	,113	,311
İsrail	Ortalama	,071	
	Std.Sapma	,199	
	Medyan	,085	
	Mod	,1170	
	Çarpıklık Katsayısı	-,371	,115
	Basıklık Katsayısı	,053	,230
İtalya	Ortalama	,075	
	Std.Sapma	,082	
	Medyan	,093	
	Mod	,117	
	Çarpıklık Katsayısı	-,296	,082
	Basıklık Katsayısı	,354	,164
Japonya	Ortalama	-,013	
	Std.Sapma	,182	
	Medyan	-,002	
	Mod	,117	
	Çarpıklık Katsayısı	-,341	,102
	Basıklık Katsayısı	,307	,204
Letonya	Ortalama	,068	
	Std.Sapma	,162	
	Medyan	,076	
	Mod	,117	
	Çarpıklık Katsayısı	-,164	,126
	Basıklık Katsayısı	,405	,251
Lüksemburg	Ortalama	,061	
	Std.Sapma	,217	
	Medyan	,070	
	Mod	,432	
	Çarpıklık Katsayısı	-,268	,131
	Basıklık Katsayısı	-,311	,260

PISA 2015’e Katılan Ülkelerin Fen Öz Yeterliği Ölçeği için Fark Puanlarına İlişkin Betimsel İstatistikler Devamı

Ülkeler	İstatistik	Değeri	Standart Hata
Meksika	Ortalama	,107	
	Std.Sapma	,171	
	Medyan	,117	
	Mod	,117	
	Çarpıklık Katsayısı	-,240	,104
	Basıklık Katsayısı	-,029	,208
Hollanda	Ortalama	,063	
	Std.Sapma	,203	
	Medyan	,085	
	Mod	,117	
	Çarpıklık Katsayısı	-,429	,120
	Basıklık Katsayısı	-,039	,240
Peru	Ortalama	,117	
	Std.Sapma	,165	
	Medyan	,117	
	Mod	,117	
	Çarpıklık Katsayısı	-,137	,116
	Basıklık Katsayısı	-,181	,232
Katar	Ortalama	,117	
	Std.Sapma	,196	
	Medyan	,117	
	Mod	,117	
	Çarpıklık Katsayısı	-,430	,090
	Basıklık Katsayısı	,156	,180
Rusya Federasyonu	Ortalama	,062	
	Std.Sapma	,207	
	Medyan	,099	
	Mod	,117	
	Çarpıklık Katsayısı	-,158	,119
	Basıklık Katsayısı	-,356	,237
Birleşik Arap Emir.	Ortalama	,122	
	Std.Sapma	,197	
	Medyan	,117	
	Mod	,117	
	Çarpıklık Katsayısı	-,521	,077
	Basıklık Katsayısı	,238	,154
Türkiye	Ortalama	,105	
	Std.Sapma	,193	
	Medyan	,117	
	Mod	,117	
	Çarpıklık Katsayısı	-,414	,116
	Basıklık Katsayısı	,290	,231

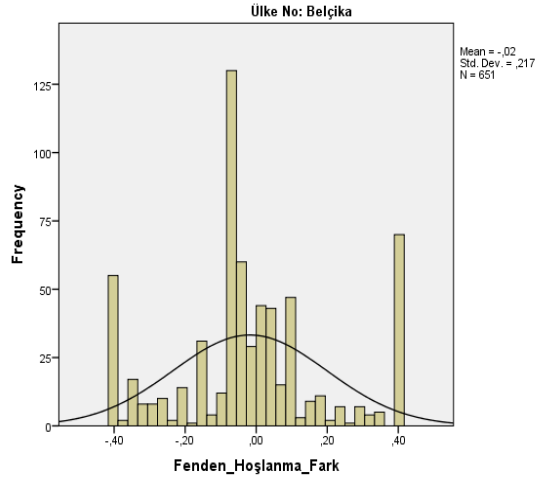
Ek 6. Fene Yönelik Tutum Ölçeklerinin Fark Puanlarına İlişkin Histogram Grafikleri

Avusturya

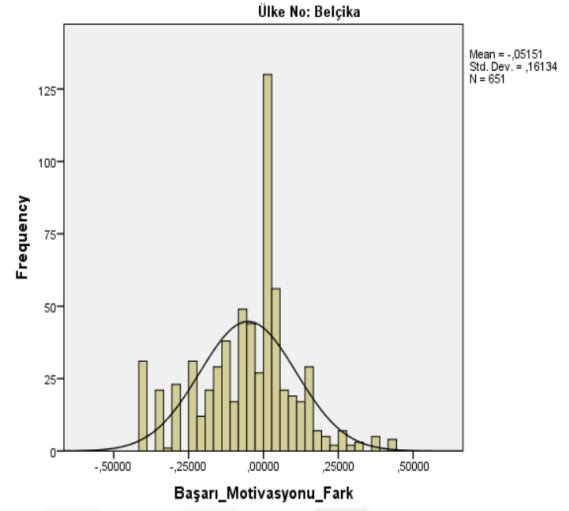


Belçika

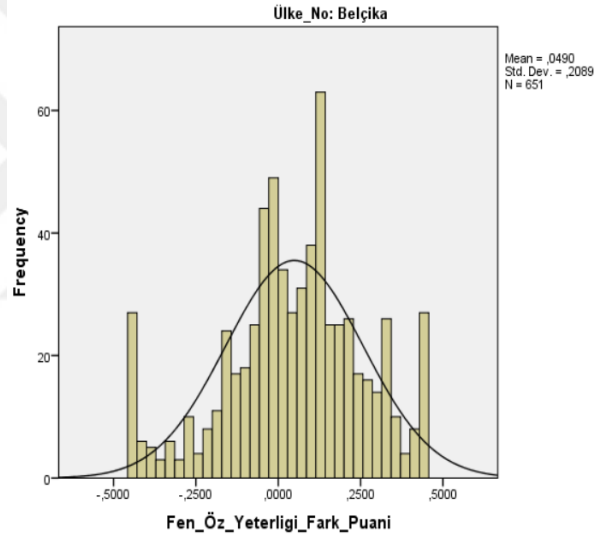
Fenden Hoşlanma Fark Puanları



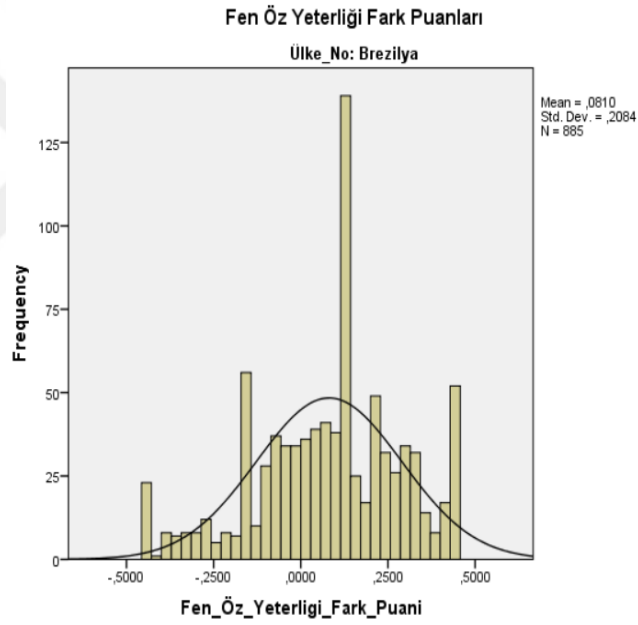
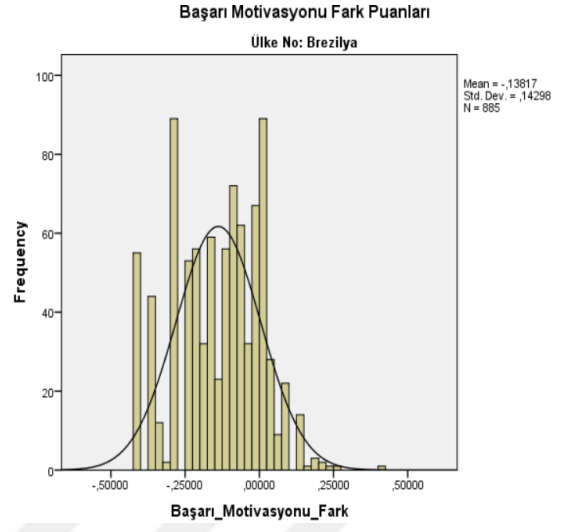
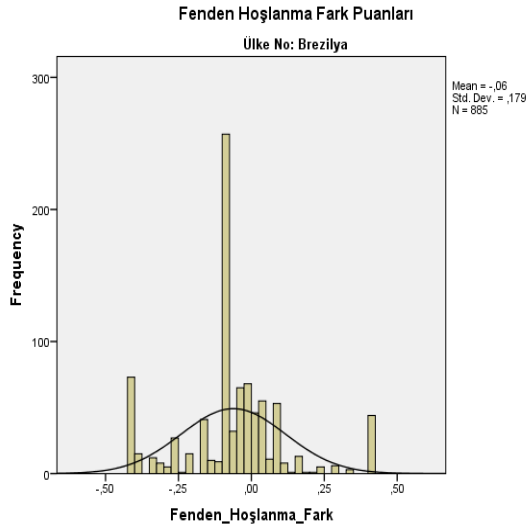
Başarı Motivasyonu Fark Puanları



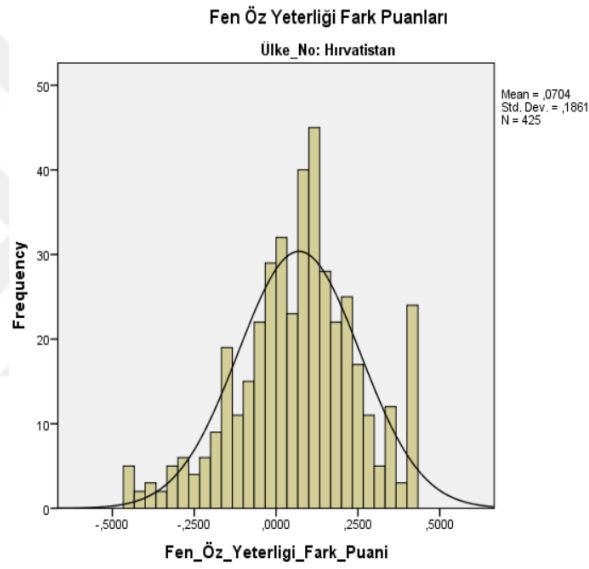
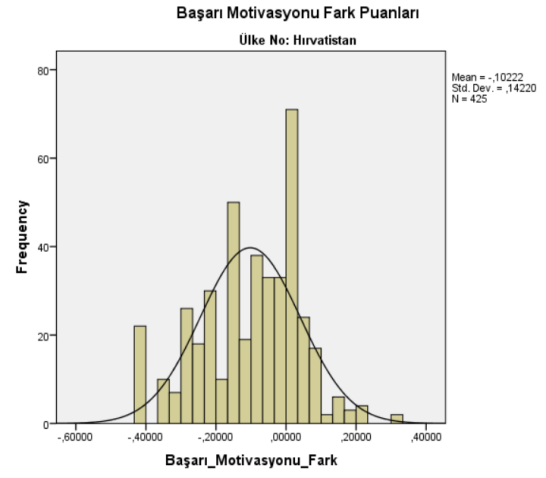
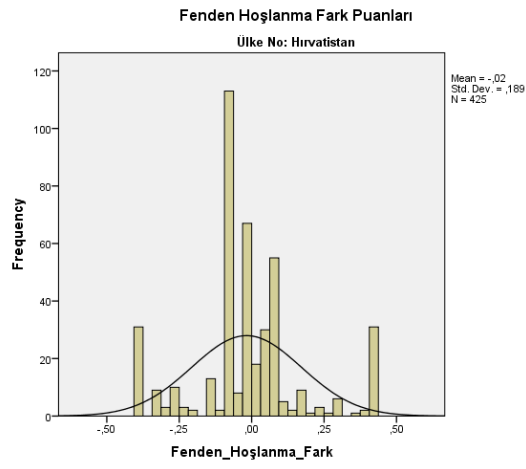
Fen Öz Yeterliği Fark Puanları



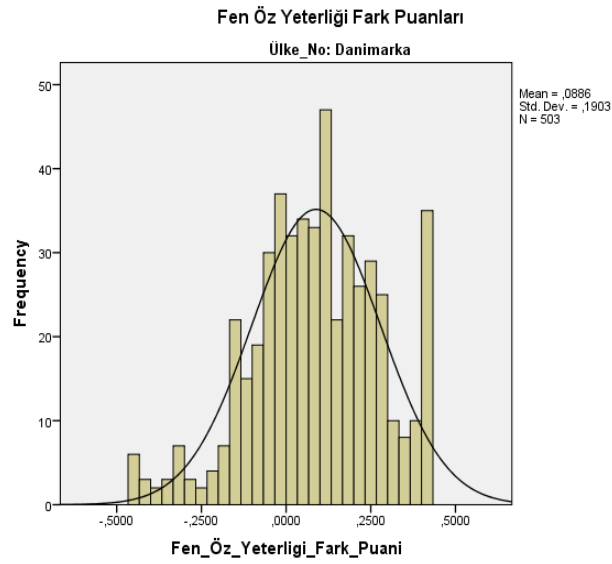
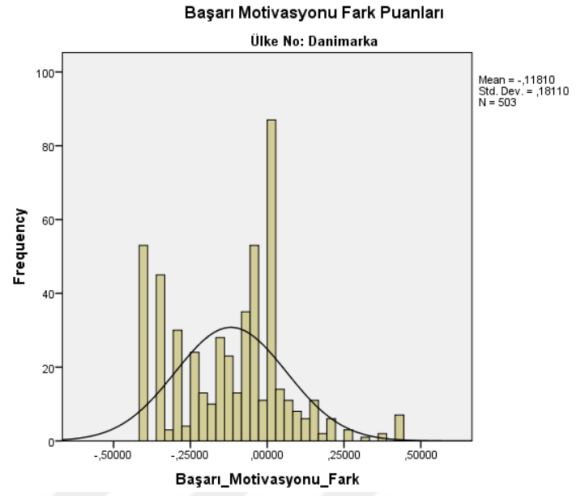
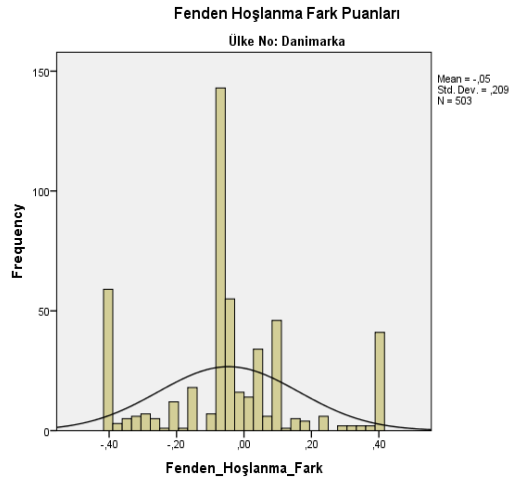
Brezilya



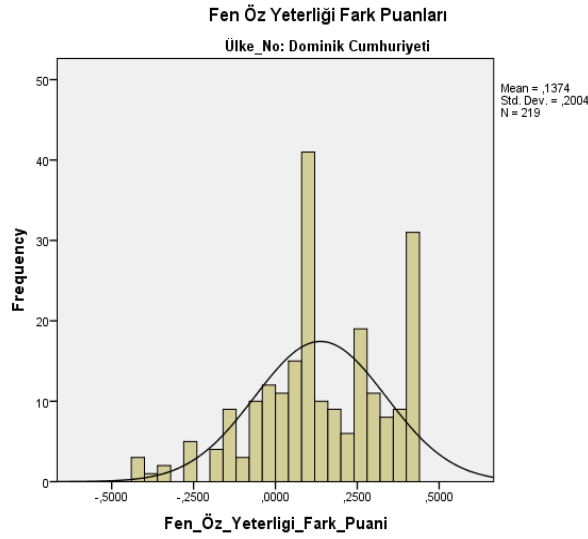
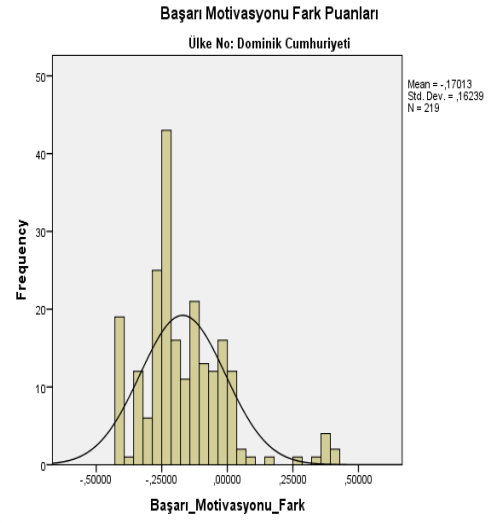
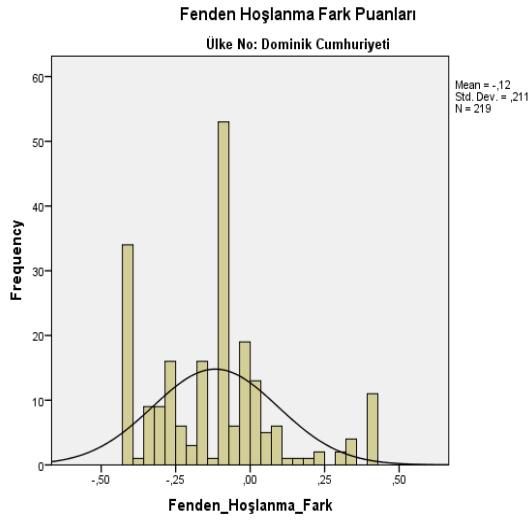
Hırvatistan



Danimarka

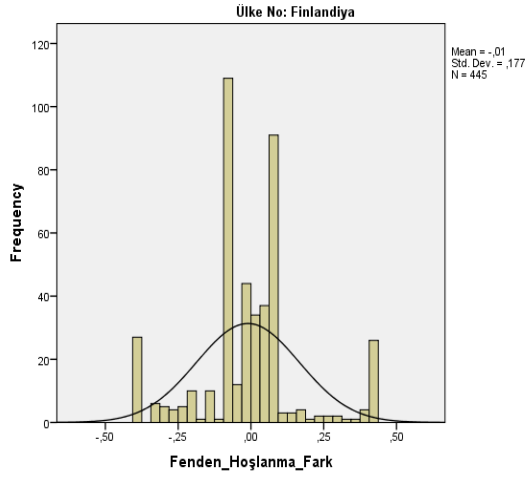


Dominik Cumhuriyeti

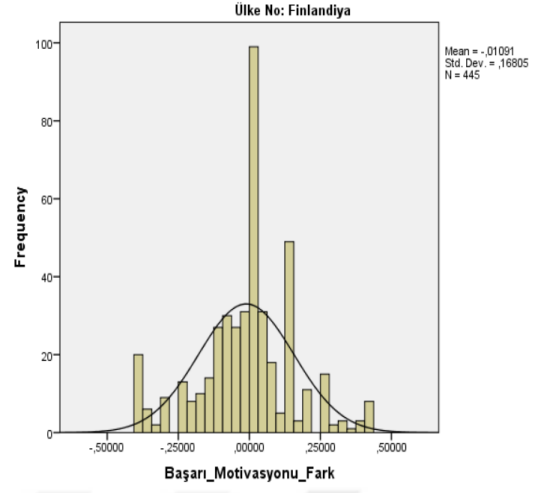


Finlandiya

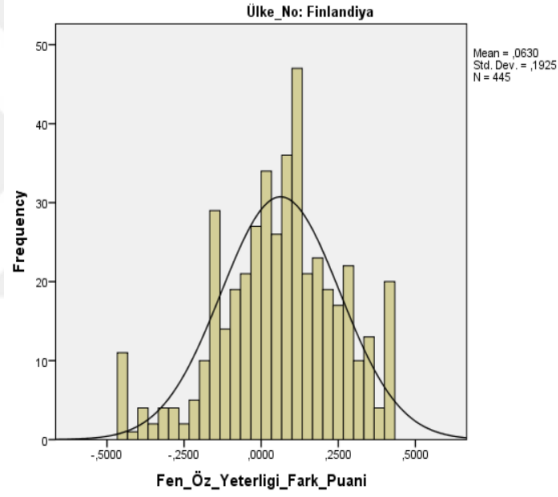
Fenden Hoşlanma Fark Puanları



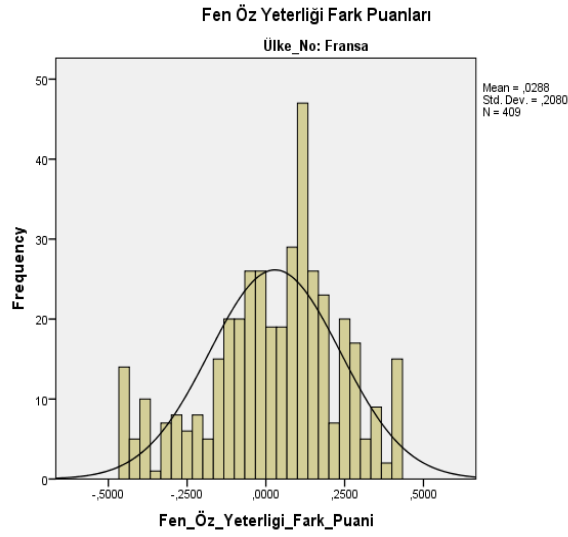
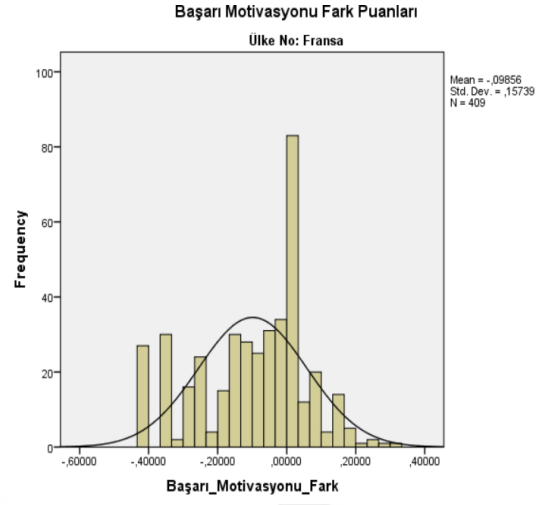
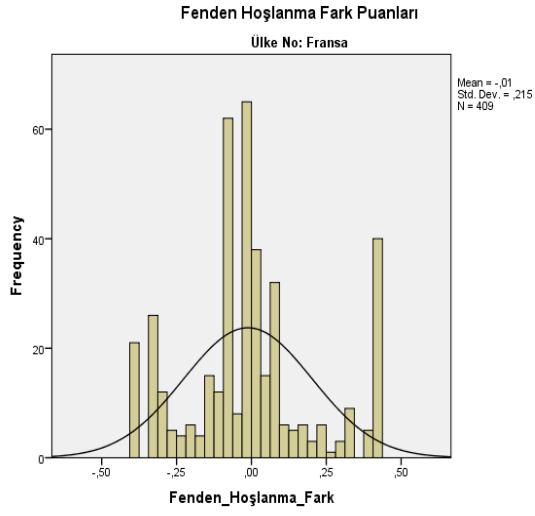
Başarı Motivasyonu Fark Puanları



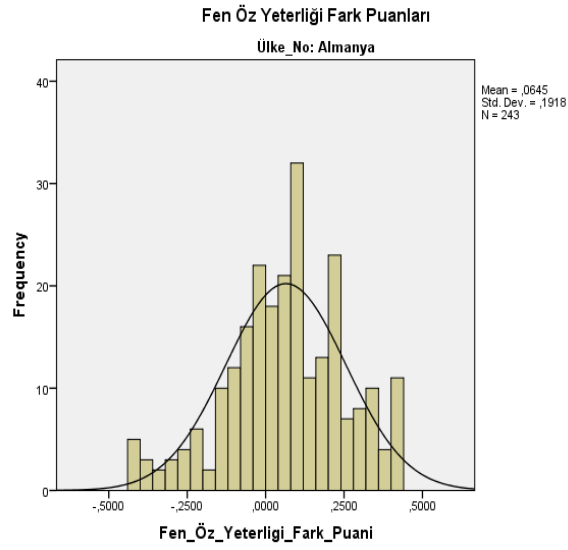
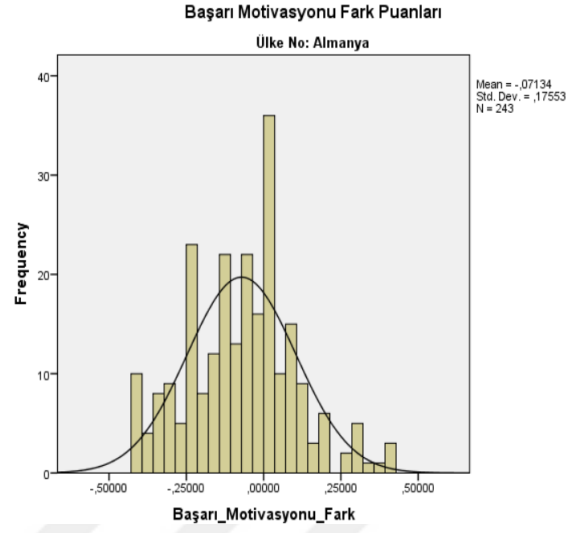
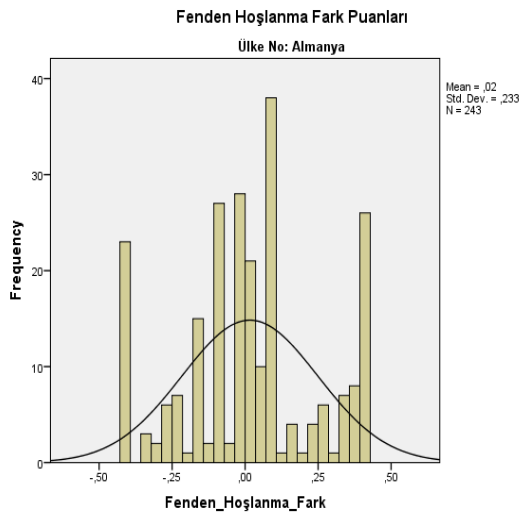
Fen Öz Yeterliği Fark Puanları



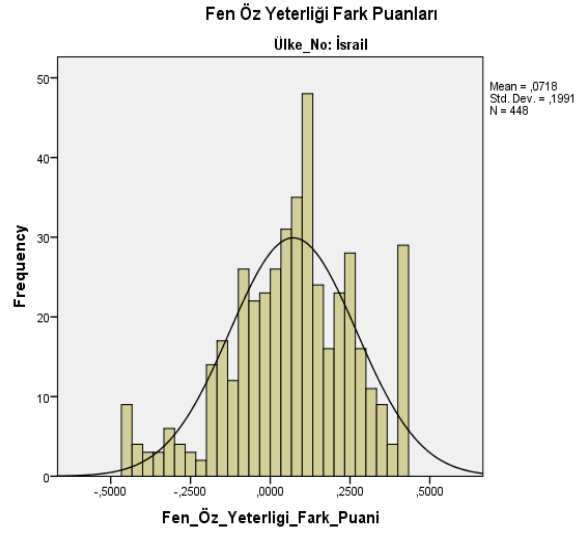
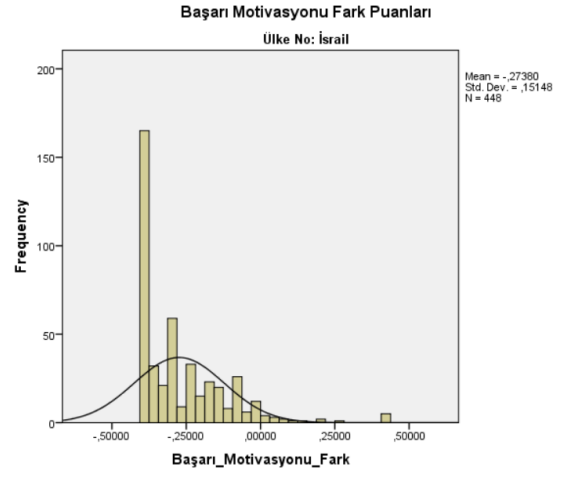
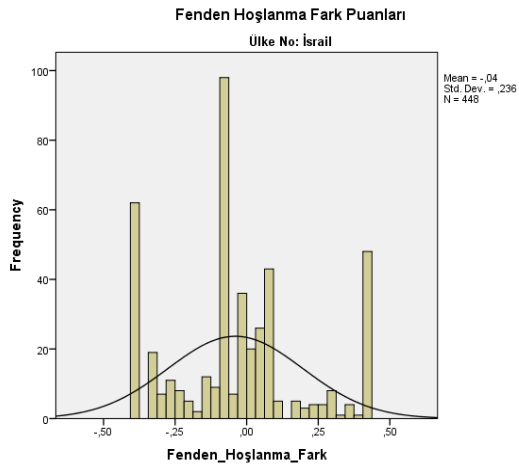
Fransa



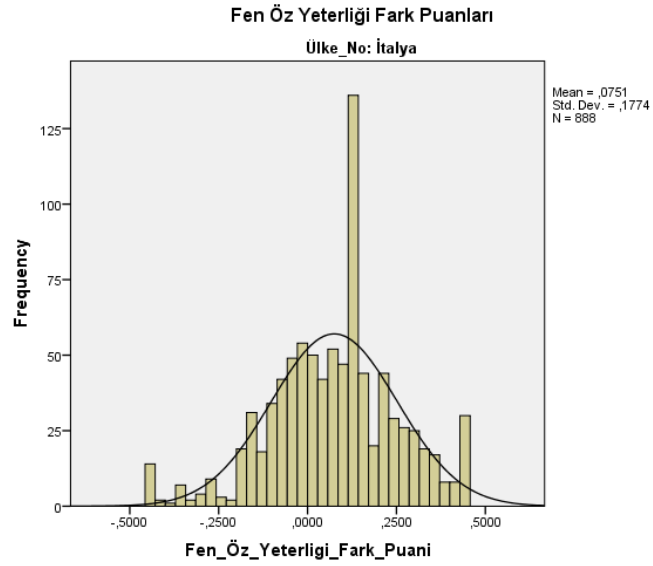
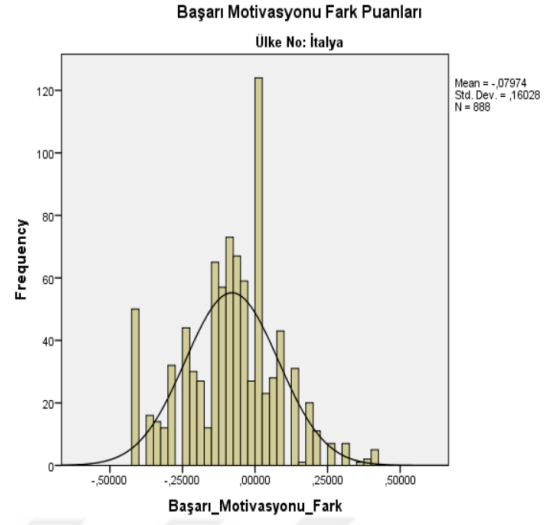
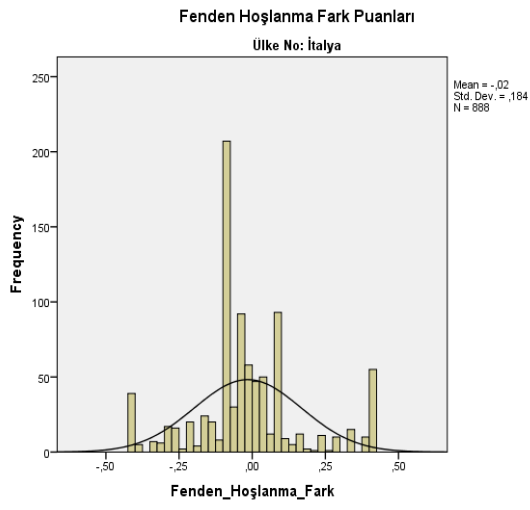
Almanya



İsrail

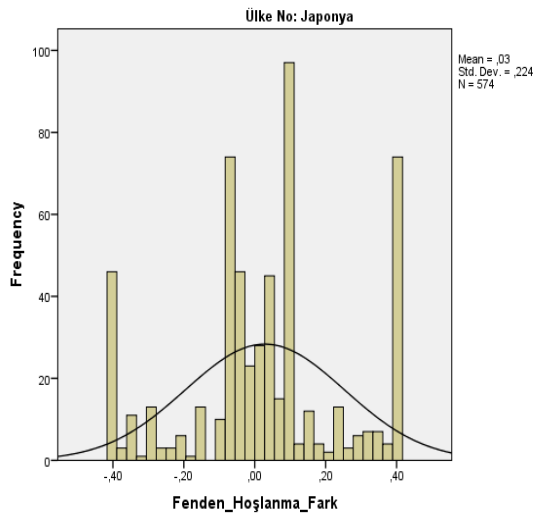


İtalya

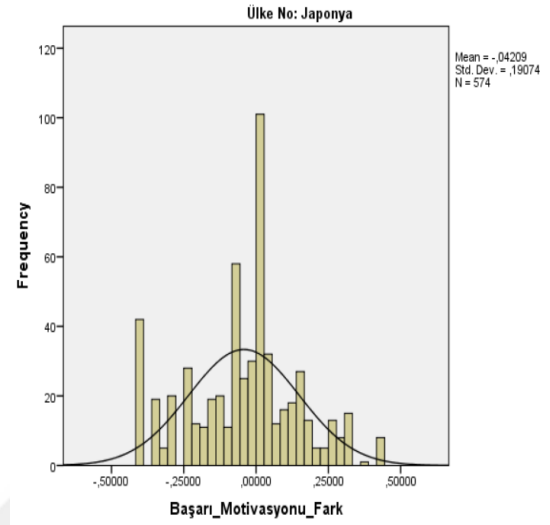


Japonya

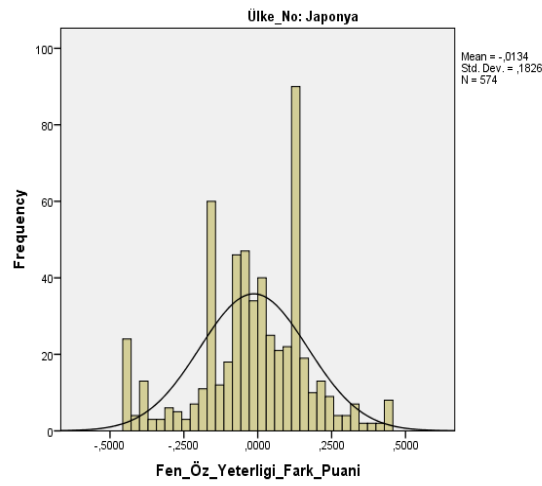
Fenden Hoşlanma Fark Puanları



Başarı Motivasyonu Fark Puanları

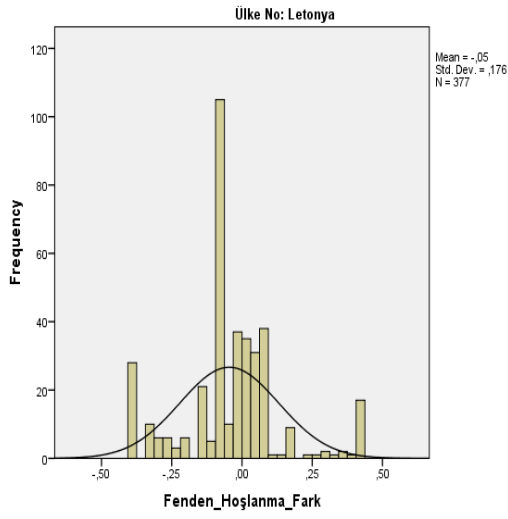


Fen Öz Yeterliği Fark Puanları

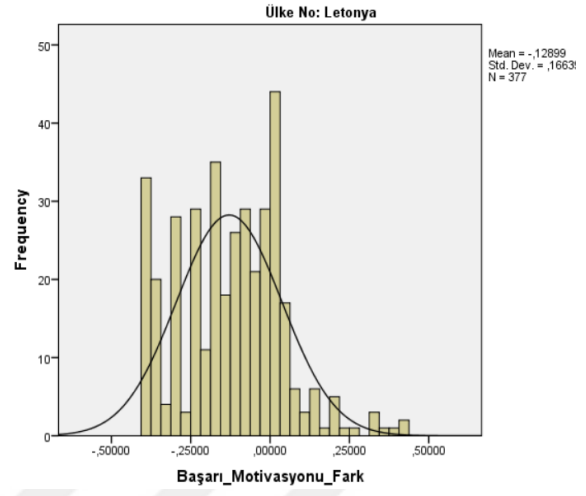


Letonya

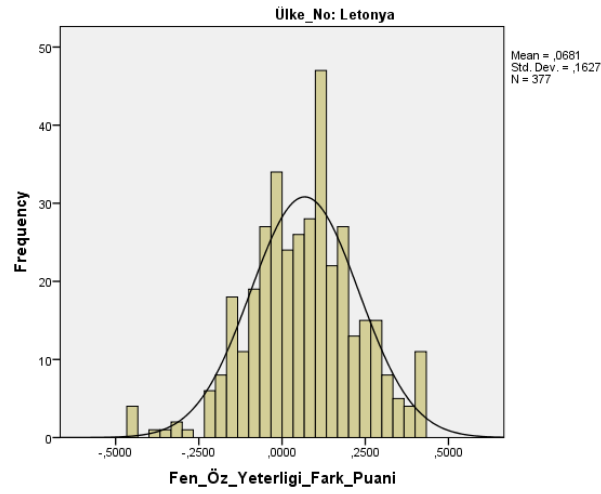
Fenden Hoşlanma Fark Puanları



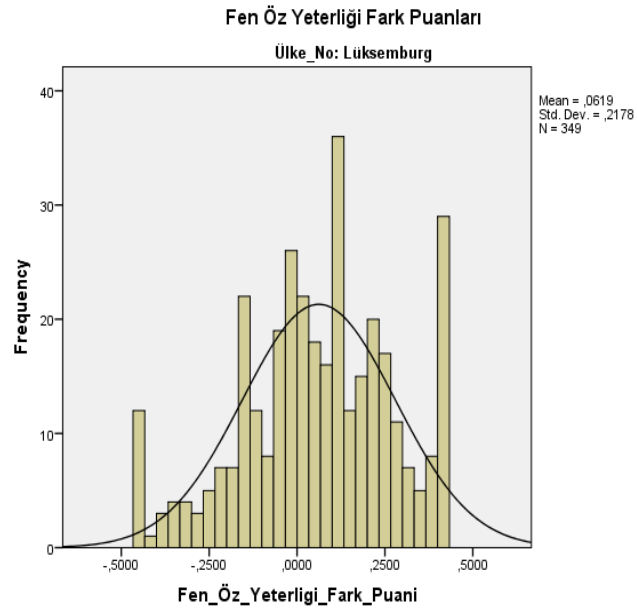
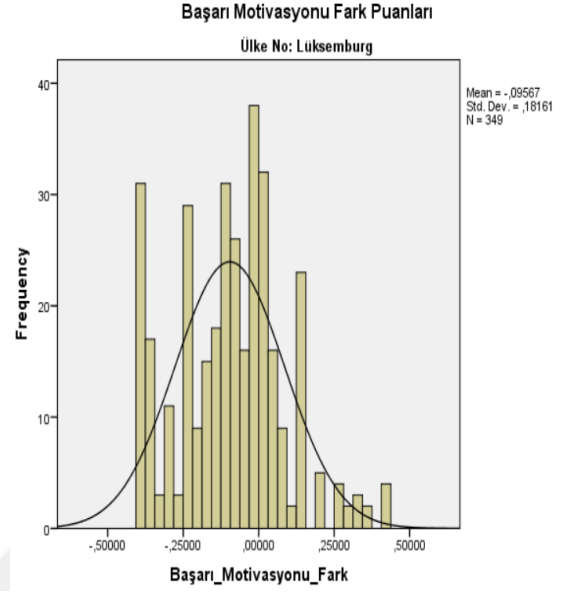
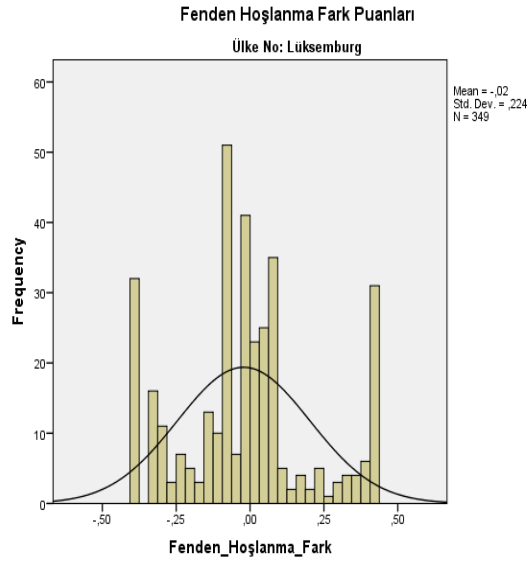
Başarı Motivasyonu Fark Puanları



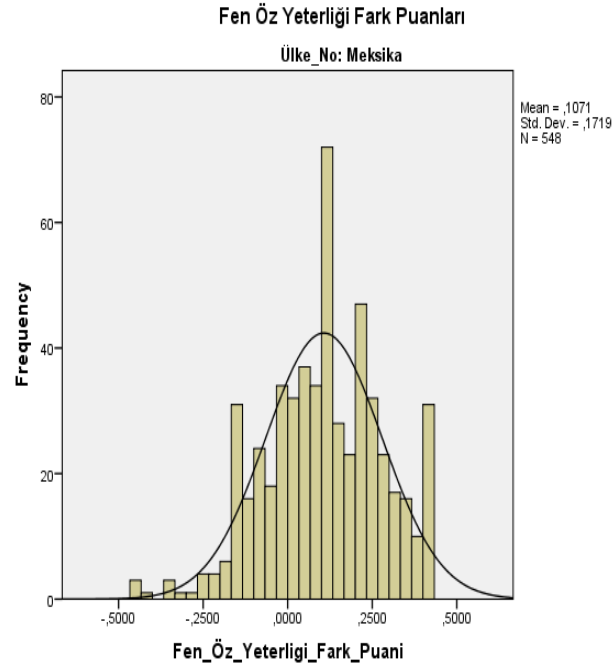
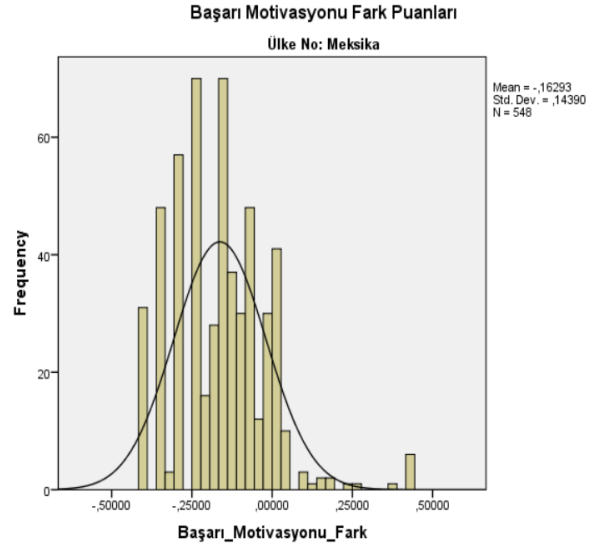
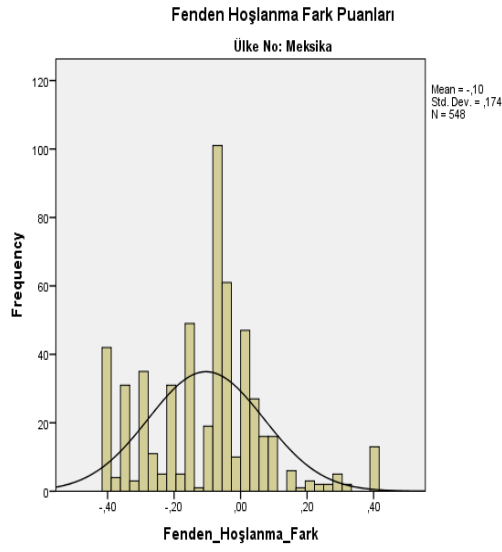
Fen Öz Yeterliği Fark Puanları



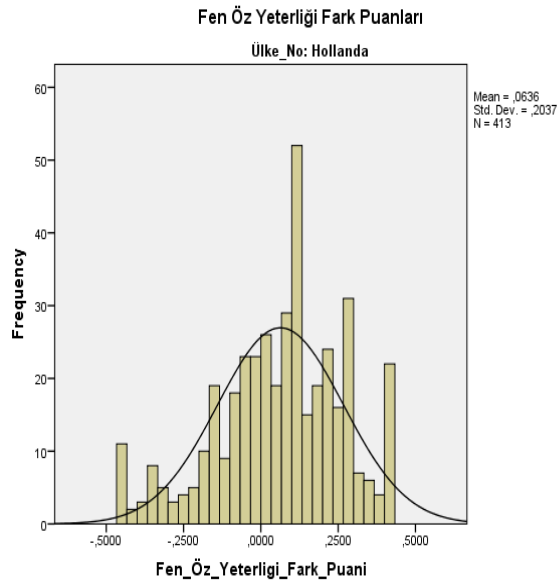
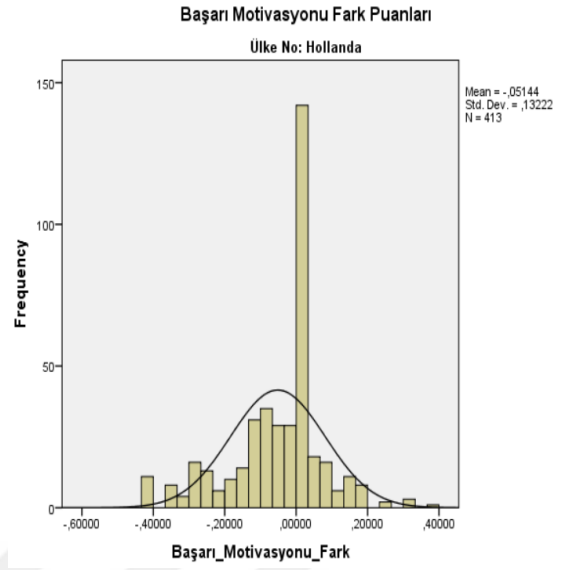
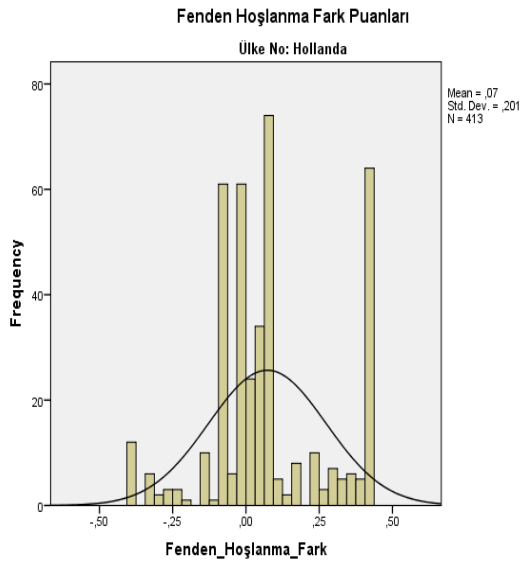
Lüksemburg



Meksika

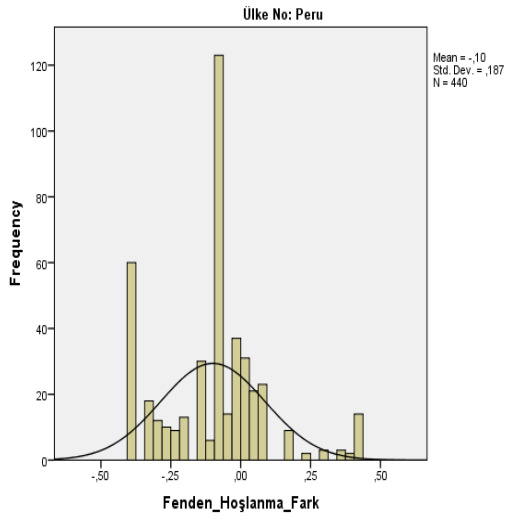


Hollanda

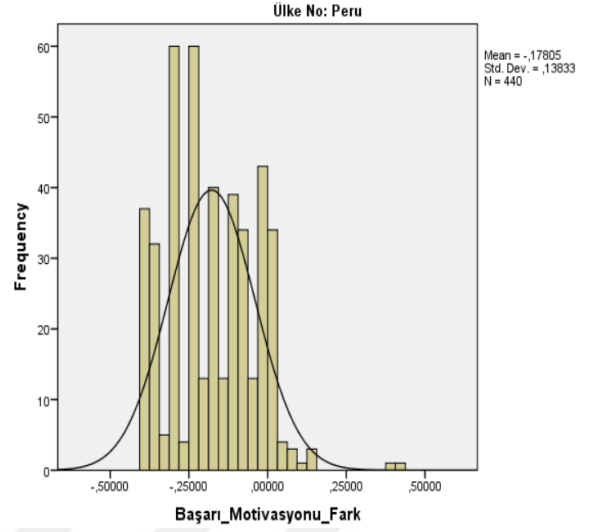


Peru

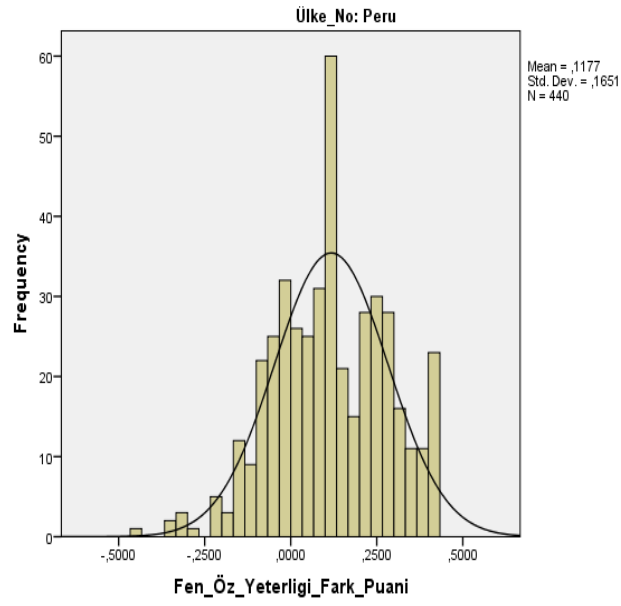
Fenden Hoşlanma Fark Puanları



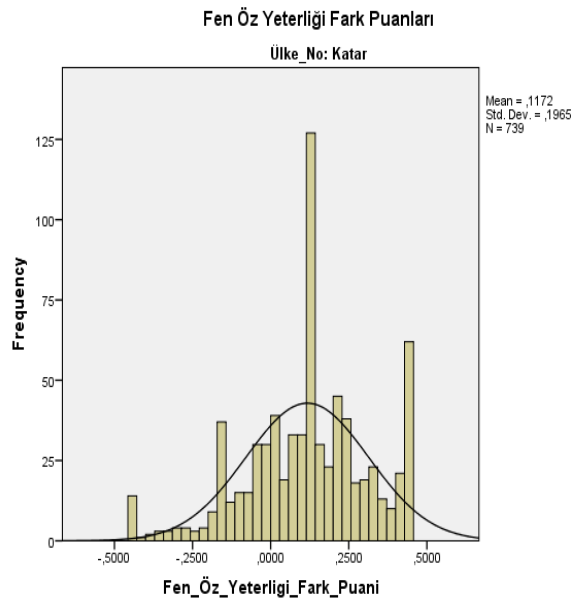
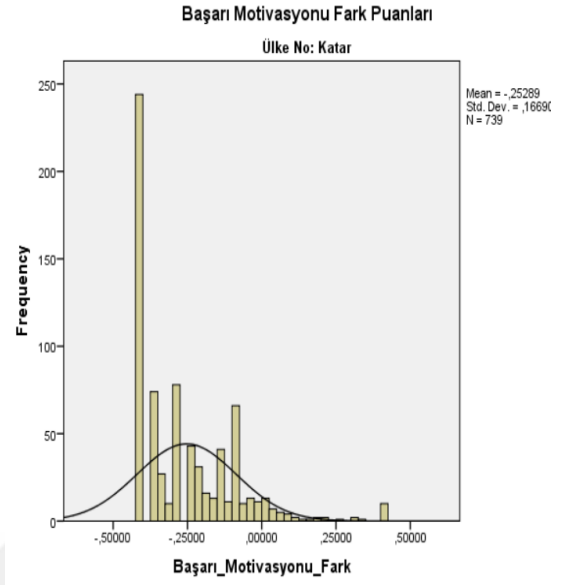
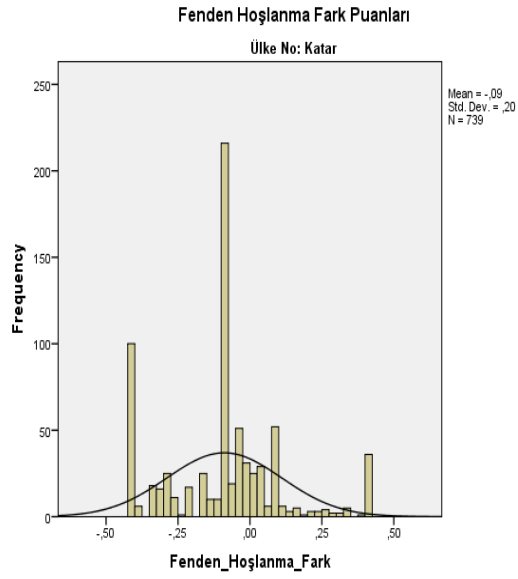
Başarı Motivasyonu Fark Puanları



Fen Öz Yeterliği Fark Puanları

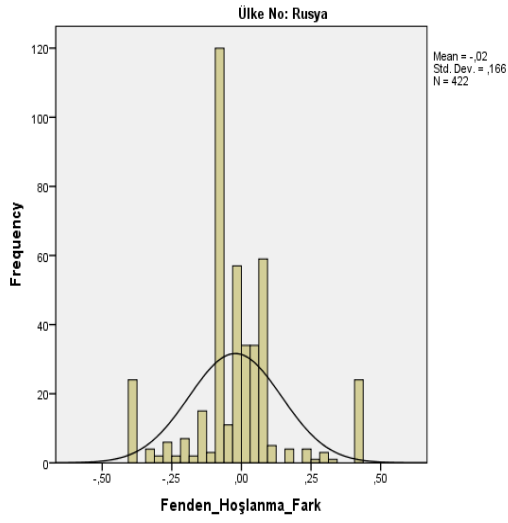


Katar

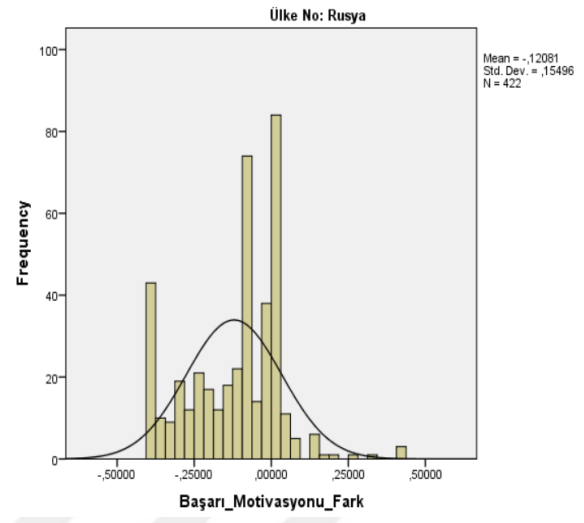


Rusya

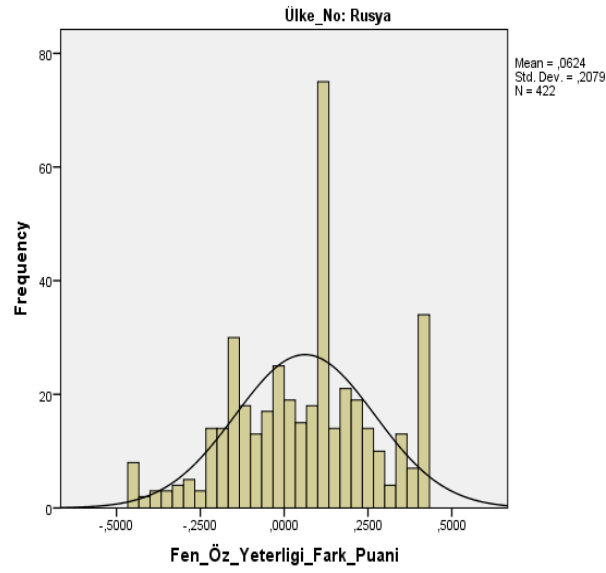
Fenden Hoşlanma Fark Puanları



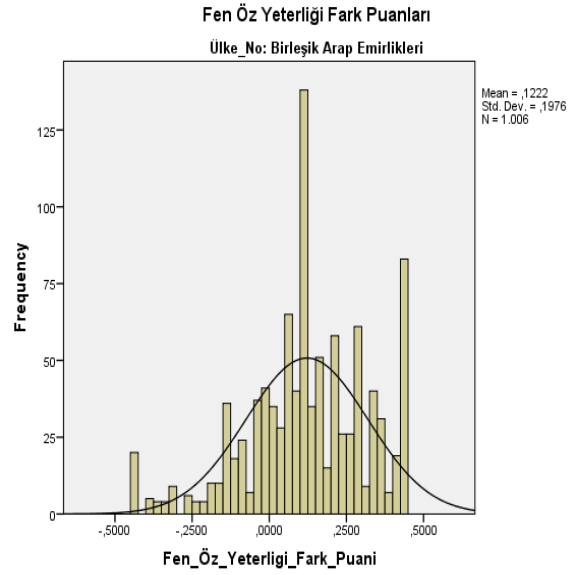
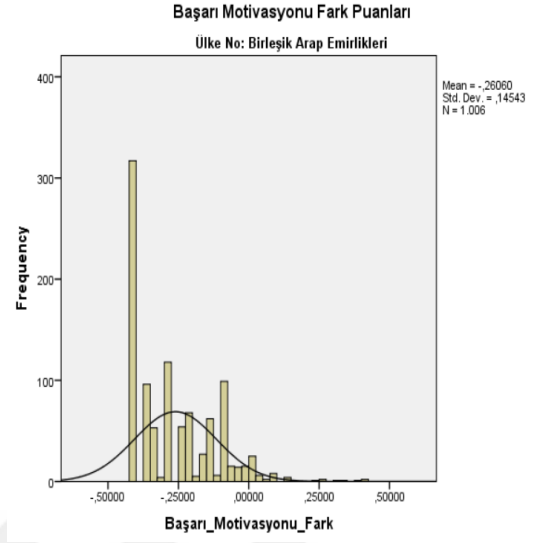
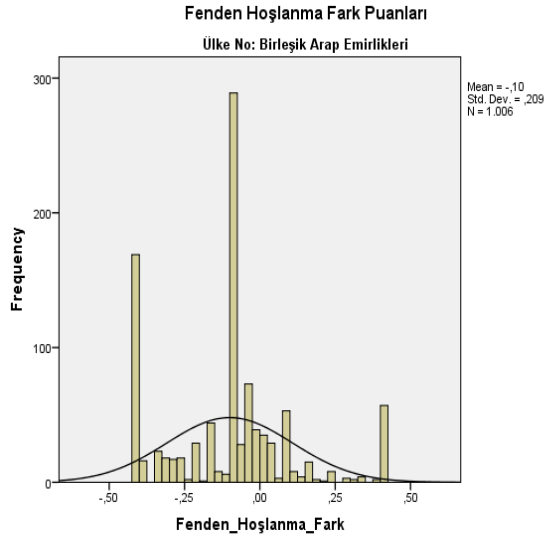
Başarı Motivasyonu Fark Puanları



Fen Öz Yeterliği Fark Puanları

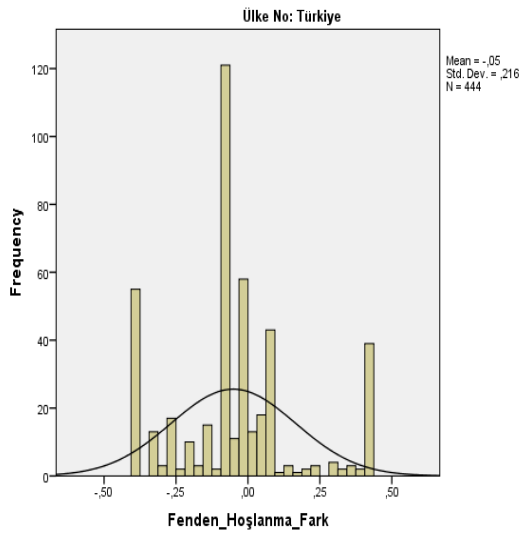


Birleşik Arap Emirlikleri

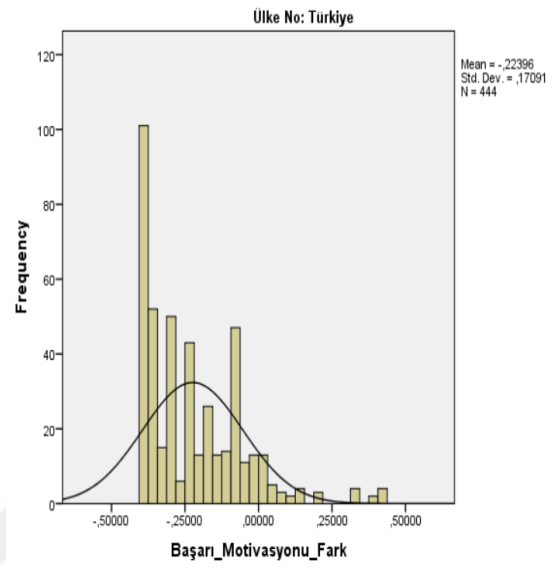


Türkiye

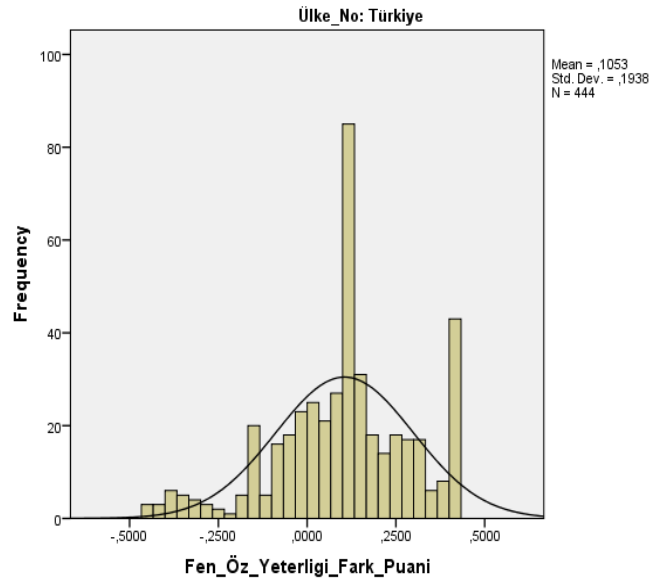
Fenden Hoşlanma Fark Puanları



Başarı Motivasyonu Fark Puanları



Fen Öz Yeterliği Fark Puanları



Ek 7. Bootstrap Örneklerinin Elde Edilmesinde Kullanılan R Kodları

```
library(ccrs)
library(haven)
library(NbClust)
Sevme <- read_sav("C:/Users/Desktop/Sevme.sav")
View(Sevme)
Sevme<-data.matrix(Sevme)
Sevme<-Sevme[,2:6]
Sevme[sample(x=nrow(Sevme),replace=TRUE),]
bootstrap<-replicate(n=200,Sevme[sample(x=nrow(Sevme),replace=TRUE),],simplify=F)
bootstrap
f<-bootstrap
n <- 10898 ; m <-5 ;q <- 4
#ccrs paketi ile veri setinin hazırlanması
ccrsdata.list <- list()
for(i in 1:200){
  ccrsdata.list[[i]] <-create.ccrsdata(f[[i]],q=q)}
#ccrs analizi
lam <- 0.8; K <- 1;
ccrs.list<-list()
for(i in 1:200){
  ccrs.list[[i]] <-ccrs(ccrsdata.list[[i]],K=K,lam=lam,
    tandem.initial=FALSE,tol = 1e-5, maxit = 50,
    trace = 1, nstart = 20, parallel=F,verbose=T)}
##Tepki Stili Tabanlı Kümeleme Sonuçları
ts <- list()
for(i in 1:200){
  ts[[i]] <- NbClust(ccrs.list[[i]]$ccrs.list$Beta[,2:4], diss = NULL, distance = "euclidean",
    min.nc = 2, max.nc = 15, method = "kmeans", index = "kl")
}
```

```

#İçerik Tabanlı Kümeleme Sonuçları

ik <- list()

for(i in 1:200){

  ik[[i]] <- NbClust(ccrs.list[[i]]$crs.list$Y.hat, diss = NULL, distance = "euclidean",
min.nc = 2, max.nc = 15, method = "kmeans", index = "kl") }

# K-ortalamlar Kümeleme Sonuçları

k <- list()

for(i in 1:200) {

  k[[i]] <- NbClust(f[[i]], diss = NULL, distance = "euclidean", min.nc = 2, max.nc = 15,
method = "kmeans", index = "kl") }

#Tepki Stili Tabanlı Kümeleme Sonuçlarını Dışarı Aktarma

l <- list()

for(i in 1:200){

  l[[i]] <-ts[[i]]$Best.nc}

write.csv(l,"C:/Users/Desktop/SevmeTS.csv")

#K-ortalamlar Kümeleme Sonuçlarını Dışarı Aktarma

m <- list()

for(i in 1:200){

  m[[i]] <-k[[i]]$Best.nc}

write.csv(m,"C:/Users/Desktop/SevmeKmeans.csv")

#İçerik Tabanlı Kümeleme Sonuçlarını Dışarı Aktarma

n <- list()

for(i in 1:200){

  n[[i]] <-ik[[i]]$Best.nc}

write.csv(n,"C:/Users/Desktop/SevmeKS.csv")

```

Ek 8. TSYDK Yöntemi ve K-ortalamlar Tekniği ile Elde Edilen Kümeleme Sonuçları için R Kodları

```
library(ccrs)

library(haven)

library(ggplot2)

library(NbClust)

Sevme <- read_sav("C:/Users/Desktop/Sevme.sav")

View(Sevme)

Sevme<-data.matrix(Sevme)

Sev<-Sevme[,2:6]

View(Sev)

n <- 10898 ; m <- 5 ; H.true <- 3 ; K.true <- 2 ; q <- 4

ccrsdata.list <-create.ccrsdata(Sev,q=q)

convert.X2F(Sev,q=q)

###CCRS

lam <- 0.8; K <- 2;

ccrs.list <- ccrs(ccrsdata.list,K=K,lam=lam, tandem.initial=FALSE,tol = 1e-5, maxit = 50,
trace = 1, nstart = 500, parallel=F,verbose=T)

###check content-based clustering result

plot(ccrs.list$crs.list)

ccrs.list$cls.cont.vec

#Tepki Stiline Göre Kümeleme Sonuçları

tskmeans <- kmeans(ccrs.list$crs.list$Beta[,2:4], centers = 3, nstart = 500)

#K-ortalamlar Kümeleme Sonuçları

final3 <- kmeans(Sev, centers = 4, nstart = 500)
```



GAZİLİ OLMAK AYRICALIKTIR...