

**ORTAK MADDELERİN DEĞİŞEN MADDE FONKSİYONU
GÖSTERİP GÖSTERMEMESİ DURUMUNDA TEST
EŞİTLEMeye ETKİSİNİN FARKLI YÖNTEMLERLE
İNCELENMESİ**

**THE STUDY OF THE EFFECT OF ANCHOR ITEMS
SHOWING OR NOT SHOWING DIFFERENTIAL ITEM
FUNCTIONING TO TEST EQUATING USING VARIOUS
METHODS**

Kadriye Belgin DEMİRUS

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme

Bilim Dalı İçin Öngördüğü

Doktora Tezi

olarak hazırlanmıştır.

2015

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼'ne,

Kadriye Belgin DEMİRUS'un hazırladığı "Ortak Maddelerin Deđişen Madde Fonksiyonu Gösterip Göstermemesi Durumunda Test Eşitlemeye Etkisinin Farklı Yöntemlerle İncelenmesi" başlıklı bu çalıřma j¼rimiz tarafından **Eđitim Bilimleri Anabilim Dalı, Eđitimde Ölçme ve Deđerlendirme Bilim Dalı'nda Doktora Tezi** olarak kabul edilmiřtir.

Başkan Prof. Dr. řener BÜYÜKÖZTÜRK _____

Üye (Danıřman) Prof. Dr. Selahattin GELBAL _____

Üye Doç. Dr. Duygu ANIL _____

Üye Yrd. Doç. Dr. Ömer Kutlu _____

Üye Yrd. Doç Dr. Sevda ÇETİN _____

ONAY

Bu tez Hacettepe Üniversitesi Lisansüstü Eđitim-Öđretim ve Sınav Yönetmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri üyeleri tarafından 24 / 04 / 2015 tarihinde uygun gör¼lmüş ve Enstitü Yönetim Kurulunca / / tarihinde kabul edilmiřtir.

Prof. Dr. Berrin AKMAN
Eđitim Bilimleri Enstitüsü M¼d¼r¼

ORTAK MADDELERİN DEĞİŞEN MADDE FONKSİYONU GÖSTERİP GÖSTERMEMESİ DURUMUNDA TEST EŞİTLEMeye ETKİSİNİN FARKLI YÖNTEMLERLE İNCELENMESİ

Kadriye Belgin DEMİRUS

ÖZ

Bu araştırmada ortak maddelerin tamamı cinsiyete göre TB-DMF'li/DMF'siz olduğunda Madde Tepki Kuramı'na dayalı yapılan eşitleme yöntemlerinin performansını karşılaştırmak amaçlanmıştır. Araştırmada DMF'li maddelerin test eşitlemeye etkisi gerçek veri üzerinden, ayrı kalibrasyon yöntemleri ve eşdeğer gruplarda ortak test deseni kullanılarak yatay eşitleme ile ortaya konulmuştur.

Araştırmada DMF analizleri "Mantel-Haenszel" yöntemi için EASYDIF programında ve "lojistik regresyon" yöntemi için Zumbo tarafından hazırlanan syntax ile SPSS'de yapılmıştır. Test eşitleme yöntemleri olarak lineer ölçek dönüştürme (moment) yöntemlerinden "ortalama-ortalama" ile "ortalama-sigma" ve karakteristik eğrisi dönüştürme yöntemlerinden "Haebara" ile "Stocking-Lord" kullanılmıştır. Eşitleme yöntemlerinin performansı RMSD eşitleme hataları hesaplanarak değerlendirilmiştir. Madde parametrelerinin ve yeteneğin kestiriminde BILOG-MG, test eşitlemede IRTEQ yazılımı işe koşulmuştur.

Çalışmanın verisi oluşturulan fen testi formlarının 1350 8.sınıf öğrencisine uygulamasından elde edilmiştir. Araştırmanın sonucunda ortak maddeler erkekler lehine TB-DMF'li olduğunda en büyük RMSD eşitleme hatasını ortalama-ortalama yöntemi, en küçük hatayı ise ortalama-sigma yöntemi üretmiştir. Ortak maddeler DMF'siz olduğunda ise en büyük hata ortalama-sigma yönteminde, en küçük RMSD eşitleme hatası karakteristik eğrisi yöntemlerinde (Stocking-Lord ve Haebara) birbirine eşit olarak elde edilmiştir.

Anahtar sözcükler: Test eşitleme, değişen madde fonksiyonu, ortak test deseni, eşitleme hatası

Danışman: Prof. Dr. Selahattin GELBAL, Hacettepe Üniversitesi, Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı

THE STUDY OF THE EFFECT OF ANCHOR ITEMS SHOWING OR NOT SHOWING DIFFERENTIAL ITEM FUNCTIONING TO TEST EQUATING USING VARIOUS METHODS

Kadriye Belgin DEMİRUS

ABSTRACT

The purpose of this study to compare the results of equating methods based on Item Response Theory when all the anchor items showing or not showing gender based uniform DIF. In the study the effect of items with DIF on test equating presented on real data with horizontal equating using separate calibration methods and equivalent groups with anchor test design.

DIF analysis conducted on EASDIF software for “Mantel-Haenszel” method and on SPSS with syntax presented by Zumbo for “logistic regression” method. For test equating methods “mean- mean” , “mean-sigma”, “Haebara” and “Stocking-Lord” were used. The performance of the equating methods were evaluated through RMSD equating errors. BILOG-MG was utilized for the prediction of item parameters and ability, IRTEQ software was utilized for test equating.

Data set for the study has been obtained from the forms of science test has been developed which applied to 1350 students in 8th grade. According to the results of the study when the anchor items showing uniform DIF favored males, mean-mean method has produced the biggest equating error while mean-sigma method has produced the smallest. When the anchor items not showing DIF the biggest equating error has been obtained from mean-sigma method and smallest equating error has been obtained from Stocking-Lord ve Haebara methods in equal to each other.

Keywords: test equating, differential item functioning, anchor test design, equating error

Advisor: Prof. Dr. Selahattin GELBAL, Hacettepe University, Department of Educational Sciences, Division of Measurement and Evaluation in Education

ETİK BEYANNAMESİ

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

Kadriye Belgin DEMİRUS

TEŞEKKÜR

Öncelikle kendisiyle yolumun Tokat'ta bir sempozyumda kesiştiği, tanıştığım ilk anda mütevazı tavırlarıyla kafamdaki "hoca" algısını değiştiren, her zaman yapıcı ve iyi niyetli konuşmalarıyla beni dimdik ayakta tutan çok değerli bilim insanı hocam, danışmanım Prof. Dr. Selahattin GELBAL'a sonsuz teşekkürlerimi sunarım.

Görüş ve önerileriyle bu çalışmaya çok değerli katkılarda bulunan jüri üyelerim Prof.Dr. Şener BÜYÜKÖZTÜRK, Doç. Dr. Duygu ANIL, Yrd. Doç. Dr. Ömer KUTLU ve Yrd. Doç Dr. Sevda ÇETİN'e çok teşekkür ederim.

Doktora öğrenimim boyunca yetişmemde emeği olan Prof. Dr. Hülya KELECİOĞLU'na, Prof. Dr. Nuray SENEMOĞLU'na, bana ders dışı değerli vakitlerini ayırmaktan çekinmeyen Doç. Dr. Nuri DOĞAN'a ve Dr. Zafer ÇEPNİ'ye, destekleriyle yanımda olan ölçmeciler arkadaşlarım Dr. Güler YAVUZ'a ve İbrahim UYSAL'a, yurt dışındaki yayınlara ulaşmamda katkı sağlayan Doç. Dr. Burcu ATAR ve Dr. Ahmet TURHAN'a, yazdıkları Easydif programıyla beni tanıştıran Dr. Isabel BENÍTEZ BAENA'ya ve tezimin uygulamasında yardımcı olan üniversite arkadaşlarıma teşekkürlerimi sunarım.

Son olarak iyi ve kötü günümde her zaman desteğini gördüğüm ilk öğretmenim canım annem Hatice'ye, babam Zeki'ye, psikoloğum olan kardeşlerim Akif ve Yavuz'a çok teşekkür ederim. Doktora kabulü aldığım ilk günün sevincini aynı heyecanla bitirdiğim gün de benimle paylaşan ve tüm bu zorlu süreç boyunca sonsuz destekçim olarak yanımda bulunan biricik eşim Oğuz'a ve tabii ki bilgisayar başında çalışırken tekmeleriyle beni motive eden canım oğlum Uğur Alp'e ne kadar teşekkür etsem azdır...

Sevgili Babaannem Kadriye evik'e

İÇİNDEKİLER

ÖZ.....	iii
ABSTRACT.....	iv
TEŞEKKÜR.....	vi
İÇİNDEKİLER.....	viii
TABLolar DİZİNİ.....	x
ŞEKİLLER DİZİNİ.....	xi
SİMGELER VE KISALTMALAR DİZİNİ.....	xii
1. GİRİŞ.....	1
1.1. Problem Durumu.....	1
1.2. Araştırmanın Amacı ve Önemi:.....	5
1.3. Problem Cümlesi:.....	6
1.3.1. Alt Problemler:.....	7
1.4. Sayıtlılar:.....	8
1.5. Sınırlılıklar:.....	8
1.7. Araştırmanın Kuramsal Temeli.....	8
2. İLGİLİ ARAŞTIRMALAR.....	40
2.1. İlgili Araştırmalar Özet.....	56
3. YÖNTEM.....	58
3.1. Araştırmanın Türü.....	58
3.2. Araştırma Verileri.....	58
3.3. Veri Toplama Aracı.....	59
3.4. Evren ve Örneklem.....	64
3.5. Araştırmanın Deseni.....	65
3.6. Kullanılan Programlar.....	65
3.7. Verilerin Analizi.....	66
3.8. Tek Boyutluluk Varsayımın Test Edilmesi.....	67
3.9. Yerel Bağımsızlık Varsayımın Test Edilmesi.....	69
3.10. Kalibrasyon ve Madde Parametreleri ile Yetenek Kestirimleri.....	70
3.11. Model Veri Uyumunun Kontrolü.....	73
3.12. A ve B Formlarının KTK ve MTK'ya Dayalı Güvenirlik Katsayıları.....	77
3.13. Eşitleme Hatasının Hesaplanması.....	77
4. BULGULAR VE TARTIŞMA.....	80
4.1. Ortak Maddelerin DMF'li Olduğu Durumda Uygulanan MTK'ya Dayalı Test Eşitleme Yöntemlerinden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular.....	81
4.2. Ortak Maddelerin DMF'siz Olduğu Durumda Uygulanan MTK'ya Dayalı Test Eşitleme Yöntemlerinden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular.....	83

4.3. Ortak Maddelerin DMF'li veya DMF'siz Olduğu Durumlarda A ve B Formlarının Ortalama-Ortalama Yönteminden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular	86
4.4. Ortak Maddelerin DMF'li veya DMF'siz Olduğu Durumlarda A ve B Formlarının Ortalama-Sigma Yönteminden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular	88
4.5. Ortak Maddelerin DMF'li veya DMF'siz Olduğu Durumlarda A ve B Formlarının Haebara'nın Karakteristik Eğrisi Dönüştürme Yönteminden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular	91
4.6. Ortak Maddelerin DMF'li veya DMF'siz Olduğu Durumlarda A ve B Formlarının Stocking-Lord'un Karakteristik Eğrisi Dönüştürme Yönteminden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular	92
4.7. Ortak Maddelerin DMF'li veya DMF'siz Olduğu Durumlarda A ve B Formlarının Ortalama-Ortalama, Ortalama, Haebara ve Stocking-Lord Yöntemlerinden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular	94
5. SONUÇ ve ÖNERİLER	96
5.1. Sonuçlar.....	96
5.2. Öneriler.....	97
5.2.1. Araştırma Sonuçlarına Dönük Öneriler.....	97
5.2.2. Diğer Araştırmacılara Dönük Öneriler	99
KAYNAKÇA.....	101
EKLER DİZİNİ	108
EK 1. A VE B FORMU CİNSİYETE DAYALI DMF-MH VE DMF-LR ANALİZ SONUÇLARI.....	109
EK 2. A VE B FORMLARININ DOĞRULAYICI FAKTÖR ANALİZİ ÖLÇME MODELLERİ.....	111
EK 3. A VE B FORMLARINA AİT KESTİRİLEN YETENEKLERİN (θ) NORMAL DAĞILIM EĞRİLERİ.....	115
EK 4. MADDE KARAKTERİSTİK EĞRİLERİ	116
EK 5. TEST BİLGİ EĞRİLERİ	122
EK 6. EŞİTLENEN A VE B FORMLARINDA MADDELERİN ÖLÇTÜKLERİ KAZANIM VE BİLİMSEL SÜREÇ BECERİLERİ	127
EK 7. UYGULAMA YAPILAN FEN ve TEKNOLOJİ TESTİ A VE B FORMLARI	128
EK 7. ARAŞTIRMA İZİN TALEBİ	140
EK 8. ORJİNALLİK RAPORU	141
ÖZGEÇMİŞ.	142

TABLolar DİZİNİ

Tablo 1.1: MH İstatistiği için i Maddesinin j Yetenek Düzeyinde Karşılaştırma Tablosu	12
Tablo 1.2: Δ -MH DMF Etki Büyüklüğü Değerlendirmesi	13
Tablo 1.3: Tek Grup Deseni	18
Tablo 1.4: Dengelenmiş Tek Grup Deseni	19
Tablo 1.5: Eşdeğer Gruplar Deseni	19
Tablo 1.6: Eşdeğer Olmayan Gruplarda Ortak Test Deseni	20
Tablo 1.7: Eşdeğer Gruplarda Ortak Test Deseni	22
Tablo 1.8: İlköğretim 6., 7. ve 8. Sınıf Düzeyi İçin “Bilimsel Süreç Becerisi” Kazanımları	39
Tablo 3.1: Gerçek Uygulamada Kullanılan A ve B Formları Madde Sayıları	60
Tablo 3.2: A ve B Formlarında Maddelerin Sınav Yılı ve Madde Numaralarına Göre Dağılımları	61
Tablo 3.3: A ve B Formu Cinsiyete Dayalı DMF-MH ve DMF-LR Analizlerine İlişkin Bulgular	62
Tablo 3.4: Eşitleme Yapılacak A ve B Formlarının Son Durumdaki Madde Sayıları	63
Tablo 3.5: DMF Analizi için Kullanılan Evren ve Örneklem Sayıları	64
Tablo 3.6: Araştırmada Kullanılan A ve B Formlarının Cinsiyete Göre Dağılımı	65
Tablo 3.7: Eşitlenecek A ve B Formlarına Ait KTK'ya Dayalı Betimsel İstatistikler	66
Tablo 3.8: A ve B Formlarına Ait Doğrulayıcı Faktör Analizi Sonuçları	68
Tablo 3.9: Farklı Yetenek Gruplarından Elde Edilen Maddeler Arası Tetrakorik Korelasyonlara Ait Betimsel İstatistikler	70
Tablo 3.10: 1PL Model için Madde Parametrelerine Ait Betimsel İstatistikler	72
Tablo 3.11: 2PL Model için Madde Parametrelerine Ait Betimsel İstatistikler	72
Tablo 3.12: 3PL Model için Madde Parametrelerine Ait Betimsel İstatistikler	72
Tablo 3.13: A ve B Formlarının 1, 2, 3PLM'e Dayalı Uyum İyiliği Sonuçları	75
Tablo 3.14: A ve B Formlarının MTK'ya Dayalı Betimsel İstatistikleri	76
Tablo 3.15: A ve B Formlarının KTK ve MTK'ya Dayalı Güvenirlik Katsayıları	77
Tablo 4.1: Ortak Maddelerin DMF'li ve DMF'siz Olduğu Durumlarda Dört Yönteme göre Yapılan Eşitlemelerin RMSD Eşitleme Hataları	80
Tablo 4.2: Ortak Maddelerin DMF'li ve DMF'siz Olduğu Durumlarda Dört Eşitleme Yöntemine göre Elde Edilen A ve B Eşitleme Katsayıları	80
Tablo 5.1: Araştırma Sonuçlarına Dayalı Önerilen Eşitleme Yöntemleri	98

ŞEKİLLER DİZİNİ

Şekil 1.1 Tek Biçimli Olmayan DMF	Şekil 1.2 Tek Biçimli DMF	10
Şekil 1.3. X ve Y Formlarının Hipotetik Test Karakteristik Eğrileri		33
Şekil 3.1: A Formu (Ortak Maddeler DMF'li) 4. Maddeye Ait 1, 2, 3 PLM Uyum İyiliği Testleri		74
Şekil 4.1: Ortak Maddeler DMF'li Olduğunda Eşitleme Hataları (RMSD)		81
Şekil 4.2: Ortak Maddeler DMF'siz Olduğunda Eşitleme Hataları (RMSD)		84
Şekil 4.3: Ortak Maddeler DMF'li veya DMF'siz Olduğunda O-O Yöntemine Göre Eşitleme Hataları (RMSD)		86
Şekil 4.4: Ortak Maddeler DMF'li veya DMF'siz Olduğunda O-S Yöntemine Göre Eşitleme Hataları (RMSD)		88
Şekil 4.5: Ortak Maddeler DMF'li veya DMF'siz Olduğunda Haebara Yöntemine Göre Eşitleme Hataları (RMSD)		91
Şekil 4.6: Ortak Maddeler DMF'li veya DMF'siz Olduğunda S-L Yöntemine Göre Eşitleme Hataları (RMSD)		92
Şekil 4.7: Ortak Maddeler DMF'li veya DMF'siz Olduğunda O-O, O-S, Haebara ve S-L Yöntemlerine Göre Eşitleme Hataları (RMSD)		94

SİMGELER VE KISALTMALAR DİZİNİ

MTK: Madde Tepki Kuramı

KTK: Klasik Test Kuramı

MKE: Madde Karakteristik Eğrisi

MH: Mantel- Haenzsel

LR: Lojistik Regresyon

O-O: Ortalama-Ortalama Test eşitleme Yöntemi

O-S: Ortalama-Sigma Test Eşitleme Yöntemi

S-L: Stocking-Lord Test Eşitleme Yöntemi

HB: Haebara Test Eşitleme Yöntemi

DMF: Değişen Madde Fonksiyonu

TB-DMF: Tek Biçimli Değişen Madde Fonksiyonu

TBO-DMF: Tek Biçimli Olmayan Değişen Madde Fonksiyonu

OKÖSYS: Orta Öğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavı

SBS: Seviye Belirleme Sınavı

PISA: Uluslararası Öğrenci Değerlendirme Programı (Programme for International Student Assessment)

PIRLS: Uluslararası Okuma Becerilerinde Gelişim Projesi (Progress in International Reading Literacy Study)

TIMMS: Uluslararası Matematik ve Fen Bilgisi Çalışması (Trends in International Mathematics and Science Study)

KPSS: Kamu Personeli Seçme Sınavı

YGS: Yükseköğretime Geçiş Sınavı

TEOG: Temel Öğretimden Ortaöğretime Geçiş

LYS: Lisans Yerleştirme Sınavı

YDS: Yabancı Dil Sınavı

ALES: Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı

1. GİRİŞ

Bu bölümde problem durumu, arařtırmanın amacı ve önemi, problem cümlesi, alt problemler, sayılılar, sınırlılıklar ve arařtırmanın kuramsal temelleri başlıklarına yer verilmiřtir.

1.1. Problem Durumu

Ülkemizde ulusal düzeyde yapılan sınavlarda (TEOG, YGS, LYS, YDS, ALES, KPSS) öğrenci veya personel seçme ve yerleřtirme amaçlanırken, uluslararası düzeyde yapılan sınavlarda (PISA, TIMMS, PIRLS) öğrenci başarısı ve tutumunu etkileyen faktörlerin ortaya konulması ve ülkeler arası karşılařtırmaların yapılması amaçlanmaktadır. Ulusal sınavlarda bireyler hakkında, uluslararası sınavlarda ülkeler hakkında önemli kararlar alınmaktadır. Bu nedenle test puanlarına dayalı yapılan yorumların gerçeęi yansıtması için geçerli ölçmelere ihtiyaç duyulmaktadır.

Geçerlik test geliřtirme ve deęerlendirmede kullanılan en temel faktördür. Test puanlarına dayalı yapılan yorumları ya da alınan kararları destekleyen teori ve kanıtların bir derecesidir (AERA, APA ve NCME, 1999). Bir test bireylerde ölçülmesi amaçlanan özellięi başka özelliklerle karıřtırmadan ne kadar doęru ölçebiliyorsa o kadar geçerlidir (Tekin,1991; McDonald, 1999). Cinsiyet, yerleřim yeri, okul türü, kültür, sosyo ekonomik düzey gibi demografik özellikler öğrenci başarısını etkileyen deęiřkenler olarak arařtırmalara konu olmaktadır. Ancak bu deęiřkenlerde belli bir alt grupta (Ör: kız-erkek) bulunan eřit yetenek düzeyindeki bireylerin puanları arasında, ölçme amacı olmamasına raęmen, gruba baęlı farklılık söz konusuysa yapılan ölçmelerin geçerli olmadığı söylenebilir. Geçerli bir testin maddelerinde ölçülmek istenilen özellik evrendeki bütün bireyler için aynı şeyi ifade etmelidir.

Sınavlarda alt gruplar arası farklılıkların gerçek yetenek farklılıęından mı yoksa testin özellięinden mi kaynaklandıęının arařtırılması yanlılık çalıřmaları ile yapılmaktadır. Yanlılık tesadüfi hatadan kaynaklanmamaktadır; çünkü tesadüfi hata grup üyelerinin hepsi için geçerlidir. Yanlılık, alt gruplardaki bireylerin test puanlarında, buldukları gruba baęlı olarak farklılık bulunmasıdır ve sistematik hata olarak ele alınır. Ölçme sonuçları sınava giren farklı gruplara karşı yansızlık

özelliđi taşınalıdır. Yansızlık geçerlik için bir kanıt teşkil etmektedir (Camilli ve Shepard, 1994; Zumbo 1999).

Bir testteki maddelerin aynı yetenek düzeyinde doğru cevaplanma olasılıkları alt gruplara göre farklılık gösteriyorsa, o maddenin deđişen madde fonksiyonuna (DMF) sahip olduđu kabul edilmektedir (Camilli ve Shepard, 1994; Embretson ve Reise, 2000). DMF ve madde yanlılıđı birbirinden farklı kavramlardır. İstatistiksel olarak farklı fonksiyonlaşan maddelerin yanlı olup olmadıklarını belirlemek için betimsel çalışmalara ihtiyaç duyulmaktadır. Bu nedenle DMF, madde yanlılıđı için gerekli bir koşuldur; ancak yanlılık kararı vermek için yeterli deđildir. Maddelerdeki DMF varlıđı, alt gruplar arası gerçek farklılıktan kaynaklanabileceđi gibi ölçülmesi istenmeyen deđişkenlerin etkisi ile madde yanlılıđının bir sonucu olarak da ortaya çıkabilmektedir (Zumbo, 1999). Testlerin psikometrik özellikleri kapsamında ele alınan DMF incelemesi geçerlik için bir kanıttır (Embretson, 2007). Aynı yetenek düzeyinde olmalarına rağmen farklı gruplardaki bireylerin aynı test maddesine verdikleri cevapların farklı olmasının nedenlerini incelemek, hem test maddesinin özelliklerinin hem de farklı gruplarda bulunan bireylerin bilişsel süreçlerinin, test alma stratejilerinin ve bilgi eksikliklerinin anlaşılabilmesini sağlayacaktır. Böylece, test puanlarını etkileyen ancak test ile ölçülmesi amaçlanmayan bozucu deđişkenlerin varlıđı tespit edilebilecektir. DMF'nin belirlenmesi, düzeltici çalışmaların yapılması, testin yanlılık içermediđinin ortaya konulması testin yapı geçerliđinin sağlanmasında ve test alma koşullarının eşdeğerliđinden emin olmada önemli görülen süreçlerdir (Camilli ve Shepard, 1994).

Ülkemizde yapılan ulusal ve uluslararası sınavlarda maddelerin yanlılık taşıyıp taşımadıđını belirlemeyi amaçlayan pek çok araştırma bulunmaktadır. Bu araştırmalarda madde yanlılıđının belirlenmesi testin geçerlik ve güvenirliliđini artırmak için önemli bir çalışma olarak görülmektedir (Yurdugöl, 2003; Bekçi, 2007; Bakan Kalaycıođlu, 2008; Asil, 2010; Çepni, 2011). Literatür incelendiđinde özellikle fen ve matematik gibi sayısal alan maddelerinde cinsiyet deđişkenine göre DMF gözlenmektedir. Araştırmalarda erkek ve kızların farklı ilgi alanları ve yeteneklerine sahip olmalarının maddelerin farklı fonksiyonlaşmasına neden olduđu belirtilmektedir. Mantıksal analiz sonucu DMF alt gruplar arası gerçek yetenek farklılıđından kaynaklanmadıđında madde gruplardan birine avantaj sağlamak ve yanlı kabul edilmektedir. Örneđin, ÖSS fen bilimleri testinde

erkekler lehine işleyen üç DMF'li fizik maddesinden bir tanesinin yanlılık gösterdiği saptanmıştır. Yanlılık nedeni olarak fizik maddesinde kullanılan otomobillere ve hız konusuna erkek öğrencilerin daha yakın olması gösterilmiştir (Bakan Kalaycıoğlu, 2008). Testin amacına uygun olmayan şekilde test maddesinin bazı karakteristik özelliklerinden dolayı erkeklerin kızlara göre maddeyi daha çok doğru cevaplama madde yanlılığını ortaya çıkarmıştır (Zumbo, 1999).

Alt gruplar arası gerçek yetenek farklılığından kaynaklanan değişen madde fonksiyonu yanlılığa neden olmamaktadır. Bir başka ifadeyle maddenin amaca hizmet ettiğini, kız ve erkek öğrenciler için aynı yeteneği ölçtüğünü göstermektedir. ÖSS matematik alt testinde bir madde kızlar lehine, iki madde erkekler lehine DMF'li bulunmuştur. Ancak matematik maddelerinde yanlılığa sebep olacak herhangi bir durum görülmemiştir. Kız öğrencilerin erkek öğrencilere göre işlemleri daha düzenli ve adım adım yapması sebebiyle bir maddenin kız öğrenciler lehine çalışmış olabileceği; iki maddenin ise geometri maddeleri olması ve görsel uzamsal konularda erkek öğrencilerin daha başarılı olmaları sebebiyle erkek öğrenciler lehine çalışmalarının beklendiği belirtilmiştir (Bakan Kalaycıoğlu, 2008).

ALES sayısal 1 testinde üç maddenin erkek öğrenciler lehine, dört maddenin ise kız öğrenciler lehine işlediği, sayısal 2 testinde ise bir maddenin erkek öğrenciler lehine, üç maddenin ise kız öğrenciler lehine işlediği bulunmuştur. Cebirsel ifadelerle soyut bir biçimde sunulmuş, algoritmik işlemlerle çözülen maddelerin kız öğrenciler lehine; gerçek hayat problemi olarak ifade edilmiş, algoritmik rutin işlemlerle çözülemeyen problemlerin ise erkekler lehine işlediği tespit edilmiştir. DMF gösteren maddelerin yanlı olup olmadıkları uzman kanılarıyla yardımıyla incelendiğinde, DMF sebebi olabilecek gerekçelerin tümü testin ölçmesi amaçlanan matematiksel/sayısal yetenek yapısı içinde kaldığı görülmüş ve yanlı olmadıkları belirtilmiştir. Literatürde erkek öğrencilerin matematiği, kız öğrencilere kıyasla, hayatlarında daha çok kullanılabilir ve daha değerli bir kavram olarak algıladıklarının rapor edildiği, bu nedenle erkeklerin gerçek hayattan alınmış uygulamalı problemlerde denk yetenek düzeylerindeki kız öğrencilere göre daha başarılı olmalarının beklendiği belirtilmiştir (Çepni, 2011). Bireylerin ya da ülkelerin geleceği hakkında önemli kararlar alınan sınavlarda testlerin geçerliğini artırmak için her sınavda DMF analizlerinin yapılarak sonraki sınavlarda yanlı çıkma olasılığı bulunan maddelerin elimine edilmesi gerekmektedir.

Yanlılık aynı testin farklı formları arasında güçlük farkı olduğunda da ortaya çıkabilmektedir. Ortaöğretime, lisans ve lisansüstü yükseköğretime giriş, yabancı dil yeterlik belirleme gibi seçme ve/veya yerleştirme gerektiren sınavlar (TEOG, YGS, LYS, YDS, ALES, KPSS) yılda bir kez ya da birden fazla uygulanmaktadır. Bununla birlikte ülkeler arası karşılaştırmaların yapıldığı uluslararası geniş ölçekli testlerde (PISA, TIMMS, PIRLS) farklı sorular içeren çok sayıda kitapçık bulunmaktadır. Aynı teste ait farklı formların kullanılması durumunda sınava başvuranların kolay/zor formdan hangisini aldığı dikkat edilmesi gereken önemli bir konudur. Çünkü zor formun sorularını çözen bireyler kolay formu çözenlere göre daha düşük puan alabilirler. Böyle bir durumda test formlarının birbiriyle karşılaştırılması güçleşmektedir. Formları ortak bir ölçeğe yerleştirmek için kullanılan istatistiksel yöntemler test eşitleme olarak adlandırılmaktadır. Eşitleme yapılarak farklı formlardan elde edilen puanlar birbiri yerine kullanılabilir. Böylece zor testi alan bireylere karşı yapılan haksızlık önlenerek, test formlarından kaynaklanan yanlılık problemi ortadan kaldırılmaktadır (Angoff, 1971; Hambleton, Swaminathan ve Rogers, 1991; Cook ve Eignor, 1991; Kolen ve Brennan, 2004; Holland ve Dorans, 2006). Eşitleme için bir formun birim sistemi diğer formun birim sistemine dönüştürülmelidir. Bu dönüştürme sonucunda formların eşdeğer olacağı belirtilmektedir. Aynı testin farklı formlarından elde edilen ölçme sonuçlarının eşitlenmesi ve eşitleme sonuçlarının doğruluğu sınavın geçerliği açısından önemlidir. Bir testin farklı formlarından elde edilen puanların eşitlenmesi sayesinde bireylerin gelişimleri, yeterlikleri belirlenerek performanslarının da karşılaştırılması mümkün olabilmektedir (Angoff, 1971).

Test maddeleri farklı şekilde çalıştığında gruplar arası puanların karşılaştırılması zorlaşmaktadır. Ortak test kullanılarak yapılacak bir eşitlemede ise, söz konusu ortak test DMF'li maddeler barındırıyorsa, bu maddeler ortak testin geçerliğini ve güvenilirliğini düşürme eğilimindedir ve grup farklılıklarının doğru biçimde ortaya konmasında ciddi tehlike oluşturmaktadır. Dolayısıyla ortak test kullanıldığında, bu testin farklı test formlarında aynı şekilde işleyip işlemediğini belirlemek için DMF analizi yapılmalıdır. Uygulamalarda ortak maddelerin gruplar arasında farklı fonksiyonlaşmaması için farklı test formları arasında yaklaşık olarak aynı pozisyonda yönetilmelidir (Cook ve Paterson, 1987).

DMF'li maddelerin eşitleme yapılacak testlerde yer alması arzu edilmeyen bir durumdur. Çünkü bu tür maddeler, test eşitlemeye olumsuz etki yapabilir ve bu durum test eşitleme hatası ile belirlenebilmektedir. Örneğin, farklı sınıf düzeylerinde en az hataya sahip dikey bir ölçek hazırlamak oldukça önemlidir. Bununla birlikte eşitleme katsayıları DMF varlığında bozulabilmektedir. Bu yüzden bu tür DMF'li maddelerin belirlenmesi ve test eşitleme katsayılarına yaptıkları etkinin en aza indirilmesi gerekmektedir (Hidalgo-Montesinos ve Lopez-Pina, 2002). DMF'li maddelerin testten çıkarılması arzu edilmeyen bir durumdur; çünkü bu durum madde sayısının azalmasına ve kapsam geçerliğinin düşmesine neden olmaktadır. Chu (2002) ve Turhan (2006) araştırmalarında DMF'li maddelerin testten çıkarılmasının eşitleme hatasını arttırdığını; bu artışın DMF'li madde sayısı arttıkça büyüdüğünü belirtmişlerdir. Chu (2002) araştırmasında özellikle kısa bir testten DMF'li maddenin silinmesinin geçerliği daha fazla zedelediği sonucuna ulaşmıştır. Dolayısıyla böyle bir durumda DMF'li maddeleri çıkarmak yerine, DMF'nin farklı yöntemlerle elde edilen eşitleme hatalarına etkisi karşılaştırılarak düşük hata veren yöntem tercih edilebilir.

1.2. Araştırmanın Amacı ve Önemi:

Bu araştırmada, ortak maddeler cinsiyete göre DMF'li veya DMF'siz olduğunda Madde Tepki Kuramı'na dayalı lineer ölçek dönüştürme (moment) yöntemlerinden "ortalama-ortalama" ile "ortalama-sigma" ve karakteristik eğrisi dönüştürme yöntemlerinden "Haebara" ile "Stocking-Lord" eşitleme yöntemlerinden hesaplanan eşitleme hatalarının karşılaştırılması amaçlanmıştır. Söz konusu dört yöntemden moment yöntemleri kolay hesaplanabilmeleri; karakteristik eğrisi yöntemleri üstünlüklere sahip olmaları nedeniyle geniş bir kullanım alanına sahiptir.

Yapılan literatür taramasında genel olarak DMF'li maddeler varlığında test eşitleme konulu araştırmaların oldukça sınırlı sayıda olduğu gözlenmektedir (Chu, 2002; Turhan, 2006; Atalay Kabasakal, 2014; Huggins, 2014). Bununla birlikte madde parametre kaymasını (Item Parameter Drift, IPD) değişen madde fonksiyonunun bir çeşidi olarak ele alıp IPD'li maddelerin madde parametre kestirimlerine, eşitleme katsayılarına, yetenek kestirimlerine etkisini farklı eşitleme

yöntemleriyle inceleyen arařtırmalar bulunmaktadır (Wells ve dię., 2002; Han, 2008; Babcock ve Albano, 2012).

Ülkemizde yapılan MTK'ya dayalı eřitleme alıřmalarının daha ok eřdeęer olmayan gruplarda ortak madde deseni ile simülasyon veri üzerinden gerekleřtirildięi görölmektedir (Kilmen, 2010; Gök, 2012; Uysal, 2014). Gerek veri üzerinden yapılan eřitleme alıřmalarında MTK'ya dayalı Rasch modeli ve 2PLM kullanılarak sözde ortak test deseni (Kelecioęlu, 1994), Rasch modeli kullanılarak ortak test deseni (řahhüseyinoęlu, 2005) ve dikey eřitleme ile (etin, 2009) alıřıldıęı; KTK'ya dayalı daha ok tek grup deseni (Bozdaę 2007; Öztürk, 2010; Kan, 2010; 2011; Mutluer, 2013) alıřmalar yapıldıęı belirlenmiřtir.

Simülasyon veri kullanmanın avantajları olduęu gibi dezavantajları da bulunmaktadır. Öncelikle oluřturulan veri üzerinde bir yanlılık olduęuna dair řüphe vardır; eęer veri belirli bir MTK modeline dayalı oluřturulduysa bu yanlılık karřılařtırılan eřitleme yöntemlerinden birisinin ierisinde yer alabilir (Hwang ve Cleary, 1986). Bununla birlikte simülasyon veri sonuçların genelleneceęi gerek veriye tam olarak benzemeyebilir. Harris ve Crouse (1993) simülasyon verinin en faydalı olduęu durumun gerek veriye en ok benzedięi durum olduęunu belirtmiřtir; bu sayede simülasyon alıřmaları gerek veri eřitleme durumlarını temsil edecek řekilde genelleřtirilebilir ve gerek veri ieren analizlerde kullanılabilir. Ancak, bu durumu saęlamak kolay olmamaktadır.

Bu arařtırmada DMF'li maddelerin test eřitlemeye etkisi, dięer arařtırmalardan farklı olarak gerek veri üzerinden, ayrı kalibrasyon yöntemleri ve eřdeęer gruplarda ortak test deseni kullanılarak, ortak maddelerin tamamı DMF'li veya DMF'siz olduęunda, yatay eřitleme ile ortaya konulmaktadır. Bu nedenle arařtırmanın test eřitleme alanında yapılan alıřmalara önemli bir katkı getireceęi düşünölmektedir.

1.3. Problem Cümlesi:

Fen ve teknoloji başarı testi formlarının MTK'ya dayalı yöntemlerle eřitlenmesinde ortak maddelerin DMF'li veya DMF'siz olması durumlarının eřitleme hatasına etkisi nasıldır?

1.3.1. Alt Problemler:

1) Ortak maddelerin DMF'li olduđu durumda uygulanan MTK'ya dayalı test eşitleme yöntemlerinde eşitleme hataları nasıl deęişim göstermektedir?

a) Ortak maddeler DMF'li olduđunda ortalama-ortalama yönteminde eşitleme hatası nasıldır?

b) Ortak maddeler DMF'li olduđunda ortalama-sigma yönteminde eşitleme hatası nasıldır?

c) Ortak maddeler DMF'li olduđunda Haebara'nın karakteristik eğrisi dönüştürme yönteminde eşitleme hatası nasıldır?

d) Ortak maddeler DMF'li olduđunda Stocking-Lord'un karakteristik eğrisi dönüştürme yönteminde eşitleme hatası nasıldır?

2) Ortak maddelerin DMF'siz olduđu durumda uygulanan MTK'ya dayalı test eşitleme yöntemlerinde eşitleme hataları nasıl deęişim göstermektedir?

a) Ortak maddeler DMF'siz olduđunda ortalama-ortalama yönteminde eşitleme hatası nasıldır?

b) Ortak maddeler DMF'siz olduđunda ortalama-sigma yönteminde eşitleme hatası nasıldır?

c) Ortak maddeler DMF'siz olduđunda Haebara'nın karakteristik eğrisi dönüştürme yönteminde eşitleme hatası nasıldır?

d) Ortak maddeler DMF'siz olduđunda Stocking-Lord'un karakteristik eğrisi dönüştürme yönteminde eşitleme hatası nasıldır?

3) Ortak maddelerin DMF'li veya DMF'siz olduđu durumlarda A ve B formlarının eşitlenmesinde kullanılan ortalama-ortalama yönteminde eşitleme hataları nasıl deęişmektedir?

4) Ortak maddelerin DMF'li veya DMF'siz olduđu durumlarda A ve B formlarının eşitlenmesinde kullanılan ortalama-sigma yönteminde eşitleme hataları nasıl deęişmektedir?

5) Ortak maddelerin DMF'li veya DMF'siz olduđu durumlarda A ve B formlarının eşitlenmesinde kullanılan Haebara'nın karakteristik eğrisi dönüştürme yönteminde eşitleme hataları nasıl deęişmektedir?

6) Ortak maddelerin DMF'li veya DMF'siz olduđu durumlarda A ve B formlarının eşitlenmesinde kullanılan Stocking-Lord'un karakteristik eğrisi dönüştürme yönteminde eşitleme hataları nasıl değişmektedir?

7) Ortak maddelerin DMF'li veya DMF'siz olduđu durumlarda A ve B formlarının eşitlenmesinde kullanılan ortalama-ortalama, ortalama-sigma, Haebara ve Stocking-Lord yöntemlerinde eşitleme hataları nasıl değişmektedir?

1.4. Sayıtlar:

Tüm öğrenciler için fen ve teknoloji testi formlarının uygulanma şartları eşittir.

1.5. Sınırlılıklar:

- 1) Fen ve teknoloji testi maddelerinin DMF analizi cinsiyet değişkeni ile sınırlıdır.
- 2) Araştırmada DMF'li maddelerin yanlılığını araştırma kapsam dışı tutulmuştur.
- 3) Araştırma MTK'ya dayalı ve parametrelerin ayrı ayrı kestirimine dayalı test eşitleme yöntemleri ile sınırlıdır.

1.7. Araştırmanın Kuramsal Temeli

A. Değişen Madde Fonksiyonu

Testle ölçülen özellikler bakımından birbiriyle benzer, bir başka ifadeyle aynı yetenek düzeyinde olan; ancak farklı gruplarda yer alan bireylerin bir maddeyi doğru cevaplama olasılıklarının farklı olması, o maddenin değişen madde fonksiyonuna sahip olduğunun göstergesidir (Hambleton ve diğ., 1991). Maddenin ölçülmek istendiği yetenek düzeyinde yapılacak bir karşılaştırmada cinsiyet, etnik grup, sosyal sınıf, yaş, bölge, yaşanan çevre vb. bakımından farklı gruplardaki cevaplayıcıların, söz konusu maddeyi doğru cevaplama olasılıklarında meydana gelen istatistiksel farklılık değişen madde fonksiyonu olduğunu göstermektedir. Dezavantajlı olduğu düşünülen grup odak (focal) grup olarak adlandırılırken, bu grubun performansı ile karşılaştırılan avantajlı grup ise referans (reference) grup olarak adlandırılmaktadır. Referans grup çoğunluğu (majority), odak grup ise azınlığı (minority) teşkil eden grup olarak da ifade edilmektedir (Camilli ve Shepard, 1994; Holland ve Wainer, 1993).

Değişen madde fonksiyonu (Differential Item Functioning, DIF), madde yanlılığı (item bias) kavramı ile çoğu zaman karıştırılmaktadır. Madde yanlılığı, istatistiksel

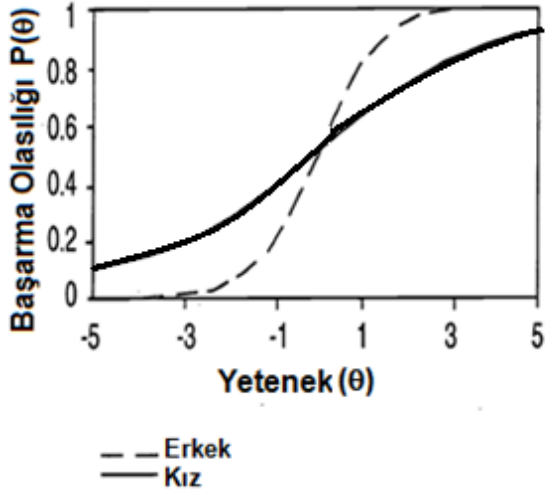
olarak DMF'li bulunan maddelerin aynı yetenek düzeyinde bulunan gruptan birine daha zor/kolay gelmesinin veya gruptan birinin avantajlı/dezavantajlı olmasının mantıksal analiz ile kanıt bulmasıdır. Bu bağlamda test performansında veya test maddelerinde görülen istatistiksel grup farklılıkları otomatik olarak "yanlılık" şeklinde ifade edilmemelidir; çünkü söz konusu puan farklılıkları, deneyim ve bilgiden kaynaklı grup farklılıklarının geçerli (gerçek) bir yansıması olabilir (Camilli ve Shepard, 1994).

Sonuç olarak hangi DMF belirleme yöntemiyle olursa olsun test ya da madde ile ölçülmek istenilen özellik bakımından gruplar arasında elde edilen DMF gerçek bir farklılıktan (madde etkisi=item impact) ya da madde yanlılığından kaynaklanabilmektedir. Madde etkisine bağlı olarak bulunan DMF'de, alt gruplar ölçülen özellik bakımından gerçekten birbirinden farklı olduğu için maddeyi doğru cevaplama olasılıkları da birbirinden farklılık göstermektedir. Dolayısıyla bir maddenin istatistiksel olarak DMF göstermesi onun yanlı olduğu anlamına gelmemektedir (Camilli ve Shepard, 1994; Zumbo, 1999). Değişen madde fonksiyonuna sahip bir madde yanlı olmayabilir; ancak yanlı bir maddenin değişen madde fonksiyonuna sahip olduğu söylenebilir.

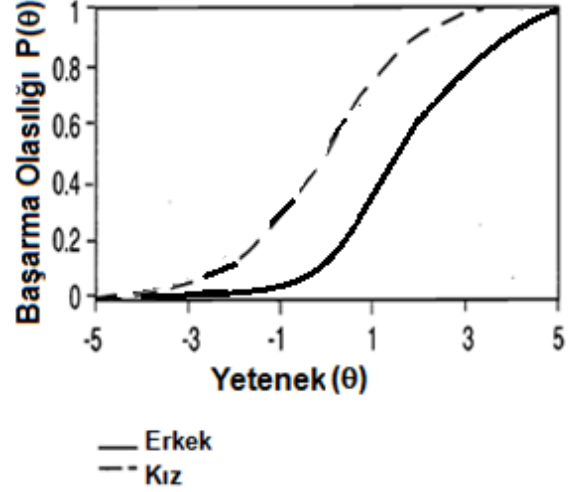
Örneğin, bir testi alan ve aynı yetenek düzeyinde bulunan bireylerin cinsiyete göre maddeyi doğru cevaplama olasılıkları farklı olsun. Bu durumda söz konusu madde cinsiyet değişkenine göre DMF içermektedir, denilir. Bu fark cinsiyet alt gruplarına dayalı gerçek bir farklılıktan kaynaklanabileceği gibi madde yanlılığından da kaynaklanabilir. Dolayısıyla betimsel verilerin toplanması yoluyla maddenin yanlı olup olmadığı araştırılmalıdır.

DMF, tek biçimli (TB-DMF) ve tek biçimli olmayan (TBO-DMF) olmak üzere iki çeşittir. Bir grup (referans grup) tüm yetenek düzeylerinde diğer gruptan (odak grup) daha büyük maddeye doğru cevap verme olasılığına sahipse o madde TB-DMF barındırmaktadır. Maddeyi doğru cevaplama olasılığı bir grup için bazı yetenek düzeylerinde diğer gruptakinden daha büyük bazı yetenek düzeylerinde ise daha küçükse o madde TBO-DMF içermektedir. TB-DMF'de alt gruplar için çizilen madde karakteristik eğrileri (MKE) çakışmazken; TBO-DMF'de ise MKE'lerin çakıştığı gözlenmektedir. TBO-DMF'ye ait bir MKE'de iki grubun bir madde üzerindeki değişen a, b ve c parametrelerine göre görece performansı gözlenmektedir. TB-DMF varlığında görülen kesişmeyen MKE'ler ise aynı a

parametresine (ayırıcılık gücüne) sahip olduklarında ortaya çıkan zorunlu bir durumdur (Ellis ve Raju, 2003; Camilli ve Shepard, 1994; Swaminathan ve Rogers, 1990).



Şekil 1.1 Tek Biçimli Olmayan DMF



Şekil 1.2 Tek Biçimli DMF

Şekil 1.1'de TBO-DMF'ye sahip bir maddeye ait karakteristik eğrisi görülmektedir. Burada tüm yetenek düzeylerinde hep aynı grubun lehine doğru cevaplama olasılığı yoktur. Düşük yetenek düzeylerinde kızlar avantajlıyken; yüksek yetenek düzeylerinde erkeklerin avantajlı olduğu gözlenmektedir. Şekil 1.2'de ise TB-DMF'ye sahip bir maddeye ait karakteristik eğrisi verilmiştir. Burada tüm yetenek düzeylerinde kızlar erkeklere göre maddeyi daha yüksek doğru cevaplama olasılığına sahiptir. Bir başka ifadeyle verilen maddede kızların lehine TB-DMF bulunmaktadır.

MKE'ler arasında kalan alan, DMF'nin düzeyi hakkında izlenim vermektedir. Eğer iki MKE kesişirse, bu alanın bir kısmı pozitif DMF, diğer kısmı ise negatif DMF olarak değerlendirilmektedir. DMF'nin bir belirtisi, belirli bir grup için hesaplanan bu net alandır. Ne var ki, dikkate değer miktarda DMF'nin, sıfıra indirildiği durumlar daha detaylı şekilde incelenmelidir; nihayetinde madde iki grup için de farklı fonksiyon göstermektedir. Bu özel durumda her iki eğrinin alanlarını pozitif olarak düşünüp genel DMF indeksini bulmak için toplamak faydalı olabilmektedir (Camilli ve Shepard, 1993).

B. Değişen Madde Fonksiyonu Belirleme Yöntemleri

DMF belirlemede kullanılan pek çok yöntem bulunmaktadır. Bu yöntemler Klasik Test Kuramı'na (KTK) ve Madde Tepki Kuramı'na dayalı yöntemler olarak sınıflandırılabilir. KTK'ya göre DMF belirleme yöntemleri olarak; Mantel-Haenszel (MH) (Holland ve Thayer, 1988), Lojistik regresyon (LR) (Swamanithan ve Rogers, 1990), SIBTEST (Shealy ve Stout, 1993) yöntemleri örnek gösterilmektedir. Lord'un ki-kare testi (Lord, 1980), Raju'nun alan ölçümleri (Raju, 1988,1990) ve Olabilirlik oranı (Thissen, Steinberg ve Wainer, 1993) yöntemleri ise MTK'ya dayalı DMF belirleme yöntemleri olarak belirtilmektedir (Camilli ve Shepard,1994; Ellis ve Raju, 2003; Holland ve Wainer, 1993).

Aşağıda bu araştırmada DMF belirlemede kullanılan MH ve LR yöntemleri açıklanmıştır.

• Mantel-Haenszel Yöntemi

Mantel-Haenszel istatistik yöntemi, sigara kullanımı ve akciğer kanseri arasındaki ilişkiyi araştıran sosyolog-matematikçi-istatistikçi Haenszel'in, örnek olay-kontrol çalışmalarını incelerken biyoistatistikçi Mantel'den geriye dönük veri analizi ile ilgili yardım istemesi ile ortaya çıkmıştır. Bu yöntemde analiz bir ya da daha çok değişken üzerine kurulduğunda ve bu değişkenler sabit kabul edildiğinde her alt grup- belli bir hastalık- maruz kalma odds oranları, hastalığa sahip olma riskini ifade etmektedir (Mannocci, 2009).

Mantel ve Haenszel (1959) tarafından geliştirilen bu yöntem daha sonra Holland ve Thayer (1988) tarafından DMF araştırmalarına uyarlanmıştır. MH istatistiği geniş kullanım alanına sahip kolay ve zahmetsiz bir DMF belirleme yöntemidir. Bu yöntemde daha önceden yetenek ölçeğinde eşleştirilmiş referans ve odak grupların madde performanslarını karşılaştırılmaktadır. Eşleştirme, kriter olarak kullanılan gözlenen toplam test puanının eşit olmasına dayalı yapılmaktadır. İki kategorili puanlamada her madde için K karşılaştırma tabloları (2x2) oluşturulmaktadır. Burada K, eşleşme değişkeni olan toplam test puanının bölüdüğü seviyelerin sayısıdır. Testin içerisindeki j puan seviyesindeki i maddesinin MH istatistiğini hesaplamada kullanılan 2x2 karşılaştırma tablosu Tablo 1.1'deki gibidir (Holland ve Thayer, 1988).

Tablo 1.1: MH İstatistiği için i Maddesinin j Yetenek Düzeyinde Karşılaştırma Tablosu

Grup	1	0	Toplam
Referans	A _j	B _j	NR _j
Odak	C _j	D _j	NO _j
Toplam	N1 _j	N0 _j	N _{..j}

Tablo 1.1’de, A_j referans grubunda maddeyi doğru cevaplayanların sayısını, B_j referans grubunda maddeyi yanlış cevaplayanların sayısını, C_j odak grubunda maddeyi doğru cevaplayanların sayısını, D_j ise odak grubunda maddeyi yanlış cevaplayanların sayısını belirtmektedir.

Verilen bir j yetenek düzeyinde referans ve odak grup için maddenin doğru cevap verilme frekanslarından yola çıkılarak hesaplanan odds oranı (α-MH) aşağıdaki eşitlikle hesaplanmaktadır:

$$\alpha_{-MH} = \frac{\sum_{j=1}^K A_j D_j / N_{..j}}{\sum_{j=1}^K B_j C_j / N_{..j}}$$

α-MH katsayısı odak ve referans gruptaki performans farklılığının derecesini ölçerek, DMF ölçüsünü göstermektedir. α-MH, 1’den uzaklaştıkça DMF derecesi artmaktadır. Bu durum aşağıdaki gibi ifade edilmektedir (Roznowski ve Reith,1999):

α-MH = 1.00 ise DMF yoktur.

α-MH < 1.00 ise madde odak grup lehine DMF içermektedir.

α-MH > 1.00 ise madde referans grup lehine DMF içermektedir.

Holland ve Thayer (1988) simetrik ölçek elde etmek amacıyla α-MH’nin logaritmik dönüşümünü önermiştir. Bu dönüşüm aşağıdaki gibi ifade edilmektedir.

$$\Delta_{-MH} = -2.35 \ln(\alpha_{MH})$$

Δ-MH’nin (MH-Delta DMF) sıfır değeri değişen madde fonksiyonu olmadığını belirtmektedir. Pozitif değerler odak grubun lehinde, negatif değerler ise referans

grubun lehinde DMF göstermektedir. DMF etki büyüklüğünü belirten Δ -MH'nin kategorilendirilmesi aşağıda Tablo 1.2'de verilmiştir (Zwick and Ercikan, 1989).

Tablo 1.2: Δ -MH DMF Etki Büyüklüğü Değerlendirmesi

Düzyey	Değer	DMF Yorumu
A	$ \Delta\text{-MH} < 1$	Yok veya ihmal edilebilir
B	$1 \leq \Delta\text{-MH} < 1,5$	Orta düzeyde, dikkat edilmeli
C	$ \Delta\text{-MH} \geq 1,5$	Yüksek düzeyde, önemli

EASYDIF (González ve diğ., 2011) yazılımı ile Mantel Haenszel istatistiği kullanılarak TB-DMF ve modifiye edilmiş Mantel-Haenszel prosedürü kullanılarak TBO-DMF analiz edilebilmektedir. Ayrıca standartlaştırma prosedürleri için gerekli istatistikleri de çıktı olarak verilmektedir. EASYDIF yazılımı adından anlaşıldığı üzere komut yazmaya gerek duymadan çeşitli seçeneklere direkt giriş yapmayı sağlayan kolay bir yazılım programıdır. Programa girilen veri dosyası sırasıyla grup numarası ve maddelere verilen cevaplar şeklinde aralarında boşluk olmadan *.txt veya *.dat formatında olmalıdır. Yazılım ile İki kategorili maddeler için Mantel-Haenszel ki-kare, Mantel-Haenszel odds oranı, MH-Delta DMF ve standart hatası, standartlaştırılmış p farkı, Delta ölçüğü için MH-Delta ve standart hatası hesaplanabilmektedir. Çok kategorili maddeler içinse, sıralı veri için MH istatistiği, p değeri ve standartlaştırılmış p değeri elde edilmektedir. Kişi sayısı için limit yoktur; madde sayısı ise 350/ (cevap kategori sayısı-1) kadar olabilmektedir. Tüm gruba, yüksek ve düşük performanslı gruplara ait istatistikler her madde için ayrı ayrı verilmektedir. Yüksek ve düşük performanslı gruplara ait χ^2 istatistikleri DMF'nin TB veya TBO olduğunu belirlemeyi sağlamaktadır. Eğer iki grup için de χ^2 olasılığı anlamlı ise TB-DMF; sadece bir grup için olasılık anlamlıysa TBO-DMF olduğu anlaşılmaktadır. Bu yazılımda ayrıca her bir yetenek düzeyinde odak ve referans grupların doğru cevaplama olasılıkları grafiksel olarak da görülebilmektedir (Padilla ve diğ., 2012).

- **Lojistik Regresyon Yöntemi**

Swaminathan ve Rogers (1990) tarafından önerilen lojistik regresyon analizinde odak ve referans grubundaki bireylerin bir maddeyi belli bir ölçüte göre doğru cevaplama olasılıkları hesaplanmaktadır. Odak ve referans grubun tanımlanması

bağımsız değişkendir ve bağımsız değişken kesikli veya sürekli veri niteliğinde olabilir. Kesikli veri türündeki maddeler bağımlı değişkeni oluşturmaktadır (Zumbo,1999).

Lojistik regresyon (LR) yönteminde, Mantel-Haenszel yöntemindeki gibi hem tek biçimli hem de tek biçimli olmayan DMF belirlenebilmektedir. Simülasyon çalışmalarında LR yönteminin MH yöntemine göre tek biçimli olmayan DMF'yi belirlemede daha güçlü olduğu elde edilmiştir. Lojistik regresyon analizinde serbestlik derecesi 2 olan χ^2 testi ile TB-DMF ve TBO-DMF'li olması muhtemel maddeler değişkenlerin sıra ile hiyerarşik biçimde modele alınmasıyla elde edilmektedir. İstatistiksel önemliliğe karar verirken manidarlık düzeyi (p) 0,05 olarak alınabilmektedir. Söz konusu değişkenler model 1 (temel model) için toplam test puanı (yetenek düzeyi), model 2 için toplam test puanı ve grup değişkeni, model 3 için ise bu iki değişkene ilaveten toplam puan-grup etkileşimi değişkenidir. LR yöntemine göre tek biçimli DMF varlığında, grup değişkeni (ör: cinsiyet) ile toplam puan değişkeni arasında herhangi bir etkileşim bulunmamaktadır ve model 2 ile model 1'in standartlaştırılmış regresyon katsayıları farkı ($\Delta R^2=R_2^2-R_1^2$) anlamlıdır. Yani; maddeyi doğru cevaplama olasılığı tüm yetenek düzeylerinde hep aynı grubun lehine daha büyüktür. Tek biçimli olmayan DMF'de ise grup ile toplam puan değişkenleri arasında etkileşim bulunmaktadır ve model 3 ile model 2'nin standartlaştırılmış regresyon katsayıları farkı ($\Delta R^2=R_3^2-R_2^2$) anlamlıdır. Bu durumda, iki grup için maddeyi doğru cevaplama olasılıkları farklılığı tüm yetenek düzeylerinde aynı değildir. Model 3 ile model 1'in standartlaştırılmış regresyon katsayıları farkı ($\Delta R^2=R_3^2-R_1^2$) anlamlı ise; ki bu fark Swaminathan ve Rogers'a (1990) göre χ^2 testine karşılık gelmektedir, hem TB hem de TBO-DMF'yi açıklamaktadır (Rogers ve Swaminathan, 1993; Zumbo, 1999; Gierl, Jodoin ve Ackerman, 2000).

Standartlaştırılmış regresyon katsayıları farkı (ΔR^2), DMF'nin derecesini vermektedir ve A, B, C olarak üç düzeyde belirlenmektedir. Literatürde kesme noktaları için farklı ölçütler alındığı gözlenmektedir. $\Delta R^2 < 0,035$ ise (A) DMF yoktur ya da ihmal edilebilir düzeydedir; $0,035 \leq \Delta R^2 < 0,070$ ise (B) orta düzeyde DMF vardır; $\Delta R^2 \geq 0,070$ ise (C) önemli düzeyde DMF vardır. Bir maddenin DMF içeren (B veya C düzeyi) bir madde olarak sınıflandırılabilmesi için ki-kare değerinin 0,05'e eşit veya 0,05'den küçük ve R^2 değerinin de en az 0,035 olması

gerekmektedir (Jodoin ve Gierl, 2001). Zumbo ve Thomas (1996)'a göre, $\Delta R^2 < 0,13$ ise (A) DMF yoktur ya da ihmal edilebilir düzeydedir, $0,13 \leq \Delta R^2 < 0,26$ ise (B) orta düzeyde DMF vardır, $\Delta R^2 \geq 0,26$ ise (C) önemli düzeyde DMF vardır; şeklinde yorumlanmaktadır. Bununla birlikte Bakan Kalaycıoğlu (2008), Berberoğlu ve Bakan Kalaycıoğlu (2008), Çepni'nin (2011) çalışmalarında $\Delta R^2 \geq 0.010$ olan maddeler B düzeyinde, $\Delta R^2 \geq 0.020$ olan maddeler C düzeyinde DMF'li kabul edilmektedir.

C. Test Eşitleme ve Koşulları

Güvenlik nedenleriyle bir testin farklı formları farklı zamanlarda uygulanabilmektedir; ancak bu durumda formların güçlükleri bakımından birbirinden farklı olabilecekleri gibi bir endişe ortaya çıkmaktadır. İstatistiksel bir süreç olan eşitleme, formların birbirleri yerine kullanılabilmesi için test formlarının puanlarını uyumlu hale getirmektedir. Eşitleme ile güçlük ve kapsam bakımından benzer şekilde oluşturulan formların güçlük farkları düzenlenmektedir. Özet olarak eşitleme, kapsamdaki farklılıkları değil; güçlükteki farklılıkları ayarlamaktadır. Bir formun birim sisteminin diğer formun birim sistemine dönüştürülmesi ve bu şekilde formlar arasında eşdeğerliğin sağlanması işlemidir. Test formlarındaki puanlar eşitlendiğinde bu formlar değiştirilebilir şekilde kullanılabilir. Test eşitleme, bir testin alternatif formlarına ait puanları raporlama ve yorumlama sürecini geliştirme potansiyeline sahiptir (Kolen ve Brennan, 2004; Angoff, 1971).

Aynı yeteneği ölçen tek boyutlu iki test formunun eşitlenebilmesi için aşağıdaki koşullara dikkat edilmelidir (Lord, 1980:199):

Eşitlik: Her bir yetenek düzeyi (Θ) için, $X(Y)$ 'nin şartlı frekans dağılımı, X 'in şartlı frekans dağılımıyla aynı olmalıdır.

Gruplararası Varyans Değişmezliği: Elde edildiği evrenden bağımsız olarak eşitleme fonksiyonu olan $X(Y)$ alt gruplarda aynı olmalıdır.

Simetriklik: Hangi formun X hangi formun Y olduğu farketmeden aynı eşitleme iki yönlü gerçekleştirilebilmelidir.

Eşitleme çalışmaları bir grup istatistiksel prosedürlerin uygulanmasını içermektedir. Bu uygulama basamakları şunlardır (Kolen ve Brennan, 2004,7):

- 1) Eşitleme amacının belirlenmesi: Eşitleme amacı açıkça belirtilmelidir.

- 2) Alternatif formların yapılandırılması: Alternatif test formları aynı kapsam ve istatistiksel özelliklerde oluşturulmalıdır.
- 3) Veri toplama deseninin seçilmesi: Eşitleme, test formlarının istatistiki olarak nasıl farklılaştığı ile ilgili bilgi sağlayan verilere ihtiyaç duyulmaktadır.
- 4) Veri toplama deseninin uygulanması: Desende belirtildiği şekilde test uygulanıp veri toplanmalıdır.
- 5) Bir veya daha fazla operasyonel eşitleme tanımının belirlenmesi: Eşitlemede formlar arasında nasıl bir ilişki varsayımı seçildiği belirlenmelidir. Örneğin, bu seçim lineer veya lineer olmayan eşitleme yöntemlerinden hangisi olduğuna karar verme şeklinde olabilir.
- 6) Bir veya daha fazla istatistiksel kestirim yöntemi seçilmesi: Lineer ve lineer olmayan eşitleme ilişkileri uygun yöntemler seçilerek kestirilmelidir.
- 7) Eşitleme sonuçlarının değerlendirilmesi: Değerlendirme; test geliştirme, testin uygulanması, istatistiksel prosedürlerin işe koşulması ve eşitleme sonuçlarının özellikleri gibi ögelerin hepsini içermektedir.

D. Test Eşitleme Çeşitleri

Aynı özelliği ölçmek amacıyla hazırlanan bir testin farklı formlarının farklı bireylere uygulanması sıklıkla rastlanan bir durumdur. Örneğin ülkemizde her yıl ya da yılda birden fazla yapılan ve belli süre geçerliği devam eden ALES, YDS, KPSS vb. sınavların farklı formları birbiri yerine kullanılabilir. Eğer bu kullanılan formların puan dağılımları aynı değilse, eşdeğer puanlar elde edilmesi gerekmektedir. Eşdeğer puanların elde edilmesinde literatürde yatay ve dikey eşitleme çeşitlerinin ele alındığı görülmektedir (Hambleton ve Swaminathan, 1985; Crocker ve Algina, 1986; Kolen ve Brennan; 2004; Holland ve Dorans, 2006). Aşağıda bu eşitleme çeşitlerinden bahsedilmektedir.

Benzer yetenek dağılımına sahip iki grubun aldığı bir testin, benzer güçlükte ve güvenilirlikte iki formu arasında yapılan eşitleme **yatay eşitleme** olarak adlandırılmaktadır. Yatay eşitlemede güvenlik vb. nedenlerle bir testin farklı formları uygulanmaktadır. Bu test formlarının birbirine paralel olmaları ve bunun ötesinde testi alan bireylerin yetenek dağılımlarının yaklaşık olarak eşit olması beklenmektedir (Hambleton ve Swaminathan, 1985; Holland ve Dorans, 2006).

Yatay eşitleme ile aynı özelliği ölçen testin, aynı test özelliklerine, aynı güçlüğü sahip olacak şekilde tasarlanan farklı formları üzerinde eşdeğer puanlar elde edilmektedir. Örneğin; aynı yetenek düzeyinde farklı bölgelerde farklı test bataryaları kullanılırsa farklı bataryalar üzerinde puanları eşitlemek gerekmektedir. Bu durumda yatay eşitleme yapılması gerekir (Crocker ve Algina, 1986).

Bununla birlikte farklı güçlük seviyelerinde başarı testi bataryaları hazırlanabilmektedir. Her seviye yetenek dağılımları birbirinden farklı bireyleri belli bir sınıf aralığında test etmektedir. Bu durumda **dikey eşitleme** söz konusudur. Kolen ve Brennan (2004: 373) dikey eşitleme yerine dikey ölçekleme teriminin kullanılmasını önermişlerdir. Çünkü dikey ölçeklemede eşitlemenin amacı olan formların birbiri yerine kullanılabilmesi söz konusu değildir. Dikey eşitlemede benzer yapıları ölçmek için hazırlanmış benzer güvenirlikte; fakat farklı güçlükteki testler aynı ölçek üzerine yerleştirilmektedir. Öğrencilerin eğitimsel başarı veya yeteneklerinin sınıf düzeyi atladıkça, bir yıldan diğer yıla, ne derece büyüdüğü; aynı testin farklı formlarının her yıl uygulanması ile belirlenebilmektedir. Bununla birlikte çok fazla öğrenmenin gerçekleştiği ve çok farklı seviyenin olduğu okul öğrenmelerinde tek bir grup test sorusu (batarya) kullanmak sıkıntılar doğurmaktadır. Şöyle ki; test bataryasının bir seviyesi üçüncü sınıf başı için uygunken; diğer ikinci seviyesi ise üçüncü sınıf sonundakiler için uygun olabilmektedir. Fakat üçüncü sınıfta yavaş ilerleme kaydeden öğrenciler için test bataryasının ikinci seviyesi, hassas bir ölçmeye izin vermeyecek kadar zor olabilmektedir. Dolayısıyla buradaki çözüm yolu, yavaş ilerleyen öğrencilere birinci seviyeyi uygulamaktır. Farklı düzeylerdeki bu öğrencilerin performansını yorumlayabilmek için iki seviyede eşdeğer puanlar elde etmek gerekmektedir. Bu işleme dikey eşitleme adı verilmektedir. Dikey eşitleme ile öğrencilerin yıldan yıla başarı gelişimi karşılaştırılabilmektedir ve yatay eşitlemeye göre daha karmaşıktır (Hambleton ve Swaminathan, 1985; Crocker ve Algina,1986; Kolen ve Brennan; 2004, Holland ve Dorans, 2006).

Genel lineer, eşit yüzdelik ve MTK'ya dayalı yöntemlerde gözlenen veya gerçek puanlara dayalı eşitleme yapılabilmektedir. Gözlenen puan eşitlemede, gerçek puanlar eşitleme fonksiyonu hesaplanmasına doğrudan dahil edilmektedir. Gözlenen puan eşitlemede puan dağılımlarının ortalama ve standart sapması belli bir evren bireyleri için eşit ayarlanmaktadır. Gerçek puan eşitleme ise gözlenen

puan eşitlemeden daha karmaşıktır. Gerçek puan yöntemleri için bir bireyin gözlenen puanı; gözlenen puan ve hatasından oluşmaktadır. Gözlenen puan yöntemleri genellikle en çok kullanılan yöntemler olmuştur; çünkü kolay ve en eski yöntemlerdendir (Kolen ve Brennan, 2004).

E. Test Eşitleme Desenleri

Madde parametreleri kestirimleri MTK'ya dayalı ölçek dönüştürme problemini karmaşıktır. Parametre kestirim süreci çok çeşitli veri toplama desenlerine bağlıdır. BILOG ve LOGIST programları parametre kestirimlerinde kullanılan programlardır ve kullanıcı başka bir değer seçmediği sürece yetenek kestirimleri ortalaması 0, standart sapması 1 olacak şekilde ölçeklenmektedir (Kolen ve Brennan, 2004:166). Test eşitleme desenleri; tek grup deseni, dengelenmiş tek grup deseni, eşdeğer grup deseni ve ortak madde deseni olmak üzere dört grupta incelenmektedir (Kolen ve Brennan, 2004; Holland ve Dorans, 2006).

• Tek Grup Deseni

Tek grup deseninde, aynı bireylere bir testin X ve Y gibi farklı iki formu uygulanmaktadır. Bu desenin avantajı tek bir grup olması nedeniyle grupların yetenek farkından kaynaklanacak hatayı önlemesidir. Dezavantajı ise formların verilmiş sırasının test performanslarına yapacağı etkidir; çünkü aynı bireyler iki formu da almaktadır. Dolayısıyla birinci formun uygulanmasından sonra yorgunluk vb. nedenlerle testi alan bireylerin performanslarında düşme yaşanabilir ve bu durum ikinci formdan elde edilecek puanları etkileyebilir. Tek grup deseni Tablo 1.3'te görülmektedir:

Tablo 1.3: Tek Grup Deseni

Evren	Örnekleme	X	Y
P	1	√	√

• Dengelenmiş Tek Grup Deseni

Tek grup desenindeki sıra etkisini kontrol edebilmek için dengelenmiş tek grup deseni geliştirilmiştir. Bu desende her bir alt gruba formlar farklı sırada uygulanmaktadır. Birinci alt grup önce X formunu (X1), daha sonra Y formunu (Y2)

alırken; ikinci alt grup önce Y formunu (Y1), daha sonra X formunu (X2) almaktadır (Kolen ve Brennan, 2004; Holland ve Dorans, 2006).

Dengelenmiş tek grup deseninde tüm bireyler için her iki formdaki parametreler birlikte kestirilebilir; çünkü iki formu alan kişiler aynıdır ve parametre kestirimlerinin aynı ölçek üzerinde olduğu varsayılır. Eğer iki formun parametreleri aynı ölçekleme kuralı kullanılarak ayrı olarak kestirilirse, aşağıda eşdeğer grup deseninde tartışılan mantık takip edilerek parametre kestirimlerinin aynı ölçek üzerinde olduğu varsayılır (Kolen ve Brennan, 2004). Dengelenmiş tek grup deseni Tablo 1.4'te gösterilmiştir:

Tablo 1.4: Dengelenmiş Tek Grup Deseni

Evren	Örneklem	X ₁	Y ₁	X ₂	Y ₂
P	1	√			√
P	2		√	√	

• Eşdeğer Gruplar Deseni

Eşdeğer (Equivalent=Random) gruplar deseninde aynı evrenden rastgele seçilmiş iki örneklem grubuna bir testin X ve Y formları uygulanmaktadır. Gruplardaki bireyler bir tane test formu aldıkları için bu desende sıra etkisi oluşmamaktadır. Dikkat edilmesi gereken nokta, örneklem gruplarını eşdeğer seçebilmektir. Eşdeğerliği sağlamak için evrenden rastgele seçim yapılabilir (Kolen ve Brennan, 2004; Holland ve Dorans, 2006). Eşdeğer gruplar deseni Tablo 1.5'te görülmektedir:

Tablo 1.5: Eşdeğer Gruplar Deseni

Evren	Örneklem	X	Y
P	1	√	
P	2		√

Eşdeğer gruplar deseninde Form X'in parametreleri Form Y'nin parametrelerinden bağımsız olarak kestirilebilmektedir. Bağımsız kestirimlerde yetenek için aynı ölçekleme kuralı (ortalama=0, standart sapma=1) kullanıldığında, başka bir dönüşüme gerek duymadan iki formun parametre kestirimlerinin aynı ölçek üzerinde olduğu varsayılmaktadır. Burada başka dönüşüme gerek duyulmadığı varsayılmaktadır; çünkü iki grup da rastgele seçilmiş eşdeğer gruplardır. Bazı nedenlerden dolayı iki form için farklı ölçekleme kuralları kullanıldığında, Θ

kestirimlerinin ortalama ve standart sapma deęerleri eřitlik (1.1c) ve (1.2b)'deki (bkz.sf.29) Θ -parametrelerinin ortalama (μ) ve standart sapma (σ) deęerleri yerine kullanılabilir (Kolen ve Brennan, 2004:166).

Eřdeęer grup deseni test eřitlemede byk rneklem istemektedir. Bu kořul saęlandığında eřdeęer grup deseni iyi bir seęenek olarak kabul edilmektedir. Bu desenin avantajı tek grup deseninde her ęrencinin iki formu aldıęı durumlarda ortaya ıkan sıra etkisinden etkilenmemesidir. Ancak gruplar aynı olmadığı iin yetenek daęılımındaki farklılıklar eřitleme srecinde yanlılık ortaya ıkarabilir (Kolen ve Brennan, 2004; Holland ve Dorans, 2006).

- **Eřdeęer Olmayan Gruplarda Ortak Test Deseni**

Eřdeęer olmayan gruplarda ortak test deseni (Non-equivalent Groups with Anchor Test Design), genellikle test gvenlięi ve dięer uygulama sorunları yznden birden fazla formun aynı test tarihinde uygulanamadıęı durumlarda kullanılmaktadır. Bu desende P ve Q gibi farklı iki evrenden seilmiř iki eřdeęer olmayan rneklem bir grup ortak maddeye sahip X ve Y formlarını almaktadır. Bu desende ortak maddeler i ve dıř olmak zere iki řekilde ifade edilebilmektedir. İ ortak maddeler, test formunun istatistiki zelliklerini ve kapsamını temsil edecek řekilde seilmektedir. Eęer i ortak maddelerin puanı test formundaki ęrenci puanına katkı yapmıyorsa, ortak maddeler dıř ortak maddeler olarak nitelendirilmektedir. Tipik olarak dıř ortak maddeler ayrı bir zamanda ayrı bir test formu olarak verilmektedir (Kolen ve Brennan, 2004; Holland ve Dorans, 2006). Eřdeęer olmayan gruplarda ortak test deseni Tablo 1.6'da grlmektedir:

Tablo 1.6: Eřdeęer Olmayan Gruplarda Ortak Test Deseni

Evren	rneklem	X	A	Y
P	1	√	√	
Q	2		√	√

P ve Q evrenleri farklı ($P \neq Q$) ve eřdeęer olmayan gruplar durumundayken ortak test olan A'nın istatistiki grevi yanlılıęı ortadan kaldırmak ve duyarlılıęı artırmaktır. A, X ve Y formlarının minyatr versiyonu olduęunda (rneęin, daha az gvenilir bir minitest iki testi birbirine baęlayan bir yapıyı lebilir) P ve Q evrenlerinin eřdeęer olmamasından kaynaklı yanlılıęı ortadan kaldırmaktadır. X ve Y formlarıyla

aynı yapıyı ölçmeyen veya onlarla yüksek korelasyon vermeyen bir ortak test (A), yanlılık giderme ve duyarlılığı artırma konusunda o kadar kullanışlı olmamaktadır. Güvenlik nedenleriyle testin birden fazla uygulamasının yapılamadığı durumlarda P ve Q evrenlerinde testi alanlardan elde edilebilecek bilgi bilişsel olmayan demografik bilgilerdir. Test eşitleme açısından bu bilgi tatmin edici bir ilişki sağlamamaktadır (Holland ve Dorans, 2006).

Eşdeğer olmayan gruplarda ortak test deseninde Form Y'nin madde parametreleri ve yetenekleri genellikle Form Y ilk uygulandığında kestirilmektedir. Sonuç olarak, Form X Form Y'ye eşitlenirken sadece Form X parametrelerinin kestirimine ihtiyaç bulunmaktadır. Çünkü, Form X'i alan bireyler Form Y'yi alanlarla eşdeğer kabul edilmemektedir. İki kestirim için parametre kestirimleri aynı ölçekte değildir. Buna karşın iki form için de ortak olan bir madde kümesi bulunmaktadır. Bu ortak maddelerin madde parametre kestirimleri ölçek dönüştürme kestirimleri için kullanılabilir (Kolen ve Brennan, 2004).

Eşdeğer olmayan gruplarda ortak test deseninde, iki örnekleme bireyler arasında küçük farklılıklar olduğunda, tüm lineer ölçekleme ve eşitleme yöntemleri benzer sonuçlar verme eğiliminde olmaktadır (aynı durum lineer olmayan ölçekleme ve eşitleme yöntemlerinde de geçerlidir). Eşdeğer olmayan gruplarda ortak test deseni neredeyse eşdeğer gruplarda ortak test deseni olduğunda ortak teste olan ihtiyaç azalmaktadır. Aksi halde iki örneklem çok farklı ise, ortak test, grupların yetenek farklılıkları ile test formlarının farklılıklarını birbirinden ayırmada kritik öneme sahiptir (Holland ve Dorans, 2006:201).

• Eşdeğer Gruplarda Ortak Test Deseni

P ve Q evrenleri birbirine eşit olduğunda ($P=Q$), eşdeğer olmayan gruplarda ortak test deseni, eşdeğer gruplarda ortak test deseni (Equivalent Groups with Anchor Test Design) adını alır. Bu durumda ortak testin rolü değişir ve grup farklılıklarını ayarlamakta değil; ilgili parametrelerin kestiriminde hassasiyet kazanmak için kullanılmaktadır. Bu özel durum için A ortak testinin, X ve Y formları ile aynı yapıyı ölçmesi ve hatta test puanı olması da gerekmemektedir. Sadece A'nın, X ve Y formlarının her ikisi ile de korelasyon vermesi önemlidir. Lord'a göre, eşdeğer gruplarda ortak test deseni, hem eşdeğer gruplar deseni hem de tek grup deseninin olumlu taraflarını almak için önerilmiştir. Lord (1950) araştırmasında, eşdeğer gruplarda ortak test deseninde ortak testin kullanılma şeklinin sonuçların

doğru şekilde elde edilmesinde dramatik etkisi olabildiğini göstermiştir. Örneğin, ortak ölçümler ile testler arasındaki korelasyona bakmaksızın bazı yöntemlerde daha kararlı kestirimler elde etmiştir (Holland ve Dorans, 2006:199). Eşdeğer gruplarda ortak test deseni Tablo 1.7’de verilmiştir:

Tablo 1.7: Eşdeğer Gruplarda Ortak Test Deseni

Evren	Örneklem	X	A	Y
P	1	√	√	
P	2		√	√

Eşdeğer grup deseni, standart eşitleme hatası ile ölçülen doğruluğa erişmek için en büyük örneklem sayısına ihtiyaç duyar. Teorik olarak ideal tasarımın dış ortak teste sahip bir eşdeğer grup deseni olduğu söylenebilir. Eğer ortak test diğerlerinden daha sonra uygulanırsa muhtemel sıra etkisi sadece ortak testi etkiler. Eğer bu etkiler ciddi boyuttaysa, ortak test ihmal edilebilir. Bu durumda geriye basit bir eşdeğer tasarım kalır. Bu noktada herhangi bir sıralama etkisi mümkün değildir. Eğer ortak test diğer iki teste dahilse (iç ortak test), kapsam ve sıralama etkileri ortaya çıkar ve incelenmeleri gerekir (Holland ve Dorans, 2006).

F. Ortak Maddelerin Özellikleri

Farklı test formlarını bağlamada kullanılan ortak (anchor= common) maddelerin sayısı eşitlemenin duyarlılığı açısından önemli görülmektedir. Ortak madde sayısı tüm testin en az %20-25'i kadar olmalıdır (Hambleton ve Swaminathan,1985; Hambleton ve diğ., 1991; Cook ve Eignor, 1991). Angoff (1971)'a göre ortak maddelerin tüm testteki madde sayısının %20'si kadar olması yeterli görülmektedir. Kolen ve Brennan (1995) ise 40 ve daha fazla maddeli testlerde ortak madde sayısının %20 olmasını önermişlerdir. Bazı araştırmalar ortak maddelerin sayısının en az 15 olması gerektiğini ileri sürmektedir (McKinley ve Recklase, 1981). Ancak, tek boyutluluk varsayımı karşılandığında daha az sayıda maddenin kabul edilebilir olduğu ifade edilmektedir (Cook ve Paterson, 1987; Kolen ve Brennan, 2004). Ortak madde sayısında artış olması hatayı azaltarak eşitleme sonuçlarını geliştirici etki yapmaktadır (Hanson ve Bèguin, 2002; Kolen ve Brennan, 2004). Holland ve Dorans (2006)'a göre de uzun ortak testler kısa ortak testlere göre test eşitleme için daha avantajlıdır; çünkü güvenilirlikleri daha yüksektir.

Grup farklılıklarını doğru bir şekilde yansıtabilmek için ortak madde kümesinin toplam test formlarını kapsam ve istatistiki özellikler açısından oransal olarak temsil etmesi gerekmektedir. Angoff'a (1971) göre, ortak madde kümesi toplam test formunun küçük bir alt grubu olmalıdır. Ortak maddelerin aynı zamanda X ve Y formlarına benzer şekilde davranması gerekmektedir. Benzer davranışı sağlamaya yardımcı olması için her ortak madde her iki formda da benzer bir yerleşime (madde numarasına) sahip olmalıdır. Ek olarak ortak maddeler X ve Y formlarında tümüyle aynı olması gerekmektedir. Örneğin, sözcük değişiklikleri olmamalıdır (Kolen ve Brennan, 2004). Bununla birlikte bazı araştırmalarda bu görüşe zıt sonuçlara ulaşılmıştır. Örneğin, Gao ve diğ. (1999) araştırmalarında ortak maddelerin kapsamı temsil etmemesinin eşitleme doğruluğunu etkilemediği sonucuna ulaşmıştır. Aynı şekilde Hannick ve Huang (2002)'in çalışmasında, ortak maddelerin istatistiksel olarak ölçeği temsil etmeme durumunun eşitlemenin standart hatasında (Standart Error of Equating- SEE) büyük bir etki yapmadığı; içerik ve güçlük bakımından testin bütününe daha yetersiz temsil gücüne sahip ortak maddelerin ölçekleme sonuçları üzerinde minimal bir etkiye sebep olduğu belirtilmiştir.

Cook ve Paterson (1987) ise yaptıkları araştırmada, ortak maddelerin bir formdan diğerine yerleri değiştirildiğinde özellikle belli tipte maddeler için (Ör. analitik maddeler) eşitleme sonuçlarının çok fazla etkilendiğini bulmuştur. Diğer madde türleri ise en azından ortalama bir etkiye neden olmuştur. Araştırmada ayrıca, ortak maddelerin kapsamla eşleştiği varsayıldığında, madde güçlük ve ayırt edicilik parametrelerinin eşitlemede çok fazla bir etki yaratmadığı belirtilmiştir. Aksine ortak maddelerin uzunluğu ve yerleşim yerlerinin daha önemli olduğu vurgulanmıştır.

Ortak testin en önemli özellikleri zamana göre eşlik, kararlılık ve eşitlenmiş iki testin puanları ile korelasyonudur. Korelasyonun mümkün olduğunca yüksek olması önemli görülmektedir. Eşitlenecek testlerle yüksek korelasyon veren ortak testin kullanılması hem elde edilen kestirimlerin hem de eşitlenmiş puanların örneklem hatalarını büyük ölçüde azaltmaktadır (Lord, 1955:194; Holland ve Dorans, 2006). Bununla birlikte iç ortak testler, dış ortak testlere göre daha avantajlıdır; çünkü diğer testlerle olan parça-bütün ilişkisi, eşitlenen testler ile korelasyonun yüksek olmasını garanti etmektedir (Holland ve Dorans, 2006).

Farklı set ortak testler farklı eşitleme sonuçlarına neden olmaktadır. Örneğin Cook, ve diğ. (1985) yaptıkları araştırmada biyoloji testinin iki farklı formunu dört farklı ortak test ile eşitlemiştir. Sonuçta, farklı ortak madde setleri kullanıldığında birbirinin aynısı eşitleme sonuçları elde edilememiştir. Bu duruma istisnai bir yaklaşım getiren Cook ve Paterson (1987) araştırmalarında, eşitleme yapılacak grupların yetenek dağılımı birbirine benzer olduğunda, seçilen farklı ortak testlerin birbirine benzer eşitleme sonuçları verdiğini bulmuşlardır. Grupların yetenek dağılımı birbirinden farklı olduğunda ise, farklı ortak testler birbirinden uzak eşitleme sonuçları elde edilmesine neden olmuştur. Dolayısıyla ortak madde setini seçerken özel dikkat gösterilmesi gerekmektedir.

G. Madde Tepki Kuramına Dayalı Test Eşitleme Yöntemleri

Literatür incelendiğinde MTK'ya dayalı test eşitleme yöntemleri üç gruba ayrılmaktadır. Bunlar, parametrelerin ayrı kestirimine dayalı lineer ölçek dönüştürme veya moment yöntemleri, madde karakteristik eğrisi (MKE) dönüşüm yöntemleri ve parametrelerin eşzamanlı kestirildiği eşzamanlı eşitleme (concurrent equating) veya eşzamanlı kalibrasyon (concurrent calibration) yöntemleridir. Lineer eşitleme yöntemleri Marco'nun (1977) ortalama-sigma, Linn ve diğ.'nin (1981) sağlam ortalama-sigma ve Loyd ve Hoover'ın (1980) ortalama-ortalama yöntemlerine sahiptir. Madde karakterisitik eğrisi (MKE) dönüştürme yöntemlerinde ise Haebara'nın (1980) ve Stocking ve Lord'un (1983) yöntemleri bulunmaktadır. Eşzaman kalibrasyon (Kim ve Kohen, 1998; Peterson ve diğ.,1983) veya diğer adıyla eşzaman ölçekleme (Chu ve Kamata, 2000) yöntemi ise diğer bir MTK'ya dayalı eşitleme yöntemidir.

Lineer ve MKE eşitleme yöntemleri iki aşamadan oluşan **ayrı eşitleme/ ayrı kalibrasyon** prosedürüne sahiptir. Birinci aşamada kestirim programının iki kez çalıştırılmasıyla iki test formundan ayrı ayrı madde parametreleri kestirilmektedir. Farklı test formlarından kestirilen bu madde parametreleri aynı ölçek üzerinde değildir. Dolayısıyla ikinci aşamada, gizil değişken/yetenek ölçeği ortalaması ve standart sapması ayarlanarak sınırlandırılmakta; madde parametre kestirimleri ortak bir ölçek üzerinde kalibre edilmektedir. Ayrı kestirim yönteminde ortak maddelerin iki gruptaki madde parametre kestirimleri ölçek dönüştürme kestiriminde kullanılmaktadır; bu şekilde bir formun madde parametreleri diğer formun madde parametrelerinin ölçeğine yerleştirilmektedir (Kolen ve Brennan,

1995). Ayrı kalibrasyon sonrası eşitleme, ortak maddelerde bireysel olarak yer alabilecek bir takım problemleri de ortaya çıkarma potansiyeline sahiptir (Hanson ve Béguin, 2002; Kolen ve Brennan, 2004). Aynı evrenden seçilmiş (eşdeğer) gruplar ortak maddeler içeren iki test formunu aldığımda madde parametrelerinin ölçeklenmesine gerek olmasa da, yine de faydalı olabilmektedir. Eşdeğer gruplarda madde parametrelerinin ölçeklenmesi, farklı örneklem gruplarının iki formu almasıyla ortaya çıkan örneklem hatasından kaynaklı kestirim hatasını azaltabilmektedir (Hanson ve Béguin, 2002).

Eşzamanlı eşitlemede ise, iki forma ait madde parametreleri tek bir form gibi düşünülerek, aynı anda programın tek bir kez çalıştırılmasıyla kestirilmektedir. Ortak maddelerin dahil olması tüm madde parametreleri ölçeğini birleştirmekte ve elde edilen madde parametreleri otomatik olarak aynı ölçek üzerine yerleşmektedir (Kolen, 2006). Kim ve Kohen (1998) eşzamanlı kalibrasyonun madde ayırt edicilik parametresini iyileştirmede daha iyi olduğunu ve ortak madde sayısı arttığında ayrı kalibrasyon sonrası eşitleme ile benzer sonuçlar ürettiğini ortaya koymuştur. Hanson ve Béguin (2002), yapılan bu çalışmanın bir sınırlılığı üzerinde durmuş ve ayrı ve eş zaman kalibrasyon için farklı MTK yazılım programları kullanıldığına dikkati çekmiştir. Hanson ve Béguin (2002), daha fazla bilgi sağlamak amacıyla yaptıkları simülasyon çalışmasında iki karakteristik eğri ve iki moment yöntemini ayrı ve eşzamanlı kalibrasyon kullanarak karşılaştırmışlardır. Araştırma sonucunda, eşzamanlı kalibrasyonun karakteristik eğrisi yöntemlerinden daha doğru ve değişmez eşitleme sonuçları verdiğini; iki moment yönteminin ise karakteristik eğrisi ve eşzamanlı kalibrasyon kadar iyi bir performansı olmadığını belirtmişlerdir. Ayrıca eşzaman kalibrasyonun üstün çıkma nedeni olarak, tek seferde iki grubun kalibrasyonunu yaparak daha büyük örnekleme sahip olmasını göstermişlerdir.

Bununla birlikte, pratikte ayrı kestirim yöntemleri daha tercih edilebilirdir; çünkü tek boyutluluk varsayımının ihlalinin MTK ölçekleme sonuçları üzerindeki etkilerini hafifletici etkiye sahip olmakla birlikte aynı maddeye ilişkin farklı madde parametre kestirimlerinin (Ör: farklı sınıf düzeylerinde) karşılaştırılmasına imkan vermektedir. Eş zamanlı kestirimde bir maddeye ilişkin tek bir parametre kestirildiğinden maddenin farklı düzeylerde nasıl davrandığını karşılaştırmak sorun olabilmektedir; buna ek olarak tek boyutluluk varsayımının karşılanma gerekliliği daha katıdır;

ancak tek bir yeteneğin ölçülmesi başarı testleri için çoğu zaman zordur. Eş zaman eşitlemede, çok sayıda madde ile çalışılması gerekliliği ve pek çok maddenin ulaşılamamış olarak kodlanması gibi durumlar nedeniyle de bazen uyum problemleri yaşanabilmektedir. Sonuç olarak eşzaman eşitleme ayrı kalibrasyon yönteminin yanında yardımcı ek bir yöntem olarak kullanılabilir (Kolen ve Brennan; 2004: 391). Aşağıda ayrı kalibrasyon yöntemlerinden iki moment yöntemi ortalama-sigma (O-S), ortalama-ortalama (O-O) ve iki karakteristik eğrisi yöntemi Haebara (HB) ve Stocking Lord'a (S-L) ilişkin açıklamalara yer verilmiştir.

- **Lineer Ölçek Dönüştürme (Moment) Yöntemleri**

Ortak madde deseninde ölçek formlarını dönüştürmenin en açık ve basit yolu ortak maddelerin madde parametre kestirimlerinin ortalama ve standart sapmalarını (1.1) ve (1.2) eşitliklerindeki parametreler yerine koymaktır. Dönüştürmeden sonra madde parametre kestirimleri sıklıkla kalibre edilmiş olarak kabul edilmektedir. Bu yöntemlere lineer ölçek dönüştürme veya moment yöntemleri adı verilmektedir.

a) Ortalama-Sigma Yöntemi

Marco (1977) tarafından önerilen ortalama-sigma yöntemi ortak maddelerden kestirilen b (güçlük) parametrelerinin ortalama ve standart sapmalarını (1.1a) ve (1.2a) eşitliklerindeki parametreler yerine konulmasına dayalı bir yöntemdir. Bu sayede lineer eşitleme için A ve B eşitleme katsayıları elde edilmektedir (Kolen ve Brennan, 2004:167).

$$A = \frac{\sigma(b_X)}{\sigma(b_Y)} \quad (1.1a)$$

$$B = \mu(b_X) - A \mu(b_Y) \quad (1.2a)$$

$\sigma(b_X)$: X formu ortak maddelerinin güçlük parametreleri standart sapması

$\sigma(b_Y)$: Y formu ortak maddelerinin güçlük parametreleri standart sapması

$\mu(b_X)$: X formu ortak maddelerinin güçlük parametreleri ortalaması

$\mu(b_Y)$: Y formu ortak maddelerinin güçlük parametreleri ortalaması

A : Eşitleme denkleminin eğimi

B : Eşitleme denklemi sabiti

Madde güçlük kestirimleri dağılımlarının ilk iki momentini kullanan simetrik bir yöntem olan ortalama-sigma yönteminde lineer dönüşüm katsayıları, A ve B elde edilerek; birinci kalibrasyondan kestirilen madde güçlük parametreleri ortalama ve standart sapması, ikinci kalibrasyondan kestirilen madde güçlük parametreleri dönüştürülmüş dağılımının ortalama ve standart sapmasına eşit olmaktadır. Bununla beraber zayıf kestirilmiş madde güçlüklerinin örneklem momentlerinin hesaplanmasında ciddi bir etkisi olabilir ve üretilen bir lineer dönüşüm madde güçlüklerinin büyük çoğunluğuna çok iyi uymayabilir. Cook ve diğ. (1979) momentlerin hesaplanmasında kullanılan madde güçlüklerinin ranjını sınırlandırarak bu probleme çözüm getirmeye çalışmıştır. Bejar ve Wingersky (1981) sağlam yöntemler ile momentlerin kestiriminde aykırı noktalara daha küçük ağırlıklar veren daha detaylı ve özenli bir yaklaşım kullanmışlardır (Stocking ve Lord, 1983: 203).

Linn ve diğ. (1980,1981) ise kestirilen madde güçlüklerinin standart hatası ile ters orantılı ağırlıklandırılmış momentleri önermişler ve bu şekilde aykırı değerlerin etkilerini azaltmaya çalışmışlardır. Madde parametre kestirimlerinin ağırlıklarını standart hatalarına göre işleyen bu yöntem **sağlam ortalama-sigma (robust mean-sigma)** yöntemi olarak adlandırılmaktadır (Stocking ve Lord, 1983: 203; Kolen ve Brennan, 2004:168). Bejar ve Wingersky (1981), standart hatalardan bağımsız olarak tüm aykırı değerleri aynı şekilde değerlendirirken; Linn ve diğ. (1980), aynı standart hataya sahip tüm noktaları aykırılık durumlarını dikkate almadan aynı şekilde değerlendirmektedir. Lord ve Stocking ise aykırı değerlerin ortalama ve standart sapma hesaplamalarına dahil edilmesini daha sağlam bir yol olarak önermişlerdir (Stocking ve Lord, 1983).

Sağlam ortalama-sigma yönteminde, ortalama-sigma yönteminden farklı olarak parametre kestirimlerinin farklı standart hatalara sahip olabileceği gerçeği ele alınmaktadır. Eşitlenecek iki testte aynı ortak maddeye ait b parametreleri çifti aralarındaki varyans büyükse düşük ağırlık, varyansları küçükse büyük ağırlık almaktadır. Daha sonra kaç tane ortak madde varsa hepsi işe koşularak ölçek ağırlıklandırması hesaplanmaktadır. Ağırlıklandırılmış madde parametre kestirimleri hesaplanarak, bu kestirimlerin ortalama ve standart sapmaları belirlenmektedir. Elde edilen verilerden hareketle son olarak α ve β katsayıları hesaplanmaktadır (Hambleton ve diğ., 1991:132-133). Stocking ve Lord (1983)

madde ayırtedicilikleri ilişkisini gözardı eden robust yöntemlerin performansından tatmin olmamış ve alternatif olarak karakteristik eğrisi yöntemlerini önermişlerdir. Gerçekten lineer ve MKE eşitleme yöntemlerinin karşılaştırıldığı araştırmalarda (Stocking ve Lord, 1983; Cohen ve Kim, 1998; Uysal, 2014) en büyük eşitleme hatasının sağlam ortalama-sigma yönteminde elde edilmesi ve bu yöntemin çok fazla tavsiye edilmemesi dikkati çeken bir nokta olarak göze çarpmaktadır.

b) Ortalama-Ortalama Yöntemi

Loyd ve Hoover (1980) tarafından tanımlanan diğer bir lineer ölçek dönüştürme yöntemi ortalama-ortalama yöntemidir. Bu yöntemde ortak maddelerden kestirilen a (ayırt edicilik) parametrelerinin ortalaması (1.1b) eşitliğindeki parametreler yerine konularak A eşitleme katsayısı elde edilmektedir. Daha sonra, ortak maddelerden kestirilen b parametrelerinin ortalaması (1.2a) eşitliğindeki parametreler yerine konularak B katsayısı elde edilmektedir (Kolen ve Brennan, 2004:167).

$$A = \frac{\mu(a_Y)}{\mu(a_X)} \quad (1.1b)$$

$$B = \mu(b_X) - A \mu(b_Y) \quad (1.2a)$$

$\mu(ax)$: X formu ortak maddelerinin ayırt edicilik parametreleri ortalaması

$\mu(ay)$: Y formu ortak maddelerinin ayırt edicilik parametreleri ortalaması

$\mu(bx)$: X formu ortak maddelerinin güçlük parametreleri ortalaması

$\mu(by)$: Y formu ortak maddelerinin güçlük parametreleri ortalaması

A : Eşitleme denkleminin eğimi

B : Eşitleme denklemi sabiti

Son olarak A ve B katsayıları (1.3)-(1.6) eşitliklerinde yerine konularak yeniden ölçeklenmiş parametre kestirimleri elde edilmektedir.

$$\theta_{X_i} = A \theta_{Y_i} + B$$

θ_{xi} : i bireyinin X testindeki yeteneği (kestirilen yetenek)

θ_{yi} : i bireyinin Y testindeki yeteneği (gerçek yetenek)

A ve B doğrusal eşitlikteki katsayılarıdır ve iki ölçekteki madde parametrelerinin ilişkileri aşağıdaki gibidir (Kolen ve Brennan, 2004):

$$a_{xj} = \frac{a_{yj}}{A} \quad (1.4)$$

$$b_{xj} = A b_{yj} + B \quad (1.5)$$

$$c_{xj} = c_{yj} \quad (1.6)$$

Burada a_{xj} , b_{xj} ve c_{xj} j maddesinin Form X üzerindeki madde parametreleridir; a_{yj} , b_{yj} ve c_{yj} ise j maddesinin Form Y üzerindeki madde parametreleridir. Şans parametresi c ise, (1.6) numaralı eşitlikte görüleceği gibi, ölçek dönüşümünden bağımsızdır.

A ve B eşitleme katsayıları θ yetenek düzeyi cinsinden aşağıda verilen (1.1c) ve (1.2b) denklemleri ile de elde edilebilmektedir:

$$A = \frac{\sigma(\theta_X)}{\sigma(\theta_Y)} \quad (1.1c)$$

$$B = \mu(\theta_X) - A\mu(\theta_Y) \quad (1.2b)$$

Kestirimler parametrelerin yerine kullanıldığında veya MTK modeli tam olarak geçerli olmadığında (1.1) ve (1.2) eşitlikleri muhakkak kullanılmak zorunda değildir. Böyle olunca ortalama-sigma ve ortalama-ortalama yöntemleri genellikle farklı sonuçlar üretmektedir. b parametresi kestirimleri a parametresi kestirimlerine göre daha kararlı olduğu için ortalama-sigma yöntemi bazen ortalama-ortalama yöntemine tercih edilmektedir. Ancak, Baker ve Al-Karni (1991), ortalamalar genellikle standart sapmalardan daha kararlı değerler ürettiği için ortalama-ortalama yöntemini daha tercih edilebilir bulmaktadır. İki moment yönteminden

hangisinin tercih edileceği konusunda yeterli araştırma bulunmamaktadır. Bu nedenle iki yöntemin de kullanılarak karşılaştırılması önerilmektedir (Kolen ve Brennan, 2004:167).

- **Karakteristik Eğrisi Dönüşüm Yöntemleri**

Daha önce belirtilen yöntemlerde doğacak olası bir problem a, b ve c parametre kestirimlerinin çeşitli kombinasyonlarının bireylerin puan aldığı yetenek ranjı boyunca neredeyse birbirinin aynı madde karakteristik eğrileri üretebilmesidir. Örneğin, iki kestirimde b parametreleri farklı olan maddeler benzer madde karakteristik eğrilerine sahip olabilmektedir. Bu durumda, ortalama-sigma yöntemi, iki kestirimdeki madde karakteristik eğrileri benzer olsa da, b parametre kestirimleri arasındaki farklılıktan aşırı şekilde etkilenmektedir. Çünkü şimdiye kadar ölçek dönüştürme yöntemlerinde tüm madde parametre kestirimleri ayrı ele alınmıştır. Bu probleme çözüm olarak Haebara (1980) ve ona benzer şekilde Stocking ve Lord (1983) tüm madde parametrelerini birlikte ele alan bir yöntem geliştirmişlerdir: Geliştirilen iki yöntem karakteristik eğrisi yöntemleri adını almaktadır (Kolen ve Brennan, 2004).

$$P_{ij}(\theta_{Xi}; a_{Xj}, b_{Xj}, c_{Xj}) = P_{ij}(A\theta_{Yi} + B; \frac{aY_j}{A}, Ab_{Yj} + B, cY_j) \quad (1.7)$$

Eşitlik 1.7'ye göre puanları raporlaştırmak için kullanılan ölçeğe bakılmaksızın verilen bir yetenek düzeyindeki bireylerin belli bir maddeyi doğru cevaplama olasılığı aynıdır. Kestirimler (1.7) eşitliğindeki parametrelerin yerine kullanıldığında, herhangi bir A ve B için tüm maddeler ve bireyler boyunca eşitlik olacağının garantisi yoktur. Bu eksiklik karakteristik eğri yöntemleri ile belirtilmektedir (Kolen ve Brennan, 2004). Karakteristik eğri yöntemleri iki ve üç parametrelili modellerle kullanım için en uygun yöntem olarak görülmektedir (Hambleton ve Swaminathan, 1985:223). Bu yöntemler madde parametrelerinin kestirilmesinde hataya daha dirençli kabul edilmektedir (Kaskowitz ve De Ayala, 2001). Ayrıca lineer eşitleme yöntemlerinden farklı olarak eşitleme katsayılarını belirlemede madde parametrelerinin tümü kullanılmaktadır.

a) Haebara'nın Karakteristik Eğri Yöntemi

Haebara (1980) tarafından kullanılan yöntemde madde karakteristik eğrileri arası fark, belli bir yetenek düzeyindeki bireyler için, her bir maddenin madde karakteristik eğrileri arasındaki farkın karelerinin toplamını ifade etmektedir.

Her θ_i için karesi alınmış farklar ve farkların maddeler üzerinden toplamı aşağıdaki gibi hesaplanabilir (Kolen ve Brennan, 2004):

$$H_{diff}(\theta_i) = \sum_{j:V} \left[p_{ij}(\theta_{xi}; \hat{a}_{Xj}, \hat{b}_{Xj}, \hat{c}_{Xj}) - p_{ij}\left(\theta_{xi}; \frac{\hat{a}_{Yj}}{A}, A\hat{b}_{Yj} + B, \hat{c}_{Yj}\right) \right]^2 \quad (1.8)$$

$p_{ij}(\theta_{xi}; \hat{a}_{Xj}, \hat{b}_{Xj}, \hat{c}_{Xj})$: X testi madde karakteristik fonksiyonu

$p_{ij}\left(\theta_{xi}; \frac{\hat{a}_{Yj}}{A}, A\hat{b}_{Yj} + B, \hat{c}_{Yj}\right)$: Y testi dönüştürülmüş madde karakteristik fonksiyonu

V : Ortak madde sayısı

A : Eşitleme denkleminin eğimi

B : Eşitleme denklemi sabiti

Eşitlik 1.8'de toplama işlemi (j:V) ortak maddeler üzerinden yapılmaktadır. Bu eşitlikte, eşitlenecek iki ölçek üzerindeki madde karakteristik eğrileri arasındaki farkın karesi alınarak toplanmaktadır. Daha sonra H_{diff} , testi alan bireyler üzerine dağıtılmaktadır. Kestirim işlemi aşağıdaki kriteri asgariye çekecek A ve B'nin bulunması ile devam etmektedir;

$$H_{crit} = \sum_i H_{diff}(\theta_i) \quad (1.9)$$

Verilen 1.9 eşitliği tüm testi alan bireyler için hesaplanmaktadır.

b) Stocking ve Lord'un Karakteristik Eğri Yöntemi

Haebara'nın yöntemi aksine Stocking ve Lord (1983) her bir maddenin madde karakteristik eğrileri toplamları farkının karesini kullanmışlardır (Kolen ve Brennan, 2004).

$$SL_{diff}(\theta_i) = \left[\sum_{j:V} p_{ij}(\theta_{xi}; \hat{a}_{xj}, \hat{b}_{xj}, \hat{c}_{xj}) - \sum_{j:V} p_{ij} \left(\theta_{xi}; \frac{\hat{a}_{yj}}{A}, A\hat{b}_{yj} + B, \hat{c}_{yj} \right) \right]^2 \quad (1.10)$$

$p_{ij}(\theta_{xi}; \hat{a}_{xj}, \hat{b}_{xj}, \hat{c}_{xj})$: X testi madde karakteristik fonksiyonu

$p_{ij} \left(\theta_{xi}; \frac{\hat{a}_{yj}}{A}, A\hat{b}_{yj} + B, \hat{c}_{yj} \right)$: Y testi dönüştürülmüş madde karakteristik fonksiyonu

V : Ortak madde sayısı

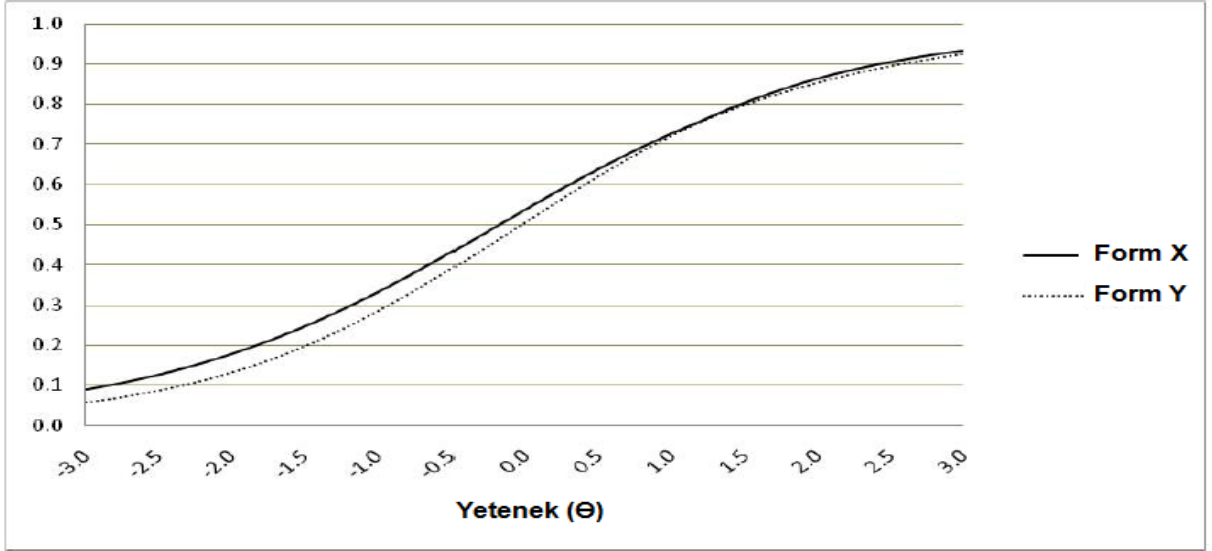
A : Eşitleme denkleminin eğimi

B : Eşitleme denkleminin sabiti

Stocking ve Lord (1983) yaklaşımında, kare alınmadan önce, her parametre kestirim kümesinin maddeleri üzerinden toplama işlemi yapılmaktadır. MTK fonksiyonu 1.11'de verilmiştir.

$$\tau(\theta_i) = \sum_j p_{ij}(\theta_i) \quad (1.11)$$

Eşitlik 1.11 test karakteristik eğrisi olarak adlandırılmaktadır. Buna göre $SL_{diff}(\theta_i)$, verilen bir θ_i için test karakteristik eğrileri arasındaki farkın karesini belirtmektedir. Aksine, $H_{diff}(\theta_i)$ verilen bir θ_i için test karakteristik eğrileri arasındaki farkın karesinin toplamıdır. Aşağıdaki şekil 1.1.'de X ve Y formlarının hipotetik test karakteristik eğrileri gösterilmiştir.



Şekil 1.3. X ve Y Formlarının Hipotetik Test Karakteristik Eğrileri

Daha sonra SL_{diff} bireyler üzerine dağıtılmaktadır. Kestirim işlemi aşağıdaki kriteri asgariye çekecek A ve B'nin hesaplanması ile devam etmektedir;

$$SL_{crit} = \sum_i SL_{diff}(\theta_i) \quad (1.12)$$

Eşitlik 1.12'deki toplam, bireyler üzerinden elde edilmektedir. Eşitlik 1.9 ve 1.12'de A ve B'yi çözmeye yaklaşımı hesaplama açısından yoğun bir iteratif yaklaşımdır.

Stocking-Lord karakteristik eğrisi yöntemi kalibrasyonlardan elde edilen bilgiyi kullanmaktadır. Her bir maddenin kalibrasyonu madde karakteristik eğrisi veya madde cevap fonksiyonunu vermektedir. Eğer ki kestirimler hatasız olsaydı, lineer dönüşüm için uygun A ve B eşitleme katsayılarının seçilmesi iki karakteristik eğrisinin kesişmesine neden olurdu (Stocking ve Lord, 1983: 203).

Stocking ve Lord (1983), madde karakteristik eğrisi yönteminin madde ayırtedicilik değerlerini dönüştürmede daha iyi sonuçlar ürettiğini ortaya koymuştur. Kaskowitz ve De Ayala (2001) araştırmalarında, madde parametre kestirim hatasının Stocking-Lord test karakteristik eğrisi yönteminin sağlamlığına etkisini incelemiştir. Simülatif verilerde madde parametre kestirim hataları 2PLM için 0.11, 0.25, 0.63; 3PLM için 0.30, 0.91 ve 2.73 olarak hesaplanmıştır. Sonuç olarak, birbirine daha uzak test karakteristik eğrilerinde ortaya çıkan daha büyük hatalar eşitleme katsayıları doğruluğunu azaltıcı etki yapmıştır. Bununla birlikte madde parametre kestirim hataları varlığında S-L yönteminin sağlamlığı oldukça kabul edilebilir bulunmuştur. Kolen ve Brennan (1995) ise karakteristik eğri dönüşüm

yöntemlerinden hem Stocking ve Lord hem de Haebara yöntemini; lineer yöntemlerden ise hem ortalama-sigma hem de ortalama-ortalama yöntemlerini önermiştir.

H. Eşitlemenin Değerlendirme Kriterleri

Gerçekleştirilen eşitlemenin ne derece kabul edilebilir olduğunu değerlendirmek için literatürde kullanılan çok çeşitli kriterler bulunmaktadır. Bu kriterlerden bazıları; gerçek-simülatif veri, zayıf eşitlik, büyük örneklem, standart hatalar, indisler, tutarlılık kararlılık incelemesi olarak sıralanmaktadır. Hangi kriterin kullanışlı olduğu konusunda kesin bir görüş bulunmamaktadır. Eşitleme yapıldığı süreçte sonuçların değerlendirilmesinde eşitleme kriterine ihtiyaç vardır. Eşitleme sonuçları duruma özeldir ve her durum için ayrı, özel eşitleme kriteri geliştirilmelidir (Harris ve Crouse, 1993). Kolen ve Brennan (2004) eşitleme koşullarından eşitlik, gruplarda varyans değişmezliği, simetriklik özelliklerinin değerlendirme kriterini geliştirmede kullanılabileceğini belirtmiş; bununla birlikte eşitlemenin standart hatasını kullanarak tesadüfi hata kestirimini değerlendirme kriteri olarak ele almıştır.

Test eşitlemede iki çeşit hata vardır: Sistemik hata (Bias) ve Tesadüfi hata (Random Error). Tesadüfi hata bireylerin puanları evren yerine örneklemden elde edildiği zaman ortaya çıkmaktadır. Tüm evrende uygulama yapıldığında tesadüfi hata ortaya çıkmayacaktır. Dolayısıyla, özellikle büyük örneklemelerde kestirilen eşitleme ilişkisindeki tesadüfi hata büyüklüğü önemsiz bir hal almaktadır. Küçük örneklemelerde ise tesadüfi hata artmaktadır. Tesadüfi hata, eşitlenmiş puanların standart sapmalarını ele alarak hesaplanan standart hatadır (Kolen ve Brennan, 2004).

Diğer bir hata çeşidi olan sistemik hata farklı durumlarda ortaya çıkabilmektedir. Örneğin, eşitleme ilişkilerinin kestiriminde kullanılan yöntem yanlış içeriyorsa sistemik hata ortaya çıkmaktadır. Kullanılan düzgünleştirme teknikleri sistemik hata ortaya koymakta; örneğin faydalı bir düzgünleştirme tekniği tesadüfi hatayı azaltırken sistemik hatayı artırabilmektedir. Yine eşitleme varsayımları sağlanmadığında ve çok boyutlu testler MTK gerçek puan ile eşitlendiğinde de sistemik hata ortaya çıkmaktadır. Bunlardan başka, eşitleme için seçilen veri toplama deseni düzgün uygulanmazsa (örneğin, eşdeğer gruplar deseninde X ve Y formu rastgele oturmuş öğrencilere verilmezse) ve eşitleme yürütülen gruptaki

bireyler eşitlenmiş formu alan bireylerden büyük ölçüde farklı ise yine sistematik hata ortaya çıkacaktır. Bununla birlikte tesadüfi hatanın aksine örneklem büyüklüğünün artması sistematik hatayı azaltmamaktadır. İyi bir model küçük sistematik ve tesadüfi hataya sahip olmalıdır. Hatanın büyük olduğu durumlarda eşitleme yapmak doğru olmayabilir (Kolen ve Brennan, 2004).

Test eşitlemeye karışan hatayı belirleyebilmek için eşitleme hatası ölçütüne ihtiyaç duyulmaktadır. Eşitleme hatası ölçütleri literatürde çok çeşitli indekslerle hesaplanabilmektedir. Bunlar; fark kareleri ortalamasının karekökü (Root Mean Square Differences, RMSD) veya hata kareleri ortalamasının karekökü (Root Mean Squared Error, RMSE), standartlaştırılmış hata kareleri ortalamasının karekökü (Standardized Root Mean Squared Error, SRMSE), ortalama kare hata (Mean Squared Error, MSE) veya toplam hata, yanlılık (Bias), ortalama mutlak fark (Mean Absolute Difference, MAD) vb. kolay yorumlanmaları bakımından avantajlı eşitleme kriterleridir (Harris ve Crouse, 1993; MacCann, 1990; Bogan ve Yen, 1983).

I. Madde Tepki Kuramı Varsayımları

Madde tepki kuramı, klasik test kuramına kıyasla son yıllarda giderek artan şekilde test geliştirme ve değerlendirme çalışmalarında kullanılmaktadır. Klasik test kuramında elde edilen madde istatistikleri gruba bağımlıdır. Örneğin, evrendeki grup ortalaması üzerinde yetenek düzeyi olan bireylerin oluşturduğu örnekleme, madde güçlüğü daha yüksek çıkmaktadır. Ayrıca bir birey aynı yeteneği ölçen farklı bir test aldığı anda yeteneği farklı çıkabilir. Çünkü elde edilen tüm parametreler gruba bağlıdır. Madde tepki kuramı ise daha spesifik, örtük modellere dayanan bir kuramdır. Burada bireylerin yeteneği gruptan bağımsız şekilde matematiksel model dayanaklı kestirilir. Ayrıca test maddelerinin, aynı yetenek düzeyinde olan; fakat farklı gruplardan gelen bireylerin cevapları açısından farklı özellikler göstermemesi beklenmektedir. Dolayısıyla MTK'da yetenek kestirimleri ve madde parametreleri gruptan bağımsızdır ve değişmez oldukları kabul edilir.

Bir bireyin bir maddeyi cevaplama davranışı ile o maddeye cevap vermesini sağlayan özellik arasında bir ilişki vardır ve bu ilişki madde karakteristik eğrisi (MKE) ile gösterilebilir. Bu eğriye göre yüksek yetenek düzeyindeki cevaplayıcıların maddeye doğru cevap verme olasılıkları daha yüksektir (Hambleton ve Swaminathan, 1985). MKE ile gösterilebilen bireyin maddeyi doğru

cevaplandırma olasılığı ile yetenek düzeyi arasında ilişki matematiksel modele dayanır. Burada yetenek ölçüleri testlerden bağımsız olarak elde edilebildiği gibi madde parametreleri de gruptan bağımsız olarak hesaplanabilmektedir. MTK'nın temel amacı da değişmez madde parametreleri ve yetenek kestirimleri elde etmektir. Bu kadar güçlü özelliklere sahip kuram KTK'dan farklı olarak bazı sınırlılıklara sahiptir. MTK'nın matematiksel modellere dayanması, güçlü ve sağlanması zor varsayımları beraberinde getirmiştir (Hambleton ve Swaminathan,1985).

Test puanlarının yorumlanmasında ve test sonuçlarının raporlamada seçilen model ile test verisi arasında yakın bir eşleşme (uyum) yakalanmalıdır. Uyum iyiliği soruşturmasında aşağıdaki şu üç yola başvurulmaktadır (Hambleton ve Swaminathan,1985:151):

- 1) Test verisinin seçilen test modelinin varsayımlarını karşılayıp karşılamadığının belirlenmesi.
- 2) MTK kullanımından beklenen sonuçların (değişmez madde ve yetenek kestirimleri) elde edilip edilmediğinin ortaya konulması.
- 3) Test verisi ve model parametrelerinden yaralanarak beklenen ve gözlenen sonuçların (örneğin, test puan dağılımları) arasındaki uyumun yüksekliğinin araştırılması.

Model seçiminde varsayımların testi, uyum iyiliği soruşturmasında kullanılan yöntemlerin ilki olarak ele alınmaktadır. Model varsayımların kontrolü aşamasında tek boyutluluk, yerel bağımsızlık, maddelerin ayırt edicilik indislerinin eşitliği (sadece 1PLM için), düşük şans başarısı ile cevaplama (1PLM ve 2PLM için), hız testi olmama varsayımları test edilmektedir.

Yerel bağımsızlık varsayımı tek boyutluluk varsayımı ile eşlenik olarak değerlendirilir. Yerel bağımsızlık varsayımına göre, bireylerin farklı maddelere vermiş oldukları cevaplar istatistiksel olarak birbirinden bağımsızdır. Bu varsayımın gerçekleşebilmesi için maddelerin cevaplanması birbirine bağlı olmamalı; cevaplar birbirinden etkilenmemelidir. Örneğin, maddenin içeriği diğer test maddelerinin cevaplanması için ipucu içermemelidir. Maddelerin sıralanışı test performansını etkilememelidir. Ayrıca, en önemlisi, test tek boyutlu olmalıdır. Eğer yerel bağımsızlık yoksa, bazı bireyler kendileriyle aynı yetenek düzeyindeki bireylerden

daha yüksek test puanları alacaktır. Dolayısıyla bireyin test performansının ölçülebilmesi için birden fazla yeteneğin hesaba katılması gerekecektir. (Hambleton ve Swaminathan,1985: 22-24).

MTK modeli test verisine uyumlu olduğunda; bireylerin yetenek kestirimleri aynı yetenek ölçeğinden elde edilir ve bireyler belli yeteneği ölçen madde havuzundan farklı maddeler cevaplasalar bile yetenekleri karşılaştırılabilir. Madde istatistikleri test maddelerinin kalibrasyonunda kullanılan birey örnekleminde bağımsız elde edilir ve yetenek kestirimlerinin doğruluğunun göstergesi yetenek ölçeğinin her noktasında elde edilir (Hambleton ve Swaminathan,1985:161-162)

Uyum iyiliği soruşturmasının üçüncü aşamasında, farklı modeller kullanılarak yapılan kestirimlerin doğruluk dereceleri kontrol edilmektedir. Bu kontrollerde farklı yollarla kestirilen MKE'ler karşılaştırılabilmektedir. Madde ve birey düzeylerinde model-test veri uyumlarının artık ve standartlaştırılmış artıkları incelenebilmektedir. Uyum bilgisi pek çok istatistikle özetlenmektedir: Verinin grafik gösterimlerinin kullanılması, yetenek kestirimleri ve test puanlarının grafiğinin incelenmesi, gerçek ve kestirilen madde ve yetenek parametrelerinin grafiğinin incelenmesi gibi. Yine beklenen ve gözlenen puan dağılımları karşılaştırılabilmekte, ki-kare istatistiği gibi pek çok istatistik ve grafik yöntemleri sonuçları yorumlamada kullanılabilmektedir. Madde sıralaması, madde biçim etkisi, sıkılma, kopya, testin hızlandırılması gibi testin uygulama etkileri kapsamında hipotezler incelenebilmektedir (Hambleton ve Swaminathan,1985:157-158).

J. Bilimsel Süreç Becerileri (BSB)

Teknolojik gelişmelere paralel olarak bilimsel bilginin her geçen gün yenilediği çağımızda, öğrencilerin bu değişme ve gelişmelere ayak uydurabilmeleri için öğretim programlarında düzenlemelere gerek duyulmaktadır. Değişen bilimsel bilginin sürekli gözden geçirilerek geliştirilmesi, teknolojik gelişmelerin takip edilmesi ve bilimsel araştırmalarda kullanılması için fen ve teknoloji öğretimi büyük öneme sahiptir. Öğrencilerin dünyayı keşfetmesi, burada gerçekleşen olayları bilimsel süreçlerle araştırıp açıklayabilmesi ve insanlar yararına bilimsel ürünler ortaya koyabilmesi için bir takım zihinsel ve psikomotor becerilere sahip olması gerekmektedir. Bu beceriler, genel olarak, araştırma-soruşturma, eleştirel düşünme, problem çözme, karar verme becerileri olarak sıralanabilir.

Fen ve teknoloji dersi öğretim programının vizyonu; bireysel farklılıkları ne olursa olsun bütün öğrencilerin fen ve teknoloji okuryazarı olarak yetişmesidir. Fen ve teknoloji okuryazarlığı, bireylerin çevreleri ve dünya hakkındaki merak duygusunu sürdürmeleri için gerekli olan fenle ilgili beceri, tutum, değer, anlayış ve bilgilerin bir bileşimidir. Fen ve teknoloji okuryazarı olan bir kişi, bilimin ve bilimsel bilginin doğasını, temel fen kavram, ilke, yasa ve kuramlarını; problemleri çözerken ve karar verirken bilimsel süreç becerilerini anlayarak uygun şekillerde kullanır; fen, teknoloji, toplum ve çevre arasındaki etkileşimleri anlar; bilimsel ve teknik psikomotor beceriler geliştirir; bilimsel tutum ve değerlere sahip olduğunu gösterir (MEB, 2006). Buradan anlaşıldığı gibi bilimsel süreç becerileri fen okuryazarlığı için kazanılması gereken becerilerdir.

Bilimsel süreçler aslında düşünmenin temelini oluşturmaktadır ve sadece fende değil diğer bilim alanlarında da problem çözmede kullanılmaktadır. Bilimsel süreçler, bilgi toplama, toplanan verileri çeşitli yöntemler kullanarak düzenleme, sıra dışı durumları açıklama ve problem çözmede kullanılan zihinsel ve bedensel becerilerdir (Carin ve Bass, 2001). Ülkemizde, Talim ve Terbiye Kurulu Başkanlığı'nca (2006) geliştirilen ve araştırmada da kullanılan İlköğretim Fen ve Teknoloji Dersi Öğretim Programı'nda yer alan bilimsel süreç becerileri Tablo 1.8'deki gibi alt boyutlara ve kazanımlara ayrılmıştır:

Tablo 1.8: İlköğretim 6., 7. ve 8. Sınıf Düzeyi İçin “Bilimsel Süreç Becerisi” Kazanımları

	Beceriler	Beceriye Yönelik Kazanım
PLANLAMA VE BAŞLAMA	Gözlem	1. Nesneleri (cisim, varlık) ve olayları duyu organlarını veya gözlem araç gereçlerini kullanarak gözlemler. 2. Bir cismin şekil, renk, büyüklük ve yüzey özellikleri gibi duyuşsal özelliklerini belirler. 3. Gözlem için uygun ve gerekli araç gereci seçip bunları beceriyle kullanır.
	Karşılaştırma-Sınıflama	4. Nesneleri sınıflandırmada kullanılacak nitel ve nicel özellikleri belirler. 5. Nesnelere veya olaylar arasındaki belirgin benzerlikleri ve farklılıkları saptar. 6. Gözlemlere dayanarak bir veya birden fazla özelliğe göre karşılaştırmalar yapar. 7. Benzerlik ve farklılıklara göre grup ve alt-gruplara ayırma şeklinde sınıflamalar yapar.
	Çıkarım Yapma	8. Olmuş olayların sebepleri hakkında gözlemlere dayanarak açıklamalar yapar.
	Tahmin	9. Gözlem, çıkarım veya deneylere dayanarak geleceğe yönelik olası sonuçlar hakkında fikir öne sürer.
	Kestirme	10. Olay ve nesnelere yönelik kütle, uzunluk, zaman, sıcaklık ve adet gibi nicelikler için uygun birimleri de belirterek yaklaşık değerler hakkında fikirler öne sürer.
	Değişkenleri Belirleme	11. Verilen bir olay veya ilişkide en belirgin bir veya birkaç değişkeni belirler. 12. Verilen bir olaydaki bağımlı değişkeni belirler. 13. Verilen bir olaydaki bağımsız değişkeni belirler. 14. Verilen bir olaydaki kontrol edilen değişkenleri belirler.
	Hipotez Kurma	15. Verilen bir olaydaki bağımsız değişkenin bağımlı değişken üzerindeki etkisini denenebilir bir önerme şeklinde ifade eder.
	Deney Tasarlama	16. Kurduğu hipotezi sınamaya yönelik bir deney önerir.
	Deney Malzemelerini Tanıma Ve Kullanma	18. Verilen malzemeleri kullanarak kurduğu hipotezi sınamaya yönelik tasarladığı deneyi gerçekleştireceği bir düzenek kurar.
	Deney Düzenegi Kurma	17. Basit araştırmalarda gerekli malzeme, araç ve gereçleri seçerek emniyetli ve etkin bir şekilde kullanır.
UYGULAMA	Değişkenleri Kontrol Etme ve Değiştirme	19. Hipotezle ilgili olan değişkenlerin dışındaki değişkenleri sabit tutar. 20. Bağımsız değişkeni değiştirerek bağımlı değişken üzerindeki etkisini belirler.
	İşlevsel Tanımlama	21. Değişkenlerin birden fazla anlama gelebileceği, sınırları tam çizilmemiş durumlarda araştırmacının amacına (hipotez) uygun değişkenleri kesin olarak ve ölçme kriteri ile birlikte tanımlar.
	Ölçme	22. Cetvel, termometre, tartı aleti ve zaman ölçer gibi ölçme araçlarını tanıır. 23. Büyüklükleri uygun ölçme araçları kullanarak belirler. 24. Büyüklükleri birimleri ile ifade eder.
	Bilgi Ve Veri Toplama	25. Değişik kaynaklardan yararlanarak bilgi (çevrede, sınıfta gözlem ve deney yaparak, fotoğraf, kitap, harita veya bilgi ve iletişim teknolojilerini kullanarak) toplar. 26. Kurduğu hipotezi sınamaya yönelik nitel veya nicel veriler toplar.
	Verileri Kaydetme	27. Gözlem ve ölçüm sonucunda elde edilen araştırmacının amacına uygun verileri yazılı ifade, resim, tablo ve çizim gibi çeşitli yöntemlerle kaydeder.
ANALİZ VE SONUÇ ÇIKARMA	Veri İşleme Ve Model Oluşturma	28. Deney ve gözlemlerden elde edilen verileri derleyip işleyerek gözlem sıklığı dağılımı, çubuk grafik, tablo gibi farklı formlarda gösterir. 29. Grafik çizmeyle ilgili kuralları uygular.
	Yorumlama Ve Sonuç Çıkarma	30. İslenen verileri ve oluşturulan modeli yorumlar. 31. Elde edilen bulgulardan desen ve ilişkilere ulaşır.

Tablo 1.8’de verilenlere göre gözlem, sınıflama-karşılaştırma, çıkarım yapma, tahmin, kestirme, değişkenleri belirleme becerileri bilimsel süreç becerilerinin “planlama ve başlama” basamağını; hipotez kurma, deney tasarlama, deney malzemeleri ve araç-gereçleri tanıma ve kullanma, deney düzeneğı kurma, değişkenleri kontrol etme ve değiştirme, işlevsel tanımlama, ölçme, bilgi ve veri toplama, verileri kaydetme becerileri “uygulama” basamağını; veri işleme ve model oluşturma, yorumlama ve sonuç çıkarma, sunma becerileri ise “analiz ve sonuç çıkarma” basamağını oluşturmaktadır.

2. İLGİLİ ARAŞTIRMALAR

Bu bölümde, ilgili alan yazında yapılmış araştırmaların özetlerine yer verilmiştir.

Baker ve Al Karni (1991) araştırmalarında, hem simülatif hem de gerçek veri kullanarak “Stocking-Lord” ve “ortalama-ortalama” yöntemlerinin madde parametre ve yetenek kestirimlerinde ne derece uyumlu sonuçlar verdiğini karşılaştırmıştır. Araştırmada 60 maddeli 15 ortak maddeli formlarda yatay ve dikey eşitleme gerçekleştirilmiş; eşitlemelerin değerlendirilme kriteri olarak RMS eşitleme hataları hesaplanmıştır. Sonuç olarak, gerçek verinin kullanıldığı durumda iki yöntemin benzer eşitleme sonuçlarına sahip olduğu gözlenmiştir. Simülatif veri kullanıldığında ise, Stocking-Lord yöntemi ortalama-ortalama yöntemine göre daha az hata üretmiştir. Bütüne bakıldığında, genel olarak pek çok amaç için ortalama-ortalama yönteminin daha kabul edilebilir eşitleme katsayıları verdiği belirtilmiştir. Stocking-Lord yönteminin ise, tipik test karakteristiklerine karşı daha az duyarlı olduğu sonucu ortaya konulmuştur. Test kalibrasyonlarının, kalibre etmesi tipik olarak sorunlu veri setleriyle ilişkili olabileceğinden şüphelenen araştırmacılara Stocking-Lord yöntemini kullanmaları önerilmiştir.

Cohen ve Kim (1998), simülatif ve çok kategorili maddelere ait veri üzerinde MTK'ya dayalı aşamalı tepki modeli (graded response model) ile çalışmıştır. Araştırmada “ortalama-ortalama”, “ortalama-sigma”, “sağlam ortalama-sigma” ve “Stocking-Lord” eşitleme yöntemleri kullanılarak ortak madde sayısı ve yetenek dağılımı değişkenlerinin eşitleme hatasına etkileri karşılaştırılmıştır. Araştırma sonuçlarında eşitleme yöntemlerinin genel olarak benzer sonuçlar ürettiği belirtilmiştir. Bununla beraber, ortak madde sayısı arttığında eşitleme hatalarının azaldığı ve ortak madde sayısındaki değişimin en az sağlam ortalama-sigma ve Stocking-Lord yöntemlerini etkilediği ve bu yöntemlerin aykırı madde parametrelerinden en az etkilendiği bulunmuştur. Ayrıca, çarpık yetenek dağılımlarında daha az eşitleme hatası elde edilmiştir. Araştırmada yer alan değişkenler için, tüm eşitleme yöntemleri içerisinde Stocking-Lord yönteminin kullanılması önerilmiştir.

Tate (2000), yapılandırılmış cevaplı ve çoktan seçmeli maddelerin birlikte bulunduğu simülasyon çalışmasında “Stocking-Lord” ve “ortalama-sigma” yöntemleri kullanarak eşitlemede kestirim yanlılığı (estimation bias) ve tesadüfi

hata deęerlerini incelemiřtir. Yapılandırılmıř cevaplı maddelerin oranı, ortak madde t¼r¼ (yapılandırılmıř/ oktan semeli), ortak madde sayısı, yeteneklerin yıldan yıla deęiřim ¼l¼s¼, řans fakt¼r¼n¼n varlıęı deęiřkenlerinin hata ¼zerindeki etkilerine bakılmıřtır. Arařtırmada iki eřitleme y¼nteminin ok benzer sonular ¼rettięi, yapılandırılmıř cevaplı ortak maddelerin sayısı arttıęında hatanın arttıęı, ortak maddelerin hepsi oktan semeli kullanıldıęında tek boyutluluk varsayımı altında geerli sonular elde edildięi; ancak tek boyutluluk bozulduęunda yanlılıęın arttıęı, tek boyutluluk varsayımının ihlalinde uygun oranda karma madde t¼r¼ kullanılmasının eřitlemeyi saęlamlařtırdıęı, karma ortak maddelerin sayısı yarıya indirildięinde A eřitleme katsayısı standart hatasının iki kat, B eřitleme katsayısının ¼ kat arttıęı, yetenek daęılımı ortalamasının yıldan yıla artması durumunda A ve B katsayıları hatalarında anlamlı bir deęiřim g¼zlenmedięi, řans fakt¼r¼ hesaba katıldıęında eřitleme hatasının arttıęı sonularına ulařılmıřtır. Ayrıca ok kategorili maddeler iin uyarlanmıř S-L ve OS y¼ntemlerinin birbirine ok benzer sonular verdięi belirtilmiřtir.

Hanson ve B¼guin (2002), sim¼lasyon alıřmasında ortak madde deseniyle MTK'ya dayalı ayrı ve eřit zamanlı madde parametre kestirimlerinin performansını d¼rt farklı durum iin karřılařtırmıřlardır. Bunlar: 1) MULTILOG ve BILOG-MG programları kullanıldıęında, 2) her bir form iin ¼rneklem b¼y¼kl¼ę¼ 1000 ve 3000 olduęunda, 3) ortak madde sayısı 10 ve 20 olduęunda, 4) iki formu eřitdeęer ve eřitdeęer olmayan gruplar aldıęında (ortalama farkı olmadıęında ve bir standart sapma ortalama farkı olduęunda) durumları olarak sıralanmıřtır. Ayrıca madde parametre kestirimleri ayrı yapıldıęında "ortalama-ortalama", "ortalama-sigma", "Stocking-Lord" ile "Haebara" y¼ntemleri kullanılmıřtır. İki kategorili puanlanan 60 maddeli veri seti 3PLM ile uyumlu kabul edilmiřtir. Arařtırma sonucunda b¼y¼k ¼rneklemde, daha ok ortak madde olduęunda ve eřitdeęer gruplarda hata daha az bulunmuřtur. BILOG-MG ve MULTILOG programları eřitdeęer olmayan gruplarda eřit zaman eřitleme hari benzer sonular ¼retmiřtir. MKE y¼ntemleri (Stocking-Lord ve Haebara), ortalama-ortalama ve ortalama-sigma y¼ntemlerine g¼re b¼y¼k ¼l¼de daha az hatalı sonular vermiřtir. Eřitdeęer olmayan gruplarda eřit zamanlı kalibrasyon ayrı kalibrasyona g¼re BILOG'da daha az hata ¼retirken MULTILOG'da tam tersi olmuřtur.

Chu (2002), “DMF Varlığında Eşdeğer Gruplarda Test Eşitleme” adlı araştırmasında hiyerarşik Rasch (bir parametrelili hiyerarşik genelleştirilmiş lineer lojistik) modeli ile parametre kestirimi ve test eşitlemede DMF bilgisini kullanmıştır. Araştırmanın amacı, simülasyonlar ile DMF varlığında hiyerarşik Rasch modelinin test eşitleme ve parametre kestirimi performanslarını incelenmektir. Hiyerarşik Rasch modelinde DMF faktörleri parametre olarak kabul edilirse, DMF'den kaynaklanan yanlılığın ortadan kaldırılabilceği ve bu sayede test eşitleme ve parametre kestiriminde performansın geliştirilebileceği belirtilmiştir. Cinsiyete göre iki alt grubun yeteneklerinin eşit ve normal dağıldığı ve sadece X formunun DMF'li madde içerdiği varsayılmıştır. DMF'li maddeler varken ve yokken tek grup eşzamanlı eşitleme ile hiyerarşik Rasch modeli performansları kıyaslanmıştır. Hiyerarşik Rasch modeli performansı, DMF'li maddelerin sayısı, DMF'li maddelerin ortak madde olup olmaması ve DMF büyüklüğü olmak üzere üç değişken açısından incelenmiştir. Araştırma sonuçlarına göre, tüm simülasyon çalışmalarında madde parametre ve yetenek kestirimleri için üç model arasında en iyi performansı hiyerarşik Rasch modeli göstermiş, bunu sırasıyla, DMF maddelerini içeren tek grup eşzamanlı eşitleme ve DMF'li maddeler silindiğinde tek grup eşzamanlı eşitleme modelleri takip etmiştir. Sonuçta test geliştirme ve parametre kestirimleri esnasında DMF bilgisinden yararlanmanın, DMF'li maddeleri silmekten veya onları görmezden gelmekten daha iyi sonuçlar üretebileceği ve kısa bir testten DMF'li maddenin silinmesinin geçerliliği daha fazla zedelediği belirtilmiştir.

Araştırmada incelenen üç değişken, hiyerarşik Rasch modeli ile DMF maddeli tek grup eşzamanlı eşitleme modelleri üzerinde farklı etkilere sahip olmuştur. İki modelin madde parametre kestirimlerinin kararlılığı ile DMF'li madde sayısı ve DMF'nin büyüklüğü değişkenleri arasında negatif bağ bulunmuştur. Önem testleri hiyerarşik Rasch modeli DMF'li parametre kestiriminin DMF büyüklüğü hariç diğer pek çok DMF etkisine karşı güçlü olduğunu göstermiştir. DMF ortak maddelerde yer aldığıında, ortak olmayan maddelerde veya her ikisinde yer almasına nazaran daha kararlı madde parametre kestirimleri elde edilmiştir. Yetenek kestirimlerinde ise (DMF büyüklüğü değişkeni hariç) farklı sonuçlar elde edilmiştir. Sadece ortak maddelerde DMF olması durumunda daha az kararlı kestirimler elde edilirken; 2

ve 5-DMF'li madde sayısına sahip desen karşılaştırmaları yetenek kestirimi açısından sonuçsuz/ yetersiz kalmıştır.

Zhang ve diğ. (2003), araştırmalarında yetenek testinden (Scholastic Aptitude Test 1:Reasoning Test) DMF'li maddelerin çıkarılmasının etkilerini incelemişlerdir. Araştırma eşitleme ile doğrudan ilişkili değildir; çünkü öncelikli amaç DMF'li maddeleri puanlamadan çıkarmaktır. DMF'li maddelerin çıkarılmasından sonra puanlar eşit yüzdelikli eşitleme yöntemiyle tekrar eşitlenmiştir. Araştırmanın sonuçlarına göre, maddelerin çıkarılması ve puanların tekrar eşitlenmesinden sonra ölçeklenmiş puan farklılıkları DMF etkisine bağlı olarak alt gruplar arasında değişim gözlenmiştir. DMF'li maddelerin çıkarılması odak grup lehine avantaj sağlamış ve referans grup puanları ise DMF'li maddelerin çıkarılmasından negatif etkilenmiştir.

Karkee ve Wright (2004) gerçek veri üzerinde yürüttükleri araştırmada 8 ile 12. sınıf arasında öğrenim gören 31.813 öğrenciye 78 maddeden oluşan çoktan seçmeli İngilizce testini uygulamış ve sonuçları ölçeklemiştir. Araştırmanın amacı, "iki moment (OO ve OS)", "iki karakteristik eğrisi (HB ve S-L)" ve "teta regresyon" yöntemlerini kullanarak eşitleme yapmak ve sonuçları karşılaştırmaktır. Eşitleme yapılırken elde edilen 3PLM'ye dayalı parametre kestirimleri, araştırmada belirtilen eşitleme katsayılarıyla, 1PLM'ye dayalı parametre kestirimlerine dönüştürülmüştür. Eşitlemenin değerlendirme kriterleri olarak; test karakteristik eğrileri, standart hata, frekans dağılımı eğrileri, test bilgi fonksiyonları, ölçek puanları ortalama ve standart sapmaları kullanılmıştır. Tercih edilmesi önerilen yöntem kullanılan eşitleme değerlendirme kriterine göre farklılık göstermiştir. Araştırma sonucunda, karakteristik eğri yöntemlerinin daha küçük eşitleme hataları (standart hata) ile moment yöntemlerine üstünlük sağladığı belirlenmiştir. S-L yöntemi, 1PLM ölçek puan dağılımı ile benzer ortalama ve standart sapmalar üretmiştir. Dolayısıyla ölçek puan dağılımı ve ortalama, st. sapma kullanımı önemliyse, S-L yönteminin kullanılması önerilmiştir. HB yöntemi ise, performans seviyeleri sınıflama yüzdeleri önemli görüldüğünde 1PLM ile S-L yönteminden daha benzer değerler üretmiştir. Bunlar dışında iki karakteristik eğrisi yönteminin farklılıkları önemsiz görülmüştür. S-L ve HB yöntemi ölçek puan dağılımı orta noktasının az yukarısında en küçük hataya sahipken, O-S yöntemi dağılımın ortasında en küçük hataya sahip bulunmuştur. Moment ve Teta regresyon yöntemleri dağılımın ortasında daha

büyük hatalar üretmiştir. Puan dağılımının iki ucunda ise O-O ve Teta regresyon yöntemleri daha küçük hataya sahip bulunmuştur. Ayrıca tüm değerlendirme kriterleri için karakteristik eğrisi yöntemleri 3PL ölçeğin 1PL ölçeğe dönüştürülmesinde en iyi sonuçları vermiştir. Bununla birlikte araştırmada ham puan dağılımının negatif çarpık olduğuna vurgu yapılarak, daha sonraki araştırmacılara normal ve diğer ham puan dağılımlarında ve daha küçük örneklemelerde ölçekleme yapmaları önerilmiştir.

Kim ve Lee (2004); Kim ve Lee (2006) araştırmalarında karma testlerde (çoktan seçmeli ve yapılandırılmış maddeli) simülatif veri kullanarak MTK'ya dayalı "iki moment yöntemi (OO ve OS)" ve "iki karakteristik eğrisi eşitleme yöntemi (HB ve SL)" hatalarını farklı koşullarda karşılaştırmıştır. Bu koşullar beş farklı eşitleme modeli (3 parametrelili lojistik model (3PLM), aşamalı tepki modeli (ATM), genelleştirilmiş kısmi puan modeli (GKPM), sınıflamalı ölçeği modeli (SÖM) ve çoktan seçmeli model (ÇSM); iki farklı örneklem büyüklüğü (500/ 3000), iki farklı yetenek dağılımı (eşdeğer / eşdeğer olmayan), üç farklı karma test formatı (iki /çok kategorili madde sayısı: 10/10, 20/5, 30/2), iki bilgisayar programı (Multilog/ Parscale), üç farklı bağlama şekli (ortak maddeler sadece iki kategorili, sadece çok kategorili ve karışık) olarak sıralanmıştır. Eşitlemelerin değerlendirme kriteri olarak kare yanlılık ve varyans hesabına dayanan ortalama kare hata (Mean Squared Error=MSE) tüm koşullar için hesaplanmıştır. Araştırma sonucunda en iyi uyuma sahip modellerin 3PLM ve ÇSM olduğu; örneklem büyüdükçe eşitleme hatasının azaldığı; eşitleme yöntemlerinden karakteristik eğrisi yöntemlerinin en doğru sonuçlara sahip olduğu bulunmuştur. Karakteristik eğri yöntemleri arasında ise Haebara'nın daha az eşitleme hatası (MSE) verdiği gözlenmiştir. Hanson ve Bèguin'in (2002) ulaştığı sonuç gibi S-L ve HB yöntemlerinde daha düşük hatanın (MSE), daha düşük varyans sebebiyle elde edildiği belirtilmiştir. Tek bir formatta ortak madde olduğunda veya sayıları arttığında eşitleme hatası MSE tüm eşitleme yöntemlerinde azalmıştır. Eşdeğer gruplarda eşdeğer olmayan gruplara göre daha küçük eşitleme hataları elde edilmiştir. Ayrıca kullanılan bilgisayar programları benzer sonuçlar vermiş; bazı durumlarda karakteristik eğri yöntemleri için hatalarda farklılık ortaya konulmuştur.

Chu ve Kamata (2005), "DMF Varlığında Test Eşitleme" adlı araştırmalarında simülatif veri ile eşdeğer gruplarda test eşitleme ve akabinde yetenek kestirimi

esnasında DMF gösteren bazı ortak maddelerden faydalanan Hiyerarşik Rasch modelini (1PL çok düzeyli madde tepki modeli) incelemiştirler. DMF faktörü bir parametre olarak ele alındığında, Hiyerarşik Rasch modelinin DMF'den kaynaklanan yanlılığı ortadan kaldırılabileceği ve yetenek kestirimi ile test eşitlemenin performansını artırabileceği belirtilmiştir. Araştırmada Hiyerarşik Rasch modeli ile yetenek kestirimlerinin ve madde güçlüklerinin doğruluk ve kararlılığının incelenmesi amaçlanmıştır. DMF'li maddelerin sayısı (2, 5 ve 10 olduğunda) ve DMF 'in büyüklüğü (MH-Delta DMF; 0.6, 0.8, 1.2 ve 2.0 olduğunda) olmak üzere iki faktör incelenmiştir. İki çoklu-grup eşzaman eşitleme (multiple-group concurrent equating) ile önerilen metodun performansı ile karşılaştırma yapılmıştır. Birinci eşzaman eşitleme temel model olarak kullanılmış ve DMF'li maddeler DMF'li olmayanlarla yer değiştirmiştir. Hiyerarşik Rasch modeli gibi aynı DMF maddelerine sahip ikinci eşzaman eşitlemede DMF görmezden gelinerek eşitleme yapılmıştır. Sonuç olarak, Hiyerarşik Rasch modeli performansının madde parametreleri ve yetenek kestiriminde hem DMF'li hem de DMF'siz maddeler varlığında çoklu-grup eşzaman eşitlemeden daha iyi olduğu ve yanlılık barındıran maddelerde kullanılmasının uygun olduğu ortaya konulmuştur.

Kim ve Kolen (2006) karma testlerde (çoktan seçmeli ve yapılandırılmış maddeli) simülatif veri kullanarak MTK'ya dayalı eşitleme yöntemlerinden "iki moment yöntemi (OO ve OS)" ve "iki karakteristik eğrisi yöntemini (HB ve S-L)" eşzaman kalibrasyon ile karşılaştırmıştır. Araştırmada standart başarı testlerini temsilen "geniş ranjlı" iki karma test ve profesyonel sertifika testlerini (screening tests) temsilen "dar ranjlı" iki karma test eşdeğer olmayan gruplarda ortak madde deseni ile eşitlenerek beş yöntemde birbiriyle kıyaslanmıştır. Burada farklı formattaki maddelerle yapılan eşitlemelerin sağlamlık dereceleri üzerine odaklanılmıştır. Bunun için format etkisi üç düzeyde incelenmiş; iki baskın yapı θ_1 ve θ_2 'nin aralarında 0.5, 0.8 ve 1.0 korelasyon gösterdiği üç durum oluşturulmuştur (θ_1 : sadece çoktan seçmeli maddelere verilen cevaplardan etkilenmiş, θ_2 : sadece yapılandırılmış maddelere verilen cevaplardan etkilenmiş varsayılmıştır). Korelasyonun 1.0 olması format etkisi yoktur, anlamına gelmiştir. Ayrıca eşitleme yapılan araştırmada ortak madde setinin kapsam, istatistiksel karakterler ve oransal olarak toplam test formlarını temsili sağlanmıştır. Testlerdeki ortak maddeler, madde havuzundan, 12'si çoktan seçmeli ve 5'i yapılandırılmış formatta

(tüm testin %33.3'ü) seçilmiştir. ICL (IRT Command Language) bilgisayar programı ile iki kategorili puanlanan maddeler için 3PLM, çok kategorili puanlanan maddeler için genelleştirilmiş kısmi puan modeli ile madde parametreleri hesaplanmış ve STUIRT programı ile eşitleme yapılmıştır. Eşitlemenin değerlendirilme kriteri olarak gözlenen puan dağılımından (Observed Score Distribution, OSD) MSE elde edilmiştir. Araştırma sonucunda, her koşulda, en doğru sonuçları beş yöntem içinde eşzaman kalibrasyon yönteminin verdiği; en büyük hataların ise moment yöntemlerinde olduğu gözlenmiştir. Format etkisi olmadığında geniş ranjlı testte en büyük hata sırasıyla; O-O, O-S, S-L, HB ve en düşük hata eşzaman eşitlemede elde edilmiştir. Format etkisi olduğunda ise tek fark olarak HB, S-L yönteminden daha büyük hata üretmiştir. Dar ranjlı testte format etkisi olmadığında en büyük hata sırasıyla; O-S, O-O, HB, S-L ve en düşük hata eşzaman eşitlemede elde edilmiştir. Format etkisi varken sıralama O-O, O-S, S-L, HB olarak değişmiştir. Sonuç olarak çok ileri düzey format etkisi olduğunda bile karakteristik eğrisi yöntemleri ve eşzaman eşitleme en sağlam yöntemler olarak çalışmıştır. O-O yönteminin en fazla hatayı verme nedeni olarak; güçlük parametreleri sayısı ve az sayıda (her madde için bir tane) ayırteçlilik parametresi eğitim kestiriminde kullanılması gösterilmiştir.

Turhan (2006), "İki Parametrelili Lojistik Çok Aşamalı (Hiyerarşik) Madde Tepki Modeli İle DMF Varlığında Dikey Eşitleme" adlı araştırmasında madde güçlüğü ve ayırıcılığı parametrelerini hesaplamış; sonrasında 500'er örneklem büyüklüğünde gerçek ve simülasyon veriden yararlanarak DMF'li maddelerle farklı sınıf düzeyleri arasında test eşitleme çalışması yapmıştır. Bağımsız değişkenler DMF'li madde sayısı (2 ve 4) ve DMF türüdür (TB/ TBO-DMF). DMF varlığında iki parametrelili lojistik çok aşamalı madde tepki modelinin performansı, sınıf düzeyindeki artış ele alınarak, geleneksel MTK ile karşılaştırılmıştır. Sınıf düzeyi arttıkça, iki parametrelili lojistik çok aşamalı madde tepki modeli geleneksel MTK'ya göre daha iyi bir dikey ölçek üretmiştir. Bazı tek biçimli olmayan DMF'li ortak maddeler örneklemesi, sınıf seviyesi değişkeni ile birlikte, iki modelde aynı ölçek sonucunu doğurmuş; ancak bazı tek biçimli ortak DMF'li ortak maddeler dikey ölçeği bozmuştur. TB, TBO ve hem TB hem de TBO-DMF'li maddelerin dikey ölçeklemeye etkisini inceleyen küçük simülasyon çalışmasında, aynı anda hem TB hem TBO-DMF'li maddeler varlığında, tek bir çeşit DMF'li maddelerin olması durumuna göre ölçek puanları

daha çok bozulmuştur. Ortak madde setindeki TBO-DMF'li maddelerin sayısı arttıkça ölçekteki bozulmanın arttığı gözlenmiştir. Araştırmada çelişkili bir sonuç bulunmuştur: TB-DMF'li ortak madde sayısı arttıkça ölçekteki bozulma azalmıştır. Bu durumun sınırlı simülasyon boyutu nedeniyle oluşmuş tesadüfi hatadan kaynaklanabileceği belirtilmiştir.

Han (2008) "Madde Parametre Kaymasının (Item Parameter Drift- IPD) Test Eşitleme ve Yetenek Kestirimleri Üzerine Etkisi" adlı araştırmasında IPD, DMF'nin bir çeşidi olarak ele alınmıştır. DMF belirleme tekniklerinin herhangi bir modifikasyona gerek duymadan IPD'yi aynı şekilde hesaplayabildiği vurgulanmıştır. Simülatif veri üzerinden yürütülen araştırmada eşitlenecek formlar 40'ar maddeden (10'u ortak) oluşturulmuştur. Tek yönlü ve çok yönlü IPD'li maddelerin ortak maddelerdeki oranı (10 ortak maddede %10(1), %20(2), %30(3), %40(4) ve %50(5) ile 5 farklı oran olduğunda) ve b parametresindeki IPD büyüklüğü (-1.00'den -0.05'e -0.05 artışla 20 farklı büyüklük) bağımsız değişkenleri ile $5 \times 20 = 100$ replikasyon çalışılmıştır. Araştırmada bağımsız değişkenlerin eşitleme katsayıları, madde parametre ve yetenek kestirimlerine olan etkilerine bakılmıştır. Eşitleme yöntemleri olarak tek yönlü IPD varlığında O-S yöntemi, çok yönlü IPD varlığında O-O, O-S ve S-L yöntemleri ve eşitlemenin değerlendirme kriteri olarak RMSE, RMSD ve BIAS kullanılmıştır. Ortak maddelerin tamamının b parametreleri form 2'de daha az olduğunda tek yönlü IPD, b parametreleri form 1 ve form 2'de farklı işaretler aldığıda çok yönlü IPD olduğu şeklinde tanımlanmıştır. Sonuç olarak tek yönlü IPD'li ortak maddeler O-S yönteminde b parametre kestirimlerini; dolayısıyla b parametre varyansları ile hesaplanan A katsayılarını büyük ölçüde etkilemiştir. Örneğin, IPD'nin büyüklüğü (1.00 olduğunda) ve IPD'li madde oranı arttığında (10 ortak maddenin 5'i= %50 IPD'li iken) A eşitleme katsayısı azalmış; çünkü IPD nedeniyle Form 2'de ortak madde b parametreleri varyansı artmıştır. B katsayısı ise IPD'li madde sayısı ve IPD büyüklüğü artınca, Form 2'de b parametre ortalamasının azalmasına bağlı olarak büyümüştür. Ortak maddelerde çok yönlü IPD'nin etkisi dört farklı koşulda incelenmiştir. Sonuçta S-L yönteminin genel olarak iyi performans gösterdiği, ortak madde sayısı arttığında O-S yönteminin S-L yöntemine yaklaşan eşitleme sonuçları verdiği belirtilmiştir. Ortak maddelerde IPD'nin bulunup bulunmaması bireylerin yetenek kestiriminde çok farklı sınıflamalar üretmiştir.

Hu ve diğ. (2008) simülasyon çalışmasında dört eşitleme yönteminden “eşzaman eşitleme”; ayrı kalibrasyonlardan “O-S”, “test karakteristik eğrisi” ve “ayarlanmış ortak madde parametre (Fixed Common Item Parameter- FCIP)” yöntemini karşılaştırmıştır. Araştırmanın bağımsız değişkeni olarak ortak maddelerde aykırı (outlier) değerlerin bulunup bulunmaması ele alınmıştır. Eşitlemenin değerlendirme kriteri olarak, eşitleme yanlılığı (bias) ve RMSE kullanılmıştır. Araştırma sonucunda, test karakteristik eğrisi ve O-S dönüşümleri en iyi eşitleme performansına sahipken, eşzaman ve FCIP kalibrasyonlarının grup eşitliği ve aykırı değerlerin sayısı ile karmaşık bir etkileşimi gözlenmiştir. Eşdeğer olmayan gruplarda ortak madde setinde aykırı değer olmadığında dört yöntemin eşit olmayan şekilde hassas oldukları belirtilmiştir. İki grubun yetenek ortalamaları aynı olduğunda dört yöntem eşdeğer hatalar üretirken; ortalamalar bir standart sapma değiştiğinde O-S ve karakteristik eğri yöntemi küçük hatalar, FCIP yöntemi orta düzey hatalar ve eşzaman kalibrasyon en büyük hataları üretmiştir.

Norman Dvorak (2009) tarafından yapılan çalışmada, Kernel eşitleme yöntemi ile S-L test karakteristik eğrisi yöntemi karşılaştırılmıştır. Araştırmada iki kategorili simülatif veri ile gerçek puan eşitleme yapılmıştır. Eşdeğer olmayan gruplarda ortak madde deseni ile yürütülen eşitlemelerde bağımsız değişkenler olarak; örneklem büyüklüğü (1000, 2000, 10000), test uzunluğu (25, 50, 75), ortalama faktör yükleri (0.50, 0.62) ve dış ortak madde yüzdesi (%10, %20, %30) ele alınmıştır. 2PLM ye dayalı BILOG MG3 ile yapılan kestirimler sonucu EQUATE 2.1 programı ile eşitlemeler gerçekleştirilmiştir. Eşitlemenin değerlendirme kriteri olarak RMSD eşitleme hataları hesaplanmıştır. Araştırma sonuçlarında, her iki eşitleme yönteminin de bağımsız değişkenlerin farklı düzeylerinde oldukça iyi bir performans göstermiştir. RMSD eşitleme hataları bu dört değişken arasında etkileşim olduğunu ortaya koymuştur. Örneğin, uzun testlerde (75 madde) örneklem büyüklüğü veya ortak maddelerin yüzdesi göz ardı edildiğinde ortalama faktör yükü 0.62 iken, Kernel eşitleme her zaman daha doğru sonuçlar vermiştir. Ortalama faktör yükü 0.50 olduğunda, büyük örneklerde ve ortak maddeler tüm testin %30'unu oluşturduğunda karakteristik eğri eşitleme yöntemleri daha doğru sonuçlar üretmiştir. Örneklem büyüklüğü ve madde sayısı en büyük değerlerini aldığı anda, S-L yöntemi Kernel ile aynı veya ondan daha doğru değerler üretmiştir; test uzunluğu ve güvenilirlik düzeyleri göz önüne alındığında ise Kernel daha

üstün bulunmuştur. Araştırmada tüm koşullar genel olarak incelendiğinde Kernel eşitleme yönteminin karakteristik eğri eşitleme yöntemine göre daha tercih edilebilir, iyi performans gösterdiği sonucuna ulaşılmıştır.

Speron (2009) simülatif veri ile çalıştığı araştırmasında, Madde Tepki Kuramı'na dayalı "iki moment (OO ve OS)" ve "iki karakteristik eğrisi (HB ve S-L)" eşitleme yöntemlerinden hesaplanan eşitleme hatalarını karşılaştırmıştır. Araştırmada, 1-0 şeklinde puanlanan 2PLM ile uyumlu simülatif veriler üzerinde "ortak maddeli eşitlenmemiş gruplar eşitleme deseni" kullanılmıştır. Eşitleme hatalarını etkileyen bağımsız değişkenler olarak; örneklem büyüklüğü, ortak maddelerin sayısı, ortak maddelerin toplam testi temsil etme derecesi ve grupların yetenek dağılımları ele alınmıştır. Araştırmanın sonucunda, karakteristik eğri yöntemleri ve O-O yöntemlerinin benzer sonuçlar verdiği ve daha az hatalı kestirime neden olduğu bulunmuştur. Bununla birlikte araştırmada O-S yönteminin diğer yöntemlerden farklı sonuçlar verdiği ve en fazla hataya sahip yöntem olduğu sonucuna ulaşılmıştır. Örneklem büyüklüğü arttıkça eşitleme hatalarının düştüğü gözlenmiştir.

Kilmen (2010) araştırmasında, Madde Tepki Kuramı'na dayalı "ortalama-ortalama", "ortalama-sigma", "Haebara" ve "Stocking-Lord" eşitleme yöntemlerinden hesaplanan eşitleme hatalarının yetenek dağılımı (benzer ve farklı yetenek dağılımı) ve örneklem büyüklüğü (500-1000 kişilik) değişkenlerine dayalı olarak karşılaştırılmasını amaçlamıştır. Araştırmada, 1-0 şeklinde puanlanan 3 parametrelili modelle uyumlu simülatif veri üzerinde "ortak maddeli eşitlenmemiş gruplar eşitleme deseni" kullanılmıştır. Araştırmanın verilerini üretmede WinGen2 programından yararlanılmış, bu program ile 600 adet 1-0 verisi oluşturmuştur. Simülasyon verilerine ait yetenek parametreleri BILOG-MG programı yardımıyla "beklenen a posteriori" (expected a posteriori) yöntemi kullanılarak kestirilmiştir. IRTEQ programı kullanılarak testlerin eşitlemesi için gerekli eşitleme denklemleri oluşturulmuştur.

Araştırma sonunda, 3 parametrelili modele uygunluk gösteren testler için; 500 ve 1000 kişilik benzer ve farklı yetenek dağılımına sahip grupların kullanıldığı durumlarda, Stocking-Lord yöntemi ile yapılan test eşitleme uygulamasının daha az hatalı eşitlemeler yaptığı bulunmuştur. Araştırmada en yüksek eşitleme hatalarını veren yöntemler ise ortalama-ortalama ve ortalama-sigma yöntemleridir.

Örneklem büyüklüğünün 1000 kişilik olduğu durumda eşitleme hataları 500 kişilik örneklemden elde edilen eşitleme hatalarına nazaran daha düşük bulunmuştur. Yetenek dağılımı değişkenine göre test eşitleme yöntemleri karşılaştırıldığında ise, benzer yetenek dağılımına sahip grupların eşitlenmesinden elde edilen eşitleme hatalarının, farklı yetenek dağılımına sahip grupların eşitlenmesinden elde edilen eşitleme hatalarına nazaran daha düşük olduğu bulunmuştur.

Lee ve Ban (2010), MTK'ya dayalı test eşitleme yöntemlerinden eş zamanlı kalibrasyon, Stocking Lord, Haebara ve yetenek dönüşümü (proficiency transformation) yöntemlerini karşılaştırmıştır. Araştırma verisi için İngilizce formlarının gerçek uygulama sonucu 3PLM ile elde edilen madde parametrelerinden faydalanılmıştır. Bunun için gerçek veriden her biri 3000 kişi olan rastgele eşdeğer gruplar seçilerek hesaplanan madde parametrelerinden simülatif veri oluşturulmuştur. Araştırmada bağımsız değişkenler olarak, örneklem büyüklüğü (500/3000), eşitlenen formların madde sayısı (25/75) ve yetenek dağılımları ($N(0,1)$, $N(.5,1)$, ve $N(1,1)$) ele alınmıştır. Yetenek kestirimleri maksimum olabilirlik (Maximum likelihood- ML) yöntemi ile gerçekleştirilmiştir (yetenek dönüşümü yöntemi için EAP, MAP ve ML karşılaştırılmıştır). Eşitlemenin değerlendirme kriteri olarak kare yanlılık ve varyans hesabına dayanan ortalama kare hata (Mean Squared Error= MSE) kullanılmıştır. Araştırma sonucunda örneklem büyüdükçe MSE eşitleme hatasının azaldığı; ayrı kalibrasyon yöntemlerinin, eşzaman kalibrasyon ve yetenek dönüşümü yöntemlerine göre daha iyi sonuçlar verdiği (üç örnekleminde aynı evrenden alındığı durum hariç); Haebara yönteminin Stocking-Lord yöntemiyle çoğu zaman eşit veya bazı koşullarda çok az farkla daha az ya da daha fazla eşitleme hatasına sahip olduğu; yetenek dönüşümü yönteminin ayrı kalibrasyon yöntemlerine göre hem madde sayısı hem de örneklem sayısı küçük olduğunda daha iyi sonuç verdiği tersi durumlarda genellikle en büyük hataya sahip olduğu bulunmuştur.

Pang ve diğ. (2010) yaptıkları araştırmada karma testlerde simülatif veri kullanarak ortak maddelerle eşdeğer olmayan gruplar deseni ile, üç sınıf düzeyinde yapılan dikey eşitlemede eşzaman kalibrasyon; ayrı kalibrasyonlardan O-S, test karakteristik eğrisi ve ayarlanmış ortak madde parametre (Fixed Common Item Parameter- FCIP) yöntemlerini karşılaştırmıştır. Bağımsız değişken olarak dört farklı test için θ ölçeğinde iki yıl arası, gerçek değişimler bilinmediğinden, beş olası

değişim durumu oluşturulmuştur (-0.3, -0.1, 0, 0.1, 0.3). Araştırmada EQAO programı ile eşitleme yapılmış ve eşitlemenin değerlendirme kriteri olarak eşitleme yanlılığı (bias) ve RMSE kullanılmıştır. Sonuç olarak eşzaman eşitleme yöntemi diğer üç yöntemden önemli şekilde büyük hatalar üretmiş; eşdeğer olmayan gruplarda grup performansı değişimi karşısında sağlam duramamıştır. Gruplarda yetenek değişimi (± 0.3) olduğunda tüm eşitleme yöntemlerinde daha büyük yanlılık elde edilmiştir. En büyük yanlılığı eşzaman eşitleme üretirken, aksine O-S, test karakteristik eğrisi ve FCIP daha küçük yanlılık üretmiştir. RMSE eşitleme hataları ise farklı koşullar için eşzaman eşitlemede geniş ranjda büyük elde edilirken; diğer yöntemlerde dar ranjda küçük hatalar elde edilmiştir.

Gök (2012) simülatif veri üzerinde çalıştığı araştırmasında, MTK'ya dayalı eşdeğer olmayan gruplarda ortak madde deseni kullanarak 2PLM ve 3PLM ile test eşitleme gerçekleştirmiştir. Araştırmada kullanılan bağımsız değişkenler örneklem büyüklüğü (1000 ve 3000), test uzunluğu (30, 60 ve 80 madde), yetenek dağılımı (benzer ve farklı yetenek dağılımları), model türü (2PLM ve 3PLM) ve test eşitleme yöntemleri (ortalama-ortalama, ortalama-standart sapma ve Stocking-Lord) olarak belirlenmiştir. Araştırmada yetenek ve madde parametreleri EAP kestirim yöntemi ile elde edilmiş ve eşitleme yanlılığı (BIAS) ile eşitleme hatası (RMSE) değerlendirme kriterleri hesaplanmıştır. Araştırmada sonuç olarak kullanılan model açısından, 3PLM'de en doğru kestirim Stocking-Lord, en hatalı kestirim ortalama-standart sapma yönteminde; 2PLM'de ise en doğru kestirim ortalama-ortalama, en hatalı kestirim ortalama-standart sapma yönteminde elde edilmiştir. 3PLM kullanıldığında en kararlı parametre kestirimlerine test uzunluğu 60'ın üzerinde ve örneklem büyüklüğü 1000'in üzerinde olduğunda bulunurken; 2PLM kullanıldığında test uzunluğu 30'dan 60'a çıktıkça hata ve yanlılık değerlerinin düştüğü gözlenmiş ve örneklem büyüklüğünün yanlılık üzerinde herhangi bir etkiye sahip olmadığı bulunmuştur. İki modelde de yetenek dağılımı benzer gruplarda, hata ve yanlılık değerinin farklı gruplara göre daha düşük olduğu ve 2PLM'nin daha az yanlı ve hatalı sonuçlar verdiği görülmüştür. Araştırmadan elde edilen sonuçlar genel olarak değerlendirildiğinde, en iyi eşitlemelerin 3000 kişilik örneklem, 80 maddelik testler, benzer yetenek dağılımına sahip gruplar, 2PLM ve SL yöntemi kullanılarak elde edilebileceği belirtilmiştir.

Meng (2012), simülatif, iki kategorili verinin kullanıldığı araştırmasında MTK'ya dayalı ortalama-sigma ve Stocking-Lord ile Kernel eşitleme yöntemlerini karşılaştırılmıştır. Araştırmada bağımsız değişkenler olarak, ortak madde uzunluğu (5 ve 15) ve yetenek dağılımının eşitleme üzerindeki etkileri incelenmiştir. Madde parametreleri 3PLM kullanılarak kestirilmiştir. Araştırma sonucunda, ortak maddelerinin sayısı arttıkça eşitleme hatasının azaldığı ve yetenek farklılıklarının eşitleme hatası üzerinde büyük etkiye sahip olduğu bulunmuştur. Şöyle ki, gruplar arası farklılık arttıkça eşitleme hatalarının da arttığı; farklı yetenek dağılımlarında S-L yönteminin OS yöntemine göre daha az hatalı sonuçlar verdiği gözlenmiştir. MTK'ya dayalı ortalama-sigma ve Stocking-Lord yöntemlerinin Kernel eşitleme yöntemine göre daha doğru sonuçlar verdiği belirtilmiştir.

Atalay Kabasakal (2014), "DMF'nin Test Eşitlemeye Etkisi" başlıklı simülatif veri ile çalıştığı araştırmasında DMF'li maddelerin bulunduğu testlerin Çok Düzeyli Madde Tepki Modeli (ÇDMTM) ve Madde Tepki Modeli'ne (MTM) dayalı ayrı ve eş zamanlı kalibrasyonlarda yapılan eşitlemeleri karşılaştırmıştır. Eşitlemede ÇDMTM için iki düzeyli 1PL-MTM, eş zaman eşitleme ile geleneksel tek grup eş zaman eşitleme ve karakteristik eğri dönüştürme yöntemlerinden S-L kullanılmıştır. Araştırmada bağımsız değişkenler olarak, test uzunluğu (20,40); örneklem büyüklüğü (500, 2000); DMF etki büyüklüğü (B ve C); DMF'nin bulunduğu test (ortakta, ortak olmayanda ve her ikisinde) olarak ele alınmıştır. DMF'li ortak madde oranı %10 ile sabit tutulmuştur. Eşitlemenin değerlendirme kriterleri olarak, madde parametreleri ve yetenek kestirimlerine ait eşitleme hataları (RMSE, BIAS ve SE) hesaplanmıştır. Eşitleme deseni olarak denk gruplarda ortak madde deseni kullanılmıştır. Sonuç olarak, örneklem büyüklüğü ve test uzunluğunun artması genel olarak MTM'ye dayalı eşitlemelerde olumlu etki yaratırken ÇDMTM bu iki koşuldaki MTM'ye göre daha az etkilenmiştir. MTM'ye dayalı yöntemler kendi içinde incelendiğinde ise DMF'li maddelerin varlığından, ayrı kalibrasyon yöntemlerinin eşzamanlı kalibrasyona göre daha çok etkilendiği; bu etkinin en çok DMF'li maddenin ortak testte olduğu ve DMF etki büyüklüğünün C olduğu koşullarda gerçekleştiği görülmüştür. S-L yönteminde DMF'li maddelerin silinmesi yetenek kestirimleri hatasını genel olarak yükseltmiştir; eşitlemede hatalardaki artışın DMF bulunan test ve DMF etki büyüklüğüne göre değiştiği ve artışın küçük örneklemde büyük örnekleme göre daha fazla olduğu sonucuna ulaşılmıştır. Tüm

simülasyon koşulları dikkate alındığında DMF'li maddelerin testten çıkarılmasının küçük örneklerde ve kısa testlerde eşitleme hatasını arttırdığı, büyük örnekler ve uzun testlerde ise eşitleme hatasını azalttığı sonucuna ulaşılmıştır

Huggins (2014), "Ortak Maddelerde DMF'nin Eşitlemede Grup Değişmezliği Üzerine Etkileri" başlıklı araştırmasında simülatif veride eşdeğer olmayan gruplarda ortak test deseni ile 3PLM'ye dayalı dayalı "iki moment (OO ve OS)" ve "iki karakteristik eğrisi (HB ve S-L)" eşitleme yöntemlerini kullanmıştır. Araştırmada DMF'li ortak maddelerin MTK gerçek puan eşitlemede grup değişmezliğine etkisini incelemek amaçlanmıştır. Eşitleme yapılacak formlar 10'u iç ortak madde olmak üzere 50'şer ikili puanlanan maddeden oluşmaktadır. Bağımlı değişken gerçek puan eşitlemede grup değişmezliğidir. Bağımsız değişkenler ise gruplar arası (odak ve referans) b parametreleri farkına dayalı DMF büyüklüğü (0.30= küçük, 0.60= orta, 0.90= büyük), ortak testte bulunan DMF'li madde sayısı (2,%20'si; 4,%40'ı; 6, %60'ı), DMF yönü (tek yönlü-iki yönlü), alt gruplarda (odak-referans) yetenek ortalamaları arası farklılık ve ortak formlarda DMF farklılığı durumlarıdır. Araştırmaya göre ortak formlarda DMF farklılığı, özellikle eşdeğer olmayan gruplarda eşitlemede ortak maddenin sadece bir formda DMF'li çıkması durumunda söz konusu olabilmektedir. Ortak formlarda DMF farklılığı olmaması ise iki formun ortak madde DMF'lerinde büyüklük, işaret için farklılık olmadığı anlamına gelmiştir. Araştırmada ortak maddelerde DMF değerlendirmesinin eşitleme değişmezliği veya eşitleme bağımlılığına etki edeceği hipotezi test edilmiştir. Eşitleme değişmezliğinin eksiliği durumunda aynı puana sahip bireylerin eşitlenmiş puanlarının farklı çıkacağı ve bu durumun alt gruplardan birine avantaj sağlayacağı belirtilmiştir. Araştırmada eşitlemelerde grup değişmezliği değerlendirme kriterleri olarak her koşul için RMSD(x), standartlaştırılmamış REMSD, RESD(x), RSD(x) indisleri hesaplanmıştır (Total ve odak-referans gruplar için). Ortak formlarda DMF farklılığı olmadığına, alt gruplarda gerçek ortalama yetenekler eşitken ve ortak maddelerde TB-DMF söz konusuyken eşitleme değişmezliğine ilişkin RMSD eşitleme hataları dört yöntemde çok benzer elde edilmiştir. S-L ve HB yöntemleri O-O ve O-S ile karşılaştırıldığında daha küçük RMSD eşitleme hataları vermiştir. O-O yöntemi ise O-S yöntemine göre daha küçük RMSD değerleri üretmiştir. Formlar arası ortak maddelerde DMF farklılığı olduğunda DMF yönü eşitleme alt gruplarda gerçek ortalama yetenekler eşitken ve

ortak maddelerde TB-DMF söz konusuysen ise eşitleme değişmezliği, özellikle DMF büyüklüğünün arttığı durumlarda, en çok O-S yönteminde bozulmuştur. O-O, HB ve S-L yöntemlerinde ise TB-DMF'li madde sayısının artması eşitleme değişmezliği hatalarında büyük değişiklik yaratmamıştır. Alt grup yetenek ortalamaları farklı olduğunda ise O-S yönteminde, iki yönlü DMF'li maddelerin büyüklüğüne bağlı olarak RMSD hataları artmıştır. Araştırmada özellikle TB-DMF'li ortak madde varlığının eşitleme bağımlılığı üzerindeki etkisinin daha büyük olduğu, DMF'nin büyüklüğünün, sayısının grup değişmezliğini negatif yönde etkilediği bulunmuştur. Araştırma bulguları karşılıklı alt gruplarda ve karşılıklı formlar arasında DMF'li ortak maddelerin değişik durumlarının eşitlemenin grup değişmezliğini bozabileceğini göstermiştir. Ortak madde düzeyinde değişmezliğin ve grup değişmezliği yoksunluğunun arasındaki bu ilişki nedeniyle uygulayıcıların ortak maddelerde belirli DMF tiplerini takip ederek eşitleme yapmaları önerilmiştir.

Uysal (2014), karma testlerde simülatif veri kullanarak yaptığı araştırmasında eşitlenmemiş gruplarda ortak madde deseni altında, MTK'ya dayalı "üç moment (OO, OS ve SOS)" ve "iki karakteristik eğrisi (HA ve S-L)" eşitleme yöntemlerinin hatalarını karşılaştırmıştır. Parametre kestirimleri iki kategorili maddeler için 3PLM, çok kategorili maddeler için GPCM (genelleştirilmiş kısmi kredi modeli) kullanılarak kestirilmiştir. Araştırmada bağımsız değişkenler olarak, ortak madde oranı (4, %10; 8, %20 ve 12, %30) ve yetenek dağılımı (normal-normal, sağa çarpık- sağa çarpık, sola çarpık-sola çarpık ve sağa çarpık-sola çarpık, sağa çarpık- normal, sola çarpık-normal)ele alınmıştır. Eşitlemenin değerlendirme kriteri olarak RMSD eşitleme hataları hesaplanmıştır. Sonuç olarak, ortak madde oranlarına ilişkin sonuçlar değerlendirildiğinde, benzer yetenek dağılımlarında genellikle %20 ve %30 ortak madde kullanıldığında hata ortalamalarının düştüğü gözlemlenmektedir. Farklı yetenek dağılımlarında ise, %10 ortak madde oranının %20 ve %30 ortak madde oranından daha düşük hata verdiği durumlara rastlanmaktadır. Bununla birlikte yetenek dağılımındaki farklılık arttıkça hata miktarının büyük oranlarda yükseldiği gözlemlenmiştir. Araştırmada grupların yetenek dağılımı aynı olduğunda ve özellikle ortak madde oranı %10'un üzerindeyken S-L ve HB yöntemlerinden elde edilen RMSD eşitleme hataları birbirine çok yaklaşmış hatta birbiriyle aynı bulunmuştur. Örneğin, N-N dağılımda %20 ortak madde oranında S-L ve HB yöntemleri birbiriyle eşit ve en düşük hatayı üretmiştir. Genel olarak bakıldığında,

SL yönteminin genellikle daha az hata verdiği görülmektedir. Benzer şekilde moment yöntemleri karşılaştırıldığında, en büyük hatalar SOS yönteminde elde edilmiştir. O-O ve O-S yöntemleri karşılaştırıldığında ise, iki grup da sağa veya sola çarpık dağılıma sahip olduğunda (tüm ortak madde oranlarında) ve normal-normal dağılımda (sadece %20 ortak maddede) O-O daha büyük hatalar üretmiştir. Yetenek dağılımındaki farklılık arttıkça hata miktarının büyük oranlarda yükseldiği gözlenmiştir.

2.1. İlgili Araştırmalar Özet

Madde Tepki Kuramı ayrı kalibrasyona dayalı “iki moment (OO ve OS)” ve “iki karakteristik eğrisi (HB ve S-L)” eşitleme yöntemlerinin DMF varlığında nasıl çalıştıklarını araştıran sınırlı araştırma bulunmaktadır (Han, 2008). DMF’li ortak maddelerin genel olarak ölçeği bozduğu; özellikle aynı anda hem TB hem TBO-DMF’li maddeler varlığında, tek bir çeşit DMF’li maddelere göre ölçek puanları daha çok bozulduğu sonucuna ulaşılmıştır (Turhan, 2006). Ayrıca hiyerarşik madde tepki modellerinin performansının madde parametreleri ve yetenek kestiriminde hem DMF’li hem de DMF’siz maddeler varlığında eşzaman eşitlemeden daha iyi olduğu belirtilmiştir (Chu, 2002).

Bununla beraber “iki moment (OO ve OS)” ve “iki karakteristik eğrisi (HB ve S-L)” eşitleme yöntemlerinden hesaplanan eşitleme hatalarının karşılaştırıldığı diğer araştırmalarda ortak maddelerin DMF’siz oldukları varsayılmaktadır. Söz konusu dört yöntemin birbiriyle veya eşzaman kalibrasyon, Kernel, teta regresyon, FCIP gibi yöntemlerle karşılaştırılmasına son yıllarda sıklıkla rastlanmaktadır. Bu araştırma sonuçlarına genel olarak bakıldığında, karakteristik eğrisi yöntemlerinin moment yöntemlerinden daha doğru sonuçlar ürettiği görülmekle birlikte; farklı koşullar için benzer sonuçlar verdiğini söyleyen araştırmalar da vardır (Cohen ve Kim,1998). Bazı istisnai durumlarda; örneğin çok kategorili maddeler için uyarlanmış S-L ve OS yöntemlerinin birbirine çok benzer sonuçlar verdiği (Tate, 2000); bazı durumlarda da O-S yönteminin en fazla hataya sahip yöntem olduğu gözlenmektedir (Speron, 2009).

Karakteristik eğrisi yöntemleri karşılaştırıldığında HB ve S-L yöntemlerinin eşitleme hatalarının bazı koşullarda çok az farklı oldukları ve çoğu zaman eşit hatalara sahip oldukları sonucuna ulaşılmıştır (Lee ve Ban, 2010). Karakteristik eğrisi yöntemleri ve eşzaman eşitleme ise en sağlam eşitleme yöntemleri olarak

çalışmıştır (Kim ve Kolen, 2006). Fakat farklı koşullarda (gruplararası yetenek değişimi büyüdüğünde) en büyük yanlılığı eşzaman eşitleme üretirken, aksine O-S, test karakteristik eğrisi ve FCIP yöntemlerinde daha küçük yanlılık ortaya konulduğu bulunmuştur (Pang ve diğ., 2010). S-L yönteminde DMF'li maddelerin silinmesi yetenek kestirimleri hatasını genel olarak yükseltmiştir, madde parametre kestirimleri hatasını ise düşürmüştür (Atalay Kabasakal, 2014).

Araştırmalarda ortak madde sayısı, ortak maddelerin toplam testi temsil etme derecesi, ortak madde türü (yapılandırılmış/çoktan seçmeli), yetenek dağılımı, test uzunluğu, parametre kestiriminde kullanılan model türü, kestirim yapılan program ve örneklem büyüklüğü sıklıkla ele alınan bağımsız değişkenler olarak göze çarpmaktadır. Bağımlı değişkenler olarak ele alınan yetenek kestirimi veya madde parametre kestirimleri doğruluk ve kararlık dereceleri, hesaplanan eşitleme hataları (RMSE, SEE, MSE, BIAS vb.) ile ortaya konulmuştur. Araştırmalarda ortak maddelerin sayısı, test uzunluğu ve örneklem büyüklüğü arttıkça eşitleme hataları azalmıştır. Eşdeğer gruplarda eşdeğer olmayan gruplara göre daha küçük eşitleme hataları elde edilmiştir. Bununla birlikte DMF'li ortak maddeler varlığında eşitlemenin grup değişmezliği de incelenen bir diğer bağımlı değişkendir (Huggins, 2014). Yapılan araştırmaların çok azında gerçek veri üzerinden çalışılmıştır (Baker ve Al Karni 1991; Karkee ve Wright, 2004; Turhan, 2006). Sonuç olarak her araştırmada ele alınan koşullar farklılaştığı için, yöntemlerden hangisinin tercih edileceğinin mevcut koşullar altında değerlendirilmesi gerekmektedir.

3. YÖNTEM

Bu bölümde araştırmanın türü, araştırma verileri, veri toplama aracı, evren ve örneklem, araştırma deseni, kullanılan programlar, verilerin toplanması ve verilerin analizi, varsayımların testi, güvenilirlik hesaplamaları başlıkları üzerinde durulmuştur.

3.1. Araştırmanın Türü

Bu araştırmada ortak maddelerin değişen madde fonksiyonu (DMF) içerip içermeme durumuna bağlı olarak, Madde Tepki Kuramı'na (MTK) dayalı test eşitleme yöntemleri karşılaştırılmaktadır. Örneklem bilgilerinin evrene genellenmesi amaçlanmamış, test eşitleme yöntemlerinde iki farklı durumda (DMF'li veya DMF'siz ortak maddeler varlığında) eşitleme hatalarının karşılaştırılması üzerine odaklanılmıştır. Bu gerekçelerle araştırma "nedensel karşılaştırma (casual comparative)" türündedir.

Nedensel karşılaştırma araştırmalarında, belli bir faktörün gözlenen sonuçlarda farklılık oluşturup oluşturmadığı incelenir; ancak incelenen faktörün sonuçlar üzerinde kesin olarak neden olduğu söylenemez. Bu araştırma olası nedenleri belirleme bakımından değerlidir (Büyüköztürk ve diğ., 2008).

Araştırma ayrıca, DMF'li madde elde etmek için oluşturulan testin psikometrik özelliklerinin incelenmesi, test maddelerinin cinsiyete göre DMF içerip içermediğinin araştırılması ve eşitlenecek formların psikometrik özelliklerinin ortaya konulması bakımından betimseldir.

3.2. Araştırma Verileri

Araştırmada iki farklı veri grubu üzerinde çalışılmıştır: 1) DMF Analizi için kullanılan veri grubu 2) Test eşitleme için kullanılan veri grubu.

Öncelikle araştırmanın birinci aşamasında DMF analizleri için Yenilik ve Eğitim Teknolojileri Müdürlüğü'nden (YEĞİTEK) elde edilen 2000-2008 yılları arasında yapılan OKS ve 2009-2012 yılları arasında yapılan SBS'ye giren 8. sınıf öğrencileri puanlarından oluşmaktadır. Dolayısıyla DMF analizi için gerekli OKS ve SBS verileri Yenilik ve Eğitim Teknolojileri Müdürlüğü'nden (YEĞİTEK), test eşitleme için yapılan gerçek uygulama verileri ise 29'ar maddeden oluşan A ve B

formlarından elde edilmiştir. Oluşturulan fen ve teknoloji testi A ve B formları Ek-8'de görülmektedir.

3.3. Veri Toplama Aracı

Ortak maddelerin DMF'li veya DMF'siz olduğu durumlarda test eşitlemesi yapmak için kullanılacak veri toplama aracının (fen ve teknoloji testi A ve B formu) hazırlanması üç aşamada gerçekleştirilmiştir:

- 1) DMF'li veya DMF'siz ortak maddelerin ve DMF'siz ortak olmayan maddelerin seçileceği madde havuzunun oluşturulması,
- 2) Madde havuzundan gerçek uygulama yapılacak DMF'li ve DMF'siz ortak maddeli paralel test formlarının elde edilmesi,
- 3) Uygulama sonunda maddelerin DMF'li ve DMF'siz olma durumlarının incelenerek eşitleme için test formlarına son halinin verilmesi.

Öncelikle eşitleme amacıyla oluşturulacak iki formun DMF'li veya DMF'siz ortak maddeleri ve DMF'siz ortak olmayan maddelerinin seçileceği madde havuzu oluşturulmuştur. Bunun için 2000-2012 yılları arasında uygulanan OKS ve SBS fen bilgisi alt testlerinde 2000-2007 yılları OKS'de çıkan 175 madde ve 2008-2012 yıllarında SBS'de çıkan 100 madde olmak üzere toplam 275 maddenin cinsiyete dayalı DMF analizi yapılmıştır. Tablo 3.3'te verilen ve her bir yıla ait evrenden tesadüfi örnekleme yoluyla seçilen $N=4000$ ile $N=20000$ arası örneklemlerden; $N=750$ ve $N=1000$ büyüklüğünde alt gruplar oluşturulmuştur. Oluşturulan alt grupların her birinde maddeler, MH ve LR yöntemleriyle DMF analizine tabi tutulmuştur. Maddelerin seçilmesinde iki yöntemde de alt gruplarda DMF'li/DMF'siz çıkma kriteri işe koşulmuştur.

Araştırmada DMF'li maddeleri belirlemede grup değişkeni olarak cinsiyet (kız-erkek) ele alınmıştır. DMF'li maddelerin aynı grup lehine çalışması beklenmiş; bu nedenle tüm DMF'li ortak maddeler erkekler lehine ve tek biçimli seçilmiştir. Bununla birlikte kullanılan MH ve LR yöntemlerinin ikisine göre de, B ve C düzeyinde DMF'li çıkan maddeler dikkate alınmıştır. LR yönteminde sadece TB-DMF içeren maddeler için ikinci model ile birinci model arasındaki ikili puanlama için hesaplanan Nagelkerke R^2 değerleri farkı (ΔR^2) ≥ 0.010 olan maddeler ile MH yöntemine göre $-1 \leq \Delta MHI < -1.5$ aralığında olan maddeler B düzeyinde DMF'li alınmıştır. Aynı şekilde LR yöntemine göre ikinci model ile birinci model

arasındaki Nagelkerke R^2 değerleri farkı (ΔR^2) ≥ 0.020 olan maddeler C düzeyinde TB-DMF'li ve MH yöntemine göre $-1.5 \leq \Delta MHI$ olan maddeler B düzeyinde referans grup olan erkekler lehine DMF'li kabul edilmiştir. Yapılan analizler sonucunda maddeler DMF'li ve DMF'siz olmak üzere gruplandırılmıştır.

İkinci aşama olarak DMF analizleri yapılan maddelerden gerçek uygulaması yapılacak A ve B formları oluşturulmuştur. B ve C düzeylerinde DMF'li maddelerin sayısı toplam incelenen madde sayısına oranla oldukça az olduğundan sadece bazı alt gruplarda B veya C düzeyinde DMF'li bulunan maddeler de ortak DMF'li madde olarak veri toplama aracına kabul edilmiştir. Ayrıca uygulama yapılacak formlarda ortak maddeler dışında kalan maddeler DMF'siz olanlardan seçilmiştir. Maddelerin kapsam bakımından paralelliğin sağlanabilmesi için ölçtükleri davranış bakımından kategorilendirilmiştir. Bu kategorilendirme yapılırken bilimsel süreç becerileri konusunda çalışma yapmış, fen bilgisi, ölçme ve değerlendirme, program geliştirme alanlarında yüksek lisans ve doktora düzeyinde uzmanlaşmış beş kişi ve iki fen bilgisi öğretmenin görüşlerine başvurulmuştur.

Uygulamanın olması gereken bir ders süresinden (40 dk) daha kısa sürebileceği göz önüne alınarak madde sayıları çok yüksek tutulmamış ve öğrencilerin tüm sorulara ulaşabilmesi hedeflenmiştir. Gerçek uygulamaya hazır formların son hali Tablo 3.1'de gösterilmektedir.

Tablo Hata! Belgede belirtilen stilde metne rastlanmadı..1: **Gerçek Uygulamada Kullanılan A ve B Formları Madde Sayıları**

Formlar	DMF'siz 15 Madde	Ortak Maddeler		DMF'siz 15 Madde	Toplam Madde Sayısı
		8 DMF'li Madde	6 DMF'siz Madde		
A	√	√	√		29
B		√	√	√	29

Tablo 3.1'e göre, 29'ar maddeli A ve B formlarında ortak olmayan DMF'siz çıkması öngörülerek formlara konulan maddelerin sayısı 15 iken, ortak maddelerin sayısı 8'i DMF'li 6'sı DMF'siz olmak üzere toplam 14'tür. Bu maddelerin OKS-SBS yıllarına ve madde numaralarına göre dağılımları Tablo 3.2'de gösterilmiştir.

Tablo 3 1: A ve B Formlarında Maddelerin Sınav Yılı ve Madde Numaralarına Göre Dağılımları

	<i>Madde No</i>	<i>Form A Yıl /Madde no</i>	<i>Form B Yıl /Madde no</i>		<i>Madde No</i>	<i>Form A Yıl /Madde no</i>	<i>Form B Yıl /Madde no</i>
DMF'siz ORTAK OLMAYAN	1	2001/21	2011/19	DMF'siz ORTAK	16	2003/4	
	2	2004/18	2001/24		17	2003/7	
	3	2000/17	2001/20		18	2010/10	
	4	2010/16	2008/8		19	2009/13	2011/14
DMF'li ORTAK	5	2001/3		DMF'siz ORTAK OLMAYAN	20	2006/11	2011/15
	6	2001/1			21	2008/14	2007/14
	7	2000/10			22	2009/11	2009/10
	8	2001/8			23	2003/18	2006/8
	9	2003/2			24	2007/21	2006/20
	10	2010/13			25	2000/21	2001/14
	11	2011/1			26	2000/16	2003/1
	12	2012/1			27	2011/11	2011/9
DMF'siz ORTAK	13	2008/17		28	2012/19	2009/18	
	14	2005/18		29	2011/5	2009/7	
	15	2005/17					

Ek-2'de ayrıntılı olarak verilen, gerçek uygulama sonucunda 29'ar maddeli A ve B formlarına ait DMF analizleri sonuçları Tablo 3.3'de özetlenmiştir.

Tablo 3 2: A ve B Formu Cinsiyete Dayalı DMF-MH ve DMF-LR Analizlerine İlişkin Bulgular

Madde No	A FORMU				B FORMU			
	DMF-MH		DMF-LR		DMF-MH		DMF-LR	
	Δmh	Düzyey	ΔR^2 ($R_2^2-R_1^2$)	Düzyey	Δmh	Düzyey	ΔR^2 ($R_2^2-R_1^2$)	Düzyey
1	0.603		0.005		0.972*	A	0.007*	A
2	-0.571		0.004		-0.336		0.001	
3	-0.883		0.007*	A	0.127		0.000	
4	0.702		0.003		-0.655		0.005	
5**	-1.233*	B	0.018*	B	-1.754*	C	0.035*	C
6**	-2.384*	C	0.047*	C	-2.317*	C	0.045*	C
7**	-1.587*	C	0.025*	C	-1.669*	C	0.029*	C
8**	-0.550		0.002		0.761		0.006	
9**	-0.249		0.001		-0.624		0.004	
10**	-1.028*	B	0.011*	B	-1.126*	B	0.012*	B
11**	0.160		0.001		-0.156		0.000	
12**	-0.077		0.000		-0.655		0.005	
13**	0.246		0.000		0.506		0.003	
14**	0.729		0.009*	A	0.880		0.008*	A
15**	0.012		0.000		1.226	B	0.016*	B
16**	-0.178		0.000		0.853		0.006	
17**	-0.084		0.000		0.124		0.000	
18**	-0.102		0.000		0.784		0.006	
19	0.807*	A	0.006*	A	1.209	B	0.012*	B
20	-0.334		0.001		0.143		0.000	
21	1.333*	B	0.009*	A	0.274		0.000	
22	0.699		0.005		0.343		0.001	
23	0.294		0.001		-0.214		0.001	
24	0.521		0.003		1.193	B	0.013*	B
25	0.960*	A	0.008*	A	0.675		0.002	
26	0.350		0.001		0.292		0.001	
27	-0.235		0.001		-1.808*	C	0.029*	C
28	0.860		0.003		0.666		0.004	
29	1.025*	B	0.011*	B	0.752		0.005	

*p<0,05 ** 5-18 maddeler ortak maddelerdir (5-12 DMF'li, 13-18 DMF'siz öngörülen maddeler)

Tablo 3.3. incelendiğinde, ortak madde olarak hazırlanan ve erkekler lehine TB-DMF'li çıkması öngörülen 8 ve 9 nolu maddeler TBO-DMF'li 11 ve 12. Nolu maddeler DMF'siz çıktıkları için A ve B formundan çıkarılmışlardır (bkz. Ek-2). Erkekler lehine TB-DMF'li ortak madde çıkması öngörülerek formlara yerleştirilen 5, 6, 7 ve 10 nolu maddeler ise beklenildiği gibi erkekler lehine tek biçimli B ve C

düzeyinde DMF'li bulunmuştur ve DMF'li ortak maddeler olarak A ve B formuna kabul edilmiştir. Bununla birlikte ortak madde olarak hazırlanan ve DMF'siz çıkması öngörülen 14. madde LR yöntemine göre A ve B formunda A düzeyinde TBO-DMF'li; 15. madde ise hem MH hem de LR yöntemine göre sadece B formunda, B düzeyinde kızlar lehine TBO-DMF'li bulunmuştur (bkz EK 2). Dolayısıyla DMF'siz ortak madde olma koşulunu sağlamadıkları için A ve B formundan çıkarılmışlardır. Ortak olmayan maddeler ise DMF'siz çıkması öngörülen maddelerden hazırlanmıştır. Uygulama sonrası yapılan analiz sonuçlarına göre bazı maddelerin bu durumu ihlal ettiği gözlenmiştir. Bu maddelerden 19, 21, 24 ve 25 sadece tek bir formda DMF verdikleri için kapsam geçerliğini düşürmemek adına formlara kabul edilmiştir. Diğer taraftan 27. maddenin B formunda MH ve LR yöntemlerine göre C düzeyinde erkekler lehine TB-DMF'li ve 29. maddenin de A formunda her iki yöneme B düzeyinde kızlar lehine TB-DMF'li bulunması ve bu iki maddenin karşılıklı formlarda kapsam eşdeğerliği göstermemesi nedeniyle A ve B formundan çıkarılmışlardır. Sonuç olarak 1,2, 3, 4,19, 20, 21, 22, 23, 24, 25, 26 ve 28 nolu maddeler DMF'siz ortak olmayan maddeler olarak A ve B formunda yer almıştır. Formlarda yapılan değişiklikler sonrası A ve B formlarının son durumu Tablo 3.4'te verilmiştir:

Tablo 3 3: Eşitleme Yapılacak A ve B Formlarının Son Durumdaki Madde Sayıları

<i>Formlar</i>	<i>DMF'siz 13 Madde</i>	<i>Ortak Maddeler</i>		<i>DMF'siz 13 Madde</i>	<i>Toplam Madde Sayısı</i>
		<i>4 DMF'li Madde</i>	<i>4 DMF'siz Madde</i>		
<i>A</i>	√	√	√		21
<i>B</i>		√	√	√	21

Tablo 3.4'e göre eşitleme yapılacak formlardan DMF'li ortak maddeli A formu ve DMF'li ortak maddeli B formu ile DMF'siz ortak maddeli A formu ve DMF'siz ortak maddeli B formununun 17'şer maddeden oluştuğu görülmektedir. Eşitleme yapmak için son hali verilen formların ortak olmayan maddelerinin ölçtükleri davranışlar EK-7'de verilmiştir.

3.4. Evren ve Örneklem

Araştırmada DMF analizi için 2000-2012 yılları arasında çalışılan her bir yıla ait evrenden tesadüfi örnekleme yoluyla seçilen örneklem sayıları aşağıda verilen Tablo 3.5'teki gibidir:

Tablo 3 4: DMF Analizi için Kullanılan Evren ve Örneklem Sayıları

<i>Sınav Türü</i>	<i>Evren</i>	<i>Örneklem</i>	<i>Sınav Türü</i>	<i>Evren</i>	<i>Örneklem</i>
2000 OKS	419216	4000	2007 OKS	818364	15000
2001 OKS	553495	5000	2008 OKS	905929	15000
2003 OKS	600289	6000	2009 SBS	1011219	20000
2004 OKS	634787	10000	2010 SBS	1008307	20000
2005 OKS	768696	10000	2011 SBS	1070148	20000
2006 OKS	798307	10000	2012 SBS	1075546	20000

Tablo 3.5'te belirtilen örneklemlerden 750 ve 1000'er kişilik alt gruplar oluşturulmuş ve maddelerin DMF tutarlılığı bakımından incelemesi yapılmıştır. Elde edilen sonuçlar farklı OKS ve SBS yıllarında yapılan Yurdugül (2003); Bekçi (2007) ve Acar'ın (2008) araştırma sonuçlarıyla tutarlı görünmektedir.

Test eşitleme çalışması için gerçek veri uygulaması tercih edilmiştir. Bu kapsamda evrenden örneklem seçme yoluna gidilmeyip, çalışma grubu üzerinde uygulama yapılmıştır. Araştırmanın çalışma grubunu 2012-2013 eğitim-öğretim yılında sekiz ilde (Ankara, Kütahya, Eskişehir, Denizli, Bursa, Kahramanmaraş, Malatya, İstanbul) 8. sınıfta öğrenim gören 1576 ulaşılabilen öğrenci oluşturmuştur. Ancak bazı okullardan gelen verinin ayrı incelenmesi sonucu, tek boyutluluk ve normallik sayıltılarını karşılamama, atlanan veya ulaşılamayan maddeler barındırma, aykırı standart değerlere sahip olma gibi nedenlerle bu örneklemler araştırmada kullanılmamıştır. Söz konusu durumlara sahip kişiler çıkarıldığında örneklem 1350 öğrenciden oluşmuştur. Tablo 3.6'da araştırma örnekleminin cinsiyete göre dağılımı verilmiştir.

Tablo 3 5: Arařtırmada Kullanılan A ve B Formlarının Cinsiyete Gre Daęılımı

<i>Cinsiyet</i>	<i>A Formu</i>		<i>B Formu</i>		<i>Toplam</i>
	<i>N</i>	<i>Yzde (%)</i>	<i>N</i>	<i>Yzde (%)</i>	
<i>Kız</i>	360	52.1	312	47.3	672
<i>Erkek</i>	331	47.9	347	52.7	678
<i>Toplam</i>	691	100.0	659	100.0	1350

Tablo 3.6'da grldę gibi A formunda %52.1 oranında kızların ve %47.9 oranında erkeklerin yer aldıęı; B formunda ise %47.3 oranında kızların ve %52.7 oranında erkeklerin bulunduęu grlmektedir. A ve B formlarında kız ve erkek oranlarının birbirine olduka yakın olduęu sylenebilir.

3.5. Arařtırmanın Deseni

Bu arařtırmada test eřitleme iin eřdeęer gruplarda uygulanan ortak test/madde deseni (equivalent groups (EQ) design with anchor test) kullanılacaktır. Bu desen iki rneklem grubu ortak bir evrenden alındıęı zaman ortaya çıkmaktadır (Holland ve Dorans, 2006). Eřdeęer gruplar ortak madde deseninde eřdeęer yetenek dzeyinde rastgele seilmiş iki grup bir testin farklı iki formunu almaktadır. ęrencilerin ęretim programına uygun olarak hazırlanmış bu formlar ortak maddeler barındırmaktadır.

3.6. Kullanılan Programlar

1) DMF gsteren maddelerin belirlenmesi amacıyla Mantel-Haenzel yntemi EASYDIF (González ve dię., 2011) programı ile; Lojistik regresyon yntemi ise, Zumbo (1999) tarafından hazırlanan syntax kullanılarak SPSS programı ile gerekleřtirilmiştir.

2) MTK varsayımlarından tek boyutluluk varsayımı iin LISREL (Jreskog ve Srborn, 1986) programı kullanılarak doęrulatory faktr analizi yapılmış, yerel baęımsızlık varsayımı iin STATISCA programı ile elde edilen tetrakorik korelasyonlardan model veri uyumu test edilmiştir.

3) BILOG-MG (Zimowski ve dię., 1996) programı ile 2 parametrelili lojistik modelin veri ile uyumlu olduęu tespit edilerek, beklenen a posteriori (Expected A Posteriori,

EAP) parametre kestirim yöntemi kullanılarak madde parametreleri kalibre edilmiş ve yetenek kestirimleri gerçekleştirilmiştir.

4) MTK'ya dayalı ortalama-ortalama, ortalama-sigma ve karakteristik eğrisi yöntemleri için IRTEQ (Han, 2009) programı kullanılarak iki form arasında eşitleme yapmak için gerekli eşitleme katsayıları (A ve B) elde edilmiştir.

5) Her bir eşitleme yöntemi için eşitleme katsayıları kullanılarak EXCEL programı ile RMSD eşitleme hataları değerlendirme kriteri olarak hesaplanmıştır.

3.7. Verilerin Analizi

Bu başlıkta test eşitleme yapılacak A ve B formlarına (ADMF'li, BDMF'li ve ADMF'siz, BDMF'siz) ait gerçekleştirilen analizlere yer verilmiştir. Tablo 3.7'de eşitleme yapılacak A ve B formlarına ait KTK'ya dayalı betimsel istatistikler görülmektedir.

Tablo 3 6: Eşitlenecek A ve B Formlarına Ait KTK'ya Dayalı Betimsel İstatistikler

	<i>A FORMU</i>		<i>B FORMU</i>	
	<i>Ortak Maddeler</i>		<i>Ortak Maddeler</i>	
	<i>DMF'li</i>	<i>DMF'siz</i>	<i>DMF'li</i>	<i>DMF'siz</i>
<i>Madde Sayısı</i>	17	17	17	17
<i>Ortak Madde Sayısı</i>	4	4	4	4
<i>N</i>	691	691	659	659
<i>A. Ortalama</i>	8.441	9.036	8.470	9.089
<i>A. Ortalama SH</i>	0.139	0.150	0.143	0.143
<i>Mod</i>	7	8	8	7
<i>Medyan</i>	8	9	7	9
<i>Standart Sapma</i>	3.674	3.953	3.681	3.693
<i>Varyans</i>	13.505	15.626	13.550	13.644
<i>Çarpıklık (SH)</i>	0.276 (0.093)	0.087 (0.093)	0.192 (0.095)	0.055 (0.095)
<i>Çarpıklık / SH</i>	0.026	0.935	2.021	0.579
<i>Basıklık (SH)</i>	-0.831(0.186)	-0.878(0.186)	-0.849(0.19)	-0.813(0.190)
<i>Minimum</i>	1	1	1	1
<i>Maksimum</i>	17	17	17	17
<i>Ortalama r_{yx} (biserial)</i>	0.416	0.483	0.427	0.435
<i>Ortalama güçlük p</i>	0.496	0.531	0.498	0.534

Tablo 3.7 incelendiğinde, eşitleme yapılacak formlardan A DMF'li ortak maddeli ile B DMF'li ortak maddeli ve A DMF'siz ortak maddeli ile B DMF'siz ortak maddeli formların toplam 17'şer maddeden oluştuğu görülmektedir. Her bir formda 4 ortak madde bulunmaktadır. Eşitlenecek formların ortalama güçlük (p) ve nokta çift serili ayırt edicilik indeksleri, r_{jx} (biserial), bakımından birbirine benzer olduğu ve 0.50 civarında elde edildikleri görülmektedir. Formların ortalama-mod-medyan değerlerinin birbirine yakın olması; çarpıklık basıklık katsayılarının sıfır civarında olması ve çarpıklık/SH ile hesaplanan z istatistiklerinin $\alpha= 0.01$ için 2.58'den küçük çıkması sebepleriyle tüm formların normal dağılıma uygun oldukları söylenebilir (Büyüköztürk, 2005).

Eşitleme yapılacak formların MTK'ya dayalı madde parametre ve yetenek kestirimleri yapılmadan önce MTK varsayımları test edilmiş ve veriye uygun MTK parametre kestirim modeline karar verilmiştir. MTK varsayımlarının testi aşağıda başlıklar halinde açıklanmıştır.

3.8. Tek Boyutluluk Varsayımın Test Edilmesi

Test eşitlemede kullanılan A ve B formlarının OKS ve SBS'de çıkmış fen ve teknoloji maddelerinden oluştuğu bilinmektedir. Söz konusu formların tek boyutlu olma hipotezinin incelenmesi için LISREL programı ile doğrulayıcı faktör analizi yapılmıştır. Burada amaç, kurulan modele ilişkin beklenen kovaryans matrisi ile örneklemden gözlenen kovaryans matrisi arasındaki uyumun çıkarımını yapmaktır. Modelin veri ile uyumunun değerlendirilmesinde çeşitli uyum ölçütleri kullanılmaktadır. Bunlar, Ki-kare (χ^2) Uyum Testi (Chi-Square Goodness of Fit), KiKare/ Serbestlik derecesi (Sd) Uyum Testi, Yaklaşık Hataların Ortalama Karekökü (Root Mean Square Error of Approximation, RMSEA), Normlaştırılmış Uyum İndeksi (Normed Fit Index, NFI), Normlaştırılmamış Uyum İndeksi (Non-Normed Fit Index, NFI), Karşılaştırmalı Uyum İndeksi (Comparative Fit Index, CFI), Fazlalık Uyum İndeksi (Incremental Fit Index, IFI), Göreli Uyum İndeksi (Relative Fit Index, RFI), Uyum İyiliği İndeksi (Goodness of Fit Index, GFI), Düzeltilmiş Uyum İyiliği İndeksi (Adjusted Goodness of Fit Index, AGFI) uyum ölçütleridir. Aşağıda Tablo 3.8'de bu uyum ölçütlerine ilişkin araştırmada elde edilen değerlere yer verilmiştir.

Tablo 3 7: A ve B Formlarına Ait Doğrulayıcı Faktör Analizi Sonuçları

<i>Model Uyum Ölçütleri</i>	<i>A FORMU</i>		<i>B FORMU</i>		<i>Beklenen Uyum Değeri</i>
	<i>A DMF'li</i>	<i>A DMF'siz</i>	<i>B DMF'li</i>	<i>B DMF'siz</i>	
χ^2 / Sd	1.771	1.881	2.275	1.603	< 3
<i>RMSEA</i>	0.033	0.036	0.044	0.030	0'a yakın
<i>NFI</i>	0.92	0.94	0.88	0.93	1'e yakın
<i>NNFI</i>	0.96	0.97	0.92	0.97	1'e yakın
<i>CFI</i>	0.97	0.97	0.93	0.97	1'e yakın
<i>IFI</i>	0.97	0.97	0.93	0.97	1'e yakın
<i>RFI</i>	0.91	0.93	0.86	0.91	1'e yakın
<i>GFI</i>	0.97	0.96	0.95	0.97	1'e yakın
<i>AGFI</i>	0.96	0.95	0.94	0.96	1'e yakın

Model veri uyumu değerlendirilmesinde Ki-kare'nin Serbestlik derecesine bölünmesiyle elde edilen χ^2/Sd oranının 3'ten küçük olması modelin veriyle uyumlu olduğuna işaret etmektedir (Carmines ve Mclver, 1981). Bazı araştırmacılar ise Ki-kare/Sd oranının 2'den küçük olmasını (Byrne, 1989), bazıları da büyük örneklerde 5'ten küçük olmasını (Kelloway,1996) kabul edilebilir uyum olarak ele almaktadır.

RMSEA değeri ideal olarak 0'a yakın istenmekle birlikte 0.05'ten küçükse mükemmel, 0.08'den küçükse iyi düzeyde (Jöreskog ve Sörbom, 1993) ve 0.10'dan küçükse zayıf düzeyde uyum olduğu kabul edilir (Tabachnick ve Fidell, 2007). NFI, NNFI, CFI, IFI, RFI, GFI ve AGFI indeksleri teorik olarak 0 ile 1 arasında değerler almakla birlikte 1'e yakın olması model veri uyumunun sağlandığına, 0.95'ten büyük değer ise mükemmel uyuma karşılık gelmektedir. Tüm formlarda sıklıkla kullanılan GFI'nın 0.95 ve üzerinde bulunması verinin modellerle mükemmel uyumuna işaret etmektedir. Ki-kare (χ^2) uyum iyiliği hipotez testinde p değerinin >0.05 ile anlamlı olmaması beklenmektedir. Ancak ki-kare uyum değerlendirmesi çok tutarlı bir değerlendirme kabul edilmemektedir; çünkü büyük örneklerde beklenen ve gözlenen kovaryans matrisleri arasındaki çok önemsiz farklar bile anlamlı çıkabilmektedir. Uyum araştırmalarında indeksler arasında da tutarsızlık olabildiği ve raporda en az dört indeksin sonuçlarına yer verilmesi gerektiği belirtilmektedir (Byrne, 1989; Kline,1998; Schumacker ve Lomax, 2004; Tabachnick ve Fidell, 2007).

Tablo 3.8'de görüldüğü gibi tüm Ki-kare/Sd oranları 3'ten küçük, RMSEA değerleri 0'a yakın ve 0.05'ten küçük; diğer indeksler ise beklenildiği gibi 1'e yakın elde edilmiştir. Her bir forma ait ölçme modeli grafikleri Ek 3'te verilmiştir. Tüm formlara ait istenilen uyum indeksi değerlerinin elde edilmesi gerekçesiyle, fen sorularından oluşan formların tek bir özelliği ölçtüğü sonucuna ulaşılmıştır.

3.9. Yerel Bağımsızlık Varsayımın Test Edilmesi

Bireylerin bir test içindeki farklı maddelere verdikleri cevapların istatistiksel olarak birbirinden bağımsız olmasına dayanan yerel bağımsızlık varsayımında, farklı yetenek gruplarına ilişkin maddeler arası korelasyonların karşılaştırılmaktadır. Buna göre yerel bağımsız bir testte, aynı yetenek düzeyindeki kişilerden oluşan alt gruplara ait maddeler arası korelasyonların tüm gruptan elde edilen korelasyonlardan daha küçük (sıfıra yakın) olması beklenmektedir. Bu varsayımdan, "tüm yetenek grupları için maddeler arası korelasyon yoktur" sonucu çıkarılmamalıdır. Madde puanları belli bir yetenek düzeyinde ilişkisizdir. Bu aynı zamanda tek boyutluluk demektir. Dolayısıyla tek boyutlu bir test aynı zamanda yerel bağımsız kabul edilebilir. Ancak yerel bağımsız bir test tek boyutlu olmayabilir (Hambleton ve Swaminathan,1985: 23-24).

Bu araştırmada A ve B formlarında tek boyutluluk varsayımı karşılanmış olmakla birlikte, bununla yakından ilişkili yerel bağımsızlık varsayımının testi yoluna da gidilmiştir. Crocker ve Algina (1986:342-343) yerel bağımsızlığın testinde, belli yetenek düzeylerinde elde edilen madde puanları arasındaki ikili korelasyonların karşılaştırılmasını önermiştir. Bu nedenle araştırmada, tüm gruptaki maddeler arası tetrakorik korelasyonlar, daha dar (homojen) yetenek dağılımında olan alt ve üst yetenek gruplarından elde edilen maddeler arası tetrakorik korelasyonlarla karşılaştırılmıştır.

Test puanı aralıklarının belirlenmesinde puan dağılımının alt ve üst uçlarındaki %27'lik puan dilimleri ölçüt olarak alınmıştır. Bunun nedeni, alt ve üst grupların hem evreni temsil etme gücünün en yüksek olduğu (örneklem güvenirligi) hem de gruplar arasındaki farkın maksimum yansıtıldığı fonksiyonun %27'lik ölçüte karşılık gelmesidir (Kelley,1939). Elde edilen korelasyonlara ilişkin betimsel istatistikler Tablo 3.9'da verilmiştir.

Tablo 3.8: Farklı Yetenek Gruplarından Elde Edilen Maddeler Arası Tetrakorik Korelasyonlara Ait Betimsel İstatistikler

	<i>Yetenek Grupları</i>	<i>Puan Aralıkları</i>	<i>N</i>	<i>Ortalama Korelasyon</i>	<i>Minimum</i>	<i>Maksimum</i>
<i>A Formu (DMF'li Ortak Maddeli)</i>	<i>Tüm</i>	1-17	691	0.233	-0.107	0.566
	<i>Üst</i>	12-17	187	-0.059	-1.000	0.573
	<i>Alt</i>	1-6	187	-0.101	-0.731	0.533
<i>B Formu (DMF'li Ortak Maddeli)</i>	<i>Tüm</i>	1-17	659	0.250	-0.007	0.472
	<i>Üst</i>	11-17	178	-0.103	-1.000	0.555
	<i>Alt</i>	1-6	178	-0.084	-0.622	0.440
<i>A Formu (DMF'siz Ortak Maddeli)</i>	<i>Tüm</i>	1-17	691	0.297	-0.027	0.574
	<i>Üst</i>	12-17	187	-0.115	-1.000	0.677
	<i>Alt</i>	1-6	187	-0.073	-0.539	0.424
<i>B Formu (DMF'siz Ortak Maddeli)</i>	<i>Tüm</i>	1-17	659	0.258	-0.007	0.472
	<i>Üst</i>	12-17	178	-0.111	-1.000	0.595
	<i>Alt</i>	1-7	178	-0.064	-0.634	0.380

Tablo 3.9 incelendiğinde formlarda homojen olan üst ve alt gruptan elde edilen korelasyonların ortalamasının heterojen olan tüm gruba ilişkin ortalamadan düşük ve sıfıra oldukça yakın olduğu gözlenmektedir. Ranj küçüldükçe korelasyonların düşmesi yerel bağımsızlığın sağlandığını desteklemektedir.

3.10. Kalibrasyon ve Madde Parametreleri ile Yetenek Kestirimleri

Araştırmada farklı gruplardan elde edilen madde parametrelerini aynı ölçeğe dönüştürmek için ayrı kalibrasyon yöntemi kullanılmıştır. Ayrı kalibrasyon sonrası eşitleme, ortak maddelerde bireysel olarak yer alabilecek bir takım problemleri de ortaya çıkarma potansiyeline sahiptir (Hanson ve Béguin, 2002; Kolen ve Brennan, 2004). Örneğin DMF'li ortak maddeler geçerlik ve güvenilirliği zedeleyici etki yapabilir. Ayrı kalibrasyonun, araştırma kapsamında ele alınan DMF'li ortak maddelerin madde parametre kestirimlerinin incelenmesine imkân tanıyacağı düşünülmüştür.

Ayrı kalibrasyon yönteminde uygun parametre sayısı belirlenerek (1, 2, 3 PLM) farklı gruplara ait madde parametreleri arasında doğrusal ilişki ortaya konulmaktadır. Elde edilen A ve B katsayıları ile iki grubun madde parametre ve yetenek kestirimleri ortak ölçeğe dönüştürülmektedir. Bu süreç şu şekilde

işlemektedir: Bireylerin maddelere verdikleri cevaplardan yola çıkılarak seçilen modele göre (1, 2, 3 PLM) a, b ve c madde parametreleri kestirilmektedir. Daha sonra MTK varsayımları ışığında madde parametrelerinden bireylerin yetenekleri kestirilmektedir. Gerçek (gözlenen) test verisi ile beklenen (kestirilen) madde parametreleri ve yetenekler arasında en yüksek uyum yakalanana kadar madde parametreleri ve yeteneklerin kestirim işlemi devam etmektedir. Yetenek kestirimi içinse kullanılacak yöntemin seçilmesi gerekmektedir. İki kategorili maddelerin puanlamasında kullanılan yetenek kestirim yöntemleri, Maksimum Olabilirlik (Maximum Likelihood, ML), Bayes modellerden Maksimum A Posteriori (Maximum A-Posteriori, MAP) ve Beklenen A Posteriori (Expected A Posteriori, EAP)'dir. Bu yöntemlerde testle ölçülmek istenen gizil değişkenin (yetenek) ortalaması sıfır, standart sapması bir olan normal dağılım gösterdiği kabul edilmektedir. Araştırmada maddelerin tamamı doğru veya yanlış yapıldığı durumlarda tüm modellere göre kestirim yapabilen, önceden belirlenmiş bir θ önsel dağılımının ortalamasını kullanan ve kolay hesaplanabilen Bayes modellerden EAP yöntemi tercih edilmiştir (Embretson ve Reise, 2000).

Swaminathan ve Gifford (1985) simülasyon ile yürüttükleri araştırmalarında 2PLM ile Bayes kestirimlerinin Maksimum Olabilirlik kestirimlerinden daha iyi olduğu sonucuna ulaşmışlardır. Bayes parametre kestirimlerinin ranj dışı kayma göstermedikleri için daha anlamlı olduğunu ve gerçek kestirilen değerler arası daha küçük ortalama kare farkına (MSD) sahip oldukları için daha doğru sonuçlar ürettiğini belirtmişlerdir.

Maddelerin 1, 2 ve 3PLM için kestirilen madde parametre değerlerine ilişkin betimsel istatistikler aşağıda Tablo 3.10, Tablo 3.11 ve Tablo 3.12'de verilmiştir.

Tablo 3.9: 1PL Model İçin Madde Parametrelerine Ait Betimsel İstatistikler

<i>Model</i>	<i>Formun Adı</i>	<i>Madde Parametresi</i>	<i>Ortalama</i>	<i>St. Sapma</i>	<i>Min.</i>	<i>Max.</i>
<i>1PLM</i>	<i>A Formu (DMF'li Ortak Maddeli)</i>	b	0.010	0.652	-1.349	1.179
	<i>B Formu (DMF'li Ortak Maddeli)</i>	b	0.002	0.731	-1.299	0.957
	<i>A Formu (DMF'siz Ortak Maddeli)</i>	b	-0.167	0.617	-1.355	0.928
	<i>B Formu (DMF'siz Ortak Maddeli)</i>	b	-0.197	0.716	-1.384	0.942

Tablo 3.10: 2PL Model İçin Madde Parametrelerine Ait Betimsel İstatistikler

<i>Model</i>	<i>Formun Adı</i>	<i>Madde Parametresi</i>	<i>Ortalama</i>	<i>St. Sapma</i>	<i>Min.</i>	<i>Max.</i>
<i>2PLM</i>	<i>A Formu (DMF'li Ortak Maddeli)</i>	a	1.012	0.498	0.385	1.936
		b	0.152	0.967	-1.839	2.446
	<i>B Formu (DMF'li Ortak Maddeli)</i>	a	0.975	0.299	0.381	1.373
		b	0.084	0.824	-1.476	1.931
	<i>A Formu (DMF'siz Ortak Maddeli)</i>	a	1.149	0.423	0.452	1.924
		b	-0.054	0.737	-1.220	1.431
<i>B Formu (DMF'siz Ortak Maddeli)</i>	a	0.998	0.286	0.412	1.396	
	b	-0.118	0.739	-1.278	1.386	

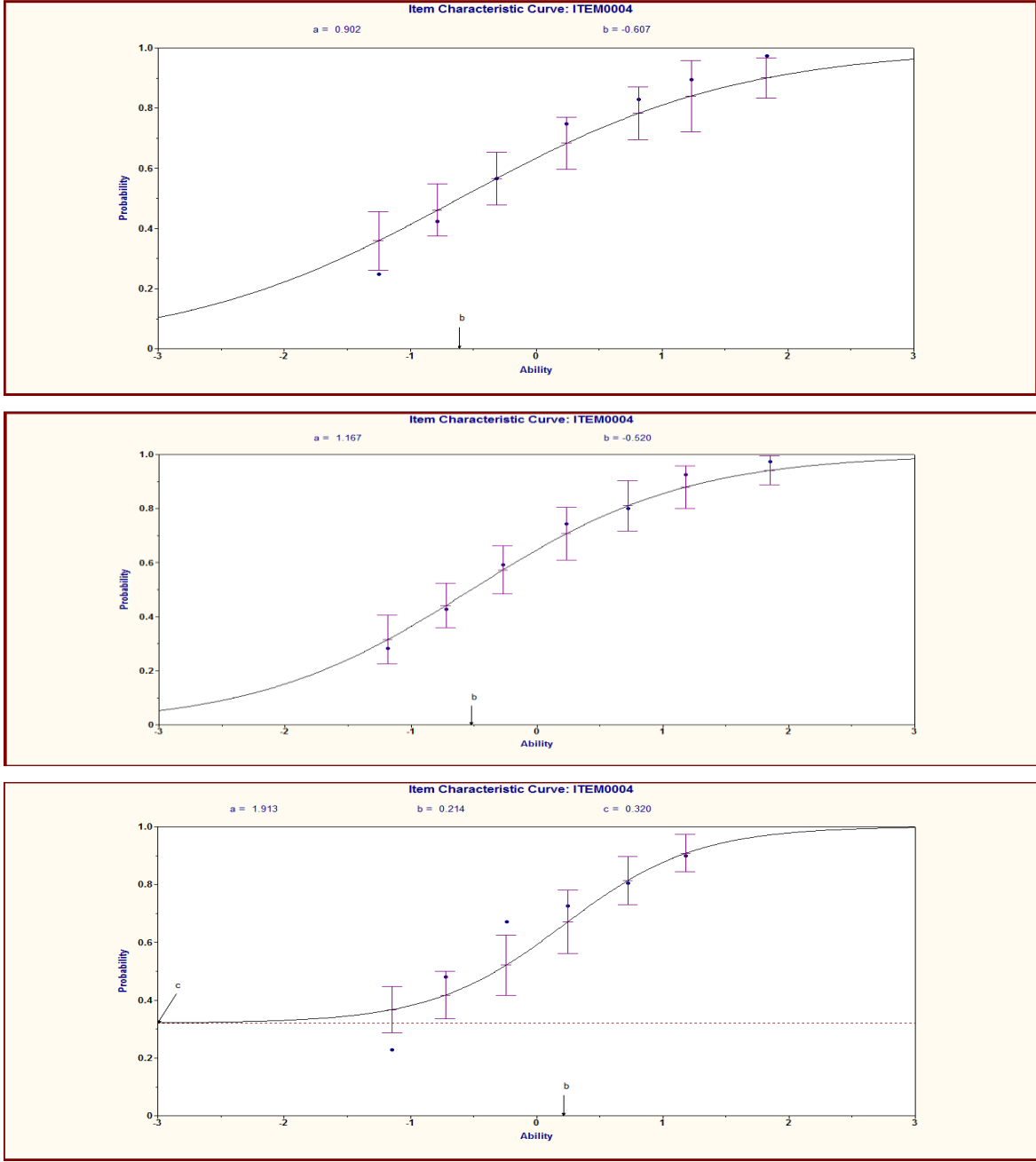
Tablo 3.11: 3PL Model İçin Madde Parametrelerine Ait Betimsel İstatistikler

<i>Model</i>	<i>Formun Adı</i>	<i>Madde Parametresi</i>	<i>Ortalama</i>	<i>St. Sapma</i>	<i>Min.</i>	<i>Max.</i>
<i>3PLM</i>	<i>A Formu (DMF'li Ortak Maddeli)</i>	a	2.018	0.652	0.812	3.120
		b	0.852	0.824	-0.176	2.954
		c	0.289	0.071	0.194	0.444
	<i>B Formu (DMF'li Ortak Maddeli)</i>	a	1.878	0.661	0.984	3.137
		b	0.765	0.711	-0.148	2.364
		c	0.279	0.072	0.170	0.435
	<i>A Formu (DMF'siz Ortak Maddeli)</i>	a	2.080	0.569	1.359	3.284
		b	0.510	0.686	-0.657	1.815
		c	0.269	0.044	0.193	0.329
	<i>B Formu (DMF'siz Ortak Maddeli)</i>	a	1.854	0.686	1.056	3.394
		b	0.585	0.726	-0.504	2.110
		c	0.289	0.068	0.175	0.396

3.11. Model Veri Uyumunun Kontrolü

Madde Tepki Kuramı modellerinin uyum iyiliği değerlendirmelerinde tek boyutluluk, yerel bağımsızlık, hız testi olmama gibi varsayımların test edilmesi dışında hangi modelin veri ile uyumlu olduğunun raporlaştırılması da gerekmektedir. Düşük model-veri uyumu varlığında yetenek ve madde parametrelerinde değişmezlikten söz edilmesi mümkün değildir. Uyum iyiliği analizinde artıkların (residuals) bilgisine başvurmak çok değerli bir yoldur. Artıkların analizi yolunda tüm yetenek düzeylerinde her bir maddenin beklenen ve gözlenen doğru cevaplama oranları arasındaki fark incelenmekte ve bu fark madde karakteristik eğrileri ile açık bir şekilde gözlenmektedir. Madde karakteristik eğrilerinde artıklar küçük ve rastgele dağılım gösterdiğinde model veri uyumunun yakalandığı sonucuna ulaşılmaktadır (Hambleton ve diğ., 1991).

Bu araştırmada A DMF'li ortak maddeli, A DMF'siz ortak maddeli, B DMF'li ortak maddeli ve B DMF'siz ortak maddeli formların tüm maddelerinde 1, 2 ve 3 parametrelili lojistik modellerden elde edilen madde karakteristik eğrileri incelenerek artıkların analizi değerlendirilmiştir. Şekil 3.1'de rastgele seçilmiş, ortak maddelerin DMF'li olduğu A formunun 4. maddesine ait sırasıyla 1, 2 ve 3 PLM için artıkları gösteren MKE'ler verilmiştir.



Şekil 3.1: A Formu (Ortak Maddeler DMF'li) 4. Maddeye Ait 1, 2, 3 PLM Uyum İyiliği Testleri

Şekil 3.1'deki grafikler incelendiğinde, madde karakteristik eğrilerinde artıkların küçük ve rastgele dağılım gösterdiği modelin 2PLM olduğu gözlenmektedir. Tüm formlara ait madde karakteristik eğrileri EK 5'te, test bilgi eğrileri ise EK 6'da sunulmuştur.

Model varsayımları testinde bir diğer yol olarak BILOG-MG programı ile 1, 2 ve 3 PLM için elde edilen χ^2 istatistiğinin anlamlılık düzeyi incelenerek testin ve maddelerin kullanılan modellerle uyumu incelenmiştir. Büyük χ^2 değerleri beklenen

(teorik) ve gözlenen (ampirik) MKE'ler arasında farkın büyüklüğüne dolayısıyla model veri uyumsuzluğuna işaret etmektedir. H_0 hipotezinin manidarlık testinde hesaplanan p değerinin 0.01'den büyük çıkması modelde beklenen ve gözlenen dağılımlar arasında uyum olduğunun göstergesi olmuştur (Hambleton ve Swaminathan, 1985).

Tablo 3.12:A ve B Formlarının 1, 2, 3PLM'e Dayalı Uyum İyiliği Sonuçları

<i>Formlar</i>	<i>Toplam χ^2 (SH) / Model İle Uyumlu Madde Sayısı</i>		
	<i>1PLM</i>	<i>2PLM</i>	<i>3PLM</i>
<i>N_A= 691, N_B= 659, N=17</i>			
<i>A Formu</i> <i>(DMF'li Ortak Maddeli)</i>	447.0 (120.0)/ 6	163.8 (122.0)*/ 17	249.2 (111.0)/ 12
<i>B Formu</i> <i>(DMF'li Ortak Maddeli)</i>	263.9 (124.0)/ 10	144.7 (126.0)*/ 17	252.8 (121.0)/ 11
<i>A Formu</i> <i>(DMF'siz Ortak Maddeli)</i>	472.3 (129.0)/ 6	249.8 (128.0)/ 13	329.2 (125.0)/ 11
<i>B Formu</i> <i>(DMF'siz Ortak Maddeli)</i>	290.6 (127.0)/ 10	180.1 (127.0)/ 15	317.5 (130.0)/ 10

*p> 0.01

Tablo 3.13'te verilen testlerin bütününe ait Ki-kare (SH) ve anlamlılık (p) değerlerine bakıldığında DMF'li ortak maddeli A ve B formunun ki-kare (p>0.01) ile 2PLM'ye göre model veri uyumu sağladığı; diğer formların sözkonusu modellerle (1PLM, 2PLM, 3PLM) uyumsuz olduğu görülmektedir. MTK model uyumu için yapılan pek çok çalışmada madde düzeyindeki uyuma bakılmaktadır. Bir testte bir tip MTK modeli ile tüm maddelerin uyumlu çıkması oldukça olası bir durumdur (Kang ve Cohen, 2007). Araştırmada maddelerin tek tek model uyumu incelendiğinde, iki parametrelili lojistik modellerle uyumlu (p>0.01) madde sayısının tüm testlerde en fazla olduğu gözlenmektedir. Yapılan incelemeler sonucunda madde parametreleri ve yeteneklerin kestirilmesinde 2PLM'nin kullanılmasının uygun olacağına karar verilmiştir.

Madde parametre ve yeteneklerinin 2PLM'ye dayalı kestirimlerinden sonra A ve B formlarından elde edilen MTK'ya dayalı betimsel istatistikler aşağıda Tablo 3.14'te verilmiştir.

Tablo 3.13: A ve B Formlarının MTK'ya Dayalı Betimsel İstatistikleri

	<i>A FORMU Ortak Maddeler</i>		<i>B FORMU Ortak Maddeler</i>	
	<i>DMF'li</i>	<i>DMF'siz</i>	<i>DMF'li</i>	<i>DMF'siz</i>
<i>Madde Sayısı</i>	17	17	17	17
<i>Ortak Madde Sayısı</i>	4	4	4	4
<i>θ Ortalama</i>	-0.001	-0.001	-0.001	-0.001
<i>θ St. Sapma</i>	0.034	0.034	0.034	0.034
<i>θ Çarpıklık (SH)</i>	0.323 (0.093)	0.224 (0.093)	0.270(0.095)	0.163(0.095)
<i>θ Çarpıklık / SH</i>	3.473*	2.408	2.842*	1.715
<i>θ Basıklık (SH)</i>	-0.749(0.186)	-0.721(0.186)	-0.664(0.190)	-0.641(0.190)
<i>μ(b) Toplam</i>	0.152	-0.053	0.084	-0.118
<i>μ(b) Ortak Madde</i>	0.164	-0.705	0.108	-0.756
<i>μ(a) Toplam</i>	1.011	1.148	0.975	0.998
<i>μ(a) Ortak Madde</i>	0.529	1.132	0.839	0.905

* $\alpha = 0.01$ için >2.58 (normal dağılım yok)

Tablo 3.14 incelendiğinde tüm formların yetenek (θ) ortalama ve standart sapmalarının eşit olduğu görülmektedir. DMF'li ortak maddeli formların normal dağılımdan saptığı, sağa çarpık dağılım gösterdikleri, DMF'siz ortak maddeli formların ise normal dağılım koşulunu sağladıkları göze çarpmaktadır. Kestirilen yeteneklere dayalı normal dağılım eğrileri EK 4'te verilmiştir.

3.12. A ve B Formlarının KTK ve MTK'ya Dayalı Güvenirlik Katsayıları

İki testi alan bireylerin puan dağılımları eşitse ve bu iki test eşit derecede güvenilirse test puanlarının karşılaştırılabilir olduğu söylenebilir (Lord, 1950). Araştırmada her bir 1-0 puanlamalı eşitlenecek form için klasik test kuramına dayalı KR-20 ve madde tepki kuramına dayalı Lord'un güvenirlik katsayıları hesaplanmıştır. MTK'ya dayalı Lord'un güvenirlik katsayısı iç tutarlılık anlamında KR-20'ye benzer şekilde bilgi vermektedir. KR-20 güvenirlik katsayısı madde varyansı/test varyansı oranından; Lord'un güvenirlik katsayısı ise gerçek puan varyansı/puan varyansı oranından hareket etmektedir. Güvenirlik analizi sonuçları Tablo 3.15'te görülmektedir.

Tablo 3.14: A ve B Formlarının KTK ve MTK'ya Dayalı Güvenirlik Katsayıları

<i>Formlar</i>	<i>KTK'ya Dayalı KR-20 Güvenirlik Katsayısı</i>	<i>MTK'ya Dayalı Lord'un Güvenirlik Katsayısı</i>		
		<i>1PLM</i>	<i>2PLM</i>	<i>3PLM</i>
<i>A Formu (DMF'li Ortak Maddeli)</i>	0.744	0.736	0.778	0.729
<i>B Formu (DMF'li Ortak Maddeli)</i>	0.752	0.742	0.762	0.719
<i>A Formu (DMF'siz Ortak Maddeli)</i>	0.792	0.782	0.800	0.772
<i>B Formu (DMF'siz Ortak Maddeli)</i>	0.756	0.748	0.766	0.737

Tablo 3.15 incelendiğinde MTK'ya dayalı Lord'un güvenirlik katsayıları en yüksek 2PLM'de elde edilirken, 1PLM ve 3 PLM'de Lord'un güvenirlik katsayıları KR-20 güvenirlik katsayılarından düşük bulunmuştur. Özdemir'in (2004) araştırmasında benzer şekilde 2PLM'de Lord'un güvenirlik katsayısı yardımıyla elde edilen güvenirliliğin, KTK'ya dayalı KR-20 güvenirliliğinden yüksek sonuçlar verdiği bulunmuştur.

3.13. Eşitleme Hatasının Hesaplanması

Bu araştırmada A ve B formlarının ortak maddeleri DMF'li ve DMF'siz olduğu durumlarda yapılan eşitlemelere karışan hata eşitlik (3.2) ile gösterilen RMSD kriteri kullanılarak hesaplanmıştır. Bu kriter bireylerin gerçek yetenek düzeyi ile almadıkları testten kestirilen yetenek düzeyleri arası farkın kare ortalamalarının kareköküne dayanmaktadır (Paek ve Young, 2005; Harris ve Crouse, 1993):

$$RMSD = \sqrt{\frac{\sum_i f_i (\theta^* - \theta)^2}{\sum_i f_i}} \quad (3.2)$$

θ = Gerçek yetenek düzeyi

θ^* = Kestirilen yetenek düzeyi

f = Yetenek düzeylerine ait frekans

Eşitlik (3.2)'de her bir bireyin kestirilen yetenek düzeyi (θ^*) ile gerçek yetenek düzeyi (θ) farkının karesi hesaplanmaktadır. Daha sonra elde edilen değerler toplanarak yetenek düzeylerine ait frekansa bölünüp, sonucun karekökleri alınmaktadır. Bulunan RMSD değerlerinin küçük olması, gerçek ve kestirilen yetenek düzeyleri arasındaki farkın da küçük olduğunu; yani eşitleme hatasının az olduğunu belirtmektedir.

IRTEQ programıyla madde parametreleri eşitlenirken iki farklı grup ele alınmıştır. Bunlar temel grup ve karşılaştırılacak gruptur. Yine aynı şekilde yeniden ölçekleme yapılacak temel test ve karşılaştırılacak test söz konusudur. Burada süreç şöyle işlemektedir:

Araştırmada temel test B formudur. Temel test formu aynı zamanda yeniden ölçeklendirilen test formudur. A formu ise hedef form ya da karşılaştırılacak test formudur. Temel test formunu alan; ancak bu formla ortak maddeler içeren A formunu almayan bireylerin gerçek yeteneklerinden (θ) yola çıkarak, almadıkları A formundaki kestirilen yetenekleri (θ^*) aşağıdaki eşitlik (3.3) ile hesaplanmıştır:

$$\theta^* = A\theta + B \quad (3.3)$$

θ = Bireyin gerçek yetenek düzeyi (B formu)

θ^* = Bireyin almadığı testteki kestirilen yetenek düzeyi (A formu)

A = Eşitleme denkleminin eğimi

B = Eşitleme denkleminin sabiti

O-O, O-S, Haebara ve S-L eşitleme yöntemlerine ait A ve B eşitleme katsayıları ile temel formu (B) alan bireylerin karşılaştırma yapılacak formdaki (A) yeni yetenekleri kestirilmiş; aralarındaki farklar ile hatalar belirlenmeye çalışılmıştır.

4. BULGULAR VE TARTIŞMA

Bu bölümde problem ve alt problemlerin sırasına göre verilmiş araştırma bulguları ve bu bulgularla ilgili değerlendirmeler yer almaktadır. Alt problemlere dayalı bulgulara geçmeden önce ortak maddelerin TB-DMF'li ve DMF'siz olduğu durumlarda ortalama-ortalama (O-O), ortalama-sigma (O-S), Haebara (HB) ve Stocking Lord (S-L) yöntemlerine göre hesaplanan RMSD eşitleme hataları bir arada Tablo 4.1'de verilmiştir.

Tablo 4.1: Ortak Maddelerin DMF'li ve DMF'siz Olduğu Durumlarda Dört Yönteme göre Yapılan Eşitlemelerin RMSD Eşitleme Hataları

<i>Eşitleme Yöntemleri</i>	<i>RMSD Eşitleme Hataları</i>	
	<i>Ortak Maddeler DMF'li</i>	<i>Ortak Maddeler DMF'siz</i>
<i>Ortalama-Ortalama</i>	0.518	0.202
<i>Ortalama-Sigma</i>	0.221	0.248
<i>Haebara</i>	0.485	0.181
<i>Stocking-Lord</i>	0.468	0.181

Ortak maddelerin TB-DMF'li ve DMF'siz olduğu durumlarda ortalama-ortalama (O-O), ortalama-sigma (O-S), Haebara (HB) ve Stocking Lord (S-L) yöntemlerine göre hesaplanan A ve B eşitleme katsayıları, ortak maddelere ait a ve b parametreleri ortalamaları, korelasyonları Tablo 4.2'de görülmektedir.

Tablo 4.2: Ortak Maddelerin DMF'li ve DMF'siz Olduğu Durumlarda Dört Eşitleme Yöntemine göre Elde Edilen A ve B Eşitleme Katsayıları

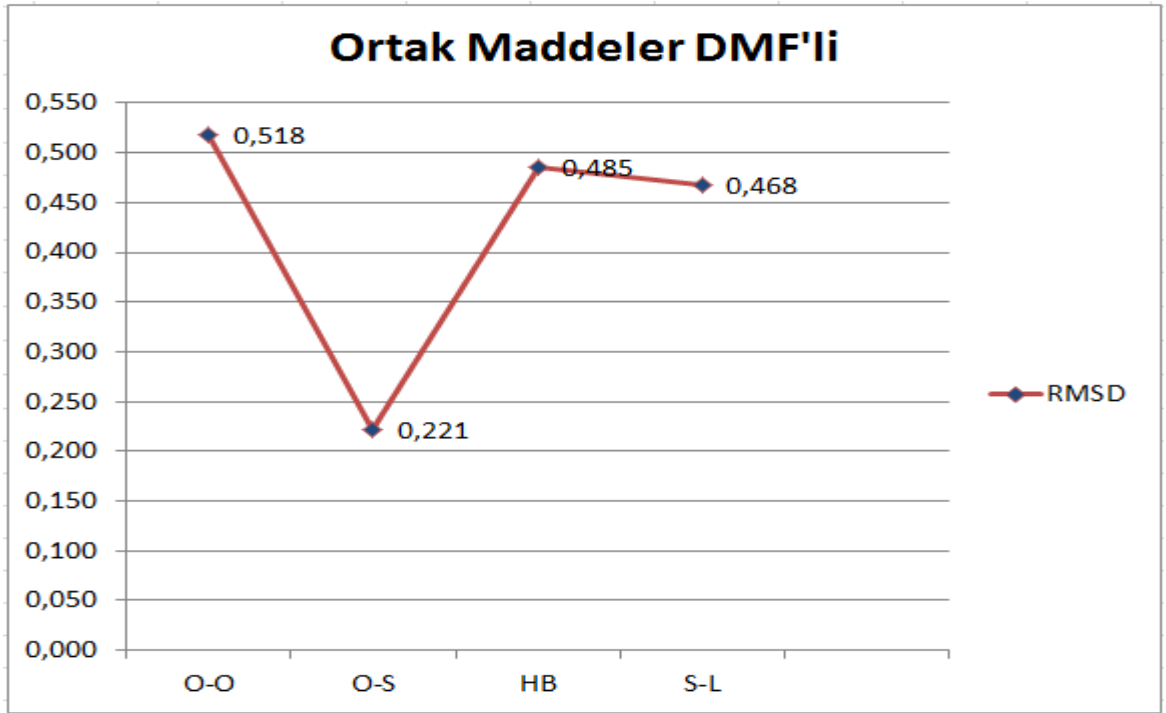
<i>Eşitleme Yöntemleri</i>	<i>Eşitleme Katsayıları</i>			
	<i>Ortak Maddeler DMF'li</i>		<i>Ortak Maddeler DMF'siz</i>	
	<i>A</i>	<i>B</i>	<i>A</i>	<i>B</i>
<i>Ortalama-Ortalama</i>	1.59	-0.01	0.80	-0.10
<i>Ortalama-Sigma</i>	1.25	0.03	0.76	-0.13
<i>Haebara</i>	1.55	0.05	0.81	-0.07
<i>Stocking-Lord</i>	1.53	-0.05	0.81	-0.07

Araştırmanın ana problemini oluşturan dört eşitleme yöntemine göre RMSD eşitleme hatalarının karşılaştırılması yapılırken Tablo 4.1'de elde edilen sonuçlar alt problemlere göre grafikler halinde verilerek somutlaştırılmıştır.

4.1. Ortak Maddelerin DMF'li Olduğu Durumda Uygulanan MTK'ya Dayalı Test Eşitleme Yöntemlerinden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular

Ortak maddelerin TB-DMF'li olduğu durumlarda her bir yöntemden elde edilen RMSD eşitleme hataları Tablo 4.1'de görülmektedir. Tablo 4.1'e göre A ve B formlarında ortak maddeler TB-DMF'li olduğunda ortalama-ortalama yöntemine dayalı eşitleme hatası 0.518, ortalama-sigma yöntemine dayalı eşitleme hatası 0.221, Haebara'nın karakteristik eğrisi dönüştürme yöntemine dayalı eşitleme hatası 0.485 ve Stocking-Lord'un karakteristik eğrisi dönüştürme yöntemine dayalı eşitleme hatası 0.468 bulunmuştur.

Aşağıda ortak maddelerin DMF'li olduğu durumda dört farklı eşitleme yöntemine göre elde edilen RMSD eşitleme hataları grafiksel olarak Şekil 4.1'de görülmektedir.



Şekil 4.1: Ortak Maddeler DMF'li Olduğunda Eşitleme Hataları (RMSD)

Şekil 4.1'de görüldüğü gibi ortak maddeler TB-DMF'li olduğunda en büyük RMSD eşitleme hatası ortalama-ortalama yönteminde elde edilirken, en küçük hata ortalama-sigma yönteminde bulunmuştur. Kolen ve Brennan (2004) b parametreleri kestirimlerinin a parametreleri kestirimlerinden daha kararlı olmaları nedeniyle O-S yöntemini O-O yöntemine göre bazen daha tercih edilebilir bulmaktadır. O-O yöntemi ölçükleri eşitlemek için direkt olarak a ve b

parametrelerinin ortalamalarını kullanmakta, A ve B eşitleme katsayılarını daha ileri rafine edememektedir (Baker ve Al-Karni, 1991). Araştırmada O-O yönteminin en büyük eşitleme hatasını verme nedeninin, ortak maddelerde a ve b parametreleri ortalamalarının A ve B formlarında gösterdiği farklılık olduğu düşünülmektedir. Diğer yöntemlerde eşitlenen iki formun ortak maddeleri arasında (O-S yönteminde b parametre standart sapmalarında, karakteristik eğri yöntemlerinde karakteristik eğrileri arasında) daha küçük farklar olmasının eşitleme hatalarının daha küçük çıkmasına neden olduğu söylenebilir.

Bununla birlikte Haebara ve Stocking-Lord karakteristik eğrisi eşitleme yöntemlerinin birbirine çok yakın eşitleme hatalarına sahip olduğu gözlenmektedir. HB yöntemi S-L yönteminden 0.017 farkla daha büyük RMSD eşitleme hatasına sahiptir. S-L yöntemi karakteristik eğrisi yöntemleri içerisinde en az hataya sahip eşitleme yöntemi olmuştur. HB ve S-L yöntemlerinin birbirine yakın eşitleme hataları vermesinin hesaplanma şekillerinin benzerliğinden kaynaklandığı söylenebilir. İki yöntem de madde karakteristik eğrileri arasındaki farka dayanmaktadır. Haebara yönteminde belli bir yetenek düzeyindeki bireyler için her bir maddenin madde karakteristik eğrileri arasındaki farkın karelerinin toplamı kullanılırken S-L yönteminde her bir maddenin madde karakteristik eğrileri toplamları farkının karesi işe koşulmaktadır.

Araştırma bulgularını Hu ve diğ.'nin (2008) yürüttüğü araştırma destekler niteliktedir. Hu ve diğ. (2008) araştırmalarında ortak maddelerde aykırı (outlier) değerlerin bulunup bulunmamasının eşitleme yanlılığı (bias) ve RMSE eşitleme hatalarına etkisine bakmıştır. Aykırı ortak madde bir formun ortak maddelerindeki tüm b parametrelerinin diğer formun b parametrelerinden iki puan düşük olması olarak tanımlanmıştır. Araştırma sonucunda, test karakteristik eğrisi ve O-S dönüşümleri en iyi eşitleme performansına sahip bulunmuştur. Eşdeğer gruplarda ortak madde setinde aykırı değer olmadığında karakteristik eğrisi ve O-S yöntemleri eşit derecede iyi performans gösterdiği belirtilmiştir. İki grubun yetenek ortalamaları aynı olduğunda dört yöntem eşdeğer hatalar üretirken; ortalamalar bir standart sapma değiştiğinde O-S ve karakteristik eğri yönteminin küçük hatalar ürettiği bulunmuştur.

Huggins'in (2014) araştırmasında ortak formlarda DMF farklılığı olup olmaması durumlarının iki moment (O-O, O-S) ve iki karakteristik eğrisi (HB, S-L)

yöntemlerinde eşitleme değişmezliğine etkisi DMF'li ortak madde sayısı, DMF büyüklüğü, DMF yönü, alt gruplarda (odak-referans) yetenek ortalamaları farklılığı koşullarında bakılmıştır. Ortak formlarda DMF farklılığı, özellikle eşdeğer olmayan gruplarda eşitlemede ortak maddenin sadece bir formda DMF'li çıkması durumunda söz konusu olabilmektedir. Ortak formlarda DMF farklılığı olmaması ise iki formun ortak madde DMF'lerinde büyüklük, işaret için farklılık olmadığı anlamına gelmiştir. Eşitleme değişmezliği (eşitleme bağımlılığı) ise aynı puana sahip farklı alt gruplarda bulunan bireylerin eşitlenmiş puanlarının da aynı olması olarak tanımlanmıştır. Araştırmada DMF'li ortak maddeler söz konusuysa bu durumun eşitleme değişmezliğini bozabileceği düşünülmüştür.

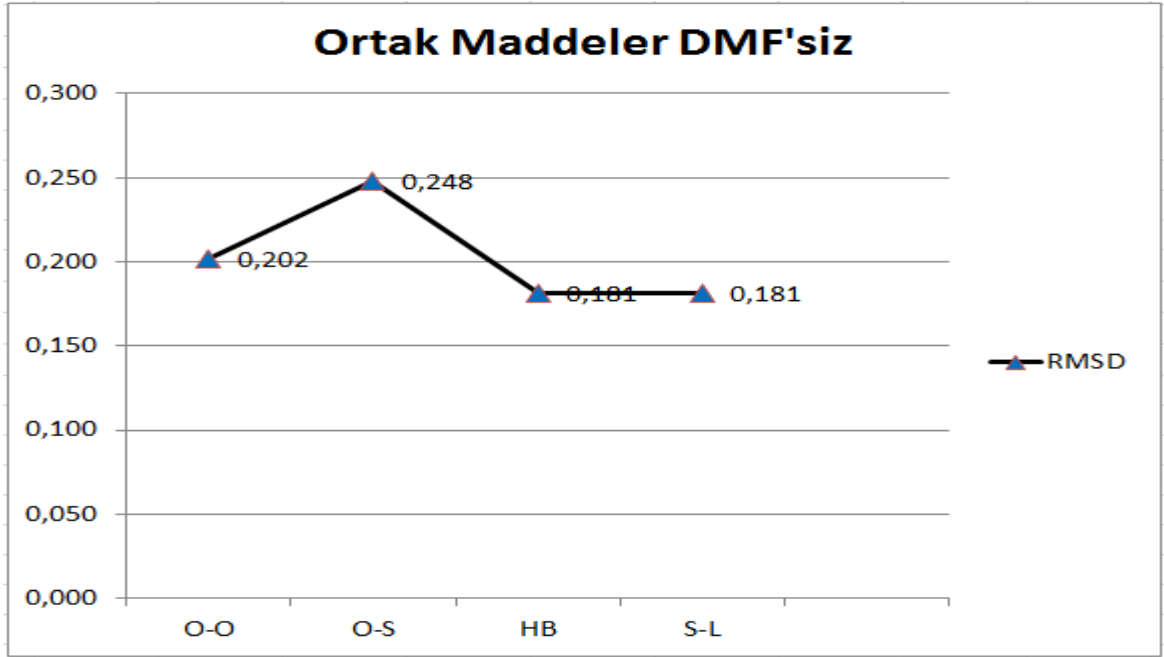
Bu araştırmada ortak formlarda herhangi bir DMF farklılığı yoktur; çünkü tüm ortak maddeler birbiriyle aynı büyüklükte ve erkekler lehine TB-DMF'lidir. Tüm DMF'li ortak maddeler referans grup (erkekler) lehine DMF'li olduğu için (-) işaretlidir. Bu koşullara benzer şekilde Huggins'in (2014) araştırmasında ortak formlarda karşılıklı farklılık olmadığı durumda, ortak maddeler TB-DMF'liyken eşitleme değişmezliğindeki en büyük bozulma ortalama REMSD hatası en büyük olan O-S yönteminde elde edilmiştir. O-S yöntemi diğer tüm koşullarda da en büyük REMSD hatasını vermiştir. Karakteristik eğri yöntemleri (HB, S-L) ise en küçük REMSD değerlerini üretmiş, en küçük hata S-L yönteminde gözlenmiştir. Huggins'in (2014) araştırmasında bu araştırmadan farklı olarak tüm ortak maddeler DMF'li değildir. Ayrıca iki araştırmada bağımlı değişkenler ve buna bağlı eşitlemenin değerlendirme kriterleri farklıdır. Bu araştırmada DMF'li ortak maddelerin test eşitlemeye etkisi RMSD eşitleme hataları bakılırken, Huggins (2014) DMF'nin eşitleme değişmezliğine etkisine bakmaktadır. Dolayısıyla iki araştırmada O-S yöntemi için elde edilen farklılığın iki araştırmada bağımlı değişkenlerin farklı olmasından kaynaklandığı söylenebilir.

4.2. Ortak Maddelerin DMF'siz Olduğu Durumda Uygulanan MTK'ya Dayalı Test Eşitleme Yöntemlerinden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular

O-O, O-S, HB ve S-L yöntemlerine göre ortak maddelerin DMF'siz olduğu durumlarda her bir yöntemden elde edilen RMSD eşitleme hataları Tablo 4.1'de verilmiştir.

Tablo 4.1'e göre A ve B formlarında ortak maddeler DMF'siz olduğunda ortalama-ortalama yöntemine dayalı eşitleme hatası 0.202, ortalama-sigma yöntemine dayalı eşitleme hatası 0.248, Haebara'nın karakteristik eğrisi dönüştürme yöntemine dayalı eşitleme hatası 0.181 ve Stocking-Lord'un karakteristik eğrisi dönüştürme yöntemine dayalı eşitleme hatası 0.181 bulunmuştur.

Aşağıda ortak maddelerin DMF'siz olduğu durumda dört farklı eşitleme yöntemine göre elde edilen RMSD eşitleme hataları Şekil 4.2'de grafiksel olarak görülmektedir.



Şekil 4.2: Ortak Maddeler DMF'siz Olduğunda Eşitleme Hataları (RMSD)

Şekil 4.2'de görüldüğü gibi ortak maddeler DMF'siz olduğunda en büyük RMSD eşitleme hatası ortalama-sigma yönteminde elde edilirken, en küçük hata karakteristik eğrisi yöntemlerinde (Haebara ve S-L) eşit değerlerde elde edilmiştir. Bununla beraber karakteristik eğrisi yöntemlerinin moment yöntemlerine göre daha doğru ve kararlı sonuçlar ürettiği gözlenmektedir. Nitekim literatürde bu bulguyu destekleyen pek çok çalışma bulunmaktadır (Baker ve Al-Karni, 1991; Hanson ve Béguin, 2002; Kim ve Cohen, 1992; Ogasawara, 2001a, 2001b; Kim ve Lee, 2004; Kim ve Lee, 2006; Karkee ve Wright, 2004; Kolen ve Brennan, 2004; Kim ve Kolen, 2006, Kilmen, 2010).

S-L ve HA yöntemleri, ortak maddelerin DMF'siz ele alındığı düşünülen diğer araştırmaların bulgularına zıtlık oluşturmayacak şekilde birbirine eşit eşitleme

hataları üretmiştir. Karakteristik eğrisi eşitleme yöntemlerinden birinin diğerine açık ve önemli bir üstünlüğü bulunmamaktadır (Karkee ve Wright, 2004:12) ve ikisinin de kullanımı önerilmektedir. Yapılan araştırmaların bazılarında Stocking-Lord yöntemi daha doğru sonuçlar üretirken (Gök, 2012; Hu ve diğ., 2008), bazılarında ise Haebara yöntemi daha az hata üreten eşitleme yöntemi olarak belirtilmiştir (Kim ve Colen, 2006). Bununla birlikte literatürde, iki yöntemin birbirine çok yakın ve hatta bazı koşullarda birbiriyle eşit hatalar ürettiği bulgusunu destekleyen araştırmalar bulunmaktadır (Way ve Tang, 1991; Hanson ve Béguin, 2002; Lee ve Ban 2010; Uysal, 2014).

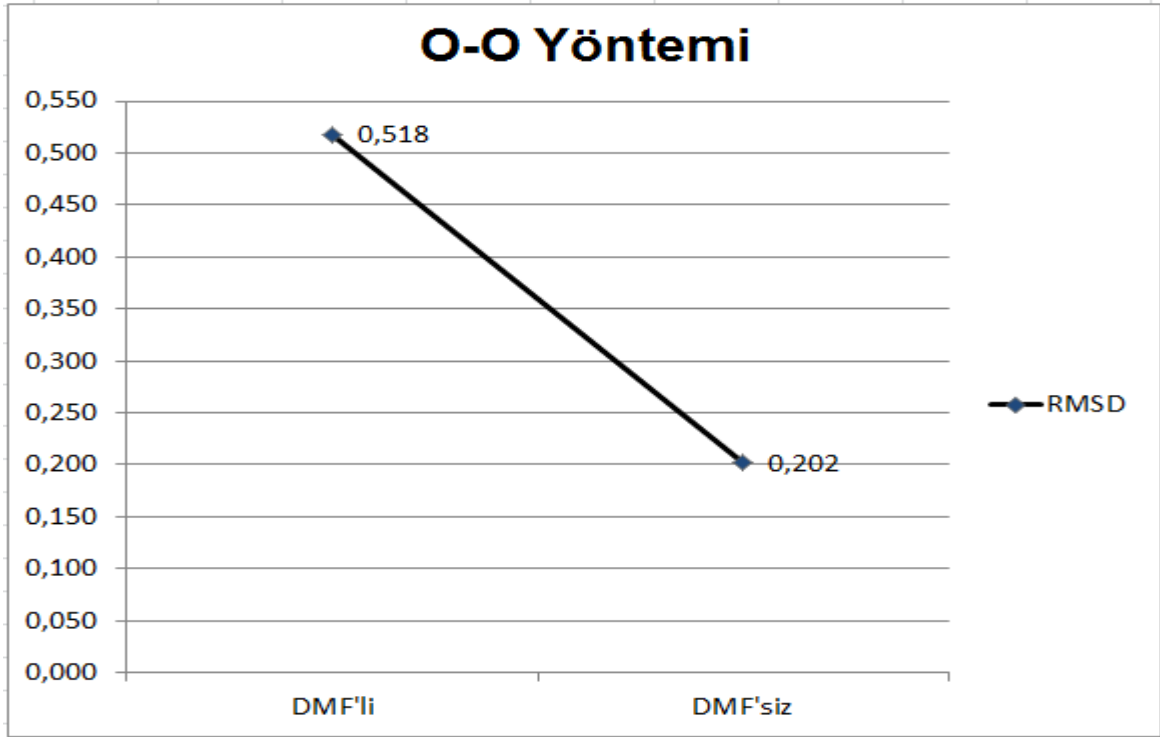
Araştırmada lineer eşitleme yöntemleri (O-O ve O-S) de birbiriyle oldukça yakın eşitleme hataları üretmiştir. Holland ve Dorans'a (2006:201) göre iki örnekteki bireyler arasında çok ufak farklılıklar olduğunda, tüm lineer ölçekleme ve eşitleme yöntemleri benzer sonuçlar verme eğiliminde olmaktadır (aynı durum lineer olmayan ölçekleme ve eşitleme yöntemlerinde de geçerlidir). Bu bakımdan eşdeğer gruplarda eşitleme ile elde edilen sonuç literatürle uyumlu görülmektedir.

Ortak maddeler DMF'siz olduğunda O-S yöntemi en büyük hatayı üretmiştir. Literatürde bu bulguyu destekleyen araştırmalara rastlanmaktadır (Speron, 2009; Kilmen, 2010, Gök, 2012). Moment yöntemleri karşılaştırıldığında ortak maddeler DMF'siz olduğunda O-O yönteminin daha doğru eşitleme yaptığı ve O-S yöntemine göre daha tercih edilebilir görünmektedir. Bunun nedeni olarak ortalamaların standart sapmaya göre daha az hassas olması gösterilebilir. Araştırma bulgusunu destekler nitelikte Baker ve Al-Karni (1991) ortalamaların genellikle standart sapmalardan daha kararlı değerler üretmeleri nedeniyle ortalama-ortalama yöntemini ortalama-sigma yöntemine göre daha tercih edilebilir bulmaktadır. Ortalama ortalama yöntemi a ve b'nin sadece ortalamalarını kullandığı için parametrelerin dağılımsal özelliklerine karşı, güçlük parametresinin standart sapmasını kullanan yöntemlere göre daha az hassas bir yöntemdir. Bununla birlikte O-O ve O-S yöntemini karşılaştıran ampirik araştırma yeterli ve ikna edici değildir; bu nedenle burada önerilen yaklaşım iki prosedürü de dikkate almak ve eşitleme gerçekleştiğinde iki yöntemin uygulanmasından kaynaklanan ham puandan ölçek puanı elde etme değişimini karşılaştırmaktır (Kolen ve Brennan, 2004:167).

4.3. Ortak Maddelerin DMF'li veya DMF'siz Olduğu Durumlarda A ve B Formlarının Ortalama-Ortalama Yönteminden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular

Tablo 4.1'e göre A ve B formlarında ortak maddeler TB-DMF'li olduğunda ortalama-ortalama yöntemine dayalı RMSD eşitleme hatası 0.518, ortak maddeler DMF'siz olduğunda ortalama-ortalama yöntemine dayalı RMSD eşitleme hatası 0.202 olarak bulunmuştur. O-O yöntemine göre elde edilen RMSD sonuçlarına ait grafiksel gösterim aşağıda Şekil 4.3'te verilmiştir.

Bu bölümde, alt problem sırasına göre verilmiş araştırma bulguları ve bu bulgularla ilgili değerlendirmeler yer almaktadır.



Şekil 4.3: Ortak Maddeler DMF'li veya DMF'siz Olduğunda O-O Yöntemine Göre Eşitleme Hataları (RMSD)

Şekil 4.3'e göre ortalama-ortalama yönteminde A ve B formlarının ortak maddeleri TB-DMF'li olduğunda eşitleme hatası RMSD, ortak maddelerin DMF'siz olduğu duruma göre 0.316 farkla daha yüksek bulunmuştur. Ortalama-ortalama yönteminde denklem eğimi (A) ayırıcılık parametrelerinin ortalaması ile hesaplanmaktadır. Daha sonra, ortak maddelerden kestirilen b parametrelerinin ortalaması ile B katsayısı elde edilmektedir. Literatürde DMF'nin eşitleme katsayılarını veya yetenek kestirimlerini etkilediğini belirten araştırmalar

bulunmaktadır (Han, 2008; Wells ve diğ., 2012). Bu sebeple aşağıda hesaplamaları verilmiş A ve B eşitleme katsayılarından hareketle O-O yönteminden elde edilen eşitleme hataları tartışılmıştır.

A ve B formlarında **DMF’li** ortak maddelerin a parametreleri ortalamaları oranlanarak elde edilen A eşitleme katsayısı ve b parametreleri ortalamalarıyla hesaplanan B eşitleme katsayısı formül değerleri yerine konulduğunda aşağıdaki gibidir:

$$\begin{array}{l}
 A = \mu(a_B) / \mu(a_A) \text{ ve,} \\
 \mu(a_B) = 0.839 \\
 \mu(a_A) = 0.529
 \end{array}
 \left. \vphantom{\begin{array}{l} A = \mu(a_B) / \mu(a_A) \text{ ve,} \\ \mu(a_B) = 0.839 \\ \mu(a_A) = 0.529 \end{array}} \right\} \mathbf{A = 1.586}
 \qquad
 \begin{array}{l}
 B = \mu(b_A) - A \cdot \mu(b_B) \text{ ve,} \\
 \mu(b_A) = 0.164 \\
 \mu(b_B) = 0.108
 \end{array}
 \left. \vphantom{\begin{array}{l} B = \mu(b_A) - A \cdot \mu(b_B) \text{ ve,} \\ \mu(b_A) = 0.164 \\ \mu(b_B) = 0.108 \end{array}} \right\} \begin{array}{l} B = 0.164 - 1.586(0.108) \\ \mathbf{B = -0.007 \approx -0.01} \end{array}$$

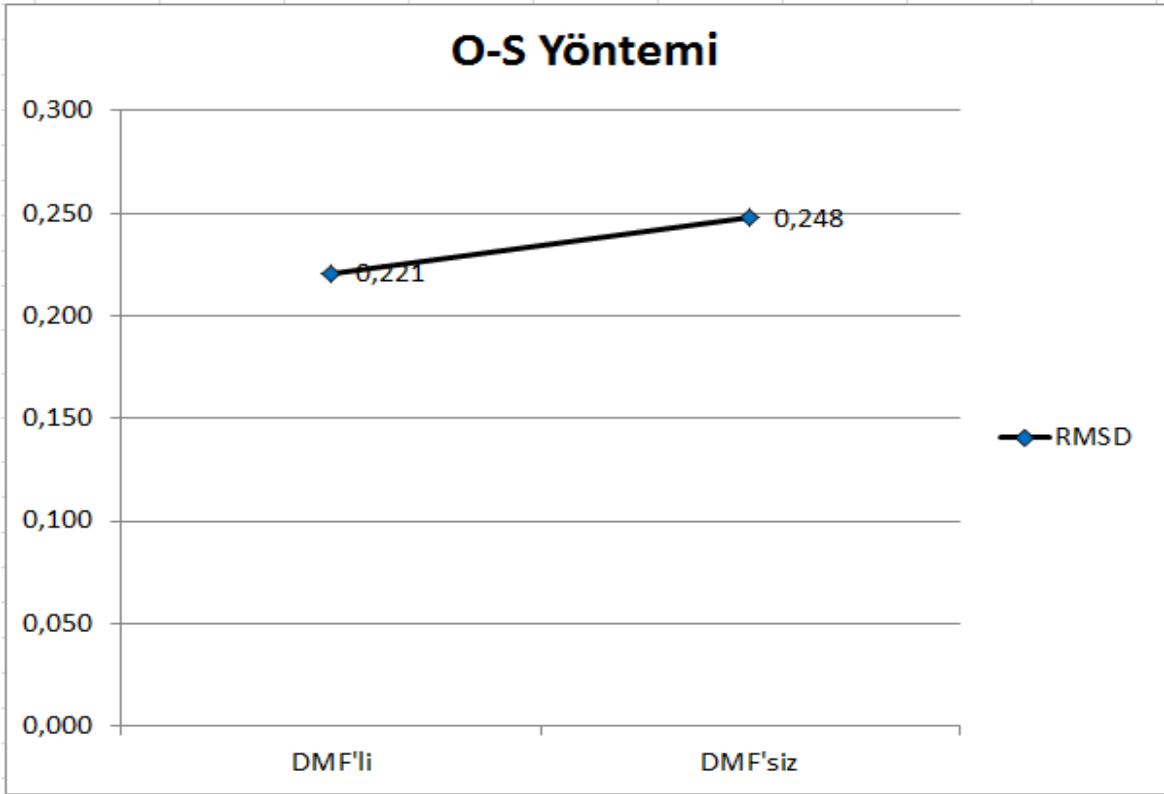
A ve B formlarında **DMF’siz** ortak maddelerden hesaplanan A ve B eşitleme katsayıları ise aşağıdaki gibidir:

$$\begin{array}{l}
 A = \mu(a_B) / \mu(a_A) \text{ ve,} \\
 \mu(a_B) = 0.905 \\
 \mu(a_A) = 1.132
 \end{array}
 \left. \vphantom{\begin{array}{l} A = \mu(a_B) / \mu(a_A) \text{ ve,} \\ \mu(a_B) = 0.905 \\ \mu(a_A) = 1.132 \end{array}} \right\} \mathbf{A = 0.80}
 \qquad
 \begin{array}{l}
 B = \mu(b_A) - A \cdot \mu(b_B) \text{ ve,} \\
 \mu(b_A) = -0.705 \\
 \mu(b_B) = -0.756
 \end{array}
 \left. \vphantom{\begin{array}{l} B = \mu(b_A) - A \cdot \mu(b_B) \text{ ve,} \\ \mu(b_A) = -0.705 \\ \mu(b_B) = -0.756 \end{array}} \right\} \begin{array}{l} B = -0.705 - 0.80 \cdot (-0.756) \\ \mathbf{B = -0.10} \end{array}$$

Ortak maddeler DMF’li veya DMF’siz olduğunda yukarıda hesaplanan A ve B katsayıları incelendiğinde, B katsayılarının (-0.01 ve -0.10) birbirine oldukça yakın oldukları göze çarpmaktadır. A katsayıları karşılaştırıldığında ise DMF’li ortak maddelerle eşitlemede A katsayısı daha büyük bulunmuştur. Bunun nedeninin, a parametreleri ortalamasının kestirilen A formunda keskin şekilde azalması olduğu söylenebilir. Bir başka ifadeyle DMF’li ortak maddelerle eşitlemede A eşitleme katsayısının büyümesi (A.θ+B) ile hesaplanan yeteneğin büyük çıkmasına neden olmaktadır. Bu durumda DMF’li formda RMSD eşitleme hatası da artmaktadır. Ortak maddeler DMF’siz olduğunda a ve b parametreleri ortalamalarının iki formda daha benzer değerlere sahip olmalarının eşitleme hatasını küçülttüğü düşünülmektedir.

4.4. Ortak Maddelerin DMF'li veya DMF'siz Olduğu Durumlarda A ve B Formlarının Ortalama-Sigma Yönteminden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular

Tablo 4.1'e göre eşitleme yapılan A ve B formlarında ortak maddeler TB-DMF'li olduğunda ortalama-sigma yöntemine dayalı RMSD eşitleme hatası 0.221 iken ortak maddeler DMF'siz olduğunda ortalama-sigma yöntemine dayalı RMSD eşitleme hatası 0.248 olarak bulunmuştur. O-S yöntemine göre elde edilen RMSD sonuçlarına ait grafiksel gösterim aşağıda Şekil 4.4.'te verilmiştir.



Şekil 4.4: Ortak Maddeler DMF'li veya DMF'siz Olduğunda O-S Yöntemine Göre Eşitleme Hataları (RMSD)

Şekil 4.4'e göre ortalama-sigma yönteminde A ve B formlarının ortak maddeleri DMF'siz olduğunda eşitleme hatası RMSD, ortak maddelerin TB-DMF'li olduğu duruma göre 0.027 farkla daha yüksek bulunmuştur.

O-S yöntemi ortak maddelerden kestirilen b (güçlük) parametrelerinin ortalama ve standart sapmalarına dayalı bir yöntemdir. A ve B formlarında **DMF'li** ortak maddelerin b parametreleri standart sapmaları oranlanarak elde edilen A eşitleme katsayısı ve b parametreleri ortalamalarıyla hesaplanan B eşitleme katsayısı formül değerleri yerine konulduğunda aşağıdaki gibi elde edilmiştir:

$A = \sigma(b_A) / \sigma(b_B)$ ve

$$\left. \begin{array}{l} \sigma(b_A) = 1.758 \\ \sigma(b_B) = 1.406 \end{array} \right\} \mathbf{A = 1.250}$$

$B = \mu(b_A) - A \cdot \mu(b_B)$ ve,

$$\left. \begin{array}{l} \mu(b_A) = 0.164 \\ \mu(b_B) = 0.108 \end{array} \right\} \begin{array}{l} B = 0.164 - 1.250 \cdot (0.108) \\ \mathbf{B = 0.029 = 0.03} \end{array}$$

A ve B formlarında **DMF'siz** ortak maddelerin b parametreleri standart sapmaları oranlanarak elde edilen A eşitleme katsayısı ve b parametreleri ortalamalarıyla hesaplanan B eşitleme katsayısı formül değerleri yerine konulduğunda aşağıdaki gibi elde edilmiştir:

$$\left. \begin{array}{l} \sigma(b_A) = 0.580 \\ \sigma(b_B) = 0.761 \end{array} \right\} \mathbf{A = 0.760} \quad \left. \begin{array}{l} \mu(b_A) = -0.705 \\ \mu(b_B) = -0.756 \end{array} \right\} \begin{array}{l} B = (-0.705) - 0.760(-0.756) \\ \mathbf{B = -0.13} \end{array}$$

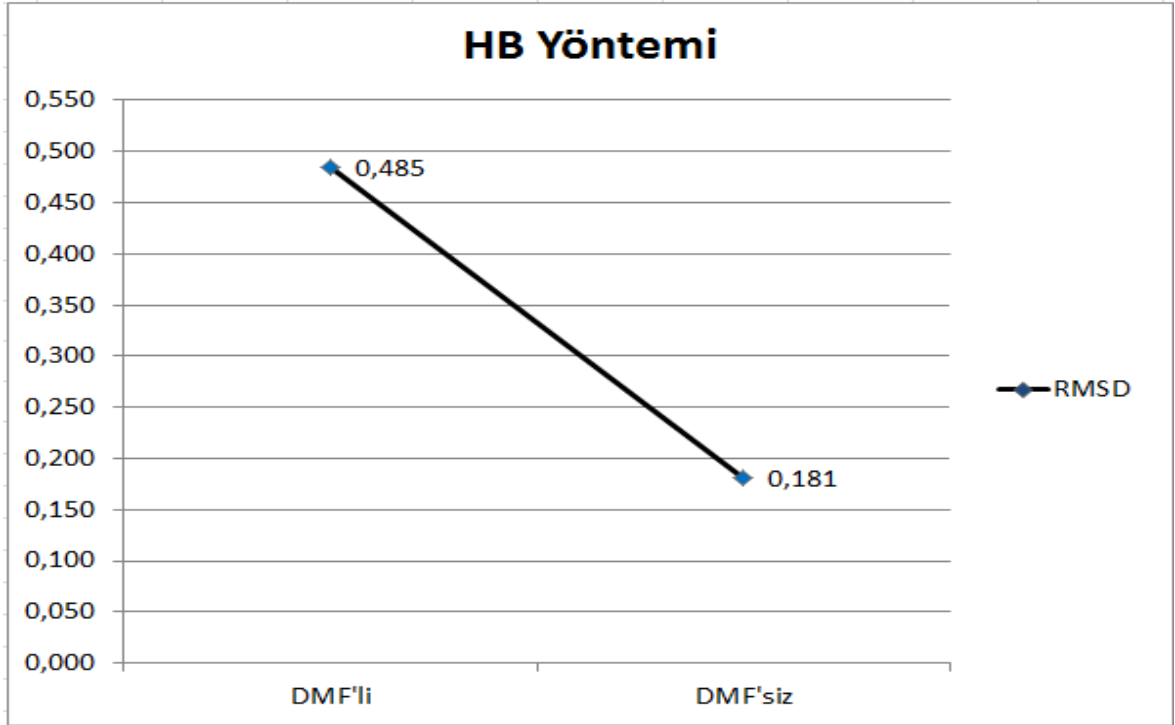
Sonuç olarak A eşitleme katsayısı ortak maddeler DMF'siz olduğunda ortak maddelerin DMF'li olduğu duruma göre azalmış; çünkü Form B'deki b parametreleri varyansı büyümüştür. Ortak maddeler DMF'li olduğunda ise Form B'deki b parametreleri varyansı küçülmüş ve A katsayısı artmıştır. B eşitleme katsayısı ise ortak maddeler DMF'siz olduğunda ortak maddelerin DMF'li olduğu duruma göre azalmıştır. DMF'siz ortak maddeli formda A ve B katsayılarının ayrı ayrı veya birlikte düşüş göstermeleri kestirilen yeteneğin (θ^*) azalmasına böylece gerçek ve kestirilen yetenekler arası farkın büyümesine neden olmaktadır. Bu durumda gerçek ve kestirilen yetenekler arası farka dayalı hesaplanan RMSD eşitleme hatası da DMF'siz ortak maddeli formda büyümektedir. O-S yönteminin DMF'siz ortak maddelerle eşitlemede daha büyük hata vermesi temel form B'deki DMF'siz ortak madde b parametreleri varyans artışına ya da form A'da kestirilen b parametreleri varyans düşmesine bağlanmıştır. Benzer şekilde Wyse ve Reckase (2012) iki karakteristik eğrisi (S-L ve HB), iki moment (O-O ve O-S) ve eş zaman eşitleme yöntemlerini karşılaştırdıkları araştırmalarında O-S yönteminin zayıf performans göstermesine neden olarak bir ortak maddenin güçlüğünde yaşanan ciddi değişimin temel formun güçlük parametreleri varyansını büyük çapta artırmasına bağlamışlardır. Aynı araştırmada problemlı maddenin testten çıkarılmasının güçlük farkını azaltarak eşitlemeyi iyileştirdiği belirtilmiştir.

Araştırmada elde edilen bulgu Han'ın (2008) araştırma bulgularıyla A katsayısı için farklılık B katsayısı için benzerlik göstermektedir. Han (2008) araştırmasında O-S yönteminde TB-DMF'nin büyüklüğü ve TB-DMF'li madde oranı arttığında (40 maddeli testte 10 ortak maddenin 5'i= %50 DMF'li iken) A eşitleme katsayısının azaldığını belirtmiş; buna neden olarak DMF nedeniyle Form 2'de ortak madde b parametreleri varyansının artmasını göstermiştir. Benzer şekilde B katsayısı TB-DMF'li madde sayısı ve TB-DMF büyüklüğü artınca, Form 2'de b parametre ortalamasının azalmasına bağlı olarak büyümüştür. Bu araştırmada da Form B'de b parametreleri ortalaması düştüğü için B katsayısı büyümüştür. Araştırmalarda A katsayılarında yaşanan farklılık; iki araştırmanın birebir aynı koşullara sahip olmamasına bağlanabilir. Örneğin Han'ın araştırmasında IPD'li ortak maddelerin hepsi Form 2'yi alanlar için Form 1'i alanlardan daha kolaydır; b parametrelerinin tamamı iki formda farklı işarettedir. Bu araştırmada ortak maddelerin b parametreleri işaretleri iki formda karışıktır (\pm işaretli). Yine iki araştırmada DMF'li ortak madde oranlarında farklılıklar bulunmaktadır. Bu araştırmada tüm ortak maddeler DMF'lidir. Han'ın araştırmasında ise ortak maddelerin %50'si DMF'liyken bulunan A katsayısı, %10'u DMF'liyken bulunan A katsayısından daha küçük elde edilmiştir. Yani araştırmada DMF'li madde oranı arttıkça ortak maddelerin güçlük parametre varyansları ve A katsayısı artmıştır. Bu durum O-S yönteminde eşitleme hatasının büyümesine neden olmuştur. Araştırmada ayrıca bu araştırmadan farklı olarak uzun test (40 maddeli) ve büyük örneklem (5000) kullanılmıştır. İki araştırmada elde edilen farklılığa Han'ın (2008) araştırmasındaki farklı simülasyon koşulları neden olduğu düşünülmektedir.

Araştırmada elde edilen çelişkili bulguya DMF'siz ortak madde b parametreleri ortalama ve standart sapmasının tüm test ile istatistiksel benzerliğinin etkisi olabileceği düşünülmüş ve bu durum hesaplanarak test edilmiştir. Tüm testin b parametreleri ortalama ve standart sapma değerleri ortak maddelerin ortalama ve standart sapması yerine konulmuştur. Ancak sonuçta DMF'siz ortak maddelerle eşitlemede RMSD eşitleme hatası değişmemiş; hatta bir miktar daha büyümüştür. Elde edilen sonuç Cook ve Paterson (1987) ile Hannick ve Huang'ın (2002) araştırmalarında belirtildiği gibi ortak maddelerin testin tümüyle birebir istatistiksel benzerliğinin eşitleme sonuçlarını etkilemediği ile tutarlıdır.

4.5. Ortak Maddelerin DMF'li veya DMF'siz Olduğu Durumlarda A ve B Formlarının Haebara'nın Karakteristik Eğrisi Dönüştürme Yönteminden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular

Tablo 4.1'e göre A ve B formlarında ortak maddeler TB-DMF'li olduğunda Haebara'nın karakteristik eğrisi dönüştürme yöntemine dayalı RMSD eşitleme hatası 0.485 bulunurken ortak maddeler DMF'siz olduğunda Haebara'nın karakteristik eğrisi dönüştürme yöntemine dayalı RMSD eşitleme hatası 0.181 olarak hesaplanmıştır. HB yönteminden elde edilen RMSD sonuçlarına ait grafiksel gösterim aşağıda Şekil 4.5'te verilmiştir.



Şekil 4.5: Ortak Maddeler DMF'li veya DMF'siz Olduğunda Haebara Yöntemine Göre Eşitleme Hataları (RMSD)

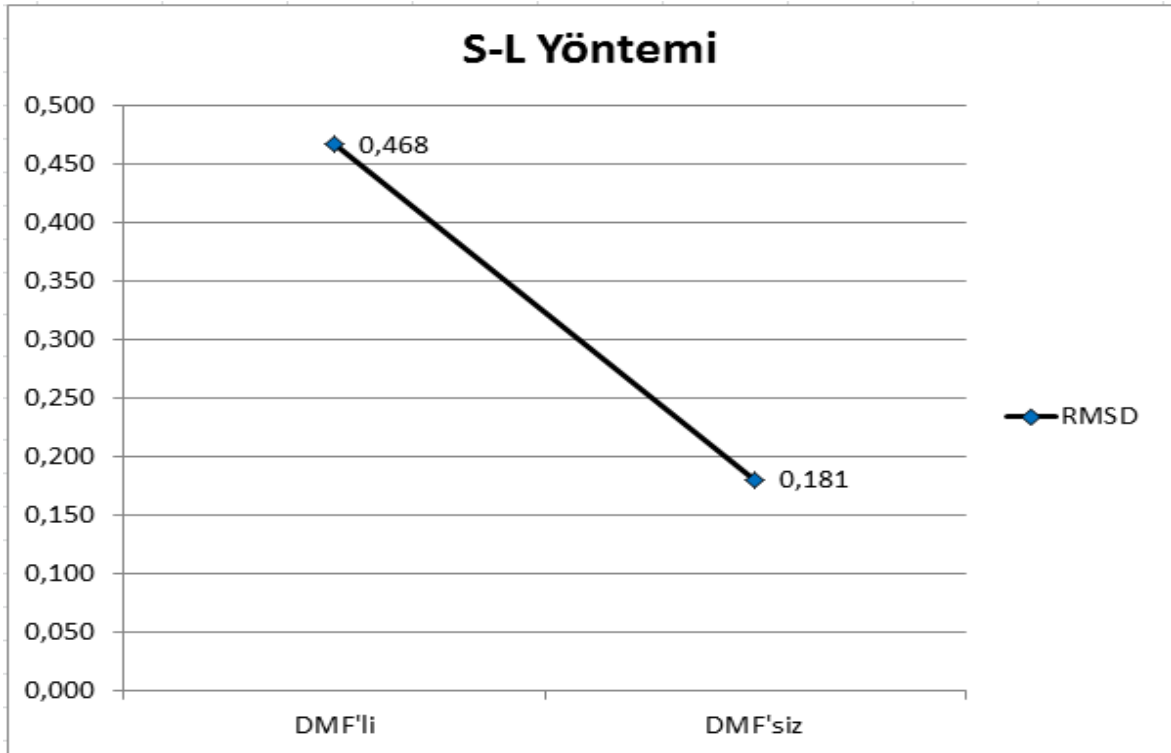
Şekil 4.5 incelendiğinde ortak maddeler TB-DMF'li olduğu durumda HB yönteminden elde edilen RMSD eşitleme hatası ortak maddelerin DMF'siz olduğu duruma kıyasla 0.304 farkla daha büyük bulunmuştur. Haebara yöntemi eşitleme yapılan A ve B formlarında her bir maddenin madde karakteristik eğrileri arası farka dayanmaktadır. Belli bir yetenek düzeyindeki bireyler için her bir maddenin madde karakteristik eğrileri arasındaki farkın karelerinin toplamı alınmaktadır (Kolen ve Brennan, 2004). Dolayısıyla Haebara yöntemiyle eşitlemeden elde edilen bu bulgu eşitleme yapılan A ve B formlarının karakteristik eğrileri arasındaki farklılık bulunması ile ilişkilendirilmektedir. Ortak maddelerin DMF'li olduğu

durumda A ve B formlarının karakteristik eğrilerinin arasında ortak maddelerin DMF'siz olduğu duruma göre daha büyük fark olduğu düşünülmektedir.

Kaskowitz ve De Ayala (2001) araştırmalarında birbirine daha uzak test karakteristik eğrilerinde daha büyük hata ortaya çıktığını ve ortaya çıkan hatanın eşitleme katsayılarının doğruluğunu azaltıcı etki yaptığını belirtmişlerdir.

4.6. Ortak Maddelerin DMF'li veya DMF'siz Olduğu Durumlarda A ve B Formlarının Stocking-Lord'un Karakteristik Eğrisi Dönüştürme Yönteminden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular

Tablo 4.1'e göre eşitleme yapılan A ve B formlarında ortak maddeler TB-DMF'li olduğunda Stocking-Lord'un karakteristik eğrisi dönüştürme yöntemine dayalı RMSD eşitleme hatası 0.468, ortak maddeler DMF'siz olduğunda Stocking-Lord'un karakteristik eğrisi dönüştürme yöntemine dayalı RMSD eşitleme hatası 0.181 olarak bulunmuştur. Stocking-Lord'un karakteristik eğrisi dönüştürme yöntemine göre elde edilen RMSD sonuçlarına ait grafiksel gösterim aşağıda Şekil 4.6'da verilmiştir.



Şekil 4.6: Ortak Maddeler DMF'li veya DMF'siz Olduğunda S-L Yöntemine Göre Eşitleme Hataları (RMSD)

Şekil 4.6 incelendiğinde ortak maddeler TB-DMF'li olduğu durumda S-L yönteminden elde edilen RMSD eşitleme hatası, ortak maddelerin DMF'siz olduğu

duruma kıyasla daha büyük bulunmuştur. S-L yönteminde elde edilen bu bulgu diğer bir karakteristik eğrisi dönüştürme yöntemi HB ile de tutarlılık göstermektedir. Bu bulgu HB yönteminde olduğu gibi eşitleme yapılan A ve B formlarının karakteristik eğrileri arasındaki farklılık bulunması ile ilişkilendirilmektedir (Kaskowitz ve De Ayala (2001)).

Chen'in (2013) araştırma bulguları bu çalışmada elde edilen bulguyu destekler niteliktedir. Chen (2013) simülatif veri ile yürüttüğü araştırmasında eşdeğer gruplarda S-L yöntemiyle eşitlemede ortak maddelerin a ve b parametrelerinde ayrı ayrı kayma (IPD) yaşanması durumlarını farklı koşullarda incelemiştir. Araştırmada DMF'nin bir çeşidi olarak ele alınan (Han, 2008) IPD'li maddelerin silinmesi veya testten çıkarılmasının eşitleme yanlılığı (BIAS) ve RMSE eşitleme hatalarına yaptığı etkiyi karşılaştırılmıştır. Sonuç olarak, IPD'li/DMF'li ortak maddelerin testte olması durumunda IPD'li madde sayısı (3, %10-8, %25) ve IPD düzeyine (a parametrelerinde ± 0.4 , b parametrelerinde ± 0.2 ve ± 0.4 kayma) bakmaksızın eşitleme yanlılığı ve RMSE eşitleme hataları yüksek bulunmuştur. IPD'li maddelerin silinmesi durumunda yanlılık ve hata değerleri düşmüştür.

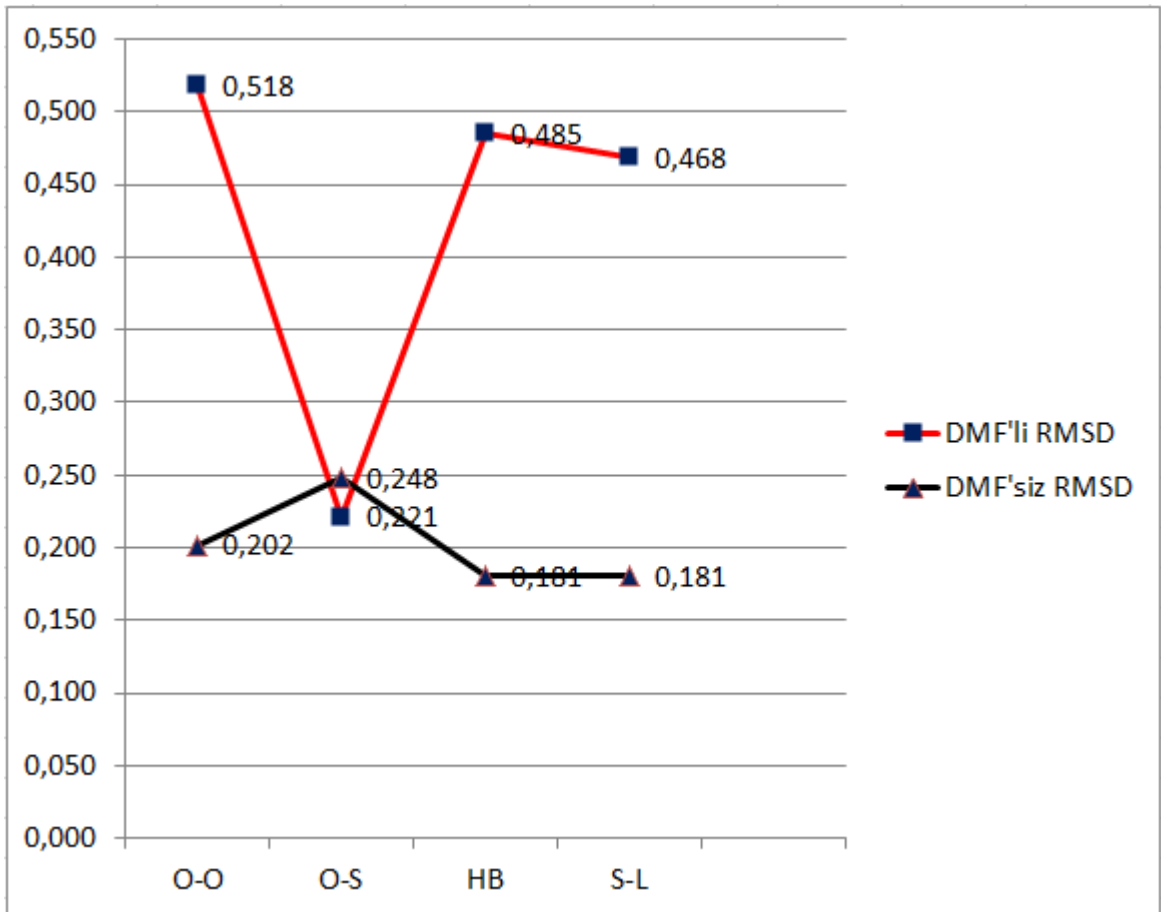
Wells ve diğ.'nin (2002) yapmış oldukları araştırmaya bulguları da elde edilen sonuçla paralellik göstermektedir. Wells ve diğ. (2002) S-L eşitleme yöntemini kullanarak farklı koşullarda a ve b parametrelerindeki kaymanın yetenek kestirimlerindeki RMSD eşitleme hatasına etkisini incelemişlerdir. Araştırma sonuçlarında S-L yönteminde a ve b parametrelerinde ayrı ayrı ve birlikte kayma olduğunda yetenek kestirimleri arasında farklılığa dayanan RMSD değerlerinin arttığı sonucuna ulaşmıştır.

Atalay Kabasakal (2014) DMF'li maddelerin testten çıkarılması durumunda S-L yöntemine göre yetenek kestirimlerinde elde edilen hatanın 24 farklı simülasyon koşulunda genel olarak arttığını bulmuştur. Bu artış DMF bulunan test ve DMF etki büyüklüğüne göre değişmektedir. Ancak bu araştırmanın koşullarıyla benzer simülasyon koşulu kıyaslanırsa, DMF'li maddeler ortak testte olduğunda, küçük örnekleme (500) ve kısa testlerde (20) bazı simülasyonlarda yetenek kestirimine ait DMF'li ortak maddelerin silinmesi BIAS ve RMSE değerini düşürmüştür. SE ise artmıştır. Madde parametrelerine dayalı hatalarda da DMF'li ortak maddelerin testten çıkarılması genel olarak SE'yi yükseltmiş; BIAS ve RMSE'yi düşürmüştür. Araştırmada kullanılan her bir simülasyon koşulunun bir uygulama olduğu

düşünüldüğünde elde edilen sonuç bu araştırma ile tamamen tutarsız değildir. Bununla birlikte iki araştırmada kullanılan DMF'li madde oranları da farklılık göstermektedir. Bu araştırmada ortak maddeler tüm testin %25'i ve tamamı DMF'li iken, diğer araştırmada tüm testin %25'i olan ortak maddelerin %10'u DMF'li seçilmiştir. Özet olarak iki araştırmanın koşulları benzer seçildiğinde tutarlılık gözlemlendiği söylenebilir.

4.7. Ortak Maddelerin DMF'li veya DMF'siz Olduğu Durumlarda A ve B Formlarının Ortalama-Ortalama, Ortalama, Haebara ve Stocking-Lord Yöntemlerinden Elde Edilen Eşitleme Hatalarına İlişkin Bulgular

Tablo 4.1'e göre A ve B formlarının ortak maddeleri TB-DMF'li ve DMF'siz olduğu durumlarda 4 farklı eşitleme yöntemine göre elde edilen eşitleme hataları (RMSD) değerleri Şekil 4.7'de gösterilmektedir.



Şekil 4.7: Ortak Maddeler DMF'li veya DMF'siz Olduğunda O-O, O-S, Haebara ve S-L Yöntemlerine Göre Eşitleme Hataları (RMSD)

Şekil 4.7 incelendiğinde ortak maddeler TB-DMF'li olduğunda O-O, HB ve S-L yöntemlerinde RMSD eşitleme hataları ortak maddelerin DMF'siz olduğu duruma göre büyük çıkma eğilimindedir. Araştırmada O-S yöntemi için çelişkili bir sonuç elde edilmiştir. Ortak maddeler DMF'siz olduğunda O-S yönteminde elde edilen eşitleme hatası az bir farkla (0.027) büyümektedir. Bu durumun nedeni O-S yönteminin b parametreleri standart sapmalarındaki değişimden etkilenmesi olarak görülmektedir. Temel form B'deki DMF'siz ortak madde b parametreleri standart sapmasının artışı ya da form A'da kestirilen b parametreleri standart sapmasının düşmesi hatanın büyümesine neden olmuştur.

Literatürde O-S yönteminde b parametrelerinin standart sapmasını kullanmanın eşitleme sonuçlarında daha az kararlı sonuçlar ürettiği belirtilmektedir. Baker ve Al-Karni (1991)'nin simülatif veri ile yürüttükleri araştırmada elde edilen bulgular ortak maddelerin DMF'siz olduğu durumda elde edilen araştırma bulguları ile paralellik göstermektedir. Söz konusu araştırmada Stocking-Lord yöntemi ortalama-ortalama yöntemine göre daha az hata üretmiştir. Bütüne bakıldığında, genel olarak pek çok amaç için ortalama-ortalama yönteminin daha kabul edilebilir eşitleme katsayıları verdiği belirtilmiştir. Stocking- Lord yönteminin ise, test karakteristiklerine karşı daha az duyarlı olduğu sonucu ortaya konulmuştur.

Araştırmada O-S yönteminde ortak maddelerin DMF'li ve DMF'siz olduğu durumlarda eşitleme hataları arası fark 0.027 ile oldukça küçük bulunmuştur. Bu küçük farklılık a ve b parametrelerindeki değişim ile açıklanmıştır. Araştırma bulgularına benzer şekilde Wells ve diğ. (2002) simülasyon veri ile eşitleme yaptıkları araştırmalarında a ve b parametrelerinde ayrı ayrı veya birlikte kayma (Item Parameter Drift, IPD) olması durumunun yetenek kestirimlerinde büyük bir etki yaratmadığı sonucunu elde etmişlerdir.

DMF'li maddelerde a ve b parametrelerinde kayma (Item Parameter Drift, IPD) beklendiği bir gerçektir ve bu durum kontrollü simülasyon koşullarında sağlanabilmektedir (Wells ve diğ., 2002). Ancak sınırlı gerçek veri ile çalışıldığı durumlarda tesadüfi hatadan kaynaklı sonuçlar elde edilebilir. Aynı durum simülasyon çalışmalarında da geçerlidir. Turhan (2006) simülasyon veri ile yürüttüğü DMF varlığında dikey eşitleme araştırmasında çelişkili sonuç elde etmiştir. TB-DMF'li ortak madde sayısı arttığında ölçekteki bozulmanın azaldığı sonucuna ulaşmış; bu durumu sınırlı simülasyon koşullarına bağlamıştır.

5. SONUÇ ve ÖNERİLER

Bu bölümde araştırmadan elde edilen sonuçlara ve bu sonuçlara bağlı olarak araştırma sonuçlarına ve diğer araştırmacılara dönük önerilere yer verilmiştir.

5.1. Sonuçlar

Bu araştırmada, ortak maddeler cinsiyete göre TB-DMF'li veya DMF'siz olduğunda Madde Tepki Kuramı'na dayalı "ortalama-ortalama", "ortalama-sigma", "Haebara" ve "Stocking-Lord" eşitleme yöntemlerinden hesaplanan RMSD eşitleme hatalarının karşılaştırılması amaçlanmıştır. Araştırma gerçek veri üzerinden yürütülmüştür. Dolayısıyla elde edilen sonuçların eşitleme yöntemlerinin performansıyla ilgili gerçeğe uygun sonuçlar oldukları söylenebilir.

Araştırmadan elde edilen sonuçları aşağıdaki gibi özetlemek mümkündür:

- Ortak maddeler erkekler lehine TB-DMF'li olduğu durumda en küçük RMSD eşitleme hatasını ortalama-sigma yöntemi üretmiştir. Ortak maddeler erkekler lehine TB-DMF'li olduğu durumda en büyük hata ise ortalama-ortalama yönteminde elde edilmiştir.
- Ortak maddeler DMF'siz olduğu durumda en küçük RMSD eşitleme hatası karakteristik eğrisi yöntemlerinde (Stocking-Lord ve Haebara) elde edilmiştir. Karakteristik eğrisi yöntemlerinde hatalar eşit bulunmuştur. Ortak maddeler DMF'siz olduğunda en büyük hata ise, ortalama-sigma yönteminde elde edilmiştir.
- Moment yöntemlerinden ortalama-ortalama yöntemi ortak maddeler erkekler lehine TB-DMF'li olduğu durumda ortak maddelerin DMF'siz olduğu duruma kıyasla daha büyük RMSD eşitleme hatasına sahip bulunmuştur.
- Moment yöntemlerinden ortalama-sigma yöntemi ortak maddeler erkekler lehine TB-DMF'li olduğu durumda ortak maddelerin DMF'siz olduğu duruma kıyasla daha küçük RMSD eşitleme hatası üretmiştir.
- Karakteristik eğrisi yöntemlerinden Haebara yöntemi, ortak maddelerin erkekler lehine TB-DMF'li olduğu durumda ortak maddelerin DMF'siz

olduđu duruma kıyasla daha büyük RMSD eşitleme hatasına sahip bulunmuştur.

- Karakteristik eğrisi yöntemlerinden Stocking- Lord yöntemi, ortak maddelerin erkekler lehine TB-DMF'li olduđu durumda ortak maddelerin DMF'siz olduđu duruma kıyasla daha büyük RMSD eşitleme hatasına sahip bulunmuştur.
- Ortak maddeler erkekler lehine TB-DMF'li olduđunda RMSD eşitleme hataları ortalama-sigma yöntemi hariç diđer tüm yöntemlerde (O-O, S-L, HA) ortak maddelerin DMF'siz olduđu duruma göre daha büyük elde edilmiştir. Ancak bu farklılık oldukça küçüktür (0.027). Ortak maddeler erkekler lehine TB-DMF'li iken O-S yönteminin en küçük RMSD eşitleme hatasına sahip olması temel formdaki b parametreleri varyansı ile açıklanmıştır.

5.2. Öneriler

5.2.1. Araştırma Sonuçlarına Dönük Öneriler

Yılda birden fazla uygulaması yapılan öğrenci veya personel seçme ve yerleştirme amaçlanan sınavların birbiriyle istatistiksel olarak eşit olması gerekmektedir. Ancak bu sayede sınavlar birbiri yerine kullanılabilir. Eşitleme çalışmaları yapılırken yanlılık potansiyeli taşıyan DMF'li maddeler üzerinde çalışılması eşitleme sürecini olumsuz etkileyecek ve sistematik biçimde belli bir grubun lehine veya aleyhine kararlar alınmasına neden olacaktır. Bu araştırma DMF'li maddelerin test eşitleme sürecine nasıl bir etki yaptığı hakkında test geliştirme konusunda çalışmalar yapan ölçme ve değerlendirme alan uzmanlarına bilgi verebilir.

Aşağıda Tablo 5.1'de araştırmadan elde edilen sonuçlar doğrultusunda verilen öneriler sunulmuştur.

Tablo 5.1: Araştırma Sonuçlarına Dayalı Önerilen Eşitleme Yöntemleri

		<i>Önerilen Eşitleme Yöntemi</i>			
		<i>Moment Yöntemleri</i>		<i>Karakteristik Eğrisi Yöntemleri</i>	
		<i>O-O</i>	<i>O-S</i>	<i>S-L</i>	<i>HB</i>
<i>Ortak Maddelerin Durumu</i>	<i>Tamamı Erkekler Lehine Tek Biçimli DMF'li</i>		√	√	
	<i>Tamamı DMF'siz</i>	√		√	√

Tablo 5.1'e göre ortak maddelere dayalı eşdeğer gruplar deseninde eşitleme yaparken ortak maddelerin tamamı erkekler lehine TB-DMF'li olduğunda;

- moment yöntemlerinde en küçük hatayı veren ortalama-sigma yönteminin kullanılması;
- karakteristik eğrisi yöntemleri uygulanacaksa en küçük hatayı veren Stocking-Lord yönteminin kullanılması önerilebilir.

Ortak maddelere dayalı eşdeğer gruplar deseninde eşitleme yaparken ortak maddelerin tamamı erkekler lehine TB-DMF'li olduğunda tüm yöntemler içinde en küçük eşitleme hatasını üreten ortalama-sigma yöntemi kullanılarak DMF'li maddeler kontrol edilebilir. Böylece büyük eşitleme hatasının önüne geçilebilir ve sınavların sonunda alınan kararların doğruluğu artırılmış olur.

Tablo 5.1'e göre ortak maddelere dayalı eşdeğer gruplar deseninde eşitleme yaparken ortak maddelerin tamamı DMF'siz olduğunda;

- moment yöntemlerinde en küçük hatayı veren ortalama-ortalama yönteminin kullanılması;
- karakteristik eğrisi yöntemleri uygulanacaksa birbiriyle eşit hata üreten Stocking-Lord veya Heabara yönteminden birinin kullanılması önerilebilir.

Ortak maddelere dayalı eşdeğer gruplar deseninde eşitleme yaparken ortak maddelerin tamamı DMF'siz olduğunda tüm yöntemler içinde birbiriyle eşit ve en küçük eşitleme hatasını üreten S-L ve HB yöntemlerinin literatürde birbirine karşı açık bir üstünlüğü bulunmamaktadır. Dolayısıyla hangisinin kullanılacağı araştırmacının tercihinine bırakılmıştır.

5.2.2. Diğer Arařtırmacılara Dönük Öneriler

Test eřitleme ve DMF konusunda gelecek arařtırmacılar için ařağıdaki öneriler sunulmuřtur.

- Bu arařtırmada TB-DMF'li ortak maddeler kullanılmıřtır. Sonraki arařtırmalarda TB ve TBO-DMF'li ortak maddeler ayrı ayrı veya birlikte düşünülerek eřitleme hatalarına bakılabilir.
- Bu arařtırmada ortak maddelerin tamamı ya DMF'li ya da DMF'siz olarak ele alınmıřtır. Bundan sonraki arařtırmalarda aynı anda hem DMF'li hem de DMF'siz ortak maddeler varlıęında eřitleme hataları incelenebilir. DMF'li ortak maddeleri çıkarmanın eřitleme hataları üzerine etkisi bakılabilir. DMF'li ortak maddelerin oranı baęımsız deęiřken olarak ele alınabilir.
- Bu arařtırmada DMF analizi cinsiyet deęiřkeni ile sınırlı tutulmuřtur; bunun dıřında yerleřim yeri, okul türü, kültür, sosyo ekonomik düzey gibi farklı demografik özellikler deęiřken olarak ele alınıp DMF analizleri gerekleřtirilebilir.
- Bu arařtırmada DMF analizleri KTK'ya dayalı yöntemlerden MH ve LR yöntemleriyle belirlenmiřtir. Sonraki arařtırmalarda MTK'ya dayalı DMF belirleme yöntemleri KTK'ya dayalı DMF belirleme yöntemleriyle birlikte kullanılabilir.
- Bu arařtırmada 17'řer maddeden oluřan kısa testlerde eřitleme alıřması yapılmıřtır. Bundan sonraki arařtırmalarda daha uzun testlerde gerek veri ile alıřılabilir. Uzun ve kısa testler birlikte ele alınıp DMF'li ortak maddelerin etkileri karřılařtırılabilir.
- Bu arařtırmanın bir sınırlılıęı olarak madde yanlılıęı iřlenmemiřtir. Ortak maddelerin deęiřen madde fonksiyonuna sahip olma durumları incelenmiřtir. Bundan sonraki arařtırmalarda ortak maddelerde "madde yanlılıęı" varlıęında test eřitleme hataları karřılařtırılabilir.
- Bu arařtırmada fen ve teknoloji bařarı testi formlarının eřitlenmesinde ortak maddelerin DMF'li veya DMF'siz olması durumlarının eřitleme hatalarına etkisi alıřılmıřtır. Bundan sonraki arařtırmalarda farklı kapsamda testlerle

alıřılarak eřitleme hatalarının llen zellik bakımından deęiřip deęiřmedięi alıřılabilir.

- Bu arařtırmada eřitleme deseni olarak kullanılan eřdeęer gruplarda ortak madde deseni yerine dięer eřitleme desenleri ile alıřılabilir.
- Bu arařtırmada gerek veri zerinde MTK'da madde parametrelerinin ayrı kalibrasyonuna dayalı karakteristik eęrisi (HB ve S-L) ve moment yntemleri (O-O, O-S) ile eřitleme yapılmıřtır. Bundan sonraki arařtırmalarda gerek veri zerinde eřzaman kalibrasyon, ok dzeyli madde tepki modelleri, Kernel, Teta regresyon, FCIP gibi KTK ve MTK'ya dayalı farklı eřitleme yntemleriyle hesaplanan eřitleme hataları karřılařtırılabilir.
- Bu arařtırmada yetenek kestirimlerindeki eřitleme hataları hesaplanmıřtır; dięer arařtırmalarda a veya b parametreleri kestirimlerindeki eřitleme hataları hesaplanarak eřitleme yntemleri karřılařtırılabilir.
- Bu arařtırmada sadece gerek veri kullanılmıřtır. Simlasyon ve gerek veri birlikte alıřılıp eřitleme yntemleri birbiriyle karřılařtırılabilir.
- Bu arařtırmada eřitlemenin deęerlendirme kriteri olarak RMSD kullanılmıřtır. Sonraki arařtırmalarda farklı deęerlendirme kriterleri (BIAS, SRMSD, MSE, SEE vb.) iře kořularak birbiriyle kıyaslanabilir.

KAYNAKÇA

- Acar, T. (2008). *Maddenin farklı fonksiyonlaşmasını belirlemede kullanılan genelleştirilmiş aşamalı doğrusal modelleme, lojistik regresyon ve olabilirlik oranı tekniklerinin karşılaştırılması*. Yayınlanmamış Doktora Tezi, Ankara: Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü.
- AERA (American Educational Research Association), APA (American Psychological Association) ve NCME (National Council on Measurement in Education) (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.
- Angoff, W.H. (1971). Scales, Norms, and Equivalent Scores. In Thorndike, R.L. (Ed.) *Educational Measurement, 508-600*. American Council on Education, US: Washington D.C.
- Atalay Kabasakal, K. (2014). *Değişen madde fonksiyonunun test eşitlemeye etkisi*. Yayınlanmamış Doktora Tezi, Ankara: Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Bekçi, B. (2007). *Orta öğretim kurumları öğrenci seçme ve yerleştirme sınavının değişen madde fonksiyonlarının cinsiyete ve okul türüne göre incelenmesi*. Yayınlanmamış Yüksek Lisans Tezi, Ankara: Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü.
- Babcock, B., & Albano, A. D. (2012). Rasch scale stability in the presence of item parameter and trait drift. *Applied Psychological Measurement, 36(7)*, 565-580.
- Bakan Kalaycıoğlu, D. (2008). *Öğrenci Seçme Sınavı'nın madde yanlılığı açısından incelenmesi*. Yayınlanmamış Doktora Tezi, Ankara: Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü.
- Baker, F.B., & Al-Karni, A. (1991). A comparison of two procedures for computing IRT equating coefficients. *Journal of Educational Measurement, 28(2)*, 147-162.
- Berberoğlu, G. ve Bakan Kalaycıoğlu, D. (2008). *The comparison of dif detection methods in the university entrance examinations in Turkey*. 6th Conference of the International Test Commission, UK: Liverpool.
- Bozdağ, S. (2007). *Şans başarısının test eşitlemeye etkisi*. Yayınlanmamış Yüksek Lisans Tezi, Mersin: Mersin Üniversitesi, Sosyal Bilimler Enstitüsü.
- Büyüköztürk, Ş. (2005). *Sosyal bilimler için veri analizi el kitabı*. 5. Baskı Ankara: Pegem Yayınları.
- Byrne, B. M. (1989). *A primer of Lisrel: Basic applications and programming for confirmatory factor analytic models*. US: Springer-Verlag, New York Inc.
- Camilli, G., & Shepard, L. A. (1994). *Methods for identifying biased test items*. Volume 4, Thousand Oaks, CA: Sage Publications.
- Carin A. A., & Bass J. E. (2001). *Teaching science as inquiry*. (8th Edition). Upper Saddle River, N.J.: Merrill/Prentice-Hall.

- Carmines, E.G., & McIver, S.P. (1981). Analyzing models with unobserved variables: Analysis of covariance structures. In G. W. Bohrnstedt and E.F. Borgatta (Eds.), *Social measurement: current issues*, 65-115. Beverly Hills, California: Sage Publications, Inc.
- Chu, K., & Kamata, A. (2000). *Nonequivalent group equating via 1 p-hgllm* (Report). US: Florida.
- Chu, K. (2002). *Equivalent group test equating with the presence of differential item functioning*. Unpublished Doctoral Dissertation. US: Florida State University.
- Chu, K., & Kamata, A. (2005). Test equating with the presence of dif. *Journal of Applied Measurement*, 6(3), 342-354.
- Cohen, A. S., & Kim, S. (1998). An investigation of linking methods under the graded response model. *Applied Measurement In Education*, 22(2), 116-130.
- Cook, L. L., Eignor, D. R., & Taft, H. L. (1985). *A comparative study of curriculum effects on the stability of IRT and conventional item parameter estimates*. Princeton NJ: Educational Testing Service.
- Cook, L. L., & Paterson, N. S. (1987). Problems related to the use of conventional and item response theory equating methods in less than optimal circumstances. *Applied Psychological Measurement*, 11, 225-244.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. FL, Orlando: Harcourt Brace Jovanovich, Inc.
- Çepni, Z. (2011). *Değişen madde fonksiyonlarının sibtest, Mantel Haenzsel, lojistik regresyon ve madde tepki kuramı yöntemleriyle incelenmesi*. Yayımlanmamış Doktora Tezi. Ankara: Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü.
- Davies, A. A., Holland P., & Thayer, D. T. (2004). *The Kernel method of test equating*. Springer, US:New York.
- Doğan, N. ve Öğretmen, T. (2006). Madde yanlılığını belirleme teknikleri arasında bir karşılaştırma. *Eğitim Araştırmaları*, 23, 94-105.
- Doğan, N. ve Öğretmen, T. (2008). Değişen madde fonksiyonunu belirlemede Mantel-Haenzsel, ki-kare ve lojistik regresyon tekniklerinin karşılaştırılması. *Eğitim ve Bilim*, 33, 100-112.
- Ellis, B. B., & Raju, N. S. (2003). *Test and item bias: What they are, what they aren't and how to detect them*. Educational Resources information center (ERIC), Chapter 7, 88-99.
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. USA: Lawrence Erlbaum Associates.
- Embretson, S. E. (2007). Construct validity: A universal validity system or just another test evaluation procedure? *Educational Researcher*, 36(8), 449-455.
- Gao, X., Hanson, B. A., & Harris, D. J. (1999). *Effect of using different common item sets under the common item non-equivalent groups design*. Paper Presented at the Annual Meeting of the American Educational Research Association, CAN: Montreal.

- González, A.; Padilla, J.L.; Hidalgo, M.D. Gómez-Benito, J. ve Benítez, I. (2011). Easydif: Software for analysing differential item functioning using the Mantel-Haenszel and standardization procedures. *Applied Psychological Measurement*, 35, 483-484.
- Gök, B. (2012). *Denk olmayan gruplarda ortak madde deseni kullanılarak madde tepki kuramına dayalı eşitleme yöntemlerinin karşılaştırılması*. Yayınlanmamış Doktora Tezi, Ankara: Hacettepe Üniversitesi.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. US: Kluwer-Nijhoff Publishing.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage Publications.
- Han, K.T. (2008). *Impact of item parameter drift on test equating and proficiency estimates*. Unpublished Doctoral Dissertation. US: University of Massachusetts.
- Han, K. T. (2009). IRTEQ: Windows application that implements IRT scaling and equating (computer programme). *Applied Psychological Measurement*, 33(6), 491-493.
- Hannick P. L., & Huang C. (2002). *Effects of decreasing the number of common items in equating link items sets*. Paper presented at AERA. US: New Orleans.
- Hanson, B.A., & Béguin, A. A. (2002). Obtaining a common scale for item response theory item parameters using separate versus concurrent estimation in the common-item equating design. *Applied Psychological Measurement*, 26(1), 3-24.
- Harris, D. J., & Crouse, J. D. (1993). A study of criteria used in equating. *Applied Measurement in Education*, 6(3), 195-240.
- Hidalgo Montesinos, M. D., & Lopez Pina, J. A. (2002). Two-stage equating in differential item functioning detection under the graded response model with the Raju area measures and Lord statistic. *Educational and Psychological Measurement*. 62(1), 32.
- Holland, P. W., & Thayer, D. T. (1988). Differential item performance and the Mantel-Haenszel procedure. In H.WainerveH.I.Braun (Eds.), *Test Validity Hillsdale*, 129-145. NJ: Lawrence Erlbaum,
- Holland, P. W., & Wainer, H. (1993). *Differential item functioning*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Holland, P. W., & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), *Educational Measurement (4th Edition)*. 187-220. Westport, CT: American Council on Education and Praeger.
- Hu, H., Rover, W. T., & Vukminovic, Z. (2008). Investigation of IRT-based equating methods in the presence of outlier common items. *Applied Psychological Measurement*. 32(4), 311-333.
- Huggins, A. C. (2014). The effect of differential item functioning in anchor items on population invariance of equating. *Educational and Psychological Measurement*. 74(4), 627-658.

- Hwang, C., & Cleary, T. A. (1986). *Comparing IRT pre-equating and section preequating: A simulation study*. Paper Presented at the Annual Meeting of the American Educational Research Association, US: San Francisco.
- Jöreskog, K.G., & Sorböm, D. (1986). *Prells a program for multivariate data screening and data summarization: A preprocessor for Lisrel*. Mooresville, Ind.: Scientific Software Inc.
- Jöreskog, K.G., & Sorböm, D. (1993). *Lisrel 8: Structural equation modeling with the simplis command language*. Lincolnwood: Scientific Software International, Inc.
- Kan, A. (2010). Test eşitleme: Aynı davranışları ölçen, farklı madde formlarına sahip testlerin istatistiksel eşitliğinin sınanması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(1), 16-21.
- Kan, A. (2011). Test eşitleme: OKS testlerinin istatistiksel eşitliğinin sınanması. *Eğitim ve Bilim*, 36(160), 38-51.
- Kang, T., & Cohen, A. S. (2007). IRT model selection methods for dichotomous items. *Applied Psychological Measurement*, 31(4), 331-358.
- Karkee, T. B., & Wright, K. R. (2004). *Evaluation of linking methods for placing three-parameter logistic item parameter estimates onto a one parameter scale*. Paper presented at the Annual Meeting of the American Educational Research, US: San Diego, California, April 16.
- Kaskowitz, G. S., & De Ayala, R. J. (2001). The effect of error in item parameter estimates on the test response function method of linking. *Applied Psychological Measurement*, 25, 39-52.
- Kelecioğlu, H. (1994). *Öğrenci seçme sınavı puanlarının eşitlenmesi üzerine bir çalışma*. Yayımlanmamış Doktora Tezi, Ankara: Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü.
- Kelley, T. L. (1939). The selection of upper and lower groups for the validation of test items. *The Journal of Educational Psychology*, 30(1), 17-24.
- Kelloway, E. K. (1996). Structural equation modelling in perspective. *Journal Of Organizational Behavior (Impact Factor: 3.85)*, 16(3): 215 - 224.
- Kim, S., & Cohen, A. S. (1998). A comparison of linking and concurrent calibration under item response theory. *Applied Psychological Measurement*, 22, 131-143.
- Kim, S., & Lee, W. C. (2004). *IRT scale linking methods for mixed-format tests*. ACT Research Report 2004-5. US: IA, ACT Inc.
- Kim, S., & Lee, W. C. (2006). An extension of four IRT linking methods for mixed-format tests. *Journal of Educational Measurement*, 43(1), 53-76.
- Kim, S., & Kolen, M. J. (2006). Robustness to format effects of IRT linking methods for mixed-format tests. *Applied Measurement in Education*, 19(4), 357-381.
- Kline, R. B. (1998). *Principles and practice of structural equation modeling*. New York: The Guilford Press.

- Kolen, M. J., & Brennan, R. L. (1995). *Test equating: Methods and practices*. New York: Springer-Verlag.
- Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling and linking: Methods and practices*. (2nd Edition). New York: Springer-Verlag.
- Kolen, M. J. (2006). Scaling and Norming. In R. L. Brennan (Ed.), *Educational Measurement. (4th Edition)*. 155-186. Westport, CT: American Council on Education and Praeger.
- Lee, W. C., & Ban, J. C. (2010). A comparison of IRT linking procedures. *Applied Measurement in Education*, 23(1), 23-48.
- Linn, R. L., Levine, M. V., Hastings, C. N., & Wardrop, J. L. (1981). Item bias in a test of reading comprehension. *Applied Psychological Measurement*, 5, 159-173.
- Lord, F. M. (1950). Notes on comparable scales for test scores. *Educational Testing Service Research Bulletin (RB) 50-48*, US: NJ, Princeton ETC.
- Lord, F. M. (1955). Equating test scores: A maximum likelihood solution. *Psychometrika*, 20, 193-200.
- Lord, F. M. (1980). Applications of item response theory to practical testing problems. Hillsdale, New Jersey: Lawrence Erlbaum Associates Publishers.
- MacCann, R. G. (1990). Derivations of observed score equating methods that cater to populations differing in ability. *Journal of Educational Statistics*, 15(2), 146-170.
- Mannocci, A. (2009). The Mantel-Haenszel procedure: 50 years of the statistical method for confounders control. *Italian Journal of Public Health*. 6(4), 338-340.
- Mantel N., & Haenszel W. M. (1959). Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of The National Cancer Institute*, 22, 719-48.
- McDonald, Roderick P. (1999). Test theory: A unified treatment. New Jersey: Lawrence Erlbaum Associates, Publishers.
- MEB (Milli Eğitim Bakanlığı) (2006). *İlköğretim fen ve teknoloji dersi (6, 7 ve 8. sınıflar) öğretim programı*. [Çevrim-içi: <http://ttkb.meb.gov.tr/www/ogretim-programlari/icerik/72>], Erişim tarihi: 15 Ocak 2013.
- Meng, Y. (2012). *Comparison of kernel equating and item response theory equating methods*. Unpublished Doctoral Dissertation. US: University of Massachusetts.
- Norman Dvorak, R. L. (2009). *A comparison of kernel equating to the test characteristic curve method*. Unpublished Doctoral Dissertation. Nebraska: University of Nebraska.
- Ogasawara, H. (2001a). Item response theory true score equating and their standard errors. *Journal Of Educational Behavioral Statistics*, 26(1), 31-50.
- Ogasawara, H. (2001b). Least square estimations of item response theory linking coefficients. *Applied Psychological Measurement*, 25(4), 3-21.

- Özdemir, D. (2004). Çoktan seçmeli testlerin klasik test teorisi ve örtük özellikler teorisine göre hesaplanan psikometrik özelliklerinin iki kategorili ve ağırlıklandırılmış puanlanması yönünden karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi, Sayı: 26*, 117-123.
- Öztürk, N. (2010). *Akademik personel ve lisansüstü eğitimi giriş sınavı puanlarının eşitlenmesi üzerine bir çalışma*. Yayımlanmamış Yüksek Lisans Tezi, Ankara: Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü.
- Paek, I., & Young, M. J. (2005). Investigation of student growth recovery in a fixed-item linking procedure with a fixed-person prior distribution for mixed-format test data. *Applied Measurement in Education, 18*(2), 199-215.
- Padilla, J. L., Hidalgo, M.D., Benítez, I., & Gómez-Benito, J.(2012). Comparison of three software programs for evaluating dif by means of the mantel-haenszel procedure: easy-dif, difas and ezdif. *Psicología, 33*, 135-156.
- Pang, X., EMadera, E., Radwan, N., & Zhang, S. (2010). *A comparison of four test equating methods*. Report Prepared for the Education Quality and Accountability Office (EQAO).
- Peterson, N.S., Cook L.L., & Stocking, M.L. (1983). IRT versus conventional equating methods: A comparative study of scale stability. *Journal of Educational Statistics, 8*(2), 137.
- Rogers, H. J., & Swaminathan, H. (1993). A comparison of logistic regression and mantel-haenszel procedures for detecting differential item functioning. *Applied Psychological Measurement, 17*, 105-116.
- Roznowski, M., & Reith J. (1999). Examining the measurement quality of tests containing differentially functioning items: Do biased items result in poor measurement? *Educational and Psychological Measurement, 59*(2), 248-269.
- Schumacker, R. E., & Lomax, R. G. (2004). *A beginner's guide to structural equation modeling. (2nd edition)*. US: Mahwah-New Jersey, Lawrence Erlbaum Associates, Inc. Publishers.
- Shealy, R., & Stout, W. F. (1993). A model-based standardization approach that separates true bias/ dif from group ability differences and detects test bias/ dtf as well as item bias/ dif. *Psychometrika, 58*, 159-194.
- Speron, E. (2009). *A comparison of metric linking procedures in item response theory*. Unpublished Doctoral Dissertation. Illinois, Chicago: University Of Illinois.
- Stocking, M.L., & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement, 7*(2), 201-210.
- Swaminathan, H., & Rogers, H. J. (1990). Detecting differential item functioning using logistic regression procedures. *Journal of Educational Measurement, 27*(4), 361-370.
- Şahhüseyinoğlu, D. (2005). *İngilizce yeterlik sınavı puanlarının üç farklı eşitleme yöntemine göre karşılaştırılması*. Yayımlanmamış Doktora Tezi, Ankara: Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü.

- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. (5th Pearson International Edition). USA: Pearson Education.
- Tate, R. (2000). Performance of a proposed method for the linking of mixed format tests with constructed response and multiple choice items. *Journal of Educational Measurement*, 37(4), 329-346.
- Tekin, H. (1991). *Eğitimde ölçme ve değerlendirme*. Ankara: Yargı Yayınları.
- Turhan, A. (2006). *Multilevel 2PL item response model vertical equating with the presence of differential item functioning*. Unpublished Doctoral Dissertation. US: Florida State University.
- Uysal, İ. (2014). *Madde tepki kuramı'na dayalı test eşitleme yöntemlerinin karma modeller üzerinde karşılaştırılması*. Yayınlanmamış Yüksek Lisans Tezi, Bolu: Abant İzzet Baysal Üniversitesi, Eğitim Bilimleri Enstitüsü.
- Way, W. D., & Tang, K. L. (1991). *A comparison of four logistic model equating methods*. Paper Presented at the Annual Meeting of the American Educational Research Association. US: Chicago.
- Wells, C. S., Subkoviak, M. J., & Serlin, R. C. (2002). The effect of item parameter drift on examinee ability estimates. *Applied Psychological Measurement*, 26(1), 77-87.
- Wyse, A. E., & Reckase M. D. (2011). A graphical approach to evaluating equating using test characteristic curves. *Applied Psychological Measurement*, 35(3), 217-234.
- Yurdugül, H. (2003). *Ortaöğretim kurumları seçme ve yerleştirme sınavının madde yanlılığı açısından incelenmesi*. Yayınlanmamış Doktora Tezi, Ankara: Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü.
- Zhang, Y., Matthews-Lopez, J., & Dorans, N.J. (2003). *Using dif dissection to assess effect of item deletion due to dif on the performance of sat i: Reasoning test sub-populations* (Report), US: New Jersey.
- Zimowski, M. F., Muraki, E., Mislevy, R. J., & Bock, R. D. (1996). *Bilog-mg: Multiple-group IRT analysis and test maintenance for binary items*. Chicago: Scientific Software International.
- Zumbo, B.D.A. (1999). *Handbook on the theory and methods of differential item functioning: Logistic regression modelling as a unitary framework for binary and likert-type item scores*. Ottawa: Directorate of Human Resources Research and Evaluation, Department of National Defense.
- Zumbo, B. D. A., & Thomas, D. R. (1996). *A measure of dif effect size using logistic regression procedures*. Paper Presented at the National Board of Medical Examiners. US, Philadelphia.
- Zwick, R., & Ercikan, K. (1989). Analysis of differential item functioning in the naep history assessment. *Journal of Educational Measurement*, 26, 55-66.

EKLER DİZİNİ

EK 1. A VE B FORMU CİNSİYETE DAYALI DMF-MH VE DMF-LR ANALİZ SONUÇLARI

A Formu Cinsiyete Dayalı DMF-MH ve DMF-LR Analiz Sonuçları

Madde No	DMF-MH			DMF-LR				
	Δmh	αMH	p	R_1^2	R_2^2	R_3^2	ΔR^2 ($R_2^2 - R_1^2$)	ΔR^2 ($R_3^2 - R_2^2$)
1	0.603	0.773	0.142	0.174	0.179	0.181	0.005	0.002
2	-0.571	1.275	0.208	0.316	0.320	0.326	0.004	0.006*
3	-0.883	0.145	0.053	0.345	0.352	0.353	0.007*	0.001
4	0.702	0.741	0.117	0.311	0.314	0.315	0.003	0.001
5	-1.233	1.690	0.001*	0.108	0.126	0.126	0.018*	0.000
6	-2.384	2.758	0.000*	0.168	0.215	0.219	0.047*	0.004
7	-1.587	1.965	0.000*	0.071	0.096	0.096	0.025*	0.000
8	-0.550	1.263	0.279	0.327	0.329	0.329	0.002	0.000
9	-0.249	1.112	0.622	0.311	0.312	0.312	0.001	0.000
10	-1.028	1.549	0.012*	0.154	0.165	0.165	0.011*	0.000
11	0.160	0.933	0.744	0.013	0.014	0.014	0.001	0.000
12	-0.077	1.033	0.957	0.014	0.014	0.014	0.000	0.000
13	0.246	0.900	0.627	0.273	0.273	0.275	0.000	0.002
14	0.729	0.733	0.106	0.198	0.207	0.209	0.009*	0.002*
15	0.012	0.994	0.957	0.196	0.196	0.196	0.000	0.000
16	-0.178	1.078	0.808	0.374	0.374	0.376	0.000	0.002
17	-0.084	1.036	0.946	0.329	0.329	0.332	0.000	0.003
18	-0.102	1.044	0.860	0.207	0.207	0.207	0.000	0.000
19	0.807	0.709	0.042*	0.061	0.067	0.075	0.006*	0.008*
20	-0.334	1.152	0.516	0.413	0.414	0.414	0.001	0.000
21	1.333*	0.566	0.005*	0.457	0.466	0.466	0.009*	0.000
22	0.699	0.740	0.130	0.364	0.369	0.370	0.005	0.001
23	0.294	0.882	0.515	0.209	0.210	0.210	0.001	0.000
24	0.521	0.801	0.251	0.288	0.291	0.291	0.003	0.000
25	0.960*	0.664	0.019*	0.113	0.121	0.121	0.008*	0.000
26	0.350	0.861	0.475	0.187	0.188	0.188	0.001	0.000
27	-0.235	1.105	0.605	0.191	0.192	0.192	0.001	0.000
28	0.860	0.693	0.075	0.450	0.453	0.453	0.003	0.000
29	1.025*	0.646	0.023	0.378	0.386	0.386	0.011*	0.000

*p<0.05

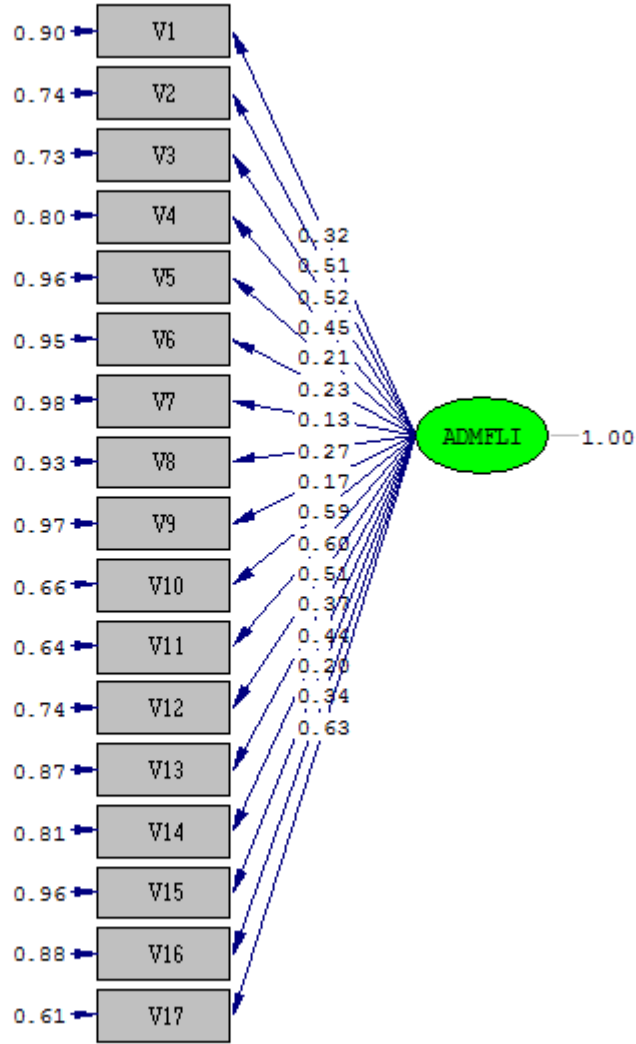
B Formu Cinsiyete Dayalı DMF-MH ve DMF-LR Analiz Sonuçları

Madde No	DMF-MH			DMF-LR				
	Δmh	αMH	p	R_1^2	R_2^2	R_3^2	ΔR^2 ($R_2^2 - R_1^2$)	ΔR^2 ($R_3^2 - R_2^2$)
1	0.972	0.661	0.047*	0.329	0.336	0.337	0.007*	0.001
2	-0.336	1.154	0.459	0.168	0.169	0.173	0.001	0.004
3	0.127	0.947	0.808	0.077	0.077	0.082	0.000	0.005
4	-0.655	1.321	0.150	0.151	0.156	0.156	0.005	0.000
5	-1.754	2.109	0.000*	0.149	0.184	0.186	0.035*	0.002
6	-2.317	2.681	0.000*	0.199	0.244	0.249	0.045*	0.005
7	-1.669	2.035	0.000*	0.132	0.161	0.163	0.029*	0.002
8	0.761	1.382	0.104	0.232	0.238	0.243	0.006	0.005*
9	-0.624	1.304	0.146	0.190	0.194	0.201	0.004	0.007*
10	-1.126	1.614	0.014*	0.364	0.373	0.374	0.012*	0.001
11	-0.156	1.068	0.758	0.032	0.032	0.035	0.000	0.003
12	-0.655	1.012	0.971	0.000	0.001	0.001	0.005	0.000
13	0.506	0.806	0.294	0.231	0.234	0.235	0.003	0.001
14	0.880	0.687	0.059	0.263	0.271	0.275	0.008*	0.004*
15	1.226*	0.593	0.003*	0.143	0.159	0.161	0.016*	0.002*
16	0.853	0.695	0.093	0.243	0.249	0.251	0.006	0.002
17	0.124	0.948	0.871	0.277	0.277	0.277	0.000	0.000
18	0.784	0.716	0.057	0.140	0.146	0.154	0.006	0.008
19	1.209	0.597	0.005*	0.256	0.268	0.269	0.012*	0.001
20	0.143	0.940	0.794	0.236	0.236	0.237	0.000	0.001
21	0.274	0.889	0.583	0.313	0.313	0.315	0.000	0.002
22	0.343	0.864	0.495	0.265	0.266	0.272	0.001	0.006
23	-0.214	1.095	0.690	0.323	0.324	0.325	0.001	0.001
24	1.193	0.601	0.004*	0.223	0.236	0.236	0.013*	0.000
25	0.675	0.750	0.147	0.328	0.330	0.330	0.002	0.000
26	0.292	0.882	0.580	0.332	0.333	0.334	0.001	0.001
27	-1.808	2.158	0.000*	0.236	0.265	0.265	0.029*	0.000
28	0.666	0.753	0.174	0.333	0.337	0.340	0.004	0.003
29	0.752	0.725	0.080	0.211	0.216	0.216	0.005	0.000

*p<0.05

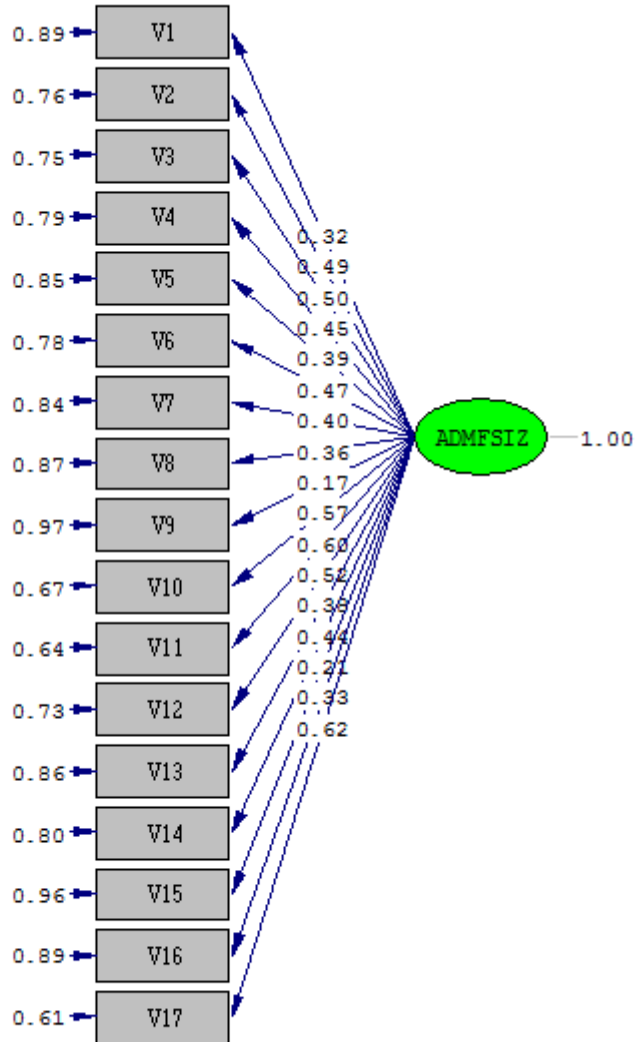
EK 2. A VE B FORMLARININ DOĞRULAYICI FAKTÖR ANALİZİ ÖLÇME MODELLERİ

A FORMU (ORTAK MADDELER DMF'Lİ) DOĞRULAYICI FAKTÖR ANALİZİ ÖLÇME MODELİ



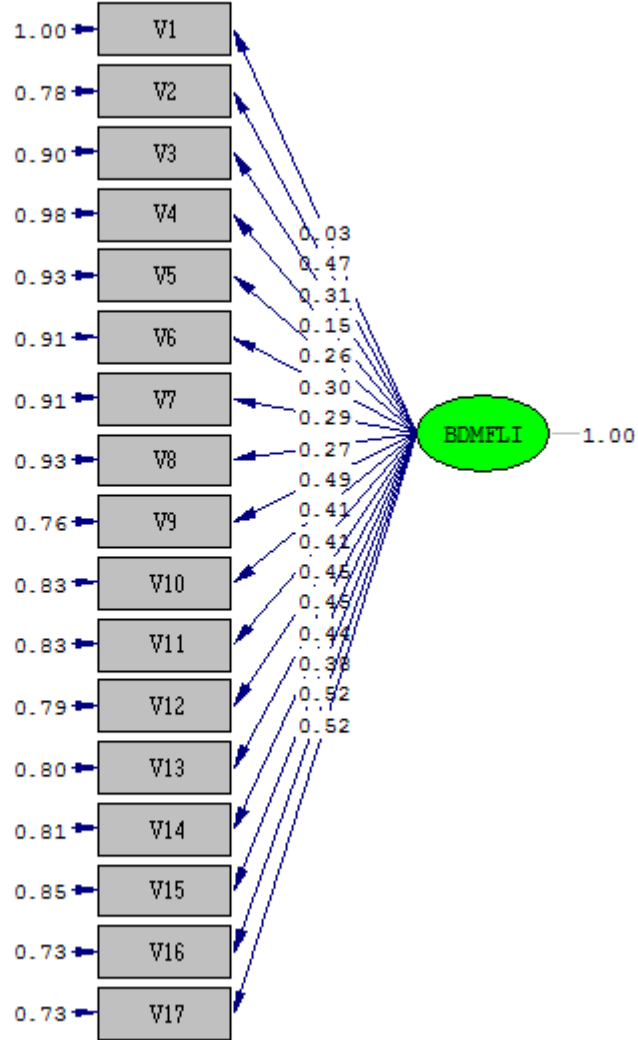
Chi-Square=210.80, df=119, P-value=0.00000, RMSEA=0.033

A FORMU (ORTAK MADDELER DMF'SİZ) DOĞRULAYICI FAKTÖR ANALİZİ ÖLÇME MODELİ



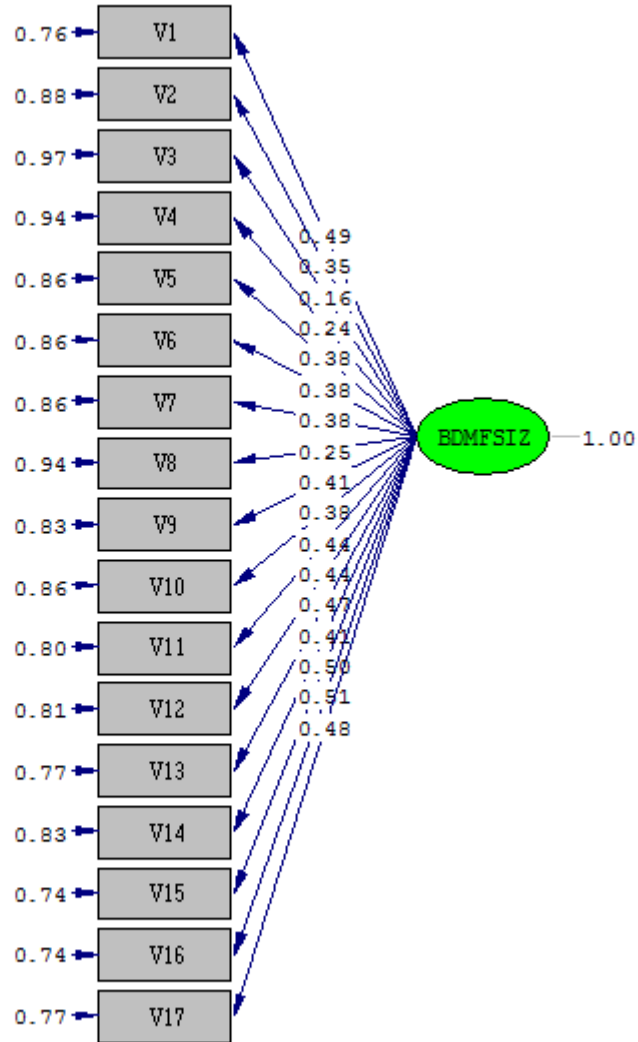
Chi-Square=223.94, df=119, P-value=0.00000, RMSEA=0.036

B FORMU (ORTAK MADDELER DMF'Lİ) DOĞRULAYICI FAKTÖR ANALİZİ ÖLÇME MODELİ



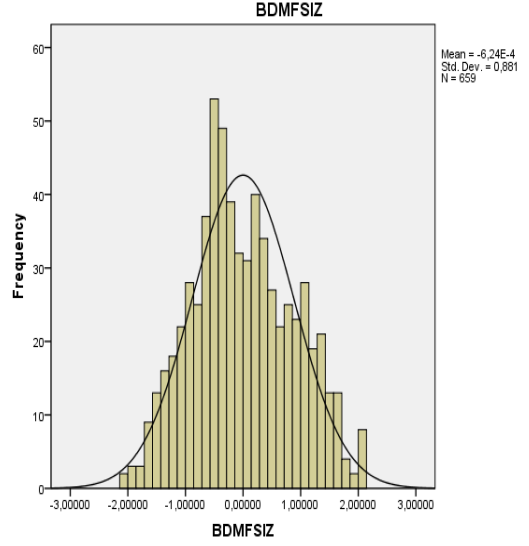
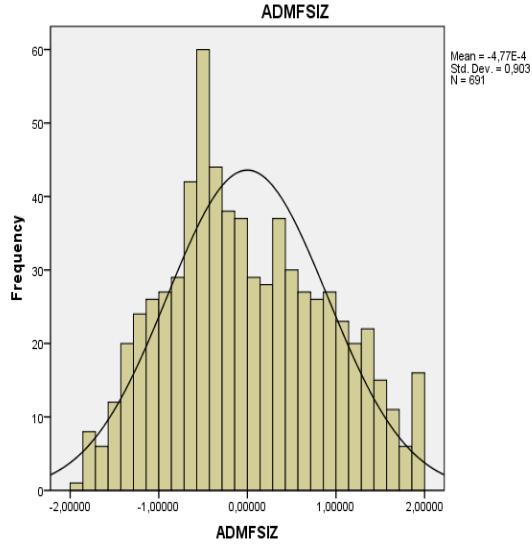
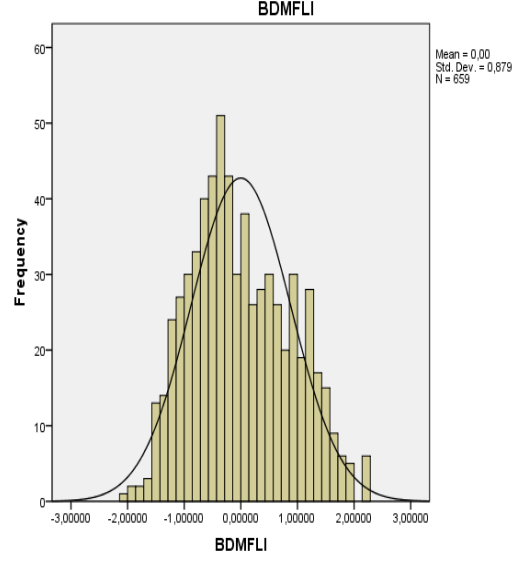
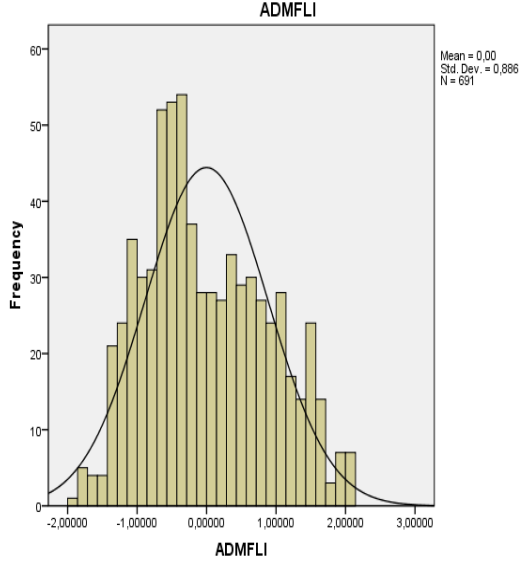
Chi-Square=270.83, df=119, P-value=0.00000, RMSEA=0.044

B FORMU (ORTAK MADDELER DMF'SİZ) DOĞRULAYICI FAKTÖR ANALİZİ ÖLÇME MODELİ



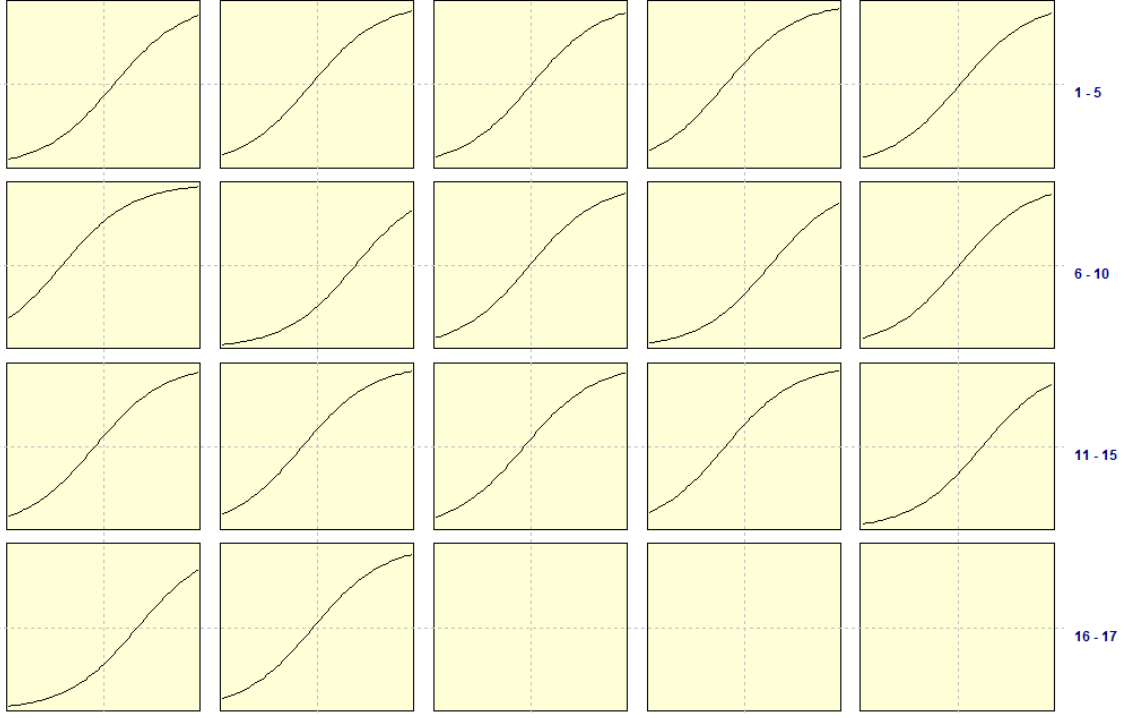
Chi-Square=190.77, df=119, P-value=0.00003, RMSEA=0.030

EK 3. A VE B FORMLARINA AİT KESTİRİLEN YETENEKLERİN (θ) NORMAL DAĞILIM EĞRİLERİ

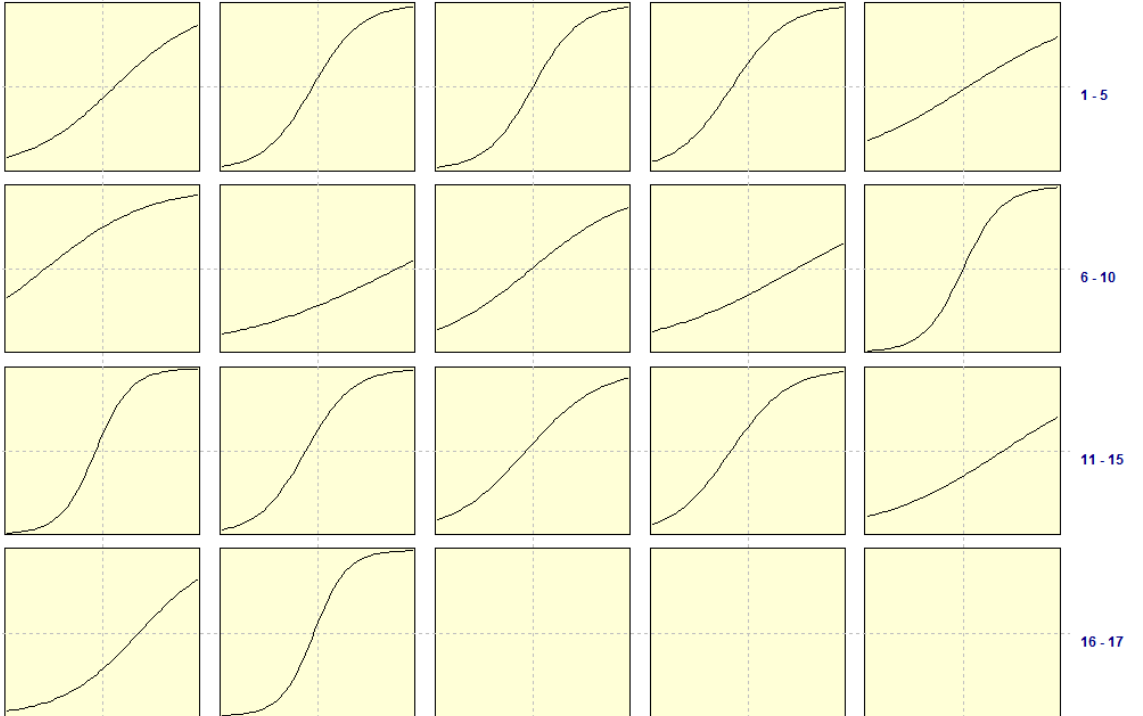


EK 4. MADDE KARAKTERİSTİK EĞRİLERİ

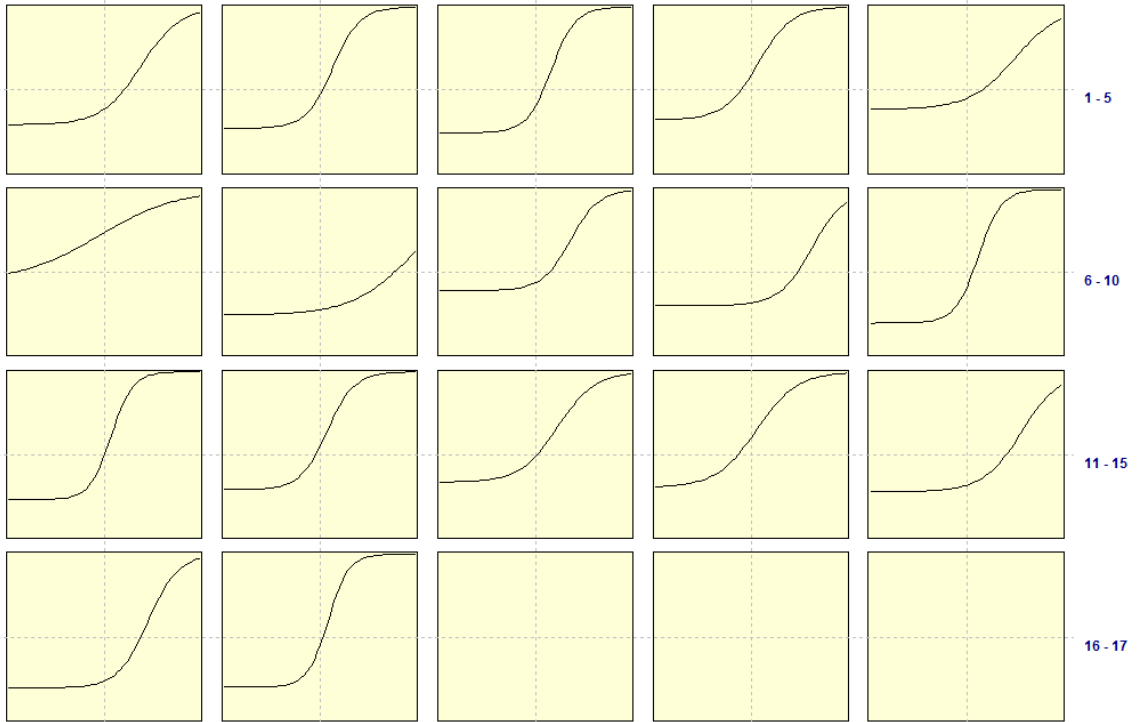
ADMFLI 1PLM MKE



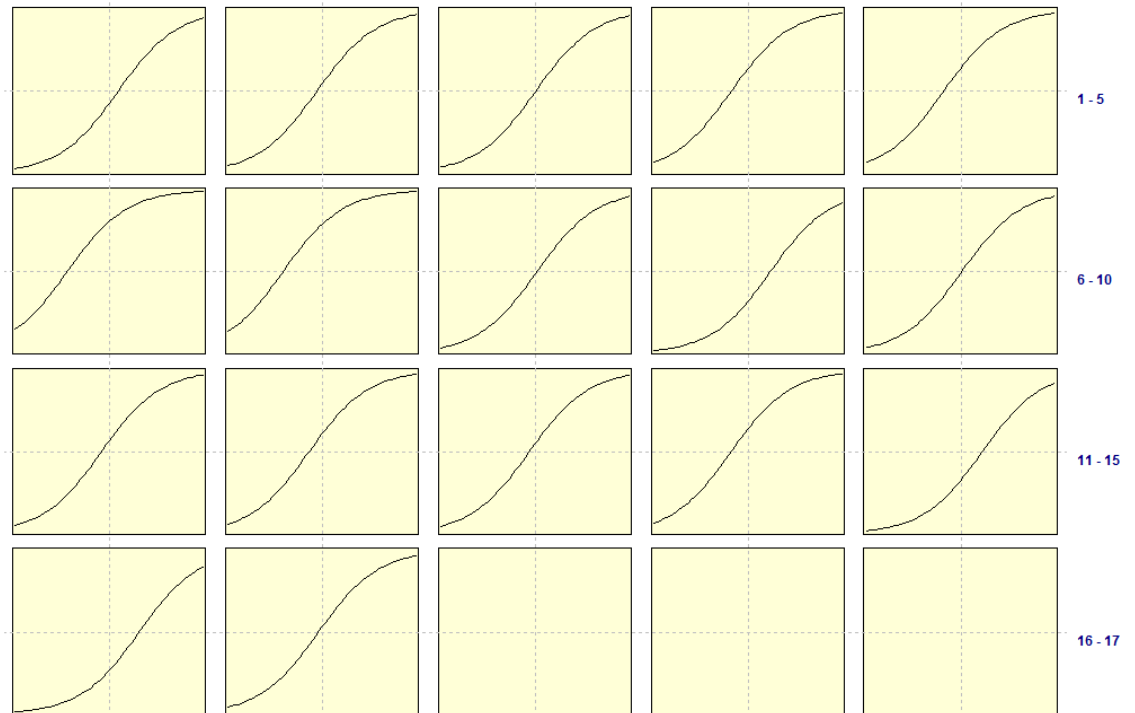
ADMFLI 2PLM MKE



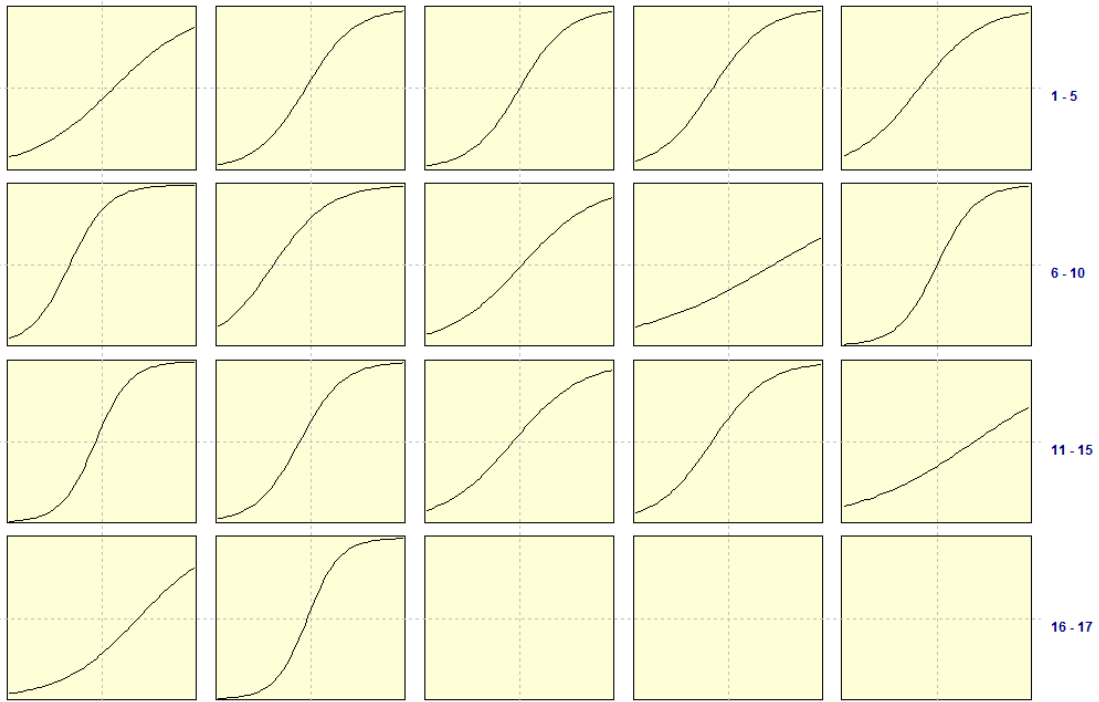
ADMFLI 3PLM MKE



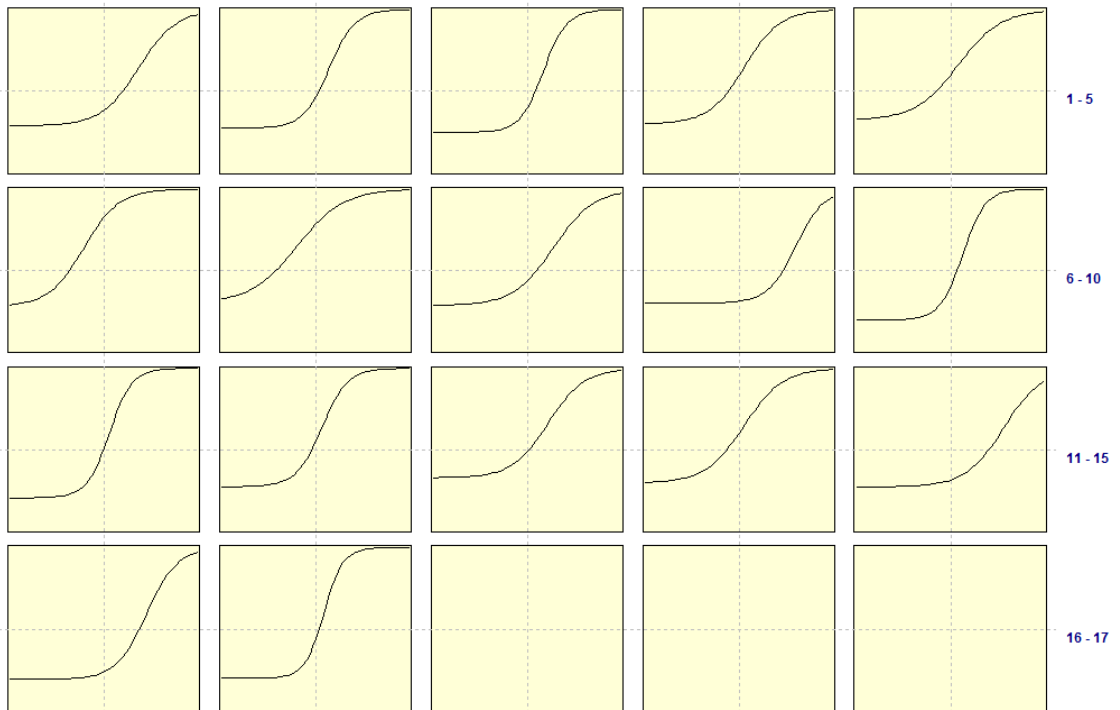
ADMFSIZ 1PLM MKE



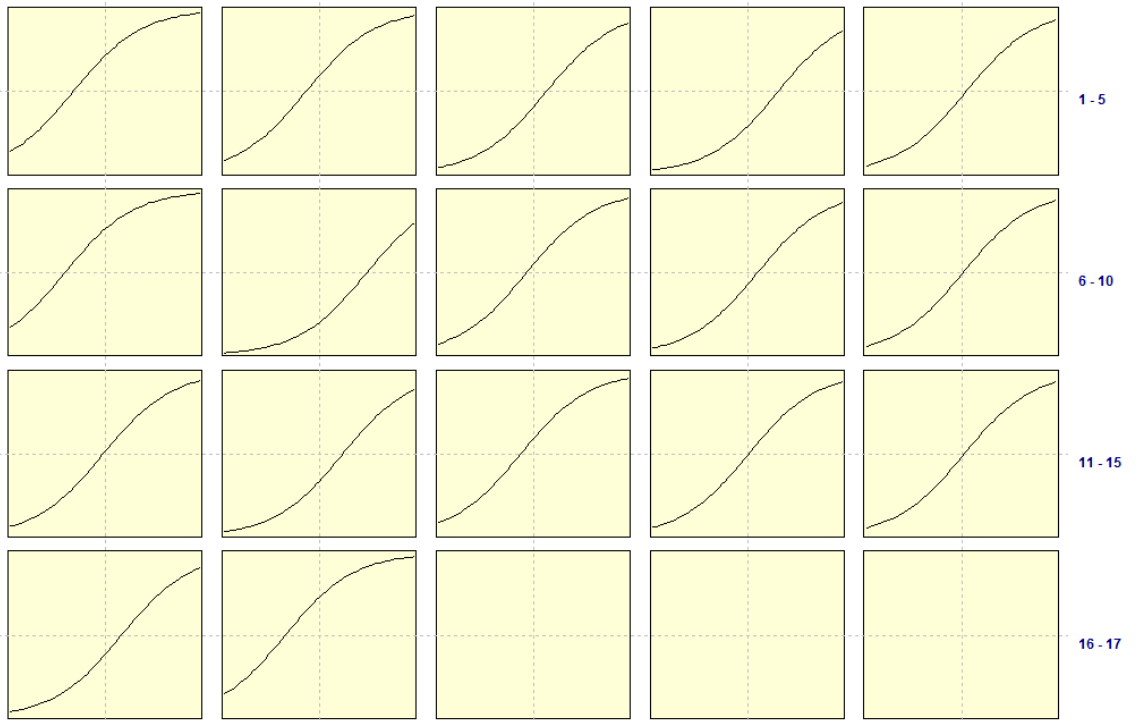
ADMFSIZ 2PLM MKE



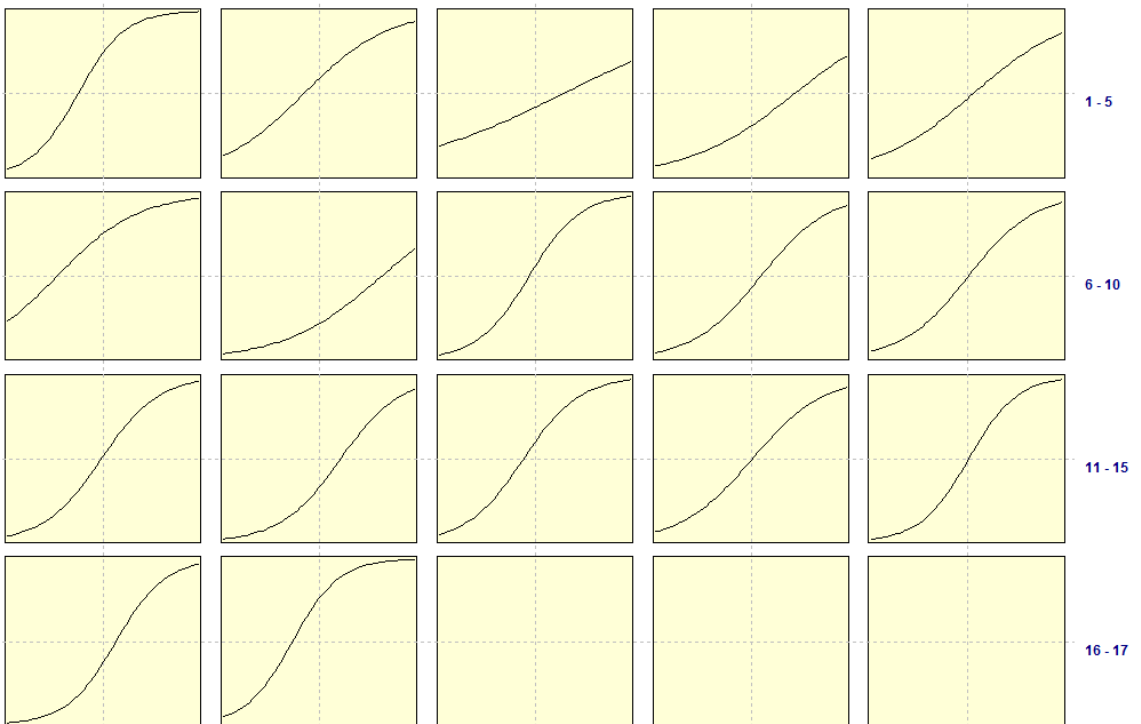
ADMFSIZ 3PLM MKE



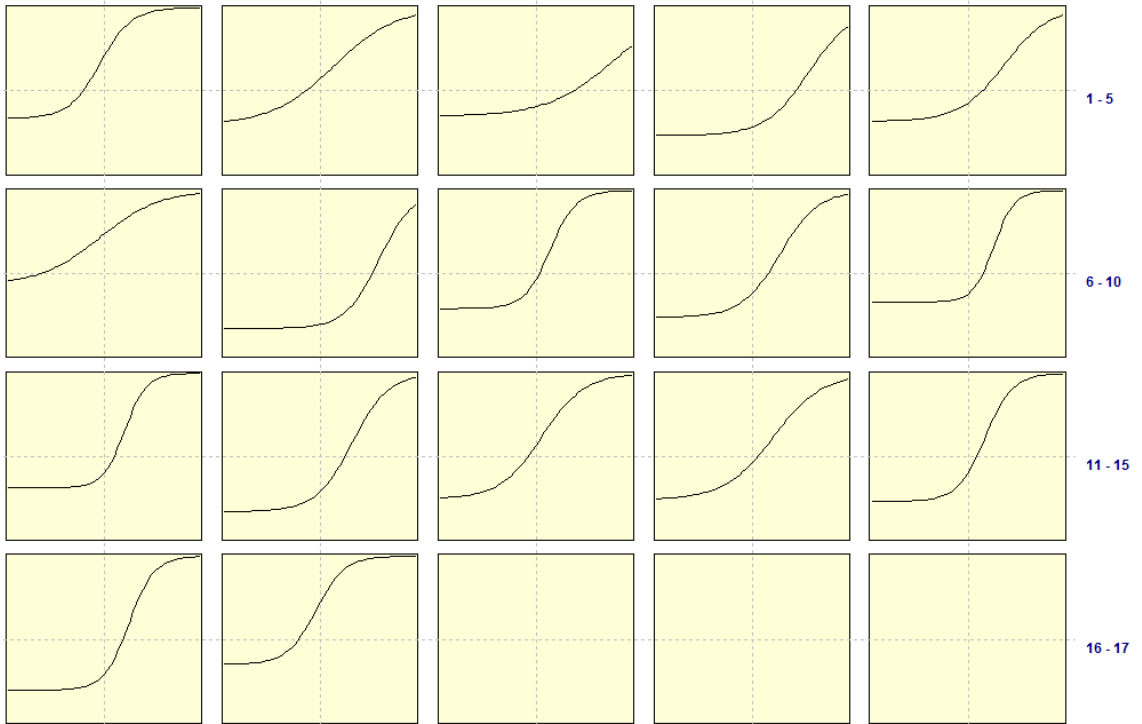
BDMFLI 1PLM MKE



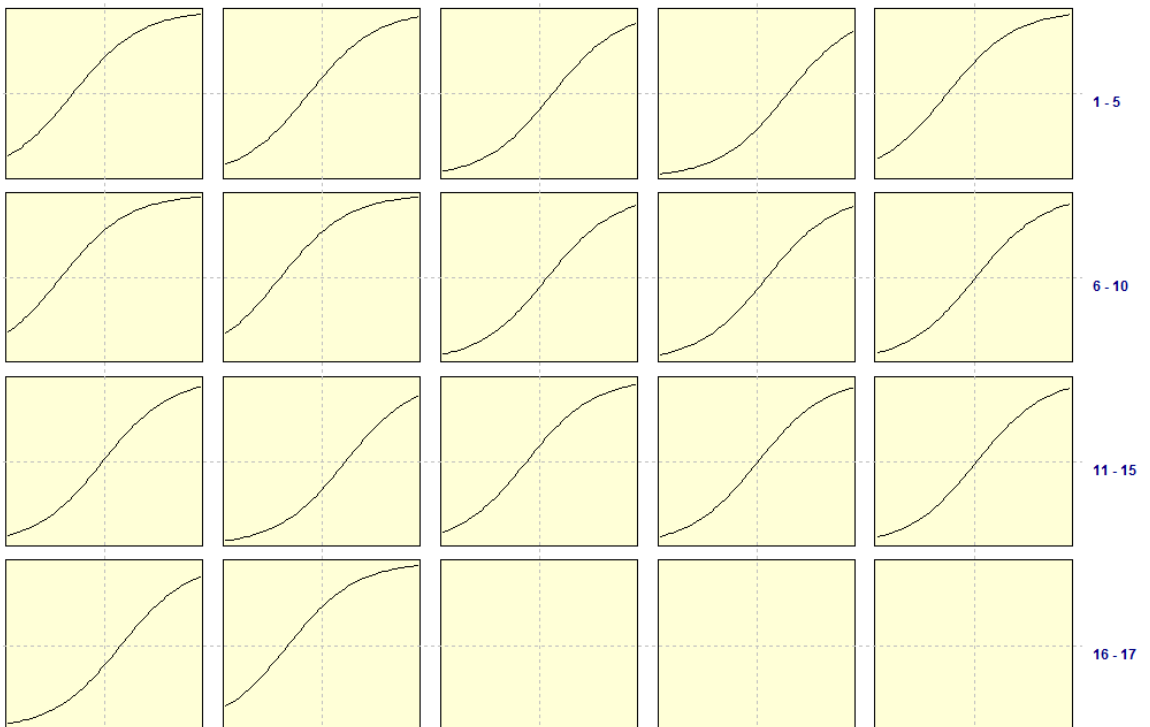
BDMFLI 2PLM MKE



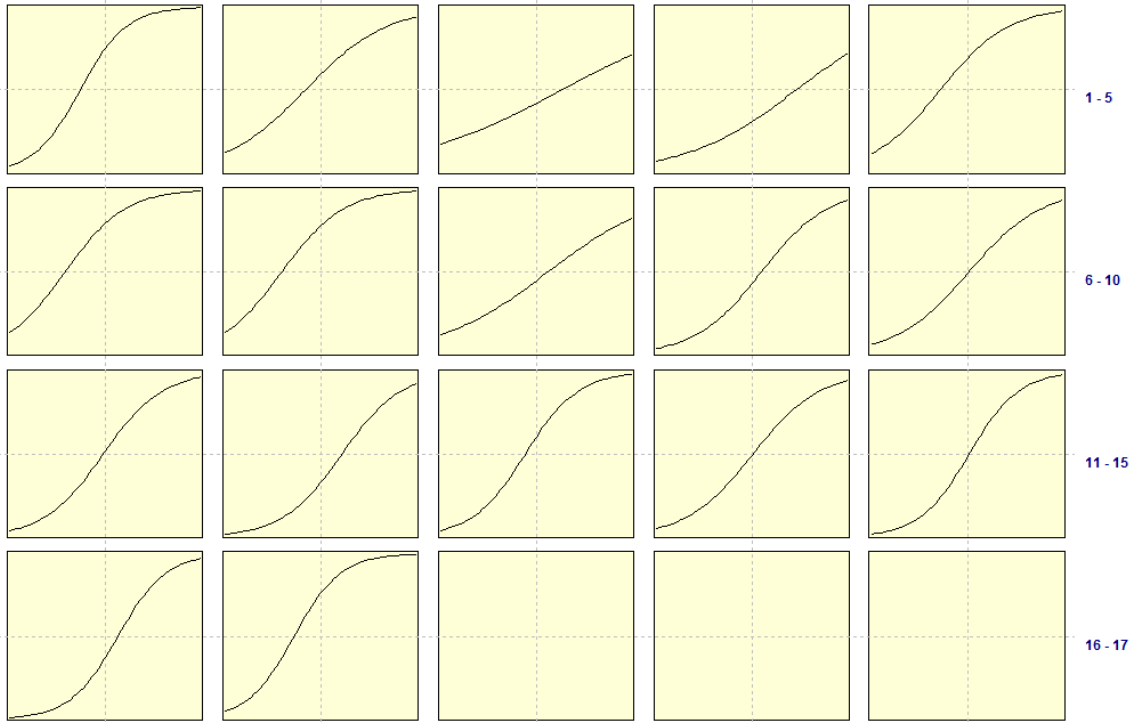
BDMFLI 3PLM MKE



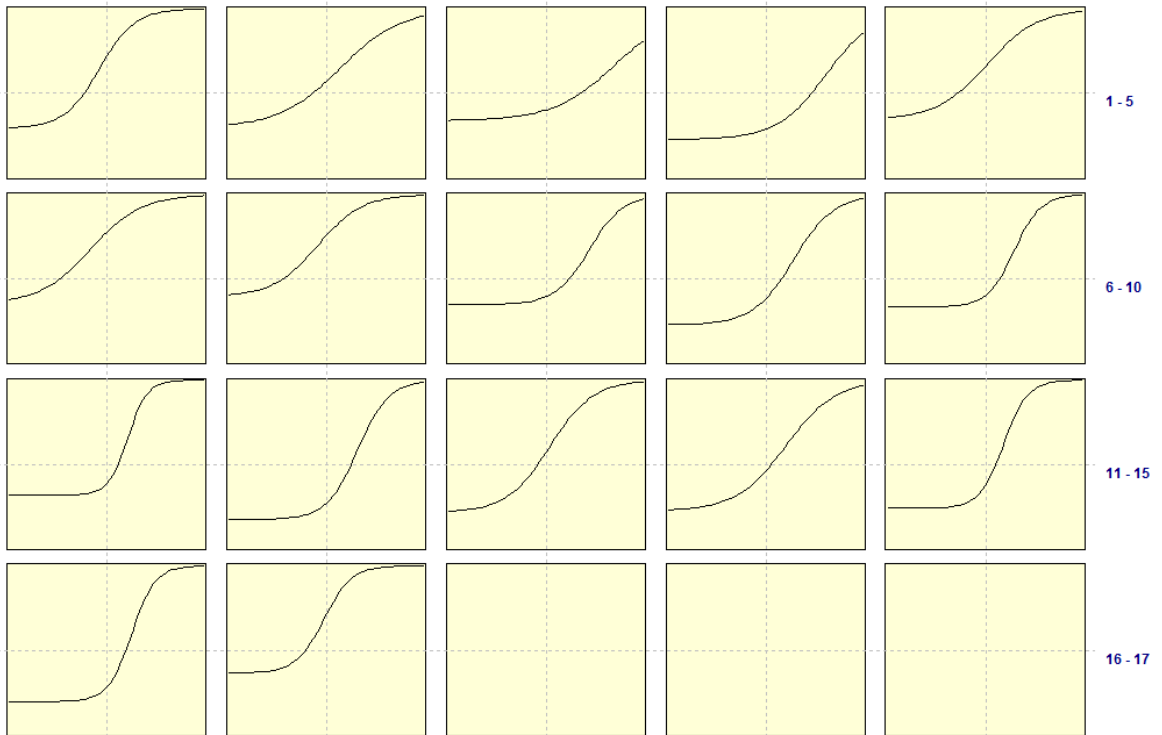
BDMFSIZ 1PLM MKE



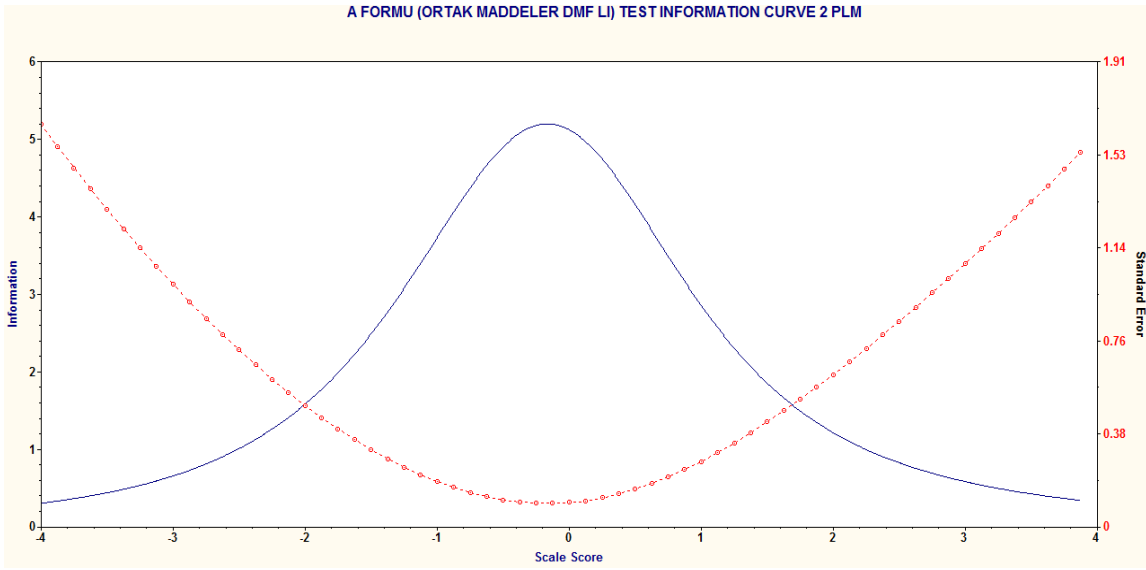
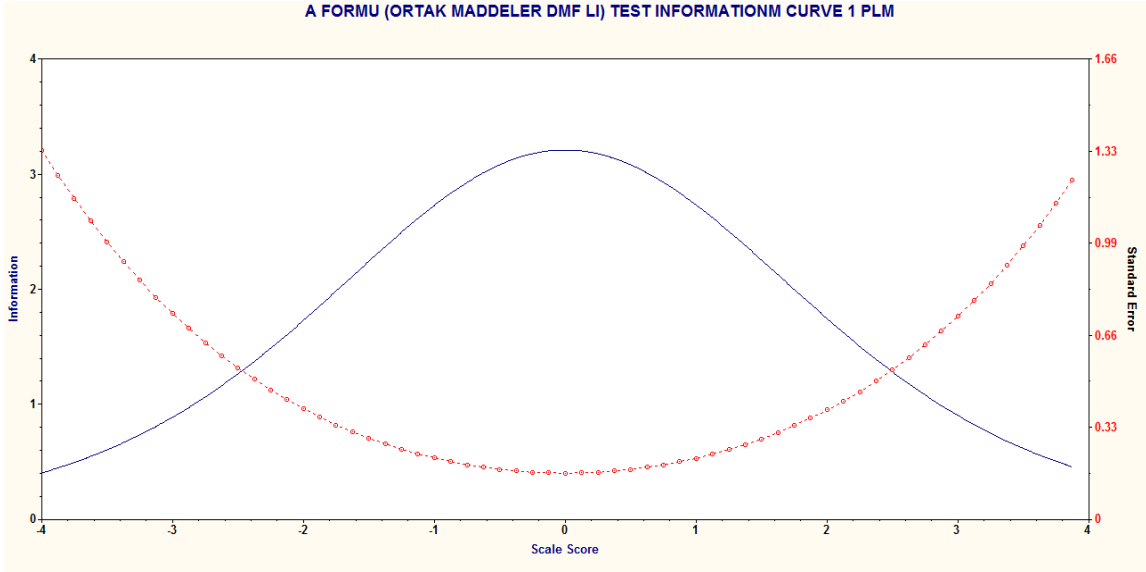
BDMFSIZ 2PLM MKE

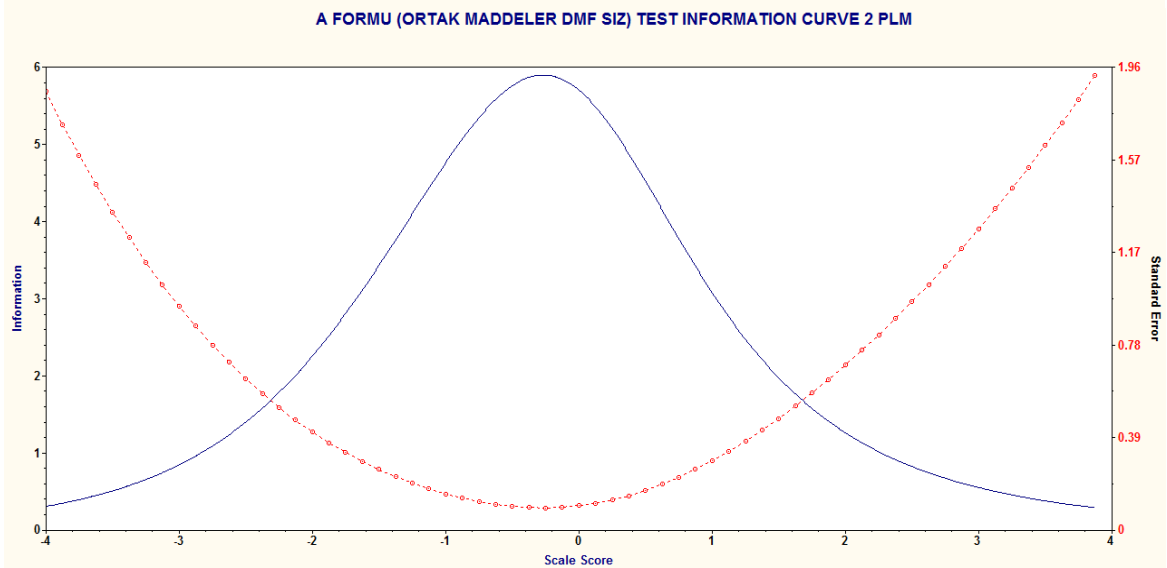
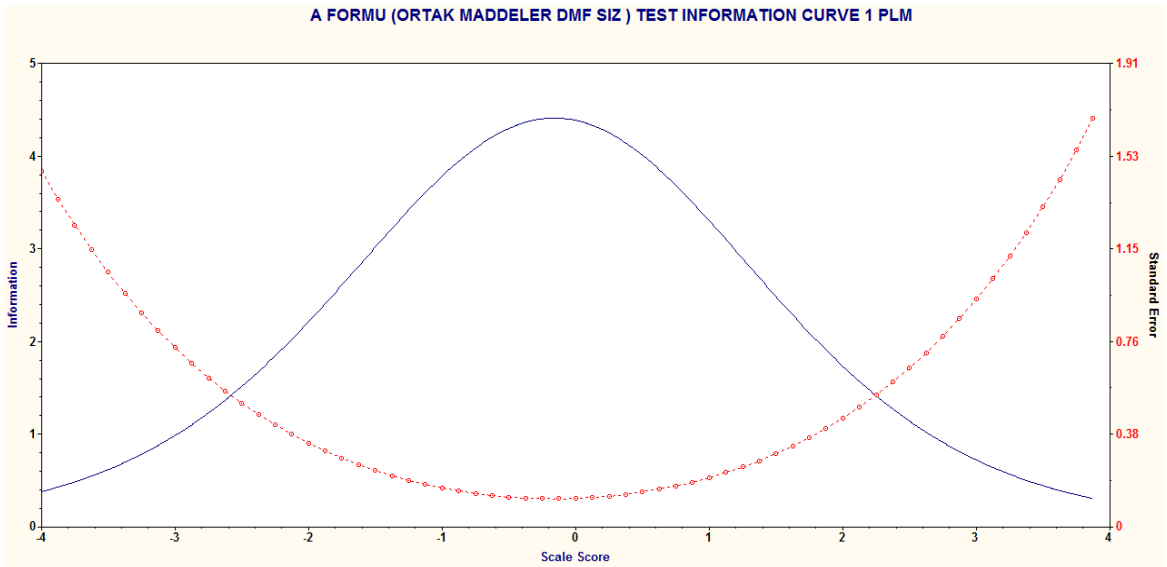
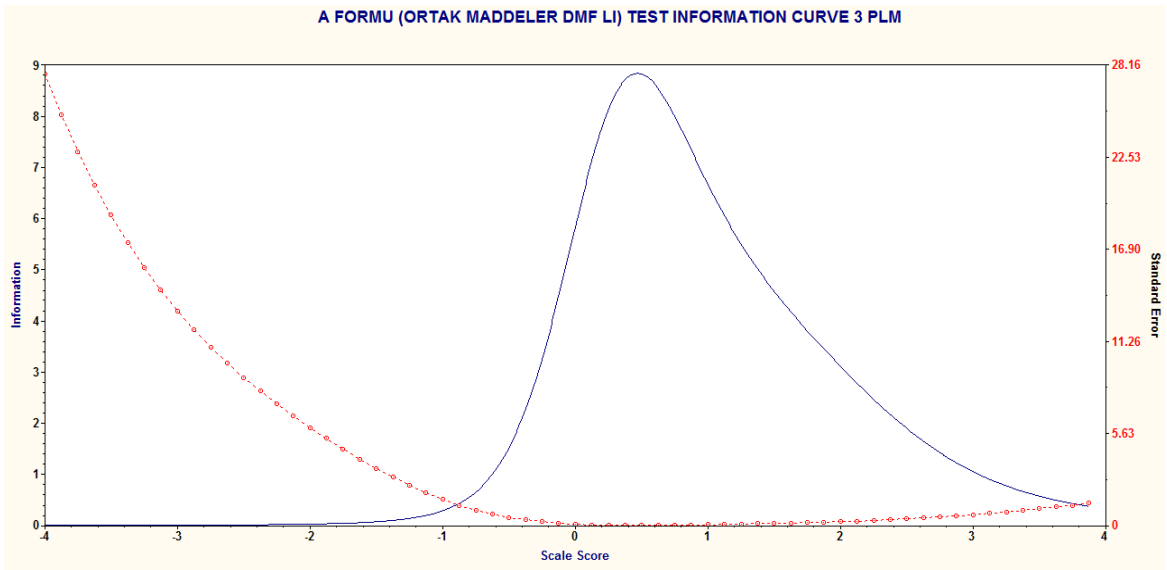


BDMFSIZ 3PLM MKE

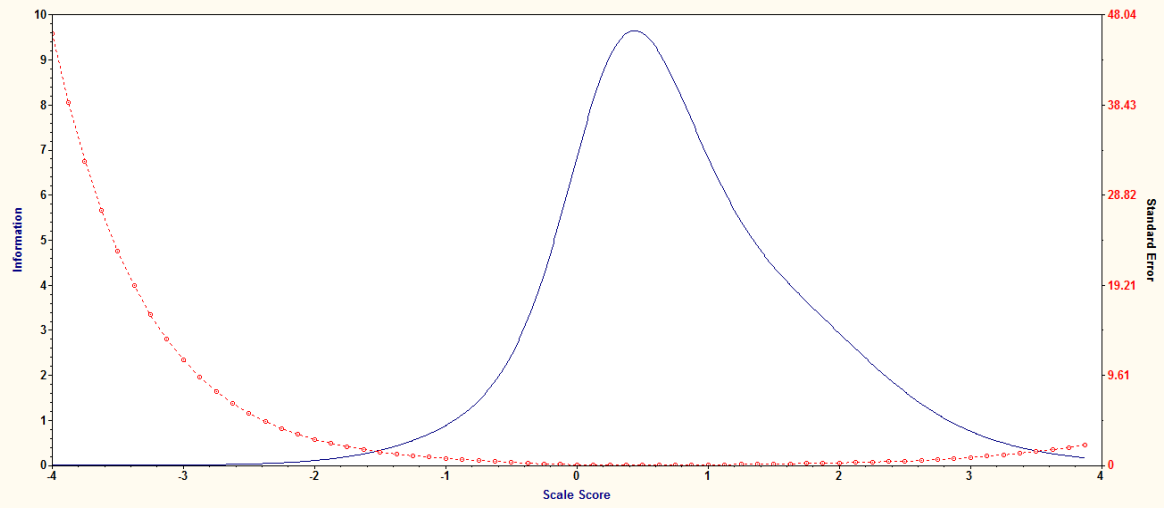


EK 5. TEST BİLGİ EĞRİLERİ

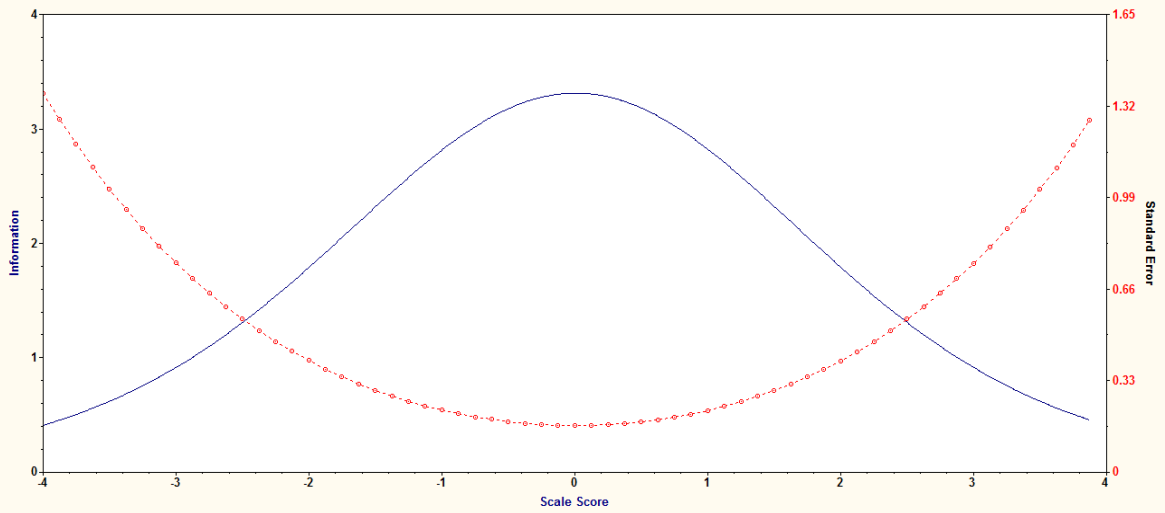




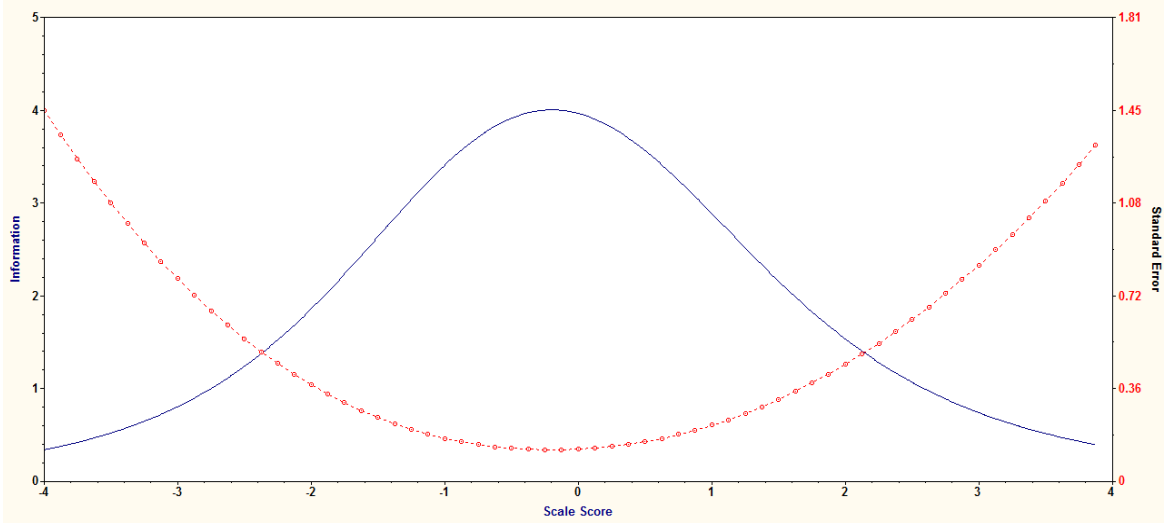
A FORMU (ORTAK MADDELER DMF SIZ) TEST INFORMATION CURVE 3 PLM



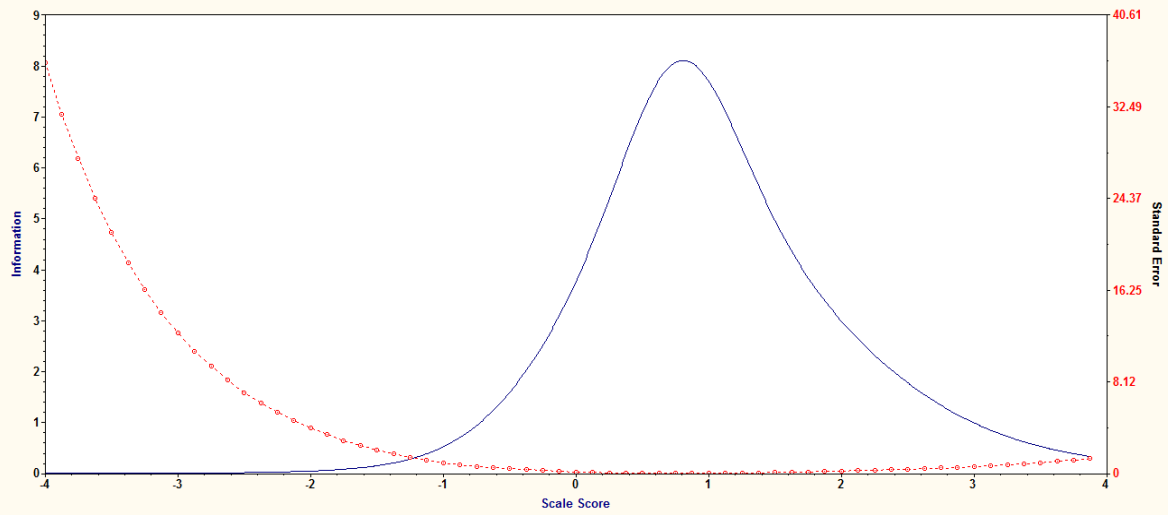
B FORMU (ORTAK MADDELER DMF LI) TEST INFORMATION CURVE 1PLM



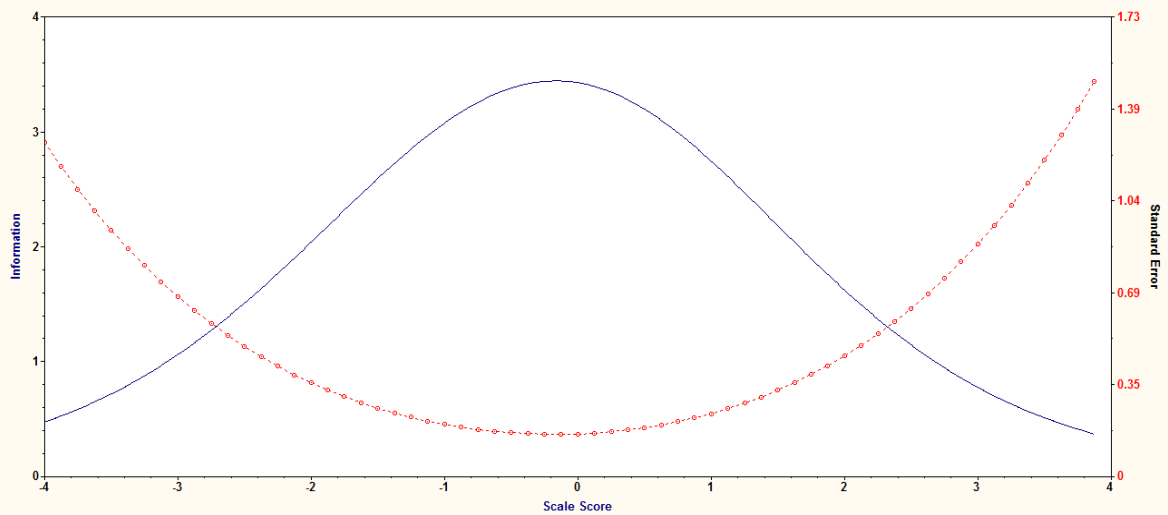
B FORMU (ORTAK MADDELER DMF LI) TEST INFORMATION CURVE 2 PLM



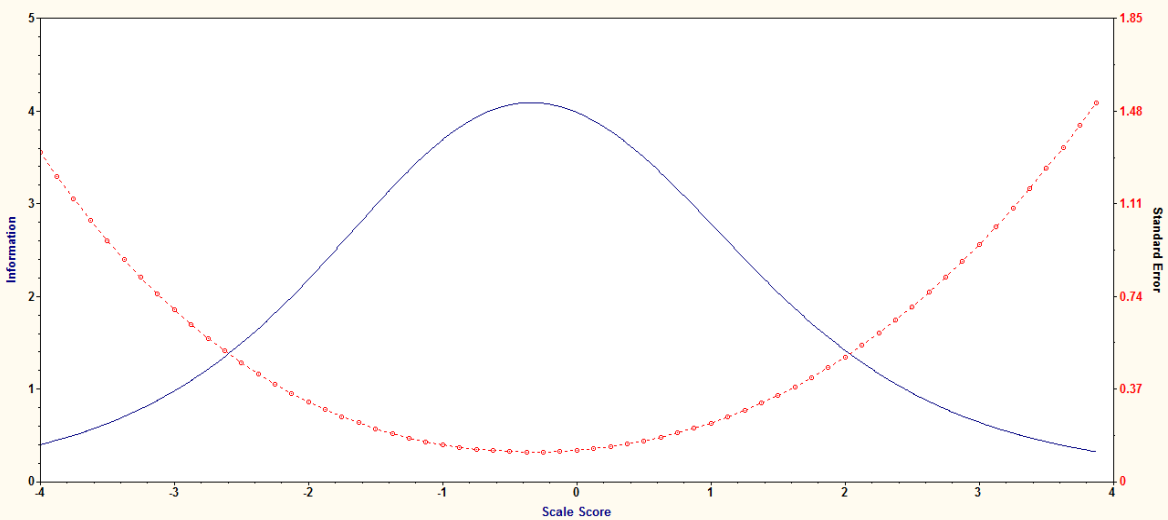
B FORMU (ORTAK MADDELER DMF Lİ) TEST INFORMATION CURVE 3 PLM



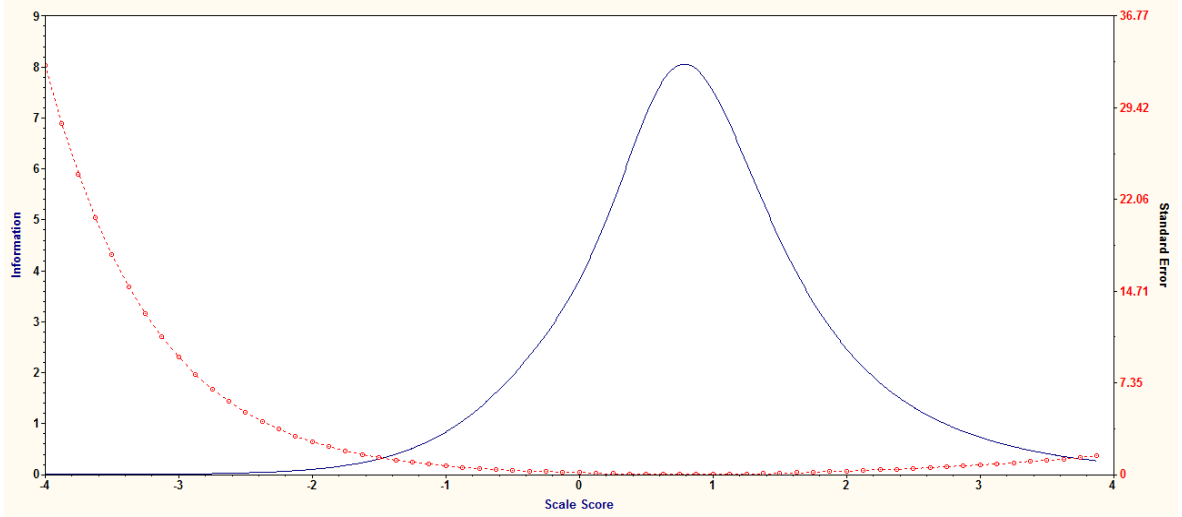
B FORMU (ORTAK MADDELER DMF SİZ) TEST INFORMATION CURVE 1PLM



B FORMU (ORTAK MADDELER DMF SİZ) TEST INFORMATION CURVE 2 PLM



B FORMU (ORTAK MADDELER DMF SIZ) TEST INFORMATION CURVE 3 PLM



**EK 6. EŞİTLENEN A VE B FORMLARINDA MADDELERİN ÖLÇÜTLERİ
KAZANIM VE BİLİMSEL SÜREÇ BECERİLERİ**

<i>FORM A</i>	<i>FORM B</i>	<i>KAZANIM VE BİLİMSEL SÜREÇ BECERİLERİ</i>
<i>Madde No</i>		
1	2	Bilimsel Süreç Becerileri/ Yorumlama ve Sonuç Çıkarma BSB.30. İşlenen verileri ve oluşturulan modeli yorumlar.
2	1	Bilimsel Süreç Becerileri/ Yorumlama ve Sonuç Çıkarma BSB.31. Elde edilen bulgulardan desen ve ilişkilere ulaşır.
3	4	8. Sınıf/ 2.Ünite/ 1.Konu: Kaldırma Kuvveti Kazanım1.9. Farklı yoğunluğa sahip sıvıların cisimlere uyguladığı kaldırma kuvvetini karşılaştırır ve sonuçları yorumlar.
4	23	Bilimsel Süreç Becerileri/ Verileri İşleme ve Model Oluşturma BSB.28. Deney ve gözlemlerden elde edilen verileri derleyip gözlem sıklığı dağılımı, çubuk grafik, tablo ve fiziksel modeller gibi farklı formlarda gösterir.
19	19	8. Sınıf/ 5. Ünite/ 5. Konu: Isınma- Soğuma Eğrileri Kazanım 6.2. Isınan-soğuyan maddelerin sıcaklık-zaman grafiklerini yorumlar.
20	20	8. Sınıf/ 5.Ünite/ 4. Konu: Erime-Donma-Buharlaştırma-Yoğuşma Isısı Kazanım 4.2. Farklı maddelerin erime/kaynama sıcaklıklarını karşılaştırır.
21	21	8. Sınıf/ 3. Ünite/ 1. Konu: Elementlerin Sınıflandırılması Kazanım: Elementlerin verilen özelliklerine göre grubunu/iyon yükünü belirtir.
22	22	8. Sınıf/ 3. Ünite/ 3. Konu: Kimyasal Tepkimeler Kazanım 3.6. Basit kimyasal tepkime denklemlerini sayma yöntemiyle denkleştirir Kazanım 3.2. Çok atomlu yaygın iyonların oluşturduğu bileşiklerin formüllerinde element atomlarının sayısını hesaplar.
23	24	Bilimsel Süreç Becerileri/ Yorumlama ve Sonuç Çıkarma BSB.30. İşlenen verileri ve oluşturulan modeli yorumlar.
24	3	8. Sınıf/ 1. Ünite/ 5. Konu: Adaptasyon ve Evrim Kazanım 5.1. Canlıların yaşadıkları çevreye adaptasyonunu örneklerle açıklar.
25	25	Bilimsel Süreç Becerileri/ Deney Tasarlama BSB. 16. Kurduğu hipotezi sınamaya yönelik bir deney önerir.
26	26	Bilimsel Süreç Becerileri/ Yorumlama ve Sonuç Çıkarma BSB.30. İşlenen verileri ve oluşturulan modeli yorumlar.
28	28	Bilimsel Süreç Becerileri/ Deney Düzenegi Kurma BSB.18. Verilen malzemeleri kullanarak kurduğu hipotezi sınamaya yönelik tasarladığı deneyi gerçekleştireceği bir düzenek kurar.

FEN VE TEKNOLOJİ TESTİ

Cinsiyeti: KIZ



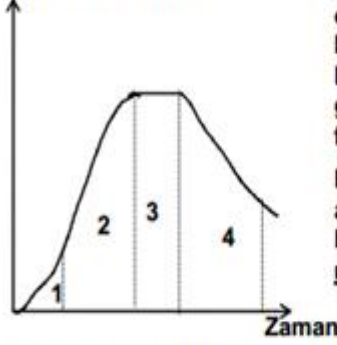
ERKEK



Lütfen tüm soruları DİKKATLİCE okuyup cevap veriniz.

SORULAR

1. Bakteri yoğunluğu

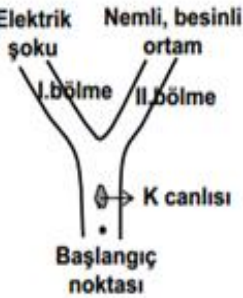


Yanda bir besin ortamında gelişen bakterilerin yoğunluğunun zamana göre değişim grafiği verilmiştir.

Bu grafiğe göre aşağıdakilerden hangisi söylenemez?

- A) 1. zaman diliminde ortamdaki besin en fazladır.
- B) 2. zaman diliminde bakteri çoğalma hızı en fazladır.
- C) 3. Zaman diliminde bakteri sayısı değişmemiştir.
- D) 4. zaman diliminde meydana gelen bakteri sayısı ölüden çoktur.

2.



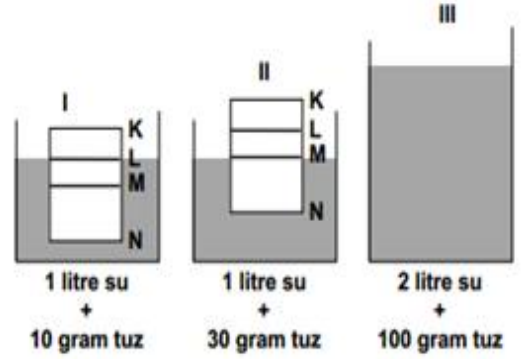
Bir bilim adamı "K canlısı öğrenme yeteneğine sahiptir." iddiasında bulunuyor. Bu iddiasının doğruluğunu ispatlamak için şekildeki düzenekle deney yapıyor.

K canlısı başlangıç noktasına konulduğunda I. bölmeye yönelmişken elektrik şokunu hissettiği için, kendine uygun olan nemli ve besinli II. bölmeye geçiyor.

Bilim adamının iddiasının doğru olabilmesi için, K canlısının tekrar tekrar başlangıç noktasına konulduğunda, hangi davranışı göstermesi beklenir?

- A) Önce II. sonra I. bölmeye geçmesi
- B) Önce I. sonra II. bölmeye geçmesi
- C) Sadece I. bölmeye geçmesi
- D) Sadece II. bölmeye geçmesi

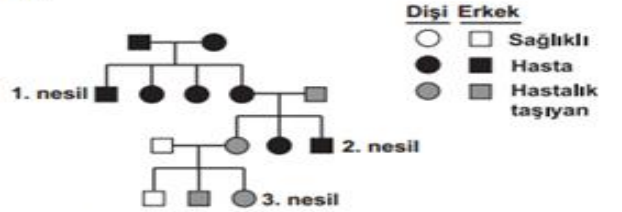
3.



I ve II nolu kaplarda bir tahta parçasının tuzlu su çözeltileri içindeki durumu verilmiştir. Aynı tahta parçası III. kaptaki çözeltiliye bırakıldığında su seviyesinin nerede olması beklenir?

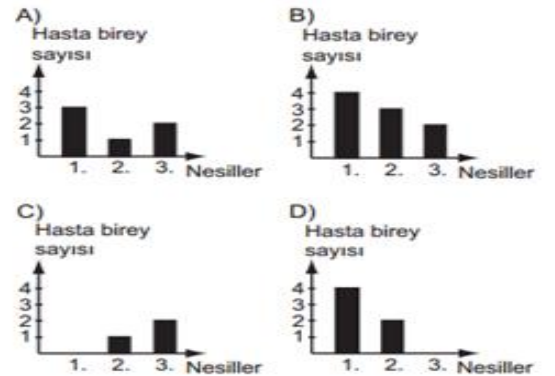
- A) M ve N arasında
- B) M ve L arasında
- C) K ve L arasında
- D) L hizasında

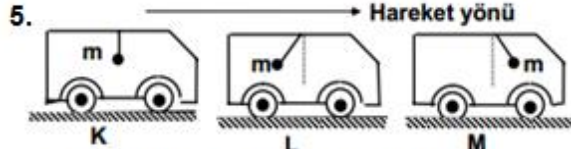
4.



Şekildeki soy ağacında 1. nesil, akraba olan bireylerin evliliğinden doğmuştur. 2. ve 3. nesil ise akraba olmayan bireylerin evliliğinden doğmuştur.

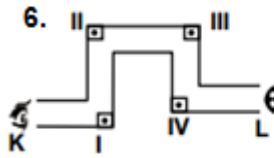
Bu soy ağacına göre, her bir nesildeki hasta birey sayısı hangi grafikte gösterilmiştir?



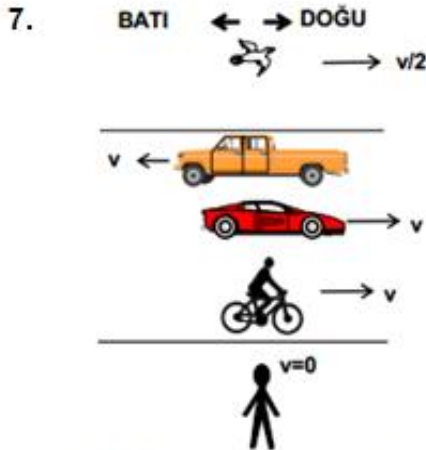
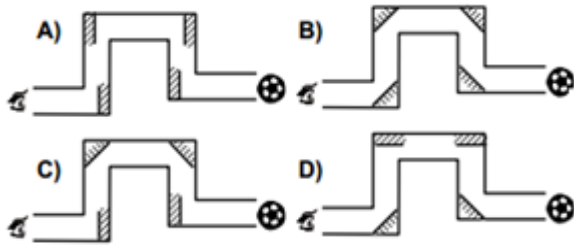


Şekilde K, L ve M araçlarının tavanlarına ipe asılan m kütleli cisimlerin bir anlık durumları görülmektedir. Araçların şekilde verilen ok yönündeki o anki hareket durumları için aşağıdakilerden hangisi söylenebilir?

	K	L	M
A)	Hareketsiz	Yavaşlamakta	Hızlanmakta
B)	Yavaşlamakta	Hızlanmakta	Yavaşlamakta
C)	Hareketsiz	Hızlanmakta	Yavaşlamakta
D)	Hızlanmakta	Yavaşlamakta	Hareketsiz

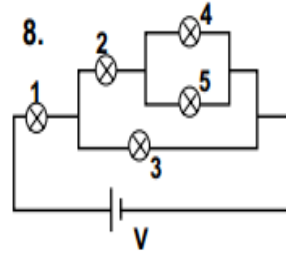


Şekildeki K noktasından bakan bir kişinin L deki topu görebilmesi için, düz aynalar I, II, III ve IV nolu bölgelere nasıl yerleştirilmelidir?



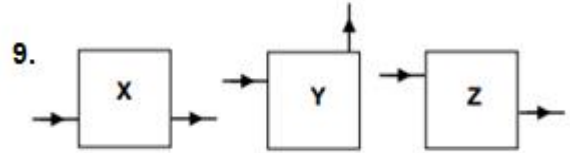
Şekildeki kuş, kamyon, otomobil ve bisikletli belirtilen yönlerde ve belirtilen hızlarla ilerlemektedirler. Çocuk ise yol kenarında durmaktadır. Buna göre, otomobilden bakan bir kişi için diğerlerinin hızlarıyla ilgili olarak aşağıdaki yargılardan hangisi yanlıştır?

- A) Kuş $3v/2$ lik hızla doğuya doğru uçmaktadır.
 B) Kamyon $2v$ hızıyla batıya doğru gitmektedir.
 C) Bisikletli durmaktadır.
 D) Yol kenarındaki çocuk batıya doğru v hızıyla gitmektedir.

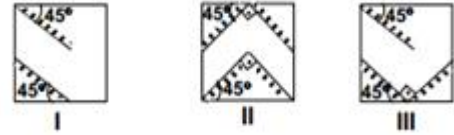


Şekilde verilen elektrik devresindeki eşdeğer ampullerden en az ışık veren iki ampul hangileridir?

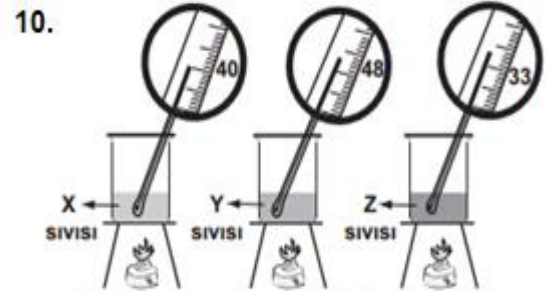
- A) 1 ve 3 B) 2 ve 3 C) 3 ve 4 D) 4 ve 5



Camdan yapılmış X, Y ve Z kutularına giren ışınların çıkışta belirtilen yolları izleyebilmeleri için düz aynalardan oluşan I, II ve III nolu sistemler hangi kutuların içinde yer almalıdır?



	X	Y	Z
A)	II	III	I
B)	I	II	III
C)	III	I	II
D)	II	I	III



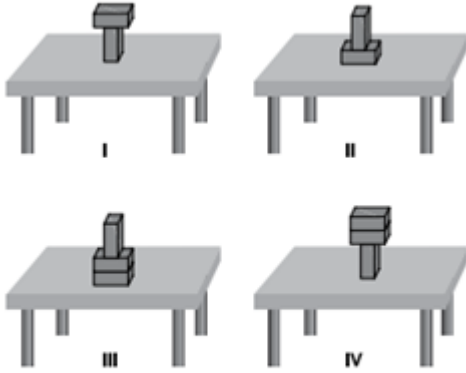
Başlangıçta sıcaklıkları 10°C olan aynı miktardaki X, Y ve Z sıvıları, eşit sürede ısıtıldığında sıvıların sıcaklıkları termometrelerdeki gibi gözleniyor.

Y ve Z'nin öz ısılan çizelgede verildiğine göre, X'in öz ısısı aşağıdakilerden hangisi olabilir? (Kaplar ve ısıtıcılar özdeşdir.)

Sıvı	Öz ısı ($\text{J/g}^{\circ}\text{C}$)
X	?
Y	0,63
Z	1,04

- A) 0,41 B) 0,79 C) 1,26 D) 1,67

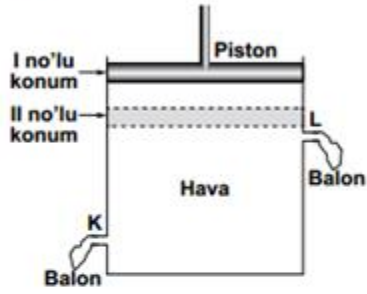
11. Özdeş tuğlalar yatay bir masa üzerine dört farklı şekilde konulmuştur.



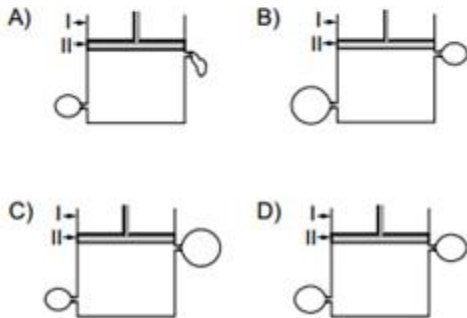
Buna göre, hangi şekildeki tuğlaların masaya uyguladığı basınç en büyüktür?

- A) I. B) II. C) III. D) IV.

12. İçi hava dolu pistonlu küçük bir kabın K ve L çıkışlarına şekildeki gibi özdeş balonlar takılıyor. Daha sonra piston, aşağı bastırılarak, I no'lu konumdan II no'lu konuma getiriliyor.



Buna göre, balonların şişme durumları aşağıdakilerden hangisindeki gibi olabilir?



- 13.

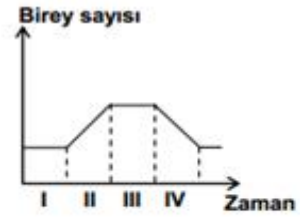


Serpil, birinde nemli, diğerinde kurutulmuş ekme bulunan iki naylon torbayı gevşek bir şekilde bağlıyor. Beş gün sonra sadece nemli ekme üzerinde küf mantarının oluştuğunu gözlüyor.

Serpil, bu gözleme göre aşağıdaki sonuçlardan hangisine ulaşır?

- A) Hava alan tüm yiyecekler çabuk bozulur.
B) Kuru ortamlar, mantar türleri için uygundur.
C) Kurutulmuş ekme daha uzun süre dayanır.
D) Küf mantarı bölünerek çoğalır.

- 14.

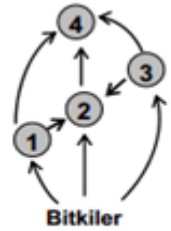


Yukarıdaki grafik, uygun bir ortama bırakılan bakteri popülasyonunun birey sayısındaki değişimi göstermektedir.

Bu grafiğe göre aşağıdakilerden hangisi söylenemez?

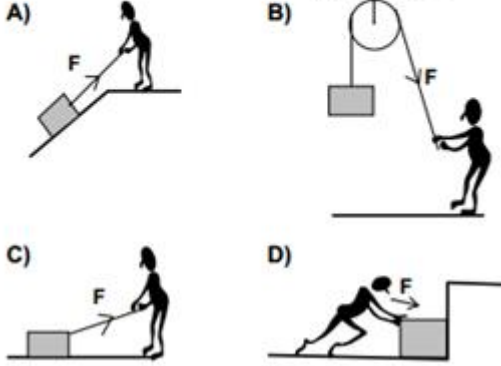
- A) IV. aralıkta ölen bakterilerin sayısı yeni oluşanlardan azdır.
B) III. aralıkta bakteri sayısı sabit kalmıştır.
C) II. aralıkta bölünme hızı kazanmıştır.
D) I. aralıkta henüz bölünme başlamamıştır.

15. Hem bitki hem de hayvan yiyen bir canlı, yandaki besin ağında kaç numaralı kısımda yer alabilir?



- A) 1 B) 2 C) 3 D) 4

16. Fiziksel anlamda iş yapılabilmesi için;
 • Kuvvet uygulanmalı,
 • Kuvvet etkisindeki cisim yol almalıdır.
 Buna göre aşağıdakilerden hangisinde kesinlikle iş yapılamaz?



17. Aşağıdaki olayların hangisi basıncın etkisiyle oluşmaz?

- A) Yükseklere çıkan kişilerde kulak ağrısının hissedilmesi
 B) Yükseklere çıkıldıkça suyun kaynama sıcaklığının azalması
 C) Açılan soda şişesinde hava kabarcıklarının oluşması
 D) Bir cismin ekvatordaki ağırlığı ile kutuplardaki ağırlığının farklı olması

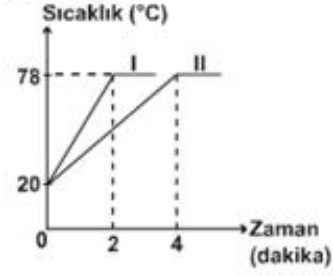
18.



Tahtadaki iyonların oluşturduğu bileşiğin formülünü yazmaya çalışan öğrencinin, zihninden geçirdiği aşağıdaki düşüncelerden hangisi yanlıştır?

- A) Katyon önce, anyon sonra yazılmalı
 B) Bileşik nötr yapıda olmalı
 C) Bileşikte pozitif ve negatif yükler birbirini dengelemeli
 D) Bileşik formülünde 2 Ca ve 1 Cl atomu olmalı

19. Özdeş I ve II kaplarında bulunan aynı sıvılar özdeş ısıtıcılarla ısıtılıyor ve aşağıdaki grafik elde ediliyor.



Grafığe göre, aşağıdakilerden hangisi söylenemez?

- A) Kaplardaki sıvıların miktarları farklıdır.
 B) I. kaptaki sıvıya daha az ısı verilmiştir.
 C) II. kaptaki sıvı daha uzun süre ısıtılmıştır.
 D) Kaplardaki sıvılara aynı miktarlarda ısı verilmiştir.

20.

Madde	Erime sıcaklığı (°C)	Kaynama sıcaklığı (°C)
X	114	186
Y	-98	65
Z	6	79
T	-111	-60

Çizelgeye göre, X, Y, Z ve T maddelerinin 1 atm basınç ve oda sıcaklığındaki (25 °C) fiziksel hâli aşağıdakilerin hangisinde doğru verilmiştir?

	X	Y	Z	T
A)	Sıvı	Gaz	Gaz	Sıvı
B)	Sıvı	Gaz	Sıvı	Gaz
C)	Katı	Sıvı	Gaz	Sıvı
D)	Katı	Sıvı	Sıvı	Gaz

21.

Atom	Elektron dizilişi		
X	2	8	2
Y	2	8	8
Z	2	8	8
T	2	8	8

"Yüksüz bir atomun en dış enerji düzeyindeki (katman) elektron sayısı, o atomun periyodik çizelgedeki grubunu belirler."

Buna göre çizelgede elektron dizilişi verilen atomlardan hangileri aynı grupta yer alır?

- A) Y ve Z B) X ve Z C) Y ve T D) X ve T

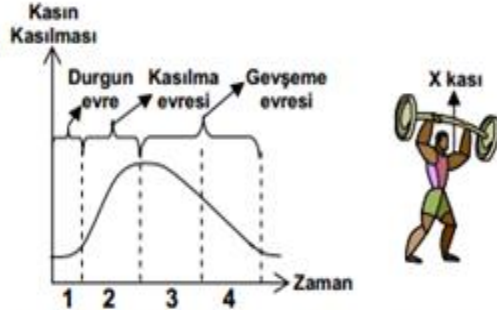
22.



Tanecik modeli verilen tepkimenin denkleşmesi için ürünler tarafına aşağıdakilerden hangisi eklenmelidir?

- A) B)
C) D)

23.



Yukarıdaki grafikte bir kasın durgun, kasılma ve gevşeme evreleri verilmiştir.

Buna göre şekildeki sporcunun X kasının durumu, grafikte hangi zaman aralığındaki gibidir?

- A) 1 B) 2 C) 3 D) 4

24. Sıcakkanlı canlılar, çeşitli davranış şekilleriyle vücut sıcaklıklarını dengede tutmaya çalışır. Sıcak havada dışarıya fazla ısı verecekleri ortam ararlar. Soğuk havada ise vücut yüzeyini azaltıp hareketlerini artırır.

Buna göre sıcakkanlı canlılar, sıcak ve soğuk havada aşağıdakilerin hangisinde verilen davranışları gösterirler?

Sıcak havada

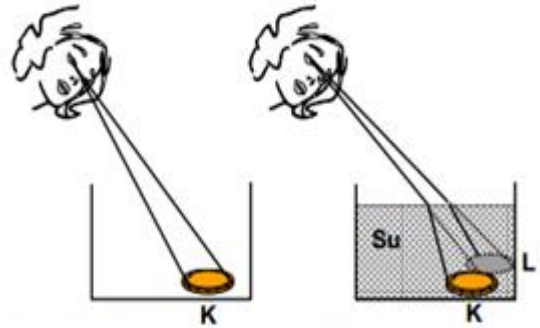
Soğuk havada

- | | |
|--------------------------|-----------------------|
| A) Sıcak yer arama | Hareketlerini azaltma |
| B) Hareketlerini azaltma | Serin yer arama |
| C) Hareketlerini artırma | Kıvrılıp büzülme |
| D) Serin yer arama | Kıvrılıp büzülme |

25. Bir öğrenci sıcaklığın çözünürlüğü etkileyen bir faktör olduğunu aşağıdaki deneylerden hangisi ile gösterebilir?

- A) 100 ml seyreltik şeker çözeltisinin sıcaklığını 25 °C den 50 °C ye çıkararak.
B) 10 gram şekeri 100 ml kaynar suda çözerek.
C) Oda sıcaklığında doyma noktasına yakın, henüz tam doymamış 100 ml şeker çözeltisini soğutarak.
D) Tuz örneğini 25 °C de, şeker örneğini 50 °C de, 100 ml suda çözerek.

26.



Boş kaptaki paraya bakan çocuk parayı bulunduğu K noktasında görür. Kaba su doldurulduğunda ise çocuk parayı L noktasında görmüş gibi görür. Çocuk bu deneyden sonra aşağıdaki yorumlardan hangisini yapabilir?

- A) Su içindeki balıklar birbirlerini buldukları yerden daha yakında görürler.
B) Su altından gökyüzüne bakan bir dalgıç uçakları buldukları yerden daha yüksekte görür.
C) İnsanlar suyun dışında zıplayan yunusları buldukları yerden daha uzakta görür.
D) Çukur aynayla elde edilen görüntü cismin aynaya uzaklığından daha yakındır.

27. Betül ve Sevgi'nin arasında geçen musluk suyu ile ilgili konuşma aşağıdaki gibidir.

Musluğumuzdan akan suyun içimi hiç hoş değil.

Sanırım suyunuz sert. Suyun sertliği giderildiğinde içimi hoş olacaktır. Bunun için ... işlemi yapılmalı.



Betül

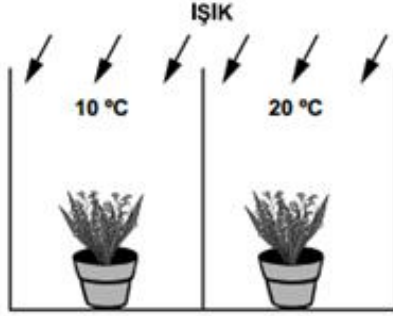


Sevgi

Sevgi'nin önerdiği işlem aşağıdakilerden hangisidir?

- A) Kaynatma B) Klorlama
C) Dinlendirme D) Süzme

28. Bir öğrenci, farklı sıcaklıkların fotosentez olayına etkisini gözlemek için özdeş bitkilerle şekildeki düzeneği hazırlıyor.



Bitkilere eşit miktarda düzenli olarak su veriliyor. Saksılar ve içindeki topraklar özdeştir.

Eğer bu öğrenci "Karanlıktaki bitki, ışık alan bitki gibi fotosentez yapar mı?" sorusuna cevap ararsa, bu düzenekte hangi değişikliği yapmalıdır?

- A) Bitkileri birbiriyle yer değiştirmelidir.
B) 10 °C'deki ortamın sıcaklığını azaltıp buradaki bitkinin ışık ve hava almasını engellemelidir.
C) Her iki ortamın sıcaklığını 10 °C'ye ayarlayıp her iki bitkiyi de ışıkta bırakmalıdır.
D) Her iki ortamın sıcaklığını 20 °C'ye ayarlayıp bitkilerden birinin ışık almasını engellemelidir.

29. Oğuzcan, Fen ve Teknoloji dersinde defterine sigortalarla ilgili aşağıdaki bilgileri not etmiştir:

- 1- Metal çiftli sigortalarda, devreden aşırı akım geçmesi sonucu metal çift ısınarak bükülür ve elektrik akımı kesilir.
- 2- Bazı sigortalar, elektrik akımını iyi ileten ama erime sıcaklığı düşük olan metallerden yapılır. Devreden aşırı akım geçtiğinde, tel hemen eriyip kopar ve elektrik akımı geçmez.
- 3- Binalardaki elektrik tesisatında tellerin aşırı ısınarak yanmasını engellemek için sigorta kullanılır.
- 4- Şehir elektriğinin binalara girdiği yerlerde sigortalar elektrik sayacı ile aynı panoda bulunabilir.

Buna göre, Oğuzcan'ın not ettiği bilgilerden hangisi sigortanın çalışma prensipleri veya güvenlik açısından önemi ile ilgili değildir?

- A) 1. B) 2. C) 3. D) 4.

TEST BİTTİ

FEN VE TEKNOLOJİ TESTİ

Cinsiyeti: KIZ



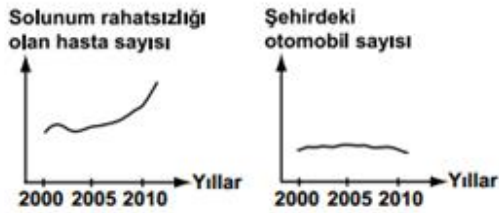
ERKEK

Lütfen tüm soruları DİKKATLİCE okuyup cevap veriniz.

SORULAR

1. Hastane kayıtlarına göre; bir şehirde hava kirliliği artışına bağlı olarak solunum rahatsızlığı olanların sayısı da artmıştır.

Bir araştırmacı, "Bu şehirde hava kirliliğine otomobil sayısındaki artışın neden olabileceği" tahmininde bulunmuştur.



Daha sonra bu konu ile ilgili olarak yukarıdaki grafikleri inceleyen araştırmacı hangi yorumu yaparsa doğru olur?

- A) Tahminim doğru, çünkü otomobil sayısı sürekli artmış.
 B) Tahminim doğru, toplu taşıma araçlarının kullanılması için çalışmalarda bulunmalıyım.
 C) Tahminim yanlış, havayı kirleten başka sebepleri araştırmalıyım.
 D) Tahminim yanlış, çünkü hasta sayısı sürekli azalmış.

2.

Toprak Derinliği (cm)	Toprak Sıcaklığı°C°	
	En Yüksek	En Düşük
0	53	-19
5	46	-5
30	31	-2
110	26	4

Belli bir bölgede toprağın çeşitli derinliklerindeki sıcaklıklar bir yıl süreyle ölçülmüştür. Derinliğe bağlı olarak ölçülen sıcaklıkların en düşük ve en yüksek değerleri tabloda verilmiştir.

Buna göre, aşağıdaki yorumlardan hangisine ulaşamaz?

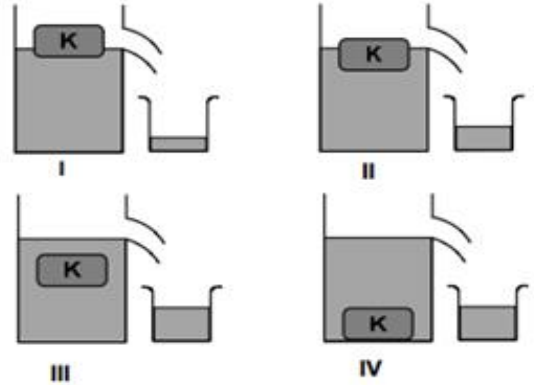
- A) Toprak derinliği arttıkça en yüksek ve en düşük sıcaklık değerleri birbirlerine yaklaşır.
 B) Toprak derinliği arttıkça en yüksek sıcaklık değerinde azalma görülür.
 C) Toprak derinliğine bağlı olarak en düşük ve en yüksek sıcaklık değerleri azalır.
 D) Toprak derinliğinin artması ile en düşük sıcaklık değerlerinde yükselme görülür.

3. Bir cins arı, yüksek sıcaklıkta büyütülürse açık renkli, gelişebileceği en düşük sıcaklıkta büyütülürse siyah renkli olur.

Aşağıdakilerden hangisi bu olayla benzerlik göstermez?

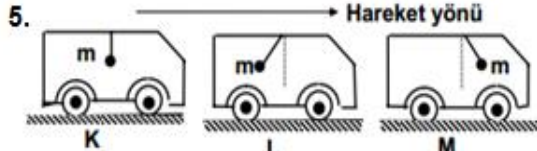
- A) Çuha çiçeği bitkisinin 15-20°C'de kırmızı çiçek açarken 30-35°C'de beyaz çiçek açması
 B) Kuzey Kutbu'na yakın bölgede yaşayan tavşanların kışın ve yazın farklı renklerde olması
 C) Afrika'da yaşayan insanların ten renginin siyah, Avrupa'da yaşayan insanların ten renginin açık olması
 D) Sirkesineklerinin 25°C'de tutulan larvalarından kıvrık kanatlı yavruların, 16°C'de tutulan larvalarından düz kanatlı yavruların ortaya çıkması

4. K cisimi, taşma seviyesine kadar doldurulmuş kaplar içerisinde farklı sıvılara bırakılıyor. K cisminin sıvılar içerisindeki denge konumları şekildeki gibi oluyor.



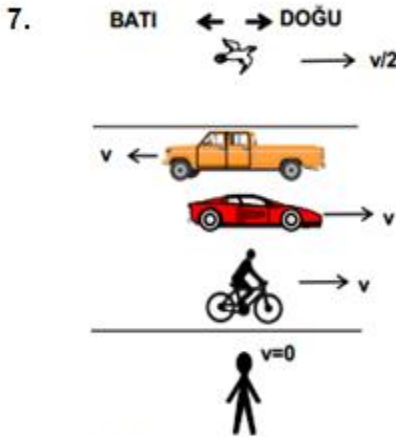
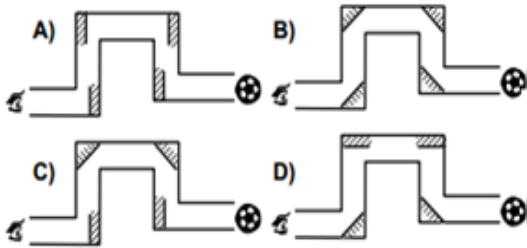
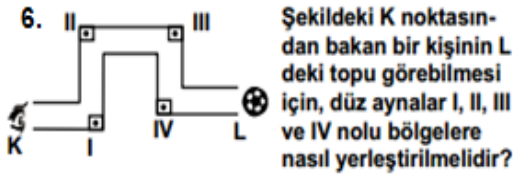
Buna göre, hangi kaptan taşan sıvının kütlesi en azdır? (K sıvılar içinde çözünmemektedir.)

- A) I B) II C) III D) IV



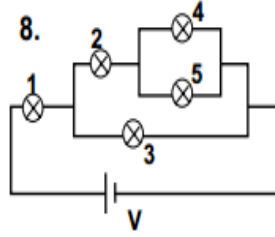
Şekilde K, L ve M araçlarının tavanlarına ipe asılan m kütleli cisimlerin bir anlık durumları görülmektedir. Araçların şekilde verilen ok yönündeki o anki hareket durumları için aşağıdakilerden hangisi söylenebilir?

- | | K | L | M |
|----|--------------|--------------|--------------|
| A) | Hareketsiz | Yavaşlamakta | Hızlanmakta |
| B) | Yavaşlamakta | Hızlanmakta | Yavaşlamakta |
| C) | Hareketsiz | Hızlanmakta | Yavaşlamakta |
| D) | Hızlanmakta | Yavaşlamakta | Hareketsiz |



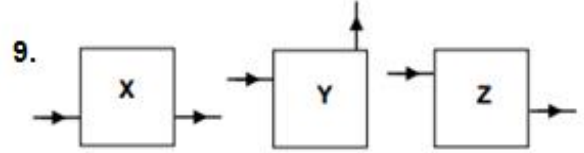
Şekildeki kuş, kamyon, otomobil ve bisikletli belirtilen yönlerde ve belirtilen hızlarla ilerlemektedirler. Çocuk ise yol kenarında durmaktadır. Buna göre, otomobilden bakan bir kişi için diğerlerinin hızlarıyla ilgili olarak aşağıdaki yargılardan hangisi yanlıştır?

- A) Kuş $3v/2$ lik hızla doğuya doğru uçmaktadır.
 B) Kamyon $2v$ hızıyla batıya doğru gitmektedir.
 C) Bisikletli durmaktadır.
 D) Yol kenarındaki çocuk batıya doğru v hızıyla gitmektedir.

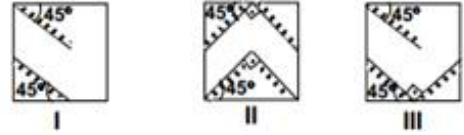


Şekilde verilen elektrik devresindeki eşdeğer ampullerden en az ışık veren iki ampul hangileridir?

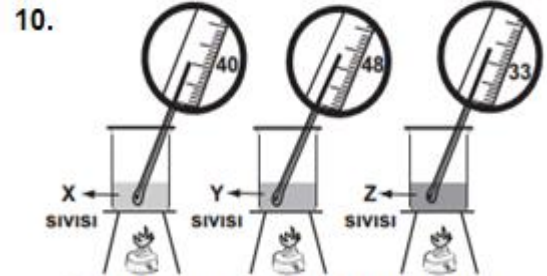
- A) 1 ve 3 B) 2 ve 3 C) 3 ve 4 D) 4 ve 5



Camdan yapılmış X, Y ve Z kutularına giren ışınların çıkışta belirtilen yolları izleyebilmeleri için düz aynalardan oluşan I, II ve III nolu sistemler hangi kutuların içinde yer almalıdır?



- | | X | Y | Z |
|----|-----|-----|-----|
| A) | II | III | I |
| B) | I | II | III |
| C) | III | I | II |
| D) | II | I | III |



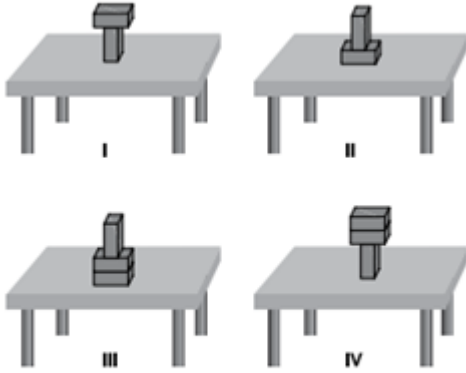
Başlangıçta sıcaklıkları 10°C olan aynı miktardaki X, Y ve Z sıvıları, eşit sürede ısıtıldığında sıvıların sıcaklıkları termometrelerdeki gibi gözleniyor.

Y ve Z'nin öz ısıları çizelgede verildiğine göre, X'in öz ısısı aşağıdakilerden hangisi olabilir? (Kaplar ve ısıtıcılar özdeştir.)

Sıvı	Öz ısı ($\text{J/g}^{\circ}\text{C}$)
X	?
Y	0,63
Z	1,04

- A) 0,41 B) 0,79 C) 1,26 D) 1,67

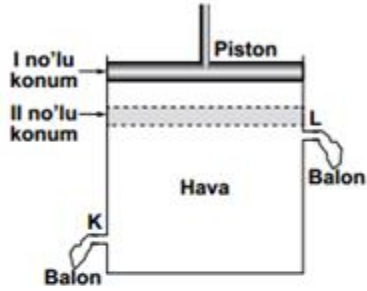
11. Özdeş tuğlalar yatay bir masa üzerine dört farklı şekilde konulmuştur.



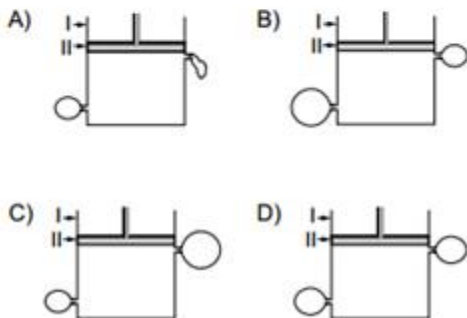
Buna göre, hangi şekildeki tuğlaların masaya uyguladığı basınç en büyüktür?

- A) I. B) II. C) III. D) IV.

12. İçi hava dolu pistonlu küçük bir kabın K ve L çıkışlarına şekildeki gibi özdeş balonlar takılıyor. Daha sonra piston, aşağı bastırılarak, I no'lu konumdan II no'lu konuma getiriliyor.



Buna göre, balonların şişme durumları aşağıdakilerden hangisindeki gibi olabilir?



- 13.

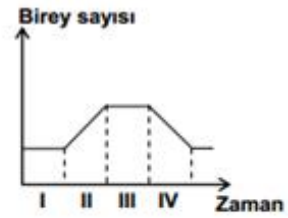


Serpil, birinde nemli, diğerinde kurutulmuş ekmek bulunan iki naylon torbayı gevşek bir şekilde bağlıyor. Beş gün sonra sadece nemli ekmek üzerinde küf mantarının oluştuğunu gözlüyor.

Serpil, bu gözleme göre aşağıdaki sonuçlardan hangisine ulaşır?

- A) Hava alan tüm yiyecekler çabuk bozulur.
B) Kuru ortamlar, mantar türleri için uygundur.
C) Kurutulmuş ekmek daha uzun süre dayanır.
D) Küf mantarı bölünerek çoğalır.

- 14.

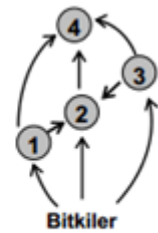


Yukarıdaki grafik, uygun bir ortama bırakılan bakteri popülasyonunun birey sayısındaki değişimi göstermektedir.

Bu grafiğe göre aşağıdakilerden hangisi söylenemez?

- A) IV. aralıkta ölen bakterilerin sayısı yeni oluşanlardan azdır.
B) III. aralıkta bakteri sayısı sabit kalmıştır.
C) II. aralıkta bölünme hızı kazanmıştır.
D) I. aralıkta henüz bölünme başlamamıştır.

15. Hem bitki hem de hayvan yiyen bir canlı, yandaki besin ağında kaç numaralı kısımda yer alabilir?

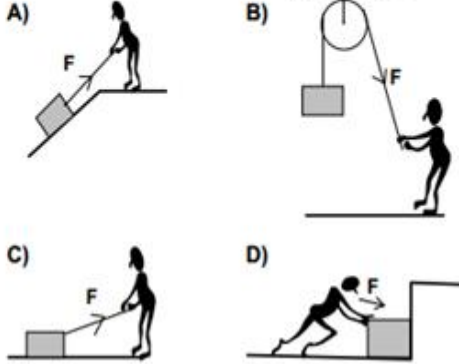


- A) 1 B) 2 C) 3 D) 4

16. Fiziksel anlamda iş yapılabilmesi için;

- Kuvvet uygulanmalı,
- Kuvvet etkisindeki cisim yol almalıdır.

Buna göre aşağıdakilerden hangisinde kesinlikle iş yapılamaz?



17. Aşağıdaki olayların hangisi basıncın etkisiyle oluşmaz?

- A) Yükseklere çıkan kişilerde kulak ağrısının hissedilmesi
 B) Yükseklere çıkıldıkça suyun kaynama sıcaklığının azalması
 C) Açılan soda şişesinde hava kabarcıklarının oluşması
 D) Bir cismin ekvatordaki ağırlığı ile kutuplardaki ağırlığının farklı olması

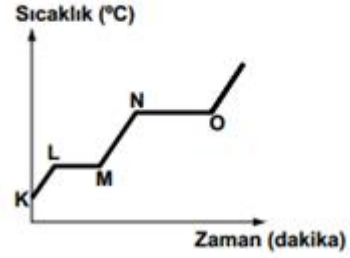
18.



Tahtadaki iyonların oluşturduğu bileşiğin formülünü yazmaya çalışan öğrencinin, zihninden geçirdiği aşağıdaki düşüncelerden hangisi yanlıştır?

- A) Katyon önce, anyon sonra yazılmalı
 B) Bileşik nötr yapıda olmalı
 C) Bileşikte pozitif ve negatif yükler birbirini dengelemeli
 D) Bileşik formülünde 2 Ca ve 1 Cl atomu olmalı

19. Saf bir maddenin ısıtılmasına ait sıcaklık - zaman grafiği aşağıda verilmiştir:



Buna göre, aşağıdakilerden hangisi doğrudur?

- A) Madde K noktasında erimeye başlamıştır.
 B) Madde L - M aralığında tamamen katı hâldedir.
 C) Madde N noktasında kaynamaya başlamıştır.
 D) Madde O noktasında tamamen sıvı hâldedir.

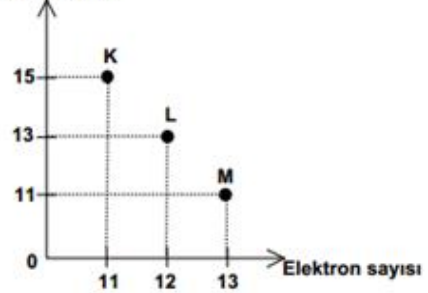
20.

Madde	Erime noktası (°C)	Kaynama noktası (°C)
K	850	1140
L	-39	357
M	-218	-183

Tabloda K, L ve M maddelerinin erime ve kaynama sıcaklıkları verilmektedir. Bu maddelerin 25°C deki hal durumları nedir?

- | | <u>K</u> | <u>L</u> | <u>M</u> |
|----|----------|----------|----------|
| A) | Sıvı | Katı | Gaz |
| B) | Katı | Sıvı | Gaz |
| C) | Gaz | Sıvı | Katı |
| D) | Katı | Gaz | Sıvı |

21. Proton sayısı



Bir atomdan oluşan iyonun yükü, proton sayısı ile elektron sayısının farkına eşittir.

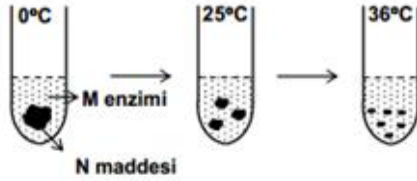
K, L ve M iyonlarının proton ile elektron sayısı arasındaki ilişki grafikteki gibi olduğuna göre, bu iyonların yükleri aşağıdakilerden hangisidir?

- | | <u>K</u> | <u>L</u> | <u>M</u> |
|----|----------|----------|----------|
| A) | +4 | +1 | -2 |
| B) | -4 | -1 | +2 |
| C) | -4 | -1 | -2 |
| D) | +4 | +1 | +2 |

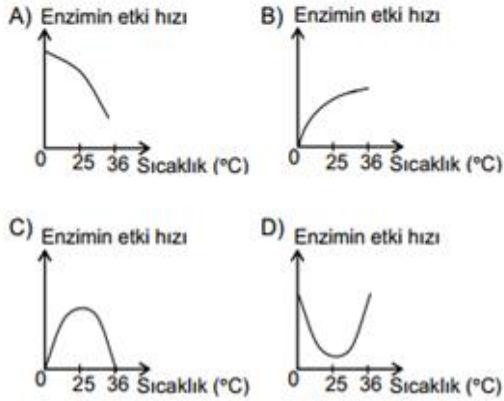
22. KHCO_3 , Na_2CO_3 , AlPO_4 bileşik formülleri için aşağıdakilerden hangisi doğrudur?

- A) Toplam atom sayıları aynıdır.
 B) Anyonlarının yükleri birbirine eşittir.
 C) Her formülde aynı sayıda katyon vardır.
 D) Her formülde aynı sayıda element vardır.

23.



Bir öğrencinin, sıcaklık arttıkça M enziminin, N maddesine etkisini gözlemek için yaptığı deney aşamaları yukarıda verilmiştir. Öğrenci, bu durumu hangi grafikte gösterebilir?



24.



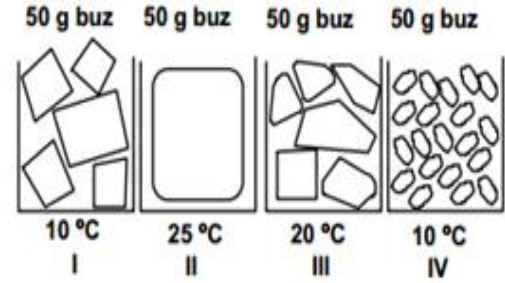
Yukarıdaki grafik çizilirken aşağıdaki bilgilerden hangileri kullanılmıştır?

- I- Bitkide terleme miktarı arttıkça, bitkinin topraktan aldığı su miktarı artar.
 II- Havanın sıcak olduğu öğle saatlerinde bitkide terleme en fazladır.
 III- Bitkinin ışık almadığı zamanlarda terleme en alt düzeydedir.

- A) I - II
 B) I - III
 C) II - III
 D) I - II - III

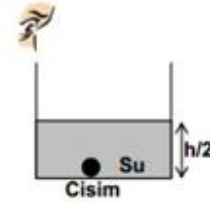
25. Bir araştırmacı, buzların şekli ile erime süreleri arasındaki ilişkiyi gözlemek istiyor.

Bunun için içerisinde değişik şekillerde buz parçaları bulunan aşağıdaki deney düzeneklerinden hangilerini kullanmalıdır?



- A) I ile IV B) I ile III C) II ile III D) II ile IV

26.



Yarıya kadar su ile dolu olan kap içindeki cismi gören gözlemci, kap tamamen su ile doldurursa cisimle ilgili aşağıdaki durumlardan hangisi ile karşılaşabilir?

- A) Cismin konumunda değişiklik gözlemez.
 B) Cismi, kendinden daha uzakta görür.
 C) Cismi göremez.
 D) Cismi, kendine daha yakın görür.

27. Kartlarda günlük yaşamda kullanılan bazı maddeler yer almaktadır.



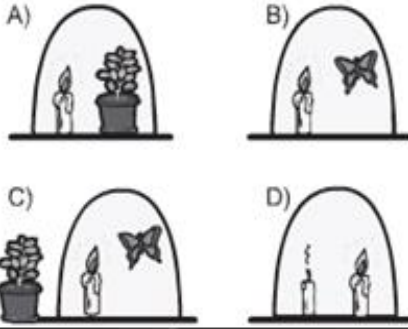
Buna göre, hangi kartlardaki maddeleri oluşturan elementler metallere, hangileri ametallere örnektir?

Metaller	Ametaller
A) 1, 2, 3	4
B) 4	1, 2, 3
C) 2, 4	1, 3
D) 1, 3	2, 4

28. Ahmet, sınıfındaki bir etkinlikte fotosentezin önemini anlatmaktadır.



Şekil-I ve şekil-II'deki gibi konusunu anlatan Ahmet'in şekil-II'deki düzeneği hangisidir?



29. Işık veren ampulün harcadığı elektrik enerjisi miktarı aşağıdakilerden hangilerine bağlıdır?

- I- Ampulün aydınlattığı alanın büyüklüğüne
II- Ampulün gücüne
III- Ampulün ışık verme süresine

- A) Yalnız II
B) I - II
C) II - III
D) I - II - III

TEST BİTTİ

EK 7. ARAŞTIRMA İZİN TALEBİ



T.C.
MİLLÎ EĞİTİM BAKANLIĞI
Temel Eğitim Genel Müdürlüğü

Sayı : 70297673/100/1057998
Konu: Araştırma İzin Talebi

22/05/2013

HACETTEPE ÜNİVERSİTESİ REKTÖRLÜĞÜNE
(Sosyal Bilimler Enstitüsü Müdürlüğü)
ANKARA

İlgi : Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Müdürlüğü'nün 20.05.2013 tarihli ve 12908312-900-2442 sayılı yazısı.

Enstitünüz Eğitim Bilimleri Ana Bilim Dalı Eğitim Ölçme ve Değerlendirme Bilim Dalı Doktora Programı öğrencilerinden Kadriye BELGİN DEMİRSU'nun, "**Değişen Madde Fonksiyonlu Maddelerin Test Eşitlemeye Etkisi**" konulu tez çalışması incelenmiştir.

Genel Müdürlüğümüze sunulan ve kayıtlarımızda muhafaza edilen araştırmanın; eğitim öğretimi aksatmayacak şekilde, **gönüllülük esasına** dayalı olarak uygulanması ve araştırma sonucunda yazılan tezin basılı ve dijital ortamda Genel Müdürlüğümüze teslim edilmesi koşulu ile araştırmaya kaynak oluşturabilecek tez çalışmasının ilgi yazıda belirtilen 8 ilde uygulanmasında bir sakınca bulunmamaktadır.

Bilgilerinizi ve gereğini rica ederim.

Dr. Şule ERŞAN
Bakan a.
Grup Başkanı

DAĞITIM

Bilgi :

Ankara, Bursa, Denizli, Eskişehir, İstanbul,
Kahramanmaraş, Kütahya, Malatya Valiliklerine
(İl Milli Eğitim Müdürlüğü)

Bu belge, 5070 sayılı Elektronik İmza Kanununun 5 inci maddesi gereğince güvenli elektronik imza ile imzalanmıştır.
Evrak teyidi <http://evraksorgu.meb.gov.tr> adresinden 55ae-bbd7-348b-8a60-ed5e kodu ile yapılabilir.

Araştırma No: 06648 Kızılay/ANKARA

Ayrıntılı bilgi için: Dr. A. ORAKCI, Sb. Md.

EK 8. ORJİNALLİK RAPORU

Turnitin Doküman Görüntüleyici - Google Chrome
https://turnitin.com/dv?s=1&o=546021574&u=1036944301&lang=tr&

thesis tez - TESLİM TARİHİ:Şub-2015 Roadmap 3 / 6

Originality GradeMark PeerMark

ORTAK MADDELERİN DEĞİŞEN MADDE KADRIYE BELGİN DEMİRUS TARAFINDAN

turnitin %5 BENZER 1 ÜZERİNDEN

**ORTAK MADDELERİN DEĞİŞEN MADDE FONKSİYONU
GÖSTERİP GÖSTERMEMESİ DURUMUNDA TEST
EŞİTLEMeye ETKİSİNİN FARKLI YÖNTEMLERLE
İNCELENMESİ**

**THE STUDY OF THE EFFECT OF ANCHOR ITEMS
SHOWING OR NOT SHOWING DIFFERANTIAL ITEM
FUNCTIONING TO TEST EQUATING USING VARIOUS
METHODS**

Kadriye Belgin DEMİRUS

Hacettepe Üniversitesi
Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin
Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme
Bilim Dalı İçin Öngördüğü
Doktora Tezi
olarak hazırlanmıştır.

2015

Tüm kaynaklar
1in16 Eşleşmesi

- www.slideshare.net %2
İnternet kaynağı - 5 url
- ardamercan925/1-program-giri %2
- mervegunn/ek-fen-giri-ekso... %1
- ardamercan925/ek-fen-giri-... %1
- turgay_maya/trkyede-uygul... %<1
- dennytobing/1/buku-jarigan-... %<1
- fr.slideshare.net %2
İnternet kaynağı
- cito.com.tr %2
İnternet kaynağı - 3 url
- egitim.cukurova.edu.tr %2
İnternet kaynağı - 2 url
- www.researchgate.net %2
İnternet kaynağı - 17 url
- www.pegem.net %1
İnternet kaynağı - 4 url
- TechKnowledge Turkey' ... %1
Öğrenci ödevleri- 32 ödevler
- Afyon Kocatepe Universit... %1
Öğrenci ödevleri- 22 ödevler

Kaynakları Çıkar

ÖZGEÇMİŞ

Kişisel Bilgiler

<i>Adı Soyadı</i>	Kadriye Belgin DEMİRUS
<i>Doğum Yeri</i>	Trabzon
<i>Doğum Tarihi</i>	1980

Eğitim Durumu

<i>Lise</i>	Balıkesir Muharrem Hasbi Koray (Süper) Lisesi	1998
<i>Lisans</i>	Balıkesir Üniversitesi Necatibey Eğitim Fakültesi/ Fen Bilgisi Öğretmenliği	2002
<i>Yüksek Lisans</i>	Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü/ Eğitim Bilimleri	2006
<i>Yabancı Dil</i>	İngilizce: Okuma (İyi). Yazma (İyi). Konuşma (İyi)	

İş Deneyimi

<i>Çalıştığı Kurumlar</i>	Şahin Ortaokulu Kazan/ANKARA	2013-halen
---------------------------	------------------------------	------------

İletişim

<i>e-Posta Adresi</i>	belgindemirus@gmail.com
<i>Jüri Tarihi</i>	24.04.2015