

ANKARA ÜNİVERSİTESİ  
EĞİTİM BİLİMLERİ ENSTİTÜSÜ  
ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI  
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME BİLİM DALI

MADDE TEPKİ KURAMI'NA DAYALI TEST EŞİTLEME  
YÖNTEMLERİNDEN KESTİRİLEN EŞİTLEME HATALARININ  
ÖRNEKLEM BÜYÜKLÜĞÜ VE YETENEK DAĞILIMINA GÖRE  
KARŞILAŞTIRILMASI

DOKTORA TEZİ

Sevilay Kilmen

Ankara  
Temmuz, 2010

ANKARA ÜNİVERSİTESİ  
EĞİTİM BİLİMLERİ ENSTİTÜSÜ  
ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI  
EĞİTİMDE ÖLÇME VE DEĞERLENDİRME BİLİM DALI

MADDE TEPKİ KURAMI'NA DAYALI TEST EŞİTLEME  
YÖNTEMLERİNDEN KESTİRİLEN EŞİTLEME HATALARININ  
ÖRNEKLEM BÜYÜKLÜĞÜ VE YETENEK DAĞILIMINA  
GÖRE KARŞILAŞTIRILMASI

DOKTORA TEZİ

Sevilay Kilmen

Danışman: Doç. Dr. Nükhet Çıkrıkçı Demirtaşlı

Ankara  
Temmuz, 2010

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼'ne;

Bu alıřma j¼rimiz tarafından ¼lme ve Deęerlendirme Anabilim Dalında DOKTORA TEZ ALIřMASI RAPORU olarak kabul edilmiřtir.

Bařkan .....  
Prof. Dr. Nizamettin KO

¼ye .....  
Prof. Dr. Ezel TAVřANCIL

¼ye .....  
Do. Dr. N¼khet IKRIKI DEMİRTAřLI (Danıřman)

¼ye .....  
Do. Dr. H¼lya KELECİOęLU

¼ye .....  
Do. Dr. Mehtap AKAN

Onay

Yukarıdaki imzaların, adı geen ¼đretim ¼yelerine ait olduęunu onaylarım.

02/07/2010

Prof. Dr. Nejla TURAL  
Enstit¼ M¼d¼r¼

## TEŐEKKÜR

Bu alıŐma boyunca, bana ve tezimin geliŐimine nemli katkılar sađlayan deđerli hocam ve danıŐmanım Do. Dr. Nkhet ıkırıı DemirtaŐlı'ya ncelikle teŐekkr bir bor bilirim.

Doktora eđitimimde beni yetiŐtiren deđerli hocalarım Prof. Dr. Nizamettin Ko'a, Prof. Dr. Ezel TavŐancıl'a, Yrd. Do. Dr. mer Kutlu'ya, Yrd. Do. Dr. may okluk'a; tezimin her aŐamasında yardım ve grŐlerini aldıđım arkadaŐlarım Devrim Erdem'e, Selda Gltekin'e ve İ. Alper Kse'ye teŐekkr ederim.

Doktora đrenimim boyunca yurt ii doktora burs olanađı sađlayan TUBİTAK'a teŐekkr ederim.

Desteklerini benden esirgemeyen anneme, babama ve kardeŐim Sonay Kaplan'a, bu alıŐmanın oluŐum srecinde zverisi, anlayıŐı ve yardımlarından dolayı sevgili eŐim Onur Kilmen'e teŐekkr bir bor bilirim.

Sevilay KİLMEN

## ÖNSÖZ

Eğitimde testler, öğrenci gelişimini izleme, belli bir programa öğrenci seçme gibi çeşitli amaçlarla kullanılmaktadır. Test gizliliği açısından, bazı durumlarda, öğrenciler, farklı zamanlarda farklı test formlarından aldıkları puanlara göre karşılaştırılabilirler. Örneğin, Türkiye’de, Üniversitelerarası Kurul Yabancı Dil Sınavı (ÜDS); Kamu Personeli Yabancı Dil Bilgisi Seviye Tespit Sınavı (KPDS); Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı (ALES) yılda iki kere uygulanan sınavlardır. Bu sınavlarda kullanılan sorular ön denemeden geçmeyen, tek kullanımlık sorular olduğu için, dönemden döneme uygulanan testlerin ortalama güçlüğünün denk olduğuna dair objektif kanıtlar bulunmamaktadır. Bu sınavlardan alınan puanlar belli bir süre geçerliğini (5 yıl) korumaktadır. Bu türden bir sınavın farklı dönemlerdeki uygulamalarından alınan aynı puanın -eşit olup olmadığı sınınamadığı halde-eşit olduğu varsayılarak, eğitim kurumlarına girişte ölçüt olarak kullanıldığı görülmektedir.

Ayrıca bu tür sınavlarda kesme puanı farklı dönemlerde uygulanan sınavlara göre değişmemektedir. Bu testlerin eşitlenmeden karşılaştırılması öğrenci başarısına ilişkin kararlara hata karışmasına neden olmaktadır. Diğer bir deyişle, eşitliği sınınamamış testlerden alınan puanların birbirinin dengi gibi kabul edilmesi, adil bir değerlendirmenin yapılmasını engellemektedir.

Bazı durumlarda öğrenciler eş zamanlı test edilseler bile, farklı test formlarından alınan puanlara göre karşılaştırılmaları gerekebilir. Milli Eğitim Bakanlığı tarafından düzenlenen Öğrenci Başarısını Belirleme Sınavı (ÖBBS) bu türden sınavlara örnek olarak verilebilir. ÖBBS’nda aynı sınıf düzeyindeki öğrencilere farklı sorulardan oluşan test formları verilmektedir. Bu sınavda, aynı yapıyı ölçen bir testin iki farklı formundan alınan aynı puan gerçekten birbirine eşit midir? ÖBBS’ndan alınan puanlar Klasik Test Kuramı’na göre belirlenmektedir. Bu kurama göre, öğrencilerin bir testten aldıkları puanlar

test maddelerinin güçlüğüne; test maddelerinin güçlüğü ise grubun başarı dağılımına bağlıdır. Bu durumda farklı güçlüklerle sahip testler farklı yetenek dağılımına sahip gruplara uygulanmışsa bu testten alınan aynı puanın birbirine denk olduğunu söylemek oldukça güçtür.

Oysa yukarıda bahsedilen durumlara ve sorulara test eşitleme yöntemleri ile çözüm getirilebilmektedir. Test eşitleme yöntemleri sayesinde 'farklı gruplara farklı test formları uygulansa bile' testlerden alınan puanlar aynı ölçek üzerine yerleştirilebilmektedir. Örneğin, ülkelerarası karşılaştırmalar yapmaya izin veren geniş ölçekli testlerde de (TIMSS, PISA, PIRLS) çok sayıda farklı kitapçık kullanılmaktadır. Bu kitapçıklarda, farklı kitapçıkları alan bireylerin puanlarının karşılaştırılmasını sağlayan ortak madde blokları kullanılmaktadır. Bu madde blokları sayesinde farklı soruları yanıtlayan bireylerin puanları birbiriyle karşılaştırılabilmektedir.

Bu araştırma test eşitlemede kullanılan yöntemlerin çeşitli koşullarda karşılaştırılmasını amaçlamaktadır. Bu araştırmanın yapılmasındaki nedenlerden biri Türkiye'de, her yıl veya yılda birden fazla uygulanarak sonuçları önemli kararlar (seçme, yeterlik belirleme) için kullanılan ve yukarıda örneği verilen ölçme uygulamalarında kullanılmayan test eşitleme uygulamalarına dikkati çekebilmektir.

## ÖZET

### MADDE TEPKİ KURAMI'NA DAYALI TEST EŞİTLEME YÖNTEMLERİNDEN KESTİRİLEN EŞİTLEME HATALARININ ÖRNEKLEM BÜYÜKLÜĞÜ VE YETENEK DAĞILIMINA GÖRE KARŞILAŞTIRILMASI

Kilmen, Sevilay

Doktora, Ölçme ve Değerlendirme Anabilim Dalı

Tez Danışmanı: Doç. Dr. Nükhet Çıkrıkçı Demirtaşlı

Temmuz, 2010, 101 sayfa

Bu araştırmada, Madde Tepki Kuramı'na dayalı "ortalama-ortalama", "ortalama-standart sapma", "Haebara" ve "Stocking-Lord" eşitleme yöntemlerinden kestirilen eşitleme hatalarının, yetenek dağılımı (benzer ve farklı yetenek dağılımı) ve örneklem büyüklüğü (500-1000 kişilik) değişkenlerine dayalı olarak karşılaştırılması amaçlanmıştır. Araştırmada, 1-0 şeklinde puanlanan 3 parametrelili modele uyumlu simülatif veriler 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 saptanmıştır. Bu bulgu literatürle de desteklenmektedir. Araştırmada en yüksek eşitleme hatalarını veren yöntemler ise ortalama-ortalama ve ortalama-standart sapma

yöntemleridir. Örneklem büyüklüğüne dayalı olarak test eşitleme hataları incelendiğinde, 1000 kişilik örneklem büyüklüğüne sahip gruplara ait eşitleme çalışmasından elde edilen eşitleme hatalarının 500 kişilik örneklemelerden elde edilen eşitleme hatalarına nazaran daha düşük olduğu saptanmıştır. 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.

Anahtar kelimeler: test eşitleme, Madde Tepki Kuramı, eşitleme hatası.



## **ABSTRACT**

### **COMPARISON OF EQUATING ERRORS ESTIMATED FROM TEST EQUATION METHODS BASED ON ITEM RESPONSE THEORY ACCORDING TO THE SAMPLE SIZE AND ABILITY DISTRIBUTION**

Kilmen, Sevilay

Doctoral, Department of Measurement and Evaluation

Advisor: Associate Professor Dr. N khet  ıkırıkçı Demirtaşı

July, 2010, 101 pages

In this study, it was aimed that equating methods of “mean-mean”, “mean-sigma”, “Heabera” and “Stocking-Lord” were compared according to the ability distribution (similar and different ability distribution) and sample size (500, 1000) variables based on Item Response Theory (IRT). Common item nonequivalent groups equating design on dichotomous simulated data adapted to 3-parameters IRT model was used. For data generation that was used in this study WinGEN2 software was used and with this software 600 dichotomously responded data matrices were generated. Ability parameters of simulated data have been estimated with the help of BILOG-MG by using expected a posteriori method. By using IRTEQ software, equating equations for test equating have been formed.

At the end of the research, in the case of groups having similar and different ability distribution with sample sizes  $N=1000$  and  $N=500$ , test equating application with “Stocking-Lord” method gave less error has been presented. This finding has supported by the literature. In this research, methods having the biggest equating errors were “mean-mean” and “mean-sigma”. When examining test equating errors based on sample size, equating errors gotten from groups with 1000 size were less than groups with 500 size

was found. When test equating methods compared according to the ability distribution, equating errors on equating groups with similar ability distribution were less than groups with different ability distribution has been found.

Key words: Test equating, Item Response Theory, equating error.

## İÇİNDEKİLER

TEŞEKKÜR.....	ii
ÖNSÖZ .....	iii
ÖZET .....	v
ABSTRACT.....	vii
ÇİZELGELER LİSTESİ .....	xi
ŞEKİLLER LİSTESİ .....	xii
EKLER.....	xiv
BÖLÜM 1 .....	1
GİRİŞ .....	1
Problem .....	1
Test Eşitlemenin Gerekli Olduğu Koşullar .....	6
Test Eşitlemenin Basamakları .....	8
Test Eşitleme Desenleri .....	8
Test Eşitleme Yöntemleri .....	16
Test Eşitlemede Hata Kavramı.....	31
İlgili Araştırmalar .....	33
Amaç .....	42
Önem.....	43
Sınırlılıklar.....	44
Tanımlar.....	45
BÖLÜM 2 .....	46
YÖNTEM.....	46
Araştırmanın Modeli.....	46
Verilerin Üretilmesi.....	46
Verilerin Analizi .....	55

BÖLÜM 3 .....	58
BULGULAR VE YORUMLAR .....	58
500 Kişilik <i>Benzer</i> Yetenek Dağılımına Sahip Gruplardan Elde Edilen Farklı Eşitleme Yöntemlerinden Kestirilen Hatalara İlişkin Bulgular .....	58
500 Kişilik <i>Farklı</i> Yetenek Dağılımına Sahip Gruplardan Elde Edilen Farklı Eşitleme Yöntemlerinden Kestirilen Hatalara İlişkin Bulgular .....	60
1000 Kişilik <i>Benzer</i> Yetenek Dağılımına Sahip Gruplardan Elde Edilen Farklı Eşitleme Yöntemlerinden Kestirilen Hatalara İlişkin Bulgular .....	62
1000 Kişilik <i>Farklı</i> Yetenek Dağılımına Sahip Gruplardan Elde Edilen Farklı Eşitleme Yöntemlerinden Kestirilen Hatalara İlişkin Bulgular .....	64
Ortalama-Ortalama Yöntemine İlişkin Bulgular .....	67
Ortalama-Standart Sapma Yöntemine İlişkin Bulgular .....	68
Haebara Yöntemine İlişkin Bulgular .....	68
Stocking-Lord Yöntemine İlişkin Bulgular .....	69
Örneklem Büyüklüğü Bakımından Test Eşitleme Hatalarının Karşılaştırılmasına İlişkin Bulgular.....	70
Yetenek Dağılımı Bakımından Test Eşitleme Hatalarının Karşılaştırılmasına İlişkin Bulgular.....	72
BÖLÜM 4 .....	75
SONUÇ VE ÖNERİLER .....	75
Sonuçlar.....	75
Öneriler .....	76
KAYNAKÇA .....	78
EKLER .....	90

## ÇİZELGELER LİSTESİ

Çizelge 1.	Araştırmada Kullanılan Verilerin Yapısı.....	47
Çizelge 2.	K Formuna Ait Maddelerin Parametre Değerleri.....	49
Çizelge 3.	L Formuna Ait Maddelerin Parametre Değerleri.....	50
Çizelge 4.	Örnekleme Büyüklüğü ve Yetenek Dağılımı Değişkenlerine Göre Kullanılan Test Formları.....	52
Çizelge 5.	N=500 Benzer Yetenek Dağılımı Koşulunda Farklı Yöntemlere Göre Yapılan Eşitlemelerden Hesaplanan Hata Miktarları.....	58
Çizelge 6.	N=500 Farklı Yetenek Dağılımı Koşulunda Farklı Yöntemlere Göre Yapılan Eşitlemelerden Hesaplanan Hata Miktarları.....	60
Çizelge 7.	N=1000 Benzer Yetenek Dağılımı Koşulunda Farklı Yöntemlere Göre Yapılan Eşitlemelerden Hesaplanan Hata Miktarları.....	62
Çizelge 8.	N=1000 Farklı Yetenek Dağılımı Koşulunda Farklı Yöntemlere Göre Yapılan Eşitlemelerden Hesaplanan Hata Miktarları.....	65

## ŞEKİLLER LİSTESİ

Şekil 1.	Test Eşitleme Yöntemlerinin Eşitlenecek Testlerin Güçlük Düzeyleri Bakımından Sınıflandırılması.....	5
Şekil 2.	Test Eşitleme Desenleri.....	9
Şekil 3.	Tek Grup Deseni.....	9
Şekil 4.	Dengelenmiş Tek Grup Deseni.....	11
Şekil 5.	İç Ortak Maddeli Eşitlenmemiş Gruplar Deseni.....	12
Şekil 6.	Dış Ortak Maddeli Eşitlenmemiş Gruplar Deseni.....	13
Şekil 7.	Klasik Test Kuramı'na Dayalı Test Eşitleme Yöntemleri.....	17
Şekil 8.	Başarılı ve Başarısız Grupta Doğrusal Eşitleme.....	20
Şekil 9.	Eşit Yüzdelikli Eşitleme.....	21
Şekil 10.	Araştırma şeması.....	53
Şekil 11.	Ortalama-Ortalama Yöntemine ait RMSD Değerleri.....	67
Şekil 12.	Ortalama-Standart Sapma Yöntemine ait RMSD Değerleri...	68
Şekil 13.	Haebara Yöntemine ait RMSD Değerleri.....	69
Şekil 14.	Stocking-Lord Yöntemine ait RMSD Değerleri.....	70

Şekil 15. Örneklem Büyüklüğü Açısından Test Eşitleme Yöntemlerinden Kestirilen Hataların (RMSD değerleri) Karşılaştırılması .....	71
Şekil 16. Yetenek Dağılımı Açısından Test Eşitleme Yöntemlerinden Kestirilen Hataların (RMSD değerleri) Karşılaştırılması .....	73

## EKLER

Ek 1.	500 Kişilik Benzer Yetenek Dağılımına Sahip Grupların Eşitlenmesinden Elde Edilen RMSD Değerleri.....	90
Ek 2.	500 Kişilik Farklı Yetenek Dağılımına Sahip Grupların Eşitlenmesinden Elde Edilen RMSD Değerleri.....	93
Ek 3.	1000 Kişilik Benzer Yetenek Dağılımına Sahip Grupların Eşitlenmesinden Elde Edilen RMSD Değerleri .....	96
Ek 4.	1000 Kişilik Farklı Yetenek Dağılımına Sahip Grupların Eşitlenmesinden Elde Edilen RMSD Değerleri .....	99



## **BÖLÜM 1**

### **GİRİŞ**

Bu bölümde, test eşitlemeye dayalı temel kavramlara ilişkin genel bilgiler verilmiş, Klasik Test Kuramı ve Madde Tepki Kuramı'na dayalı test eşitleme yöntemleri tanıtılmış, araştırmanın problemi açıklanmış, amaç, önem, sınırlılıklar ve tanımlara yer verilmiştir.

#### **Problem**

Eğitim ve psikolojide testler, eğitim ihtiyaçlarının ve kalitesinin belirlenmesinde, eğitim politikalarına yön vermede, belli bir işe eleman, belli bir kuruma öğrenci seçme ve yerleştirmede, rehberlik ve klinik hizmetlerin yürütülmesinde yaygın olarak kullanılmaktadır. Testler, bu amaçlar çerçevesinde, birey içi ve bireylerarası farkları duyarlı bir şekilde ortaya çıkarmayı amaçlar. Özellikle günümüzde, testlerin ülke çapında ortaöğretim kurumlarına ve yükseköğretim programlarına girişte kullanılması, bu testlerin sonuçlarına göre verilecek kararların önemini daha da arttırmıştır. Öğrencilerin testlerden aldığı puan, gelecekteki iş ve meslek yaşamını belirleyecek ve iyi bir iş sahibi olmasını sağlayacak bir ölçü haline gelmiştir. Bunun bir sonucu olarak, testlere olan ilgi giderek artmıştır.

Testlerin bireylerle ilgili önemli ve kritik kararlar almada kullanılması için, testlerin bazı özellikleri taşıması gerekmektedir. Örneğin, test puanlarının geçerli ve güvenilir çıkarsamalar yapmaya olanak tanınması, testlerin duyarlı ölçmeler yapması, yanlı kararlara yol açmaması, güvenilir ve adil bir değerlendirme yapılabilmesi için testlerin gizliliğinin sağlanması bunlardan bazılarıdır. Bu çerçevede, uygulamada, bu özelliklerden ödün vermemek için bazı yollar izlenir. Örneğin, bazı test uygulamalarında, öğrencilerin kopya çekme girişimlerini azaltmak için, aynı oturumda birden fazla test formu

kullanılır. Farklı zamanlarda yapılan test uygulamalarında, test gizliliğini sağlamak için bir testin paralel formları kullanılmaktadır. Bu tür durumlar, her yıl veya yılda birden fazla uygulanan, öğrenci öğrenmelerini izleyen ve değerlendiren, belli bir alandaki yeterliklerin düzeyini saptamaya çalışan ya da belli bir alandaki yeterliğe dayalı olarak bireyleri bir işe veya bir programa seçmede kullanılan testler için söz konusudur. Bu noktada bireyler hakkında verilecek gelişimi değerlendirme veya seçme kararlarının geçerli ve tutarlı olabilmesi için her yıl veya yılda birden fazla aynı amaçla uygulanan testlerden elde edilen puanların birbirleri ile karşılaştırılabilir hale dönüştürülmesi gerekir.

Klasik Test Kuramı'nda bir testin farklı formlarından alınan puanların birbiri ile karşılaştırılabilmesi için, test formlarının paralel olması gerekmektedir. Pratikte paralel testlerin oluşturulması ise oldukça zordur (Aiken, 2000). Klasik Test Kuramı'na göre, ölçmenin standart hatası test güvenilirliğinin ve varyansın bir fonksiyonudur. Testi alan tüm katılımcılar için eşit olduğu varsayılır. Testler paralel olsa bile, testi alanların yetenek düzeyi birbirinden farklı olacağından ve farklı hata kaynaklarından etkilenme düzeyleri de farklı olacağından, test puanları farklı miktarlarda hataya sahip olacaktır. Ayrıca, bir test zor ise, testi alan kişi düşük yeteneğe sahipmiş gibi; eğer test kolay bir test ise, testi alan kişi yüksek yetenekli gibi görünebilir. Katılımcıların yeteneğini test maddelerinin güçlük düzeyi belirlemektedir. Çünkü Klasik Test Kuramı'nda maddelerin ayırt ediciliği ve güçlüğü belli bir grup için geçerlidir. Bu nedenle farklı testleri alan kişileri Klasik Test Kuramı'na göre karşılaştırmak oldukça güçtür (Hambleton, Swaminathan ve Rogers, 1991). Bu durumda, aynı amaçla, aynı kapsama hizmet eden farklı test formlarından elde edilen puanların birbirine dönüşümünü sağlayan işlemlere ihtiyaç duyulmaktadır.

Test puanlarının birbirine dönüşümü ile ilgili olarak literatürde çeşitli kavramların kullanıldığı görülmektedir. İlişkilendirme (linking), bir testten elde edilen puanların diğer testin puanına dönüştürme sürecine verilen genel bir isimdir. Linn (1993), ilişkilendirme yöntemlerini beş kategoride ele almaktadır. Bunlar:

1. Eşitleme (equating): aynı yapıyı aynı yolla ölçen testlerin ilişkilendirilmesinde kullanılır.
2. Kalibrasyon (calibration): aynı yapıyı farklı yollarla ölçen testlerin ilişkilendirilmesinde kullanılır.
3. İstatistiksel moderasyon (statistical moderation): aynı yapıyı ölçmeyen testlerin, bir dış ölçme sonucu ile istatistiksel olarak ilişkilendirilmesinde kullanılır.
4. Kestirim (prediction): aynı yapıyı ölçmeyen fakat puanları arasında ampirik ilişkilerin kestirilebildiği testlerin ilişkilendirilmesinde kullanılır.
5. Sosyal moderasyon (social moderation): aynı yapıyı ölçmeyen test puanlarının yargıya dayalı olarak ilişkilendirilmesinde kullanılır.

Yukarıda, Linn (1993) tarafından yapılan sınıflamaya göre, test eşitleme (test equating), ilişkilendirilecek testlerin aynı yapıyı ölçmesi ve aynı yolla ölçmesi yönüyle diğer ilişkilendirme yöntemlerinden ayrılmaktadır. Buna göre test eşitleme, aynı yapıyı ölçen farklı test formlarından elde edilen puanların birbirine dönüşümünü sağlayan istatistiksel bir süreçtir. Eşitlenmiş puanlar, testin ne zaman ve kime uygulandığına bakılmaksızın aynı anlama gelir. Böylece bir test formundan elde edilen puanlar, diğer test formundan elde edilen puanlarla karşılaştırılabilmektedir (Kolen ve Brennan, 2004).

Angoff (1971), test eşitlemeyi “bir test formunun birim sisteminin diğer test formunun birim sistemine dönüştürme süreci” olarak tanımlamaktadır. Angoff (1971) bu tanımlamayı açıklamak için Celsius ve Fahrenheit ölçeklerinde kullanılan birim dönüştürmesini örnek olarak göstermiştir. Celsius ve Fahrenheit ölçekleri aynı yapıyı (sıcaklığı) ölçmektedirler. Ancak, farklı ölçekler kullanıldığında farklı sonuçlar elde edilmektedir. Sonuçların farklı olması, ölçek birimlerinin farklı olmasından kaynaklanmaktadır. Bu nedenle sıcaklık ölçen ölçekler arasında doğrusal bir dönüşüm yapılmaktadır. Bu dönüşüm Eşitlik 1 ve 2’de verilmektedir:

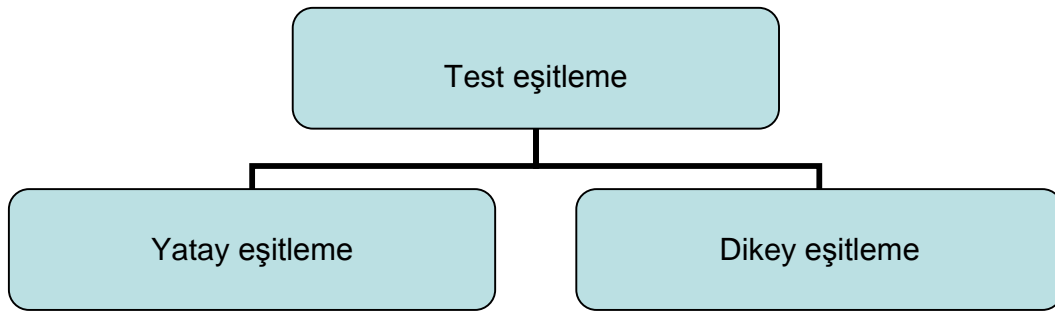
$$C = (F - 32) / 2 \quad (1)$$

$$F = (C * 2) + 32 \quad (2)$$

Eşitlik 1 ve 2’de görüldüğü gibi, havanın sıcaklık değeri, Celsius ölçeğinin birimleri Fahrenheit ölçeğinin birimlerine dönüştürülerek Fahrenheit cinsinden veya Fahrenheit ölçeğinin birimleri Celsius ölçeğinin birimlerine dönüştürülerek Celsius ölçeği cinsinden ifade edilebilmektedir. Böylece, farklı ölçeklerle ölçülen iki farklı odanın sıcaklığı birbiri ile kıyaslanabilmektedir.

Eğitimdeki ölçmelerde de test eşitleme sürecinde aynı yapıyı ölçen iki ölçek bulunmaktadır. Test eşitlemede, aynı yapıyı ölçen iki ölçekten alınan puanlar arasında istatistiksel ilişki kurulmaktadır (Felan, 2002). X ve Y, aynı yapıyı ölçen iki test formu olmak üzere, X formundan Y formuna puan dönüşümünde kullanılan fonksiyon, Y formundan X formuna puan dönüşümünde de kullanılabilir. Bu özelliğe test eşitlemenin simetriklik özelliği denilmektedir. Ayrıca, test eşitleme sonucunda elde edilen puan dönüşümünün, eşitleme denkleminin oluşturulduğu gruba bakılmaksızın tüm gruplar için aynı olması beklenmektedir. Test eşitlemenin bu özelliği de gruptan bağımsızlık özelliği olarak anılmaktadır (Angoff, 1971; Kolen ve Brennan, 2004; Lord 1980).

Lord (1980) test eşitlemenin gerçekleştirilebilmesi için, eşitleme yapılacak testlerin güçlüklerinin eşit olması gerektiğini öne sürmüştür. Ancak bu kural Madde Tepki Kuramı’nın değişmezlik özelliği sayesinde aşılabilmektedir. Çünkü Madde Tepki Kuramı’na göre yetenek kestirimi, bir kişinin zor ya da kolay testi almasına bağlı değildir. Günümüzde, güçlükleri eşit olmayan testler arasında da eşitlemeler yapılabilmektedir. Bu bağlamda test eşitleme yöntemleri, eşitlenecek testlerin güçlükleri bakımından ikiye ayrılmaktadır. Bu yöntemler Şekil 1’de gösterilmektedir:



Şekil 1. Test Eşitleme Yöntemlerinin **Eşitlenecek Testlerin Güçlük Düzeyleri** Bakımından Sınıflandırılması

Benzer güçlük düzeyine sahip aynı yapıyı ölçen iki testi eşitleme, yatay eşitleme (horizontal equating) olarak adlandırılmaktadır. Yatay eşitleme, daha çok farklı dönemlerde ve gizliliğin gerekli olduğu bir testin farklı formları söz konusu olduğunda kullanılır ve bu formların paralel olması beklenir. Yatay eşitleme, benzer yetenek grubundaki bireyler üzerinde uygulanır (örneğin aynı sınıf düzeyi). Farklı yetenek düzeyindeki kişilere uygulanan ve aynı yapıyı ölçen farklı güçlük düzeyindeki iki testi (örneğin farklı sınıflardaki kişilerin aldıkları testler) eşitleme ise dikey eşitleme (vertical equating) olarak adlandırılmaktadır. Öğrenci başarı gelişiminin yıllara göre izlenmesi bu kapsamda yer almaktadır (Aiken, 2000; Crocker ve Algina; 1986; Felan, 2002; Hambleton ve Swaminathan, 1989).

Test eşitleme doğrudan ve dolaylı olmak üzere iki şekilde kullanılabilir. Doğrudan test eşitleme (direct equating), bir testteki puanının doğrudan diğer testteki puana çevrilebilmesi anlamına gelmektedir. Dolaylı test eşitlemede (indirect equating) ise, bir testten elde edilen puanlar, üçüncü bir test kullanılarak, diğer testin puanlarına çevrilmektedir (Holland, 2007). Kolen ve Brennan (2004), dolaylı eşitlemeyi *zincirleme eşitleme* (chain equating) olarak da tanımlamaktadır.

Test eşitleme çalışmaları ortak kişiler ve ortak maddeler üzerinden yapılabilir. Ortak kişi eşitlemede (person anchoring), aynı kişiler aynı özelliği ölçen farklı testleri almaktadırlar. Bu testlerde ortak madde yer

almamaktadır. Bu durumda eşitleme çalışmaları kişiler üzerinden yapılmaktadır. Ortak madde ile eşitlemede (item anchoring) ise, bir testin farklı formlarını alanlar, bu formlar içerisinde yer alan ortak maddeleri de yanıtlamaktadırlar. Bu durumda eşitleme çalışmaları ortak maddeler kullanılarak yapılmaktadır (Yu ve Osborn Popp, 2005).

### **Test Eşitlemenin Gerekli Olduğu Koşullar**

Test eşitleme süreci aşağıdaki test koşullarında kullanılabilir:

- ÖBBS sınavı gibi, öğrencilere eşzamanlı olarak bir testin farklı formlarının verildiği durumlarda, zor testi alan bireylerin, kolay testi alan bireylere göre daha düşük puan almaları olasıdır. Bu durum farklı test formlarından elde edilen puanların birbiriyle karşılaştırılabilmesini güçleştirmektedir. Test eşitleme, zor testi alan bireylere karşı olası haksızlıkları engellemekte, test formlarından kaynaklanan yanlılık problemlerini ortadan kaldırmaktadır (Angoff, 1971; Cook ve Eignor, 1991; Hambleton, Swaminathan ve Rogers, 1991). Örneğin, ülkelerarası karşılaştırmalar yapmaya izin veren TIMMS (Trends in International Mathematics and Science Studies), PIRLS (Progress in International Reading Literacy Study) ve PISA (Programme for International Student Assessment) gibi geniş ölçekli testlerde de farklı sorular içeren çok sayıda kitapçık kullanılmaktadır. Test eşitleme çalışmaları sonucunda, farklı soruları yanıtlayan bireylerin puanları birbiriyle karşılaştırılabilmektedir.
- Bazı ölçme koşulları bir testin her dönemde farklı formlarının uygulanmasını gerektirebilir. Bu durumda farklı kişilerin farklı dönemlerde aldıkları puanların test eşitleme çalışmalarıyla karşılaştırma yapmaya uygun hale getirilmesi gerekmektedir. Örneğin, Üniversitelerarası Kurul Yabancı Dil Sınavı (ÜDS); Kamu Personeli Yabancı Dil Bilgisi Seviye Tespit Sınavı (KPDS); Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı (ALES) yılda iki kere uygulanan

sınavlardır. Bu sınavlara farklı zamanlarda katılarak alınan aynı puanların testin güçlüğünü gözetererek birbirine denk kabul etmek için bu puanların aynı ölçek üzerine yerleştirilmesi, diğer bir deyişle puanların aynı ölçeğe dönüştürülerek eşitlenmesi gerekmektedir.

- Test eşitleme öğrencilerin ölçülen özellik bakımından gerçek gelişimini izlemek için kullanılır. Bu türden durumlarda kullanılan test eşitleme yöntemine dikey eşitleme denilmektedir (Angoff, 1971; Crocker ve Algina, 1986).
- Ön test-son test kontrol gruplu desenlerin kullanıldığı deneysel çalışmalarda ön test etkisini ortadan kaldırmak amacıyla kullanılan paralel formlardan elde edilen puanların birbiriyle kıyaslanabilmesi için kullanılabilir.
- Test eşitlemenin bir diğer avantajı ise, bir teste ait çok sayıda maddenin parametrelerinin tek oturumda kestirilmesinde imkan tanımasıdır. Bazı oturumlarda testteki madde sayısı her bireye uygulanamayacak kadar fazla olur. Bu durumda, soruların uygulandığı yanıtlayıcı grubundaki her birey belli bir test deseni kapsamında oluşturulmuş soru kümelerine ve ortak maddelere yanıt verirler. Bu türden test uygulamalarına eksik test deseni (incomplete test design) adı verilir (Demirtaşlı ve Arıkan, 2008; Verhelst, 2004). Böylece, bir kişinin yanıtlanmadığı maddelerdeki performansı ortak maddelerle kolaylıkla belirlenebilir. Bu türden uygulamalar, Madde Tepki Kuramı'nın (MTK) değişmezlik özelliğine (invariance) dayanan ve araştırmacılara/öğreticilere pratik kolaylıklar sağlayan uygulamalardır. Örneğin 75 maddelik bir testin madde parametreleri kestirilmek istensin. Testin çok uzun bir test olması ve bu yüzden test alanlarda bıkkınlık ve dikkat kaybına yol açmaması için maddeler A, B ve C olmak üzere üç madde setine/test kitapçığına bölünür. Birbirinden farklı iki gruptan, birinci gruba A ve B madde seti; ikinci gruba ise A ve C madde seti uygulanır. A madde seti, iki farklı grubun madde

parametrelerinin birbirine dönüşümünü sağlayan ortak madde setidir (Crocker ve Algina, 1986).

### **Test Eşitlemenin Basamakları**

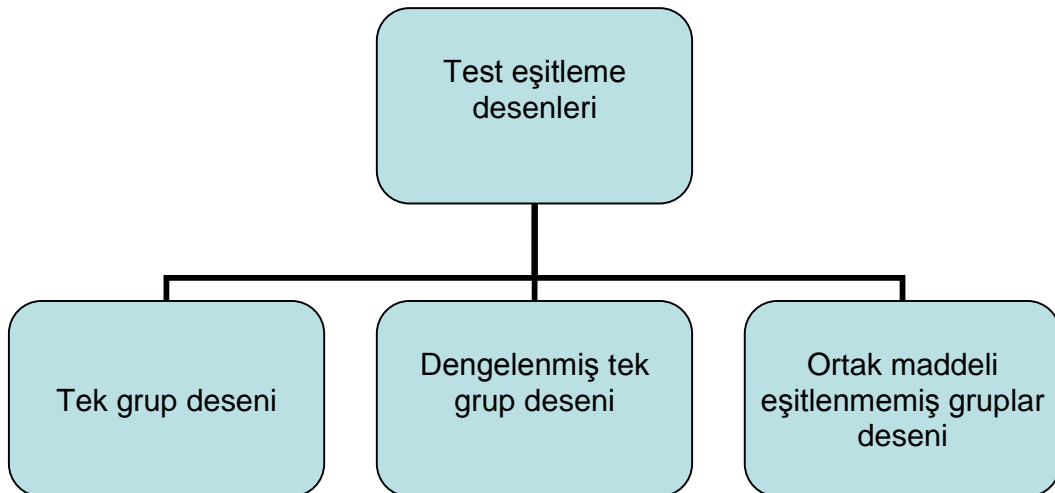
Kolen ve Brennan (2004)'e göre eşitleme yedi basamaktan oluşan bir süreçtir. Bu aşamalar aşağıdaki gibidir:

1. Eşitleme amacına karar verilmesi
2. Alternatif formların oluşturulması (Test formları benzer içeriğe sahip olmalıdır)
3. Veri toplama deseninin seçimi
4. Veri toplama deseninin uygulanması (testin veri toplama desenine göre uygulanması ve puanların elde edilmesi)
5. Eşitlemenin bir veya daha fazla işlevsel tanımının seçilmesi (eşitleme, kestirim yapılacak test formları arasındaki ilişki türlerine ilişkin olarak bir seçim yapmayı gerektirir. Örneğin, doğrusal veya doğrusal olmayan eşitleme yöntemlerinin uygulanıp uygulanmayacağına karar verilmesi)
6. Bir veya daha fazla eşitleme ilişkilerini kestirim yönteminin seçilmesi (eşitlemede çeşitli yöntemler bulunmaktadır. Örneğin doğrusal eşitleme ilişkileri Tucker ve Levine yöntemleri kullanılarak kestirilebilir)
7. Eşitleme sonuçlarının değerlendirilmesi.

### **Test Eşitleme Desenleri**

Test eşitleme çalışması yapmak isteyen bir araştırmacı öncelikle, testi alan gruba/grupları ve testte ortak madde olup olmasını dikkate alarak kullanacağı eşitleme desenine karar vermelidir. Test eşitleme desenleri Şekil 2'de gösterilmektedir.

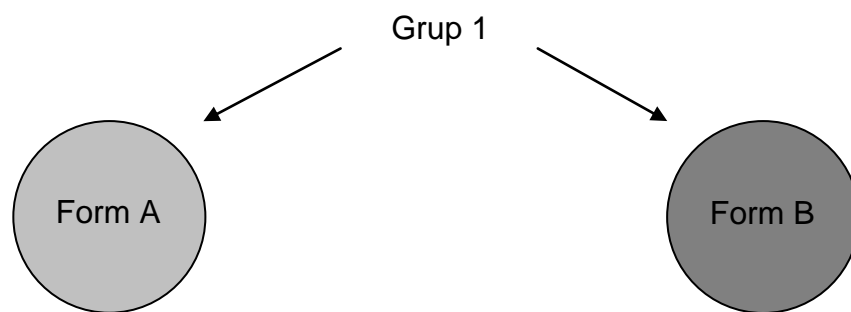




Şekil 2. Test Eşitleme Desenleri

### Tek grup deseni (Single group design)

Tek grup deseninde iki farklı test formunu (A, B) alan tek bir grup bulunmaktadır. Testleri alan grubun aynı olması, grup değişkenini sabit tutmaktadır. Böylece test formları arasındaki güçlük farkları gruba bağlı olmadan değerlendirilmektedir. Bu güçlük farklarından yola çıkılarak testler arasında eşitleme denklemi elde edilmektedir. Bu desende grupların hedef evreni temsil etmelerine gerek yoktur. Gruplar hedef evrenden daha başarılı ya da daha başarısız olabilirler (Cook ve Eignor, 1991; Hambleton ve Swaminathan, 1989; Kolen, 1988; Livingston, 2004). Tek grup deseni Şekil 3'te özetlenmektedir:

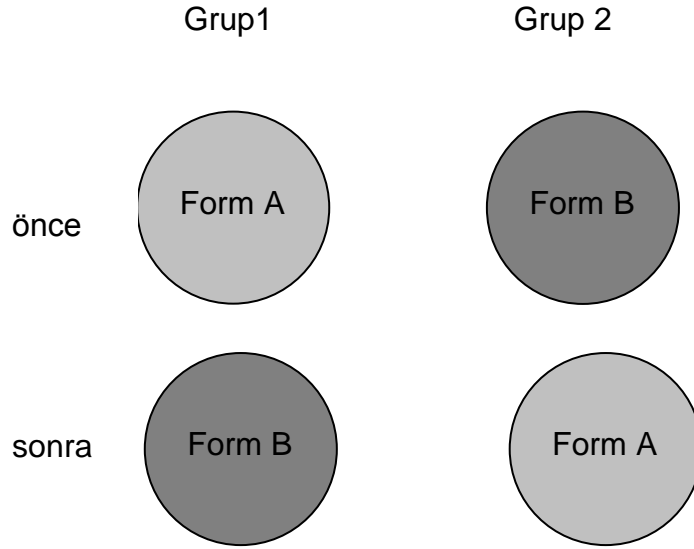


Şekil 3. Tek Grup Deseni

Şekil 3 incelendiğinde, aynı kişilerin iki farklı test formunu aldıkları görülmektedir. Bu desenin kullanımında, grup değişkeni sabit tutulsa da, bazı problemler yaşanmaktadır. Örneğin, iki testi almak öğrenciyi yorabilir. Sonra verilen form, önce verilen forma göre öğrencilere daha zor gelebilir. Öğrencilere verilen bu iki forma ait maddelerdeki benzerlik öğrenci performansını yükseltebilir. Böyle durumlarda, testler arasındaki güçlük farklarını yorumlamak güçleşir (Cook ve Eignor, 1991; Hambleton ve Swaminathan, 1989; Kolen, 1988; Livingston, 2004).

### **Dengelenmiş tek grup deseni (Counterbalanced single group design)**

Tek grup desenindeki aynı formu üst üste almanın getireceği öğrenme/deneyim problemlerini ortadan kaldırmak için literatürde dengelenmiş tek grup deseninin kullanılması önerilmektedir. Örneğin, kağıt kalem testi olarak verilen bir X testi ve bilgisayar ortamında uygulanan Y testi olsun. Bu testlerin aynı zamanda aynı sınıfın içinde uygulanması zordur. Bunun yerine test alacak gruptaki bireyler ikiye ayrılarak, farklı test koşullarına farklı sıralarla atanırlar (Kolen, 2007). Bu desende grupların yetenek bakımından birbirine olabildiğince benzer olması istenir. Ayrıca tek grup deseninde olduğu gibi bu desende de grupların hedef evreni yansıtmalarına gerek yoktur. Gruplar hedef evrenden daha başarılı ya da daha başarısız olabilirler (Cook ve Eignor, 1991; Livingston, 2004). Şekil 4'te dengelenmiş tek grup desenine dayalı olarak yapılan eşitleme çalışması gösterilmektedir:



Şekil 4. Dengelenmiş Tek Grup Deseni

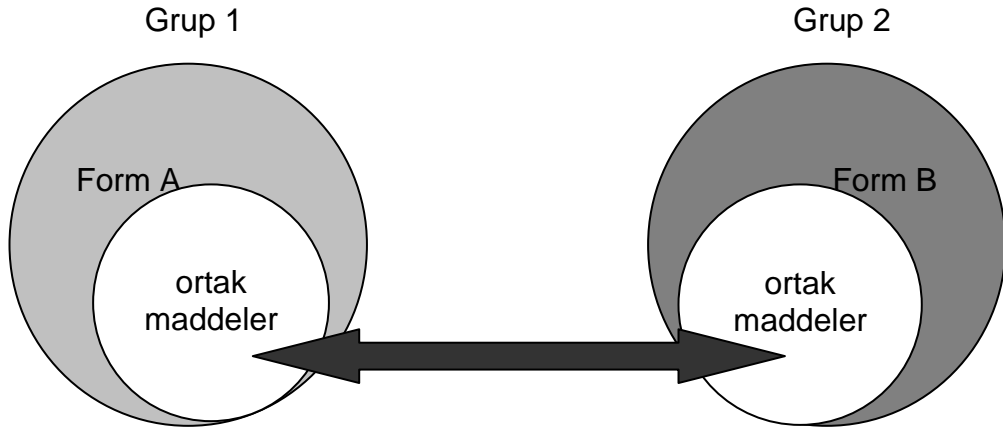
Şekil 4 incelendiğinde testleri alacak öğrenci grubunun ikiye ayrıldığı görülmektedir. Form A ve Form B bir testin iki formu olmak üzere; bir grup önce Form A'yı sonra Form B'yi; diğer grup ise önce Form B'yi sonra Form A'yı almaktadır. Böylece, sıra etkisi kontrol edilmektedir.

**Ortak maddeli eşitlenmemiş gruplar deseni (Common items nonequivalent group design)**

Bir testin iki farklı formunda yer alan maddelere ortak (common, anchor) maddeler adı verilmektedir. Bu desende ise iki farklı grup farklı testleri (A, B) almaktadırlar. Gruplar sadece tek bir formu yanıtlamaktadırlar. Bu desende grupların eşitlenmesi gerekmez. MTK'ya göre eşitlemede, ortak maddeler, madde ve yetenek parametrelerinin aynı ölçek üzerine yerleştirilmesinde kullanılır. Diğer bir deyişle eşitleme denkleminin oluşmasında kullanılmaktadırlar (Cook ve Eignor, 1991; Vale, 1986). Bu desende her iki grubun ortalama ve standart sapmalarının farklılaşmasının bir önemi yoktur. Bu bağlamda eşitleme çalışmalarının amacı, her iki testten elde edilen madde parametrelerini ortak bir ölçeğe yerleştirmek ve her iki testten elde edilen puanları karşılaştırmaktır (Embretson ve Reise, 2000).

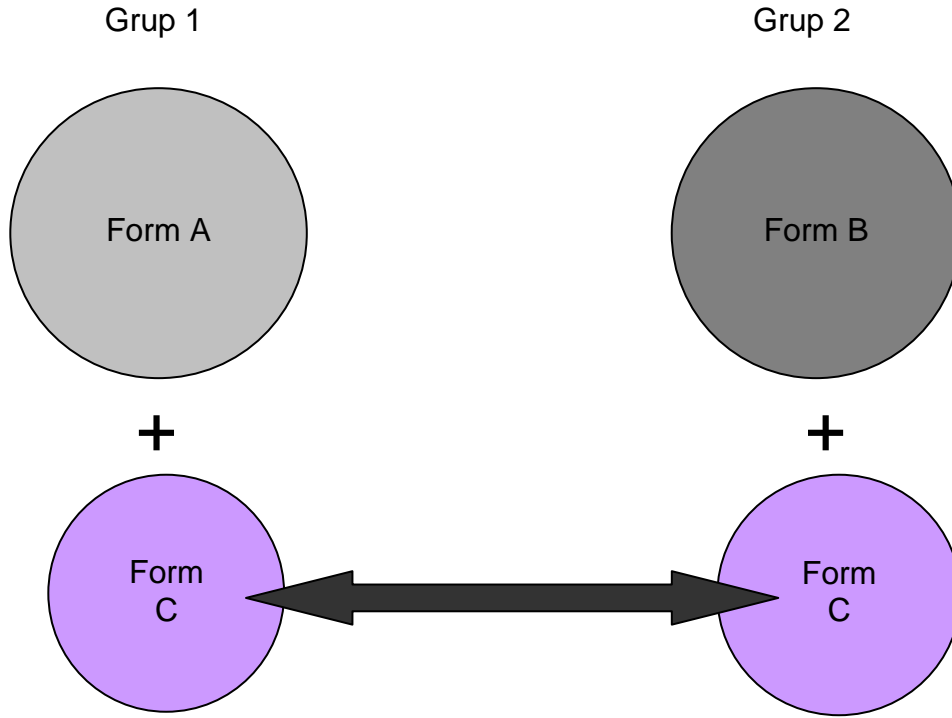
Ortak maddeli eşitlenmemiş gruplar deseni, ortak maddelerin puanlamaya katılma durumuna göre ikiye ayrılmaktadır. Ortak test puanlarının, testin tamamından alınan puana katkıda bulunması iç ortak maddelere (internal anchor) işaret etmektedir. İç ortak maddeler testteki diğer maddelerle karşık durumdadır. Testteki ortak maddelerden alınan puanların test puanlarına karışmadığı durumlarda dış ortak maddeler (external anchor) söz konusudur. Dış ortak maddeler testten ayrı uygulanır (Kolen, 1988).

Şekil 5'te iç ortak maddeli eşitlenmemiş gruplar test deseni gösterilmektedir.



Şekil 5. İç Ortak Maddeli Eşitlenmemiş Gruplar Deseni

Şekil 5 incelendiğinde, testleri alan iki farklı öğrenci grubunun olduğu görülmektedir. Bir grup Form A'yı; diğer grup ise Form B'yi almaktadır. Bu desende test eşitleme çalışmaları farklı test formlarında yer alan "iç ortak maddeler" sayesinde yapılabilmektedir. Şekil 6'da dış ortak maddeli test eşitleme deseni gösterilmektedir:



Şekil 6. Dış Ortak Maddeli Eşitlenmemiş Gruplar Deseni

Şekil 6 incelendiğinde, bir grubun Form A ve Form C'yi; diğer grubun ise Form B ve Form C'yi aldığı görülmektedir. Form C dış ortak madde setidir. Ortak maddeler iç ortak maddeli desenden farklı olarak ayrı bir test olarak uygulanmakta ve puanlamaya katılmamaktadırlar. Bu desende test eşitleme denklemi Form C'ye verilen tepkilerden yola çıkılarak oluşturulmaktadır.

Ortak maddeli eşitlenmemiş gruplar deseni, test eşitlemede sıklıkla kullanılan ve araştırmacılara pratik avantajlar sağlayan bir desendir. Örneğin, tek oturumda farklı test formlarının kullanıldığı sınavların eşitlenmesi bu desene göre yapılır. Ayrıca bu desen, farklı yıllarda farklı formların kullanılmasıyla gerçekleşen daha karmaşık durumlarda, farklı test formlarının eşitlenmesinin yanı sıra, farklı yıllara ait puanların eşitlenmesinde de kullanılır (Michalides, 2003).

Bu desende kullanılan ortak maddelerin belli başlı özellikleri vardır. Bu özelliklerden biri, ortak maddelerin eşitlenmek istenen test formlarının içeriğini ve özelliklerini yansıtmasıdır. Test formu ile ortak madde seti

olabildiğince özdeş olmalıdır. Tek soru köküne bağlı maddeler ortak madde olarak kullanılmamalıdır. Örneğin, paragraf sorularında tek bir uyarıcıya bağlı maddeler yer almaktadır. Bu maddeler kendi aralarında yüksek korelasyona sahip olabilirler. Bu durum Madde Tepki Kuramı'nın temel varsayımlarından biri olan yerel bağımsızlığın ihlaline yol açmaktadır. Bu nedenle bu tip maddeler ortak madde setine dahil edilmemelidir (Cook ve Eignor, 1991; Muraki, Hombo ve Lee, 2000; Yu ve Osborn Popp, 2005).

Ortak maddeli test eşitleme yöntemi ile ilgili olarak bir çok araştırma yapılmıştır. Bu araştırmalarda ortak madde sayısının, ortak madde kapsamının ve güçlüğünün test eşitleme hatası üzerindeki etkileri çalışılmıştır. Ortak maddelerle ilgili olarak yapılan araştırmalar aşağıda verilmektedir.

Klein ve Jarjoura (1985) tarafından yapılan araştırmada testin bütünü yansıtan maddelerden oluşan ortak madde seti ile yapılan eşitlemelerin, testin bütünü yansıtmayan maddelerden oluşan ortak madde seti ile yapılan eşitlemelere nazaran daha az eşitleme hatası ile sonuçlandığı saptanmıştır. Yang (1997) tarafından yapılan araştırmada ise, ortak madde güçlüğünün testin güçlüğünü yansıttığı durumlarda eşitlemenin daha doğru yapıldığı sonucuna ulaşılmıştır.

Meyers, Miller ve Way (2009) tarafından yapılan çalışmada, ortak maddelerin test içindeki yerinin değişmesinin eşitleme hataları üzerindeki etkisi araştırılmıştır. Bir parametrelili modele dayalı veriler üzerinde yapılan çalışmada, ortak maddelerin test içindeki yerinin değişmesinin, maddelerin güçlüklüğü üzerinde etkili bir faktör olduğu saptanmıştır. Bu durumun test eşitleme sonuçlarını olumsuz etkilediği bulunmuştur.

Ortak maddelerle ilgili bir başka araştırma konusu ise ortak madde sayısıdır. Literatürde ortak madde sayısının çalışıldığı çok sayıda araştırma bulunmaktadır. Bu araştırmalardan biri Hills, Subhiyah ve Hirsch (1988) tarafından yapılmıştır. Bu araştırmada, 5, 10, 15, 20, 25 ve 30 maddelik ortak testler oluşturulmuş ve eşitleme hataları gözlemlenmiştir. 5 maddelik ortak

test için eşitleme hataları diğerlerine nazaran daha yüksek çıkarken, diğerleri için benzer eşitleme hataları elde edilmiştir. Cohen ve Kim (1998) tarafından yapılan araştırmada ise, ortak madde sayısı 5, 10, 20 ve 30 olarak değişimlenmiştir. Araştırma sonucunda, ortak madde sayısının eşitleme hatalarını etkilemediği saptanmıştır.

Suanthong (1998) tarafından yapılan çalışmada, ortak madde sayısı 10, 20, 30 madde olarak değişimlenerek eşitleme hataları gözlenmiştir. Araştırmada, ortak maddeli eşitlenmemiş gruplar üzerinde bir parametrelili modele dayalı eşitlemeler yapılmıştır. Araştırma sonucunda, 100 maddeli test için, 20 ortak maddenin kullanıldığı durumlarda, en iyi eşitlemenin yapıldığı gözlenmiştir.

Kaskowitz (1998) tarafından yapılan araştırmada, 5, 15 ve 25 maddelik ortak madde setlerinin kullanıldığı durumlarda ortak madde sayısının artması eşitleme hatasını düşürmüştür. Bastari (2000) tarafından yapılan bir başka araştırmada ortak madde setinde yer alan maddelerin sayısının artmasının eşitleme hatasını azalmasına neden olduğu saptanmıştır. Hanson ve Beguin (2002) tarafından yapılan araştırmada, 60 maddelik bir test için, 10 ve 20 maddelik ortak madde setleri kullanılarak test eşitleme hataları gözlenmiştir. Araştırma sonucunda, 20 maddelik ortak madde setinin kullanıldığı durumlarda test eşitleme hatasının daha düşük olduğu saptanmıştır.

Kim ve Cohen (2002) tarafından yapılan ve simülasyon verisinin kullanıldığı araştırmada ise, dereceli tepki modeli (graded response model) altında, 30 maddelik bir test için, ortak test uzunluğunun (10 madde - 20 madde) eşitleme hatasına etkisi araştırılmıştır. Araştırmada, 20 maddelik ortak madde setinin kullanıldığı durumlarda test eşitleme hatasının daha düşük olduğu saptanmıştır. Kaskowitz (1998), Bastari (2000), Kim ve Cohen (2002), Hanson ve Beguin (2002) tarafından yapılan araştırmaların ortak bulgusu ortak madde sayısı arttıkça eşitleme hatasının düştüğüdür.

Yukarıda belirtilen araştırmalar genel olarak incelendiğinde, ortak maddeli

desen kullanıldığında, ortak madde sayısının artmasının eşitleme hatalarını düşürdüğü şeklinde genel bir sonuç çıkarılabilir (Bastari, 2000; Hanson ve Beguin, 2002; Kaskowitz, 1998; Kim ve Cohen, 2002). Ancak bu sonuçla çelişen sonuçlar vermiş araştırmalar da bulunmaktadır (Cohen ve Kim, 1998; Hills, Subhiyah ve Hirsch, 1988; Suanthong, 1998). Yukarıda belirtilen araştırmalar, farklı testler, farklı eşitleme yöntemleri, farklı örneklemeler ve farklı bilgisayar programları kullanılarak gerçekleştirildiğinden dolayı, bu araştırmalardan elde edilen bulguların tüm durumlara genellenebilmesi oldukça güçtür. Ortak maddelerin sayısına ilişkin olarak ortak sonuç, test madde sayısının % 20'sinden az olmamasıdır (Budescu, 1985; Cook ve Eignor, 1991; Jodoin ve Davey, 2003; Li, Lissitz ve Yang, 1999; Muraki, Hombo ve Lee, 2000; Skyes, Hou, Hanson ve Wang, 2002).

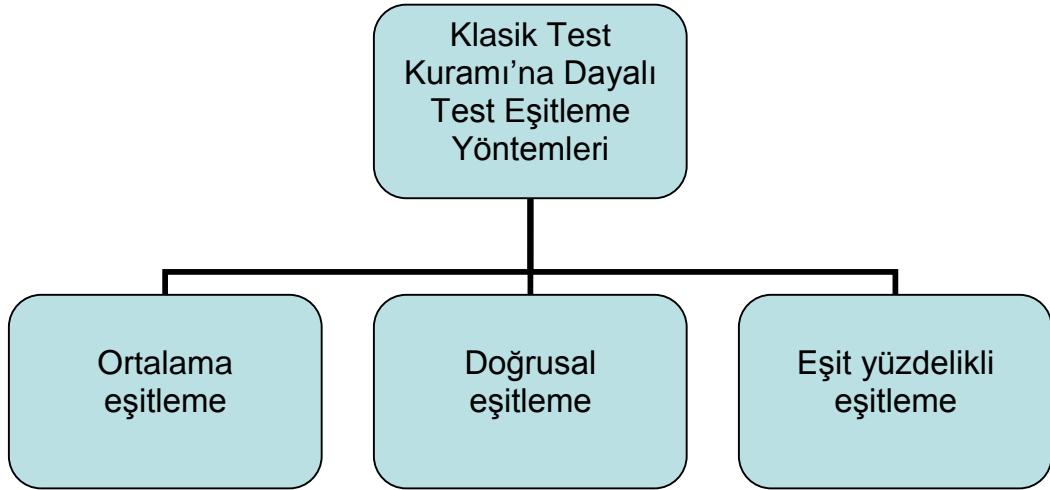
### **Test Eşitleme Yöntemleri**

Test eşitleme yöntemleri dayandığı kurama dayalı olarak ikiye ayrılmaktadır. Bunlar, Klasik Test Kuramı'na ve Madde Tepki Kuramı'na dayalı test eşitleme yöntemleridir. Aşağıda, her iki kurama dayalı eşitleme yöntemleri tanıtılmaktadır.

#### **Klasik Test Kuramı'na dayalı test eşitleme yöntemleri**

Klasik Test Kuramı'na dayalı test eşitleme yöntemleri üç kategoride ele alınabilir (Kolen, 2007).





Şekil 7. Klasik Test Kuramı'na Dayalı Test Eşitleme Yöntemleri

**Ortalama eşitleme.** Klasik Test Kuramı'na dayalı olarak geliştirilen test eşitleme yöntemlerinden en az parametre kullanılanı, ortalama eşitlemedir. Ortalama eşitlemede X ve Y bir testin iki farklı formu olmak üzere, X ve Y formunun ortalamaları arasındaki fark sabit kabul edilir. Eşitlenen puanlar aynı başarı düzeyine işaret etmektedir. Ortalama eşitleme Eşitlik 3 ve 4 ile ifade edilebilir (Kolen, 2007):

$$X - Ortalama(X) = Y - Ortalama(Y) \quad (3)$$

$$X = Y - Ortalama(Y) + Ortalama(X) \quad (4)$$

Örneğin, X formunun ortalaması 80; Y formunun ortalaması 75 olsun. Eşitlikte yerine konacak olursa;

$$X = Y - 75 + 80 \text{ olur.} \quad (5)$$

Y formundan alınan 80 puanın X formundaki karşılığı Eşitlik 5 kullanılarak bulunabilir.

$$X = 80 - 75 + 80$$

$$X = 85$$

**Doğrusal eşitleme.** Doğrusal eşitleme yöntemleri, bir testin paralel

formlarının güvenilirliğinin ve standart sapmasının eşit olduğu durumlarda kullanılır. İki test formundan elde edilen puanların standart sapmalarının eşit olduğu durumlarda birbirine denk olduğunu ifade etmektedir. Diğer bir deyişle, doğrusal eşitleme testi alan grupların yetenek düzeylerinin aynı olduğu durumlarda kullanılması uygundur. X ve Y formlarını alan grupların standart sapması ( $S_s$ ) ve ortalamaları hesaplanır. İki grubun puanları aşağıdaki eşitlik kullanılarak eşitlenir (Angoff, 1971):

$$\frac{Y - ortalama(Y)}{S_{S_y}} = \frac{X - ortalama(X)}{S_{S_x}} \quad (6)$$

Bu eşitlik tekrar düzenlenecek olursa;

$$Y = \underbrace{\frac{S_{S_y}}{S_{S_x}}}_A X + \underbrace{ortalama(Y) - \frac{S_{S_y}}{S_{S_x}} ortalama(X)}_B \Rightarrow Y = AX + B \quad (7)$$

Eşitlik 7'de, A, dönüşüm doğrusunun eğimi olmakla birlikte  $\frac{S_{S_y}}{S_{S_x}}$  oranına eşittir. B ise  $ortalama(Y) - \frac{S_{S_y}}{S_{S_x}} ortalama(X)$ 'e eşittir. Bu dönüşüm eşitliği regresyon eşitliğinin aksine simetriktir. Aynı eşitlik X formundan Y formuna ya da Y formundan X formuna puan dönüştürmede kullanılabilir (Angoff, 1971).

Eğer Y formu bir Z referans ölçeğine dönüştürülecekse yeni formül aşağıdaki gibi olacaktır:

$$Z = A'Y + B' \quad (8)$$

Eşitlik 8'de Y yerine  $AX+B$  konulduğunda;

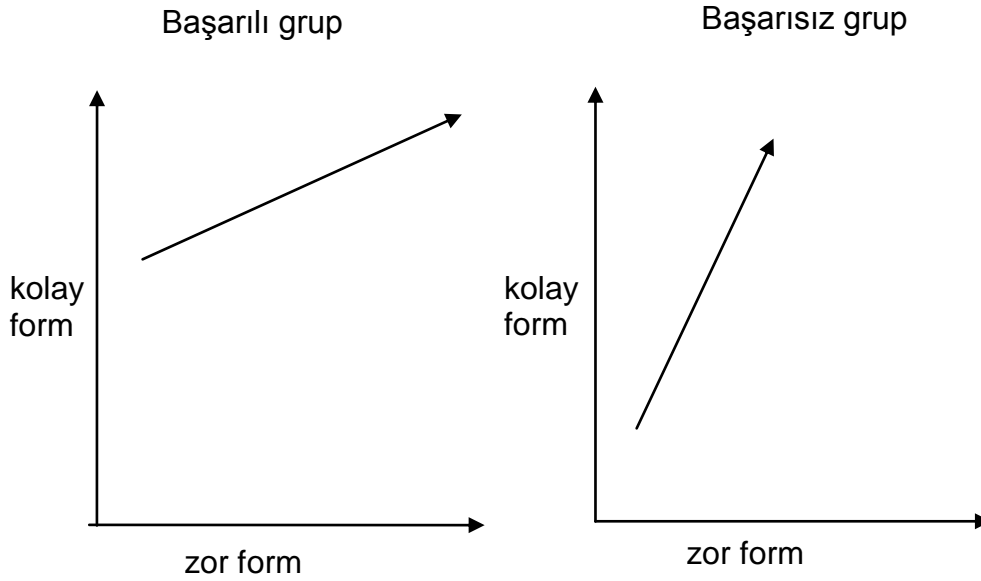
$$Z = A'(AX + B) + B' \text{ olur.} \quad (9)$$

Oluşan yeni denklem sayesinde Z puanlarının X puanlarına dönüşümü sağlanır. Oluşan denklemin eğimi ve kesişim noktası Eşitlik 10 üzerinde gösterilmektedir (Angoff, 1971).

$$Z = \frac{A'A}{A''}X + \frac{A'B + B'}{B''} \quad (10)$$

Ortalama eşitlemede sadece test formlarının ortalamaları kullanılırken; doğrusal eşitlemede ortalama ve standart sapmalar kullanılır. Eşitlenecek formların standart sapmaları eşitse, ortalama eşitleme ve doğrusal eşitleme yöntemleri aynı sonucu verir (Kolen, 1988).

Doğrusal eşitleme yönteminin bir takım sınırlılıkları da bulunmaktadır. Örneğin, bir testin iki farklı formundan elde edilen puanların ranji değişiyorsa, bir formdaki puanın diğer formdaki karşılığı bulunmayabilir. Örneğin 100 maddelik iki test formu doğrusal eşitleme ile eşitlendiğinde, zor olan X formundan 99 puan alan bireyin, kolay olan Y formundaki puan karşılığı 103 puan olarak hesaplanabilir. Doğrusal eşitlemenin bir diğer sınırlılığı da eşitleme sonuçlarının gruba bağımlı olmasıdır (Livingston, 2004). Aşağıdaki şekilde farklı yetenek düzeylerine sahip iki gruba ait eşitleme doğruları gösterilmektedir.

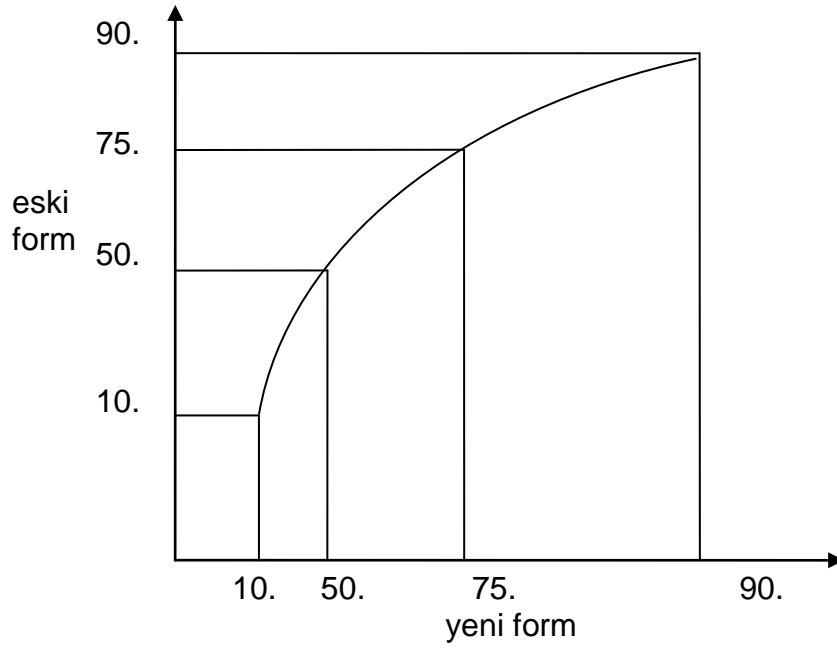


Şekil 8. Başarılı ve Başarısız Grupta Doğrusal Eşitleme.  
(Livingston (2004)'ten uyarlanmıştır)

Şekil 8 incelendiğinde başarılı gruba ait doğrusal eşitleme denkleminin, başarısız gruptan elde edilen denklemden oldukça farklı olduğu görülmektedir. Bu durum test eşitlemenin gruptan bağımsızlık özelliğini zedelemektedir.

**Eşit yüzdellikli eşitleme.** Klasik Test Kuramı'na dayalı test eşitleme yöntemlerinden biri de eşit yüzdellikli eşitlemedir. Eşit yüzdellikli eşitlemede, her bir formdaki maddeler yüzdellik sırasına çevrilirler. Sonra, iki formdaki eşdeğer puanlar, ilk formdaki p. yüzdelliği, ikinci formdaki p. yüzdellikteki puana eşitleyerek hazırlanır (Aiken, 2000).

Eşit yüzdellikli eşitlemede, aynı yüzdellik sırasındaki puanlar belli bir grup için aynı düzeydeki performans düzeyleri olarak düşünülmektedir. Diğer bir deyişle, A ve B bir testin iki farklı formu olmak üzere; A formunun puan dağılımının belli bir yüzdesindeki puan; B formunun puan dağılımının aynı yüzdesine denk gelen puana eşittir. Örneğin, 50. yüzdellik dilimde A formundaki 20 puan, B formundaki 50. yüzdellik dilime denk gelen 17 puana eşitlenebilmektedir (Kolen, 1988). Şekil 9'da eşit yüzdellikli eşitlemede puan dönüşümü bir örnek üzerinde gösterilmektedir.



Şekil 9. Eşit Yüzdellikli Eşitleme.  
(Livingston (2004)'ten uyarlanmıştır)

Şekil 9 incelendiğinde, eski ve yeni test formundan elde edilen puanlar arasında eğrisel bir ilişkinin olduğu görülmektedir. Her bir puan yüzdesine karşılık gelen puan çiftlerinin bulunmasıyla test formları eşitlenmektedir.

Ortalama eşitleme ve doğrusal eşitleme yöntemleri ile karşılaştırıldığında, eşit yüzdellikli eşitleme yöntemi daha büyük örneklemelere ulaşılmasını gerektirmektedir. Eğer testlere ait puan dağılımları benzerse, doğrusal eşitleme ve eşit yüzdellikli eşitleme aynı sonucu verir (Felan, 2002; Kolen, 1988). Küçük örneklemelerde, test formlarına ait dağılım düzensizlik gösterebilir. Daha büyük örneklemelerde bu dağılımın daha düzenli olacağı varsayılır. Bu durumda puanlar düzeltilerek (smoothing) kullanılır. Böylece daha düzenli dağılan bir dağılım elde edilerek puanlar eşitlenir (Livingston, 2004).

### **Madde Tepki Kuramı'na dayalı test eşitleme yöntemleri**

Madde Tepki Kuramı'na göre, bireylerin test maddesi üzerindeki performansı ile bu performansının altında yatan yetenek düzeyi arasındaki

ilişki monotonik artan fonksiyonla tanımlanabilir. Bu fonksiyona *madde karakteristik fonksiyonu* adı verilir. Bu fonksiyona göre yetenek arttıkça maddeyi yanıtlama olasılığı artar (Hambleton, Swaminathan ve Rogers, 1991).

Madde Tepki Kuramı'nda madde ve yetenek parametreleri aynı ölçek üzerinde yer almaktadır. Örneğin Madde Tepki Kuramı'na göre kestirilen yetenek düzeyi 0,05 olsun. Araştırmacı, herhangi örtük özellik dağılımını referans almadan, bu bireyin ortalama ve ortalamanın altında yer alan güçlüklerdeki maddeleri çözme şansının %50'den fazla olduğunu ifade edebilir. Bu bireyin ortalamanın üstünde yer alan güçlük düzeyindeki problemleri çözme şansı da düşüktür (Embretson ve Reise, 2000).

Madde Tepki Kuramı'nda bir kişinin bir testten aldığı puanlar matematiksel modeller kullanılarak belirlenmektedir. Madde Tepki Kuramı'nda kullanılan modeller temel olarak ikiye ayrılmaktadır. Bunlar tek boyutlu Madde Tepki Kuramı modelleri ve çok boyutlu Madde Tepki Kuramı modelleridir. Tek boyutlu Madde Tepki Kuramı modelleri maddelerle tek bir örtük özelliği ölçmektedir. Çok boyutlu Madde Tepki Kuramı modelleri ise birden fazla örtük özelliğin ölçülmesinde kullanılır. Madde Tepki Kuramı modelleri test için kullanılan puanlama türüne göre; 1-0 şeklinde puanlanan modeller (binary or dichotomous) ve çoklu kategorilerde puanlanan modeller (polytomous) olmak üzere ikiye ayrılabilir (Ackerman, 1992, 1994,1996; Embretson, 1996, 1997; Embretson ve Reise, 2000; Reckase, 1985, 1997).

Madde Tepki Kuramı'nda yukarıda belirtilen modellerin kullanılması için çeşitli varsayımların karşılanması gerekmektedir. Tek boyutlu Madde Tepki Kuramı çerçevesindeki temel varsayım, testle ölçülen yeteneğin tek olmasıdır. Bu özellik "tek boyutluluk" varsayımı olarak adlandırılır. Bu varsayım tam olarak karşılanmayabilir. Çünkü bilişsel ve kişisel faktörler test performansını etkiler. Kişisel faktörler; motivasyon düzeyi, test kaygısı, hızlı çalışma becerisi ve tahmine giderek yanıt verme eğilimi olabilir. Bilişsel beceriler baskın olan yeteneğe işaret etmektedir. Eğer çalışmada birden çok

faktör söz konusu ise çok boyutlu Madde Tepki Kuramı modelleri kullanılması önerilmektedir (Hambleton, Swaminathan ve Rogers, 1991).

Diğer bir varsayım ise, yerel bağımsızlık varsayımdır. Yerel bağımsızlık testle ölçülen özellik sabit tutulduğunda bir bireyin iki maddeye verdiği tepkinin istatistiksel olarak birbirinden bağımsız olması demektir. Diğer bir deyişle, farklı maddelere verilen tepkilerin arasında ilişkinin bulunmaması anlamına gelir. Yerel bağımsızlık ve tek boyutluluk ilişkili kavramlardır (Hambleton, Swaminathan ve Rogers, 1991). Bir testteki maddeler evrenin tamamı için istatistiksel olarak bağımlı ise test tek boyutludur. Tek bir boyut varsa, gizil yetenek bakımından homojen olan her bir alt evrende maddeler istatistiksel olarak bağımsızdır. Bu bağımsızlık gizil yetenek ölçeğindeki belli bir yetenek düzeyinde bulunan katılımcılardan oluşan bir alt evren için tanımlanır ve yerel bağımsızlık olarak adlandırılır (Crocker ve Algina, 1986).

Madde Tepki Kuramı'ndaki varsayımlar karşılandığında ve model veri uyumu sağlandığında, kişilerin yetenek parametreleri madde havuzundan alınan herhangi bir madde setinden bağımsız olarak kestirilebilmektedir. Diğer bir deyişle bir kişinin zor ya da kolay testi alması sorun değildir. Aynı şekilde madde parametreleri de aynı evreden gelen herhangi bir alt gruptan bağımsız olarak kestirilebilir. Maddelerin karşılaştırılmasındaki değişmezlik, maddeler arasındaki farklılıkların kişilere bağlı olmadığı anlamına gelmektedir. Buna Madde Tepki Kuramı'nın değişmezlik (invariance) özelliği de denilmektedir (Cook ve Eignor, 1991; Embretson ve Reise, 2000).

Tek boyutluluk ve yerel bağımsızlık varsayımlarının karşılanmasıyla Madde Tepki Kuramı'na dayalı modeller kullanılabilir. Madde Tepki Kuramı'na ilişkin olarak 1-0 şeklinde puanlanan verilerde kullanılmak üzere 3 model geliştirilmiştir. Bunlar;

- 1 parametrelili lojistik model,
- 2 parametrelili lojistik model,
- 3 parametrelili lojistik modeldir.

**1 parametrelili lojistik model.** Bu model, bir madde parametresini barındırdığı için 1 parametrelili model (1PL) olarak bilinir. Maddeler sadece güçlük bakımından farklılaşır. Ayırıcılık parametresi tüm maddeler için eşit kabul edilir. Maddeye doğru cevap verme olasılığı 0.50 olduğunda, eğri bükülme noktasından yön değiştirmeye başlar (Crocker ve Algina, 1986; Embretson ve Reise, 2000; Hambleton ve Swaminathan, 1989). 1 parametrelili modelde s kişisinin i maddesine doğru cevap verme olasılığı:

$X_{is}$  : s bireyinin i maddesine verdiği cevap (1 veya 0)

$\theta_s$  : s kişisinin yetenek düzeyi

$b_i$  : i maddesinin güçlüğü

$D$  : 1,7 olmak üzere;

$$P(X_{is} = 1 | \theta_s, b_i) = \frac{e^{D(\theta_s - b_i)}}{1 + e^{D(\theta_s - b_i)}} \quad \text{şeklinde ifade edilebilir.} \quad (11)$$

**2 parametrelili lojistik model.** İki parametrelili modelde, maddeye doğru cevap verme olasılığının hesaplanmasında güçlük parametresinin yanı sıra madde ayırıcılık parametresi de kullanılır. Bu durumda, s kişisinin i maddesine doğru cevap verme olasılığı:

$X_{is}$  : s bireyinin i maddesine verdiği cevap (1 veya 0)

$\theta_s$  : s kişisinin yetenek düzeyi

$b_i$  : i maddesinin güçlüğü

$a_i$  : i maddesinin ayırt ediciliği

$D$  : 1,7 olmak üzere;

$$P(X_{is} = 1 | \theta_s, a_i, b_i) = \frac{e^{Da_i(\theta_s - b_i)}}{1 + e^{Da_i(\theta_s - b_i)}} \quad \text{şeklinde ifade edilebilir.} \quad (12)$$

Eşitlik 12'de, 1 parametrelili lojistik modelden farklı olarak farklı ayırıcılığı sahip maddelerin ayırıcılığını gösteren  $a_i$ 'nin bulunmasıdır. Maddelerin eşit



olarak örtük özellik ile ilişkili olmadığı durumlarda, iki parametrelili lojistik modelin kullanılması uygundur (Crocker ve Algina, 1986; Embretson ve Reise, 2000; Hambleton ve Swaminathan, 1989).

**3 parametrelili lojistik model.** 3 parametrelili modelde iki parametrelili modelden farklı olarak, şans parametresi de hesaba katılarak maddeyi doğru cevaplama olasılıkları hesaplanır. Eklenen bu parametre nedeniyle madde karakteristik eğrileri 0 noktasından başlamamaktadır. Örneğin çoktan seçmeli bir testte az bilenlerin de maddeye doğru cevap verme olasılığı bulunduğundan, madde karakteristik eğrisi 0 noktasının üzerinden başlamaktadır. Üç parametrelili lojistik modelde, s bireyinin i maddesine doğru cevap verme olasılığı, madde güçlük, ayırtıcılık ve şans parametresi hesaba katılarak aşağıdaki Eşitlik 13 ile bulunur (Crocker ve Algina, 1986; Embretson Reise, 2000; Hambleton ve Swaminathan, 1989):

$X_{is}$  : s bireyinin i maddesine verdiği cevap (1 veya 0)

$\theta_s$  : s kişinin özellikteki düzeyi

$b_i$  : i maddesinin güçlüğü

$a_i$  : i maddesinin ayırt ediciliği

$c_i$  : i maddesi için şans başarısı

$D$  : 1,7 olmak üzere

$$P(X_{is} = 1 | \theta_s, a_i, b_i, c_i) = c_i + (1 - c_i) \frac{e^{Da_i(\theta_s - b_i)}}{1 + e^{Da_i(\theta_s - b_i)}} \text{ şeklinde ifade edilebilir. (13)}$$

Madde Tepki Kuramı'na dayalı test eşitleme yöntemleri parametrelerin kestirimine dayalı olarak iki kategoride ele alınabilmektedir. Bunlar *eşzamanlı* (concurrent) eşitleme yöntemi ve parametrelerin *ayrı* kestirildiği (separate) eşitleme yöntemleridir. Eşzamanlı eşitleme yönteminde ortak maddelere sahip iki teste ait madde parametreleri aynı anda kestirilir. İki teste ait parametrelerin ayrı kestirildiği durumda ise test eşitleme çalışması iki aşamadan oluşur. Önce, ortak maddelere sahip iki teste ait madde parametreleri, kullanılan bilgisayar yazılımı üzerinde iki ayrı analizle kestirilir.

Kestirilen parametreleri iki farklı ölçeğe aittir ve aynı ölçek üzerinde gösterilemezler. İkinci aşamada ayrı ayrı kestirilen madde parametrelerini aynı ölçek üzerine yerleştirmeyi sağlayacak denklem elde edilir. Bu denklem bir ölçekten alınan puanın diğer ölçekteki karşılığını bulmaya yarayan doğrusal dönüşüm denklemdir (Hanson ve Beguin, 2002; Kim ve Cohen, 2002; Kolen ve Brennan, 2004; Speron, 2009; Vale, 1986).

Madde Tepki Kuramı'na dayalı olarak yapılan test eşitleme çalışmalarında, testlerden alınan puanların aynı ölçekte yer almasını sağlamak için ortak maddeler kullanılır. Söz konusu denklem aşağıda verilmiştir (Cook ve Eignor, 1991).

$$\theta^* = A\theta + B \quad (14)$$

$\theta$  : Kişinin yetenek düzeyi;

$\theta^*$ : Kişinin, girmedığı sınavdan alacağı kuramsal yetenek düzeyi

A : Eşitleme denkleminin eğimi

B : Eşitleme denkleminin sabiti.

Burada A, doğrusal dönüştürmenin eğimi, B ise sabittir. Benzer olarak güçlük, şans ve ayırıcılık parametreleri de doğrusal dönüştürülür (Cook ve Eignor, 1991):

$$a_i^* = \frac{1}{A} a_i \quad (15)$$

$$b_i^* = Ab_i + B \quad (16)$$

$$c_i^* = c_i \quad (17)$$

$a$  : Ayırıcılık parametresi

$a_i^*$  : Dönüştürülmüş ayırıcılık parametresi

$b$  : Güçlük parametresi

$b_i^*$  : Dönüştürülmüş güçlük parametresi

$c$  : Şans parametresi

$c_i^*$  : Dönüştürülmüş şans parametresi

A : Eşitleme denkleminin eğimi

B : Eşitleme denkleminin sabiti

Eşitleme yöntemleri ile eşitleme denklemine ait eğim (A) ve sabit (B) bulunur. Madde Tepki Kuramı'na dayalı olarak ortak maddeli test eşitleme yönteminin kullanıldığı durumlarda, ölçek dönüşümü için çeşitli yöntemler geliştirilmiştir. Literatürde sıklıkla kullanılan test eşitleme yöntemleri aşağıda açıklanmaktadır.

### Moment yöntemleri

Moment yöntemleri literatürde ortalama-ortalama ve ortalama-standart sapma yöntemleri olmak üzere iki kategoride ele alınmaktadır. Aşağıda bu yöntemler açıklanmaktadır.

**Ortalama-ortalama yöntemi.** Loyd ve Hoover (1980) tarafından tanımlanan ve güçlük parametresinin yanında ayıricılık parametresinin ölçek puan dönüşümünde kullanıldığı yöntemdir. İ ve J ölçeklerine ilişkin olarak ayıricılık ve güçlük parametrelerine dayalı ortalamalarla oluşturulan formül aşağıdaki gibidir (Kolen ve Brennan, 2004) :

$$A = \frac{\mu(\alpha_i)}{\mu(\alpha_j)} \quad (18)$$

$$B = \mu(b_j) - A\mu(b_i) \quad (19)$$

$\mu(\alpha_i)$  : İ ölçeğine ait ayıricılık parametrelerinin ortalaması

$\mu(\alpha_j)$  : J ölçeğine ait ayıricılık parametrelerinin ortalaması

$\mu(b_i)$  : İ ölçeğine ait güçlük parametrelerinin ortalaması

$\mu(b_j)$  : J ölçeğine ait güçlük parametrelerinin ortalaması

A : Eşitleme denklemindeki eğim

B : Eşitleme denklemindeki sabit

Eşitlik 18 ve 19 incelendiğinde, test eşitlemede kullanılacak olan denklemin eğiminin, İ testine ait ayırıcılık değerinin ortalamasının, J testine ait ayırıcılık değerinin ortalamasına oranı olduğu görülmektedir. Denklemden kullanılacak denklemin sabiti (B) ise, J testinden elde edilen güçlük ortalamasının, İ testinden elde edilen güçlük ortalamasının eğimle çarpımından elde edilen değerden farklıdır.

**Ortalama-standart sapma yöntemi.** Marco (1977) tarafından tanımlanan dönüştürme sürecidir. Bu yöntemde, A ve B'nin hesaplanmasında güçlük parametresinin standart sapması ve ortalaması kullanılır. İ ve J ölçeklerine ilişkin güçlük parametresine ait ortalama ve standart sapma değerleri ile oluşturulan formül aşağıdaki gibidir (Kolen ve Brennan, 2004) :

$$A = \frac{\sigma(b_j)}{\sigma(b_i)} \quad (20)$$

$$B = \mu(b_j) - A\mu(b_i) \quad (21)$$

$\sigma(b_i)$  : İ ölçeğine ait güçlük parametrelerinin standart sapması

$\sigma(b_j)$  : J ölçeğine ait güçlük parametrelerinin standart sapması

$\mu(b_i)$  : İ ölçeğine ait güçlük parametrelerinin ortalaması

$\mu(b_j)$  : J ölçeğine ait güçlük parametrelerinin ortalaması

A : Eşitleme denklemindeki eğim

B : Eşitleme denklemindeki sabit

Eşitlik 20 ve 21 incelendiğinde, test eşitlemede kullanılacak olan denklemin eğiminin (A), J testine ait güçlük parametresinin standart sapmasının, İ testine ait güçlük parametresinin standart sapmasına oranı olduğu görülmektedir. Eşitlikte kullanılacak denklemin sabiti (B) ise, ortalama-ortalama yönteminde olduğu gibi, J testinden elde edilen güçlük

ortalamasının,  $\bar{I}$  testinden elde edilen güçlük ortalamasının eğimle çarpımından elde edilen değerden farkıdır. Ortalama-standart sapma yönteminin ortalama-ortalama yönteminden farkı, denklem eğiminin hesaplanış biçimidir. Ortalama-ortalama yönteminde denklem eğiminin saptanmasında ayırıcılık parametresinin ortalaması; ortalama-standart sapma yönteminde ise güçlük parametresinin standart sapması kullanılmaktadır.

### Karakteristik eğri yöntemleri

Karakteristik eğri yöntemleri iteratif bir süreçte gerçekleşir ve üç basamaktan oluşmaktadır:

- Ölçek transformasyon sabitlerinin kestirimi,
- Test (veya madde) karakteristik eğrileri arasındaki farkın hesaplanması,
- Bu farklılıkların minimize edilip edilmediğinin saptanması.

Karakteristik eğri yöntemlerinde iki yaklaşım bulunmaktadır. Bunlar, Haebara yaklaşımı ile Stocking-Lord yaklaşımıdır.

**Haebara Yaklaşımı.** Bu fonksiyon, yönteme adını veren Haebara (1980) tarafından geliştirilmiştir. Belli bir yetenek düzeyindeki katılımcılar için, madde karakteristik eğrileri arasındaki fark, her bir maddeye ait madde karakteristik eğrileri arasındaki farkın karelerinin toplamıdır. (Kolen ve Brennan, 2004; Raju ve Arenson, 2002). Bu yöntemin matematiksel ifadesi Eşitlik 22'de verilmektedir.

$$Hdiff(\theta_i) = \sum_{j \in V} \left[ p_{i_j}(\theta_{j_i}; a_{j_j}, b_{j_j}, c_{j_j}) - p_{i_j}(\theta_{j_i}; \frac{a_{i_j}}{A}, Ab_{i_j} + B, c_{i_j}) \right]^2 \quad (22)$$

$$Hcrit = \sum_i Hdiff(\theta_i) \quad (23)$$

A : Eşitleme denkleminin eğimi

B : Eşitleme denkleminin sabiti

$p_{ij}(\theta_{ji}; a_{ji}, b_{ji}, c_{ji})$ : Madde karakteristik fonksiyonu

$p_{ij}(\theta_{ji}; \frac{a_{ji}}{A}, Ab_{ji} + B, c_{ji})$ : Eşitlenmiş madde karakteristik fonksiyonu

**Stocking-Lord Yaklaşımı.** Stocking-Lord (1983) Haebara yaklaşımından farklı olarak, belli bir yetenek düzeyindeki katılımcılar için, madde karakteristik eğrileri arasındaki fark, her bir maddeye ait madde karakteristik eğrileri arasındaki farkın toplamının karesidir (Kolen ve Brennan, 2004). Modelin matematiksel ifadesi Eşitlik 24'te verilmektedir.

$$SLdiff = \left[ \sum_{j:V} p_{ij}(\theta_{ji}, a_{jj}, b_{jj}, c_{jj}) - p_{ij}(\theta_{ji}; \frac{a_{ij}}{A}, Ab_{ij} + B, c_{ij}) \right]^2 \quad (24)$$

$$SLcrit = \sum_i SLdiff(\theta_i) \quad (25)$$

A: Eşitleme denkleminin eğimi

B: Eşitleme denkleminin sabiti

$p_{ij}(\theta_{ji}; a_{ji}, b_{ji}, c_{ji})$ : Madde karakteristik fonksiyonu

$p_{ij}(\theta_{ji}; \frac{a_{ji}}{A}, Ab_{ji} + B, c_{ji})$ : Eşitlenmiş madde karakteristik fonksiyonu

Haebara ile Stocking-Lord yaklaşımlarında  $H_{crit}$  (Eşitlik 23) ve  $SL_{crit}$  (Eşitlik 25) değerlerini en aza indiren A ve B değerleri saptanmaya çalışılır. Bu süreçte, ortalama-ortalama ya da ortalama-standart sapma yöntemleriyle elde edilen A ve B değerleri başlangıç değerleri olarak minimizasyon çalışmaları için formülde yerine konulur. Karakteristik eğri yöntemi altında yer alan, Haebara ile Stocking-Lord tarafından tanımlanan iteratif minimizasyon algoritmaları bilgisayar programları yardımıyla gerçekleştirilmekte, A ve B değerleri kolaylıkla hesaplanabilmektedir (Kolen ve Brennan, 2004). ST

(Hanson ve Zeng, 2004), POLYST (Kim ve Kolen, 2003), STUIRT (Kim ve Kolen, 2004), PİE (Hanson ve Zeng, 1995) ve IRTEQ (Han, 2007b) bu programlardan bazılarıdır.

### Test Eşitlemede Hata Kavramı

Test eşitlemede hata kavramı bireyin yetenek düzeyi ile almadığı test için kestirilen yetenek düzeyi arasındaki farkla açıklanmaktadır. Hatadan arınık bir eşitlemede farklı testlerle elde edilen yetenek düzeylerinin eşit olması beklenir. Bu durum Eşitlik 26 ile açıklanabilir (Cook ve Eignor, 1991):

$$a_i * (\theta_a * -b_i *) = \frac{1}{A} a_i (A\theta_a + B - [Ab_i + B]) \quad (26)$$

$$= \frac{1}{A} a_i (A\theta_a + B - Ab_i - B)$$

$$= \frac{1}{A} a_i A(\theta_a - b_i)$$

$$a_i * (\theta_a * -b_i *) = a_i (\theta_a - b_i)$$

$$P(\theta) = P(\theta^*)$$

$a_i^*$  : Dönüştürülmüş ayırıcılık parametresi

$b$  : Güçlük parametresi

$b_i^*$  : Dönüştürülmüş güçlük parametresi

$c$  : Şans parametresi

$c_i^*$  : Dönüştürülmüş şans parametresi

$A$  : Eşitleme denkleminin eğimi

$B$  : Eşitleme denkleminin sabiti

$P(\theta)$  : Bireyin maddeyi doğru yanıtlama olasılığı

$P(\theta^*)$ : Bireyin parametreleri dönüştürülmüş maddeyi doğru yanıtlama olasılığı

Bireylere uygulanan test formundan kestirilen yetenek düzeyi ile bireyin almadığı test formuna ilişkin dönüştürülmüş yetenek düzeyinin birbirine eşit çıkmaması eşitleme hatasından kaynaklanmaktadır. Eşitleme hataları seçkisiz eşitleme hatası ve sistematik eşitleme hatası olarak ikiye ayrılmaktadır.

**Seçkisiz eşitleme hatası:** Bu eşitleme hatası evrenden örneklem alındığında söz konusu olan hata türüdür. Küçük örneklem gruplarının kullanılması durumunda bu sorunla karşılaşılmaktadır. Örneklemdeki birey sayısının artırılmasıyla bu hatalar azaltılabilir (Kolen, 1988).

**Sistematik eşitleme hatası:** Uygun test eşitleme yönteminin ve deseninin kullanılmadığı veya varsayımların ihlal edildiği durumlarda ortaya çıkmaktadır. Örneğin, tek grup kullanılarak yapılan test eşitleme çalışmalarında yorgunluk ve sıra etkisi kontrol edilmiyorsa sistematik hatanın varlığından söz edilebilir. Bunlara ek olarak ortak maddelere ait madde karakteristik eğrileri bir uygulamadan diğerine farklılaşıyorsa sistematik hata meydana gelebilir (Embretson ve Reise, 2000; Kolen, 1988).

Test eşitleme çalışmalarında kullanılabilecek birçok test eşitleme yöntemi bulunmaktadır. Bu yöntemlerle yapılan eşitlemeler yukarıda belirtilen nedenlerden dolayı bir miktar hata barındırmaktadır. Bu araştırmada, test eşitleme konusunda yapılmış araştırmalardan farklı olarak, 500 ve 1000 kişilik örneklem büyüklükleri, normal ve çarpık yetenek dağılımları ve test eşitleme yöntemleri birlikte ele alınmış; söz konusu değişkenler için test eşitleme hatalarının karşılaştırılması amaçlanmıştır. Pratikte test eşitleme çalışmaları yapacak kişiler için çeşitli örneklem büyüklüğü ile yetenek dağılımlarına uygun olan eşitleme yönteminin kesin olarak bilinmemesi nedeniyle, söz konusu koşullar altında bu yöntemlerin karşılaştırılması zorunluluğu doğmuştur. Bu bağlamda, araştırmamızın problemini farklı gruplara



uygulanmış bir testin farklı formlarının “ortalama-ortalama, ortalama-standart sapma, Haebara ve Stocking-Lord” eşitleme yöntemlerine göre eşitlenmesi ve bu eşitleme yöntemlerinden elde edilen sonuçların, örneklem büyüklüğü ve grupların yetenek düzeyi dağılımı değişkenlerine dayalı olarak karşılaştırılması oluşturmaktadır.

### **İlgili Araştırmalar**

Test eşitleme konusunda yapılan araştırmalar incelendiğinde, çok çeşitli açılardan bu konunun ele alındığı görülmüştür. Test eşitleme ile ilgili araştırmalar aşağıda verilmektedir.

#### **Klasik Test Kuramı Ve Madde Tepki Kuramı'na Dayalı Test Eşitleme Yöntemlerinin Eşitleme Hataları Bakımından Karşılaştırıldığı Araştırmalar**

Literatürde Klasik Test Kuramı ve Madde Tepki Kuramı'na dayalı eşitleme yöntemlerinin karşılaştırılmasına yönelik birçok araştırma yapılmıştır. Petersen, Cook ve Stocking (1983) tarafından yapılan araştırmada, Klasik Test Kuramı'na dayalı eşitleme yöntemleri ve Madde Tepki Kuramı'na dayalı eşzamanlı eşitleme yöntemi karşılaştırılmıştır. Araştırma sonucunda, Madde Tepki Kuramı'na dayalı eşitleme yönteminin diğer yöntemlere nazaran daha iyi sonuçlar verdiği saptanmıştır.

Caldwell (1984) tarafından yapılan araştırmada, ortak maddeli testler kullanılarak Klasik Test Kuramı'na ve Rasch modeline dayalı eşitleme yöntemleri karşılaştırılmıştır. Araştırmada kullanılan ortak maddeler orta ve uç güçlüklerde kümelenmişlerdir. Bu madde setlerinin kullanıldığı testlerden elde edilen eşitleme sonuçları kıyaslandığında, en iyi eşitleme sonuçları veren madde setinin orta güçlükte olduğu saptanmıştır. Eşitleme yöntemleri karşılaştırıldığında ise Rasch modelinin Klasik Test Kuramı'na dayalı eşitleme yöntemlerine nazaran daha iyi eşitleme sonuçları verdiği bulunmuştur.

Hills, Subhiyah ve Hirsch (1988) tarafından yapılan arařtırmada, Florida's Statewide Student Assessment Test Part II (SSAT-II) verileri kullanılmıřtır. 1984 ve 1986 yıllarında dzenlenen sınavlarda yer alan ortak maddeler sayesinde testlerin eřitlenmesi saęlanmıřtır. Arařtırmada, Klasik Test Kuramı'na dayalı doęrusal eřitleme, Madde Tepki Kuramı'na dayalı 1 ve 3 parametrelili lojistik modeller iin eřzamanlı eřitleme, 3 parametrelili lojistik model doęrusal eřitleme yontemleri karřılařtırılmıřtır. Arařtırma sonularına gre, testlerin glk ve gvenirliklerinin benzer olduęu durumlarda sz konusu beř eřitleme yonteminin benzer sonular verdięi saptanmıřtır. Arařtırmanın bir dięer bulgusu da ortak madde sayısı ile ilgilidir. Arařtırmada 10, 15, 20, 25 ve 30 maddelik hazırlanmıř ortak testlerden benzer eřitleme hataları elde edilmiřtir.

Kelecioęlu (1994) tarafından yapılan arařtırmada 1990, 1991 ve 1992 yılları arasında yapılan ęrenci Seme Sınavı (SS) puanlarını, Klasik Test Kuramı ve Madde Tepki Kuramı'na dayalı eřitleme yontemlerine gre eřitlemek amalanmıřtır. Bu arařtırma, Trkiye'de test eřitlemeye iliřkin olarak yapılan alıřmaların ilki olmuřtur. Arařtırmanın rneklemini ilgili yıllarda SS sınavına katılan 4000 aday oluřturmaktadır. SS testleri szde ortak (quasi anchor) maddelerle eřitlenmiřtir. Arařtırmanın sonucunda SS sınavlarının Trke ve Fen Bilimleri alt testleri iin en uygun yontemlerin Klasik Test Kuramı'na dayalı yontemler olduęu; Sosyal Bilimler ve Matematik alt testleri iin Madde Tepki Kuramı'na dayalı eřitlemenin uygun olduęu saptanmıřtır.

Yang (1997) tarafından, ortak maddeli eřitlenmemiř gruplar eřitleme deseninin kullanıldıęı arařtırmada, Klasik Test Kuramı'na ve Madde Tepki Kuramı'na dayalı eřitleme yontemleri karřılařtırılmıřtır. Arařtırmada ortak maddelerin glęnn testin glęn yansıtıęı durumlarda eřitlemenin daha doęru yapıldıęı sonucuna ulařılmıřtır. Bu bulgulara ek olarak, Madde Tepki Kuramı'na dayalı eřitleme yontemlerinin, dięer eřitleme yontemlerine gre daha hatasız sonular verdięi saptanmıřtır.

Suanthong (1998) tarafından yapılan, simlasyon verilerine dayalı

araştırmada, çeşitli faktörlerin eşitleme sonuçlarına etkileri araştırılmıştır. Bunlar, madde güçlüğü'nün standart sapması (.68, .95, .99), testteki madde sayısı (50,100, 200 madde), ortak madde sayısı (10, 20, 30 madde) ve örneklem büyüklüğüdür (100, 300, 500 kişi). Araştırmada ortak maddeli eşitlenmemiş gruplar üzerinde Rasch modeline dayalı eşitlemeler yapılmıştır. Bu eşitlemeler sonucunda, .99 güçlük sapmasına sahip, 100 maddeli test için, 20 ortak maddenin kullanıldığı 100 kişilik örneklemelerde, en iyi eşitlemenin yapıldığı gözlenmiştir.

Chen (2001) tarafından yapılan bir araştırmada, Iowa Tests of Basic Skills (ITBS) test bataryasının kısa ve uzun formu beş ilişkilendirme yöntemine göre ilişkilendirilmeye çalışılmıştır. Bu yöntemler Klasik Test Kuramı ve Madde Tepki Kuramı'na dayalı eşitleme yöntemleridir. Araştırma sonucunda en iyi eşitleme yöntemlerinin Madde Tepki Kuramı'na dayalı yöntemler olduğu saptanmıştır.

Şahhüseyinoğlu (2005) tarafından yapılan başka bir test eşitleme araştırmasında, 2000, 2001 ve 2002 yıllarında Hacettepe Üniversitesi'nce düzenlenen İngilizce Yeterlik Sınavı puanları üzerinde, Madde Tepki Kuramı ve Klasik Test Kuramı'na dayalı eşitleme yöntemleri karşılaştırılmıştır. Araştırma verilerini 2000 yılında sınava giren 2233; 2001 yılına sınava giren 1989 ve 2002 yılında sınava giren 2033 kişiye ait sınav sonuçları oluşturmuştur. Araştırma sonucunda Rasch modelinin ve Klasik Test Kuramı'na dayalı yöntemlerden eşit yüzdellik yönteminin, eşitleme için en uygun yöntemler olduğu bulunmuştur.

Bozdağ (2007) tarafından yapılan araştırmada, şans başarısının test eşitlemeye etkisi araştırılmıştır. Tek grup desende düzenlenen çalışmada klasik eşitleme yöntemlerinden, doğrusal eşitleme ve eşit yüzdellikli eşitleme yöntemleri kullanılmıştır. Araştırma verileri, 2005–2006 eğitim-öğretim yılında Mersin ili sınırları içinde yer alan farklı ilköğretim okullarında öğrenim gören 1031 8. sınıf öğrencisinden elde edilmiştir. Araştırmanın sonucunda, şanstın arındırılmamış test puanlarının eşitlenmesinde doğrusal eşitlemenin, şanstın arındırılmış test puanlarının eşitlenmesinde ise eşit yüzdellikli eşitlemenin

uygun yöntemler olduğu bulunmuştur. Tüm hatalar göz önünde bulundurulduğunda, şanstan arındırılmış eşit yüzdelikli eşitleme yönteminin en az hata ile sonuçlanan yöntem olduğu görülmüştür.

Cho (2007) tarafından yapılan bir araştırmada, çoklu kategorilerde puanlanan yazma testleri Klasik Test Kuramı'na dayalı eşit yüzdelikli eşitleme yöntemi ve Madde Tepki Kuramı'na dayalı yöntemlerle eşitlenmiştir. Araştırmada ortak maddeli eşitlenmemiş gruplar deseni kullanılmıştır. Araştırmada Klasik Test Kuramı'na dayalı eşit yüzdelikli eşitleme yönteminin, Madde Tepki Kuramı'na dayalı eşitleme yöntemlerine göre daha hatasız eşitlemeler yaptığı saptanmıştır.

Klasik Test ve Madde Tepki Kuramı'na dayalı eşitleme yöntemlerinin karşılaştırıldığı araştırmalarda çoğunlukla Madde Tepki Kuramı'na dayalı yöntemlerin daha az eşitleme hatasıyla sonuçlandığını saptayan araştırmaların olduğu görülmekle birlikte, Klasik Test Kuramı'na dayalı test eşitleme yöntemlerinin daha iyi sonuçlar verdiğini gösteren araştırmalar da bulunmaktadır.

### **Madde Tepki Kuramı'na Dayalı Test Eşitleme Yöntemlerinin Eşitleme Hataları Bakımından Karşılaştırıldığı Araştırmalar**

Madde Tepki Kuramı'na dayalı test eşitleme çalışmaları, bir testin farklı formlarına ait parametrelerin hem eş zamanlı kestirilmesiyle hem de ayrı ayrı kestirilmesiyle gerçekleştirilebilir. Hangi yöntemin daha az test eşitleme hatası verdiği, bir çok test eşitleme araştırmalarına konu olmuştur.

Literatürde eş zamanlı ve parametrelerin ayrı ayrı kestirildiği eşitleme yöntemlerinin karşılaştırılmasına yönelik araştırmalardan biri Wingersky, Cook, ve Eignor (1987) tarafından yapılmıştır. Bu araştırmada LOGIST (Wingersky, Barton ve Lord, 1982) programı kullanılarak parametrelerin eşzamanlı kestirildiği eşitleme yöntemi ile ayrı ayrı kestirildiği eşitleme

yöntemleri karşılaştırılmıştır. Bu araştırmanın sonunda eşzamanlı eşitleme yönteminin daha az eşitleme hatasına yol açtığı saptanmıştır.

Stocking ve Lord (1983) tarafından yapılan araştırmada Stocking-Lord yöntemi ile ortalama-standart sapma yöntemi karşılaştırılmıştır. Araştırma sonucunda Stocking-Lord yöntemin daha iyi sonuçlar verdiği saptanmıştır. Baker ve Al-Karni (1991) tarafından yapılan araştırmada karakteristik eğri yöntemleri ile ortalama-ortalama yöntemi karşılaştırılmıştır. Araştırmanın sonucunda karakteristik eğri eşitleme yöntemlerinin daha az eşitleme hatası verdiği saptanmıştır. Bu bulguya, Hung, Wu ve Chen (1991), Way ve Tang (1991) tarafından yapılan çalışmalarda da rastlanmıştır.

Cohen ve Kim (1998) tarafından yapılan araştırmada Madde Tepki Kuramı'na dayalı dereceli tepki modeli (graded response model) kullanılarak ortalama-ortalama, ortalama-standart sapma ve Stocking-Lord yöntemi karşılaştırılmıştır. Araştırmada simülatif veriler kullanılmıştır. Araştırma sonucunda, farklı koşullar altında yukarıda yer alan eşitleme yöntemlerinin benzer sonuçlar verdiği saptanmıştır.

Kim ve Cohen (1998) tarafından yapılan araştırmada 500'er kişilik gruplara 50 maddelik testler ve 4 ortak madde seti (5, 10, 25, ve 50 maddelik) uygulanmıştır. Araştırmada Stocking-Lord yöntemi ve eşzamanlı eşitleme yöntemi karşılaştırılmıştır. Araştırma sonunda, Stocking-Lord yönteminin az sayıda ortak madde içeren testlerin eşitlenmesinde daha iyi sonuçlar verdiği, daha fazla ortak madde içeren testlerin eşitlenmesinde ise bu yöntemlerin benzer sonuçlar verdiği sonucuna ulaşılmıştır.

Bastari (2000) tarafından yapılan araştırmada iki ve çok kategorili yanıtlanan soruların birlikte yer aldığı farklı test formlarının çeşitli deęişmeler açısından karşılaştırılması amaçlanmıştır. 1-0 şeklinde puanlanan veriler için 3 parametrelili model; çok kategorili veriler için ise yapılandırılmış tepki modeli kullanılmıştır. Parametre kestirimlerinde MULTLOG programı kullanılmıştır. Araştırmada test uzunluğu, farklı formatlardaki maddelerin oranı ve ortak test

uzunluđu test eşitleme yöntemlerini etkileyebilecek faktörler olarak düşünölmüştür. Araştırma sonunda, testteki ve ortak madde setindeki madde sayısının artmasının eşitleme hatalarını düşürdüđu saptanmıştır. İki ve çok kategorili maddelerin birlikte yer aldığı testlerde iki kategorili puanlanan maddelerin ađırlının artması eşitleme hatalarını düşüren diđer bir faktördür.

Tsai, Hanson, Kolen ve Forsyth (2001) tarafından yapılan bir araştırmada, ortak maddeli eşitlenmemiş gruplar deseninde, eşzamanlı eşitleme ile Stocking-Lord yöntemi karşılaştırılmıştır. 1493 ve 1793 kişiden oluşan eşitlenmemiş iki grup üzerinde yürütölen çalışmada, aynı içerikte, benzer istatistiksel özellikleri taşıyan ve ortak maddelere sahip iki test formu kullanılmıştır. Araştırma sonunda, eşzamanlı olarak yapılan eşitlemelerin az hatalı eşitleme sonucu verdiği saptanmıştır.

Hanson ve Béguin (2002) tarafından yapılan araştırmada, 60 maddeden oluşan, 3 parametrelili modele uyumlu 1-0 şeklinde puanlanan simölasyon verisi kullanılmıştır. Araştırmada, parametrelerin ayrı ayrı kestirildiđi eşitleme yöntemleri (Stocking-Lord, Haebara, ortalama-ortalama ve ortalama-standart sapma) ile eşzamanlı eşitleme (concurrent calibration) yöntemi karşılaştırılmıştır. Araştırmada 10'ar ve 20'şer maddelik ortak madde setleri kullanılmıştır. Araştırma sonucunda, tüm durumlar için eşzamanlı eşitleme yönteminin az hatalı eşitlemeler yaptığı saptanmıştır. Parametrelerin ayrı ayrı kestirildiđi yöntemler karşılaştırıldığında, Haebara ve Stocking-Lord yöntemlerinin, moment yöntemlerinden (ortalama-ortalama, ortalama-standart sapma) daha az eşitleme hatasıyla sonuçlandıđı görölmüştür. Moment yöntemleri birbiriyle kıyaslandıđında ise, ortalama-ortalama yönteminin ortalama-standart sapma yöntemine nazaran daha düşük eşitleme hatasına sahip olduđu gözlenmiştir. Büyük örneklemin kullanılmasının ve 10 maddelik ortak madde seti yerine 20 maddelik ortak madde setinin kullanılmasının eşitleme hatasını düşürdüđu de bulgular arasındadır.

Kim ve Cohen (2002) tarafından yapılan ve simölasyon verisinin

kullanıldığı araştırmada, dereceli tepki modeli (graded response model) altında, karakteristik eğri yöntemi ve eşzamanlı eşitleme yöntemi kullanılmıştır. Ortak madde sayısı (10-20) test eşitleme sonucunu etkileyen bir faktör olarak ele alınmıştır. Araştırma sonunda eşzamanlı kestirimin yapıldığı durumlarda, eşitleme hatalarının azaldığı sonucuna ulaşılmıştır. Ayrıca ortak madde sayısı arttıkça eşitleme hatalarının da azaldığı gözlenmiştir.

Kim ve Lee (2004) tarafından yapılan simülasyon araştırmasında üç parametrelili modele ve genelleştirilmiş kısmi kredi modeline (generalized partial credit model) göre oluşturulmuş karışık formattaki testler, karakteristik eğri yöntemleri (Stocking-Lord ve Haebara) ve moment yöntemlerine (ortalama-ortalama ve ortalama-standart sapma) dayalı olarak eşitlenmiştir. Araştırma sonucunda karakteristik eğri yöntemlerinin moment yöntemlerine nazaran daha az eşitleme sahip olduğu saptanmıştır.

Karkee ve Wright (2004) tarafından yapılan araştırmada, 31.813 öğrenciye ait test verileri kullanılarak 3 parametrelili modele göre kestirilmiş parametreler, 1 parametrelili modele göre elde edilen parametrelere eşitlenmiştir. Bu çalışmada, karakteristik eğri yöntemleri (Stocking-Lord ve Haebara) ile moment yöntemleri (ortalama-ortalama ve ortalama-standart sapma) karşılaştırılmıştır. Araştırma sonunda en hatasız eşitlemelerin Stocking-Lord yöntemine göre elde edildiği saptanmıştır. Ortalama-ortalama yöntemi ile ortalama-standart sapma yöntemleri karakteristik eğri yöntemlerine göre, daha yüksek eşitleme hatasına sahiptir.

Kim ve Kolen (2006) tarafından yapılan araştırmada, ikili ve çoklu kategorilerde puanlanan maddelerin birlikte kullanıldığı verilerde eşitleme yöntemleri karşılaştırılmıştır. Simülasyon verisi üzerinde yürütülen bu çalışmada, eşitleme yöntemleri olarak karakteristik eğri yöntemleri (Stocking-Lord ve Haebara), moment yöntemleri (ortalama-ortalama ve ortalama-standart sapma) ve eşzamanlı eşitleme yöntemi kullanılmıştır. Araştırma sonucunda, en iyi eşitleme sonuçlarını veren yöntemin eşzamanlı eşitleme

yöntemi olduğu bulunmuştur. Araştırmanın bir diğer bulgusu da, karakteristik eğri yöntemlerinin, moment yöntemlerine nazaran daha iyi eşitleme sonuçları verdiği'dir. En fazla eşitleme hatası ile sonuçlanan yöntem ise ortalama-ortalama yöntemidir.

Speron (2009) tarafından yapılan çalışmada, 2 parametrelî modele uyumlu veriler üzerinde, ortalama-ortalama, ortalama-standart sapma, Stocking-Lord ve Haebara yöntemleri karşılaştırılmaktadır. Araştırma sonunda, ortalama-ortalama, Stocking-Lord ve Haebara yöntemlerinin benzer sonuçlar verdiği bulunmuştur. Ortalama-standart sapma yönteminin ise en fazla eşitleme hatasına sahip olduğu saptanmıştır.

Lee ve Ban (2010) tarafından yapılan araştırmada eş zamanlı eşitleme yöntemi, Stocking-Lord yöntemi ve Haebara yöntemi karşılaştırılmıştır. Araştırma sonucunda, eşzamanlı eşitleme yönteminin diğer yöntemlere nazaran daha fazla eşitleme hatası verdiği saptanmıştır. Haebara ve Stocking-Lord yöntemleri karşılaştırıldığında, Haebara yönteminin Stocking-Lord yöntemine nazaran daha az eşitleme hatasına sahip olduğu saptanmıştır.

### **Test Eşitleme Hatalarının Örneklem Büyüklüğüne Göre Karşılaştırıldığı Araştırmalar**

Yotinprasert (1986) tarafından yapılan çalışmada, örneklem büyüklüğü açısından Klasik Test Kuramı'na ve Rasch modeline dayalı eşitleme yöntemleri karşılaştırılmıştır. Araştırma sonunda, örneklem büyüklüğü arttıkça eşitleme hatalarının azaldığı sonucuna ulaşılmıştır. Küçük örneklem kullanıldığında, Rasch yönteminin Klasik Test Kurama dayalı eşitleme yöntemine göre daha çok hata içerdiği saptanmıştır. Suanthong (1998) tarafından 1 parametrelî modele dayalı eşitlemelerin yapıldığı araştırmada ise, araştırmada kullanılan 100, 300, 500 kişilik örneklem büyüklüklerinden 100 kişilik örneklemde, en az hatayla eşitlemenin yapıldığı gözlenmiştir.



Bastari (2000) tarafından hem 1-0 verisinin hem de çoklu kategorilerde puanlanan verilerin bulunduğu karışık test formatında (mix format) yapılan arařtırmalarda 400, 1000 ve 3000 kiřilik örneklem büyüklüğünde gruplar oluşturulmuřtur. Arařtırma sonunda, örneklem büyüklüğü arttikça eřitleme hatalarının azaldığı gözlenmiřtir. Tate (2000) tarafından yapılan arařtırmada da benzer bir sonuç elde edilmiřtir.

Kim ve Cohen (2002) tarafından yapılan arařtırmada 300 ve 1000 kiřilik örneklem büyüklüklerinde eřitleme hataları incelenmiřtir. Arařtırma sonucunda, 1000 kiřilik örneklem büyüklüklerinden elde edilen eřitleme hatalarının, 300 kiřilik örneklem büyüklüklerine nazaran daha düşük olduđu saptanmıřtır. Harris (1993) ve Yang (1997) tarafından yapılan arařtırmalarda da benzer sonuçlar elde edilmiřtir.

Hanson ve Beguin (2002) tarafından yapılan arařtırmada, 1000 ve 3000 kiřilik örneklem grupları oluşturulmuřtur. Arařtırma sonunda, 3000 kiřilik örneklem büyüklüklerinden elde edilen eřitleme hatalarının, 1000 kiřilik örneklem büyüklüklerine nazaran daha düşük olduđu saptanmıřtır.

Çetin (2009) tarafından yapılan arařtırmada, 2005 ÖBBS sınavına giren 6., 7. ve 8. sınıf öğrencilerine ait veriler üzerinde dikey ölçeklemeler yapılmıřtır. Bu arařtırma sonucunda, 250, 500 ve 1000 kiřilik örneklem büyüklüklerinden, en az hata veren örneklem büyüklüğünün 1000 kiřilik olduđu saptanmıřtır.

Speron (2009) tarafından 2 parametrelili modele uyumlu veriler üzerinde yapılan arařtırmada örneklem büyüklüğü (250, 1000) test eřitleme hatalarını etkileyen bir deęiřken olarak ele alınmıřtır. Arařtırma sonucunda, 1000 kiřilik örneklem büyüklüklerinden elde edilen eřitleme hatalarının, 250 kiřilik örneklem büyüklüğünden elde edilenlere nazaran daha düşük olduđu saptanmıřtır.

Lee ve Ban (2010) tarafından yapılan test eřitleme arařtırmasında örneklem büyüklüğü 500 ve 3000 olmak üzere iki kategoride ele alınmıřtır.

Araştırma sonucunda 3000 kişilik örneklemelerden elde edilen eşitleme hatalarının, 500 kişilik örneklemelerden elde edilenlere nazaran daha düşük olduğu saptanmıştır.

### **Test Eşitleme Hatalarının Yetenek Dağılımına Göre Karşılaştırıldığı Araştırmalar**

Hanson ve Beguin (2002) tarafından yapılan araştırmada, 1000 ve 3000 kişilik örneklem büyüklüğüne sahip ortalaması 0, standart sapması 1 olan dağılım ve ortalaması 1 standart sapması 1 olan dağılıma sahip iki grubun eşitlenmesinden elde edilen eşitleme hataları karşılaştırılmıştır. Araştırma sonucunda, grupların yetenek dağılımının benzer olduğu koşullarda eşitleme hatalarını düştüğü saptanmıştır.

Kim ve Cohen (2002) tarafından dereceli tepki modelinin kullanıldığı eşitlemelerde de benzer bir sonuç elde edilmiştir. 300 ve 1000 örneklem büyüklüklerinde, ortalaması 0, standart sapması 1 olan dağılım ve ortalaması 1 standart sapması 1 olan dağılıma sahip iki grubun puanları eşitlenmiştir. Araştırmanın sonucunda benzer yetenek dağılımına sahip grupların eşitlenmesinden elde edilen eşitleme hatalarının, farklı yetenek dağılımına sahip grupların eşitlenmesinden elde edilen hatalara nazaran daha düşük olduğu saptanmıştır. Bastari (2000) ve Kim ve Lee (2006) tarafından yapılan araştırmalarda da benzer bir sonuç elde edilmiştir.

### **Amaç**

Bu araştırmanın genel amacı, farklı dağılım ve örneklem büyüklüğü koşulları altında, Madde Tepki Kuramı'na dayalı test eşitleme yöntemlerini (ortalama-ortalama, ortalama-standart sapma, Haebara ve Stocking-Lord) eşitleme hatalarına göre karşılaştırmak ve bu yöntemlerden en uygun olanını önermektir. Bu genel amaç doğrultusunda aşağıdaki sorulara yanıt aranmıştır:

1. 500 kişilik örneklem büyüklüğünde ve *benzer* yetenek dağılımına sahip gruplarda, 3 parametrelili lojistik modele uyumlu farklı test formlarının eşitlenmesinde en az hata veren test eşitleme yöntemi hangisidir?
2. 500 kişilik örneklem büyüklüğünde ve *farklı* yetenek dağılımına sahip gruplarda, 3 parametrelili lojistik modele uyumlu farklı test formlarının eşitlenmesinde en az hata veren test eşitleme yöntemi hangisidir?
3. 1000 kişilik örneklem büyüklüğünde ve *benzer* yetenek dağılımına sahip gruplarda, 3 parametrelili lojistik modele uyumlu farklı test formlarının eşitlenmesinde en az hata veren test eşitleme yöntemi hangisidir?
4. 1000 kişilik örneklem büyüklüğünde ve *farklı* yetenek dağılımına sahip gruplarda, 3 parametrelili lojistik modele uyumlu farklı test formlarının eşitlenmesinde en az hata veren test eşitleme yöntemi hangisidir?

### Önem

Test eşitleme, bir testin farklı formlarının bulunduğu durumlarda kullanılan bir süreçtir. Bu farklı test formları farklı öğrencilere eşzamanlı olarak veya farklı zamanlarda uygulanabilir. Farklı zamanlarda uygulanan sınavlarda, test sorularının gizliliğini sağlamak ve testin güvenilirliğini arttırmak açısından aynı sorular sorulmamaktadır. Farklı test formlarının uygulandığı durumlarda da test formlarının güçlüklerinin farklı olması öğrencileri değerlendirirken ve aralarında karşılaştırmalar yaparken bir takım problemlere sebep olabilmektedir. Örneğin, MEB tarafından hazırlanarak uygulanan Öğrenci Başarısını Belirleme Sınavı (ÖBBS)'nda, aynı sınıf düzeyindeki kişilere farklı test formları verilmekte ve öğrenciler farklı test formlarından aldıkları puanlara göre karşılaştırılmaktadır. Öğrencilerin kopya girişiminde bulunmalarını engellemek için gerçekleşen bu uygulama, öğrencilerin adil bir şekilde karşılaştırılmasına engel olabilir. Bu durum, katılımcılar için büyük önem taşıyan testlerden elde edilen sonuçlarının güvenilirliğini

zedelemektedir. Eğitim kararlarında kullanılan testlerin sonuçlarına dayalı değerlendirmenin adil olabilmesi için gereken çözüm yolu test eşitlemedir.

Test eşitleme çalışmalarında kullanılabilecek farklı eşitleme yöntemleri bulunmakla birlikte, bu yöntemlerle elde edilen test eşitleme sonuçları çeşitli değişkenlerden etkilenebilir. Bu araştırmada test eşitlemede kullanılan grupların yetenek dağılımı ve örneklem büyüklüğü test eşitleme hatalarında farklılık yaratabilecek değişkenler olarak ele alınmıştır. Yetenek dağılımı ve örneklem büyüklüğünün değiştiği koşullarda farklı test eşitleme yöntemlerinden elde edilen test eşitleme hatalarının karşılaştırılması ve en az hata ile sonuçlanan eşitleme yönteminin saptanması amaçlanmaktadır. Bu araştırmanın sonuçlarının, farklı test formlarının eşitlenmesinde, uygun eşitleme yönteminin seçimi için araştırmacılara, öğretmenlere ve geniş ölçekli test uygulamalarını yürüten ölçme ve değerlendirme kurumlarına/merkezlerine yararlı olacağı düşünülmektedir. Bu çalışmanın aynı zamanda hatadan en az arınık eşitleme sonuçları elde edebilmek için gerekli koşulları belirlemeye yönelik yapılan kuramsal araştırmalara da katkıda bulunacağı düşünülmektedir. Ayrıca, araştırmacının bu çalışmayı seçmesinin bir diğer amacı Türkiye’de özellikle ard arda yıllarda veya yılda birden çok defa uygulanan ve önemli kararlara (üniversiteye giriş, öğretmeliğe atanma, dil yeterliğini belirleme) dayanak olan test puanlarının adil olarak kullanımı için, test eşitleme çalışmalarının önemine dikkati çekebilmektir.

### **Sınırlılıklar**

Her araştırmada olduğu gibi, bu araştırmanın bir takım sınırlılıkları bulunmaktadır. Test eşitleme çalışmaları, Madde Tepki Kuramı’na ve Klasik Test Kuramı’na dayalı olarak yapılabilmektedir. Yapılan araştırmalar, Madde Tepki Kuramı’na dayalı eşitleme yöntemlerinin Klasik Test Kuramı’na dayalı eşitleme yöntemlerine göre daha az eşitleme hatası ile sonuçlandığını göstermektedir. Bu araştırma, test eşitleme çalışmalarında kullanılması önerilen Madde Tepki Kuramı’na dayalı test eşitleme yöntemleriyle sınırlıdır.

Ayrıca bu arařtırmada 1-0 řeklinde puanlanan Madde Tepki Kuramı modellerinden 3 parametrelili modeller ele alınmıř, diđer modeller arařtırma kapsamı dıřında bırakılmıřtır.

Madde Tepki Kuramı'na dayalı test eřitleme yöntemleri parametrelerin eřit zamanlı kestirimine dayanan test eřitleme yöntemleri ve parametrelerin ayrı ayrı kestirimine dayanan test eřitleme yöntemleri olmak üzere ikiye ayrılmaktadır. Bu arařtırma, parametrelerin ayrı ayrı kestirimine dayanan test eřitleme yöntemleri ile sınırlandırılmıřtır.

Bu arařtırma, on maddesi ortak olan elli maddelik test ile sınırlı tutulmuřtur. Literatürde, test eřitlemede kullanılan veriler, 1-0 řeklinde puanlanan veriler, çoklu kategorilerde puanlanan veriler ve 1-0 ile çoklu kategorilerde puanlamanın birlikte ele alındıđı karıřık formatta (mix format) yer alan veriler olarak sınıflandırılabilir. Bu arařtırma, gerçek test kořullarında en sık kullanılan veriler olması gerekçesiyle 1-0 řeklinde puanlanan verilerle sınırlandırılmıřtır.

Bu arařtırmada yetenek dađılım deđiřkeni, ortalaması 0 standart sapması 1 olan dađılım ile ortalaması 1 standart sapması 1 olan dađılım ele alınmıřtır. Örneklem büyüklüđü deđiřkeni ise, zaman, maliyet, ulařılabilirlik nedeniyle gerçek arařtırmalarda daha sık görüleceđi düşünölen 500 ve 1000 kiřilik örneklemlele sınırlı tutulmuřtur.

### Tanımlar

**Benzer yetenek dađılımı:** Her ikisi de 0 ortalama, 1 standart sapma ile dađılım gösteren veri seti.

**Farklı yetenek dađılımı:** Biri 0 ortalama, 1 standart sapma; diđer 1 ortalama, 1 standart sapma ile dađılım gösteren iki veri seti.

## **BÖLÜM 2**

### **YÖNTEM**

Bu bölümde, araştırmanın modeli, verilerin üretilmesi ve verilerin analizi açıklanmaktadır.

#### **Araştırmanın Modeli**

Bu araştırmada, örneklem büyüklüğü ve yetenek dağılımı değişkenlerine bağlı olarak, Madde Tepki Kuramı'na dayalı eşitleme yöntemleri karşılaştırılmaya çalışılmıştır. Böylece en az hatayı verecek eşitleme koşulları saptanarak, bu konuda yapılan kuramsal çalışmalara katkıda bulunulması amaçlanmıştır. Araştırma, bu yönüyle temel araştırma niteliği taşımaktadır.

#### **Verilerin Üretilmesi**

İstatistiksel problemlere ilişkin olarak simülasyon verisi, 1900'lü yılların başlarında kullanılmaya başlanmıştır. Bilgisayar teknolojilerindeki ilerlemelerle, istatistiksel problemleri çözmeye ilişkin popüler ve formal bir araştırma yöntemi haline gelmiştir (Harwell, Stone, Hsu ve Kirisci; 1996). Bu araştırmada test eşitleme yöntemlerine ait eşitleme hataları, örneklem büyüklüğü ve yetenek dağılımı değişkenlerine göre karşılaştırılmıştır. Araştırmada ele alınan yetenek dağılımı koşullarını sağlayabilecek örneklemi belirlemek gerçek verilerle güçtür. Bu nedenle araştırmada simülasyon programı ile üretilen simülatif veriler kullanılmıştır. Literatürde bu türden test eşitleme çalışmalarında (Bastari, 2000; Custer, Omar and Pomplun, 2006; Hanson ve Béguin, 2002; Han, Kolen ve Pohlmann, 1997; Hu, 2004; Kim ve Cohen, 2002; Kim ve Kolen, 2006; Kim ve Lee, 2004; Wang, Hanson ve Haris, 2000) simülatif verilerle çalışıldığı görülmektedir.

Literatürde GENIRV (Baker, 1988) ve WinGen2 (Han, 2007a, Han ve

Hambleton, 2007) isimli Madde Tepki Kuramı'na dayalı simülasyon verisi üreten programlar bulunmaktadır. WinGen2 bilgisayar programı farklı koşullarda, Madde Tepki Kuramı modellerine ilişkin olarak 1, 2, ve 3 parametrelili lojistik modeller, çok boyutlu modeller, çok kategorili modeller ve parametrik olmayan modeller için çeşitli örneklem büyüklüğü ve yetenek dağılımlarında veriler üretebilmektedir (Han, 2007a). Bu araştırmada, ücretsiz olarak edinilmesi nedeniyle, WinGen2 simülasyon veri üretme programından yararlanılmıştır.

Veri üretimi üç aşamada gerçekleştirilmiştir:

- Test formlarının yapısının tanımlandığı aşama
- Katılımcıların özelliklerinin tanımlandığı aşama
- Test formları ve katılımcı özelliklerinin birleştirilip 1-0 veri setlerinin oluşturulduğu aşama.

Veri üretme aşamalarının ilki test formlarının yapısının tanımlandığı aşamadır. Bu aşamada araştırmada kullanılacak Madde Tepki Kuramı modeline karar verilmiş, teste özgü madde sayısı ve ortak madde sayısı belirlenmiştir. Araştırmada, çoktan seçmeli testlerin sıklıkla kullanılması nedeniyle 3 parametrelili modeller araştırma kapsamına alınmış, 1 ve 2 parametrelili modeller araştırma kapsamı dışında bırakılmıştır. Bu nedenle araştırmada kullanılmak üzere 3 parametrelili lojistik modellere uyumlu simülatif test formları oluşturulmuştur. Araştırmada kullanılan test formları Çizelge 1'de verilmiştir:

Çizelge 1. Araştırmada Kullanılan Verilerin Yapısı

3 parametrelili modele dayalı üretilen verilerin madde sayısı	
Form K	50 madde (40 + 10 madde ortak)
Form L	50 madde (40 + 10 madde ortak)

Çizelge 1’de görüldüğü gibi, tek boyutluluk özelliği taşıyan K ve L adlarında iki adet test formu oluşturulmuştur. K ve L formu, aynı yapıyı ölçen bir testin farklı formları olarak tasarlanmış olup, 3 parametrelili lojistik modele uyumlu testlerdir.

K ve L formları 50’şer maddeden oluşmaktadır. Bu araştırma, ortak maddeli eşitlenmemiş gruplar eşitleme desenine dayalı olarak kurgulanmıştır. Bu nedenle, K ve L formlarında yer alan ilk 40 madde teste özgü maddeler olup her iki testte de madde parametre değerleri değişmektedir. Bu test formlarındaki veriler madde güçlükleri (b) bakımından ortalaması 0 olan madde dağılıma sahip olacak şekilde düzenlenmiştir. Ayırıcılık parametresi (a) 1-2 değerleri arasında değişmektedir. Şans parametresi (c) ise, 5 seçenekli testler için pratikte sık rastlanabilecek aralık olan 0,2 ile 0,3 değeri arasında değişmektedir. Araştırmada kullanılan test formlarına ait veriler benzer güçlük ve ayırt edicilik ortalamaları gösterdiğinden, bu test formları arasında yatay eşitlemeler yapılmıştır. Son on madde de ise madde parametreleri sabit tutularak, her iki formdan elde edilen puanların birbirine dönüşümünü sağlayacak olan “ortak madde seti” oluşturulmuştur. Farklı gruplar tarafından sabit parametre değerlerine sahip maddelere verilen tepkiler, farklı test formlarından alınan puanların birbirine dönüşümünü sağlayacak olan denklemin oluşturulmasında kullanılmıştır.

Her iki test formunda da yer alan 10 maddelik ortak madde seti, teste özgü olan 40 madde ile aynı özellikleri taşımaktadır. Ortak madde setine ait veriler madde güçlükleri (b) bakımından ortalaması 0 olan madde dağılıma sahip olacak şekilde düzenlenmiştir. Ayırıcılık parametresi (a) 1-2 değerleri arasında, şans parametresi (c) ise 0,2 ile 0,3 değeri arasında değişecek şekilde oluşturulmuştur. Bunun nedeni, literatürde önerildiği gibi, teste özgü maddelerle ortak madde setinde yer alan madde parametrelerinin benzer psikometrik özellikte olmasını sağlamaktır.

Veri üretme sürecinin birinci aşamasında oluşturulan test formlarına ilişkin madde parametre değerleri Çizelge 2 ve Çizelge 3’te verilmektedir.



Çizelge 2. K Formuna Ait Maddelerin Parametre Değerleri

<b>Madde no</b>	<b>a</b>	<b>B</b>	<b>c</b>	<b>Madde no</b>	<b>a</b>	<b>b</b>	<b>c</b>
<b>1</b>	1,25	-1,02	0,28	<b>26</b>	1,67	-2,37	0,23
<b>2</b>	1,35	2,66	0,24	<b>27</b>	1,92	-1,76	0,21
<b>3</b>	1,69	2,36	0,28	<b>28</b>	1,99	-1,61	0,25
<b>4</b>	1,18	-0,95	0,20	<b>29</b>	1,17	2,95	0,25
<b>5</b>	1,80	0,41	0,28	<b>30</b>	1,29	-0,53	0,25
<b>6</b>	1,19	0,81	0,22	<b>31</b>	1,98	-2,19	0,23
<b>7</b>	1,70	2,82	0,26	<b>32</b>	1,77	0,59	0,29
<b>8</b>	1,69	0,27	0,22	<b>33</b>	1,99	-1,49	0,25
<b>9</b>	1,65	-2,13	0,27	<b>34</b>	1,10	0,95	0,20
<b>10</b>	1,80	1,37	0,26	<b>35</b>	1,44	-0,72	0,26
<b>11</b>	1,82	1,72	0,29	<b>36</b>	1,06	0,58	0,20
<b>12</b>	1,93	-0,97	0,27	<b>37</b>	1,83	-0,32	0,20
<b>13</b>	1,92	-1,56	0,20	<b>38</b>	1,17	-2,64	0,20
<b>14</b>	1,84	-0,19	0,21	<b>39</b>	1,48	0,85	0,22
<b>15</b>	1,00	0,02	0,29	<b>40</b>	1,21	-1,81	0,28
<b>16</b>	1,24	0,66	0,26	<b>41</b>	1,35	2,34	0,26
<b>17</b>	1,86	1,96	0,29	<b>42</b>	1,84	0,45	0,20
<b>18</b>	1,56	0,18	0,23	<b>43</b>	1,23	2,50	0,23
<b>19</b>	1,08	0,83	0,23	<b>44</b>	1,20	0,03	0,20
<b>20</b>	1,34	-1,59	0,27	<b>45</b>	1,12	-0,38	0,29
<b>21</b>	1,42	2,17	0,21	<b>46</b>	1,50	-2,01	0,28
<b>22</b>	1,22	2,52	0,29	<b>47</b>	1,37	-1,32	0,20
<b>23</b>	1,10	2,33	0,20	<b>48</b>	1,98	2,53	0,27
<b>24</b>	1,83	-2,57	0,28	<b>49</b>	1,29	0,61	0,23
<b>25</b>	1,32	1,43	0,22	<b>50</b>	1,77	-2,19	0,27

Çizelge 3. L Formuna Ait Maddelerin Parametre Değerleri

	<b>a</b>	<b>b</b>	<b>c</b>	<b>Madde no</b>	<b>a</b>	<b>b</b>	<b>c</b>
<b>1</b>	1,98	-2,58	0,24	<b>26</b>	1,08	1,94	0,28
<b>2</b>	1,93	2,87	0,26	<b>27</b>	1,44	0,25	0,28
<b>3</b>	1,04	-2,10	0,25	<b>28</b>	1,76	-1,89	0,23
<b>4</b>	1,64	1,12	0,24	<b>29</b>	1,09	-1,38	0,23
<b>5</b>	1,90	0,17	0,21	<b>30</b>	1,18	0,06	0,26
<b>6</b>	1,01	1,31	0,22	<b>31</b>	1,10	-1,59	0,27
<b>7</b>	1,85	1,16	0,22	<b>32</b>	1,11	-1,07	0,25
<b>8</b>	1,48	-2,02	0,28	<b>33</b>	1,30	-0,12	0,27
<b>9</b>	1,45	-0,32	0,24	<b>34</b>	1,75	0,61	0,24
<b>10</b>	1,22	0,23	0,20	<b>35</b>	1,59	-1,66	0,22
<b>11</b>	1,80	-1,87	0,28	<b>36</b>	1,81	2,45	0,29
<b>12</b>	1,47	1,38	0,24	<b>37</b>	1,95	-1,51	0,27
<b>13</b>	1,06	-1,01	0,24	<b>38</b>	1,31	0,50	0,29
<b>14</b>	1,3	0,86	0,23	<b>39</b>	1,92	0,48	0,24
<b>15</b>	1,59	-2,37	0,29	<b>40</b>	1,23	-1,55	0,22
<b>16</b>	1,138	1,74	0,24	<b>41</b>	1,35	2,34	0,26
<b>17</b>	1,05	-1,05	0,28	<b>42</b>	1,84	0,45	0,20
<b>18</b>	1,33	-2,17	0,27	<b>43</b>	1,23	2,50	0,23
<b>19</b>	1,41	2,18	0,25	<b>44</b>	1,20	0,03	0,20
<b>20</b>	1,59	2,63	0,25	<b>45</b>	1,12	-0,38	0,29
<b>21</b>	1,84	2,38	0,24	<b>46</b>	1,50	-2,01	0,28
<b>22</b>	1,93	-0,48	0,29	<b>47</b>	1,37	-1,32	0,20
<b>23</b>	1,73	-2,02	0,22	<b>48</b>	1,98	2,53	0,27
<b>24</b>	1,60	-0,12	0,29	<b>49</b>	1,29	0,61	0,23
<b>25</b>	1,43	1,80	0,24	<b>50</b>	1,77	-2,19	0,27

Çizelge 2 ve Çizelge 3 incelendiğinde, testin son 10 maddesine (41-50 nolu maddeler) ait madde ayırıcılık parametresi, madde güçlük parametresi ve şans parametresi değerlerinin her iki test formunda da sabit olduğu görülmektedir. Bunun amacı, her iki test formunda yer alan ortak maddelere verilen tepkileri kullanarak, farklı gruplara uygulanan testlerden (K ve L testleri) kestirilen yetenek parametrelerini ortak bir ölçeğe yerleştirmeyi sağlayacak olan denklemdeki A (eğim) ve B (sabit) değerlerini saptamaktır.

Veri üretiminin ikinci aşamasında ise, katılımcılara ait bilgiler belirlenmiştir. Bu aşamada, katılımcılara ait örneklem büyüklükleri ve yetenek dağılımları tanımlanmıştır. Bu araştırmada örneklem büyüklüğü test eşitleme sonuçlarını etkileyebilecek bir değişken olarak düşünülmüş ve iki düzeyde ele alınmıştır. Buna göre, örneklem büyüklüğü bakımından 500 ve 1000 kişilik gruplar oluşturulmuştur. Araştırmada, örneklem büyüklüğünün yanında, yetenek dağılımı da değişken olarak ele alınmıştır. Yetenek dağılımları açısından veriler iki koşul altında düzenlemiştir. Normal dağılıma sahip  $[N(0,1)-N(0,1)]$  ve birbirinden farklı yetenek dağılımlarına sahip  $[N(0,1)-N(1,1)]$  gruplar oluşturulmuştur. Buna göre;

- 500 kişilik ortalaması 0, standart sapması 1 olan iki grup  $[N(0,1)]$
- 500 kişilik ortalaması 1, standart sapması 1 olan bir grup  $[N(1,1)]$
- 1000 kişilik ortalaması 0, standart sapması 1 olan iki grup  $[N(0,1)]$
- 1000 kişilik ortalaması 1, standart sapması 1 olan bir grup  $[N(1,1)]$  olmak üzere altı grup oluşturulmuştur.

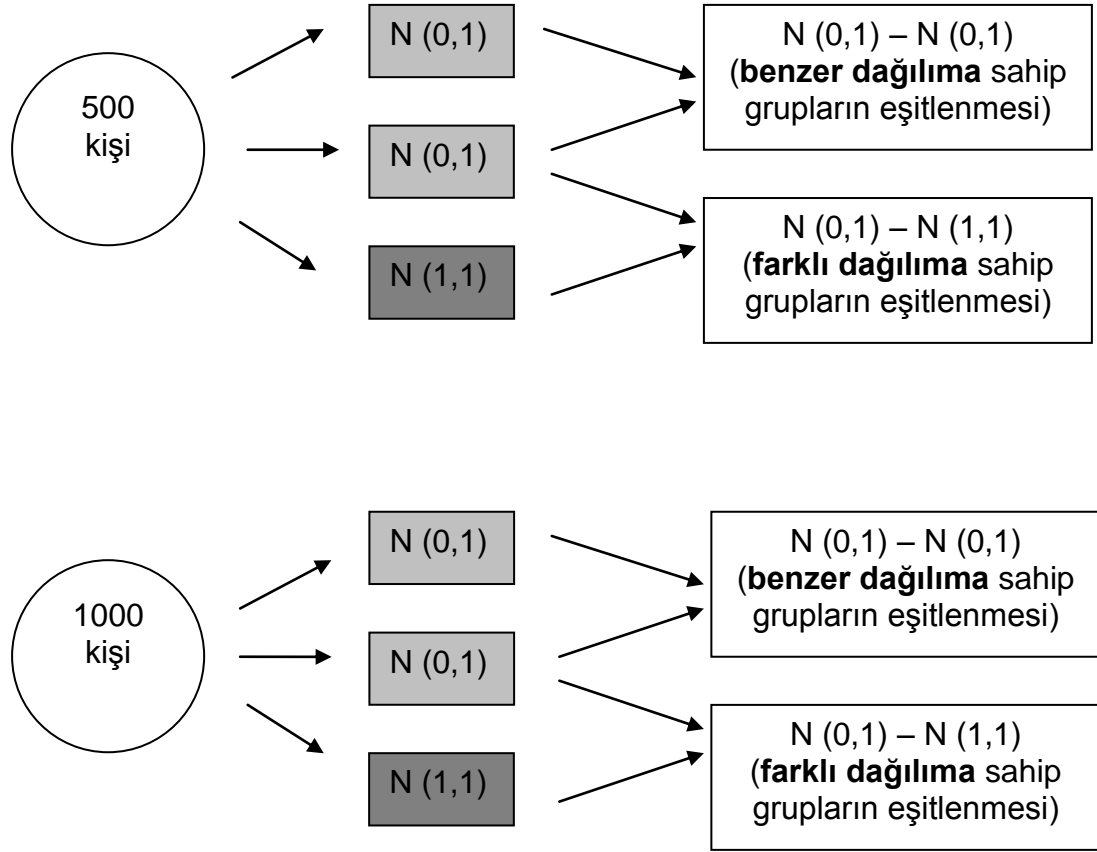
Üçüncü aşama ise veri üretiminin son aşaması olmakla birlikte, 1-0 verilerinin oluşturulduğu aşamadır. Bu aşamada, birinci aşamada oluşturulan test formları ve ikinci aşamada oluşturulan gruplar eşleştirilmiştir. Bu eşleştirme Çizelge 4'te gösterilmektedir:

Çizelge 4. Örneklem Büyüklüğü ve Yetenek Dağılımı Değişkenlerine Göre Kullanılan Test Formları

Örneklem büyüklüğü	Yetenek dağılımı	Test formu	Replikasyon sayısı
500	N(0,1)	K	100
	N (0,1)	L	100
	N (1,1)	K	100
1000	N(0,1)	K	100
	N (0,1)	L	100
	N (1,1)	K	100
Toplam: 6 grup veri			Toplam: 600 veri seti

Çizelge 4 incelendiğinde, örneklem büyüklüğü ve yetenek dağılımı değişkenleri göz önüne alınarak oluşturulan altı grubun, K ve L test formlarından birini aldığı görülmektedir. Yapılan simülasyon araştırmalarında bulguların genellenebilirliğini arttırmak amacıyla söz konusu araştırma koşulları için replikasyonlar yapılmaktadır. Harwell, Stone, Hsu, and Kirisci (1996), Madde Tepki Kuramı'na ilişkin çalışmalarda, 25 replikasyonun yeterli olacağını belirtmektedir. Literatür incelendiğinde, test eşitlemeye ilişkin olarak yapılan simülasyon araştırmalarında replikasyon sayısının 20 ile 100 arasında değiştiği görülmüştür (Hanson ve Beguin, 2002; Kim ve Cohen, 1998; Kim ve Cohen, 2002; Kim ve Kolen, 2006; Lee ve Ban, 2010; Wolkowitz, 2008; Yi, 2008). Buna göre, çalışmanın bu aşamasında, yukarıda açıklanan altı veri grubunun her biri için 100 replikasyon olmak üzere toplam 600 replikasyon yapılmıştır. Böylece, araştırmada kullanılmak üzere toplam 600 adet 1-0 verisi elde edilmiştir.

Araştırmada üretilen veriler ve test eşitleme koşulları Şekil 10'da özetlenmektedir:



Şekil 10. Araştırma Şeması

Şekil 10 incelendiğinde, yukarıda belirtilen altı grup veri setinin ikiyeşerli olarak kullanılması ile dört adet test eşileme koşulu oluşturulduğu görülmektedir. Bu koşullar aşağıda açıklanmaktadır:

- 1) **500'er kişilik benzer yetenek dağılımına sahip iki grup için 3 parametrelili lojistik modele uyumlu test verilerinin üretildiği koşul:** K testini alan grup için, 500 kişilik örneklem büyüklüğüne sahip ortalaması 0, standart sapması 1 olan yetenek dağılım özelliği taşıyan 100 replikasyon ve L testini alan grup için 500 kişilik örneklem büyüklüğüne sahip, yetenek ortalaması 0, standart sapması 1 olan yetenek dağılımı gösteren 100 replikasyon olmak üzere toplam 200 veri seti üretilmiştir. 200 set 1-0 verisi ikiyeşerli olarak eşitlenmiş, bu eşitlemeler sonucunda her bir test eşileme yöntemi için 100 adet olmak üzere toplam 400 adet eşileme denklemi elde edilmiştir.

- 2) **500'er kişilik farklı yetenek dağılımına sahip iki grup için 3 parametrelili lojistik modele uyumlu test verilerinin üretildiği koşul:** K testini alan grup için, 500 kişilik örneklem büyüklüğüne sahip ortalaması 0, standart sapması 1 olan yetenek dağılım özelliği taşıyan 100 replikasyon ve L testini alan grup için 500 kişilik örneklem büyüklüğüne sahip, yetenek ölçüleri ortalaması 1, standart sapması 1 olan yetenek dağılımı gösteren 100 replikasyon olmak üzere toplam 200 veri seti üretilmiştir. 200 veri seti ikiyeşerli olarak eşitlenmiş, bu eşitlemeler sonucunda her bir test eşitleme yöntemi için 100 adet olmak üzere toplam 400 adet eşitleme denklemi elde edilmiştir.
- 3) **1000'er kişilik benzer yetenek dağılımına sahip iki grup için 3 parametrelili lojistik modele uyumlu test verilerinin üretildiği koşul:** K testini alan grup için, 1000 kişilik örneklem büyüklüğüne sahip ortalaması 0, standart sapması 1 olan yetenek dağılım özelliği taşıyan 100 replikasyon ve L testini alan grup için 1000 kişilik örneklem büyüklüğüne sahip, ortalaması 0, standart sapması 1 olan yetenek dağılımı gösteren 100 replikasyon olmak üzere toplam 200 veri seti üretilmiştir. 200 veri seti ikiyeşerli olarak eşitlenmiş, bu eşitlemeler sonucunda her bir test eşitleme yöntemi için 100 adet olmak üzere toplam 400 adet eşitleme denklemi elde edilmiştir.
- 4) **1000'er kişilik farklı yetenek dağılımına sahip iki grup için 3 parametrelili lojistik modele uyumlu test verilerinin üretildiği koşul:** K testini alan grup için, 1000 kişilik örneklem büyüklüğüne sahip ortalaması 0, standart sapması 1 olan yetenek dağılım özelliği taşıyan 100 replikasyon ve L testini alan grup için 1000 kişilik örneklem büyüklüğüne sahip ortalaması 1, standart sapması 1 olan yetenek dağılımı gösteren 100 replikasyon olmak üzere toplam 200 veri seti üretilmiştir. 200 veri seti ikiyeşerli olarak eşitlenmiş, bu eşitlemeler sonucunda her bir test eşitleme yöntemi için 100 adet olmak üzere toplam 400 adet eşitleme denklemi elde edilmiştir.

## Verilerin Analizi

Bu araştırma için gerekli veriler WinGen2 bilgisayar programı yardımıyla üretilmiştir. Bu verilerin 1-0 şeklinde puanlanması nedeniyle, araştırma verilerine ait madde ve yetenek parametrelerinin kestiriminde BILOG-MG 3.0 (Zimowski, Muraki, Mislevy ve Bock, 1996) programı kullanılmıştır. WinGen2 programı ile üretilen 1-0 verileri “wgr” uzantılı dosyalardır. BILOG-MG programına aktarılabilmesi için bu veriler, “dat” uzantılı dosyalara dönüştürülmüşlerdir.

BILOG programına aktarılan veriler için bir sonraki aşama, yetenek kestirim yöntemine karar verilmesidir. Madde Tepki Kuramı'na dayalı olarak bir çok parametre kestirim yöntemi kullanılmaktadır. Bunlardan bazıları, maksimum olabilirlik (Maximum Likelihood Estimation, MLE), beklenen a posteriori (Expected A-Posteriori, EAP) ve maksimum a posterioridir (Maximum A-Posteriori, MAP) (Baker ve Kim, 2004; Embretson ve Reise, 2000; Tong, 2005).

Maksimum olabilirlik yöntemi, yanıtlayıcıların tüm yanıtları doğru ya da yanlış yanıtlaması halinde  $\pm$  sonsuzda yetenek kestirimleri yapmaktadır. Bu durum yeteneklerin birbiriyle karşılaştırılmalarını zorlaştırmaktadır. Bu nedenle bu yöntem tercih edilmemektedir. Maksimum olabilirliğe dayalı yetenek kestiriminin aksine, maksimum a posteriori ve beklenen a posteriori yöntemleri yanıtlayıcıların tüm yanıtları doğru ya da yanlış yanıtlaması halinde sonlu yetenek kestirimi sağlayabilmektedir. Her iki yöntemde de bir öncül dağılım kullanılmaktadır. Aralarındaki fark, maksimum a posteriori yönteminde öncül dağılımın tepe değerinden, beklenen a posteriori yönteminde ise ortalamasından yararlanılarak yetenek kestirimlerinin elde edilmesidir (Embretson ve Reise, 2000).

Bu çalışmada test eşitleme çalışmalarında kullanılmak üzere 600 adet 1-0 verisi üretilmiştir. BILOG-MG programında –kullanıcı farklı bir komut vermedikçe- yetenek değerleri, ortalaması 0; standart sapması 1 olacak

şekilde kestirilmektedir (Kolen ve Brennan, 2004). Bu araştırmada bilinen evren parametreleri ile çalışıldığından, yetenek kestirimleri için öncül dağılımlar bilinmektedir. Bu nedenle bu araştırmadaki yetenek kestiriminde, öncül dağılımın kullanıldığı beklenen a posteriori yetenek kestiriminden yararlanılmıştır. Her bir dağılımın ortalaması ve standart sapması, o dağılıma ait kestirimlerde öncül dağılım olarak kullanılmıştır. Her bir 1-0 veri dosyasına, sahip olduğu örneklem büyüklüğü ve yetenek dağılımını yansıtan etiketler verilmiştir. Bu verilere ait yetenek parametreleri her bir veri dosyası için ayrı ayrı olmak üzere toplam 600 adet sintaks dosyası hazırlanarak kestirilmiştir. Sintaks dosyalarında,

- Ortalaması 0, standart sapması 1 olan yetenek dağılımına sahip veriler için, BILOG-MG yetenek kestirimlerinde, ortalama 0'a; standart sapma ise 1'e;
- Ortalaması 1, standart sapması 1 olan yetenek dağılımına sahip veriler için, BILOG-MG yetenek kestirimlerinde, ortalama ve standart sapma 1'e sabitlenmiştir.

BILOG-MG programı ile kestirilen madde parametreleri "par" uzantılı dosyalar olduğundan Excel programından yararlanılarak "wgi" dosyalarına dönüştürülmüş ve IRTEQ (Han, 2007b) test eşitleme programına aktarılmıştır. Bu program yardımıyla bir testin iki farklı formu (K ve L) arasındaki eşitleme denklemi oluşturulmuştur. Elde edilen bu denklem sayesinde kişinin katılmadığı sınavdan elde edeceği kuramsal yetenek değeri ( $\Theta^*$ ) kestirilmiştir.

Bu araştırmada çeşitli koşullar altında (örneklem büyüklüğü ve yetenek dağılımı), "ortalama-ortalama, ortalama-standart sapma, Stocking-Lord ve Haebara" eşitleme yöntemlerinden kestirilen eşitleme hataları karşılaştırılmıştır. Test eşitleme yöntemlerinden kestirilen eşitleme hataları karşılaştırılırken bir ölçüte ihtiyaç duyulur. Karşılaştırma ölçütü olarak da literatürde çeşitli indeksler kullanılmaktadır. Bu indekslerden en yaygın kullanılanı "en küçük fark kareleri ortalaması (root mean square difference,



RMSD)” dır (Harris ve Crouse, 1993; Peak ve Young, 2005). Bu arařtırmada test eřitleme yntemlerinden elde edilen eřitleme hatalarının karřılařtırılmasında lut olarak, “en kck fark kareleri ortalaması” kullanılmıřtır. Bu yntemin matematiksel gsterimi ařađıdaki gibidir:

- $\theta^*$  : Kestirilmiş yetenek dzeyi  
 $\theta$  : Gerek yetenek dzeyi  
 $f$  : Belli bir yetenek dzeyinin frekansı

$$RMSD = \sqrt{\frac{\sum_i f_i (\theta^* - \theta)^2}{\sum_i f_i}} \quad (27)$$

Eřitlik 27 incelendiđinde, en kck fark kareleri ortalamasından hesaplanan deđerin, tm bireyler iin, bireylere uygulanmıř testten hesaplanan yetenek dzeyi ile test eřitlemenin yapıldıđı ancak bireylerin almadıđı test formundan kestirilen yetenek dzeyi arasındaki farkın karelerinin toplamının, toplam birey sayısına oranının karekk olduđu grlmektedir. Bu deđerin dřk olması hesaplanan yetenek dzeyi ile kestirilen yetenek dzeyi arasındaki farkın kck olduđuna iřaret etmektedir. Diđer bir deyiřle, bu deđerin dřk olması az hatalı eřitlemeler yapıldıđının bir gstergesidir.

## BÖLÜM 3

### BULGULAR VE YORUMLAR

Bu bölümde, araştırmanın amaçları çerçevesinde elde edilen bulgular özetlenmektedir.

#### 500 Kişilik *Benzer Yetenek Dağılımına Sahip Gruplardan Elde Edilen Farklı Eşitleme Yöntemlerinden Kestirilen Hatalara İlişkin Bulgular*

500 kişilik benzer yetenek dağılımına sahip gruplardan elde edilen eşitleme hatalarına ilişkin bulgular Çizelge 5'te gösterilmektedir:

Çizelge 5. N=500 Benzer Yetenek Dağılımı Koşulunda Farklı Yöntemlere Göre Yapılan Eşitlemelerden Hesaplanan Hata Miktarları

N	Dağılım	Test eşitleme yöntemi				
		O-O	O-S	H	S-L	
500	[N(0,1)-N(0,1)]	A	1,002	0,979	1,016	1,015
		B	-0,007	-0,013	-0,023	-0,023
		RMSD ortalaması	0,120	0,111	0,096	<b>0,091</b>

**O-O:** Ortalama-ortalama yöntemi

**O-S:** Ortalama-standart sapma yöntemi

**H** : Heabara yöntemi

**S-L:** Stocking-Lord yöntemi

**A:** Eşitleme denkleminin eğimi

**B:** Eşitleme denkleminin sabiti

**N:** Örneklem büyüklüğü

**RMSD:** Eşitleme hatası

Çizelge 5 incelendiğinde, 500 kişilik örneklem büyüklüğünde ve benzer

yetenek dağılımı gösteren gruplar için 3 parametrelili lojistik modele uyumlu iki test formunun eşitlenmesinde karşılaştırılan dört yöntem arasında en az eşitleme hatası veren test eşitleme yönteminin karakteristik eğri yaklaşımına dayalı geliştirilen “Stocking-Lord” yöntemi olduğu görülmektedir. Bu yöntemi sırasıyla Heabara ve ortalama-standart sapma yöntemi izlemektedir. Söz konusu koşul için en fazla hata içeren yöntem ise “ortalama-ortalama” yöntemidir.

Stocking ve Lord (1983) tarafından yapılan araştırmada karakteristik eğri yöntemleri ile ortalama-standart sapma yöntemi karşılaştırılmıştır. Araştırma sonucunda Haebara ve Stocking-Lord yöntemlerinin ortalama-standart sapma yöntemlerine nazaran daha iyi sonuçlar verdiği saptanmıştır. Stocking ve Lord (1983) tarafından yapılan araştırmanın bulguları ile bu araştırmanın bulguları benzerlik göstermektedir. Baker ve Al-Karni (1991) tarafından yapılan araştırmada da benzer bir sonuç elde edilmiştir. Ogasawara (2001) tarafından yapılan başka bir araştırmada hem 2 hem de 3 parametrelili modeller için Haebara ve Stocking-Lord yöntemlerinin moment yöntemlerine (ortalama-ortalama, ortalama-standart sapma) nazaran daha iyi eşitlemeler yaptığı sonucuna ulaşılmıştır. Baker ve Al-Karni (1991) ve Ogasawara (2001) tarafından yapılan araştırmaların bulguları ile bu araştırmanın bulguları benzerlik göstermektedir.

Cohen ve Kim (1998) tarafından yapılan araştırmada, dereceli tepki modeline (graded response model) uyumlu simülatif veriler kullanılarak ortalama-ortalama, ortalama-standart sapma, Haebara ve Stocking-Lord yöntemleri karşılaştırılmıştır. Araştırma sonucunda, yukarıda yer alan bulgulara benzer sonuçlar elde edilmiştir.

Hanson ve Beguin (2002) tarafından yapılan araştırmada, Haebara ve Stocking-Lord yöntemlerinin, moment yöntemlerinden (ortalama-ortalama, ortalama-standart sapma) daha az eşitleme hatasıyla sonuçlandığı görülmüştür. Kim ve Kolen (2006) tarafından yapılan araştırmada çoktan seçmeli maddelerin ve öğrencinin yanıtını yapılandırdığı (constructed response model) açık uçlu yanıt vermeye uygun maddelerin birlikte

kullanıldığı karışık test formatında hazırlanmış çok boyutlu testlerin eşitlenmesi araştırılmıştır. Araştırmada Stocking-Lord yöntemine dayalı olarak yapılan eşitlemelerin, ortalama-ortalama yöntemine dayalı olarak yapılan eşitlemelere nazaran daha düşük hata ile sonuçlandığı saptanmıştır. Bu bulguya, Hung, Wu ve Chen (1991), Way ve Tang (1991), Karkee ve Wright (2004), Kaskowitz ve De Ayala (2001), Kim ve Lee (2004), Kim ve Lee (2006), Kim ve Kolen (2004) ve Kim ve Song (2004) tarafından yapılan çalışmalarda da rastlanmaktadır.

### 500 Kişilik Farklı Yetenek Dağılımına Sahip Gruplardan Elde Edilen Farklı Eşitleme Yöntemlerinden Kestirilen Hatalara İlişkin Bulgular

500 kişilik farklı yetenek dağılımına sahip gruplardan elde edilen eşitleme hatalarına ilişkin bulgular Çizelge 6'da gösterilmektedir:

Çizelge 6. N=500 Farklı Yetenek Dağılımı Koşulunda Farklı Yöntemlere Göre Yapılan Eşitlemelerden Hesaplanan Hata Miktarları

N	Dağılım		Test eşitleme yöntemi			
			O-O	O-S	H	S-L
500	[N(0,1)-N(1,1)]	A	1,005	1,172	1,079	1,089
		B	-0,086	-0,045	-0,065	-0,064
		RMSD ortalaması	0,170	0,240	0,168	<b>0,150</b>

**O-O:** Ortalama-ortalama yöntemi

**O-S:** Ortalama-standart sapma yöntemi

**H** : Heabara yöntemi

**S-L:** Stocking-Lord yöntemi

**A:** Eşitleme denkleminin eğimi

**B:** Eşitleme denkleminin sabiti

**N:** Örneklem büyüklüğü

**RMSD:** Eşitleme hatası

Çizelge 6 incelendiğinde, 500 kişilik farklı yetenek dağılımına sahip gruplarda, 3 parametrelili lojistik modele uyumlu farklı test formlarının

eşitlenmesinde kullanılabilir en uygun test eşitleme yönteminin karakteristik eğri yaklaşımına dayalı geliştirilen “Stocking-Lord” yöntemi olduğu görülmektedir. Bu yöntemi sırasıyla Heabara ve ortalama-ortalama yöntemi izlemektedir. Söz konusu koşul için en fazla hata içeren yöntem ise “ortalama-standart sapma” yöntemidir.

Stocking ve Lord (1983) tarafından yapılan araştırmada karakteristik eğri yöntemleri ile ortalama-standart sapma yöntemi karşılaştırılmıştır. Araştırma sonucunda Haebara ve Stocking-Lord yöntemlerinin ortalama-standart sapma yöntemlerine nazaran daha iyi sonuçlar verdiği saptanmıştır. Baker ve Al-Karni (1991) ve Ogasawara (2001) tarafından yapılan araştırmalarda da benzer bulgular elde edilmiştir. Cohen ve Kim (1998) tarafından yapılan araştırmada, dereceli tepki modeline (graded response model) uyumlu simülatif veriler kullanılarak ortalama-ortalama, ortalama-standart sapma, Haebara ve Stocking-Lord karakteristik eğri eşitleme yöntemleri karşılaştırılmıştır. Araştırma sonucunda, yukarıda yer alan bulgulara benzer sonuçlar elde edilmiştir.

Hanson ve Beguin (2002) tarafından yapılan araştırmada, Haebara ve Stocking-Lord yöntemlerinin, moment yöntemlerinden (ortalama-ortalama, ortalama-standart sapma) daha az eşitleme hatasıyla sonuçlandığı görülmüştür. Kim ve Kolen (2006) tarafından yapılan araştırmada çoktan seçmeli maddelerin ve öğrencinin yanıtını yapılandığı (constructed response model) açık uçlu yanıt vermeye uygun maddelerin birlikte kullanıldığı karışık test formatında hazırlanmış çok boyutlu testlerin eşitlenmesi araştırılmıştır. Araştırmada Stocking-Lord yöntemine dayalı olarak yapılan eşitlemelerin, ortalama-ortalama yöntemine dayalı olarak yapılan eşitlemelere nazaran daha düşük hata ile sonuçlandığı saptanmıştır. Bu bulguya, Hung, Wu ve Chen (1991), Way ve Tang (1991), Karkee ve Wright (2004), Kaskowitz ve De Ayala (2001), Kim ve Lee (2004), Kim ve Lee (2006), Kim ve Kolen (2004) ve Kim ve Song (2004) tarafından yapılan çalışmalarda da rastlanmaktadır.

Bu arařtırmada, en yksek RMSD deęerlerinin ise ortalama-standart sapma yntemine ait olduęu gzlenmiřtir. Speron (2009) tarafından yapılan alıřmada benzer bir sonu elde edilmiřtir. Baker ve Al-Karni (1991) tarafından yapılan arařtırmada, ortalama-ortalama ynteminin, ortalama-standart sapma yntemine gre daha iyi eřitleme sonucu verdięi belirtilmektedir. Bunun nedeni olarak, ortalamaların, standart sapmaya gre daha kararlı bir parametre olması gsterilmektedir.

### 1000 Kiřilik Benzer Yetenek Daęılımına Sahip Graplardan Elde Edilen Farklı Eřitleme Yntemlerinden Kestirilen Hatalara İliřkin Bulgular

1000 kiřilik benzer yetenek daęılımına sahip gruplardan elde edilen eřitleme hatalarına iliřkin bulgular izelge 7’de gsterilmektedir:

izelge 7. N=1000 Benzer Yetenek Daęılımı Kořulunda Farklı Yntemlere Gre Yapılan Eřitlemelerden Hesaplanan Hata Miktarları

N	Daęılım	Test eřitleme yntemi				
		O-O	O-S	H	S-L	
1000	[N(0,1)-N(0,1)]	A	1,008	1,007	1,003	1,009
		B	0,129	-0,011	-0,007	-0,008
		RMSD ortalaması	0,090	0,098	0,086	<b>0,077</b>

**O-O:** Ortalama-ortalama yntemi

**O-S:** Ortalama-standart sapma yntemi

**H** : Heabara yntemi

**S-L:** Stocking-Lord yntemi

**A:** Eřitleme denkleminin eęimi

**B:** Eřitleme denkleminin sabiti

**N:** rneklem byklę

**RMSD:** Eřitleme hatası

izelge 7 incelendięinde, 1000 kiřilik benzer yetenek daęılımına sahip gruplarda, 3 parametrelili lojistik modele uyumlu farklı test formlarının eřitlenmesinde kullanılabilir en uygun test eřitleme ynteminin karakteristik eęri yaklařımına dayalı geliřtirilen ‘‘Stocking-Lord’’ yntemi

olduđu grlmektedir. Bu yntemi sırasıyla Heabara ve ortalama-ortalama yntemi izlemektedir. Sz konusu kořul iin en fazla hata ieren yntem ise "ortalama-standart sapma" yntemidir.

Stocking ve Lord (1983) tarafından yapılan arařtırmada karakteristik eđri yntemleri ile ortalama-standart sapma yntemi karřılařtırılmıřtır. Arařtırma sonucunda Haebara ve Stocking-Lord yntemlerinin ortalama-standart sapma yntemlerine nazaran daha iyi sonular verdiđi saptanmıřtır. Stocking ve Lord (1983) tarafından yapılan arařtırmanın bulguları ile bu arařtırmanın bulguları benzerlik gstermektedir.

Baker ve Al-Karni (1991) tarafından yapılan arařtırmada da benzer bir sonu elde edilmiřtir. Ogasawara (2001) tarafından yapılan bařka bir arařtırmada hem 2 hem de 3 parametrelili modeller iin Haebara ve Stocking-Lord yntemlerinin moment yntemlerine (ortalama-ortalama, ortalama-standart sapma) nazaran daha iyi eřitlemeler yaptığı sonucuna ulařılmıřtır. Baker ve Al-Karni (1991) ve Ogasawara (2001) tarafından yapılan arařtırmaların bulguları ile bu arařtırmanın bulguları benzerlik gstermektedir.

Cohen ve Kim (1998) tarafından yapılan arařtırmada, dereceli tepki modeline (graded response model) uyumlu simlatif veriler kullanılarak ortalama-ortalama, ortalama-standart sapma, Haebara ve Stocking-Lord karakteristik eđri eřitleme yntemleri karřılařtırılmıřtır. Arařtırma sonucunda, farklı kořullar altında yukarıda yer alan bulgulara benzer sonular elde edilmiřtir.

Hanson ve Beguin (2002) tarafından yapılan arařtırmada, Haebara ve Stocking-Lord yntemlerinin, moment yntemlerinden (ortalama-ortalama, ortalama-standart sapma) daha az eřitleme hatasıyla sonulandıđı grlmřtr. Hanson ve Beguin (2002) tarafından yapılan arařtırmanın sonularıyla, bu arařtırmanın sonuları tutarlılık gstermektedir.

Kim ve Kolen (2006) tarafından yapılan arařtırmada oktan semeli maddelerin ve ğrencinin yanıtını yapılandırdığı (constructed response model) açık uçlu yanıt vermeye uygun maddelerin birlikte kullanıldığı karışık test formatında hazırlanmış ok boyutlu testlerin eřitlenmesi arařtırılmıştır. Arařtırmada Stocking-Lord yöntemine dayalı olarak yapılan eřitlemelerin, ortalama-ortalama yöntemine dayalı olarak yapılan eřitlemelere nazaran daha düşük hata ile sonuçlandığı saptanmıştır. Bu bulguya, Hung, Wu ve Chen (1991), Way ve Tang (1991), Karkee ve Wright (2004), Kaskowitz ve De Ayala (2001), Kim ve Lee (2004), Kim ve Lee (2006), Kim ve Kolen (2004) ve Kim ve Song (2004) tarafından yapılan alıřmalarda da rastlanmaktadır.

Bu arařtırmada, en yüksek RMSD deęerlerinin ise ortalama-standart sapma yöntemine ait olduęu gözlenmiştir. Speron (2009) tarafından yapılan alıřmada benzer bir sonuç elde edilmiştir. Baker ve Al-Karni (1991) tarafından yapılan arařtırmada, ortalama-ortalama yönteminin, ortalama-standart sapma yöntemine göre daha iyi eřitleme sonucu verdięi belirtilmektedir. Bunun nedeni olarak, ortalamaların, standart sapmaya göre daha kararlı bir parametre olması gösterilmektedir.

### **1000 Kiřilik *Farklı* Yetenek Daęılımına Sahip Gruplardan Elde Edilen Farklı Eřitleme Yöntemlerinden Kestirilen Hatalara İliřkin Bulgular**

1000 kiřilik örneklem büyüklüğünde ve farklı yetenek daęılımından elde edilen eřitleme hatalarına iliřkin bulgular izelge 8'de gösterilmektedir:



Çizelge 8. N=1000 *Farklı Yetenek Dağılımı Koşulunda Farklı Yöntemlere Göre Yapılan Eşitlemelerden Hesaplanan Hata Miktarları*

N	Dağılım		Test eşitleme yöntemi			
			O-O	O-S	H	S-L
1000	[N(0,1)-N(1,1)]	A	1,023	1,142	1,069	1,056
		B	-0,054	-0,023	-0,051	-0,046
		RMSD ortalaması	0,126	0,189	0,120	<b>0,117</b>

**O-O:** Ortalama-ortalama yöntemi

**O-S:** Ortalama-standart sapma yöntemi

**H** : Heabara yöntemi

**S-L:** Stocking-Lord yöntemi

**A:** Eşitleme denkleminin eğimi

**B:** Eşitleme denkleminin sabiti

**N:** Örneklem büyüklüğü

**RMSD:** Eşitleme hatası

Çizelge 8 incelendiğinde, 1000 kişilik farklı yetenek dağılımına sahip gruplarda, 3 parametrelili lojistik modele uyumlu farklı test formlarının eşitlenmesinde kullanılabilecek en uygun test eşitleme yönteminin karakteristik eğri yaklaşımına dayalı geliştirilen “Stocking-Lord” yöntemi olduğu görülmektedir. Bu yöntemi sırasıyla Heabara ve ortalama-ortalama yöntemi izlemektedir. Söz konusu koşul için en fazla hata içeren yöntem ise “ortalama-standart sapma” yöntemidir.

Stocking ve Lord (1983) tarafından yapılan araştırmada Heabara ve Stocking-Lord yöntemlerinin ortalama-standart sapma yöntemlerine nazaran daha iyi sonuçlar verdiği saptanmıştır. Stocking ve Lord (1983) tarafından yapılan araştırmanın bulguları ile bu araştırmanın bulguları benzerlik göstermektedir. Baker ve Al-Karni (1991) tarafından yapılan araştırmada da benzer bir sonuç elde edilmiştir. Ogasawara (2001) tarafından yapılan başka bir araştırmada hem 2 hem de 3 parametrelili modeller için Heabara ve Stocking-Lord yöntemlerinin moment yöntemlerine (ortalama-ortalama, ortalama-standart sapma) nazaran daha iyi eşitlemeler yaptığı sonucuna ulaşılmıştır. Baker ve Al-Karni (1991) ve Ogasawara (2001) tarafından

yapılan arařtırmaların bulguları ile bu arařtırmanın bulguları benzerlik göstermektedir.

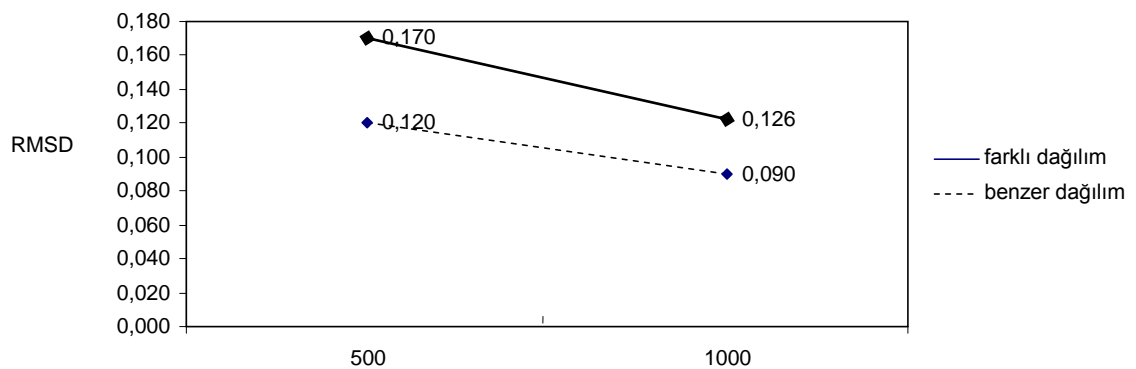
Cohen ve Kim (1998) tarafından yapılan arařtırmada, dereceli tepki modeline (graded response model) uyumlu simülatif veriler kullanılarak ortalama-ortalama, ortalama-standart sapma, Haebara ve Stocking-Lord karakteristik eğri eşitleme yöntemleri karşılaştırılmıştır. Arařtırma sonucunda, farklı koşullar altında yukarıda yer alan bulgulara benzer sonuçlar elde edilmiştir. Hanson ve Beguin (2002) tarafından yapılan arařtırmada, Haebara ve Stocking-Lord yöntemlerinin, moment yöntemlerinden (ortalama-ortalama, ortalama-standart sapma) daha az eşitleme hatasıyla sonuçlandığı görülmüştür.

Kim ve Kolen (2006) tarafından yapılan arařtırmada çoktan seçmeli maddelerin ve öğrencinin yanıtını yapılandırdığı (constructed response model) açık uçlu yanıt vermeye uygun maddelerin birlikte kullanıldığı karışık test formatında hazırlanmış çok boyutlu testlerin eşitlenmesi arařtırılmıştır. Arařtırmada Stocking-Lord yöntemine dayalı olarak yapılan eşitlemelerin, ortalama-ortalama yöntemine dayalı olarak yapılan eşitlemelere nazaran daha düşük hata ile sonuçlandığı saptanmıştır. Benzer bulguya, Hung, Wu ve Chen (1991), Way ve Tang (1991), Karkee ve Wright (2004), Kaskowitz ve De Ayala (2001), Kim ve Lee (2004), Kim ve Lee (2006), Kim ve Kolen (2004) ve Kim ve Song (2004) tarafından yapılan çalışmalarda da rastlanmaktadır.

Bu arařtırmada, en yüksek RMSD değerlerinin ise ortalama-standart sapma yöntemine ait olduğu gözlenmiştir. Speron (2009) tarafından yapılan çalışmada benzer bir sonuç elde edilmiştir. Baker ve Al-Karni (1991) tarafından yapılan arařtırmada, ortalama-ortalama yönteminin, ortalama-standart sapma yöntemine göre daha iyi eşitleme sonucu verdiği belirtilmektedir. Bunun nedeni olarak, ortalamaların, standart sapmaya göre daha kararlı bir parametre olması gösterilmektedir.

## Ortalama-Ortalama Yöntemine İlişkin Bulgular

Şekil 11’de 500 ve 1000 kişilik örneklem büyüklüklerine sahip; benzer ve farklı yetenek dağılımları gösteren gruplara ait ortalama-ortalama yöntemine dayalı test eşitleme çalışmasından elde edilen RMSD ortalamaları gösterilmektedir.

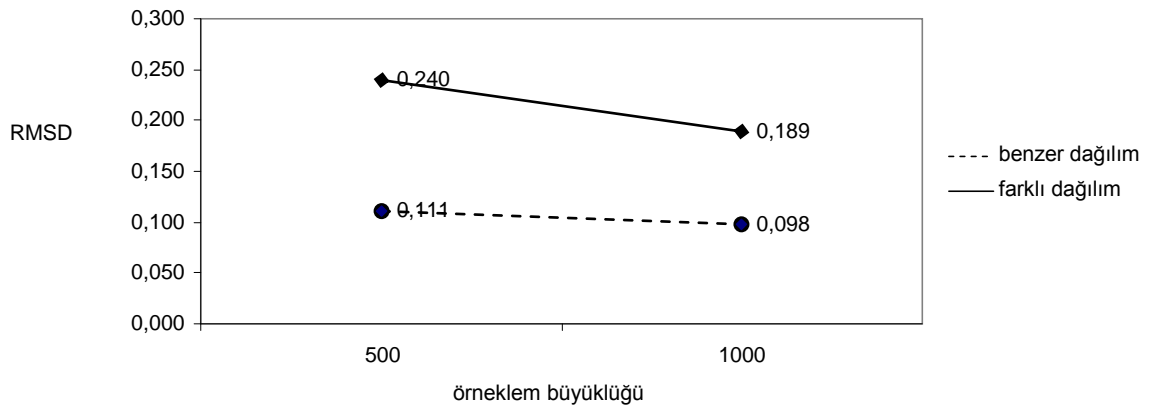


Şekil 11. Ortalama-Ortalama Yöntemine ait RMSD Değerleri

Şekil 11 incelendiğinde, hem benzer yetenek dağılımına sahip gruplar  $[N(0,1) - N(0,1)]$  için hem de farklı yetenek dağılımına sahip gruplar  $[N(0,1) - N(1,1)]$  için ortalama–ortalama yöntemine ait RMSD ortalamasının örneklem büyüklüğü arttıkça azalma eğilimi gösterdiği görülmektedir. Yetenek dağılımı bakımından RMSD ortalamaları incelendiğinde ise, farklı yetenek dağılımlarına sahip gruplara ait RMSD ortalamasının, benzer yetenek dağılımına sahip gruplara göre daha yüksek olduğu görülmektedir. Diğer bir deyişle, eşitleme yapılan grupların dağılımları benzer bir dağılım gösterdiğinde RMSD ortalamaları hem 500 hem de 1000 kişilik gruplar için azalma eğilimi göstermiştir.

## Ortalama-Standart Sapma Yöntemine İlişkin Bulgular

Şekil 12'de 500 ve 1000 kişilik örneklem büyüklüklerine sahip, benzer ve farklı yetenek dağılımları gösteren gruplara ait ortalama-standart sapma yöntemine dayalı test eşitleme çalışmasından elde edilen RMSD ortalamaları gösterilmektedir.

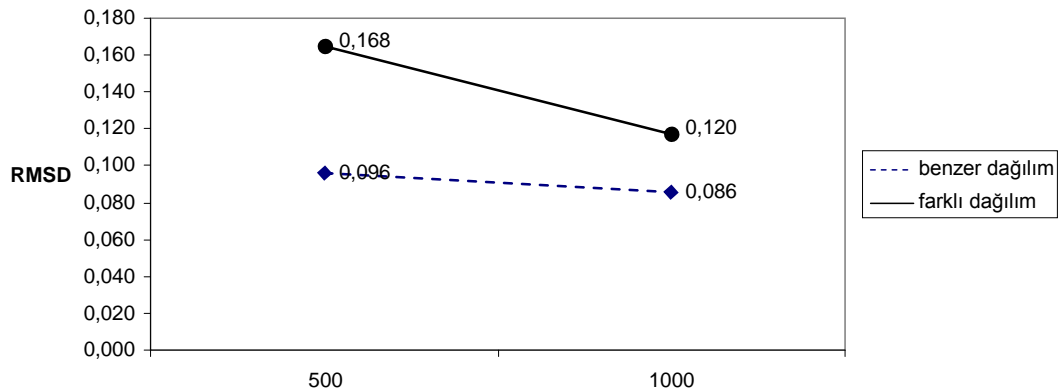


Şekil 12. Ortalama-Standart Sapma Yöntemine ait RMSD Değerleri

Şekil 12 incelendiğinde, ortalama-standart sapma yöntemine ait RMSD ortalamalarının örneklem büyüklüğü arttıkça azalma eğilimi gösterdikleri görülmektedir. Yetenek dağılımı bakımından 500 ve 1000'er kişilik örneklem büyüklüklerine sahip grupların eşitlenmesinden elde edilen RMSD değerleri incelendiğinde ise, farklı yetenek dağılımlarına sahip gruplara ait değerlerin, benzer yetenek dağılımına sahip gruplara göre daha yüksek olduğu görülmektedir. Diğer bir deyişle, eşitleme yapılan grupların dağılımları benzerlik gösterdikçe iki form arasındaki eşitleme hataları (RMSD değerleri) hem 500 hem de 1000 kişilik gruplar için azalma eğilimi göstermiştir.

## Haebara Yöntemine İlişkin Bulgular

Şekil 13'te 500 ve 1000 kişilik örneklem büyüklüklerine sahip; benzer ve farklı yetenek dağılımları gösteren gruplara ait Haebara yöntemine dayalı test eşitleme çalışmasından elde edilen RMSD değerleri gösterilmektedir:

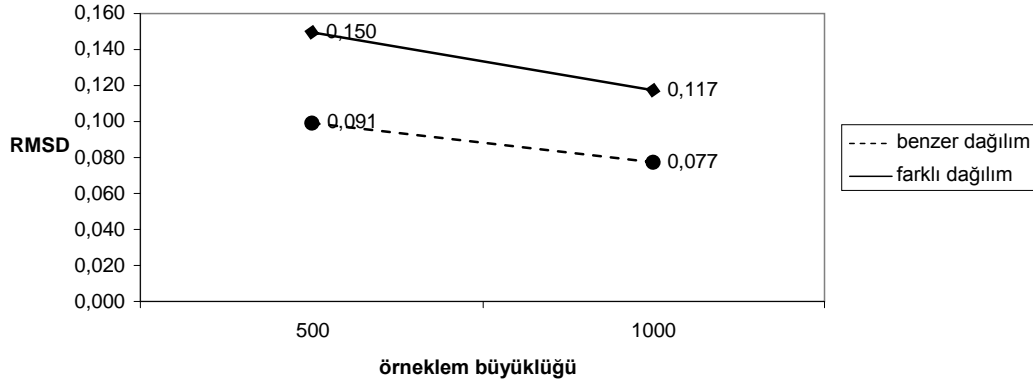


Şekil 13. Haebara Yöntemine ait RMSD Değerleri

Şekil 13 incelendiğinde, Haebara yöntemine ait RMSD ortalamalarının örneklem büyüklüğü arttıkça azalma eğilimi gösterdikleri görülmektedir. Yetenek dağılımı bakımından RMSD değerleri incelendiğinde ise, farklı yetenek dağılımlarına sahip gruplara  $[N(0,1) - N(1,1)]$  ait eşitleme hatalarının benzer yetenek dağılımına sahip gruplara  $[N(0,1) - N(0,1)]$  göre daha yüksek olduğu görülmektedir. Diğer bir deyişle, eşitleme yapılan grupların dağılımları benzerlik gösterdikçe iki form arasındaki eşitleme hataları (RMSD değerleri) hem 500 hem de 1000 kişilik gruplar için azalma eğilimi göstermiştir.

### Stocking-Lord Yöntemine İlişkin Bulgular

Şekil 14'te 500 ve 1000 kişilik örneklem büyüklüklerine sahip; benzer ve farklı yetenek dağılımları gösteren gruplara ait Stocking-Lord yöntemine dayalı test eşitleme çalışmasından elde edilen RMSD değerleri gösterilmektedir.

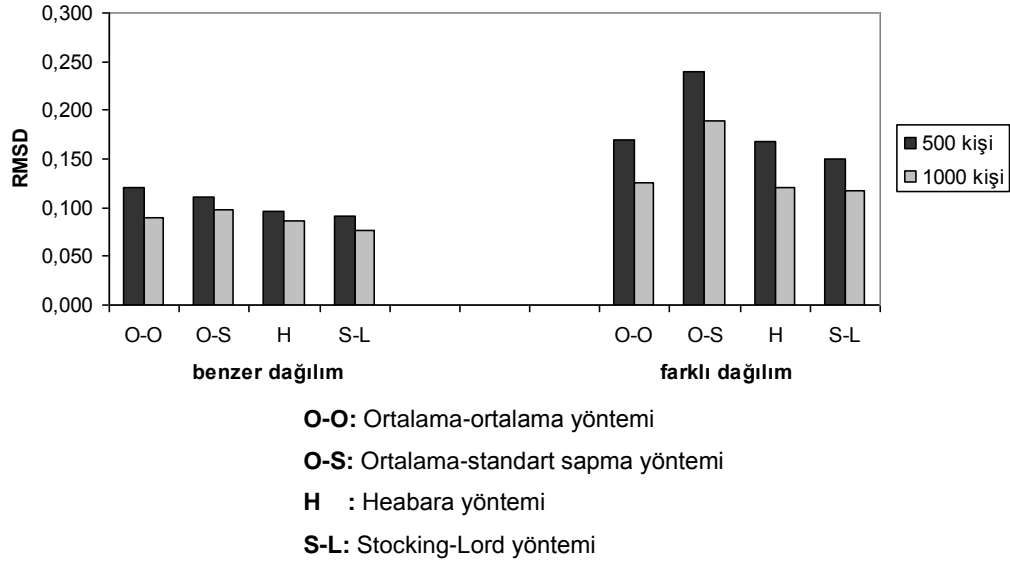


Şekil 14. Stocking-Lord Yöntemine ait RMSD Değerleri

Şekil 14 incelendiğinde, Stocking ve Lord yöntemine ait RMSD değerlerinin 1000 kişilik örneklerde daha düşük değerler verdiği; 500 kişilik örnek büyüklüklerinde ise bu değerlerin daha yüksek olduğu görülmektedir. Bu durum hem benzer yetenek dağılımına sahip gruplar  $[N(0,1) - N(0,1)]$  için hem de farklı yetenek dağılımına sahip gruplar  $[N(0,1) - N(1,1)]$  için geçerlidir. Yetenek dağılımı bakımından elde edilen RMSD değerleri incelendiğinde ise, farklı yetenek dağılımlarına sahip gruplara ait değerlerin, benzer yetenek değerlerine sahip gruplara göre daha yüksek olduğu görülmektedir. Diğer bir deyişle, eşitleme yapılan grupların dağılımları benzer olduğunda RMSD değerleri hem 500 hem de 1000 kişilik gruplar için azalma eğilimi göstermiştir.

### Örneklem Büyüklüğü Bakımından Test Eşitleme Hatalarının Karşılaştırılmasına İlişkin Bulgular

Örneklem büyüklüğü açısından test eşitleme yöntemlerinden kestirilen eşitleme hatalarının karşılaştırılmasına yönelik bulgular Şekil 15'te gösterilmektedir.



Şekil 15. Örneklem Büyüklüğü Açısından Test Eşitleme Yöntemlerinden Kestirilen Hataların (RMSD değerleri) Karşılaştırılması

Şekil 15 incelendiğinde, hem benzer hem de farklı dağılımlar için, 500 kişilik gruplardan elde edilen RMSD ortalamalarının 1000 kişilik örneklere nazaran daha yüksek değerler aldığı görülmektedir. Diğer bir deyişle örneklem büyüklüğü arttıkça RMSD ortalamalarında azalma gözlenmiştir. Bu durum benzer ve farklı yetenek dağılım koşullarının her ikisinde de gözlenmektedir.

Yotinprasert (1986) tarafından, Rasch modeline dayalı eşitleme yöntemlerinin karşılaştırıldığı araştırmada örneklem büyüklüğü arttıkça eşitleme hatalarının azaldığı sonucuna ulaşılmıştır. Speron (2009) tarafından 2 parametrelili modele uyumlu veriler üzerinde yapılan araştırmada 1000 kişilik örneklemlerden elde edilen eşitleme hatalarının 250 kişilik örneklem büyüklüğünden elde edilenlere oranla daha düşük olduğu saptanmıştır. Harris (1993), Hanson ve Beguin (2002), Kim ve Cohen (2002) ve Yang (1997) tarafından yapılan araştırmalarda da benzer sonuçlar elde edilmiştir. Bastari (2000), Tate (2000) tarafından hem 1-0 verisinin hem de çoklu kategorilerde puanlanan verilerin bulunduğu karışık test formatında (mix format) yapılan araştırmalarda da benzer bir sonuç elde edilmiştir.

Çetin (2009) tarafından yapılan arařtırmada, 2005 ÖBBS sınavına giren 6., 7. ve 8. sınıf öğrencilerine ait veriler üzerinde dikey ölçeklemeler yapılmıřtır. Bu arařtırma sonucunda, 250, 500 ve 1000 kiřilik örneklem büyüklüklerinden, en az hata veren örneklem büyüklüğünün 1000 kiřilik olduđu saptanmıřtır.

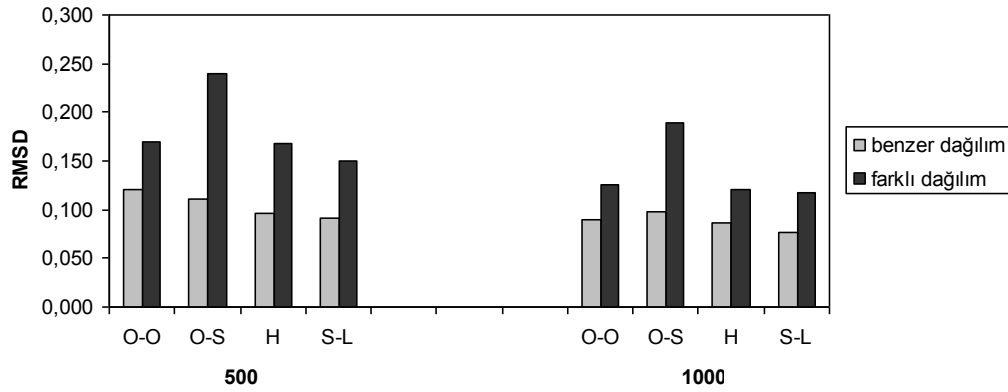
Lee ve Ban (2010) tarafından yapılan test eřitleme arařtırmasında örneklem büyüklüğü 500 ve 3000 olmak üzere iki kategoride ele alınmıřtır. Arařtırma sonucunda 3000 kiřilik örneklemelerden elde edilen eřitleme hatalarının, 500 kiřilik örneklemelerden elde edilenlere nazaran daha düşük olduđu saptanmıřtır.

Bu bulguların aksine, Suanthong (1998) tarafından 1 parametrelili modele dayalı eřitlemelerin yapıldığı arařtırmada ise, arařtırmada kullanılan 100, 300, 500 kiřilik örneklem büyüklüklerinden 100 kiřilik örneklemelerde, en az hatayla eřitlemenin yapıldığı gözlenmiřtir. Buna göre bu arařtırmanın bulguları Yang (1997), Bastari (2000), Çetin (2009); Hanson ve Beguin (2002), Harris (1993), Kim ve Cohen (2002), Lee ve Ban (2010), Speron (2009), Tate (2000) ve Yotinprasert (1986)'in bulguları ile tutarlılık göstermektedir.

### **Yetenek Dağılımı Bakımından Test Eřitleme Hatalarının Karşılaştırılmasına İliřkin Bulgular**

Yetenek dağılımı bakımından test eřitleme yöntemlerinden kestirilen eřitleme hatalarının karşılaştırılmasına yönelik bulgular Őekil 16'da gösterilmektedir.





**O-O:** Ortalama-ortalama yöntemi

**O-S:** Ortalama-standart sapma yöntemi

**H :** Heabara yöntemi

**S-L:** Stocking-Lord yöntemi

Şekil 16. Yetenek Dağılımı Açısından Test Eşitleme Yöntemlerinden Kestirilen Hataların (RMSD değerleri) Karşılaştırılması

Şekil 16 incelendiğinde, hem 500 hem de 1000 kişilik gruplar için, yetenek dağılım değişkeni bakımından benzer yetenek dağılımına sahip grupların farklı yetenek dağılımlarına nazaran daha düşük eşitleme hatasına sahip olduğu bulunmuştur. Diğer bir deyişle iki test formunun eşitlendiği iki grubun yetenek dağılımı farklılaştıkça RMSD ortalamaları/eşitleme hataları artma eğilimi göstermiştir. Hanson ve Beguin (2002) tarafından yapılan araştırmada, 1000 ve 3000 kişilik örneklem büyüklüğüne sahip ortalaması 0, standart sapması 1 olan dağılım ve ortalaması 1 standart sapması 1 olan dağılıma sahip iki grubun eşitlenmesinden elde edilen bulgular da bu araştırmanın bulgularıyla benzerlik göstermektedir.

Kim ve Cohen (2002) tarafından dereceli tepki modelinin kullanıldığı eşitlemelerde de benzer bir sonuç elde edilmiştir. 300 ve 1000 örneklem büyüklüklerinde, ortalaması 0, standart sapması 1 olan dağılım ve ortalaması 1 standart sapması 1 olan dağılıma sahip iki grubun puanları eşitlenmiştir. Araştırmanın sonucunda benzer yetenek dağılımına sahip grupların eşitlenmesinden elde edilen eşitleme hatalarının, farklı yetenek dağılımına

sahip grupların eşitlenmesinden elde edilen hatalara nazaran daha düşük olduğu saptanmıştır. Bastari (2000) ve Kim ve Lee (2006) tarafından yapılan araştırmalarda da benzer bir sonuç elde edilmiştir. Grupların yetenek dağılımının benzer olması eşitleme hatalarını düşürmüştür. Buna göre, yetenek dağılımına ilişkin temel bulgu literatürle desteklenmektedir.

## BÖLÜM 4

### SONUÇ VE ÖNERİLER

Bu bölümde, araştırmadan elde edilen sonuçlara ve bu sonuçlara dayalı olarak sunulabilecek önerilere yer verilmiştir.

#### Sonuçlar

Bu araştırmada farklı örneklem büyüklüğü ve yetenek dağılımı koşullarında, Madde Tepki Kuramı'na dayalı ortalama-ortalama, ortalama-standart sapma, Haebara ve Stocking-Lord yöntemlerinin karşılaştırılması amaçlanmıştır. Bu araştırmanın sonuçları belirtilen koşullar altında aşağıda özetlenmektedir:

500 ve 1000 kişilik örneklem büyüklüklerine sahip gruplardan elde edilen RMSD değerleri incelendiğinde, 1000 kişilik gruplardan elde edilen RMSD değerleri ortalamasının 500 kişilik örneklem büyüklüklerine sahip gruplardan elde edilen RMSD değerlerine nazaran daha düşük olduğu saptanmıştır. Diğer bir deyişle örneklem büyüklüğü arttıkça RMSD değerlerinin ortalaması düşmüştür.

Yetenek dağılımı değişkenine göre test eşitleme yöntemleri karşılaştırıldığında, benzer yetenek dağılımına sahip gruplara ait eşitleme çalışmalarından elde edilen RMSD değerleri ortalamasının farklı yetenek dağılımlarına sahip gruplardan elde edilen RMSD değerlerine nazaran daha düşük olduğu saptanmıştır. Diğer bir deyişle eşitleme yapılacak grupların yetenek dağılımları birbirine benzedikçe, RMSD değerlerinin ortalaması düşmüştür.

Hem benzer yetenek dağılımı hem de farklı yetenek dağılımı gösteren gruplar için, 500 ve 1000 kişilik örneklem büyüklüklerinde, Stocking-Lord

yönteminin diğer eşitleme yöntemlerine nazaran daha az eşitleme hatası içerdiği saptanmıştır. Ortalama-ortalama yöntemi ile ortalama-standart sapma yöntemi, tüm koşullarda, karakteristik eğri yöntemlerine nazaran daha fazla eşitleme hatası içermektedir. Bu araştırmanın bulgularına dayanarak, bu çalışmada kullanılan koşullar içinde, en iyi eşitleme koşulunun 1000 kişilik örneklem büyüklüğünde benzer yetenek dağılımına sahip gruplarda, Stocking-Lord eşitleme yönteminin kullanılması ile elde edilebildiği ifade edilebilir.

### Öneriler

Bu çalışmadan elde edilen bulgulara dayanarak, pratikte test eşitleme çalışmaları için Stocking-Lord test eşitleme yönteminin kullanılması ve 500 yerine 1000 kişilik örneklem büyüklüğünün tercih edilmesi önerilebilir. Bundan sonra test eşitleme alanında yapılabilecek araştırmalar için ise şu öneriler sunulabilir:

- Bu çalışmada eşitleme yöntemlerinin karşılaştırılmasında simülatif veriler kullanılmıştır. Bu eşitleme yöntemleri gerçek veriler üzerinde de karşılaştırılabilir.
- Bu çalışmada 1-0 şeklinde puanlanan veriler üzerinde eşitleme yöntemleri karşılaştırılmıştır. Çok kategorili puanlanan veya 1-0 ve çoklu kategorilerde puanlamanın birlikte yapıldığı veriler üzerinde çalışılarak benzer değişkenlerin eşitleme hataları ile ilişkileri araştırılabilir.
- Bu çalışmadan elde edilen bulgular, 500 ve 1000 kişilik örneklem gruplarıyla sınırlıdır. Daha küçük ve daha büyük örneklemelerde eşitleme yöntemleri karşılaştırılabilir.
- Bu çalışmada, Madde Tepki Kuramı'na dayalı eşitleme yöntemleri ele alınmış, Klasik Test Kuramı'na dayalı eşitleme yöntemleri araştırma

kapsamı dışında bırakılmıştır. Çeşitli koşullar altında Madde Tepki Kuramı ve Klasik Test Kuramı'na dayalı test eşitleme yöntemleri karşılaştırılabilir.

- Bu araştırmada tek boyutlu Madde Tepki Kuramı modelleri ile sınırlıdır. Çok boyutlu Madde Tepki Kuramı modelleri üzerinde eşitleme çalışmaları yapılabilir.
- Araştırmada benzer güçlükteki testlerin ortak maddelerle eşitlenmesine yönelik bir çalışma yapılmıştır. Aynı araştırma farklı güçlüklerdeki testler üzerinde tekrarlanabilir.
- Bu araştırma yetenek ve madde parametrelerinin ayrı ayrı kestirildiği ortalama-ortalama, ortalama-standart sapma, Haebara ile Stocking-Lord yöntemleri ile sınırlandırılmış; söz konusu parametrelerin eşzamanlı kestirimiyle elde edilen test eşitleme hataları bu araştırmanın kapsamı dışında bırakılmıştır. Çeşitli koşullarda parametrelerin eşzamanlı ve ayrı ayrı kestirildiği eşitleme yöntemlerine ait hatalar incelenebilir.
- Bu araştırmada test eşitleme hatalarını etkileyebileceği düşünülen örneklem büyüklüğü ve yetenek dağılımı gibi etkenler ele alınmıştır. Bu değişkenlerin yanı sıra, ortak madde uzunluğu, test uzunluğu, yetenek kestiriminde kullanılan bilgisayar programı gibi farklı değişkenlerin eşitleme hataları üzerindeki etkileri çalışılabilir.

## KAYNAKÇA

- Angoff, W. H. (1971). "Scales, Norms and Equivalent Scores". In Thorndike, R. L. (Ed.), *Educational Measurement* (p. 509-600). Washington: American Council on Education.
- Aiken, L. R. (2000). *Psychological Testing and Assessment*. Boston: Allyn and Bacon.
- Ackerman, T. (1992) A Didactic Explanation of Item Bias, Item Impact, and Item Validity from a Multidimensional Perspective. *Journal of Educational Measurement*, 29 (1), 67-91.
- Ackerman, T. (1994). Creating A Test Information Profile for a Two-Dimensional Latent Space. *Applied Psychological Measurement*, 18 (3), 257-276.
- Ackerman, T. (1996). Graphical Representation of Multidimensional Item Response Theory Analyses. *Applied Psychological Measurement*, 20 (4) 311-330.
- Baker, F. B. (1988). GENIRV: A Program to Generate Item Response Vectors [Computer program]. Madison: University of Wisconsin, Department of Educational Psychology, Laboratory of Experimental Design.
- Baker, F.B. and Al-Karni, A. (1991). A Comparison of Two Procedures for Computing IRT Equating Coefficients. *Journal of Educational Measurement*, 28 (2), 147-162.
- Baker, F. B. and Kim, S. (2004). *Item Response Theory: Parameter Estimation Techniques*. New York: Marcel Dekker.

- Bastari, B. (2000). *Linking Multiple Choice and Constructed Response Items to a Common Proficiency Scale*. Unpublished doctoral dissertation, University of Massachusetts, Amherst.
- Bozdağ, S. (2007). *Şans Başarısının Test Eşitlemeye Etkisi*. Yayımlanmamış yüksek lisans tezi, Mersin Üniversitesi, Sosyal Bilimler Enstitüsü, Mersin.
- Budescu, D. (1985). Efficiency of Linear Equating as a Function of the Length of the Anchor Test. *Journal of Educational Statistics*, 12 (1), 33-43.
- Caldwell, L. J. (1984). *A Comparison of Equating Error in Linear and Rasch Model Test Equating Method*. Unpublished doctoral dissertation, Florida State University, Tallahassee.
- Chen, H. W. (2001). *Calibration of the ITBS Test Battery to the Complete Test Battery: A Comparison Five Linking Methods*. Unpublished doctoral dissertation, University of Iowa, Iowa City.
- Cho, Y. (2007). *Comparison of Bootstrap Standard Errors of Equating Using IRT and Equipercentile Methods with Polytomously Scored Items Under the Common Item Nonequivalent Groups Design*. Unpublished doctoral dissertation, University of Iowa, Iowa City.
- Cohen, A. S. and Kim, S. H. (1998). An Investigation of Linking Methods Under the Graded Response Model. *Applied Psychological Measurement*, 22 (2), 116-130.
- Cook L. L. and Eignor R. E. (1991). An NCME Instructional Module on IRT Equating Methods. *Instructional Topics in Educational Measurement. Educational Measurement: Issues and Practice*, 10 (1), 37–45.

- Crocker, L. and Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. New York: Holt, Rinehart and Winston.
- Custer, M., Omar, M. H. and Pomplun, M (2006). Vertical Scaling with the Rasch Model Utilizing Default and Tight Convergence Settings with WINSTEPS and BILOG-MG. *Applied Measurement in Education*, 19, (2), 133-149.
- Çetin, E. (2009). *Dikey Ölçeklemede Klasik Test ve Madde Tepki Kuramına Dayalı Yöntemlerin Karşılaştırılması*. Yayımlanmamış yüksek lisans tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Demirtaşlı, N. ve Arıkan, S. (2008). *CITO Türkiye Öğrenci İzleme Sistemi (ÖİS) ÖİS’de MTK Uygulamaları*. 1. Eğitimde ve Psikolojide Ölçme ve Değerlendirme Kongresi’nde sunuldu, Ankara.
- Embretson, S. (1996) The New Rules of Measurement. *Psychological Assessment*, 8 (4), 341-349.
- Embretson, S. (1997) Multicomponent Response Models. In W. van der Linden and R. Hambleton (Eds), *Handbook of Modern Item Response Theory* (p. 305-321). Springer-Verlag: New York.
- Embretson, S. E. and Reise, S. P. (2000). *Item Response Theory for Psychologists*. London: Lawrence Erlbaum Associates, Publishers.
- Felan, G. D. (2002). Test Equating: Mean, Linear, Equipercentile and Item Response Theory. Paper presented at the *Annual Meeting of the Southwest Educational Research Association*, Austin.
- Haebara, T. (1980). Equating Logistic Ability Scales by a Weighted Least Squares Method. *Japanese Psychological Research*, 22 (3), 144-149.
- Hambleton, R. K. and Swaminathan, H. (1989). *Item Response Theory*,



*Principles and Applications*. USA: Kluwer Nijhoff Publishing.

Hambleton, R. K., Swaminathan, H. and Rogers, H. (1991). *Fundamentals of Item Response Theory*. Newbury Park CA: Sage.

Han, K. T. (2007a). WinGen: Windows Software That Generates IRT Parameters and Item Responses. *Applied Psychological Measurement*, 31 (5), 457-459.

Han, K. T. (2007b). IRTEQ: Windows Application That Implements IRT Scaling and Equating [Computer Program]. Web: <http://www.umass.edu/remp/software/irteq> adresinden 5 Mayıs 2008'de alınmıştır.

Han, K. T., and Hambleton, R. K. (2007). User's Manual: WinGen 2 (*Center for Educational Assessment Report No. 642*). Amherst, MA: University of Massachusetts.

Han, K. T., Kolen, M. and Pohlmann, J. (1997). A Comparison Among IRT True- and Observed- Score Equatings and Traditional Equipercetile Equating. *Applied Measurement in Education*, 10 (2), 105-121.

Hanson, B. A. and 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.

Hanson, B. A. and Zeng, L. (1995a). PIE. A Computer Program for IRT Equating [Computer software]. Iowa City: ACT.

Harris, D. J. (1993, April). *Practical Issue in Equating*. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, G. A.

- Harris, D. J. and Crouse, J. D. (1993). A Study of Criteria Used in Equating. *Applied Psychological Measurement*, 6 (3), 195-240.
- Harwell, M., Stone, C. A., Hsu, T. and Kirisci, L. (1996). Monte Carlo Studies in Item Response Theory. *Applied Psychological Measurement*, 20 (2),101-125.
- Hills, J. R., Subhiyah, R. G. and Hirsch, T. M. (1988). Equating Minimum-Competency Tests: Comparisons of Methods. *Journal of Educational Measurement*, 25 (3), 221-231.
- Holland, P. W. (2007). Data Collection Designs and Linking Procedures. In Dorans, N. J., Pommerich, M., Holland, P. W. (Eds.) *Linking and Aligning Scores and Scales*. USA: Springer.
- Hu, H. (2004). *Investigation of IRT Based Equating Methods in the Presence of Outliers*. Unpublished doctoral dissertation, University of Alberta, Alberta.
- Hung, P., Wu, Y., and Chen, Y. (1991). *IRT Item Parameter Linking: Relevant Issues for the Purpose of Item Banking*. Paper presented at the International Academic Symposium on Psychological Measurement, Taiwan.
- Jodoin, M. G., and Davey, T. (2003). *A multidimensional simulation approach to investigate the robustness of IRT common item equating*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.
- Karkee, T. B. And Wright, K. R. (2004, April). Evaluation of Linking Methods for Placing Three Parameter Logistic Item Parameter Estimates Onto Rasch Scale. Paper presented at the Meeting of the American

Educational Research, San Diego, California.

- Kaskowitz, G. S. (1998). *The Effect of Error in Item Parameter Estimates on Linking and Equating with the IRT Test Characteristic Curve Method*. Unpublished doctoral dissertation, University of Nebraska, Omaha.
- Kaskowitz, G. S. and 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 (1), 39-52.
- Kelecioğlu, H. (1994). *Öğrenci Seçme Sınavı Puanlarının Eşitlenmesi Üzerine Bir Çalışma*. Yayımlanmamış doktora tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.
- Kim, S. H. and Cohen, A. S. (2002). A Comparison of Linking and Concurrent Calibration Under the Graded Response Model. *Applied Psychological Measurement*, 26 (1), 25–41.
- Kim, S. and Kolen, M.J. (2003). POLYST, A Computer Program for Polytomous IRT Scale Transformation. [Computer software]. Iowa City: ACT.
- Kim, S. and Kolen, M.J. (2004). STUIRT A Computer Program for Scale Transformation under Unidimensional Item Response Theory Models. [Computer software]. Iowa City: ACT.
- Kim, S. and Kolen, M.J. (2004). *Optimally Defining Criterion Functions for the Characteristic Curve Procedures in the IRT Scale Linking*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Kim, S. and Kolen, M. J. (2006). Robustness of Format Effects of IRT Linking Methods for Mixed Format Tests. *Applied Measurement in Education*, 19 (4), 357-381.

- Kim, S., and Lee, W. (2004). *IRT Scale Linking Methods for Mixed-Format Tests* (ACT Research Report 2004–5). Iowa City, IA: ACT, Inc.
- Kim, S. and 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., and Song, M.-Y. (2004). *Least Squares Estimation of IRT Scale Linking Coefficients under the Graded Response Model*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Klein, L. W. and Jarjura, D. (1985). Importance of Content Representation for Common Item Equating with Nonrandom Groups. *Journal of Educational Measurement*, 22 (3), 197-206.
- Kolen, M. J. (1988). Traditional Equating Methodology. *Educational Measurement: Issues and Practice*, 7 (4), 29-36.
- Kolen, M. J. (2007). Data Collection Designs and Linking Procedures. In Dorans, N. J.; Pommerich, M. Holland, P. W. (Eds.), *Linking and Aligning Scores and Scales*. USA: Springer.
- Kolen, M. J. and Brennan, R. L. (2004). *Test Equating, Scalling, and Linking*. (second edition). USA: Springer.
- Lee, W. C. and Ban, J. C. (2010). A Comparison of IRT Linking Procedures. *Applied Measurement in Education*, 23 (1), 23-48.
- Li, Y.H., Lissitz, R.W., and Yang, Y. (1999). *Estimating IRT Equating Coefficients for Tests with Polytomously and Dichotomously Scored Items*. Paper presented at the annual meeting of the National Council on Measurement in Education, Montreal, Quebec, Canada.

- Linacre, J. M. and Wight, B. D. (1995). *A User's Guide to BIGSTEPS: Rasch Model computer program*. Chicago: MESA Press.
- Linn, R. L. (1993). Linking Results of Distinct Assessments. *Applied Measurement in Education*, 6 (1), 83–102.
- Livingston, S. (2004) *Equating Test Scores (without IRT)*. New Jersey: *Educational Testing Service*.
- Lord, M. F. (1980). *Application of Item Response Theory to Practical Testing Problems*. New Jersey: Lawrence Erlbaum Associates Publishers.
- Loyd, B. H., and Hoover, H. D. (1980). Vertical Equating Using the Rasch Model. *Journal of Educational Measurement*, 17 (3), 179-193.
- Marco, G. L. (1977). Item Characteristic Curve Solutions to Three Intractable Testing Problems. *Journal of Educational Measurement*, 14 (2), 139-160.
- Meyers, J. L., Miller, E. G. and Way, W. D. (2009). Item Position and Item Difficulty Change in an IRT-Based Common Item Equating Design. *Applied Measurement in Education*, 22 (1), 38-60.
- Michalides, M. P. (2003). *Effects of Common Item Selection on the Accuracy of Item Response Theory Test Equating with Nonequivalent Groups*. Unpublished doctoral dissertation, Stanford University, California.
- Muraki, E. And Bock, D. (2003). *PARSCALE User Guide*. Chicago, IL: Scientific Software International, Inc.
- Muraki, E., Hombo, C. M. and Lee, Y. W. (2000). Equating and Linking of Performance Assessments. *Applied Psychological Measurement*, 24 (4), 325-337.

- Ogasawara, H. (2001). Standard Errors of Item Response Theory Equating/Linking by Response Function Methods. *Applied Psychological Measurement*, 25 (1), 53–67.
- Oshima, T. C, Davey, T.C., and Lee, K. (2000). Multidimensional Linking: Four Practical Approaches. *Journal of Educational Measurement*, 37 (4), 357-373.
- Peak, I. and 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 (1), 199-215.
- Peterson, N. S., Cook, L. L. and Stocking, M. L. (1983). IRT Versus Conventional Equating Methods: A Comparative Study of Scale Stability. *Journal of Educational Statistics*, 8 (2), 137-156.
- Raju, N. S., and Arenson, E. A. (April, 2002). *Developing a Common Metric in Item Response Theory: An Area-Minimization Approach*. Paper presented at the annual meeting of the National Council on Measurement in Education, New Orleans, LA.
- Reckase, M. (1985) The Difficulty of Test Items That Measure More Than One Ability. *Applied Psychological Measurement*, 9 (4), 401-412.
- Reckase, M. (1997) A Linear Logistic Multidimensional Model for Dichotomous Item Response Data. In W. van der Linden and R. Hambleton (Eds.) *Handbook Of Modern Item Response Theory* (p.271-286). Springer, New York, USA.
- Sykes, R. C., Hou, L., Hanson, B. and Wang, Z. (2002). Multidimensionality and the Equating of a Mixed-Format Math Examination. Paper presented at the Annual Meeting of the National Council on

Measurement in Education, New Orleans.

Stocking, M. L. and Lord, F. M. (1983). Developing a Common Metric in Item Response Theory. *Applied Psychological Measurement*, 7 (2), 201-210.

Suanthong, S. (1998). *An Investigation of Factors Affecting Test Equating in Latent Trait Theory*. University of North Texas. Unpublished doctoral dissertation, University of North Texas, Denton.

Speron, E. (2009). A Comparison of Metric Linking Procedures in Item Response Theory. Unpublished doctoral dissertation, Illinois Institute of Technology, Chicago, Illinois.

Şahhüseyinoğlu, D. (2005). *İngilizce Yeterlik Sınavı Puanlarının Üç Farklı Eşitleme Yöntemine Göre Karşılaştırılması*. Yayımlanmamış doktora tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.

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.

Tong, Y. (2005). *Comparisons of Methodologies and Results in Vertical Scaling for Educational Achievement Tests*. Unpublished doctoral Dissertation, The University of Iowa, Iowa City.

Tsai, T. H., Hanson, B. A., Kolen, M. J. and Forsyth, R. A. (2001). A Comparison of Bootstrap Standard Errors of IRT Equating Methods for the Common-Item Nonequivalent Groups Design. *Applied Measurement in Education*, 14 (1), 17–30.

Vale, C. D. (1986). Linking Item Parameters Onto a Common Scale. *Applied Psychological Measurement*, 10 (4), 333-344.

- Velhelst, N. D. (2004). Reference Supplement to The Preliminary Pilot Version of The Manuel for Relating Language Examinations to the Common Eoropan Framework Of Reference for Languages: Learning, Teaching, Assessment, Section G: Item Response Theory. DGIV/EDU/LANG (2004) 13, Council of Europe Language Policy Division, Strasbourg.
- Wang, T., Hanson, B. and Haris, D. (2000). The Effectiveness of Circular Equating as a Criterion for Evaluating Equating. *Applied Psychological Measurement*, 24 (3), 195-210.
- Way, W. D., and Tang, K. L. (1991, April 4-6). *A Comparison of Four Logistic Model Equating Methods*. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago, IL.
- Wingersky, M. S., Barton, M. A. and Lord, R. M. (1982). *LOGIST User's Guide*. Princeton, NJ: Educational Testing Service.
- Wingersky, M. S., Cook, L. L. and Eignor, D. R. (1987). Specifying The Characteristics of Linking Items Used For Item Response Theory Item Calibration. Princeton, NJ: ETS.
- Wolkowitz, A. A. (2008). *A Comparison of Classical Test Theory and Item Response Theory Methods for Equating Number-Right Scored to Formula Scored Assessments*. Unpublished doctoral dissertation, University of Kansas, Kansas.
- Yang, W. L. (1997). *The Effects of Content Homogeneity and Equating Method on the Accuracy of Common Item Test Equating*. Unpublished doctoral dissertation, Michigan State University, Michigan.



- Yi, (2008). Mixed-Format Test Equating: Effects Of Test Dimensionality and Common-Item Sets. Unpublished doctoral dissertation, University of Maryland, Maryland.
- Yotinprasert, S. (1986), *The Effect of Sample Size on Error Produced by Tucker and Rasch Equating Methods under Common Items Nonrandom Grups Design*. Unpublished doctoral dissertation, Florida State University, Tallahassee.
- Yu, C. H. and Osborn Popp, S. H. (2005). Test Equating by Common Items and Common Subjects: Concepts and Applications. *Practical Assessment Research & Evaluation*, 10 (4). <http://pareonline.net/getvn.asp?v=10&n=4>. adresinden 03 Mart 2009 tarihinde erişilmiştir.
- Zimowski, M. F., Muraki, E., Mislevy, R. J. and Bock, R. D. (1996). *BLOG-MG: Multiple-Group IRT Analysis and Test Maintenance for Binary Items*. Chicago: Scientific Software International.

## EKLER

## Ek 1. 500 Kişilik Benzer Yetenek Dağılımına Sahip Grupların Eşitlenmesinden Elde Edilen RMSD Değerleri

	A	B	RMSD	A	B	RMSD	A	B	RMSD	A	B	RMSD
1	1,060	-0,050	<b>0,076</b>	0,990	-0,070	<b>0,071</b>	1,000	-0,010	<b>0,010</b>	0,990	-0,020	<b>0,022</b>
2	1,150	0,050	<b>0,153</b>	1,040	0,030	<b>0,049</b>	1,080	0,050	<b>0,092</b>	1,100	0,040	<b>0,105</b>
3	1,090	-0,050	<b>0,101</b>	1,030	-0,070	<b>0,076</b>	1,050	-0,030	<b>0,057</b>	1,070	-0,040	<b>0,079</b>
4	0,940	-0,160	<b>0,170</b>	1,080	-0,130	<b>0,151</b>	0,960	-0,030	<b>0,049</b>	0,950	-0,020	<b>0,052</b>
5	0,930	0,020	<b>0,070</b>	0,910	0,010	<b>0,087</b>	1,000	-0,050	<b>0,050</b>	1,010	-0,050	<b>0,051</b>
6	1,060	0,050	<b>0,076</b>	1,040	0,040	<b>0,055</b>	1,020	0,050	<b>0,054</b>	1,000	0,040	<b>0,040</b>
7	0,970	0,080	<b>0,085</b>	0,840	0,040	<b>0,159</b>	0,930	0,030	<b>0,074</b>	0,860	0,040	<b>0,140</b>
8	0,970	-0,040	<b>0,067</b>	0,980	-0,030	<b>0,036</b>	0,950	-0,030	<b>0,056</b>	0,930	-0,040	<b>0,078</b>
9	0,960	-0,150	<b>0,154</b>	0,830	-0,170	<b>0,236</b>	0,920	-0,140	<b>0,160</b>	0,910	-0,150	<b>0,173</b>
10	1,040	0,010	<b>0,040</b>	1,160	0,040	<b>0,160</b>	1,070	-0,040	<b>0,079</b>	1,100	-0,040	<b>0,105</b>
11	0,990	-0,120	<b>0,120</b>	1,040	-0,110	<b>0,116</b>	0,970	-0,100	<b>0,104</b>	0,990	-0,120	<b>0,120</b>
12	1,080	0,160	<b>0,178</b>	0,990	0,120	<b>0,120</b>	1,140	0,060	<b>0,148</b>	1,130	0,040	<b>0,132</b>
13	1,000	0,130	<b>0,130</b>	1,030	0,140	<b>0,143</b>	1,010	0,130	<b>0,130</b>	1,010	0,140	<b>0,140</b>
14	1,050	0,030	<b>0,057</b>	0,960	0,010	<b>0,040</b>	1,030	-0,030	<b>0,042</b>	1,060	-0,030	<b>0,065</b>
15	1,140	-0,070	<b>0,152</b>	1,110	-0,070	<b>0,036</b>	1,080	-0,050	<b>0,057</b>	1,100	-0,050	<b>0,079</b>
16	0,910	-0,060	<b>0,105</b>	0,920	-0,060	<b>0,098</b>	0,960	-0,100	<b>0,107</b>	0,970	-0,110	<b>0,114</b>
17	0,950	0,100	<b>0,111</b>	0,970	0,110	<b>0,114</b>	0,940	0,140	<b>0,152</b>	0,900	0,140	<b>0,170</b>
18	0,990	-0,010	<b>0,014</b>	1,020	0,000	<b>0,019</b>	0,990	-0,020	<b>0,022</b>	0,980	0,000	<b>0,019</b>
19	0,960	-0,110	<b>0,117</b>	0,910	-0,120	<b>0,148</b>	0,930	-0,080	<b>0,105</b>	0,920	-0,070	<b>0,104</b>
20	1,010	-0,100	<b>0,100</b>	1,080	-0,090	<b>0,119</b>	1,030	-0,040	<b>0,049</b>	1,010	-0,030	<b>0,032</b>
21	0,990	-0,110	<b>0,110</b>	0,910	-0,130	<b>0,156</b>	0,950	-0,150	<b>0,157</b>	0,970	-0,180	<b>0,182</b>
22	1,040	-0,020	<b>0,044</b>	1,020	-0,020	<b>0,028</b>	1,010	-0,080	<b>0,081</b>	1,000	0,070	<b>0,070</b>
23	1,000	0,000	<b>0,000</b>	0,950	-0,010	<b>0,049</b>	0,970	-0,020	<b>0,035</b>	0,990	0,010	<b>0,014</b>
24	0,980	0,180	<b>0,181</b>	0,850	0,130	<b>0,194</b>	0,990	0,110	<b>0,110</b>	1,010	0,110	<b>0,110</b>
25	0,940	0,110	<b>0,124</b>	0,910	0,100	<b>0,132</b>	0,980	0,080	<b>0,082</b>	0,940	0,060	<b>0,083</b>
26	1,070	-0,090	<b>0,113</b>	1,070	-0,090	<b>0,113</b>	1,140	-0,090	<b>0,162</b>	1,120	-0,090	<b>0,146</b>
27	1,230	0,050	<b>0,227</b>	1,270	0,060	<b>0,267</b>	1,190	0,060	<b>0,192</b>	1,130	0,070	<b>0,144</b>
28	1,030	-0,190	<b>0,192</b>	0,970	-0,200	<b>0,202</b>	0,980	-0,190	<b>0,191</b>	0,950	-0,180	<b>0,186</b>
29	1,160	-0,010	<b>0,154</b>	1,280	0,030	<b>0,270</b>	1,170	0,060	<b>0,174</b>	1,150	0,080	<b>0,164</b>
30	1,120	0,040	<b>0,122</b>	0,940	-0,010	<b>0,058</b>	1,060	-0,050	<b>0,076</b>	1,070	-0,060	<b>0,090</b>
31	0,940	-0,020	<b>0,061</b>	0,870	-0,040	<b>0,132</b>	0,920	-0,080	<b>0,111</b>	0,880	-0,070	<b>0,135</b>

32	1,020	0,080	<b>0,082</b>	0,980	0,070	<b>0,073</b>	0,950	0,100	<b>0,111</b>	0,930	0,120	<b>0,138</b>
33	0,980	-0,070	<b>0,073</b>	1,040	-0,060	<b>0,071</b>	1,040	-0,110	<b>0,117</b>	1,020	-0,100	<b>0,102</b>
34	0,990	-0,040	<b>0,041</b>	0,950	-0,050	<b>0,069</b>	0,940	0,030	<b>0,065</b>	0,950	0,030	<b>0,057</b>
35	1,060	-0,040	<b>0,070</b>	0,990	-0,050	<b>0,051</b>	0,980	-0,020	<b>0,028</b>	0,940	-0,010	<b>0,059</b>
36	0,970	-0,090	<b>0,095</b>	0,960	-0,090	<b>0,098</b>	1,000	-0,130	<b>0,130</b>	1,000	-0,140	<b>0,140</b>
37	1,100	0,080	<b>0,125</b>	1,110	0,080	<b>0,133</b>	1,110	0,080	<b>0,133</b>	1,080	0,090	<b>0,139</b>
38	0,900	-0,010	<b>0,097</b>	0,790	0,040	<b>0,207</b>	0,920	-0,040	<b>0,087</b>	0,890	-0,030	<b>0,110</b>
39	1,050	-0,110	<b>0,120</b>	1,040	-0,110	<b>0,116</b>	1,050	-0,090	<b>0,101</b>	1,050	-0,080	<b>0,093</b>
40	1,040	-0,130	<b>0,136</b>	1,250	-0,070	<b>0,249</b>	1,030	0,000	<b>0,029</b>	1,010	0,030	<b>0,031</b>
41	0,950	0,030	<b>0,057</b>	0,990	0,040	<b>0,041</b>	1,010	0,030	<b>0,031</b>	1,010	-0,010	<b>0,014</b>
42	0,960	0,110	<b>0,117</b>	1,010	0,120	<b>0,120</b>	1,070	0,050	<b>0,084</b>	1,060	0,050	<b>0,076</b>
43	1,080	0,040	<b>0,087</b>	0,920	-0,020	<b>0,080</b>	1,050	-0,070	<b>0,085</b>	1,070	-0,070	<b>0,097</b>
44	1,070	0,020	<b>0,071</b>	1,040	0,020	<b>0,043</b>	1,060	-0,020	<b>0,061</b>	1,040	-0,010	<b>0,039</b>
45	1,110	-0,090	<b>0,139</b>	1,110	-0,090	<b>0,139</b>	1,190	-0,140	<b>0,231</b>	1,180	-0,160	<b>0,236</b>
46	1,050	-0,080	<b>0,093</b>	1,030	-0,090	<b>0,095</b>	0,960	-0,040	<b>0,056</b>	0,960	-0,040	<b>0,056</b>
47	0,980	-0,150	<b>0,151</b>	0,960	-0,150	<b>0,155</b>	0,970	-0,150	<b>0,153</b>	0,970	-0,140	<b>0,143</b>
48	1,010	-0,030	<b>0,032</b>	0,990	-0,030	<b>0,032</b>	1,010	-0,010	<b>0,014</b>	1,000	0,000	<b>0,000</b>
49	0,960	0,040	<b>0,056</b>	0,860	0,010	<b>0,135</b>	0,940	0,020	<b>0,061</b>	0,950	0,020	<b>0,051</b>
50	1,140	-0,010	<b>0,135</b>	1,210	0,000	<b>0,203</b>	1,220	-0,100	<b>0,235</b>	1,230	-0,110	<b>0,248</b>
51	1,020	-0,090	<b>0,092</b>	0,910	-0,100	<b>0,132</b>	1,010	-0,150	<b>0,150</b>	1,040	-0,160	<b>0,165</b>
52	0,910	0,020	<b>0,089</b>	0,910	0,030	<b>0,092</b>	0,930	0,000	<b>0,068</b>	0,950	-0,020	<b>0,052</b>
53	1,220	-0,060	<b>0,221</b>	1,060	-0,090	<b>0,107</b>	1,210	-0,120	<b>0,236</b>	1,210	-0,120	<b>0,236</b>
54	0,900	-0,020	<b>0,980</b>	0,770	-0,050	<b>0,226</b>	0,920	-0,090	<b>0,118</b>	0,940	-0,100	<b>0,115</b>
55	0,950	0,160	<b>0,167</b>	0,910	0,150	<b>0,173</b>	1,000	0,060	<b>0,060</b>	1,030	0,040	<b>0,049</b>
56	0,930	0,050	<b>0,084</b>	0,850	0,030	<b>0,148</b>	1,070	0,000	<b>0,068</b>	1,110	-0,010	<b>0,107</b>
57	0,990	0,110	<b>0,110</b>	0,900	0,090	<b>0,132</b>	1,000	0,030	<b>0,030</b>	1,020	0,010	<b>0,022</b>
58	1,000	0,010	<b>0,010</b>	0,910	-0,020	<b>0,088</b>	1,010	-0,040	<b>0,041</b>	1,000	-0,030	<b>0,030</b>
59	0,110	-0,160	<b>0,192</b>	0,980	-0,180	<b>0,181</b>	0,960	-0,080	<b>0,089</b>	0,980	-0,070	<b>0,073</b>
60	1,050	-0,080	<b>0,093</b>	1,020	-0,090	<b>0,092</b>	1,030	-0,080	<b>0,085</b>	1,060	-0,090	<b>0,107</b>
61	1,010	-0,030	<b>0,032</b>	0,960	-0,050	<b>0,063</b>	1,000	-0,010	<b>0,010</b>	0,970	-0,020	<b>0,035</b>
62	0,900	-0,080	<b>0,125</b>	0,790	-0,110	<b>0,229</b>	0,920	-0,110	<b>0,134</b>	0,910	-0,090	<b>0,125</b>
63	0,970	0,020	<b>0,035</b>	0,950	0,010	<b>0,049</b>	1,020	-0,040	<b>0,044</b>	1,030	-0,040	<b>0,049</b>
64	1,010	-0,140	<b>0,140</b>	1,100	-0,130	<b>0,161</b>	1,020	-0,120	<b>0,121</b>	0,990	-0,100	<b>0,100</b>
65	1,000	0,130	<b>0,130</b>	0,940	0,110	<b>0,124</b>	1,010	0,120	<b>0,120</b>	1,030	0,130	<b>0,133</b>
66	0,960	0,090	<b>0,098</b>	1,020	0,110	<b>0,111</b>	0,990	0,080	<b>0,081</b>	0,980	0,070	<b>0,073</b>

67	0,990	0,090	<b>0,091</b>	0,860	0,050	<b>0,143</b>	0,930	0,110	<b>0,129</b>	0,940	0,100	<b>0,115</b>
68	1,060	-0,020	<b>1,061</b>	1,010	-0,040	<b>0,041</b>	1,050	0,010	<b>0,049</b>	1,040	0,010	<b>0,039</b>
69	1,000	0,030	<b>0,030</b>	1,040	0,040	<b>0,056</b>	0,980	0,040	<b>0,044</b>	1,010	0,040	<b>0,041</b>
70	1,020	-0,110	<b>0,112</b>	1,050	-0,100	<b>0,111</b>	0,980	-0,090	<b>0,092</b>	1,000	-0,080	<b>0,080</b>
71	0,970	-0,070	<b>0,076</b>	0,840	-0,100	<b>0,183</b>	0,900	-0,090	<b>0,132</b>	0,900	-0,080	<b>0,125</b>
72	0,980	0,050	<b>0,054</b>	1,060	0,070	<b>0,091</b>	1,030	0,050	<b>0,058</b>	1,030	0,060	<b>0,068</b>
73	0,850	-0,070	<b>0,160</b>	1,020	-0,040	<b>0,044</b>	0,920	0,030	<b>0,083</b>	0,900	0,030	<b>0,101</b>
74	0,950	-0,060	<b>0,077</b>	0,930	-0,070	<b>0,097</b>	0,940	-0,060	<b>0,083</b>	0,930	-0,050	<b>0,084</b>
75	1,050	0,080	<b>0,094</b>	1,100	0,100	<b>0,139</b>	1,080	0,070	<b>0,104</b>	1,080	0,050	<b>0,092</b>
76	1,020	0,050	<b>0,054</b>	0,920	0,020	<b>0,080</b>	0,990	0,020	<b>0,022</b>	1,030	0,020	<b>0,035</b>
77	1,110	0,120	<b>0,160</b>	0,890	0,050	<b>0,117</b>	1,030	-0,020	<b>0,035</b>	1,050	-0,050	<b>0,070</b>
78	0,970	-0,040	<b>0,049</b>	0,098	-0,030	<b>0,036</b>	1,000	-0,070	<b>0,070</b>	1,010	-0,070	<b>0,071</b>
79	1,130	0,030	<b>0,129</b>	1,030	0,010	<b>0,031</b>	1,120	0,000	<b>0,116</b>	1,130	0,010	<b>0,126</b>
80	0,870	0,080	<b>0,148</b>	0,860	0,070	<b>0,151</b>	0,840	0,100	<b>0,183</b>	0,810	0,080	<b>0,198</b>
81	0,910	0,080	<b>0,118</b>	0,810	0,050	<b>0,189</b>	0,990	-0,040	<b>0,041</b>	0,960	-0,040	<b>0,056</b>
82	1,060	-0,020	<b>0,061</b>	1,040	-0,030	<b>0,049</b>	1,080	-0,010	<b>0,078</b>	1,100	0,010	<b>0,097</b>
83	0,950	-0,010	<b>0,049</b>	1,110	0,040	<b>0,113</b>	1,050	-0,030	<b>0,057</b>	1,040	-0,040	<b>0,056</b>
84	0,940	-0,100	<b>0,116</b>	1,020	-0,080	<b>0,082</b>	0,960	-0,050	<b>0,063</b>	0,950	-0,040	<b>0,063</b>
85	1,030	0,040	<b>0,049</b>	0,980	0,030	<b>0,036</b>	1,020	0,000	<b>0,019</b>	1,030	0,000	<b>0,029</b>
86	1,010	0,040	<b>0,041</b>	0,990	0,030	<b>0,032</b>	0,990	0,030	<b>0,032</b>	0,980	0,040	<b>0,044</b>
87	1,050	-0,140	<b>0,148</b>	1,040	-0,140	<b>0,145</b>	1,070	-0,160	<b>0,174</b>	1,090	-0,150	<b>0,173</b>
88	0,990	-0,110	<b>0,110</b>	1,070	-0,090	<b>0,113</b>	1,080	-0,090	<b>0,119</b>	1,070	-0,100	<b>0,121</b>
89	0,910	0,120	<b>0,147</b>	0,980	0,140	<b>0,141</b>	0,950	0,160	<b>0,167</b>	0,950	0,170	<b>0,177</b>
90	1,020	0,080	<b>0,082</b>	1,010	0,080	<b>0,081</b>	1,010	0,060	<b>0,061</b>	1,000	0,040	<b>0,040</b>
91	1,140	0,030	<b>0,139</b>	0,980	-0,010	<b>0,022</b>	1,090	-0,060	<b>0,106</b>	1,110	-0,070	<b>0,127</b>
92	1,040	0,160	<b>0,165</b>	1,040	0,160	<b>0,165</b>	1,040	0,130	<b>0,136</b>	1,020	0,130	<b>0,147</b>
93	1,050	-0,020	<b>0,052</b>	0,990	-0,030	<b>0,032</b>	1,070	-0,120	<b>0,138</b>	1,090	-0,140	<b>0,165</b>
94	1,120	0,000	<b>0,116</b>	1,090	-0,010	<b>0,088</b>	1,110	-0,020	<b>0,108</b>	1,100	-0,020	<b>0,099</b>
95	1,040	-0,060	<b>0,071</b>	1,010	-0,070	<b>0,071</b>	1,230	-0,130	<b>0,257</b>	1,230	-0,140	<b>0,262</b>
96	0,950	0,080	<b>0,093</b>	0,910	0,070	<b>0,112</b>	0,920	0,090	<b>0,119</b>	0,920	0,100	<b>0,126</b>
97	0,990	-0,070	<b>0,071</b>	0,980	-0,070	<b>0,073</b>	1,050	-0,150	<b>0,158</b>	1,070	-0,150	<b>0,164</b>
98	0,980	-0,120	<b>0,122</b>	0,950	-0,120	<b>0,129</b>	1,030	-0,180	<b>0,182</b>	1,050	-0,170	<b>0,177</b>
99	1,030	-0,050	<b>0,058</b>	1,120	-0,020	<b>0,118</b>	1,110	-0,080	<b>0,133</b>	1,070	-0,080	<b>0,104</b>
100	0,940	0,080	<b>0,099</b>	0,880	0,060	<b>0,130</b>	0,970	0,020	<b>0,035</b>	0,970	0,000	<b>0,029</b>

Ek 2. 500 Kişilik Farklı Yetenek Dağılımına Sahip Grupların Eşitlenmesinden Elde Edilen RMSD Değerleri

	A	B	RMSD	A	B	RMSD	A	B	RMSD	A	B	RMSD
1	1,150	-0,170	<b>0,223</b>	1,450	-0,110	<b>0,448</b>	1,130	-0,030	<b>0,129</b>	1,100	0,000	<b>0,097</b>
2	1,130	0,080	<b>0,149</b>	1,060	0,060	<b>0,083</b>	1,090	0,020	<b>0,890</b>	1,110	0,030	<b>0,110</b>
3	1,150	-0,150	<b>0,209</b>	1,220	-0,130	<b>0,249</b>	1,200	-0,170	<b>0,257</b>	1,190	-0,120	<b>0,219</b>
4	0,990	-0,150	<b>0,150</b>	1,110	-0,130	<b>0,167</b>	1,050	-0,130	<b>0,139</b>	1,050	-0,120	<b>0,129</b>
5	0,910	0,040	<b>0,095</b>	1,050	0,080	<b>0,093</b>	1,040	0,000	<b>0,038</b>	1,080	-0,010	<b>0,078</b>
6	0,960	-0,050	<b>0,063</b>	1,190	0,020	<b>0,184</b>	1,100	-0,010	<b>0,097</b>	1,070	0,020	<b>0,070</b>
7	0,970	0,090	<b>0,095</b>	0,960	0,080	<b>0,089</b>	0,900	0,110	<b>0,146</b>	0,890	0,090	<b>0,139</b>
8	1,040	-0,160	<b>0,164</b>	0,960	-0,170	<b>0,174</b>	0,970	-0,120	<b>0,124</b>	0,980	-0,120	<b>0,122</b>
9	1,060	-0,020	<b>0,061</b>	1,210	0,020	<b>0,202</b>	1,130	0,000	<b>0,124</b>	1,070	0,030	<b>0,073</b>
10	0,960	-0,080	<b>0,089</b>	0,930	-0,080	<b>0,104</b>	0,920	-0,070	<b>0,104</b>	1,000	-0,130	<b>0,130</b>
11	1,110	-0,300	<b>0,318</b>	1,650	-0,220	<b>0,665</b>	1,230	-0,170	<b>0,279</b>	1,200	-0,130	<b>0,232</b>
12	0,990	0,060	<b>0,061</b>	0,970	0,050	<b>0,058</b>	0,980	0,040	<b>0,044</b>	0,990	0,030	<b>0,032</b>
13	0,960	0,070	<b>0,080</b>	1,080	0,110	<b>0,134</b>	1,020	0,020	<b>0,028</b>	1,030	0,020	<b>0,035</b>
14	1,180	-0,030	<b>0,176</b>	1,100	-0,050	<b>0,109</b>	1,200	-0,130	<b>0,233</b>	1,230	-0,150	<b>0,268</b>
15	1,180	-0,190	<b>0,258</b>	1,240	-0,180	<b>0,293</b>	1,210	-0,170	<b>0,264</b>	1,200	-0,120	<b>0,227</b>
16	0,610	-0,030	<b>0,380</b>	0,870	0,030	<b>0,129</b>	0,860	-0,030	<b>0,329</b>	0,960	-0,090	<b>0,098</b>
17	1,070	-0,030	<b>0,074</b>	1,420	0,090	<b>0,415</b>	1,150	0,100	<b>0,176</b>	1,060	0,140	<b>0,151</b>
18	0,970	-0,030	<b>0,042</b>	1,090	0,010	<b>0,087</b>	1,050	-0,040	<b>0,062</b>	1,050	-0,040	<b>-0,062</b>
19	1,040	-0,110	<b>0,117</b>	1,310	-0,030	<b>0,301</b>	1,170	-0,010	<b>0,164</b>	1,140	0,040	<b>0,141</b>
20	1,030	-0,130	<b>0,133</b>	1,210	-0,090	<b>0,222</b>	1,160	-0,080	<b>0,174</b>	1,150	-0,050	<b>0,153</b>
21	1,070	-0,090	<b>0,117</b>	1,100	-0,080	<b>0,125</b>	1,070	-0,090	<b>0,112</b>	1,110	-0,110	<b>0,153</b>
22	1,130	-0,340	<b>0,362</b>	1,660	-0,210	<b>0,671</b>	1,300	-0,240	<b>0,376</b>	1,240	-0,220	<b>0,320</b>
23	1,110	0,070	<b>0,127</b>	1,140	0,070	<b>0,152</b>	1,120	0,030	<b>0,120</b>	1,120	0,080	<b>0,141</b>
24	0,970	0,140	<b>0,143</b>	1,030	0,190	<b>0,192</b>	1,020	0,150	<b>0,151</b>	1,070	0,140	<b>0,155</b>
25	0,980	-0,070	<b>0,073</b>	1,100	-0,030	<b>0,101</b>	1,050	-0,030	<b>0,057</b>	1,010	-0,040	<b>0,040</b>
26	0,060	-0,010	<b>0,398</b>	1,220	-0,090	<b>0,231</b>	1,030	-0,050	<b>0,058</b>	1,100	-0,080	<b>0,125</b>
27	1,110	0,010	<b>0,107</b>	1,320	0,060	<b>0,315</b>	1,130	0,060	<b>0,139</b>	1,050	0,090	<b>0,102</b>
28	0,910	-0,130	<b>0,156</b>	0,970	-0,120	<b>0,123</b>	0,940	-0,190	<b>0,199</b>	1,040	-0,240	<b>0,243</b>
29	0,990	-0,200	<b>0,200</b>	1,720	0,030	<b>0,696</b>	1,280	-0,040	<b>0,273</b>	1,290	-0,050	<b>0,285</b>
30	1,080	-0,220	<b>0,233</b>	1,240	-0,170	<b>0,287</b>	1,080	-0,170	<b>0,187</b>	1,100	-0,170	<b>0,195</b>
31	1,170	-0,290	<b>0,333</b>	1,280	-0,260	<b>0,374</b>	1,140	-0,260	<b>0,272</b>	1,130	-0,240	<b>0,270</b>

32	1,080	-0,010	<b>0,078</b>	1,090	-0,010	<b>0,087</b>	1,030	-0,070	<b>0,076</b>	1,090	-0,070	<b>0,110</b>
33	0,920	0,060	<b>0,098</b>	0,920	0,060	<b>0,098</b>	0,920	0,020	<b>0,079</b>	0,960	-0,020	<b>0,043</b>
34	0,680	-0,220	<b>0,379</b>	1,080	-0,130	<b>0,151</b>	1,020	-0,070	<b>0,072</b>	1,100	-0,130	<b>0,162</b>
35	1,050	-0,230	<b>0,235</b>	1,290	-0,170	<b>0,327</b>	1,080	-0,130	<b>0,151</b>	1,010	-0,080	<b>0,130</b>
36	0,590	-0,040	<b>0,398</b>	0,940	0,050	<b>0,076</b>	0,930	-0,010	<b>0,068</b>	1,090	-0,130	<b>0,156</b>
37	1,040	0,070	<b>0,080</b>	1,210	0,130	<b>0,241</b>	1,090	0,120	<b>0,148</b>	1,030	0,140	<b>0,143</b>
38	1,030	-0,110	<b>0,114</b>	1,010	-0,120	<b>0,120</b>	1,030	-0,160	<b>0,163</b>	1,120	-0,210	<b>0,240</b>
39	1,040	-0,240	<b>0,243</b>	1,320	-0,170	<b>0,353</b>	1,060	-0,170	<b>0,180</b>	1,140	-0,220	<b>0,258</b>
40	0,980	0,110	<b>0,112</b>	1,030	0,120	<b>0,123</b>	0,990	0,120	<b>0,120</b>	0,970	0,130	<b>0,133</b>
41	0,990	0,090	<b>0,091</b>	0,970	0,080	<b>0,085</b>	1,020	0,050	<b>0,054</b>	1,080	-0,020	<b>0,080</b>
42	1,000	-0,200	<b>0,200</b>	1,320	-0,110	<b>0,327</b>	1,190	-0,210	<b>0,278</b>	1,210	-0,180	<b>0,278</b>
43	0,920	0,030	<b>0,083</b>	1,010	0,060	<b>0,061</b>	0,940	0,030	<b>0,065</b>	0,970	0,060	<b>0,066</b>
44	1,130	-0,250	<b>0,279</b>	1,550	-0,140	<b>0,549</b>	1,280	-0,160	<b>0,314</b>	1,230	-0,110	<b>0,247</b>
45	1,030	-0,080	<b>0,085</b>	1,070	-0,070	<b>0,097</b>	1,110	-0,050	<b>0,117</b>	1,140	-0,090	<b>0,162</b>
46	1,080	-0,240	<b>0,252</b>	1,700	-0,040	<b>0,676</b>	1,210	-0,050	<b>0,208</b>	1,160	-0,010	<b>0,155</b>
47	0,990	-0,170	<b>0,170</b>	0,980	-0,170	<b>0,171</b>	0,980	-0,140	<b>0,141</b>	0,970	-0,120	<b>0,123</b>
48	1,040	-0,130	<b>0,135</b>	1,070	-0,130	<b>0,146</b>	1,050	-0,180	<b>0,186</b>	1,120	-0,190	<b>0,222</b>
49	1,020	0,080	<b>0,082</b>	1,090	0,100	<b>0,132</b>	1,010	0,110	<b>0,110</b>	1,040	0,080	<b>0,089</b>
50	1,130	-0,140	<b>0,188</b>	1,170	-0,130	<b>0,209</b>	1,200	-0,190	<b>0,270</b>	1,250	-0,260	<b>0,354</b>
51	1,120	-0,280	<b>0,302</b>	1,190	-0,270	<b>0,325</b>	1,110	-0,290	<b>0,308</b>	1,140	-0,270	<b>0,301</b>
52	1,160	-0,230	<b>0,277</b>	1,500	-0,150	<b>0,505</b>	1,250	-0,130	<b>0,274</b>	1,200	-0,110	<b>0,222</b>
53	1,000	0,130	<b>0,130</b>	0,830	0,090	<b>0,187</b>	0,870	0,080	<b>0,148</b>	0,980	0,010	<b>0,021</b>
54	0,970	-0,050	<b>0,058</b>	0,900	-0,060	<b>0,113</b>	0,980	-0,170	<b>0,171</b>	1,110	-0,270	<b>0,290</b>
55	0,950	0,070	<b>0,085</b>	1,100	0,120	<b>0,154</b>	0,980	0,070	<b>0,073</b>	1,040	0,040	<b>0,055</b>
56	1,040	-0,040	<b>0,055</b>	1,220	0,010	<b>0,212</b>	1,110	0,010	<b>0,106</b>	1,070	0,040	<b>0,078</b>
57	1,140	-0,260	<b>0,293</b>	1,160	-0,250	<b>0,293</b>	1,110	-0,250	<b>0,272</b>	1,140	-0,240	<b>0,275</b>
58	0,940	-0,070	<b>0,091</b>	1,050	-0,030	<b>0,057</b>	0,990	-0,070	<b>0,071</b>	0,980	-0,080	<b>0,082</b>
59	0,990	0,030	<b>0,032</b>	0,910	0,020	<b>0,089</b>	0,910	0,030	<b>0,092</b>	0,890	0,080	<b>0,082</b>
60	1,080	-0,160	<b>0,178</b>	1,150	-0,140	<b>0,201</b>	1,080	-0,120	<b>0,142</b>	1,100	-0,090	<b>0,104</b>
61	1,140	0,050	<b>0,144</b>	1,180	0,070	<b>0,187</b>	1,170	0,020	<b>0,165</b>	1,170	0,000	<b>0,164</b>
62	0,950	0,000	<b>0,048</b>	0,890	-0,010	<b>0,106</b>	0,990	-0,100	<b>0,100</b>	1,020	-0,020	<b>0,027</b>
63	0,960	-0,060	<b>0,071</b>	1,190	-0,010	<b>0,183</b>	1,100	-0,030	<b>0,100</b>	1,060	0,020	<b>0,061</b>
64	1,090	-0,260	<b>0,274</b>	1,330	-0,220	<b>0,386</b>	1,240	-0,300	<b>0,378</b>	1,150	-0,220	<b>0,263</b>
65	0,930	0,070	<b>0,097</b>	1,040	0,110	<b>0,116</b>	0,990	0,090	<b>0,090</b>	0,970	0,100	<b>0,104</b>
66	0,790	0,070	<b>0,214</b>	1,110	0,180	<b>0,208</b>	1,060	0,110	<b>0,124</b>	1,110	0,080	<b>0,132</b>

67	1,010	-0,130	<b>0,130</b>	1,130	-0,100	<b>0,160</b>	1,020	-0,040	<b>0,044</b>	0,990	0,010	<b>0,013</b>
68	1,110	-0,070	<b>0,127</b>	1,030	-0,090	<b>0,095</b>	1,070	-0,060	<b>0,090</b>	1,020	-0,010	<b>0,022</b>
69	0,910	-0,010	<b>0,087</b>	1,240	0,090	<b>0,248</b>	1,020	0,040	<b>0,044</b>	0,950	0,080	<b>0,093</b>
70	1,060	-0,090	<b>0,106</b>	1,220	-0,060	<b>0,220</b>	1,050	-0,020	<b>0,052</b>	1,030	0,000	<b>0,028</b>
71	1,060	-0,160	<b>0,170</b>	0,910	-0,190	<b>0,208</b>	0,960	-0,170	<b>0,174</b>	0,950	-0,100	<b>0,110</b>
72	0,930	-0,130	<b>0,146</b>	1,440	0,010	<b>0,424</b>	1,200	-0,090	<b>0,212</b>	1,180	-0,030	<b>0,176</b>
73	1,020	-0,230	<b>0,230</b>	1,280	-0,170	<b>0,318</b>	1,070	-0,130	<b>0,146</b>	1,040	-0,070	<b>0,080</b>
74	1,020	-0,060	<b>0,063</b>	0,960	-0,070	<b>0,079</b>	0,930	-0,090	<b>0,112</b>	0,970	-0,090	<b>0,066</b>
75	1,090	0,050	<b>0,100</b>	1,000	0,020	<b>0,020</b>	1,080	-0,010	<b>0,077</b>	1,140	-0,060	<b>0,147</b>
76	1,010	-0,110	<b>0,140</b>	1,120	-0,080	<b>0,140</b>	0,980	-0,100	<b>0,101</b>	1,040	-0,110	<b>0,116</b>
77	1,040	-0,120	<b>0,126</b>	1,130	-0,090	<b>0,154</b>	1,180	-0,180	<b>0,250</b>	1,210	-0,220	<b>0,299</b>
78	1,060	-0,230	<b>0,237</b>	1,270	-0,190	<b>0,321</b>	1,140	-0,220	<b>0,257</b>	1,150	-0,200	<b>0,246</b>
79	1,090	0,060	<b>0,105</b>	1,030	0,090	<b>0,049</b>	1,050	0,020	<b>0,052</b>	1,060	0,010	<b>0,058</b>
80	0,920	-0,250	<b>0,261</b>	1,250	-0,140	<b>0,277</b>	1,050	-0,120	<b>0,129</b>	1,030	-0,100	<b>0,104</b>
81	0,990	-0,220	<b>0,220</b>	1,440	-0,080	<b>0,431</b>	1,130	-0,160	<b>0,203</b>	1,030	-0,120	<b>0,123</b>
82	1,040	-0,060	<b>0,071</b>	1,230	-0,020	<b>0,222</b>	1,120	-0,040	<b>0,122</b>	1,130	-0,030	<b>0,128</b>
83	0,950	0,180	<b>0,186</b>	0,870	0,160	<b>0,203</b>	0,920	0,100	<b>0,126</b>	1,040	0,030	<b>0,048</b>
84	0,990	-0,170	<b>0,170</b>	1,150	-0,140	<b>0,201</b>	1,080	-0,180	<b>0,195</b>	1,090	-0,150	<b>0,173</b>
85	1,010	0,120	<b>0,120</b>	0,910	0,080	<b>0,118</b>	0,960	0,040	<b>0,055</b>	1,050	-0,010	<b>0,049</b>
86	0,680	0,010	<b>0,308</b>	0,850	0,080	<b>0,165</b>	0,920	0,070	<b>0,104</b>	1,080	-0,010	<b>0,077</b>
87	1,120	-0,220	<b>0,248</b>	1,240	-0,200	<b>0,305</b>	1,160	-0,200	<b>0,252</b>	1,160	-0,160	<b>0,222</b>
88	0,990	-0,500	<b>0,500</b>	1,930	-0,310	<b>0,949</b>	1,300	-0,250	<b>0,382</b>	1,220	-0,150	<b>0,259</b>
89	0,970	-0,060	<b>0,066</b>	1,240	0,020	<b>0,231</b>	1,090	-0,020	<b>0,088</b>	1,080	0,000	<b>0,076</b>
90	1,150	-0,250	<b>0,288</b>	1,270	-0,220	<b>0,340</b>	1,270	-0,240	<b>0,354</b>	1,310	-0,280	<b>0,382</b>
91	1,210	-0,360	<b>0,413</b>	1,780	-0,210	<b>0,783</b>	1,370	-0,230	<b>0,425</b>	1,400	-0,270	<b>0,471</b>
92	1,050	0,100	<b>0,111</b>	1,220	0,160	<b>0,265</b>	1,130	0,140	<b>0,188</b>	1,080	0,150	<b>0,168</b>
93	1,060	-0,120	<b>0,133</b>	1,380	-0,050	<b>0,369</b>	1,210	-0,080	<b>0,217</b>	1,200	-0,090	<b>0,212</b>
94	1,090	-0,090	<b>0,125</b>	1,600	0,050	<b>0,582</b>	1,320	-0,010	<b>0,309</b>	1,210	0,020	<b>0,204</b>
95	1,080	-0,050	<b>0,092</b>	1,030	-0,060	<b>0,066</b>	1,040	-0,040	<b>0,055</b>	1,060	-0,060	<b>0,083</b>
96	1,070	-0,180	<b>0,192</b>	1,530	-0,060	<b>0,514</b>	1,180	-0,100	<b>0,200</b>	1,160	-0,080	<b>0,173</b>
97	1,000	-0,060	<b>0,060</b>	1,150	-0,020	<b>0,145</b>	1,070	-0,050	<b>0,083</b>	1,090	-0,080	<b>0,117</b>
98	0,650	-0,060	<b>0,342</b>	0,870	-0,020	<b>0,126</b>	0,880	-0,070	<b>0,135</b>	1,010	-0,130	<b>0,130</b>
99	1,050	-0,280	<b>0,284</b>	1,130	-0,260	<b>0,288</b>	1,100	-0,280	<b>0,296</b>	1,100	-0,280	<b>0,296</b>
100	0,970	0,180	<b>0,182</b>	0,960	0,180	<b>0,184</b>	1,010	0,140	<b>0,140</b>	0,980	0,120	<b>0,121</b>

Ek 3. 1000 Kişilik Benzer Yetenek Dağılımına Sahip Grupların Eşitlenmesinden Elde Edilen RMSD Değerleri

	A	B	RMSD	A	B	RMSD	A	B	RMSD	A	B	RMSD
1	0,880	-0,010	<b>0,116</b>	0,900	0,000	<b>0,097</b>	0,930	0,010	<b>0,068</b>	0,930	0,020	<b>0,071</b>
2	1,030	0,110	<b>0,113</b>	0,990	0,100	<b>0,100</b>	1,040	0,050	<b>0,063</b>	1,040	0,040	<b>0,055</b>
3	0,980	-0,070	<b>0,072</b>	1,070	-0,050	<b>0,100</b>	1,010	-0,040	<b>0,041</b>	1,010	-0,040	<b>0,041</b>
4	1,000	-0,010	<b>0,010</b>	1,180	0,040	<b>0,178</b>	1,030	0,090	<b>0,049</b>	1,010	0,100	<b>0,100</b>
5	1,090	0,030	<b>0,091</b>	0,980	0,000	<b>0,019</b>	1,020	-0,050	<b>0,053</b>	1,040	-0,060	<b>0,071</b>
6	1,070	0,010	<b>0,068</b>	0,930	-0,030	<b>0,074</b>	1,070	0,070	<b>0,097</b>	1,090	-0,060	<b>0,105</b>
7	1,060	-0,100	<b>0,115</b>	0,990	-0,110	<b>0,110</b>	1,060	-0,100	<b>0,115</b>	1,040	-0,090	<b>0,097</b>
8	0,910	-0,020	<b>0,089</b>	0,930	-0,020	<b>0,129</b>	0,940	0,020	<b>0,061</b>	0,920	0,010	<b>0,077</b>
9	0,900	0,010	<b>0,096</b>	0,940	0,020	<b>0,061</b>	0,890	0,020	<b>0,107</b>	0,900	0,010	<b>0,096</b>
10	1,000	0,120	<b>0,120</b>	1,010	0,120	<b>0,120</b>	0,980	0,090	<b>0,092</b>	1,000	0,090	<b>0,090</b>
11	0,950	0,040	<b>0,062</b>	0,970	0,040	<b>0,049</b>	0,940	0,040	<b>0,070</b>	0,960	0,030	<b>0,065</b>
12	0,960	-0,020	<b>0,043</b>	1,030	0,000	<b>0,028</b>	0,990	-0,020	<b>0,022</b>	1,000	-0,010	<b>0,010</b>
13	1,150	0,000	<b>0,144</b>	1,210	0,020	<b>0,202</b>	1,100	-0,010	<b>0,096</b>	1,120	-0,010	<b>0,115</b>
14	1,040	-0,110	<b>0,116</b>	1,120	-0,090	<b>0,146</b>	1,010	0,000	<b>0,009</b>	0,970	0,000	<b>0,028</b>
15	1,020	0,100	<b>0,101</b>	0,990	0,100	<b>0,100</b>	0,090	0,020	<b>0,089</b>	1,120	0,020	<b>0,117</b>
16	1,000	-0,060	<b>0,060</b>	1,000	-0,060	<b>0,060</b>	0,980	-0,050	<b>0,053</b>	0,960	-0,050	<b>0,063</b>
17	0,900	0,030	<b>0,048</b>	0,960	0,040	<b>0,055</b>	0,960	0,010	<b>0,039</b>	0,960	0,020	<b>0,043</b>
18	0,940	0,030	<b>0,065</b>	0,930	0,030	<b>0,073</b>	0,970	0,040	<b>0,049</b>	0,960	0,050	<b>0,055</b>
19	0,990	0,000	<b>0,009</b>	0,860	-0,040	<b>0,140</b>	0,970	-0,040	<b>0,049</b>	0,990	-0,040	<b>0,041</b>
20	1,050	0,050	<b>0,069</b>	1,060	0,050	<b>0,076</b>	1,090	0,060	<b>0,105</b>	1,080	0,060	<b>0,097</b>
21	1,130	0,040	<b>0,131</b>	1,020	0,000	<b>0,019</b>	1,100	0,040	<b>0,104</b>	1,080	0,050	<b>0,091</b>
22	0,940	0,100	<b>0,115</b>	0,970	0,110	<b>0,113</b>	1,010	0,020	<b>0,022</b>	1,020	-0,010	<b>0,053</b>
23	1,060	-0,050	<b>0,050</b>	1,120	-0,040	<b>0,122</b>	1,080	-0,060	<b>0,097</b>	1,060	-0,070	<b>0,090</b>
24	1,040	0,140	<b>0,145</b>	1,020	0,130	<b>0,131</b>	1,020	0,130	<b>0,131</b>	0,980	0,130	<b>0,131</b>
25	0,960	-0,100	<b>0,107</b>	0,970	-0,100	<b>0,104</b>	0,970	-0,110	<b>0,113</b>	0,960	-0,110	<b>0,116</b>
26	1,060	0,070	<b>0,090</b>	1,010	0,060	<b>0,060</b>	1,010	0,100	<b>0,100</b>	1,000	0,120	<b>0,120</b>
27	0,900	0,130	<b>0,161</b>	0,790	0,080	<b>0,217</b>	0,960	0,050	<b>0,063</b>	0,960	0,050	<b>0,063</b>
28	1,020	14,000	<b>0,141</b>	0,990	0,130	<b>0,130</b>	1,020	0,130	<b>0,131</b>	1,020	0,130	<b>0,131</b>
29	0,980	-0,080	<b>0,082</b>	0,990	-0,070	<b>0,070</b>	0,990	-0,100	<b>0,100</b>	0,990	-0,100	<b>0,100</b>
30	1,110	0,000	<b>0,106</b>	1,070	-0,010	<b>0,068</b>	1,140	-0,010	<b>0,168</b>	1,130	0,010	<b>0,143</b>
31	0,970	0,040	<b>0,049</b>	1,040	0,060	<b>0,071</b>	0,980	0,060	<b>0,063</b>	0,980	0,070	<b>0,072</b>



32	1,090	0,030	<b>0,049</b>	1,090	0,030	<b>0,049</b>	1,050	0,040	<b>0,062</b>	1,050	0,040	<b>0,062</b>
33	0,950	-0,090	<b>0,102</b>	1,040	-0,060	<b>0,071</b>	0,950	-0,030	<b>0,056</b>	0,940	-0,040	<b>0,070</b>
34	0,990	-0,120	<b>0,120</b>	1,020	-0,120	<b>0,121</b>	1,040	-0,060	<b>0,071</b>	1,030	-0,050	<b>0,057</b>
35	0,890	0,040	<b>0,113</b>	0,880	0,040	<b>0,122</b>	0,890	0,080	<b>0,132</b>	0,910	0,100	<b>0,132</b>
36	1,060	-0,030	<b>0,065</b>	1,090	-0,020	<b>0,089</b>	1,070	0,020	<b>0,070</b>	1,050	0,030	<b>0,056</b>
37	1,110	-0,010	<b>0,106</b>	1,100	-0,010	<b>0,097</b>	1,050	0,060	<b>0,076</b>	1,040	0,060	<b>0,071</b>
38	1,020	-0,070	<b>0,072</b>	0,930	-0,080	<b>0,104</b>	0,970	-0,060	<b>0,066</b>	0,960	-0,040	<b>0,055</b>
39	1,010	-0,030	<b>0,031</b>	1,000	-0,030	<b>0,030</b>	1,020	0,000	<b>0,019</b>	1,000	0,010	<b>0,010</b>
40	1,030	-0,060	<b>0,066</b>	1,040	-0,060	<b>0,071</b>	0,990	-0,070	<b>0,070</b>	0,980	-0,070	<b>0,072</b>
41	1,020	-0,210	<b>0,210</b>	1,060	-0,200	<b>0,208</b>	1,020	-0,180	<b>0,181</b>	1,010	-0,180	<b>0,180</b>
42	1,030	-0,030	<b>0,041</b>	1,040	-0,020	<b>0,203</b>	1,040	-0,020	<b>0,043</b>	1,030	-0,030	<b>0,041</b>
43	0,950	0,070	<b>0,085</b>	0,890	0,050	<b>0,117</b>	1,010	0,050	<b>0,050</b>	0,990	0,060	<b>0,060</b>
44	0,940	-0,040	<b>0,070</b>	0,990	-0,030	<b>0,031</b>	1,020	-0,060	<b>0,063</b>	1,040	-0,060	<b>0,071</b>
45	1,050	-0,150	<b>0,157</b>	1,100	-0,140	<b>0,170</b>	1,010	-0,070	<b>0,070</b>	0,990	-0,060	<b>0,060</b>
46	0,980	0,000	<b>0,019</b>	1,020	0,010	<b>0,021</b>	1,000	0,020	<b>0,020</b>	1,000	0,030	<b>0,030</b>
47	1,010	0,060	<b>0,060</b>	1,050	0,080	<b>0,093</b>	1,090	0,040	<b>0,095</b>	1,090	0,040	<b>0,095</b>
48	1,030	-0,030	<b>0,041</b>	0,990	-0,040	<b>0,041</b>	1,030	-0,080	<b>0,085</b>	1,040	-0,080	<b>0,088</b>
49	0,950	0,000	<b>0,048</b>	0,950	0,000	<b>0,048</b>	0,970	-0,030	<b>0,041</b>	0,980	-0,040	<b>0,044</b>
50	1,030	-0,140	<b>0,142</b>	1,280	-0,080	<b>0,281</b>	1,050	-0,010	<b>0,049</b>	1,050	0,010	<b>0,049</b>
51	1,040	0,040	<b>0,055</b>	1,110	0,050	<b>0,117</b>	1,090	0,040	<b>0,095</b>	1,110	0,030	<b>0,110</b>
52	0,970	-0,080	<b>0,085</b>	1,000	-0,070	<b>0,070</b>	0,980	-0,030	<b>0,035</b>	0,970	-0,010	<b>0,030</b>
53	1,010	0,010	<b>0,138</b>	0,990	0,000	<b>0,009</b>	1,020	-0,050	<b>0,053</b>	1,030	-0,040	<b>0,066</b>
54	1,040	-0,030	<b>0,048</b>	1,020	-0,030	<b>0,035</b>	1,000	-0,010	<b>0,010</b>	0,980	-0,010	<b>0,021</b>
55	0,990	-0,060	<b>0,060</b>	1,080	-0,040	<b>0,086</b>	1,050	-0,050	<b>0,069</b>	1,030	-0,060	<b>0,066</b>
56	1,000	0,050	<b>0,050</b>	0,950	0,040	<b>0,062</b>	0,990	-0,020	<b>0,022</b>	0,990	-0,040	<b>0,041</b>
57	1,040	-0,080	<b>0,088</b>	0,970	-0,100	<b>0,104</b>	0,970	-0,090	<b>0,094</b>	0,960	-0,100	<b>0,107</b>
58	1,040	-0,030	<b>0,048</b>	1,010	-0,040	<b>0,041</b>	1,040	-0,090	<b>0,097</b>	1,040	-0,090	<b>0,097</b>
59	1,010	-0,040	<b>0,041</b>	0,880	-0,070	<b>0,135</b>	0,980	-0,030	<b>0,035</b>	1,000	-0,020	<b>0,020</b>
60	1,170	-0,110	<b>0,197</b>	1,070	-0,130	<b>0,146</b>	1,130	-0,090	<b>0,154</b>	1,120	-0,070	<b>0,135</b>
61	0,990	-0,070	<b>0,070</b>	0,940	-0,080	<b>0,098</b>	0,970	-0,010	<b>0,030</b>	0,950	-0,020	<b>0,052</b>
62	0,990	0,000	<b>0,057</b>	0,840	-0,040	<b>0,159</b>	0,990	-0,060	<b>0,060</b>	1,010	-0,070	<b>0,070</b>
63	1,070	-0,050	<b>0,084</b>	1,030	-0,060	<b>0,066</b>	0,970	-0,040	<b>0,049</b>	0,960	-0,040	<b>0,055</b>
64	1,020	-0,100	<b>0,101</b>	0,950	-0,120	<b>0,129</b>	1,060	-0,130	<b>0,142</b>	1,060	-0,140	<b>0,151</b>
65	1,050	0,080	<b>0,093</b>	0,990	0,060	<b>0,060</b>	1,030	0,030	<b>0,041</b>	1,010	0,040	<b>0,041</b>
66	1,100	0,050	<b>0,108</b>	1,130	0,060	<b>0,139</b>	1,120	0,000	<b>0,115</b>	1,100	0,000	<b>0,096</b>

67	1,070	0,050	<b>0,084</b>	1,060	0,040	<b>0,070</b>	1,110	-0,080	<b>0,132</b>	1,120	-0,110	<b>0,159</b>
68	1,070	0,020	<b>0,070</b>	1,090	0,030	<b>0,091</b>	1,050	0,040	<b>0,062</b>	1,030	0,050	<b>0,057</b>
69	1,000	-0,050	<b>0,050</b>	1,060	-0,040	<b>0,070</b>	0,980	-0,010	<b>0,021</b>	0,980	-0,010	<b>0,021</b>
70	1,020	0,080	<b>0,082</b>	0,990	0,070	<b>0,070</b>	1,100	0,090	<b>0,131</b>	1,100	0,100	<b>0,138</b>
71	1,080	-0,060	<b>0,097</b>	0,980	-0,090	<b>0,092</b>	1,010	-0,070	<b>0,070</b>	1,010	-0,050	<b>0,050</b>
72	1,010	-0,110	<b>0,110</b>	1,030	-0,110	<b>0,113</b>	1,040	-0,130	<b>0,135</b>	1,040	-0,120	<b>0,126</b>
73	0,920	-0,010	<b>0,077</b>	0,980	0,010	<b>0,021</b>	0,960	-0,040	<b>0,055</b>	0,980	-0,040	<b>0,044</b>
74	1,000	-0,010	<b>0,010</b>	1,100	0,010	<b>0,097</b>	1,060	0,020	<b>0,061</b>	1,050	0,020	<b>0,052</b>
75	0,900	0,190	<b>0,213</b>	0,870	0,170	<b>0,211</b>	0,910	0,120	<b>0,148</b>	0,930	0,120	<b>0,137</b>
76	0,960	-0,050	<b>0,063</b>	1,180	-0,010	<b>0,173</b>	1,020	-0,030	<b>0,035</b>	1,000	-0,040	<b>0,040</b>
77	1,000	0,070	<b>0,070</b>	1,010	0,080	<b>0,080</b>	1,020	0,100	<b>0,152</b>	1,020	0,110	<b>0,111</b>
78	1,140	0,040	<b>0,140</b>	1,040	0,010	<b>0,039</b>	1,070	0,010	<b>0,120</b>	1,070	0,020	<b>0,070</b>
79	0,970	-0,050	<b>0,057</b>	1,050	-0,030	<b>0,056</b>	0,980	0,070	<b>0,072</b>	0,960	-0,070	<b>0,076</b>
80	0,940	-0,030	<b>0,065</b>	0,910	-0,040	<b>0,095</b>	1,000	-0,070	<b>0,070</b>	1,020	-0,070	<b>0,072</b>
81	1,020	-0,150	<b>0,151</b>	1,120	-0,130	<b>0,174</b>	1,030	-0,090	<b>0,094</b>	1,000	-0,080	<b>0,080</b>
82	0,990	0,030	<b>0,031</b>	1,000	0,030	<b>0,030</b>	0,990	0,060	<b>0,060</b>	1,000	0,060	<b>0,060</b>
83	1,030	-0,020	<b>0,035</b>	0,980	-0,030	<b>0,035</b>	1,050	-0,050	<b>0,069</b>	1,070	-0,060	<b>0,090</b>
84	0,900	0,010	<b>0,097</b>	0,920	0,020	<b>0,079</b>	0,940	0,020	<b>0,061</b>	0,940	0,010	<b>0,058</b>
85	1,120	0,030	<b>0,119</b>	1,000	0,000	<b>0,000</b>	1,040	0,060	<b>1,071</b>	1,030	0,050	<b>0,057</b>
86	1,120	0,070	<b>0,135</b>	0,980	0,040	<b>0,044</b>	1,100	0,040	<b>0,104</b>	1,110	0,050	<b>0,117</b>
87	0,940	-0,120	<b>0,133</b>	0,980	-0,110	<b>0,111</b>	0,900	-0,060	<b>0,113</b>	0,880	-0,060	<b>0,130</b>
88	1,120	-0,170	<b>0,205</b>	1,060	-0,180	<b>0,189</b>	1,000	-0,100	<b>0,100</b>	0,950	-0,080	<b>0,093</b>
89	0,970	0,050	<b>0,057</b>	0,920	0,030	<b>0,151</b>	0,990	0,010	<b>0,013</b>	0,950	0,010	<b>0,049</b>
90	1,040	0,120	<b>0,126</b>	0,940	0,080	<b>0,098</b>	0,980	0,060	<b>0,063</b>	0,990	0,070	<b>0,070</b>
91	0,870	0,120	<b>0,173</b>	0,880	0,120	<b>0,166</b>	0,960	0,060	<b>0,071</b>	0,980	0,050	<b>0,053</b>
92	0,950	-0,070	<b>0,085</b>	1,110	-0,030	<b>0,110</b>	1,010	-0,040	<b>0,041</b>	1,000	-0,020	<b>0,020</b>
93	0,850	-0,140	<b>0,201</b>	0,880	-0,130	<b>0,174</b>	0,870	-0,120	<b>0,173</b>	0,870	-0,120	<b>0,173</b>
94	1,070	0,110	<b>0,120</b>	1,060	0,100	<b>0,115</b>	1,070	0,070	<b>0,097</b>	1,080	0,070	<b>0,104</b>
95	1,010	0,040	<b>0,041</b>	0,850	0,000	<b>0,144</b>	0,960	0,020	<b>0,039</b>	0,960	0,010	<b>0,039</b>
96	1,060	-0,090	<b>0,107</b>	1,150	-0,060	<b>0,156</b>	1,030	-0,050	<b>0,057</b>	1,020	-0,040	<b>0,044</b>
97	1,000	-0,060	<b>0,060</b>	0,990	-0,070	<b>0,070</b>	1,060	-0,100	<b>0,115</b>	1,050	-0,100	<b>0,111</b>
98	1,050	-0,010	<b>0,049</b>	1,060	-0,010	<b>0,058</b>	1,050	0,010	<b>0,049</b>	1,050	0,030	<b>0,056</b>
99	0,900	-0,110	<b>0,146</b>	0,910	-0,110	<b>0,140</b>	0,910	-0,050	<b>0,100</b>	0,900	-0,050	<b>0,108</b>
100	0,970	0,010	<b>0,030</b>	1,000	0,010	<b>0,010</b>	0,990	0,020	<b>0,022</b>	0,980	0,020	<b>0,027</b>

Ek 4. 1000 Kişilik Farklı Yetenek Dağılımına Sahip Grupların Eşitlenmesinden Elde Edilen RMSD Değerleri

	A	B	RMSD	A	B	RMSD	A	B	RMSD	A	B	RMSD
1	0,970	-0,010	<b>0,030</b>	1,030	0,010	<b>0,030</b>	1,000	-0,040	<b>0,040</b>	1,030	-0,030	<b>0,041</b>
2	1,090	-0,030	<b>0,091</b>	1,120	-0,030	<b>0,119</b>	1,110	-0,090	<b>0,138</b>	1,090	-0,080	<b>0,117</b>
3	0,980	-0,020	<b>0,027</b>	1,090	0,000	<b>0,086</b>	1,040	-0,060	<b>0,088</b>	1,100	-0,060	<b>0,113</b>
4	1,020	-0,010	<b>0,021</b>	1,260	0,050	<b>0,255</b>	1,140	-0,010	<b>0,135</b>	1,120	0,030	<b>0,119</b>
5	0,970	-0,030	<b>0,041</b>	1,330	0,070	<b>0,325</b>	1,060	0,000	<b>0,057</b>	1,110	-0,040	<b>0,113</b>
6	1,000	0,070	<b>0,070</b>	0,910	0,040	<b>0,095</b>	0,930	0,000	<b>0,067</b>	1,000	-0,020	<b>0,020</b>
7	1,100	-0,230	<b>0,249</b>	1,320	-0,190	<b>0,362</b>	1,180	-0,120	<b>0,210</b>	1,120	-0,070	<b>0,135</b>
8	1,010	-0,140	<b>0,140</b>	1,160	-0,100	<b>0,183</b>	1,090	-0,100	<b>0,132</b>	1,080	-0,100	<b>0,126</b>
9	0,960	0,050	<b>0,063</b>	1,010	0,070	<b>0,070</b>	0,990	-0,010	<b>0,013</b>	1,010	0,000	<b>0,009</b>
10	0,980	-0,030	<b>0,053</b>	1,130	0,020	<b>0,126</b>	1,050	-0,050	<b>0,069</b>	1,080	-0,040	<b>0,086</b>
11	1,000	-0,010	<b>0,010</b>	0,930	-0,030	<b>0,073</b>	0,970	-0,080	<b>0,085</b>	1,040	-0,130	<b>0,135</b>
12	1,100	-0,110	<b>0,146</b>	1,350	-0,040	<b>0,339</b>	1,190	-0,100	<b>0,208</b>	1,190	-0,080	<b>0,199</b>
13	1,080	0,040	<b>0,086</b>	1,110	0,050	<b>0,117</b>	1,100	0,020	<b>0,222</b>	1,090	0,050	<b>0,100</b>
14	1,070	0,020	<b>0,070</b>	0,990	0,010	<b>0,013</b>	1,000	0,030	<b>0,030</b>	1,000	0,040	<b>0,040</b>
15	1,110	-0,020	<b>0,107</b>	1,090	-0,030	<b>0,091</b>	1,090	-0,040	<b>0,095</b>	1,100	-0,020	<b>0,098</b>
16	1,110	-0,080	<b>0,132</b>	1,190	-0,060	<b>0,192</b>	1,140	-0,080	<b>0,156</b>	1,130	-0,090	<b>0,154</b>
17	0,910	-0,220	<b>0,236</b>	1,180	-0,170	<b>0,242</b>	1,080	-0,240	<b>0,252</b>	1,110	-0,210	<b>0,262</b>
18	1,020	-0,130	<b>0,131</b>	1,060	-0,120	<b>0,133</b>	1,040	-0,050	<b>0,063</b>	0,990	-0,040	<b>0,041</b>
19	0,940	-0,040	<b>0,070</b>	1,070	0,000	<b>0,067</b>	0,950	-0,010	<b>0,049</b>	0,990	-0,020	<b>0,022</b>
20	0,970	-0,010	<b>0,030</b>	1,410	0,120	<b>0,413</b>	1,120	0,050	<b>0,126</b>	1,160	0,050	<b>0,162</b>
21	0,950	0,050	<b>0,069</b>	0,950	0,050	<b>0,089</b>	0,960	0,040	<b>0,055</b>	0,010	0,010	<b>0,013</b>
22	1,000	0,090	<b>0,090</b>	1,000	0,090	<b>0,090</b>	1,030	-0,010	<b>0,030</b>	1,120	-0,060	<b>0,130</b>
23	1,100	-0,140	<b>0,169</b>	1,420	-0,080	<b>0,412</b>	1,230	-0,070	<b>0,232</b>	1,160	-0,040	<b>0,159</b>
24	1,170	-0,040	<b>0,168</b>	1,250	-0,020	<b>0,241</b>	1,160	0,020	<b>0,155</b>	1,080	0,040	<b>0,086</b>
25	1,010	-0,330	<b>0,330</b>	1,650	-0,210	<b>0,660</b>	1,190	-0,200	<b>0,271</b>	1,120	-0,160	<b>0,197</b>
26	1,090	0,020	<b>0,089</b>	1,170	0,040	<b>0,168</b>	1,140	0,000	<b>0,134</b>	1,150	0,020	<b>0,145</b>
27	1,000	0,000	<b>0,000</b>	1,120	0,060	<b>0,130</b>	1,060	0,010	<b>0,058</b>	1,070	0,000	<b>0,067</b>
28	1,080	-0,030	<b>0,082</b>	1,080	-0,030	<b>0,082</b>	1,080	-0,030	<b>0,082</b>	1,060	-0,020	<b>0,061</b>
29	0,990	-0,060	<b>0,060</b>	0,960	-0,060	<b>0,060</b>	0,980	-0,100	<b>0,101</b>	1,010	-0,120	<b>0,120</b>
30	1,180	-0,040	<b>0,178</b>	1,320	0,000	<b>0,308</b>	1,240	-0,040	<b>0,234</b>	0,180	0,010	<b>0,173</b>
31	0,930	0,150	<b>0,164</b>	0,930	0,150	<b>0,164</b>	0,930	0,090	<b>0,112</b>	0,970	0,070	<b>0,075</b>

32	1,070	0,090	<b>0,112</b>	1,130	0,110	<b>0,167</b>	1,080	0,040	<b>0,086</b>	1,090	0,050	<b>0,100</b>
33	0,920	0,010	<b>0,078</b>	1,040	0,050	<b>0,063</b>	0,990	0,000	<b>0,009</b>	1,000	-0,010	<b>0,010</b>
34	0,950	-0,090	<b>0,102</b>	1,010	-0,080	<b>0,080</b>	0,950	-0,070	<b>0,085</b>	0,930	-0,050	<b>0,084</b>
35	0,940	-0,060	<b>0,083</b>	0,970	-0,060	<b>0,066</b>	0,940	-0,010	<b>0,125</b>	1,010	-0,090	<b>0,090</b>
36	1,110	-0,320	<b>0,337</b>	1,450	-0,240	<b>0,496</b>	1,180	-0,160	<b>0,236</b>	1,110	-0,090	<b>0,139</b>
37	1,130	-0,090	<b>0,154</b>	1,240	-0,060	<b>0,239</b>	1,210	-0,100	<b>0,226</b>	1,190	-0,080	<b>0,200</b>
38	1,090	-0,120	<b>0,219</b>	1,190	-0,110	<b>0,214</b>	1,100	-0,130	<b>0,102</b>	1,030	-0,080	<b>0,085</b>
39	1,070	-0,120	<b>0,137</b>	1,170	-0,090	<b>0,187</b>	1,120	-0,120	<b>0,166</b>	1,090	-0,010	<b>0,132</b>
40	0,960	0,090	<b>0,097</b>	0,970	0,090	<b>0,094</b>	0,990	0,010	<b>0,013</b>	1,030	-0,040	<b>0,049</b>
41	0,930	-0,150	<b>0,164</b>	0,940	-0,150	<b>0,160</b>	0,990	-0,170	<b>0,170</b>	1,000	-0,170	<b>0,170</b>
42	1,100	-0,100	<b>0,139</b>	1,150	-0,080	<b>0,165</b>	1,140	-0,100	<b>0,168</b>	1,140	-0,110	<b>0,170</b>
43	1,020	0,070	<b>0,072</b>	0,920	0,040	<b>0,087</b>	0,970	0,080	<b>0,085</b>	0,960	0,070	<b>0,079</b>
44	1,130	-0,180	<b>0,219</b>	1,320	-0,140	<b>0,339</b>	1,190	-0,170	<b>0,250</b>	1,190	-0,150	<b>0,237</b>
45	0,940	0,000	<b>0,058</b>	0,990	0,010	<b>0,013</b>	0,990	-0,010	<b>0,013</b>	0,990	0,020	<b>0,022</b>
46	1,100	-0,090	<b>0,132</b>	1,230	-0,060	<b>0,230</b>	1,120	-0,080	<b>0,140</b>	1,130	-0,070	<b>0,143</b>
47	0,940	0,100	<b>0,115</b>	0,970	0,110	<b>0,113</b>	0,980	0,080	<b>0,082</b>	1,070	0,040	<b>0,078</b>
48	0,930	0,090	<b>0,112</b>	0,920	0,090	<b>0,118</b>	0,920	0,020	<b>0,079</b>	0,970	-0,020	<b>0,035</b>
49	1,010	0,010	<b>0,013</b>	1,220	0,080	<b>0,227</b>	1,030	0,030	<b>0,041</b>	1,060	-0,020	<b>0,061</b>
50	0,990	-0,300	<b>0,300</b>	1,540	-0,160	<b>0,546</b>	1,230	-0,180	<b>0,289</b>	1,190	-0,140	<b>0,230</b>
51	1,020	-0,040	<b>0,400</b>	1,310	0,030	<b>0,301</b>	1,180	-0,010	<b>0,174</b>	1,140	0,020	<b>0,136</b>
52	1,110	-0,440	<b>0,452</b>	1,870	-0,230	<b>0,872</b>	1,180	-0,140	<b>0,223</b>	1,150	-0,100	<b>0,176</b>
53	1,000	0,060	<b>0,060</b>	0,890	0,030	<b>0,110</b>	0,930	0,010	<b>0,068</b>	1,000	-0,020	<b>0,020</b>
54	1,060	-0,130	<b>0,142</b>	1,410	-0,050	<b>0,399</b>	1,200	-0,130	<b>0,233</b>	1,180	-0,100	<b>0,200</b>
55	1,120	-0,100	<b>0,153</b>	1,320	-0,050	<b>0,313</b>	1,250	-0,130	<b>0,274</b>	1,230	-0,130	<b>0,257</b>
56	0,890	-0,020	<b>0,108</b>	1,290	0,110	<b>0,301</b>	1,030	0,010	<b>0,030</b>	1,080	-0,050	<b>0,092</b>
57	1,010	-0,040	<b>0,104</b>	1,080	-0,010	<b>0,078</b>	1,050	-0,060	<b>0,077</b>	1,060	-0,080	<b>0,098</b>
58	1,100	-0,010	<b>0,097</b>	1,030	-0,030	<b>0,041</b>	1,070	-0,010	<b>0,120</b>	1,150	-0,130	<b>0,194</b>
59	1,000	-0,030	<b>0,030</b>	0,970	-0,040	<b>0,049</b>	1,010	-0,060	<b>0,060</b>	1,050	-0,070	<b>0,085</b>
60	1,070	-0,190	<b>0,201</b>	1,110	-0,180	<b>0,209</b>	1,110	-0,230	<b>0,253</b>	1,160	-0,240	<b>0,285</b>
61	1,120	-0,220	<b>0,248</b>	1,280	-0,180	<b>0,325</b>	1,160	-0,150	<b>0,215</b>	1,070	-0,100	<b>0,120</b>
62	1,040	-0,220	<b>0,223</b>	1,420	-0,110	<b>0,420</b>	1,170	-0,120	<b>0,203</b>	1,180	-0,080	<b>0,191</b>
63	1,110	-0,170	<b>0,200</b>	1,450	-0,070	<b>0,440</b>	1,130	-0,090	<b>0,154</b>	1,100	-0,080	<b>0,125</b>
64	1,000	-0,070	<b>0,070</b>	1,250	-0,010	<b>0,241</b>	1,050	0,000	<b>0,048</b>	1,100	-0,050	<b>0,108</b>
65	1,070	-0,030	<b>0,074</b>	1,140	-0,020	<b>0,136</b>	1,080	-0,070	<b>0,104</b>	1,070	-0,070	<b>0,097</b>
66	1,030	0,020	<b>0,035</b>	1,090	0,030	<b>0,092</b>	1,070	0,010	<b>0,068</b>	1,060	0,010	<b>0,098</b>

67	0,990	-0,050	<b>0,050</b>	1,250	0,030	<b>0,243</b>	1,160	-0,070	<b>0,169</b>	1,180	-0,080	<b>0,191</b>
68	0,940	0,040	<b>0,070</b>	0,940	0,040	<b>0,070</b>	0,970	0,020	<b>0,035</b>	1,010	-0,020	<b>0,022</b>
69	0,980	-0,160	<b>0,161</b>	1,220	-0,100	<b>0,235</b>	1,080	-0,180	<b>0,195</b>	1,120	-0,200	<b>0,231</b>
70	1,110	-0,030	<b>0,110</b>	1,190	-0,010	<b>0,183</b>	1,130	0,000	<b>0,125</b>	1,130	0,010	<b>0,126</b>
71	1,110	-0,070	<b>0,127</b>	1,190	-0,050	<b>0,190</b>	1,140	-0,090	<b>0,162</b>	1,180	-0,090	<b>0,195</b>
72	0,960	-0,070	<b>0,079</b>	1,020	-0,050	<b>0,053</b>	1,020	-0,120	<b>0,121</b>	1,040	0,110	<b>0,116</b>
73	0,930	0,010	<b>0,068</b>	1,020	0,030	<b>0,035</b>	0,970	-0,040	<b>0,049</b>	1,020	-0,050	<b>0,053</b>
74	0,990	-0,050	<b>0,050</b>	1,120	-0,010	<b>0,116</b>	1,120	-0,070	<b>0,135</b>	1,140	-0,070	<b>0,152</b>
75	0,980	0,160	<b>0,161</b>	0,900	0,130	<b>0,162</b>	0,930	0,090	<b>0,112</b>	1,000	0,060	<b>0,060</b>
76	0,990	-0,220	<b>0,220</b>	1,220	-0,180	<b>0,278</b>	1,040	-0,160	<b>0,164</b>	1,030	-0,190	<b>0,192</b>
77	1,020	0,060	<b>0,063</b>	1,080	0,080	<b>0,111</b>	1,060	0,070	<b>0,090</b>	1,030	0,070	<b>0,075</b>
78	1,140	-0,020	<b>0,136</b>	1,250	0,000	<b>0,241</b>	1,190	-0,020	<b>0,184</b>	1,170	0,020	<b>0,165</b>
79	0,990	-0,260	<b>0,260</b>	1,450	-0,130	<b>0,454</b>	1,140	-0,240	<b>0,275</b>	1,160	-0,240	<b>0,285</b>
80	0,910	-0,060	<b>0,105</b>	0,940	-0,050	<b>0,076</b>	0,920	-0,100	<b>0,126</b>	0,960	-0,080	<b>0,088</b>
81	0,930	-0,040	<b>0,078</b>	0,900	-0,050	<b>0,108</b>	0,920	-0,070	<b>0,104</b>	0,990	-0,110	<b>0,110</b>
82	1,010	-0,100	<b>0,100</b>	1,120	-0,070	<b>0,135</b>	1,080	-0,130	<b>0,151</b>	1,110	-0,120	<b>0,160</b>
83	1,070	0,050	<b>0,084</b>	0,990	0,030	<b>0,031</b>	1,030	0,030	<b>0,041</b>	1,030	0,010	<b>0,030</b>
84	0,990	-0,080	<b>0,080</b>	1,130	-0,040	<b>0,131</b>	1,100	-0,060	<b>0,113</b>	1,100	-0,070	<b>0,119</b>
85	1,170	-0,100	<b>0,192</b>	1,190	-0,090	<b>0,204</b>	1,150	-0,040	<b>0,150</b>	1,090	-0,020	<b>0,089</b>
86	1,100	0,070	<b>0,119</b>	1,020	0,050	<b>0,053</b>	1,030	0,070	<b>0,075</b>	1,040	0,070	<b>0,079</b>
87	0,950	-0,120	<b>0,129</b>	0,940	-0,120	<b>0,133</b>	0,960	-0,140	<b>0,145</b>	0,960	-0,130	<b>0,135</b>
88	1,030	-0,070	<b>0,075</b>	1,090	-0,060	<b>0,105</b>	1,040	-0,060	<b>0,071</b>	0,990	-0,030	<b>0,031</b>
89	1,060	-0,060	<b>0,083</b>	1,270	0,010	<b>0,261</b>	1,120	0,000	<b>0,116</b>	1,120	-0,010	<b>0,116</b>
90	0,960	0,020	<b>0,098</b>	1,030	0,040	<b>0,049</b>	1,020	0,000	<b>0,019</b>	0,980	0,130	<b>0,131</b>
91	0,890	0,200	<b>0,108</b>	0,910	0,210	<b>0,227</b>	0,930	0,160	<b>0,173</b>	0,980	0,130	<b>0,131</b>
92	1,030	-0,170	<b>0,172</b>	1,390	-0,090	<b>0,387</b>	1,190	-0,090	<b>0,204</b>	1,110	-0,010	<b>0,106</b>
93	0,980	-0,180	<b>0,181</b>	1,090	-0,160	<b>0,182</b>	1,090	-0,240	<b>0,255</b>	1,100	-0,200	<b>0,222</b>
94	1,080	-0,010	<b>0,078</b>	1,210	0,030	<b>0,205</b>	1,140	-0,040	<b>0,141</b>	1,150	-0,030	<b>0,148</b>
95	1,010	-0,030	<b>0,031</b>	1,080	-0,020	<b>0,079</b>	1,030	-0,020	<b>0,035</b>	1,000	0,020	<b>0,020</b>
96	1,000	0,080	<b>0,080</b>	1,020	0,080	<b>0,082</b>	1,040	0,080	<b>0,088</b>	1,040	0,070	<b>0,079</b>
97	0,910	0,090	<b>0,125</b>	0,940	0,100	<b>0,115</b>	0,930	0,060	<b>0,090</b>	0,990	0,010	<b>0,013</b>
98	1,040	-0,010	<b>0,039</b>	1,130	0,020	<b>0,127</b>	1,080	-0,040	<b>0,087</b>	1,100	-0,040	<b>0,104</b>
99	1,000	-0,050	<b>0,050</b>	1,070	-0,030	<b>0,074</b>	1,050	-0,060	<b>0,077</b>	1,050	-0,060	<b>0,077</b>
100	1,120	-0,100	<b>0,153</b>	1,040	-0,120	<b>0,126</b>	1,040	-0,130	<b>0,135</b>	1,050	-0,130	<b>0,138</b>