



**TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**



**PARALEL TEST OLUŞTURMADA ORTAK SORU SEÇİMİNİ
ETKİLEYEN FAKTÖRLERİN BENZETİM ÇALIŞMASI İLE
DEĞERLENDİRİLMESİ**

Pervin DEMİR

**BİYOİSTATİSTİK ANABİLİM DALI
YÜKSEK LİSANS TEZİ**

**DANIŞMAN
Prof. Dr. Atilla Halil ELHAN**

2014 - ANKARA

**TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

**PARALEL TEST OLUŞTURMADA ORTAK SORU SEÇİMİNİ
ETKİLEYEN FAKTÖRLERİN BENZETİM ÇALIŞMASI İLE
DEĞERLENDİRİLMESİ**

Pervin DEMİR

**BİYOİSTATİSTİK ANABİLİM DALI
YÜKSEK LİSANS TEZİ**

**DANIŞMAN
Prof. Dr. Atilla Halil ELHAN**

2014 - ANKARA

Ankara Üniversitesi Sağlık Bilimleri Enstitüsü

Biyoistatistik Tezli Yüksek Lisans **Programı**

çerçevesinde yürütölmüş olan bu çalışma, aşğıdaki jüri tarafından
Yüksek Lisans **Tezi** olarak kabul edilmiştir.

Tez Savunma Tarihi: 07/11/2014

Prof. Dr. Atilla Halil ELHAN
Ankara Üniversitesi
Tıp Faköltesi
Biyoistatistik Anabilim Dalı
Jüri Başkanı
Tez Danışmanı

Prof. Dr. Yasemin YAVUZ
Ankara Üniversitesi
Tıp Faköltesi
Biyoistatistik Anabilim Dalı

Yrd. Doç. Dr. Selcen YÜKSEL
Yıldırım Beyazıt Üniversitesi
Tıp Faköltesi
Biyoistatistik Anabilim Dalı

İÇİNDEKİLER

Kabul ve Onay	ii
İçindekiler	iii
Önsöz	v
Simgeler ve Kısaltmalar	vi
Şekiller	vii
Çizelgeler	viii
1. GİRİŞ	1
1.1 Araştırmanın Konusu, Amacı ve Önemi	1
1.2 Test Eşitleme	3
1.3 Test Eşitleme Desenleri	6
1.3.1 Tek Grup Deseni (The Single Group Design)	6
1.3.2 Dengelenmiş Desen (The Counterbalanced Design)	6
1.3.3 Eşdeğer/ Denk Gruplar Deseni (The Equivalent Groups Design)	7
1.3.4 Denk Olmayan Gruplarda Ortak Madde Deseni (The Nonequivalent Groups Anchor Design)	7
1.4 Test Eşitleme Yöntemleri	10
1.4.1 Klasik Test Teorisine Dayalı Eşitleme Yöntemleri	10
1.4.1.1 Ortalama Eşitleme Yöntemi (Mean Equating Method)	10
1.4.1.2 Doğrusal Eşitleme Yöntemi (Linear Equating Method)	11
1.4.1.3 Eşit Yüzdelikli Eşitleme Yöntemi (Equipercentile Equating Method)	12
1.4.2 Madde Yanıt Teorisine Dayalı Eşitleme Yöntemleri	12
1.4.2.1 Eş Zamanlı Eşitleme Yöntemi	15
1.4.2.2 Parametre Kestirimlerinin Ayrı Ayrı Yapıldığı Eşitleme Yöntemi	16
1.4.2.2.1 Moment Yöntemleri	17
1.4.2.2.2 Karakteristik Eğri Yöntemleri	18
2. GEREÇ VE YÖNTEM	20
2.1 Benzetim Çalışması	20
2.1.1 Benzetim Koşulları	20

2.1.2	Benzetim Çalışmasında Kullanılan Programlar Hakkında Bilgi	21
2.1.3	Verilerin Türetilmesi ve Test Eşitleme Adımları	22
2.1.4	Benzetim Sonuçlarının Değerlendirilmesi	26
3.	BULGULAR	29
3.1	Elde Edilen Eşitleme Sonuçlarının, Soru Sayısı ve Ortak Madde Yüzdesi Sabit Tutularak, Ortak Maddelerin Ölçekteki Konumlarına Göre Değerlendirilmesi	35
3.2	Elde Edilen Eşitleme Sonuçlarının, Soru Sayısı Sabit Tutularak, Ortak Madde Yüzdesi ve Ortak Maddelerin Ölçekteki Konumlarına Göre Değerlendirilmesi	36
3.3	Elde Edilen Eşitleme Sonuçlarının, Ortak Madde Yüzdesi Sabit Tutularak, Soru Sayısı ve Ortak Maddelerin Ölçekteki Konumlarına Göre Değerlendirilmesi	37
3.4	Elde Edilen Eşitleme Sonuçlarının, Ortak Maddelerin Ölçekteki Konumları Sabit Tutularak, Soru Sayısı ve Ortak Madde Yüzdesine Göre Değerlendirilmesi	38
4.	TARTIŞMA	40
5.	SONUÇ VE ÖNERİLER	44
	ÖZET	45
	SUMMARY	46
	KAYNAKLAR	47
	EK	53
	ÖZGEÇMİŞ	60

ÖNSÖZ

Yüksek lisans eğitimime başladığım ilk günden itibaren ilgi ve desteğini benden esirgemeyen, çalışmalarımı yönlendiren ve çalışma disiplini örnek aldığım danışman hocam Prof. Dr. Atilla Halil ELHAN'a,

Her çıkmaza düştüğümde bana yol gösteren, cesaret veren ve yardımını hiçbir zaman esirgemeyen Doç. Dr. Derya GÖKMEN ÖZTUNA'ya,

Yüksek lisans eğitimimde bilgilerinden yararlandığım Ankara Üniversitesi Biyoistatistik Anabilim Dalı değerli hocalarıma,

Hem akademik hem de manevi anlamda desteklerini her zaman hissettiğim, kendimi geliştirmemde büyük katkıları olan Prof. Dr. S. Yavuz SANİSOĞLU, Yrd. Doç. Dr. Mesut AKYOL ve umutsuzluğa düştüğümde bana ümit verici sözleriyle destek veren Yrd. Doç. Dr. Selcen YÜKSEL'e,

Stresli anlarımda destek olan asistan arkadaşım Afra ALKAN ve Yıldırım Beyazıt Üniversitesi Biyoistatistik Anabilim Dalı çalışma arkadaşlarıma,

Eğitim hayatım boyunca maddi ve manevi desteklerini esirgemeyen aileme ve dostlarıma bana karşı gösterdikleri sabır ve anlayış için teşekkür ederim.

SİMGELER ve KISALTMALAR

β	Madde zorluk düzeyi
EÇO	En çok olabilirlik
HKO	Hata Kareler Ortalaması
KKM	Kısmi Kredi Modeli
KTT	Klasik Test Teorisi
MC	Monte Carlo
MYT	Madde Yanıt Teorisi
Ort	Ortalama
RM	Rasch Modeli
SSM	Sıralı Sonuçlu Model
SS	Standart Sapma
θ	Bireyin incelenen özellik düzeyi

ŞEKİLLER

Şekil 1.1. Ortak birey ve ortak madde eşitleme senaryoları	5
Şekil 1.2. Tek grup deseni	6
Şekil 1.3. Dengelenmiş desen	6
Şekil 1.4. Eşdeğer/denk gruplar deseni	7
Şekil 1.5. İç ortak madde deseni	8
Şekil 1.6. Dış ortak madde deseni	8
Şekil 1.7. Eş zamanlı eşitleme yöntemine ilişkin örnek veri yapısı	15
Şekil 3.1. Seçilen ortak maddelerin ölçekteki konumlarına göre eşitleme sonucunda elde edilen β kestirimleri için HKO değerlerinin dağılımı	35
Şekil 3.2. Seçilen ortak maddelerin yüzde ve ölçekteki konumlarına göre eşitleme sonucunda elde edilen β kestirimleri için HKO değerlerinin dağılımı	37
Şekil 3.3. Soru sayısı ve ortak maddelerin ölçekteki konumlarına göre eşitleme sonucunda elde edilen β kestirimleri için HKO değerlerinin dağılımı	38
Şekil 3.4. Soru sayısı ve seçilen ortak madde yüzdesine göre eşitleme sonucunda elde edilen β kestirimleri için HKO değerlerinin dağılımı	39

ÇİZELGELER

Çizelge 2.1. Benzetim Koşulları	21
Çizelge 2.2. İlk senaryo için birinci benzetim sonucunda elde edilen yanıt deseni	25
Çizelge 2.3. İlk senaryo için birinci benzetim sonucunda elde edilen Set 2’de yer alan maddelerin orijinal, eşitlenmiş ve eşitlenmemiş madde zorluk kestirimleri	26
Çizelge 2.4. Performans ölçütleri	27
Çizelge 3.1. Eşitleme yapıldığında ve eşitleme yapılmadığında elde edilen β kestirimleri için yanlılık değerlerine ait tanımlayıcı istatistikler	30
Çizelge 3.2. Eşitleme sonucunda elde edilen β kestirimleri için yanlılık değerlerine ait tanımlayıcı istatistikler	31
Çizelge 3.3. Eşitleme yapıldığında ve eşitleme yapılmadığında elde edilen β kestirimleri için HKO değerlerine ait tanımlayıcı istatistikler	33
Çizelge 3.4. Eşitleme sonucunda elde edilen β kestirimleri için HKO değerlerine ait tanımlayıcı istatistikler	34

1. GİRİŞ

1.1 Araştırmanın Konusu, Amacı ve Önemi

Bilimsel çalışmalarda ölçme önemlidir. Ölçme, genel anlamıyla herhangi bir niteliği gözlemek ve gözlem sonuçlarını sayılarla veya başka sembollerle ifade etmektir (Turgut, 1977). Ölçek, belirli bir özelliği ölçmek üzere geliştirilmiş, psikometrik özellikleri (geçerlik, güvenirlik, değişime duyarlılık) belirlenmiş ölçme aracı olarak tanımlanabilir. Ölçekler özellikle sağlık bilimleri, eğitim ve sosyal bilimler alanlarında doğrudan ölçüm yapılamayan, gizli (latent) özelliklerin ölçülmesinde kullanılır (Yüksel, 2012).

Ölçeklerde yer alan maddelerin analizinde klasik test teorisi (KTT) ve madde yanıt teorisi (MYT) yaygın olarak kullanılan iki teoridir. Tek parametrelili MYT'e benzer yapı gösteren Rasch modelleri (RM); yaşam kalitesi, tutum veya özürllülük düzeyinin değerlendirildiği ölçeklerin analizlerinde kullanılmaktadır. RM, incelenen veri seti iki sonuçlu maddelerden oluşuyorsa İkili RM veya RM; veri seti çok sonuçlu maddelerden oluşuyorsa, maddelerin kategorilerinin özelliklerine bağlı olarak Kısmi Kredi Modeli (KKM) veya Sıralı Sonuçlu Model (SSM) olarak isimlendirilir. Eşik değerleri arasındaki uzaklıklar her madde için aynı olduğunda KKM, SSM'ye; maddeler iki sonuçlu olduğunda ise RM'ye indirgenmektedir.

RM kullanılarak gerçekleştirilen analizlerin, eğitim, psikoloji, pazarlama, ekonomi, sosyoloji gibi farklı alanlarda olduğu gibi sağlık alanında da ölçek geliştirmek amacıyla kullanımı standart hale gelmiştir.

Sağlık alanında hastalara uygulanan tedavilerin/yöntemlerin etkinliğinin değerlendirilmesinde farklı ölçme araçları/testler/formlar kullanılmaktadır. Hastaların zamana bağlı değişimlerinin ortaya konması veya aynı amaca yönelik

farklı ölçek uygulanan bireylerin karşılaştırılabilmesi için testi oluşturan soruların parametrelerinin (güçlük ve ayırt edicilik düzeylerinin) benzer olduğu testler ya da aynı ölçeğe sahip paralel (eşdeğer) testler oluşturulmalıdır.

Paralel testlerin oluşturulması aşamasında, testin belirli sayıda ortak madde/soru içermesi istenir. Ortak olan maddelerin seçimi, hastaların incelenen özellik düzeylerinin tahmininin standart hatası/güvenirliği üzerinde önemli bir etkinliğe sahiptir. Ortak madde seçiminde kaç maddenin ortak olması gerektiği ve bu maddelerin neye göre belirleneceği bilinmemektedir. Ortak madde setinde yer alan soru sayısı, sorunun incelenen özellik düzeyi eksenini üzerindeki konumu, testte yer alan toplam madde sayısı ve maddelerin kategori sayısı, elde edilecek tahminlerin doğruluğunu etkileyebilecek faktörler arasında yer alır. Paralel test oluşturmada ortak madde kullanımından bahseden çalışmalar olmasına rağmen, ortak madde sayısının ve özelliklerinin detaylı olarak tartışıldığı çalışmalar oldukça azdır (Szabo, 2006). Yapılan çalışmalardan, farklı testlerden elde edilen skorların karşılaştırılmasında bireyler ve maddeler üzerinden test eşitleme yapılmasının gerekliliği anlaşılmaktadır (Yu ve Popp, 2005).

Bu çalışmanın amacı, iki sonuçlu yanıt kategorisine sahip maddelerden oluşan iki farklı soru bankası için, paralel test oluşturmada test eşitleme (anchoring) için kullanılacak en iyi ortak madde setinin belirlenmesinde etkili olan faktörlerin bir benzetim çalışması ile incelenmesidir. Bu amaçla, ortak madde seçimine yönelik farklı senaryolar tanımlanacak ve uygulamada bu sürecin işleyişi araştırılacaktır.

Literatürde test eşitleme aşamalarını tanımlayan ve test eşitleme yöntemlerini karşılaştıran çalışmalar bulunmaktadır (Kolen, 1988; Wang ve Kolen, 1994; Rapp, 1999; Luo ve ark., 2001; Kim ve Lee, 2004; Yu ve Popp, 2005; Livingston, 2004; Kan, 2010; Kilmen, 2010; Gök, 2012). Ancak test eşitlemede kullanılacak ortak madde sayısının ve özelliklerinin belirlenmesini değerlendiren çalışmalar oldukça azdır (Yang ve Houang, 1996; Yang, 2000; Cohen ve Kim, 1998; Suanthong, 1998; Meyers ve ark., 2009; Wei, 2010; Li, 2012). Test eşitleme üzerine yapılan çalışmalar

içerisinde RM'den yararlanılan çalışmalar da yer almaktadır (Caldwell, 1984; Suanthong, 1998; Şahhüseyinoğlu, 2006).

Ülkemizde test eşitlemede KTT ve MYT yöntemlerini inceleyen ve karşılaştıran çalışmalar (Bozdağ ve Kan, 2010; Kan, 2010; Kilmen, 2010; Kan, 2011; Öztürk ve Anıl, 2012; Gök, 2012) olmasına rağmen, RM kullanılarak ortak madde seçiminin değerlendirilmesine ilişkin yapılmış çalışmaya rastlanmamıştır. Çalışmanın orijinalliği, öngörülecek ortak madde seçimini belirleme süreci ile uygulamada test eşitleme ve bireylerin zamana bağlı değişimlerini değerlendirme üzerine yeni bir bakış açısı getirecek olmasıdır.

1.2 Test Eşitleme

Aynı gizli özelliği ölçmek üzere geliştirilmiş farklı testler farklı gruplara uygulanarak bireylerin incelenen özellik düzeyleri karşılaştırılmak istenebilir. Böyle bir durumda gruplardan birine uygulanacak testin diğer gruba uygulanan testten daha kolay ya da zor olmaması gerekmektedir. Bir cevaplayıcıya diğer bir cevaplayıcıdan daha zor bir test formu uygulamayı engellemek ve aynı testin birden fazla formu üzerinde karşılaştırılabilir puanlar elde etmek için test formlarının eşitlenmesine ihtiyaç duyulur (Cook ve Eignor, 1991).

Test eşitleme; benzer içerik ve güçlük düzeyinde geliştirilen iki ya da daha fazla testi tek bir ölçeğe yerleştirmek amacıyla kullanılan ve testlerden elde edilen puanların birbirlerinin yerine kullanılmasını sağlayan istatistiksel bir yöntemdir (Kolen ve Brennan, 2004). Angoff (1984), test eşitlemeyi bir formun birim sistemini diğer formun birim sistemine dönüştürme olarak tanımlamaktadır. Angoff (1984) çalışmasında, Celsius (C) ve Fahrenheit (F) ölçeklerinde sıcaklık ölçümlerinin doğrusal bir dönüşüm ile birbirlerine dönüştürülmesinden bahsetmiştir (Eşitlik 1.1). Bu dönüşüm ile iki farklı ölçek birimleri birbirlerine dönüştürülerek, farklı ölçekler ile ölçülen sıcaklığın birbiri ile kıyaslanabilmesi sağlanmaktadır.

$$C = \frac{(F - 32)}{2} \quad \text{veya} \quad F = (C * 2) + 32 \quad (1.1)$$

Test eşitleme, eşitlenecek testlerin güçlük düzeylerine göre yatay (horizontal) ve dikey (vertical) eşitleme olarak adlandırılmaktadır. **Yatay eşitleme**, incelenen özellik düzeyi benzer olan bireylere uygulanan benzer güçlük düzeyindeki aynı yapıyı ölçen testler ile; **dikey eşitleme** ise, incelenen özellik düzeyi farklı olan bireylere uygulanan farklı güçlük düzeyindeki aynı yapıyı ölçen formlar ile yapılmaktadır. Örneğin; Ankara Üniversitesi ve Yıldırım Beyazıt Üniversitesi Tıp Fakültesi 1. sınıf öğrencilerinin dönem başarılarını karşılaştırmak amacıyla her iki fakülte öğrencilerine farklı testlerin uygulanması ile yatay eşitleme, Ankara Üniversitesi Tıp Fakültesi 2. sınıf öğrencilerinin dönem başarılarını bir önceki sene ile karşılaştırmak amacıyla öğrencilere farklı testlerin uygulanması ile de dikey eşitleme yapılmış olmaktadır.

Test eşitlemenin yapılabilmesi için testlerin belirli koşulları sağlaması gerekmektedir. Bunlar;

Eşitlik: Lord (1980)'a göre; farklı incelenen özellik düzeylerindeki cevaplayıcılar için X ve Y testleri denk ise, cevaplayıcıların X veya Y testini almaları farklılık oluşturmamalıdır. Uygulamada Lord'un eşitlik koşulunun sağlanması zordur.

Simetriklik: X testinden Y testine dönüştürülen puanlar, Y testinden X testine dönüştürülen puanlara eşit olmalıdır (Lord, 1980; Angoff, 1984).

Aynı yapıyı ölçme: Testlerin aynı gizli özelliği ölçmek amacıyla geliştirilmiş olması gerekmektedir. Testler tek boyutlu bir yapıya sahip olmalıdır.

Gruptan bağımsızlık: Testlerin eşitlenmesi ile elde edilen skorlar tüm gruplar için aynı olmalıdır (Lord, 1980; Angoff, 1984; Kolen ve Brennan, 2004).

Test eşitleme; ikinci testten elde edilen puanların birinci testin puanlarına çevrilmesi ile **doğrudan** ya da ikinci testten elde edilen puanların üçüncü bir testin kullanımıyla birinci testin puanlarına çevrilmesi ile **dolaylı** olarak gerçekleştirilebilmektedir.

Yu ve Popp (2005); test eşitlemenin ortak birey ve ortak madde üzerinden gerçekleştirilebileceğini belirtmişlerdir. **Ortak birey eşitleme** (common subject equating) senaryosunda, aynı bireylere aynı yapıyı ölçen, ortak madde içermeyen farklı testler uygulanmaktadır (Şekil 1.1.a). **Ortak madde eşitleme** (common item equating) senaryosunda ise, bireyler aynı yapıyı ölçen ve ortak madde içeren (3. ve 6. maddeler) farklı testler almaktadırlar (Şekil 1.1.b). Ortak madde eşitleme; eş zamanlı ve ayrı ayrı olarak iki farklı şekilde yapılabilir.

Bireyler	Test X Maddeler				Test Y Maddeler			
	1	2	3	4	1	2	3	4
1								
2								
3								

(a) Ortak birey eşitleme

	Bireyler	Maddeler					
		1	2	3	4	5	6
Test X	1						
	2						
	3						
Test Y	1						
	2						
	3						

(b) Ortak madde eşitleme

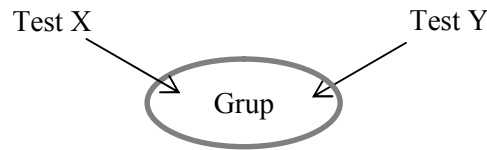
Şekil 1.1. Ortak birey ve ortak madde eşitleme senaryoları

1.3 Test Eşitleme Desenleri

Eşitleme çalışmalarında testlerin ve grupların sayısına göre farklı veri toplama desenleri belirlenmektedir. Veri toplama deseni “eşitleme deseni” olarak adlandırılmaktadır (Kolen ve Brennan, 2004). Literatürde çeşitli eşitleme desenleri tanımlanmıştır (Livingston, 2004).

1.3.1 Tek Grup Deseni (The Single Group Design)

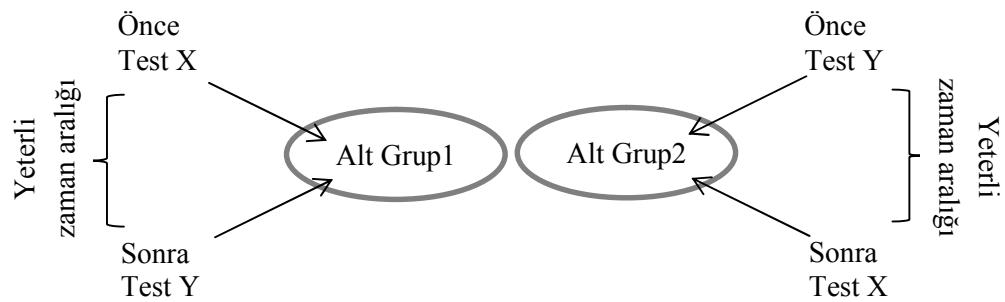
Eşitlenecek X ve Y testleri aynı gruba uygulanmakta ve testler bireylerin incelenen özellik düzeylerinden etkilenmemektedirler (Şekil 1.2). Veriliş sırasından kaynaklı hatalar oluşabileceğinden dolayı aynı gruba X ve Y testlerinin verilmesi nadiren kullanılmaktadır.



Şekil 1.2. Tek grup deseni

1.3.2 Dengelenmiş Desen (The Counterbalanced Design)

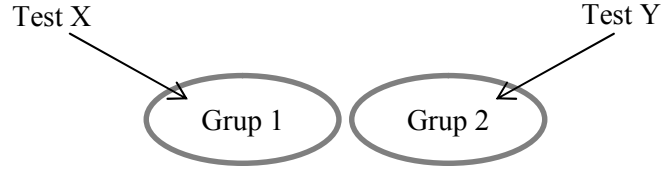
Tek grup deseninde testlerin veriliş sırasından kaynaklanan hatayı kontrol etmek için kullanılan desendir. Bu desende, grup ikiye ayrılır ve testler sıra ile uygulanır (Şekil 1.3).



Şekil 1.3. Dengelenmiş desen

1.3.3 Eşdeğer/ Denk Gruplar Deseni (The Equivalent Groups Design)

Aynı yapıyı ölçen farklı iki test incelenen özellik düzeyi denk olan gruplara uygulanır. Elde edilen eşitleme sonuçlarının doğruluğunu artırmak için test uygulanan birey sayısının fazla olmasının gerekliliği bu desenin sınırlılığıdır. Denk gruplar deseni Şekil 1.4'te gösterilmektedir.



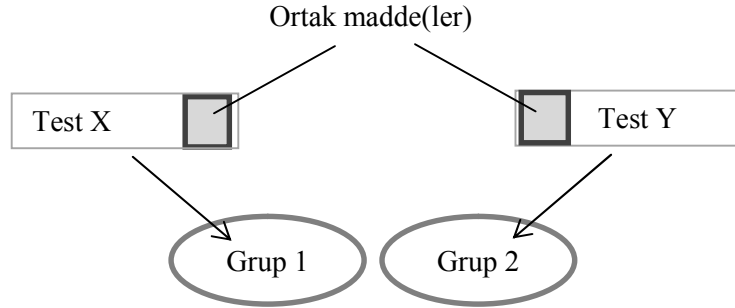
Şekil 1.4. Eşdeğer/denk gruplar deseni

1.3.4 Denk Olmayan Gruplarda Ortak Madde Deseni (The Nonequivalent Groups Anchor Design)

Aynı yapıyı ölçen iki farklı testin her biri farklı gruplara uygulanmaktadır. Denk olmayan gruplara uygulanan testler ortak maddeler içermektedir. Testlerde yer alan ortak maddeler sayesinde puanların aynı ölçek üzerinde yer alması sağlanır. Bu desen; farklı testlerin uygulandığı durumlarda elde edilen puanların karşılaştırılmasının yanında, farklı zamanlarda farklı testlerin uygulanması sonucunda zamana bağlı değişimin ve testlerden elde edilen puanların karşılaştırılmasına da olanak tanımaktadır. Bu avantajları sayesinde en yaygın kullanılan eşitleme desendir. Bu çalışmada da ortak madde deseni kullanılmıştır.

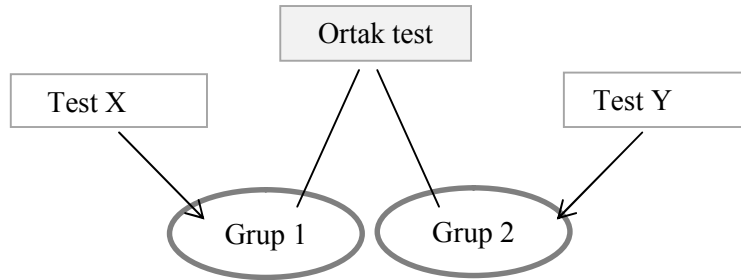
Test eşitleme yapılırken ortak maddeler toplam puana dahil edilebileceği gibi toplam puana dahil edilmeden de kullanılabilir. Denk olmayan gruplarda ortak madde deseni, ortak maddelerin toplam puana dahil edilme durumlarına göre iç ortak madde deseni (The internal-anchor design) ve dış ortak madde deseni (The external-anchor design) olmak üzere ikiye ayrılmaktadır.

İç ortak madde deseninde, gruplara uygulanan testler belirli sayıda ortak madde içermektedirler. Testlerde yer alan ortak maddeler toplam puana dahil edilmektedir (Şekil 1.5).



Şekil 1.5. İç ortak madde deseni

Dış ortak madde deseninde ise, gruplara testler uygulanırken ortak maddeler ayrı bir test gibi ele alınmaktadır. Üçüncü test olarak ele alınan ortak maddeler toplam puana dahil edilmemektedir (Şekil 1.6).



Şekil 1.6. Dış ortak madde deseni

Ortak madde deseni uygulanan çalışmalarda ortak madde seçiminde maddelerin test formları ile özdeş olmasına dikkat edilmelidir. Ortak madde seçimi ve sayısı eşitleme sonucunu etkilemesi muhtemel bir durumdur.

Szabo (2006) tarafından yapılan çalışmada, iki farklı testte yer alan maddelerin aynı ölçek üzerinde yer alması için testler arasında ortak maddeler ile bağlantı kurulması gerektiği belirtilmiştir. Bu çalışmada; soru bankası oluşturmada başarılı bir test eşitleme için gerekli olan ortak madde sayı ve zorluk düzeyi deneysel veri ile araştırılmıştır. Eski ve yeni testler arasında uygun test eşitleme için kaç ortak madde

alınması gerektiğinin ve maddelerin seçimine nasıl karar verileceğinin sorgulanması gerektiği ifade edilmiştir.

Wei (2010) yaptığı çalışmada, ortak maddelerin temsil edebilirliğinin ihmal edilmesinin, eşitleme sonuçlarını ve dengelenmemiş gruplar ortak madde deseninde yıllara göre ölçek skorlarının değişmezliğini nasıl etkilediğini incelemiştir. Çalışmanın sonucunda, bu varsayımın ihlal edilmesinin ölçek skorları üzerinde anlamlı bir etkisi olduğu belirtilmiştir.

Li (2012) tarafından yapılan çalışmada, ortak maddelerin etkisi belirli koşullar altında benzetim çalışması ile incelenmiştir. Madde parametre sapmasının büyüklüğü, ortak madde sayısı, ortak madde setinde yer alan sapan çoktan seçmeli madde sayısı, eski ve yeni test alan grupların incelenen özellik dağılımları benzetim koşulları olarak alınmıştır. Çalışma sonucunda, madde sapma büyüklüklerinin ve ortak madde sayısının eşitleme sonuçlarını büyük ölçüde etkilediği, grupların incelenen özellik dağılımlarının etkisinin ise çok az olduğu belirtilmiştir. Çalışma belirli koşullar altında gerçekleştirildiğinden tüm durumlara genellemek mümkün değildir.

Ortak madde sayısının artması durumunda, eşitleme hatasının azaldığını belirten çalışmaların (Bastari, 2000; Kaskowitz ve De Ayala, 2001; Kim ve Cohen, 2002) yanında ortak madde sayısının etkisi olmadığının belirtildiği çalışmalar da yer almaktadır (Cohen ve Kim, 1998). Ortak maddelerin güçlük düzeylerinin testin güçlüğünü yansıtması durumunda daha az hata ile eşitleme sonuçlarının elde edildiğini belirten çalışmalar bulunmaktadır (Caldwell, 1984; Yang, 2000).

Angoff (1984), ortak madde sayısının tüm testlerde yer alan madde sayısının %20'si kadar olması gerektiğini belirtirken; Kolen ve Brennan (2004), tek boyutluluk varsayımı sağlandığı sürece daha az sayıda ortak madde ile de test eşitleme yapılabileceğini ifade etmektedir.

Literatürde yer alan çalışmalarda elde edilen sonuçlar farklı yöntem, test ve örneklemeler için gerçekleştirilmiştir. Bu çalışmalardan, ortak madde seçiminin ve sayısının test eşitleme sonuçları üzerinde etkili olduğu anlaşılmaktadır. Ancak elde edilecek tahminleri etkilemesi muhtemel olan faktörleri bir arada değerlendirmek gerekmektedir.

Bu çalışmada, test eşitleme işleminde en iyi ortak madde setinin belirlenmesinde etkili olan bazı faktörler bir arada incelenmiştir.

1.4 Test Eşitleme Yöntemleri

Testlerin eşitlenmesinde farklı varsayımlar altında farklı eşitleme yöntemleri kullanılmaktadır. Geleneksel yöntem olan KTT'ne dayalı yöntemlerin yanında MYT'ne dayalı yöntemler de test eşitleme için kullanılan yöntemler arasında yer almaktadır. MYT bazı önemli avantajlara sahip olmasına rağmen, KTT yöntemleri daha basittir (Livingston, 2004).

1.4.1 Klasik Test Teorisine Dayalı Eşitleme Yöntemleri

KTT ile yapılan eşitleme yöntemleri üzerinde çalışılan gruplara (örnekleme) bağımlıdır. Bu nedenle gruplar aynı evrenden olmalıdırlar.

1.4.1.1 Ortalama Eşitleme Yöntemi (Mean Equating Method)

Farklı testlere ait ortalamalar denk olacak şekilde eşitleme yapılmaktadır. Testlerin standart sapmalarının aynı olduğu varsayımı altında kullanılmaktadır. Eşitlenecek testlerin güçlükleri arasındaki farkın, puan ölçeği boyunca sabit olduğu kabul edilir. Uygulanan X testinden alınan bir puanın, Y testindeki karşılığı (Y^*) 1.2 eşitliği yardımı ile hesaplanmaktadır.

$$Y^* = X - m_X + m_Y \quad (1.2)$$

X : X testinden alınan puan

m_X : X testinden elde edilen puanlara ait ortalama

m_Y : Y testinden elde edilen puanlara ait ortalama

Y^* : X testinden alınan bir puanın Y testindeki karşılığı

1.4.1.2 Doğrusal Eşitleme Yöntemi (Linear Equating Method)

Farklı testleri alan grupların, aynı incelenen özellik düzeyinde olması halinde uygulanmaktadır. Aşağıda belirtilen eşitlik 1.3 ile X testinden alınan puanların Y testindeki karşılığı (Y^*) elde edilmektedir.

$$Y^* = \frac{s_Y}{s_X} X + m_Y - \frac{s_Y}{s_X} m_X \quad (1.3)$$

X : X testinden alınan puan

m_X : X testinden elde edilen puanlara ait ortalama

m_Y : Y testinden elde edilen puanlara ait ortalama

s_X : X testinden elde edilen puanlara ait standart sapma

s_Y : Y testinden elde edilen puanlara ait standart sapma

Y^* : X testinden alınan bir puanın Y testindeki karşılığı

Doğrusal eşitlemede, sonuçlar gruplara bağlıdır (Livinston, 2004). Eşitleme yapılacak bir testten elde edilen puanların çok düşük ya da yüksek olması, eşitleme sonrası sınırı aşan puanların elde edilmesine neden olabilmektedir.

1.4.1.3 Eşit Yüzdelikli Eşitleme Yöntemi (Equipercetile Equating Method)

İki testin aynı puan dağılımına sahip olmadığı durumlarda eşit yüzdelikli eşitleme yöntemi kullanılmaktadır. Örneklemin küçük olduğu durumlarda dağılımdan aşırı sapmalar olabileceğinden kullanışsız olduğu da belirtilmektedir.

Eşit yüzdelikli eşitlemede; her iki test için elde edilen puanların yüzdelik sıraları bulunur. Test Y için test X ile aynı yüzdeliğe karşılık gelen değerler (Y^*) belirlenir. Bu durumda test skorları eşitlenmiş olur. Eşit yüzdelikli eşitleme için en basit formül eşitlik 1.4’de verilmektedir.

$$Y^* = X_{\text{altpuan}} + \frac{Y_{\text{yüzdelik}} - X_{\text{altyüzdelik}}}{X_{\text{üstyüzdelik}} - X_{\text{altyüzdelik}}} (X_{\text{üstpuan}} - X_{\text{altpuan}}) \quad (1.4)$$

1.4.2 Madde Yanıt Teorisine Dayalı Eşitleme Yöntemleri

KTT ile yapılan test eşitlemede; farklı testler farklı bireylere uygulandığında ve bireylerin performansları ham puanları ile ifade edildiği durumda, birey incelenen özelliğinin testten elde edilen toplam puanlara bağlı olarak karşılaştırılamaması, sınırlılığa neden olmaktadır. Kolay bir testte yüksek skor alan incelenen özellik düzeyi düşük olan bir birey, daha zor bir testte daha düşük skor alan incelenen özellik düzeyi yüksek olan bir birey ile karşılaştırılmış olabilir (Luo ve ark., 2001).

KTT ile test eşitleme yalnızca bireyler aynı evrenden alındığı durumda tartışılabilir. MYT’de yer alan RM bireylerin aynı evrenden seçilme şartını ortadan kaldırır. Bu sayede, bireylerin performanslarının karşılaştırılması farklı seviyeler için de mümkün olmaktadır. MYT’ne dayalı eşitleme yöntemleri eski teste ait ham veriler olmadığı durumda dahi uygulanabilir (Luo ve ark., 2001).

MYT ile madde ve incelenen özellik parametreleri aynı ölçek üzerine yerleştirilmektedir. KTT’nde madde zorluk parametresi ve madde ayırt edicilik

parametresi testlerin uygulandığı gruplara bağımlı iken MYT’nde parametre tahminleri gruplardan bağımsızdır. Benzer şekilde, MYT’nde bireylerin incelenen özellik düzeyleri testte yer alan maddelerden bağımsızdır.

MYT modellerinin uygulanabilmesi için tek boyutluluk, yerel bağımsızlık ve değişmezlik varsayımlarının sağlanması gerekmektedir.

- Tek boyutluluk (Unidimensionality); testler ile tek bir özelliğin ölçülmesi, incelenen özellik düzeyinin başka bir özellikten etkilenmemesidir.
- Yerel bağımsızlık (Local Independence); bireylerin incelenen özellik düzeyleri sabit tutulduğunda, maddelere verdikleri yanıtların istatistiksel olarak bağımsız olmasıdır.
- Değişmezlik ise maddeler arasındaki farklılığın bireylerden bağımsız olmasıdır.

İkili yanıt kategorisine sahip testler için parametre sayılarına göre RM, iki ve üç parametrelili lojistik modeller kullanılmaktadır.

Rasch Model:

Danimarkalı matematikçi George Rasch tarafından 1960 yılında geliştirilmiştir (Chung, 2010). Bireyin bir maddeye “x” yanıtını verme olasılığı, bireyin incelenen özellik düzeyi (θ) ile madde zorluk düzeyi (β) arasındaki farkın lojistik fonksiyonu olarak tanımlanır. n. bireyin i. maddeye x yanıtını verme olasılığı 1.5’de verilen eşitlik ile hesaplanmaktadır.

$$P(X = x_{ni} | q_n) = \frac{\exp[x_{ni}(q_n - b_i)]}{1 + \exp(q_n - b_i)} \quad , \quad x_{ni}=0,1 \quad (1.5)$$

θ_n : n. bireyin incelenen özellik düzeyi

β_i : i. maddenin zorluk düzeyi

x_{ni} : n. bireyin i. maddeye verdiği yanıt

RM yalnızca madde zorluk parametresini içerdiğinden tek parametrelili model olarak da adlandırılmaktadır. Bu modelde ayırt edicilik parametresi tüm maddeler için eşit ve “1” kabul edilir.

İki ve Üç Parametrelili Lojistik Model:

Birnbaum 1968 yılında iki ve üç parametrelili modeli önermiştir. Madde zorluk parametresine ek olarak madde ayırt edicilik parametresinin yer aldığı iki parametrelili lojistik model eşitlik 1.6’da verilmiştir.

$$P(X = x_{ni} | q_n) = \frac{\exp[a_i x_{ni} (q_n - b_i)]}{1 + \exp[a_i (q_n - b_i)]}, \quad x_{ni}=0,1 \quad (1.6)$$

a_i : i. maddenin ayırt edicilik parametresi

Üç parametrelili modelde, madde zorluk ve ayırt edicilik parametrelerine ek olarak şansa bağlı tahmin parametresi de modelde yer almaktadır. İlgili model eşitlik 1.7’de gösterilmiştir.

$$P(X = x_{ni} | q_n) = c_i + (1 - c_i) \frac{\exp[a_i x_{ni} (q_n - b_i)]}{1 + \exp[a_i (q_n - b_i)]}, \quad x_{ni}=0,1 \quad (1.7)$$

c_i : i. maddenin şansa bağlı tahmin parametresi

Her madde için $a_i = 1$ ve $c_i = 0$ olduğu durumda iki ve üç parametrelili model, RM’e eşit olmaktadır (Chung, 2010).

Modellerde belirtilen madde zorluk parametresi (item difficulty parameter); bir maddeye doğru yanıt verenlerin sayısının, o maddeyi yanıtlayanların sayısına oranıdır. Madde ayırt edicilik parametresi (item discrimination parameter); bir

maddenin, o maddeyle ölçülmek istenen özelliğe sahip olan ile olmayanı ayırt edebilmesidir. Şansa bağlı tahmin parametresi (guessing parameter) ise; bir maddenin doğru yanıtlanması için gereken en düşük madde zorluk değeridir.

Belirtilen MYT modellerinden biri olan RM kullanımı, sadece aynı test uygulanan grupta bireylerin performansını karşılaştırmaz, aynı zamanda farklı gruplarda ya da farklı testler uygulanan bireylerin karşılaştırılmasına da izin vermektedir (Luo ve ark., 2001). Ancak denk olmayan gruplara uygulanan farklı testlerden elde edilen parametre tahminlerini aynı ölçek üzerine yerleştirebilmek için uygun dönüşümlerin yapılması gerekmektedir. MYT’ne dayalı eşitleme yöntemleri parametre kestirimlerine bağlı olarak ikiye ayrılmaktadır. Birincisi, “eş zamanlı eşitleme (Concurrent Calibration)”, ikincisi ise “ayrı ayrı eşitleme (Separate Calibration)” yöntemleridir.

1.4.2.1 Eş Zamanlı Eşitleme Yöntemi

Eş zamanlı eşitleme yönteminde, ortak madde içeren X ve Y testleri için parametre kestirimleri eş zamanlı olarak tek bir analiz ile elde edilir. Eş zamanlı eşitleme için örnek veri yapısı Şekil 1.7’de gösterilmektedir. İlgili örnekte; X testi 1, 2 ve 8. maddelerden, Y testi 4, 5 ve 6. maddelerden oluşmakta, 3 ve 7. maddeler ise her iki testte yer alan ortak maddeleri oluşturmaktadır.

	Bireyler	Maddeler							
		1	2	3	4	5	6	7	8
Test X	1								
	2								
	3								
Test Y	1								
	2								
	3								

Şekil 1.7. Eş zamanlı eşitleme yöntemine ilişkin örnek veri yapısı

1.4.2.2 Parametre Kestirimlerinin Ayır Ayır Yapıldığı Eşitleme Yöntemi

İki farklı test için parametre kestirimleri ayrı ayrı elde edilir. Elde edilen kestirim değerlerinin aynı ölçek üzerinde tanımlanması, ortak maddeler kullanılarak aşağıdaki eşitliklerde gösterilen denklemler ile gerçekleştirilir (Cook ve Eignor, 1991). Eşitlik 1.8, X testindeki n. bireyin θ değerinin Y testindeki karşılığını vermektedir.

$$q_{Y_n}^* = Aq_{X_n} + B \quad (1.8)$$

$$b_{Y_i}^* = Ab_{X_i} + B \quad (1.9)$$

$$a_{Y_i}^* = \frac{1}{A} a_{X_i} \quad (1.10)$$

$$c_{Y_i}^* = c_{X_i} \quad (1.11)$$

q_{X_n} : X testinde n. bireyin incelenen özellik düzeyi

$q_{Y_n}^*$: X testinde n. bireyin incelenen özellik düzeyinin Y testindeki karşılığı

A : Eşitleme denkleminin eğimi

B : Eşitleme denkleminin sabiti

b_{X_i} : X testinde i. maddenin zorluk düzeyi

$b_{Y_i}^*$: X testinde i. maddenin zorluk düzeyinin Y testindeki karşılığı

a_{X_i} : X testinde i. maddenin ayırt edicilik parametresi

$a_{Y_i}^*$: X testinde i. maddenin ayırt edicilik parametresinin Y testindeki karşılığı

c_{X_i} : X testinde i. maddenin şansa bağlı tahmin parametresi

$c_{Y_i}^*$: X testinde i. maddenin şansa bağlı tahmin parametresinin Y testindeki karşılığı

MYT'nde parametre kestiriminin ayrı ayrı yapılması ile elde edilen kestirimlerin aynı ölçekte yer alması için oluşturulan dönüşüm denklemlerinde belirtilen A ve B katsayılarının hesaplanmasında kullanılan dönüşüm yöntemleri ikiye ayrılmaktadır. Bunlar; Moment yöntemleri ve Karakteristik eğri dönüştürme yöntemleridir.

1.4.2.2.1 Moment Yöntemleri

Moment yöntemleri, ortalama-standart sapma ve ortalama-ortalama yöntemlerinden oluşmaktadır.

Ortalama-standart sapma yöntemi

Marco tarafından 1977 yılında tanımlanmıştır (Weeks, 2010). A ve B katsayılarının hesaplanmasında X ve Y testlerinin ortak maddelerine ait zorluk parametresinin ortalama ve standart sapması kullanılmaktadır.

$$A = \frac{s(b_Y)}{s(b_X)} \quad (1.12)$$

$$B = m(b_Y) - Am(b_X) \quad (1.13)$$

$s(b_Y)$: Y testinde ortak maddelere ait zorluk parametrelerinin standart sapması

$s(b_X)$: X testinde ortak maddelere ait zorluk parametrelerinin standart sapması

$m(b_Y)$: Y testinde ortak maddelere ait zorluk parametrelerinin ortalaması

$m(b_X)$: X testinde ortak maddelere ait zorluk parametrelerinin ortalaması

Ortalama-ortalama yöntemi

Loyd ve Hoover tarafından 1980 yılında tanımlanmıştır (Weeks, 2010). Bu yöntemde, A ve B katsayılarının elde edilmesinde X ve Y testlerinde ortak maddelerden elde edilen madde zorluk parametresi ile birlikte madde ayırt edicilik parametresinin ortalamalarından da yararlanılmaktadır.

$$A = \frac{m(a_X)}{m(a_Y)} \quad (1.14)$$

$$B = m(b_Y) - Am(b_X) \quad (1.15)$$

$m(a_X)$: X testinde ortak maddelere ait ayırt edicilik parametrelerinin ortalaması

$m(a_Y)$: Y testinde ortak maddelere ait ayırt edicilik parametrelerinin ortalaması

1.4.2.2.2 Karakteristik Eğri Yöntemleri

Madde karakteristik eğrisi, testte yer alan maddenin doğru yanıtlanma olasılığı ile bireyin incelenen özellik düzeyi arasındaki ilişkiyi göstermede kullanılmaktadır. Bu eğriyi belirlemede çeşitli modellerden yararlanılmaktadır. Karakteristik eğrilerin kullanıldığı test eşitlemede, madde parametre kestirimlerinin aynı anda değerlendirilmesine olanak tanınmaktadır. Test eşitleme için öncelikle A ve B katsayıları elde edilir. Testler için elde edilen karakteristik eğriler arasındaki fark minimize edilir. Literatürde kullanılan iki farklı karakteristik eğri yöntemi bulunmaktadır. Bunlar; Haebara ve Stocking-Lord yaklaşımlarıdır (Kim ve Lee, 2004).

Haebara Yaklaşımı

Dönüştürme için Haebara tarafından 1980 yılında tanımlanan fonksiyon kullanılmaktadır (Weeks, 2010). Bu yöntemde, belirli incelenen özellik düzeyine sahip bireyler için madde karakteristik eğrileri arasındaki fark, her bir maddeye ait madde karakteristik eğrileri arasındaki farkın **karelerinin toplamıdır** (Kolen ve Brennan, 2004). İlgili fonksiyon, eşitlik 1.16'da verilmiştir.

$$Hdiff(q_n) = \sum_{i=1}^k [p_{in}(q_n; a_{X_i}, b_{X_i}, c_{X_i}) - p_{in}(q_n; a_{Y_i}^*, b_{Y_i}^*, c_{Y_i})]^2 \quad (1.16)$$

Madde karakteristik
fonksiyonu

Eşitlenmiş
madde karakteristik
fonksiyonu

k: ortak madde sayısı

$$a_{Y_i}^* = \frac{1}{A} a_{X_i}$$

$$b_{Y_i}^* = Ab_{X_i} + B$$

Stocking-Lord Yaklaşımı

Stocking ve Lord tarafından 1983 yılında tanımlanan yöntemde (Weeks, 2010), belirli incelenen özellik düzeyine sahip bireyler için madde karakteristik eğrileri arasındaki fark, her bir maddeye ait madde karakteristik eğrileri arasındaki farkın **toplamının karesidir** (Kolen ve Brennan, 2004). İlgili fonksiyon eşitlik 1.17’ de belirtilmiştir.

$$Hdiff(q_n) = \left[\sum_{i=1}^k p_{in}(q_n; a_{x_i}, b_{x_i}, c_{x_i}) - p_{in}(q_n; a_{y_i}^*, b_{y_i}^*, c_{y_i}^*) \right]^2 \quad (1.17)$$

Madde karakteristik eğri yöntemlerinin her ikisinde de elde edilen fonksiyonları minimum yapan A ve B katsayılarının hesaplanmasında, iterasyon için başlangıç değeri olarak moment yöntemlerinden elde edilen katsayılar kullanılabilmektedir.

Test eşitleme yöntemlerine ilişkin literatürde yapılmış çalışmalar incelendiğinde; MYT’ne dayalı yöntemlerin KTT’ne dayalı yöntemlerden daha iyi sonuçlar verdiğini ifade eden çalışmalar bulunmaktadır (Kelecioğlu, 1994; Yang, 2000; Chen, 2001; Luo ve ark., 2001). MYT yöntemleri ile ilgili yapılan çalışmaların bazılarında, karakteristik eğri yöntemlerinin karşılaştırılan diğer yöntem/lerden daha az hata içerdiği belirtilmiştir (Stocking ve Lord, 1983; Kim ve Cohen, 1996; . Hanson ve Beguin, 2002; Kim ve Lee, 2004; Kim ve Kolen, 2006; Kilmen, 2010).

Eşitleme çalışmalarında rastgele ya da sistematik hatalar yapılabilmektedir (Kolen, 1988). Örneklem seçiminden kaynaklanan hatalar rastgele hatalardır. Sistematik hatalar ise; seçilen eşitleme yöntem/lerinin varsayımları sağlanmadan kullanılması durumunda ortaya çıkabilen hatalardır. Ortak madde seçiminin kullanıldığı eşitleme çalışmalarında ortak madde seçiminden kaynaklı hatalar da oluşabilmektedir. Gerçekleştirilen eşitleme çalışmalarında ilgili hataların en aza indirilmesine dikkat edilmelidir.

2. GEREÇ VE YÖNTEM

2.1 Benzetim Çalışması

Bu çalışmada, iki sonuçlu yanıt kategorisine sahip maddelerden oluşan iki farklı soru bankası için, ortak madde deseni kullanılarak test eşitleme işleminde en iyi ortak madde setinin belirlenmesinde etkili olan faktörler incelenmiştir. Bu amaçla, ortak madde seçimine yönelik farklı senaryolar tanımlanmıştır.

2.1.1 Benzetim Koşulları

Aynı amaca yönelik hazırlanan testlerin denk olmayan gruplara uygulanması durumunda, testlerden elde edilen puanların karşılaştırılması amacıyla ortak maddelerden yararlanılmaktadır. Literatürde konu ile ilgili yapılan çalışmalar dikkate alınarak;

- ▼ ortak madde setinde yer alan soru sayısı,
- ▼ sorunun incelenen özellik düzeyi eksenini üzerindeki konumu,
- ▼ testte yer alan toplam madde sayısı,
- ▼ maddelerin kategori sayısı
- ▼ örneklem büyüklüğü

elde edilecek tahminlerin doğruluğunu etkileyebilecek faktörler olarak tanımlanmıştır.

Sağlık alanında yapılan çalışmalarda büyük örneklemeler ile çalışmanın her zaman mümkün olmadığı göz önünde bulundurularak, örneklem büyüklüğü her test için 500

birey ile sabitlenmiştir. Testler 20 ve 50 maddeden oluşacak şekilde belirlenmiştir. Ortak madde sayısı literatürde yapılan çalışmalar dikkate alınarak toplam madde sayısının % 10 ve % 20'si olarak tasarlanmıştır (Angoff, 1984; Kolen ve Brennan, 2004). Test eşitlemede kullanılacak ortak maddeler için belirlenen konumlar Çizelge 2.1'de belirtilmiştir. Oluşturulan 16 farklı senaryonun her biri için 1000 Monte Carlo (MC) tekrarı yapılmıştır.

Çizelge 2.1. Benzetim Koşulları

Soru sayısı	Test eşitlemede kullanılacak ortak madde sayısı	Test eşitlemede kullanılacak ortak maddelerin incelenen özellik üzerindeki konumu
20	2 (%10)	En kolay
		En zor
		Orta düzey
		En kolay- en zor
20	4 (%20)	En kolay
		En zor
		Orta düzey
		En kolay- en zor- 1. ve 3. çeyreklik
50	5 (%10)	En kolay
		En zor
		Orta düzey
		En kolay- en zor-1., 2. ve 3. çeyreklik
50	10 (%20)	En kolay
		En zor
		Orta düzey
		En kolay- en zor- 1., 2. ve 3. çeyreklik

2.1.2 Benzetim Çalışmasında Kullanılan Programlar Hakkında Bilgi

Belirlenen benzetim koşullarının değerlendirilmesi amacıyla verilerin türetilmesinde, parametre kestirimlerinin elde edilmesinde ve test eşitlemenin gerçekleştirilmesinde R 3.1.1 programından yararlanılmıştır. R programı, istatistiksel hesaplamalar yapmak ve grafikler çizmek için yazılım ortamı sağlayan; UNIX, Windows ve MacOS ortamlarında çalıştırılabilen açık kaynak kodlu bir programdır (R Development Core Team, 2011). R programında hazır kütüphaneler kullanılabildiği gibi çalışmaya özel kodlamalar da yapılabildiğinden son zamanlarda yaygın olarak kullanılmaktadır. Bu çalışmada, veri türetme ve parametre kestiriminde “eRm” (Mair ve ark., 2014), test

eşitlemenin gerçekleştirilmesinde ise “plink” (Weeks, 2010) kütüphanelerinden yararlanılmıştır. Elde edilen sonuçlara ait istatistiksel hesaplamalar ve grafikler için IBM SPSS Statistics 21.0 (IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.) ve Microsoft-Excel 2007 programları kullanılmıştır.

2.1.3 Verilerin Türetilmesi ve Test Eşitleme Adımları

Benzetim çalışmasında yanıt deseninin oluşturulabilmesi için birey ve madde parametreleri Uniform dağılımından $U[-3,+3]$ üretilmiştir. Böylece birey ve madde parametrelerinin benzer değişim aralığına sahip olması sağlanmıştır. Daha sonra bireylere ait iki sonuçlu yanıt desenlerinin türetilmesinde eşitlik 1.5’de tanımlanan RM’den yararlanılmıştır. Yanıt deseni ve parametre kestirimlerinde R programlamada yer alan “eRm” paketi kullanılmıştır. eRm kütüphanesinde yanıt desenine ait madde zorluk tahminlerinin kestiriminde koşullu en çok olabilirlik (EÇO), birey incelenen özellik düzeylerinin kestiriminde ise klasik EÇO kestirimi kullanılmaktadır. Koşullu EÇO kestirimi ile elde edilen parametre kestirimleri eşitlik 2.1’de de belirtildiği gibi bireylerden bağımsız olarak elde edilmektedir (Erdoğan, 2012; Yüksel, 2012).

$$L = \frac{\exp(-\sum_i b_i s_i)}{\prod_r \sum_{x|r} \exp(-\sum_i x_i b_i)} \quad (2.1)$$

s_i : i. maddeye ait ham skor

x_i : i. maddeye ait yanıt deseni

r : Bireylere ait ham skor

Ayrı ayrı gerçekleştirilen parametre kestirimlerinden yararlanılarak testler arasında eşitleme yapılırken “plink” paketinden yararlanılmıştır. Test eşitlemede MYT’ne

dayalı bir yaklaşım olan Stocking-Lord dönüşümü ile elde edilen A ve B katsayıları kullanılmıştır.

Çalışmalarda, belirli amaca yönelik geliştirilen bir testin uygulandığı grup ile aynı amaca yönelik ama farklı sorulardan oluşan yeni testin uygulanacağı grubun skorlarının karşılaştırılabilmesi için ortak madde/ler kullanılabilir. Bu tez çalışmasında, test skorları için en iyi eşitlemeyi sağlayacak ortak madde seti belirlenmeye çalışılmıştır. Bu amaçla, ilk olarak birinci test 500 bireye uygulanır (Set 1). Bireylerin incelen özellik düzeyleri ve madde zorluk tahminleri belirlenir. Aynı amaca yönelik geliştirilecek olan ikinci test için birinci testten bazı sorular ortak alınır ve ikinci test farklı 500 bireye uygulanır (Set 2). İki testten elde edilen skorların aynı ölçek üzerinde yer alması için Set 1’den elde edilen ortak maddelere ait madde zorlukları eşitlenerek Set 2 için ilgili kestirimler yapılır. Böylece testlere ait madde zorluklarının yanında bireylerin incelenen özellik düzeyleri de düzeltmeler yapılarak aynı ölçek üzerinde yer alır.

Uygulamada ortak madde seçimi yapılırken, en iyi eşitleme için ilk testten hangi maddelerin alınması gerektiğini belirlemeye yönelik gerçekleştirilen benzetim çalışmasının aşamaları aşağıda belirtildiği gibidir. Bu aşamalarda, testlerde soru sayısının 20, ortak madde yüzdesinin 10 ve en kolay zorluk düzeyinde yer alan maddelerin ortak alındığı ilk senaryo üzerinden örnek verilmiştir.

1. Aşama: Her bir senaryoda belirlenen sayıda madde için 1000 bireyin iki sonuçlu yanıt desenleri RM ile türetilmiştir (Orijinal veri seti). Türetilen orijinal setten β ve θ parametreleri kestirilmiştir. Elde edilen β kestirimleri orijinal madde zorluk düzeyleri olarak tanımlanmıştır.

Örnek: İlk senaryo için 1000 bireyin 38 maddeye ait yanıt deseni türetilmiştir. Tüm madde ve bireylere ait parametre kestirimleri kaydedilmiştir.

2. Aşama: Orijinal veri setinden Set1 ve Set2 olmak üzere iki farklı veri seti çekilmiştir. Set1 oluşturulurken; orijinal veri setinde yer alan ilk 500 birey ve formlarda yer alan soru sayısına (20 veya 50) göre belirtilen sayıda madde rastgele alınmıştır. Set1 için belirlenen maddelere ait orijinal β kestirimleri kaydedilmiştir.

Örnek: Set 1 oluşturulurken, 38 maddeden rastgele 20 madde seçilmiştir (1, 4, 6, 8, 10, 11, 13, 16, 17, 18, 19, 21, 23, 24, 29, 31, 32, 33, 34, 35). İlgili maddelere ait madde zorluk düzeyleri kaydedilmiştir.

3. Aşama: Set2 oluşturulurken; Set1'e rastgele alınmış olan maddeler dışındaki maddeler ile ölçekteki konumu dikkate alınarak Set1'den alınan ortak maddeler ve orijinal veri setinde yer alan son 500 bireye ilişkin yanıt deseni alınmıştır. Belirlenen madde ve bireyler için Set 2'ye ait parametre kestirimleri RM ile elde edilmiştir. Set 2 için elde edilen β kestirimleri eşitlenmemiş madde zorluk düzeyleri olarak kaydedilmiştir.

Örnek: Set 2 oluşturulurken, kalan 18 madde (2, 3, 5, 7, 9, 12, 14, 15, 20, 22, 25, 26, 27, 28, 30, 36, 37, 38) ile Set 1'de madde zorluk düzeyi en düşük olan (en kolay) iki madde (29, 32) ve son 500 bireye ait yanıt deseni alınmıştır (Çizelge 2.2). Tanımlanan Set 2'ye ait 20 madde için eşitlenmemiş parametre kestirimleri elde edilmiştir.

4. Aşama: Bu aşamaya kadar ortak madde içeren Set1 ve Set2 için yanıt desenleri, birey ve madde zorluk kestirimleri orijinal veri seti yardımı ile elde edilmiştir. Aynı amaca yönelik olarak geliştirilmiş olduğu varsayılan ve ortak madde içeren iki farklı sette yer alan maddelerin aynı ölçekte yer almasını sağlamak için Stocking-Lord dönüşüm yöntemiyle elde edilen A ve B katsayıları yardımıyla setler arası eşitleme gerçekleştirilmiştir.

Örnek: Set 1 ve Set 2’de yer alan iki ortak madde (29, 32) yardımıyla Stocking-Lord eşitleme yöntemi kullanılarak, Set 2 madde zorluk kestirimleri Set 1’e göre düzenlenmiştir. Böylece iki sette yer alan maddeler aynı ölçekte tanımlanmıştır.

Çizelge 2.2. İlk senaryo için birinci benzetim sonucunda elde edilen yanıt deseni

Bireyler	Maddeler																				
	1	4	6	.	.	.	33	34	35	29	32	2	3	5	.	.	.	36	37	38	
1	1	0	1	.	.	.	0	1	0	0	1	0	1	0	.	.	.	0	1	0	
2	0	0	0	.	.	.	0	0	0	0	0	1	1	0	.	.	.	0	1	0	
3	1	0	1	.	.	.	0	1	0	0	0	1	1	0	.	.	.	1	0	0	
.	.	.	.	Set 1				.	.	.	.	.	.	.	.	.	.	.	.		
.	.	.	.					.	.	.	.	.	.	.	.	.	.	.	.		
.	.	.	.					.	.	.	.	.	.	.	.	.	.	.	.		
498	1	0	0	.	.	.	0	1	0	0	0	1	1	0	.	.	.	0	0	0	
499	1	1	1	.	.	.	0	1	1	0	1	0	1	1	.	.	.	1	1	0	
500	1	0	1	.	.	.	1	1	1	0	0	1	1	1	.	.	.	1	1	0	
501	1	0	1	.	.	.	1	1	0	0	0	0	1	0	.	.	.	0	1	1	
502	1	0	0	.	.	.	0	1	1	1	1	1	1	0	.	.	.	1	1	0	
503	1	1	1	.	.	.	1	1	0	0	0	1	1	0	.	.	.	1	1	0	
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Set 2				.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.					.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.					.	.	.
998	1	1	1	.	.	.	0	1	0	0	0	0	1	0	.	.	.	1	1	0	
999	1	1	1	.	.	.	1	1	0	0	0	1	1	0	.	.	.	1	1	1	
1000	1	1	1	.	.	.	0	1	0	1	1	1	1	1	.	.	.	1	1	1	

5. Aşama: Set1 ve Set2 için gerçekleştirilen eşitleme sonucunda elde edilen eşitlenmiş madde zorluk düzeyleri ile orijinal setten elde edilen madde zorluk düzeyleri karşılaştırılmıştır. Böylece elde edilen analizler sonucunda; test eşitlemek için uygun madde sayısı ve zorluk düzeyi belirlenmeye çalışılmıştır.

Örnek: Set 2’de yer alan 20 madde için eşitleme sonucunda elde edilen düzenlenmiş madde zorluk kestirimleri ile aynı maddeler için orijinal setten elde edilen madde zorluk kestirimleri karşılaştırılmıştır (Çizelge 2.3).

Çizelge 2.3. İlk senaryo için birinci benzetim sonucunda elde edilen Set 2’de yer alan maddelerin orijinal, eşitlenmiş ve eşitlenmemiş madde zorluk kestirimleri

Madde	Orijinal β	Eşitlenmiş β	Fark	Eşitlenmemiş β	Fark
2	1,913	2,014	0,101	2,197	0,284
3	2,805	2,747	-0,058	2,930	0,125
5	-1,032	-1,201	-0,170	-1,018	0,013
7	1,764	1,907	0,143	2,090	0,326
9	-1,046	-0,931	0,115	-0,748	0,298
12	-0,365	-0,306	0,059	-0,123	0,242
14	-0,906	-0,842	0,063	-0,659	0,246
15	-0,898	-0,887	0,012	-0,704	0,195
20	3,177	3,236	0,059	3,419	0,242
22	-0,942	-0,828	0,115	-0,645	0,298
25	-2,472	-2,468	0,004	-2,285	0,187
26	0,265	0,304	0,039	0,487	0,222
27	-1,695	-1,739	-0,044	-1,556	0,139
28	-1,084	-1,080	0,004	-0,897	0,187
29	-2,084	-2,048	0,036	-1,865	0,219
30	-2,794	-2,705	0,089	-2,523	0,272
32	-2,214	-2,250	-0,036	-2,067	0,147
36	1,185	1,321	0,135	1,503	0,318
37	2,288	2,431	0,143	2,613	0,326
38	-0,301	-0,334	-0,033	-1,151	0,150

2.1.4 Benzetim Sonuçlarının Değerlendirilmesi

Set1 ve Set2 için belirtilen koşullarda gerçekleştirilen eşitleme sonucunda elde edilen β kestirimlerinin orijinal setten elde edilen kestirimlerle karşılaştırılması ile eşitlemenin hangi durumda en iyi şekilde gerçekleştiği değerlendirilebilir. Parametre kestirimlerine ait sonuçların karşılaştırılmasında; farklı senaryolar için istatistiksel yöntemlerin performanslarının değerlendirilmesinde kullanılan yanlılık ve hata kareler ortalaması (HKO) ölçütlerinden yararlanılmıştır. Yanlılık bir doğruluk ölçütü iken, HKO hem doğruluk hem de etkinliği değerlendiren bir ölçüttür (Burton ve ark., 2006). Yanlılık; kestirilen parametre değerleri ile gerçek parametre değerleri arasındaki farkın toplamının tekrar sayısına bölünmesi ile elde edilmektedir. HKO; kestirilen parametre değerleri ile gerçek parametre değerleri arasındaki farkın kareleri toplamının tekrar sayısına bölünmesi ile hesaplanmaktadır.

Çalışma kapsamında eşitleme sonuçlarının, senaryolara göre değerlendirilmesi amacıyla yanlılık ve HKO, Çizelge 2.4’de belirtildiği şekilde elde edilmiştir.

Çizelge 2.4. Performans ölçütleri

Değerlendirme ölçütü	Eşitlik
Yanlılık (Bias)	$\frac{\sum_{k=1}^T (\beta_{\text{eşitlenmiş}_k} - \beta_{\text{orijinal}_k})}{T}$
HKO (Mean Square Error)	$\frac{\sum_{k=1}^T (\beta_{\text{eşitlenmiş}_k} - \beta_{\text{orijinal}_k})^2}{T}$

$\beta_{\text{eşitlenmiş}_k}$: k. maddenin eşitleme sonucunda elde edilen madde zorluk kestirimi

$\beta_{\text{orijinal}_k}$: k. maddenin orijinal setten elde edilen madde zorluk kestirimi

T : tekrar sayısı

Belirtilen senaryolarda her madde için elde edilen yanlılık ve HKO değerlerinin ortalaması alınmıştır (Eşitlik 2.2).

$$\text{Ortalama Yanlılık} = \frac{\sum_{i=1}^I \text{Yanlılık}(\beta_i)}{I} \quad \text{ve} \quad \text{Ortalama HKO} = \frac{\sum_{i=1}^I \text{HKO}(\beta_i)}{I} \quad (2.2)$$

I: Toplam madde sayısı

Ortalama yanlılık ve HKO değerlerinin azalması, eşitleme sonuçlarının daha yansız ve doğru yapıldığını ifade etmektedir. Belirlenen tüm koşullar için madde parametrelerine ait ortalama yanlılık ve ortalama HKO sonuçları, eşitleme performanslarının değerlendirilmesi amacıyla hesaplanmıştır.

Eşitleme yapılmadığında elde edilen parametre kestirimlerine ait değerlendirme ölçütleri (yanlılık ve HKO), eşitlemenin yapıldığı durum için Çizelge 2.4’de tanımlanan eşitlikler yardımıyla elde edilmiştir.

Tanımlanan tüm senaryolarda, eşitleme yapıldığında ve eşitleme yapılmadığında elde edilen madde zorluk kestirimleri için yanlılık ve HKO değerlerinin karşılaştırılmasında İki Eş Arasındaki Farkın Önemlilik Testi (Paired Samples T Test) kullanılmıştır. İstatistiksel anlamlılık düzeyi $p<0.05$ olarak kabul edilmiştir.

3. BULGULAR

Benzetim çalışmasında tanımlanan 16 farklı senaryo (2 farklı soru sayısı*2 farklı ortak madde sayısı*Ortak maddelerin ölçekteki 4 farklı konumu) için elde edilen eşitleme sonuçlarının değerlendirilmesi amacıyla hesaplanan ortalama yanlılık ve HKO değerlerinin 0'a yaklaşması istenmektedir. Böylece ortak madde seçimine yönelik en iyi senaryoya ilişkin önerilerde bulunulabilecektir.

Her bir senaryoda, eşitleme yapıldığında ve eşitleme yapılmadığında elde edilen β kestirimlerine ait ortalama yanlılık değerleri Çizelge 3.1'de verilmiştir.

Çizelge 3.1'de yer alan sonuçlar incelendiğinde, eşitleme yapıldığında elde edilen yanlılık değerlerinin eşitlemenin yapılmadığı durumunda elde edilen yanlılık değerlerine göre sıfıra daha yakın olduğu belirlenmiştir. İki durum için hesaplanan yanlılık değerlerine ait farklılıkların; soru sayısının 20, ortak madde yüzdesinin 10 ve ortak maddelerin orta zorluk düzeyinden alındığı senaryo dışındaki tüm senaryolar için istatistiksel olarak anlamlı olduğu tespit edilmiştir ($p<0,05$). Farklılığın istatistiksel olarak anlamlı olmadığı senaryoda, yanlılığa ait sapmanın eşitleme yapılmadığı durumda daha yüksek olduğu görülmüştür.

Eşitleme sonucunda elde edilen β kestirimleri için yanlılık değerlerine ait tanımlayıcı istatistikler Çizelge 3.2'de belirtilmiştir. Tüm senaryolarda değerlerin sıfıra oldukça yakın olduğu belirlenmiştir. Yanlılığın en düşük olduğu senaryoların; soru sayısının 50, ortak madde yüzdesinin 10 ve ortak maddelerin orta zorluk düzeyinden seçildiği senaryo 11 ile soru sayısının 50 ve ortak madde yüzdesinin 20 olduğu senaryo 13, 15 ve 16 olduğu görülmektedir.

Çizelge 3.1. Eşitleme yapıldığında ve eşitleme yapılmadığında elde edilen β kestirimleri için yanlılık değerlerine ait tanımlayıcı istatistikler

Soru sayısı	Ortak madde sayısı	Ortak maddelerin incelenen özellik üzerindeki konumu	Eşitleme yapıldığında Ort $\pm$ SS	Eşitleme yapılmadığında Ort $\pm$ SS	p
20	2 (%10)	En kolay	0,001 $\pm$ 0,005	0,261 $\pm$ 0,009	<0,001
		En zor	-0,005 $\pm$ 0,004	-0,265 $\pm$ 0,009	<0,001
		Orta düzey	0,001 $\pm$ 0,004	-0,007 $\pm$ 0,007	<0,001
		En kolay- en zor	-0,001 $\pm$ 0,006	-0,009 $\pm$ 0,008	<0,001
20	4 (%20)	En kolay	0,003 $\pm$ 0,004	0,456 $\pm$ 0,007	<0,001
		En zor	-0,005 $\pm$ 0,004	-0,457 $\pm$ 0,008	<0,001
		Orta düzey	0,001 $\pm$ 0,003	-0,001 $\pm$ 0,006	0,064
		En kolay- en zor- 1. ve 3. çeyreklik	-0,002 $\pm$ 0,004	0,007 $\pm$ 0,008	<0,001
50	5 (%10)	En kolay	0,004 $\pm$ 0,004	0,265 $\pm$ 0,006	<0,001
		En zor	-0,005 $\pm$ 0,004	-0,267 $\pm$ 0,007	<0,001
		Orta düzey	0,000 $\pm$ 0,004	-0,002 $\pm$ 0,006	<0,001
		En kolay- en zor-1., 2. ve 3. çeyreklik	-0,001 $\pm$ 0,004	-0,002 $\pm$ 0,006	0,006
50	10 (%20)	En kolay	0,000 $\pm$ 0,004	0,474 $\pm$ 0,005	<0,001
		En zor	-0,003 $\pm$ 0,004	-0,475 $\pm$ 0,006	<0,001
		Orta düzey	0,000 $\pm$ 0,004	-0,001 $\pm$ 0,005	0,026
		En kolay- en zor- 1., 2. ve 3. çeyreklik	0,000 $\pm$ 0,004	-0,004 $\pm$ 0,005	<0,001

Ort $\pm$ SS: Ortalama $\pm$ Standart sapma

Çizelge 3.2. Eşitleme sonucunda elde edilen β kestirimleri için yanlılık değerlerine ait tanımlayıcı istatistikler

Soru sayısı	Ortak madde sayısı	Ortak maddelerin incelenen özellik üzerindeki konumu	Senaryo	Ort $\pm$ SS	Ortanca (min; maks)
20	2 (%10)	En kolay	S1	0,001 $\pm$ 0,005	0,000 (-0,011; 0,011)
		En zor	S2	-0,005 $\pm$ 0,004	-0,005 (-0,012; 0,006)
		Orta düzey	S3	0,001 $\pm$ 0,004	0,001 (-0,005; 0,011)
		En kolay- en zor	S4	-0,001 $\pm$ 0,006	-0,002 (-0,015; 0,012)
20	4 (%20)	En kolay	S5	0,003 $\pm$ 0,004	0,002 (-0,005; 0,013)
		En zor	S6	-0,005 $\pm$ 0,004	-0,004 (-0,013; 0,006)
		Orta düzey	S7	0,001 $\pm$ 0,003	0,001 (-0,006; 0,009)
		En kolay- en zor- 1. ve 3. çeyreklik	S8	-0,002 $\pm$ 0,004	-0,003 (-0,011; 0,008)
50	5 (%10)	En kolay	S9	0,004 $\pm$ 0,004	0,004 (-0,007; 0,013)
		En zor	S10	-0,005 $\pm$ 0,004	-0,005 (-0,018; 0,006)
		Orta düzey	S11	0,000 $\pm$ 0,004	0,000 (-0,011; 0,009)
		En kolay- en zor-1., 2. ve 3. çeyreklik	S12	-0,001 $\pm$ 0,004	0,000 (-0,010; 0,010)
50	10 (%20)	En kolay	S13	0,000 $\pm$ 0,004	0,000 (-0,008; 0,009)
		En zor	S14	-0,003 $\pm$ 0,004	-0,003 (-0,014; 0,009)
		Orta düzey	S15	0,000 $\pm$ 0,004	0,000 (-0,008; 0,009)
		En kolay- en zor- 1., 2. ve 3. çeyreklik	S16	0,000 $\pm$ 0,004	0,000 (-0,009; 0,010)

Ort $\pm$ SS: Ortalama $\pm$ Standart sapma
min; maks: : Minimum; maksimum

Eşitleme yapıldığında ve eşitleme yapılmadığında elde edilen β kestirimlerine ait HKO değerleri Çizelge 3.3'te belirtilmiştir. Tüm senaryolarda, eşitleme yapıldığı durumda elde edilen β kestirimlerine ait HKO değerlerinin, eşitlemenin yapılmadığı durumda elde edilen β kestirimlerine ait HKO değerlerinden istatistiksel olarak anlamlı düzeyde daha düşük olduğu tespit edilmiştir ($p<0,001$).

Eşitleme sonucunda elde edilen β kestirimleri için HKO değerlerine ait tanımlayıcı istatistikler Çizelge 3.4'de verilmiştir. Değerlerin tüm senaryolarda sıfıra oldukça yakın olduğu belirlenmiştir.

Çizelge 3.4 incelendiğinde; soru sayısının 20, ortak madde yüzdesinin 10 ve ortak maddelerin incelenen özellik üzerindeki konumunun kolay olduğu senaryo 1'de en yüksek HKO değerinin elde edildiği görülmüştür.

HKO değerleri incelendiğinde; madde sayısının 50, ortak madde yüzdesinin 20 ve ortak madde zorluk düzeyinin orta düzey olduğu senaryo 15'te en düşük değerin elde edildiği tespit edilmiştir. Bir sonraki en düşük HKO değerleri senaryo 7, 11 ve 13'e aittir.

Çizelge 3.3. Eşitleme yapıldığında ve eşitleme yapılmadığında elde edilen β kestirimleri için HKO değerlerine ait tanımlayıcı istatistikler

Soru sayısı	Ortak madde sayısı	Ortak maddelerin incelenen özellik üzerindeki konumu	Eşitleme yapıldığında Ort $\pm$ SS	Eşitleme yapılmadığında Ort $\pm$ SS	p
20	2 (%10)	En kolay	0,019 $\pm$ 0,004	0,140 $\pm$ 0,005	<0,001
		En zor	0,014 $\pm$ 0,001	0,145 $\pm$ 0,006	<0,001
		Orta düzey	0,011 $\pm$ 0,001	0,061 $\pm$ 0,003	<0,001
		En kolay- en zor	0,015 $\pm$ 0,002	0,073 $\pm$ 0,003	<0,001
20	4 (%20)	En kolay	0,011 $\pm$ 0,001	0,273 $\pm$ 0,008	<0,001
		En zor	0,011 $\pm$ 0,001	0,273 $\pm$ 0,008	<0,001
		Orta düzey	0,009 $\pm$ 0,001	0,046 $\pm$ 0,002	<0,001
		En kolay- en zor- 1. ve 3. çeyreklik	0,012 $\pm$ 0,002	0,057 $\pm$ 0,003	<0,001
50	5 (%10)	En kolay	0,011 $\pm$ 0,001	0,107 $\pm$ 0,004	<0,001
		En zor	0,011 $\pm$ 0,001	0,106 $\pm$ 0,004	<0,001
		Orta düzey	0,009 $\pm$ 0,001	0,029 $\pm$ 0,002	<0,001
		En kolay- en zor-1., 2. ve 3. çeyreklik	0,010 $\pm$ 0,001	0,032 $\pm$ 0,002	<0,001
50	10 (%20)	En kolay	0,009 $\pm$ 0,001	0,258 $\pm$ 0,005	<0,001
		En zor	0,010 $\pm$ 0,001	0,258 $\pm$ 0,006	<0,001
		Orta düzey	0,008 $\pm$ 0,001	0,023 $\pm$ 0,001	<0,001
		En kolay- en zor- 1., 2. ve 3. çeyreklik	0,010 $\pm$ 0,001	0,027 $\pm$ 0,001	<0,001

Ort $\pm$ SS: Ortalama $\pm$ Standart sapma

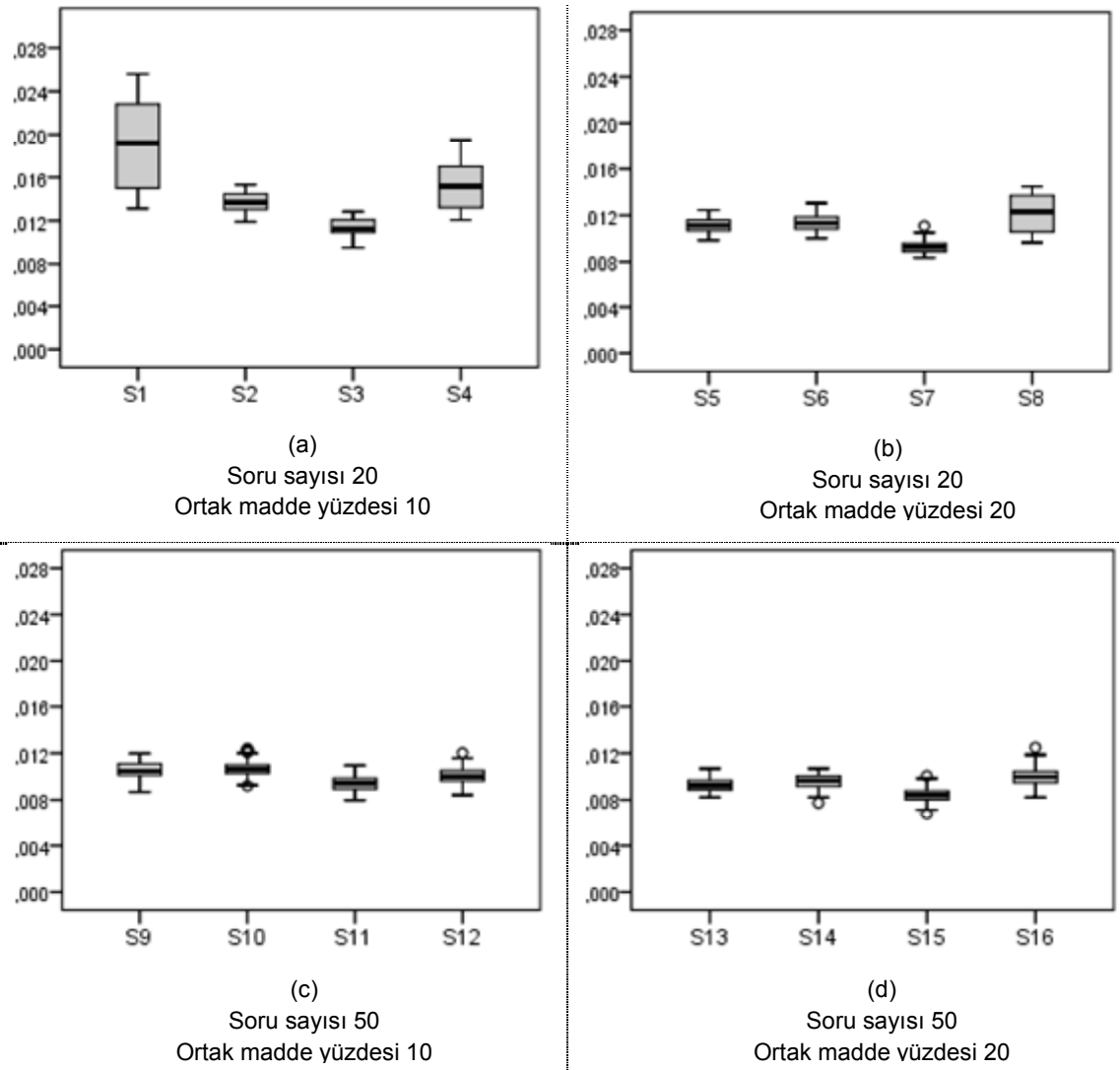
Çizelge 3.4. Eşitleme sonucunda elde edilen β kestirimleri için HKO değerlerine ait tanımlayıcı istatistikler

Soru sayısı	Ortak madde sayısı	Ortak maddelerin incelenen özellik üzerindeki konumu	Senaryo	Ort $\pm$ SS	Ortanca (min; maks)
20	2 (%10)	En kolay	S1	0,019 $\pm$ 0,004	0,019 (0,013; 0,026)
		En zor	S2	0,014 $\pm$ 0,001	0,014 (0,012; 0,015)
		Orta düzey	S3	0,011 $\pm$ 0,001	0,011 (0,009; 0,013)
		En kolay- en zor	S4	0,015 $\pm$ 0,002	0,015 (0,012; 0,019)
20	4 (%20)	En kolay	S5	0,011 $\pm$ 0,001	0,011 (0,010; 0,012)
		En zor	S6	0,011 $\pm$ 0,001	0,011 (0,010; 0,013)
		Orta düzey	S7	0,009 $\pm$ 0,001	0,009 (0,008; 0,011)
		En kolay- en zor- 1. ve 3. çeyreklik	S8	0,012 $\pm$ 0,002	0,012 (0,010; 0,014)
50	5 (%10)	En kolay	S9	0,011 $\pm$ 0,001	0,010 (0,009; 0,012)
		En zor	S10	0,011 $\pm$ 0,001	0,011 (0,009; 0,012)
		Orta düzey	S11	0,009 $\pm$ 0,001	0,009 (0,008; 0,011)
		En kolay- en zor-1., 2. ve 3. çeyreklik	S12	0,010 $\pm$ 0,001	0,010 (0,008; 0,012)
50	10 (%20)	En kolay	S13	0,009 $\pm$ 0,001	0,009 (0,008; 0,011)
		En zor	S14	0,010 $\pm$ 0,001	0,010 (0,008; 0,011)
		Orta düzey	S15	0,008 $\pm$ 0,001	0,008 (0,007; 0,010)
		En kolay- en zor- 1., 2. ve 3. çeyreklik	S16	0,010 $\pm$ 0,001	0,010 (0,008; 0,013)

Ort $\pm$ SS: Ortalama $\pm$ Standart sapma
min; maks: : Minimum; maksimum

3.1 Elde Edilen Eşitleme Sonuçlarının, Soru Sayısı ve Ortak Madde Yüzdesi Sabit Tutularak, Ortak Maddelerin Ölçekteki Konumlarına Göre Değerlendirilmesi

Soru sayısının 20 ve ortak madde sayısının 2 (%10) olduğu durumda; en yüksek HKO değeri ortak maddelerin ölçekteki konumunun en kolay olduğu senaryoda (senaryo 1) ve en düşük HKO değeri ise ortak madde zorluk düzeyinin orta düzeyde alındığı senaryoda (senaryo 3) elde edilmiştir (Şekil 3.1.a).



Şekil 3.1. Seçilen ortak maddelerin ölçekteki konumlarına göre eşitleme sonucunda elde edilen β kestirimleri için HKO değerlerinin dağılımı

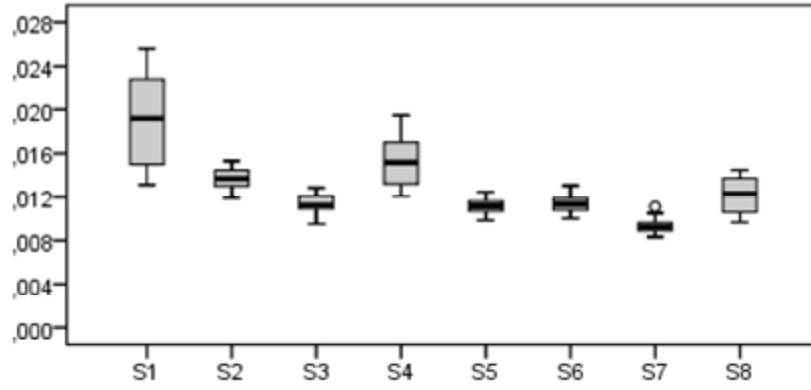
Soru sayısının 20 ve ortak madde sayısının 4 (%20) (Şekil 3.1.b), soru sayısının 50 ve ortak madde sayısının 5 (%10) (Şekil 3.1.c), soru sayısının 50 ve ortak madde sayısının 10 (%20) (Şekil 3.1.d) alındığı durumlarda, en düşük HKO değerlerinin ortak madde seçiminde orta zorluk düzeyinde yer alan maddelerin kullanıldığı senaryolarda görüldüğü tespit edilmiştir (sırasıyla, senaryo 7, 11 ve 15).

3.2 Elde Edilen Eşitleme Sonuçlarının, Soru Sayısı Sabit Tutularak, Ortak Madde Yüzdesi ve Ortak Maddelerin Ölçekteki Konumlarına Göre Değerlendirilmesi

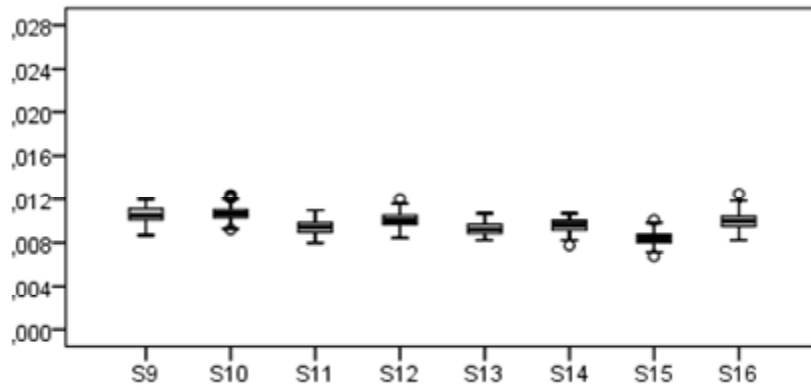
Testlerde yer alan soru sayısı 20 iken, en yüksek HKO değerine sahip olan senaryonun 1. senaryo olduğu belirlenmiştir. Şekil 3.2.a’da görüldüğü gibi soru sayısının 20 olduğu durumda en düşük HKO, ortak madde yüzdesinin 20 ve madde zorluk düzeyinin orta düzey olduğu 7. senaryoda elde edilmiştir. Bir sonraki en düşük HKO değeri ise senaryo 3, 5 ve 6’ da görülmüştür.

Soru sayısının 50 olduğu durumda; en düşük HKO değeri, ortak madde yüzdesinin 20 ve madde zorluk düzeyinin orta düzey olarak alındığı senaryo 15’de görülmektedir (Şekil 3.2.b). En yüksek HKO değeri ise ortak madde yüzdesinin 10 ve ortak madde seçiminin en kolay ile en zor olarak belirlendiği senaryolarda elde edildiği belirlenmiştir (sırasıyla, senaryo 9 ve 10).

Soru sayısı sabit tutularak yapılan değerlendirmeler sonucunda; her iki soru sayısı durumu için HKO değerlerinin en düşük olduğu senaryoların ortak madde yüzdesinin 20 olduğu ve orta zorluk düzeyinde ortak maddelerin seçildiği senaryolarda elde edildiği görülmektedir.



(a) Soru sayısı 20



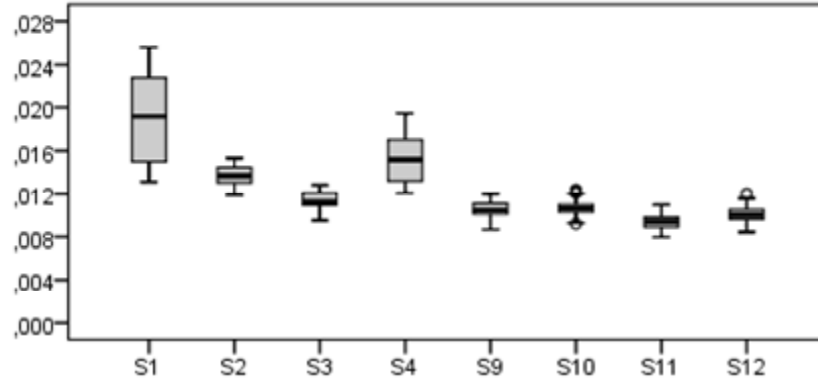
(b) Soru sayısı 50

Şekil 3.2. Seçilen ortak maddelerin yüzde ve ölçekteki konumlarına göre eşitleme sonucunda elde edilen β kestirimleri için HKO değerlerinin dağılımı

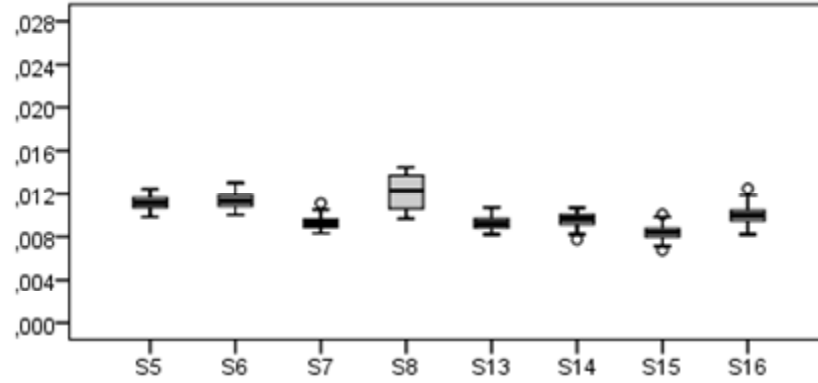
3.3 Elde Edilen Eşitleme Sonuçlarının, Ortak Madde Yüzdesi Sabit Tutularak, Soru Sayısı ve Ortak Maddelerin Ölçekteki Konumlarına Göre Değerlendirilmesi

Ortak madde yüzdesinin 10 olarak alındığı durumlar için soru sayısı ve ortak maddelerin ölçekte yer aldıkları konumlar birlikte değerlendirildiğinde, en düşük HKO değeri soru sayısının 50 ve ortak maddelerin orta zorluk düzeyinde olduğu senaryo 11’de elde edilmiştir (Şekil 3.3.a).

Şekil 3.3'te yer alan grafikler incelendiğinde, her iki durum içinde soru sayısının fazla olduğu ve orta zorluk düzeyinde ortak madde alındığı durumlarda en güvenilir eşitleme sonucunun elde edildiği tespit edilmiştir.



(a) Ortak madde yüzdesi 10



(b) Ortak madde yüzdesi 20

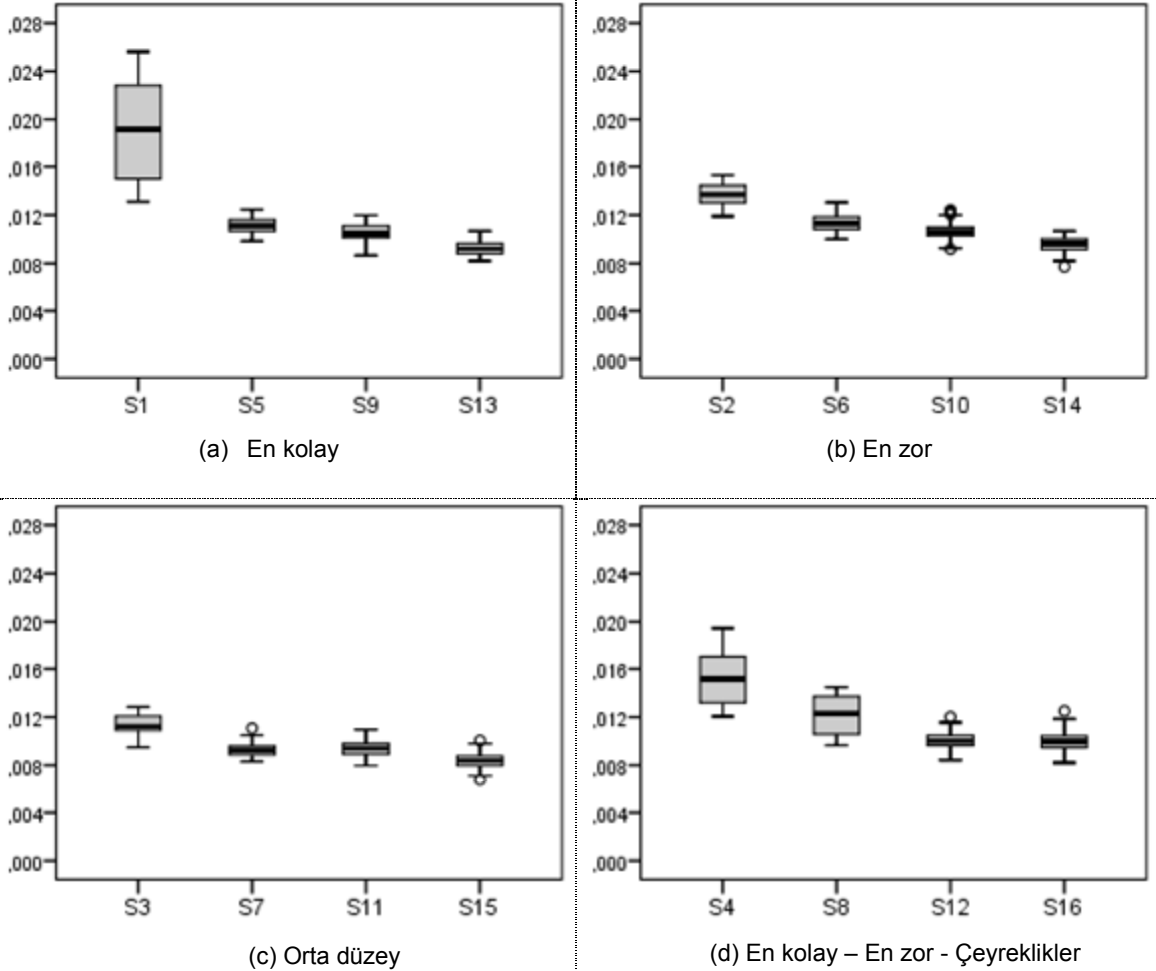
Şekil 3.3. Soru sayısı ve ortak maddelerin ölçekteki konumlarına göre eşitleme sonucunda elde edilen β kestirimleri için HKO değerlerinin dağılımı

3.4 Elde Edilen Eşitleme Sonuçlarının, Ortak Maddelerin Ölçekteki Konumları Sabit Tutularak, Soru Sayısı ve Ortak Madde Yüzdesine Göre Değerlendirilmesi

Seçilen ortak madde zorluk düzeyinin en kolay olduğu durumda, en düşük HKO değerinin soru sayısı ve ortak madde yüzdesinin en yüksek olduğu senaryo 13 için

elde edildiği görülmektedir (Şekil 3.4.a). Ortak maddelerin cetveldeki konumunun en zor olarak belirlendiği senaryolar arasında ise en doğru eşitleme sonucu, soru sayısının 50 ve ortak madde yüzdesinin 20 olduğu senaryo 14’de belirlenmiştir (Şekil 3.4.b).

Orta düzey zorlukta olan maddelerin yer aldığı senaryolar arasında, HKO en düşük olan senaryonun belirlenen en yüksek soru sayısı ve ortak madde sayısı ile senaryo 15 olduğu görülmektedir (Şekil 3.4.c). Ortak maddelerin cetveldeki konumu için belirlenen son durumda ise en düşük HKO değeri senaryo 12 ve 16’da elde edilmiştir (Şekil 3.4.d).



Şekil 3.4. Soru sayısı ve seçilen ortak madde yüzdesine göre eşitleme sonucunda elde edilen β kestirimleri için HKO değerlerinin dağılımı

4. TARTIŞMA

Bireylerin incelenen özellik düzeylerinin değerlendirilmesi amacıyla farklı ölçme araçlarının kullanılması durumunda test eşitlemenin gerekliliği bilinmektedir. Test eşitlemede yararlanılan yöntemlere yönelik çalışmalar değerlendirildiğinde, karakteristik eğri yöntemi olan Stocking-Lord yaklaşımının kullanıldığı çalışmaların daha güvenilir sonuçlar verdiği belirtilmektedir (Kim ve Lee, 2004; Kim ve Kolen, 2006). Bu tez çalışması kapsamında da belirtilen yöntem kullanılarak değerlendirilen senaryolar için elde edilen eşitleme hataları sıfıra oldukça yakın değerlerdir. Ayrıca, eşitleme sonucunda elde edilen kestirim değerlerine ait hatalar, eşitlemenin yapılmadığı durumda elde edilen değerlerden anlamlı düzeyde daha düşüktür.

Literatürde denk olmayan gruplarda ortak madde deseninin yaygın olarak kullanıldığı görülmektedir (Kolen ve Brennan, 2004). Bu çalışmada da denk olmayan gruplarda ortak madde deseninden yararlanılmıştır. Böylece test eşitleme için tanımlanan en iyi ortak madde seti seçimi, farklı gruplara aynı amaca yönelik uygulanan farklı testler için de genellenebilecektir.

Paralel testlerin oluşturulmasında ortak madde seçiminin önemli olduğu vurgulanmaktadır (Yang ve Houang, 1996; Szabo, 2006; Li, 2012; Meyers ve ark, 2012). Cohen ve Kim (1998) tarafından yapılan çalışmada, farklı ortak madde sayıları kullanılarak yapılan analizler sonucunda ortak madde sayısının eşitleme hataları üzerinde etkisi olmadığı saptanmıştır. Ancak sonraki yıllarda yapılan çalışmalarda, ortak madde sayısının artmasının eşitleme sonuçlarının hatasını düşürdüğü belirtilmektedir (Bastari, 2000; Hanson ve Beguin, 2002; Kim ve Cohen, 2002).

Benzer şekilde; bu çalışma kapsamında ortak madde yüzdesinin 20 olduğu senaryolar için elde edilen HKO değerlerinin, ortak madde yüzdesinin 10 olduğu senaryolarda elde edilen HKO değerlerine göre daha düşük olduğu görülmektedir.

Literatürde ortak madde sayısının yanında konumunun ve temsil edebilirliğinin de önemli olduğunu belirten çalışmalar yer almaktadır (Wei, 2010). Li (2012) yaptığı çalışmada ortak madde sayısına ek olarak ortak maddelerin sapmalarının ve yanıt kategorilerinin etkisini de incelemiştir. Çalışma sonunda ortak madde sayı ve sapma büyüklüklerinin sonuçlar üzerinde etkili olduğunu vurgulamıştır.

Bu çalışmada da benzer şekilde soru sayısının 20 ve 50 olarak sabit tutulduğu durumlarda, ortak madde yüzdesinin 20 olduğu ve ortak madde zorluk düzeyinin orta düzey olarak alındığı senaryolarda, eşitlemelerin daha düşük HKO ile yapıldığı görülmüştür. Bu sonuç, ortak madde sayısı yanında ortak madde seçiminde madde zorluk düzeyinin de etkili olduğunu göstermektedir.

Soru sayısı ve ortak madde yüzdesinin sabit tutulduğu durumda en iyi eşitleme sonucunun, ortak maddelerin seçiminde orta zorluk düzeyine sahip maddelerin kullanıldığı senaryolar olduğu belirlenmiştir. Bu sonuç, Kilmen (2010)'in tez çalışmasında belirttiği, Caldwell (1984) tarafından yapılan çalışmanın sonucu ile benzerdir. İlgili çalışmada, KTT ve RM'e dayalı eşitleme sonuçları karşılaştırılırken ortak madde seçimleri orta ve uç güçlüklerde alınmış, orta güçlük düzeyinde yer alan maddeler ile yapılan eşitlemelerin daha iyi sonuçlar verdiği belirtilmiştir.

Ortak madde yüzdesinin sabit olduğu senaryolarda ise, soru sayısının 50 ve ortak maddenin ölçekteki konumunun orta düzey olduğu senaryolarda en güvenilir eşitlemenin yapıldığı tespit edilmiştir.

Suanthong 1998 yılında yapmış olduğu doktora tezinde, simülasyon verilerini kullanarak 81 farklı durumda eşitleme sonuçlarını değerlendirmiştir (3 farklı madde güçlük standart sapması (0,68; 0,95; 0,99), testlerde 3 farklı madde sayısı (50, 100,

200), 3 farklı ortak madde sayısı (10, 20, 30) ve 3 farklı örneklem büyüklüğü (100, 300, 500)). Güçlük sapması en yüksek olan, 100 maddeli ve 20 ortak maddenin alındığı 100 kişilik örneklemde en iyi eşitleme sonucunun elde edildiği belirtilmiştir. İlgili çalışmada bu tez çalışmasında olduğu gibi RM'den yararlanılmıştır. Ancak ortak maddelerin ölçek üzerindeki konumları değerlendirilmemiştir.

Ortak maddeler için tanımlanan farklı konumlar sabit tutulup, soru sayısı ve ortak madde yüzdesinin değerlendirilmesi durumunda, tüm konumlar için soru sayısı ve ortak madde yüzdesinin yüksek olduğu durumun daha güvenilir eşitleme sonucu verdiği belirlenmiştir. Gök (2012) çalışmasında, ortak madde deseni kullanarak MYT'ne dayalı yöntemleri karşılaştırırken farklı koşullarda (testlerdeki madde sayısı, örneklem büyüklüğü, incelenen özellik dağılımı ve veri türetilen model) sonuçları değerlendirmiş ve genel olarak testte yer alan madde sayısının artmasının sonuçlar üzerinde olumlu etkisi olduğunu belirtmiştir.

Testte yer alan soru sayısı ile ortak madde yüzdesinin yüksek olduğu ve ortak maddelerin orta zorluk düzeyinden alındığı durumda en iyi eşitleme sonucunun elde edildiği görülmüştür. Elde edilen bu sonuçlar tez çalışması kapsamında belirlenen koşul ve yöntemler ile sınırlıdır.

Literatürde yapılan çalışmalarda genellikle yöntem karşılaştırmalarına yönelik olarak farklı durumlar farklı modeller altında incelenmiştir. Genel olarak elde edilen çıkarım, ortak madde sayısının artmasının daha iyi eşitleme sonuçlarının elde edilmesinde önemli olduğudur. Bu çalışmada RM ile parametre kestirimleri elde edilerek ortak madde seçimi üzerine farklı senaryolar incelenmiştir. Caldwell (1984) tarafından RM kullanılarak yapılan eşitlemede, ortak madde zorluk düzeyinin orta güçlük düzeyinden olduğu durumda daha iyi sonuçların elde edildiğinin belirtildiği çalışma ile tez çalışmasının sonuçları benzerdir.

Yapılan her çalışmada olduğu gibi bu çalışmada da bazı kısıtlılıklar bulunmaktadır. Çalışmada belirtilen benzetim koşulları ile elde edilen 16 farklı senaryo

incelenmiştir. Bu benzetim koşullarında tanımlanan ortak madde yüzdesi artırılabilir, farklı dağılım, örneklem büyüklüğü ve soru sayısı ile çalışılabilir. Böylece testlerde yer alan soru sayıları ve ortak madde sayıları için farklı durumlar değerlendirilebilir. Çalışmada eşitlenecek testlerin uygulandığı örneklem büyüklüğü sabit alınmıştır. İki yanıtı RM'den yararlanılarak yanıt deseni türetilmiştir. Senaryolar farklı örneklem büyüklükleri için incelenebileceği gibi farklı yanıt desenleri için de incelenebilir.

5. SONUÇ VE ÖNERİLER

Sağlık alanında olduğu gibi birçok alanda (eğitim, psikoloji, v.b) incelenen özellik düzeylerinin ölçülmesi amacıyla bireylere uygulanan testlerden elde edilen puanlar karşılaştırılabilir. Bu karşılaştırmalar, farklı gruplara uygulanan farklı testler için olabileceği gibi bireylerin zamana bağlı değişimlerini değerlendirme amacıyla da gerçekleştirilebilir. Farklı testlerden elde edilen skorların karşılaştırılabilmesi için paralel testler oluşturulması gerekmektedir. Böylece, gruplardan birine uygulanan testin diğer gruba uygulanan testten daha kolay ya da zor olması engellenmektedir.

Paralel testlerin oluşturulmasında ortak madde setlerinin kullanımı oldukça yaygındır. Ortak madde setlerinin belirlenmesinde, testlerde yer alan soru sayısı, soruların yanıt kategorileri, testlerin uygulandığı birey sayısı ve ortak madde sayısı ile birlikte ortak maddelerin zorluk düzeylerinin de dikkate alınması gerekmektedir. Bu tez çalışması kapsamında, paralel test oluşturmak için test eşitlemede kullanılacak en iyi ortak madde setinin tanımlanması amacıyla farklı senaryolar incelenmiştir. Gerçekleştirilen benzetim çalışması sonucunda, testte yer alan soru sayısı ile ortak madde yüzdesinin yüksek olduğu ve ortak maddelerin orta zorluk düzeyinden alındığı durumda en iyi eşitleme sonucunun elde edildiği görülmüştür.

Elde edilen bulgular sonucunda, iki sonuçlu yanıt kategorisine sahip farklı testlerin eşitlenmesinde orta zorluk düzeyinde yer alan maddelerin ortak alınmasının daha güvenilir eşitleme sonuçları verdiği söylenebilir.

ÖZET

Paralel Test Oluşturmada Ortak Soru Seçimini Etkileyen Faktörlerin Benzetim Çalışması ile Değerlendirilmesi

Aynı örtük özelliğin ölçümünde kullanılan farklı test sonuçlarının karşılaştırılmasında paralel testler oluşturulması önemlidir. Paralel testlerin oluşturulması ile elde edilecek tahminlerin doğruluğunu etkileyebilecek çeşitli faktörler bulunmaktadır.

Bu çalışmanın amacı, iki sonuçlu yanıt kategorisine sahip maddelerden oluşan farklı büyüklükteki soru bankaları için, paralel test oluşturmada test eşitleme (anchoring) için kullanılacak en iyi ortak madde setinin belirlenmesinde etkili olan faktörlerin benzetim çalışması ile incelenmesidir. Bu amaca yönelik olarak; Rasch Model ile test eşitleme işleminde ortak madde seçiminin etkisi 16 farklı senaryo (2 farklı soru sayısı*2 farklı ortak madde yüzdesi*Ortak maddelerin incelenen özellik eksenindeki 4 farklı konumu) için incelenmiştir. Çalışmada orijinal set için 1000 bireyin iki sonuçlu yanıt desenleri R programında yer alan “eRm” paketi kullanılarak türetilmiştir. Belirtilen koşullarda oluşturulacak iki farklı setin eşitlenmesinde madde yanıt teorisi eşitleme yöntemlerinden olan Stocking-Lord yönteminden yararlanılmıştır. Tanımlanan her senaryo için orijinal madde zorluk düzeyleri ile eşitleme sonucunda elde edilen madde zorluk düzeyleri, yanlışlık ve hata kareler ortalaması hesaplanarak karşılaştırılmıştır.

Çalışmada ortak madde setinde yer alan maddelerin zorluk düzeyleri sabit tutulduğunda, ortak madde sayısının artması ile daha güvenilir eşitleme sonuçları elde edilmiştir. Ortak madde zorluk düzeyinin orta düzey alınması durumunda daha az hata ile eşitleme yapıldığı tespit edilmiştir. Sonuç olarak, ortak madde setinde yer alan soru sayısının yüksek olduğu ve orta zorluk düzeyinde yer alan maddelerin eşitlemede kullanılması ile daha güvenilir test eşitleme sonuçlarının elde edildiği söylenebilir.

Anahtar sözcükler: İkili Rasch model, madde yanıt teorisi, ortak madde sayısı, paralel test, test eşitleme

SUMMARY

Evaluation the factors affecting common item selection to develop parallel test by using simulation study

Parallel tests are important to compare the results of different tests used to measure the same latent feature. There are many factors which affect the accuracy of estimates obtained by forming parallel test.

The aim of this study was to investigate the factors that affect the selection of the best common item set used to construct a parallel test for dichotomous item banks of different sizes, by a simulation study. For this purpose, the effect of the common item selection in test equating with Rasch Model was examined for 16 different scenarios (2 different number of items*2 different percentage of common items*4 different item difficulty of common item set). In this study, dichotomous response patterns of 1000 persons for original set was derived using “eRm” package in R program. Stocking-Lord method was used to equate two different sets which were formed under the specified conditions. For each scenario, the item difficulty estimates obtained from the original data set and test equating were compared by calculating the bias and mean square of error.

When the level of item difficulty in common item set was constant, more reliable equating results were obtained while the number of common items increased. In the case that the common item difficulty was moderate, the test equating results had less error. As a result, one can get more reliable test equating results when the number of items in the common item set is increasing and the common item difficulty is moderate.

Keywords: Dichotomous Rasch model, item response theory, number of anchored items, parallel tests, test equating

KAYNAKLAR

- AINOL MADZIAH, A., NOOR LIDE, A. K. (2006). Classical and Rasch Analyses of Dichotomously Scored Reading Comprehension Test Items. *Malaysian Journal of ELT Research*, 2(1): 1-20.
- ALBANO, A. (2014). R Package “Equate”: Observed-Score Linking and Equating. February 22, 2014. <http://www.r-project.org>.
- ANGOFF, W. H. (1984). Scales, Norms and Equivalent Scores. *Educational Testing Service*. Princeton, New Jersey.
- BASTARI, B. (2000). Linking multiple-choice and constructed-response items to a common proficiency scale. University of Massachusetts, Amherst, January 1, 2000 (<http://scholarworks.umass.edu/dissertations/AAI9960735>).
- BATTAUZ, M. (2014). R Package “equateIRT”: Direct, chain and average equating coefficients with standart errors using IRT methods. July 2, 2014.
- BAYKUL, Yaşar. (2000). EĞİTİMDE VE PSİKOLOJİDE ÖLÇME: Klasik Test Teorisi ve Uygulaması. Hacettepe Üniversitesi Eğitim Fakültesi, ÖSYM Yayınları, Ankara 2000.
- BEJAR, I. I., WINGERSKY, M. S. (1981). An Application of Item Response Theory to Equating the Test of Standard Written English. *ETS RR No: 81-35*, New York.
- BOZDAĞ, S., KAN, A. (2010). Şans Başarısının Test Eşitlemeye Etkisi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi (H. U. Journal of Education)*, 39: 91-108.
- BRAMLEY, T. (2006). Equating methods used in KS3 Science and English. Paper for the NAA technical seminar, Oxford, 23-24 March 2006.
- BURTON, A., ALTMAN, D. G., ROYSTON, P., HOLDER, R. L. (2006). The design of simulation studies in medical statistics. *Statistics in Medicine*, Published online 31 August 2006 in Wiley InterScience, 25: 4279-4292.
- CALDWELL, L. J. (1984). A Comparison of Equating Error in lineer and Rasch Model Test equating Method.
- CHEN, H. W. (2001). Calibration of the ITBS Test Battery to the Complete Test Battery: A comparison Five Linking Methods. University of Iowa, ISSN: 049189270.
- CHUNG, W. W. (2010). Recent Developments in Rasch Measurement. The Hong Kong Institute of Education. This paper was firstly presented in a lecture of the Chair Professors Public Lecture series of The Hong Kong Institute of Education on 28 January 2010.

- COHEN, A. S., KIM, S. H. (1998). An Investigation of Linking Methods Under the Graded Response Model. *Applied Psychological Measurement*, 22(2): 116-130.
- COOK, L. L., EIGNOR, D. R. (1991). IRT Equating Methods. ITEMS: Instructional Topics in Educational Measurement. *Educational Testing Service*. Princeton, New Jersey.
- DEMİR, S., GÜLER, N. (2014). Ortak maddeli denk olmayan gruplar desenine ilişkin test eşitleme çalışması. *International Journal of Human Sciences*, 11(2): 190-208.
doi: 10.14687/ijhs.v11i2.2870.
- ERDOĞAN, B. (2012). Çoklu Atama Yöntemlerinin Rasch Modelleri İçin Performansının Benzetim Çalışması İle İncelenmesi. Ankara Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı, Doktora Tezi.
- FISHER, G.H., MOLENAAR, I. (1995). Rasch Models. Foundations, Recent Developments and Applications. Springer, New York.
- GAO, H. (2004). The Effect of Different Anchor Tests on the Accuracy of Test Equating for Test Adaptation. A PhD Thesis at The Faculty of the College of Education of Ohio University.
- GÖK, B. (2012). Denk Olmayan Gruplarda Ortak Madde Deseni Kullanılarak Madde Tepki Kuramına Dayalı Eşitleme Yöntemlerinin Karşılaştırılması. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Eğitimde Ölçme ve Değerlendirme Bilim Dalı, Doktora Tezi, Ankara, 2012.
- HANSON, B. A., BEGUIN, 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.
doi: 10.1177/0146621602026001001.
- HEINE, J. H. (2014). R Package “pairwise”: Rasch Model Parameters by Pairwise Algorithm. April 14, 2014.
- How much item drift is too much? (2014). <http://www.rasch.org/rmt/rmt273a.htm>. Erişim tarihi: 29.04.2014, 16:00.
- IBM Corp. Released 2012. (2012). IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.
- KAN, A. (2010). Test Eşitleme: Aynı Davranışları Ölçen, Farklı Madde Formlarına Sahip Testlerin İstatistiksel Eşitliğinin Sınanması. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 1(1): 16-21.
- KAN, A. (2011). Test Eşitleme: OKS Testlerinin İstatistiksel Eşitliğinin Sınanması. *Eğitim ve Bilim (Education and Science)*, 36(160).
- KASKOWITZ, G. S., DE AYALA, R. J. (2001). The Effect of Error in Item Parameter Estimates on the Test Response Function Method of Linking. *Applied Psychological Measurement*, 25(1): 39-52.

- KELECİOĞLU, H. (1994). Öğrenci Seçme Sınavı Puanlarının Eşitlenmesi üzerine Bir Çalışma. Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Doktora Tezi, Ağustos 1994, Ankara.
- KİLMEN, S. (2010). Madde Tepki Kuramı'na Dayalı Test Eşitleme Yöntemlerinden Kestirilen Eşitleme Hatalarının Örneklem Büyüklüğü ve Yetenek Dağılımına Göre Karşılaştırılması. Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Eğitimde Ölçme ve Değerlendirme Bilim Dalı, Doktora Tezi, Temmuz 2010.
- KIM, D., BARTON, K., CHOI, S. (2010). Sample Size Impact on Screening Methods in the Rasch Model. *Presented at the American Educational Research Association*, Denver, CO.
- KIM, S. H., COHEN, A. S. (1996). Comparison of Linking and Concurrent Calibration Under Item Response Theory. *Paper presented at the annual meeting of the American Educational Research at New York*, April 1996.
- KIM, S. H., COHEN, A. S. (2002). A comparison of Linking and Concurrent Calibration Under the Graded Response Model. *Applied Psychological Measurement*, 26(1): 25-41. doi: 10.1177/0146621602026001002.
- KIM, S., KOLEN, M. J. (2006). Robustness to Format Effects of IRT Mixed-Format Tests. *Applied Measurement in Education*, 19(4): 357-381. doi: 10.1207/s15324818ame1904_7.
- KIM, S., LEE, W.C. (2004). IRT Scale Linking Methods for Mixed-Format Tests. *ACT Research Report Series*. December 2004.
- KOLEN, M. J. (1988). An NCME Instructional Module on Traditional Equating Methodology. *ITEMS: Instructional Topics in Educational Measurement*, Winter 1988. p: 29-36.
- KOLEN, M.J., & BRENNAN, R.L. (1987). Linear Equating Models for the Common Item Nonequivalent Populations Design. *Applied Psychological Measurement*, 11(1): 263-277.
- KOLEN, M. J., BRENNAN, R. L. (2004). Test equating, scaling, and linking: Methods and practices (2 nd Ed). New York: Springer-Vergal.
- LAU, C. A., ARCE, A. J. (2011). Comparing Methods for Detecting Unstable Anchor Items with Net DIF and Global DIF Conceptions. *Annual Meeting of the American Educational Research Association*, New Orleans, Louisiana, April 2011.
- LI, Y. (2012). Examining the Impact of Drifted Polytomous Anchor Items on Test Characteristic Curve (TCC) Linking and IRT True Score Equating. Research Report, *ETS RR- 12-09*, Princeton, New Jersey, May 2012.
- LIAO, C. (2010). TOEIC Listening and Reading Test Scale Anchoring Study.ETS, January 2010.
- Linking Tests with the Rasch Model. <http://www.rasch.org/rmt/rmt111k.htm>. Erişim tarihi: 29.04.2014, 15:42.

- LIU, J., SINHARAY, S., HOLLAND, P.W., FIEGENBAUM, M., CURLEY, E. (2009). The Effects of Different Types of Anchor Tests on Observed Score Equating. *ETS*, Princeton, New Jersey, December 2009. <http://www.ets.org/research/contact.html>.
- LIVINGSTON, S. A. (2004). Equating Test Scores (Without IRT). *Educational Testing Service*, New Jersey.
- LORD, F. M. (1980). Applications of item response theory to practical testing problems. New Jersey, Lawrence Erlbaum.
- LUO, G., SEOW, A. & CHIN C. L. (2001). Linking and Anchoring Techniques in Test Equating Using The Rasch Model. Conference paper, Danson, Myles (Ed), 5th CAA International Computer Assisted Assessment Conference, 2-3 July 2001, Loughborough, UK. <https://dspace.lboro.ac.uk/2134/1817>.
- MAIR, P., HATZINGER, R., MAIER, M.J. (2014). R Package “eRm”: Extended Rasch Modeling. February 5, 2014. <http://r-forge.r-project.org/projects/erm/>.
- MASTERS, G. N. (1985). Common - Person Equating with the Rasch Model. *Applied Psychological Measurement*, 9(1): 73-82.
- MEYERS, J. L., MILLER, E.G., 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, doi: 10.1080/08957340802558342.
- MEYERS, J.L., MURPHY, S., GOODMAN, J., TURHAN, A. (2012). The Impact of Item Position Change on Item Parameters and Common Equating Results under the 3PL Model. Annual Meeting of the National Council on Measurement in Education, Vancouver, B.C., April 2012.
- MITTELHAEUSER, M.A., BEGUIN, A.A., SIJTSMA, K. (2011). Comparing the Effectiveness of Different Linking Designs: the Internal Anchor versus the External Anchor and Pre-test Data. Measurement and Research Department Reports, 2011-1. Cito Arnhem.
- MOSES, T., DENG, W., ZHANG, Y. (2010). The Use of Two Anchors in Nonequivalent Groups With anchor Test (NEAT) Equating. *ETS RR-10-23*, Princeton, New Jersey.
- MURPHY, S., LITTLE, I., FAN, M., LIN, C-H., KIRKPATRICK. (2010). The Impact of Different Anchor Stability Methods on Equating Results and Student Performance. *Paper presented at the annual meeting of the National Council of Measurement in Education*, Denver, CO, May 2010.
- ÖZTUNA, D. (2008). Kas-İskelet sistemi Sorunlarının Özürüllük Değerlendiriminde Bilgisayar Uyarlamalı Test Yönteminin Uygulanması. Ankara Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı, Doktora Tezi.
- ÖZTÜRK, N., ANIL, D. (2012). Akademik Personel ve Lisansüstü Eğitimi Giriş Sınavı Puanlarının eşitlenmesi Üzerine Bir Çalışma. *Eğitim ve Bilim (Education and Science)*, 37(165): 180-193.

- PANG, X., MADERA, E., RADWAN, N., ZHANG, S. (2010). A Comparison of Four Test Equating Methods. *EQAO Research*.
- PHILLIPS, S.E. (1983). Comparison of Equipercetile and Item Response Theory Equating When the Scaling Test Method Is Applied to a Multilevel Achievement Battery. *Applied Psychological Measurement*, 7(3): 267-281, Summer 1983.
- RAPP, J. (1999). Linear and Equipercetile Methods for Equating PET. National Institute for Testing and Evaluation, December 1999.
- RIZOPOULOS, D. (2006). ltm: An R Package for Latent Variable Modeling and Item Response Theory Analyses. *Journal of Statistical Software*, November 2006, 17(5).
- RIZOPOULOS, D. (2013). R Package “ltm”: Latent Trait Models Under IRT. December 20, 2013. <http://rwiki.sciviews.org/doku.php?id=packages:cran:ltm>.
- SHIN, S. (2009). How to Treat Omitted Responses in Rasch Model-Based Equating. *Practical Assessment, Research & Evaluation*, 14(1), ISSN: 1531-7714.
- SINHARAY, S., HOLLAND, P. (2006). The Correlation Between the Scores of a Test and an Anchor Test. *ETS: RR-06-04*.
- STOCKING, M. L., LORD, F. M. (1983). Developing a Common Metric in Item Response Theory. *Applied Psychological Measurement*, 7(2): 201-210, April 1983, doi: 10.1177/014662168300700208.
- SUANTHONG, S. (1998). An Investigation of Factors Affecting Test Equating in Latent trait Theory. PhD Thesis. University of North Texas, Denton, August 1998. “digital.library.unt.edu/ark:/67531/metadc278079/”, erişim tarihi: 28.09.2014, 12:42.
- SUNNASSEE, D. (2011). Conditions Affecting the Accuracy of Classical Equating Methods for Small Samples Under the NEAT Design: A Simulation Study .
- SZABÓ, G. (2006). Anchors aweigh! An Analysis of the Impact of Anchor Item’s Number and Difficulty Range on Item Difficulty Calibrations. In M. Nikolov & J. Horváth (Eds.), *UPRT, 2006: Empirical Studies In English Applied Linguistics*, p: 249-262.
- ŞAHHÜSEYİNOĞLU, D. (2006). İngilizce Yeterlik Sınavı Puanlarının Üç Farklı Eşitleme Yöntemine Göre Karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 31(1): 115-125.
- TOPCZEWSKI, A., CUI, Z., WOODRUFF, D., CHEN, H., FANG, Y. (2013). A Comparison of Four Linear Equating Methods for the Common-Item Nonequivalent Groups Design Using Simulation Methods. *ACT Research Report Series*.
- TURGUT, F. (1971). Şans Başarısının Test Puanlarına Etkisi. ODTÜ Yayınları, Ankara.

- TURGUT, F. (1977). Eğitimde Ölçme ve Değerlendirme Metodları. Ankara: Nüve Matbaası.
- TURGUT, F. (1979). "Ölçekleme, Norm Hazırlama, Normları Eşitleme ve Ayarlama". Test Geliştirme Teknikleri Basılmamış Ders Notları, Ankara.
- VALE, C. D. (1986). Linking Item Parameters onto a Common Scale. *Applied Psychological Measurement*, 10(4): 333-344, December 1986.
- WANG, T., KOLEN, M.J. (1994). A Quadratic Curve Equating Method to Equate the First Three Moments in Equipercentile Equating. *ACT Research Report Series*, 94-2, June 1994.
- WEEKS, J.P. (2010). plink: an r Package for Linking Mixed-format Tests Using IRT-Based Methods. <http://CRAN.R-project.org/package=plink>.
- WEI, H. (2010). Impact of Non-representative Anchor Items on Scale Stability. *Paper presented at the annual meeting of the National Council on Measurement in Education*, Denver.
- WOLDBECK, T. (1998). Basic Concepts in modern Methods of Test Equating. Paper presented at the annual meeting of the Southwest Psychological Association, New Orleans.
- WRIGHT, B. D., MASTERS, G. N. (1982). Rating Scale Analysis Rasch Measurement. Chicago, IL: MESA PRESS.
- YANG, W. L. (2000). The Effects of Content Homogeneity and Equating Method On the Accuracy of Common-Item Test Equating. *ETS, Paper prepared for the presentation at the Annual Meeting of the American Educational Research Association*, New Orleans.
- YANG, W.L., HOUANG, R.T. (1996). The Effect of Anchor Length and Equating Method on Accuracy of Test Equating: Comparisons of Linear and IRT-Based Equating Using Anchor Item Design. *Paper Presented at the AERA Annual Conference, Division D*. New York.
- YEN, W. (2002). Scaling and Equating. New York State Technical Conference, *ETS*.
- YI, Q., HANSON, B., WIDIATMO, H., HARRIS, D. (2000). Procedures of screening for "affected" common items in common - item nonequivalent groups design. *Paper presented at the annual meeting of National Council on Measurement in Education*, New Orleans.
- YU, C. H., OSBORN POPP, S. E. (2005). Test Equating by Common Items and Common Subjects: Concepts and Applications. *Practical Assessment Research & Evaluation*, 10(4), ISSN 1531-7714, May 2005.
- YÜKSEL, S. (2012). Ölçeklerde Saptanan Madde İşlev Farklılığının Karma Rasch Modelleri İle İncelenmesi. Ankara Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı, Doktora Tezi.
- ZHU, W. (1998). Test Equating: What, why, how? *Research. Quarterly for the Exercise and Sport*. 69(1): 11-23.

EK

Benzetim çalışmasında tanımlanan ilk senaryonun gerçekleştirilmesine yönelik kullanılan fonksiyon

Çalışmada yararlanılan R paketlerinin yüklenmesi

```
install.packages ("eRm")
```

```
install.packages ("plink")
```

```
library (eRm)
```

```
library (plink)
```

#Orijinal set için madde ve birey sayısının tanımlanması

```
np=1000 # birey sayısı
```

```
ni=38 # madde sayısı
```

```
cur=1
```

```
repeat{
```

#1000 birey ve 38 madde için yanıt deseninin türetilmesi ve parametre kestirimlerinin elde edilmesi

```
ppar=runif(np,-3,3) #bireyin incelen özellik düzeyi
```

```
ipar=runif(ni,-3,3) #madde zorluk düzeyi
```

```
write.table(ppar,file=paste(toString(cur),"_ppar.txt"),sep="",col.names=FALSE,  
row.names=FALSE)
```

```
write.table(ipar,file=paste(toString(cur),"_ipar.txt"),sep="",col.names=FALSE,  
row.names=FALSE)
```

```
Full_data=sim.rasch(ppar,ipar)
```

```
write.table(Full_data,file=paste(toString(cur),"_Full_data.txt"),sep="",  
col.names=FALSE, row.names=FALSE)
```

```
FD=RM(Full_data)
```

```
FD_Beta_rm=cbind(FD$betapar,FD$se.beta)
write.table(FD_Beta_rm[,1],file=paste(toString(cur),"_Full_data_Beta.txt"),
col.names=FALSE, row.names=FALSE)
write.table(FD_Beta_rm[,2],file=paste(toString(cur),"_Full_data_Beta_se.txt"),
col.names=FALSE, row.names=FALSE)
```

```
p.s<-person.parameter(FD)
extreme=matrix(p.s$pers.ex,nrow=length(p.s$pers.ex),ncol=1)
extreme=cbind(extreme)
  id<-1:np
  if (dim(extreme)[1]!=0){
    id_y<-id[-extreme[,1]]
    pers<-cbind(id_y,p.s$theta.table[[1]])
    pers.se<-cbind(id_y,unlist(p.s$se.theta))
    pers.extrm=cbind(extreme[,1],NA)
    pers.extrm.se=cbind(extreme[,1],NA)
    pers=rbind(pers,pers.extrm)
    pers.se=rbind(pers.se,pers.extrm.se)
    pers=pers[order(pers[,1]),]
    pers.se=pers.se[order(pers.se[,1]),]
  }else{
    id_y<-id
    pers<-cbind(id_y,p.s$theta.table[[1]])
    pers.se<-cbind(id_y,unlist(p.s$se.theta))
    pers=pers[order(pers[,1]),]
    pers.se=pers.se[order(pers.se[,1]),]
  }
```

```
Full_data_Theta_rm=cbind(pers[,2],cbind(pers.se[,2]))
```

```
write.table(Full_data_Theta_rm[,1],file=paste(toString(cur),"_Full_data_Theta.txt"),
col.names=FALSE, row.names=FALSE)
```

```
write.table(Full_data_Theta_rm[,2],file=paste(toString(cur),
"_Full_data_Theta_se.txt"),col.names=FALSE, row.names=FALSE)
```

#Set 1 ve Set 2'de yer alan maddelerin tanımlanması

```
Full_data_Items=sample(1:ni,38,replace=FALSE)
```

```
write.table(Full_data_Items,file=paste(toString(cur),"_Full_data_Items.txt"),
col.names=FALSE, row.names=FALSE)
```

```
Set1_Items=Full_data_Items[1:20]
```

```
write.table(Set1_Items,file=paste(toString(cur),"_Set1_Items.txt"),
col.names=FALSE, row.names=FALSE)
```

```
Set2_Items=Full_data_Items[21:38]
```

```
write.table(Set2_Items,file=paste(toString(cur),"_Set2_Items.txt"),
col.names=FALSE, row.names=FALSE)
```

#Set 1'in tanımlanması ve madde zorluk kestirimlerinin kaydedilmesi

```
Set1=Full_data[c(1:500),c(Set1_Items)]
```

```
write.table(Set1,file=paste(toString(cur),"_Set1.txt"),sep=" ",row.names=FALSE,
col.names=FALSE)
```

```
Set1_Beta=c()
```

```
for(i in 1:20){
```

```
    Set1_Beta[i]=FD$betapar[Set1_Items[i]]
```

```
}
```

```
write.table(Set1_Beta,file=paste(toString(cur),"_Set1_Beta.txt"),sep=" ",
```

```
col.names=FALSE, row.names=FALSE)
```

```
Set1_Beta_se=c()
```

```
for(i in 1:20){
```



```

        Set1_Beta_se[i]=FD$se.beta[Set1_Items[i]]
    }
write.table(Set1_Beta_se,file=paste(toString(cur),"_Set1_Beta_se.txt"),sep="",
col.names=FALSE, row.names=FALSE)

#Ortak madde setinin tanımlanması
common_item=c()
common_item=which(Set1_Beta%in%sort(Set1_Beta)[1:2])
common_item_FD=c(Set1_Items[common_item[1]],Set1_Items[common_item[2]])
write.table(common_item_FD,file=paste(toString(cur),"_common_item_FD.txt"),
sep="",row.names=FALSE,col.names=FALSE)

#Set 2'nin tanımlanması ve parametre kestirimlerinin elde edilmesi
Set2=Full_data[c(501:np),c(Set2_Items,common_item_FD[1],
common_item_FD[2])]
write.table(Set2,file=paste(toString(cur),"_Set2.txt"),sep="",row.names=FALSE,
col.names=FALSE)

S2=RM(Set2)

S2_Beta_rm=cbind(S2$betapar,S2$se.beta)
write.table(S2_Beta_rm[,1],file=paste(toString(cur),"_Set2_Beta.txt"),
col.names=FALSE, row.names=FALSE)
write.table(S2_Beta_rm[,2],file=paste(toString(cur),"_Set2_Beta_se.txt"),
col.names=FALSE, row.names=FALSE)

p.s2<-person.parameter(S2)
extreme_2=matrix(p.s2$pers.ex,nrow=length(p.s2$pers.ex),ncol=1)
extreme_2=cbind(extreme_2)
id<-((np/2)+1):np
if (dim(extreme_2)[1]!=0){
    id_y<-id[-extreme_2[,1]]

```

```

pers_2<-cbind(id_y,p.s2$theta.table[[1]])
pers_2.se<-cbind(id_y,unlist(p.s2$se.theta))
pers_2.extrm=cbind(extreme_2[,1],NA)
pers_2.extrm.se=cbind(extreme_2[,1],NA)
pers_2=rbind(pers_2,pers_2.extrm)
pers_2.se=rbind(pers_2.se,pers_2.extrm.se)
pers_2=pers_2[order(pers_2[,1]),]
pers_2.se=pers_2.se[order(pers_2.se[,1]),]
} else {
  id_y<-id
  pers_2<-cbind(id_y,p.s2$theta.table[[1]])
  pers_2.se<-cbind(id_y,unlist(p.s2$se.theta))
  pers_2=pers_2[order(pers_2[,1]),]
  pers_2.se=pers_2.se[order(pers_2.se[,1]),]
}

```

```

S2_Theta_rm=cbind(pers_2[,2],cbind(pers_2.se[,2]))
write.table(S2_Theta_rm[,1],file=paste(toString(cur),"_Set2_Theta.txt"),
col.names=FALSE, row.names=FALSE)
write.table(S2_Theta_rm[,2],file=paste(toString(cur),"_Set2_Theta_se.txt"),
col.names=FALSE, row.names=FALSE)

```

#Set 2'nin ortak maddeler yardımı ile Set 1'e eşitlenmesi

```

disc=c(rep(1,20))
asym=c(rep(0,20))

```

```

x_1=Set1_Beta
Set2_Beta=c()
for(i in 1:20){
  Set2_Beta[i]=S2$betapar[i]
}
x_2=Set2_Beta

```

```
cat1=c(rep(2,20))
```

```
cat2=c(rep(2,20))
```

```
pm1=as.poly.mod(20)
```

```
pm2=as.poly.mod(20)
```

```
common=matrix(c(common_item,19:20),2,2)
```

```
pars=as.irt.pars(x=list(x_1,x_2),common=common,cat=list(cat1,cat2),
```

```
poly.mod=list(pm1,pm2))
```

```
sonuc=plink(pars,rescale="SL")
```

```
S2_Beta_equate=link.pars(sonuc)$group2
```

```
write.table(S2_Beta_equate[,2],file=paste(toString(cur),"_Set2_Beta_equate.txt"),
```

```
col.names=FALSE, row.names=FALSE)
```

#Orijinal setten ve eşitleme sonucunda elde edilen madde zorluk kestirimlerinin düzenlenmesi

```
final_items=c(Set1_Items[-c(common_item[1],common_item[2])],Set2_Items,
```

```
common_item_FD[1],common_item_FD[2])
```

```
write.table(final_items,file=paste(toString(cur),"_final_items.txt"),
```

```
col.names=FALSE, row.names=FALSE)
```

```
Full_Beta=c()
```

```
for(i in 1:ni){
```

```
    Full_Beta[i]=FD$betapar[final_items[i]]
```

```
}
```

```
write.table(Full_Beta,file=paste(toString(cur),"_Full_Beta.txt"),sep="",
```

```
col.names=FALSE, row.names=FALSE)
```

```
Set2_Beta_equate=S2_Beta_equate[,2]
```

```
Full_Equate_Beta=c(Full_Beta[1:18],Set2_Beta_equate)
```

```
write.table(Full_Equate_Beta,file=paste(toString(cur),"_Full_Equate_Beta.txt"),
```

```
sep="",col.names=FALSE, row.names=FALSE)
```

```
Final_Beta=data.frame(final_items,Full_Beta,Full_Equate_Beta)
```

```
Unified_Beta=Final_Beta[order(final_items),]
```

```
write.table(Unified_Beta,file=paste(toString(cur),"_Unified_Beta.txt"),
```

```
sep=" ",col.names=FALSE, row.names=FALSE)
```

```
cur=cur+1
```

```
if(cur>1000)break()
```

```
}
```

ÖZGEÇMİŞ

I- Bireysel Bilgiler

Adı ve Soyadı : Pervin DEMİR
Doğum yeri, tarihi : Keçiören, 07.10.1986
Uyruğu : T.C.
Medeni durumu : Bekar
İletişim Adresi ve Telefonu : Yıldırım Beyazıt Üniversitesi
Tıp Fakültesi Biyoistatistik AD.
Çankırı Cad. Çiçek Sok. No:3 Altındağ/ ANKARA
0312- 324 15 55 / 1130

II- Eğitimi

Yüksek Lisans : Ankara Üniversitesi, Tıp Fakültesi, Biyoistatistik AD,
Mezuniyet yılı: -
Lisans : Ankara Üniversitesi, Fen Fakültesi, İstatistik Bölümü (Bölüm 3.sü),
Mezuniyet yılı: 21.06.2009, Not Ort: 3,47 (4'lük sistem)
Lise : Fatih Sultan Mehmet Lisesi (YDA), Fen Bilimleri, Hazırlık +3 yıl
Mezuniyet yılı: 15.06.2004, Not Ort: 4,36 (5'lik sistem)

Yabancı Dili: İngilizce (2012 İlkbahar ÜDS: 60)

III- Ünvanları

2012 - : Araştırma Görevlisi; YBU, Tıp Fakültesi, Biyoistatistik AD

IV- Mesleki Deneyimi

2012 - : Araştırma Görevlisi; YBU, Tıp Fakültesi, Biyoistatistik AD

V- Üye Olduğu Bilimsel Kuruluşlar

Biyoistatistik Derneği

VI- Bilimsel Çalışmalar

Yayınların Listesi

1. Özelçi, R., Dilbaz, B., Akdağ Çırık, D., Zengin, T., **Demir, P.**, Yeğenoğlu, B. (2014). Post Menopozal Kadınlarda Sağ Ve Sol Femur Kemik Mineral Yoğunluğu Arasındaki Farkın Değerlendirilmesi. Jinekoloji - Obstetrik ve Neonatoloji Tıp Dergisi (The Journal of Gynecology - Obstetrics and Neonatology) 2014; 11(2): 37 – 40.
2. Argüder, E., Karalezli, A., Hezer, H., Kılıç, H., Er, M., Hasanoğlu, H. C., **Demir, P.** (2013). Sigara Bırakma Başarısını Etkileyen Faktörler. Türk Toraks Dergisi, 2013, 14(3): 081-087. Doi: 10.5152/ttd.2013.18.
3. Kaymaz, D., Ergün, P., Candemir, İ., Utku, E., Demir, N., Şengül, F., Egesel, N., **Demir, P.** (2013). Türk Toraks 2013, 61(4): 295-302.
4. Kaymaz, D., Ergün, P., Günay, E., Aktaş, S., Mendil, N., Battal, S., Utku, E., **Demir, P.** (2011). Hospital at home for patients with acute exacerbations of chronic obstructive pulmonary disease; will it be an effective home care model?. European Respiratory Journal (ERJ), September 1, 2011. Vol.38, no. Suppl 55, p 1246.

Ulusal ve Uluslararası Kongrelerde Sunulan Bildiriler ve Posterler

1. **Demir, P.**, Elhan, A. H., Öztuna, D. (2014). Paralel Test Oluşturmada Ortak Soru Seçimini Etkileyen Faktörlerin Benzetim Çalışması ile Değerlendirilmesi. XVI. Ulusal Biyoistatistik Kongresi, Bildiri Özetleri Kitabı S08, s: 13, Side, ANTALYA, 10 – 12 Eylül 2014. (Sözlü Bildiri).
2. **Demir, P.**, Aktaş, H., Bayraklı, A., Yüksel, S. (2014). Queuing Theory: Adaptation of the Theory on Lunch Service System For Yildirim Beyazıt University (YBU). XXVII. International Biometric Conference (IBC), Florence, ITALY, 6-11 July 2014 (Poster Presentation)

3. **Demir, P.**, Yüksel, S., Elhan, A. H., Sanisoğlu, S.Y., Akyol, M. (2013). Yapay Madde İşlev Farklılığına Neden Olan Faktörlerin Benzetim Çalışması İle Değerlendirilmesi. Uluslararası Katılımlı 15. Ulusal Biyoistatistik Kongresi, Bildiri Özetleri Kitabı S15, s: 24, Didim/AYDIN, 20-23 Ağustos 2013. (Sözlü Bildiri)
4. Simsek Utku, E., Yılmaz Kaymaz, D., Ergun, P., **Demir, P.**, Mendil Alptekinoğlu, N., Candemir, İ., Demir, N. (2013). Relationship between London chest daily living activity scale and i-BODE index in COPD. European Respiratory Society Annual Congress 2013, P3704.
5. Karalezli, A., Kılıç, H., Hezer, H., Babaoğlu, E., Argüder, E., Er, M., **Demir, P.**, Hasanoğlu, H. C. (2013). Sigara Bırakma Polikliniğine Başvuran Olgularda Kronik Obstrüktif Akciğer Hastalığı Farkındalık Oranı. Türkiye Solunum Araştırmaları Derneği 35. Ulusal Kongresi, SS098, Çeşme/ İZMİR, 2-6 Ekim 2013. (Sözlü Bildiri)
6. Hezer, H. Karalezli, A., Kılıç, H., Argüder, E., Er, M., Şentürk, A., Hasanoğlu, H. C., **Demir, P.** (2013). Sigarayı Bırakma Oranları Üzerine Anksiyete ve Depresyon Skoru Ölçeklerinin Katkısı. Türkiye Solunum Araştırmaları Derneği 35. Ulusal Kongresi, TP136, Çeşme/ İZMİR, 2-6 Ekim 2013.
7. **Demir, P.**, Elhan, A. H., Öztuna, D. (2012). Genellenebilirlik Kuramı. Uluslararası Katılımlı XIV. Ulusal Biyoistatistik Kongresi, Bildiri Özetleri Kitabı P10, s: 85, KAYSERİ, 4-7 Eylül 2012. (Poster Sunumu)
8. Kaymaz, D., Ergün, P., Utku, E., Battal, S., Candemir, İ., Mendil, N., Taşdemir, F., Sengül, F., Egesel, N., **Demir, P.** (2012). Prevalence and the association of anaemia with clinical and functional parameters in patients with COPD who referred to a pulmonary rehabilitation programme. European Respiratory Society Annual Congress 2012, P3550.
9. Türkoğlu Selçuk, N., Ergün, P., Kaymaz, D., Egesel, N., **Demir, P.** (2011). Pulmoner Rehabilitasyon Programlarında Hasta Uyumsuzluğu Nedenleri, Türk Toraks Derneği 13. Yıllık Kongresi, Tartışmalı Poster Sunusu 292, İSTANBUL, 5 – 9 Mayıs 2011.

10. Ergün, P., Kaymaz, D., Ertürk, H., Türkoğlu Selçuk, N., Demir, N., Şengül, F., **Demir, P.** (2011). Periferik Kas Kaybı Olan Orta- İleri Evre Koah'lı Olgularda Nmes Ve Endurans Egzersiz Eğitim Etkinliği; Kuadriceps Kası Kesitsel Alanın Bilgisayarlı Tomografi İle Değerlendirilmesi. Türk Toraks Derneği 13. Yıllık Kongresi, Sözlü sunum 646, İSTANBUL, 5 – 9 Mayıs 2011. (3.lük Bildiri Ödülü) (Sözlü Bildiri)
11. Kaymaz, D., Ergün, P., Günay, E., Aktaş, E., Mendil, N., Battal, S., Utku, E., **Demir, P.** (2011). KOAH akut ataklı olgularda evde hastane uygulaması etkin bir evde sağlık hizmet sunum modeli midir?. Türk Toraks Derneği 14. Yıllık Kongresi, Sözlü sunum 236, ANTALYA, 13-17 Nisan 2011. (1.lük Bildiri Ödülü) (Sözlü Bildiri)

VII- Bilimsel Etkinlikleri

Projeler

1. "Romatoid Artritli Hastalarda Bilgisayar Uyarlamalı Test Yöntemi İle Özürlülük Değerlendirmesi" (2010-2012 , 42.841 EUR), TÜBİTAK 1001 Araştırma Projeleri, **Bursiyer.**

Seminerler

1. Test Eşitleme ve Test Eşitlemede Kullanılan Yöntemler, 11.04.2014, Yıldırım Beyazıt Üniversitesi, Tıp Fakültesi, Biyoistatistik AD.
2. Gerçek ve Yapay Madde İşlev Farklılığı, 11.06.2013, Hacettepe Üniversitesi, Tıp Fakültesi, Biyoistatistik AD.
3. Genellenebilirlik Kuramı, 06.03.2012, Ankara Üniversitesi, Tıp Fakültesi, Biyoistatistik AD.

VIII- Diğer Bilgiler

Katıldığı Diğer Kurslar, Toplantılar

1. XIII. Ulusal Biyoistatistik Kongresi, Patalya Termal Hotel, Kızılcahamam/ANKARA, 12-14.09.2011 (Katılımcı)
2. Rumm2030 Programında Rasch Analizi Kursu Katılım Belgesi
Introductory Rasch Analysis Ankara University Faculty of Medicine
Department of Biostatistics' Psychometric Laboratory, 30.09.2011 –
02.10.2011.
3. 'Mükemmellik Modeli' Altı Sigma Semineri Katılım Belgesi SPAC 6 Sigma
Danışmanlık, 24.12.2009
4. Stajyer Katılımcı Belgesi Türkiye Cumhuriyeti Merkez Bankası(TCMB), 02 –
06.02.2009
5. Bilgisayar İşletmenliği (Operatörlüğü) Sertifikası MEB, 11.10.2008 –
31.01.2009
6. VI. İstatistik Öğrenci Kolokyum Katılım Belgesi Çukurova Ünv., 18 –
19.05.2009
7. V. İstatistik Öğrenci Kolokyum Katılım Belgesi Selçuk Ünv.
'Dünyada ve Türkiye' de Kuş Gribi ve İnsanlar Üzerindeki Etkisi' başlıklı
sunum ile katıldım., 6 – 7.05.2008
8. Tezhip Kursu Bitirme Belgesi MEB, 10.11.2010 – 08.06.2011