



TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ



ÖLÇME EŞDEĞERLİĞİNİN
ÇOKLU GRUP UYGULAMALARI VE
RASCH ANALİZİ İLE DEĞERLENDİRİLMESİ

Serhat HAYME

BİYOİSTATİSTİK ANABİLİM DALI
YÜKSEK LİSANS TEZİ

DANIŞMAN

Doç. Dr. Derya ÖZTUNA

2015- ANKARA

**TÜRKİYE CUMHURİYETİ
ANKARA ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

**ÖLÇME EŞDEĞERLİĞİNİN
ÇOKLU GRUP UYGULAMALARI VE
RASCH ANALİZİ İLE DEĞERLENDİRİLMESİ**

**Serhat
HAYME**

**BIYOİSTATİSTİK
ANABİLİM DALI
YÜKSEK LİSANS TEZİ**

**DANIŞMAN
Doç. Dr. Derya Öztuna**

2015- ANKARA

Ankara Üniversitesi Sağlık Bilimleri Enstitüsü

Biyoistatistik Yüksek Lisans **Programı**

çerçevesinde yürütülmüş olan bu çalışma aşağıdaki jüri tarafından

Yüksek Lisans **Tezi** olarak kabul edilmiştir.

Tez Savunma Tarihi: 27/04/2015

Prof.Dr.Atilla Halil ELHAN
Ankara Üniversitesi
Tıp Fakültesi
Biyoistatistik Anabilim Dalı
Jüri Başkanı

Doç.Dr.Derya ÖZTUNA
Ankara Üniversitesi
Tıp Fakültesi
Biyoistatistik Anabilim Dalı
Tez Danışmanı

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

İçindekiler

Kabul ve Onay	ii
İçindekiler	iii
Önsöz	iv
Simgeler Ve Kısaltmalar	v
Şekiller	vi
Çizelgeler	vii
1. GİRİŞ	1
1.1 Araştırmanın Konusu, Amacı ve Önemi	1
1.2 Ölçme Eşdeğerliği	3
1.3 Yapısal Eşitlik Modellemesi	4
1.3.1 Doğrulayıcı Faktör Analizi (DFA)	6
1.3.2 Çoklu Grup Doğrulayıcı Faktör Analizi	12
1.3.3 Ölçme Eşdeğerliğinin Sağlanamama Nedenleri ve Uygulanacak Stratejiler	19
1.4 Rasch Analizi	21
1.5 ÇG-DFA ile Rasch Analizi Yaklaşımlarının Karşılaştırılması	24
2. GEREÇ VE YÖNTEM	27
3. BULGULAR	31
3.1 ÇG-DFA Sonuçları	31
3.1.1 Cinsiyet Bazında ÇG-DFA Sonuçları	31
3.1.2 Hastalık Süresi Bazında ÇG-DFA Sonuçları	37
3.2 Rasch Analizi Sonuçları	43
4. TARTIŞMA	46
5. SONUÇ VE ÖNERİLER	50
ÖZET	51
SUMMARY	52
KAYNAKLAR	53
ÖZGEÇMİŞ	55
EKLER	57

ÖNSÖZ

Akademik ve insani yönden kendime örnek aldığım, tez süresince karşılaştığım tüm zorlukları aşmamda bana yardımcı olan ve bu sıkıntılı süreçteki en büyük destekçim olan değerli danışmanım Doç. Dr. Derya ÖZTUNA' YA,

Lisansüstü eğitimim boyunca bilgi birikimi ve yönlendirmeleriyle bana rehber olan, çalışma disiplini örnek aldığım hocam Prof. Dr. Atilla Halil ELHAN'A,

Yüksek lisansa başladığımdan bu yana yetişmemde emeği bulunan, bilgi ve tecrübelerinden yararlandığım ve desteklerini her zaman hissettiğim hocalarım Prof. Dr. Yasemin YAVUZ, Doç.Dr. Serdal Kenan KÖSE ve Yrd.Doç.Dr. Beyza Doğanay ERDOĞAN'A,

Tez süreci gibi sıkıntılı bir dönemi rahat atlatmam için her türlü fedakarlığı gösteren oda arkadaşlarım ve çalışma arkadaşlarıma,

Destegini ve yardımını benden esirgemeyen arkadaşım Arş. Gör. Pervin DEMİR'E, tez çalışmama veri sağladığı için Dr. O.Tolga KASKATI'ya

Her zaman yanımda olduklarını bildiğim, benden hiçbir fedakarlığı esirgemeyen ailem ve dostlarıma bana karşı gösterdiğiniz ilgi ve anlayış için teşekkürü borç bilirim.

SİMGELER VE KISALTMALAR

AFA	Açıklayıcı Faktör Analizi
ÇG-DFA	Çoklu Grup Doğrulayıcı Faktör Analizi
DFA	Doğrulayıcı Faktör Analizi
DMF	Değişen Madde Fonksiyonu
KKM	Kısmi Kredi Modeli
KTT	Klasik Test Teorisi
KUİ	Karşılaştırmalı Uyum İndeksi
MİF	Madde İşlev Farklılığı
MSDA	Modifiye Sağlık Değerlendirme Anketi
MYT	Madde Yanıt Teorisi
RA	Romatoid Artrit
RM	Rasch Model
SDA	Sağlık Değerlendirme Anketi
SSM	Sıralı Sonuçlu Model
TLI	Tucker-Lewis İndeksi
YEM	Yapısal Eşitlik Modellemesi
YHKOK	Yaklaşımın Hata Kareler Ortalaması Karekökü
θ	İncelenen Özellik

ŞEKİLLER

Şekil 1.1. İki faktörlü bir ölçme modeli (kırmızı) ve yapısal model (yeşil)	5
Şekil 1.2. Ölçme modelinin grafiksel gösterimi	7
Şekil 1.3. Ölçme modelinin matematiksel gösterimi	10
Şekil 1.4. Farklı ölçme eşdeğerliği gösterimleri (a) Güçlü değişmezliğin sağlanması durumu, (b) Zayıf değişmezliğin sağlanmaması durumu, (c) Güçlü değişmezliğin sağlanmaması durumu	17
Şekil 3.1. “h4-Yatağa yatıp, kalkabiliyor musunuz” maddesinin cinsiyete göre MİF incelemesi	45
Şekil 3.2. “h6- Dolu bir fincanı veya bardağı ağızınıza götürebiliyor musunuz” maddesinin hastalık süresine göre MİF incelemesi	45

ÇİZELGELER

Çizelge 3.1 Cinsiyet bazında her bir grup için DFA'ya ilişkin uyum iyiliği istatistikleri	31
Çizelge 3.2 Cinsiyet bazında ÇG-DFA'ya ilişkin uyum iyiliği istatistikleri ve karşılaştırma sonuçları	32
Çizelge 3.3 Yapısal değişmezlik modeline göre cinsiyetler bazında faktör yükleri ve eşik değerleri	33
Çizelge 3.4 Zayıf değişmezlik modeline göre cinsiyetler bazında faktör yükleri ve eşik değerleri	34
Çizelge 3.5 Güçlü değişmezlik modeline göre cinsiyetler bazında faktör yükleri ve eşik değerleri	36
Çizelge 3.6 Hastalık süresi bazında her bir grup için DFA'ya ilişkin uyum iyiliği istatistikleri	37
Çizelge 3.7 Hastalık süresi bazında ÇG-DFA'ya ilişkin uyum iyiliği istatistikleri ve karşılaştırma sonuçları	38
Çizelge 3.8 Yapısal değişmezlik modeline göre hastalık süresi bazında faktör yükleri ve eşik değerleri	39
Çizelge 3.9 Zayıf değişmezlik modeline göre hastalık süresi bazında faktör yükleri ve eşik değerleri	40
Çizelge 3.10 Güçlü değişmezlik modeline göre hastalık süresi bazında faktör yükleri ve eşik değerleri	42
Çizelge 3.11 Cinsiyet bazında MİF sonuçları	43
Çizelge 3.12 Hastalık süresi bazında MİF sonuçları	44

1. GİRİŞ

1.1 Araştırmanın Konusu, Amacı ve Önemi

Günlük yaşıntıda ve farklı disiplinlerde gerçekleştirilen bilimsel araştırmalarda ölçme önemli bir yer tutmaktadır. Ölçme, genel anlamıyla herhangi bir özelliği gözlemek ve gözlem sonuçlarını sayılarla ve sembollerle ifade etmektir (Turgut, 1977). Bir bireyin boyunun/kilosunun ölçülmesi fiziksel bir ölçme olup; “doğrudan ölçme” kapsamına girerken, aynı bireyin yetenek seviyesinin ölçülmesi ise psikolojik ölçme olup; “dolaylı ölçme” kapsamına girmektedir. Doğrudan ölçmede, bir bireyi bir kez ölçmek yeterli olabilirken; dolaylı ölçmede tekrarlanan her bir ölçümde farklı sonuçlar elde edilebilmektedir. Çünkü dolaylı ölçmede, doğrudan ölçme için söz konusu olan hata kaynaklarından (ölçücü, ölçme aracı ve ölçülen özellik) başka, asıl ölçme konusu olan davranış ile gerçekte gözlenen davranış arasında kurulan ilişki de bir hata kaynağı olabilir (Tekin, 2002; Öztuna, 2008).

Ölçme araçları, ölçülen büyüklüğün sayılarla ifade edilmesini; böylelikle gözlemin daha duyarlı yapılarak objektif ve standart ölçümler elde edilmesini sağlar. Ölçek ise, doğrudan ölçüm yapılamayan, gizli (latent) özelliklerin ölçülmesinde kullanılan, psikometrik özellikleri (geçerlilik, güvenilirlik, kullanışlılık, değişime duyarlılık) belirlenmiş ölçme aracı olarak tanımlanabilir. Geçerlilik, ölçme aracının ölçülecek özelliği, tam ve doğru olarak, başka özelliklerle karıştırmadan ölçebilmesidir.

Geçerlilik kapsamında dikkat edilmesi gereken bir diğer önemli nokta, ölçme aracı farklı özellikteki gruplara uygulandığı durumda ortaya çıkar. Bahsedilen bu gruplara örnek olarak “cinsiyet”, “eğitim durumu” ve “gelir düzeyi” gibi demografik özellikler verilebileceği gibi, “hastalık grupları”, “hastalık süresi” gibi tanısallık özellikler ya da kültürlerarası geçerliliğin değerlendirildiği bir çalışmada “ülke”, “coğrafi bölge” gibi özellikler verilebilir.

Ölçme aracı farklı gruplarda yer alan bireylere uygulandığında, gruplar arası karşılaştırmaların anlamlı olabilmesi için ölçme aracının ölçtüğü yapının eşdeğer olması gereklidir. Ölçme eşdeğerliği olarak adlandırılan bu kavram, herhangi bir bireyin belirli bir gözlenen puana sahip olma olasılığının hangi grupta yer aldığından bağımsız olma durumudur (Somer ve ark., 2009). Örneğin aynı özürlülük düzeyine sahip iki birey olsun. Bu bireylerin ölçme aracından elde edilen puanları, bireylerden birinin erkek diğerkinin kadın olması nedeniyle farklılık göstermemelidir. Ölçme aracında yer alan herhangi bir madde için ölçme eşdeğerliliğinin sağlanmaması durumunda, ölçme aracının cinsiyete göre farklı işleyişe sahip olduğu yorumu yapılır ve bu durum ölçme aracının geçerliliğini zedeler.

Ölçme eşdeğerliği, farklı yaklaşımlar kullanılarak değerlendirilmesine rağmen, yapısal eşitlik modellemesi (YEM) kapsamında “Çoklu Grup Doğrulayıcı Faktör Analizi (ÇG-DFA)” ve madde yanıt teorisi (MYT) yaklaşımı kapsamında ise “Rasch Analizi” yaygın olarak kullanılan yöntemlerdir.

Ölçme eşdeğerliğinin değerlendirilmesinde faktör analizi kapsamında en çok uygulanan yaklaşım, ölçme aracının bütününden elde edilen puanların bireylere ilişkin alt gruplarda karşılaştırılması biçiminde olmaktadır. Ancak bu yaklaşım; faktör yüklerinin, faktörler arası korelasyonun ve hata varyanslarının gruplarda aynı olması varsayımlarının test edildiği ÇG-DFA ile karşılaştırıldığında teorik olarak yetersiz ve yüzeysel bir değerlendirme sağlamaktadır. Sağlık alanında ölçme eşdeğerliğinin, madde işlev farklılığı (MİF) kapsamında değerlendirilmesinde Rasch analizinin yaygın olarak kullanıldığı bilinmektedir.

Bu çalışmada, ölçme eşdeğerliğinin değerlendirilmesinde “ÇG-DFA” ve “Rasch Analizi”nin benzerlikleri ve farklılıkları irdelenecektir. Sağlık alanında kullanılan ve geçerli olduğu gösterilen bir ölçme aracının ölçme eşdeğerliği, incelenecek bir demografik özellik ya da hasta karakteristiği bakımından hem ÇG-DFA ile hem de Rasch Analizi ile değerlendirilecektir.

1.2 Ölçme Eşdeğerliği

Ölçme aracı farklı gruplarda yer alan bireylere uygulandığında, gruplar arası karşılaştırmaların anlamlı olabilmesi için ölçme aracının ölçtüğü yapının eşdeğer olması gereklidir.

Drasgow ve Kanfer (1985)'e göre, eğer incelenen özellik (θ) bakımından benzer puanlara sahip bireyler, madde ya da alt ölçek düzeyinde beklenen ham puan ya da gerçek puana sahiplerse, ilgili ölçme aracı ya da alt ölçek, gruplar ya da kitleler arasında ölçme eşdeğerliğine (measurement equivalence) sahiptir denir (Raju et.al., 2002).

Daha basit bir ifade ile ölçme eşdeğerliği, herhangi bir bireyin belirli bir gözlenen puana sahip olma olasılığının hangi grupta yer aldığından bağımsız olma durumudur. Bu tanımdan hareketle ölçme eşdeğerliğinin sağlandığı koşulda, farklı gruplarda yer alan aynı gerçek puana sahip olan bireyler aynı gözlenen puana sahip olacaktır. Bu koşulun sağlanmadığı durumda yapılacak olan grup karşılaştırmalarından elde edilecek farklılıkların ölçmedeki bir yanlılıktan mı yoksa gerçek grup farklılıklarından kaynaklandığını yorumlamak sorunlu olabilmektedir (Somer ve ark., 2009).

Ölçme eşdeğerliğinin değerlendirilmesinde, yaygın ve popüler iki yaklaşım bulunmaktadır. Bu yaklaşımlardan birisi YEM kapsamında ÇG-DFA'dır. Diğeri ise MYT'ye dayalı olarak MİF'in değerlendirildiği Rasch analizidir. MYT kapsamında MİF, farklı gruplardaki bireylerin θ düzeyleri sabit tutulduğunda, bu bireylerin madde(ler) üzerindeki başarı olasılıklarının veya olabirliklerinin farklı olmasıdır. Diğeri bir ifadeyle, θ boyunca belirli düzeylere sahip olan; fakat farklı kitlelerden gelen bireylerin bir maddeye yanıt verme olasılıklarının farklı olmasıdır (Öztuna, 2008). YEM kapsamında ise ölçme eşdeğerliği, kovaryans yapılarının karşılaştırılan gruplar arasında eşdeğer olmasını ifade etmektedir ve ÇG-DFA ile değerlendirilir.

1.3 Yapısal Eşitlik Modellemesi

Çoklu grup uygulamaları, yapısal eşitlik modellemesi (YEM) kapsamında yer alan analiz yöntemleridir.

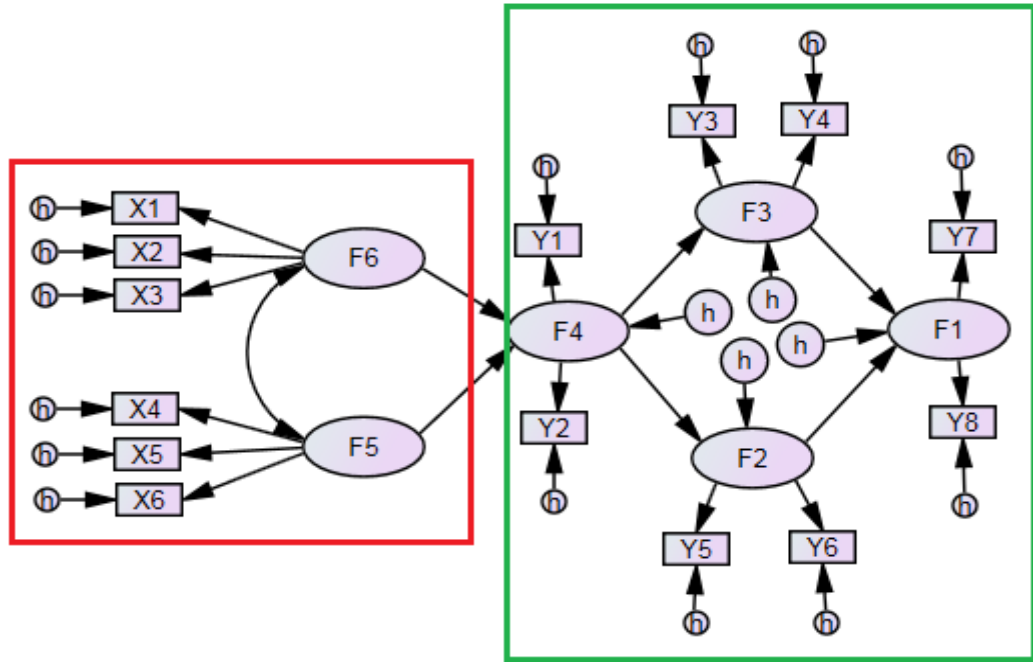
YEM çalışmalarının en belirgin özelliği tümüyle kurama dayalı olmasıdır. Yapılacak analizlerde amaç, araştırmacının kendisinin oluşturduğu ya da daha önceden var olan kuramsal bir yapının sorgulanması denilebilir. Araştırmanın ilk aşamasında bu kuramsal çerçeve ortaya konmakta, değişkenler arasındaki ilişki örüntüsü belirlenmekte ve tümdengelimci bir yaklaşım ile bu kuramsal yapının veriler tarafından doğrulanıp doğrulanmadığı test edilmektedir. Tüm YEM çalışmaları, özü itibarıyla doğrulayıcı özelliğe sahiptir.

Eğitim bilimleri, psikoloji, tıp gibi alanlarda birçok teori ve model direkt olarak gözlenemeyen (örtük) değişkenler aracılığıyla teorik bir yapıya uygun şekilde ortaya konmaktadır. Örtük değişken(ler) ve gözlenen değişkenler arasındaki ilişkileri ortaya koyabilme açısından YEM yaygın olarak kullanılmaktadır.

Modelleme açısından YEM çalışmaları temel olarak 3'e ayrılır (Şimşek, 2007):

- a-** Belirlenmiş bir modelin veri tarafından doğrulanıp doğrulanmadığını test etmeye yönelik “***Doğrulayıcı modelleme stratejisi***”,
- b-** Bir dizi değişken ele alındığında modeller arasından en çok hangisinin veri tarafından desteklendiğini belirlemeye yönelik “***Alternatif modeller stratejisi***”,
- c-** Bir dizi değişken arasındaki ilişkileri en iyi açıkladığı varsayılan bir modelin test edilmesi ve analiz sonuçlarına dayanarak modelin geliştirilmesi yönünde iyileştirmeler yapılmasını hedefleyen “***Model geliştirme stratejisi***”

YEM uygulamalarının genellikle iki parçası bulunmaktadır: ölçme modeli ve yapısal model. Modelin “ölçüm” kısmını oluşturan ölçme modeli, hangi faktörlerin gözlenen değişkenlere dayalı ölçüldüğünü ve bu ölçmelerin geçerlilik ve güvenilirliklerinin ne olduğunu tanımlamaktadır. Modelin “teorik” kısmını oluşturan yapısal model ise, faktörler arasındaki dolaylı ve doğrudan ilişkileri tanımlamakta ayrıca açıklanan ve açıklanamayan varyans miktarını göstermektedir (Öğretmen, 2006).



Şekil 1.1. İki faktörlü bir ölçme modeli (kırmızı) ve yapısal model (yeşil)

Şekil 1.1’de, iki faktörlü bir ölçme modeli (kırmızı alan) ve bu faktörler arasındaki ilişki yapısını gösteren yapısal model (yeşil alan) örneklenmiştir.

YEM kapsamında yaygın olarak kullanılan analiz türleri; doğrulayıcı faktör analizi (DFA), path (yol) modelleri, YEM uygulamaları, aracılık / düzenleyicilik (moderator & mediator) analizleri, gizli büyüme eğrileri (latent growth curve) modelleri, çoklu grup uygulamaları (multi-group analysis), çoklu nitelik çoklu yöntem (multi-trait multi-method) uygulamalarıdır.

Çoklu grup uygulamaları, DFA kapsamında, path analizi kapsamında veya YEM analizleri kapsamında uygulanabilmektedir. Örneğin DFA kapsamında; (içsel) yapı

geçerliliği değerlendirilen bir ölçme aracının, farklı alt gruplarda ya da kültürlerde benzer işleyişe sahip olup olmadığı değerlendirilebileceği gibi, elde edilen bir path modelinin veri setinde yer alan alt gruplar bazında geçerli olup olmadığı incelenebilir.

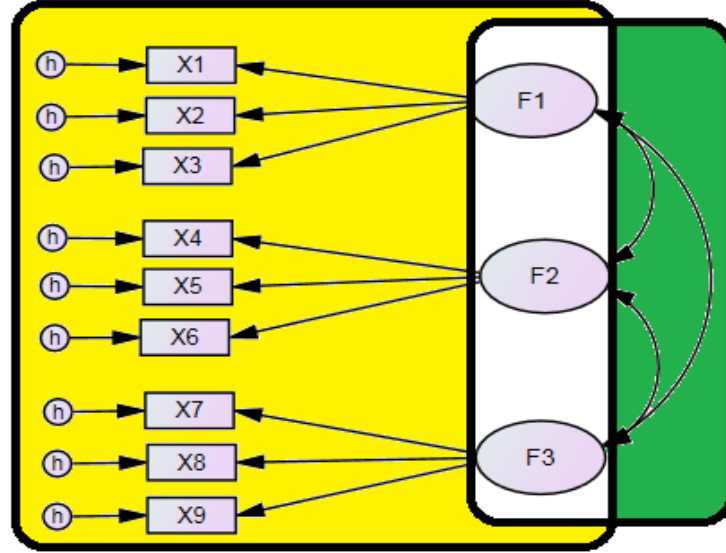
Tez çalışmasının amacı doğrultusunda, DFA kapsamında çoklu grup uygulamaları detaylı olarak incelenecektir.

1.3.1 Doğrulayıcı Faktör Analizi (DFA)

Faktör analizi, birbirleriyle ilişkili veri yapılarını birbirinden bağımsız ve daha az sayıda yeni veri yapılarına dönüştürmek, bir oluşumu, nedeni açıkladıkları varsayılan değişkenleri gruplayarak ortak faktörleri ortaya koymak, bir oluşumu etkileyen değişkenleri gruplamak, majör ve minör faktörleri tanımlamak amacıyla başvuru bir yöntemdir (Özdamar, 2002). Faktör analizi, açıklayıcı ve doğrulayıcı olmak üzere iki başlıkta incelenebilir.

Açıklayıcı faktör analizi (AFA), temel bir madde setinden yola çıkarak; herhangi bir önsel bilgi olmaksızın bu maddelerden oluşturulacak gizli yapı (faktör) sayısının belirlenmesinde kullanılan istatistiksel bir yöntemdir. DFA’da ise temel hipotez, “maddeler ile ait oldukları faktörler arasında bir ilişki vardır” biçiminde kurulur. Daha sonra, teorik bilgi, deneysel araştırma ya da her ikisi de kullanılarak ilişki yapısına ilişkin önsel bilgi ışığında yukarıdaki hipotez istatistiksel olarak test edilir. Klasik istatistiksel yöntemler, analizin önemliliğini belirlemek için tek bir istatistiksel test kullanırken; DFA verinin modele uyumunu belirlemek için çeşitli istatistiksel testlerden yararlanır. Modelin uyum istatistikleri yardımıyla faktörlerin gerçekten bu maddelerden oluşup oluşmadığına karar verilir. Sonuç olarak YEM’e dayalı DFA’nın amacı, önceden belirlenen bu ilişki örüntülerinin veri tarafından doğrulanıp doğrulanmadığını ortaya koymaktır (Öztuna, 2008).

YEM terminolojisinde yer alan kavramlar, bir ölçme modeli üzerinden (Şekil 1.2) açıklanabilir. Şekil 1.2’de altı farklı öge bulunmaktadır.



Şekil 1.2. Ölçme modelinin grafiksel gösterimi

Birincisi, **örtük değişken (latent variable) veya faktör** olarak adlandırılan, gözlenemeyen ve ancak bazı göstergeler (indicator) aracılığıyla ölçülebildikleri varsayılan teorik yapılardır. Örtük değişkenler YEM uygulamalarında oval veya dairelerle gösterilir.

İkincisi, **gözlenen değişkenler (observed variables)** olup; faktörleri ölçmede kullanılan göstergelerdir. Gözlenen değişkenler genellikle ölçme araçlarında yer alan maddeler olup; YEM çalışmalarında dikdörtgen veya karelerle gösterilir. Örtük değişkenler tamamen teorik yapılar oldukları için belirli bir ölçme birimine sahip olamazlar, bu nedenle ölçme modelleri test edilirken her birisini en iyi tanımladığı düşünülen bir gözlenen değişkene sabitlenirler. Bu değişkene referans değişken (reference variable) adı verilir. Araştırmacı teorik bilgisi çerçevesinde bu değişkenin hangisi olduğuna kendisi karar verebileceği gibi, araştırma sürecinde DFA’dan önce AFA uygulanmışsa her bir faktörde en yüksek faktör yüküne sahip olan değişken referans değişken olarak kullanılabilir.

Üçüncü öge ise, bu iki tür değişken arasındaki ilişkiyi gösteren *tek yönlü oklardır*. YEM içerisindeki her türlü çalışmada tek yönlü oklar, tek yönlü doğrusal ilişkiyi gösterir. Bu şekilde bakıldığında, ölçme modellerinde örtük değişkenlerin gözlenen değişkenleri yordadığı şeklinde bir hipotezin varlığını görürüz. Her bir ok, regresyon katsayılarına denk gelecek olan bir hipotez anlamına gelmektedir ve analiz sonuçlarına göre her bir katsayının anlamlı olup olmadığı yani her bir örtük değişkenin kendi gözlenen değişkenlerini anlamlı bir şekilde yordayıp yordamadığına bakılır. Anlamlı olması, her bir maddenin kendi örtük değişkeninin iyi bir temsilcisi olduğuna ilişkin bilgi verir.

Ölçme modelinde dördüncü öge, h harfi ile temsil edilen ve her bir gözlenen değişkende örtük değişken tarafından *açıklanamayan varyansı ya da hatayı (error)* ifade eden ögedir. Bu öge, her bir gözlenen değişkende söz konusu ölçme modeli ile açıklanamayan bir özelliğin var olduğunun bir göstergesidir.

Beşinci öge ise, hatadan gözlenen değişkene doğru giden *tek yönlü oklardır*. Bu anlamda, her gözlenen değişken kendi hatası ile ilişkilidir ve her gözlenen değişkende açıklanamayan bir şeyler varsa bunun söz konusu değişkenle bağlantılı olması beklenir.

Altıncı öge ise, örtük değişkenler arasındaki ilişkiyi gösteren *eğik ve iki yönlü oktur*. Bu oklar, korelasyon veya kovaryans (birlikte değişim) değeriyle eşdeğer olup; tek yönlü oklarda olduğu gibi değişkenler arasında bir neden sonuç (yordama) ilişkisini göstermez; yani ilişkinin yönü belli değildir (Şimşek, 2007).

Şekil 1.2’de gösterilen model, matematiksel olarak aşağıdaki gibi ifade edilir:

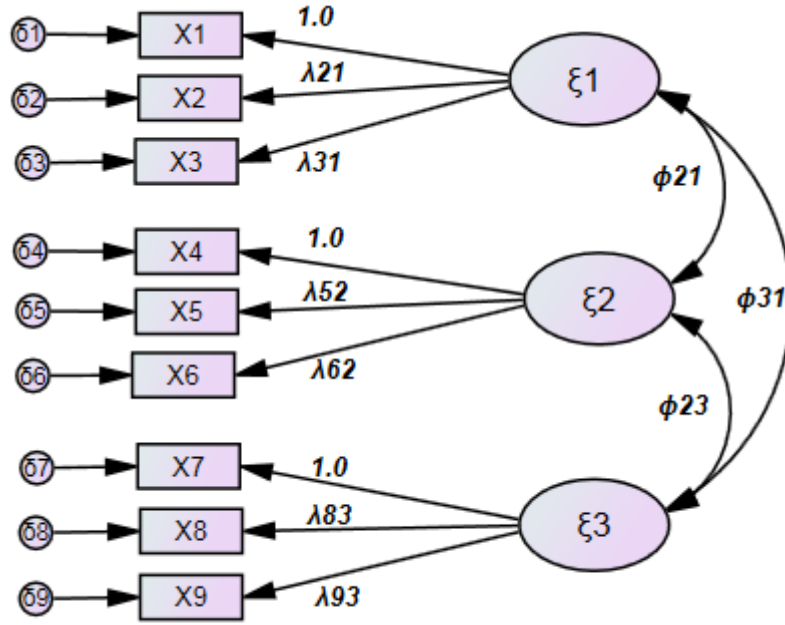
$$x = \Lambda_x \xi + \delta \quad (Eşitlik 1.1)$$

Doğrusal DFA modelinde, x ; $qx1$ boyutunda gözlenen değişken matrisi, ξ ; $nx1$ boyutunda örtük (gizli) değişken vektörü, Λ_x ; q sayıda gözlenen değişkeni n sayıda faktöre bağlayan $q \times n$ boyutunda regresyon katsayısı ya da faktör yükü matrisi (Λ_x matrisinin elemanları λ değerleridir) ve δ ; $qx1$ boyutunda ölçüm hatası ya da artıklar matrisidir. Kovaryans yapılarına ek olarak, regresyon sabitlerinin de eşitliği test edilmek istendiğinde (yani, gözlenen puan ortalamaları ile de ilgileniliyorsa) Eşitlik 1.1, Eşitlik 1.2'deki gibi ifade edilir ve bu eşitlikte τ ; $qx1$ boyutunda regresyon sabiti matrisidir:

$$x = \tau + \Lambda_x \xi + \delta \quad (Eşitlik 1.2)$$

DFA'da, gözlenen değişkenler nicel değişken olduğunda regresyon sabiti hesaplanırken; gözlenen değişkenler sınıflanabilir / sıralanabilir nitel değişken olduğunda regresyon sabiti yerine eşik değerleri söz konusu olur. Eşik değeri, nitel değişken açısından bir kategoriden diğerine geçişte, gizli değişken skoru 0 olduğunda, gizli değişken / faktörün beklenen değeridir. k düzeyli herhangi bir nitel değişken için, " $k-1$ " adet eşik değeri bulunmaktadır. Örneğin, 5 kategorili (0-4) Likert türünde bir madde için 4 adet eşik değeri vardır. İlk eşik değeri, bir bireyin ilgili değişkende 0 kategorisi yerine 1 kategorisini seçmesinin daha olası olduğu noktadaki beklenen değeri ifade etmektedir. İkinci eşik değeri, bireyin ilgili değişkende 1 kategorisi yerine 2 kategorisini seçmesi olasılığının daha yüksek olduğu noktadaki beklenen değeri ifade etmektedir. Benzer şekilde, dördüncü eşik değeri ise; bireyin 3 kategorisi yerine 4 kategorisini seçmesinin daha muhtemel olduğu noktadaki beklenen değere karşılık gelmektedir. Gözlenen değişkenlerin nitel olduğu durumda regresyon sabitlerinin eşitliği testi, eşik değerlerinin eşitliği testine dönüşür.

Şekil 1.2'de gösterilen model, yukarıdaki matematiksel ifadeler cinsinden Şekil 1.3'te sunulmuştur:



Şekil 1.3. Ölçme modelinin matematiksel gösterimi

Ölçüm hatalarının sıfır ortalamaya sahip olduğu varsayımı altında, Eşitlik 1.1’de yer alan x ’in beklenen değeri aşağıdaki gibidir:

$$E(x) = \Lambda_x \xi \quad (Eşitlik 1.3)$$

Ayrıca, ölçüm hatalarının birbirinden ve faktörlerden bağımsız olduğu varsayımı altında gözlenen değişkenler arasındaki varyans-kovaryans matrisi (Σ_x) aşağıdaki gibi ifade edilir:

$$\Sigma_x = \Lambda_x \Phi \Lambda_x' + \Theta_\delta \quad (Eşitlik 1.4)$$

Λ_x' ; Λ_x ’in devrik dönüşümü, Φ ; faktör değişkenlerinin (ξ) varyans-kovaryans matrisi, Θ_δ ; ölçüm hata varyanslarının köşegen matrisidir. Eşitlik 1.4’ün sol tarafındaki ifade, verilen bir kitle için varyans-kovaryans matrisini ifade eder. DFA uygulamalarında, faktör yükleri matrisinin (Λ_x) elemanları üzerine bazı kısıtlılıklar getirilir, bu kısıtlılıklar olmaksızın elde edilen eşitlik, AFA için bir istatistiksel

modeldir. Birden fazla kitle olduğunda, her bir kitle için birer tane Λ_x , Φ ve Θ_δ olacaktır (Raju ve ark., 2002).

DFA’da parametre kestirimlerinde en sık kullanılan yöntem, en çok olabilirlik yöntemidir (maximum likelihood method, ML). Ancak verilerin normal dağılım göstermediği veya kategorik değişkenlerin olduğu durumda en çok olabilirlik yönteminden başka bir yöntemin seçilmesi gerekmektedir. Bu yöntemler, ağırlıklandırılmış en küçük kareler (Weighted Least Squares, WLS), robust ağırlıklandırılmış en küçük kareler (Robust Weighted Least Squares, WLSMV) veya ağırlıklandırılmamış en küçük kareler (Unweighted Least Squares, ULS) yöntemleridir.

DFA’da model uyumunu değerlendirmede uyum istatistiklerinden yararlanılır. Uyum istatistikleri, modelin kabul edilip edilemeyeceğine ilişkin sınır değerler kullanılarak yorumlanır. En yaygın kullanılan uyum iyiliği istatistikleri, uyum iyiliği indeksi (Uİİ) [*Goodness of Fit Index-(GFI)*], düzeltilmiş uyum iyiliği indeksi (DUİİ) [*Adjusted Goodness of Fit Index (AGFI)*], karşılaştırmalı uyum indeksi (KUİ) [*Comparative Fit Index (CFI)*], Tucker-Lewis indeksi (TLİ) [*Tucker-Lewis Index (TLI)*], Yaklaşımın Hata Kareler Ortalaması Karekökü (YHKOK) [*Root Mean Square Error of Approximation (RMSEA)*], Artık Kareler Ortalaması Karekökü (AKOK) [*Root Mean Square Residual (RMSR)*] ve Ağırlıklandırılmış Ortalama Artıkların Karekökü (AOAK) [*Weighted Root Mean Residuals (WRMR)*]’dir. Bunların ilk dördünün değerlerinin 0.90’dan büyük olması, kabul edilebilir bir uyum, 0.95’den büyük olmaları ise iyi bir uyum olduğunun göstergesi olarak kabul edilir. YHKOK, AKOK ve AOAK’da ise söz konusu değerlerin 0.05’in altında olması iyi bir uyum, 0.08’in altında olması ise kabul edilebilir bir uyum olduğunu gösterir (Öztuna, 2008).

Model uyumunu değerlendirmede kullanılan bir diğer ölçü ki-kare testidir. Gözlenen ve kestirilen varyans-kovaryans matrisi arasında farklılık olup olmadığını test eder. İstatistiksel olarak anlamlı olmayan bir ki-kare değeri, her iki matrisin benzer

olduğunu ve mükemmel bir uyum olduğunu göstermektedir. Ki-kare değeri örneklem büyüklüğüne karşı hassas olduğu için, bu duruma çözüm olarak ki-kare değerinin serbestlik derecesine oranı bir kriter olarak ele alınmaktadır. χ^2/df oranı 5'ten küçük oldukça ve '0' a yaklaştıkça iyi uyumun göstergesi olarak yorumlanabilmektedir. Ancak örneklem büyüdükçe serbestlik derecesi de büyüyeceğinden bu oran artacaktır. Böyle durumlarda birden fazla uyum ölçütü kullanmanın geçerli sonuçlar vereceği belirtilmektedir (Uyar, 2011).

1.3.2 Çoklu Grup Doğrulayıcı Faktör Analizi

ÇG-DFA, ölçme modelleri söz konusu olduğunda, bir ölçeğe ait ölçme modelinin, bir başka deyişle faktör yapısının, birden fazla grupta aynı olup olmadığını test etmede kullanılır ve bu analizlere 'kovaryans matrislerinin eşitliği' testi de denir (Öğretmen, 2006). Bahsedilen bu gruplara örnek olarak "cinsiyet", "eğitim durumu" ve "gelir düzeyi" gibi demografik özellikler verilebileceği gibi, "hastalık grupları", "hastalık süresi" gibi tanısal özellikler ya da kültürlerarası geçerliliğin değerlendirildiği bir çalışmada "ülke", "coğrafi bölge" gibi özellikler verilebilir.

Bir ölçme modelinin farklı gruplarda karşılaştırılmasından önce her grupta geçerli bir modelin belirlenmesi gerekir. Eğer modelin tümüne ilişkin bir karşılaştırma yapılır ve sonuçta her grupta modeldeki parametrelerin aynı değerlere sahip olduğu kanıtlanırsa modelin gruplarda değişmez olduğuna ilişkin bir kanıt elde edilir (Şimşek, 2007).

ÇG-DFA kapsamında ölçme eşdeğerliği, aşamalı 4 hiyerarşik modelin test edilmesini içermektedir.

a- Yapısal değişmezlik (configural invariance)

Ölçme eşdeğerliğinin en temel düzeyi olan yapısal değişmezlikte, ölçme modelinin gruplar arasında aynı olup olmadığı test edilir. Yapısal

değişmezliğe ilişkin kanıt elde edilmesi, ölçme aracındaki maddelerin gruplar arasında aynı yapıyı ölçtüğü anlamına gelir. Bu aşamada, faktör yükleri, regresyon sabitleri / eşik değerleri ve hata varyanslarının serbest tahmin edilmesine izin verilerek, yalnızca gruplar için faktör sayısı ve yüklenme örüntüsü sınırlandırılmaktadır (Somer ve ark., 2009). Yapısal değişmezlik, diğer değişmezlik modellerinin test edilebilmesi açısından temel (başlangıç) model rolüne sahiptir. Örneğin, her bir faktörde 5'er maddenin yer aldığı 3 faktörlü bir ölçme modeli olsun. Yapısal değişmezlik, bu maddelerin ilgili faktörlere yüklenme durumlarını gösteren ölçme modelinin gruplar bazında farklılık göstermemesi temeline dayanır. Yani faktör 1'in 3.maddesi incelenen grup değişkeninin her bir kategorisi için ilgili faktörde yer almalıdır.

Eşitlik 1.4'te yer alan her bir matrisin eşitliğinin test edilebilmesi için, yapısal değişmezliğin sağlanması gerekir.

b- Zayıf / metrik değişmezlik (weak / metric invariance)

Zayıf değişmezlik modelinde, faktör yüklerinin gruplar arasında değişmez olup olmadığı test edilmektedir. Faktör yükleri eşit olduğunda, örtük değişkenlerin ölçüm birimi benzerdir. Zayıf değişmezlik test edilirken faktör sayısı ve yüklenme örüntüsü ile birlikte faktör yükleri de sınırlandırılmaktadır (Somer ve ark., 2009). İki grubun olduğu bir örnek için, zayıf değişmezlik test edilirken gruplardan birisi (örn: grup 2) için model eşitliklerinden faktör yükleri silinerek diğer grubun (örn: grup 1) parametre değerlerine sabitlenir.

Metrik değişmezlik, Eşitlik 1.4'te Λ_x matrisinin gruplar arasında eşitliğinin testini ifade eder. Metrik değişmezlik sağlandığında, Eşitlik 1.3'de tanımlanan beklenen değer de gruplarda benzer olacaktır. Bu durumda, benzer faktör skorları matrisine (ξ) sahip farklı gruptan iki birey, aynı beklenen ham puan vektörüne sahip olacaktır.

c- Güçlü / skalar değişmezlik (strong / scalar invariance)

Güçlü değişmezlik modelinde, grupların faktör puanı sıfır olduğunda elde edilen regresyon sabitinin eşit olup olmadığı test edilir. Eğer gözlenen değişken kategorik bir değişken ise, regresyon sabiti yerine eşik değerlerinin eşitliği test edilir. Güçlü değişmezlik test edilirken, faktör örüntüsü ve faktör yüklerine ek olarak regresyon sabiti / eşik değerleri de sınırlandırılmaktadır. (Somer ve ark., 2009). İki grubun olduğu bir örnek için, güçlü değişmezlik test edilirken gruplardan birisi (örn: grup 2) için model eşitliklerinden faktör yükleri ve regresyon sabitleri / eşik değerleri silinerek diğer grubun (örn: grup 1) parametre değerlerine sabitlenir. Hem faktör yükleri hem de sabit değerler / eşik değerleri birbirine eşit olduğunda, farklı gruplardan elde edilen puanlar hem aynı ölçüm birimine hem de aynı orjine sahip olurlar. Böylece, faktör ortalamaları gruplar arasında karşılaştırılabilir olur (Chen, 2008).

d- Katı değişmezlik (strict değişmezlik)

Faktör yüklerinin ve regresyon sabitlerinin /eşik değerlerinin yanında, gruplar arasında hata varyanslarının eşitliğini test eder. Hata varyansları gözlenen değişkenler arasındaki korelasyonların büyüklüğünü etkilemektedir ve katı değişmezlik gözlenen değişkenler arasındaki korelasyonların gruplar arasında karşılaştırılabilmesi için gereklidir (Uyar, 2011). İki grubun olduğu bir örnek için, katı değişmezlik test edilirken gruplardan birisi (örn: grup 2) için model eşitliklerinden faktör yükleri ve regresyon sabitlerine / eşik değerlerine ek olarak hata varyansları da silinerek diğer grubun (örn: grup 1) parametre değerlerine sabitlenir.

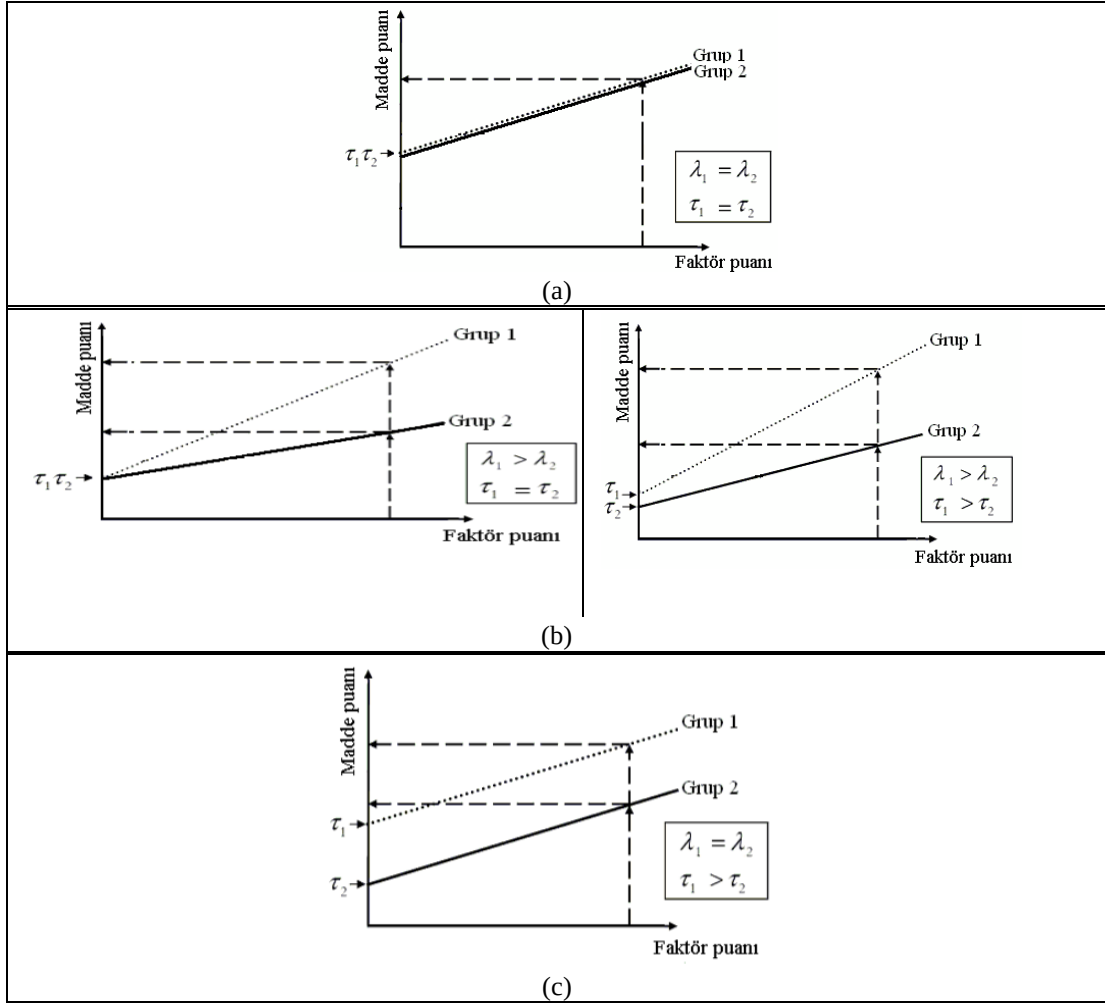
Bu değişmezlik, Eşitlik 1.4'te yer alan Θ_{δ} matrislerinin eşitliğini test eder. Bu matrislerin gruplar arasında benzer olup olmadığının incelenmesi, ölçüm araçlarının güvenilirliği açısından uygulayıcılara önemli bilgiler sağlar (Raju et.al., 2002).

Farklı koşullar altında ölçme eşdeğerliğinin grafiksel gösterimleri Şekil 1.4'te sunulmuştur. Şekil 1.4 (a)'da, hem faktör yüklerinin hem de regresyon sabitlerinin iki grup için benzer olduğu durum örneklenmiştir. Buna göre, herhangi bir faktör puanı için, madde puanlarının her iki grupta da aynı olması, regresyon sabitlerinin ve faktör yüklerinin eşit olmasından kaynaklanmaktadır. Şekil 1.4 (b)'de faktör yükleri eşitsizliğinin ölçme değişmezliği üzerindeki etkisi gösterilmiştir. Faktör yüklerinin gruplar arasında eşitsizliği, maddelerin faktör puanlarına farklı ölçme birimleriyle kalibre edildiğini gösterir. Madde puanındaki bir birimlik değişme grupların faktör puanlarındaki aynı miktar değişmeye karşılık gelmezse, eğimler aynı olmayacağından gruplar için regresyon doğruları farklı olur. Şekil 1.4 (b)'de θ 'nın ne seviyede eksik tahmin edildiği θ düzeyine göre değişmektedir. Yüksek yetenek seviyeleri için daha fazla tahmin hatası yapılmaktadır. Bu durum **tek biçimli olmayan yanlılık** olarak adlandırılmaktadır. Şekil 1.4 (c)'de, gruplar arasında faktör yükleri aynı iken regresyon sabitlerinin farklı olduğu görülmektedir. Şekle göre aynı faktör puanına sahip bireylerden grup 1'de olan bireyler grup 2'de olan bireylerden daha yüksek puan elde edeceklerdir. Regresyon sabitleri farklı olduğundan **tek biçimli yanlılıktan** söz edilir (Asil, 2010).

Ölçme eşdeğerliği, ölçeğin değişmezliğinin (tüm maddelerin eş zamanlı olarak tümel bir test ile değerlendirilmesi) test edilmesinin yanı sıra ölçekte yer alan her bir maddenin değişmezliğini test etmede kullanılabilir. Maddeler, ölçüm değişmezliği standartlarını sağladığı zaman, “değişmez” olarak değerlendirilir, aksi durumda ölçüm yanlılığı yarattıkları söylenir. Ölçme aracında bazı maddeler değişmezlik standartlarını yerine getirirken, bazıları bu standartların dışında kalabilir (Chen, 2008). Pratikte, ölçme aracının gruplar arasında tam değişmezlik (full measurement invariance) standartlarını karşılaması her zaman mümkün olmadığı, bu durumlarda en azından kısmi ölçme değişmezliğine (partial measurement invariance) yer verilmesi gerektiği belirtilmektedir.

Ölçme eşdeğerliğinin herhangi bir tür için bozulumu söz konusu olduğunda, iki yaklaşım vardır. İlki, en son model uyumunun sağlandığı durumun raporlanmasıdır.

İkincisi ise, model uyumunun bozulmasına neden olan madde(ler) belirlenmeye çalışılarak, modelde bu maddelere ilişkin parametreler serbest bırakılarak model uyumu iyileştirmeye çalışılır. Örneğin zayıf değişmezliğin test edilmesi durumunda, yapısal değişmezlik modeli ile kıyaslandığında, modelde kötüleşme olduğu belirlenirse, ilk yaklaşım kapsamında; “incelenen ölçme aracı için yapısal değişmezlik sağlanmıştır” şeklinde bir raporlama yapılarak süreç tamamlanır. İkinci yaklaşımda ise, model uyumunun bozulmasına neden olan madde(ler) belirlenmeye çalışılır. Bu kapsamda program tarafından maddelerin faktör yüklerinin serbest bırakılmasına ilişkin önerilen düzeltme indisleri içerisinden hangi madde ya da maddelerin faktör yükleri serbest bırakıldığında modelde ne miktarda bir iyileşme yarattıkları değerlendirilerek, sırasıyla bu düzeltme işlemleri gerçekleştirilir. Serbest bırakılan faktör yükleri hiyerarşik modellerle test edilerek, model uyumunda iyileşme yaratılmaya çalışılır. Model uyumunun sağlandığı (modelde kötüleşme olmadığı duruma kadar) faktör yüklerinin serbest bırakılmasına izin verilir. En son aşamada ise, faktör yükleri serbest bırakılan maddeler haricinde incelenen ölçme aracı için zayıf değişmezlik sağlanmıştır ya da serbest bırakılan ve bırakılmayan tüm maddeler üzerinden kısmi zayıf değişmezlik sağlanmıştır biçiminde raporlama yapılır. Bu aşamada faktör yükleri serbest bırakılan maddelerle beraber, sonraki değişmezlik türlerinin test edilmesine aşamasına geçilebilir.



Şekil 1.4. Farklı ölçme eşdeğerliği gösterimleri (a) Güçlü değişmezliğin sağlanması durumu, (b) Zayıf değişmezliğin sağlanmaması durumu, (c) Güçlü değişmezliğin sağlanmaması durumu

ÇG-DFA, sağlık bilimlerinin yanı sıra, eğitim ve sosyal bilimler alanlarında yapılan çalışmalarda da uygulama alanı bulmuştur.

Munyombwe ve ark. (2015) tarafından yapılan bir çalışmada, farklı çalışmalardan elde edilen veri bir araya getirilerek, tümel bir veri analizinin gerçekleştirilmesi için, ölçülen özelliğin çalışmalar üzerinden aynı anlama gelmesi gerekliliği vurgulanmıştır. Bu kapsamda, ilgili çalışmanın amacı Genel Sağlık Anketinin, iki inme çalışmasında ölçüm değişmezliğinin araştırılmasıdır. Ölçüm değişmezliğinin yapısal, zayıf ve güçlü değişmezlik kapsamında değerlendirildiği ve tüm değişmezlik türlerinin doğrulandığı çalışmanın bulguları, iki farklı veri seti üzerinden tümel analizlerin yapılabileceğini ortaya koymuştur (Munyombwe et.al, 2015).

2014 yılında Andre S. tarafından yapılan bir çalışmada, politik doğruluğun (politik doğruluk, Avrupa’da göçmenlerin politik entegrasyonunun bir ölçüsü olarak görülmektedir) değerlendirilmesinde bu özelliğin çoklu boyutları olduğu ve YEM kapsamında modellenebileceği düşüncesinden hareketle mevcut dört ölçme değişmezliği türüne ilişkin modellerin kültürler arasında eşitliği ÇG-DFA ile test edilmiştir. Elde edilen sonuçlarda, politik doğruluğun ülkeler arasında karşılaştırılabileceği, ancak Avrupa Birliği ülkeleri arasında karşılaştırma yaparken dikkatli olunması gerektiği ortaya konmuştur. Bu durumun nedeninin zayıf değişmezliğin kısmi olarak sağlanmasından kaynaklı olduğu belirtilmiştir (Andre, 2014).

Costa ve ark. 2014 yılında yaptıkları bir çalışmada, EORTC QLQ-C30 ölçeğinin ölçme değişmezliğini, yapısal, metrik, sayısal kapsamda kanser türlerinde, ÇG-DFA ile değerlendirmişlerdir. Prostat kanseri grubu haricinde, tüm gruplarda model uyumunun yeterli bulunduğu çalışmada, meme kanseri grubunda tek bir madde için sayısal değişmezliğin sağlanmadığı ve bu duruma neden olan faktörün tedavi durumu (başlangıç, tedavi sırasında ve tedavi sonunda) değişkeni olduğu sonucuna ulaşılmıştır (Costa et.al, 2014).

Yan ve ark. (2014), sağlık durumunu değerlendiren ölçme araçlarında ölçüm değişmezliğinin bir uygulamasını yapmak ve önemine ilişkin farkındalığı arttırmak için yaptıkları çalışmada, “Astım Kontrol Anketi”nin yaş ve cinsiyet değişkenleri bazında değişmezliğini ÇG-DFA ile incelemişlerdir (Yan et.al, 2014).

Asil (2010) tarafından yapılan doktora tezinde, Uluslararası Öğrenci Değerlendirme Programı 2006 kapsamında uygulanan öğrenci anketinin kültürler ve diller arası eşdeğerliği, Avustralya, Yeni Zelanda, Amerika Birleşik Devletleri ve Türkiye örneklemeleri üzerinde karşılaştırmalı olarak incelenmiştir. Araştırmanın birinci aşamasında, öğrenci anketinin faktör yapısı DFA ile incelenmiştir. İkinci aşamada ise, Ortalama ve Kovaryans Yapıları metoduna dayanan ÇG-DFA kullanılarak anket maddelerinin farklı kültürler ve diller arasında Değişen Madde Fonksiyonu (DMF)

gösterip göstermediği araştırılmıştır. Son aşamada, uzman görüşlerine başvurularak, maddelerin DMF göstermesinin nedenleri belirlenmeye çalışılmıştır. DFA sonuçları ölçme modelinin tüm örneklemelerde aynı faktör yapısına sahip olduğunu göstermiştir. ÇG-DFA bulguları ülkeler arasında değişen fonksiyon gösteren maddelerin olduğunu ortaya koymuştur. Ülkeler arasında dil ve kültürel farklılıklar arttıkça DMF gösteren maddelerin sayısının da arttığı gözlenmiştir. Maddelerin DMF göstermesinin asıl nedenlerinin çeviri problemleri ve kültürel farklılıklar olduğu sonucuna varılmıştır (Asil, 2010).

1.3.3 Ölçme Eşdeğerliğinin Sağlanamama Nedenleri ve Uygulanacak Stratejiler

Ölçek puanları gruplar arasında karşılaştırıldığında, yapıların eşitliğini yani ölçme eşdeğerliğini etkileyen birçok neden olabilir. Bu nedenler ölçüm değişmezliği türlerine göre farklılık gösterir.

Yapısal değişmezliğin sağlanamaması, ölçme aracı ile ölçülmesi hedeflenen yapının, bir kültürden diğer bir kültüre aktarılması sırasında ortaya çıkabilir. Çünkü incelenen yapı, bir kültürden diğerine farklılık gösterebilir.

Zayıf değişmezliğin bozulumu farklı nedenlerle olabilir. İlki, bir ölçek bir kültürden diğer kültüre uyarlandığında ortaya çıkar, çünkü incelenen yapının tanımlamaları ve anlamları kültürler arasında tam olarak örtüşmeyebilir. Sonuç olarak, madde içeriği bir kültürde diğerine kıyasla daha uygun olabilir. İkinci neden, çeviri hatalarıdır. Maddeler bir kültürden diğerine çevrilirken, anlamları değişebilir. Üçüncü neden, özellikle uç değerlerin kullanılması ya da göz ardı edilme stratejileri kapsamında yanıt setlerinin farklılaşmasıdır.

Güçlü değişmezliğin bozulumunun da farklı nedenleri vardır. İlki, sosyal farklılıklar nedeniyle bir gruptaki bireylerin diğer gruptakilere göre tutarlı bir şekilde daha

yüksek ya da daha düşük değerlendirme yapmasıdır. İkinci neden, bir grubun kendi yetersizliklerinden duydukları endişe nedeniyle, θ değerlendirmelerinde beklenenden daha farklı bir yanıt stratejisi geliştirmeleridir. Üçüncü neden, farklı kültür gruplarında yer alan bireylerin kendileri ile ilgili değerlendirme yapma süreçlerinde referans gördükleri gruplara göre gerçeği yansıtmayan yanıtlara sahip olabilmeleridir (Chen, 2008).

Ölçüm değişmezliğinin bozulumu durumunda uygulanması gereken yaklaşımlar, henüz netleşmemiş olup, farklı görüşler bulunmaktadır. Uygulamada yaygın olarak kullanılan 4 farklı yaklaşım bulunmaktadır.

İlk yaklaşım, ölçüm değişmezliğini bozan maddelerin çıkartılmasıdır. Bu yaklaşımda, farklı gruplar için ölçeğin farklı versiyonları ortaya çıkar. Aynı zamanda, incelenen yapının kapsamının tam sağlanamamasına neden olur (kapsam geçerliliği). **İkinci yaklaşım**, ölçüm değişmezliği gösteren maddelerin tümü ölçekte yer alır ve toplam/ortalama skorlar değişmezlik özelliği gösteren ve göstermeyen maddelerden elde edilir. Bu yaklaşımın varsayımı, değişmezlik özelliğini sağlamayan maddelerin grup karşılaştırmalarında düşük yanlılığa neden olacağıdır. Ancak bu varsayım, güvenilir değildir. **Üçüncü yaklaşım**, kısmi ölçüm değişmezliğidir. Bu yaklaşım değişmez maddelerin gruplar arasında eşitliği kısıtlamasına sahipken; değişmezlik özelliğini bozan maddelerin de gruplarda farklı olmasına izin verir. Değişmezlik özelliğini bozan maddeler üzerinde herhangi bir kısıtlama olmadığı için, toplam/ortalama skor yaklaşımına göre daha düşük istatistiksel yanlılığa neden olur. **Dördüncü yaklaşım**, ölçüm değişmezliği sağlanmadığı durumda grup karşılaştırması yapılmamasıdır (Chen, 2008).

1.4 Rasch Analizi

Sağlık alanında hastaların θ düzeylerinin belirlenmesi için geliştirilen ölçme araçları, geçerli değişim skorlarının ve parametrik istatistiklerin elde edilmesinin gerekli olduğu klinik denemelerde yaygın olarak kullanılır. Rasch analizi, veri seti Rasch modelinin beklentilerine uyum sağladığı durumda sıralı ölçekle elde edilmiş yanıtların, aralık değişkenine dönüşümünü sağlayarak, bireylerin doğru bir şekilde değerlendirilmesine yardımcı olur (Öztuna, 2008).

Rasch modeli (İkili Rasch modeli), Georg Rasch tarafından 1960 yılında, iki sonuçlu (evet/hayır, var/yok, katılıyorum/katılmıyorum vb.) maddeler için geliştirilmiş olup; bu modelde bireyin bir maddeye doğru yanıt verme olasılığı, bireyin θ düzeyi ile maddenin zorluğu arasındaki farkın lojistik fonksiyonu olarak tanımlanmıştır. İkili Rasch Modeline kısaca Rasch Model de denilmektedir (RM). İki sonuçlu maddeler, tek basamaklı maddeler olarak da adlandırılmaktadır. Eğer bireyler bu tek basamağı başarı ile gerçekleştirirlerse 1 puan, gerçekleştiremezlerse 0 puan almaktadırlar. Bu modelde madde zorluk parametresinin (β) yanı sıra bireyin θ düzeyi de kestirilmektedir. İncelenen özellik düzeyi θ_i ve madde zorluk düzeyi β_j olmak üzere i. bireyin j.inci maddeye x yanıtı verme olasılığı aşağıdaki gibi hesaplanmaktadır (Yüksel, 2012):

$$P(X=x_{ij} / \theta_i) = P_{ij} = \frac{\exp[x_{ij}(\theta_i - \beta_j)]}{1 + \exp(\theta_i - \beta_j)}, x=0,1 \quad (\text{Eşitlik 1.5})$$

RM, ilk olarak iki sonuçlu maddeler için geliştirilmiş olsa da, çok sonuçlu maddeler için de kullanılabilecek modellere sahiptir. Bunlardan ilki kısmi kredi modeli (KKM), ikincisi sıralı sonuçlu model (SSM)'dir. SSM ve KKM arasındaki en temel farklılık, SSM'de aynı yanıt formatının ölçme aracındaki tüm maddeler için geçerli olduğunun ve eşik değerleri arasındaki farklılığın tüm maddeler için eşit olduğunun varsayılmasıdır. Örneğin, birinci ve ikinci kategoriye birbirinden ayıran eşik değeri ile ikinci ve üçüncü kategoriye birbirinden ayıran eşik değeri arasındaki uzaklık, tüm maddeler için eşittir.

Rasch modelinin önemli bir özelliği, gözlenen ham puanların, doğrusal olmayan bir dönüşüm kullanarak θ puanlarının tahmin edilmesi için yeterli bir istatistik olmasıdır. Madde yanıt teorisinin diğer modellerinde, ham puan yeterli istatistik değildir (Öztuna, 2008).

Rasch analizi farklı durumlarda kullanım alanına sahiptir. Bunlar; yeni bir ölçeğin geliştirilmesi, var olan bir ölçme aracının psikometrik özelliklerinin değerlendirilmesi, bilgisayar uyarlamalı test için soru bankalarının oluşturulmasıdır. Bu alanların her biri için aşağıdaki standart protokolün yerine getirilmesi gerekmektedir:

1. Tek boyutluluk için ölçme aracının iç tutarlılığının test edilmesi
2. Ölçme aracındaki maddelerin modele uyumunun incelenmesi
3. Maddelerin değişmezlik özelliğini sağlayıp sağlamadığının test edilmesi
4. Maddelere ilişkin eşik değerlerinin sıralı olup olmadığının test edilmesi
5. MİF'in test edilmesi

Standart protokol kapsamında ilk 4 aşamanın değerlendirildiği ve psikometrik özellikleri gösterilmiş ölçme araçlarında, ölçme eşdeğerliği MİF kapsamında değerlendirilir. Rasch analizi kapsamında MİF gösteren maddelerin belirlenmesi için kullanılan en yaygın yaklaşım RM'de bulunan artıkların kullanılmasıdır. Ölçeği yanıtlayan n . kişi, i . iki sonuçlu maddeye yanıt vermiş olsun. Bu yanıt için RM'den bulunan beklenen değer ve varyans da sırasıyla $E[y_{ni}]$ ve $V[y_{ni}]$ olsun. Bu yanıtla ilişkin standartlaştırılmış artık değeri aşağıdaki gibi olacaktır:

$$Z_{ni} = \frac{y_{ni} - E[y_{ni}]}{\sqrt{V[y_{ni}]}} \quad (\text{Eşitlik 1.6})$$

Standartlaştırılmış artıklar bireylerin maddelere verdikleri bütün yanıtlar için hesaplanır. Daha sonra bireyler θ kestirimlerine göre, her sınıfta en az 50 birey olacak şekilde, G tane sınıf aralığına (class interval) bölünür. Her madde için standartlaştırılmış artık değerleri bağımlı değişken, bulunan G tane sınıf aralığı ve

grup faktörü bağımsız değişkenler olarak alınıp, iki yönlü varyans analizi gerçekleştirilir. Varyans analizi sonucunda iki çeşit MİF bulmak mümkündür. Gruplar, bir maddeye verilen yanıtlar bakımından tutarlı sistematik bir farklılık gösteriyorsa “tekdüze (uniform) MİF”, göstermiyorsa “tekdüze olmayan (non uniform) MİF” den bahsedilir. Diğer bir ifadeyle, ölçme aracındaki herhangi bir madde için bir grup, θ düzeyi boyunca diğerinden anlamlı derecede daha yüksek/düşük skora sahip ise tekdüze MİF olduğu söylenir. Diğer yandan, ölçme aracındaki herhangi bir madde için θ ’nın belli bir değerine kadar bir grup, o değerden sonra da diğer grup anlamlı derecede daha yüksek/düşük skorlara sahip ise tekdüze olmayan MİF’ den bahsedilir.

Tekdüze MİF yalnızca grup faktörünün varyans analizi sonucunda anlamlı bulunması, tekdüze olmayan MİF ise varyans analizi sonucunda grup faktörü ve θ düzeyi sınıfları arasındaki etkileşim teriminin anlamlı bulunması ile ortaya konabilir (Erdoğan, 2012). Tekdüze MİF olduğunda, her bir grup için madde parametreleri ayrı ayrı kalibre edilir. Tekdüze olmayan MİF için genellikle sorunlu madde ölçme aracından çıkarılır.

Gökçe ve ark. (2011) tarafından yapılan bir çalışmanın amacı, öğrencileri sınıf ortamında motive eden iç ve dış motivasyon faktörlerinin belirlenmesinde kullanılan Harter tarafından geliştirilen ölçeğin, Türkiye’de ilköğretim 3-8. sınıf düzeyindeki öğrencilere adaptasyonunun yapılması, geçerlilik ve güvenirliğinin değerlendirilmesi ve okul türü, sınıf düzeyi, okulun maddi statüsü, cinsiyet, anne eğitim düzeyi, baba eğitim düzeyi ve ailenin maddi durumu bakımından MİF olup olmadığının analiz edilmesidir. Bu çalışma kapsamında iç motivasyon bileşenindeki bir maddenin okul türü, sınıf düzeyi, okulun maddi statüsü ve ailenin ekonomik durumu bakımından; dış motivasyon bileşeninde yer alan bir başka maddenin ise okul türü ve sınıf düzeyine göre farklı işleve sahip olduğu belirlenmiştir.

Elhan ve ark. (2008) tarafından yapılan çalışmada, bel ağrılı hastaların özürüllüğün değerlendirilmesinde kullanılmak üzere geliştirilen ölçeklerde yer alan maddeler yaş

ve cinsiyet değişkenleri için MİF gösterip göstermedikleri açısından değerlendirilmiştir. “Vücut fonksiyonları” alt ölçeğinde yer alan maddeler yaş ve cinsiyet değişkenleri bakımından MİF göstermezken; “aktivite ve katılım” alt ölçeğinde yer alan bir madde yaş açısından tekdüze MİF göstermiş; yaşlı hastaların genç hastalara göre bu maddede daha fazla zorluk çektikleri belirlenmiştir.

1.5 ÇG-DFA ile Rasch Analizi Yaklaşımlarının Karşılaştırılması

Ölçme eşdeğerliğinin incelenmesinde en yaygın kullanılan ve çalışma kapsamında teorik alt yapısı detaylı olarak incelenen ÇG-DFA ve Rasch Analizi işlevsel olarak benzerlik ve farklılıklara sahiptir. Bu kapsamda farklı çalışmalarda bu iki yaklaşımın karşılaştırılmasına ilişkin sonuçlar mevcuttur.

Gruplar arası ölçme eşdeğerliğinin gösterilmesinde ve yapısal eşdeğerliğin incelenmesinde ÇG-DFA uygun ve önerilen bir metot olup çok boyutlu ölçeklere uygulanabilmesi, çok kategorili verilerin analizinde kullanılabilmesi ve hem tekdüze hem de tekdüze olmayan değişen madde fonksiyonu (DMF)’yi belirleyebilmesi gibi birçok yönden diğer metotlara tercih edilmektedir. Bu yöntemle ortak ve kültüre özgü maddeler tespit edilebilmektedir DMF analizlerinde ÇG-DFA’nın MYT yaklaşımı ile aynı etkililiğe sahip olduğu hatta verinin çok kategorili olduğu durumlarda daha iyi performans gösterdiği simülasyon çalışmalarıyla kanıtlanmıştır (Asil, 2010).

Raju ve ark. (2002) tarafından gerçekleştirilen çalışmada, ölçme eşdeğerliği açısından DFA ve MYT perspektifi arasındaki benzerlikler aşağıdaki gibi özetlenmiştir:

- ≈ Her iki yaklaşım da, gizli bir değişken ile bir grup gözlenen değişken (maddeler) arasındaki ilişkiyi değerlendirir.

- ≈ Her ikisi de, gizli değişken üzerinde aynı θ değerine sahip farklı populasyonlarda yer alan bireylerin madde/alt ölçek düzeyinde gerçek skorlarının ne derece benzer olduğunu inceler. Bu benzerlik, iki yaklaşım arasındaki en önemli benzerliktir ve klasik test teorisi (KTT)'deki paralellik tanımına kavramsal olarak benzerdir. KTT'de, paralellik koşulu, iki testte gerçek skorların eşitliğini gerektirir, bu yaklaşımlarda ise θ sabit olduğunda, iki populasyon için gerçek skorların eşitliği incelenir.
- ≈ Ölçüm değişmezliğinin DFA ve MYT tanımları, iki populasyon için θ düzeyinde skorların dağılımının benzerliği anlamına gelmez. Gerçekte, gizli değişken dağılımları, farklıdır. Dağılımlar arası bu ortalama farklılık, yaygın olarak “etki (impact)” terimine karşılık gelir.
- ≈ Ölçüm değişmezliği sağlanmadığında, her iki yaklaşım da problemin düzeyini ve kaynağını bulmak amacıyla kullanılabilir. MYT kapsamında, her bir madde MİF açısından değerlendirilir ve bu değerlendirme MYT perspektifinde klasik MİF incelemesidir. Ancak DFA kapsamında, değişmezliğin sağlanmama nedenlerinin incelenmesinde önce, önerilen model her bir grup için uyum iyiliği testleri ile değerlendirilir.
- ≈ Ölçüm değişmezliğinin sağlanamaması durumunda, madde yanıt fonksiyonları, hem MYT hem de DFA kapsamlarında çok değerli bilgi sağlar. Bu grafikler, ölçüm değişmezliğinin sağlanamama kapsamı ve konumu hakkında, uygulayıcılara önemli yarar sağlar.

Aynı makale kapsamında DFA ve MYT perspektifi arasındaki farklılıklar da aşağıdaki gibi özetlenmiştir:

- ≠ Gizli değişken ile (madde/alt ölçek düzeyinde) gerçek skor arasındaki ilişki, DFA kapsamında doğrusal iken; MYT perspektifinde doğrusal değildir.

- ≠ Gözlenen değişkenler iki sonuçlu olduğunda, lojistik regresyon modeli, gizli değişken ile gözlenen değişken arasındaki ilişkiyi açıklamada doğrusal regresyon modelinden daha uygundur. Bu kapsamda, ölçüm değişmezliği için MYT temelli yaklaşımlar, DFA yaklaşımına tercih edilir. Bununla beraber, bir madde ya da ölçüm değişkeni için olası skor sayısı arttıkça, DFA kapsamında kullanılan doğrusal model eşit derecede kabul edilebilir olacaktır.
- ≠ DFA yaklaşımı, eş zamanlı olarak çok sayıda gizli değişken ve çoklu grupların olduğu durumların üstesinden gelmede daha avantajlıdır. Diğer taraftan, MYT temelli MİF ya da ölçüm değişmezliği yaklaşımının büyük kısmı, tek boyutlu ölçekler ile sınırlandırılmıştır.
- ≠ DFA kapsamında ölçüm değişmezliğinin en katı formunda, madde hata varyanslarının, popülasyonlarda eşit olması gerekir. Jöreskog, madde hata varyanslarının eşitliğinin test edilmesini önerse de, bu gerekliliğin uygulamada karşılanması zordur. MYT perspektifinde, madde düzeyinde varyansa ilişkin açık bir tartışma bulunmamaktadır, çünkü θ 'ya bağlıdır. MYT'nde önemli bir kavram, θ ile ilişkili olan ölçümün standart hatasıdır. Bu değer, θ 'nın bir fonksiyonu olduğu için, bireyden bireye farklılık gösterebilir. Bu özellik, MYT'nin, KTT'ne önemli bir üstünlüğüdür.
- ≠ MYT'de; seçilen θ değeri için bir bireyin bir maddenin herhangi bir kategorisini seçme olasılığı bilinir. Bu olasılık değerleri, DFA kapsamında mevcut değildir.

2. GEREÇ VE YÖNTEM

Bu tez çalışmasının amacı; ölçme eşdeğerliğinin, YEM kapsamında “ÇG-DFA” ve MYT yaklaşımı kapsamında ise “Rasch Analizi” ile incelenmesidir.

Bu amaçla ilk olarak uygulamalarda kullanılacak ölçme aracı belirlenmiştir. Uygulama kapsamında sağlık alanında kullanılan ve geçerli olduğu gösterilen ölçme aracı olarak Modifiye Sağlık Değerlendirme Anketi (MSDA) seçilmiştir.

MSDA, Pincus ve arkadaşları tarafından orijinal Sağlık Değerlendirme Anketi (SDA)’nden geliştirilmiş bir ölçektir. SDA, 8 bölüm (giyinip-kuşanma, doğrulma, yemek yeme, yürüme, hijyen, uzanma, kavrama ve günlük işler) şeklinde 20 maddeden oluşan, günlük yaşam aktivitelerini değerlendiren bir fiziksel özrürlük ölçeğidir. Her madde 0–3 arası (0: hiç zorluk çekmeden yapıyorum, 3: hiç yapamıyorum) puanlanmaktadır (Küçükdeveci ve ark, 2004). MSDA ise, 20 maddeli SDA’da yer alan her bir bölümden birer sorunun seçilmesi sonucunda oluşturulan 8 maddeli bir ölçektir (Pincus et.al., 1983). MSDA’nın psikometrik özellikleri, romatoid artrit (RA) hastaları üzerinde gerçekleştirilen bir çalışma ile değerlendirilmiş ve geçerli olduğu sonucuna ulaşılmıştır (Wolfe, 2001).

Uygulama verisi olarak, daha önce gerçekleştirilen bir TÜBİTAK projesi kapsamında RA’lı hastalardan elde edilen veriler kullanılmıştır (Romatoid Artritli Hastalarda Bilgisayar Uyarlamalı Test Yöntemi İle Özürlülük Değerlendirmesi. TÜBİTAK 1001 Araştırma Projeleri, 109S342, 2010-2012). Bu çalışmada yer alan toplam 299 hasta, 20 maddeden oluşan SDA’ya yanıt vermişlerdir. Bu maddeler içerisinde MSDA’da yer alan maddeler seçilmiştir. Hastaların 77’si (%26) kadın, 222’si (%74) erkektir. Hastaların yaş ortalaması ($\pm$ standart sapma) 52.3 ± 11.5 (minimum 18, maksimum 82) olup; ortalama RA hastalık süresi ise 11.3 ± 8.0 yıldır.

Ölçme eşdeğerliğinin değerlendirilmesinde grup değişkeni olarak demografik bir özellik olan cinsiyet ve hasta karakteristiği olarak hastalık süresi (≤ 10 yıl, >10 yıl) değişkenleri seçilmiştir.

ÇG-DFA

Ölçme eşdeğerliğinin ÇG-DFA kapsamında incelenmesinde; yapısal değişmezlik, zayıf değişmezlik ve güçlü değişmezlik modelleri test edilmiştir. Güçlü değişmezliğin değerlendirilmesinde, gözlenen değişkenler nitel değişkenler olduğu için regresyon sabiti yerine eşik değerlerinin eşitliği test edilmiştir. Literatürde de belirtildiği gibi katı değişmezliğin sağlanması zorlayıcı bir varsayım içerdiğinden karşılanması her zaman mümkün olmadığı için, çalışma kapsamında değerlendirilmemiştir.

DFA uygulamasında parametre tahmininde, MSDA ölçeğinin kategorileri sıralanabilir nitel (ordinal) düzeyinde olduğu için en çok olabilirlik yöntemi yerine robust ağırlıklandırılmış en küçük kareler yöntemi kullanılmıştır.

Model uyumu ki-kare testi, YHKOK, KUI ve TLİ istatistikleri ile incelenmiştir.

Her bir alt grupta DFA sonuçları değerlendirilirken ve ölçme eşdeğerliği kapsamında yapısal değişmezlik test edilirken; ki-kare/serbestlik derecesi [χ^2/sd] değerinin 3 ve altı olması iyi uyum olarak kabul edilmiştir. Ölçme eşdeğerliği kapsamında modeller arası karşılaştırmalarda program tarafından hesaplanan ve karşılaştırması yapılan modellerden elde edilen ki-kare istatistiklerine dayalı “ki-kare fark” değerlerinin istatistiksel anlamlılıkları değerlendirilmiştir. İstatistiksel olarak anlamlı olmayan “ki-kare fark” değeri, test edilen modelin geçerli olduğunu; diğer bir ifade ile test edilen ölçüm değişmezliği türünün sağlandığını göstermektedir.

Her bir model için YHKOK değeri 0.05'in altında ise iyi bir uyum, 0.08'in altında olması ise kabul edilebilir bir uyum olarak değerlendirilmiştir (Öztuna, 2008). KUI ve TLI indisleri bazında yapılan karşılaştırmalarda da modeller arası farklılık değerlerinin 0.01'in altında olması test edilen modelin geçerli olduğunu ifade etmektedir (Cheung ve Rensvold, 2002).

MSDA'nın ölçme eşdeğerliği test edilirken sırasıyla aşağıdaki adımlar izlenmiştir.

- 1- Temel model olarak grup değişkeninin her bir kategorisi bazında ayrı ayrı DFA yapılmıştır. DFA sonucunda model uyum istatistikleri incelenerek; model ile veri arasındaki uyum değerlendirilmiştir. Uyumun yetersiz olduğu durumda düzeltme indisleri yardımıyla model uyumunda iyileştirmeler yapılmıştır. Düzeltme indisleri içerisinde gözlenen değişkenlere ilişkin hata varyanslarını ilişkilendiren indisler göz önünde bulundurulmuştur. Bu kapsamda önemli bir diğer nokta, grup değişkeninin her bir kategorisi için önerilen ortak düzeltme indisleri içerisinde en yüksek değere sahip olanların, klinik anlamlılıkları da dikkate alınarak modele eklenmesi durumudur.
- 2- Ölçme eşdeğerliğinin incelenmesinde hiyerarşik analizlerin ilk aşaması olarak yapısal değişmezlik test edilmiştir. Yapısal değişmezlik test edilirken faktör sayısı ve gözlenen değişkenlerin örtük değişkene yüklenme örüntüsü sınırlandırılmıştır. Başka bir ifade ile bu aşamada grupların faktör yapılarının eşdeğer olup olmadığı aynı model içerisinde test edilmiştir. Elde edilen modelin uyumu, uyum istatistikleri ile değerlendirilmiştir.
- 3- Yapısal değişmezliğin sağlanması durumunda, analizin ikinci aşaması olan zayıf değişmezlik test edilmiştir. Zayıf değişmezlikte yapısal değişmezliğe ek olarak grupların örtük değişkene ilişkin ölçme biriminin benzer olup olmadığı test edilmiştir. Zayıf değişmezlik test edilirken, faktör sayısı ve yüklenme örüntüsü ile birlikte faktör yükleri de sınırlandırılmıştır. Daha sonra zayıf değişmezlik modelinin χ^2 , KUI ve TLI değerlerinden yapısal değişmezlik

modeli için elde edilen χ^2 , KUI ve TLI değerleri çıkarılarak zayıf değişmezlik modeli için ölçme eşdeğerliği incelenmiştir. Elde edilen $\Delta\chi^2$, ΔKUI ve ΔTLI değerleri modelde anlamlı düzeyde bir kötüleşme olmadığını gösteriyorsa, bu sonuç gruplar arasında zayıf değişmezliğin sağlandığını yani faktör yüklerinin iki grup için eşdeğer olduğunu ifade eder.

- 4- Analizin üçüncü aşaması olan güçlü değişmezlik test edilirken, ilk iki aşamada yapılan sınırlandırmalara ek olarak eşik değerleri de sınırlandırılmıştır. Daha sonra güçlü değişmezlik modelinin χ^2 , KUI ve TLI değerlerinden zayıf değişmezlik modeli için elde edilen χ^2 , KUI ve TLI değerleri çıkarılarak güçlü değişmezlik modeli için ölçme eşdeğerliği incelenmiştir. Elde edilen $\Delta\chi^2$, ΔKUI ve ΔTLI değerleri modelde anlamlı düzeyde bir kötüleşme olmadığını gösteriyorsa, bu sonuç gruplar arasında güçlü değişmezliğin sağlandığına yani faktör yüklerine ek olarak eşik değerlerinin de iki grup için eşdeğer olduğunu gösterir.

ÇG-DFA analizleri için MPlus 6.1 deneme versiyonu (Muthen ve Muthen, 2006) kullanılmıştır.

Rasch Analizi

MSDA, psikometrik özellikleri gösterilen bir ölçek olduğu için Rasch analizi kapsamında sadece MİF açısından değerlendirilmiştir. Rasch analizinde MİF iki yönlü varyans analizi kullanılarak değerlendirilmiş; gruplar, bir maddeye verilen yanıtlar bakımından tutarlı sistematik bir farklılık gösteriyorsa “tekdüze (uniform) MİF”, göstermiyorsa “tekdüze olmayan (non-uniform) MİF” olduğu yorumuna ulaşılmıştır. Ayrıca MİF olması durumunda madde zorluk parametreleri iki grup için ayrı ayrı elde edilecek, MİF varlığı grafiksel olarak da değerlendirilecektir. Rasch Analizi için RUMM 2020 programı kullanılmıştır.

3. BULGULAR

Çalışma kapsamında elde edilen sonuçlar ÇG-DFA ve Rasch analizi olarak iki başlık altında incelenecektir.

3.1 ÇG-DFA Sonuçları

ÇG-DFA sonuçları cinsiyet ve hastalık süresi (≤ 10 yıl, >10 yıl) değişkenleri bazında ayrı ayrı değerlendirilmiştir.

3.1.1 Cinsiyet Bazında ÇG-DFA Sonuçları

Kadın ve erkek hasta gruplarında MSDA'nın ölçme eşdeğerliği test edilirken önce temel model olarak kadın ve erkek hasta grupları için ayrı ayrı DFA yapılmıştır. Kadın ve erkek grupları için yapılan DFA'ya ilişkin model uyum indeksleri Çizelge 3.1'de **Model 1** olarak verilmiştir.

Çizelge 3.1 Cinsiyet bazında her bir grup için DFA'ya ilişkin uyum iyiliği istatistikleri

	<i>Grup</i>	<i>Ki-Kare (χ^2)</i>	<i>sd</i>	<i>χ^2/sd</i>	<i>YHKOK</i>	<i>KUI</i>	<i>TLI</i>
Model 1	Kadın	50.89	20	2.54	0.142	0.985	0.980
	Erkek	92.09	20	4.60	0.127	0.984	0.977
Model 2	Kadın	31.33	18	1.74	0.098	0.994	0.990
	Erkek	50.26	18	2.79	0.090	0.993	0.989

Çizelge 3.1'den görülebileceği gibi, her iki cinsiyet grubu için de ölçme modelinin uyum indeks değerleri, verinin modele orta düzeyde uyum gösterdiğini ifade etmektedir. Bu durumda program tarafından önerilen, maddelerin hata varyanslarının ilişkilendirilmesine yönelik olan düzeltme indisleri incelenmiştir. Düzeltme indislerini seçerken, kadın ve erkek grupları için ortak olmalarına ve yüksek değere sahip olmalarına dikkat edilerek, klinik anlamlılıkları da göz önünde bulundurulmuştur. Bu kapsamda h1 (Ayakkabı bağlamak ve düğme iliklemek dahil,

kendi kendinize giyinebiliyor musunuz) ile h10 (Kendi kendinize yıkanıp, kurulabiliyor musunuz), h4 (Yatağa yatıp, kalkabiliyor musunuz) ile h8 (Dışarıda, düz bir zemin üzerinde yürüyebiliyor musunuz) maddelerinin hataları ilişkilendirilerek, **Model 2** sonuçları elde edilmiştir. Çizelge 3.1'den Model 2'ye ilişkin uyum istatistikleri incelendiğinde, uyum değerlerinin iyi olduğu ve kadın ile erkek hastalar için MSDA ölçme modelinin geçerli olduğu sonucuna ulaşılmıştır. Gruplar bazında DFA sonuçları yeterli bulunduğundan, ölçme eşdeğerliğini incelemek üzere hiyerarşik analizlere geçilmiştir.

Ölçme eşdeğerliğinin incelenmesi amacıyla yapılan hiyerarşik analizlerin ilk aşaması olan yapısal değişmezlik test edilirken faktör sayısı ve gözlenen değişkenlerin faktöre yüklenme örüntüsü üzerine sınırlama getirilmiştir. Diğer bir deyişle, bu aşamada kadın ve erkek hasta gruplarında faktör yapılarının benzer olup olmadığı aynı model içerisinde eşzamanlı olarak test edilmiştir. Yapısal değişmezlik modeline göre cinsiyetler bazında faktör yükleri ve eşik değerleri Çizelge 3.3'de sunulmuştur. Çizelge 3.3'den görülebileceği gibi, bu aşamada faktör yükleri ve eşik değerleri üzerinde bir sınırlama bulunmamaktadır.

Çizelge 3.2'ye göre, söz konusu model uyum istatistikleri incelendiğinde $YHKOK=0.083$, $KUI=0.993$, $TLI=0.991$ ve $\chi^2/sd=2.04$ olup; yapısal değişmezliğin sağlandığını kanıtlamaktadır.

Çizelge 3.2 Cinsiyet bazında ÇG-DFA'ya ilişkin uyum iyiliği istatistikleri ve karşılaştırma sonuçları

<i>Model</i>	<i>YHKOK</i>	<i>$\Delta\chi^2$ (Δsd)</i>	<i>P</i>	<i>ΔKUI</i>	<i>ΔTLI</i>
Yapısal Değişmezlik	0.083	-	-	-	-
Zayıf Değişmezlik	0.066	6.067 (7)	0.532	0.002	0.003
Güçlü Değişmezlik	0.063	12.971 (8)	0.113	0	0.001

Analizin ikinci aşamasında test edilen zayıf değişmezlikte, yapısal değişmezliğe ek olarak grupların faktöre ilişkin ölçme biriminin benzer olup olmadığı test edilmiştir. Zayıf değişmezlik test edilirken, faktör sayısı ve yüklenme örüntüsü ile birlikte faktör yükleri de sınırlandırılmıştır. Zayıf değişmezlik modeline göre cinsiyetler bazında faktör yükleri ve eşik değerleri Çizelge 3.4'te sunulmuştur. Çizelge 3.4'ten

görülebileceği gibi, bu aşamada faktör yükleri birbirine eşit iken; eşik değerleri üzerinde bir sınırlama bulunmamaktadır.

Çizelge 3.3 Yapısal değişmezlik modeline göre cinsiyetler bazında faktör yükleri ve eşik değerleri

	Kadın			Erkek		
<i>Faktör: MSDA</i>	Tahmin	SH	p	Tahmin	SH	P
h1- Ayakkabı bağlamak ve düğme iliklemek dahil, kendi kendinize giyinebiliyor musunuz?	1.000	-	-	1.000	-	-
h4- Yatağa yatıp, kalkabiliyor musunuz?	0.835	0.056	<0.001	0.939	0.160	<0.001
h6- Dolu bir fincanı veya bardağı ağzınıza götürebiliyor musunuz?	0.929	0.043	<0.001	0.960	0.106	<0.001
h8- Dışarıda, düz bir zemin üzerinde yürüyebiliyor musunuz?	0.863	0.063	<0.001	0.999	0.165	<0.001
h10- Kendi kendinize yıkanıp, kurulatabiliyor musunuz?	0.973	0.016	<0.001	0.955	0.060	<0.001
h14- Eğilip yerden bir giysiye alabiliyor musunuz?	0.893	0.040	<0.001	1.107	0.106	<0.001
h17- Muslukları açıp kapatabiliyor musunuz?	0.868	0.051	<0.001	0.995	0.152	<0.001
h19- Arabaya binip inebiliyor musunuz?	0.922	0.037	<0.001	1.093	0.157	<0.001
Eşik Değerleri						
h1_1	0.213	0.144	0.139	0.680	0.274	0.013
h1_2	1.177	0.182	<0.001	1.177	0.182	<0.001
h1_3	1.713	0.251	<0.001	1.713	0.251	<0.001
h4_1	0.644	0.154	<0.001	0.830	0.266	0.002
h4_2	1.626	0.238	<0.001	1.626	0.238	<0.001
h6_1	0.528	0.150	<0.001	0.892	0.226	<0.001
h6_2	1.455	0.196	<0.001	1.455	0.196	<0.001
h6_3	2.023	0.308	<0.001	2.023	0.308	<0.001
h8_1	0.959	0.169	<0.001	1.071	0.264	<0.001
h8_2	1.763	0.261	<0.001	1.763	0.261	<0.001
h10_1	0.384	0.147	0.009	0.710	0.246	0.004
h10_2	1.141	0.173	<0.001	1.141	0.173	<0.001
h10_3	1.527	0.220	<0.001	1.527	0.220	<0.001
h14_1	0.528	0.150	<0.001	0.766	0.282	0.007
h14_2	1.454	0.188	<0.001	1.454	0.188	<0.001
h14_3	1.888	0.289	<0.001	1.888	0.289	<0.001
h17_1	0.685	0.156	<0.001	0.860	0.260	0.001
h17_2	1.366	0.202	<0.001	1.366	0.202	<0.001
h17_3	1.825	0.255	<0.001	1.825	0.255	<0.001
h19_1	0.770	0.159	<0.001	0.915	0.280	0.001
h19_2	1.515	0.207	<0.001	1.515	0.207	<0.001
h19_3	1.977	0.288	<0.001	1.977	0.288	<0.001

SH: Standart hata; Tahmin, faktör yükünü ve eşik değerleri ifade etmektedir.

h4 ve h8 maddelerinde sadece 0, 1 ve 2 kategorilerinde yanıt veren hastalar olduğu için, bu iki madde için eşik değer sayısı ikidir.

Çizelge 3.4 Zayıf değişmezlik modeline göre cinsiyetler bazında faktör yükleri ve eşik değerleri

	Kadın			Erkek		
<i>Faktör: MSDA</i>	Tahmin	SH	p	Tahmin	SH	P
h1- Ayakkabı bağlamak ve düğme ilikleme dahil, kendi kendinize giyinebiliyor musunuz?	1.000	-	-	1.000	-	-
h4- Yatağa yatıp, kalkabiliyor musunuz?	0.842	0.053	<0.001	0.842	0.053	<0.001
h6- Dolu bir fincanı veya bardağı ağızınıza götürebiliyor musunuz?	0.936	0.038	<0.001	0.936	0.038	<0.001
h8- Dışarıda, düz bir zemin üzerinde yürüyebiliyor musunuz?	0.870	0.059	<0.001	0.870	0.059	<0.001
h10- Kendi kendinize yıkanıp, kurulanabiliyor musunuz?	0.972	0.015	<0.001	0.972	0.015	<0.001
h14- Eğilip yerden bir giysiye alabiliyor musunuz?	0.907	0.036	<0.001	0.907	0.036	<0.001
h17- Muslukları açıp kapatabiliyor musunuz?	0.877	0.048	<0.001	0.877	0.048	<0.001
h19- Arabaya binip inebiliyor musunuz?	0.933	0.033	<0.001	0.933	0.033	<0.001
Eşik Değerleri						
h1_1	0.213	0.144	0.139	0.652	0.273	0.017
h1_2	1.233	0.181	<0.001	1.233	0.181	<0.001
h1_3	1.837	0.245	<0.001	1.837	0.245	<0.001
h4_1	0.644	0.154	<0.001	0.754	0.197	<0.001
h4_2	1.605	0.189	<0.001	1.605	0.189	<0.001
h6_1	0.528	0.150	<0.001	0.887	0.211	<0.001
h6_2	1.525	0.194	<0.001	1.525	0.194	<0.001
h6_3	2.146	0.304	<0.001	2.146	0.304	<0.001
h8_1	0.959	0.169	<0.001	0.974	0.179	<0.001
h8_2	1.700	0.227	<0.001	1.700	0.227	<0.001
h10_1	0.384	0.147	0.009	0.706	0.247	0.004
h10_2	1.209	0.171	<0.001	1.209	0.171	<0.001
h10_3	1.667	0.216	<0.001	1.667	0.216	<0.001
h14_1	0.528	0.150	<0.001	0.605	0.250	0.015
h14_2	1.297	0.167	<0.001	1.297	0.167	<0.001
h14_3	1.720	0.228	<0.001	1.720	0.228	<0.001
h17_1	0.685	0.156	<0.001	0.764	0.214	0.001
h17_2	1.303	0.158	<0.001	1.303	0.158	<0.001
h17_3	1.783	0.214	<0.001	1.783	0.214	<0.001
h19_1	0.770	0.159	<0.001	0.783	0.224	0.001
h19_2	1.408	0.171	<0.001	1.408	0.171	<0.001
h19_3	1.870	0.259	<0.001	1.870	0.259	<0.001

SH: Standart hata; Tahmin, faktör yükünü ve eşik değerleri ifade etmektedir.

h4 ve h8 maddelerinde sadece 0, 1 ve 2 kategorilerinde yanıt veren hastalar olduğu için, bu iki madde için eşik değer sayısı ikidir.

Zayıf değişmezlik modeli için ölçme eşdeğerliği, zayıf değişmezlik modelinin χ^2 , KUI ve TLI değerlerinden yapısal değişmezlik modeli için elde edilenler çıkarılarak incelenmiştir. Elde edilen $\Delta\chi^2$, ΔCFI ve ΔTLI değerleri modelde anlamlı düzeyde bir kötüleşme olmadığını göstermektedir (Çizelge 3.2). Bu sonuç gruplar arasında zayıf değişmezliğin sağlandığına yani faktör yüklerinin kadın ve erkekler için eşdeğer olduğuna işaret etmektedir.

Analizin üçüncü aşaması olan güçlü değişmezlik test edilirken, ilk iki aşamada yapılan sınırlandırmalara ek olarak eşik değerleri de sınırlandırılmıştır. Güçlü değişmezlik sağlandığında hem faktör yükleri hem eşik değerleri birbirine eşit olduğunda, farklı gruplardan elde edilen puanlar hem aynı ölçüm birimine hem de aynı orjine sahip olurlar. Güçlü değişmezlik modeline göre cinsiyetler bazında faktör yükleri ve eşik değerleri Çizelge 3.5'te sunulmuştur. Çizelge 3.5'ten de görülebileceği gibi, bu aşamada hem faktör yükleri hem de eşik değerleri gruplar bazında birbirine eşittir.

Güçlü değişmezlik modeli için ölçme eşdeğerliği, güçlü değişmezlik modelinin χ^2 , KUI ve TLI değerlerinden zayıf değişmezlik modeli için elde edilenler çıkarılarak incelenmiştir. Elde edilen $\Delta\chi^2$, ΔCFI ve ΔTLI değerleri modelde anlamlı düzeyde bir kötüleşme olmadığını göstermektedir (Çizelge 3.2). Bu sonuç gruplar arasında güçlü değişmezliğin sağlandığına yani faktör yüklerinin ve eşik değerlerinin kadın ve erkekler için eşdeğer olduğunu göstermektedir.

Hiyerarşik analiz sonuçları genel olarak değerlendirildiğinde, MSDA'da faktör yapısı ve örüntüsünün, faktör yüklerinin ve eşik değerlerinin kadın ve erkek hasta grupları için eşdeğer olduğu görülmektedir.

Çizelge 3.5 Güçlü değişmezlik modeline göre cinsiyetler bazında faktör yükleri ve eşik değerleri

	Kadın			Erkek		
Faktör: MSDA	Tahmin	SH	p	Tahmin	SH	P
h1- Ayakkabı bağlamak ve düğme iliklemek dahil, kendi kendinize giyinebiliyor musunuz?	1.000	-	-	1.000	-	-
h4- Yatağa yatıp, kalkabiliyor musunuz?	0.841	0.053	<0.001	0.841	0.053	<0.001
h6- Dolu bir fincanı veya bardağı ağızınıza götürebiliyor musunuz?	0.932	0.037	<0.001	0.932	0.037	<0.001
h8- Dışarıda, düz bir zemin üzerinde yürüyebiliyor musunuz?	0.871	0.058	<0.001	0.871	0.058	<0.001
h10- Kendi kendinize yıkayıp, kurulanabiliyor musunuz?	0.972	0.015	<0.001	0.972	0.015	<0.001
h14- Eğilip yerden bir giysiye alabiliyor musunuz?	0.908	0.036	<0.001	0.908	0.036	<0.001
h17- Muslukları açıp kapatabiliyor musunuz?	0.878	0.048	<0.001	0.878	0.048	<0.001
h19- Arabaya binip inebiliyor musunuz?	0.934	0.034	<0.001	0.934	0.034	<0.001
Eşik Değerleri						
h1_1	0.360	0.141	0.011	0.360	0.141	0.011
h1_2	1.156	0.164	<0.001	1.156	0.164	<0.001
h1_3	1.860	0.242	<0.001	1.860	0.242	<0.001
h4_1	0.610	0.122	<0.001	0.610	0.122	<0.001
h4_2	1.647	0.191	<0.001	1.647	0.191	<0.001
h6_1	0.675	0.133	<0.001	0.675	0.133	<0.001
h6_2	1.514	0.195	<0.001	1.514	0.195	<0.001
h6_3	2.219	0.288	<0.001	2.219	0.288	<0.001
h8_1	0.885	0.132	<0.001	0.885	0.132	<0.001
h8_2	1.757	0.227	<0.001	1.757	0.227	<0.001
h10_1	0.464	0.136	<0.001	0.464	0.136	0.001
h10_2	1.131	0.155	<0.001	1.131	0.155	<0.001
h10_3	1.680	0.218	<0.001	1.680	0.218	<0.001
h14_1	0.418	0.128	0.001	0.418	0.128	0.001
h14_2	1.264	0.167	<0.001	1.264	0.167	<0.001
h14_3	1.777	0.230	<0.001	1.777	0.230	<0.001
h17_1	0.617	0.125	<0.001	0.617	0.125	<0.001
h17_2	1.281	0.156	<0.001	1.281	0.156	<0.001
h17_3	1.846	0.218	<0.001	1.846	0.218	<0.001
h19_1	0.638	0.134	<0.001	0.638	0.134	<0.001
h19_2	1.389	0.172	<0.001	1.389	0.172	<0.001
h19_3	1.956	0.261	<0.001	1.956	0.261	<0.001

SH: Standart hata; Tahmin, faktör yükünü ve eşik değerleri ifade etmektedir.

h4 ve h8 maddelerinde sadece 0, 1 ve 2 kategorilerinde yanıt veren hastalar olduğu için, bu iki madde için eşik değer sayısı ikidir.

3.1.2 Hastalık Süresi Bazında ÇG-DFA Sonuçları

Hastalık süresi değişkeninin, “ ≤ 10 yıl” , “ > 10 yıl” gibi iki grupta incelendiği MSDA’nın ölçme eşdeğerliği test edilirken önce temel model olarak bu iki grup için ayrı ayrı DFA yapılmıştır. Bu gruplar için yapılan DFA’ya ilişkin model uyum indeksleri Çizelge 3.6 ’da **Model 3** olarak verilmiştir.

Çizelge 3.6 Hastalık süresi bazında her bir grup için DFA’ya ilişkin uyum iyiliği istatistikleri

	<i>Grup</i>	<i>Ki-Kare (χ^2)</i>	<i>sd</i>	<i>χ^2/sd</i>	<i>YHKOK</i>	<i>KUİ</i>	<i>TLİ</i>
Model 3	≤ 10 Yıl	85.88	20	4.29	0.139	0.983	0.976
	> 10 Yıl	35.95	17	2.11	0.081	0.995	0.992
Model 4	≤ 10 Yıl	60.34	20	3.02	0.125	0.990	0.987
	> 10 Yıl	38.97	17	2.29	0.100	0.995	0.991

Çizelge 3.6’dan görülebileceği gibi, her iki hastalık süresi grubu için de ölçme modelinin uyum indeks değerleri, verinin modele orta düzeyde uyum gösterdiğini ifade etmektedir. Bu durumda program tarafından önerilen, maddelerin hata varyanslarının ilişkilendirilmesine yönelik olan düzeltme indisleri incelenmiştir. Düzeltme indislerini seçerken, hastalık süresi grupları için ortak olmalarına ve yüksek değere sahip olmalarına dikkat edilerek, klinik anlamlılıkları da göz önünde bulundurulmuştur. Bu kapsamda h4 (Yatağa yatıp, kalkabiliyor musunuz) ile h8 (Dışarıda, düz bir zemin üzerinde yürüyebiliyor musunuz), h1 (Ayakkabı bağlamak ve düğme iliklemek dahil, kendi kendinize giyinebiliyor musunuz) ile h10 (Kendi kendinize yıkanıp, kurulanabiliyor musunuz), h8 (Dışarıda, düz bir zemin üzerinde yürüyebiliyor musunuz) ile h14 (Eğilip yerden bir giysiyi alabiliyor musunuz) maddelerinin hataları ilişkilendirilerek, **Model 4** sonuçları elde edilmiştir. Çizelge 3.6’dan Model 4’e ilişkin uyum istatistikleri incelendiğinde, uyum değerlerinin iyi olduğu ve hastalık süresi grupları için MSDA ölçme modelinin geçerli olduğu sonucuna ulaşılmıştır. Gruplar bazında DFA sonuçları yeterli bulunduğundan, ölçme eşdeğerliğini incelemek üzere hiyerarşik analizlere geçilmiştir.

Ölçme eşdeğerliğinin incelenmesi amacıyla yapılan hiyerarşik analizlerin ilk aşaması olan yapısal değişmezlik test edilirken faktör sayısı ve gözlenen değişkenlerin faktöre yüklenme örüntüsü sınırlandırılarak, hastalık süresi gruplarında faktör

yapılarının benzer olup olmadığı aynı model içerisinde eşzamanlı olarak test edilmiştir. Yapısal değişmezlik modeline göre hastalık süresi grupları bazında faktör yükleri ve eşik değerleri Çizelge 3.8’de sunulmuştur. Çizelge 3.8’den görülebileceği gibi, bu aşamada faktör yükleri ve eşik değerleri üzerinde bir sınırlama bulunmamaktadır.

Çizelge 3.7’ye göre, söz konusu model uyum istatistikleri incelendiğinde $YHKOK=0.083$, $KUI=0.995$, $TLI=0.991$ ve $\chi^2/sd=2.04$ olup; yapısal değişmezliğin sağlandığını kanıtlamaktadır.

Çizelge 3.7 Hastalık süresi bazında ÇG-DFA’ya ilişkin uyum iyiliği istatistikleri ve karşılaştırma sonuçları

<i>Model</i>	<i>YHKOK</i>	<i>$\Delta\chi^2 (\Delta sd)$</i>	<i>p</i>	<i>ΔKUI</i>	<i>ΔTLI</i>
Yapısal Değişmezlik	0.083	-	-	-	-
Zayıf Değişmezlik	0.067	7.501 (7)	0.38	0.001	0.002
Güçlü Değişmezlik	0.065	13.458 (8)	0.097	0	0.001

Hastalık süresi grupları arasında yapılan analizin ikinci aşamasında test edilen zayıf değişmezlikte, cinsiyet gruplarına göre yapılan işlemlerin aynısı hastalık süresi değişkeni için de tekrarlanır. Zayıf değişmezlik test edilirken, faktör sayısı ve yüklenme örüntüsü ile birlikte faktör yükleri de sınırlandırılmıştır. Zayıf değişmezlik modeline göre hastalık süresi grupları bazında faktör yükleri ve eşik değerleri Çizelge 3.9’da sunulmuştur. Çizelge 3.9’dan görülebileceği gibi, bu aşamada faktör yükleri birbirine eşit iken; eşik değerleri üzerinde bir sınırlama bulunmamaktadır.

Zayıf değişmezlik modeli için ölçme eşdeğerliği, zayıf değişmezlik modelinin χ^2 , KUI ve TLI değerlerinden yapısal değişmezlik modeli için elde edilenler çıkarılarak incelenmiştir. Elde edilen $\Delta\chi^2$, ΔCFI ve ΔTLI değerleri modelde anlamlı düzeyde bir kötüleşme olmadığını göstermektedir (Çizelge 3.7). Bu sonuç gruplar arasında zayıf değişmezliğin sağlandığına yani faktör yüklerinin hastalık süresi grupları arasında benzer olduğunu göstermektedir.

Çizelge 3.8 Yapısal değişmezlik modeline göre hastalık süresi bazında faktör yükleri ve eşik değerleri

	Hastalık süresi ≤10 yıl			Hastalık süresi >10 yıl		
<i>Faktör: MSDA</i>	Tahmin	SH	p	Tahmin	SH	P
h1- Ayakkabı bağlamak ve düğme iliklemek dahil, kendi kendinize giyinebiliyor musunuz?	1.000	-	-	1.000	-	-
h4- Yatağa yatıp, kalkabiliyor musunuz?	0.875	0.036	<0.001	1.129	0.198	<0.001
h6- Dolu bir fincanı veya bardağı ağızınıza götürebiliyor musunuz?	0.892	0.043	<0.001	1.094	0.128	<0.001
h8- Dışarıda, düz bir zemin üzerinde yürüyebiliyor musunuz?	0.827	0.069	<0.001	0.978	0.176	<0.001
h10- Kendi kendinize yıkayıp, kurulanabiliyor musunuz?	0.942	0.026	<0.001	1.134	0.114	<0.001
h14- Eğilip yerden bir giysiye alabiliyor musunuz?	0.977	0.032	<0.001	0.998	0.138	<0.001
h17- Muslukları açıp kapatabiliyor musunuz?	0.947	0.046	<0.001	0.942	0.124	<0.001
h19- Arabaya binip inebiliyor musunuz?	0.911	0.041	<0.001	1.228	0.190	<0.001
<i>Eşik Değerleri</i>						
h1_1	0.284	0.098	0.004	0.276	0.274	0.314
h1_2	1.111	0.117	<0.001	1.111	0.117	<0.001
h1_3	1.875	0.186	<0.001	1.875	0.186	<0.001
h4_1	0.665	0.104	<0.001	0.616	0.300	0.040
h4_2	1.986	0.209	<0.001	1.986	0.209	<0.001
h6_1	0.593	0.103	<0.001	0.755	0.228	0.001
h6_2	1.620	0.153	<0.001	1.620	0.153	<0.001
h6_3	2.509	0.301	<0.001	2.509	0.301	<0.001
h8_1	0.999	0.116	<0.001	0.776	0.234	<0.001
h8_2	1.890	0.194	<0.001	1.890	0.194	<0.001
h10_1	0.393	0.099	<0.001	0.452	0.270	0.094
h10_2	1.173	0.124	<0.001	1.173	0.124	<0.001
h10_3	1.790	0.169	<0.001	1.790	0.169	<0.001
h14_1	0.393	0.099	<0.001	0.361	0.266	0.174
h14_2	1.349	0.125	<0.001	1.349	0.125	<0.001
h14_3	1.960	0.207	<0.001	1.960	0.207	<0.001
h17_1	0.508	0.101	<0.001	0.724	0.190	<0.001
h17_2	1.353	0.130	<0.001	1.353	0.130	<0.001
h17_3	1.983	0.194	<0.001	1.983	0.194	<0.001
h19_1	0.593	0.103	<0.001	0.744	0.273	0.006
h19_2	1.584	0.155	<0.001	1.584	0.155	<0.001
h19_3	2.328	0.237	<0.001	2.328	0.237	<0.001

SH: Standart hata; Tahmin, faktör yükünü ve eşik değerleri ifade etmektedir.

h4 ve h8 maddelerinde sadece 0, 1 ve 2 kategorilerinde yanıt veren hastalar olduğu için, bu iki madde için eşik değer sayısı ikidir.

Çizelge 3.9 Zayıf değişmezlik modeline göre hastalık süresi bazında faktör yükleri ve eşik değerleri

	Hastalık süresi ≤10 yıl			Hastalık süresi >10 yıl		
<i>Faktör: MSDA</i>	Tahmin	SH	p	Tahmin	SH	P
h1- Ayakkabı bağlamak ve düğme iliklemek dahil, kendi kendinize giyinebiliyor musunuz?	1.000	-	-	1.000	-	-
h4- Yatağa yatıp, kalkabiliyor musunuz?	0.885	0.035	<0.001	0.885	0.035	<0.001
h6- Dolu bir fincanı veya bardağı ağzınıza götürebiliyor musunuz?	0.900	0.039	<0.001	0.900	0.039	<0.001
h8- Dışarıda, düz bir zemin üzerinde yürüyebiliyor musunuz?	0.832	0.064	<0.001	0.832	0.064	<0.001
h10- Kendi kendinize yıkayıp, kurulanabiliyor musunuz?	0.952	0.024	<0.001	0.952	0.024	<0.001
h14- Eğilip yerden bir giysiye alabiliyor musunuz?	0.975	0.029	<0.001	0.975	0.029	<0.001
h17- Muslukları açıp kapatabiliyor musunuz?	0.943	0.042	<0.001	0.943	0.042	<0.001
h19- Arabaya binip inebiliyor musunuz?	0.924	0.037	<0.001	0.924	0.037	<0.001
<i>Eşik Değerleri</i>						
h1_1	0.284	0.098	0.004	0.130	0.291	0.655
h1_2	1.136	0.119	<0.001	1.136	0.119	<0.001
h1_3	1.990	0.172	<0.001	1.990	0.172	<0.001
h4_1	0.665	0.104	<0.001	0.421	0.229	0.066
h4_2	1.891	0.165	<0.001	1.891	0.165	<0.001
h6_1	0.593	0.103	<0.001	0.595	0.201	0.003
h6_2	1.571	0.133	<0.001	1.571	0.133	<0.001
h6_3	2.468	0.248	<0.001	2.468	0.248	<0.001
h8_1	0.999	0.116	<0.001	0.659	0.178	<0.001
h8_2	1.879	0.163	<0.001	1.879	0.163	<0.001
h10_1	0.393	0.099	<0.001	0.273	0.248	0.270
h10_2	1.126	0.115	<0.001	1.126	0.115	<0.001
h10_3	1.754	0.151	<0.001	1.754	0.151	<0.001
h14_1	0.393	0.099	<0.001	0.234	0.273	0.391
h14_2	1.375	0.120	<0.001	1.375	0.120	<0.001
h14_3	2.102	0.184	<0.001	2.102	0.184	<0.001
h17_1	0.508	0.101	<0.001	0.712	0.191	<0.001
h17_2	1.409	0.123	<0.001	1.409	0.123	<0.001
h17_3	2.111	0.174	<0.001	2.111	0.174	<0.001
h19_1	0.593	0.103	<0.001	0.512	0.213	0.016
h19_2	1.463	0.136	<0.001	1.463	0.136	<0.001
h19_3	2.166	0.181	<0.001	2.166	0.181	<0.001

SH: Standart hata; Tahmin, faktör yükünü ve eşik değerleri ifade etmektedir. h4 ve h8 maddelerinde sadece 0, 1 ve 2 kategorilerinde yanıt veren hastalar olduğu için, bu iki madde için eşik değer sayısı ikidir.

Analizin üçüncü aşaması olan güçlü değişmezlik test edilirken, ilk iki aşamada yapılan sınırlandırmalara ek olarak eşik değerleri de sınırlandırılmıştır. Güçlü değişmezlik modeline göre hastalık süresi grupları bazında faktör yükleri ve eşik değerleri Çizelge 3.10'da sunulmuştur. Çizelge 3.10'da görülebileceği gibi, bu aşamada hem faktör yükleri hem de eşik değerleri gruplar bazında birbirine eşittir.

Güçlü değişmezlik modeli için ölçme eşdeğerliği, güçlü değişmezlik modelinin χ^2 , KUI ve TLI değerlerinden zayıf değişmezlik modeli için elde edilenler çıkarılarak incelenmiştir. Elde edilen $\Delta\chi^2$, ΔCFI ve ΔTLI değerleri modelde anlamlı düzeyde bir kötüleşme olmadığını göstermektedir (Çizelge 3.7). Bu sonuç gruplar arasında güçlü değişmezliğin sağlandığına yani faktör yüklerinin ve eşik değerlerinin hastalık süresi gruplar arasında eşdeğer olduğunu ifade etmektedir.

Hiyerarşik analiz sonuçları tekrar genel olarak değerlendirildiğinde, MSDA'da faktör yapısı ve örüntüsünün, faktör yüklerinin ve eşik değerlerinin hastalık süresi grupları için eşdeğer olduğu görülmektedir.

Çizelge 3.10. Güçlü değişmezlik modeline göre hastalık süresi bazında faktör yükleri ve eşik değerleri

	Hastalık süresi ≤10 yıl			Hastalık süresi >10 yıl		
<i>Faktör: MSDA</i>	Tahmin	SH	p	Tahmin	SH	P
h1- Ayakkabı bağlamak ve düğme iliklemek dahil, kendi kendinize giyinebiliyor musunuz?	1.000	-	-	1.000	-	-
h4- Yatağa yatıp, kalkabiliyor musunuz?	0.890	0.035	<0.001	0.890	0.035	<0.001
h6- Dolu bir fincanı veya bardağı ağzınıza götürebiliyor musunuz?	0.899	0.039	<0.001	0.899	0.039	<0.001
h8- Dışarıda, düz bir zemin üzerinde yürüyebiliyor musunuz?	0.849	0.062	<0.001	0.849	0.062	<0.001
h10- Kendi kendinize yıkayıp, kurulanabiliyor musunuz?	0.952	0.024	<0.001	0.952	0.024	<0.001
h14- Eğilip yerden bir giysiye alabiliyor musunuz?	0.977	0.029	<0.001	0.977	0.029	<0.001
h17- Muslukları açıp kapatabiliyor musunuz?	0.939	0.042	<0.001	0.939	0.042	<0.001
h19- Arabaya binip inebiliyor musunuz?	0.925	0.037	<0.001	0.925	0.037	<0.001
<i>Eşik Değerleri</i>						
h1_1	0.273	0.095	0.004	0.273	0.095	0.004
h1_2	1.155	0.111	<0.001	1.155	0.111	<0.001
h1_3	1.991	0.171	<0.001	1.991	0.171	<0.001
h4_1	0.611	0.095	<0.001	0.611	0.095	<0.001
h4_2	1.889	0.160	<0.001	1.889	0.160	<0.001
h6_1	0.620	0.092	<0.001	0.620	0.092	<0.001
h6_2	1.573	0.131	<0.001	1.573	0.131	<0.001
h6_3	2.433	0.231	<0.001	2.433	0.231	<0.001
h8_1	0.894	0.100	<0.001	0.894	0.100	<0.001
h8_2	1.893	0.162	<0.001	1.893	0.162	<0.001
h10_1	0.387	0.094	<0.001	0.387	0.094	<0.001
h10_2	1.142	0.107	<0.001	1.142	0.107	<0.001
h10_3	1.756	0.150	<0.001	1.756	0.150	<0.001
h14_1	0.377	0.094	<0.001	0.377	0.094	<0.001
h14_2	1.392	0.118	<0.001	1.392	0.118	<0.001
h14_3	2.095	0.177	<0.001	2.095	0.177	<0.001
h17_1	0.600	0.092	<0.001	0.600	0.092	<0.001
h17_2	1.418	0.120	<0.001	1.418	0.120	<0.001
h17_3	2.102	0.171	<0.001	2.102	0.171	<0.001
h19_1	0.594	0.093	<0.001	0.594	0.093	<0.001
h19_2	1.468	0.135	<0.001	1.468	0.135	<0.001
h19_3	2.148	0.175	<0.001	2.148	0.175	<0.001

SH: Standart hata; Tahmin, faktör yükünü ve eşik değerleri ifade etmektedir.

h4 ve h8 maddelerinde sadece 0, 1 ve 2 kategorilerinde yanıt veren hastalar olduğu için, bu iki madde için eşik değer sayısı ikidir.

3.2 Rasch Analizi Sonuçları

MSDA’nde yer alan maddelerin MİF açısından incelenmesinde maddeler sıralanabilir nitel (ordinal) olduğu için kısmi kredi modeli kullanılarak madde parametreleri tahmin edilmiştir. Rasch modelinde sadece madde zorluk parametresi söz konusu olduğu için, MİF incelemesi bu parametreye dayalı olarak gerçekleştirilmiştir.

Cinsiyet ve hastalık süresi açısından MİF sonuçları sırasıyla Çizelge 3.11 ve Çizelge 3.12’de sunulmuştur. Her iki Çizelge da incelendiğinde, hem “grup” faktörü hem de “sınıf aralığı*grup” etkileşimi faktörü için hiçbir maddenin p değerinin, Bonferroni düzeltmesine göre ($p=0.003$) istatistiksel olarak anlamlı olmadığı görülmektedir. Bu durumda, hem cinsiyet için hem de hastalık süresi için hiçbir maddenin MİF göstermediği sonucuna ulaşılmıştır.

Çizelge 3.11 Cinsiyet bazında MİF sonuçları

Maddeler	Cinsiyet				Sınıf Aralığı * Cinsiyet			
	KT	F	sd	p	KT	F	sd	P
h1-Ayakkabı bağlamak ve düğme iliklemek dahil, kendi kendinize giyinebiliyor musunuz?	0.823	1.273	1	0.261	0.451	0.698	3	0.554
h4-Yatağa yatıp, kalkabiliyor musunuz?	0.000	0.000	1	0.995	0.175	0.174	3	0.914
h6-Dolu bir fincanı veya bardağı ağzınıza götürebiliyor musunuz?	0.897	0.894	1	0.346	1.362	1.358	3	0.258
h8-Dışarıda, düz bir zemin üzerinde yürüyebiliyor musunuz?	0.227	0.205	1	0.651	0.188	0.170	3	0.917
h10-Kendi kendinize yıkanıp, kurulanabiliyor musunuz?	0.458	0.597	1	0.441	0.338	0.441	3	0.724
h14-Eğilip yerden bir giysiyi alabiliyor musunuz?	1.161	1.415	1	0.236	1.047	1.276	3	0.285
h17-Muslukları açıp kapatabiliyor musunuz?	0.000	0.000	1	0.987	0.669	0.730	3	0.535
h19-Arabaya binip inebiliyor musunuz?	2.220	2.661	1	0.105	1.716	2.057	3	0.108

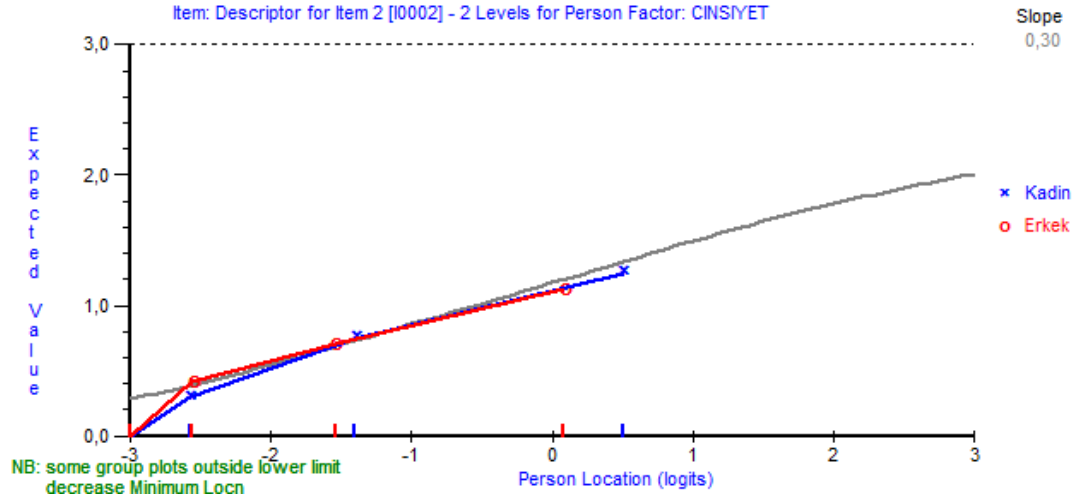
KT: Kareler toplamı, F: İki yönlü varyans analizinde hesapla bulunan F istatistiği, sd: serbestlik derecesi

Çizelge 3.12 Hastalık süresi bazında MİF sonuçları

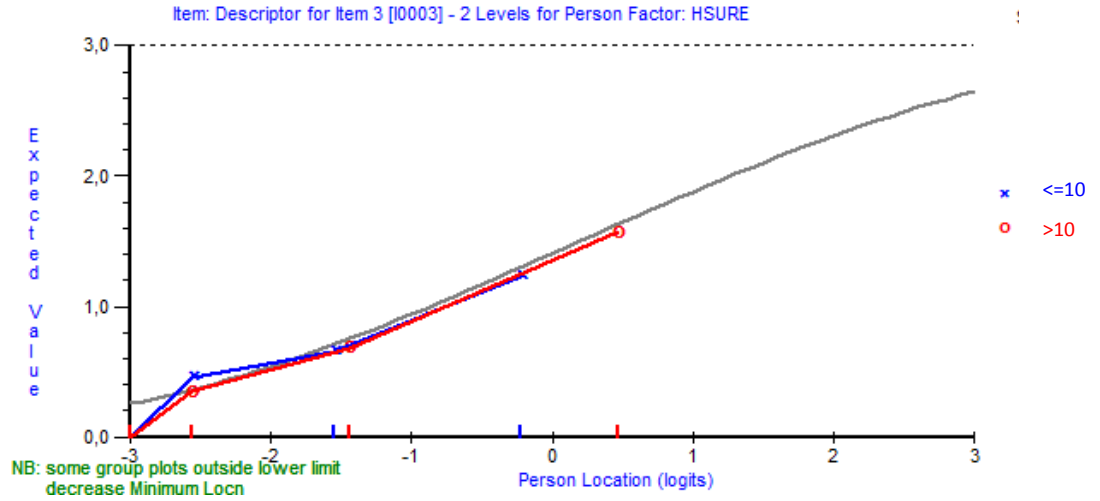
Maddeler	Hastalık süresi				Sınıf Aralığı * Hastalık Süresi			
	KT	F	sd	p	KT	F	sd	P
H31-Ayakkabı bağlamak ve düğme iliklemek dahil, kendi kendinize giyinebiliyor musunuz?	0.012	0.019	1	0.890	1.355	2.136	3	0.098
h4-Yatağa yatıp, kalkabiliyor musunuz?	2.887	2.955	1	0.088	0.765	0.783	3	0.505
h6-Dolu bir fincanı veya bardağı ağzınıza götürebiliyor musunuz?	0.121	0.117	1	0.732	0.146	0.142	3	0.935
h8-Dışarıda, düz bir zemin üzerinde yürüyebiliyor musunuz?	1.732	1.583	1	0.210	0.311	0.285	3	0.837
h10-Kendi kendinize yıkanıp, kurulabiliyor musunuz?	0.262	0.342	1	0.559	0.467	0.609	3	0.610
h14-Eğilip yerden bir giysiyi alabiliyor musunuz?	0.081	0.096	1	0.757	0.257	0.305	3	0.822
h17-Muslukları açıp kapatabiliyor musunuz?	6.836	7.787	1	0.006	0.487	0.555	3	0.646
h19-Arabaya binip inebiliyor musunuz?	0.551	0.638	1	0.426	0.700	0.811	3	0.489

KT: Kareler toplamı, F: İki yönlü varyans analizinde hesapla bulunan F istatistiği, sd: serbestlik derecesi

MSDA’da yer alan “h4-Yatağa yatıp, kalkabiliyor musunuz” maddesinin cinsiyete göre, “h6- Dolu bir fincanı veya bardağı ağzınıza götürebiliyor musunuz” maddesinin hastalık süresine göre MİF göstermediği grafiksel olarak örneklendirilmiştir (Şekil 3.1 ve Şekil 3.2). Şekil 3.1 ve Şekil 3.2 incelendiğinde, x eksenini 0 düzeyi, y eksenini bu maddede ilgili kategoride yanıt verme olasılığı olmak üzere, cinsiyet ya da hastalık süresi grupları için madde karakteristik eğrileri görülmektedir. Bu eğrilerin çakışık ve paralel olması MİF olmadığını ifade etmektedir.



Şekil 3.1. “h4-Yatağa yatıp, kalkabiliyor musunuz” maddesinin cinsiyete göre MİF incelemesi



Şekil 3.2. “h6- Dolu bir fincanı veya bardağı ağzınıza götürebiliyor musunuz” maddesinin hastalık süresine göre MİF incelemesi

4. TARTIŞMA

Sağlık alanında yapılan çalışmalarda, yaşam kalitesi, özürlülük, fonksiyonellik, işlevsellik, çok sayıda maddeden oluşan ölçme araçları ile ölçülen gizli yapılardır. Ölçme araçlarının psikometrik özelliklerinin değerlendirilmesi, ölçülen bu gizli yapı cinsinden bireylerin θ düzeylerinin doğru bir şekilde belirlenebilmesi üzerinde önemli bir etkiye sahiptir. İçerik geçerliği, kriter geçerliği, yapı geçerliği, içsel tutarlılık, tekrar edilebilirlik, değişime duyarlılık ve yorumlanabilirlik bu özelliklerden bazılarıdır.

Bireylerin θ düzeylerinin tahmin edilmesinde kullanılan ölçme araçları, farklı gruplarda yer alan bireylere uygulandığında, gizli özellik bakımından gruplar arası karşılaştırmaların anlamlı olabilmesi için ölçme aracının ölçtüğü yapının gruplarda eşdeğer olması gereklidir. Raju ve arkadaşlarının (2002) ifade ettiği gibi örtük özellik üzerinde aynı konumda olan kişilerin aynı gerçek puana sahip olması ölçme eşdeğerliğini sağlayan bir koşuldur (Raju et.al., 2002). Ölçme eşdeğerliği olarak adlandırılan bu kavramın sağlanamadığı durumlarda ölçme aracı geçerli olarak değerlendirilemez.

Ölçme eşdeğerliğinin incelenmesinde YEM uygulamaları kapsamında ÇG-DFA ve MYT kapsamında Rasch Analizi yaygın kullanılan yöntemler arasındadır. ÇG-DFA'da ölçüm değişmezliği; faktör yüklerinin, regresyon sabitlerinin / eşik değerlerinin ve hata varyanslarının gruplarda eşitliği; Rasch analizinde madde zorluk parametrelerinin eşitliği ile incelenir.

ÇG-DFA, faktör yapısının benzerliğinin (faktör sayısı), faktör yüklerinin (ölçeğin birimi), sabit değerlerin (ölçeğin orijini), artık varyansının (ölçeğin kesinliği) eşitliğini ve incelenen yapının diğer özelliklerini test etmek için kullanılabilir (Chen, 2008). ÇG-DFA, geleneksel olarak, temelde tüm parametrelerin serbest bırakıldığı model ile işleme başlar. Sonrasında her bir parametrenin gruplar benzer olduğu

kısıtlılığına dayalı model ile devam eder. Amaç, hangi parametrelerin gruplar arasında değişmez olduğunu saptamaktır (Costa ve ark., 2014). Rasch analizi ise ölçme aracında yer alan maddelerin zorluk parametrelerinin eşitliğinin MİF kapsamında test edilmesi temeline dayanır.

Bu tez çalışması kapsamında sağlık alanında kullanılan ve geçerli olduğu modern psikometrik yöntemler ile gösterilen MSDA'nın, cinsiyet ve hastalık süresi bakımından ölçüm değişmezliği gösterip göstermediği ÇG-DFA ve Rasch Analizi ile incelenmiştir.

Ölçme eşdeğerliğinin hiyerarşik olarak test edilmesinden önce, her bir grup bazında gerçekleştirilen DFA'da, model uyumunu arttırmak için düzeltme indislerinden yararlanılmış ve modelde, belirli bir maddeler için hata varyansları ilişkilendirilmiştir. Byrne ve arkadaşları tarafından yapılan bir çalışmaya göre, temel modelde bu tür düzeltmelerin yapılmasının başlangıç düzeyindeki model uyumunu arttırmak açısından önerildiği belirtilmektedir (Somer ve ark., 2009).

ÇG-DFA'da cinsiyet bazında değerlendirmeler 77 kadın ve 222 erkek hasta üzerinde gerçekleştirilmiştir. DFA için örneklem büyüklüğünün yüksek olması gerekmektedir. DFA kapsamında yapılan testler (özellikle de ki-kare testi), örneklem büyüklüğüne duyarlı olduğu için bu analizlerde örneklem büyüklüğünün yüksek olması istenir. Ayrıca, benzer büyüklükteki gruplar anlamlı yorumlar yapılabilmesi için önemlidir. Bununla beraber ÇG-DFA'da, örneklem büyüklüğünün gruplarda eşit olması koşulu bulunmamakta, eşit olmayan gruplar olduğunda da bu analiz gerçekleştirilebilmektedir (Use of CFA with Multiple Groups).

ÇG-DFA kapsamında ölçme eşdeğerliği yapısal, zayıf ve güçlü değişmezlik modelleri bazında değerlendirilmiştir. Sağlanması daha zor olan katı değişmezlik bu çalışma kapsamında incelenmemiştir. Asil (2010) tarafından yapılan bir çalışmada, gruplar arasında, regresyon sabitlerinin ve faktör yüklerinin değişmezliği sağlandığında test puanlarının geçerli bir şekilde karşılaştırılabileceği belirtilmiştir.

Bununla beraber hata varyanslarının değişmezliğinin test edilmesinin ölçme değişmezliği için ideal olsa da uygulamada zor olduğu ve birçok araştırmada gerekli bulunmadığından bahsedilmiştir. Çünkü hata varyansının sistematik hata ve tesadüfi hatadan oluştuğu ve sistematik hatanın grup üyeliğinden kaynaklanması durumunda regresyon doğrularının grup üyeleri için farklı olacağı düşünülmektedir (Asil, 2010).

Diğer yandan değişmezlik türlerinin ne detayda inceleneceği konusunda ortak bir fikir olmamakla beraber, yapısal, zayıf (en azından kısmi değişmezlik) ve güçlü değişmezliğin test edilmesinin minimum gereksinim olduğu belirtilmiştir (Costa ve ark., 2014). Sadece zayıf değişmezliğin sağlanması durumunda ölçüm değişmezliğinin olduğu farklı yazarlar tarafından ifade edilmiştir (Raju ve ark., 2002). Ayrıca, değişmezlik özelliğinin bozulmasının, grup karşılaştırmalarında elde edilen sonuçlar üzerinde önemli bir etkiye sahip olduğu belirtilmiştir. Ölçüm değişmezliği sağlanmadığında, ölçümden kaynaklanan hataların, gerçek olmayan grup farklılıklarına neden olabileceği ya da ölçümden kaynaklanan hatalar nedeniyle net olarak belirlenemeyen gerçek grup farklılıklarının gözden kaçabileceği vurgulanmıştır (Chen, 2008).

Ölçme eşdeğerliğinin incelenmesi bakımından iki yöntem arasında benzerlik ve farklılıklar bulunmaktadır. Temelde her iki yöntem de, gizli değişken ile gözlenen değişkenler arasındaki ilişkiyi değerlendirir. Farklılık, gizli değişken ile faktör arasındaki ilişkinin ÇG-DFA’da doğrusal; Rasch analizi kapsamında doğrusal olmamasıdır. Bunun bir sonucu olarak θ ’nın orta düzeylerinde bireylerin gerçek puanlarını tahmin etmede iki yöntem yakın sonuçlar verirken; uç düzeylerde MYT modelleri daha hassas ve daha doğru kestirimlerde bulunmaktadır. Dolayısıyla MYT modelleri ölçme eşdeğerliğini θ ’nın uç düzeylerinde regresyon modellerine dayalı YEM modellerine göre daha doğru olarak test etmektedir (Somer ve ark., 2009). ÇG-DFA, eşzamanlı olarak çok sayıda gizli değişken olduğunda avantajlı iken; Rasch analizinin temel varsayımı incelenen yapının tek boyutlu olmasıdır. Bu durumda çok boyutlu Rasch modellerinin (multidimensional Rasch models) kullanılması uygun olabilir. Bu ayrışmalara rağmen, her iki yaklaşım da ölçüm değişmezliğinin

bozulumu durumunda, bu problemin kaynağını bulma amacıyla kullanılabilir (Raju et.al., 2002).

Somer ve arkadaşları tarafından yapılan bir çalışmada, test geliştirme çalışmalarında maddelerin gruplara ilişkin yanlılık özellikleri tek tek incelenmek istenildiğinde çalışmalara MYT modelleri ile başlamanın daha pratik ve uygun olacağı, madde analizleri sonucu oluşturulmuş ölçek ya da alt ölçeklerin incelenmesi aşamasında ise YEM yöntemlerinden yararlanmanın daha uygun olacağı bildirilmiştir. Bu bağlamda yapılan çalışmanın özelliklerine ve amacına göre tek tek MYT ya da YEM yöntemleri kullanılabileceği gibi, her iki yöntemden birlikte birbirini destekleyici şekilde yararlanmanın da uygun olabileceği vurgulanmıştır (Somer ve ark., 2009).

Bu tez çalışması kapsamında da MSDA'nın cinsiyet ve hastalık süresi bakımından ölçme eşdeğerliğine sahip olduğu her iki yöntem ile de doğrulanmıştır. Yöntemlerden hangisinin tercih edilmesi gerektiği, daha kapsamlı olarak gerçekleştirilecek benzetim çalışmaları ile ortaya konabilir.

5. SONUÇ VE ÖNERİLER

Sağlık alanında olduğu gibi birçok alanda (eğitim, psikoloji, v.b), bireylerin θ düzeyleri; geçerli, güvenilir ve değişime duyarlı ölçme araçları ile doğru bir şekilde belirlenebilir. Geçerlilik kapsamında değerlendirilmesi gereken bir diğer önemli husus, benzer özellikteki bireylerin ölçme aracında yer alan maddelere benzer yanıtlar vermesi, yani bireylerin θ düzeyleri sabit tutulduğunda bu bireylerin madde(ler) üzerindeki başarı olasılıklarının veya olabirliklerinin benzer olup olmadığının incelenmesidir.

Ölçme eşdeğerliğinin değerlendirilmesinde ÇG-DFA ve Rasch Analizi yaygın kullanılan yöntemler arasındadır. Bu yöntemlerin işleyiş açısından bazı benzerlik ve farklılıkları bulunmaktadır.

Bu tez çalışmasında elde edilen sonuçlar irdelendiğinde MSDA'nın cinsiyet ve hastalık süresi bakımından ölçme eşdeğerliğine sahip olduğu her iki yöntem ile de doğrulanmıştır. Ölçme araçları, farklı gruplarda yer alan bireylere uygulandığında, gruplar arası karşılaştırmaların anlamlı olabilmesi için ölçme aracının ölçtüğü yapının gruplarda eşdeğer olduğunu test eden bu iki yöntemin birlikte kullanılmasının, elde edilen sonuçların karşılaştırılabilmesine olanak sağlaması bakımından önemli olduğu düşünülmektedir.

ÖZET

Ölçme Eşdeğerliğinin Çoklu Grup Uygulamaları ve Rasch Analizi ile Değerlendirilmesi

Ölçek, doğrudan ölçüm yapılamayan, gizli (latent) özelliklerin ölçülmesinde kullanılan, psikometrik özellikleri (geçerlik, güvenirlik, kullanışlılık, değişime duyarlılık) belirlenmiş ölçme aracı olarak tanımlanabilir. Geçerlilik, ölçme aracının ölçülecek özelliği, tam ve doğru olarak, başka özelliklerle karıştırmadan ölçebilmesidir. Geçerlilik kapsamında değerlendirilmesi gereken bir diğer önemli husus, benzer özellikteki bireylerin ölçme aracında yer alan maddelere benzer yanıtlar vermesidir. Ölçme eşdeğerliği olarak tanımlanan bu özellik, farklı gruplardaki bireylerin incelenen özellik düzeyleri sabit tutulduğunda madde(ler) üzerindeki başarı olasılıklarının veya olabilirliklerinin benzer olmasıdır.

Ölçme eşdeğerliği, farklı yaklaşımlar kullanılarak değerlendirilmesine rağmen, faktör analizi kapsamında “Çoklu Grup Doğrulamalı Faktör Analizi (ÇG-DFA)” ve modern test teorisi kapsamında ise “Rasch Analizi” yaygın olarak kullanılan yöntemlerdir. Bu tez çalışması kapsamında sağlık alanında geçerli bir ölçüm aracı olan Modifiye Sağlık Değerlendirme Anketi (MSDA)’nin cinsiyet ve hastalık süresi (≤ 10 yıl, > 10 yıl) yönünden ölçme eşdeğerliğine sahip olup olmadığı ÇG-DFA ve Rasch analizi ile değerlendirilmiştir. Elde edilen sonuçlara göre, MSDA’nın cinsiyet ve hastalık süresi bakımından ölçme eşdeğerliğine sahip olduğu her iki yöntem ile de doğrulanmıştır.

İşleyiş açısından benzer ve farklı yönleri bulunan bu iki yöntemin, ölçme eşdeğerliğinin test edilmesinde birlikte kullanılmasının, elde edilen sonuçların karşılaştırılabilmesine olanak sağlaması bakımından önemli olduğu düşünülmektedir.

Anahtar Kelimeler: Ölçme eşdeğerliği, madde işlev farklılığı, çoklu grup doğrulamalı faktör analizi, Rasch analizi

SUMMARY

The Evaluation of Measurement Equivalence via Multiple Group Analysis and Rasch Analysis

The scale can be defined as a measurement tool used in the assessing the latent characteristics. Its psychometric properties; those are validity, reliability, usefulness and responsiveness, are confirmed. Validity is the ability of the measurement tool to completely and accurately measure the property to be scaled without interfering with other properties. The other important thing that should be evaluated in terms of validity is whether individuals with the same latent trait respond to the items in the same manner. This feature, defined as the measurement equivalence, means that for the individuals with the same latent trait from different groups, the probability of success or likelihood on the item(s) will be same.

Measurement equivalence, although assessed using different approaches, the scope of factor analysis "Multiple Group Confirmatory Factor Analysis (MG-CFA)" and the scope of the modern test theory "Rasch Analysis" are commonly used methods. In this study, the measurement equivalence of the Modified Health Assessment Questionnaire (MHAQ), which is a valid measurement tool in the health field, has been evaluated by "MG-CFA" and by "Rasch Analysis" in terms of gender and disease duration (≤ 10 years, > 10 years). According to the results obtained, it has been confirmed with both methods that MHAQ has measurement equivalence for both gender and disease duration.

Although MG-CFA and Rasch analysis have similar and different aspects, using them together to test the measurement equivalence is considered important in terms enabling us to make a comparison between the results obtained.

Key Words: Measurement equivalence, differential item functioning, multiple-group confirmatory factor analysis, Rasch analysis

KAYNAKLAR

- ANDRÉ S. (2014). Does Trust Mean the Same for Migrants and Natives? Testing Measurement Models of Political Trust with Multi-group Confirmatory Factor Analysis. *Social Indicators Research*, 115(3): 963-82,
- ASİL, M. (2010). Uluslararası Öğrenci Değerlendirme Programı (PISA) 2006 Öğrenci Anketinin Kùltürler Arası Eşdeğerliğinin İncelenmesi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı, Doktora Tezi.
- CHEN, F.F. (2008). What Happens If We Compare Chopsticks With Forks? The Impact of Making Inappropriate Comparisons in Cross-Cultural Research. *Journal of Personality and Social Psychology*, Vol. 95, No. 5, 1005–1018.
- CHEUNG, G.W., RENSVOLD, R.B. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance. *Structural Equation Modeling*, 9(2) : 233–255.
- COSTA, D. S. J., AARONSON, N. K., FAYERS, P. M., PALLANT, J. F., VELİKOVA, G., KİNG M. T. (2014). Testing the measurement invariance of the EORTC QLQ-C30 across primary cancer sites using multi-group confirmatory factor analysis. Switzerland: Springer International Publishing,
- ELHAN, A.H., ÖZTUNA, D., KUTLAY, Ş., KÜÇÜKDEVECİ, A.A., TENNANT A. (2008). An initial application of computerized adaptive testing (CAT) for measuring disability in patients with low back pain. *BMC Musculoskeletal Disorders*, 9:166.
- ERDOĞAN, B. (2012). Çoklu Atama Yöntemlerinin Rasch Modelleri İçin Performansının Benzetim Çalışması İle İncelenmesi. Ankara Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı, Doktora Tezi.
- GÖKÇE, E., ÖZTUNA, D., ELHAN, A.H. (2011). Adaptation of Harter’s Scale of Intrinsic versus Extrinsic Motivational Orientation in Classroom to the Primary Schools in Turkey. *Egit Arast*; 42:79-94.
- KÜÇÜKDEVECİ, A.A., ŞAHİN, H., ATAMAN, S., GRİFFİTHS, B., TENNANT, A.(2004). Issues in Cross-Cultural Validity: Example from The Adaptation, Reliability, and Validity Testing of a Turkish Version of the Stanford Health Assessment Questionnaire. *Arthritis Rheum.*,51(1):14-9.
- MUNYOMBWE, T., WEST, R.M , HİLL, K. . Testing measurement invariance of the GHQ-28 in stroke patients. *Qual Life Res*. 2015 Jan 22. [Epub ahead of print]
- MUTHEN&MUTHEN (2006). MPlus, Version 6. Los Angeles, CA.
- ÖĞRETMEN, T. (2006). Uluslararası Okuma Becerilerinde Gelişim Projesi (PIRLS) 2001 Testinde Psikometrik Özelliklerinin İncelenmesi: Türkiye Amerika Birleşik Devletleri Örneği, Hacettepe Üniversitesi,Sosyal Bilimler Enstitüsü, Eğitim Bilimleri Anabilim Dalı, Doktora Tezi
- ÖZDAMAR, K. (2002).Paket Programları İle İstatistiksel Veri Analizi- 2 (Çok Değişkenli Analizler), Kaan Kitapevi Eskişehir.
- ÖZTUNA, D. (2008). Kas-İskelet sistemi Sorunlarının Özürlülük Değerlendiriminde Bilgisayar Uyarlamalı Test Yönteminin Uygulanması. Ankara Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı, Doktora Tezi.

- PİNCUS, T., SUMMEY, J.A., SORACI, S. A. , WALLSTON, K. A. (1983). Hummon NP. Assessment of patient satisfaction in activities of daily living using a modified Stanford Health Assessment Questionnaire. *Arthritis Rheum*, 26: 1346-53.
- RAJU, N.S., LAFFİTTE, L.J., BYRNE, B.M. (2002). Measurement Equivalence: A Comparison of Methods Based on Confirmatory Factor Analysis and Item Response Theory. *Journal of Applied Psychology*, Vol. 87, No. 3, 517–529.
- SİMSEK, Ö.F. (2007). Yapısal Eşitlik Modellemesine Giriş: Temel İlkeler ve LISREL Uygulamaları. İstanbul: Ekinoks Yayıncılık.
- SOMER, O., KORKMAZ, M., DURAL, S., CAN, S. (2009). Ölçme Eşdeğerliğinin Yapısal Eşitlik Modellemesi ve Madde Cevap Kuramı Kapsamında İncelenmesi. *Türk Psikoloji Dergisi*, 24(64): 61-75,.
- TEKİN, H. (2002). Eğitimde Ölçme ve Değerlendirme. Ankara: Yargı Yayıncılık.
- TURGUT, F. (1977). Eğitimde Ölçme ve Değerlendirme Metodları. Ankara: Nüve Matbaası.
- Use of CFA with Multiple Groups. Erişim Tarihi: (02.03.2015) <http://ebooks.narotama.ac.id/files/Confirmatory%20Factor%20Analysis/Chapter%205%20Use%20of%20Confi%20rmatory%20Factor%20Analysis%20with%20Multiple%20Groups.pdf>
- UYAR, Ş. (2011). PISA 2009 Türkiye Örnekleminde Öğrenme Stratejileri Modelinin Farklı Gruplarda Ölçme Değişmezliğinin İncelenmesi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı, Yüksek Lisans Tezi.
- WOLFE, F. (2001). Which HAQ is best? A comparison of the HAQ, MHAQ and RA-HAQ, a difficult 8 item HAQ (DHAQ), and a rescored 20 item HAQ (HAQ20): analyses in 2,491 rheumatoid arthritis patients following leflunomide initiation. *J Rheumatol.*, 28(5):982-9.
- WU, AMERY, D., Lİ, ZHEN & ZUMBO, BRUNO, D. (2007). Decoding the Meaning of Factorial Invariance and Updating the Practice of Multi-group Confirmatory Factor Analysis: A Demonstration with TIMSS Data. *Practical Assessment Research & Evaluation*, Vol. 12, No. 3, ISSN 1531-7714.
- YAN, Y., WU, W., STRUNK, R.C., GARBUTT, J. (2014). Use of factor analysis models to evaluate measurement invariance property of the Asthma Control Questionnaire (ACQ). *Qual Life Res.*, 23: 509-13.
- YÜKSEL, S. (2012). Ölçeklerde Saptanan Madde İşlev Farklılığının Karma Rasch Modelleri İle İncelenmesi. Ankara Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı, Doktora Tezi.

ÖZGEÇMİŞ

1.Bireysel Bilgiler

Adı ve Soyadı : Serhat Hayme
Doğum Yeri ve Tarihi : Çanakkale - 06.09.1986
Uyruğu : T.C.
Medeni Durumu : Bekar
İletişim Adresi ve Telefonu : Ankara Üniversitesi Tıp Fakültesi Biyoistatistik AD.
 Morfoloji Kampüsü 06700 Sıhhiye/ANKARA
 Tel: 05054529330

II- Eğitimi

Yüksek Lisans

Ankara Üniversitesi Tıp Fakültesi Biyoistatistik A.D
 Başlangıç/Bitiş Yılı:2012/2015

Lisans

Uludağ Üniversitesi Fen-Edebiyat Fakültesi Matematik Bölümü
 Mezuniyet yılı: 2010

Lise

Ziya Gökalp Y.D.A. Lisesi
 Mezuniyet Yılı: 2004

Yabancı Dili: İngilizce (ÜDS 75)

III-Ünvanları

2009 - 2010: Kültür Koleji ve Eğitim Kurumlarında Matematik Öğretmeni
 2010- 2011: Murat Eğitim Kurumlarında Matematik Öğretmeni
 2014 -2015: Araştırma Görevlisi; Ankara Üniversitesi Tıp Fakültesi, Biyoistatistik
 Anabilim Dalı

IV- Mesleki Deneyimi

ANKARA ÜNİVERSİTESİ TIP FAKÜLTESİ BİYOİSTATİSTİK ANABİLİM DALI

Pozisyon: 2014 yılından itibaren “Araştırma Görevlisi“ olarak çalışmaktayım.

Özet İş Tanımı: Tıp Fakültesi 2. ve 3. sınıf öğrencilerine Biyoistatistik dersi kapsamında uygulama yaptırılması, Ankara Üniversitesi ve diğer üniversite, kurumlardan gelen araştırmacıların çalışmalarının istatistiksel analizlerinin yapılması, istatistiksel danışmanlık verilmesi

Kültür Eğitim Kurumları

Pozisyon: 2009-2010 yılları arasında bu kurumda, ortaöğretim ve ilköğretim 2.kademe gruplarında matematik öğretmeni olarak çalıştım.

Özet İş Tanımı: Öğrencileri ÖSYM sınavlarına hazırlamak ve lise matematik derslerinde gelişim göstermelerini sağlamak, eğitim koçluğu yapmak.

Murat Eğitim Kurumları

Pozisyon: 2010-2011 yılları arasında KPSS grupları matematik öğretmeni olarak çalıştım.

Özet İş Tanımı: KPSS sınavına hazırlanan öğrencileri sınava yönelik hazırlamak ve danışmanlık etmek.

V- Üye Olduğu Bilimsel Kuruluşlar

- 1)Biyoistatistik Derneği
- 2)Türk Matematik Derneği

VI- Bilimsel İlgi Alanları

Ulusal Kongrelerde Sunulan Bildiriler ve Posterler:

1964-1994 Tarihleri Arasında Bildirilen Kızamık Vakalarından Olan Ölümünün Zamana Bağlı İncelenmesi (XIV. Uluslararası Katılımlı Biyoistatistik Kongresi Kayseri -2013)

Seminerler

Temel Bileşenler Lojistik Regresyon Analizi (Hacettepe Üniversitesi,Tıp Fak., Biyoistatistik A.B.D)

VII- Diğer

Sertifika ve Başarı Belgeleri

- 1-Bilgisayar İşletmenliği Sertifikası
- 2.İleri Excel Kullanıcılığı Sertifikası
- 3.Yapısal Eşitlik Modelleri (YEM) Başarı belgesi /Eskişehir Osmangazi Üniversitesi
- 4.XIV. Uluslararası Katılımlı Biyoistatistik Kongresi Poster Bildirim 3.'lüğü

EKLER**EK-1**

Tez çalışmasında kullanılan veri seti ile ilgili izin yazısı

23/03/2015

İLGİLİ MAKAMA,

TÜBİTAK 1001 kapsamında 2010-2012 yılları arasında gerçekleştirilen ve yürütücülüğünü yapmış olduğum “109S342” kodlu ve “ Romatoid Artritli Hastalarda Bilgisayar Uyarlamalı Test Yöntemi ile Özürlülük Değerlendirmesi” başlıklı proje kapsamında romatoid artrit hastalarından elde edilen Sağlık Değerlendirme Anketi verilerinin, Anabilim Dalımız Yüksek Lisans öğrencilerinden Serhat HAYME tarafından gerçekleştirilecek “Madde İşlev Farklılığının Çoklu Grup Uygulamaları ve Rasch Analizi ile Değerlendirilmesi” başlıklı tez çalışması için kullanılmasına izin veriyorum.

Saygılarımla



Prof. Dr. Atilla Halil ELHAN

Adres: Ankara Üniversitesi Tıp Fakültesi

Biyoistatistik Anabilim Dalı

Telefon: 5958196

E-mail: ahelhan@yahoo.com