

## KLASİK TEST KURAMI VE ÖRTÜK ÖZELLİKLER KURAMININ ÖRNEKLEMLER BAĞLAMINDA KARŞILAŞTIRILMASI

### COMPARISON OF CLASSICAL TEST THEORY AND LATENT TRAITS THEORY BY SAMPLES

Nuri DOĞAN\* A.Ata TEZBAŞARAN\*\*

**ÖZET:** Bu çalışmada, örtük özellikler kuramının, klasik kurama üstünlüğü olarak kabul edilen madde parametrelerinin değişmezliğini sağlayıp sağlamadığı incelenmiştir. Elde edilen sonuçlara göre, örtük özellikler kuramının madde parametrelerinin değişmezlik özelliğini sağlayamadığı; aksine, klasik test kuramının örtük özellikler kuramı kadar kararlı sonuçlar verdiği görülmüştür.

**Anahtar Sözcükler:** klasik test kuramı, örtük özellikler kuramı, parametrelerin değişmezliği

**ABSTRACT:** In this study, it is examined whether the latent traits theory provides the invariance of item parameters, which is accepted as the advantage of latent traits theory over classical theory. Finally, it is found that the latent traits theory could not provide the invariance of item parameters; on the contrary, the classical test theory brings outcomes as invariance as the latent traits theory does.

**Key Words:** classical test theory, latent Traits Theory, invariance of parameters

#### 1. GİRİŞ

Duyarlı ölçmeler yapılabilmesi için nitelikli ölçme araçları kullanılmalıdır. Eğitim ve psikoloji alanlarında nitelikli ölçme aracı elde etmek amacıyla çeşitli kuramlar geliştirilmiş ve kullanılmıştır. Bunlardan klasik test kuramı (KTK) ve örtük özellikler kuramı (ÖÖK) yaygın olarak kullanılmaktadır (Keeves 1988).

##### 1.1. Test Geliştirme

Testlerin, alınacak kararlar için uygunluğu ve yeterince duyarlı ölçme yapıp yapmadığını belirlemek için, geçerlik, güvenilirlik, kullanışlılık ve diğer bakımlardan sağlamlığı saptanmalıdır. Bir

ölçeğin sağlamlığı, psikometrik özelliklerinin yeterliği ile ölçülür. Aracın psikometrik özellikleri, test ve madde analizleri ile belirlenir.

İstenilir nitelikte bir test oluşturmak kaliteli madde seçme yoluyla sağlanabilir. Maddenin kalitesini belirlemenin yolu madde analizidir. Madde analizi; istenilir nitelikte bir test oluşturma için maddelere ait istatistik özelliklerin belirlenmesi, madde düzeltme ve madde seçme (ayıklama) işlemlerini kapsar. Temel yöntem bütün kuramlar için bu olduğundan, madde analizi test geliştirme için bu olduğundan, madde analizi test geliştirme için temel kabul edilebilir.

Birçok test geliştirme uzmanına göre, madde analizinin hareket noktası madde gücü, ayırıcılığı ve cevapların seçeneklere dağılımıdır. Bunlar, bir maddenin işleyişi ile ilgili bilinmesi zorunlu göstergelerdir.

##### 1.2. Klasik Test Kuramı

Ölçmenin amacı, ölçülen özelliğin gerçek değerine yaklaşmaktır. Çünkü, ölçmelere çeşitli hatalar karışır. Bu nedenle gerçek değeri elde etmek pratikte sağlanamaz; gözlenen puanlardan kestirilmeye çalışılır. Klasik test kuramı gerçek puanı gözlenen puandan kestirmeye çalışan kuramlardan biridir. Klasik test kuramına gerçek puan modeli de denilmektedir. Klasik test kuramının temel denklemi şöyle ifade edilmektedir.

\* Araştırma Görevlisi, Hacettepe Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü – Ankara

\*\* Doç.Dr. Hacettepe Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü – Ankara

X: Gözlenen puan

T: Gerçek puan

E: Random hatalar

Kuramın varsayımı, bir bireyin özellikleri sabittir ve gözlenen puanlardaki değişim, yorgunluk veya şans gibi çeşitli faktörlerin sonucu olan random hatalardan kaynaklanır (Crocer & Algina 1986).

Klasik test kuramında test ve madde istatistiklerinin hareket noktası madde güçlük düzeyi ( $p$ ) ve madde ayıricılığıdır ( $r_{jx}$ ). Bu istatistikler yardımıyla teste ait özellikler kestirilebilir (Osterlid 1989).

Klasik test kuramı bir takım sınırlılıklara sahiptir. Bunlardan ikisi, temel sınırlılık olarak ifade edilebilir.

1. Klasik test kuramında madde ve test istatistikleri elde edildiği örnekleme bağımlıdır. Bir örneklemedeki yetenek düzeyi ve dağılımı, klasik test kuramının madde güçlük indeksi ve madde ayıricılık indeksi gibi madde istatistiklerini ve test istatistiklerini önemli ölçüde etkilemektedir.
2. Bireylerin yetenekleri kendilerine uygulanan maddelere bağımlıdır. Bir birey, aynı özelliği ölçen farklı güçlükteki testlerden farklı puanlar alabilir (Hambleton & Swaminathan 1985; Yu 1999).

### 1.3. Örtük Özellikler Kuramı

Klasik test kuramının sınırlılıklarını giderebileceği iddia edilen bir model olarak 1930'lu yılların sonunda örtük özellikler kuramı ortaya atılmıştır. Tucker 1940'da örtük özellik kuramının en önemli kavramlarından "madde karakteristik eğrisi" kavramını ilk defa kullanmıştır. Ancak, Lord'un 1950'li yılların başındaki çalışmaları örtük özellikler kuramının başlangıcı kabul edilir. Lord, ilk kez bu kuram doğrultusunda "normal ogive" modelini geliştirip, başarı ve yetenek testlerinde kullanmıştır. 1950'li yılların sonunda ise Birnbaum, "normal ogive" modelinin yerine bu alanda çalışmaları kolaylaştıran "lojistik" modeli geliştirmiştir. 1950-1970 yılları arasında, matematiksel işlemlerinin karmaşıklığı nedeniyle kuram yavaş gelişmiş, ancak bilgisayar teknolojisindeki

gelişmelere paralel olarak kuramın kullanımı ve gelişimi de hızlanmıştır (Hambleton & Swaminathan 1985:4-8; Baker 2001:4).

Örtük özellikler kuramı, örneklemeden bağımsız madde ölçeklemesini ve maddeden bağımsız olarak yetenek kestirimini matematiksel modellerle olanaklı hale getirmek iddiasındadır. Kuramın temeli, bireylerin yeteneklerini gösteren psikolojik uzayla, bu uzayı temsil etmek amacıyla hazırlanan maddeler arasındaki bağıntılardır. Kurama göre, kişilerin özellikleri bir  $q$  uzayından oluşur ve çalışılan özelliğin ölçülmesi amacıyla hazırlanan maddelerden oluşturulmuş testin ise  $q$  uzayını temsil ettiği kabul edilir.

Örtük özellikler kuramında bireyin performansı, yetenek ( $q$ ) faktörüyle belirlenmektedir ve bireyin performansı ile yetenek ( $q$ ) arasındaki ilişki madde karakteristik fonksiyonu (MKF) adı verilen fonksiyonla ifade edilmektedir. Madde karakteristik fonksiyonu, madde puanının  $q$  vektörü üzerindeki regresyonudur ve  $P(q)$  ile gösterilir.  $P(q)$ , tanım gereği bir gruptan diğerine değişmez. Böylece  $P(q)$ 'yı belirleyen parametreler de sabit parametreler olarak kalır (Lord & Novic 1968; Yu 1999).

Örtük özellikler kuramının varsayımları; normal dağılım, tek boyutluluk ve yerel bağımsızlıktır. Veriler bu varsayımları sağladığında örtük özellikler kuramını kullanmanın doğru olacağı ve avantajların gerçekleşeceği belirtilmektedir. Ancak bu varsayımların sağlanması çok zordur.

Örtük özellikler kuramında kullanılan modellerdeki farklılık, madde karakteristik eğrisinin formülasyonundan veya puanlamadan ileri gelmektedir. Tek boyutlu ve 1-0 şeklinde puanlanan verilerde kullanılan modeller için, madde karakteristik fonksiyonlarının içerdiği madde parametrelerinin sayısına göre tanımlar yapılmıştır. Bir, iki ve üç parametre kullanan farklı modeller bulunmaktadır (Hambleton & Swaminathan 1985; Hambleton & diğerleri 1991).

Bir parametrelili model, madde tepki işleyişini betimlemek için bir parametrenin yeterli olduğunu varsayar. Bu da madde güçlüğünü ( $p$ ) karşılayan "b" parametresidir. Maddenin en iyi ölçtüğü yete-

nek düzeyini de gösterir. Bir parametreliliğin en önemli sınırlılığı bütün maddelerin eşit ayırıcılıkta olduğu varsayımdır. İki parametreliliğin bir parametreliliğe ikinci parametreyi eklemiş olmasıdır. Bu da "a" parametresidir ve madde ayırıcılığına (r<sub>yx</sub>) karşılık gelir. İki parametreliliğin en önemli sınırlılığı şansa doğru cevap verme olasılığını dikkate almamasıdır. Üç parametreliliğin lojistik model, bir ve iki parametreliliğe parametresi de denilen "c" parametresini eklemiştir. Şansın etkili olduğu testlerin analiz edilmesinde üç parametreliliğin kullanılması önerilmektedir (Lord ve Novic 1968; Hambleton & Swaminathan 1985)

Varsayımlar ve model-veri uyumu sağlandığı takdirde örtük özellikler kuramını klasik test kuramından üstün kılan bazı avantajların olduğu belirtilmektedir. Örtük özellikler kuramının avantajları içinde parametrelerin değişmezliği önemli yer tutar. Kuramın diğer avantajlarının parametrelerin değişmezliği avantajına bağlı olduğu söylenebilir. Değişmezlik iki açıdan ele alınabilir. Birincisi; aynı özelliği ölçmeye yönelik olarak hazırlanmış olan farklı maddelere verilen tepkilere dayalı olarak kestirilen yetenek parametreleri değişmezdir. İkincisi; aynı maddelerin farklı bireylere uygulanmasıyla elde edilen madde parametreleri değişmezdir (Hambelton & Swaminathan 1985).

Bu avantajların kuramın yapısından geldiği iddia edilmektedir. Bir kişinin soruya doğru cevap verme olasılığı, madde karakteristik eğrisinin şekline bağlı olduğundan, kişinin yeteneği grubun yetenek dağılımından bağımsız olacaktır. Aynı şekilde madde parametrelerinin değeri de madde karakteristik eğrisinin şekline bağlı olduğundan madde parametreleri de gruptan bağımsız olarak kestirilebilir.

Klasik kuramın bir uzantısı olması, işlemlerin zorluğu ve karşılanması zor varsayımlarının bulunması Örtük özellikler kuramının önemli sınırlılıklarındandır.

Psikoloji ve eğitim gibi davranış bilimlerinin hem kuramsal hem de deneysel taraflarından söz edilebilir. Bir bilim dalı olarak eğitimbilim içinde-

ki kuramların da deneysel yollarla kanıtlanması gerekmektedir.

Örtük özellikler kuramının en önemli avantajı olan parametrelerin değişmezliği, ilgili araştırmalarda, yeterli kanıtlar sağlanmadığı için (Fan 1998; Hambleton ve diğerleri 1991; Somer 1998; Stage 1998; Nartgün 2002), tartışmalı bir konu olmaya devam etmektedir. Parametrelerin değişmezlik özelliğinin denetlenmesine yönelik araştırmalar yapılması önemli bir ihtiyaç olarak görülmektedir.

Bu çalışmada, örtük özellikler kuramının gruptan bağımsız olarak madde parametrelerini kestirme avantajı, örneklem bağlamında incelenmiştir. Böylece, örtük özellikler kuramının parametrelerin değişmezliği bakımından, klasik test kuramından daha iyi işleyip işlemediği saptanabilir.

Bu amaçla;

1. Farklı örneklemelerden klasik test kuramına göre elde edilen p istatistikleri ile evren parametreleri arasındaki,
2. Farklı örneklemelerden klasik test kuramına göre elde edilen r<sub>yx</sub> istatistikleri ile evren parametreleri arasındaki,
3. Farklı örneklemelerden ve evrenden örtük özellikler kuramına göre elde edilen 1, 2 ve 3 parametreliliğe ait b parametreleri arasındaki,
4. Farklı örneklemelerden ve evrenden örtük özellikler kuramına göre elde edilen 2 ve 3 parametreliliğe ait a parametreleri arasındaki ilişkileri saptamak araştırmanın kapsamına alınmıştır.

## 2. YÖNTEM

Araştırma, klasik test kuramı ve örtük özellikler kuramını gerçek verilere bağlı olarak karşılaştıran bir çalışmadır.

Araştırmanın evrenini, Milli Eğitim Bakanlığı Orta Öğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavının 2001 yılı uygulamasına katılan 553281 kişi oluşturmaktadır. Çalışmada, evreni oluşturan cevaplayıcıların sözel bölüm puanları dikkate alınmıştır.

Arařtırma, evrenden seilen beř (5) farklı örneklem üzerinden yürütülmüřtür. Seilen örneklem; random, saęa arpık, sola arpık, basık ve normal daęılım gösteren örneklemlerdir. Örneklem seilirken, evreni temsil edebilirlikleri dikkate alınmıřtır. Bu amala, seilen örneklemlerin büyüklüęü,  $\alpha=0,05$  ve ortalamadan sapma (hořgörü düzeyi) 0,21 alınarak hesaplanmıřtır. Örneklem büyüklüklerinin ilgili formülle (ıngı 1994) hesaplanan deęerden (22336) küçük olmamasına özen gösterilmiřtir.

Örneklem seçiminde random ve kasıtlı örneklem teknikleri kullanılmıřtır. Random örneklem SPSS paket programı yardımıyla elde edilmiřtir. Kasıtlı örneklem ise, amalanan daęılımı saęlayan verilerin örneklem alınmasıyla saęlanmıřtır.

### 2.1. Arařtırma Verileri

Arařtırmada kullanılan veriler, Milli Eęitim Bakanlıęı'nın uyguladıęı, Orta Öęretim Kurumları Öęrenci Seme ve Yerleřtirme Sınavı 2001 yılı sonuçlarıdır. Sınavda, "akademik yeteneęi ölçme aęırlıklı Türke, matematik, fen ve sosyal bilgiler" alanlarını kapsayan, dört seenekli 100 soru bulunmaktadır. Alanlara ait bilgiler 25'er soruyla ölçülmektedir. Testin aynı zamanda sözel ve sayısal akademik yeteneęi ölçen bir test olduęu öne sürülmektedir (MEB 2001b).

### 2.2. Verilerin Analizi

Verilerin analizi üç adımda gerekleřtirilmiřtir.

**Birinci adım:** Evren ve örneklem için betimsel istatistikler hesaplanmıřtır.

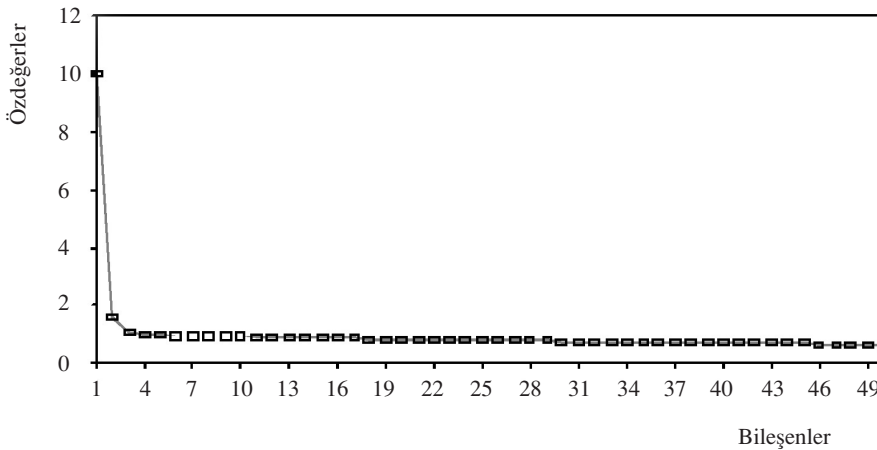
**Tablo 2.1.** Evren ve Örneklemere ait Betimsel İstatistikler

İstatistikler	Evren	Random	SaęaÇ.	Sola Ç.	Basık	Normal
N	552725	27501	23531	23531	24691	29244
Ortalama	26,43	26,40	15,87	35,13	25,00	25,49
Medyan	26,00	26,00	15,00	36,00	25,00	25,00
Mod	24,00	22,00	11,00	40,00	25,00	25,00
S. Sapma	9,05	9,08	6,06	6,06	11,07	6,87
Varyans	81,86	82,38	36,77	36,77	122,65	47,20
arpıklık	0,20	0,20	1,00	-1,00	0,00	0,00
Basıklık	-0,73	-0,73	1,37	1,37	-1,07	0,00
KR-21	0,87	0,87	0,72	0,73	0,92	0,75

Betimsel istatistikler incelendięinde (Tablo 2.1) amalanan daęımların yeterince saęlandıęı söylenebilir. arpık örneklemde, yapılacak karřılařtırmaların isabetlilięini artırmak için arpıklık mutlak deęerleri (1,001) ve basıklık deęerleri (1,370) eřit tutulmuřtur. Basık daęımda ise arpıklık katsayısının 0 olması saęlanmıřtır.

**İkinci adım:** Evrendeki verilerin örtük özellikler kuramı ile alıřmaya uygunluęu test edilmiřtir. Bu amala tek boyutluluk, yerel baęımsızlık ve daęılımın normallięine bakılmıřtır. Sözel puanlara özdeęer-bileřen grafięine (řekil 2.1) göre, testin tek boyutlu olduęu söylenebilir. 1. faktörden sonra özdeęerler birden bire düşmekte, dięer faktörlerin özdeęerleri arasındaki farklar ise azalmaktadır.

Testin tek boyutlu ıkması, örtük özellikler kuramının varsayımlarından tek boyutluluk ve yerel baęımsızlık varsayımlarının karřılandıęını kabul etmek için yeterli bir sonuçtur. Yerel baęımsızlık, tek boyutluluęun bir sonucu olarak görülmektedir. Bu nedenle tek boyutluluk saęlandıęında yerel ba-



**řekil 1.** Sözel Testine Ait Özdeęer - Bileřen Grafięi

ğimsizliğin da sağlandığı kabul edilmektedir (Lord 1980:19, Hambleton ve Swaminathan 1985:25).

Çalışma evreninin, örtük özellikler kuramının normallik varsayımını yeterince karşılamadığını belirtmek gerekir. Ancak, gerçek verilerde normallik varsayımını sağlamanın güçlüğü birçok araştırmacı tarafından kabul edilmiştir. Bu varsayım sağlanmadığı durumlarda da araştırmaya devam etmenin mümkün olduğu bildirilmektedir. Çünkü gerçek veriler, genel olarak hipotetik varsayımları sağlamaktan uzaktır. Kuramların gerçek durumlarda sınanması esasa alındığından, birçok araştırmacı, bu varsayım sağlanmadığı halde araştırmalarına devam etmiştir (Lord 1980; Hambleton ve Swaminathan 1985; Gelbal 1994; Kelecioğlu 2001).

**Üçüncü adım:** Madde istatistikleri ve parametreleri hesaplanmıştır. Araştırma problemlerine bağlı olarak madde parametreleri, madde istatistikleri ve evren değerleri arasındaki ilişki için Spearman'ın sıra farkları korelasyonu, iki korelasyon farkının Fisher z'si ile test edilmesi tekniklerinden yararlanılmıştır.

### 3. BULGULAR VE YORUM

Bulgular araştırma problemlerine bağlı olarak sıralanmıştır.

1- Farklı örneklemelerden klasik kurama göre elde edilen madde güçlük indeksleri ile evrene ait madde güçlük parametreleri arasındaki ilişkiler araştırılmış ve sonuçlar Tablo 3.1'de verilmiştir.

Tablo 3.1'de, örneklemeler ile evren arasındaki korelasyonların 1,00 ile 0,95 arasında değiştiği gözlenmektedir. Korelasyonların hepsi pozitif ve

**Tablo 3.1.** Klasik Kurama Göre Evren Ve Örneklemeler İçin Elde Edilen Madde Güçlük (p) Değerleri Arasındaki Korelasyonlar

Örneklemeler	Evren
Random	1,000**
Basık	0,986**
Sağa çarpık	0,945**
Sola çarpık	0,974**
Normal	0,984**

\*\*p < 0,01

$\alpha = 0,01$  düzeyinde manidardır. Random örnekleme ait p değerleri ile evrene ait p değerleri arasında mükemmel bir ilişki vardır. Çarpık dağılımlarda ilişkinin düştüğü söylenebilir.

Evren ve örnekleme ait p değerleri için hesaplanan korelasyonlar arasındaki farkların manidarlığına ilişkin bilgiler Tablo 3.2'de verilmiştir.

**Tablo 3.2.** Klasik Kurama Göre Evren ve Örneklemeler İçin Elde Edilen Madde Güçlük (p) Değerleri Arasındaki Korelasyonların Farkları

Örneklemeler	1	2	3	4
Random (1)	-	-	-	-
Basık (2)	0,014*	-	-	-
Sağa çarpık (3)	0,055**	0,041**	-	-
Sola çarpık (4)	0,026*	0,012	-0,029*	-
Normal (5)	0,016*	0,002	-0,039**	-0,01

\* p < 0,05 \*\* p < 0,01

Tablo 3.2'ye göre, çarpık örneklemeler için elde edilen korelasyonlarla diğer örneklemeler için elde edilen korelasyonlar arasındaki farkların çoğu manidardır. Özellikle sağa çarpık örnekleme diğer örneklemelere ait korelasyon farkları  $\alpha = 0,01$  düzeyinde manidar bulunmuştur. Bu bulguya göre, madde güçlüğünün örneklem dağılımının çarpıklığından, özellikle sağa çarpıklığından etkilendiği yorumu yapılabilir.

2. Farklı örneklemelerden klasik kurama göre elde edilen madde ayırıcılık istatistikleri ile evrene ait madde ayırıcılık parametreleri arasındaki ilişkilere bakılmış ve sonuçlar Tablo 3.3'de verilmiştir.

**Tablo 3.3.** Klasik Kurama Göre Evren ve Örneklemeler İçin Elde Edilen Madde Ayırıcılık ( $r_{jx}$ ) Değerleri Arasındaki Korelasyonlar

Örneklemeler	Evren
Random	0,998**
Basık	0,927**
Sağa çarpık	0,678**
Sola çarpık	0,945**
Normal	0,884**

\*\*p < 0,01

Tablo 3.3 de, korelasyonların 0,998 ile 0,678 arasında değiştiği görülmektedir. Korelasyonların hepsi pozitif yönde ve  $\alpha = 0,01$  düzeyinde manidar bulunmuştur. En küçük korelasyon sağa çarpık örnekleme ait ayırıcılık değerleri ile evren ayırıcılık parametreleri arasında; en büyük korelasyon random örnekleme ait ayırıcılık değerleri ile evren ayırıcılık parametreleri arasında gözlenmektedir. Korelasyon değerleri sadece çarpık dağılımlara gidildikçe düşmemektedir. Normal dağılımlı örneklem için elde edilen madde ayırıcılık değerleri ile evren için elde edilen madde ayırıcılık değerleri arasındaki korelasyonlar da düşmektedir. Ayırıcılık değerlerine ait korelasyonlardaki değişkenlik p değerleri için elde edilen korelasyonlardaki değişkenlikten daha fazladır.

Evren ve örneklere ait ayırıcılık güçlerine ait korelasyonlar arasındaki farkların manidarlığına ilişkin bilgiler Tablo 3.4'de bulunmaktadır.

Tablo 3.4'e göre, evren ile örneklere ait madde ayırıcılık değerleri arasında hesaplanan korelasyonlar arasındaki farklardan çoğu  $\alpha = 0,05$  ve  $\alpha = 0,01$  düzeylerinde manidardır. Basık ile sola çarpık örnekleme ait korelasyonların farkı ve basık ile normal dağılımlı örnekleme ait korelasyonların farkı manidar bulunmamış; normal dağılımlı örnekleme sola çarpık örnekleme ait korelasyonların farkı  $\alpha = 0,05$  düzeyinde; diğer korelasyon farkları  $\alpha = 0,01$  düzeyinde manidar bulunmuştur. Bu bulgulara göre, klasik kurama göre farklı örneklemelerden elde edilen  $r_{jx}$ 'ler arasındaki korelasyonlar değişme göstermektedir. Diğer bir deyişle klasik kurama göre elde edilen  $r_{jx}$  değerleri örneklem değişmelerinden etkilenmektedir.

**Tablo 3.4:** Klasik Kurama Göre Evren ve Örneklem İçin Elde Edilen Madde Güçlük (p) Değerleri Arasındaki Korelasyonların Farkları

Örneklem	1	2	3	4
Random (1)	-	-	-	-
Basık (2)	0,071**	-	-	-
Sağa çarpık (3)	0,320**	0,249**	-	-
Sola çarpık (4)	0,053**	-0,018	-0,267**	-
Normal (5)	0,114**	0,043	-0,206**	0,061*

\* p < 0,05 \*\* p < 0,01

3- Farklı örneklemelerden örtük özellikler kuramının 1, 2 ve 3 parametrelili modellerine göre elde edilen b parametreleri ile evrene ait b parametreleri arasındaki ilişkilere bakılmış ve sonuçlar Tablo 3.5'de verilmiştir.

Tablo 3.5'e göre, örtük özellikler kuramı 1, 2 ve 3 parametrelili modellere bağlı olarak evren ve örneklem için hesaplanan b parametreleri arasındaki korelasyonların hepsi pozitif yönde ve  $\alpha = 0,01$  düzeyinde manidardır. Korelasyonlar 0,998 ile 0,868 arasında değişme göstermektedir. En yüksek korelasyon 1 parametrelili modele göre hesaplanan evren ve random örnekleme ait b değerleri arasında bulunmuştur. En düşük korelasyon ise 3 parametrelili modele göre hesaplanan evren b parametreleri ile 1 parametrelili modele göre hesaplanan sağa çarpık örnekleme ait b parametreleri arasında bulunmuştur. Özellikle çarpık dağılımlar için hesaplanan b değerleri ile evren için hesaplanan b değerleri arasındaki korelasyonların düştüğü gözlenmektedir. Dikkati çeken nokta, korelasyonların büyüklükleri örneklere bağlı olarak değişmekle birlikte, bu değişme sistematik değildir. Bununla birlikte, korelasyon değerlerinin modellere bağlı olarak sistematik bir değişme gösterdikleri söylenebilir. Buna göre, evren ve örneklem için aynı modellerle elde edilen b parametreleri arasında korelasyonlar yüksek bulunurken, farklı modellerden elde edilen b parametreleri arasındaki korelasyonların düştüğü söylenebilir.

Örtük özellikler kuramı 1, 2 ve 3 parametrelili modellere bağlı olarak, evren ve örneklem için hesaplanan b parametrelerine ait korelasyon katsayıları arasındaki farkların çoğunun  $\alpha = 0,01$  ve  $\alpha = 0,05$  düzeyinde manidar olduğu tespit edilmiştir (EK 1). Ancak, 1 parametrelili modelden 3 parametrelili modele gidildikçe korelasyon farklarının manidarlığında azalma olduğu; 3 parametrelili modelde parametrelerin değişmezliğinin diğer modellere göre daha fazla sağlandığı söylenebilir. Korelasyonlar ve korelasyon farkları incelendiğinde evren ve örneklem için değişmeyi azaltmak amacıyla aynı modelin kullanılmasının yararlı olacağı görülmektedir. Bu bulgulardan b parametresinin örnek-

**Tablo 3.5.** 1, 2 ve 3 Parametrelili Modellere Göre Evren ve Örneklem İçin Elde Edilen b Parametreleri Arasındaki Korelasyonlar

Örneklem	Evren 1P b	Evren 2P b	Evren 3P b
Random 1P b	0,998**	0,982**	0,944**
Random 2P b	0,983**	0,996**	0,952**
Random 3P b	0,934**	0,956**	0,980**
Basık 1P b	0,993**	0,975**	0,942**
Basık 2P b	0,988**	0,981**	0,952**
Basık 3P b	0,962**	0,956**	0,949**
Sağa Ç. 1P b	0,950**	0,928**	0,868**
Sağa Ç. 2P b	0,958**	0,944**	0,938**
Sağa Ç. 3P b	0,951**	0,944**	0,953**
Sola Ç. 1P b	0,979**	0,954**	0,933**
Sola Ç. 2P b	0,901**	0,877**	0,896**
Sola Ç. 3P b	0,938**	0,956**	0,950**
Normal 1P b	0,989**	0,966**	0,940**
Normal 2P b	0,967**	0,962**	0,944**
Normal 3P b	0,954**	0,965**	0,972**

\*\* p&lt;0,01

lemlere ve modellere göre farklılaştığı yorumunu çıkarmak mümkündür.

4- Farklı örneklemelerden örtük özellikler kuramının 2 ve 3 parametrelili modellerine göre elde edilen a parametreleri ile evrene ait a parametreleri arasındaki ilişkiler araştırılmış ve sonuçlar Tablo 3.6'da verilmiştir.

**Tablo 3.7.** 2 Parametrelili Modele Göre Evren ve 2, 3 Parametrelili Modellere Göre Örneklem İçin Elde Edilen a Değerleri Arasındaki Korelasyonların Farkları

Örneklem	1	2	3	4	5	6	7	8	9
Random 2P a (1)	-	-	-	-	-	-	-	-	-
Random 3P a (2)	0,523**	-	-	-	-	-	-	-	-
Basık 2P a (3)	0,082**	-0,441**	-	-	-	-	-	-	-
Basık 3P a (4)	0,257**	-0,266*	0,175**	-	-	-	-	-	-
Sağa çarpık 2P a (5)	0,257**	-0,266*	0,175**	0	-	-	-	-	-
Sağa çarpık 3P a (6)	0,664**	0,141	0,582**	0,407**	0,407**	-	-	-	-
Sola çarpık 2P a (7)	0,032**	-0,491**	-0,05*	-0,225**	-0,225**	-0,632**	-	-	-
Sola çarpık 3P a (8)	0,242**	-0,281*	0,16**	-0,015	-0,015	-0,422**	0,21**	-	-
Normal 2P a (9)	0,101**	-0,422**	0,019	-0,156*	-0,156*	-0,563**	0,069**	-0,141*	-
Normal 3P a (10)	0,616**	0,093	0,534**	0,359**	0,359**	-0,048	0,584**	0,374**	0,515**

\* p&lt;0,05

\*\* p&lt;0,01

**Tablo 3.6.** Örtük Özellikler Kuramı 2 ve 3 Parametrelili Modellere Göre Evren ve Örneklem İçin Elde Edilen a Parametreleri Arasındaki Korelasyonlar

Örneklem	Evren 2P a	Evren 3P a
Random 2P a	0,988**	0,353*
Random 3P a	0,465**	0,989**
Basık 2P a	0,906**	0,073
Basık 3P a	0,731**	0,792**
Sağa çarpık 2P a	0,731**	-0,216
Sağa çarpık 3P a	0,324*	0,501**
Sola çarpık 2P a	0,956**	0,525**
Sola çarpık 3P a	0,746**	0,791**
Normal 2P a	0,887**	0,030
Normal 3P a	0,372**	0,827**

\*\* p &lt; 0,01

\* p &lt; 0,05

Tablo 3.6'da, 2 ve 3 parametrelili modeller için evren ve örneklemere ait a parametreleri arasındaki korelasyonların 0,989'dan -0,216 aralığında; geniş bir ranjda değiştiği gözlenmektedir. Örneklemeler değiştikçe korelasyonlarda da değişme gözlenmektedir. 3 parametrelili model için manidar olmayan korelasyonlar elde edilmiştir. Farklı modeller için elde edilen a değerleri arasındaki korelasyonların da düştüğü gözlenmektedir.

2 ve 3 parametrelili modele göre evren ve örneklemeler için elde edilen a değerleri arasındaki korelasyonların farklarının manidarlığına ilişkin bilgiler Tablo 3.7 ve 3.8'de verilmiştir.

**Tablo 3.7.** 3 Parametrelili Modele Göre Evren ve 2, 3 Parametrelili Modellere Göre Örneklem İçin Elde Edilen  $a$  Değerleri Arasındaki Korelasyonların Farkları

Örneklem	1	2	3	4	5	6	7	8	9
Random 2P a (1)	-	-	-	-	-	-	-	-	-
Random 3P a (2)	-0,636**	-	-	-	-	-	-	-	-
Basık 2P a (3)	0,28*	0,916**	-	-	-	-	-	-	-
Basık 3P a (4)	-0,439**	0,197**	-0,719**	-	-	-	-	-	-
Sağa çarpık 2P a (5)	0,569**	1,205**	0,289	1,008**	-	-	-	-	-
Sağa çarpık 3P a (6)	-0,148	0,488**	-0,428*	0,291**	-0,717**	-	-	-	-
Sola çarpık 2P a (7)	-0,172	0,464**	-0,452**	0,267*	-0,741**	-0,024	-	-	-
Sola çarpık 3P a (8)	-0,438**	0,198**	-0,718**	0,001	-1,007**	-0,29**	-0,266**	-	-
Normal 2P a (9)	0,323	0,959**	0,043	0,762**	-0,246	0,471**	0,495**	0,761**	-
Normal 3P a (10)	-0,474**	0,162**	-0,754**	-0,035	-1,043**	-0,326**	-0,302**	-0,036	-0,797**

\*  $p < 0,05$ \*\*  $p < 0,01$ 

Tabloya 3.7'ye göre, 2 parametrelili model için hesaplanan evren  $a$  parametreleri ile örneklemle ait  $a$  değerleri arasındaki korelasyonların farklarının 30'u  $\alpha = 0,01$  ve 7'si  $\alpha = 0,05$  düzeylerinde manidardır. Korelasyon farklarından 6 sı manidar bulunmamıştır. Bu bulguya bağlı olarak, örneklem bazında hesaplanan  $a$  değerlerinin evren parametresi ile korelasyonlarında büyük değişimler olduğu söylenebilir.

Tablo 3.8'de, 3 parametrelili model bazında evren ve örneklem için hesaplanan  $a$  değerleri arasındaki korelasyonlara ait farkların 32'si  $\alpha = 0,01$  ve 3'ü  $\alpha = 0,05$  düzeylerinde manidar bulunmuştur. Korelasyon farklarından 10'u manidar bulunmamıştır.

Korelasyonların düşüklüğü ve farkların manidarlığı,  $a$  parametresinin örneklem değişimlerinden fazlaca etkilendiğini göstermektedir.

#### 4. SONUÇ VE ÖNERİLER

Elde edilen bulgulara bağlı olarak aşağıdaki sonuçlara ulaşılmıştır.

**4.1.** Klasik test kuramına dayalı olarak, evren için elde edilen  $p$  (madde güçlük) değerleriyle, örneklem için elde edilen  $p$  değerleri arasında ilişkiler pozitif yönde manidar bulunmuştur. Korelasyonlar arası farkların çoğu manidar bulunurken bir kısmı manidar çıkmamıştır. Bu sonuçlardan klasik

kurama göre elde edilen  $p$  değerlerinin örneklem değıştikçe değışme gösterdiği ve özellikle dağılımın çarpıklığından daha fazla etkilendiği söylenebilir.

**4.2.** Klasik test kuramına dayalı olarak evren için elde edilen  $r_{jx}$  (madde ayırıcılık gücü) değerleriyle, örneklem için elde edilen  $r_{jx}$  değerleri arasında pozitif yönde manidar ilişkiler bulunmuştur. Ancak,  $r_{jx}$  değerlerine ilişkin korelasyonların,  $p$  değerlerine ilişkin korelasyonlardan görece daha düşük olduğu söylenebilir. Örneklem dağılımı sağa çarpık olduğu zaman  $r_{jx}$  için elde edilen ilişkinin oldukça düştüğü gözlenmektedir. Korelasyonlar arası farklar incelendiğinde tamamına yakınının manidar olduğu görülmektedir. Buna göre klasik kurama göre elde edilen madde  $r_{jx}$  değerlerinin örneklem değışikliklerine bağlı olarak değıştiği ve dağılımın çarpıklığından etkilendiği söylenebilir.

**4.3.** Örtük özellikler kuramının 1, 2, 3 parametrelili modellerine göre, evren için elde edilen  $b$  (madde güçlük) parametreleri ile örneklem için elde edilen  $b$  parametreleri arasında bulunan ilişkiler pozitif yönde manidardır. Korelasyon değerlerine bakıldığında elde edilen  $b$  parametrelerinin dağılımın çarpıklığından daha fazla etkilendiği gözlenmektedir. Bu bakımdan  $b$  parametrelerinin klasik kuramdaki  $p$  istatistiklerine benzediği söy-



lenebilir. Aynı zamanda, b parametrelerini hesaplarlarken aynı örtük kuram modelini kullanmanın korelasyonları görece yükselttiği de gözlenmektedir.

**4.4.** Örtük özellikler kuramının 2 ve 3 parametrelili modellerine göre evren için elde edilen madde a (ayırıcılık gücü) parametreleri ile örneklem için elde edilen a parametreleri arasında bulunan ilişkilerin çoğu pozitif yönde manidardır. Ancak manidar olmayan ilişkiler de gözlenmektedir. Korelasyonlar, parametre hesaplanırken aynı örtük kuram modelinin kullanılmasıyla görece artmaktadır. Aynı örtük kuram modelinin kullanılması gereği b parametrelerine göre a parametrelerinde daha etkili olarak göze çarpmaktadır. Ayrıca korelasyonlar arasındaki farkların bir çoğunun çeşitli düzeylerde manidar bulunması ilişkilerdeki değişkenliği destekleyen bir sonuç olarak ele alınabilir. Bu sonuçlara göre, a parametresinin örneklem ve model değişmelerinden etkilendiği söylenebilir.

Örtük özellikler kuramına göre, evren ve örneklem için elde edilen a ve b parametreleri arasında hesaplanan korelasyonların aralarındaki farkların manidar çıkması, kuramın parametrelerin değişmezlik özelliğini sağlayamadığının göstergesi sayılabilir. Özellikle a parametreleri yönünden değişmezliğin sağlanmadığı daha belirgindir. Bu sonuçlara göre, evrene ait gerçek veriler için örtük özellikler kuramının klasik kurama göre avantajı olarak belirtilen değişmezlik özelliğini sağlayamadığını belirtmek yeterli olacaktır. Bu sonuçlar, Fan (1998) ve Nartgün (2002) 'nün ulaştığı sonuçları desteklemektedir. Söz konusu çalışmalarda da örtük özellik kuramının parametrelerin değişmezlik özelliğini sağlayamadığı, bu bakımdan klasik kuramdan daha avantajlı olmadığı sonucuna ulaşılmıştır.

Farklı örneklemelerden klasik kurama göre elde edilen madde güçlük ve ayırıcılık değerleri belli ölçülerde değişkenlik göstermektedir. Buna göre klasik kurama göre elde edilen madde güçlük ve ayırıcılık değerlerinin örneklem değişmelerinden etkilendiği söylenebilir. Ancak bu, klasik kuram için beklenen bir durumdur. Klasik kurama göre elde edilen madde istatistiklerinin örnekleme bağ-

lı olarak değişmesi beklenir. Dikkat çeken nokta klasik kurama göre elde edilen madde güçlük ve ayırıcılık gücü değerleri arasındaki korelasyonlarda var olan değişimin, örtük özellikler kuramına göre elde edilen madde parametreleri arasındaki korelasyonlarda var olan değişimlerden yüksek olmamasıdır.

Her iki kuram için, elde edilen madde istatistikleri ve parametrelerinin örneklem değişmelerinden etkilendiği; dolayısıyla, parametre değişmezliği açısından kuramların herhangi birinin avantaj sağlayamadığı, madde güçlük değerlerinin genel olarak madde ayırıcılık değerlerinden yüksek korelasyon verdiği gözlenmektedir. Elde edilen korelasyonların manidarlığı dikkate alındığında, klasik kuramın değişmezliği sağlamasına rağmen örtük kuramın özellikle a parametreleri açısından değişmezliği sağlayamadığı; korelasyonlar arası farklar dikkate alındığında her iki kuramın da değişmezliği sağlayamadığı; madde ayırıcılık gücü değerleri arasındaki korelasyonların madde güçlük değerleri arasındaki korelasyonlardan düşük olması bakımından iki kuramın birbirine benzediği söylenebilir.

## KAYNAKLAR

- Crocker L. ve J. (1986). *Introduction to classical and modern test theory*. Florida: Harcourt Brace Jovanovich College Publishers.
- Baker, F. (2001). *The basics of Item response theory*. ERIC clearinghouse on assessment and evaluation. <<http://ericae.net/irt/>> (2001, Nisan 4)
- Çıngı, H. (1994). *Örnekleme kuramı*. (2. Baskı). Ankara: H. Ü. Fen Fakültesi Basımevi.
- Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their Item-person statistics. *Educational and psychological measurement*, 58, 3, 357-381
- GELBAL, S. (1994). P madde güçlük indeksi ile rash modelinin bir parametresi ve bunlara dayalı yetenek ölçüleri üzerine bir karşılaştırma. Yayımlanmamış Doktora Tezi Hacettepe Üniversitesi, Ankara.
- Hambleton, R. K., H. S. ve J. H. R. (1991). *Fundamentals of item response theory*. Boston: sage publications.
- Hambleton, R. K. ve H. (1985). *Item response theory: principles and application*. Boston: kluwer-nijhoff publishing.

- Keeves, J. P. (Ed.). (1998). *Educational research, methodology and measurement: An international Handbook*. Oxford: Pergamon Pres.
- Kelecioğlu, H. (2001). Örtük Özellikler Teorisindeki b ve a parametreleri ile klasik test teorisindeki p ve r istatistikleri arasındaki ilişki. *H. Ü. Eğitim Fakültesi Dergisi*, 20, 104-110
- Lord, M.F ve R. M. N. (1968). *Statistical theories of mental test scores*. New York: Addison-Wesley Publishing Company.
- Lord, M. F. (1980). *Application of item response theory to practical testing problem*. New Jersey: Lawrence Erlbaum Associates Publishers.
- M. E. B. *Orta öğretim kurumları öğrenci seçme ve yerleştirme sınavı kılavuzu*. Ankara: M.E.B.
- Nartgün, Z. (2002). Aynı tutumu ölçmeye yönelik likert tipi ölçek ile metrik ölçeğin madde ve ölçek özelliklerinin klasik test kuramı ve örtük özellikler kuramına göre incelenmesi. Yayımlanmamış Doktora Tezi, Hacettepe Üniversitesi, Ankara.
- Osterlind, S. J. (1989). *Constructing test items*. Boston: Kluwer Academic Publishers.
- Somer, O. (1998). Kişilik testlerinde klasik ve modern test kuramları ile madde analizi. *Türk Psikoloji Dergisi*, 13, 41, 1-15
- Stage, C. (1998). *Item analysis based on item response theory and on classical test theory: a comparison*. The international SweSAT conference, San Diego.
- Yu. C. H. (Haziran 1999) *True score and item response theory*. <[http://seamonkey.ed.asu.edu/~alex/computer/sas/math\\_reality.htm](http://seamonkey.ed.asu.edu/~alex/computer/sas/math_reality.htm)> (1 Nisan 2002)