

T. C.  
ANKARA ÜNİVERSİTESİ  
EĞİTİM BİLİMLERİ ENSTİTÜSÜ  
EĞİTİMDE PSİKOLOJİK HİZMETLER ANABİLİM DALI  
ÖLÇME VE DEĞERLENDİRME BİLİM DALI

BEDEN EĞİTİMİ DERSİ İÇİN HAZIRLANAN TUTUM ÖLÇEĞİNİN  
PSİKOMETRİK KURAMLAR AÇISINDAN İNCELENMESİ

DOKTORA TEZİ

Mehmet KARAKILIÇ

Ankara  
Haziran, 2009

## **BÖLÜM I**

### **GİRİŞ**

Bu bölümde, problem, amaç, önem, sayıtlılar, sınırlılıklar ve tanımlar ile ilgili açıklamalara yer verilmiştir.

#### **Problem**

İnsan davranışlarını geliştiren bir sistem olarak görülen eğitim ve öğretim sürecinde yürütülen etkinliklerin vazgeçilemez ve tamamlayıcı en önemli parçası ölçme ve değerlendirmedir. Ölçme, bireylerin ya da nesnelerin belirli özelliklere sahip olup olmadığının, sahipse sahip oluş derecesinin belirlenerek sonuçların sembollerle ifade edilmesi olarak tanımlanır (Lord ve Novick, 1968). Eğitim sistemi içerisinde bireylerin ilgi, yetenek, başarı vb. özelliklerinin belirlenmesi eğitimde ölçme ve değerlendirme çalışmalarının temelini oluşturur.

Ölçme ve değerlendirme uygulamaları, öğrencilerin bilişsel, duyuşsal ve psikomotor becerilerinin ele alındığı, öğrenci gelişiminin farklı boyutlarını kapsayan bir süreci ifade eder. Bu becerilerin ölçümü, bireylerin eğitim alanında ya da psikolojik olarak o andaki durumu hakkında açıklayıcı bilgiler ortaya koyar. Gallahue ve Donnelly (2003) bu ölçümlerin aynı zamanda, bireylerin gelecekteki eğitim tercihlerinin belirlenmesi ve psikolojik yapılarının geleceğe uyarlanması aşamasında önemli bilgiler ortaya koyabileceğini belirtmektedir.

Demirhan (2006), eğitim sisteminde bireylere kazandırılmak istenen bilişsel davranışların bilgiden kaynaklanan zihinsel yeteneklerin ürünü olduğunu, duyuşsal davranışların duygu ve değerlerle ilgili olduğunu, psikomotor davranışların ise zihinsel düşüncelerin kas koordinasyonu ile ifade edildiğini belirtmektedir.

Eğitim ve psikolojide ölçme çalışmalarında yer alan duyuşsal özelliklerden biri tutumlardır. Tutum, "çeşitli kuram, kavram, nesne, durum ya

da kişilere karşı öğrenilmiş, olumlu veya olumsuz tepkide bulunma eğilimi” (Oppenheim, 2000, Tezbaşaran, 1997), veya “diğer kişilere ve nesnelere ilişkin düşüncelere, duygulara ve hareketlere yön veren öğrenilmiş öz eğilimler bütünüdür” şeklinde tanımlanmaktadır (Baron ve Byrne, 1997; Middlebrook, 1974:575; Turgut, 1984:154). Ajzen’e (1993) göre ise tutum, farklı tercihler ya da karışıklıklar karşısındaki utanma derecesi olarak tanımlanmıştır. Bu tanımlar, tutumun tanımı, doğası ve ölçülmesine yönelik farklı yaklaşımların ürünü olarak ortaya atılmıştır. Bunlardan biri üç bileşen yaklaşımıdır (Anderson, 1983; Fazio, 1986).

Üç bileşen yaklaşımına göre tutumlar, belirli tepki sınıfları ile belirli uyarın sınıflarına tepkide bulunma eğilimleridir. Tutumların ölçülmesine yönelik yapılan araştırmalarda önemli ve etkin bir yer tutan üç bileşen yaklaşımında bireylerin tutum objesine yönelik bilgileri tutumun bilişsel bileşenini, bireylerin duyguları tutumun duyuşsal bileşenini ve bireylerin tutumlarını ortaya koyan fiziksel hareketleri de tutumun devimsel bileşenini oluşturur (Eagly ve Chaiken, 1995; Baron ve Byrne, 1997).

Belirli bir uyarana karşı olan bilişsel tepkiler, kişinin uyarın nesnesi hakkındaki bilgileri ve inançlarıdır. Duygusal tepkiler, kişinin o nesne hakkında hissettikleridir. Davranışsal tepkiler ise açık davranışlardır. Bu üç tepki birbiriyle tutarsız olabilir. Bu yaklaşım bir kişinin tek bir nesneye karşı olan tutumunun üç ögesi arasındaki farkın çok büyük olması nedeni ile eleştirilere uğramıştır. Tutumların davranışları ön görüp göremeyeceği günümüzde hala tartışılmasına rağmen tutumların davranıştan önce oluştuğu ve insan davranışının ortaya çıkmasına öncülük ettiği görüşü yaygındır (Baysal, 1981). Bu görüşe göre, tutum sonucu bireyler davranışlara yönelirler, bir başka deyişle, her tutumun altında bir davranış eğilimi yatar. Aynı zamanda, tutumun ortaya çıkmasına yol açtığı duygu, düşünce ve davranış eğilimleri bütünleşir. Tutum doğrudan gözlenemez ama insanların davranışları değerlendirilerek onların sahip oldukları tutumun hangi yönde veya nasıl bir nitelik taşıdığı anlaşılabilir.

Tutum-davranış bağlantısını güçlendiren birçok değişken vardır. Bunlardan birincisi, bir nesneyle yaşanan doğrudan deneyim sayesinde oluşan tutumlar, doğrudan deneyimle oluşmayan tutumlara göre, bu nesneyle ilgili davranışlarla daha bağlantılı görünmektedir. Bu tür tutumlar daha güçlüdürler çünkü kesinlik taşırlar, daha erişilebilirlerdir ve tutum nesnesinin sunumunun ardından hemen aktive olabilirler. Örneğin, Kağıtçıbaşı (1999) en temel tutumların çoğunun çocukluk yaşlarında oluştuğunu belirtmektedir. Anne ve babaların çocukları istediği gibi davrandığında onlara gülümsemeleri ve ödüllendirmeleri, istemedikleri bir davranışı sergilediklerinde ise onları cezalandırmaları çocuklarda kalıcı olumlu veya olumsuz birçok tutumun gelişmesinde etkili olabilmektedir. Tutumlar bu şekilde, kasıtlı yaptırım olmadığında da taklit yoluyla gelişebilir. Çocukların yakın çevrelerindeki akrabaları, öğretmenleri, arkadaşları ve hatta ünlü kişiler onların tutum kazanmalarında etkili olabilir. Yine benzer şekilde, beden eğitimi dersinde öğretilen fiziksel hareketleri diğer öğrenciler kadar iyi yapamayan bir öğrencinin öğretmen ve arkadaşları tarafından olumsuz şekilde eleştirilmesi, o öğrencinin daha sonraki dönemlerde arkadaşları ile birlikte katılacağı fiziksel aktivitelere ya da oyunlara karşı olumsuz bir yaklaşım göstermesine neden olabilir. İkinci olarak, kararlı olan tutumlar, olmayanlara oranla daha fazla tutum-davranış tutarlılığı gösterirler. Çünkü tutumun ölçümü ve davranışın ölçümü arasında geçen zaman arttıkça tutum-davranış bağlantısının gücü daha az olacaktır. Ayrıca, tutum ve davranış birlikte ölçüldüğünde kararlı tutumlarda bağlantı güçlü olacaktır (Baysal, 1981; Inceoğlu, 1993).

Toplumda insanlar, tutumu günlük yaşamda diğer insanları eleştirme gücü olarak görürler; fakat psikometrik açıdan tutum henüz tam olarak anlaşılmiş değildir. Tutum, bireylerin doğrudan ölçülemeyen özelliklerinden birisi olması nedeni ile ilgi, görüş ve inanç kavramları ile benzerlik göstermekle beraber temelde bu kavramlardan bazı yönleriyle önemli farklılıklar da göstermektedir. İlgi, görüş ve inanç kavramları arasındaki benzerlikler, bu kavramların tutum kavramı yerine kullanılmasına neden olmaktadır. İlgi, bireyin herhangi bir çaba sarfetmeksizin, dikkat ettiği,

gözlediği, üzerinde düşündüğü ve keyif alarak yaptığı şeyler (Özgüven, 1998) olarak tanımlanmıştır. Bu tanımlamada önemli olan nokta, ilginin kişiye, objeye ya da etkinliklere yönelik bir zorlama içermeden kendiliğinden gelişen bir davranış olması, tercihlere bağlı olması ve bireylere yapmaktan keyif veren etkinlikler içermesidir. Ancak tutumlar, Turgut (1986) tarafından yapılan daha kesin bir tanımlama ile ilgi, görüş ve inançlardan ayrılmaktadır. Bir şeyle ilgilenen kişi, düşüncelerinin ya da tepkilerinin olumlu olup olmadığına bakmaksızın, zamanının çoğunu ilgilendiği şeyle veya onunla ilişkili bulunduğu şeylerle uğraşmakla veya düşünmekle geçirir. İnanç kavramı ise bir önermenin doğru kabul edilmesi ile ilgilidir. İnanç kavramında bir şeye karşı ya da ondan yana olmak yoktur. Bu yönü ile inanç tutumdan farklılaşmaktadır (Turgut, 1984). Tutumlarla görüşler de benzerlik göstermektedir, ancak görüşler genellenebilirlik derecesi ve ölçme tekniği bakımından tutumlardan farklıdır. Görüşler belirli oluşumlara veya durumlara gösterilen özgül tepkilerdir. Tutumlar ise daha geniş çapta olaylar grubuna veya insan topluluklarına karşı bireyin tepkilerindeki etkileri bakımından daha geneldir. Ayrıca, insanlar görüşlerinin farkındadırlar ama tutumlarının tam olarak farkında olmayabilirler (Tezbaşaran, 1997).

Fazio (1986) tutumların inançlar tarafından belirlendiğini, inançların ise bir kişinin bir nesneye karşı olan tutumunu belirlediğini ifade etmektedir. Örneğin; bir öğrencinin bir derse katılmakla elde edeceği yararlarla ilişkin güçlü inancı, o derse yönelik olumlu tutum geliştirmesini sağlar. Kişinin bir nesneye karşı olan olumlu ya da olumsuz inancı o nesneye yönelik tutumunun olumlu ya da olumsuz olmasını belirler. Davranışsal tepkiler genelde tutumları belirlemek için kullanılır. Bunlar, sözel ya da sözel olmayan tepkiler olabilir. Bir olaya ya da bir şeye yönelik olarak ortaya konulan sözel tepkilerden tutumları elde edebiliriz. Örneğin; bireyin beden eğitimine yönelik tutumu, bu konuda sarf ettiği değerlendirme cümlelerinden ve ifadelerinden anlaşılabilir. Sözel olmayan tepkiler; algılama tepkilerini, fizyolojik reaksiyonları ve motor tepkileri içerir. Beden eğitimi öğretmeni egzersiz yaptırdığı sırada, öğrencinin jest ve mimiklerinden ya da vücut dili hareketlerinden derse yönelik tutumunu belirleyebilir (Demirhan, 2006).

İnançlar tarafından belirlendiği vurgulanan tutumlarda değişiklik gözleminin zor olduğu da belirtilmektedir (Silverman ve Subramaniam, 1999). Kişinin yumuşak huyluluğunun, duruma bağlı olarak olumlu ya da olumsuz tutum değişikliği yapmasına neden olabileceği, durum olumlu deneyimler içeriyorsa istenen tutumun (olumlu tutum) gelişeceği, istenmeyen tutumun ise (olumsuz tutum), durumun yeterince olumlu deneyim içermemesinden kaynaklanabileceği de vurgulanmaktadır.

Tutum bir davranışa başlamayı, devam etmeyi ya da başarıya ulaşmayı etkileyen önemli bir etkidir (Silverman ve Subramaniam, 1999). Bu açıdan tutumlar, eğitim sisteminde de birçok disiplin için önemli bir kavramdır. Bireylerin bilişsel davranışlarıyla birlikte duyuşsal özelliklerinin de başarıyı yordayıcı güce sahip olması, özellikle eğitim alanında hedef alınan davranışların bir kısmının tutumlar üzerinde etkili olması, öğrencilerin okula, derslere, öğretmenlerine ve arkadaşlarına, okul etkinliklerine yönelik başarılarını etkilemesi, eğitim uygulamalarında tutumların ölçülmesini bir zorunluluk haline getirmektedir (Turgut, 1984).

Öğrencilerin derslere ve meslek alanlarına yönelik tutumlarının belirlenmesi, onların gelecekteki meslek alanlarının belirlenmesi ve yaşam tarzlarının ortaya çıkarılması açısından önemlidir. Bu tür bir yaklaşım, öğrencilerin eğitim sistemi içerisinde başarılı olabilmesi açısından da önemlidir. Öğrenci başarısı üzerine yapılan araştırmalar (Silverman, 1993), öğretme ve öğrenmenin karmaşık doğasını anlamaya yardımcı olur çünkü eğitim sistemi içerisinde yer alan müfredatın önemli birer parçası olan derslere yönelik öğrencilerin olumlu tutum sergilemesi, öğrencilerin bu derslerdeki başarılarını etkileyen önemli bir unsur olarak görülmektedir. Öğrencilerin tutumlarının araştırmalara dahil edilmesi, öğrencilerin öğrenme sürecine bütünsel bir bakış açısı getirmiştir. Öğrencilerin tutum bilgilerinin elde edilmesi hem öğretmenler hem de öğrencilere öğrenme sürecinde yardım edebilir. Birçok alanda öğrencilerin derslere yönelik tutumlarının belirlenmesi için yapılan araştırmaların temel amacı, bu öğrencilerin derslere

yönelik tutumlarını belirlemek ve varsa olumsuz tutumların olumluya dönüştürülmesiyle, derslerdeki başarılarının artırılmasıdır (Arul, 2002).

Tutum olumludan olumsuza doğru giden uzun bir süreçtir. Uzun süreli ya da geliştirilmiş iyi deneyim bireyin derse katılma isteğinin gelişmesini sağlar. Bu durumda süreç olumlu yöne doğru bir eğilim gösterir. Sürecin olumlu tarafına yaklaşması kişinin isteğinin ve hislerinin yoğunluğuna bağlıdır. Eğer etkinliklerin çoğunda tutum etkiliyse, çocukların nasıl algıladığını, nasıl hissettiğini, tutumu nasıl değerlendirdiğini bilmek öğretmenler ve ebeveynler için derslere karşı olumlu yaklaşım geliştirmenin anahtarı olabilir (Silverman ve Subramaniam, 1999). Öğretmenler, öğrencileri için doğru fırsatları ve deneyimleri sağlamada çok önemli bir role sahiptir.

Öğrencilerin öğretim programında yer alan derslere yönelik tutumlarını incelemeye yönelik pek çok araştırma yapılmıştır. Bu çalışmalar, temel bilimler olarak bilinen Fen Bilgisi, Matematik (Nuhoğlu, 2008; Yenilmez, 2007) ve öğretim programında yer alan Coğrafya, (Alım, 2008) İngilizce (Erdem, 2007) gibi derslerle birlikte, öğrencilerin sosyal ve fiziki gelişimlerini önemli ölçüde etkileyen Resim (Beykal, 2005) ve Beden Eğitimi (Özer ve Aktop, 2003; Demirhan ve Altay, 2001) derslerine yönelik tutumlarını belirleyen önemli araştırmalardır.

Bu araştırmada KTK ve ÖÖK'nın psikometrik özellikleri beden eğitimi dersine yönelik bir tutum ölçeği hazırlama çalışması kapsamında inceleneceği ve karşılaştırılacağı için aşağıda, beden eğitimi dersine yönelik öğrenci tutumlarının incelenmesinin önemi vurgulanacak ve bu derse yönelik tutumları ortaya koyan araştırmaların bulgularına değinilecektir.

### **Beden Eğitimi ve Tutumlar**

Beden eğitimi, beden etkinlikleri yoluyla, bireyin vücut gelişimine yardım ederek uyumunu sağlayan eğitimin bir parçasıdır ve insan bütünü oluşturarak fiziksel, duygusal, zihinsel ve sosyal niteliklerin, bulunduğu yaşın ve genetik kapasitenin gerektirdiği güce ulaştırılması için, çoğu kez rekabet

olmaksızın yapılan etkinliklerin tümü olarak tanımlanmaktadır (Bowyer, 1996). Burada sözü edilen beden etkinlikleri, her bireyin kişisel ve sağlık özelliği ölçüsünde katılacağı, kendini ifade edebileceği etkinlikler olup, vücudun yapı ve fonksiyonunu geliştirebilmeyi, eklem ve kasların kontrolü ve dengeli bir biçimde gelişmelerini sağlamayı, okul çağı sonrası iş ve rekreatif faaliyetlerin süresi içinde harcanan fiziksel eforu en ekonomik şekilde kullanmasını öğreten faaliyet sistemini ifade etmektedir. Spor ise, birey ya da grupların sağlık, eğlence veya gösteri amacı ile yaptıkları, fiziksel ve/veya zihinsel aktivite gerektiren, bazı kurallar içinde uygulanan organize edilmiş oyun olarak tanımlanır (Bompa, 2002).

Beden eğitimi, her bireye yönelik kişisel gelişim sağlama ve sağlıklı yaşam becerisi kazandırma etkinliklerini içermesiyle, yukarıda verilen spor kavramının tanımından farklı olarak tüm bireylere ulaşabilme özelliği taşır. Fiziksel etkinliklere düzenli olarak katılmanın bireylerin fiziksel ve zihinsel sağlıklarına önemli faydalarının olduğu (McKenzie, 2003) bilinmektedir. Fiziksel etkinliklere katılmak sadece, fiziksel ve zihinsel sağlığı geliştirmek açısından değil, toplumsal açıdan halk sağlığı bilincinin yerleşmesi için de önemlidir. ABD’de 1996 yılında Halk Sağlığı Bölümü’nün (Department of Health and Human Service) yaptığı araştırma, ulusal gençliğin sağlık durumlarının beden eğitimi derslerine katılımları ile ilişkili olduğunu ortaya koymuştur. Bununla beraber, bireylerin fiziksel etkinliklere katılmak için ortaya koydukları istek ve arzunun en önemli anahtarı, bireylerin tutumları olarak tanımlanmaktadır (Carlson, 1995; Sanders ve Graham, 1995; Solmon ve Carter, 1995; McKenzie, Alcaraz, ve Sallis, 1994; Tjeerdsma, Rink, ve Graham, 1994). Beden eğitime ve fiziksel aktiviteye karşı olumlu tutum geliştirmek, gençler arasında fiziksel olarak aktif bir yaşam tarzı geliştirmek için önemlidir (Graham, 1995).

Fiziksel etkinliklere yönelik tutumları olumlu yönde etkilemenin, buna bağlı olarak toplumsal açıdan halk sağlığı bilincinin yerleşmesinin ve özellikle çocukların düzenli olarak fiziksel etkinliklere katılmasına fırsat sağlamanın



(Gallahue & Donnelly, 2003; Graham ve ark., 1998) günümüzdeki ve gelecekteki en önemli yolu okullardaki beden eğitimi dersleridir.

Öğrenciler beden eğitimine, farklı sahip oldukları farklı görüş ve düşünceler çerçevesinde şekillendirdikleri tutumlar, istekler ve karakter yapılarıyla başlarlar (Gallehue, 1996). Erken yaşlarda şekillenmeye başlayan fiziksel aktiviteye yönelik tutumlar, ilerleyen yaşlarda özellikle ders öğretmeni ve sınıf çevresi gibi durumsal şartlardan, öğrencilerin daha önce beden eğitimi ile ilgili yaşadıkları deneyimlerden, beden eğitimi dersini algılayışlarından, okudukları sınıf düzeylerinden, yetenek düzeylerinden, cinsiyet farklılıklarından, sosyal sınıflarından ya da etnik kökenlerinden fazlaca etkilenir (Brophy ve Good, 1986; Aicinena, 1991; Lee, Kang ve Hume, 1999; Silverman ve Subramaniam, 1999; Gallahue & Donnelly, 2003; Portman, 2003). Örneğin, beden eğitimine yönelik olumlu tutumu olan bir öğrenci, öğrenme ortamını hem öğretmenleri hem de kendisi için eğlenceli bir deneyim haline getirir. Carlson (1995) ve Graham ve ark. (1998) yaptıkları araştırmalarda, bu belirlemeyi destekleyen sonuçlar elde etmiş, beden eğitimi dersine yönelik olumsuz tutuma sahip olan öğrencilerin bu dersti hayatlarındaki bir ihtiyacı karşılamayan, gereksiz bir ders olarak gördüklerini ortaya koymuştur. Buna ek olarak, beden eğitimi öğretmenin öğrencilerin beden eğitimi dersine yönelik olumlu tutum geliştirmelerinin anahtarı olduğunu, düşük yetenek düzeyindeki öğrencilerin beden eğitimi dersinde, takım oyunlarına katılmadıkları ve yeteneksizlikleri ile alay edildiği için kendilerini sınıftan soyutlanmış olarak gördüklerini, olumsuz tutuma sahip öğrencilerin beden eğitimi dersine katılımı rahatsız ve güvensiz olarak nitelendirdiklerini ve olumsuz tutumlar nedeniyle öğretme ve öğrenme sürecinin zorlaştığını ileri sürmüşlerdir.

Tutumların öğretme ve öğrenme süreci açısından önemi göz önünde bulundurulduğunda; beden eğitimi dersine olumlu tutum ile başlayan öğrencilerin okuldaki eğitimleri süresince bu derste daha başarılı oldukları (Carlson, 1995; Sanders & Graham, 1995; Solmon & Carter, 1995; McKenzie, Alcaraz, & Sallis, 1994; Tjeerdsma, Rink, & Graham, 1994)

gözlenmektedir. Bununla birlikte, beden eğitimi dersinde yapılan fiziksel etkinliklere yönelik olumlu tutumları olan bireylerin okul dışındaki fiziksel etkinliklere katılma eğilimlerinin daha fazla olduğu ve belirli bir dönemde fiziksel etkinliğe yönelik olumlu tutumların gelişmiş olmasının yaşam boyu fiziksel etkinliklere katılmanın temelini oluşturarak çocukluk çağlarında fiziksel olarak aktif olan bireylerin daha sonra fiziksel etkinliklere katılan aktif yetişkinler olma olasılığının diğerlerine göre çok daha yüksek olduğu da (Chung & Phillips 2002; Greenwood & Stillwell 2001; Portman, 2003) çeşitli araştırmalarla belirlenmiştir. Bu nedenle, öğrencilerin beden eğitimi dersine yönelik olumlu tutumlar geliştirmesini sağlamak beden eğitimi programlarının amaçları arasında yer alır ve beden eğitimi alanında öğrencilerin dersteki başarılarını büyük ölçüde etkileyen müfredata yönelik tutumlarının incelenmesi oldukça önemlidir. Çünkü öğrencilerin beden eğitimi ile ilgili deneyimleri, bu ders hakkında ne düşündükleri, ne hissettikleri ve bu ders ile ilgili bilgileri daha sonraki dönemlerde beden eğitimi derslerinin tüm öğrenciler için değerli bir deneyim olmasını sağlayabilecek önemli bir faktördür (Graham, 1995).

Luke ve Sinclair (1991)'ın öğrencilerin dersteki başarılarını büyük ölçüde etkileyen müfredata yönelik tutumlarını, çeşitli değişkenler açısından incelemeyi amaçlayan çalışmalarında; müfredatın içeriği, öğretmen davranışları, sınıf atmosferi, öğrencilerin benlik algıları ve tesisler olmak üzere beş önemli faktörün tutumları etkilediğini belirlediklerini ileri sürmektedirler. Scantling ve ark. (1995) ve Greenockle ve ark. (1990) çalışmalarında da benzer sonuçlara ulaşmış, etkinlikten sıkılma, sınıfta terlemekten çekinme, ders için soyunup giyinmek istememe ve sporcu öğrencilerin öncelikli olduğuna ilişkin algıların tutumları etkileyen faktörler olduğunu belirlediklerini ifade etmişlerdir.

Öğrencilerin beden eğitimine yönelik tutumları ve algıları, yukarıda sözü edilenler dışında pek çok araştırmayla (Carlson, 1995; Barney, 2003; Bowyer, 1996; Stewart, Green ve Huelskamp, 1991; Tannehil ve Zakrajek, 1993; Graham ve ark., 1998; Gallahue & Donnelly, 2003; Portman, 2003)

yaş, cinsiyet ve sınıf düzeyi gibi değişkenler açısından da incelenmiştir. Bu araştırmaların bir kısmında okunan sınıf ve cinsiyet bakımından belirgin farklar bulunmazken, diğer araştırmalarda erkek öğrencilerin daha olumlu bir tutuma sahip olup daha zor ve riskli hareketlerin yapıldığı etkinlikleri tercih ettiği, kız çocukların ise estetik duyguları ön plana çıkaran hareketlerden hoşlandığı ve sınıf düzeyine paralel olarak beden eğitimine yönelik tutumların olumlu yönde artış gösterdiği sonucuna ulaşılmıştır.

Beden eğitimi dersine yönelik tutumlarının ölçülmesine ilişkin yapılan araştırmalarda farklı sonuçlar elde edilmesine karşın, yapılan bu araştırmalar özetle şu ortak sonuçları ortaya koymuştur (Strand ve Scantling, 1994):

- Genellikle öğretim programının içeriği, öğretmenlerin davranışları, sınıf atmosferi ve öğrencilerin benlik algıları, öğrencilerin beden eğitimi dersine yönelik tutumlarında belirleyici etkenlerden biridir.
- Öğrenciler, beden eğitimi derslerine öğretim programı zorunluluğu dışında gönüllü olarak katıldığı zamanlarda, genellikle memnundur.
- Öğrencilere karar verme fırsatlarının verilmesi, beden eğitimi dersine yönelik daha olumlu tutumların gelişmesinde etkili olabilecek etmenler arasındadır.
- Öğrenciler, beden eğitimi derslerinde kalabalık sınıflardan, diğer öğrencilerle kıyaslanmaktan, bireysel olarak yetenek düzeylerinin puanlanmasından, giyinip soyunmaktan ve sınıf içerisinde mahçup durumda bırakılmaktan hoşlanmamaktadır.
- Öğrenciler, beden eğitimi dersinde puanlama ölçütlerinin sadece yetenek düzeyi olarak belirlenmesinden hoşlanmadıklarını belirtmişlerdir.
- Beden eğitimi dersleri, öğrencilerin isteklerini ve tercihlerini de dikkate alan çeşitli etkinlik ve aktivitelerden oluşmalıdır.

- Erkek öğrenciler beden eğitimi derslerinde yarışmacı karaktere sahip zorlu aktiviteler yapmayı tercih etmekte, kız öğrenciler ise daha çok artistik değerler üzerinde durmaktadır.
- Öğrenciler fiziksel uygunluk ve sosyal gelişime değer vermelerine karşın, sadece fiziksel uygunluk için yapılan aktiviteleri sevmediklerini belirtmişlerdir.
- Öğrenciler, beden eğitimi dersini öğretim programının önemli bir parçası olarak görmektedir.

Yukarıda ortak bulguları özetlenen araştırmalarla elde edilenler dışındaki farklı bulguların (Silverman ve Subramaniam, 1999; Oppenheim, 2000; Ryan, Fleming ve Maina, 2003; Subramaniam ve Silverman, 2005) nedenlerinin;

- beden eğitime yönelik tutumların ölçülmesi için geliştirilen ölçme araçlarının sağlam kuramsal bir temele dayandırılmaması,
- tutumların bazı çalışmalarda tek bileşenli bir yapıda incelenmesinin sadece duyuşsal boyutu ele alarak tutum ölçümüne sınırlama getirmesi,
- tutumların bazı çalışmalarda iki bileşenli bir yapıda incelenmesinin bilişsel ve duyuşsal boyutları ele alıp devimsel bileşeni göz ardı etmesiyle yine tutum ölçümüne sınırlama getirmesi ve
- böylelikle tutumun temel yapısını ortaya koyan üç bileşen kuramının ve güncel psikometrik tekniklerin göz ardı edilmesi olabileceği belirtilmiştir.

Öğrencilerin beden eğitime yönelik tutumlarının ölçüldüğü çalışmalarda genel olarak KTK kapsamında nicel ve nitel ölçme yöntemleri kullanılmıştır. Son yıllarda bu çalışmalara ÖÖK temelinde geliştirilen ölçme araçları da dahil edilmiştir. Bundan sonraki bölümlerde beden eğitimi ve diğer tüm alanlarda tutumların ölçülmesinde kullanılan yöntemlere ve bu yöntemlerin birbirlerine göre farklılıkları ve üstünlüklerine yer verilmektedir.

## Tutumların Ölçülmesi

Tutumların ölçülmesinde değişik amaçlar vardır. Bu amaçlardan ilki bireylerin belirli tutum ve değerlerinin belirlenmesidir, ikinci amaç bireylerin gözlenen tutum ve değer yargılarını etkileyen aile ve genel çevre faktörlerinin incelenmesidir. Üçüncü amaç ise, kişilik ölçekleri ile birlikte davranışı etkileyen önemli bir faktör olarak bireyin uyum problemlerinin teşhis edilmesidir (Özgüven, 1998:91).

Bütün psikometrik yapılarda olduğu gibi bireylerin tutumlarını doğrudan gözlemek olanaksızdır. Tutumların fiziksel boyutunun olmaması ve soyut kavram olması nedeniyle ölçeklenmesi oldukça güçtür. Bireylere herhangi bir tutum objesi ya da konusu ile ilgili tutumları sorulduğunda, genel olarak tam bir cevap veremezler, yüzeysel ya da doğru olmayan ifadeler kullanırlar. Thurstone ve Chave (1970) tutumların kanılar aracılığı ile ölçülebileceği görüşünü savunmuştur. Diğer araştırmacılar, kanıların her zaman tutumları ifade etmediklerini ve bireylerin söyledikleri ile yaptıklarının farklı olabileceğini belirtmişlerdir. Buna rağmen, davranışların bireylerin kendilerine özgü tutumlardan kaynaklandığı kanısı oldukça yaygındır. Bu kanı tutumların ölçülmesine temel oluşturmaktadır (Özgüven, 1994; Köklü, 1995; Tavşancıl, 2002).

Tutumların ölçülmesinde farklı yöntemler kullanılmaktadır. Bu yöntemler, görünen davranışın gözlenmesine dayanan ölçmeler, bireylerin kendi ifadelerine dayanan ölçmeler (ölçekler), bireyin kısmen yapılandırılmış bir uyarıcıyı yorumlama şekline göre tutum ölçümü (yansıtma yöntemleri), bireylere verilen bazı objektif iş ya da görevleri yerine getirmesine dayanan ölçümler ve bireylerin fizyolojik tepkilerine dayanan ölçmelerdir (fizyolojik ölçme yöntemleri) (Baysal, 1981:55).

Tutumların ölçülmesinde benimsenen görünen davranışın gözlenmesine dayanan yaklaşımlardan gözlem tekniğinde, tutum ögesinin bulunduğu ortamda bireylerin davranışları dolaysız olarak gözlenir. Bireylerin gözlem sürecinde sergiledikleri davranışların hangilerinin tutum ögesine

yönelik olduğu kesin olarak belirlenmelidir. Gözlenen davranışların büyük bir bölümü tutum ögesinden farklı bir nedenden kaynaklanıyor olabilir (Özguven, 1994). Bir öğrencinin beden eğitimi dersine devamlılığı ve derste yapılan etkinliklere katılması, bu öğrencinin beden eğitimi dersine yönelik tutumunun kesin bir göstergesi olmayabilir. Bu davranışların temelinde farklı nedenler olabilir. Gözlem tekniği uzun zamana dayalı bir ölçüm tekniği olduğu için çok tercih edilen bir yöntem değildir.

Bireylerin bir tutum ögesine yönelik tutumlarına ilişkin dolaysız sorular sorarak sözlü ve yazılı cevaplar da alınabilir. Ancak bu yöntem güvenilir sonuçlar vermeyebilir. Sözlü ve yazılı ölçme araçlarında yaşanan temel sorunlar bu yöntemin uygulamasındaki genel sakıncalardan birisidir. Bu yöntemde özellikle teknik sorunlar öne çıkmaktadır. Ölçme hataları ya da soruların anlaşılmasındaki güçlük, tutumların ölçülmesindeki en temel sorun olarak karşımıza çıkmaktadır. Ayrıca, bu tür çalışmalarda güvenilirliğin anlaşılmayan sorular, dikkatsiz yanıtlayıcılar ve yetersiz bilgi işlem uygulayıcıları nedeniyle düşük olması da bu yöntemin önemli bir diğer kısıtlılığıdır.

Tutumların ölçülmesinde nicel yöntemi kullanan araştırmacı veri toplamak için kendi araçlarını geliştirebilir ya da geliştirilen araçları kullanabilir. Nitel yöntemle öğrencilerin tutumlarını belirlemek için anket, gözlem, görüşme, dokümanlar ya da kritik vaka sorgulama raporları veri kaynağı olarak kullanılabilir. Nicel yöntem kullanımı, öncelikle araçların geliştirilmesi ve geçerliğinin saptanması ile başlar. Psikometrik kuram ya da test kuramı ölçme araçlarının geçerliğinin sınanmasında kullanılır (Lord, 1980). Araçların güvenilirlik ve geçerliği tutumun nicel yöntemler ile ölçülmesinde anahtar parçalardır.

Tutumların ölçülmesinde kullanılan yöntemlerden biri olan tutum ölçekleri, tutumların ölçülmesinde en çok kullanılan araçlardır. Tutum ölçülürken araştırma konusu olan tutum objesi ile ilgili, cümle, sıfat ya da ifadeler dizisi olan bir liste hazırlanır. Bireylerin bu uyarıcılara gerçek duyguları doğrultusunda tepkide bulunmaları istenir. Bu uyarıcılar listesi ölçek

olarak adlandırılır. Başka bir deyişle, tutum ölçekleri bireylerin iç dünyasını ortaya çıkarmak için oluşturulmuş bir dizi ifadeye bireyin cevap vermesi için hazırlanmış araçlardır. Tutum ölçümünde sonuç, bireyin duygularının yoğunluğunun tutum objesinin lehinde mi yoksa aleyhinde mi olduğunu yansıtmalıdır. Bir grubun tamamının tutumlarını araştırdığımızda bireylerin tek tek tutumları hakkında bilgi edinmeye gerek yoktur. Tepkiler ortak tutulabilir. Bu uygulama, tutumların daha gerçekçi bir şekilde ifade edilebilmeleri olasılığını artırır (Tavşancıl, 2002).

Tutumların ölçülmesi ile ilgili çalışmalar ve bu konudaki gelişmeler incelendiğinde bazı temel yaklaşımlar ve bu yaklaşımlar çerçevesinde geliştirilmiş ölçekler (ölçme araçları) görülmektedir. Bunlar arasında Bogardus'un Toplumsal Uzaklık Ölçeği, Thurstone Eşit Görünümlü Aralıklar Ölçeği, Likert'in Dereceleme Toplamları ile Ölçekleme Tekniği, Guttman'ın Yığışımlı (Birikimli) Ölçekleme Tekniği, Osgood Duygusal Anlam Ölçeği standartlaştırılmış ölçme teknikleridir. Bu ölçeklerden Bogardus, Thurstone, Likert ve Guttman ölçekleri cümlelerden/maddelerden, Osgood Duygusal Anlam Ölçeği ise sıfatlardan oluşmaktadır (Tavşancıl, 2002).

İlk tutum ölçme tekniği olan toplumsal uzaklık ölçeği, Bogardus tarafından 1925 yılında insanların diğer dil, din, ve ırklardan olan insanlarla olan ilişkileri incelenmek amacı ile geliştirilmiştir. Ölçeğin temel mantığı, eğer bir birey başka dil, din, ve ırktan birisi ile aynı mahallede yaşamayı kabul ediyorsa toplum içerisindeki sosyal yaşantı çerçevesindeki diğer etkinlik ve oluşumlarda da bu insanla beraber yaşamayı kabul ediyor demektir. Ölçekteki maddeler çok yakın toplumsal ilişkiyi benimseme eğiliminden, toplumsal ilişkiden kaçınma eğilimine doğru sistematik olarak sıralanmaktadır (Arul, 2002; Tavşancıl, 2002).

Bogardus Toplumsal Uzaklık Ölçeği, uygulama ve planlama kolaylığı nedeniyle etnik tutumların ölçülmesinde kullanılan bir ölçektir (Sencer ve Sencer, 1978). Tüm formlarında, ırklar ve tüm etnik gruplar için verilen toplam puanların ortalamaları, en az 1.00, en çok 7.00 arasında değişen ırksal uzaklık katsayısını (Racial Distance Quotient) verir. Ölçek tutumların

ölçülmesine ilişkin geliştirilen ilk olmasına karşın, bu ölçeğin geçerliği ve güvenilirliği konusunda önemli şüpheler ve ciddi tartışmalar vardır (Tavşancıl, 2002).

Thurstone sosyal tutumların ölçülebileceğini savunan ilk bilim adamıdır. Tek boyutlu ölçek geliştirmek için farklı teknikler bulmuştur. Bu teknikler içinde en yaygın olarak kullanılanları, çiftli karşılaştırmalar tekniği ve eşit görünen aralıklar ölçeğidir (Tezbaşaran, 1997). Çiftli karşılaştırmalar tekniğinde ölçülmek istenen tutum objesine ilişkin çok sayıda ifade oluşturulur ve bunlar çiftler halinde düzenlenir, daha sonra bu ifadelere olası karşılaştırma sayısı  $C = [N(N-1)]/2$  eşitliği ile belirlenir. Bu aşamadan sonra, ifadelerin ölçek değerlerini belirleyecek karşılaştırmalar elde etmek için hakemler grubu oluşturulur. Hakemler her ifade çiftini okuyup karşılaştırır ve birini seçer, bundan sonra her ifadenin tüm ötekiler karşısında aldığı olumlu yargıların oranı hesaplanır. Her madde için bu oranlar toplanır ve ölçek değerlerine dönüştürülür, ölçeğin iç tutarlılık analizini yapmak için maddeler arasındaki mutlak ortalama fark hesaplanır ve son olarak ölçek, tutumları ölçülmek istenen bireylere verilerek her ifadeye katılıp katılmadıklarını belirtmeleri istenir. Cevaplayanların ölçek puanı, verdikleri olumlu tepkinin ortancası olarak hesaplanır (Tavşancıl, 2002).

Thurstone din, ölüm cezası, doğum kontrolü gibi konulara yönelik tutumları ölçebilmek için çalışmalar yapmış ve eşit görünen aralıklar tekniğini geliştirmiştir. Bu teknik objelerin sıralı değerlendirmesini vermesi yanında, ölçek üzerindeki herhangi iki ölçüm arasındaki uzaklığa ilişkin yargılara ulaşmayı sağlamaktadır (Özgüven, 1994:342; Tavşancıl, 2002:124). Geliştirilen bu ölçek ile;

- belirli bir bireyin söz konusu tutum ölçeğindeki yeri,
- kabul edebileceği tutumların değişim aralığı,
- belirli bir grupta, o grup için belirlenen frekans dağılımına göre ölçekteki her maddenin onaylanma oranı ve



- belirli bir grup için, belirlenen frekans dağılımına göre o grubun tutumlarının çeşitliliği

gibi özellikler belirlenebilmektedir (Baysal, 1981:75).

Thurstone ölçeğine yönelik eleştiriler vardır. Bu eleştirilerden ilki, Thurstone tutum ölçeği geliştirme çalışmalarının yoğun bir çaba ve ağır bir uğraşıya gereksinim duymasındır. İkinci eleştiri, maddelerin ölçek değerinin jüri üyelerinin yargılarına göre saptanması işlemidir. Bu işlemde jüri üyelerinin, yaptıkları derecelendirmeye kendi kişisel tutumlarını da yansıtmaya olasılığının bulunmasıdır. Thurstone ölçeğine yapılan bir diğer eleştiri de, bu ölçeğin farklı değerler sistemine bağlı kültürlerde uygulandığında, maddelerin anlam ve sıralarında önemli değişimler olmasıdır (Özgüven, 1994).

Sıfatlardan oluşan Osgood Duygusal Anlam Ölçeği, bireylerin sahip olduğu tutumu hakkında taşıdığı anlamsal incelenmesi temeline dayanmaktadır. Duygusal Anlam Ölçeği farklı kültür ve yaş gruplarına uygulanan ekonomik ve kullanımı kolay bir ölçme aracı olmasına karşın, ölçeğin farklı kültürlerde uygulanabilmesi için farklı kültürlere özgü geçerliğin saptanması ve sıfatlar için standart bir yapı oluşturulması gerektiğinden geliştirilmesi zor ve zahmetlidir (Kağıtçıbaşı, 1999).

Guttman'ın ikinci dünya savaşı sırasında askerlerin tutumlarını ölçmek için geliştirdiği tekniğe Yığılımlı (Birikimli) Ölçekleme Tekniği adı verilmiştir. Bu teknikte geliştirilen ölçeklerde, tek boyutluluğun ve güvenilirliğin sağlanabilmesi için bir maddeyi en üst düzeyde red ettiğini gösteren bir maddeyi işaretleyen bireyin diğer maddelerde de red ettiğini gösteren maddeleri işaretlemesi gereklidir. Bu ölçekte, bireylerin ölçek maddelerine verdiği cevapların, ölçeğin diğer maddelerinde de aynı yönde olması beklenir. Tek boyutluluk bakımından güvenilir sonuçlar ortaya koyması ve tutumlardaki değişmeye duyarlı olmasına karşın, maddelerin seçiminde sağduyuya dayanması, madde seçiminde çok katı ve sınırlayıcı olması ve analiz süresinin uzun ve karmaşık olması nedeniyle eleştirilmiştir (Sencer ve Sencer, 1978; Özgüven, 1994).

Rensis Likert, 1932 yılında bir anlamda Thurstone ölçeğine yöneltilen eleştirileri ortadan kaldırmak amacı ile Likert tipi tutum ölçeğini geliştirmiştir. Tutumları ölçülecek bireylerin tepkide bulunacakları ifadeler bulunan Likert ölçeği, deneklerin ön planda tutulduğu ölçekleme yaklaşımının önemli bir örneğidir. Bu tip ölçekte, denekler benimsedikleri ifadeleri seçmek yerine, verilen her ifadeye ne ölçüde katılıp katılmadıklarını derecelendirirler (Özgül, 1994:344). Toplamalı sıralama tekniğine sahip olan Likert ölçeği, tutumların ölçülmesinde en yaygın olarak kullanılan ölçektir.

Likert ölçeğinin, Guttman ve Thurstone ölçeklerine göre avantajı, geliştirilmesinin daha kolay olmasının yanında çok çeşitli tutum objelerine ve durumlarına uyum sağlayabilmesidir. Bu, tutumun ölçülebilen boyutlarından hem yönünü hem de derecesini hesaplayabilme kolaylığı sağlar. Bu özellikleriyle bu çalışmada tutum ölçeği hazırlama çalışması kapsamında Likert yöntemi tercih edilmiştir.

Likert tutum ölçeği genel olarak üç temel varsayıma dayanmaktadır. Bunlardan birincisi, ölçekteki her maddenin bir tutum boyutuyla monotonik bir ilişkiye sahip olmasıdır. Likert ölçeğinde olumlu maddelere “Tamamen Katılıyorum” tepkisi cevaplayıcıların aşırı olumlu tutum içerisinde olduğunu ve o madde için en yüksek puanı alacağını, “Hiç Katılmıyorum” tepkisi ise cevaplayıcıların aşırı olumsuz bir tutum içerisinde olduğunu ve o madde için en düşük puanı alacağını gösterir. İkinci varsayım, madde cevaplayıcıları doğrultusunun ölçülen tutumla monotonik bir ilişki içerisinde olmasıdır. Üçüncü varsayım ise, ölçekteki maddelerin ilgilenilen tutumu ölçtüğüdür. Bu, doğrusal olarak bir araya getirilen maddelerin tek bir faktör altında toplandığı ve bu maddelerin toplamının her bir maddede kapsanan önemli bilginin tümünü içerdiğidir (Tezbaşaran, 1997:9).

Tutumların ölçülmesinde de diğer psikolojik özelliklerin ölçülmesindeki gibi kullanılan ölçek ve ölçülen özelliklerle ilgili ilkeler vardır. Ölçeklerle elde edilebilecek ölçme sonuçları ancak bu ilkeler karşılandığı sürece geçerli olacaktır. Bu ilkelerden birincisi sürekliliktir. Ölçekler ölçülen niteliğin kategorik değil sürekli bir değişken olduğunu kabul eder. Tutum ölçeklerinde

birbirine zıt iki uç arasındaki çeşitli derecelendirmelerin sonsuz derecede birbirlerini izleyen küçük birimlere ayrılması durumunda bile sürekliliğin korunacağı varsayılır. İkinci ilke tek boyutluluktur. Ölçeklerde ölçülen niteliğin elden geldiğince tek bir özelliğe bağlı kalması öngörülür. Ancak tutum ölçeklerinde tek bir boyutun ölçüldüğünü düşünmek zordur. Tek boyutluluk, ölçekteki maddelerin içtutarlılık değerleri araştırılarak incelenebilir. Diğer bir ilke de doğrusallıktır. Bir sayıtlı olarak ölçülen niteliğin bir doğru çizgi biçiminde sürekli olduğu ve aralıkların birimlerle gösterildiği kabul edilir (Köklü, 1995).

Yukarıda bahsedilen ölçeklerin geliştirilmesinde temel alınan, Klasik Test Kuramı (KTK)'nın da incelenmesi tutum ölçeklerinin geliştirilmesindeki temel esasların daha iyi anlaşılmasını sağlayacaktır. Bundan sonraki bölümde KTK ve KTK'nın sınırlılıklarına çözüm bulmak amacı ile geliştirilen Örtük Özellikler Kuramı (ÖÖK) incelenecektir.

### **Klasik Test Kuramı**

KTK, psikoloji ve eğitim alanlarında var olan bazı gerçek yapıları ölçme yolu ile açıklayarak kuramsal yapılar şeklinde ifade eder. Bu test kuramı, gerçek puan kavramı olarak da nitelendirilmektedir (Baykul, 2000).

KTK'da, eğitim ve psikoloji alanları içerisinde var olan ve açıklanması istenen yapılar, ölçme yöntemleri kullanılarak kuramsal yapılara bağlanır ve açıklanmaya çalışılır. Bu yapılırken kullanılan ölçme yöntemleri, ölçmek istediğimiz özelliğin gerçek değerini vermez. Çünkü ölçmeye karışan birtakım hatalar, ölçmek istediğimiz özelliğin gerçek değerini doğrudan elde etmemize olanak tanımaz ve bu nedenle, elde edilen gözlenen puanlar yardımıyla ölçülmek istenen özellikle ilgili kestirme yapılmaya çalışılır. KTK'da, ölçülmek istenen özelliğe ait gerçek değer, gerçek puan olarak adlandırılır ve ölçme yoluyla doğrudan elde edilemeyen gerçek puan, birtakım varsayımlarla gözlenen puanlardan kestirilir. Bu nedenle KTK, gerçek puan kuramı olarak da bilinir (Baykul, 2000:97).

KTK'da bireyin bir testteki gözlenen puanı, gerçek puanın ve hata puanlarının toplamından oluşmaktadır. Kuramda, ölçme hatalarının tesadüfi olduğu ve bu hataların gerçek puanlarla, başka bir değişkenin gerçek puanıyla ve farklı ölçmelere ait hata puanlarıyla korelasyon vermediği varsayılmaktadır (Lord ve Novick, 1968; Hambleton ve Swaminathan 1985). Ölçmenin hatasız olması beklenemez ve bu durum KTK ile hatayı en aza indirerek gözlenen puanın gerçek puana yakın bir değer alması ile olası bir durumdur. Bu amaca ulaşabilmek için, KTK aynı özelliğin birden fazla ölçülmesi yolunu kullanır. Çünkü hata, ortalaması sıfır olan tesadüfi bir değişkendir ve ancak sonsuz sayıda ölçme yapıldığı zaman bu değer sıfır olur. Bu sebeple sonsuz sayıda yapılan ölçme sonucunda elde edilen kişinin gözlenen puanının ortalaması, o kişinin gerçek puanına eşit olur. KTK'da bireylerin belirli bir psikolojik özelliğe sahip oluş derecelerinin belirlenmesi, her bireyin, o özelliği ölçmek üzere hazırlanmış olan ölçme aracının maddelerine vermiş oldukları tepkilerin toplamının alınması ile elde edilir. Yani bireyin ölçme aracından almış olduğu ham puan, onun ölçülen özelliğe sahip oluş derecesinin bir göstergesidir. Bu açıklamaya göre, ölçme aracından yüksek puan almış olan bir bireyin, ölçülmek istenen özelliğe yüksek düzeyde sahip olduğu, düşük puan almış olan bir bireyin ise, ölçülmek istenen özelliğe düşük düzeyde sahip olduğu ifade edilebilir. Buna ek olarak, madde ayırtecdilik gücü indeksi, madde güçlük indeksi gibi madde düzeyindeki bilgilerin elde edilmesine de imkan sağlamaktadır. Hem test hem de madde düzeyindeki bilgilerin elde edilmesi için kuram kapsamında değişik yöntem ve teknikler geliştirilmiştir (Fan, 1998).

Ondokuzuncu yüzyılın başlarından beri ölçeklerin geliştirilmesi, geliştirilen ölçek ve testlerin puanlanması ve yorumlanması için kullanılan KTK'nın bir takım yapısal sınırlılıkları olduğu bilinmektedir.

KTK'nın kullanımındaki sınırlılıkların çoğu, KTK ile geliştirilen testlerin psikometrik özelliklerinin (madde ayırtecdilik gücü, madde puanları ortalaması, bireylerin test puanları, geçerlik ve güvenilirlik, vs.) testin uygulandığı guruba bağımlı olmasından kaynaklanmaktadır (sample

dependency). Bu sınırlılığa ek olarak, bireylerin test ya da ölçek puanları, uygulanan testin maddelerine bağımlıdır (test dependency). Ayrıca, KTK kapsamında geliştirilen testlerde orta yetenek grubundaki bireyler temel olarak alındığı için, üst ve alt yetenek düzeyindeki bireylere yönelik doğru yetenek kestirimlerinin yapılmasına olanak vermemektedir. Ölçme hatalarına ilişkin varyansın tüm cevaplayıcılar için aynı kalması ise KTK'nın diğer bir sınırlılığıdır (Andrich, 1982).

KTK'nın yapısında bulunan sınırlılıkların bu kuram çerçevesinde geliştirilen test ve ölçeklerde ortaya çıkmasıyla ölçeklerle ölçülmek istenen özelliğe ilişkin yapılan ölçümlerdeki hata payı oranlarının yüksek çıkması ve elde edilen sonuçların etkin kullanımının güçleşmesi, alternatif kuramların geliştirilmesi çalışmalarına ışık tutmuş ve sonucunda ÖÖK geliştirilmiştir (Andrich, 1982).

### **Örtük Özellikler Kuramı**

Bu kuramla ilgili çalışmalar 1920'li yılların başlarında Thurstone, Thorndike, Symonds, ve Horst tarafından yapılmış ve daha sonraki yıllarda ise Lawley, Lord ve Novick ÖÖK ile ilgili çalışmalara devam ederek kuramın geliştirilmesinde önemli çalışmalar yapmıştır (Thissen, 1991).

Tutum ölçeklerinin geliştirilmesinde kullanılan modern ölçme kuramlarından birisi olan ÖÖK sıklıkla Madde Tepki Kuramı (MTK), Güçlü Gerçek Puan Kuramı ya da Modern Zihinsel Test Kuramı olarak da bilinmektedir (Andrich, 1982).

ÖÖK bireyin bir testle ölçülmek istenen özelliğe sahip oluş derecesi ile bu testin bir maddesine verdiği tepki arasındaki ilişkiyi ortaya koyar. KTK'dan farklı olarak bireyin maddeye cevap verirken gösterdiği davranışları hakkında daha güçlü varsayımlar ortaya koyar (Stocking 1997).

ÖÖK temel varsayımı; testi alan kişinin o teste ilişkin performansının; bireyin özellikler olarak adlandırılan yeteneklerinin tanımlanması ve bu özelliklerle ilişkili olarak bireylerin yetenek puanlarının

tahmin edilerek, madde ve test performanslarının açıklanıp, kestirilebileceği esasına dayanmaktadır (Hambelton ve Swaminathan, 1985).

ÖÖK, bireylerin yeteneklerinin gösterildiği psikolojik uzay ile bu uzayı temsil etmek amacıyla hazırlanan madde uzayı arasındaki bağıntılara dayanır.  $k$  tane özellik ve bu özelliklerin her birinin en az bir madde tarafından temsil edildiği  $n$  maddeden oluşan bir test düşünülürse,  $k$  tane özelliğin tümünü örten ve her biri bir özelliğe karşı gelen, özelliklerden maddelere bir bağıntı tanımlanmış olur. Bağıntıda yer alan  $k$  tane özelliğin her biri, özellikler kümesinin elemanları olup bireylerin bir özelliğini ifade eder ve bir psikolojik boyut olarak kabul edilir ve bu nedenle bu özelliklere "örtük özellikler" adı verilir. Ölçmeye konu olabilecek her bir birey, her biri bir özelliği temsil eden  $k$  boyutlu bir uzaydaki konumu açısından düşünüldüğünde bu uzay "örtük özellikler uzayını" (latent space) oluşturur.  $\theta$  örtük özellikler uzayı  $\theta \equiv (\theta_1, \theta_2, \dots, \theta_k)$  vektörü ile gösterilebilir (Gelbal, 1994; Lord ve Novick 1968, 359).

ÖÖK içerisinde bulunan modellerin birçoğu, bireyin bir maddeye ilişkin performansının belirleyicisi olan tek bir örtük özelliğin var olduğunu savunmaktadır. ÖÖK bireyin ölçülen özelliği ile maddeye ilişkin gösterdiği performans arasındaki ilişkinin matematiksel olarak ortaya konulabileceğini varsaymaktadır. Bu matematiksel tanım madde karakteristik fonksiyonu ve eğrisi olarak adlandırılır. İki kategorili puanlamaya uygun olan maddeler için madde karakteristik fonksiyonu, belirli bir yetenek düzeyindeki bireyin maddeyi doğru cevaplama olasılığını matematiksel olarak göstermektedir. Bu şartlı olasılık madde parametrelerinin bir fonksiyonudur. Genel olarak, bu şartlı olasılığı göstermek amacıyla seçilen matematiksel fonksiyon ya normal ogive ya da lojistik ogive dir (Stocking, 1997).

Ölçülen örtük özellik tek boyutlu ise madde karakteristik fonksiyonu madde karakteristik eğrisi adını alır. Madde karakteristik fonksiyonu, madde puanının  $\theta$  vektörü üzerindeki regresyonudur ve  $P_i(\theta)$  ile gösterilir.  $P_i(\theta)$  fonksiyonu gözlenen madde tepkisinden hareketle gözlenemeyen özelliklerin kestirilmesinde kullanılır. Doğrusal olmayan (Nonlinear) bir

eğridir ve tam örtük uzay tanımlanmışsa madde karakteristik eğrisi farklı gruplarda değişmezliğini korur. Madde karakteristik eğrisi, belirli bir yetenek düzeyindeki kişinin bir maddeyi doğru cevaplandırma olasılığı hakkında bilgi verir (Hambelton ve Swaminathan, 1985). Madde karakteristik fonksiyonu örtük özellikler uzayının tek boyutlu olması halinde monoton artan bir fonksiyondur ve bu eğri "madde karakteristik eğrisi" olarak adlandırılır (Baykul, 1979).

ÖÖK'da madde karakteristik fonksiyonlarının matematiksel yapılarına bağlı olarak farklı modeller ortaya çıkmıştır. Bu modellerin kullanılmasında örtük özellikler kuramının bazı varsayımlarının karşılanması gerekmektedir. Bu varsayımlar, dağılımın normal olması, tek boyutlu olması ve yerel bağımsızlıktır. Bu varsayımlar karşılandıktan sonra model-veri uyumunun test edilmesi gerekmektedir.

Kategorik puanlamaya elverişli ölçme araçlarının geliştirilmesi ve puanların yorumlanmasında 1, 2 ve 3 parametrelili normal ogive ve lojistik modeller kullanılmaktadır. Normal ogive modellerdeki matematiksel işlem güçlükleri daha çok lojistik modellerin kullanılmasını gündeme getirmiştir.

**Örtük Özellikler Kuramı Modelleri:** Bu modellerden bir parametrelili model için madde karakteristik fonksiyonu aşağıdaki eşitlik kullanılarak elde edilir.

$P_i(\theta)$ :  $\theta$  yetenek düzeyindeki bir bireyin maddeyi doğru cevaplandırma olasılığı,

D : Sabit değer, 1.7

$b_i$  : Maddenin güçlük düzeyini göstermek üzere,

$$P_i(\theta) = \frac{e^{Da(\theta - b_i)}}{1 + e^{Da(\theta - b_i)}} \quad (i = 1, 2, 3, \dots, n) \quad (1)$$

Bir parametrelili model kullanılarak her madde için kestirilen değer madde güçlük indeksi olan "b" parametresidir. Bu model aynı zamanda

Rasch modeli olarak da bilinir. Aralarındaki fark bir parametrelili modelde kullanılan  $a_i$  parametresi sabit bir değer olan 1 olarak kabul edilirken, Rasch modelinde  $a_i$  değerlerinin ortalamaları alınarak sabitlemiştir. Bu modelde bütün maddelerin eşit madde ayırt edicilik gücüne sahip olduğu düşünülür. Bir maddenin  $b$  parametresi, maddenin 0.50 doğru cevaplandırma olasılığına  $\theta$  yetenek ölçeği üzerinde karşılık gelen değerdir. Bir parametrelili modelde maddenin  $b$  parametresinin değeri arttıkça, bireylerin o maddeye ilişkin doğru cevaplandırma yetenek düzeyi de artmaktadır. Herhangi bir yetenek düzeyindeki grubun puanlarını  $x=0$  ve  $Sx=1$  olarak transfer edilirse uygulamada bu parametrenin -2 ile +2 aralığında değer aldığı görülür. Daha kolay olan maddelerde bu değer -2 ye zor maddelerde ise değer +2 ye yaklaşır. Madde güçlük indeksi 0.00 olan maddeler orta güçlükte olan maddeler olarak kabul edilir (Hambleton ve Swaminathan, 1985).

Bu modelin en önemli sınırlılığı, bütün maddelerin eşit ayırt edicilik düzeyine sahip olduğu varsayımdır. Maddelerin aynı ayırt edicilik gücüne sahip olması aynı şekilde sahip olan madde karakteristik eğrilerine sahip olmasına neden olmaktadır. Bu durum kuramsal olarak ifade edildiğinde herhangi bir problem ortaya çıkmazken uygulama açısından ölçme araçlarındaki bütün maddelerin aynı ayırt edicilik gücüne sahip olması sık rastlanan bir durum değildir. Bu nedenle, madde karakteristik fonksiyonları hesaplanırken maddelerin ayırt edicilik gücünün de ( $a$  parametresi) dikkate alınması gerekliliği vardır.

Bir parametrelili modeldeki bu sınırlılığı ortadan kaldırmak için farklı bir modelin geliştirilmesi gerekliliği ortaya çıkmıştır. Bu nedenle iki parametrelili lojistik model geliştirilmiştir. İki parametrelili lojistik modelinde, madde karakteristik fonksiyonu aşağıdaki eşitlik yardımıyla elde edilir.

$$P_i(\theta) = \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad (i = 1, 2, 3, \dots, n) \quad (2)$$

Eşitlikte ifade edilen  $P_i(\theta)$  ve  $b_i$  parametreleri Eşitlik 5'te verilen anlamlarını korumaktadır. Bu modelde bir parametrelili modelden farklı olarak



her madde için ayrı bir “a” parametresinin elde edilmesidir. İki parametrelili modelde maddeler sahip oldukları “b” değerlerinin yanı sıra “a” değerleri bakımından da değişkenlik göstermektedir. a parametresi kuramsal olarak “-∞ ile +∞” arasında değişen değerler alır. Ancak uygulamada, a parametresine ilişkin değerlerin daha çok 0.75 ile 2.50 arasında değişen değerler aldığı gözlenmektedir. Buna göre, herhangi bir maddenin “a” parametresine ilişkin değeri yükseldikçe o maddenin ayırt edicilik gücü de artmaktadır.

ÖÖK kapsamında geliştirilen bir parametrelili modelde olduğu gibi iki parametrelili modelinde sınırlılıkları vardır. İki parametrelili modelin sınırlılığını bireylerin soruları şans faktörüne dayalı olarak cevaplama durumunda, bu faktörün madde parametreleri üzerindeki etkisini göz ardı etmesidir. Bu nedenle şans faktörü iki parametrelili modele eklenmiş ve “c” parametresi olarak ifade edilen bir değer elde edilebileceği üç parametrelili model geliştirilmiştir.

Üç parametrelili modelde madde karakteristik fonksiyonu aşağıdaki eşitlik ile ifade edilir.

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}} \quad (i = 1, 2, 3, \dots, n) \quad (3)$$

$c_i$  parametresi bir maddenin doğru cevaplandırılabilmesi için gereken en düşük yetenek seviyesini gösterir. Bu nedenle  $c_i$  parametresi şans başarısını ifade eden bir parametre olarak da adlandırılır. Bir ve iki parametrelili modellerde bu değer 0 olarak kabul edilir (Hambelton ve Swaminathan, 1985).

ÖÖK temel olarak test maddelerinin 0 (sıfır) ve 1 tarzında puanlandığı yetenek ölçümlerinde madde ve yetenek parametrelerinin kestirilmesine yönelik modellere ağırlık verilmiştir. Sıfır ve 1'in dışında puanlama yapılan yetenek, tutum ve kişilik ölçümlerinde madde ve yetenek parametrelerinin kestirimi için kullanılacak modeller üzerine de yapılan ilk

çalışmalar Samejima (1972) ile başlamıştır.

Yetenek ölçmelerinde 0 (sıfır) ve 1 'in dışında puanlama yapılan test maddelerinde, maddeye verilecek tepki basamaklandırılmıştır. Birinci basamak cevaplandırıldığında bir değerlendirme yapılır, en son basamak doğru cevaplandırıldığında tam puanlama yapılır. İki basamaklı bir madde de birinci basamak yanlış cevaplandırıldığında 0 (sıfır) puan verilir, birinci basamak yanlış cevaplandırıldığında ikinci basamakta doğru cevap alınmaz. Bu durumda üç puan kategorisi bulunmaktadır (0, 1 ve 2). Başarıyla tamamlanan basamak sayısı arttıkça kategori puanı da artar, böylece yüksek kategori puanı daha yüksek yetenek düzeyini gösterir.

Test maddelerine verilecek tepkinin kredilendirildiği durumlarda kullanılmak üzere Masters tarafından "kısmi puanlama" modeli (Partial Credit) geliştirilmiştir (De Ayala, 1993).

Masters'ın Kısmi Puanlama Modeli Rach Modelinin uzantısıdır. Belirli yetenek seviyesinde i maddesinin her bir cevap kategorisinin cevaplandırılma olasılığı.

$$P_{nxi}(Q_n) = \frac{\exp \left[ \sum_{k=0}^{xi} (Q_n - b_{ik}) \right]}{\sum_{j=0}^{mi} \exp \left[ \sum_{k=0}^j (Q_n - b_{ik}) \right]} \quad (4)$$

eşitliğiyle belirlenir (De Ayala, 1993: 175). Bu model de madde basamakları sıra ile tamamlanır, basamakların eşit güçlükte olması gerekmediği gibi güçlüklerine göre sıralanması da gerekmez. Kısmi puanlama modelinde bütün maddelerin bireyleri ayırt etmede eşit ayırt edicilik gücüne sahip olduğu varsayılır. Kısmi puanlama modeli Likert tipi tutum ölçekleri için yetersiz kalmaktadır.

Andrich tarafından "Dereceleme Ölçeği" (Rating Scale) modeli Likert tipi tutum ölçekleri için geliştirilmiştir. Andrich "Dereceleme Ölçeği" modelinde kısmi puanlama modelinden yararlanmıştır (De Ayala, 1993).

Likert tipi ölçekte 5-7 sabit noktalı ölçek kullanılmaktadır. Her bir maddedeki kategorilerin göreceli güçlüğü'nün bütün maddelerde aynı olduğu varsayılır. Ayrıca maddelerin bireyleri ayırt etmede eşit ayırt edicilik gücüne sahip olduğu da varsayılır.

$i$  maddesindeki her bir kategorinin cevaplandırılma olasılığı

$$P_{nxi}(Q_n) = \frac{\exp[k_x + K_i(Q_n - b_i)]}{\sum_{j=0}^m \exp[k_x + j(Q_n - b_i)]} \quad (5)$$

$b_i$  ; maddenin ölçek değeri, diğer bir deyişle maddenin ölçtüğü tutum düzeyi

$k_i$ ; geçilen eşik sayısı

$Q_n$  ; bireylerin tutum düzeyi

$X_i$ ; 0 olduğunda  $k_x=0$  olur, diğer durumlarda

$x=1, 2, 3, \dots, m$

$T_k$  ;  $m+1$  sıralı kategoriyi birbirinden ayırt eden eşik değer

Andrich bütün maddelerde kategori eşik değerinin aynı olduğunu sadece her bir madde için (bi) ölçek değerinin kestirilmesi gerektiğini savunmuştur (De Ayala, 1993).

**Aşamalı Tepki Modeli.** Kısmi puanlama ve dereceleme ölçeği modelinin dışında Samejima (1972) tarafından Aşamalı Tepki Modeli (ATM) geliştirilmiştir. ATM iki parametrelili logistik modelin uzantısıdır. Model de sınır veya eşik kavramı önem kazanmaktadır. Bireyin  $k$ 'inci veya daha düşük kategoriyi işretleme olasılığı belirlenmeye çalışılır.

Maddeye verilecek olası cevap kategorileri sıralanmıştır.  $M_{g+1}$  kategori bulunduğundan, puanlar  $X_g = 0, 1, \dots, m_g$  şeklindedir. Her bir maddeye verilen tepki  $X_g$  değeri ile ifade edilir.  $i$  bireyinin  $n$  maddeye verdiği cevap örüntüsü  $V = (X_1, X_2, X_3, \dots, X_n)$  vektörü ile gösterilir. Bireyin maddedeki herhangi bir kategoriyi veya ondan daha yüksek kategoriyi cevaplama

olasılığı  $P_{xg}$ ; ile gösterilir ve aşağıdaki eşitlik kullanılarak hesaplanır (De Ayala 1993).

$$P(x_{pi}) = \frac{\exp \sum_{i=0}^x a_i (\theta_p - b_i)}{\sum_{x=0}^m \exp \sum_{k=0}^x a_i (\theta_p - b_i)} \quad (6)$$

k'inci kategorinin işaretlenme olasılığından daha sonraki kategorinin cevaplandırılma olasılığı çıkarıldığında maddedeki belirli bir kategorinin cevaplandırılma olasılığı elde edilir.

$$P_{xg}(Q) = P_{xg}^*(Q) - P_{(xg+1)}^*(Q) > 0 \text{ ' dir.} \quad (7)$$

Ölçekte bulunan herhangi bir maddenin ölçeğin bütünüyle yapılan ölçmenin doğruluğuna katkısı madde bilgi fonksiyonu ile belirlenir (item information function) (Hambleton ve Swaminathan 1985). Test maddesinin herhangi bir kategorisinin maddeyle yapılan ölçmenin doğruluğuna, güvenilirliğine katkısı kategori bilgi fonksiyon ile belirlenir. Kategori bilgi fonksiyonlarının toplanmasıyla madde bilgi fonksiyonu elde edilir.

Madde bilgi fonksiyonu  $\theta$  yeteneğine ait fonksiyondur. Her bir yetenek seviyesinde madde ve test bilgi fonksiyonu hesaplanır. Madde bilgi fonksiyonu maddenin ölçtüğü yetenek düzeyi (b) bireylerin yeteneğine yakın veya eşitse madde hakkında daha fazla bilgi verir. Ayrıca maddenin ayırt edicilik değeri yükseldikçe ve ölçmenin standart hatası azaldıkça madde bilgi fonksiyonu değeri artar. Bilgi fonksiyonu ölçmenin standart hatası ile ters orantılıdır (De Ayala,1993).

$$I(\theta_m) = \sum_{x=0}^{m_t} \frac{((P_{xi}(\theta) - (P_{(x+1)}(\theta)))^2}{((P_{xi}(\theta) - (P_{(x+1)}(\theta)))^2} \quad (8)$$

SE ( $\theta$ );  $\theta$  yetenek seviyesinde yetenek kestiriminin standart hatası KTK'da güvenilirlik katsayısının fonksiyonu olan ölçmenin standart hatası ölçmenin yapıldığı gruba bağlıdır (Hambleton ve Swaminathan, 1985).

$$SE^*(\theta) = \frac{1}{\sqrt{I(\theta)}} \quad (9)$$

Madde ve test bilgi fonksiyonu örneklemeden bağımsız olarak elde edilir ve her bir yetenek düzeyinde ölçmenin standart hatasını yansıtmaktadır. Bu nedenle madde ve test bilgi fonksiyonu KTK'ndaki güvenilirlik ve standart hata kavramlarına alternatif olarak kabul edilir (Berberoğlu, 1988).

Tek boyutlu çok kategorili modeller sıralı ve sıralı olmayan modeller olmak üzere sınıflandırılmaktadır. Sıralı çok kategorili modeller içinde Samejima (1969) tarafından geliştirilen "Aşamalı Tepki Modeli" (Graded Response Model), Masters (1982) tarafından geliştirilen "Kısmi Puan Modeli", Andersen (1987) tarafından geliştirilen "Oranlı Ölçek Modeli", Verhelset, Glas ve Vries (1988) tarafından geliştirilen "Adımlar Modeli", Tutz (1990) tarafından geliştirilen "Ardışıklık Modeli" ve Muraki (1992) tarafından geliştirilen "Genelleştirilmiş Kısmi Puan Modeli" yer almaktadır (De Ayala, 1993; Lord ve Stocking 1997).

ATM sıralı çoklu kategorilerle uğraşan matematiksel modellerin bir kolunu temsil etmektedir. Bu sıralı kategoriler öğrencilerin performansını değerlendirmede kullanılan A, B, C, D, E, ve F gibi harf derecelerini, Likert tipi ölçeklerde kullanılan kesinlikle katılıyorum, katılıyorum, katılmıyorum ve hiç katılmıyorum derecelemelerini veya bir soruyu puanlamada cevaplayıcıların erişti derecesiyle uyum içerisinde verilen kısmi puan gibi oranları içermektedir (De Ayala, 1993).

Model, ÖÖK'nın bilinen diğer varsayımlarına ek olarak, cevaplayıcıların cevapladığı kategorilerin sıralanabilir olduğunu varsaymaktadır. Model yardımıyla, cevaplayıcıların verdiği tepkinin doğru ya da yanlış olduğunu belirlemekten çok verilen tepkiden mümkün olduğunca daha fazla bilgi elde edilmeye çalışılmaktadır (Hambleton ve Swaminathon, 1985).

ATM, bir bireyin  $k$  veya daha yüksek bir kategoriye cevaplama olasılığını,  $k$ 'dan düşük kategorileri cevaplama olasılığına göre belirler. Bu modelde cevaplayıcı  $m_i+1$  kategoriye ayrılmış bir  $i$  maddesini tepki.  $m_i$ ,  $i$  maddesindeki kategori sınırlarının sayısı olarak ifade edilir.  $i$  maddesinin her bir kategorisi ile ilgili olarak  $x_i$ , 0, 1, ...,  $m_i$  tamsayı değeriyle kategori puanı olarak adlandırılır. Yüksek kategoriler yüksek yeteneği gösterir (De Ayala, 1993).

Likert tipi tutum ölçeği geliştirme çalışmaları KTK çerçevesinde uzun yıllar devam etmiş daha sonra örneklemden bağımsız test geliştirme amacı gündeme gelmiştir. Bu amaç doğrultusunda Reise ve Yu 1990 yılında 25 maddelik bir test kullanarak ATM üzerine yaptıkları çalışmada örneklem büyüklüğünün yetenek kestirimlerini etkilemediğini, buna rağmen madde parametrelerinin kestirimini oldukça etkilediğini bulmuşlardır. Araştırmacılar ATM ile madde parametrelerine ilişkin doğru ve kesin kestirimlerin yapılabilmesi için örneklem grubunun en az 500 denekten oluşması gerektiğini belirtmişlerdir.

ATM'nin tutumların ölçülmesinde kullanıldığı diğer bir çalışmada Koch (1983) öğretmenlerin okul yöneticilerinin iletişim becerilerine karşı tutumlarını incelemiştir. Samejima (1969) yılında ATM'ni sözel olmayan usa vurma yeteneklerinin ölçülmesine uyarlamıştır. Dodd (1984) Aşamalı Tepki ve Kısmi Puanlama Modellerini tutumların ölçülmesi çalışmalarında kullanmıştır. Bu modellerin tutumların ölçülmesinde kullanıldığı araştırmalar geleneksel çözümlene yöntemlerine paralel yüksek ilişkili sonuçlar vermiştir. Bu araştırmalar aynı zamanda tutumların ölçülmesinde her bir birey için sonuçların kesinliği ve örneklemden bağımsız ölçek geliştirme avantajlarını ortaya koymuştur.

Yapılacak araştırmalarla, öğrencilerin olumlu ya da olumsuz tutumları ortaya konulabilir. Bu araştırmalarla, özellikle olumsuz tutumların kaynağı belirlenebilir ve tutumları olumsuz etkileyen faktörler ortadan kaldırılabilir veya en aza indirilmesi sağlanabilir. Olumsuz tutumların olumluya dönüşmesi ile Beden Eğitimi derslerine katılım artırılabilecek ve bu, gelecekte fiziksel olarak

eđitilmiş sađlıklı bireylerin topluma kazandırılmasını sađlayacaktır. Öğrencilerin tutumlarının, öğretme-öđrenme sürecinde oynadıđı rolü anlayabilmek ve yakalayabilmek için tutumları ölçmenin çok büyük bir önemi vardır.

Öđrencilerin beden eđitimine yönelik tutumları ile çok ilgilenilmesine karşı çok az bilgiye ulaşılmıştır. Ölçme araçlarının geliştirilmesinde yararlanılan ölçme kuramlarının gelişmesiyle ve bu araçların beden eđitimi araştırmalarında kullanılmasıyla daha fazla bilgi edinilebilir. Beden eđitimi alanı daha çok tutum üzerine olan araştırmalara eğilim göstermektedir. Bunun, uygulamalı ölçmenin ve beden eđitimi öğretiminin üzerinde önemli bir etkisi vardır. Ölçmenin ilkelerini alanyazındaki araştırmalarla birleştirmek beden eđitimi dinamiklerinin daha iyi anlaşılmasını sađlar.

Beden eđitimine ve fiziksel aktiviteye yönelik tutumları belirlemeyi amaçlayan araştırmalarda kullanılan ölçme araçlarının geçerlik ve güvenilirliğine ilişkin alanyazında yeterli bilgi olmaması, bu konuya odaklanmış araştırmaların sonuçları arasında gözlenen tutarsızlıkların nedeni olarak görölmektedir. Araştırmalardan elde edilen sonuçların farklı çıkmasının başka bir nedeni ise yirmi yıl önce geliştirilmiş araçlara olan güven ile ilgilidir. Karmaşık psikometrik analizlerin olmadığı bir zamanda geliştirilen araçları kullanan araştırmacılar belki de çok fazla geçerli sonuçlar elde edemeyeceklerdir. Ayrıca önceki araştırmalarda beden eđitimine karşı öğrencilerin tutumunu etkileyen ebeveyn etkileri, medya, beceri seviyesi, spor geçmişı göz önüne alınmamıştır. Öğrencilerin beden eđitimine karşı olan tutumu çok boyutlu bakış açısıyla ele alınmalıdır. Tutumun çok boyutlu bakış açısını yakalamak için araştırmacının bu bakış açısından tutumun kavramsal tanımlarını yapması gerekir. Bu kavramsal tanımlama tutum araçlarının geliştirilmesinde vazgeçilmez bir öncü olacaktır (Graham, 1995).

Beden eđitimi'ne yönelik tutumların ölçülmesinde kullanılan tutum ölçeklerinin KTK'na dayalı olarak geliştirilmiş olması ve bu ölçme araçlarının geçerlik ve güvenilirliklerinde sınırlılıklar bulunması nedeni ile ÖÖK'nın kullanıldığı ve bu yöntemlerin psikometrik varsayımlarını karşılayan aynı

zamanda klasik ölçme kuramlarının da varsayımlarını destekleyen yeni tutum ölçeklerinin geliştirilmesi gerekliliği ortaya çıkmıştır. Burada bahsedilen ÖÖK'nın KTK'ya göre tercih edilme nedenlerinin iyi anlaşılması gereklidir. Aşağıda bu konu ile ilgili açıklayıcı bilgiler verilmiş ve araştırma problemine ilişkin genel bir çerçeve oluşturulmuştur.

KTK'da genel olarak maddeler seçilirken, madde güçlük indeksi değerleri ve madde-toplam korelasyonları kullanılır. Test geliştirme sürecinde madde güçlük indeksleri ve madde-toplam korelasyonları en yüksek maddeler iç tutarlılık değeri yüksek bir ölçek geliştirmek ve hata oluşumunu en aza indirmek için seçilir. Buna ek olarak toplam puan dağılımları elde edilerek bu değerler araştırmacının hedeflediği dağılım puanları ile karşılaştırılır. Daha sonra bazı maddelerde değişiklikler yapılarak araştırmacının hedeflediği puan dağılımı ile en yakın puanların elde edilebileceği nihai madde havuzu oluşturulur. Madde güçlük değerleri sıklıkla madde eksiltme esnasında kullanılır.

ÖÖK'nın KTK karşısında birtakım avantajları vardır. Bunlar temel olarak test geliştirme yöntemleri ve ölçmede eşitlik ilkesinin korunmasıdır. Maddelere ve tutumları ölçülen gruba bağlı olan KTK'ndan farklı olarak ÖÖK da teste dahil edilen ve çıkarılan maddelerin her birisinin ayrı ayrı değerlendirilebilmesidir. Buna ek olarak ÖÖK araştırmacılara durumsal olarak test bilgi fonksiyonu ile ölçmenin standart hatasını hesaplama olanağı tanırken, KTK bütün yetenek düzeyleri için tek bir ortalama standart hata hesaplama olanağı tanımaktadır (Hulin, Drasgow & Parsons, 1983). Bu, araştırmacıya belirlenen yetenek/özelliik aralıklarında maksimum ölçme imkanı sağlayacak maddeleri seçme imkanı sağlar (Heinen, 1996).

KTK istatistikleri örnekleme bağımlıdır. Grubun ölçülen özelliğindeki ortalama puanlardaki değişimlere paralel olarak değişiklikler gösterirler. Bu nedenle test geliştiren araştırmacılar madde kalibrasyonu için kullanacakları örneklem grubunu seçerken oldukça dikkatli davranmalıdırlar. Eğer ölçümleme (kalibrasyon) örnekleme ölçüt örnekleme grubundan farklı olursa testin psikometrik özellikleri de dramatik olarak değişiklik gösterir.



Buna ek olarak ÖÖK'nin KTK'na göre avantajlarından biri ÖÖK'nin arařtırmacıya güvenilirliđin artırılması için daha fazla bilgi sađlamasıdır. Bu, her bir madde için ayrı ayrı psikometrik özelliklerin ortaya konması ile gerçekleştirilir. Daha önce de belirtildiđi gibi, ÖÖK sıklıkla Güçlü Gerçek Puan Kuramı ya da Modern Zihinsel Test Kuramı olarak adlandırılmasının nedeni ÖÖK'nin daha güncel bir kuramsal yapı olması ve açıklıđa kavuşturulmamıř denenceleri açıklayabilecek bir yapıya sahip olmasıdır.

KTK ve ÖÖK çođunlukla aynı problemler için kullanılırlar fakat temelde farklı kuramsal yapılarıdır ve farklı yöntemleri kullanılırlar. Andrich, (1982) iki kuramın genel olarak tutarlı ve modern olmasına karřın belirli noktalarda birbirlerinden farklılık gösterdiđini belirtmektedir. Bu noktalar;

- Birçok durumda ÖÖK, KTK'na göre daha güçlü bulgular ortaya koyar ve daha güçlü varsayımlar yapılmasına olanak verir. Özellikle hatanın hesaplanmasında daha güçlü bulgular ortaya koyar. Ancak bu ÖÖK varsayımlarının tam olarak karřılanması ile mümkün olabilir.
- KTK bulguları önemli uygulama sonuçları vermesine karřın, ÖÖK'nin model bađımlı yapısı KTK'nın analog yapısına karřın avantajlar sađlamaktadır
- KTK'nın yararlarından birisi test puanlarının hesaplanması ve açıklanabilmesindeki kolaylıktır. Buna karřın ÖÖK daha karmařık deđerlendirme süreçleri içerir.
- ÖÖK maddelerin ve bireylerin ölçeklenmesinde gelişim göstermiştir. Bu özellik ÖÖK modellerine bađlıdır. Ancak çođu model madde güçlüđünü ve bireylerin yetenek/tutum düzeylerini aynı skala üzerinde ölçeklemektedir. Bu sayede madde güçlük deđerleri ve bireylerin yetenek/tutum düzeyleri gerçek anlamda karřılařtırılabilmektedir.

Beden eđitimi'ne yönelik tutumların ölçülmesi amacıyla yapılan çalışmalarda; a)KTK'na dayalı olarak geliştirilmiş ve kullanılan ölçeklerde,

daha önce de sözü edilen geçerlik ve güvenilirlik sınırlılıkları, modern psikometrik analizlerin kullanılmaması, b) tutum ölçeği geliştirme çalışmalarında tutum kuramının sağlam temellere dayanmaması ve tutumun temel yapısını ortaya koyan tutum kuramının göz ardı edilmesi gibi önemli sorunlar bulunmaktadır. Bu sorunlara çözüm bulmak amacıyla yapısal olarak daha güçlü tutum ölçeğinin geliştirilmesi hedeflenmiştir. Bu noktada, yapısal olarak daha güçlü bir ölçeğin geliştirilmesi aşamasında kullanılacak kuramsal yapılar ve bu yapılarla bağlı yöntemler önem kazanmaktadır. Bu nedenle, KTK ve ÖÖK'nin beden eğitimi dersine yönelik tutum ölçeği geliştirme çalışması kapsamında incelenerek karşılaştırılması, tutum ölçeği geliştirme çalışmalarında kullanılan kuramsal yapıların ortak güçlü yanlarının, birbirlerine olan yarar ve sakıncalarının ortaya konması açısından gerekli görülmüştür.

### **Amaç**

Bu çalışmanın amacı, KTK ve ÖÖK'na göre psikometrik özelliklerin, İlköğretim 6-7-8. sınıf öğrencilerinin (12-15 yaş) beden eğitimi dersine yönelik tutumlarını ölçen, Likert tipi bir ölçek hazırlama çalışması kapsamında incelenmesi ve karşılaştırılmasıdır.

Bu genel amaç doğrultusunda aşağıda belirtilen sorulara cevap aranmıştır.

1. Korelasyona dayalı analiz, içtutarlılık ölçütüne dayalı analiz ve Aşamalı Tepki Modeli'ne dayalı olarak kestirilen madde ayırıcılık gücü değerleri arasında ne düzeyde bir ilişki vardır?,
2. Korelasyona dayalı analiz, içtutarlılık ölçütüne dayalı analiz ve Aşamalı Tepki Modeli'ne dayalı olarak hazırlanan ölçeklerin güvenilirlikleri ne düzeydedir?,

3. Korelasyona dayalı analiz, içtutarlılık ölçütüne dayalı analiz ve Aşamalı Tepki Modeli'ne dayalı olarak hazırlanan ölçeklerin test bilgi fonksiyonları nedir?,
4. Klasik Test Kuramı'na ve Örtük Özellikler Kuramı'na göre hazırlanan asıl ölçeğin geçerlik düzeyi nedir?,
5. Klasik Test Kuramı ve Örtük Özellikler Kuramı'na göre hazırlanan asıl ölçeğin güvenirlik düzeyi nedir?

### **Önem**

Bu çalışma kapsamında beden eğitimi dersine yönelik tutumları ölçen bir tutum ölçeği geliştirilirken KTK'nın madde yanlılıkları ÖÖK ile aşılmaya çalışılmıştır. KTK'nda madde parametreleri örneklem grubunun tepkileri doğrultusunda yanlılık göstermektedir. ÖÖK'da ise örneklem grubundan bağımsız olarak madde parametreleri kestirilebilmektedir. Çalışma kapsamında KTK ve ÖÖK, madde ve ölçek parametreleri kapsamında karşılaştırılmış ve bu iki kuram test geliştirme kapsamında incelenmiştir. Bu tür bir çalışma, tutum ölçeği geliştirme çalışmalarında geçerlik ve güvenirlik kestirimlerinin doğru yapılabilmesi ve yapısal olarak daha güçlü ölçme araçlarının geliştirilmesi açısından çok önemlidir. Bu nedenle oluşturulan tutum ölçeğinin geçerlik ve güvenirlik çalışmaları her iki kuramın ortak güçlü yönlerinden faydalanılarak yapılmıştır.

Tutumlar genellikle çocukluk döneminde şekillenmeye başlar ve deneyim, pekiştirme, taklit ve sosyal öğrenme yoluyla yerleşirler. Büyümeyle paralel olarak, özellikle ergenlik döneminde sosyal etkilerin rolü artmaktadır. Tutumların büyük bölümü bireylerde genellikle 12-30 yaş aralığında son şeklini almaktadır. Bu dönem kritik dönem olarak da adlandırılmaktadır (Tavşancıl, 2002). Kritik dönem olarak adlandırılan ve tutumların büyük ölçüde şekillendiği bu yaş aralığı içerisinde özellikle 12-15 yaş grubu öğrencilerin Beden Eğitimi dersine yönelik tutumlarını ölçen bir ölçeğin geliştirilmesi ve tutumlarının bu ölçek yoluyla ortaya konularak gelecekteki

olumsuz yaklaşımların birçok spor dalının başlangıç yaşı olan bu dönemde yapılan bu çalışma ile Beden Eğitimi dersine ilişkin tutumların belirlenmesine, öğrencilerin bu derse yönelik olan olumsuz tutumları varsa, bunlara ilişkin iyileştirici çözümler üretilmesine, öğrenme ve öğretme sürecine ve alanyazına katkıda bulunacağı düşünülmektedir.

### **Sınırlılıklar**

Bu araştırma çerçevesinde çalışma grubu olarak sadece 2005-2006 eğitim öğretim yılının bahar döneminde öğrenim gören 12-15 yaş arası ilköğretim 6, 7, ve 8. sınıf öğrencileri alınmıştır. Çalışma grubuna 12-15 yaş grubu öğrencilerin dahil edilme nedeni, bu yaş grubunun birçok takım sporu (basketbol, voleybol vb.) ve bireysel sporlar (boks, taekwondo, bisiklet vb.) için aktif spora başlama yaşı olarak belirlenmiş olmasıdır (Bompa, 1999: 35; Dick, 2002: 303; Bompa, 2000: 11). Ayrıca, araştırma dahilinde tutum ölçeklerinin madde analizlerinin ÖÖK kapsamında yapılmasında politomik verilere uygulanabilen ÖÖK modellerinden sadece Aşamalı Tepki Modeli kullanılmıştır (Hambleton ve Swaminathan, 1985).

### **Tanım**

Tutum; “Bireyin kendine ya da çevresindeki herhangi bir toplumsal konu, obje ya da olaya karşı deneyim, inanç ve bilgilere dayanarak örgütlediği bilişsel, duygusal ve davranışsal tepki ön eğilimi” olarak tanımlanabilir.

### **KISALTMALAR**

|             |       |                            |
|-------------|-------|----------------------------|
| <b>KTK</b>  | ..... | Klasik Test Kuramı         |
| <b>ÖÖK</b>  | ..... | Örtük Özellikler Kuramı    |
| <b>ATM</b>  | ..... | Aşamalı Tepki Modeli       |
| <b>BE</b>   | ..... | Beden Eğitimi              |
| <b>BED</b>  | ..... | Beden Eğitimi Dersi        |
| <b>BETÖ</b> | ..... | Beden Eğitimi Tutum Ölçeği |

## BÖLÜM II

### İLGİLİ ARAŞTIRMALAR

#### Yurt İçinde Yapılan Araştırmalar

Baykul (1979) yılında yaptığı araştırmasında KTK'na göre geliştirdiği bir matematik yetenek testini, ÖÖK'nın üç parametrelili lojistik modeli ile geliştirdiği bir matematik yetenek testinin psikometrik özellikleri ile karşılaştırmıştır. Araştırma bulgularına göre puanların şans başarısından arındırılmış ve arındırılmamış hallerinin her ikisinde de üç parametrelili lojistik model ile elde edilen puanlardan kestirilen KR-20 güvenilirlik katsayısının KTK'na göre hesaplanan değerden daha yüksek olduğu bulunmuştur. Ancak üç parametrelili lojistik model ile elde edilen puanlardan kestirilen KR-21 güvenilirlik değerleri ile KTK'na göre hesaplanan değerler arasında manidar bir farkın olmadığı sonucuna ulaşılmıştır. KTK ve ÖÖK'ndan elde edilen puanların genel yetenek puanları ile olan ilişki düzeyi incelendiğinde, KTK'ndan elde edilen değerlerin daha yüksek olduğu ancak ders başarısı ile olan ilişki düzeyinde anlamlı bir farklılık olmadığı sonucuna ulaşılmıştır.

Berberoğlu (1988) yaptığı çalışmada Öğrenci Seçme Sınavı alt testlerinin hazırlanmasında Rasch modelinin kullanılmasının sağlayacağı faydaları ortaya koymayı amaçlamıştır. Bu amaç doğrultusunda 1986 Öğrenci Seçme Sınavı sonuçlarına göre üç ayrı yetenek düzeyini temsil eden 3390 kişilik bir örneklem grubu kullanılmıştır. Araştırmada ilk olarak kullanılan puanların Rasch modelinin sayıtlılarını karşılayıp karşılamadığı üzerinde çalışmalar yapılmıştır. Bunun yanında, Rasch modelinden yararlanılmasının örneklemden bağımsız madde ölçekleme olanağı verip vermediği incelenmiş ve farklı örneklem gruplarından Rasch modeli kullanılarak elde edilen madde güçlük değerleri karşılaştırılmıştır. Araştırma bulgularına göre Rasch modeli kullanılarak örneklemden bağımsız olarak madde parametrelerinin kestirilebileceği sonucuna ulaşılmıştır.

Gelbal (1994) örtük özellikler kapsamındaki Rasch modeli ve KTK ile geliştirilen madde ve yetenek parametrelerinin benzerliklerini incelediği araştırmasında, madde güçlük indekslerini birim değerlere dönüştürerek, z ve b parametresi arasındaki ilişkilere bakmıştır. Rasch modeli, KTK ve Rasch modelinde b parametresi yerine madde güçlük indekslerinin birim normal değerleri kullanılarak kestirilen yetenek ölçütleri arasındaki benzerlikler üzerine çalışmıştır. Rasch modeli ve KTK'na göre test puanlarının normal dağıldığı ve dağılmadığı durumlarda madde ve yetenek parametreleri kestirilmiştir. Araştırma bulgularına göre KTK'nın p madde güçlük indeksi ile Rasch modeliyle kestirilen b parametresi arasında benzerlikler olduğu ve normal dağılımın sağlanamadığı durumlarda bu benzerliğin fazla etkilenmediği bulunmuştur. Buna ek olarak, kestirilen yetenek ölçütlerine göre, KTK ve Rasch modeli arasında benzerlikler olduğu tespit edilmiştir.

Başarı (1997) ÖÖK'nın ATM ve KTK'nın dereceleme toplamları ile ölçekleme tekniğine dayalı olarak geliştirilen Likert tipi matematik dersine yönelik tutum ölçeğinin her iki yöntemle kestirilen madde ve tutum parametreleri arasındaki ilişkileri incelemiştir. Matematik dersine yönelik tutum ölçeğinin madde parametreleri bakımından ÖÖK ve KTK'na göre kestirilen parametreler arasında yüksek ilişki olduğu bulunmuştur. Buna ek olarak, her iki kurama göre ölçeklerin yeterli güvenilirlik değerine ve ölçüt grup geçerliğine sahip olduğu ortaya konmuştur.

Benzer bir çalışmada Yıldız (1998) bir grup bireyin aynı objeye karşı tutumlarını ölçmeye yönelik olarak hazırlanmış Likert tipi ölçek ile metrik ölçeğin psikometrik özelliklerini karşılaştırmıştır. Bu çalışmada ölçeğin psikometrik özelliklerinin (madde ayırıcılık gücü, bireylerin tutum ölçüleri, güvenilirlik ve geçerlik) ölçekler bakımından benzerlikler gösterdiği bulunmuştur. Likert tipi tutum ölçeği ile metrik tutum ölçeği arasında psikometrik özelliklerin ortaya konulması bakımından benzerlikler olduğu belirlenmiştir.

Anıl (2002) deneme uygulamasının yapılmadığı durumlarda uzman tahminlerine göre madde özellikleri tahmin edilerek KTK ve ÖÖK'nın iki

parametrelili logistik modeline dayalı hesaplanan madde ve test özelliklerini tahmin etme gücünün nasıl olduğunu belirlemeye çalışmıştır. Bu amaçla Milli Eğitim Bakanlığı'nın 1999-2000 öğretim yılı Orta Öğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavı'na giren yansız olarak seçilmiş 9914 öğrencinin matematik alt testine verdiği tepki araştırmada kullanılmıştır. Araştırmadan elde edilen bulgulara göre Orta Öğretim Kurumları Öğrenci Seçme ve Yerleştirme sınavı matematik alt testine ait madde güçlük indekslerinin uzman tahminlerine dayalı kestirimlerinin KTK'na dayalı madde güçlük indekslerini yordadığı, gözlenen değerlerden elde edilen güvenilirlik katsayı değerlerinin uzman değerlerinden hesaplanan güvenilirlik katsayı değerine göre daha yüksek olduğu bulunmuştur. Buna ek olarak, farklı yetenek düzeylerine sahip olan bireyler için uzman tahmin değerlerine dayalı olarak elde edilen parametre değerlerinin ÖÖK'na dayalı olarak elde edilen parametre değerlerinden daha fazla bilgi vermede yeterli olmadığı ortaya konmuştur.

Nartgün (2002) aynı tutumu ölçmeye yönelik Likert tipi ölçek ile metrik ölçeğin (Matematikle İlgili Düşünceler Ölçeği) KTK ve ÖÖK kapsamında değişik yöntemlerle kestirilen madde ve ölçek özelliklerinin karşılaştırılması amacıyla bir araştırma yapmıştır. Bu amaçla Baykul (1990) tarafından geliştirilen Matematikle İlgili Düşünceler Ölçeğinin Likert ve metrik olmak üzere iki farklı formu 1550 ortaöğrenim öğrencisine uygulanmıştır. Araştırmada elde edilen bulgular ışığında Likert tipi ölçek ve metrik ölçekte yer alan maddelerin KTK ve ÖÖK göre kestirilen madde ayırıcılık güçlerinin birbirine benzediği, ölçeklerde yer alan maddelerin ölçtükleri tutum düzeylerinin her iki kurama göre kestirilen değerlerinin birbirleri ile büyük benzerlikler gösterdiği, öğrencilerin ilgili ölçeklerden belirtilen kuramlara göre kestirilen tutum ölçülerinin büyük benzerlik gösterdiği, her iki ölçeğin tutum düzeyleri farklı olduğu düşünülen grupları iyi ayırabildiği ve ölçüt grup geçerliğine sahip olduğu, Matematikle İlgili Düşünceler Ölçeği'nin Likert formundaki maddelerin hem KTK hem de ÖÖK bakımından, metrik formunda yer alan maddelerin ise KTK bakımından değişmezlik özelliklerini, alt örneklemelerin tüm gruptan farklılaştığı ölçüde kaybettikleri ortaya konmuştur.

Demirtaşlı (2002) çalışmasında Raven Standart İlerlemeli Matrisler Testine (RSİMT) ait madde parametrelerini KTK ve ÖÖK modellerini kullanarak kestirmiş ve görgül yaklaşımla karşılaştırmıştır. Araştırmadan elde edilen bulgular ÖÖK ve KTK'na dayalı olarak kestirilen madde güçlük indekslerinin birbirleri ile karşılaştırılabilir olduklarını ve bu tür karşılaştırmalarda özellikle üç parametrelili modelin KTK ile en yüksek ilişki düzeyini sağladığını ortaya koymuştur. Bununla birlikte iki kurama dayalı olarak kestirilen madde ayırıcılık değerleri arasındaki ilişki oldukça düşük değerli ve negatif bulunmuş, bu parametreler her iki kuram ile karşılaştırılabilir düzeyde bulunmamıştır. ÖÖK ve KTK kullanılarak kestirilen madde güçlük ölçüleri örneklemeler arası değişmezlik özelliği göstermiş ancak madde ayırıcılık ölçüleri örneklemeler arasında farklı değerler ortaya koymuştur. Görgül verilere dayalı olarak yürütülen çalışmanın bulguları genel olarak değerlendirildiğinde KTK'nın değişmez madde ölçüleri ortaya koymada başarısız olmadığı bulunmuştur. Diğer yandan, bu araştırmanın sonuçları Raven Standart İlerlemeli Matrisler Testi'nin kullanılabilirliği, bilgisayarlı bireyselleştirilmiş test uygulaması, test eşitlemesi gibi bazı uygulamalarda özellikle iki parametrelili modelin kullanılabilirliği yönünde bulgular ortaya koymuştur.

Doğan ve Tezbaşaran (2003) ÖÖK'nin KTK karşısında üstünlüğü olarak kabul edilen madde parametrelerinin değişmezliğini sağlayıp sağlamadığını incelemiştir. Araştırma kapsamında Milli Eğitim Bakanlığı Orta Öğretim Kurumları Öğrenci Seçme ve Yerleştirme Sınavının 2001 yılı uygulamasına katılan 553281 kişinin yer aldığı sınavın "akademik yeteneği ölçme ağırlıklı Türkçe, Matematik, Fen ve Sosyal bilgiler" alanlarını kapsayan dört seçenekli yüz soru araştırma verisini oluşturmuştur. Verilerin analizi üç aşamada yapılmıştır. Birinci aşamada evren ve örneklemeler için betimsel istatistikler hesaplanmıştır. Çarpık örneklemelerde yapılacak karşılaştırmaların isabetliliğini artırmak için çarpıklık mutlak değerleri ile basıklık değerleri eşit tutulmuştur. Basık dağılımda ise çarpıklık katsayısı 0 olarak sabitlenmiştir. Araştırmanın ikinci aşamasında evrendeki verilerin ÖÖK ile çalışmaya uygun olup olmadığı test edilmiştir. Bu amaçla tek boyutluluk, yerel bağımsızlık ve



dağılımın normalliği incelenmiştir. Üçüncü aşamada ise madde istatistikleri ve parametreleri hesaplanmıştır. Araştırma problemine bağlı olarak madde parametreleri, madde istatistikleri ve evren değerleri arasındaki ilişki için Spearman'ın sıra farkları korelasyonu kullanılmış, iki korelasyon farkı Fisher z testi ile test edilmiştir.

Araştırma sonuçlarına göre KTK'na dayalı olarak elde edilen p değerlerinin örneklem sayısı arttıkça farklılık gösterdiği özellikle dağılımın çarpıklığından daha fazla etkilendiği, KTK'na göre elde edilen r<sub>yx</sub> değerlerinin örneklem değişikliklerine bağlı olarak değiştiği ve dağılımın çarpıklığından etkilendiği ortaya konmuş, genel olarak KTK ve ÖÖK için elde edilen istatistikler ve parametrelerin örneklem değişimlerinden etkilendiği bulunmuştur. Parametre değişmezliği açısından kuramların herhangi birisinin avantaj sağlamadığı, KTK'nın değişmezlik ölçütlerinin karşılamasına rağmen ÖÖK'nın a parametreleri açısından değişmezlik ölçütlerini sağlayamadığı bulunmuştur.

Tezbaşaran ve Kelecioğlu (2004) madde-ölçek korelasyonlarına, alt-üst grup ortalamalarına ve aşamalı tepki modeline göre geliştirilen sigaraya ilişkin tutum ölçeğinin madde ve ölçek özelliklerini incelemişlerdir. Araştırma kapsamında sigaraya ilişkin tutumu ölçen 41 maddeli beşli Likert tipi ölçeğin madde ve ölçek özellikleri üç ayrı teknikte incelenmiştir. 41 maddelik deneme ölçeğinden madde-ölçek korelasyonlarına, alt-üst grup ortalamalarına ve ÖÖK'na dayalı aşamalı tepki modeli (graded response model) kullanılarak kestirilen madde ayırıcılık güçlerine göre eşit sayıda maddeden oluşan üç ayrı ölçek geliştirilmiştir. Araştırma sonucunda her üç teknikte seçilen maddelerin %94'ü ortak madde olarak bulunmuştur. Deneme ölçeğinin ve seçilen ölçeğin maddelerden ortak olanların a, r, ve t değerleri arasındaki sıra farkları korelasyon katsayıları 0.52 ile 0.88 arasında değiştiği bulunmuştur. Her üç teknikte elde edilen ölçeklerin Cronbach Alfa güvenilirlik katsayıları ve marjinal güvenilirlik katsayıları birbirine eşit olarak bulunmuştur (0.92 ve 0.95). Deneme ölçeğinin ve üç tekniğe göre seçilen ölçeklerin test bilgi fonksiyonları

incelenmiş ve ATM'nin tutum düzeylerine göre verdiği bilgi diğer iki teknikle oluşturulan ölçeklerden daha yüksek bulunmuştur.

Kan (2006) KTK'na ve ÖÖK'na göre kestirilen madde parametrelerini karşılaştırmak amacı ile yaptığı çalışmasında OKS (Orta Öğrenim Kurumları Sınavı) Türkçe alt testine ait maddeleri KTK ve ÖÖK'na göre analiz etmiştir. Araştırmada temel olarak KTK ve ÖÖK'na göre kestirilen madde istatistikleri arasındaki ilişkinin incelenmesi amaçlanmıştır. Araştırmada veri toplama aracı olarak OKS Türkçe alt testi kullanılmış ve araştırma bu sınavı alan 553108 öğrenci üzerinde yapılmıştır. Toplanan verilerin öncelikle ÖÖK'nın varsayımlarını karşılayıp karşılamadığı test edilmiş ve her iki kurama göre madde istatistiklerinin elde edilmesi için BILOG ve ITEMAN programları kullanılmıştır. Her iki kurama göre elde edilen madde istatistikleri arasındaki ilişkinin incelenmesi amacıyla Spearman Brown sıra farkları korelasyon katsayısı kullanılmıştır. Araştırma sonucunda her iki kurama göre kestirilen madde parametreleri ve istatistiklerinin oldukça benzer oldukları bulunmuştur.

Çetin ve Kelecioğlu (2007) çalışmalarında seçmeli testlerin alternatif bir uygulama yöntemi olan eleme puanlamasının psikometrik özelliklerinin geleneksel uygulama ile KTK ve ÖÖK bağlamında karşılaştırılmasını amaçlamışlardır. Araştırmanın örneklemini Hacettepe Üniversitesi'nin çeşitli fakültelerinde İngilizce hazırlık öğrenimi gören 370 öğrenci oluşturmuştur. Öğrencilerin, yarısına öncelikle eleme puanlaması diğer yarısına geleneksel uygulamadan daha sonra eleme puanlaması yaptırılarak veriler elde edilmiştir. KTK'na dayalı olarak, testin uzunluk verimliliği, güvenilirliği ve geçerliği karşılaştırılmıştır. ÖÖK'na dayalı geleneksel puanlama ile eleme puanlamasının göreceli verimlilikleri, marjinal güvenilirlikleri ve geçerlikleri karşılaştırılmıştır. Araştırma sonuçlarına göre KTK'na dayalı karşılaştırmalarda, geleneksel puanlama ile karşılaştırılmasında, eleme puanlamasının uzunluk olarak 1.88 kat daha verimli olduğu bulunmuştur. Eleme puanlaması (0.84) güvenilirlik yönünden geleneksel puanlamadan (0.73) daha yüksek bulunmuştur. Geleneksel puanlama ile eleme puanlamasının uyum geçerlikleri arasındaki fark anlamlı bulunmamıştır.

ÖÖK'na dayalı olarak, tüm yetenek düzeylerinde eleme puanlamasının geleneksel puanlamaya göre daha verimli olduğu bulunmuştur (1.42 kattan, 19.44 kata kadar). ÖÖK'na dayalı olarak elde edilen marjinal güvenilirlikler yönünden eleme puanlamasının geleneksel puanlamadan daha yüksek katsayılar verdiği gözlenmiştir. ÖÖK'na dayalı olarak eleme puanlamasının geleneksel puanlamadan daha yüksek geçerlik değerlerine sahip olduğu belirlenmiştir.

Gözen Çıtak (2007) KTK ve ÖÖK'na göre çoktan seçmeli testlerde farklı puanlama yöntemlerini karşılaştırdığı çalışmasında, çoktan seçmeli bir testte yer alan madde seçeneklerinin iki kategorili (1 – 0 ) puanlama, uzman yargısına dayalı seçenek ağırlıklandırma yöntemleri ile puanlaması durumunda test güvenirliliğinin ve geçerliğinin belirlenmesini amaçlamıştır. Çalışmada çoktan seçmeli türde 18 maddeden oluşan sözel yetenek testi 1593 üniversite öğrencisine uygulanmış, ÖÖK'nda 1 – 0 puanlamanın kullanıldığı durumda, yetenek ölçeği üzerindeki parametrelerin ağırlıklı puanlamanın kullanıldığı duruma göre daha doğru kestirdiği ve bu puanlama yönteminin test geçerliği açısından da daha etkili olduğu sonucuna ulaşılmıştır. KTK'nda ise deneysel ağırlıklandırmanın kullanıldığı durumda güvenirliliğin daha yüksek kestirildiği ancak tüm puanlama yöntemlerinde test geçerliğinin değişmediği gözlenmiştir.

Çelen (2008) KTK ve ÖÖK yöntemleri ile geliştirilen iki testin geçerlik ve güvenirliliğini karşılaştırmıştır. Bu çalışmada araştırmacı, test geliştirme sürecinde deneme uygulaması aşamasından sonra madde seçimini KTK ve ÖÖK doğrultusunda yaparak geliştirilen iki testin psikometrik özelliklerinin karşılaştırılmasını amaçlamıştır. Araştırma kapsamında veriler 6, 7 ve 8. sınıf öğrencilerinden seçilerek oluşturulan çalışma grubundan, araştırmacı tarafından geliştirilen sözel yetenek testi kullanılarak elde edilmiştir. Madde seçiminin KTK ve ÖÖK'na göre yapılması sonucu oluşturulan nihai testlerin uygulanması ile elde edilen sonuçlar karşılaştırılmış ve elde edilen bulguların, ÖÖK'nın farklı yollardan aynı sonuçlara ulaşmanın bir diğer yolu olduğu görüşünü destekler nitelikte bulunmuştur. Elde edilen bulgulara dayanarak

geçerli ve güvenilir bir ölçme aracı geliştirmek için seçilecek kuramın değil, seçilen kuramın gereklerinin yerine getirilmesinin önemli olduğu genel sonucuna ulaşmıştır.

### **Yurt Dışında Yapılan Çalışmalar**

Douglass, Kahavari, ve Farber (1979) KTK ve ÖÖK'nın Rasch modelinin karşılaştırıldığı bir çalışma yapmışlardır. Bu çalışmada 373 denek MacAndrew alkolizm ölçeği ve alkol kullanım sıklığı ve miktarını belirleyen Khavari Alkol Testini almışlardır. MacAndrew testine verilen tepki hem KTK hem de Rasch modeli ile analize tabi tutulmuştur. Analiz sonucunda bazı maddelerin Rasch modeli ile bazılarının ise KTK ile uyum gösterdiği sonucuna varılmıştır.

Goldman ve Raju (1986) bir ve iki parametrelili modelin bir tutum ölçeğinin ölçümleme (kalibrasyon) için uygun olup olmadığını araştırmıştır. Bu çalışmada ayrıca örneklem büyüklüğünün madde ve yetenek parametrelerinin kestirilmesine etkisi olup olmadığı araştırılmıştır. Üçbin deneğin SRA tutum ölçeğine verdiği tepki ÖÖK'nın bir ve iki parametrelili modelleri kullanılarak analiz edilmiştir. Her iki model kullanılarak analiz yapılan her örneklem boyutu için (n=1000, 500 ve 250) yetenek kestirimleri evren için yapılan yetenek kestirimleri ile yüksek ilişki değerleri ortaya koymuştur. Madde güçlük değerleri için denek sayısının 250 olması bir parametrelili model ile kestirimler yapabilmek için yeterli olduğu bulunmuştur. Denek sayısının 500 olduğu durumlar için ise madde parametrelerinin iki parametrelili modelin uygun olduğu sonucuna varılmıştır. Madde ayırt edicilik kestirimlerinin iki parametrelili model ile yapılabilmesi için örneklem sayısının en az bin olması gerektiği bulunmuştur.

Cole, Wood ve Dunn (1991) yaptıkları çalışmada, KTK'nı ve ÖÖK'nı psikomotor beceri verileri bağlamında karşılaştırmışlardır. Çalışmada Motor Gelişim Testi ile elde edilen veriler kullanılmıştır. Çalışma kapsamında Motor Gelişim Testi'nden iki kategorili olarak (dikotomik) elde edilen verilere ÖÖK'nın bir parametrelili, iki parametrelili ve üç parametrelili lojistik

modellerinden hangisinin uyumlu olduđu arařtırılmıřtır. alıřmanın bulguları, KTK ve ÖÖK'nın madde ve testlere iliřkin kestirilen madde güçlük deęerleri ve madde ayırtedicilik gücü deęerleri baęlamında paralel sonuçlar ortaya koymuřtur. Buna ek olarak, ÖÖK'nın yetenek kestirimlerinde daha fazla açıklayıcı bilgiler ortaya koyduđu bulunmuřtur. İki parametrelili ÖÖK modelinin Motor Geliřim Testi'nden elde edilen verilerin analiz edilmesi için en yüksek model veri uyum deęeri verdięi bulunmuřtur. Sonuç olarak ÖÖK'nın uygulamalı beden eęitiminin zorlayıcı ölçme taleplerini karřılayabilecek bir yapıda bulgular ortaya koyduđu bulunmuřtur.

Stage (1997) yaptıęı alıřmada SweSAT alt testi olan READ bölümünün madde parametrelerini hem ön deneme formunda hem de nihai formda KTK ve ÖÖK çerçevesinde karřılařtırmıřtır. Bu amaçla arařtırmacı 16 maddelik READ bölümünün 1996 yılında uygulanmıř öntesti ve 1997 yılında uygulanan nihai test maddelerine iliřkin  $p$  ve  $r_{(bis)}$  deęerlerini KTK ve  $a$ ,  $b$ , ve  $c$  parametrelerini de ÖÖK dahilinde elde ederek birbirleri ile karřılařtırmıřtır. Bu alıřma sonucunda her iki kuram çerçevesinde yapılan madde analizlerinin tatmin edici olduđu sonucuna ulařılmıřtır. Ön deneme uygulaması ve nihai formdan KTK ile elde edilen madde güçlük deęerleri arasında 0.78, ÖÖK ile elde edilen madde güçlük deęerleri arasında 0.55 düzeyinde anlamlı bir iliřki bulunmuřtur. Benzer řekilde ön deneme uygulaması ve nihai formdan KTK ile elde edilen madde ayırtedicilik deęerleri arasında 0.66 ve ÖÖK ile elde edilen madde ayırtedicilik deęerleri arasında 0.54 düzeyinde anlamlı bir iliřki bulunmuřtur. Öntest ile elde edilen madde parametrelerinin KTK aısından daha tutarlı sonuçlar ortaya koyduđu bulgusuna ulařılmıřtır. Öntest ile nihai uygulama arasında çok az maddenin madde parametrelerinde deęiřim olduđu ve bu deęiřime ÖÖK'nın daha hassas olduđu bulgusuna da ulařılmıřtır. Genel olarak alıřma bulgularının her iki kuramın madde parametrelerinin incelenmesi sürecinde benzer sonuçlar ortaya koyduđu bulgusuna ulařılmıřtır. alıřma sonucunda görülen farklılıkların ön denemeye katılan cevaplayıcı sayılarının nihai formu cevaplayanların sayılarından yüksek olmasından kaynaklanabileceęi yorumu getirilmiřtir.

Fan (1998) çalışmasında, KTK'na göre geliştirilen kırksekiz maddelik sözel ve altmış maddelik matematik alt boyutları olan Teksas Akademik Beceri Değerlendirme Ölçeği'ne (TASS; Texas Assesment of Academic Skills) yüzdoksanüçbin ikiyüzkırk dördüncü sınıf öğrencisi tarafından verilen yanıtları, beş ana örneklemden seçilen biner kişilik yirmi örneklem grubundan elde edilen verilerle KTK ve ÖÖK'na göre analiz ederek madde istatistiklerini karşılaştırmıştır. Bu çalışmada, KTK'na göre kestirilen madde ayırıcılık gücü ve madde güçlük indeksleri ile ÖÖK'na göre kestirilen a ve b parametreleri arasında yüksek korelasyonlar bulunmuştur. Buna ek olarak, ham puanlar ile yetenek puanlarının da yüksek düzeyde ilişkili olduğu ve KTK ve ÖÖK'na göre kestirilen madde ve ölçek parametrelerinin birbirleri ile tutarlı bulgular ortaya koyduğu sonucuna ulaşılmıştır. Madde güçlük indekslerinin farklı örneklem gruplarında değişmezlik özelliklerini koruduğu, ancak madde ayırıcılık gücü değerlerinin ise değişmezlik özelliklerini farklı örneklem grupları için koruyamadığı sonucuna da ulaşılmıştır.

Subramaniam ve Silverman (2000), Beden Eğitime yönelik tutumların ölçülmesi için bir tutum ölçeği geliştirme çalışması yapmışlardır. Bu çalışmanın amacı Beden Eğitime yönelik tutumları ölçen bir tutum ölçeği geliştirmek ve geçerlik ve güvenirlik çalışmalarına ilişkin psikometrik verileri incelemektir. Bu çalışma temel olarak dört aşamada gerçekleştirilmiştir. Birinci aşamada Beden Eğitime yönelik olarak orta dereceli okullarda okuyan öğrencilerin bu derse yönelik olarak ortaya koydukları tutum ifadeleri oluşturulmuş ve bu ifadeler alandaki uzmanlar tarafından değerlendirilerek başlangıçta 30 maddelik bir madde havuzu ile çalışmaya başlanmıştır. Bu çalışma esnasında tutumun işlevsel tanımı yapılmış ve iki boyuttan oluştuğu belirtilmiştir. Bu boyutlar bilişsel ve duyuşsal olarak belirtilmiştir. Belirlenen 30 maddenin onikisi negatif maddeler olarak oluşturulmuştur. Birinci aşamada toplam 110 öğrenci üzerinde çalışma yapılmıştır. İkinci aşamada belirlenen 30 madde 33 öğrenci üzerinde 5'li Likert ölçeği şeklinde uygulanmış ve maddelerin psikometrik özellikleri değerlendirilerek nihai olarak 20 maddenin Beden Eğitime yönelik tutumların ölçülmesinde kullanılabileceği bulgusuna ulaşılmıştır. Üçüncü aşamada belirlenen ölçeğin yapı geçerliği çalışması

yapılmıştır. Bu aşamada yapı geçerliğinin sınanmasında önceden denence ile belirlenen iki faktörlü yapı Açıklayıcı Faktör Analizi yöntemi kullanılmıştır. Son aşamada 20 maddelik nihai tutum ölçeğinin güvenirlik çalışması 6., 7. ve 8. sınıflarda eğitim gören 995 öğrenci üzerinde yapılmıştır. Bu çalışmalar sonucunda ölçeğin belirlenen iki bileşenli yapıda geçerli ve güvenilir bulgular ortaya koyduğu ve denencede belirlenen iki faktörlü yapının gözlenen veri ile uyumlu olduğu sonucuna varılmıştır.

Macdonald ve Paunonen'in (2002) yaptıkları çalışma, Fan'ın (1998) çalışmasının genişletilmiş bir tekrarı niteliğindedir. Bu çalışmada, madde ve denek istatistiklerinin karşılaştırılması, KTK ve ÖÖK kapsamında madde istatistiklerinin farklı örneklem gruplarındaki değişmezliği ve belirli örneklem grupları için KTK ve ÖÖK kapsamında kestirilen madde ve denek istatistiklerinin doğruluğunun tespiti amaçlanmıştır. Çalışma kapsamında bin kişilik denek grubundan Monte Carlo simülasyon yöntemi kullanılarak farklı örneklem grupları oluşturulmuş ve bu gruplar üzerinden KTK ve ÖÖK ile madde ve test istatistikleri elde edilmiştir. Bu çalışmalar sonucunda, KTK ve ÖÖK'nın madde güçlük değerleri bakımından benzer sonuçlar ortaya koyduğu ancak madde ayırteçicilik gücü değerleri bakımından test maddelerinin yapısal özelliklerine bağlı olarak farklılaşma olduğu bulunmuştur. Yetenek kestirimleri bakımından, her iki kuram dahilinde ulaşılan sonuçların doğru olduğu bulunmuştur. Buna ek olarak, test madde güçlük değerleri bakımından her iki kuramında birbirine paralel ve tutarlı sonuçlar ortaya koyduğu, ancak madde ayırteçicilik gücü değerleri bakımından sadece ÖÖK kapsamında tüm örneklem gruplarında yapılan analizlerin doğru sonuçlar ortaya koyduğu ve KTK'nın sadece bazı örneklem gruplarında doğru madde ayırteçicilik gücü değeri kestirimi yaptığı sonucuna ulaşılmıştır.

Wiberg (2004) İsveç sürücü ehliyeti yazılı sınavı üzerinde bir çalışma yürütmüştür. Bu çalışmada araştırmacı ÖÖK'nın bir, iki ve üç parametreliliği logistik modellerinin hangisinin testin maddelerinin analizinde kullanılmasını daha uygun olacağını belirlemeyi ve belirlenen bu model ile elde edilen test

ve madde parametrelerini KTK ile elde edilen madde parametreleri ile karşılaştırmayı amaçlamıştır. Çalışmada çoktan seçmeli sorularından oluşan altmışbeş maddelik bir test kullanılmıştır. Araştırma amacı kapsamında hangi logistik modelin test verilerine daha uygun olduğunun bulunması için testin 2004 yılında 5404 kişiye uygulanması ile elde edilen veriler üzerinden ÖÖK'nın sayıltılarının karşılanması düzeyi araştırılmıştır. Bu değerlendirme sonucunda üç parametrelili logistik modelin test parametrelerinin analizi için en uygun model olduğu sonucuna ulaşılmıştır. Üç parametrelili logistik model ile elde edilen a ve b parametreleri ile KTK ile elde edilen p ve  $r_{(pbis)}$  madde parametreleri karşılaştırılmıştır. Bu çalışma sonucunda her iki kuramın da madde parametreleri açısından benzer bulgular ortaya koyduğu sonucuna ulaşılmıştır.

Lin (2008) yaptığı çalışmada Ağırlıklandırılmış Sapma Modeli kullanarak geliştirilen iki formun paralellik derecesini KTK ve ÖÖK çerçevesinde incelemeyi amaçlamıştır. Altıyüz maddelik ilk madde havuzundan altmış maddelik iki form oluşturulmuştur. Bu çalışmada testlerin psikometrik yapılarının incelenmesinde KTK ve ÖÖK'nın iki modeli kullanılmıştır. Psikometrik yapıların incelenmesinde kullanılan bu üç model test uyarlama uygunluğu, ortalama test tekrarlama oranları, içerik paralelligi ve istatistiksel paralellik açısından karşılaştırılmıştır. Araştırma sonucunda elde edilen bulgular KTK kadar ÖÖK'nın da yeterli sonuçlar ortaya koyduğu bulunmuştur. Paralel formlardaki soru yapılarının orta düzeyde ve üzerinde olduğu durumlarda KTK'nın test paralelligi düzeyinde daha tatmin edici bulgular ortaya koyduğu sonucuna ulaşılmıştır.

Alanyazındaki yurtiçi çalışmalar incelendiğinde genel olarak yapılan araştırmalarda ÖÖK ve KTK'nın ölçek geliştirme çalışmalarında kullanıldıklarında karşılaşılan madde parametrelerinin kestirimi, yapılan çalışmalar dahilinde geliştirilen ya da başta KTK çerçevesinde geliştirilmiş ve uzun yıllar boyunca kullanılan bazı testlerin geçerlik ve güvenilirlikleri her iki kuram çerçevesinde ele alınmıştır. Bu çalışmaların büyük bir bölümünde test geliştirme çalışmalarında ve kullanılan testlerin madde parametrelerinin



kestiriminde hem ÖÖK modellerinin hem de KTK'nın benzer ve tutarlı bulgular ortaya koyduğu görülmektedir. Buna ek olarak geçerlik ve güvenirlik çalışmalarında da bu iki kuramın birbirleri ile tutarlı sonuçlar ortaya koyduğu görülmektedir. Elde edilen bulgular doğrultusunda test geliştirme çalışmalarında KTK ve ÖÖK'nin birlikte ve ayrı ayrı kullanılabilirlikleri, test ya da ölçeklerdeki maddeler ve bireyler hakkında daha fazla bilgi elde etmek isteyen araştırmacılar için ÖÖK'nin KTK'na göre daha çok bilgi sağladığı yönünde bulgulara ulaşılmıştır.

Uluslararası alanyazın incelendiğinde, beden eğitimi'ne yönelik tutumların ölçülmesi için tutum ölçeklerinin KTK ve ÖÖK yöntemlerinin birlikte kullanıldığı kısıtlı sayıdaki çalışmalara karşın bu iki kuramın birlikte kullanıldığı farklı alanlardaki tutum ölçeği geliştirme çalışmalarına rastlanmıştır (Strong, Bren & Lejuez, 2004; Dodeen ve Johanson, 2003; Laidlaw, Power, Schmidt & WHOQOL-OLD, 2007; Waugh, 2007). Söz konusu çalışmalarda da, bu çalışmada olduğu gibi tutum ölçeklerinin KTK ve ÖÖK'nin varsayımlarının aynı anda karşılandığı ve psikometrik özelliklerinin bu iki kuram çerçevesinde incelendiği görülmüştür. Psikometrik performans açısından incelendiğinde geliştirilen ölçeklerin hem KTK hem de ÖÖK kuramı açısından önemli değerler ortaya koyduğu gözlenmiştir. Bu çalışmalar sonucunda KTK ve ÖÖK'nin ölçek geliştirme çalışmalarında birlikte kullanılabilirliği sonucuna varılmıştır. Bu iki kuramın birlikte kullanılmasının özellikle geliştirilen tutum ölçeklerindeki geçerlik ve güvenirlik çalışmalarının daha güçlü bir şekilde yapılacağını, geçerlik ve güvenirlik bakımından daha güvenilir ve güçlü bulgular elde edileceği sonucuna ulaşılmıştır. KTK'nın başlangıçta özellikle karmaşık tutum yapılarının ortaya konulmasında önemli bir kuram olduğu, karmaşık yapıların belirlenmesiyle bundan sonra yapılacak çalışmaların ve psikometrik özelliklerin incelenmesi aşamasında ÖÖK'nin önemli bir rol üstlendiği ortaya konmuştur.

Yurtdışında yapılan çalışmalarda da yurtiçinde yapılan çalışmalara benzer şekilde ÖÖK ve KTK'nın ölçek geliştirme çalışmalarında kullanıldıklarında karşılaşılan madde parametrelerinin kestirimi, çalışma

dahilinde geliştirilen ya da KTK çerçevesinde geliştirilmiş ve halen kullanılan bazı testlerin geçerlik ve güvenirlikleri her iki kuram çerçevesinde ele alınmıştır. Bu çalışmalarda da yurtiçinde yapılan araştırma bulgularını destekleyen sonuçlar elde edilmiştir.

## BÖLÜM III

### YÖNTEM

Bu bölümde araştırmanın türü, çalışma grubu, ölçeklerin geliştirilmesi, verilerin toplanması ve çözümlenmesi açıklanmaktadır.

#### Araştırmanın Modeli

Bu araştırmada KTK ve ÖÖK'na dayalı olarak İlköğretim 6, 7, ve 8. sınıf öğrencilerinin Beden Eğitimi Dersi'ne yönelik tutumlarını ölçen Likert tipi bir tutum ölçeği geliştirilmiş ve geliştirilen ölçeğin psikometrik özellikleri bu iki kuram ile incelenmiştir. Yapılan analizler doğrultusunda beden eğitimi'ne yönelik tutumları ölçen Likert tipi tutum ölçeğinin psikometrik özellikleri ortaya konulmuş, geçerlik ve güvenilirliği her iki kuram çerçevesinde ele alınmıştır. Belirtilen nedenler ile bu çalışma tarama modelindedir.

#### Çalışma Grubu

Bu araştırma çerçevesinde, Ankara ilinde yer alan ilçelerinden (Altındağ, Çankaya, Yenimahalle, Etimesgut, Keçiören ve Mamak) birer ilköğretim okulu seçilmiş ve bu ilköğretim okullarının 6, 7 ve 8. sınıflarından da 926 öğrenci çalışmaya dahil edilmiştir. ÖÖK'nın Aşamalı Tepki Modelinin varsayımlarının karşılanabilmesi ve doğru ve kesin parametre kestirimlerinin yapılabilmesi için en az 500 öğrenciden oluşan bir çalışma grubu oluşturulması hedefi gerçekleştirilmiştir (Goldman ve Raju 1986; Reise ve YU, 1990). Deneme uygulamasından elde edilen cevap kağıtları üzerinde yapılan ön inceleme sonucunda 139 öğrencinin cevap kağıtlarını gelişigüzel olarak işaretlediği görülmüştür. Bu cevap kağıtlarının büyük bir bölümünde tüm tepkilerin aynı kategori altında işaretlendiği ya da maddelerin büyük bir bölümünün işaretlenmeden boş bırakıldığı görülmüştür. ÖÖK'nda madde ve yetenek kestirimi aşamalarında kullanılan "en yüksek olabilirlik" (maximum marginal likelihood) yöntemi mükemmel ya da sıfır puanlarda tutarlı olmayan cevap örüntüsünde kestirimde bulunamadığı için gelişigüzel cevaplanan 139 cevap kağıdı analize dahil edilmemiştir. Sonuç olarak toplam 787 cevap

kağıdı ölçek geliştirme çalışması analizlerine dahil edilmiştir. Analize dahil edilen cevap kağıtlarına ilişkin öğrencilerin okullara ve sınıf düzeylerine göre dağılımı Çizelge 1’de sunulmuştur.

**Çizelge 1. Deneme Formu Uygulamasında Analize Dahil Edilen Cevap Kağıtlarına İlişkin Öğrencilerin Okullara/Sınıf Düzeylerine Göre Dağılımı**

| Okullar                          | Öğrenci Sayısı |          |          |        |
|----------------------------------|----------------|----------|----------|--------|
|                                  | 6. Sınıf       | 7. Sınıf | 8. Sınıf | Toplam |
| Demetevler İlköğretim Okulu      | 46             | 37       | 35       | 118    |
| Ergazi İlköğretim Okulu          | 32             | 29       | 34       | 95     |
| Barışyolu İlköğretim Okulu       | 46             | 41       | 43       | 130    |
| Kazım Karabekir İlköğretim Okulu | 36             | 28       | 32       | 96     |
| Nurçin Sayan İlköğretim Okulu    | 28             | 37       | 31       | 96     |
| Aydınlıkevler İlköğretim Okulu   | 39             | 47       | 36       | 122    |
| Etimesgut İlköğretim Okulu       | 55             | 39       | 36       | 130    |
| Genel Toplam                     | 282            | 258      | 247      | 787    |

Çalışma kapsamında deneme formu 787 öğrenciye uygulanarak KTK ve ÖÖK’nın psikometrik özelliklerinin incelenmesi çalışmaları yürütülmüştür. Bu çalışmalar sonucunda her iki kuram kapsamında belirlenen ortak otuz madde ile oluşturulan asıl ölçeğin madde parametrelerinin kestirilmesi ve incelenmesi çalışmaları nihai ölçeğin 426 (213 kız ve 213 erkek) öğrenciye tekrar uygulanması ile yürütülmüştür. Asıl ölçeğin uygulandığı öğrencilerin okullara ve sınıf düzeylerine göre dağılımı Çizelge 2’de sunulmuştur.

**Çizelge 2. Asıl Ölçeğin Uygulandığı Öğrencilerin Okullara/Sınıf Düzeylerine Göre Dağılımı**

| Okullar                          | Öğrenci Sayısı |          |          |        |
|----------------------------------|----------------|----------|----------|--------|
|                                  | 6. Sınıf       | 7. Sınıf | 8. Sınıf | Toplam |
| Demetevler İlköğretim Okulu      | 36             | 29       | 41       | 106    |
| Ergazi İlköğretim Okulu          | 32             | 25       | 34       | 91     |
| Barışyolu İlköğretim Okulu       | 33             | 52       | 23       | 108    |
| Kazım Karabekir İlköğretim Okulu | 41             | 38       | 42       | 121    |
| Genel Toplam                     | 142            | 144      | 140      | 426    |

## Ölçeğin Geliştirilmesi

Bireylerin Beden Eğitimi dersine yönelik tutumlarını belirlemede kullanılacak tutum ölçeğinin oluşturulmasında ilköğretim altıncı, yedinci ve sekizinci sınıf öğrencilerinden Beden Eğitimi dersine yönelik olarak duygu, düşünce ve davranışlarını ifade eden birer paragraf yazmaları istenmiş ve bu paragraflar içerik analizi yöntemi ile çözümlenerek tutum ifadesi olan cümleler oluşturulmuştur. Ayrıca tutum cümlelerinin oluşturulması aşamasında alanyazında yer alan araştırmalar da incelenmiştir. Bu kapsamda Özer ve Aktop'un (2003) İlköğretim öğrencileri için hazırlanmış bir Beden Eğitimi dersi tutum ölçeğinin adaptasyonu çalışması ve Demirhan ve Altay'ın (2001) lise birinci sınıf öğrencileri için Beden Eğitimi'ne İlişkin Tutum Ölçeği geliştirilmesi çalışmaları incelenmiş ve tutum cümlelerinin oluşturulmasında bu çalışmalardaki tutum cümlelerinden de faydalanılmıştır. Tutum cümlelerinin dil ve anlatım kurallarına, tutum ifadesi yazma tekniklerine uygunluğu ve ifadelerin 12-15 yaş grubu öğrencilerin anlayabilecekleri ölçüde yalın ve anlaşılır nitelikte olup olmadıklarının belirlenmesi için ölçme değerlendirme uzmanları, Beden Eğitimi dersi öğretmenleri ve Beden Eğitimi alanında akademik çalışmalar yürüten Beden Eğitimi Öğretmenliği alanındaki öğretim üyelerinin görüşlerine başvurulmuştur.

Üç bileşen yaklaşımına göre bireylerin tutuma yönelik bilgileri tutumun bilişsel bileşenini, bireylerin duyguları tutumun duyuşsal bileşenini ve bireylerin tutumlarını ortaya koyan fiziksel hareketleri de tutumun edimsel bileşenini oluşturur (Baron ve Byrne, 1997). Üç bileşen yaklaşımında tutum, "çeşitli kurum, kavram, nesne, durum ya da kişilere karşı öğrenilmiş, olumlu veya olumsuz tepkide bulunma eğilimi" (Oppenheim, 2000), veya "diğer kişilere ve nesnelere ilişkin düşüncelere, duygulara ve hareketlere yön veren öğrenilmiş öz eğilimler bütünü" dür şeklinde tanımlanmıştır (Baron ve Byrne, 1997; Middlebrook, 1974:575; Turgut, 1984:154). Beden eğitimi dersine yönelik tutum ölçeğinin deneme formunda yer alacak tutum ifadelerinin seçilmesinde üç bileşen yaklaşımı temel alınmıştır. Bu bağlamda, deneme

formunda tutumun duyuşsal bileşenini içeren 17 madde, bilişsel bileşenini içeren 17 madde ve devimsel bileşenini içeren 16 madde ölçekte yer almıştır.

Yapılan çalışmalar sonucunda toplam elli tutum cümlesi oluşturulmuştur. Bu tutum cümlelerinin 25 tanesi olumlu tutum ifadesinden 25 tanesi de olumsuz tutum ifadesinden oluşmuştur. Tutum cümlelerinin puanlanmasında olumlu tutumlar için oluşturulan cümleler 5 “Tamamen Katılıyorum”, 4 “Katılıyorum”, 3 “Kararsızım”, 2 “Katılmıyorum”, ve 1 “Hiç Katılmıyorum” şeklinde puanlanmıştır. Olumsuz cümlelerde ise tersine puanlama yöntemi uygulanmıştır.

Araştırmanın kapsamında iki farklı kurama göre Beden Eğitimi dersine yönelik bir tutum ölçeği geliştirilmesi ve psikometrik özelliklerinin incelenmesi amaçlandığından deneme uygulaması yapılan tutum ölçeğinin verilerinin çözümlenmesinde farklı teknikler kullanılmıştır.

Likert tipi tutum maddelerinin analiz edilmesinde farklı madde analizi teknikleri kullanılmaktadır. Korelasyona dayalı analiz ve içtutarlılık ölçütüne dayalı (t-test) analiz teknikleri en sık kullanılan madde analizi teknikleridir (Tezbaşaran, 1997).

**Korelasyona Dayalı Analiz:** Bu analiz tekniğinde, maddelerin ayırt edicilik gücü, her bir tutum cümlesi için elde edilen madde puanları ile ölçek puanları arasındaki korelasyon katsayısı ile ifade edilir (madde-ölçek korelasyonu). Madde ve ölçek puanları arasındaki korelasyon katsayılarının hesaplanması için Pearson Momentler Çarpımı Korelasyonu kullanılır ve aşağıda verilen denklemlerle ifade edilir.

$$r_{xy} = \frac{\sum_{i=1}^n X.Y - \frac{(\sum_{i=1}^n X)(\sum_{i=1}^n Y)}{n}}{\sqrt{\sum_{i=1}^n X^2 - \frac{(\sum_{i=1}^n X)^2}{n}} \sqrt{\sum_{i=1}^n Y^2 - \frac{(\sum_{i=1}^n Y)^2}{n}}} \quad (10)$$

Verilen formülle bulunan katsayı -1.00 ile +1.00 arasında bir değer alır ve işareti eksi olan ve sıfır değerine yakın bulunan maddeler belirtilen tutumu ölçmek için yetersiz olarak sınıflandırılır ve ölçeğe dahil edilmez. Katsayı değeri pozitif ve yüksek olan maddeler ise ölçekte kullanılabilecek maddeler olarak yorumlanır ve ölçeğe dahil edilir. Maddelere yönelik olarak elde edilen korelasyon katsayılarının pozitif ve yüksek olması aynı zamanda ölçeğin geçerlik ve güvenilirliğinin yüksek olmasının da ön koşulu olmaktadır (Tezbaşaran, 1997).

**İçtutarlılık Ölçütüne (t-test) Dayalı Analiz:** İçtutarlılık ölçütüne dayalı analiz tekniğinde her bir bireyin ölçekteki maddelere verdiği tepkiler toplanarak bir toplam puan elde edilir. Bireylerin ölçek puanları belirlendikten sonra bu puanlar büyükten küçüğe doğru sıralanır. Elde edilen bu puan dağılımının üst ve alt %27'lik grupları seçilir. Bu gruplar ölçülmek istenen tutumla ilgili olumlu ve olumsuz yönde olanları belirtmektedir. Daha sonra üst ve alt gruplardaki bireylerin her bir madde için madde puanları ortalamaları belirlenir ve bu ortalamalar arasındaki farkların anlamlı olup olmadığı denklem (2) de belirtilen t-testi ile belirlenir.

$$t = \frac{\bar{X}_{\ddot{u}} - \bar{X}_a}{\sqrt{\left(\frac{S_{\ddot{u}}^2}{n_{\ddot{u}}}\right) + \left(\frac{S_a^2}{n_a}\right)}} \quad (11)$$

*ü: üst grup, a: alt grup*

*$\bar{X}$  : Grubun madde puanları aritmetik ortalaması*

*$S^2$ : Grubun madde puanları varyansı ve*

*n: Gruptaki cevaplayıcı sayısı*

Bu işlem sonucunda t değeri manidar olan maddeler, olumlu tutum ile olumsuz tutuma sahip olan bireyleri iyi ayırt edebilen maddeler olarak kabul edilir ve ölçeğin nihai formunda kullanılmak üzere seçilir (Tezbaşaran, 1997).

KTK çerçevesinde Likert tipi tutum ölçeği geliştirme çalışmalarında, madde analizi yöntemleri kullanılarak yüksek ayırtedicilik gücüne sahip maddelerin nihai formda yer almak üzere seçilmesi gerekir. Bu çalışmada, korelasyona dayalı analiz ve iç tutarlılık ölçütüne dayalı (t-test) madde analiz teknikleri kullanılmıştır.

KTK ve ÖÖK'nda verilerin analizi yapılmadan önce elde edilen verilerin kuramın temel varsayımlarını yerine getirip getirmediği ve verilerin analiz için seçilen modelle uyumlu olup olmadığının incelenmesi gerekmektedir.

KTK'da ölçek kapsamında olan maddelerin ilgilenilen tutumu ölçmesi gerekir. Başka bir deyişle, doğrusal olarak bir araya getirilen maddelerin tek bir faktörle ilişkili olduğu ve bu maddelerin toplamının maddelerle kapsanan bilginin tamamını içermesidir (Tezbaşaran, 1997:9). Buna ek olarak, bu kuramda ölçme hatalarının tesadüfi olduğu ve bu hataların gerçek puanlarla, başka bir değişkenin gerçek puanıyla ve farklı ölçmelere ait hata puanlarıyla korelasyon vermediği varsayılmaktadır (Lord ve Novick, 1968).

Örtük özellikler kuramında ise üç temel varsayımı vardır. Bunlar; normal dağılım, tek boyutluluk ve yerel bağımsızlıktır.

**Normal dağılım;** ölçülen özelliğin ölçmenin yapıldığı grupta  $\mu= 0$ ,  $\sigma=1$  olan normal dağılım gösterdiği varsayımdır. Test maddelerinden elde edilen verilerin normal dağılım özellikleri gösterip göstermediği  $\chi^2$  istatistiği ile test edilebilir (Hambleton, Swaminathan ve Rogers, 1991). Ayrıca normal dağılım özelliğinin test edilmesinde Kolmogorov-Smirnow uyum iyiliği istatistiksel yöntemi de kullanılır. Buna ek olarak, ölçme aracından elde edilen verilerin aritmetik ortalama, mod ve ortancasının birbirine yakın değerler göstermesi ya da eşit olması da puan dağılımının simetrik olduğu hakkında bilgi verir. Ölçek puanları dağılımının çarpıklık ve basıklık katsayılarının sıfıra eşit olması dağılımın normalliği konusunda bilgi veren parametrelerdir.



**Tek boyutluluk;** örtük özellikler kuramının bütün modellerinde testle ölçülen özelliğin tek boyutlu olması istenir. Bir ölçme aracının tek boyutlu olup olmadığının belirlenmesi için farklı yöntemler kullanılabilir. Bu yöntemlerden temel olanı faktör analizidir (Hambleton ve Swaminathan 1985: 156).

Hambleton ve Swaminathan (1985:21) tek boyutluluğun yoklanmasında kullanılan faktör analizi sonuçlarında maddelerin bir boyut altında toplandığı görüldüğünde testin tek boyutlu olarak kabul edilebileceğini belirtmiştir. Benzer şekilde Lord (1980:21) tek boyutluluğun belirlenmesinde, faktör analizinde birinci faktör yüküne ait öz değerinin (eigen value), ikinci faktör yüküne ait öz değerden çok farklı olmasının ve ikinci faktör yüküne ait öz değer diğerlerinden çok farklı olmamasının tek boyutluluğa ölçü sayılabileceğini savunmuştur. Buna ek olarak ölçek maddeleri arasındaki yüksek korelasyon bulunmasının ölçeğin tek boyutluluğuna dair ipucu verdiğine yönünde görüşler de bulunmaktadır (Crocker ve Algina, 1986:343).

**Yerel bağımsızlık;** maddelerin birbirinden bağımsız olarak cevaplandırılmasıdır. Başka bir deyişle, maddelerden birine verilen doğru veya yanlış cevabın başka maddeye verilecek cevabı etkilememesidir (Gelbal, 1994:33). İki maddenin birlikte doğru (ya da yanlış, doğru-yanlış veya yanlış-doğru) cevaplandırılma olasılığı, maddenin ayrı ayrı doğru (ya da yanlış, doğru-yanlış veya yanlış-doğru) cevaplandırılma olasılıklarının çarpımına eşitse maddelerin yerel bağımsızlığa sahip olduğu savunabilir.

Lord (1980) tek boyutluluk varsayımının karşılanması durumunda belli bir yetenek düzeyindeki bireylerin maddelere verdiği tepkiler arasındaki korelasyonun sıfır olduğunu ve bu nedenle tek boyutlu olan ölçeğin aynı zamanda yerel bağımsızlık varsayımını da karşıladığını belirtmiştir.

Hambleton ve Swaminathan (1985) yerel bağımsızlığın tek boyutlulukla ilişkili olması nedeniyle faktör analizi tekniği ile de maddelerin yerel bağımsızlığa sahip olup olmadığının test edilebileceğini savunmuştur.

Yerel bağımsızlık sayıltısı belirli yetenek seviyesindeki bireylerin maddelere verdikleri tepkinin istatistiksel bağımsızlığı anlamına geldiğinden bireylerin test performansını açıklayacak tek bir yeteneğin olması gerekmektedir. Tek boyutluluk varsayımının karşılanması yerel bağımsızlık varsayımının da karşılandığının göstergesi olarak kabul edilmektedir (Hambleton ve Swaminathan, 1985). Bu nedenle tek boyutluluğun ortaya konmasında kullanılacak olan faktör analizi tekniği ile elde edilecek sonuçlar yerel bağımsızlık varsayımının karşılanıp karşılanmadığı konusunda da bilgi verecektir.

**Verileri model uyumu;** örtük özellikler kuramsal yaklaşımda test puanlarından yararlanılarak bireylerin yetenek ölçümleri çeşitli matematiksel fonksiyonlar yardımıyla kestirilmeye çalışılır. Kullanılan matematiksel fonksiyonlardan biri lojistik fonksiyonlardır. Lojistik Regresyon analizinde açıklayıcı değişkenlerin  $\beta$  katsayılarını çeşitli teknikler kullanılarak kestirilmeye çalışılır. Kestirilen  $\beta$  değerleri ile kurulan modelin verilerle uyumlu olup olmadığının incelenmesi gerekmektedir. Model ile verilerin uyumlu olmamasına; uygun olmayan açıklayıcı değişkenler, veriler içerisindeki aykırı değerler, ve verilerin toplandığı ölçeklerin yanlış olması neden olabilir. Modelin geçerliliğini test etmek amacıyla uyum testi yapılmalıdır (Tatlıdil, 1992:231).

Verilerin modelle olan uyumunu test etmek amacıyla  $\chi^2$ ,  $-2 \log \lambda$  - istatistikleri kullanılabilir. Bunların yanı sıra varyans analizi tekniğinin kullanılabilmesine dair görüşler bulunmaktadır (Hambleton ve Swaminathan 1985).

KTK ve ÖÖK çerçevesinde ölçeğin geliştirilebilmesi için gerekli olan bu kuramlara ait varsayımların karşılanıp karşılanmadığı test edilmiştir. Aşağıda bu çalışmalarla ilgili analiz sonuçları değerlendirilmiştir.

## Normal Dağılım

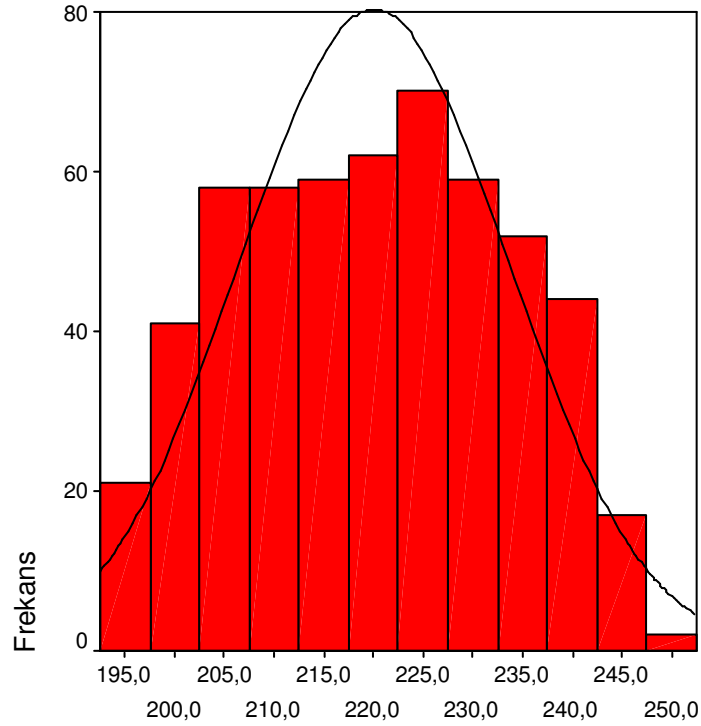
Örnekleme grubu üzerinde ölçülen özelliklerin evrende normal dağılım göstermesi ÖÖK kadar (Hambleton, Swaminathan ve Rogers; 1991) KTK için de önemli varsayımlardan biridir (Crocker ve Algina, 1986). Bu bakımdan normal dağılım varsayımının karşılanması ölçeğin psikometrik özelliklerinin kestirilmesinde her iki kuram için verilerin sağlıklı olarak değerlendirilebilmesi açısından önemlidir. Verilerin normal dağılım gösterip göstermediğine ilişkin analizlerde betimsel istatistikler ve Kolmogorov-Smirnow Uyum İstatistiği kullanılmıştır. Elde edilen bulgular Çizelge 3’de verilmiştir.

**Çizelge 3. Beden Eğitimi ve Spor Dersine Yönelik Tutum Ölçeğinin 787 Öğrenciye Uygulanması Sonucu Elde Edilen Puanlara İlişkin Betimsel İstatistikler**

| Betimsel İstatistikler |        |
|------------------------|--------|
| En Küçük Değer         | 58     |
| En Büyük Değer         | 248    |
| Aritmetik Ortalama     | 204.28 |
| Tepe Değer             | 232    |
| Ortanca                | 210    |
| Standart Kayma         | 29.38  |
| Basıklık               | -.99   |
| Kayıklık               | .007   |

Çizelge incelendiğinde ölçekten elde edilen puanların oldukça geniş bir dağılıma sahip oldukları görülmektedir. Puan dağılımlarına ilişkin standart kayma değerinin yüksek oluşu bu açıklamayı doğrulamaktadır. Ölçekten elde edilen puanlara ilişkin basıklık değerinin negatif bir değer göstermesi puanlara ilişkin dağılımın basık olduğunu göstermektedir. Kayıklık katsayısının pozitif değer alması puanlara ilişkin dağılımın hafif sağa kayık olduğunu göstermektedir.

Buna ek olarak ölçekten elde edilen puanların aritmetik ortalama, ortanca ve mod değerlerinin birbirlerine yakın nitelikte bulunmuştur. Ölçekten elde edilen puanlara ilişkin frekans dağılımlarından yararlanılarak sütun grafiği elde edilmiştir. Bu grafik Şekil 1’de verilmiştir.



**Şekil 1.** Ölçek Puanlarının Dağılımı

Ölçekten elde edilen puanların normal dağılım gösterip göstermediğine ilişkin yapılan Kolmogorov-Smirnov Uyum İyiliği istatistiği sonucunda dağılımın 0.05 anlamlılık düzeyinde normal olduğu belirlenmiştir ( $z=.86$ ,  $p>.05$ ).

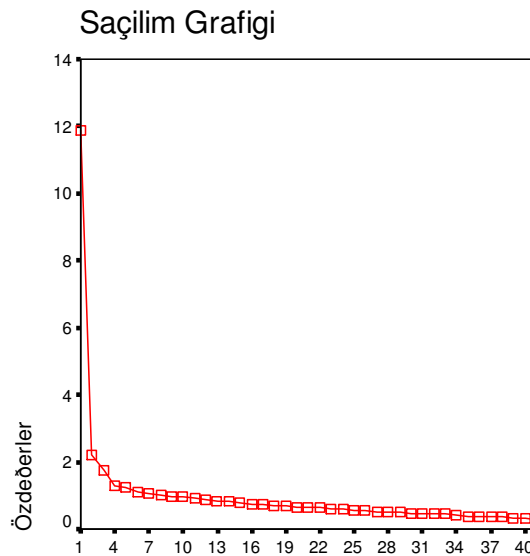
### **Tek Boyutluluk**

Tek boyutluluğun ölçülmesinde kullanılan farklı yöntemler bulunmaktadır. Bu yöntemlerden ilki maddeler arası  $\rho_{hi}$  veya tetrakorik korelasyon matrisine dayalı özdeğerler grafiği kullanımıdır. İkinci yöntem, test verilerine dayalı olarak maddeler arasındaki korelasyon matrisi ile elde edilen öz değerler grafiği ile test edilmesiyle eşit sayıda değişken ve örneklem büyüklüğüne sahip tesadüfi verilere dayalı özdeğerler grafiğinin karşılaştırılmasıdır. Diğer bir yöntem ise, doğrusal olmayan faktör analizi teknikleri ve kalanlar (residual) analizlerinden yararlanmaktır. Kullanılan başka bir teknikte yetenek düzeyinde farklı olan bireyler için elde edilen maddeler arasındaki varyans, kovaryans ve korelasyon matrislerinin

karşılaştırması teknikleri kullanarak tek boyutluluk ve yerel bağımsızlık varsayımları test edilebilmektedir (Hambleton ve Swaminathan, 1985:21; Hambleton ve Jones, 2003).

Beden Eğitimi dersine yönelik olarak öğrencilerin tutumunu ölçmek amacıyla hazırlanan ölçeğin tek boyutlu olup olmadığını belirlemek amacıyla ön denemeden elde edilen veriler kullanılarak yapılan faktör analizi sonucunda faktör yüklerinin beş grupta toplandığı görülmüştür; birinci faktörün öz değeri (eigen value) 14.01, ikinci faktörün özdeğeri 2.79, üçüncününki 1.89 dördüncü faktörün öz değeri 1.61 ve son faktörün öz değeri 1.35'dir. Açıklanan toplam varyans %43,20'dir. Toplam varyansın yüzde 28.02'si birinci faktör, 5.57'si ikinci faktör, 3.78'i üçüncü faktör, 3.22'si dördüncü faktör, 2.69'u ise beşinci faktör tarafından açıklanmaktadır. Faktör analizi sonucunda maddelerin aldıkları özdeğerler Ek-2'de verilmiştir.

Faktör yüklerine ait özdeğerlerde; birinci faktöre ait özdeğer 14.01 iken, ikinci faktörünün özdeğerinin 2.79 düşmesi ve maddelerin büyük bir çoğunluğunun birinci boyutta yüksek yük vermesinden dolayı ölçeğin tek boyutlu olduğu kararına varılmıştır. Yapılan faktör analizine ilişkin saçılım grafiği aşağıda verilmiştir.



**Şekil 2.** Temel Bileşenler Analizine İlişkin Grafik

## **Yerel Bağımsızlık**

Hambleton ve Swaminathan'ın (1985) yerel bağımsızlığın tek boyutlulukla ilişkili olduğu, bu nedenle faktör analizi tekniği ile maddelerin yerel bağımsızlığa sahip olup olmadığının test edilebileceği görüşü doğrultusunda faktör analizi sonuçlarından yararlanarak maddelerin yerel bağımsızlığa sahip olup olmadığı değerlendirilmiştir. Örtük Özellikler kuramına göre ölçeği ya da testi cevaplayan grubun farklı test maddelerine verdiği cevaplar arasındaki ilişkilerin kaynağı maddeye ilişkin performans etkileyen yetenekler olarak tanımlanmaktadır. Yetenek sabit olarak düşünüldüğünde, bireylerin maddelere verdiği yanıtların da bağımsız olması beklenir. Buna ek olarak faktör analizinde, iki değişken arasında ilişki olduğunda, bu iki değişken arasında ortak bir özelliğin olduğu anlamına gelir. Bu iki değişkenin arasında ilişki olmasının nedeni olan faktörün sabitlenmesi ve buna bağlı olarak bu iki değişken arasındaki ilişkinin ortadan kalkması faktör analizinin temelini oluşturur.

Ön deneme sonucu elde edilen verilere uygulanan faktör analizi sonucunda maddelerin büyük bir çoğunluğunun bir boyut altında toplandığı görülmüştür. Diğer bir deyişle ölçekteki maddeler arasında istatistiksel bir ilişki söz konusudur, bu durumda belirli tutum seviyesindeki bireylerin maddelere verdikleri tepkinin istatistiksel olarak bağımsız olması beklenir. Bu sonuçtan yerel bağımsızlık sayılıısının yerine getirildiğini düşünülmüştür.

Ön denemede kullanılan elli tutum cümlesinden elde edilen verilerden örtük özellikler kuramının gerektirdiği varsayımları yerine getiren tutum cümleleri belirlendikten sonra bu cümlelerin modelle olan uyumu incelenmiştir.

## **Veri-Model Uyumu**

Madde ve tutum düzeyi kestiriminde kullanılan model ile verilerin uyum iyiliğinin test edilmesi gerekmektedir. Lojistik modelin uyum iyiliğini test edilmesinde tüm değişkenleri içeren model ile kestirilen modele ilişkin

olabilirlik oran değerlerinin farkına dayanan  $-2 \log \lambda$  istatistiği kullanılabilir.  $-2 \log \lambda$ 'nın kay-kare dağılımı gösterdiği kabul edilmektedir (Tatlıdil, 1992:231; Hambleton ve Swaminathan, 1985). Bu çalışmada tutum ölçekleri gibi çok kategorili (polinomik) ölçme araçlarından elde edilen verilerin analizinde kullanılan MULTILOG 7.0 analiz programı da model veri uyumunun bir göstergesi olan  $-2 \log \lambda$  istatistiğinden faydalanmaktadır.

Tek boyut altında yüksek faktör yükü veren tutum cümlelerinden elde edilen verilerin aşamalı tepki modeliyle uyumunu gösteren  $-2 \log \lambda$  değeri MULTILOG 7.0 da yapılan analiz sonucu 53721.20 olarak elde edilmiştir. Bulunan  $-2 \log \lambda$  değeri anlamlıdır. Bu sonuçtan modele giren değişkenlerin her birinin açıklayıcı değişken olduğu sonucuna varılmıştır.

### **Ölçeğin Klasik Test Kuramıyla Geliştirilmesi**

Elli maddelik deneme uygulamasından elde edilen verilere tek boyutluluk sayıtlısının sınanması aşamasında uygulanan faktör analizi sonucunda elli tutum cümlesinden 42'sinin birinci boyut altında yüksek faktör yüküne sahip olduğu görülmüştür. Maddelerin faktör yükleri Ek-2 'de verilmiştir. KTK'na dayalı Likert tipi ölçek geliştirme çalışmasının ilk aşamasında ön deneme sonucunda elde edilen verilerin madde-ölçek korelasyon değerleri Pearson Momentler Çarpımı Korelasyonu ile hesaplanmıştır. Elde edilen korelasyon katsayıları incelenmiş ve korelasyon katsayıları 0.40'ın altında olan ve işareti eksi olan maddeler çıkarılmıştır. Bu analizler sonunda işareti eksi olan bir madde (madde16) ve 0.40'ın altında korelasyon katsayı değerine sahip maddeler (madde 14, 19, 20, 23, 42, ve 46 ve 50) çıkarılarak korelasyona dayalı analiz sonucu 42 maddeye ulaşılmıştır. Deneme uygulaması sonucu elde edilen madde-ölçek korelasyon değerleri Çizelge 4'te verilmiştir. Madde ölçek korelasyon değerlerinin ortalaması 0.36 olarak bulunmuştur.

**Çizelge 4. Elli Maddelik Beden Eğitimi ve Spor Dersi Tutum Ölçeğinin Ön Deneme Formunun Uygulanması Sonucu Elde Edilen Maddelerin Ölçek Puanı ile Olan Korelasyonları**

| Madde No | r     | Madde No | r    |
|----------|-------|----------|------|
| 1        | 0,56  | 26       | 0,52 |
| 2        | 0,53  | 27       | 0,59 |
| 3        | 0,59  | 28       | 0,47 |
| 4        | 0,48  | 29       | 0,65 |
| 5        | 0,54  | 30       | 0,64 |
| 6        | 0,48  | 31       | 0,55 |
| 7        | 0,50  | 32       | 0,55 |
| 8        | 0,47  | 33       | 0,59 |
| 9        | 0,55  | 34       | 0,53 |
| 10       | 0,41  | 35       | 0,58 |
| 11       | 0,55  | 36       | 0,67 |
| 12       | 0,49  | 37       | 0,48 |
| 13       | 0,49  | 38       | 0,53 |
| 14       | 0,34  | 39       | 0,54 |
| 15       | 0,47  | 40       | 0,52 |
| 16       | -0,22 | 41       | 0,57 |
| 17       | 0,41  | 42       | 0,39 |
| 18       | 0,57  | 43       | 0,64 |
| 19       | 0,36  | 44       | 0,61 |
| 20       | 0,20  | 45       | 0,46 |
| 21       | 0,41  | 46       | 0,39 |
| 22       | 0,57  | 47       | 0,62 |
| 23       | 0,36  | 48       | 0,59 |
| 24       | 0,61  | 49       | 0,63 |
| 25       | 0,63  | 50       | 0,31 |

\*\*P<.01

KTk'ında maddelerin analiz edilmesinde kullanılan diğer yöntem olan iç tutarlılık ölçütüne dayalı analiz dahilinde birinci boyut altında yüksek faktör yükü veren maddelerden elde edilen puanlar %27'lik alt ve üst gruplara ayrılarak madde analizi yapılmıştır. Bu analiz dahilinde üst %27'lik dilimde bulunan grubun (N=218) ölçekten elde ettikleri puan ortalamaları  $233.83 \pm 6.01$  ve alt %27'lik dilimdeki grubun (N=218) ölçek puanları ortalamaları  $165.94 \pm 23.96$  olarak bulunmuştur. Madde analizi sırasında alt ve üst puan gruplarında bulunan bireylerin cümlelerdeki kategorilere verdikleri tepkiden yararlanarak "t" istatistiği hesaplanmıştır. Madde analizine katılan cümlelerin "t" istatistiği değerleri Çizelge 5'te verilmiştir.



**Çizelge 5. Ön Denemede Kullanılan Tutum Cümlelerinin Madde Analizi  
Sonucu Elde Edilen "t" İstatistikleri ve Eta Kare Değerleri**

| Madde No |     | x    | ss   | t       | $\eta^2$ | Madde No |     | x    | ss   | t       | $\eta^2$ |
|----------|-----|------|------|---------|----------|----------|-----|------|------|---------|----------|
| 1        | Üst | 4.93 | 0.27 | 14.50** | 0.412    | 26       | Üst | 4.69 | 0.68 | 13.45** | 0.376    |
|          | Alt | 4.03 | 0.97 |         |          |          | Alt | 2.86 | 1.42 |         |          |
| 2        | Üst | 4.83 | 0.49 | 13.00** | 0.360    | 27       | Üst | 4.86 | 0.58 | 13.62** | 0.382    |
|          | Alt | 3.62 | 1.11 |         |          |          | Alt | 3.29 | 1.26 |         |          |
| 3        | Üst | 4.78 | 0.55 | 14.06** | 0.397    | 28       | Üst | 4.95 | 0.31 | 15.39** | 0.441    |
|          | Alt | 3.54 | 1.17 |         |          |          | Alt | 3.86 | 1.32 |         |          |
| 4        | Üst | 4.68 | 0.81 | 12.53** | 0.344    | 29       | Üst | 4.78 | 0.62 | 18.49** | 0.533    |
|          | Alt | 3.58 | 1.30 |         |          |          | Alt | 2.99 | 1.29 |         |          |
| 5        | Üst | 4.79 | 0.63 | 12.97** | 0.359    | 30       | Üst | 4.83 | 0.44 | 22.21** | 0.622    |
|          | Alt | 3.57 | 1.21 |         |          |          | Alt | 3.02 | 1.32 |         |          |
| 6        | Üst | 4.69 | 0.82 | 12.92** | 0.358    | 31       | Üst | 4.77 | 0.70 | 14.59** | 0.415    |
|          | Alt | 3.35 | 1.26 |         |          |          | Alt | 3.18 | 1.42 |         |          |
| 7        | Üst | 4.88 | 0.46 | 12.75** | 0.351    | 32       | Üst | 4.86 | 0.42 | 14.58** | 0.415    |
|          | Alt | 3.77 | 1.17 |         |          |          | Alt | 3.38 | 1.32 |         |          |
| 8        | Üst | 4.60 | 0.91 | 13.13** | 0.365    | 33       | Üst | 4.98 | 0.13 | 16.04** | 0.462    |
|          | Alt | 3.14 | 1.35 |         |          |          | Alt | 3.76 | 1.3  |         |          |
| 9        | Üst | 4.92 | 0.49 | 14.25** | 0.404    | 34       | Üst | 4.57 | 0.91 | 13.84** | 0.390    |
|          | Alt | 3.71 | 1.31 |         |          |          | Alt | 2.85 | 1.34 |         |          |
| 10       | Üst | 4.86 | 0.46 | 10.44** | 0.266    | 35       | Üst | 4.87 | 0.50 | 16.84** | 0.486    |
|          | Alt | 3.98 | 1.15 |         |          |          | Alt | 3.50 | 1.39 |         |          |
| 11       | Üst | 4.71 | 0.87 | 15.69** | 0.451    | 36       | Üst | 4.81 | 0.50 | 19.99** | 0.571    |
|          | Alt | 2.94 | 1.39 |         |          |          | Alt | 2.77 | 1.24 |         |          |
| 12       | Üst | 4.77 | 0.72 | 11.98** | 0.324    | 37       | Üst | 4.94 | 0.28 | 11.13** | 0.292    |
|          | Alt | 3.61 | 1.21 |         |          |          | Alt | 3.63 | 1.26 |         |          |
| 13       | Üst | 4.43 | 1.11 | 11.76** | 0.316    | 38       | Üst | 4.37 | 1.22 | 13.41** | 0.375    |
|          | Alt | 2.84 | 1.41 |         |          |          | Alt | 2.62 | 1.24 |         |          |
| 14       | Üst | 4.54 | 0.87 | 9.20**  | 0.220    | 39       | Üst | 4.69 | 0.75 | 13.09** | 0.364    |
|          | Alt | 3.39 | 1.28 |         |          |          | Alt | 3.19 | 1.34 |         |          |
| 15       | Üst | 4.79 | 0.69 | 10.04** | 0.251    | 40       | Üst | 4.92 | 0.37 | 13.39** | 0.374    |
|          | Alt | 3.61 | 1.27 |         |          |          | Alt | 3.73 | 1.19 |         |          |
| 16       | Üst | 2.12 | 1.56 | 11.92** | 0.321    | 41       | Üst | 4.93 | 0.26 | 15.10** | 0.432    |
|          | Alt | 2.84 | 1.34 |         |          |          | Alt | 3.63 | 1.34 |         |          |
| 17       | Üst | 4.79 | 0.78 | 16.04** | 0.462    | 42       | Üst | 4.68 | 0.76 | 9.81**  | 0.243    |
|          | Alt | 3.84 | 1.28 |         |          |          | Alt | 2.56 | 1.33 |         |          |
| 18       | Üst | 4.75 | 0.80 | 10.44** | 0.266    | 43       | Üst | 4.40 | 1.14 | 18.31** | 0.528    |
|          | Alt | 3.03 | 1.34 |         |          |          | Alt | 2.99 | 1.33 |         |          |
| 19       | Üst | 3.97 | 1.29 | 11.01** | 0.288    | 44       | Üst | 4.86 | 0.38 | 16.29** | 0.469    |
|          | Alt | 2.56 | 1.45 |         |          |          | Alt | 3.14 | 1.33 |         |          |
| 20       | Üst | 4.62 | 0.87 | 11.61** | 0.310    | 45       | Üst | 4.89 | 0.48 | 10.21** | 0.258    |
|          | Alt | 4.18 | 1.10 |         |          |          | Alt | 3.49 | 1.26 |         |          |
| 21       | Üst | 4.88 | 0.42 | 11.55** | 0.308    | 46       | Üst | 4.69 | 0.71 | 8.16**  | 0.182    |
|          | Alt | 3.39 | 1.30 |         |          |          | Alt | 3.25 | 1.43 |         |          |
| 22       | Üst | 4.61 | 0.83 | 11.61** | 0.310    | 47       | Üst | 4.43 | 1.12 | 10.11** | 0.254    |
|          | Alt | 2.85 | 1.36 |         |          |          | Alt | 3.17 | 1.37 |         |          |
| 23       | Üst | 3.75 | 1.38 | 9.24**  | 0.222    | 48       | Üst | 4.84 | 0.54 | 14.62** | 0.416    |
|          | Alt | 2.37 | 1.35 |         |          |          | Alt | 3.02 | 1.34 |         |          |
| 24       | Üst | 4.94 | 0.40 | 18.85** | 0.542    | 49       | Üst | 4.78 | .73  | 18.27** | 0.527    |
|          | Alt | 3.53 | 1.30 |         |          |          | Alt | 3.01 | 1.40 |         |          |
| 25       | Üst | 4.92 | 0.43 | 16.44** | 0.474    | 50       | Üst | 4.96 | .30  | 11.14** | 0.293    |
|          | Alt | 3.50 | 1.28 |         |          |          | Alt | 3.76 | 1.34 |         |          |

\*\*p<.01

Çizelge 5'ten de görüleceği gibi, 50 tutum cümlesinin tamamının "t" değeri 424 serbestlik derecesinde ve 0.01 manidarlık düzeyinde anlamlı bulunmuştur. Diğer bir deyişle bu tutum cümleleri olumlu tutuma sahip bireyleri olumsuz tutuma sahip bireylerden ayırt edebilmektedir. Bu cümlelerin yeterli madde ayırt edicilik gücüne sahip olduğu sonucuna varılmıştır. Buna ek olarak hesaplan t değerlerinin etki büyüklüklerinin incelenmesi amacıyla eta kare ( $\eta^2$ ) değerleri de tabloda verilmiştir.

### **Ölçeğin Aşamalı Tepki Modeline Göre Geliştirilmesi**

Örtük özellikler kuramına dayalı olarak Beden Eğitimi dersine yönelik tutum ölçeğinin geliştirilmesi için 787 öğrencinin verdiği yanıtlar MULTILOG 7.0 (Thissen, 1991) istatistiksel programında yer alan Samejima'nın "Aşamalı Tepki" (Graded Response) modeli kullanılarak analiz edilmiştir. Bu analizler sonucunda 50 tutum cümlesine ait madde ayırt edicilik değeri (a), her cümle için kategori eşik değeri (b), standart hatalar, madde bilgi fonksiyonları ve tutum düzeyi kestirimleri elde edilmiştir.

### **Ayırt Edicilik Gücü İndeksi**

Modelle uyumu yüksek ve yeterli olan elli tutum cümlesinin ayırt edicilik gücü incelenmiştir. Tutum cümlelerinin ayırt edicilik gücü değerleri Çizelge 6 'da verilmiştir.

**Çizelge 6. Ön Denemede Kullanılan Tutum Cümlelerinin Aşamalı Tepki Modeliyle Kestirilen Madde Ayırt Edicilik Değerleri**

| Madde No | a    | Madde No | a    |
|----------|------|----------|------|
| 1        | 1.28 | 26       | 1.76 |
| 2        | 1.67 | 27       | 2.07 |
| 3        | 1.50 | 28       | 1.74 |
| 4        | 1.15 | 29       | 1.35 |
| 5        | 1.26 | 30       | 2.50 |
| 6        | 1.15 | 31       | 1.66 |
| 7        | 1.24 | 32       | 1.62 |
| 8        | 1.31 | 33       | 1.53 |
| 9        | 1.05 | 34       | 1.43 |
| 10       | 0.97 | 35       | 1.41 |
| 11       | 1.88 | 36       | 2.68 |
| 12       | 1.14 | 37       | 1.07 |
| 13       | 1.06 | 38       | 0.62 |
| 14       | 0.13 | 39       | 2.06 |
| 15       | 1.90 | 40       | 1.67 |
| 16       | 1.89 | 41       | 2.25 |
| 17       | 1.95 | 42       | 0.62 |
| 18       | 1.00 | 43       | 2.05 |
| 19       | 0.74 | 44       | 2.38 |
| 20       | 1.16 | 45       | 0.54 |
| 21       | 1.03 | 46       | 1.04 |
| 22       | 1.04 | 47       | 1.01 |
| 23       | 1.16 | 48       | 1.98 |
| 24       | 2.07 | 49       | 1.89 |
| 25       | 1.66 | 50       | 1.13 |

Çizelge 6'da görüldüğü gibi cümlelerin ayırt edicilik gücü değerleri 0.13 ile 2.68 aralığında değişmektedir. ATM'nde madde ayırt edicilik güçleri  $-\infty$  ve  $+\infty$  arasında değişen değerler alır. ATM'inde değeri sıfırın altında olan maddeler ölçeğe dahil edilmez. İki ve üç parametrelili lojistik modellerde ayırt edicilik gücü değeri 1.00 civarında olan maddelerin ayırt edicilik gücü değerinin yeterli olduğu kabul edilir (Hambleton ve Swaminathan, 1985; Hambleton, Swaminathan ve Rogers, 1991). Bu çalışmada maddeler ayırt edicilik güçlerine göre seçilirken  $a_j$  değeri 1.00 ve üzerinde olan maddeler alınmıştır. Bu eleme sonrasında  $a_j$  değeri 1.00'in altında olan 10, 14, 19, 38, 42 ve 45 no'lu maddeler ölçekten çıkarılmış ve ATM ile yapılan analizler sonucunda 44 madde kalmıştır.

KTK ve ÖÖK dahilinde yapılan korelasyona dayalı analiz, iç tutarlılık ölçütüne dayalı analiz ve ATM'ne dayalı madde analizleri sonucunda her üç yöntem dahilinde ortak olan maddeler Çizelge 7'de verilmiştir. Buna göre, her üç yöntemle ortak olarak seçilen madde sayısının 40 olduğu, yani üç yöntem kapsamında ortak madde sayısı oranının %80 olduğu görülmüştür.

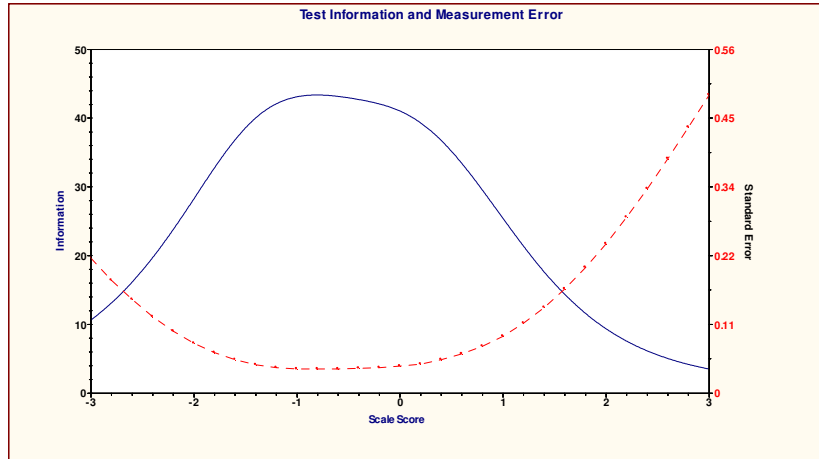
**Çizelge 7. Korelasyona Dayalı Analiz, İç Tutarlılık Ölçütüne Dayalı (t-test) Madde Analizi ve Aşamalı Tepki Modeli ile Seçilen Ortak Maddeler**

| Madde No | r     | t     | a    | Madde No | r  | t    | a     |      |   |
|----------|-------|-------|------|----------|----|------|-------|------|---|
| 1        | 0.56  | 14.50 | 1.28 | ✓        | 26 | 0.52 | 13.45 | 1.76 | ✓ |
| 2        | 0.53  | 13.00 | 1.67 | ✓        | 27 | 0.59 | 13.62 | 2.07 | ✓ |
| 3        | 0.59  | 14.06 | 1.50 | ✓        | 28 | 0.62 | 15.39 | 1.74 | ✓ |
| 4        | 0.48  | 12.53 | 1.15 | ✓        | 29 | 0.65 | 18.49 | 1.35 | ✓ |
| 5        | 0.54  | 12.97 | 1.26 | ✓        | 30 | 0.64 | 22.21 | 2.50 | ✓ |
| 6        | 0.48  | 12.92 | 1.15 | ✓        | 31 | 0.55 | 14.59 | 1.66 | ✓ |
| 7        | 0.50  | 12.75 | 1.24 | ✓        | 32 | 0.55 | 14.58 | 1.62 | ✓ |
| 8        | 0.57  | 13.13 | 1.31 | ✓        | 33 | 0.59 | 16.04 | 1.53 | ✓ |
| 9        | 0.55  | 14.25 | 1.05 | ✓        | 34 | 0.53 | 13.84 | 1.43 | ✓ |
| 10       | 0.41  | 10.44 | 0.97 | ✓        | 35 | 0.58 | 16.84 | 1.41 | ✓ |
| 11       | 0.55  | 15.69 | 1.88 | ✓        | 36 | 0.67 | 19.99 | 2.68 | ✓ |
| 12       | 0.49  | 11.98 | 1.14 | ✓        | 37 | 0.48 | 11.13 | 1.07 | ✓ |
| 13       | 0.49  | 11.76 | 1.06 | ✓        | 38 | 0.53 | 13.41 | 0.62 | ✓ |
| 14       | 0.34  | 9.20  | 0.13 | ✓        | 39 | 0.54 | 13.09 | 2.06 | ✓ |
| 15       | 0.47  | 10.04 | 1.90 | ✓        | 40 | 0.52 | 13.39 | 1.67 | ✓ |
| 16       | -0.22 | 11.92 | 1.89 | ✓        | 41 | 0.57 | 15.10 | 2.25 | ✓ |
| 17       | 0.57  | 16.04 | 1.95 | ✓        | 42 | 0.39 | 9.81  | 0.62 | ✓ |
| 18       | 0.47  | 10.44 | 1.00 | ✓        | 43 | 0.64 | 18.31 | 2.05 | ✓ |
| 19       | 0.36  | 11.01 | 0.74 | ✓        | 44 | 0.61 | 16.29 | 2.38 | ✓ |
| 20       | 0.20  | 11.61 | 1.16 | ✓        | 45 | 0.46 | 10.21 | 0.54 | ✓ |
| 21       | 0.41  | 11.55 | 1.03 | ✓        | 46 | 0.39 | 8.16  | 1.04 | ✓ |
| 22       | 0.41  | 11.61 | 1.04 | ✓        | 47 | 0.47 | 10.11 | 1.01 | ✓ |
| 23       | 0.36  | 9.24  | 1.16 | ✓        | 48 | 0.59 | 14.62 | 1.98 | ✓ |
| 24       | 0.61  | 18.85 | 2.07 | ✓        | 49 | 0.63 | 18.27 | 1.89 | ✓ |
| 25       | 0.63  | 16.44 | 1.66 | ✓        | 50 | 0.31 | 11.14 | 1.13 | ✓ |

### Madde ve Test Bilgi Fonksiyonu

ÖÖK'na dayalı olarak tutum ölçeği geliştirme çalışması sırasında madde bilgi fonksiyonu değerleri incelenmiştir.

Bireyin yeteneğini veya tutum düzeyini kestirmede her bir yetenek veya tutum düzeyinde maddenin testin doğruluğuna katkısı madde bilgi fonksiyonu olarak tanımlanır (Koch, 1983). Madde bilgi fonksiyonu KTK'daki maddenin testin bütünüyle olan korelasyonundan farklıdır. Çünkü madde bilgi fonksiyonu diğer maddelerden bağımsızdır. Madde bilgi fonksiyonu ölçülen özelliğe bağlıdır. Yani ölçmenin doğruluğu farklı tutum düzeylerinde farklılık gösterir.

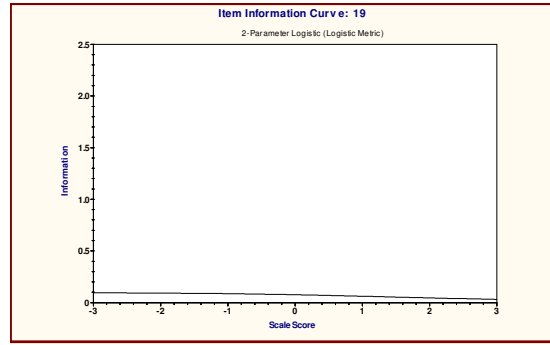


**Şekil 3.** 44 Maddeye İlişkin Elde Edilen Madde Bilgi Fonksiyonu Grafiği

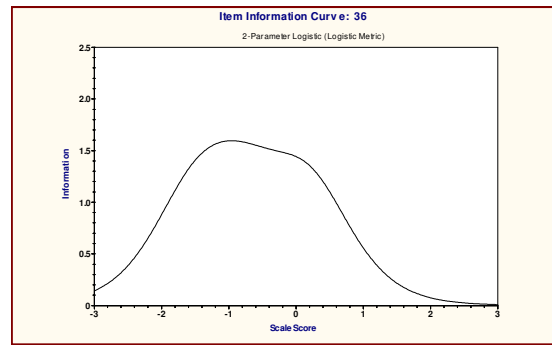
Analizler sonucunda seçilen 44 maddenin test bilgi fonksiyonu Şekil 3'te verilmiştir.

Şekil 3'ten de görüleceği gibi, maddelerin (-1 ve +1) tutum düzeyi aralığında olan bireylerin tutumları hakkında daha fazla bilgi verdiği anlaşılmıştır. Bu tutum düzeyleri aralığında en fazla bilgi veren maddelerin belirlenmesi için cümlelerin madde bilgi fonksiyonları incelenmiştir. (-1, +1) tutum düzeyi aralığının her bir noktasında 0.4 ile 2.4 aralığında madde bilgi fonksiyonu değerine sahip maddeler arasından bilgi fonksiyonu değeri yüksek olanlar seçilmiştir. Maddeler arasından seçim yapılırken olumlu tutum ifadesi içeren cümleler ile olumsuz tutum ifadesi içeren cümle sayısının birbirine eşit olmasına özen gösterilmiştir. Bu analiz sonrasında madde bilgi fonksiyonu düşük olan 4, 6, 9, 12, 13, 18, 20, 21, 22, 23, 37, 46, 47 ve 50 nolu maddeler ölçekten çıkarılmış ve geriye 30 madde kalmıştır.

Madde bilgi fonksiyonu düşük (madde 19) ve yüksek (madde 36) olan maddelerin madde bilgi fonksiyonu grafikleri Şekil 4 ve Şekil 5'te örnek olarak verilmiştir.



**Şekil 4.** Madde Bilgi Fonksiyonu Grafiği (Düşük Ayırtedicilik Değerli Madde)



**Şekil 5.** Madde Bilgi Fonksiyonu Grafiği (Yüksek Ayırtedicilik Değerli Madde)

ATM'inde madde bilgi fonksiyonlarının incelenmesi ve yüksek madde bilgi fonksiyonu veren otuz maddenin seçilmesi nedeni ile madde seçim tekniklerine göre madde parametrelerinin karşılaştırılmasında doğru ve güvenilir sonuçların elde edilmesi için madde sayısının sabitlenmesi düşünülmüştür. Bu seçim, madde ölçek korelasyonları ve t değerleri büyükten küçüğe sıralandıktan sonra ilk 30 madde alınarak yapılmıştır. Bu seçim sonrasında korelasyona dayalı analiz, içtutarlılık ölçütüne dayalı analiz ve ATM ile oluşturulan ölçeklerden seçilen otuz maddeden yirmisekiz'inin ortak olduğu, yani ortak madde oranının yaklaşık %94 olduğu görülmüştür.

### Verilerin Çözümlemesi

ÖÖK'nda yer alan Samejima'nın aşamalı tepki modeliyle madde parametreleri ve bireylerin tutum düzeylerinin kestirilebilmesi için

MULTILOG 7.0 (Thissen 1988) programından yararlanılmıştır. Analizler sonucunda her cümlenin ayırteçilik değeri (a), kategori eşik değeri (b) ve standart hataları, madde bilgi fonksiyonu ve ölçek bilgi fonksiyonu değeri, ölçeğin marjinal güvenilirliği ve bireylerin tutum düzeyi sonuçları elde edilmiştir.

KTK'ında kullanılan maddelerin ölçeğin bütünüyle olan korelasyonu incelenmiş ve maddelerin ayırteçilik gücü hakkında sayısal değerler elde edilmiştir. ATM ile kestirilen madde ayırteçilik gücü (a) değerleri ile madde-ölçek puanı korelasyonu ile belirlenen madde ayırteçilik gücü değerleri arasındaki ilişki Spearman Sıra Farkları Korelasyon katsayısıyla belirlenmiştir. Bulunan korelasyon katsayısının anlamlılığı Fisher'in z istatistiği ile test edilmiştir.

ATM'nde her bir cümlenin kategori güçlük düzeylerinin aritmetik ortalaması alınarak tutum cümlesinin ölçtüğü tutum düzeyi (b) değeri hesaplanmıştır. Benzer şekilde cevaplayıcıların maddelere verdikleri tepkinin, ham puanların aritmetik ortalaması alınarak maddenin ölçtüğü ortalama tutum düzeyi değeri elde edilmiştir. Daha sonra, her tutum ifadesinin iki farklı yöntemle belirlenen tutum düzeyi değerleri arasındaki korelasyon incelenmiştir. Bulunan Spearman Sıra Farkları Korelasyon katsayısının anlamlılığı Fisher'in z istatistiği ile test edilmiştir.

Klasik test kuramında dayalı olarak geliştirilen Beden Eğitimi Dersi (BED) tutum ölçeğinin iç tutarlılık güvenilirliği Cronbach  $\alpha$  katsayısı ile incelenmiştir. Cronbach  $\alpha$  katsayısının hesaplanması sırasında

$$\alpha = \frac{k}{k-1} \left[ 1 - \frac{\sum_{i=1}^k \sigma_i^2}{\sigma_x^2} \right] \quad (12)$$

eşitliğinden yararlanılmıştır (Crocker ve Algina 1986:138). Örtük özellikler kuramına dayalı olarak geliştirilen ölçeğin iç tutarlılık güvenliğinin Cronbach  $\alpha$  katsayısı ile hesaplanması sırasında ölçekteki maddelerin

varyans deęerleri z deęerine çevrilmiştir. Ölçek varyansı ise model yardımıyla kestirilen tutum puanlarından hesaplanmıştır (Hambleton, Swaminathan ve Rogers, 1991).

ATM ile ölçeğin marjinal güvenilirlik katsayısı ve test bilgi fonksiyonu deęerleri Multilog 7.0 programının uygulanması sonucu elde edilmiştir. Elde edilen nihai ölçeğin yapı geçerliğinin incelenmesi amacıyla Temel Bileşenler Analizi uygulanmıştır.

Madde parametrelerinin incelenmesi amacıyla elde edilen nihai ölçek 213 kız ve 213 erkek öğrenciye uygulanmış ve elde edilen veriler üzerinden KTK ve ÖÖK çerçevesinde ATM kullanılarak madde ayırt edicilik deęerleri elde edilmiştir.

### **Geçerlik ve Güvenirlik**

**Geçerlik:** Geliştirilen bir ölçme aracının geçerliğinin sınanmasında kullanılan yöntemler kapsam geçerliği, bir ölçüte dayalı geçerlik ve yapı geçerliği'dir (Özgüven, 1998). Kapsam geçerliğinin sınanmasında genellikle alandaki uzman olan kişilerin görüşlerinden faydalanılır. Bu çalışma sonrasında ölçekteki maddelerin hangilerinin ölçülmek istenen özellięi daha net olarak yansıttığı ile ilgili uzmanların belirttięi ortak görüşler dikkate alınarak ölçeğin nihai formu oluşturulur. Bunun yanında ölçekten elde edilen puanların tutum boyutunun hangi bölgesinde ve ne kadar genişlikte puanlar oluşturduęu da kapsam geçerlilięi için bir gösterge olarak kabul edilebilir.

Ölçüte dayalı geçerliğinin sınanmasında kullanılan yöntemlerden biri olan ayırt etmeye dayalı yaklaşımda ölçülmek istenen tutumla ilgili olarak olumlu ve olumsuz tutuma sahip olduęu bilinen iki farklı gruba aynı ölçek uygulanır. Bu gruplardan elde edilen ölçek puanları arasında anlamlı bir farkın olup olmadıęı t istatistięi ile belirlenir. Aradaki farkın anlamlı olması ölçeğin farklı tutumlara sahip bireyleri birbirlerinden ayırt edebildięini, yani farklı tutum düzeylerini ölçebildięini göstermektedir. Bu sonuç ölçeğin ölçüte dayalı geçerliğinin bir kanıtı sayılabilir.



Bu çalışma dahilinde elde edilen nihai ölçme aracının geçerliğinin sınanmasında kapsam geçerliği, yapı geçerliği ve ölçüte dayalı geçerlik türleri kullanılmıştır. Kapsam geçerliği için ölçme aracı geliştirilirken ölçme ve değerlendirme alanından ve beden eğitimi alanından uzmanlar tarafından maddeler incelenmiş ve yapılan eleştiriler doğrultusunda ölçekte yer alan maddeler yeniden düzenlenmiştir. Ölçüte dayalı geçerlik çalışması için ayırt etmeye dayalı yaklaşım dahilinde beden eğitimi'ne karşı olumlu tutuma sahip oldukları kabul edilen (Özer ve Aktop, 2003) spor lisanslı öğrenciler ile herhangi bir spor branşından lisanslı olmayan öğrencilerin tutum ölçeğinden aldıkları ortalama puanlar karşılaştırılmış ve puanlar arasında anlamlı bir farkın olup olmadığı araştırılmıştır. Yapı geçerliği için de Temel Bileşenler Analizi yapılmıştır. Ölçek maddelerinin tek bir faktör yapısında aldıkları özdeğerler incelenmiş ve ölçeğin yapısal olarak tek boyutlu olduğu ortaya konmuştur.

**Güvenirlilik:** KTK'nda güvenirlilik kavramı gözlenen puanlar ile gerçek puanların bir korelasyonunun fonksiyonu olarak tanımlanır. KTK'nda güvenirliliğin hesaplanmasında farklı yöntemler kullanılmaktadır. Bu yöntemlerle belirlenen güvenirlilikler, test tekrar-test güvenirliliği, paralel testler güvenirliliği, iki yarı güvenirliliği, Cronbach  $\alpha$  güvenirliliği, KR-20 ve KR-21 güvenirliliği ile varyans analizi yöntemi ile güvenirliliğin sınanması olarak sıralanabilir.

Tutum ölçeklerinin güvenirliliklerinin hesaplanmasında en çok kullanılan yöntem maddeler arası korelasyonlara dayalı olarak elde edilen Cronbach Alpha katsayılarının hesaplanmasıdır. Tutum ölçekleri gibi çok kategorili ölçeklerin (politomik) güvenirliliklerinin hesaplanmasında Cronbach Alpha katsayıları kullanılmaktadır (Keeves, 1988). Likert ölçekleri kullanılarak yapılan çoğu araştırmada da bu yöntem kullanılmaktadır.

Madde bilgi fonksiyonu  $\theta$  yeteneğinin fonksiyonudur. Her bir yetenek seviyesinde madde ve test bilgi fonksiyonu hesaplanır. Madde bilgi fonksiyonu maddenin ölçtüğü yetenek düzeyi (b) bireylerin yeteneğine yakın veya eşitse madde hakkında daha fazla bilgi verir. Ayrıca maddenin

ayırteçicilik değeri yükseldikçe ve ölçmenin standart hatası azaldıkça madde bilgi fonksiyonu değeri artar. Bilgi fonksiyonu ölçmenin standart hatası ile ters orantılıdır.

Madde ve test bilgi fonksiyonu birey örnekleminde bağımsız olarak elde edilmektedir, ayrıca her bir yetenek seviyesinde ölçmenin standart hatasını temsil etmektedir. Bu özelliğinden dolayı madde ve test bilgi fonksiyonu klasik test kuramındaki güvenilirlik ve standart hata kavramlarına alternatif olarak görülmektedir (Berberoğlu, 1988:28).

Buna ek olarak ölçekteki maddelerin madde bilgi fonksiyonlarının toplanmasıyla elde edilen test bilgi fonksiyonu bütün yetenek seviyesinde tek bir dağılım veriyorsa, bütün yetenek seviyelerindeki test bilgi fonksiyonunun ortalamasının "güvenirlik" katsayısını belirlediği ve bu güvenilirlik katsayısının "marjinal güvenilirlik" olarak adlandırılmaktadır.

Bu çalışma kapsamında kullanılan ölçeğin tutum ölçeği olması nedeni ile KTK kapsamında maddeler arası korelasyonlara dayalı olarak elde edilen Cronbach Alpha Katsayısı ve ÖÖK kapsamında marjinal güvenilirlik katsayıları kullanılmıştır.

## BÖLÜM IV

### BULGULAR VE YORUMLAR

Bu bölümde araştırmanın bulguları açıklanmış, bulgulara dayalı yorumlara yer verilmiştir.

#### **Korelasyona Dayalı Analiz İç Tutarlılık Ölçütüne Dayalı Analiz ve Aşamalı Tepki Modeli'ne Göre Geliştirilen Ölçeklerin Madde Ayırtedicilik Gücü Değerleri Arasındaki İlişkinin İncelenmesi**

Elli maddelik beden eğitimi'ne yönelik tutum ölçeğinin deneme uygulaması sonucunda elde edilen verilere KTK çerçevesinde korelasyona dayalı analiz ve içtutarlılık ölçütüne dayalı analiz teknikleri kullanılarak madde analizleri yapılmıştır. ÖÖK çerçevesinde ise ATM kullanılarak madde analizleri yapılmıştır.

Deneme uygulaması sonucu her üç yöntem ile oluşturulan ölçeklere ilişkin kestirilen madde ayırtedicilik değerleri arasındaki korelasyon katsayıları Çizelge 8'de verilmiştir.

**Çizelge 8. Deneme Uygulaması Sonucu Farklı Madde Analizi Yöntemleri İle Elde Edilen Madde Ayırtedicilik Değerlerine İlişkin Korelasyon Katsayıları (40 madde)**

|                    | <b>r Değerleri</b> | <b>t Değerleri</b> |
|--------------------|--------------------|--------------------|
| <b>t Değerleri</b> | 0.961**            |                    |
| <b>a Değerleri</b> | 0.827**            | 0.865**            |

\*\* p<.01

Çizelge 8 incelendiğinde, korelasyona dayalı analiz yöntemi ile seçilen maddelerden elde edilen madde ayırtedicilik gücü değerleri ile iç tutarlılık yöntemi ile seçilen maddelerden elde edilen madde ayırtedicilik gücü değerleri arasında yüksek düzeyde ( $r=0.961$ ) pozitif ve anlamlı bir ilişki bulunmuştur. Korelasyona dayalı analiz yöntemi ile seçilen maddelerden elde edilen madde ayırtedicilik gücü değerleri ile ATM ile elde edilen madde ayırtedicilik gücü değerleri arasında da pozitif ve anlamlı ( $r=0.827$ ) ilişki

bulunmuştur. İçtutarlılık ölçütüne dayalı analiz yöntemi ile seçilen maddelerden elde edilen madde ayırteçicilik gücü değeri ile ATM ile seçilen maddelerden elde edilen madde ayırteçicilik gücü değeri arasında diğerlerine göre biraz daha düşük olmakla beraber pozitif ve anlamlı ( $r=0.865$ ) bir ilişki gözlenmiştir. Elde edilen bu bulgular Tezbaşaran ve Kelecioğlu'nun (2004) yapmış olduğu çalışmanın bulgularını desteklemektedir. Tezbaşaran ve Kelecioğlu da yaptıkları çalışmada deneme formundan madde-ölçek korelasyonları, alt-üst grup tekniği ve ATM'ne dayalı olarak seçilen maddelerden kestirilen madde ayırteçicilik gücü değeri arasındaki ilişkiyi incelemiştir. Yapılan çalışmada madde-ölçek korelasyonları tekniği ve alt-üst grup tekniği'ne dayalı olarak kestirilen madde ayırteçicilik değerlerine ilişkin korelasyon katsayısının diğerlerine oranla daha yüksek olduğu ( $r=0.95$ ) bulunmuştur. Benzer şekilde, alt-üst grup tekniği ve ATM'ne dayalı olarak kestirilen madde ayırteçicilik gücü değeri arasında diğerlerine oranla daha düşük bir korelasyon değeri elde edilmiştir ( $r=0.86$ ). Bu çalışmada da sonuçlar benzer niteliktedir. Bu bulgu, deneme formundan elde edilen madde ayırteçicilik gücü değerlerinin her üç yöntem için de benzerlik gösterdiği şeklinde yorumlanabilir.

Deneme ölçeğinden korelasyona dayalı analiz yöntemi, içtutarlılık ölçütüne dayalı analiz yöntemi ve ATM'ne dayalı analiz yöntemi ile seçilen maddelerle oluşturulan üç farklı ölçeğin ortak otuz maddesi ile yapılan incelemede ölçek maddelerinin  $r$ ,  $t$  ve  $a$  değerleri arasındaki korelasyon değerleri Çizelge 9'da verilmiştir.

**Çizelge 9. Deneme Uygulaması Sonucu Farklı Madde Analizi Yöntemleri İle Elde Edilen Madde Ayırteçicilik Değerlerine İlişkin Korelasyon Katsayıları (30 madde)**

|                    | <b>r Değerleri</b> | <b>t Değerleri</b> |
|--------------------|--------------------|--------------------|
| <b>t Değerleri</b> | 0.884**            |                    |
| <b>a Değerleri</b> | 0.637**            | 0.660**            |

\*\*  $p < .01$

Çizelge incelendiğinde deneme ölçeğinden korelasyona dayalı analiz yöntemi, içtutarlılık ölçütüne dayalı analiz yöntemi ve ATM'ne dayalı analiz yöntemi ile seçilen maddelerle oluşturulan üç farklı ölçeğin ortak otuz maddesi ile yapılan incelemede ölçek maddelerinin r, t ve a değerleri arasındaki korelasyonun deneme ölçeği ile elde edilen korelasyon değerlerinden daha düşük olduğu gözlenmiştir. En yüksek korelasyon r ve t değerleri arasında bulunurken, en düşük korelasyon değeri t ve a değerleri arasında gözlenmiştir. Elde edilen bulgular Tezbaşaran ve Kelecioğlu'nun (2004) yapmış olduğu çalışmanın bulguları ile paralellik göstermektedir. İlgili çalışmada, ölçek maddelerinden elde edilen korelasyon değeri incelenmiş ve deneme ölçeğinden elde edilen r, t ve a değerleri arasındaki korelasyonun, ortak otuz maddeden elde edilen korelasyon değerlerinden daha yüksek olduğu bulunmuştur. Araştırmacılar bu bulguyu, korelasyon değerlerinin madde sayısından etkilenmesinden ve tüm maddelerin r, t ve a değerleri arasındaki benzerliğin seçilen maddelerin r, t ve a değerlerinden daha yüksek olduğu şeklinde yorumlamışlardır. Benzer şekilde, yapılan çalışmadan elde edilen bulgu, korelasyon değerlerinin madde sayısından etkilendiği şeklinde yorumlanabilir.

### **Üç Farklı Yöntem İle Geliştirilmiş Tutum Ölçeklerinin Güvenirlikleri Arasındaki İlişkinin İncelenmesi**

Korelasyona dayalı analiz yöntemi, içtutarlılık ölçütüne dayalı analiz yöntemi ve ATM'ne dayalı analiz yöntemleri ile seçilen maddelerle oluşturulan üç farklı ölçeğin güvenilirlik katsayılarının incelenmesi amacıyla, bu üç yöntemle elde edilen ölçeklerin Cronbach- $\alpha$  ve Marjinal Güvenirlik değerleri Çizelge 10'da verilmiştir.

**Çizelge 10. Deneme Ölçeği ve Üç Ayrı Teknikle Elde Edilen Ölçeklerin Cronbach- $\alpha$  ve Marjinal Güvenirlik Katsayıları**

| Ölçekler                            | Cronbach- $\alpha$ | Marjinal Güvenirlik |
|-------------------------------------|--------------------|---------------------|
| Deneme Formu                        | 0,937              | 0,967               |
| Korelasyona Dayalı Analiz           | 0,939              | 0,965               |
| İçtutarlılık Ölçütüne Dayalı Analiz | 0,941              | 0,964               |
| Aşamalı Tepki Modeli                | 0,932              | 0,964               |

Ölçeğin deneme formu ve üç farklı yöntemle seçilen maddelerle oluşturulan ölçeklerin Cronbach- $\alpha$  ve marjinal güvenirlikleri yüksek bulunmuştur. Buna ek olarak, deneme formu ve diğer ölçeklerin Cronbach- $\alpha$  ve marjinal güvenirlikleri arasında da önemli sayılabilecek bir farklılık bulunmamıştır. Elde edilen bu bulgular, Çetin ve Kelecioğlu (2007) ve Tezbaşaran ve Kelecioğlu'nun (2004) çalışmaları ile paralellik göstermektedir. Yapılan çalışmalarda da benzer şekilde, ölçeğin deneme formu ve diğer üç teknik kullanılarak oluşturulan ölçeklerin Cronbach- $\alpha$  ve marjinal güvenirlikleri 0,92 ile 0,96 aralığında oldukça yüksek bulunmuştur. Elde edilen bu bulgu, korelasyona dayalı analiz, iç tutarlılık ölçütüne dayalı analiz ve ATM'ne dayalı analiz yöntemleri ile güvenirlikleri yüksek ölçeklerin oluşturulabileceği ve uygulanan analiz yöntemleri arasında ölçeklerin güvenirlikleri bakımından farklılıkların olmadığı şeklinde yorumlanabilir.

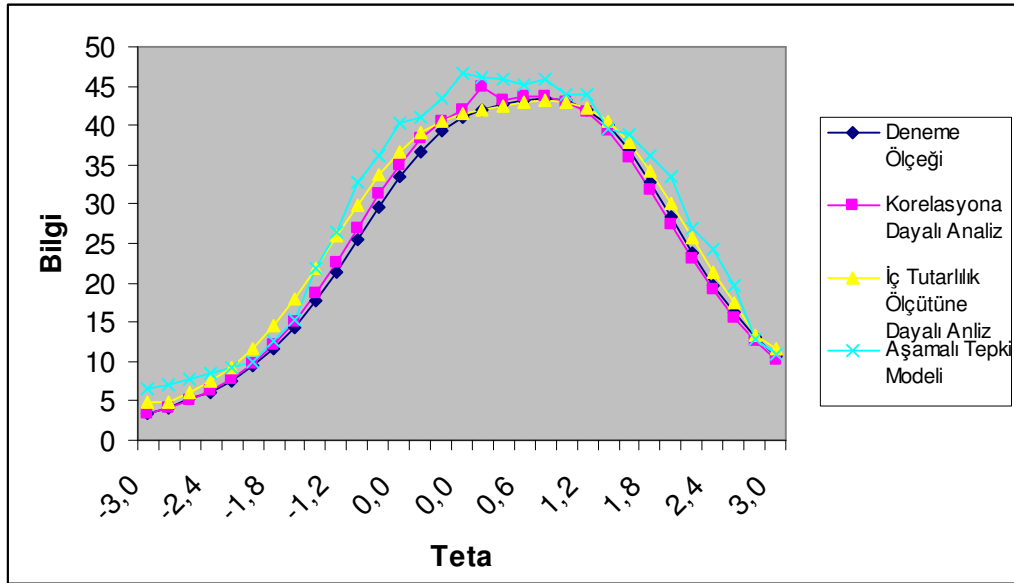
**Üç Farklı Yöntem İle Geliştirilmiş Tutum Ölçeklerinin Test Bilgi Fonksiyonları Arasındaki İlişkinin İncelenmesi**

Deneme ölçeği ve korelasyona dayalı analiz, içtutarlılık ölçütüne dayalı analiz ve ATM'ne dayalı analiz yöntemleri ile geliştirilen ölçeklerin tutum düzeylerinin (-3; +3) arasındaki test bilgi fonksiyonu değerleri ve fonksiyonlara ilişkin standart hata değerleri Çizelge 11'de verilmiştir.

**Çizelge 11. Deneme Ölçeği ve Üç Farklı Yöntem ile Geliştirilen Ölçeklerin Tutum Düzeylerine Göre Test Bilgi Fonksiyonu (TBF) Değerleri ve TBF'na İlişkin Standart Hata (Sh) Değerleri**

| Tutum Düzeyi | Ölçeğin Deneme Formu |      | Korelasyona Dayalı Analiz |      | İç Tutarlılık Ölçütüne Dayalı Analiz |      | Aşamalı Tepki Modeli |      |
|--------------|----------------------|------|---------------------------|------|--------------------------------------|------|----------------------|------|
|              | TBF                  | Sh   | TBF                       | Sh   | TBF                                  | Sh   | TBF                  | Sh   |
| -3.0         | 3.49                 | 0.53 | 3.50                      | 0.53 | 4.97                                 | 0.49 | 6.52                 | 0.39 |
| -2.8         | 4.16                 | 0.49 | 4.20                      | 0.49 | 4.95                                 | 0.45 | 7.16                 | 0.37 |
| -2.6         | 5.30                 | 0.44 | 5.10                      | 0.44 | 6.05                                 | 0.41 | 7.83                 | 0.36 |
| -2.4         | 6.14                 | 0.40 | 6.28                      | 0.44 | 7.47                                 | 0.37 | 8.54                 | 0.34 |
| -2.2         | 7.55                 | 0.36 | 7.80                      | 0.40 | 9.30                                 | 0.33 | 9.26                 | 0.33 |
| -2.0         | 9.35                 | 0.32 | 9.73                      | 0.39 | 11.61                                | 0.29 | 9.97                 | 0.32 |
| -1.8         | 11.59                | 0.29 | 12.15                     | 0.36 | 14.46                                | 0.26 | 12.67                | 0.31 |
| -1.6         | 14.34                | 0.26 | 15.12                     | 0.32 | 17.86                                | 0.24 | 15.32                | 0.30 |
| -1.4         | 17.61                | 0.24 | 18.64                     | 0.29 | 21.73                                | 0.21 | 21.89                | 0.29 |
| -1.2         | 21.35                | 0.21 | 22.64                     | 0.6  | 25.87                                | 0.20 | 26.38                | 0.28 |
| -1.0         | 25.40                | 0.19 | 26.92                     | 0.23 | 29.97                                | 0.18 | 32.77                | 0.28 |
| -0.8         | 29.52                | 0.18 | 31.19                     | 0.21 | 33.69                                | 0.17 | 36.06                | 0.28 |
| -0.6         | 33.40                | 0.17 | 35.07                     | 0.19 | 36.73                                | 0.17 | 40.26                | 0.28 |
| -0.4         | 36.74                | 0.16 | 38.26                     | 0.18 | 38.98                                | 0.16 | 41.13                | 0.27 |
| -0.2         | 39.30                | 0.16 | 40.54                     | 0.17 | 40.49                                | 0.16 | 43.40                | 0.27 |
| 0.0          | 41.02                | 0.16 | 41.96                     | 0.16 | 41.43                                | 0.16 | 46.65                | 0.27 |
| 0.2          | 42.05                | 0.15 | 44.80                     | 0.16 | 42.03                                | 0.15 | 46.15                | 0.27 |
| 0.4          | 42.69                | 0.15 | 43.31                     | 0.15 | 42.51                                | 0.15 | 45.98                | 0.27 |
| 0.6          | 43.15                | 0.15 | 43.61                     | 0.15 | 42.92                                | 0.15 | 45.15                | 0.27 |
| 0.8          | 43.36                | 0.15 | 43.60                     | 0.15 | 43.14                                | 0.15 | 45.88                | 0.27 |
| 1.0          | 43.10                | 0.15 | 43.06                     | 0.15 | 42.96                                | 0.15 | 43.91                | 0.27 |
| 1.2          | 42.07                | 0.15 | 41.72                     | 0.15 | 42.12                                | 0.15 | 43.83                | 0.27 |
| 1.4          | 40.02                | 0.16 | 39.34                     | 0.15 | 40.42                                | 0.16 | 39.66                | 0.27 |
| 1.6          | 36.85                | 0.17 | 35.94                     | 0.16 | 37.75                                | 0.16 | 38.91                | 0.27 |
| 1.8          | 32.77                | 0.18 | 31.79                     | 0.16 | 34.18                                | 0.17 | 36.06                | 0.27 |
| 2.0          | 28.28                | 0.18 | 27.34                     | 0.17 | 29.99                                | 0.18 | 33.58                | 0.28 |
| 2.2          | 23.82                | 0.21 | 22.99                     | 0.18 | 25.61                                | 0.20 | 26.99                | 0.29 |
| 2.4          | 19.73                | 0.23 | 19.20                     | 0.19 | 21.38                                | 0.22 | 24.31                | 0.30 |
| 2.6          | 16.15                | 0.25 | 15.55                     | 0.21 | 17.58                                | 0.24 | 19.56                | 0.31 |
| 2.8          | 13.12                | 0.27 | 12.63                     | 0.23 | 13.32                                | 0.26 | 12.77                | 0.32 |
| 3.0          | 10.61                | 0.31 | 10.22                     | 0.28 | 11.63                                | 0.29 | 10.97                | 0.33 |

Ölçeğin deneme formu ve korelasyona dayalı analiz, iç tutarlılık ölçütüne dayalı analiz ve ATM'ne dayalı analiz yöntemleri ile geliştirilen ölçeklerin tutum düzeylerinin (-3; +3) arasındaki test bilgi fonksiyonu grafikleri Şekil 6'da verilmiştir.



**Şekil 6.** Ölçeğin Deneme Formu ve Korelasyona Dayalı Analiz, İç Tutarlılık Ölçütüne Dayalı Analiz Ve ATM'ne Dayalı Analiz Yöntemleri İle Geliştirilen Ölçeklerin Test Bilgi Fonksiyonu Grafikleri

Test Bilgi Fonksiyonlarına ilişkin Çizelge 11 ve Şekil 6 incelendiğinde ATM ile geliştirilen ölçeğin tutum düzeylerine göre ölçeğin deneme formu ve diğer iki yöntem ile geliştirilen ölçeklere göre görece olarak yüksek bilgi verdiği, korelasyona dayalı analiz ve içtutarlılık ölçütüne dayalı analiz yöntemleri ile geliştirilen ölçeklerin ise deneme ölçeği ile yakın düzeyde bilgi verdikleri görülmüştür. Elde edilen bu bulgu Tezbaşaran ve Kelecioğlu'nun (2004) yapmış oldukları çalışmanın bulguları ile paralellik göstermektedir. İlgili çalışmada da deneme ölçeği ile korelasyona dayalı analiz ve içtutarlılık ölçütüne dayalı analiz yöntemleri ile geliştirilen ölçeklerin verdikleri bilgi düzeylerinin birbirlerine daha yakın olduğu, ATM ile geliştirilen ölçeğin verdiği bilgi düzeyinin ise ölçeğin deneme formu ve diğer iki yöntem ile geliştirilmiş ölçeklerin verdiği bilgi düzeyinden daha yüksek olduğu bulunmuştur.

Bundan sonraki bölümlerde korelasyona dayalı analiz, içtutarlılık ölçütüne dayalı analiz yöntemleri ve ATM ile geliştirilen ölçeklerden her üç



yöntem için ortak olan ve Çizege 7'de belirtilen otuz maddeden oluşan nihai ölçeğin tekrar uygulanmasıyla elde edilen verilerden ölçeğin psikometrik özelliklerinin incelenmesine ve ölçeğin geçerlik ve güvenirliğine ilişkin bulgular verilmiştir.

### **Otuz Maddelik Nihai Ölçeğin Madde Parametrelerinin İncelenmesi**

Otuz maddelik nihai ölçeğin 213 erkek 213 kız toplam 426 ilköğretim 6., 7., ve 8. sınıf öğrencisine uygulanmasıyla elde edilen verilere KTK kapsamında madde analizi tekniği uygulanmıştır. Madde analizi sırasında maddelerin ölçek puanı ile olan Pearson momentler çarpımı korelasyonu incelenmiştir. Çizelge 12'de maddelerin ölçekle olan korelasyon değerleri verilmiştir.

**Çizelge 12. Beden Eğitimi Dersi Tutum Ölçeğindeki Toplam Ölçek Puanı ile Olan Korelasyon Değerleri**

| <b>Madde No</b> | <b>r</b> | <b>Madde No</b> | <b>r</b> |
|-----------------|----------|-----------------|----------|
| 1               | 0.58     | 16              | 0.69     |
| 2               | 0.61     | 17              | 0.73     |
| 3               | 0.72     | 18              | 0.71     |
| 4               | 0.60     | 19              | 0.61     |
| 5               | 0.61     | 20              | 0.53     |
| 6               | 0.79     | 21              | 0.51     |
| 7               | 0.62     | 22              | 0.70     |
| 8               | 0.63     | 23              | 0.68     |
| 9               | 0.67     | 24              | 0.75     |
| 10              | 0.58     | 25              | 0.66     |
| 11              | 0.78     | 26              | 0.68     |
| 12              | 0.76     | 27              | 0.70     |
| 13              | 0.61     | 28              | 0.71     |
| 14              | 0.61     | 29              | 0.63     |
| 15              | 0.62     | 30              | 0.71     |

Çizelge 12 incelendiğinde maddelerin ayırt edicilik değerlerinin 0.51 ile 0.79 arasında değiştiği görülmektedir. Madde ayırt edicilik değerinin 1'e yaklaşması durumunda maddenin testin bütünüyle ölçülen özelliği ölçtüğü ve testle ölçülen özelliğe sahip olanlar ile olmayanları ayırt edebildiği sonucuna varılır. Maddelerin ayırt edicilik değerlerinin ortalaması 0.66'dır. Bu durumda

ölçekteki maddelerin yüksek ayırt edicilik gücüne sahip olduğu sonucuna varılmıştır.

Aynı ölçekte yer alan maddelerin ikinci uygulama sonucu elde edilen verilerden yararlanılarak ATM'yle a (madde ayırt edicilik) parametresi kestirilmiştir. Çizelge 13'de maddelerin kestirilen a parametre değerleri yer almaktadır.

Çizelge 13 incelendiğinde ölçekteki maddelerin ayırt edicilik değerlerinin 1.34-3.65 arasında olduğu görülmektedir. Maddelerin ayırt edicilik değerlerinin ortalaması ise 2.11'dir. Bu durumda maddelerin Çizelge 11 de verilen kendi b değerleri etrafında, dar bir aralıkta yüksek ayırt edicilik değerine sahip olduğu görülmektedir.

**Çizelge 13. Beden Eğitimi Dersi Tutum Ölçeğinde Yer Alan Maddelerin Aşamalı Tepki Modeliyle Kestirilen "a" Parametre Değerleri**

| Madde No | a    | Madde No | a    |
|----------|------|----------|------|
| 1        | 1.86 | 16       | 1.89 |
| 2        | 1.78 | 17       | 2.73 |
| 3        | 2.43 | 18       | 2.49 |
| 4        | 1.59 | 19       | 2.10 |
| 5        | 1.34 | 20       | 1.88 |
| 6        | 3.65 | 21       | 1.58 |
| 7        | 1.65 | 22       | 2.24 |
| 8        | 2.58 | 23       | 1.94 |
| 9        | 2.36 | 24       | 2.10 |
| 10       | 1.47 | 25       | 1.36 |
| 11       | 3.36 | 26       | 1.76 |
| 12       | 3.32 | 27       | 2.19 |
| 13       | 2.21 | 28       | 1.94 |
| 14       | 1.78 | 29       | 1.70 |
| 15       | 2.24 | 30       | 2.51 |

KTK kapsamındaki madde analiziyle elde edilen madde ölçek korelasyonları ile ATM ile kestirilen madde ayırt edicilik değerleri (a) arasındaki ilişki Spearman Sıra Farkları Korelasyon katsayısı ile incelenmiştir

ve 0.72 olarak bulunmuştur. Madde analiziyle belirlenen madde-ölçek korelasyonu ile ATM'yle kestirilen madde ayırt edicilik ( $a$ ) değerleri arasında yüksek ilişki bulunmaktadır.

Bulunan korelasyon katsayısı Fisher'in  $z$  istatistiği ile test edildiğinde anlamlı bulunmuştur. Bu durumda her iki yöntemle kestirilen madde ayırt edicilik değerinin yakın değerler olduğu sonucuna varılmıştır. Çalışma sonucunda KTK ve ÖÖK ile elde edilen madde ayırt edicilik değerleri arasında yüksek ve anlamlı bir ilişkinin olduğu şeklindeki bulgu Kan (2006), Tezbaşaran (2003), Demirtaşlı (2002), Nartgün (2002) ve Erden'in (1997) çalışmalarında elde ettiği bulgular ile paralellik göstermektedir. İlgili çalışmalarda da KTK ve ÖÖK'na göre kestirilen madde ayırt değerleri arasında 0.80 ile 0.96 arasında değişen yüksek ilişki elde edilmiştir. Buna ek olarak elde edilen bulgular Stage (1997) ve Wiberg'in çalışmalarında elde ettiği bulgularla da paralellik göstermektedir. Stage (1997) çalışmasında KTK ile elde edilen madde ayırt edicilik değerleri ile ÖÖK ile elde edilen madde ayırt edicilik değerleri arasında 0.66 düzeyinde anlamlı korelasyon değeri bulunmuştur. Aynı şekilde Wiberg'de (2004) her iki kuram ile elde edilen madde ayırt edicilik değerleri arasında 0.75 düzeyinde pozitif ve anlamlı korelasyon katsayısı bulunmuştur. Bu bulgular elde edilen madde ayırt edicilik değerlerinin birbirleri ile tutarlı olduğu, farklı kuramlara göre kestirilen madde ayırt edicilik değerleri bakımından kuramlar arasında büyük farkların olmadığı şeklinde yorumlanabilir. KTK çerçevesinde maddelerin ölçeğin tümüyle olan korelasyonlarının yüksek çıkması bu maddelerin ölçülen tutumla ilgili olarak yüksek ayırt edicilik gücü değerlerine sahip olduğunun bir göstergesi olarak kabul edilmektedir.

KTK kapsamında bulunan madde ayırt edicilik değerine ilişkin katsayı -1.00 ile +1.00 arasında bir değer alır ve işareti eksi olan ve sıfır değerine yakın bulunan maddeler belirtilen tutumu ölçmek için yetersiz olarak sınıflandırılır. Katsayı değeri pozitif ve yüksek olan maddelerin yüksek ayırt edicilik değerine sahip oldukları kabul edilir. ÖÖK'nda da madde ayırt edicilik gücü değerleri hesaplanırken  $a_j$  parametrelerinin daha çok 0.75 ile

2.50 arasında deęişen deęerler aldıęı gözlenmektedir. Buna göre, herhangi bir maddenin  $a_j$  parametresine iliřkin deęeri yükseldikçe o maddenin ayırtedicilik gücü de artmaktadır. Belirtilen kuramsal temellere dayanarak, hem KTK ile elde edilen madde ölçek korelasyonları hem de ÖÖK ile elde edilen  $a_j$  deęerleri arasında yüksek ayırtedicilik deęerleri elde edilmesi temelinde, bu iki deęer arasında pozitif ve yüksek korelasyon katsayısı elde edilmesi beklenir. Bu çalışmada elde edilen 0.72 deęerindeki pozitif ve yüksek korelasyon katsayısı kuramsal gereklilięin karşılandığını, yani her iki kuram dahilinde yüksek ayırt edicilik gücü deęerlerinin elde edildiğini göstermektedir.

KTK kapsamında Beden Eğitimi dersi'ne yönelik tutum ölçeğinde yer alan maddelerin ölçtüęü ortalama tutum düzeyini belirlemek amacıyla madde puanlarının aritmetik ortalaması alınmıştır. Çizelge 14'te ölçekteki maddelerin ortalama deęerleri yer almaktadır.

**Çizelge 14. Beden Eğitimi ve Spor Dersi Tutum Ölçeğindeki Maddelerin Ortalama Deęerleri**

| Madde No | x    | Madde No | x    |
|----------|------|----------|------|
| 1        | 4.48 | 16       | 3.93 |
| 2        | 4.22 | 17       | 3.98 |
| 3        | 4.16 | 18       | 4.12 |
| 4        | 4.13 | 19       | 4.37 |
| 5        | 4.17 | 20       | 3.71 |
| 6        | 4.02 | 21       | 4.19 |
| 7        | 4.32 | 22       | 3.79 |
| 8        | 3.87 | 23       | 4.28 |
| 9        | 3.82 | 24       | 3.94 |
| 10       | 4.31 | 25       | 4.33 |
| 11       | 4.23 | 26       | 4.28 |
| 12       | 3.77 | 27       | 3.69 |
| 13       | 4.07 | 28       | 4.00 |
| 14       | 4.41 | 29       | 3.93 |
| 15       | 3.88 | 30       | 3.90 |

Klasik Test Kuramı kapsamında maddelerin ölçtüęü ortalama tutum düzeyleri incelendiğinde 3.69 - 4.48 arasında deęiřtięi görülmüřtür. Tutum düzeylerinin ortalaması 4.00 olarak bulunmuřtur. Maddelerin daha çok

olumlu tutumu belirten *katılıyorum* kategorisinde yığılma gösterdiği görülmüştür.

KTK'na dayalı analizler sırasında her madde için bir tane maddenin ölçtüğü tutum düzeyi değeri elde edilirken ATM için Multilog 7.0'da yapılan analizlerde her madde için kategori sayısının bir eksiği kadar tercih indeksi olarak kullanılan b değeri elde edilmiştir. Uygulanan ölçek beş kategorili olduğu için dört ayrı b değeri elde edilmiştir. KTK kapsamında elde edilen maddenin ölçtüğü tutum düzeyi değeri ile ATM'yle elde edilen b değerleri arasındaki ilişkiyi inceleyebilmek amacıyla ATM'nde her madde için elde edilen b değerlerinin aritmetik ortalaması alınmıştır.

İki ayrı yöntemle bulunan değerler arasındaki ilişkiyi inceleyebilmek için bu yola gidilmiştir. Ölçekte yer alan maddelerin ATM yardımıyla kestirilen (b) değerleri Çizelge 15'de verilmiştir.

**Çizelge 15. Maddelerin Aşamalı Tepki Modeli Yardımıyla Kestirilen “b” Değerleri**

| Madde No | b     | Madde No | b     |
|----------|-------|----------|-------|
| 1        | -.92  | 16       | -.44  |
| 2        | -.78  | 17       | -.75  |
| 3        | -.95  | 18       | -.82  |
| 4        | -.90  | 19       | -.33  |
| 5        | -.71  | 20       | -.69  |
| 6        | -.96  | 21       | -.27  |
| 7        | -.45  | 22       | -.77  |
| 8        | -.86  | 23       | -.15  |
| 9        | -.94  | 24       | -.91  |
| 10       | -.37  | 25       | -1.04 |
| 11       | -.69  | 26       | -.14  |
| 12       | -.53  | 27       | -.52  |
| 13       | -1.28 | 28       | -.72  |
| 14       | -.41  | 29       | -.35  |
| 15       | -.41  | 30       | -.84  |

Çizelge 15 incelendiğinde bulunan b değerlerinin tamamının negatif (-) değer aldığı görülmektedir. Tutum ölçekleri için bulunan b değerinin

“tercih indeksi” olarak kullanılması daha doğrudur. Bulunan b değerlerinin negatif (-) değer alması bu maddelerin daha çok tercih edildiğini yani popüler maddeler olduğunu gösterir. Pozitif b değeri alan maddeler ise tercih edilmeyen maddelerdir (Ellis, Minsel ve Becker, 1989). Ölçekten elde edilen maddelerin tamamının (-) negatif değer almış olması bu maddelerin tamamının tercih edilen maddeler olduğunun bir göstergesi olarak kabul edilebilir.

KTK’ndaki madde analizi tekniğiyle elde edilen maddelerin ölçtüğü ortalama tutum düzeyi ile ATM’yle kestirilen b değeri arasında -0.81 korelasyon bulunmuştur. Diğer bir deyişle her iki yöntemle elde edilen maddelerin ölçtükleri tutum düzeyi değerleri arasında ters yönde yüksek düzeyde bir ilişki bulunmaktadır. Elde edilen bulgular Başarır’ın (1997) ve Nartgün’ün (2002) yapmış olduğu çalışmanın bulgularıyla benzer niteliktedir. Başarır (1997) çalışmasında KTK’ndaki madde analizi tekniği ile elde edilen maddelerin ölçtüğü ortalama tutum düzeyi ile ATM ile kestirilen b değerleri arasında -.94’lük korelasyon değeri bulmuştur. KTK’da maddelerin ölçtüğü tutum düzeyi parametresinin hesaplanması sırasında madde puanlarının aritmetik ortalaması alınmıştır. Bu çalışma kapsamında bulunan korelasyon değeri Başarır’ın (1997) ve Nartgün’ün (2002) çalışmalarında bulunduğu korelasyon değerinden daha düşüktür. Bunun nedeni şu şekilde açıklanabilir; Beden Eğitimi’ne yönelik tutumların ölçülmesine ilişkin ölçeğin uygulandığı grubun ortalama tutum düzeyleri Başarır’ın (1997) uygulama yaptığı grubun ortalama tutum düzeyinden daha yüksektir. Ayrıca, Başarır’ın (1997) çalışmasında nihai ölçeğin uygulanması sonucu elde edilen b parametrelerinin bir kısmı pozitif değer almıştır. Bu çalışmada ise b değerlerinin tamamı negatif değer almıştır. KTK’ında madde parametreleri (madde ayırtedicilik değerleri ve ölçek değerleri) örneklem grubuna bağlı olduğundan (sample dependency) tutum düzeyleri yüksek olan bu grubun ortalama tutum düzeyleri ile ATK ile elde edilen b değerleri arasında diğer çalışmalarda olduğu gibi negatif yönde bir ilişki ortaya çıkmıştır. Bulunan korelasyon katsayısı Fisher’in z istatistiği ile test edildiğinde 0.01 düzeyinde anlamlı bulunmuştur. Elde edilen bu

bulgular, ölçekte yer alan maddelerin KTK ve ÖÖK kapsamında kestirilen değerlerinin birbirleri ile benzer olduğunu şeklinde yorumlanabilir.

Asıl ölçeğin uygulanması ile elde edilen verilerden kestirilen madde parametrelerinin, tutum düzeylerinin farklılaşmasına paralel olarak değişim gösterip göstermediğinin belirlenmesi için, 30 maddelik BETÖ'nin 426 öğrenciye uygulanması ile elde edilen verilerden kızlar, erkekler, alt grup ve üst grup olmak üzere dört ayrı grup oluşturulmuştur. Alt ve üst gruplar oluşturulurken toplam puanlar üzerinden alt ve üst %27'lik puan dağılımları alınmıştır. Daha sonra, dört ayrı grubun a ve b parametrelerinin ölçeğin tümünden elde edilen verilerle kestirilmiş a ve b parametreleri ile ilişkisi incelenmiştir.

Farklı grupların verilerinden elde edilen a parametreleri ve b parametreleri ortalamaları Çizelge 16'da verilmiştir.

Çizelge 16 incelendiğinde farklı gruplardan elde edilen b değerlerinin de tamamının eksi (-) değer aldığı, yani tercih edilen madde olma özelliklerini farklı gruplar için de korudukları görülmektedir.

Tablo 16'daki kızlar, erkekler, alt grup ve üst grup olarak belirlenen farklı grupların a değerlerine ilişkin aritmetik ortalamalar sırasıyla kızlar için 2.04, erkekler için 1.99, üst grup için 1.99 ve alt grup için 1.82 olarak bulunmuştur. Tüm grup için kestirilen a değerlerine ilişkin aritmetik ortalama ise 2.13 olarak bulunmuştur.

**Çizelge 16. Farklı Grupların Maddelerinden Aşamalı Tepki Modeli Yardımıyla Kestirilen a Değerleri ve b Değerleri Ortalamaları**

| Madde No | Kızlar a Değerleri | Kızlar b Ortalamalar | Erkekler a Değerleri | Erkekler b Ortalamalar | Üst Grup a Değerleri | Üst Grup b Ortalamalar | Alt Grup a Değerleri | Alt Grup b Ortalamalar |
|----------|--------------------|----------------------|----------------------|------------------------|----------------------|------------------------|----------------------|------------------------|
| 1        | 1.79               | -0.82                | 1.69                 | -1.04                  | 1.86                 | -1.50                  | 1.90                 | -1.00                  |
| 2        | 1.69               | -0.86                | 1.80                 | -0.71                  | 1.78                 | -0.66                  | 1.41                 | -0.55                  |
| 3        | 2.44               | -0.79                | 2.02                 | -1.13                  | 2.43                 | -0.53                  | 2.01                 | -0.76                  |
| 4        | 1.70               | -0.92                | 1.65                 | -0.88                  | 1.59                 | -1.18                  | 1.29                 | -0.9                   |
| 5        | 1.40               | -0.68                | 1.21                 | -0.74                  | 1.34                 | -0.94                  | 1.78                 | -0.52                  |
| 6        | 3.02               | -1.06                | 2.84                 | -0.89                  | 3.65                 | -0.69                  | 3.14                 | -0.88                  |
| 7        | 1.51               | -0.48                | 1.51                 | -0.38                  | 1.65                 | -0.46                  | 1.95                 | -0.16                  |
| 8        | 1.91               | -0.83                | 2.36                 | -0.87                  | 2.58                 | -0.76                  | 2.05                 | -0.89                  |
| 9        | 2.16               | -0.85                | 2.29                 | -1.02                  | 1.36                 | -1.51                  | 1.92                 | -0.91                  |
| 10       | 1.51               | -0.57                | 1.97                 | -0.22                  | 0.59                 | -0.75                  | 1.99                 | -0.04                  |
| 11       | 2.44               | -0.70                | 2.29                 | -0.73                  | 3.36                 | -0.74                  | 2.16                 | -0.58                  |
| 12       | 3.01               | -0.56                | 2.25                 | -0.48                  | 3.32                 | -0.42                  | 2.94                 | -0.31                  |
| 13       | 1.44               | -1.25                | 1.93                 | -1.29                  | 4.90                 | -1.03                  | 1.98                 | -2.25                  |
| 14       | 1.86               | -0.64                | 2.23                 | -0.23                  | 1.78                 | -0.39                  | 1.03                 | -0.12                  |
| 15       | 1.74               | -0.44                | 2.22                 | -0.36                  | 3.24                 | -0.61                  | 2.13                 | -0.05                  |
| 16       | 1.57               | -0.45                | 1.94                 | -0.39                  | 1.89                 | -0.17                  | 1.08                 | -0.13                  |
| 17       | 2.12               | -0.77                | 2.14                 | -0.73                  | 2.73                 | -1.11                  | 2.62                 | -0.77                  |
| 18       | 3.02               | -0.67                | 2.22                 | -0.98                  | 1.49                 | -1.02                  | 1.98                 | -0.77                  |
| 19       | 1.89               | -0.30                | 2.32                 | -0.36                  | 1.10                 | -0.42                  | 1.95                 | -0.42                  |
| 20       | 2.38               | -0.54                | 1.98                 | -0.81                  | 1.88                 | -0.67                  | 1.35                 | -0.44                  |
| 21       | 2.04               | -0.45                | 2.03                 | -0.11                  | 1.58                 | -0.49                  | 1.94                 | -0.47                  |
| 22       | 2.22               | -0.79                | 2.20                 | -0.78                  | 0.89                 | -1.31                  | 1.4                  | -0.61                  |
| 23       | 1.69               | -0.21                | 1.68                 | -0.08                  | 0.94                 | -0.13                  | 1.93                 | -1.45                  |
| 24       | 2.11               | -0.85                | 1.88                 | -0.96                  | 2.10                 | -0.69                  | 1.90                 | -1.10                  |
| 25       | 1.84               | -1.94                | 1.32                 | -1.24                  | 1.20                 | -1.61                  | 1.34                 | -2.24                  |
| 26       | 1.63               | -0.61                | 1.85                 | -0.19                  | 0.76                 | -0.93                  | 1.41                 | -1.32                  |
| 27       | 2.02               | -0.63                | 1.98                 | -0.43                  | 2.19                 | -0.92                  | 1.72                 | -0.20                  |
| 28       | 2.24               | -0.62                | 1.99                 | -0.81                  | 1.94                 | -0.68                  | 1.18                 | -0.52                  |
| 29       | 2.07               | -0.32                | 1.77                 | -0.38                  | 0.96                 | -0.43                  | 1.05                 | -0.14                  |
| 30       | 2.98               | -0.86                | 2.4                  | -0.83                  | 1.51                 | -1.07                  | 2.25                 | -0.83                  |

Farklı gruplardan ÖÖK ile elde edilen a değerleri arasındaki ilişkinin belirlenmesi için hesaplanan korelasyon katsayıları Çizelge 17'de verilmiştir.



**Çizelge 17. Farklı Grupların Maddelerinden Aşamalı Tepki Modeli Yardımıyla Kestirilen a Değerleri Arasındaki Korelasyon Değerleri**

|         | a TOPAM | a ÜST   | a ALT   | a ERKEK |
|---------|---------|---------|---------|---------|
| a ÜST   | 0.598*  |         |         |         |
| a ALT   | 0.669** | 0.435** |         |         |
| a ERKEK | 0.774** | 0.345   | 0.544** |         |
| a KIZ   | 0.687** | 0.292   | 0.333   | 0.670** |

\*\*p < 0.01, \*p < 0.05

Çizelge 17 incelendiğinde farklı gruplardan elde edilen a değerleri ile tüm gruptan elde edilen a değerleri arasında, üst gruptan elde edilen a değerlerinin diğer gruplara oranla daha düşük düzeyde anlamlı bir ilişki ortaya koyduğu bulunmuştur. Üst gruptan elde edilen a değerleri ile tüm gruptan elde edilen a değerleri arasında da pozitif ve anlamlı ilişki bulunmuştur. Ancak bu ilişki diğer gruplardan elde edilen a değerleri kadar yüksek düzeyde bir ilişki değildir. Bu sonuç MTK ile seçilen maddelerin farklı gruplar için de benzer a parametreleri gösterdiği ancak ölçekte yer alan maddelerin madde ayırt edicilik gücü değerleri bakımından alt örneklemelerin tüm gruptan farklılaşması ölçüsünde değişmezlik özelliğini kaybettiği şeklinde yorumlanabilir. Bu sonuçlar Fan (1998), Demirtaşlı (2002), Nartgün (2002), Doğan ve Tezbaşaran'ın (2003), Weiberg (2004) ve Stage'in (1997) çalışmalarında elde ettiği bulgularla paralellik göstermektedir. Bu çalışmalarda da madde ayırt edicilik gücü değerlerinin değişmezliğinin özellikle alt ve üst örneklem gruplarında sağlanamadığına ilişkin bulgular elde edilmiştir.

Farklı gruplardan ATM ile elde edilen b değerleri arasındaki ilişkinin belirlenmesi için hesaplanan korelasyon katsayıları Çizelge 18'de verilmiştir.

**Çizelge 18. Farklı Grupların Maddelerinden Aşamalı Tepki Modeli Yardımıyla Kestirilen b Değerleri Arasındaki Korelasyon Değerleri**

|         | b TOPLAM | b UST   | b ALT   | b KIZ   |
|---------|----------|---------|---------|---------|
| b UST   | 0.626**  |         |         |         |
| b ALT   | 0.596**  | 0.574** |         |         |
| b KIZ   | 0.896**  | 0.712** | 0.640** |         |
| b ERKEK | 0.961**  | 0.632** | 0.592** | 0.796** |

\*\*p < 0.01

Çizelge 18 incelendiğinde farklı gruplardan elde edilen b değerleri ile tüm gruptan elde edilen b değerleri arasında anlamlı bir ilişki bulunmuştur. Bulunan b değerlerinin tamamının yüksek düzeyde bir korelasyon katsayısı vermediği görülmüştür. Bu sonuç ÖÖK ile seçilen maddelerin farklı gruplar için de benzer b parametreleri gösterdiği, ancak, ölçekte yer alan maddelerin ölçtükleri tutum düzeyleri bakımından, örneklem değişikliklerine paralel olarak değiştiği şeklinde yorumlanabilir. Elde edilen bulgularda alt gruplardan elde edilen b değerlerinin farklılaşması, KTK'nda madde parametrelerinin örneklem grubuna bağlı olmasından kaynaklandığı, kuramsal temeline dayalı olarak yorumlanabilir (Hambleton ve Jones, 2003). Ayrıca, kız ve erkek öğrencilerine ait b değerlerinin de yüksek ve anlamlı olarak bulunması genellikle kız çocukların beden eğitimi dersine yönelik tutumlarının erkek çocuklarına göre daha olumsuz olduğuna yönelik araştırma bulgularının (Portman, 2003; Silverman ve Swaminathan, 2005) aksine araştırmaya katılan kız öğrenci grubunun beden eğitimi dersine yönelik tutumlarının erkek öğrencilere yakın olması ile açıklanabilir. Çalışma sonucu elde edilen b değerlerine ilişkin bulgularda her iki kuram açısından da anlamlı ilişkilerin bulunması KTK ve ÖÖK açısından benzer sonuçların ortaya koyduğu şeklinde yorumlanabilir. Bu bulgular Kan (2006), Tezbaşaran ve Kelecioğlu (2004), Nartgün (2002), Başarır (1997) ve Fan'ın (1998) yapmış olduğu çalışmaların sonuçları ile paralellik göstermektedir. Stage (1997)'in yaptığı çalışmasının sonuçlarına göre, özellikle farklı karakteristik yapıları bulunan (otantik) ve büyük örneklem grupları üzerinde yürütülen bu tür çalışmalarda yüksek madde değişmezliği değerlerine ulaşmanın ve belirgin bir ÖÖK'na yönelik avantaj bulguları elde etmenin zorlaştığı şeklinde bulunmuştur.

Çalışma sonucunda oluşturulan ölçeğin uygulanması ile maddelere ilişkin kestirilen madde ayırıcılık gücü indeksleri ve madde güçlük indekslerinin birbirleri ile benzer bir yapı oluşturduğu yani benzer sonuçlar ortaya koyduğu görülmektedir. Bunun dışında madde ayırıcılık gücü değerlerine ilişkin korelasyonların madde güçlük değerlerine ilişkin korelasyonlardan göreceli olarak biraz daha düşük olduğu şeklinde yorumlanabilir. Bu sonuçlar Kan (2006), Tezbaşaran ve Kelecioğlu (2004),

Tezbaşaran (2003), Nartgün (2002), Erden (1997) ve Fan'ın (1998) elde ettiği araştırma bulgularını destekler niteliktedir. İlgili çalışmalarda test maddelerine ait KTK'na ilişkin  $p$  ve  $r_{jx}$  istatistikleri ile ÖÖK kapsamında  $a$  ve  $b$  parametreleri arasındaki ilişkiler incelenmiş ve yüksek korelasyonlar bulunmuştur. Bu sonuçlar ışığında, madde istatistikleri ve parametreleri açısından uygulamada elde edilen sonuçlar büyük oranda benzerlik göstermektedir.

**Klasik Test Kuramı ve Örtük Özellikler Kuramına dayalı olarak geliştirilen Beden Eğitimi dersine yönelik tutum ölçeğinin ölçüte dayalı geçerlik ve yapı geçerliğinin belirlenmesi.**

Likert tipi tutum ölçeklerinin geçerliklerinin belirlenmesi için kullanılan yöntemler yapı, kapsam ve ölçüte dayalı geçerliklerdir (Tezbaşaran, 1997). Ölçeğin yapı geçerliğinin sınanmasında ölçülmek istenen tutum ile ilgili kuramsal yapı tanımlanır ve bu yapıya ilişkin sınanabilir denenceler çıkartılır. Çıkartılan bu denenceler üzerinde deneysel ve istatistiksel çalışmalar yürütülür (Tekin, 1993). Ölçeğin boyutlarının belirlenmesi amacı ile faktör analizi uygulanır ve ölçme aracının esas olarak alınan kuramsal temele ilişkin hangi boyutlarda ölçme yaptığı ortaya çıkarılabilir (Tezbaşaran, 1997). Bu çalışmaların yanında yapı geçerliğinin sınanması aşamasında ölçek maddelerinin alandaki uzmanlar tarafından incelenmesi, cevaplayıcıların tepki sürelerinin incelenmesi ve ölçek puanlarının kararlılık düzeylerinin saptanmasında kullanılan yöntemler de kullanılabilir.

Bu çalışmada Beden Eğitimi dersine yönelik tutumları ölçmeye yönelik olarak geliştirilen ölçeğin 12-15 yaş aralığındaki ilköğretim öğrencilerine yönelik olarak geliştirilmesi nedeni ile ölçeğin geçerliğinin belirlenmesinde ölçüte dayalı geçerlik ve yapı geçerliği yöntemleri kullanılmıştır.

KTK kapsamındaki dereceleme toplamlarıyla ölçekleme tekniğine dayalı olarak geliştirilen Beden Eğitimi dersine yönelik tutum ölçeğinin ölçüt grup geçerliğini belirlemek amacıyla Beden Eğitimi dersine yönelik olumlu tutuma sahip olduğu kabul edilen (Kangalgil, Hünük, ve Demirhan, 2003)

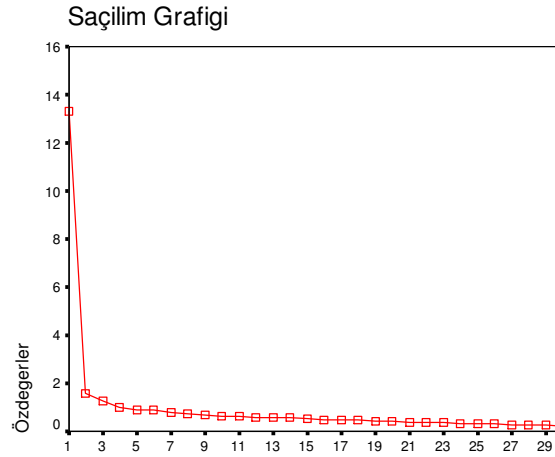
118 spor lisanslı ilköğretim öğrencisinin ham puanlarının aritmetik ortalaması ile Beden Eğitimi dersine karşı Lisanslı sporculara göre olumsuz tutuma sahip olduğu düşünülen 163 ilköğretim öğrencisinin aynı ölçekten elde ettiği ham puanların aritmetik ortalaması arasında anlamlı bir farklılık olup olmadığı bağımsız gruplara uygulanan "t" istatistiği ile test edilmiş ve 279 serbestlik derecesi ve 0.01 manidarlık düzeyinde t değeri anlamlı bulunmuştur ( $t=7.311$ ,  $p<.01$ ). İki grubun tutum puanları ortalamaları arasında anlamlı bir fark vardır. Spor lisanslı öğrencilerin ortalama tutum puanları 141.35 ve puan dağılımına ilişkin standart sapma değeri 3.61 olarak bulunurken spor lisanslı olmayan öğrencilerin ortalama tutum puanları 129.12 ve puan dağılımına ilişkin standart sapma değeri 17.91 olarak bulunmuştur. Bu sonuçtan ölçeğin önceden farklı tutumlara sahip olduğu kabul edilen grupları ayırt ederek Beden Eğitimi dersine yönelik tutumu ölçmede geçerliliğe sahip olduğu görülmüştür.

Benzer şekilde iki farklı gruptan elde edilen verilerden yararlanılarak aşamalı tepki modeliyle bireylerin tutum puanları kestirilmiştir. Her bir gruptaki bireylerin kestirilen tutum düzeylerinin aritmetik ortalamaları arasında anlamlı bir fark olup olmadığı "t" istatistiği ile incelenmiştir. 279 serbestlik derecesi ve 0.01 manidarlık düzeyinde t değeri anlamlı bulunmuştur ( $t=5.083$ ,  $p<.001$ ). Spor lisanslı öğrencilerin ortalama tutum puanları ATM'nde 1.69 ve puan dağılımına ilişkin standart sapma değeri 0.69 olarak bulunurken spor lisanslı olmayan öğrencilerin ortalama tutum puanları 1.11 ve puan dağılımına ilişkin standart sapma değeri 0.94 olarak bulunmuştur. Farklı tutuma sahip olduğu kabul edilen iki öğrenci grubunun kestirilen tutum puanları ortalamaları arasında fark bulunmuştur.

Elde edilen otuz maddelik tutum ölçeğinin yapı geçerliğinin belirlenmesi amacıyla Temel Bileşenler Analizi uygulanmıştır. Bu aşamanın ilk adımı olarak ölçme aracının uygulanması sonucu elde edilen verilerin faktörleşmeye uygunluğu incelenmiş ve 0.01 manidarlık düzeyinde KMO (Kaiser-Meyer-Olkin) değeri 0.960, Bartlett testi değeri 7043.994 olarak bulunmuştur. Temel Bileşenler Analizi uygulamasında ölçme aracının tek

boyutlu olup olmadığının sınanmasında saçılım grafiğinden de yararlanılmıştır. Şekilde dikey eksen özdeğer miktarlarını, yatay eksen ise faktörleri gösterir. Şekil faktörlerin öz değerleriyle eşleştirilmesi sonucunda bulunan noktaların birleştirilmesi ile elde edilir. Şekilde yüksek ivmeli, hızlı düşüşlerin yaşandığı faktör önemli faktör sayısını verir. Yatay çizgiler faktörlerin getirdikleri ek varyansların katkılarının birbirine yakın olduğunu gösterir (Büyüköztürk, 2002: 479).

Temel Bileşenler Analizi sonucunda elde edilen saçılım grafiği Şekil 7'de verilmiştir.



**Şekil 7.** Temel Bileşenler Analizi Sonucu Elde Edilen Grafik

Saçılım grafiği incelendiğinde birinci faktörden sonra keskin bir düşüşün olduğu görülmektedir. Bu düşüş ölçme aracının tek bir boyuta sahip olduğunun bir göstergesi olarak kabul edilebilir. Yapılan Temel Bileşenler Analizi sonucunda elde edilen üç faktörün açıkladığı toplam varyans %53.99 olarak bulunmuştur. Analiz sonucunda birinci faktörün toplam varyansın %18.86'sını, ikinci faktörün %17.57'sini ve üçüncü faktörün de %17.55'ini açıkladığı görülmüştür.

Elde edilen verilerden her iki yöntemle madde seçimi yapılarak ortak otuz tutum ifadesinden oluşan BE dersine yönelik tutum ölçeğinin geçerli olduğu sonucuna ulaşılmıştır.

Ölçekte yer alan maddelerin dağılımları ve faktör yükleri Çizelge 19'da verilmiştir.

**Çizelge 19. Ölçekte Yer Alan Maddelerin Dağılımları ve Faktör Yükleri**

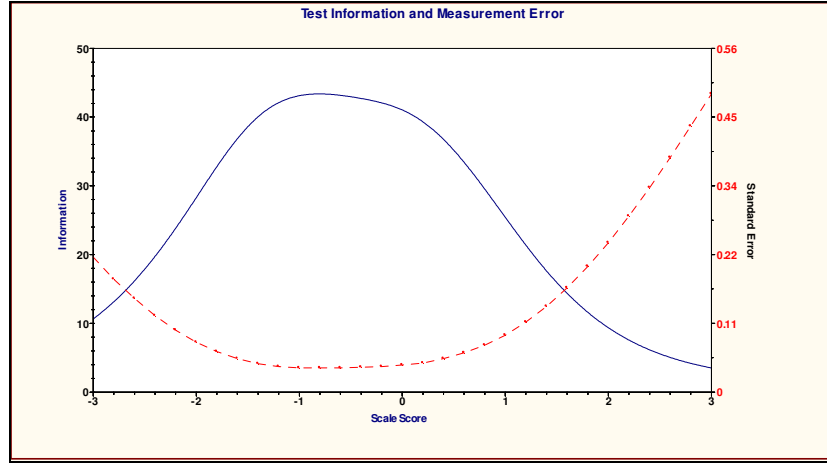
| Alt Başlıklar | Madde Sayısı | Faktör Yükleri  | $\alpha$ |
|---------------|--------------|-----------------|----------|
| Bilişsel      | 11           | 0.76-0.49 arası | 0.91     |
| Duyuşsal      | 9            | 0.65-0.46 arası | 0.88     |
| Devimsel      | 10           | 0.69-0.49 arası | 0.87     |

Yapılan çalışmalar sonucu elde edilen BETÖ'nin otuz maddesi bulunmaktadır. Bu maddelerin onbeşi olumlu, onbeşi ise olumsuz tutum ifadelerinden oluşmaktadır. Ölçekten alınabilecek en düşük puan otuz en yüksek puan ise yüzelli'dir.

#### **Likert Tipi Beden Eğitimi Dersine Yönelik Tutum Ölçeğinin Güvenirliğinin Belirlenmesi**

Bu çalışmada Beden Eğitimi dersine yönelik tutumu ölçmek amacıyla geliştirilen KTK kapsamındaki dereceleme toplamıyla ölçekleme tekniği ve ÖÖK kapsamında ATM'yle madde seçimi yapıp ortak otuz tutum cümlesinden oluşan nihai ölçeğin 426 ilköğretim öğrencisine uygulanması sonucu elde edilen ham puanlarından hesaplanan Cronbach  $\alpha$  katsayısı .937 olarak bulunmuştur. Elde edilen  $\alpha$  katsayısı ile ölçeğin kendi içinde tutarlı ölçme yaptığı ve ölçeğin tek bir özelliği ölçtüğü sonucuna varılmıştır.

Aşamalı tepki modelinde her bir tutum seviyesinde ölçekle yapılan ölçmenin doğruluğu hakkında bilgi veren test bilgi fonksiyonu değerleri Çizelge 20'de verilmiştir. ATM'yle her bir tutum düzeyinde kestirilen test bilgi fonksiyonu grafiği Şekil 8'de verilmiştir.



**Şekil 8.** Aşamalı Tepki Modeliyle Her Bir Tutum Düzeyinde Kestirilen Test Bilgi Fonksiyonu Grafiği

**Çizelge 20.** Aşamalı Tepki Modeliyle Her bir Tutum Düzeyinde Kestirilen Test Bilgi Fonksiyonu Değeri ve Standart Hataları

| $\theta$ | Test Bilgi Fonksiyonu Değeri | Standart Hata |
|----------|------------------------------|---------------|
| -3.0     | 10.61                        | .31           |
| -1.6     | 13.12                        | .27           |
| -1.4     | 40.03                        | .16           |
| 0        | 42.08                        | .15           |
| 0.2      | 39.31                        | .16           |
| 1.6      | 36.75                        | .16           |
| 1.8      | 11,59                        | .29           |
| 3.0      | 9.35                         | .32           |

Çizelge 20'deki değerler ve Şekil 8'deki test bilgi fonksiyonu dağılımı incelendiğinde ölçeğin -1.4 ve +1.6 tutum düzeyi aralığında bireylerin tutum düzeylerini daha az hatayla ölçtüğü sonucuna varılabilir. Bu aralıkta bulunan maddelerin test bilgi fonksiyonu değerleri yüksek ve standart hata değerleri daha düşük çıkmıştır. Test bilgi fonksiyonu değerleri yüksek olan maddelerin ayırt edicilik değerleri de yüksek olmaktadır (Koch, 1983; Hambleton, Swaminathan and Rogers, 1991:91).

Maddelerin ayırt edicilik değeri ile madde ve test bilgi fonksiyonu arasında pozitif ilişki bulunmaktadır (Koch, 1983). Ölçek maddelerin kestirilen ayırt edicilik değerlerin Çizelge 13'te görüldüğü gibi 1.34-3.65

arasında deęiřtięi deęerlendirildięinde, maddelerin bireylerin tutum dzeyleri hakkında daha fazla ve aıklayıcı bilgiler verdięi ve lekte yapılan lmenin daha az hatalı gerekleřmesine olanak saęladıęı sonucuna varılmıřtır.

K'nda standart hata, bir cevaplayıcı grubundan kestirilen parametre deęerinin barındırdıęı eksik bilgi olarak kabul edilmekte, bunun tersi ise test bilgi fonksiyonu ile tanımlanmaktadır.  $\theta$  yetenek dzeyi zerindeki "standart hata" ile "bilgi" arasındaki iliřki K kapsamındaki gvenirlięin en yakın tahminin elde edilmesinde kullanılmakta, kestirilen deęer "Marjinal Gvenirlik" olarak adlandırılmakta ve bu deęer test bilgi fonksiyonu ve standart hataya dayalı olarak farklı  $\theta$  dzeyleri iin kestirilen bir ortalama olarak tanımlanmaktadır (Thissen, 1991). Bu arařtırma kapsamında ATM'ne dayalı olarak gvenirlik kestirimi iin marjinal gvenirlik katsayısından yararlanılmıřtır.

lęin ATM'yle kestirilen marjinal gvenirlięi 0.96'dır. Btn tutum dzeylerindeki ortalama gvenirlik katsayısı olan bu deęer lkle yapılan lmenin gvenirlięinin yksek olduęunu gstermektedir.

Benzer gvenirlik sonuları elen (2008), Kan (2006), Tezbařaran ve Kelecioęlu (2004), Kelecioęlu (2004), Nartgn (2002), Anıl (2002), Bařarır (1997) ve Stage'in (1997) yapmıř olduęu alıřmalarda da elde edilmiřtir. Bu alıřmalarda da arařtırma amacı dahilinde geliřtirilen lme aralarından KTK ve K erevesinde elde edilen Cronbach Alpha ve Marjinal Gvenirlik deęerlerinin olduka yksek ve birbirine yakın deęerler verdięi sonucuna ulařılmıřtır. zellikle her iki kuramın sayıltılarını saęlayan maddelerde bu deęerlerin ok daha yksek olduęu grlmektedir. Deneme uygulaması sonucunda KTK ve K dahilinde belirlenen  farklı yntem ile yapılan madde analizleri sonucu otuz maddeden yirmisekizinin ortak olduęu yani ortak madde oranının yaklařık %94 olduęu belirlenmiřtir. Bu sonularla, lme aralarının geliřtirilmesinde her iki kuramın birlikte kullanılmasının zellikle leklerin gvenirlikleri aısından nemli bir katkı saęladıęı řeklinde yorumlanabilir.



## BÖLÜM V

### SONUÇ VE ÖNERİLER

Araştırmanın bu bölümünde, araştırma verilerinin çözümlenmesiyle elde edilen bulgulara dayalı sonuçlara yer verilmiş ve öneriler sunulmuştur.

#### Sonuçlar

Bu çalışmanın amacı, KTK ve ÖÖK'na göre psikometrik özelliklerin, İlköğretim 6-7-8. sınıf öğrencilerinin (12-15 yaş) Beden Eğitimi dersine yönelik tutumlarını ölçen, Likert tipi bir ölçek hazırlama çalışması kapsamında incelenmesi ve karşılaştırılmasıdır. Çalışmanın amacı dahilinde Beden Eğitimi dersine yönelik tutumları ölçen tutum ölçeğinin üç farklı yöntem ile hazırlanmasıyla elde edilen ölçeklerden her iki kuram ile madde ve analizlerin yapılması ve psikometrik özelliklerin incelenmesiyle aşağıda belirtilen sonuçlara ulaşılmıştır.

1. Ölçeğin deneme formundan korelasyona dayalı analiz, iç tutarlılık ölçütüne dayalı analiz ve ATM'ne dayalı olarak kestirilen madde ayırtedicilik değerleri arasındaki ilişkinin incelenmesi amacıyla korelasyon katsayısı hesaplanmıştır. Korelasyona dayalı analiz yöntemi ile seçilen maddelerden elde edilen madde ayırtedicilik değerleri ile iç tutarlılık yöntemi ile seçilen maddelerden elde edilen madde ayırtedicilik değerleri arasında yüksek düzeyde pozitif ve anlamlı bir ilişki bulunmuştur. Korelasyona dayalı analiz yöntemi ile seçilen maddelerden elde edilen madde ayırtedicilik değerleri ile ATM ile elde edilen madde ayırtedicilik değerleri arasında da pozitif ve anlamlı ilişki bulunmuştur. İç tutarlılık ölçütüne dayalı analiz yöntemi ile seçilen maddelerden elde edilen madde ayırtedicilik gücü değerleri ile ATM ile seçilen maddelerden elde edilen madde ayırtedicilik gücü değerleri arasında diğerlerine göre biraz daha düşük olmakla beraber pozitif ve anlamlı bir ilişki gözlenmiştir.

2. KTK ve ÖÖK'nın ATM kullanılarak beden eğitimi dersine yönelik tutumların ölçülmesi için yapılan analizler sonucunda her iki kuramın sayıltılarını karşılayan ortak maddeler seçilerek otuz maddelik beden eğitimi

dersine yönelik tutumları ölçen bir ölçek oluşturulmuştur. Her iki kuramın sayıltılarını karşılayan maddelerin yaklaşık %94'ünün ortak olduğu görülmüştür.

3. Korelasyona dayalı analiz yöntemi, iç tutarlılık ölçütüne dayalı analiz yöntemi ve ATM'ne dayalı analiz yöntemleri ile seçilen maddelerle oluşturulan üç farklı ölçeğin güvenilirlik katsayılarının incelenmesi amacıyla, bu üç yöntemle hazırlanan ölçeklerin Cronbach- $\alpha$  ve Marjinal Güvenirlik değerleri kestirilmiştir. Ölçeğin deneme formunun ve üç farklı yöntemle seçilen maddelerle oluşturulan ölçeklerin Alpha ve marjinal güvenirlikleri yüksek bulunmuştur. Buna ek olarak, deneme ölçeği ve diğer ölçeklerin Cronbach- $\alpha$  ve marjinal güvenirlikleri arasında da önemli sayılabilecek bir farklılık bulunmamıştır.

4. Korelasyona dayalı analiz yöntemi, iç tutarlılık ölçütüne dayalı analiz yöntemi ve ATM'ne dayalı analiz yöntemleri ile seçilen maddelerle oluşturulan üç farklı ölçeğin Test Bilgi Fonksiyonları incelenmiş ve ATM ile geliştirilen ölçeğin diğer iki yöntem ile geliştirilmiş ölçeklerden daha fazla bilgi verdiği görülmüştür.

5. Beden Eğitimi dersine yönelik tutumu ölçmek amacıyla hazırlanan Likert tipi tutum ölçeğinin uygulanmasıyla elde edilen verilere dayanılarak, KTK'ndaki madde analizi tekniğiyle elde edilen madde ayırt edicilik değeri ve maddenin ölçtüğü ortalama tutum düzeyi değerleri ile ATM'yle kestirilen maddelerin (a) ve (b) parametreleri arasında ilişki olup olmadığı incelenmiştir. KTK kapsamındaki madde analiziyle elde edilen madde ölçek korelasyonları ile ATM ile kestirilen madde ayırt edicilik değerleri (a) arasında pozitif ve anlamlı ilişki bulunmuştur.

6. KTK kapsamındaki madde analizi tekniği ile elde edilen maddenin ölçtüğü ortalama tutum düzeyi değeri ile ATM'yle kestirilen (b) parametreleri arasında negatif ve anlamlı ilişki bulunmuştur.

7. Beden Eğitimi dersine yönelik tutumların ölçülmesinde kullanılan ölçekte yer alan maddelerin KTK ve ÖÖK bakımından değişmezlik özelliğini kaybedip kaybetmediğinin belirlenmesi amacıyla ölçeğin uygulandığı 426

kişilik öğrenci grubundan üst grup, alt grup , kız öğrenciler ve erkek öğrenciler olarak dört ayrı alt çalışma grubu oluşturulmuştur. Bu gruplar için ve tüm grup için a ve b parametreleri hesaplanmıştır. Kestirilen bu parametreler arasında pozitif ve anlamlı bir ilişki bulunmuştur. Ancak ilişki düzeyinin alt gruplar ve üst gruplar için daha düşük olduğu görülmüştür.

8. Her iki yöntemle kestirilen  $\alpha$  katsayılarından ölçeğin güvenilir ölçmeler yaptığı sonucuna varılmıştır. Buna ek olarak ATM'yle kestirilen her bir yetenek seviyesindeki test bilgi fonksiyonlarının aritmetik ortalamalarıyla marjinal güvenilirlik değeri elde edilmiştir. Ölçeğin marjinal güvenilirliği de yüksek bulunmuştur. Elde edilen güvenilirlik katsayılarının birbirine yakın değerler olduğu görülmüştür.

9. KTK kapsamındaki dereceleme toplamlarıyla ölçekleme tekniği ile ATM'ne dayalı oluşturulan ölçeklerin madde parametreleri ve test istatistikleri birbirine yakın sonuçlar vermektedir.

10. ÖÖK kapsamında kestirilen madde parametreleri araştırma grubundan, bireylerin kestirilen tutum değerlerinin örneklemden bağımsız olma özelliği göz önünde bulundurularak ATM'nin tutum ölçeklerinin geliştirilmesi ve analizlerde kullanılabileceği sonucuna ulaşılmıştır.

11. Psikometrik özellikler incelendiğinde geliştirilen ölçeğin hem KTK hem de ÖÖK açısından önemli değerler ortaya koyduğu gözlenmiştir. Bu çalışma sonucunda KTK ve ÖÖK'nin ölçek geliştirme çalışmalarında birlikte kullanılabileceği sonucuna varılmıştır. Bu iki kuramın birlikte kullanılmasının özellikle geliştirilen tutum ölçeklerindeki geçerlik ve güvenilirlik çalışmalarının daha güçlü bir şekilde yapılacağını ve geçerlik ve güvenilirlik bakımından daha güvenilir ve güçlü bulgular elde edileceği sonucuna ulaşılmıştır. Ayrıca, çalışma bulgularına göre ATM'nin KTK kapsamında kullanılan yöntemlere göre daha fazla bilgi vermesi nedeni ile tercih edilebileceği sonucuna da ulaşılmıştır.

12. KTK ve ÖÖK çerçevesinde geliştirilen 30 maddelik BE dersine yönelik tutum ölçeğinin geçerli ve güvenilirliğinin yüksek olduğu ve bu ölçme aracının Beden Eğitimi dersine yönelik tutumların ölçülmesinde kullanılabileceği sonucuna ulaşılmıştır.

### Öneriler

1. Bu çalışmada KTK ve ÖÖK'nin psikometrik özellikleri beden eğitimi dersine yönelik bir tutum ölçeği geliştirme aşamasında incelenmiş ve karşılaştırılmıştır. Çalışmanın sonuçları her iki kuramında psikometrik özelliklerinin benzer olduğu yönünde bulgular ortaya koymuştur. Bu noktada ölçek geliştirme çalışması yapacak araştırmacılar için her iki kuramdan birinin kullanılabileceği, buna ek olarak ölçek geliştirme aşamasında madde ve test için daha çok açıklayıcı bilgiler isteyen araştırmacılara ÖÖK'ni tercih etmeleri önerilebilir.

2. Araştırma grubunun farklı tutum düzeylerindeki maddelerin kestirilen parametrelerinin değişip değişmediği incelenirken üst grup, alt grup, kızlar ve erkekler şeklinde belirlenen alt örneklerle çalışılmıştır. Ortalama puan seviyesinde olan gruplara ilişkin parametre kestirimleri ve bu parametrelerin tüm grup ve diğer belirlenen gruplardan elde edilen parametrelerle olan ilişkileri incelenmemiştir. Üst grup ve alt grup dışındaki puan dağılımlarında madde parametrelerinin üst ve alt gruplardan elde edilen parametrelerle ve tüm gruptan elde edilen madde parametreleriyle nasıl bir ilişkide olduğunun incelenmesiyle farklı tutum düzeylerindeki gruplarda madde parametrelerinin değişip değişmediğine ilişkin daha kapsamlı bilgilere ulaşılabilir.

3. Araştırmada farklı grupların madde parametreleri incelenirken sadece ÖÖK ile elde edilen parametreler kullanılmıştır. KTK çerçevesinde farklı gruplardan elde edilen madde özellikleri incelenmemiştir. KTK ile elde edilen madde parametrelerinin farklı gruplardaki yapısı ve ÖÖK'yla farklı gruplardan elde edilen madde parametreleri incelenerek iki kuram çerçevesinde madde parametrelerinin değişmezliğinin kuramsal boyutlarda karşılaştırılması mümkün olabilir.

## KAYNAKÇA

- Aicinena, S. (1991). The Teacher and Student Attitudes Towards Physical Education. *The Physical Educator*. Vol. 48, pp. 28-32
- Ajzen, I. (1993). *Attitude Theory and the Attitude-Behaviour Relation*. Hillside, NJ. Lawrence Erlbaum Associates, Inc.
- Alım, M. (2008). Lise Öğrencilerinin Coğrafya Dersine Yönelik Tutumları. *Doğu Coğrafya Dergisi*. 19, , 25-32.
- Anderson, J. R. (1983) 'A spreading activation theory of memory', *Journal of Verbal Learning and Verbal Behavior*, 22, 261-295.
- Anastasi, A. (1988). *Psychological Testing*. New York: McMillan Publishing Company.
- Andrich, D. (1982). An Index of Person Separation in Latent Trait Theory, The Traditional KR20 Index, and The Guttman Scale Response Pattern. *Education Research and Perspectives*, 9, 95-104.
- Anıl, D. (2002). *Deneme Uygulamasının Yapılmadığı Durumlarda Madde ve Test Parametrelerinin Klasik ve Örtük Özellikler Test Teorilerine Göre Kestirilmesi*. Yayınlanmamış Doktora Tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Arul, M. J. (2002). Measurement of Attitudes. <http://www.geocities.com/Athens/5503/atti2-a.html>.
- Barney, D. (2003). Factors That Impact Middle School Student's Attitudes and Perceptions in Physical Education. *Research Quarterly for Exercise and Sport*, 74 (Suppl. 1), A-36.
- Baron, A. R. ve Byrne, D. (1997). *Social Psychology*. A Viacom Company, Massachusetts.
- Başarır, D. E. (1997). *Örtük Özellikler ve Klasik Test Teorisi Yaklaşımına Dayalı Olarak Geliştirilen Likert Tipi Tutum Ölçeğinin Psikometrik Özelliklerinin Karşılaştırılması*. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü. Yayınlanmamış Doktora Tezi.

- Baykul, Y. (1979). *Örtük Özellikler ve Klasik Test Kuramları Üzerine Bir Karşılaştırma*. Yayınlanmamış Doktora Tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Baykul, Y. (2000). *Eğitimde ve Psikolojide Ölçme. Klasik Test Teorisi ve Uygulaması*. ÖSYM Yayınları, Ankara.
- Baysal, A. C. (1981). *Sosyal ve Örgütsel Psikolojide Tutumlar*. İstanbul: İstanbul Üniversitesi İşletme Fakültesi.
- Berberoğlu, G. (1988). *Seçme Amacıyla Kullanılan Testlerde Rasch Modelinin Katkıları*. Yayınlanmamış Doktora Tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Beykal, O. F. (2005). Resim-İş Dersine İlişkin Tutumların Ölçülmesi. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*.18, 26-38.
- Bompa, T. O. (2000). *Total Training for Young Champions: Proven Conditioning Programs for Athletes Ages 6 to 18*. Human Kinetics, Champaign. IL.
- Bompa, T. O. (2002). *Periodization: Theory and Methodology of Training*. Fourth Edition. Human Kinetics, Champaign, IL.
- Bowyer, G. R. (1996). Student Perceptions of Physical Education. *Journal Of Physical Education, Recreation and Dance*. 67, 23-26.
- Brophy, J. E. ve Good, T. L. (1986). Teacher Behavior and Student Achievement. *Handbook of Research on Teaching*. (3rd Edition), pp.328-375.
- Büyüköztürk, Ş. (2002). Faktör Analizi: Temel Kavramlar ve Ölçek Geliştirmede Kullanımı. *Kuram ve Uygulamada Eğitim Yönetimi*. 32: 470-483.
- Carlson, T. B. (1995). We Hate Gym: Student Alienation From Physical Education. *Journal of Teaching in Physical Education*. Vol.14, pp.467-477.

- Chung, M. and Phillips, D. A. (2002). The Relationship Between Attitude Toward Physical Education and Leisure-Time Exercise in High School Students, *Physical Educator*. No. 59, pp. 126–138.
- Cole, E., Wood, T. M. and Dunn, J. M. (1991). Item Response Theory: A Useful Test Theory for Adapted Physical Education. *Adapted Physical Activity Quarterly*, Vol. 8, No. 4. pp. 216-231.
- Crocker, L. ve Algina, J. (1986). *Introduction To Classical and Modern Test Theory*, Orlando, Harcourt Brace Jovanovich Inc.
- Çelen, Ü. (2008). Klasik Test Kuramı ve Madde Tepki Kuramı Yöntemleriyle Geliştirilen İki Testin Geçerlik ve Güvenirliğinin Karşılaştırılması. *Eğitim Online*. 7 (3). 758-768.
- Çetin, B. ve Kelecioğlu, H. (2007). Geleneksel Yöntemle ve Eleme Yöntemiyle Puanlanan Çoktan Seçmeli Testlerin Psikometrik Özelliklerinin İncelenmesi. *Eurasian Journal of Educational Research*, 29, pp, 25-34
- De Ayala, R. J. (1993). An Introduction to Polytomous Item Response Theory Models. *Measurement and Evaluation in Counselling and Development*. Vol.25, pp. 172-190.
- De Ayala, R. J., Dodd, B. G., ve Koch, W. R. (1992b). A Comparison of the Partial Credit and Graded Response Models in Computerized Adaptive Testing. *Applied Measurement in Education*. Vol. 5, pp.17-34.
- Demirhan, G. (2006). *Spor Eğitiminin Temelleri*. Spor Yayınevi ve Kitapevi-Bağırhan Yayınları, Ankara.
- Demirhan, G. ve Altay, F. (2001). Lise Birinci Sınıf Öğrencilerinin Beden Eğitimi ve Spora İlişkin Tutum Ölçeği II. *Spor Bilimleri Dergisi*. C.12. S.2. 9-20.
- Demirtaşlı, N. Ç. (2002). A Study of Raven Standard Progressive Matrices Test's Item Measures Under Classical and Item Response Models:

An Empirical Comparison. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*. s. 71-79.

- Dick, F. W. (2002). *Sports Training Principles*. Fourth Edition. A&C Black Publishers Ltd. London.
- Dooden, H. and Johanson, G. A. (2003). An Analysis of Sex-Related Differential Item Functioning in Attitude Assessment. *Assessment & Evaluation in Higher Education*, Vol. 28, No.2, 129-134.
- Doğan, N. ve Tezbaşaran, A. (2003). Klasik Test Kuramı ve Örtük Özellikler Kuramının Örneklemeler Bağlamında Karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*. s. 58-67.
- Douglass, F. M., Khavari, K. A. and Farber, P. D. (1979). A Comparison of Classical and Latent Trait Item Analysis Procedures. *Educational and Psychological Measurement*, Vol. 39, No. 2, 337-352.
- Eagly, A., & Chaiken, S. (1995). Attitude strength, attitude structure and resistance to change. In R. Petty and J. Kosnik (Eds.), *Attitude Strength*. (pp. 413-432). Mahwah, NJ: Erlbaum.
- Ellis, B. B., Minsel, B., ve Becker, P. (1989). Evaluations of Attitude Survey Translations: An Investigation using Item Response Theory. *International Journal of Psychology*. Vol. 24, pp. 665-684.
- Erdem, D. (2007). İngilizce Dersine Yönelik Bir Tutum Ölçeği Geliştirme Çalışması. *Eğitim Araştırmaları*. 28, 45-54.
- Fan, X. (1998). Item Response Theory and Classical Test Theory: an Empirical Comparison of Their Item-Person Statistics. *Educational and Psychological Measurement*. Vol.58, No.3, pp. 357-381.
- Fazio, R. H. (1986). How do attitudes guide behavior? In R. M. Sorrentino & E. T. Higgins (Eds.), *The handbook of motivation and cognition: Foundations of social behavior* (pp 204-243). New York: Guilford Press.



- Gallahue, D. L. (1996). *Developmental Physical Education for Today's Children*. (3rd Ed). Dubuque, IA: Brown & Benchmark.
- Gallahue, D. L. and Donnelly, F. C. (2003). *Developmental Physical Education for All Children (4th ed.)*, Human Kinetics, Champaign, IL (2003).
- Gelbal, S. (1994). *P Madde Güçlük İndeksi İle Rasch Modelinin b Parametresi ve Bunlara Dayalı Yetenek Ölçüleri Üzerine Bir Karşılaştırma*. Yayınlanmamış Doktora Tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Greenokle, K. M., Lee, A. A. and Lomax, R. (1990). The Relationship Between Selected Student CHARACTERISTICS AND Activity Patterns in a Required High School Physical Education Class. *Research Quarterly for exercise and Sport*. 61, 59-69.
- Greenwood, M. and Stillwell, J. (2001). Activity Preferences of Middle School Physical Education Students, *Physical Educator*. No. 58, pp. 26–29.
- Goldman, S. H. ve Nambury, S. R. (1986). Recovery of One and Two Parameter Logistic Item Parameters: An Empirical Study. *Educational and Psychological Measurement*, Vol. 46, No. 1, 11-21.
- Gözen, G. (2007). *Klasik Test ve Madde Tepki Kuramlarına Göre Çoktan Seçmeli Testlerde Farklı Puanlama Yöntemlerinin Karşılaştırılması*. Yayınlanmamış Doktora Tezi. Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Graham, G. (1995). Physical Education Through Students' Eyes and in Students' Voices. Introduction. *Journal of Teaching Physical Education*. Vol.14, pp.364-371.
- Graham, G., Holt/Hale, S. A. ve Parker, M. (1998). *Children Moving: A Reflexive Approach to Teaching Physical Education*. (4th Edition). Mountain View, CA: Mayfield.
- Hambleton, R.K. ve Swaminathan, H. (1985). *Item Response Theory: Principles and Application*, Boston, Kluwer Academic Publishers Group.

- Hambleton, R. K., Swaminathan, H. Ve Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Newbury Park, CA: Sage.
- Hambleton, R. K. ve Jones, R. W. (2003). Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. *Instructional Topics in Educational Measurement*. Vol.2. pp. 253 – 262.
- Heinen, T. (1996). *Latent Class and Discrete Latent Trait Models: Similarities and Differences*. Thousand Oaks. California: Sage.
- Hulin, C. L., Drasgow, F. and Parsons, C. K. (1983). *Item Response Theory*. Homewood, Illinois: Dow Jones-Irwin.
- Ikulayo, P. B. (1983). Attitudes of Girls Towards Physical Education. *Physical Education Review*. Vol.6, No.1, pp.24-26.
- İnceođlu, M. (1993). *Tutum Algı İletişim*. Ankara:Verso Yayıncılık.
- Kađıtçıbaşı, Ç. (1999). Yeni İnsan ve İnsanlar. 10. Baskı. Sosyal Psikoloji Dizisi: 1, İstanbul: Evrim Basım Yayım ve Dađıtım.
- Kan, A. (2006). Klasik Test Teorisine ve Örtük Özellikler Teorisine Göre Kestirilen Madde Parametrelerinin Karşılaştırılması üzerine Ampirik Bir Çalışma. *Mersin Üniversitesi Eğitim Bilimleri Dergisi*, Cilt. 2, Sayı. 2, ss. 227-235.
- Keeves, J. P. (1988). *Educational Research, Methodology and Measurement: An International Handbook*. Oxford: Pergamon Press.
- Keleciođlu, H. (2006). Item Characteristics Analysis of Attitudes Towards Smoking Scale By Log-linear Models and t Test. *Social Behaviour and Personality*, 34 (7), 853-864.
- Kline, P. (1986). *A Handbook of Test Construction*. USA: Methuen & Co.
- Koch, W. R. (1983). Likert Scaling Using the Graded Response Latent Trait Model. *Applied Psychological Measurement*. Vol.7, pp.15-32.

- Köklü, N. (1995). Tutumların Ölçülmesi ve Likert Tipi Ölçeklerde Kullanılan Seçenekler. *Ankara Üniversitesi Eğitim Bilimleri Dergisi*. Vol.28, No.2, s.81-93.
- Laidlaw, K., Power, M. J., Schmidt, S. and The WHOQOL-OLD Group. (2007). The Attitudes to Ageinf Questionnaire (AAQ): Developmnt and Psychometric Properties. *International Journal of Geriatric Pschology*, 22, 367-379.
- Lee, H. J., Kang, S. and Hume, D. (1999). Students'Attitudes Toward Physical Education in a Junior High School. *Research Quarterly for Exercise and Sport*. 70 (Suppl. 1), A-94.
- Lin, C. J. (2008). Comparisons Between Classical Test Theory and Item Response Theory in Automated Assembly of Parallel Test Forms. *The Journal of Technology, Learning, and Assessment*. Vol. 6, No. 8, pp. 3-42.
- Lord, F. M. (1980). *Applications of Item Response Theory to Practical Testing Problems*. Hillsdale, NJ: Erlbaum.
- Lord, F. ve Novick, R. M. (1968). *Statistical Theories of Mental Test Scores*, New York, Addison Wesley Publishing Company.
- Luke, M. D. and Sinclair, G. D. (1991). Gender Differences in Adolescents' Attitudes toward School Physical Education. *Journal Of Teaching Physical Education*. 11, 31-46.
- Macdonald, P. ve Paunonen, S. V. (2002). A Monte Carlo Comparison of Item and Person Statistics Based on Item Response Theory versus Classical Test Theory. *Educational and Psychological Measurement*. Vol. 62 No. 6, pp; 921-943.
- McKenzie, T. L. (2003). Health-Related Physical Education: Physical, activity fitness, and wellness. In: S.J. Silverman and C.D. Ennis, Editors, *Student Learning in Physical Education: Applying Research to Enhance Instruction*, Human Kinetics, Champaign, IL (2003), pp. 207–226.

- McKenzie, T. L., Alcaraz, J. E., & Sallis, J. F. (1994). Assessing children's liking for activity units in an elementary school physical education curriculum. *Journal of Teaching in Physical Education*, 13, 206-215.
- Middlebrook, N.P. (1974). *Social Psychology and Modern Life*. Alfred A. Knoff, New York.
- Nartgün, Z. (2002). *Aynı Tutumu Ölçmeye Yönelik Likert Tipi Ölçek İle Metrik Ölçeğin Madde ve Ölçek Özelliklerinin Klasik Test Kuramı ve Örtük Özellikler Kuramına Göre İncelenmesi*. Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Yayınlanmamış Doktora Tezi.
- Nuhoğlu, H. (2008). İlköğretim Fen ve Teknoloji Dersine Yönelik Bir Tutum Ölçeğinin Geliştirilmesi. *Eğitim Araştırmaları*. 7(3), 627-639.
- Oppenheim, A. N. (2000). *Questionnaire Design and Attitude Measurement*, Basic Book Inc. New York.
- Özbal, N. ve Çetingöz, D. (2006). Cinsiyet, Sınıf Düzeyi ve Başarı Durumlarına Göre İlköğretim İkinci Kademe Öğrencilerinin Sosyal Bilimler Dersine Yönelik Tutumları. *Çağdaş Eğitim Dergisi*. 31(327), 22-28.
- Portman, P. A. (2003). Are Physical Education Classes Encouraging Students to Be Physically Active? Experience of Ninth Graders in Their Last Semester of Required Physical Education, *Physical Educator*. Vol. 60, No. 3, pp. 150–160.
- Ryan, S., Fleming, D. and Maina, M. (2003). Attitudes of Middle School Students Toward Their Physical Education Teachers and Classes, *Physical Educator*. Vol. 60, No. 2, pp. 28–42.
- Özer, D. ve Aktop, A. (2003). İlköğretim Öğrencileri İçin Hazırlanmış Bir Beden Eğitimi Dersi Tutum Ölçeğinin Adaptasyonu. *Spor Bilimleri Dergisi*. 14, (2). s. 67-82.
- Özgüven, İ. E. (1994). *Psikolojik Testler*. Ankara: PDREM Yayınları.
- Özgüven, İ. E. (1994). *Bireyi Tanıma Teknikleri*. Ankara: PDREM Yayınları.

- Reise, S. P. ve Yu, J. (1990). Parameter Recovery in the Graded Response Model Using MULTILOG. *Journal of Educational Measurement*. Vol.27, pp.133-144.
- Samejima, A. F. (1972). A General Model for Free –Response Data. *Psychometrica Monograpy Supplement*. Vol. 37, No. 1, Part 2.
- Sanders, S., & Graham, G. (1995). Kindergarten children's initial experience in physical education: The relentless persistence for play clashes with the zone of acceptable responses. *Journal of Teaching in Physical Education*, 14, 372-383.
- Scantling, E., Strand, B., Lackey, D. and McAleese, W. (1995). An Analysis of Physical Education Avoidence. *The Physical Educator*. 52, pp. 197-201.
- Sencer, M. Ve Sencer, Y. (1978). *Toplumsal Arařtırmalarda Yönetimbilim*. Ankara: TODAİE Yayınları, No:172.
- Silverman, S. (1993). Student Characteristics, Practice, and Achievement in Physical Education. *Journal of Educational Research*. Vol.87, pp.54-61.
- Silverman, S. ve Subramaniam, P. R. (1999). Student Attitude Toward Physical Education and Physical activity: A Review of Measurement Issues and Outcomes. *Journal of Teaching Physical Education*. Vol.19, pp. 97-125.
- Silverman, S. ve Subramaniam, P. R. (2005). Middle School Students' Attitudes Toward Physical Education. *Journal of Teaching Physical Education*. Vol.21, pp. 84-101.
- Solmon, M. A., & Carter, 1. A. (1995). Kindergarten and first-grade students' perceptions of physical education in one teachers' classes. *Elementary School Journal*, 95, 335-365.
- Stage, C. (1997). A Comparison Between Item Analysis Based on Item Response Theory and Classical Test Theory: A Study of the SweSAT Subtest READ. *Educational Research*. Pp. 2-22.

- Steward, M. J., Gren, S. R. and Huelskamp, J. (1991). Secondary Student Attitudes Toward Physical Education. *The Physical Educator*. Vol. 51, No. 3. pp. 119-129.
- Strand, B. Ve Scantling, E. (1994). An Analysis of Secondary Student Preferences Towards Physical Education. *Physical Educator*. Vol. 51, No.3, pp.119-130.
- Strong, D. R., Bren, R. B. and Lejuez, C. W. (2004). Using Item Response Theory to Examine Gambig Attitudes and Beliefs. *Personality and Individual Differences*, 36, 1515-1529.
- Stocking, M. L. (1997). *Item Response Theory. Educational Research, Methodology and Measurement: An International Handbook*. Second Edition. Elsevier Science Ltd. USA.
- Subramaniam, P. R. and Silverman, S. (2000). Validation of Scores From an Instrument Assessing Student Attitude Toward Physical Education. *Measurement in Physical Education and Exercise Science*. 4 (1), 29-43.
- Tannehill, D. and Zakrajsek, D. (1993). Student Attitudes Toward Physical Education: A Multicultural Study. *Journal of Teaching Physical Education*. Vol. 13, pp. 78-84.
- Tatlıdil, H. (1992). *Uygulamalı Çok Değişkenli İstatistiksel Analiz*. Ankara
- Tavşancıl, E. (2002). *Tutumların Ölçülmesi ve SPSS ile Veri Analizi*. Ankara: Nobel Yayınevi.
- Thurstone, L. L. ve Chave, E. J. (1970). *Attitude Measurement*. Chicago: McNally and Company.
- Tekin, H. (1993). *Eğitimde Ölçme ve Değerlendirme, Gözden Geçirilmiş Sekizinci Baskı*, Yargı Yayınları, Ankara.
- Tezbaşaran, A. (1997). *Likert Tipi Ölçek Geliştirme Kılavuzu*. (2. Baskı) Ankara: Türk Psikologlar Derneği Yayınları.

- Tezbaşaran, A. ve Kelecioğlu, H. (2004). *Madde-Ölçek Korelasyonlarına, Alt-Üst Grup Ortalamalarına ve Aşamalı Tepki Modeline Göre Geliştirilen Sigaraya İlişkin Tutum Ölçeğinin Madde ve Ölçek Özelliklerinin İncelenmesi*. XIII. Ulusal Eğitim Bilimleri Kurultayı, 6-9 Temmuz 2004. İnönü Üniversitesi, Eğitim Bilimleri Fakültesi, Malatya.
- Thissen, D. (1991). *MULTILOG 6.0 User's Guide*. Scientific Software Inc. Chicago, USA.
- Tjeerdsman, B. L., Rink, J. E., & Graham, G. (1996). Student's perceptions, Values, and beliefs prior to during, and after badminton instruction. *Journal of Teaching in Physical Education*, 15, 464-476.
- Turgut, F. (1984). *Eğitimde Ölçme ve Değerlendirme Metodları*. Saydam Matbaası, Ankara.
- Waugh, R. F. (2002). Measuring Self-Reported Studying and Learning for University Students: Linking Attitudes and Behaviours on The Same Scales. *British Journal of Educational Psychology*, 72, 573-604.
- Yenilmez, K. (2007). İlköğretim Öğrencilerinin Matematik Dersine Yönelik Tutumları. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi*. 23, 51-59.
- Yıldız, Y. (1998). *Aynı Tutumu Ölçmeye Yönelik Likert Tipi Ölçek İle Metrik Ölçeğin Psikometrik Özelliklerinin Karşılaştırılması*. Yayımlanmamış Doktora Tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.

## EKLER

## Ek 1.

## BEDEN EĞİTİMİ DERSİNE YÖNELİK TUTUM ÖLÇEĞİ

Bu ölçek ilköğretim 6, 7, 8'inci sınıf öğrencilerinin Beden Eğitimi dersine yönelik tutumlarını belirlemek için hazırlanmıştır. Maddeleri cevaplayanın kimliği değil verilen cevaplar önemlidir. Lütfen ölçek üzerine ad ve soyadınızı yazmayınız. Her bir ifade için o maddenin karşısında bulunan kutulara (X) işareti koyarak **Tamamen Katılıyorum**, **Katılıyorum**, **Kararsızım**, **Katılmıyorum** veya **Hiç Katılmıyorum** seçeneklerinden birisini işaretleyiniz. Her ifade için sadece bir işaretleme yapınız.

SINIFI: CİNSİYETİ : Erkek Kız Doğum Tarihi: ...../...../.....

| ÖRNEK İŞARETLEME!           | Tamamen Katılıyorum | Katılıyorum | Kararsızım | Katılmıyorum | Hiç Katılmıyorum |
|-----------------------------|---------------------|-------------|------------|--------------|------------------|
| Müzik dersinden hoşlanırım. |                     | X           |            |              |                  |

|   | Tamamen Katılıyorum | Katılıyorum | Kararsızım | Katılmıyorum | Hiç Katılmıyorum |
|---|---------------------|-------------|------------|--------------|------------------|
| 1. Beden Eğitimi dersinden hoşlanırım.  |                     |             |            |              |                  |
| 2. Beden Eğitimi derslerinde keyifsiz olurum  |                     |             |            |              |                  |
| 3. Beden Eğitimi dersinde canım sıkılır.  |                     |             |            |              |                  |
| 4. Beden Eğitimi dersinde zilin çalmasını dört gözle beklerim.  |                     |             |            |              |                  |
| 5. Beden Eğitimi sevdiğim dersler arasındadır.  |                     |             |            |              |                  |
| 6. Beden Eğitimi dersinde kendimi rahat hissedirim.   |                     |             |            |              |                  |
| 7. Beden Eğitimi dersine girmek istemem.  |                     |             |            |              |                  |
| 8. Beden Eğitimi dersinde yapılan hareketleri isteyerek yapmam.   |                     |             |            |              |                  |
| 9. Beden Eğitimi sıkıcı bir derstir.  |                     |             |            |              |                  |
| 10. Boş zamanlarımda arkadaşlarımla Beden Eğitimi dersinde öğrendiğimiz oyunları oynamaktan hoşlanırım. |                     |             |            |              |                  |



|  | Tamamen Katılıyorum | Katılıyorum | Kararsızım | Katılmıyorum | Hiç Katılmıyorum |
|--|---------------------|-------------|------------|--------------|------------------|
| 11. Beden Eğitimi ders saatlerinin artırılmasını isterim.  |                     |             |            |              |                  |
| 12. Beden Eğitimi dersi ilgimi çekmez.   |                     |             |            |              |                  |
| 13. Beden Eğitimi dersinde elimden gelenin en iyisini yapmaya gayret ederim.                     |                     |             |            |              |                  |
| 14. Okulda Beden Eğitimi dersinin başlamasını dört gözle beklerim.                               |                     |             |            |              |                  |
| 15. Beden Eğitimi dersindeki fiziksel aktiviteleri yapmaktan hoşlanırım.                         |                     |             |            |              |                  |
| 16. Beden Eğitimi dersindeki hareketleri yapmak içimden gelmez.                                  |                     |             |            |              |                  |
| 17. Beden Eğitimi dersi müfredattan çıkarılmalıdır.  |                     |             |            |              |                  |
| 18. Beden Eğitimi dersinin çok faydalı olduğunu düşünürüm.                                       |                     |             |            |              |                  |
| 19. Beden Eğitimi dersinden nefret ederim.   |                     |             |            |              |                  |
| 20. Beden Eğitimi dersi insanı dinlendirir.  |                     |             |            |              |                  |
| 21. Beden Eğitimi dersine katılmak vakit kaybıdır.   |                     |             |            |              |                  |
| 22. Beden Eğitimi dersinin olduğu gün okula gitmek için heyecan duyarım.                         |                     |             |            |              |                  |
| 23. Beden Eğitimi gereksiz bir derstir.  |                     |             |            |              |                  |
| 24. Beden Eğitimi dersine katılmanın kendini dinç hissetmek için iyi bir yol olduğunu düşünürüm. |                     |             |            |              |                  |
| 25. Beden Eğitimi dersi olduğu günler okula gitmek istemem.                                      |                     |             |            |              |                  |
| 26. Sağlıklı bir insan olabilmek için Beden Eğitimi derslerine katılmak gerekir                  |                     |             |            |              |                  |
| 27. Beden Eğitimi dersinden bir önceki dersin hemen bitmesini isterim.                           |                     |             |            |              |                  |
| 28. Beden Eğitimi dersinin olduğu gün neşeli olurum.   |                     |             |            |              |                  |
| 29. Beden Eğitimi dersinin önemli olduğunu düşünürüm.  |                     |             |            |              |                  |
| 30. Beden Eğitimi dersine girmek yerine başka şeyler yapmak isterim.                             |                     |             |            |              |                  |

## EK 2. Deneme Formundaki Maddelere İlişkin Faktör Yükleri

| Madde No | Faktörler |        |        |        |        |
|----------|-----------|--------|--------|--------|--------|
|          | 1         | 2      | 3      | 4      | 5      |
| M01      | 0,566     | 0,036  | 0,376  | 0,061  | -0,120 |
| M02      | 0,558     | -0,309 | 0,202  | 0,156  | -0,059 |
| M03      | 0,613     | -0,298 | 0,153  | 0,118  | -0,040 |
| M04      | 0,501     | -0,361 | 0,207  | 0,091  | -0,018 |
| M05      | 0,553     | -0,061 | 0,363  | 0,034  | 0,071  |
| M06      | 0,473     | 0,016  | 0,312  | 0,150  | 0,009  |
| M07      | 0,525     | -0,400 | 0,123  | 0,043  | -0,043 |
| M08      | 0,467     | -0,206 | -0,057 | 0,370  | -0,159 |
| M09      | 0,562     | -0,138 | 0,261  | 0,100  | 0,185  |
| M10      | 0,419     | -0,399 | 0,028  | 0,178  | 0,006  |
| M11      | 0,544     | 0,121  | 0,425  | -0,237 | -0,037 |
| M12      | 0,504     | -0,324 | 0,154  | 0,075  | 0,098  |
| M13      | 0,453     | 0,273  | 0,116  | 0,282  | 0,036  |
| M14      | 0,238     | -0,172 | -0,065 | 0,413  | -0,274 |
| M15      | 0,487     | -0,375 | 0,076  | 0,077  | 0,032  |
| M16      | -0,297    | 0,242  | -0,063 | -0,071 | 0,248  |
| M17      | 0,428     | -0,347 | 0,012  | -0,125 | 0,085  |
| M18      | 0,562     | 0,147  | 0,268  | -0,181 | -0,041 |
| M19      | 0,323     | 0,311  | -0,002 | 0,154  | -0,340 |
| M20      | 0,169     | 0,092  | 0,096  | -0,049 | 0,393  |
| M21      | 0,585     | 0,251  | 0,200  | 0,245  | -0,027 |
| M22      | 0,483     | 0,334  | -0,114 | 0,327  | -0,089 |
| M23      | 0,312     | 0,450  | -0,124 | 0,318  | -0,096 |
| M24      | 0,629     | -0,200 | 0,001  | -0,016 | 0,129  |
| M25      | 0,635     | 0,074  | 0,050  | -0,012 | 0,229  |
| M26      | 0,485     | 0,311  | -0,152 | 0,172  | -0,066 |
| M27      | 0,595     | -0,094 | -0,162 | -0,044 | -0,109 |
| M28      | 0,364     | 0,151  | -0,142 | 0,066  | 0,236  |
| M29      | 0,644     | 0,231  | 0,114  | -0,258 | -0,099 |
| M30      | 0,638     | 0,084  | -0,182 | 0,167  | -0,031 |
| M31      | 0,560     | -0,092 | -0,322 | 0,144  | -0,182 |
| M32      | 0,542     | 0,104  | -0,268 | -0,020 | 0,243  |
| M33      | 0,613     | -0,207 | -0,297 | -0,153 | 0,024  |
| M34      | 0,513     | 0,209  | -0,144 | 0,166  | 0,098  |
| M35      | 0,601     | -0,257 | -0,270 | -0,195 | 0,070  |
| M36      | 0,658     | 0,309  | -0,025 | -0,189 | -0,135 |
| M37      | 0,629     | -0,212 | -0,282 | -0,233 | -0,008 |
| M38      | 0,460     | 0,231  | 0,050  | -0,113 | -0,065 |
| M39      | 0,518     | 0,189  | -0,110 | 0,191  | 0,394  |
| M40      | 0,555     | -0,133 | -0,277 | -0,169 | 0,104  |
| M41      | 0,511     | 0,156  | -0,115 | 0,051  | 0,432  |
| M42      | 0,550     | 0,364  | 0,204  | -0,347 | -0,149 |
| M43      | 0,378     | 0,218  | 0,230  | -0,362 | -0,013 |
| M44      | 0,641     | 0,247  | 0,064  | -0,147 | -0,013 |
| M45      | 0,631     | -0,181 | -0,188 | -0,197 | -0,035 |
| M46      | 0,448     | 0,218  | -0,096 | 0,121  | -0,013 |
| M47      | 0,377     | -0,025 | -0,228 | -0,157 | -0,342 |
| M48      | 0,612     | 0,173  | -0,180 | -0,085 | -0,026 |
| M49      | 0,603     | -0,074 | -0,221 | -0,182 | -0,173 |
| M50      | 0,656     | 0,032  | 0,048  | -0,121 | 0,079  |

### EK 3. Nihai Formdaki Maddelere İlişkin Faktör Yükleri

| Madde No | Faktörler |        |        |
|----------|-----------|--------|--------|
|          | 1         | 2      | 3      |
| M01      | 0,657     | 0,410  | 0,075  |
| M02      | 0,669     | 0,344  | -0,128 |
| M03      | 0,734     | 0,305  | -0,156 |
| M04      | 0,591     | 0,410  | -0,134 |
| M05      | 0,608     | 0,362  | 0,115  |
| M06      | 0,605     | 0,310  | 0,074  |
| M07      | 0,669     | 0,337  | -0,194 |
| M08      | 0,628     | 0,168  | -0,065 |
| M09      | 0,663     | 0,209  | 0,217  |
| M10      | 0,533     | 0,067  | -0,347 |
| M11      | 0,737     | 0,037  | -0,223 |
| M12      | 0,644     | -0,140 | 0,237  |
| M13      | 0,714     | -0,041 | -0,095 |
| M14      | 0,563     | -0,188 | 0,133  |
| M15      | 0,722     | 0,001  | 0,067  |
| M16      | 0,765     | -0,172 | 0,020  |
| M17      | 0,670     | -0,197 | -0,219 |
| M18      | 0,656     | -0,324 | 0,037  |
| M19      | 0,709     | -0,234 | -0,346 |
| M20      | 0,625     | -0,135 | 0,166  |
| M21      | 0,690     | -0,320 | -0,267 |
| M22      | 0,770     | -0,115 | 0,138  |
| M23      | 0,735     | -0,241 | -0,227 |
| M24      | 0,646     | -0,141 | 0,313  |
| M25      | 0,663     | -0,242 | -0,103 |
| M26      | 0,585     | -0,178 | 0,339  |
| M27      | 0,507     | 0,072  | 0,392  |
| M28      | 0,709     | 0,037  | 0,275  |
| M29      | 0,726     | -0,104 | 0,207  |
| M30      | 0,705     | -0,140 | -0,142 |

#### EK 4. MULTILOG 7.0 Programı İle Yapılan Madde Analizleri Sonuçları

MULTILOG--FOR MULTIPLE CATEGORICAL ITEM RESPONSE DATA--VERSION 7.0.3

MULTILOG for Windows 7.00.2327.2

Created on: 19 March 2006, 13:45:16

>PROBLEM RANDOM,

INDIVIDUAL,  
DATA = 'E:\doktora\_tez\memo\_tez3.DAT',

NITEMS = 50,

NGROUPS = 1,

NEXAMINEES = 787,

NCHARS = 26;

DATA FILE NAME IS  
E:\DOKTORA\_TEZ\MEMO\_TEZ3.DAT

TYPE OF INPUT:  
INDIVIDUAL RESPONSE VECTORS

>TEST ALL,

GRADED,

NC = (5(0)50);

---> FIND WARNING: 1 RECORDS NOT CONTAINING > IN COLUMN 1 HAVE BEEN  
SKIPPED

NUMBER OF CODES 6

123450

VECTOR OF CATEGORIES FOR CODE=1

11

VECTOR OF CATEGORIES FOR CODE=2

22

VECTOR OF CATEGORIES FOR CODE=3

33

VECTOR OF CATEGORIES FOR CODE=4

44

VECTOR OF CATEGORIES FOR CODE=5

55

VECTOR OF CATEGORIES FOR CODE=0

00  
(26A1,1X,50A1)

MULTILOG--FOR MULTIPLE CATEGORICAL ITEM RESPONSE DATA--VERSION 7.0.3  
MULTILOG for Windows 7.00.2327.2

Created on: 19 March 2006, 13:45:16

DATA PARAMETERS:

- NUMBER OF LINES IN THE DATA FILE: 787
- NUMBER OF CATEGORICAL-RESPONSE ITEMS: 50
- NUMBER OF CONTINUOUS-RESPONSE ITEMS, AND/OR GROUPS: 1
- TOTAL NUMBER OF "ITEMS" (INCLUDING GROUPS): 51
- NUMBER OF CHARACTERS IN ID FIELDS: 26
- MAXIMUM NUMBER OF RESPONSE-CODES FOR ANY ITEM: 6
- THE MISSING VALUE CODE FOR CONTINUOUS DATA: 9.0000
- THE DATA WILL BE STORED IN MEMORY

ESTIMATION PARAMETERS:

- THE ITEMS WILL BE CALIBRATED--
- BY MARGINAL MAXIMUM LIKELIHOOD ESTIMATION
- MAXIMUM NUMBER OF EM CYCLES PERMITTED: 25
- NUMBER OF PARAMETER-SEGMENTS USED IS: 50
- NUMBER OF FREE PARAMETERS IS: 250
- MAXIMUM NUMBER OF M-STEP ITERATIONS IS 4 TIMES
- THE NUMBER OF PARAMETERS IN THE SEGMENT
- NUMBER OF QUADRATURE POINTS IS: 19
- THE M-STEP CONVERGENCE CRITERION IS: 0.000100
- THE EM-CYCLE CONVERGENCE CRITERION IS: 0.001000
- THE RK CONTROL PARAMETER (FOR THE M-STEPS) IS: 0.9000
- THE RM CONTROL PARAMETER (FOR THE M-STEPS) IS: 1.0000
- THE MAXIMUM ACCELERATION PERMITTED IS: 0.0000
- THETA-GROUP LOCATIONS WILL REMAIN UNCHANGED

QUADRATURE POINTS FOR MML,  
AT THETA:

|        |        |        |        |        |        |        |        |        |       |       |
|--------|--------|--------|--------|--------|--------|--------|--------|--------|-------|-------|
| -4.500 | -4.000 | -3.500 | -3.000 | -2.500 | -2.000 | -1.500 | -1.000 | -0.500 | 0.000 | 0.500 |
|        | 1.000  | 1.500  | 2.000  | 2.500  | 3.000  | 3.500  | 4.000  | 4.500  |       |       |

MULTILOG for Windows 7.00.2327.2

READING DATA...

KEY-

CODE CATEGORY

```

1 1111111111111111111111111111111111111111111111111111111111111111
2 2222222222222222222222222222222222222222222222222222222222222222
3 3333333333333333333333333333333333333333333333333333333333333333
4 4444444444444444444444444444444444444444444444444444444444444444
5 55555555555555555555555555555555555555555555555555555555555555
0 0000000000000000000000000000000000000000000000000000000000000000

```

FORMAT FOR DATA-

(26A1,1X,50A1)

FIRST OBSERVATION AS READ-

ID 0001 5,00 2,00 1994  
 ITEMS 5555555555555555155452555555555515555455455555525545  
 NORML 0.000

ERROR MESSAGE FROM SUBROUTINE INVSD: MATRIX IS SINGULAR

FINISHED CYCLE 25  
 MAXIMUM INTERCYCLE PARAMETER CHANGE= 0.30860 P( 75)

ITEM SUMMARY

MULTILOG for Windows 7.00.2327.2

ITEM 1: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 1 1.28 (0.14)  
 B(1) 2 -2.15 (0.28)  
 B(2) 3 -1.37 (0.18)  
 B(3) 4 -0.78 (0.13)  
 B(4) 5 0.63 (0.08)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.179 0.216 0.256 0.298 0.340 0.381 0.418 0.448  
 -1.4 - 0.0 0.473 0.490 0.502 0.508 0.511 0.511 0.509 0.505  
 0.2 - 1.6 0.500 0.492 0.480 0.464 0.442 0.414 0.380 0.341  
 1.8 - 3.0 0.299 0.258 0.218 0.182 0.149 0.121 0.097

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 32 21 51 92 160  
 OBS. PROP. 0.0899 0.0590 0.1433 0.2584 0.4494  
 EXP. PROP. 0.1883 0.0890 0.1776 0.2499 0.2952

ITEM 2: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 6 1.67 (0.14)  
 B(1) 7 -1.96 (0.25)  
 B(2) 8 -1.17 (0.15)  
 B(3) 9 -0.60 (0.11)  
 B(4) 10 0.62 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.370 0.471 0.581 0.689 0.785 0.859 0.910 0.941  
 -1.4 - 0.0 0.959 0.968 0.969 0.962 0.946 0.925 0.904 0.889  
 0.2 - 1.6 0.880 0.869 0.842 0.790 0.710 0.611 0.503 0.400  
 1.8 - 3.0 0.308 0.232 0.171 0.125 0.090 0.064 0.046

b OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 23 44 68 256 396  
 OBS. PROP. 0.0292 0.0559 0.0864 0.3253 0.5032  
 EXP. PROP. 0.0835 0.1184 0.1328 0.3376 0.3277

## ITEM 3: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 11 1.50 (0.14)  
 B(1) 12 -2.07 (0.32)  
 B(2) 13 -1.29 (0.21)  
 B(3) 14 -0.81 (0.16)  
 B(4) 15 0.37 (0.09)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.319 0.370 0.419 0.464 0.501 0.531 0.552 0.566  
 -1.4 - 0.0 0.575 0.578 0.578 0.575 0.569 0.562 0.554 0.543  
 0.2 - 1.6 0.528 0.506 0.475 0.437 0.391 0.342 0.293 0.245  
 1.8 - 3.0 0.201 0.163 0.130 0.103 0.081 0.063 0.049

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 36 48 51 202 450  
 OBS. PROP. 0.0457 0.0610 0.0648 0.2567 0.5718  
 EXP. PROP. 0.1027 0.1087 0.0955 0.2837 0.4094

## ITEM 4: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 16 1.15 (0.15)  
 B(1) 17 -1.85 (0.26)  
 B(2) 18 -1.31 (0.18)  
 B(3) 19 -0.78 (0.13)  
 B(4) 20 0.34 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.167 0.193 0.221 0.248 0.275 0.300 0.322 0.341  
 -1.4 - 0.0 0.356 0.368 0.376 0.382 0.385 0.386 0.386 0.383  
 0.2 - 1.6 0.377 0.369 0.357 0.342 0.323 0.301 0.276 0.249  
 1.8 - 3.0 0.221 0.194 0.168 0.143 0.121 0.101 0.084

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 40 41 44 56 175  
 OBS. PROP. 0.1124 0.1152 0.1236 0.1573 0.4916  
 EXP. PROP. 0.2039 0.1524 0.1416 0.1564 0.3458

## ITEM 5: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 21 1.26 (0.13)  
 B(1) 22 -1.79 (0.29)  
 B(2) 23 -1.17 (0.20)  
 B(3) 24 -0.55 (0.14)  
 B(4) 25 0.69 (0.09)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.243 0.287 0.333 0.377 0.419 0.455 0.484 0.506  
 -1.4 - 0.0 0.521 0.530 0.534 0.534 0.532 0.527 0.521 0.515  
 0.2 - 1.6 0.507 0.496 0.480 0.458 0.429 0.393 0.352 0.308  
 1.8 - 3.0 0.264 0.222 0.184 0.150 0.121 0.097 0.077

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 53 49 79 226 380

OBS. PROP. 0.0673 0.0623 0.1004 0.2872 0.4828  
 EXP. PROP. 0.1401 0.0974 0.1319 0.2949 0.3358

ITEM 6: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 26 1.15 (0.15)  
 B(1) 27 -1.76 (0.25)  
 B(2) 28 -1.40 (0.20)  
 B(3) 29 -0.98 (0.15)  
 B(4) 30 0.29 (0.08)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.167 0.193 0.221 0.248 0.275 0.300 0.322 0.341  
 -1.4 - 0.0 0.356 0.368 0.376 0.382 0.385 0.386 0.386 0.383  
 0.2 - 1.6 0.377 0.369 0.357 0.342 0.323 0.301 0.276 0.249  
 1.8 - 3.0 0.221 0.194 0.168 0.143 0.121 0.101 0.084

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 40 41 44 56 175  
 OBS. PROP. 0.1124 0.1152 0.1236 0.1573 0.4916  
 EXP. PROP. 0.2039 0.1524 0.1416 0.1564 0.3458

ITEM 7: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 31 1.24 (0.13)  
 B(1) 32 -1.53 (0.26)  
 B(2) 33 -0.89 (0.18)  
 B(3) 34 -0.24 (0.12)  
 B(4) 35 0.85 (0.10)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.185 0.221 0.259 0.298 0.337 0.373 0.405 0.432  
 -1.4 - 0.0 0.453 0.469 0.479 0.484 0.487 0.486 0.484 0.480  
 0.2 - 1.6 0.474 0.466 0.454 0.437 0.415 0.388 0.355 0.319  
 1.8 - 3.0 0.280 0.242 0.205 0.172 0.142 0.116 0.094

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 77 65 95 201 349  
 OBS. PROP. 0.0978 0.0826 0.1207 0.2554 0.4435  
 EXP. PROP. 0.1839 0.1152 0.1444 0.2495 0.3069

ITEM 8: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 36 1.31 (0.17)  
 B(1) 37 -1.55 (0.22)  
 B(2) 38 -1.15 (0.16)  
 B(3) 39 -0.83 (0.13)  
 B(4) 40 0.09 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.202 0.279 0.378 0.498 0.632 0.771 0.898 1.000  
 -1.4 - 0.0 1.069 1.106 1.116 1.109 1.091 1.071 1.048 1.010  
 0.2 - 1.6 0.947 0.850 0.727 0.593 0.464 0.350 0.258 0.186  
 1.8 - 3.0 0.132 0.093 0.065 0.045 0.031 0.021 0.015



OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 36 28 32 166 525  
 OBS. PROP. 0.0457 0.0356 0.0407 0.2109 0.6671  
 EXP. PROP. 0.1287 0.0708 0.0693 0.2592 0.4721

## ITEM 9: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 41 1.05 (0.13)  
 B(1) 42 -2.40 (0.34)  
 B(2) 43 -1.83 (0.26)  
 B(3) 44 -1.23 (0.19)  
 B(4) 45 0.04 (0.10)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.163 0.190 0.217 0.245 0.273 0.299 0.322 0.342  
 -1.4 - 0.0 0.359 0.371 0.379 0.385 0.388 0.388 0.387 0.383  
 0.2 - 1.6 0.378 0.370 0.359 0.346 0.328 0.307 0.283 0.256  
 1.8 - 3.0 0.229 0.201 0.175 0.149 0.127 0.106 0.088

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 48 31 44 92 183  
 OBS. PROP. 0.1206 0.0779 0.1106 0.2312 0.4598  
 EXP. PROP. 0.2086 0.0992 0.1268 0.2304 0.3351

## ITEM 10: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 46 0.97 (0.15)  
 B(1) 47 -1.22 (0.20)  
 B(2) 48 -0.66 (0.14)  
 B(3) 49 -0.05 (0.09)  
 B(4) 50 0.45 (0.08)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.145 0.160 0.174 0.189 0.202 0.214 0.225 0.233  
 -1.4 - 0.0 0.241 0.246 0.251 0.254 0.255 0.256 0.255 0.254  
 0.2 - 1.6 0.251 0.247 0.241 0.234 0.225 0.214 0.202 0.189  
 1.8 - 3.0 0.175 0.160 0.145 0.130 0.116 0.102 0.090

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 55 51 40 72 180  
 OBS. PROP. 0.1382 0.1281 0.1005 0.1809 0.4523  
 EXP. PROP. 0.2186 0.1565 0.1077 0.1744 0.3427

## ITEM 11: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 51 1.88 (0.14)  
 B(1) 52 -1.97 (0.30)  
 B(2) 53 -1.34 (0.21)  
 B(3) 54 -0.76 (0.15)  
 B(4) 55 0.33 (0.08)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.230 0.317 0.427 0.558 0.700 0.841 0.964 1.057  
 -1.4 - 0.0 1.116 1.144 1.149 1.139 1.121 1.101 1.077 1.040  
 0.2 - 1.6 0.975 0.877 0.751 0.612 0.477 0.360 0.263 0.189  
 1.8 - 3.0 0.134 0.094 0.065 0.045 0.031 0.021 0.015

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 12 10 18 71 245  
 OBS. PROP. 0.0337 0.0281 0.0506 0.1994 0.6882  
 EXP. PROP. 0.1157 0.0648 0.0963 0.2554 0.4677

ITEM 12: 5 GRADED CATEGORIES  
 P(#) ESTIMATE (S.E.)

A 56 1.14 (0.12)  
 B(1) 57 -1.57 (0.30)  
 B(2) 58 -0.65 (0.19)  
 B(3) 59 -0.01 (0.13)  
 B(4) 60 0.91 (0.11)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.164 0.187 0.210 0.233 0.255 0.275 0.293 0.307  
 -1.4 - 0.0 0.320 0.329 0.336 0.340 0.343 0.345 0.344 0.343  
 0.2 - 1.6 0.339 0.334 0.326 0.316 0.303 0.287 0.269 0.248  
 1.8 - 3.0 0.226 0.203 0.180 0.157 0.136 0.117 0.099

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 96 98 95 155 343  
 OBS. PROP. 0.1220 0.1245 0.1207 0.1970 0.4358  
 EXP. PROP. 0.2055 0.1598 0.1319 0.1871 0.3157

ITEM 13: 5 GRADED CATEGORIES  
 P(#) ESTIMATE (S.E.)

A 61 1.06 (0.12)  
 B(1) 62 -2.32 (0.34)  
 B(2) 63 -1.21 (0.21)  
 B(3) 64 -0.46 (0.14)  
 B(4) 65 0.85 (0.11)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.145 0.160 0.174 0.189 0.202 0.214 0.225 0.233  
 -1.4 - 0.0 0.241 0.246 0.251 0.254 0.255 0.256 0.255 0.254  
 0.2 - 1.6 0.251 0.247 0.241 0.234 0.225 0.214 0.202 0.189  
 1.8 - 3.0 0.175 0.160 0.145 0.130 0.116 0.102 0.090

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 55 51 40 72 180  
 OBS. PROP. 0.1382 0.1281 0.1005 0.1809 0.4523  
 EXP. PROP. 0.2186 0.1565 0.1077 0.1744 0.3427

ITEM 14: 5 GRADED CATEGORIES  
 P(#) ESTIMATE (S.E.)

A 71 0.13 (0.11)  
 B(1) 72 -3.84 (3.97)  
 B(2) 73 3.02 (2.16)

B(3) 74 7.52 (5.75)  
 B(4) 75 12.90 (\*\*\*\*)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.005 0.005 0.005 0.005 0.005 0.005 0.005 0.005  
 -1.4 - 0.0 0.005 0.005 0.005 0.005 0.005 0.005 0.005 0.005  
 0.2 - 1.6 0.005 0.005 0.005 0.005 0.005 0.005 0.005 0.005  
 1.8 - 3.0 0.005 0.005 0.005 0.005 0.005 0.005 0.005

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 288 178 105 93 123  
 OBS. PROP. 0.3659 0.2262 0.1334 0.1182 0.1563  
 EXP. PROP. 0.3801 0.2150 0.1280 0.1153 0.1616

ITEM 15: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 74 1.74 (0.15)  
 B(1) 82 -1.23 (0.18)  
 B(2) 83 -0.67 (0.12)  
 B(3) 84 -0.13 (0.09)  
 B(4) 85 0.57 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.128 0.175 0.236 0.312 0.404 0.507 0.616 0.719  
 -1.4 - 0.0 0.808 0.876 0.921 0.948 0.961 0.965 0.961 0.952  
 0.2 - 1.6 0.933 0.901 0.850 0.777 0.684 0.579 0.472 0.372  
 1.8 - 3.0 0.286 0.215 0.159 0.116 0.084 0.060 0.043

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 65 66 97 155 404  
 OBS. PROP. 0.0826 0.0839 0.1233 0.1970 0.5133  
 EXP. PROP. 0.1904 0.1248 0.1488 0.1939 0.3421

ITEM 16: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 76 1.89 (0.15)  
 B(1) 77 -1.95 (0.34)  
 B(2) 78 -1.58 (0.27)  
 B(3) 79 -0.91 (0.18)  
 B(4) 80 -0.11 (0.11)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.230 0.317 0.427 0.558 0.700 0.841 0.964 1.057  
 -1.4 - 0.0 1.116 1.144 1.149 1.139 1.121 1.101 1.077 1.040  
 0.2 - 1.6 0.975 0.877 0.751 0.612 0.477 0.360 0.263 0.189  
 1.8 - 3.0 0.134 0.094 0.065 0.045 0.031 0.021 0.015

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 12 10 18 71 245  
 OBS. PROP. 0.0337 0.0281 0.0506 0.1994 0.6882  
 EXP. PROP. 0.1157 0.0648 0.0963 0.2554 0.4677

## ITEM 17: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 81 1.95 (0.15)  
 B(1) 82 -1.23 (0.18)  
 B(2) 83 -0.67 (0.12)  
 B(3) 84 -0.13 (0.09)  
 B(4) 85 0.57 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.128 0.175 0.236 0.312 0.404 0.507 0.616 0.719  
 -1.4 - 0.0 0.808 0.876 0.921 0.948 0.961 0.965 0.961 0.952  
 0.2 - 1.6 0.933 0.901 0.850 0.777 0.684 0.579 0.472 0.372  
 1.8 - 3.0 0.286 0.215 0.159 0.116 0.084 0.060 0.043

## OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 65 66 97 155 404  
 OBS. PROP. 0.0826 0.0839 0.1233 0.1970 0.5133  
 EXP. PROP. 0.1904 0.1248 0.1488 0.1939 0.3421

## ITEM 18: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 86 1.00 (0.12)  
 B(1) 87 -1.36 (0.33)  
 B(2) 88 -0.20 (0.19)  
 B(3) 89 0.87 (0.14)  
 B(4) 90 1.82 (0.21)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.145 0.160 0.174 0.189 0.202 0.214 0.225 0.233  
 -1.4 - 0.0 0.241 0.246 0.251 0.254 0.255 0.256 0.255 0.254  
 0.2 - 1.6 0.251 0.247 0.241 0.234 0.225 0.214 0.202 0.189  
 1.8 - 3.0 0.175 0.160 0.145 0.130 0.116 0.102 0.090

## OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 55 51 40 72 180  
 OBS. PROP. 0.1382 0.1281 0.1005 0.1809 0.4523  
 EXP. PROP. 0.2186 0.1565 0.1077 0.1744 0.3427

## ITEM 19: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 91 0.74 (0.12)  
 B(1) 92 -5.58 (1.50)  
 B(2) 93 -4.59 (1.21)  
 B(3) 94 -3.12 (0.87)  
 B(4) 95 -0.75 (0.35)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.093 0.093 0.093 0.092 0.092 0.091 0.090 0.089  
 -1.4 - 0.0 0.088 0.087 0.086 0.084 0.082 0.080 0.078 0.075  
 0.2 - 1.6 0.073 0.070 0.067 0.064 0.061 0.058 0.054 0.051  
 1.8 - 3.0 0.048 0.045 0.041 0.038 0.036 0.033 0.030

## OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 29 19 50 161 528  
 OBS. PROP. 0.0368 0.0241 0.0635 0.2046 0.6709

EXP. PROP. 0.0523 0.0331 0.0828 0.2380 0.5938

ITEM 20: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 96 1.16 (0.17)  
 B(1) 97 -1.68 (0.22)  
 B(2) 98 -1.07 (0.15)  
 B(3) 99 -0.40 (0.09)  
 B(4) 100 0.46 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.167 0.193 0.220 0.248 0.275 0.300 0.323 0.342  
 -1.4 - 0.0 0.357 0.369 0.377 0.383 0.386 0.387 0.385 0.382  
 0.2 - 1.6 0.377 0.369 0.358 0.343 0.325 0.304 0.280 0.253  
 1.8 - 3.0 0.226 0.198 0.172 0.147 0.124 0.104 0.087

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 45 45 36 87 185  
 OBS. PROP. 0.1131 0.1131 0.0905 0.2186 0.4648  
 EXP. PROP. 0.2046 0.1456 0.0983 0.2124 0.3391

ITEM 21: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 101 1.03 (0.13)  
 B(1) 102 -1.41 (0.22)  
 B(2) 103 -0.66 (0.14)  
 B(3) 104 -0.11 (0.10)  
 B(4) 105 0.95 (0.10)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.167 0.193 0.220 0.248 0.275 0.300 0.323 0.342  
 -1.4 - 0.0 0.357 0.369 0.377 0.383 0.386 0.387 0.385 0.382  
 0.2 - 1.6 0.377 0.369 0.358 0.343 0.325 0.304 0.280 0.253  
 1.8 - 3.0 0.226 0.198 0.172 0.147 0.124 0.104 0.087

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 45 45 36 87 185  
 OBS. PROP. 0.1131 0.1131 0.0905 0.2186 0.4648  
 EXP. PROP. 0.2046 0.1456 0.0983 0.2124 0.3391

ITEM 22: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 106 1.04 (0.10)  
 B(1) 107 -1.69 (0.43)  
 B(2) 108 0.07 (0.20)  
 B(3) 109 1.12 (0.18)  
 B(4) 110 2.56 (0.34)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.167 0.193 0.220 0.248 0.275 0.300 0.323 0.342  
 -1.4 - 0.0 0.357 0.369 0.377 0.383 0.386 0.387 0.385 0.382  
 0.2 - 1.6 0.377 0.369 0.358 0.343 0.325 0.304 0.280 0.253  
 1.8 - 3.0 0.226 0.198 0.172 0.147 0.124 0.104 0.087

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 45 45 36 87 185  
 OBS. PROP. 0.1131 0.1131 0.0905 0.2186 0.4648  
 EXP. PROP. 0.2046 0.1456 0.0983 0.2124 0.3391

## ITEM 23: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 111 1.16 (0.21)  
 B(1) 112 -1.34 (0.18)  
 B(2) 113 -0.89 (0.12)  
 B(3) 114 -0.58 (0.09)  
 B(4) 115 0.09 (0.06)

## @THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.179 0.216 0.256 0.298 0.340 0.381 0.418 0.448  
 -1.4 - 0.0 0.473 0.490 0.502 0.508 0.511 0.511 0.509 0.505  
 0.2 - 1.6 0.500 0.492 0.480 0.464 0.442 0.414 0.380 0.341  
 1.8 - 3.0 0.299 0.258 0.218 0.182 0.149 0.121 0.097

## OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 32 21 51 92 160  
 OBS. PROP. 0.0899 0.0590 0.1433 0.2584 0.4494  
 EXP. PROP. 0.1883 0.0890 0.1776 0.2499 0.2952

## ITEM 24: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 116 2.07 (0.17)  
 B(1) 117 -1.42 (0.17)  
 B(2) 118 -0.96 (0.12)  
 B(3) 119 -0.59 (0.09)  
 B(4) 120 0.22 (0.06)

## @THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.124 0.195 0.301 0.454 0.661 0.917 1.195 1.448  
 -1.4 - 0.0 1.635 1.744 1.790 1.793 1.765 1.720 1.680 1.643  
 0.2 - 1.6 1.570 1.419 1.188 0.920 0.666 0.458 0.304 0.197  
 1.8 - 3.0 0.126 0.079 0.050 0.031 0.019 0.012 0.008

## OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 29 33 45 175 505  
 OBS. PROP. 0.0368 0.0419 0.0572 0.2224 0.6417  
 EXP. PROP. 0.1275 0.0921 0.0967 0.2546 0.4291

## ITEM 25: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 121 1.66 (0.13)  
 B(1) 122 -1.41 (0.23)  
 B(2) 123 -0.64 (0.15)  
 B(3) 124 -0.06 (0.11)  
 B(4) 125 0.94 (0.10)

## @THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.166 0.200 0.237 0.277 0.318 0.357 0.394 0.425  
 -1.4 - 0.0 0.451 0.470 0.484 0.493 0.499 0.501 0.500 0.498  
 0.2 - 1.6 0.493 0.485 0.475 0.460 0.439 0.413 0.380 0.344

1.8 - 3.0 0.304 0.264 0.225 0.188 0.156 0.127 0.103

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 82 83 94 193 335  
 OBS. PROP. 0.1042 0.1055 0.1194 0.2452 0.4257  
 EXP. PROP. 0.2022 0.1479 0.1363 0.2269 0.2868

ITEM 26: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 126 1.76 (0.17)  
 B(1) 127 -1.25 (0.17)  
 B(2) 128 -0.88 (0.13)  
 B(3) 129 -0.37 (0.09)  
 B(4) 130 0.39 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.109 0.161 0.234 0.334 0.465 0.625 0.806 0.986  
 -1.4 - 0.0 1.142 1.255 1.321 1.350 1.355 1.345 1.327 1.302  
 0.2 - 1.6 1.263 1.194 1.081 0.926 0.749 0.576 0.424 0.302  
 1.8 - 3.0 0.210 0.144 0.098 0.065 0.044 0.029 0.019

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 48 36 80 175 448  
 OBS. PROP. 0.0610 0.0457 0.1017 0.2224 0.5693  
 EXP. PROP. 0.1693 0.0815 0.1380 0.2305 0.3807

ITEM 27: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 131 2.07 (0.15)  
 B(1) 132 -2.09 (0.28)  
 B(2) 133 -1.63 (0.23)  
 B(3) 134 -1.25 (0.19)  
 B(4) 135 -0.16 (0.10)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.066 0.107 0.172 0.272 0.421 0.630 0.901 1.209  
 -1.4 - 0.0 1.507 1.737 1.873 1.927 1.927 1.897 1.854 1.799  
 0.2 - 1.6 1.692 1.496 1.220 0.919 0.647 0.434 0.281 0.178  
 1.8 - 3.0 0.111 0.069 0.042 0.026 0.016 0.010 0.006

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 17 14 23 62 240  
 OBS. PROP. 0.0478 0.0393 0.0646 0.1742 0.6742  
 EXP. PROP. 0.1682 0.0724 0.1043 0.2181 0.4369

ITEM 28: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 136 1.74 (0.15)  
 B(1) 137 -1.30 (0.16)  
 B(2) 138 -0.72 (0.10)  
 B(3) 139 -0.17 (0.07)  
 B(4) 140 0.55 (0.06)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.227 0.308 0.409 0.527 0.654 0.779 0.888 0.971  
 -1.4 - 0.0 1.025 1.054 1.066 1.068 1.066 1.061 1.050 1.032  
 0.2 - 1.6 1.000 0.945 0.861 0.752 0.628 0.503 0.388 0.291  
 1.8 - 3.0 0.214 0.154 0.110 0.078 0.055 0.038 0.027

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 14 16 39 72 215  
 OBS. PROP. 0.0393 0.0449 0.1096 0.2022 0.6039  
 EXP. PROP. 0.1215 0.0950 0.1722 0.2194 0.3919

ITEM 29: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 141 1.35 (0.16)  
 B(1) 142 -1.35 (0.18)  
 B(2) 143 -0.80 (0.12)  
 B(3) 144 -0.15 (0.08)  
 B(4) 145 0.66 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.301 0.362 0.424 0.485 0.539 0.584 0.618 0.640  
 -1.4 - 0.0 0.652 0.656 0.652 0.643 0.627 0.602 0.568 0.523  
 0.2 - 1.6 0.469 0.409 0.348 0.289 0.235 0.188 0.148 0.115  
 1.8 - 3.0 0.089 0.068 0.052 0.039 0.030 0.023 0.017

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 18 9 16 58 297  
 OBS. PROP. 0.0452 0.0226 0.0402 0.1457 0.7462  
 EXP. PROP. 0.1107 0.0441 0.0697 0.1924 0.5832

ITEM 30: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 146 2.50 (0.15)  
 B(1) 147 -1.30 (0.19)  
 B(2) 148 -0.77 (0.13)  
 B(3) 149 -0.32 (0.10)  
 B(4) 150 0.65 (0.08)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.071 0.114 0.181 0.284 0.434 0.644 0.910 1.208  
 -1.4 - 0.0 1.490 1.705 1.829 1.873 1.864 1.826 1.758 1.629  
 0.2 - 1.6 1.416 1.138 0.848 0.595 0.398 0.259 0.165 0.103  
 1.8 - 3.0 0.064 0.040 0.025 0.015 0.009 0.006 0.004

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 20 15 12 60 291  
 OBS. PROP. 0.0503 0.0377 0.0302 0.1508 0.7312  
 EXP. PROP. 0.1663 0.0743 0.0519 0.1920 0.5155

ITEM 31: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)  
 A 151 1.66 (0.15)  
 B(1) 152 -1.61 (0.21)  
 B(2) 153 -1.14 (0.16)  
 B(3) 154 -0.64 (0.12)



B(4) 155 0.38 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

|        |      |       |       |       |       |       |       |       |       |
|--------|------|-------|-------|-------|-------|-------|-------|-------|-------|
| -3.0 - | -1.6 | 0.227 | 0.296 | 0.377 | 0.469 | 0.565 | 0.657 | 0.737 | 0.800 |
| -1.4 - | 0.0  | 0.842 | 0.867 | 0.876 | 0.875 | 0.867 | 0.855 | 0.841 | 0.825 |
| 0.2 -  | 1.6  | 0.803 | 0.766 | 0.711 | 0.636 | 0.548 | 0.455 | 0.366 | 0.287 |
| 1.8 -  | 3.0  | 0.220 | 0.165 | 0.123 | 0.090 | 0.066 | 0.048 | 0.035 |       |

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 41 35 59 200 452  
 OBS. PROP. 0.0521 0.0445 0.0750 0.2541 0.5743  
 EXP. PROP. 0.1331 0.0815 0.1125 0.2790 0.3938

ITEM 32: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

|      |     |       |        |
|------|-----|-------|--------|
| A    | 156 | 1.62  | (0.22) |
| B(1) | 157 | -1.33 | (0.18) |
| B(2) | 158 | -1.08 | (0.14) |
| B(3) | 159 | -0.81 | (0.11) |
| B(4) | 160 | -0.06 | (0.06) |

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

|        |      |       |       |       |       |       |       |       |       |
|--------|------|-------|-------|-------|-------|-------|-------|-------|-------|
| -3.0 - | -1.6 | 0.134 | 0.175 | 0.227 | 0.289 | 0.361 | 0.439 | 0.519 | 0.594 |
| -1.4 - | 0.0  | 0.660 | 0.711 | 0.747 | 0.767 | 0.776 | 0.775 | 0.769 | 0.759 |
| 0.2 -  | 1.6  | 0.743 | 0.721 | 0.686 | 0.638 | 0.575 | 0.501 | 0.423 | 0.347 |
| 1.8 -  | 3.0  | 0.278 | 0.217 | 0.168 | 0.127 | 0.096 | 0.072 | 0.053 |       |

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 37 22 36 102 201  
 OBS. PROP. 0.0930 0.0553 0.0905 0.2563 0.5050  
 EXP. PROP. 0.2059 0.0849 0.1129 0.2502 0.3461

ITEM 33: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

|      |     |       |        |
|------|-----|-------|--------|
| A    | 161 | 1.53  | (0.13) |
| B(1) | 162 | -1.41 | (0.22) |
| B(2) | 163 | -0.75 | (0.15) |
| B(3) | 164 | 0.00  | (0.10) |
| B(4) | 165 | 0.83  | (0.09) |

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

|        |      |       |       |       |       |       |       |       |       |
|--------|------|-------|-------|-------|-------|-------|-------|-------|-------|
| -3.0 - | -1.6 | 0.144 | 0.183 | 0.230 | 0.283 | 0.343 | 0.405 | 0.466 | 0.522 |
| -1.4 - | 0.0  | 0.570 | 0.607 | 0.634 | 0.651 | 0.660 | 0.664 | 0.663 | 0.659 |
| 0.2 -  | 1.6  | 0.652 | 0.641 | 0.626 | 0.603 | 0.571 | 0.528 | 0.476 | 0.417 |
| 1.8 -  | 3.0  | 0.356 | 0.296 | 0.241 | 0.193 | 0.152 | 0.119 | 0.092 |       |

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 33 31 44 90 158  
 OBS. PROP. 0.0927 0.0871 0.1236 0.2528 0.4438  
 EXP. PROP. 0.2089 0.1318 0.1458 0.2338 0.2796

ITEM 34: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

|      |     |       |        |
|------|-----|-------|--------|
| A    | 166 | 1.43  | (0.17) |
| B(1) | 167 | -1.30 | (0.16) |

B(2) 168 -0.97 (0.13)  
 B(3) 169 -0.61 (0.10)  
 B(4) 170 0.12 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.138 0.175 0.219 0.269 0.326 0.385 0.444 0.500  
 -1.4 - 0.0 0.547 0.585 0.613 0.631 0.641 0.645 0.645 0.641  
 0.2 - 1.6 0.634 0.622 0.605 0.580 0.544 0.500 0.447 0.389  
 1.8 - 3.0 0.330 0.274 0.223 0.178 0.140 0.110 0.085

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 35 27 44 83 167  
 OBS. PROP. 0.0983 0.0758 0.1236 0.2331 0.4691  
 EXP. PROP. 0.2180 0.1160 0.1509 0.2193 0.2959

ITEM 35: 5 GRADED CATEGORIES  
 P(#) ESTIMATE (S.E.)  
 A 171 1.41 (0.16)  
 B(1) 172 -1.27 (0.15)  
 B(2) 173 -0.52 (0.09)  
 B(3) 174 0.00 (0.07)  
 B(4) 175 0.71 (0.06)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.392 0.450 0.504 0.550 0.586 0.612 0.628 0.636  
 -1.4 - 0.0 0.638 0.636 0.631 0.622 0.610 0.593 0.567 0.532  
 0.2 - 1.6 0.487 0.434 0.377 0.319 0.264 0.215 0.172 0.135  
 1.8 - 3.0 0.106 0.082 0.063 0.048 0.037 0.028 0.021

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 10 8 18 59 261  
 OBS. PROP. 0.0281 0.0225 0.0506 0.1657 0.7331  
 EXP. PROP. 0.0796 0.0497 0.0949 0.2252 0.5506

ITEM 36: 5 GRADED CATEGORIES  
 P(#) ESTIMATE (S.E.)  
 A 176 2.68 (0.17)  
 B(1) 177 -1.43 (0.17)  
 B(2) 178 -1.06 (0.13)  
 B(3) 179 -0.72 (0.10)  
 B(4) 180 0.14 (0.06)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.055 0.094 0.159 0.266 0.435 0.687 1.031 1.438  
 -1.4 - 0.0 1.834 2.130 2.293 2.352 2.349 2.310 2.231 2.071  
 0.2 - 1.6 1.783 1.395 0.995 0.661 0.417 0.254 0.152 0.090  
 1.8 - 3.0 0.053 0.031 0.018 0.010 0.006 0.004 0.002

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 17 15 23 54 289  
 OBS. PROP. 0.0427 0.0377 0.0578 0.1357 0.7261  
 EXP. PROP. 0.1560 0.0755 0.0946 0.1697 0.5042

ITEM 37: 5 GRADED CATEGORIES  
 P(#) ESTIMATE (S.E.)

A 181 1.07 (0.12)  
 B(1) 182 -1.23 (0.22)  
 B(2) 183 -0.50 (0.14)  
 B(3) 184 0.12 (0.10)  
 B(4) 185 1.00 (0.10)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.190 0.221 0.254 0.287 0.318 0.347 0.372 0.392  
 -1.4 - 0.0 0.407 0.418 0.426 0.430 0.432 0.433 0.432 0.430  
 0.2 - 1.6 0.426 0.420 0.409 0.395 0.376 0.352 0.324 0.293  
 1.8 - 3.0 0.260 0.227 0.195 0.166 0.139 0.115 0.095

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 32 31 66 61 166  
 OBS. PROP. 0.0899 0.0871 0.1854 0.1713 0.4663  
 EXP. PROP. 0.1805 0.1259 0.2171 0.1633 0.3132

ITEM 38: 5 GRADED CATEGORIES  
 P(#) ESTIMATE (S.E.)

A 186 0.62 (0.13)  
 B(1) 187 -1.72 (0.24)  
 B(2) 188 -0.99 (0.16)  
 B(3) 189 -0.25 (0.10)  
 B(4) 190 0.73 (0.08)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.103 0.108 0.112 0.116 0.120 0.123 0.126 0.128  
 -1.4 - 0.0 0.130 0.131 0.132 0.133 0.133 0.134 0.133 0.133  
 0.2 - 1.6 0.132 0.131 0.129 0.127 0.125 0.122 0.118 0.114  
 1.8 - 3.0 0.110 0.105 0.100 0.095 0.090 0.084 0.078

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 62 54 59 60 163  
 OBS. PROP. 0.1558 0.1357 0.1482 0.1508 0.4095  
 EXP. PROP. 0.2180 0.1583 0.1536 0.1429 0.3272

ITEM 39: 5 GRADED CATEGORIES  
 P(#) ESTIMATE (S.E.)

A 191 2.06 (0.16)  
 B(1) 192 -1.64 (0.22)  
 B(2) 193 -1.27 (0.16)  
 B(3) 194 -0.81 (0.12)  
 B(4) 195 0.09 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.092 0.134 0.192 0.273 0.380 0.513 0.668 0.833  
 -1.4 - 0.0 0.987 1.111 1.195 1.241 1.257 1.255 1.236 1.198  
 0.2 - 1.6 1.131 1.025 0.885 0.726 0.567 0.425 0.309 0.219  
 1.8 - 3.0 0.153 0.105 0.072 0.049 0.033 0.022 0.015

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 30 12 39 69 248  
 OBS. PROP. 0.0754 0.0302 0.0980 0.1734 0.6231  
 EXP. PROP. 0.1964 0.0498 0.1292 0.1869 0.4377

## ITEM 40: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 196 1.67 (0.15)  
 B(1) 197 -1.97 (0.26)  
 B(2) 198 -1.37 (0.18)  
 B(3) 199 -0.88 (0.14)  
 B(4) 200 0.07 (0.08)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.161 0.217 0.289 0.376 0.477 0.584 0.690 0.783  
 -1.4 - 0.0 0.856 0.906 0.937 0.953 0.960 0.960 0.956 0.946  
 0.2 - 1.6 0.927 0.894 0.842 0.767 0.673 0.568 0.461 0.363  
 1.8 - 3.0 0.278 0.208 0.154 0.112 0.081 0.058 0.042

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 25 29 52 83 209  
 OBS. PROP. 0.0628 0.0729 0.1307 0.2085 0.5251  
 EXP. PROP. 0.1645 0.1259 0.1647 0.1975 0.3474

## ITEM 41: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 201 2.25 (0.14)  
 B(1) 202 -1.16 (0.17)  
 B(2) 203 -0.36 (0.11)  
 B(3) 204 0.26 (0.08)  
 B(4) 205 0.72 (0.08)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.124 0.188 0.280 0.408 0.576 0.780 1.001 1.207  
 -1.4 - 0.0 1.366 1.464 1.509 1.519 1.515 1.510 1.503 1.490  
 0.2 - 1.6 1.465 1.413 1.312 1.151 0.946 0.730 0.534 0.375  
 1.8 - 3.0 0.256 0.171 0.113 0.074 0.048 0.031 0.020

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 14 16 46 78 202  
 OBS. PROP. 0.0393 0.0449 0.1292 0.2191 0.5674  
 EXP. PROP. 0.1430 0.0993 0.1947 0.2213 0.3418

## ITEM 42: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 206 0.62 (0.11)  
 B(1) 207 -1.93 (0.35)  
 B(2) 208 -0.88 (0.21)  
 B(3) 209 0.06 (0.13)  
 B(4) 210 0.91 (0.12)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.103 0.108 0.112 0.116 0.120 0.123 0.126 0.128  
 -1.4 - 0.0 0.130 0.131 0.132 0.133 0.133 0.134 0.133 0.133  
 0.2 - 1.6 0.132 0.131 0.129 0.127 0.125 0.122 0.118 0.114  
 1.8 - 3.0 0.110 0.105 0.100 0.095 0.090 0.084 0.078

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 62 54 59 60 163  
 OBS. PROP. 0.1558 0.1357 0.1482 0.1508 0.4095

EXP. PROP. 0.2180 0.1583 0.1536 0.1429 0.3272

ITEM 43: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 211 2.05 (0.17)  
 B(1) 212 -1.43 (0.17)  
 B(2) 213 -0.84 (0.12)  
 B(3) 214 -0.32 (0.08)  
 B(4) 215 0.50 (0.06)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.136 0.198 0.284 0.398 0.541 0.706 0.879 1.036  
 -1.4 - 0.0 1.157 1.236 1.282 1.307 1.320 1.322 1.313 1.295  
 0.2 - 1.6 1.262 1.204 1.107 0.969 0.803 0.631 0.474 0.344  
 1.8 - 3.0 0.243 0.168 0.115 0.078 0.052 0.035 0.024

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 20 34 40 84 220  
 OBS. PROP. 0.0503 0.0854 0.1005 0.2111 0.5528  
 EXP. PROP. 0.1510 0.1495 0.1319 0.2080 0.3596

ITEM 44: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 216 2.38 (0.15)  
 B(1) 217 -1.44 (0.19)  
 B(2) 218 -1.09 (0.14)  
 B(3) 219 -0.57 (0.10)  
 B(4) 220 0.22 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.090 0.138 0.210 0.313 0.456 0.642 0.864 1.097  
 -1.4 - 0.0 1.307 1.462 1.552 1.588 1.591 1.574 1.543 1.491  
 0.2 - 1.6 1.395 1.239 1.030 0.802 0.590 0.416 0.283 0.189  
 1.8 - 3.0 0.124 0.081 0.052 0.034 0.022 0.014 0.009

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 22 13 32 73 258  
 OBS. PROP. 0.0553 0.0327 0.0804 0.1834 0.6482  
 EXP. PROP. 0.1675 0.0619 0.1193 0.2056 0.4457

ITEM 45: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 221 0.54 (0.13)  
 B(1) 222 -1.87 (0.29)  
 B(2) 223 -1.16 (0.21)  
 B(3) 224 -0.49 (0.15)  
 B(4) 225 0.63 (0.10)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.103 0.108 0.112 0.116 0.120 0.123 0.126 0.128  
 -1.4 - 0.0 0.130 0.131 0.132 0.133 0.133 0.134 0.133 0.133  
 0.2 - 1.6 0.132 0.131 0.129 0.127 0.125 0.122 0.118 0.114  
 1.8 - 3.0 0.110 0.105 0.100 0.095 0.090 0.084 0.078

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5

OBS. FREQ. 62 54 59 60 163  
 OBS. PROP. 0.1558 0.1357 0.1482 0.1508 0.4095  
 EXP. PROP. 0.2180 0.1583 0.1536 0.1429 0.3272

## ITEM 46: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 226 1.04 (0.12)  
 B(1) 227 -1.95 (0.32)  
 B(2) 228 -1.07 (0.21)  
 B(3) 229 -0.16 (0.13)  
 B(4) 230 0.87 (0.11)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.214 0.238 0.262 0.283 0.302 0.318 0.331 0.341  
 -1.4 - 0.0 0.348 0.353 0.356 0.357 0.358 0.357 0.356 0.353  
 0.2 - 1.6 0.349 0.343 0.335 0.324 0.311 0.294 0.274 0.252  
 1.8 - 3.0 0.228 0.204 0.179 0.156 0.134 0.115 0.097

## OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5

OBS. FREQ. 64 73 126 176 348  
 OBS. PROP. 0.0813 0.0928 0.1601 0.2236 0.4422  
 EXP. PROP. 0.1515 0.1325 0.1820 0.2140 0.3199

## ITEM 47: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 231 1.01 (0.16)  
 B(1) 232 -1.32 (0.18)  
 B(2) 233 -0.73 (0.12)  
 B(3) 234 -0.25 (0.09)  
 B(4) 235 0.62 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.234 0.253 0.269 0.283 0.294 0.302 0.309 0.313  
 -1.4 - 0.0 0.316 0.317 0.318 0.318 0.317 0.315 0.313 0.309  
 0.2 - 1.6 0.303 0.295 0.286 0.273 0.258 0.241 0.222 0.202  
 1.8 - 3.0 0.182 0.161 0.142 0.123 0.106 0.091 0.077

## OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN

CATEGORY(K): 1 2 3 4 5

OBS. FREQ. 22 20 51 76 187  
 OBS. PROP. 0.0618 0.0562 0.1433 0.2135 0.5253  
 EXP. PROP. 0.1213 0.0917 0.1911 0.2192 0.3767

## ITEM 48: 5 GRADED CATEGORIES

P(#) ESTIMATE (S.E.)

A 236 1.98 (0.15)  
 B(1) 237 -1.08 (0.15)  
 B(2) 238 -0.71 (0.12)  
 B(3) 239 -0.21 (0.08)  
 B(4) 240 0.59 (0.07)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)

-3.0 - -1.6 0.091 0.129 0.182 0.252 0.343 0.455 0.585 0.723  
 -1.4 - 0.0 0.854 0.964 1.044 1.092 1.114 1.119 1.111 1.097  
 0.2 - 1.6 1.075 1.043 0.988 0.903 0.790 0.659 0.526 0.404  
 1.8 - 3.0 0.301 0.220 0.157 0.111 0.078 0.054 0.038

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 70 43 88 188 398  
 OBS. PROP. 0.0889 0.0546 0.1118 0.2389 0.5057  
 EXP. PROP. 0.2133 0.0867 0.1384 0.2313 0.3303

ITEM 49: 5 GRADED CATEGORIES  
 P(#) ESTIMATE (S.E.)

A 241 1.89 (0.23)  
 B(1) 242 -1.31 (0.15)  
 B(2) 243 -1.12 (0.12)  
 B(3) 244 -0.80 (0.09)  
 B(4) 245 -0.11 (0.06)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.120 0.167 0.230 0.311 0.411 0.528 0.653 0.775  
 -1.4 - 0.0 0.882 0.963 1.015 1.043 1.053 1.052 1.043 1.031  
 0.2 - 1.6 1.015 0.994 0.959 0.902 0.816 0.707 0.585 0.465  
 1.8 - 3.0 0.358 0.267 0.196 0.141 0.101 0.071 0.050

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 25 20 43 91 177  
 OBS. PROP. 0.0702 0.0562 0.1208 0.2556 0.4972  
 EXP. PROP. 0.1881 0.0971 0.1656 0.2488 0.3003

ITEM 50: 5 GRADED CATEGORIES  
 P(#) ESTIMATE (S.E.)

A 246 1.13 (0.00)  
 B(1) 247 -1.39 (0.00)  
 B(2) 248 -0.41 (0.00)  
 B(3) 249 0.41 (0.00)  
 B(4) 250 1.39 (0.00)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 0.167 0.193 0.221 0.248 0.275 0.300 0.322 0.341  
 -1.4 - 0.0 0.356 0.368 0.376 0.382 0.385 0.386 0.386 0.383  
 0.2 - 1.6 0.377 0.369 0.357 0.342 0.323 0.301 0.276 0.249  
 1.8 - 3.0 0.221 0.194 0.168 0.143 0.121 0.101 0.084

OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN  
 CATEGORY(K): 1 2 3 4 5  
 OBS. FREQ. 40 41 44 56 175  
 OBS. PROP. 0.1124 0.1152 0.1236 0.1573 0.4916  
 EXP. PROP. 0.2039 0.1524 0.1416 0.1564 0.3458

ITEM 51: GRP1, N[MU: 0.00 SIGMA: 1.00]  
 P#;(S.E.): 252; (0.00) 253; (0.00)

@THETA: INFORMATION: (Theta values increase in steps of 0.2)  
 -3.0 - -1.6 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000  
 -1.4 - 0.0 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000  
 0.2 - 1.6 1.000 1.000 1.000 1.000 1.000 1.000 1.000 1.000  
 1.8 - 3.0 1.000 1.000 1.000 1.000 1.000 1.000 1.000

## TOTAL TEST INFORMATION

## @THETA: INFORMATION:

-3.0 - -1.6 10.613 13.118 16.149 19.730 23.825 28.280 32.778 36.850  
-1.4 - 0.0 40.027 42.079 43.100 43.361 43.150 42.696 42.051 41.022  
0.2 - 1.6 39.308 36.747 33.405 29.520 25.400 21.351 17.616 14.347  
1.8 - 3.0 11.597 9.352 7.556 6.138 5.030 4.166 3.495

## @THETA: POSTERIOR STANDARD DEVIATION:

-3.0 - -1.6 0.307 0.276 0.249 0.225 0.205 0.188 0.175 0.165  
-1.4 - 0.0 0.158 0.154 0.152 0.152 0.152 0.153 0.154 0.156  
0.2 - 1.6 0.159 0.165 0.173 0.184 0.198 0.216 0.238 0.264  
1.8 - 3.0 0.294 0.327 0.364 0.404 0.446 0.490 0.535

MARGINAL RELIABILITY: 0.9374

NEGATIVE TWICE THE LOGLIKELIHOOD= 53721.2  
(CHI-SQUARE FOR SEVERAL TIMES MORE EXAMINEES THAN CELLS)