



## ÖĞRETMENLİK MESLEĞİ SORUN ALGISI ÖLÇEĞİNİN GEÇERLİK VE GÜVENİRLİĞİNİN KLASİK TEST KURAMINA VE MADDE TEPKİ KURAMINA DAYALI OLARAK KARŞILAŞTIRILMASI

Erol KARACA\*

### Öz

Bu çalışmanın temel amacı Klasik Test Kuramı (KTK) ve Madde Tepki Kuramı (MTK) altında Öğretmenlik Mesleği Sorun Algısı Ölçeğinin (ÖMSAÖ) kestirilen madde ve test parametrelerinin karşılaştırılarak, araştırmacılara hangi kuram altında kestirilen parametrelerin daha keskin ve güvenilir olduğunu sunmaktır. Bu amaçla çalışmada, öğretmenlerin mesleki sorunlarının saptanmasında kullanılan Likert tipi dereceleme ölçeğinden elde edilen ölçmelerin KTK ve MTK dayalı olarak geçerlik ve güvenilirlik düzeyleri saptanmış ve birbirleriyle karşılaştırılmıştır. Araştırmanın çalışma grubunu ise, Milli Eğitim Bakanlığı'na bağlı ilköğretim ve ortaöğretim kurumlarında görev yapan 6 ildeki (Kütahya (160), Bursa (38), Burdur (31), Ankara (25), Eskişehir (5) ve Balıkesir (3) öğretmenlerden basit tesadüfi örneklem yoluyla seçilmiş 274 öğretmen oluşturmaktadır. 12 ölçek ise, yönergeye uygun şekilde doldurulmadığından analiz dışında tutulmuştur. Öğretmenlerin mesleki sorunlarının ölçülmesi amacıyla 24 maddeden oluşan 5'li Likert tipi bir ölçek kullanılmıştır. Ölçek, KTK ve MTK yöntemiyle puanlanmış ve elde edilen puanların karşılaştırması yapılmıştır. Sonuç olarak her iki kuram altında kestirilen madde ve test parametrelerinin yüksek düzeyde ilişkili olduğu, model varsayımların karşılandığı ve sınırlılıkların göz önüne alındığı durumlarda her iki kuramın da kullanılabilceği sonucuna ulaşılmıştır.

**Anahtar Kelimeler:** Öğretmenlik mesleği sorun algısı, klasik test kuramı, madde tepki kuramı, aşamalı tepki modeli ve test geliştirme.

### THE COMPARISON OF THE RELIABILITY AND VALIDITY OF THE SCALE OF THE PROBLEM PERCEPTION OF TEACHING PROFESSION BASED ON THE CLASSICAL TEST THEORY AND ITEM RESPONSE THEORY

### Abstract

The primary objective of the present research was to investigate the test and item parameters Scale of the Problem Perception of Teaching Profession (SPPTP) under Classical Test Theory (CTT) and Item Response Theory (IRT) and submit the best model which was fitted to real data. For this purpose in this study, the levels of reliability and validity of the measurements which were obtained from the Likert type, rating scale used for determining the perceptions of the teachers towards the professional problems based on CTT and IRT and these levels were compared among them. The research was carried out with 262 teachers, working in primary and secondary schools depending on the Ministry of education in 6 cities (Kütahya (160), Bursa (38), Burdur (31), Ankara (25), Eskişehir (5) ve Balıkesir (3), Turkey, in the 2009. Twelve questionnaires were omitted, because they were not filled out according to the instructions. Participants were 262 teachers working in these cities. By the aim of the measuring, the teachers' professional problems, a 5-point Likert-type scale that was composed of 24 items, was used. Scoring has been made according to CTT and IRT, and then scores have been compared. As a result, it can be concluded that item and test parameters were highly correlated and researchers could use two methods if their assumptions were met and limitations were considered.

**Keywords:** The problem perception of teaching profession, classical test theory, item response theory, graded response model and test development.

\* Dr., Anadolu Üniversitesi Sağlık Bilimleri Fakültesi Sosyal Hizmet Bölümü Öğretim Üyesi



## GİRİŞ

Geniş anlamı ile ölçme, herhangi bir niteliğin gözlenmesi ve gözlem sonucunun sayılarla ya da başka sembollerle ifade edilmesidir (Tekin, 1977; Gelbal, 2004). Genel olarak değerlendirme ise, ölçme sonuçlarının bir ölçütü veya ölçütler takımıyla kıyaslanıp bir karara varılması sürecidir (Gelbal, 2004; Demirel, 2005). Başka bir ifade ile değerlendirme, ölçme sonuçlarını bir ölçüte vurarak ölçülen nitelik hakkında bir değer yargısına varma sürecidir (Gelbal, 1994; Gelbal, 2004). Değerlendirmenin en iyi şekilde yapılabilmesi için ölçmelerde gözlenen bir özelliğin gerçek değerini bulmak önemlidir. Böyle olmakla birlikte ölçmeye karışan çeşitli hatalar yüzünden gerçek değer doğrudan elde edilemez ve gözlenen ölçme sonuçları yardımıyla kestirilmeye çalışılır. Bu kestirim test geliştirmede nihai testin istatistiklerinin kestirilmesi aşamasında yapılır.

Test geliştirme, öğrenci grubuna uygulanan test maddelerinin yeterince güvenilir ve geçerli olup olmadığı, iyi çalışıp çalışmadığı, bilenler ile bilmeyenleri ayırıp ayırmadığı, güçlüğü ve kolaylığı gibi özelliklerin belirlenip, bu özelliklerden yararlanılarak istenilen niteliklerde bir ölçme aracının hazırlanması işidir (Gelbal, 2004; Demirel 2005). Test geliştirme süreci, geliştirilen testin gerçek amaç için kullanılmasına kadar bir takım aşamaların yerine getirilmesini gerektiren bir süreçtir. Bir testin geliştirilmesiyle ilgili temel aşamalar, aşağıdaki gibi sıralanabilir (Baykul, 2000):

- Testin kimlere, niçin uygulanacağını ve uygulamadan elde edilen puanların hangi amaç ya da amaçlarla kullanılacağını saptama,
- Test ile hangi özelliklerin ölçüleceğini belirleme,
- Testte kullanılacak ve ölçülecek özelliğe en uygun madde tiplerini belirleme ve bu yönde madde yazma ya da sağlama,
- Test maddelerini dil, psikometrik ve bilimsel denetimden geçirme, test maddelerini düzeltme, amaca uygun maddeleri belirleme,
- Deneme formunu hazırlama,
- Deneme uygulamasını gerçekleştirme,
- Deneme uygulaması sonuçlarını puanlama, test istatistiklerini hesaplama, maddeleri analiz etme, madde seçimi ve nihai testi oluşturma,
- Nihai testin istatistiklerinin kestirilmesidir.

Bir testin geliştirilmesiyle ilgili bu temel aşamalardan yararlanarak geliştirilen testlerle yapılan ölçmelerdeki hata miktarı azalacak ve bu ölçme sonuçları kullanılarak alınan kararlarda isabetlilik derecesi artacaktır (Gelbal, 2004). Öğrencilere uygulanan testin, maddeleri ve seçenekleri üzerinde yapılan çalışmaya madde analizi denilmektedir. Madde analizi, madde istatistiklerinin hesaplanması, doğrudan teste konulabilecek maddelerin seçilmesi, düzeltilerek teste konulabilecek maddelerin belirlenmesi ve bu maddeler üzerinde yapılacak düzeltme çalışmalarının ne doğrultuda olacağını saptanması, teste konulması mümkün olmayan maddelerin ayıklanması amacıyla yapılır (Baykul, 2000). Başka bir deyişle madde analizi ile madde ayırıcılık gücü indeksine ve testin amacına göre istenen güçlük düzeyinde maddelerin seçilerek, güvenilir ve geçerli bir testin oluşturulması hedeflenir. İstenilen özelliklerdeki maddelerin seçilerek, istenilen özellikte test oluşturulması amacıyla yapılan madde analiziyle maddenin özellikleri olarak kabul edilen, maddeyi betimleyen istatistikler hesaplanır. Bu istatistikler, test kuramında kabul edilen yöntemlere göre değişirler.



Test geliştirmede madde ve test istatistiklerini kestirme aşamasında kullanılan başlıca iki yöntem vardır. Bunlardan birincisi “Klasik Test Kuramı” (KTK), ikincisi ise “Örtük Özellikler Kuramı” ya da “Madde Tepki Kuramı” (MTK)’dır (Baykul, 2000).

Spearman tarafından 1900’lü yılların başında geliştirilen ve psikometri alanında test ve ölçek geliştirme çalışmalarının büyük bir kısmının kuramsal dayanağını oluşturan KTK, testten alınan gözlenen puanı, gerçek puan ve ölçme hatası ile açıklayan basit bir kuramdır. KTK dayalı yöntemler, testi uygulama ve parametreleri kestirme kolaylığı, daha az varsayım gerektirmesi ve üzerinde daha çok çalışılmış olmasından dolayı MTK’ya dayanan yöntemlere göre daha yaygın kullanılmaktadır (Hambleton ve Swaminathan, 1989; Kelecioğlu, 2013: 104; Köse, 2015: 185). Ancak KTK’ya dayalı olarak test geliştirme ve uygulamasında elde edilen, madde ve test istatistiklerinin, testin uygulandığı grubun yetenek düzeyine göre değişmesi, alternatif kuram arayışlarına yol açan önemli bir sınırlılık olmuştur. Bu kuramın bu sınırlılıklarına alternatif olarak 20. yüzyılın ortalarında adını duyuran, “Örtük Özellikler Kuramı” ya da “Madde Tepki Kuramı” olarak adlandırılan yeni bir kuram geliştirilmiştir.

Spearman tarafından temel taşları oluşturulan KTK’ya alternatif olarak geliştirilen MTK’ya; 1946 yılında Tucker, MTK’nin en önemli kavramlarından biri olan “Madde Özellik Eğrisini”, 1950’li yıllarda Lord “Normal Ogive Modeli”ni, 1970’li yıllarda Rasch “Rasch Modeli”ni geliştirerek destek vermişlerdir (Erkin, 1994). Tüm bu gelişmeler sonucunda madde ve test istatistiklerinin kestiriminde yeni bir yöntem olarak MTK’ya dayalı çalışma sayısında bir artış görülmektedir (Çıkrıkçı-Demirtaşlı, 1998). Günümüzde KTK yaygın olarak kullanılmakla birlikte, MTK giderek daha yaygın ve tercih edilir olmaya başlamıştır (Hambleton, 1994; Reise, Ainsworth ve Haviland, 2005).

KTK ve MTK kuramları karşılaştırıldığında ortaya çıkan sonuçlar Tablo 2’de verilmiştir (Baykul, 2000; Çıkrıkçı -Demirtaşlı, 1998; Kelecioğlu, 2001):

**Tablo 2: KTK ile MTK’nin Karşılaştırması.**

<b>Klasik Test Kuramı</b>	<b>Madde Tepki Kuramı</b>
Gerçek puan kavramı gözlenemez	Örtük özelliktir ki bunda da gerçek puan kavramı gözlenemez ancak uyarıcılara verilen tepkilerden dolaylı olarak ölçülür.
Farklı gruplara ait sonuçların karşılaştırılması söz konusu olmaz.	Farklı gruplara ait sonuçların karşılaştırılması mümkündür.
Örneklem sayısının çok büyük olmaması nedeniyle uygulanması daha pratiktir.	Örneklem sayısının çok büyük olması nedeniyle sınıfta uygulanması çok mümkün olmadığından uygulanması pratik değildir.
Madde seçimi yapılırken güvenilirliğin ve geçerliğin yüksek olması istenir.	Ölçüt olarak verinin modele uygunluğuna bakılır.
Doğrusal modelleri kullanır. Kestirim yöntemi olarak madde ve test parametrelerini kullanır.	Doğrusal olmayan modelleri kullanır. Kestirim yöntemi olarak madde ve test parametrelerini kullanır.
Parametre kestirimleri; model hatasını minimize eden regresyonda kullanılan “en küçük kareler yöntemi” adı verilen yöntem ile elde edilir.	Parametre kestirimleri; bilgi kuramına dayalıdır ve bilgiyi maksimize eden “en çok olabilirlik” (likelihood) kestirim yöntemi ile elde edilir.
Normallik varsayımına dayalıdır ve bu varsayım istatistikte yalnızca doğrusal modellerde geçerlidir.	Normallik varsayımı yoktur.
KTK’na dayalı yöntemler, testi uygulama ve parametreleri kestirme kolaylığı, daha az varsayım gerektirmesi ve üzerinde daha çok çalışılmış	KTK’na göre daha az kullanılmaktadır. Ancak 1950’li yıllardan itibaren geliştirilen MTK’na dayalı yöntemler de giderek yaygınlık kazanmaktadır.



olmasından dolayı daha yaygın kullanılmaktadır.	
Bireylerin puanları, onlara uygulanan teste bağımlıdır. Örneğin, kolay bir testten 50 puan alan bir öğrenci ile daha zor bir testten 50 puan alan öğrencinin puanlarının karşılaştırılması mümkün değildir.	Bireylerin belli bir alandaki doğrudan gözlenemeyen yetenekleri ya da özellikleri ile bu alanı yoklayan sorulardan oluşan test maddelerine verdikleri cevaplar arasında bir ilişki vardır ve bu ilişki matematiksel olarak ifade edilebilir. ÖÖT'ne göre geliştirilen testlerden elde edilen yetenek ölçüleri, bireylere uygulanan testlerden bağımsız olarak elde edilebilmektedir. Bu özellik, test puanları eşitlendiğinde, bireylerin yeteneklerinin gruptan bağımsız olarak karşılaştırılmasını sağlamaktadır.
Elde edilen madde istatistikleri, testin uygulandığı grubun özelliklerine bağlıdır.	Madde özellikleri matematiksel modeller esas alınarak hesaplandığından, testin uygulandığı gruba bağlı olmadığı iddia edilmektedir. Bu durum, bir testin MTK'na göre bir kez ölçeklendikten sonra bu maddelerin özellikleri değişmediğinden pek çok kez kullanılmasına olanak sağlar. Ancak bu değişmezliğin sağlanması madde parametrelerinin elde edilmesinde yapılan deneme uygulamasının ve bu uygulamanın yapıldığı grubun bazı şartları sağlanmasına bağlıdır.
Madde parametreleri gruptan bağımsız değildir. Dolayısıyla testlerin geçerlik ve güvenilirlikleri MTK'ya göre kararlı olmadığı için istenen yetenek düzeyindeki bireylerin seçimindeki geçerlilik göreceli olarak MTK'ya göre daha düşüktür.	Madde parametrelerinin gruptan bağımsız olması, testlerin geçerlik ve güvenilirliklerinin daha kararlı olarak belirlenebilmesini sağlamaktadır. Ayrıca istenen yetenek düzeyindeki bireylerin seçimi, KTK'na göre daha geçerli olarak yapılabilmektedir. Ancak belli bir matematiksel modele bağlı olduğundan, KTK'na göre daha güçlü ve yerine getirilmesi daha güç varsayımları vardır.
KTK'na dayalı madde analizinde 1 ve 0 ile puanlanan maddeler için madde güçlük indeksi (p), maddeyi doğru cevaplandıranların tüm cevaplayıcılara oranı olarak tanımlanır ve bu oran da 0 ve 1 arasında değer alır. Madde kolaylaştıkça güçlük indeksinin değeri büyür yani 1'e yaklaşır. Madde zorlaştıkça da madde güçlük indeksinin değeri küçülür yani 0'a yaklaşır. Madde güçlü indeksi aynı zamanda o maddeden alınan puanların aritmetik ortalamasını ifade eder. Bu indeks maddeyi cevaplandıranlardan tesadüfi olarak seçilen bir bireyin o maddeyi doğru cevaplandırma olasılığını da belirtir. Madde güçlük indeksi olarak p değerleri yerine standart z puanları da kullanılabilir. Bir maddeye cevap veren bireylerin yetenekleri, çok düşük e çok yüksek düzeyler arasında değişen bir dağılım gösterir ve bu dağılımın normal olduğu kabul edilir.	Normal dağılım varsayımı altında KTK'ndan MTK'na bir geçiş mümkündür. MTK'nda yer alan ve madde indeksine karşılık gelen b parametresi de, b değerleri ile O aynı ölçekte olduğundan örtük özelliğin düzeyini gösterir ve (-&, +&) aralığında değerler aldığı kabul edilir. Bu parametre, maddenin doğru cevaplandırılma olasılığının 0.50 olduğu noktadaki yetenek düzeyini belirtir. B değeri madde kolaylaştıkça düşer, madde zorlaştıkça yükselir.
KTK'nda madde ayırıcılık gücü indeksi, maddenin, sorunun ölçtüğü özelliği kazanmış bireyleri diğerlerinden ayırma gücünü gösterir ve maddenin geçerliliğini ifade eder. Bu değer, madde puanları ile test puanlarının korelasyonudur.	MTK'nda madde ayırıcılık gücü indeksine karşılık gelen a parametresi de maddenin kalitesini, yani elde edilen O yetenek ölçüsünün gerçek O hakkında ne kadar bilgi vermekte olduğunu gösterir. Bu bakımdan bir tür geçerlik ölçüsüdür.
KTK ve MTK'nın ortak varsayımı ise test puanlarının normal dağılım göstermesidir. Bu varsayım iki kuramın ortak yanını oluşturmaktadır. Normal dağılım şartının sağlanması durumunda MTK'ndaki a ve b	



parametreleri ile KTK'ndaki madde güçlük indeksi p ve madde ayıricılık gücü indeksi r arasında geçiş sağlanabilmektedir. Yani dağılım normal olduğunda KTK'ndaki madde istatistiklerinden MTK'ndaki b ve a parametrelerini kestirmek mümkün olabilmektedir.

KTK, MTK ve Genellenebilirlik Kuramının üçü de ölçme araçlarının ve elde edilen ölçme sonuçlarının hatasız olması ile ilgilenmektedir.

MTK'da, eğitimsel ve psikolojik test verilerinin analizinde kullanılan modeller, gözlenen parametrelerle gözlenemeyen parametreler yani test maddesine verilen cevap ile ölçme konusu olan psikolojik özellik arasındaki ilişkiyi tanımlayan olasılık modelleridir. Burada sözü edilen olasılık boyutu, modelin matematiksel formunun getirdiği bir boyuttur (Çıkrıkçı-Demirtaşlı, 1998).

MTK'nın ortaya çıktığı günden günümüze değin sayıları giderek artan çeşitli modeller geliştirilmiştir. Modeller, Rasch Modeli, Samejima Derecelendirilmiş Model gibi modeli geliştiren kişinin adıyla anıldığı gibi Normal Ogive Model, Lojistik Model gibi modelin matematiksel fonksiyonu esas alınarak da isimlendirilmektedir. Değişik adlar alan Kuramdaki her model; cevaplayıcının maddedeki edimi ile psikolojik özellik (yetenek) arasındaki ilişkiyi veren matematiksel bir eşitlikten ve MTK çerçevesinde belirlenmiş, tanımlanmış temel varsayımlardan oluşur (Hambleton ve Swaminahan, 1985; Çıkrıkçı-Demirtaşlı, 1998).

Her ne kadar yapılan çalışmaların büyük bir çoğunluğu ikili puanlanan modellerden elde edilen veri analizine yoğunlaşmış olsa da; kişilik, tutum ve algı gibi psikolojik özellikleri ölçen ve çoklu puanlanan maddelere sahip ölçme araçları için çoklu puanlanan madde tepki modelleri geliştirilmiştir. Bu modeller, cevaplayıcının yetenek düzeyi ile belli bir kategoride tepki verme arasında doğrusal olmayan ilişkiler kuran modellerdir (Emretson ve Reise, 2000). Günümüzde yaygın olarak kullanılan Samejima (1969)'nın aşamalı tepki modeli de bu modellerden biridir. Bu modelin en büyük avantajı, ikili puanlanan modellerle karşılaştırıldığında bireylerin yeteneği hakkında daha fazla bilgi alınabilmesidir (Koch, 1983; Köse, 2015: 186; Uyar, Gübeş ve Kelecioğlu, 2013: 126).

MTK'ya dayalı ölçme kuram ve uygulamaları, sayılan sınırlılıklarına rağmen temelinde matematiğin getirdiği kuramsal zenginliği barındırdığı için gelişmeye devam etmektedir. Bu gelişmeler çerçevesinde, ülkemizde bu kuramla yapılacak psikometri ile ilgili çalışma ve araştırmaların özellikle psikoloji ve eğitim boyutunda hız ve ağırlık kazanması beklenmektedir.

Çalışmanın amacı, öğretmenlerin mesleki sorunlarının saptanmasında kullanılan Likert tipi Öğretmenlik Mesleği Sorun Algısı Ölçeğinden (ÖMSAÖ) elde edilen test istatistiklerinin Klasik Test Kuramı (KTK) ve Madde Tepki Kuramına (MTK) dayalı olarak geçerlik ve güvenilirlik düzeylerini saptamak ve birbiriyle karşılaştırmaktır. Bu genel amaç doğrultusunda aşağıdaki alt amaçlara yanıt aranmıştır:

- KTK'ya göre hesaplanan ve MTK'ya dayalı olarak kestirilen mesleki sorun algısı puanları arasında anlamlı bir ilişki var mıdır?
- KTK ve MTK'ya dayalı olarak yapılan madde analizleri sonucunda, düzeltilmeden tekrar kullanılmaması gereken madde sayısı kaçtır?
- Madde analizi sonucunda testte kalan maddeler üzerinden KTK ve MTK'ya dayalı olarak elde edilen mesleki sorun algısı puanları arasında anlamlı bir ilişki var mıdır?



## YÖNTEM

**Araştırma Modeli:** Araştırma, mesleki sorun algılarının değerlendirilmesi çerçevesinde puanlayıcıların tutarlılığı ile maddelerin güçlüğüne hesaplanmasında KTK ve MTK'nın yarar ve sakıncalarının ortaya konmasına, kuramlardan hangisinin daha çok bilgi verdiğinin saptanmasına ve iki kurama dayalı olarak elde edilen geçerlik ve güvenilirlik katsayılarının karşılaştırmasına dayalıdır. Bu bağlamda yapılan çalışma, karşılaştırmalı, betimsel düzeyde kuramsal bir araştırmadır.

**Çalışma Grubu:** Araştırmanın çalışma grubunu ise, Milli Eğitim Bakanlığı'na bağlı ilköğretim ve ortaöğretim kurumlarında görev yapan 6 ildeki öğretmenlerden basit tesadüfi örneklem yoluyla seçilmiş 274 öğretmen oluşturmaktadır. 12 ölçek ise, yönergeye uygun şekilde doldurulmadığından analiz dışında tutulmuştur. Katılımcıların % 33.20'si (87) sınıf öğretmeni, %17.90'ı (47) matematik öğretmeni, %13.40'ı (35) Türkçe öğretmeni, % 6.90'ı (18) tarih öğretmeni, % 5.70'i (15) kimya öğretmeni, % 4.60'ı (12) İngilizce öğretmeni, % 3.10'u (8) fizik öğretmeni, % 2.70'i (7) biyoloji öğretmeni, % 1.50'si (4) sosyal bilgiler öğretmeni, Geriye kalan 23 (%9.07) öğretmenin branşlarının ise, din kültürü ve ahlak bilgisi (4), görsel sanatlar (2), Türk dili ve edebiyatı (1), coğrafya (3), felsefe (3), rehberlik (2), beden eğitimi (1), bilgisayar teknolojisi (3), elektrik (1), teknoloji ve tasarım (3) olduğu görülmektedir. Buna göre uygulanan toplam 274 anketin, 262'sinden elde edilen veriler analize dâhil edilmiştir.

**Ölçme Aracı:** Araştırmada öğretmenlik mesleği ve sorunları ile ilgili alanyazının taraması ile yurtdışında ve yurtiçinde öğretmenlik mesleği ile ilgili araştırmalardaki ölçme araçları ve çeşitli mesleklerde karşılaşılan sorunlarla ilgili ölçeklerin gözden geçirilmesi sonucu hazırlanan Karaca (2011) tarafından geliştirilen 24 maddeden oluşan ÖMSAÖ kullanılmıştır. Ölçekteki algı maddelerine verilecek tepkiler için beşli dereceleme tercih edilmiştir. Katılımcılardan ölçekte yer alan her bir algı ifadesini “*kesinlikle katılmıyorum – katılmıyorum – kararsızım – katılıyorum - kesinlikle katılıyorum*” gibi beş kategoriden biriyle sınıflaması istenmiştir. Her katılımcı için toplam puanın elde edilebilmesi için, en olumlu kategoriye 5 puan, en olumsuz kategoriye 1 puan verilerek toplanan cevaplar 1–5 arasında puanlanmıştır (Turgut, 1977: 10, 11). Ölçekten alınabilecek en düşük puan 24, en yüksek puan ise 120'dir. ÖMSAÖ'nin öğretmenlerin mesleki sorunlarına ilişkin algı maddeleri yedi boyutta yer almıştır. Boyutlar, maddelerin içerdiği anlamlar dikkate alınarak adlandırılmaya çalışılmıştır. Birinci boyuta, “Çalışma Koşulları ve Ücret”; ikinci boyuta, “Çevresel Koşullar”; üçüncü boyuta, “Mesleğin Yürütümü Koşulları”; dördüncü boyuta, “İşe Yerleştirme ve İstihdam Politikası”; beşinci boyuta, “Eğitim Politikası”; altıncı boyuta “Yaşam Kalitesi” ve yedinci boyuta ise “Yasal Düzenlemeler” adı verilmiştir.

**İşlem:** Ölçek araştırmacı tarafından gerekli izinler alınarak uygulanmıştır. Uygulama öncesinde araştırmanın amacı öğretmenlere açıklanmış ve konu ile ilgili sorular yanıtlanmıştır. Ölçeğin uygulanması 30-45 dakika almaktadır. Uygulama sonucunda toplanan veriler SPSS 12.0 programına aktarılmış ve veriler kontrol edilmiştir. Kontrol işlemi sonucunda, 12 öğretmenin ölçeği eksik doldurması nedeniyle cevapları veri grubundan çıkarılmıştır.



**Verilerin Analizi:** ÖMSAÖ'den elde edilen verilerin analizinde KTK altında yapılan analizler SPSS ve LISREL yazılım programı, MTK altında yapılan analizler ise MULTILOG programı ile analiz edilmiştir. KTK altında, madde ayırıcılık parametresi için, madde toplam korelasyonlarından, madde güçlük parametresi için madde ortalamalarından yararlanılmıştır. ÖMSAÖ'nün iç tutarlık anlamında güvenilirliği için  $\alpha$  katsayısı esas alınmıştır. MTK altında madde ve test parametrelerinin kestiriminde, farklı ayırıcılık parametrelerine sahip maddeleri barındıran ölçekler için uygun olan ve en çok kullanılan MTK modeli olan Smejima (1969)'ın Aşamalı Tepki Modeli (ATM) kullanılmıştır.

## BULGULAR VE YORUM

Araştırmanın amacına uygun olarak öncelikle KTK altında ölçekteki maddelerin madde güçlük ve ayırt edicilik değerleri hesaplanmış ve Tablo 1'de gösterilmiştir.

**Tablo 1:** KTK'ya Göre Hesaplanan Madde Güçlük ve Madde Ayırıcılık Parametre Değerleri.

Madde	$\alpha$	b
1	0,880	4,13
2	0,880	3,89
3	0,882	3,55
4	0,879	3,91
5	0,881	3,71
6	0,879	4,26
7	0,881	4,03
8	0,879	4,19
9	0,880	4,10
10	0,881	4,05
11	0,878	4,18
12	0,881	3,80
13	0,879	4,11
14	0,881	3,92
15	0,879	4,24
16	0,879	4,08
17	0,880	4,03
18	0,879	4,05
19	0,881	3,98
20	0,878	4,23
21	0,881	4,04
22	0,880	4,19



## AKADEMİK BAKIŞ DERGİSİ

Sayı: 64 Kasım – Aralık 2017

Uluslararası Hakemli Sosyal Bilimler E-Dergisi

ISSN:1694-528X Calal-Abad Uluslararası Üniversitesi,

Türk Dünyası Kırgız – Türk Sosyal Bilimler Enstitüsü

Calal-Abad – KIRGIZİSTAN

<http://www.akademikbakis.org>



23	0,880	3,73
24	0,881	3,90

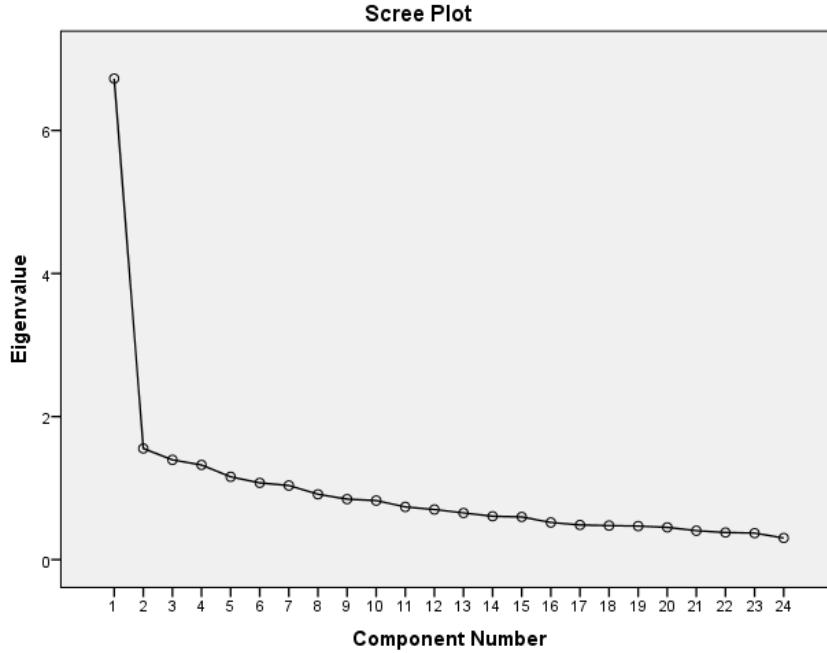
Dereceli puanlanan maddelerde madde güçlüğü tepkilerin ortalaması olarak kabul edilir. KTK’da madde ayırt edicilik değerleri ise, madde toplam korelasyonlarına karşılık gelmektedir. Yüksek ayırt ediciliğe sahip maddeler, ölçeğin ölçtüğü özelliği temsil ettiğini göstermektedir. Ölçeğin iç tutarlık katsayısı ise  $\alpha$  katsayısı ile hesaplanmıştır.

Tablo 1’e göre ölçekteki maddelerin ortalama güçlüğü 4.01’dir. Bu değer ortalama güçlüğü karşılık gelmektedir. Ölçeğin iç tutarlık anlamında hesaplanan  $\alpha$  katsayısı ise 0.89’dur. KTK altında analiz edilen 24 maddenin madde ayırt edicilik değerleri incelendiğinde ise, 24 maddenin tümünün ayırt edicilik değerinin 0,80’in üzerinde olduğu görülmektedir. Bu bulgu ölçekteki tüm maddelerin ölçeğin ölçtüğü özelliği temsil etmede yeterli olduğunu ortaya koymaktadır.

Araştırmanın amacına uygun olarak ikinci aşamada KTK üzerinde önemli avantajları olmasına rağmen, MTK’nın tek boyutluluk ve yerel bağımsızlık varsayımlarını sağlayıp sağlamadığı incelenmiştir. Bu amaçla veri setine sırasıyla şu işlemler uygulanmıştır:

- Açıklayıcı Faktör Analizi (AFA): Faktör analizi, değişkenlerin faktörlerle nasıl ilişkili olduğunu keşfetmek ya da değişkenlerin bir faktör altında olup olmadığını doğrulamak amacıyla kullanılır. Araştırmanın temel amacı keşfetmek ise AFA kullanılmalıdır (Schumacker ve Lomax, 2004). AFA sonucu elde edilen özdeğeri 1’den büyük yedi faktörden ilkinin tek başına toplam varyansın %28,17’sini açıkladığı görülmüştür. İlk faktörün açıkladığı varyansın %20 ve üzerinde olduğu durumda, veri grubundaki başat boyut olduğuna karar verilebilir (Lee, 1995). Şekil 1’de özdeğer grafiği ile veri grubunun faktör yapısı gösterilmektedir. Özdeğer grafiği özellikle başat faktörün varlığını ortaya çıkarmada etkili bir yöntemdir (Hambleton ve Swaminathan, 1985). Bu grafik ise baskın bir şekilde tek faktörde düşüş sergilemiştir. Döndürülmüş bileşenler matrisindeki faktör yükleri incelendiğinde (0,45-0,60) algı ölçeğini oluşturan maddelerin tek faktör altında toplanabileceği gözlenmiştir.





**Şekil 1: Özdeğer Grafiği.**

- Doğrulayıcı Faktör Analizi (DFA): Araştırmanın temel amacı doğrulamak ise DFA kullanılmalıdır (Schumacker ve Lomax, 2004). Bu çalışmada ölçeğin tek boyutlu yapıyı doğrulayıp doğrulamadığını incelemek amacıyla aynı zamanda LISREL yazılım programında DFA yapılmıştır. DFA sonuçlarına göre model-veri uyumu Yaklaşık Hataların Ortalama Karakökü (RMSEA), Standardize Edilmiş Artık Ortalamaların Karekökü (RMR), İyi Uyum İndeksi (GFI), Düzenlenmiş İyi Uyum İndeksi (AGFI), Normleştirilmemiş Uyum İndeksi (NNFI) ve Karşılaştırmalı Uyum İndeksi (CFI) değerleri kullanılarak değerlendirilmiştir. RMSEA ve RMR 0-1 arasında değer almaktadır. RMSEA ve RMR'nin aldığı değer 0.05'e eşit veya daha küçük ise uyum mükemmel, 0.08 ve altında ise uyum kabul edilebilir ve 0.10 ve daha büyük ise uyum kötüdür (Çokluk, Şekercioğlu ve Büyüköztürk, 2010). Diğer uyum indekslerinden GFI ve AGFI, 0-1 arasında değişen değerler alır. GFI ve AGFI değeri 0.95 ve üzerinde ise uyum mükemmel, 0.90-0.94 arasında ise uyum kabul edilebilirdir (Schumacker ve Lomax, 2004; Hooper, Caughlan ve Mullen, 2008). NNFI ve CFI değerleri 0.95'in üzerinde ise uyum mükemmel, .90'ın üzerinde ise uyum kabul edilebilirdir (Sümer, 2000). Modele ait bu indeksler birlikte değerlendirildiğinde tek faktör modeli kabul edilmiştir (Kay-Kare= 190.99; RMSEA=0.053; RMR=0.053; GFI=0.89; AGFI=0.85; NNFI=0.89; CFI=0.91). AFA ve DFA sonucu elde edilen tüm bulgulara dayalı olarak ölçeğin tek boyutlu bir yapıya sahip olduğu söylenebilir.

MTK'ya göre yerel bağımsızlık varsayımının ihlali aynı zamanda tek boyutluluk varsayımının da ihlali demektir. Bu durumda tek boyutluluk varsayımını sağlayan algı ölçeğindeki 24 maddenin yerel bağımsızlık varsayımını da sağladığı kabul edilmiştir.

Araştırmanın amacına uygun olarak MTK'nın tek boyutluluk ve yerel bağımsızlık varsayımlarını sağladığının saptanmasının ardından ÖMSAÖ'nden elde edilen veriler, ATM



altında madde ve test parametrelerinin kestirilmesi için MULTILOG (Thissen, 1991) programıyla analiz edilmiştir. ATM'ye dayalı olarak kestirilen madde parametreleri ise Tablo 2'de gösterilmiştir.

**Tablo 2: Aşamalı Modele Göre Kestirilen Parametre Değerleri.**

Madde	$\alpha_i$ (SE)	$\beta_1$ (SE)	$\beta_2$ (SE)	$\beta_3$ (SE)	$\beta_4$ (SE)
1	0.06 (0.00)	-7.90 (0.00)	19.28 (0.00)	136.04 (0.00)	1.39 (0.00)
2	0.21 (0.51)	-6.01 (4.59)	-1.62 (2.18)	2.12 (2.38)	6.68 (****)
3	0.18 (0.19)	-6.86 (9.90)	-191 (3.04)	2.36 (3.24)	7.58 (****)
4	1.32 (0.21)	-3.72 (0.82)	-1.99 (0.35)	-1.17 (0.24)	0.38 (0.16)
5	1.14 (0.19)	-3.30 (0.65)	-1.93 (0.37)	-1.13 (0.25)	1.15 (0.23)
6	0.85 (0.17)	-3.61 (0.78)	-1.65 (0.42)	-0.40 (0.24)	1.83 (0.38)
7	1.18 (0.19)	-3.64 (0.71)	-1.84 (0.35)	-0.99 (0.23)	1.04 (0.21)
8	1.19 (0.19)	-2.58 (0.50)	-1.44 (0.29)	-0.59 (0.19)	1.11 (0.22)
9	1.75 (0.23)	-3.29 (0.64)	-1.92 (0.29)	-1.39 (0.22)	0.31 (0.13)
10	1.47 (0.23)	-3.04 (0.97)	-1.32 (0.24)	-0.98 (0.20)	0.34 (0.15)
11	1.78 (0.23)	-2.67 (0.48)	-1.83 (0.27)	-1.14 (0.18)	0.29 (0.13)
12	1.41 (0.23)	-3.11 (0.57)	-1.76 (0.31)	-1.27 (0.23)	0.52 (0.16)
13	1.23 (0.19)	-3.20 (0.63)	-1.93 (0.36)	-1.06 (0.23)	0.44 (0.17)
14	1.69 (0.23)	-2.94 (0.54)	-1.91 (0.29)	-1.23 (0.19)	0.42 (0.13)
15	1.25 (0.19)	-2.94 (0.58)	-1.67 (0.30)	-0.72 (0.19)	1.05 (0.21)
16	1.43 (0.21)	-3.76 (0.81)	-1.91 (0.32)	-1.20 (0.22)	0.57 (0.16)
17	1.22 (0.19)	-2.69 (0.52)	-1.68 (0.33)	-0.94 (0.23)	0.62 (0.19)
18	1.53 (0.22)	-2.96 (0.51)	-2.31 (0.36)	-1.46 (0.24)	0.38 (0.14)
19	1.50 (0.21)	-2.66 (0.43)	-1.94 (0.31)	-1.19 (0.21)	0.60 (0.16)
20	1.39 (0.20)	-3.19 (0.63)	-2.06 (0.36)	-1.10 (0.21)	0.77 (0.18)
21	1.61 (.23)	-3.02 (0.58)	-1.78 (0.28)	-0.97 (0.18)	0.59 (0.15)
22	1.39 (0.21)	-3.14 (0.65)	-1.72 (0.32)	-0.80 (0.18)	0.41 (0.17)
23	2.06 (0.28)	-2.66 (0.47)	-1.92 (0.28)	-1.35 (0.18)	0.43 (0.13)
24	1.44 (0.22)	-2.87 (0.54)	-1.72 (0.32)	-1.17 (0.24)	0.55 (0.16)

MTK altında madde ayırıcılık parametresi maddenin örtük özellik sürekliliğinde cevaplayıcıları yeteneklerine göre ayırabilme gücü olarak ifade edilir. Yüksek ayırıcılığa sahip maddeler bireyleri yeteneklerine göre daha iyi ayıran maddeler iken düşük ayırıcılığa sahip maddeler ise bireyleri yetenek düzeyine göre ayırmada yetersiz maddelerdir. Bunun yanında yüksek ayırıcılığa sahip maddeler test bilgi fonksiyonuna daha fazla katkı yapan maddelerdir (Hambleton vd., 1991). Ayırt edicilik parametre değerleri 0.01-0.34 arasında ise madde çok düşük ayırıcılığa; 0.35-0.64 arasında ise düşük ayırıcılığa; 0.65-1.34 arasında ise orta ayırıcılığa; 1.35-1.69 arasında ise yüksek ayırıcılığa ve 1.70 ve daha yukarı bir değer almışsa çok yüksek ayırıcılığa sahiptir (Baker, 2001: 35).

Tablo 2'ye göre 1., 2. ve 3. maddeler düşük ayırıcılığa; 5., 6., 7., 8., 13., 15., ve 17. maddeler orta ayırıcılığa; 4., 10., 12., 14., 16., 18., 19., 20., 21., 22. ve 24. maddeler çok yüksek ayırıcılığa ve 9., 11. ve 23. maddeler çok yüksek ayırıcılığa sahiptir. Tüm maddeler içerisinde ise 23. maddenin en yüksek eğim parametresine ( $\alpha_i=2.06$ ) ve 1. maddenin ise en düşük eğim parametresine ( $\alpha_i=0.06$ ) sahip olduğu görülmektedir. KTK ve ATM altında elde edilen madde ayırıcılık parametreleri arasındaki ilişki test edilmiş ve  $r=0,80$  ( $p<.01$ )'lık yüksek bir

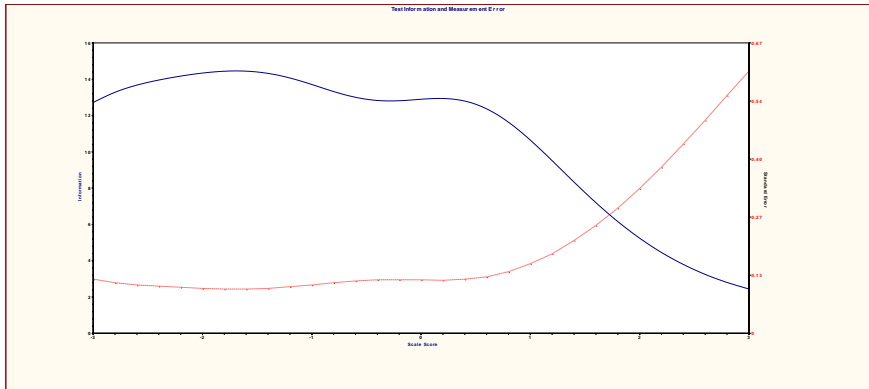


ilişki bulunmuştur. Her iki kuram altında hesaplanan bu ayırıcılık değerlerinin yüksek olması beklenen bir durumdur. Bunun nedeni olarak, ölçekteki maddelerin başat bir boyut altında toplanması gösterilebilir (Köse, 2015: 190).

Tablo 2’de aynı zamanda kategoriler arası eşik değerlerine de ( $\beta_i$ ) yer verilmiştir. ATM’de kategoriler arasında sıralı olan kategoriler arası eşik değerleri, katılımcıların 0.50 olasılıkla eşik değerinin üzerinde cevap vermek için gerekli olan algı düzeyini gösterir (Matteucci ve Stracqualursi, 2006). Buna göre bir katılımcının ilk kategoriye cevaplaması için düşük yetenek ya da algıya sahip olması yeterli olmakta iken, bir üst kategori için daha yüksek algı gerekmektedir. Tablo 2 incelendiğinde, kategoriler arası eşik değerlerinin en düşükten en yükseğe doğru sıralanmış olduğu görülmektedir Buna göre her madde için ( $\beta_1$ ) en düşük eşik parametresine, ( $\beta_4$ ) ise en yüksek eşik parametresine sahiptir. Örneğin; 1. maddenin

$\beta_1 = -7.90$  ve  $\beta_4 = 1.39$  olarak hesaplanmıştır. Bu maddenin 1. kategorisinde %50 olasılıkla tepki vermek için gerekli yetenek düzeyi  $\theta = -7.90$  olarak, 5. Kategoride ise %50 olasılıkla tepki vermek için gerekli yetenek düzeyi ise  $\theta = -1.39$  olarak hesaplanmıştır. Ayrıca eşik parametre değerlerinin çoğunun negatif değerler almasına dayalı olarak kategorilerdeki cevapların daha çok algının düşük düzeyi ( $\theta < 0$ ) tarafından desteklendiği söylenebilir.

Araştırmanın amacına uygun olarak ölçekte bulunan maddelerin test bilgi fonksiyonuna yaptıkları katkılar da hesaplanmıştır. Şekil 2’de gösterilen test bilgi fonksiyonu incelendiğinde, ölçekteki bilginin en çok -3—1.6 yetenek ya da tutum düzeyleri arasında elde edildiğini göstermektedir. Madde kalibrasyon sonuçları incelendiğinde ise, ölçeğin en fazla bilgiyi -1.6 yetenek düzeyinde (bilgi değeri=9.556) verdiği görülmüştür. Ölçeğin verdiği bilgi özellikle 1.6 düzeyinden sonra azalmıştır. Bu durum standart hatanın da artmasına bağlıdır. ATM altında hesaplanan marjinal güvenirlilik değeri 0,86 olarak hesaplanmıştır.



Şekil 2: Test Bilgi Fonksiyonu.

## SONUÇ VE ÖNERİLER

Bireylerin çeşitli objelere, fikirlere ve olaylara ne şekilde tepkide bulunacağını büyük ölçüde tutumları ve algıları tarafından yapılandırıldığı bilinmektedir. Bir bireyin bir konuya yönelik algısının bilinmesi, gelecekte nasıl davranacağı ile ilgili bir kestirmeye olanak tanımaktadır.



Bu nedenle ölçülmesi büyük önem taşımaktadır. Hiç kuşkusuz herhangi bir konuya yönelik algının iyi bir şekilde saptanabilmesi, bu amaçla geliştirilecek olan ölçme araçlarının da nitelikli olmasını gerektirmektedir. Bu bağlamda öğretmenlerin mesleki sorunlarına yönelik algılarının belirlenmesine yönelik ölçme araçlarının niteliğini artırmaya dönük bu çalışmada, ÖMSAÖ'nin maddeleri hem KTK altında hem de çok kategorili MTK modellerinden Likert tipi ölçekler için uygun olan Aşamalı Tepki Modeline göre incelenmiştir. Ölçeğin tek boyutluluğunu belirlemek amacıyla açımlayıcı ve doğrulayıcı faktör analizi yapılmış ve yapılan analizler sonucunda ölçeğin tek boyutlu bir yapıya sahip olduğu görülmüştür. Tek boyutluluk varsayımını sağlayan maddelerin aynı zamanda MTK'nın önemli ilkelerinden biri olan yerel bağımsızlık varsayımını da sağladığı kabul edilmiştir.

Hem KTK hem de MTK altında kestirilen madde ayırıcılık parametreleri incelendiğinde KTK altında kestirilen 24 maddenin Koch (1983)ün 0.60 kriterini karşıladığı sonucuna varılmıştır. MTK altında ATM'ye dayalı olarak madde ayırıcılık parametreleri incelendiğinde ise ilk üç madde hariç ölçekteki diğer maddelerin orta, yüksek ve çok yüksek ayırıcılığa sahip olduğu görülmektedir. Ayrıca madde ayırıcılık parametrelerinin 0.06-2.06 arasında değişim göstermesi de ölçekteki maddelerin farklı yetenek düzeyindeki bireylere hitap edebildiği sonucunu ortaya koymaktadır. Her iki kuram altında elde edilen güvenilirlik değerleri de birbirine yakındır. KTK altında 0,89, MTK altında ise 0,86 güvenilirlik değeri elde edilmiştir. Bu durum kuramlar arasında kestirilen parametreler arasındaki uyuma işaret etmektedir.

Araştırma bulguları, ÖMSAÖ maddelerinin ayırt ediciliğe sahip olduğunu ve önemli bir çoğunluğunun ise yüksek ve çok yüksek ayırt ediciliğe sahip olduğunu ortaya koymuştur. Bu bağlamda öğretmenlerin mesleki sorunlarına yönelik algılarını saptamaya dönük ÖMSAÖ ile elde edilen bulguların geçerli ve güvenilir olduğu söylenebilir.

## KAYNAKÇA

Baker, F. B. (2001). *The basics of item response theory*. ERIC Clearinghouse on Assessment and Evaluation (ERIC Document Reproduction Service No. ED458219).

Çokluk, Ö., Şekercioğlu, G. ve Büyüköztürk, Ş. (2010). *Sosyal Bilimler İçin Çok Değişkenli İstatistik SPP ve LISREL Uygulamaları* (1. Baskı), Ankara: Pegem A Yayıncılık.

Hambleton, R. K., Swaminathan, H. and Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Sage Publications, London.

Karaca, E. (2011). Investigation of the reliability and validity of the perception of teaching profession problems scale with a structural equation model: a research on teachers. *Education*, 132(2), 262-272.

Koch, W. R. (1983). Likert Scaling using the graded response latent trait model. *Applied Psychological Measurement*, 7 (1), 15-32.

Matteucci, M. and Stracqualursi, L. (2006). Student assessment via graded response model. *Statistica*, 4, 435-447.

Schumacker Randy E. and Lomax Richard G. (2004). *A Beginner's Guide to Structural Equation Modeling (second edition)*, London: Lawrence Erlbaum Associates Publishers, 79-122.

Thissen, D. (1988). *MULTILOG (Computer program)*. Mooresville, IN: Scientific Software.



Turgut, M. F. (1977). Tutumların ölçülmesi, *Eğitimde Ölçme Teknikleri Ders Notu*. Ankara: Hacettepe Üniversitesi, No:7.

#### **İNTERNET KAYNAKLARI**

Hooper, D., Caughlan, J., and Mullen, M.R. (2008). Structural Equation Modeling: Guidelines for Determining Model Fit. *Electronic Journal Of Business Research Methods*. 6 (1), 53-60. <https://tr.scribd.com/document/142037436/Structural-Equation-Modelling-Guidelines-for-Determining-Model-Fit>. Erişim Tarihi: Eylül 15, 2016.

Smejima, F. (1969). Estimation of latent ability using a response pattern of graded scores (Psychometric monograph No.17). Richmond, VA: Psychometric society. <https://www.psychometricsociety.org/sites/default/files/pdf/MN17.pdf>. Eylül 15, 2016.