

Sendikaya Bağlılık Ölçeği Test ve Madde Puanlarının Madde Tepki Kuramına Dayalı Aşamalı Tepki Modeli ile İncelenmesi

Examining Test and Item Scores of the Commitment Scale to the Union via Graded Response Model Based on Item Response Theory

Dr. Öğr. Üyesi Erol Karaca

Başvuru Tarihi: 14.11.2016

Kabul Tarihi: 29.11.2018

Öz

Bu çalışmanın temel amacı, Sendikaya Bağlılık Ölçeğinden (SBÖ) elde edilen madde puanlarını Madde Tepki Kuramı (MTK) modeli olan Aşamalı Tepki Modeli (ATM) ile incelemektir. Araştırma verilerinin toplanmasında Gordon ve arkadaşları (1980) tarafından geliştirilen, Bilgin (2003) tarafından Türkçe uyarlama çalışması yapılan SBÖ kullanılmıştır. Araştırmanın çalışma grubunu 2009 yılında Kütahya, İstanbul, Eskişehir, Bursa, İzmir ve Balıkesir illerinde görev yapan basit tesadüfi örneklem yoluyla seçilen 353 öğretmen oluşturmaktadır. SBÖ'de sendikal bağlılığa yönelik tutuma ilişkin 15 madde bulunmaktadır ve bu maddelere öğretmenlerin "kesinlikle katılmıyorum (1)" ve "kesinlikle katılıyorum (5)" gibi arasında tepkide bulunmaları istenmiştir. Tutum maddeleri, çok kategorili MTK modellerinden likert tipi ölçekler için uygun olan ATM kullanılarak incelenmiştir. Araştırma verilerinin analizinde öncelikle veri setinin MTK'nın tek boyutluluk ve yerel bağımsızlık varsayımlarını sağlayıp sağlamadığı incelenmiştir. Araştırma verilerinin analizinin ikinci aşamasında ise MULTILOG programında ATM seçilerek madde kalibrasyonu yapılmıştır. Araştırmanın sonucunda, veri setinin MTK'nın tek boyutluluk ve yerel bağımsızlık varsayımlarını sağladığı ve model-veri uyumunun sağlandığı görülmüştür. Araştırma bulgularına göre sendikaya bağlılık açısından öğretmenleri en iyi ayırt eden ve dolayısıyla en çok bilgi veren

maddenin 4. madde, öğretmenleri en azayırt eden ve dolayısıyla en az bilgi veren maddenin ise 14. madde olduğu görülmektedir.

Anahtar Kelimeler: Madde Tepki Kuramı, Test Geliştirme, Aşamalı Tepki Modeli, Tutum

Abstract

The main aim of this study is to examine the item scores obtained from 'Commitment Scale to the Union' with the Graded Response Model (GRM), which is a model of Item Response Theory. Within the data collection process 'Commitment Scale to the Union' was used, which is developed by Gordon and colleagues (1980) and adapted to Turkish by Bilgin (2003). The research group of the study consist of 353 teachers working in different provinces of Turkey at primary and secondary schools in Kütahya (159), İstanbul (74), Eskişehir (62), Bursa (35), İzmir (12) and Balıkesir (11) in 2009 and they were selected by simple random sampling. There are 15 items in the scale related to commitment of union and these items were requested to answered with responses such as strongly agree (5) and strongly disagree (1), ordinal scores ranging from 1 to 5 as categories. The items of attitude were calibrated with GRM, which is appropriate to Likert type data. Within the analysis

of the research, it is examined primarily whether the data set provides the one dimensional and local independence hypothesis or not. In the second stage of the analysis, by GRM chosen from MULTLOG (Thissen, 1991) program, item calibration was implemented. According to the results, it is seen that data set provides the GRM's one dimensional and local independence hypothesis and model-data consistency was provided. According to the findings, item 4 has the highest slope parameter for teachers and gives the maximum information about the commitment to the union; item 14 has the lowest slope parameter and gives the minimum information about teachers' commitment to the union.

Keywords: *Item Response Theory, Test Development, Graded Response Model, Attitude*

Giriş

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 ölçütle 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 yönetime göre değişirler. Zira 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)'dir (Baykul, 2000).

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 (Kelecioğlu, 2001). 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 XX. Yüzyılın ortalarında adını duyuran, "Örtük Özellikler Kuramı" ya da "Madde Tepki Kuramı" olarak adlandırılan yeni bir kuram geliştirilmiştir. Günümüzde KTK yaygın olarak kullanılmakla birlikte, MTK giderek daha yaygın ve tercih edilir olmaya başlamıştır (Kelecioğlu, 2001; Reise & Ainsworth & Haviland, 2005).

KTK'nın odaklandığı temel durum madde güclüğü ve ayırt edicilik parametresidir (Doğan ve Tezbaşaran, 2003; Uyar & Gübeş & Kelecioğlu, 2013). Örtük özellikleri ölçmek için kullanılan ölçme araçlarında da emelde maddelerin birbiriyle ve ölçmek istediği özellik ile ilişkisi incelenir. MTK, bir dizi matematiksel modellerle KTK'nın açıklayamadığı, bireylerden bağımsız madde ölçeklemesi ve maddelerden bağımsız yetenek ölçüleri elde etme kolaylığını getirmiştir. Başka bir deyişle, yetenek hem seçilen maddelerden hem de testi cevaplayan diğer kişilerin performansından bağımsızdır. MTK'nın KTK'dan en önemli farklı yönü test maddelerinin özelliklerini veren madde parametreleri kestirimlerinin bir cevaplayıcı grubundan diğerine değişmemesi; cevaplayıcıya ait yetenek parametrelerini kestirimlerinin bir grup madde örnekleminde diğerine değişmeden kalmasıdır (Lord & Novick, 1968; Hambleton, 1989; Çıkrıkçı-Demirtaşlı, 1998).

MTK'da ikili ve çoklu puanlanan maddeler için çeşitli modeller geliştirilmiştir. 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 & Swaminathan, 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 (Embretson & 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; Uyar & Gübeş & Kelecioğlu, 2013).

Çalışmanın amacı, öğretmenlerin sendikal bağlılık düzeylerinin saptanmasında ve yine öğretmenlerin sendikal bağlılıkları ile sendikalılığın önemine inanmaları, sendikaya sadakatleri, sendikaya karşı sorumlulukları ve sendika için çaba sarf etme konusunda gönüllü olmaları arasındaki ilişkiyi araştırmada kullanılan Sendikal Bağlılık Ölçeğinden (SBÖ) elde edilen madde cevaplarını, MTK modellerinden çok kategorili sıralı puanlanan maddeler için geliştirilmiş olan Aşamalı Tepki Modeli ile değerlendirmektir.

Yöntem

Araştırmanın bu kısmında araştırma modeli, çalışma grubu, veri toplama araçları, işlem yolu ve verilerin analizi başlığı altında istatistiksel yöntem ve teknikler açıklanmıştır.

Araştırmanın Türü

Bu araştırmada, SBÖ'nden elde edilen madde puanları Aşamalı Tepki Modeli ile incelenmiştir. Araştırma, var olan yöntem ve tekniklerin, gerçek veri üzerinden sınanması yapıldığından betimsel araştırma niteliğindedir.

Çalışma Grubu

Araştırmanın çalışma grubunu ise, Kütahya, İstanbul, Eskişehir, Bursa, İzmir ve Balıkesir illerinde 2009 yılında görev yapan basit tesadüfi örneklem yoluyla seçilen 378 öğretmen oluşturmaktadır. Ancak uygulama sonrası, Yönergeye uygun bir şekilde doldurulmadığı saptanan 16 anket ile sendika üyesi olmayan öğretmenler tarafından doldurulduğu saptanan 9 anket analize dahil edilmeyerek, araştırma 353 öğretmen ile yürütülmüştür.

Veri Toplama Araçları

Bu çalışmada; öğretmenlerin sendikal bağlılıklarını etkileyen faktörlerin belirlenmesi amacıyla Gordon ve arkadaşları (1980) tarafından geliştirilen, Bilgin (2003) tarafından Türkçe uyarlama çalışması yapılan SBÖ kullanılmıştır. SBÖ'de, sendikal bağlılığa yönelik tutuma ilişkin 15 madde bulunmaktadır. Araştırmada, tutum maddelerine verilecek tepkiler için beşli dereceleme tercih edilmiştir. Cevaplayıcılardan ölçekte yer alan her bir tutum ifadesini "*kesinlikle katılmıyorum - katılmıyorum - kararsızım - katılıyorum - kesinlikle katılıyorum*" gibi beş kategoriden biriyle sınıflaması istenmiştir. Her cevaplayıcı 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).

Verilerin Analizi

Araştırma verilerinin analizinde öncelikle veri setinin 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 & Lomax, 2004). AFA sonucu elde edilen özdeğeri 1'den büyük dört faktörden ilkinin tek başına toplam varyansın %61'ini açıkladığı görülmüştür. Yamaç-Eğim grafiği ise baskın bir şekilde tek faktörde düşüş sergilemiştir. Döndürülmüş bileşenler matrisindeki faktör yükleri incelendiğinde (0.49-0.80) tutum ölçeğini oluşturan maddelerin tek faktör altında toplanabileceği gözlenmiştir.

- Doğrulayıcı Faktör Analizi (DFA): Araştırmanın temel amacı doğrulamak ise DFA kullanılmalıdır (Schumacker & 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), Standartize Edilmiş Artık Ortalamaların Karekökü (RMR), İyi Uyum İndeksi (GFI), Düzenlenmiş İyi Uyum İndeksi (AGFI), Normlaştırılmamış 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 & 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 & Lomax, 2004; Hooper & Caughlan & 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=0.60; RMR=0.82; GFI=0.92; AGFI=0.89; NNFI=0.95; CFI=0.97). AFA ve DFA sonucu elde edilen tüm bulgulara dayalı olarak ölçeğin tek boyutlu bir yapıya sahip olduğu söylenebilir.

Yukarıda da değinildiği üzere, MTK'ya göre yerel bağımsızlık varsayımının ihlali aynı zamanda tek boyutluluk varsayımının da ihlali demektir. Bu durumda tek boyutluluk varsayımını sağlayan tutum ölçeğindeki 15 maddenin yerel bağımsızlık varsayımını da sağladığı kabul edilmiştir.

Araştırma verilerinin analizinin ikinci aşamasında ise MULTILOG (Thissen, 1991) programında Aşamalı Tepki Modeli (Graded Response Model) seçilerek madde kalibrasyonu yapılmıştır. Modele ait madde parametreleri ise Marjinal Maximum Likelihood (MML) Tahmin Yöntemi ile kestirilmiştir.

Bulgular

Bu kısımda araştırmanın amacına göre elde edilen bulgulara yer verilmektedir.

SBÖ'den elde edilen verilerin aşamalı tepki modeli ile kalibre edilmesi sonucu negatif log likelihood ($-2*LL$) değeri 9890.6 olarak bulunmuştur. Maksimum like-

Tablo 1. Kategorilere Ait Gözlenen ve Beklenen Frekans Oranları

Maddeler		1	2	3	4	5
1	f	100	100	100	53	0
	g	0.2833	0.2833	0.2833	0.1501	0.0000
	b	0.2837	0.2766	0.2823	0.1537	0.0037
	fark	0.0004	-0.0067	-0.001	0.0036	0.0037
2	f	40	40	40	40	33
	g	0.2073	0.2073	0.2073	0.2073	0.1710
	b	0.2128	0.2020	0.2014	0.2042	0.1796
	fark	0.0055	-0.0053	-0.0059	0.0031	0.0086
3	f	36	36	35	35	35
	g	0.2034	0.2034	0.1977	0.1977	0.1977
	b	0.2141	0.1998	0.1909	0.1928	0.2029
	fark	0.0107	-0.0036	-0.0068	-0.0049	0.0052
4	f	59	134	36	87	37
	g	0.1671	0.3796	0.1020	0.2465	0.1048
	b	0.1952	0.3562	0.0889	0.2326	0.1272
	fark	0.0281	-0.00234	-0.0131	0.0139	0.0224
5	f	33	104	54	98	64
	g	0.0935	0.2946	0.1530	0.2776	0.1813
	b	0.1108	0.3005	0.1400	0.2535	0.1952
	fark	0.0173	0.0059	-0.013	-0.0241	0.0139
6	f	61	152	64	51	25
	g	0.1728	0.4306	0.1813	0.1445	0.0708
	b	0.1947	0.3977	0.1654	0.1513	0.0908
	fark	0.0219	-0.0329	0.0159	0.0068	0.02
7	f	17	63	65	141	66
	g	0.0483	0.1790	0.1847	0.4006	0.1875
	b	0.0671	0.1758	0.1762	0.3873	0.1937
	fark	0.0188	-0.0032	-0.0085	-0.0133	0.0062
8	f	38	124	71	92	28
	g	0.1076	0.3513	0.2011	0.2606	0.0793
	b	0.1282	0.3418	0.1772	0.2539	0.0988
	fark	0.0206	-0.0095	-0.0239	-0.0067	0.0195
9	f	24	99	55	128	42
	g	0.0680	0.2805	0.1558	0.3626	0.1331
	b	0.0843	0.2871	0.1396	0.3383	0.1507
	fark	0.0163	0.0066	-0.0162	-0.0243	0.0176
10	f	13	38	49	152	101
	g	0.0368	0.1076	0.1388	0.4306	0.2861
	b	0.0567	0.1094	0.1351	0.4167	0.2822
	fark	0.0199	0.0018	-0.0037	-0.0139	-0.0039
11	f	32	131	60	102	28
	g	0.0907	0.3711	0.1700	0.2890	0.0793
	b	0.1101	0.3668	0.1489	0.2780	0.0963
	fark	0.0194	-0.0043	-0.0211	-0.011	0.017
12	f	16	76	80	133	48
	g	0.0453	0.2153	0.2266	0.3768	0.1360
	b	0.0590	0.2084	0.2191	0.3713	0.1422
	fark	0.0137	-0.0069	-0.0075	-0.0055	0.0062
13	f	54	169	73	51	6
	g	0.1530	0.4788	0.2068	0.1445	0.0170
	b	0.1741	0.4464	0.1923	0.1637	0.0235
	fark	0.0211	-0.0324	-0.0145	0.0192	0.0065
14	f	23	86	89	121	34
	g	0.0652	0.2436	0.2521	0.3428	0.0963
	b	0.0750	0.2354	0.2459	0.3401	0.1036
	fark	0.0098	-0.0082	-0.0062	-0.0027	0.0073
15	f	15	33	76	158	71
	g	0.0425	0.0935	0.2153	0.4476	0.2021
	b	0.0638	0.0932	0.2060	0.4325	0.2046
	fark	0.0213	-0.0003	-0.0093	-0.0151	0.0025

f=frekans, g=gözlenen, b=beklenen

likelihood kestiriminde negatif loglikelihood değeri, verinin modelden uzaklaşma derecesini göstermektedir (Embretson ve Reise, 2000). Ölçeğe ait marjinal güvenilirlik katsayısı 0.85 olarak elde edilmiştir. Marjinal güvenilirlik, tüm bağlılık düzeyindeki öğretmenlerden tahmin edilen koşullu standart hataların ortalaması olarak ele alınan toplam güvenilirliği temsil etmektedir.

Maddelerin veriye uyum düzeyleri, gözlenen ve beklenen oranlar arasındaki fark aracılığı ile incelenebilir. Gözlenen ve beklenen oranlar arasındaki farklar “artık” olarak da isimlendirilir (Embretson & Reise, 2000). Artıkların sifıra yaklaştıkça, modelin veri ile uyum gösterdiği kabul edilmektedir.

Tablo 1’de kategorilere yanıt veren öğretmen sayıları ile gözlenen-beklenen frekans oranları ve frekans oranları arasındaki fark ya da artık değerleri verilmiştir.

Tablo 1’de yer alan 15 maddenin her bir kategorisine ait beklenen ve gözlenen oranlar arasındaki farklar incelendiğinde, bütün artıkların .01 değerinden daha küçük olduğu görülmektedir. Bu bulguya dayalı olarak ATM’nin veriye uyum gösterdiği söylenebilir.

Diğer bir araştırma bulgusu ise Tablo 2’de de görülmekte olan ATM’ye dayalı olarak kestirilen madde parametreleridir.

Tablo 2. Aşamalı Modele Göre Kestirilen Parametre Değerleri

Madde	α_i (SE)	β_1 (SE)	β_2 (SE)	β_3 (SE)	β_4 (SE)
1	0.31 (0.15)	-3.03 (1.30)	0.79 (1.39)	5.48 (2.53)	18.12 (****)
2	0.28 (0.24)	-4.78 (4.51)	-1.26 (1.26)	1.73 (2.11)	5.55 (5.05)
3	0.35 (0.20)	-3.84 (2.36)	-1.03 (0.72)	1.26 (1.16)	4.04 (2.30)
4	1.83 (0.20)	-1.18 (0.17)	0.18 (0.09)	0.49 (0.11)	1.58 (0.20)
5	1.44 (0.18)	-1.92 (0.26)	-0.35 (0.12)	0.20 (0.12)	1.33 (0.20)
6	1.74 (0.19)	-1.21 (0.16)	0.33 (0.11)	0.98 (0.13)	1.90 (0.24)
7	0.22 (0.14)	-11.95 (7.75)	-5.17 (3.72)	-1.49 (1.79)	6.49 (4.47)
8	1.74 (0.20)	-1.61 (0.20)	-0.10 (0.09)	0.53 (0.11)	1.83 (0.23)
9	1.48 (0.18)	-2.14 (0.29)	-0.50 (0.12)	0.04 (0.29)	1.59 (0.22)
10	0.25 (0.12)	-11.41 (6.26)	-6.57 (3.55)	-3.43 (1.91)	3.81 (2.16)
11	1.64 (0.20)	-1.79 (0.23)	-0.08 (0.09)	0.46 (0.11)	1.91 (0.24)
12	0.23 (0.14)	-12.32 (7.50)	-4.50 (2.76)	-0.24 (5.08)	8.01 (4.90)
13	1.76 (0.20)	-1.32 (0.19)	0.43 (0.10)	1.25 (0.16)	2.88 (0.44)
14	0.20 (0.22)	-12.55 (****)	-4.00 (4.73)	1.13 (2.11)	10.78 (****)
15	0.22 (0.16)	-12.40 (9.78)	-7.78 (6.35)	-2.61 (2.47)	6.28 (5.36)

MTK altında madde ayırıcılık parametresi maddenin örtük özellik sürekliliğinde cevaplayıcıları yeteneklerine göre ayırabilme gücü olarak ifade edilir. Dolayısıyla madde ayırıcılık parametresi, madde ile ölçülmek istenen yetenek arasındaki güçlü ilişkiyi

yansıtır. Bu bağlamda pozitif ve yüksek eğim değerleri tercih 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 ayrırcılığa sahip maddeler test bilgi fonksiyonuna daha fazla katkı yapan maddelerdir (Hambleton & Swaminathan & Rogers, 1991). Ayırt edicilik parametre değerleri 0.01-0.34 arasında ise madde çok düşük ayrırcılığa; 0.35-0.64 arasında ise düşük ayrırcılığa; 0.65-1.34 arasında ise orta ayrırcılığa; 1.35-1.69 arasında ise yüksek ayrırcılığa ve 1.70 ve daha yukarı bir değer almışsa çok yüksek ayrırcılığa sahiptir (Baker, 2001).

Tablo 2'ye göre eğitim parametrelerinin 0.20-1.83 arasında değerler aldığı; 1., 2., 3., 7., 10., 12., 14. ve 15. maddeler düşük ayrırcılığa, 5., 9. ve 11. maddeler yüksek ayrırcılığa ve 4., 6., 8. ve 13. maddeler çok yüksek ayrırcılığa sahiptir. Tüm maddeler içerisinde ise 4. maddenin en yüksek eğitim parametresine ($a_i = 1.83$) ve 14. maddenin ise en düşük eğitim parametresine ($a_i = 0.20$) sahip olduğu görülmektedir.

Maddeyi tanımlayan bir diğer parametre madde güçlük ya da lokasyon parametresidir. Madde güçlük parametresi, (b_i), madde güçlüğüünün ölçüsü, davranışın veya tutumun frekansı olarak bilgi veren parametredir (Rubio ve diğerleri, 2007). ATM'de ölçekteki her bir madde (i), bir tane madde ayrırcılık parametresi, a_i ve $j=1, \dots, m$, tane kategori eşik parametresi (β_{ij}) ile tanımlanır. ATM cevaplayıcının belirli bir kategori ve daha üzerinde tepki verme olasılığı için $m-1$ tane eşik tepki fonksiyonunu tahmin eder. 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 & 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 (Madera, 2003). Tablo 2'de madde ayrırcılık parametresine ilaveten kategoriler arası madde eşik parametresine de (β_i) yer verilmiştir. Buna göre her madde için (β_1) en düşük eşik parametresine, (β_4) ise en yüksek eşik parametresine sahiptir. Örneğin; 1. maddenin $\beta_1 = -3.03$ ve $\beta_4 = 18.12$ olarak hesaplanmıştır. Bu maddenin 1. kategorisinde %50 olasılıkla tepki vermek için gerekli yetenek düzeyi $\theta = -3.03$ olarak, 5. Kategoride ise %50

olasılıkla tepki vermek için gerekli yetenek düzeyi ise $\theta = 18.12$ hesaplanmıştır. Ayrıca eşik parametre değerlerinin çoğunun negatif değerler almasına dayalı olarak kategorilerdeki cevapların daha çok algının düşük düzeyi ($\theta < 0$) tarafından desteklendiği söylenebilir. Kategori arası eşik parametreleri, -12.55 ile 18.12 değerleri arasında değişmektedir. Eşik parametre değerlerinin çoğunun negatif değerler almasına dayalı olarak kategorilerdeki cevapların daha çok tutumun düşük düzeyi ($\theta < 0$) tarafından desteklendiği söylenebilir.

Tablo 2'ye göre 14. maddede birinci kategoriden ikinci kategoriye geçiş için eşik parametresi, -12 yetenek düzeyinin altında (-12.55) kestirilmiştir. Bu durum, ilk kategoriye az kişinin yanıtlamasından ve maddenin ölçmek istediği özellikle ilişkili olmamasından kaynaklanmış olabilir (Embretson & Reise, 2000). Nitekim Tablo 1 incelendiğinde, 14. maddenin birinci kategorisine cevap veren öğretmen sayısının (23) diğer kategorilere cevap veren öğretmen sayısından oldukça az olduğu görülmektedir. Öğretmenlerin sadece % 6.5'i ilk kategoriye seçmiştir. Tablo 2'ye göre 14. maddeye ait ayırt edicilik parametresinin düşük olması (0.20) da bu durumun oluşmasının bir diğer sebebi olarak gösterilebilir.

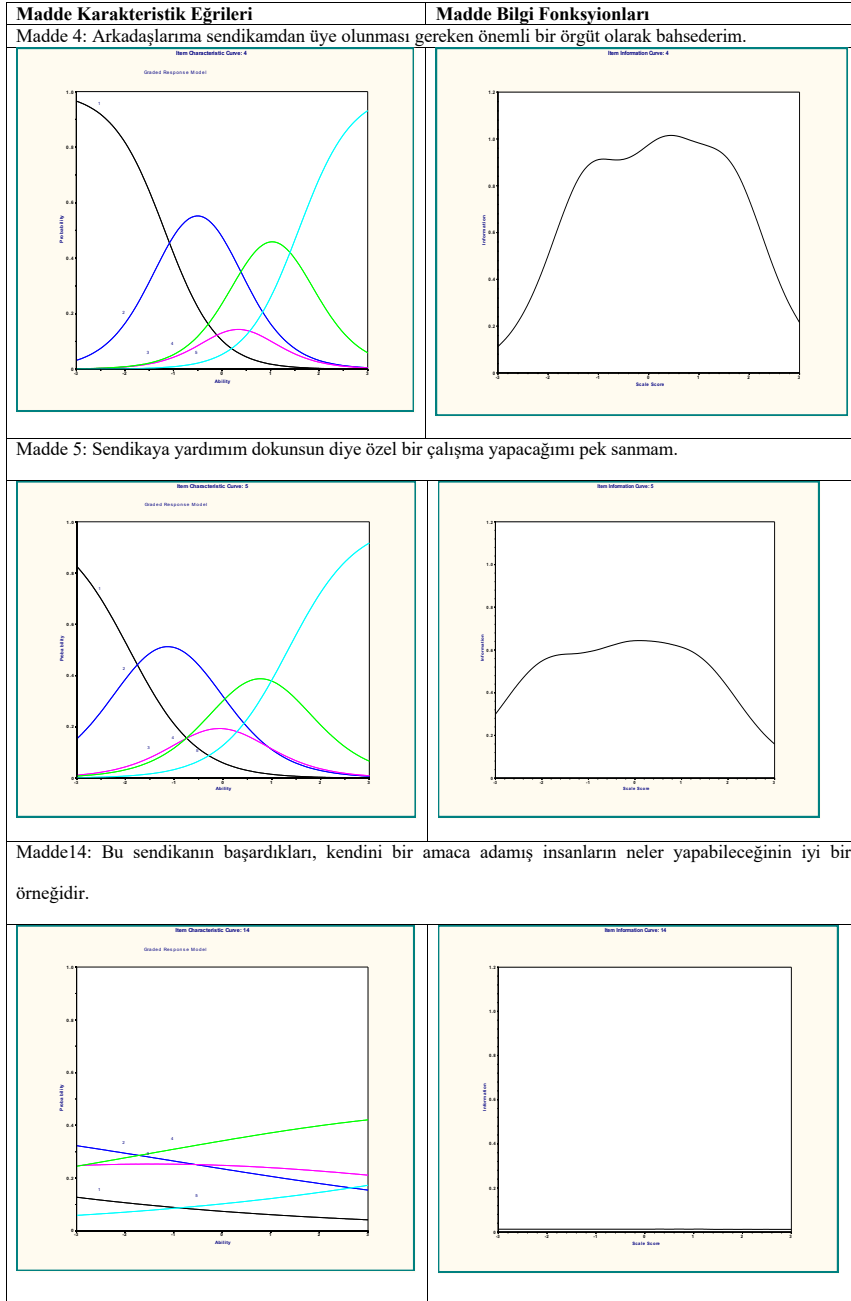
MTK güçlü bir test kuramı olmasına rağmen model veri uyumunun sağlanmasını gerektirir. Ancak bu konuda genel kabul görmüş bir model veri uyumu testi bulunmamaktadır (Gray-Little & Williams & Hancock, 1997). Bu çalışmada model veri uyumu açısından madde parametrelerinin değişmezlik özelliğinin test edilmesi yöntemine başvurulmuştur. Bunun için örneklem tesadüfi olarak ikiye ayrılmış ve madde parametreleri bu gruplar üzerinden kestirilmiştir. Madde ayrırcılık parametreleri arasındaki korelasyon $r=0,90$ ($p<0,01$) olarak hesaplanmıştır. Madde eşik parametreleri arasındaki korelasyon ise

β_1 için $r=0,96$, β_2 için $r=1,00$, β_3 için $r=0,92$ ve β_4 için $r=0,94$ ($p<0,01$)'dir. Bu bulgular ölçekte bulunan maddelerin değişmezlik özelliğinin bulunduğunu ve ATM'nin veri grubuna uygunluğunu desteklemektedir (Köse, 2015).

Maddelerin istatistiksel özelliklerine ilişkin bilgi edinmenin bir diğer yolu madde karakteristik eğrilerini ve bilgi fonksiyonu grafiklerini incelemektir. Bu

eğriler, her bir yanıtama düzeyinin tek başına yeteneğin bir fonksiyonu olduğunu göstermektedir. Aynı zamanda madde bilgi fonksiyonları, madde karakteristik eğrilerindeki farklılıklar sonucu o maddenin ölçmenin kesinliğine olan etkisini de göstermektedir (Uyar & Gübeş & Kelecioğlu, 2013, s. 132).

Şekil 1'de ölçekteki üç maddeye ait kategori karakteristik eğrileri verilmiştir. Bu maddelerden, 4. madde en yüksek eğime ve ayırıcılığa ($\alpha_i = 1.83$); 5. madde yüksek düzeyde bir eğime ($\alpha_i = 1.44$) ve 14. madde en düşük eğime sahip olması ($\alpha_i = 0,20$) nedeni ile seçilmiş ve örnek oluşturması açısından grafikleri verilmiştir.

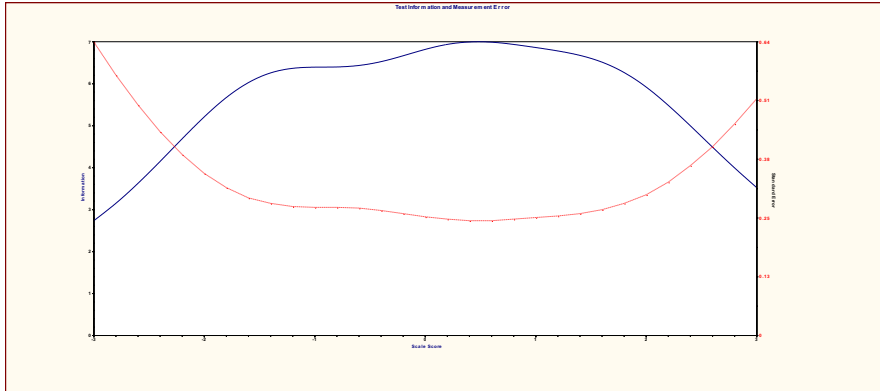


Şekil 1. SBÖ'de Yer Alan Maddelerin Farklı Eğim ve Kategori Eşik Parametrelerine Göre Karakteristik Eğrileri

Ölçekteki maddelere verilecek tepkiler için 5'li derecelendirme tercih edilmiştir. Cevaplayıcılardan ölçekte yer alan her bir maddeyi ve dolayısıyla Şekil 1'deki maddeleri "kesinlikle katılmıyorum - katılmıyorum - kararsızım - katılıyorum - kesinlikle katılıyorum" gibi beş kategoriden biriyle sınıflaması istenmiştir. Her cevaplayıcı 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). Bu bağlamda 4. maddeye ait madde karakteristik eğrisinde 5. kategori monoton artan bir eğim sergilemekte iken, 1. kategori monoton azalan bir eğim sergilemektedir. Bu sonuç, yetenek düzeyi arttıkça o kategorinin seçilme olasılığının düşük olduğunu göstermektedir. Geriye kalan 2., 3. ve 4. kategoriler ise düşük ve orta düzeydeki yeteneklerde artan bir eğim, yüksek yetenek düzeylerinde ise azalan bir eğim sergilemektedir. En yüksek eğitim parametre değerine sahip olan 4. madde, aynı zamanda en çok bilgiyi veren maddedir. Benzer şekil-

de yüksek ayırıcılığa sahip 5. maddeye ait madde karakteristik eğrisinde de 5. kategori monoton artan bir eğim, 1. kategori ise monoton azalan bir eğim sergilemektedir. Yine 5. maddede 5. ve 1. kategori dışındaki kategorilerde düşük ve orta düzeydeki yeteneklerde artan bir eğim, yüksek yetenek düzeylerinde ise azalan bir eğim sergilendiği görülmektedir. Düşük eğitim parametresine sahip 14. maddeye ait madde karakteristik eğrisinde ise, kategorilerin çok geniş bir yetenek alanında cevaplanma olasılığına sahip olduğu görülmektedir. En yüksek eğitim parametre değerine sahip olan madde en çok bilgiyi verirken, en düşük eğitim parametre değerine sahip olan madde ise en çok bilgiyi düşük tutum düzeyinde vermektedir (Uyar & Gübeş & Keleşoğlu, 2013). Bu bağlamda 14. madde, en çok bilgiyi düşük tutum düzeyinde vermektedir.

Madde bilgi fonksiyonlarının yanı sıra Şekil 2'de görülen tüm maddelerden elde edilen toplam bilginin yer aldığı test bilgi fonksiyonu incelenmiştir.



Şekil 2. Aşamalı Tepki Modeline Göre Elde Edilen Test Bilgi Fonksiyonu

Şekil 2 ölçekteki bilginin en çok 0.2 ile 1.6 yetenek ya da tutum düzeyleri arasında elde edildiğini göstermektedir. Madde kalibrasyon sonuçları incelendiğinde, testin fazla bilgiyi 1.6 yetenek düzeyinde (bilgi değeri=6.993) verdiği görülmüştür. Testin verdiği bilgi özellikle 1.8 yetenek düzeyinden sonra azalmıştır. Bu durum standart hatanın da artmasına neden olmuştur.

Sonuç ve Öneriler

Bireylerin çeşitli objelere, fikirlere ve olaylara ne şekilde tepkide bulunacağını büyük ölçüde tutumlar tarafından yapılandırıldığı bilinmektedir. Bir bireyin tutumunun bilinmesi, gelecekte nasıl davranacağı ile ilgili bir kestirmeye olanak tanımaktadır. Bu nedenle ölçülmesi büyük önem taşımaktadır. Bu bağlam-

da öğretmenlerin sendikalara yönelik tutumlarının belirlenmesi de son derece önemlidir. Hiç kuşkusuz tutumların belirlenmesi önemli olduğu kadar, bu tutumların belirlenmesine dönük ölçme araçlarının da nitelikli olması önemlidir. Bu bağlamda öğretmenlerin sendikal bağlılıklarını belirlemeye yönelik ölçme araçlarının niteliğini artırmaya dönük bu çalışmada, SBÖ'nin maddeleri 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.

Araştırma bulguları, gözlenen ve beklenen frekans oranları arasındaki farkın küçük olduğunu ortaya koymuştur. Bu sonuca göre Aşamalı Tepki Modeli veriye ve SBÖ maddelerinin seçiminde ve değerlendirilmesi için uygun bir modeldir.

Araştırma bulguları, SBÖ maddelerinin ayırt ediciliğe sahip olduğunu ve önemli bir çoğunluğunun ise yüksek ve çok yüksek ayırt ediciliğe sahip olduğunu ortaya koymuştur. Bu bağlamda öğretmenlerin sendikal bağlılıklarını ve sendikal bağlılıkları ile sendikalılığın önemine inanmaları, sendikaya sadakatleri, sendikaya karşı sorumlulukları ve sendika için çaba sarf etme konusunda gönüllü olmaları arasındaki ilişkiyi araştırmaya dönük SBÖ 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.
- Baykul, Y. (1979). Örtük Özellikler ve Klasik Test Kuramları Üzerine Bir Karşılaştırma. Yayınlanmamış Doktora Tezi, H.Ü. Sosyal Bilimler Enstitüsü, Ankara.
- Baykul, Y. (2000). *Eğitimde ve Psikolojide Ölçme: Klasik Test Teorisine ve Uygulaması*. Ankara: ÖSYM Yayınları.

Bilgin, L. (2003). Sendikaya Bağlılık Ölçeğinin Türkçe Uyarlaması. *D.E.Ü. Sosyal Bilimler Enstitüsü Dergisi*, 5 (4), 12-31.

Çıkrıkçı-Demirtaşlı N. (1998). Test geliştirmede yeni yaklaşımlar: örtük özellikler kuramı -temel özellikleri, varsayımları, model ve sınırlılıkları. *Ankara Üniversitesi Eğitim Fakültesi Dergisi*, 2 (28), 161-173.

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

Demirel, Ö. (2005). *Öğretme Sanatı*. 8. Baskı, Ankara: Pegem A Yayıncılık.

Doğan, N. ve Tezbaşaran, A. A. (2003). Klasik test kuramı ve örtük özellikler kuramının örneklem bağlamında karşılaştırılması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 25, 58-67.

Embretson, S. E. & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Mahwah, NJ: Erlbaum.

Gelbal, S. (1994). *p Madde Güçlük İndeksi ile Rasch Modelinin b Parametresi ve Bunlara Dayalı Yetenek Ölçüleri Üzerine Bir Karşılaştırma*. Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Yayınlanmamış Yüksek Lisans Tezi.

Gelbal, S. (2004). Ölçme ve değerlendirme. Öğretmen Adayları İçin Konu Anlatımlı KPSS (Editör Demirel, Ö.). Geliştirilmiş 20. Baskı, Ankara: Pegem A Yayıncılık.

Gordon, M. E. & Philpot, J. W. & Burt, R. E. & Thompson, C. A. & Spiller, W. E. (1980). Commitment to the union: development of a measure and an examination of its correlates. *Journal of Applied Psychology*, 65, 479-499.

Gray-Little, B & Williams, V. S. L. & Hancock, T. D. (1997). An item response theory analysis of the Rosenberg Self Esteem Scale. *Personality and Social Psychology Bulletin*, 23, 443-451.

- Hambleton, R. K. & Swaminathan, H. (1985). *Item Response Theory*. Kluwer-Nijhoff Publishing, Boston, USA.
- Hambleton, R. K. & Swaminathan, H. (1989). *Item Response Theory. Principles and Application*. Kluwer-Nijhoff Publishing, Boston, USA.
- Hambleton, R. K. & Swaminathan, H. & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. Sage Publications, London.
- Hooper, D. & Caughlan, J. & Mullen, M. R. (2008). Structural equation modeling: guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6 (1), 53-60. Retrieved Kasım 12, 2010, Web: <http://arrow.dit.ie/cgi/viewcontent.cgi?article=1001&context=buschmanart>.
- Kelecioğlu, H. (2001). Örtük özellikler teorisindeki b ve a parametreleri ile klasik test teorisindeki p ve r istatistikleri arasındaki ilişki. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 20: 104-110.
- Koch, W. R. (1983). Likert Scaling using the graded response latent trait model. *Applied Psychological Measurement*, 7 (1), 15-32.
- Köse, İ. A. (2015). Aşamalı tepki modeli ve klasik test kuramı altında elde edilen test ve madde parametrelerinin karşılaştırılması. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 15 (2), 184-197.
- Lord, F. M. & Novick, M. R. (1968). *Statistical Theories of Mental Test Scores*, New York: Addison Wesley.
- Matteucci, M. & Stracqualursi, L. (2006). Student assessment via graded response model. *Statistica*, 4, 435-447.
- Reise, S. P. & Ainsworth, A. T. & Haviland, M. G. (2005). Item response theory. Fundamentals, applications, and promise in psychological research. *Current Directions in Psychological Science*, 14, 2, 95-101.
- Rubio, V.J.& Aguado, D. & Hontangas, P.M. & Hernandez, J.M. (2007). Psychometric properties of an emotional adjustment measure. *European Journal of Psychological Assessment*, 23(1), 39-46.
- Smejima, F. (1969). Estimation of latent ability using a response pattern of graded scores (Psychometric monograph No.17). Richmond, VA: Psychometric society. Retrieved from <http://www.psychometrika.org/journal/online/MN17.pdf>.
- Schumacker R. E. & Lomax R. G. (2004). *A beginner's Guide to Structural Equation Modeling* (second edition), Londra: Lawrence Erlbaum Associates Publishers, 79-122.
- Sümer, N. (2000). Yapısal eşitlik modelleri: temel kavramlar ve örnek uygulama, *Türk Psikoloji Yazıları*, 3 (6), 49-73.
- Tekin, H. (1977). *Eğitimde Ölçme ve Değerlendirme*. Ankara: Mars Matbaası.
- Thissen, D. (1991). *MULTILOG User's Guide (Version 6.0)*. Mooresville, IN: Scientific Software.
- Turgut, M. F. (1977). *Tutumların Ölçülmesi*, Eğitimde Ölçme Teknikleri Ders Notu. No:7.
- Uyar, Ş. & Gübeş, N. Ö. & Kelecioğlu, H. (2013). PISA 2009 tutum anketi madde puanlarının aşamalı tepki modeli ile incelenmesi, *Eğitim ve Öğretim Araştırmaları Dergisi*, 2 (4): 125-134.