

Çok Boyutlu Madde Tepki Kuramı¹

İbrahim Alper KÖSE *

Abant İzzet Baysal Üniversitesi

Özet

Eğitim ve psikolojide test alanların verdiği tepkilerden en doğru ve kullanışlı sonuçlar elde etmeyi amaçlamak bu alanlarda çalışanların temel noktası olmuştur. Bu amaca ulaşmak için birçok kuram geliştirilmiştir. Klasik test kuramı, madde tepki kuramı...vb örnek olarak verilebilir. Madde tepki kuramı bireyin performansının altında yatan örtük özelliği açıklamak için kullanılan güçlü bir model olmasına rağmen, tek boyutluluk varsayımının karşılanmadığı durumlarda tek boyutlu modeli, çok boyutlu test verisine uygulamak yetenek ve madde parametreleri kestirimlerinde geçerlik sorunlarını beraberinde getirecek ve model-veri uyumunda önemli problemler ortaya çıkaracaktır. Bunun yanında ölçme araçlarına dayalı olarak verilen çeşitli kararların (seçme, yerleştirme, akademik başarının değerlendirilmesi vb.) isabetliliği açısından doğru modelleme son derece önemlidir. Bu nedenlerle bu çalışmanın amacı, çok boyutlu madde tepki kuramının ayrıntılı olarak tanıtılması, tek boyutlu modeller karşısında alana getirdiği yenilikler ve avantajların tartışılmasıdır.

Anahtar Sözcükler: Ölçme Kuramı, Madde Tepki Kuramı, Çok Boyutlu Madde Tepki Kuramı

Abstract

Providing the most useful and valid inferences from examinee responses is the basic point for researchers, who study at the education and psychology area. For this aim, many theories have been developed such as classic test theory, item response theory...ect. Item response theory is the robust theory to explain latent ability on examinee performance but in the case of circumstances in which unidimensionality assumption cannot be met, application of unidimensional models to multidimensional data matrix, brings validity problems on estimation of ability and item parameters and also cause model-data fit problems. Furthermore, correct modeling as a basis for measurement tools is very fundamental for various decisions (election, placement, assessment of achievement). For these reasons, the purpose of this study is to introduce multidimensional item response theory (MIRT) and to discuss advantages and innovations of MIRT across unidimensional IRT models in the frame of literature.

Keywords: Measurement Theory, Item Response Theory, Multidimensional Item Response Theory

Eğitim ve psikolojide çeşitli amaçlarla bireyler hakkında kararlar verilmektedir. Bu kararların doğru ve güvenilir olmaları amacıyla eğitim ve psikoloji alanındaki araştırmacılar, davranışların daha geçerli ve güvenilir ölçümlerini sağlayacak yeni yaklaşımlar geliştirmektedirler. Bu yaklaşımlar geliştirilirken ölçülecek özelliği nicelleştirecek ve ortaya konan niceliğin anlamını yorumlayabilecek güçlü ölçme kuramlarına ihtiyaç duyulmuştur. Bu amaçla, araştırmacılar, geçerli ve güvenilir sonuçlar veren güçlü ölçme kuramlarını geliştirmektedirler.

Tek boyutlu Madde Tepki Kuramı (MTK) pratik test durumlarına Klasik Test Kuramı (KTK)'ndan daha iyi bir çerçeve sunmasına rağmen, kuramsal ve görgül kanıtlar tepki verilerinin çoğunun tek boyutluluk varsayımına tamamen uymadığını göstermektedir. Bunun yanında tek boyutlu MTK'da tek bir maddenin sadece bir örtük özelliği ölçtüğünü varsaymak veri grubu ile tek boyutlu modellerin uyum sorunlarını da beraberinde getirmektedir. Buna ek olarak çok boyutlu veri

¹ Bu çalışma "Madde Tepki Kuramına Dayalı Tek Boyutlu ve Çok Boyutlu Modellerin Test Uzunluğu ve Örnekleme Büyüklüğü Açısından Karşılaştırılması" isimli tez çalışmasından özetlenerek hazırlanmıştır.

* Yrd. Doç. Dr., Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Eğitim Bilimleri Bölümü, i.alper.kose@gmail.com

grubu için tek bir özellik boyutunun olduğunun varsaymak, MTK'nın değişmezlik özelliğini de tehdit etmektedir (Ackerman, 1994).

Tek bir örtük özelliği ölçmek için çok boyutlu maddeler kullanmak testin yapı geçerliliğini düşürmektedir. Çünkü testin yapı geçerliliği kuramsal yapıya ve teste özgü ölçme aracına veya işlemine bağlıdır. Tek boyutlu modellerin kullanılması ile tek boyutluluk varsayımının ihlali sonucu çok boyutlu örtük uzay ile hedeflenen tek boyutlu örtük uzay arasında birebir eşleme yapılamayacaktır. Testin yapı geçerliğinde bilişsel süreçlerde tek boyutlu modellemeden çok boyutlu modellemeye geçiş ile birlikte bu sorunlar ortadan kalkacaktır. Eğitim ve psikolojide test alanların verdiği tepkilerden en doğru ve kullanışlı sonuçlar elde etmeyi amaçlamak bu alanlarda çalışanların temel odak noktası olmuştur. Bunun yanında ölçme araçlarına dayalı olarak verilen çeşitli kararların (seçme, yerleştirme, akademik başarının değerlendirilmesi vb.) isabetliliği açısından doğru modelleme son derece önemlidir. Bu noktadan hareketle bu çalışmanın temel amacı çok boyutlu MTK'nın ayrıntılı olarak tanıtılması, tek boyutlu modeller karşısında eğitim ve psikoloji alanına getirdiği yenilikleri ve avantajları ilgili alanyazın çerçevesinde tartışmaktır.

Ölçme tarihinde, bir testi yanıtlayan bireyin performansının altında yatan örtük özelliği açıklamak için geliştirilen başlıca kuram KTK'dır. KTK, testten alınan puanı, gerçek puan ve ölçme hatasının toplamı ile açıklayan basit bir modeldir. Buna karşın kuram, pratik test durumlarında istenmeyen özellikleri de taşımaktadır. Test ve madde özelliklerinin gruba bağlı olması, bireyin performansı hakkında bilginin sadece testin tümü üzerinden verilmesi ve madde bazında cevaplayıcı performansı hakkındaki bilginin ise kaybolması kuramın önemli sınırlılıklarındandır (Hambleton, Swaminathan ve Rogers, 1991).

Psikolojik ölçme alanında son yüzyılın en önemli gelişmelerinden birisi MTK'dır. MTK, kişinin ölçülen özellikteki yerini, test maddelerine verdiği yanıtları kullanarak KTK'nın zayıf varsayımlarına karşı geliştirdiği güçlü varsayımlarla matematiksel bir model ile açıklayan modern test kuramıdır (Bobcock, 2009). Kuramın önceki ölçme kuramlarına göre en önemli avantajları; madde güçlük değerlerinin ve yetenek düzeylerinin aynı ölçek düzeyinde yer alması (Spencer, 2004), cevaplayıcının yeteneğini daha keskin (precision of measurement) ve daha küçük ölçme hataları ile ortaya koyması, elde edilen madde parametrelerinin aynı evrenden gelmiş farklı örneklemelerde değişmezlik özelliğini koruması ve tahmin edilen yetenek düzeylerinin farklı alt test maddelerinde de değişmez olarak kalmasıdır. Ayrıca MTK'da cevaplayıcıların yetenek düzeylerinin karşılaştırılması daha keskin ve daha kullanışlıdır. Bunun nedeni KTK'dan farklı olarak, testin tümünden elde edilen ölçme hatasının bütün cevaplayıcılar için aynı olmadığını varsaymasıdır (Lee, 2007).

MTK, geliştirilmesinden günümüze kadar eğitim ve psikoloji alanındaki ölçmelerde sıklıkla kullanılmıştır. Bunun yanında MTK, test geliştirme, test puanlarını eşitleme, madde yanlılıklarını belirleme, bilgisayar ortamında bireyselleştirilmiş testlerde cevaplayıcının yeteneğinin kestirilmesi gibi birçok ölçme alanına yenilikler ve uygulanabilir çözümler getirmiştir. Buna karşın MTK, güçlü matematiksel ve istatistiksel varsayımlara dayanması ve bu varsayımların karşılanması ve model veri uyumunun sağlanması durumunda işlerlik kazanabilmektedir (Zhao, 2008). Kısa zaman öncesine kadar en çok kullanılan MTK modeli olan tek boyutlu MTK, cevaplayıcının bir grup test maddesi karşındaki performansını, tek bir yetenek veya özellik ile açıklamaktaydı. Farklı türdeki testler ile yapılan araştırmalar kuramın bu varsayımının tamamen karşılanmasının oldukça güç olduğunu ortaya koymuştur. Özellikle başarı testlerinin birden fazla boyutu ölçtüğü artık bilinen bir gerçektir. Böyle çok boyutlu testlerin olması, kuramın tek boyutluluk varsayımının ihlallerini ortaya çıkarmaktadır (Pomplun, 1988).

Tek boyutluluk varsayımının ihlali ile yapılan kestirimlerde sadece başat boyut kestirilmekte, diğer boyut hakkındaki bilgi ise kaybolmaktadır (Kreiter, 1993). Bunun yanında Ackerman (1992, 1994a, 1994b), tek boyutluluk varsayımı ihlal edildiğinde test alanların tek bir özellik üzerinde sıralanamayacağını belirtmiştir. Bunun nedeni olarak da çok boyutlu örtük uzay ile cevaplayıcının tek bir yeteneği ile bire bir eşleme yapılamayacağını göstermiştir. Tek bir örtük özelliği ölçen bir testi geliştirmenin neredeyse imkânsız oluşu (Lee, 2007), bilgi kaybının en aza indirilmesi ve bunların sonucu olarak daha geçerli ve güvenilir ölçmeler yapabilmek amacıyla iki veya daha fazla yetenek gerektiren testler için çok boyutlu MTK modelleri geliştirilmiştir.

Çok Boyutlu Madde Tepki Kuramı

Basit istatistiksel modeller, karmaşık modellerin temelini oluşturmaktadır. Bu durum MTK modelleri için de geçerlidir. Tek boyutlu MTK modelleri ile çok boyutlu MTK modelleri arasındaki ilişki tek boyutluluk varsayımının ihlaline dayanır. Örnek olarak bir matematik sorusunu alalım.

“Ayşe ile ablasının cevizlerinin toplamı 86’dır. Ayşe’nin cevizlerinin 2 katının 13 fazlası, ablasının cevizlerinin 3 katına eşittir. Her birinin kaç ceviz vardır?”

Bu tür sorularda cevaplayıcıdan önce matematik problemini okuyup anlaması, daha sonra da matematiksel işlemlerle bu problemi çözmesi beklenir. Bu tür problemler, hem okuduğunu anlama hem de işlem yeteneklerini beraber gerektirmektedir. Bu gibi durumlarda tek boyutluluk varsayımının karşılanması oldukça zordur. Bunun sonucu olarak test verisini doğru bir şekilde temsil edecek farklı modellere ihtiyaç duyulmaktadır. Tek boyutlu MTK’nın bu sınırlılığına karşı geliştirilmiş modeller genel olarak çok boyutlu MTK modelleri olarak adlandırılır (Smith, 2009).

Çok boyutlu MTK, yapılar veya boyutlar olarak kavramsallaştırılan iki veya daha fazla örtük değişken ile test alanın belirli bir test maddesini doğru olarak cevaplama olasılığı arasındaki ilişkiyi matematiksel bir model ile açıklayan modellerdir. Bir başka ifade ile tek boyutlu MTK modellerinin çok boyutluluğa uyarlanmasıdır (Ackerman, Gierl ve Walker, 2003). Çok boyutlu MTK, faktör analizinin veya yapısal eşitlik modellemesinin özel bir durumu veya tek boyutlu madde tepki kuramının bir uzantısı olarak da düşünülebilir. Faktör analizinin bir uzantısı olarak Christofferson (1975, Aktaran; Ackerman, Gierl ve Walker, 2003; McDonald, 2000) bir grup örtük değişkeni, v ile tanımlanmış ve bu değişkeni;

$$v_i = \lambda' f + \delta_i \quad \text{şeklinde ifade etmiştir. Eşitlikte;}$$

$\lambda' = [\lambda_1, \lambda_2, \dots, \lambda_n]$ ortak faktör yükleri matrisini (common factor loadings Matrices),

f.....ortak faktör vektörünü (common factor vector),

δ_i i'nci tekil faktörü (unique factor), ifade etmektedir.

Bu model her i maddesi için, maddenin doğru cevaplanmasından sorumlu bir örtük değişkenin varlığını varsaymaktadır. Bu örtük değişken sürekli ve normal dağılan bir değişkendir. Cevaplayıcının yeterliği belli bir eşik değerinin, t_i , üzerinde veya eşitse, madde doğru cevaplanır, aksi durumlarda cevaplayıcı maddeyi yanlış yanıtlayacaktır. 1-0 şeklinde puanlanan her i maddesi için cevaplayıcının tepkisi,

$$U_i = 1, \text{ eğer } v_i \geq t_i \text{ ise veya}$$

$$U_i = 0, \text{ eğer } v_i < t_i \text{ şeklinde gösterilir.}$$

i maddesini doğru yanıtlayan cevaplayıcıların oranı (p değeri veya güçlük düzeyi) normal dağılım eğrisinde eşik değerinin, t_i , üzerinde kalan alanın oranı olarak,

$$p_i = N(t_i)$$

ifade edilebilir. Eşitlikte N normal ogive fonksiyonunu temsil etmektedir. Bu sonuç k-boyutlu normal ogive fonksiyonunun gelişmesine neden olmuştur. Fonksiyon;

$$P = \{U_i = 1 | \theta_1 \dots \theta_k\} = N\{\beta_{i0} + \beta' \theta_i\} = N\{\beta_{i0} + \beta_{i1} \theta_1 + \beta_{i2} \theta_2 + \dots \beta_{ik} \theta_k\}$$

şeklindedir. Eşitlik 6’da;

$$\beta_{i0} = \frac{t_i}{\sqrt{\psi_i}} \text{ madde güçlük parametresini,}$$

ve k’ncı boyut için;

$$\beta_i = \frac{\lambda_i}{\sqrt{\psi_i}} \text{ madde ayırıcılık parametresini göstermektedir.}$$

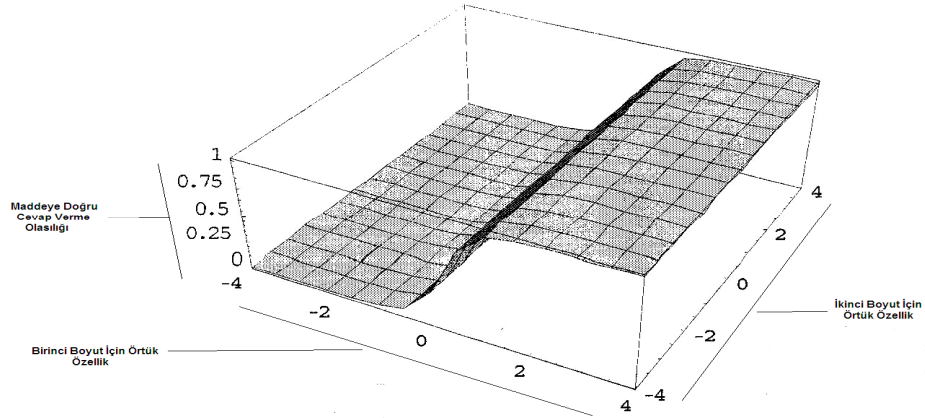
Burada ψ_i açıklanan madde varyansını veya ortak faktörün (communality) 1'den çıkarılmasını ifade eder. Eşitlik 6 incelendiğinde Lord (1980; Aktaran; Ackerman, Gierl ve Walker, 2003)'un tek boyutlu madde parametreleri;

$$P\{U_i = 1|\theta\} = N\{a_i(\theta - b_i)\} = N\{a_i\theta - a_i b_i\}$$

ile ilişkili olduğu görülmektedir. Bu eşitlikte β_{i0} , $-a_i b_i$ ifadesine ve β_i 'de madde ayırıcılık parametresine (a_i) karşılık gelmektedir. Bu iki yaklaşımın bir uzantısı olarak iki boyutlu normal ogive modeli

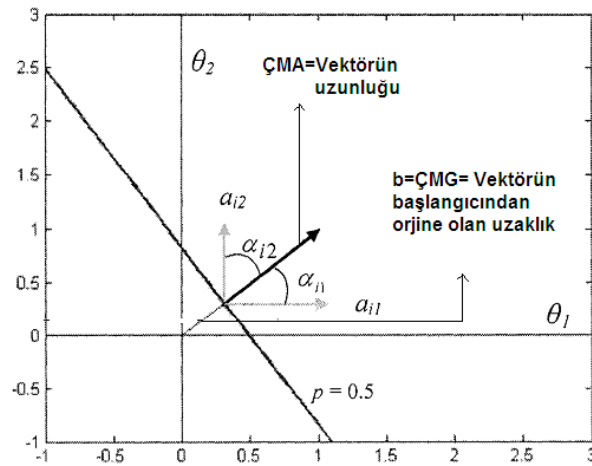
$$P\{U_i = 1|\theta_1, \theta_2\} = \frac{1}{1 + e^{-1.7(a_1\theta_1 + a_2\theta_2 + d)}}$$

(d.....)



Şekil 1. İki Boyutlu Bir Maddenin Madde Karakteristik Yüzeyi ($\alpha_1=2.0$, $\alpha_2=0$, $d=0.5$)

Tek boyutlu MTK'da madde karakteristik eğrilerinin yerini çok boyutlu MTK'da madde karakteristik yüzeylerinin alması, bazı terminolojilerin de değişmesi anlamına gelmektedir. Bu kavramları Reckase (1985, 1997) çok boyutlu madde güçlüğü-ÇMG (multidimensional item difficulty) ve çok boyutlu madde ayırıcılığı-ÇMA (multidimensional item discrimination) olarak ifade etmiştir. ÇMG ve ÇMA'in her ikisi de çok boyutlu maddeleri özetleyen değerlerdir.



Şekil 2. Madde Vektör Grafiği

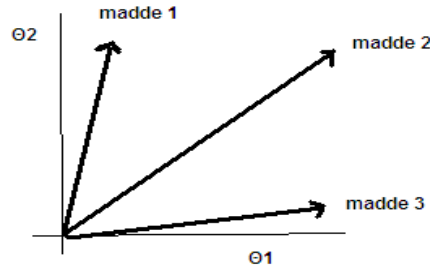
ÇMA (Şekil 2.), maksimum ayırıcılık parametresi olarak da bilinen çok boyutlu madde ayırt ediciliği, tek boyutlu MTK'da madde ayırt ediciliğine karşılık gelmektedir. Tek boyutlu modellerdeki

madde ayırt ediciliğinden farklı olarak çok boyutlu MTK’da madde ayırcılığı vektör (\vec{a}) olarak ifade edilir. Bu vektör her iki boyut için ortak madde ayırt edicilik gücü olarak tanımlanır (Smith, 2009). Matematiksel ifadesi;

$$\text{ÇMA} = \sqrt{\sum_{k=1}^m a_{ik}} \quad \text{veya} \quad \text{ÇMA} = \sqrt{a_{i1}^2 + a_{i2}^2} \quad \text{‘dir.}$$

Bu vektörel uzunluk maksimum ayırcılığı temsil eder. İfadedeki a_1 ve a_2 her iki örtük özellik için madde ayırcılık değerlerini ifade etmektedir (Zhang, 2008).

ÇMA, maddenin toplam ayırt ediciliğini ifade eden bir vektör olduğuna göre, bu vektörün bir açısıl yönünün (direction) de olması gereklidir. Bu açı değeri θ_1 eksenine yakın ise birincil olarak θ_1 yeteneğini (Şekil 3.-madde 3), θ_2 eksenine yakın maddelerin ise θ_2 yeteneğini (Şekil 3.-madde 1), birincil olarak ölçtüğü söylenebilir. θ_1 ve θ_2 eksenleri arasında 45^0 ’lik açı yapan maddeler ise θ_1 ve θ_2 yeteneklerinin her ikisini de eşit olarak ölçen maddelerdir (Şekil 3.-madde2).



Şekil 3. Çok Boyutlu Farklı Örtük Özellik Kompozisyonundaki Maddelerin Grafikselleştirilmesi

Reckase (1997) ve Zhang (2008), çok boyutlu madde ayırt ediciliği vektörünün yönünü

$$\alpha_i = \arccos \frac{a_{i1}}{\text{ÇMA}} \quad \text{olarak tanımlamıştır.}$$

Örnek olarak iki boyutlu bir maddenin ayırcılık parametreleri;

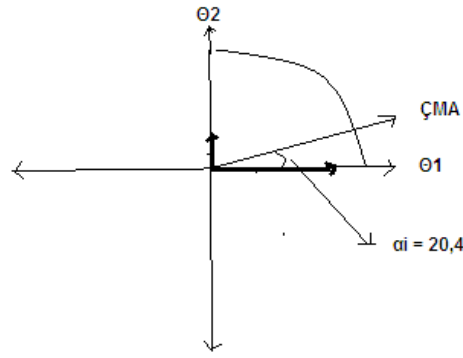
$a_1 = 0.652$ ve

$a_2 = 0.245$ olarak hesaplanmış olsun. Hesaplanan a_1 değeri birinci boyuta ait ayırcılık parametresini, a_2 ise ikinci boyuta ait ayırcılık parametresini ifade etmektedir. Bu maddeyi temsil eden toplam madde ayırt edicilik parametresi;

$\text{ÇMA} = \sqrt{a_{i1}^2 + a_{i2}^2}$ ‘dir ve $\text{ÇMA} = \sqrt{0.652^2 + 0.245^2} = 0.696$ olarak hesaplanır. Bu vektörün yönü ise;

$$\alpha_i = \arccos \frac{a_{i1}}{\text{ÇMA}} = \arccos \frac{0,652}{0,696} = 20,4 \text{ bulunur.}$$

Hesaplanan bu değer maddeyi temsil eden madde ayırcılık parametresinin birinci örtük özellik ile 20, 4 derecelik bir açı yaptığını göstermektedir. Yani madde büyük oranda birinci örtük özelliği ölçmektedir (Şekil 4).



Şekil 4. ÇMA'nın Grafik Üzerinde Gösterimi

ÇMG (Şekil 4), tek boyutlu MTK'da madde güçlüğüne (b) karşılık gelen, en iyi ayırtıcılık gücünü veren ve orijinden θ uzayına en dik eğim noktasına uzaklıktır. Parametrenin matematiksel ifadesi;

$$\text{ÇMG} = \frac{-d_i}{\text{ÇMA}} \text{ şeklindedir.}$$

Formüldeki d_i i maddesinin yer parametresini temsil etmektedir. Formülden elde edilen ve orijinden $p=.50$ noktasına olan uzaklığın ifadesi olan ÇMG'nün işareti maddenin bağıl güçlüğünü (relative difficulty) ifade eder. Negatif ÇMG değeri maddenin bağıl olarak daha kolay olduğunu, pozitif ÇMG değeri maddenin bağıl olarak daha zor olduğunu ifade etmektedir (Kao, 2007).

Çok Boyutlu MTK'da Bilgi Fonksiyonu

Çok boyutlu MTK'da bilgi fonksiyonu, tek boyutlu MTK'daki bilgi fonksiyonunun bir uzantısı olup, fonksiyonun matematiksel gösterimine bilginin doğrultusunun (direction of information) eklenmesi ile gösterilmektedir. Çok boyutlu bilginin matematiksel ifadesi;

$$\text{Çok Boyutlu Bilgi} = P_i(\theta) [1 - P_i(\theta)] \left(\sum_{k=1}^m \alpha_{ik} \cos \alpha_{ik} \right)^2 \text{ şeklindedir. Formülde;}$$

$P_i(\theta)$ θ yetenek düzeyindeki cevaplayıcının i maddesine doğru cevap verme olasılığını, α_{ik}örtük özellik kompozisyonundaki i maddesini temsil eden vektördür. Bu vektör Θ_1 eksenine ile yapılan açı ile temsil edilir (Ackerman, 2005).

Çok Boyutlu Madde Tepki Kuramı Analizlerinde Kullanılan Yazılımlar

Artan ihtiyaçlar doğrultusunda, MTK parametre tahminleri için bir çok bilgisayar programı geliştirilmiştir. Çok boyutlu MTK modellerinde kullanılan programlar MIRTE, MAXLOG, TESTFACT ve NOHARM'dır. Bu programlar hem tek boyutlu, hem de çok boyutlu modeller için analiz yapabilme esnekliğine sahiptir. Bu programlar da kullandıkları tahmin yöntemleri ve yetenek kestirimi yapıp yapamamalarına bağlı olarak birbirinden farklılaşmaktadır (Lee, 2007).

TESTFACT programı, marjinal en çok olabilirlik algoritmasını kullanarak tam bilgi madde faktör analizi ile madde parametrelerini tahmin etmektedir. Program madde güçlük, ayırt edicilik ve yetenek parametresini kestirebilmekte fakat şans parametresini kestirememektedir. NOHARM programı en küçük kareler ve tetrakorik korelasyon algoritması ile faktör analizine yaklaştırma tekniği ile madde parametrelerini kestirebilmektedir. NOHARM programı sadece madde güçlük ve ayırt edicilik parametresini kestirebilmekte, şans ve yetenek parametresini ise kestirememektedir (Bobcock, 2009).

Tek ve Çok Boyutlu MTK'nın Karşılaştırmasını Temel Alan Araştırmalar

Chang (1992)'ın yaptığı araştırmada, Woodcock-Johnson Psycho-Educational Battery-Revised testinin okuma becerisi ve matematik alt test verileri ile suni veri kullanılmış ve yetenek

kestirimleri tek ve çok boyutlu madde tepki kuramı modelleri altında incelenmiştir. Araştırma sonucunda, çok boyutlu MTK ile elde edilen sonuçların kuramsal ve ampirik olarak daha büyük ölçme keskinliğine sahip olduğu belirlenmiştir. Bunun yanında çok boyutlu kuram altında elde edilen yetenek kestirimlerinin, tek boyutlu kuram altında elde edilen yetenek kestirimlerine kıyasla, daha düşük standart hatalarının olduğu ve daha keskin yetenek kestirimlerinin elde edildiği vurgulanmıştır.

Spencer (2004) suni veri üzerinde yapmış olduğu çalışmada, tek ve çok boyutlu madde tepki kuramı modelleri altında model ve yetenek parametre kestirimlerinin karşılaştırmasını yapmıştır. Bu araştırmada ayrıca kullanılan yazılımlar da karşılaştırılmıştır. Araştırma sonucunda, cevaplayıcıların çok boyutlu örtük özellikleri, çok boyutlu ölçme modeli ile tek boyutlu ölçme modeline göre daha hassas ölçümlere ulaşıldığı belirlenmiştir. Ayrıca çok boyutlu veri matrisine, çok boyutlu ölçme modelinin daha iyi uyum gösterdiği belirlenmiştir.

de la Tore ve Patz (2005), araştırmalarında telafisel tipte 3-parametrelili çok boyutlu MTK modelinin, tek boyutlu MTK'ya göre daha iyi kestirimlerde bulunup bulunamayacağını test etmişlerdir. Araştırma sonucunda, örtük özellikler arası ilişki büyüklüğü 0.00 olduğunda çok boyutlu MTK'nın daha iyi kestirimlerde bulunduğu, ilişki büyüklüğü 1.00'a doğru yaklaştığında ise tek boyutlu MTK'nın tercih edilmesi gerektiği vurgulanmıştır.

Seungho Yang (2007) yapmış olduğu araştırmada tek boyutluluk varsayımının ihlalinin parametre tahmini ve model veri uyumuna olan etkisini tek ve çok boyutlu modeller altında incelemiştir. Araştırma sonucunda, yetenek parametresi kestiriminde en büyük etkiyi test uzunluğunun yaptığı, örneklem büyüklüğünün ise madde güçlük parametresi kestiriminde etkisi olduğu gözlenmiştir. Diğer araştırmalardan farklı olarak testlerde örtük özellikler arasındaki ilişki orta düzeyden yüksek düzeye kadar olan durumlarda çok boyutlu MTK modellerinin daha kesin kestirimlerde bulunduğu, ancak örtük özellikler arası ilişki 0.0'a yaklaştıkça tek boyutlu Rasch modelinin daha kesin kestirimlerde bulunduğu ortaya konmuştur.

Yukarıda kısaca özetlenen araştırmalar ve ilgili alanyazın göstermiştir ki, çok boyutlu MTK ile ilgili uygulamalar eğitim ve psikoloji alanındaki yerini gittikçe artırmaktadır. Bunun yanında testlerin veya testlerde bulunan maddelerin sadece tek bir örtük özelliği ölçtüğünü varsaymak ve bu varsayım altında ölçmeler yaparak birey hakkında karar vermek artık daha tartışılır hale gelmiştir. Buna karşın tek boyutlu MTK'nın bu karşılanması oldukça güç olan varsayımına karşın çok boyutlu MTK modellerinin kullanılması durumlarında daha keskin, geçerli ve güvenilir sonuçlar alındığı belirlenmiştir. Ancak çok boyutlu MTK'da cevaplayıcının maddeye doğru cevap verme olasılığını grafiksel olarak gösteren yazılımların hala çalışma aşamasında olması ve araştırmacıların kullanımına hizmet edememesi bu konudaki çalışmaların temel eksikliği olmaya devam etmektedir.

Tartışma

Eğitim ve psikoloji alanında tek boyutlu MTK modelleri yaygın olarak kullanılmasına rağmen, kuramın tek boyutluluk varsayımının özellikle başarı ve yetenek testlerinde tamamıyla karşılanması oldukça güçtür (Hambleton, Swaminathan ve Rogers, 1991; Reckase, 1997). Tek boyutluluk varsayımının karşılanmadığı durumlarda tek boyutlu modeli, çok boyutlu test verisine uygulamak yetenek ve madde parametreleri kestirimlerinde geçerlik sorunlarını beraberinde getirecek ve model-veri uyumunda önemli problemler ortaya çıkaracaktır. Ackerman (1994a-b)'ın belirttiği üzere cevaplayıcının performansını etkileyen birden fazla örtük özellik varsa, çok boyutlu MTK modelleri kullanılmalıdır. MTK modelleri için yapılan çalışmalarda (Drasgow ve Parsons, 1983, Hambleton, 1969; Aktaran; Kreiter, 1993) orta derecede çok boyutlu verilere (moderate degrees of multidimensionality), tek boyutlu modellerin uygulanması güçlü bir yapı göstermesine rağmen, veri grubunda çok boyutluluk arttıkça bilgi kaybına ve yetenek kestirimlerinin yanlış yorumlanmasına yol açmaktadır. Tek boyutluluk varsayımının ihlali model veri uyumsuzluğunu da beraberinde getirmektedir. Tek boyutlu modeller altında yapılan model veri uyumu testlerinde ortaya çıkan uyumsuzlukların en büyük nedeni test verisinin çok boyutluluğudur. Yukarıda alanyazında çok boyutlu modellerin uygulanması hakkında özetlenen araştırmalar göstermiştir ki, testlerin çok boyutlu olduğu durumlarda yetenek ve madde parametrelerinin kestirimi daha geçerli ve güvenilir sonuçlar vermiştir.

1987-2006 yılları arasında Amerika Birleşik Devletleri Merkez Bankası Başkanı olarak çalışan Alan GREENSPAN 2007 yılında yazmış olduğu ‘Türbülans Çağı (The Age of Turbulance)’ adlı kitabının bir bölümünde, geliştirilen ekonomik modelleri tartışmış, daha spesifik ve parametre sayısının fazla olduğu modellerin daha geçerli çıkarsamalarda bulunduğunu belirtmiştir. Paralel bir düşünce ile Walker ve Beretvas (2003) tek ve çok boyutlu madde tepki kuramı modellerini karşılaştırmış ve daha kompleks modellerin daha az hatalı, daha güvenilir ve model veri uyumu daha yüksek olduğunu ortaya koymuştur. Eğitim alanında araştırma yapanların amacı, bireyin verdiği tepkilerden yararlanarak, bireyi karakterize eden en geçerli ve en güvenilir sonuçlara ulaşmaktır. Bireyin verdiği tepkilerin en doğru biçimde analiz edilmemesi, birey hakkında verilecek yanlış kararlara da temel hazırlayacaktır.

Psikometri alanındaki en büyük gelişmelerden biri de MTK’ya dayalı çalışmalardır. Kuramın tek boyutlu modelleri eğitim araştırmacılarına uzun yıllar hizmet etmiş ve etmeye devam etmektedir. Kuramın çok boyutlu uzantısı olan çok boyutlu MTK, elektronik yazılımların ilerlemesi ile alanda hizmet etmeye başlamış, artan bilgi ve teknoloji ile araştırmacıların beklentilerini karşısında gelişmeye devam edecektir.

Kaynaklar

- Ackerman, T.A. (1992). *Assessing construct validity using multidimensional item response theory*. Paper Presented at the Annual Meeting of American Educational Research Association. San Fransisco, CA, USA.
- Ackerman, T.A.(1994a). *Graphical Representation of Multidimension-al Item Response Theory Analyses*. Paper Presented at the Annual Meeting of American Educational Research Association. New Orleans, LA.
- Ackerman, T.A. (1994b). Using multidimensional item response theory to understand what items and tests are measuring. *Applied Measurement in Education*, 7(4), 255-278. Web: <http://www.dx.doi.org/10.1207> adresinden 5 Eylül 2008 tarihinde alınmıştır.
- Ackerman, T.A., Gierl, M.J., Walker, C.M. (2003). Using multidimensional item response theory to evaluate educational and psychological tests. *Educational Measurement: Issues and Practice: MIRT Instructional Module*.
- Ackerman, T.A. (2005). Multidimensional item response theory modeling. In J.J. McArdle (Ed). *Contemporary Psychometrics* (p.3-24). Web:<http://www.questia.com/reader/printPaginator/1580> adresinden 13 Mayıs 2009’da alınmıştır.
- Antal, T. (2007). On multidimensional item response theory a coordinate free approach. *Electronic Journal of Statistics*, 1, 290-306.
- Bobcock, B.G.E.(2009). *Estimating a Noncompensatory IRT Model Using a modified Metropolis algorithm*. Unpublished Doctoral Dissertation. The University of Minesota.
- Bock, R.D. (1997). A brief history of item response theory. *Educational Measurement: Issues and Practice*. Winter 1997.
- Bock, D.R., Thissen, D. ve Zimowski, M.F. (1997). IRT estimation of domain scores. *Journal of Educational Measurement*, 34 (3), 197-211.
- Bolt, D.M. ve Lall, V.F. (2003). Estimation of compensatory and non-compensatory multidimensional item response models using Markov Chain Monte Carlo. *Applied Psychological Measurement*, 27, 395. Web: <http://apm.sagepub.com> adresinden 3 Nisan 2008’de alınmıştır.
- Chang, Y.W. (1992). *A comparison of unidimensional and multidimensional IRT approaches to test information in a test battery*. Unpublished Doctoral Dissertation. University of Minnesota. de la Tore, J. ve Patz, R.J. (2005). Making the most of what we have: A practical application of multidimensional item response theory in test ccoring. *Journal of Educational and Behavioral Statistics*, 30(3), 295-311. Web: <http://www.jstor.org/stable/3701380> adresinden 12 Ocak 2009 tarihinde alınmıştır.
- Drasgow ve Parsons (1983). Application of unidimensional item response theory models to multidimensional data. *Applied Psychological Measurement*, 7, 189-199.
- Embretson, S.E. ve Reise, S.P. (2000). *Item Response Theory For Psychologists*. Lawrence Erlbaum Associate, Inc.
- Greenspan, A. (2007). *The age of turbulence*. The Penguin Pres. New York, 2007.
- Hambleton, R.K. ve Swaminathan, H. (1989). *Item Response Theory. Principles And Applications*. Kluwer-Nijhoff Publishing. Boston-USA.
- Hambleton, R.K. (1994). Item response theory: A broad psychometric framework for measurement advances. *Psicothema*, 6, 535-536.

- Hambleton, R. K., Swaminathan, H. and Rogers, H. (1991). *Fundamentals of Item Response Theory*. Newbury Park CA: Sage.
- Kao, S.C. (2007). *The new goodness of fit index multidimensional item response model*. Unpublished Doctoral Dissertation. Michigan State University.
- Kreiter, C.D. (1993). *An empirical investigation of compensatory and noncompensatory test items in simulated and real data*. Unpublished Doctoral Dissertation. The University of Iowa.
- Lee, S. H. (2007). *Multiitem response theory: A SAS MDIRT MACRO and empirical study of PIAT MATH Test* Unpublished Doctoral Dissertation. The University of Oklahoma.
- Li, Y. H. ve Schafer, W.D (2005). Trait parameter recovery using multidimensional computerized adaptive testing in reading and mathematics. *Applied Psychological Measurement*, 29, 3-25. Web: <http://www.sagepub.com> adresinden 3 Nisan 2008 tarihinde alınmıştır.
- McDonald, R.P. (1982). Linear versus models in item response theory *Applied Psychological Measurement*, 6, 379-396.
- McDonald, R.P. (2000). A basis for multidimensional item response theory. *Applied Psychological Measurement*, 24, 99. <http://www.sagepub.com> adresinden 3 Nisan 2008 tarihinde alınmıştır.
- Pomplun, M.R. (1988). *Effects of local dependence in achievement tests on IRT ability estimation*. Unpublished Doctoral Dissertation. The Florida State University.
- Reckase, M.D. (1997). *Models for multidimensional tests and hierarchically structured training materials*. Technical Report. The American College Testing Program. Iowa City, Iowa.
- Seungho Yang, M. A. (2007). *A Comparison of unidimensional and multidimensional rasch models using parameter estimates and fit indices when assumption of unidimensionality is violated*. Unpublished Doctoral Dissertation. The Ohio State University
- Sijtsma, K. ve Junker, B.W. (2006). Item response theory: past performance. Present developments and future expectations. *Behaviormetrika*, 1, 75-102.
- Smith, J. (2009). *Some issues in item response theory: Dimensionality assessment and models for Guessing*. Unpublished Doctoral Dissertation. University of South California.
- Spencer, G.S. (2004). *The strength of multidimensional item response theory in exploring construct space that is multidimensional and correlated*. unpublished doctoral dissertation. Brigham Young University.
- Walker, C.M. ve Beretvas, S.N. (2003). Comparing multidimensional and unidimensional proficiency classifications: multidimensional IRT As a diagnostic aid. *Journal of Educational Measurement*, 40 (3), 255-275.
- Zhang, B. (2008). Application of unidimensional item response models to tests with items sensitive to secondary dimension. *The Journal of Experimental Education*, 77 (2), 147-166.
- Zhao, Y. (2008). *Approaches for addressing the fit of item response theory models to educational test data*. Unpublished Doctoral Dissertation. University of Massachusetts Amherst.