

# ÖĞRENCİ YETENEĞİNİN KESTİRİMİNDE BİLİŞSEL TANI MODELLERİ VE UYGULAMALARI

T. Oğuz BAŞOKÇU\*

## ÖZET

Bilişsel Tanı Modelleri (BTM), temelinde örtük sınıf analizi olan yaklaşımlardır. Örtük sınıf analizi, çok değişkenli kategorik bir veri kullanarak ve birbiriyle ilişkili durumlardan yararlanarak alt gruplar belirleyen istatistiksel bir yöntemdir. BTM ise cevaplayıcıda belirli bir bilginin yapısını ya da bir becerinin gelişimini, cevaplayıcının bilişsel düzeydeki güçlü ve zayıf yönlerini dikkate alarak hesaplamak amacıyla geliştirilmiştir (Leighton ve Gierl, 2007). Bu modellere göre öğrencilerin testteki maddelere verdikleri cevaplar, onların ait oldukları örtük sınıflarının bir vektörüdür. Bu nedenle modeller, maddelere verilen cevaplardan yola çıkarak öğrencilerin örtük sınıflarını belirlemeyi amaçlar. Bu amaç doğrultusunda geliştirilen farklı modeller vardır. Bu çalışmada BTM olarak adlandırılan bu modellerin temel özellikleri, yapıları ve kullanılan modellerin uygulamaları hakkında bilgiler verilecektir.

**Anahtar Sözcükler:** Bilişsel Tanı Modelleri, Çok Boyutlu Madde Tepki Kuramı, Örtük Sınıf Analizi, Q Matrisi

## THE COGNITIVE DIAGNOSTIC MODELS FOR ESTIMATING STUDENTS' ABILITY AND THEIR APPLICATIONS

## ABSTRACT

Cognitive Diagnostic Models (CDM) are based on latent class analysis. Such analysis is a statistical method that determines subclasses by using multiple variable categorical data and making use of mutually related cases. CDM, on the other hand, is developed to calculate the structure of a certain knowledge or the development of a certain capacity by taking into account both strengths and weaknesses of the respondent in cognitive terms (Leighton ve Gierl, 2007). According to these models, replies of the students to the options of the test compose a vector of latent classes they belong. Thus, these models, through replies to the options, aim to determine the latent classes of students. To this end, various models are introduced. The aim of this paper is to study basic qualities, structures and applications of CDM.

**Keywords:** Cognitive Diagnostic Models, Multidimensional IRT, Latent Class Analyze, Q Matrix

---

\* Yrd. Doç. Dr. T. Oğuz BAŞOKÇU, Ege Üniversitesi, Eğitim Bilimleri Fakültesi, Eğitimde Ölçme ve Değerlendirme ABD. tahsin.oguz.basokcu@ege.edu.tr

## 1. GİRİŞ

Bilişsel Tanı Modelleri (BTM terimiyle ifade edilen modellere literatürde; bilişsel psikometrik modeller (cognitive psychometric models) (Rupp, 2007), bilişsel tanı modelleri (cognitive diagnosis models) (Nichols, Chipman ve Brennan, 1995), sınırlandırılmış örtük sınıf modelleri (restricted latent class models) (Haertel, 1989), çoklu sınıflandırma modelleri (multiple classification models) veya yapısal madde tepki kuramı modelleri (structured item response theory models) (Rupp ve Mislevy, 2007) gibi farklı isimler verilmiştir.

Bütün bu farklı tanımlamaların her biri, modellerin belirli özelliklerine yönelik çağrışımları ifade etmektedir. Bazı tanımlamalar modellerin teorik alt yapısına göre belirlenirken bazıları modellerin kendine özgü amacını ifade etmektedir. Diğerlerinde ise modellerin istatistiksel özellikleri dikkate alınmıştır. Çalışmada kullanılan Bilişsel Tanı Modeli teriminin temelinde ise modelin ayırt edici özelliğinin yerleştirme, kabul ya da sertifikasyondan çok “tanı” olduğunun düşünülmesi yatmaktadır.

### 1.1. Bilişsel Tanı Modellerinin Genel Özellikleri

BTM'nin amacı, iki kategorili niteliklerin sıralanışına dayanan örtük kategorilere göre cevaplayıcıları sınıflamaktır. Bu modellerde uzmanlığı belirleyen bir vektör olan örtük değişkenler, öğrenciye yönelik tanının altında yatan nitelik setini tanımlar. Burada “nitelik” kavramı vasıf, görev, alt görev, bilişsel süreç veya beceri olarak tanımlanabilir. (Tatsuoka, 1995a, s.330).

BTM uygulayarak geliştirilmiş bir testte, bir toplam puan ya da toplam alt ölçek puanları yerine, sınavı alan her bir bireyin spesifik olarak kapsamda hangi becerilere sahip olduğu ve hangileri konusunda eksiklikleri bulunduğu belirlendiği bir ölçüm yapılıdır. Bu anlamda BTM ile geliştirilmiş testler sadece değerlendirme sürecine değil aynı zamanda her bir öğrencinin eğitim ihtiyacını belirleme konusuna da hizmet eder (Cheng ve Chang, 2007).

Örtük özellik kuramı ve BTM, Madde Tepki Kuramı (MTK) modelleri ile yakından ilişkilidir. BTM içindeki birçok parametre kestirimi, MTK fonksiyonlarından yararlanarak geliştirilmiştir. Fakat BTM modelleri çok boyutlu yapısı dolayısıyla MTK'dan farklı özellikler göstermektedir. MTK modelleri bireylerin maddelere verdikleri cevaplara dayanarak bireyin örtük özelliğini kestirmek konusunda oldukça başarılı araçlardır. MTK modelleri madde ve yetenek parametrelerini sürekli bir ölçek üzerinde kestirirken “a”, ayıt edicilik, “b”, güçlük ve “c” tahmin parametrelerini hesaplar. Fakat MTK modelleri madde güçlüğünün nasıl meydana geldiği konusunda bir fikir vermez. Gerçekte ise eğitim, temel olarak MTK modellerinin bize sunduklarının ötesinde öğrenme alanının yapısını, öğrencinin psikolojik yapısını ve maddede yoklanan görevi çözebilme becerisinin altındaki psikolojiyi doğru analiz edebilmeyi hedefler (Almond, Steinberg ve Mislevy, 1993). BTM, bu açığı gidermeye odaklanmakta ve öğrencinin maddeyi çözerken kullandığı bilişsel süreçleri de dikkate alarak test cevaplama örüntüsünü öğrencinin çok boyutlu profilini çıkartmak amacıyla kullanmaktadır.

MTK'da öğrencinin yeteneği genel yetenek parametresi olan  $\theta$  ile tahminlenir.  $\theta$  ve gözlenen puanlar belirli bir öğrencinin performansını ve grup içindeki sırasını kestirme imkânı tanır. BTM ise bunların dışında öğrencinin zihinsel süreci ve performansını altında yatan bilginin yapısını da ortaya çıkartabilir. BTM uygulamalarında her test birden fazla özelliği ölçer ve her madde birden fazla özellik ile ilişkilendirilebilir. Bu şekilde öğretmen sınav sonunda tek bir puan vermek yerine BTM ile öğrenci performansını, sahip olduğu becerileri de göstererek düzenleyebilir. Bu durumda öğretmen öğrencinin bir beceri profilini çıkartıp eksik olan becerilerini belirleyerek bu becerilerin üzerine eğilebilir. Diğer taraftan, eğitim sürecinin güçlü ve zayıf yönlerini ortaya çıkartarak daha iyi öğrenme stratejileri düzenlenebilir.

Görüldüğü gibi BTM öğrencilerin örtük ölçekteki yetenek düzeylerine odaklanmaz, her bir bilişsel öge üzerindeki öğrenci performansını ve bu ögeye sahip olma olasılığını belirler. Elde edilen olasılıklar öğrencinin uzmanlaştığı becerilerin profilini çıkartacak şekilde dönüştürülür. BTM'de de tıpkı yapısal eşitlik modellerinde olduğu gibi örtük değişken ile maddeler arasındaki ilişkilerin analizden önce belirlenmesi gerekmektedir. Modellerde bu süreç Q matrisi olarak adlandırılan madde özellik matrisiyle tanımlanır.

#### 1.1.1. Q Matrisi

Q matrisi, bir test için nitelik ya da beceri vektörlerinin sütunlarda, maddelerin ise satırlarda temsil edildiği bir örüntüdür. Nitelikler, alan uzmanları tarafından belirlenen özellikler, prosedürler, buluş yöntemleri, stratejiler, beceriler ve başka bilgi bileşenleri olabilir. Q matrisi 1-0 şeklinde ikili numaralar kullanarak özelliğin maddede bulunup bulunmadığını gösterir. Bu numaralandırma ilk önce Fisher (1973) tarafından "ağırlıklandırılma" şeklinde tanımlanmış ve bir k niteliğinin bir i maddesinde bulunduğu durumda 1, bulunmadığı durumda ise 0 şeklinde kodlandığını belirtmiştir.

$J \times K$  matrisini  $Q = [q_{jk}]$  0 ve 1 için şu şekilde yazabiliriz;

$$q_{jk} = \begin{cases} 1 & \text{eğer } k \text{ özelliği } , j \text{ maddesinde bulunuyorsa} \\ 0 & \text{eğer } k \text{ özelliği } , j \text{ maddesinde bulunmuyorsa} \end{cases}$$

Tablo 1.1'de 3 nitelik ve 5 madde için hazırlanmış örnek bir Q matrisi görülmektedir. Q matrisinde görüldüğü gibi 1. maddeyi doğru cevaplamak için ilk niteliğe sahip olmak gerekmektedir. İkinci madde için ise 1. ve 2. niteliklere öğrencilerin sahip olması gerekmektedir.

Maddeler	$\alpha_1$	$\alpha_2$	$\alpha_3$
1	1	0	0
2	1	1	0
3	1	1	1
4	0	1	1
5	0	0	1

Ölçülen niteliklerle birlikte düzenlenen Q matrisi, madde güçlüğü'nün psikometrik model ve bilişsel süreçle arasındaki bağlantısını kurar. Modelin seçimi ve Q matrisi birbirine eşit sınırlardan oluşan, örtük sınıf cevaplama olasılıklarını belirlemeyi sağlar.

Son dönem yapılan araştırmalarda Q matrisinin doğruluk miktarını ve etkisini ölçmeye yönelik sağlam kanıtlar sunan metotlar geliştirilmektedir. Henson (2004), Q matrisin kalitesinin, öğrencilerin beceri profilinin kestirimini doğrudan belirlediğini ifade etmektedir. Bu nedenle, Q matrisin yapısı ve veri model uyumu “tanısal analiz” için çok önemli bir nitelik taşımaktadır (de la Torre & Douglas, 2004).

### **1.1.2. BTM Hata Kaynakları**

BTM’de testi alan bireyin yanlış sınıflandırılmasına neden olan farklı hata kaynakları bulunmaktadır. Maddenin yapılandırılması, Q matris ile ilişkilendirilmesi veya cevaplanması aşamasında oluşabilecek bu hata kaynakları modellerin yapısı incelenirken ve modellere ilişkin parametreler belirlenirken öncelikli önem taşımaktadır. DiBello ve arkadaşları (2007) sınıflamalarda oluşabilecek hatalarla ilgili dört önemli faktöre işaret etmektedir.

**Strateji:** Cevaplayıcının maddeye cevap verirken Q matris tarafından listelenen özelliklerin dışında farklı bir yöntem kullanması, BTM içinde belirlenen cevaplayıcı profilinin hatalı olmasına neden olmaktadır.

**Tamlık (Completeness):** Bir test maddesinin Q matris tarafından belirlenen özelliklerden farklı bir özelliğe daha sahip olması durumudur. Bu durumda Q matrisin tamlık özelliği göstermediği söylenebilir.

**Pozitiflik düzeyi (Positivity):** Bazı durumlarda özelliğe sahip olan bir birey o özellik altında tanımlanan bir maddeyi yanlış cevaplayabilir, benzer şekilde özelliğe sahip olmayan bir birey o özellik altında tanımlanan bir maddeyi doğru cevaplayabilir. Bu sonucu doğuran temel neden Q matris tarafından belirlenen özelliklerin temelde bir genellemeye dayanmasıdır. Özellikler girift yapılar değildir ve alt gruplara ayrılma potansiyeli taşımaktadırlar. Bu durumda sınıflama sonucunda belli bir özelliğe sahip olduğuna karar verilen bir bireyin, o özelliğin alt becerilerinden bir ya da birkaçına sahip olmaması mümkündür. Genel olarak bireyin özelliğe sahip olma düzeyini ikili (dikotomus) bir değişken olarak tanımlayan modellerde pozitiflik düzeyi bir hata kaynağı olarak gözlenmektedir.

**Kaydırma (Slip):** Birey tarafından gerçekleşen rastsal hatalar kaydırma olarak tanımlanır. Kaydırma arta kalan ve sistematik olmayan bütün hataları içeren kategoridir.

### **1.2. BTM ve Tarihsel Gelişimi**

BTM’nin temelinde madde ve bilgi arasındaki etkileşim vardır. Bu nedenle madde-bilgi etkileşimi üzerinde duran yaklaşımları özetlemek BTM’nin yapısını tanımlamayı kolaylaştıracaktır.

BTM’nin çıkış noktası Lineer Lojistik Test Modeli (LLTM) ile (Fischer, 1973, 1976, 1977; Spada, 1977) Tatsuoka ve Tatsuoka’nın (1995) Rule Space modelidir. LLTM ise temel Rasch modelinin, madde cevaplama aşamasına karışan bilişsel süreçleri de içerecek şekilde genişletilmiş halidir. (DiBello, Roussos & Stout, 2007).

Temel fonksiyonu MTK’ya dayanan BTM’nin bu anlamda MTK’ya parametrik olmayan ve çok boyutlu bir içerikle yaklaştığı görülmektedir. Bu anlamda geleneksel tek boyutlu MTK’nın çok boyutlu biçimde ele alınışını ve Madde Tepki Fonksiyonu

(MTF) üzerindeki değişiklikleri incelemek BTM'nin genel yapısını ortaya koymaya yardımcı olacaktır.

### 1.2.1. Tek Boyutlu Lojistik Modeller (1PL,2PL&3PL ve LLTM)

3PL modele ait MTF (Madde Tepki Fonksiyonu) aşağıdaki gibidir;

$$P(X_{ij} = 1 | \theta_j) = c_i + \frac{1 - c_i}{1 + \exp[-1.7a_i(\theta_j - b_i)]}$$

$\theta_j = j$  öğrencisinin maddenin tek boyutlu olarak ölçtüğü düşünülen yeteneğinin düzeyini;  $a_i = i$  maddesine ait ayırt edicilik parametresini ve MTF'nin maksimum eğimini;  $c_i =$  tahmin parametresini yani düşük yetenek düzeyindeki bir öğrencinin maddeyi doğru cevaplama olasılığını, MTF'nin düşük asimptotunu;  $b_i = i$  maddesinin güçlük parametresini tanımlar.  $c_i = 0$  olduğu durum 2PL modelidir.  $a_i = 1$  ve  $c_i = 0$  olduğu durum Rash modelini gösterirken ise 1PL modelde ise  $a$  parametresi maddelerin ayırtedicilik parametrelerinin ortalamaları alınarak sabitlenir. 1PL, 2PL ve 3PL modeller tek boyutludurlar. Bu nedenle herhangi bir çoklu yetenek setiyle doğrudan ilişki kurabilme imkânları yoktur.

Yukarıda genel modele ilişkin formülde görüldüğü gibi tek bir olası strateji vardır ve her madde için tek bir beceri bulunmaktadır ve bu beceri bütün maddeler için geçerlidir. Beceriye ilişkin başarılı performans aynı zamanda maddeye ilişkin başarılı performans anlamına gelmektedir ve maddeyi çözmek için gerekli olan bütün becerinin var olduğunun kabulü olarak tanımlanabilir (Q'nun eksiksizliği). Tam pozitiflik, beceriye sahip olan kişilerin maddeyi doğru, sahip olmayanların ise yanlış işaretlemesi durumudur. Bir beceriye sahip olan birey bazen maddeyi yanlış cevaplayabilir, aynı şekilde beceriye sahip olmayan birey de maddeyi doğru cevaplayabilir. Bu durum tam pozitifliğin var olmadığı anlamına gelir. Tam pozitiflikten kaymalara çoğunlukla becerinin tam olarak tanımlanamaması sebep olur. Tam pozitiflik eksikliğinin üstesinden beceriye sahip olma düzeyinin sürekli bir değişken haline dönüştürülmesi ve lojistik bir MTF üzerinde beceriye sahip olma düzeyinin doğru cevap verebilme olasılığıyla tanımlanması yoluyla gelinmeye çalışılmaktadır. Rastsal kaydırma "slip" parametresinin olasılığı ise "0" kabul edilir.

Değinilmesi gereken diğer bir model ise LLTM'dir. Bu model aslında yetenek değişkeninin çoklu bilişsel bir becerinin vektörü olarak tanımlamaz ve her bir öğrenci için bilişsel becerilerin tanısı mümkün değildir; ama madde güçlüğü analizinde güçlüğün faktörlerine yönelmek bir "bilişsel işlem" olarak düşünülebilir ve tarihsel bir basamak olarak BTM'nin çok boyutlu MTK modellerinden evrimini gösterebilir. LLTM mantıksal bir modelle, madde güçlük parametresini, güçlük temelli bilişsel yeteneklere ayırarak modelin bütününe ait bir güçlük parametresini tanımlar. LLTM her bir becerinin birden çok boyutu temsil ettiği çok boyutlu madde tepki modellerine benzer. Fakat güçlük parametresi her bir beceri için madde özelinde değildir, bu modelde her bir özelliğin testin tümüne ilişkin güçlük düzeyi bulunur. LLTM her bir madde için "bilişsel işlem" listesini barındıran bir Q matrisine sahiptir (Fischer, 1983). Sonuç olarak LLTM de yetenek tek boyutlu olsa da Q matrisi kullanarak Rasch modeli

gibi tek boyutlu modellerle bütünüyle tanılayıcı olan modeller arasındaki tarihsel yarı yolu gösteren bir yerde durmaktadır.

Fischer bu modele ilişkin MTF'nin Rasch'ın IPL modelinin sınırlandırılmış bir hali olduğunu söylemektedir. Madde güçlüğünün lineer ifadesini bir anlamda bilişsel işlemin madde güçlüğünü kısıtlaması olarak tanımlamaktadır. IPLM ilişkin çalışmalarda “bilişsel işlemler” modelin içinde yer almaktadır. Fischer bu bilişsel işlemlerin, öğrencinin yeteneğinin değil maddenin niteliğinin göstergesi olduğunu söylemektedir. LLTM Q matrisi ile madde güçlüğünün bilişsel belirteçleri tespit edildiğinde, doğal olarak sonraki adımın tümünde ölçülen yeteneğin her bir madde güçlüğü belirteci için faktörlerine bölünmesidir. Bu aşama Tatsuoka tarafından gerçekleştirilmiş ve her bir madde güçlüğüne ait bilişsel faktör haritalanarak bu faktörün gerektirdiği öğrenci yeteneğiyle madde performansı arasındaki ilişki bir bir gösterilmiştir. LLTM doğrudan öğrenci yetenek düzeyi ile madde arasındaki ilişkiyi gösteremez, fakat içinde Q matrisi bulunan ilk model olarak alandaki önemli adımlardan biridir.

LLTM'ye ait MTF, IPLM ile aynıdır sadece bi parametresi bulunur ve şöyle tanımlanır;

$$b_i = \sum_{k=1}^K q_{ik} \eta_k + c$$

$q_{ik} = i$  maddesinin çözümünde  $k$  yeteneğinin bir faktör olduğu durumun frekansdır  $\eta_k = k$  yeteneğinin güçlük özelliğidir ve  $c =$  ölçeğin orijiniinde belirlenebilecek keyfi bir sabittir.  $q_{ik}$  değişkeni her zaman ikilidir ve becerinin madde güçlüğünden etkilenip etkilenmediğine işaret eder.

Bu modelde de görüldüğü gibi tek bir olası strateji bulunmaktadır. Her ne kadar model bir madde için herhangi bir sayıda yeteneğe izin verse de ayrıştırılmış yeteneklere ilişkin bir performansın modellenmesi mümkün değildir. Tek bir örtük özellik kullanarak model, başarıyla bütün yeteneklere ilişkin düzeyleri tek boyutlu bir parametreye indirerek, yeteneklerin kombinasyonuna ilişkin bir değer verir. Bu  $\theta$  öğrencinin bütün yetenek düzeylerinin ağırlıklandırılmış ortalaması olarak kabul edilebilir.

### 1.2.2. Denkleştirici Çok-boyutlu MTK Modeli (Compensatory multidimensional IRT model- MIRT-C)

Bu model ilk kez Reckase ve McKinley (1991) tarafından ortaya atılmıştır. Fakat modelin bilişsel tanı alanında popülerlik kazanması Adams'ın (1997) çalışmalarıyla gerçekleşmiştir. Adams Rasch modelini temel alarak bütün ai parametrelerinin 1'e eşit olduğu, bütün ci parametrelerinin sıfıra eşit olduğu ve Q matrisinin hangi becerinin hangi maddeyle ilişkili olduğunu tanımladığı bir model geliştirmiştir. MIRT-C modeli için MTF aşağıdaki gibi yazılır;

$$P(X_{ij} = 1 | \theta_j) = c_i + (1 - c_i) \left[ 1 + e^{-1.7 \|\mathbf{a}_i\| \left( \frac{\mathbf{a}_i^T \theta_j}{\|\mathbf{a}_i\|} - b_i \right)} \right]^{-1}$$

$\theta_j = j$  öğrencisine ilişkin yeteneğe sahip olma düzeyinin vektörüdür,  $j$  öğrencisinin maddenin doğru cevaplanması için gerekli olan yeteneğe sahip olma düzeyinin göstergesidir;  $a_i = i$  maddesinin duyarlı olduğu bütün becerilere ilişkin ayırt edicilik parametresinin vektörüdür.  $c_i =$  tahmin parametresidir ve madde için gerekli olan bütün yeteneklerde düşük düzeyde olan bir öğrencinin maddeyi doğru cevaplama olasılığıdır.

$b_i = i$  maddesi için güçlük parametresidir ve  $\|\underline{a}_i\| = (\sum_{k=1}^K a_{ik}^2)^{1/2}$  dir.

Maddenin denkleştirici niteliği MTF'ye bakıldığında açıkça görülmektedir. MTF,  $c_i = 0$ ,  $b_i = 0$  ve  $a_1 = a_2 = 1$  olduğu iki boyut için oluşan durumda aşağıdaki gibidir;

$$P(X_{ij} = 1 \mid \theta_1, \theta_2) = \frac{1}{1 + e^{-1.7(\theta_1 \theta_2)}}$$

Eğer  $\theta_1$  çok düşük bir değer alırsa, örneğin  $\theta_1 = -2$  gibi bu durumda  $\theta_2$  bu düşük değeri telafi edici şekilde çok yüksek bir değer olarak  $\theta_2 = +2$  olur. Sonuçta  $P(X_{ij} = 1 \mid \theta_1, \theta_2) = 0,5$  olur ve bu değer  $\theta_1 = \theta_2 = 0$  olduğu durumla aynıdır. Açıkça görülmektedir ki bu modelde bir beceriye ilişkin yüksek düzey, bir başka beceriye ilişkin düşük düzeyle telafi edilmekte ve ikisi içinde orta düzey bir etki elde edilmektedir (Dibello ve ark. 2007). Bilişsel tanı alanı için bu model genel olarak tek boyutlu MTK modelinin geliştirilmiş ve çok boyutlu analize uygun hale getirilmiş durumudur. Fakat becerilere yönelik telafi edici doğası nedeniyle bu model birçok yönden bilişsel alanda tanı koymaya yönelik amaçları karşılayamamaktadır. Özellikle madde performansının bilişsel analizinde, genellikle o maddeye yönelik başarılı performans için maddenin gerektirdiği bütün becerilere sahip olunmasının gerekliliği vardır. Bu bağlamda bilişsel varsayımları karşılamak amacıyla, bilişsel perspektife daha uygun birleştirilmiş bir modele ihtiyaç duyulmaktadır.

### 1.2.3. Denkleştirici olmayan Çok-boyutlu MTK Modeli (Noncompensatory MIRT model- MIRT-NC)

Denkleştirici olmayan Çok-boyutlu MTK Modeli Sympson (1977) tarafından geliştirilmiştir. Bu modelde bilgi yapısı çok boyutludur, aynı zamanda becerilerle madde performans etkileşimi denkleştirilmemektedir. Bununla birlikte bir öğrencinin maddede aranan herhangi bir beceri için sahip olduğu düşük düzey, doğru cevap olasılığının da düşük olmasına neden olur, yani öğrencinin diğer becerilere yüksek düzeyde sahip olması modelde dikkate alınmaz. Bu model ilişkin MTF aşağıdaki gibidir;

$$P(X_{ij} = 1 \mid \underline{\theta}_j) = \prod_{k=1}^K \frac{1}{1 + e^{-1.7 a_{ik} (\theta_{jk} - \theta_{ik})}}$$

$\theta_{jk} = j$  öğrencisinin  $k$  yeteneğindeki düzeyi,  $a_{ik} = i$  maddesinin  $k$  yeteneği üzerindeki ayırt edicilik parametresi,  $b_{ik} = i$  maddesinin  $k$  yeteneği üzerindeki zorluk düzeyidir (Sympson, 1977).

Model tam olarak telafi edici olmadığından Dibello (2007) model için “denkleştirici olmayan” terimini kullanmıştır, fakat aynı zamanda bu model için “birleşik”

(conjunctive) terimi de kullanılmaktadır. Modelin tekrar formülasyonu denkleştirici olmayan doğasını ortaya koymaktadır, bu formüldeki bütün terimler 2PL modelin MTF gibidir;

$$Y_{ikj} = \begin{cases} 1 & \text{eğer } j \text{ öğrencisi } i \text{ maddesindeki } k \text{ becerisini başarıyla yaptıysa} \\ 0 & \text{diğer durumlar} \end{cases}$$

Bu durumda;

$$P(X_{ij} = 1 | \theta_j) = \prod_{k=1}^K P(Y_{ikj} = 1 | \theta_{jk})$$

Bu modelin kullanımı sadece öğrencilerin beklenen bütün becerilere sahip olması gerektiği durumlarda söz konusu olmaktadır, çünkü bu analizde becerilerin düzeyleri bir birinden bağımsız değerlendirilmektedir. Bu bağımsızlık nedeniyle bütün becerilerin başarıyla gerçekleştirilme olasılığı, her bir becerinin başarısının koşullu olasılıklarının çarpımı olmaktadır (Dibello ve ark., 2007).

#### 1.2.4. Çok bileşenli Örtük Özellik Modeli (Multicomponent latent trait model-MLTM)

Bu model Denkleştirici olmayan çok boyutlu MTK modelinin bir uzantısı olarak kabul edilebilir. Model Whitely (1980) ve Embretson (1985, 1997) tarafından geliştirilmiştir. MIRT-NC gibi her bir yeteneğin öğrenci tarafından başarıyla olmasını belirleyen çok boyutlu bir MTF vardır. Model bu MTF çarpımlarını bir madde için gerekli olan bütün becerilerin öğrenci tarafından başarıyla olmasını belirlemek amacıyla simültane olarak kullanılmaktadır. Ek olarak MLTM bir madde için öğrencinin başarılı performans olasılığını, öğrencinin bütün becerilere sahip olup olmadığı koşuluna göre modeller. MLTM'de bir maddeye ilişkin öğrenci performansının modellenmesi, öğrencinin Q-becerileri dışındaki performansının modellenmesini de içerir (Whitely, 1980).

$Y_{ikj}$  nin  $j$  öğrencisinin  $i$  maddesindeki  $k$  becerisini konusundaki başarısının ikili göstergesi olduğu durum için modelin MTF aşağıdaki gibidir;

$$P(X_{ij} = 1 | \theta_j) = \prod_{k=1}^K P(Y_{ikj} = 1 | \theta_{jk}) + g \left[ 1 - \prod_{k=1}^K P(Y_{ikj} = 1 | \theta_{jk}) \right]$$

Bu fonksiyonda  $\theta_{jk} = k$  becerisine  $j$  öğrencisinin sahip olma düzeyi,  $a = P(X_{ij} = 1 | \prod_{k=1}^K Y_{ikj} = 1)$ ,  $g = P(X_{ij} = 1 | \prod_{k=1}^K Y_{ikj} = 0)$ ,  $P(Y_{ikj}=1|\theta_{jk}) = \frac{1}{1+e^{-(\theta_{jk}-b_{ik})}}$  (1parametrelili lojistik fonksiyon) ve  $b_{ik} = k$  becerisine ilişkin  $i$  maddesinin güçlük düzeyini ifade eder (Embretson, 1985, 1997).

Bu modelin genel yapısında, öğrencinin beceriye ilişkin başarısına bağlı iki olası strateji bulunmaktadır. İlk strateji öğrencinin maddeyi gerekli becerilere sahip olarak yanıtlaması ve ikinci strateji ise öğrencinin madde için gerekli becerilerden en az bir tanesine sahip olmadan yanıtlamaya çalışması yani tahmin (guessing) etmesidir. İlk strateji öğrencinin madde için gerekli bütün becerilere sahip olma olasılığını verirken ikincisi öğrencide gerekli becerilerin bulunmaması durumunun olasılığıdır. MLTM'de maddenin cevaplanmasında gerekli becerinin etkisi öğrencinin  $\theta_{jk}$  parametresinde



doğrudan temsil edilmektedir. Sonuç olarak bu modelde beceri üzerindeki performans maddeyi doğru cevaplama performansına göre modellenmez (Dibello ve ark. 2007).

### 1.2.5. Sınırlandırılmış Örtük Sınıf Modeli (Restricted latent class model- RLCM)

Bu modelin “sınırlandırılmış” olarak adlandırılmasının temel nedeni kullanılan tanımlayıcı örtük cevap vektörlerinin sayısının ve türünün modelde yer alan Q matrisi tarafından sınırlandırılmış olmasıdır.

Bu modelin bir başka ayırt edici niteliği ise öğrencinin aranan beceri sahip olma düzeyinin daha önce bahsedilen modellerdeki gibi sürekli bir değişken niteliğinde tanımlanmamasıdır. Bunun yerine öğrenci yeteneği bir  $K - \alpha$  boyutsal vektörü olarak 0 ve 1’lerle ifade edilmektedir. Bu vektör öğrenciyi her bir beceri için sahip (1) ya da sahip değil (0) şeklinde tanımlamaktadır. Modele ilişkin MTF aşağıdaki gibidir;

$$P(X_{ij} = 1 | \alpha_j) = \pi_i \prod_{k=1}^K \alpha_{jk}^{q_{ik}} \prod_{k=1}^{K-1} (1 - \alpha_{jk}^{q_{ik}})$$

Bilişsel olarak model basitçe şunu söylemektedir; eğer bir öğrenci maddenin ölçtüğü becerilerin tamamına sahipse bu öğrencinin o maddeye doğru cevap verme olasılığı 1 (ya da çok yakın) dır. Diğer yandan eğer öğrenci maddenin ölçtüğü özelliklere sahip değilse maddeyi doğru cevaplama olasılığı çok düşük ve farklıdır. Bu nedenle RLCM de eğer bir öğrenci maddeyle ölçülen becerilerden herhangi birine ya da birkaçına sahip değilse bu öğrencinin o maddeyi doğru cevaplama olasılığının olmadığı anlamına gelir. Bu kabul modelin istatistiksel olarak kullanışlı olmasının önemli nedenlerinden biridir (Haertel, 1984, 1990).

Tarihsel olarak RLCM Macready ve Dayton (1977) tarafından ileri sürülen modelin bir uzantısı niteliği taşımaktadır. Macready ve Dayton’ın modeli benzer bir MTF’na sahip olmasına rağmen bütün maddelerin aynı özelliği ölçtüğü varsayımına dayanmaktadır bu nedenle de öğrencileri bütün becerilere sahip olanlar ve en az birine sahip olmayanlar olarak belirlenen iki kategori arasında sınıflamaktadır. Bu modellerin en yaygın kullanılanı ve ön plana çıkanı Junker ve Sijtsma tarafından geliştirilen DINA (Deterministic Input Noisy And gate) modelidir. DINA modelin yapısı ve özellikleri ilerleyen bölümlerde ayrıntılı biçimde aktarılacaktır.

### 1.2.6. Rule Space Modeli

Tatsuoka ve Tatsuoka’nın (1982) geliştirdiği Rule Space (Kural Düzlemi) yaklaşımı, öğrencinin test maddelerindeki performanslarından yola çıkarak “özellik”(attribute) olarak tanımlanan spesifik bilişsel becerilerini tanımlamayı amaçlamaktadır. Bu model istatistiksel bir yaklaşımla öğrencilerin cevaplama örüntülerini kullanarak onları özelliğe sahip ya da değil biçiminde tanımlanan sınıflama kalıplarına yerleştirmektedir. Bu istatistiksel sınıflama modeli, “Rule Space” olarak adlandırılan iki boyutlu Kartezyen sınıflama uzayı yardımıyla hesaplanır. Bu yaklaşım theta ( $\theta$ ) ve zeta ( $\zeta$ ) olarak belirlenen iki eksen aracılığıyla simgelenir (Gierl, 2007). Modelde  $\theta$ , MTF aracılığıyla hesaplanan bireyin yetenek parametresi iken  $\zeta$ , Tatsuoka (1984) tarafından atipik olarak tanımlanan cevaplama örüntüsünü ifade eder. Tam olarak  $\zeta$ , zor maddelerin doğru cevaplanıp kolay maddelerin cevaplanmadığı durumları tanımlayan indekstir.

Rule Space yaklaşımında Q matris tarafından belirlenen cevaplama örüntüsü ile öğrencinin cevaplama örüntüsü karşılaştırılır. Q matris tarafından belirlenen örüntü “ideal” modeli verirken öğrencinin cevaplama düzeneği gözlenen örüntüyü tanımlamaktadır. Bu iki örüntünün karşılaştırılması modele ismini veren “Rule” yapısını oluşturmaktadır. Her bir “Rule” için olasılık dağılımı aşağıdaki formül ile hesaplanır;

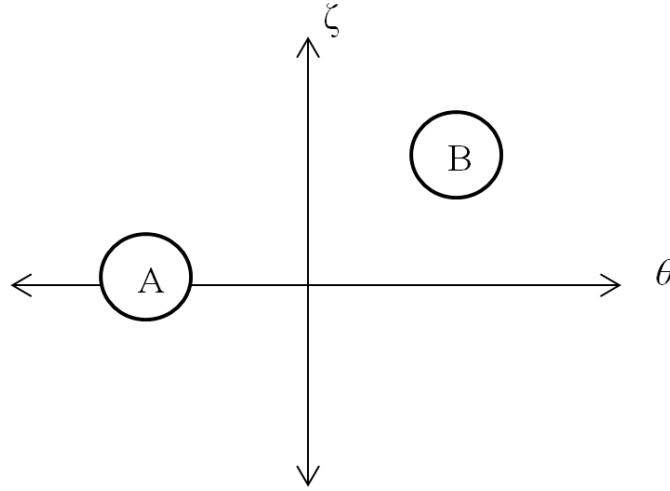
$$\sum_{s=0}^S \left[ \sum_{\sum u_i=1} \prod_{i=1}^n p_i^{u_i} (1 - p_i)^{1-u_i} \right]$$

$$\text{ortalama: } \mu_R = \sum_{i=1}^t p_i + \sum_{i=t+1}^n q_i$$

$$\text{varyans: } \sigma_R^2 = \sum_{i=1}^n p_i q_i$$

Yukarıdaki formülde kaydırma sayısı  $s$ ,  $u_i$  ise  $i$  maddesi için kaydırma gözlendiğinde 1 gözlenmediğinde ise 0 değeri almaktadır. Bununla birlikte formülde  $n$  madde sayısını  $t$  ise toplam puanı ifade etmektedir (Tatsuoka & Tatsuoka, 1987; Tatsuoka, 1995).

İdeal örüntü kullanılarak Tatsuoka tarafından ortaya atılan “Bilgi düzeyleri” (knowledge states) tanımlanır ve bu bilgi düzeylerinin Kural düzlemi olarak tanımlanan analitik düzlemdeki konumları theta ve zeta parametrelerine göre belirlenir. Aşağıda iki bilgi düzeyi için belirlenen konumları gösteren grafik yer almaktadır;

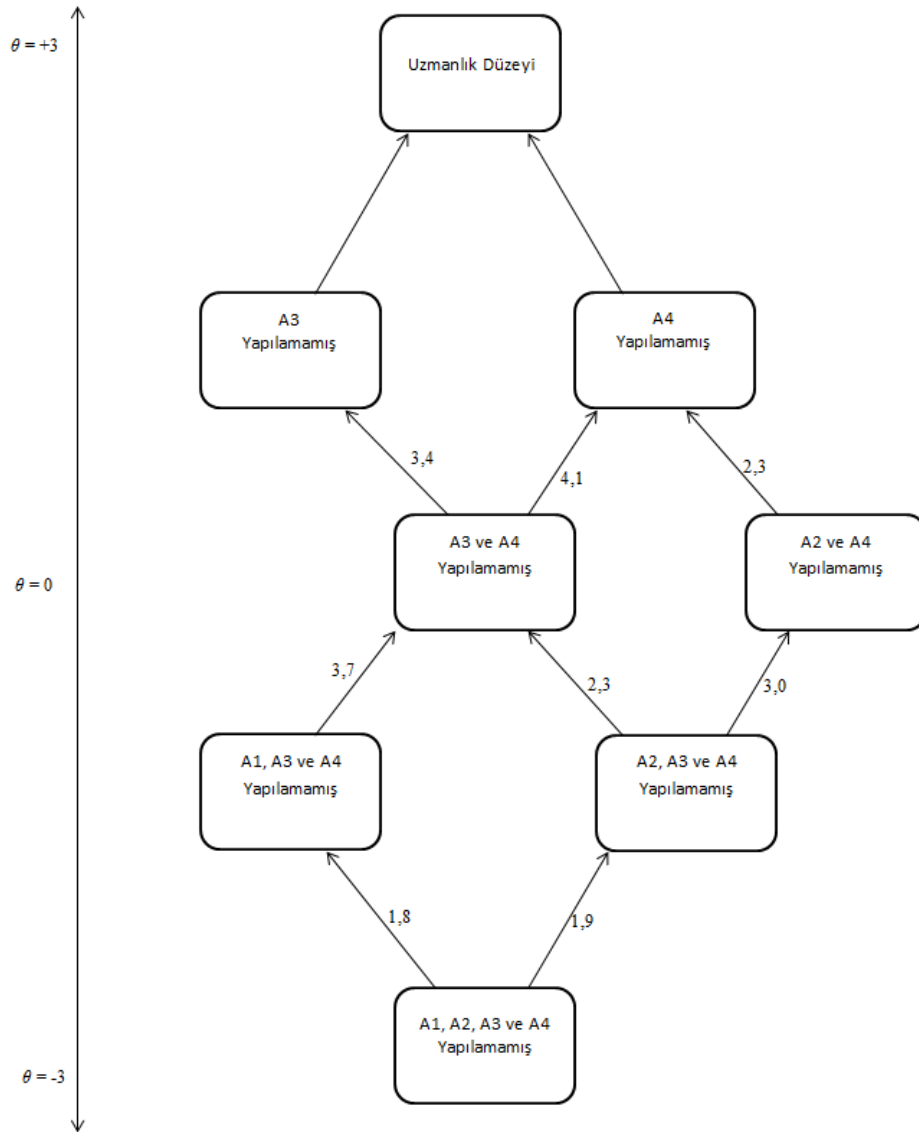


Şekil 1. İki boyutlu Kural Düzleminde yer alan iki bilgi düzeyi

Şekil 1’de görüldüğü gibi B bilgi düzeyinin A bilgi düzeyine göre daha yüksek bir yetenek seviyesinde olduğu görülmektedir. Bununla birlikte B bilgi düzeyinde çok daha

fazla atipik cevaplama görüldüğü belirlenmiştir. Bu durum gibi birçok cevaplama örüntüsü Kural düzlemi üstünde yerleştirilebilmekte ve bu modelle bu bilgi düzeyleri için  $\theta$  ve  $\zeta$  parametreleri belirlenebilmektedir.

Bilgi düzeyleri ve maddeler arasındaki ilişkiler “Rule Space” Mahalanobis uzaklıkları yardımıyla belirlenmektedir. Aşağıdaki şekilde 4 özellik (A1, A2, A3, A4) için oluşturulmuş bir teste ilişkin “bilgi düzeylerinin” nasıl belirlendiği gösterilmektedir;



Şekil 2. Sekiz Bilgi Düzeyini Gösteren Diyagram

Şekil 2’de görüldüğü gibi “bilgi düzeyleri” öğrencilerin hangi özelliklere sahip olduklarını göstermekte ve aynı zamanda herhangi bir bilgi düzeyindeki öğrencinin bir sonraki aşamada hangi özelliği öğrenmesi gerektiği hakkında da bilgi vermektedir. Örneğin A1, A2 ve A4 özelliklerine sahip olmayan öğrenci ilk aşamada A1 özelliğini öğrenmelidir. Daha sonraki aşamada ise öğrencinin A3 ve A4 özellikleri arasında tercih yapması söz konusudur, ama burada da Mahalanobis uzaklıklarına göre (3,4) A4’ü kazanması yani A3’ün bilinmediği düzeye çıkması gerekmektedir. Tatsuoka ve Tatsuoka (1997) bu yaklaşımın CAT (Computer Adaptive Tests) uygulamaları için basamakların belirlenmesi ve program öğretiminde özellikler arasındaki hiyerarşinin ortaya çıkartılması aşamalarında da kullanılabileceğini belirtmişlerdir.

Rule Space modeli bilişsel tanı yaklaşımı ve bilgisayar uygulamalı testler için çığır açıcı bir nitelik taşımaktadır. Yaklaşım bu yapıyla test toplam puanları dışında testte ölçülen özelliklerle öğrencilerin cevaplama örüntüleri arasındaki ilişkiyi açıklanabilir kılmayı hedeflemiştir. Fakat bütün bunların yanında Rule Space modeli testte ölçülen özellikler ile maddeler ve maddelerin birbirleriyle ilişkileri konusunda bir bilgi vermemektedir. Bu durum öğrencinin bilişsel profilini eksiksiz belirleme konusunda bazı sıkıntılara yol açmaktadır.

#### 1.2.7. DINA Model

Haertel (1989) tarafından geliştirilen DINA model, diğer modeller gibi ikili yetenek modellerine benzeyen bir örtük sınıf analizidir. DINA model çoğu BTM’de görüldüğü gibi madde özellik ilişkisini temel alır. Modelin iyi işleyebilmesi, bir maddenin doğru cevaplanması için gerekli olan özelliklerin doğru belirlenmesine bağlıdır. Bununla birlikte DINA modelin geliştirilen G-DINA ve HO-DINA gibi modifikasyonlarıyla bir maddeyle ilişkili birden çok özelliğin gerek ağırlandırılarak gerekse hiyerarşik olarak analize katılabilme imkânı bulunmaktadır (de la Torre, 2010; de la Torre ve Douglas, 2004).

DINA model cevaplayıcının gözlenen yeteneğinin altında yatan örtük özelliği ortaya çıkartmayı amaçlamaktadır. Bu anlamda model, örtük özellikle gözlenen özellik arasındaki ilişkiyi olasılıkla temellendirmekte ve her madde için iki madde parametresinin sınıflanmasını sağlamaktadır. Bunlar  $s$  “kaydırma” (slip) ve  $g$  “tahmin” (guess) parametreleridir.

$$s_j = P[Y_{ij} = 0 | \eta_{ij} = 1] \text{ ve}$$
$$g_j = P[Y_{ij} = 1 | \eta_{ij} = 0],$$

Yukarıdaki formülde  $s_j$  örtük özelliğe sahip bireyin  $j$  maddesine yanlış cevap verme olasılığını gösteren durumu (yanlış pozitif olasılık), ve  $g_j$  ise örtük sınıfa sahip olmayan bireyin doğru cevap verme olasılığı durumunu (doğru pozitif olasılık) ifade etmektedir.  $s_j$  parametresi ne kadar düşük olursa aranan özelliklere sahip bireylerin doğru cevap verme olasılığı o kadar artmaktadır (de la Torre, 2009).

Bununla birlikte  $g_j$  parametresine tahmin parametresi de denir. Maris (1999), alternatif olarak  $g_j$  parametresini zihinsel tahmin yeteneğine başarıyla kullanmak olarak da

açıklar. Tahmin parametresi bu anlamda MTK'daki şans parametresinden farklı bir yapıdadır. Modelde  $g$  parametresi sadece maddeyi doğru cevaplamak için gerekli olan özelliğe sahip olmayan bireyin maddeyi doğru cevaplama anlamına gelmemektedir. Aynı zamanda cevaplayıcının maddeyi doğru cevaplamak için gerekli olduğu düşünülen özellikler dışında başka özellikleri kullanarak maddeyi doğru cevaplama anlamına da gelmektedir. Bu nedenle  $Q$  matris tarafından tanımlanmayan farklı özelliklerin de bir maddenin doğru cevaplanması için yeterli olabileceğini göstermektedir. Bir madde için "1" değerine yakın çıkan  $g$  parametresi sadece o maddeyi özelliğe sahip olmayan bireylerin cevapladığı anlamına gelmez, aynı zamanda maddeyi doğru cevaplamak için gerekli olan bazı özelliklerin belirlenmediği şekilde de yorumlanabilir.

DINA modelde madde tepki fonksiyonu aşağıdaki biçimdedir;

$$P(\alpha_{ij}^*) = \begin{cases} g_j & \text{if } \alpha_{ij}^* < 1_{K_j} \\ 1 - s_j & \text{otherwise} \end{cases}$$

Fonksiyonda  $K_j^*$  uzunluğunda,  $1_{K_j}^*$  'nin bir kişinin vektörü olduğu durumda,  $g_j$ ,  $j$  maddesi için gerekli olan özelliklerden en az birine sahip olmamasına rağmen doğru tahminde bulunan bireyin olasılığını vermektedir.  $1 - s_j$  ise madde için gerekli bütün özelliklere sahip olmasına rağmen kaydırma yapmadığı halde maddeyi yanlış cevaplayan bireylerin olasılığını ifade etmektedir (de la Torre, Hong, & Deng, 2010).

DINA modelde koşullu dağılan madde cevap değişkeni  $Y_{ij}$  aynı zamanda  $\alpha_{ij}$  den  $\eta_{ij}$  ye de bağlıdır. Bu DINA modelin olasılık fonksiyonunun bir uzantısıdır, konular arasındaki bağımsızlığı gösteren koşullu bağımsızlık şu şekilde yazılabilir;

$$L(s, g; \alpha) = \prod_{i=1}^N \prod_{j=1}^I [s_j^{1-y_{ij}} (1 - s_j)^{y_{ij}}]^{\eta_{ij}} [g_j^{y_{ij}} (1 - g_j)^{1-y_{ij}}]^{1-\eta_{ij}}$$

DINA modelde  $s$  ve  $g$  parametreleri madde düzeyinde ortaya çıkmaktadır. Her madde popülasyonu iki sınıfa bölmekte ve aynı sınıfa düşen öğrencilerin o maddeye doğru cevap verme olasılıkları eşit olmaktadır. Öğrencilerin sınav performansı testte yoklanan niteliğin vektörünün tam bir göstergesi değildir. Bu nedenle olasılık temelinde bir model ancak  $s$  ve  $g$ 'nin olasılığını görmeye izin verir. "Kaydırma" durumu öğrencinin maddede aranan niteliğe sahip olmasına rağmen alt görevi doğru cevaplayamadığı ya da maddeyi doğru cevaplayamadığında gerçekleşir. "Tahmin" ise öğrencinin maddede aranan özelliklerden birine ya da birkaçına sahip olmamasına rağmen alt görevleri tamamlaması veya maddeyi doğru cevaplama durumudur. Seçilecek model belirlenirken bu nedenle  $s$  ve  $g$  parametrelerinin alt görev düzeyinde mi yoksa madde düzeyinde mi gerçekleştiğine karar verilmesi gerekmektedir.

DINA modelde öğrenciler her madde için temel iki sınıfta yer alır. Bu sınıflardan ilki yokluk sınıfı (null class) yani beklenen hiçbir beceriye sahip olmayan öğrencilerin oluşturduğu grup ve diğeri de tam sınıf (full class) yani bütün becerilere sahip olan

öğrencilerin sınıfıdır. DINA modelde tek bir beceriyi bile kaçırın biri yokluk sınıfı (null class) içinde yer alır. Bir maddenin doğru cevaplanması için gerekli bütün becerilere sahip olan bireyin maddeyi doğru cevaplama olasılığını gösteren fonksiyon aşağıdaki gibidir;

$$P[Y_{ij} = 1 | \eta_{ij}, s_j, g_j] = (1 - s_j)^{\eta_{ij}} g_j^{1-\eta_{ij}}$$

P aranan bütün becerilere sahip olan öğrencinin maddeyi doğru cevaplama olasılığıdır.  $\eta_{ij}$ ,  $\alpha$  tarafından belirlenen örtük cevaplama ve  $i$  inci konunun niteliği ve  $q_j$  nin vektörüdür. Q matrisinin  $j$  inci maddesine tekabül eden sıra şu şekilde gösterilebilir;

$$\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}}$$

Tatsuoka (1982)  $\alpha_i = (\alpha_{i1}, \dots, \alpha_{iK})$  yi “knowledge states” bilgi durumları olarak tanımlar. Burada  $\alpha_{iK} = 0$  veya  $1$  olması  $i$  öğrencisinin  $k$ ;  $\eta_i = (\eta_{i1}, \dots, \eta_{ij})$  niteliğine sahip olup olmamasına bağlıdır.  $j =$  toplam madde sayısını göstermektedir, bu maddeler  $i$  öğrencisinin istenilen niteliklere sahip olup olmadığını ve  $Y_{ij}$  gözlenen puanının belirlenmesini sağlar. Belirli bir  $k$  niteliğini için  $2k$  kadar olası bilgi modeli yani örtük sınıf vardır.

DINA model geliştirilen ya da uygulanan bir test için belirlenen  $k$  sayıdaki özelliği temel olarak  $2^k$  sayıda örtük sınıf belirlemektedir. Örneğin bir testte sadece 3 özelliğin ölçüldüğü düşünülürse bu durumda testi alan bireyler toplam 8 örtük sınıf içinde sınıflandırılır. 3 özellik için olası sınıflar “000”, “100”, “010”, “001”, “110”, “011”, “101” ve “111” şeklinde sıralanır. Hiçbir özelliğe sahip olmayanlar ilk sınıfa dahil olurken sadece birinci ve üçüncü özelliğe sahip olan bireyler 7. sınıfa yerleşirler.

DINA modelin bir cevaplayıcının bir özelliğe sahip olup olmadığını belirleme aşaması olasılık temelli bir süreçtir. Öğrencinin özellik bakımından 0 sınıfta mı yoksa 1 sınıfta mı olacağı bir olasılık değeridir. Bu değer araştırmacı tarafından değiştirilebileceği gibi genellikle kullanımda .50 eşiği temel alınır. Öğrencinin özelliğe sahip olma olasılığı .50 değerinin altında kalırsa 0 sınıfına, üstünde ya da aynı değeri alırsa 1 sınıfına dahil olur. DINA model öğrenciye ilişkin  $\alpha$  kestirimini yaparken o özelliği temsil eden maddelerin doğru cevaplanma oranını temel almaz. Öğrencinin özelliğe sahip olma olasılığı özelliği temsil eden maddelerin güçleriyle ilgilidir (de la Torre 2008c).

### 1.2.8. G-DINA Model

G-DINA model DINA modelin geliştirilmiş halidir. Bu model de birçok bilişsel tanı modeli gibi  $J \times K$  Q matrisine dayanır. G-DINA model de örtük sınıfları  $2^{K_j}$  sayıda örtük gruba ayırır. Her örtük grup  $\alpha_{ij}^*$  ile gösterilen bir beceri vektörüne indirgenir. Her örtük grubun kendisine ait  $P(\alpha_{ij}^*)$  ile ifade edilen başarı olasılığı vardır (de la Torre, 2008a).

G-DINA için  $P(\alpha_{ij}^*)$  temelli olan orijinal formül, her bir spesifik beceri ve bu becerilerin birbirleriyle etkileşimi toplam etkilerine göre parçalara ayrılabilir. Aşağıda G-DINA model için olasılık formülü verilmiştir;

$$P(\alpha_{ij}^*) = \delta_{j0} + \sum_{k=1}^{K_j^*} \delta_{jk} \alpha_{1k} + \sum_{k=k+1}^{K_j^*} \sum_{k=1}^{K_j^*-1} \delta_{jkk'} \alpha_{1k} \alpha_{1k'} \dots + \delta_{j12 \dots K_j^*} \prod_{k=1}^{K_j^*} \alpha_{1k}$$

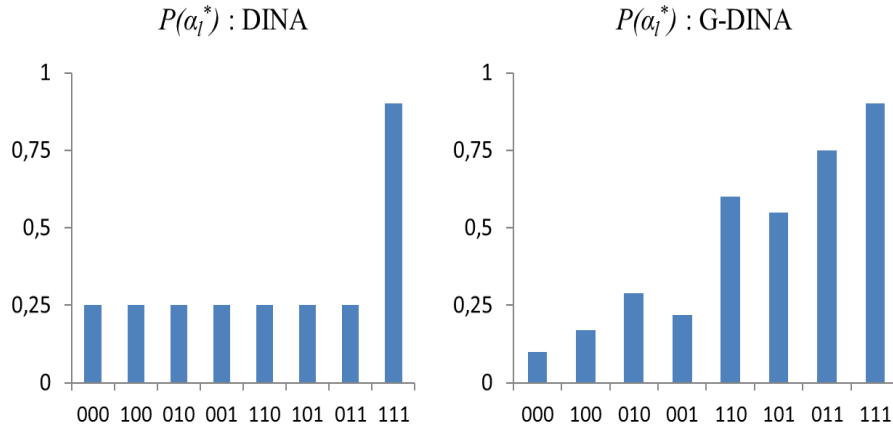
$\delta_{j0}$  = j maddesinin kesişimi

$\delta_{jk}$  =  $\alpha_k$  üstündeki temel etki

$\delta_{jkk'}$  =  $\alpha_k$  ve  $\alpha_{k'}$  nın birbirleriyle etkileşiminin etkisi

$\delta_{j12 \dots K_j^*}$  =  $\alpha_1, \dots, \alpha_{K_j^*}$  bağlı karşılıklı etkileşimin etkisi anlamına gelmektedir.

G-DINA modelin tahminleme kodları EM algoritmasının bir uygulamasıdır. Analiz sürecinde öncelikle  $P(\alpha_{ij}^*)$  değerleri standart hatalarıyla birlikte hesaplanır. Daha sonra yeteneklerin sonsal (posterior) olasılıkları belirlenir ve bu olasılıklara göre öğrencilerin örtük sınıfları ile madde ve teste ilişkin uyum istatistikleri hesaplanır.



Şekil 3. DINA ve G-DINA için Örtük Sınıf Olasılıkları

Yukarıdaki grafikte DINA model ve G-DINA model için başarı olasılığı dağılımının becerilere sahip olma düzeylerine göre nasıl olduğu görülmektedir. Üç nitelik için DINA modelde maddeyi doğru cevaplama olasılığı sadece öğrencinin üç niteliğin tamamına sahip olduğu durumda maksimum olmakta ve diğer her durum için olasılık minimum düzeyde kalmaktadır. G-DINA modelde ise her bir niteliğin doğru cevaplanma olasılığına katkısı farklıdır ve öğrencinin bu niteliklerden herhangi birine ya da bir kaçına sahip olması durumunda maddeyi doğru cevaplama olasılığı niteliğin ağırlığına göre değişmektedir. G-DINA model de her bir madde için

cevaplayıcının sahip olabileceği her bir  $P(\alpha_{ij}^*)$  durumuna ilişkin olasılık hesaplamaktadır.

### 1.2.9. NIDA Model

NIDA model “Noisy inputs, Deterministic “And” Gate” kelimelerinden gelmektedir. Bu model amaç olarak diğer örtük sınıf modellerine benzer ama bununla birlikte BTM arasında yapı olarak en karmaşığdır. Örneğin Birleştirilmiş model ve Reparametrize model NIDA modelden geliştirilmiştir. DINA modelden farklı olarak NIDA model kaydırma ve tahmin parametrelerini madde/görev düzeyi yerine beceri seviyesi üzerinden kestirir.

DINA modelin aksine NIDA modelde  $s$  ve  $g$  parametreleri alt görev düzeyinde oluşur. NIDA modelde doğru cevaplanan bir madde bütün alt görevlerin doğru yanıtlandığının göstergesidir. Bununla birlikte  $s$  ve  $g$  öğrencinin nitelik profiline göre her bir alt görevde yer alabilir. Eğer bir madde için bir nitelik gerekiyorsa o niteliğe sahip olan öğrenci kaydırma yapmaması şartıyla alt görevleri de doğru cevaplayacaktır. Bu nedenle nida modelde  $s$  ve  $g$  parametreleri her bir maddeye değil her bir niteliğe bağlıdır (de la Torre & Douglas, 2008).

NIDA model aşağıdaki gibi ifade edilebilir;

$$P[Y_{ij} = 1 \mid \alpha, s, g] = \prod_{k=1}^K \left[ (1 - s_k)^{\alpha_{ik}} g_j^{1-\alpha_{jk}} \right]^{Q_{jk}}$$

DINA modelden farklı olarak NIDA model testi alan bireyin bir beceriye ait soruyu doğru cevaplayamaması durumunda bu bireyin aynı beceriyi ölçen soruları yanlış cevaplama olasılığının doğru cevaplama olasılığından daha yüksek olduğunu kabul eder. DINA modelde örtük cevaplama değişkeni  $\eta_{ij}$  ile ifade edilir. Her bir  $k$  becerisi için “slip” kaydırma parametresi aşağıdaki gibidir;

$$s_k = P[\eta_{ijk} = 0 \mid \alpha_{ik} = 1, Q_{jk} = 1]$$

ve “ $g$ ” tahmin parametresi aşağıdaki gibi ifade edilir;

$$g_k = P[\eta_{ijk} = 0 \mid \alpha_{ik} = 0, Q_{jk} = 1]$$

NIDA modelde “eksiksiz” olarak tanımlanan fazladan bir indeks daha vardır;

$$P[\eta_{ijk} = 1 \mid \alpha_{ik} = a, Q_{jk} = 0] = 1$$

Bu parametre  $k$  becerisine sahip olan ama bu beceriye sahip oluş biçimi  $Q$  matrisi içinde tanımlanamayan durumlar için  $\alpha_{ij}$  değerini temsil eder.

DINA modelle NIDA model arasında birçok benzerlikte bulunmaktadır. İki modelde monotonluk ve koşullu bağımsızlık varsayımları altında görev performansını temel alan



rastlantısal Bilişsel modellerdir (Junker, 2001b). Bu iki modelde performans ve bir grup madde ya da görev arasında stokastik bağlayıcı bir ilişkiyi varsayan bir stratejiyi benimsemektedir. Bununla beraber madde veya görev performansı ile ilişkili olan becerilerin mutlaka madde veya görevi doğru cevaplama olasılığını maksimum düzeyde temsil etmesi gerekmektedir.

### 1.2.10. HO-DINA Model

Bilişsel tanı koymak amacıyla örtük değişkeni tam olarak belirlemede hem maddeyi doğru cevaplamak için gerekli yetenek örüntüsünün koşullu olasılığı hem de yetenek örüntüsünün ortak dağılımı gereklidir. DINA model bir koşullu olasılık tanımlamaktadır fakat bu model analizleri sonucunda belirlenen ortak dağılım çok sınıflı bir nitelik taşımaktadır. Sonraki dağılım da ise her yetenek örüntüsü tek bir kategoriye tekabül eder. Bu nedenle, EM algoritmasının tamamlanması aşamasında, olası yetenek örüntüsü sayısına bağlı olan posterior olasılıklar  $2^K-1$  tane marjinalizasyon veya kestirim basamağı üretir ve bu  $\sum_{i=1}^I p(\alpha_i|X) = 1$  örüntüyü gösterir. Bütün olası örtük sınıfların temsil edildiği bir modele genel olarak “doymuş” model adı verilmektedir ve bu model ortak dağılımın en genel formülasyonudur. Fakat eğer beceri sayısı biraz fazlaysa en çok olabilirlik kestirimleri nerdeyse tamamen yavaşlamakta ve doymuş modele ilişkin kestirimler tamamen imkânsız olmasa da uygulanması çok zor bir hâl almaktadır.

De la Torre ve Douglas (2004) bu soruna ilişkin çözüm önerileri sunmuşlardır. DINA modele ilişkin hesap hatalarını azaltmak amacıyla yeteneklerin ortak dağılımında bir modifikasyon yapmışlardır. Bu yaklaşım bir anlamda bilgi düzeyi olarak adlandırılabilir ve gözlenen  $\alpha$  bileşenini genel zekâ kavramıyla ilişkilendirmektedir. Ek olarak aslında bilişsel tanı amacı taşıyan çoğu test sadece birkaç genel yeteneği ölçmektedir. Bu nedenle de la Torre (2009)  $\alpha$ 'nın bileşenlerinin koşullu bağımsızlığı sağlandığının varsayıldığı durumlarda yeteneklerin ortak dağılımı olarak alınmasını ve  $\theta$ 'nın genel yüksek düzey (higher-order) örtük özellik olarak tanımlanmasını önermektedir. Bu amaçla önerilen  $\alpha$ 'nın  $\theta$  üzerindeki olasılık dağılımına ilişkin formül aşağıdaki gibidir;

$$P(\alpha|\theta) = \prod_{k=1}^K P(\alpha_k|\theta) = \prod_{k=1}^K \frac{\exp(\lambda_{0k} + \lambda_1 \theta)}{1 + \exp(\lambda_{0k} + \lambda_1 \theta)}$$

Eğitim uygulamalarında, yüksek düzey örtük özellik  $\theta$  genel olarak özel bir alana ilişkin beceri düzeyi ya da yetkinlik olarak tanımlanabilir. Yüksek düzey örtük özelliğin DINA model içinde tanımlanabileceği bir model olarak HO-DINA model sunulmuştur. Bu formülasyon problemin karmaşıklığından kaynaklanan hesaplama hatalarını belirgin biçimde azaltmaktadır. Doymuş modele ilişkin  $2^K-1$  tane parametre tahmininin yerine bu modelde sadece  $K+1$  adet ( $K$  kesişim ve 1 eğitim parametresi) parametre bulunmaktadır. Sonuç olarak modelde yeteneklerin ortak dağılımı altında  $K$  parametre sayısı katlanarak değil doğrusal olarak artmaktadır. HO-DINA model de her ne kadar madde parametre sayısı kolay hesaplanabilir oranda düşse de bu formülasyon için EM algoritmasıyla hesap yapmak mümkün olmamaktadır. Bu nedenle, de la Torre ve Douglas(2004) HO-DINA model için Monte Carlo Markov Zinciri kullanarak güvenilir parametre tahminleri yapmanın yöntemini sunmuşlardır.

**1.2.11. DINO model (Deterministic-Input, Noisy-Or)**

Bu model RCLM içinde yer alan DINA modele çok benzemektedir. Aralarındaki fark ise DINA modeldeki “bağlayıcı” (“and”) yapı bu modelde “ayırıcı” (“or”) olmaktadır. Bu genel olarak şu anlama gelmektedir; DINA modelde bir maddenin ölçtüğü özellikler birbirine bağlıdır. Bu modelde cevaplayıcının bir maddeyi doğru cevaplama olasılığının yüksek olması için madde için gerekli becerilerin hepsine sahip olması gerekmektedir ama DINO modelde maddenin gerektirdiği becerilerden sadece birine sahip olması bile cevaplayıcının maddeyi doğru cevaplama olasılığının yüksek olması için yeterlidir (bu modelin “veya” kısmını ifade etmektedir). Bunun dışında iki modelin madde ve yetenek parametreleri nerdeyse aynıdır, fakat MTF, DINO model için ayırıcı karakteri taşımaktadır.

$$P(X_{ij} = 1 | \alpha_j) = \pi_i \left[ \prod_{k=1}^K (1 - \alpha_{jk})^{q_{ik}} \right] r_i \prod_{k=1}^K (1 - \alpha_{jk})^{q_{ik}}$$

Formülde  $\pi_i = i$  maddesi için gerekli becerilerden en az bir tanesine sahip olan öğrencinin maddeyi doğru cevaplama olasılığı,  $r_i = i$  maddesi için gerekli hiçbir beceriye sahip olmayan öğrencinin maddeyi doğru cevaplama olasılığıdır (Templin ve Henson, 2006).

**1.2.12. Birleşik Model (Unified Model)**

Birleşik model, MTK gibi geliştirilen bir modeldir ve madde cevapları ve bu cevapların altında yatan yetenek arasındaki stokastik ilişkiyi ifade eder (DiBello, Stout and Roussos, 1995). Birleştirilmiş modelin temelinde Tatsuoka'nın Rule Space modeli ve örtük sınıf tepki modeli vardır. Birleştirilmiş model fonksiyonu aşağıdaki gibidir:

$$P(X_i = 1 | \alpha_i, \theta_j) = d_i \prod_{k=1}^K \pi_{ik}^{\alpha_{jk} \cdot q_{ik}} r_{ik}^{(1 - \alpha_{jk}) \cdot q_{ik}} P_{c_i}(\theta_j) + (1 + d_i) P_{b_i}(\theta_j)$$

Bu fonksiyonda “kaydırma” ve “tahmin” olarak adlandırılan iki parametre bulunmaktadır.

$\pi_{jk} = P(j \text{ maddesindeki } k \text{ becerisini doğru yanıtladı/ } k \text{ becerisine sahip)}$

$r_{jk} = P(j \text{ maddesindeki } k \text{ becerisini doğru yanıtladı/ } k \text{ becerisine sahip değil)}$

Bununla birlikte denklemdaki  $c_i$  ve  $b_i$  parametreleri MTK Rasch modelindeki güçlük parametrelerinin Q matrisi tarafından belirlenmiş halidir. Denklemden yer alan  $d_i$  parametresi ise diğer stratejiler arasından Q matrisine dayanan doğru stratejiyi seçme olasılığına işaret eder.

Birleştirilmiş model, Bilişsel modeller arasında Q matrisin testte yer alan bütün bilişsel gereklilikleri karşılamadığını kabul eden ilk modeldir. Bu modelde test model uyumunun artırılması için Q matrisi dışında kalan başka örtük yeteneklerinde farklı parametrelerle temsil edilip denklemin içinde yer alması gerektiği ortaya koyulmuştur. Bu model, sınıflama geçerliliğinin ancak gerçek test ve tekrar test verilerinin uyumuyla sağlandığı görüşüne dayanmaktadır. Birleştirilmiş modelin sınıflama geçerliği konusundaki en önemli eksiği ise denklemin içinde yer alan bütün parametrelerin istatistiksel olarak hesaplanamamasıdır.

### 1.2.13. Reparametrize Birleşik Model (Reparameterized Unified Model - RUM)

Orijinal Birleşik model kavramsal olarak heyecan verici olsa da madde parametrelerinin tanımlanamaz olması ve kısıtlı parametrelerle istatistiksel kestirimlere imkan vermemesinden dolayı işlevsel olmamıştır. Bu nedenle Hartz (2002) tarafından daha esnek bir model olan RUM geliştirilmiştir.

HYBRID model gibi RUM da RLCM'nin bir uzantısı sayılabilir. HYBRID model RLCM'nin bilişsel temelle öğrencileri örtük sınıflar üzerinde sınıflanmalarının tanımlanması mantığında bir değişiklik yapmaz. Bunun yerine zor sınıflanan öğrencileri tek boyutlu MTK çatısı altında toplayarak RLCM'nin bilgi yapısına ilişkin varsayımlarını yumuşatır. Diğer yandan RUM, örtük sınıflamanın bilişsel anlamına bilgi yapısı temelli bir yaklaşım geliştirmiş ve sınıflama sürecini daha detaylı hale getirecek bilişsel modellemeyi amaçlamıştır (Hartz ve Roussos. 2005).

RUM konusunda çalışan yazarlar bilişsel düzeyde basitleştirilmiş tek boyutlu MTK modelleriyle çözülemeyen bilişsel temelli yüzlerce düğümden oluşan uzman sistemleri arasındaki bağlantıyı kurmayı amaçlamışlardır. BTM'nin kapsamının dışında RUM ek olarak, MLTM ile GLTM'nin karmaşık kapsamı ile model arasında bir bağlantı kurmaya çalışarak yeni modeli istatistiksel olarak daha kolay işlenebilir ama bilişsel olarak HYBRID ve RLCM'den daha karmaşık bir yapıya çevirmiştir.

MLTM ve GLTM de olduğu gibi, RUM'de tek bir beceriyle öğrenci performansını modellemeyi amaçlar. Fakat RUM sürekli olan beceri düzeyi parametrelerini kesikli duruma çevirir. RLCM'de ki gibi RUM da cevaplayıcıların beceriye ilişkin performansları bir lojistik fonksiyonla (sürekli nitelikte beceri düzeyi) belirlemek yerine, madde parametrelerini cevaplayıcının beceri düzeyini ikili düzeyde ele alınarak belirler.

RLCM'den farklı olarak RUM (MLTM'deki gibi) Q matrisin tam olarak maddelerin gerektirdiği becerileri karşılamadığı durumları da kapsar. Q matris tarafından tanımlanmayan becerileri RUM tek boyutlu MTK yöntemleri ile sürekli örtük özellikler olarak modeller. RUM için MTF aşağıdaki gibidir;

$$P(X_{ij} = 1 | \alpha_j, \eta_j) = P\left(\prod_{k=1}^K Y_{ikj} = 1 | \alpha_j\right) P_{c_i}(\eta_j)$$

$$P_{c_i}(\eta_j) = \frac{1}{1 + \exp\{-1.7[\eta_j - (-c_i)]\}} \text{ olduğu durum için } -c_i \text{ güçlük parametresi ve } \eta_j$$

yetenek parametresidir.  $\eta_j$  parametresi cevaplayıcının Q matrisinde tanımlanmayan becerilerinin tekboyutlu fonksiyonudur,  $c_i$  ise Hartz ve DiBello'nun tanımladığı "eksiksizlik parametresi"dir ve becerilerin kapsamının Q matriste ne ölçüde temsil edildiğinin bir göstergesi olarak tanımlanabilir.

### 1.2.14. Fusion Model

Fusion Model RUM'un Hartz tarafından geliştirilmiş ve kestirimler için hiyerarşik yapıda Bayesian yöntemi kullanan bir uzantısıdır. Fusion Modele ait MTF aşağıdaki gibidir;

$$P(X_{ij} = 1 | \alpha_j, \theta_j) = \pi_i^* \prod_{k=1}^K r_{ik}^{*(1-\alpha_{jk})q_{ik}} P_{c_i}(\theta_j)$$

Formülde  $X_{ij}$ ,  $j$  cevaplayıcısının  $i$  maddesine yanıtı,  $\alpha_j$ ,  $j$  cevaplayıcısının  $Q$  matrisi tarafından belirlenmiş  $K$  becerisine ilişkin bilişsel beceri düzeyinin vektörünü,  $\theta_j =$ ,  $j$  cevaplayıcısının  $Q$  matrisi tarafından belirlenemeyen testle ilişkili beceri veya yeteneğini,  $\pi_i^* = i$ . madde için gerekli bütün becerilere sahip olan cevaplayıcının  $i$  maddesini doğru cevaplama olasılığı,  $r_i^* = i$  maddesi için  $k$  becerisine ilişkin ( $Q$  matrisi tarafından tanımlanan) ayırt edicilik parametresini,  $q_{ik} = i$  maddesi için  $k$  becerisinin gerekli olup olmamasını,  $c_i = Q$  matrisi tarafından belirlenen becerilere ek olarak maddeyi doğru cevaplama için gerekli olan  $\theta_j$  düzeyini,  $P_{c_i}(\theta_j) = Q$  matrisi tarafından tanımlanmayan becerilerin doğru uygulanma olasılığını ifade eder (Montero, D., Molfils, Wang, Yen, Julian, ve Moody, 2003).

Modelde genel olarak Hartz (2002),  $\pi_{jk}^* = \prod_{k=1}^k \pi_{jk}$  ve  $r_{jk}^* = \frac{r_{j,k}}{\pi_{j,k}}$  yi Bayesian Markov Monte Carlo Zinciri yöntemi kullanarak reparametrize etmiştir. RUM'da  $\pi_{jk}^*$ ,  $Q$  temelli şartlı madde güçlüğüyken  $r_{jk}^*$  ise  $j$  maddesi tarafından temsil edilen beceriler hakkında elde edilen bilgiyi göstermektedir. Bunun yanında bu kestirim için "Arpeggio" isimli bilgisayar programını geliştirmiştir (Hartz, 2002).

Bütün diğer IRT modelleri gibi IRT temelli beceri tanılama modelleri de gözlenen  $j$  cevaplayıcısının  $i$  maddesini cevaplama olasılığını cevaplayıcının yetenek parametreleri ve madde parametrelerini hesaplayarak tanımlamaktadır. Sembolik olarak, bu olasılık  $P(X_{ij} = x | \underline{\theta}_j, \underline{\beta}_i)$  şeklinde tanımlanır. Bu formülde  $X_{ij} = x$ ,  $j$  cevaplayıcısının  $i$  maddesine verdiği cevabı ( $x=1$ , doğru cevaplama,  $x=0$  yanlış cevaplama),  $\underline{\theta}_j$ ,  $j$  cevaplayıcısının yetenek parametrelerinin vektörünü,  $\underline{\beta}_i$  ise  $i$  maddesinin parametrelerinin vektörünü ifade etmektedir.

Daha önce geliştirilen olasılık temelli istatistiksel modeller klasik test teorisine göre araştırmacıların veriden daha detaylı bilgi elde etmesine olanak sağlamaktadır (Mislevy, 1995). Yeni modellerden biri olan Fusion model ise iki temel noktada bilgi vermektedir (DiBello ve diğerleri, 1995; Hartz ve Roussos, 2002,2005). Birincisi cevaplayıcıların bilişsel profili hakkında detaylı bir bilgi vermesi ikincisi ise test ve maddelerin tanılama kapasitesi hakkındaki parametrelerdir.

Diğer IRT modelleri gibi Fusion Model de cevaplayıcının bir maddeyi cevaplama olasılığını yetenek ve madde parametrelerini kullanarak belirlemeye çalışır. Model madde cevap olasılıkları ile test geliştirenler tarafından belirlenen özellik setleri arasında bağlar kurarak cevaplayıcıların özelliklere sahip olup olmadıklarını belirlemektedir. Özellikler ve maddeler arasındaki bilişsel ilişki 1-0 şeklinde belirlenen  $Q$  matris ile tanımlanır (Tatsuoka, 1983).

Beceri ve madde arasındaki ilişkinin doğası gereği Fusion model maddenin doğru cevaplanması için bireyin sahip olması gereken özelliklerle tanımlanan özellikler

arasında bir bağ kurar. Bu durumda model doğru cevaplamanın maddeyle ilişkili bütün özelliklere sahip olmakla ilişkili olduğunu varsayar. Bu bir madde için yüksek doğru cevaplama olasılığının o maddeyle ilişkili özelliklere yüksek düzeyde sahip olma gerekliliğidir.

Fusion model madde cevap fonksiyonu aşağıdaki gibidir:

$$P(X_{ij} = 1 | \alpha_j, \theta_j) = \pi_i^* \prod_{k=1}^K r_{ik}^{*(1-\alpha_{jk}) \times Q_{ik}} p_{ci}(\theta_i)$$

Fusion model  $\alpha_j$  ve  $\theta_j$  olmak üzere iki yetenek parametresi içerir. Bu parametrelerden  $\alpha_j$  belirlenen özellikler ile ilgili bireyin yetenek düzeyini temsil eder,  $\theta_j$  ise bireyde var olan fakat Q matris tarafından belirlenmemiş teste ölçülen özellikle ilişkili artık yeteneği gösterir. Yetenek parametrelerine ek olarak Fusion Modelde maddelere ilişkin  $r_{ik}^*$ ,  $\pi_i^*$  ve  $c_i$  parametreleri vardır.

Modelde  $\pi_i^*$  parametresi cevaplayıcının  $i$  maddesi için Q matris tarafından belirlenmiş bütün özelliklere sahip olduğu durumda maddeyi çözerken bu özellikleri doğru kullanabilme olasılığını vermektedir. Bu olasılık Q matris için madde güçlüğü ile yakından ilişkilidir ve 0 ile 1 arasında değer alır.

Modelde  $r_{ik}^*$  parametresi  $i$  maddesindeki  $k$  özelliğinin tanılama kapasitesinin bir göstergesidir. Eğer madde için  $k$  özelliği yüksek gereklilik gösteriyorsa bu durumda  $r_{ik}^*$  değeri düşük olur. Bu değerın sifıra yakın olması  $i$  maddesinin  $k$  özelliği için tanılama düzeyinde iyi tanımlandığını göstermektedir. Bu anlamda  $r_{ik}^*$  parametresi testin tanılama kapasitesinin belirlenmesi anlamında kritik bir önem taşımaktadır.

Model de  $c_i$  parametresi madde cevap fonksiyonunun Q matris tarafından belirlenmeyen özelliklerle ilişkisinin bir göstergesidir. Bu anlamda düşük  $c_i$  değeri maddenin daha çok  $\theta_j$  ile ilişkili olduğunu yani Q matris tarafından tanımlanma düzeyinde eksiklik olduğunu göstermektedir. Görüldüğü gibi  $c_i$  değeri Q matrisin tamlığı-bütünlüğü hakkında bilgi vermektedir.

$r_{ik}^*$  değerinin düşük olması (0 -0,5) maddenin yüksek ayırt ediciliğe sahip olduğuna işaret etmektedir.  $c_i$  değerinin düşük olması ise (0-1,5) maddenin Q matris tarafından eksiklik tanımlandığını göstermektedir (Louis, Roussos ve Sarah M. Hartz, 2007).

## 2. YÖNTEM

Araştırma, modellerin genel özelliklerini açıklamanın yanında bazı modellere ilişkin analiz sonuçlarının karşılaştırılmasını da içermektedir. Analiz yapılacak modeller uygulama sıklıklarına ve yazılım desteklerine göre belirlenmiştir. Araştırmalarda yaygın olarak tercih edilen ve yazılım desteği bulunan modeller arasından DINA, G-DINA, Ho-DINA ve Fusion model seçilmiş ve tahminledikleri parametreler ve öğrencilere ait yetenek kestirimleri bakımından karşılaştırılmıştır.

Analizi yapılacak veri simülasyon yöntemiyle üretilmiştir. Simülasyon datası Tablo 2’de verilen Q matrisine göre oluşturulmuştur.

**Tablo 2.**  
*Simülasyon Datası Q Matrisi*

Maddeler	Özellikler					Maddeler	Özellikler				
	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$		$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_5$
1	1	0	0	0	0	16	0	1	0	1	0
2	0	1	0	0	0	17	0	1	0	0	1
3	0	0	1	0	0	18	0	0	1	1	0
4	0	0	0	1	0	19	0	0	1	0	1
5	0	0	0	0	1	20	0	0	0	1	1
6	1	0	0	0	0	21	1	1	1	0	0
7	0	1	0	0	0	22	1	1	0	1	0
8	0	0	1	0	0	23	1	1	0	0	1
9	0	0	0	1	0	24	1	0	1	1	0
10	0	0	0	0	1	25	1	0	1	0	1
11	1	1	0	0	0	26	1	0	0	1	1
12	1	0	1	0	0	27	0	1	1	1	0
13	1	0	0	1	0	28	0	1	1	0	1
14	1	0	0	0	1	29	0	1	0	1	1
15	0	1	1	0	0	30	0	0	1	1	1

Araştırma için belirlenen Q matris incelendiğinde özellikleri tek başlarına ölçen 10 madde, iki özelliği aynı anda ölçen 10 madde ve 3 özelliği aynı anda ölçen 10 madde bulunduğu görülmektedir. Bununla birlikte her bir özellik toplam 12’şer madde ile ilişkilendirilmiştir.

Araştırmada simülasyon çalışması kapsamında Ox Edit programı kullanılarak iki data üretilmiştir. Bu datalar 30 madde ve 5 özellik için toplam 2000 kişinin cevaplama örüntüsünü içermektedir. Aynı zamanda modellerin hata ve uyum parametrelerinin değişikliklerini gözleyebilmek amacıyla üretilen dataların ilkinde DINA modele göre g ve s parametreleri “0,00” değerine, ikinci datada ise 0,27 ile 0,30 değerleri arasına sabitlenmiştir. Madde parametrelerinin sıfır olması veri ile model arasındaki uyumun mükemmel olduğuna yani belirlenen Q matrisin hatasız olduğuna işaret etmektedir. Bununla birlikte simüle edilen ikinci data için belirlenen 0,30 değerleri ise model data uyumu için kabul edilebilir sınır olarak gösterilmektedir (Li, 2008).

### 3. BULGULAR ve YORUM

#### 3.1. Modeller için madde parametrelerinin karşılaştırılması

Araştırma kapsamında ele alınan modellere ilişkin madde parametreleri karşılaştırılırken maddenin Q matris tarafından tanımlanma düzeyini gösteren parametreler incelenmiştir. Tablo 3 ve 4 DINA, G-DINA, Ho-DINA ve Fusion model analizleriyle belirlenen madde parametreleri verilmiştir.

**Tablo 3.**  
*DINA Parametrelerinin "0,00" Olduğu Model*

	DINA		G-DINA		HO-DINA		FUSION					
	g	1-s	g	1-s	g	1-s	$\pi_i^*$	$r_1^*$	$r_2^*$	$r_3^*$	$r_4^*$	$r_5^*$
1	0,00	1,00	0,01	1,00	0,00	1,00	0,00	0,00	-	-	-	-
2	0,00	1,00	0,01	1,00	0,00	1,00	0,00	-	0,00	-	-	-
3	0,00	1,00	0,01	1,00	0,00	1,00	0,00	-	-	0,00	-	-
4	0,00	1,00	0,01	1,00	0,00	1,00	0,00	-	-	-	0,00	-
5	0,00	1,00	0,01	1,00	0,00	1,00	0,00	-	-	-	-	0,00
6	0,00	1,00	0,01	1,00	0,00	1,00	0,00	0,00	-	-	-	-
7	0,00	1,00	0,01	1,00	0,00	1,00	0,00	-	0,00	-	-	-
8	0,00	1,00	0,01	1,00	0,00	1,00	0,00	-	-	0,00	-	-
9	0,00	1,00	0,01	1,00	0,00	1,00	0,00	-	-	-	0,00	-
10	0,00	1,00	0,01	1,00	0,00	1,00	0,00	-	-	-	-	0,00
11	0,00	1,00	-	1,00	0,00	1,00	0,00	0,00	0,00	-	-	-
12	0,00	1,00	-	1,00	0,00	1,00	0,00	0,00	-	0,00	-	-
13	0,00	1,00	-	1,00	0,00	1,00	0,00	0,00	-	-	0,00	-
14	0,00	1,00	-	1,00	0,00	1,00	0,00	0,00	-	-	-	0,00
15	0,00	1,00	-	1,00	0,00	1,00	0,00	-	0,00	0,00	-	-
16	0,00	1,00	-	1,00	0,00	1,00	0,00	-	0,00	-	0,00	-
17	0,00	1,00	-	1,00	0,00	1,00	0,00	-	0,00	-	-	0,00
18	0,00	1,00	-	1,00	0,00	1,00	0,00	-	-	0,00	0,00	-
19	0,00	1,00	-	1,00	0,00	1,00	0,00	-	-	0,00	-	0,00
20	0,00	1,00	-	1,00	0,00	1,00	0,00	-	-	-	0,00	0,00
21	0,00	1,00	-	1,00	0,00	1,00	0,00	0,00	0,00	0,00	-	-
22	0,00	1,00	-	1,00	0,00	1,00	0,00	0,01	0,00	-	0,00	-
23	0,00	1,00	-	1,00	0,00	1,00	0,00	0,00	0,01	-	-	0,00
24	0,00	1,00	-	1,00	0,00	1,00	0,01	0,01	-	0,01	0,01	-
25	0,00	1,00	-	1,00	0,00	1,00	0,00	0,00	-	0,00	-	0,00
26	0,00	1,00	-	1,00	0,00	1,00	0,00	0,00	-	-	0,00	0,00
27	0,00	1,00	-	1,00	0,00	1,00	0,01	-	0,00	0,00	0,00	-
28	0,00	1,00	-	1,00	0,00	1,00	0,01	-	0,01	0,00	-	0,01
29	0,00	1,00	-	1,00	0,00	1,00	0,01	-	0,00	-	0,01	0,00
30	0,00	1,00	-	1,00	0,00	1,00	0,01	-	-	0,00	0,00	0,00

Tablo 3 incelendiğinde bütün modeller için parametrelerin aynı ve model uyumunun oldukça yüksek olduğu görülmektedir. Tabloda G-DINA model için g parametresi 8. maddeden sonra hesaplanmamıştır. Bunun nedeni araştırmanın G-DINA modele ilişkin bölümde de belirtildiği gibi modelin her madde için tek bir parametre hesaplamak yerine tek bir madde için her örtük sınıfa ait bir g parametresi hesaplamasıdır. Bu nedenle tek bir özelliği ölçen maddelerde maddeyle ilişkili özelliğe sahip olmayanların maddeyi cevaplama olasılıkları diğer modellerde yer alan g parametresiyle aynı anlamı taşımaktadır. Benzer durum Fusion modelde de gözlenmektedir. Fusion model  $r_i^*$  parametresi diğer modellerin g parametresine benzemektedir, fakat bu parametre madde düzeyinde değil özellik düzeyinde gözlenmektedir. Diğer yandan s parametresinin yorumlanmasında G-DINA model ve Fusion model de farklılaşma olmamaktadır. Bu parametre madde için gerekli bütün özelliklere sahip olan bireylerin

maddeyi doğru cevaplama olasılıklarını verdiği için G- DINA model de madde için gerekli bütün özelliklere sahip bireyler için hesaplanan parametre DINA ve Ho-DINA için hesaplanan parametreyle aynı anlama gelmektedir. Fusion model de ise  $\pi_i^*$  parametresi madde için gerekli bütün özelliklere sahip olan bireyin bu özellikleri kullanma olasılığını verdiği için bu değer DINA model için belirlenen (1-s) parametresiyle aynı anlamı taşımaktadır.

Tablo 4'te  $g$  ve  $s$  değerlerinin 0,30'a sabitlendiği data kullanılarak yapılan analizler sonucunda modellere ilişkin parametrelerin karşılaştırılması verilmiştir.

**Tablo 4.**  
*DINA Parametrelerinin "0,30" olduğu Model*

	DINA		G-DINA		HO-DINA		FUSION					
	$g$	1-s	$g$	1-s	$g$	1-s	$\pi_i^*$	$r_1^*$	$r_2^*$	$r_3^*$	$r_4^*$	$r_5^*$
1	0,34	0,72	0,34	0,71	0,35	0,72	0,65	0,47	-	-	-	-
2	0,28	0,70	0,27	0,70	0,28	0,70	0,67	-	0,38	-	-	-
3	0,30	0,68	0,29	0,69	0,30	0,68	0,63	-	-	0,42	-	-
4	0,32	0,67	0,31	0,66	0,32	0,68	0,68	-	-	-	0,53	-
5	0,28	0,68	0,28	0,69	0,28	0,68	0,72	-	-	-	-	0,46
6	0,33	0,70	0,34	0,69	0,34	0,70	0,65	0,41	-	-	-	-
7	0,28	0,71	0,27	0,70	0,28	0,70	0,67	-	0,38	-	-	-
8	0,31	0,69	0,31	0,69	0,31	0,69	0,65	-	-	0,41	-	-
9	0,30	0,68	0,29	0,67	0,31	0,68	0,69	-	-	-	0,51	-
10	0,27	0,70	0,28	0,69	0,27	0,69	0,73	-	-	-	-	0,46
11	0,32	0,71	-	0,72	0,32	0,72	0,59	0,63	0,58	-	-	-
12	0,30	0,71	-	0,72	0,30	0,70	0,56	0,54	-	0,65	-	-
13	0,31	0,75	-	0,73	0,31	0,74	0,67	0,49	-	-	0,56	-
14	0,31	0,72	-	0,72	0,31	0,73	0,63	0,52	-	-	-	0,68
15	0,30	0,68	-	0,66	0,30	0,67	0,58	-	0,56	0,56	-	-
16	0,29	0,71	-	0,70	0,29	0,71	0,65	-	0,58	-	0,55	-
17	0,29	0,67	-	0,67	0,29	0,67	0,63	-	0,59	-	-	0,57
18	0,27	0,72	-	0,71	0,28	0,72	0,66	-	-	0,47	0,55	-
19	0,28	0,70	-	0,71	0,28	0,70	0,64	-	-	0,57	-	0,56
20	0,30	0,69	-	0,68	0,30	0,69	0,68	-	-	-	0,59	0,61
21	0,29	0,68	-	0,69	0,29	0,69	0,52	0,75	0,60	0,66	-	-
22	0,30	0,76	-	0,74	0,30	0,76	0,58	0,69	0,70	-	0,65	-
23	0,29	0,67	-	0,66	0,30	0,68	0,49	0,62	0,80	-	-	0,85
24	0,30	0,71	-	0,71	0,30	0,71	0,51	0,71	-	0,76	0,75	-
25	0,30	0,69	-	0,66	0,30	0,71	0,55	0,59	-	0,73	-	0,76
26	0,29	0,70	-	0,68	0,29	0,70	0,61	0,69	-	-	0,64	0,68
27	0,29	0,66	-	0,64	0,28	0,64	0,50	-	0,77	0,70	0,71	-
28	0,30	0,73	-	0,75	0,29	0,73	0,56	-	0,68	0,77	-	0,66
29	0,27	0,69	-	0,65	0,28	0,69	0,58	-	0,70	-	0,68	0,65
30	0,29	0,65	-	0,63	0,29	0,65	0,56	-	-	0,80	0,56	0,79

Tablo 4'te görüldüğü gibi DINA, G-DINA ve HO-DINA madde parametreleri birbirine yakın değerler alırken Fusion model tarafından belirlenen madde parametreleri daha



düşük değerler almaktadır. Bununla birlikte her iki örneklem için Fusion modelin tahminlediği ci parametresi analiz sonuçlarında 10.00 değerini almıştır. Bu değer Q matrisin tanımlanması konusunda herhangi bir eksikliğin olmadığını göstermektedir. Fusion model madde parametreleri model veri uyumu konusunda daha duyarlı değerler ürettiğinden ve daha çok hata kaynağını parametreleri kestirirken kullandığından modelde gözlenen bozulmadan daha çok etkilenmektedir.

### 3.2. Modellerin Yetenek kestirimlerine ilişkin analiz sonuçları

BTM modellerinde öğrenciler özelliklere sahip olma durumlarına göre bir yetenek sınıfına atanırlar. Tablo 5 araştırma kapsamında incelenen modellerin öğrencileri örtük sınıfa atama olasılıklarını vermektedir.

**Tablo 5.**  
*Örtük Sınıflar ve Posterior Olasılıkları*

Örtük Sınıflar	Parametrelerin "0,00" Olduğu Model				Parametrelerin "0,30" Olduğu Model			
	DIN A	G-DINA	HO-DINA	FUSI ON	DIN A	G-DINA	HO-DINA	FUSI ON
"00000"	,033	,033	,033	,032	,029	,032	,030	,031
"10000"	,027	,027	,027	,027	,026	,026	,026	,033
"01000"	,031	,031	,031	,031	,033	,035	,033	,033
"00100"	,028	,028	,028	,028	,026	,031	,026	,039
"00010"	,034	,034	,034	,033	,034	,026	,030	,019
"00001"	,038	,038	,038	,038	,032	,021	,030	,014
"11000"	,033	,033	,033	,031	,025	,026	,026	,037
"10100"	,029	,029	,029	,029	,020	,014	,020	,051
"10010"	,025	,025	,025	,025	,023	,025	,025	,022
"10001"	,030	,030	,030	,030	,025	,022	,025	,019
"01100"	,031	,031	,031	,031	,040	,037	,038	,046
"01010"	,028	,028	,028	,028	,040	,045	,042	,022
"01001"	,029	,029	,029	,029	,027	,019	,027	,017
"00110"	,035	,035	,035	,036	,058	,059	,056	,028
"00101"	,032	,032	,032	,031	,050	,036	,047	,021
"00011"	,032	,032	,032	,032	,033	,036	,034	,010
"11100"	,030	,030	,030	,029	,022	,020	,022	,064
"11010"	,034	,034	,034	,034	,028	,032	,029	,030
"11001"	,034	,034	,034	,034	,028	,027	,028	,025
"10110"	,030	,030	,030	,030	,019	,018	,021	,040
"10101"	,033	,033	,033	,032	,023	,026	,022	,032
"10011"	,028	,028	,028	,028	,023	,026	,024	,014
"01110"	,028	,028	,028	,028	,030	,035	,031	,036
"01101"	,028	,028	,028	,027	,022	,022	,021	,026
"01011"	,038	,038	,038	,037	,044	,050	,039	,017
"00111"	,033	,033	,033	,033	,033	,033	,033	,017
"11110"	,029	,029	,029	,028	,026	,029	,026	,058
"11101"	,031	,031	,031	,030	,040	,037	,041	,050
"11011"	,030	,030	,030	,029	,030	,033	,032	,025
"10111"	,041	,041	,041	,043	,040	,042	,040	,032
"01111"	,029	,029	,029	,029	,036	,039	,037	,030
"11111"	,037	,037	,037	,037	,038	,041	,037	,060

Tablo 5 incelendiğinde parametrelerin “0,0”a eşitlendiği örneklem için modellerin örtük sınıflar için belirlediği olasılıkların birbirleriyle aynı oldukları görülmektedir. Bununla birlikte madde parametrelerin “0,30” olduğu örneklem için modellerin olasılık kestirimlerinin değiştiği gözlenmektedir. Özellikle Fusion model ve diğer modellerin örtük sınıflar için kestirdikleri olasılıklarda farklılaşmalar bulunmaktadır. Tablo 6, örneklerde özelliklere sahip olan bireylerin oranlarının modeller tarafından kestirimlerini vermektedir.

**Tablo 6.**  
*Modellerin Özellik Dağılım Kestirimleri*

$\alpha$	Parametrelerin "0,00" Olduğu Model				Parametrelerin "0,30" Olduğu Model			
	DINA	G-DINA	HO-DINA	FUSION	DINA	G-DINA	HO-DINA	FUSION
1	0,4970	0,4970	0,4970	0,4970	0,4341	0,4452	0,4055	0,5955
2	0,4950	0,4950	0,4950	0,4950	0,5078	0,5284	0,5115	0,5905
3	0,5005	0,5005	0,5005	0,5005	0,5212	0,5192	0,5295	0,664
4	0,5075	0,5075	0,5075	0,5075	0,5329	0,5688	0,5255	0,4485
5	0,5175	0,5175	0,5175	0,5175	0,5221	0,5105	0,5280	0,3955

Tablo 6 incelendiğinde model parametrelerinin “0,00”a eşitlendiği örneklemde özelliklere sahip olma oranlarını bütün modellerin tam olarak aynı kestirdikleri gözlenmektedir. Madde parametrelerinin 0,30 değerine sahiplendiği örneklem incelendiğinde ise özellikle Fusion model kestirimlerinin diğer modellere göre farklılaştığı görülmektedir. Örneğin DINA, G-DINA ve HO-DINA analizleri, model için tanımlanan 5. özellik için grubun %51-52 tarafından sahip olduğunu belirlerken aynı oran Fusion Model için %39 değerini almaktadır. Bu durum Fusion modellerin diğer modellere göre daha az kişinin 5. özelliğe sahip olduğunu belirlediği anlamına gelmektedir. Bununla birlikte Fusion model 1. 2. ve 3. özelliklere sahip olma oranlarını diğer modellerden daha yüksek kestirmiştir. Bu durum madde parametrelerinde modele ilişkin hata değerlerinin artmasının, farklı modellerde farklı yetenek kestirimi ve öğrenci profilleri hakkında benzer olmayan sonuçlara ulaşılabileceğini göstermektedir.

#### 4. SONUÇ

Araştırma genel olarak BTM'nin yapısını, gelişimini ve temel özelliklerini ortaya koymayı amaçlamaktadır. Modellerin genel yapısına bakıldığında her tür özelliği ölçmek için geliştirilen testlerde kullanılabilmesi görülmektedir. Bununla birlikte özellikle çok boyutlu testlerde kullanılabilmesi klasik yaklaşımların sınırlılıklarına sahip olmadığını göstermektedir. Çok boyutlu IRT gibi klasik test teorisinin sınırlılıklarını taşımayan yaklaşımlardan üstün yanı sıra ise her bir maddenin birden çok boyutla ilişkilendirilmesine izin vermesidir. Bu anlamda BTM herhangi bir madde ve ölçülen özellikler arasında hiçbir teorik ön koşul ilişkisi bulunmamaktadır.

BTM yaklaşımının birçok avantajlarının yanında belli sınırlılıkları olduğu da araştırma kapsamında görülmektedir. Her şeyden önce analizin yapısal modelini oluşturan Q matrisin eksiksizliğinin istatistiksel olarak belirlenememesi en önemli sınırlılık olarak

göze çarpmaktadır. Bununla birlikte modellerin öğrenci yetenek puanlarını ikili aralıkta belirlemesi bir başka sınırlık olarak görülebilir. Öğrencilerin özelliğe sahip olma durumlarının sürekli bir değişken olarak tanımlanmaması bazı durumlarda karar aşamasında yanlış sonuçlar doğurduğu görülmektedir. Örneğin öğrencinin herhangi bir özelliğe sahip olma olasılığı 0,5 olarak belirlendiğinde –bu olasılık değeri bütün modellerde istenilen değerlerle tanımlanabilir- o özelliğe sahip olma olasılığı 0,51 olan öğrenci içinde 0,99 olan öğrenci içinde özelliğe sahip (1) kararı verilmektedir. Fakat araştırma kapsamında incelendiği gibi model veri uyumun sağlandığı durumlarda öğrenci hakkında verilen kararların bütün modeller için benzer olduğu gözlenmektedir.

BTM yaklaşımının psikometri alanda giderek yaygınlaşması ve modeller hakkındaki çalışmaların artması şu an sınırlılık olarak görülen birçok sorunun zaman içinde çözüleceğine işaret etmektedir. Özellikle son yıllarda modellere ilişkin bilgisayar yazılımı desteğinin artması ve araştırmacıların uygulamaya dönük çalışmalara eğilmesi hangi modellerin daha pratik ve kararlı oldukları yönündeki literatür bilgisini arttırmaktadır.

#### KAYNAKLAR

- Adams, R. J., Wilson, M. (1996). Formulating the Rasch model as a mixed coefficients multinomial logit. In: Engelhard, G., Wilson, M. (Ed), *Objective Measurement III: Theory in Practice*. Ablex, Norwood, NJ.
- Adams, R. J., Wilson, M. & Wang, W.C. (1997). The multidimensional random coefficients multinomial logit model. *Applied Psychological Measurement* 21, 1–23.
- Almond, R. G., Steinberg, L. S. & Mislevy, R. J. (2003). *A framework for reusing assessment components*. In: Yanai, H., Okada, A., Shigemasu, K., Kano, Y., Meulman, J.J. (Ed.), *New Developments in Psychometrics* (s. 281–288). Tokyo: Springer.
- Cheng Y. & Chang H. (2007). *The modified maximum global discrimination index method for cognitive diagnostic computerized adaptive testing*. Presented at the CAT and Cognitive Structure Paper Session, June 7.
- de la Torre, J. (2009). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics*, 34, 115–130.
- de la Torre, J. & Douglas, J. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika*, 69 (3), 333-353.
- DiBello, L. V., Stout, W. F., & Roussos, L. A. (1995). *Unified cognitive/psychometric diagnostic assessment likelihood-based classification techniques*. In P.D. Nichols, S. F. Chipman, and R. L. Brennan (Ed.), *Cognitively Diagnostic Assessment* (s. 327-361). Hillsdale, NJ: Lawrence Erlbaum Associates
- Dibello, L. V. RoussosL. A. & Stout, W. (2007). *Review of cognitively diagnostic assessment and a summary of psychometric models*. Rao, C. Sinharay, S. (Ed.) *Handbook of Statistics, Psychometrics*, 26, Amsterdam: North-Holland.
- Embretson, S. E. (1985). *Multicomponent latent trait models for test design*. In: Embretson, S.E. (Ed.), *Test Design: Developments in Psychology and Psychometrics*(s. 195–218). New York: Academic Press.
- Embretson, S. E. (1997). *Multicomponent response models*. In: van der Linden, W.J.,

- Hambleton, R.L. (Ed.), Handbook of Modern Item Response Theory (s.305–321). New York: Springer.
- Fischer, G. H. (1976). 'Some probabilistic models for measuring change', in Advances in Psychological and Educational Measurement (s.107-112). D. De Gruijter & L. van der Kamp (Ed), Bern: Huber,.
- Fischer, G. H. (1973). The linear logistic model as an instrument in educational research. *Acta Psychologica* 37, 359–374.
- Fischer, G. H. (1983). Logistic latent trait models with linear constraints. *Psychometrika*, 48, 3–26.
- Gierl, M. J. (2007). Making diagnostic inferences about cognitive attributes using Rule-Space model and attribute hierarchy method. *Journal of Educational Measurement*, 44, 325–340
- Gitomer, D. H. & Yamamoto, K. (1991). Performance modeling that integrates latent trait and class theory. *Journal of Educational Measurement* 28, 173–189.
- Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement*, 26, 333-352.
- Haertel, E.H. (1984). An application of latent class models to assessment data. *Applied Psychological Measurement*, 8, 333–346.
- Haertel, E. H. (1990). Continuous and discrete latent structure models of item response data. *Psychometrika* 55, 477–494.
- Hartz, S. (2002). *Skills diagnosis: Theory and practice. user manual for Arpeggio software*. Princeton, NJ: ETS.
- Hartz, S.M., Roussos, L.A. (2005). *The Fusion Model for skills diagnosis: Blending theory with practice*. ETS Research Report, Educational Testing Service, Princeton, NJ.
- Henson, R. A., Roussos, L., Templin, J. L. (2004). *Cognitive diagnostic “fit” indices*. Yayınlanmamış ETS Proje Raporu, Princeton, NJ.
- Leighton, J. P. & Gierl M. J. (2007). *Why Cognitive Diagnostic Assessment?* Leighton, J. P. Gierl M. J. (Ed). Cognitive Diagnostic Assessment for Education. New York: Cambridge University Press.
- Li, F. (2008). *A modified higher-order DINA model for detecting differential item functioning and differential attribute functioning*. Yayınlanmamış Doktora Tezi, The University of Georgia,
- Louis A. Roussos, L. V. D., William Stout, & Sarah M. Hartz, R. A. H., Jonathan L. Templin. (2007). *The Fusion model skills diagnosis system*. In J. Gierl (Ed.), Cognitive Diagnostic Assessment for Education Theory and Applications ( 275-318). New York: Cambridge University Press
- Macready, G. B., Dayton, C. M. (1977). The use of probabilistic models in the assessment of mastery. *Journal of Educational Statistics* 2, 99–120.
- Maris, E. (1999). Estimating multiple classification latent class models. *Psychometrika*, 64, 187-212.
- Montero, D., Molfils, L., Wang, J., Yen, W., Julian, M. & Moody, M. (2003). *Investigation of the application of cognitive diagnostic testing to an end-of-course- high school examination*, Presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, IL, April 24.
- Nichols, P. D., Chipman, S. F., & Brennan, R. L. (Ed.). (1995). *Cognitively diagnostic assessment*. Hillsdale, New Jersey: Lawrence Erlbaum.

- Reckase, M. D., McKinley, R.L. (1991). The discriminating power of items that measure more than one dimension. *Applied Psychological Measurement*, 15, 361–373.
- Rupp, A. A. (2007). The answer is in the question: A guide for describing and investigating the conceptual foundations and statistical properties of cognitive psychometric models. *International Journal of Testing*, 7 (2), 95-125
- Rupp, A. A., & Mislevy, R. J. (2007). *Cognitive psychology as it applies to diagnostic assessment*. J. Leighton (Ed.), *Cognitive diagnostic assessment in education: Theory and practice*. Cambridge, UK: Cambridge University Press.
- Spada, H. (1977). 'Logistic models of learning and thought', in *structural models of thinking and learning*. H. Spada & W. F. Kempf (eds), Vienna: Huber, , 227-62
- Sympson, J. B. (1977). *A model for testing with multidimensional items*. In: Weiss, D.J. (Ed.), *Proceedings of the 1977 Computerized Adaptive Testing Conference*. University of Minnesota, Department of Psychology, Psychometric Methods Program, Minneapolis, 82–88.
- Tatsuoka, K. K. & Tatsuoka, M. M. (1984). Bug distribution and pattern classification. *Psychometrika* 52(2), 193-206.
- Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*, 20, 345–354.
- Tatsuoka, K. K. (1984). Caution indices based on item response theory. *Psychometrika*, 49(1), 95-110.
- Tatsuoka, K. K. (1990). *Toward an integration of item-response theory and cognitive error diagnoses*. In: Frederiksen, N., Glaser, R.L., Lesgold, A.M., Shafto, M.G. (Ed.), *Diagnostic Monitoring of Skill and Knowledge Acquisition*. Erlbaum, Hillsdale, NJ.
- Tatsuoka, K.K. (1995). *Cognitive assessment, An introduction to the Rule Space Method*, Routledge, New York: Taylor & Francis Group.
- Tatsuoka, K. K. (1995). *Architecture of knowledge structures and cognitive diagnosis: A statistical pattern recognition and classification approach*. In P.D. Nichols, S. F. Chipman, and R. L. Brennan (Ed.), *Cognitively Diagnostic Assessment* (p. 327-361). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Templin, J. L. & Henson, R.A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods*, 11 (3), 287-305
- von Davier, M. (2005). *A general diagnostic model applied to language testing data*. ETS Araştırma Raporu: RR-05-16, Educational Testing Service, Princeton, NJ.
- von Davier, M., DiBello, L., Yamamoto, K. (2006). *A shortlist of models for cognitive diagnosis*. Klieme, E., Leutner, D. (Eds.), *Assessment of Competencies in Educational Contexts*. Bern: Hogrefe & Huber Publishers,.
- von Davier, M., Yamamoto, K. (2004). *A class of models for cognitive diagnosis*. ETS Spearman Konferansında bildiri olarak sunulmuştur, The Inn at Penn, Philadelphia, PA, Invitational Conference.
- Whitley, S. E. (1980). Multicomponent latent trait models for ability tests. *Psychometrika*, 45, 479–494.

## EXTENDED ABSTRACT

### Introduction

In this study, it is aimed to give information about the Cognitive Diagnostic Models (CDM) approaches' cognitive profiles, basic qualities, structures that intended to determine the students' and also the practices of used models. For this aim, the general qualities of CDM models, the structure of Q matrix that is used for establishing the relationship of item qualities and the sources of error. Cognitive Diagnosis Models are given different names in literature such as cognitive psychometric models (Rupp, 2007); cognitive diagnosis models (Nichols, Chipman, and Brennan, 1995), restricted latent class models (Haertel, 1989), multiple classification models or structured item response theory models (Rupp and Mislevy, 2007). The aim of Cognitive Diagnosis Models is to classify the responders according to latent categories based on the placement of two category qualities. In these models, latent models as the determiners of expertise identify the quality sets that lie behind the diagnosis of students. Here, quality can be described as qualification, duty, sub-duty, cognitive process or skill (Tatsuoka, 1995a, p. 330).

Latent Traits Theory and Cognitive Diagnosis Model (BTM) is closely related to Item Response Theory (MTK). Many parameter estimations in BTM are developed by using the functions of MTK. However, BTM models indicate different qualities from MTK due to their multi-dimensional structure. MTK models are so successful tools in predicting the latent qualities of an individual based on the responses they give to the items. MTK models calculate the "a" distinctiveness, "b" difficulty, "c" prediction parameters while predicting the item and skill parameters on a continuous scale but they do not give any idea about how item difficulty occurs. In fact education apart from MTK models, aims at the structure of learning field, the psychological structure of a student and the psychology that lie behind the ability to solve the test that is evaluated in the item (Almond, Steinberg, Mislevy, 1993). BTM focuses on this depict and by taking the cognitive processes that a student use in solving the item into consideration BTM use the response pattern of a student to reveal out the multidimensional profiles of students.

BTM does not focus on the skill levels of students in latent scale; it determines the performance of a student and the probability to have this factor on each cognitive factors. The resulting probabilities are changed into profiles of students' expertise skills. Like structural equation models in BTM the relation of latent variables and items should be determined before analysis. The process in the models are identified as item quality matrix named as Q matrix. In the historical development of CDM, there have been various models. In the basis of these models, one-dimensional logistic models are found. Basic information about One Parameter Logistic Model (OPLM), 2PL&3PL and LLTM, Compensatory multidimensional IRT model- MIRT-C, No compensatory MIRT model- MIRT-NC, Multicomponent latent trait model- MLTM, Restricted latent class model- RLCM, Rule Space Model, DINA Model, G-DINA Model, NIDA Model, HO-DINA Model, DINO model, Unified Model, Reparameterized Unified Model – RUM, Fusion Model that mention item and skill is given. In this part, the development stages of the models, statistical structures and usage areas are identified.

It is aimed to compare the widespread used and supported via software models among the explained models. For this aim, DINA, G-DINA, Ho-DINA and Fusion models are chosen and compared according to predicted parameters and skill predictions that belong to the students.

### **Method**

Q matrix is prepared for the comparison of chosen BTM models. There have been 10 items that measure only one quality in determined qualities of Q matrix, 10 items that measure two qualities at the same time and 10 items that measure 2 qualities at the same time. Each quality is related to 12 items total. Simulation data are produced by using Ox-Edit program. These data compose response patters of 2000 people for 30 items and 5 qualities. At the same time, in order to observe the changes of model error and compliance parameter, in the first data according to Dina model, g and s parameters are fixed to “0,00” value, in the second data they are fixed to 0,27 and 0,30 values. For analyzing the DINA, G-DİNA and Ho- DINA models, the codes that were prepared for each models in Ox-edit program are used. For Fusion model, Arpeggio Suit program is used.

### **Findings**

#### *Comparison of item parameters for models*

As the g and s parameters are fixed to “0,00” value in the analysis, it is seen that parameters are the same for whole models and model fit is rather high. As the g and s parameters are fixed to “0,30” value in the analysis, DINA, G-DINA and HO-DINA item parameters’ values are nearly the same but the item parameters that are determined by Fusion model take lower values. Also, for both sample, c parameter that is predicted by Fusion model, took 10.00 values after the analysis. This value indicates no deficiency in the identification of Q matrix. As Fusion model produces more reactive values about item parameters and model data fix and uses error sources for prediction the parameters observed is more affected.

#### *The Analysis Results Of Models’ Skill Predictions*

In the analysis, for the samples that the parameters are equaled to “0,0”, the models’ probabilities are seen as equal. But the probable predictions of models for the samples whose parameters are fixed to “0,30” change. Especially, difference is found in the probabilities for latent class of Fusion model and the other models. In the analysis, it is observed that the whole models predicted the rates of having qualities in the sample that was fixed to “0,00” value as the same. Fusion model predictions differ from the other models when the sample with “0,30” value is examined. For example, for 5th quality DINA, G-DINA and HO-DINA analysis determine that group has 51-52% and Fusion model determines 39% value for the group. This means that Fusion model determines less people having 5th quality. But the Fusion model predicted that 1<sup>st</sup>, 2nd and 3rd qualities’ rate is higher. This shows that different results can be gained about the increase of error values, different skills prediction in different models and student profiles.

### **Results**

In this study it is seen that BTM has many advantages and limitations. The most striking limitation is statistically non detection of Q matrix’s completeness. The other

limitation is the model's determination of skill scores in dual range. Not determining the students' having qualities as continuous variable emerges out false results during the decision process. For example, when the probability for a student for having a quality is determined as 0,5, it means that the student who has the probability for having the quality is 0,51 or 0,99 has got that quality(1). As examined in this study's context, in the model data fit conditions the decisions for students are the same. BTM approach's increasingly getting widespread in the field of psychometrics and the studies' increase about models indicate that whole trouble due to limitations will be solved in time. Especially in recent years increase in computer software programs and practical studies indicate which model is practical and decisive.