

Demir, E. K., Koç, N. (2018). DINA model ile geliştirilen bir testin psikometrik özelliklerinin belirlenmesi. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 18 (1), 130-156.

Geliş Tarihi:26/10/2017

Kabul Tarihi: 24/01/2018

DINA MODEL İLE GELİŞTİRİLEN BİR TESTİN PSİKOMETRİK ÖZELLİKLERİNİN BELİRLENMESİ *

Elif Kübra DEMİR**

Nizamettin KOÇ***

ÖZET

Bu araştırmada Bilişsel Tanı Modellerinden biri olan DINA model ile test geliştirme süreci gerçekleştirilmiş ve teste ait psikometrik özellikler belirlenmiştir. Geliştirilen test, 7. sınıf Fen ve Teknoloji dersine ait Potansiyel ve Kinetik Enerji konu alanını kapsamaktadır. Bu amaçla belirlenen beş uzman konu alanına ilişkin dört temel özellik belirlenmiştir. Maddeleri hazırlayan uzman grubundan bağımsız alanında doktora düzeyinde eğitim almış beş uzmandan maddeler ile özellikleri ilişkilendirmeleri istenmiş ve hazırlanan Q-matrisler karşılaştırılarak hemfikir olunan 65 madde testin deneme uygulamasına konmuştur. Ölçme aracının deneme uygulaması 504 kişilik öğrenci grubuna uygulanmıştır. Deneme uygulaması için maddelere ait g parametrelerinin aritmetik ortalaması 0.42, s parametrelerinin ortalaması 0.30 ve δ parametrelerine ait değerlerinin ortalaması ise 0.28 olarak hesaplanmıştır. DINA modelde elde edilen g, s ve δ parametrelerine göre 25 maddelik testler oluşturulmuş ve mevcut veri üzerinden analiz edilmiştir. Testlerin madde parametreleri ve model veri uyumları karşılaştırıldığında literatürün desteklediği gibi nihai testin δ parametresi dikkate alınarak oluşturulmasına karar verilmiştir. Nihai test 270 kişilik öğrenci grubuna uygulanmıştır. Nihai teste ilişkin g, s ve δ parametrelerinin ortalaması ise sırasıyla 0.39, 0.26 ve 0.34 olarak hesaplanmıştır.

Anahtar Kelimeler: Bilişsel Tanı Modelleri, BTM, DINA, Q-matris, Test geliştirme

DETERMINING PSYCHOMETRIC PROPERTIES OF A TEST DEVELOPED BY DINA MODEL

ABSTRACT

In this research, test developing process was realized by means of DINA model as one of the Cognitive Diagnostic Models, and psychometric properties of the test were determined. The test developed includes Potential and Kinetic Energy subject area of 7th Grade Science and Technology Course. Items of 15 latent classes constituted as a result of these attributes were written by an expert group consisting of 5 teachers of Science and Technology experienced in their field. Five experts with doctor's degree in their field and as independent of the expert group preparing the items were asked to associate the items and attributes, and the Q-matrices prepared were compared, 65 agreeable items were included in the trial practice of the test. The trial practice of the test was applied to a student group of 504 persons consisting of 7th and 8th grades. For the trial test, the mean of g, s, δ parameters of the items was calculated as 0.42, 0.30, 0.28, respectively. Upon comparison of the item parameters and the test-level fit statistics of the tests, it was concluded to perform the final test by considering the δ parameter as supported by the literature. The final test was applied to a group of 270 students from 7th and 8th grades. As a result of the analysis, the mean of g, s and δ parameters of the final test were calculated respectively as 0.39, 0.26 and 0.34.

Key Words: Cognitive Diagnostic Models, CDM, DINA, Q-matrix, Test Development

* Aynı isimli yüksek lisans tez çalışmasından üretilmiştir.

**Arş. Gör., Ege Üniversitesi, Eğitim Fakültesi, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı, elif.kubra.demir@ege.edu.tr

***Prof. Dr., Ankara Üniversitesi, Eğitim Bilimleri Fakültesi, Eğitimde Ölçme ve Değerlendirme Anabilim Dalı Emekli Öğretim Üyesi, nkoc@ankara.edu.tr

1.GİRİŞ

Bilişsel Tanı Modellerinde (BTM) eğitimde ölçmek istediğimiz özelliklerin ve geliştirilen başarı testlerinin birden fazla boyutlu olduğu ve dolayısıyla klasik test kuramının tek boyutlu yapısının bu ölçme araçlarının analizinde yetersiz kalacağı öngörülmüştür. BTM öğrencilerde belirli bir bilginin yapısını ya da bir becerinin gelişimini, öğrencilerin bilişsel düzeydeki güçlü ve zayıf yönlerini dikkate alarak çok boyutlu hesaplama yapabilmek amacıyla geliştirilmiştir. Bu nedenle BTM testin birbiriyle ilişkili birden fazla özelliği ölçtüğü durumlar için daha uygundur. Testteki her madde bu özellikleri ya da bilişsel bileşenleri ölçecek biçimde düzenlenir ve testte ölçülen birden fazla boyutu ölçebilir. BTM öğrencilerin örtük ölçekteki yetenek düzeylerine odaklanmaz, her bir bilişsel öge üzerindeki öğrenci performansını belirler. Elde edilen olasılıklar öğrencinin uzmanlaştığı becerilerin profilini çıkartacak şekilde dönüştürülebilir (Leighton ve Gierl, 2007). BTM ile geliştirilmiş bir testte, bir toplam puan veya toplam alt ölçek puanı yerine, testi alan bireylerin test kapsamındaki becerilerden hangilerine sahip olup olmadıkları belirlenir. Bu belirlemenin yanı sıra ölçülmek istenen özellikler ile bu özellikleri içeren maddeler arasındaki ilişki belirlenmeye çalışılır.

Test sonuçlarından yola çıkarak yapılan tanılamaların objektif olması ve bireylerin yeterli düzeylerini ve farklılıklarını daha keskin şekilde belirlemesi beklenir. Bilişsel tanı modelleri, bireylerin yetenek ve özelliklerini tanımlayarak profillerinin çıkarılmasını sağlar. Bu anlamda bilişsel tanı modelleriyle geliştirilmiş testler sadece değerlendirme sürecine değil aynı zamanda her bir öğrencinin eğitim ihtiyacını belirleme konusuna da hizmet eder (Cheng ve Chang, 2007).

BTM ile öğrencinin zihinsel sürecinin ve performansının altında yatan bilginin yapısı ortaya çıkartılabilir. 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 noktaları belirleyerek bu noktalar üzerine eğilebilir. Eğitim sürecinin güçlü ve zayıf yönlerini ortaya çıkartarak daha doğru öğrenme stratejileri düzenlenebilir.

Bir testi %50 doğruluk oranı ile yanıtlamanın pek çok yolu vardır. Örneğin 20 maddeden oluşan ve her doğru yanıtın 1 puan olduğu bir yabancı dil testinde 10 sorunun dilbilgisi, 10 sorunun ise algılama ile ilgili olduğu düşünülürse bu testten 10 puan almanın 184756 yolu vardır. Bu demek oluyor ki 184756 kişinin aynı yetenek düzeyinde olduğu kabul edildiği halde her birinin yetenek ve birikimi birbirlerinden farklıdır. Tüm dilbilgisi sorularını doğru yanıtlayan öğrenci ile hiçbir dilbilgisi sorusunu doğru yanıtlayamayan öğrencinin yeterlilikleri aynı sayılacağı gibi her iki birey için de aynı öğrenme etkinlikleri yapılacak ve bu durum öğrenim sürecine ve öğrencilerin eksikliklerinin giderilmesine yön veremeyecektir. Madde tepki kuramına göre aynı yetenek düzeyine sahip olan bu öğrenciler aslında farklı yeteneklere sahiptirler.

Jang'a (2008) göre, BTM öğrenmeyi değerlendirmeye değil, öğrenme için değerlendirmeye hizmet eder. Öğrenmenin değerlendirilmesi, öğrencilerin mevcut program çerçevesinde sahip oldukları kazanım sayısının belirlenmesini ve bir sonraki eğitim aşaması için hazırbuluşluklarının ortaya konmasını sağlar. Öğrenme için değerlendirme ise öğretmenlerin öğrenime yön verebilmelerini ve sınıf içi öğrenim durumunu düzenleyebilmelerini sağlar. Öğretmenler, öğrencilerin profillerine göre öğrenme etkinliklerini yeniden düzenleyebilir, yaklaşımını değiştirebilir ve öğrencilerin

eksikliklerine yönelik takviye edici öğrenme ortamları hazırlayabilirler. Bilişsel tanı modellerinin bu yaklaşımı öğrencileri de olumlu yönde etkilemektedir. Öğrencilere değerlendirmenin bir amaç değil araç olduğu düşüncesi verilmekte ve bu durum öğrenmeye devam ederek eksikliklerini gidermeleri konusunda pekiştirici olmaktadır. Öğrencilerin kendi durumlarını ve yeterlik düzeylerini ayrıntılı şekilde görebilir olmaları kendi öğrenmelerinin değerlendiricisi olmalarını sağlamakta ve mevcut birikimleriyle yeni öğrenmelerine yön verebilmektedirler.

Araştırma kapsamında bilişsel tanı modellerinden DINA modelin seçilmiş olması DINA modelin diğer modellere göre kuramsal alt yapısının kuvvetli olması, yetenek kestirimde başarılı olması ve modelle ilgili çalışmaların yapılmış olmasıdır. Ancak model genel olarak simülasyon çalışmalarıyla ve Q-matrisi hazır birkaç test üzerinden denenmiştir. Aynı zamanda bu testlerin çoğu matematik ve yabancı dil (İngilizce) alanı ile sınırlıdır. Test geliştirmede öğrencilerin yeteneklerinin kestirilerek profillerinin belirlenmesine ve buna bağlı olarak eğitim ihtiyaçlarının düzenlenmesine olan ihtiyaç giderek artmaktadır. Bunun yanı sıra DINA modelin gelişimi, uygulanabilirliği konusunda gerçek verilere dayanan çalışmaların sınırlı oluşu bir diğer problem durumudur. Bu nedenle bütün süreçleri ile DINA modele göre geliştirilen ve gerçek uygulama verisi kullanılan fen alanındaki bir test ile modele ilişkin araştırma yapılmak istenmektedir.

BTM yapısı gereği birden fazla özelliğin ölçüldüğü testler için daha uygundur. BTM ile test geliştirme sürecinde, testteki her bir madde ölçülmek istenen özelliklerin ve bilişsel yapıların temsilcisidir. Testteki her madde bu özelliklerden birini veya birkaçını ölçecek şekilde düzenlenir. BTM ile test geliştirme sürecinde maddelerin ölçülecek özelliklerle eşleştirilmesi için Q-matris kullanılır. Q-matris terimi ilk defa Tatsuoka (1991) tarafından kullanılmıştır ve maddeler ile maddelere ait özelliklerin dağılımını göstermektedir (Kato, 2009).

Q-matriste sütunlar özellikleri, satırlar ise maddeleri temsil etmektedir. BTM uygulamalarında kılavuzluk yapan Q-matris, 1-0 şekilde kodlama yapılarak oluşturulur. Bu kodlama “ağırlıklandırma” şeklinde tanımlanmış ve bir k niteliğinin bir i maddesinde bulunduğu durumda 1, bulunmadığı durumda 0 şeklinde kodlandığını belirtmiştir (Fischer, 1973).

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

q_{jk} ; k özelliği j maddesinde bulunuyorsa; 1,

k özelliği j maddesinde bulunmuyorsa; 0 şeklinde gösterilebilir.

Burada “özellik” kavramı nitelik, vasıf, görev, alt görev, bilişsel süreç veya beceri olarak tanımlanabilir (Tatsuoka, 1995). Q-matriste tanımlanan k tane özellik için 2^k tane örtük sınıf belirlenir. Tablo 1’de örnek bir Q matris verilmiştir.

Tablo 1.*Q- Matris*

Maddeler	Özellikler			
	α_1	α_2	α_3	α_4
1	0	1	0	0
2	1	1	0	0
3	0	1	1	0
4	0	1	1	1
5	1	0	1	1
6	1	1	1	1

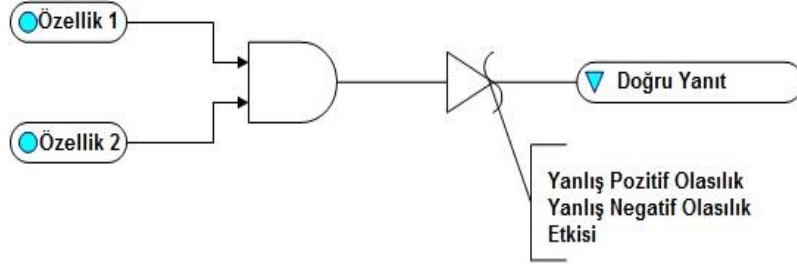
Tablo 1.'de verilen Q-matriste 4 adet özellik bulunduğundan 2^4 tane yani 16 adet örtük sınıf belirlenir. Bu örtük sınıflar (0000), (1000), (0100), (0010), (0001), (1100), (1010), (1001), (0110), (0101), (0011), (1110), (1101), (1011), (0111), (1111) şeklinde belirlenmektedir. Örtük sınıflar öğrencinin hangi özelliklere sahip olup hangilerine sahip olmadıklarını ifade eden gruplardır. (1000) örtük sınıfına dâhil olan öğrencinin ölçülen dört özellikten sadece α_1 olarak tanımlanan ilk özelliğe sahip olduğu anlaşılır.

Q-matrisin hazırlanma süreci bir maddeyi doğru yanıtlamak için hangi özellik veya özelliklerin gerekli olduğunun belirlendiği hipotetik bir süreçtir. BTM'de maddeler ve özellikler arasındaki ilişki ve öğrencilerin dâhil oldukları örtük sınıflar Q matris ile belirlendiğinden Q matrisin doğru oluşturulması oldukça önemlidir. Bu nedenle Q matris konu alanına hâkim uzmanlarca hazırlanmalıdır. Henson (2004), Q matrisin kalitesinin, öğrencilerin beceri profilinin kestirimini doğrudan belirlediğini ifade etmektedir. Bu nedenle, bu çalışmalar "tanısal analiz" için çok önemli bir nitelik taşımaktadır. Son dönem yapılan araştırmalarda Q matrisinin doğruluk miktarını ve etkisini ölçmeye yönelik sağlam kanıtlar sunan metotlar geliştirilmektedir aynı zamanda BTM'ne ilişkin parametreler de Q matris geçerliğine ilişkin kanıtlar sunmaktadır. (de la Torre ve Douglas, 2004).

1.1. DINA Model

DINA modelin temelleri Macready ve Dayton (1977), Tatsuoka (1983) ve Haertel (1989) tarafından atılmıştır. Haertel (1989) geliştirdiği modele "İkili Yetenek Modeli" adını vermiştir. Junker ve Sijtsma (2001) model ile ilgili çalışmaları ilerleterek modele DINA adını vermişlerdir. DINA "Deterministic Input Noisy And Gate" ifadesinin baş harflerini ifade etmektedir. "Deterministic input" terimi bir maddenin doğru yanıtlanabilmesi için gerekli olan özelliklerin o maddeyi yanıtlayan bireylerin örtük özelliklerinde tam olarak tanımlanıyor olduğunu, özelliğe sahip olma durumunun "1" olmama durumunun "0" ile gösterildiğini ifade etmektedir (Rupp ve Templin, 2008). Bir mühendislik terimi olan "and-gate" ise dilimizde de "ve kapısı" olarak isimlendirilen bir çeşit devredir. Ve Kapısı, sadece tüm girişleri 1 ise 1 verir, diğer tüm hallerde 0 verir. DINA modelde tek bir özelliğe sahip maddeler oluşturulabileceği gibi bir madde birden fazla özellik tarafından da temsil edilebilir. Modelde özelliklerle ilişkilendirilen maddelerin dağılımını gösteren

Q-matrisi hazırlamak gerekmektedir.



Şekil 1. DINA Modelin Biçimsel Gösterimi

DINA model, örtük özellik ve ilgili gözlenen özellik arasındaki ilişkiyi olasılık temeli olarak modellemekte ve her madde için s ve g parametresi olarak iki ayrı madde parametresi üretmektedir. s parametresi, İngilizce “*slip*” kelimesinin baş harfini ifade etmektedir. Bu parametre dilimizde *kaydırma parametresi* olarak da adlandırılır. g parametresi ise “*guess*” kelimesinin baş harfini belirtmektedir ve dilimizde *tahmin parametresi* olarak da isimlendirilir.

s parametresi, bireyin j maddesini doğru yanıtlayabilmek için gerekli özellik veya özelliklere sahip olmasına rağmen maddeyi yanlış yanıtlamasını ifade eder. Bu durum yanlış pozitif olasılık olarak tanımlanır. Maddeye ilişkin s parametresinin değeri ne kadar düşük olursa aranan özelliklere sahip bireylerin maddeyi doğru yanıtlama olasılığı o kadar artar. g parametresi ise bireyin j maddesini doğru yanıtlayabilmek için gerekli özellik veya özelliklere sahip olmamasına rağmen maddeyi doğru yanıtlamasını ifade eder. Bu durum doğru pozitif olasılık olarak tanımlanır. g parametresinin değeri ne kadar yüksek ise maddeyi doğru yanıtlamak için gerekli özelliğe sahip olmayan bireylerin maddeyi doğru yanıtlama olasılıkları artar. Benzer şekilde, g parametresinin değeri ne kadar düşükse maddenin sadece gerekli özelliğe sahip bireyler tarafından doğru yanıtlanma olasılığı o kadar artar (Zhang, 2006).

Maris (1999), g parametresinin alternatif bir yorumunu yapmış ve g parametresinin bireyin gerekli özelliğe sahip olmamasına rağmen zihinsel beceri ile maddeyi doğru yanıtlayabilmesi olarak tanımlamıştır. Bu durumda g parametresi, doğru yanıt için gerekli olduğu düşünülen özellikler dışında başka özellikleri kullanarak da maddenin doğru yanıtlandığı anlamına gelmektedir. Bu durum aynı zamanda Q matris geçerliliği olarak da düşünülebilir. Bir madde için “1” değerine yakın çıkan g parametresi maddenin doğru yanıtlanabilmesi için gerekli özelliğe sahip olmayan öğrenciler tarafından doğru yanıtlandığını belirttiği gibi aynı zamanda maddeyi doğru yanıtlamak için gerekli olan bazı özelliklerin belirlenemediği şeklinde de yorumlanabilir. Bu durumda Q matriste maddeye ilişkin olarak belirlenen özellikler dışında başka özelliklerin de madde ile ilişkili olduğu düşünülebilir. Tahmin parametresi bu anlamda MTK’daki şans parametresinden farklı bir yapıdadır. (de la Torre ve Douglas, 2004).

de la Torre (2008)’e göre madde ayırt edicilik indeksi δ_j (delta), $1 - s_j - g_j$ formülüyle hesaplanabilmektedir. Modelde s ve g parametreleri 0 ile 1 arasında değer alabildiğinden δ_j katsayısı $-1 \geq \delta_j \leq 1$ arasında değerler almaktadır. Örtük özelliğe göre öğrencileri

mükemmel düzeyde ayırt eden bir madde için $\delta_j = 1$ olmalıdır ve bu değer sıfıra yaklaştıkça maddenin ayırt edicilik gücü düşmektedir. DINA modelde hesaplanan δ_j parametresi konusunda tam bir ölçüt bulunmamakla birlikte daha yüksek δ_j değeri daha ayırıcı bir maddeyi işaret etmektedir. s parametresi kullanılarak elde edilen $1 - s_j$ değerinin yüksek olması maddeyi doğru yanıtlamak için gerekli özelliğe sahip bireyler tarafından kaydırma yapmadan doğru yanıtlanma olasılığının yüksek olduğunu ifade eder. Bu nedenle s veya $1 - s_j$ değerinin de madde ayırt ediciliği için kullanılabilmesi belirtilmektedir (de la Torre ve Lee, 2010). DINA modelde madde ayırt ediciliği için genellikle δ parametresi dikkate alınmakla birlikte g ve s parametrelerine dayalı olarak da madde ayırt ediciliğinin belirlenmesi test geliştirme sürecinde dikkat edilmesi gereken bir durumdur. Bu nedenle DINA modele göre test geliştirmede δ , s ve g parametreleri dikkate alınarak seçilen maddeler ile oluşturulan testlerin madde ve model istatistiklerinin karşılaştırılması uygundur.

2. YÖNTEM

2.1. Araştırma Modeli

Araştırmada, DINA modele göre test geliştirme çalışması yapılarak oluşturulan testin psikometrik özelliklerinin belirlenmesi ve modelin kuramsal temeline katkı sağlanması amaçlanmıştır. Bu yönüyle araştırma temel araştırma türünde betimsel bir çalışmadır.

2.2. Çalışma Grubu

Çalışma 7. ve 8. sınıf öğrencileri ile yürütülmüştür. Deneme uygulaması 504 kişilik çalışma grubu üzerinden yürütülmüştür. Testin nihai hali ise 270 kişilik öğrenci grubuna uygulanmıştır.

2.3. Veri Toplama Aracı

Bilişsel tanı modelleri kapsamında DINA modele göre alan uzmanları tarafından hazırlanan maddeler doğrultusunda araştırmacı tarafından geliştirilen başarı testi için 7. sınıf Fen ve Teknoloji dersi Potansiyel ve Kinetik Enerji konusu seçilmiştir. Konu alanının 7. ve 8. sınıf Fen ve Teknoloji dersi kapsamında olması, Seviye Belirleme Sınavı (SBS) gibi geniş ölçekli sınavlarda ağırlık verilen önemli bir konu olması testin hem 7. hem de 8. sınıflara uygulanabilir olmasını sağlamıştır. Bu araştırmanın gereği olarak geliştirilen başarı testi için DINA modele uygun olarak test geliştirme aşamalarının gerekleri yerine getirilmeye çalışılmış ve aşağıdaki adımlar incelenmiştir.

2.3.1. Testin Amacı

Araştırmada kullanılmak üzere geliştirilen test, ilköğretim 7. ve 8. sınıf öğrencilerinin potansiyel ve kinetik enerji konu alanında sahip oldukları özellikleri ve dâhil oldukları örtük sınıfları belirlemeyi amaçlayan bir başarı testidir.

2.3.2. Özelliklerin belirlenmesi

Öğrencilerin Potansiyel ve Kinetik Enerji konu alanı kapsamında sahip olması beklenen özellikleri belirlemek için 5 uzmanın görüşü alınmıştır. Görüşü alınan uzman grup 3 yıldan fazla süredir Millî Eğitim Bakanlığı'na bağlı ilköğretim okullarında görev yapan Fen ve Teknoloji öğretmenlerinden oluşmaktadır. Konu alanı ile ilgili olarak öğrencilerin sahip olması gereken 4 temel özellik uzman grubun ortak kararı ile belirlenmiştir. Bu özellikler şöyle sıralanmıştır;

1. Kütle potansiyel ve kinetik enerjideki rolü (α_1 , kütle)
2. Yüksekliğin potansiyel ve kinetik enerjideki rolü (α_2 , yükseklik)
3. Hızın potansiyel ve kinetik enerjideki rolü (α_3 , hız)
4. Sürtünme kuvvetinin potansiyel ve kinetik enerjideki rolü (α_4 , sürtünme kuvveti)

Belirlenen 4 temel özelliğe sahip olan bireylerin potansiyel ve kinetik enerji konusuna hâkim olması beklenmektedir.

2.3.3. Örtük Özellik Grubu Oluşturma

Öğrencilerin sahip olması gereken özellikler ve sayısı belirlendikten sonra örtük özellik grupları oluşturulur. Belirlenen 4 özellik için toplam 16 kombinasyon mevcuttur. Örtük özellik sınıfları 2^k formülü ile hesaplanırsa da test geliştirilirken hiçbir özelliğe sahip olmayan bir maddenin teste yer alması mümkün olmadığından, hiçbir özelliğe sahip olmayan (0000) örtük sınıf da oluşmayacaktır. Bu durumda test geliştirilirken oluşturulması gereken örtük sınıflar 2^k-1 formülü ile hesaplanabilir ve maddelerin temsile edeceği 15 örtük sınıf belirlenir. Tablo 2.'de özellikler ile örtük sınıflar verilmiştir.

Tablo 2.

Teste Ait Özellik – Örtük Sınıf Matrisi

Özellik	Örtük sınıf	Özellik	Örtük sınıf
α_1	1000	α_2, α_4	0101
α_2	0100	α_3, α_4	0011
α_3	0010	$\alpha_1, \alpha_2, \alpha_3$	1110
α_4	0001	$\alpha_1, \alpha_2, \alpha_4$	1101
α_1, α_2	1100	$\alpha_1, \alpha_3, \alpha_4$	1011
α_1, α_3	1010	$\alpha_2, \alpha_3, \alpha_4$	0111
α_1, α_4	1001	$\alpha_1, \alpha_2, \alpha_3, \alpha_4$	1111
α_2, α_3	0110		

2.3.4. Madde Türünün ve Sayısının Belirlenmesi

DINA model henüz açık suçlu sorularda kullanılmamış ve analizleri çalışılmamıştır. Geliştirilen teste sınıf düzeyi de göz önünde bulundurularak 4 seçenekli çoktan seçmeli madde tipi kullanılması uygun bulunmuştur. Örtük sınıf sayısı dikkate alındığında her örtük sınıfa ait 5 madde yazılması planlanmıştır. Beş kişiden oluşan uzman grubun her bir örtük sınıfa ait madde yazması istenmiştir. Testte 15 örtük sınıf bulunduğundan 75 maddelik madde havuzu oluşturulmuştur.

2.3.5. Q- Matrisin Hazırlanması

DINA modele göre test geliştirilirken bireylerin sahip olduğu örtük özelliklerin saptanabilmesi ve analizin doğru çalışabilmesi için maddelerin sahip olduğu özelliklerin doğru şekilde belirlenmesi gerekir. Maddeler ile özelliklerin ilişkilendirilmesi için uzman görüşüne başvurulmuştur. Uzman grup, doktora düzeyinde eğitim almış iki Ölçme ve Değerlendirme uzmanı ile Fen Bilgisi Öğretmenliği alanında görev yapan üç öğretim görevlisinden oluşan beş kişilik bir gruptur. Q-matrisin belirlenmesinde bilgi ve deneyimlerine başvuru uzman grup maddeleri hazırlayan uzman gruptan bağımsızdır.

Uzman gruba DINA model ve Q-matris ile ilgili bilgilendirme yapılmıştır. Aynı zamanda grup testin kapsamı, belirlenen özellikler ve kendilerinden beklenen madde-özellik

ilişkilendirmesi konusunda yeterince bilgilendirilerek Q matris oluşturma ile ilgili örnek uygulamalar yapılmıştır. Uzmanlardan bir maddenin doğru yanıtlanabilmesi için öğrencinin test kapsamında belirlenen 4 özellikten hangisi veya hangilerine sahip olması gerektiğinin dikkate alınması beklenmektedir. Her bir uzmanın 75 maddeden oluşan testi inceleyerek Q matris oluşturmaları istenmiştir. Oluşturulan 5 adet Q matris incelenerek uzmanların madde-özellik ilişkilendirmesinde hemfikir oldukları 65 madde seçilmiştir. Seçilen maddeler üç ölçme ve değerlendirme uzmanı tarafından ölçme ilkeleri ve bilimsel doğruluk açısından tekrar incelenmiş ve bir Türkçe öğretmeni tarafından da dil ve anlatım bakımından gerekli düzeltmeler yapılmıştır. DINA model parametreleri maddenin özellikleri ile ilgili bilgi vermenin yanı sıra Q-matris ile tanımlanan özelliklerin maddeyi temsil etme düzeyi ile ilgili de bilgi vermektedir. Tablo 3.'te uzmanlar tarafından belirlenen özelliklerin ilişkilendirildiği maddeler ve özellikle ilişkili toplam madde sayıları verilmiştir.

Tablo 3.*Madde – Özellik İlişkisi*

Özellik	Maddeler	N
Kütlenin enerjideki rolü (α_1)	1,2,3,4,8,9,10,16,17,19,20,21,23,24,26,28,30,31,32,33,34,35,36,37,38,39,41,42,44,46,47,49,51,52,53,54,55,57,60,61,63	41
Yüksekliğin enerjideki rolü (α_2)	1,3,5,6,11,12,13,14,16,19,20,22,,25,26,27,28,29,40,30,31,33,35,36,38,42,43,44,45,46,51,52,53,54,55,57,58,60,61,63,64,65	40
Hızın enerjideki rolü (α_3)	2,4,5,7,11,13,15,16,17,18,19,22,23,24,27,28,29,30,31,34,39,40,41,44,45,46,48,49,50,53,54,55,57,58,60,61,63,65	38
Sürtünme kuvvetinin enerjideki rolü (α_4)	6,7,8,9,10,12,14,17,18,20,23,24,26,27,28,29,31,34,37,40,42,43,44,45,47,48,50,51,52,53,55,56,57,59,60,62,63,65	38

Tablo 3. incelendiğinde özelliklerle ilişkili maddeler görülmektedir. Özelliklerle ilişkili madde sayısı dikkate alındığında α_1 özelliğinin toplam 41 madde ile temsil edildiği görülmektedir. Aynı şekilde α_2 özelliği 40 adet, α_3 özelliği 38 adet ve α_4 özelliği de 38 adet madde ile ilişkilendirilmiştir. Deneme uygulaması sonrasında madde seçimi yapılmıştır ve 25 maddeden oluşan testin nihai formu oluşturulmuştur

2.4. Verilerin Analizi

Veriler OxEdit programında analiz edilmiştir. Öğrencilerin 1-0 şeklinde kodlanmış yanıtları ile teste ilişkin Q-matrix kullanılarak program çalıştırılır. Program, DINA modelle ilişkin özelliklerin sonsal dağılımlarını, öğrencilerin örtük sınıflarını, model ve madde uyum indekslerini ve madde parametrelerini ve bu parametrelere ait standart hataları vermektedir.

3. BULGULAR

3.1. Deneme Uygulamasına Ait Bulgular

Deneme uygulamasına ait test puanları ortalaması 34.27'dir. Ortanca (34.00) ve tepe değeri (31.00) de aritmetik ortalamaya çok yakın değerler almıştır. Bu değerler dağılımın normal dağılıma yakın olduğunu gösterir. Öğrencilerin testten aldıkları puanlar 11.00 ile 60.00 arasında değiştiğinden dağılımın dizi genişliği 49.00 olarak hesaplanmıştır. Testin standart sapması ise 10.60 olarak hesaplanmıştır. Testin deneme uygulaması sonucu elde edilen KR-20 iç tutarlılık katsayısı 0.88 olarak bulunmuştur. Çarpıklık (0.136) ve basıklık (0.109) katsayıları incelendiğinde ise dağılımın normalden bir miktar sivri ve bir miktar sağa çarpık olduğu anlaşılmaktadır. Dağılıma ilişkin Kolmogorov-Smirnov testi yapılmıştır. Test sonucuna göre z değeri 1.342, anlamlılık değeri ise 0.055 olarak bulunmuştur, $p > 0.05$ olduğundan dağılımın normal olduğu yorumu yapılabilir. Deneme uygulaması sonucunda dağılıma ilişkin elde edilen tüm bilgiler Tablo 4.'te verilmiştir.

Tablo 4.

Deneme Uygulaması Sonucu Elde Edilen Test İstatistikleri

İstatistik	Değer	İstatistik	Değer
Aritmetik Ortalama	34.27	En Düşük Puan	11.00
Ortanca	34.00	En Yüksek Puan	60.00
Tepe Değer	31.00	KR-20	0.88
Standart Hata	0.47		
Dizi Genişliği	49.00		
Standart Sapma	10.60		
Çarpıklık	0.14		
Basıklık	0.11		

3.1.1. Deneme Uygulaması DINA Model - Veri Uyum İstatistikleri Model Veri Uyumu

Modele ait model veri uyum istatistik değerleri hesaplanmıştır. Bu değerler Tablo 5.'de verilmiştir. Teste ilişkin madde uyum istatistikleri ise Tablo 6.'da verilmiştir. Bu değerler karşılaştırmalı veriler olduğundan ve DINA modele göre analiz edilmiş madde sayıları ve örneklem büyüklüğü ilişkili ölçüt olabilecek çalışmalar bulunmadığından karşılaştırma yapılmamıştır. Bu alanda yapılacak olan çalışmalara katkı sağlaması amacıyla uyum istatistikleri Tablo 5. ve Tablo 6.'da verilmiştir.

Tablo 5.

Model Veri Uyumu İstatistikleri

-2LL	41909.06
AIC	42199.06
BIC	42810.18

3.1.2. DINA Model Madde Parametreleri ve Madde Seçimi**Tablo 6.***Madde Veri Uyumu İstatistikleri*

	Prop	Z (corr)	Log (OR)
Mean Abs. Dev.	0.0046	0.2614	0.2546
Max Abs Dev.	0.0161	0.2865	1.2019
SE Max Abs Dev	0.0222	0.0449	0.2005

DINA modelde maddelere ilişkin g ve s parametreleri ile bu parametrelerin standart hatalarına ulaşılmıştır. Tablo 7.'de maddelere ilişkin g, s değerleri ile bu değerlerin standart hataları verilmiştir. Aynı zamanda DINA model parametreleri ile KTK parametrelerinin farklı olduğunun incelenebilmesi için KTK'ya ait madde güçlük indeksi (p_j), madde ayırt edicilik indeksi (r_{jx}) ve madde test korelasyonu (r_j) verilmiştir.

Tablo 7.*DINA Model Madde Parametreleri*

Madde No	g	SE (g)	s	SE (s)	δ	p_j	r_{jx}	r_j
1	0.47	0.03	0.22	0.03	0.31	0.61	0.23	0.18
2	0.34	0.03	0.18	0.03	0.48	0.69	0.26	0.24
3	0.59	0.03	0.12	0.03	0.29	0.55	0.6	0.49
4	0.59	0.03	0.15	0.03	0.26	0.82	0.39	0.4
5	0.52	0.03	0.45	0.04	0.03	0.69	0.37	0.32
6	0.39	0.03	0.3	0.03	0.31	0.72	0.26	0.28
7	0.43	0.03	0.33	0.04	0.24	0.71	0.52	0.43
8	0.4	0.03	0.46	0.04	0.14	0.73	0.46	0.4
9	0.4	0.03	0.45	0.04	0.15	0.51	0.25	0.25
10	0.24	0.03	0.59	0.04	0.17	0.28	0.35	0.34
11	0.51	0.03	0.57	0.03	-0.07	0.35	-0.08	-0.09
12	0.23	0.03	0.5	0.04	0.27	0.22	0.03	0.07
13	0.46	0.03	0.16	0.03	0.38	0.64	0.53	0.48
14	0.5	0.03	0.24	0.03	0.26	0.76	0.44	0.45
15	0.41	0.04	0.26	0.03	0.32	0.69	0.32	0.3
16	0.4	0.03	0.2	0.03	0.4	0.44	-0.05	-0.01
17	0.41	0.03	0.71	0.04	-0.11	0.6	0.47	0.39
18	0.5	0.03	0.19	0.03	0.31	0.44	0.22	0.19
19	0.37	0.03	0.26	0.04	0.37	0.61	0.63	0.55
20	0.41	0.03	0.45	0.04	0.15	0.41	0.44	0.39
21	0.32	0.03	0.36	0.04	0.32	0.77	0.38	0.37
22	0.17	0.03	0.1	0.02	0.74	0.25	0.27	0.26
23	0.28	0.02	0.18	0.03	0.54	0.66	0.62	0.51
24	0.36	0.03	0.19	0.03	0.45	0.6	0.51	0.41
25	0.31	0.03	0.17	0.03	0.52	0.62	0.55	0.46
26	0.32	0.03	0.38	0.04	0.3	0.52	0.53	0.4
27	0.34	0.03	0.3	0.04	0.36	0.68	0.58	0.5
28	0.47	0.03	0.11	0.03	0.42	0.7	0.39	0.36
29	0.43	0.03	0.23	0.03	0.34	0.61	0.55	0.45
30	0.46	0.03	0.47	0.04	0.07	0.3	0.23	0.21

Tablo 7. devamı

31	0.32	0.03	0.37	0.04	0.3	0.54	0.35	0.25
32	0.3	0.03	0.26	0.03	0.43	0.61	0.23	0.18
33	0.45	0.03	0.12	0.03	0.43	0.32	0.17	0.18
34	0.41	0.03	0.17	0.03	0.42	0.53	0.42	0.34
35	0.38	0.03	0.35	0.04	0.27	0.29	0.33	0.3
36	0.26	0.03	0.32	0.04	0.42	0.38	0.49	0.39
37	0.43	0.03	0.19	0.03	0.38	0.29	0.21	0.2
38	0.38	0.03	0.36	0.04	0.26	0.63	0.44	0.36
39	0.43	0.03	0.16	0.03	0.4	0.62	0.25	0.19
40	0.5	0.03	0.23	0.03	0.27	0.7	0.51	0.44
41	0.42	0.03	0.41	0.04	0.16	0.61	0.52	0.46
42	0.46	0.03	0.29	0.04	0.24	0.24	0.01	0.01
43	0.47	0.03	0.2	0.03	0.33	0.52	0.54	0.44
44	0.54	0.03	0.13	0.03	0.33	0.46	0.49	0.44
45	0.41	0.03	0.27	0.04	0.32	0.45	0.18	0.18

Maddelere ilişkin g parametreleri incelendiğinde g parametrelerinin ortalaması 0.42, ortanca değeri 0.42 ve tepe değeri ise 0.41 olarak hesaplanmıştır. En düşük g parametresine sahip madde 62. maddedir ve aldığı değer $g=0.13$ olarak belirlenmiştir. En yüksek g parametresine sahip madde ise 61. maddedir ve aldığı değer $g=0.63$ olarak belirlenmiştir. Bu durumda g parametre değerleri 0.13 ile 0.63 arasında değişmektedir. Q-matris incelendiğinde en yüksek g parametresine sahip olan 61. Maddenin ait olduğu örtük sınıf (1 1 1 0) sınıfıdır ve bu durumda $\alpha_1, \alpha_2, \alpha_3$ özellikleri ile ilişkilidir. Maddenin yüksek g parametresine sahip olması α_1, α_2 ve α_3 özelliklerine sahip olmayan öğrenciler tarafından da maddenin doğru yanıtlanma olasılığının yüksek olduğunu ifade eder. Düşük g parametre değerine sahip olan 62. Madde ise α_4 özelliği ile ilişkili ve (0 0 0 1) örtük sınıfına dâhildir. Maddenin g parametresinin düşük değer alması α_4 özelliğine sahip olmayan öğrenciler tarafından maddenin doğru yanıtlanma olasılığının düşük olduğunu ifade etmektedir. Aynı zamanda maddenin doğru yanıtlanması için gerekli olan özellik dışında başka bir özellik ile ilişkili olmadığını belirtmektedir. Bu ifade maddenin Q-matriste son derece doğru tanımlandığının da göstergesi olarak kabul edilebilir.

Deneme uygulamasına ait sj parametreleri incelendiğinde sj parametrelerinin ortalaması 0.30, ortanca değeri 0.26, ve tepe değeri ise 0.18 olarak hesaplanmıştır. Maddelere ilişkin s parametreleri incelendiğinde en düşük s parametresine sahip madde 22. maddedir ve aldığı değer $s=0.10$ olarak belirlenmiştir. En yüksek s parametresine sahip madde ise 17. maddedir ve aldığı değer $s=0.73$ olarak belirlenmiştir. Bu durumda sj parametresine ait en düşük değer 0.10 iken en yüksek değer de 0.71 olarak gözlemlenmiştir. En düşük s parametre değerine sahip olan 22. Madde (0 1 1 0) örtük sınıfını temsil eder ve dolayısıyla α_2 ve α_3 özellikleri ile ilgilidir. Bu durumda α_2 ve α_3 özelliklerine sahip olan öğrencilerin maddeyi doğru yanıtlanma olasılıkları yüksektir. En yüksek s parametresine sahip olan 17. Madde ise (1 0 1 1) örtük sınıfına dâhildir ve α_1, α_3 ve α_4 özellikleri ile ilişkilidir. Maddeye ilişkin s parametresinin yüksek olması maddenin α_1, α_3 ve α_4 özelliklerine sahip olan öğrenciler tarafından da yanlış yanıtlanma olasılığının yüksek olduğunu ifade etmektedir.

Deneme uygulamasına ait δj değerleri incelendiğinde bu değerlerin ortalaması 0.28, ortanca değeri 0.30 ve tepe değeri ise 0.26 olarak hesaplanmıştır. δj parametresine ilişkin en düşük değer -0.11 iken en yüksek değer de 0.74 olarak belirlenmiştir. Maddenin δj

değeri yükseldikçe ayırt edicilik gücü artmaktadır. En yüksek δ değerine sahip olan ve ayırt ediciliği en yüksek olan madde 22. madde iken en düşük δ değerine sahip olan 17. maddenin ise ayırt edicilik gücü en düşük olan madde olarak değerlendirilebilir. En nitelikli ve en niteliksiz madde belirlemede s ve δ parametreleri aynı maddeleri belirlemiştir.

Madde ayırt ediciliği dikkate alındığında DINA model ile KTK parametreleri arasında farklılıklar görülmektedir. DINA modele göre δ , s ve g değerleri dikkate alındığında ayırt ediciliği en yüksek madde olan 22. M-maddenin KTK'ya göre madde ayırt edicilik indeksi (r_{jx}) 0.27, madde test korelasyonu 0.26 olarak hesaplanmıştır. Benzer şekilde DINA model parametrelerine göre ayırt ediciliği en düşük olan 17. maddenin KTK'ya göre madde ayırt edicilik indeksi (r_{jx}) 0.47, madde test korelasyonu 0.39 olarak hesaplanmıştır.

Madde güçlüğüne ilişkin durum incelendiğinde de benzer bulgulara rastlanmıştır. Düşük g (0.13) ve yüksek s (0.55) parametre değerlerine sahip 62. maddenin KTK'ya göre madde güçlük indeksi değeri (p) 0.46 olarak hesaplanmıştır. Madde DINA modele göre zor maddelerden biri olarak yorumlanabilirken KTK'ya göre orta güçlüğü yakın bir maddedir. Bu durumda DINA model ile KTK madde istatistiklerinin farklılık gösterdikleri ifade edilebilir

3.2. g , s ve δ Parametreleri Dikkate Alınarak Oluşturulan Testlere İlişkin Test ve Madde İstatistikleri

DINA modele ilişkin g parametresinin değeri ne kadar düşük olursa maddeyi doğru yanıtlayabilmek için gerekli olan özelliğe sahip olmayan bireyler tarafından doğru yanıtlanma olasılığı da o kadar düşüktür. Dört seçenekli bir teste öğrencilerin maddeye ilişkin özelliğe sahip olmasa da doğru yanıt bulabilme olasılığı %25'tir. Seçeneklerden birini veya birkaçını eleyebilen bir öğrenci için ise bu olasılık daha da artmaktadır. DINA modelde genel kabul görmüş bir kesme değeri olmamakla birlikte beş seçenekli maddelerden oluşan test için g parametresi 0.35 değerinin altındaki maddeler nitelikli maddeler olarak değerlendirilebilir (Li, 2008). Geliştirilen test dört seçenekli bir test olduğu için ve doğru yanıt ulaşma olasılığı beş seçenekli bir testten %5 daha fazla olduğu için g parametresi 0.40 değerinin altında olan maddelerden ilk 25 madde seçilmiştir.

Literatürde s parametresinin tek başına da madde ayırt ediciliğini belirlemede yeterli olduğuna ilişkin çalışmalar vardır. Bu çalışmalarda s parametresi için 0.30 değerinin altındaki maddelerin ayırt edicilik güçlerinin yüksek olduğu ifade edilmiştir (Li, 2008). Geliştirilen testte maddelere ilişkin s parametreleri incelendiğinde 65 madde içinde 40 maddenin s parametresi değerinin 0.30 olduğu görülmüştür. Bu durum testteki maddelerin nitelikli maddeler olduğuna dair fikir sahibi olunmasını sağlar. Maddelere ilişkin s parametreleri dikkate alınarak en düşük s değerine sahip ilk 25 madde seçilmiştir.

DINA modelde hesaplanan δ_j parametresi konusunda tam bir ölçüt bulunmamakla birlikte daha yüksek δ_j değeri daha ayırıcı bir maddeyi işaret etmektedir. Bununla birlikte yapılan çalışmalarda δ değeri 0.30 ve üzerinde olan maddelerin ayırt edicilik güçlerinin yeterli olduğu ileri sürülmüştür (Li ve Oranje, 2007). Teste ilişkin g ve s parametreleri için kullanılacak 0.40 ve 0.30 kesme değerleri δ 'ya ait formülde kullanıldığında da δ değerine ilişkin kesme değerinin 0.30 olarak alınması beklenmektedir. Maddeler δ_j değerlerine göre sıralanmış ve değerleri 0.30 ve üzerinde olan ilk 25 madde seçilmiştir.

g , s ve δ parametrelerine göre 25 madde ile oluşturulan testler için mevcut Q-matris düzenlenerek her bir teste ait Q-matris hazırlanmıştır. Q-matrisler hazırlandıktan sonra testte ölçülen dört özelliğin temsil edildiği madde sayıları incelenmiştir. Testte ölçülen her bir özelliğin birbirine yakın sayıda maddelerde temsil edilmesi istenir. Böylece ölçülen özellikler arasında denge sağlanmış olur. Her hangi bir testte bu örüntü sağlanamıyorsa testin kapsam geçerliğine ilişkin sorun teşkil eder. g , s ve δ parametrelerine göre oluşturulan testlere ait özellik madde sayısı ilişkisi Tablo 8.'de verilmiştir.

Tablo 8.

g , s ve δ Parametrelerine Göre Oluşturulan Testlerde Özelliklerin Temsil Edildiği Madde Sayısı

Özellikler	Parametrelere Göre Oluşturulan Testlerde Özellikleri Temsil Eden Madde Sayısı		
	g	s	δ
α_1	16	16	15
α_2	14	13	13
α_3	11	17	16
α_4	15	14	13

Oluşturulan testlerdeki özellikler ile temsil edildikleri madde sayıları dikkate alındığında her üç testte de özelliklerin dengeli dağıldığı ve bu nedenle kapsam bakımından üç testin de uygun testler olduğu ifade edilebilir. g , s ve δ parametrelerine göre oluşturulan testler mevcut veri üzerinden düzenlenen Q matrisleri ile analiz edilmiştir. Analizler sonunda testlere ilişkin model- veri uyum istatistikleri karşılaştırılmıştır. Tablo 9'da testlere ait model veri uyum değerleri verilmiştir.

Tablo 9.

g , s ve δ Parametrelerine Göre Oluşturulan Testlerin Model Veri Uyum İstatistikleri

Model- Veri Uyum İstatistikleri	g	s	δ
-2LL	16192,06	16472.17	16089.23
AIC	16322,06	16602.17	16219.23
BIC	16596,01	16876.11	16493.18

Tablo 9. incelendiğinde -2LL, AIC ve BIC değerlerinin her üçünde de en küçük değer δ parametresine göre oluşturulan teste aittir. Bu durumda δ parametresine göre oluşturulan testin model veri uyumu diğer testlere göre daha yüksektir. Analiz sonucunda her bir teste özelliklerin gözlenme olasılıkları belirlenmektedir. Testlere göre özellik olasılık değerleri Tablo 10.'da verilmiştir. Tablo.10 incelendiğinde değerlerin testlere göre farklılıkları görülmektedir.

Tablo 10.*g, s ve δ Parametrelerine Göre Oluşturulan Testlerde Özelliklerin Gözlenme Olasılıkları*

Özellikler	Parametrelere Göre Oluşturulan Testlerde Özelliklerin Gözlenme Olasılıkları		
	<i>g</i>	<i>s</i>	δ
α_1	0.74	0.53	0.53
α_2	0.48	0.56	0.65
α_3	0.81	0.60	0.73
α_4	0.54	0.50	0.65

Analiz sonucunda her teste ait elde edilen madde parametreleri incelenmiştir. Madde parametrelerine ilişkin değerler Tablo. 11’de verilmiştir.

Tablo 11.*g, s ve δ Parametrelerine Göre Oluşturulan Testlerin Madde Parametreleri*

<i>g</i> Parametresine Göre Oluşturulan Teste İlişkin Değerler				<i>s</i> Parametresine Göre Oluşturulan Teste İlişkin Değerler				δ Parametresine Göre Oluşturulan Teste İlişkin Değerler			
Madde No	<i>g</i>	<i>s</i>	δ	Madde No	<i>g</i>	<i>s</i>	δ	Madde No	<i>g</i>	<i>s</i>	δ
2	0.30	0.32	0.37	1	0.48	0.23	0.29	2	0.46	0.30	0.24
6	0.56	0.40	0.04	2	0.40	0.43	0.17	13	0.49	0.36	0.15
8	0.34	0.32	0.33	3	0.41	0.33	0.25	15	0.56	0.64	-0.21
10	0.42	0.62	-0.04	4	0.31	0.44	0.25	16	0.53	0.24	0.23
12	0.36	0.61	0.02	13	0.48	0.29	0.22	19	0.45	0.33	0.23
16	0.56	0.40	0.05	16	0.50	0.27	0.23	21	0.20	0.36	0.44
19	0.38	0.17	0.45	18	0.39	0.27	0.33	22	0.37	0.26	0.37
21	0.26	0.31	0.43	22	0.42	0.42	0.16	23	0.38	0.19	0.42
22	0.47	0.39	0.14	23	0.47	0.26	0.26	24	0.39	0.20	0.42
23	0.33	0.31	0.36	24	0.36	0.14	0.50	25	0.28	0.28	0.44
24	0.41	0.37	0.22	25	0.17	0.21	0.62	27	0.40	0.13	0.47
25	0.28	0.26	0.46	28	0.37	0.13	0.49	28	0.44	0.09	0.47
26	0.30	0.21	0.49	29	0.51	0.39	0.10	29	0.28	0.53	0.19
27	0.43	0.11	0.46	33	0.58	0.35	0.07	32	0.37	0.30	0.33
31	0.51	0.50	-0.02	34	0.47	0.19	0.34	33	0.46	0.15	0.39
32	0.46	0.70	-0.16	37	0.37	0.24	0.39	34	0.38	0.17	0.45
34	0.44	0.37	0.18	39	0.49	0.08	0.42	36	0.25	0.28	0.46
35	0.26	0.19	0.55	40	0.52	0.13	0.35	37	0.59	0.16	0.24
36	0.37	0.19	0.44	43	0.47	0.46	0.07	39	0.41	0.21	0.38
38	0.54	0.50	-0.04	44	0.45	0.37	0.19	43	0.42	0.42	0.16
50	0.63	0.18	0.19	46	0.35	0.54	0.11	44	0.33	0.25	0.42
52	0.59	0.15	0.26	47	0.50	0.12	0.38	45	0.37	0.28	0.35
59	0.15	0.24	0.61	48	0.53	0.32	0.27	47	0.47	0.21	0.32
62	0.23	0.28	0.48	59	0.21	0.35	0.56	48	0.18	0.30	0.52
65	0.36	0.16	0.48	63	0.61	0.29	0.10	59	0.21	0.29	0.50
\bar{x}	0.40	0.33	0.27		0.43	0.29	0.27		0.39	0.28	0.33
Ortanca	0.38	0.31	0.33		0.47	0.27	0.25		0.39	0.28	0.38
En Düşük	0.15	0.11	-0.16		0.17	0.08	0.07		0.18	0.09	-0.21
En Yüksek	0.63	0.70	0.61		0.61	0.54	0.62		0.59	0.64	0.52

Geliştirilen testin deneme uygulamasından sonra DINA model hata parametrelerinden *g*, *s* ve δ değeri en iyi olan 25 madde belirlenmiştir. Testlerin analizleri sonucunda her üç testin madde parametreleri verilmiştir. Aynı zamanda bu değerlere ilişkin aritmetik

yoluyla doğru yanıtlanma olasılığının da yüksek olması anlamına gelmektedir. Bu durumda g parametresi yüksek olan maddelerin daha kolay olduğu ifade edilebilir. Aynı zamanda literatürde g parametresi yüksek ve s parametresi düşük maddelerin kolay maddeler olduğu belirtilmiştir (Zhang, 2006). Maddelerden birbirine yakın veya aynı g değerine sahip olanlardan s değeri daha küçük olan madde daha kolay olarak yorumlanmış ve maddeler teste yerleştirilmiştir ve nihai test 270 öğrenciye uygulanmıştır.

Q-matriste madde - özellik eşleştirmeleri incelendiğinde özelliklerin temsil edildiği madde sayısına da ulaşılabilir. Tablo 12’de ayrıca verilen bu bilgiye göre testte α_1 özelliği 15 adet, α_2 özelliği 13 adet, α_3 özelliği 16 adet ve α_4 özelliği 13 adet madde tarafından temsil edilmiştir. Bu durumda testte özelliklerin dengeli dağıldığı ileri sürülebilir.

Tablo 12.
Nihai Testte Özelliklerin Temsil Edildiği Madde Sayısı

Özellik	Madde Sayısı
α_1	15
α_2	13
α_3	16
α_4	13

Tablo 13.
Madde Veri Uyumu İstatistikleri

	Prop	Z (corr)	Log (OR))
Mean Abs. Dev.	0.0072	0.0665	0.2848
Max Abs Dev.	0.0174	0.2730	1.1393
SE Max Abs Dev	0.0280	0.0612	0.2644

3.3.2. Nihai Teste İlişkin Madde Parametreleri

Nihai testte ilişkin g (tahmin), s (kaydırma) ve δ (1-s-g) parametreleri ve bu parametrelere ait standart hatalar elde edilmiştir. Maddelere ait parametreler Tablo 14.’de verilmiştir. Madde parametrelerine ilişkin yorumlar her madde için benzerdir. Aynı yorumların tekrarlanmaması adına madde parametrelerine ait genel bir açıklama yapılmış ve önemli görülen maddeler incelenmiştir.

Tablo 14.
Nihai Test Madde Parametreleri

Madde No	g	SE (g)	s	SE (s)	δ
1	0.26	0.05	0.48	0.04	0.25
2	0.52	0.05	0.34	0.04	0.14
3	0.48	0.09	0.33	0.04	0.19
4	0.45	0.04	0.17	0.05	0.37
5	0.38	0.04	0.10	0.03	0.52
6	0.38	0.05	0.34	0.03	0.27
7	0.49	0.04	0.31	0.04	0.20
8	0.58	0.04	0.23	0.04	0.18
9	0.45	0.04	0.08	0.03	0.46
10	0.64	0.04	0.32	0.04	0.05
11	0.47	0.04	0.28	0.03	0.26
12	0.28	0.04	0.05	0.04	0.67
13	0.40	0.04	0.11	0.02	0.48
14	0.46	0.04	0.20	0.04	0.34
15	0.34	0.04	0.16	0.04	0.50
16	0.19	0.04	0.33	0.04	0.48
17	0.36	0.04	0.59	0.05	0.04
18	0.21	0.04	0.57	0.05	0.23
19	0.46	0.04	0.37	0.04	0.17
20	0.14	0.04	0.19	0.04	0.67
21	0.45	0.04	0.04	0.05	0.51
22	0.32	0.04	0.42	0.05	0.26
23	0.13	0.04	0.18	0.03	0.70
24	0.47	0.04	0.10	0.05	0.43
25	0.47	0.04	0.30	0.04	0.23
\bar{x}	0.39	0.04	0.26	0.04	0.34
Ortanca	0.45	0.04	0.28	0.04	0.27
Tepe Değer	0.45	0.04	0.10	0.04	0.23
En Düşük	0.13	0.04	0.04	0.02	0.04
En Yüksek	0.64	0.09	0.59	0.05	0.70

Nihai teste ilişkin g, s ve δ parametrelerinin ortalaması ise sırasıyla 0.39, 0.26 ve 0.34 olarak hesaplanmıştır. Bu değerler ile nihai test sonuçları karşılaştırıldığında maddelerin parametreleri örnekleme göre değişiklik gösterse de ortalama değerlerin birbirine çok yakın olduğu görülmektedir.

DINA modelle geliştirilen test yok denecek kadar azdır. Genellikle çalışmalar mevcut testlerdeki maddelere ait Q matris belirlenerek yapılmaktadır. Nihai teste ilişkin madde parametrelerinin karşılaştırılabilir olmasını sağlamak için yapılan bir çalışmadan örnek vermek yararlı olacaktır Lee, Park ve Taylan (2011) yaptıkları çalışmada TIMSS maddelerinden 25 madde belirleyerek Minnesota örnekleminde DINA modele göre analiz etmişlerdir. Analiz sonucunda maddelerin ortalama g değerleri 0.48, ortalama s değerleri 0.28 ve ortalama δ değerleri 0.24 olarak hesaplanmıştır. Bu durum dikkate alındığında nihai teste ilişkin madde parametrelerin daha iyi olduğu görülmektedir.

3.3.3. Nihai Teste İlişkin Örtük Sınıf Bilgileri

DINA modelde madde parametrelerinin yanı sıra örtük özelliklerin örnekleme görülme sıklıklarına, örtük sınıfların sonsal dağılımlarına ve her bir bireyin ait olduğu örtük sınıf bilgisine ulaşılır. Test kapsamında belirlenen dört özellikten her birinin örnekleme gözlenme sıklığı Tablo 15.'te verilmiştir.

Tablo 15.
Nihai Testte Özelliklerin Gözlenme Olasılıkları

Özellikler	Gözlenme Olasılıkları
α_1	0.51
α_2	0.56
α_3	0.73
α_4	0.51

Tablo 15. incelendiğinde α_1 özelliğinin gözlenme sıklığının 0.51 olduğu anlaşılmaktadır. Bu durumda grubun %51'i α_1 özelliğine sahiptir. Aynı şekilde α_2 özelliğinin görülme sıklığı 0.56, α_3 özelliğinin görülme sıklığı 0.73 ve α_4 özelliğinin görülme sıklığı ise 0.51 olarak belirlenmiştir. Bu durumda grupta en çok α_3 özelliğine sahip olma oranı yüksektir. En çok gözlenen ikinci özellik α_2 özelliğidir, α_1 ve α_4 özelliklerine sahip olunma oranları ise birbirine eşittir ve diğer özelliklere göre daha az gözlenmektedir.

Modele ilişkin analiz sonuçlarından bir diğeri ise örtük sınıfların sonsal olasılıklarıdır. Örtük sınıfların sonsal olasılıkları test ve uygulanan grup göz önünde bulundurularak belirlenmekte ve örtük sınıfların oluşma olasılığına ilişkin değerleri ifade etmektedir. Nihai teste ilişkin örtük sınıf ve sonsal olasılıkları Tablo 16'da verilmiştir.

Tablo 16.
Örtük Sınıfların Sonsal Olasılıkları

Örtük Sınıf	Olasılık Değeri	Örtük Sınıf	Olasılık Değeri
(0000)	0.152	(0110)	0.0362
(1000)	0.0931	(0101)	0.0271
(0100)	0.0373	(0011)	0.0189
(0010)	0.0931	(1110)	0.0411
(0001)	0.0922	(1101)	0.0001
(1100)	0.0355	(1011)	0.0385
(1010)	0.0305	(0111)	0.0268
(1001)	0.0263	(1111)	0.2954

Tablo 16. incelendiğinde grupta (0000) örtük sınıfının gözlenme olasılığı 0.1520'dir. Benzer şekilde en düşük gözlenme olasılığı (1101) örtük sınıfına aittir. Bu örtük sınıfın gözlenme olasılığı 0.0001 olarak hesaplanmıştır. En çok gözlenmesi beklenen örtük sınıf ise 0.2954 olasılık değeri ile (1111) örtük sınıfıdır. Bu değerler sadece grup ve test özelliklerine göre örtük sınıfların gözlenme olasılıklarıdır. Örtük sınıflara ait birey sayıları bu değerlerle doğrudan orantılı olarak gözlenmeyebilir.

BTM'nin en önemli işlevi öğrencilerin profillerini belirlemek olduğundan analiz sonucunda her bir öğrencinin ait olduğu örtük sınıf bilgisine ulaşılır. Tablo 17.'de ise her bir örtük sınıfa düşen birey sayısı verilmiştir.

Tablo 17.
Örtük Sınıflara Ait Birey Sayısı

Örtük Sınıf	Birey Sayısı (N)	Örtük Sınıf	Birey Sayısı (N)
0 0 0 0	49	0 1 1 0	12
1 0 0 0	2	0 1 0 1	4
0 1 0 0	9	0 0 1 1	2
0 0 1 0	33	1 1 1 0	14
0 0 0 1	23	1 1 0 1	0
1 1 0 0	2	1 0 1 1	2
1 0 1 0	1	0 1 1 1	7
1 0 0 1	0	1 1 1 1	110

Tablo 17. incelendiğinde en çok öğrenciyi barındıran örtük sınıf (1111) sınıfıdır. Nihai testin 270 öğrenciyeye uygulandığı dikkate alındığında grubun yaklaşık % 41'i testte ölçülen özelliklerin hepsine sahiptir. Grupta hiçbir özelliğe sahip olmadığı belirlenen ve (0000) örtük sınıfına dâhil edilen 49 öğrenci bulunmaktadır. Hiçbir özelliğe sahip olmayan öğrenciler grubun yaklaşık %18'ini oluşturmaktadır. Bununla birlikte bazı örtük sınıflarda çok az öğrenci bulunmaktadır. Aynı zamanda grupta (1001) ve (1101) örtük sınıflarına dâhil olan hiçbir öğrenci yoktur. Bu durumun anlamı grupta sadece α_1 ve α_4 özelliği ile sadece α_1 , α_2 ve α_4 özelliğine sahip öğrenci bulunmamaktadır. BTM modellerinin amaçlarından biri de özelliklerin birbiri ile ilişkisini açıklayabilmek ve bilişsel süreçlerin işleyişine ışık tutabilmektedir. Ancak tek bir uygulama ile böyle bir ilişki kurmaya çalışmak yanıltıcıdır. Benzer sonuçların tekrarlandığı varsayılırsa neden bazı örtük sınıfların oluşmadığı ve bazılarının gözlenme olasılıklarının daha yüksek olduğu araştırılabilir. Özellikler arasındaki bağ kurulabilir ve böylece öğrencilerin bu özelliklere nasıl sahip olduğu ve özellikler arasındaki ilişki ortaya çıkartılabilir.

3.3.4. Nihai Teste İlişkin Geçerlik ve Güvenirlik Bilgileri

DINA model ile hazırlanan testte uzman görüşü ile konu alanına ait özellikler belirlenir. Çalışmada konu alanına ilişkin özelliklerin belirlenmesinde her biri üç yıldan fazla mesleki tecrübeye sahip beş adet Fen ve Teknoloji öğretmeninin görüşü alınmıştır. Özelliklere ilişkin yazılan maddeler ise üçü ölçme ve değerlendirme uzmanı, dördü ise Fen ve Teknoloji öğretmeni olan yedi kişilik başka bir uzman grup tarafından değerlendirilerek Q matris oluşturulmuştur. Nihai teste seçilen maddelere ilişkin Q matris incelendiğinde belirlenen özelliklerin her birinin temsil edildiği görülmektedir. Testte α_1 özelliği 15 adet, α_2 özelliği 13 adet, α_3 özelliği 16 adet ve α_4 özelliği 13 adet madde tarafından temsil edilmiştir. DINA modele göre testte her bir örtük sınıfı temsil eden maddenin bulunma zorunluluğu yoktur ancak oluşturulan Q matris incelenirse seçilen maddelerin dağılımı neredeyse tüm örtük sınıfları kapsamaktadır. DINA modele göre geliştirilen test belirlenen özellikleri temel aldığından ve uzman görüşü ile hazırlandığından testin kapsam geçerliğini sağladığı ifade edilebilir.

DINA model toplam puan üzerinden değerlendirme yapmadığından ve aynı puana sahip olan öğrencilerin profillerinin farklı olabileceğini ileri sürdüğünden geliştirilen testten elde edilen bilgiler KTK'ye göre elde edilen bilgilerle karşılaştırılmamıştır. Öğrencilerin dönem veya yılsonu notları, ara sınav ya da SBS puanları DINA modele göre geliştirilen bir test için dış geçerlik ölçütü olarak kullanılamaz.

DINA modelde test geliştirme çalışmaları henüz oldukça sınırlıdır. Bu durum geliştirilen teste ait ölçüt geçerliğini sınırlandırmaktadır. Yapılan çalışmaların çoğu simülasyon çalışmalarından veya daha önce geliştirilmiş testlere ait Q-matris oluşturularak yapılmış çalışmalardan oluşmaktadır. Gerçek uygulama verisine dayanan araştırmaların madde parametreleri ile geliştirilen teste ait parametreler karşılaştırılmıştır. Zhang (2006) TIMSS maddelerini kullanarak oluşturduğu 14 maddelik testi DINA model ile analiz etmiştir. Analiz sonucunda maddelere ait ortalama g değerleri 0.44, ortalama s değerleri 0.23 ve ortalama δ değerleri 0.33 olarak hesaplanmıştır. Başokçu (2011) yaptığı çalışmada önceden geliştirilmiş ve lisans öğrencilerine yönelik ölçme ve değerlendirme testine ait maddeleri uzman görüşü ile birlikte değerlendirmiştir. Testte yer alan özellikler belirlenerek teste ait Q-matris hazırlanmıştır. 50 maddeden oluşan testin DINA modele göre analizi sonucunda maddelere ait g değerlerinin ortalaması 0.52, s değerlerinin ortalaması 0.25 ve δ değerlerinin ortalaması ise 0.23 olarak hesaplanmıştır. Lee ve diğerler (2011) ise TIMSS maddelerinin içinden 25 madde belirleyerek Q matris oluşturmuşlardır ve maddeleri Minnesota örnekleminde DINA modele göre analiz etmişlerdir. Analiz sonucunda maddelerin ortalama g değerleri 0.48, ortalama s değerleri 0.28 ve ortalama δ değerleri 0.24 olarak hesaplanmıştır. Yapılan çalışmalara ait madde parametreleri incelendiğinde geliştirilen teste ait maddelerin parametre değerlerinin iyi çalıştığı görülmektedir. Nihai teste ilişkin g , s ve δ parametrelerinin ortalaması ise sırasıyla 0.39, 0.26 ve 0.34 olarak hesaplanmıştır.

Bir ölçme aracının ve bu araçtan elde edilecek sonuçların kabul görebilmesi için ölçme aracının geçerli olmasının yanı sıra güvenilir olması da gerekmektedir. DINA modele göre analiz yapıldığında teste ait bir güvenilirlik indeksi elde edilmemektedir. Bu nedenle teste ait bir iç tutarlık katsayısı olan KR-20 hesaplanmıştır. Teste ait KR-20 değeri 0.82'dir. Hesaplanan iç tutarlılık katsayısı için de genel kabul en az .70 olmasıdır (Nunnally, 1978; Nitko, 2001). Teste ait KR-20 değeri 0.70 değerinden yüksektir ve bu durum testin güvenilirliğine ilişkin bir kanıt oluşturmaktadır.

4.TARTIŞMA ve SONUÇ

Çalışmada araştırmacı tarafından uzman görüşleri yardımıyla Fen ve Teknoloji dersi kapsamında Potansiyel ve Kinetik Enerji konu alanına yönelik test geliştirilmiştir. Testin deneme uygulaması 65 maddeden oluşmaktadır ve 504 öğrenciye uygulanmıştır. Deneme uygulamasının yapıldığı çalışma grubunun %51'ini 7. sınıf öğrencileri, %49'unu ise 8. Sınıf öğrencileri oluşturmaktadır. Testin DINA modele göre analiz edilmesiyle madde parametrelerine ulaşılmıştır. Deneme uygulaması için maddelere ait g parametrelerinin aritmetik ortalaması 0.42, ortanca değeri olarak hesaplanmıştır. Maddelerin s parametrelerinin ortalaması 0.30 iken δ parametrelerine ait değerlerinin ortalaması ise 0.28 olarak hesaplanmıştır.

DINA modelde madde ayırt ediciliği genellikle δ parametresi ile belirlense de maddelere ait s parametresinin de ayırt edicilik için kullanılması gerektiğini vurgulayan çalışmalar vardır. Aynı zamanda madde seçiminde g parametresinin de önemli olduğu belirtilmiştir (Li, 2008). Bu nedenle deneme uygulamasındaki verilere dayanarak g , s ve δ parametrelerini dikkate alan ve 25 maddeden oluşan üç ayrı test oluşturulmuştur. Testlerin her birinde dört örtük özelliğin de yeterli sayıda temsil edildiği gözlenmiştir. Her bir test deneme uygulamasına ait veri üzerinden analiz edilmiştir. Testlerin model veri uyum indeksleri ve madde veri uyum indeksleri karşılaştırıldığında en iyi değerlere sahip olan testin maddelere ait δ parametreleri dikkate alınarak oluşturulan test olduğu

belirtilmiştir. Her üç teste ait madde parametreleri de incelenmiştir. Madde parametrelerinin ortalama değerleri dikkate alındığında en iyi değerler sahip olan testin yine δ parametreleri dikkate alınarak oluşturulan test olduğu gözlenmiştir. Bu nedenle nihai test için δ_j parametreleri dikkate alınarak oluşturulan testin kullanılmasına karar verilmiştir.

Nihai test için seçilen maddeler genel olarak kolaydan zora doğru sıralanmıştır. Test 270 kişilik bir öğrenci grubuna uygulanmıştır ve bu grubun %52'sini 7. Sınıf öğrencileri, %48'ini ise 8. Sınıf öğrencileri oluşturmaktadır. Bu anlamda deneme uygulamasına benzer bir grup üzerinde çalışılmıştır. Maddelerin yeri değiştiğinden Q-matris yeniden düzenlenerek test verileri analiz edilmiştir. Analiz sonucunda teste ait model veri uyum indeksleri ile madde veri uyum indekslerine ulaşılmıştır.

Test kapsamında belirlenen dört özellikten her birinin örnekleme gözlenme sıklığı incelendiğinde α_1 özelliğinin gözlenme sıklığı 0.51, α_2 özelliğinin görülme sıklığı 0.56, α_3 özelliğinin görülme sıklığı 0.73 ve α_4 özelliğinin görülme sıklığı ise 0.51 olarak belirlenmiştir. Analiz sonucunda her bir öğrencinin ait olduğu örtük sınıf bilgisine ulaşılmıştır. En çok öğrenciyi barındıran örtük sınıf (1111) sınıfıdır. Grubun yaklaşık % 41'i teste ölçülen özelliklerin hepsine sahiptir. Grupta hiçbir özelliğe sahip olmadığı belirlenen ve (0000) örtük sınıfına dâhil edilen 49 öğrenci bulunmaktadır. Hiçbir özelliğe sahip olmayan öğrenciler grubun yaklaşık %18'ini oluşturmaktadır. Bazı örtük sınıflarda ise çok az öğrenci bulunmaktadır. Nihai uygulamada (1001) ve (1101) örtük sınıflarına dâhil olan hiçbir öğrenciye rastlanmamıştır.

Nihai teste ilişkin g , s ve δ parametrelerinin ortalaması ise sırasıyla 0.39, 0.26 ve 0.34 olarak hesaplanmıştır. Daha önce Zhang (2006), Başokçu (2011), Lee ve diğerleri (2011) tarafından yapılan çalışmalar incelenmiş ve nihai teste ilişkin parametrelerle karşılaştırılmıştır. Bu karşılaştırma sonucunda nihai teste ait madde parametreleri ortalamalarının diğer çalışmalardaki değerlerden daha iyi olduğu gözlenmiştir.

Testte yoklanan özellikler ile bu özellikleri temsil eden madde sayıları incelenmiştir. α_1 özelliğinin 15 adet, α_2 özelliği 13 adet, α_3 özelliğinin 16 adet ve α_4 özelliğinin 13 adet madde tarafından temsil edildiği belirlenmiştir. DINA modele göre geliştirilen bir test konu alanı ile ilgili belirlenen özellikleri temel alarak geliştirildiğinden ve uzman görüşü ile hazırlandığından testin kapsam geçerliğini sağladığı ifade edilmiştir. Teste ilişkin bir iç tutarlık katsayısı olan KR-20 hesaplanmıştır. Teste ait KR-20 değeri 0.82'di Teste ait KR-20 değeri 0.70 değerinden yüksektir ve bu durum testin güvenilirliğine ilişkin bir kanıt oluşturmaktadır.

KAYNAKÇA

- Başokçu, T. O. (2011). *Bağıl ve mutlak değerlendirme ile dina modele göre yapılan sınıflamaların geçerliğinin karşılaştırılması*. Yayınlanmamış Doktora Tezi, Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara.
- Cheng Y. ve 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*, Haziran 7.
- de la Torre, J. (2008). An empirically-based method of Q-matrix validation for the DINA model: Development and applications. *Journal of Educational Measurement*, 45, 343–362.
- de la Torre, J. ve Douglas, J. A. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika* 69, 333–353
- de la Torre, J. ve Lee, Y. S. (2010). A note on Invariance of the DINA Model Parameters, *Journal of Educational Measurement*, 47(1), 115-127
- Fischer, G. H. (1973). The linear logistic model as an instrument in educational research. *Acta Psychologica* 37, 359–374.
- Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement*, 26, 333-352.
- Henson, R. (2004). *Test discrimination and test construction for cognitive diagnostic models*. Yayınlanmamış doktora tezi, University of Illinois.
- Jang, E. E. (2008). A framework for cognitive diagnostic assessment. *Natural language processing for diagnostic language assessment* (pp. 117-131).
- Junker, B. W. ve Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25(3), 258-272.
- Kato, K. (2009). *Improving efficiency of cognitive diagnosis by using diagnostic items and adaptive testing*. Yayınlanmamış Doktora Tezi, The University of Minnesota.
- Lee, Y., Park, Y. ve Taylan, D. (2011). A cognitive diagnostic modeling of attribute mastery in massachusetts, minnesota, and the u.s. national sample using the TIMSS 2007. *International Journal of Testing*, 11: 144–177, 2011
- Leighton, J. P. ve Gierl M. J. (2007). *Why cognitive diagnostic assessment?* Leighton, J. P. Gierl M. J. (Editörler). Cognitive diagnostic assessment for education. New York: Cambridge University Press.
- Li, D., ve Oranje, A. (2007). *Estimation of standard error of regression effects in latent regression models using Binder's linearization* (ETS Research Rep. No. RR-07-09). Princeton, NJ: ETS.
- 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.

- 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.
- Nitko, A. J. (2001). *Educational assessment of students*. NJ: Merrill: Uper Saddle River.
- No Child Left Behind Act of 2001, Pub. L. No. 107{110, 115 Stat. 1425 (2002).
- Nunnally , J. (1 978). *Psychometric theory*. New York: McGraw-Hill.
- Rupp, A. A. ve Templin, J. (2008). The effects of q-matrix misspecification on parameter estimates and classification accuracy in the DINA Model. *Educational and Psychological Measurement*, 68. No 1, ss;78-96.
- Nunnally , J. (1978). *Psychometric Theory*. New York: McGraw-Hill.
- Rupp, A. A. ve Templin, J. (2008). The effects of q-matrix misspecification on parameter estimates and classification accuracy in the DINA Model. *Educational and Psychological Measurement*, 68. No 1, ss;78-96.
- Tatsuoka, K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement* 20, 345–354.
- Tatsuoka, K. (1991). *Boolean algebra applied to determination of universal set of knowledge states*. Araştırma Raporu ONR-1. Princeton NJ: Educational Testing Service
- Tatsuoka, K. (1995). *Architecture of knowledge structures and cognitive diagnosis: A statistical pattern recognition and classification approach*. In P.D. Nichols, S. F.
- Zhang, W. (2006). *Detecting differential item functioning using the DINA Model*. Yayınlanmamış Doktora Tezi. The University of North Carolina at Greensboro.

EXTENDED ABSTRACT

1. Introduction

This study aimed to determine the psychometric properties of the test and to contribute to the theoretical basis of the model by conducting the test development according to the DINA model. In line with this general aim, it was objectived to examine the test development process in detail with the DINA model, analyze the results, make inferences about the DINA model parameters and determine the ability estimation of the students.

2. Method

Within the scope of cognitive diagnostic models, the Potential and Kinetic Energy topic within the 7th grade Science and Technology course was selected for the achievement test developed in the direction of the items prepared by field experts according to the DINA model. The fact that the subject area is included in the 7th and 8th grade Science and Technology curriculum and it is an important subject that is in large-scale exams ensured that the test is applicable to both grade levels. Each of the 7th and 8th grades consists of 11 classes. The developed test participated in 504 trial applications and 270 students in the final application.

The developed achievement test aims to determine the attributes of the students in the subject area and the latent classes which they belong. The opinions of the five experts were consulted to identify attributes within the concept of the subject area. The expert group consists of the teachers of science and technology who work in the primary schools affiliated to the Ministry of National Education for more than three years. The four basic attributes that students should have in relation to the subject area were determined by the common decision of the expert group, and these attributes are listed as follows;

1. The role of mass in potential and kinetic energy (α_1 , mass)
2. The role of height in potential and kinetic energy (α_2 , height)
3. The role of velocity in potential and kinetic energy (α_3 , velocity)
4. The role of friction force in potential and kinetic energy (α_4 , friction force)

It is expected that students with the four basic characteristics determined will be able to master the topic of potential and kinetic energy. A total of 16 combinations are available for these four attributes. The 15 latent classes are determined because it is not possible for an item that does not have any attributes to be included in the test. When the number of latent classes is taken into consideration, it is planned to write five items belonging to each latent class. The five-person expert group has written items for each latent class. As there are 15 latent classes in the test, the item pool consists of 75 items. It was found appropriate to use a four-option multiple choice item type with considering the grade level.

Q- Matrix Development: Expert opinion has been consulted in order to link the attributes with the items. The five-person expert group consists of two measurement and assessment experts with doctoral degree and three instructors of science teaching. The expert group consulted for determining the Q-matrix and the expert group that wrote the items were independent of each other. Each expert was asked to form a Q matrix by examining the test consisting of 75 items. By examining 5 Q matrices generated, 65 items that experts agreed on items-attributed relations were selected. The selected items were re-examined by three measurement and evaluation experts in terms of measurement

principles and scientific accuracy, and necessary corrections were made by a Turkish teacher regarding language and expression.

3. Findings and Discussion

As a result of the trial application, when the g parameters of the items were examined the mean of the g parameters was calculated as 0.42, the median value as 0.42 and the mod value as 0.41. The item with the lowest g parameter is item 62, and its value was determined as $g = 0.13$. The item with the highest g parameter is item 61 and its value was determined as $g = 0.63$. In this case, g parameter values range from 0.13 to 0.63. When the Q-matrix is examined, it is the latent class (1 1 1 0) belonging to the item with the highest g parameter, and in this case, the item is related to the properties α_1 , α_2 , α_3 . Having a high g parameter of the item means that the probability of correct answer of the item by the students who do not have the attributes of α_1 , α_2 and α_3 is high. The item with the lowest g parameter value is related to the α_4 attribute and is included in the latent class (0 0 0 1). The low value of the parameter g indicates that the probability of correct answer is low by the student who does not have the attributes α_4 . At the same time it also states that the item is not related to any other attributes other than the one required for the correct answer. This statement can be acceptable as an indication that the Q-matrix is highly accurate.

The mean of the s_j parameters was calculated as 0.30, median value 0.26, and mod value 0.18. The item with the lowest s parameter is item 22 and its value was determined as $s = 0.10$. The item with the highest s parameter is the item17 and the value obtained was determined as $s = 0.71$. In this case, the lowest value of s_j parameter is 0.10 while the highest value is 0.71. The item with the lowest s parameter value represents the latent class (0 1 1 0) and thereby is related to the α_2 and α_3 attributes. In this case, students with α_2 and α_3 attributes are more likely to answer correctly to the item. The item with the highest s parameter (1 0 1 1) belongs to the latent class and is related to the α_1 , α_3 and α_4 attributes. The high s parameter implies a high likelihood of false answers by students with α_1 , α_3 , and α_4 attributes.

When the δ_j values of the trial application were examined, the mean of these values was calculated as 0.28, the median value as 0.30 and the mode value as 0.26. The lowest value for the δ_j parameter was -0.11 while the highest value was determined as 0.74. As the value of the item in the DINA model increases, the discriminating power increases. However, when the DINA model and CTT item discrimination values are considered, there are differences between the parameters. According to the DINA model, the item with a δ value of 0.74 and the highest discriminating power; according to the CTT the item discrimination index (r_{jx}) was calculated to be 0.27 and the item test correlation was calculated to be 0.26. Similarly, according to the DINA model parameters, the lowest discrimination value with -0.11; according to the CTT item discrimination index was 0.47 and item test correlation was 0.39. Similar findings were also encountered when the situation related to item difficulty was examined. For an item with low g (0.13) and high s (0.55) parameter values, the item difficulty index value was calculated as 0.46 according to the CTT. The item can be interpreted as one of the difficult items according to the DINA model, but it is a medium difficulty item according to the CTT. In this case, it can be stated that the item statistics of DINA model and CTT differ and do not completely overlap their meanings.

It is necessary to select the item for the final test after the trial testing. The parameters of the DINA model will be taken into consideration when selecting the item, but there is no definite information on which parameter should be referenced in the literature. The best 25 items were determined for each according to the values of g , s and δ from the DINA model parameters. Thus, three different tests were developed taking into account the values of g , s and δ . When the items are examined, 13 of the items included in the two tests formed according to the parameters g and δ are common. It was seen that 18 of the items included in the two tests formed according to the parameters s and δ were common when the determined items were examined. The three tests were repeatedly applied and parameters were compared. When the means of the error parameters for all three tests are taken into consideration, it is seen that the test with the best item parameter values is the test which is made by considering the δ parameter. According to other tests, the mean of g and s parameters is low and the means of δ parameters are high of the test. At the same time, in the model data fit statistics, the test that is produced according to the δ parameter which has the highest compatibility with the model. It has been decided that the test to be used for the final application should be made according to the δ parameter, which has been proposed by most of the studies in the literature to take the δ parameter into account. Thus, 25 items with the highest δ value in the trial test were combined and the final test developed. The items in the final test were applied to the students from easy to difficult, based on DINA model parameters. The fact that the g parameter of an item according to the DINA model is high means that the possibility of correct answered of the item through estimation is also high. In this case, it can be pointed out that the items with high g parameter are easier. The final form of the test was applied to a group of 270 students.

When the item parameters for the final test are examined, it is seen that the mean of the g parameters is 0.39. The mean value of the g parameters can be interpreted fairly well since there is a 0.25 chance that anyone who has no idea of the item can predict the correct answer for a four-option item. The median and mod values belong to the g parameter values of the items were found to be 0.45. In the test, the item with the lowest g parameter has 0.13 value while the item with the highest g parameter has 0.64 value. The fact that the g parameters for the items are low makes it possible to assert that the item is correctly identified in the Q-matrix and provides a proof related to Q-matrix validity. When the s parameters of the items were examined, it was observed that values changed between 0.04 and 0.59. The mean, median and mode of s parameters were determined as 0.26, 0.28 and 0.10 respectively. As it is stated that the s parameter means the accurate answer to an item by individuals who has the attributes belonging to that item, it is also related to the discriminatory power of the items. When considered the parameter value is between 0 and 1, it can be stated that the mean s value for the test is quite good. When δ parameters are analyzed, the average δ values of the items in the test are observed as 0.34. At the same time, the parameter has a median of 0.27 and a mode of 0.23 value. As the δ parameter ranges from -1 to 1 theoretically, it is stated that the discriminatory power of the items with the value over 0.30 is high. The average of δ parameters of the items is 0.34, and accordingly it may imply that the discriminatory power of the test is high.

The occurrence frequency of latent attributes in the sample, posterior distributions of latent classes and latent class of each individual are obtained in DINA model. When the findings are examined, it is understood that the frequency of observation of α_1 is 0.51. In this case, 51% of the group has the α_1 attribute. Similarly, the frequency of occurrence of the α_2 attribute was determined as 0.56, the frequency of the α_3 attribute was

determined as 0.73, and the frequency of the α_4 attribute was determined as 0.51. The latent class including the most students is class (1111) with a value of 0.41. In this case, about 41% of the group has all of the attributes measured in the test. There are 49 students in the group who are determined to have no attributes and are included in the (0000) latent class. It constitutes approximately 18% of the group of students who have no attributes. However, there are very few students in some latent classes. At the same time, there are no students in the group included in the (1001) and (1101) latent classes. This situation' means that in the meaning group there are not only students with the attributes of α_1 , α_2 and α_4 with only α_1 and α_4 attributes.

One of the aims of CDM models is to be able to explain the relation of the attributes and the way of learning and to shed light on the functioning of the cognitive processes. For instance, if there are no students with only α_1 and α_4 attributes in the group, this situation can give an idea that these attributes are learned together with other attributes to the researchers. However, it is misleading to try to establish such a relationship with a single application. Supposing that the similar results are repeated, it can be investigated why some latent classes do not occur and the probability of observation of some is higher. The attributes can be linked so that students can find out how they have these attributes and how they relate to attributes, consecutiveness, and so on.