

Matematik Sorularının Bilişsel Tanı Modeli ile İncelenmesi: Bir Q-Matriks Çalışması

Aygıl TAKIR, Atatürk Öğretmen Akademisi, ORCID ID: 0000-0003-3042-7585

Hasan ÖZDER, Atatürk Öğretmen Akademisi, ORCID ID: 0000-0003-1094-3590

Osman CANKOY, Orta Doğu Teknik Üniversitesi Kuzey Kıbrıs Kampüsü, ORCID ID: 0000-0002-4765-9297

Öne Çıkanlar

- Bu çalışma, Kuzey Kıbrıs'ta uygulanan Kolej Giriş Sınavı (KGS) matematik alt testinin bilişsel yapısını, DINA modeli ile incelemiştir.
- Maddeler, ölçülmek istenen becerilerle iyi uyum sağlamış ve test yapısı geçerlidir.
- Oluşturulan Q matris, testin bilişsel yapısını doğru yansıtmaktadır.
- DINA modeline dayalı betimsel istatistikler, öğrencilerin farklı bilişsel özelliklerdeki yeterlik düzeylerinin değişkenlik gösterdiğini ortaya koymuştur.

Öz

Öğrencilerin bilişsel becerilerinin doğru ve kapsamlı bir şekilde değerlendirilmesi, hem bireysel öğrenme süreçlerinin desteklenmesi hem de eğitim sisteminin kalitesinin artırılması açısından büyük önem taşımaktadır. Bu doğrultuda, eğitim politikalarının geliştirilmesi ve öğretim programlarının iyileştirilmesi için etkili değerlendirme araçlarının kullanılması gerekmektedir. Bu çalışma, 2021 yılında 1.833 beşinci sınıf ilkökul öğrencisine uygulanan KGS'deki matematik alt testinin psikometrik özelliklerini ortaya koymayı amaçlamaktadır. Analizlerde, DINA bilişsel tanı modeli tercih edilmiştir. DINA modelinin teorik altyapısı, diğer modellere kıyasla hem basit hem de sağlam bir yapı sunmaktadır. Testte yer alan 23 matematik maddesi için, uzman görüşlerine dayanarak 15 bilişsel özellik içeren bir Q matrisi oluşturulmuştur. Elde edilen bulgular, g ve s parametrelerinin ortalama değerlerinin sırasıyla 0,25 ve 0,13 olduğunu göstermektedir. Ayrıca, beşinci sınıf öğrencilerinin δ değerlerinin ortalamasının yüksek olduğu (0,73) belirlenmiştir. Beceri örüntülerine göre, öğrencilerin %13'ünün KGS'de ölçülen tüm özelliklere hâkim olduğu tespit edilmiştir. Q matrisin hazırlanması, bilişsel tanı modellerinin etkin kullanımı açısından kritik bir aşamadır. Bu çalışmada bilişsel özelliklerin ve Q matrisin oluşturulmasında uzman görüşlerinden yararlanılmıştır. Q matrisi değerlendirmek amacıyla kullanılan (1-s) değerleri, tüm maddeler için yüksek bulunmuştur. Bu durum, kazanımlar ile madde başarıları arasındaki ilişkinin büyük ölçüde doğru kurulduğunu göstermekte olup, Q matrisinin geçerliğine güçlü bir destek sağlamaktadır.

Anahtar Kelimeler: Bilişsel Tanı Modeli, Kolej Giriş Sınavı, DINA, Matematik Değerlendirmesi.

Önerilen Atıf

Takir, A., Cankoy, O. & Özder, H. (2025). Matematik Sorularının Bilişsel Tanı Modeli ile İncelenmesi: Bir Q-Matriks çalışması. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 26(2), 759-796. DOI: 10.17679/inuefd.1594482



İnönü Üniversitesi
Eğitim Fakültesi Dergisi
Cilt 26, Sayı 2, 2025
ss. 759-796
[DOI](#)
10.17679/inuefd.1594482

Makale Türü
Araştırma Makalesi

Gönderim Tarihi
05.12.2024

Kabul Tarihi
03.06.2025

1. Giriş

Matematikte öğrencilerin performansını etkileyebilecek birçok bilişsel beceri bulunmaktadır. Bunlar arasında problemi tanımlama, problemi sadeleştirme ve temsilleştirme, örüntü ve ilişkileri belirleme ile hesaplama yapma gibi beceriler sayılabilir (Burkhardt & Swan, 2012). Bu bilişsel beceriler dört alana ayrılabilir: temel sayılar (sayıları sayı doğrusu üzerinde konumlandırma, az sayıdaki nesneyi doğru bir şekilde tahmin etme, sembolleri kullanma vb.), bellekten geri çağırma ve işleme (zihinsel hesaplamalar yapma, teorem ve formülleri hatırlama, işlem basamaklarını anımsama vb.), akıl yürütme (karar verme, genelleme yapma vb.) ve görsel-uzamsal organizasyon (matematiksel nesne temsillerinin uzamsal düzenini anlama ve kullanma, grafik ve tabloları analiz etme vb.) (Karagiannakis ve ark., 2016). Ayrıca, birçok karmaşık matematiksel beceri birden fazla bilişsel alanla ilişkilendirilebilmektedir.

Bu bilişsel alanlara ek olarak, karmaşık bağlantılar kurma, esnek problem çözme stratejileri geliştirme ve mantıklı argümanlar oluşturma gibi üst düzey bilişsel beceriler de matematikte başarı açısından büyük önem taşımaktadır. Bu beceriler, öğrencilerin alışılmadık problemlere yaratıcı biçimde yaklaşımlarını, gerekçeli çıkarımlar yapmalarını ve farklı çözüm yollarının etkililiğini değerlendirmelerini mümkün kılar. Bu durum, üst düzey düşünme becerilerinin hem öğretim programı tasarımına hem de değerlendirme uygulamalarına entegre edilmesinin gerekliliğini ortaya koymaktadır (Hidayat, 2020). Üst düzey bilişsel beceriler öğrenciler arasında önemli düzeyde farklılık gösterdiğinden, bu becerilerin değerlendirilmesi kişiselleştirilmiş bir yaklaşımı gerekli kılmaktadır. Kişiselleştirilmiş değerlendirme yöntemleri yalnızca bireysel öğrencilerin öğrenme profillerini doğru biçimde analiz etmek için değil, aynı zamanda gelişmiş bilişsel becerileri destekleyecek hedefe yönelik öğretim stratejileri geliştirmek açısından da kritik öneme sahiptir. Bu bağlamda, bilişsel tanı modelleri (BTM; Cognitive Diagnostic Models), öğrencilerin üst düzey düşünme becerilerine ilişkin güçlü ve zayıf yönlerini belirlemede değerli bir kuramsal çerçeve sunmakta ve geleneksel değerlendirme yaklaşımlarına kıyasla daha ayrıntılı bilgiler sağlamaktadır.

BTM'ler, öğrencilerin bilişsel becerileri hakkında kapsamlı veriler sunar. Öğrencilerin bilişsel açıdan güçlü ve zayıf yönlerini ortaya koyarak bilgi yapılarının ve bilişsel yeteneklerinin anlaşılmasına katkıda bulunur (Leighton & Gierl, 2007). Psikometrik temelli bu modeller, öğrencileri birbirleriyle kıyaslamaktan ziyade bireysel düzeyde değerlendirir. BTM'ler, belirli bir alanda gerekli olan beceri bileşenlerini ölçerek, öğrencilerin öğrenme süreçlerini desteklemek amacıyla biliş kuramı ile ölçme kuramını bütünleştirir.

Bilişsel Tanı Modelleri (BTM), esas olarak sınıf içi öğretim ve öğrenme süreçlerine odaklanmaktadır (Junker & Sijtsma, 2001; Tatsuoka, 1985). Geleneksel ölçme kuramı, bireylerin bilişsel becerilerini değerlendirme noktasında yeterince ayrıntılı veri sunamamaktadır (Li, Ren & Wei, 2021). Oysa BTM'ler, bilişsel süreçler ile ölçme yaklaşımlarını bir araya getirerek değerlendirme süreçlerine önemli katkılar sağlamaktadır. Bu kuramsal yaklaşım, öğrencilerin bilişsel işleme süreçlerine ilişkin potansiyel özelliklerini ortaya koymanın yanı sıra, bireye özgü hedeflenmiş çözümler ve stratejik yönlendirme imkânı da sunmaktadır. De la Torre (2009), BTM'leri, öğrencilerin ayrıntılı becerileri edinip edinmediklerini değerlendirmek amacıyla geliştirilen gizil değişken modelleri olarak tanımlamaktadır. Başka bir ifadeyle, BTM'ler, bir testteki soruları yanıtlatabilmek için gerekli olan birçok ayrıntılı bilişsel özelliğin varlığını ya da yokluğunu belirlemek üzere tasarlanmıştır. "Özellik" kavramı, bilişsel tanılama temelli değerlendirmelerin merkezinde yer almaktadır (Leighton & Gierl, 2007). Özellikler, belirli hedefler veya bilişsel beceriler olarak tanımlanır. BTM'ler, sınava giren bireyin sahip olduğu ya da olmadığı düşünülen özellik kombinasyonlarını belirlemek amacıyla kullanılabilir (Su, 2013). Özelliğe sahip olma durumu "tam sınıf", sahip olmama durumu ise "yokluk sınıfı" şeklinde adlandırılmakta ve bu durum ikili gizil değişken vektörleriyle temsil edilmektedir. Bu sayede, öğrenciler için yalnızca toplam bir puan elde etmek yerine, hangi özelliklere sahip olduklarını gösteren ayrıntılı bir profil oluşturulabilmektedir. Bu çerçevede, BTM'lerin öğrencilerin özelliklerini değerlendirmeye yönelik kuramsal bir yapı sunduğu söylenebilir.

Bilişsel Tanı Modelleri (BTM), geleneksel değerlendirme yaklaşımlarının sunduğu toplam puanların ötesine geçerek, sınava giren bireylerin belirli alt becerileri ne düzeyde edindikleri ya da edinemedikleri konusunda daha nitelikli geri bildirimler sunmaktadır (DeCarlo, 2010). Bu modeller, bireylerin bilişsel açıdan güçlü ve zayıf yönlerini belirlemeye olanak tanır; elde edilen bu veriler, sınıf içi öğretimin şekillendirilmesinde ve öğrenci öğreniminin desteklenmesinde etkin bir şekilde kullanılabilir. Yeterlik temelli modeller, bireylerin belirli bir alanda yeterli düzeye ulaşmış olup olmadığını belirlemeye yönelik istatistiksel bir çerçeve sunar. Bu modeller, özellikle bir bireyin yeterlilik sınavını başarıyla tamamlayıp tamamlayamayacağını ya da bir sonraki akademik düzeye geçmeye hazır olup olmadığını değerlendirmek gibi yüksek riskli karar süreçlerinde önemli bir rol oynamaktadır. Ancak, günümüzde uygulanan pek çok modern değerlendirme yaklaşımında, test maddeleri yalnızca tek bir beceriyi değil, birden fazla ve ayrıntılı bilişsel özelliğin uygulanmasını gerektirecek şekilde yapılandırılmaktadır. Bu karmaşık yapı, geleneksel modellerinin sınırlılıklarını gözler önüne sermekte ve öğrencilerin bilgi düzeylerinin çok boyutlu doğasını yakalayabilecek daha gelişmiş ve tanısal yaklaşımlara duyulan gereksinimi açıkça ortaya koymaktadır (Zhang ve ark., 2023).

Herhangi bir bilişsel tanı modeli (BTM) kapsamında, madde-özellik ilişkisinin modellenmesini sağlayan Q matrisi (Tatsuoka, 1983) kritik bir yapısal bileşendir. Q matrisi, test maddelerinde gözlenen doğru yanıtların arkasında yatan bilişsel özellikleri (genellikle kavramlar ve prosedürler olarak tanımlanan) detaylı bir biçimde tanımlar ve sistematize eder (Clements ve ark., 2020). Bu matris, her bir test maddesi ile bu maddenin doğru çözümü için gereksinim duyulan bilişsel özellikler arasındaki doğrudan ilişkiyi ortaya koymakta; ayrıca, değerlendirme sürecinde yer alan akıl yürütme, içerik bilgisi ve bilişsel işleme bileşenlerinin entegrasyonunu sağlamaktadır (Tatsuoka, 2009). Son dönemde gerçekleştirilen çalışmalarda, özellik tanımlama aşamasındaki potansiyel ön yargıların ve ölçüm hatalarının minimize edilmesi amacıyla, hem uzman görüşlerine hem de veri odaklı algoritmik yöntemlere dayalı hibrit yaklaşımların Q matrisi geliştirilmesinde belirleyici rol oynadığı vurgulanmaktadır (Wang ve ark., 2023).

Q matrisi, değerlendirme sürecinde açıklık ve doğruluk sağlamada kritik bir rol üstlenir. Bu açıklık, öğrencilerin öğrenme ilerlemesini izlemek ve öğretim stratejilerini şekillendirmek için önemli olan araçların geliştirilmesinde temel bir gerekliliktir. Örneğin, biçimlendirici değerlendirmeler, madde geliştirme süreçleri (maddelerin hedeflenen içerik ve becerileri ölçme yeterliliğinin değerlendirilmesi) ve müfredat uyumu analizleri (içerik ve becerilerin belirli sınıf veya ders standartları ile tutarlılığının doğrulanması) bu kapsamda sayılabilir (de la Torre ve ark., 2016). Öğeler ile bilişsel özellikler arasındaki ilişkiyi kurmak için uzmanların özellikleri tanımlaması gerekirken birlikte, güncel gelişmeler, Q matrisinin oluşturulmasında hem uzman görüşlerinin hem de istatistiksel yöntemlerin entegrasyonunun daha isabetli tanısal değerlendirmelere olanak tanıdığını ortaya koymaktadır (Wang ve ark., 2023). Bu bütünsel yaklaşım, Q matrisinin öğrenci bilgisinin çok katmanlı ve karmaşık yapısını daha güvenilir biçimde temsil etmesini sağlamakta ve bilişsel tanı modellerinin geçerliliğini artırmaktadır.

Q matrisinde, satırlar (j) test maddelerini, sütunlar (k) ise ölçülmesi hedeflenen alt becerileri temsil eder; böylece matris, $j \times k$ boyutlarında bir yapıya sahiptir. Bir madde belirli bir bilişsel özelliği ölçüyorsa, ilgili matris hücresine "1" değeri atanırken, o özellik madde için gerekli değilse "0" değeri kullanılır. Her test maddesi, tanımlanmış özelliklerden bir veya birkaçını ölçmek üzere tasarlanmıştır. Q matrisi, test maddeleri ile bilişsel özellikler arasındaki ilişkiyi ortaya koyan bir bağlantı ya da köprü işlevi görür. Q matrisinin doğru ve tutarlı biçimde geliştirilmesi, herhangi bir bilişsel tanı modelinin etkin ve güvenilir uygulanabilmesi için kritik bir aşamadır.

Bir Q matrisi aracılığıyla, tanımlanan k bilişsel özellik temelinde 2^k adet örtük sınıf oluşturulabilir. Örtük sınıflar, bireylerin hangi özelliklere sahip olduğunu veya olmadığını temsil eder ve bu sınıflar, bireylerin doğru yanıt vermesi beklenen maddelerle ilişkilendirilir (Yakar ve ark., 2021). Her bir özellik modeli, farklı yapısal özelliklere sahip ayrı bir örtük sınıf olarak

tanımlanır (de la Torre, Hong & Deng, 2010). Örneğin, $k = 3$ olduğunda oluşan sekiz örtük sınıf Tablo 1’de sunulmaktadır.

Tablo 1.

Örtük Sınıflar ve Anlamları ($k=3$)

Örtük Sınıflar	Bireylerin Özellikleri
(000)	Hiçbir özelliğe sahip olmayanlar
(100), (010), (001)	Bir özelliğe sahip olanlar
(110), (101), (011)	İki özelliğe sahip olanlar
(111)	Üç özelliğe sahip olanlar

Farklı amaçlar doğrultusunda çeşitli çalışmalarda kullanılan çok sayıda BTM bulunmaktadır. Bu modeller, bir maddenin ölçtüğü özellikler arasındaki tamamlayıcı olan ya da tamamlayıcı olmayan ilişkileri yansıtmaya biçimlerine göre birbirinden ayrılır. *Tamamlayıcı olan modellerde*, bir özelliğin gerektirdiği becerilere sahip olma, aynı maddede ölçülen diğer özelliklerin eksikliğini dengeleyebilir. Buna karşılık, *tamamlayıcı olmayan modellerde* ise bir maddeyi doğru yanıtlatabilmek için tüm ilgili özelliklerin edinilmiş olması gerekir (de la Torre, 2011). Tüm BTM’lerde, testi oluşturan özelliklerin doğru şekilde belirlenmesi ve bireylerin bu özelliklere göre doğru şekilde tanılanması için maddeler ile özellikler arasında kesin ve doğru eşleştirmelerin yapılması büyük önem taşımaktadır.

Bazı BTM örnekleri arasında de la Torre (2008) tarafından geliştirilen DINA (Deterministic Input Noisy “And” gate), de la Torre’nin (2011) sunduğu G-DINA (Generalized DINA), Templin ve Henson’ın (2006) geliştirdiği DINO (Deterministic Input Noisy “Or” gate), de la Torre ve Douglas (2004) ile de la Torre ve Minchen’in (2014) önerdiği HO-DINA (Higher-Order DINA) ve Roussos ve arkadaşlarının (2007) geliştirdiği NC-RRUM (non-compensatory reduced re-parameterized unified model) sayılabilir. DINA modeli tamamlayıcı olmayan bir modeldir. Bu, bir alt becerinin kazanılmasının, test maddeleriyle ölçülen başka bir alt becerinin eksikliğini telafi edemeyeceği anlamına gelir (George & Robitzsch, 2015). DINA, sınava girenlerin bir maddeyi etkili bir şekilde çözmek için gerekli tüm becerilere sahip olması gerektiğini varsayan en basit modellerden biridir. Bu nedenle, tamamlayıcı olmayan bir modeldir; gerekli bir özelliğin eksikliği, diğer özelliklerin varlığıyla telafi edilemez (de la Torre & Minchen, 2014). Başka bir deyişle, DINA modelinde bir özelliğin eksikliği, gerekli tüm özelliklerin eksikliğine eş değerdir (Henson & Douglas, 2005).

Sınava girenlerin özellikleri ile madde özellikleri arasındaki etkileşim, örtük (latent) tepki değişkenini tanımlar. Bu değişken aynı zamanda ideal tepki olarak da bilinir (Tatsuoka, 1995) ve şu şekilde ifade edilir:

$$\eta_{ij} = \prod \alpha_{ik}^{q_{jk}}$$

Burada, α_{ik} i’inci sınava girenin k’inci ($k = 1, \dots, K$) özelliği (alt becerisi) kazandığını, q_{jk} ise Q-matrisinde j’inci maddenin k’inci özellikle ilişkilendirilme durumunu gösterir. Yukarıdaki eşitliğe göre, $\eta_{ij} = 1$ durumu yalnızca sınava giren i’nin j’inci madde için gerekli tüm özelliklerde sahip olması durumunda gerçekleşir.

DINA, *kaydırma* (s_j) ve *tahmin* (g_i) parametreleriyle tanımlanır (Doust ve ark., 2021). Kaydırma parametresi (s_j), gerekli tüm özelliklere (becerilere) sahip olan bir bireyin ilgili maddeyi yanlış yanıtlaması durumunu ifade eder. Diğer yandan, tahmin parametresi (g_i), gerekli özelliklere sahip olmayan bireyin maddeyi doğru yanıtlaması durumunu temsil eder. Junker ve Sijtsma’ya (2001) göre, bu parametreler sırasıyla “yanlış negatif” ve “yanlış pozitif” oranlarını yansıtır. Kaydırma ve tahmin parametreleri aşağıdaki şekilde formüle edilir:

$$s_j = P(X_{ij} = 0 | \eta_{ij} = 1),$$

$$g_j = P(X_{ij} = 1 | \eta_{ij} = 0),$$

Doğru yanıt verme olasılığı aşağıdaki şekilde ifade edilir:

$$P(X_{ij} = 1 | \alpha_i, s_j, g_j) = (1 - s_j)^{\eta_{ij}} (g_j)^{1 - \eta_{ij}}$$

Burada, g_j ve s_j tahmin ve kaydırma parametreleridir (Minchen & de la Torre, 2018).

Bir maddeye doğru yanıt verilebilmesi için, sınava giren kişinin kaydırma durumunu önlemek adına gerekli tüm becerilere sahip olması ve tahmin yoluyla doğru yanıt verme olasılığını ortadan kaldıracak şekilde gerekli becerilerden en az birine sahip olmaması gerekir. Kaydırma ve tahmin parametrelerinin modele dâhil edilmediği durumlarda, bir maddeye doğru yanıt verme olasılığı yalnızca 0 ya da 1 olur.

Templin ve Henson (2010), s ve g parametrelerini Tablo 2’de gösterildiği biçimde tanımlamıştır.

Tablo 2.

DINA Modeldeki Tepki Olasılıkları

	$x_{ij}=1$	$x_{ij}=0$
$\eta_{ij} = 1$	$1-s_j$	s_j
$\eta_{ij} = 0$	g_j	$1-g_j$

Madde parametre analizi, her bir madde için bağımsız olarak s ve g indekslerini tahmin etmek için madde ayırt edicilik indeksi (Item Discrimination Index - IDI, δ) içerir ($\delta = 1 - g - s$). δ , g ve s parametrelerinin etkisi olmaksızın, sınava giren bireylerin bir maddeyi ne ölçüde doğru cevaplayabileceğini gösterir.

DINA modeli, bilişsel tanı değerlendirmelerinde birçok önemli avantaj sunmaktadır. En belirgin güçlü yönlerinden biri, öğrencilerin belirli bilişsel özelliklere hâkimiyet durumlarına göre detaylı profillemeye yapabilmesidir. Bu sayede, bir öğrencinin hangi alt becerilere sahip olduğu belirlenerek, öğretmenlere öğrencilerin güçlü ve zayıf yönleri hakkında değerli bilgiler sağlanmakta ve daha hedefe yönelik, etkili öğretim stratejilerinin geliştirilmesine olanak tanınmaktadır. Ayrıca, DINA modeli çoktan seçmeli sorular da dâhil olmak üzere farklı soru türlerini işleyebilme esnekliğine sahiptir; bu da modelin çeşitli değerlendirme biçimlerine uyarlanabilirliğini artırmakta ve uygulama alanını genişletmektedir (Liu & Kang, 2023). Bunun yanında, model parametrelerinin tahmin edilmesini kolaylaştıran ve DINA modelinin araştırmacılar ile uygulayıcılar tarafından daha erişilebilir olmasını sağlayan çeşitli yazılım paketleri mevcuttur (de la Torre ve ark., 2022). Son olarak, DINA modeli, öğrencilerin bilişsel özellikler açısından yeteneklerinin ayrıntılı analizini sunarak, eğitimcilerin bireysel öğrenme ihtiyaçlarını daha derinlemesine kavramalarına ve böylece öğrencilerin öğrenme süreçlerini daha etkin desteklemelerine imkân tanır (Wang ve ark., 2023). Tüm bu avantajlar, DINA modelini hem öğrenci öğrenme çıktıları değerlendirme hem de bu çıktıları iyileştirmek için güçlü bir araç konumuna getirmektedir.

Bilişsel tanı modellerinin (BTM) pratik uygulamaları, dil testi (Li & Suen, 2013; Sawaki ve ark., 2009), psikoloji (Tatsuoka, Varadi & Jaeger, 2012; Templin & Henson, 2006) ve matematik (DeCarlo, 2011; Whitely & Schneider, 1981; Parlak, 2017; Birenbaum ve ark., 2004; Choi, Lee & Park, 2015; Doğan & Tatsuoka, 2008) gibi farklı alanlarda yaygın olarak görülmektedir. Matematik becerileri arasında sıkı ön koşul ilişkilerinin bulunması, bu becerilerin ölçümünde BTM uygulanmasını zorunlu kılmaktadır. Bu bağlamda, eğitimciler, müfredat geliştiriciler ve politika yapımcıların, öğrencilerin matematik alanında neleri bildiklerini ve yapabildiklerini, ayrıca hangi alan, kavram ve konulara daha fazla odaklanılması gerektiğini doğru şekilde anlamaları büyük önem taşımaktadır.

Alan yazınında, TIMSS (Uluslararası Matematik ve Fen Eğilimleri Araştırması) matematik testlerinin bilişsel analizine yönelik çeşitli çalışmalar bulunmaktadır (Birenbaum ve ark., 2004; Doğan & Tatsuoka, 2008; Choi ve ark., 2015; Lee & Park, 2015; Parlak, 2018; Terzi & Şen, 2019; Evran, 2019). Bu çalışmaların bir kısmı, farklı ülkelerdeki öğrencilerin bilişsel profillerini karşılaştırmalı olarak incelemiştir (Birenbaum ve ark., 2004; Doğan & Tatsuoka, 2008; Choi, Lee & Park, 2015). Söz konusu araştırmaların çoğunda, farklı eğitim bağlamlarında bilişsel tanı verilerinin analizinde DINA modeli tercih edilmektedir. Bu modelin, öğrencilerin bilişsel güçlü ve zayıf yönlerine ilişkin ayrıntılı bilgi sağladığı ve geleneksel değerlendirme yöntemlerine kıyasla öğrenci performansının daha derinlemesine anlaşılmasına olanak tanıdığı ortaya konmuştur. Özellikle Choi, Lee ve Park (2015) tarafından gerçekleştirilen bir çalışmada, 2003 yılı TIMSS 8. sınıf matematik değerlendirme verileri DINA modeli aracılığıyla analiz edilmiştir. Çalışmanın temel amacı, öğrenci performansına ilişkin daha spesifik ve eyleme geçirilebilir tanısallık bilgileri elde etmektir. Bulgular, modelin özellikle ABD ve Kore öğrencilerinin performans karşılaştırmalarında değerli tanısallık bilgileri sağladığını ve DINA modeline ait ayırım endeksinin yorumlama süreçlerine katkı sunduğunu göstermiştir.

Gündüz ve Çakan (2020) tarafından gerçekleştirilen bir çalışmada, Türk TIMSS 2015 örnekleminde elde edilen matematik testi verileri kullanılarak öğrencilerin çeşitli özelliklere göre sınıflandırılması incelenmiştir. Araştırmada, DINA modeli ile birlikte DINO modeli de uygulanmıştır. İçerik alanları (sayılar, veri gösterimi, geometrik şekiller ve ölçüler), bilişsel alanın düzeyleri (bilgi, uygulama, akıl yürütme), bu iki alanın birleşimi ve temel bileşen analiziyle elde edilen özelliklere dayalı olarak dört farklı Q matrisi oluşturulmuştur. Söz konusu Q matrisleri kullanılarak gerçekleştirilen sınıflandırmaların doğruluk ve tutarlılıkları analiz edilmiştir. Elde edilen bulgular, içerik alanlarının bilişsel düzeylerle birleştirildiği Q matrisinin sınıflandırma doğruluğu açısından en iyi performansı sağladığını ortaya koymuştur.

Parlak (2023) ise bilgi ve becerilere odaklanarak 8. sınıf öğrencilerinin matematikteki güçlü ve zayıf yönlerini araştırmayı amaçlamıştır. Bu çalışmada, TIMSS 2015 8. sınıf matematik testi öğrenci yanıt kalıpları ile uzmanlar tarafından geliştirilen bir Q matrisi kullanılarak Türkiye ve Singapur'daki öğrencilerin verileri karşılaştırılmıştır. DINA modeli, öğrencilerin genel puan ve sıralamalarından bağımsız olarak güçlü ve zayıf yönlerini değerlendirmek üzere uygulanmıştır. Öne çıkan bulgular arasında, Singapur öğrencilerinin yüksek genel puanlarına rağmen dönüşüm geometrisi alanında zayıf performans sergilemeleri ve Türk öğrencilerin mantıksal akıl yürütme ile açık uçlu görevlerde zorluk yaşamaları yer almaktadır.

Bu çalışmalar, öğrencilerin bilişsel becerilerinin anlaşılmasının etkili öğretim ve öğrenim süreçleri için kritik öneme sahip olduğu özellikle matematik alanında, DINA modelinin eğitim araştırmalarında kullanımının giderek yaygınlaştığını göstermektedir.

1.1. Problem Durumu

Kıbrıs Türk eğitim sisteminde, ilköğretim eğitimi tamamlayan öğrenciler, ortaöğretim aşamasında genellikle Türk Maarif Kolejlerini tercih etmektedir. Bu okullar, aynı çatı altında hem ortaokul hem de lise eğitimi sunmaktadır. Ancak bu okullara kayıt olabilmek için öğrencilerin, beşinci sınıf seviyesindeki ilköğretim öğrencilerine yönelik müfredat temelli performans testi olan Kolej Giriş Sınavı'na (KGS) girmeleri zorunludur. Maarif Kolejlerinde eğitim dili İngilizce olup, devlet tarafından finanse edilmeleri nedeniyle herhangi bir ücret talep edilmemektedir. Lise eğitimi başarıyla tamamlayan öğrenciler, Kuzey Kıbrıs, Türkiye ve diğer ülkelerdeki üniversitelerde eğitimlerine devam edebilmektedir. Söz konusu kolejler, Kuzey Kıbrıs'ta prestijli kurumlar olarak kabul edilmekte ve öğrenciler genellikle 4. sınıftan itibaren uzun ve maliyetli bir hazırlık sürecine başlamaktadır.

KGS, her yıl genellikle beş ay arayla uygulanan iki modülden oluşmaktadır: KGS1 ve KGS2. Öğrencilerin başarıları, her iki modülden alınan puanların ortalaması üzerinden değerlendirilir. Ancak COVID-19 salgını nedeniyle, 2021 yılında sınav yalnızca bir kez gerçekleştirilmiştir. Sınav,

Matematik (27 soru), Türkçe (27 soru), Fen ve Teknoloji (14 soru), Sosyal Bilimler (10 soru) ve İngilizce (22 soru) alt testlerinden oluşmakta olup, her biri dört şıklı çoktan seçmeli ve üç yanıltıcı seçeneğin bulunduğu sorulardan meydana gelmektedir.

Bu çalışma, Bilişsel Tanı Modelleri (BTM), DINA modeli kullanılarak KGS'nin bilişsel yapısını incelemeyi amaçlamaktadır. Başka bir ifadeyle, çalışma KGS matematik alt testinin psikometrik özelliklerinin tanımlanmasında DINA modelinin uygulanma yollarını araştırmaktadır. Matematik alt testiyle ilişkili bilişsel özelliklerin analizi yoluyla, testin öğrencilerin farklı matematik alanlarındaki yeterliliklerini ne ölçüde yansıttığı ortaya konmaktadır. Çalışmada, KGS matematik alt testi için toplam 21 bilişsel özellik belirlenmiştir. Literatürde özellik sayısı konusunda farklı görüşler bulunmakla birlikte, Lee ve arkadaşları (2011) 15 özellik önerirken, Tatsuoka (1984, 1990) en az 8 özellik tanımlanmasının yeterli olduğunu belirtmiştir. 21 özelliğin tamamının kullanılması daha geniş kapsamlı ve zaman alıcı bir araştırma gerektirdiğinden, bu çalışma kapsamında Doğal Sayılar, Doğal Sayılarla Dört İşlem, Doğal Sayılar Problemleri ve Geometri alanlarını kapsayan 15 özellik üzerinde durulmuştur.

Araştırmanın temel amacı, KGS matematik madde parametrelerini tahmin etmek ve 5. sınıf öğrencilerinin KGS matematik testindeki bilişsel özellikler bakımından yeterlilik düzeylerini incelemektir. BTM'ler, pratik test kısıtlamaları içinde çok boyutlu yapıları ölçme imkânı sağlar ve matematik bilgisinin çok spesifik biçimde değerlendirilmesine olanak tanır (de la Torre ve ark., 2016). Bu bağlamda, çalışmada DINA modeli tercih edilmesinin sebebi DINA modelin hem teorik altyapısının sağlam oluşu hem de hem de bilişsel tanı analizinde güvenilir sonuçlar sunmasıdır. Ayrıca, DINA modeli öğrencilerin bilişsel becerilerini ayrıntılı şekilde ortaya koyarak bireysel matematiksel yeterliliklerin değerlendirilmesine imkân tanır. Q matrisi ise, uzman görüşü temel alınarak araştırmacılar tarafından oluşturulmuştur. Uzman görüşü, BTM uygulamalarında Q matrislerinin hazırlanmasında hâlâ en yaygın ve kabul gören yöntemdir. Bu sayede belirlenen bilişsel özelliklerin, amaçlanan öğrenme çıktılarıyla uyumlu olması sağlanmış ve analizlerin matematik programına dayanması güçlendirilmiştir.

Matematikte öğrencilerin öğrenme eksikliklerini belirlemek büyük bir önem taşır. Beceri profili, öğrencilerin alan bazındaki güçlü ve zayıf yönlerini ortaya koyarak, öğretmenlerin bireysel bilişsel ihtiyaçlara yönelik daha etkili ve hedeflenmiş öğretim yöntemleri geliştirmesine olanak sağlar. Daha önce de belirtildiği üzere, Bilişsel Tanı Modelleri (BTM'ler) öğrencilerin öğrenme süreçlerini ve ilerlemelerini detaylı biçimde ölçerek, geleneksel puanlama sistemlerinden çok daha kapsamlı bilgiler sunar (de la Torre, 2009, 2011). Bu bağlamda, bu çalışmanın bulguları öğretmenlerin öğrencilerin çeşitli öğrenme gereksinimlerini daha iyi karşılamak üzere öğretimlerini bireyselleştirmelerine ve böylece matematik başarısını artırmalarına katkı sağlayabilir.

Çalışma, DINA modelini yüksek risk taşıyan sınavlar bağlamında, özellikle Kuzey Kıbrıs'taki 5. sınıf öğrencilerinin girdiği KGS sınavı üzerinden incelemesi nedeniyle önemlidir. KGS, öğrencilerin akademik hayatlarında kritik bir dönemeç olmasının yanı sıra, gelecekteki eğitim fırsatlarını belirlemede de önemli bir rol oynar. Bu araştırma, DINA modelini kullanarak KGS matematik sınavındaki öğrenci performansını analiz etmeyi, sınavda yer alan bilişsel süreçleri detaylı şekilde anlamayı ve mevcut değerlendirme aracının öğrencilerin gerçek yeteneklerini ne ölçüde yansıttığını ortaya koymayı hedeflemektedir.

Araştırmanın önemi, Kuzey Kıbrıs'taki sınav sistemlerinin gelecekteki tasarım ve reform süreçlerine yapacağı katkılarda yatmaktadır. Dünya genelinde eğitim sistemleri, veri temelli ve bireyselleştirilmiş öğrenme yaklaşımlarına doğru ilerlerken, sadece toplam puana dayalı değerlendirmelerin ötesinde, öğrenciye özgü ayrıntılı geri bildirim sunan yaklaşımlara ihtiyaç artmaktadır. DINA modeli, öğrencilerin hangi bilişsel becerilerde güçlü ya da zayıf olduğunu ayrıntılı şekilde ortaya koyarak, eğitimcilere öğrenme boşluklarını hedef alan daha etkili müdahaleler geliştirme imkânı sağlar. Bu sayede, öğrencilerin yeteneklerinin ayrıntılı

değerlendirilmesi, eğitimde daha incelikli yaklaşımların benimsenmesine ve erken yaşlarda öğrenme eksikliklerinin giderilmesine katkıda bulunur. Ayrıca, bu çalışmanın sonuçları yalnızca bireysel sınıflarda değil, eğitim sisteminin tamamında karar alma süreçlerine rehberlik edebilir. Böylelikle, öğrencilerin genel başarısını artıracak daha etkili ve kişiselleştirilmiş eğitim uygulamalarının geliştirilmesine katkı sağlanabilir. Uzun vadede, bu durum politika yapıcıların geleneksel değerlendirme çerçevelerini yeniden gözden geçirerek, daha tanısal ve esnek öğrenci değerlendirme yöntemlerini araştırmalarına olanak tanıyabilir.

2. Yöntem

Bu çalışma, 2021 yılında 1.833 5. sınıf ilköğretim öğrencisine uygulanan KGS (Kolej Giriş Sınavı) matematik alt testinin psikometrik özelliklerini incelemek amacıyla yürütülmüştür. Bu bağlamda, test sonuçları veri kaynağı olarak kullanılmıştır. Ayrıca, Q matrisi oluşturulmasında uzman görüşlerinden yararlanılmıştır.

2.1. Veri Toplama Araçları

Veri toplama amacıyla, 2021 yılında uygulanmış olan KGS matematik testi kullanılmıştır. Söz konusu test, toplamda 17 sayfadan oluşmaktadır. Matematik testine ait soru örnekleri ve bu soruların özellikleri Tablo 3'te sunulmuştur.

Tablo 3.

KGS 2021 Matematik Testindeki Soruların Örnekleri ve Özellikleri

Soru	Öğrenme Çıktıları
Aşağıda verilen işlem şemasına göre, tüm işlemlerin doğru olabilmesi için "K" ile belirtilen kutuya hangi sayı yazılmalıdır? $\begin{array}{r} 12 \\ + \\ K \\ \hline = \\ \begin{array}{r} x \\ 4 \\ \hline = \\ \end{array} \\ - 40 \\ \hline = 80 \end{array}$ A) 2 B) 8 C) 12 D) 18	- Doğal sayılarla toplama ve çıkarma işlemleri yapabilmek. - En fazla üç basamaklı doğal sayılarla çarpma işlemi yapabilmek. - Çarpma ve bölme arasındaki ilişkiyi anlayarak eksik elemanı (çarpan, bölüm veya bölme) bulabilmek.
AA iki basamaklı bir doğal sayıdır. $AA + AA + AA = 132$ ise, $4 + (4 \div A)$ işleminin sonucu kaç eştir? A) 2 B) 3 C) 4 D) 5	- Doğal sayılarla toplama ve çıkarma işlemleri yapabilmek. - Doğal sayıların (dokuz basamağa kadar) basamak ve basamak değerlerini belirleyebilmek. - En fazla üç basamaklı doğal sayılarla çarpma işlemi yapabilmek.
Aşağıda Fatma'nın 4 adımda yapmış olduğu bölme işlemi verilmiştir. Buna göre Fatma'nın, yapmış olduğu bölme işlemi için ne söyleyebiliriz? 1. Adım $\begin{array}{r} 848 \overline{)16} \\ - 80 \\ \hline 4 \end{array}$ 2. Adım $\begin{array}{r} 848 \overline{)16} \\ - 80 \\ \hline 48 \end{array}$ 3. Adım $\begin{array}{r} 848 \overline{)16} \\ - 80 \\ \hline 48 \\ - 32 \\ \hline 16 \end{array}$ 4. Adım $\begin{array}{r} 848 \overline{)16} \\ - 80 \\ \hline 48 \\ - 32 \\ \hline 16 \\ - 16 \\ \hline 0 \end{array}$ A) Bölme işlemini hatasız yapmıştır. B) İlk hatasını 2. adımda yapmıştır. C) İlk hatasını 3. adımda yapmıştır. D) İlk hatasını 4. adımda yapmıştır.	En fazla dört basamaklı bir doğal sayıyı, en fazla iki basamaklı bir doğal sayıya bölebilmek.

2.2. Süreç ve Veri Analizi

Bu çalışma, DINA modelini kullanarak bilişsel tanı modeli oluşturmak için Shi ve arkadaşları (2021) tarafından önerilen adımları takip etmiştir. Bu adımlar, Q-matrisinin değerlendirilmesi, modelin kalibrasyonu, model uyumunun incelenmesi, madde düzeyinde tanılama analizi, sınıflandırma güvenilirliğinin test edilmesi ve elde edilen sonuçların raporlanması ile görselleştirilmesini kapsamaktadır.

2.2.1. Veri ve Q Matris'in Hazırlanması

Q-matris, 23 matematik sorusu ve 15 bilişsel özellik için oluşturulmuştur. DINA modeli analizinde, “doğal sayılar”, “doğal sayılarla dört işlem” ve “doğal sayı problemleri” alanlarındaki öğrenme hedefleri bilişsel özellik olarak esas alınmıştır. Araştırmacılar, testteki her soruyu detaylı şekilde incelemiş ve 5. sınıf matematik müfredatını gözden geçirerek soruların kapsadığı bilişsel özellikleri belirlemiştir. Bu çalışma sonucunda, Q-matrisinin ilk versiyonu hazırlanmıştır.

De la Torre ve Minchen (2014) tarafından belirtildiği üzere, istatistiksel yöntemler Q-matrisin geliştirilmesine katkı sağlasa da atrisin uygunluğunun alan uzmanları tarafından titizlikle değerlendirilmesi gerekmektedir. Bu kapsamda, Q-matrisin ilk versiyonu, matematik öğretimi alanından üç ve ölçme değerlendirme alanından ise bir uzman olmak üzere toplam dört uzmana iletilmiştir.

Uzmanlardan, Q-matrisindeki her bilişsel özellik için yorum ve önerilerde bulunmaları istenmiş; ayrıca, her özelliğin KGS testi kapsamında ne derece ölçüldüğünü derecelendirmeleri istenmiştir. Bazı uzmanlar, dikkate alınması gereken ek özellikler de önermiştir.

Bu aşamanın tamamlanmasının ardından, her bir madde, uzman görüşleri doğrultusunda belirlenen 15 bilişsel özellik ile ilişkilendirilerek Q-matris oluşturulmuştur. Maddeleri temsil eden özellikler ve sıklık bilgileri Tablo 4 ve Tablo 5'te sunulmaktadır.

Table 4.

Özellikler ve Sıklıkları

Özellikler	Sıklıkları
Özellik 1. Doğal sayıyı en yakın yüzlere yuvarlama	1
Özellik 2. En fazla 9 basamaklı doğal sayıları yazma ve okuma	2
Özellik 3. Doğal sayıların basamak değerlerini belirleme (en fazla dokuz basamağa kadar)	3
Özellik 4. Doğal sayılarla toplama ve çıkarma işlemleri yapma	12
Özellik 5. Toplama ve çıkarma işlemlerinin sonuçlarını tahmin etme	2
Özellik 6. En fazla iki basamaklı doğal sayılarla çarpma işlemi yapma	8
Özellik 7. En fazla dört basamaklı bir doğal sayıyı en fazla iki basamaklı bir doğal sayıya bölme	7
Özellik 8. Zihinden çarpma ve bölme işlemleri için uygun strateji kullanma	1
Özellik 9. Bölme işleminde kalanı yorumlama	1
Özellik 10. Çarpma ve bölme arasındaki ilişkiyi anlayarak eksik elemanı (çarpan, bölüm veya bölme) bulma	3
Özellik 11. En fazla iki işlemi içeren parantezli ifadelerin sonucunu bulma	2
Özellik 12. Dört işlemle ilgili metin problemleri çözme	8
Özellik 13. Bir noktanın başka bir noktaya göre konumunu ifade etme	1
Özellik 14. Dikdörtgen, paralelkenar, eşkenar dörtgen ve yamuk temel elemanlarını tanımlama ve çizme	1
Özellik 15. Üçgen ve dörtgenin iç açılarının toplamını belirleme ve eksik açıyı bulma	1

Toplamda 23 maddeden oluşan test için 54 ilişkilendirme yapılmıştır. On madde üç özelliğe, on bir madde iki özelliğe ve iki madde ise bir özelliğe bağlıdır.

Tablo 5'te görüldüğü üzere, Q matrisi her bir maddenin hangi bilişsel özellikleri değerlendirdiğini göstermektedir. Bir hücrede 1 değeri bulunması, ilgili maddenin söz konusu bilişsel özelliği ölçtüğünü; 0 değeri ise ölçmediğini belirtmektedir.

Table 5.

Uzmanlar Tarafından Oluşturulan Q Matrisi

Madde No - Özellik	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0
4	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
7	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0
8	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0
9	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
11	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0
12	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0
13	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
14	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
16	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0
17	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
19	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0
20	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0
21	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
22	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0
23	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0

2.3. Veri Analizi

DINA modeli kullanılarak, Ox Edit programlama ortamında (Ma & de la Torre, 2016) veri analizi gerçekleştirilmiştir. DINA modeli ile analiz yapabilmek için gerekli kodlar programa yüklenmiştir. Bu kodlar, programın Maksimum-Beklenti (EM) algoritması kullanarak DINA model parametrelerinin tahmin edilmesini sağlamaktadır. Araştırmada, s ve g parametreleri farklı sınır değerleri üzerinden değerlendirilmiştir [örneğin, 0,05 - 0,40 (Henson & Douglas, 2005); 0,25 - 0,15 (Rupp & Templin, 2008)]. Eldeki çalışmada, s ve g parametreleri düşük ve yüksek seviyeler olarak yorumlanmış; 0,45 ve üzeri yüksek, 0,25 ve altı ise düşük seviye olarak belirlenmiştir.

2.3.1. Model Veri Uyum İstatistikleri

DINA modelinin veri uyumu, “*Test-Level Fit Statistics*” analizi ile değerlendirilmiştir. Görelî model uyumu ise genellikle geleneksel bilgi temelli ölçütler kullanılarak incelenmiş olup, bu ölçütler arasında AIC (Akaike Bilgi Kriteri; Akaike, 1974) ve BIC (Bayesyen Bilgi Kriteri; Schwarz, 1978) yer almaktadır.

3. Bulgular

Tablo 6, DINA modeli kullanılarak analiz edilen 5. sınıf öğrencilerinin madde parametre tahmin sonuçlarını sunmaktadır.

DINA modelindeki *s* ve *g* parametrelerinin ortalama değerleri sırasıyla 0,25 ve 0,13 olarak belirlenmiştir. Bu parametrelerden *g*’nin ortalama puanı, 5. sınıf öğrencilerinin %25’inin, tüm özelliklere sahip olmadan maddeleri tahmin yoluyla yanıtladığını göstermektedir. Cevap hakkında hiçbir fikri olmayan bir öğrencinin doğru cevabı tahmin etme olasılığı 0,25 olarak hesaplanmıştır. Bu nedenle, *g* parametrelerinin ortalama değeri genel ortalama olarak kabul edilebilir.

Analiz sonucunda, *g* değerlerinin 0 ile 0,387 arasında değişim gösterdiği görülmektedir. Ayrıca, 2, 3, 9, 12 ve 15 numaralı maddelerin *g* parametre değerleri yüksek bulunmuştur. Bu durum, öğrencinin yeterli bilgiye sahip olmamasına rağmen bu maddeleri doğru yanıtlama olasılığının yüksek olduğunu göstermektedir. Yeterli bilgiye sahip olmayan öğrencilerin bu soruları doğru yanıtlayabilme ihtimalinin orta düzeyde olduğu şeklinde yorumlanabilir.

Tablo 6.

Madde Parametreleri

Madde no	<i>g</i>	SE (<i>g</i>)	<i>s</i>	SE (<i>s</i>)	δ	1- <i>s</i>
1	0.02	0.02	0.14	0.01	0.84	0.86
2	0.19	0.02	0.15	0.01	0.66	0.85
3	0.36	0.02	0.17	0.01	0.47	0.83
4	0.29	0.02	0.14	0.01	0.57	0.86
5	0.32	0.02	0.16	0.01	0.51	0.84
6	0.34	0.02	0.16	0.01	0.50	0.84
7	0.26	0.02	0.00	0.02	0.74	1.00
8	0.39	0.02	0.20	0.01	0.41	0.80
9	0.35	0.02	0.19	0.01	0.46	0.81
10	0.34	0.02	0.20	0.01	0.47	0.80
11	0.00	0.03	0.00	0.02	1.00	1.00
12	0.29	0.02	0.00	0.02	0.71	1.00
13	0.00	0.03	0.00	0.02	1.00	1.00
14	0.35	0.02	0.13	0.01	0.53	0.87
15	0.27	0.02	0.20	0.01	0.53	0.80
16	0.00	0.03	0.00	0.02	1.00	1.00
17	0.34	0.02	0.20	0.01	0.46	0.80
18	0.36	0.02	0.19	0.01	0.45	0.81
19	0.28	0.02	0.14	0.01	0.58	0.86
20	0.24	0.02	0.13	0.01	0.63	0.87
21	0.30	0.02	0.17	0.01	0.54	0.83
22	0.36	0.02	0.22	0.01	0.43	0.78
23	0.18	0.02	0.21	0.01	0.61	0.79

s parametresinin her madde için 0 ile 0,22 arasında değiştiği gözlemlenmiştir. En yüksek *s* parametresi değeri 22. maddeden (0,22) elde edilmiştir. Diğer maddelerde ise *s* parametre değerleri genel olarak düşük bulunmuştur. En düşük *s* parametrelerine sahip maddeler 7, 11, 12, 13 ve 16 olarak belirlenmiştir. Q matrisi incelendiğinde, bu maddelerin sırasıyla “doğal sayılarla toplama ve çıkarma işlemleri yapma” ve “en fazla iki basamaklı doğal sayılarla çarpma işlemleri yapma” özellikleriyle daha yoğun şekilde ilişkili olduğu görülmüştür.

De la Torre (2008) tarafından belirtildiği üzere, 1-s değerlerinin sıfıra yakın olması, Q matris tarafından tanımlanan özelliklerin maddeler için tam olarak temsil edilmediğine işaret eder. Bu bağlamda, 1-s değerleri, Q matris ile değerlendirme aracındaki maddeler arasındaki uyum derecesinin göstergesi olarak yorumlanmaktadır. Mevcut analizde, maddeler için 1-s değerleri 0,78 ile 1 arasında değişmiş ve 1-s değerlerinin ortalaması 0,73 olarak bulunmuştur. Bu sonuç, Q matrisinin maddeleri doğru cevaba ulaşmak için gerekli özelliklere uygun şekilde atamasını büyük oranda sağladığını göstermektedir. Tüm parametreler değerlendirildiğinde, en ideal maddelerin 1, 11, 13, 16 ve 23. maddeler olduğu söylenebilir.

3.1. Özelliklerin Olasılıkları

Özellik olasılığı, inceleme katılımcılarının sahip olduğu olasılık özelliklerinin sayısını gösteren “A” ile ifade edilen 2^A şeklinde tanımlanmaktadır. Bu çalışmada 15 özellik bulunduğu için 2^{15} örtük sınıf mevcuttur. Özellik olasılıkları Tablo 7’de sunulmuştur.

Analiz sonuçları, en yüksek özellik kazanım olasılığının “doğal sayılarla toplama ve çıkarma işlemleri yapma”da (%0,83) olduğunu göstermiştir. Bu durum, 5. sınıf öğrencilerinin %83’ünün bu özelliği yüksek düzeyde kazandığını ortaya koymaktadır. En düşük kazanım oranı ise “bir noktanın konumunu başka bir noktaya göre yön ve birim kullanarak ifade etme” özelliğine (%59) aittir.

Tablo 7.

Özelliklerin Olasılıkları

Özellikler	Olasılıkları
Özellik 1. Doğal sayıyı en yakın yüzlere yuvarlama	0.660
Özellik 2. En fazla 9 basamaklı doğal sayıları yazma ve okuma	0.784
Özellik 3. Doğal sayıların basamak değerlerini belirleme (en fazla dokuz basamağa kadar)	0.767
Özellik 4. Doğal sayılarla toplama ve çıkarma işlemleri yapma	0.829
Özellik 5. Toplama ve çıkarma işlemlerinin sonuçlarını tahmin etme	0.624
Özellik 6. En fazla iki basamaklı doğal sayılarla çarpma işlemi yapma	0.771
Özellik 7. En fazla dört basamaklı bir doğal sayıyı en fazla iki basamaklı bir doğal sayıya bölme	0.872
Özellik 8. Zihinden çarpma ve bölme işlemleri için uygun strateji kullanma	0.620
Özellik 9. Bölme işleminde kalanı yorumlama	0.610
Özellik 10. Çarpma ve bölme arasındaki ilişkiyi anlayarak eksik elemanı (çarpan, bölüm veya bölme) bulma	0.685
Özellik 11. En fazla iki işlemi içeren parantezli ifadelerin sonucunu bulma	0.717
Özellik 12. Dört işlemle ilgili metin problemleri çözme	0.665
Özellik 13. Bir noktanın başka bir noktaya göre konumunu ifade etme	0.590
Özellik 14. Dikdörtgen, paralelkenar, eşkenar dörtgen ve yamuk temel elemanlarını tanımlama ve çizme	0.601
Özellik 15. Üçgen ve dörtgenin iç açılarının toplamını belirleme ve eksik açıyı bulma	0.638

3.2. Özelliklerin Sınıf Olasılıkları

Tablo 8, 5. sınıf öğrencilerine ait örnek matematik beceri sınıf profillerini göstermektedir. (111111111111111) biçimindeki özellik deseni, çalışmada yer alan 15 beceriyi temsil etmektedir. Bu çalışmada kullanılan Q matrisi yapısına göre, 2^{15} olası örtük sınıf bulunmaktadır.

Tablo 8.

Örnek Örtük Sınıf Olasılıkları

Örtük Sınıflar															Olasılık
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.13
1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0.05
1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0.03
1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0.03
1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0.02

Beceri sınıf profili, 5. sınıf öğrencilerinin %12,7'sinin tüm becerilere sahip olduğunu göstermektedir. Bu desen, 5. sınıf öğrencilerinin yaklaşık %13'ünün 15 becerinin tamamında yeterli gösterdiğini ortaya koymaktadır. Diğer bir özellik deseni olan (11111110111111) ise öğrencilerin en az bir beceride yeterli olmadığını göstermektedir. Bu desen, 5. sınıf öğrencilerinin yaklaşık %5'inin 14 beceride yeterli gösterdiğini ve yalnızca bir beceride yeterli göstermediğini ortaya koymaktadır.

4. Tartışma & Sonuç

Bu çalışmanın amacı, Kuzey Kıbrıs'ta uygulanan KGS matematik alt testinin bilişsel yapısını, bir BTM olan DINA modeli aracılığıyla incelemektir. Daha önce belirtildiği üzere, çalışmada kullanılan bilişsel özellikler uzman görüşleri doğrultusunda belirlenmiştir. Uzmanlar, KGS'nin program temelli bir ölçme yapısı olmasına rağmen, bazı öğrenme alanlarının (örneğin, ölçme ve veri) testte temsil edilmemesini eleştirmiştir. Bu durum, testte yer alan hiçbir sorunun bu öğrenme alanlarındaki özelliklerle ilişkilendirilmemesiyle sonuçlanmıştır. Söz konusu bulgu, test geliştirme sürecinde kapsam geçerliğine daha fazla önem verilmesi gerektiğine işaret etmektedir. Bu tür bir tutarsızlık, daha kolay değerlendirilebilen içerik alanlarının tercih edilmesi, test geliştirme sürecindeki zaman kısıtlamaları ya da sınıf içi öğretim ile sistem düzeyindeki değerlendirme öncelikleri arasındaki uyumsuzluktan kaynaklanıyor olabilir.

Yapılan analizler, maddelere ait (1-s) değerlerinin oldukça yüksek olduğunu göstermiştir. Bu bulgu, maddelerin belirlenen bilişsel özelliklerle doğru şekilde eşleştirildiğini ve oluşturulan Q matrisin geçerliğini desteklemektedir. DINA modeline dayalı betimsel istatistikler, öğrencilerin farklı bilişsel özelliklerdeki yeterli düzeylerinin değişkenlik gösterdiğini ortaya koymuştur. Örneğin, “doğal sayılarla toplama ve çıkarma işlemi yapma” (Özellik 4), en yüksek yeterli olasılığına sahipken (0.83); “bir noktanın başka bir noktaya göre konumunu yön ve birim kullanarak ifade etme” (Özellik 13), en düşük yeterli olasılığına sahiptir (0.59). Bu farklılıklar; bazı becerilere öğretim sürecinde yeterince odaklanılmaması, bazı bilişsel özelliklerin görece daha karmaşık olması ya da öğrencilerin belirli madde formatlarına aşinalık düzeylerinin düşük olması gibi etkenlerle açıklanabilir.

Bu bulgular, öğrencilerin öğrenme süreçlerine ilişkin ayrıntılı tanısıl bilgiler sunma noktasında BTM'lerin değerini pekiştirmektedir. Matematik disiplini, çoklu bilişsel süreçleri içeren bir alan olup, BTM'ler bu süreçleri geleneksel toplam puanlardan daha ayrıntılı ve anlamlı biçimde değerlendirme imkânı sunmaktadır. Her ne kadar DINA modeli ikili yapıyla (yeterlik var/yok) çalışsa da eğitimsel değerlendirmelerde çok boyutlu yapıları analiz etmek açısından

oldukça işlevsel olduğu görülmektedir. Bununla birlikte, yakın tarihli çalışmalar (örn. Delafontaine ve ark., 2022; Kalkan & Toprak, 2022), sonuçların sağlamlığını artırmak ve model-veri uyumunu daha kapsamlı biçimde incelemek adına birden fazla BTM'nin kullanılmasını önermektedir.

Bir diğer önemli husus ise Q matrisinin geliştirilmesidir. Bu çalışmada Q matrisi uzman görüşüne dayalı olarak oluşturulmuş olsa da literatür bu sürecin karmaşık ve hataya açık olduğunu ve doğruluğu önemli ölçüde etkilediğini vurgulamaktadır (Li ve ark., 2021). Bu karmaşıklık, bilişsel özelliklerin soyut doğasından, uzman yorumlarındaki tutarsızlıklardan ya da maddelerin özelliklerle eşleştirilmesindeki öznel yaklaşımlardan kaynaklanabilir. Güncel araştırmalar, daha güvenilir Q matrisleri oluşturmak için uzman görüşünün istatistiksel iyileştirme yöntemleriyle birleştirilmesini önermektedir. Buna ek olarak, Stone testi ya da "Posterior Predictive Model Checking – PPMC" gibi tahmin yöntemleri, model uyumunun geçerlik analizinde kullanılabilir. Wu ve ark. (2020) ise özellikle yüksek zorluk düzeyine sahip testlerde, bilişsel özelliklerin ayrıntı düzeyinin tanısıl çıkarımların doğruluğunu nasıl etkilediğine dikkat çekmektedir.

Bilişsel Tanı modellerinin, DINA modeli de dâhil olmak üzere, en büyük avantajlarından biri, öğrenci gelişimini ayrıntılı geri bildirimlerle izleyebilme yeteneğidir. Bu modeller, büyük ölçekli değerlendirmelerde kullanıldığında, yıllar içindeki performans desenlerini ortaya çıkarabilir ve öğrencilerin spesifik beceri profillerine dayalı olarak öğretim kararlarının alınmasına olanak sağlar. Bu özellikleriyle BTM'ler, özellikle biçimlendirici değerlendirme ve eğitim politikalarının izlenmesi için oldukça uygundur. Bu bağlamda, DINA modeli yorumlanabilirliği ve sadeliği ile öne çıkarak hem eğitim araştırmacıları hem de uygulayıcılar için pratik bir tercih olarak görülmektedir.

Özetle, bu çalışma DINA modelinin KGS matematik alt testinin ayrıntılı bilişsel tanısıl analizinde sahip olduğu potansiyeli ortaya koymuştur. Sonuçlar, BTM'lerin öğrencilerin güçlü ve zayıf yönlerini özellik düzeyinde belirlemede etkin olduğunu göstermekte olup, bu durum hem öğretim tasarımı hem de politika kararları açısından kritik öneme sahiptir. Güncel literatür aynı zamanda tanısıl doğruluğu ve yorumlanabilirliği artırmak için Q matrisin titizlikle oluşturulması ve uygun model seçimlerinin önemine vurgu yapmaktadır. Bu bulgular, yüksek riskli sınavlar arasında yer alan KGS gibi testlerde BTM'lerin kullanımının, daha hedefe yönelik ve adil eğitim destek stratejilerinin geliştirilmesini teşvik edebileceğini göstermektedir.

Çalışmanın bulguları, Kuzey Kıbrıs Eğitim Sistemi'ne özgü birkaç pratik öneri sunmaktadır. Birincisi, sınav geliştiricileri ve müfredat planlayıcılarının, ulusal matematik müfredatı ile KGS gibi zorlu sınavlar arasındaki uyumu artırmaları gerekmektedir. Müfredatın bazı alanlarında (örneğin ölçme ve veri) soruların bulunmaması, bu alanlarda güçlü olan öğrencilerin dezavantajlı duruma düşmesine yol açabilecek bir içerik geçerliği eksikliğine işaret etmektedir. İkincisi, okullar ve eğitim yetkilileri, BTM gibi tanısıl değerlendirme araçlarını yalnızca yüksek riskli sınavlar için değil, aynı zamanda düzenli sınıf içi değerlendirme süreçlerinin bir parçası olarak kullanmayı düşünebilirler. Bu yaklaşım, öğretmenlerin öğrencilerin belirli bilişsel becerilerdeki yeterliklerini daha iyi izlemelerine ve buna göre öğretim planlama ve uygulamalarını uyarlamalarına olanak tanıyarak, bölgedeki mevcut maliyetli ve yüksek baskılı sınav hazırlık süreçlerine olan bağımlılığı azaltabilir.

Etik Komite Onayı: Bu araştırma için Doğu Akdeniz Üniversitesi Bilimsel Araştırma ve Yayın Etiği Kurulundan (06.12.2022-ETK00-2022-0283) etik izin alınmıştır.

Hakem Değerlendirmesi: Dış bağımsız.

Yazar Katkıları: Yazarlar; bu makalenin araştırılması, yazarlığı ve yayımlanması için eşit düzeyde katkı sağlamışlardır.

Çıkar Çatışması: Yazar(lar); bu makalenin araştırılması, yazarlığı ve yayımlanmasına ilişkin herhangi bir potansiyel çıkar çatışması beyan etmemiştir.

Finansal Destek: Yazar(lar); bu makalenin araştırılması, yazarlığı ve yayımlanması için herhangi bir finansal destek almamıştır.

Yapay Zekâ Kullanımı Bildirimi: Yazar (lar), bu makalenin araştırılması, yazarlığı ve yayınlanması için Grammarly ve DeepL yapay zekâ araç/araçlarından faydalanmıştır.

Kaynakça

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. <https://doi.org/10.1109/tac.1974.1100705>
- Birenbaum, M., Tatsuoka, C., & Yamada, T. (2004). Diagnostic assessment in TIMSS-R: Between countries and within-country comparisons of eighth graders' mathematics performance. *Studies in Educational Evaluation*, 30(2), 151-173. <https://doi.org/10.1016/j.stueduc.2004.06.004>
- Burkhardt, H., & Swan, M. (2012). Designing assessment of performance in mathematics. *Educational Designer*, 2(5), 1-40.
- Choi, K. M., Lee, Y. S., & Park, Y. S. (2015). What CDM can tell about what students have learned: An analysis of TIMSS eighth grade mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(6), 1563-1577. <https://doi.org/10.22034/EMES.2021.244398>
- Clements, D. H., Banse, H., Sarama, J., Tatsuoka, C., Joswick, C., Hudyma, A., ... & Tatsuoka, K. K. (2022). Young children's actions on length measurement tasks: strategies and cognitive attributes. *Mathematical thinking and learning*, 24(3), 181-202. <https://doi.org/10.1080/10986065.2020.1843231>
- DeCarlo, L. T. (2010). On the statistical and theoretical basis of signal detection theory and extensions: Unequal variance, random coefficient, and mixture models. *Journal of Mathematical Psychology*, 54(3), 304–313. <https://doi.org/10.1016/j.jmp.2010.01.001>
- DeCarlo, L.T. (2011). On the analysis of fraction subtraction data: The DINA model, classification, latent class sizes, and the q-matrix. *Applied Psychological Measurement*, 35(1), 8-26. <https://doi.org/10.1177/0146621610377081>
- Delafontaine, A., Tang, C. Y., & De Boeck, P. (2022). Cross-cultural item analysis using cognitive diagnosis models: A Q-matrix approach. *International Journal of Testing*, 22(2), 145–165. <https://doi.org/10.1080/15305058.2021.1972641>
- 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(4), 343–362. <https://doi.org/10.1111/j.1745-3984.2008.00069.x>
- de La Torre, J. (2009). A cognitive diagnosis model for cognitively based multiple-choice options. *Applied Psychological Measurement*, 33(3), 163–183. <https://doi.org/10.1177/0146621608320523>
- de la Torre, J. (2009). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics*, 34(1), 115–130. <https://doi.org/10.3102/1076998607309474>
- de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika*, 76(2), 179–199. <https://doi.org/10.1007/s11336-011-9207-7>
- de la Torre, J., & Douglas, J. A. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika*, 69(3), 333– 353. <https://doi.org/10.1007/BF02295640>
- de la Torre, J., & Douglas, J. A. (2008). Model evaluation and multiple strategies in cognitive diagnosis: An analysis of fraction subtraction data. *Psychometrika*, 73(4), 595-624. <https://doi.org/10.1007/s11336-008-9063-2>
- de la Torre, J., & Minchen, N. (2014). Cognitively diagnostic assessments and the cognitive diagnosis model framework. *Psicología Educativa*, 20(2), 89–97. <https://doi.org/10.1016/j.pse.2014.11.001>

- de la Torre, J., Chiu, C.-Y., & Lee, Y.-S. (2022). Advances in cognitive diagnostic models: New techniques for applying the DINA model in large-scale assessments. *Psychometrika*, 87(4), 901–926. <https://doi.org/10.1007/s11336-022-09863-3>
- Dogan, E., & Tatsuoka, K. (2008). An international comparison using a diagnostic testing model: Turkish students' profile of mathematical skills on TIMSS-R. *Educational Studies in Mathematics*, 68(3), 263–272. <https://doi.org/10.1007/s10649-007-9099-8>
- Doust, A. R., Khan, W. A., & Alghafri, M. S. (2021). Application of DINA model to assess Omani and Iranian fourth-grade students' mathematics' attributes. *North American Academic Research (Naar) Journal*, 4(10), 220–237. <https://doi.org/10.5281/zenodo.5639812>
- Henson, R., & Douglas, J. (2005). Test construction for cognitive diagnosis. *Applied Psychological Measurement*, 29(4), 262–277. <https://doi.org/10.1177/0146621604272623>
- Evrar, D. (2019). An application of cognitive diagnosis modeling in TIMSS: a comparison of Intuitive definitions of Q-matrices. *International Journal of Modern Education Studies*, 3(1), p.04. <https://doi.org/10.51383/ijonmes.2019.33>
- George, A. C., & Robitzsch, A. (2015). Cognitive diagnosis models in R: A didactic. *The Quantitative Methods for Psychology*, 11(3), 189–205. <https://doi.org/10.20982/tqmp.11.3.p189>
- Gündüz, T., & Çakan, M. (2020). TIMSS 2015 Türkiye örnekleminde matematik başarı testine dayalı olarak öğrencilerin bilişsel tanı modelleri ile sınıflandırılması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 14(2), 1040–1079.
- Hidajat, F. A. (2021). Students creative thinking profile as a high order thinking in the improvement of mathematics learning. *European Journal of Educational Research*, 10(3), 1247–1258.
- Junker, B.W. & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25(3), 258–272. <https://doi.org/10.1177/01466210122032064>
- Kalkan, Ö., & Toprak, M. (2022). The impact of Q-matrix refinement methods on the diagnostic accuracy of CDMs. *Educational Measurement: Issues and Practice*, 41(1), 52–64. <https://doi.org/10.1111/emip.12464>
- Karagiannakis, G. N., Baccaglini-Frank, A. E., & Roussos, P. (2016). Detecting strengths and weaknesses in learning mathematics through a model classifying mathematical skills. *Australian Journal of Learning Difficulties*, 21(2), 115–141. <https://doi.org/10.1080/19404158.2017.1289963>
- Lee, Y. S., Park, Y. S., & Taylan, D. (2011). Cognitive diagnostic modeling of attribute mastery in Massachusetts, Minnesota, and the US national sample using the TIMSS 2007. *International Journal of Testing*, 11(2), 144–177. <https://doi.org/10.1080/15305058.2010.534571>
- Leighton, J. & Gierl, M., (eds.). (2007). *Cognitive diagnostic assessment for education: theory and applications*. Cambridge, MA: Cambridge University Press. <https://doi.org/10.1017/CBO9780511611186>
- Li, X. & Wang, W. C. (2015). Assessment of differential item functioning under cognitive diagnosis models: The DINA model example. *Journal of Educational Measurement*, 52(1), 28–54. <https://www.jstor.org/stable/43940553>

- Li, L., An, Y., Ren, J., & Wei, X. (2021). Research on the cognitive diagnosis of Chinese listening comprehension ability based on the G-DINA model. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.714568>
- Li, H. & Suen, H. K. (2013). Constructing and validating a q-matrix for cognitive diagnostic analyses of a reading test, *Educational Assessment*, 18(1), 1-25. <https://doi.org/10.1080/10627197.2013.76152>
- Liu, J., & Kang, T. (2023). Advances in cognitive diagnostic models: The role of DINA in educational assessment. *Educational Assessment*, 29(1), 45–67. <https://doi.org/10.1007/s11336-022-09762-9>
- Ma, W., & de la Torre, J. (2016). A sequential cognitive diagnosis model for polytomous responses. *British Journal of Mathematical and Statistical Psychology*, 69(3), 253-275. <https://doi.org/10.1111/bmsp.12070>
- Parlak, B. (2018). *Investigation of the TIMMS 2015 mathematics achievement in eight grade students with cognitive diagnostic model*. (Publication No.494335) (Phd thesis, Hacettepe university) Council of Higher Education National Thesis Center).
- Parlak, B. (2023). Matematik performansının bilişsel tanı modeli ile değerlendirilmesi: TIMSS Türkiye ve Singapur örneği. *Milli Eğitim Dergisi*, 52(1), 413-436.
- Roussos, L. A., DiBello, L. V., Stout, W., Hartz, S. M., Henson, R. A., & Templin, J. L. (2007). *The Fusion Model Skills Diagnosis system*. In Cambridge University Press eBooks (pp. 275–318). <https://doi.org/10.1017/cbo9780511611186.010>
- Rupp, A. A. & Templin, J. L. (2008). Unique characteristics of diagnostic classification models: A comprehensive review of the current state of the art. *Measurement*, 6(4), 219-262. <https://doi.org/10.1080/15366360802490866>
- Sawaki, Y., Kim, H. J., & Gentile, C. (2009). Q-Matrix construction: Defining the link between constructs and test items in large-scale reading and listening comprehension assessments. *Language Assessment Quarterly*, 6(3), 190–209. <https://doi.org/10.1080/15434300902801917>
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461-464.
- Shi, Q., Ma, W., Robitzsch, A., Sorrel, M. A., & Man, K. (2021). Cognitively diagnostic analysis using the G-DINA model in R. *Psych*, 3(4), 812–835. <https://doi.org/10.3390/psych3040052>
- Su, Y. L. (2013). *Cognitive diagnostic analysis using hierarchically structured skills*. The University of Iowa.
- Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*, 20(4), 345-354. <https://doi.org/10.1111/j.1745-3984.1983.tb00212.x>
- Tatsuoka, K. (1985). A probabilistic model for diagnosing misconceptions in the pattern classification approach. *Journal of Educational Statistics*, 10(1), 55–73. <https://doi.org/10.2307/1164930>
- Tatsuoka, K. K. (1995). Architecture of knowledge structures and cognitive diagnosis: A statistical pattern recognition and classification approach. *Cognitively Diagnostic Assessment*, 327-359.
- Tatsuoka, K.K. (2009). *Cognitive assessment—an introduction to the rule space method*. Routledge, New York. <https://doi.org/10.4324/9780203883372>

- Tatsuoka, C., Varadi, F., & Jaeger, J. (2013). Latent partially ordered classification models and normal mixtures. *Journal of Educational and Behavioral Statistics*, 38(3), 267-294. <https://doi.org/10.3102/1076998612458318>
- Tatsuoka, C., Clements, D. H., Sarama, J., Izsák, A., Orrill, C. H., de la Torre, J., ... & Tatsuoka, K. (2016). Chapter 4. *Journal for Research in Mathematics Education. Monograph*, 15, 73-96.
- Templin, J. L., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological methods*, 11(3), 287. <https://doi.org/10.1037/1082-989X.11.3.287>
- Terzi, R., & Sen, S. (2019). A nondiagnostic assessment for diagnostic purposes: Q-Matrix validation and Item-Based Model fit evaluation for the TIMSS 2011 assessment. *SAGE Open*, 9(1). <https://doi.org/10.1177/2158244019832684>
- Yakar, L., Dogan, N. Senol, D., & Yuksel, N. S. (2021). Monitoring student achievement with cognitive diagnosis model. *Journal of Measurement and Evaluation in Education and Psychology*, 12(3), 303-320. <https://doi.org/10.21031/epod.903084>
- Wang, L., Li, Y., & Zhang, J. (2023). A new approach to Q-matrix construction for cognitive diagnostic models: Combining expert knowledge with statistical methods. *Psychometrika*, 88(1), 123–145.
- Whitely, S. E., & Schneider, L.M. (1981). Information structure for geometric analogies: A test theory approach. *Applied Psychological Measurement*, 5, 383-397. <https://doi.org/10.1177/014662168100500312>
- Wu, H., Liu, Y., & Zhang, X. (2020). The effects of attribute granularity on the estimation quality of cognitive diagnosis models. *Applied Psychological Measurement*, 44(7), 546–561. <https://doi.org/10.1177/0146621620917447>
- Zhang, S., Liu, J., & Ying, Z. (2023). *Statistical applications to cognitive diagnostic testing*. Annual Review of Statistics and its Application, 10(1), 143–165. <https://doi.org/10.1146/annurev-statistics-033021-111803>

Examining Mathematics Questions with a Cognitive Diagnostic Model: A Q-Matrix Study

Aygil TAKIR, Atatürk Teacher Training Academy, ORCID ID: 0000-0003-3042-7585

Hasan ÖZDER, Atatürk Teacher Training Academy, ORCID ID: 0000-0003-1094-3590

Osman CANKOY, Middle East Technical University North Cyprus Campus, ORCID ID: 0000-0002-4765-9297

Highlights

- This study examined the cognitive structure of the mathematics subtest of the College Entrance Exam (KGS) administered in Northern Cyprus using the DINA model.
- The items aligned well with the skills to be measured, and the test structure is reliable.
- The constructed Q-matrix accurately reflects the cognitive structure of the test.
- Descriptive statistics based on the DINA model revealed variability in students' mastery levels across different cognitive attributes.

Abstract

Accurate and comprehensive assessment of students' cognitive abilities is not only important to support individual learning processes, but also to improve the quality of an education system. In this context, it is essential to use assessment tools that can help shape educational policy and improve curricula. This study was conducted to determine the psychometric characteristics of the CEE mathematics test administered to 1,833 fifth-grade primary school students in 2021. A cognitive diagnostic model, namely DINA, was used in this study to achieve the purpose. The theoretical infrastructure of the DINA model is easy and strong compared to other models. The Q-matrix was constructed for 23 mathematics questions and 15 attributes were determined based on expert opinions. The results showed that the mean values of the parameters g and s were 0.25 and 0.13, respectively. The mean of the δ values of the fifth-grade students was high (0.73). Skill patterns showed that 13% of fifth-grade students possessed all the attributes included in CEE. The construction of the Q-matrix is a very important step in the use of CDMs. The cognitive attributes and the Q-matrix were determined by expert opinion in this study. The (1-s) values indicated the validity of the Q-matrix. For all of the items, 1-s values appeared to be quite high. This showed that the attainments and item-attainment relationship were done as accurately as possible. These results can be interpreted as the validity of the Q-matrix. These results showed that the DINA model provides a detailed conclusion about students' attributes.

Keywords: Cognitive Diagnosis Model, College Entrance Examination, DINA, Mathematics Assessment



Inonu University
Journal of the Faculty of
Education
Vol 26, No 2, 2025
pp. 759-796
[DOI](#)
10.17679/inuefd.1594482

[Article Type](#)
Research Article

[Received](#)
05.12.2024

[Accepted](#)
03.06.2025

Suggested Citation

Takir, A., Cankoy, O. & Özder, H. (2025). Examining Mathematics Questions with a Cognitive Diagnostic Model: A Q-Matrix Study. *İnönü University Journal of the Faculty of Education*, 26(2), 759-796. DOI: 10.17679/inuefd.1594482

1. Introduction

Many cognitive skills can affect student performance in mathematics, such as recognizing the problem, simplifying and representing the problem, looking for patterns and relationships, and performing calculations (Burkhardt & Swan, 2012). These cognitive skills can be divided into four domains: a) core numbers (placing numbers on number lines, estimating a small number of objects accurately, managing symbols, etc.), b) memory retrieval and processing (performing mental calculations, remembering theorems and formulas, remembering procedures, etc.), c) reasoning (decision-making, generalizing, etc.), and d) visual-spatial organizations (understanding and utilizing the spatial arrangement of mathematical object representations, analyzing graphs and tables. etc.) (Karagiannakis et al., 2016). Furthermore, many complex mathematical skills may be associated with more than one domain.

In addition to these cognitive domains, higher-order cognitive skills such as making complex connections, developing flexible problem-solving strategies and constructing logical arguments are crucial for success in mathematics. These skills enable students to approach unfamiliar problems with creativity, justify their arguments and evaluate the effectiveness of different solutions. This emphasises the importance of integrating higher-order thinking into both curriculum design and assessment practices (Hidajat, 2020). Since higher-order cognitive skills vary widely among learners, the assessment of these skills requires a personalized approach. Personalized assessment methods are important not only to accurately capture the unique learning profiles of individual students but also to develop targeted instructional strategies that promote advanced cognitive development. In this context, cognitive diagnostic models (CDMs) provide a valuable framework for identifying specific strengths and weaknesses in students' higher-order thinking and offer more nuanced information than traditional assessment approaches.

CDMs provide information about students' cognitive skills. They show students' cognitive strengths and weaknesses (Leighton & Gierl, 2007) and help to understand their knowledge structures and abilities. Psychometric models consider students individually, rather than referring to other students. CDMs measure components necessary for in a particular domain and integrate the theories of cognition and measurement to support students learning.

CDMs are mainly related to classroom teaching and learning, and processes (Junker & Sijtsma, 2001; Tatsuoka, 1985). Traditional measurement theory is not fine-grained enough to evaluate cognitive skills (Li, Ren & Wei, 2021). However, CDMs combine a cognitive process and measurement method and provide advantages in evaluation. The application of the theory not only reveals the potential characteristics of learners in cognitive processing but also provides targeted remedies and strategic guidance for individuals. de la Torre (2009) defines CDMs as latent variable models developed primarily for assessing student mastery and non-mastery set of fine-grained skills. In other words, CDMs are developed to identify the presence or absence of multiple fine-grained skills required for solving problems on a test. An *attribute* is a very important concept in cognitive diagnostic assessment (Leighton & Gierl, 2007). Attributes have been defined as specific objectives or skills. CDMs can be used to identify the skill combinations that the examinee is likely to either possess or not possess (Su, 2013). The presence and absence of skills are referred to as skills mastery and non-mastery, respectively, and are represented by a vector of binary latent variables. Thus, instead of a single score, a profile can be generated for a student to indicate which skills are mastered and whether they are mastered or not. Therefore, it can be concluded that CDMs provide the framework that intends to evaluate students' attributes.

CDMs offer more useful feedback regarding test takers' mastery or non-mastery of specific sub-skills than the total scores derived from traditional assessment procedures (DeCarlo, 2010). These models help identify individuals' specific strengths and weaknesses, which can be used to inform classroom instruction and support student learning. Mastery

testing and mastery-based models provide a statistical framework for making informed decisions about whether an individual has achieved sufficient competence in a specific domain. These models are particularly valuable in high-stakes contexts, such as determining whether an individual should pass a licensure examination or is ready to advance to the next academic level. However, in many modern assessments, test items are designed to require the application of multiple, fine-grained cognitive attributes rather than a single skill. This complexity highlights the limitations of traditional mastery models and underscores the need for more nuanced diagnostic approaches that can capture the multifaceted nature of student knowledge (Zhang et al., 2023).

For any CDM, it is important to specify an item-attribute incidence matrix called the *Q-matrix* (Tatsuoka, 1983). Q-matrix involves the identification of concepts and procedures, known as cognitive attributes in general, requisite to observing the accurate performance on individual test items (Clements et al., 2020). Q-matrix shows the relationship between the items in a test and the attributes required to respond to the items correctly. It also expresses how each item on an assessment measures one or more attributes. Therefore, it explicitly coordinates the components of reasoning, content knowledge, and processing skills on which an assessment is based (Tatsuoka, 2009). Recent studies have highlighted the importance of improving the Q-matrix with data-driven techniques alongside expert knowledge, to reduce potential biases and errors in the attribute identification process (Wang et al., 2023).

The Q-matrix plays a crucial role in ensuring clarity in the assessment process. This clarity is essential for tools that track student progress and provide insights for instructional strategies (e.g., formative assessments), item creation (e.g., determining if items assess the intended content and skills), and curriculum alignment (e.g., verifying if the content and skills align with required standards for a specific grade or course) (de la Torre et al., 2016). Establishing the connection between items and attributes requires experts to identify the attributes. However, recent advances suggest that incorporating both expert judgments and statistical methods in Q-matrix construction can lead to more accurate diagnostic assessments (Wang et al., 2023). This dual approach ensures that the Q-matrix more reliably reflects the complex nature of student knowledge and enhances the validity of cognitive diagnostic models.

In the Q-matrix, row (j) of the matrix represents test items, and column (k) represents the sub-skills to be measured. Therefore, the Q-matrix results in a $j \times k$ matrix. If an item measures an attribute, “1” is used in the matrix cell. If the attribute is not required by the item “0” is used in the matrix cell. Each item in the test is arranged to measure one or more of the specified attributes. Q-matrix can be considered as a kind of association or bridge that describes the relationship between test items and attributes. The development of the Q-matrix is one of the most important steps in using any CDM.

By a Q-matrix, 2^k latent classes can be formed for k cognitive attributes defined. The latent classes show which attributes individuals have and which they do not together with questions expected to be answered correctly by individuals (Yakar et al., 2021). Each of the attribute patterns is a separate latent class and they have different structures (de la Torre, Hong & Deng, 2010). For $k = 3$, the eight latent classes are shown in Table 1.

Table 1.*Latent Classes and Their Meanings (for k=3)*

Latent Classes	Indicating individuals who are...
(000)	non-masters of any of the attributes
(100), (010), (001)	masters in one of the attributes
(110), (101), (011)	masters in two of the attributes
(111)	masters in all the attributes

There are a large number of CDMs that have been used in different studies for different purposes. CDMs are differentiated by reflecting *compensatory* or *non-compensatory* relationships between attributes of an item. In a compensatory model, the mastery of one attribute can compensate for the lack of other attributes measured by the same item. On the contrary, in non-compensatory models, the mastery of all the attributes is required to answer an item correctly (de la Torre, 2011). For all CDMs, it is very important to determine the attributes included in the test and to make the correct associations between items and attributes to diagnose individuals correctly.

Some CDMs include (a) the deterministic input noisy “and” gate (DINA) (de la Torre, 2008), (b) generalized DINA (G-DINA) (de la Torre, 2011), (c) deterministic input noisy “or” gate (DINO) (Templin & Henson, 2006), (d) higher-order DINA (HO-DINA) (de la Torre & Douglas, 2004; de la Torre & Minchen, 2014), and (e) non-compensatory reduced re-parameterized unified model (NC-RRUM) (Roussos et al., 2007). The DINA (de la Torre, 2008) is a non-compensatory model, meaning that, under its assumptions, the mastery of one sub-skill cannot compensate for the lack of another sub-skill measured by the test items (George & Robitzsch, 2015). The DINA model is one of the simplest CDMs, assuming examinees must possess all the skills required to effectively solve an item, making it a non-compensatory model; lacking one required attribute cannot be made up for by the presence of other attributes (de la Torre & Minchen, 2014). In other words, in the DINA model, the absence of one attribute is equivalent to the absence of all required attributes (Henson & Douglas, 2005).

The interaction between the examinee attributes and item specification defines the latent response variable. This is also known as the ideal response (Tatsuoka, 1995) and is defined as

$$\eta_{ij} = \prod \alpha_{ik}^{q_{jk}}$$

where α_{ik} is the mastery of attribute k ($k = 1, \dots, K$) for examinee i and q_{jk} is the state of the j^{th} item and k^{th} attribute in the Q-matrix. According to the above equality, $\eta_{ij} = 1$ if and only if examinee i has mastered all the required attributes for item j .

The DINA parameterizes slipping (s_j) and guessing (g_j) parameters (Doust et al., 2021). s_j happens when examinees who possess all the required skills for an item miss the item. On the other hand, g_j happens when the examinees lack the required skills and still answer the item correctly. According to Junker and Sijtsma (2001), these parameters are simply “*false negative and false positive rates*,” respectively. The slip and guessing parameters are given by the following formulas, respectively.

$$s_j = P(X_{ij} = 0 | \eta_{ij} = 1), \text{ and}$$

$$g_j = P(X_{ij} = 1 | \eta_{ij} = 0),$$

The probability of a correct response can be written as:

$$P(X_{ij} = 1 | \alpha_i, s_j, g_j) = (1 - s_j)^{\eta_{ij}} (g_j)^{1 - \eta_{ij}}$$

where g_j and s_j are the guessing and slipping item parameters (Minchen & de la Torre, 2018).

Answering an item correctly requires that the examinee possesses all the necessary skills to avoid slipping and does not have at least one of the required skills to guess correctly. It is important to note that, in the absence of guessing and slipping parameters, the model probability of a correct response to an item is either 0 or 1. Templin and Henson (2010) described the s and g parameters as shown in Table 2.

Table 2.

Response Probabilities in the DINA Model

	$x_{ij}=1$	$x_{ij}=0$
$\eta_{ij} = 1$	$1-s_j$	s_j
$\eta_{ij} = 0$	g_j	$1-g_j$

Item parameter analysis contains an *item discrimination index* (IDI- δ) to estimate the s and g index, for each item independently (i.e., $\delta = 1 - g - s$). The δ indicates how much an item would be answered correctly by examinees without the effect of g and s parameters.

The DINA model offers several advantages that make it highly useful in cognitive diagnostic assessments. One of its key strengths is the ability to profile students based on their mastery of specific cognitive attributes. By identifying which sub-skills a student has mastered, the DINA model provides teachers with valuable insights into students' strengths and weaknesses, allowing for more targeted and effective instructional strategies. Another significant advantage is the model's flexibility in handling a wide range of item types, including multiple-choice questions. This makes it adaptable to different forms of assessments, broadening its applicability (Liu & Kang, 2023). Additionally, the DINA model is supported by various software packages, which facilitate the estimation of model parameters and make the application of the model more accessible to researchers and practitioners (de la Torre et al., 2022). Finally, by offering a detailed analysis of students' abilities across different cognitive attributes, the DINA model allows educators to gain a deeper understanding of individual learning needs, helping to better support student progress (Wang et al., 2023). These advantages make the DINA model a powerful tool for both assessing and improving student learning outcomes.

Practical applications of CDMs exist in different areas of assessment, such as language testing [Li & Suen, 2013; Sawaki et al., 2009], psychology [Tatsuoka, Varadi, & Jaeger, 2012; Templin & Henson, 2006] and mathematics [DeCarlo, 2011; Whitely & Schneider 1981; Parlak, 2017; Birenbaum et al., 2004; Choi, Lee & Park, 2015; Dogan & Tatsuoka, 2008]. Having a tight prerequisite relationship of mathematical skills makes it necessary to apply the CDM in testing these skills. In other words, it is important that educators, curriculum developers, and policymakers understand what students know and can do in mathematics and what areas, concepts, and topics need more focus and effort.

There exists a body of research studies that focus on the cognitive analysis of TIMSS (Trends in International Mathematics and Science Study) mathematics tests [Birenbaum et al., 2004; Dogan & Tatsuoka, 2008; Choi et al., 2015; Lee & Park, 2015; Parlak, 2018; Terzi & Sen, 2019; Evran, 2019]. Some of these studies compared the cognitive analysis of students in different countries [Birenbaum et al., 2004; Dogan & Tatsuoka, 2008; Choi, Lee & Park, 2015]. Most of these studies use the DINA model to analyze cognitive diagnostic data in different educational contexts. They demonstrate the effectiveness of the model as it provides detailed insights into students' cognitive strengths and weaknesses and enables a more nuanced understanding of student performance compared to traditional assessment methods. The study conducted by Choi, Lee and Park (2015) aimed to analyze the results of the 2003 TIMSS 8th grade mathematics assessment using the DINA model to provide more specific and actionable

information about student performance. The study found that the DINA model provides valuable diagnostic information, particularly when comparing student performance between the US and Korea, with the DINA discrimination index offering further interpretive possibilities.

In another study conducted by Gündüz and Çakan (2020), the aim was to investigate the classification of students based on various attributes using mathematics test data from the Turkish TIMSS 2015 sample. The DINA model (together with the DINO model) was used in the study. Four different Q matrices were created based on the content domain (numbers, data representation, geometric shapes and measures), the levels of the cognitive domain (knowledge, application, reasoning), a combination of both domains, and the attributes derived through principal component analysis. The study analyzed the accuracy and consistency of the classifications made using these Q-matrices. The results showed that the Q-matrix using the levels of the content area as an attribute dimension provided the best classification.

In her study, Parlak (2023) aimed to investigate the strengths and weaknesses of Year 8 students in mathematics, focusing on knowledge and skills. Using student response patterns from the TIMSS 2015 8th grade mathematics test and a Q-matrix developed by experts, the study compared data from students in Turkey and Singapore. The DINA model was used to assess students' strengths and weaknesses, regardless of overall score and ranking. Key findings include Singapore's poor performance in transformation geometry despite high overall scores and Turkish students' difficulties in logical reasoning and open-ended tasks.

These studies highlight the increasing use of the DINA model in educational research, particularly in mathematics, where understanding students' cognitive skills is critical to effective teaching and learning.

1.1. Problem Statement

In the Turkish Cypriot education system, following primary school education, families tend to prefer Türk Maarif Colleges for secondary education. These schools provide both secondary and high school education under the same roof. However, enrolment is conditional: students who would like to study in these English-medium schools need to take the College Entrance Examination (CEE), a curriculum-orientated performance test administered to fifth-grade primary school students. In Maarif Colleges, the medium of instruction is English. These schools do not charge any fees since they are governed by the state. At the end of their high school education, students can study in universities in Northern Cyprus, Turkey and other countries. These colleges are quite prestigious in Northern Cyprus; many students go through a long and expensive preparation process usually starting from 4th grade.

Each year the CEE is held in two different modules, namely, CEE1 and CEE2. The modules are generally administered five months apart. The mean score of the results of both modules is considered in determining student success. Due to the COVID-19 pandemic, the exam was held once in 2021. The CEE comprises several sub-tests, including Mathematics (27 questions), Turkish Language (27 questions), Science and Technology (14 questions), Social Sciences (10 questions) and English Language (22 questions). It is a multiple-choice exam which asks students to choose the correct answer among a set of four different alternatives that include three distractors.

The current study aims to investigate the cognitive structure of the CEE by using a CDM, namely DINA. More specifically, this study investigates the ways of using the DINA model in defining psychometric characteristics of the CEE mathematics test. By analysing the cognitive attributes associated with the mathematics sub-test, the study explores how well the test reflects students' proficiency in various mathematical areas. In this study, 21 attributes for the CEE mathematics sub-test were determined. In the related literature, there are different opinions about the number of attributes. According to Lee et al. (2011), the number of attributes should be set to 15. On the other hand, Tatsuka (1984, 1990) suggested that at least 8 attributes

should be defined for determining students' skills. Since applying all 21 attributes would take more time and require a broader research project, the researchers focused on 15 attributes including Natural Numbers, Four Operations with Natural Numbers, Natural Numbers Problems and Geometry topics within the scope of this study. The objectives of the study are a) to estimate the CEE mathematics item parameters and b) to study 5th-grade students' mathematics attribute mastery levels in the CEE mathematics test.

CDMs measure the multidimensional constructs within practical testing constraints and support highly specific assessments of mathematical knowledge (de la Torre et al, 2016). The DINA model was chosen from the CDMs for the purposes of this research. Its theoretical infrastructure is simple and robust, making it a reliable choice for cognitive diagnostic analysis. Furthermore, the DINA model allows for a highly specific and detailed assessment of students' cognitive abilities, offering insight into their individual mathematical competencies. In this study, the Q matrix was developed by the researchers based on expert opinions as expert opinion is still the main method for creating Q-matrices in CDM applications. The involvement of expert opinion ensures the alignment of the attributes with the intended learning outcomes, making the analysis more grounded in the actual curriculum.

It is very important to determine the learning deficiencies of students in mathematics. The attribute profile identifies the students' domain-specific strengths and weaknesses and can help the teachers make instructional decisions to better address the students' specific cognitive needs. As discussed before, CDMs can provide more detailed information in the form of scores which allow for effective measurement of student learning and progress, design better instruction, and address individual and group needs (de la Torre, 2009, 2011). Thus, the findings of this study can help teachers tailor instruction to better meet the diverse needs of students, improving their learning outcomes in mathematics.

This study is of particular importance as it examines the DINA model in the context of high-stakes testing, specifically the CEE for 5th grade students in Northern Cyprus. The CEE is not only a critical milestone in students' academic careers, but also plays a crucial role in determining their future educational opportunities. By applying the DINA model to analyze student performance in the CEE mathematics test, this study aims to provide a deeper understanding of the cognitive processes involved in the exam and how well the current assessment reflects students' actual abilities.

The significance lies in the potential impact on the future design and reform of testing procedures in Northern Cyprus. As education systems around the world move towards data-driven, personalized approaches to learning, there is a growing need for assessments that provide more than a summative score. The DINA model provides a detailed breakdown of the specific cognitive attributes that students have mastered or are struggling with, allowing educators to target areas of weakness and develop more targeted interventions. This shift in assessing students' abilities could lead to a more nuanced approach to educational testing and help close students' learning gaps earlier in their educational careers.

Furthermore, the results of this study could inform decision-making not only in individual classrooms, but also across the education system, helping to develop more effective, personalized educational practices that can improve overall student success. In the long term, this could help policymakers rethink traditional assessment frameworks and explore more diagnostic and flexible methods of student assessment.

2. Method

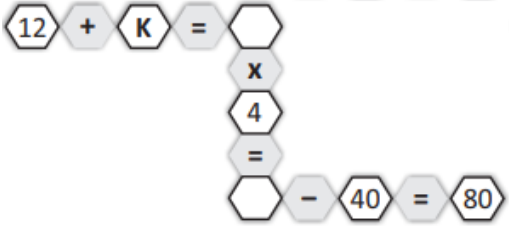
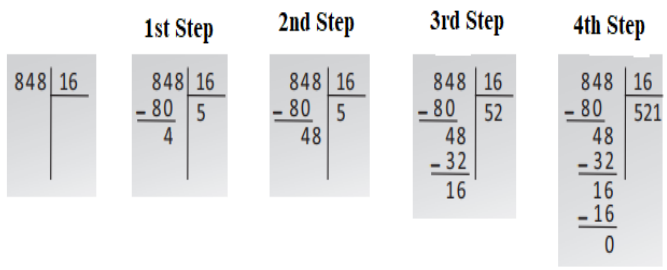
The current study was conducted in order to define the psychometric characteristics of the CEE mathematics sub-test administered to 1,833 5th-grade primary school students in 2021. Thus, the test results were used as survey data. In addition, expert opinions were used to create the Q-matrix model.

2.1. Data Collection Tools

The CEE 2021 mathematics test was used for data collection. The test was administered in 2021 and consisted of 17 pages. The sample of the questions and their attributes in the mathematics test are presented in Table 3.

Table 3.

Sample of the Mathematics Questions in CEE 2021

Question	Attributes
According to the given figure, which number should be written instead of "K" to satisfy all operations?  A) 2 B) 8 C) 12 D) 18	- Performing addition and subtraction with natural numbers. - Doing multiplication with at most 3-digit natural numbers. - Finding a missing element (multiplier, quotient, or division) by understanding the relationship between multiplication and division.
AA is a two-digit natural number. If $AA + AA + AA = 132$, Find $4 + (4 \div A)$. A) 2 B) 3 C) 4 D) 5	- Performing addition and subtraction with natural numbers. - Specifying the digits and place values of natural numbers (up to nine digits). - Doing multiplication with at most 3-digit natural numbers.
Fatma does the following division in 4 steps. What can be said about Fatma's work?  A) She divided without any error. B) She made the first mistake in step 2. C) She made the first mistake in step 3. D) She made the first mistake in step 4.	Dividing a natural number of up to four digits into a natural number of up to two digits.

2.2. Process and Data Analysis

The study was conducted using the steps involved in creating a CDM using the DINA model. The researchers used the steps recommended by Shi et al. (2021), namely Q-matrix evaluation, CDM calibration, model fit evaluation, item diagnostic testing, classification reliability testing, and result presentation and visualization.

2.2.1. Data and Q-Matrix Preparation

The Q-matrix was created for 23 mathematics questions and 15 attributes. While creating the Q-matrix for DINA model analysis, the objectives of Natural Numbers, Four Operations with Natural Numbers, and Natural Number Problems were considered as cognitive attributes. Researchers carefully discussed every question in the test and determined cognitive attributes representing topics by examining the 5th-grade mathematics curriculum. The researchers constructed the first version of the Q-matrix after this study. According to de la

Torre and Minchen (2014), although statistical procedures could provide some valuable insights into the Q-matrix, the appropriateness of the Q-matrix should be carefully assessed by domain experts. For this purpose, the first version of the Q-matrix was sent to four experts, three working in the field of mathematics education and one in measurement and evaluation.

The experts were requested to provide comments and suggestions for each attribute in the Q-matrix. They were asked to rate how well each skill was measured in CEE. Additionally, the experts offered comments and suggestions for each attribute and, in some instances, proposed additional attributes for consideration.

After this step, a Q-matrix was created by associating each item with 15 attributes determined by the experts. The attributes and their frequency for representing items and Q-matrix are given in Table 4 and Table 5, respectively.

Table 4.

Attributes and Frequencies

Attributes	Frequencies
Attribute 1. Rounding a natural number to the nearest hundreds	1
Attribute 2. Writing and reading at most 9-digit natural numbers.	2
Attribute 3. Specifying the digits and place values of natural numbers (up to nine digits).	3
Attribute 4. Performing addition and subtraction with natural numbers.	12
Attribute 5. Guessing the results of addition and subtraction operations.	2
Attribute 6. Doing multiplication with at most two-digit natural numbers.	8
Attribute 7. Dividing a natural number of up to four digits into a natural number of up to two digits.	7
Attribute 8. Using an appropriate strategy for mental multiplication and division.	1
Attribute 9. Interprets the remainder in a division problem.	1
Attribute 10. Finding a missing element (multiplier, quotient, or division) by understanding the relationship between multiplication and division.	3
Attribute 11. Find the result of parenthetical expressions containing up to two operations.	2
Attribute 12. Solving word problems on four operations.	8
Attribute 13. Expressing the position of a point relative to another point	1
Attribute 14. Identify and draw the basic elements of a rectangle, parallelogram, rhombus and trapezoid.	1
Attribute 15. Determining the sum of the measures of the interior angles of a triangle and quadrilateral and finding the missing angle.	1

A total of 54 associations were made for the 23 items. Ten items were associated with three, eleven items with two, and two items with one attribute in the test.

As shown in Table 5, the Q-matrix indicates the attributes measured by each item. A cell with the value 1 means that the corresponding item measures the corresponding attribute, and a cell with the value 0 means the opposite.

Table 5.*Q-matrix Constructed by the Experts*

Item No - Attributes	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0
4	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
7	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0
8	0	0	0	1	0	1	0	1	0	0	0	0	0	0	0
9	0	0	0	1	0	1	0	0	0	0	0	1	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
11	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0
12	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0
13	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0
14	0	1	0	0	0	0	1	0	1	0	0	1	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
16	0	0	1	0	0	0	0	0	0	1	0	1	0	0	0
17	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
19	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0
20	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0
21	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0
22	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0
23	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0

2.3. Data Analysis

The CDM package through the Ox Edit programming using DINA (Ma & de la Torre, 2016) was used to analyze the data. The necessary syntax codes were loaded to be able to analyze the data by using the DINA model. These syntax codes allow estimating DINA model parameters by using the program's expectation-maximization EM algorithm. *s* and *g* parameters were taken at different limit values [i.e. 0.05 - 0.40, Henson & Douglas (2005); 0.25 - 0.15, Rupp & Templin (2008)]. In the current study, *s* and *g* parameters levels were interpreted as low and high (0.45 and above was determined as high level, and 0.25 and below as low level).

2.3.1. Model Data Fit Statistics

Model data compatibility of the DINA model was carried out with Test-Level Fit Statistics analysis. Relative model fit has typically been evaluated using conventional information-based indices, namely, the AIC-Akaike's information criterion (Akaike, 1974), the BIC-Bayesian information criterion (Schwarz, 1978).

3. Findings

Table 6 presents 5th-grade students' item parameter estimation results which were analyzed by using DINA.

The average values of the g and s parameters in the DINA model were 0.25 and 0.13, respectively. The mean score of the g parameter revealed that 25% of 5th grade students answered the items by guessing parameters without possessing all attributes. The probability of predicting the correct answer by a student who has no idea about the answer was 0.25. Therefore, it can be said that the mean value of the g parameters is the average value.

The g parameter values varied between 0 and 0.387. Besides, the g parameter values of items 2, 3, 9, 12 and 15 were high. In other words, the probability of answering an item correctly is high even though a student does not have enough information to respond to these items. It can be interpreted that students who do not have enough information have a moderate probability of answering these questions correctly.

Table 6.

Item Parameter Estimates

Item no	g	SE (g)	s	SE (s)	δ	1- s
1	0.02	0.02	0.14	0.01	0.84	0.86
2	0.19	0.02	0.15	0.01	0.66	0.85
3	0.36	0.02	0.17	0.01	0.47	0.83
4	0.29	0.02	0.14	0.01	0.57	0.86
5	0.32	0.02	0.16	0.01	0.51	0.84
6	0.34	0.02	0.16	0.01	0.50	0.84
7	0.26	0.02	0.00	0.02	0.74	1.00
8	0.39	0.02	0.20	0.01	0.41	0.80
9	0.35	0.02	0.19	0.01	0.46	0.81
10	0.34	0.02	0.20	0.01	0.47	0.80
11	0.00	0.03	0.00	0.02	1.00	1.00
12	0.29	0.02	0.00	0.02	0.71	1.00
13	0.00	0.03	0.00	0.02	1.00	1.00
14	0.35	0.02	0.13	0.01	0.53	0.87
15	0.27	0.02	0.20	0.01	0.53	0.80
16	0.00	0.03	0.00	0.02	1.00	1.00
17	0.34	0.02	0.20	0.01	0.46	0.80
18	0.36	0.02	0.19	0.01	0.45	0.81
19	0.28	0.02	0.14	0.01	0.58	0.86
20	0.24	0.02	0.13	0.01	0.63	0.87
21	0.30	0.02	0.17	0.01	0.54	0.83
22	0.36	0.02	0.22	0.01	0.43	0.78
23	0.18	0.02	0.21	0.01	0.61	0.79

It was observed that the s parameter varied between 0 and 0.22 for each item. The highest s parameter value was obtained from item 22 (0.22). For other items, the s parameter

values were low. Items with the lowest s parameters were 7, 11, 12, 13 and 16. When the Q-matrix was examined, it was observed that these items were associated heavily with *"performing addition and subtraction with natural numbers"* and *"doing multiplication with at most two-digit natural numbers"* attributes, respectively.

According to de la Torre (2008), 1- s values closer to 0 indicate a misrepresentation of the attributes of the elements defined by the Q-matrix. In this regard, it means that it is interpreted as an indicator of the degree of correspondence between the Q-matrix and the items of the assessment instrument. The current analysis indicated that the 1- s values for the items varied between 0.78 and 1. The mean for the 1- s values was 0.73. This indicates that the Q-matrix enables a correct assignment of the items to the attributes required for a correct answer. When all parameters are evaluated, it can be said that the most ideal items were the 1st, 11th, 13th, 16th, and 23rd items.

3.1. Skills Probability

The attribute probability follows 2^A , where "A" demonstrates the number of probability attributes that examinees possess. The number of attributes in the present study was 15, which means that there are 2^{15} latent classes for 15 attributes. The probability of the attributes are presented in Table 7.

Results of the analysis showed that the highest attribute mastery probability was *"performing addition and subtraction with natural numbers"* (0.83). It can be concluded that 83% of the 5th-grade students showed a high mastery proportion of this attribute. The lowest mastery proportion was attributed to *"expressing the position of a point relative to another point using direction and units"* (59%).

Table 7.

Skills Probabilities

Attributes	Probabilities
Attribute 1. Rounding a natural number to the nearest hundreds	0.660
Attribute 2. Writing and reading at most 9-digit natural numbers.	0.784
Attribute 3. Specifying the digits and place values of natural numbers (up to nine digits).	0.767
Attribute 4. Performing addition and subtraction with natural numbers.	0.829
Attribute 5. Guessing the results of addition and subtraction operations.	0.624
Attribute 6. Doing multiplication with at most two-digit natural numbers.	0.771
Attribute 7. Dividing a natural number of up to four digits into a natural number of up to two digits.	0.872
Attribute 8. Using an appropriate strategy for mental multiplication and division.	0.620
Attribute 9. Interprets the remainder in a division problem.	0.610
Attribute 10. Finding a missing element (multiplier, quotient, or division) by understanding the relationship between multiplication and division.	0.685
Attribute 11. Find the result of parenthetical expressions containing up to two operations.	0.717
Attribute 12. Solving word problems on four operations.	0.665
Attribute 13. Express the position of a point relative to another point using direction and units	0.590
Attribute 14. Identify and draw the basic elements of rectangle, parallelogram, rhombus and trapezoid.	0.601
Attribute 15. Determines the sum of the measures of the interior angles of triangles and quadrilaterals and finds the missing angle	0.638

3.2. Skills Class Profiles

Table 8 shows a sample of 5th-grade students' mathematics skill class profiles. The attribute pattern (111111111111111) represents the number of skills in the study (15 skills). There are 2¹⁵ possible latent classes for the Q-matrix configuration in this study.

Table 8.

Sample Latent Class Probabilities

Mastery Pattern															Probability
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.13
1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0.05
1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0.03
1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0.03
1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0.02

The skill class profile showed that 0.127 of 5th-grade students possessed all of the attributes. This pattern showed that %13 5th grade students possessed all 15 attributes. The other attribute pattern (111111110111111) indicated that the 5th-grade students lack at least one skill. This pattern showed that %5 of 5th-grade students possessed 14 skills and that they did not master one.

4. Discussion & Conclusion

The aim of this study was to investigate the cognitive structure of the CEE maths subtest in Northern Cyprus using a CDM, specifically the DINA model. As mentioned earlier, the attributes were determined by expert opinions. One of the major criticisms of the experts was that despite the curriculum-based nature of the CEE, certain learning areas of the maths curriculum (such as measurement and data) were not represented in the test. Consequently, no tasks were linked to attributes from these content areas. This emphasizes the need to pay more attention to content validity in test construction. Consequently, no items were linked to attributes from these content areas. Possible reasons for this discrepancy could be the tendency to favor easier-to-assess content areas or time constraints in test development. In addition, there could be a discrepancy between the implementation of the curriculum in the classrooms and the assessment priorities at the system level.

The analysis showed that the (1-s) values of the items were relatively high, indicating that the items were accurately matched to the intended attributes, supporting the validity of the constructed Q-matrix. The descriptive statistics based on the DINA model showed different probabilities of students' mastery of the different attributes. For example, attribute 4 had the highest probability of mastery (0.83), while attribute 13 had the lowest (0.59), suggesting different levels of student performance. This variation may be attributed to factors such as insufficient instructional emphasis on certain skills, cognitive complexity of specific attributes, or a lack of student familiarity with particular item formats.

These findings reinforce the value of CDMs in providing fine-grained diagnostic information about student learning. Mathematics as a discipline involves multiple cognitive processes, and the CDMs provide a way to assess these processes in a more nuanced and informative way than traditional total scores. The DINA model, despite its dichotomous nature (mastery vs. non-mastery), proves useful for analysing multidimensional constructs within educational assessments. However, recent studies (e.g. Delafontaine et al., 2022; Kalkan & Toprak, 2022)

recommend including multiple CDMs to improve the robustness of the results and to investigate model-data fit more comprehensively.

Another important aspect is the development of the Q-matrix. While this study relied on expert judgment, the literature emphasizes that the creation of the Q-matrix is a complex and error-prone process that significantly affects diagnostic accuracy (Li et al., 2021). This complexity may stem from the abstract nature of cognitive attributes, inconsistencies in expert interpretations, or the inherent subjectivity in mapping items to attributes. Recent research suggests combining expert opinion with statistical refinement methods to construct more reliable Q-matrices. In addition, estimation methods such as the Stone test or Posterior Predictive Model Checking (PPMC) can further validate model fitting. Wu et al. (2020) also emphasize the importance of attribute granularity and how it affects the accuracy of diagnostic inferences, especially for tests with a high degree of difficulty.

One of the greatest strengths of CDMs, including the DINA model, is their ability to monitor student development by providing detailed feedback. When applied to large-scale assessments, these models can reveal patterns of performance across years and inform instructional decisions based on students' specific skill profiles. This makes CDMs particularly suitable for formative assessment and educational policy monitoring. In this context, the DINA model is characterized by its interpretability and simplicity, making it a practical choice for educational researchers and practitioners alike.

To summarise, this study has demonstrated the potential of the DINA model for a detailed cognitive diagnostic analysis of the CEE mathematics subtest. The results demonstrate the effectiveness of CDMs in identifying students' strengths and weaknesses at the attribute level, which is crucial for both instructional design and policy decisions. Recent literature also emphasizes the importance of rigorous Q-matrix construction and model selection to improve diagnostic accuracy and interpretability. These findings suggest that the inclusion of CDMs in high-stakes tests such as the CEE could promote more targeted and equitable educational support strategies.

The findings of this study suggest several practical recommendations that are specific to the Turkish Cypriot education system. Firstly, exam developers and curriculum planners in Northern Cyprus should ensure better alignment between the national maths curriculum and challenging examinations such as the CEE. The lack of questions on certain areas of the curriculum (e.g. measurement and data) indicates a gap in content validity that could penalize students whose strengths lie in these areas. Secondly, schools and education authorities may consider using diagnostic assessment tools such as CDMs as part of formative assessment practices, not only for high-stakes tests but also for regular classroom assessments. This would help teachers to better monitor students' mastery of specific cognitive skills and adjust planning and delivery of instruction accordingly, reducing reliance on costly and high-pressure exam preparation processes currently prevalent in the region.

Ethics Committee Approval: Ethical permission was obtained from the institution Eastern Mediterranean University (06.12.2022-ETK00-2022-0283) for this research.

Peer-review: Externally peer-review.

Author Contributions: The authors contributed equally to the research, authorship, and publication of this article.

Conflict of Interest: The author(s) declare no potential conflicts of interest related to the research, authorship and publication of this article.

Financial Disclosure: The author(s) has received no financial support for the research, authorship and publication of this article.

Notice of Use of Artificial Intelligence: The author(s) utilised Grammarly and DeepL artificial intelligence tool(s) for the research, authorship and publication of this article.

References

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. <https://doi.org/10.1109/tac.1974.1100705>
- Birenbaum, M., Tatsuoka, C., & Yamada, T. (2004). Diagnostic assessment in TIMSS-R: Between countries and within-country comparisons of eighth graders' mathematics performance. *Studies in Educational Evaluation*, 30(2), 151-173. <https://doi.org/10.1016/j.stueduc.2004.06.004>
- Burkhardt, H., & Swan, M. (2012). Designing assessment of performance in mathematics. *Educational Designer*, 2(5), 1-40.
- Choi, K. M., Lee, Y. S., & Park, Y. S. (2015). What CDM can tell about what students have learned: An analysis of TIMSS eighth grade mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(6), 1563-1577. <https://doi.org/10.22034/EMES.2021.244398>
- Clements, D. H., Banse, H., Sarama, J., Tatsuoka, C., Joswick, C., Hudyma, A., ... & Tatsuoka, K. K. (2022). Young children's actions on length measurement tasks: strategies and cognitive attributes. *Mathematical thinking and learning*, 24(3), 181-202. <https://doi.org/10.1080/10986065.2020.1843231>
- DeCarlo, L. T. (2010). On the statistical and theoretical basis of signal detection theory and extensions: Unequal variance, random coefficient, and mixture models. *Journal of Mathematical Psychology*, 54(3), 304–313. <https://doi.org/10.1016/j.jmp.2010.01.001>
- DeCarlo, L.T. (2011). On the analysis of fraction subtraction data: The DINA model, classification, latent class sizes, and the q-matrix. *Applied Psychological Measurement*, 35(1), 8-26. <https://doi.org/10.1177/0146621610377081>
- Delafontaine, A., Tang, C. Y., & De Boeck, P. (2022). Cross-cultural item analysis using cognitive diagnosis models: A Q-matrix approach. *International Journal of Testing*, 22(2), 145–165. <https://doi.org/10.1080/15305058.2021.1972641>
- 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(4), 343–362. <https://doi.org/10.1111/j.1745-3984.2008.00069.x>
- de La Torre, J. (2009). A cognitive diagnosis model for cognitively based multiple-choice options. *Applied Psychological Measurement*, 33(3), 163–183. <https://doi.org/10.1177/0146621608320523>
- de la Torre, J. (2009). DINA model and parameter estimation: A didactic. *Journal of Educational and Behavioral Statistics*, 34(1), 115–130. <https://doi.org/10.3102/1076998607309474>
- de la Torre, J. (2011). The generalized DINA model framework. *Psychometrika*, 76(2), 179–199. <https://doi.org/10.1007/s11336-011-9207-7>
- de la Torre, J., & Douglas, J. A. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika*, 69(3), 333– 353. <https://doi.org/10.1007/BF02295640>
- de la Torre, J., & Douglas, J. A. (2008). Model evaluation and multiple strategies in cognitive diagnosis: An analysis of fraction subtraction data. *Psychometrika*, 73(4), 595-624. <https://doi.org/10.1007/s11336-008-9063-2>
- de la Torre, J., & Minchen, N. (2014). Cognitively diagnostic assessments and the cognitive diagnosis model framework. *Psicología Educativa*, 20(2), 89–97. <https://doi.org/10.1016/j.pse.2014.11.001>

- de la Torre, J., Chiu, C.-Y., & Lee, Y.-S. (2022). Advances in cognitive diagnostic models: New techniques for applying the DINA model in large-scale assessments. *Psychometrika*, 87(4), 901–926. <https://doi.org/10.1007/s11336-022-09863-3>
- Dogan, E., & Tatsuoka, K. (2008). An international comparison using a diagnostic testing model: Turkish students' profile of mathematical skills on TIMSS-R. *Educational Studies in Mathematics*, 68(3), 263–272. <https://doi.org/10.1007/s10649-007-9099-8>
- Doust, A. R., Khan, W. A., & Alghafri, M. S. (2021). Application of DINA model to assess Omani and Iranian fourth-grade students' mathematics' attributes. *North American Academic Research (Naar) Journal*, 4(10), 220–237. <https://doi.org/10.5281/zenodo.5639812>
- Henson, R., & Douglas, J. (2005). Test construction for cognitive diagnosis. *Applied Psychological Measurement*, 29(4), 262–277. <https://doi.org/10.1177/0146621604272623>
- Evrar, D. (2019). An application of cognitive diagnosis modeling in TIMSS: a comparison of Intuitive definitions of Q-matrices. *International Journal of Modern Education Studies*, 3(1), p.04. <https://doi.org/10.51383/ijonmes.2019.33>
- George, A. C., & Robitzsch, A. (2015). Cognitive diagnosis models in R: A didactic. *The Quantitative Methods for Psychology*, 11(3), 189–205. <https://doi.org/10.20982/tqmp.11.3.p189>
- Gündüz, T., & Çakan, M. (2020). TIMSS 2015 Türkiye örnekleminde matematik başarı testine dayalı olarak öğrencilerin bilişsel tanı modelleri ile sınıflandırılması. *Necatibey Eğitim Fakültesi Elektronik Fen ve Matematik Eğitimi Dergisi*, 14(2), 1040–1079.
- Hidajat, F. A. (2021). Students creative thinking profile as a high order thinking in the improvement of mathematics learning. *European Journal of Educational Research*, 10(3), 1247–1258.
- Junker, B.W. & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25(3), 258–272. <https://doi.org/10.1177/01466210122032064>
- Kalkan, Ö., & Toprak, M. (2022). The impact of Q-matrix refinement methods on the diagnostic accuracy of CDMs. *Educational Measurement: Issues and Practice*, 41(1), 52–64. <https://doi.org/10.1111/emip.12464>
- Karagiannakis, G. N., Baccaglini-Frank, A. E., & Roussos, P. (2016). Detecting strengths and weaknesses in learning mathematics through a model classifying mathematical skills. *Australian Journal of Learning Difficulties*, 21(2), 115–141. <https://doi.org/10.1080/19404158.2017.1289963>
- Lee, Y. S., Park, Y. S., & Taylan, D. (2011). Cognitive diagnostic modeling of attribute mastery in Massachusetts, Minnesota, and the US national sample using the TIMSS 2007. *International Journal of Testing*, 11(2), 144–177. <https://doi.org/10.1080/15305058.2010.534571>
- Leighton, J. & Gierl, M., (eds.). (2007). *Cognitive diagnostic assessment for education: theory and applications*. Cambridge, MA: Cambridge University Press. <https://doi.org/10.1017/CBO9780511611186>
- Li, X. & Wang, W. C. (2015). Assessment of differential item functioning under cognitive diagnosis models: The DINA model example. *Journal of Educational Measurement*, 52(1), 28–54. <https://www.jstor.org/stable/43940553>

- Li, L., An, Y., Ren, J., & Wei, X. (2021). Research on the cognitive diagnosis of Chinese listening comprehension ability based on the G-DINA model. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.714568>
- Li, H. & Suen, H. K. (2013). Constructing and validating a q-matrix for cognitive diagnostic analyses of a reading test, *Educational Assessment*, 18(1), 1-25. <https://doi.org/10.1080/10627197.2013.76152>
- Liu, J., & Kang, T. (2023). Advances in cognitive diagnostic models: The role of DINA in educational assessment. *Educational Assessment*, 29(1), 45–67. <https://doi.org/10.1007/s11336-022-09762-9>
- Ma, W., & de la Torre, J. (2016). A sequential cognitive diagnosis model for polytomous responses. *British Journal of Mathematical and Statistical Psychology*, 69(3), 253-275. <https://doi.org/10.1111/bmsp.12070>
- Parlak, B. (2018). *Investigation of the TIMMS 2015 mathematics achievement in eight grade students with cognitive diagnostic model*. (Publication No.494335) (Phd thesis, Hacettepe university) Council of Higher Education National Thesis Center).
- Parlak, B. (2023). Matematik performansının bilişsel tanı modeli ile değerlendirilmesi: TIMSS Türkiye ve Singapur örneği. *Milli Eğitim Dergisi*, 52(1), 413-436.
- Roussos, L. A., DiBello, L. V., Stout, W., Hartz, S. M., Henson, R. A., & Templin, J. L. (2007). *The Fusion Model Skills Diagnosis system*. In Cambridge University Press eBooks (pp. 275–318). <https://doi.org/10.1017/cbo9780511611186.010>
- Rupp, A. A. & Templin, J. L. (2008). Unique characteristics of diagnostic classification models: A comprehensive review of the current state of the art. *Measurement*, 6(4), 219-262. <https://doi.org/10.1080/15366360802490866>
- Sawaki, Y., Kim, H. J., & Gentile, C. (2009). Q-Matrix construction: Defining the link between constructs and test items in large-scale reading and listening comprehension assessments. *Language Assessment Quarterly*, 6(3), 190–209. <https://doi.org/10.1080/15434300902801917>
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6, 461-464.
- Shi, Q., Ma, W., Robitzsch, A., Sorrel, M. A., & Man, K. (2021). Cognitively diagnostic analysis using the G-DINA model in R. *Psych*, 3(4), 812–835. <https://doi.org/10.3390/psych3040052>
- Su, Y. L. (2013). *Cognitive diagnostic analysis using hierarchically structured skills*. The University of Iowa.
- Tatsuoka, K. K. (1983). Rule space: An approach for dealing with misconceptions based on item response theory. *Journal of Educational Measurement*, 20(4), 345-354. <https://doi.org/10.1111/j.1745-3984.1983.tb00212.x>
- Tatsuoka, K. (1985). A probabilistic model for diagnosing misconceptions in the pattern classification approach. *Journal of Educational Statistics*, 10(1), 55–73. <https://doi.org/10.2307/1164930>
- Tatsuoka, K. K. (1995). Architecture of knowledge structures and cognitive diagnosis: A statistical pattern recognition and classification approach. *Cognitively Diagnostic Assessment*, 327-359.
- Tatsuoka, K.K. (2009). *Cognitive assessment—an introduction to the rule space method*. Routledge, New York. <https://doi.org/10.4324/9780203883372>

- Tatsuoka, C., Varadi, F., & Jaeger, J. (2013). Latent partially ordered classification models and normal mixtures. *Journal of Educational and Behavioral Statistics*, 38(3), 267-294. <https://doi.org/10.3102/1076998612458318>
- Tatsuoka, C., Clements, D. H., Sarama, J., Izsák, A., Orrill, C. H., de la Torre, J., ... & Tatsuoka, K. (2016). Chapter 4. *Journal for Research in Mathematics Education. Monograph*, 15, 73-96.
- Templin, J. L., & Henson, R. A. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological methods*, 11(3), 287. <https://doi.org/10.1037/1082-989X.11.3.287>
- Terzi, R., & Sen, S. (2019). A nondiagnostic assessment for diagnostic purposes: Q-Matrix validation and Item-Based Model fit evaluation for the TIMSS 2011 assessment. *SAGE Open*, 9(1). <https://doi.org/10.1177/2158244019832684>
- Yakar, L., Dogan, N. Senol, D., & Yuksel, N. S. (2021). Monitoring student achievement with cognitive diagnosis model. *Journal of Measurement and Evaluation in Education and Psychology*, 12(3), 303-320. <https://doi.org/10.21031/epod.903084>
- Wang, L., Li, Y., & Zhang, J. (2023). A new approach to Q-matrix construction for cognitive diagnostic models: Combining expert knowledge with statistical methods. *Psychometrika*, 88(1), 123–145.
- Whitely, S. E., & Schneider, L.M. (1981). Information structure for geometric analogies: A test theory approach. *Applied Psychological Measurement*, 5, 383-397. <https://doi.org/10.1177/014662168100500312>
- Wu, H., Liu, Y., & Zhang, X. (2020). The effects of attribute granularity on the estimation quality of cognitive diagnosis models. *Applied Psychological Measurement*, 44(7), 546–561. <https://doi.org/10.1177/0146621620917447>
- Zhang, S., Liu, J., & Ying, Z. (2023). *Statistical applications to cognitive diagnostic testing*. Annual Review of Statistics and its Application, 10(1), 143–165. <https://doi.org/10.1146/annurev-statistics-033021-111803>

İletişim/Correspondence

Dr. Aygıl TAKIR
aygil.takir@aoa.edu.tr

Prof. Dr. Osman CANKOY
cankoy@metu.edu.tr

Doç. Dr. Hasan ÖZDER
hasan.ozder@aoa.edu.tr