



Yanıt Sürelerinin Karma Modellemeye Dâhil Edilmesi PISA 2022 Matematik Okuryazarlığı Örtük Gruplarının Belirlenmesini Nasıl Etkiler? *

Halime Yıldırım Hoş¹, Menekşe Uysal Saraç²

Öz

Bu çalışmada, PISA 2022 Türkiye uygulamasında matematik okuryazarlığı testini yanıtlayan öğrencilerin yetenek ve yetenek dışı örtük sınıfları, karma madde yanıt modeliyle belirlenmiştir. Jeon ve De Boeck'un (2019) önerdiği karma madde tepki kuramı modelleme yaklaşımı doğrultusunda, yanıt süreleri ile madde zorluğu ve başarı olasılıkları arasındaki ilişkiler incelenmiştir. Çalışmada hiyerarşik olarak dört farklı modelin karşılaştırması yapılmıştır. Bu modellerin ilki tek sınıflı iki parametrelili madde yanıt teorisi (2PL IRT) modeli (Model I), ikincisi ise (Model II) yetenek sınıfı ile tahmin sınıfı olarak adlandırılan ve başarı olasılığı 0.25'e sabitlenen iki sınıflı modeldir. Diğer iki sınıflı modelde (Model III), tahmin grubunun başarı olasılığı serbestçe kestirilmiştir. Son model, Jeon ve De Boeck'un (2019) önerdiği yaklaşım doğrultusunda yanıt süresi bilgisini kovaryant olarak içeren iki sınıflı bir modeldir (Model IV). Analiz sonucunda, Model IV (yanıt süresinin kovaryant olarak dâhil edildiği iki sınıflı model) en iyi uyumu sağlayan model olarak belirlenmiştir. Yetenek dışı sınıfta ortalama madde yanıt süreleri ve başarı olasılıkları düşük olma eğilimindeyken, yetenek sınıfında bu değerlerin daha yüksek olduğu gözlenmiştir. Bununla birlikte, zamanı daha etkili kullanan (başarı olasılıkları yüksek) yetenek sınıfı, kolay maddelere hızlı yanıt verirken, zor maddelere daha fazla zaman ayırarak başarılı olmuştur. Yetenek dışı sınıfın genel olarak düşük performansı dikkat çekmiş olup kolay maddelerde daha hızlı yanıt vermeleri başarısızlıkla sonuçlanmış, buna karşılık, zor maddelere daha fazla zaman ayırarak kısmen başarılı oldukları belirlenmiştir. Bu grubun, daha yüzeysel bir yaklaşım benimsediği düşünülmektedir. Tüm maddelerde yetenek sınıfına göre daha hızlı yanıt veren, ancak zor maddelerde daha fazla süre harcayarak dikkatli olma eğiliminde olan bir tür madde yanıtlama stratejisini kullanmaktadır.

Anahtar Kelimeler

Madde yanıt süreleri
Bilgi geri çağırma stratejisi
Hızlı tahmin davranışı
Karma madde tepki kuramı
PISA

Makale Hakkında

Gönderim Tarihi: 06.10.2024
Kabul Tarihi: 24.01.2025
Elektronik Yayın Tarihi: 03.03.2025

DOI: 10.15390/EB.2025.14125

* Bu çalışmanın bir bölümü 4-6 Ekim 2024 tarihleri arasında düzenlenen Uluslararası Ölçme, Seçme ve Yerleştirme Sempozyumu'nda sözlü bildiri olarak sunulmuştur.

¹ İstanbul Medeniyet Üniversitesi, Eğitim Bilimleri Fakültesi, Eğitim Bilimleri Bölümü, Türkiye, halime.yldrm@gmail.com

² Çankırı Karatekin Üniversitesi, İnsan ve Toplum Bilimleri Fakültesi, Eğitim Bilimleri Bölümü, Türkiye, menekseysl@gmail.com

Giriş

Yeteneği ölçmek için sıklıkla kullanılan bilişsel testlerde, ölçülen aslında performans seviyeleridir ve süreçler bu performans sonucuna ulaşmada yer alan faaliyetlerdir (De Boeck ve Jeon, 2019). Özellikle küçük ölçekli sınavlar genellikle yetenek ölçümlerine odaklanırken madde yanıt süresi ile ilgili ölçüm sağlamazlar. Süreçler hakkında bilgi toplamadan da yeteneği ölçmek mümkündür, ancak bu sadece performans düzeyini belirtir. Performans önemli bir faktör olmakla birlikte yanıt süreçlerine dair ek veri (madde yanıt stratejisi, süresi vb.) toplamak performans düzeyini göstermenin ötesine geçer. Bu veri, öğrencilerin nasıl düşündüklerini ve hangi stratejileri kullandıklarının anlaşılmasına yardımcı olabilir. Yanıt süreçleri ile ilgili bu veri bize açıklama, bir şeyin nasıl meydana geldiğine dair bir anlatı sağlar. Dolayısıyla öğrenci yanıt stratejisini anlamamıza yardımcı olurken aynı zamanda müdahale ve iyileştirmeler konusunda da fikir verir.

Bilgisayar tabanlı, geniş ölçekli örneğin PISA, TIMSS gibi uluslararası sınavlara giren bireylerin davranışları hakkında otomatik olarak veri toplamak amacıyla kullanılan süreç verisi giderek daha yaygın hale gelmektedir (Anghel, Khorramdel ve von Davier, 2024). Bu tür bilgiler, sınav katılımcılarının teste veya test maddelerine yanıt verme süresini, neye tıkladıklarını, yazma sıralarını vb. içerir. Bu bilgiler, katılımcıların bir teste yaklaşımının altında yatan bilişsel süreçlere ilişkin mevcut bir teoriye dayanıyorsa madde tasarımının iyileştirilmesine (Brückner ve Pellegrino, 2017), sınav katılımcılarının katılımının belirlenmesine ve yapı hakkında çıkarımlarda bulunulmasına yardımcı olabilir (Oranje, Gorin, Jia, Kerr, Ercikan ve Pellegrino, 2017). Bu şekilde, süreç verisi değerlendirmenin geçerliğine dair kanıt sağlayabilir (AERA, APA ve NCME, 2014). Ayrıca, farklı test alma stratejileri ışığında başarı farklılıklarının anlaşılmasına ve yeniden çerçevelendirilmesine yardımcı olabilir (Pohl, Ulitzsch ve von Davier, 2021).

Değerlendirme süreçlerinde bireylerin yeteneklerinin doğru kestirilmesi üzerinde etkili çok sayıda faktör vardır. Erwin ve Wise'a (2002) göre, düşük çaba, bir kişinin yeteneğinin doğru bir şekilde kestirilmesinin önündeki en belirgin engeldir. Cevaplayıcı bir sınav aldığı anda, kendisine not vb. bir katkısı yoksa sınav esnasındaki çabalamaya eksikliği, sınavın geçerliğine bir tehdit olarak algılanmaktadır (Baumert ve Demmrich, 2001; Eklöf, 2010; Finn, 2015; Wise, Pastor ve Kong, 2009). Sonuçta, bir testte iyi performans göstermek için cevaplayıcının bilgi ve becerisinin yeterli düzeyde olmasına ve teste aktif şekilde katılıp ilgilenecek kadar motivasyona ihtiyacı vardır (Eklöf, 2010). Haladyna ve Downing'e (2004) göre aynı yeterlik düzeyindeki iki öğrenciden daha düşük motivasyona sahip olanın yetenek kestiriminin diğerinden daha düşük olması, aslında bu iki öğrenci arasındaki yeterlik farkını değil, motivasyonları arasındaki farkı yansıtmaktadır. Burada sözü edilen sınav motivasyonu başarı motivasyonunun özel bir halidir ve sınava giren kişinin iyi bir performans gösterme motivasyonudur (Baumert ve Demmrich, 2001; Eklöf, 2010). Sınav alma motivasyonu test maddeleri üzerinde çalışma, çaba gösterme ve çözüme ısrarcı olmaktır. Bu motivasyon, cevaplayıcıların sınavı tamamlamak için gösterdikleri çaba ve kullandıkları madde yanıt stratejileri gibi bilişsel stratejilere de yansır (Baumert ve Demmrich, 2001; Brophy ve Ames, 2005). PISA'da veya diğer benzer araştırmalarda, öğrenciler genellikle matematik, okuma ve fen bilimleri alanlarındaki yeterliklerini ölçmek üzere tasarlanmış bir teste girmeye davet edilir. Öğrencilerin katılımı zorunlu değildir ve performanslarının kendileri için doğrudan bir sonucu yoktur: bu nedenle test, katılımcı düzeyinde düşük riskli bir değerlendirmedir (Baumert ve Demmrich, 2001; Finn, 2015). Bu sınavlar, öğrenciye not ya da bireysel bir geri bildirim sunmadığı için, öğrencinin test alma motivasyonunun düşük olması beklenmektedir. Dolayısıyla süreç verisi sağlayan bu sınavlarda öğrenci çabasına, dolaylı olarak motivasyonuna ve kullandıkları yanıt stratejisine ilişkin çıkarım yapmak oldukça önemlidir.

Sınav katılımcılarının farklı madde-yanıt stratejileri, eğitim araştırmacılarının uzunca bir süredir ilgisini çekmektedir. Örneğin Mitlevy ve Verhelst (1990), bireylerin bir sınava girerken birbirinden bağımsız sonlu sayıda madde-çözüm stratejisinden birini uyguladığını ve bu stratejinin tepki örüntülerine dayalı olarak kestirildiği bir model önermiştir. Yamamoto (1989), benzer bir sonlu karma yapısına sahip olan ve bireylerin çoklu madde çözüm stratejilerini ayırt etmek için de kullanılabilen bir HYBRID modeli önermiştir. Araştırmacılar, zaman kısıtlamaları altında uygulanan

testler sırasında sıklıkla ortaya çıkan, sınav katılımcılarının tahmin stratejisine özellikle dikkat etmişlerdir (Jeon ve De Boeck, 2019). Mislevy ve Verhelst (1990) ve Yamamoto (1997) sonrasında modellerini rastgele tahmin (random guessing) madde-yanıt stratejisini araştırmak için uygulamışlardır. Bir dizi araştırmacı, hız testlerinde zaman baskısı altındaki katılımcıların tahmin stratejilerini yakalamak için çeşitli modelleme yaklaşımları sunmuştur (Bolt, Cohen ve Wollack, 2002; Cao ve Stokes, 2008; Chang, Tsai ve Hsu, 2014; Wang ve Xu, 2015). Düşük riskli değerlendirmelerde motivasyonsuz sınav katılımcıları da tahmin stratejisi uygulayabilir (Pokropek, 2016; Wise ve Kong, 2005; Wise ve DeMars, 2006). Bu çalışmalarda tahmin etme, genellikle kişinin yeteneğine dayanmayan bir madde yanıt stratejisi olarak tanımlanır. Bu modellerde tahmin bir çözüm stratejisine dayanmadığı için genellikle hızlı bir süreç olduğu, tahmin stratejisinde doğruluğun şans eseri beklenen doğruluk seviyesinde ya da altında olması ve normal bir madde çözüm stratejisinden ayırt edilebilen yalnızca bir tür tahmin stratejisi olduğu varsayılır (Jeon ve De Boeck, 2019). Hızlı tahmin etmenin altında yatan temel fikir, sınav katılımcılarının rastgele bir yanıtı hızlı bir şekilde girerek uygulanan bir maddeden kurtulmak istemeleridir. Buna karşılık, sınava katılanlar bilgi, beceri ve yeteneklerini kullanarak maddeyi okuyacak, madde zorluğunu aşarak bir yanıt göndererek çözüm davranışı da sergileyebilirler. Çünkü, kavramsal olarak, çözüm davranışı hızlı tahmin etme davranışından daha fazla zaman gerektirir (Wise, 2019).

Yanıt stratejilerinin, yalnızca madde yanıtlarının doğruluklarıyla değil, aynı zamanda yanıt süresine dayalı olarak da karakterize edildiği düşünülebilir (Jeon ve De Boeck, 2019). Büyük ölçekli, düşük riskli değerlendirmelerde sınav katılımcısının sınavdan kopması; teste katılımının azalması, test süresi boyunca performansın düşmesi olasılığı da dâhil olmak üzere bir çok faktör geçerlik için bir tehdit oluşturmaktadır. Bu durum, araştırmalarda hızlı tahmin konusuna büyük ilgi gösterilmesine neden olmuştur (Maddox, 2023). Örneğin madde yanıtları ve sürelerine bakılarak düşük başarı olasılığına sahip bir grubun hızlı ya da kısa sürede yanıt vermesi o grubun "hızlı tahmin stratejisi" olarak adlandırılmasına neden olabilir. Ancak zor bir maddeye hızlı yanıt vermek, kişinin olumsuz bir olaydan kaçınma motivasyonunun, hoş olmayan bir durumdan uzak durarak öz saygısını koruma isteğinin, çaba eksikliğinin veya testlerle başa çıkmaya yönelik sığ bir yaklaşımı benimsediğinin göstergesi olabilir (Sideridis, Tsaousis ve Al-Harbi, 2022). Bu durumda, yanıt süreleri "tahmin stratejisi" sınıfının kestirilmesinde bilgi verici bir kovaryant değişken ya da yordayıcı olduğu anlamına gelir. Ancak Jeon ve De Boeck'un (2019) da öne sürdüğü gibi yanıt süreleri ortak değişken olarak kullanıldığında, yeteneğe dayalı kestirimin yapıldığı olağan (regular) sınıfın -yetenek sınıfı (ability)- tek alternatifi bu hızlı tahmin sınıfı olmayabilir.

Alternatif bir stratejide katılımcılar, oldukça kolay maddelere hızlı ve doğru yanıtlar verecektir; ancak hem bilgi hem de muhakeme süreçlerine ihtiyaç duyulan daha zor maddelerde daha fazla zaman harcayacaklardır. Bu strateji "bilgiyi geri çağırma" olarak adlandırılmaktadır (Jeon ve De Boeck, 2019; Sideridis vd., 2022). Bilgiyi geri çağırma stratejisi, özellikle daha kolay maddeler için nispeten kısa yanıt süresi ve yüksek başarı olasılıkları, daha zor maddelerde ise nispeten daha fazla zaman gerektirecek ve düşük bir başarı oranına sahip olacaktır; çünkü doğru yanıt almak bilgi almanın ötesinde muhakeme süreçlerini de gerektirmektedir. Bloom'un taksonomisinde (Bloom, Engelhart, Furst, Hill ve Krathwohl, 1956) bilgi, yetenek veya beceri gelişiminin en düşük seviyesi olarak kabul edilir; dolayısıyla, gelişimin daha ileri aşamalarındaki deneklerin sadece hazır bilginin geri getirilmesinden daha karmaşık stratejiler kullanması beklenebilir. Bilginin kaynağından bağımsız olarak, bu bilgi, bilgi tabanının (knowledge base) yüzeyinde mevcut olabilir; dolayısıyla, bu bilgiye dayalı stratejide hızlı ve doğru cevap verme olasılığı yüksek olacaktır. Diğer durumlarda, kişinin diğer bilgi parçalarından yaptığı çıkarımlara dayalı olarak bir yanıtın oluşturulması veya yanlış yanıtların elenmesi gerekir. Bu, daha derin bir işleme türünü gerektirir. Çünkü doğru yanıt kişinin bilgi tabanının yüzeyinde mevcut değildir. Söz konusu "yüzey ve derin" terimleri alanyazında öğrenme ve anlama ile ilişkilendirilmektedir (Entwistle ve Peterson, 2004; De Jong ve Ferguson-Hessler, 1996). Bilgi geri çağırma ile ilgili öne çıkan diğer kavram ise otomatik ve kontrollü işlemdir (Shiffrin ve Schneider, 1977). Hazır bilginin alınması otomatik işleme ile tutarlıyken, kontrollü işleme kasıtlıdır, çaba gerektirir. Kontrollü işlemlerde çıkarımlar veya uygulamalar yapmak için daha fazla zaman gerekir. Bilgi geri çağırma stratejisi prensipte yüzeysel bilginin yanı sıra derin bilgi ile de kullanılabilir, ancak mevcut çalışmada daha çok kontrollü işlemeden ziyade otomatik işleme dayalı yorumlar yapılmıştır.

Bir soru çözüm stratejisi olan bilgiyi geri çağırma genellikle kişi olağan yeteneğini kullanmadığında, yanıt vermek için mutlaka hızlı bir süre beklenmediğinde ya da doğru cevap verme oranı şansa bağlı olandan daha yüksek olduğunda ortaya çıkabilir (Jeon ve De Boeck, 2019). Tahmin ise, bireyin yeteneğine dayanmayan bir madde çözme stratejisi olarak tanımlanır. Ayrıca, önceki araştırmalarda şu varsayımlar belirtilmiştir (Bolt vd., 2002; Chang vd., 2014; Wise ve DeMars, 2006; Wise ve Kong, 2005): (1) Tahmin genellikle hızlı bir süreçtir çünkü yeteneğe dayalı çözüm stratejisine kıyasla nispeten kısa bir zaman harcanır; (2) Tahmin stratejisindeki doğru cevap verme oranı, genellikle şansa bağlı olarak beklenen doğruluk seviyesinde ya da bu seviyenin altında olur; (3) Sadece bir tür tahmin stratejisi vardır ve bu strateji, düzenli bir madde çözüm stratejisinden ayırt edilebilir.

Sonuç olarak akademik başarıyı anlamak ve artırmak bireylerin test alma bağlamlarında, yani bir sınav esnasında hangi madde-yanıt stratejisi türünü kullandıklarını analiz etmeyi ve tanımlamayı gerektirir (Jeon ve De Boeck, 2019; Sideridis vd., 2022). Bu noktada önemli faktörlerden biri maddeye yanıt vermek için harcanan zamanın rolünü anlamaktır. Eğitimsel testlerdeki dijital dönüşüm, teknolojinin büyük ölçekli değerlendirmeleri geliştirmesi için birçok yeni fırsat sunmuştur. Bunlar arasında test katılımcısının yanıt süreçlerine ilişkin süreç verisini rutin olarak toplama ve kullanma potansiyeli de yer almaktadır. Süreç verisi uzun zamandır değerlendirmelerde önemli bir geçerlik kaynağı olarak kabul edilmektedir. Ayrıca değerlendirme döngüsü boyunca birden fazla amaç için kullanılmaktadır (Maddox, 2023). Cevaplayıcıların maddelerde geçirdiği süreler; gösterdiği çaba, test alma motivasyonu ve madde yanıt stratejisi vb. pek çok durumun işaretçisi olabilirken, bu veri yorumlanırken ya da analiz edilirken madde güçlüğünün de dikkate alınması önemlidir.

Yanıt Sürelerinin Karma Modellemeye Dahil Edilmesi

Bu çalışmada, bireylerin yanıt süresi (YS) verisinin karma modele dâhil edilmesi ile kestirilen örtük sınıflar ve yanıt süreleri aracılığıyla öğrenci çabasının yorumlanması amaçlanmaktadır. Çalışmada Jeon ve De Boeck (2019)'un modeli temel alınarak geliştirilen, öğrenci yanıtları ve YS'leri bir araya getiren model kullanılmıştır. Belirli bir grup (örtük sınıf) için, 2PL modeli uygulanarak, j kişinin i maddesine doğru yanıt verme olasılığı Y_{ij} aşağıdaki gibi kestirilir:

$$P(Y_{ij}=1|a_i, b_i, \theta_j, R_g) = \frac{e^{a_i(\theta_j - b_i)}}{1 + e^{a_i(\theta_j - b_i)}}$$

Denklemden ikili puanlanan bir maddede, başarılı olma olasılığı, a ayırt edicilik parametresinin, b madde zorluğunun ve θ örtük yeteneğe ilişkin kişi kestiriminin bir fonksiyonudur. R_g ise olağan grubu göstermektedir. Analizde kullanılan Mplus yazılımında, ayırt edicilik ve madde güçlüğü parametreleri, faktör yükleri kullanılarak madde ayırt ediciliklerine dönüştürülen madde faktör analizi kullanılarak kestirilmektedir:

$$a_i = \lambda_i \sqrt{f_{var}}$$

Denklemden λ_i madde faktör yükünü göstermektedir. $\sqrt{f_{var}}$ ise örtük faktörün varyansıdır ve model tanımlama amaçlı bire sabitlendiğinde faktör yükü madde ayırt ediciliğine eşittir. Madde güçlüğünü hesaplamak için ise eşik değerleri kullanılır:

$$\beta_i = \frac{\tau_i}{\lambda_i \sqrt{f_{var}}}$$

Denklemden τ_i , 2PL modeldeki eşik kestirimidir. Diğerleri bir önceki denklemde olduğu gibi λ_i madde faktör yükünü, $\sqrt{f_{var}}$ ise örtük faktörün varyansıdır.

Çalışmada ikincil sınıf (secondary class S_g) yeteneğin kestirimde kullanılmadığı yetenek dışı (non-ability) sınıftır. Bu sınıf için parametre kestirimi şu şekilde yapılır:

$$P(Y_{ij}=1|\delta_i, S_g) = \frac{e^{(-\delta_i)}}{1 + e^{(-\delta_i)}}$$

Denklemden δ_i , ikincil sınıf için sabit (intercept) parametresidir ve ilgili parametrenin eksi log değeridir. Böylece, j kişinin S_g (ikincil) sınıftaki başarı olasılığı, δ_i madde güçlüğünün bir fonksiyonudur ancak bu eşitlikte olasılık fonksiyonuna yetenek parametresi (θ_j) dâhil edilmez. Bireyin ikincil, yani yetenek dışı örtük sınıfa atanma olasılığı multinomial regresyon modeli kullanılarak aşağıdaki gibi kestirilir:

$$P(S_g) = \frac{e^{(\gamma_0 - \sum_i \gamma_1 * RT_i)}}{\sum_{u=1}^{1+s} e^{(\gamma_0 - \sum_i \gamma_1 * RT_i)}}$$

Son olarak bu denklemden γ_0 ve γ_1 sırasıyla sabit ve eğim parametreleridir. Burada dikkat çeken parametre γ_1 parametresidir. Bu parametre yanıt sürelerinin S_g sınıfının oluşmasına katkı sağlayıp sağlamadığını gösterir. Başka bir ifadeyle soru üzerinde daha fazla/az zaman harcamanın bireyin S_g sınıfına atanmasında önemli bir bilgi verip vermediğini yorumlamamızda yardımcı olur. Eğer γ_1 katsayıları tüm maddelerde negatifse, bu hızlı tahmin veya hızlı yanıt veren (fast responder) bir grubun göstergesi olabilir. γ_1 katsayısının pozitif olması, bir madde üzerinde daha fazla zaman harcamanın S_g grubuna atanma olasılığını artırdığını ve bunun tüm maddelerde gerçekleşmesi durumunda muhtemelen (zamanın niteliksel yönlerini bilmiyoruz) düşünceli ve dikkatli bir gruptandırma işaret edeceğini gösterir. S_g grubuna atanmanın yalnızca YS 'lere dayandığı ve yetenekten bağımsız olduğu göz önüne alındığında, bu gizil grupların "Yetenek dışı" gruplar olarak adlandırılmasının nedeni; yetenek seviyelerine sahip olmamaları değil, belirli gizil sınıfa (S_g) atanmalarında yetenek parametresinin kullanılmamasıdır. Yetenek dışı/ikincil sınıf terimi, belirli bir yetenek türünü ölçmek için tasarlanmış test maddeleri üzerinde beklenen şekilde çalışan bir yanıtlayıcı sınıfına karşı olarak kullanılmaktadır (Jeon ve De Boeck, 2019). Standart madde yanıt analizinde, tüm test katılımcılarının test maddelerine yeteneklerine dayalı olarak yanıt verecekleri varsayılır. Bu anlamda, test maddelerini çözerken yeteneğe dayanmayan bir katılımcı sınıfı (bu çalışmada varsayılan ve incelenen), "olağan/düzenli" yetenek sınıfına karşı "ikincil" bir sınıf olarak görülmektedir.

Bu çalışma, Jeon ve De Boeck (2019) tarafından önerilen karma modelleme yaklaşımı kullanılarak, PISA 2022 matematik okuryazarlığı testini alan bir grup Türk öğrencinin yetenek ve yetenek dışı örtük sınıflarını belirlemeyi amaçlamaktadır. Araştırma, bireylerin yetenek dışı bir değişken olan yanıt sürelerine göre nasıl sınıflandırılabilirliğini inceleyerek bu alandaki modelleme çalışmalarına katkı sağlamayı hedeflemektedir. Bu çalışmada kurgulanan modelde, madde zorluğu ile yanıt süresi eğimleri arasındaki pozitif ilişki, bilgiyi geri çağırma stratejisini kullanan örtük bir sınıfa işaret edebilir. Öte yandan, düşük başarı olasılıkları ve hızlı madde yanıtlarıyla karakterize edilen sınıf, hızlı tahmin stratejisini kullanan örtük bir grubu da yansıtabilir. Bununla birlikte, model sonuçları beklenen sınıfların dışında, Jeon ve De Boeck (2019) ile Sideridis ve diğerleri (2022), çalışmalarında olduğu gibi farklı örtük sınıflar da ortaya çıkarabilir.

Yöntem

Araştırmanın Modeli

Bu araştırma, PISA 2022 Türkiye uygulamasında matematik okuryazarlığı testine katılan öğrencilerin, soru çözme stratejilerine bağlı olarak olağan ve ikincil sınıflar şeklinde modellenerek sınıflandırılmasını amaçlamaktadır. Bu amaçla, veriye en iyi uyum sağlayan modelleme çerçevesi temel alınarak özelleşen örtük sınıfların belirlenmesi ve yanıt süreleri ile madde zorlukları arasındaki ilişkilerin incelenmesi açısından betimsel bir araştırmadır.

Katılımcılar

Bu çalışma, PISA 2022 uygulaması kapsamında elde edilen Türkiye verisine odaklanmaktadır. PISA, dünya çapında zorunlu eğitimin sonuna gelen ve örgün eğitime devam eden 15 yaş grubu öğrencileri hedef almaktadır. Araştırmanın örneklemini oluşturulurken seçkisiz tabakalama yöntemi kullanılmaktadır ve örnekleme yer alacak tabakalara, Uluslararası Merkez ile PISA'ya katılan ülkeler birlikte karar vermektedir. İki aşamalı bir örnekleme sürecinin uygulandığı PISA araştırmasında, ilk aşamada Uluslararası Merkez her ülke için katılımcı okulları rastgele seçmektedir. İkinci aşamada ise, değerlendirmenin yapılacağı öğrenciler, katılımcı ülkeler tarafından Maple isimli bilgisayar programı yardımıyla yine rastgele belirlenmektedir (Millî Eğitim Bakanlığı, 2022). Bu çalışmanın örneklemini, PISA 2022 uygulamasında Türkiye'de en fazla katılımcı tarafından yanıtlanan 13, 14, 20 ve 24. kitapçıkları alan toplam 1180 öğrenciden oluşmaktadır.

Verinin Elde Edilmesi ve Ölçme Araçları

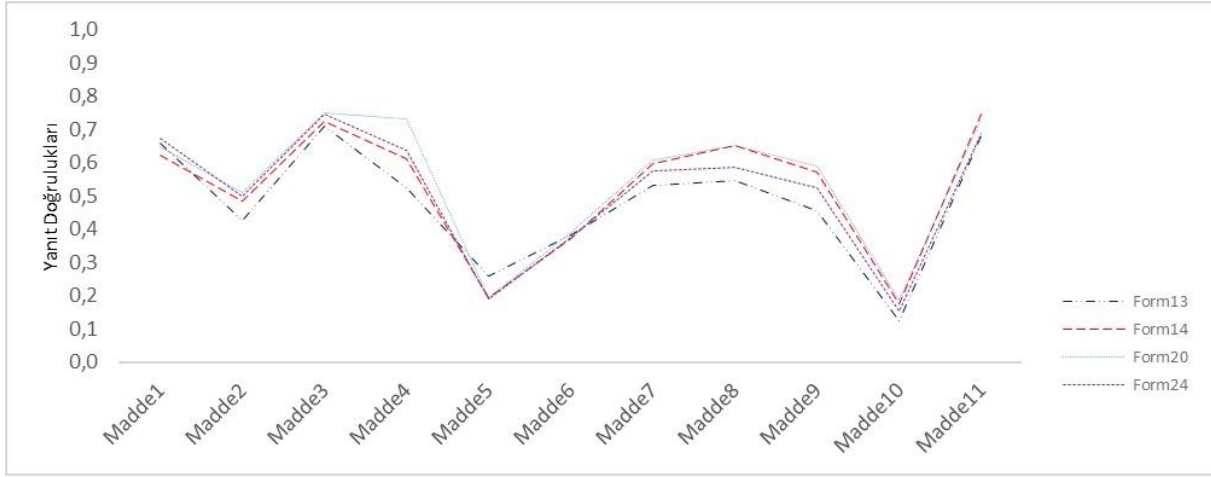
PISA kapsamında matematik okuryazarlığı, bireyin matematiksel düşünme yeteneğini kullanarak, gerçek yaşamda karşılaşılan problemleri çözme becerisi olarak tanımlanmaktadır. Bu çalışma, PISA 2022 bilişsel maddelerini içermekte olan ve erişime açık olan bir veri seti üzerinde gerçekleştirilmiştir. Kullanılan veri, OECD'nin veri tabanından (<https://www.oecd.org/en/data/datasets/pisa-2022-database.html>) indirilmiştir.

Test maddeleri, bilgi düzeyinden uygulama düzeyine kadar uzanan bir aralıkta, kolay ve zor maddelerin bir karışımından oluşmaktadır. Çalışmaya on bir çoktan seçmeli test maddesi dâhil edilmiştir. Sınav katılımcılarının büyük bir çoğunluğu tüm test maddelerine yanıt vermiştir. Atlanan yanıt oranı oldukça düşüktür ve veri setindeki maddelerin %90'ında erişilemeyen maddeler oranı %1'in altındadır.

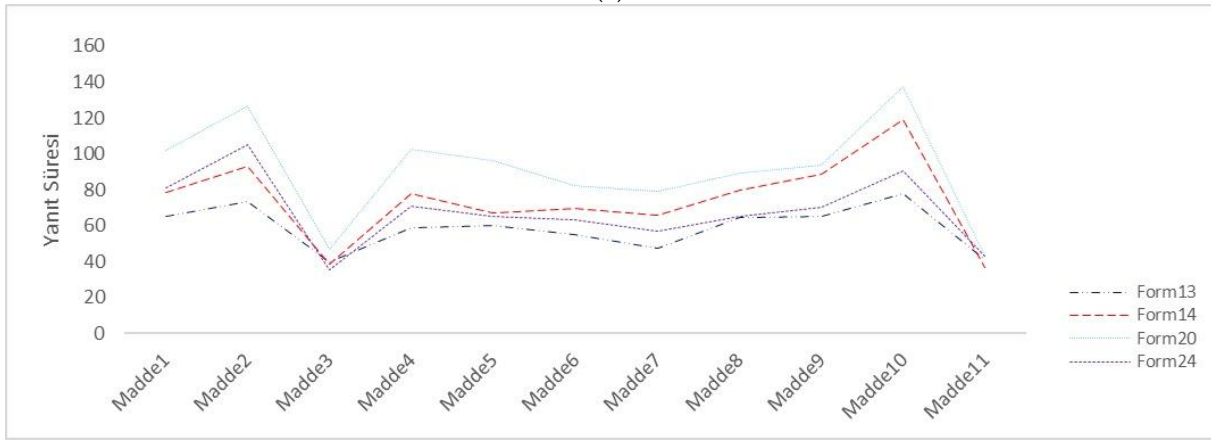
Veri Analizi

Çalışmada yer alan çoktan seçmeli maddeler yanıt doğruluğu ve yanıt süreleri açısından incelendiğinde, genelinde ortalama yanıt süresi 40.03 ile 106.049 (saniye) arasında değişmektedir (ortalama 72.42 ve medyan 63.93). Madde yanıt sürelerine ilişkin frekans dağılımı Ek 1'de verilmiştir. Tüm test maddeleri için ortalama doğru yanıt oranı yaklaşık %51.5 olup, %2 ile %73 arasında değişmiştir. Sınav katılımcılarının çoğunluğu tüm test maddelerine yanıt vermiştir. Atlanan yanıt oranı oldukça düşüktür, ayrıca veri setinde, maddelerin %90'ında erişilemeyen maddeler oranı %1'in altındadır.

Analize dâhil edilen formlardaki maddelerin yanıt süreleri ve yanıt doğruluklarının dağılımları gösterilmektedir (bkz. Şekil 1). Şekil 1(a), yanıt doğrulukları maddenin testteki konumlarına göre belirgin bir örüntü sergilemediğini ortaya koymaktadır. Şekil 1(b) incelendiğinde, maddelerin farklı konumlarda olmasına rağmen yanıt süresi dağılımlarının birbirine benzediği görülmektedir. Yanıt süresi dağılımları, hızlı yanıtlar içeren bir yanıt sınıfının varlığını açıkça gösteren iki modlu bir yapı sergilememektedir. Bu durum, yanıtların bir kısmının tahmin olabileceğini, ancak önemli bir kısmının bu şekilde olmadığını düşündürmektedir (Jeon ve De Boeck, 2019).



(a)



(b)

Not: Bu görsellerde yer alan madde sıralamaları, sadece veri etiketlemesinin bir sonucu olup, kitapçıklardaki gerçek yerleşimi göstermemektedir.

Şekil 1. Kitapçıklarda Yer Alan Maddelerin Yanıt Süreleri ve Yanıt Doğruluklarının Dağılımı

Veri analizinden önce, dağılımlardaki çarpıklığı gidermek amacıyla ham yanıt sürelerinin log dönüşümleri gerçekleştirilmiştir. Elde edilen log dönüşümlü yanıt süresi dağılımının medyanı/ortalaması 4.10/4.02, maksimum değeri ise 8.20 olarak belirlenmiştir. Ayrıca, log-dönüştürülmüş yanıt sürelerine kişi içi ve madde içi merkezileştirme uygulanmıştır. Yanıt sürelerine çift merkezileştirilme (double centered) yapılarak maddeler arasındaki zaman yoğunluğuna ve bireysel hız farklarına bağlı değişimleri azaltmak amaçlanmıştır. Çift merkezileştirilmiş yanıt süreleri; $RT_{ip}^{DC} = RT_{ip} - \overline{RT}_{.p}^{WP} - \overline{RT}_{.i}^{WI} + \overline{RT}_{..}$ formülü ile hesaplanmaktadır (Jeon and De Boeck, 2019).

Çalışma kapsamında hiyerarşik olarak dört modelin karşılaştırılması yapılmıştır. İlk olarak, tek sınıflı iki parametrelili madde yanıt teorisi modeli (Model I) temel alınarak analiz yapılmıştır. İkinci model (Model II), yetenek sınıfı ile yetenek dışı (ikincil) sınıftan oluşmaktadır. Model II'deki yetenek dışı sınıfta başarı olasılığı yanıt seçeneklerinin sayısının tersi olan 0.25'e sabitlenen rastgele tahmin (random guessing) grubudur. Diğer model (Model III), her iki sınıfta da eşik değerlerinin serbestçe kestirildiği yetenek ve yetenek dışı sınıflardan oluşan iki sınıflı modeldir. Son olarak, Jeon ve De Boeck (2019) 'un önerdiği yaklaşım doğrultusunda dördüncü model (Model IV) ilki yetenek sınıfı; ikincisi ise yanıt süresinin kovaryant olarak eklendiği yetenek dışı sınıftan oluşan iki sınıflı modeldir.

Model değerlendirmeleri, Akaike Bilgi Kriteri (AIC), Bayesyen Bilgi Kriteri (BIC) (Schwarz, 1978), Düzeltilmiş Bayesyen Bilgi Kriteri (SABIC) ve Ölçeklendirilmiş Olabilirlik Oran Testi (SLRTS) (Lo, Mendell & Rubin, 2001) kullanılarak gerçekleştirilmiştir. Model tahminleri, Mplus yazılımı ile yapılmıştır.

Bulgular

Öncelikle sırasıyla Model I, Model II, Model III ve Model IV'e göre veri analiz edilerek incelenmiş ve elde edilen model uyum indeksleri Tablo 1'de sunulmuştur.

Tablo 1. Kestirilen Modeller için Uyum İndeksleri

Model	LL	N.Par	c	M-Comp	-2*LL	SLRTS	AIC	BIC	SABIC
Model I: Tek sınıflı	7120.087	22	1.010	-	-	-	14284.1	14395.7	14325.8
Model II: İki sınıflı (sabit)	7120.091	23	0.967	-	-	-	14286.1	14402.8	14329.7
Model III: İki sınıflı (serbest)	7093.054	34	1.010	M3-M2	54.074	49.038*	14254.1	14426.5	14318.5
Model IV: İki sınıflı (yanıt süresi kovaryantı ile)	6748.443	44	1.075	M4-M3	689.22	533.06*	13584.8	13806.9	13667.1

Not: $p < 0.001$ anlamlı. LL = Olabilirlik Oranı, AIC = Akaike Bilgi Kriteri, BIC = Bayesian Bilgi Kriteri, SABIC = Örneklem Boyutuna Göre Düzeltilmiş BIC, SLRTS = Ölçeklendirilmiş Olabilirlik Oranı Test İstatistiği, c = ölçeklendirme düzeltme faktörü

Model uyum istatistikleri değerlendirildiğinde, Model III (serbest parametrelili iki sınıf modeli) LL = 7093.054 değeri ile Model II'ye kıyasla daha iyi bir uyum sağladığı görülmektedir. Ayrıca, AIC = 14254.1 ve BIC = 14426.5 değerlerindeki azalma, bu modelin sabit iki sınıf modeline göre veriye daha uyumlu olduğunu göstermektedir. SLRTS testi sonuçları ise $\chi^2(11) = 49.04$, $p < .001$ ile iki model arasındaki farkın istatistiksel olarak anlamlı olduğunu ortaya koymaktadır.

Son aşamada, yanıt sürelerinin dâhil edildiği iki sınıflı model, LL = 6748.443 değeriyle en iyi uyumu sağlamaktadır. Bu modelin AIC = 13584.8, BIC = 13806.9 ve SABIC = 13667.1 değerlerindeki önemli düşüşler, uyumda kaydedilen belirgin bir iyileşmeyi göstermektedir. SLRTS testi sonuçları ise $\chi^2(10) = 533.06$, $p < .001$ ile Model III'e kıyasla anlamlı bir gelişme olduğunu ortaya koymaktadır.

Sonuç olarak, Model IV tüm uyum kriterlerinde en yüksek performansı sergilemekte ve yanıt sürelerinin analize dâhil edilmesi, modelin veriye olan uyumunu anlamlı bir şekilde iyileştirmektedir.

Yanıt Sürelerinin Kovaryant Olarak Dahil Edildiği İki Sınıflı Model

Yanıt sürelerinin kovaryant olarak dâhil edildiği iki sınıflı (yetenek-olağan R_g ve Yetenek dışı-ikincil S_g sınıf) Model IV'e ilişkin sonsal olasılıkların ortalamaları ve entropi değerleri Tablo 2'de sunulmuştur.

Tablo 2. Model IV için Sonsal Olasılıkların Ortalamaları ve Entropi Değeri

Örtük Sınıflar	Sonsal Olasılıkların Ortalaması (AvePP)		Entropi	En Muhtemel Sınıf Üyeliği n (%)
	Yetenek dışı sınıf	Yetenek sınıfı		
Yetenek dışı sınıf	0.943	0.057	0.814	551 (%48)
Yetenek sınıfı	0.049	0.951		597 (%52)

Tabloda sunulan sonsal olasılıkların ortalamaları ve entropi değeri incelendiğinde, AvePP'nin 0.90'ın üzerinde olduğu ve entropi değerinin 0.814'e ulaştığı gözlenmektedir. Bu durum, örtük sınıfların güçlü bir şekilde ayrıldığı ve modelin sınıflama belirsizliğinin düşük olduğunu göstermektedir. Muhtemel sınıf üyeliği dağılımına bakıldığında, Sınıf 1'de 551 kişi (%48) ve Sınıf 2'de 597 kişi (%52) yer almaktadır. Bu dengeli dağılım, her iki sınıfın belirgin şekilde tanımlandığını ve modelin sağlam bir şekilde işlediğini ortaya koymaktadır.

Matematik okuryazarlığı maddelerine ilişkin yetenek ve yetenek dışı sınıfların karşılaştırmalı parametre kestirimleri Tablo 3'te verilmiştir.

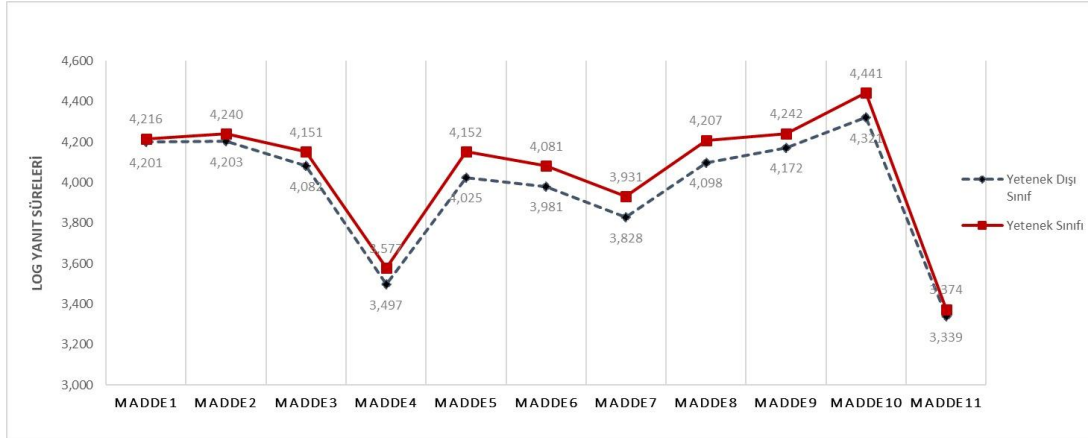
Tablo 3. Yetenek ve Yetenek Dışı Sınıflar için Madde Parametre Tahminleri

Matematik Okuryazarlığı	Yetenek sınıfı				Yetenek dışı sınıf			
	Eğim $\lambda_i^{R_s}$		Eşik değeri $\tau_i^{R_s}$		Eşik değeri $\delta_i^{S_s}$		YS Eğim $\gamma_i^{S_s}$	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE
Madde1	0.669**	0.243	-1.332**	0.118	-0.002	0.098	-0.011	0.249
Madde2	0.989**	0.286	-0.866**	0.122	1.205**	0.121	-1.404**	0.189
Madde3	0.738*	0.300	-1.984**	0.16	0.691**	0.119	-0.697**	0.206
Madde4	-0.054	0.261	-1.798**	0.131	-0.315**	0.099	1.734**	0.288
Madde5	0.229	0.188	1.262**	0.111	1.389**	0.121	-0.055	0.242
Madde6	0.395	0.227	-1.277**	0.108	0.686**	0.104	-0.481*	0.237
Madde7	0.464**	0.163	0.189	0.092	0.889**	0.106	-1.008**	0.239
Madde8	1.142*	0.498	-2.562**	0.287	1.036**	0.127	0.807**	0.218
Madde9	0.452	0.271	-1.717**	0.134	1.675**	0.165	-1.745**	0.256
Madde10	1.545**	0.483	1.297**	0.232	6.551**	2.232	-1.624**	0.23
Madde11	0.799	0.374	-2.528**	0.225	0.146	0.105	-0.398*	0.152

Not. YS = Yanıt Süresi. *p<.05, **p<.001.

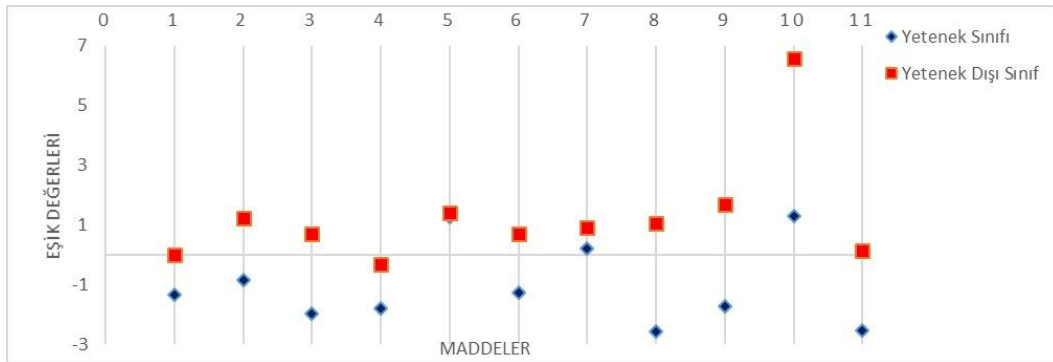
Yetenek sınıfı için kestirilen eşik parametreleri incelendiğinde, maddelerin büyük çoğunluğunda (madde 5, 7 ve 10 hariç) eşik değerlerinin negatif olduğu gözlenmiştir. Bu, yetenek sınıfındaki katılımcıların maddeleri genellikle kolay bulduğunu göstermektedir. Öte yandan, yetenek dışı sınıf için kestirilen eşik parametrelerinde maddelerin çoğunda (madde 1 ve 11 manidar değil, madde 4 ise negatif) pozitif değerler aldığı görülmektedir. Bu bulgu, yetenek dışı sınıftaki katılımcıların birçok maddeyi zor bulduğunu işaret etmektedir. Örneğin, madde 1'in eşik değeri (-0.002), bu maddenin yetenek dışı katılımcılar tarafından ortalama zorlukta algılandığını gösterirken, madde 2 (1.205), madde 9 (1.675) ve özellikle madde 10 (6.551) gibi yüksek pozitif eşik değerleri, bu maddelerin yetenek dışı katılımcılar için oldukça zorlayıcı olduğunu göstermektedir.

Yetenek dışı sınıftaki madde yanıt sürelerine ait eğim katsayıları değerlendirildiğinde, negatif eğim katsayıları dikkat çekmektedir (bkz. Tablo 3). Bu durum, katılımcıların ilgili maddelerde daha fazla zaman harcamalarının, yetenek dışı sınıfa ait olma olasılıklarını azalttığını göstermektedir. Örneğin, bir öğrenci madde 10'a daha fazla zaman harcadıkça, ikincil sınıfa veya yetenek dışı sınıfa ait olma olasılığı düşecektir. Detaylı incelenecek olursa, madde 10'un yüksek bir eşik değeri ve negatif bir YS eğim katsayısı ile dikkat çektiği görülmektedir. Bu durum, öğrencilerin bu maddeyi çok zor bulduklarını ancak hızlı bir şekilde yanıtlamanın (daha kısa sürede) yetenek dışı sınıfta yer alma olasılığını artırdığını göstermektedir. Aynı şekilde, yüksek bir eşik değeri ve negatif bir YS eğim katsayısına sahip (γ_i) madde 9 için de, yetenek dışı sınıftaki öğrencilerin bu maddeyi zor bulduğunu ve maddede harcanan süre azaldıkça yetenek dışı sınıfta yer alma olasılığının arttığı söylenebilir.

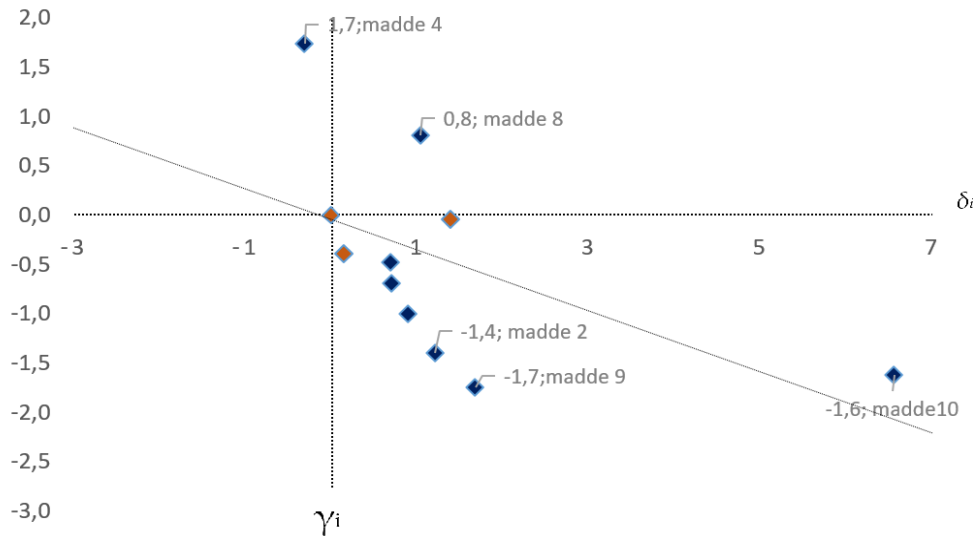


Şekil 2. Log-Yanıt Süreleri

Şekil 2’de, yetenek ve yetenek dışı örtük sınıflar için ortalama log madde yanıt süreleri verilmiştir. İki örtük sınıfın yanıt sürelerinin dağılımları incelendiğinde, maddelerin tamamında, yetenek sınıfının ortalama log madde yanıt sürelerinin daha yüksek olduğu görülmektedir. Yanıt süreleri arasında büyük farklılıklar olmamasına rağmen, yetenek dışı sınıf katılımcılarının tüm maddelerde daha düşük ortalama log yanıt süreleri kullandığı dikkat çekmektedir. Bu durum, her iki grup arasında benzer eğilimler olsa da, yetenek sınıfının maddelerde daha fazla zaman harcadığını göstermektedir. Öte yandan, yetenek dışı grup için daha kısa yanıt süreleri, bu grubun daha düşük bilişsel katılım sergilediğini ya da tahmine dayalı yanıt stratejileri kullandığını gösterebilir.



(a)



(b)

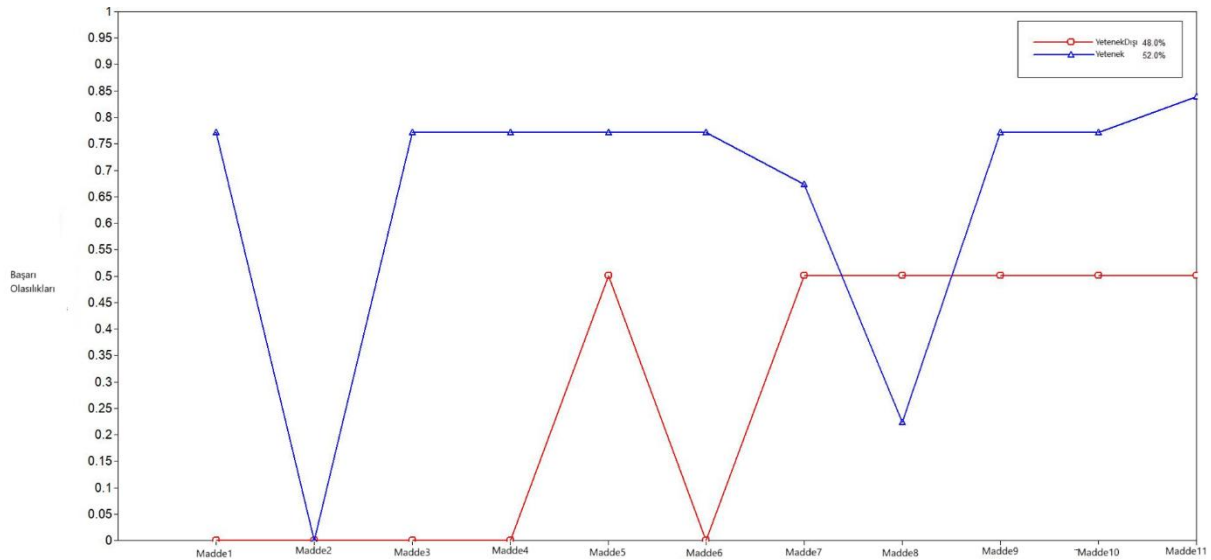
Şekil 3. (a) Yetenek ve Yetenek Dışı Gruplar için Eşik Parametrelerinin Dağılımı ve (b) Yetenek Dışı Gruptaki Maddeler için Eşik Parametreleri (δ_i) ve Tepki Süresi Eğimleri (γ_i)

Yetenek ve yetenek dışı gruplar için maddelerin eşik değerleri verilmiştir (bkz. Şekil 3 (a)). Yetenek sınıfına ait eşik değerleri büyük ölçüde negatif eksen tarafında bir dağılım gösterirken, yetenek dışı grubun eşik değerleri pozitif eksen tarafında farklılaşmakta ve bazı maddelerde (özellikle madde 10) çok daha yüksek değerler sergilemektedir. Yetenek dışı grup, belirli maddelerde (2, 3, 10) yetenek sınıfına kıyasla belirgin bir zorluk yaşarken; diğer maddelerde (5, 7) yetenekli sınıfa daha yakın bir zorluk yaşamıştır. Bu durum, her iki grup arasındaki güçlük farklarının madde bazında nasıl değiştiğini ortaya koymaktadır.

Yetenek dışı sınıf için madde yanıt sürelerine ait eğim katsayıları ile madde eşik parametreleri arasındaki ilişkiyi gösteren saçılma diyagramı incelendiğinde (Şekil 3(b)), maddelerin büyük çoğunluğundaki pozitif eksen yönündeki yüksek eşik değerlerine karşılık gelen negatif eğim katsayıları dikkat çekmektedir. Özellikle, madde bazında en yüksek eşik parametresine sahip olan madde 10 (en sağdaki nokta) için negatif madde yanıt süresi eğim katsayısı belirlenmiştir. Bu bağlamda, madde 10'un en yüksek eşik değerine rağmen negatif yanıt süresi eğim katsayısına sahip olması, öğrencilerin bu maddeyi çözme sürecinde daha az zaman harcadıklarında yetenek dışı sınıfta yer alma olasılıklarının arttığı şeklinde yorumlanabilir.

Yetenek dışı sınıfta madde yanıt süreleri eğim katsayıları ile eşik parametreleri arasındaki korelasyon katsayısı incelendiğinde, negatif yönde manidar korelasyon katsayısı dikkat çekmektedir (Spearman rho= -0.682, <.05). Bu sonuç, hızlı yanıt veren grubun bilgi geri çağırma stratejisini kullanan bir grup olmadığını açıkça ortaya koymaktadır. Jeon ve De Boeck (2019)'un çalışmasında belirtildiği gibi, " madde zorluğu ile yanıt süresinin bilgi geri çağırma sınıfına ait olma marjinal olasılığı üzerindeki etkisi arasında pozitif bir ilişki bekliyoruz. " (s. 698)

Bu çalışmada yetenek dışı sınıfta, genel/baskın olarak hızlı yanıt verme stratejisini benimseyen, negatif eğim katsayılarının tüm maddelerde gözlenmediği (madde 4 ve madde 8 pozitif eğim) için bu sınıf hızlı tahmin grubu olarak isimlendirilmemiştir.



Şekil 4. Yetenek Dışı ve Yetenek Sınıflarındaki Katılımcıların Maddeler için Başarı Olasılıkları

Yetenek ve yetenek dışı sınıfa ilişkin başarı olasılıkları incelendiğinde, yetenek sınıfının doğru yanıt olma olasılıkları genel olarak (madde 2 ve 8 hariç) daha yüksek olduğu görülmektedir. Maddelerin büyük çoğunluğunda yetenek sınıfının doğru yanıt olma olasılığı %65 ile %85 arasında değişmektedir. Özellikle 1., 3., 4., 5., 6., 7., 9., 10. ve 11. maddelerde başarı olasılığı daha belirgindir; bu da yetenek sınıfının belirli matematiksel okuryazarlık konularında daha yetkin olduğunu ve bu maddelere daha iyi yanıt verme eğiliminde bulunduğunu ortaya koymaktadır.

Yetenek sınıfının yanıt sürelerinin, maddelerin zorluk düzeyleriyle de ilişkili olduğu gözlenmiştir. Özellikle, bu grup en kolay maddelerde (madde 4 ve madde 11) daha az zaman harcayarak başarılı olmuştur. Aynı şekilde, en zor maddelere (madde 5 ve madde 10) daha fazla zaman ayırarak yine başarı elde etmişlerdir. Yüksek başarı olasılıklarına sahip bu öğrencilerin kolay maddelere hızlı, zor maddelere ise daha yavaş yanıt vermesi, yanıt süresi ile madde zorluğu arasındaki etkileşimli ilişkiyi göstermektedir (bkz. Şekil 3 (a) ve Şekil 4). Ayrıca, bilgiyi etkili bir şekilde kullanan ve sınav süresini verimli değerlendiren bir örtük sınıfa işaret edebilir. İlginç bir şekilde, öğrenciler madde 7'yi zor bulmalarına rağmen bu maddeye daha az zaman ayırmış ve bu durum başarı olasılığında düşüşe neden olmuştur.

Yetenek dışı sınıfın başarı olasılıkları genel olarak düşük seviyelerde kalmaktadır; birçok maddede doğru yanıt olma olasılığı %20'nin altında yer almaktadır. Özellikle 1., 2., 3., 4. ve 6. maddelerde bu sınıfın başarı olasılığı neredeyse sıfırdır. İlgi çekici olarak, yetenek dışı sınıf nispeten kolay (sınıfa koşullu olarak kolay) madde 4'ü yanıtlarken görece daha düşük madde yanıt süresi kullanmışlar ve bu maddedeki başarı olasılıkları düşmüştür. Benzer bir örüntü, 3. ve 6. maddede de gözlenmiştir. Diğer maddelere kıyasla daha kolay bulunan bu maddelerde en düşük madde yanıt süreleri harcanmış, ancak başarı olasılıkları oldukça düşük kalmıştır. Daha zor maddelerde (5., 9. ve 10.) ise madde yanıt sürelerinde gözlemlenen artışla birlikte başarı olasılıklarında da bir artış gözlenmiştir. Bu durum orta güçlükteki nispeten kolay maddelerde sürenin etkisiz kullanımına işaret ederken daha zor maddelerde süreyi daha etkili kullanarak kısmen başarılı (başarı olasılığı yaklaşık olarak .50) olan örtük bir sınıfa işaret etmektedir.

Madde güçlükleri ve yanıt sürelerine ilişkin bu ayrıntılı inceleme, negatif yanıt süresi eğimleri birleştirildiğinde, "hızlı yanıt veren ancak süreyi etkili kullanamayan" bir yanıt stratejisine sahip örtük sınıfın var olabileceğini düşündürmektedir. Öğrenciler, görece kolay maddelerde düşünmeden hızlı yanıt vermeye çalışabilirler. Bu, yüzeysel bir değerlendirme ile sonuçlanabilir ve bu da yanlış yanıtlara yol açabilir. Daha zor maddelerde ise, öğrencilerin geçirdikleri zamanın artmasıyla birlikte kısmen de olsa başarılarının artması, onların sorulara daha derinlemesine ve dikkatli bir şekilde yaklaşma çabalarını yansıtabilir.

Tartışma, Sonuç ve Öneriler

Bu çalışmada PISA 2022 Türkiye verisi kullanılarak, matematik okuryazarlığı testine katılan öğrencilerin yetenek ve yetenek dışı örtük sınıflarını belirlemeyi amaçlayan karma bir modelleme yaklaşımı uygulanmıştır. Motivasyonu düşük katılımcıların, özellikle düşük riskli sınavlarda tahmin stratejilerine başvurdukları, araştırmacılar tarafından sıklıkla vurgulanmaktadır (Wise ve DeMars, 2006; Wise ve Kong, 2005). Bu çalışmada, PISA gibi düşük riskli ve katılımcılar için doğrudan sonucu olmayan bir uygulamada, madde yanıt sürelerinin bilgilendirici doğası kullanılarak farklı madde yanıt stratejileri sergileyen örtük sınıfların varlığı incelenmiştir.

Çalışmada, yanıt süreleriyle ilişkili birey içi değişkenliğe de izin veren bir sonlu-karma madde yanıt modeli kullanılmıştır. Bu modelde, ikincil sınıfın hızlı ya da yavaş bir sınıf olması gerekmez ve doğası, önceden tanımlanmış değildir. Joen ve De Boeck (2019), Sideridis ve Alahmadi (2022) ve Sideridis ve diğerleri (2022) çalışmalarında da vurguladıkları gibi, potansiyel bir ikincil sınıfın doğası, bu sınıfın özellikleri, madde başarı oranları ve yanıt süresi ortak değişkeninin etkisi üzerinden veriden elde edilmiştir. Böylece, ikincil sınıfın doğası ampirik bir meseledir.

Bu yaklaşım, daha önceki karma modelleme çalışmalarında görülen sınıf tanımlamalarından farklıdır; örneğin, Meyer'in (2010) çalışmasında sınıfların doğası önceden belirlenmiştir. Burada bir sınıf hızlı tahmin sınıfı, diğer sınıf ise çözüm davranışı (yetenek temelli) sınıfı olarak tanımlanmıştır. Sınıflar, çözüm davranışı sınıfının yanıt süresi dağılımındaki ortalamanın hızlı tahmin sınıfından daha yüksek olmasıyla ayrılmıştır. Diğer taraftan, Wang ve Xu'un (2015) çalışmasında da benzer şekilde sınıfların doğasının önceden belirlendiği iki sınıflı bir karma model uygulanmıştır. Yazarlar, bir sınıfı maddeye özgü başarı olasılığına sahip sınıf olarak tanımlarken, diğer sınıfı tahmin grubu olarak tanımlamışlardır. Mevcut çalışmada ise, sınıf yapıları daha esnek bir şekilde veri odaklı olarak ortaya çıkmıştır ve ikincil sınıfın özellikleri veriyile şekillendirilmiştir.

Bu amaçla, Jeon ve De Boeck (2019) tarafından önerilen modeller test edilmiştir. Özel olarak, yanıt sürelerinin dâhil edildiği iki sınıflı model (Model IV) en iyi uyum gösterirken, yanıt sürelerinin kovaryant değişken olarak kullanılması örtük sınıflar arasındaki farklılıkları daha iyi anlamamıza olanak tanımıştır. Model IV'ten elde edilen örtük sınıflara ilişkin parametre kestirimleri ile bu sınıflardaki madde yanıt süreleri arasındaki ilişkiler, madde yanıt stratejilerine dair önemli ipuçları sunmuştur.

Bu çalışmada, yetenek sınıfı daha yüksek başarıya sahip ve daha uzun yanıt süreleri kullanan bir grup olarak bulunmuştur. Özellikle, yetenek sınıfının kolay maddelerde daha az zaman harcayarak başarı elde etmesi, bu grubun hızlı düşünme becerilerinin yanı sıra konuya hâkimiyetini de yansıtır. Zor maddelere daha fazla zaman ayırarak başarılı olmaları, yanıt süresi ile madde zorluğu arasında etkileşimli bir ilişkiyi göstermektedir. Bu bulgu, yüksek yeterlik düzeyine sahip öğrencilerin bilgi kullanımında etkin ve stratejik olduklarını gösterebilir. Ancak, bu yorumlar spekülâtif oldukları için dikkatle değerlendirilmelidir.

Bu çalışmada, ikincil yani yetenek dışı sınıf, yetenek sınıfına kıyasla daha düşük başarıya sahip ve daha kısa yanıt süreleri kullanan bir gruptur. Ancak, Sideridis ve Alahmadi'nin (2022) çalışmasında, yanıt süreleri tarafından tanımlanan gizil sınıfın, yetenek sınıfına göre daha fazla zaman harcadığı ve daha yüksek başarı düzeyine ulaştığı gözlenmiştir. Aynı çalışmada, özellikle matematik gibi analitik becerilerin başarılı olmak için kritik bir öneme sahip olduğu durumlarda, ek sürenin diğer konularda olduğu kadar faydalı olmayacağı vurgulanmıştır. Ayrıca, ilgili çalışmada sürenin öğrenci performansına etkisinin içerik türüne göre değişebileceği de bulunmuştur. Bu sonuçlar, yanıt sürelerinin, içerik zorluğunun öğrenci performansını değerlendirme sürecinde önemli değişkenler olduğunu göstermektedir. Mevcut çalışmada da özellikle birçok maddede doğru yanıtlama oranı %20'nin altında kalırken, en kolay maddelerde bile başarı oranlarının düşük olması, bu grubun konuya dair yeterli bilgiye (hatta temel bilgilere) sahip olmadığını gösterebilir. Zor maddelerde ise daha fazla zaman harcamalarına rağmen başarı oranlarının kısmen artması, bu grubun daha zor sorulara karşı daha fazla çaba gösterdiğini, ancak soruyu çözmek için yeterli bilgi ve yeterlikten yoksun olduklarını yansıtabilir.

Sideridis ve Alahmadi (2022) yaptıkları çalışmada, daha düşük performans gösteren grupların, zor maddelere daha fazla zaman harcadıkları ancak yine de yüksek yetenekli gruba kıyasla daha düşük başarı elde ettikleri sonucuna varmışlardır. İlgili çalışmada, düşük başarılı bireylerin bu maddeler için harcadıkları ek sürenin, mevcut beceri düzeylerinin ötesinde bir içeriğe işaret ettiği ve bu nedenle faydalı olmadığı öne sürülmüştür. Söz konusu çalışmada da madde çözüme stratejileri ile ilgili yapılan yorumların hipotetik/spekülâtif olduğu önemle vurgulanmıştır.

Sonuç olarak, bu çalışma, öğrencilerin bilgileri ve yanıt stratejileri arasındaki ilişkiyi anlamak için önemli ipuçları sunmakta ve eğitimde bu süreçlerin dikkate alınmasının gerekliliğini açıkça ortaya koymaktadır. Ayrıca Amerikan Eğitim Araştırmaları Derneği, Amerikan Psikoloji Derneği ve Eğitimde Ölçme Ulusal Konseyi tarafından geliştirilen Eğitimsel ve Psikolojik Test Standartları çaba ölçümlerinin toplanmasını ve raporlanmasını ve bu ölçümlerin test puanlarının yorumlanmasında kullanılmasını önermektedir (Uluslararası Test Komisyonu, 2013).

Büyük ölçekli dijital sınavlar ve e-sınavlar, geniş bir öğrenci kitlesine ulaşarak veri toplama açısından önemli fırsatlar sunmaktadır. Bu uygulamalar, öğrenci performansını değerlendirmek için çok sayıda veri kaynağı sağlarken bu verinin analizi, eğitim sistemlerinin daha iyi anlaşılmasına katkı sağlayabilir. Özellikle, PISA ve TIMSS gibi uluslararası uygulamalara benzer şekilde, ulusal düzeydeki e-sınav uygulamalarında da madde yanıt süresi verisinin toplanması, öğrenci performansını değerlendirmek ve yanıt stratejilerini anlamak açısından kritik bir öneme sahiptir. Bu veri, eğitimcilerin öğrenci başarılarını daha iyi anlamalarına olanak tanıyabilir (Jeon ve De Boeck, 2019).

Öđrenci motivasyonunun ölçülmesi, bir testin bilgiyi mi yoksa motivasyonu mu ölçtüđünü belirlemede kritik bir öneme sahiptir (Eklöf, 2010). Test performansının yalnızca bilgi düzeyini yansıtmakla kalmayıp, aynı zamanda öğrencilerin test sırasında kullandığı stratejiler ve motivasyonel durumları da gösterebileceđi unutulmamalıdır. Dijital dönüşüm ve teknoloji, büyük ölçekli değerlendirmeleri geliştirmek ve dönüştürmek adına birçok fırsat sunmaktadır. Yanıt sürelerinin sistematik bir şekilde izlenmesi ve analiz edilmesi, bu fırsatların başında gelir. Bu bağlamda, yanıt süreleri, büyük ölçekli testlerin kalitesini ve güvenilirliğini artırmak için güçlü bir araç olabilir (Maddox, 2023).

Çalışma kapsamında, yetenek dışı sınıfın karakterize edildiđi model kestirimleri en fazla iki sınıfla sınırlandırılmıştır. Bunun temel sebebi, yorumlama ve parametre kestirimlerinde karşılaşılabilecek zorluklardır. İlgili model karşılaştırmalarının iki sınıfla sınırlandırıldığı Sideridis ve Alahmadi (2022) çalışmasında da benzer zorluklardan söz edilerek, sınıfların karakterize edilmesi ve yorumlama kolaylığı sebebiyle bu yaklaşım tercih edilmiştir. Sınıfların örtük doğası düşünöldüğünde, bir yetenek sınıfının var olduđu ancak iki, üç ya da dört yetenek dışı sınıfının da olabileceđi dikkate alınarak üç ve daha fazla sınıflı model karşılaştırmalarına yer verilebilir. Bu açıdan, yapılacak çalışmalar için birden fazla yetenek dışı sınıfın var olması durumunda bu sınıfların yanıtlama stratejilerinin nasıl deđişeceđinin belirlenmesi önemli görölmektedir.

Yanıt sürelerinin analiz edilmesi, öğrencilerin stratejilerini daha iyi anlamamıza olanak tanımaktadır. Örneđin, öğrencilerin yanıt verirken daha fazla zaman harcaması, bilgi geri çağırma stratejilerine işaret edebilir (Jeon ve De Boeck, 2019). Öte yandan, daha kısa yanıt süreleri, gözden geçirme, kaçınma motivasyonu veya öğrenilmiş çaresizlik gibi psikolojik durumları gösterebilir (Abramson, Metalsky ve Alloy, 1989; Seligman, 1972). Daha sonra yapılacak çalışmalarda bu çalışmada kullanılan model geliştirilebilir. Örneđin, yanıt stratejisi sınıflarının doğasını araştırmak için diđer davranışsal, demografik ya da psikolojik bilgileri modele dahil etmek faydalı olabilir. Çalışmada kullanılan veri geniş ölçekli ancak öğrenciler açısından düşük riskli bir veridir. Öğrenciler için seçme, yerleştirme ya da puan/not gibi bir motivasyonun olduđu bir sınavda analizin tekrar edilmesiyle elde edilen strateji sınıfları tekrar incelenebilir.

Kaynakça

- Abramson, L. Y., Metalsky, G. I. ve Alloy, L. B. (1989). Hopelessness depression: a theory-based subtype of depression. *Psychological Review* 96(2), 358-372. doi:10.1037/0033-295X.96.2.358
- AERA, APA ve NCME. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Anghel, E., Khorramdel, L. ve von Davier, M. (2024). The use of process data in large-scale assessments: A literature review. *Large-scale Assessments in Education*, 12(1), 13. doi:10.1186/s40536-024-00202-1.
- Baumert, J. ve Demmrich, A. (2001). Test motivation in the assessment of student skills: the effects of incentives on motivation and performance. *European Journal of Psychology of Education*, 16(3), 441-462.
- Bloom, B., Engelhart, M., Furst, E., Hill, W. ve Krathwohl, D. (1956). *Taxonomy of educational objectives: The classification of educational goals handbook I: Cognitive domain*. New York: David McKay Company.
- Bolt, D. M., Cohen, A. S. ve Wollack, J. A. (2002). Item parameter estimation under conditions of test speededness: Application of a mixture Rasch model with ordinal constraints. *Journal of Educational Measurement*, 39(4), 331-348. doi:10.1111/j.1745-3984.2002.tb01146.x
- Brophy, J. ve Ames, C. (2005). *NAEP testing for twelfth graders: Motivational issues*. Washington, DC: National Assessment Governing Board.
- Brückner, S. ve Pellegrino, J. W. (2017). Contributions of response processes analysis to the validation of an assessment of higher education students' competence in business and economics. B. D. Zumbo ve A. M. Hubley (Ed.), *Understanding and investigating response processes in validation research* içinde (s. 31-35). New York: Springer International Publishing.
- Cao, Y. ve Stokes, L. (2008). Modeling response times in test-taking: Applications and developments. *Journal of Educational Measurement*, 45(2), 135-153.
- Chang, Y. C., Tsai, C. C. ve Hsu, H. C. (2014). The impact of guessing strategies on item response theory model parameters. *Educational and Psychological Measurement*, 74(1), 69-85.
- De Boeck P. ve Jeon M. (2019) An overview of models for response times and processes in cognitive tests. *Frontiers in Psychology*, 10, 102. doi:10.3389/fpsyg.2019.00102
- De Jong, T. ve Ferguson-Hessler, M. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31(2), 105-113.
- Eklöf, H. (2010). Skill and will: Test-taking motivation and assessment quality. *Assessment in Education Principles Policy Practice*, 17(4), 345-356.
- Entwistle, N. ve Peterson, E. (2004). Conceptions of learning and knowledge in higher education: Relationships with study behavior and inferences of learning environments. *International Journal of Educational Research*, 41, 407-428. doi:10.1080/0969594X.2010.516569
- Erwin, T. D. ve Steven L. W. (2002). A scholar-practitioner model for assessment. *Building a scholarship of assessment* içinde (s. 67-81). San Francisco: Jossey-Bass.
- Finn, B. (2015). *Measuring motivation in low-stakes assessments (research report RR-15-19)*. Princeton, NJ: Educational Testing Service.
- Haladyna, T. M. ve Downing, S. M. (2004). Construct-irrelevant variance in high-stakes testing. *Educational Measurement: Issues and Practice*, 23(1), 17-27. doi:10.1111/j.1745-3992.2004.tb00149.x
- ITC. (2013). *International guidelines for test use*. <http://www.intestcom.org/Guidelines/Test+Use.php> adresinden erişildi.
- Jeon, M. ve De Boeck, P. (2019). An analysis of an item-response strategy based on knowledge retrieval. *Behavior Research Methods*, 51, 697-719. doi:10.3758/s13428-018-1064-1
- Lo, Y., Mendell, N. R. ve Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika*, 88(3), 767-778. doi:10.1093/biomet/88.3.767

- Maddox, B. (2023). The uses of process data in large-scale educational assessments. *OECD Education Working Papers*, 286, Paris: OECD Publishing. [doi:10.1787/5d9009ff-en](https://doi.org/10.1787/5d9009ff-en).
- Meyer, J. (2010). A mixture Rasch model with item response time components. *Applied Psychological Measurement*, 34(7), 521-538.
- Millî Eğitim Bakanlığı. (2022). *PISA 2022 uluslararası öğrenci değerlendirme programı*. https://pisa.meb.gov.tr/meb_iys_dosyalar/2022_01/26105818_PISA_2022_TanYtYm_KitapcYYY.pdf adresinden erişildi.
- Mislevy, R. J. ve Verhelst, N. (1990). Modeling item responses when different subjects employ different solution strategies. *Psychometrika*, 55(2), 195-215.
- Oranje, A., Gorin, J., Jia, Y., Kerr, D., Ercikan, K. ve Pellegrino, J. W. (2017). Collecting, analysing, and interpreting response time, eye tracking and log data. K. Erickan ve J. W. Pellegrino (Ed.), *Validation of score meaning for the next generation of assessments* içinde (s. 39-51). Mount Royal, NJ: National Council on Measurement in Education.
- Pohl, S., Ullrich, E. ve von Davier, M. (2021). Reframing rankings in educational assessments. *Science*, 372(6540), 338-340.
- Pokropek, A. (2016). Grade of membership response time model for detecting guessing behaviors. *Journal of Educational and Behavioral Statistics*, 41(3), 300-325. [doi:10.3102/1076998616636618](https://doi.org/10.3102/1076998616636618).
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461-464.
- Seligman, M. E. (1972). Learned helplessness. *Annual Review of Medicine*, 23, 407-412. [doi:10.1146/annurev.me.23.020172.002203](https://doi.org/10.1146/annurev.me.23.020172.002203)
- Shiffrin, R. M. ve Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84(2), 127-190.
- Sideridis, G. ve Alahmadi, M. T. S. (2022). The role of response times on the measurement of mental ability. *Frontiers in Psychology*, 13. [doi:10.3389/fpsyg.2022.892317](https://doi.org/10.3389/fpsyg.2022.892317)
- Sideridis, G., Tsaousis, I. ve Al-Harbi, K. (2022). Identifying ability and nonability groups: incorporating response times using mixture modeling. *Educational and Psychological Measurement*, 82(6), 1087-1106.
- Wang, C. G. ve Xu, G. (2015). A mixture hierarchical model for response times and response accuracy. *British Journal of Mathematical and Statistical Psychology*, 68(3), 456-477.
- Wise, S. L. (2019). An information-based approach to identifying rapid-guessing thresholds. *Applied Measurement in Education*, 32(4), 325-336, [doi:10.1080/08957347.2019.1660350](https://doi.org/10.1080/08957347.2019.1660350)
- Wise, S. L. ve Demars, C. E. (2006). An application of Item response time: The effort moderated IRT model. *Journal of Educational Measurement*, 43(1), 19-38.
- Wise, S. L. ve Kong, X. (2005). Response time effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education*, 18, 163-183. [doi:10.1207/s15324818ame1802_2](https://doi.org/10.1207/s15324818ame1802_2)
- Wise, S. L., Pastor, D. A. ve Kong, X. J. (2009). Correlates of rapid-guessing behavior in low-stakes testing: Implications for test development and measurement practice. *Applied Measurement in Education*, 22(2), 185-205.
- Yamamoto, K. H. (1997). Modeling the effects of test length and test time on parameter estimation using the HYBRID model. J. Rost ve R. Langeheine (Ed.), *Applications of latent trait and latent class models in the social sciences* içinde (s. 89-98). New York: Waxmann Verlag GmbH.

Ek 1. Matematik Okuryazarlığı Maddelerindeki Yanıt Sürelerinin Frekans Dağılımı

