

Citation / Atıf: Büyükkatak, E. (2025). Bireye uyarlanmış testler üzerine kavramsal bir inceleme: Gelişimi, yöntemi, avantajları ve dezavantajları. *Journal of Management Theory and Practices Research*, 6(1), 1–12. <https://doi.org/10.71284/jmtpr.1626781>

RESEARCH ARTICLE / ARAŞTIRMA MAKALESİ

BİREYE UYARLANMIŞ TESTLER ÜZERİNE KAVRAMSAL BİR İNCELEME: GELİŞİMİ, YÖNTEMİ, AVANTAJLARI VE DEZAVANTAJLARI

*A CONCEPTUAL REVIEW OF COMPUTER ADAPTED TESTS (CAT):
DEVELOPMENT, METHODS, ADVANTAGES AND DISADVANTAGES*

Emrah Büyükkatak  ¹

Özet

Teknolojik gelişmeler son çeyrek yüzyılda çok hızlı bir şekilde ilerlemiş, bilgi ve iletişim teknolojilerinin gelişmesiyle bilgisayarlar amaçlarımızın çoğuna ulaşmada en yararlı kolaylaştırıcı haline gelmiştir. Bu gelişmeler her sektörü etkilediği gibi eğitim ortamında, özellikle de öğrenme, test etme ve değerlendirme gibi alanlarda değişikliklere yol açmıştır. Bu doğrultuda testler ve testlerin uygulanma yöntemleri de gelişerek çeşitlenmiştir. Bilgisayarların artan uygunluğu ve hesaplama yeteneği sebebiyle, bilgisayarlı değerlendirme testlerde daha belirgin bir rol kazanmıştır. Bilgisayar teknolojisinin eğitimde kullanımının artmasıyla birlikte yakın dönemde seçme, yerleştirme ve yetenek düzeyinin belirlenmesi gibi çeşitli maksatlarla bilgisayar ortamında bireye uyarlanan testler kullanılmaktadır. Bu çalışma bilgisayar ortamında bireye uyarlanan testler üzerine kavramsal bir araştırma yaparak, bu testlerin gelişiminin, yönteminin, avantajlarının ve dezavantajlarının incelenmesini amaçlamaktadır. Bu amaç kapsamında, bilgisayar ortamında bireye uyarlanan testlerin özelliklerine ve uygulanmasına yönelik kavramsal ve kuramsal yapısı incelenmiştir. Bilgisayar ortamında bireye uyarlanan testlerin kullanımı farklı etkilere sahiptir ve birçok avantajları ile birlikte dezavantajları da bulunmaktadır. Bir testin uyarlanabilir formatta uygulanması istenildiğinde bilgisayar ortamında bireye uyarlanan testlerin olumlu yönleri kadar olumsuz yönlerinin de dikkate alınması gerekmektedir. Eğitim sistemindeki sınavlarda bu tür testlerin de kullanılması, mili test sistemlerinin bu standartlara göre yapılandırılması, uygulama becerisi, teknik donanım, pedagojik anlayış ve politik desteğin bir arada olması önerilmektedir. Bu çalışma ortaya koyduğu sonuçlarla akademik literatüre birçok katkı sağlarken, politika yapıcılara ve uygulayıcılara da rehberlik edebilecektir.

Anahtar Kelimeler: Test, Uyarlanmış Test, Bilgisayar Ortamında Bireye Uyarlanmış Test

JEL kodları: I20

¹ Dr., e-mail: ebuyukatak.academic@gmail.com



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Telif Hakkı & Lisans | Copyright & License

The copyrights of the studies published in our journal belong to our journal and are published as open access under CC-BY-NC-ND license.

Dergimiz yayımlanan çalışmaların telif hakları dergimiz ait olup, CC-BY-NC-ND lisansı altında açık erişim olarak yayımlanmaktadır.

Abstract

Technological developments have progressed rapidly in the last quarter century, and with the development of information and communication technologies, computers have become the most useful facilitators in achieving most of our goals. These developments have affected every sector, as well as causing changes in the educational environment, especially in areas such as learning, testing and evaluation. Accordingly, tests and test application methods have also developed and diversified. Due to the increasing convenience and computational ability of computers, computerized assessment has gained a more prominent role in tests. With the increasing use of computer technology in education, computer adapted tests (CAT) have been used for various purposes such as selection, placement and determination of talent level in the recent period. This study aims to conduct a conceptual research on computer-adapted tests and to examine the development, method, advantages and disadvantages of these tests. Within the scope of this purpose, the conceptual and theoretical structure of the characteristics and application of tests adapted to individuals in a computer environment were examined. The use of computer-adapted tests has different effects and has many advantages as well as disadvantages. When a test is to be applied in an adaptive format, the negative aspects of computer-adapted tests should be taken into consideration as well as their positive aspects. It is recommended that such tests be incorporated into examinations within the education system, and that national testing systems be structured according to these standards, with implementation skills, technical infrastructure, pedagogical understanding, and policy support all working together. While this study contributes to the academic literature with its results, it may also guide policy makers and practitioners.

Keywords: Test, Adaptive Test, Computerized Adaptive Test

JEL codes: I20

1. GİRİŞ

Son yıllarda, bilgi ve iletişim teknolojisi alanında birçok ilerleme olmuştur. Bilgisayarların ve elektronik cihazların kullanımının tüm dünyada yaygınlaşmasıyla, bilgi ve iletişim teknolojilerinin gelişmesi her sektörü etkilediği gibi eğitim ortamının farklı alanlarını da etkilemiştir. Bilgisayar teknolojilerindeki bu gelişmeler öğrenme, test etme ve değerlendirme gibi eğitim ortamlarının birçok bölümünde değişikliklere ve gelişmelere yol açmıştır (Bennett, 2002; Pommerich, 2004). Özellikle dünya test endüstrisi de son çeyrek yüzyılda büyük değişiklikler göstermektedir. 60 yıldan daha uzun süredir yaygın olarak kullanılan klasik test teorisi yöntemleri, yerini yeni test yöntemlerine (Bilgisayar Destekli Uyarlanabilir Test, Genellenebilirlik Kuramı, Yapı Geçerliği Modelleri, Bayesyen Test Teorileri, Doğrudan Ölçüm Modelleri), en önemlisi de Madde Tepki Kuramına (MTK) bırakmaktadır.

Bilgisayarlar ve çevrimiçi ölçmeler, bireylerin her alandaki yeterliliğini değerlendirmek için giderek daha fazla kullanılmaktadır (Fleming ve Hipple, 2004). Ölçümdeki uygulamalar ve tercihler de bu değişikliklerden en çok etkilenen ve bunlara hızla uyum sağlayanlardandır (Bennett, 2002). Bilgisayar teknolojisinin eğitimde kullanılmasına paralel olarak testler ve testlerin uygulanma yöntemleri de gelişerek çeşitlenmiştir. Modern bilgisayarların artan uygunluğu ve hesaplama yeteneği nedeniyle, bilgisayarlı değerlendirme testlerde daha belirgin bir rol kazanmıştır. Bilgisayar ortamında bireye uyarlanan testler (BOBUT), son dönemde psikometri ve eğitim alanlarında, önemli ve yaygın olarak kullanılan bir araştırma ve uygulama konusu olmuştur (Magis, Yan ve von Davier, 2017).

1970'lerin başlarında ortaya çıkan bilgisayarlı testler (Drasgow, 2002) ile birlikte lisanslama, sertifikalandırma, kabul ve psikolojik testler gibi büyük ölçekli test programlarında bilgisayarlı testler uygulanmaya başlamıştır (Kim ve Huynh, 2007). Bu gelişmelerin sonucunda TOEFL ve IELTS gibi dil yeterlilik ve SAT, GRE ve GMAT gibi akademik yeterlilik standart testleri bilgisayarlı formatlarda uygulanmaktadır. Böylelikle bilgisayar teknolojisinin eğitimde kullanımının artmasıyla birlikte son 30 yıldır, seçme, yerleştirme, yetenek düzeyinin belirlenmesi gibi çeşitli maksatlarla BOBUT (Computerized Adaptive Test-CAT) kullanılmaktadır.

Uyarlanabilir test, farklı test sorularının her bir bireyin ölçülen özellikteki durumuna bağlı olarak farklı bireylere uygulandığı bir testtir. Geleneksel testlerden farklı olarak uyarlanabilir testlerde, tüm adaylara aynı sabit öge kümesi (kâğıt-kalem testinde olduğu gibi) uygulanmaz. Bu sebeple uyarlanabilir test yanıtı bağlı, programlanmış, bireyselleştirilmiş ve ardışık test olarak da adlandırılır. BOBUT bilgisayarlar tarafından etkili ve verimli bir şekilde yönetilmek üzere psikolojik ve eğitimsel testlerin yeniden tasarlanması ile oluşmuştur. Bu testlerde amaç, her sınava giren kişi için, aynı anda o kişiyi özellik açısından en etkili ve verimli şekilde ölçen test soruları kümesini seçmektir.

Bilgisayar ortamında bireye uyarlanan testler, psikometrik özellikleri belirlenmiş bir maddehavuzundan, her birey için tüm maddelerin aynı olmadığı ve yanıtlayıcıların beceri düzeylerine uygun maddelerin seçilmesiyle oluşturulan testler olarak tanımlanmaktadır (Weiss, 2004). Özellikle, her sınava girenin yeteneğine göre test sorularının ayarlayan bilgisayar uyarlamalı test, hem seçim hem de değerlendirme amaçları için benimsenmektedir. Bireye uyarlanan testlerin popüleritesindeki artış büyük olasılıkla artan test güvenliği, anında puan raporlaması, geliştirilmiş verimlilik ve üstün ölçüm hassasiyeti gibi avantajlara atfedilebilir (Zickar vd., 1999). Bilgisayar ortamında bireye uyarlanan testlere olan bu yeni güven, uyarlamalı testlerin özellikle bir seçim aracı olarak kullanılmasının etkinliğini ortaya koymaktadır (Tonidandel vd., 2002). Bu testler, bir öğrencinin önceki öğelerdeki performansına göre test öğelerinin zorluğunu dinamik olarak seçmede algoritma kullanarak testleri öğrencilere göre uyarlama fırsatları sunar ve gerçek bilgi düzeyine yakınsar (Collares ve Cecilio-Fernandes, 2019).

Bilgisayar ortamında bireye uyarlanan testler kâğıt-kalem testlerinden birçok yönden farklıdır. Bu farklılıklar, bir sınava giren kişinin test içinde uygulama becerisiyle ve kişinin puanını hesaplamak için kullanılan prosedürle ilgilidir. Geleneksel ve uyarlanabilir testler arasındaki diğer farklılıklar, uyarlanabilir bir teste nasıl başlanacağı ve daha sonra hangi öğenin uygulanacağı kararlarından kaynaklanmaktadır. Bazen uyarlanabilir test tasarımcıları bir teste kolay veya zor bir öğeyle başlamayı seçebilir. Benzer şekilde, test geliştiricisi sonraki öğelerin zorluğunu değiştirebilir. Bilgisayar ortamında bireye uyarlanan testlerin, kâğıt tabanlı testlerle karşılaştırıldığında test güvenilirliğini arttırdığına, öğrenci motivasyonunu ve katılımını iyileştirdiğine dair kanıtlar bulunmaktadır (Martin ve Lazendic, 2018). Ayrıca BOBUT ile kâğıt-kalem testinde kestirilen yeterlik (başarı) düzeyleri göz önüne alındığında BOBUT yeterlik kestirimlerinin güvenilirliğini (standart hata değerinin düşük olması) artırdığı ve kullanılan soru sayısında önemli derecede tasarruf sağlandığı görülmektedir (Özbaşı ve Demirtaşlı, 2015). Bu bağlamda, bilgisayar ortamında bireye uyarlanan testlerin hâlihazırda kullanılan kâğıt kalem testlerine göre önemli üstünlükleri bulunmaktadır (Montgomery & Cutler, 2013; Wainer vd., 2000; Weiss, 2004). Bu avantajların yanı sıra bireye uyarlanmış testlerin dezavantajları ve çözülmesi beklenen sorunları da bulunmaktadır.

Bu araştırmanın temel amacı bilgisayar ortamında bireye uyarlanan testler üzerine kavramsal bir araştırma yaparak, bu testlerin gelişiminin, yönteminin, avantajlarının ve dezavantajlarının incelenmesidir. Bu çalışma uygulayıcılara ve akademik literatüre birçok katkı sağlaması beklenmektedir. Öncelikle literatürde bireye uyarlanan testler ile ilgili olarak çoğunlukla simülatif çalışmalar ve denemeler, test stratejilerinin karşılaştırılması, değişik alanlarda uygulanabilirliğinin incelenmesi gibi özellikle testlerin uygulanmasına yönelik araştırmalar (Fergadiotis vd., 2021; Özbaşı ve Demirtaşlı, 2015; Gür ve Karabay, 2018; Gao vd., 2023; Kezer ve Koç, 2014; Dumas vd., 2012; Rice vd., 2022) olmakla birlikte eğitim alanında kavramsal incelenmesi üzerine yapılan çalışmalar sınırlıdır ve genellikle kıyaslama ve üstünlüklerine yönelik görüşlerden oluşmaktadır (Rezaie ve Golshan, 2015; Kalender, 2004; Alam, 2022; Doğan, 2022). Bu açıdan değerlendirildiğinde, literatürde kavramsal ve kuramsal sınırlı sayıda araştırmalar olmasından dolayı bu araştırma, kavramların daha iyi anlaşılmasına yönelik alan yazına katkı sağlayacak ve bir yenilik getirecektir. Ayrıca avantajları ve dezavantajları ortaya konarak, hâlihazırda birçok ülkede ulusal düzeyde yapılan sınavlarda yaygın olarak kullanılmayan ve klasik yöntem olan kâğıt-kalem testlerine göre daha az maliyet ve zaman harcanarak uygulanan bilgisayar ortamında bireye uyarlanan testlere yönelik araştırma yapılmasına imkân sağlayacaktır. Böylelikle bu testlerin ülkelerin gündemine

ve eğitim politikalarına girmesine yönelik uygulayıcılara ve eğitim sistemlerine birçok katkı sağlayacaktır.

2. BİLGİSAYAR ORTAMINDA BİREYE UYARLANAN TESTLER (BOBUT)

2.1. Bilgisayar Ortamında Bireye Uyarlanan Testlerin (BOBUT) Tarihsel Gelişimi

Bireye uyarlanmış test düşüncesi 20.yy başlarında ortaya çıkmıştır. Alfred Binet ve Theodore Simon tarafından 1905 yılında eğitim alamayan, zekâ seviyesi düşük bireyleri tespit etmek için tasarlanan Stanford-Binet Zekâ Testi olarak kullanıma sunulan zekâ testi, bireye uyarlanmış testler için ilk başlangıç olarak kabul edilmektedir (Weiss, 1983). Bu test, sabit dallandırma kuralı, değişken başlangıç noktası ve değişken sonlandırma kuralı ile çalışan mekanik bir dallandırma stratejisine sahiptir (Kezer, 2013; Özgüven, 2007; Turgut ve Baykul, 2010). Binet testi, farklı sayıda maddelerin farklı bireylere uygulanması için değişken sonlandırma kuralına sahiptir (Weiss, 1983). Zekâ katsayısı (Intelligence Quotient-IQ) ile akranların seviyelerinin kıyaslandığı bu test, problem çözme, dikkat ve hatırlama gibi birçok alanı ölçmek için dizayn edilmiştir. Testten ortaya çıkan sonuç bireyin zekâ katsayısıdır ve eğitime alınıp alınamayacağına bu zekâ katsayısına göre karar verilmiştir. Genellikle çocukların zekâ ölçümünde kullanılmış olsa da Binet tarafından yapılan iki güncelleme ile yetişkin grup için de IQ testi olarak geliştirilmiş ve sonuç olarak 11 ayrı yaş grubuna ölçüm yapabilen bir zekâ testi meydana getirilmiştir (Kuzucu, 2024). Lewis M. Terman tarafından Binet'in ölümünden sonra test güncellenmiş ve günümüze kadar kullanılması sağlanmıştır (Thorndike, 1991).

Binet'nin ardından 1950'lerin sonuna kadar bireye uyarlanmış testlerle ilgili önemli bir gelişme olmamıştır. Ancak deneme amaçlı ve yalnızca bireye uyarlanmış test yaklaşımlarının özelliklerinin belirlenmesi amacıyla çalışmalarda kullanılmak üzere testler yapılmıştır. Bununla birlikte 1960'lı yıllarda bireye uyarlanmış testlere Lord çalışmalarıyla katkılar sağlamıştır. 1971 yılında bilgisayar ortamında uygulanmayan ve kendi kendine puanlanabilen esnek düzeyli test (selfscoring flexilevel test) olarak adlandırılan bireye uyarlanmış test Lord tarafından oluşturulmuştur (Lord, 1971).

Bilinen ilk bilgisayar ortamında bireye uyarlanan testler ABD'deki Brigham Young Üniversitesi'nde Larson ve Madsen (1985) tarafından uygulanmıştır. Ardından 1990'lar boyunca daha fazla uyarlanmış testler oluşturma ve geliştirme konusunda çalışmalar yapılmıştır. Bu süreçte bazı standartlaştırılmış testler bilgisayara uyarlanabilir test formatında uygulanmıştır. 1998 yılında Yabancı Dil Olarak İngilizce Testi (TOEFL) bilgisayara uyarlanabilir test formatını kullanmıştır (Mojarrad vd., 2013). Bilgisayar ortamında bireye uyarlanan test uygulamaları, ABD Silahlı Kuvvetler Mesleki Yetenek Testi, Yabancı Dil Olarak İngilizce Testi (TOEFL), Lisansüstü Kayıt Sınavı (GRE) ve Uluslararası Öğrenci Değerlendirme Programı (PISA) (Schleicher, 2019) gibi yüksek riskli kabul testlerinde de daha yaygın hale gelmektedir. Bu gelişmelere paralel olarak, BOBUT'u deneyimleyerek sınava girenlerin sayısı tüm dünyada artmaktadır.

2.2. Bilgisayar Ortamında Bireye Uyarlanan Testlerin (BOBUT) Çalışma Prensibi

Bilgisayar ortamında bireye uyarlanan testler, katılımcıların yetenek düzeylerine göre uyarlanan bir tür bilgisayar tabanlı testtir. Bilgisayar ortamında uyarlanabilir testin temel prensibi basittir: Teste tabi tutulan öğrenci için çok zor veya çok kolay sorular sormaktan kaçınır. Yetenekli öğrencilerin kolay sorulara doğru cevap vereceğinden ve zorluk yaşayan öğrencilerin de zor sorularla karşılaşacaklarından ve bu tür cevaplardan nispeten az şey öğrenileceğinden öğrenciyi zorlayan ancak bunaltmayan soruların uygulanmasıyla çok daha fazlası öğrenileceği değerlendirilmektedir (Davey, 2011). Bu soruların doğru şekilde belirlenmesi ve sunulması her bilgisayar ortamında uyarlanabilir testin amacıdır.

Bilgisayara uyarlanabilir testlerde genellikle bir madde havuzu kullanılır. Bu madde havuzu, farklı zorluk seviyelerine sahip çok sayıda test maddesini içerir ve test maddeleri buradan seçilmektedir (Larson, 1999). BOBUT'un başarısı, temel test taslağının tüm alanlarını orantılı olarak kapsayan, değişken zorluk

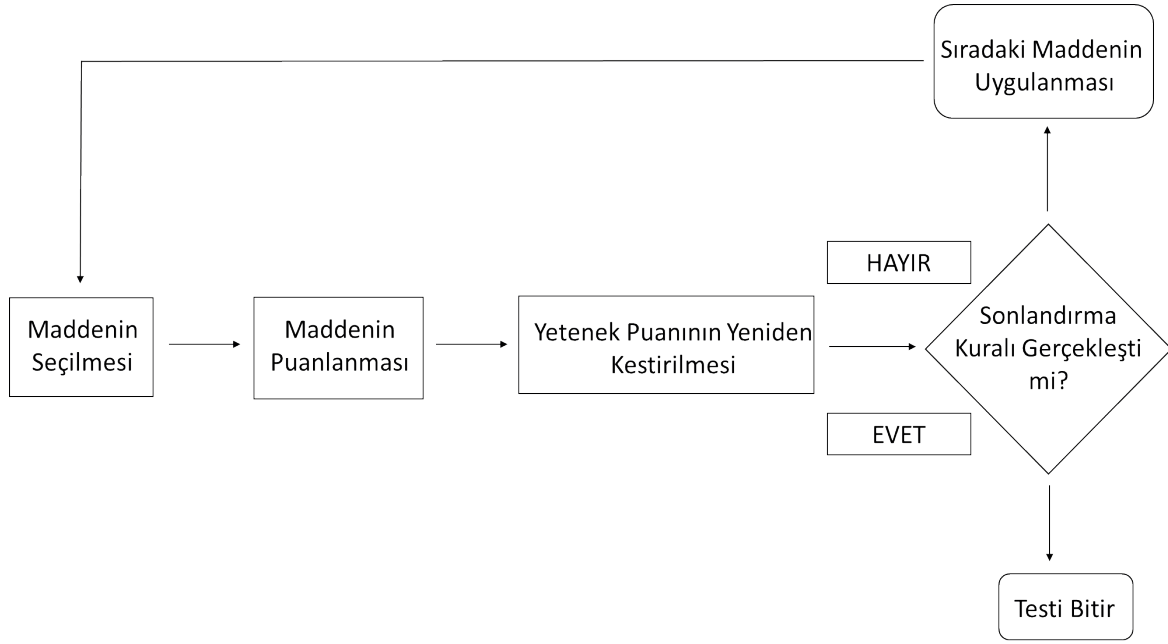
derecesine sahip ayarlanmış büyük bir madde havuzuna bağlıdır ve yüksek riskli değerlendirmeler için en az 1.000 madde önerilir (Wise ve Kingsbury, 2000). Bu amaçla, madde havuzu geliştirmede uluslararası bir iş birliği çerçevesi son derece yararlı hale gelmektedir. Ayrıca iyi geliştirilmiş BOBUT, madde seçimi algoritmaları ile vasıtasıyla azaltılan madde aşırı maruz kalma potansiyel sorununu karşılamak için büyük bir madde havuzuna ihtiyaç duyulmaktadır.

BOBUT çalışma prensibinde iki önemli nokta vardır: Soru seçimi ve puan tahmini (Davey, 2011). Her ikisi de bir soru sorulduğunda ve yanıtlandığında tekrarlanır. İlk olarak öğrencinin performans düzeyi hakkında hâlihazırda bilinenler göz önüne alındığında uygulanacak en uygun soru belirlenir ve soru (madde) seçimi yapılır. Ardından öğrencinin puanını veya performans tahminini hassaslaştırmak için daha önce yanıtlanan soruya/sorulara verilen yanıtlar kullanılır. Bu şekilde daha sonra sorulan maddelerin daha uygun olması sağlanır. Bilgisayara uyarlamalı testlerde, katılımcıların önceki maddelere verdikleri yanıtlar kendilerine daha sonra hangi maddelerin uygulanması gerektiğini belirler. Eğer sınav giren bir birey maddeye doğru cevap verirse, bir sonraki madde daha karmaşık ve zor olacaktır. Öte yandan eğer sınav giren kişi maddeye yanlış cevap verirse bir sonraki madde daha kolay olacaktır. Bu süreç, sınav giren kişinin yetenek düzeyi tahmin edilene kadar devam edecektir. Diğer bir ifadeyle bir veya daha fazla madde bir sınav giren kişiye uygulanır ve bu maddeler doğru veya yanlış olarak puanlanır. Sınav giren kişinin yanıtlarına dayanarak, bilinen zorluk ve ayırt ediciliğe sahip maddeler içeren madde havuzundan maddeler seçilir. Eğer bir yanıt yanlışsa, seçilen bir sonraki madde daha kolay olur veya bir yanıt doğruysa, bir sonraki madde daha zor olur. Yeni maddeler uygulanır, puanlanır ve sonraki maddeler, bireyin maddelere verdiği yanıtlara ve bazı durumlarda önceki madde yanıtlarıyla belirlenen birey için özellik seviyelerine ilişkin tahminlere dayanarak madde havuzundan seçilir. Bu şekilde maddelerin özellikleri, test süreci boyunca sınav giren kişinin özelliklerine uyarlanmış olur.

BOBUT için önemli diğer bir nokta ise testin uygulanması esnasında gerçekleştirilen yeterlik kestirimlerindeki kuramın belirlenmesi ve güvenilir sonuçlara ulaşılmasıdır. Alanyazında birçok BOBUT uygulaması (GRE, SAT, GMAT, TOEFL) Madde Tepki Kuramı esas alınarak gerçekleştirilmektedir. Bazı uyarlanabilir testler Klasik Test Kuramı (KTK) esas alınarak uygulansa da (Rudner ve Guo, 2011), MTK'nın sağladığı bazı avantajlar (bireye göre yeterlik ve hata kestirimi, madde ve test parametreleri değişmezliği) ile katılımcılara ve uygulayıcılara birtakım kolaylıklar sağlanmasından dolayı BOBUT uygulamalarında KTK yerine MTK'yı tercih edilmektedir.

2.3. Bilgisayar Ortamında Bireye Uyarlanan Testlerin (BOBUT) Yöntemi ve Adımları

BOBUT çalışma prensibinde sınav girenlerin havuzdan seçilen maddelere verdiği yanıtların kaydedilerek düzenli olarak güncellendiği bu yetenek düzeyi tetadır (θ). Maddeler bireyin yetenek (θ :teta) düzeyine uygun olarak seçilmektedir. Bu madde seçimi çeşitli yöntemlerle ve farklı şekilde yapılmaktadır. İlk aşamada katılımcının madde havuzundaki maddelerden seçilen orta zorluktaki maddeye yanıt vermesi beklenir. Bireyin maddelere verdiği cevaplarla değişen θ değeri, sonraki maddenin seçiminde önemli rol oynar. Katılımcının cevabına göre bireyin θ düzeyi tahmini elde edildikten sonra, madde havuzundaki hangi maddenin bireyin θ değerini tahmini için maksimum bilgi sağlayacağına karar verilir. Madde seçiminde temel prensip, katılımcının doğru cevapladığı maddeden sonra daha zor maddelerle, yanlış cevapladığı maddeden sonra ise daha kolay maddelerle karşılaşmasını sağlamaktır (Gür ve Hamide, 2020). Katılımcıların her bir maddeye verdikleri yanıtlardan sonra θ düzeyleri yeniden hesaplanır ve yeni bir θ tahmini yapılır. Sonlandırma kuralları uygulandığında test sona ermektedir (Bulut ve Kan, 2012). BOBUT süreci belirlenen sayıda madde uygulanana ya da belirlenen bir hata değerine ulaşana kadar devam eder ve böylece bireyin yetenek puanı ortaya çıkar.



Şekil 1. BOBUT Adımları

Kaynak: Magis, Yan ve Davier, 2017.

BOBUT uygulaması genellikle belirli adımları içeren tekrarlamalı bir işlemdir. Henüz uygulanmamış olan tüm maddeler, tahmin edilen yetenek seviyesi göz önüne alındığında, hangisinin daha sonra uygulanacak en iyi madde olacağını belirlemek için değerlendirilir. Yeteneğin en doğru olarak kestirilmesi ve analizi için BOBUT'ta takip edilmesi gereken temel bazı adımlar bulunmaktadır. BOBUT, farklı yeteneklerin ölçüldüğü ve geniş madde havuzdan ilk maddenin seçilmesiyle başlatma kuralına uygun olarak başlar. Ardından Şekil-1'de verilen döngüyle devam eder.

BOBUT uygulamasına başlamadan önce ilk yapılması gereken adım başlatma kuralının belirlenmesidir. Başlatma kuralı çoğunlukla orta güçlükte bir sorunun seçilmesi olarak uygulanır. Ancak bazı uygulamalarda katılımcılara ait önsel yetenek dağılımları göz önüne alınarak da başlatma kuralı belirlenebilir (Segall, 1993). Yetenek tahmini önceki performanslara göre birey bazında yapılabileceği gibi, grubun ortalaması alınarak da her katılımcı için aynı başlangıç yetenek tahmini de kullanılmaktadır (Kezer, 2013). Başlatma kuralının belirlenmesinin ardından bu kurala uygun ilk madde uygulanır ve maddeye verilen yanıtla göre yetenek düzeyi yeniden kestirilir. Yetenek düzeyinin kestirimi Maksimum Olabilirlik (MO), Beklenen Sonsal Dağılım (BSD) veya Modal Sonsal Dağılım (MSD) gibi yöntemler ile kestirim yapılır (& Şenel, 2024). Maksimum Olabilirlik Yöntemi, Beklenen Sonsal Dağılım (BSD) bir bireyin verdiği cevaplara en yüksek olasılıkla neden olacak yetenek düzeyini (θ) bulmayı amaçlamaktadır. Bireyin yetenek düzeyini, önsel bilgiye (prior) dayanarak sonsallık (posterior) dağılımının beklenen değeri olarak tahmin eder. Modal Sonsal Dağılım (MSD) ise en yüksek noktası (en olası θ değeri) alınarak bireyin yetenek düzeyi tahmin edilir. Her yetenek kestiriminin ardından testin ne zaman sonlandırılacağını belirleyen süreç olan sonlandırma kuralının sağlanıp sağlanmadığı kontrol edilir. Yanıtlanan maddenin ardından kestirime ait standart hata belli bir noktaya kadar bir değer almışsa veya madde sayısı daha önceden belirlenmiş bir sayıya ulaşmışsa test sonlandırılır. Sonlandırma kuralının karşılanması durumunda BOBUT uygulaması son bulur. Eğer test sonlandırılmıyorsa, bir sonraki madde tekrar aynı yöntemlerle belirlenir ve süreç yeniden başlar. Bu döngü sonlandırma kuralı karşılandığında veya uygulanacak madde bitinceye kadar devam eder (Reckase, 2011).

2.3.1. Madde Havuzu

Madde havuzu, tüm uyarlanabilir test yönetim sistemlerinin merkezinde yer almaktadır. Havuz, sınav sırasında herhangi bir öğrenciye yalnızca bir kısmı sunulacak olan organize bir soru koleksiyonudur. Bir soru havuzunun önemli özellikleri boyut, kompozisyon, yapı ve soruların uyarlandığı modeldir (Davey ve Pitoniak, 2011). Yüksek kaliteli değerlendirme öğeleri oluşturmak, BOBUT'u mümkün kılmak için yeterli psikometrik özelliklere sahip bir madde havuzu oluşturmak kritik bir adımdır, ancak karmaşık bir süreçtir (Koller, Levenson ve Glück, 2017). Maddelerin, yüksek derecede ayırt edici gücü korunurken, içerik ve zorluk gibi çok çeşitli örnekleme özelliklerinin kapsanması gerekmektedir. Düşük kaliteli maddeler, istenmeyen psikometrik özelliklere yol açabilir ve değerlendirmenin bütünlüğünü etkileyebilir (Downing, 2005).

Havuz büyüklüğü uygun bir şekilde test uzunluğunun katları ile ölçülür. Eğer bir testte her öğrenciye 30 soru soruluyorsa, 300 soruluk bir havuzda 10 test uzunluğu bulunur (Fetzer vd., 2008). Uygulamada bir havuz için en iyi boyut çeşitli faktörlere bağlıdır. Havuzun boyutu koşullara (kullanıma hazır soru sayısı) göre kısıtlanabilir ve hatta belirlenebilir. Aynı zamanda, havuzun boyutu bir testin gözlemlenen özelliklerinin çoğunu belirler (verimlilik ve güvenlik, testlerin ortak soruları paylaşan farklı öğrencilere ne ölçüde uygulandığı). Havuzun bileşimi içerdiği soruların türüne göre belirlenir. Soruların formatı, içeriği ve performans özelliklerinin tümü havuzun oluşumuna katkıda bulunur.

Havuzdaki tüm soruların MTK modeli kapsamında uygun veya uyarlanmış olması gerekir. Uyarlanabilir testler, iki temel işlevi gerçekleştirmek için arka planda çalışan MTK'ya dayanmaktadır (Baker, 2001). Birincisi, tamamen farklı sınavlara giren öğrenciler arasında karşılaştırılabilir puanlar üretmektir. İkincisi, belirli bir öğrenciye, testin başlarında gösterdiği performans düzeyi göz önüne alındığında, ona uygulanması en uygun olan soruları belirlemektir. Aralarında önemli farklar bulunan çeşitli MTK modelleri kullanılabilir. Bununla birlikte, tüm modeller, bir havuza dâhil edilmeden önce her soru için önemli yanıt verisi örneklerinin mevcut olmasını gerektirir. Bu nedenle, ilk maddeye ilişkin verilerin, test programının faaliyete geçmesinden önce toplanmalı ve uyarlanmalıdır.

2.3.2. Madde Seçimi

Çoğu BOBUT sistemi, belirli kurallara veya kısıtlamalara uygun olarak, testi alan bireyi en iyi şekilde ölçmek amacıyla soruları seçer. Uyarlanabilir test sistemlerindeki farklılıkların çoğu bu kuralların doğasındaki en iyi tanımından kaynaklanmaktadır. Bu farklılıklar üç ana kategoride toplanabilir. Bunlardan ilki, soruların havuzdan tek tek mi, yoksa önceden düzenlenmiş setler veya paketler halinde mi alındığıyla ilgilidir. İster tek tek, ister grup halinde seçilsin, sorular bir testin bir veya daha fazla ölçüm amacına veya hedefine ulaşmasına yardımcı olacak şekilde seçilir. Bunlardan en önemli amaçlardan birisi de her öğrencinin ölçüldüğü hassasiyeti en üst düzeye çıkarmaktır (Davey, 2011).

Seçim, ölçüm hassasiyetinin ötesindeki değerlendirmelere tabi olmalıdır. İlk husus, testin ölçmeyi amaçladığı madde veya içeriğe yöneliktir. İçerik gereklilikleri uzun yıllardır geleneksel kâğıt biçimindeki testler aracılığıyla karakterize edilebilir. Bu şekilde uygun bir testin içermesi gereken çeşitli içerik alanları, çeşitli formatlardaki soruların sayısı ve bunların oranı belirtilmektedir. Uyarlanabilir bir testte madde oluşturma ve seçim görevi, geleneksel bir test oluşturma görevinden daha zordur çünkü revizyon olasılığı olmadan anında gerçekleşmesi gerekir (Parshall vd., 2002).

2.4. Bilgisayar Ortamında Bireye Uyarlanan Testlerin (BOBUT) Avantajları

Bilgisayar ortamında bireye uyarlanan testlerin diğer geleneksel test yöntemlerindeki türlere göre birçok avantajı bulunmaktadır. Öncelikli olarak verimlilik bilgisayar ortamında bireye uyarlanan testlerin başlıca avantajlarından. BOBUT'lar diğer geleneksel test yöntemlerine kıyasla maliyet tasarrufu sağlar. Geleneksel yöntemlerde, farklı öğrenci gruplarına farklı testler verilir ve farklı testleri hazırlamak ve puanlamak çok zaman almaktadır (Rezaie ve Golshan, 2015). Ancak BOBUT'ların kullanımı test ha-

zırlama ve puanlama için gereken süreyi azaltır ve sonuçların tutarlılığını artırır. BOBUT uygulaması esneklik. Ne cevap kâğıtlarına ne de eğitimli test yöneticilerine ihtiyaç duyulmaktadır ve böylece test yöneticisi farklılıkları, ölçüm hatasına neden olan bir faktör olarak ortadan kaldırılmıştır.

Bilgisayara uyarlanmış testlerin hazırlanması gibi uygulanması da kâğıt- kalem testlerine kıyasla daha fazla zaman tasarrufu sağlamaktadır. Bununla birlikte bilgisayara uyarlanmış testlerde kullanılan soru sayısı ve bu soruları cevaplamak için gereken süre, geleneksel kâğıt-kalem testlerinden de daha azdır. BOBUT'larda testler talep üzerine verilir ve puanlar anında elde edilir. Bu durum sınava girenleri motive edebilir (Linacre, 2000). Sınava giren kişinin bir sonraki bölüme geçmeden önce diğerlerinin bitirmesini beklemek zorunda kalmaması testlerin hızlı olduğunun bir başka kanıtıdır. Cevaplayıcıların maddelerin kendi seviyelerine uygun olduğunu bilmesi ve kendi hızında sınavı yönetmesi sınav kaygılarını azaltabilir. Sınavı cevaplama süresinin kısa olması, sınava giren kişinin test sonuçlarını önemli ölçüde etkileyebilecek bir faktör olan sınav yorgunluğunun azalmasına da etki etmektedir.

Her test, her sınava girenin seviyesine uyarlandığı için, geleneksel testlere kıyasla bilgisayar uyarlamalı testlerden daha fazla bilgi toplanmaktadır. Kâğıt-kalem testlerinde cevaplayıcılara ait test sonuçlarından yüksek düzeyde bilgi elde etmek daha zordur (van der Linden, 2005). BOBUT'lar çok sayıda öğrencinin hemen farklı sınıflara yerleştirilmesi gerektiğinde çok faydalı olabilir. Çünkü esnek seviye stratejisi sayesinde, sınava girenlerin kendileri için çok zor veya çok kolay olan çok sayıda soruyu cevaplamalarına gerek kalmaz ve daha fazla ölçüm hassasiyetine sahiptir. Çok daha az madde ile bireylerin yeteneği etkin bir biçimde kestirebilmektedir (Ali ve Chang, 2014). Sınava katılan tüm bireylerin puanları hem doğru cevaplanan soruların yüzdesine hem de bu soruların zorluk seviyesine göre belirlenmektedir. İki sınava giren eşit oranda soruyu doğru cevaplıyorsa, daha zor soruları cevaplayan daha yüksek puan almaktadır (Economides ve Roupas, 2007).

Katılımcıların birbirlerinden kopya çekmesinin mümkün olmaması, basılı test kitapçıkları bulunmaması veya soruları önceden ele geçirme gibi güvenlik sorunları yaşanmaması sonucu test güvenliği daha iyi sağlanmaktadır. Geleneksel yöntemlerde uygulanan testlerde iki test arasında bazen birkaç ayı geçen minimum bir süre bulunmaktadır. Bilgisayar ortamında bireye uyarlanan testler için iki test arasında böyle bir zorunluluk yoktur. Her katılımcıya farklı sorular sorulacağından, daha kısa zamanda teste yeniden katılabilir ve bu durum daha sık ölçüm alma imkânı sağlamaktadır (Akyıldız, 2014).

2.5. Bilgisayar Ortamında Bireye Uyarlanan Testlerin (BOBUT) Dezavantajları

Diğer test türleri gibi, bilgisayar ortamında bireye uyarlanan testlerin de kendilerine özgü özel sınırlamaları ve dezavantajları vardır. Bu nedenle, test yöneticileri bu testleri kullanmak istediklerinde bu sınırlamaların farkında olmalıdır. BOBUT test uygulama prosedürleri kâğıt-kalem testlerine kıyasla farklıdır ve bu bazı sınavlar için sorunlu olabilir. Ayrıca BOBUT tüm dersler ve beceriler için geçerli değildir. Kâğıt-kalem testinde kullanılan madde parametreleri, bazen bilgisayar uyarlamalı testte geçerli olmayabilir. Çoğu BOBUT bir MTK'ya dayanmaktadır, ancak MTK tüm becerilere, tüm madde türlerine ve açık uçlu maddelere uygulanamamaktadır.

BOBUT'ların bir diğer önemli dezavantajı ise bir sınava giren kişinin geri dönüp cevapları değiştirmesine izin verilmemesidir. Çünkü bir sonraki maddeler daha önce cevaplanan maddelere göre seçilmektedir. Ayrıca madde tasarımı ve ayarlanması BOBUT'ların başarısını etkileyen önemli bir faktördür. Maddeler zorluk/yetenek ölçeğinde uygun şekilde ayarlanmazsa, test ne geçerli ne de güvenilir olacaktır. Her zorluk seviyesinde, o seviyede tekrarlı ölçümler yapabilmek için birkaç madde kullanılmalıdır ve BOBUT geniş madde havuzu ile çalışmaktadır (Veldkamp ve Linden, 2010). Madde havuzunun oluşturulabilmesi için geniş gruptan elde edilmiş çok sayıda maddeye ihtiyaç duyulmaktadır.

Büyük gruplar için yapılacak test uygulaması için bilgisayar donanımı sınırlılığı oluşabilir. Ayrıca donanım sınırlamaları, bilgisayar tarafından yönetilebilecek öğe türlerini kısıtlayabilir. Bazen ayrıntılı sanat eserleri ve grafikler veya kapsamlı okuma pasajları içeren öğelerin sunulması zor olabilmektedir.

Bununla birlikte uygulamanın bilgisayarda yapılması nedeniyle bilgisayar kullanımına yönelik olarak olumsuz tutum ve kaygıya sahip katılımcılar için uygulama sıkıntısı ortaya çıkabilmektedir (Tian vd., 2007). Bilgisayar ortamında bireye uyarlanan testler, sınav uygulayacak tesisin çok sayıda yeterli sınav bilgisayarına sahip olması ve sınava girenlerin en azından biraz da olsa bilgisayar okuryazarı olması durumunda yönetilebilir. Aksi durum büyük bir dezavantaja sebep olabilmektedir.

Ayrıca sınava girenlerin algıları veya tutumları ile ilgili de bazı dezavantajlara sahiptir. Sınava giren her kişinin farklı sorular alması nedeniyle eşitsizlikler olduğu algılanabilir. BOBUT'un daha sonra bir dizi kolay soru seçeceği düşüncesi ile zeki bir katılımcı kasıtlı olarak ilk soruları kaçırabilir. Bu durum sınava giren kişinin tahmini yeteneğinin en yüksek yetenek seviyesi olmamasına yol açabilir.

3. SONUÇ ve TARTIŞMA

Teknolojik gelişmeler son yüzyıldan bu yana çok hızlı bir şekilde ilerlemiş, bilgisayarlar amaçlarımızın çoğuna ulaşmada en yararlı kolaylaştırıcı haline gelmiştir. Eğitimin her aşamasında kullanılmasıyla birlikte test alanında da genişlemeye ve yeniliğe yol açan bir araç olmuştur. Bu gelişmelerin de etkisiyle günümüzde bilgisayar ortamında bireye uyarlanan testler (BOBUT), test sürecini yenilikçi, esnek, bireyselleştirilmiş, verimli ve hızlı hale getirmektedir.

Bilgisayar ortamında bireye uyarlanan testler birçok açıdan diğer geleneksel yöntemlerle uygulanan testlere göre farklılık göstermektedir. Bu farklılıkların çoğu tekniktir, inceliklidir ve çoğu öğrenci ve puan kullanıcısı için görünmezdir. Ancak bu farklılıkların, üretilen puanların kalitesi, anlamı ve kullanışlılığı üzerinde önemli etkileri bulunmaktadır. Uyarlanabilir test, deneyimli ve yetenekli kişiler tarafından en iyi şekilde kullanılan karmaşık bir araçtır. Bu deneyimli ve uzman kişilerin test sürecinde yer alması, soruların yanıtlanıp yanıtlanmayacağı ve nasıl yanıtlanacağı konusunda gerekli uzmanlığın gerçekten uygulanıp uygulanmadığı ortaya koymaktadır. Ancak herhangi bir uyarlanabilir test sisteminin yeteneklerini ve sınırlamalarını tam olarak anlamak ve bu konuları da dikkate almak gerekmektedir.

Bilgisayar ortamında bireye uyarlanan testlerin kullanımı farklı etkilere sahiptir ve birçok avantajları bulunmaktadır. Çok sayıda öğrenciyi değerlendirmek uzun zaman almaktadır ve sınava girenlere doğru ve hızlı sonuçlar sağlamak zordur. BOBUT bu konuda akademik kurumlara yardımcı olabilmektedir. BOBUT kullanımıyla, değerlendirme sonuçları çok hızlı bir şekilde elde edilebilir, insan kaynaklarına olan ihtiyaç azaltılabilir ve değerlendirme daha verimli olabilir.

Bu testlerin kullanımı öğretmenler ve öğrenciler için de önemli sonuçlara sahip olabilir. Bilgisayar okuryazarlığı olan öğretmenler ve öğrenciler bilgisayar teknolojisini kullanabilir ve yenilikçi olabilirler. Ancak öğretmenler ve öğrenciler, öğretim ve test alanındaki yeni gelişmelerin farkında olmalı ve becerilerini geliştirmelidir. Özellikle bilgisayar ortamında bireye uyarlanan testlerin kullanımıyla öğrenciler kendilerini istedikleri zaman ve kendi hızlarında değerlendirecektir. Puanlama ve değerlendirme zaman alıcı bir süreç olmayacaktır. Ancak teknolojinin etkin kullanılmasıyla uygulanan bu testlerin avantajları olduğu kadar dezavantajları da vardır. Yetkililer bir testi uyarlanabilir formatta uygulamak istediklerinde bilgisayar uyarlamalı bu testin olumlu yönleri kadar olumsuz yönlerinin de farkında olmalıdırlar.

Uygulayıcılar ve politika yapıcılar açısından eğitim ve teknik yeterliliğin sağlanması, altyapı ve erişim desteğinin geliştirilmesi, öğrenci yanıtlarının ve profillerinin korunmasına yönelik önlemlerin alınması, yasal ve yönetsel alt yapının geliştirilmesi, eğitim sisteminde resmi sınavlarda bu tür testlerin de kullanılması, milli test sistemlerinin bu standartlara göre yapılandırılması önerilmektedir.

Bireye uyarlanmış testlerin gelecekteki araştırmalar için uyarlama algoritmalarının geliştirilmesi, çok boyutlu yetenek modellerinin kullanılması, farklı yaş grupları, sosyoekonomik seviyeler ve dijital yeterlilik düzeylerine göre bireye uyarlanmış testlerin etkililiğinin karşılaştırılması, dil ve kültürler arasındaki uyumun incelenmesi ve yapay zeka entegrasyonunun sağlanması önerilebilir.

KAYNAKÇA

- Akyıldız, M. (2014). *Online ölçme ve değerlendirme, bireye uyarlanmış testlerde imkanlar, ufuklar*. Anadolu Üniversitesi, Açıköğretim Fakültesi, Yaygın Öğretim Bölümü.
- Alam, A. (2022). Cloud-based e-learning: Scaffolding the environment for adaptive e-learning ecosystem based on cloud computing infrastructure. In *Computer Communication, Networking and IoT: Proceedings of 5th ICICC 2021, Volume 2* (pp. 1–9). Springer Nature Singapore.
- Ali, U. S., & Chang, H. (2014). An item-driven adaptive design for calibrating pretest items. *ETS Research Report Series*, 2014(1), 1–12. <https://doi.org/10.1002/ets2.12044>
- Baker, F. B. (2001). *The basics of item response theory* (2nd ed.). ERIC Clearinghouse on Assessment and Evaluation, University of Maryland.
- Baykul, Y. (2000). *Eğitimde ve psikolojide ölçme: Klasik test teorisi ve uygulaması*. ÖSYM Yayınları.
- Bennett, R. E. (2002). Inexorable and inevitable: The continuing story of technology and assessment. *The Journal of Technology, Learning, and Assessment*, 1(1), 1–23.
- Bulut, O., & Kan, A. (2012). Application of computerized adaptive testing to entrance examination for graduate studies in Turkey. *Eurasian Journal of Educational Research*, 49, 61–80.
- Büyükkıdık, S., & Yörü, F. G. A. (2024). Çok aşamalı testlerin panel deseni, modül uzunluğu, örneklem büyüklüğü ve yetenek parametresi kestirim yöntemleri açısından farklı koşullar altında karşılaştırılması. *Boğaziçi University Journal of Education*, 41(2), 9–27.
- Collares, C. F., & Cecilio-Fernandes, D. (2019). When I say ... computerised adaptive testing. *Medical Education*, 53(2), 115–116.
- Davey, T. (2011). *A guide to computer adaptive testing systems*. Council of Chief State School Officers.
- Davey, T., & Pitoniak, M. J. (2011). Designing computerized adaptive tests. In *Handbook of test development* (pp. 557–588). Routledge.
- Doğan, A. G. (2022). Zihinsel hız: Kavramsal gelişimi ve önemine dair bir derleme çalışması. *Sosyal Bilimler Araştırmaları Dergisi*, 12(3).
- Downing, S. M. (2005). The effects of violating standard item writing principles on tests and students. *Advances in Health Sciences Education*, 10, 133–143.
- Drasgow, F. (2002). The work ahead: A psychometric infrastructure for computerized adaptive tests. In C. N. Mills et al. (Eds.), *Computer-based testing* (pp. 67–88). Lawrence Erlbaum.
- Dumas, H. M., et al. (2012). Computer adaptive test performance in children with and without disabilities. *Disability and Rehabilitation*, 34(5), 393–401.
- Economides, A. A., & Roupas, C. (2007). Evaluation of computer adaptive testing systems. *International Journal of Web-Based Learning and Teaching Technologies*, 2(1), 70–87.
- Fergadiotis, G., et al. (2021). Computer adaptive testing for the assessment of anomia severity. *Seminars in Speech and Language*, 42(3), 180–191.
- Fetzer, M., et al. (2008). Computer adaptive testing (CAT) in an employment context. *PreVisor's PreView*.
- Fleming, S., & Hipple, D. (2004). Foreign language distance education at the University of Hawai'i. In C. A. Spreen (Ed.), *New technologies and language learning* (pp. 13–54). University of Hawai'i.
- Gao, I., et al. (2023). Adaptive testing of computer vision models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 4003–4014).
- Gür, R., & Hamide, D. G. (2020). The effect of item exposure control methods on measurement precision. *Eğitim ve Bilim*, 45(202).
- Gür, R., & Karabay, E. (2018). Simülasyon ortamında bireye uyarlanmış test uygulaması. *Journal of Theoretical Educational Science*, 11(2), 201–228.

- Kalender, İ. (2004). Bilgisayar ortamında bireyselleştirilmiş testlerin eğitimde kullanımı. *XIII. Ulusal Eğitim Bilimleri Kurultayı*, İnönü Üniversitesi.
- Kezer, F. (2013). *Bilgisayar ortamında bireye uyarlanmış test stratejilerinin karşılaştırılması* (Yayımlanmamış doktora tezi). Ankara Üniversitesi.
- Kezer, F., & Koç, N. (2014). Bilgisayar ortamında bireye uyarlanmış test stratejilerinin karşılaştırılması. *Eğitim Bilimleri Araştırmaları Dergisi*, 4(1), 145–174.
- Kim, D. H., & Huynh, H. (2007). Comparability of computer and paper-and-pencil versions. *Journal of Technology, Learning, and Assessment*, 6(4).
- Koller, I., Levenson, M. R., & Glück, J. (2017). What do you think you are measuring? *Frontiers in Psychology*, 8, 126.
- Kuzucu, R. E. (2024). *Türkiye’de bilgisayar ortamında bireye uyarlanmış test (BOBUT) çalışmaları* (Yayımlanmamış yüksek lisans tezi). Harran Üniversitesi.
- Larson, J. (1999). Considerations for testing reading proficiency via CAT. *Studies in Language Testing*, 10, 71–90.
- Larson, J. W., & Madsen, H. S. (1984). Computerized adaptive language testing. *Calico Journal*, 2(3), 32.
- Linacre, J. M. (2000). Computer-adaptive testing: A methodology whose time has come. In S. Chae et al. (Eds.), *Development of computerized middle school achievement test*. Komesa Press.
- Lord, F. M. (1971). The self-scoring flexilevel test. *Journal of Educational Measurement*, 8(3), 147–151.
- Magis, D., Yan, D. A., & van der Linden, W. J. (2017). *Computerized adaptive and multistage testing with R*. Springer.
- Martin, A. J., & Lazendic, G. (2018). Computer-adaptive testing: Implications. *Journal of Educational Psychology*, 10(1), 27–45.
- Mojarrad, H., et al. (2013). CBA vs. PPBA: EFL learners’ reading. *International Journal of Language Learning and Applied Linguistics World*, 4(4), 418–428.
- Montgomery, J. M., & Cutler, J. (2013). CAT for public opinion surveys. *Political Analysis*, 21(2), 172–192.
- Özbaşı, D., & Demirtaşlı, N.** (2015). Bilgisayar okuryazarlığı testinin BOBUT olarak geliştirilmesi. *Journal of Measurement and Evaluation in Education and Psychology*, 6(2).
- Özgüven, İ. E.** (2007). *Psikolojik testler* (4. Baskı). Nobel Yayın Dağıtım.
- Parshall, C., et al. (2002). *Practical considerations in computerized testing*. Springer.
- Pommerich, M. (2004). Mode effects for passage-based tests. *The Journal of Technology, Learning, and Assessment*, 2(6), 3–44.
- Reckase, D. M. (2011). Computerized adaptive assessment. In *The Road Ahead for State Assessments*.
- Rezaie, M., & Golshan, M. (2015). CAT: Advantages and limitations. *International Journal of Educational Investigations*, 2(5), 128–137.
- Rice, N., et al. (2022). Development of a computer adaptive progress test. *Computers and Education: Artificial Intelligence*, 3, 100083.
- Rudner, L., & Guo, F. (2011). CAT for small scale programs. *Graduate Management Admission Council*.
- Schleicher, A. (2019). *PISA 2018: Insights and interpretations*. OECD Publishing.
- Segall, D. O. (1993). Score equating analyses of the CAT-ASVAB. *Defense Advisory Committee Briefing*.
- Şenel, S.** (2024). Comparison of CAT procedures at low ability levels. *Bartın University Journal of Faculty of Education*, 13(3), 547–559.
- Thorndike, R. M., et al. (1991). *Measurement and evaluation in psychology and education*. Macmillan.
- Tian, J., et al. (2007). Introduction to computerized adaptive testing. *US-China Education Review*, 4(1), 72–81.
- Tonidandel, S., et al. (2002). Impact of CAT characteristics. *Journal of Applied Psychology*, 87(2), 320.

- Turgut, M. F., & Baykul, Y. (2010). *Eğitimde ölçme ve değerlendirme* (Vol. 2). Pegem Akademi.
- van der Linden, W. J. (2005). A comparison of item-selection methods. *Law School Admission Council Report: 04-02*.
- Veldkamp, B. P., & van der Linden, W. J. (2009). Designing item pools. In *Elements of Adaptive Testing* (pp. 231–245). Springer.
- Wainer, H., et al. (2000). *Computerized adaptive testing: A primer*. Routledge.
- Wang, T., & Kolen, M. J. (2001). Evaluating comparability in CAT. *Journal of Educational Measurement*, 38(1), 19–49.
- Weiss, D. J. (1983). Latent trait theory and adaptive testing. In D. J. Weiss (Ed.), *New Horizons in Testing* (pp. 5–7). Academic Press.
- Weiss, D. J. (2004). CAT in counseling and education. *Measurement and Evaluation in Counseling and Development*, 37(2), 70–84.
- Wise, S. L., & Kingsbury, G. G. (2000). Developing and maintaining a CAT program. *Psicológica*, 21(1), 135–155.
- Zickar, M. J., et al. (1999). Development of a computerized selection system. In *Innovations in Computerized Assessment* (pp. 7–33). Psychology Press.