



Motorlu Taşıtlar Sürücü Kursiyerleri Sınavının Simülatif Bilgisayar Ortamında Bireye Uyarlanmış Test Uygulaması

Simulative Computerized Adaptive Motor Vehicle Driving License Exam Test Administration

Recep GÜR*

Ersoy KARABAY**

Received: 23 February 2015

Accepted: 05 February 2018

ABSTRACT: The aim of this study is to compare simulatively administration of Motor Vehicle Driving Trainees Exam Test as computerized adaptive test [CAT] with paper-pencil test method. This study is conducted by using basic research model. Research population consists of the candidates that attended B type certificate exam at October 23rd, 2010 and February 12nd, 2011, and research sample consists of 5000 candidates being chosen randomly among those candidates. Data analysis is done via XCalibre 4.2 and CATSim programs. It is found in the study that 50 items are applied to candidates in Motor Vehicle Driving Trainees paper-pencil test while minimum 27, maximum 50 and mean 32.52 items are applied simulative computerized adaptive Motor Vehicle Driving Trainees Exam test. When it is considered that in paper-pencil test each candidate encounters 50 items, it is found that candidates encounter less item at the rate of 34.96% in simulative computerized adaptive Motor Vehicle Driving Trainees Exam test in compare with paper –pencil test application. Theta levels being estimated in Motor Vehicle Driving Trainees Exam Test between CAT and paper-pencil test is $r=.94$, $p<.01$. This shows that these applications estimate similar theta levels.

Keywords: motor vehicle, driving test, computerized adaptive test.

ÖZ: Motorlu Taşıtlar Sürücü Kursiyerleri Sınavı'nın Bilgisayar Ortamında Bireye Uyarlanmış Test uygulaması ile kâğıt kalem test yönteminin simülatif olarak karşılaştırılmasının amaçlandığı bu çalışma temel araştırma modelindedir. Araştırmanın evrenini 23.10.2010 ve 12.02.2011 tarihlerinde yapılan B sertifika türü sürücü sınavına katılan adaylar; örneklemini ise bu adaylar arasından tesadüfi olarak seçilen 5000'er aday oluşturmaktadır. Verilerin analizinde XCalibre 4.2 ve CATSim programlarından yararlanılmıştır. Araştırma sonucunda, Motorlu Taşıtlar Sürücü Kursiyerleri Sınavı kâğıt kalem testinde adaylara 50 madde uygulanmakta iken, simülatif Bilgisayar Ortamında Bireye Uyarlanmış Test uygulaması sonucunda adaylara minimum 27, maksimum 50 ve ortalama 32.52 madde uygulanmıştır. Dolayısıyla yetenek kestirimi için, simülatif BOBUT uygulamasında adaylara uygulanan ortalama madde sayısı bakımından, Motorlu Taşıtlar Sürücü Kursiyerleri Sınavı kâğıt kalem testindeki madde sayısına göre %34.96 oranında ekonomiklik sağlanmıştır. Motorlu Taşıtlar Sürücü Kursiyerleri Sınavı'nın simülatif Bilgisayar Ortamında Bireye Uyarlanmış Test uygulaması ile kâğıt kalem testi uygulamasından elde edilen yetenek düzeyleri arasında ($r=.94$, $p<.01$) elde edilmesi uygulamaların benzer yetenek kestiriminde bulunduğunu göstermektedir.

Anahtar kelimeler: motorlu taşıt, sürücü sınavı, bilgisayar ortamında bireye uyarlanmış test.

* Asst. Prof. Dr., Erzincan University, Erzincan, Turkey, rgur@erzincan.edu.tr

** Corresponding Author: Asst. Prof. Dr., Hasan Kalyoncu University, Gaziantep, Turkey, ekarabay@gmail.com

Citation Information

Gür, R. & Karabay, E. (2018). Motorlu taşıtlar sürücü kursiyerleri sınavının simülatif bilgisayar ortamında bireye uyarlanmış test uygulaması. *Kuramsal Eğitim Bilim Dergisi [Journal of Theoretical Educational Science]*, 11(2), 201-228.

Giriş

Günümüz dünyasında teknoloji özellikle de bilgisayar teknolojisi hayatın hemen hemen her alanına girmiş bulunmaktadır. İnsanlar teknolojik gelişmeler sayesinde daha önce aylar süren çalışmaları birkaç saat bazen de saniyeler içinde yapabilmektedirler. Bilgisayar teknolojisindeki bu gelişmelerden sınavların da etkilenmesi kaçınılmazdır. Kâğıt-kalem sınavlarıyla çok zahmetli olan ve uzun süren çalışmalar elektronik sınav (e-sınav) uygulamalarıyla çok daha kısa sürede daha sorunsuz bir şekilde sonuçlanabilmektedir. Milli Eğitim Bakanlığı [MEB], bu gelişmelerden geri kalmamak anlamında Motorlu Taşıt Sürücü Adayı Sınavı'nda [MTSAS] e-sınav uygulamalarına başlamıştır. Ancak bu Elektronik-sınav uygulamaları henüz bilgisayar ortamında bireye uyarlanmış test [BOBUT] olarak yapılmamaktadır.

MTSAS, karayollarında seyreden araçları kullanacak adaylara yönelik yazılı ve uygulama sınavlarını kapsamaktadır (MEB, 2014a). Bu iki sınav sonucundaki performanslarına göre adayların sürücü ehliyeti alıp almayacaklarına karar verilmektedir. MTSAS kâğıt-kalem sınavı olarak Türkiye genelinde, e-sınav olarak ise sadece Ankara'da Beşevler Otelcilik ve Turizm Meslek Lisesi'nde bir salonda yapılmaktadır. MEB tarafından 2015 yılına kadar MTSAS yılda dokuz kez uygulanmakta iken, 2015 yılından itibaren Motorlu Taşıt Sürücü Kursiyerleri Sınavı [MTSKS] adıyla yılda altı kez uygulanmaktadır. MTSKS'de ilk yardım, trafik ve çevre bilgisi, araç tekniği olmak üzere üç temel alandan sırasıyla 13, 27 ve 10 soru olmak üzere toplamda 50 soru sorulmakta ve kâğıt-kalem sınavını tamamlamak için 60 dakika süre verilmektedir. MEB tarafından yapılan e-sınav uygulamalarında ise 60 soru için 45 dakika süre verilmektedir (MEB, 2014b).

Ülkelerin sürücü ehliyeti sınavları genelde teori ve uygulama olmak üzere iki bölümden oluşmaktadır. Teori bölümünden başarılı olan adaylar uygulama sınavına girerek bu bölümde de başarılı olmaları halinde sürücü ehliyeti almaya hak kazanmaktadırlar. Bu araştırmanın konusunu teori testleri ilgilendirdiği için bazı ülkelerin sürücü ehliyeti ile ilgili bilgiler aşağıda verilmiştir.

İngiltere sürücü ehliyeti sınavlarına 17 yaşını dolduranlar girebilmektedir. Teori sınav çoktan seçmeli test ve tehlike algı testi (hazard perception test) olmak üzere iki bölümden oluşmaktadır. İki testten de ayrı ayrı başarılı olmak gerekmektedir. Bu testler bilgisayar ortamında da yapılabilmektedir (www.gov.uk, 2014). Almanya'da sürücü ehliyeti almak için diğer ülkelerde olduğu gibi teori ve uygulama olmak üzere iki aşamalı sınavda başarılı olmak gerekmektedir. Test çoktan seçmeli maddelerden oluşmakta ve bilgisayar ortamında uygulanmaktadır. Almanya'daki sınavlarda yanlış yanıtlardan dolayı ceza puanı verilmektedir ve 10 ceza puanı alan aday başarısız sayılmaktadır (www.berlin.angloinfo.com, 2014). İtalya'da da teori ve uygulama olmak üzere iki sınavdan başarılı olmak gerekmektedir. 18 yaşını dolduranların girebildiği sınavda çoktan seçmeli maddeler sorulmaktadır (www.italy.angloinfo.com, 2014). Finlandiya'da ise teori ve uygulama sınavları aynı merkezde yapılmaktadır ve teori sınavları bilgisayar ortamında yapılmaktadır (www.expat-finland.com, 2014).

Avustralya'da tam bir sürücü belgesi sahibi olmanın ilk basamağı olan sürücü bilgi testi bilgisayarda yürütülen ve yaklaşık 600 soru içeren bir soru bankasından rastgele seçilen 45 sorudan oluşan bir sınavdır. Her soru sırayla ekrana gelmektedir. Bir soruyu yanıtlamak için ekrandaki üç seçenektan birine dokunulması gerekmektedir.

Sınav genel sürücü bilgisi ve yol güvenliği konularını içermektedir (www.rms.nsw.gov.au, 2014).

Yukarıda örnek olarak verilen ülkelerin tamamında sürücü ehliyeti teorik sınavları bilgisayar ortamında yapılmaktadır. Türkiye'nin de e-sınav uygulamalarını yaygınlaştırmasının gerekli olduğu düşünülmektedir. Ayrıca, bilgisayar ortamında yapılan sınavlar ve BOBUT uygulamaları için geniş madde havuzlarına ihtiyaç duyulmaktadır (Embretson & Reise, 2000). Bu yüzden Türkiye'de MTSKS uygulamalarında kullanılan madde havuzunun genişletilmesi gerekli görülmektedir.

BOBUT uygulamaları, Madde Tepki Kuramı'nın (MTK) değişmezlik özelliğini kullanarak, her birey için iyi ölçme yapabilen maddelerin madde havuzundan seçilerek bireye sunulmasını sağlayan bir algoritma ile çalışmaktadır (Embretson & Reise, 2000). Dolayısıyla kâğıt kalem testlerine göre, daha az madde ile daha güvenilir kestirimler yapabileceği tanınmaktadır. MTSKS'nin BOBUT olarak uygulanması, bireysel farklılıkları gözetenek, bireyin yetenek düzeyine en uygun maddelerin bireye uygulanması ile gerçekleştirilen ölçme yaklaşımı olanağı sağlayacaktır.

BOBUT, temel olarak uygulayıcıların gereksinimlerini karşılamak ve bireylerin yetenek düzeylerini belirlemek için oluşturulmuştur. Bireylere kendi yetenek düzeylerine göre çok kolay bir ölçme aracı uygulamak, zaman kaybının yanı sıra dikkatsizce hata yapılması ve kritik maddelere bilerek yanlış yanıtların seçilmesi gibi istenmeyen birey davranışlarına neden olabilmektedir. Diğer taraftan bireyin yeteneğinin çok üzerinde ölçmeler yapan bir testi uygulamak da bilgi verici olmayan sonuçları ortaya çıkaracaktır. Çünkü bireyler maddeleri yanıtlamak için ciddi bir girişimde bulunmayacak, tahmin yoluyla, seçenekler yardımıyla ya da istenmeyen davranışlara yönelerek doğru yanıtları bulmaya çalışacaklardır. Bu nedenle bireyin ölçülen özellikteki yerine bağlı olarak, bireylere test uygulaması yapmak daha güvenilir ve geçerli ölçmeler yapmaya olanak sağlamaktadır (Kalender, 2011; Öztuna, 2008). BOBUT uygulamaları ehliyet sınavlarında, askeri eğitim sınavlarında, özel sektördeki iş başvuru sınavlarında, lise sonrası eğitim için giriş sınavlarında ve sertifika sınavlarında yapılmaktadır (Russo, 2002; Trotter, 2001).

Türkiye'de yapılan araştırmalarda, genelde sürücü eğitimi programının değerlendirilmesi (Vursavaş, 2004), Türkiye ile Avrupa Birliği ülkelerinde uygulamalı ve teorik eğitim için ayrılan süreler (Güler, 2010; Selbes, 2008), trafik kazalarında önemli olan faktörler (Özdemir, 2010), Avrupa Birliği ülkeleri ile Türkiye'de okullarda verilen trafik güvenliği eğitimlerinin karşılaştırılması (Payam, 2012), Türkiye'deki trafik güvenliği eğitiminin bütünlüklü değerlendirilmesi (İnal, 2014) konularına odaklanılmıştır.

BOBUT uygulamaları Türkiye'de özellikle ulusal düzeyde yapılan sınavlarda bugüne kadar yaygınlaşmamıştır. Donanım ve yazılım gereksiniminin yanında ulusal sınavlarda maddelerin önceden denenerek madde havuzlarının oluşturulmamasının bu durumda rol oynadığı düşünülmektedir. MTSKS BOBUT uygulamasının sağlayacağı avantajlar göz önüne alındığında, adayların daha kısa sürede, daha az madde ile geçerli ve güvenilir yetenek düzeyinin belirlenmesi için, alınması gereken önlemler ya da yapılması gereken yatırımların neler olması gerektiği hususunda MEB'e önemli katkılar sağlayacağı düşünülmektedir. Ayrıca literatürde, MTSKS simülatif BOBUT uygulanması ile ilgili bir çalışmaya rastlanamamıştır. Bu çalışmanın literatürdeki söz

konusu eksikliği gidermeye katkı sağlayacağı düşünülmektedir. Ulusal düzeydeki sınavlarda büyük madde havuzları oluşturarak BOBUT uygulamasına geçilmesinin ve bu uygulamanın yaygınlaştırılmasının gerekli olduğu düşünülmektedir.

MTSKS uygulamalarının BOBUT olarak yapılması hem daha güvenilir sınavlar yapılmasına hem de daha kısa sürede daha iyi yetenek kestirimi yapılmasına imkan tanıyacaktır. Ayrıca MTSKS'nin BOBUT olarak uygulanması ile sürücü adaylarının yetenekleri kâğıt-kalem testine göre daha az hata ile kestirileceğinden dolayı gerçekten hak edenlerin sürücü ehliyeti almasına olanak sağlayacağı düşünülmektedir. Bu düşünceden hareketle, MTSKS'nin BOBUT olarak uygulanacağı simülatif bir çalışma yapılmasının gerekli olduğu düşünülmektedir.

Bu araştırmanın amacı, Motorlu Taşıtlı Sürücü Kursiyerleri Sınavı'nın (MTSKS) Bilgisayar Ortamında Bireye Uyarlanmış Test (BOBUT) uygulaması ile MTSKS kâğıt kalem test yönteminin simülatif olarak karşılaştırılmasıdır. Bu doğrultuda BOBUT uygulamasında madde sayısının dağılımı, BOBUT ve kâğıt kalem testi uygulamalarında kestirilen yetenek parametreleri arasında manidar bir ilişki olup olmadığı incelenmiştir. Bu amaç doğrultusunda aşağıdaki sorulara yanıt aranmıştır:

1. MTSKS simülatif BOBUT uygulamasında yetenek kestirimi için adaylara;
 - a) Minimum kaç madde uygulanmıştır?
 - b) Maksimum kaç madde uygulanmıştır?
 - c) MTSKS kâğıt kalem testindeki madde sayısı ile karşılaştırıldığında uygulanan ortalama madde sayısında ne derecede azalma sağlamaktadır?
2. MTSKS simülatif BOBUT ile kâğıt kalem testi uygulamalarında kestirilen yetenek parametreleri arasında manidar bir ilişki var mıdır?

Yöntem

Bu bölümde araştırmanın modeli, evren ve örneklem, veri toplama araçları, verilerin toplanması ve analizine yer verilmiştir.

Araştırmanın Modeli

Motorlu Taşıtlı Sürücü Kursiyerleri Sınavı'nın (MTSKS) Bilgisayar Ortamında Bireye Uyarlanmış Test (BOBUT) uygulaması ile MTSKS kâğıt kalem test yönteminin simülatif olarak karşılaştırılmasının amaçlandığı bu çalışma temel araştırma modelindedir. Çünkü bu çalışma mevcut kuramsal bilginin gelişmesine ve genişlemesine katkıda bulunmanın yanı sıra uygulamaya da katkı getirecektir (Karasar, 2012).

Evren ve Örneklem

Bu çalışmanın evrenini 23.10.2010 ve 12.02.2011 tarihlerinde yapılan MTSAS uygulamalarına katılan adaylar oluşturmaktadır. Her iki uygulama için B sertifika ve T kitapçık türünü alan adaylar araştırmaya dâhil edilmiştir. Araştırmaya dâhil edilen adayların dağılımı Tablo 1'de verilmiştir.

Tablo 1

2010 ve 2011 MTSAS Uygulamalarına Katılan Adayların Dağılımı

	2010	2011
İlk Yardım	65.789	69.157
Motor ve Araç Tekniği	67.139	69.630
Trafik ve Çevre	68.346	74.995

Tablo 1’de gösterilen adaylar arasından her bir uygulama ve alt test için 5000’er aday tesadüfi örnekleme yöntemi kullanılarak belirlenmiştir. Araştırma kapsamında MTSKS simülatif BOBUT uygulaması için simülatif veri üretiminde, aynı 5000’er adayın betimsel istatistiklerinden faydalanılmıştır.

Veriler ve Toplanması

Çalışma kapsamında kullanılan veriler, 2010 ve 2011 yılında yapılan MTSAS B sertifika türü T kitapçığını alan adaylar için tesadüfi örnekleme yöntemiyle seçilen 5000’er adayın verdikleri yanıtlar üzerinden oluşturulmuştur. 2010 ve 2011 yıllarında yapılan MTSAS, ilk yardım, motor ve trafik alt testlerinden oluşmakta ve bu alt testler sırasıyla 30, 40 ve 50 olmak üzere toplamda 120 çoktan seçmeli maddeden oluşmaktadır. Her bir madde 4 seçenekten oluşmaktadır.

2010 ve 2011 yıllarında yapılan iki MTSAS uygulamasındaki maddeler ve adayların bu maddelere verdikleri yanıtlar Milli Eğitim Bakanlığı’ndan alınmıştır. Madde havuzunu oluşturmak için 2010 ve 2011 yıllarında yapılan iki MTSAS uygulamasındaki toplam 240 madde kullanılmıştır. Önce tek boyutluluk, yerel bağımsızlık, parametrelerin değişmezliği test edilmiş, sonra kullanılacak lojistik modele karar verilmiştir. Madde parametreleri ve bireylerin yetenek düzeyleri kestirildikten sonra modelle uyumlu maddelerle madde havuzu oluşturulmuştur.

Simülatif Verilerin Üretilmesi

Araştırmanın amacı doğrultusunda MTSKS simülatif BOBUT uygulaması için simülatif veriden faydalanılmıştır. Simülatif verilerin üretilmesi için CATSim programından faydalanılmıştır. Simülatif BOBUT uygulaması için, 2010 ve 2011 yıllarında MTSAS’ı alan 5000’er adaya ilişkin yetenek ortalamalarının dağılımına ve iki parametrelili lojistik (2PL) modele uyum gösteren 230 maddenin parametrelerine ilişkin dağılıma benzer olacak şekilde dağılım gösteren 5000 kişilik simülatif veri ve 230 maddelik simülatif madde havuzu kullanılmıştır. Gerçek uygulama sonucunda elde edilen betimsel istatistiklere dayalı olarak üretilen simülatif verilerin betimsel istatistikleri Tablo 2’de verilmiştir.

Tablo 2

Gerçek Uygulama Sonucunda Elde Edilen Betimsel İstatistiklere Dayalı Olarak Üretilen Simülatif Verilerin Betimsel İstatistikleri

	MTSAS'a Katılan Bireylerin Yetenek Düzeyleri	Madde Havuzu		Simülatif Bireylerin Yetenek Düzeyleri	Simülatif Madde Havuzu	
		a	b		a	b
N	5000	230	230	5000	230	230
\bar{x}	-.02	.61	-1.35	-.02	.61	-1.35
K_y	-.20	.11	1.15	.00	.19	1.12
B_s	-.29	-.48	2.72	.015	-.08	1.06

K_y : Çarpıklık katsayısı B_s : Basıklık katsayısı

Tablo 2 incelendiğinde, MTSAS uygulamalarına katılan bireylerin yetenek düzeyleri ortalamalarına ait $K_y = -.20$, $B_s = -.29$ katsayılarının (Leech, Barrett & Morgan'a (2005) göre, K_y ve B_s katsayıları ± 1 aralığında olduğu için normal dağılım) yanı sıra $\bar{x} = -.02$ (Ek A, Şekil 2) temel alınarak Ek A'da yer verilen Şekil 3'teki gibi $\bar{x} = -.02$ ve normal dağılım gösteren 5000 kişilik bir simülatif veri oluşturulmuştur.

Tablo 2'de madde havuzundaki maddelerle paralel olacak şekilde, EK B'de yer verilen Şekil 4'teki gibi dağılım gösteren a parametreleri için $K_y = .11$, $B_s = -.48$ (K_y ve B_s katsayıları ± 1 aralığında olduğu için normal dağılım) ve $\bar{x} = .61$; Şekil 6'daki gibi dağılım gösteren b parametreleri için $K_y = 1.15$, $B_s = 2.72$ ($K_y > 0$ olduğu için sağa çarpık ve $B_s > 0$ olduğu için sivri bir dağılım) ve $\bar{x} = -1.35$ değerleri temel alınarak normal dağılım özelliği gösteren a parametreleri (Şekil 5) ve sağa çarpık ve sivri bir dağılım özelliği gösteren b parametrelerine (Şekil 7) sahip 230 maddelik sanal bir madde havuzu oluşturulduğu görülmektedir.

EK B'de yer verilen Şekil 8'deki simülatif madde havuzundaki 230 maddeye ait a ve b parametrelerine ilişkin saçılma diyagramı incelendiğinde, a parametrelerinin .40 ile .70 arasında; b parametrelerinin ise -3.00 ile -1.00 arasında yoğunluk gösterdiği; bu saçılma diyagramına paralel olarak EK C'deki madde havuzunun bilgi fonksiyonu (Şekil 9) ve madde havuzunun standart hata fonksiyonu (Şekil 10) incelendiğinde -3.00 ile -1.00 aralığında madde havuzunun maksimum bilgi verdiği ve minimum hata gösterdiği söylenebilir.

Verilerin Analizi

2010 ve 2011 yıllarında ilk yardım, motor ve trafik alt testlerinden elde edilen toplam puana ait betimsel istatistikler Tablo 3'te verilmiştir.

Tablo 3
Betimsel Test İstatistikleri

	2010			2011		
	İlk Yardım	Motor	Trafik	İlk Yardım	Motor	Trafik
Ortalama	22.61	26.80	36.63	22.52	30.46	37.31
Ortanca	24.00	27.00	38.00	23.00	31.00	38.00
Tepe Değer (Mod)	26.00	29.00	44.00	25.00	32.00	41.00
Varyans	21.89	39.91	58.84	22.91	36.89	52.11
Standart Sapma	4.68	6.32	7.67	4.79	6.07	7.22
Minimum Değer	0.00	1.00	0.00	0.00	0.00	2.00
Maksimum Değer	30.00	40.00	50.00	30.00	40.00	50.00
Ranj	30.00	39.00	50.00	30.00	40.00	48.00
Çarpıklık	-.92	-.63	-.85	-.84	-1.01	-.91
Basıklık	.74	.24	.83	.80	1.77	1.02
KR-20	.82	.84	.88	.82	.85	.87

Tablo 3 incelendiğinde, 2011 motor ve trafik alt testleri dışında çarpıklık ve basıklık katsayılarının ± 1 aralığında ve tepe değer, medyan, aritmetik ortalamalarının birbirine yakın değerler olması tek değişkenli normalliğin sağlandığına ilişkin bilgi vermektedir (Leech, Barrett, & Morgan, 2005). Varyans, standart sapma ve ranj değerleri incelendiğinde, alt testleri alan bireylerin heterojen grup oldukları, her bir alt test için en yüksek puan alan bireylerin tüm maddeleri doğru yanıtladığı; en düşük puan alan bireylerin ise 2010 motor alt testi (minimum 1 madde) ve 2011 trafik alt testi (minimum 2 madde) dışında diğer alt testlerde hiçbir maddeyi doğru yanıtlayamadığı söylenebilir. Madde ve birey parametreleri kestirilip verilerin MTK'ya uygunluğu (tek boyutluluk, yerel bağımsızlık, hız testi kontrolü, model veri uyumu, madde ve yetenek parametrelerinin değişmezliği) incelendikten sonra, BOBUT uygulaması simülatif olarak gerçekleştirilmiştir.

Tek boyutluluk, maddelerin yalnızca tek bir yeteneği veya başat bir faktörü ölçmesi şeklinde tanımlanabilir (Hambleton, Swaminathan, & Rogers, 1991). Uygulamada tek boyutluluk varsayımının karşılanması oldukça zordur; çünkü yanıtlayıcının test performansına heyecan, motivasyon gibi diğer faktörler de etki etmektedir. Bu nedenle, bu varsayımın karşılanması için yanıtlayıcının test performansı üzerinde etkili olan başat bir bileşenin varlığı aranmaktadır. Bu başat bileşen de testin ölçtüğü tek yetenek olarak adlandırılmaktadır (Hambleton & Swaminathan, 1985).

Testin tek boyutlu olup olmadığını belirlemek için tetrakorik korelasyon matrisine dayalı temel bileşenler analizi yapılmıştır. Öncelikle temel bileşenler analizinin varsayımları incelenmiştir. Veri seti büyüklüğünün temel bileşenler analizine uygun olduğuna karar vermek için KMO testi kullanılmıştır. 2010 "ilk yardım", "motor

ve araç tekniği”, “trafik ve çevre” için sırasıyla KMO ve Bartlett testi sonuçları ; .93 ($\chi^2 = 16214.662, p < .01$), .92 ($\chi^2 = 19115.88, p < .01$), .93 ($\chi^2 = 28569.28, p < .01$); 2011 “ilk yardım”, “motor ve araç tekniği”, “trafik ve çevre” için KMO ve Bartlett testi sonuçları sırasıyla .92 ($\chi^2 = 14280.04, p < .01$), .93 ($\chi^2 = 22303.94, p < .01$) ve .94 ($\chi^2 = 26833.41, p < .01$) olarak hesaplanmıştır. KMO değerleri .90’dan büyük olduğu için, veri yapısının temel bileşenler analizi yapabilmek için, mükemmel derecede yeterli olduğu söylenebilir (Leech, Barrett, & Morgan, 2005).

Veri setinin normallik varsayımını karşılayıp karşılamadığını kontrol etmek amacıyla Bartlett testi sonuçları incelendiğinde ise, elde edilen ki-kare değerlerinin .01 düzeyinde manidar olduğu görülmektedir. Bu sonuç, verilerin çok değişkenli normal dağılım özelliği gösterdiği ve dolayısıyla temel bileşenler analizinin bir diğer varsayımının karşılandığı anlamına gelmektedir. Elde edilen sonuçlar, verilere temel bileşenler analizi uygulamanın uygun olduğunu göstermektedir. Verilerin tek boyutluluk açısından incelenmesi için Tablo 4’te bileşenlere ilişkin özdeğerler ve açıklanan varyans oranları verilmiştir.

Tablo 4

Bileşenlere İlişkin Özdeğerler ve Açıklanan Varyans Oranları

	Alt Testler	Bileşenler	Özdeğer	Açıklanan Varyans (%)
2010	İlk yardım	1	5.15	17.18
		2	1.70	5.68
	Motor ve Araç Tekniği	1	5.61	14.02
		2	1.89	4.79
	Trafik ve Çevre	1	7.07	14.13
		2	2.05	4.09
2011	İlk yardım	1	4.88	16.28
		2	1.44	4.81
	Motor ve Araç Tekniği	1	6.22	15.55
		2	1.77	4.42
	Trafik ve Çevre	1	7.03	14.05
		2	2.06	4.12

Tablo 4 incelendiğinde, açıklanan varyans oranları düşük olmasına rağmen, birinci faktörden sonra diğer faktörlerin toplam varyans yüzdesine yaptığı katkı azalmaktadır. Ayrıca birinci bileşendeki özdeğerler ile ikinci bileşendeki özdeğerler arasındaki fark 3-3.5 kattan fazla olmasının yanı sıra birinci faktörden sonra diğer faktörlerin toplam varyans yüzdesine yaptığı katkı azaldığı için testlerin tek boyutlu olduğu sonucuna ulaşılmıştır (Lord, 1980). Testin tek boyutlu olması, yerel bağımsızlık

varsayımının da kabul edilmesi için yeterli bir sonuçtur (Hambleton & Swaminathan, 1985). Yerel bağımsızlık, belli bir yetenek düzeyindeki bir bireyin, bir maddeye verdiği yanıtın, diğer maddelere verdiği yanıtlardan etkilenmemesidir (Embretson & Reise, 2000).

Yanıtlayıcılar tarafından (5000'er aday), 2010 ve 2011 "ilk yardım", "motor ve araç tekniği" ve "trafik ve çevre" alt testleri için yanıtlanmış maddelerin sayısı incelendiğinde, adayların her bir alt testteki maddelerin %95'inden daha fazlasını yanıtladığı (kayıp veri oranının her bir alt test için %5'ten az olduğu) görülmektedir. Hemen hemen bütün yanıtlayıcılar, maddelerin neredeyse hepsini yanıtladıysa, hızın test performansı üzerinde önemsiz bir faktör olduğu varsayılır (Hambleton & Swaminathan, 1985). Dolayısıyla alt testlerin hız testleri olmadığı sonucuna ulaşılmıştır.

Model veri uyumunu değerlendirmek için, her bir alt testin tek boyut altında toplanan trafik ve çevre 50, motor ve araç tekniği 40 ve ilk yardım 30 maddeye ait üç model altında (1PL, 2PL ve 3PL) madde parametreleri kestirimi için XCalibre 4.2 programı kullanılmıştır. Model veri uyumunu sınamada -2loglikelihood istatistiği ve ki-kare istatistiğinden yararlanılmıştır. Yapılan analizler sonucunda veriyle en uyumlu modelin 2 parametrelili lojistik model olduğu bulunmuştur. Tablo 5'te iki parametrelili lojistik modele ve alt testlere göre uyumlu madde sayıları verilmiştir.

Tablo 5

İki Parametrelili Lojistik Modele Uyumlu Madde Sayısı

	2010	2011	Toplam
İlk yardım	29	29	58
Motor ve Araç Tekniği	39	39	78
Trafik ve Çevre	48	46	94
Toplam	116	114	230

Tablo 5 incelendiğinde; 2010 yılında 116, 2011 yılında 114; trafik ve çevre 94, motor ve araç tekniği 78 ve ilk yardım 58 olmak üzere madde havuzunda toplamda 230 madde olduğu görülmüştür. Analizlere 240 madde yerine madde havuzunda yer alan 230 madde ile devam edilmiştir.

Madde parametrelerinin değişmezliğinin incelenmesi için adayların alt testlerden aldıkları toplam puan kullanılarak, alt ve üst yetenek grupları oluşturulmuştur. En yüksek puan alan %27'lik grup ($n_1=1350$) aday üst grup olarak değerlendirilirken, en düşük puan alan %27'lik grup ($n_2=1350$) aday da alt grup olarak değerlendirilmiştir. Alt ve üst gruplardan elde edilen madde parametreleri (a ve b) arasındaki ilişki ise, Spearman rho korelasyon katsayısı ile hesaplanmıştır. Tablo 6'da alt ve üst gruptan elde edilen madde parametreleri arasındaki ilişkiler gösterilmiştir.

Tablo 6

Alt ve Üst Gruptan Kestirilen Madde Parametreleri Arasındaki Spearman Rho Korelasyon Katsayıları (n₁ = 1350; n₂ = 1350)

		Altgrup	Üst Grup			
			2010		2011	
			a	b	a	b
2010	İlk yardım	a	.85*	-		
		b	-	.35		
	Motor	a	.60*	-		
		b	-	.75*		
	Trafik	a	.58*	-		
		b	-	.46*		
2011	İlk yardım	a			.89*	-
		b			-	.42*
	Motor	a			.81*	-
		b			-	.39*
	Trafik	a			.75*	-
		b			-	.62*

(* $p < .05$)

Tablo 6 incelendiğinde, alt ve üst gruptan elde edilen a parametreleri $r=.58$ ile $r=.89$ arasında değişmekte olup orta ve yüksek düzeyde ilişki göstermektedir ($p<.05$). Madde güçlük parametreleri için ise, 2010 ilkyardım alt ve üst gruptan elde edilen değer ($r=.35$, $p>.05$) dışında, $r=.39$ ile $r=.75$ arasında değişmekte olup orta ve yüksek düzeyde ilişki göstermektedir ($p<.05$). Bu durumda 2010 ilkyardım alt ve üst gruptan elde edilen b parametresi dışında, a ve b parametreleri için değişmezlik sağlandığı söylenebilir.

Farklı madde örneklemelerinden yetenek parametrelerinin kestirilmesi için ise, trafik 50, motor 40 ve ilkyardım 30 maddeden oluşan 2010 ve 2011 alt testleri ilk yarı ve ikinci yarı olmak üzere, iki bölüme ayrılmıştır. Bireylerin ayrı ayrı ilk yarı ve ikinci yarı madde örneklemelerinden yetenek parametreleri kestirilmiştir. Her iki madde örnekleminde kestirilen yetenek parametreleri arasındaki ilişkinin saptanması için Pearson Momentler Çarpımı korelasyon katsayısı hesaplanmıştır. Tablo 7’de ilk yarı ve son yarı madde örneklemelerinden kestirilen yetenek düzeyleri arasındaki Pearson momentler çarpımı korelasyon katsayıları ($n=5000$) gösterilmiştir.

Tablo 7

İlk Yarı ve Son Yarı Madde Örneklerinden Kestirilen Yetenek Düzeyleri Arasındaki Pearson Momentler Çarpımı Korelasyon Katsayıları (N = 5000)

		Son Yarı					
		2010			2011		
İlk Yarı		ilkyardım	motor	trafik	ilkyardım	motor	trafik
2010	<i>İlk yardım</i>	.72**	-	-			
	<i>Motor</i>	-	.69**	-			
	<i>Trafik</i>	-	-	.73**			
2011	<i>İlk yardım</i>				.68**	-	-
	<i>Motor</i>				-	.71**	-
	<i>Trafik</i>				-	-	.72**

(** $p < .01$)

Tablo 7 incelendiğinde, ilk yarı ve son yarılar arasındaki korelasyon katsayıları $r = .68$ ile $r = .73$ arasında değişmekte olup orta ve yüksek düzeyde ilişki göstermektedir ($p < .01$). Dolayısıyla ilk yarı ve son yarılardan ayrı ayrı kestirilen yetenek düzeyleri arasında orta ve yüksek düzeyde manidar ilişki bulunduğu için, ilgili testlere ait farklı maddelerden elde edilen yetenek düzeylerinin değişmezlik özelliğine sahip olduğunu göstermektedir. Bir başka ifadeyle, yetenek kestiriminin madde örneklemeden bağımsız olduğu söylenebilir.

Kullanılan alt testlere ilişkin güvenilirliğin belirlenmesinde, iç tutarlılığının göstergesi olan Kuder-Richardson 20 (KR-20) güvenilirlik katsayıları hesaplanmıştır. KR-20 güvenilirlik katsayıları sırasıyla 2010 ilk yardım için .82, motor ve araç tekniği için .84, trafik ve çevre için .88; 2011 ilk yardım için .82, motor ve araç tekniği için .85 ve trafik ve çevre için ise .87 olarak hesaplanmıştır. Dolayısıyla her bir alt testin güvenilirlik düzeyinin yüksek olduğu söylenebilir.

MTSKS simülatif BOBUT uygulaması için 5000 simülatif bireyin, simülatif olarak oluşturulan 230 madde havuzundan minimum 5 madde, maksimum 50 madde yanıtlaması koşulu aranmıştır. Daha önce üç ayrı test toplamında 120 maddeden oluşan MTSAS 2013 yılında toplamda tek bir teste indirilmiş ve madde sayısı da 60'a düşürülmüştür. MTSAS, 2014 yılında yeniden düzenlenerek 50 maddeden oluşan (trafik ve çevre 27, ilk yardım 13, motor ve araç tekniği 10 madde) tek bir test olarak uygulanmaya başlanmıştır (MEB, 2014a). 2015 yılında MTSKS adıyla (trafik ve çevre 27, ilk yardım 13, araç tekniği 10 madde) tek bir test olarak uygulanmaktadır (MEB, 2014b). Dolayısıyla MTSKS uygulamalarında trafik, ilkyardım ve motor olmak üzere üç farklı alan bulunduğu için adayların her üç alandan da maddelerle karşılaşması amacıyla (trafik $27/50 = .54$; ilk yardım $13/50 = .26$ ve motor $10/50 = .20$) oranlarında içerik dengelenmesi (content balancing) kısıtlaması yapılmış ve en az 5 madde yanıtlaması koşuluyla adayların her bir alt testteki maddelerle karşılaşması sağlanmıştır.

Ayrıca güncel MTSKS uygulamalarında 50 madde kullanıldığı için maksimum 50 madde yanıtlanması koşulu aranmıştır.

MTSKS simülatif BOBUT uygulaması için başlangıç yetenek düzeyi $\theta=0$, yetenek kestirim yöntemi sonsal maksimum kestirim yöntemi (Maximum A Posteriori-MAP, sonlandırma kuralı olarak standart hatanın .30'un altında olması stratejisi kullanılmıştır. Maddelere tümüyle doğru ya da tümüyle yanlış yanıt veren bireyler bulunduğu için yetenek kestirim yöntemlerinden maksimum likelihood kestirim yöntemi yerine MAP tercih edilmiştir. Kezer (2014) de başlangıç yetenek düzeyi 0, sonlandırma kuralı $SE<.30$, MAP yetenek kestirim yöntemi stratejisi seçildiğinde kestirilen yetenek düzeyi ile kâğıt kalem testi arasında ($r=.95$) pozitif yönde yüksek bir korelasyon bulunduğunu ($p<.01$) ifade etmiştir.

MTSKS simülatif BOBUT uygulaması için CATSim programı kullanılmıştır. MTSKS simülatif BOBUT uygulaması ve kâğıt kalem testi uygulamasından kestirilen yetenek düzeyleri arasındaki ilişkiyi incelemek için ise, SPSS 20.0 versiyonunda Pearson momentler çarpımı korelasyon katsayısı hesaplanmıştır.

Bulgular

Bu bölümde araştırma sorularına yanıt aramak için 5000 simülatif bireyin, simülatif olarak oluşturulan 230 madde havuzundan minimum 5 madde, maksimum 50 madde yanıtlanması koşulunun yanı sıra başlangıç yetenek düzeyi $\theta=0$, yetenek kestirim yöntemi MAP, sonlandırma kuralı olarak standart hatanın .30'un altında olması durumunda MTSKS simülatif BOBUT uygulaması yapılarak analizler sonucunda elde edilen bulgulara aşağıda yer verilmiştir.

MTSKS Simülatif Bobut Uygulamasında Madde Sayısı Dağılımına İlişkin

Bulgular

MTSKS simülatif BOBUT uygulaması sonucunda bireylerin yanıtladıkları madde sayılarına ilişkin dağılım Tablo 8'de verilmiştir.

Tablo 8 incelendiğinde, adayların madde havuzunda bulunan 230 maddeden farklı sayıda maddelerle karşılaştıkları görülmektedir. Yapılan simülatif BOBUT uygulamasında yetenek kestirimi için;

- En az madde ile testi bitiren 375 aday (%7.50) 27 madde ile karşılaşmıştır.
- En çok madde ile testi bitiren 148 aday (%2.96) 50 maddeye yanıt vermiştir.
- Adaylar ortalama 32.52 madde yanıtlarken, kâğıt kalem testinde ise yetenek kestirimi için 50 madde yanıtlamaları gerekmektedir. Dolayısıyla yetenek kestirimi için, simülatif BOBUT uygulamasında adaylara uygulanan ortalama madde sayısı bakımından, MTSKS kâğıt kalem testindeki madde sayısına göre %34.96 oranında ekonomiklik sağlanmıştır.

MTSKS simülatif BOBUT uygulamasında, adayların en az 5 madde yanıtlamaları koşulu aranmıştır. Dolayısıyla ilk maddeden itibaren yeteneğin kestirildiği ve aynı sonlandırma kurallarının korunduğu bir durumda adayların karşılaştıkları madde sayılarının daha da düşeceği düşünülebilir.

Tablo 8

MTSKS Simülatif Bobut Uygulamasında Madde Sayısı Dağılımı

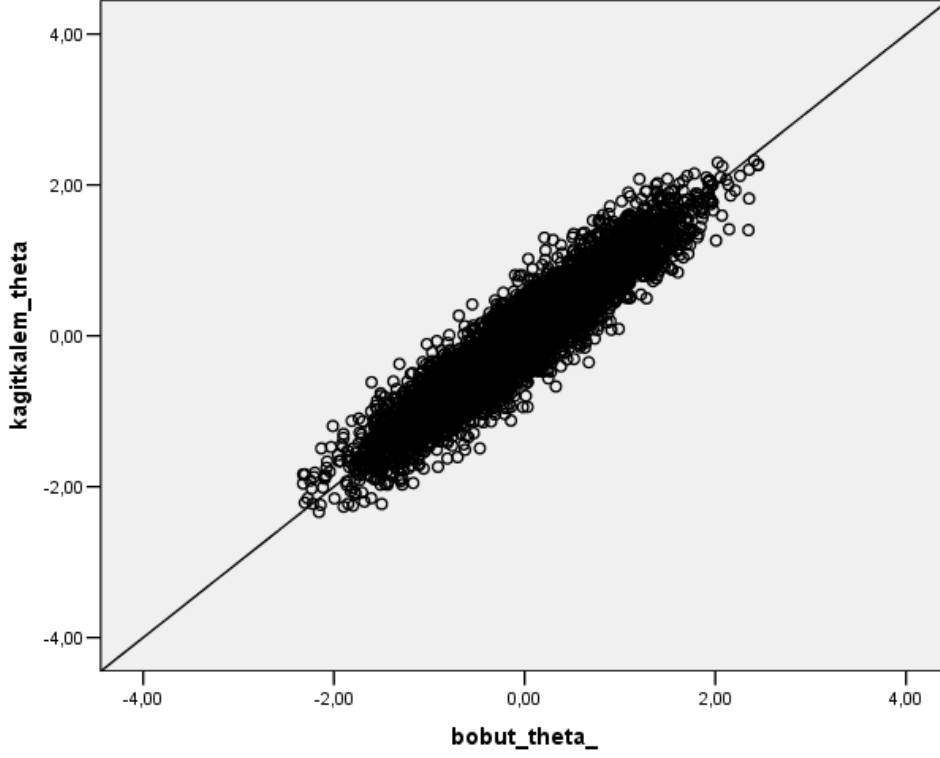
<i>Madde Sayısı</i>	<i>f</i>	Σf	<i>%</i>	$\Sigma\%$
27	375	375	7.50	7.50
28	1456	1831	29.12	36.62
29	482	2313	9.64	46.26
30	311	2624	6.22	52.48
31	287	2911	5.74	58.22
32	353	3264	7.06	65.28
33	183	3447	3.66	68.94
34	214	3661	4.28	73.22
35	155	3816	3.10	76.32
36	101	3917	2.02	78.34
37	124	4041	2.48	80.82
38	168	4209	3.36	84.18
39	97	4306	1.94	86.12
40	68	4374	1.36	87.48
41	80	4454	1.60	89.08
42	132	4586	2.64	91.72
43	35	4621	0.70	92.42
44	61	4682	1.22	93.64
45	38	4720	0.76	94.40
46	45	4765	0.90	95.30
47	40	4805	0.80	96.10
48	33	4838	0.66	96.76
49	14	4852	0.28	97.04
50	148	5000	2.96	100.00

MTSKS Simülatif BOBUT ile Kâğıt Kalem Testi Uygulamalarında Kestirilen Yetenek Düzeyleri Arasındaki Korelasyona İlişkin Bulgular

Adayların simülatif BOBUT uygulaması ve kâğıt kalem testi uygulaması ile kestirilen yetenek parametreleri arasındaki ilişki incelenmiştir. Kâğıt kalem testi uygulamasından elde edilen puanlar iki parametrelili lojistik modele göre kalibre edilmiştir. Kâğıt kalem test uygulaması sonucu adayların kestirilen yetenek parametrelerinin -3.50 ile 2.00 arasında değiştiği görülmektedir. Yetenek kestirimlerine ilişkin aritmetik ortalama -.02, standart sapma ise .94 olarak elde edilmiştir. Adaylara ilişkin kestirilen yetenek parametreleri -2.32 ile 2.45 arasında değişmektedir. Yetenek parametrelerine ilişkin aritmetik ortalama -.02, standart sapma .01 olarak saptanmıştır.

İki uygulamayı da alan adayların yetenek düzeylerine ilişkin saçılma diyagramı Şekil 1'de verilmiştir.

Şekil 1. Kâğıt Kalem Testi ve Simülatif Bobut Uygulamalarından Kestirilen Yetenek Düzeylerine İlişkin Saçılma Diyagramı



Kâğıt kalem testi ve simülatif BOBUT uygulamalarında kestirilen yetenek düzeyleri arasında $r=.94$ düzeyinde pozitif yönde yüksek düzeyde bir ilişki elde edilmiştir ($p<.01$). Kâğıt kalem testi ve simülatif BOBUT uygulamalarında kestirilen yetenek düzeyleri arasındaki manidar yüksek korelasyon, uygulamaların benzer yetenek kestiriminde bulunduğunu göstermektedir. Dolayısıyla adayların daha az (ortalama 32.52) madde yanıtladığı simülatif BOBUT uygulaması ile adayların 50 madde yanıtlamak zorunda oldukları kâğıt kalem testi uygulamasından elde edilen yetenek düzeyleri arasında pozitif yönde yüksek korelasyon katsayıları elde edilmesi uygulamaların benzer yetenek kestiriminde bulunduğunu göstermektedir.

Sonuç ve Tartışma

Bu araştırmada Türkiye'de ulusal düzeyde yapılan MTSKS e-sınav uygulamalarının BOBUT olarak yapılması durumunda nasıl bir sonuç alınabileceğini belirlemek amacıyla araştırma simülasyon çalışmasına dayalı olarak gerçekleştirilmiştir. Bu amaç doğrultusunda MTSKS'de yer alan maddelerin hitap ettiği yetenek düzeyleri ile adayların yetenek düzeylerinin örtüşmediği, maddelerin düşük yetenek düzeylerinde daha çok bilgi verdiği sonucuna ulaşılmıştır. Bunun sebebi olarak MEB tarafından yapılan sınavlarda, her bir yetenek düzeyine hitap edecek şekilde yeterli sayıda ve eşit olarak dağılım gösteren ve çok sayıda yüksek derecede ayırt edici maddelerden oluşan bir madde havuzundan maddelerin seçilmemesinden kaynaklandığı söylenebilir.

Dolayısıyla uç yetenek düzeylerindeki bireylerin yetenek düzeyleri geçerli ve güvenilir bir şekilde kestirilememektedir.

Klasik testler, genellikle orta düzeyde yeteneğe sahip olan bireyler için daha doğru sonuçlar verirken; BOBUT yeteneğin geniş bir yetenek aralığında doğru sonuçlar vermektedir (De Ayala 2009; Weiss, 1985). Ayrıca çok sayıda yüksek derecede ayırt ediciliğe sahip, güçlük-yetenek düzeyinde eşit biçimde temsil edilen maddelerden oluşması madde havuzunun etkililiğini göstermektedir (Veldkamp & Linden, 2010; Segall, 2003). Dolayısıyla bireysel farklılıkları gözetenek, bireye kendi yetenek düzeyine en uygun maddelerin uygulanması ile gerçekleştirilen ölçme yaklaşımı için geniş madde havuzuna ihtiyaç vardır (Magis & Raîche, 2012).

MTSKS kâğıt kalem testi ve simülatif BOBUT uygulamalarında kestirilen yetenek düzeyleri arasında pozitif yönde yüksek bir ilişki bulunmuştur. Bu sonuca göre, MTSKS'nin kâğıt kalem testi olarak uygulanması ile BOBUT olarak uygulanması arasında yetenek kestirimi açısından çok büyük farkın olmadığı sonucuna ulaşılmıştır. Literatürdeki BOBUT çalışmalarında, BOBUT yönteminin kâğıt kalem testleri yerine kullanılabilirliğini ortaya koymak adına, genellikle geleneksel kâğıt kalem testi ile BOBUT uygulamalarından elde edilen yetenek parametreleri arasındaki ilişkiye bakılmış ve yüksek korelasyonlar (.74 ile .98 arasında) elde edilmiştir (Bulut & Kan, 2012; Cömert, 2008; Evans, 2010; İşeri, 2002; Kalender, 2011; Kaptan, 1993; Kaskatı, 2011; Kezer, 2014).

MTSKS simülatif BOBUT uygulaması ve kâğıt kalem testiyle yetenek kestirimi için gerekli madde sayıları karşılaştırıldığında, simülatif BOBUT uygulamasında madde sayısında önemli ölçüde azalma sağlanmıştır. Bu sonuca göre, MTSKS'nin BOBUT olarak uygulanması ile daha az sayıda madde ile daha kısa sürede uygulanabileceği sonucuna ulaşılmıştır. Çünkü BOBUT avantajlarından biri, çok daha az maddeyle geleneksel testlerin ortaya koyduğu niteliklere göre, daha geçerli ve güvenilir yetenek kestirimlerine olanak sağlayan ölçmeler yapmasıdır. BOBUT uygulamalarında yeterli doğruluğa erişmek için mümkün olan en az sayıda maddeye gereksinim vardır. Böylelikle testuygulama süresi de kısalmaktadır. Ayrıca daha kısa uygulama süresi, bireyin test sonuçlarını etkileyebilecek bir faktör olan bıkkınlığı ve buna bağlı olası dikkat bozukluklarını azaltmaktadır (Embertson & Reise, 2000; Hambleton & Swaminathan, 1985; Hambleton, Swaminathan, & Rogers, 1991).

Öneriler

- EK D'deki 2010-2011 MTSAS alt testlerinin bilgi fonksiyonları ve alt testlere ilişkin ölçmenin koşullu standart hata fonksiyonları matrisleri incelendiğinde alt testlerin düşük yetenek düzeyine hitap ettiği görülmektedir. Dolayısıyla madde havuzunda genelde düşük yetenek düzeyine hitap eden maddelerin bulunduğu tespit edilmiştir (EK C Şekil 9 ve Şekil 10). Ayrıca EK C Şekil 11 simülatif madde havuzundaki 230 maddeye ait b parametreleri ve 5000 simülatif bireye ilişkin θ dağılımları incelendiğinde, maddelerin her bir yetenek düzeyine hitap edecek şekilde yeterli sayıda ve eşit olarak dağılım göstermediği söylenebilir. MTSKS uygulamalarında maddeler her bir yetenek düzeyine hitap edecek şekilde yeterli sayıda ve eşit olarak dağılım gösterir ve yüksek derecede ayırt edici çok sayıda madde içeren havuzla çalışıldığı takdirde BOBUT uygulamalarının daha iyi sonuçlar vereceği söylenebilir.

• MTSKS'nin BOBUT uygulanması ile adaylar kendi yetenek düzeylerinin çok altında ya da çok üzerinde maddelerle vakit harcamayacakları için zaman bakımından ekonomiklik sağlayacağı düşünülmektedir. Bu nedenle bu sınavın BOBUT olarak uygulanması önerilmektedir.

• MTSKS madde havuzunun yeniden gözden geçirilerek daha üst yetenek düzeylerinde yüksek bilgi veren maddelerin de havuza eklenmesi gerektiği düşünülmektedir. Bu sayede hem sınav süresi daha da azaltılabilir hem de yetenek kestirimi daha hatasız yapılabilir.

• BOBUT olarak yapılması düşünülen başka sınavlar uygulamaya başlamadan önce simülatif çalışmalar yapılabilir. Böylece test merkezleri karşılaşılabilecekleri zorlukları önceden tespit etme olanağı yakalayabilirler.

• Bu araştırmanın sınırlılığı olarak simülatif BOBUT uygulaması için başlangıç yetenek düzeyi $\theta=0$, yetenek kestirim yöntemi MAP, sonlandırma kuralı olarak standart hatanın .30'un altında olma stratejisi kullanılmıştır. Kâğıt-kalem testi ile farklı stratejiler kullanılarak yapılan simülatif BOBUT uygulamalarından elde edilen yetenek kestirimleri karşılaştırılabilir.

Summary

Purpose and Significance: In this study, it is aimed to compare simulatively administration of Motor Vehicle Driving Trainees Exam Test (MTSKS) as computerized adaptive test with paper-pencil test method. Accordingly, it is examined that distribution of item number in application of simulative computerized adaptive MTSKS test, whether there is a significant relationship between applications of theta parameters being estimated in simulative computerized adaptive MTSKS test adaptive and paper-pencil test or not. Adaptive test applications do not prevail until today especially in national exams in Turkey. It is thought that by composing large item pool computerized adaptive test should be imposed and be generalized in national exam. In literature, no study was encountered about application of MTSKS as a computerized adaptive test in Turkey. It is thought that this study contributes to fill the gap about this subject in literature.

In order to reach the purpose of this current study, the researchers have attempted to address the following research questions:

1. For estimating candidates' theta levels in application of computerized adaptive MTSKS test;
 - a) What is the minimum number of items applied?
 - b) What is the maximum number of items applied?
 - c) How much does simulative computerized adaptive MTSKS reduce the average number of items applied compared to the number of items in the paper-pencil test?
2. Is there a significant relationship between theta levels being estimated in application of MTSKS paper-pencil test and in application of simulative computerized adaptive MTSKS test?

Methods: This research, aiming to compare simulatively administration of MTSKS as computerized adaptive test with paper-pencil test method, is a basic research model. Because this study contributes to the development and expansion of existing theoretical knowledge as well as to the application (Karasar, 2012). Population of this study consists of candidates that attended the applications of MTSKS at October 23rd, 2010 and February 12nd, 2011. Data being used within the scope of study was formed by responses being given by randomly sampled 5000 candidates taking B type certificate T booklet in MTSKS being applied in 2010 and 2011. For the purpose of the study, simulated data was utilized for simulated BOBUT application. For the simulated BOBUT application, a 5000-person simulated data and a 230-item simulated item pool were used, which are similar to the distribution of the skill averages of 5000 people who took the MTSAS in 2010 and 2011, and the distribution of the parameters of 230 items that fit 2PL models. Once the item and individual parameters were predicted and the fitness of the data to the IRT (single dimensionality, local independence, speed test control, model data consistency, and invariance of item and ability parameters) was tested, the simulative BOBUT application was applied. XCalibre 4.2 and CATSim programs are used for data analysis.

Results: In the research it is found that 375 candidates (7.50%) completed the test with minimum number of items (27 items) while 148 candidates (2.96%) completed the test with maximum number of items (50 items). Candidates encountered average 32.52 items in simulative computerized adaptive test. When it is considered that in paper-pencil test each candidates encounters 50 items, it is found that candidates encounter less item at the rate of 34.96% in simulative computerized adaptive MTSKS test in compare with paper –pencil test application. Additionally it is found that high positive relationship at the level of $r=.94$ is obtained between theta levels being estimated in application of MTSKS paper-pencil test and in application of simulative computerized adaptive MTSKS test ($p<.01$). This result shows that these applications estimate similar theta levels. Thus, with simulated BOBUT practice, where candidates responded with fewer (average 32.52) items, achieving high correlation coefficients in the positive direction between the ability levels of the paper-pencil test that candidates have to answer 50 items suggests that similar skills are estimated.

Discussion and Conclusions: In research it is found that because theta levels being addressed by items in item pool do not correspond to theta levels of candidates, items give more information at lower theta levels. It is also seen that by application of computerized adaptive MTSKS test, test can be administered in a short time with less number of items. It is found that paper-pencil test and computerized adaptive test are highly related. This shows that MTSKS can be administered in the form of computerized adaptive test. According to this result, the application of MTSKS as adaptive test has achieved the result that it can be applied in a shorter time with fewer items. Because one of the advantages of adaptive test is that it makes more accurate measurements that allow superior, reliable ability estimates from the qualities set forth by traditional tests with fewer items. It is suggested that more items appealing to high theta levels should be added to item pool, by this way each candidate can encounter with items compatible with their theta levels.

Kaynakça

- Bulut, O. & Kan, A. (2012) Application of computerized adaptive testing to entrance examination for graduate studies in Turkey. *Eğitim Araştırmaları-Eurasian Journal of Educational Research*, 49, 61-80.
- Cömert, M. (2008). *Bireye uyarlanmış bilgisayar destekli ölçme ve değerlendirme yazılımı geliştirilmesi* (Yayınlanmamış yüksek lisans tezi). Bahçeşehir Üniversitesi, İstanbul.
- De Ayala, R. J. (2009). *The theory and practice of item response theory*. NY: The Guildford Press.
- Embretson, S.E. & Reise, S.P. (2000). *Item response theory for psychologists*. New Jersey: Lawrence Erlbaum Associates.
- Evans, J. J. (2010). *Comparability of examinee proficiency scores on computer adaptive tests using real and simulated data* (Unpublished doctoral dissertation). The State University of New Jersey, USA.
- Güler, K. (2010). *Türkiye'deki sürücü eğitimi ve sınav sistemlerinin diğer ülkelerle karşılaştırılması ve yeni bir model önerisi* (Yayınlanmamış yüksek lisans tezi). Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- Hambleton, R. K. & Swaminathan, H. (1985). *Item response theory: principles and applications*. USA: Kluwer Nijhoff Publishing.
- Hambleton, R. K., Swaminathan, H. & Rogers, H. J. (1991). *Fundamentals of item response theory*. California: Sage Publications Inc.
- İnal, K. (2014). *Türkiye'de trafik güvenliği eğitiminin bütünlüklü değerlendirilmesi* (Yayınlanmamış doktora tezi). Ankara Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- İşeri, A. I. (2002). *Assessment of students' mathematics achievement through computer adaptive testing procedures* (Unpublished doctoral dissertation). Middle East Technical University, Ankara.
- Kalender, İ. (2011). *Effects of different computerized adaptive testing strategies on recovery of ability* (Yayınlanmamış doktora tezi). Middle East Technical University, Ankara.
- Kaptan, F. (1993). *Yetenek kestiriminde adaptive (bireyselleştirilmiş) test uygulaması ile geleneksel kâğıt-kalem testi uygulamasının karşılaştırılması* (Yayınlanmamış doktora tezi). Hacettepe Üniversitesi, Ankara.
- Karasar, N. (2012). *Bilimsel araştırma yöntemleri*. Ankara: Nobel.
- Kaskatı, O. T. (2011). *Rasch modelleri kullanarak romatoid artirit hastaları özürüllük değerlendirmesi için bilgisayar uyarlamalı test yöntemi geliştirilmesi* (Yayınlanmamış doktora tezi). Ankara Üniversitesi, Ankara.
- Kezer, F. (2014). Bilgisayar ortamında bireye uyarlanmış test stratejilerinin karşılaştırılması. *Eğitim Bilimleri Araştırma Dergisi Uluslararası E-Dergi*, 4(1), 143-174.
- Leech, N. L., Barrett, K.C. & Morgan, G. A. (2005). *Spss for intermediate statistics: Use and interpretation*. London: Lawrence Erlbaum Associates Inc.

- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. New Jersey: N.J. Lawrence Erlbaum Associates Publishers.
- Magis, D. & Raîche, G. (2012). Random generation of response patterns under computerized adaptive testing with the R package catR. *Journal of Statistical Software*, 48(8), 1-31.
- MEB (2014a). Motorlu taşıt sürücü adayları sınavı e-kılavuzu. http://www.meb.gov.tr/sinavlar/dokumanlar/2014/kilavuz/mtsatsat_klvz_2014.pdf adresinden 29.10.2014 tarihinde alınmıştır.
- MEB (2014b). Motorlu taşıt sürücü kursiyerleri sınavı e-kılavuzu. http://www.meb.gov.tr/meb_sinavindex.php?KATEGORI=10-11-12-24&B=2 adresinden 15.12.2014 tarihinde alınmıştır.
- Özdemir, S. (2010). *Türkiye’de örgün eğitim sistemi içerisinde trafik eğitiminin durumu: Avrupa ve dünya ülkeleri ışığında geliştirilmesi* (Yayınlanmamış yüksek lisans tezi). Polis Akademisi Başkanlığı Güvenlik Bilimleri Enstitüsü, Ankara.
- Öztuna, D. (2008). *Kas-iskelet sistemi sorunlarının özür lülük değerlendirmesinde bilgisayar uyarlamalı test yönteminin uygulanması* (Yayınlanmamış doktora tezi). Ankara Üniversitesi, Ankara.
- Payam, M. M. (2012). *İlk ve ortaöğretim okullarında trafik eğitiminde temel sorunlar: Siirt ili örneği* (Yayınlanmamış doktora tezi). Polis Akademisi Başkanlığı Güvenlik Bilimleri Enstitüsü, Ankara.
- Russo, A. (2002). Mixing technology and testing: Computer-based assessments lend flexibility, quick turnaround and lower costs, supporters say. *School Administrator*.
- Segall, D. O. (2003). Computerized adaptive testing. *Encyclopedia of social measurement*. <http://www.academicpress.com/refer/measure/> adresinden 16.12.2014 tarihinde alınmıştır.
- Selbes, C. (2008). *Türkiye’de sürücü eğitimi müfredat programı ile sürücü belgesi alma prosedürünün Avrupa Birliği ülkeleri ile karşılaştırılması* (Yayınlanmamış yüksek lisans tezi). Gazi Üniversitesi Fen Bilimleri Enstitüsü, Ankara.
- Trotter, A. (2001). Testing computerized exam. *Education week*, 20(37), 30-35. <http://www.edweek.org/ew/ewstory.cmf?slug=37online.h20>, adresinden 18.12.2014 tarihinde alınmıştır.
- Veldkamp, B. P. & Linden, W. J. (2010). Designing item pools for adaptive testing. Linden W. J. & Glas C.A.W. (Eds.) *Elements of adaptive testing*. New York: Springer.
- Vursavaş, F. (2004). *Sürücü eğitim programının değerlendirilmesi* (Yayınlanmamış yüksek lisans tezi). Ankara Üniversitesi, Ankara.
- Web:<http://berlin.angloinfo.com/information/transport/driving-licences/the-driving-test> adresinden 18.12.2014 tarihinde alınmıştır.
- Web:https://www.expats-finland.com/living_in_finland/driving.html adresinden 18.12.2014 tarihinde alınmıştır.
- Web:<https://www.gov.uk/driving-theory-test> adresinden 18.12.2014 tarihinde alınmıştır.

Web:<http://italy.angloinfo.com/transport/driving-licences/getting-a-licence> adresinden 18.12.2014 tarihinde alınmıştır.

Web:http://www.rms.nsw.gov.au/documents/roads/licence/road_users_handbook-english.pdf adresinden 18.12.2014 tarihinde alınmıştır.

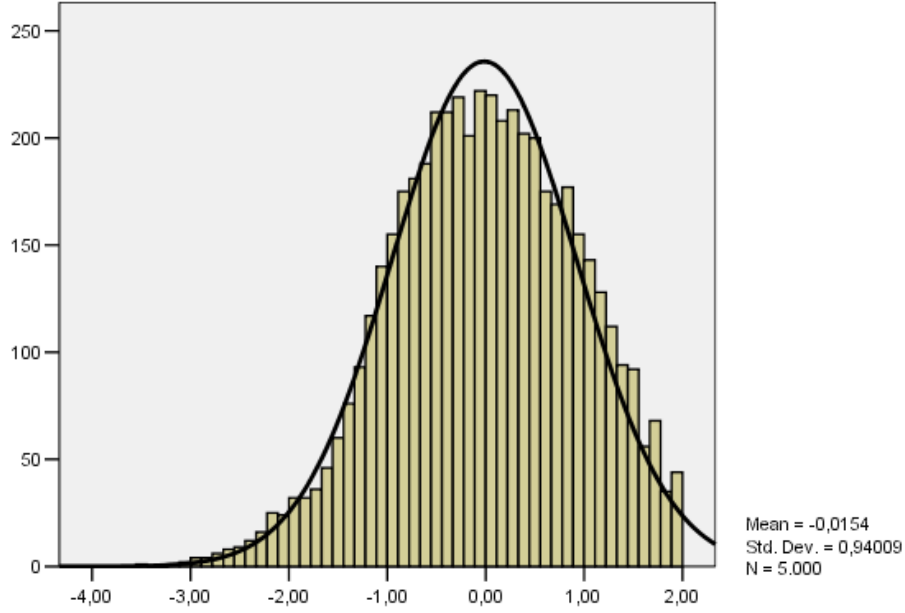
Weiss, D.J. (1985). Adaptive testing by computer. *Journal of Consulting and Clinical Psychology*, 53(6), 774-789.

EKLER

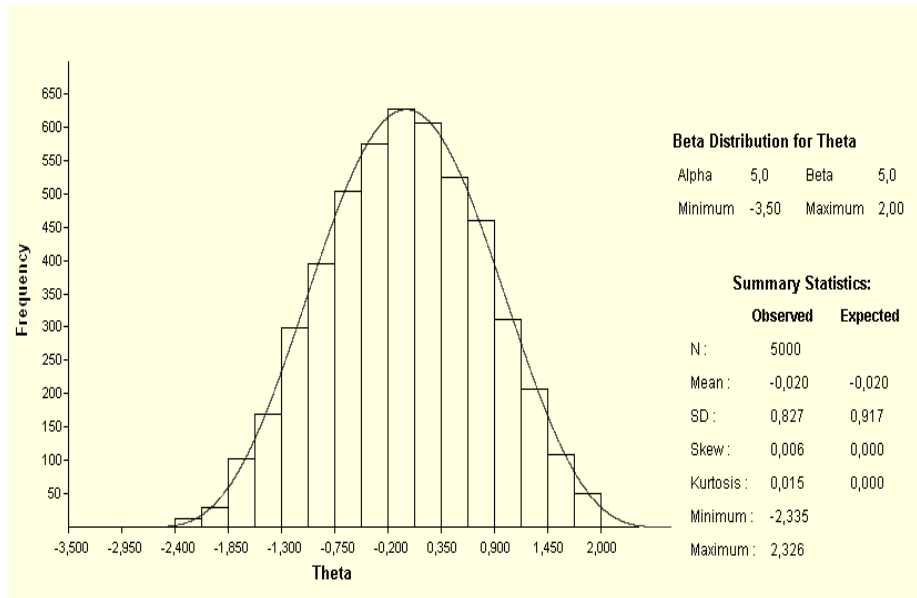
Ek A

Yetenek Düzeylerine İlişkin Dağılımlar

Şekil 2. 2010-2011 Motorlu Taşıtlar Sürücü Adayı Sınavını Alan 5000 Bireye İlişkin Yetenek Ortalamalarının Dağılımı



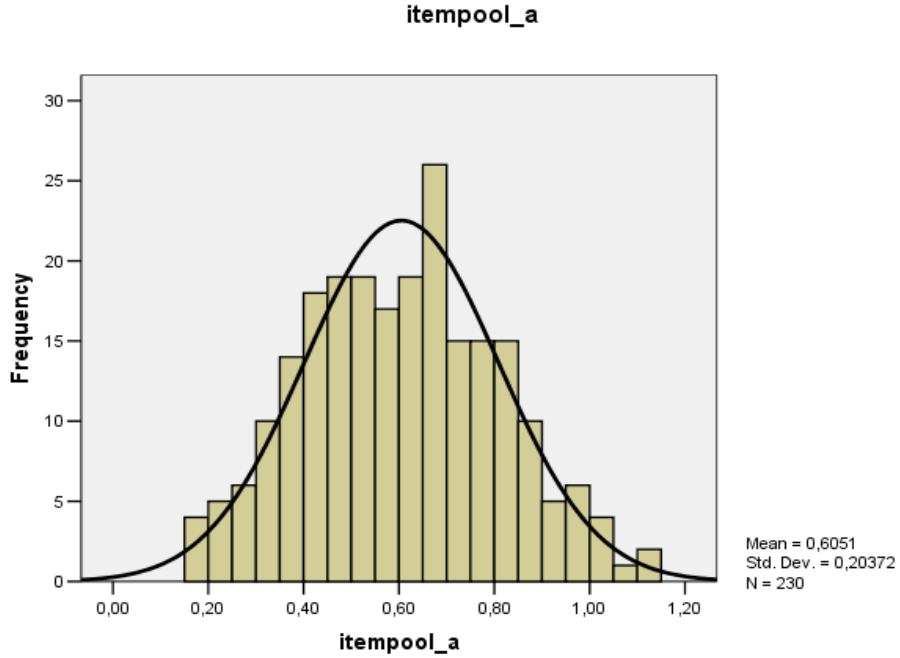
Şekil 3. Simülatif 5000 Bireye İlişkin Yetenek Dağılımı



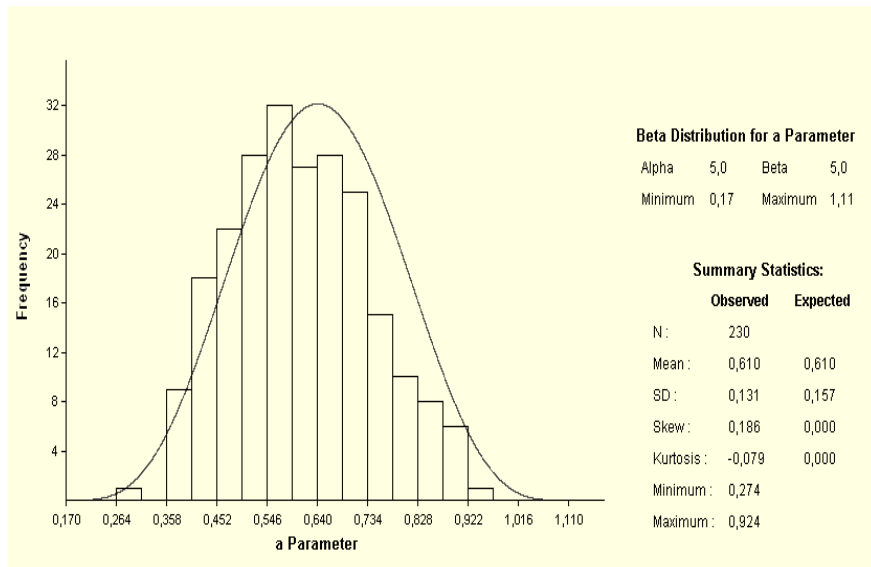
Ek B

Madde Parametrelerine İlişkin İlişkin Dağılımlar

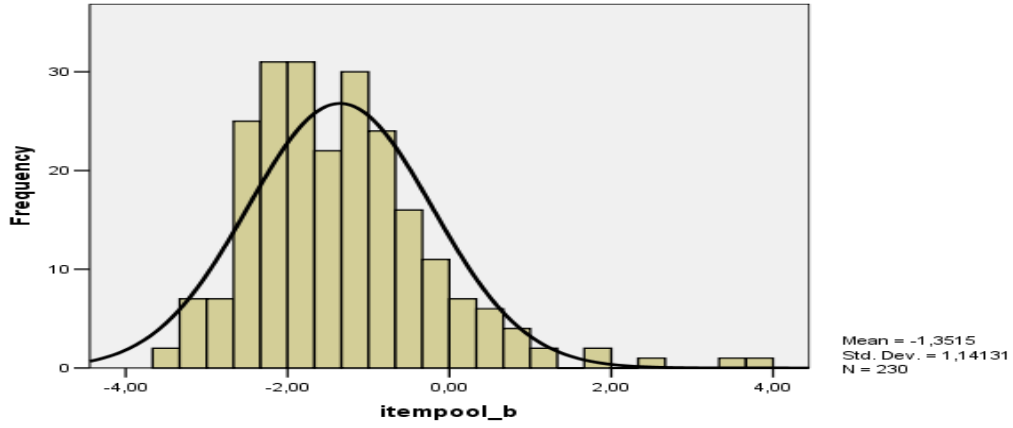
Şekil 4. 2010-2011 Motorlu Taşıtlar Sürücü Adayı Sınavındaki 2PL Modele Uyum Gösteren Maddelerle Oluşturulan Madde Havuzundaki 230 Maddeye Ait a Parametrelerine İlişkin Dağılım



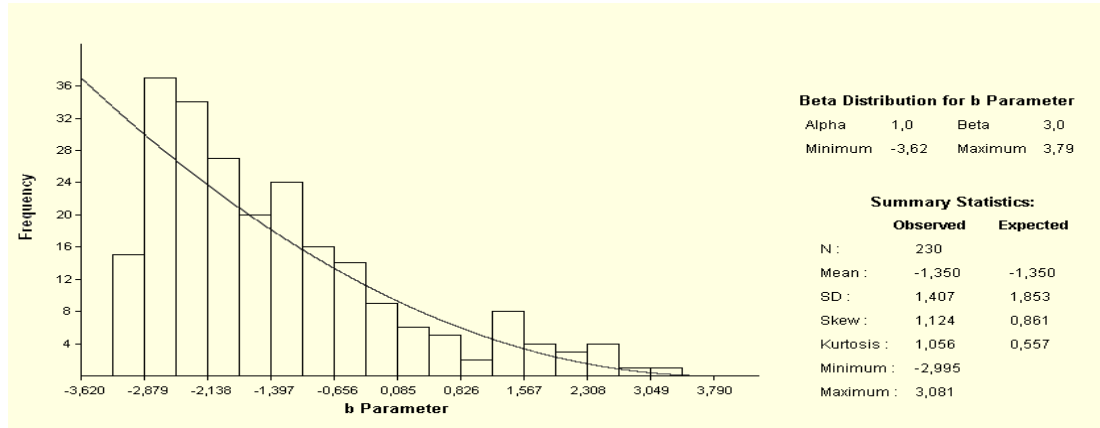
Şekil 5. Simülatif Madde Havuzundaki 230 Maddeye Ait a Parametrelerine İlişkin Dağılım



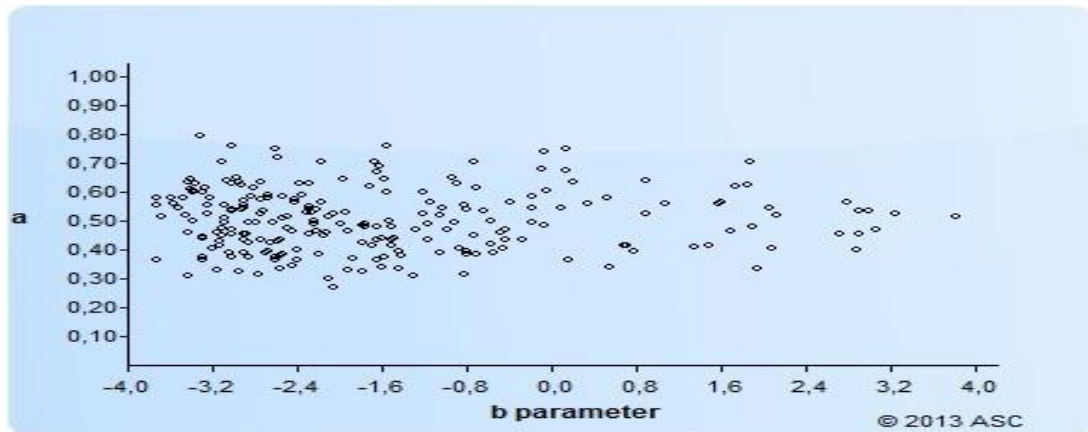
Şekil 6. 2010-2011 Motorlu Taşıtlar Sürücü Adayı Sınavındaki 2PL Modele Uyum Gösteren Maddelerle Oluşturulan Madde Havuzundaki 230 Maddeye Ait b Parametrelerine İlişkin Dağılım



Şekil 7. Simülatif Madde Havuzundaki 230 Maddeye Ait b Parametrelerine İlişkin Dağılım



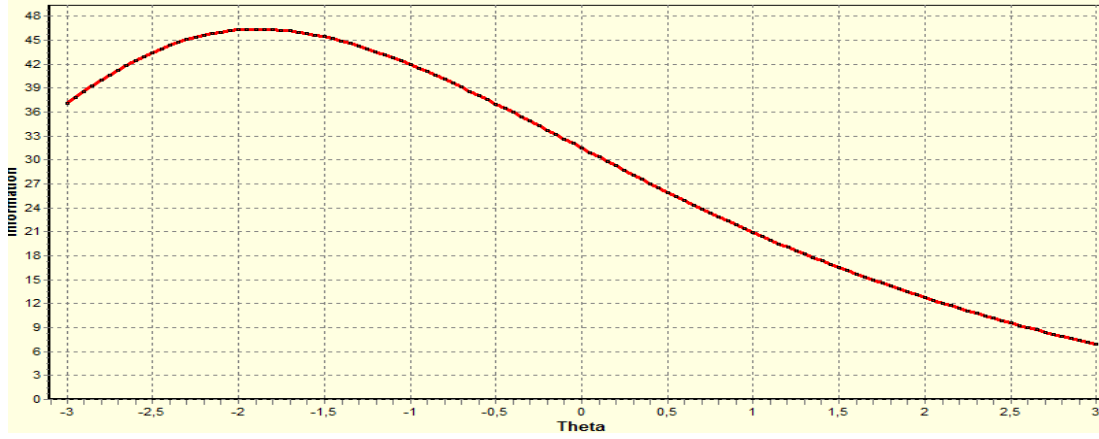
Şekil 8. Simülatif Madde Havuzundaki 230 Maddeye Ait a ve b Parametrelerine İlişkin Saçılma Diyagramı



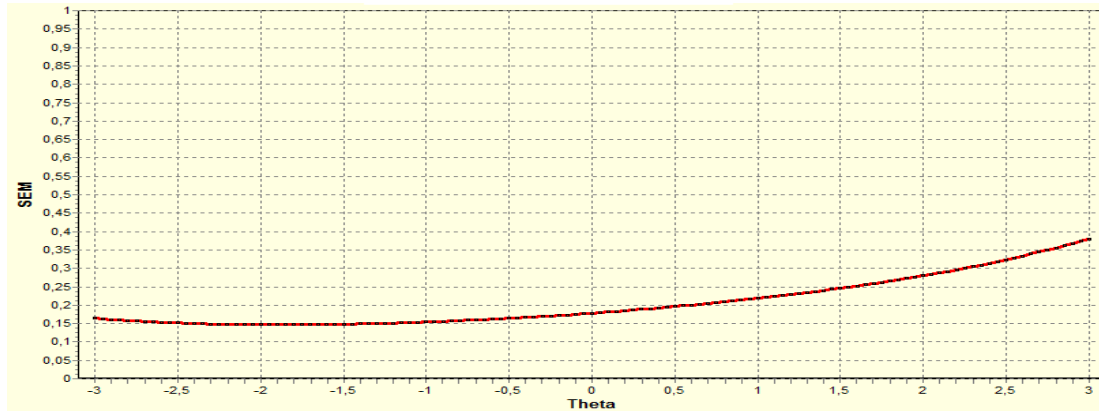
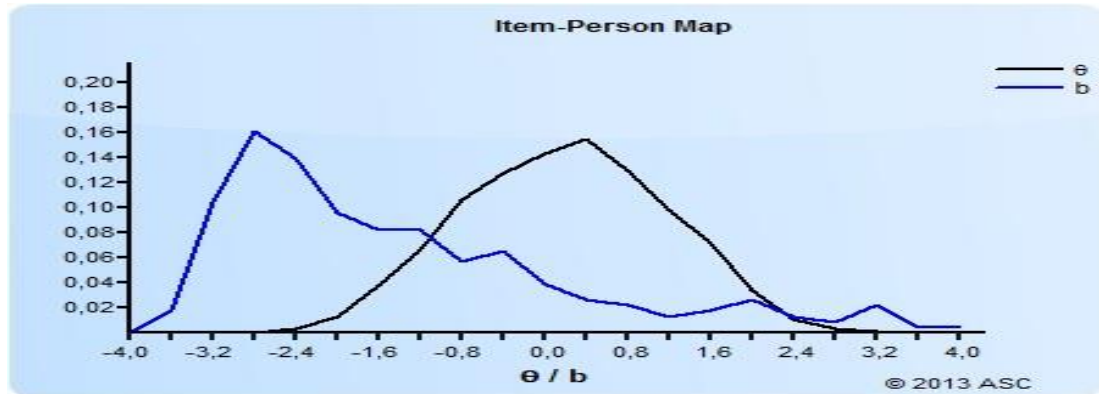
Ek C

Simülatif Madde Havuzuna İlişkin Bilgi Fonksiyonu ve Standart Hata Fonksiyonu ve θ/b Dağılımları

Şekil 9. Madde Havuzunun Bilgi Fonksiyonu



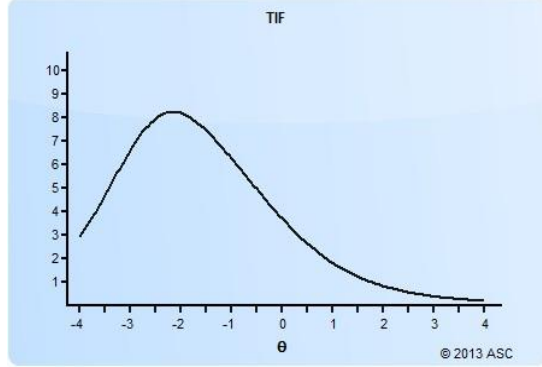
Şekil 10. Madde Havuzunun Standart Hata Fonksiyonu

Şekil 11. Simülatif Madde Havuzundaki 230 Maddeye Ait b Parametreleri ve 5000 Simülatif Bireye İlişkin θ Dağılımı

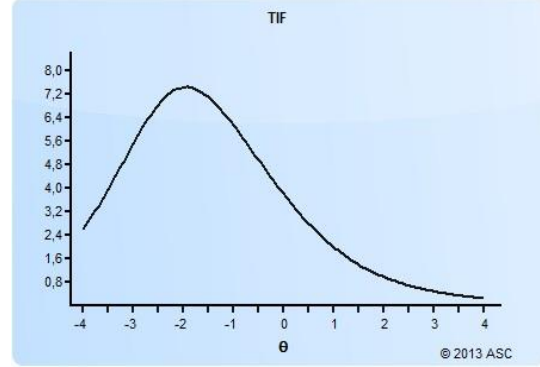
Ek D

2010-2011 MTSAS Alt Testlerinin Bilgi Fonksiyonları Ölçmenin Koşullu Standart Hata Fonksiyonları Matrisleri

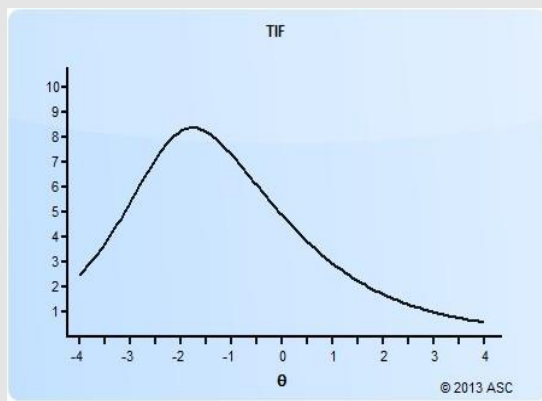
Şekil 12. 2010-2011 Mtsas Alt Testlerinin Bilgi Fonksiyonları Matrisi



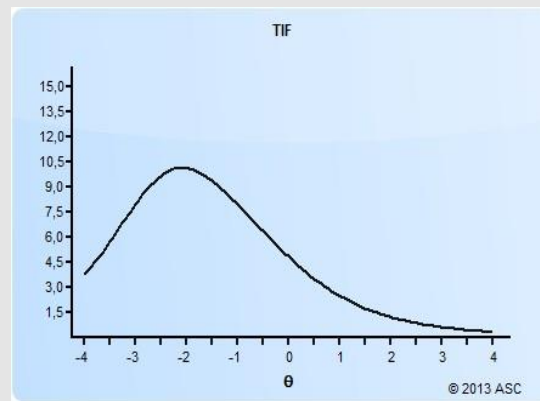
2010-İlkyardım Alt Testi
 $I_{max.}(\theta = -2.15) = 8.24$



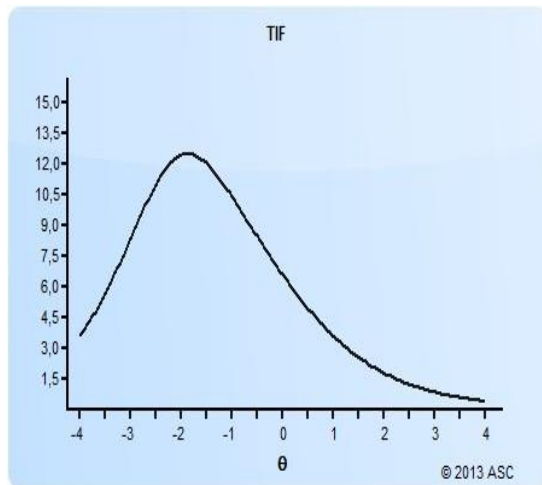
2011-İlkyardım Alt Testi
 $I_{max.}(\theta = -1.95) = 7.44$



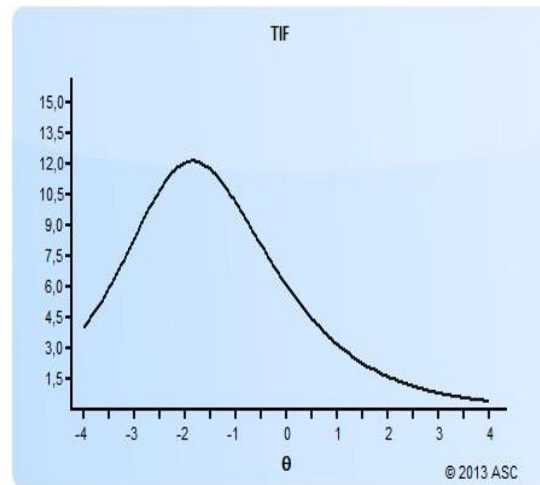
2010-Motor Alt Testi
 $I_{max.}(\theta = -1.75) = 8.38$



2011-Motor Alt Testi
 $I_{max.}(\theta = -2.10) = 10.14$

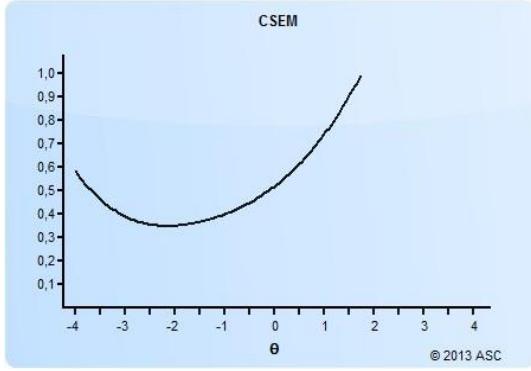


2010-Trafik Alt Testi
 $I_{max.}(\theta = -1.85) = 12.51$

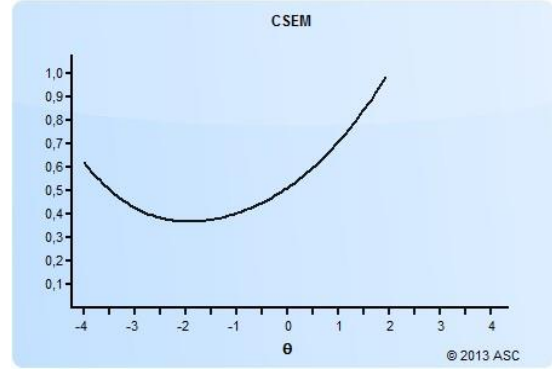


2011-Trafik Alt Testi
 $I_{max.}(\theta = -1.85) = 12.15$

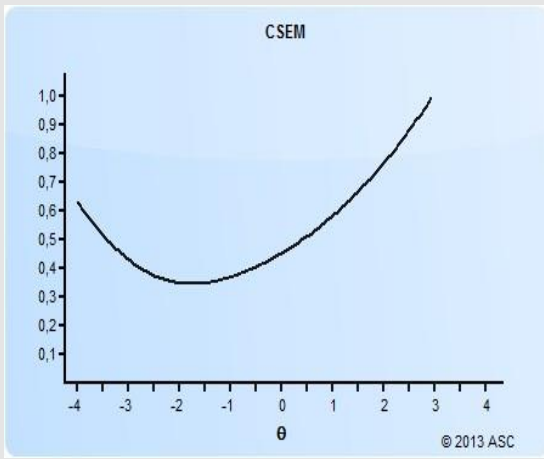
Şekil 13. 2010-2011 Mtsas Alt Testlerine İlişkin Ölçmenin Koşullu Standart Hata Fonksiyonları Matrisi



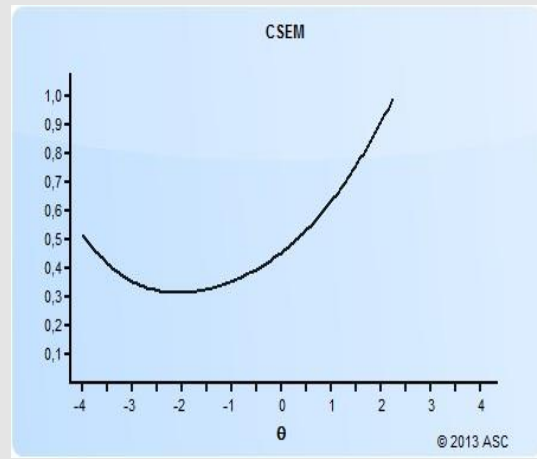
2010-İlkyardım Alt Testi
 $CSEM(\theta = -2.15) = .35$



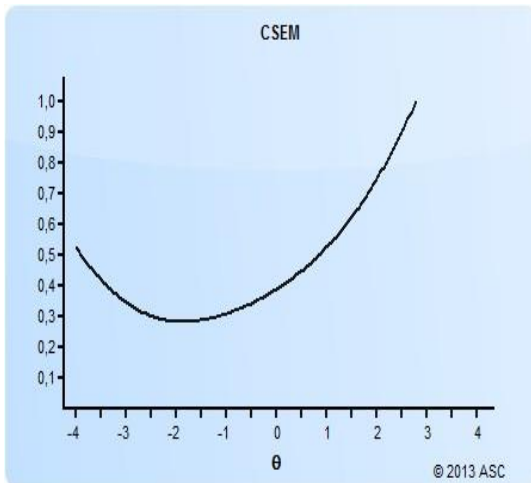
2011-İlkyardım Alt Testi
 $CSEM(\theta = -1.95) = .37$



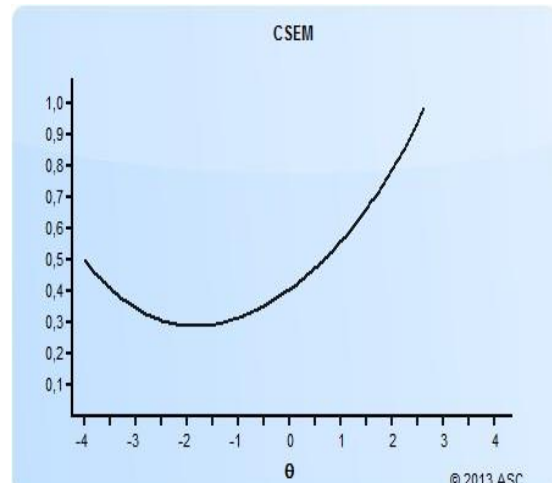
2010-Motor Alt Testi
 $CSEM(\theta = -1.75) = .35$



2011-Motor Alt Testi
 $CSEM(\theta = -2.10) = .31$



2010-Trafik Alt Testi
 $CSEM(\theta = -1.85) = .283$



2011-Trafik Alt Testi
 $CSEM(\theta = -1.85) = .287$



This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0). For further information, you can refer to <https://creativecommons.org/licenses/by-nc-sa/4.0/>