

**BİREYSELLEŐTİRİLMİŐ BİLGİSAYARLI TEST
UYGULAMALARINDA MADDE KULLANIM SIKLIĐI
KONTROL YÖNTEMLERİNİN İNCELENMESİ**

**INVESTIGATION OF ITEM EXPOSURE CONTROL METHODS
IN COMPUTERIZED ADAPTIVE TESTING**

Nagihan BOZTUNÇ ÖZTÜRK

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı İçin

Öngördüğü

Doktora Tezi


olarak hazırlanmıştır.

2014


Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼'ne,

Nagihan BOZTUNÇ ÖZT¼RK¼'n¼n hazırladıđı "Bireyselleřtirilmiř Bilgisayarlı Test Uygulamalarında Madde Kullanım Sıklıđı Kontrol Y¼ntemlerinin İncelenmesi" bařlıklı bu alıřma j¼rimiz tarafından **Eđitim Bilimleri Anabilim Dalı, Eđitimde Ölme ve Deđerlendirme Bilim Dalı'nda Doktora Tezi** olarak kabul edilmiřtir.

Bařkan


Prof. Dr. Selahattin GELBAL


¼ye (Danıřman)


Do. Dr. Nuri DOĐAN

¼ye


Prof. Dr. Mehtap AKAN

¼ye


Prof. Dr. H¼lya KELECİOĐLU

¼ye


Yrd. Do. Dr. Derya OBANOĐLU AKTAN

ONAY

Bu tez Hacettepe ¼niversitesi Lisans¼st¼ Eđitim-¼đretim ve Sınav Y¼netmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri ¼yeleri tarafından 29/01/2014 tarihinde uygun g¼r¼lm¼ř ve Enstit¼ Y¼netim Kurulunca/...../..... tarihinde kabul edilmiřtir.

Prof. Dr. Berrin AKMAN
Eđitim Bilimleri Enstit¼s¼ M¼d¼r¼

BİREYSELLEŞTİRİLMİŞ BİLGİSAYARLI TEST UYGULAMALARINDA MADDE KULLANIM SIKLIĞI KONTROL YÖNTEMLERİNİN İNCELENMESİ

Nagihan BOZTUNÇ ÖZTÜRK

ÖZ

Bireyselleştirilmiş bilgisayarlı test (BBT) uygulamalarının gün geçtikçe daha fazla tercih edilmesiyle birlikte testlerin geçerliğini tehlikeye düşürebilecek bazı sorunların da beraberinde geldiği görülmektedir. Uygulamalarda madde havuzunun çarpık kullanılması, sık kullanılan maddelerin kişiler tarafından hatırlanabilmesi, madde havuzu kullanımı sürekliliğinin sağlanamaması gibi durumların meydana gelmesi BBT sürecine yeni bir bileşenin eklenmesine sebep olmuştur. Bu nedenle test güvenliğini sağlayabilmek ve madde havuzunun kullanımını daha verimli hale getirebilmek için madde kullanım sıklığı kontrol yöntemleri geliştirilmiştir. Bu araştırmada da, çeşitli madde kullanım sıklığı kontrol yöntemlerinin, madde seçme yöntemi ve madde havuzu özelliklerine göre ölçme kesinliğine ve test güvenliğine olan etkisinin incelenmesi amaçlanmaktadır.

Araştırmada madde kullanım sıklığı kontrol yöntemleri olarak Bilgi Koşullu Tesadüfi (5 ve 10 madde), Sympon-Hetter ve Azalarak Kaybolma yöntemleri kullanılmıştır. Ayrıca madde kullanım sıklığı kontrolünün yapılmadığı koşul da referans oluşturması için araştırmaya dahil edilmiştir. Madde seçme yöntemlerinden ise Fisher'in En Çok Bilgi, a-Tabakalama ve Aşamalı En Çok Bilgi Oranı yöntemleri kullanılmıştır.

Simülasyon çalışması olarak yürütülen bu araştırmada 1000 bireye ait yetenek parametreleri ortalaması 0, standart sapması 1 olan normal dağılımdan elde edilmiştir. Madde havuzları oluşturulurken; her iki madde havuzunda da a parametresi [0.50;2.00], c parametresi [0.05;0.20] aralığında tek biçimli dağılıma sahip olacak şekilde türetilmiştir. b parametresi ise orta güçlükteki madde havuzunda [-3.00;+3.00] aralığında tek biçimli dağılımdan, yüksek güçlükteki madde havuzunda ise ortalaması 2, standart sapması 1.5 olan normal dağılımdan türetilmiştir. İlk madde seçiminde bireylerin yetenek düzeyleri sıfır olarak tanımlanmış, yetenek kestiriminde ise Beklenen Sonsal Dağılım kestirim yöntemi kullanılmıştır. Sonlandırma kuralı olarak

sabit test uzunluęu seilmiř ve bu uzunluk 25 olarak belirlenmiřtir. BBT uygulaması bir günde 1000 kiřinin uygulamaya katılması řeklinde simüle edilmiřtir.

Ölme kesinlięi iin uyum katsayısı, RMSE, yanlılık ve ortalama mutlak fark deęerleri hesaplanırken; test gvenlięi iin madde kullanım sıklıęı oranlarının daęılımı ve oranların ortalaması, standart sapması, en yksek madde kullanım sıklıęı oranı, ölklenmiř ki-kare istatistięi, F oranı ve test akıřması hesaplanmıřtır.

Fisher'in En ok Bilgi ve Ařamalı En ok Bilgi Oranı madde seme yntemlerinin kullanıldıęı kořullarda madde havuzu kullanımı aısından iki yntem yakın sonular vermiřtir. a-Tabakalama yntemi ise, madde kullanım sıklıęının kontrol edilmedięi durumda bile madde havuzu kullanımını dengelemiř ve en ok madde kullanım sıklıęı oranını, χ^2 ve test akıřma oranını azaltmıřtır. a-Tabakalama madde seme yntemi ve Azalarak Kaybolma madde kullanım sıklıęı kontrol yntemlerinin beraber kullanıldıęı durumlarda madde havuzunun kullanımındaki arpıklık ve test akıřması dięer kořullara gre daha dřk olmuřtur.

Arařtırmanın sonularına gre, madde kullanım sıklıęı kontrol yntemleri kullanıldıęında ölme kesinlięini belirlemede kullanılan gstergelere ait deęerlerde byk farkların bulunmadıęı tespit edilmiřtir. Test gvenlięine bakıldıęında ise genel olarak Azalarak Kaybolma madde kullanım sıklıęı kontrol ynteminin, madde havuzunun kullanımındaki arpıklıęın ve test akıřmasının azaltılması aısından dięer yntemlere gre daha iyi sonular verdięi grlmřtr.

Anahtar szckler: Bireyselleřtirilmiř bilgisayarlı test, madde kullanım sıklıęı kontrol yntemleri, bilgi kořullu tesadfi yntem, Sympson-Hetter yntemi, azalarak kaybolma yntemi, ölme kesinlięi, test gvenlięi

Danıřman: Do. Dr. Nuri DOęAN, Hacettepe niversitesi, Eęitim Bilimleri Anabilim Dalı, Eęitimde Ölme ve Deęerlendirme Bilim Dalı

INVESTIGATION OF ITEM EXPOSURE CONTROL METHODS IN COMPUTERIZED ADAPTIVE TESTING

Nagihan BOZTUNÇ ÖZTÜRK

ABSTRACT

It has been seen that some issues that could imperil the validity of the tests are arisen after the Computerized Adaptive Testing (CAT) administrations have been preferred recently. These issues such as skewness of the item pool utilization, the possibility of remembering the over-exposed items by examinees, not being able to progress the sustainability of item pool are caused to add new a component to CAT. Therefore item exposure control methods have been developed in order to ensure test security and to make the use of item pool more efficient. This study aims at investigating the various item exposure control methods in CAT under dichotomously IRT model by using different item pool characteristics and item selection methods in terms of measurement precision and test security.

The exposure control methods used in this study were the randomesque method (with 5 and 10 items), the Sympson-Hetter method and the Fade Away method. Each of these methods were compared with no-exposure control condition. Maximum Fisher Information, a-Stratification and Gradual Maximum Information Ratio methods were used as item selection methods.

This study was a simulation study. 1.000 simulees were generated from a standard normal $N(0,1)$ distribution. a parameters were generated from a uniform distribution with the range of $[0.50;2.00]$ and c parameters were generated from a uniform distribution with the range of $[0.05;0.20]$ both medium and hard item pools. b parameters were generated from a uniform distribution with the range of $[-3.00;+3.00]$ for medium item pool, and from a standard normal $N(2,1.5)$ distribution for hard item pool. The initial theta estimate was assigned zero for each simulee and Expected A Posteriori was used for ability estimation. The fixed test length rule was adopted to terminate the CAT and the length of the test was fixed at 25 items. The simulation was designed as 1000 simulees attended CAT in one day.

Fidelity, RMSE, bias and average absolute difference statistics were calculating for precision of measurement; moreover item exposure rates and these rates' means and standart deviations, maximum exposure rates, scaled chi-square statistic, F ratio and test overlap were also calculated for test security.

When Maximum Fisher Information and Gradual Maximum Information Ratio item selection methods were used, the results of these two methods was close with regard to item pool utilization under all conditions. On the other hand, a-stratification method balanced utilization of items and also reduced maximum item exposure rate, χ^2 and test overlap statistics in case no item exposure methods were used. In case of using a-stratification as item selection method and fade away method as item exposure method together, this combined method had the lowest skewness of item pools utilization and test overlap compared to all other conditions.

According to the results of this study, it has been verified that there was no difference in the value of indicators used to identify the precision of measurement when exposure control methods were utilized. Fade Away method as item exposure control method gave better results in regard to reducing the skewness of the item pool utilization and test overlap compared to others.

Keywords: Computerized adaptive testing, item exposure control methods, randomesque method, Sympson-Hetter method, fade away method, precision of measurement, test security

Advisor: Assoc. Prof. Dr. Nuri DOĞAN, Hacettepe University, Department of Educational Sciences, Division of Educational Measurement and Evaluation

ETİK BEYANNAMESİ

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü, tez yazım kurallarına uygun olarak hazırladığım bu tez çalışmada,

- tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu,
- başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- atıfta bulunduğum eserlerin tümünü kaynak olarak gösterdiğimi,
- kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- ve bu tezin herhangi bir bölümünü bu üniversitede veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı

beyan ederim.

Nagihan BOZTUNÇ ÖZTÜRK

TEŞEKKÜR

Doktora tezi çalışmam süresince; aklıma takılan her soruyu rahatça tartışabildiğim, yoğun olduğu anlarda dahi vakit ayırarak bir çözüm sunabilen, bilimsel desteğinin her zaman yanımda olduğunu hissettiğim değerli danışmanım Doç. Dr. Nuri DOĞAN'a; lisansüstü eğitimimin her aşamasında öğrenciden ziyade bir meslektaşı olduğumu hissettiren, düşüncelerine ve önerilerine çok değer verdiğim sevgili hocam Prof. Dr. Hülya KELECİOĞLU'na ve Prof. Dr. Selahattin GELBAL'a; görüş ve önerileriyle tez çalışmama katkı sağlayan Prof. Dr. Mehtap ÇAKAN'a ve Yrd. Doç. Dr. Derya ÇOBANOĞLU AKTAN'a;

Tezime getirdiği önerilerle katkısını eksik etmeyen canım dostum Dr. Melek Gülşah EROĞLU'na ve sadece akademik hayatta değil her zaman yanımda olduğunu hissettiğim canım dostum Dr. Gülşen TAŞDELEN TEKER'e; tezimi okuyup verdikleri öneriler için canım arkadaşlarım Dr. Kübra ATALAY KABASAKAL'a, Leyla YILMAZ FINDIK'a ve Suzan Beyza KAPTI'ya; Matlab programı ile veri analizinde yardımcı olan değerli arkadaşım Fatih Burak ŞAHİN'e;

Hayatımın her döneminde yanımda olan, sevgisini ve desteğini benden esirgemeyen, en önemlisi de bana her durumda güvenen anneme, babama ve kardeşlerime;

Her konuda olduğu gibi akademik çalışmalarım boyunca da desteğini hiçbir zaman esirgemeyen ve yeri geldiğinde bana destek olabilmek için kendinden ödün veren yol arkadaşım sevgili eşim Korhan ÖZTÜRK'e;

Doktora öğrenimim boyunca yurt içi doktora burs olanağı sağlayan TÜBİTAK'a;

teşekkürü bir borç bilirim.

İÇİNDEKİLER

ÖZ	iii
ABSTRACT	v
ETİK BEYANNAMESİ	vii
TEŞEKKÜR.....	viii
İÇİNDEKİLER.....	ix
TABLolar DİZİNİ	xii
ŞEKİLLER DİZİNİ	xiii
SİMGELER VE KISALTMALAR DİZİNİ	xiii
1. GİRİŞ	1
1.1. Problem Durumu	1
1.2. Araştırmanın Amacı ve Önemi.....	4
1.3. Problem Cümlesi	5
1.3.1. Alt Problemler	5
1.4. Sınırlılıklar.....	8
1.5. Araştırmanın Kuramsal Temeli	8
1.5.1. BBT'nin Temel Bileşenleri.....	10
1.5.1.1. Madde Havuzu	10
1.5.1.2. Madde Tepki Modeli.....	11
1.5.1.3. Test Başlangıç Yöntemi	12
1.5.1.4. Madde Seçme Yöntemi	13
1.5.1.5. Yetenek Kestirimi Yöntemi	14
1.5.1.6. Sonlandırma Kuralı	15
1.5.1.7. Kapsam Dengelenmesi	16
1.5.1.8. Madde Kullanım Sıklığı Kontrolü	17
1.5.2. Madde Kullanım Sıklığı Kontrol Yöntemleri	18
1.5.2.1. Tesadüfi Seçme(Randomization) Yöntemleri	19
1.5.2.2. Koşullu Seçme Yöntemleri (Conditional Selection Strategies)	19
1.5.2.3. Tabakalı Yöntemler (Stratified Strategies)	20
1.5.2.4. Birleştirilmiş Yöntemler (Combined Strategies)	20
1.5.2.5. Çok Aşamalı Bireyselleştirilmiş Test Desenleri (Multiple Stage Adaptive Test Designs)	21
2. İLGİLİ ARAŞTIRMALAR.....	22
2.1. İlgili Araştırmalar Özet	29
3. YÖNTEM.....	31
3.1. Araştırmanın Türü.....	31
3.2. Verilerin Türetilmesi.....	31
3.2.1. Örneklem Gruplarının Oluşturulması	31
3.2.2. Madde Havuzlarının Oluşturulması.....	32

3.3. BBT Koşulları.....	33
3.3.1. Madde Kullanım Sıklığı Kontrol Yöntemleri	34
3.3.1.1. Bilgi Koşullu Tesadüfi (BKT) Yöntemi (Randlesque Method)	34
3.3.1.2. Sympson-Hetter (SH) Yöntemi (Sympson-Hetter Method).....	34
3.3.1.3. Azalarak Kaybolma (AK) Yöntemi (Fade Away Method).....	36
3.3.2. Madde Seçme Yöntemleri	37
3.3.2.1. Fisher'ın En Çok Bilgi Yöntemi.....	37
3.3.2.2. a-Tabakalama Yöntemi	38
3.3.2.3. Aşamalı En Çok Bilgi Oranı (AEBO)Yöntemi	41
3.3.3. a-Tabakalama Yönteminin Belirlenmesi	42
3.3.4. Hedeflenen Kullanım Sıklıklarının Hesaplanması	43
3.3.4.1. Sympson-Hetter Yöntemi	43
3.3.4.2. Azalarak Kaybolma Yöntemi	43
3.4. Verilerin Analizi.....	43
4. BULGULAR ve TARTIŞMA	48
4.1. Birinci Alt Probleme İlişkin Bulgular	48
4.2. İkinci Alt Probleme İlişkin Bulgular	52
4.3. Üçüncü Alt Probleme İlişkin Bulgular	56
4.4. Dördüncü Alt Probleme İlişkin Bulgular.....	61
4.5. Beşinci Alt Probleme İlişkin Bulgular	65
4.6. Altıncı Alt Probleme İlişkin Bulgular	69
4.7. Tartışma Ve Yorum.....	74
5. SONUÇ ve ÖNERİLER	77
5.1. Sonuçlar	77
5.1.1. Alt Problemlere Ait Sonuçlar	77
5.1.1.1. Birinci Alt Probleme Ait Sonuçlar.....	77
5.1.1.2. İkinci Alt Probleme Ait Sonuçlar	78
5.1.1.3. Üçüncü Alt Probleme Ait Sonuçlar	78
5.1.1.4. Dördüncü Alt Probleme Ait Sonuçlar.....	79
5.1.1.5. Beşinci Alt Probleme Ait Sonuçlar.....	80
5.1.1.6. Altıncı Alt Probleme Ait Sonuçlar	81
5.1.2. Madde Havuzu Özellikleri Temelinde Sonuçlar	82
5.1.2.1. Orta Güçlük Düzeyinde Olan Madde Havuzuna Ait Sonuçlar	82
5.1.2.2. Yüksek Güçlük Düzeyinde Olan Madde Havuzuna Ait Sonuçlar.....	83
5.2. Öneriler.....	84
5.2.1. Uygulamaya Dönük Öneriler.....	84
5.2.2. Araştırmaya Dönük Öneriler	85
KAYNAKÇA.....	87
EKLER DİZİNİ	91
EK-1: Araştırmanın Yürütüldüğü Grubun Dağılım Grafikleri	92
EK-2: Madde Havuzlarındaki Maddelerin b Parametrelerinin a Parametrelerine Göre Dağılım Grafikleri	93
EK-3: BBT Uygulamalarındaki Madde Kullanım Sıklığı Oranı Grafikleri	94

EK-4: Madde Havuzu Güçlüğü Orta Düzeyde Olduğunda Madde Seçme Yöntemleri Bazında Madde Kullanım Sıklığı Kontrol Yöntemlerinin Ölçme Kesinliği Açısından Karşılaştırılma Tablosu	109
EK-5: Madde Havuzu Güçlüğü Orta Düzeyde Olduğunda Madde Seçme Yöntemleri Bazında Madde Kullanım Sıklığı Kontrol Yöntemlerinin Test Güvenliği Açısından Karşılaştırılma Tablosu	110
EK-6: Madde Havuzu Güçlüğü Yüksek Düzeyde Olduğunda Madde Seçme Yöntemleri Bazında Madde Kullanım Sıklığı Kontrol Yöntemlerinin Ölçme Kesinliği Açısından Karşılaştırılma Tablosu	111
EK-7: Madde Havuzu Güçlüğü Yüksek Düzeyde Olduğunda Madde Seçme Yöntemleri Bazında Madde Kullanım Sıklığı Kontrol Yöntemlerinin Test Güvenliği Açısından Karşılaştırılma Tablosu	112
ÖZGEÇMİŞ	113

TABLolar DİZİNİ

Tablo 1.1. Arařtırma Kapsamında Kullanılan Yöntemler ve Madde Havuzu Özellikleri.....	8
Tablo 1.2. BBT'de Madde Kullanım Sıklığı Kontrolü İçin Önerilen Bazı İlkeler....	11
Tablo 4.1. Birinci Alt Probleme Ait Ortalama Uyum, RMSE, Yanlılık ve OMF Deęerleri.....	48
Tablo 4.2. Birinci Alt Probleme Ait Test Güvenlięi İçin Kullanılan Göstergeler	49
Tablo 4.3. Birinci Alt Probleme Ait Ortalama F Oranları	51
Tablo 4.4. İkinci Alt Probleme Ait Ortalama Uyum, RMSE, Yanlılık ve OMF Deęerleri.....	52
Tablo 4.5. İkinci Alt Probleme Ait Test Güvenlięi İçin Kullanılan Göstergeler.....	54
Tablo 4.6. İkinci Alt Probleme Ait Ortalama F Oranları.....	55
Tablo 4.7. Üçüncü Alt Probleme Ait Ortalama Uyum, RMSE, Yanlılık ve OMF Deęerleri.....	57
Tablo 4.8. Üçüncü Alt Probleme Ait Test Güvenlięi İçin Kullanılan Göstergeler..	58
Tablo 4.9. Üçüncü Alt Probleme Ait Ortalama F Oranları.....	60
Tablo 4.10. Dördüncü Alt Probleme Ait Ortalama Uyum, RMSE, Yanlılık ve OMF Deęerleri.....	61
Tablo 4.11. Dördüncü Alt Probleme Ait Test Güvenlięi İçin Kullanılan Göstergeler	63
Tablo 4.12. Dördüncü Alt Probleme Ait Ortalama F Oranları	64
Tablo 4.13. Beřinci Alt Probleme Ait Ortalama Uyum, RMSE, Yanlılık ve OMF Deęerleri.....	66
Tablo 4.14. Beřinci Alt Probleme Ait Test Güvenlięi İçin Kullanılan Göstergeler ..	67
Tablo 4.15. Beřinci Alt Probleme Ait Ortalama F Oranları	69
Tablo 4.16. Altıncı Alt Probleme Ait Ortalama Uyum, RMSE, Yanlılık ve OMF Deęerleri	70
Tablo 4.17. Altıncı Alt Probleme Ait Test Güvenlięi İçin Kullanılan Göstergeler....	71
Tablo 4.18. Altıncı Alt Probleme Ait Ortalama F Oranları	73

ŞEKİLLER DİZİNİ

Şekil 1.1.	BBT Algoritmasının Örnek Akış Şeması.....	9
Şekil 3.1.	Orta Güçlük Düzeyindeki Madde Havuzunun Test Bilgi Fonksiyonu.	32
Şekil 3.2.	Yüksek Güçlük Düzeyindeki Madde Havuzunun Test Bilgi Fonksiyonu.....	33
Şekil 3.3.	a-Tabakalama Yönteminin Şekille Gösterilmesi	39

SİMGELER VE KISALTMALAR DİZİNİ

BBT: Bireyselleştirilmiş Bilgisayarlı Test

3 PLM: 3 Parametrelili Lojistik Model

FEB Yöntemi: Fisher'in En Çok Bilgi Yöntemi

ESD: En Çok Sonsal Dağılım

BSD: Beklenen Sonsal Dağılım

AEBO Yöntemi: Aşamalı En Çok Bilgi Oranı Yöntemi

KOD: Madde Kullanım Sıklığı Kontrolünün Yapılmadığı Durum

BKT Yöntemi: Bilgi Koşullu Tesadüfi Yöntemi

SH Yöntemi: Sympson-Hetter Yöntemi

AK Yöntemi: Azalarak Kaybolma Yöntemi

OMF: Ortalama Mutlak Fark

RMSE: Hataların Ortalama Karekökü

ko: Ortalama Madde Kullanım Sıklığı Oranı

1. GİRİŞ

Bu bölümde; araştırmanın temelini oluşturan problem durumu, araştırmanın amacı ve önemi, problem cümlesi, alt problemler, sınırlılıklar ve araştırmanın kuramsal temeli ile ilgili bilgiler yer almaktadır.

1.1. Problem Durumu

Bireyselleştirilmiş bilgisayarlı test (BBT) uygulamaları, bilgisayar ve psikometri alanındaki gelişmeler ile birlikte her geçen gün daha fazla tercih edilmektedir. Testlerin uygulanmasının ve puanlanmasının hızlı olması, testte bireyin yetenek düzeyine göre maddelerin yer alması, istenilen zamanda uygulanabilmesi, testlerin kısalması gibi özelliklerinden dolayı BBT geleneksel kağıt kalem testlerine göre daha cazip bir seçenek haline gelmektedir (Rudner, 1998).

GMAT (Graduate Management Admission Test), GRE (Graduate Record Examination), TOEFL (Test of English as a Foreign Language) gibi uygulamalarda da BBT kullanılmaktadır. Bireyler hakkında önemli kararların alınmasında kullanılan bu uygulamaların, geleneksel kağıt kalem testlerinde olduğu gibi, güvenilir ve geçerli olması gerekmektedir. BBT uygulamalarının yaygınlaşması ile birlikte geçerliği tehlikeye düşürebilecek bazı sorunların da beraberinde geldiği görülmektedir. Testlerin güvenliği (security) konusu tam bu noktada gündeme gelmektedir (Chang ve Twu, 1998; Georgiadou, Triantafillou ve Economide, 2007; Davey ve Nering, 2002, Akt. Barrada, Olea, Ponsoda ve Abad, 2009; Lee ve Dodd, 2012).

BBT uygulamalarında kullanılan madde havuzlarında çok sayıda madde yer almaktadır. Fakat bazı durumlarda kullanım sıklığının bazı maddelerde oldukça yüksek, bazı maddelerde ise düşük olduğu görülmektedir. Böyle durumlarla karşılaşıldığında bireylerin sık kullanılan maddeleri hatırlama olasılıklarının yüksek olduğu düşünülebilir. Eğer sık kullanılan maddeler bireyler tarafından hatırlanır ve paylaşılırsa testin geçerliği de tehlikeye girebilir (Georgiadou, Triantafillou ve

Economide, 2007; Lee ve Dodd, 2012). Örneğin bir birey katıldığı BBT uygulamasında aldığı maddeleri, bir süre sonra katıldığı BBT uygulamasında da tekrar alırsa bireyin elde ettiği puan gerçek bilgisinden mi kaynaklanıyor, yoksa maddeleri hatırlamasından mı kaynaklanıyor sorusunun cevabı bulunamaz. Ayrıca maddelerin hatırlanması madde havuzunda yer alan maddelerin farklı bireylerle paylaşılması gibi tehlikeleri de beraberinde getirmektedir. Bu durum test güvenliğini olumsuz yönde etkileyecektir.

İyi bir madde havuzunun geliştirilmesi uzun ve zahmetli bir süreç gerektirmektedir. Bu nedenle madde havuzunun verimli bir şekilde kullanılmayıp sadece belli bir yüzdesinin kullanılması test geliştiricilerinin istemediği bir durumdur (Revuelta ve Ponsoda, 1998). Belirtilen nedenlerle test güvenliğini sağlayabilmek ve madde havuzunun kullanımını daha verimli hale getirebilmek için bazı yöntemler geliştirilmiştir. Bu yöntemlere madde kullanım sıklığı kontrol yöntemleri denilmektedir. Madde kullanım sıklığı yöntemleri, uygulamada karşılaşılan sorunlar nedeniyle zamanla BBT'nin temel bileşenleri arasına girmektedir (Davis, 2002; Boyd, 2003).

Madde havuzunda yer alan maddelerin uygulamaya katılacak bireyler tarafından bilinmesi, tekrar uygulamaya girecek olan bireylerin aynı maddeleri alması, benzer yeterlik düzeylerindeki bireylere aynı maddelerin uygulanması, madde havuzunun düşük bir yüzdesinin kullanılması, madde havuzunda sık kullanılan maddelerin ifşa edilerek farklı uygulamalarda yer alması gibi olayların kontrol edilebilmesi için geliştirilen madde kullanım sıklığı kontrol yöntemlerine yenilerinin eklenmesi devam etmektedir.

Literatüre katılan her bir yeni yöneme ait sonuçların mevcut yöntemlerin sonuçlarıyla karşılaştırılması gerekmektedir. Madde seçme yöntemi, maddelerin psikometrik özellikleri, madde havuzu büyüklüğü, sonlandırma kuralı, örneklem dağılımı gibi BBT uygulamalarında değişebilen faktörlerde madde kullanım sıklığı yöntemlerinin nasıl sonuçlar vereceği araştırılması gereken bir problem olarak kendini göstermektedir.

BBT uygulamalarının ve madde kullanım sıklığı kontrolünün belki de en önemli noktasında madde havuzu yer almaktadır. Çünkü iyi bir uygulamanın yapılabilmesi madde havuzundaki maddelerin özelliklerine bağlıyken, madde kullanım sıklığını etkileyen faktörlerden biri de madde havuzunda bulunan diğer maddelerin psikometrik özellikleridir (Revuelta ve Ponsoda,1998). Psikometrik açıdan ise maddeler, sahip oldukları madde parametreleri ve yetenek düzeyine karşılık gelen madde bilgisine dayanarak değerlendirilirler. Yetenek ölçeği boyunca maddelerin istenen dağılımı testin amacına bağlı olarak değişkenlik gösterir. Başarı testleri için, madde havuzu genel olarak çok kolaydan çok zora doğru her türlü maddeyi içermesi gerekirken, b parametresinin ideal dağılımı ise tek biçimli dağılım (uniform distribution) olmalıdır. Belirli bir kesme noktasının altında ve üstünde yer alan bireyleri ayırt etmeyi amaçlayan ölçüt referanslı testler için ise madde havuzundaki maddelerin çoğu kesme noktası civarında en çok bilgiyi vermeyi sağlayacak madde güçlük değerlerine sahip olmalıdır (Boyd, 2003). Bu durumda farklı amaçlarla kullanılacak farklı psikometrik özelliklere sahip olan madde havuzlarında madde kullanım sıklığı kontrol yöntemlerinin nasıl sonuçlar vereceğinin araştırılması gerekmektedir.

Madde kullanım sıklığı kontrol yöntemlerinin çalışma prensiplerine bakıldığında, bu yöntemlerin madde seçme yöntemleriyle beraber çalıştığı ancak bu yöntemlere bir çeşit kısıtlamalar getirdiği görülmektedir. Örneğin, madde kullanım sıklığının kontrol edilmediği durumda bireye uygulanması için seçilen bir madde, madde kullanım sıklığı kontrolü yapıldığında, kullanılan kontrol yönteminin özelliğine göre belirli ölçütler ile kıyaslanıp, uygulanmayabilir. Sonrasında ise bireye uygulanması için farklı bir madde seçilir. Bu süreç, bireye uygulanacak maddenin belirlenmesine kadar devam eder. Kısacası, bir madde kullanım sıklığı kontrol yöntemi farklı madde seçme yöntemleri ile birlikte kullanıldığında, madde seçme yönteminin algoritmasına bağlı olarak, farklı madde havuzu kullanımı sonuçları verecektir (Han, 2012). Bu durum da BBT uygulamasındaki madde seçme yönteminin değişmesi madde kullanım sıklığını ne derece etkilediği sorusunu akla getirmektedir.

Madde kullanım sıklığının kontrol edilmesine odaklanırken, ölçme kesinliğinin de göz ardı edilmemesi gerekmektedir. Madde kullanım sıklığı kontrol yöntemleri, test güvenliği ile ölçme kesinliğini dengede tutmalıdır (Boyd, 2003; Boyd, Dodd ve Fitzpatrick, 2013). Bu nedenle madde kullanım sıklığı kontrol yöntemleri karşılaştırılırken hem elde edilen yetenek düzeyi kestirimlerinin kesinliği hem de madde havuzunun kullanımı birlikte ele alınmalıdır.

1.2. Araştırmanın Amacı ve Önemi

Ülkemizde BBT uygulamaları ile ilgili yapılan araştırmalar oldukça sınırlı olup yapılan araştırmalarda; geleneksel kağıt kalem testleri ile BBT uygulamalarından elde edilen puanlar arasında anlamlı farkın olup olmadığının incelendiği, ölçme kesinliği açısından farklı başlangıç kurallarının, madde seçme yöntemlerinin, yetenek kestirim yöntemlerinin ve test sonlandırma kurallarının karşılaştırıldığı görülmektedir (Kaptan, 1993; Kalender, 2011; Sulak, 2013; Eroğlu, 2013). Yurtdışında yapılan araştırmalarda ise daha çok BBT bileşenlerine ait yeni yöntemlerin geliştirilmesine odaklandığı görülmüştür. Madde seçme ve madde kullanım sıklığı kontrolü gibi konularda araştırmacılar tarafından yeni yöntemler önerilmektedir (Finkelman, Nering ve Roussos, 2009; Han, 2009; Han, 2012). Geliştirilen bu yöntemlerin; farklı puanlanan madde türlerinde, farklı psikometrik özelliklere sahip madde havuzlarında, farklı yetenek dağılımına sahip gruplarda test edilmesi, BBT uygulamalarındaki kullanımlarıyla ilgili bilgiler vermesi açısından önemli görülmektedir. Yeni yöntemlerin test edilmesiyle güvenilirliği ve geçerliği daha yüksek olan BBT uygulamalarının yapılması hedeflenmektedir.

BBT'nin temel bileşenleri arasında yer almaya başlayan madde kullanım sıklığı kontrolü, IACAT'e (International Association for Computerized Adaptive Testing) göre; BBT'nin güncel konuları arasında yer almaktadır (International Association for Computerized Adaptive Testing, 2013). Ölçme kesinliğinde önemli farklar oluşturmadan madde havuzu kullanımının sürdürülebilir olması madde kullanım sıklığının hedefleri arasında yer almaktadır. Madde havuzunun kullanımının düzenlenmesi, testler arasındaki madde çakışmasının azaltılması ve BBT

uygulamalarının daha geçerli olması için madde kullanım sıklığı kontrol yöntemlerinin geliştirilmesi devam etmektedir.

Gelişen teknolojinin sağladığı imkânlar ve zaman kavramının hayatımızdaki öneminin artması ile beraber BBT uygulamalarının avantajları göz önüne alındığında; ülkemizde de geleneksel kâğıt kalem uygulamalarının yerini BBT uygulamalarının alacağı günlerin çok da uzak olmadığı söylenebilir. Bu durum, ülkemizde BBT uygulamalarının bileşenleri ile ilgili çalışmaların yapılmasının gerekliliğini ortaya koymaktadır. Yapılan bu araştırmada çeşitli durumlar altında, farklı madde kullanım sıklığı kontrol yöntemlerinin, ölçme kesinliği ve testin güvenliği açısından nasıl sonuçlar verdiğinin incelenmesi amaçlanmıştır. Bu sayede 1-0 puanlanan maddelerin yer aldığı BBT uygulamalarında madde kullanım sıklığının kullanılıp kullanılmaması durumunda elde edilecek sonuçlar hakkında bilgi elde edilmiştir.

Farklı madde özelliklerine sahip madde havuzlarında, farklı madde seçme yöntemleri ile farklı madde kullanım sıklığı yöntemlerinin incelenmesi hem ulusal hem de uluslararası literatüre getireceği katkı açısından oldukça önemlidir. Ayrıca araştırmamızın madde kullanım sıklığı kontrol yöntemleri ile ilgili ülkemizde yapılan ilk araştırma olması da ayrı bir önem taşımaktadır.

1.3. Problem Cümlesi

Bireyselleştirilmiş bilgisayarlı test uygulamalarında madde kullanım sıklığı kontrol yöntemlerinin kullanımı; farklı madde seçme yöntemlerine ve farklı madde havuzu özelliklerine göre, ölçme kesinliğini ve testin güvenliğini nasıl etkilemektedir?

1.3.1. Alt Problemler

- 1) BBT uygulamasında madde seçme yöntemi Fisher'ın En Çok Bilgi Yöntemi (FEB) ve madde havuzu orta güçlük düzeyinde (*b parametrelerinin ortalaması 0.06, standart sapması 1.7*) olduğunda ölçme kesinliği ve test güvenliği;
 - a) Madde kullanım sıklığı kontrolü yapılmadığında ve,

- b) Bilgi Koşullu Tesadüfi-5 madde,
 - c) Bilgi Koşullu Tesadüfi-10 madde,
 - d) Sympson-Hetter,
 - e) Azalarak Kaybolma madde kullanım sıklığı kontrol yöntemleri kullanıldığında nasıl değişmektedir?
- 2) BBT uygulamasında madde seçme yöntemi Fisher'ın En Çok Bilgi Yöntemi (FEB) ve madde havuzu yüksek güçlük düzeyinde (*b parametrelerinin ortalaması 2.06, standart sapması 1.5*) olduğunda ölçme kesinliği ve test güvenliği;
- a) Madde kullanım sıklığı kontrolü yapılmadığında ve,
 - b) Bilgi Koşullu Tesadüfi-5 madde,
 - c) Bilgi Koşullu Tesadüfi-10 madde,
 - d) Sympson-Hetter,
 - e) Azalarak Kaybolma madde kullanım sıklığı kontrol yöntemleri kullanıldığında nasıl değişmektedir?
- 3) BBT uygulamasında madde seçme yöntemi a-Tabakalama Yöntemi ve madde havuzu orta güçlük düzeyinde (*b parametrelerinin ortalaması 0.06, standart sapması 1.7*) olduğunda ölçme kesinliği ve test güvenliği;
- a) Madde kullanım sıklığı kontrolü yapılmadığında ve,
 - b) Bilgi Koşullu Tesadüfi-5 madde,
 - c) Bilgi Koşullu Tesadüfi-10 madde,
 - d) Sympson-Hetter,
 - e) Azalarak Kaybolma madde kullanım sıklığı kontrol yöntemleri kullanıldığında nasıl değişmektedir?

- 4) BBT uygulamasında madde seçme yöntemi a-Tabakalama Yöntemi ve madde havuzu yüksek güçlük düzeyinde (*b parametrelerinin ortalaması 2.06, standart sapması 1.5*) olduğunda ölçme kesinliği ve test güvenliği;
- Madde kullanım sıklığı kontrolü yapılmadığında ve,
 - Bilgi Koşullu Tesadüfi-5 madde,
 - Bilgi Koşullu Tesadüfi-10 madde,
 - Sympson-Hetter,
 - Azalarak Kaybolma madde kullanım sıklığı kontrol yöntemleri kullanıldığında nasıl değişmektedir?
- 5) BBT uygulamasında madde seçme yöntemi Aşamalı En Çok Bilgi Oranı (AEBO) ve madde havuzu orta güçlük düzeyinde (*b parametrelerinin ortalaması 0.06, standart sapması 1.7*) olduğunda ölçme kesinliği ve test güvenliği;
- Madde kullanım sıklığı kontrolü yapılmadığında ve,
 - Bilgi Koşullu Tesadüfi-5 madde,
 - Bilgi Koşullu Tesadüfi-10 madde,
 - Sympson-Hetter,
 - Azalarak Kaybolma madde kullanım sıklığı kontrol yöntemlerine göre nasıl değişmektedir?
- 6) BBT uygulamasında madde seçme yöntemi Aşamalı En Çok Bilgi Oranı (AEBO) ve madde havuzu yüksek güçlük düzeyinde (*b parametrelerinin ortalaması 2.06, standart sapması 1.5*) olduğunda ölçme kesinliği ve test güvenliği;
- Madde kullanım sıklığı kontrolü yapılmadığında ve,
 - Bilgi Koşullu Tesadüfi-5 madde,
 - Bilgi Koşullu Tesadüfi-10 madde,

d) Sympson-Hetter,

e) Azalarak Kaybolma madde kullanım sıklığı kontrol yöntemleri kullanıldığında nasıl değişmektedir?

Problem durumunun çözümlenmesi için kullanılan madde kullanım sıklığı kontrol yöntemleri, madde seçme yöntemleri ve madde havuzlarına ait özellikler Tablo 1.1.'de özetlenmiştir.

Tablo 1.1. Araştırma Kapsamında Kullanılan Yöntemler ve Madde Havuzu Özellikleri

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>Madde Seçme Yöntemleri</i>	<i>Madde Havuzu Özellikleri</i>
<i>Madde Kullanım Sıklığı Kontrolünün Yapılmadığı Durum</i>	<i>Fisher'in En Çok Bilgi Yöntemi</i>	<i>Orta Güçlük Düzeyi</i>
<i>Bilgi Koşullu Tesadüfi-5 Madde</i>	<i>a-Tabakalama Yöntemi</i>	<i>Yüksek Güçlük Düzeyi</i>
<i>Bilgi Koşullu Tesadüfi-10 Madde</i>	<i>Aşamalı En Çok Bilgi Oranı Yöntemi</i>	
<i>Sympson-Hetter</i>		
<i>Azalarak Kaybolma</i>		

1.4. Sınırlılıklar

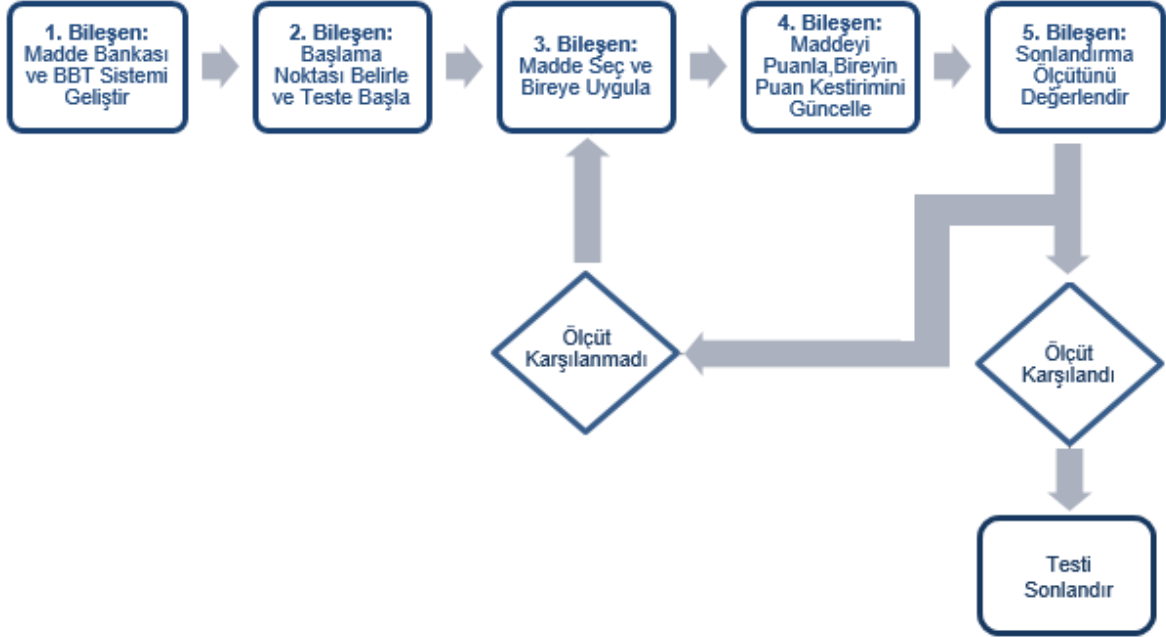
Araştırma,

1. 1-0 puanlama yöntemine dayanması,
2. 3 PLM'ye dayalı MTK modelinin kullanılması,
3. Beklenen Sonsal Dağılım (BSD) yetenek kestirim yönteminin kullanılması,
4. Test başlangıcı için yetenek düzeyinin 0 seçilmesi,
5. Sonlandırma kuralının sabit test uzunluğu (25 madde) seçilmesi ile sınırlıdır.

1.5. Araştırmanın Kuramsal Temeli

BBT'nin temel amacı, her bir birey için bireyselleştirilmiş testler oluşturarak, geleneksel kağıt kalem testlerine göre bireylerin yetenek düzeylerinin daha yüksek kesinlikle kestirebilmektir. Bu sebeple, geniş bir madde havuzundan bireyin mevcut yetenek düzeyine uygun olan maddeler seçilir (Chang, 2004). Bireyin yetenek düzeyine uygun olarak seçilen madde, birey tarafından doğru cevaplanırsa bir

sonraki maddenin güçlük düzeyi yükselir. Eğer madde yanlış cevaplanırsa bir sonraki maddenin güçlük düzeyi düşer. Bu süreç bireyin yetenek kestirimi sonlandırılana kadar devam eder. Thompson ve Weiss (2011), bu süreci Şekil 1.1.'deki gibi özetlemişlerdir.



Şekil 1.1. BBT Algoritmasının Örnek Akış Şeması

Şekil 1.1.'de gösterildiği gibi, bir BBT uygulaması ilk iki bileşeni gerçekleştirilmiş olarak başlar, daha sonra üçüncü, dördüncü ve beşinci bileşenler uygulanır. Bir birey BBT uygulamasına başladığında, bilgisayara madde bankası önceden yüklenmiş (1.Bileşen) ve birey için belirli bir başlama noktası tespit edilmiş olmalıdır (2.Bileşen). Başlangıç noktası için ilk madde seçilir. Madde cevaplandıktan sonra puanlanır ve bireyin yetenek düzeyi kestirilir. Daha sonra sonlandırma kuralı değerlendirilir. Eğer sonlandırma kuralı karşılanmamışsa madde havuzundan yeni bir madde seçilir (3.Bileşen), birey yeni maddeyi cevaplar, daha sonra yeteneği güncellenir (4.Bileşen) ve sonlandırma ölçütü tekrar değerlendirilir (5.Bileşen) (Thompson ve Weiss, 2011). BBT uygulaması için belirlenen sonlandırma ölçütü karşılanana kadar bu süreç devam eder ve ölçüt karşılandığında test sonlandırılır.

1.5.1. BBT'nin Temel Bileşenleri

BBT'nin temel bileşenleri, araştırmacılar tarafından farklı sayıda başlık altında toplanabilmektedir. Kingsbury ve Zara (1989); BBT'nin temel bileşenlerini altı başlık altında toplamıştır. Bunlar; (a) madde havuzu, (b) madde cevap modeli, (c) test başlangıç yöntemi, (d) madde seçme yöntemi, (e) yetenek kestirimi yöntemi ve (f) sonlandırma kuralıdır. Fakat BBT programlarının pratikte kullanılmasında karşılaşılabilecek durumlardan dolayı, iki bileşen daha ilave edilmektedir. Bu bileşenler kapsam dengelenmesi (content balancing) ve madde kullanım sıklığı kontrolü (item exposure control) bileşenleridir (Davis, 2002; Boyd, 2003).

1.5.1.1. Madde Havuzu

BBT uygulaması geniş bir madde topluluğundan, bireyin yetenek düzeyine uygun olan maddelerin seçilmesiyle oluşur. Bu geniş madde topluluğu madde havuzu olarak adlandırılır.

Madde havuzunun kalitesinin iyi olması, bireyselleştirme algoritmasının da iyi çalışması anlamına gelmektedir. Eğer bir madde havuzu sınırlı sayıda ya da düşük kaliteli maddelerden oluşursa, en iyi ve en gelişmiş bireyselleştirme programı bile işlevini yerine getiremez. Bireyselleştirilmiş testlerin kağıt kalem testlerine göre avantajlarını elde edebilmek için madde havuzunun farklı yetenek düzeylerine uygun yüksek nitelikte maddeler içermesi gerekmektedir (Flaughner, 2000). Reckase (2010) mümkün olan en iyi madde havuzunu, madde seçme algoritması tarafından seçilen maddenin uygulanabildiği havuz olarak tanımlamıştır.

Bireylerin yetenek düzeylerine uygun testlerin hazırlanabilmesi için yeterli sayıda maddeye sahip iyi tasarlanmış madde havuzlarına ihtiyaç duyulmaktadır (Gu ve Reckase, 2007). Way (1998), madde kullanım sıklığı ve test çakışması (test overlap) indekslerinin de madde havuzunun büyüklüğünün belirlenmesinde önemli olduğunu belirtmiş ve madde havuzu büyüklüğü ile ilgili yapılan çalışmaları inceleyip aşağıdaki tabloyu oluşturmuştur.

Tablo 1.2. BBT’de Madde Kullanım Sıklığı Kontrolü İçin Önerilen Bazı İlkeler

<i>Sınav Türü</i>	<i>Madde Havuzunun Sınav Uzunluğuna Oranı</i>	<i>Ortalama Madde Kullanım Sıklığı Oranı</i>	<i>Ortalama Çakışma Yüzdesi (Genel)</i>	<i>Ortalama Çakışma Yüzdesi (En Çok Koşullu)</i>
Seçme ve Yerleştirme Sınavları*	12	0.08-0.12	%10-15	%30
Lisans/Sertifika	6-8	0.10-0.15	%15-20	%40

* Seçme ve yerleştirme (high stake) sınavları için ilkeler sabit uzunluklu BBT’ye göre verilmiştir.

Kaynak: Way, W. D.. (2008). Protecting the Integrity of Computerized Testing Item Pools. *Educational Measurement: Issues and Practice*, 17-27

1.5.1.2. Madde Tepki Modeli

BBT’nin genel mantığında, bireylerin yetenek düzeylerine uygun olan maddelerin uygulanması yer almaktadır. Madde ve yetenek parametrelerini aynı ölçekte değerlendirebilmek ve bireylere en uygun maddeyi seçebilmek için çoğu BBT uygulamasında Madde Tepki Kuramı’na (MTK) dayalı kestirimler yapılmaktadır. MTK test geliştirme, madde analizi ve bireylerin puanlanması için birçok avantajı ile güçlü bir psikometrik paradigmadır (Thompson ve Weiss, 2011).

BBT uygulamalarında en uygun MTK modelinin 3 parametrelili lojistik model (3PLM) olduğu belirtilmektedir (Green, Bock, Humphreys, Linn ve Reckase, 1984, Lord, 1980, Weiss, 1983, Akt. Hambleton, Swaminathan ve Rogers, 1991).

3 PLM’nin matematiksel formülü aşağıdaki gibidir ($i = 1, 2, \dots, n$);

$$P_i(\theta) = c_i + (1 - c_i) \frac{e^{Da_i(\theta - b_i)}}{1 + e^{Da_i(\theta - b_i)}}$$

Formülde;

$P_i(\theta) = \theta$ yetenek düzeyindeki bir bireyin i maddesini doğru cevaplama olasılığını,

b_i = madde güçlük parametresini,

a_i = madde ayırt edicilik parametresini,

$D = 1.7$ (ölçekleme faktörü) değerini temsil etmektedir.

c_i parametresi ise madde karakteristik eğrisinde düşük yetenekli bireylerin bir maddeyi doğru cevaplama olasılığını (madde karakteristik eğrisinin düşük asimptotunu) göstermektedir. c_i parametresi, tahmin etmenin de test performansına bir faktör olarak dahil edilmesini sağlamıştır (Hambleton ve Swaminathan, 1985).

BBT uygulamalarında 1-0 puanlanan maddelerin yanı sıra çoklu puanlanan maddelerin (Davis, 2002; Pastor, Dodd ve Chang, 2002; Burt, Davis ve Dodd, 2003; Davis ve Dodd, 2005; Macken-Ruiz, 2008; Lee ve Dodd, 2012) ve ortak köklü maddelerin (testlet) (Boyd, 2003; Boyd, Dodd ve Fitzpatrick, 2013) yer aldığı madde havuzlarıyla da araştırmalar yapılmaktadır. Çoklu puanlanan maddelerin yer aldığı araştırmalarda çoklu puanlanan MTK modellerinden Kısmi Bilgi ve Genelleştirilmiş Bilgi Puan Modellerinin; ortak köklü maddelerin yer aldığı araştırmalarda 3 Parametrelili Lojistik Ortak Köklü Madde Tepki Modelinin (3-parameter logistic testlet response theory model) ve Kısmi Bilgi modellerinin kullanıldığı görülmektedir.

1.5.1.3. Test Başlangıç Yöntemi

Bir BBT uygulamasının ilk maddesinin belirlenmesinde çok çeşitli başlangıç noktaları seçilebilir. Ancak ortalama düzeye karşılık gelen sabit bir değerin alınması daha basittir. MTK kullanıldığında bu değer genellikle 0.0'dır. Bireylerin aynı θ başlangıç düzeyi ile teste başlamaları, tüm bireylerin aynı maddeyi ilk madde olarak alacağı anlamına gelmektedir. Bu durumu ortadan kaldırmak için başlangıç maddesi -0.5 ile 0.5 değerleri arasından seçilebilir (Thompson ve Weiss, 2011).

BBT uygulamalarının amacı, testi her birey için olabildiğince bireyselleştirmektir. Bireyler hakkında ön bilginin olması, bireyselleştirmenin daha kolay yapılmasını sağlayacaktır. Örneğin bir eğitim süreci içinde öğrencilere birden fazla BBT uygulaması yapılıyorsa, ilk uygulamanın sonucu sonraki uygulama için ideal bir başlangıç noktası olarak kullanılabilir. Diğer bir seçenek ise birey hakkında çeşitli bilgilerin toplanmasıdır. Örneğin; bir eğitim süreci sonucunda sertifika verilecekse;

yapılacak test ile ilişkisi olduğu gösterilen, eğitim sürecinde elde edilmiş olan akademik not ortalaması (grade-point average), BBT uygulamasında bir başlangıç değeri olarak kullanılabilir (Thompson ve Weiss, 2011).

1.5.1.4. Madde Seçme Yöntemi

MTK temelli BBT uygulamalarının madde seçme sürecinde; bir sonraki madde, bireyin geçici yetenek düzeyine uygun olarak seçilmektedir. Bir BBT uygulamasının başlangıcında, bir ya da iki madde uygulandıktan sonra, bir sonraki madde, adım kuralı (step rule) tarafından seçilir. Örneğin; eğer ilk madde doğru cevaplanmışsa, bireyin asıl önsel yetenek düzeyi ($\hat{\theta}$) bir miktar (örneğin 0.50) yükseltilir; eğer ilk madde yanlış cevaplanmışsa, asıl yetenek düzeyi ($\hat{\theta}$) bir miktar düşürülür. Her bir maddenin uygulanması ve puanlanmasının ardından, bir sonraki maddenin seçilmesi için yeni yetenek düzeyi ($\hat{\theta}$) kullanılır. Madde havuzundan seçim yapılırken maddenin bireye önceden uygulanmamış olmasına dikkat edilir (Weiss, 2004). Madde seçme yöntemleri zamanla değişik kategoriler altında sınıflandırılmıştır.

Kingsbury ve Zara (1989), madde seçme yöntemlerini önceden yapılandırılmış madde seçme yöntemleri (prestructured item selection procedures) ve değişken adımsal madde seçme yöntemleri (variable stepsize procedures) olarak iki kategoriye ayırmıştır. Bu iki kategori arasındaki en önemli fark sürecin bireyselleştirilmesine izin verme miktarıdır. Önceden yapılandırılmış madde seçme yöntemlerinde, madde seçme algoritmasını tanımlamak için önceden tanımlanmış madde havuzları bulunmaktadır. Değişken adımsal madde seçme yöntemleri ise test boyunca birbirini etkileyen çeşitli dallanmalar ile bireylere sınırsız yollar sunmaktadır.

Weiss ve Kingsbury (1984), BBT uygulamasında iki adet etkili madde seçme yöntemi olduğunu belirtmiştir. Bunlar, En Çok Bilgi ve Bayes madde seçme yöntemleridir. En Çok Bilgi madde seçme yöntemi, bireyin son yetenek düzeyinde en çok bilgi veren maddenin seçilmesini sağlarken, Bayes madde seçme yöntemi

ise yetenek kestiriminin beklenen sonsal dağılımının varyansını en küçük yapan maddenin seçilmesini amaçlamaktadır.

Van der Linden ve Pashley (2000) ise madde seçme yöntemlerini, klasik madde seçme yöntemleri ve modern seçme yöntemleri olarak iki başlık altında toplamıştır. Buna göre; En Çok Bilgi Ölçütü madde seçme yöntemi ve Owen'ın Bayes madde seçme yöntemi klasik madde seçme yöntemleri arasında yer alırken, En Çok Global Bilgiyi Verme Ölçütü (Maximum Global-Information Criterion), Olabilirlikle Ağırlıklandırılmış Bilgi Ölçütü (Likelihood-Weighted Information Criterion), Tam Bayes Ölçütü (Fully Bayesian Criteria), Önsel Bir Gözlem ile Bayes Ölçütü (Bayesian Criteria with an Empirical Prior) ve Tesadüfi Madde Parametreleri ile Bayes Ölçütü (Bayesian Criteria with Random Item Parameter) yöntemleri modern seçme yöntemleri arasında yer almaktadır.

1.5.1.5. Yetenek Kestirimi Yöntemi

BBT uygulamalarında yetenek kestirim süreci üç farklı aşamaya ayrılabilir. Bunlar; (1) madde seçme sürecine başlamak, (2) test sürecinde bireyin yeteneğine uygun madde seçimini sağlamak ve (3) test sonunda bireyin puanını belirlemek için hesaplanan yetenek kestirimleridir. Bu üç aşamanın da kendine özgü gereklilikleri ve sınırlılıkları bulunmaktadır (van der Linden ve Pashley, 2000).

Bireylerin yetenek kestiriminin yapıldığı bu üç aşamada kullanılacak çeşitli yetenek kestirimi yöntemleri bulunmaktadır. En Çok Olabilirlik Yöntemi (Maximum Likelihood, EÇO), En Çok Sonsal Dağılım (Maximum A Posteriori, ESD) ve Beklenen Sonsal Dağılım (Expected A Posteriori, BSD) yöntemleri üç temel yetenek kestirim yöntemidir.

En Çok Olabilirlik yöntemi, bireyin madde cevap örüntüsüne ait en yüksek olabilirlik (likelihood) değerini veren θ 'nın araştırılmasına dayanan bir yöntemdir. Olası tüm θ değerleri için olabilirlikler hesaplanır ve en yüksek olabilirlik değerini veren θ seçilir. En Çok Olabilirlik yönteminin olumlu yanları arasında yanlı olmaması ve hata değerlerinin normal dağılması yer alır. Ancak tüm sorular doğru ya da yanlış

cevaplandığında kestirim yapamaması, uç durumların (aberrance) varlığında ve kısa test uzunluklarında hatalı kestirimlerin olabilmesi En Çok Olabilirlik yönteminin sınırlılıklarındandır (Embretson ve Reise, 2000).

Bireylerin yetenek düzeylerinin kestiriminde önsel dağılımlar kullanıldığında bu yöntem, En Çok Sonsal Dağılım yöntemi adını alır. Bir Bayes kestirim yöntemi olan ESD; sonsal dağılımı en büyük yapmaya dayanan yetenek düzeyi kestirimi elde etmek için parametre bilgisi ve gözlenen log-olabilirlik fonksiyonunu beraber kullanan tekrarlı bir yöntemdir. Yöntemin üstün yanları arasında tüm soruları doğru ya da yanlış cevaplayan kişilerin yetenek düzeylerini kestirebilmesi ve önsel dağılıma ait bilgiyi yetenek kestirimi işlemine dâhil ederek yetenek kestiriminin hassasiyetini arttırması yer alır. Ancak bu yöntem özellikle madde sayısının az olduğu testlerde yanlış kestirim yapar. Yetenek kestiriminde kullanılan önsel dağılımlar yanlış olursa, yetenek kestirimi ciddi şekilde yanlış ve yanıltıcı olabilir. Testte yer alan maddelerin ayırt edicilikleri düşük ise önsel dağılım daha önemli hale gelir ve önsel dağılımın bireylerin yetenek kestirimindeki etkisi artar (Embretson ve Reise, 2000).

En Çok Olabilirlik ve ESD yöntemlerinin aksine BSD yöntemi tekrarlı olmayan yetenek kestirim sürecine sahiptir. BSD, sonsal dağılımın ortalamasından faydalanarak türetilmiş bir Bayes kestirimidir. BSD kestirimi ile elde edilen yetenek kestirimi ve standart hata değerleri ESD ile benzerdir. Aralarındaki tek fark, ESD yönteminin sonsal dağılımın modunu, BSD yönteminin ise sonsal dağılımın ortalamasını kullanmasıdır. BSD'nin ESD'ye göre avantajı tekrarlı bir sürece sahip olmamasıdır. Bu sayede yöntem daha hızlı hesaplama yapmaktadır. En Çok Olabilirlik yönteminin aksine, BSD, tüm soruları doğru ya da yanlış cevaplasa bile tüm bireylerin tüm cevap örüntüleri için bir yetenek kestirimi yapabilmektedir (Embretson ve Reise, 2000).

1.5.1.6. Sonlandırma Kuralı

BBT uygulamalarının sonlandırılmasında yaygın olarak kullanılan iki kural vardır. Bunlar sabit uzunluk ve değişen uzunluk sonlandırma kurallarıdır. Sabit uzunluk

test sonlandırma kuralında, her bireye aynı sayıda maddeden oluşan testlerin uygulanması gerekmektedir. Sabit uzunluk testlerin bir sonucu olarak, bireyler arasındaki ölçme kesinliği (precision of measurement) değişmektedir. Değişen uzunluk test sonlandırma kuralında ise; hedeflenen kesinlik düzeyi sağlanana kadar maddelerin uygulanması devam etmektedir. Sonlandırma kuralının seçimi; testin amacına, madde havuzunun özelliğine ve işlevsel kısıtlamalara bağlıdır (Segall, 2005).

Testlerin amaçlarına bakıldığında; bazı testlerin seçme ya da sınıflama, bazı testlerin ise her bir bireyi olabildiğince iyi ölçme amacı taşıdıkları görülmektedir. Bu bağlamda BBT'de iki farklı sonlandırma kuralı kullanılır.

Sınıflama amaçlı testlerde bir bireyin puanı belirli bir kesme değeri ile karşılaştırılır. Bu işlemdeki amaç bireyleri olabildiğince doğru ayırmaktır. Bu bağlamda sonlandırma kuralı olarak hem yetenek düzeyi ($\hat{\theta}$) hem de onunla beraber bulunan standart hatası kullanılır. Her bir maddenin uygulanmasından sonra önceden belirlenen yetenek düzeyi ve standart hata değerlerine ulaşıp ulaşılmadığı değerlendirilir. BBT uygulamasının amacı sınıflama olmadığına ise önceden belirlenmiş olan standart hata düzeyi kullanılır ve tüm bireylerin puanları hemen hemen aynı doğrulukta hesaplanır. Uygulama, önceden belirlenmiş olan standart hata düzeyine ulaşıldığında sonlandırılır (Weiss, 2004).

1.5.1.7. Kapsam Dengelenmesi

BBT uygulamasında en çok araştırılan ve geliştirilen testler başarı testleridir. Bazı başarı alanları hem tek boyutlu hem de göreceli olarak homojen olup, bazıları da göreceli olarak tek boyutlu olmasına rağmen iki ya da daha fazla kapsam alanı içermektedir. Örneğin, ilkokul düzeyindeki temel aritmetik işlemleri (toplama, çıkarma, çarpma ve bölme) tek bir boyutta ölçeklenebilirken, her bir işlem farklı bir ölçme amacına hizmet etmektedir. Bu işlemlerin tek bir ölçekte ölçeklenebilmesi MTK kullanılan BBT uygulaması ile yapılabilir. Ancak işlemlerin zorluk düzeylerinin farklı olması sebebiyle bireylerin aldığı testlerdeki ağırlıklandırmalar farklı olabilir (Örneğin; yüksek yetenek düzeyine sahip öğrenciler, daha çok bölme işlemi olan

maddeleri alma eğilimindeyken; düşük yetenek düzeyine sahip öğrenciler, daha çok toplama işlemi olan maddeleri alma eğiliminde olabilirler). Kısacası, öğrenciler tek bir ölçekte ölçülürken farklı kapsamdaki testleri alabilirler (Weiss, 2004).

Kapsam dengesinin sağlanması için çeşitli yöntemler kullanılmaktadır. Bu yöntemler, madde seçme sürecinde maddelerin kapsam kategorileri dikkate alınarak en çok bilgi madde seçme yönteminin tekrar düzeltilmesiyle uygulanır. Bireyin mevcut yetenek düzeyinden ($\hat{\theta}$) en çok bilgiyi sağlayan madde seçilir. Eğer seçilen madde bireyin testindeki bir kapsam alanında daha az kullanılmışsa madde uygulanır. Eğer değilse, sonraki en çok bilgiyi veren madde kapsam hedeflerine uygun olarak uygulanır ve süreç uygun kapsam hedefi sağlanana kadar devam eder (Weiss, 2004).

Kapsam dengelenmesinin yapıldığı testler, kapsam dengelenmesinin yapılmadığı testlere göre daha uzun olurlar. Yetenek, kişilik ve tercih ölçümleri kapsamda görece olarak daha homojen ve tek boyutlu olduğu için kapsam dengelenmesi gerektirmezken, başarı testleri için kapsam dengelenmesi gereklidir (Weiss, 2004).

1.5.1.8. Madde Kullanım Sıklığı Kontrolü

Madde kullanım sıklığı kontrolü, aynı maddelerin aşırı derecede kullanılmasını (over-exposure) engelleyerek, madde havuzunun bütünlüğünü ve madde havuzundaki maddelerin gizliliğini korumaya odaklanır (Davis ve Dodd, 2005).

Farklı yetenek düzeyindeki bireylere farklı maddelerin uygulandığı BBT’de, bireylerin yetenek dağılımları ve madde havuzundaki bilgi yapısı arasındaki ilişkiye bağlı olarak maddeler farklı oranlarda kullanılabilir (International Association for Computerized Adaptive Testing, 2013). Madde seçme yöntemlerinin kullanımı, bilgi ve kestiriminin standart hatası arasında var olan ters ilişki nedeniyle ölçme kesinliğini maksimum yapar. Ancak ölçme kesinliğini maksimum yapacak maddeler seçildiğinde, bazı maddelerin tüm bireylere uygulandığı; bazı maddelerin ise hiç bir bireye uygulanmadığı görülür ve bu da düzgün olmayan madde kullanım sıklığı oranları meydana getirir (Pastor, Dodd ve Chang, 2002). Bu duruma örnek olarak

Hulin, Drasgow ve Parsons'ın 1983 yılında yaptığı çalışmanın sonuçları verilebilir. Hulin, Drasgow ve Parsons (1983) simülasyon çalışmasında, madde seçme yöntemi olarak En Çok Bilgi yöntemi kullandığında 260 maddeden 141 tanesinin hiç uygulanmadığını saptamıştır (Akt. Revuelta ve Ponsoda, 1998).

Ayrıca aynı madde havuzundan oluşturulan bireyselleştirilmiş testlerle birden çok kez değerlendirilen bir bireye aynı sorular sorulabilir ve birey de doğru cevapları öğrenebilir. En sık kullanılan maddeler popüler hale gelerek orijinal olan psikometrik özelliklerini kaybederler. Bu durum da, testin geçerliğinin düşmesine sebep olur. Bir test geliştiricisi için geniş bir madde havuzu oluşturup bu maddelerin büyük bir yüzdesinin kullanılmaması istenmeyen bir durumdur (Revuelta ve Ponsoda, 1998).

Revuelta ve Ponsoda (1998) bir maddenin kullanılma oranının; (a) maddenin psikometrik özelliklerine, (b) havuzda uygulanabilir olan diğer maddelere ve (c) bireylerin yetenek dağılımlarına bağlı olduğunu belirtmişlerdir.

Stocking ve Lewis (2000), madde kullanım sıklığı ve ölçme kesinliği arasındaki ilişkiyi balona benzetmişlerdir. Bir taraftan bastırılan balon, tepki olarak diğer taraftan şişer. Bu nedenle madde kullanım sıklığı kontrol edilirken ölçülen bireyin yetenek kestiriminin kesinliği düşer. İyi bir kullanım sıklığı kontrol yönteminin amacı bu ilişkiye karşı gelerek; ölçme kesinliğini düşürmeden, madde havuzunun daha dengeli kullanılmasını sağlamaktır (Pastor, Dodd ve Chang, 2002).

Bu araştırmada madde kullanım sıklığı kontrol yöntemlerini incelemek amaçlandığı için bundan sonraki bölümde madde kullanım sıklığı kontrol yöntemlerinin çeşitleri hakkında genel bilgi yer almaktadır.

1.5.2. Madde Kullanım Sıklığı Kontrol Yöntemleri

Madde kullanım sıklığı kontrol yöntemleri çeşitli kategorilerde sınıflandırılmaktadır. Way (1998), bu yöntemleri tesadüfi seçme (randomization) ve koşullu seçme (conditional selection) olarak iki kategoriye ayırırken; yeni yöntemlerin de geliştirilmesiyle birlikte Georgiadou, Triantafillou ve Economide (2007), tesadüfi

seçme (randomization), koşullu seçme (conditional selection strategies), tabakalı (stratified strategies), birleştirilmiş (combined strategies) ve çok aşamalı bireyselleştirilmiş test desenleri (multiple stage adaptive test designs) yöntemleri olarak beş kategoriye ayırmıştır. Georgiadou, Triantafillou ve Economide (2007)'nin çalışmasına dayanarak, aşağıda madde kullanım sıklığı kontrol yöntemlerinin yer aldığı kategoriler hakkında genel bir bilgi verilmiştir.

1.5.2.1. Tesadüfi Seçme(Randomization) Yöntemleri

Tesadüfi seçme kategorisi altında yer alan tüm yöntemler, maksimum bilginin en uygun (optimal) düzeyi civarındaki bir grup madde arasından tesadüfi olarak seçilen bir maddenin uygulanma frekansını kontrol etmeye çalışmaktadır. Bu yöntemler, başlangıç maddelerinden sonra bireylerin yetenek düzeylerinin önemli ölçüde ayrışacağı ve daha sonraki maddelerin önemli ölçüde değişeceği varsayımına dayanmaktadır. 5-4-3-2-1 stratejisi (5-4-3-2-1 Strategy), Bilgi Koşullu Tesadüfi stratejisi (Randomesque Strategy), Bilgi 4 prosedürü (INFO4 Procedure), 0.10 logit içinden (within 0.10 logits) stratejisi ve kademeli (progressive) stratejisi yöntemleri bu yöntemler arasında yer almaktadır.

1.5.2.2. Koşullu Seçme Yöntemleri (Conditional Selection Strategies)

Bu kategori altında yer alan yöntemler, kullanım sıklığı kontrol parametresi kullanarak maddenin kullanım sıklığı oranını kontrol eder. Test öncesinde bir maksimum değer (r) belirlenir ve madde seçildiğinde uygulanıp uygulanmayacağı, maddenin kullanım sıklığı kontrol parametresine bağlı olur. Kullanım sıklığı kontrol parametresi sık kullanılan maddeler için önceden belirlenmiş olan kullanım sıklığı oranı kadar düşük olabilir. Bu durum da maddenin kullanımını kısıtlar. Seyrek kullanılan maddeler için, kullanım sıklığı kontrol parametreleri 1.0 kadar yüksek olabilir. Bu da maddenin seçildiğinde neredeyse her zaman bireye uygulanabileceğini gösterir. Bu parametrelerin değerleri, BBT uygulamasından önce, karışık bir dizi tekrarlı simülasyon sonrasında belirlenir. Bu kategori altında yer alan yöntemler; Sympson-Hetter (SH) Stratejisi (the Sympson-Hetter strategy),

Geniřletilmiř SH Stratejisi (Extended Sympson-Hetter Strategy), Davey ve Parshall Stratejisi (Davey-Parshall Strategy), Stocking ve Lewis ok Terimli Stratejisi (Stocking and Lewis Multinomial Strategy), Kısıtlı En ok Bilgi Stratejisi (Restricted Maximum Information Strategy), SH Kořullu Prosedürü Stratejisi (Sympson- Hetter Conditional Procedure), Stocking ve Lewis Kestirilen Yetenek Düzeyine Kořullandırma (Stocking and Lewis conditioning on estimated ability), Hedeflenen Kullanım Sıklığı Kontrol Stratejisi (Targeted Exposure Control Strategy), Chen ve Lei Stratejisi (Chen and Lei Strategy) ve van der Linden ve Veldkamp Gölge Test Yaklařımı (van der Linden and Veldkamp Shadow Test Approach) yöntemleridir.

1.5.2.3. Tabakalı Yöntemler (Stratified Strategies)

Tabakalı yöntemlerin temel amacı, testin en faydalı (beneficial) noktasında kullanılan maddenin kullanım sıklığını sınırlamaktır. Bu yöntemler; a-Tabakalama (a-STR) Yöntemi (a-Stratified Strategy), b parametresinin bloke edildiđi a-STR (the a-Stratified strategy with b-blocking), tabaka boyunca eřit olmayan madde kullanım sıklığı a-STR BBT (a-STR CAT with Unequal Item Exposure across Strata), Kapsam Dengelenmeli a-STR BBT Deseni (a-STR CAT Design with content Blocking), ok Boyutlu tabakalama (Multi-dimensional stratification) ve 0-1 Tabakalama stratejisi (the 0-1 stratification strategy) yöntemleridir.

1.5.2.4. Birleřtirilmiř Yöntemler (Combined Strategies)

ođu arařtırmacı madde kullanım sıklığını kontrol etmek için bir yöntem kullanmak yerine farklı yöntemleri bir araya getirerek daha güçlü yöntemler geliřtirmeyi amaçlamıřlardır. Kademeli Kısıtlı strateji (Progressive Restricted strategy), Nering, Davey ve Thompson Hibrid stratejisi (Nering, Davey and Thompson's Hybrid strategy), Eggen'in stratejisi (Eggen's strategy), Kapsam Dengelenmesi Yapılan a-STR'yle SH'nin Birleřtirilmesi (Incorporation of the SH into a-STR with Content Blocking), a-STR'nin SH ile Birleřmesi (Combination of the a-STR with the SH stratejisi), Gölge Testi Kullanan a-STR BBT'de Kapsam Kısıtlaması (Content Constraints in a-STR CAT using a shadow Test) yöntemleri birleřtirilmiř yöntemler arasında yer almaktadır.

1.5.2.5. Çok Aşamalı Bireyselleştirilmiş Test Desenleri (Multiple Stage Adaptive Test Designs)

Madde kullanım sıklığını kontrol etmek için geliştirilen bir diğer alternatif yöntem, ön düzenlemeli (preconstructing) bireyselleştirilmiş testlerin oluşturulmasıdır. Bu amaç doğrultusunda, bazı araştırmacılar tarafından çok aşamalı bilgisayarda bireyselleştirilmiş testler (ÇABBT) geliştirilmektedir. ÇABBT, BBT uygulamalarında olduğu gibi bireyin yeteneğine uyum sağlarken; önceden seçilmiş (pre-selected) maddelerin kullanım sıklığı ve paralel test formlarının test uzmanı tarafından incelenmesi gibi kağıt kalem testinin faydalarını da sağlamaktadır. Bu nedenle ÇABBT, hem standart BBT uygulaması hem de kağıt kalem testinin özelliklerini biraraya getirmektedir (Armstrong ve Edmonds, 2004, Akt. Georgiadou, Triantafillou ve Economide, 2007). Bu kategori altında yer alan yöntemler; Bireyselleştirilmiş Bilgisayarlı Sıralı Test (Computerized Adaptive Sequential Testing), Bireyselleştirilmiş Çoklu Aşamalı Madde Demetleri (Adaptive Multi-stage Item Bundles) ve Çoklu FormYapıları (Multiple Forms Structures) yöntemleridir.

2. İLGİLİ ARAŞTIRMALAR

Chang ve Twu (1998), çalışmalarında 5-4-3-2-1 tesadüfi seçim, Sympson-Hetter, Davey ve Parshall, Stocking ve Lewis koşulsuz çok terimli ve Stocking ve Lewis koşullu çok terimli yöntemlerini karşılaştırmışlardır. Her bir kullanım sıklığı kontrol yöntemi madde seçme yöntemiyle birleştirilmiş ve BBT deseni bu çalışma üzerine dayandırılarak ilerletilmiştir. Bu yöntemlerin artıları ve eksiklikleri, farklı madde havuzu büyüklükleri ve farklı istenen en yüksek kullanım sıklığı (desired maksimum exposure) oranları altında incelenmiş ve test güvenliği, madde çakışması oranı ve koşullu ölçmenin standart hatası (conditional standard errors of measurement) açısından değerlendirilmiştir. Çalışmada 360 ve 720 farklı maddeden oluşan madde havuzları kullanılmıştır. Büyük madde havuzu küçük madde havuzunu kapsamaktadır. Beklenen maksimum kullanım sıklığı oranları da 0.10 ve 0.20 olarak belirlenmiştir. Yeteneğin başlangıç noktası 0 alınmış, yetenek kestirimi için en çok olabilirlik kestirimi yöntemi, madde seçimi için ise en çok bilgi madde seçme yöntemi kullanılmıştır. Kapsam dengelenmesinde Kingsbury ve Zara mekanizması tercih edilmiştir. Test uzunluğu 30 maddeye sabitlenerek, madde parametrelerinin kestirimi 3 PLM'ye göre yapılmıştır. Araştırma sonucunda, sadece madde havuzunun büyüklüğünün test güvenliğini garanti etmek için yeterli olmadığı sonucuna ulaşılmıştır. Stocking ve Lewis koşullu çok terimli yönteminin en düşük test çakışması oranlarını üreterek gözlenen kullanım sıklığı oranının kontrol edilmesinde en iyi sonucu verdiği görülmüştür.

Revuelta ve Ponsoda (1998) çalışmasında madde kullanım sıklığı kontrolü için iki yeni yöntem önermiştir. Önerilen Kademeli Yöntem (PR) ve Kısıtlı En Çok Bilgi Yöntemi (Rk); En Çok Bilgi (MI), Bir Parametre (IP), McBride ve Martin (MM), Bilgi Koşullu Tesadüfi (RA), Sympson ve Hetter (SH) ile Tesadüfi Madde Seçme yöntemlerini test kesinliği ve madde kullanım sıklığı değişkenleri açısından karşılaştırmışlardır. Çalışmada iki farklı durum raporlanmıştır. Birincisinde tüm yöntemler gerçek madde bankası ile sabit ve değişen uzunluklu test durumlarında, ikincisinde ise MI, SH, PR ve Rk yöntemlerinin farklı test uzunlukları ve maddelerin

ayirt ediciliklerinin farklı olduğu durumlarda karşılaştırılmıştır. Birinci çalışmada sabit uzunluklu test olarak 35 madde ve değişken test uzunluğu için yeteneğin standart hatasının 0.22'den daha düşük olduğu ya da maksimum 50 madde uzunluğuna ulaşıldığında durdurma kuralı kullanılmıştır. Çalışmada gerçek yetenek düzeyleri normal dağılan 2000 birey simüle edilmiştir. Başlangıç kuralı olarak bireylerin yetenekleri sıfıra eşitlenmiştir. Madde havuzunda 210 madde yer almıştır. Sonuçlar beş tekrarın ortalaması alınarak oluşturulmuştur. İkinci çalışmada ise üç madde ayirt edicilik dağılımı (lognormal ortalama 0.93, lognormal ortalama 1.07 ve lognormal ortalama 1.30), üç test uzunluğu (20, 40 ve 60) ve iki madde havuzu büyüklüğü (500 ve 1000) kullanılmıştır. Bireylerin sayısı ve başlangıç kuralı birinci çalışmadaki ile aynıdır. Her bir kombinasyon için madde havuzu oluşturulmuş ve b parametresi normal dağılımdan (0,1), c parametresi ise Beta dağılımından (5,17) türetilmiştir. Araştırma sonucunda Rk Yönteminin maksimum kullanım sıklığı oranlarını düşürmede ve PR yönteminin de kullanılmamış maddelerin sayısını düşürmede yararlı olduğu görülmüştür. Bu nedenle; test kesinliğinde ciddi düşüşler olmadan madde kullanım sıklığını kontrol etmede faydalı olabilecek Kademeli Kısıtlı birleştirilmiş madde kullanım sıklığı yöntemi önerilmiştir.

Davis (2002) çalışmasında; altı farklı madde kullanım sıklığı kontrol yöntemini [Bilgi Koşullu Tesadüfi (3 ve 6 maddeli), 0.10 logit içinden (3 ve 6 maddeli), Sympon-Hetter, koşullu Sympon-Hetter, a-Tabakalama ve geliştirilmiş a-Tabakalama yöntemleri] üç tane çoklu puanlanan MTK modeli [Kısmi bilgi (partial credit), genelleştirilmiş kısmi bilgi (generalized partial credit) ve aşamalı tepki (graded response) modelleri] altında incelemiştir. Madde kullanım sıklığı oranlarını kontrol etmek için, kullanım sıklığı kontrolünün yapılmadığı durum referans (baseline) olarak alınmıştır. Araştırmanın sonuçları üç MTK modelinde de dikkat çekecek derecede benzer bulunmuştur. Bilgi Koşullu Tesadüfi-6 madde ve 0.10 logit içinden-6 madde yöntemleri ölçme kesinliği ve uygulama kolaylığı açısından kullanım sıklığı oranlarını kontrol etmede en iyi sonuçları sağlamışlardır. Bilgi Koşullu Tesadüfi-3 madde ve 0.10 logit içinden-3 madde yöntemleri; kullanım sıklığının kontrolü, çakışma ve madde havuzu kullanımı oranlarının istenen

düzyeyde olmasında etkili olamamıştır. Sympson-Hetter ve koşullu Sympson-Hetter yöntemlerinin uygulaması zor olmuş ve zaman almıştır. Hedeflenen düzeyde kullanım sıklığı oranlarını kontrol etmelerine rağmen (Sympson-Hetter için), madde çakışması ve madde havuzu kullanımı açısından beklenen performans sağlanamamıştır. a-Tabakalama ve geliştirilmiş a-Tabakalama yöntemleri tüm değişkenlerde şaşırtıcı bir şekilde zayıf performanslar göstermiştir.

Pastor, Dodd ve Chang (2002) çalışmalarında çoklu puanlanan maddelerden oluşan bir madde havuzunda madde kullanım sıklığı kontrol yöntemlerini karşılaştırmışlardır. Genelleştirilmiş kısmi bilgi modeli kullanılarak kalibre edilen iki farklı büyüklükteki madde havuzunda beş farklı madde kullanım sıklığı kontrol yöntemi (a-Tabakalama, Sympson-Hetter, Geliştirilmiş Tabakalama, Koşullu Sympson-Hetter ve Koşullu Geliştirilmiş Tabakalama) incelenmiştir. Çalışmada yer alan üç örneklemden ikisi (8000 ve 15000 birey) kullanım sıklığı kontrol parametrelerini tespit etmek, biri ise BBT simülasyonu için (1000 birey) kullanılmıştır. 60 ve 100 maddeden oluşan iki madde havuzu 1994 NAEP örneklemindeki parametrelerin değerlerine göre oluşturulmuştur. Tüm BBT durumları için test uzunlukları 15 madde olarak belirlenmiştir. Madde seçiminde en çok madde bilgi fonksiyonu ve yetenek kestirimi için en çok olabilirlik kestirim yöntemi kullanılmıştır. Başlangıç yetenek kestirimi tüm durumlarda tüm bireyler için sifıra eşitlenmiştir. Simülasyon çalışması sonucunda, a-Tabakalama deseni, kullanım sıklığı kontrolü olmadığı durumla karşılaştırıldığında madde kullanım sıklığını ve madde çakışmasını azaltabildiği, havuz kullanımını arttırabildiği, ölçme kesinliğini ise küçük çapta azalttığı görülmüştür. Sympson-Hetter ve koşullu Sympson-Hetter gibi çok sınırlayıcı kullanım sıklığı kontrol yöntemlerinin kullanılması, ölçme kesinliği pahasına kullanım sıklığını iyi kontrol etmiştir. Sonuç olarak kullanım sıklığı kontrol parametrelerinin yakınsaması, çok sınırlayıcı kullanım sıklığı kontrol algoritmalarının bazıları için sorunlu olmuştur. Daha çok basite indirgeyen kullanım sıklığı kontrol yöntemlerinin kullanımı, özellikle test uzunluğunun madde havuzu büyüklüğüne olan oranı büyük olduğunda, tavsiye edilmiştir.

Burt, Davis ve Dodd (2003), çalışmalarında genelleştirilmiş kısmi bilgi modeline dayalı BBT uygulaması kullanarak üç madde kullanım sıklığı kontrol yöntemini (Simpson-Hetter, değiştirilmiş 0.10 logit içinden ve Bilgi Koşullu Tesadüfi yöntemleri) incelemişlerdir. Değiştirilmiş 0.10 logit içinden ve Bilgi Koşullu Tesadüfi yöntemleri için 3 ya da 6 maddeden oluşan gruplar kullanılmıştır. Simpson-Hetter yöntemi için hedeflenen kullanım sıklığı (exposure kontrol target) 0.29 olarak alınmıştır. Madde havuzu NAEP 1996 fen testinden 210 madde kullanılarak hazırlanmıştır. Kullanım sıklığı kontrol parametrelerinin üretilmesi için (Simpson-Hetter yöntemi için) 8000 birey ve BBT simülasyonlarında kullanılması için de 1000 birey simüle edilmiştir. Başlangıç yetenek düzeyi sıfıra eşitlenmiş, yetenek kestirimi yöntemi için EÇO, kapsam dengelenmesi için Kingsbury ve Zara'nın kapsam dengeleme yöntemi kullanılmış ve testler 20 maddeye sabitlenmiştir. Madde kullanım sıklığı kontrol yöntemlerinin karşılaştırılması için, madde kullanım sıklığı kontrolünün yapılmadığı durum referans olarak çalışmaya dâhil edilmiştir. Sonuçta; Simpson-Hetter yönteminin, maksimum kullanım sıklığı oranlarını daha iyi kontrol etmesi açısından diğer yöntemlerden daha iyi performans sergilediği, 6 maddeli Bilgi Koşullu Tesadüfi yönteminin ise yetenek kestiriminde en yüksek ortalama standart hataya sahip olurken, diğer yöntemlere göre madde havuzunu daha çok kullandığı görülmüştür.

French ve Thompson (2003) çalışmalarında, yeni bir yöntem olan hedeflenen kullanım sıklığı kontrol yöntemi (targeted exposure control, TEC) ve Simpson-Hetter (SH) yöntemi ile b parametresinin bloke edildiği a-Tabakalama yöntemlerinin değişken uzunluklu bir BBT uygulamasında madde havuzunun kullanımı, test uzunluğu ve test güvenilirliğine etkilerini araştırmışlardır. Performanslar madde havuzunun kullanımı, ölçme kesinliği, test çakışması ve uygulamanın kolaylığı açısından değerlendirilmiştir. Çalışmada test uzunluğu en yüksek 20 ve en düşük 7 olarak belirlenmiştir. SH ve TEC yöntemlerinden kullanım sıklığı parametrelerinin kestirilmesinde $N(0,1)$ dağılımından 4000 birey seçilmiş ve 125 tekrarla süreç tamamlanmıştır. Madde havuzu 269 maddeden oluşmaktadır. Madde parametreleri 3PLM'ye dayalı olarak Marjinal En Çok Olabilirlik Kestirimi kullanılarak kestirilmiştir.

BBT uygulamasında kullanılan örneklem büyüklüğü 10.000 olarak belirlenmiştir. SH yönteminde madde seçimi olarak en çok bilgi seçimi yöntemi kullanılmıştır. Araştırma sonuçlarına göre; güvenilirlik, yanlılık ve RMSE (Hataların Ortalama Karekökü) değerlerinde tüm yöntemler benzer sonuçlar vermiştir. SH yöntemi en kısa test uzunluğu ortalamasını vermiş, TEC yöntemi ise madde havuzu kullanımı, test çakışması, ki-kare istatistiği ve en yüksek kullanım sıklığı oranı açısından en iyi sonuçları vermiştir. Ancak uç noktalardaki bireyler için hiçbir yöntem yeterli derecede iyi sonuç vermemiştir.

Davis ve Dodd (2005) araştırmasında kullanım sıklığı kontrol yöntemlerini çoklu puanlanan maddelerden oluşan bir madde havuzunda test etmiştir. Kısmi bilgi modelin kullanıldığı BBT uygulamasında dört madde kullanım sıklığı yöntemi incelenmiştir. Kullanım sıklığının kontrol edilmediği durum referans olarak alınarak, Kingsbury-Zara, değiştirilmiş 0.10 logit içinden, Sympton-Hetter ve koşullu Sympton-Hetter yöntemleri kullanılmıştır. Kingsbury-Zara ve değiştirilmiş 0.10 logit içinden yöntemleri 3 veya 6 aday madde durumu ile uygulanmıştır. Madde havuzu için 157 tane çoklu puanlanan geniş ölçekli bir testin gerçek öğrenci verilerinden elde edilen madde parametreleri kullanılmıştır. BBT simülasyonu 20 maddelik sabit uzunluklu bir test olarak alınmıştır. Başlangıç yeteneği sıfıra ayarlanmış, madde seçiminde en çok bilgi (maximum information), yetenek kestiriminde EÇÖ, kapsam geçerliğinin sağlanmasında ise Kingsbury-Zara (1989) yöntemi kullanılmıştır. Araştırma sonucunda elde edilen bulgulara göre, 6 aday madde kullanılan Kingsbury-Zara ve değiştirilmiş 0.10 logit içinden yöntemlerinin kullanım sıklığı oranı, çakışma oranı ve madde havuzu kullanımı açısından koşullu Sympton-Hetter kadar iyi performans sergilemiştir.

Han (2009), çalışmasında kendisi tarafından geliştirilen Aşamalı En Çok Bilgi Oranı madde seçme yaklaşımının etkililiğini test etmeyi amaçlamıştır. Bunun için bir ay boyunca uygulanan bir BBT uygulaması simüle edilmiştir. Çalışma sonucunda; madde kullanım sıklığı oranı, test bilgisi, madde havuzu kullanımı ile yanlılık ve hatalardaki kestirim yeterliği değerleri incelenmiştir. Madde havuzu GMAT madde

havuzundan 500 çoktan seçmeli soru seçilerek oluşturulmuştur. Çalışma sonuçlarının genellenebilirliğini arttırmak için kapsam dengelenmesi göz ardı edilmiştir. Simülasyon çalışması standart normal dağılımdan ($N(0,1)$), toplamda 10000 birey ile yürütülmüştür. Her bir bireye 40 madde uygulanmıştır. 20 gün için her bir günde 500 bireye test iki zaman aralığında, bir zaman aralığında 250 kişi olacak şekilde uygulanmıştır. Her bir zaman aralığından sonra madde kullanım bilgisi güncellenmiştir. Madde seçme yöntemi olarak; yetenek başlangıç değerine yakın b parametrelerinin yer aldığı beş madde arasından seçme yöntemi (Yöntem1), Fisher'in En Çok Bilgi yöntemi (FEB) (Yöntem2), FEB ile Azalarak Kaybolma (AK) madde kullanım sıklığı yönteminin birlikte kullanıldığı yöntem (Yöntem3), Aşamalı En Çok Olabilirlik Oranı (AEBO) (Yöntem4) ve AEBO ile AK madde kullanım sıklığı yönteminin birlikte kullanıldığı yöntem (Yöntem5) seçilmiştir. Bireylere başlangıç olarak -0.5 ve 0.5 arasındaki madde güçlük parametresine sahip maddeler uygulanmış, yetenek kestirimi yöntemi olarak ESD kullanılmıştır. Beş madde seçim yönteminin değerlendirilmesinde yetenek kestirimi performansına ve madde havuzu kullanımına odaklanılmıştır. Her bir yöntem 100 kere tekrar edilmiş ve medyan değerleri rapor edilmiştir. Çalışmanın sonucunda yeni AEBO yaklaşımı BBT'nin erken aşamalarında FEB temelli yöntemlere göre çok az da olsa daha düşük SEE'ye sahip olmuştur. Madde havuzunun kullanımının etkililiğine bakıldığında, Yöntem 5'in madde kullanım sıklığı kontrolünde diğer yöntemlere göre açık ara farkla daha iyi sonuçlar elde ettiği görülmüştür. Yöntem 2 ve Yöntem 4 kullanıldığında, 500 maddenin aşağı yukarı 150'si kullanılmamıştır. Bu da madde havuzunun gerçek büyüklüğünün 350 madde olduğunu göstermektedir. Yöntem 5 kullanıldığında madde kullanım sıklığı oranları 0.12'den küçük olmuştur. Bu yüzden diğer çalışma yöntemleri ile karşılaştırıldığında Yöntem 5 madde havuzu kullanımında açık bir şekilde daha verimli ve etkili olmuştur. Yöntem 2 ve Yöntem 4 ise benzer sonuçlar vermiştir.

Lee ve Dodd (2012) çalışmalarında madde havuzunun özellikleri ve yetenek dağılımının çeşitli kombinasyonları altında kısmi bilgi modeline dayalı BBT'de madde kullanım sıklığı kontrol yöntemlerini incelemiştir. Madde havuzu özellikleri

(kolay, orta ve zor madde havuzu için 120 madde), iki yetenek dağılımı (normal dağılımı ve negatif çarpık) ve üç kullanım sıklığı kontrol yöntemi (Bilgi Koşullu Tesadüfi yöntemi, kademeli-kısıtlı yöntemi ve en çok bilgi yöntemi) olmak üzere üç değişken manipüle edilmiştir. Araştırmada betimsel istatistikler, bilinen ve kestirilen yetenek düzeyleri arasındaki korelasyon, RMSE ve ortalama mutlak fark (average absolute difference) gibi ölçme kesinliği göstergeleri, kullanım sıklığı oranları, madde kullanımı ve madde çakışması hesaplanmıştır. Beklendiği gibi orta zorluk düzeyindeki havuz hem kolay hem de zor madde havuzlarından daha iyi performans göstermiştir. Kademeli-kısıtlı yönteminin maksimum kullanım sıklığı oranları, ortalama madde çakışması ve havuz kullanımı açısından hem Bilgi Koşullu Tesadüfi yöntemi hem de en çok bilgi yöntemine göre daha iyi çalışmıştır. Uyumun sağlanmadığı durumda kolay madde ile negatif çarpık örneklemin verisi en kötü performansı göstermiştir.

Han (2012) çalışmasında Verimli Dengelenmiş Bilgi (efficiency balanced information) adlı yeni bir madde seçme yöntemi önererek, Fisher'in En Çok Bilgi yöntemi ve tabakalama yöntemleri ile bu yöntemi karşılaştırmıştır. Çalışmada iki tane simülasyon durumu oluşturulmuştur. Bunlardan ilki sabit uzunluklu BBT, ikincisi ise değişken uzunluklu BBT uygulaması için kullanılmıştır. Çalışmada Verimli Dengelenmiş Bilgi, Fisher'in En Çok Bilgi Yöntemi, b parametresinin bloke edildiği a-Tabakalama ve b Eşleme (b Matching) madde seçme yöntemleri ile Sympton-Hetter (hedeflenen kullanım sıklığı=0.20) ve Azalarak Kaybolma (AK) madde kullanım sıklığı kontrol yöntemleri kullanılmıştır. Araştırmada her bir durum için yetenek kestiriminin standart hatası, ölçmenin koşullu standart hatası, koşullu yanlılık, madde kullanım oranları ve test çakışması hesaplanmıştır. Simülasyon çalışmalarında yükseköğretim için BBT'nin 20 günlük uygulamaları taklit edilmiş ve uygulama sonrasında kestirimin ve ölçmenin koşullu standart hatası ile madde havuzunun kullanımı incelenmiştir. Çalışmada standart normal dağılımdan 10000 birey üretilmiş ve madde havuzunun oluşturulmasında GMAT (Graduate Management Admission Test) madde bankasından 500 çoktan seçmeli matematik maddesi çekilmiştir. Maddelerin ortalaması 0 ve standart sapmaları 1 civarında

elde edilmiştir. Madde havuzundaki maddelerin b ve c parametreleri arasında anlamlı bir korelasyon bulunmazken, a ve b parametreleri arasında zayıf bir korelasyon bulunmuştur ($r=0.367$). İlk madde için bireylerin başlangıç thetaları -0.5 ile 0.5 aralığında tesadüfi olarak seçilmiş ve bireylerin yetenek kestiriminde EÇO yöntemi kullanılmıştır. Tüm BBT uygulamaları her bir durum altında 25 kez tekrar edilmiş ve analizler için ortalama değerler alınmıştır. Birinci simülasyon çalışmasında test uzunluğu 30 madde ile sabitlenerek; ikinci simülasyon çalışmasında ise standart hatanın 0.3'e ulaştığı, test uzunluğu minimum 10- maksimum 40 madde ile sınırlandırılarak uygulamalar sonlandırılmıştır. Araştırma sonuçlarına göre, yeni yöntem madde havuzunun tabakalandırılması gerekliliğini ortadan kaldırırken, düşük a parametresi değerlerine sahip maddeleri faydalı hale getirmiştir. Bununla beraber bu yeni yöntemin, Azalarak Kaybolma yöntemi ile beraber kullanıldığı her iki simülasyon durumunda da madde havuzu kullanımının, test kesinliğinde kayıp olmaksızın, arzu edilen bir düzeyde olduğu sonucuna ulaşılmıştır.

2.1. İlgili Araştırmalar Özeti

İlgili araştırmalara bakıldığında; madde kullanım sıklığı kontrol yöntemlerinin incelenmesinde farklı koşulların bir araya getirildiği görülmektedir. Ancak araştırmalarda ilk madde seçilirken b parametresinin sifıra yakın olması, madde seçme yöntemi olarak En Çok Bilgi yönteminin kullanılması, yetenek kestiriminde En Çok Olabilirlik kestirim yönteminin kullanılması, gerçek verilerden elde edilen parametreler kullanıldığı için çalışmalara kapsam dengelenmesinin katılması gibi koşulların sıklıkla tercih edildiği görülmüştür.

Madde havuzu özelliğinin değiştirilip madde kullanım sıklığı kontrol yöntemlerinin test edildiği üç araştırmaya ulaşılmıştır. Bu üç araştırmanın ikisinde farklı madde havuzu büyüklükleri, birinde ise farklı dağılım özelliğine sahip madde havuzlarında madde kullanım sıklığı kontrol yöntemleri incelenmiştir. Farklı madde havuzu büyüklüklerinde yöntemlerin test edildiği araştırmalardan biri 1-0 puanlanan, diğeri ise çoklu puanlanan maddelerden oluşturulan madde havuzlarını kullanmıştır. 1-0

puanlanan maddelerden oluşan madde havuzu ile yapılan çalışmada, madde kullanım sıklığı kontrol yöntemlerinin madde havuzu büyüklüklerinden farklı şekillerde etkilendiği görülmüştür. Çoklu puanlanan maddelerin oluşturduğu madde havuzu ile yapılan çalışmada, madde havuzu büyüklüğünün madde kullanım sıklığı kontrol yöntemlerinde göz önünde bulundurulması gereken iki faktörden biri olduğu belirtilmiştir. Farklı dağılım özelliğine sahip madde havuzunda yapılan çalışmada ise çoklu puanlanan maddelerden oluşan bir madde havuzu kullanılmıştır. Çalışma sonucunda, madde havuzunun güçlük düzeyi yükseldiğinde ya da düştüğünde En Çok Bilgi yöntemi ve Bilgi Koşullu Tesadüfi-6 madde yöntemlerinde kullanılmayan madde sayısının arttığı görülmüştür.

Madde kullanım sıklığı yöntemlerine bakıldığında; tesadüfi seçme yöntemleri içinde yer alan madde kullanım sıklığı kontrol yöntemlerinin 1-0 puanlanan madde havuzlarında genel olarak diğer yöntemlere göre madde kullanım sıklığı kontrol etmedeki başarısının daha az olduğu, Sympson-Hetter ve türevi yöntemlerin madde havuzunun kullanımında başarılı olmasına rağmen, zahmetli bir sürece sahip olduğu, a-Tabakalama madde seçme yönteminin madde kullanım sıklığını kontrol etmede diğer madde seçme yöntemlerine göre daha başarılı olduğu sonuçlarına ulaşılmıştır.

Araştırmalar incelendiğinde; Bilgi Koşullu Tesadüfi yönteminin genellikle 3 ve 6 madde ile kullanıldığı, farklı güçlük düzeylerine sahip 1-0 puanlanan madde havuzlarında madde kullanım sıklığı yöntemlerinin incelenmediği ve Han (2009) tarafından geliştirilen Azalarak Kaybolma yönteminin araştırmacının kendisi dışında başka araştırmacılar tarafından çalışılmadığı görülmüştür. Bununla beraber, Azalarak Kaybolma yöntemi ile yapılan çalışmalarda tek oturumlu ve madde kullanım sıklığı kontrolünün gerçek zamanlı yapılması koşullarının araştırılmadığı ve madde kullanım sıklığının incelendiği araştırmalarda madde seçme yöntemi olarak Aşamalı En Çok Bilgi Oranı yönteminin de geliştiricisi (Han, 2009) dışında başka çalışmalarda kullanılmadığı tespit edilmiştir.

3. YÖNTEM

3.1. Araştırmanın Türü

Bu araştırmada madde kullanım sıklığı yöntemlerinin, farklı madde seçme yöntemleri ve farklı özellikteki madde havuzlarında kullanıldığında ölçme kesinliği ve test güvenliğinin değişimi incelenmiştir. Araştırmada mevcut teorik bilgiye yenilerinin eklenmesi amaçlandığı için temel araştırma niteliği taşımaktadır (Karasar, 2009).

3.2. Verilerin Türetilmesi

Literatürde BBT uygulamaları için kullanılan çeşitli simülasyon programları (CATsim, Firestar, catR) yer almaktadır. Bu araştırmada veri üretmek ve BBT uygulamasını gerçekleştirmek için SimulCAT simülasyon programı (Han, 2011) kullanılmıştır.

3.2.1. Örneklem Gruplarının Oluşturulması

Araştırmada iki farklı örneklem kullanılmıştır. Bu örneklemelerden biri BBT uygulamasının simülasyonunda kullanılan, diğeri ise Sympson-Hetter madde kullanım sıklığı kontrol yönteminde madde kullanım sıklığı kontrol parametrelerinin hesaplanmasında kullanılan örneklemelerdir.

BBT Örnelemi; BBT uygulamasının yapıldığı örneklem grubunda yer alan bireylerin yetenek dağılımları, ortalaması 0 ve standart sapması 1 olan normal dağılımdan ($N(0,1)$) türetilmiştir. Örneklem büyüklüğü 1000 olarak belirlenmiştir.

Sympson-Hetter Örnelemi; Madde kullanım sıklığı kontrol parametrelerinin elde edildiği tekrarlı (iterasyonel) simülasyon aşamasında yer alan örneklemin, BBT uygulamasında kullanılan örnekleme aynı dağılıma sahip olmasına dikkat edilmiştir. Bu nedenle örneklem grubunda yer alan bireylerin yetenek dağılımları, ortalaması 0 ve standart sapması 1 olan normal dağılımdan ($N(0,1)$) türetilmiştir. Örneklem büyüklüğü 10.000 olarak belirlenmiştir.

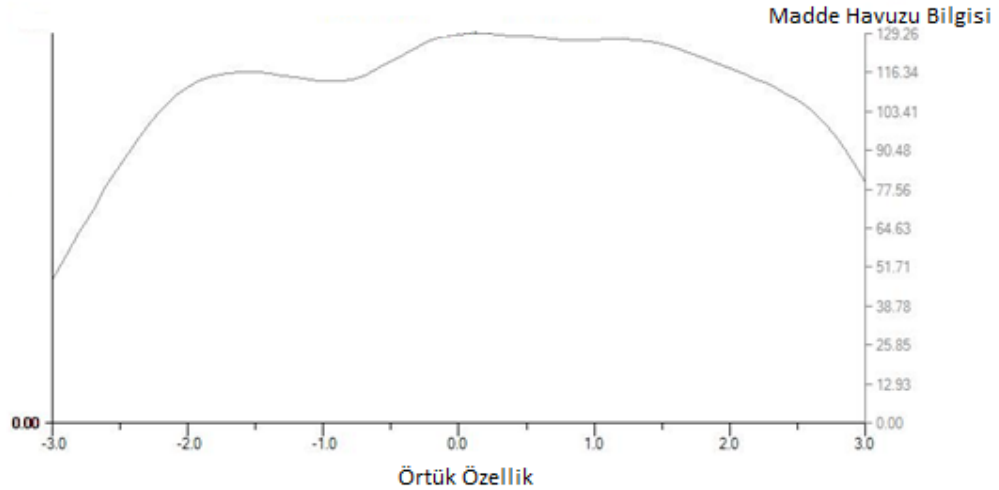
Örneklemlerin dağılımları ile ilgili daha detaylı bilgi için EK-1'de yer alan grafikler incelenebilir.

3.2.2. Madde Havuzlarının Oluşturulması

Araştırmada iki madde havuzu kullanılmıştır. Bu madde havuzlarının a ve c parametrelerinin dağılımları aynı iken b parametrelerinin dağılımları farklılaşmıştır. Birinci madde havuzu, başarı testlerini temsil etmek amacıyla b parametresinin tek biçimli dağıldığı orta güçlük düzeyine sahip olan bir havuzken; ikinci madde havuzu ise ölçüt referanslı testleri temsil etmek amacıyla b parametrenin normal dağılıma sahip olduğu yüksek güçlük düzeyine sahip bir havuzdur.

Birinci Madde Havuzunun Özellikleri;

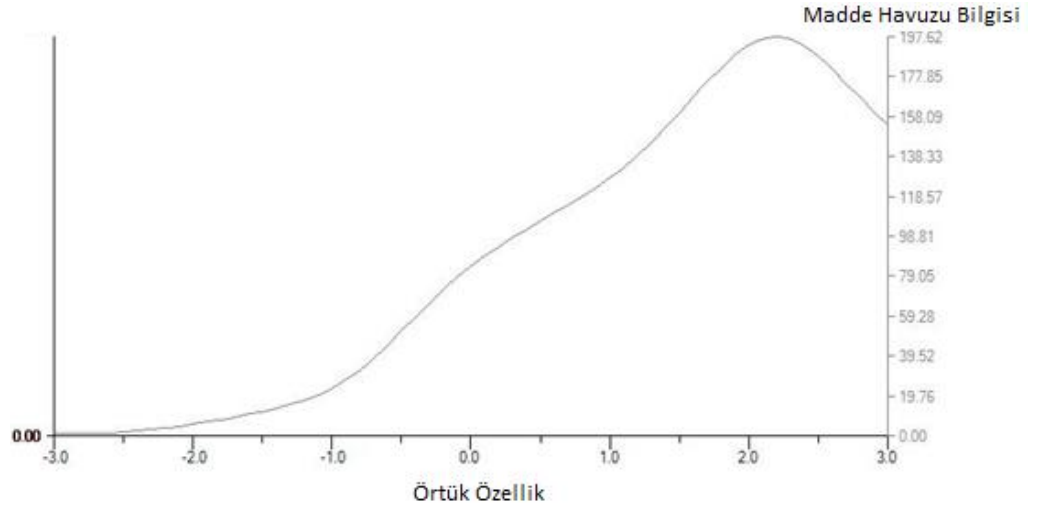
Birinci madde havuzunda kullanılacak maddelerin tüm parametreleri tek biçimli (uniform) dağılıma sahip olacak şekilde; a parametresi [0.50;2.00], b parametresi [-3.00;3.00] ve c parametresi [0.05;0.20] aralıklarında belirlenmiştir. Madde havuzunda 500 madde yer almıştır.



Şekil 3.1. Orta Güçlük Düzeyindeki Madde Havuzunun Test Bilgi Fonksiyonu

İkinci Madde Havuzunun Özellikleri;

İkinci madde havuzunda kullanılacak maddelerin güçlük parametresi ortalaması 2 ve standart sapması 1.5 olan normal dağılıma ($N(2;1.5)$) sahip olacak şekilde belirlenirken, a ve c parametreleri tek biçimli dağılıma sahip olacak şekilde a parametresi [0.50;2.00] ve c parametresi [0.05;0.20] aralıklarında belirlenmiştir. Madde havuzunda 500 madde yer almıştır.



Şekil 3.2. Yüksek Güçlük Düzeyindeki Madde Havuzunun Test Bilgi Fonksiyonu

3.3. BBT Koşulları

Çalışmada 1-0 puanlamaya dayalı olan BBT uygulaması simüle edilmiştir. Bu nedenle MTK modeli olarak BBT için en uygun model olarak belirtilen 3PLM kullanılmıştır. İlk madde seçiminde bireylerin yetenek düzeyleri sıfır olarak tanımlanmıştır. Yetenek kestiriminde ise BSD kestirim yöntemi kullanılmıştır. EÇO'dan farklı olarak BSD kestirimi bireylerin cevaplarının tümü doğru ya da tümü yanlış olduğunda dahi yetenek kestirimi sağlayabilmektedir (Embretson ve Reise, 2000). Sonlandırma kuralı olarak sabit test uzunluğu seçilmiş ve bu uzunluk 25 olarak belirlenmiştir. BBT uygulaması, bir günde 1000 kişinin uygulamaya katılması şeklinde simüle edilmiştir. Madde seçme yöntemi olarak araştırmmanın alt problemlerini oluşturan, Fisher'ın En Çok Bilgi (FEB), a-Tabakalama ve Aşamalı En

Çok Bilgi Oranı (AEBO) yöntemleri kullanılmıştır. Madde kullanım sıklığı yöntemlerinden ise Bilgi Koşullu Tesadüfi yöntemi (5 ve 10 maddeli), Sympon-Hetter yöntemi ve Azalarak Kaybolma yöntemi seçilmiştir.

Harwell (1996), örneklem yanlılığının ortadan kaldırılması için en az 25 tekrarın kullanılmasını önermektedir (Akt.Evans, 2010). Bu nedenle çalışmada 25 tekrar kullanılmıştır. Elde edilen sonuçların ortalaması alınarak araştırmanın analizleri yapılmıştır.

Araştırmada kullanılan madde kullanım sıklığı kontrolü ve madde seçme yöntemlerine ait teorik bilgiler aşağıda detaylı olarak verilmiştir.

3.3.1. Madde Kullanım Sıklığı Kontrol Yöntemleri

3.3.1.1. Bilgi Koşullu Tesadüfi (BKT) Yöntemi (Randomesque Method)

Madde seçiminde, en çok bilgi veren maddelerin çok fazla kullanılmasından kaçınmak için Kingsbury ve Zara (1989), tek bir en çok bilgi veren maddenin seçilmesi yerine, en çok bilgi veren maddelerin oluşturduğu madde setlerinden (2 maddeden 10 maddeye kadar) tesadüfi olarak madde seçmeye dayalı bir yöntem olan Bilgi Koşullu Tesadüfi yöntemini önermişlerdir (Kingsbury ve Zara, 1989; Davis ve Dodd, 2005; Han, 2011). Bu yöntemde birey için en çok bilgi verdiği belirlenen madde setinden uygulanması için bir madde tesadüfi olarak seçilir ve diğer maddeler havuza geri gönderilir. Bu işlem BBT uygulaması tamamlanana kadar devam eder (Davis ve Dodd, 2005; Macken-Ruiz, 2008).

Bu yöntem tüm madde kullanım sıklığı oranlarını hedeflenen oranda etkili bir şekilde sınırlamayabilir ama benzer yeterlikte testi alan bireyler için aynı maddenin tekrar bir şekilde kullanılmasını önleyebilir (Han, 2011).

3.3.1.2. Sympon-Hetter (SH) Yöntemi (Sympon-Hetter Method)

Sympon ve Hetter (1985) tarafından geliştirilen olasılığa dayalı olan bu yöntemde, uygulanacak maddenin olasılığı $P(A)$, madde seçme algoritması tarafından seçilebilecek maddenin olasılığından $P(S)$ ayırt edilir. Diğer bir deyişle Sympon-

Hetter yöntemi uygulanacak bir maddenin koşullu olasılığını $P(A|S)$ verir. $P(A)$ 'yı istenilen hedef düzeyinde tutmak için, $P(A|S)$ tekrarlı simülasyonlardan türetilen hedef $P(A)$ 'da elde edilir (Han, 2011). Yöntem, madde kullanım sıklığı kontrol parametrelerinin elde edildiği birinci faz ve gerçek BBT uygulamasına geçildiği ikinci faz olmak üzere iki fazlı bir uygulamayı içerir. Yöntemin ilk fazında elde edilen madde kullanım sıklığı kontrol parametreleri, ikinci fazda madde kullanım sıklığını kontrol etmede kullanılır.

Yöntemin birinci fazında, hedeflenen kullanım sıklığı oranına (target exposure control rate, r) yakınsamak için belirli sayıda tekrar yapılır. Bu fazda, her bir tekrarda Sympson-Hetter örneklemindeki bireylere BBT uygulanmasından oluşur ve uygulamada kullanılan her bir maddenin kullanılma sayısı tutulur. Bu sayı, uygulama için seçilen maddenin kaç kere seçildiğini gösteren sayının toplam örneklem büyüklüğüne bölünmesiyle, her bir madde için seçilme olasılığını $P(S)$ hesaplamakta kullanılır. Bir maddenin kullanım sıklığı kontrol parametresi, aşağıdaki kurallara göre maddeye ait $P(S)$ 'nin doğrudan bir sonucu olarak belirlenir;

$$\text{Eğer } P(S) > r, \text{ ise } K_i = r/P(S),$$

$$\text{Eğer } P(S) \leq r, \text{ ise } K_i = 1.0$$

Denklemden r hedeflenen kullanım sıklığı oranını ve K_i maddenin kullanım sıklığı kontrol parametresini göstermektedir. Eğer bir maddenin K_i 'si tek biçimli tesadüfi bir sayı değerini aşarsa, o madde uygulanabilir. Yüksek bir K_i değeri, maddenin seçildiğinde büyük bir olasılıkla uygulanacağını gösterir. Düşük bir K_i değeri ise madde seçilirse uygulanma şansının düşük olduğunu gösterir. İlk tekrarda, madde seçilirse uygulanabileceği gibi, tüm maddeler için K_i değerleri 1.0'e eşitlenir, dolayısıyla kullanım sıklığı oranları için bir referans üretilir. Uygulamada kullanılması için en azından bir maddenin uygun olmasını sağlamak için, her bir tekrardaki en büyük K_i değerlerine sahip yirmi maddenin K_i değerleri otomatik olarak 1.0'e eşitlenir. Bu maddeler aynı zamanda BBT'nin kapsam gerekliliklerini

göz önünde bulundurmak için seçilmiştir. Bu sayede tam bir BBT uzunluğu için her zaman yeterli sayıda madde bulunur.

Sympson-Hetter yönteminin ikinci fazında, birinci fazın sonunda elde edilen madde kullanım sıklığı kontrol parametreleri BBT örneklemindeki bireylere maddelerin uygulanmasını düzenlemek için kullanılır. Madde seçilir, ama uygulanmadan önce maddenin K_i değeri tek biçimli dağılımdan tesadüfi bir sayı çekilerek değerlendirilir. Eğer bir maddenin K_i değeri tek biçimli dağılımdan seçilen tesadüfi sayının değerini aşarsa madde uygulanır. Eğer aşmazsa seçilen madde mevcut birey için bloke edilir ve diğer madde seçilir (Davis ve Dodd, 2005).

Süreç en son madde uygulanana kadar bu şekilde devam eder. Dikkat edilmesi gereken nokta kullanım sıklığı parametrelerinin hesaplandığı havuza özgü olmasıdır. Yani kullanım sıklığı parametreleri madde havuzunda bir değişiklik olduğunda tekrar hesaplanmalıdır (Han, 2011).

3.3.1.3. Azalarak Kaybolma (AK) Yöntemi (Fade Away Method)

Bugünün bilgisayar teknolojisi, test merkezlerindeki ana bilgisayar sunucuları ve alıcı bilgisayarların; BBT uygulamaları öncesinde, sırasında ve/veya sonrasında madde kullanımını da içeren çeşitli test bilgilerini yeniden şekillendirmek için iletişim kurmasını mümkün hale getirmiştir. Ana sunucuda yer alan tüm madde kullanım bilgisi, alıcı bilgisayarlar tarafından her bir BBT uygulaması sırasında ya da sonrasında çevrimiçi ağ vasıtasıyla düzenli olarak güncellenebilir. Dahası her bir alıcı bilgisayar, bir sonraki test uygulaması başlamadan önce sunucudaki madde kullanım bilgisinin güncellemesini yapabilir. Çevrimiçi teknolojisi BBT sisteminde, tekrarlı simülasyonlar içeren Sympson-Hetter yöntemi gibi yöntemlerdeki madde kullanım sıklığının kestirilmesi ihtiyacını önleyerek, madde kullanım sıklığı kontrolü için madde kullanım sıklığı bilgisinin eş zamanlı olarak kullanılmasına imkân sağlar (Han, 2011).

Bu madde kullanım sıklığı kontrol yönteminde, havuzdaki seçilebilir her bir madde için madde seçme ölçütü güncellenen mevcut kullanım sıklığı oranı ve hedeflenen

kullanım sıklığı oranı arasındaki oran ile ters bir şekilde ağırlıklandırılır. Örneğin aşağıdaki eşitlik Fisher'in En Çok Bilgi yöntemine aittir;

$$I_i[\theta_{m-1}] \frac{U_i}{C}$$

BBT, mutlak madde kullanım limiti (birinci kullanım sıklığı bileşeninin) olan C'yi en çok yapan maddeyi bulmaya çalışmaktadır. U_i i maddesinin uygulanması için madde kullanımı olmuştur. Yeni yöntem ile nadiren kullanılan maddelerin daha sık kullanılması, aşırı derecede kullanılan maddelerin ise madde seçiminde muhtemel olarak "azalarak ortadan kaybolması" beklenir. Bu nedenle bu yöntem Azalarak Kaybolma Yöntemi olarak adlandırılmıştır (Han, 2011).

3.3.2. Madde Seçme Yöntemleri

3.3.2.1. Fisher'in En Çok Bilgi Yöntemi

BBT uygulamalarında, en yaygın kullanılan ve muhtemelen en eski madde seçme yöntemi Fisher'in En Çok Bilgi yöntemidir. Temel olarak bu yöntem $I[\hat{\theta}_{m-1}]$ değerini en çok yapan x maddesinin bulunmasını içerir [$\hat{\theta}$ bireyin geçici yetenek düzeyi (interim proficiency) kestirimi ve $m-1$ o ana kadar uygulanan maddelerin sayısı (Weiss, 1982, Akt. Han, 2011)]. Madde karakteristiği 3PLM ile tanımlanmış olan klasik bir çoktan seçmeli madde havuzu ele alındığında Fisher'in En Çok Bilgi yöntemine dayalı madde seçme yöntemi, aşağıdaki denkleme göre en yüksek bilgi veren maddeyi bulmaya çalışır;

$$I_i[\theta_{m-1}] = \frac{(Da_i)^2(1 - c_i)}{[c_i + e^{Da_i(\theta_{m-1}-b_i)}][1 + e^{-Da_i(\theta_{m-1}-b_i)}]^2}$$

Bu eşitlikte a_i , b_i ve c_i sırasıyla 3 PLM'de ayırt edicilik, güçlük ve pseudo-tahmin parametreleridir ve D 1.702 değerine sahip olan bir ölçekleme sabitidir. BBT uygulamasında her bir birey için test bilgisini en çok yapacak sonuçların elde edilmesinde etkili olduğundan, Fisher'in En Çok Bilgi yöntemi yaygın olarak kullanılmaktadır. Ancak testin başlangıç aşamalarında, örneğin beş veya daha az madde uygulandığında, geçici yetenek düzeyi (interim proficiency) kestirimleri nadir

olarak doğru hesaplanır. Bu nedenle testin başlangıcında Fisher'in En Çok Bilgi yöntemine göre madde seçimi geçici yetenek düzeyi kestirimlerinde çok bilgi sağlama eğiliminde değildir. Bu yönteminin bir diğer sorunu ise a parametresi yüksek olan maddeleri, düşük olan maddelere göre daha çok seçme eğiliminde olmasıdır. Bu nedenle bu yöntemin kullanılması madde havuzunun sürdürülebilirliğinde ciddi bir sorun yaratmaktadır (Han, 2011).

3.3.2.2. a-Tabakalama Yöntemi

Fisher'in En Çok Bilgi yöntemi, bireyin yetenek düzeyinde en çok bilgi veren maddenin seçilmesini amaçlar. Bunun için, $c=0$ olarak alındığında, bireyin gerçek yetenek düzeyine en yakın b parametresi değerine ve yüksek a parametresi değerine sahip maddenin seçilmesi sağlanır. Ancak BBT uygulamasının başlangıcında bireyin gerçek yetenek düzeyi bilinmediğinden, en çok bilginin sağlanabilmesi için a parametresi yüksek olan maddeler seçilir. Madde havuzunda maddelerin a parametresi değerlerine göre düzeylere ayrılmasıyla; düşük a parametresine sahip maddeler testin başlangıç aşamalarında, yüksek a parametresine sahip maddeler testin son aşamalarında uygulanabilir.

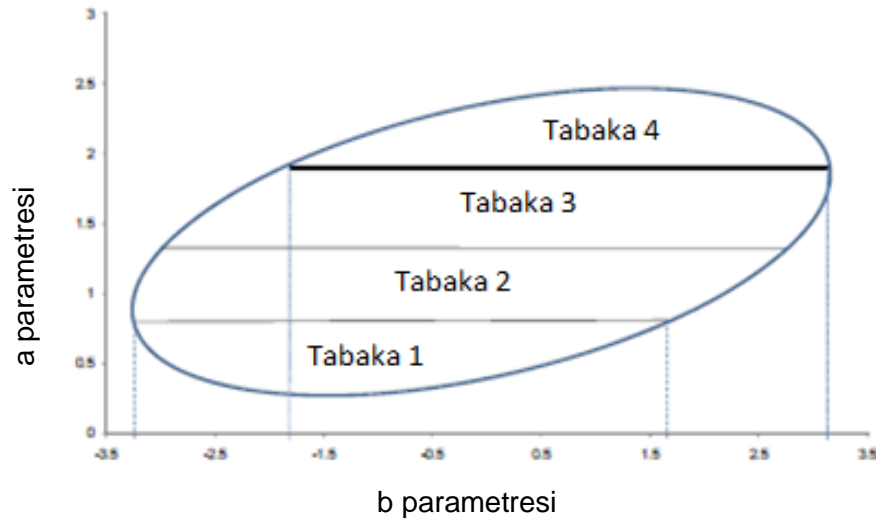
Madde havuzunun tabakalanması aynı zamanda madde kullanım sıklığı oranlarını da etkiler. En çok bilgi madde seçme yöntemi kullanıldığında madde kullanım sıklığı oranlarının düzensiz bir şekilde dağılmasının ana sebebi, düşük a parametresi değerlerine sahip maddeler yerine yüksek a parametresi değerlerine sahip maddelerin seçilmesidir.

a parametresi değerleri benzer olan maddelerin gruplanması ve maddelerin her bir aşamada bir grup içerisinde seçilmesiyle, kullanım sıklığı oranları eşit olarak dağılabilir. Bu yüzden madde havuzunun tabakalanması, hem yüksek a parametrelerine sahip maddelerin kullanım sıklığı oranlarını düşürür hem de düşük a parametrelerine sahip maddelerin kullanılma sıklığı oranlarını yükseltir (Chang ve Ying, 1999).

Chang ve Ying (1999), a-Tabakalama yönteminin aşamalarını basit olarak aşağıdaki gibi özetlemişlerdir;

- 1) Madde havuzu, maddelerin a parametresi değerlerine göre K tane düzeye bölünür,
- 2) Test K aşamaya bölünür,
- 3) K.aşamada b ve yetenek düzeyi arasındaki benzerliğe dayanarak k. düzeyden n_k seçilir ve daha sonra uygulanır ($(n_1 + \dots + n_k)$ test uzunluğuna eşit olduğu unutulmamalı),
- 4) $k = 1, 2, \dots, K$ olacak şekilde 3. basamak tekrar edilir.

Kalinowski (2009) çalışmasında, a-Tabakalama yöntemini aşağıdaki şekille göstermiştir.



Şekil 3.3. a-Tabakalama Yönteminin Şekille Gösterilmesi

Düzeylerin sayısına (K), karar vermede çeşitli faktörler dikkate alınmalıdır. Birinci faktör, düzeylerdeki a değerlerinin değişimidir. Eğer madde havuzları benzer a parametresi değerlerine sahip maddelerden oluşuyorsa, tabakalama etkisi küçülür ve bu yüzden daha az düzey gerekir. Ancak eğer madde bankası tamamen farklı a parametresi değerlerine sahip maddelerden oluşuyorsa, göreceli olarak daha geniş

düzey sayısı gerekmektedir. a parametresi değerlerinin dağılımı ve en uygun tabaka sayısı arasındaki tam ilişkinin araştırılmaya ihtiyacı vardır. İkinci faktör, beklenen yetenek düzeyi değerlerinin ranjıyla eşleşecek b parametrelerinin ranjının iyi olacağı madde havuzunun zenginliğine karar vermektir. Her bir düzeyde, b parametreleri yeteri düzeyde geniş ranjlara sahip olmalıdır. Zengin madde bankaları daha çok düzeye tabakalandırılabilirler. Üçüncü ve dördüncü faktörler test uzunluğu ve madde havuzunun büyüklüğüdür. Eğer madde havuzu yeterli büyüklüğe sahipse, K değeri test uzunluğuna yaklaşabilir (Chang ve Ying, 1999).

Her bir düzeyin büyüklüğünün aynı şekilde belirlenmesi gerekmektedir. Normal bir şekilde düzeylerde uygulanan maddelerin sayısı oransal olmalıdır. Bu durum farklı düzeylerdeki maddelerin kullanım sıklığı oranlarının benzer olacağını garanti eder. Genellikle ilk düzey hariç, diğer tüm düzeylerde madde sayılarının benzer olması gerekir. Bu yüzden eğer L test uzunluğu ise, her bir düzeyde uygulanacak maddelerin sayısı aşağı yukarı L/K olmalıdır (Chang ve Ying, 1999).

a -Tabakalama yöntemi bireylerin yetenek düzeylerinin her bir aşamada uygun maddeler ile yakın olarak eşleşebileceğini varsayar. Diğer bir deyişle, k . aşamada yetenek düzeyine ($\hat{\theta}$) yeteri kadar yakın olan b değerlerine sahip uygun maddeler olmalıdır. Bu durum b dağılımının takabalamadan etkilenmemesini gerektirmektedir. Yani a ve b arasında bir ilişki olmamalıdır. Ancak bu durum uygulamada nadiren ele alınır. Aslında a ve b parametreleri sıklıkla pozitif bir ilişki kestirir (Lord ve Wingersky, 1984, Akt. Chang, Qian ve Ying, 2001).

Chang, Qian ve Ying (2001), bu soruna bir çözüm oluşturmak için a -Tabakalama yöntemine b parametresini de dâhil etmişlerdir. b parametresinin bloke edilmesiyle oluşturulan a -Tabakalama yönteminde temel fikir, her bir tabaka için b 'nin dengelenmiş dağılımının farklı bireylerin yetenek düzeylerine iyi eşleştirme yapılmasını sağlamaktır. Yöntemi önerdikleri çalışmanın sonuçlarına bakıldığında; a ve b parametrelerinin ilişkisiz olduğu ideal bir madde havuzu için a -Tabakalama yönteminin iyi performans gösterdiğinin, ancak a ve b parametreleri arasında ilişki olduğunda problemlerin meydana geldiği görülmüştür (Chang, Qian ve Ying, 2001).

3.3.2.3. Aşamalı En Çok Bilgi Oranı (AEBO)Yöntemi

Fisher'in En Çok Bilgi yönteminin basit, açık/anlaşılır ve etkili olmasına rağmen iki önemli sorunu bulunmaktadır. Bunlar;

1. Madde kullanım sıklığı oranlarını kontrol etmede duyarlı olmaması,
2. Testin başlangıcındaki geçici yetenek düzeylerinin doğru olarak kestirilmesinde etkili olmamasıdır.

Birinci problem çeşitli madde kullanım sıklığı kontrol yöntemleriyle çözülebilir. Ancak ikinci problemin çözümü daha zordur.

BBT uygulamasının başlarında yüksek a parametrelili maddelerin seçilmesini engellemek için geliştirilen a -Tabakalama yöntemi ise madde kullanım sıklığı oranlarını kontrol etmede kullanışlı ve etkili olmasına rağmen; genel test bilgisi Fisher'in En Çok Bilgi madde seçme yöntemine göre daha düşük olma eğilimindedir. Buna ek olarak, tabaka sayısına karar vermede belirsizlikler olabilir.

Bu sebeplerden ötürü Han (2009) Aşamalı En Çok Bilgi Oranı yöntemini geliştirmiştir. Fisher bilgisi, yetenek düzeyi ölçeğindeki belirli bir noktadaki bir maddenin etkililiğinin (item effectiveness) ölçüsü olarak görülebilir. Bir maddenin verimliliği (efficiency of item), belirli bir yetenek düzeyi değerindeki Fisher bilgisinin, yetenek düzeyi ölçeği karşısında en çok bilgiye oranlanmasıyla değerlendirilir. Bu da aşağıdaki eşitlikle verilebilir;

$$\frac{I_x[\hat{\theta}_{m-1}]}{I_x[\theta^*]}$$

θ^* yetenek düzeyi noktasında bilginin maksimum olduğu noktadır. c parametresi sıfıra eşit olduğunda, θ^* b_x 'e eşit olur. Eğer $c_x \neq 0$ olursa θ^* Birnbaum'un çözümüyle kolayca hesaplanabilir;

$$\theta_x^* = b_x + \frac{1}{Da_x} \log \left(\frac{1 + \sqrt{1 + 8c_x}}{2} \right)$$

Han (2009) önerdiği madde seçme yönteminde, BBT uygulamasının erken aşamalarında potansiyel en çok bilgiye (madde verimliliği, item efficiency) ek olarak, geçici bir $\hat{\theta}$ ile beklenen bilginin oranı kullanılmıştır. Yeni yaklaşımda madde etkililiği (item effectiveness) en önemli ölçüt olarak değerlendirilir ve aşağıdaki eşitliği maksimum yapan maddeyi arar;

$$\frac{I_x[\hat{\theta}_{m-1}]}{I_x[\theta^*]} \left(1 - \frac{m}{M}\right) + I_x[\hat{\theta}_{m-1}] \frac{m}{M}$$

Eşitlikteki M test uzunluğunu ve m ise o zamana kadar uygulanan maddelerin sayısının bir fazlasını temsil etmektedir. Eşitliğin ilk kısmı madde yeterliği (item efficiency) terimi ve ikinci kısmı ise Fisher bilgi terimidir. Bu eşitlik Fisher bilgi terimi çarpanı parantezine alınırsa aşağıda gibi de verilebilir (Han, 2009);

$$I_x[\hat{\theta}_{m-1}] \frac{M - m(1 - I_x[\theta^*])}{I_x[\theta^*]M}$$

3.3.3. a-Tabakalama Yönteminin Belirlenmesi

Araştırmada kullanılacak a-Tabakalama yönteminde b parametresinin bloke edilip edilmeyeceğine karar vermek için, a ve b parametreleri arasındaki korelasyonlar hesaplanmış, a ve b parametrelerinin dağılımını gösteren grafikler çizilmiştir.

Madde güçlüğü orta ve yüksek düzeylerde olan madde havuzlarında korelasyon değerleri sırasıyla; -0.00286 ve 0.015086 olarak elde edilmiştir. Buna ek olarak her bir madde havuzunun, b parametrelerinin a parametrelerine göre değişimini gösteren grafikler de EK-2’de verilmiştir.

Hem a ve b parametrelerinin dağılımları hem de bu parametreler arasındaki korelasyon değerleri incelendiğinde parametreler arasında bir ilişki olmadığı görülmüştür. Bu nedenle araştırmada a-Tabakalama yönteminin kullanılmasına karar verilmiştir. a-Tabakalama yönteminin kullanılması sırasında madde havuzları dört tabakaya ayrılmıştır.

3.3.4. Hedeflenen Kullanım Sıklıklarının Hesaplanması

Araştırmada madde kullanım sıklığının kontrol edilmesinde dört yöntem kullanılmıştır. Bu yöntemlerden ikisinde (Simpson-Hetter ve Azalarak Kaybolma yöntemleri) hedeflenen kullanım sıklığı, uygulama öncesinde belirtilmektedir. Hedeflenen kullanım sıklığı, geniş ölçekli testlerde genellikle kullanılan bir değer olan 0.20'ye ayarlanmıştır (Way, 1998).

3.3.4.1. Simpson-Hetter Yöntemi

Simpson-Hetter yönteminin birinci fazında madde kullanım sıklığı kontrol parametrelerinin üretilmesinde tekrarlı simülasyon çalışması yapılmıştır. Her bir alt problem için maddelere ait kullanım sıklığı kontrol parametreleri elde edilmiştir. Kullanım sıklığı kontrol parametrelerinin elde edilme sürecinde, madde havuzunda yer alan her bir maddenin kullanım sıklığı kontrol parametresi 0.20 hedeflenen kullanım sıklığı ölçütüne göre hesaplanmıştır.

3.3.4.2. Azalarak Kaybolma Yöntemi

BBT uygulamasının simüle edilmesinde bir günde 1000 kişinin uygulamaya katıldığı bir durum simüle edilmiştir. Böyle bir durumda da Azalarak Kaybolma yöntemi kullanılırken madde kullanım bilgisinin güncellenmesi gerçek zaman seçeneği tercih edilerek yapılmıştır.

3.4. Verilerin Analizi

BBT simülasyonu sonrası verilerin analizi, hem ölçme kesinliği hem de testin güvenliği açısından yapılmıştır. Bu nedenle araştırma kapsamında yer alan her bir durum için ölçme kesinliğinin değerlendirilmesinde, *uyum (fidelity) katsayısı*, *RMSE (root mean squared error)*, *yanlılık (BIAS)* ve *ortalama mutlak fark (OMF, average absolute difference)*; testin güvenliğinin değerlendirilmesinde de, *madde kullanım sıklığı oranları*, *madde kullanım sıklığı ortalamaları*, *madde kullanım sıklığı standart sapmaları*, *en yüksek kullanım sıklığı oranları*, *kullanım sıklığı oranı 0 olan maddelerin yüzdeleri*, *ölçeklenmiş ki-kare istatistikleri*, *F oranları* ve *test çakışma oranları* hesaplanmıştır.

Test akışma oranının hesaplanması dıřındaki tm hesaplamalar, bir hesap tablosu programı olan Excel’de yapılmıřtır. Test akışma oranları da, Matlab programı kullanılarak hesaplanmıřtır.

lme Kesinliđi İin;

lme kesinliđinin deđerlendirilmesinde uyum, RMSE, yanlılık ve ortalama mutlak fark deđerleri 25 tekrar iin de ayrı ayrı hesaplanmıř olup elde edilen sonuların aritmetik ortalamaları bulgular kısmında verilmiřtir.

Uyum (Fidelity); Simlasyon alıřmalarında tretilen deđerler ile simlasyondan elde edilen kestirimler iin korelasyonlar sıklıkla “uyum (fidelity)” olarak raporlanır (Urry,1970, Vale ve Weiss,1975, Akt. Leung, Chang ve Hau, 2002). Yksek uyum katsayısı, test sonularıyla daha gvenilir kararlar vermek ve daha iyi madde seme yntemini belirlemek iin kullanılır (Leung, Chang ve Hau, 2002). Uyum katsayısını hesaplamak iin *Pearson Momentler arpımı Korelasyonu* kullanılmıřtır.

RMSE;

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{\theta}_i - \theta_i)^2}{n}}$$

Yanlılık;

$$Yanlılık = \frac{\sum_{i=1}^n (\hat{\theta}_i - \theta_i)}{n}$$

Ortalama Mutlak Fark (OMF);

$$OMF = \frac{\sum_{i=1}^n |\hat{\theta}_i - \theta_i|}{n}$$

RMSE, yanlılık ve ortalama mutlak fark deđerlerinin hesaplanmasında kullanılan formllerde yer alan;

$\hat{\theta}_i$ = i bireyinin kestirilen yetenek düzeyini,

θ_i = i bireyinin gerçek yetenek düzeyini,

n = örneklem büyüklüğünü temsil etmektedir.

Test Güvenliği İçin;

Madde Kullanım Sıklığı Oranı; Bir maddenin kullanım sıklığı oranı, maddenin uygulanma sayısının toplam örneklem sayısına bölünmesi ile elde edilir. m bireylerin sayısını gösterdiğinde, j . maddenin gözlenen kullanım sıklığı oranı (ko_j) aşağıdaki gibi hesaplanır;

$$ko_j = \frac{j. maddenin kullanılma sayısı}{m}$$

Araştırma kapsamında da madde havuzlarında yer alan her bir maddenin her bir 25 tekrardaki madde kullanım sıklığı oranları hesaplanmış ve bu değerlerin aritmetik ortalaması alınmıştır. Ortalama madde kullanım sıklığı oranları dağılımı raporlaştırılırken 25 tekrarın hiçbirinde bir kez bile kullanılmamış madde sayısı $ko = 0$ olarak gösterilmiştir. Ortalama kullanım sıklığı oranı 0'dan farklı olan maddeler ise 0.10'luk kategorik gruplar oluşturularak raporlaştırılmıştır. En yüksek madde kullanım sıklığı oranı da, ortalama madde kullanım sıklığı oranları arasındaki en yüksek madde kullanım sıklığı oranı seçilerek belirlenmiştir.

Madde kullanım sıklığı oranlarının ortalaması ve standart sapması için maddelerin kullanım sıklığı oranlarının her bir 25 tekrardaki ortalaması ve standart sapması bulunarak bu değerlerin aritmetik ortalaması verilmiştir.

Bunlara ek olarak madde kullanım sıklığı oranlarının dağılımlarının daha detaylı görülebilmesi için her bir alt problem altında yer alan her bir yöntem için grafikler oluşturulmuştur. Bu grafiklerde her bir maddenin 25 tekrarda hesaplanan madde kullanım sıklığı oranlarının aritmetik ortalama değerleri kullanılmıştır.

Ölçeklenmiş Ki-Kare İstatistiği (Scaled chi-square statistic); Madde kullanım sıklığı oranlarının eşitliğini belirlemek için hedefsel bir dağılım gerekmektedir. İdeal olanı,

tüm maddelerin aynı kullanım sıklığı oranlarına sahip olması ve oran dağılımının tekdüze olmasıdır. Bu yüzden tüm maddeler için beklenen tek biçimli dağılım oranı; N madde havuzunun büyüklüğünü ve L testin uzunluğunu gösterirse, $\overline{ko_j} = L/N$ olarak hesaplanır. Veri frekansını analiz eden Pearson'ın χ^2 istatistiği taklit edilerek (mimicking), gözlenen ve beklenen kullanım sıklığı oranlarının benzerliğini ölçmek için χ^2 aşağıdaki gibi düzenlenmiştir;

$$\chi^2 = \sum_{j=1}^N \frac{(ko_j - \overline{ko_j})^2}{\overline{ko_j}}$$

Bu denklem gözlenen ve beklenen madde kullanım sıklığı oranları arasındaki farklılığı gösterir. Bir madde kullanım sıklığı yönteminin başlangıçtaki amacı madde havuzundaki tüm maddelerin en iyi şekilde kullanılmasını sağlamaktır. Eğer bir yöntem düşük χ^2 değerine sahipse, madde havuzundaki çoğu maddenin kullanıldığı söylenebilir (Chang ve Ying, 1999).

Araştırmada her bir alt problem altında yer alan her bir yöntemin 25 tekrardaki χ^2 değerleri hesaplanmıştır. Bulgular kısmında da elde edilen bu değerlerin aritmetik ortalaması verilmiştir.

F Oranı; İki yöntemin madde kullanım sıklığı oranlarını karşılaştırmak için onların χ^2 ölçülerinin karşılaştırılması gerekmektedir. Bu karşılaştırma ölçüsü aşağıdaki gibi tanımlanır;

$$F_{yöntem1,yöntem2} = \chi_{yöntem1}^2 / \chi_{yöntem2}^2$$

Eğer $F_{yöntem1,yöntem2} < 1$ ise, madde kullanım sıklığı oranlarının genel dengesi açısından Yöntem 1'in Yöntem 2'ye göre daha üstün olduğu sonucuna ulaşılır (Chang ve Ying, 1999). $F_{yöntem1,yöntem2} > 1$ olduğu durumlarda ise, Yöntem 2'nin Yöntem 1'e göre daha üstün olduğu sonucuna ulaşılır.

Araştırmada F oranları hesaplanırken daha iyi sonuç vermesi beklenen madde kullanım sıklığı kontrol yöntemi, işlemin pay kısmına yazılmıştır. Bu nedenle bulgular kısmında verilen ortalama F oranları tablosunun ilk satırı en iyi sonuç

vermesi beklenen yöntemden (AK yöntemi) en kötü sonuç vermesi beklenen yöntem (BKT-5) doğru sıralanmıştır. Tablonun sütunlarında bulunan yöntemler F oranının pay kısmına, satırlarında bulunan yöntemler ise payda kısmına yazılmıştır. Her bir alt problem altında yer alan yöntemlerin birbiriyle karşılaştırılmasında 25 tekrardaki F oranları hesaplanmıştır. Bulgular kısmında da elde edilen bu değerlerin aritmetik ortalaması verilmiştir.

Test Çakışması Oranı (Test Overlap Rate); Madde kullanım sıklığı kontrolü için bir başka önemli özet indeks de test çakışma oranıdır. Bu indeks, tesadüfi olarak seçilen iki birey (m) için çakışan maddelerin beklenen sayısının test uzunluğuna (L) bölünmesiyle bulunur. Bu yüzden çakışma oranı aşağıdaki basamaklar takip edilerek hesaplanır;

1. Bireylerin her bir $m(m - 1)/2$ çifti için ortak maddelerin sayısı hesaplanır,
2. Tüm $m(m - 1)/2$ hesabı toplanır,
3. Toplam sayı $Lm(m - 1)/2$ 'ye bölünür.

İdeal durumda tesadüfi olarak seçilen herhangi iki bireye ait testler arasındaki ortak maddelerin sayısının minimum olması gerekir (Chang ve Ying, 1999). Yüksek madde çakışma oranları, madde kullanım sıklığı oranlarının aşırı derecede çarpık olduğunun bir kanıtıdır. Eğer madde havuzundaki her bir madde eşit bir seçilme olasılığına sahipse, bireyler arasındaki ortak maddelerin sayısı minimum olacaktır (Chang ve Zhang, 2002). Sabit uzunluklu bir test için beklenen test çakışması oranının en alt sınırı, L testin uzunluğu ve N de havuzun büyüklüğü olmak üzere, L/N oranı kadardır (Chang ve Zhang, 2002).

Araştırmada her bir alt problem altında yer alan her bir yöntemin 25 tekrardaki test çakışması oranı hesaplanmıştır. Bulgular kısmında da elde edilen bu değerlerin aritmetik ortalaması verilmiştir.

4. BULGULAR ve TARTIŞMA

Bu bölümde, sırasıyla her bir alt probleme ait bulgular verilmiş ve bulgulara ilişkin yorumlar yapılmıştır.

4.1. Birinci Alt Probleme İlişkin Bulgular

Araştırmanın birinci alt problemi kapsamında; “BBT uygulamasında madde seçme yöntemi *Fisher’ın En Çok Bilgi Yöntemi* ve *madde havuzu orta güçlük düzeyinde* olduğunda ölçme kesinliğinin ve test güvenliğinin; madde kullanım sıklığı kontrolünün yapılmadığı durum ve Bilgi Koşullu Tesadüfi (5 madde), Bilgi Koşullu Tesadüfi (10 madde), Sympson-Hetter ve Azalarak Kaybolma madde kullanım sıklığı kontrol yöntemlerine göre değişimi” incelenmiştir.

Ölçme kesinliğinin değerlendirilmesinde kullanılan uyum katsayısı, RMSE, yanlılık ve OMF’ye ait ortalama değerler; madde kullanım sıklığının kontrol edilmediği durum ve her bir madde kullanım sıklığı kontrolü yöntemi için Tablo 4.1.’de verilmiştir.

Tablo 4.1. Birinci Alt Probleme Ait Ortalama Uyum, RMSE, Yanlılık ve OMF Değerleri*

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>Uyum</i>	<i>RMSE</i>	<i>Yanlılık</i>	<i>OMF</i>
<i>Kontrol Olmadığında</i>	0.9853	0.1748	-0.0026	0.1348
<i>Bilgi Koşullu Tesadüfi (5 Madde)</i>	0.9851	0.1758	-0.0028	0.1356
<i>Bilgi Koşullu Tesadüfi (10 Madde)</i>	0.9845	0.1793	-0.0025	0.1380
<i>Sympson-Hetter</i>	0.9826	0.1900	-0.0037	0.1469
<i>Azalarak Kaybolma</i>	0.9823	0.1918	-0.0013	0.1486

* Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

Elde edilen sonuçlar incelendiğinde; uyum katsayısı değerlerinin tüm durumlar için 0.98 civarında olduğu görülmektedir. RMSE değerleri 0.1748 ve 0.1918 aralığında değişkenlik göstermektedir. Yanlılık değerleri her durumda sıfıra oldukça yakındır. Ortalama mutlak fark indeksine bakıldığında ise, gerçek ve kestirilen yetenek düzeyi değerlerine ait en düşük mutlak fark ortalamasının 0.1348 ile madde kullanım sıklığının kontrol edilmediği durumda, en yüksek mutlak fark

ortalamasının 0.1486 ile madde kullanım sıklığı kontrol yöntemi olarak Azalarak Kaybolma yöntemi kullanıldığında elde edildiği görülmektedir.

Birinci alt probleme ait ölçme kesinliği genel olarak değerlendirildiğinde, madde kullanım sıklığının kontrol edildiği ve kontrol edilmediği durumlar arasında büyük farkların olmadığı görülmektedir.

Test güvenliğinin değerlendirilmesinde kullanılan ortalama madde kullanım sıklığı oranlarının dağılımları, bu oranların ortalaması ve ortalama standart sapması, en yüksek ortalama madde kullanım sıklığı oranı, hiçbir tekrarda uygulanmayan maddelerin yüzdesi, maddelerin kullanımlarına ait ortalama χ^2 değeri ve ortalama test çakışması oranı; madde kullanım sıklığının kontrol edilmediği durum ve madde kullanım sıklığının kontrol edildiği her bir yöntem için Tablo 4.2.'de verilmiştir.

Tablo 4.2. Birinci Alt Probleme Ait Test Güvenliği İçin Kullanılan Göstergeler*

<i>Madde Kullanım Sıklığı Oranı (ko)</i>	<i>KOD</i>	<i>BKT5</i>	<i>BKT10</i>	<i>SH</i>	<i>AK</i>
0.9 < ko ≤ 1.0	1	0	0	0	0
0.8 < ko ≤ 0.9	0	0	0	0	0
0.7 < ko ≤ 0.8	0	0	0	0	0
0.6 < ko ≤ 0.7	1	0	0	0	0
0.5 < ko ≤ 0.6	1	5	0	0	0
0.4 < ko ≤ 0.5	7	6	8	0	0
0.3 < ko ≤ 0.4	14	15	16	0	0
0.2 < ko ≤ 0.3	25	24	21	1	0
0.1 < ko ≤ 0.2	41	36	44	121	134
0.0 < ko ≤ 0.1	117	144	169	128	176
ko = 0	293	270	242	250	190
<i>SS_{ko}</i>	0.109803	0.104997	0.09911	0.075925	0.059845
<i>En Yüksek ko</i>	1.000	0.565	0.469	0.201	0.173
% (ko = 0)	58.6	54.0	48.4	50.0	38.0
χ^2	120.5686	110.2456	98.22934	57.64689	35.81398
<i>Test Çakışma Oranı</i>	0.290	0.270	0.246	0.164	0.121

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

**Madde kullanım sıklığı oranlarının ortalaması koşulların tümünde 0.05 olarak bulunmuştur.

***KOD: Kontrolün Olmadığı Durum, BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympson-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

Tablo 4.2. incelendiğinde; BBT uygulamalarının hiçbir tekrarında bir kez bile kullanılmayan maddelerin sayısının 293 (madde kullanım sıklığının kontrol edilmediği durum) ile 190 (Azalarak Kaybolma yönteminin kullanıldığı durum) arasında değişkenlik gösterdiği görülmektedir. Bu durum madde havuzu büyüklüğünün %58.6'sı ile % 38.0'ine denk düşmektedir. Ayrıca Sympson-Hetter ve Azalarak Kaybolma yöntemlerinde kullanılan hedeflenen kullanım sıklığı oranının 0.20 olduğu göz önüne alınırsa; madde kullanım sıklığının kontrol edilmediği durum ile Bilgi Koşullu Tesadüfi yönteminin kullanıldığı her iki durum için de 0.20 oranını aşan maddelerin mevcut olduğu görülmektedir. Sympson-Hetter yönteminde ise hedeflenen kullanım sıklığı işleme katılmasına rağmen 1 madde 0.20 oranını az da olsa (0.201) aşmıştır.

Madde kullanım sıklığı oranlarına ait dağılımların incelenmesi için standart sapma değerlerine bakıldığında; en yüksek standart sapma değerinin madde kullanım sıklığının kontrol edilmediği duruma ait olduğu görülürken, en düşük standart sapma değerinin Azalarak Kaybolma yönteminin kullanıldığı duruma ait olduğu görülmektedir. Buna göre; madde kullanım sıklığının kontrol edilmediği durum için belirli maddelerin daha sıklıkla kullanıldığı ve bu nedenle de aralarındaki farkın arttığı; Azalarak Kaybolma yöntemi için ise çeşitli maddelerin uygulamalarda kullanıldığı ve bu nedenle de madde kullanım sıklığı oranlarının birbirine yakın olduğu sonucuna varılabilir. Bu sonucu, madde havuzunun kullanımına ait genel bilgiyi veren χ^2 analizi sonucu da desteklemektedir. Bu analiz sonucuna göre; Azalarak Kaybolma yönteminin kullanıldığı durumda madde havuzunun kullanımı ($\chi^2 = 35.81398$), madde kullanım sıklığı kontrolünün yapıldığı diğer tüm yöntemler ve madde kullanım sıklığının kontrol edilmediği duruma göre, ideale en yakın olanıdır. Madde havuzunun kullanımına ait daha detaylı inceleme yapmak için EK-3 bölümünde yer alan her bir durum için ortalama madde kullanım sıklığı oranlarını gösteren grafikler incelenebilir.

En yüksek madde kullanım sıklığı oranı incelendiğinde; madde kullanım sıklığının kontrol edilmediği durumun en yüksek (1.000), Azalarak Kaybolma yönteminin ise en düşük (0.173) değere sahip olduğu görülmektedir.

Test çakışma oranları incelendiğinde; en çok çakışma oranının madde kullanım sıklığının kontrol edilmediği durumda (0.290), en az çakışma oranının da Azalarak Kaybolma yönteminde (0.121) olduğu görülmektedir. BKT yöntemlerinin, test çakışma oranını azaltmakta pek de etkili olmadığı söylenebilir. Çalışmada beklenen test çakışması oranı 0.05'tir (25/500). Fakat hiçbir madde kullanım sıklığı kontrol yöntemi bu değere ulaşamamıştır.

Madde kullanım sıklığı oranlarının karşılaştırılabilmesi için hesaplanan ortalama F oranları madde kullanım sıklığının kontrol edilmediği durum ve her bir madde kullanım sıklığı kontrolü yöntemi için Tablo 4.3.'te verilmiştir.

Tablo 4.3. Birinci Alt Probleme Ait Ortalama F Oranları*

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>AK</i>	<i>SH</i>	<i>BKT10</i>	<i>BKT5</i>
<i>Kontrol Olmadığında</i>	0.30	0.48	0.81	0.91
<i>Bilgi Koşullu Tesadüfi (5 Madde)</i>	0.32	0.52	0.89	
<i>Bilgi Koşullu Tesadüfi (10 Madde)</i>	0.36	0.59		
<i>Sympson-Hetter</i>	0.62			

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

**BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympson-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

Tablo 4.3. incelendiğinde; çarpıklığın görece olarak azalması en çok $F_{AK,KOD}$ 'da olmuştur. $1 - F_{AK,KOD}$ işlemi yapıldığında Azalarak Kaybolma yönteminin, madde kullanım sıklığının kontrol edilmediği duruma göre çarpıklığın azaltılmasındaki oranın 0.70 olduğu görülmektedir. Çarpıklığın görece olarak azalması ise en az $F_{BKT5,KOD}$ oranında olmuştur. Bu değere göre, BKT-5 madde yönteminin, madde kullanım sıklığının kontrol edilmediği duruma göre çarpıklığın azaltılmasındaki oran sadece 0.09 olmuştur.

Birinci alt probleme ait test güvenliği genel olarak değerlendirildiğinde, madde kullanım sıklığının Azalarak Kaybolma yöntemi ile kontrol edildiği durumun diğer durumlara göre iyi sonuçlar verdiği görülmektedir.

Sonuç olarak, Azalarak Kaybolma yöntemi madde havuzunun kullanımı ve madde çakışması açısından diğer madde kullanım sıklığı kontrol yöntemlerine göre daha iyi sonuçlar verirken; uyum, RMSE, yanlılık ve ortalama mutlak fark gibi ölçme kesinliğinin göstergesi olan değerlerden, diğer yöntemlerden çok farklı olmamasıyla birlikte, az da olsa ödün vermiştir.

4.2. İkinci Alt Probleme İlişkin Bulgular

Araştırmanın ikinci alt problemi kapsamında; “BBT uygulamasında madde seçme yöntemi **Fisher’ın En Çok Bilgi Yöntemi** ve **madde havuzu yüksek güçlük düzeyinde** olduğunda ölçme kesinliğinin ve test güvenliğinin; madde kullanım sıklığı kontrolünün yapılmadığı durum ve Bilgi Koşullu Tesadüfi (5 madde), Bilgi Koşullu Tesadüfi (10 madde), Sympson-Hetter ve Azalarak Kaybolma madde kullanım sıklığı kontrol yöntemlerine göre değişimi” incelenmiştir.

Ölçme kesinliğinin değerlendirilmesinde kullanılan uyum katsayısı, RMSE, yanlılık ve OMF’a ait ortalama değerler madde kullanım sıklığının kontrol edilmediği durum ve her bir madde kullanım sıklığı kontrolü yöntemi için Tablo 4.4.’te verilmiştir.

Tablo 4.4. İkinci Alt Probleme Ait Ortalama Uyum, RMSE, Yanlılık ve OMF Değerleri*

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>Uyum</i>	<i>RMSE</i>	<i>Yanlılık</i>	<i>OMF</i>
<i>Kontrol Olmadığında</i>	0.9773	0.2168	-0.0008	0.1627
<i>Bilgi Koşullu Tesadüfi (5 Madde)</i>	0.9769	0.2185	-0.0020	0.1642
<i>Bilgi Koşullu Tesadüfi (10 Madde)</i>	0.9756	0.2247	-0.0010	0.1679
<i>Sympson-Hetter</i>	0.9637	0.2734	0.0007	0.2040
<i>Azalarak Kaybolma</i>	0.9693	0.2518	-0.0009	0.1902

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

Tablo 4.4. incelendiğinde; uyum katsayısı değerleri 0.9773 ile 0.9637 arasında değişkenlik göstermektedir. En yüksek RMSE değerinin 0.2734 (Sympson-Hetter

yöntemi), en düşük RMSE değerinin ise 0.2168 (madde kullanım sıklığı kontrolünün yapılmadığı durum) olduğu görülmektedir. Yanlılık değerleri her durumda sıfıra oldukça yakındır. Ortalama mutlak fark indeksine bakıldığında ise, gerçek ve kestirilen yetenek düzeyi değerlerine ait en düşük mutlak fark ortalamasının 0.1627 ile madde kullanım sıklığının kontrol edilmediği durumda, en yüksek mutlak fark ortalamasının 0.2040 ile madde kullanım sıklığı kontrol yöntemi olarak Sympson-Hetter yöntemi kullanıldığında elde edildiği görülmektedir.

İkinci alt probleme ait ölçme kesinliği genel olarak değerlendirildiğinde, madde kullanım sıklığının kontrol edildiği ve kontrol edilmediği durumlar arasında büyük farkların olmadığı görülmektedir.

Test güvenliğinin değerlendirilmesinde kullanılan ortalama madde kullanım sıklığı oranlarının dağılımları, bu oranların ortalaması ve ortalama standart sapması, en yüksek ortalama madde kullanım sıklığı oranı, hiçbir tekrarda uygulanmayan maddelerin yüzdesi, maddelerin kullanımlarına ait ortalama χ^2 değeri ve ortalama test çakışması oranı; madde kullanım sıklığının kontrol edilmediği durum ve madde kullanım sıklığının kontrol edildiği her bir yöntem için Tablo 4.5.'te verilmiştir.

Tablo 4.5. İkinci Alt Probleme Ait Test Güvenliği İçin Kullanılan Göstergeler*

<i>Madde Kullanım Sıklığı Oranı (ko)</i>	<i>KOD</i>	<i>BKT5</i>	<i>BKT10</i>	<i>SH</i>	<i>AK</i>
$0.9 < ko \leq 1.0$	1	0	0	0	0
$0.8 < ko \leq 0.9$	0	0	0	0	0
$0.7 < ko \leq 0.8$	0	0	0	0	0
$0.6 < ko \leq 0.7$	1	2	0	0	0
$0.5 < ko \leq 0.6$	4	8	6	0	0
$0.4 < ko \leq 0.5$	12	8	10	0	0
$0.3 < ko \leq 0.4$	17	15	19	0	0
$0.2 < ko \leq 0.3$	15	16	15	0	0
$0.1 < ko \leq 0.2$	25	29	29	129	128
$0.0 < ko \leq 0.1$	106	116	132	103	138
ko = 0	319	306	289	268	234
<i>SS_{ko}</i>	0.122773	0.11909	0.114736	0.07824	0.069864
<i>En Yüksek ko</i>	1.000	0.652	0.565	0.197	0.195
<i>% (ko = 0)</i>	63.8	61.2	57.8	53.6	46.8
χ^2	150.7341	141.8257	131.6453	61.2148	48.81045
<i>Test Çakışma Oranı</i>	0.351	0.333	0.313	0.172	0.147

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

**Madde kullanım sıklığı oranlarının ortalaması koşulların tümünde 0.05 olarak bulunmuştur.

***KOD: Kontrolün Olmadığı Durum, BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympson-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

Tablo 4.5.'e bakıldığında; BBT uygulamalarının hiç bir tekrarında bir kez bile kullanılmayan maddelerin sayısının 319 (madde kullanım sıklığının kontrol edilmediği durum) ile 234 (Azalarak Kaybolma yönteminin kullanıldığı durum) arasında değişkenlik gösterdiği görülmektedir. Bu durum madde havuzu büyüklüğünün %63.8'i ile % 46.8'ine denk düşmektedir. Sympson-Hetter ve Azalarak Kaybolma yöntemlerinde kullanılan hedeflenen kullanım sıklığı oranının 0.20 olduğu göz önüne alınırsa; madde kullanım sıklığının kontrol edilmediği durum ile Bilgi Koşullu Tesadüfi yönteminin kullanıldığı her iki durum için de 0.20 oranını aşan maddelerin mevcut olduğu görülmektedir.

Madde kullanım sıklığı oranlarına ait dağılımların incelenmesi için standart sapma değerlerine bakıldığında; en yüksek standart sapma değerinin madde kullanım sıklığının kontrol edilmediği duruma ait olduğu (0.122773) görülürken, en düşük

standart sapma deęerinin Azalarak Kaybolma yönteminin kullanıldığı duruma ait olduğu (0.069864) görölmektedir. Bu sonucu, madde havuzunun kullanımına ait genel bilgiyi veren χ^2 analizi sonuçları incelendiğinde de; Azalarak Kaybolma yönteminin kullanıldığı durumda madde havuzunun kullanımının ($\chi^2 = 48.81045$), madde kullanım sıklığı kontrolünün yapıldığı diğer tüm yöntemler ve madde kullanım sıklığının kontrol edilmediği duruma göre, ideale en yakın yöntem olduğu görölmektedir. Madde havuzunun kullanımına ait daha detaylı inceleme yapmak için EK-3 bölümünde yer alan her bir durum için madde kullanım sıklığı oranlarını gösteren grafikler incelenebilir.

En yüksek madde kullanım sıklığı oranı incelendiğinde; madde kullanım sıklığının kontrol edilmediği durumun en yüksek (1.000), Azalarak Kaybolma yönteminin en düşük (0.195) değere sahip olduğu görölmektedir.

Test çakışma oranları incelendiğinde; en çok çakışma oranının madde kullanım sıklığının kontrol edilmediği durumda (0.351), en az çakışma oranının da Azalarak Kaybolma yönteminde (0.147) olduğu görölmektedir. Bilgi Koşullu Tesadüfi yöntemlerinin, test çakışma oranını azaltmakta pek de etkili olmadığı söylenebilir. Çalışmada beklenen test çakışması oranı 0.05'tir (25/500). Fakat hiçbir madde kullanım sıklığı kontrol yöntemi bu değere ulaşamamıştır.

Madde kullanım sıklığı oranlarının karşılaştırılabilmesi hesaplanan ortalama F oranları madde kullanım sıklığının kontrol edilmediği durum ve her bir madde kullanım sıklığı kontrolü yöntemi için Tablo 4.6.'da verilmiştir.

Tablo 4.6. İkinci Alt Probleme Ait Ortalama F Oranları*

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>AK</i>	<i>SH</i>	<i>BKT10</i>	<i>BKT5</i>
<i>Kontrol Olmadığında</i>	0.32	0.41	0.87	0.94
<i>Bilgi Koşullu Tesadüfi (5 Madde)</i>	0.34	0.43	0.93	
<i>Bilgi Koşullu Tesadüfi (10 Madde)</i>	0.37	0.47		
<i>Sympson-Hetter</i>	0.80			

*Değerlerin tümü 25 tekrardan elde edilen verilen ortalaması hesaplanarak elde edilmiştir.

** BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympson-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

Tablo 4.6.'ya bakıldığında; çarpıklığın görece olarak azalması en çok $F_{AK,KOD}$ 'da olmuştur. $1 - F_{AK,KOD}$ işlemi yapıldığında Azalarak Kaybolma yönteminin, madde kullanım sıklığının kontrol edilmediği duruma göre çarpıklığın azaltılmasındaki oranın 0.68 olduğu görülmektedir. Çarpıklığın görece olarak azalması ise en az $F_{BKT5,KOD}$ oranında olmuştur. Bu değere göre, BKT-5 madde yönteminin, madde kullanım sıklığının kontrol edilmediği duruma göre çarpıklığın azaltılmasındaki oran sadece 0.06 olmuştur.

İkinci alt probleme ait test güvenliği genel olarak değerlendirildiğinde, madde kullanım sıklığı kontrolünün olmadığı durumun ve Bilgi Koşullu Tesadüfi yöntemlerinin (5 ve 10 madde) birbirine çok yakın sonuçlar verdikleri görülmektedir. Azalarak Kaybolma yönteminin ise hem kullanılmayan maddelerin sayısı hem de madde havuzunun kullanımı açısından Sympson-Hetter yöntemine göre daha iyi sonuçlar görülmektedir.

Sonuç olarak, Azalarak Kaybolma yöntemi madde havuzunun kullanımı ve madde çakışması açısından diğer madde kullanım sıklığı kontrol yöntemlerine göre daha iyi sonuçlar verirken; uyum, RMSE ve ortalama mutlak fark gibi ölçme kesinliğinin göstergesi olan değerlerde, çok büyük fark olmasa da, ödün vermiştir. Yanlılık açısından ise bu madde kullanım sıklığı yöntemi ile sifıra en yakın değer elde edilmiştir.

4.3. Üçüncü Alt Probleme İlişkin Bulgular

Araştırmanın üçüncü alt problemi kapsamında; “BBT uygulamasında madde seçme yöntemi **a-Tabakalama Yöntemi** ve **madde havuzu orta güçlük düzeyinde** olduğunda ölçme kesinliğinin ve test güvenliğinin; madde kullanım sıklığı kontrolünün yapılmadığı durum ve Bilgi Koşullu Tesadüfi (5 madde), Bilgi Koşullu Tesadüfi (10 madde), Sympson-Hetter ve Azalarak Kaybolma madde kullanım sıklığı kontrol yöntemlerine göre değişimi” incelenmiştir.

Ölçme kesinliğinin değerlendirilmesinde kullanılan uyum katsayısı, RMSE, yanlılık ve OMF'a ait ortalama değerler madde kullanım sıklığının kontrol edilmediği durum ve her bir madde kullanım sıklığı kontrolü yöntemi için Tablo 4.7.'de verilmiştir.

Tablo 4.7. Üçüncü Alt Probleme Ait Ortalama Uyum, RMSE, Yanlılık ve OMF Değerleri*

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>Uyum</i>	<i>RMSE</i>	<i>Yanlılık</i>	<i>OMF</i>
<i>Kontrol Olmadığında</i>	0.9769	0.2186	-0.0022	0.1696
<i>Bilgi Koşullu Tesadüfi (5 Madde)</i>	0.9765	0.2209	-0.0029	0.1712
<i>Bilgi Koşullu Tesadüfi (10 Madde)</i>	0.9755	0.2250	-0.0019	0.1740
<i>Sympson-Hetter</i>	0.9769	0.2188	-0.0012	0.1698
<i>Azalarak Kaybolma</i>	0.9756	0.2248	-0.0016	0.1748

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

Ölçme kesinliği ile ilgili sonuçlar incelendiğinde; uyum katsayısı değerlerinin 0.97 civarında olduğu görülmektedir. Buna göre en yüksek uyumu, madde kullanım sıklığının kontrol edilmediği durum (0.9769) ile Sympson-Hetter yöntemi (0.9769) göstermiştir. RMSE değerleri 0.2186 ile 0.2250 ranjında değişmektedir. Yanlılık değerleri her durumda sıfıra oldukça yakındır. Ortalama mutlak fark indeksine bakıldığında ise, gerçek ve kestirilen yetenek düzeyi değerlerine ait en düşük mutlak fark ortalamasının 0.1696 ile madde kullanım sıklığının kontrol edilmediği durumda, en yüksek mutlak fark ortalamasının 0.1748 ile madde kullanım sıklığı kontrol yöntemi olarak Azalarak Kaybolma yöntemi kullanıldığında elde edildiği görülmektedir.

Üçüncü alt probleme ait ölçme kesinliği genel olarak değerlendirildiğinde, madde kullanım sıklığının kontrol edildiği durum ile Sympson-Hetter yönteminin kullanıldığı duruma ait değerlerin birbirlerine oldukça yakın oldukları görülmektedir. Diğer durumlarda da, madde kullanım sıklığının kontrol edilmediği durum ile arasında büyük farklar olmamıştır.

Test güvenliğinin değerlendirilmesinde kullanılan ortalama madde kullanım sıklığı oranlarının dağılımları, bu oranların ortalaması ve ortalama standart sapması, en yüksek ortalama madde kullanım sıklığı oranı, hiçbir tekrarda uygulanmayan

maddelerin yüzdesi, maddelerin kullanımlarına ait ortalama χ^2 değeri ve ortalama test çakışması oranı; madde kullanım sıklığının kontrol edilmediği durum ve madde kullanım sıklığının kontrol edildiği her bir yöntem için Tablo 4.8.'de verilmiştir.

Tablo 4.8. Üçüncü Alt Probleme Ait Test Güvenliği İçin Kullanılan Göstergeler*

<i>Madde Kullanım Sıklığı Oranı (ko)</i>	<i>KOD</i>	<i>BKT5</i>	<i>BKT10</i>	<i>SH</i>	<i>AK</i>
$0.9 < ko \leq 1.0$	1	0	0	0	0
$0.8 < ko \leq 0.9$	0	0	0	0	0
$0.7 < ko \leq 0.8$	0	0	0	0	0
$0.6 < ko \leq 0.7$	1	0	0	0	0
$0.5 < ko \leq 0.6$	0	0	0	0	0
$0.4 < ko \leq 0.5$	2	0	0	0	0
$0.3 < ko \leq 0.4$	1	2	0	0	0
$0.2 < ko \leq 0.3$	6	3	9	0	3
$0.1 < ko \leq 0.2$	83	109	108	106	87
$0.0 < ko \leq 0.1$	300	338	341	326	410
$ko = 0$	106	48	42	68	0
SS_{ko}	0.079811	0.055333	0.052863	0.058551	0.048849
<i>En Yüksek ko</i>	1.000	0.304	0.290	0.197	0.295
$\% (ko = 0)$	21.2	9.6	8.4	13.6	0
χ^2	63.69785	30.61927	27.94699	34.28453	23.86452
<i>Test Çakışma Oranı</i>	0.177	0.110	0.105	0.118	0.097

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

**Madde kullanım sıklığı oranlarının ortalaması koşulların tümünde 0.05 olarak bulunmuştur.

***KOD: Kontrolün Olmadığı Durum, BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympon-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

Tablo 4.8. incelendiğinde; BBT uygulamalarının hiç bir tekrarında bir kez bile kullanılmayan maddelerin sayısının en yüksek 106 (madde kullanım sıklığının kontrol edilmediği durum) olduğu görülmektedir. Elde edilen sonuçlara göre; madde kullanım sıklığı kontrolünün yapılmadığı durumda bile bu değer madde havuzu büyüklüğünün %21.2'sine denk düşmektedir. Azalarak Kaybolma yönteminin kullanıldığı durum da ise her madde BBT uygulamalarının en az bir tekrarında en az bir kere kullanılmıştır. Ayrıca Bilgi Koşullu Tesadüfi (5 ve 10

madde) yöntemlerinde hiçbir tekrarda kullanılmayan maddelerin sayısı (sırasıyla 48 ve 42), Sympson-Hetter yönteminde hiçbir tekrarda kullanılmayan madde sayısından (68) daha az olmuştur. Buna ek olarak bu yöntemlerde yüksek madde kullanım sıklığı oranlarının sırasıyla 0.304 ve 0.290 olduğu görülmektedir. Bu değerler ile Sympson-Hetter ve Azalarak Kaybolma yöntemlerinde kullanılan hedeflenen kullanım sıklığı oranına yaklaştıkları görülmektedir. Azalarak Kaybolma yönteminde 0.20 kullanım sıklığı oranı ile işlem yapılmasına karşın bu değeri aştığı görülmektedir (0.295).

Madde kullanım sıklığı oranlarına ait dağılımların incelenmesi için ortalama standart sapma değerlerine bakıldığında; en yüksek standart sapma değerinin madde kullanım sıklığının kontrol edilmediği duruma ait olduğu (0.079811) görülürken, en düşük standart sapma değerinin Azalarak Kaybolma yönteminin kullanıldığı duruma ait olduğu (0.048849) görülmektedir. χ^2 analizi sonucuna göre; en düşük değer Azalarak Kaybolma yöntemine ait olduğu ($\chi^2 = 23.86452$) görülmektedir. Bununla birlikte Bilgi Koşullu Tesadüfi (5 ve 10 madde) yöntemleri ile Sympson-Hetter yöntemine ait χ^2 değerlerinin de hem birbirine hem de Azalarak Kaybolma yöntemine yakın olduğu sonucu elde edilmiştir. Buna göre özellikle BKT 10 madde ve Azalarak Kaybolma yöntemlerinin madde havuzu kullanımındaki çarpıklık birbirine yakındır. Madde havuzunun kullanımına ait daha detaylı inceleme yapmak için EK-3 bölümünde yer alan her bir durum için madde kullanım sıklığı oranlarını gösteren grafikler incelenebilir.

En yüksek madde kullanım sıklığı oranları incelendiğinde; madde kullanım sıklığının kontrol edilmediği durumun en yüksek (1.000), Sympson-Hetter yönteminin ise en düşük (0.197) değere sahip olduğu görülmektedir.

Test çakışma oranları incelendiğinde; en çok çakışma oranının madde kullanım sıklığının kontrol edilmediği duruma (0.177), en az çakışma oranının da Azalarak Kaybolma yöntemine (0.097) ait olduğu görülmektedir. Bilgi Koşullu Tesadüfi yöntemleri, Sympson-Hetter yöntemine göre madde çakışmasını daha iyi kontrol

etmiştir. Çakışma değerlerinin küçük çıkmasına rağmen, hiçbir madde kullanım sıklığı kontrol yöntemi beklenen çakışma oranına (0.05) ulaşamamıştır.

Madde kullanım sıklığı oranlarının karşılaştırılabilmesi hesaplanan ortalama F oranları madde kullanım sıklığının kontrol edilmediği durum ve her bir madde kullanım sıklığı kontrolü yöntemi için Tablo 4.9.'da verilmiştir.

Tablo 4.9. Üçüncü Alt Probleme Ait Ortalama F Oranları*

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>AK</i>	<i>SH</i>	<i>BKT10</i>	<i>BKT5</i>
<i>Kontrol Olmadığında</i>	0.37	0.54	0.44	0.48
<i>Bilgi Koşullu Tesadüfi (5 Madde)</i>	0.78	1.12	0.91	
<i>Bilgi Koşullu Tesadüfi (10 Madde)</i>	0.85	1.23		
<i>Sympson-Hetter</i>	0.70			

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

**BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympson-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

Tablo 4.9. incelendiğinde; çarpıklığın görece olarak azalması en çok $F_{AK,KOD}$ 'da olmuştur. $1 - F_{AK,KOD}$ işlemi yapıldığında Azalarak Kaybolma yönteminin, madde kullanım sıklığının kontrol edilmediği duruma göre çarpıklığın azaltılmasındaki oranın 0.63 olduğu görülmektedir. Çarpıklığın görece olarak azalması ise en az $F_{BKT10,BKT5}$ oranında olmuştur. Bu değere göre, BKT-10 madde yönteminin, BKT-5 madde yöntemine göre çarpıklığın azaltılmasındaki oranı sadece 0.09 olmuştur. Ayrıca Sympson-Hetter yönteminin hem BKT-5 madde hem de BKT-10 madde yöntemlerine göre madde havuzunun kullanımındaki çarpıklığı görece olarak arttırdığı görülmektedir ($F_{SH,BKT5} > 1, F_{SH,BKT10} > 1$).

Üçüncü alt probleme ait test güvenliği genel olarak değerlendirildiğinde, madde kullanım sıklığının kontrol edilmediği durumlar da bile a-Tabakalama yönteminin oldukça iyi sonuçlar verdiği görülmektedir. Madde kullanım sıklığı kontrolü yöntemlerinin analizlere dahil edilmesiyle madde havuzunu en iyi şekilde Azalarak Kaybolma yönteminin kullandığı görülmüştür. Buna ek olarak BKT yöntemlerinin de madde havuzu kullanımının oldukça iyi olduğu görülmektedir. Hedeflenen kullanma sıklığı oranına bakıldığında da bir tek Sympson-Hetter yöntemi madde kullanım

sıklığı oranlarında 0.20 değerini aşmamıştır. Genel olarak madde kullanım sıklığı kontrol yöntemlerinin kullanılmasına bakıldığında; yöntemlerin madde havuzunun kullanımını daha normalleştirip test çakışmasını azaltırken, az da olsa ölçme kesinliğinden ödün verilmesine sebep olduğu görülmektedir.

4.4. Dördüncü Alt Probleme İlişkin Bulgular

Araştırmanın dördüncü alt problemi kapsamında; “BBT uygulamasında madde seçme yöntemi *a-Tabakalama Yöntemi* ve *madde havuzu yüksek güçlük düzeyinde* olduğunda ölçme kesinliğinin ve test güvenliğinin; madde kullanım sıklığı kontrolünün yapılmadığı durum ve Bilgi Koşullu Tesadüfi (5 madde), Bilgi Koşullu Tesadüfi (10 madde), Sympson-Hetter ve Azalarak Kaybolma madde kullanım sıklığı kontrol yöntemlerine göre değişimi” incelenmiştir.

Ölçme kesinliğinin değerlendirilmesinde kullanılan uyum katsayısı, RMSE, yanlılık ve OMF’a ait ortalama değerler madde kullanım sıklığının kontrol edilmediği durum ve her bir madde kullanım sıklığı kontrolü yöntemi için Tablo 4.10’da verilmiştir.

Tablo 4.10. Dördüncü Alt Probleme Ait Ortalama Uyum, RMSE, Yanlılık ve OMF Değerleri*

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>Uyum</i>	<i>RMSE</i>	<i>Yanlılık</i>	<i>OMF</i>
<i>Kontrol Olmadığında</i>	0.9713	0.2436	-0.0010	0.1862
<i>Bilgi Koşullu Tesadüfi (5 Madde)</i>	0.9686	0.2546	-0.0019	0.1930
<i>Bilgi Koşullu Tesadüfi (10 Madde)</i>	0.9645	0.2704	-0.0028	0.2043
<i>Sympson-Hetter</i>	0.9614	0.2816	-0.0044	0.2128
<i>Azalarak Kaybolma</i>	0.9584	0.2919	-0.0019	0.2164

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

Tablo 4.10.’a bakıldığında; uyum katsayısı değerlerinin 0.9713 ile 0.9584 arasında değişkenlik gösterdiği görülmektedir. En düşük uyum katsayısına Azalarak Kaybolma yöntemi sahip olmuştur. En yüksek RMSE değerinin 0.2919 (Azalarak Kaybolma yöntemi), en düşük ortalama RMSE değerinin ise 0.2436 (madde kullanım sıklığı kontrolünün yapılmadığı durumu) olduğu görülmektedir. Ortalama yanlılık değerleri her durumda sıfıra oldukça yakın olmakla birlikte en yüksek değer -0.0044 ile Sympson-Hetter yöntemine aittir. Ortalama mutlak fark indeksine

bakıldığında ise, gerçek ve kestirilen yetenek düzeyi değerlerine ait en düşük mutlak fark ortalamasının 0.1862 ile madde kullanım sıklığının kontrol edilmediği durumda, en yüksek mutlak fark ortalamasının 0.2164 ile madde kullanım sıklığı kontrol yöntemi olarak Azalarak Kaybolma yöntemi kullanıldığında elde edildiği görülmektedir.

Dördüncü alt probleme ait ölçme kesinliği genel olarak değerlendirildiğinde, madde kullanım sıklığının kontrol edildiği ve kontrol edilmediği durumlar arasında büyük farkların olmadığı görülmektedir.

Test güvenliğinin değerlendirilmesinde kullanılan ortalama madde kullanım sıklığı oranlarının dağılımları, bu oranların ortalaması ve ortalama standart sapması, en yüksek ortalama madde kullanım sıklığı oranı, hiçbir tekrarda uygulanmayan maddelerin yüzdesi, maddelerin kullanımlarına ait ortalama χ^2 değeri ve ortalama test çakışması oranı; madde kullanım sıklığının kontrol edilmediği durum ve madde kullanım sıklığının kontrol edildiği her bir yöntem için Tablo 4.11.'de verilmiştir.

Tablo 4.11. Dördüncü Alt Probleme Ait Test Güvenliği İçin Kullanılan Göstergeler*

<i>Madde Kullanım Sıklığı Oranı (ko)</i>	<i>KOD</i>	<i>BKT5</i>	<i>BKT10</i>	<i>SH</i>	<i>AK</i>
$0.9 < ko \leq 1.0$	1	0	0	0	0
$0.8 < ko \leq 0.9$	0	0	0	0	0
$0.7 < ko \leq 0.8$	0	0	0	0	0
$0.6 < ko \leq 0.7$	0	0	0	0	0
$0.5 < ko \leq 0.6$	1	2	0	0	1
$0.4 < ko \leq 0.5$	10	3	0	0	0
$0.3 < ko \leq 0.4$	14	17	15	0	2
$0.2 < ko \leq 0.3$	14	18	35	0	10
$0.1 < ko \leq 0.2$	45	50	45	122	72
$0.0 < ko \leq 0.1$	210	264	267	225	415
<i>ko = 0</i>	205	146	138	153	0
<i>SS_{ko}</i>	0.106896	0.094107	0.088362	0.076282	0.062294
<i>En Yüksek ko</i>	1.000	0.511	0.393	0.199	0.518
<i>% (ko = 0)</i>	41.0	29.2	27.6	30.6	0.0
χ^2	114.2708	88.56513	78.0812	58.19094	38.80823
<i>Test Çakışma Oranı</i>	0.278	0.226	0.205	0.166	0.127

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

**Madde kullanım sıklığı oranlarının ortalaması koşulların tümünde 0.05 olarak bulunmuştur.

***KOD: Kontrolün Olmadığı Durum, BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympson-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

Tablo 4.11.'e bakıldığında; BBT uygulamalarının hiç bir tekrarında bir kez bile kullanılmayan maddelerin sayısının en yüksek 205 (madde kullanım sıklığının kontrol edilmediği durum) olduğu, Azalarak Kaybolma yönteminin kullanıldığı durumda ise her maddenin uygulamaların en az bir tekrarında kullanıldığı görülmektedir. Diğer yöntemler de ise uygulamaların hiç birinde kullanılmayan maddelerin sayısı 100'ün üstündedir. Sympson-Hetter ve Azalarak Kaybolma yöntemlerinde kullanılan hedeflenen kullanım sıklığı oranının 0.20 olduğu göz önüne alınırsa; madde kullanım sıklığının kontrol edilmediği durum ile Bilgi Koşullu Tesadüfi yönteminin kullanıldığı her iki durum için de 0.20 oranını aşan maddelerin mevcut olduğu görülmektedir. Bunun yanı sıra Azalarak Kaybolma yönteminde kullanım sıklığı oranları için bir hedef değer olmasına rağmen bu değer aşılmıştır (0.518).

Madde kullanım sıklığı oranlarına ait dağılımların incelenmesi için standart sapma değerlerine bakıldığında; en yüksek standart sapma değerinin madde kullanım sıklığının kontrol edilmediği duruma ait olduğu (0.106896) görülürken, en düşük standart sapma değerinin Azalarak Kaybolma yönteminin kullanıldığı duruma ait olduğu (0.062294) görülmektedir. χ^2 analizi sonucuna göre; en düşük değerin Azalarak Kaybolma yöntemine ait olduğu ($\chi^2 = 38.80823$) görülmektedir. Sympson-Hetter yönteminin madde havuzu kullanımında, Bilgi Koşullu Tesadüfi (5 ve 10 madde) yöntemlerine göre daha iyi olduğu görülmekteyken, uygulamaların hiçbirinde kullanılmayan madde sayılarına bakıldığında daha çok maddenin hiç kullanılmadığı görülmektedir. Madde havuzunun kullanımına ait daha detaylı inceleme yapmak için EK-3 bölümünde yer alan her bir durum için madde kullanım sıklığı oranlarını gösteren grafikler incelenebilir.

En yüksek madde kullanım sıklığı oranları incelendiğinde; madde kullanım sıklığının kontrol edilmediği durumunun en yüksek (1.000), Sympson-Hetter yönteminin ise en düşük (0.199) değere sahip olduğu görülmektedir.

Test çakışma oranları incelendiğinde; en çok çakışma oranının madde kullanım sıklığının kontrol edilmediği duruma (0.278), en az çakışma oranının da Azalarak Kaybolma yönteminde (0.127) olduğu görülmektedir. Beklenen test çakışması oranına ($25/500=0.05$) en çok Azalarak Kaybolma yöntemi yaklaşmıştır.

Madde kullanım sıklığı oranlarının karşılaştırılabilmesi hesaplanan ortalama F oranları madde kullanım sıklığının kontrol edilmediği durum ve her bir madde kullanım sıklığı kontrolü yöntemi için Tablo 4.12.'de verilmiştir.

Tablo 4.12. Dördüncü Alt Probleme Ait Ortalama F Oranları*

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>AK</i>	<i>SH</i>	<i>BKT10</i>	<i>BKT5</i>
<i>Kontrol Olmadığında</i>	0.34	0.51	0.68	0.78
<i>Bilgi Koşullu Tesadüfi (5 Madde)</i>	0.44	0.66	0.88	
<i>Bilgi Koşullu Tesadüfi (10 Madde)</i>	0.50	0.75		
<i>Sympson-Hetter</i>	0.67			

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

** BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympson-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntem

Tablo 4.12. incelendiğinde; çarpıklığın görece olarak azalması en çok $F_{AK,KOD}$ 'da olmuştur. $1 - F_{AK,KOD}$ işlemi yapıldığında Azalarak Kaybolma yönteminde, madde kullanım sıklığının kontrol edilmediği duruma göre çarpıklığın azaltılmasındaki oranın 0.66 olduğu görülmektedir. Çarpıklığın görece olarak azalması ise en az $F_{BKT10,BKT5}$ oranında olmuştur. Bu değere göre, BKT-10 madde yönteminin, BKT-5 madde yöntemine göre çarpıklığın azaltılmasındaki oranı sadece 0.12 olmuştur.

Dördüncü alt probleme ait test güvenliği genel olarak değerlendirildiğinde, uygulanan gruptan farklı özelliklere sahip olan bir madde havuzu kullanılmasına rağmen a-Tabakalama yönteminin madde kullanım sıklığı kontrol yöntemlerini kullanmadığında bile oldukça iyi sonuçlar verdiği görülmektedir. Madde kullanım sıklığı kontrolü yöntemlerinin analizlere dahil edilmesiyle madde havuzunu en iyi Azalarak Kaybolma yönteminin kullandığı görülmüştür. Fakat Azalarak Kaybolma yönteminin a-Tabakalama yöntemi ile kullanımında hedeflenen kullanma sıklığını ölçüt almakta başarısız olduğu görülmektedir. BKT-10 madde yönteminin ise madde havuzu kullanımının iyi olduğu görülmektedir. Hedeflenen kullanma sıklığı oranına bakıldığında da bir tek Sympon-Hetter yöntemi madde kullanım sıklığı oranlarında 0.20 değerini aşmamıştır.

Genel olarak bakıldığında madde kullanım sıklığı kontrol yöntemlerinin kullanımı madde havuzunun kullanımını ve test çakışmasını düzeltirken, çok büyük fark olmasa da, ölçme kesinliğinden ödün verilmesine sebep olmuştur.

4.5. Beşinci Alt Probleme İlişkin Bulgular

Araştırmanın beşinci alt problemi kapsamında; “BBT uygulamasında madde seçme yöntemi **Aşamalı En Çok Bilgi Oranı** ve **madde havuzu orta güçlük düzeyinde** olduğunda ölçmenin kesinliğinin ve test güvenliğinin; madde kullanım sıklığı kontrolünün yapılmadığı durum ve Bilgi Koşullu Tesadüfi (5 madde), Bilgi Koşullu Tesadüfi (10 madde), Sympon-Hetter ve Azalarak Kaybolma madde kullanım sıklığı kontrol yöntemlerine göre değişimi” incelenmiştir.

Ölçme kesinliğinin değerlendirilmesinde kullanılan uyum katsayısı, RMSE, yanlılık ve OMF'a ait ortalama değerler madde kullanım sıklığının kontrol edilmediği durum ve her bir madde kullanım sıklığı kontrolü yöntemi için Tablo 4.13.'te verilmiştir.

Tablo 4.13. Beşinci Alt Probleme Ait Ortalama Uyum, RMSE, Yanlılık ve OMF Değerleri*

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>Uyum</i>	<i>RMSE</i>	<i>Yanlılık</i>	<i>OMF</i>
<i>Kontrol Olmadığında</i>	0.9854	0.1744	-0.0024	0.1346
<i>Bilgi Koşullu Tesadüfi (5 Madde)</i>	0.9851	0.1763	-0.0034	0.1354
<i>Bilgi Koşullu Tesadüfi (10 Madde)</i>	0.9845	0.1797	-0.0028	0.1381
<i>Sympson-Hetter</i>	0.9830	0.1880	-0.0023	0.1453
<i>Azalarak Kaybolma</i>	0.9823	0.1919	-0.0029	0.1486

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

Tablo 4.13. incelendiğinde; uyum katsayısı değerlerinin 0.98 civarında olduğu görülmektedir. En yüksek RMSE değerinin 0.1919 (Azalarak Kaybolma yöntemi), en düşük ortalama RMSE değerinin ise 0.1744 (madde kullanım sıklığı kontrolünün yapılmadığı durumu) olmuştur. Ortalama yanlılık değerleri her durumda sıfıra oldukça yakın olmakla birlikte en yüksek değer -0.0034 ile BKT-5 madde yöntemine aittir. Ortalama mutlak fark indeksine bakıldığında ise, gerçek ve kestirilen yetenek düzeyi değerlerine ait en düşük mutlak fark ortalamasının 0.1346 ile madde kullanım sıklığının kontrol edilmediği durumda, en yüksek mutlak fark ortalamasının 0.1486 ile madde kullanım sıklığı kontrol yöntemi olarak Azalarak Kaybolma yöntemi kullanıldığında elde edildiği görülmektedir. Beşinci alt probleme ait ölçme kesinliği genel olarak değerlendirildiğinde, madde kullanım sıklığının kontrol edildiği ve kontrol edilmediği durumlar arasında büyük farkların olmadığı görülmektedir.

Test güvenliğinin değerlendirilmesinde kullanılan ortalama madde kullanım sıklığı oranlarının dağılımları, bu oranların ortalaması ve ortalama standart sapması, en yüksek ortalama madde kullanım sıklığı oranı, hiçbir tekrarda uygulanmayan maddelerin yüzdesi, maddelerin kullanımlarına ait ortalama χ^2 değeri ve ortalama

test çakışması oranı; madde kullanım sıklığının kontrol edilmediği durum ve madde kullanım sıklığının kontrol edildiği her bir yöntem için Tablo 4.14.'te verilmiştir.

Tablo 4.14. Beşinci Alt Probleme Ait Test Güvenliği İçin Kullanılan Göstergeler*

<i>Madde Kullanım Sıklığı Oranı (ko)</i>	<i>KOD</i>	<i>BKT5</i>	<i>BKT10</i>	<i>SH</i>	<i>AK</i>
$0.9 < ko \leq 1.0$	1	0	0	0	0
$0.8 < ko \leq 0.9$	0	0	0	0	0
$0.7 < ko \leq 0.8$	0	0	0	0	0
$0.6 < ko \leq 0.7$	1	0	0	0	0
$0.5 < ko \leq 0.6$	1	1	0	0	0
$0.4 < ko \leq 0.5$	7	8	3	0	0
$0.3 < ko \leq 0.4$	16	18	20	0	0
$0.2 < ko \leq 0.3$	23	22	27	0	0
$0.1 < ko \leq 0.2$	40	39	44	122	133
$0.0 < ko \leq 0.1$	119	142	165	135	199
ko = 0	292	270	241	243	168
<i>SS_{ko}</i>	0.109442	0.10224	0.09419	0.07544	0.059324
<i>En Yüksek ko</i>	1.000	0.509	0.438	0.199	0.174
<i>% (ko = 0)</i>	58.4	54.0	48.2	48.6	33.6
χ^2	119.7769	104.533	88.72043	56.9125	35.19322
<i>Test Çakışma Oranı</i>	0.289	0.258	0.227	0.163	0.120

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

**Madde kullanım sıklığı oranlarının ortalaması koşulların tümünde 0.05 olarak bulunmuştur.

***KOD: Kontrolün Olmadığı Durum, BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympon-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

Tablo 4.14. incelendiğinde; BBT uygulamalarının hiç bir tekrarında bir kez bile kullanılmayan maddelerin sayısının en yüksek 292 (madde kullanım sıklığının kontrol edilmediği durum), en düşük ise 168 (Azalarak Kaybolma yöntemi) olduğu görülmektedir. Bu durum madde havuzu büyüklüğünün %58.4'ü ile % 33.6'sına denk gelmektedir. Azalarak Kaybolma yöntemi dışındaki diğer durumlarda BBT uygulamasının hiç bir tekrarında hiç birinde kullanılmayan maddelerin sayısı 200'ün üstündedir. Sympon-Hetter ve Azalarak Kaybolma yöntemlerinde kullanılan hedeflenen kullanım sıklığı oranının 0.20 olduğu göz önüne alınırsa; madde kullanım sıklığının kontrol edilmediği durum ile Bilgi Koşullu Tesadüfi

yönteminin kullanıldığı her iki durum için de 0.20 oranını aşan maddelerin mevcut olduğu görülmektedir.

Madde kullanım sıklığı oranlarına ait dağılımların incelenmesi için standart sapma değerlerine bakıldığında; en yüksek standart sapma değerinin madde kullanım sıklığının kontrol edilmediği duruma ait olduğu (0.109442) görülürken, en düşük standart sapma değerinin Azalarak Kaybolma yönteminin kullanıldığı duruma ait olduğu (0.059324) görülmektedir. χ^2 analizi sonucuna göre de en düşük değer Azalarak Kaybolma yöntemine ait olduğu ($\chi^2 = 35.19322$) görülmektedir. Madde kullanım sıklığı kontrol edilmediğinde de χ^2 değeri en yüksek değerini ($\chi^2 = 119.7769$) almaktadır. Bu değer neredeyse Azalarak Kaybolma yöntemiyle elde edilen χ^2 değerinin 3 katıdır. Yani Aşamalı En Çok Bilgi Oranı madde seçme yöntemi herhangi bir madde kullanım sıklığı kontrolü yöntemi kullanılmadan uygulamaya alındığında madde havuzunun kullanımında oldukça çarpık bir dağılım elde edilmiştir. Madde havuzunun kullanımına ait daha detaylı inceleme yapmak için EK-3 bölümünde yer alan her bir durum için madde kullanım sıklığı oranlarını gösteren grafikler incelenebilir.

En yüksek madde kullanım sıklığı oranları incelendiğinde; madde kullanım sıklığının kontrol edilmediği durumunun en yüksek (1.000), Azalarak Kaybolma yönteminin ise en düşük (0.174) değere sahip olduğu görülmektedir.

Test çakışma oranları incelendiğinde; en çok çakışma oranının madde kullanım sıklığının kontrol edilmediği duruma (0.289), en az çakışma oranının da Azalarak Kaybolma yönteminde (0.120) olduğu görülmektedir. Bilgi Koşullu Tesadüfi yöntemlerinin, test çakışma oranını azaltmakta pek de etkili olmadığı söylenebilir. Çalışmada beklenen test çakışması oranı 0.05'tir (25/500). Fakat hiçbir madde kullanım sıklığı kontrol yöntemi bu değere ulaşamamıştır.

Madde kullanım sıklığı oranlarının karşılaştırılabilmesi için hesaplanan ortalama F oranları, madde kullanım sıklığının kontrol edilmediği durum ve her bir madde kullanım sıklığı kontrolü yöntemi için Tablo 4.15.'te verilmiştir.

Tablo 4.15. Beşinci Alt Probleme Ait Ortalama F Oranları*

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>AK</i>	<i>SH</i>	<i>BKT10</i>	<i>BKT5</i>
<i>Kontrol Olmadığında</i>	0.29	0.48	0.74	0.87
<i>Bilgi Koşullu Tesadüfi (5 Madde)</i>	0.34	0.54	0.85	
<i>Bilgi Koşullu Tesadüfi (10 Madde)</i>	0.40	0.64		
<i>Sympson-Hetter</i>	0.62			

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

**BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympson-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

Tablo 4.15. incelendiğinde; çarpıklığın görece olarak azalması en çok $F_{AK,KOD}$ 'da olmuştur. $1 - F_{AK,KOD}$ işlemi yapıldığında Azalarak Kaybolma yönteminin, madde kullanım sıklığının kontrol edilmediği duruma göre çarpıklığın azaltılmasındaki oranın 0.71 olduğu görülmektedir. Çarpıklığın görece olarak azalması ise en az $F_{BTK5,KOD}$ oranında olmuştur. Bu değere göre, BKT-5 madde yönteminin, madde kullanım sıklığının kontrol edilmediği duruma göre çarpıklığın azaltılmasındaki oran sadece 0.13 olmuştur.

Beşinci alt probleme ait test güvenliği genel olarak değerlendirildiğinde, Azalarak Kaybolma yönteminin hem madde havuzunun kullanımı açısından hem de en yüksek madde kullanma sıklığı oranı açısından diğer yöntemlere göre daha iyi sonuçlar verdiği görülmüştür.

Sonuç olarak, Azalarak Kaybolma yöntemi madde havuzunun kullanımı ve madde çakışması açısından diğer madde kullanım sıklığı kontrol yöntemlerine göre daha iyi sonuçlar verirken; uyum, RMSE, yanlılık ve ortalama mutlak fark gibi ölçme kesinliğinin göstergesi olan değerlerden, diğer yöntemlerden çok farklı olmamasıyla birlikte, az da olsa ödün vermiştir.

4.6. Altıncı Alt Probleme İlişkin Bulgular

Araştırmanın altıncı alt problemi kapsamında; “BBT uygulamasında madde seçme yöntemi **Aşamalı En Çok Bilgi Oranı** ve **madde havuzu yüksek güçlük düzeyinde** olduğunda ölçmenin kesinliğinin ve test güvenliğinin; madde kullanım

sıklığı kontrolünün yapılmadığı durum ve Bilgi Koşullu Tesadüfi (5 madde), Bilgi Koşullu Tesadüfi (10 madde), Sympson-Hetter ve Azalarak Kaybolma madde kullanım sıklığı kontrol yöntemlerine göre değişimi” incelenmiştir.

Ölçme kesinliğinin değerlendirilmesinde kullanılan uyum katsayısı, RMSE, yanlılık ve OMF’a ait ortalama değerler madde kullanım sıklığının kontrol edilmediği durum ve her bir madde kullanım sıklığı kontrolü yöntemi için Tablo 4.16.’da verilmiştir.

Tablo 4.16. Altıncı Alt Probleme Ait Ortalama Uyum, RMSE, Yanlılık ve OMF Değerleri *

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>Uyum</i>	<i>RMSE</i>	<i>Yanlılık</i>	<i>OMF</i>
<i>Kontrol Olmadığında</i>	0.9775	0.2161	-0.0013	0.1623
<i>Bilgi Koşullu Tesadüfi (5 Madde)</i>	0.9766	0.2202	-0.0028	0.1653
<i>Bilgi Koşullu Tesadüfi (10 Madde)</i>	0.9757	0.2243	-0.0019	0.1684
<i>Sympson-Hetter</i>	0.9649	0.2688	-0.0005	0.2006
<i>Azalarak Kaybolma</i>	0.9689	0.2533	-0.0018	0.1904

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

Tablo 4.16. incelendiğinde; uyum katsayısı değerleri 0.9775 ile 0.9649 arasında değişkenlik göstermektedir. En yüksek RMSE değerinin 0.2688 (Sympson-Hetter yöntemi), en düşük ortalama RMSE değerinin ise 0.2161 (madde kullanım sıklığı kontrolünün yapılmadığı durumu) olduğu görülmektedir. Yanlılık değerleri her durumda sıfıra oldukça yakındır. Ortalama mutlak fark indeksine bakıldığında ise, gerçek ve kestirilen theta değerlerine ait en düşük mutlak fark ortalamasının 0.1623 ile madde kullanım sıklığının kontrol edilmediği durumda, en yüksek mutlak fark ortalamasının 0.2006 ile madde kullanım sıklığı kontrol yöntemi olarak Sympson-Hetter yöntemi kullanıldığında elde edildiği görülmektedir.

Altıncı alt probleme ait ölçme kesinliği genel olarak değerlendirildiğinde, madde kullanım sıklığının kontrol edildiği ve kontrol edilmediği durumlar arasında büyük farkların olmadığı görülmektedir.

Test güvenliğinin değerlendirilmesinde kullanılan ortalama madde kullanım sıklığı oranlarının dağılımları, bu oranların ortalaması ve ortalama standart sapması, en

yüksek ortalama madde kullanım sıklığı oranı, hiçbir tekrarda uygulanmayan maddelerin yüzdesi, maddelerin kullanımlarına ait ortalama χ^2 değeri ve ortalama test çakışması oranı; madde kullanım sıklığının kontrol edilmediği durum ve madde kullanım sıklığının kontrol edildiği her bir yöntem için Tablo 4.17.'de verilmiştir.

Tablo 4.17. Altıncı Alt Probleme Ait Test Güvenliği İçin Kullanılan Göstergeler*

<i>Madde Kullanım Sıklığı Oranı (ko)</i>	<i>KOD</i>	<i>BKT5</i>	<i>BKT10</i>	<i>SH</i>	<i>AK</i>
$0.9 < ko \leq 1.0$	1	0	0	0	0
$0.8 < ko \leq 0.9$	0	0	0	0	0
$0.7 < ko \leq 0.8$	0	0	0	0	0
$0.6 < ko \leq 0.7$	1	0	0	0	0
$0.5 < ko \leq 0.6$	3	4	0	0	0
$0.4 < ko \leq 0.5$	13	15	14	0	0
$0.3 < ko \leq 0.4$	16	12	23	0	0
$0.2 < ko \leq 0.3$	19	23	19	0	0
$0.1 < ko \leq 0.2$	24	25	25	128	124
$0.0 < ko \leq 0.1$	102	117	131	99	142
ko = 0	321	304	288	273	234
SS_{ko}	0.121133	0.115063	0.108593	0.078516	0.070344
En Yüksek ko	1.000	0.565	0.475	0.198	0.196
% (ko = 0)	64.2	60.8	57.6	54.6	46.8
χ^2	146.7332	132.3965	117.9268	61.64717	49.48222
Test Çakışma Oranı	0.343	0.314	0.285	0.172	0.148

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

**Madde kullanım sıklığı oranlarının ortalaması koşulların tümünde 0.05 olarak bulunmuştur.

***KOD: Kontrolün Olmadığı Durum, BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympon-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

Tablo 4.17.'ye bakıldığında; BBT uygulamalarının hiç bir tekrarında bir kez bile kullanılmayan maddelerin sayısının 321 (madde kullanım sıklığının kontrol edilmediği durum) ile 234 (Azalarak Kaybolma yönteminin kullanıldığı durum) arasında değişkenlik gösterdiği görülmektedir. Bu durum madde havuzu büyüklüğünün %64.2'si ile %46.8'ine denk düşmektedir. Sympon-Hetter ve Azalarak Kaybolma yöntemlerinde kullanılan hedeflenen kullanım sıklığı oranının 0.20 olduğu göz önüne alınırsa; madde kullanım sıklığının kontrol edilmediği durum

ile Bilgi Koşullu Tesadüfi yönteminin kullanıldığı her iki durum için de 0.20 oranını aşan maddelerin mevcut olduğu görülmektedir.

Madde kullanım sıklığı oranlarına ait dağılımların incelenmesi için standart sapma değerlerine bakıldığında; en yüksek standart sapma değerinin madde kullanım sıklığının kontrol edilmediği duruma ait olduğu (0.121133) görülürken, en düşük standart sapma değerinin Azalarak Kaybolma yönteminin kullanıldığı duruma ait olduğu (0.070344) görülmektedir. Bu sonucu, madde havuzunun kullanımına ait genel bilgiyi veren χ^2 analizi sonuçları incelendiğinde de; Azalarak Kaybolma yönteminin kullanıldığı durumda madde havuzunun kullanımının ($\chi^2 = 49.48222$), madde kullanım sıklığı kontrolünün yapıldığı diğer tüm yöntemler ve madde kullanım sıklığının kontrol edilmediği duruma göre, ideale en yakın yöntem olduğu görülmektedir. Madde havuzunun kullanımına ait daha detaylı inceleme yapmak için EK-3 bölümünde yer alan her bir durum için madde kullanım sıklığı oranlarını gösteren grafikler incelenebilir.

En yüksek madde kullanım sıklığı oranı incelendiğinde; madde kullanım sıklığının kontrol edilmediği durumunun en yüksek (1.000), Azalarak Kaybolma yönteminin en düşük (0.196) değere sahip olduğu görülmektedir.

Test çakışma oranları incelendiğinde; en çok çakışma oranının madde kullanım sıklığının kontrol edilmediği duruma (0.343), en az çakışma oranının da Azalarak Kaybolma yönteminde (0.148) olduğu görülmektedir. BKT-5 madde yönteminin, test çakışma oranını azaltmakta pek de etkili olmadığı söylenebilir. Hiçbir madde kullanım sıklığı kontrol yöntemi beklenen test çakışması oranına ($25/500=0.05$) ulaşamazken, Azalarak Kaybolma yöntemi bu değere en çok yaklaşan yöntem olmuştur.

Madde kullanım sıklığı oranlarının karşılaştırılabilmesi hesaplanan ortalama F oranları madde kullanım sıklığının kontrol edilmediği durum ve her bir madde kullanım sıklığı kontrolü yöntemi için Tablo 4.18.'te verilmiştir.

Tablo 4.18. Altıncı Alt Probleme Ait Ortalama F Oranları*

<i>Madde Kullanım Sıklığı Yöntemleri</i>	<i>AK</i>	<i>SH</i>	<i>BKT10</i>	<i>BKT5</i>
<i>Kontrol Olmadığında</i>	0.34	0.42	0.80	0.90
<i>Bilgi Koşullu Tesadüfi (5 Madde)</i>	0.37	0.47	0.89	
<i>Bilgi Koşullu Tesadüfi (10 Madde)</i>	0.42	0.52		
<i>Sympson-Hetter</i>	0.80			

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

**BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympson-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

Tablo 4.18. incelendiğinde; çarpıklığın görece olarak azalması en çok $F_{AK,KOD}$ 'da olmuştur. $1 - F_{AK,KOD}$ işlemi yapıldığında Azalarak Kaybolma yönteminin, madde kullanım sıklığının kontrol edilmediği duruma göre çarpıklığın azaltılmasındaki oranın 0.66 olduğu görülmektedir. Çarpıklığın görece olarak azalması ise en az $F_{BKT5,KOD}$ oranında olmuştur. Bu değere göre, BKT-5 madde yönteminin, madde kullanım sıklığının kontrol edilmediği duruma göre çarpıklığın azaltılmasındaki oranın sadece 0.10 olmuştur.

Altıncı alt probleme ait test güvenliği genel olarak değerlendirildiğinde, madde kullanım sıklığı kontrolünün olmadığı durum ile Bilgi Koşullu Tesadüfi yöntemlerinin (5 ve 10 madde) birbirine çok yakın sonuçlar verdikleri görülmektedir. Azalarak Kaybolma yönteminin ise hem kullanılmayan maddelerin sayısı hem de madde havuzunun kullanımı açısından Sympson-Hetter yöntemine göre daha iyi sonuçlar verdiği görülmektedir.

Sonuç olarak, Azalarak Kaybolma yöntemi madde havuzunun kullanımı ve test çakışması açısından diğer madde kullanım sıklığı kontrol yöntemlerine göre daha iyi sonuçlar verirken; uyum, RMSE ve ortalama mutlak fark gibi ölçme kesinliğinin göstergesi olan değerlerde çok büyük fark olmasa da ödün vermiştir. Yanlılık açısından ise bu madde kullanım sıklığı yöntemi ile sıfıra en yakın üçüncü değer elde edilmiştir. Uyum, RMSE ve ortalama mutlak fark gibi ölçme kesinliğinin göstergesi olan değerlerde en düşük değerlere Sympson-Hetter yöntemi sahip olmuştur. Yanlılık açısından ise sıfıra en yakın değeri elde etmiştir.

4.7. Tartışma Ve Yorum

Araştırmanın bulgularına genel olarak bakıldığında farklı ortalama güçlük düzeylerine sahip olan madde havuzlarında madde kullanım sıklığı kontrol yöntemlerinin birbirine göre göstermiş olduğu performansların benzer oldukları görülmektedir. Aralarındaki en önemli fark yüksek güçlükteki madde havuzunda $k_0 = 0$ olan maddelerin sayısının fazla olmasıdır. Ayrıca ölçme kesinliğinin göstergesi olan değerlerin orta güçlükteki madde havuzlarında daha iyi oldukları görülmüştür. Bu sonuç da beklenen bir sonuç olmuştur. Çünkü uygulamada $N(0,1)$ normal dağılımdan türetilen bir örneklem kullanılmıştır. Bu nedenle ortalama güçlüğü orta düzeyde olan madde havuzu örnekleme daha çok hitap etmiştir. Ayrıca yüksek güçlük düzeyine sahip madde havuzu incelendiğinde; kullanılmayan maddelerin genellikle zor maddeler olduğu görülmüştür (Kullanılmayan maddelerin b parametresi ortalamaları 2.50 ile 3.55 arasında değişmiştir).

Literatüre bakıldığında madde kullanım sıklığı kontrol yöntemlerinin farklı özelliklere sahip olan 1-0 puanlanan maddelerin yer aldığı madde havuzlarındaki performanslarının değişimi ile ilgili bir araştırmaya rastlanmamıştır. Çoklu puanlanan maddelerin yer aldığı, farklı ortalama güçlüğüne sahip madde havuzlarında madde kullanım sıklığı kontrol yöntemlerinin yapıldığı bir araştırmaya rastlanmıştır. Lee ve Dodd'un (2012) yapmış olduğu bu araştırmada üç farklı madde havuzu (kolay, orta, zor) ve iki farklı örneklem grubu (negatif çarpık, normal) kullanılmış ve madde kullanım sıklığı kontrol edilmiştir. Araştırma sonucunda orta güçlükteki madde havuzunun ölçme kesinliği, maksimum madde kullanım sıklığı oranı ve madde havuzunun kullanımı açısından, yeterli dağılımına bakılmaksızın, daha iyi sonuçlar verdiği görülmüştür. Ayrıca normal dağılıma sahip gruba bakıldığında, madde havuzunun güçlük düzeyi yükseldiğinde En Çok Bilgi yöntemi ve Bilgi Koşullu Tesadüfi-6 madde yöntemlerinde kullanılmayan madde sayısının arttığı görülmüştür. Bu araştırma çoklu puanlanan maddelerin yer aldığı madde havuzlarında yürütülmüştür. Literatüre bakıldığında da; 1-0 puanlanan maddeler ve çoklu puanlanan maddelerde madde kullanım sıklığı kontrol

yöntemlerinin madde kullanım sıklığı açısından farklı sonuçlar verdiği görülmektedir (Boyd, 2003; Boyd, Dodd ve Choi, 2010, Akt. Lee ve Dodd, 2012). Buna rağmen madde havuzu yüksek güçlük düzeyinde olduğunda, kullanılmayan madde sayısının bazı yöntemlerde artış göstermesi açısından sonuçlar benzerdir.

Madde kullanım sıklığı kontrol yöntemlerine bakıldığında; her bir madde seçme yönteminde ve her bir madde havuzunda genellikle en iyi sonuçları veren yöntemin Azalarak Kaybolma yönteminin olduğu görülmektedir. Han (2009) tarafından geliştirilen bu madde kullanım sıklığı kontrol yöntemiyle yapılan araştırmalar yine sadece Han tarafından yapılmıştır. Han (2009) çalışmasında, bilgi fonksiyonunun $\theta = 1$ civarında en yoğun olduğu bir madde havuzunda farklı madde seçme ve kullanım sıklığı kontrol yöntemlerini incelemiştir. Araştırmanın sonuçlarında; Aşamalı En Çok Bilgi Oranı (AEBO) yöntemi ile Azalarak Kaybolma (AK) yönteminin birlikte kullanıldığında, Fisher'in En Çok Bilgi (FEB) yöntemi ile Azalarak Kaybolma yönteminin birlikte kullanıldığı duruma göre madde havuzunun verimli kullanılması açısından daha iyi sonuçlar vermiştir. Ancak FEB ve AEBO yöntemlerinin madde kullanım sıklığının 0.20 ile sınırlandırıldığı durumlarda ise benzer sonuçlar vermiştir. Yapılan bu araştırmada ise FEB ve AEBO yöntemlerinin Azalarak Kaybolma madde kullanım sıklığı kontrolü yöntemiyle birlikte kullanıldığında birbirine yakın sonuçlar verdiği görülmüştür. Bu nedenle iki araştırmanın bulguları birbirini farklılaşmaktadır.

Madde havuzunun güçlüğü orta düzeyde olduğunda, FEB ve a-Tabakalama yönteminin KOD, SH ve AK yönteminde elde edilen uyum katsayıları, Han (2012)'in araştırması ile benzer sonuçlara sahiptir. Madde havuzunun kullanımına bakıldığında ise, test çakışması açısından araştırma ile benzer sonuçların elde edildiği görülürken, hedeflenen kullanım sıklığına uyma açısından farklı sonuçlar elde edilmiştir. Han'ın araştırmasında hem FEB hem de a-Tabakalama yönteminin kullanıldığı her iki durumda da SH yöntemi hedeflenen kullanım sıklığını aşarken Azalarak Kaybolma yöntemi bu hedefi aşmamıştır. Bu araştırmada ise Azalarak Kaybolma yönteminin a-Tabakalama yöntemi ile birlikte kullanıldığı durumda

hedeflenen kullanım sıklığını aştığı görülürken, SH yöntemi birinci alt problem hariç diğer durumların tümünde hedeflenen kullanım sıklığına uyum sağlamıştır.

Araştırmada a-Tabakalama yönteminin madde kullanım sıklığının kontrol edilmediği durum ile Sympson-Hetter yöntemiyle beraber kullanıldığı durum karşılaştırıldığında; hem orta hem de zor güçlükteki madde havuzlarında, ölçme kesinliğinde çok büyük değişikliklerin olmaması, madde havuzunun daha verimli kullanılması, hedeflenen kullanma sıklığına uyulması ve test çakışmasının azaltılması açısından a-Tabakalama + Sympson-Hetter yönteminin daha iyi olduğu görülmüştür. Bu bulgu; Leung, Chang ve Hau (2002)'nin bulguları ile desteklenmektedir.

Her iki madde havuzunda da madde kullanım sıklığının kontrol edilmediği durumda, madde havuzunun kullanımı ve test çakışması açısından en iyi sonucu a-Tabakalama yöntemi vermiştir. Fisher'in En Çok Bilgi yöntemi ve Aşamalı En Çok Bilgi Oranı yöntemleri ise birbirine yakın sonuçlar vermişlerdir. Chang ve Ying (1999)'in çalışmasında da, a-Tabakalama yöntemi test çakışması ve madde havuzunun kullanımı açısından Fisher'in En Çok Bilgi yöntemine göre daha iyi sonuçlar vermiştir.

5. SONUÇ ve ÖNERİLER

5.1. Sonuçlar

Araştırmaya ait sonuçlar, hem alt problemlere ait sonuçlar hem de madde havuzu özellikleri temelinde sonuçlar olarak iki başlık altında verilmiştir.

5.1.1. Alt Problemlere Ait Sonuçlar

5.1.1.1. Birinci Alt Probleme Ait Sonuçlar

Madde kullanım sıklığı kontrol yöntemlerinin kullanılmasıyla, madde kullanım sıklığı kontrolünün yapılmadığı duruma göre; uyum, RMSE, yanlılık ve OMF değerlerinde az da olsa düşüşler olmuştur. Ancak değerler arasında büyük farklar olmamıştır. Bu sebeple, madde kullanım sıklığı kontrol yöntemlerinin ölçme kesinliğini olumsuz yönde etkilemeyeceği sonucuna varılmıştır.

Madde kullanım sıklığı kontrol yöntemleri karşılaştırıldığında ise, hem madde havuzunda yer alan maddeleri farklı oranlarda kullanabilmesi hem de madde çakışması açısından Azalarak Kaybolma yönteminin test güvenliğini sağlamada oldukça başarılı olduğu bulunmuştur. F oranlarına bakıldığında da, Azalarak Kaybolma yönteminin madde kullanım sıklığının kontrol edilmediği duruma göre, madde havuzunun kullanımındaki çarpıklığı görece olarak en çok azaltan yöntem olduğu sonucuna ulaşılmıştır.

Bilgi Koşullu Tesadüfi yönteminde farklı madde seti sayıları kullanılmasına rağmen (5 ve 10 madde), test güvenliğini gösteren sonuçlardaki değişim yüksek olmamıştır. Madde kullanım sıklığının kontrol edilmediği durum ile Bilgi Koşullu Tesadüfi yöntemler karşılaştırıldığında, hem madde havuzunun kullanımı hem de test çakışması açısından Bilgi Koşullu Tesadüfi yönteminin (5 ve 10 madde) iyileştirici sonuçlarının düşük olduğu söylenebilir.

Sympson-Hetter yönteminin ise test güvenliği açısından Azalarak Kaybolma yönteminden sonra en iyi ikinci yöntem olduğu sonucuna ulaşılmıştır.

5.1.1.2. İkinci Alt Probleme Ait Sonuçlar

Ölçme kesinliğinin gözlenmesinde kullanılan göstergelere göre, uyum ve RMSE değerlerinde en çok farklılık Sympson-Hetter yönteminde olmuştur. Yöntemler arasında en yanlış sonucu BKT-5 madde yöntemi verirken, gerçek ve kestirilen theta değerlerine ait en yüksek mutlak fark ortalamasını yine Sympson-Hetter yöntemi vermiştir. Elde edilen bulgular sonucunda madde kullanım sıklığının kontrol edilmediği ve madde kullanım sıklığının kontrol edildiği durumlar arasında büyük farkların olmadığı sonucuna ulaşılmıştır.

Madde havuzunun kullanılması açısından Bilgi Koşullu Tesadüfi yöntemlerinin iyi sonuçlar vermediği bulunmuştur. Bu yöntemlere ait madde kullanım sıklığı oranlarının standart sapma değerleri ve χ^2 değerleri, madde kullanım sıklığının kontrol edilmediği duruma ait değerlere yakın olduğu sonucuna ulaşılmıştır. Bilgi Koşullu Tesadüfi yöntemlerinin, madde kullanım sıklığının kontrol edilmediği durumuna göre tek üstünlüğü en yüksek madde kullanım sıklığı oranlarını düşürmesi olmuştur.

Madde havuzundaki farklı maddelerin kullanılması ve maddelerin çakışması açısından Sympson-Hetter ve Azalarak Kaybolma yöntemleri birbirlerine yakın sonuçlar vermiştir. Buna göre, her iki yöntemde de hedeflenen kullanım sıklığı oranı aşılmamış ve kullanılmayan maddelerin de kullanılması sağlanmıştır. Fakat ölçme kesinliğini gösteren analizlere bakıldığında, değerler arasında büyük farklar olmamasına rağmen, Azalarak Kaybolma yöntemi Sympson-Hetter yöntemine göre biraz daha iyi sonuçlar vermiştir.

5.1.1.3. Üçüncü Alt Probleme Ait Sonuçlar

Madde kullanım sıklığı kontrol yöntemlerinin kullanılmasıyla, madde kullanım sıklığı kontrolünün yapılmadığı duruma göre; uyum, RMSE, yanlışlık ve OMF değerlerinde az da olsa düşüşler olmuştur. Ancak değerler arasında büyük farkların olmadığı bulunmuştur. Bu sebeple, madde kullanım sıklığı kontrol yöntemlerinin ölçme kesinliğini olumsuz yönde etkilemeyeceği sonucuna varılmıştır.

BKT-5 madde ve BKT-10 madde yöntemleri, madde kullanım sıklığı oranlarının dağılımı açısından benzer sonuçlar vermiştir. Bununla birlikte madde havuzunun kullanımındaki çarpıklık ve en yüksek madde kullanım sıklığı oranı değeri her iki yöntem için de birbirine oldukça yakındır. Testlerin çakışma oranları incelendiğinde de yöntemin her iki durumda da yakın sonuçlar verdiği gözlenmiştir. Buna göre, madde seçme yöntemi olarak a-Tabakalama yöntemi kullanıldığında Bilgi Koşullu Tesadüfi yönteminde kullanılan madde seti büyüklüğü sonuçlarda çok büyük farklar yaratmamıştır.

Sympson-Hetter yöntemi ise, hedeflenen kullanım sıklığı oranına sadık kalmasına rağmen, Bilgi Koşullu Tesadüfi yöntemine (5 ve 10 madde) göre daha kötü sonuçlar vermiştir.

Azalarak Kaybolma yönteminin kullanılmasıyla madde havuzunda yer alan tüm maddelerin en az bir tekrarda uygulandığı ve madde kullanım sıklığı oranlarının birbirine yakın olduğu görülmektedir. Bunlara rağmen, Azalarak Kaybolma yöntemi hedeflenen kullanım sıklığı oranını aşmıştır.

Elde edilen bulgulara göre, Azalarak Kaybolma yöntemi diğer yöntemlere nazaran daha iyi sonuçlar vermiştir. BKT-10 madde yönteminin de madde havuzunun kullanılması ve madde çakışması açısından Azalarak Kaybolma yöntemine yakın sonuçlar verdiği bulunmuştur.

5.1.1.4. Dördüncü Alt Probleme Ait Sonuçlar

Ölçme kesinliğinin gözlenmesinde kullanılan göstergelere göre, uyum ve RMSE değerlerinde en çok farklılık Azalarak Kaybolma yönteminde elde edilmiştir. Yöntemler arasında en yanlış sonucu Sympson-Hetter yöntemi verirken, gerçek ve kestirilen theta değerlerine ait en yüksek mutlak fark ortalamasının yine Azalarak Kaybolma yöntemine ait olduğu görülmektedir. Elde edilen bulgular sonucunda madde kullanım sıklığının kontrol edilmediği ve edildiği durumlar arasında büyük farkların olmadığı sonucuna ulaşılmıştır.

Madde kullanım sıklığının kontrol edilmediği durumda; madde havuzunun kullanımındaki çarpıklığın, test çakışmasının ve madde kullanım sıklığı oranlarına ait standart sapmanın madde kullanım sıklığının kontrol edildiği durumlara kıyasla daha yüksek olduğu bulunmuştur. Bu durumda madde havuzunun daha verimli kullanılabilmesi için madde kullanım sıklığı kontrol yöntemlerinin kullanılması gerekliliği sonucuna ulaşılmıştır.

BKT-5 madde yöntemi kullanıldığında madde kullanım sıklığının kontrol edilmediği duruma göre; madde havuzunun kullanımındaki çarpıklık azalmış, kullanılan madde sayısı artmış ve en yüksek madde kullanım sıklığı oranı düşmüştür. Ancak test çakışmasında büyük bir farklılık olmamıştır.

BKT-10 madde yöntemi, madde kullanımının kontrol edilmediği duruma göre daha iyi sonuçlar verirken, BKT-5 madde yöntemi ile elde edilen değerlere yakın sonuçlar vermiştir. Bu yöntem ile en yüksek madde kullanım sıklığı oranı daha da düşmüştür.

Sympson-Hetter yöntemi kullanılarak hedeflenen madde kullanım sıklığı oranına sadık kalınmıştır.

Her bir yöntemin sonucu incelendiğinde madde havuzunun kullanımı ve test çakışması açısından en iyi sonuçları Azalarak Kaybolma yöntemi vermiştir. Madde havuzu güçlüğü uygulamanın yapıldığı gruba göre yüksek düzeyde olsa da, Azalarak Kaybolma yöntemi kullanılarak madde havuzundaki tüm maddelerin en az bir tekrarda kullanılması sağlanmıştır. Fakat Azalarak Kaybolma yöntemi ile hedeflenen madde kullanım sıklığı oranı aşılmıştır. Azalarak Kaybolma yönteminin kullanılmasıyla madde kullanım sıklığının kontrol edilmediği duruma göre madde havuzunun kullanımındaki çarpıklık oldukça azalmıştır.

5.1.1.5. Beşinci Alt Probleme Ait Sonuçlar

Madde kullanım sıklığı kontrol yöntemlerinin kullanılmasıyla, madde kullanım sıklığı kontrolünün yapılmadığı duruma göre; uyum, RMSE, yanlılık ve OMF değerlerinde az da olsa düşüşler olmuştur. Ancak değerler arasında büyük farkların olmadığı

görülmüştür. Bu sebeple, madde kullanım sıklığı kontrol yöntemlerinin ölçme kesinliğini olumsuz yönde etkilemeyeceği sonucuna varılmıştır.

Madde havuzunun daha verimli kullanılması için madde kullanım sıklığı kontrol yöntemleri incelendiğinde, BKT-5 madde yöntemi madde havuzunun kullanımındaki çarpıklık ve madde çakışması açısından, madde kullanım sıklığının kontrol edilmediği duruma yakın sonuçlar vermiştir. Bilgi Koşullu Tesadüfi yöntemindeki madde seti 10 yapıldığında en yüksek madde kullanım sıklığı oranı, χ^2 değeri ve kullanılmayan madde sayısı değerleri düşmüştür.

Sympson-Hetter yöntemi ile BKT-10 madde yöntemi, en az bir kere bile kullanılmayan madde sayısı açısından benzer sonuçlar vermesine rağmen, Sympson-Hetter yöntemi kullanılarak madde havuzundaki çarpıklık BKT-10 madde yöntemine göre azalmıştır. Ayrıca Sympson-Hetter yönteminin kullanılmasıyla test çakışma oranı da daha düşük elde edilmiştir.

Azalarak Kaybolma yöntemi, madde kullanım sıklığının kontrol edilmediği duruma göre, madde havuzundaki çarpıklığı en çok azaltan yöntem olmuştur. Madde kullanım sıklığı yöntemleri karşılaştırıldığında, madde havuzunun kullanımı açısından Azalarak Kaybolma yönteminin diğer yöntemlere göre daha iyi sonuçlar verdiği sonucuna ulaşılmıştır.

5.1.1.6. Altıncı Alt Probleme Ait Sonuçlar

Ölçme kesinliğinin gözlenmesinde kullanılan göstergelere bakıldığında, uyum ve RMSE değerlerinde en çok farklılık Sympson-Hetter yönteminde elde edilmiştir. Yöntemler arasında en yanlı sonucu BKT-5 madde yöntemi verirken, gerçek ve kestirilen theta değerlerine ait en yüksek mutlak fark ortalaması yine Sympson-Hetter yöntemine ait olmuştur. Elde edilen bulgular sonucunda madde kullanım sıklığının kontrol edilmediği ve edildiği durumlar arasında büyük farkların olmadığı görülmüştür.

BKT-5 madde yöntemi kullanıldığında, madde havuzunun kullanımındaki çarpıklık madde kullanım sıklığının kontrol edilmediği duruma göre, görece olarak düşük bir

oranda azalmıştır. Ayrıca kullanılmayan madde yüzdesi ve madde kullanım sıklığı oranlarının dağılımı da, madde kullanım sıklığının kontrol edilmediği durumdaki değer ve dağılım ile oldukça yakın olmuştur. BKT-10 madde yöntemi de, BKT-5 madde yönteminde olduğu gibi, madde kullanım sıklığının kontrol edilmediği duruma göre çok farklı sonuçlar vermemiştir.

Sympson-Hetter ve Azalarak Kaybolma yöntemlerinin kullanılmasıyla hedeflenen kullanım sıklığı oranına sadık kalınmıştır. Madde havuzunun kullanımı açısından Azalarak Kaybolma yöntemi daha iyi sonuçlar vermiştir. Azalarak Kaybolma yöntemi, Sympson-Hetter yöntemine göre çarpıklığın azalmasında görece olarak daha etkili olmuştur. Ayrıca Azalarak Kaybolma yöntemi en düşük test çakışma oranına sahip olmuştur.

Tüm bu değişkenler ve ölçme kesinliğini temsil eden değişkenler beraber değerlendirildiğinde, Azalarak Kaybolma yönteminin diğer madde kullanım sıklığı kontrol yöntemlerine göre daha iyi sonuçlar verdiği sonucuna ulaşılmıştır.

5.1.2. Madde Havuzu Özellikleri Temelinde Sonuçlar

5.1.2.1. Orta Güçlük Düzeyinde Olan Madde Havuzuna Ait Sonuçlar

Madde seçme yöntemleri bazında madde kullanım sıklığı kontrol yöntemleri karşılaştırıldığında ölçme kesinliğinin; genel olarak Aşamalı En Çok Bilgi Oranı yönteminin madde kullanım sıklığı kontrolünün yapılmadığı durumda daha iyi olduğu sonucuna varılmıştır. Fisher'in En Çok Bilgi yöntemi ile Sympson-Hetter yönteminin birlikte kullanıldığı durumda ise en yanlı değere ulaşılmıştır.

Madde kullanım sıklığı kontrol yöntemlerine test güvenliği açısından genel olarak bakıldığında tüm madde seçme yöntemlerinde; madde kullanım sıklığı kontrolünün yapılmadığı koşullarda madde havuzunun kullanımının en çarpık olduğu, test çakışması oranının ise en yüksek olduğu sonucuna varılırken; Azalarak Kaybolma yönteminin kullanıldığı koşullarda madde havuzunun iyi kullanıldığı ve test çakışması oranının düştüğü sonucuna varılmıştır. Ancak a-Tabakalama yöntemi ile

Azalarak Kaybolma yönteminin birlikte kullanılması durumunda hedeflenen kullanım sıklığı oranının aşıldığı sonucu elde edilmiştir.

Aşamalı En Çok Bilgi Oranı yönteminin hem ölçme kesinliği hem de madde havuzunun kullanılmasında Fisher'in En Çok Bilgi yöntemine oldukça benzer şekilde çalıştığı bulunmuştur. Madde seçme yöntemi olarak a-Tabakalama yöntemi seçildiğinde, madde kullanım sıklığının kontrol edilmemesi durumunda bile, Fisher'in En Çok Bilgi yöntemi ve Aşamalı En Çok Bilgi Oranı yöntemlerinin Azalarak Kaybolma yöntemiyle beraber kullanıldığı durumlardan bile daha fazla maddenin kullanılması sağlanmıştır.

Bilgi Koşullu Tesadüfi yönteminin (5 ve 10 madde), Fisher'in En Çok Bilgi yöntemi ve Aşamalı En Çok Bilgi Oranı yöntemi ile birlikte kullanıldığında, madde havuzunun kullanımı ve test çakışması açısından Sympson-Hetter yönteminden daha kötü sonuçlar vermiştir. Fakat bu sonuç, Bilgi Koşullu Tesadüfi yönteminin a-Tabakalama ile birleştirilip kullanıldığında değişmiş olup Bilgi Koşullu Tesadüfi yöntemlerinin madde havuzunu daha iyi kullanması ve test çakışmasını azaltması ile sonuçlanmıştır.

Sonuç olarak, madde seçme yöntemleri ve madde kullanım sıklığı kontrol yöntemlerinin birleştirilmesiyle oluşturulan birleştirilmiş yöntemlerden en iyisinin a-Tabakalama + Azalarak Kaybolma yöntemi olduğu sonucuna ulaşılmıştır. Çünkü bu yöntem diğer yöntemlere göre, ölçme kesinliğinden çok az ödün verip test güvenliğini oldukça yükseltmiştir.

Madde seçme yöntemleri bazında madde kullanım sıklığı kontrol yöntemlerinin ölçme kesinliği ve test güvenliği açısından daha detaylı karşılaştırılabilmesi için EK-4 ve EK-5'teki yer alan tablolar incelenebilir.

5.1.2.2. Yüksek Güçlük Düzeyinde Olan Madde Havuzuna Ait Sonuçlar

Madde kullanım sıklığı kontrol yöntemlerine ölçme kesinliği açısından genel olarak bakıldığında; tüm madde seçme yöntemlerinde madde kullanım sıklığı kontrolünün yapılmadığı koşullarda en iyi değerlerin elde edildiği görülürken; Fisher'in En Çok

Bilgi ve Aşamalı En Çok Bilgi Oranı yöntemlerinde Sympson-Hetter yönteminin kullanıldığı, a-Tabakalama yönteminde ise Azalarak Kaybolma yönteminin kullanıldığı koşullarda en düşük değerler elde edilmiştir.

Madde kullanım sıklığı kontrol yöntemlerine test güvenliği açısından genel olarak bakıldığında tüm madde seçme yöntemlerinde; madde kullanım sıklığı kontrolünün yapılmadığı koşullarda madde havuzunun kullanımının en çarpık olduğu, test çakışması oranının ise en yüksek olduğu sonucuna varılırken; Azalarak Kaybolma yönteminin kullanıldığı koşullar madde havuzunun iyi kullanıldığı ve test çakışması oranının düştüğü sonucuna varılmıştır. Ancak a-Tabakalama yöntemi ile Azalarak Kaybolma yönteminin birlikte kullanılması durumunda hedeflenen kullanım sıklığı oranının aşıldığı sonucuna varılmıştır.

Sonuç olarak; madde kullanım sıklığı kontrolünün yapıldığında az da olsa ölçme kesinliğinden feragat edildiği, ama test güvenliği açısından madde havuzunun kullanımında önemli değişiklikler olduğu ve test çakışmasının da azaldığı sonucuna varılmıştır. Bu durumda; madde seçme yöntemleri ve madde kullanım sıklığı kontrol yöntemlerinin birleştirilmesiyle oluşturulan birleştirilmiş yöntemler arasında en iyi a-Tabakalama yöntemi ile Azalarak Kaybolma yöntemi ile birleşiminin ölçme kesinliğinden az da olsa ödün vererek, hem madde havuzunun kullanımı hem de test çakışması açısından oldukça iyi sonuçlar verdiği sonucuna ulaşılmıştır.

Madde seçme yöntemleri bazında madde kullanım sıklığı kontrol yöntemlerinin ölçme kesinliği ve test güvenliği açısından daha detaylı karşılaştırılabilmesi için EK-6 ve EK-7'de yer alan tablolar incelenebilir.

5.2. Öneriler

5.2.1. Uygulamaya Dönük Öneriler

Araştırmadan elde edilen bulgulara dayanarak;

- 1) BBT uygulamalarında test güvenliğinin sağlanması için madde kullanım sıklığı kontrol yöntemlerinin kullanılması önerilebilir.

- 2) BBT uygulamasında madde kullanım sıklığı kontrol yöntemi kullanılmayacaksa, test güvenliği açısından hem orta hem de yüksek güçlük düzeyindeki madde havuzlarında madde seçme yöntemi olarak a-Tabakalama yönteminin kullanılması önerilebilir.
- 3) BBT uygulamasında madde kullanım sıklığı kontrol yöntemi kullanılacaksa, Azalarak Kaybolma yönteminin kullanılması önerilebilir.
- 4) Madde havuzunun güçlüğünün uygulanan gruba göre orta veya yüksek düzeyde olduğunda, a-Tabakalama madde seçme yöntemi ve Azalarak Kaybolma madde kullanım sıklığı kontrol yönteminin beraber kullanılması önerilebilir.
- 5) BBT uygulamasında hedeflenen madde kullanım sıklığı oranının kullanıldığı Sympton-Hetter yöntemi ile Azalarak Kaybolma yöntemi kullanılabilirlik açısından karşılaştırıldığında Azalarak Kaybolma yönteminin tercih edilmesi önerilebilir.

5.2.2. Araştırmaya Dönük Öneriler

- 1) Araştırmada madde seçme yöntemleri olarak sadece Fisher'in En Çok Bilgi yöntemi, a-Tabakalama yöntemi ve Aşamalı En Çok Bilgi Oranı yöntemi kullanılmıştır. Benzer bir araştırma farklı madde seçme yöntemleri kullanılarak yapılabilir.
- 2) Araştırmada madde kullanım sıklığı kontrol yöntemleri olarak sadece Bilgi Koşullu Tesadüfi (5 ve 10 madde), Sympton-Hetter ve Azalarak Kaybolma yöntemleri kullanılmıştır. Benzer bir araştırma farklı madde kullanım sıklığı kontrol yöntemleri kullanılarak yapılabilir.
- 3) Bu araştırmada sadece madde havuzunun b parametresi değiştirilerek farklı madde havuzları oluşturulmuştur. a ve c parametrelerinin de değiştirilmesiyle oluşturulan madde havuzları benzer bir araştırmada kullanılabilir.

- 4) Farklı madde havuzu büyüklüklerinde madde kullanım sıklığı kontrol yöntemleri karşılaştırılabilir.
- 5) Araştırmada istenilen durumlar simüle edilip analizler yapılmıştır. Gerçek uygulamaların verileri kullanılarak madde kullanım sıklığı kontrol yöntemleri karşılaştırılabilir.
- 6) Araştırmada madde kapsamı dengesi ele alınmamıştır. Kapsam dengesinin de işleme dahil edildiği uygulamalarda madde kullanım sıklığı kontrol yöntemleri incelenebilir.

KAYNAKÇA

- Barrada, J. R., Olea, J., Ponsoda, V., & Abad, F. J. (2009). Test Overlap Rate and Item Exposure Rate as Indicators of Test Security in CATs. *Proceedings of the 2009 GMAC Conference on Computerized Adaptive Testing* .
- Boyd, A. M., Dodd, B., & Fitzpatrick, S. (2013). A Comparison of Exposure Control Procedures in CAT Systems Based on Different Measurement Models for Testlets. *Applied Measurement in Education*, 113–135.
- Boyd, M. A. (2003). *Strategies for Controlling Testlet Exposure Rates in Computerized Adaptive Testing Systems*. Unpublished Doctoral Thesis, The University of Texas, Austin.
- Burt, W., Davis, L. L., & Dodd, B. G. (2003). A Comparison of Item Exposure Control Procedures Using a CAT System Based on the Generalized Partial Credit Model. *Annual meeting of the American Educational Research Association*. Chicago.
- Chang, H. H., & Zhang, J. (2002). Hypergeometric Family and Item Overlap Rates In Computerized Adaptive Testing. *Psychometrika* , 387-398.
- Chang, H.-H. (2004). Understanding Computerized Adaptive Testing. D. Kaplan içinde, *The SAGE Handbook of Quantitative Methodology for the Social Sciences* (s. 117-133). Thousand Oaks: Sage Yayınları.
- Chang, H.-H., & Ying, Z. (1999). a-Stratified Multistage Computerized Adaptive Testing. *Applied Psychological Measurement*, 211-222.
- Chang, H.-H., Qian, J., & Ying, Z. (2001). a-Stratified Multistage Computerized Adaptive Testing with b Blocking. *Applied Psychological Measurement*, 333-341.
- Chang, S.-W., & Twu, B.-Y. (1998). *A Comparative Study of Item Exposure Control Methods in CAT*. ACT Research Report Series.
- Davis, L. L. (2002). *Strategies for Controlling Item Exposure in Computerized Adaptive Testing with Polytomously Scored Items*. Unpublished Doctoral Thesis, The University of Texas, Austin.
- Davis, L. L., & Dodd, B. G. (2005). *Strategies for Controlling Item Exposure in Computerized Adaptive Testing with Partial Credit Model*. Pearson Educational Measurement.
- Embretson, S. E., & Reise, S. P. (2000). *Item Response Theory for Psychologists*. Taylor & Francis.
- Erođlu, M. G. (2013). Bireyselleřtirilmiř Bilgisayarlı Test Uygulamalarında Farklı Sonlandırma Kurallarının Ölçme Kesinliđi ve Test Uzunluđu Açısından Karřılařtırılması. Yayınlanmamıř Doktora Tezi, Hacettepe Üniversitesi, Ankara.

- Evans, J. J. (2010). *Comparability of Examinee Proficiency Scores on Computer Adaptive Tests Using Real and Simulated Data*. The State University of New Jersey: New Jersey.
- Finkelman, M., Nering, M. L., & Roussos, L. A. (2009). A Conditional Exposure Control Method for Multidimensional Adaptive Testing. *Journal of Educational Measurement*, 84-103.
- Flaugher, R. (2000). Item Pools. H. Wainer içinde, *Computerized Adaptive Testing: A Primer Second Edition* (s. 37-59). New Jersey: Lawrence Erlbaum Associates, Publishers.
- French, B. F., & Thompson, T. D. (2003). The Evaluation of Exposure Control Procedures for an Operational CAT. *The annual meeting of the American Educational Research Association*. Chicago.
- Georgiadou, E., Triantafillou, E., & Economide, A. A. (2007). A Review of Item Exposure Control Strategies for Computerized Adaptive Testing Developed from 1983 to 2005. *The Journal of Technology, Learning, and Assessment*, 3-37.
- Gu, L., & Reckase, M. D. (2007). Designing Optimal Item Pools for Computerized Adaptive Tests with Sympson-Hetter Exposure Control. *Proceedings of the 2007 GMAC Conference on Computerized Adaptive Testing*.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory Principles and Applications*. Hingham: Kluwer Nijhoff Publishing.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of Item Response Theory*. SAGE Publications .
- Han, K. C. (2009). A Gradual Maximum Information Ratio Approach to Item Selection in Computerized Adaptive Testing. *Proceedings of the 2009 GMAC Conference on Computerized Adaptive Testing* (s. 1-15). Virginia: Graduate Management Admission Council.
- Han, K. C. (2011). *User's Manual: SimulCAT*. Graduate Management Admission Council.
- Han, K. T. (2012). An Efficiency Balanced Information Criterion for Item Selection in Computerized Adaptive Testing. *Journal of Educational Measurement*, 225–246.
- Kalender, İ. (2011). *Effects of Different Computerized Adaptive Testing Strategies on Recovery of Ability*. Unpublished Doctoral Thesis, Middle East Technical University, Ankara.
- Kalinowski, K. E. (2009). *Stratified Item Selection and Exposure Control In Unidimensional Adaptive Testing Of Two-dimensional Data*. Unpublished Doctoral Thesis, University of North Texas.

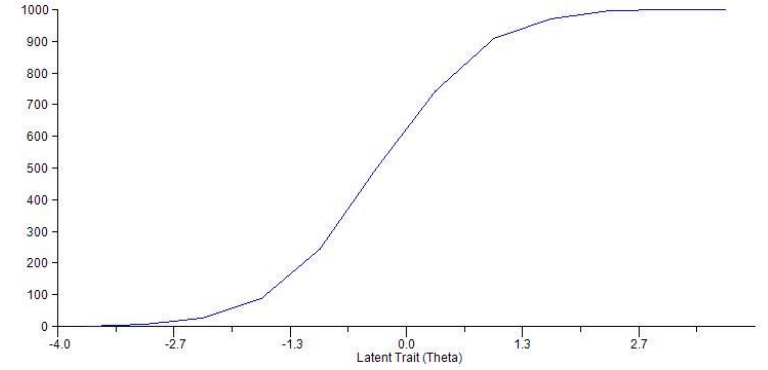
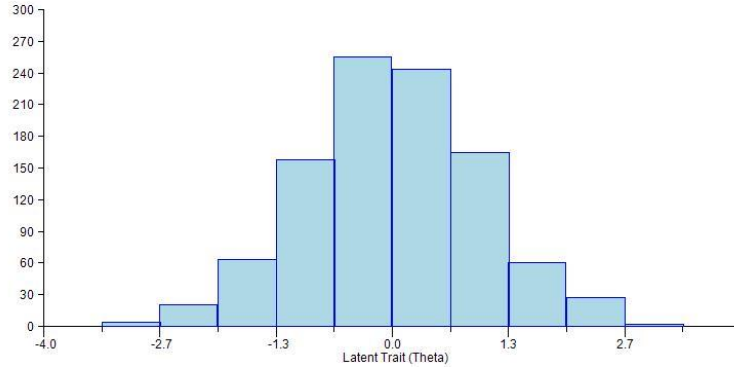
- Kaptan, F. (1993). Yetenek Kestiriminde Adaptive (Bireyselleştirilmiş) Test Uygulaması ile Geleneksel Kağıt Kalem Testi Uygulamasının Karşılaştırılması. Yayınlanmamış Doktora Tezi, Hacettepe Üniversitesi, Ankara.
- Karasar, N. (2009). Bilimsel Araştırma Yöntemi. Ankara: Nobel Basımevi.
- Kingsbury, G. G., & Zara, A. R. (1989). Procedures for Selecting Items for Computerized Adaptive Tests. *Applied Measurement in Education* , 359-375.
- Lee, H. Y., & Dodd, B. G. (2012). Comparisson of Exposure Controls, Item Pool Characterisric, and Population Distiributions for CAT Using the Partial Credit Model. *Educational and Psychological Measurement*, 159-175.
- Leung, C.-K., Chang, H.-H., & Hau, K.-T. (2002). Item Selection in Computerized Adaptive Testing: Improving the a-Stratified Design with the Sympson-Hetter Algorithm. *Applied Psychological Measurement*, 376-392.
- Macken-Ruiz, C. L. (2008). *A Comparison of Multi-stage and Computerized Adaptive Tests Based on the Generalized Partial Credit Model*. Unpublished Doctoral Thesis, University of Texas, Austin.
- Pastor, D. A., Dodd, B. G., & Chang, H.-H. (2002). A Comparison of Item Selection Techniques and Exposure Control Mechanisms in CATs Using the Generalized Partial Credit Model. *Applied Psychological Measurement* , 147-163.
- Reckase, M. D. (2010). Designing item pools to optimize the functioning of a computerized adaptive test. *Psychological Test and Assessment Modeling*, 127-141.
- Revuelta, J., & Ponsoda, V. (1998). A Comparison of Item Exposure Control Methods in Computerized Adaptive Testing. *Journal of Educational Measurement*, 311-327.
- Rudner, L. M. (1998). *An On-line, Interactive, Computer Adaptive Testing Tutorial*. Ocak 03, 2014 tarihinde <http://echo.edres.org:8080/scripts/cat/catdemo.htm> adresinden alındı
- Segall, D. O. (2005). Computerized Adaptive Testing. K. Kempf-Leonard içinde, *Encyclopedia of Social Measurement* (s. 429-438). Texas: Elsevier.
- Some Current Issues in CAT*. (tarih yok). 03 15, 2013 tarihinde International Association for Computerized Adaptive Testing: <http://iacat.org/node/491> adresinden alındı
- Stocking, M. L., & Lewis, C. (2000). Methods of Controlling the Exposure of Items in CAT. W. J. van der Linden, & G. A. Glas içinde, *Computerized Adaptive Testing: Theory and Practice* (s. 163-182). Kluwer Academic Publishers .
- Sulak, S. (2013). Bireyselleştirilmiş Bilgisayarlı Test Uygulamalarında Kullanılan Madde Seçme Yöntemlerinin Karşılaştırılması. Yayınlanmamış Doktora Tezi, Hacettepe Üniversitesi, Ankara.

- Thompson, N. A., & Weiss, D. J. (2011). A Framework for the Development of Computerized Adaptive Tests. *Practical Assessment, Research & Evaluation.*, 1-9.
- van der Linden, W. J., & Pashley, P. J. (2000). Item Selection and Ability Estimation in Adaptive Testing. W. J. van der Linden, & G. A. Glas içinde, *Computerized Adaptive Testing: Theory and Practice* (s. 1-25). Kluwer Academic Publishers.
- Way, W. D. (1998). Protecting the Integrity of Computerized Testing Item Pools. *Educational Measurement: Issues and Practice*, 17-27.
- Weiss, D. J. (2004). Computerized Adaptive Testing for Effective and Efficient Measurement in Counseling and Education. *Measurement and Evaluation in Counseling and Development*, 70-84.
- Weiss, D. J., & Kingsbury, G. G. (1984). Application of Computerized Adaptive Testing to Educational Problems. *Journal of Educational Measurement*, 361-375.

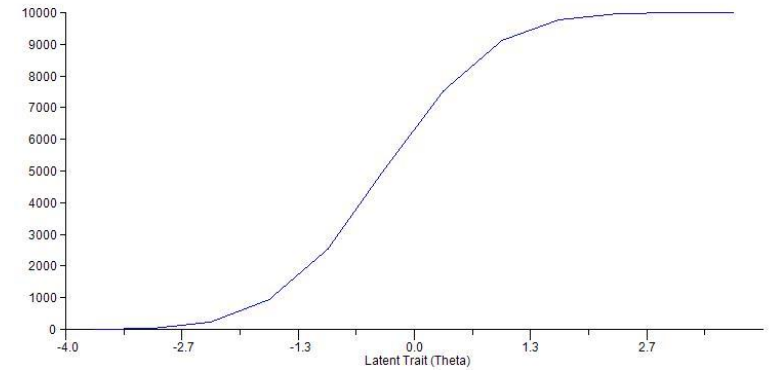
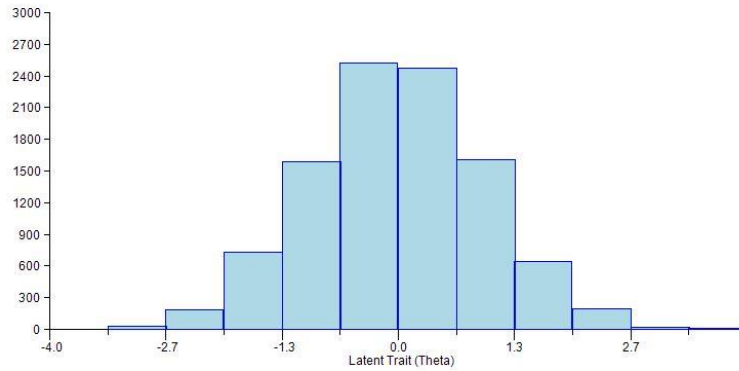
EKLER DİZİNİ

EK-1: Araştırmanın Yürütüldüğü Grubun Dağılım Grafikleri

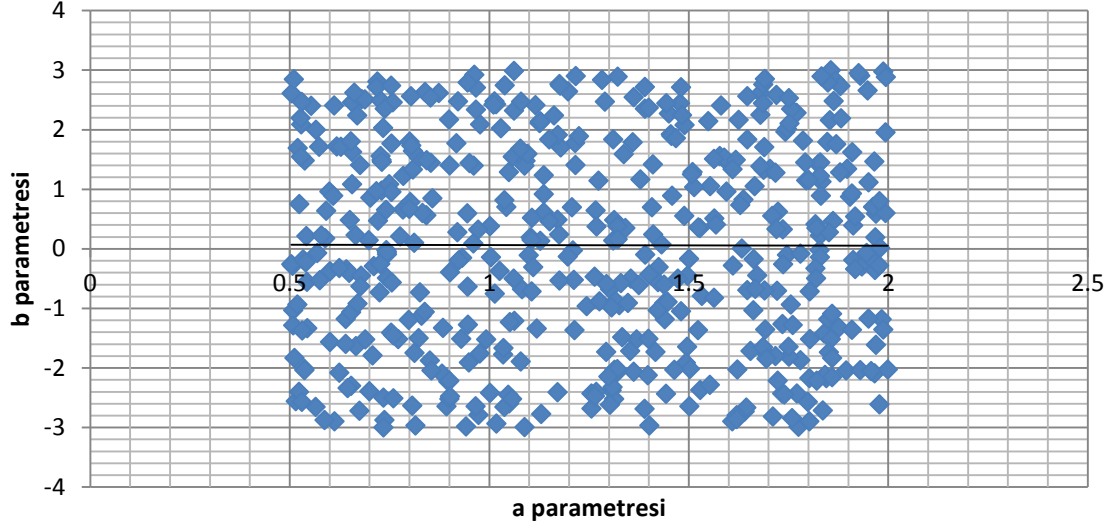
1. BBT Örnekleme



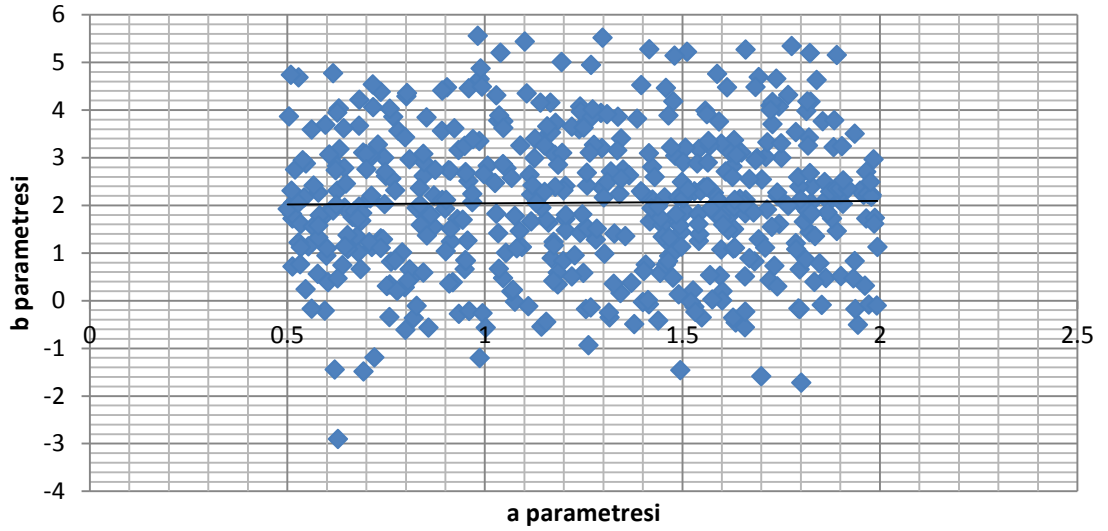
2. Sympson-Hetter Örnekleme



EK-2: Madde Havuzlarındaki Maddelerin b Parametrelerinin a Parametrelerine Göre Dağılım Grafikleri



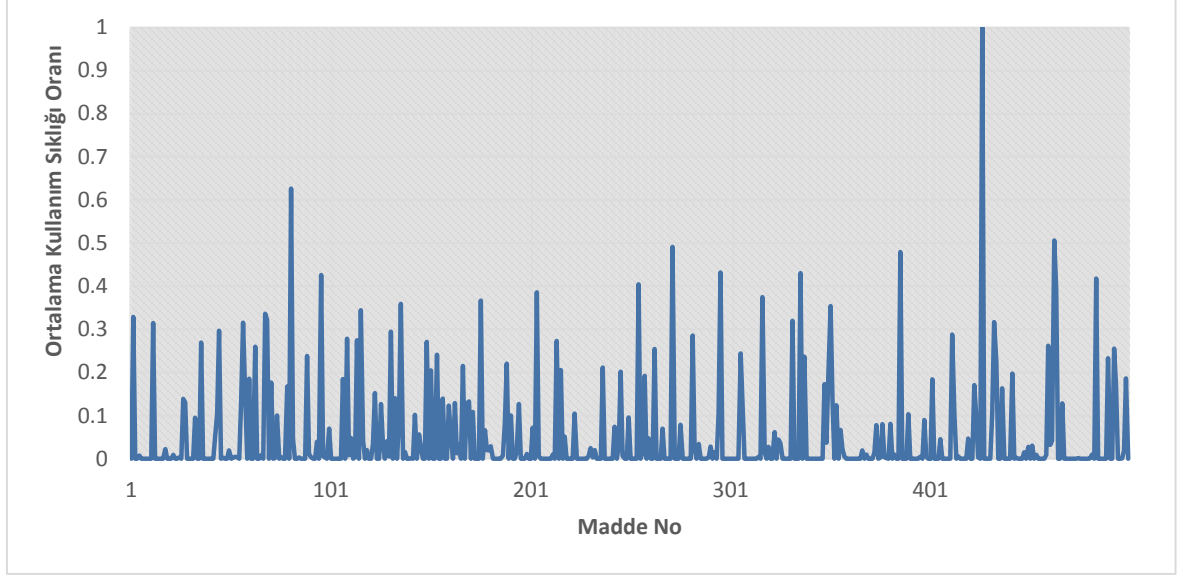
Şekil EK-2. 1. Madde güclüğü orta düzeyde olan madde havuzuna ait a ve b parametrelerinin dağılımı grafiği



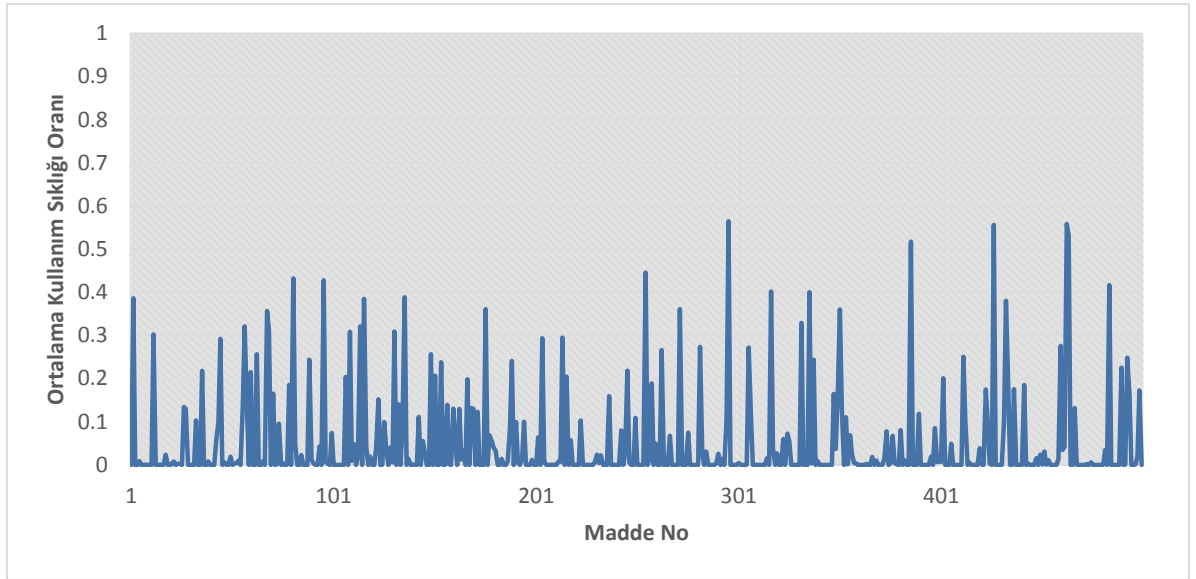
Şekil EK-2. 2. Madde güclüğü yüksek düzeyde olan madde havuzuna ait a ve b parametrelerinin dağılımı grafiği

EK-3: BBT Uygulamalarındaki Madde Kullanım Sıklığı Oranı Grafikleri

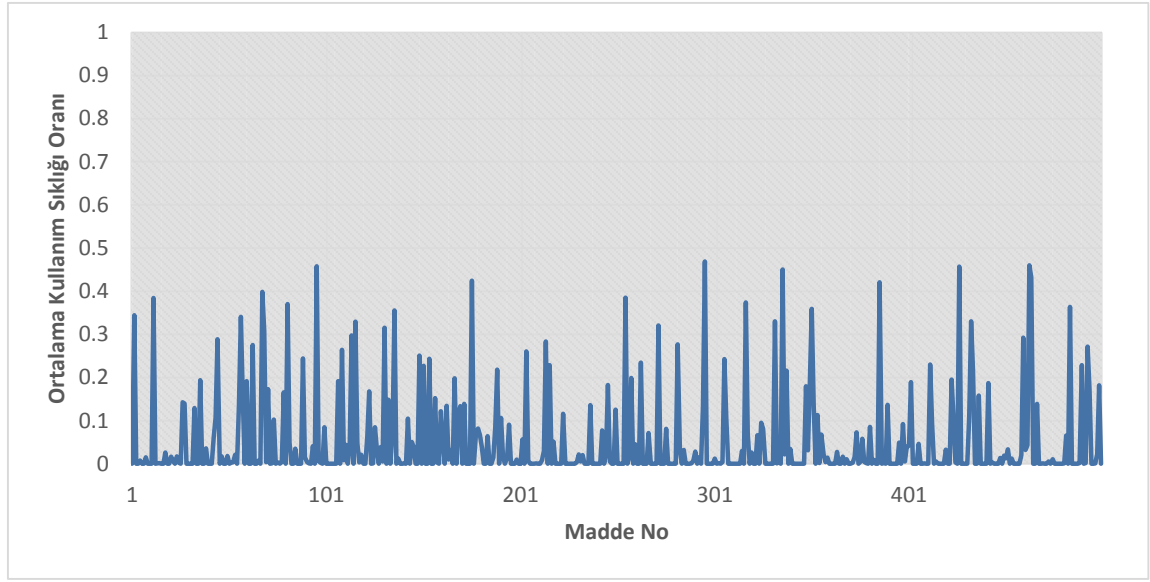
1. Birinci Alt Probleme Ait Madde Kullanım Sıklığı Oranı Grafikleri



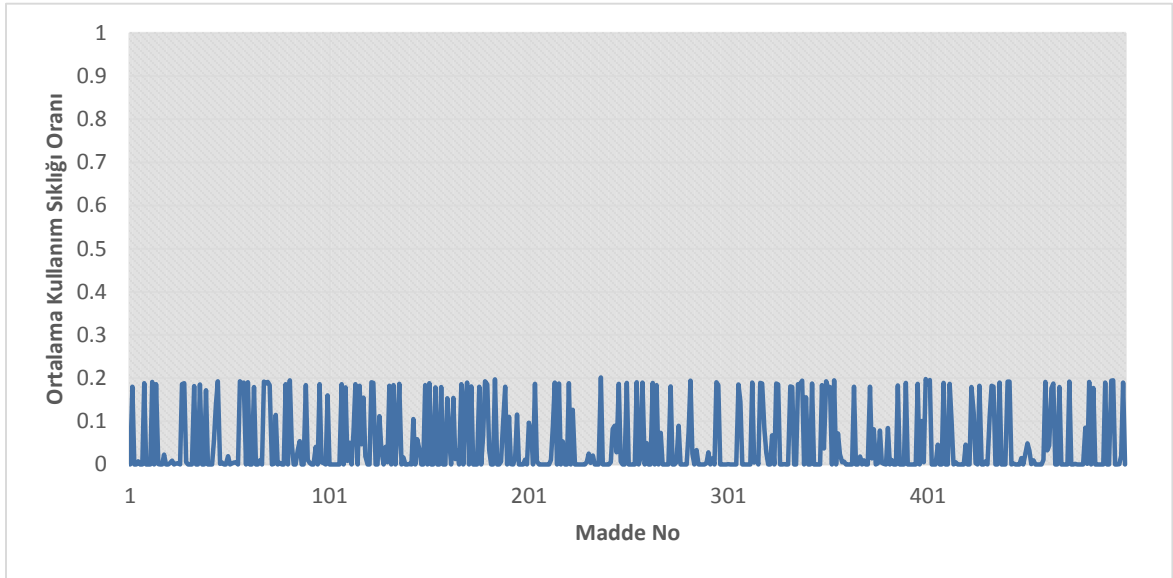
Şekil EK-3.1.1. Madde Kullanım Sıklığı Kontrolü Yapılmadığında



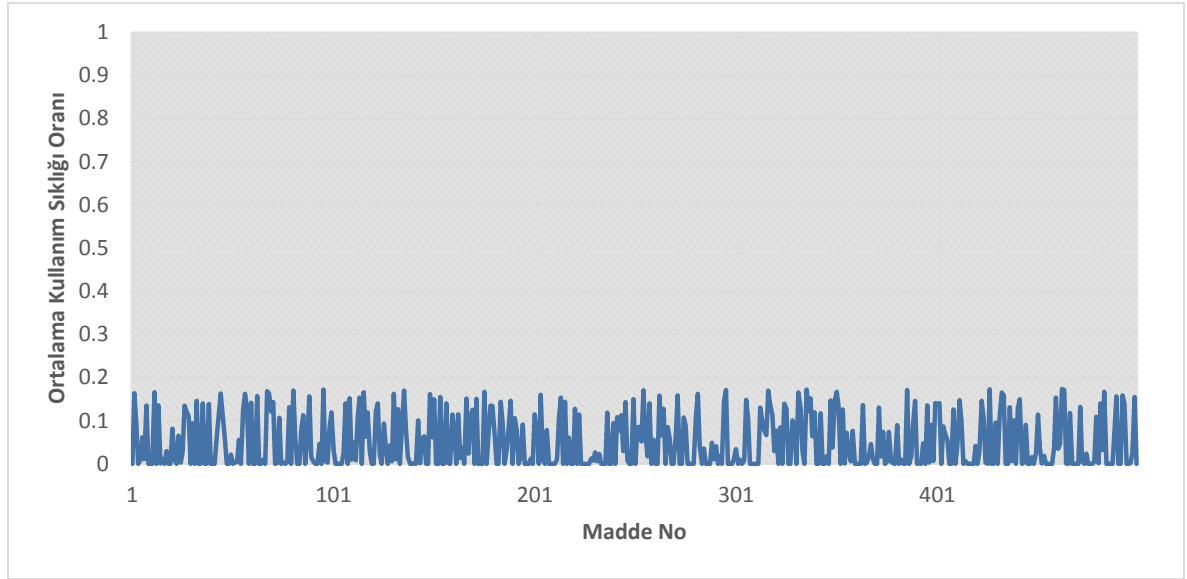
Şekil EK-3.1.2. BKT-5 Madde Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında



Şekil EK-3.1.3. BKT-10 Madde Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

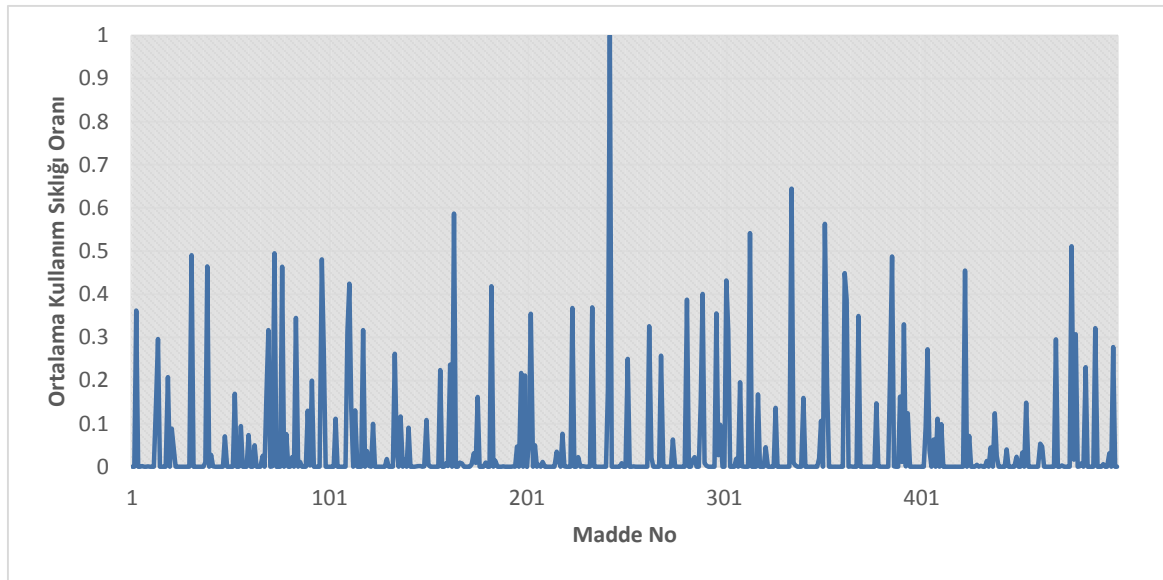


Şekil EK-3.1.4. Sympton-Hetter Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

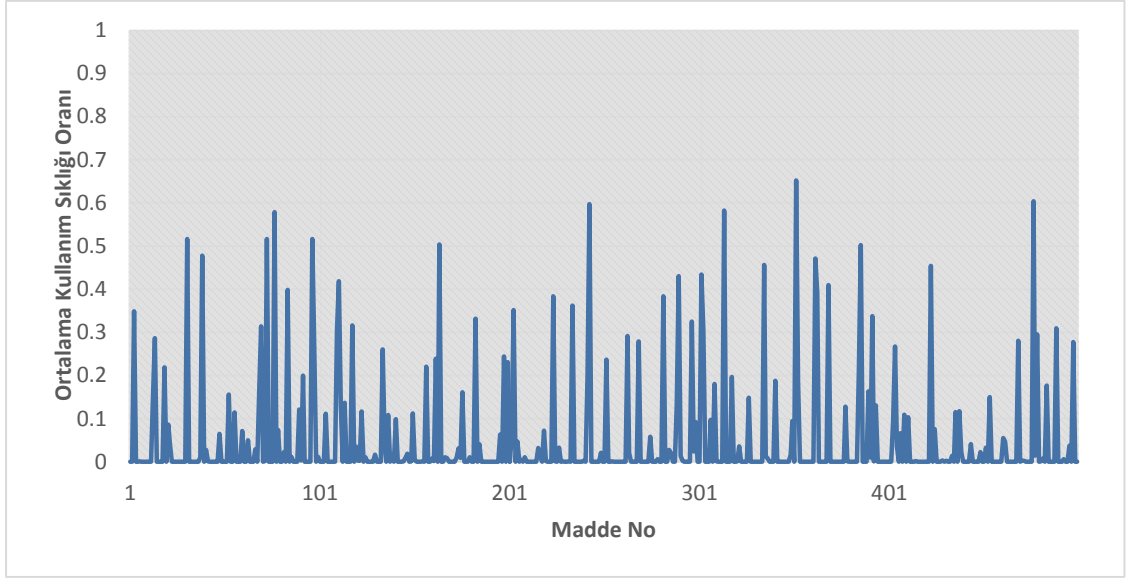


Şekil EK-3.1.5. Azalarak Kaybolma Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

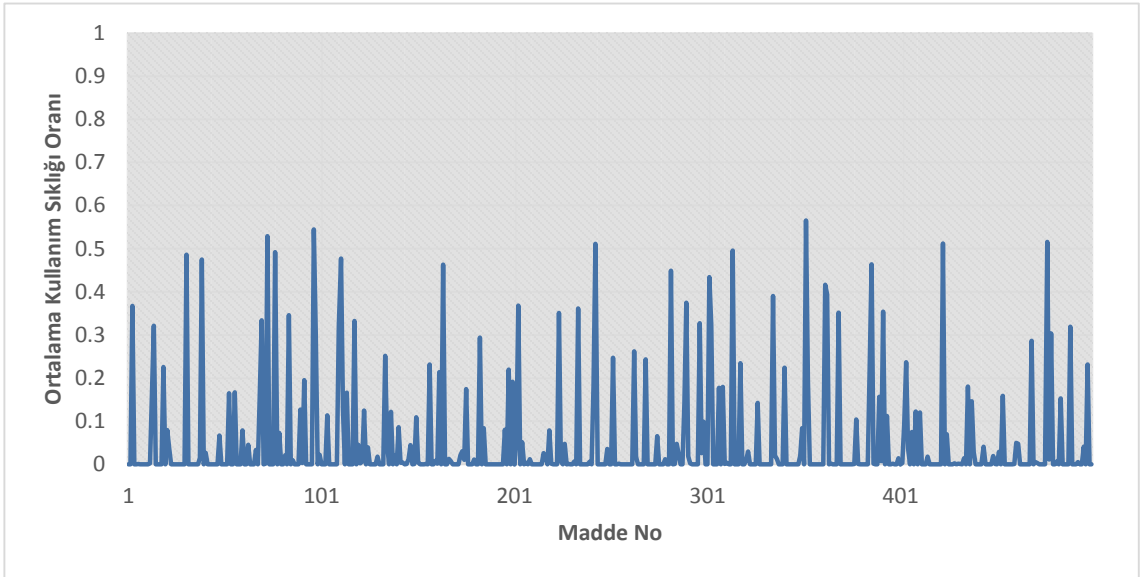
2.İkinci Alt Probleme Ait Madde Kullanım Sıklığı Oranı Grafikleri



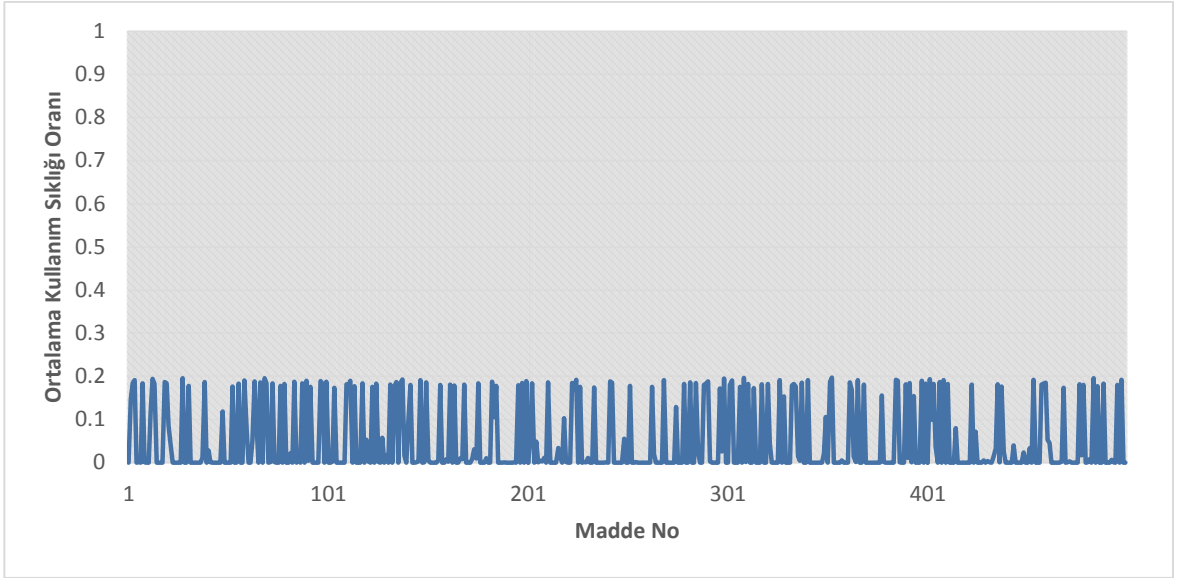
Şekil EK-3.2.1. Madde Kullanım Sıklığı Kontrolü Yapılmadığında



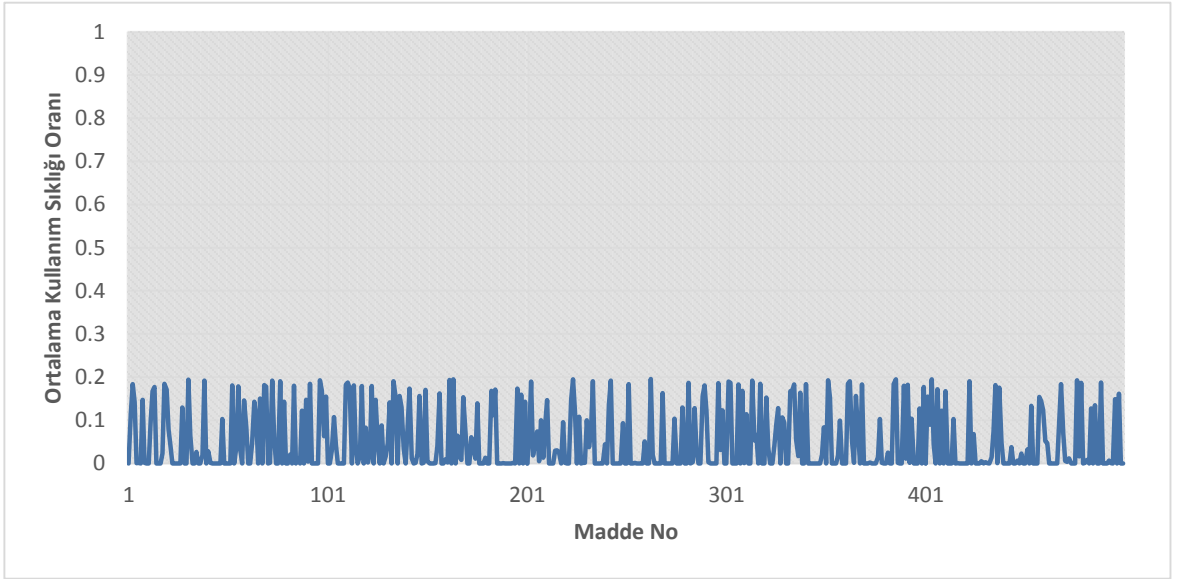
Şekil EK-3.2.2. BKT-5 Madde Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında



Şekil EK-3.2.3. BKT-10 Madde Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

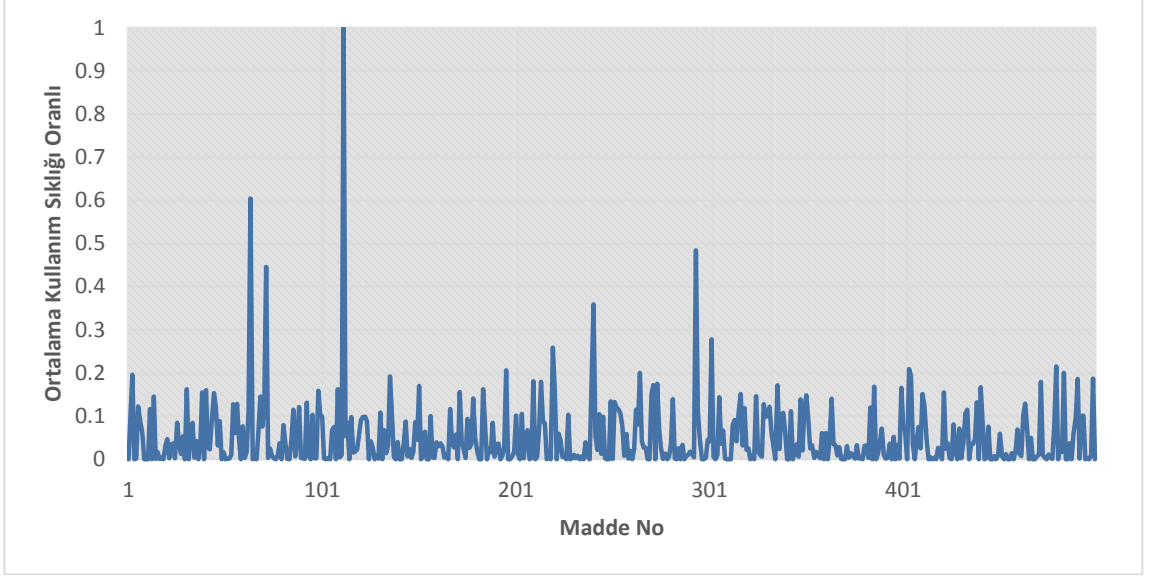


Şekil EK-3.2.4. Symphon-Hetter Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

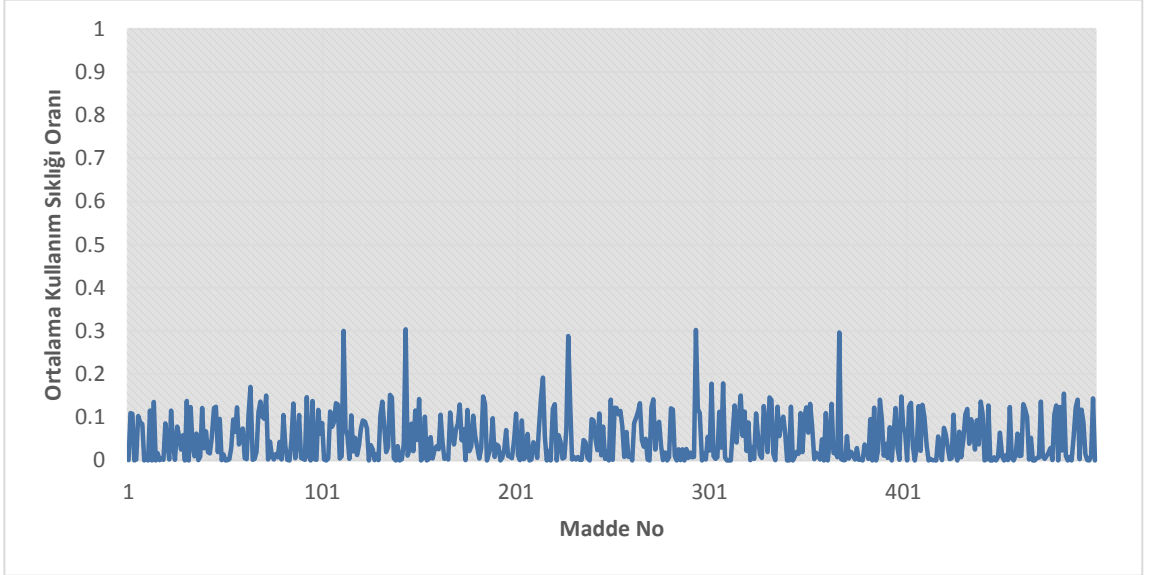


Şekil EK-3.2.5. Azalarak Kaybolma Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

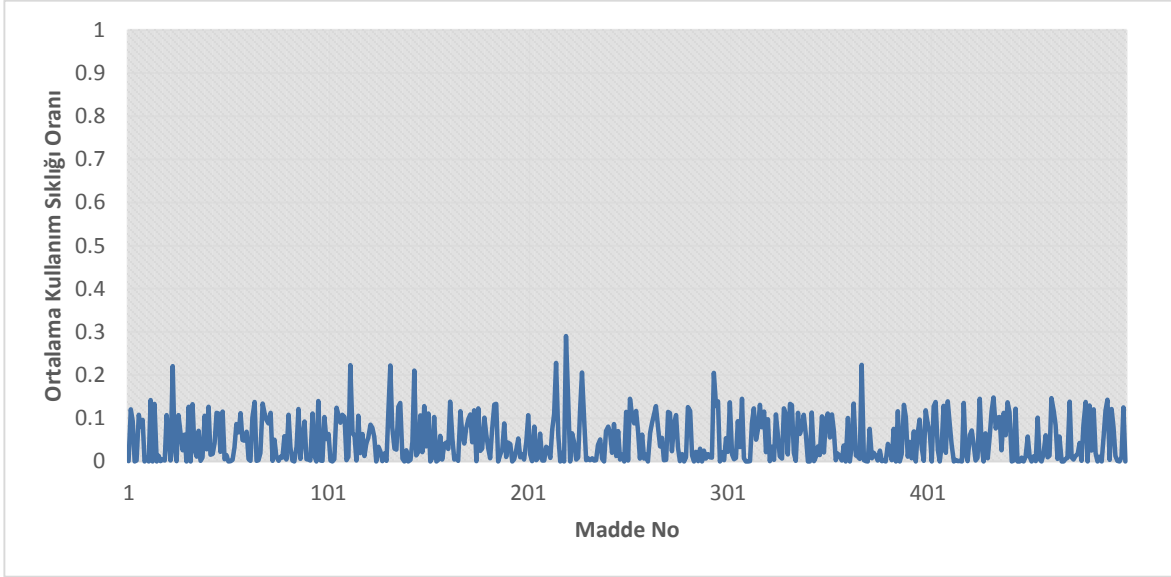
3. Üçüncü Alt Probleme Ait Madde Kullanım Sıklığı Oranı Grafikleri



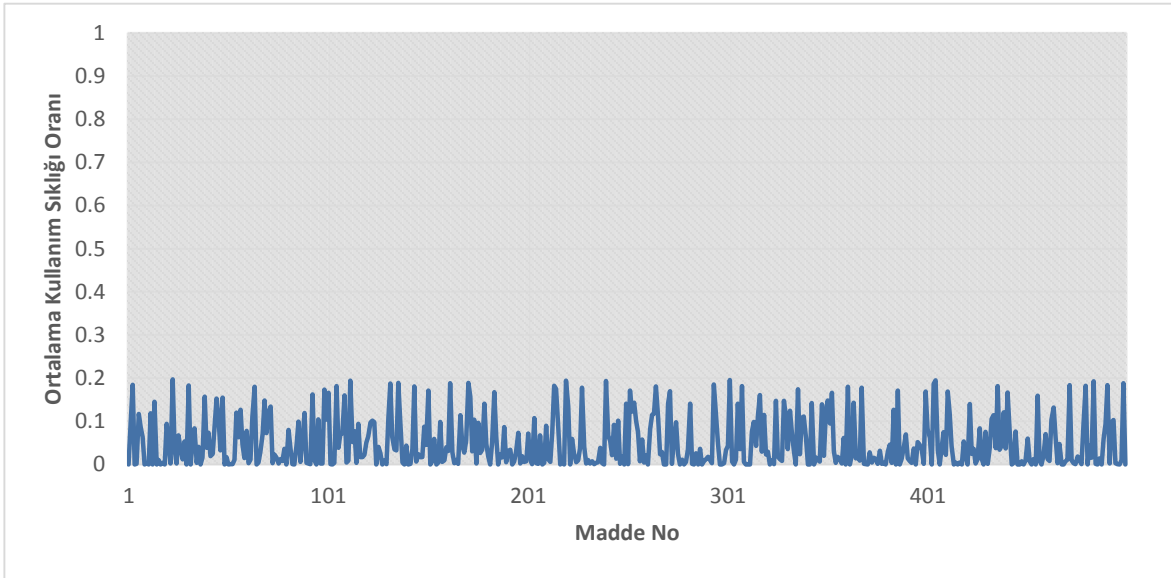
Şekil EK-3.3.1. Madde Kullanım Sıklığı Kontrolü Yapılmadığında



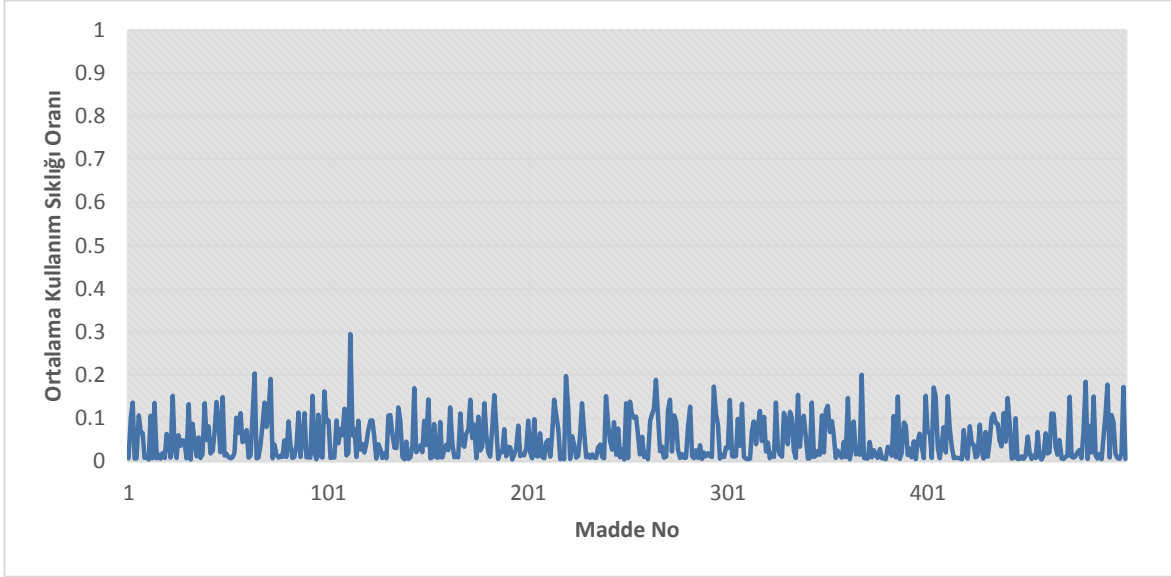
Şekil EK-3.3.2. BKT-5 Madde Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında



Şekil EK-3.3.3. BKT-10 Madde Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

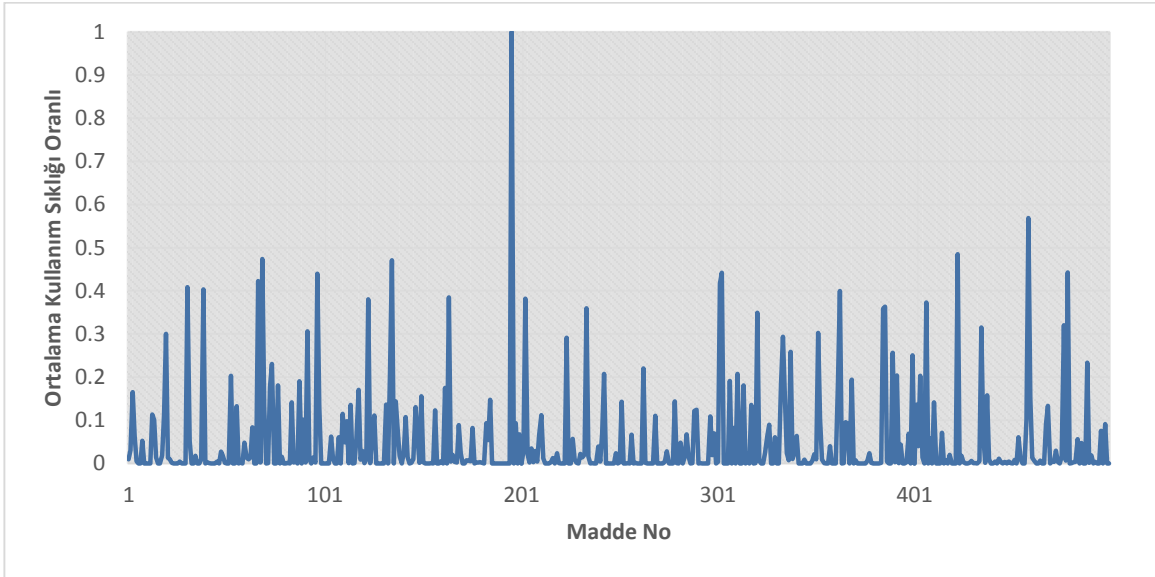


Şekil EK-3.3.4. Sympton-Hetter Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

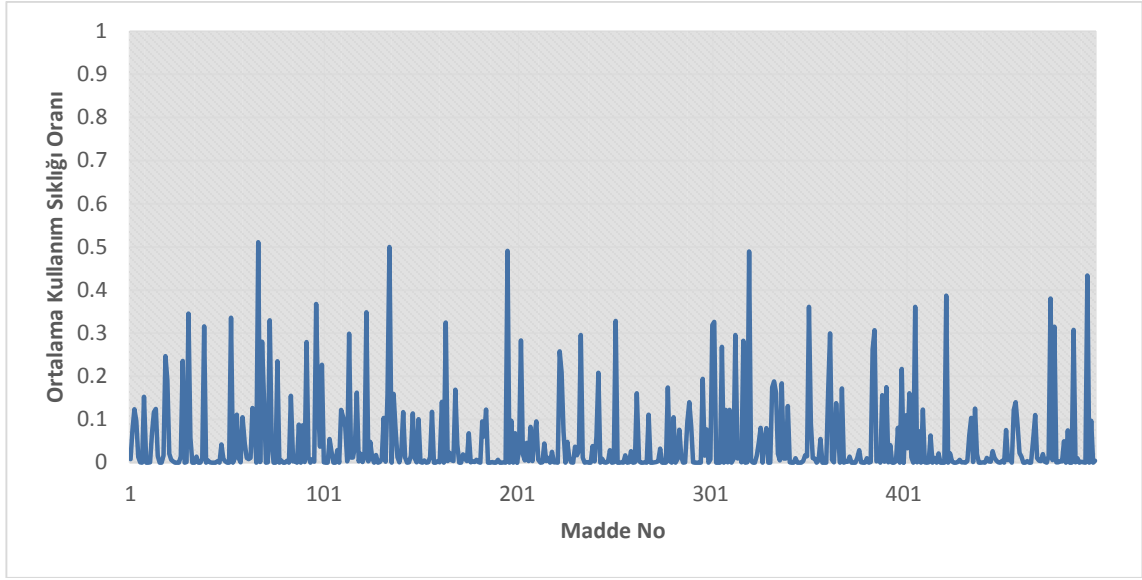


Şekil EK-3.3.5. Azalarak Kaybolma Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

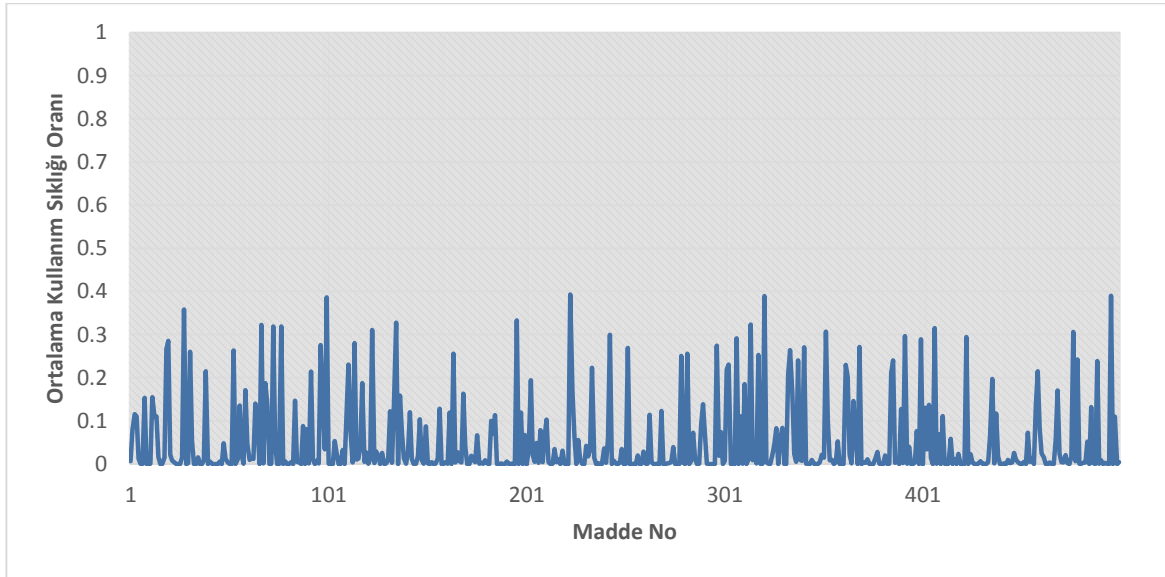
4. Dördüncü Alt Probleme Ait Madde Kullanım Sıklığı Oranı Grafikleri



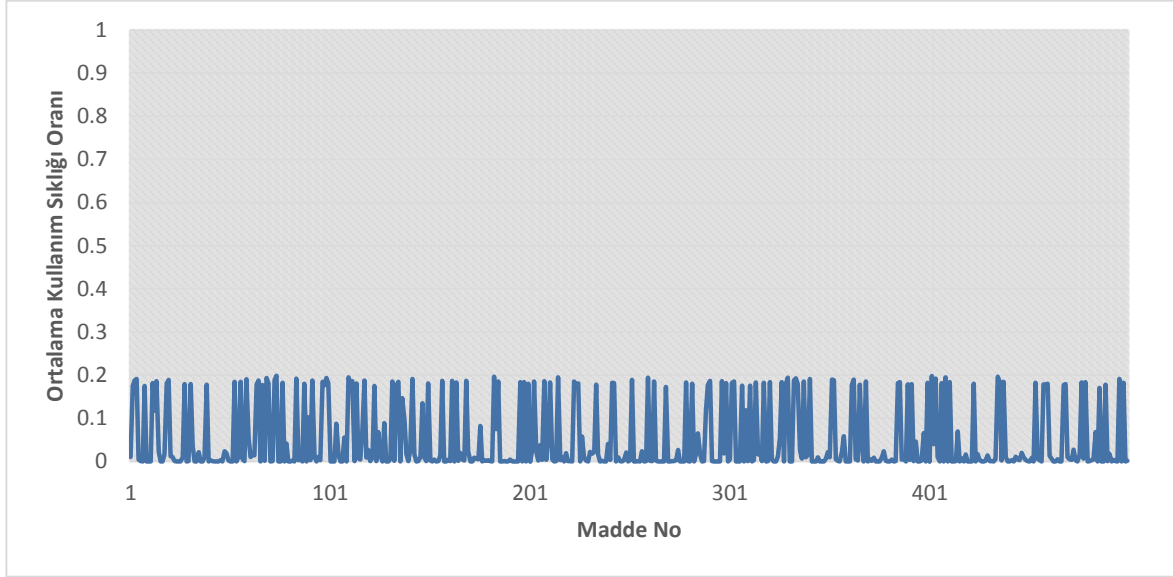
Şekil EK-3.4.1. Madde Kullanım Sıklığı Kontrolü Yapılmadığında



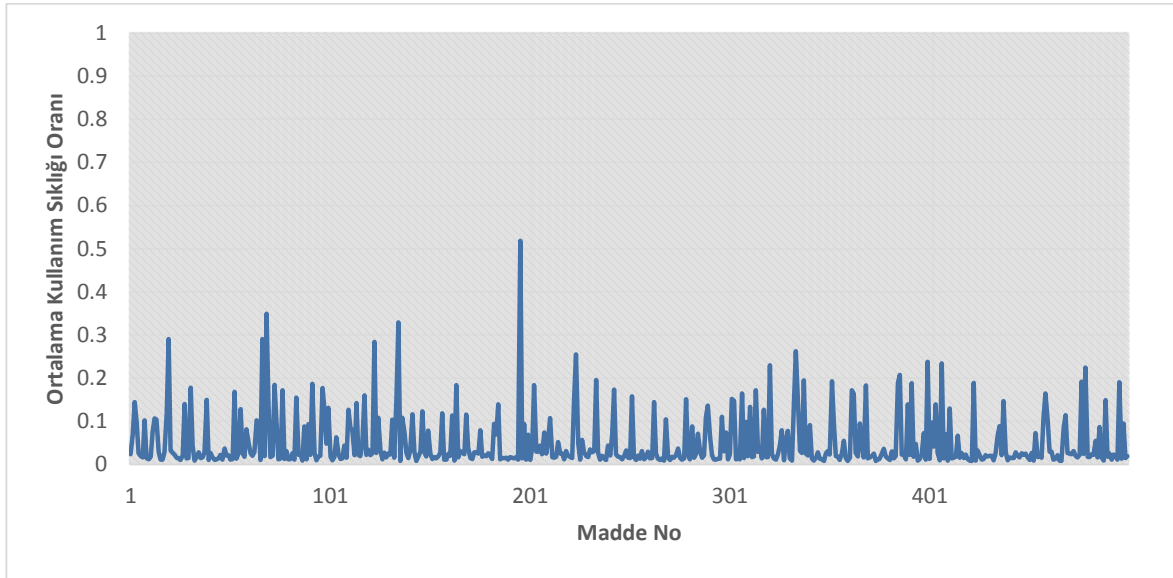
Şekil EK-3.4.2. BKT-5 Madde Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında



Şekil EK-3.4.3. BKT-10 Madde Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

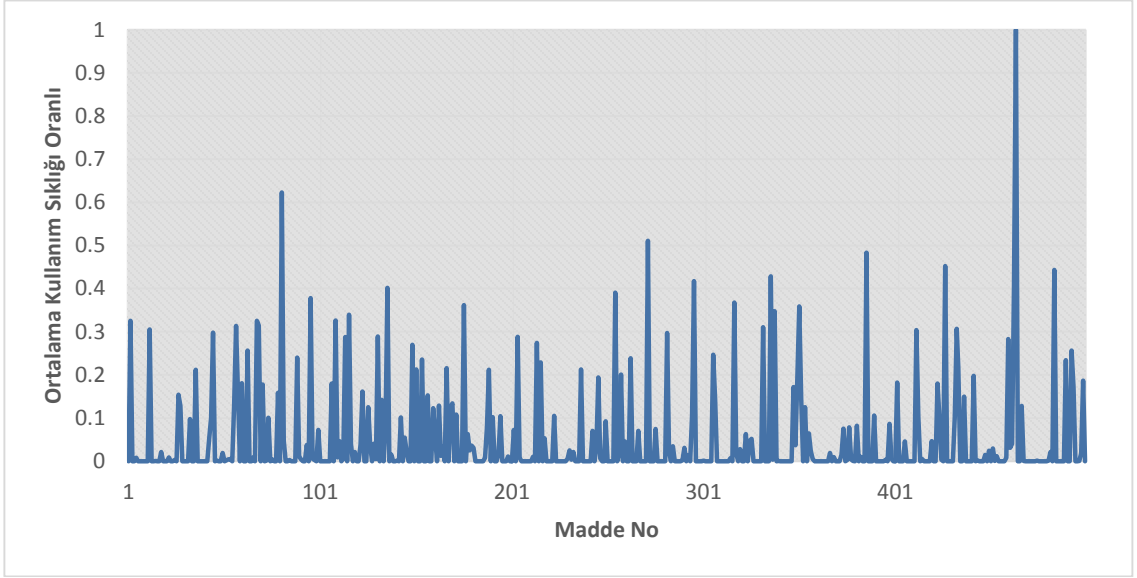


Şekil EK-3.4.4. Symptom-Hetter Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

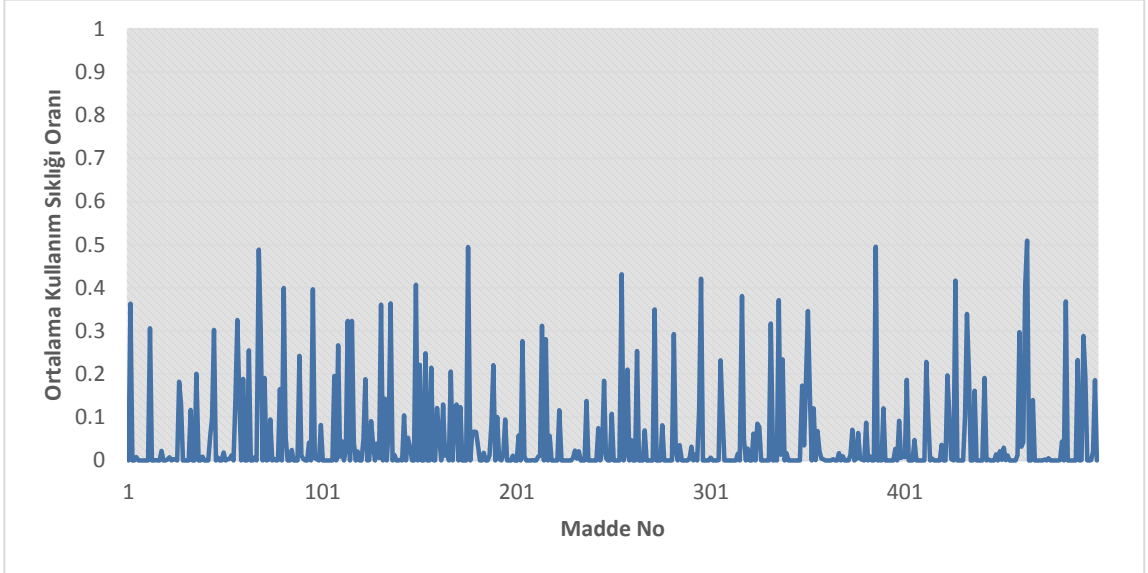


Şekil EK-3.4.5. Azalarak Kaybolma Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

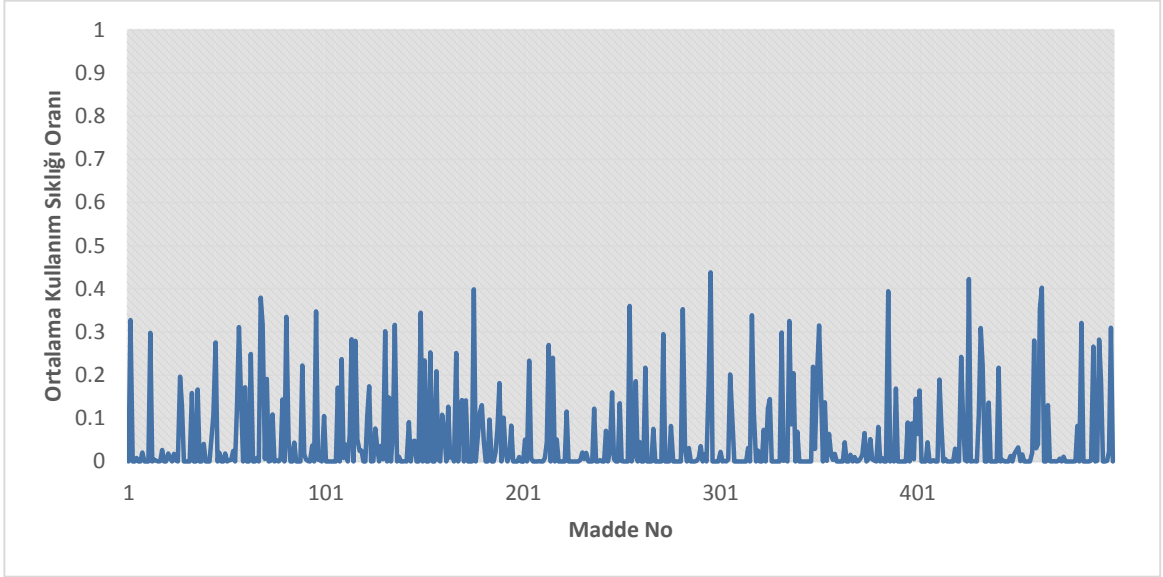
5. Beşinci Alt Probleme Ait Madde Kullanım Sıklığı Oranı Grafikleri



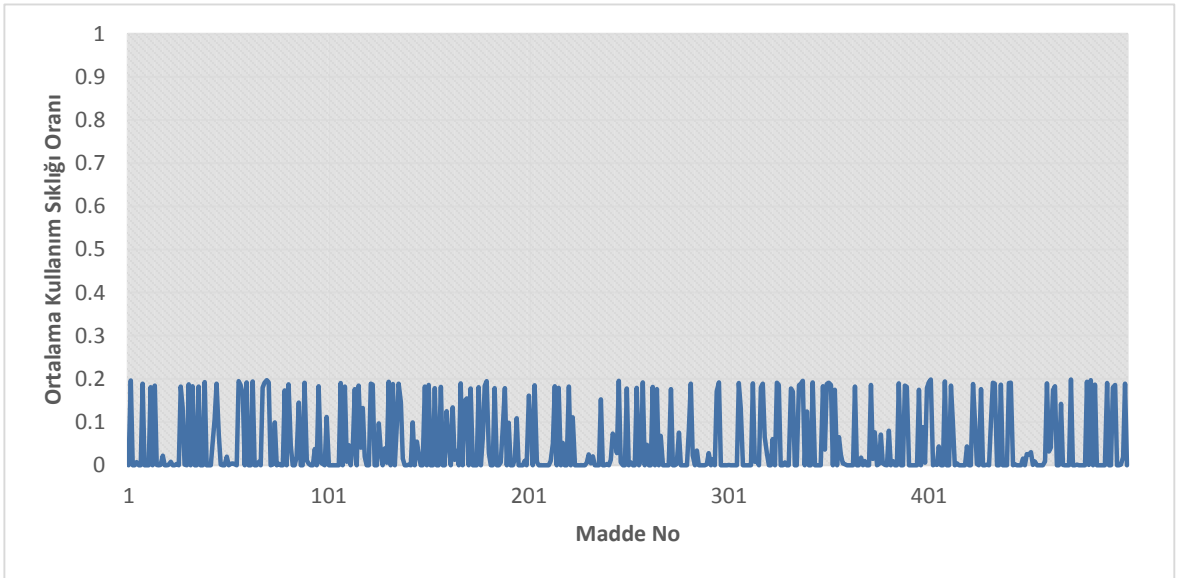
Şekil EK-3.5.1. Madde Kullanım Sıklığı Kontrolü Yapılmadığında



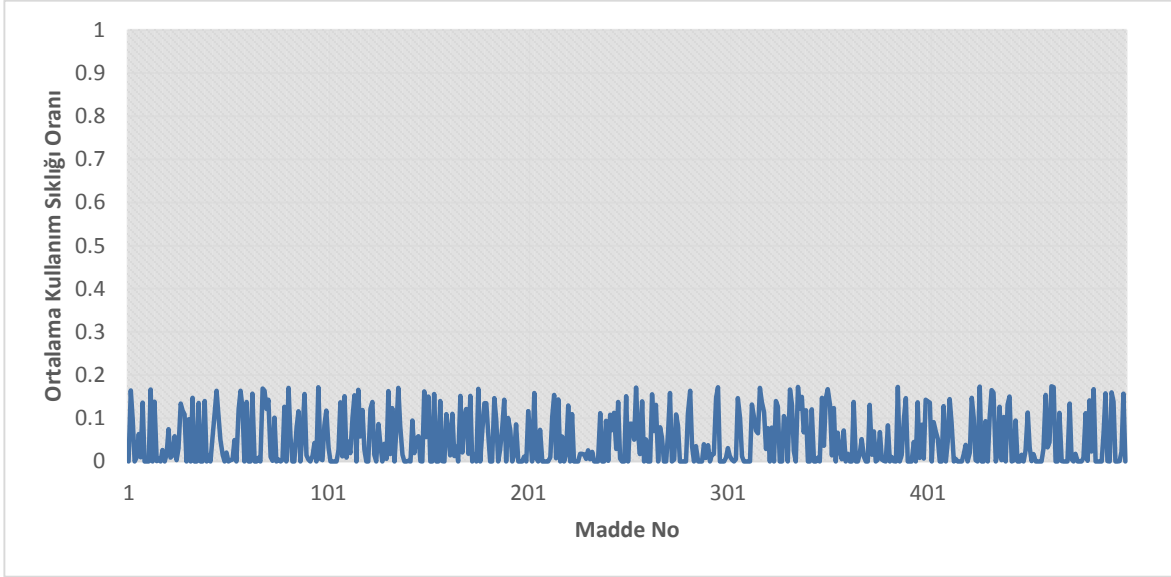
Şekil EK-3.5.2. BKT-5 Madde Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında



Şekil EK-3.5.3. BKT-10 Madde Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

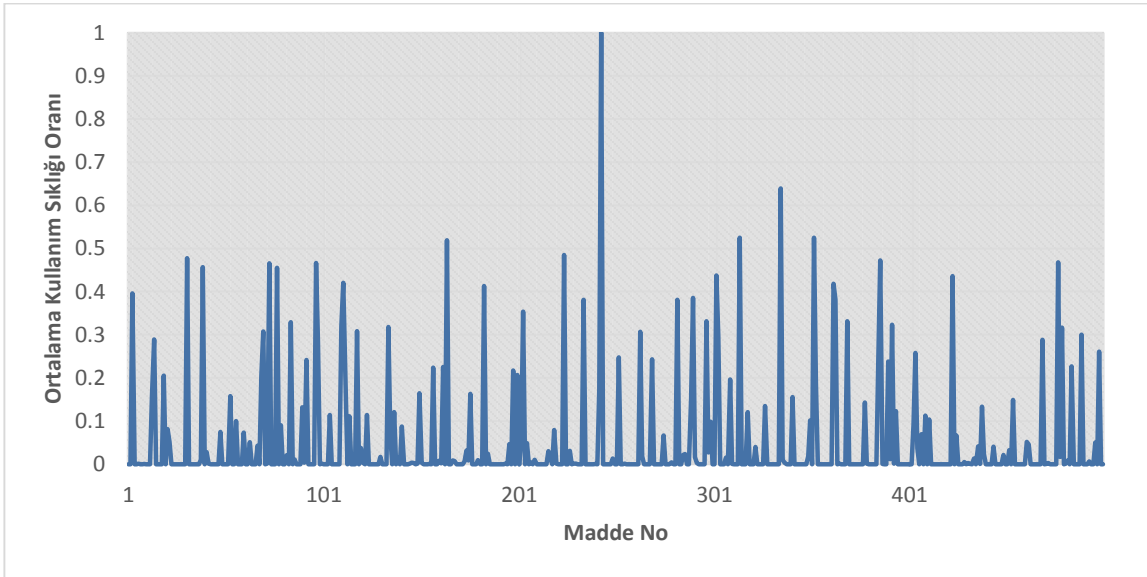


Şekil EK-3.5.4. Sympton-Hetter Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

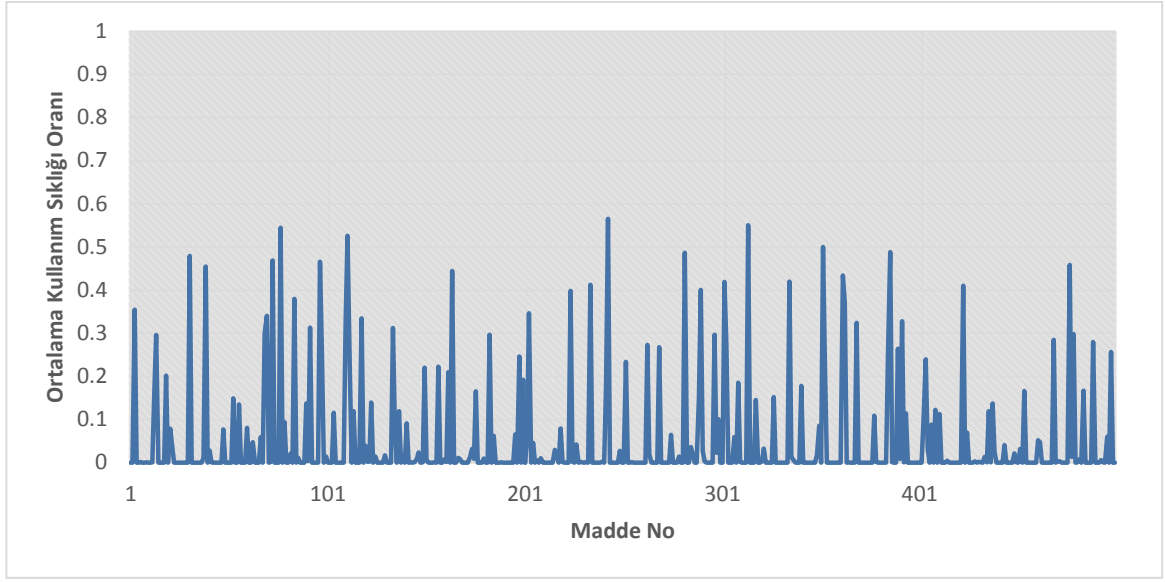


Şekil EK-3.5.5. Azalarak Kaybolma Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

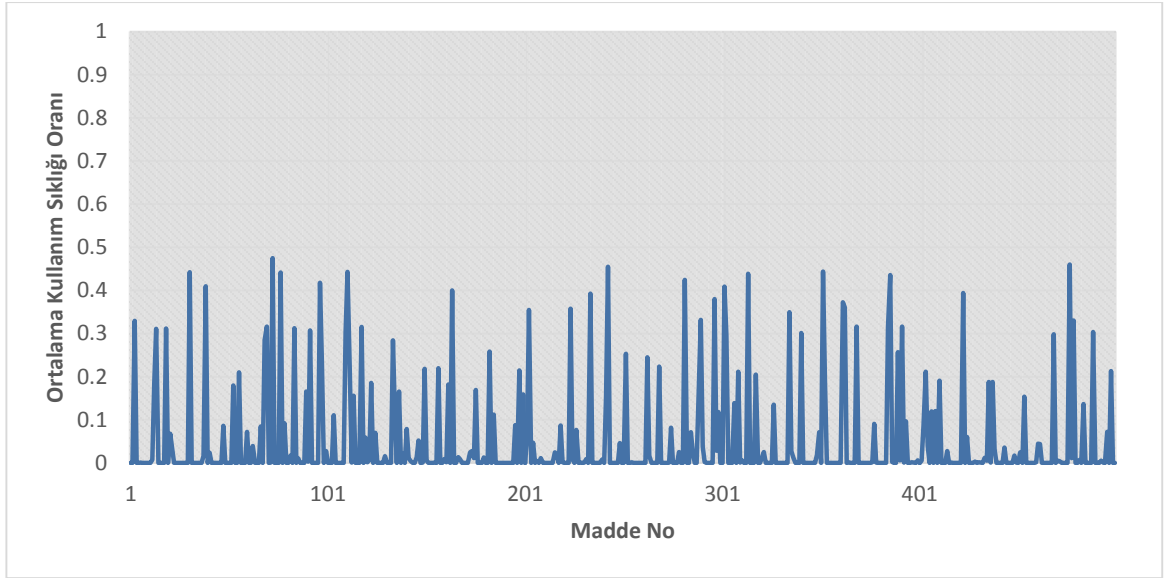
6. Altıncı Alt Probleme Ait Madde Kullanım Sıklığı Oranı Grafikleri



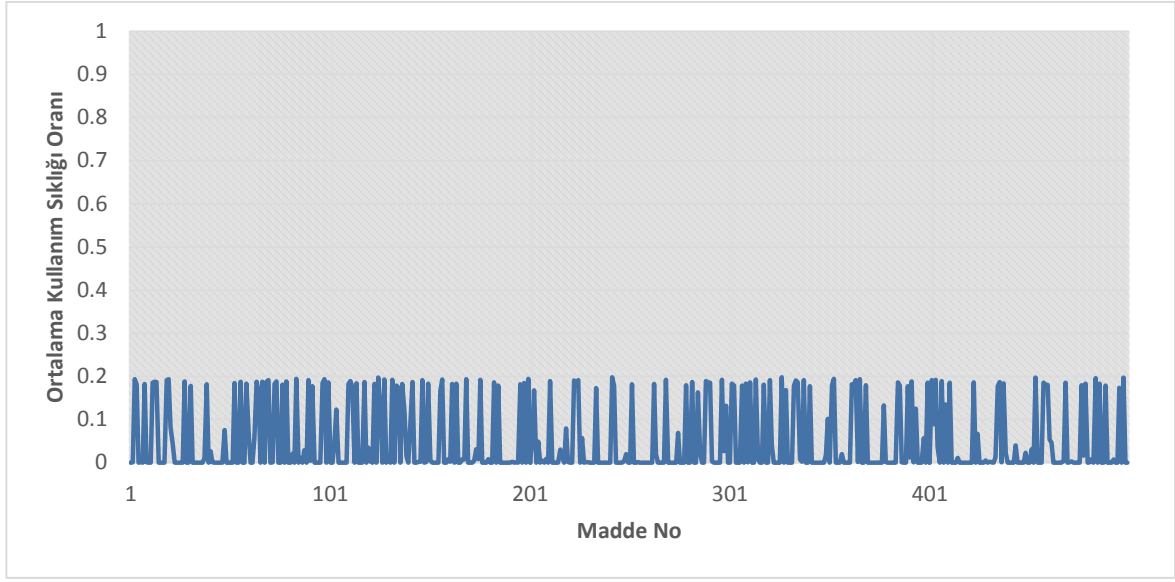
Şekil EK-3.6.1. Madde Kullanım Sıklığı Kontrolü Yapılmadığında



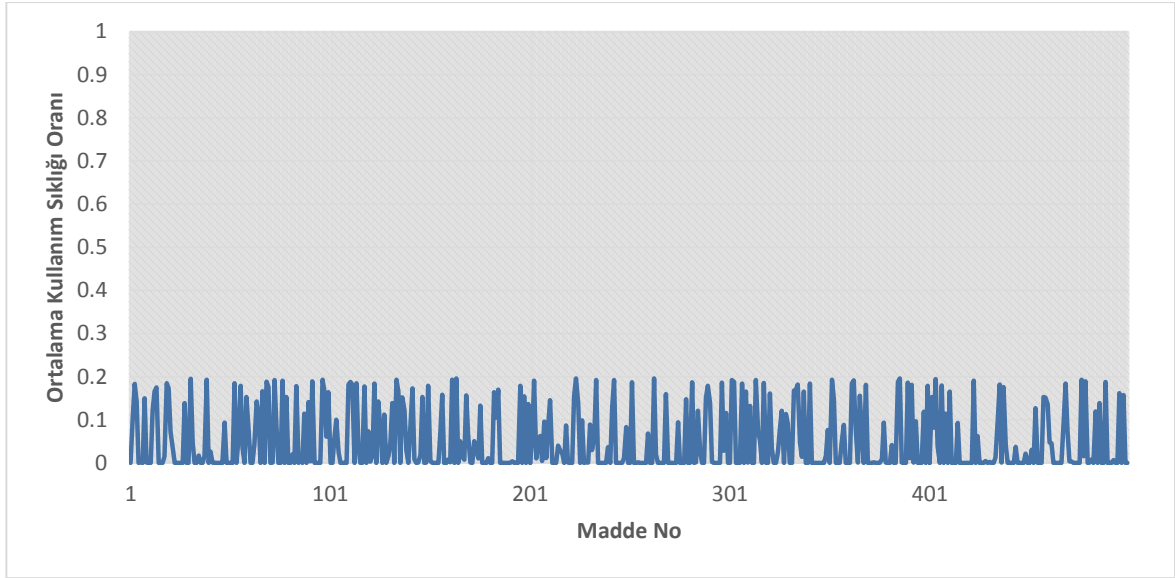
Şekil EK-3.6.2. BKT-5 Madde Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında



Şekil EK-3.6.3. BKT-10 Madde Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında



Şekil EK-3.6.4. Symptom-Hetter Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında



Şekil EK-3.6.5. Azalarak Kaybolma Madde Kullanım Sıklığı Kontrol Yöntemi Kullanıldığında

EK-4: Madde Havuzu Güçlüğü Orta Düzeyde Olduğunda Madde Seçme Yöntemleri Bazında Madde Kullanım Sıklığı Kontrol Yöntemlerinin Ölçme Kesinliği Açısından Karşılaştırılma Tablosu*

<i>Madde Seçme Yöntemi</i>	<i>Madde Kullanım Sıklığı Kontrol Yöntemi</i>	<i>Uyum</i>	<i>RMSE</i>	<i>Yanlılık</i>	<i>OMF</i>
<i>Fisher'in En Çok Bilgi Yöntemi</i>	Kontrol Olmadığında	0.9853	0.1748	-0.0026	0.1348
	Bilgi Koşullu Tesadüfi (5 Madde)	0.9851	0.1758	-0.0028	0.1356
	Bilgi Koşullu Tesadüfi (10 Madde)	0.9845	0.1793	-0.0025	0.1380
	Sympson-Hetter	0.9826	0.1900	-0.0037	0.1469
	Azalarak Kaybolma	0.9823	0.1918	-0.0013	0.1486
<i>a-Tabakalama Yöntemi</i>	Kontrol Olmadığında	0.9769	0.2186	-0.0022	0.1696
	Bilgi Koşullu Tesadüfi (5 Madde)	0.9765	0.2209	-0.0029	0.1712
	Bilgi Koşullu Tesadüfi (10 Madde)	0.9755	0.2250	-0.0019	0.1740
	Sympson-Hetter	0.9769	0.2188	-0.0012	0.1698
	Azalarak Kaybolma	0.9756	0.2248	-0.0016	0.1748
<i>Aşamalı En Çok Bilgi Oranı Yöntemi</i>	Kontrol Olmadığında	0.9854	0.1744	-0.0024	0.1346
	Bilgi Koşullu Tesadüfi (5 Madde)	0.9851	0.1763	-0.0034	0.1354
	Bilgi Koşullu Tesadüfi (10 Madde)	0.9845	0.1797	-0.0028	0.1381
	Sympson-Hetter	0.9830	0.1880	-0.0023	0.1453
	Azalarak Kaybolma	0.9823	0.1919	-0.0029	0.1486

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

EK-5: Madde Havuzu Güçlüğü Orta Düzeyde Olduğunda Madde Seçme Yöntemleri Bazında Madde Kullanım Sıklığı Kontrol Yöntemlerinin Test Güvenliği Açısından Karşılaştırılma Tablosu*

Madde Kullanım Sıklığı Oranı (ko)	Fisher'in En Çok Bilgi Yöntemi					a-Tabakalama Yöntemi					Aşamalı En Çok Bilgi Oranı Yöntemi				
	KOD	BKT5	BKT10	SH	AK	KOD	BKT5	BKT10	SH	AK	KOD	BKT5	BKT10	SH	AK
$0.9 < ko \leq 1.0$	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0
$0.8 < ko \leq 0.9$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$0.7 < ko \leq 0.8$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$0.6 < ko \leq 0.7$	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0
$0.5 < ko \leq 0.6$	1	5	0	0	0	0	0	0	0	0	1	1	0	0	0
$0.4 < ko \leq 0.5$	7	6	8	0	0	2	0	0	0	0	7	8	3	0	0
$0.3 < ko \leq 0.4$	14	15	16	0	0	1	2	0	0	0	16	18	20	0	0
$0.2 < ko \leq 0.3$	25	24	21	1	0	6	3	9	0	3	23	22	27	0	0
$0.1 < ko \leq 0.2$	41	36	44	121	134	83	109	108	106	87	40	39	44	122	133
$ko = 0$	293	270	242	250	190	106	48	42	68	0	292	270	241	243	168
SS_{ko}	0.110	0.105	0.099	0.076	0.06	0.080	0.055	0.053	0.059	0.049	0.109	0.102	0.094	0.075	0.059
En Yüksek ko	1.000	0.565	0.469	0.201	0.173	1.000	0.304	0.290	0.197	0.295	1.000	0.509	0.438	0.199	0.174
% ($ko = 0$)	58.6	54	48.4	50	38	21.2	9.6	8.4	13.6	0	58.4	54	48.2	48.6	33.6
χ^2	120.569	110.246	98.229	57.647	35.814	63.698	30.619	27.947	34.285	23.865	119.777	104.533	88.720	56.913	35.193
Test Çıkışma Oranı	0.290	0.270	0.246	0.164	0.121	0.177	0.110	0.105	0.118	0.097	0.289	0.258	0.227	0.163	0.120

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

** Madde kullanım sıklığı oranlarının ortalaması koşulların tümünde 0.05 olarak bulunmuştur.

***KOD: Kontrolün Olmadığı Durum, BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympson-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

EK-6: Madde Havuzu Güçlüğü Yüksek Düzeyde Olduğunda Madde Seçme Yöntemleri Bazında Madde Kullanım Sıklığı Kontrol Yöntemlerinin Ölçme Kesinliği Açısından Karşılaştırılma Tablosu*

<i>Madde Seçme Yöntemi</i>	<i>Madde Kullanım Sıklığı Kontrol Yöntemi</i>	<i>Uyum</i>	<i>RMSE</i>	<i>Yanlılık</i>	<i>OMF</i>
Fisher'in En Çok Bilgi Yöntemi	Kontrol Olmadığında	0.9773	0.2168	-0.0008	0.1627
	Bilgi Koşullu Tesadüfi (5 Madde)	0.9769	0.2185	-0.0020	0.1642
	Bilgi Koşullu Tesadüfi (10 Madde)	0.9756	0.2247	-0.0010	0.1679
	Sympson-Hetter	0.9637	0.2734	0.0007	0.2040
	Azalarak Kaybolma	0.9693	0.2518	-0.0009	0.1902
a-Tabakalama Yöntemi	Kontrol Olmadığında	0.9713	0.2436	-0.0010	0.1862
	Bilgi Koşullu Tesadüfi (5 Madde)	0.9686	0.2546	-0.0019	0.1930
	Bilgi Koşullu Tesadüfi (10 Madde)	0.9645	0.2704	-0.0028	0.2043
	Sympson-Hetter	0.9614	0.2816	-0.0044	0.2128
	Azalarak Kaybolma	0.9584	0.2919	-0.0019	0.2164
Aşamalı En Çok Bilgi Oranı Yöntemi	Kontrol Olmadığında	0.9775	0.2161	-0.0013	0.1623
	Bilgi Koşullu Tesadüfi (5 Madde)	0.9766	0.2202	-0.0028	0.1653
	Bilgi Koşullu Tesadüfi (10 Madde)	0.9757	0.2243	-0.0019	0.1684
	Sympson-Hetter	0.9649	0.2688	-0.0005	0.2006
	Azalarak Kaybolma	0.9689	0.2533	-0.0018	0.1904

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

EK-7: Madde Havuzu Güçlüğü Yüksek Düzeyde Olduğunda Madde Seçme Yöntemleri Bazında Madde Kullanım Sıklığı Kontrol Yöntemlerinin Test Güvenliği Açısından Karşılaştırılma Tablosu*

Madde Kullanım Sıklığı Oranı (ko)	Fisher'in En Çok Bilgi Yöntemi					a-Tabakalama Yöntemi					Aşamalı En Çok Bilgi Oranı Yöntemi				
	KOD	BKT5	BKT10	SH	AK	KOD	BKT5	BKT10	SH	AK	KOD	BKT5	BKT10	SH	AK
$0.9 < ko \leq 1.0$	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0
$0.8 < ko \leq 0.9$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$0.7 < ko \leq 0.8$	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$0.6 < ko \leq 0.7$	1	2	0	0	0	0	0	0	0	0	1	0	0	0	0
$0.5 < ko \leq 0.6$	4	8	6	0	0	1	2	0	0	1	3	4	0	0	0
$0.4 < ko \leq 0.5$	12	8	10	0	0	10	3	0	0	0	13	15	14	0	0
$0.3 < ko \leq 0.4$	17	15	19	0	0	14	17	15	0	2	16	12	23	0	0
$0.2 < ko \leq 0.3$	15	16	15	0	0	14	18	35	0	10	19	23	19	0	0
$0.1 < ko \leq 0.2$	25	29	29	129	128	45	50	45	122	72	24	25	25	128	124
$0.0 < ko \leq 0.1$	106	116	132	103	138	210	264	267	225	415	102	117	131	99	142
$ko = 0$	319	306	289	268	234	205	146	138	153	0	321	304	288	273	234
SS_{ko}	0.123	0.119	0.115	0.078	0.070	0.107	0.094	0.088	0.076	0.062	0.121	0.115	0.109	0.079	0.070
En Yüksek ko	1.000	0.652	0.565	0.197	0.195	1.000	0.511	0.393	0.199	0.518	1.000	0.565	0.475	0.198	0.196
% ($ko = 0$)	63.8	61.2	57.8	53.6	46.8	41	29.2	27.6	30.6	0	64.2	60.8	57.6	54.6	46.8
χ^2	150.734	141.826	131.645	61.215	48.810	114.271	88.565	78.081	58.191	38.808	146.733	132.396	117.927	61.647	49.482
Test Çakışma Oranı	0.351	0.333	0.313	0.172	0.147	0.278	0.226	0.205	0.166	0.127	0.343	0.314	0.285	0.172	0.148

*Değerlerin tümü 25 tekrardan elde edilen verilerin ortalaması hesaplanarak elde edilmiştir.

**Madde kullanım sıklığı oranlarının ortalaması koşulların tümünde 0.05 olarak bulunmuştur.

***KOD: Kontrolün Olmadığı Durum, BKT5: Bilgi Koşullu Tesadüfi (5 Madde)Yöntemi, BKT10: Bilgi Koşullu Tesadüfi (10 Madde) Yöntemi, SH: Sympson-Hetter Yöntemi, AK: Azalarak Kaybolma Yöntemi

ÖZGEÇMİŞ

Adı Soyadı	Nagihan BOZTUNÇ ÖZTÜRK
Doğum Yeri	KIRKLARELİ
Doğum Yılı	1985
Medeni Hali	Evli

Eğitim ve Akademik Durumu

Lise	Babaeski Anadolu Lisesi, Kırklareli	1996-2003
Lisans	Marmara Üniversitesi Atatürk Eğitim Fakültesi Fizik Öğretmenliği	2003-2008
Yüksek Lisans	Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı	2008-2010
Doktora	Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü Eğitim Bilimleri Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim Dalı	2010-2014
Yabancı Dil	İngilizce	ÜDS 82.5
İş Deneyimi	Hacettepe Üniversitesi Öğrenci İşleri Daire Başkanlığı Görev: Eğitim Öğretim Planlamacısı	2013- (devam ediyor)
	SEBİT Eğitim ve Bilgi Teknolojileri Görev: Ölçme ve Değerlendirme Uzmanı	2010-2013