



**BECERİ VE KAVRAM YANILGILARININ TANIMLANMASINDA
BİLİŞSEL TANI MODELİ: SISM MODEL PERFORMANSININ
DEĞERLENDİRİLMESİ VE DINA, BUG-DINO MODELLERİ İLE
KARŞILAŞTIRILMASI**

Büşra Eren

DOKTORA TEZİ

EĞİTİM BİLİMLERİ ANA BİLİM DALI

GAZİ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

MAYIS, 2023

TELİF HAKKI VE TEZ FOTOKOPİ İZİN FORMU

Bu tezin tüm hakları saklıdır. Kaynak göstermek koşuluyla tezin teslim tarihinden itibaren 24 (yirmi dört) ay sonra tezden fotokopi çekilebilir.

YAZARIN

Adı : Büşra

Soyadı : EREN

Bölümü : Eğitimde Ölçme ve Değerlendirme

İmza :

Teslim tarihi :

TEZİN

Türkçe Adı : Beceri ve Kavram Yanılgılarının Tanımlanmasında Bilişsel Tanı Modeli: SISM Model Performansının Değerlendirilmesi ve DINA, BUG-DINO Modelleri ile Karşılaştırılması

İngilizce Adı : Cognitive Diagnostic Model for Identifying Skills and Misconceptions: Evaluation of SISM Model Performance and Comparison with DINA, BUG-DINO Models

ETİK İLKELERE UYGUNLUK BEYANI

Tez yazma sürecinde bilimsel ve etik ilkelere uyduđumu, yararlandıđım tüm kaynakları kaynak gösterme ilkelerine uygun olarak kaynakçada belirttiđimi ve bu bölümler dışındaki tüm ifadelerin şahsıma ait olduđunu beyan ederim.

Yazar Adı Soyadı: Büşra EREN

İmza:

JÜRİ ONAY SAYFASI

Büşra EREN tarafından hazırlanan “Becerî ve Kavram Yanılgılarının Tanımlanmasında Bilişsel Tamı Modeli: SISM Model Performansının Değerlendirilmesi ve DINA, BUG-DINO Modelleri ile Karşılaştırılması” adlı tez çalışması aşağıdaki jüri tarafından oy birliği ile Gazi Üniversitesi Eğitim Bilimleri Anabilim Dalı’nda Doktora tezi olarak kabul edilmiştir.

Danışman: (Prof. Dr. Şeref TAN)

(Eğitimde Ölçme ve Değerlendirme, Gazi Üniversitesi)

Başkan: (Prof. Dr. İsmail KARAKAYA)

(Eğitimde Ölçme ve Değerlendirme, Gazi Üniversitesi)

Üye: (Prof. Dr. Nilüfer KAHRAMAN)

(Eğitimde Ölçme ve Değerlendirme, Gazi Üniversitesi)

Üye: (Prof. Dr. Duygu ANIL)

(Eğitimde Ölçme ve Değerlendirme, Hacettepe Üniversitesi)

Üye: (Dr. Öğr. Üyesi Mahmut Sami KOYUNCU)

(Eğitimde Ölçme ve Değerlendirme, Afyon Kocatepe Üniversitesi)

Tez Savunma Tarihi: 05/05/2023

Bu tezin Eğitim Bilimleri Anabilim Dalı’nda Doktora tezi olması için şartları yerine getirdiğini onaylıyorum.

Prof. Dr. Şaban ÇETİN

Eğitim Bilimleri Enstitüsü Müdürü



27.02.2023'te kaybettiğim sevgili babaannem Sevim KARADUMAN'a...

TEŞEKKÜR

Doktora eğitimim boyunca akademik olarak gelişmemde değerli katkılar sunan, birlikte çalışmaktan onur duyduğum değerli danışmanım ve kıymetli hocam Prof. Dr. Şeref TAN'a, Tezimin şekillenmesinde bilgi ve görüşleri ile destek olan, tez izleme komitemde ve tez savunma jürimde yer alan değerli hocalarım Prof. Dr. Nilüfer KAHRAMAN ve Prof. Dr. Duygu ANIL'a,

Tez savunma jürime katılarak görüş ve önerileriyle tezime önemli katkılar sağlamalarından mutluluk duyduğum değerleri hocalarım Prof. Dr. İsmail KARAKAYA ve Dr. Öğr. Üyesi Mahmut Sami KOYUNCU'ya,

Tez süresince gerek sağlamış olduğu yardımlar gerek sormuş olduğum sorular noktasında nezaketini esirgemeyen Prof. Dr. Jimmy de la TORRE'a,

Tezimin niteliğine katkılar sunan kıymetli hocalarım ve sevgili meslektaşlarım Arş. Gör. Dr. Tuba GÜNDÜZ, Dr. Öğr. Üyesi Ayşenur ERDEMİR, Öğr. Gör. Ayça BEYHAN ve Dr. Semih AŞİRET'e,

Doktora öğrenimim boyunca uzaktan yürütmek durumunda kaldığım bu süreci destekleri ile kolaylaştıran sevgili arkadaşım Arş. Gör. Duygu Diğdem SİLİK ve Gazi Üniversitesi Eğitimde Ölçme ve Değerlendirme Bilim Dalı'nda görev yapan sevgili hocam Arş. Gör. Dr. Esra OYAR'a,

Bu süreci destekleriyle kolaylaştıran sevgili mesai arkadaşlarıma, özellikle sevgili Tuğba NİKSARLI ve Orçun DÜVEN'e,

Doktora öğrenimim boyunca 2211-Yurt İçi Doktora Burs Programı kapsamında maddi destekler sunan TÜBİTAK'a,

Yaşamımın her alanında maddi manevi tüm desteđi üzerimden eksik etmeyen beni bugünlere getiren annem Aysun KARADUMAN, babam İsmet KARADUMAN ve sevgili kardeşim Can KARADUMAN'a,

Öğrencilik hayatımı tamamladığım yıllara yayılmış zorlu bir sürecin üstesinden gelmemde her zaman sabır ve hoşgörüsüyle güç verip doktora sürecimi olabildiğince kolaylaştıran kıymetli eşim Alptekin EREN'e,

Desteklerini esirgemeyen, daima yanımda olan adını sayamadığım sevgili arkadaşlarıma ve bugünlere gelmemde üzerimde emeđi olan tüm hocalarıma teşekkürü bir borç bilirim.

Büşra EREN

**BECERİ VE KAVRAM YANILGILARININ TANIMLANMASINDA
BİLİŞSEL TANI MODELİ: SISM MODEL PERFORMANSININ
DEĞERLENDİRİLMESİ VE DINA, BUG-DINO MODELLERİ İLE
KARŞILAŞTIRILMASI**

(Doktora Tezi)

Büşra Eren

GAZİ ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

Mayıs 2023

ÖZ

Bu çalışmanın amacı, bilişsel tanı modellerinde beceri ve kavram yanılıgılarını eş zamanlı olarak ölçen modelin performansını farklı koşullar altında karşılaştırarak modelin hangi koşullar altında benzer ya da farklı sonuçlar verdiğini göstermek ve modeli beceriler ile kavram yanılıgılarını ayrı ayrı tanımlayan modellerin performanslarıyla karşılaştırmaktır. Becerileri ve kavram yanılıgılarını eş zamanlı olarak ölçen modelin performansına ilişkin kapsamlı bilgilere ulaşmak amacıyla örneklem büyüklüğü, test uzunluğu, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısı olmak üzere farklı simülasyon çalışmalarında sıklıkla kullanılan dört faktör çalışmaya dahil edilmiştir. Olası bilgi durumlarına dayalı olarak ana simülasyon tasarımı modelin performansını dört faktörün birleştirilmesiyle 72 simülasyon koşulu altında karşılaştırmıştır. Çalışmanın birinci ve ikinci alt problemleri kapsamında modele ilişkin madde parametre tahminlerine ait “mutlak sapma” ve “ortalama mutlak sapma” değerleriyle birlikte model kapsamında yer alan dört madde parametresine ilişkin ortalama standart hata değerleri tüm koşullar için özetlenerek sunulmuştur. Çalışmanın üçüncü alt problemi kapsamında ise becerileri ve kavram yanılıgılarını eş zamanlı olarak

ölçen modelin sınıflama doğruluğu performansı, beceri ve kavram yanılgılarını ayrı ayrı ölçen modellerle karşılaştırılmıştır. Birinci ve ikinci alt problemlerden elde edilen bulgular test uzunluğuyla örneklem büyüklüğünün nispeten arttığı ve madde ayırt edicilik düzeyinin yüksek olduğu koşullar altında dört madde parametresine ait tahminlerin beklenti maksimizasyon algoritmasıyla doğru bir şekilde tahmin edilebildiğini göstermiştir. Üçüncü alt problemde elde edilen bulgular ise simülasyon çalışması sonucu beceri ve kavram yanılgılarını eş zamanlı olarak ölçen modelin dört faktörde değişimlenen koşullarda diğer modellere göre daha yüksek sınıflama doğruluğu performansı sergilediğini ortaya koymuştur. Bu modeli sırasıyla kavram yanılgılarının ölçen model ve becerileri ölçen model takip etmiştir. Beceri ve kavram yanılgılarını eş zamanlı olarak ölçen modelin kısa testlerin kullanıldığı koşullarda dahi kabul edilebilir sınıflama doğruluğu değerleri aldığı ve madde ayırt edicilik düzeyi arttıkça sınıflama doğruluğu değerlerinin arttığı görülmüştür. Ancak ölçülen nitelik sayısının artmasıyla birlikte modelin sınıflama doğruluğu performansı daha az karmaşık olan Q-matrislerin kullanıldığı koşullara kıyasla azalmıştır. Beceri ve kavram yanılgılarının eş zamanlı olarak ölçülmesi için tasarlanmış modele ait analizlerden elde edilen sonuçlar ümit verici olsa da, literatürde yeni bir model olarak yer aldığından elde edilen bulgulardan genellenebilir sonuçlara ulaşamamıştır. Bu nedenle modelin güvenilir kullanımları sağlanmadan önce doğasını daha iyi anlamak adına simülasyon çalışmalarının tasarlanması ya da gerçek veri setlerinin analizleriyle ek çalışmaların yürütülmesine ihtiyaç vardır. Ancak bu sayede becerilerin ve kavram yanılgılarının bilişsel tanısı, eğitim öğretim reformlarına yardımcı olabilir.

Anahtar Kelimeler : Bilişsel Tanı Modelleri, SISM, Kavram Yanılgıları
Sayfa Adedi : xvii + 172
Danışman : Prof. Dr. Şeref TAN

**COGNITIVE DIAGNOSTIC MODEL FOR IDENTIFYING SKILLS
AND MISCONCEPTIONS: EVALUATION OF SISM MODEL
PERFORMANCE AND COMPARISON WITH DINA, BUG-DINO
MODELS (Ph.D Thesis)**

Büşra Eren

GAZI UNIVERSITY

GRADUATE SCHOOL OF EDUCATIONAL SCIENCES

May 2023

ABSTRACT

The aim of this study is to compare the performance of model, which can simultaneously measure skills and misconceptions in cognitive diagnosis models, under different conditions in order to determine which conditions the model produces similar or different results, and to compare model to the performance of models that define skills and misconceptions separately. Four factors, including sample size, test length, item discrimination level, and number of attributes measured, which are often utilized in various simulation studies, were included in the study to obtain comprehensive information on the model performance of simultaneously measure skills and misconceptions. The main simulation design compared the performance of model under 72 simulation conditions based on possible knowledge states by combining four factors. For all conditions within the scope of the study's first and second sub-problems, the "absolute deviation" and "mean absolute deviation" values of the item parameter estimations related to the model, as well as the mean standard error values of the four item parameters included in the model, were summarized. The classification accuracy performance of the model that measures skills and misconceptions simultaneously was compared with the models measuring skill and misconceptions separately in the scope of the study's third sub-problem. The results of the first and second sub-problems demonstrated that the expectation maximization algorithm could accurately estimate the four item parameters under conditions of relatively increased test length and sample size and high item discrimination level. The results of the third sub-problem demonstrated that model for

simultaneously identifying skills and misconceptions outperformed other models in classification accuracy under conditions that varied across four factors. This model was followed by the model measuring misconceptions and the model measuring skills, respectively. It has been observed that the model, which simultaneously measures skills and misconceptions, has acceptable classification accuracy values even under the conditions where short tests are used, and the classification accuracy values increase as the item discrimination level increases. However, as number of attributes measured increased, the model's classification accuracy performance deteriorated when compared to conditions where less complex Q-matrices were used. Although the results of the model designed to measure skills and misconceptions simultaneously are promising, generalizable results could not be reached from the findings, since it is a new model in the literature. As a result, before the model can be used reliably, simulation studies or additional studies involving the analysis of real data sets are required to better understand its nature. Only in this manner can cognitive diagnosis of skills and misconceptions support education reforms.



Key Words : Cognitive Diagnosis Models, SISM, Misconceptions
Page Number : xvii + 172
Supervisor : Prof. Dr. Şeref TAN

İÇİNDEKİLER

TELİF HAKKI VE TEZ FOTOKOPİ İZİN FORMU	i
ETİK İLKELERE UYGUNLUK BEYANI	ii
JÜRİ ONAY SAYFASI	iii
TEŞEKKÜR	iv
ÖZ	vii
ABSTRACT	ix
İÇİNDEKİLER	xi
TABLolar LİSTESİ	xiv
ŞEKİLLER LİSTESİ	xvi
SİMGELER VE KISALTMALAR LİSTESİ	xvii
BÖLÜM I	1
GİRİŞ	1
1.1. Problem Durumu	1
1.2. Araştırmanın Amacı	4
1.3. Araştırmanın Önemi	4
1.4. Problem Cümlesi	6
1.4.1. Alt problemler	6
1.5. Sınırlılıklar	7
1.6. Sayılıtlar	7
BÖLÜM II	8
KURAMSAL ÇERÇEVE	8
2.1. Araştırmanın Kuramsal Temelleri	8

2.1.1. Bilişsel Tanı Değerlendirmeleri	8
2.1.2. Nitelikler ve Hiyerarşi	12
2.1.3. Q-matris	15
2.1.4. Bilişsel Tanı Modelleri	17
2.1.4.1. DINA (Deterministic Input Noisy “and” Gate)	19
2.1.4.2. BUG-DINO (“BUG”- Deterministic Input Noisy “or” Gate)	21
2.1.4.3. SISM (Simultaneously Identifying Skills “and” Misconceptions Model)	22
2.1.4.3.1.SISM’e İlişkin Özel Durumlar	24
2.1.5. Kavram Yanılgıları	25
2.2. İlgili Araştırmalar	27
BÖLÜM III	35
YÖNTEM	35
3.1. Araştırmanın Modeli	35
3.2. Verilerin Üretimi	36
3.2.1. Simülasyon Koşulları	36
3.2.1.1. Örneklem Büyüklüğü	36
3.2.1.2. Test Uzunluğu	37
3.2.1.3. Madde Ayırt Edicilik Düzeyi	38
3.2.1.4. Ölçülen Nitelik Sayısı	39
3.3. Verilerin Analizi	46
3.3.1. Değerlendirme Kriteri	48
BÖLÜM IV	51
BULGULAR VE YORUM	51
4.1. Birinci Alt Probleme İlişkin Bulgular	51
4.2. İkinci Alt Probleme İlişkin Bulgular	70
4.3. Üçüncü Alt Probleme İlişkin Bulgular	79
BÖLÜM V	107
SONUÇ VE TARTIŞMA	107

5.1. Sonular	107
5.1.1. Birinci Alt Probleme İlişkin Sonular	108
5.1.2. İkinci Alt Probleme İlişkin Sonular	111
5.1.3. Üüncü Alt Probleme İlişkin Sonular	116
5.2. Öneriler	123
5.2.1. Uygulayıcılara Yönelik Öneriler	123
5.2.2. Araştırmacılara Yönelik Öneriler	124
KAYNAKLAR	128
EKLER	141
EK 1. 3 Beceri ve 3 Kavram Yanılgısı için Olası Nitelik Sınıfları	142
EK 2. 6 Nitelik için 20 Maddelik Q-matris	143
EK 3. 250 Kişilik Veri Seti için Örnek Olarak Oluşturulan Kod	144
EK 4. 8 ve 10 Nitelikli Q-matrisler için “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerlerine İlişkin Tablolar	157
EK 5. 6, 8 ve 10 Nitelikli Q-matrislerin Kullanıldığı Durumlarda Dört Madde Parametresine ait “Hataların Ortalama Karekökü” Değerleri	169
EK 6. 6, 8 ve 10 Nitelikli Q-matrisler için Faktörlerde Değişimlenen Her Bir Koşulda Dört Parametreye Ait Tüm Maddeler Arasındaki Ortalama Standart Hata Değerleri	170

TABLolar LİSTESİ

Tablo 1. Örnek Q-Matrisi	16
Tablo 2. Bilişsel Tanı Modellerinin Sınıflandırılması	17
Tablo 3. Bilişsel Tanı Değerlendirmelerinde Kavram Yanılgılarını İçeren Modellerin Sınıflandırılması	18
Tablo 4. SISM’de Bireylerin Farklı Başarı Olasılıklarına Ait Dört Madde Parametresi....	23
Tablo 5. 4 Kavram Yanılgısı ve 4 Beceri Arasındaki Hiyerarşik İlişki	42
Tablo 6. 4 Beceri ve 4 Kavram Yanılgısı için Olası Nitelik Sınıfları	43
Tablo 7. 8 Nitelik için 20 Maddelik Q-matris	45
Tablo 8. Simülasyon Çalışması için Değişimlenen Koşullar	47
Tablo 9. 6 Niteliği Ölçen Yüksek Ayırt Ediciliğe Sahip 20 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma Değerleri”	53
Tablo 10. 6 Niteliği Ölçen Düşük Ayırt Ediciliğe Sahip 20 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma Değerleri”	55
Tablo 11. 6 Niteliği Ölçen Yüksek Ayırt Ediciliğe Sahip 40 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma Değerleri”	57
Tablo 12. 6 Niteliği Ölçen Düşük Ayırt Ediciliğe Sahip 40 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma Değerleri”	59

Tablo 13. 6 Niteliği Ölçen Yüksek Ayırt Ediciliğe Sahip 60 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma Değerleri”	61
Tablo 14. 6 Niteliği Ölçen Düşük Ayırt Ediciliğe Sahip 60 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma Değerleri”	63
Tablo 15. 6 Nitelikli 20 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri	80
Tablo 16. 6 Nitelikli 40 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri	82
Tablo 17. 6 Nitelikli 60 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri	84
Tablo 18. 8 Nitelikli 20 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri	87
Tablo 19. 8 Nitelikli 40 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri	90
Tablo 20. 8 Nitelikli 60 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri	93
Tablo 21. 10 Nitelikli 20 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri	96
Tablo 22. 10 Nitelikli 40 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri	99
Tablo 23. 10 Nitelikli 60 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri	103

ŞEKİLLER LİSTESİ

Şekil 1. 3 boyutlu faktör analizi ve madde tepki kuramı modellerinin, 3 boyutlu bilişsel tanı modeli ile karşılaştırılması	10
Şekil 2. Nitelik hiyerarşileri.....	14
Şekil 3. 3 beceri ve 3 kavram yanılığısı için hiyerarşik nitelik yapısı	40
Şekil 4. 4 beceri ve 4 kavram yanılığısı için hiyerarşik nitelik yapısı	41
Şekil 5. 5 beceri ve 5 kavram yanılığısı için hiyerarşik nitelik yapısı	42
Şekil 6. 6 nitelikli Q-matriste dört madde parametresine ait farklı simülasyon koşullarından elde edilen “ortalama mutlak sapma” değerleri.....	65
Şekil 7. 8 nitelikli Q-matriste dört madde parametresine ait farklı simülasyon koşullarından elde edilen “ortalama mutlak sapma” değerleri.....	66
Şekil 8. 10 nitelikli Q-matriste dört madde parametresine ait farklı simülasyon koşullarından elde edilen “ortalama mutlak sapma” değerleri.....	67
Şekil 9. 6 nitelikli Q-matriste dört madde parametresine ait farklı simülasyon koşullarından elde edilen “hataların ortalama karekökü” değerleri	71
Şekil 10. 8 nitelikli Q-matriste dört madde parametresine ait farklı simülasyon koşullarından elde edilen “hataların ortalama karekökü” değerleri	72
Şekil 11. 10 nitelikli Q-matriste dört madde parametresine ait farklı simülasyon koşullarından elde edilen “hataların ortalama karekökü” değerleri	73

SİMGELER VE KISALTMALAR LİSTESİ

ACCR	Attribute Correct Classification Rate
ACDM	Additive-Cognitive Diagnosis Model
BUG-DINO	“BUG”- Deterministic Input Noisy “or” Gate
BUG-DINO-CR	“BUG”- Deterministic Input Noisy “or” Gate- Constructed-Response
BUG-DINO-Step	“BUG”- Deterministic Input Noisy “or” Gate- Step
CR	Constructed-Response
C-RUM	Compensatory-Reparameterized Unified Model
DINA	Deterministic Input Noisy “and” Gate
DINA-CR	Deterministic Input Noisy “and” Gate- Constructed-Response
DINA-Step	Deterministic Input Noisy “and” Gate-Step
DINO	Deterministic Input Noisy “or” Gate
GDCM-MC	Generalized Diagnostic Classification Models for Multiple Choise
G-DINA	Generalized DINA
GDM	General Diagnostic Model
HDCM	Hierarchical Diagnostic Classification Model

HO-DINA	Hierarchical DINA Model
MAD	Mean Absolute Deviation
LCDM	Log-linear CDM
LLM	Linear Logistic Model
MC	Multiple Choice
MC-DINA	Multiple Choice-Deterministic Input Noisy “and” Gate
MC-DINA-H	Multiple Choice-Deterministic Input Noisy “and” Gate-Hierarchical
MC-DINA-S	Multiple Choice-Deterministic Input Noisy “and” Gate Measuring Subskills
MC-DINA-M	Multiple Choice-Deterministic Input Noisy “and” Gate Measuring Misconceptions
MC-M-DINO1	Multiple Choice-Misconception-Deterministic Input Noisy “or” Gate1
MC-M-DINO2	Multiple Choice-Misconception-Deterministic Input Noisy “or” Gate2
NC-RUM	Noncompensatory Reparameterized Unified Model
NIDA	Noisy Inputs, Deterministic “and” Gate
NIDO	Noisy Input, Deterministic “or” Gate Model
NR IRT	Nominal Response Item Response Theory
PCCR	Pattern Correct Classification Rate
RMSE	Root Mean Square Error
rSISM	Reduced Simultaneously Identifying Skills “and” Misconceptions Model
SCIM	Scaling Individuals “and” Classifying Misconceptions

SISM Simultaneously Identifying Skills “and”
Misconceptions Model

TIMSS Trends in Mathematics and Science Study



BÖLÜM I

GİRİŞ

Bu bölümde çalışmaya ait problem durumu, amaç, önem, problem cümlesi, sınırlılık ve sayılılar olmak üzere altı başlık yer almaktadır.

1.1. Problem Durumu

Geleneksel eğitim sistemlerinde yer alan ezberci ve kalıplaşmış fikirlere sahip bireylerden uzaklaşan nitelikli insan gücü gelişimi, eğitimin temelinde yer alan hedeflerden biridir. Başka bir deyişle birçoğu okullarda planlanan eğitim sistemleri, süreci verimli bir şekilde inşa ederek nitelikli öğrenmelere sahip bireyler yetiştirmeyi amaçlamaktadır. Bu sistem içerisinde bireyin istendik nitelikleri yeterli düzeyde kazanabilmesi için öğretim sürecinin ayrılmaz bir parçası olan ölçme - değerlendirme uygulamalarının güvenilir ve geçerli koşullar altında yürütülmesi sağlanmalıdır. Bu nedenle meydana gelen eksikliklerin, programda yer alan güçlü ve zayıf yönlerin belirlenebilmesi adına bireylere ait niteliklere ilişkin geri bildirimlere ihtiyaç vardır (Ayan, 2018). Bu ihtiyacı karşılamak için Ackerman (1996) değişkenlerin tek bir boyutla ifade edilemeyeceği başarı testlerinin tasarlandığına dikkat çekmiştir.

Eğitim ve psikolojide yer alan bu testler her ne kadar geniş bir alanı kapsamış olsa da testler değerlendirmelerde içeriği tam anlamıyla yansıtamadığı için bireylerin başarı düzeylerini sınırlı bir şekilde özetleyen toplam puanlar rapor edilmektedir (Ardıç, 2020). Klasik test kuramı veya madde tepki kuramında bir maddeye yanıt verirken bireyin hangi bilişsel süreçleri kullandığı bilinmeden puanlar raporlanmaktadır. Psikometri literatüründe toplam puan raporlamayla bireyleri tek bir yeterlik sürekliliği içerisinde yerleştirmeyi amaçlayan

değerlendirmelerin yeterli tanı bilgisi sağlamadığı vurgulanmaktadır (de la Torre, 2009b). Bu kuramları temele alan geleneksel değerlendirmeler matematiksel alt yapıları sayesinde bireyleri karşılaştırma ve bireylerin gelecekteki başarılarını tahmin etmede avantajlı gözükse de uzmanlara tanımlanmış bir alanda bireylerin güçlü ve zayıf yönlerine ilişkin bilgi verememektedir (de la Torre & Karelitz, 2009). Geçtiğimiz on yıl içerisinde bireylerin bilişsel süreçlerinin anlaşılması için alternatif olarak geliştirilen tanı sınıflama modelleri olarak da isimlendirilen bilişsel tanı modellerinin uygulaması hızlı bir şekilde artmış ve eğitim uzmanları, sınıf içi eğitim ve öğretimi geliştirmek için bireylerin yeterlikleri hakkında tanı bilgileri sağlayan biçimlendirici değerlendirmelere yönelmişlerdir. Bu tür biçimlendirici değerlendirmeler bilişsel tanı değerlendirmeleri olarak adlandırılır (de la Torre & Minchen, 2014). Bilişsel tanı değerlendirmelerinin en ayırt edici özelliği, bireylerin çoklu niteliklerindeki yeterlikler üzerine daha detaylı geri bildirim sağlayan profiller sunmasıdır. Bu çerçevede bireylere başarılı oldukları niteliklerin yanı sıra henüz yeterince başarılı olmadıkları niteliklerle ilgili geri bildirim olanağı sağlanmaktadır. Geri bildirimlerle öğrenilen ya da öğrenilemeyen durumların tespiti yapıldığı sürece bireyler kendi gelişim süreçlerini takip edebilecekler ve bu sayede eksikliklerini tamamlayarak ilgili üniteye istenilen öğrenme düzeyine ulaşacaklardır. Ayrıca dersin belirli bir ünitesinde öğretilmesi planlanan birtakım semboller ve kavramlar bu dersin devamı niteliğindeki sonraki ünitelerin işleniş sırasında birer köprü işlevi kuracaktır. Kavram ve sembollerde meydana gelebilecek eksiklikler öğretim süreci devam ederken giderilmezse, bireyler ileriki öğrenmelerinde birtakım engellerle karşılaşacaktır.

Öğretimin amacı tüm bireylerin en iyi öğrenme süreçlerini deneyimlemelerini sağlamaktır. Ancak bireylerin çoğunun öğretim süreci devam ederken almış oldukları derslerdeki çeşitli konu alanlarına ilişkin öğrenmelerde güçlükler veya başarısızlıklar yaşadığı kabul edilmektedir (Tall & Razali, 1993). Öğrenme zorluklarının nedenlerinden biri olarak kavramlar arası ilişkilerin çok güçlü olması sunulabilir (Niss, 1999). Bu süreçte sıklıkla karşılaşılan öğrenme engelleri kavram yanılgılarıdır. Kavram yanılgıları bireylerin çeşitli fenomen kavramları veya bu alanlara ilişkin uzman bilgisinden farklı ön bilgileri olarak görülmektedir. Kavram yanılgıları, bireylerin dünyayla veya sınıftaki etkileşimleriyle önceki öğrenmelerinden ve deneyimlerinden ya da kavramların benzer kullanımlarından kaynaklanabilir. Öğretim sürecinin doğru ve eksiksiz bir şekilde devam edebilmesi için bireyler tarafından kavramlara doğru anlamların yüklenmesi ve kavram yanılgılarının ortadan kaldırılması gerekmektedir (Sever, Budak & Yalçınkaya, 2009).

Bahsedilen bu kavram yanılgıları kalıcı ve geleneksel öğretim yoluyla düzeltilmeye karşı dirençli olabilmektedir. Bu direnç nedeniyle sistematik hatalar üretebileceğinden ve öğrenmeyi engelleyebileceğinden, öğretim öncesi ve öğretim süreci boyunca bu yanılgılar üzerinde durulmalıdır. Bilişsel tanı modelleri, öğrenme ve öğretme süreciyle ilgili daha ayrıntılı bilgiler sağladığından eğitimcilerin bu sorunların üstesinden gelmelerine yardımcı olmak ve bireylerin öğrenmelerini geliştirmek adına geleneksel yaklaşımlara ek olarak kavram yanılgılarının belirlenmesinde sunulabilecek çözüm yollarından biridir. Ancak uzun yıllar literatürde yer alan mevcut bilişsel tanı modelleri nitelikleri oluşturan becerilerin ve kavram yanılgılarının varlığını ya da yokluğunu eş zamanlı olarak belirleyebilecek şekilde kullanılamamıştır. Yani bilişsel tanı modellerinin çoğu becerileri veya kavram yanılgılarını ayrı ayrı tanımlamak için tasarlanırken, iki durumu aynı anda içerecek şekilde tasarlanamamıştır. Son yıllarda alandaki bu boşluğu doldurmak için yeni bir bilişsel tanı modeli (SISM- Simultaneously Identifying Skills “and” Misconceptions Model) geliştirilmiştir (Kuo, Chen & de la Torre, 2018). Model, bireylerde bir arada bulunabilecek beceri ve kavram yanılgılarını eş zamanlı olarak belirleyebilen bir modeldir. Bilişsel tanı değerlendirmeleri kapsamında eğitimciler ve bireylere verilecek olan geri bildirimlerin eksiksiz ve pratik olarak faydalı olabilmesi adına kullanılan modellerin sadece bireylerin sahip olduğu beceriler hakkında değil, aynı zamanda bireylerde var olan kavram yanılgıları hakkında da bilgi vermesi gerekmektedir. Çünkü bir maddeye verilen yanlış bir yanıt bireyin gerekli becerilere sahip olmaması, bireyde bazı kavram yanılgılarının var olması veya her ikisini içeren durumların varlığına işaret edebilir.

Literatür incelendiğinde SISM’in performansına ilişkin sınırlı sayıda çalışmaya rastlanmıştır (Kuo vd., 2018). Bu nedenle modelin geçerliğine ilişkin kanıtların sağlanmasına yönelik daha çok çalışmaya ihtiyaç olduğu kanısı dikkate alındığında test tasarımı (Q-matris yapısı) ve veri gereksinimleri (örneklem büyüklüğü, test uzunluğu ve madde ayırt edicilik düzeyi gibi) açısından modelin daha iyi anlaşılabilmesi için farklı simülasyon koşulları altında performansının değerlendirilmesinin gerekli olduğu düşünülmüştür. Çünkü model performansı, becerileri ve kavram yanılgılarını ölçmeyi amaçlayan test geliştiriciler için bir ön koşuldur. Bu amaçla bilişsel tanı değerlendirmeleri kapsamında ya da klasik test kuramı gibi farklı kuramlar kapsamında geliştirilen ancak bilişsel tanı değerlendirmelerine uyarlanan bir testte, kullanılacak tanısalsınıflama modelinin becerileri ve kavram yanılgılarını güvenilir bir şekilde belirleyebilmesi beklenmektedir (Schellman, 2021). Yukarıda sıralanan durumlar bu çalışmanın yapılmasının gerekçesini oluşturmaktadır.

1.2. Araştırmanın Amacı

Çalışmanın amacı SISM'in performansını örneklem büyüklüğü, test uzunluğu, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısı olmak üzere farklı koşullar altında karşılaştırarak modelin hangi koşullar altında benzer ya da farklı sonuçlar verdiğini göstermek ve SISM'i, becerilerle kavram yanlışlarını ayrı ayrı tanımlayan (sırasıyla DINA ve BUG-DINO) modellerin performanslarıyla karşılaştırmaktır.

1.3. Araştırmanın Önemi

Ölçme ve değerlendirme uygulamaları eğitim sistemi içerisinde uzun yıllardır sonuç odaklı bir şekilde işlevini sürdürmektedir. Bu durum sınıf ortamının çoğu zaman baskılayıcı bir yarış ortamına dönüşmesine sebep olmaktadır (Ayan, 2018). Amerika'da "Hiçbir Çocuk Geride Kalmayın" projesi (2002) ile birlikte yarış ortamının beraberinde getirdiği başarısızlık duygusu nedeniyle mutsuz olan bireylerin ihtiyaçlarına dikkat çekilmiş ve eğitimin niteliğini arttırmak amacıyla girişimlerde bulunulmuştur. Böylelikle bireylerin güçlü ve zayıf yönlerine daha çok odaklanan eğitim faaliyetleriyle birlikte biçimlendirici değerlendirmeye ilişkin yeni arayışlar ortaya çıkmıştır (Cheng, 2010; de la Torre, 2009b). Bu ihtiyacı karşılayan bilişsel tanı modellerinin gelişimiyle birlikte bireylerin güçlü ve zayıf yönlerinin daha detaylı bir şekilde tanımlanmasıyla değerlendirme sürecinin bireyselliği artırılmıştır (Başokçu, 2011).

Dünyada meydana gelen bu değişimleri eğitim sistemlerinin içerisine entegre edebilmek için öncelikle bilişsel tanı değerlendirmeleri kapsamında kullanılan modellerin özelliklerinin farklı simülasyon koşulları altında incelenerek, hangi koşullarda daha yüksek performans gösterdiğinin belirlenmesi gerekmektedir (Cui, Gierl & Chang, 2012). Literatürde bilişsel tanı değerlendirmeleri kapsamında elde edilen bulguların geçerli ve güvenilir olmasıyla doğru model seçiminin özellikle sınıflama doğruluğu üzerinde önemli bir etkisinin olduğu araştırmacılar tarafından ifade edilmiştir (de la Torre & Douglas, 2004; Lee & Sawaki, 2009). de la Torre, Hong ve Deng (2010) çalışmalarında madde sayısı, Q-matris, nitelik sayısı vb. arasındaki ilişkilerin incelenmesinin modelin sınıflama doğruluğu performansında etkili olduğunu ve bu ilişkilerin araştırılması gerektiğini vurgulamışlardır. Bu amaçla çalışma, bilişsel tanı değerlendirmelerinde kullanılan modellerden biri olan SISM'in faktörlerde değişimlenen farklı koşullar altındaki (örneklem büyüklüğü, test uzunluğu, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısı) performansını değerlendirmekle birlikte

beceri (DINA) ve kavram yanılgısı (BUG-DINO) modelleriyle karşılaştırarak modelin sınıflama doğruluğu performansını incelemiştir.

Model karşılaştırmalarının sunulmasının amacı, beceri ve kavram yanılgılarının farklı modellerle ayrı ayrı tanımlanmalarından kaynaklı olarak ortaya çıkabilecek bilgi kayıplarıdır. Örneğin bireyler tarafından doğru ve yanlış seçeneği seçme olasılıkları, bireylerin beceriye sahip oldukları ve kavram yanılgılarının da var olduğu durumlar değerlendirildiğinde beceri ya da kavram yanılgılarını ayrı ayrı ölçen modellerin kullanımıyla ayırt edilemez. Bununla birlikte uygulamada, bazen doğru çözümler üretebilen bazı bireyler için kavram yanılgıları varlığını sürdürebilir. Buna ek olarak son çalışmalar, bilimsel olmayan kavramlar gibi bireylerin kavram yanılgılarının kavramsal bir değişim meydana geldikten ve doğru yanıtlar üretildikten sonra bile bilimsel kavramlarla bir arada bulunabileceğini göstermiştir (Kuo vd., 2018).

Ortaya çıkabilecek olası problem durumlarının çözümüne yönelik olarak geçerli ve güvenilir çıkarımlar için beceri ve kavram yanılgılarının belirlenmesi noktasında bu çalışmada performansı değerlendirilen SISM, uzmanların karşılaşılabileceği bilgi kayıplarının giderilmesi noktasında katkı sağlayacaktır. Ayrıca bu çalışma model karşılaştırmalarıyla çalışma kapsamında yer verilen çok sayıda değişimlenen koşul için ortaya çıkan sorunlara katkılar sunabilir. Değişimlenen koşullarla model tarafından üretilen madde parametrelerinin hangi durumlarda daha tutarlı sonuçlar verdiği belirlenerek bilişsel tanı değerlendirmeleri kapsamında model ya da modellere ilişkin özelliklerin daha kolay anlaşılması ve bunlara ilişkin sınırlılıkların belirlenmesi noktasında araştırmacıların ya da eğitim paydaşlarının zihinlerindeki soru işaretlerinin giderileceği düşünülmektedir.

1.4. Problem Cümlesi

Bilişsel tanı değerlendirmesi kapsamında kullanılan modellerden SISM'in örneklem büyüklüğü, test uzunluğu, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısı olmak üzere dört faktöre ait değişimlenen koşullarda madde parametrelerine ilişkin elde edilen ortalama mutlak sapma ve hata değerleriyle birlikte bu modelin DINA ve BUG-DINO modellerle karşılaştırıldığı durumlardan elde edilen sınıflama doğruluğu değerleri nasıl değişmektedir?

1.4.1. Alt problemler

1. SISM'e göre madde parametrelerine ait "ortalama mutlak sapma" değerleri,
 - 1.1. Örneklem büyüklüğü (250, 500, 1500 ve 3000),
 - 1.2. Test uzunluğu (20, 40 ve 60 madde),
 - 1.3. Madde ayırt edicilik düzeyi (yüksek ayırt edicilik ve düşük ayırt edicilik) ve
 - 1.4. Ölçülen nitelik sayısına (6, 8 ve 10 nitelik) göre nasıl değişmektedir?
- 2.SISM'e göre madde parametrelerine ait "hataların ortalama karekökü" değerleri,
 - 2.1. Örneklem büyüklüğü (250, 500, 1500 ve 3000),
 - 2.2. Test uzunluğu (20, 40 ve 60 madde),
 - 2.3. Madde ayırt edicilik düzeyi (yüksek ayırt edicilik ve düşük ayırt edicilik) ve
 - 2.4. Ölçülen nitelik sayısına (6, 8 ve 10 nitelik) göre nasıl değişmektedir?
- 3.Becerileri ölçen DINA, kavram yanılgılarını ölçen BUG-DINO ve hem becerileri hem de kavram yanılgılarını ölçen SISM'in sınıflama doğruluğu değerleri,
 - 3.1. Örneklem büyüklüğü (250, 500, 1500 ve 3000),
 - 3.2. Test uzunluğu (20, 40 ve 60 madde),
 - 3.3. Madde ayırt edicilik düzeyi (yüksek ayırt edicilik ve düşük ayırt edicilik) ve
 - 3.4. Ölçülen nitelik sayısına (6, 8 ve 10 nitelik) göre nasıl değişmektedir?

1.5. Sınırlılıklar

1. Çalışma örneklem büyüklüğü (250, 500, 2500 ve 3000), test uzunluğu (20, 40 ve 60 madde), madde ayırt edicilik düzeyi (yüksek ayırt edicilik ve düşük ayırt edicilik) ve ölçülen nitelik sayısı (6, 8 ve 10 nitelik) olarak tanımlanan dört faktöre ait değişimlenen koşullarla sınırlıdır.

2. Çalışma SISM ile ilişkili Q-matrislerinin oluşturulması sürecinde becerilerin ve kavram yanılgılarının kendi içinde (küme-içi) bağımlılıklarının olmadığı yani beceriler ve kavram yanılgıları arasında dış bağımlılık ilişkisinin kurulduğu durumlarla sınırlıdır.

3. Çalışma SISM'in daha az karmaşık olan modellerle karşılaştırıldığı durumlarda beceri modeli için DINA, kavram yanılgısı modeli için BUG-DINO ile sınırlıdır. Bu üç model iç içe geçmiş yani yuvalanmış modeller oldukları için becerileri içeren DINA ve kavram yanılgılarını içeren BUG-DINO modellerinde doğaları gereği SISM tarafından hesaplanan dört madde parametresinin tümü hesaplanamamıştır. Bu nedenle üçüncü araştırma problemi kapsamında model karşılaştırmaları sınıflama doğruluğu performansı üzerinden değerlendirilmiştir. Ayrıca çalışmanın amacı kapsamında SISM'in sınıflama doğruluğu performansının DINA ve BUG-DINO modellerinin sınıflama doğruluğu performanslarıyla karşılaştırıldığı durumlar dikkate alındığında çalışmada yer alan bulgular DINA model için sadece beceriler, BUG-DINO model için sadece kavram yanılgıları ve SISM için hem beceriler hem de kavram yanılgılarını içeren sınıflama doğruluğu performanslarıyla sınırlıdır.

1.6. Sayıtlar

1. Bilişsel tanı değerlendirmeleri kapsamında beceri ve kavram yanılgılarının eş zamanlı olarak ölçüldüğü modellerden SISM dikkate alındığında, model kapsamında Q-matris geçerliğinin istatistiksel olarak analizi mümkün olmadığından Q-matris geçerliğinin testteki nitelikleri yansıttığı ve simülasyon programının gerçek durumlara uygun veriler ürettiği varsayılmıştır.

BÖLÜM II

KURAMSAL ÇERÇEVE

Bu bölümde çalışmanın kuramsal çerçevesi ve ilgili literatür hakkında detaylı bilgilere yer verilmiştir. Kuramsal çerçevede “Bilişsel Tanı Değerlendirmeleri”, “Nitelik ve Hiyerarşiler”, “Q-matris”, “Bilişsel Tanı Modelleri” ve “Kavram Yanılgılarına İlişkin Bilgiler” başlıkları yer almaktadır. Ayrıca bu çalışmada yer alan “Bilişsel Tanı Modelleri” genel başlığı altında, çalışmada kullanılan modellere (*Deterministic-Input, Noisy “and” Gate (DINA) Model*, “BUG”- *Deterministic Input Noisy “or” Gate Model (BUG-DINO)*, *Simultaneously Identifying Skills “and” Misconceptions Model (SISM)*) ilişkin bilgiler sunulmuştur.

İlgili araştırmalar bölümü beceri ve kavram yanılgılarının belirlenmesine ilişkin bilişsel tanı modellerinin kullanıldığı araştırmalara yer vermiştir.

2. 1. Araştırmanın Kuramsal Temelleri

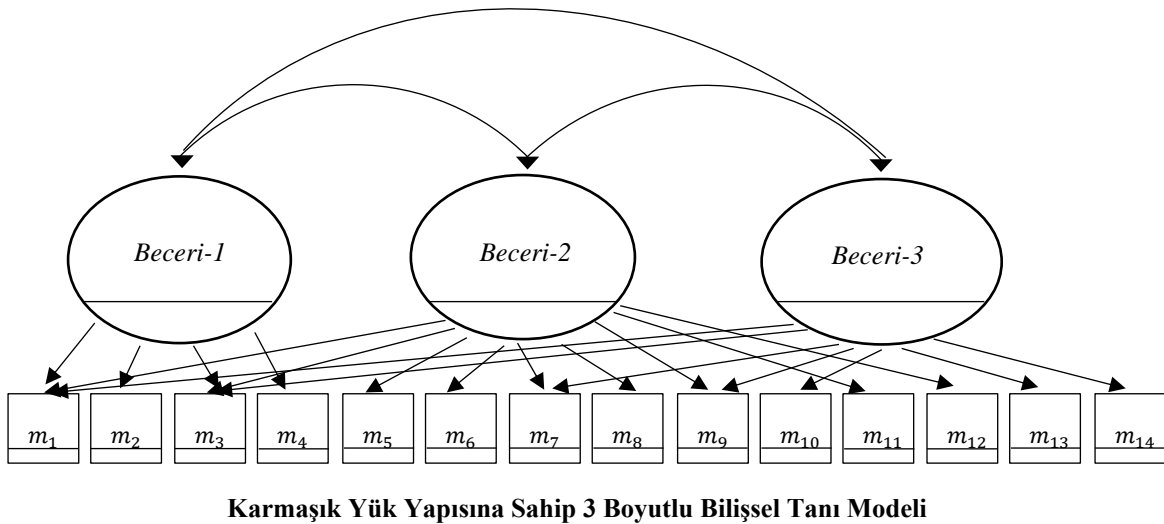
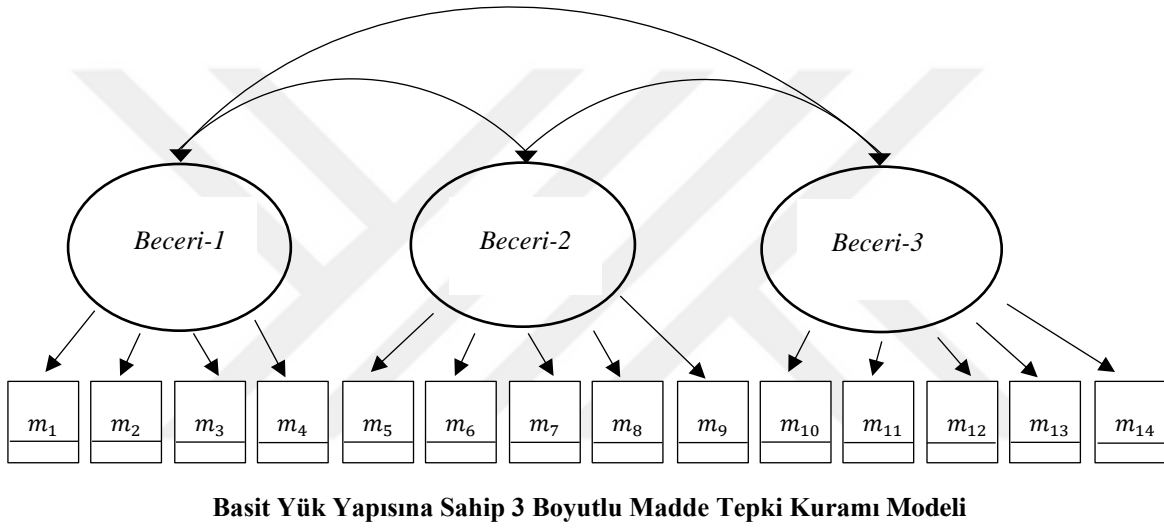
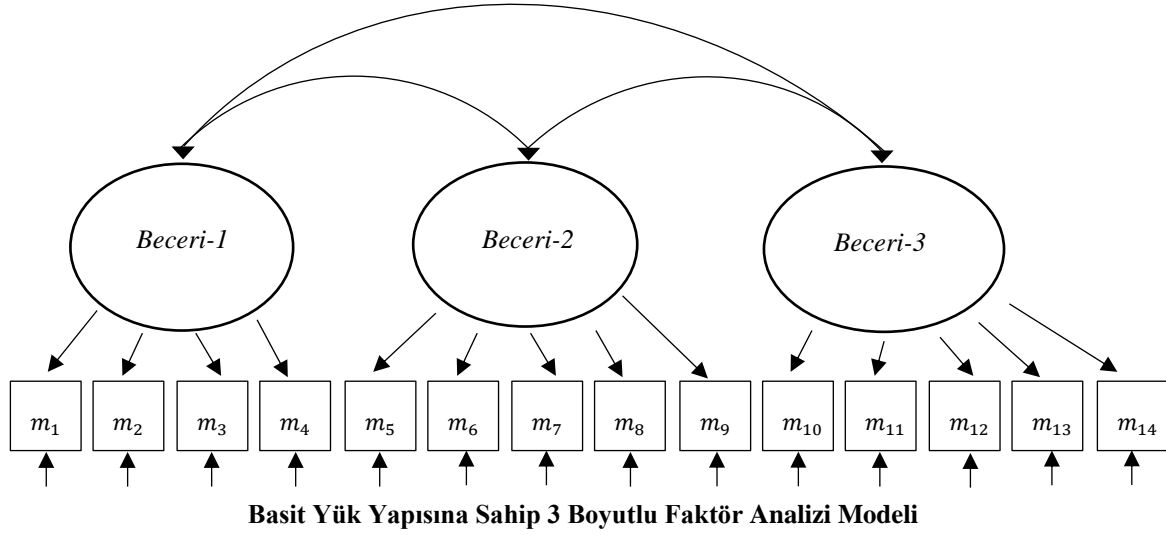
2. 1. 1. Bilişsel Tanı Değerlendirmeleri

Literatürde mevcut testlerin çoğu klasik test kuramı ya da madde tepki kuramı temelinde geliştirilmiştir. Bu testler genellikle bireylerin yeteneğini veya başarısını temsil eden toplam bir test puanı sağlamasına ek olarak bireyi diğer bireylere veya belirli standartlara göre sıralamak için kullanılır. Bu yönüyle bu değerlendirmeler sınıf içi eğitimin gelişimine yeterli bir katkı sağlayamamaktadır (de la Torre, 2008). Aynı test puanlarına sahip bireyler farklı bilgi yapılarına sahip olabilir. Bu nedenle toplam test puanlarının eğitimde temel olarak kullanılması yararlı olmaz. Çünkü bireylerin iç bilgi yapılarındaki farklılıklar eğitim ve öğretimin güvenilir temelini oluşturmalıdır (Li vd., 2020).

Bu doğrultuda geleneksel tek boyutlu klasik test kuramı ve madde tepki kuramının aksine son yıllarda özellikle sınıf içi değerlendirmelerde tanı değerlendirme raporları ya da tanısız geri bildirim için “Tanı Sınıflama Modelleri” olarak bilinen bir dizi bilişsel tanı modeli geliştirilmiştir (Rupp, Templin & Henson, 2010). Leighton, Gierl ve Hunka (2004) bu modellerle beceriler olarak adlandırılan bilişsel gereksinimlerin, testte yer alan maddeler yardımıyla birleştirilerek sınıflamaya dayalı karar verme amacıyla kullanılabileceğini iddia etmişlerdir. Bahsedilen sınıflamalara ilişkin kararlar verilirken bilişsel tanı modelleri, gizil sınıf sayısının kaç olduğunun belirlenmesinin yanında bireylerin hangi becerilere sahip olduğunun belirlenmesi ve maddeler tarafından ölçülen beceri kombinasyonlarını bireylerin maddelere vermiş oldukları yanıtlar üzerinden tahmin etmesi bakımından doğrulayıcıdır (Templin & Hoffman, 2013).

Doğrulayıcı doğasıyla bir ya da daha fazla gizil değişkeni ölçen bu modeller, faktör analizi ve madde tepki kuramı modellerine benzemektedir. Bilişsel tanı değerlendirmeleri kapsamında yer alan bu modellerde gizil değişkenlerin kategorik olduğu varsayılmaktadır. Madde tepki kuramı ve faktör analizi, doğrulayıcı olmaları yönünden bilişsel tanı modellerine benzese de gizil değişkenlerin sürekli olmaları yönünden bilişsel tanı modellerinden ayrılmaktadır (Sen & Bradshaw, 2017). Ayrıca bilişsel tanı modellerinde madde-nitelik ilişkisi önceden belirlenmektedir. Bu yönüyle bilişsel tanı modelleri madde tepki kuramı modellerinden farklılık göstermektedir (Liu, Quian, Luo & Woo, 2018).

Bilişsel tanı modelleri yaptıkları sınıflamalar sayesinde bireyler için tanı bilgisine dayalı pedagojik olarak daha ayrıntılı sonuçlar sağlayabilir (de la Torre, 2009b). Çok boyutlu madde tepki kuramı modelleri kullanılarak da bireylere ilişkin tanı bilgisi sağlamak mümkündür ancak bu modellerin temelinde yatan değerlendirme tasarımları tanılayıcı bilgiler sağlama noktasında yetersiz kalmaktadır (Bradshaw & Templin, 2014). Bahsedilen yetersizlikler nedeniyle araştırmacılar testlerden elde edilen bilgileri kullanarak bireylere çok boyutlu geri bildirimlerin nasıl sağlanacağına ilişkin bir süredir çalışmalarını sürdürmektedir. Ancak her bir boyutu ölçmede çok sayıda maddeye ihtiyaç duyulduğu için pratik test koşulları altında birden fazla sürekli becerinin ölçülmesi zordur. Verilen bilgiler faktör analizi, madde tepki kuramı ve bilişsel tanı modellerinin belirli yönler bakımından birbirinden farklılaştığını göstermektedir. Rupp ve Templin (2008) bu farklılaşmayı aşağıda sunulan Şekil 1 üzerinden açıklamışlardır.



Şekil 1. 3 boyutlu faktör analizi ve madde tepki kuramı modellerinin, 3 boyutlu bilişsel tanı modeli ile karşılaştırılması. Rupp, A., & Templin, J. (2008). Unique characteristics of diagnostic models: A review of the current state-of-the-art. *Measurement*, 6(4), 219–262. <https://doi.org/10.1080/15366360802490866>

Şekil 1, bilişsel tanı modelinin çok boyutlu faktör analizi ve çok boyutlu madde tepki kuramı modellerinin sahip oldukları özelliklere göre temel özelliklerini özetlemektedir. Burada kullanılan bilişsel tanı modeli, karmaşık yük yapısına sahip üç boyutlu telafi edici bir modeldir. Bu model daha basit yapılara sahip olan üç boyutlu faktör analizi ve madde tepki kuramı modelleriyle karşılaştırılmaktadır.

Şekil 1 incelendiğinde gizil beceri değişkenleri arasındaki iki yönlü oklar, ikili korelasyonları; gizil beceri değişkenlerinden madde değişkenlerine giden tek yönlü oklar, 1'in Q-matrisi girişlerine karşılık gelen yükleri; 3 boyutlu faktör analizi modelinde dikdörtgenleri gösteren en alttaki oklar ölçme hatasını; 3 boyutlu madde tepki kuramı ve bilişsel tanı modellerinde yer alan dikdörtgenlerdeki yatay çizgiler gizil eşikleri (threshold-yani ikili gözlenen değişkenler) ve bilişsel tanı modelinde dairelerdeki yatay çizgiler de gizil eşikleri gösterir (yani ikili gizil beceri değişkenleri). Yatay çizgilerin konumu basitlik amacıyla aynıdır ancak farklı marjinal yüzdelerin madde değişkenleri için doğru olduğunu ve gizil beceri değişkenleri için farklı marjinal yeterlik oranlarını gösterebilmektedir. Bu durum faktör analizi ve madde tepki kuramı modelleriyle bilişsel tanı modelleri arasındaki temel farkın bilişsel tanı modellerinin kategorik gizil değişkenler içerirken, faktör analizi ve madde tepki kuramı modellerinin sürekli gizil değişkenler içerdiğinin bir göstergesidir.

Madde tepki kuramı modellerinde olduğu gibi her bireyi çok boyutlu sürekli bir alana yerleştirmek yerine, bilişsel tanı modelleri kullanılarak bireyler kabaca (beceriye sahip ya da değil) her bir beceriye göre sınıflandırılabilir. Bu yönelim eğitim alanında gerçekleştirilen ölçmelerdeki mevcut manzaranın büyük bir içerik alanı için tek bir puanla sonuçlanan değerlendirmeler oluşturma ve yönetme sürecinden ziyade, birden çok alt alan için yeterlik kararları verebilecek değerlendirmelere doğru kaydığının bir göstergesidir (Liu, Huggins-Manley & Bulut, 2018). Bu taleplere yanıt olarak psikometri alanındaki uzmanlar, eğitim değerlendirme rutinlerinin bireylerin belirli beceriler üzerindeki performanslarını gösteren daha ayrıntılı bir değerlendirmeye dönüşmesini sağlayan araçlar geliştirmişlerdir (Sen & Bradshaw, 2017). Tanısal çerçeve içinde bir ölçme aracı geliştirmek şüphesiz bu bilgileri elde etmek için örnek bir yaklaşımdır. Bu çerçeve bireyler hakkında çok boyutlu bir yaklaşım sunduğundan sınıflama tabanlı geri bildirim sağlayan etkili bir yoldur (Rupp vd., 2010). Böylece bireye ait performans, bireyin sahip olduğu ya da olmadığı becerilerle birlikte sunulabilir. Bu durum bireye ilişkin tanılayıcı bir profil sunmaktadır. Tanılayıcı profil yardımıyla eğitim süreci bireyselleştirilmiş şekilde düzenlenebilir.

Genel anlamda uzmanlar bahsedilen modellerle ilişkilendirilen bir maddenin doğru yanıtlanabilmesi için hangi becerilerle ilişki kurulacağı belirlendiği hipotezlere ihtiyaç duyarlar. Bu hipotezlerin arka planında nitelikler ve maddeler arasındaki ilişkiyi açıklayan Q-matrisler yer almaktadır. Çalışmanın bir sonraki bölümünde bilişsel tanı modelleriyle ilişkili olan nitelikler, nitelikler arası ilişkilerin sunulduğu hiyerarşi türleri ve Q-matrise ilişkin genel bilgilere yer verilmiştir.

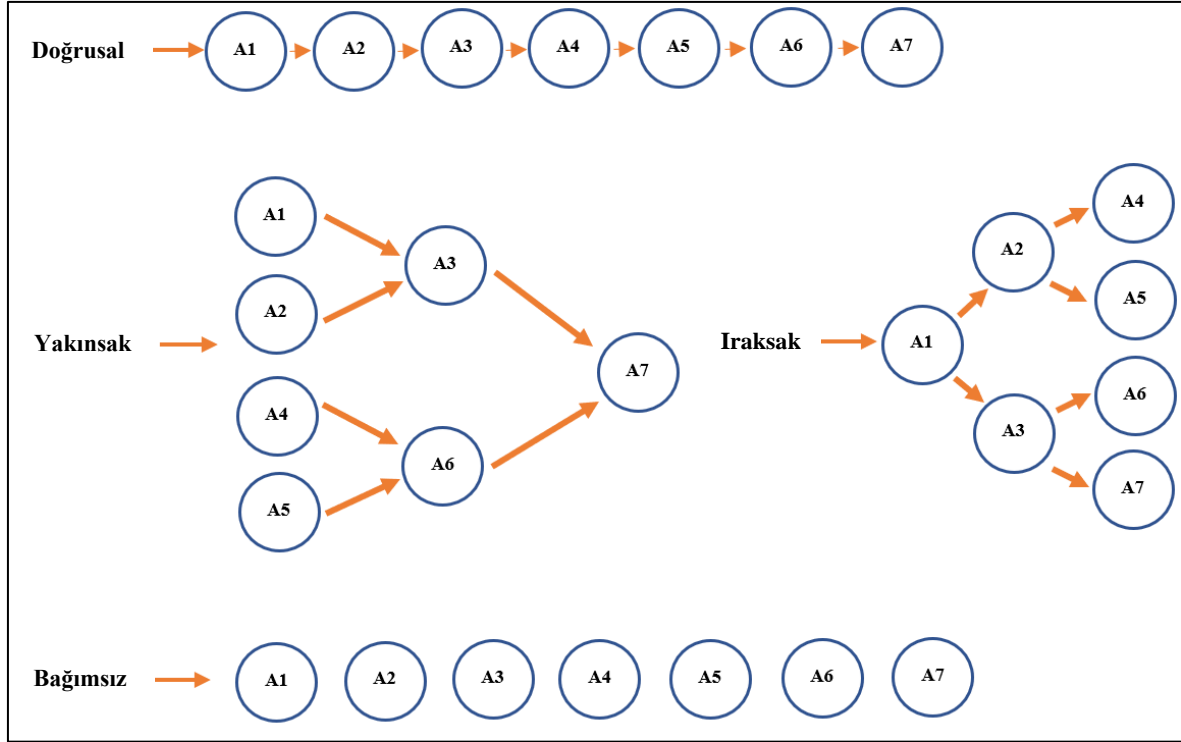
2.1.2. Nitelikler ve Hiyerarşi

Bilişsel tanı değerlendirmelerinde özellik ya da nitelikler olarak adlandırılan bir dizi ayrık gizil değişken bireylerin sahip olduğu/oldukları beceriler, süreçler, kavram yanılgıları ya da hatalar gibi doğrudan gözlemlenemeyen karakteristiklerdir (Elbulok, 2021; Ravand & Baghaei, 2020). Literatür incelendiğinde en sık kullanılan terimlerin beceri ve nitelik olduğu görülmektedir (Rupp vd., 2010). Kavramsal olarak nitelik, bireyin bir maddeyle ilgili problemi çözmesinde ihtiyaç duyduğu karakter, beceri ya da bir yapıdır. Psikometrik açıdan ise istatistiksel modeldeki gizil bir değişken olarak tanımlanır (Ma & de la Torre, 2019a).

Araştırmacılar niteliklerin üretken kurallar, proje türleri, genel bilişsel görevler veya öne sürülen bilgi ve düşünme becerileri (Tatsuoka, Corter & Tatsuoka, 2004); bir bireyin bir test maddesini çözmek için sahip olması gereken prosedürler, beceriler, süreçler, stratejiler ve bilginin tanımı (Dogan & Tatsuoka, 2008) veya belirli bir görevi tamamlamak için gerekli olan işlem becerileri ve bilgi yapısı (Leighton & Gierl, 2007) olduğunu ifade etmişlerdir. Literatür incelendiğinde ölçme alanındaki bilişsel niteliklere ilişkin tek tip bir tanımın yapılamadığı görülmekle birlikte günümüz literatüründe en güncel tanımlardan biri olarak Wu, Wu, Chang, Kang ve Zhang (2020) tarafından yapılmış olan *-bilişsel nitelik, belirli bir standart temelinde bireylerin bilgi durumlarını daha kesin olarak anlamak için bir sınıflama yolu olarak ele alınabilir-* tanımı sunulabilir.

Literatürde sunulan tanımlar dikkate alındığında nitelik teriminin yaygın olarak kullanılan beceri terimini de içine alan daha genel bir kavram olarak tanımlandığı görülmektedir. Çalışmada bireylerin bilgi durumlarını daha iyi anlamak amacıyla hedeflenen beceri, görev, özellik ve kavram yanılgısı gibi tüm durumları kapsayan daha genel bir sınıflama yolu olarak performansı değerlendirilen SISM için “nitelik” teriminin kullanılması uygun görülürken; DINA model için “beceri” terimi, BUG-DINO model için “kavram yanılgıları” terimi kullanılmıştır.

Bilişsel tanı modellerinin temelinde yer alan Q-matris oluşturma işlemlerine geçmeden önce gerekli olan adımlardan biri, ölçülmesi hedeflenen niteliklerin belirlenmesinin yanı sıra - *eğer varsa*- belirlenen nitelikler arasındaki olası ilişkileri tanımlamaktır (Ayan, 2018; Koyuncu, 2020). Nitelikler arası ilişkilendirme “nitelik hiyerarşisi” olarak adlandırılmaktadır. Nitelik hiyerarşileri, bir niteliğin başka bir niteliğin ön koşulu olduğu durumlara işaret etmektedir. Öğretim süreci genel olarak sırayla ve bir önceki adımın üzerine inşa edildiği için böyle bir varsayım eğitim bağlamında yaygın olarak kullanılmaktadır. Bu işlem dolaylı olarak örnekleme gözlenmesi beklenen nitelik profillerinin (gizil sınıfların) hangisi olduğu yönünde bir hipotez sunmaktadır. Örneğin nitelik 1 nitelik 2’ye sahip olmak için ön koşul ise *-bir diğer ifadeyle nitelik 1 genel olarak nitelik 2 kazanılmadan önce kazanılması gerekli bir nitelik ise-* oluşacak olası tüm gizil sınıflar içerisinde 2. niteliğe sahip olup, 1. niteliğe sahip olmamayı gerektiren profillerde mantıksal olarak evrenden hiç kimse bulunmayacaktır. Böyle bir durum pratikte mümkün değildir. Başka bir deyişle hipotezin önerdiği hiyerarşi eğer doğru ise bu gizil sınıf boş kalacaktır. Bu durum, arasında belirli bir hiyerarşi bulunan niteliklerin kombinasyonlarını ölçen test maddelerinin kullanılmayacağı anlamını taşımamaktadır. Aksine tanılayıcı durum belirlemede bu tür maddeler, bireylerin niteliklere sahip olup olmadıklarına ilişkin fazladan bilgiler elde edilmesine yardımcı olmaktadır ve kapsam geçerliği açısından gereklidir (Rupp vd., 2010). Bilişsel tanı modellerinde ölçme aracıyla ölçülen nitelikler hiyerarşik bir yapıya sahip olduğunda var olan yapı göz ardı edilmemelidir. Bu çalışma kapsamında beceriler ve kavram yanılgıları arasındaki hiyerarşik ilişkiler dikkate alınmıştır. Beceri ve kavram yanılgılarının oluşturduğu hiyerarşinin göz ardı edilmesi bilişsel tanı modelinin uygun ve kullanışlı olmamasına neden olmaktadır (Bradshaw & Templin, 2014). Nitelik hiyerarşilerine ilişkin literatürde yer alan sınıflamalardan biri Şekil 2’de aşağıda sunulmuştur.



Şekil 2. Nitelik hiyerarşileri. Sun, J., Xin, T., Zhang, S., & de la Torre, J. (2013). A polytomous extension of the generalized distance discriminating method. *Applied Psychological Measurement*, 37(7), 503-521. <https://doi.org/10.1177/0146621613487254>

Şekil 2 incelendiğinde örneğin *doğrusal hiyerarşi* türünde, niteliklerin tek bir zincir şeklinde ardışık olarak sıralandığı görülmektedir. Bireyin zincirin sonunda yer alan niteliği kazanabilmesi için şekilde sunulan diğer 6 niteliğe daha önce sahip olması gerekmektedir. Bu durumun bir örneği matematikte yer alan dört işlem becerileridir. Bireylerin sırasıyla toplama, çıkarma, çarpma ve bölme olmak üzere dört aşamalı bir zinciri takip etmeleri gerekmektedir.

Yakınsak hiyerarşi ise zincirde yer alan herhangi bir niteliğin kazanılması için bir ya da birden çok niteliğin kazanılması gerektiğine vurgu yapmaktadır. Zincirin 2. adımında yer alan A6 niteliğine sahip olan bir birey aynı zamanda A4 ve A5 niteliklerine de sahip olmalıdır. Ancak A1, A2, A3 ve A7 niteliğine sahip olması gerekmemektedir. Yani birey bazı durumlarda önceki niteliklerden sadece birine ya da bu niteliklerden hepsine sahip olabilir.

Ölçülen nitelikler yukarıda bahsedilen dört farklı ilişkisel zincire bağlıdır. Bu nedenle kestirimler sonucu bireylerden elde edilen gizil sınıf sayıları da değişmektedir (Ayan, 2018). 1-0 şeklinde kodlanan K tane niteliğin birbirinden bağımsız olduğu durumlarda bireylerin toplam gizil sınıf sayısı 2^K iken madde sayısı $2^K - 1$ 'dir. Bu nitelik profil tahminleri hangi

niteliklerin bireylerde bulunduğunu ya da bulunmadığını özetlemektedir. Şekil 2 dikkate alındığında nitelikler arası ilişkinin bağımsız olduğu hiyerarşi türünde $K = 7$ olduğu için toplam gizil sınıf sayısı $2^7 = 128$ ve madde sayısı ise $2^7 - 1 = 127$ 'dir (Koyuncu, 2020).

Araştırmacıların ilgilendikleri nitelikler konu alanının kapsamı, araştırmacının ilgilendiği yanıtlama süreciyle birlikte bu sürecin nasıl parçalardan oluşacağına ilişkin tanımlamalarla ilişkilidir. Bu nedenle araştırmacılardan niteliklerle ilişkili olan konu alanlarının kapsamını daraltarak nitelikleri özelleştirmeleri beklenmektedir. Ancak bu işlem nitelik sayısının artışına sebep olabilir. Nitelik sayısındaki artışla birlikte madde parametre tahmini güçleşmekte ve bireylere ilişkin sınıflamalarda hata yapma olasılığı artmaktadır (Kunina-Habenicht, Rupp & Wilhelm, 2012). Bir çözüm önerisi olarak araştırmacılara geniş kapsamlı görevlerde nitelikleri geniş bir aralıkta tanımlamaları tavsiye edilirken kapsamı dar görevlerde ise niteliklere ilişkin ayrıntılı tanımlama yapmaları tavsiye edilir (Rupp vd., 2010).

2. 1. 3. Q-matris

Bilişsel tanı değerlendirmeleri, modelin altında yatan çeşitli varsayımlara bakılmaksızın bireyler hakkında tanısal bilgi sağlamak için Q-matris adı verilen ortak bir bileşene sahiptir. Çoğu zaman bilişsel tanı modelleri sürecin yalnızca psikometrik bileşenine atıfta bulunmak için kullanılır. Bununla birlikte daha bütünsel bir bakış açısı, bilişsel tanı modelleme sürecinin bir parçası olarak Q-matrisini içermektedir (de la Torre, 2008). Bu matris bireylerde yer alan nitelikler ve bu niteliklerin ölçülmesine ilişkin tasarlanmış maddeler arasında ilişkiler kurmaktadır. Bu durum faktör yük yapısının önceden belirlendiği doğrulayıcı faktör analizine benzemektedir (Liu & Cheng, 2018).

Teste dahil olan bilişsel süreçlerin tam olarak anlaşılabilmesi nedeniyle gerçek Q-matrisi bilinmemektedir (De Carlo, 2012). Ancak Q-matris kullanılarak bir maddeyi doğru yanıtlamak için hangi nitelik veya niteliklerin gerekli olduğu özetlenebilir. Bu matriste yer alan satırlar maddeleri, sütunlar ise nitelikleri temsil etmektedir (de la Torre, 2009b). Fischer (1973) tarafından sunulan $J \times K$ matrisi 1 ve 0 olarak kodlanan madde ve niteliklerden oluşmaktadır. Matriste bulunan k niteliği, bir i maddesinde bulunduğu anda "1", bulunmadığında ise "0" olarak kodlanmaktadır. Bu kodlamalar maddeler tarafından ölçülen nitelikler kuramsal bir gerekçeyle desteklendikten sonra, analizlere başlamadan önce belirlenmelidir (Rupp, 2007). Aşağıda maddelerin doğru yanıtlanabilmesinde sahip

olunması gereken niteliklere ilişkin bir örnek sunulmuştur. Bu örnekte 3 nitelik 5 çoktan seçmeli test maddesiyle ölçülmüştür.

Tablo 1

Örnek Q-Matrisi

	Nitelik 1 (α_1)	Nitelik 2 (α_2)	Nitelik 3 (α_3)
Madde 1	0	0	1
Madde 2	0	1	1
Madde 3	0	0	1
Madde 4	1	1	1
Madde 5	0	1	1

Tablo 1 incelendiğinde 4. maddenin doğru bir şekilde yanıtlanabilmesi için tüm niteliklere sahip olunması gerekmektedir. Ayrıca madde 3'ü doğru olarak yanıtlayan bir bireyin sadece 3. niteliğe sahip olması gerektiği görülmektedir. Q-matrisleri bir maddeye doğru yanıt vermek için gerekli olan nitelikler ve bu nitelikleri ölçmek için tasarlanan maddeler arasındaki bağlantıları içerdiğinden bilişsel tanı değerlendirmeleri kapsamında önemli bir yere sahiptir. Bu uygulamaların geçerliği K boyutlu gizil nitelik vektörünün, bireylerin atandıkları sınıfları doğru bir şekilde belirleyip belirlemediğine bağlıdır. Bu durumda madde puanlarının koşullu dağılımları, niteliklerin etkisi için ayarlama yapıldıktan sonra birbirinden bağımsız olmaktadır. Bu özellik genellikle yerel bağımsızlık olarak adlandırılmaktadır (Rupp vd., 2010). Yerel bağımsızlık varsayımı, K niteliklerinin $\alpha_1, \alpha_2, \dots, \alpha_K$ 'nın tüm gizil uzayı kapsadığı varsayımına eşdeğerdir. Yani bilişsel tanı değerlendirmesinde hiçbir gizil nitelik gözden kaçırılmamış veya atlanmamıştır. Bu varsayım temelinde modelleme sürecinin ayrılmaz bir parçası olarak Q-matrisinin rolü, çıkarımların geçerliği söz konusu olduğunda daha da kritik hale gelir. Bunun nedeni hem bilişsel tanı modeli hem de Q-matrisinin, tanısal modelleme bağlamı altında potansiyel olarak model-veri uyumsuzluğuna neden olmasıdır. Teste dahil olan bilişsel süreçlerin tam olarak anlaşılabilmesi model-veri uyumsuzluğunun altında yatan sebeplerden biri olarak gösterilebilir (Chen, de la Torre & Zhang, 2013). Bu nedenle uygulamalarda çoğu durumda alan uzmanları tarafından Q-matris hatalı bir şekilde tanımlanmakta ve bilişsel tanı modellerinden yapılan çıkarımların geçerliği araştırmacılar arasında tartışmaları gündeme getirmektedir. Q-matriste yer alan nitelik tanımlamalarının geçerliği sağlanmadan gerçekleştirilen herhangi bir model analizi güvenilir olmaz.

2.1.4. Bilişsel Tanı Modelleri

Literatür incelendiğinde bilişsel tanı modellerine yönelik farklı sınıflamaların olduğu görülmektedir. Ancak modeller arasındaki geleneksel ayrımlar *-bir modeldeki bağlantı fonksiyonununun değiştirilmesi sonucu başka bir modelin parametreleştirilmesine neden olduğundan-* net değildir. Bilişsel tanı değerlendirmeleri literatürüyle referans kolaylığı sağlamak adına Tablo 2’deki sınıflama Ravand ve Baghaei (2020) tarafından önerilmektedir.

Tablo 2

Bilişsel Tanı Modellerinin Sınıflandırılması

Bilişsel Tanı Modeli Türü		Örnekler	Yazar(lar)
Özel	Ayrık	DINO (Deterministic-Input, Noisy “or” Gate Model)	(Templin & Henson, 2006)
		NIDO (Noisy Input, Deterministic “or” Gate Model)	
	Birleşik/Bağlayıcı	DINA (Deterministic-Input, Noisy “and” Gate Model)	(Junker & Sijtsma, 2001)
		NIDA (Noisy Inputs, Deterministic “and” Gate)	(DiBello, Stout & Roussos, 1995; Hartz, 2002)
		A-CDM (Additive-Cognitive Diagnosis Model)	(de la Torre, 2011)
	Ek/İlave	C-RUM (Compensatory Reparameterized Unified Model)	(DiBello vd., 1995; Hartz, 2002)
		NC-RUM (Noncompensatory Reparameterized Unified Model)	
		LLM (Linear Logistic Model)	(Maris, 1999)
		HO-DINA (Hierarchical DINA Model)	(de la Torre, 2008)
		GDM (General Diagnostic Model)	(Henson, Templin & Willse, 2009; Von Davier, 2005)
LCDM (Log-linear CDM)			
Genel	Ayrık, Bağlayıcı ve İlave	G-DINA (Generalized DINA)	(de la Torre, 2011)
		HDCM (Hierarchical Diagnostic Classification Model)	(Templin & Bradshaw, 2013)
	Hiyerarşik		

Ravand, H., & Baghaei, P. (2020) Diagnostic Classification Models: Recent Developments, Practical Issues, and Prospects. *International Journal of Testing*, 20(1), 24-56. <https://doi.org/10.1080/15305058.2019.1588278>

Tablo 2’de yer alan sınıflamada kavram yanlışlarına ait modellere yer verilmediği görülmektedir. Bilişsel tanı modelleri kapsamında kavram yanlışlarının ölçülmesine ilişkin literatürde yer alan modeller incelendiğinde bu tabloya, kavram yanlışlarıyla ilgili modellerin dahil edildiği aşağıdaki gibi bir tablonun eklenmesi önerilebilir.

Tablo 3

Bilişsel Tanı Değerlendirmelerinde Kavram Yanlışlarını İçeren Modellerin Sınıflandırılması

Ölçülen Nitelik Türü	Modeller	Yazar(lar)
Kavram Yanılgısı	BUG-DINO (“BUG”- Deterministic Input Noisy “or” Gate)	(Kuo, Chen, Yang & Mok, 2016)
	BUG-DINO-CR (“BUG”- Deterministic Input Noisy “or” Gate- Constructed-Response)	(Kuo, Chen, Yang & Mok, 2016)
	BUG-DINO-Step (“BUG”- Deterministic Input Noisy “or” Gate-Step)	(Kuo, Chen, Yang & Mok, 2016)
	MC-DINO1 (Multiple Choice-Misconception-Deterministic Input Noisy “or” Gate1)	(Ozaki, Sugawara & Arai, 2020)
	MC-DINO2 (Multiple Choice-Misconception-Deterministic Input Noisy “or” Gate2)	(Ozaki, Sugawara & Arai, 2020)
	MC-DINA-M (Multiple Choice-Deterministic Input Noisy “and” Gate Measuring Misconceptions)	(Elbulok, 2021)
	Beceri + Kavram Yanılgısı	SCIM (Scaling Individuals “and” Classifying Misconceptions)
GDCM-MC (Generalized Diagnostic Classification Models for Multiple Choise)		(DiBello, Henson & Stout, 2015)
SISM (Simultaneously Identifying Skills “and” Misconceptions Model)		(Kuo, Chen & de la Torre, 2018)
MC-DINA-H (Multiple Choice-Deterministic Input Noisy “and” Gate)		(Elbulok, 2021)

Bu bölümde sadece çalışma kapsamında kullanılan beceri ve kavram yanlışlarıyla ilişkili modellere yer verilmiştir.

2.1.4.1. DINA (Deterministic Input Noisy “and” Gate)

Başlangıcı Macready ve Dayton (1977) gibi araştırmacılara uzanan DINA, madde tepki kuramının bir uzantısı olarak Haertel (1989) tarafından ikili (1-0) yetenek modeli olarak tanımlanmıştır. Junker ve Sijtsma (2001) ise modeli literatürde DINA olarak isimlendirmişlerdir (Rupp & Templin, 2008).

Madde tepki kuramı bireylerin sürekli bir dağılım gösteren becerilere sahip olduğunu varsayarken, ikili yetenek modellerine benzeyen DINA ise bireyleri becerilerle önceden ilişkilendirilmiş gizil sınıflar altında kategorize etmektedir (Başokçu, 2011).

DINA'nın bağlayıcı doğası nedeniyle bir bireyin, maddeyi doğru yanıtlayabilmesi için hedeflenen tüm becerilere sahip olması gerekmektedir. Birey ancak tüm becerilere sahip olduğunda “1” puan alır ve birey becerinin varlığına ya da yokluğuna göre “Başarılı sınıf” olarak adlandırılan gizil bir sınıfa atanır. Aksi takdirde birey en az bir beceriye sahip olmadığına bile, hiçbir beceriye sahip olmayan bir bireyle aynı yetenek düzeyinde kabul edilerek “0” puan alacaktır. “Başarısız sınıf” olarak adlandırılan gizil bir sınıfa atanacaktır (Ömür-Sünbül, 2013). Bu yönüyle DINA sahip olduğu beceri hiyerarşisi nedeniyle telafi edici olmayan bir modeldir.

DINA'da maddeler bir ya da birden fazla beceriyle ilişkilendirilerek çok boyutlu bir yapı sergilediğinden bireyler aynı maddeye yanıt vermiş olsalar dahi bireylerden biri başarısız (0) sınıfa atanırken, diğer bireyin başarılı sınıfa (1) atanması olasıdır.

DINA'nın diğer modellere kıyasla en önemli farkı, çoklu beceri gerektiren testlerde en az bir beceriye sahip olmayan bireylerin oluşturduğu başarısız sınıftaki her bir madde için gerekli beceriye sahip olmayan bir bireyi içeren “katı” bir model olmasıdır.

Bu modele göre i . bireyin j . maddeye vermiş olduğu doğru yanıt olasılığı Eşitlik 1'deki formülle tanımlanmaktadır:

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

Eşitlik 1'deki η_{ij} , bireyleri iki gruba ayıran deterministik gizil yanıt değişkenidir ve matematiksel olarak Eşitlik 2'deki gibi ifade edilir:

$$\eta_{ij} = \prod_{k=1}^K \alpha_{ik}^{q_{jk}} \quad (2)$$

$\eta_{ij} = 1$ olduğunda, i . bireyin j . maddede yer alan tüm becerilere sahip olduğu söylenebilir. Aksi bir durumda ise $\eta_{ij} = 0$ 'dır. Olasılık temelli DINA, ilgilenilen beceri sayısı dikkate

alınmaksızın testte yer alan maddelerin her biri için madde düzeyinde iki ayrı madde parametresi üretmektedir. Bu parametreler iki grupta (başarılı sınıf/başarısız sınıf) yer alan bireyler tarafından maddeye doğru yanıt verme olasılığı için iki farklı hata olasılığı olan yanlış pozitif olasılık ve yanlış negatif olasılıkla belirlenmektedir (Rupp & Templin, 2008). Yanlış pozitif olasılık olarak bilinen “s parametresi”, bireyin maddeyi doğru yanıtlamak için gerekli olan beceriye sahip olsa dahi maddeyi yanlış yanıtlama olasılığı olarak tanımlanır. s parametresinin olasılık değeri ne kadar küçük olursa bireylerin ilgili maddeyi doğru yanıtlama olasılığı o kadar artacaktır (Kalkan, 2016; Koyuncu, 2020).

Birey maddeyle ölçülmesi hedeflenen beceri ya da becerilere sahip olmadığı halde maddeyi doğru olarak yanıtlamışsa elde edilen olasılık değerleri, yanlış negatif olasılık olarak bilinen “g parametre” değerleridir. Yani bu durum maddenin gerekli beceriler dışındaki başka becerilerle doğru yanıtlandığı anlamına gelmektedir. Başka becerilerin doğru yanıtlama durumu üzerindeki etkisi Q-matrisin geçerliğine ilişkin bilgiler verebilir. Bu amaçla g parametresinin 1’e yakın değerleri iki şekilde yorumlanabilir:

- 1 - Maddenin gerektirdiği beceriye sahip olmayan bireyler maddeyi doğru yanıtlayabilir,
- 2 - Maddenin gerektirdiği bazı beceriler eksik belirlenmiştir ki bu durum Q-matrisi için bir geçerlik tehdidi oluşturmaktadır. Bu nedenle g parametresine ilişkin elde edilen değerlerin küçük olması istenen bir durumdur. g parametresine ait olasılık değerleri ne kadar küçükse maddenin yanıtlanması için gerekli olan beceriye sahip bireyler tarafından doğru yanıtlanma olasılığı o kadar artacaktır (Kalkan, 2016). g parametresi bu yönüyle madde tepki kuramındaki şans parametresinden ayrılmaktadır (de la Torre & Douglas, 2004).

s ve g parametreleri aşağıda sunulan Eşitlik 3’teki gibi formülleştirilmiştir:

$$\begin{aligned} s_j &= P(X_{ij} = 0 \mid \eta_{ij} = 1) \\ g_j &= P(X_{ij} = 1 \mid \eta_{ij} = 0) \end{aligned} \quad (3)$$

Bireyin belirli bir beceriye sahip olma olasılığı beceriyle ilişkili olan maddelerin güçleriyle ilgilidir (de la Torre, 2008). Madde güçlüğünün yorumlanmasında s ve g parametreleri araştırmacılara yol göstermektedir. Maddenin doğru yanıtının tahminle bulunması ne kadar güçleşirse, madde güçlüğü ciddi derecede artacaktır. Madde güçlüğünün artmasıyla madde zorlaşır, böylelikle maddelerde s parametrelerine ait olasılık değerleri yüksek iken g parametrelerine ait olasılık değerleri küçük çıkacaktır.

s ve g parametrelerine ait olasılık değerlerinin her ikisi de küçük değerler alıyorsa, ilgili maddenin orta güçlükte olduğu söylenebilir. Ayrıca yüksek g parametre olasılık değerleri, maddelerin belirlenen beceriler dışında matriste daha önce tanımlanmamış başka becerilerle de doğru yanıtlanabileceğine işaret etmektedir. Bu durumda tanımlanamayan becerilerin neler olduğu araştırılmalı ve geçerlik endişesi nedeniyle Q-matris yeniden düzenlenmelidir (Kalkan, 2016).

2.1.4.2. BUG-DINO (“BUG”- Deterministic Input Noisy “or” Gate)

Kavramlar arası ilişkilerin güçlü ve çok boyutlu olduğu çeşitli alan derslerinde kavram yanılgıları sistematik hatalara neden olabileceğinden bireyler maddelere yanlış yanıtlar verebilmektedir (Bradshaw & Templin, 2014). Örneğin DINA'nın gerektirdiği koşulları sağlayamayan bireyler maddelere doğru yanıt veremeyeceklerdir. Halbuki birey, maddenin ilişkilendirildiği bilgilere sahip olmakla birlikte bireyde kavram yanılgıları da var olabilir. Bu problemin önüne geçmek için eğer bireyde gerçekte birtakım kavram yanılgıları varsa bu yanılgıların belirlenerek düzeltilmesi gerekmektedir (Kuo, Chen, Yang & Mok, 2016). Kavram yanılgılarını belirleme işi istatistiksel modeller yardımıyla yürütülmektedir. Bu amaçla kullanılan modellerden biri BUG-DINO'dur. BUG-DINO doğası gereği ayrık bir modeldir. Modelde yer alan “Hata/Yanlış-(BUG)” terimi, bir bireyde kavram yanılgılarının var olduğu ya da bireyin bir problemi çözmek için gerekli olan becerilerden hiçbirine sahip olmadığı anlamına gelmektedir. BUG-DINO modelinin ayrık doğası bireyde kavram yanılgılarından en az birinin var olması durumunda belirli bir maddenin doğru yanıtlanamayacağını varsaymaktadır. Bu durumun aksine bireyin madde için gerekli beceriye sahip ancak kavram yanılgısının var olmadığı durumlarda maddeye doğru yanıt vermesi beklenmektedir.

BUG-DINO, her madde için s ve g parametrelerinin kullanıldığı DINO'nun bir modifikasyonudur. İki model aynı formüllere sahip olsa da bireylerin verdikleri yanlış yanıtlardan kaynaklı olarak ortaya çıkabilecek kavram yanılgılarının yorumlanabilmesi için BUG-DINO tahmin ve kayma parametrelerinin ilişkilerini tersine çevirmektedir. Birey *i* tarafından *j* maddesinin doğru yanıtlanma olasılığı kullanılarak, model için tanımlanan matematiksel formül Eşitlik 4'de aşağıda sunulmuştur:

$$P(X_{ij} = 1 | \beta_i) = (1 - s_j)^{1-r_{ij}} g_j^{r_{ij}} \quad (4)$$

Eşitlik 4’de yer alan r_{ij} , bireyleri iki gruba ayıran deterministik gizil yanıt değişkenidir. Matematiksel olarak “ $r_{ij} = 1 - \prod_{l=1}^L (1 - \beta_{il})^{q_{jl}} \cdot r_{ij}$ ” şeklinde tanımlandığında bu durum, i . bireyde j maddesiyle ölçülen kavram yanlışlarından en az birinin varlığına işaret etmektedir. Aksi bir durumda $r_{ij} = 0$ değerini almaktadır. $s_j = P(X_{ij} = 0 | r_{ij} = 0)$ madde parametresi, $r_{ij} = 0$ olduğunda j maddesine yanlış yanıt verme olasılığıdır.

Madde parametresi $g_j = P(X_{ij} = 1 | r_{ij} = 1)$, $r_{ij} = 1$ olduğunda j maddesine doğru yanıt verme olasılığıdır. β_i , i bireyinin kavram yanlışlığı örüntüsünü temsil etmektedir.

2.1.4.3. SISM (Simultaneously Identifying Skills “and” Misconceptions Model)

DINA ve BUG-DINO modelleri sırasıyla becerileri ve kavram yanlışlarını tanımlamak için ayrı ayrı kullanılabilir. Bununla birlikte bireylerin sadece becerilerini tanımlamak, bireylerde var olabilecek kavram yanlışlarını göz ardı edebilmektedir. Bu durum tam tersi için de geçerlidir. Bu nedenle bireylere ilişkin daha eksiksiz tanı bilgileri sağlamak adına sahip olunan beceri ve var olabilecek kavram yanlışlarını aynı anda tanımlamak gerekmektedir. Bazı kavram yanlışları (eğitim-öğretimden kaynaklananlar vb.) bazı becerilerle bir arada bulunmamaktadır. Yani bu yanlışlar belirli becerilerle ilişkilidir. Bu nedenle bireyler bu becerilere ve kavram yanlışlarına aynı anda sahip olamamaktadır. Diğer kavram yanlışları (sezgisel düşünmeden kaynaklanan, bilimsel olmayan kavramlar vb.) bazı becerilerle bir arada bulunabilmektedir. Bu durum kavram yanlışlarının bazı becerilerle ilişkili olmadığı anlamına gelmektedir. Böylece bireylerde bazı kavram yanlışları varken, yine de hedeflenen beceriler kazanılabilmektedir (Kuo vd., 2018).

Bu bölüm çeşitli parametrelerle sınırlandırıldığında DINA ve BUG-DINO olmak üzere iki farklı bilişsel tanı modeli üretebilen doymuş bir model olan SISM’i açıklamaktadır. Literatür incelendiğinde DINA ve BUG-DINO modellerinde bireyler iki gruba ayrıldığından her madde iki parametreye sahiptir. Ancak SISM diğer modellerden farklı olarak bireyleri dört gruba ayırarak her bir birey grubunun maddeye doğru yanıt verme olasılığının sunulduğu dört madde parametresine sahiptir (Kuo vd., 2018; Schellman, 2021).

Bunlar:

- a) Ölçülen tüm becerilere sahip olan ve kavram yanılgıları olmayan bireylerin başarı olasılığı (h_j),
- b) Ölçülen tüm becerilere sahip olan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı (ω_j),
- c) Ölçülen tüm becerilere sahip olmayan ve kavram yanılgıları olmayan bireylerin başarı olasılığı (g_j) ve
- d) Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı (ε_j).

Yukarıda sunulan dört farklı başarı olasılığı okuyucuya kolaylık sağlaması açısından Tablo 4'te özetlenmiştir.

Tablo 4

SISM'de Bireylerin Farklı Başarı Olasılıklarına Ait Dört Madde Parametresi

	Kavram Yanılgısı Yok	Kavram Yanılgısı Var
Becerilerin Tümüne Sahip	h_j	ω_j
Becerilerin Tümüne Sahip Değil	g_j	ε_j

Bu model bireylerin sahip olduğu niteliklerin temelinde kavram yanılgılarının var olabileceğini savunduğundan bireyin sahip olduğu nitelik, kavram yanılgısı ve becerinin birleşiminden oluşmaktadır.

Modele ait madde tepki fonksiyonu dört madde parametresi gerektirmektedir. a_i 'nin ilk K_S maddelerinin becerilere karşılık geldiği ve son $K - K_S$ 'nin kavram yanılgılarına karşılık geldiği varsayılmaktadır.

SISM'in madde tepki fonksiyonu Eşitlik 5'de tanımlanmıştır:

$$P(X_{ij} = 1|a_i) = h_j^{\eta_{ij}(1-\nu_{ij})} \omega_j^{\eta_{ij}\nu_{ij}} g_j^{(1-\eta_{ij})(1-\nu_{ij})} \varepsilon_j^{(1-\eta_{ij})\nu_{ij}} \quad (5)$$

Eşitlik 5 incelendiğinde η_{ij} ve γ_{ij} parametreleri sırasıyla $\eta_{ij} = \prod_{k=1}^{K_S} a_{ik}^{q_{jk}}$ ve $\gamma_{ij} = \prod_{k=K_S+1}^K (1 - a_{ik})^{q_{jk}}$ olarak tanımlanan deterministik gizil tepki değişkenleridir. SISM, bireyleri η_{ij} ve γ_{ij} 'nin farklı kombinasyonlarına göre dört gizil gruba ayırmaktadır.

Dört parametre matematiksel olarak Eşitlik 6'daki gibi tanımlanmaktadır:

$$\begin{aligned} h_j &= P(X_{ij} = 1 | \eta_{ij} = 1, \gamma_{ij} = 0), \\ \omega_j &= P(X_{ij} = 1 | \eta_{ij} = 1, \gamma_{ij} = 1), \\ g_j &= P(X_{ij} = 1 | \eta_{ij} = 0, \gamma_{ij} = 0), \\ \varepsilon_j &= P(X_{ij} = 1 | \eta_{ij} = 0, \gamma_{ij} = 1). \end{aligned} \quad (6)$$

SISM'e ait madde parametrelerini tahmin etmek için "beklenti-maksimizasyon algoritması" geliştirilmiştir. Bu algoritmanın hesaplama detaylarına Kuo vd. (2018)'in çalışmalarından ulaşılabilir.

2.1.4.3.1 SISM'e İlişkin Özel Durumlar

Bir madde aynı anda beceri ve kavram yanılgılarını ölçüyorsa ve bu yanılgılardan bazıları ölçülen tüm becerilerle ilişkili değilse, SISM'de dört madde parametresine ihtiyaç duyulmaktadır. Bununla birlikte ölçülen tüm becerilere sahip olan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığını yansıtan " ω_j " madde parametresi modelden çıkarılabilmektedir. Böyle bir durumda SISM indirgenmiş bir model haline gelmekte ve rSISM (Reduced Simultaneously Identifying Skills "and" Misconceptions Model) olarak adlandırılmaktadır.

SISM'de bir madde yalnızca beceriyi ölçüyor ancak herhangi bir kavram yanılgısını ölçmüyorsa, γ_{ij} 'nin değeri her zaman 0'a eşittir ve SISM'in formülü DINA'ya eşdeğer olan $P(X_{ij} = 1 | a_i) = h_j^{\eta_{ij}} g_j^{(1-\eta_{ij})}$ olarak yazılabilmektedir. Aksine bir madde yalnızca kavram yanılgılarını ölçüyorsa η_{ij} 'nin değeri her zaman 1'e eşittir ve SISM'in formülü BUG-DINO'ya eşdeğer olan $P(X_{ij} = 1 | a_i) = h_j^{(1-\gamma_{ij})} \omega_j^{\gamma_{ij}}$ olarak yazılabilir. Bu nedenle DINA ve BUG-DINO, γ_{ij} ve η_{ij} sırasıyla 0 ve 1 olarak ayarlandığında SISM'den elde edilebilir. Ancak bazı durumlarda kavram yanılgıları ve beceriler bağımsız olsa da diğerleri birbirleriyle ilişkili olabilir. Bu nedenle Kuo vd. (2018) olası olmayan bilgi durumlarını modelde sifira sabitlemişlerdir.

2.1.5. Kavram Yanılgıları

Birey günlük yaşamında karşılaştığı olgu ya da olaylarla kavramlara ilişkin anlam ve açıklamaları yorumlar ve bilgiyi yapılandırır. Bireyin öğrenme ortamında kavramlara ilişkin bilimsel olmayan açıklamaları varsa bilgiyi yapılandırma süreci olumsuz olarak etkilenmektedir (Şenyiğit, 2020). Bu durumun ortadan kaldırılabilmesi için kavramlara ilişkin olumsuz düşüncelerin tespit edilerek giderilmesi gerekmektedir.

Öğretim süreci içerisinde bireyin sahip olduğu düşünceler, öğreticinin sunduklarıyla ilişki kurarken karmaşık bir boyut kazanabilir ve zamanla düşüncelerle bilimsel görüşlerin yer değiştirmesi noktasında birey direnç gösterebilir (Mintzes, 1984). Yaşanan bu durum eski bilgiler ve yeni bilgiler arası ilişkiyle kurulan anlamlı öğrenmelerde, meydana gelen kopukluk veya yanlışlıkları içermekte ve kavramlara ilişkin birtakım yanılgılara sebep olmaktadır (Koçak, 2020). Kavramlara ilişkin yanılgılar literatürde farklı şekillerde tanımlanmaktadır. Bazı araştırmalarda bireyin sistemli şekilde hatalar ürettiği bir kavrayış olarak tanımlanırken (Zembar, 2008b; Akt. Koçak, 2020); Yenilmez ve Yaşa (2008), Çakır ve Yürük (1999) çalışmalarında kavram yanılgılarını öğrenme sürecini etkileyen engeller olarak görmüşler ve bu kavramı bilimsel gerçeklerle tutarsız bireyin deneyimleri sonucu oluşan, kavramlara ilişkin öğretim sürecini olumsuz etkileyen bilgiler olarak ifade etmişlerdir.

Yanılgıların tespiti ve tespit sonrası yapılacak doğru değerlendirmeler için öncelikle belirli kavramlar (hata, kavram karmaşası gibi) ve yanılgılar arasındaki farkın iyi anlaşılması gerekmektedir. Özellikle yanılgıların, birey yanıtlarına karışan rastgele hatalardan farklı olduğu hususu önemli bir konudur. Kavram yanılgıları ve rastgele hatalar arasındaki farkın bir örneği çıkarma alıştırmaları olan “ $26-17=9$ ” üzerinden sunulmuştur. Birey doğru yanıt olan “9” yerine “8” yanıtını verdiğinde eğer hata benzer çıkarma alıştırmalarıyla ilişkili maddeler için tekrarlanmıyorsa rastgele bir hatadır. Ancak aynı örnek için başka bir birey, ilk sütunda çıkarmak için büyük alt rakam 7’yi üste ve daha küçük üst rakam 6’yı alta çevirerek 11 yanıtını vermiş ve bu durum diğer maddeler için tekrarlanmışsa bu noktada bir hata örüntüsü oluşacak ve bu örüntü “büyükten küçüğe” kavram yanılgısı olarak adlandırılacaktır (Elbulok, 2021). Görüldüğü üzere kavram yanılgısıyla hata birbirinden farklıdır. Hata, birey tarafından istenmeden yapılan bir davranış ya da yanılgı şeklindedir. Birey kavram ile ilgili bilgi eksikliğine sahip olduğu için yanlış bir yanıt vermişse bu bir kavram yanılgısı değildir. Aksine birey ilgili kavramda yanılgıya sahip olursa düşüncelerinin doğru olduğuna ilişkin açıklamalar sunarak direnç gösterebilir. Bu gibi durumlarda kavram yanılgılarının varlığı

göz ardı edilmemelidir (Ural, 2017; Akt. Parlak, 2019). Özetle bireylerde var olan kavram yanlışlarının tümünün kavrama ilişkin edinilmiş hatalar olduğunu ifade etmek mümkündür ancak bireyde var olan hataların her koşulda kavram yanlışlığı olmadığını da vurgulamak gerekmektedir (Bingölbali & Özmantar, 2009).

Kavram yanlışlığıyla karıştırılan bir diğer kavram ise kavram kargaşasıdır. Belirli bir kavrama ilişkin yapılan tanımlamalarda, birden fazla kelimenin kullanımı ya da bir kelimenin birden çok kavram için kullanımı kavram kargaşasını ortaya çıkarabilir ve bu durumun ortadan kaldırılması çoğu zaman mümkün olmamaktadır. Bireyler kavramlara ilişkin belirli bilgi şemalarına sahip olabilir. Bireye ait bilgi şemaları doğru bilgiyi, hatayı, kavram kargaşalarını ya da kavram yanlışlarını içerebilir. Öğreticilerin, bireylerin kavramlarla ilişkili olan bilgi şemalarından haberdar olmaları öğrenme sürecini planlama aşamasında büyük önem taşımaktadır. Bu nedenle öğretmenler kavram öğretiminde bu durumun bilincinde olacak şekilde bireylerin zihinsel süreçlerinde yaşam boyu edindikleri deneyimlerle ilgili kavramlara ilişkin açıklamaların hangi durumda olduğuyla ilgili bilgi sahibi olmalıdır (Alwan, 2011). Daha sonra bu süreci iyi planlamalı, kavramlara ilişkin tanımlamalar ve açıklamalarda olabildiğince net olmalıdır (Ülgen, 2004). Ancak unutmamak gerekir ki kavram yanlışları her zaman bireylerin zihinsel süreçlerinden kaynaklı olarak oluşmamaktadır. Bazı durumlarda öğretici, öğrenme ortamı ve bireylerin bu ortama ilişkin tutumları da kavram yanlışlarının oluşmasına neden olmaktadır (Koçak, 2020). Bu yanlışların düzeltilmesi gerekmektedir. Aksi takdirde kavramsal öğrenmenin meydana gelmesi mümkün olmayacaktır (Driver & Erickson, 1983). Bu durum öğretim sürecinde kavram yanlışlarına verilmesi gereken dikkatin neden önemli olduğunu vurgular niteliktedir.

Literatürde kavram yanlışlarının önemi çeşitli çalışmalarla desteklenmiştir. Örneğin Schulte (2001) çalışmasında kavram yanlışlarının önemine vurgu yaparak, yanlışların oluşturduğu çeşitli engellerin varlığında bireylerin öğretim süreci içerisinde karşılaşılabilecekleri problem durumlarına ilişkin önlemler almak amacıyla ilköğretim sürecinden başlayarak temel kavramların bireylere doğru bir şekilde aktarılmasının önemli olduğunu ifade etmiştir. Eğer aktarım doğru bir şekilde yapılmazsa bireylerin sahip oldukları kavram yanlışları öğretim sonucu kazanılan yeni kavramlarla ilişki kurmada ya da ilişki kursa dahi amacı eski bilgilerle yeni bilgileri birleştirmek olan anlamlı öğrenmelerin sağlanması bakımından çeşitli sorunlara yol açacaktır. Ortaya çıkan yanlışların doğru bir şekilde düzeltilmesi yani eski öğrenmelerle daha sonraki öğrenmeler arasında kurulan

ilişkinin anlamlı olabilmesi için öncelikle kavram yanlışlarının tespit edilmesi gerekmektedir. Zembat (2008) kavram yanlışlarının ne olduklarının listelenmesinden ziyade bu yanlışlara neden olan durumların incelenerek analiz edilmesi, gerekli çıkarımlardan elde edilen bilgilerle öğretim sürecinin desteklenmesi ve gerektiğinde avantaja çevrilmesi noktasında öğreticilere önemli görevler düştüğünü ifade etmiştir (Akt. Koçak, 2020). Baki (2006) ise bireyde var olan kavram yanlışısındaki çelişkinin öğretici tarafından bireye sunulmasıyla birey zihnindeki bilginin çeşitli öğretim stratejileriyle yapılandırılması gerektiğini vurgulamıştır.

Kavram yanlışlarının tespiti ve giderilmesinde kavram haritaları, anlam çözümleme tabloları, v diyagramları, onarım teorisi ve kavramsal değişim yaklaşımları gibi çeşitli yöntemler kullanılmaktadır (Anıl & Küçüközer, 2010; Smith, Blakeslee & Anderson, 1993). Bu yöntemlerden kavram haritaları sayesinde özellikle kavramsal düzeydeki öğrenmeler için eski ve yeni kavramlar arasında ilişkiler kurularak bireylerin kavramlara ilişkin eksiklikleri belirlenerek öğrenmeleri desteklenmektedir (Chen, Lin & Nien, 2014). Ayrıca kavramlara ilişkin yanlışların zamanında giderilebilmesi adına *-öğretim süreci devam ederken ya da öğretim süreci sonunda-* ilgili kavrama ilişkin edinilen bilgilerde ne tür hataların olduğuna yönelik kapsamlı çıkarımlarla birlikte bireylerin düşüncelerinin değiştirilip sağlamaştırılması noktasında önlemler alınabilir (McIntosh & Draper, 2001). Kavramsal değişimde bireylere kavram ile ilgili kendi fikirlerini sunabilecekleri bir tartışma ortamının oluşturulması gerekmektedir (Smith vd., 1993). Böylelikle bireylere direnç gösterdikleri kavramlara ilişkin yanlışları düzeltebilmek adına zihinlerinde değişimin oluşması için imkân verilmektedir (Pines & West, 1986). Bu imkân sayesinde bireylerden kavramsal değişimin bir parçası olarak kendilerinde var olan bilginin yetersizliğinden rahatsızlık duyarak, kavramlar arası ilişkileri düzenleyip yeniden organize etmeleri beklenmektedir (Posner, Strike, Hewson & Gertzog, 1982).

2.2. İlgili Araştırmalar

Klasik test kuramı ve tek boyutlu madde tepki kuramı gibi geleneksel psikometrik yöntemler tek bir sürekli gizil yeteneğin ölçülmesine odaklandığından bu yöntemler kavram yanlışlarıyla ilişkili iki stratejiyi kullanmaktadır. Bunlar:

- 1) Kavram yanılgısının ölçülmesine ilişkin testte yer alan bir maddede ilgili seçeneğin bireyler tarafından kaç kez seçildiğinin hesaplanmasıyla her bir yanılgı bir alt puanla değerlendirilerek çeldirici analizi yapılmaktadır (Khazanov, 2009).
- 2) Madde tepki kuramı kapsamında kavram yanılgılarının tanısında madde karakteristik eğrileri kullanılmaktadır. Bu eğriler bireysel olarak madde ve birey bazında incelenmektedir (Sadler, 1998).

Yukarıda verilen stratejilere ilişkin çeşitli sınırlamalar bulunmaktadır. Örneğin kavram yanılgısı az sayıda maddeyle ölçüldüğünde ortaya çıkan yanılgılar güvenilir olabilmektedir (Haberman, Sinharay & Puhan, 2009). Öğretime ilişkin kararların alınmasında alt puanlar kullanılmaktadır. Alt puanlara ilişkin güvenilir olmayan kararlar öğretim stratejileriyle kaynaklarını yanlış yönlendirebilmekte ve kavram yanılgısına ilişkin alt puanların kullanımı bir bireyde kavram yanılgısının var olup olmadığı sorusuna doğrudan yanıt veremeyebilmektedir.

Geleneksel yöntemlerde genel yetenek tahminlerinde değişken, hatayı ölçmek ve maddelerin genel yeteneği ölçmek için ne kadar iyi işlev gördüğü hakkında bilgi sağlamak açısından avantajlara sahip olsa da madde tepki kuramı kavram yanılgılarının ölçülmesi ya da testin/maddenin kavram yanılgılarının ölçülmesindeki etkililiğini değerlendirmesi noktasında faydalı değildir. Madde tepki kuramında madde karakteristik eğrilerinin incelenmesi bireylerde hangi kavram yanılgılarının var olduğuna ilişkin karar vermek ya da tanımlamalar yapmak için maddeler arasında toplanan bilgileri kullanmaz. Ayrıca kavram yanılgılarının belirlenmesinde kullanılan çok boyutlu madde tepki kuramı modelleri, parametre tahmini ve çoklu gizil değişkenlerin ölçeklendirilmesi noktasında bilişsel tanı modellerine kıyasla çok fazla veri gerektirir. Bu nedenle hem klasik test kuramı hem de madde tepki kuramına ilişkin yöntemler, kavram yanılgılarıyla ilgili açıklayıcı bilgiler sağlayabilir ancak bazı durumlarda bu yöntemler bireylerde kavram yanılgılarının var olup olmadığını doğrudan belirleyemeyebilir ya da bir maddenin kavram yanılgısını ne kadar iyi ölçtüğü noktasında doğru bilgiler sunamayabilir. Yanılgıların tanısı öğretim uygulamalarını bilgilendirmek için önemlidir ve maddelerden elde edilen bilgiler kavram yanılgılarının geçerli tanısını veren testler için önem arz etmektedir. Bu nedenle geleneksel yöntemlere göre daha yeni olan bilişsel tanı modelleri kavram yanılgılarının modellenmesi noktasında avantajlar sağlamaktadır. Gözden kaçırabilecek problem çözme becerileri ve bilişsel süreçler gibi madde performansına ilişkin bilgiler sunan bu yeni modeller, sağladıkları profil merkezli yaklaşım sayesinde bireyde var olabilecek kavram yanılgılarına ilişkin potansiyel

bilgiler sunduğundan (de la Torre, 2009b) kullanılan bilişsel tanı modellerinin performansının incelemesinin gerekli olduğu düşünülmüştür. Bu nedenle bu bölümde bilişsel tanı değerlendirmelerinde bilişsel tanı modelleri kapsamında kavram yanlışlarının incelendiği çalışmalara yer verilmiştir.

Kavram yanlışlarının bilişsel tanı modellerindeki gelişimi 1970'li yıllara dayansa da bu modellerin sınırlı kullanımları olmuştur (Brown & Burton, 1978). Bunlardan biri Lee, Park ve Taylan (2011) tarafından gerçekleştirilen çalışmada görülmektedir. Araştırmacılar çoktan seçmeli testlerdeki kavram yanlışlarının, bilişsel tanı modellerinde kullanılan çeldiricilerin genel analizleriyle değerli bilgiler sağlayabileceğini vurgulamışlardır. “Trends in Mathematics and Science Study (TIMSS)” 2007 veri seti üzerinde yürütülen çalışmada, bireylerin ilgili niteliklere sahip olup olmadıklarını modellemek için DINA modeli kullanılmış ve çeldirici analizleri yapılmıştır. Çeldirici analizlerinin amacı, modele dayalı olarak niteliklere sahip olan bireylerin bu nitelikleri gerektiren bazı maddelere neden yanlış yanıt verdiklerini değerlendirmektir. Sonuçlar gerekli niteliklere sahip olan bireylerin gerekli niteliklere sahip olmayan bireylere göre belirli çeldiricileri seçme olasılıklarının daha yüksek olduğunu göstermiştir. Araştırmacıların bulguları, çeldiricileri bilişsel tanı modelleriyle değerlendirmenin önemi için daha fazla gerekçe sağlamıştır. Mevcut çalışmada önerilen analizler, çeldiricileri kısmi bilgi gerektirenlerden (beceriler) ve hatalı öğrenilmiş becerileri gösterenlerden (kavram yanlışları) daha fazla ayırt etmeyi amaçlamıştır.

Bu sınırlı kullanımların aksine son zamanlarda popüler hale gelen kavram yanlışlarına ilişkin artan ilgi, genişleyerek bilişsel tanı modelleri kapsamında çeşitli araştırmalara konu olmaya başlamıştır. Bradshaw ve Templin (2014) tarafından tanıtılan “Scaling Individuals “and” Classifying Misconceptions (SCIM)” olarak ifade edilen model bu araştırmalardan biridir. Bu model SISM gibi ikili yanıtlara ait verileri değil, seçilen belirli yanıtlara ilişkin verileri kullanmaktadır. Madde tepki kuramıyla bilişsel tanı modellerinin sentezlenmesiyle tasarlanan nominal bir tepki modelidir ve gizil becerilerin ölçümüne ek olarak kavram yanlışlarını belirlemek için kullanılmaktadır. Bu çalışmada model tanıtımı yapıldıktan sonra “Markov Zinciri Monte Carlo Tahmin Algoritması” kullanılarak bir simülasyon çalışmasıyla modelin özellikleri değerlendirilmiştir. Ardından SICM'in uygulanabilirliği gerçek bir veri seti üzerinde incelenmiştir. Simülasyon çalışması kapsamında kavram yanlışlarının sayısı (3 veya 6), örneklem büyüklüğü (3.000 veya 10.000), test uzunluğu (30 veya 60 madde) ile yetenek ve kavram yanlışları için ana etkilerin büyüklüğü (düşük veya yüksek) değişimlenmiştir. Birey yeteneği standart bir normal dağılımdan örneklenirken

kavram yanılgıları arasındaki korelasyonlar, ilgili ancak farklı yapıları yansıtmak için ,5 olarak sabitlenmiştir. Araştırmacılar örneklem büyüklüğü arttığında ve test tasarımının karmaşıklığı azaldığında parametre tahminlerinin azaldığını bulmuşlardır. Ayrıca sonuçlar bireylerin kavram yanılgılarını belirleme ve birey yeteneklerini eş zamanlı olarak ölçeklendirmek için maddeler ve yapılar arasındaki ilişkinin güçlü olması gerektiğini göstermiştir. Çalışmanın ikinci aşamasında araştırmacılar SCIM modeli 10.000 liseliden alınan yanıtları içeren ampirik verilere uyarlamışlardır. Bu aşamada SICM ve bir madde tepki kuramı modeli (Nominal Response Item Response Theory - NR IRT) kullanarak verileri analiz etmişlerdir. Çalışma, ampirik veri analizinde yalnızca bir madde tepki kuramı yaklaşımı kullanmak yerine bilişsel tanı modeli ve madde tepki kuramı yaklaşımlarını birleştirmenin faydalarını göstermiştir. Bilişsel tanı modeli yalnızca bir madde tepki kuramı yaklaşımı kullanıldığında, kaybolan maddeler ve bireyler hakkında ek bilgiler sağlamıştır. Ayrıca sonuçlar SICM'in spesifik test tasarım özellikleriyle birleştğinde tipik olarak mevcut modelleme ve test prosedürleri tarafından sağlanan ölçeklendirilmiş bir yetenek tahmini sağlamanın yanı sıra kavram yanılgılarını detaylandıran tanı puan raporlarını etkinleştirebileceğini vurgulamaktadır. Uygulamada ise kavram yanılgısı tanılarının iyileştirme amacıyla kullanılabilmesi ifade edilmiştir. Ancak SICM modeli birçok parametreyi ve iki farklı birey parametre tahminini içerdiğinden büyük örneklem değerlerine sahiptir. Araştırmacılar bu durumun biçimlendirici geri bildirim istendiği ve büyük örneklem elde edilemediği birçok test durumunda modelin uygulanabilirliğini sınırlandırabileceğini ifade etmişlerdir.

DiBello, Henson ve Stout (2015) çalışmalarında bilişsel tanı değerlendirmeleri kapsamında çoktan seçmeli maddeleri analiz eden bir model (Generalized Diagnostic Classification Models for Multiple Choice/GDCM-MC) geliştirmişlerdir. Bu model beceri ve kavram yanılgılarının bir arada yer aldığı genişletilmiş bir Q-matris içermektedir. Yazarlar ayrıca, 3-4 beceri ile 3-6 kavram yanılgısını ölçen 10 ile 30 madde içeren dört Q-matris tasarımı kullanıldığında kısa testlerle GDCM-MC'nin performansını değerlendirmek için bir simülasyon çalışması yürütmüşlerdir. Örneklem büyüklüğü (500, 1000 veya 2000) ve Q-matris tasarımını, toplam 12 koşul için değiştirmişlerdir. Simülasyon çalışması sonuçları, daha iyi test koşullarında GDCM-MC'nin madde parametre tahminlerinin azaldığını ve sınıflama doğruluğu değerlerinin arttığını göstermiştir.

Shear ve Roussos (2017) ise çalışmalarında DiBello ve diğerleri (2015) tarafından geliştirilen GDCM-MC modelini bireylerin geometrideki kavram yanlışlarını belirlemek için klasik test kuramı tabanlı kesme puanlarını kullanmak üzere tasarlanmış çeldirici yaklaşımına sahip bir test olan Tanısal Geometri Değerlendirme Proje'sinden alınan alt veri kümesine uyarlamışlardır. Araştırmacılar GDCM-MC analizlerinden elde edilen sınıflamaların klasik test kuramı tabanlı kesme puanlarını desteklediğini ve kavram yanlışlarını belirlemede geleneksel kavram envanterlerinden alınan verilerin kullanımının, bilişsel tanı modelleri kapsamında başarılı bir uyarlama olduğu bulgusuna ulaşmışlardır.

Kuo vd. (2016) çalışmalarında hem yapılandırılmış yanıtli (constructed-response/CR) hem çoktan seçmeli (multiple-choice/MC) maddelerle oluşturulmuş testler için yeni bilişsel tanı modellerini (DINA-CR, DINA-Step, BUG-DINO, BUG-DINO-CR ve BUG-DINO-Step) tanıtmışlardır. Bu modellerin verileri analiz etmek için nasıl kullanılacağı ve bireylerin matematik alanındaki beceri ve kavram yanlışlarını nasıl tanımlayacağı gösterilmiştir. Yeni modellerin kullanımının gösterilmesi amacıyla gerçek verilerin analizi, dört Tayvanlı ilkokuldan rastgele seçilen 497 kişilik 6. sınıf öğrencilerinin sekiz maddeye verdikleri yanıtlar üzerinden incelenmiştir. Sonuçlar DINA-CR/BUG-DINO-CR modellerinin beceri ve kavram yanlışlarının tanısında, DINA/BUG-DINO modellerinden daha iyi olduğunu ve DINA-Step/BUG-DINO-Step modellerinin beceri ve kavram yanlışlarının tanısında, DINA-CR/BUG-DINO-CR modellerinden daha iyi olduğunu göstermiştir. Bu çalışma önerilen bilişsel tanı modellerinin becerileri ve kavram yanlışlarını belirlemede geleneksel bilişsel tanı modellerinden daha etkili olduğuna dair kanıtlar sağlamaktadır. Pratikte bir test yapılandırılmış yanıtli maddeleri içeriyorsa DINA-CR/BUG-DINO-CR ve DINA-Step/BUG-DINO-Step modelleri, DINA/BUG-DINO modellerinden daha iyi tanı performans sergilemektedir. Son olarak araştırmacılar bilişsel tanı modellerini verimli bir şekilde kullanmak ve puanlamak için bilgisayarların kullanımını savunmaktadır. Bununla birlikte bilişsel tanı için bilgisayarlarda uyarlanabilir testlerin uygulanmasının, önerilen modeller için uygun madde seçim algoritmalarının yanı sıra daha büyük bir madde havuzunu gerektireceğinin unutulmaması gerektiği çalışmada vurgulanmıştır.

Kuo vd. (2018) tarafından bireylerin sahip oldukları becerileri ve kavram yanlışlarını eş zamanlı olarak belirleyebilen yeni bir bilişsel tanı modeli geliştirilmiştir. Ayrıca model parametrelerini tahmin etmek için beklenti maksimizasyon algoritmasının kullanılması önerilmiştir. Önerilen model ve algoritmanın yaşayabilirliğini değerlendirmek için bir simülasyon çalışması yapılmıştır. Gerçek veriler, bir uyarlama çalışmasıyla önerilen

modelin uygulanabilirliğini göstermek ve mevcut bilişsel tanı modelleriyle karşılaştırılmasını sağlamak için analiz edilmiştir. Model bağlamında doğru sınıflama oranlarının nasıl arttırılabileceğini belirlemek için araştırmacılar tarafından gerçek veri tabanlı bir simülasyon çalışması yürütülmüştür. Simülasyon çalışması doğru parametre tahminlerinin elde edilebildiğini özellikle örneklem büyüklüğünün nispeten büyük olduğu, testin uzun olduğu ve madde ayırt edicilik düzeyinin yüksek olduğu koşullarda SISM'in güvenilir olduğunu göstermiştir. SISM'den elde edilen beceri ve kavram yanılgılarının sınıflama doğruluğu değerlerinin ise orta ile yüksek arasında değiştiği gözlenmiştir. Gerçek veri analizinden elde edilen bulgular beceri ve kavram yanılgılarına ilişkin SISM ile elde edilen sınıflama doğruluğu değerlerinin, DINA ve BUG-DINO modellerinden elde edilen sınıflama doğruluğu değerlerinden daha yüksek olduğunu göstermiştir. Çalışma bulgularında vurgulanan önemli bir nokta ise SISM'in kısa testlerin kullanıldığı koşullarda dahi kabul edilebilir bir sınıflama doğruluğu sağlamasıdır. Araştırmacılar gerçek veri tabanlı simülasyon çalışmasının sonuçlarına dikkat çekerek test uzunluğunun artması ve daha iyi tasarlanmış bir Q-matrisi kullanılarak sınıflama doğruluğu değerlerinin arttırılabileceğini savunmuşlardır. Son olarak çalışmada problem çözmek için çeşitli stratejiler kullanan bireyler farklı beceri ve kavram yanılgısı alt kümelerini içerebildiklerinden, bilişsel tanı modellerine problem çözenin birden fazla yolu dahil edilebilirse bireylere ve öğreticilere sağlanan tanısal geri bildiriminin daha yararlı olacağı vurgulanmıştır.

Ozaki, Sugawara ve Arai (2020) çalışmalarında çoktan seçmeli veriler için bireylerin kavram yanılgılarını tahmin edebilen modellere ihtiyaç duyulduğunu vurgulamışlar ve iki yeni istatistiksel model geliştirmişlerdir. Bu modeller kavram yanılgılarını tahmin etmede çoktan seçmeli maddeler için DINO modelleri olarak adlandırılan MC-M-DINO1 (Multiple Choice-Misconception-Deterministic Input Noisy “or” Gate1) ve MC-M-DINO2 (Multiple Choice-Misconception-Deterministic Input Noisy “or” Gate2)'dir. Çalışmada üç model (MC-M-DINO1, MC-M-DINO2 ve BUG-DINO) bir simülasyon çalışmasında karşılaştırılmış ve geliştirilen iki yeni model “Okuma Becerileri” testinin uygulandığı gerçek veriler üzerinden analiz edilmiştir. Bulgular geliştirilen iki yeni modele ilişkin parametre değerlerinin, BUG-DINO modelden elde edilen parametre değerlerine kıyasla daha yüksek bir performansa sahip olduğunu göstermiştir.

Norris (2021) çalışmasında Fizik alanında “Kuvvet Değerlendirmesi” konusunda bireylerin bilgi ve kavram yanılgılarının ölçülebilmesi amacıyla sırasıyla iki parametrelilikli lojistik model ve DINA modelinin birey yanıtlarına uyduğu bir test geliştirmiştir. Araştırmacı bilgiyi

sürekli bir gizil yapı olarak ve kavram yanılgılarını bir dizi ayrı beceri olarak ölçen bilişsel tanı değerlendirmeleri için yeni bir test formatının etkililiğini araştırdığını ifade etmiştir. Geliştirilen test Fizik alanında üniversitede en az iki yarıyılı tamamlamış 449 lisans öğrenimi gören bireylerin kavram yanılgı profillerini ve bilgi puanlarını tahmin etmek için kullanılmıştır. Bu çalışma, farklı bilgi puanlarına ve farklı Fizik eğitimi seviyelerine sahip bireyler için 6 kavram yanılgısının var olma olasılıklarını karşılaştırmıştır. Elde edilen bulgular, 6 kavram yanılgısının her birinin var olduğu ve var olmadığı bireyler arasında bilgi puanlarında anlamlı farklılık gösterirken yüksek bilgi puanına sahip bireylerin büyük bir kısmında hala kavram yanılgılarının var olduğunu göstermiştir. Özellikle daha ileri düzeydeki Fizik derslerini tamamlayan bireylerde, tamamlamayan bireylere göre kavram yanılgılarının var olma olasılığı daha düşükken, yine de tüm eğitim seviyelerindeki bireylerde hala kavram yanılgılarının var olduğu bulgusuna ulaşılmıştır. Araştırmacı bu testle, bireylerin kavram yanılgılarını teşhis etme sorununa önceki çalışmalardan farklı bir şekilde yaklaştığını ifade etmiştir. Yeni, daha karmaşık bir ölçme modeli oluşturmak yerine bilgiyi ve kavram yanılgılarını ölçmek için ayrı test maddelerini kullanan karmaşık bir test yapısı önermiştir.

Elbulok (2021) çalışmasında Lee ve Corter (2011) tarafından sunulan Bayes Ağlarıyla literatüre sunulan kavramsal katkıları, beceriler ve kavram yanılgılarını kullanarak bilişsel tanı modelleri çerçevesine genişletmiştir. Bunun için de la Torre (2009b) tarafından geliştirilen modeli (Multiple Choice-Deterministic Inputs Noisy “and” Gate/MC-DINA) daha da geliştirerek beceri ve kavram yanılgılarını eş zamanlı olarak ölçmek için yeni bir model (Multiple Choice-Deterministic Inputs Noisy “and” Gate-Hierarchical/MC-DINA-H) önermiştir. Araştırmacı aynı zamanda bu modelin *-yanıtlar seçilen belirli yanıtları içerecek şekilde genişletildiğinde-* SISM’in özel bir durumu olduğuna vurgu yapmıştır. Model performansı, örneklem büyüklüğü ve test uzunluğu gibi farklı simülasyon koşulları altında değerlendirilmiş ve MC-DINA-H daha az karmaşık olan farklı modellerle (DINA, BUG-DINO, MC-DINA-S (MC-DINA measuring subskills), MC-DINA-M (MC-DINA measuring misconceptions) karşılaştırıldığında modelin nitelikleri daha doğru sınıflandırdığı bulgusuna ulaşılmıştır.

Schellman (2021) çalışmasında dört yıllık hibe destekli bir projenin kavram yanılgılarının tanısı için özel olarak geliştirilmiş verilerini kullanarak veri setleri üretmiş ve örneklem büyüklüğü, ölçülen nitelik sayısı, test uzunluğu ve madde ayırt edicilik düzeyi olmak üzere çeşitli koşullarda madde parametresi tahmin doğruluğu, sınıflama doğruluğu ve sınıflama

güvenirliđi dikkate alındığında BUG-DINO modelinin performansını deđerlendirmiřtir. alıřmanın sonuçları test uzunluđu ve madde ayırt edicilik düzeyi arttıķa BUG-DINO modelinin performansının arttıđını ancak ölçülen nitelik sayısı arttıķa model performansının bundan olumsuz yönde etkilendiđini ortaya koymuřtur. Ayrıca bu modelde madde ayırt edicilik düzeyi yüksek olduđunda elde edilen bulgular, 350 ve 1000 kişilik örneklem büyüklüđünde karşılařtırılabilir tahmin dođruluđu ve sınıflama güvenilirliđi deđerleri sunmuřtur.

Biliřsel tanı deđerlendirmeleri kapsamında kavram yanılgılarına iliřkin literatür deđerlendirildiđinde yukarıda örnekleri sunulan alıřmaların birkaç alıřma hari (Elbulok, 2021; Kuo vd., 2018) sadece kavram yanılgılarına odaklandıđı görölmektedir.



BÖLÜM III

YÖNTEM

Çalışmanın bu bölümünde araştırmanın modeli, verilerin üretimi ve verilerin analizine ilişkin bilgilere yer verilmiştir.

3.1. Araştırmanın Modeli

Çalışmada, beceri ve kavram yanılgılarının varlığını eş zamanlı olarak ölçmede Kuo vd. (2018) tarafından geliştirilen SISM'in performansını değerlendirmek amacıyla çalışma koşullarının oluşturulması için dört faktörün değişimlendiği veriler üretilmiştir. Bu çalışma üretilen verinin kullanımı ve model için ideal durumların tahmini açısından bir simülasyon çalışmasıdır. Simülasyon çalışmasının amacı modelin pratikte kullanımı için ele alınması gereken test uzunluğu, örneklem büyüklüğü, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısı gibi soruları yanıtlamaktır. Bu koşullar dikkate alınarak çalışmada modelin geçerliğine ilişkin kanıtlar toplama noktasında madde parametrelerinin tahmininde modelin genel performansı değerlendirildikten sonra SISM'in kullanıldığı durumlarla becerileri ve kavram yanılgılarını ayrı ayrı ölçen DINA ve BUG-DINO modellerinin sınıflama doğrulukları farklı koşullar altında karşılaştırılmış ve elde edilen bulgular tüm yönleriyle incelenerek raporlanmıştır. Çalışma bu yönüyle betimsel bir çalışma olarak da ele alınabilir (Fraenkel, Wallen & Hyun, 2012).

3.2. Verilerin Üretimi

Bu bölüm çalışma kapsamında verilerin üretimi başlığı altında yanıt örüntülerinin çeşitli faktörler altında değişimlendiği simülasyon koşullarıyla Q-matrise ilişkin özelliklerin tanımlandığı bilgileri sunmaktadır.

3.2.1. Simülasyon Koşulları

SISM'in performansına ilişkin kapsamlı bilgilere ulaşmak amacıyla örneklem büyüklüğü, test uzunluğu, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısı olmak üzere farklı simülasyon çalışmalarında (de la Torre & Douglas, 2004; Elbulok, 2021; Kuo vd., 2018; Ozaki vd., 2020; Schellman, 2021) sıklıkla kullanılan dört faktör çalışmaya dahil edilmiştir. Bu koşullar tahmin algoritmasının doğru olup olmadığını belirleme ve büyük örneklem için dikkatlice oluşturulmuş testlerde kullanıldığında modelin performansı (yani madde parametre tahminleri ve sınıflama doğruluğu değerleri) ilgili bilgi toplama fırsatı vermektedir (Bradshaw, 2011). Bu koşullara göre oluşturulan yanıt örüntülerinin dikkate alınması noktasında ilgili literatüre ek olarak, veri üretimi ve analizlerine ilişkin süreçler dikkate alınarak harcanacak zaman değerlendirilmiştir. Aşağıda sırasıyla çalışmada kullanılan faktörlere ilişkin koşullar özetlenmiştir.

3.2.1.1. Örneklem Büyüklüğü

Parametre tahminlerini etkileyen önemli faktörlerden biri örneklem büyüklüğüdür. Bu nedenle örneklem büyüklüğü çalışmaya dahil edilmiştir. Örneklem büyüklüğünü belirlemede ilk olarak çalışma kapsamına dahil edilen DINA modeline ilişkin literatürde yer alan çalışmalar incelenmiştir. Odabaş (2016) çalışmasında 500, 1000, 5000 örneklem büyüklüğü; Uyumaz (2016) çalışmasında 250, 500, 1000 örneklem büyüklüğü; Qin vd. (2015) çalışmalarında 500, 1000, 2000, 4000 örneklem büyüklüğü; Ömür-Sünbül ve Kan (2015) çalışmalarında 200, 500, 1000 ve 5000 örneklem büyüklüğü; Ömür-Sünbül (2013) çalışmasında 200, 500, 1000, 5000 örneklem büyüklüğü ve de la Torre vd. (2010) çalışmalarında 1000, 2000, 4000 örneklem büyüklüğüyle çalışmışlardır. Örneklem büyüklüğünün belirlenebilmesi için becerileri ölçen bilişsel tanı modellerine ek olarak daha yeni olan kavram yanılıklarını ölçen modellere ilişkin çalışmalar incelenmiştir.

Literatürde kavram yanılgılarının ölçülmesinde kullanılan modellere ilişkin örneklem büyüklükleri incelendiğinde Bradshaw ve Templin (2014) çalışmalarında kavram yanılgılarının ölçülmesi amacıyla geliştirdikleri SCIM için 3000 ve 10000 örneklem büyüklüğü belirlerken; SISM'in geliştirildiği çalışmada örneklem büyüklüğü 500, 1000 ve 2000 olarak belirlenmiştir (Kuo vd., 2018). Çalışmada kullanılan BUG-DINO modele ilişkin literatürde yer alan araştırmalar incelendiğinde Schellman (2021) çalışmasında BUG-DINO model performansının değerlendirilmesinde 500, 1000 ve 2000 örneklem büyüklüğü; Elbulok (2021) beceri ve kavram yanılgılarının ölçülmesi için geliştirdiği yeni modelin (MC-DINA-H) performansını BUG-DINO ve DINA modelleriyle karşılaştırmada 500, 1000, 2000 ve 5000 örneklem büyüklüğü; Ozaki vd. (2020) BUG-DINO ve geliştirdiği diğer üç modelin performansını değerlendirmede örneklem büyüklüğünü 250, 500 ve 1000 olarak belirlemişlerdir.

Kavram yanılgısı ve becerilerle kavram yanılgılarını ölçen bu daha yeni modellere ilişkin literatürde, uyum indekslerinin bilişsel tanı modellerini veya Q-matrislerini seçmede iyi performans göstermeleri için gerekli olan örneklem büyüklüğü ihtiyacı hakkında az şey bilinse de (Lei & Li, 2016) örneklem büyüklüğünün göreceli olarak artmasıyla birlikte parametrelere ilişkin tahmin hatalarının azaldığı unutulmamalıdır. Bu çalışma kapsamında diğer çalışmalarla karşılaştırmalar sunmak adına daha yeni olan bu modellere ait parametre tahminlerinin ve sınıflama doğruluklarının tutarlılığının daha küçük örneklemelerde nasıl değiştiğine ilişkin incelemeler yapmak amaçlandığından örneklem büyüklüğü faktörünün alt sınırı 250 olarak seçilmiş ve bu faktöre ait diğer koşullar 500, 1500 ve 3000 olarak değişimlenmiştir.

3.2.1.2. Test Uzunluğu

Bilişsel tanı değerlendirmeleri kapsamında test uzunluğuna ilişkin literatür gözden geçirildiğinde beceri modelleri için test uzunluğunu, Odabaş (2016) çalışmasında 20 madde; Uyumaz (2016) çalışmasında 30 madde; Ömür-Sünbül ve Kan (2015) çalışmalarında 15, 30 ve 45 madde; MacDonald (2013) çalışmasında 20, 40 ve 60 madde; Ömür-Sünbül (2013) çalışmasında 20 madde; Templin ve Henson (2006) çalışmalarında 40 madde olarak belirlemişlerdir. Bilişsel tanı modelleri kapsamında kavram yanılgılarına ilişkin literatür gözden geçirildiğinde Schellman (2021) çalışmasında 20 ve 40 madde; Elbulok (2021) çalışmasında 15, 30 ve 60 madde; Ozaki vd. (2020) çalışmalarında 10, 20 ve 30 madde;

Bradshaw ve Templin (2014) çalışmalarında test uzunluğunu 30 ve 60 madde olarak belirlemişlerdir. Son olarak SISM'in model performansını değerlendiren çalışma dikkate alındığında araştırmacıların madde sayısını 20 ve 40 olarak belirledikleri görülmüştür (Kuo vd., 2018). İlgili literatür gözden geçirildiğinde test uzunlukları 15 ile 60 madde arasında değişmektedir.

Baker (1993) çalışmasında test uzunluğunun artmasıyla birlikte madde parametrelerine ilişkin tahmin hatalarının azaldığını ifade etmektedir. Bu sayede Q-matriste yapılabilecek olası hatalı tanımlamalar, hatalı tanımlamaların olmadığı diğer maddeler tarafından telafi edileceğinden bireylere ilişkin elde edilen sınıflama doğruluğu değerleri artacaktır. Bu nedenle çalışmada olası hatalı tanımlamaların önüne geçmek adına test uzunluğu faktöründe uzun test koşulu için 60 madde üst sınır olarak kabul edilmiştir. Sırasıyla beceri, kavram yanılığsı ve hem beceri hem de kavram yanılığsını eş zamanlı olarak ölçen modeller değerlendirildiğinde literatürde yer alan bu sınır değerlerinin sınıf içi ve büyük ölçekli sınavlara ait madde sayılarının etkisi de göz önüne alındığında test uzunluğunun model performansı üzerindeki etkisini incelemek adına kısa ve orta test uzunluğu koşulları için sırasıyla 20 ve 40 madde olacak şekilde toplamda test uzunluğu faktörü için üç koşul değişimlenmiştir.

3.2.1.3. Madde Ayırt Edicilik Düzeyi

Rupp vd. (2010) çalışmalarında ifade ettikleri gibi literatürde yer alan çalışmalar incelendiğinde (Akbay & de la Torre, 2020; Cui vd., 2012; Schellman, 2021) madde ayırt edicilik düzeyini değerlendirmede madde ayırt edicilik indeksinin kullanıldığı görülmektedir. İstatistiksel bilgiler arttıkça bireylere ilişkin elde edilen model veri uyumuyla birlikte tahminlerdeki değişim azalmakta ve daha kesin bilişsel tanılar elde edilmektedir. Madde ayırt edicilik düzeyiyle sınıflama doğruluğu değerlerinin doğrusal bir ilişki göstermesi beklendiğinden ve madde ayırt edicilik düzeyinin sınıflama doğruluğu değerleri üzerinde önemli bir etkisi olduğu göz önüne alındığında bu çalışmada, madde ayırt edicilik düzeyi değişimlenen faktörlerden biri olarak ele alınmıştır. Bilişsel tanı değerlendirmeleri kapsamında faktörlere ilişkin değişimlenen koşullar olarak düşük ve yüksek madde ayırt edicilik düzeyleri dikkate alındığında, yüksek ayırt ediciliğe sahip maddelere $h_j:95$, $\omega_j:15$, $g_j:35$, $\epsilon_j:05$ ve düşük ayırt ediciliğe sahip maddelere $h_j:85$, $\omega_j:20$, $g_j:40$, $\epsilon_j:10$ değerleri atanmıştır. Bu değerlerin seçiminde literatürde yer alan de la Torre ve Huo (2014), de la

Torre ve Lee (2010) ile birlikte Kuo vd. (2018) tarafından gerçekleştirilen çalışmalarla uyum dikkate alınmıştır.

3.2.1.4. Ölçülen Nitelik Sayısı

Sınıflama doğruluğunu etkileyen önemli faktörlerden biri olan beceri sayısının kaç olması gerektiğiyle ilgili bilişsel tanı değerlendirmeleri literatürü incelendiğinde beceri modelleri için bu sayının 3 ile 8 arasında değiştiği görülmektedir (Akbay, Terzi, Kaplan & Karaarslan, 2018; Ardıç, 2020; de la Torre & Douglas, 2004; de la Torre & Lee, 2010; Kalkan, 2016; Ömür-Sünbül, 2013; Ömür-Sünbül & Kan, 2015; Rupp & Templin, 2008; Sen & Cohen, 2021; Tu, Wang, Chai, Douglas & Chang, 2019; Uyumaz, 2016; Yavuz, 2021).

Çalışmada SISM'in performansının değerlendirilerek, beceri ve kavram yanılgılarını ayrı ayrı tanımlayan mevcut bilişsel tanı modellerinin sınıflama performanslarının karşılaştırılması amaçlandığından beceri sayılarına ek olarak literatürde yer alan bilişsel tanı değerlendirmelerine ilişkin kavram yanılgısı modelleri kapsamında kullanılan Q-matrisler ve matrislerde yer alan maddelerin ilişkilendirildiği kavram yanılgısı sayısına ilişkin çalışmalar incelenmiştir. Bunun için ilk olarak SISM'in geliştirildiği mevcut makale ele alınmıştır. Kuo vd. (2018) geliştirdikleri modelin farklı koşullardaki performansını değerlendirmede beceri ve kavram yanılgısı sayısını sabit tutmuşlar ve beceri sayısını 4, kavram yanılgısı sayısını 3 olarak belirlemişlerdir. SISM'in model performansına ilişkin literatürde sınırlı sayıda çalışma (Kuo vd., 2018) yer aldığından kavram yanılgısı sayısına ilişkin karar vermede bilişsel tanı değerlendirmeleri kapsamında kavram yanılgılarının tanımlanmasına yönelik geliştirilmiş modeller ve bu modellerin performanslarının incelenmesinde kullanılan kavram yanılgısı sayılarına ilişkin sınır değerler incelenmiştir. Schellman (2021) çalışmasında yeniden parametrelendirdiği BUG-DINO modelinin performansını farklı koşullar altında değerlendirmede yedi farklı Q-matris tasarımı kullanarak kavram yanılgısı sayısını 2 ve 4 olarak; Norris (2021) çalışmasında Newton yasaları hakkındaki beceri ve kavram yanılgılarını ölçmek için DINA modelini kullanarak bilişsel bir tanı testi geliştirmiş ve bireylerin sahip oldukları kavram yanılgılarını incelemiştir. Bu kapsamda çalışmasında kavram yanılgısı sayısını 6 olarak; Elbulok (2021) çalışmasında beceri ve kavram yanılgılarını eş zamanlı olarak ölçen MC-DINA-H'in performansını DINA ve BUG-DINO modelleriyle karşılaştırırken kavram yanılgısı sayısını 4 olarak; Ozaki vd. (2020) çalışmalarında kavram yanılgılarının belirlenmesinde yeni

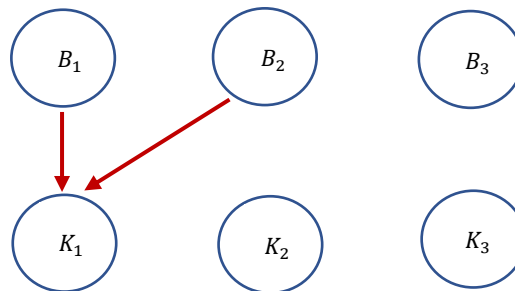
geliştirdikleri iki modelin (MC-DINO1 ve MC-DINO2) performansını BUG-DINO modelinin performansı ile karşılaştırmada kavram yanlışlığı sayısını 5 olarak; Kuo vd. (2016) çalışmalarında hem çoktan seçmeli hem de yapılandırılmış yanıtlı maddeler için geliştirdikleri DINA-CR, DINA-Step, BUG-DINO, BUG-DINO-CR ve BUG-DINO-Step modellerinin performanslarını değerlendirmede kavram yanlışlığı sayısını 4 olarak; DiBello vd. (2015) çoktan seçmeli testlerde seçeneğe dayalı puanlama için geliştirdikleri GDCM-MC'nin performansını değerlendirirken kavram yanlışlığı sayısını 6 olarak; Bradshaw ve Templin (2014) çalışmalarında kavram yanlışlıklarını ölçmek amacıyla geliştirmiş oldukları SCIM'in performansını değerlendirmede kavram yanlışlığı sayısını 3 ve 6 olarak belirlemişlerdir. Genel olarak çalışmalarda kullanım sıklığı değerlendirildiğinde kavram yanlışlıklarının sayısının 3 ile 8 arasında değiştiği görülmektedir. Literatür ile olan uyum da göz önüne alınarak bu çalışmada becerilerin sayısı 3, 4 ve 5; kavram yanlışlıklarının sayısı ise 3, 4 ve 5 olmak üzere değiştirilerek, SISM için çalışmadaki nitelik sayısı;

6 Nitelik = Beceriler (3) + Kavram yanlışlıkları (3),

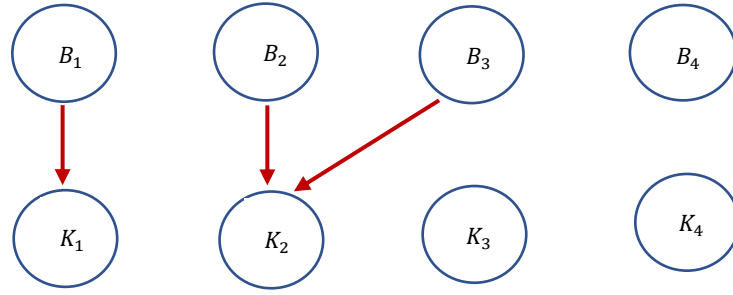
8 Nitelik = Beceriler (4) + Kavram yanlışlıkları (4),

10 Nitelik = Beceriler (5) + Kavram yanlışlıkları (5) olacak şekilde belirlenmiştir.

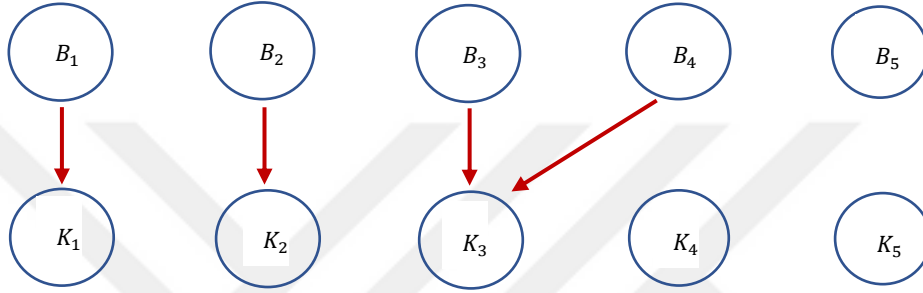
de la Torre ve Douglas (2004) uygulamalarda sayıların küçük görüldüğünün altını çizse de bu çalışmada 6, 8 ve 10 nitelikli Q-matrislerin tercih edilmesiyle modelin uzun çalışma süreleri dikkate alındığında nitelik sayısındaki artışın SISM'in performansına etkisi değerlendirilmek istenmiştir. Çalışmada 6, 8 ve 10 nitelik için beceriler ve kavram yanlışlıkları arasında oluşturulan hiyerarşik yapı Şekil 3, 4 ve 5'te özetlenmiştir. Oluşturulan bu yapı, kavram yanlışlıkları ve becerilerin bir arada bulunduğu durumlar dikkate alındığında daha düşük düzeyli bir kavram yanlışlığının ortaya çıkabilmesi için hangi üst düzey beceriye sahip olunmaması gerektiğini tanımlamaktadır (Elbulok, 2021).



Şekil 3. 3 beceri ve 3 kavram yanlışlığı için hiyerarşik nitelik yapısı



Şekil 4. 4 beceri ve 4 kavram yanılması için hiyerarşik nitelik yapısı



Şekil 5. 5 beceri ve 5 kavram yanılması için hiyerarşik nitelik yapısı

Üç hiyerarşik nitelik yapısından örneğin Şekil 4 incelendiğinde, kavram yanılması K_1 'in B_1 becerisiyle ilişkili olduğu ancak B_2 , B_3 ve B_4 becerileriyle ilişkili olmadığı ayrıca kavram yanılması K_3 ve K_4 'ün de herhangi bir beceriyle ilişkili olmadığı görülmektedir. Aynı durum Şekil 3 ve Şekil 5 için benzer bir mantıkla özetlenebilir. Bu tür ilişkiler değerlendirildiğinde, beceriler ve kavram yanılmalarını kapsayan hiyerarşik yapı ya da yapıların doğru bir şekilde tanımlanabilmesi için olası olmayan bilgi durumlarının modelden çıkarılması gerekmektedir. Olası olmayan bilgi durumları nitelikler (beceriler ve kavram yanılmaları) arası bağımlılıklar (dış bağımlılık) veya niteliklerin kendi içerisindeki (küme içi bağımlılık) bağımlılıklardan kaynaklanabilir.

SISM doğası gereği karmaşık bir yapıya sahip olduğundan bu çalışma kapsamında Elbulok (2021) tarafından yapılan çalışmada olduğu gibi kavram yanılmaları ve beceriler arasında sadece dış bağımlılık ilişkisi kurulmuş olup becerilerin ve kavram yanılmalarının küme içi bağımlılıklarının olmadığı varsayılmıştır. Dış bağımlılık ilişkisinin mantığını özetlemek amacıyla Şekil 4'teki beceriler ve kavram yanılmalarının ele alındığı hiyerarşik ilişkilerin dikkate alınmasıyla oluşturulan alt matris Tablo 5'te sunulmuştur.

Tablo 5

4 Kavram Yanılgısı ve 4 Beceri Arasındaki Hiyerarşik İlişki

	B_1	B_2	B_3	B_4	K_1	K_2	K_3	K_4
B_1	1	0	0	0	1	0	0	0
B_2	0	1	0	0	0	1	0	0
B_3	0	0	1	0	0	1	0	0
B_4	0	0	0	1	0	0	0	0
K_1	0	0	0	0	1	0	0	0
K_2	0	0	0	0	0	1	0	0
K_3	0	0	0	0	0	0	1	0
K_4	0	0	0	0	0	0	0	1

Tablo 5 incelendiğinde gölgeli hücrelerde yer alan “1” değeri, ilgili sütundaki kavram yanılgısının ilgili satırda yer alan becerinin eksikliğine bağlı olduğunu ifade etmektedir. Örneğin birey B_1 becerisine sahip değilse (1), K_1 kavram yanılgısı ortaya çıkabilir (1).

Yukarıda örneği sunulan hiyerarşik ilişkiler dikkate alınarak imkânsız bilgi durumları saptanmış ve bu durumlar tahmin edilen parametrelerdeki azalış dikkate alındığında mevcut çalışma dışında tutulmuştur. Bunun için Şekil 4’te yer alan dış bağımlılık ilişkileri dikkate alınarak Tablo 6 oluşturulmuştur. Bu tablo imkânsız bilgi durumları harici çalışmaya dahil edilen olası bilgi durumlarını özetlemektedir. Yani sahip olunan beceriler örüntüsüyle birlikte meydana gelebilecek tüm olası kavram yanılgısı kombinasyonlarını özetlemektedir. Beceriler ($B = 4$) için tüm bu olası kombinasyonlar ($2^4 = 16$) tabloda “Sınıf” olarak tanımlanmıştır. “0” değeri ilgili sütunda yer alan kavram yanılgısının var olmadığı; “*” simgesi ise ilgili kavram yanılgısının 0/1 olabileceği yani var ya da yok olabileceği şeklinde tanımlanmıştır. Şekil 3 ve Şekil 5 ile ilişkilendirilen tüm olası bilgi durumları EK 1’de sunulmuştur.

Tablo 6

4 Beceri ve 4 Kavram Yanılgısı için Olası Nitelik Sınıfları

Sınıf	Nitelikler								Olası Bilgi Durumları
	B_1	B_2	B_3	B_4	K_1	K_2	K_3	K_4	
1	1	1	1	1	0	0	*	*	$2^2 = 4$
2	1	1	1	0	0	0	*	*	$2^2 = 4$
3	1	1	0	1	0	0	*	*	$2^2 = 4$
4	1	0	1	1	0	0	*	*	$2^2 = 4$
5	0	1	1	1	*	0	*	*	$2^3 = 8$
6	1	1	0	0	0	0	*	*	$2^3 = 8$
7	0	0	1	1	*	0	*	*	$2^3 = 8$
8	1	0	0	1	0	*	*	*	$2^3 = 8$
9	0	1	1	0	*	0	*	*	$2^3 = 8$
10	1	0	1	0	0	0	*	*	$2^3 = 8$
11	0	1	0	1	*	0	*	*	$2^3 = 8$
12	1	0	0	0	0	*	*	*	$2^3 = 8$
13	0	1	0	0	*	0	*	*	$2^3 = 8$
14	0	0	1	0	*	0	*	*	$2^3 = 8$
15	0	0	0	1	*	*	*	*	$2^4 = 16$
16	0	0	0	0	*	*	*	*	$2^4 = 16$
									128

Tabloda yer alan B_1, B_2, B_3, B_4 sırasıyla 1., 2., 3. ve 4. becerileri; K_1, K_2, K_3, K_4 sırasıyla 1., 2., 3. ve 4. kavram yanılgılarını temsil etmektedir.

Kavram yanılgıları ve beceriler arasındaki ilişkilerin dikkate alınmadığı bir durumda Şekil 4 incelendiğinde nitelik sayısı 8 olduğu için olası nitelik örüntülerinin maksimum değeri 2^8 yani 256 olacaktır. Ancak çalışmada kavram yanılgılarıyla beceriler arasında dış bağımlılık ilişkileri oluşturulduğundan bazı nitelik örüntüleri mantıksız kabul edilmiştir. Örneğin B_1 becerisine sahip olan bir bireyde K_1 kavram yanılgısının var olmaması gerekmektedir. Bu nedenle hem B_1 becerisi hem de K_1 kavram yanılgısını içeren bilgi durumları imkânsız hale geleceğinden bu bilgi durumlarının modelden kaldırılması gerekmektedir. Bununla birlikte hem B_2 becerisi ve K_2 kavram yanılgısı, hem de B_3 becerisi ve K_2 kavram yanılgısını içeren olası bilgi durumları benzer bir mantıkla modelden kaldırılmıştır. Tablo 6 incelendiğinde nitelikler arasında oluşturulan hiyerarşik yapıyla olası bilgi durumları azalmış ve 128 olası bilgi durumu elde edilmiştir. Şekil 3 ve Şekil 5 ile ilişkilendirilen hiyerarşik yapılar için sırasıyla 40 ve 388 olası bilgi durumu vardır. Kavram yanılgıları ve beceriler arasında dış bağımlılık ilişkisi bulunduğundan bu ilişkiler modelde bir arada yer alamamaktadır.

Nitelikler arasındaki dış bağımlılıkların belirlenmesiyle oluşturulan Q-matris, sahip olunan beceriler örüntüsüyle birlikte meydana gelebilecek tüm olası kavram yanılgılarının potansiyel bilgi durumlarına ilişkin kombinasyonlarını sunmaktadır. Bu kombinasyonlar dışında kalan imkânsız bilgi durumlarının modelden kaldırılması madde parametrelerine ait tahminlerde azalışlar sağlayacağından (Elbulok, 2021; Tu vd., 2019) nitelikler arasındaki dış bağımlılıkların dikkate alındığı Q-matrislerin kullanılması gerektiği ifade edilmiştir. Bu

nedenle çalışmada SISM için test uzunluğu faktöründe değişimlenen koşullar (20, 40 ve 60 madde) dikkate alınarak Şekil 3, 4 ve 5'teki hiyerarşik yapılarla çoktan seçmeli maddeler için ikili puanlanan (1-0) Q-matrisler geliştirilmiştir. 40 ve 60 maddelik Q-matrisler, 20 maddelik Q-matrisin sırasıyla 2 ve 3 katına çıkarılmasıyla oluşturulmuştur.

Çeşitli araştırmalar kapsamında Q-matrislerin oluşturulmasında Q-matris yapısının karmaşıklığının artması, nitelik sayısındaki artışla ifade edilmektedir. Bunun için her bir niteliği ölçen aynı sayıda maddeyle birlikte birden fazla niteliği ölçen maddelerin de sayılarının eşit olması gerektiği vurgulanmıştır (de la Torre, 2008; de la Torre vd., 2010; Kuo vd., 2018; Yavuz, 2021). Kuo vd. (2018) çalışmalarında SISM için her bir niteliğin eşit sayıda maddeyle (7 madde) ölçüldüğü bir simülasyon çalışması tasarlamışlardır. Bu çalışma kapsamında ise 6, 8 ve 10 nitelikli Q-matrislerde matrislerin temsil etmiş oldukları nitelik sayılarıyla birlikte bu nitelikleri ölçen madde sayıları sabit tutulmamıştır. Çünkü operasyonel testlerin test taslaklarının gereksinimlerini karşılamak için oluşturulduğuna dikkat etmek önemlidir. Bu nedenle maddeler her zaman bu tür düzenli bir örüntüye sahip olmayabilir (Sen & Cohen, 2021).

Bireylerin performansının açıklanmasını kolaylaştırmak adına ayrıntılı bilgiler sağlayan başarılı bir bilişsel tanı uygulaması, test geliştirme aşamaları Q-matris özelliklerinde belirtildiği gibi bilişsel tanı modellerinin özelliklerini taşıyorsa daha başarılı olabilir. Diğer bir deyişle, teoriler veya hipotezler aracılığıyla verinin yapısını belirleyen bilişsel bir modele ihtiyaç vardır ve bu modelin test geliştirilmeden önce belirlenmesi gerekmektedir (Rupp & Templin, 2008). Ancak Gierl, Alves ve Majeau (2010) tarafından ifade edildiği gibi birçok bilişsel tanı modeli uygulamasının simülasyon çalışmaları ya da daha önce geliştirilmiş olan klasik test kuramı veya madde tepki kuramı temelli test verilerine uyarlandığı görülmektedir. Gerek kavram yanılığısı testlerinin bilişsel tanı modelleri kapsamında geliştirilmesinin zorluğu gerek uyarlama çalışmalarının varlığı göz önüne alındığında her niteliğin eşit sayıda maddeyle ölçülmesi ya da birden fazla niteliğin eşit sayıda maddeyle ölçülmesi pek olası görünmemektedir. Bu amaçla Kuo vd. (2018) tarafından geliştirilen SISM'in kusursuz durumlar dışındaki performansının değerlendirilmesinin gerekli olduğu düşünüldüğünden çalışmada nitelik sayısındaki artışla birlikte, nitelikleri ölçen madde sayısı ve birden fazla niteliği ölçen madde sayısı sabit tutulmadan Q-matriste oluşan gerçek karmaşıklığın etkisiyle SISM'in madde parametre tahminleri ve sınıflama doğruluğundaki performansı incelenmiştir.

Bu durumun örnekleri literatürde Sen ve Cohen (2021) ile Ozaki (2015) gibi araştırmacıların çalışmalarında da görülmektedir. Çalışmalarda bazı maddeler az sayıda nitelikle ilişkilendirilirken diğer maddelerin daha çok sayıda niteliklerle ilişkilendirildiği ve birden fazla niteliğin farklı sayıda maddeyle ölçüldüğü Q-matrisler kullanılmıştır.

Şekil 4'teki hiyerarşik yapının dikkate alındığı 8 nitelik için oluşturulan 20 maddelik Q-matris Tablo 7'de aşağıda sunulmuştur. Diğer iki farklı nitelik yapısı için oluşturulan matrislere EK 2'de yer verilmiştir.

Tablo 7

8 Nitelik için 20 Maddelik Q-matris

Maddeler	Nitelikler							
	Beceriler				Kavram Yanılgıları			
	B_1	B_2	B_3	B_4	K_1	K_2	K_3	K_4
1	1	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0
4	0	0	0	1	0	0	0	0
5	0	0	0	0	1	0	0	0
6	0	0	0	0	0	1	0	0
7	0	0	0	0	0	0	1	0
8	0	0	0	0	0	0	0	1
9	1	0	0	0	1	0	0	0
10	0	1	0	0	0	1	0	0
11	0	1	1	0	0	1	0	0
12	1	1	1	0	1	1	0	0
13	1	0	1	1	1	1	1	0
14	0	0	1	0	0	1	0	0
15	0	0	0	1	0	0	1	1
16	0	1	0	1	0	0	1	1
17	1	0	0	1	1	0	1	1
18	1	0	1	0	1	0	1	1
19	0	1	0	0	0	0	1	1
20	0	1	1	1	1	1	0	0

Tabloda yer alan B_1, B_2, B_3, B_4 sırasıyla 1., 2., 3. ve 4. becerileri; K_1, K_2, K_3, K_4 sırasıyla 1., 2., 3. ve 4. kavram yanılgılarını temsil etmektedir.

Tablo 7'de yer alan her bir beceri ve kavram yanılgısının en az 3 maddeyle ölçülmesine dikkat edilmiştir. Bu tabloda 20 maddelik test uzunluğu için B_1, B_2, B_3 ve B_4 sırasıyla 6, 7, 7 ve 6 madde; K_1, K_2, K_3 ve K_4 sırasıyla 7, 7, 7 ve 6 maddeyle ölçülmektedir. Ayrıca 2 nitelik 3 madde; 3 nitelik 3 madde; 4 nitelik 1 madde; 5 nitelik 4 madde ve 6 nitelik 1 maddeyle ölçülmektedir.

Nitelikler arasında kurulan dış bağımlılık ilişkileriyle oluşturulan Q-matrisler, becerilerle ilişkili ya da ilişkili olmayan kavram yanılgılarını özetlemek için genel bir çerçeve sağlamaktadır. Ancak bu matrislerle ilgili dikkat edilmesi gereken nokta içeriklerinde yer

alan becerilerin, kavram yanılgılarıyla ilişkili olup olmamasına göre dört madde parametresinin hepsine sahip olmamasıdır.

8 nitelik içeren 20 maddelik Q-matrisin oluşturulma süreci, örnek olarak aşağıda sunulmuştur.

1. 1-2-3-4. maddeler sadece becerileri ölçmektedir (beceri modeli varken, kavram yanılgısı modeli yoktur). Bu nedenle tüm kavram yanılgıları “0”dır. Burada kavram yanılgılarını ölçmeyen modele SISM yerleştirilmiştir. Modele karşılık gelen parametreler h_j ve g_j parametreleridir.
2. 5-6-7 ve 8. maddeler kavram yanılgılarını ölçen maddelerdir (kavram yanılgısı modeli varken, beceri modeli yoktur). Bu nedenle tüm beceriler “0”dır. Burada becerileri ölçmeyen modele SISM yerleştirilmiştir. Modele karşılık gelen parametreler h_j ve ω_j parametreleridir.
3. 9-10-11-12-13 ve 14. maddeler bağımlı becerileri ve kavram yanılgılarını ölçen maddelerdir. Bu maddelere h_j , g_j ve ϵ_j parametreleriyle sınırlandırılmış SISM yerleştirilmiştir. Yalnızca ölçülen K_3 kavram yanılgısının ölçülen becerilerden bağımsız olduğu maddeler (yani Madde 15, 16, 17, 18 ve 19) ve ölçülen K_2 kavram yanılgısının aynı zamanda ölçülen becerilerden bağımsız olduğu madde 20’ye, SISM yerleştirilmiştir. Bu sayede modelden tüm parametreler elde edilmiştir (h_j , ω_j , g_j ve ϵ_j).

Çalışmada olası bilgi durumlarının kullanıldığı EK 2’de sunulan Q-matrislerle Şekil 3, 4 ve 5’teki hiyerarşik yapılar dikkate alınarak bir simülasyon çalışması gerçekleştirilmiştir.

3.3. Verilerin Analizi

Çalışma kapsamında belirlenen faktörlere ilişkin koşulların analizinde Rstudio 4.2.1 versiyonu kullanılmıştır. Kuo vd. (2018)’nin çalışmalarından elde edilen Github ile birlikte “CDM” paketi (Robitzsch, Kiefer, George, Uenlue & Robitzsch, 2020) 8.2-6 versiyonu ve “G-DINA” paketi (Ma & de la Torre, 2020) 2.9.3 versiyonuyla koşullara ilişkin kodlar yazılarak analizler gerçekleştirilmiş ve SISM’e ilişkin madde parametrelerinin tahmininde “beklenti-maksimizasyon algoritması” kullanılmıştır. Tablo 8, çalışma kapsamında yer alan faktörlere ilişkin koşulları özetlemektedir.

Tablo 8

Simülasyon Çalışması için Değişimlenen Koşullar

Koşullar	Koşulların Sayısı	Koşulların Özellikleri
Örneklem Büyüklüğü	4	<ul style="list-style-type: none"> • 250 • 500 • 1500 • 3000
Test Uzunluğu	3	<ul style="list-style-type: none"> • 20 madde • 40 madde • 60 madde
Madde Ayırt Edicilik Düzeyi	2	<ul style="list-style-type: none"> • Yüksek ayırt ediciliğe sahip maddeler için: $h_j: ,95, \omega_j: ,15, g_j: ,35, \epsilon_j: ,05$ • Düşük ayırt ediciliğe sahip maddeler için: $h_j: ,85, \omega_j: ,20, g_j: ,40, \epsilon_j: ,10$
Ölçülen Nitelik Sayısı	3	<ul style="list-style-type: none"> • 6 (3 beceri + 3 kavram yanılıgısı) • 8 (4beceri + 4 kavram yanılıgısı) • 10 (5 beceri + 5 kavram yanılıgısı)

Olası bilgi durumlarına dayalı olarak test maddelerinde daha gerçekçi yanıtları yansıtmak için yanıt örüntüleri simüle edilmiş ve ana simülasyon tasarımı, yapılacak değerlendirmelerin genelliği için SISM'i dört faktörün (örneklem büyüklüğü, test uzunluğu, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısı) birleştirilmesiyle 72 (4x3x2x3) simülasyon koşulu altında karşılaştırarak her bir koşul için 100 iterasyon gerçekleştirmiştir. Simülasyon koşullarının her biri için başlangıçta SISM kullanılarak 250 kişilik veri seti oluşturulmuş ve her bir grup için doğru yanıt verme olasılıklarını sunan dört madde parametresi farklı koşulların yer aldığı çalışmada sabit tutulmuştur. Çalışmada 250 kişilik veri seti için örnek olarak oluşturulan kod (EK 3) üzerinde dört faktöre ilişkin koşullar değiştirilerek analiz edilmiştir. Olası olmayan her bir bilgi durumuna ait olasılıklar sıfıra sabitlenerek modelden kaldırılmış ve olası bilgi durumlarıyla analizler gerçekleştirilmiştir.

3.3.1. Değerlendirme Kriteri

Simülasyon çalışmalarında madde parametre tahminleri, model ya da modeller tarafından tahmin edilen parametrelere ilişkin değerlerin birey yanıtlarını üretmek için kullanılan olasılık ağırlıklarının ne kadar iyi üretildiğine dair fikirler verdiğinden (Elbulok, 2021) bu çalışma kapsamında SISM'e ait madde parametre tahminlerine ilişkin sonuçların elde edilmesi aşamasında çeşitli değerlendirme kriterleri hesaplanmıştır. Bu amaçla dört farklı madde parametresi için “Mean Absolute Deviation: MAD” ve “Root Mean Square Error: RMSE” olarak literatürde isimlendirilen, çalışmada ise “ortalama mutlak sapma” değerleri ile “hataların ortalama karekökü” değerleri olarak ifade edilen bu değerlendirme kriterlerinin seçilen faktörlere ait her bir koşulda 100 iterasyondan elde edilen tahminleri, aşağıda sunulan eşitlikler yardımıyla hesaplanmıştır.

Ortalama Mutlak Sapma

$$\text{Ortalama Mutlak Sapma}_{h_j} = \frac{1}{J \cdot R} \sum_{j=1}^J \sum_{r=1}^R (h_{jr} - \hat{h}_{jr}) \quad (7)$$

$$\text{Ortalama Mutlak Sapma}_{g_j} = \frac{1}{J \cdot R} \sum_{j=1}^J \sum_{r=1}^R (g_{jr} - \hat{g}_{jr}) \quad (8)$$

$$\text{Ortalama Mutlak Sapma}_{\omega_j} = \frac{1}{J \cdot R} \sum_{j=1}^J \sum_{r=1}^R (\omega_{jr} - \hat{\omega}_{jr}) \quad (9)$$

$$\text{Ortalama Mutlak Sapma}_{\epsilon_j} = \frac{1}{J \cdot R} \sum_{j=1}^J \sum_{r=1}^R (\epsilon_{jr} - \hat{\epsilon}_{jr}) \quad (10)$$

Hataların Ortalama Karekökü

$$\text{Hataların Ortalama Karekökü}_{h_j} = \frac{1}{J} \sum_j \sqrt{\frac{1}{R} \sum_{r=1}^R (h_{jr} - \hat{h}_{jr})^2} \quad (11)$$

$$\text{Hataların Ortalama Karekökü}_{g_j} = \frac{1}{J} \sum_j \sqrt{\frac{1}{R} \sum_{r=1}^R (g_{jr} - \hat{g}_{jr})^2} \quad (12)$$

$$\text{Hataların Ortalama Karekökü}_{\omega_j} = \frac{1}{J} \sum_j \sqrt{\frac{1}{R} \sum_{r=1}^R (\omega_{jr} - \hat{\omega}_{jr})^2} \quad (13)$$

$$\text{Hataların Ortalama Karekökü}_{\epsilon_j} = \frac{1}{J} \sum_j \sqrt{\frac{1}{R} \sum_{r=1}^R (\epsilon_{jr} - \hat{\epsilon}_{jr})^2} \quad (14)$$

Verilen eşitlikler incelendiğinde “ortalama mutlak sapma” ve “hataların ortalama karekökü” değerlerine ilişkin hesaplamalarda;

h_j, g_j, ω_j ve ϵ_j parametreleri r . iterasyondaki j maddesine ait gerçek parametre değerlerine, $\hat{h}_j, \hat{g}_j, \hat{\omega}_j$ ve $\hat{\epsilon}_j$ parametreleri ise r . iterasyondaki j maddesine ait tahmini parametre değerlerine,

$R =$ İterasyon sayısına,

$J =$ Madde sayısına karşılık gelmektedir.

“Ortalama mutlak sapma” ve “hataların ortalama karekökü” değerleri, dört madde parametresi için gerçek parametre değerleriyle tahmini parametre değerleri arasındaki sapmaların ortalama büyüklüğünü ifade etmektedir. Modelin öngördüğü madde ilişkilendirmeleriyle, gözlenen madde ilişkilendirmeleri karşılaştırılırken bu değerler ne kadar küçük ise elde edilen tahminlerin doğruluğu daha yüksek değerler olacaktır (Akbat & de la Torre, 2020; Chen, Xin, Wang & Chang, 2012; Henson, Templin & Willse, 2009). MacDonald (2013) “hataların ortalama karekökü” değerlerinin “0” olduğu durumları gözlenen gizil parametrenin hatasızlığı olarak tanımlamış ancak pratikte hiçbir parametre tahmininin hatasız olmadığını vurgulamıştır.

Çalışma kapsamında “ortalama mutlak sapma” ve “hataların ortalama karekökü” değerleri hesaplandıktan sonra seçilen faktörlerde yer alan her bir değişimlenen koşula ilişkin elde edilen standart hata değerleri dört madde parametresi için hesaplanarak tablolandırılmıştır. Ayrıca çalışmanın amacı doğrultusunda beceriler ve kavram yanılgılarını eş zamanlı olarak ölçen SISM’e ait sınıflama doğruluğu performansı beceri ve kavram yanılgısı modelleriyle karşılaştırılmıştır. Cui vd. (2012) çalışmalarında sınıflama doğruluğunu, bireylerin bir testte yer alan maddelere vermiş oldukları yanıt örüntüleri üzerinden ait oldukları sınıf ile gerçekte buldukları sınıfın birbiriyle uyuma derecesi olarak ifade etmişlerdir.

Sınıflama doğruluklarının incelenmesinde, doymuş model olan SISM ile üretilen veriler kullanılarak bu verilere indirgenmiş modeller olan DINA ve BUG-DINO modellerinin yerleştirilmesiyle Gao, Miller ve Liu (2017) tarafından gerçekleştirilen çalışmadaki gibi yanlış model kullanımının sınıflama doğruluğu performansına etkisi incelenmiştir. Bu kapsamda 6, 8 ve 10 nitelikli Q-matrisler SISM için kullanılırken beceri ve kavram yanılgısı modelleri için bu matrislerden iki alt matris çıkarılmıştır (Kuo vd., 2018). Yani SISM’de kullanılan 6 nitelikli Q-matris örneği için Q-matrisin 3 beceri içeren kısmı DINA için

kullanılırken, 3 kavram yanılığısı içeren kısmı ise BUG-DINO için kullanılmıştır. Q-matrislerin belirlenmesinden sonra tüm durumlar için faktörlerdeki her bir koşulun 100 iterasyonundan elde edilen sınıflama doğruluğu değerleriyle beceri ve kavram yanılığısı modellerine ait sınıflama doğruluğu değerleri karşılaştırılmıştır. Literatürde yer alan çalışmalar dikkate alınarak (Chen vd., 2012; Liang, Lu, Zhang & Shi, 2022; Xi, Tu & Cai, 2022; Zhan, Jiao & Liao, 2018; Zhan, Man & Malone, 2022) farklı modeller arasında sınıflama doğruluğu değerlerinin karşılaştırılabilmesi için “Pattern Correct Classification Rate: PCCR” ve “Attribute Correct Classification Rate: ACCR” olarak isimlendirilen ve çalışmada sırasıyla “örüntü doğru sınıflama oranı” ve “nitelik doğru sınıflama oranı” olarak ifade edilen indeksler aşağıda sunulan eşitlikler yardımıyla hesaplanmıştır.

Örüntü Doğru Sınıflama Oranı

Bilişsel tanı değerlendirmeleri kapsamında sunulan bilişsel tanı profillerine ilişkin tahminlerin doğruluğu için “örüntü doğru sınıflama oranı”, Eşitlik 15 yardımıyla aşağıdaki gibi hesaplanmıştır.

$$\text{Örüntü Doğru Sınıflama Oranı} = \frac{\sum_{i=1}^N R_i}{N} = \frac{\sum_{i=1}^N (I(\hat{\alpha}_i, \alpha_i))}{N} \quad (15)$$

Eşitlikte α_i ve $\hat{\alpha}_i$, sırasıyla i 'ninci bireye ait gerçek ve tahmini bilgi durumlarını göstermektedir. $\hat{\alpha}_i$ 'nin α_i 'ye eşit olduğu durumlarda $R_i = 1$ 'e eşit olur ve model birey ya da bireyleri doğru bir şekilde sınıflandırmaktadır (Chen vd., 2012).

Nitelik Doğru Sınıflama Oranı

Modelde yer alan her bir niteliğin tahmin doğruluğunu gösteren “nitelik doğru sınıflama oranı” değerleri Eşitlik 16 yardımıyla aşağıdaki gibi hesaplanmıştır.

$$\text{Nitelik Doğru Sınıflama Oranı}_k = \frac{\sum_{i=1}^N A_{ik}}{N} = \frac{\sum_{i=1}^N (I(\hat{\alpha}_{ik}, \alpha_{ik}))}{N} \quad (16)$$

Eşitlikte $\hat{\alpha}_{ik}$ ve α_{ik} eşit olduğu durumlarda $A_{ik} = 1$ 'e eşittir ve birey ya da bireyler k . nitelikte doğru bir şekilde sınıflandırılır. Her ne kadar bu iki indekse ait kesin sınır değerleri bulunmasa da indekslere ait değerler ne kadar büyükse, tüm modelin ve her bir niteliğin doğru sınıflama oranları o kadar yüksektir (Chen vd., 2012). Bu kapsamda çalışmada farklı koşullardan elde edilen sınıflama doğruluğu değerleri göreceli olarak karşılaştırılarak değerlendirilmiştir.

BÖLÜM IV

BULGULAR VE YORUM

Bu bölümde çalışmada ifade edilen “Bilişsel tanı değerlendirmesi kapsamında kullanılan modellerden SISM’in, dört faktöre ait değişimlenen koşullarda madde parametrelerine ilişkin elde edilen ortalama mutlak sapma ve hata değerleriyle birlikte bu modelin DINA ve BUG-DINO modelleriyle karşılaştırıldığı durumlardan elde edilen sınıflama doğruluğu değerleri nasıl değişmektedir?” problem cümlesine yönelik oluşturulan her bir alt problem ve bu alt problemlere ilişkin bulgular ve yorumlara yer verilmiştir.

4.1. Birinci Alt Probleme İlişkin Bulgular

1. SISM’e göre madde parametrelerine ait “ortalama mutlak sapma” değerleri,
 - 1.1. Örneklem büyüklüğü (250, 500, 1500 ve 3000),
 - 1.2. Test uzunluğu (20, 40 ve 60 madde),
 - 1.3. Madde ayırt edicilik düzeyi (yüksek ayırt edicilik ve düşük ayırt edicilik) ve
 - 1.4. Ölçülen nitelik sayısına (6, 8 ve 10 nitelik) göre nasıl değişmektedir?

Birinci alt problem kapsamında SISM’den elde edilen madde parametreleri arasındaki “mutlak sapma” ve bunların ortalamasını sunan “ortalama mutlak sapma” değerleri, çalışmanın amacı doğrultusunda örneklem büyüklüğü, test uzunluğu, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısı olmak üzere dört faktörün her bir koşulu için 100 iterasyonun analizinden elde edilmiş ve bulgular tablolarla özetlenmiştir. 8 ve 10 nitelikli Q-matrisler için “mutlak sapma” ve “ortalama mutlak sapma” değerlerine ilişkin elde edilen tablolara EK 4’te yer verilmiştir. 6 nitelikli Q-matristen elde edilen bulgulara benzer olacak şekilde 8 ve 10 nitelikli Q-matrisler özetlenebilir. Bu bölümde okuyucuya kolaylık sağlaması

açısından sadece 6 nitelikli Q-matris'ten elde edilen tablolara yer verilmiştir. 6, 8 ve 10 nitelikli Q-matrisler kullanıldığında bu tablolardan faktörlerde değişimlenen her bir koşula ait “ortalama mutlak sapma” değerleri alınmış ve grafiklerle sunularak karşılaştırılmıştır.

SISM'de yer alan ilgili eşitliklerde, madde parametrelerine ait değerler 0 veya 1'e eşitlendiğinde beceriler ve kavram yanılgılarını eş zamanlı olarak tanımlayamayan yani sadece becerileri ya da kavram yanılgılarını ölçen DINA ve BUG-DINO modellerinin elde edilebildiği ve buna bağlı olarak modelde dört farklı başarı olasılığını yansıtan madde parametrelerinin hepsinin hesaplanamadığı ifade edilmiştir (Kuo vd. 2018). Tablolar incelendiğinde örneğin DINA model sadece becerileri ölçebildiği için modele karşılık gelen madde parametreleri h_j ve g_j parametreleridir. Bu parametreler DINA model için iki farklı başarı olasılığının hesaplandığının bir göstergesidir. h_j madde parametresi ölçülen tüm becerilere sahip olan ve kavram yanılgıları olmayan bireylerin başarı olasılığını yansıtırken, g_j madde parametresi ölçülen tüm becerilere sahip olmayan ve kavram yanılgıları olmayan bireylerin başarı olasılığını yansıtmaktadır. DINA modelin her iki madde parametresi de kavram yanılgılarının var olmadığını varsaymaktadır. Bu sebeple maddelerde kavram yanılgılarının varlığını ölçen ve sırasıyla “Ölçülen tüm becerilere sahip olan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı (ω_j)” ve “ölçülen tüm becerilere sahip olmayan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı (ε_j)” olarak isimlendirilen iki madde parametresine ilişkin değerler DINA modelin doğası gereği hesaplanamamaktadır.

Tablo 9 ve Tablo 10, 6 niteliği ölçen sırasıyla madde ayırt edicilik düzeyinin yüksek ve düşük olduğu 20 maddelik bir test için farklı örneklem büyüklüğü koşullarında gerçek ve tahmini madde parametreleri arasındaki “mutlak sapma” ve “ortalama mutlak sapma” değerlerini sunmaktadır.

Tablo 9

6 Niteliği Ölçen Yüksek Ayırt Ediciliğe Sahip 20 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,15	,18	,20	,20	,09	,07	,04	,04	-	-	-	-	-	-	-	-
2	,06	,06	,05	,05	,10	,08	,06	,06	-	-	-	-	-	-	-	-
3	,03	,02	,02	,02	,05	,03	,02	,02	-	-	-	-	-	-	-	-
4	,03	,03	,03	,03	-	-	-	-	,03	,03	,01	,01	-	-	-	-
5	,03	,02	,01	,01	-	-	-	-	,05	,03	,02	,02	-	-	-	-
6	,03	,02	,02	,02	-	-	-	-	,06	,05	,04	,04	-	-	-	-
7	,05	,04	,03	,03	,13	,13	,13	,13	-	-	-	-	,12	,18	,15	,15
8	,06	,05	,04	,04	,09	,08	,06	,06	-	-	-	-	,04	,03	,01	,01
9	,33	,25	,20	,20	,12	,12	,12	,12	-	-	-	-	,04	,03	,02	,02
10	,45	,45	,46	,46	,12	,12	,12	,12	-	-	-	-	,04	,03	,02	,02
11	,07	,05	,03	,03	,10	,10	,10	,10	-	-	-	-	,20	,20	,20	,20
12	,07	,06	,03	,03	,10	,10	,10	,10	-	-	-	-	,20	,20	,20	,20
13	,34	,31	,29	,29	,14	,14	,14	,14	-	-	-	-	,04	,04	,02	,02
14	,42	,41	,41	,41	,09	,09	,09	,09	,12	,11	,11	,11	,06	,07	,08	,08
15	,10	,11	,13	,13	,12	,12	,07	,07	,05	,03	,03	,03	,06	,04	,02	,02
16	,09	,07	,04	,04	,11	,11	,10	,10	,05	,05	,05	,05	,05	,20	,19	,20
17	,07	,06	,04	,04	,09	,08	,06	,06	,11	,11	,11	,11	,09	,09	,09	,09
18	,05	,03	,02	,02	,11	,07	,04	,04	,05	,04	,03	,03	,07	,07	,07	,07
19	,08	,07	,04	,04	,12	,10	,10	,10	,05	,04	,02	,02	,19	,20	,20	,20
20	,04	,03	,02	,02	,07	,04	,03	,03	,05	,03	,02	,02	,04	,03	,01	,01
Ortalama Mutlak Sapma	,13	,12	,11	,10	,10	,09	,08	,08	,06	,05	,04	,04	,10	,10	,09	,09

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı.

Tablo 9’da yer alan dört faktöre ait 6 nitelik, 20 madde ve yüksek ayırt edicilik koşulları dikkate alındığında farklı örneklem büyüklüğü koşulları için “mutlak sapma” değerlerinin dört farklı madde parametresinde sırasıyla ,01 ile ,46 (h_j); ,02 ile ,14 (g_j); ,01 ile ,12 (ω_j) ve ,01 ile ,20 (ϵ_j) arasında değiştiği görülmektedir. Ölçülen tüm becerilere sahip olan ve kavram yanılgıları olmayan bireylerin başarı olasılığını yansıtan h_j madde parametresinden elde edilen “mutlak sapma” değerleri, diğer üç madde parametresine göre daha yüksektir. Bu parametreye ilişkin gerçek ve tahmini madde parametreleri arasındaki en büyük sapmaya sahip maddelerin 9., 10., 13. ve 14. maddeler olduğu görülmektedir. “Mutlak sapma” değerlerine yönelik farklı örneklem büyüklüğü koşulları genel olarak değerlendirildiğinde h_j madde parametresine ait “mutlak sapma” değerlerinin, örneklem büyüklüğünün 250 olduğu koşulda daha yüksek olduğu ve örneklem büyüklüğü arttıkça monoton olmayan artış ya da azalışlar olsa da bu değerlerin genel olarak azaldığı görülmektedir. Örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda ise “mutlak sapma” değerlerinin birkaç madde dışında (1., 10. ve 15. maddeler) genel olarak birbirine benzer olduğu görülmektedir.

Tabloda yer alan g_j , ω_j ve ϵ_j madde parametreleri incelendiğinde g_j madde parametresinin örneklem büyüklüğü artışıyla birlikte “mutlak sapma” değerlerinde ,01 ile ,07 arasında değişen azalışlar olduğu görülmektedir. Bazı maddelerin (7., 9., 10., 11., 12., 13. ve 14. maddeler) “mutlak sapma” değerleri her bir örneklem büyüklüğü koşulu için aynıdır. ω_j ve ϵ_j madde parametrelerine ilişkin “mutlak sapma” değerleri incelendiğinde örneklem büyüklüğünün artmasıyla birlikte ,01 ile ,06 arasında değişen artış ya da azalışlar olduğu görülmektedir. Bu iki madde parametresi için bazı maddeler hariç (7., 14., 16. ve 19. maddeler) örneklem büyüklüğü arttıkça “mutlak sapma” değerlerinin azaldığı, 1500 ve 3000 örneklem büyüklüğü koşullarında “mutlak sapma” değerlerinin değişmediği görülmektedir. Tabloda yer alan “ortalama mutlak sapma” değerleri incelendiğinde, farklı örneklem büyüklüklerinde bu değerlerin dört madde parametresi için ,04 ile ,13 arasında değiştiği ve örneklem büyüklüğü artışıyla birlikte elde edilen “ortalama mutlak sapma” değerlerinin azaldığı görülmektedir. Bu iyileşme h_j madde parametresi için ,01’lik bir azalışa karşılık gelmektedir. ϵ_j parametresi hariç g_j ve ω_j parametrelerinde 250 ve 500 örneklem büyüklüğü koşullarında madde parametrelerine ait değerler ,01’lik bir azalışa karşılık gelirken; örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda g_j , ω_j ve ϵ_j parametrelerine ait “ortalama mutlak sapma” değerlerinin değişmediği görülmektedir.

Tablo 10

6 Niteliği Ölçen Düşük Ayırt Ediciliğe Sahip 20 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,10	,13	,15	,15	,12	,10	,07	,03	-	-	-	-	-	-	-	-
2	,10	,08	,09	,10	,17	,13	,14	,16	-	-	-	-	-	-	-	-
3	,08	,04	,02	,02	,09	,06	,03	,02	-	-	-	-	-	-	-	-
4	,04	,03	,03	,03	-	-	-	-	,04	,04	,02	,02	-	-	-	-
5	,06	,03	,02	,01	-	-	-	-	,07	,05	,03	,02	-	-	-	-
6	,06	,04	,02	,02	-	-	-	-	,08	,05	,04	,03	-	-	-	-
7	,09	,06	,04	,03	,13	,17	,14	,14	-	-	-	-	,13	,16	,14	,12
8	,10	,09	,08	,10	,16	,13	,15	,15	-	-	-	-	,13	,07	,08	,07
9	,44	,28	,16	,13	,22	,22	,20	,19	-	-	-	-	,07	,05	,03	,02
10	,35	,32	,30	,30	,15	,15	,14	,14	-	-	-	-	,09	,07	,04	,03
11	,14	,12	,07	,06	,11	,11	,13	,13	-	-	-	-	,19	,19	,17	,17
12	,15	,13	,08	,05	,10	,11	,13	,13	-	-	-	-	,20	,19	,17	,17
13	,35	,24	,21	,21	,15	,17	,16	,16	-	-	-	-	,08	,07	,04	,03
14	,39	,31	,32	,33	,10	,10	,09	,09	,13	,11	,11	,11	,09	,08	,08	,07
15	,13	,11	,11	,10	,16	,14	,10	,06	,09	,04	,03	,03	,10	,07	,06	,03
16	,18	,14	,09	,07	,13	,12	,11	,10	,07	,06	,06	,06	,06	,25	,19	,19
17	,10	,10	,09	,10	,18	,15	,12	,12	,13	,13	,13	,13	,10	,10	,08	,08
18	,10	,07	,04	,04	,21	,11	,06	,04	,08	,05	,03	,03	,09	,07	,07	,07
19	,19	,13	,13	,12	,15	,14	,11	,11	,10	,06	,06	,06	,21	,17	,19	,19
20	,08	,06	,03	,02	,13	,07	,04	,03	,08	,05	,03	,02	,08	,05	,03	,02
Ortalama Mutlak Sapma	,16	,13	,11	,10	,15	,13	,11	,11	,09	,06	,05	,05	,13	,11	,10	,09

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı.

Tablo 10’da yer alan faktörlere ait 6 nitelik, 20 madde ve düşük ayırt edicilik koşulları dikkate alındığında farklı örneklem büyüklüğü koşulları için “mutlak sapma” değerlerinin dört farklı madde parametresinde sırasıyla ,01 ile ,44 (h_j); ,02 ile ,22 (g_j); ,02 ile ,13 (ω_j) ve ,02 ile ,25 (ϵ_j) arasında değiştiği görülmektedir. h_j madde parametresinden elde edilen “mutlak sapma” değerleri diğer üç madde parametresine göre daha yüksektir. Bu parametreye ilişkin gerçek ve tahmini madde parametreleri arasındaki en büyük sapmaya sahip maddelerin 9., 10., 13. ve 14. maddeler olduğu görülmektedir. “Mutlak sapma” değerlerine yönelik farklı örneklem büyüklüğü koşulları genel olarak değerlendirildiğinde h_j madde parametresine ait bu değerler, örneklem büyüklüğünün 250 olduğu koşulda daha yüksektir ve örneklem büyüklüğü artışıyla birlikte bazı maddeler hariç (1., 2., 8. ve 14. maddeler) ,01 ile ,31 arasında değişen bir azalış sergilemektedir. Tabloda yer alan g_j , ω_j ve ϵ_j madde parametreleri incelendiğinde g_j madde parametresinin örneklem büyüklüğü artışıyla birlikte “mutlak sapma” değerlerinde ,01 ile ,17 arasında değişen azalışlar olduğu görülmektedir. ω_j madde parametresine ilişkin “mutlak sapma” değerleri incelendiğinde

örneklem büyüklüğünün artmasıyla birlikte ,01 ile ,06 arasında azalışlar olduğu; ϵ_j madde parametresi için bazı maddeler hariç (7., 8. ve 16. maddeler) örneklem büyüklüğü arttıkça bu değerlerin azaldığı görülmektedir. “Ortalama mutlak sapma” değerleri incelendiğinde, farklı örneklem büyüklüklerinde bu değerlerin dört madde parametresi için ,05 ile ,16 arasında değiştiği ve örneklem büyüklüğü artışıyla birlikte dört madde parametresi için elde edilen “ortalama mutlak sapma” değerlerinin azaldığı görülmektedir. Ancak örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda g_j ve ω_j madde parametrelerine ait değerler değişmemiştir.

Aşağıda sunulan Tablo 11 ve Tablo 12, 6 niteliği ölçen madde ayırt edicilik düzeyinin yüksek ve düşük olduğu 40 maddelik bir test için farklı örneklem büyüklüğü koşullarında gerçek ve tahmini madde parametreleri arasındaki “mutlak sapma” ve “ortalama mutlak sapma” değerlerini sunmaktadır.

Tablo 11

6 Niteliği Ölçen Yüksek Ayırt Ediciliğe Sahip 40 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,13	,16	,16	,17	,06	,05	,04	,04	-	-	-	-	-	-	-	-
2	,03	,02	,01	,01	,04	,03	,02	,01	-	-	-	-	-	-	-	-
3	,03	,02	,01	,01	,04	,03	,01	,01	-	-	-	-	-	-	-	-
4	,03	,02	,02	,02	-	-	-	-	,03	,02	,01	,01	-	-	-	-
5	,02	,01	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
6	,02	,01	,01	,01	-	-	-	-	,04	,03	,02	,02	-	-	-	-
7	,03	,02	,02	,02	,07	,06	,05	,06	-	-	-	-	,08	,07	,03	,03
8	,03	,02	,01	,01	,06	,04	,03	,03	-	-	-	-	,03	,02	,01	,01
9	,32	,25	,24	,23	,11	,12	,12	,12	,12	-	-	-	-	,04	,03	,01
10	,48	,49	,49	,49	,11	,12	,12	,12	-	-	-	-	,03	,02	,01	,01
11	,05	,03	,02	,01	,10	,09	,09	,09	-	-	-	-	,20	,21	,21	,21
12	,05	,04	,02	,01	,09	,09	,09	,09	-	-	-	-	,21	,21	,21	,21
13	,37	,31	,30	,28	,12	,14	,12	,12	-	-	-	-	,04	,03	,02	,01
14	,43	,40	,40	,40	,08	,09	,08	,09	,12	,11	,12	,11	,07	,07	,07	,08
15	,09	,10	,12	,12	,08	,06	,05	,05	,04	,04	,03	,03	,04	,03	,02	,01
16	,05	,04	,02	,01	,11	,11	,10	,11	,06	,05	,05	,05	,20	,20	,20	,19
17	,03	,02	,01	,01	,06	,04	,02	,02	,11	,10	,10	,10	,10	,09	,10	,10
18	,04	,02	,01	,01	,08	,05	,03	,02	,05	,04	,03	,03	,03	,07	,07	,07
19	,06	,04	,02	,02	,09	,10	,10	,10	,09	,04	,03	,02	,01	,22	,20	,20
20	,03	,02	,01	,01	,05	,04	,02	,02	,05	,03	,02	,01	,03	,02	,01	,01
21	,13	,16	,17	,17	,07	,05	,04	,03	-	-	-	-	-	-	-	-
22	,03	,02	,01	,01	,04	,03	,02	,01	-	-	-	-	-	-	-	-
23	,03	,01	,01	,01	,04	,03	,02	,01	-	-	-	-	-	-	-	-
24	,03	,02	,02	,02	-	-	-	-	,03	,02	,01	,01	-	-	-	-
25	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
26	,02	,01	,01	,01	-	-	-	-	,03	,03	,02	,02	-	-	-	-
27	,04	,03	,02	,02	,07	,06	,06	,06	-	-	-	-	,07	,06	,04	,03
28	,03	,02	,01	,01	,06	,04	,03	,03	-	-	-	-	,03	,02	,01	,01
29	,33	,25	,24	,23	,11	,12	,12	,12	-	-	-	-	,04	,02	,01	,01
30	,48	,49	,49	,49	,12	,12	,12	,12	-	-	-	-	,03	,03	,02	,01
31	,06	,03	,02	,01	,10	,09	,09	,10	-	-	-	-	,20	,21	,21	,20
32	,05	,03	,02	,01	,10	,09	,09	,10	-	-	-	-	,21	,21	,21	,20
33	,36	,31	,29	,29	,12	,12	,13	,13	-	-	-	-	,04	,03	,02	,02
34	,43	,41	,40	,40	,09	,08	,09	,09	,12	,12	,11	,11	,07	,07	,07	,08
35	,09	,11	,12	,13	,08	,06	,05	,05	,04	,03	,03	,03	,04	,03	,02	,01
36	,06	,04	,02	,02	,11	,11	,10	,11	,05	,05	,05	,05	,19	,19	,20	,19
37	,03	,02	,01	,01	,06	,04	,02	,02	,10	,10	,10	,10	,09	,10	,09	,10
38	,04	,02	,01	,01	,07	,05	,03	,02	,04	,04	,03	,03	,07	,07	,07	,07
39	,05	,03	,02	,01	,11	,10	,10	,10	,05	,03	,02	,01	,20	,21	,20	,20
40	,03	,02	,01	,01	,05	,04	,02	,02	,05	,03	,02	,01	,03	,02	,01	,01
Ortalama Mutlak Sapma	,12	,10	,10	,09	,08	,07	,07	,06	,05	,05	,04	,04	,10	,09	,08	,08

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.

Tablo 11’de yer alan faktörlere ait 6 nitelik, 40 madde ve yüksek ayırt edicilik koşulları dikkate alındığında farklı örneklem büyüklüğü koşulları için “mutlak sapma” değerlerinin dört farklı madde parametresinde sırasıyla ,00 ile ,49 (h_j); ,01 ile ,14 (g_j); ,01 ile ,12 (ω_j) ve ,01 ile ,22 (ϵ_j) arasında değiştiği görülmektedir. h_j madde parametresinde sapmalar diğer üç madde parametresine göre daha yüksektir. Bu parametreye ilişkin gerçek ve tahmini madde

parametreleri arasındaki en büyük sapmaya sahip maddelerin 9., 10., 13., 14., 29., 30., 33. ve 34. maddeler olduğu görülmektedir. “Mutlak sapma” değerlerine yönelik farklı örneklem büyüklüğü koşulları genel olarak değerlendirildiğinde h_j madde parametresine ait “mutlak sapma” değerlerinin örneklem büyüklüğünün 250 olduğu koşulda daha büyük olduğu ve örneklem büyüklüğü arttıkça “mutlak sapma” değerlerinin genel olarak azaldığı görülmektedir. Örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda ise bu değerlerin genel olarak birbirine benzer olduğu ancak bazı maddelerde “mutlak sapma” değerlerinin her bir örneklem büyüklüğü koşulunda örneklem büyüklükleri arttıkça ,01 ile ,02 arasında değişen artış ve azalışlara sahip bulgusuna ulaşılmıştır.

Tabloda yer alan g_j , ω_j ve ϵ_j madde parametreleri incelendiğinde g_j parametresinin örneklem büyüklüğü artışıyla birlikte ,01 ile ,06 arasında “mutlak sapma” değerlerinde artış ve azalışların olduğu görülmektedir. Madde 12 ve madde 30’a ait bu değerler her bir örneklem büyüklüğü koşulu için aynıdır. ω_j madde parametresine ilişkin “mutlak sapma” değerleri incelendiğinde 36. ve 37. maddelerde bu değerlerin değişmediği, ϵ_j madde parametresinde ise 12., 18. ve 38. maddelerde bu değerlerin değişmediği görülmektedir. Bu maddeler haricinde örneklem büyüklüğünün artmasıyla birlikte her iki madde parametresine ilişkin “mutlak sapma” değerlerinde ,01 ile ,05 arasında değişen artış ya da azalışlar olduğu görülmektedir. Bu iki madde parametresinden ω_j için madde 14; ϵ_j için bazı maddeler (11., 17., 19., 31., 34., 36., 37. ve 39. maddeler) ,01 ile ,02 arasında değişen küçük sapmalara sahiptir. Genel olarak değerlendirildiğinde örneklem büyüklüğü arttıkça “mutlak sapma” değerlerinin azaldığı, 1500 ve 3000 örneklem büyüklüğü koşullarında bu değerlerin bazı maddeler haricinde genel olarak değişmediği görülmektedir.

“Ortalama mutlak sapma” değerleri incelendiğinde, farklı örneklem büyüklüklerinde bu değerlerin dört madde parametresi için ,04 ile ,12 arasında değiştiği ve örneklem büyüklüğü artışıyla birlikte azaldığı görülmektedir. Sadece ω_j ve ϵ_j madde parametrelerinde örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda bu parametrelere ait “ortalama mutlak sapma” değerleri aynıdır.

Tablo 12

6 Niteliği Ölçen Düşük Ayırt Ediciliğe Sahip 40 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,07	,07	,05	,04	,08	,08	,08	,08	-	-	-	-	-	-	-	-
2	,09	,11	,13	,13	,08	,08	,09	,09	-	-	-	-	-	-	-	-
3	,13	,14	,14	,13	,14	,14	,12	,10	-	-	-	-	-	-	-	-
4	,04	,04	,04	,04	-	-	-	-	,03	,02	,02	,01	-	-	-	-
5	,04	,03	,01	,01	-	-	-	-	,05	,03	,02	,02	-	-	-	-
6	,03	,02	,01	,01	-	-	-	-	,04	,03	,02	,01	-	-	-	-
7	,05	,05	,03	,03	,08	,06	,04	,03	-	-	-	-	,08	,06	,05	,04
8	,06	,05	,03	,02	,08	,09	,06	,05	-	-	-	-	,06	,07	,08	,08
9	,46	,53	,57	,59	,22	,21	,21	,21	-	-	-	-	,06	,04	,03	,02
10	,36	,28	,14	,12	,11	,09	,07	,07	-	-	-	-	,07	,05	,03	,03
11	,12	,14	,12	,10	,10	,10	,10	,10	-	-	-	-	,20	,20	,20	,20
12	,12	,13	,11	,10	,10	,10	,10	,10	-	-	-	-	,20	,20	,20	,20
13	,52	,58	,63	,64	,12	,09	,08	,08	-	-	-	-	,06	,05	,04	,04
14	,50	,53	,58	,58	,09	,09	,07	,07	,12	,11	,13	,13	,08	,06	,05	,03
15	,07	,06	,04	,03	,09	,08	,08	,08	,06	,04	,02	,02	,06	,04	,02	,02
16	,17	,19	,17	,15	,10	,10	,07	,07	,10	,10	,10	,10	,24	,26	,25	,24
17	,08	,09	,10	,10	,10	,09	,08	,08	,10	,09	,07	,06	,09	,08	,07	,07
18	,10	,11	,08	,06	,16	,16	,09	,08	,07	,07	,06	,04	,07	,08	,07	,07
19	,16	,19	,18	,16	,14	,12	,12	,12	,07	,06	,05	,05	,21	,22	,20	,18
20	,11	,13	,11	,09	,14	,15	,11	,09	,10	,08	,07	,05	,06	,05	,04	,04
21	,07	,07	,05	,04	,08	,08	,08	,08	-	-	-	-	-	-	-	-
22	,10	,11	,13	,13	,08	,09	,09	,09	-	-	-	-	-	-	-	-
23	,14	,15	,14	,13	,14	,14	,12	,10	-	-	-	-	-	-	-	-
24	,04	,04	,04	,04	-	-	-	-	,03	,02	,02	,01	-	-	-	-
25	,03	,02	,01	,01	-	-	-	-	,04	,03	,03	,02	-	-	-	-
26	,03	,02	,01	,01	-	-	-	-	,04	,03	,02	,01	-	-	-	-
27	,07	,04	,03	,03	,07	,07	,04	,03	-	-	-	-	,08	,06	,06	,04
28	,07	,05	,02	,02	,08	,08	,06	,05	-	-	-	-	,07	,07	,08	,08
29	,47	,52	,57	,59	,22	,21	,21	,22	-	-	-	-	,06	,04	,03	,02
30	,36	,27	,14	,12	,11	,10	,07	,07	-	-	-	-	,07	,05	,03	,03
31	,11	,14	,12	,10	,10	,10	,10	,10	-	-	-	-	,20	,20	,20	,20
32	,10	,13	,11	,10	,10	,10	,10	,11	-	-	-	-	,20	,20	,20	,19
33	,51	,58	,63	,64	,12	,10	,08	,08	-	-	-	-	,06	,05	,05	,04
34	,49	,54	,57	,58	,09	,09	,07	,07	,11	,12	,13	,13	,08	,07	,05	,03
35	,08	,06	,04	,03	,11	,08	,08	,08	,06	,04	,02	,02	,06	,04	,03	,02
36	,17	,20	,17	,15	,10	,09	,07	,07	,09	,10	,10	,10	,24	,26	,25	,24
37	,09	,09	,10	,10	,11	,10	,09	,08	,10	,08	,07	,06	,09	,08	,08	,08
38	,10	,11	,08	,06	,16	,17	,10	,06	,07	,08	,05	,05	,08	,08	,07	,07
39	,16	,20	,19	,16	,13	,12	,11	,12	,06	,05	,05	,05	,23	,23	,20	,18
40	,10	,12	,11	,10	,14	,14	,10	,09	,08	,09	,06	,04	,27	,05	,04	,04
Ortalama Mutlak Sapma	,16	,17	,16	,16	,11	,11	,09	,09	,07	,06	,06	,05	,11	,10	,10	,09

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.

Tablo 12’de yer alan faktörlere ait 6 nitelik, 40 madde ve düşük ayırt edicilik koşulları dikkate alındığında farklı örneklem büyüklüğü koşulları için “mutlak sapma” değerlerinin dört farklı madde parametresinde sırasıyla ,01 ile ,64 (h_j); ,03 ile ,22 (g_j); ,01 ile ,13 (ω_j) ve ,02 ile ,27 (ϵ_j) arasında değiştiği görülmektedir. h_j madde parametresinde sapmalar, diğer üç madde parametresine göre daha yüksektir. Bu parametreye ilişkin gerçek ve tahmini madde

parametreleri arasındaki en büyük sapmaya sahip maddelerin 9., 10., 13., 14., 29., 30., 33. ve 34. maddeler olduğu görülmektedir. “Mutlak sapma” değerlerine yönelik farklı örneklem büyüklüğü koşulları genel olarak değerlendirildiğinde, h_j madde parametresine ait bu değerlerin örneklem büyüklüğünün 250 olduğu koşulda daha büyük olduğu ve örneklem büyüklüğü artışıyla birlikte “mutlak sapma” değerlerinin bazı maddeler haricinde (madde 4 ve madde 24) ,01 ile ,24 arasında bir artış ya da azalış sergilediği görülmektedir. Tabloda yer alan g_j , ω_j ve ϵ_j madde parametrelerinden g_j madde parametresi dikkate alındığında bazı maddelerde (8., 29. ve 39. maddeler gibi) kendi içinde örneklem büyüklüğü artışıyla birlikte artış ve azalışlar olsa da madde 2, madde 22 ve madde 32 hariç diğer maddelerin “mutlak sapma” değerlerinde ,01 ile ,10 arasında değişen azalışların olduğu görülmektedir. Ayrıca bazı maddelere (1., 11., 12., 21. ve 31. maddeler) ait “mutlak sapma” değerleri örneklem büyüklüğü artsa da aynı değerleri almıştır. ω_j madde parametresine ilişkin değerler incelendiğinde bazı maddeler (14., 38. ve 40. maddeler gibi) kendi içinde örneklem büyüklüğü artışıyla birlikte artış ve azalışlar sergilese de diğer maddelerin “mutlak sapma” değerlerinde örneklem büyüklüğü artışıyla birlikte ,01 ile ,05 arasında değişen artış ya da azalışların olduğu görülmektedir. Ayrıca madde 16’ya ait “mutlak sapma” değerleri örneklem büyüklüğü artsa da aynı değerleri almıştır. ϵ_j madde parametresine ilişkin değerler incelendiğinde bazı maddeler kendi içinde (16., 18. ve 19. maddeler gibi) örneklem büyüklüğü artışıyla birlikte artış ve azalışlar sergilese de diğer maddelerin “mutlak sapma” değerlerinde örneklem büyüklüğü artışıyla birlikte ,01 ile ,23 arasında değişen artış ya da azalışların olduğu görülmektedir. Ayrıca 11., 12., 31. ve 32. maddeler örneklem büyüklüğü artışıyla birlikte aynı “mutlak sapma” değerlerini almıştır.

“Ortalama mutlak sapma” değerleri incelendiğinde, farklı örneklem büyüklüklerinde bu değerlerin dört madde parametresi için ,05 ile ,17 arasında değiştiği ve örneklem büyüklüğü artışıyla birlikte ω_j ve ϵ_j madde parametreleri için elde edilen değerlerin azaldığı görülmektedir. h_j ve g_j madde parametreleri ise örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda aynı değerleri almıştır.

Aşağıda sunulan Tablo 13 ve Tablo 14, sırasıyla madde ayırt edicilik düzeyinin yüksek ve düşük olduğu 6 niteliği ölçen 60 maddelik bir test için farklı örneklem büyüklüğü koşullarında gerçek ve tahmini madde parametreleri arasındaki “mutlak sapma” ve “ortalama mutlak sapma” değerlerini sunmaktadır.

Tablo 13

6 Niteliği Ölçen Yüksek Ayırt Ediciliğe Sahip 60 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,10	,13	,15	,15	,06	,04	,03	,03	-	-	-	-	-	-	-	-
2	,03	,02	,01	,01	,04	,03	,01	,01	-	-	-	-	-	-	-	-
3	,04	,03	,01	,01	,05	,05	,01	,01	-	-	-	-	-	-	-	-
4	,02	,02	,01	,01	-	-	-	-	,02	,02	,01	,01	-	-	-	-
5	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
6	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
7	,03	,02	,01	,01	,05	,04	,03	,03	-	-	-	-	,04	,04	,02	,02
8	,03	,02	,01	,01	,05	,03	,02	,02	-	-	-	-	,03	,02	,01	,01
9	,41	,34	,29	,28	,12	,11	,12	,12	-	-	-	-	,03	,02	,01	,01
10	,50	,51	,50	,50	,11	,12	,12	,12	-	-	-	-	,03	,02	,01	,01
11	,05	,04	,02	,01	,10	,09	,09	,09	-	-	-	-	,20	,21	,21	,21
12	,05	,05	,02	,01	,09	,09	,10	,09	-	-	-	-	,21	,21	,20	,21
13	,45	,37	,32	,31	,11	,11	,12	,12	-	-	-	-	,03	,02	,02	,01
14	,47	,44	,41	,41	,08	,08	,08	,08	,12	,12	,12	,12	,06	,06	,07	,08
15	,06	,08	,11	,11	,07	,05	,04	,04	,04	,03	,03	,03	,03	,02	,01	,01
16	,05	,05	,02	,01	,10	,10	,11	,10	,06	,06	,05	,05	,20	,20	,19	,20
17	,02	,01	,01	,01	,05	,04	,02	,01	,10	,10	,10	,10	,10	,10	,10	,10
18	,03	,04	,02	,01	,09	,07	,03	,02	,04	,04	,03	,03	,07	,06	,07	,07
19	,04	,05	,02	,01	,10	,10	,09	,10	,05	,03	,02	,01	,21	,22	,21	,20
20	,03	,03	,01	,01	,05	,06	,02	,01	,05	,03	,01	,01	,03	,02	,01	,01
21	,10	,13	,15	,15	,06	,04	,03	,03	-	-	-	-	-	-	-	-
22	,03	,01	,01	,01	,03	,03	,01	,01	-	-	-	-	-	-	-	-
23	,04	,03	,01	,01	,05	,04	,01	,01	-	-	-	-	-	-	-	-
24	,02	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
25	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
26	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
27	,02	,02	,01	,01	,06	,04	,04	,03	-	-	-	-	,05	,04	,02	,02
28	,02	,02	,01	,01	,05	,04	,02	,02	-	-	-	-	,03	,02	,01	,01
29	,41	,34	,29	,28	,12	,12	,12	,12	-	-	-	-	,03	,02	,01	,01
30	,50	,51	,50	,50	,11	,12	,12	,12	-	-	-	-	,03	,02	,01	,01
31	,06	,04	,02	,01	,09	,09	,09	,09	-	-	-	-	,21	,21	,21	,21
32	,05	,04	,02	,01	,09	,09	,09	,10	-	-	-	-	,21	,21	,21	,20
33	,45	,37	,32	,31	,12	,12	,12	,12	-	-	-	-	,03	,03	,01	,01
34	,47	,44	,41	,41	,08	,09	,08	,08	,12	,11	,12	,12	,05	,06	,07	,08
35	,06	,08	,11	,11	,08	,05	,04	,04	,04	,03	,03	,03	,03	,02	,01	,01
36	,05	,05	,02	,01	,10	,10	,11	,10	,06	,06	,05	,05	,20	,21	,19	,20
37	,02	,02	,01	,01	,05	,03	,02	,01	,10	,10	,10	,10	,09	,10	,10	,10
38	,03	,04	,01	,01	,07	,07	,03	,02	,04	,04	,03	,03	,07	,07	,07	,07
39	,05	,05	,02	,01	,11	,10	,10	,10	,04	,03	,02	,01	,20	,21	,20	,20
40	,04	,03	,01	,01	,05	,05	,02	,01	,05	,03	,02	,01	,03	,02	,01	,01
41	,09	,13	,15	,15	,07	,04	,03	,03	-	-	-	-	-	-	-	-
42	,03	,02	,01	,01	,04	,03	,02	,01	-	-	-	-	-	-	-	-
43	,04	,03	,01	,01	,04	,05	,01	,01	-	-	-	-	-	-	-	-
44	,02	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
45	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
46	,02	,01	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
47	,02	,02	,01	,01	,06	,05	,03	,03	-	-	-	-	,04	,04	,02	,02
48	,03	,02	,01	,01	,05	,04	,02	,02	-	-	-	-	,03	,02	,01	,01
49	,41	,34	,29	,28	,11	,11	,12	,12	-	-	-	-	,03	,02	,01	,01
50	,50	,51	,50	,50	,12	,12	,11	,11	-	-	-	-	,03	,02	,01	,01
51	,05	,04	,02	,01	,09	,09	,10	,09	-	-	-	-	,21	,21	,20	,21
52	,05	,05	,02	,01	,10	,09	,09	,09	-	-	-	-	,20	,21	,21	,21
53	,45	,38	,32	,31	,12	,12	,12	,12	-	-	-	-	,03	,03	,01	,01
54	,48	,44	,40	,41	,08	,08	,08	,08	,12	,12	,12	,12	,06	,07	,07	,08
55	,05	,08	,11	,11	,07	,05	,04	,04	,04	,03	,03	,03	,03	,02	,01	,01
56	,05	,05	,02	,01	,10	,10	,10	,10	,06	,06	,05	,05	,20	,20	,20	,20
57	,03	,02	,01	,01	,06	,04	,02	,01	,10	,10	,10	,10	,10	,10	,10	,10
58	,03	,04	,01	,01	,08	,07	,03	,02	,04	,04	,03	,03	,07	,06	,07	,07
59	,04	,05	,02	,01	,11	,10	,10	,09	,04	,03	,02	,01	,20	,22	,20	,21
60	,03	,03	,01	,01	,06	,05	,02	,02	,04	,03	,02	,01	,03	,01	,01	,01
Ortalama Mutlak Sapma	,12	,11	,10	,10	,08	,07	,06	,06	,05	,05	,04	,04	,09	,09	,08	,08

Tablo 13'te yer alan faktörlere ait 6 nitelik, 60 madde ve yüksek ayırt edicilik koşulları dikkate alındığında farklı örneklem büyüklüğü koşulları için “mutlak sapma” değerlerinin dört farklı madde parametresinde sırasıyla ,00 ile ,51 (h_j); ,01 ile ,12 (g_j); ,01 ile ,12 (ω_j) ve ,01 ile ,22 (ϵ_j) arasında değiştiği görülmektedir. h_j madde parametresinde sapmalar, diğer üç madde parametresine göre daha yüksektir. Bu parametreye ilişkin gerçek ve tahmini madde parametreleri arasındaki en büyük sapmaya sahip maddelerin 9., 10., 13., 14., 29., 30., 33., 34., 49., 50., 53. ve 54. maddeler olduğu görülmektedir. “Mutlak sapma” değerlerine yönelik farklı örneklem büyüklüğü koşulları genel olarak değerlendirildiğinde h_j madde parametresine ait bu değerlerin örneklem büyüklüğünün 250 olduğu koşulda bazı maddeler hariç (madde 1 ve madde 15 gibi) daha yüksek olduğu ve örneklem büyüklüğü arttıkça madde parametresinde artış ya da azalışlar olsa da genel anlamda bu değerlerin azaldığı görülmektedir. Örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda ise değerler genel olarak birbirine benzerdir. 1., 15., 21., 25., 41. ve 55. maddeler her bir örneklem büyüklüğü koşulunda örneklem büyüklükleri arttıkça ,01 ile ,06 arasında değişen daha büyük “mutlak sapma” değerlerine sahiptir. Tabloda yer alan g_j , ω_j ve ϵ_j madde parametreleri incelendiğinde g_j madde parametresine ait değerlerde, örneklem büyüklüğü artışıyla birlikte ,01 ile ,07 arasında artış ve azalışların olduğu görülmektedir. Madde 14, 29, 31, 33, 53, 54 ve 56'ya ait değerler ise her bir örneklem büyüklüğü koşulu için aynıdır. ω_j madde parametresine ilişkin değerler incelendiğinde 34. madde örneklem büyüklüğü artışıyla birlikte artış ve azalışlar sergilese de diğer maddelerin “mutlak sapma” değerlerinde örneklem büyüklüğü artışıyla birlikte ,01 ile ,04 arasında değişen artış ya da azalışların olduğu görülmektedir. Ayrıca 14., 17., 37. ve 57. maddeler örneklem büyüklüğü artışıyla birlikte aynı “mutlak sapma” değerlerini almıştır. Genel olarak değerlendirildiğinde örneklem büyüklüğü arttıkça “mutlak sapma” değerlerinin azaldığı, 1500 ve 3000 örneklem büyüklüğü koşullarında değerlerin bazı maddeler haricinde genel olarak değişmediği görülmektedir.

“Ortalama mutlak sapma” değerleri incelendiğinde, farklı örneklem büyüklüklerinde dört madde parametresi için bu değerlerin ,04 ile ,12 arasında değiştiği ve örneklem büyüklüğü artışıyla birlikte azaldığı görülmektedir. Örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda ise dört madde parametresine ait “ortalama mutlak sapma” değerleri aynıdır.

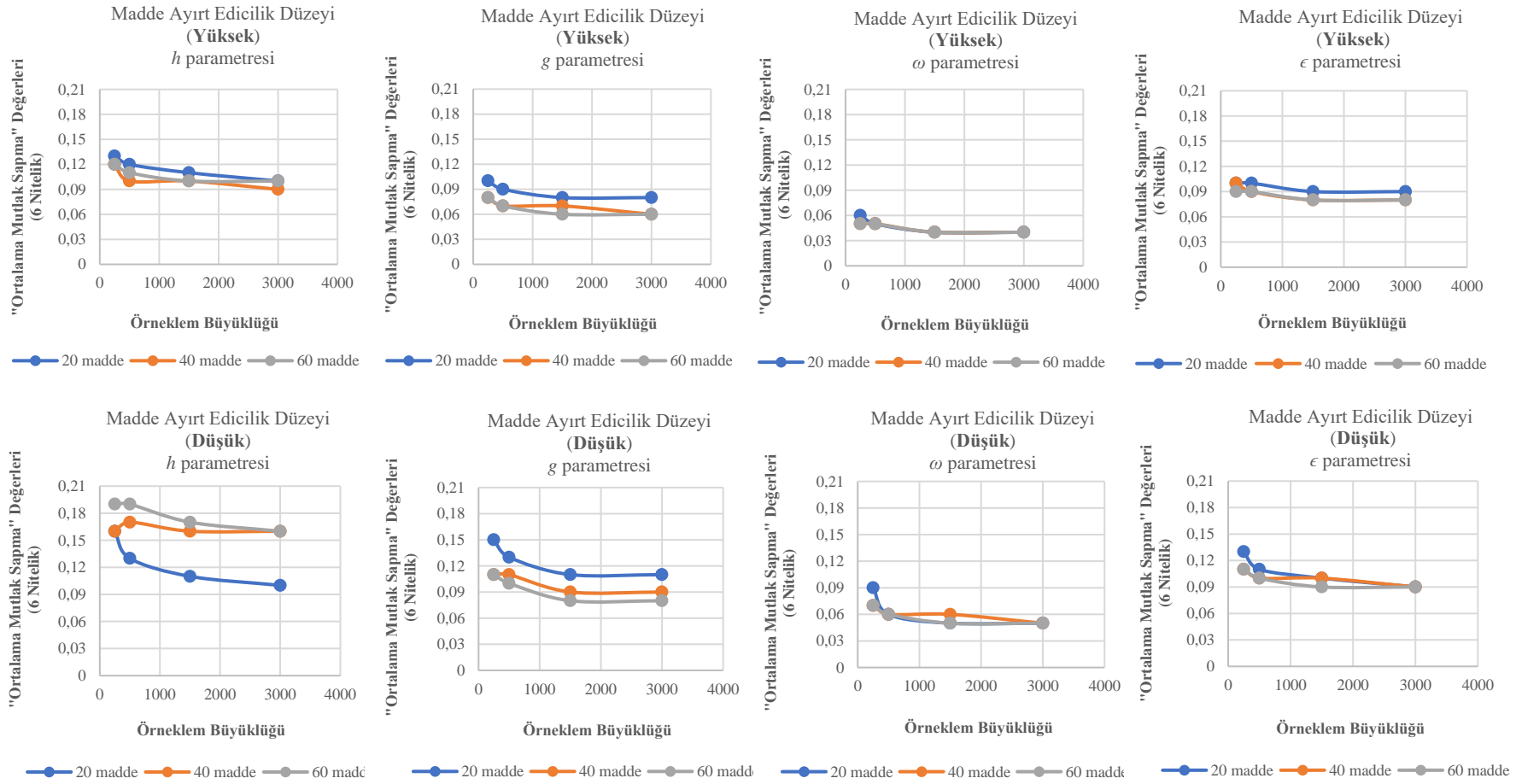
Tablo 14

6 Niteliği Ölçen Düşük Ayırt Ediciliğe Sahip 60 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,04	,05	,02	,01	,08	,07	,07	,07	-	-	-	-	-	-	-	-
2	,12	,11	,14	,15	,08	,07	,08	,09	-	-	-	-	-	-	-	-
3	,16	,17	,15	,15	,16	,13	,07	,06	-	-	-	-	-	-	-	-
4	,05	,05	,05	,05	-	-	-	-	,03	,03	,01	,01	-	-	-	-
5	,03	,02	,01	,01	-	-	-	-	,04	,03	,02	,02	-	-	-	-
6	,03	,02	,01	,01	-	-	-	-	,03	,03	,01	,01	-	-	-	-
7	,04	,04	,02	,02	,05	,06	,03	,03	-	-	-	-	,05	,04	,03	,03
8	,05	,04	,02	,01	,08	,06	,07	,07	-	-	-	-	,08	,08	,09	,10
9	,65	,60	,71	,73	,21	,21	,20	,20	-	-	-	-	,05	,04	,02	,02
10	,33	,23	,06	,02	,09	,07	,07	,06	-	-	-	-	,05	,05	,03	,03
11	,13	,15	,13	,12	,10	,10	,10	,10	-	-	-	-	,20	,20	,20	,20
12	,14	,15	,12	,11	,10	,09	,10	,11	-	-	-	-	,21	,21	,20	,19
13	,66	,67	,67	,70	,10	,08	,08	,07	-	-	-	-	,06	,05	,04	,04
14	,58	,60	,60	,62	,09	,07	,07	,06	,12	,13	,13	,14	,07	,06	,04	,04
15	,05	,04	,02	,02	,07	,07	,07	,07	,06	,03	,02	,02	,04	,03	,02	,02
16	,15	,21	,13	,09	,09	,09	,08	,08	,10	,10	,10	,10	,25	,26	,23	,22
17	,09	,10	,11	,12	,09	,08	,07	,07	,09	,08	,07	,07	,09	,08	,07	,07
18	,12	,11	,06	,06	,18	,15	,07	,04	,08	,07	,05	,04	,07	,08	,07	,06
19	,19	,22	,19	,18	,14	,13	,13	,13	,05	,05	,05	,05	,22	,23	,18	,17
20	,13	,14	,11	,10	,15	,14	,06	,05	,09	,08	,05	,03	,06	,05	,04	,04
21	,04	,04	,03	,02	,08	,06	,07	,07	-	-	-	-	-	-	-	-
22	,11	,12	,14	,15	,08	,08	,08	,09	-	-	-	-	-	-	-	-
23	,17	,18	,14	,14	,16	,15	,07	,06	-	-	-	-	-	-	-	-
24	,05	,05	,05	,05	-	-	-	-	,03	,02	,01	,01	-	-	-	-
25	,03	,02	,01	,01	-	-	-	-	,03	,03	,02	,02	-	-	-	-
26	,03	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
27	,05	,03	,02	,02	,05	,06	,04	,03	-	-	-	-	,04	,04	,04	,03
28	,05	,04	,02	,01	,07	,07	,07	,06	-	-	-	-	,09	,08	,10	,10
29	,64	,61	,71	,73	,21	,21	,20	,20	-	-	-	-	,06	,03	,02	,02
30	,33	,24	,07	,02	,09	,07	,06	,06	-	-	-	-	,06	,04	,03	,03
31	,13	,14	,13	,12	,09	,10	,10	,10	-	-	-	-	,21	,20	,20	,20
32	,13	,15	,13	,12	,10	,10	,10	,10	-	-	-	-	,20	,20	,20	,20
33	,66	,67	,67	,70	,10	,08	,08	,07	-	-	-	-	,06	,05	,04	,04
34	,58	,59	,60	,62	,09	,08	,06	,06	,12	,12	,14	,14	,07	,06	,04	,03
35	,05	,04	,02	,02	,08	,06	,07	,07	,05	,04	,02	,02	,04	,03	,02	,02
36	,16	,20	,13	,09	,10	,09	,08	,08	,10	,10	,10	,10	,24	,26	,23	,22
37	,08	,10	,12	,12	,09	,08	,07	,07	,09	,08	,07	,07	,09	,07	,07	,07
38	,13	,11	,07	,06	,17	,14	,06	,04	,07	,07	,05	,04	,07	,06	,06	,06
39	,19	,21	,18	,18	,14	,12	,14	,14	,06	,05	,05	,05	,24	,23	,17	,16
40	,13	,15	,11	,10	,16	,13	,07	,05	,10	,09	,04	,03	,06	,05	,04	,04
41	,04	,04	,03	,01	,07	,07	,07	,08	-	-	-	-	-	-	-	-
42	,12	,12	,14	,15	,08	,08	,08	,08	-	-	-	-	-	-	-	-
43	,16	,18	,15	,15	,16	,14	,08	,07	-	-	-	-	-	-	-	-
44	,05	,05	,05	,05	-	-	-	-	,03	,02	,01	,01	-	-	-	-
45	,03	,02	,01	,01	-	-	-	-	,04	,03	,02	,02	-	-	-	-
46	,03	,02	,01	,01	-	-	-	-	,04	,02	,01	,01	-	-	-	-
47	,05	,03	,02	,02	,06	,05	,04	,03	-	-	-	-	,04	,05	,03	,03
48	,05	,04	,02	,01	,07	,07	,06	,06	-	-	-	-	,09	,08	,09	,10
49	,65	,61	,71	,73	,21	,21	,20	,20	-	-	-	-	,05	,04	,02	,02
50	,34	,24	,06	,02	,09	,08	,07	,06	-	-	-	-	,05	,05	,03	,03
51	,14	,15	,12	,12	,09	,10	,10	,11	-	-	-	-	,21	,20	,20	,19
52	,14	,14	,14	,12	,09	,10	,10	,10	-	-	-	-	,21	,20	,20	,20
53	,66	,67	,66	,70	,10	,09	,08	,08	-	-	-	-	,06	,05	,04	,04
54	,58	,59	,60	,62	,09	,08	,07	,06	,12	,13	,13	,14	,07	,06	,04	,03
55	,05	,04	,02	,02	,09	,07	,07	,07	,05	,04	,02	,02	,03	,03	,02	,02
56	,14	,20	,13	,09	,10	,09	,08	,08	,11	,10	,10	,10	,24	,26	,23	,22
57	,09	,10	,11	,12	,08	,07	,07	,07	,08	,08	,07	,06	,10	,07	,07	,07
58	,13	,12	,06	,06	,18	,14	,06	,04	,07	,07	,05	,04	,08	,07	,06	,06
59	,18	,22	,19	,18	,14	,13	,14	,13	,06	,05	,05	,05	,24	,23	,17	,17
60	,13	,14	,10	,10	,17	,13	,07	,05	,09	,08	,05	,03	,06	,05	,04	,04
Ortalama Mutlak Sapma	,19	,19	,17	,16	,11	,10	,08	,08	,07	,06	,05	,05	,11	,10	,09	,09

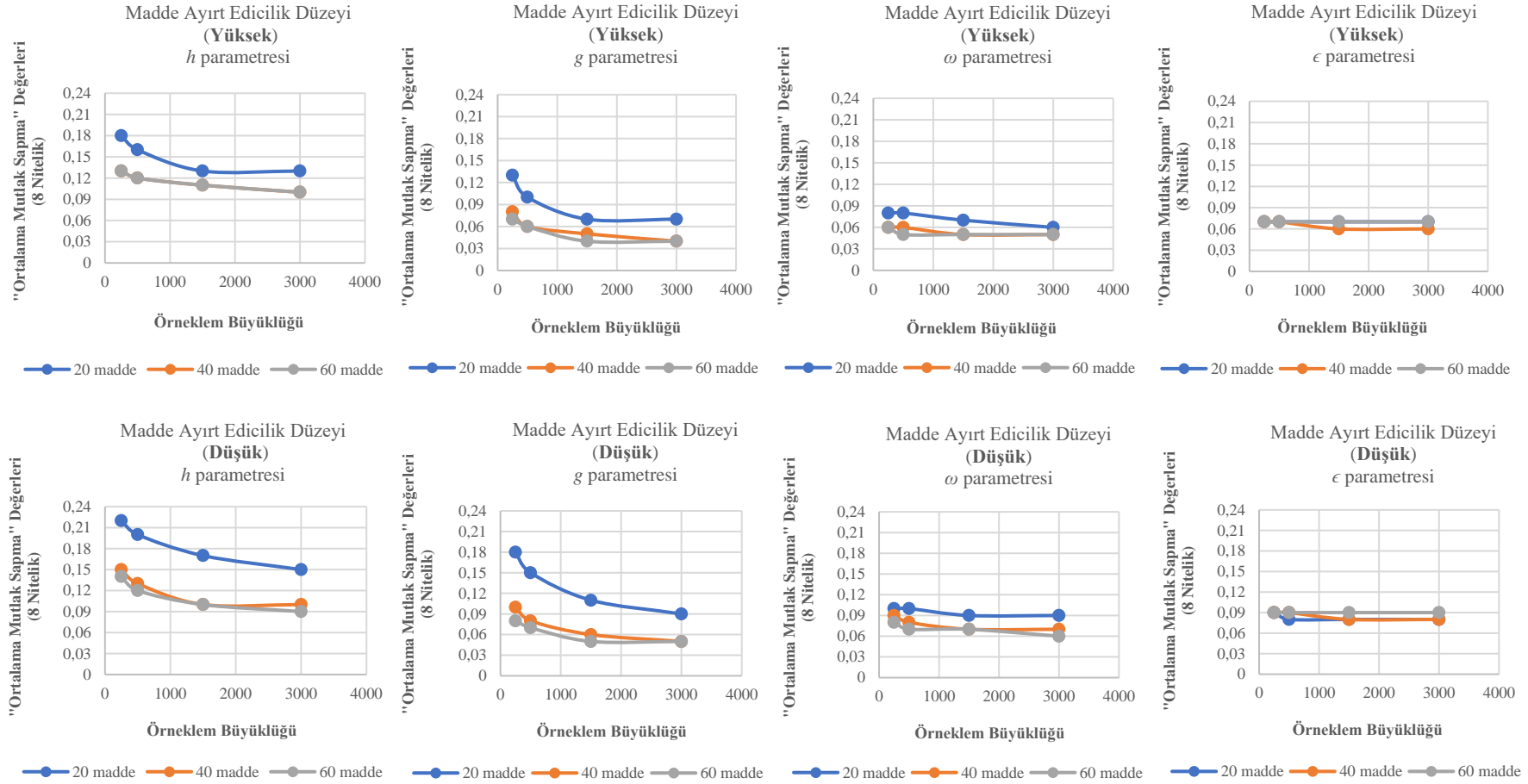
Tablo 14’te yer alan faktörlere ait 6 nitelik, 60 madde ve düşük ayırt edicilik koşulları dikkate alındığında farklı örneklem büyüklüğü koşulları için “mutlak sapma” değerlerinin dört farklı madde parametresinde sırasıyla ,01 ile ,73 (h_j); ,03 ile ,21 (g_j); ,01 ile ,14 (ω_j) ve ,02 ile ,26 (ϵ_j) arasında değiştiği görülmektedir. h_j madde parametresinde sapmalar, diğer üç madde parametresine göre daha fazladır. Bu parametreye ilişkin gerçek ve tahmini madde parametreleri arasındaki en büyük sapmaya sahip maddelerin 9., 10., 13., 14., 29., 30., 33., 34., 49., 50., 53. ve 54. maddeler olduğu görülmektedir. “Mutlak sapma” değerlerine yönelik farklı örneklem büyüklüğü koşulları genel olarak değerlendirildiğinde h_j madde parametresine ait değerlerin örneklem büyüklüğünün 250 olduğu koşulda daha yüksek olduğu ve örneklem büyüklüğü artışıyla birlikte değerlerin ,01 ile ,32 arasında bir artış ya da azalış sergilediği görülmektedir. Ayrıca 4., 24. ve 44. maddeler örneklem büyüklüğü artışıyla birlikte aynı “mutlak sapma” değerlerini almıştır. Tabloda yer alan g_j , ω_j ve ϵ_j madde parametreleri incelendiğinde g_j parametresinin örneklem büyüklüğü artışıyla birlikte bazı maddeler (2., 8., 12., 29. ve 59. maddeler gibi) hariç diğer maddelerin “mutlak sapma” değerlerinde ,01 ile ,14 arasında değişen azalışların olduğu görülmektedir. g_j madde parametresi için 11., 15., 32. ve 42. maddeler her bir örneklem büyüklüğü koşulunda aynı değerlere sahiptir. ω_j madde parametresine ilişkin değerler incelendiğinde maddelerin “mutlak sapma” değerlerinde örneklem büyüklüğü artışıyla birlikte ,01 ile ,07 arasında değişen artış ya da azalışların olduğu görülmektedir. Ayrıca 16., 19. ve 36. maddelere ait “mutlak sapma” değerleri örneklem büyüklüğü artsa da aynı değerleri almıştır. ϵ_j madde parametresine ilişkin değerler incelendiğinde bazı maddeler kendi içinde (18., 28. ve 48. maddeler gibi) örneklem büyüklüğü artışıyla birlikte artış ve azalışlar sergilese de diğer maddelerin “mutlak sapma” değerlerinde örneklem büyüklüğü artışıyla birlikte ,01 ile ,08 arasında değişen artış ya da azalışların olduğu görülmektedir. Ayrıca 11. ve 32. maddelere ait “mutlak sapma” değerleri örneklem büyüklüğü artsa da aynı değerleri almıştır. “Ortalama mutlak sapma” değerleri incelendiğinde, farklı örneklem büyüklüklerinde bu değerlerin dört madde parametresi için ,05 ile ,19 arasında değiştiği ve örneklem büyüklüğü artışıyla birlikte parametre tahminlerine ait sapmalarda azalışların olduğu görülmektedir. h_j madde parametresi hariç diğer üç madde parametresine ait değerler incelendiğinde ise örneklem büyüklüğünün 1500 ile 3000 olduğu koşullarda bu değerlerin aynı kaldığı görülmektedir.

Seçilen faktörler altında değişimlenen her bir koşulu özetlemek ve 6, 8, 10 nitelikli Q-matrisler kullanıldığında “ortalama mutlak sapma” değerlerinin genel bir karşılaştırmasını yapabilmek adına oluşturulan grafikler aşağıda sunulmuştur.



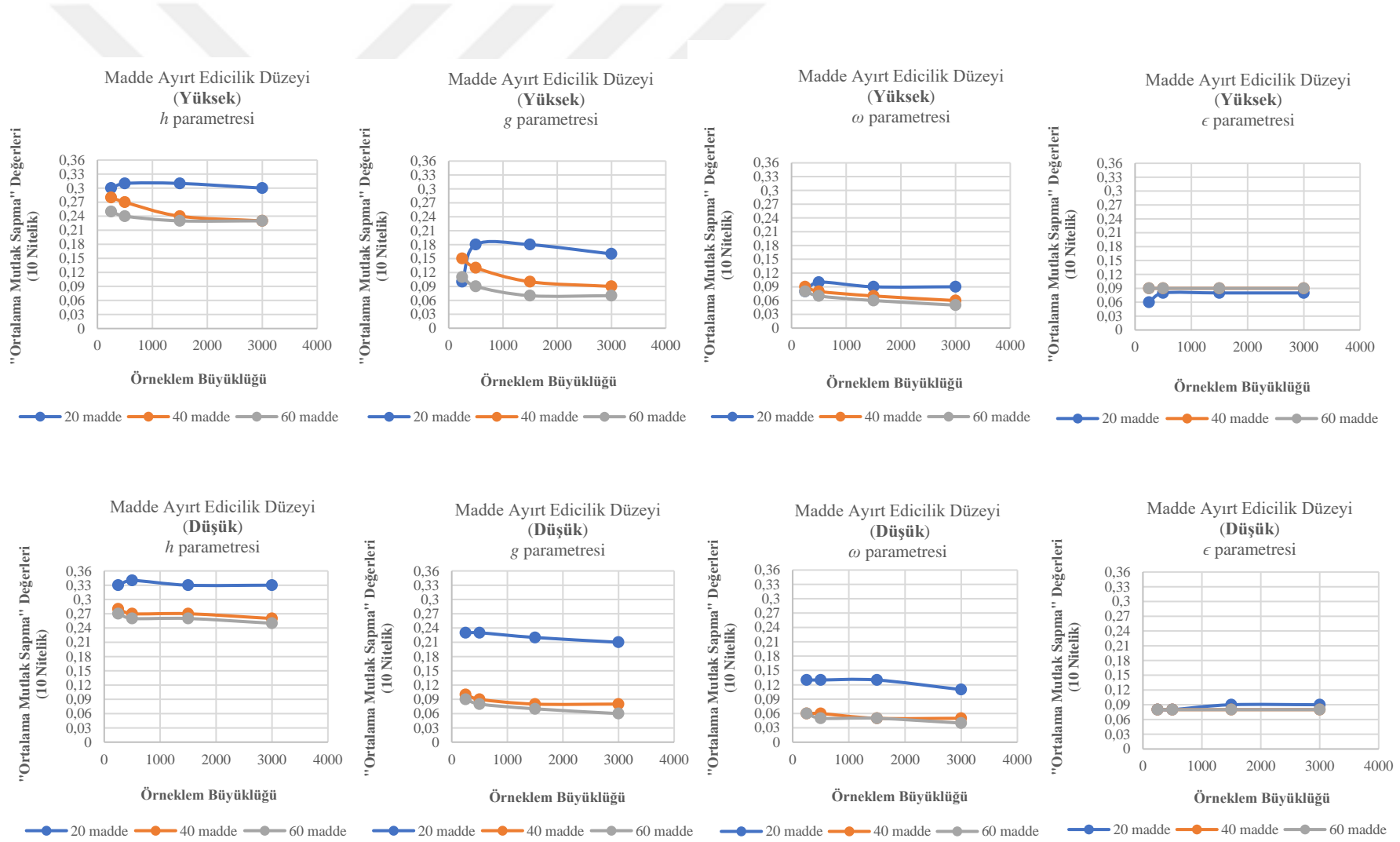
Şekil 6. 6 nitelikli Q-matriste dört madde parametresine ait farklı simülasyon koşullarından elde edilen “ortalama mutlak sapma” değerleri

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.



Şekil 7. 8 nitelikli Q-matriste dört madde parametresine ait farklı simülasyon koşullarından elde edilen “ortalama mutlak sapma” değerleri

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.



Şekil 8. 10 nitelikli Q-matriste dört madde parametresine ait farklı simülasyon koşullarından elde edilen “ortalama mutlak sapma” değerleri h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.

Şekiller incelendiğinde 6 nitelikli Q-matrisin kullanıldığı koşulda SISM'e ait dört madde parametresi için diğer koşullar sabit tutulduğunda, örneklem büyüklüğü artışıyla birlikte madde parametre tahminlerinin yani “ortalama mutlak sapma” değerlerinin azaldığı görülmektedir. Özellikle ölçülen tüm becerilere sahip olan ve kavram yanılgıları olmayan bireylerin başarı olasılığını ifade eden (h_j) madde parametresi hariç g_j , ω_j ve ϵ_j madde parametrelerine ait 1500 ve 3000 örneklem büyüklüğü koşullarından elde edilen “ortalama mutlak sapma” değerleri oldukça benzerdir. Çalışmada yer alan test uzunluğu faktörü incelendiğinde “*test uzunluğunun 60*” ve “*madde ayırt edicilik düzeyinin düşük*” olduğu koşulu içeren 6 nitelikli Q-matris için h_j madde parametresine ait “ortalama mutlak sapma” değerleri 20 ve 40 maddelik test uzunluğu koşullarıyla kıyaslandığında bu değerlerde artışların olduğu görülmüştür. Benzer bir durum madde ayırt edicilik düzeyinin düşük, test uzunluğunun 40 ve örneklem büyüklüğünün 250'den 500'e çıkarıldığı koşulda da mevcuttur. “Ortalama mutlak sapma” değerlerine ilişkin monoton olmayan artış ya da azalışlar olduğu görülse de genel olarak incelendiğinde test uzunluğunun artmasıyla birlikte, madde parametrelerine ait bu değerlerin azaldığı görülmektedir. Ancak test uzunluğunun 40 ve 60 madde olduğu koşullarda ise madde parametrelerine ait değerler hemen hemen aynıdır. Ayrıca daha yüksek madde ayırt edicilik düzeyiyle bu değerler azalırken madde ayırt edicilik düzeyinin düştüğü koşullarda artış göstererek madde parametrelerine ait sapmaların arttığı bulgusuna ulaşılmıştır. g_j , ω_j ve ϵ_j madde parametrelerine ait değerler ise aynı kalmış ya da küçük değişimlerle de olsa azalış göstermiştir.

8 nitelikli Q-matrisin kullanıldığı koşulda SISM'e ait dört madde parametresi için diğer koşullar sabit tutulduğunda, örneklem büyüklüğü artışıyla birlikte “ortalama mutlak sapma” değerlerinde genel olarak azalmalar olduğu görülmektedir. 6 nitelikli Q-matriste olduğu gibi, 8 nitelikli Q-matriste de madde ayırt edicilik düzeyinin yüksek olduğu koşulda h_j madde parametresi hariç 1500 ve 3000 örneklem büyüklüğü koşullarından elde edilen “ortalama mutlak sapma” değerleri birbirine daha yakındır. Test uzunluğu faktörü incelendiğinde değişimlenen 20, 40 ve 60 maddelik koşullar için test uzunluğunun artmasıyla birlikte madde parametrelerine ait “ortalama mutlak sapma” değerlerinin genel olarak azaldığı ve özellikle test uzunluğunun 20 maddeden 40 maddeye çıkarıldığı koşulda madde parametrelerine ait sapma değerlerindeki azalmaların daha fazla olduğu görülmektedir. Ancak madde ayırt edicilik düzeyinin yüksek ve düşük olduğu her iki koşulda da test uzunluğunun 40 maddeden 60 maddeye çıkarıldığı koşulda ϵ_j madde parametresine ait “ortalama mutlak sapma” değerleri bir miktar artmıştır. 6 nitelikli Q-matriste olduğu gibi 8 nitelikli Q-matriste de

monoton olmayan artış ya da azalışlar olduğu görülse de bu değerlerin daha yüksek madde ayırt edicilik düzeyiyle azalırken, madde ayırt edicilik düzeyinin düştüğü koşulda arttığı bulgusuna ulaşılmıştır.

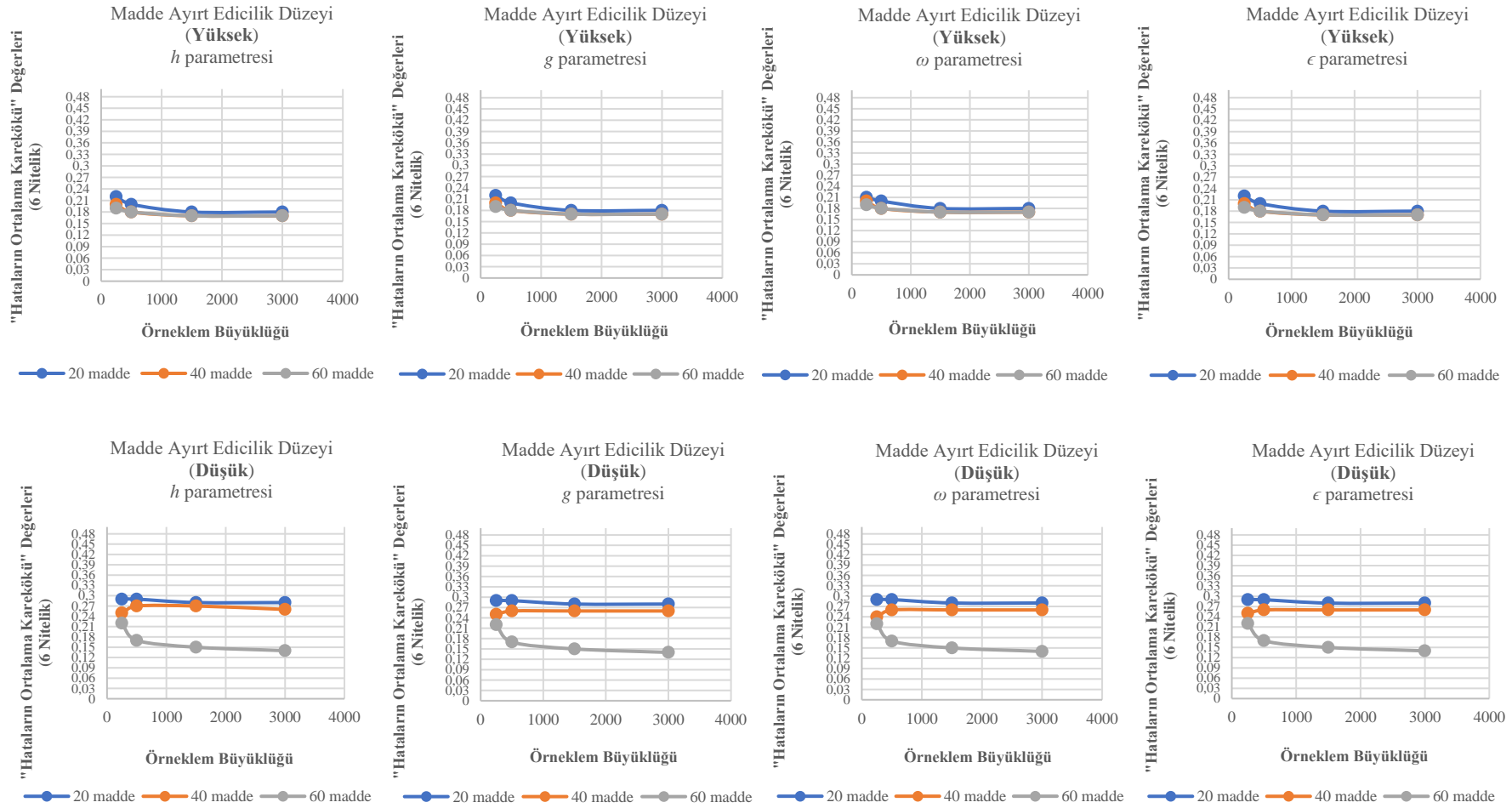
10 nitelikli Q-matrisin kullanıldığı koşulda SISM'e ait dört madde parametresi için diğer koşullar sabit tutulduğunda, örneklem büyüklüğünün artışıyla birlikte özellikle örneklem büyüklüğünün 250 ve 500 olduğu koşullarda monoton olmayan artış ve azalışların olduğu görülmektedir. Bu duruma benzer şekilde test uzunluğu faktörü incelendiğinde madde ayırt edicilik düzeyinin yüksek ve düşük olduğu her iki koşulda da değişimlenen 20, 40 ve 60 maddelik test uzunluklarında madde sayısının artmasıyla birlikte madde parametrelerine ait "ortalama mutlak sapma" değerlerinin h_j , g_j ve ω_j parametreleri için azaldığı ancak ϵ_j madde parametresine ait "ortalama mutlak sapma" değerlerinin 20 maddelik test uzunluğu koşulunda daha küçük sapmalara sahip olduğu görülmektedir. Ayrıca bu değerlerin daha yüksek madde ayırt edicilik düzeyiyle azalırken, madde ayırt edicilik düzeyinin düştüğü koşulda artış gösterdiği bulgusuna ulaşılmıştır.

Farklı faktörlerde 6, 8 ve 10 nitelikli Q-matrislerin kullanıldığı koşullardan elde edilen "ortalama mutlak sapma" değerlerinin ölçülen nitelik sayısının artmasıyla birlikte arttığı görülmektedir. Ancak ölçülen nitelik sayısının artmasıyla birlikte 6 ve 8 nitelikli Q-matrislerin kullanıldığı koşullardan elde edilen madde parametrelerine ait değerlere ilişkin, h_j madde parametresine ait değerlerdeki artışlar g_j , ω_j ve ϵ_j madde parametrelerine göre nispeten daha fazladır. Ayrıca g_j , ω_j ve ϵ_j madde parametrelerine ait değerlerin birbirlerine daha yakın oldukları görülürken 10 nitelikli Q-matrisin kullanıldığı koşulda madde parametrelerinden elde edilen "ortalama mutlak sapma" değerlerinin özellikle h_j madde parametresinde 6 ve 8 nitelikli Q-matrislerin kullanıldığı koşullara kıyasla daha fazla artış gösterdiği ve g_j , ω_j ve ϵ_j madde parametrelerinin bu durumlardan daha küçük artışlarla etkilendiği bulgusuna ulaşılmıştır.

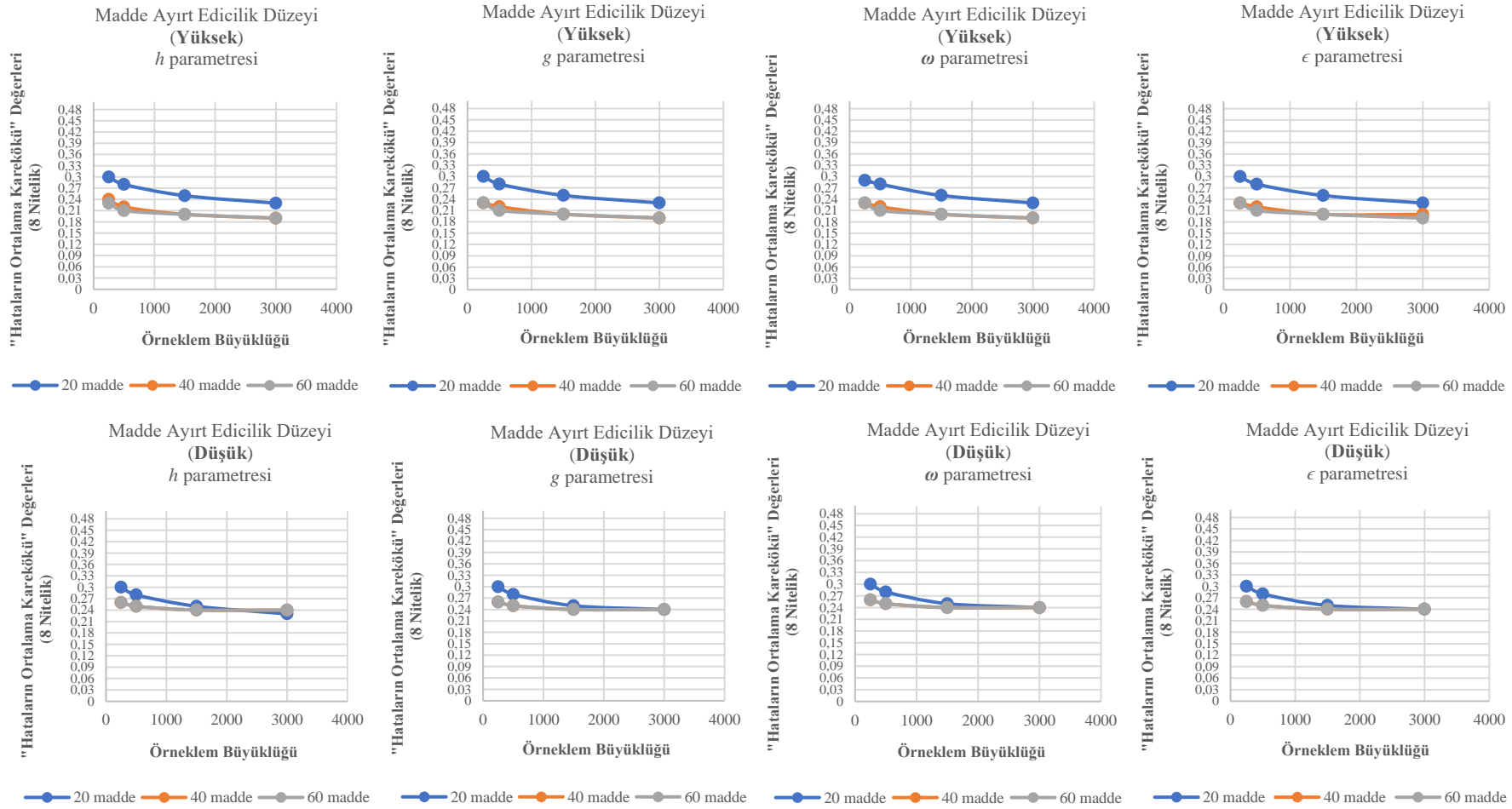
4.2. İkinci Alt Probleme İlişkin Bulgular

2. SISM'e göre madde parametrelerine ait "hataların ortalama karekökü" değerleri,
 - 2.1. Örneklem büyüklüğü (250, 500, 1500 ve 3000),
 - 2.2. Test uzunluğu (20, 40 ve 60 madde),
 - 2.3. Madde ayırt edicilik düzeyi (yüksek ayırt edicilik ve düşük ayırt edicilik) ve
 - 2.4. Ölçülen nitelik sayısına (6, 8 ve 10 nitelik) göre nasıl değişmektedir?

İkinci alt problem kapsamında SISM'den elde edilen madde parametrelerine ilişkin "hataların ortalama karekökü" değerleri çalışmanın amacı doğrultusunda örneklem büyüklüğü, test uzunluğu, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısı olmak üzere dört faktörün her bir koşulu için 100 iterasyonun analizinden elde edilmiş ve EK 5'te sunulmuştur. Çalışmanın bu bölümünde 6, 8 ve 10 nitelikli Q-matrislerin kullanıldığı koşullar için faktörlerde değişimlenen her bir koşulda EK 5'te yer alan dört madde parametresine ait "hataların ortalama karekökü" değerleri kullanılarak grafikler oluşturulmuş ve her bir koşul için aşağıda sunulan şekillerle özetlenmiştir.

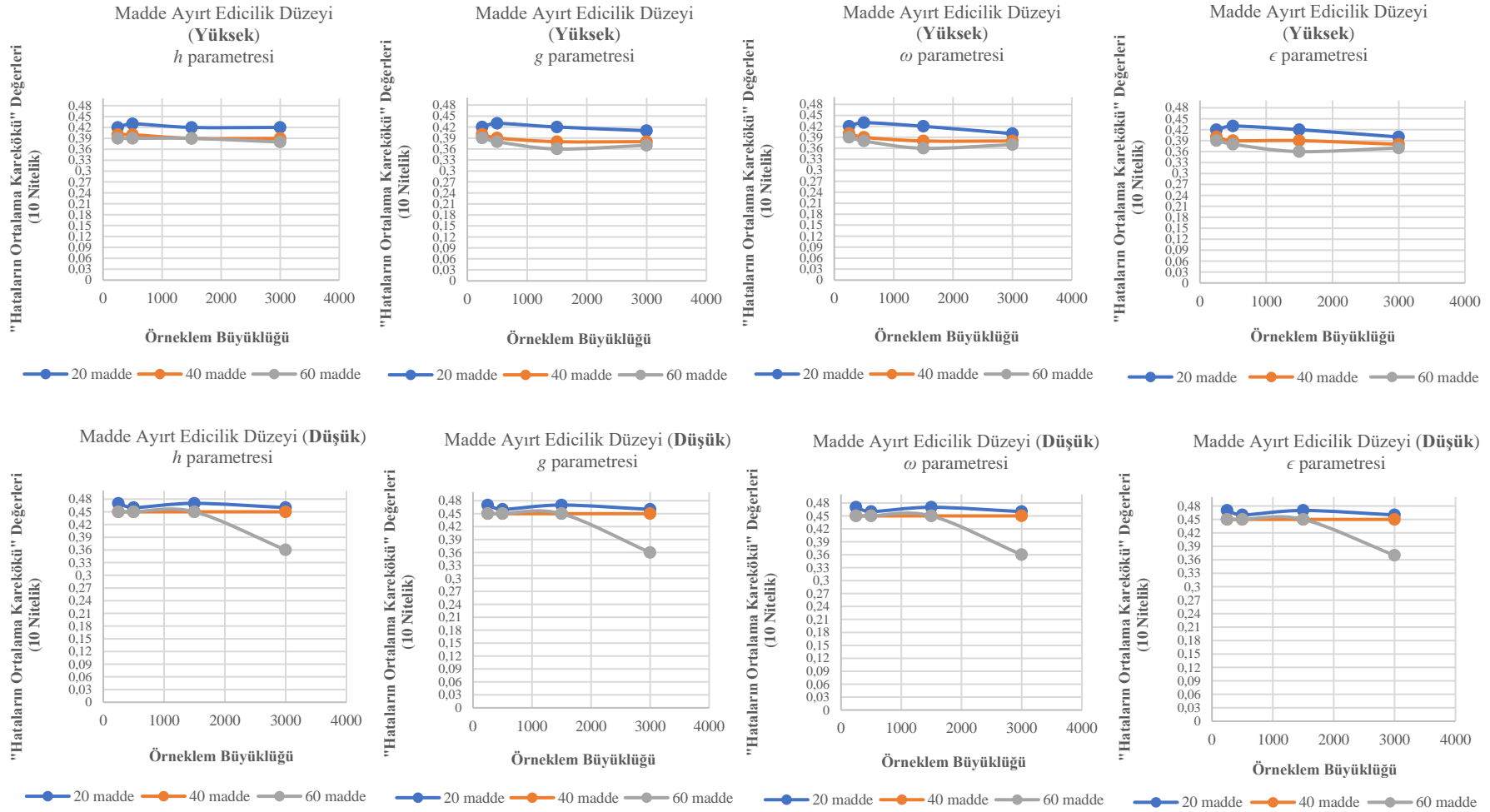


Şekil 9. 6 nitelikli Q-matriste dört madde parametresine ait farklı simülasyon koşullarından elde edilen “hataların ortalama karekökü” değerleri
 h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.



Şekil 10. 8 nitelikli Q-matriste dört madde parametresine ait farklı simülasyon koşullarından elde edilen “hataların ortalama karekökü” değerleri

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.



Şekil 11. 10 nitelikli Q-matriste dört madde parametresine ait farklı simülasyon koşullarından elde edilen "hataların ortalama karekökü" değerleri

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.

Şekiller incelendiğinde 6 nitelikli Q-matrisin kullanıldığı koşulda SISM'e ait dört madde parametresine ilişkin "hataların ortalama karekökü" değerleri incelendiğinde madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu değerlerin yaklaşık olarak ,17 ile ,22; madde ayırt edicilik düzeyinin düşük olduğu koşulda yaklaşık olarak ,14 ile ,29 arasında değiştiği görülmektedir. Madde ayırt edicilik düzeyinin yüksek olduğu koşulda, test uzunluğu 20 maddeden 40 maddeye çıkarıldığında bu değerlerin azaldığı ancak test uzunluğunun 40 maddeden 60 maddeye çıkarıldığı koşulda ise bu değerlerin benzer olduğu görülmektedir. Madde ayırt edicilik düzeyinin yüksek olduğu koşulda küçük örneklemelerden büyük örneklemelere artışla birlikte madde parametre tahminlerinin azaldığı görülse de, örneklem büyüklüğünün 1500'den 3000'e çıkarıldığı koşulda dört madde parametre değerinin de aynı kaldığı ve modele ait madde parametre değerlerinde ek bir azalışın olmadığı görülmektedir. Madde ayırt edicilik düzeyinin düşük olduğu ve madde ayırt edicilik düzeyinin yüksek olduğu koşullar karşılaştırıldığında, madde ayırt edicilik düzeyi faktörünün madde parametre tahminlerini etkilediği yani madde ayırt edicilik düzeyi arttıkça madde parametrelerine ait tahminlerinin azaldığı bulgusuna ulaşılmıştır. "*Test uzunluğu 20*madde ayırt edicilik düzeyi düşük*" ile "*Test uzunluğu 40*madde ayırt edicilik düzeyi düşük*" koşullarına ilişkin madde parametrelerine ait "hataların ortalama karekökü" değerlerinin ise madde ayırt edicilik düzeyinin yüksek olduğu koşula kıyasla monoton olmayan artış ve azalışlara sahip olduğu ve bu değerlerin nispeten daha yüksek olduğu görülmektedir.

8 nitelikli Q-matrisin kullanıldığı koşulda SISM'e ait dört madde parametresine ilişkin "hataların ortalama karekökü" değerleri incelendiğinde madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu değerlerin yaklaşık olarak ,19 ile ,30; madde ayırt edicilik düzeyinin düşük olduğu koşulda yaklaşık olarak ,24 ile ,30 arasında değiştiği görülmekle birlikte 6 nitelikli Q-matrisin kullanıldığı koşula benzer şekilde madde ayırt edicilik düzeyinin yüksek olduğu koşulda test uzunluğu ve örneklem büyüklüğü arttıkça bu değerlerin azaldığı görülmektedir. Madde ayırt edicilik düzeyinin düşük ve madde ayırt edicilik düzeyinin yüksek olduğu koşullar karşılaştırıldığında madde ayırt edicilik düzeyi faktörünün parametre tahminlerini etkilediği yani madde ayırt edicilik düzeyi arttıkça madde parametrelerine ait tahminlerinin azaldığı bulgusuna ulaşılmıştır. Madde ayırt edicilik düzeyinin düşük olduğu koşul incelendiğinde ise test uzunluğunu 40 maddeden 60 maddeye ve örneklem büyüklüğünü 1500'den 3000'e arttırmanın dört madde parametresi için "hataların ortalama karekökü" değerlerinde ek bir azalış sağlamadığı görülmektedir.

10 nitelikli Q-matrisin kullanıldığı koşulda SISM'e ait dört madde parametresine ilişkin "hataların ortalama karekökü" değerleri incelendiğinde madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu değerlerin yaklaşık olarak ,36 ile ,43; madde ayırt edicilik düzeyinin düşük olduğu koşulda yaklaşık olarak ,36 ile ,47 arasında değiştiği görülmektedir. Bu matris, 6 ve 8 nitelikli Q-matrislerin kullanıldığı koşullarda olduğu gibi örneklem büyüklüğü artışıyla birlikte madde parametre tahminlerine ait değerlerde azalmalar olduğunu gösterse de madde ayırt edicilik düzeyinin yüksek ve düşük olduğu 20 maddelik test uzunluğunda küçük örneklemelerde monoton olmayan artış ya da azalışların da olduğu görülmektedir. Benzer bir durum 40 maddelik test uzunluğu için de geçerlidir. Madde ayırt edicilik düzeyinin düşük, test uzunluğunun 40 madde olduğu koşulda 250, 500 ve 1500 örneklem büyüklüklerinde dört madde parametresi aynı değerleri alırken, 3000 örneklem büyüklüğünde ise madde parametre tahminlerine ait değerlerde azalmalar olduğu görülmektedir. Diğer koşulların sabit tutulduğu test uzunluğu faktörü incelendiğinde ise değişimlenen 20, 40 ve 60 maddelik koşullar için test uzunluğu arttıkça madde parametrelerine ait "hataların ortalama karekökü" değerleri genel olarak azalmıştır. Ayrıca bu değerlerin daha yüksek madde ayırt edicilik düzeyiyle azalırken madde ayırt edicilik düzeyinin azaldığı koşulda artış gösterdiği bulgusuna ulaşılmıştır.

Özetle ölçülen nitelik sayısının madde parametrelerine ait tahminlerin doğruluğu üzerindeki etkisini incelemek adına, "ortalama mutlak sapma" ve "hataların ortalama karekökü" değerleri incelendiğinde 6, 8 ve 10 nitelik için diğer koşullardan bağımsız olarak, ölçülen nitelik sayısının artmasıyla bu değerlerin arttığı bulgusuna ulaşılmıştır. Bu bulgulara benzer şekilde üç matrisin kullanıldığı koşullarda madde ayırt edicilik düzeyi düştükçe madde parametrelerinden elde edilen "ortalama mutlak sapma" ve "hataların ortalama karekökü" değerlerinin arttığı görülmektedir.

Çalışmanın birinci ve ikinci alt problemi kapsamında "mutlak sapma", "ortalama mutlak sapma" ve "hataların ortalama karekökü" değerleri incelendikten sonra 6, 8 ve 10 nitelikli Q-matrisler için faktörlerde değişimlenen her bir koşulda dört madde parametresine ait tüm maddeler arasındaki ortalama standart hata değerleri EK 6'da sunulmuştur.

6 nitelikten oluşan Q-matrisin kullanıldığı koşulda örneklem büyüklüğü ve test uzunluğu faktörlerinde dört madde parametresine ait tüm maddeler arasındaki ortalama standart hata değerleri, madde ayırt edicilik düzeyinin düşük olduğu koşulda ,01 ile ,42; madde ayırt edicilik düzeyinin yüksek olduğu koşulda ise ,01 ile ,34 arasında değişmektedir. Dört madde

parametresi ayrı ayrı incelendiğinde h_j madde parametresinin standart hata değerlerinin örneklem büyüklüğünün 250 olduğu koşulda diğer koşullara göre daha büyük olduğu ve madde ayırt edicilik düzeyinin düşük olduğu koşulda test uzunluğundaki artışla birlikte standart hata değerlerinin *-test uzunluğunun 40 olduğu koşul hariç-* arttığı görülmektedir. Madde ayırt edicilik düzeyinin yüksek olduğu koşulda test uzunluğu arttıkça bu parametreye ait standart hata değerlerinde madde ayırt edicilik düzeyinin düşük olduğu koşul kadar olmasa da, artışların olduğu bulgusuna ulaşılmıştır. h_j madde parametresine ait en büyük standart hata değeri “örneklem büyüklüğü 250*madde ayırt edicilik düzeyi düşük*test uzunluğu 60 (,42)” koşulu için elde edilmiştir. Örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda h_j madde parametresine ait standart hata değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,01 ile ,02; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,02 ile ,03 arasında değiştiği görülmüştür. Nispeten daha büyük olan bu örneklemelerde, test uzunluğundaki artışla birlikte h_j madde parametresine ait standart hata değerleri bazı koşullarda benzer kalsa da bu değerlerin ,01’lik azalışlara karşılık geldiği bulgusuna ulaşılmıştır. g_j , ω_j ve ϵ_j madde parametreleri incelendiğinde standart hata değerlerinin, örneklem büyüklüğünün 1500 ve 3000 olduğu koşullara kıyasla, örneklem büyüklüğünün 250 ve 500 olduğu koşullarda nispeten daha büyük olduğu görülmektedir. Bu durum özellikle küçük örneklem için test uzunluğundaki artışla birlikte ortaya çıkmaktadır. Her üç madde parametresi için en büyük standart hata değeri “örneklem büyüklüğü 250*madde ayırt edicilik düzeyi düşük*test uzunluğu 60 (g_j : ,25; ω_j : ,13 ve ϵ_j : ,07)” koşulu için elde edilmiştir. Standart hata değerleri örneklem büyüklüğünün 1500 ve 3000’e çıkarıldığı koşullarda azalma eğilimi göstermiştir. Bu değerler madde ayırt edicilik düzeyinin yüksek ve düşük olduğu koşullarda ,01 ile ,04 arasında değişmiştir. Test uzunluğundaki artışın, nispeten daha büyük olan iki örneklem büyüklüğü koşullarında h_j madde parametresinde olduğu gibi g_j , ω_j ve ϵ_j madde parametrelerine ait standart hata değerlerinde ,01’lik azalışlara karşılık geldiği bulgusuna ulaşılmıştır.

8 nitelikten oluşan Q-matrisin kullanıldığı koşulda örneklem büyüklüğü ve test uzunluğu faktörlerinde dört madde parametresine ait tüm maddeler arasındaki ortalama standart hata değerleri madde ayırt edicilik düzeyinin düşük olduğu koşulda ,02 ile ,62; madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,01 ile ,34 arasında değişmektedir. Dört madde parametresi ayrı ayrı incelendiğinde bu parametrelerden h_j madde parametresinin standart hata değerlerinin örneklem büyüklüğünün 250 olduğu koşulda daha büyük olduğu ve özellikle madde ayırt edicilik düzeyinin düşük olduğu koşullarda test uzunluğundaki artışla

birlikte standart hata deęerlerinin arttıęı grlmektedir. Madde ayırt edicilik dzeyinin yksek olduęu kořulda da bu parametreye ait standart hata deęerlerinde test uzunluęu arttıka madde ayırt edicilik dzeyinin dřk olduęu kořul kadar olmasa da artıřların olduęu bulgusuna ulařılmıřtır. h_j madde parametresine ait en byk standart hata deęeri “*rneklem byklę 250*madde ayırt edicilik dzeyi dřk*test uzunluęu 60 (,62)*” kořulu iin elde edilmiřtir. rneklem byklęnn 1500 ve 3000 olduęu kořullarda bu parametreye ait standart hata deęerlerinin madde ayırt edicilik dzeyinin yksek olduęu kořulda ,01 ile ,02; madde ayırt edicilik dzeyinin dřk olduęu kořulda ise ,02 ile ,04 arasında deęiřtięi grlmektedir. g_j , ω_j ve ϵ_j madde parametreleri incelendięinde rneklem byklęnn 250 olduęu kořulda dięer rneklem byklklerine kıyasla standart hata deęerlerinin nispeten daha byk olduęu grlmektedir. Bu durum zellikle madde ayırt edicilik dzeyinin dřk olduęu kořulda, test uzunluęundaki artıřla birlikte ortaya ıkmaktadır. Her  madde parametresi iin en byk standart hata deęeri “*rneklem byklę 250*madde ayırt edicilik dzeyi dřk*test uzunluęu 60 (g_j : ,14; ω_j : ,12 ve ϵ_j : ,35)*” kořulu iin elde edilmiřtir. Standart hata deęerleri rneklem byklęnn 1500 ve 3000’e ıkarıldıęı kořullarda azalma eęilimi gstermektedir. Bu deęerler madde ayırt edicilik dzeyinin yksek olduęu kořulda ,01 ile ,04; madde ayırt edicilik dzeyinin dřk olduęu kořulda ,01 ile ,07 arasında deęiřmektedir. Test uzunluęundaki artıřın, nispeten daha byk olan iki rneklem byklę kořullarında h_j madde parametresinde olduęu gibi g_j , ω_j ve ϵ_j madde parametrelerine ait standart hata deęerlerinde ,01’lik azalıřlara karřılık geldięi bulgusuna ulařılmıřtır.

10 nitelikten oluřan Q-matrisin kullanıldıęı kořulda rneklem byklę ve test uzunluęu faktrlerinde drt madde parametresine ait tm maddeler arasındaki ortalama standart hata deęerleri madde ayırt edicilik dzeyinin dřk olduęu kořulda ,01 ile ,85; madde ayırt edicilik dzeyinin yksek olduęu kořulda ,01 ile ,68 arasında deęiřmektedir. Drt madde parametresi ayrı ayrı incelendięinde bu parametrelerden h_j madde parametresine ait standart hata deęerlerinin rneklem byklęnn 250 olduęu kořulda daha byk olduęu, zellikle madde ayırt edicilik dzeyinin dřk olduęu kořulda test uzunluęundaki artıřla birlikte standart hata deęerlerinin arttıęı grlmektedir. Madde ayırt edicilik dzeyinin yksek olduęu kořulda da bu parametreye ait standart hata deęerlerinde test uzunluęu arttıka madde ayırt edicilik dzeyinin dřk olduęu kořul kadar olmasa da artıřların olduęu bulgusuna ulařılmıřtır. h_j madde parametresine ait en byk standart hata deęeri “*rneklem byklę 250*madde ayırt edicilik dzeyi dřk*test uzunluęu 60 (,85)*” kořulu iin elde

edilmiştir. Örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda h_j parametresine ait standart hata değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,01 ile ,03; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,02 ile ,04 arasında değiştiği görülmektedir. g_j , ω_j ve ϵ_j madde parametreleri incelendiğinde standart hata değerlerinin örneklem büyüklüğünün 250 ve 500 olduğu koşullarda diğer örneklem büyüklüklerine kıyasla nispeten daha büyük olduğu görülmektedir. Bu durum özellikle madde ayırt edicilik düzeyinin düşük olduğu koşulda test uzunluğundaki artışla birlikte ortaya çıkmaktadır. Her üç madde parametresi için en büyük standart hata değeri “*örneklem büyüklüğü 250*madde ayırt edicilik düzeyi düşük*test uzunluğu 60* ($g_j: ,18$; $w_j: ,23$ ve $\epsilon_j: ,11$)” koşulu için elde edilmiştir. Standart hata değerleri örneklem büyüklüğünün 1500 ve 3000’e çıkarıldığı koşullarda azalma eğilimi göstermektedir. Bu değerlerin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,01 ile ,04; madde ayırt edicilik düzeyinin düşük olduğu koşulda ,01 ile ,04 arasında değiştiği görülmektedir. Test uzunluğundaki artışın, nispeten daha büyük olan iki örneklem büyüklüğü koşullarında dört madde parametresine ait standart hata değerlerinde ,01’lik azalışlara karşılık geldiği bulgusuna ulaşılmıştır.

Özetle 10 nitelikli Q-matrisin kullanıldığı koşuldan elde edilen bulgular, 6 ve 8 nitelikli Q-matrisin kullanıldığı koşullarla benzerlik göstermektedir. Madde ayırt edicilik düzeyi, örneklem büyüklüğü ve test uzunluğundaki artış üç Q-matrisin kullanıldığı koşullar için de örneklem büyüklüğünün 250 olduğu koşul haricinde daha küçük standart hata değerlerini ortaya koymuştur. Burada dikkat edilmesi gereken nokta ölçülen nitelik sayısı arttıkça benzer koşullarda, ortalama standart hata değerlerinin artma eğiliminde olduğudur.

4.3. Üçüncü Alt Probleme İlişkin Bulgular

3. Becerileri ölçen DINA, kavram yanlışlarını ölçen BUG-DINO ve hem becerileri hem de kavram yanlışlarını ölçen SISM'in sınıflama doğruluğu değerleri,
 - 3.1. Örneklem büyüklüğü (250, 500, 1500 ve 3000),
 - 3.2. Test uzunluğu (20, 40 ve 60 madde),
 - 3.3. Madde ayırt edicilik düzeyi (yüksek ayırt edicilik ve düşük ayırt edicilik) ve
 - 3.4. Ölçülen nitelik sayısına (6, 8 ve 10 nitelik) göre nasıl değişmektedir?

Üçüncü alt problem kapsamında elde edilen sınıflama doğruluğu değerleri örneklem büyüklüğü, test uzunluğu, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısı olmak üzere dört faktörün her bir koşulu için 100 iterasyonun analizinden elde edilmiş ve SISM ile birlikte diğer karşılaştırma modelleri (Beceri modeli (DINA) ve kavram yanlışlığı modeli (BUG-DINO)) için sunulmuştur.

Tablo 15-23 6, 8 ve 10 nitelikli Q-matrislerin kullanıldığı koşullarda yüksek ve düşük ayırt ediciliğe sahip 20, 40 ve 60 maddelik bir testin farklı örneklem büyüklüğü koşullarından elde edilen sınıflama doğruluğu değerlerini sunmaktadır. Tablolarda yer alan “nitelik doğru sınıflama oranı” her bir beceri ya da kavram yanlışlığı düzeyinde doğru sınıflama oranını; “ortalama nitelik doğru sınıflama oranı” ise beceri, kavram yanlışlığı ya da her ikisinin ortalamasını veren sınıflama oranını ve “örüntü doğru sınıflama oranı” da niteliklerle ilişkili bilişsel tanı profiline ait tahminlerin doğruluğunu özetlemektedir. Tablolarda dikkat edilmesi gereken nokta, SISM için hem beceriler hem de kavram yanlışlıklarına ait sınıflama doğruluğu değerleri sunulurken; DINA için beceriler ve BUG-DINO için kavram yanlışlıklarına ait sınıflama doğruluğu değerlerinin sunulmuş olmasıdır. Ayrıca çalışmada yer alan üç model tarafından kullanılan Q-matrislerdeki nitelik sayısı farklılaştığından tüm nitelikler düzeyinde bu modeller, aynı bilişsel örüntülere sahip olamayacağından DINA ve BUG-DINO modelleri için bilişsel tanı profiline ilişkin “örüntü doğru sınıflama oranı” değerleri hesaplanamamıştır. Bu değerler SISM'in farklı faktörlerde değişimlenen koşullarına ilişkin elde edilerek karşılaştırmalar bu kapsamda özetlenmiştir. Model karşılaştırmaları öncelikle faktörler altında değişimlenen koşullara ilişkin SISM'in sınıflama doğruluğu performansını sunmuştur. Daha sonra modelin, becerileri sınıflama performansı yönünden DINA; kavram yanlışlıklarını sınıflama performansı yönünden BUG-DINO ile farklı koşullarda karşılaştırıldığı bulgular özetlenmiştir.

Tablo 15

6 Nitelikli 20 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri

Nitelik Doğru Sınıflama Oranı										
	Örneklem büyüklüğü	Modeller	B_1	B_2	B_3	K_1	K_2	K_3	Ortalama Nitelik Doğru Sınıflama Oranı	Örüntü Doğru Sınıflama Oranı
Madde Ayırt Edicilik Düzeyi Yüksek (20 Madde)	250	DINA	,78	,75	,76	-	-	-	,76	-
		BUG-DINO	-	-	-	,93	,91	,90	,91	-
		SISM	,73	,83	,87	,94	,93	,90	,87	,47
	500	DINA	,78	,75	,76	-	-	-	,76	-
		BUG-DINO	-	-	-	,94	,93	,90	,92	-
		SISM	,73	,85	,88	,94	,93	,91	,87	,48
	1500	DINA	,78	,74	,76	-	-	-	,76	-
		BUG-DINO	-	-	-	,93	,92	,90	,92	-
		SISM	,70	,86	,89	,94	,94	,91	,87	,49
3000	DINA	,78	,74	,77	-	-	-	,76	-	
	BUG-DINO	-	-	-	,95	,93	,91	,93	-	
	SISM	,70	,87	,90	,95	,94	,91	,88	,50	
Madde Ayırt Edicilik Düzeyi Düşük (20 Madde)	250	DINA	,65	,58	,61	-	-	-	,62	-
		BUG-DINO	-	-	-	,88	,85	,83	,85	-
		SISM	,67	,70	,73	,89	,84	,82	,77	,25
	500	DINA	,65	,59	,61	-	-	-	,62	-
		BUG-DINO	-	-	-	,89	,85	,86	,86	-
		SISM	,65	,71	,77	,90	,86	,83	,79	,27
	1500	DINA	,65	,62	,63	-	-	-	,63	-
		BUG-DINO	-	-	-	,90	,86	,83	,86	-
		SISM	,64	,69	,81	,91	,87	,84	,79	,29
3000	DINA	,67	,64	,66	-	-	-	,66	-	
	BUG-DINO	-	-	-	,90	,87	,84	,87	-	
	SISM	,64	,68	,82	,91	,88	,85	,80	,29	

B_1, B_2, B_3 sırasıyla 1., 2. ve 3. becerileri; K_1, K_2, K_3 sırasıyla 1., 2. ve 3. kavram yanılgılarını temsil etmektedir.

6 nitelik içeren Q-matris kullanıldığında 20 maddeden oluşan bir test için Tablo 15'te yer alan SISM'e ait beceri ve kavram yanılgısı düzeylerinde ($B_1, B_2, B_3; K_1, K_2, K_3$) sınıflama doğruluğu değerleri incelendiğinde örneklem büyüklüğünden bağımsız olarak her bir beceri ve kavram yanılgısı düzeyinde "nitelik doğru sınıflama oranı" değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,70 ile ,95; madde ayırt edicilik düzeyinin düşük olduğu koşulda ,64 ile ,91 arasında değiştiği görülmektedir. "Ortalama nitelik doğru sınıflama oranı" değerleri incelendiğinde ise madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu değerlerin ,87 ile ,88; "örüntü doğru sınıflama oranı" değerlerinin ise ,47 ile ,50 arasında değiştiği görülmektedir. Madde ayırt edicilik düzeyinin düşük olduğu koşulda ise

“ortalama nitelik doğru sınıflama oranı” değerlerinin ,77 ile ,80; “örüntü doğru sınıflama oranı” değerlerinin ise ,25 ile ,29 arasında değiştiği görülmektedir. Bu durum madde ayırt edicilik düzeyindeki artışların daha yüksek sınıflama doğruluğu değerleri verdiği bir göstergesidir. Tabloda örneklem büyüklüğü artışıyla birlikte madde ayırt edicilik düzeyinin yüksek ve düşük olduğu koşullarda SISM’e ait “nitelik doğru sınıflama oranı” ve “örüntü doğru sınıflama oranı” değerlerinin genel olarak arttığı; model kapsamında “ortalama nitelik doğru sınıflama oranı” değerlerinin ise 250, 500 ve 1500 örneklem büyüklüğü koşullarında aynı kaldığı ancak 3000 örneklem büyüklüğü koşulunda arttığı görülmektedir. Genel olarak değerlendirildiğinde madde ayırt edicilik düzeyinin yüksek ve test uzunluğunun 20 madde olduğu koşullarda küçük bir örneklem büyüklüğüne rağmen SISM’in “ortalama nitelik doğru sınıflama oranı” değerlerinin *-B₁ becerisi hariç-* ,80 üzerinde değerler aldığı bulgusuna ulaşılmıştır.

SISM ve DINA modellerinin beceri düzeyinde (B_1, B_2, B_3) sınıflama doğruluğu performanslarını yansıtan “nitelik doğru sınıflama oranı” değerleri karşılaştırıldığında örneklem büyüklüğünden bağımsız olarak madde ayırt edicilik düzeyinin yüksek ve düşük olduğu koşullarda SISM’in *-B₁ becerisi hariç-* diğer becerilerde DINA modelden daha yüksek sınıflama doğruluğu değerlerine sahip olduğu bulgusuna ulaşılmıştır. DINA’nın sınıflama doğruluğu performansının özellikle madde ayırt edicilik düzeyinin düşük olduğu koşulda gittikçe azaldığı ancak örneklem büyüklüğündeki artışlarla birlikte sınıflama doğruluğu performansında kısmen de olsa artışların olduğu bulgusuna ulaşılmıştır.

Kavram yanılgılarını ölçen BUG-DINO modeliyle SISM’in kavram yanılgıları düzeyinde (K_1, K_2, K_3) sınıflama doğruluğu performanslarını yansıtan “nitelik doğru sınıflama oranı” değerleri karşılaştırıldığında örneklem büyüklüğünden bağımsız olarak madde ayırt edicilik düzeyinin yüksek ve düşük olduğu koşullarda BUG-DINO’nun SISM kadar iyi bir sınıflama doğruluğu performansı sergilediği görülmekle birlikte SISM’in az da olsa sınıflama doğruluğu performansının daha yüksek olduğu bulgusuna ulaşılmıştır. Ayrıca BUG-DINO’nun çok küçük örneklemelerde dahi kavram yanılgılarını ölçmede yüksek sınıflama doğruluğu değerleri aldığı ve örneklem büyüklüğündeki artışlarla birlikte sınıflama doğruluğu performansında kısmen de olsa artışların olduğu görülmektedir.

Tablo 16

6 Nitelikli 40 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri

		Nitelik Doğru Sınıflama Oranı								
Örneklem büyüklüğü	Modeller	B_1	B_2	B_3	K_1	K_2	K_3	Ortalama Nitelik Doğru Sınıflama Oranı	Örüntü Doğru Sınıflama Oranı	
Madde Ayırt Edicilik Düzeyi Yüksek (40 Madde)	250	<i>DINA</i>	,83	,85	,88	-	-	-	,85	-
		<i>BUG-DINO</i>	-	-	-	,97	,97	,96	,97	-
		<i>SISM</i>	,81	,92	,90	,97	,98	,97	,92	,63
	500	<i>DINA</i>	,85	,87	,88	-	-	-	,87	-
		<i>BUG-DINO</i>	-	-	-	,98	,97	,97	,97	-
		<i>SISM</i>	,78	,94	,94	,98	,98	,97	,93	,66
	1500	<i>DINA</i>	,85	,87	,89	-	-	-	,87	-
		<i>BUG-DINO</i>	-	-	-	,98	,98	,97	,98	-
		<i>SISM</i>	,78	,96	,96	,97	,98	,97	,94	,68
	3000	<i>DINA</i>	,86	,88	,89	-	-	-	,90	-
		<i>BUG-DINO</i>	-	-	-	,98	,98	,97	,98	-
		<i>SISM</i>	,79	,96	,96	,98	,98	,97	,94	,68
Madde Ayırt Edicilik Düzeyi Düşük (40 Madde)	250	<i>DINA</i>	,73	,65	,70	-	-	-	,69	-
		<i>BUG-DINO</i>	-	-	-	,95	,92	,90	,92	-
		<i>SISM</i>	,77	,73	,65	,94	,92	,91	,82	,33
	500	<i>DINA</i>	,73	,65	,70	-	-	-	,69	-
		<i>BUG-DINO</i>	-	-	-	,95	,93	,91	,93	-
		<i>SISM</i>	,77	,73	,66	,95	,93	,92	,83	,33
	1500	<i>DINA</i>	,74	,67	,70	-	-	-	,70	-
		<i>BUG-DINO</i>	-	-	-	,96	,93	,91	,93	-
		<i>SISM</i>	,78	,71	,66	,95	,93	,93	,83	,34
	3000	<i>DINA</i>	,73	,67	,71	-	-	-	,70	-
		<i>BUG-DINO</i>	-	-	-	,96	,93	,91	,93	-
		<i>SISM</i>	,79	,71	,66	,95	,93	,93	,83	,34

B_1, B_2, B_3 sırasıyla 1., 2. ve 3. becerileri; K_1, K_2, K_3 sırasıyla 1., 2. ve 3. kavram yanlışlarını temsil etmektedir.

Tablo 16’da yer alan 40 maddeden oluşan bir test için SISM’e ait beceri ve kavram yanlışlığı düzeylerinde ($B_1, B_2, B_3; K_1, K_2, K_3$) sınıflama doğruluğu değerleri incelendiğinde, örneklem büyüklüğünden bağımsız olarak her bir beceri ve kavram yanlışlığı düzeyinde “nitelik doğru sınıflama oranı” değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,79 ile ,98; madde ayırt edicilik düzeyinin düşük olduğu koşulda ,65 ile ,95 arasında değiştiği görülmektedir. “Ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde ise madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu değerlerin ,92 ile ,94; “örüntü doğru sınıflama oranı” değerlerinin ise ,63 ile ,68 arasında değiştiği görülmektedir. Madde ayırt

edicilik düzeyinin düşük olduğu koşullarda “ortalama nitelik doğru sınıflama oranı” değerlerinin ,82 ile ,83; “örüntü doğru sınıflama oranı” değerlerinin ise ,33 ile ,34 arasında değiştiği görülmektedir. Bu durum madde ayırt edicilik düzeyindeki artışların daha yüksek sınıflama doğruluğu değerleri verdiğinin bir göstergesidir. Tablo 16’da örneklem büyüklüğü artışıyla birlikte madde ayırt edicilik düzeyinin yüksek olduğu koşullarda SISM’e ait “nitelik doğru sınıflama oranı” ve “örüntü doğru sınıflama oranı” değerlerinin 250, 500 ve 1500 örneklem büyüklüğü koşullarında arttığı görülmektedir ancak örneklem büyüklüğünü 1500’den 3000’e çıkarmak modelin sınıflama doğruluğu performansına ek bir katkı getirmemiştir. Genel olarak değerlendirildiğinde madde ayırt edicilik düzeyinin yüksek ve test uzunluğunun 20 maddeden 40 maddeye çıkarıldığı koşullarda SISM’in – *Örneklem büyüklüğünün 500, 1500 ve 3000 olduğu koşullarda B₁ becerisine ait “nitelik doğru sınıflama oranı” değerleri hariç-* küçük bir örneklem büyüklüğüne rağmen sınıflama doğruluğu değerlerinin daha da artarak ,80 üzerinde değerler aldığı bulgusuna ulaşılmıştır.

SISM ve DINA modellerinin beceri düzeyinde (B₁, B₂, B₃) sınıflama doğruluğu performanslarını yansıtan “nitelik doğru sınıflama oranı” değerleri test uzunluğunun 20 madde olduğu koşulla karşılaştırıldığında, 40 maddelik test uzunluğunda örneklem büyüklüğünden bağımsız olarak DINA’nın sınıflama doğruluğu değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,83 ile ,89; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,65 ile ,74 arasında değiştiği görülmektedir. Test uzunluğu faktöründe özellikle madde ayırt edicilik düzeyinin yüksek olduğu koşul dikkate alındığında, sınıflama doğruluğu değerlerindeki artışlar DINA modelin performansında artışların olduğunu göstermektedir. Ayrıca DINA modelde, örneklem büyüklüğü artışı sınıflama doğruluğu performansında artışlar sağlamaktadır. Ancak tablo incelendiğinde madde ayırt edicilik düzeyinin yüksek olduğu koşulda SISM’in -B₁ becerisi hariç- DINA modelden daha yüksek sınıflama doğruluğu değerlerine sahip olduğu görülürken, madde ayırt edicilik düzeyinin düşük olduğu koşulda SISM’in -B₃ becerisi hariç- DINA modelden daha yüksek sınıflama doğruluğu değerlerine sahip olduğu görülmektedir.

Kavram yanılgılarını ölçen BUG-DINO modeli ile SISM’in kavram yanılgıları düzeyinde (K₁, K₂, K₃) sınıflama doğruluğu performanslarını yansıtan “nitelik doğru sınıflama oranı” değerleri test uzunluğunun 20 madde olduğu koşulla karşılaştırıldığında, 40 maddelik test uzunluğunda örneklem büyüklüğünden bağımsız olarak madde ayırt edicilik düzeyinin yüksek ve düşük olduğu koşullarda BUG-DINO’nun sınıflama doğruluğu değerlerinin daha da arttığı ve SISM kadar iyi bir sınıflama doğruluğu performansı

sergilediği görülmektedir. Ancak SISM'in çok az da olsa BUG-DINO modele kıyasla sınıflama doğruluğu performansının daha yüksek olduğu bulgusuna ulaşılmıştır. Ayrıca BUG-DINO'nun çok küçük örneklerde dahi kavram yanlışlarını ölçmede yüksek sınıflama doğruluğu değerleri aldığı ve örneklem büyüklüğündeki artışlarla birlikte sınıflama doğruluğu performansında kısmen de olsa artışların olduğu görülmektedir.

Tablo 17

6 Nitelikli 60 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri

		Nitelik Doğru Sınıflama Oranı								
Örneklem büyüklüğü	Modeller	B_1	B_2	B_3	K_1	K_2	K_3	Ortalama Nitelik Doğru Sınıflama Oranı	Örüntü Doğru Sınıflama Oranı	
Madde Ayrırt Edicilik Düzeyi Yüksek (60 Madde)	250	<i>DINA</i>	,87	,93	,92	-	-	-	,91	
		<i>BUG-DINO</i>	-	-	-	,98	,98	,97	,98	--
		<i>SISM</i>	,86	,96	,91	,99	,99	,98	,95	,72
	500	<i>DINA</i>	,87	,93	,93	-	-	-	,91	-
		<i>BUG-DINO</i>	-	-	-	,99	,99	,98	,99	-
		<i>SISM</i>	,83	,97	,94	,99	,99	,98	,95	,73
	1500	<i>DINA</i>	,88	,93	,93	-	-	-	,92	-
		<i>BUG-DINO</i>	-	-	-	,99	,99	,98	,99	-
		<i>SISM</i>	,81	,98	,97	,99	,99	,98	,95	,74
	3000	<i>DINA</i>	,88	,93	,93	-	-	-	,92	-
		<i>BUG-DINO</i>	-	-	-	,99	,99	,98	,99	-
		<i>SISM</i>	,81	,98	,98	,99	,99	,98	,95	,75
Madde Ayrırt Edicilik Düzeyi Düşük (60 Madde)	250	<i>DINA</i>	,80	,68	,77	-	-	-	,75	-
		<i>BUG-DINO</i>	-	-	-	,96	,95	,94	,95	--
		<i>SISM</i>	,83	,73	,60	,95	,96	,96	,84	,34
	500	<i>DINA</i>	,80	,69	,76	-	-	-	,75	-
		<i>BUG-DINO</i>	-	-	-	,96	,96	,94	,96	-
		<i>SISM</i>	,84	,70	,64	,95	,96	,96	,84	,37
	1500	<i>DINA</i>	,80	,71	,77	-	-	-	,76	-
		<i>BUG-DINO</i>	-	-	-	,97	,96	,95	,96	-
		<i>SISM</i>	,84	,70	,64	,97	,96	,96	,84	,37
	3000	<i>DINA</i>	,80	,72	,77	-	-	-	,76	-
		<i>BUG-DINO</i>	-	-	-	,97	,96	,95	,96	-
		<i>SISM</i>	,85	,70	,70	,97	,96	,96	,86	,40

B_1, B_2, B_3 sırasıyla 1., 2. ve 3. becerileri; K_1, K_2, K_3 sırasıyla 1., 2. ve 3. kavram yanlışlarını temsil etmektedir.

Tablo 17’de yer alan 60 maddeden oluşan bir test için SISM’e ait beceri ve kavram yanılıgısı düzeylerinde ($B_1, B_2, B_3; K_1, K_2, K_3$) sınıflama dođruluđu deđerleri incelendiđinde, örneklem büyüklüğünden bađımsız olarak her bir beceri ve kavram yanılıgısı düzeyinde “nitelik dođru sınıflama oranı” deđerlerinin madde ayırt edicilik düzeyinin yüksek olduđu kođuşulda ,81 ile ,99; madde ayırt edicilik düzeyinin düşük olduđu kođuşulda ise ,60 ile ,97 arasında deđiđuştiđi görülmektedir. “Ortalama nitelik dođru sınıflama oranı” deđerleri incelendiđinde ise madde ayırt edicilik düzeyinin yüksek olduđu kođuşulda bu deđerlerin ,95 olduđu; madde ayırt edicilik düzeyinin düşük olduđu kođuşulda ise bu deđerlerinin ,84 ile ,86 arasında deđiđuştiđi görülmektedir. “Örüntü dođru sınıflama oranı” deđerlerinin madde ayırt edicilik düzeyinin yüksek olduđu kođuşulda ,72 ile ,75; madde ayırt edicilik düzeyinin düşük olduđu kođuşulda ise ,34 ile ,40 arasında deđiđuştiđi görülmektedir. Bu durum madde ayırt edicilik düzeyindeki artışların daha yüksek sınıflama dođruluđu deđerleri verdiđinin bir göstergesidir. Madde ayırt edicilik düzeyinin yüksek olduđu kođuşulda SISM’in “ortalama nitelik dođru sınıflama oranı” deđerleri incelendiđinde sınıflama dođruluđu deđerlerinin tüm kođuşullarda aynı olduđu yani örneklem büyüklüğünün sınıflama dođruluđu deđerlerinde ek bir artış sađlamadıđı bulgusuna ulaşılmıđutır. Madde ayırt edicilik düzeyinin düşük olduđu kođuşulda incelendiđinde SISM’e ait sınıflama dođruluđu deđerleri madde ayırt edicilik düzeyinin yüksek olduđu kođuşula göre azalmıđutır. Madde ayırt edicilik düzeyinin düşük olduđu kođuşulda örneklem büyüklüđu artışıyla birlikte sınıflama dođruluđu deđerleri 250, 500 ve 1500 örneklem büyüklüklerinde aynı kalırken, 3000 örneklem büyüklüđu kođuşulunda artmıđutır.

SISM ve DINA modellerinin beceri düzeyinde (B_1, B_2, B_3) sınıflama dođruluđu performanslarını yansıtan “nitelik dođru sınıflama oranı” deđerleri test uzunluđunun 20 ve 40 madde olduđu kođuşullarla karşılaştıırıldıđında, 60 maddelik test uzunluđunda örneklem büyüklüğünden bađımsız madde ayırt edicilik düzeyinin yüksek olduđu kođuşulda DINA’nın sınıflama performanslarının ,87 ile ,93 arasında; madde ayırt edicilik düzeyinin düşük olduđu kođuşulda ise ,68 ile ,80 arasında deđiđuştiđi görülmektedir. Test uzunluđu faktöründeki artışlar, özellikle madde ayırt edicilik düzeyinin yüksek olduđu kođuşul dikkate alındıđında DINA modelin performansında artışların olduđunu göstermektedir. Ayrıca DINA modelde “nitelik dođru sınıflama oranı” deđerlerinin, 250 ve 500 örneklem büyüklüklerinde artışlar göstererek sınıflama performansında artışlar sađladıđı ancak 1500 ve 3000 örneklem büyüklüklerinde hem “nitelik dođru sınıflama oranı” deđerlerinin hem de “ortalama nitelik dođru sınıflama oranı” deđerlerinin aynı deđerleri almasıyla örneklem büyüklüđu faktörünün sınıflama dođruluđu deđerlerinde ek bir artış sađlamadıđı bulgusuna ulaşılmıđutır.

Test uzunluęu faktöründeki artışla birlikte DINA model performansı artsa da, örneklem büyüklüęü ve yüksek madde ayırt edicilik düzeyi koşullarında SISM'in, $-B_1$ becerisi hariç- DINA modelden daha yüksek sınıflama doğruluęu değerlerine sahip olduęu görülürken; örneklem büyüklüęü ve düşük madde ayırt edicilik düzeyi koşullarında ise SISM'in, $-B_3$ becerisi hariç- DINA modelden daha yüksek sınıflama doğruluęu değerlerine sahip olduęu görülmektedir.

Kavram yanılgılarını ölçen BUG-DINO modelinin kavram yanılgıları düzeyinde (K_1, K_2, K_3) sınıflama doğruluęu performanslarını yansıtan “nitelik doğru sınıflama oranı” değerleri test uzunluęunun 20 ve 40 madde olduęu koşullarla karşılaştırıldığında, 60 maddelik test uzunluęunda örneklem büyüklüęünden bağımsız olarak madde ayırt edicilik düzeyinin yüksek olduęu koşulda BUG-DINO'nun sınıflama doğruluęu değerlerinin ,97 ile ,99; madde ayırt edicilik düzeyinin düşük olduęu koşulda ise ,94 ile ,97 arasında deęiştiiği görülmektedir. Bu durum madde ayırt edicilik düzeyindeki artışla birlikte sınıflama doğruluęu değerlerinin arttıęının bir göstergesidir. Modele ait “ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde madde ayırt edicilik düzeyinin yüksek olduęu koşulda bu değerler, örneklem büyüklüęü faktörünün sınıflama doğruluęu değerlerinde *-örneklem büyüklüęünün 250'den 500'e çıkarıldığı koşul haricinde-* ek bir artış sağlamadığını göstermektedir. Madde ayırt edicilik düzeyinin yüksek olduęu koşulda BUG-DINO model SISM ile karşılaştırıldığında, tek koşul haricinde (*Örneklem büyüklüęü 250*) iki modelin “nitelik doğru sınıflama oranı” değerlerinin aynı olduęu görülmektedir. Madde ayırt edicilik düzeyinin düşük olduęu koşul dikkate alındığında ise SISM ve BUG-DINO modellerinin artış ve azalışlarla birbirlerine yakın sınıflama doğruluęu performansı sergiledikleri bulgusuna ulaşılmıştır.

Tablo 18

8 Nitelikli 20 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri

Nitelik Doğru Sınıflama Oranı												
Örneklem büyüklüğü	Modeller	B_1	B_2	B_3	B_4	K_1	K_2	K_3	K_4	Ortalama Nitelik Doğru Sınıflama Oranı	Örüntü Doğru Sınıflama Oranı	
Madde Ayırt Ediciliği Yüksek (20 Madde)	250	DINA	,65	,66	,63	,69	-	-	-	-	,66	-
		BUG-DINO	-	-	-	-	,90	,92	,88	,89	,89	-
		SISM	,80	,73	,85	,78	,91	,95	,90	,90	,85	,30
	500	DINA	,68	,68	,66	,71	-	-	-	-	,68	-
		BUG-DINO	-	-	-	-	,91	,93	,89	,89	,90	-
		SISM	,80	,70	,87	,78	,91	,95	,89	,89	,85	,30
	1500	DINA	,71	,71	,70	,75	-	-	-	-	,72	-
		BUG-DINO	-	-	-	-	,91	,94	,90	,89	,91	-
		SISM	,80	,79	,82	,80	,92	,95	,91	,90	,86	,32
	3000	DINA	,74	,75	,74	,79	-	-	-	-	,75	-
		BUG-DINO	-	-	-	-	,91	,95	,90	,90	,92	-
		SISM	,80	,82	,81	,80	,91	,95	,91	,89	,86	,33
Madde Ayırt Ediciliği Düşük (20 Madde)	250	DINA	,57	,55	,54	,56	-	-	-	-	,55	-
		BUG-DINO	-	-	-	-	,83	,86	,81	,81	,83	-
		SISM	,73	,72	,73	,73	,85	,89	,81	,82	,78	,16
	500	DINA	,58	,57	,56	,59	-	-	-	-	,58	-
		BUG-DINO	-	-	-	-	,85	,88	,81	,82	,84	-
		SISM	,72	,71	,73	,73	,85	,91	,81	,82	,78	,16
	1500	DINA	,62	,62	,61	,66	-	-	-	-	,63	-
		BUG-DINO	-	-	-	-	,86	,90	,82	,82	,85	-
		SISM	,71	,67	,75	,71	,86	,91	,81	,82	,78	,16
	3000	DINA	,67	,66	,67	,70	-	-	-	-	,67	-
		BUG-DINO	-	-	-	-	,87	,90	,82	,83	,86	-
		SISM	,70	,67	,77	,71	,87	,92	,83	,83	,78	,16

B_1, B_2, B_3, B_4 sırasıyla 1., 2., 3. ve 4. becerileri; K_1, K_2, K_3, K_4 sırasıyla 1., 2., 3. ve 4. kavram yanlışlarını temsil etmektedir.

Ölçülen nitelik sayısının arttırıldığı 8 nitelik içeren bir Q-matris kullanıldığında 20 maddeden oluşan bir test için Tablo 18’de yer alan SISM’e ait beceri ve kavram yanlışlığı düzeylerinde ($B_1, B_2, B_3, B_4; K_1, K_2, K_3, K_4$) sınıflama doğruluğu değerleri incelendiğinde, örneklem büyüklüğünden bağımsız olarak her bir beceri ve kavram yanlışlığı düzeyinde “nitelik doğru sınıflama oranı” değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,70 ile ,95; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,66 ile ,92 arasında değiştiği görülmektedir. “Ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu değerlerin ,85 ile

,86 arasında deđiřtiđi ve ölçülen nitelik sayısının artmasıyla birlikte örneklem büyüklüğünden bağımsız olarak SISM'in sınıflama doğruluđu performansında azalmalar olduđu görölmektedir. 250 ve 500 örneklem büyüklüđu kořullarında “ortalama nitelik dođru sınıflama oranı” deđerlerinin aynı olduđu ancak örneklem büyüklüğünün 1500 olduđu kořulda modele ait sınıflama doğruluđu performansında artışlar olduđu bulgusuna ulařılmıştır. Örneklem büyüklüğünün 1500'den 3000'e çıkarıldıđı kořulda ise örneklem büyüklüđu faktörünün sınıflama doğruluđu deđerlerinde ek bir artış sađlamadıđı görölmektedir. “Örüntü dođru sınıflama oranı” deđerleri incelendiđinde bu deđerlerin ,30 ile ,33 arasında deđiřtiđi ve ölçülen nitelik sayısının artmasıyla azalarak modele ait sınıflama doğruluđu performansının azaldıđı görölmektedir. Madde ayırt edicilik düzeyinin düşük olduđu kořulda ise “ortalama nitelik dođru sınıflama oranı” deđerlerinin ,78; “örüntü dođru sınıflama oranı” deđerlerinin ,16 olduđu görölmektedir. Tabloda yer alan bu bulgular madde ayırt edicilik düzeyinin düşük olduđu kořulda sınıflama doğruluđu performansının örneklem büyüklüđu faktöründen etkilenmediđini ve tüm kořullar için aynı deđerleri aldıđını göstermektedir. Madde ayırt edicilik düzeyinin yüksek olduđu kořulda olduđu gibi madde ayırt edicilik düzeyinin düşük olduđu kořulda da SISM'in sınıflama doğruluđu performansının ölçülen nitelik sayısının artmasından olumsuz yönde etkilendiđi bulgusuna ulařılmıştır. Genel olarak deđerlendirildiđinde madde ayırt edicilik düzeyinin yüksek ve test uzunluğunun 20 madde olduđu kořulda küçük bir örneklem büyüklüđüne rađmen SISM, 6 nitelikli Q-matrisin kullanıldıđı kořula kıyasla daha düşük sınıflama doğruluđu performansı sergilese de özellikle madde ayırt edicilik düzeyinin yüksek olduđu kořulda “ortalama nitelik dođru sınıflama oranı” deđerleri incelendiđinde modelin ,80 üzerinde deđerler alarak iyi bir sınıflama doğruluđu performansı sergilediđi söylenebilir.

SISM ve DINA modellerinin beceri düzeyinde (B_1, B_2, B_3, B_4) sınıflama doğruluđu performanslarını yansıtan “nitelik dođru sınıflama oranı” deđerleri karřılařtırıldıđında ölçülen nitelik sayısının arttırılarak 8 nitelikli Q-matrisin kullanıldıđı kořulda örneklem büyüklüğünden bağımsız madde ayırt edicilik düzeyinin yüksek olduđu kořulda, DINA'nın sınıflama doğruluđu performanslarının ,63 ile ,79; madde ayırt edicilik düzeyinin düşük olduđu kořulda ise ,54 ile ,70 arasında deđiřtiđi görölmektedir. DINA'nın sınıflama doğruluđu performansının özellikle madde ayırt edicilik düzeyinin düşük olduđu kořulda gittikçe azaldıđı, örneklem büyüklüğündeki artışlarla birlikte sınıflama doğruluđu performansında artışlar olsa da madde ayırt edicilik düzeyinin yüksek ve düşük olduđu her iki kořulda da, SISM'in DINA modelden daha yüksek sınıflama doğruluđu deđerlerine sahip

olduđu grlmektedir. Ayrıca SISM'e gre DINA modelin sınıflama dođruluđu performansının llen nitelik sayısından daha fazla etkilendiđi sylenbilir.

8 niteliđi len Q-matrisin kullanıldıđı madde ayırt edicilik dzeyinin yksek olduđu koşulda, kavram yanılgılarını len BUG-DINO modelinin kavram yanılgısı dzeyinde (K_1 , K_2 , K_3 , K_4) "nitelik dođru sınıflama oranı" deđerlerinin ,88 ile ,95; madde ayırt edicilik dzeyinin dşk olduđu koşulda ise ,81 ile ,90 arasında deđiştii grlmektedir. Bu durum rneklem byklđ faktrnden bađımsız, madde ayırt edicilik dzeyindeki artıřla birlikte sınıflama dođruluđu deđerlerinin arttıđının bir gstergesidir. Modele ait "ortalama nitelik dođru sınıflama oranı" deđerleri incelendiđinde madde ayırt edicilik dzeyinin yksek ve dşk olduđu koşullar kapsamında BUG-DINO'nun performansı farklı rneklem byklđ koşullarında artıřlar sađlamıřtır. Madde ayırt edicilik dzeyinin yksek olduđu farklı rneklem byklđ koşullarında BUG-DINO modelle SISM'e ait "nitelik dođru sınıflama oranı" deđerlerinin birbirine yakın deđerler aldıđı grlmektedir. Ayrıca llen nitelik sayısının artmasıyla BUG-DINO modelin sınıflama dođruluđu performansı azalsa da "ortalama nitelik dođru sınıflama oranı" deđerleri incelendiđinde madde ayırt edicilik dzeyinin yksek ve dşk olduđu her iki koşulda da modelin performansının ,80'in zerinde iyi sınıflama dođruluđu deđerlerine sahip olduđu sylenbilir.

Tablo 19

8 Nitelikli 40 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri

		Nitelik Doğru Sınıflama Oranı										
Örneklem büyüklüğü	Modeller	B ₁	B ₂	B ₃	B ₄	K ₁	K ₂	K ₃	K ₄	Ortalama Nitelik Doğru Sınıflama Oranı	Örüntü Doğru Sınıflama Oranı	
Madde Ayırt Ediciliği Yüksek (40 Madde)	250	DINA	,72	,77	,69	,78	-	-	-	-	,74	-
		BUG-DINO	-	-	-	-	,96	,97	,93	,93	,95	-
		SISM	,89	,87	,92	,87	,97	,98	,94	,94	,92	,54
	500	DINA	,72	,77	,69	,80	-	-	-	-	,74	-
		BUG-DINO	-	-	-	-	,96	,98	,95	,94	,95	-
		SISM	,89	,85	,92	,88	,97	,98	,95	,95	,92	,54
	1500	DINA	,76	,81	,73	,82	-	-	-	-	,78	-
		BUG-DINO	-	-	-	-	,97	,98	,95	,95	,96	-
		SISM	,91	,91	,93	,89	,98	,98	,95	,95	,93	,55
	3000	DINA	,81	,80	,85	,84	-	-	-	-	,82	-
		BUG-DINO	-	-	-	-	,97	,98	,95	,95	,96	-
		SISM	,91	,91	,93	,90	,98	,99	,96	,95	,94	,56
Madde Ayırt Ediciliği Düşük (40 Madde)	250	DINA	,59	,60	,58	,64	-	-	-	-	,60	-
		BUG-DINO	-	-	-	-	,92	,95	,85	,86	,89	-
		SISM	,76	,75	,81	,74	,93	,95	,85	,86	,83	,25
	500	DINA	,60	,62	,59	,65	-	-	-	-	,62	-
		BUG-DINO	-	-	-	-	,93	,96	,87	,87	,91	-
		SISM	,78	,72	,85	,75	,94	,96	,87	,87	,84	,26
	1500	DINA	,63	,69	,61	,67	-	-	-	-	,65	-
		BUG-DINO	-	-	-	-	,94	,96	,87	,87	,91	-
		SISM	,79	,70	,87	,76	,94	,96	,88	,89	,85	,27
	3000	DINA	,66	,67	,64	,69	-	-	-	-	,67	-
		BUG-DINO	-	-	-	-	,94	,96	,87	,88	,91	-
		SISM	,80	,68	,88	,77	,94	,96	,88	,89	,85	,28

B₁, B₂, B₃, B₄ sırasıyla 1., 2., 3. ve 4. becerileri; K₁, K₂, K₃, K₄ sırasıyla 1., 2., 3. ve 4. kavram yanlışlarını temsil etmektedir.

Tablo 19’da yer alan 40 maddeden oluşan bir test için SISM’e ait sınıflama doğruluğu değerleri incelendiğinde, örneklem büyüklüğünden bağımsız olarak her bir beceri ve kavram yanlışlığı düzeyinde (B₁, B₂, B₃, B₄; K₁, K₂, K₃, K₄) “nitelik doğru sınıflama oranı” değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,85 ile ,99; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,68 ile ,96 arasında değiştiği görülmektedir. “Ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde ise madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu değerlerin ,92 ile ,94; madde ayırt edicilik düzeyinin düşük olduğu koşullarda ise ,83 ile ,85 arasında değiştiği görülmektedir. “Örüntü doğru sınıflama oranı” değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,54 ile

,56; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,25 ile ,28 arasında değiştiği görülmektedir. Bu durum madde ayırt edicilik düzeyindeki artışların daha yüksek sınıflama doğruluğu değerleri verdiğinin bir göstergesidir. Madde ayırt edicilik düzeyinin yüksek olduğu koşulda SISM'in "ortalama nitelik doğru sınıflama oranı" değerleri incelendiğinde sınıflama doğruluğu değerlerinin örneklem büyüklüğü artışıyla birlikte arttığı görülmektedir. Madde ayırt edicilik düzeyinin düşük olduğu koşul incelendiğinde ise örneklem büyüklüğü faktöründen bağımsız SISM'e ait sınıflama doğruluğu değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşula göre azaldığı bulgusuna ulaşılmıştır. Örneklem büyüklüğü artışıyla birlikte "ortalama nitelik doğru sınıflama oranı" değerlerinin *-3000 örneklem büyüklüğü haricinde-* arttığı görülmektedir. Bu durum madde ayırt edicilik düzeyinin düşük olduğu koşulda örneklem büyüklüğünü 1500'den 3000'e çıkarmanın modele ait sınıflama doğruluğu performansına ek bir katkı getirmediğinin bir göstergesidir. Genel olarak değerlendirildiğinde SISM, madde ayırt edicilik düzeyinin yüksek ve test uzunluğunun 40 madde olduğu koşullarda küçük bir örneklem büyüklüğüne rağmen ölçülen nitelik sayısının artmasıyla daha düşük sınıflama doğruluğu performansı gösterse de "ortalama nitelik doğru sınıflama oranı" değerleri incelendiğinde modelin ,80 üzerinde iyi bir sınıflama doğruluğu performansı sergilediği söylenebilir.

SISM ve DINA modellerinin beceri düzeyinde (B_1, B_2, B_3, B_4) "nitelik doğru sınıflama oranı" değerleri test uzunluğunun 20 madde olduğu koşulla karşılaştırıldığında, 40 maddelik test uzunluğunda örneklem büyüklüğünden bağımsız madde ayırt edicilik düzeyinin yüksek olduğu koşulda DINA'nın sınıflama doğruluğu performanslarının ,69 ile ,85; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,58 ile ,69 arasında değiştiği görülmektedir. Örneklem büyüklüğünden bağımsız olarak test uzunluğunu 20 maddeden 40 maddeye çıkarmanın, özellikle madde ayırt edicilik düzeyinin yüksek olduğu koşullar dikkate alındığında DINA modelin performansına önemli katkılar sunduğu söylenebilir. Örneklem büyüklüğü artışıyla birlikte madde ayırt edicilik düzeyinin yüksek olduğu koşulda DINA modelde 250 ve 500 örneklem büyüklüğü koşullarında "ortalama nitelik doğru sınıflama oranı" değerleri aynı kalsa da örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda bu değerlerin arttığı ve örneklem büyüklüğünün 3000 olduğu koşulda ise bu değerlerin ,80'in üzerine çıktığı bulgusuna ulaşılmıştır. Her ne kadar test uzunluğu faktöründeki artışla birlikte DINA model performansı artsa da, örneklem büyüklüğü ve madde ayırt edicilik düzeyi faktörlerinde genel olarak SISM'in DINA modelden daha yüksek "nitelik doğru sınıflama oranı" değerlerine sahip olduğu görülmektedir. Ayrıca ölçülen nitelik sayısının

artmasıyla birlikte SISM'e göre DINA modelin sınıflama doğruluğu performansının ölçülen nitelik sayısından daha fazla etkilendiği söylenebilir. Ancak test uzunluğunun artmasıyla birlikte ölçülen nitelik sayısının modele ait sınıflama doğruluklarına etkisi, test uzunluğunun 20 madde olduğu koşula kıyasla 40 maddelik test uzunluğu koşulunda azalmıştır.

8 niteliği ölçen Q-matris kullanıldığında kavram yanılgılarını ölçen BUG-DINO modelinin 40 maddelik test uzunluğunda örneklem büyüklüğünden bağımsız ve madde ayırt edicilik düzeyinin yüksek olduğu koşulda kavram yanılgısı düzeyinde (K_1, K_2, K_3, K_4) “nitelik doğru sınıflama oranı” değerlerinin ,93 ile ,98; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,85 ile ,96 arasında değiştiği görülmektedir. Bu durum madde ayırt edicilik düzeyindeki artışla birlikte sınıflama doğruluğu değerlerinin arttığının bir göstergesidir. Modele ait “ortalama nitelik doğru sınıflama oranı” değerleri ise madde ayırt edicilik düzeyinin yüksek olduğu koşul kapsamında örneklem büyüklüğünü 1500'den 3000'e çıkarmanın model performansında ek bir artış sağlamadığını göstermektedir. Madde ayırt edicilik düzeyinin yüksek olduğu koşulda örneklem büyüklüğü artışıyla birlikte BUG-DINO model SISM ile karşılaştırıldığında “nitelik doğru sınıflama oranı” değerlerinin kavram yanılgısı düzeyinde birbirine yakın değerler aldığı görülsede genel olarak değerlendirildiğinde farklı koşullar altında bazı kavram yanılgılarında SISM'in sınıflama doğruluğu performansının daha yüksek olduğu görülmektedir. Benzer bir durum madde ayırt edicilik düzeyinin düşük olduğu koşul için de geçerlidir. Ayrıca test uzunluğunu 20 maddeden 40 maddeye çıkarmanın BUG-DINO model performansında madde ayırt edicilik düzeyinin yüksek ve düşük olduğu koşullarda sınıflama doğruluğu performansını arttırdığı görülmektedir. Ölçülen nitelik sayısının artmasıyla “ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde -“*madde ayırt edicilik düzeyi düşük*örneklem büyüklüğü 250 (,89)*” koşulu hariç- tüm koşullarda BUG-DINO'nun sınıflama doğruluğu performansı, 6 nitelikli Q-matrise göre azalsa da modelin küçük örneklemelerde dahi kavram yanılgılarını ölçmede halen ,90 üzeri sınıflama doğruluğu değerleri olarak iyi bir performans sergilediği söylenebilir.

Tablo 20

8 Nitelikli 60 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri

Nitelik Doğru Sınıflama Oranı												
	Örneklem büyüklüğü	Modeller	B ₁	B ₂	B ₃	B ₄	K ₁	K ₂	K ₃	K ₄	Ortalama Nitelik Doğru Sınıflama Oranı	Örüntü Doğru Sınıflama Oranı
BUG-DINO	-	-	-	-	,97	,98	,96	,96	,97	-		
SISM	,93	,91	,95	,91	,98	,99	,97	,97	,95	,68		
500	DINA	,87	,90	,83	,89	-	-	-	-	,87	-	
	BUG-DINO	-	-	-	-	,98	,97	,97	,96	,97	-	
	SISM	,93	,91	,96	,91	,98	,99	,97	,97	,95	,68	
1500	DINA	,88	,90	,86	,89	-	-	-	-	,88	-	
	BUG-DINO	-	-	-	-	,98	,97	,97	,97	,97	-	
	SISM	,93	,91	,96	,96	,99	,99	,97	,97	,95	,68	
3000	DINA	,88	,91	,87	,89	-	-	-	-	,89	-	
	BUG-DINO	-	-	-	-	,98	,97	,97	,97	,97	-	
	SISM	,94	,87	,97	,92	,99	,99	,98	,97	,95	,69	
Madde Ayırt Ediciliği Düşük (60 Madde)	250	DINA	,60	,62	,59	,69	-	-	-	-	,60	-
		BUG-DINO	-	-	-	-	,95	,97	,91	,91	,93	-
		SISM	,83	,78	,89	,79	,96	,97	,91	,92	,88	,37
	500	DINA	,62	,63	,60	,70	-	-	-	-	,64	-
		BUG-DINO	-	-	-	-	,96	,97	,92	,92	,94	-
		SISM	,84	,74	,90	,79	,96	,98	,92	,92	,89	,38
	1500	DINA	,65	,66	,63	,71	-	-	-	-	,66	-
		BUG-DINO	-	-	-	-	,96	,98	,93	,93	,95	-
		SISM	,85	,70	,91	,81	,96	,98	,93	,93	,89	,38
	3000	DINA	,69	,70	,67	,73	-	-	-	-	,70	-
		BUG-DINO	-	-	-	-	,97	,98	,93	,93	,95	-
		SISM	,86	,71	,92	,81	,97	,98	,94	,94	,89	,38

B₁, B₂, B₃, B₄ sırasıyla 1., 2., 3. ve 4. becerileri; K₁, K₂, K₃, K₄ sırasıyla 1., 2., 3. ve 4. kavram yanlışlarını temsil etmektedir.

Tablo 20’de yer alan 60 maddeden oluşan bir test için SISM’e ait sınıflama doğruluğu değerleri incelendiğinde, örneklem büyüklüğünden bağımsız olarak her bir beceri ve kavram yanlışlığı düzeyinde (B₁, B₂, B₃, B₄; K₁, K₂, K₃, K₄) “nitelik doğru sınıflama oranı” değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,91 ile ,99; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,70 ile ,98 arasında değiştiği görülmektedir. “Ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde ise madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu değerlerin ,95 olduğu; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,88 ile ,89 arasında değiştiği görülmektedir. Madde ayırt edicilik düzeyinin yüksek olduğu koşulda “örüntü doğru sınıflama oranı” değerlerinin ,68 ile ,69;

madde ayırt edicilik düzeyinin düşük olduğu koşulda ise bu değerlerin ,37 ile ,38 arasında değiştiği görülmektedir. Bu durum madde ayırt edicilik düzeyindeki artışların daha yüksek sınıflama doğruluğu değerleri verdiği bir göstergesidir. Madde ayırt edicilik düzeyinin yüksek olduğu koşulda SISM'in "ortalama nitelik doğru sınıflama oranı" değerleri incelendiğinde bu değerlerin aynı olduğu ve örneklem büyüklüğü faktöründeki artışların SISM'in sınıflama doğruluğu performansına ek bir katkı sağlamadığı görülmektedir. Madde ayırt edicilik düzeyinin düşük olduğu koşulda incelendiğinde örneklem büyüklüğü faktöründen bağımsız SISM'e ait sınıflama doğruluğu değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşula göre azaldığı bulgusuna ulaşılmıştır. Bu kapsamda örneklem büyüklüğünün 250'den 500'e çıkarıldığı koşulda modelin sınıflama doğruluğu performansında bir artış görülse de 500, 1500 ve 3000 örneklem büyüklüğü koşullarında modele ait sınıflama doğruluğu performanslarının aynı olduğu ve örneklem büyüklüğü artışının ek bir katkı sağlamadığı bulgusuna ulaşılmıştır. 8 nitelikli Q-matrisin kullanıldığı 60 maddelik test uzunluğu koşulu, 20 ve 40 maddelik test uzunluğu koşullarıyla karşılaştırıldığında SISM'e ait "nitelik doğru sınıflama oranı" değerleriyle "ortalama nitelik doğru sınıflama oranı" değerlerinin artış sağladığı bulgusuna ulaşılmıştır. 8 nitelikli Q-matrisin ve 6 nitelikli Q-matrisin kullanıldığı koşullar genel olarak değerlendirildiğinde test uzunluğunun 60 madde, madde ayırt edicilik düzeyinin yüksek olduğu tüm örneklem büyüklüğü koşullarında SISM'in aynı "ortalama nitelik doğru sınıflama oranı" değerlerine sahip olduğu bulgusuna ulaşılmıştır. Ayrıca her iki matris için de 60 maddelik test uzunluğunda ölçülen nitelik sayısının artmasıyla birlikte "ortalama nitelik doğru sınıflama oranı" değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,90 üzerinde yüksek sınıflama doğruluğu değerleri aldığı görülmektedir.

8 nitelikli Q-matrisin kullanıldığı DINA modelde 60 madde için beceri düzeyinde (B_1 , B_2 , B_3 , B_4) "nitelik doğru sınıflama oranı" değerleri incelendiğinde, örneklem büyüklüğünden bağımsız olarak madde ayırt edicilik düzeyinin yüksek olduğu koşulda sınıflama doğruluğu değerlerinin ,82 ile ,91; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,59 ile ,73 arasında değiştiği görülmektedir. "Ortalama nitelik doğru sınıflama oranı" değerleri incelendiğinde madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu değerlerin ,87 ile ,89; madde ayırt edicilik düzeyinin düşük olduğu koşulda bu değerlerin ,60 ile ,70 arasında değiştiği görülmektedir. DINA modelde test uzunluğunun 60 madde olduğu koşulla test uzunluğunun 20 ve 40 madde olduğu koşullar karşılaştırıldığında, örneklem büyüklüğünden bağımsız olarak test uzunluğu artışının özellikle madde ayırt edicilik düzeyinin yüksek

olduğu koşul dikkate alındığında sınıflama doğruluğu performansına önemli katkılar sağladığı ve model performansının arttığı bulgusuna ulaşılmıştır. DINA modelde madde ayırt edicilik düzeyinin yüksek olduğu koşulda 250 ve 500 örneklem büyüklüğünde “ortalama nitelik doğru sınıflama oranı” değerleri aynı kalsa da örneklem büyüklüğünün artışıyla, örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda bu değerlerin arttığı bulgusuna ulaşılmıştır. Madde ayırt edicilik düzeyinin düşük olduğu tüm koşullarda ise örneklem büyüklüğünün artışıyla bu değerlerin arttığı görülmektedir. Her ne kadar test uzunluğu faktöründeki artışla birlikte DINA model performansı artsa da örneklem büyüklüğü ve madde ayırt edicilik düzeyi faktörlerinde genel olarak SISM’in, DINA modelden daha yüksek sınıflama doğruluğu değerlerine sahip olduğu görülmektedir. 6 nitelikli Q-matrisle 8 nitelikli Q-matrisin kullanıldığı koşullar karşılaştırıldığında, DINA modelin ölçülen nitelik sayısının artmasıyla sınıflama doğruluğu performansının azaldığı görülmektedir. Ayrıca aynı koşullar altında SISM ile karşılaştırıldığında DINA modelin sınıflama doğruluğu performansının, ölçülen nitelik sayısından daha fazla etkilendiği söylenebilir.

8 niteliği ölçen Q-matris kullanıldığında kavram yanılgılarını ölçen BUG-DINO modelinin 60 maddelik test uzunluğunda örneklem büyüklüğünden bağımsız ve madde ayırt edicilik düzeyinin yüksek olduğu koşulda kavram yanılgısı düzeyinde (K_1, K_2, K_3, K_4) “nitelik doğru sınıflama oranı” değerleri ,96 ile ,98; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,91 ile ,98 arasında değişmektedir. Modele ait “ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde ise bu değerlerin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,97 olduğu ve madde ayırt edicilik düzeyinin düşük olduğu koşulda ,93 ile ,95 arasında değiştiği görülmektedir. Bu durum madde ayırt edicilik düzeyindeki artışla birlikte sınıflama doğruluğu değerlerinin arttığının bir göstergesidir. Madde ayırt edicilik düzeyinin yüksek olduğu koşul kapsamında BUG-DINO modeline ait “ortalama nitelik doğru sınıflama oranı” değerleri tüm örneklem büyüklüklerinde aynıdır. BUG-DINO model SISM’le karşılaştırıldığında “nitelik doğru sınıflama oranı” değerlerinin kavram yanılgısı düzeyinde birbirine yakın değerler aldığı görülse de genel olarak değerlendirildiğinde farklı koşullar altında bazı kavram yanılgılarında SISM’in sınıflama doğruluğu performansının daha yüksek olduğu görülmektedir. Benzer bir durum madde ayırt edicilik düzeyinin düşük olduğu koşullar için de geçerlidir. Ayrıca 8 nitelikli Q-matrisin kullanıldığı koşulda test uzunluğunu 20 ve 40 maddeden 60 maddeye çıkarmanın madde ayırt edicilik düzeyinin yüksek ve düşük olduğu her iki koşulda da BUG-DINO’nun sınıflama doğruluğu

performansını arttırdığı bulgusuna ulaşılmıştır. Ölçülen nitelik sayısının artmasıyla, tüm koşullarda BUG-DINO'nun sınıflama doğruluğu performansı 6 nitelikli Q-matrise göre azalsa da, modelin çok küçük örneklerde dahi kavram yanlışlarını ölçmede halen ,90 üzeri sınıflama doğruluğu değerleri alarak yüksek bir performans sergilediği söylenebilir.

Tablo 21

10 Nitelikli 20 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri

		Nitelik Doğru Sınıflama Oranı										Ortalama Nitelik Doğru Sınıflama Oranı	Örüntü Doğru Sınıflama Oranı	
Örneklem büyüklüğü	Modeller	B ₁	B ₂	B ₃	B ₄	B ₅	K ₁	K ₂	K ₃	K ₄	K ₅			
Madde Ayırt Edicilik Düzeyi Yüksek (20 Madde)	250	DINA	,63	,62	,60	,66	,68	-	-	-	-	-	,65	-
		BUG-DINO	-	-	-	-	-	,84	,82	,81	,84	,81	,82	-
		SISM	,80	,80	,79	,80	,81	,81	,90	,90	,89	,90	,84	,20
	500	DINA	,64	,66	,62	,67	,70	-	-	-	-	-	,66	-
		BUG-DINO	-	-	-	-	-	,89	,89	,90	,89	,89	,89	-
		SISM	,80	,80	,80	,79	,80	,89	,89	,90	,89	,89	,85	,20
	1500	DINA	,67	,68	,66	,69	,71	-	-	-	-	-	,68	-
		BUG-DINO	-	-	-	-	-	,90	,90	,90	,89	,89	,90	-
		SISM	,80	,79	,79	,81	,80	,90	,90	,91	,89	,90	,86	,22
3000	DINA	,71	,70	,69	,69	,71	-	-	-	-	-	,70	-	
	BUG-DINO	-	-	-	-	-	,89	,90	,90	,90	,89	,91	-	
	SISM	,80	,80	,79	,81	,81	,90	,91	,90	,89	,90	,86	,23	
Madde Ayırt Edicilik Düzeyi Düşük (20 Madde)	250	DINA	,54	,53	,51	,54	,52	-	-	-	-	-	,52	-
		BUG-DINO	-	-	-	-	-	,83	,82	,81	,81	,81	,81	-
		SISM	,70	,70	,72	,70	,71	,82	,82	,82	,83	,81	,75	,08
	500	DINA	,56	,54	,52	,54	,54	-	-	-	-	-	,54	-
		BUG-DINO	-	-	-	-	-	,82	,82	,83	,82	,82	,82	-
		SISM	,72	,73	,72	,72	,72	,82	,83	,82	,82	,83	,77	,09
	1500	DINA	,57	,55	,54	,57	,56	-	-	-	-	-	,56	-
		BUG-DINO	-	-	-	-	-	,82	,83	,83	,82	,83	,83	-
		SISM	,72	,73	,72	,73	,72	,82	,82	,83	,83	,83	,78	,10
3000	DINA	,58	,57	,55	,58	,57	-	-	-	-	-	,57	-	
	BUG-DINO	-	-	-	-	-	,83	,82	,82	,83	,83	,83	-	
	SISM	,73	,73	,72	,72	,73	,83	,83	,83	,82	,83	,78	,10	

B₁, B₂, B₃, B₄, B₅ sırasıyla 1., 2., 3., 4. ve 5. becerileri; K₁, K₂, K₃, K₄, K₅ sırasıyla 1., 2., 3., 4. ve 5. kavram yanlışlarını temsil etmektedir.

Ölçülen nitelik sayısının arttırıldığı 10 nitelik içeren bir Q-matris kullanıldığında 20 maddeden oluşan test için Tablo 21’de yer alan SISM’e ait beceri ve kavram yanılığısı düzeylerinde ($B_1, B_2, B_3, B_4, B_5; K_1, K_2, K_3, K_4, K_5$) sınıflama doğruluğu değerleri incelendiğinde örneklem büyüklüğünden bağımsız olarak her bir beceri ve kavram yanılığısı düzeyinde “nitelik doğru sınıflama oranı” değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,79 ile ,91; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,70 ile ,83 arasında değiştiği görülmektedir. “Ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu değerlerin ,84 ile ,86 ve madde ayırt edicilik düzeyinin düşük olduğu koşulda ,75 ile ,78 arasında değiştiği görülmektedir. Madde ayırt edicilik düzeyinin yüksek ve düşük olduğu her iki koşulda da ölçülen nitelik sayısının artmasıyla birlikte örneklem büyüklüğünden bağımsız olarak SISM’in sınıflama doğruluğu performansının azaldığı görülmektedir. 10 nitelikli Q-matrisin kullanıldığı, madde ayırt edicilik düzeyinin yüksek olduğu ve farklı örneklem büyüklüğü koşullarında SISM’in sınıflama doğruluğu performansında artışlar olduğu ancak örneklem büyüklüğünün 1500’den 3000’e çıkarıldığı koşulda örneklem büyüklüğü faktörünün sınıflama doğruluğu değerlerinde ek bir artış sağlamadığı görülmektedir. “Örüntü doğru sınıflama oranı” değerleri incelendiğinde ise madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu değerlerin ,20 ile ,23; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,08 ile ,10 arasında değiştiği görülmektedir. Ölçülen nitelik sayısının artmasıyla bu değerlerin azalarak model performansının gittikçe azaldığı bulgusuna ulaşılmıştır. Genel olarak değerlendirildiğinde 10 nitelikli Q-matrisin kullanıldığı ve madde ayırt edicilik düzeyinin yüksek olduğu koşullarla 6 ve 8 nitelikli Q-matrislerin kullanıldığı koşullardan elde edilen bulgular karşılaştırıldığında, ölçülen nitelik sayısının artmasıyla test uzunluğunun 20 madde olduğu koşul altında SISM’in daha düşük sınıflama doğruluğu performansı sergilediği görülmektedir.

10 nitelikli Q-matrisin kullanıldığı DINA modelde 20 maddelik testte beceri düzeyinde (B_1, B_2, B_3, B_4, B_5) “nitelik doğru sınıflama oranı” değerleri incelendiğinde, örneklem büyüklüğünden bağımsız olarak madde ayırt edicilik düzeyinin yüksek olduğu koşulda sınıflama doğruluğu değerlerinin ,60 ile ,71; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,51 ile ,58 arasında değiştiği görülmektedir. Modele ait “ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde ise bu değerlerin madde ayırt edicilik düzeyi yüksek olduğunda ,65 ile ,70; madde ayırt edicilik düzeyi düşük olduğunda ,52 ile ,57 arasında değiştiği görülmektedir. Modele ait sınıflama doğruluğu performansının özellikle

madde ayırt edicilik düzeyinin düşük olduğu koşulda madde ayırt edicilik düzeyinin yüksek olduğu koşula kıyasla gittikçe azaldığı bulgusuna ulaşılmıştır. Ancak madde ayırt edicilik düzeyinin yüksek ve düşük olduğu her iki koşulda da örneklem büyüklüğündeki artışlarla birlikte sınıflama doğruluğu performansında artışların olduğu görülmektedir. Tablo 21’de yer alan koşullar değerlendirildiğinde SISM’in DINA modelden daha yüksek sınıflama doğruluğu değerlerine sahip olduğu ve ölçülen nitelik sayısının artmasıyla birlikte DINA’nın SISM’e göre sınıflama doğruluğu performansının ölçülen nitelik sayısından daha fazla etkilendiği söylenebilir.

10 niteliği ölçen Q-matris kullanıldığında kavram yanılgılarını ölçen BUG-DINO modelinin 20 maddelik test uzunluğunda örneklem büyüklüğünden bağımsız ve madde ayırt edicilik düzeyinin yüksek olduğu koşulda kavram yanılgıları düzeyinde (K_1, K_2, K_3, K_4, K_5) “nitelik doğru sınıflama oranı” değerleri ,81 ile ,90; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,81 ile ,83 arasında değişmektedir. Modele ait “ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde ise bu değerlerin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,82 ile ,91 ve madde ayırt edicilik düzeyinin düşük olduğu koşulda ,81 ile ,83 arasında değiştiği görülmektedir. Bu durum madde ayırt edicilik düzeyindeki artışla birlikte sınıflama doğruluğu değerlerinin arttığının bir göstergesidir. BUG-DINO’nun madde ayırt edicilik düzeyinin yüksek olduğu koşulda, örneklem büyüklüğünün artışıyla birlikte “ortalama nitelik doğru sınıflama oranı” değerlerinin arttığı görülmektedir. Madde ayırt edicilik düzeyinin yüksek olduğu koşulda örneklem büyüklüğü artışıyla birlikte BUG-DINO model SISM ile karşılaştırıldığında “nitelik doğru sınıflama oranı” değerlerinin $-K_1$ kavram yanılgısı düzeyinde, örneklem büyüklüğünün 250 olduğu koşul hariç- kavram yanılgısı düzeyinde birbirine yakın değerler aldığı görülse de farklı koşullarda bazı kavram yanılgılarında SISM’in sınıflama doğruluğu performansının BUG-DINO modele göre daha yüksek olduğu bulgusuna ulaşılmıştır. Benzer bir durum *-örneğin K_3 kavram yanılgısı düzeyinde örneklem büyüklüğünün 500 olduğu koşulu hariç-* madde ayırt edicilik düzeyinin düşük olduğu koşullar için de geçerlidir. Ölçülen nitelik sayısının artmasıyla tüm koşullarda BUG-DINO’nun sınıflama doğruluğu performansı 6 ve 8 nitelikli Q-matrislere göre azalsa da, madde ayırt edicilik düzeyinin yüksek ya da düşük olduğu her iki koşulda da modelin çok küçük örneklemelerde dahi kavram yanılgılarını ölçmede ,80 üzeri sınıflama doğruluğu değerleri alarak iyi bir performans sergilediği söylenebilir.

Tablo 22

10 Nitelikli 40 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri

		Nitelik Doğru Sınıflama Oranı											Ortalama Nitelik Doğru Sınıflama Oranı	Örüntü Doğru Sınıflama Oranı
Örneklem büyüklüğü	Modeller	B_1	B_2	B_3	B_4	B_5	K_1	K_2	K_3	K_4	K_5			
Madde Ayırt Edicilik Düzeyi Yüksek (40 Madde)	250	<i>DINA</i>	,70	,75	,66	,74	,74	-	-	-	-	-	,72	-
		<i>BUG-DINO</i>	-	-	-	-	-	,93	,92	,91	,94	,94	,92	-
		<i>SISM</i>	,88	,87	,87	,86	,90	,90	,93	,93	,94	,94	,90	,40
	500	<i>DINA</i>	,70	,76	,69	,77	,76	-	-	-	-	-	,73	-
		<i>BUG-DINO</i>	-	-	-	-	-	,93	,93	,93	,94	,95	,93	-
		<i>SISM</i>	,86	,87	,87	,87	,91	,93	,93	,93	,95	,95	,90	,41
	1500	<i>DINA</i>	,72	,77	,71	,77	,77	-	-	-	-	-	,75	-
		<i>BUG-DINO</i>	-	-	-	-	-	,94	,93	,92	,95	,95	,94	-
		<i>SISM</i>	,87	,87	,88	,87	,91	,94	,93	,94	,95	,95	,91	,42
	3000	<i>DINA</i>	,73	,77	,72	,77	,78	-	-	-	-	-	,75	-
		<i>BUG-DINO</i>	-	-	-	-	-	,94	,94	,93	,95	,95	,94	-
		<i>SISM</i>	,88	,88	,87	,90	,91	,95	,94	,94	,95	,95	,92	,46
Madde Ayırt Edicilik Düzeyi Düşük (40 Madde)	250	<i>DINA</i>	,56	,58	,56	,62	,60	-	-	-	-	-	,58	-
		<i>BUG-DINO</i>	-	-	-	-	-	,82	,83	,82	,86	,86	,83	-
		<i>SISM</i>	,74	,74	,73	,73	,76	,84	,84	,83	,87	,87	,80	,11
	500	<i>DINA</i>	,58	,61	,57	,63	,59	-	-	-	-	-	,60	-
		<i>BUG-DINO</i>	-	-	-	-	-	,84	,83	,83	,86	,86	,84	-
		<i>SISM</i>	,74	,73	,73	,75	,76	,85	,84	,84	,87	,87	,81	,12
	1500	<i>DINA</i>	,61	,64	,59	,65	,62	-	-	-	-	-	,62	-
		<i>BUG-DINO</i>	-	-	-	-	-	,85	,84	,84	,88	,87	,86	-
		<i>SISM</i>	,77	,65	,81	,76	,78	,87	,86	,86	,88	,89	,81	,12
	3000	<i>DINA</i>	,63	,65	,62	,66	,64	-	-	-	-	-	,64	-
		<i>BUG-DINO</i>	-	-	-	-	-	,86	,85	,85	,88	,88	,86	-
		<i>SISM</i>	,78	,66	,83	,77	,78	,88	,87	,87	,88	,89	,82	,13

B_1, B_2, B_3, B_4, B_5 sırasıyla 1., 2., 3., 4. ve 5. becerileri; K_1, K_2, K_3, K_4, K_5 sırasıyla 1., 2., 3., 4. ve 5. kavram yanılgılarını temsil etmektedir.

Tablo 22’de yer alan 40 maddeden oluşan bir test için SISM’e ait beceri ve kavram yanılgısı düzeylerinde ($B_1, B_2, B_3, B_4, B_5; K_1, K_2, K_3, K_4, K_5$) sınıflama doğruluğu değerleri incelendiğinde, örneklem büyüklüğünden bağımsız olarak her bir beceri ve kavram yanılgısı kapsamında “nitelik doğru sınıflama oranı” değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,86 ile ,95; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,66 ile ,89 arasında değiştiği görülmektedir. “Ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde ise madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu

değerlerin ,90 ile ,92; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,80 ile ,82 arasında değiştiği görülmektedir. “Örüntü doğru sınıflama oranı” değerlerinin ise madde ayırt edicilik düzeyi yüksek olduğunda ,40 ile ,46; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,11 ile ,13 arasında değiştiği görülmektedir. Bu durum madde ayırt edicilik düzeyindeki artışların daha yüksek sınıflama doğruluğu değerleri verdiğinin bir göstergesidir. Madde ayırt edicilik düzeyinin yüksek olduğu koşulda SISM’in “ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde örneklem büyüklüğünün 250 ve 500 olduğu koşullarda değerlerin aynı olduğu ancak örneklem büyüklüğünün 1500’den 3000’e çıkarıldığı koşulda sınıflama doğruluğu değerlerinin arttığı görülmektedir. Madde ayırt edicilik düzeyinin düşük olduğu koşulda incelendiğinde örneklem büyüklüğü faktöründen bağımsız olarak SISM’e ait sınıflama doğruluğu değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşula göre azaldığı bulgusuna ulaşılmıştır. Madde ayırt edicilik düzeyinin düşük olduğu koşulda sınıflama doğruluğu performanslarının örneklem büyüklüğünün artışıyla birlikte arttığı görülmektedir. 10 nitelikli Q-matrisin kullanıldığı ve madde ayırt edicilik düzeyinin yüksek olduğu koşullarla 6 ve 8 nitelikli Q-matrislerin kullanıldığı koşullardan elde edilen bulgular karşılaştırıldığında, test uzunluğunun 20 ve 40 madde olduğu koşullar altında ölçülen nitelik sayısının artmasıyla SISM’in daha düşük sınıflama doğruluğu performansı sergilediği bulgusuna ulaşılmıştır. Ancak madde ayırt edicilik düzeyinin yüksek ve düşük olduğu her iki koşulda da “ortalama nitelik doğru sınıflama oranı” değerlerinin ,80 ve üzerinde olduğu görülmektedir.

10 nitelikli Q-matrisin kullanıldığı DINA modelde 40 madde için beceri düzeyinde (B_1, B_2, B_3, B_4, B_5) “nitelik doğru sınıflama oranı” değerleri incelendiğinde, örneklem büyüklüğünden bağımsız olarak madde ayırt edicilik düzeyinin yüksek olduğu koşulda sınıflama doğruluğu değerlerinin ,66 ile ,78; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,56 ile ,66 arasında değiştiği görülmektedir. Modele ait “ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde ise bu değerlerin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,72 ile ,75; madde ayırt edicilik düzeyinin düşük olduğu koşulda ,58 ile ,64 arasında değiştiği görülmektedir. DINA modelde test uzunluğunun 40 madde olduğu koşulla test uzunluğunun 20 madde olduğu koşul karşılaştırıldığında örneklem büyüklüğünden bağımsız olarak test uzunluğu artışının özellikle madde ayırt edicilik düzeyinin yüksek olduğu koşullar dikkate alındığında DINA modelin sınıflama doğruluğu performansına önemli katkılar sunduğu ve model performansının arttığı görülmektedir. DINA modelde madde ayırt edicilik düzeyinin yüksek olduğu farklı örneklem büyüklüğü

koşulları altında örneklem büyüklüğünü 250'den sırasıyla 500 ve 1500'e çıkarmanın modele ait sınıflama doğruluğu performansında artışlar sağladığı görülmektedir. Ancak örneklem büyüklüğünü 1500'den 3000'e çıkarmanın 10 nitelikli Q-matrisin kullanıldığı koşulda modelin sınıflama doğruluğu performansında ek bir artış sağlamadığı bulgusuna ulaşılmıştır. Test uzunluğu faktöründeki artışla birlikte DINA model performansı artsa da örneklem büyüklüğü ve madde ayırt edicilik düzeyi faktörlerinde SISM'in, DINA modelden daha yüksek sınıflama doğruluğu değerlerine sahip olduğu görülmektedir. 6 ve 8 nitelikli Q-matrislerle 10 nitelikli Q-matrisin kullanıldığı koşullar karşılaştırıldığında DINA modelin ölçülen nitelik sayısının artmasıyla sınıflama doğruluğu performansının azaldığı bulgusuna ulaşılmıştır. Ayrıca aynı koşullar altında SISM ile DINA modelleri karşılaştırıldığında ölçülen nitelik sayısının artmasıyla birlikte DINA modelin sınıflama doğruluğu performansının azaldığı ve SISM'den daha düşük sınıflama doğruluğu performansı sergilediği söylenebilir.

10 niteliği ölçen Q-matris kullanıldığında kavram yanılığını ölçen BUG-DINO modelinin 40 maddelik test uzunluğunda örneklem büyüklüğünden bağımsız ve madde ayırt edicilik düzeyinin yüksek olduğu koşulda kavram yanılığı düzeyinde (K_1, K_2, K_3, K_4, K_5) “nitelik doğru sınıflama oranı” değerleri ,91 ile ,95; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,82 ile ,88 arasında değişmektedir. Modele ait “ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde ise bu değerlerin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,92 ile ,94 ve madde ayırt edicilik düzeyinin düşük olduğu koşulda ,83 ile ,86 arasında değiştiği görülmektedir. Bu durum madde ayırt edicilik düzeyindeki artışla birlikte sınıflama doğruluğu değerlerinin arttığının bir göstergesidir. BUG-DINO'nun madde ayırt edicilik düzeyinin yüksek ve örneklem büyüklüğünün 250'den sırasıyla 500 ve 1500'e çıkarıldığı koşullarda arttığı ancak örneklem büyüklüğünün 1500'den 3000'e çıkarıldığı koşulda “ortalama nitelik doğru sınıflama oranı” değerlerinin artmadığı yani modelin sınıflama doğruluğu performansında ek bir artış sağlanmadığı bulgusuna ulaşılmıştır. Madde ayırt edicilik düzeyinin yüksek olduğu koşulda örneklem büyüklüğü artışıyla birlikte BUG-DINO model SISM ile karşılaştırıldığında “nitelik doğru sınıflama oranı” değerlerinin kavram yanılığı düzeyinde $-K_1$ kavram yanılığı düzeyinde örneklem büyüklüğünün 250 olduğu koşul hariç- birbirine yakın değerler aldığı görülse de farklı koşullarda bazı kavram yanılığlarında SISM'in genel anlamda sınıflama doğruluğu performansının daha yüksek olduğu bulgusuna ulaşılmıştır. Madde ayırt edicilik düzeyinin düşük olduğu koşulda da “nitelik doğru sınıflama oranı” değerleri kavram yanılığı

düzeyinde birbirine yakın değerler olsa da farklı koşullarda bazı kavram yanlışlarında SISM'in genel anlamda sınıflama doğruluğu performansının daha yüksek olduğu görülmektedir. Ayrıca 10 nitelikli Q-matrisin kullanıldığı koşulda test uzunluğunu 20 maddeden 40 maddeye çıkarmanın madde ayırt edicilik düzeyinin yüksek ve düşük olduğu her iki koşulda da BUG-DINO'nun sınıflama doğruluğu performansını arttırdığı görülmektedir. Ölçülen nitelik sayısının artmasıyla tüm koşullarda BUG-DINO'nun sınıflama doğruluğu performansı 6 ve 8 nitelikli Q-matrislere göre azalsa da, modelin çok küçük örneklerde dahi kavram yanlışlarını ölçmede halen ,80 üzerinde sınıflama doğruluğu değerleri alarak iyi bir performans sergilediği söylenebilir.



Tablo 23

10 Nitelikli 60 Maddelik Bir Testin SISM, DINA ve BUG-DINO Modelleri Kullanıldığında Değişimlenen Koşullar Altındaki Sınıflama Doğruluğu Değerleri

		Nitelik Doğru Sınıflama Oranı												
Örneklem büyüklüğü	Modeller	B ₁	B ₂	B ₃	B ₄	B ₅	K ₁	K ₂	K ₃	K ₄	K ₅	Ortalama Nitelik Doğru Sınıflama Oranı	Örüntü Doğru Sınıflama Oranı	
Madde Ayırt Edicilik Düzeyi Yüksek (60 Madde)	250	<i>DINA</i>	,84	,90	,81	,88	,87	-	-	-	-	-	,86	-
		<i>BUG-DINO</i>	-	-	-	-	-	,95	,96	,96	,97	,96	,96	-
		<i>SISM</i>	,89	,88	,90	,90	,88	,96	,96	,97	,97	,97	,93	,48
	500	<i>DINA</i>	,86	,89	,85	,86	,90	-	-	-	-	-	,87	-
		<i>BUG-DINO</i>	-	-	-	-	-	,96	,96	,96	,97	,97	,96	-
		<i>SISM</i>	,90	,87	,92	,91	,91	,96	,96	,96	,97	,97	,93	,50
	1500	<i>DINA</i>	,85	,88	,90	,86	,91	-	-	-	-	-	,88	-
		<i>BUG-DINO</i>	-	-	-	-	-	,96	,96	,96	,97	,97	,96	-
		<i>SISM</i>	,91	,92	,93	,94	,91	,97	,96	,96	,97	,97	,94	,50
	3000	<i>DINA</i>	,88	,87	,87	,89	,91	-	-	-	-	-	,88	-
		<i>BUG-DINO</i>	-	-	-	-	-	,96	,96	,96	,97	,97	,96	-
		<i>SISM</i>	,89	,87	,93	,91	,91	,96	,96	,97	,98	,98	,95	,52
Madde Ayırt Edicilik Düzeyi Düşük (60 Madde)	250	<i>DINA</i>	,59	,60	,59	,67	,61	-	-	-	-	-	,60	-
		<i>BUG-DINO</i>	-	-	-	-	-	,91	,90	,90	,92	,92	,91	-
		<i>SISM</i>	,77	,76	,78	,77	,80	,90	,89	,90	,92	,93	,84	,18
	500	<i>DINA</i>	,61	,61	,60	,68	,62	-	-	-	-	-	,62	-
		<i>BUG-DINO</i>	-	-	-	-	-	,90	,90	,90	,93	,93	,92	-
		<i>SISM</i>	,77	,76	,77	,78	,80	,91	,90	,90	,93	,93	,85	,19
	1500	<i>DINA</i>	,62	,64	,62	,68	,63	-	-	-	-	-	,64	-
		<i>BUG-DINO</i>	-	-	-	-	-	,92	,91	,91	,93	,93	,92	-
		<i>SISM</i>	,78	,78	,77	,79	,80	,91	,92	,92	,93	,93	,86	,20
	3000	<i>DINA</i>	,63	,65	,65	,68	,64	-	-	-	-	-	,65	-
		<i>BUG-DINO</i>	-	-	-	-	-	,92	,92	,91	,93	,93	,92	-
		<i>SISM</i>	,78	,77	,79	,80	,81	,92	,92	,91	,93	,93	,86	,20

B₁, B₂, B₃, B₄, B₅ sırasıyla 1., 2., 3., 4. ve 5. becerileri; K₁, K₂, K₃, K₄, K₅ sırasıyla 1., 2, 3., 4. ve 5. kavram yanılıgılarını temsil etmektedir.

Tablo 23'te yer alan 60 maddeden oluşan bir test için SISM'e ait beceri ve kavram yanılıgısı düzeyinde (B₁, B₂, B₃, B₄, B₅; K₁, K₂, K₃, K₄, K₅) sınıflama doğruluğu değerleri incelendiğinde, örneklem büyüklüğünden bağımsız olarak her bir beceri ve kavram yanılıgısı kapsamında "nitelik doğru sınıflama oranı" değerlerinin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,87 ile ,98; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,76 ile ,93 arasında değiştiği görülmektedir. "Ortalama nitelik doğru sınıflama oranı" değerleri incelendiğinde ise madde ayırt edicilik düzeyinin yüksek olduğu koşulda bu değerlerin ,93 ile ,95; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,84 ile ,86

arasında deęiřtięi grlmektedir. “rnt doęru sınıflama oranı” deęerlerinin ise madde ayırt edicilik dzeyinin yksek olduęu kořulda ,48 ile ,52; madde ayırt edicilik dzeyinin dřk olduęu kořulda ,18 ile ,20 arasında deęiřtięi grlmektedir. Bu durum madde ayırt edicilik dzeyindeki artıřların daha yksek sınıflama doęruluęu deęerleri verdięinin bir gstergesidir. Madde ayırt edicilik dzeyinin yksek olduęu kořulda SISM’in “ortalama nitelik doęru sınıflama oranı” deęerleri incelendięinde sınıflama doęruluęu deęerlerinin 250 ve 500 rneklem byklkleri iin aynı olduęu; rneklem byklę kořulu 500’den sırasıyla 1500 ve 3000’e ıkarıldıęında ise SISM’in sınıflama doęruluęu performansında artıřların olduęu grlmektedir. Madde ayırt edicilik dzeyinin dřk olduęu kořulda incelendięinde rneklem byklę faktrnden baęımsız SISM’e ait sınıflama doęruluęu deęerlerinin madde ayırt edicilik dzeyinin yksek olduęu kořula gre azaldıęı bulgusuna ulařılmıřtır. Bu kořul kapsamında rneklem byklęnn 250’den sırasıyla 500 ve 1500’e ıkarıldıęı kořullarda modelin sınıflama doęruluęu performansında artıřlar olduęu grlmektedir. Ancak rneklem byklęn 1500’den 3000’e ıkarmanın modelin sınıflama doęruluęu performansında ek bir artıř saęlamadıęı bulgusuna ulařılmıřtır.

Genel olarak deęerlendirildięinde 10 nitelikli Q-matrisin kullanıldıęı ve madde ayırt edicilik dzeyinin yksek olduęu kořulla 6 ve 8 nitelikli Q-matrislerin kullanıldıęı kořullardan elde edilen bulgular karřılařtırıldıęında, test uzunluęunun 20 ve 40 madde olduęu kořullar altında llen nitelik sayısının artmasıyla SISM’in daha dřk sınıflama doęruluęu performansı sergiledięi bulgusuna ulařılmıřtır. Ancak “ortalama nitelik doęru sınıflama oranı” deęerleri incelendięinde bu deęerlerin ,80 zerinde iyi sınıflama doęruluęu deęerleri sunduęu grlmektedir. Test uzunluęunun 60 madde ve madde ayırt edicilik dzeyinin yksek olduęu tm rneklem byklę kořullarında, 10 nitelikli Q-matrisle 6 ve 8 nitelikli Q-matrisler karřılařtırıldıęında llen nitelik sayısının artmasıyla SISM’in “ortalama nitelik doęru sınıflama oranı” deęerlerinin azaldıęı bulgusuna ulařılmıřtır.

10 nitelikli Q-matrisin kullanıldıęı DINA modelde 60 madde iin beceri dzeyinde (B_1, B_2, B_3, B_4, B_5) “nitelik doęru sınıflama oranı” deęerleri incelendięinde, rneklem byklęnden baęımsız olarak madde ayırt edicilik dzeyinin yksek olduęu kořulda sınıflama doęruluęu deęerlerinin ,81 ile ,91; madde ayırt edicilik dzeyinin dřk olduęu kořulda ise ,59 ile ,68 arasında deęiřtięi grlmektedir. “Ortalama nitelik doęru sınıflama oranı” deęerleri incelendięinde ise madde ayırt edicilik dzeyinin yksek olduęu kořulda bu deęerlerin ,86 ile ,88; madde ayırt edicilik dzeyinin dřk olduęu kořullarda ise ,60 ile ,65 arasında deęiřtięi grlmektedir. DINA modelde test uzunluęunun 60 madde olduęu

koşulla test uzunluğunun 20 ve 40 madde olduğu koşullar karşılaştırıldığında örneklem büyüklüğünden bağımsız olarak test uzunluğu artışının özellikle madde ayırt edicilik düzeyinin yüksek olduğu koşul dikkate alındığında DINA modelin sınıflama doğruluğu performansına önemli katkılar sunduğu ve model performansının arttığı görülmektedir. DINA modelde 250 ve 500 örneklem büyüklüğü koşullarında “ortalama nitelik doğru sınıflama oranı” değerlerinde artışlar olsa da örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda bu değerlerin modelin performansına ek bir katkı sağlamadığı görülmektedir. Her ne kadar test uzunluğu faktöründeki artışla birlikte DINA model performansı artsa da örneklem büyüklüğü ve madde ayırt edicilik düzeyi faktörlerinde *-örneğin B₂ beceri düzeyinde örneklem büyüklüğünün 250 olduğu koşul hariç-* SISM’in, DINA modelden daha yüksek sınıflama doğruluğu değerlerine sahip olduğu görülmektedir. 6 nitelikli Q-matrisle 8 nitelikli Q-matrisin kullanıldığı koşullar karşılaştırıldığında DINA modelin ölçülen nitelik sayısının artmasıyla sınıflama doğruluğu performansının azaldığı bulgusuna ulaşılmıştır. Ayrıca aynı koşullar altında SISM ile karşılaştırıldığında DINA modelin sınıflama doğruluğu performansının, ölçülen nitelik sayısından daha fazla etkilendiği söylenebilir.

10 niteliği ölçen Q-matris kullanıldığında kavram yanılgılarını ölçen BUG-DINO modelinin 60 maddelik test uzunluğunda örneklem büyüklüğünden bağımsız ve madde ayırt edicilik düzeyinin yüksek olduğu koşulda kavram yanılgısı düzeyinde (K₁, K₂, K₃, K₄, K₅) “nitelik doğru sınıflama oranı” değerleri ,95 ile ,97; madde ayırt edicilik düzeyinin düşük olduğu koşulda ise ,90 ile ,93 arasında değişmektedir. Modele ait “ortalama nitelik doğru sınıflama oranı” değerleri incelendiğinde ise bu değerlerin madde ayırt edicilik düzeyinin yüksek olduğu koşulda ,96 olduğu ve madde ayırt edicilik düzeyinin düşük olduğu koşulda ,91 ile ,92 arasında değiştiği görülmektedir. Bu durum madde ayırt edicilik düzeyindeki artışla birlikte sınıflama doğruluğu değerlerinin arttığının bir göstergesidir. BUG-DINO’nun madde ayırt edicilik düzeyinin yüksek olduğu koşul kapsamında, elde edilen sınıflama doğruluğu değerlerinin örneklem büyüklüğünden etkilenmediği bulgusuna ulaşılmıştır. Madde ayırt edicilik düzeyinin yüksek olduğu koşulda örneklem büyüklüğü artışıyla birlikte BUG-DINO model SISM ile karşılaştırıldığında “nitelik doğru sınıflama oranı” değerlerinin kavram yanılgısı düzeyinde birbirine yakın değerler aldığı görülse de genel olarak değerlendirildiğinde farklı koşullarda bazı kavram yanılgılarında SISM’in sınıflama doğruluğu performansının daha yüksek olduğu görülmektedir. Madde ayırt edicilik düzeyinin düşük olduğu koşulda da “nitelik doğru sınıflama oranı” değerleri kavram

yanılgısı düzeyinde -örneğin K_1 kavram yanılgısı düzeyinde örneklem büyüklüğünün 250 olduğu koşul hariç- birbirine yakın değerler olsa da farklı koşullarda bazı kavram yanılgılarında SISM'in genel anlamda sınıflama doğruluğu performansının daha yüksek olduğu görülmektedir. Ayrıca 10 nitelikli Q-matrisin kullanıldığı koşulda test uzunluğunu 20 ve 40 maddeden 60 maddeye çıkarmanın, madde ayırt edicilik düzeyinin yüksek ve düşük olduğu her iki koşulda da BUG-DINO'nun sınıflama doğruluğu performansını arttırdığı görülmektedir. Ölçülen nitelik sayısının artmasıyla tüm koşullarda BUG-DINO'nun sınıflama doğruluğu performansı 6 ve 8 nitelikli Q-matrislere göre azalsa da modelin çok küçük örneklemelerde dahi kavram yanılgılarını ölçmede halen ,90 üzeri sınıflama doğruluğu değerleri alarak yüksek bir performans sergilediği söylenebilir.

Tüm tablolar genel olarak değerlendirildiğinde 6, 8 ve 10 nitelikli Q-matrisler kullanıldığında DINA ve BUG-DINO modelleri için en yüksek sınıflama doğruluğu değerlerinin, 6 nitelikli Q-matrisin kullanıldığı 60 maddelik bir test için örneklem büyüklüğü artışıyla birlikte madde ayırt edicilik düzeyinin yüksek olduğu koşuldanda elde edildiği görülmektedir. SISM'de ise 6 ve 8 nitelikli Q-matrislerin kullanıldığı koşullarda 60 maddelik bir test için örneklem büyüklüğü artışıyla birlikte madde ayırt edicilik düzeyinin yüksek olduğu koşullar birbirine benzerdir. Her iki Q-matris için de SISM için en yüksek sınıflama doğruluğu performansı bu koşullar altında elde edilmiştir.

BÖLÜM V

SONUÇ VE TARTIŞMA

Bu bölümde çalışmanın sonuçlarıyla birlikte uygulayıcılara ve araştırmacılara yönelik önerilere değinilmiştir.

5.1. Sonuçlar

Çalışma kapsamında SISM'in performansının, farklı simülasyon koşulları altında karşılaştırılarak hangi koşullarda benzer ya da farklı sonuçlar verdiğinin belirlenmesi amaçlanmıştır. Araştırmanın birinci alt problemi kapsamında SISM'e göre tahmin edilen madde parametrelerine ait "mutlak sapma" ve "ortalama mutlak sapma" değerleriyle birlikte araştırmanın ikinci alt problemi kapsamında tahmin hatasının büyüklüğünü hesaplamak için "hataların ortalama karekökü" değerleri örneklem büyüklüğü, test uzunluğu, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısına göre dört faktörde değişimlenen 72 koşul altında karşılaştırılarak modelin genel performansı değerlendirilmiştir. Araştırmanın birinci ve ikinci alt problemlerine ek olarak ilgili bölümün sonunda dört madde parametresine ilişkin ortalama standart hata değerleri de tüm koşullar için özetlenerek sunulmuştur. Becerileri ve kavram yanılgılarını eş zamanlı olarak ölçen bu modelin, kavram yanılgıları ve beceriler için elde edilen hata terimlerinde yer alan ilişkiyi dikkate almasıyla becerilerin ve kavram yanılgılarının farklı modellerle ayrı ayrı tanımlanmalarından kaynaklı olarak ortaya çıkabilecek bilgi kayıplarının önüne geçmesi ve madde parametrelerinde daha az hata üretmesi beklenmektedir. Bu nedenle çalışmada modelin genel performansının değerlendirilmesine ek olarak, araştırmanın üçüncü alt problemi kapsamında SISM'in beceri ve kavram yanılgılarını ayrı ayrı ölçen DINA ve BUG-DINO'ya kıyasla sınıflama

performansında artışlar olması beklendiğinden üç modelin sınıflama doğruluğu değerlerinden elde edilen sonuçlar faktörlerde değişimlenen koşullar için özetlenmiştir. Çalışma kapsamında performansı değerlendirilen SISM'e yönelik literatürde sınırlı sayıda çalışma (Kuo vd., 2018) yer aldığından bu çalışmaya ek olarak, bilişsel tanı değerlendirmeleri literatüründe yer alan modeller kapsamında simülasyon koşullarından elde edilen bulgular genel olarak tartışılmış ve sonuç-tartışma bölümü, alt problemler kapsamında maddeler halinde özetlenmiştir.

5.1.1. Birinci Alt Probleme İlişkin Sonuç ve Tartışma

- Birinci alt problem kapsamında 6, 8 ve 10 niteliği içeren Q-matrisler kullanıldığında tablolarda yer alan “mutlak sapma” değerleri genel olarak değerlendirildiğinde, h_j madde parametresi diğer üç madde parametresine göre daha yüksek “mutlak sapma” değerleri almıştır. Test uzunluğu, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısı sabit tutulduğunda “mutlak sapma” değerleri örneklem büyüklüğünün 250 olduğu koşulda artmış ve örneklem büyüklüğü arttıkça monoton olmayan artış ya da azalışlar olsa da genel olarak h_j madde parametresinde sapmalar azalmıştır. Örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda ise bu parametreye ait sapmaların benzer olduğu görülmüştür. g_j madde parametresine ait “mutlak sapma” değerleri örneklem büyüklüğü arttıkça monoton olmayan artış ve azalışlar sergilemiştir. Ancak sapmalar madde ayırt edicilik düzeyinin yüksek olduğu koşulda, madde ayırt edicilik düzeyinin düşük olduğu koşula kıyasla daha azdır. Benzer bir durum ω_j ve ϵ_j madde parametreleri için de geçerlidir. Genel olarak değerlendirildiğinde örneklem büyüklüğünden bağımsız olarak madde ayırt edicilik düzeyinin yüksek ve düşük olduğu koşullarda, madde sayısını arttırmak “mutlak sapma” değerlerini azaltırken; ölçülen nitelik sayısını arttırmak maddelere ait “mutlak sapma” değerlerini arttırmıştır. Kuo vd. (2018) tarafından yapılan çalışmada yer alan “mutlak sapma” değerlerine ait tablolar incelendiğinde örneklem büyüklüğü artışıyla, özellikle madde ayırt edicilik düzeyinin düşük olduğu koşulda çok küçük de olsa monoton olmayan artış ya da azalışların olduğu görülmektedir. Çalışmadan elde edilen bulgular Kuo vd. (2018) tarafından yapılan çalışmanın bulgularıyla tutarlıdır.

- Çalışmada mutlak sapma değerlerinin ortalamasını veren “ortalama mutlak sapma” değerlerine ait sonuçlar 6, 8 ve 10 nitelikli Q-matrislerin kullanıldığı koşullarda SISM’e ait dört madde parametresi için diğer koşullar sabit tutulduğunda, örneklem büyüklüğü artışıyla birlikte monoton olmayan artış ya da azalışlar sergilese de bu madde parametrelerine ait “ortalama mutlak sapma” değerlerinin genel olarak azaldığını göstermiştir. Ancak bu azalışların g_j , ω_j ve ϵ_j olmak üzere üç madde parametresinde, örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda oldukça benzer olduğu ve modele ek bir katkı sağlamadığı söylenebilir. Bu bulgu Kuo vd. (2018) tarafından yapılan çalışmadan elde edilen bulgularla tutarlıdır. Araştırmacıların çalışmalarında örneklem büyüklüğünün 1000 ve 2000 olduğu koşullarda madde parametrelerine ait “ortalama mutlak sapma” değerlerinin değişmeyerek, modele ek bir katkı sağlamadığı görülmektedir.
- Test uzunluğu faktörü genel olarak değerlendirildiğinde h_j madde parametresi için bazı durumlar hariç (örneğin “*test uzunluğu 60*” ve “*madde ayırt edicilik düzeyi düşük*”) “ortalama mutlak sapma” değerlerinin azaldığı görülmüştür. Değerlerin genel olarak daha yüksek madde ayırt edicilik düzeyiyle azalırken madde ayırt edicilik düzeyinin düştüğü koşulda artış göstererek madde parametre tahminlerinin arttığı sonucuna ulaşılmıştır. g_j , ω_j ve ϵ_j madde parametrelerine ait değerlerin ise bazı koşullarda küçük değişimlerle de olsa monoton olmayan artış ya da azalışlar sergilediği ancak bu değerlerin genel olarak azaldığı görülmüştür. Özellikle 10 nitelikli Q-matrisin kullanıldığı koşula kıyasla 6 ve 8 nitelikli Q-matrislerin kullanıldığı test uzunluğunun 20 maddeden 40 maddeye çıkarıldığı koşullarda, dört madde parametresine ait değerlerde sapmalara ilişkin azalışların daha fazla olduğu görülmüştür. Bu çalışmanın bulguları Chen ve de la Torre (2013) tarafından gerçekleştirilen çalışmada test uzunluğunun artırılıp daha sade Q-matrislerin kullanıldığı koşullarda “ortalama mutlak sapma” değerlerinin azaldığını vurguladıkları bulgularla tutarlıdır. Ayrıca Kuo vd. (2018)’nin çalışmasında yer alan “ortalama mutlak sapma” değerleri incelendiğinde bu değerler faktörlerde değişimlenen koşullar altında ,00 ile ,01 arasında değişmiştir. Araştırmacılar bu çalışmanın bulgularıyla benzer şekilde test uzunluğunun 20 maddeden 40 maddeye çıkarılmasının madde parametrelerine ait “ortalama mutlak sapma” değerlerini azalttığını raporlamışlardır.

- Çalışmada test uzunluğunun 40 maddeden 60 maddeye çıkarıldığı koşulda ise madde parametre tahminlerine ilişkin değerlerde azalışlar olsa da sonuçlar, bu azalışların oldukça küçük olduğunu göstermiştir. İstisnai bir durum 10 nitelikli Q-matris için ortaya çıkmıştır. Bu matrisin kullanıldığı koşulda test uzunluğu faktörü incelendiğinde madde ayırt edicilik düzeyinin yüksek ve düşük olduğu her iki koşulda da değişimlenen 20, 40 ve 60 maddelik test uzunluklarında madde sayısının artmasıyla birlikte madde parametrelerine ait “ortalama mutlak sapma” değerlerinin h_j , g_j ve ω_j parametreleri için azaldığı ancak ϵ_j madde parametresine ait “ortalama mutlak sapma” değerlerinin 20 maddelik test uzunluğunda daha düşük sapmalar sergilediği görülmüştür. Hem madde sayısının hem de ölçülen nitelik sayısının artmasıyla birlikte modele eklenen yeni parametrelerin tahmin edilecek parametre sayısını arttırması yönünden bu durum ϵ_j parametresine ait “ortalama mutlak sapma” değerlerinde olumsuz bir etki yaratmış olabilir. Benzer bir durum Elbulok (2021) tarafından gerçekleştirilen çalışmanın bulgularından elde edilmiştir. Araştırmacı MC-DINA-H modelinin performansını değerlendirdiği çalışmada test uzunluğu faktöründeki koşulu 15 maddeden 30 maddeye çıkarmanın, elde edilen “ortalama mutlak sapma” değerlerini arttırdığını raporlamıştır.
- Çalışma kapsamında yer alan bir diğer faktör, ölçülen nitelik sayısıdır. Ölçülen nitelik sayısının artmasıyla birlikte madde parametrelerine ilişkin “ortalama mutlak sapma” değerlerine ait tahminlerin arttığı görülmüştür. Ancak ölçülen nitelik sayısının artmasıyla birlikte 6 ve 8 nitelikli Q-matrislerin kullanıldığı koşullardan elde edilen madde parametrelerine ait değerlere ilişkin, h_j madde parametresine ait değerlerdeki artışların g_j , ω_j ve ϵ_j madde parametrelerine göre nispeten daha fazla olduğu görülmüştür. g_j , ω_j ve ϵ_j madde parametrelerine ait değerler birbirlerine daha yakınken, 10 nitelikli Q-matrisin kullanıldığı koşulda madde parametrelerinden elde edilen “ortalama mutlak sapma” değerleri özellikle h_j madde parametresinde 6 ve 8 nitelikli Q-matrislerin kullanıldığı koşullara kıyasla daha fazla artmış ve üç madde parametresi bu durumlardan daha küçük artışlarla etkilenmiştir.

Genel olarak SISM'in dört madde parametresindeki azalışları "ortalama mutlak sapma" değerleriyle her bir koşul için karşılaştırıldığında; h_j madde parametresi için "6 nitelikli Q-matris*madde ayırt edicilik düzeyi yüksek*test uzunluğu 40 madde*örneklem büyüklüğü 3000", g_j madde parametresi için "6 nitelikli Q-matris*madde ayırt edicilik düzeyi yüksek*test uzunluğu 40 madde*örneklem büyüklüğü 3000", ω_j madde parametresi için "6 nitelikli Q-matris*madde ayırt edicilik düzeyi yüksek*test uzunluğu 40 madde*örneklem büyüklüğü 1500" ve ϵ_j madde parametresi için "8 nitelikli Q-matris*madde ayırt edicilik düzeyi yüksek*test uzunluğu 40 madde*örneklem büyüklüğü 1500" koşulu için daha küçük "ortalama mutlak sapma" değerleriyle sonuçlandığı gözlenmiştir. 6, 8 ve 10 nitelikli Q-matrislerdeki bu parametrelere ait "ortalama mutlak sapma" değerlerindeki artışlar ölçülen nitelik sayısının artmasıyla ilişkilendirilebilir. 6 ve 8 nitelikli Q-matrislerin kullanıldığı koşullardan elde edilen değerler daha yakın olsa da genel olarak değerlendirildiğinde 8 ve 10 nitelikli Q-matrislerin kullanıldığı koşullara kıyasla 6 nitelikli Q-matrisin kullanıldığı koşulda, SISM daha düşük sapma değerleriyle kestirim yapmıştır. Ayrıca tüm koşullar için "ortalama mutlak sapma" değerleri incelendiğinde madde sayısını 40 maddeden 60 maddeye çıkarmanın madde parametre tahminlerine ait değerlerde ek bir azalış sağlamadığı söylenebilir. Dört madde parametresi için örneklem büyüklüğü faktörü tüm koşullarda değerlendirildiğinde nispeten daha büyük olan örneklem büyüklüğü koşullarında (1500 ve 3000) bu parametrelerin daha düşük "ortalama mutlak sapma" değerleri aldığı görülmüştür. Ayrıca madde ayırt edicilik düzeyinin yüksek olduğu koşulda ise madde parametrelerine ait "ortalama mutlak sapma" değerlerinin azaldığı söylenebilir.

Literatürde yer alan bilişsel tanı modelleri genel olarak incelendiğinde, modellerin özellikle Q-matriste yer alan nitelik sayısındaki artışlarla ölçülen nitelik sayısından etkilendiği görülmüştür (Bradshaw & Templin, 2014; Schellman, 2021; Sen & Cohen, 2021). Örneğin Romero, Ordoñez, Ponsoa ve Revuelta (2014) tasarladıkları simülasyon çalışmasını gerçek veri analiziyle destekleyerek "En Küçük Kareler Uzaklık Modeli" altında Q-matris hatalı tanımlamalarını belirlemek amacıyla "ortalama mutlak sapma" değerlerini incelemiş ve bu değerlerin ölçülen nitelik sayısının artmasından olumsuz yönde etkilendiğini ifade etmişlerdir. Elde edilen bulgular bu çalışmanın bulgularıyla tutarlıdır.

5.1.2. İkinci Alt Probleme İlişkin Sonuç ve Tartışma

- İkinci alt problem kapsamında değişimlenen faktörlerin madde parametrelerine ait tahminlerin doğruluğu üzerindeki etkisini incelemek amacıyla 6, 8 ve 10 niteliği içeren Q-matrislerin kullanıldığı koşullarda “hataların ortalama karekökü” değerleri, “ortalama mutlak sapma” değerlerinde monoton olmayan küçük artış ya da azalışların var olması sebebiyle incelenmiştir. Test uzunluğu faktöründe değişimlenen 20, 40 ve 60 maddelik koşullar için test uzunluğu arttıkça madde parametrelerine ait bu değerlerin, madde ayırt edicilik düzeyinin düşük olduğu bazı koşullarda monoton olmayan artış ya da azalışlar sergilese de madde ayırt edicilik düzeyinin yüksek olduğu tüm koşullarda azaldığı görülmüştür. Ardıç (2020) çalışmasında, simülasyon ve gerçek veri analizleriyle birlikte bilişsel tanı modelleri ile üç boyutlu madde tepki kuramı modelinin sınıflama doğruluğu performanslarını incelemiştir. Çeşitli simülasyon koşulları altında, test uzunluğunun artmasıyla “hataların ortalama karekökü” değerlerinin azaldığı bulgusuna ulaşılmıştır. Ardıç (2020) tarafından yapılan çalışmaya ek olarak literatürde yer alan çalışmalar test uzunluğunun artmasıyla “hataların ortalama karekökü” değerlerinin azaldığını raporlamıştır (Bradshaw & Templin, 2014; Chen & de la Torre, 2018; Sen & Cohen, 2021; Wang, Song, Cheng & Ding, 2019). Çalışmanın bulguları, literatürde yer alan bilişsel tanı modellerine ait çalışmaların bulgularıyla tutarlıdır.
- Test uzunluğuyla benzer bir durum örneklem büyüklüğü koşulu için de geçerlidir. Madde ayırt edicilik düzeyinin yüksek olduğu koşulda küçük örneklemelerden büyük örneklemelere artışla birlikte “hataların ortalama karekökü” değerlerinde azalışlar olduğu görülmüştür. Bu bulgu Yamaguchi (2020) tarafından kullanılan DINA modelinin çoktan seçmeli madde versiyonu için (MC-DINA) Bayes formülasyonu önerdiği çalışmasında değişimlenen test uzunluğu ve örneklem büyüklüğünün arttığı koşullarda “hataların ortalama karekökü” değerlerinin azaldığı bulgularla tutarlıdır. Sen ve Cohen (2021) çalışmalarında, bilişsel tanı değerlendirmeleri kapsamında kullanılan modellerden DINA, DINO ve C-RUM modellerinin örneklem büyüklüğü artışıyla birlikte madde parametrelerine ait “hataların ortalama karekökü” değerlerinin azaldığını vurgulamışlardır. Ayrıca Huang (2017) çalışmasında daha uzun testler ve daha büyük örneklem büyüklükleriyle madde parametre tahminlerinin azaldığını ifade etmiştir. Ancak bu çalışmada örneklem büyüklüğünün 1500’den 3000’e çıkarıldığı koşulda dört madde parametre tahminine ait “hataların ortalama

karekökü” değerlerinin aynı kalarak ek bir azalma olmadığı sonucuna ulaşılmıştır. Ayrıca “hataların ortalama karekökü” değerlerinin daha yüksek madde ayırt edicilik düzeyiyle azalırken, madde ayırt edicilik düzeyinin düştüğü koşulda artış gösterdiği görülmüştür. Çalışmanın bulguları bu yönüyle Akbay (2016)’ın çalışma bulgularıyla tutarlıdır. Çalışmada örneklem büyüklüğünden bağımsız olarak madde ayırt edicilik düzeyi düşükçe “hataların ortalama karekökü” değerlerinin arttığı bulgusuna ulaşılmıştır. Literatürde yer alan çalışmalarla desteklendiği gibi bu çalışmadan elde edilen bulgular madde ayırt edicilik düzeyi faktörünün “hataların ortalama karekökü” değerleri üzerinde etkisi olduğunu ortaya koymaktadır.

- Çalışmada diğer faktörlerden bağımsız olarak ölçülen nitelik sayısının arttırıldığı koşullarda “hataların ortalama karekökü” değerlerinde artışların olduğu sonucuna ulaşılmıştır. Bu kapsamda SISM’in dört madde parametresi tahminlerine ait değerlerdeki azalışları her bir koşul için karşılaştırıldığında h_j, g_j, ω_j ve ϵ_j madde parametreleri için “6 nitelikli Q-matris*madde ayırt edicilik düzeyi yüksek*test uzunluğu 40 madde*örneklem büyüklüğü 1500” koşulunda daha küçük “hataların ortalama karekökü” değerleriyle sonuçlandırıldığı gözlenmiştir. 6, 8 ve 10 nitelikli Q-matrislerdeki bu parametrelere ait hata değerlerindeki artışlar ölçülen nitelik sayısının artmasıyla ilişkilendirilebilir. Yani model 6 nitelikli Q-matrisin kullanıldığı koşulda en düşük hatayla kestirim yapmıştır. Çalışmanın bulguları Sen ve Cohen (2021) tarafından gerçekleştirilen çalışma bulgularıyla tutarlıdır. Araştırmacılar ölçülen nitelik sayısının arttırıldığı koşulda “hataların ortalama karekökü” değerlerinde sapmaların arttığını vurgulamışlardır.
- Çalışmada SISM’e ait dört madde parametresi için farklı faktörlerde değişimlenen koşullardan elde edilen bulgular, “hataların ortalama karekökü” ve “ortalama mutlak sapma” değerlerinin örneklem büyüklüğü ve ölçülen nitelik sayısı bakımından farklılıklar gösterdiğini ortaya koymuştur. “Hataların ortalama karekökü” değerleri “ortalama mutlak sapma değerlerine göre daha büyük değerler almıştır. Bu durum 6, 8 ve 10 nitelikli Q-matrislerdeki karmaşıklığın artmasıyla birlikte “ortalama mutlak sapma” değerlerindeki artışlarla ya da tahmin varyansının büyüklüğüyle ilişkilendirilebilir (DeMars, 2003). “Hataların ortalama karekökü” değerleri daha büyük değerler olsa da dört madde parametresi için genel olarak “ortalama mutlak sapma” değerlerinden daha tutarlı kestirimler yapmıştır. Bu bulgu Sen ve Cohen (2021) ile Harwel (2018) tarafından gerçekleştirilen çalışmanın bulgularıyla

tutarlıdır. Ayrıca bu iki değer madde ayırt edicilik düzeyinin düşük olduğu test uzunluğu koşulu (20 madde) altında monoton olmayan artış ya da azalışlar sergilemiştir. Bu bulgu beceri ve kavram yanılgılarının tanımlanması için Elbulok (2021) tarafından geliştirilen MC-DINA-H modelinin performansının değerlendirildiği çalışmanın bulgularıyla tutarlıdır. Araştırmacı test uzunluğunun 15 maddeden 30 maddeye çıkarıldığı koşulda “ortalama mutlak sapma” değerleriyle “hataların ortalama karekökü” değerlerinin arttığını ancak test uzunluğunun 30 maddeden 60 maddeye çıkarıldığı koşulda bu değerlerin azaldığını vurgulamıştır. Elbulok (2021) tarafından gerçekleştirilen çalışmada vurgulandığı gibi madde ayırt edicilik düzeyinin düşük olduğu koşulda test uzunluğu artışıyla meydana gelen bu monoton olmama durumu modele yeni parametrelerin eklenerek tahmin edilecek parametre sayısını artırması yönünden olumsuz etki ya da bireyleri ilgili niteliklere sahip olma noktasında daha iyi sınıflandırabilmek için madde artışının modele daha çok veri sağlaması yönünden olumlu olmak üzere iki farklı etkiye sahip olmasından kaynaklanabilir.

- 1. ve 2. araştırma problemlerinde dört madde parametresine ait “ortalama mutlak sapma” ve “hataların ortalama karekökü” değerleri incelendiğinde tahminlerin doğruluğuna ilişkin değerlerin, diğer faktörlere kıyasla göreceli olarak en çok ölçülen nitelik sayısından etkilendiği söylenebilir. Bu çalışmanın bulguları Bradshaw ve Templin (2014) tarafından geliştirilen Hiyerarşik Tanı Sınıflama Model’inin literatüre kazandırıldığı çalışmanın bulgularıyla tutarlıdır. Araştırmacıların çalışmalarında ifade ettikleri gibi testte yer alan nitelik sayısının artışıyla modelde kullanılan Q-matrisin karmaşıklığı ve testin boyutsallığının artmasıyla birlikte yapısal modelin karmaşıklığı da artacağından bu beklenen bir durumdur. Ayrıca bu çalışma kapsamında yer alan madde ayırt edicilik düzeyi ve test uzunluğu faktörlerinin, madde parametrelerine ilişkin tahminlerin doğruluğunda örneklem büyüklüğüne göre göreceli olarak daha etkili olduğu söylenebilir. Buradan örneklem büyüklüğünün önemsiz olduğu anlamı çıkarılmamalıdır. Modelden elde edilen parametrelerin tutarlı tahminler yapması ve modelin tanımlanabilir olabilmesi için yeterli verinin araştırmacılar tarafından toplanması gerekmektedir (Sen & Cohen, 2021).

Çalışmanın birinci ve ikinci alt problemi kapsamında “mutlak sapma”, “ortalama mutlak sapma” ve “hataların ortalama karekökü” değerleri incelendikten sonra 6, 8 ve 10 nitelikli Q-matrisler için faktörlerde değişimlenen her bir koşulda dört madde parametresine ait tüm maddeler arasındaki ortalama standart hata değerleri özetlenmiştir. Dört madde parametresi için standart hata değerlerinin sunulmasının amacı standart hata değerlerinin 0'a yakınlığı ölçüsünde madde parametrelerine ait kestirimlerin doğru olmasıdır (Rupp vd. 2010).

- Ortalama standart hata değerleri incelendiğinde madde ayırt edicilik düzeyi dışındaki diğer tüm faktörler sabit tutulduğunda; madde ayırt edicilik düzeyinin düşük olduğu koşulda, üç Q-matris için de standart hata değerlerinin arttığı görülmüştür.
- Dört madde parametresi incelendiğinde h_j madde parametresinin standart hata değerleri örneklem büyüklüğünün 250 olduğu koşulda daha büyüktür. Özellikle madde ayırt edicilik düzeyinin düşük olduğu koşullarda test uzunluğundaki artışla birlikte standart hata değerleri artmıştır. Madde ayırt edicilik düzeyinin yüksek ve örneklem büyüklüğünün 250 olduğu koşullarda h_j madde parametresine ait standart hata değerlerinde test uzunluğunun artmasıyla birlikte madde ayırt edicilik düzeyinin düşük olduğu koşul kadar olmasa da artışların olduğu görülmüştür. Üç Q-matriste de h_j madde parametresine ait en büyük standart hata değerleri “*örneklem büyüklüğü 250*madde ayırt edicilik düzeyi düşük*test uzunluğu 60*” koşulu altında elde edilmiştir. Örneklem büyüklüğünün 1500 ve 3000 olduğu koşullarda ise h_j madde parametresine ait standart hata değerlerinin oldukça azalarak 0'a yaklaştığı ve madde ayırt edicilik düzeyinin yüksek olduğu koşulda, madde ayırt edicilik düzeyinin düşük olduğu koşula göre daha küçük değerler aldığı görülmüştür. Ayrıca 1500 ve 3000 örneklem büyüklüklerinde, standart hata değerlerinin en küçük değerleri aldığı ve test uzunluğundaki artışla birlikte standart hata değerlerinin üç Q-matrisin kullanıldığı koşullarda benzer olduğu sonucuna ulaşılmıştır. Üç Q-matriste g_j , ω_j ve ϵ_j madde parametreleri incelendiğinde örneklem büyüklüğünün 250 olduğu koşulda diğer örneklem büyüklüklerine kıyasla standart hata değerlerinin daha büyük olduğu görülmüştür. Bu durum özellikle madde ayırt edicilik düzeyinin düşük olduğu koşulda test uzunluğundaki artışla birlikte ortaya çıkmıştır. Her üç madde parametresi için en büyük standart hata değeri “*örneklem büyüklüğü 250*madde ayırt edicilik düzeyi düşük*test uzunluğu 60*” koşulu altında elde edilmiştir. g_j , ω_j ve ϵ_j madde parametreleri için standart hata değerleri örneklem büyüklüğünün 1500 ve 3000'e çıkarıldığı koşullarda azalma eğilimi göstermiştir. Ancak nispeten daha

büyük olan iki örneklem büyüklüğü koşulu için madde ayırt edicilik düzeyinin yüksek ve düşük olduğu her iki koşulda da test uzunluğundaki artışların standart hata değerlerini etkilemediği sonucuna ulaşılmıştır.

- 10 nitelikli Q-matrisin kullanıldığı koşuldan elde edilen bulgular 6 ve 8 nitelikli Q-matrisin kullanıldığı koşullarla benzerlik göstermiştir. Madde ayırt edicilik düzeyi, örneklem büyüklüğü ve test uzunluğundaki artış üç Q-matris koşulu için de örneklem büyüklüğünün 250 olduğu koşul haricinde, daha küçük standart hata değerlerini ortaya koymuştur. Ölçülen nitelik sayısı arttıkça özellikle örneklem büyüklüğünün 250 olduğu koşulda test uzunluğunun artmasıyla birlikte ortalama standart hata değerlerinin arttığı sonucuna ulaşılmıştır. Standart hataların madde ayırt edicilik düzeyinin yüksek ve örneklem büyüklüğünün 1500 ve 3000 olduğu koşullar altında daha uzun test koşullarında güvenilir olduğu söylenebilir.

Özetle çalışmanın sonuçları “ortalama mutlak sapma” ile “hataların ortalama karekökü” ve madde parametrelerine karşılık gelen ortalama standart hata değerleri dikkate alınarak, test uzunluğuyla örneklem büyüklüğünün nispeten arttığı ve madde ayırt edicilik düzeyinin yüksek olduğu koşullar incelendiğinde beklenti maksimizasyon algoritması kullanılarak gerçekleştirilen dört madde parametresine ait tahminlerin bu algoritmayla doğru bir şekilde tahmin edilebildiğini göstermiştir. Kuo vd. (2018) tarafından gerçekleştirilen çalışmada vurgulandığı gibi modelden yapılabilecek çıkarımların geçerliği noktasında bu sonuç, yukarıda sıralanan koşullar altında modelin güvenilir olabileceğinin bir göstergesidir.

5.1.3. Üçüncü Alt Probleme İlişkin Sonuç ve Tartışma

- Beceri ve kavram yanlılığı düzeyinde sınıflama doğruluğu değerlerine ilişkin SISM'in genel performansı değerlendirildiğinde 6, 8 ve 10 nitelikli Q-matrislerin kullanıldığı koşullarda, örneklem büyüklüğü ve test uzunluğu faktörlerinden bağımsız olarak madde ayırt edicilik düzeyinin yüksek olduğu koşulda modelin beklendiği gibi sınıflama doğruluğu performansı artmıştır. Ancak ölçülen nitelik sayısı arttıkça modelin sınıflama doğruluğu performansı azalmıştır.
- 6 nitelikli Q-matrisin kullanıldığı 20 maddelik bir testte madde ayırt edicilik düzeyinin yüksek ve düşük olduğu koşullarda SISM'e ait “nitelik doğru sınıflama oranı” değerleri ile “örüntü doğru sınıflama oranı” değerleri örneklem büyüklüğü artışıyla birlikte artmıştır. Model madde ayırt edicilik düzeyinin yüksek olduğu

koşulda $-B_1$ becerisi hariç, ,80 üzerinde sınıflama doğruluğu değerleri almıştır. Örneklem büyüklüğü artışıyla birlikte “ortalama nitelik doğru sınıflama oranı” değerlerinin örneklem büyüklüğünün 1500’den 3000’e çıktığı koşulda arttığı görülmüştür. Genel olarak değerlendirildiğinde test uzunluğunun kısa (20 madde) olduğu küçük bir örneklem büyüklüğüne rağmen madde ayırt edicilik düzeyinin yüksek olduğu koşulda modelin iyi sınıflama doğruluğu performansı sergilediği söylenebilir. Ayrıca örneklem büyüklüğünden bağımsız olarak madde ayırt edicilik düzeyinin yüksek olduğu ve test uzunluğunun arttığı koşullarda (40 ve 60 madde) modele ait sınıflama doğruluğu performansının gittikçe arttığı görülmüştür. Test uzunluğunun 60 madde olduğu koşulda ise tüm örneklem büyüklüklerinde, “ortalama nitelik doğru sınıflama oranı” değerleri aynıdır. Bu koşul altında örneklem büyüklüğü artışının modelin sınıflama doğruluğu performansına ek bir katkı sağlamadığı söylenebilir.

- 8 nitelikli Q-matrisin kullanıldığı 20 maddelik bir testte madde ayırt edicilik düzeyinin yüksek ve düşük olduğu koşullarda SISM’e ait “nitelik doğru sınıflama oranı” değerleri örneklem büyüklüğü artışıyla birlikte genel olarak artmıştır. Ancak sınıflama doğruluğu performansı 6 nitelikli Q-matrisin kullanıldığı koşula göre azalma eğilimi göstermiştir. Örneklem büyüklüğü artışıyla birlikte “ortalama nitelik doğru sınıflama oranı” değerleri örneklem büyüklüğünün 250 ve 500 olduğu koşullarda aynı değerleri almıştır. Örneklem büyüklüğünün 1500’e çıkarıldığı koşulda modelin sınıflama doğruluğu performansı artmıştır. Örneklem büyüklüğünün 3000 olduğu koşul ise örneklem büyüklüğü faktörü model performansına ek bir katkı sağlamamıştır. “Örüntü doğru sınıflama oranı” değerlerinin de örneklem büyüklüğü artışıyla birlikte arttığı ancak ölçülen nitelik sayısının artması sebebiyle madde ayırt edicilik düzeyinin yüksek ve düşük olduğu her iki koşulda da bu değerlerin 6 nitelikli Q-matrisin kullanıldığı koşula göre azalarak, modele ait sınıflama doğruluğu performansının azaldığı görülmüştür. Genel olarak değerlendirildiğinde 6 nitelikli Q-matrisin kullanıldığı koşul kadar olmasa da test uzunluğunun kısa (20 madde) olduğu küçük bir örneklem büyüklüğüne rağmen madde ayırt edicilik düzeyinin yüksek olduğu koşulda modelin iyi bir sınıflama doğruluğu performansı sergilediği söylenebilir. Ölçülen nitelik sayısının artmasıyla azalan sınıflama doğruluğu performansı, örneklem büyüklüğünden bağımsız olarak madde ayırt edicilik düzeyinin yüksek olduğu ve

test uzunluğunun arttığı 40 maddelik bir testte ,80 üzerine çıkararak 8 nitelikli Q-matrisin kullanıldığı koşulda da iyi bir sınıflama doğruluğu performansı sergilemiştir. Madde sayısının 60 olduğu ve her iki Q-matrisin kullanıldığı koşullarda “ortalama nitelik doğru sınıflama oranı” değerleri aynıdır. Bu durum her iki Q-matriste de örneklem büyüklüğü artışının modelin sınıflama doğruluğu performansına ek bir katkı sağlamadığının bir göstergesidir. Test uzunluğundaki artışın, ölçülen nitelik sayısının artmasıyla modelde azalan sınıflama doğruluğu performansını tolere ettiği söylenebilir.

- 10 nitelikli Q-matrisin kullanıldığı 20 maddelik bir testte madde ayırt edicilik düzeyinin yüksek ve düşük olduğu koşullarda SISM’e ait “nitelik doğru sınıflama oranı” değerleri örneklem büyüklüğü artışıyla birlikte genel olarak değerlendirildiğinde, bu değerler küçük de olsa artışlar göstermiş ancak ölçülen nitelik sayısının giderek artması sebebiyle modele ait sınıflama doğruluğu performansının azaldığı görülmüştür. Örneklem büyüklüğünün artmasıyla birlikte 8 nitelikli Q-matrisin kullanıldığı koşula benzer şekilde sınıflama doğruluğu performansının arttığı görülse de bulgular örneklem büyüklüğünün 1500’den 3000’e çıkarılmasının modelin performansına ek bir katkı sağlamadığını göstermiştir. “Örüntü doğru sınıflama oranı” değerlerinin de örneklem büyüklüğü artışıyla birlikte arttığı ancak 6 ve 8 nitelikli Q-matrislerin kullanıldığı koşullara kıyasla madde ayırt edicilik düzeyinin yüksek ve düşük olduğu her iki koşulda da modele ait sınıflama doğruluğu performansının ciddi şekilde azaldığı görülmüştür. Tüm koşullar genel olarak değerlendirildiğinde “örüntü doğru sınıflama oranı” değerlerinin “nitelik doğru sınıflama oranı” ve “ortalama nitelik doğru sınıflama oranı” değerlerine göre ölçülen nitelik sayısından daha fazla etkilendiği söylenebilir. Elde edilen bulgu Chen vd. (2012) ve Schellman (2021) tarafından gerçekleştirilen çalışmaların bulgularıyla tutarlıdır. Schellman (2021) BUG-DINO modelle kavram yanılgılarını belirlediği çalışmada kavram yanılgısı düzeyinde elde edilen sınıflama doğruluğu değerlerinin örüntü doğru sınıflama oranı değerlerinden daha yüksek olduğunu raporlamıştır. Chen vd. (2012) ise “örüntü doğru sınıflama oranı” değerlerinin “nitelik doğru sınıflama oranı” değerlerinden daha düşük çıkmasını örüntünün testle ölçülen nispeten daha çok sayıda niteliğin bulunmasıyla ilişkilendirmiştir.

- Sınıflama doğruluğu değerleri ölçülen nitelik sayısından etkilense de test uzunluğunun kısa (20 madde) ve madde ayırt edicilik düzeyinin yüksek olduğu koşullarda SISM'in çok küçük bir örneklem büyüklüğüne sahipken bile ,79 ile ,90 arasında “nitelik doğru sınıflama oranı” değerleri alarak sınıflama doğruluğu performansının arttığı görülmüştür. Örneklem büyüklüğü artışı sonucu “nitelik doğru sınıflama oranı” değerlerindeki artışlarla birlikte “ortalama nitelik doğru sınıflama oranı” ve “örüntü doğru sınıflama oranı” değerlerinin de artarak modelin sınıflama doğruluğu performansının arttığı söylenebilir. Ölçülen nitelik sayısının artmasıyla azalan sınıflama doğruluğu performansı, örneklem büyüklüğünden bağımsız olarak madde ayırt edicilik düzeyinin yüksek olduğu ve test uzunluğunun arttığı koşullarda (40 ve 60 madde) iyi bir sınıflama doğruluğu performansı sergilemiştir. 6 ve 8 nitelikli Q-matrislerden farklı olarak özellikle test uzunluğu faktöründeki madde sayısının 40 maddeden 60 maddeye çıkarıldığı koşulda, örneklem büyüklüğü artışıyla birlikte modele ait sınıflama doğruluğu performansının arttığı görülmüştür. Test uzunluğundaki artışın, ölçülen nitelik sayısının artmasıyla modelde azalan sınıflama doğruluğu performansını tolere ettiği söylenebilir.
- DINA modele ait sınıflama doğruluğu performansı genel olarak değerlendirildiğinde ölçülen nitelik sayısından bağımsız olarak, madde ayırt edicilik düzeyinin yüksek olduğu koşulda örneklem büyüklüğü ve test uzunluğu faktörlerindeki artışlarla birlikte model performansının arttığı görülmektedir. Bu çalışmadan elde edilen bulgular test uzunluğunun, sınıflama doğruluğu üzerindeki etkisini inceleyen Ömür-Sünbül (2013) tarafından gerçekleştirilen çalışmanın bulgularıyla kısmen tutarlıdır. Araştırmacı tarafından DINA modele ait örneklem büyüklüğü faktörünün sınıflama doğruluğu değerlerini etkilemediği ancak test uzunluğu faktörünün sınıflama doğruluğu değerlerini etkilediği vurgulanmıştır. Bu çalışmadan elde edilen bulgular incelendiğinde ise DINA model kapsamında örneklem büyüklüğü faktörünün “nitelik doğru sınıflama oranı” değerleriyle “ortalama nitelik doğru sınıflama oranı” değerleri üzerinde etkisinin olduğu söylenebilir. 6, 8 ve 10 nitelikli Q-matrislerin kullanıldığı koşullarda ise DINA model ve SISM'in sınıflama doğruluğu performansı beceri düzeylerinde “nitelik doğru sınıflama oranı” değerleri dikkate alınarak incelendiğinde, DINA modelinin tüm koşullar altında genel olarak SISM'den daha düşük sınıflama doğruluğu performansı sergilediği görülmüştür. Çalışma bulguları Kuo vd. (2018) tarafından gerçekleştirilen çalışmanın bulgularıyla

tutarlıdır. Ayrıca DINA modelin kendi içindeki sınıflama doğruluğu performansının ölçülen nitelik sayısının artmasıyla birlikte azaldığı ancak daha sade Q-matrislerin kullanılarak test uzunluğunun arttırıldığı koşullarda sınıflama doğruluğu performansının arttığı bulgusuna ulaşılmıştır.

- BUG-DINO modele ait sınıflama doğruluğu performansı değerlendirildiğinde ölçülen nitelik sayısından bağımsız olarak, madde ayırt edicilik düzeyinin yüksek olduğu koşulda örneklem büyüklüğü ve test uzunluğu faktörlerindeki artışlarla birlikte modelin sınıflama doğruluğu performansı artarken ölçülen nitelik sayısının arttığı koşullarda modele ait sınıflama doğruluğu performansının azaldığı görülmüştür. Bu beklenen bir durumdur (Schellman, 2021). Üç farklı Q-matrisin kullanıldığı ve madde ayırt edicilik düzeyinin yüksek olduğu koşul genel olarak değerlendirildiğinde, 20 maddelik test koşulunda BUG-DINO model için en yüksek sınıflama doğruluğu değerlerinin 3000 örneklem büyüklüğünden elde edildiği görülmüştür. Test uzunluğunun 20 maddeden 40 maddeye çıkarıldığı koşulda ise örneklem büyüklüğünü 1500'den 3000'e çıkarmanın modelin sınıflama doğruluğu performansına ek bir katkı sağlamadığı görülmüştür. Son olarak test uzunluğunun 60 madde olduğu koşulda ise BUG-DINO modelinin sınıflama doğruluğu performansı tüm örneklem büyüklüğü koşullarında aynı kalarak çok küçük bir örneklem büyüklüğünde dahi yüksek sınıflama doğruluğu değerleri almıştır. Literatürde yer alan çalışmalarla desteklendiği gibi (Schellman, 2021) BUG-DINO modeli potansiyel olarak örneklem büyüklüğünün küçük olduğu koşullarda pratik kullanımlar için ilgi çekicidir. Çalışma sonucu dikkat çeken bir bulgu, ölçülen nitelik sayısının arttırıldığı tüm koşullarda BUG-DINO modelinin kavram yanlışlarını belirlemedeki sınıflama doğruluğu performansının, çok az miktarda SISM'den düşük olsa da neredeyse SISM kadar iyi olduğudur. Elde edilen bu bulgu BUG-DINO model dikkate alındığında yanlış model kullanımının kavram yanlışlığı düzeyinde elde edilen sınıflama doğruluklarını belirgin şekilde etkilemediğini göstermektedir. Bununla birlikte çalışma örneklem büyüklüğüne kıyasla test uzunluğundaki artışın sınıflama doğruluğu değerlerini daha fazla arttırdığını ortaya koymuştur. Bu bulgular Gao vd. (2017) tarafından gerçekleştirilen çalışmanın amaçlarından biri olan G-DINA modele göre üretilen verilere A-CDM ve DINA modelleri yerleştirilerek yanlış model kullanımının sınıflama doğruluğu performansına etkisini inceledikleri çalışmanın bulgularıyla tutarlıdır.

Özetle, üçüncü alt problem kapsamında elde edilen bulgular simülasyon çalışması sonucu doymuş model SISM'in dört faktörde değişimlenen koşullarda diğer modellere göre daha kararlı ve daha yüksek sınıflama doğruluğu performansı sergilediğini göstermiştir. Yani beceriler ve kavram yanılgılarının eş zamanlı olarak ölçülmesi, becerilerin ya da kavram yanılgılarının modelden çıkarılmasına kıyasla daha yüksek sınıflama doğruluğu değerleri sağlamaktadır. Bu bulgu temel niteliklerin dahil edilmediği ya da ilgili olmayan niteliklerin modellere dahil edildiği bilişsel tanı modellerinin, sınıflama doğruluğu değerlerini etkileyebileceğini göstermiştir (Elbulok, 2021; Rupp & Templin, 2008). Literatürde yer alan araştırmalar bilişsel tanı değerlendirmelerinde beceriler ve kavram yanılgılarının eş zamanlı olarak ölçülmesinin önemi vurgulamaktadır. Bu sayede eş zamanlı ölçmeler sadece becerileri ya da sadece kavram yanılgılarını ölçen modellerin kullanımıyla bireylerin başarılı ya da başarısız oldukları noktalarda gözden kaçırılacak stratejilerin belirlenmesi aşamasında araştırmacılara yardımcı olabilir (Elbulok, 2021).

Sınıflama doğruluğu performansı genel olarak değerlendirildiğinde SISM'i sırasıyla BUG-DINO ve DINA modeller takip etmiştir. Test uzunluğunun artmasıyla birlikte tüm modellere ait sınıflama doğruluğu değerleri artış göstermiştir. Bilişsel tanı modelleri literatüründe bu bulgu, Ardıç (2020) tarafından gerçekleştirilen çalışmada birey parametrelerinin sınıflama doğruluğu değerlerini bilişsel tanı modelleri ve çok boyutlu madde tepki kuramı modelleriyle tahmin ettiği ve modellerin test uzunluğundaki artışla birlikte sınıflama doğruluğu performansının arttığı bulgusuna ulaştığı çalışmayla tutarlıdır. Test uzunluğu artışının sınıflama doğruluğu performansındaki etkisine ek olarak Bradshaw ve Templin (2014) ile Wang, Song, Chen, Meng ve Ding (2015) çalışmalarında ölçülen nitelik sayısının da sınıflama doğruluğu performansını etkilediğini vurgulamışlardır. Bu çalışmadan elde edilen bulgular ölçülen nitelik sayısının artmasıyla üç modele ait sınıflama doğruluğu performanslarının azaldığını göstermiştir. Yani daha sade yapıdaki Q-matrislerin kullanımı sınıflama doğruluğu performansını arttırmaktadır (Chen & de la Torre, 2013; Wang vd., 2019). Özellikle diğer modellerle karşılaştırıldığında DINA modelin ölçülen nitelik sayısından daha fazla etkilendiği söylenebilir.

Çalışma kapsamında vurgulanması gereken önemli noktalardan biri SISM'in kısa testlerin kullanıldığı koşullarda dahi kabul edilebilir sınıflama doğruluğu değerleri almış olmasıdır (Kuo vd., 2018).

Çalışmada madde ayırt edicilik düzeyi arttıkça sınıflama doğruluğu performansının arttığı görülmektedir. Bu durum madde ayırt edicilik düzeyinin artmasıyla istatistiksel bilginin artmasından kaynaklanmaktadır. İstatistiksel bilgi arttıkça parametre tahminlerindeki değişkenlik azalacak ve parametrelere ilişkin daha kesin sonuçlar elde edilecektir. Bu nedenle madde ayırt edicilik düzeyinin artmasıyla sınıflama doğruluğu performansının artması beklenen bir durumdur (Schellman, 2021). Literatürde birçok araştırmayla desteklendiği gibi (Akbaş, 2016; Liu, Huggins-Manley & Bradshaw, 2017) bulgular, madde ayırt edicilik düzeyinin sınıflama doğruluğu performansı üzerinde etkisi olduğunu ortaya koymaktadır.

Çalışmada üç model için de örneklem büyüklüğüne kıyasla test uzunluğu artışıyla sınıflama doğruluğu performansının daha çok arttığı yani sınıflama doğruluğu performansının test uzunluğundaki artıştan göreceli olarak daha fazla etkilendiği sonucuna ulaşılmıştır. Bu durum Bradshaw (2011), Gao vd. (2017), Madison ve Bradshaw (2015) ile Yakar (2017) tarafından gerçekleştirilen çalışmalardan elde edilen bulgularla tutarlıdır. Test uzunluğuyla ilişkili olan bulgular bireylerin sınıflama doğruluğu değerlerini etkili bir şekilde arttırmak için hangi faktörlerin değişimleneceği noktasında uygulayıcılara yardımcı olmak adına bir yönlendirme sağlayabilir (Gao vd., 2017).

Bu çalışmanın sınırlılığı literatürde SISM için geliştirilmiş herhangi bir Q-matris geçerlik yöntemi bulunmadığından, kullanılan Q-matrislerin geçerliği üzerine herhangi bir çalışmanın yapılamamış olmasıdır. Pratikte gerçek Q-matrisi bilinmemektedir ve tasarım sürecinde Q-matrisin yanlış belirlenme olasılığı vardır (Gao vd. 2017). Q-matrisindeki hatalı tanımlamalar madde parametrelerine ait tahminlerin doğruluğunu etkilemektedir. Ancak çalışmadan elde edilen bulgular literatürden elde edilen bulgularla desteklendiğinde örneklem büyüklüğü ve test uzunluğunun fazla olduğu koşullarda bireylerin ilgili beceri ya da kavram yanlışlarında sınıflama doğruluğu performansının arttığı görülmektedir. Özellikle test uzunluğu faktörü dikkate alındığında hatalı tanımlanan maddelerdeki olası olmayan yanıt örüntüleri bu maddelerle aynı nitelikleri ölçen olası hatalı tanımlama bulunmayan diğer maddeler tarafından telafi edilecek ve bu faktörde meydana gelen artışlarla sınıflama doğruluğu performansı artacaktır (Uyumaz, 2016). Böyle bir durum sınıflamalar üzerindeki olumsuz etkiyi ortadan kaldırmaktadır.

Tüm durumlar genel olarak değerlendirildiğinde çalışmanın sonuçları, bilişsel tanı değerlendirmelerinde kullanılan modellere ilişkin literatürde yer alan birkaç ortak temayı desteklemektedir. Madde ayırt edicilik düzeyi, test uzunluğu ve örneklem büyüklüğü artışıyla birlikte çoğu psikometrik modelde olduğu gibi SISM'in sınıflama doğruluğu performansının bu koşullarda arttığı ancak ölçülen nitelik sayısı arttıkça modellere ait sınıflama doğruluğu performanslarının azaldığı görülmüştür.

5.2. Öneriler

Bu bölümde, çalışmanın amacı kapsamında sahadaki uygulayıcılara ve araştırmacılara yönelik önerilere yer verilmiştir.

5.2.1. Uygulayıcılara Yönelik Öneriler

Kavram yanılgılarının giderilmesinde etkileşimli öğrenme ya da akıl yürütmeyi kullanan öğretim planlarının başarısının yanı sıra, unutulmaması gereken önemli noktalardan biri araştırmacıların olması gerekenden daha fazla çaba göstererek bu modellerin gelişimine katkılar sunması ve yeni yaklaşımlar sağlamasıdır. Ancak bu sayede becerilerin tanımlanmasına ek olarak kavram yanılgılarının tanısı, eğitim öğretim reformlarına yardımcı olabilir. Bu kapsamda alan yazında kaydedilecek olan ilerlemeler için uygulayıcılara yönelik öneriler aşağıda sunulmuştur.

- Araştırmacıların ilgilendikleri nitelikler konu alanının kapsamı ve araştırmacının ilgilendiği yanıtlama süreciyle birlikte bu sürecin nasıl parçalardan oluşacağına dair tanımlamalarla ilişkilidir. Bu nedenle araştırmacılardan niteliklerle ilişkili olan konu alanlarının kapsamını daraltarak nitelikleri özelleştirmeleri beklenmektedir. SISM kapsamında ise beceriler ve kavram yanılgıları eş zamanlı olarak ölçüldüğünden modelde yer alan nitelik sayısının artmasıyla madde parametre tahmini güçleşebilir ya da modele ilişkin sınıflamalarda hata yapma olasılığı artabilir. Bu nedenle araştırmacıların nitelik hiyerarşisindeki beceri ve kavram yanılgıları arasındaki bağımlılık ilişkilerini, tanısal durumlara ilişkin tanımlamaların yapıldığı aşamada belirlemesi ve niteliklerin geniş bir aralıkta tanımlanması önerilebilir. Bu durumun bireylere ait tanı profillerinin sayısını azaltarak model karmaşıklığının azalmasına yardımcı olacağı düşünülmektedir.

- Beceri modellerine ek olarak kavram yanılgılarının tanısına yönelik katkı getiren bu modellerin gelişimini engelleyebilecek olası nedenlerden biri, kavram yanılgılarını ve becerileri belirlemedeki zorluktur. Araştırmacılar bu durumun Q-matrislerin gelişimi için büyüyen bir endişe olduğunu vurgulamışlardır (Elbulok, 2021). Bu endişeleri azaltabilmek adına araştırmacıların içerik uzmanlarıyla çalışarak bilişsel tanı değerlendirmeleri tasarımları gerekmektedir. Ayrıca becerilerle birlikte kavram yanılgılarının daha doğru ve anlamlı bir şekilde ölçüldüğü modellerin kullanımının yaygınlaştırılması araştırmacılar tarafından sağlanmalı ve bu modellere ilişkin testlerin geliştirilmesi noktasında çeldiriciler oluşturulurken mutlaka kavram yanılgıları dikkate alınarak hazırlanmalıdır.
- Çalışmada kavram yanılgılarının belirlenmesinde BUG-DINO modelinin neredeyse SISM kadar iyi bir performans sergilediği görülmüştür. Bu nedenle sınıflama güvenilirliğini arttırmak ve bireylerin kavram yanılgılarına ait gizil sınıflara doğru bir şekilde yerleştirilebilmelerini sağlamak adına modelleri kullanacak araştırmacılara bu iki modelin birlikte kullanımının uygun olduğu önerilebilir.
- Çalışmada dört madde parametresi için elde edilen değerlerin doğruluğu standart hataların sıfıra yakınlığı ölçüsünde doğrudur. Bu nedenle gerçek veri uygulamalarında dört parametreye ait ortalama standart hata değerlerinin 0'dan ciddi ölçüde farklılaştığı durumlarda araştırmacılara, ölçtükleri niteliklerin tanımlanmasına ilişkin tekrardan uzman görüşleri almaları ve analizleri tekrarlamaları önerilebilir.

5.2.2. Araştırmacılara Yönelik Öneriler

SISM'e ait analizlerden elde edilen sonuçlar ümit verici olsa da modelin güvenilir kullanımları sağlanmadan önce doğasını daha iyi anlamak adına ek çalışmaların yürütülmesine ihtiyaç vardır (Kuo vd., 2018). Bu kapsamda araştırmacılara yönelik öneriler aşağıda sunulmuştur.

- SISM çoktan seçmeli maddelerin 1-0 şeklinde puanlandığı Q-matrislerin kullanıldığı durumlar için uygundur. Literatür incelendiğinde çoktan seçmeli maddelerin sağlayacağı bilgilerin sınırlı olduğu bunun yerine bireylerin güçlü ve zayıf yönleri hakkında daha fazla bilgi sağlama yararına sahip olduğu düşünülen açık uçlu maddelerin daha etkili bir ölçme aracı olduğunu savunan çalışmalar vardır (Kuo vd.,

2016). Modelin yapılandırılmış yanıtli maddelerin kullanımı için genişletilmesiyle yanıtliın çok kategorili bir şekilde tanımlandığı Q-matris yapılarıyla, benzer bir çalışma arařtırmacılar tarafından tekrarlanabilir. Ya da çoktan seçmeli ve açık uçlu maddelerin kullanıldığı durumlarda model ve geliştirilmiş versiyonunun performansı karşılařtırılmalı olarak incelenebilir.

- Çalışmada SISM'in örneklem büyüklüğü, test uzunluęu, madde ayırt edicilik düzeyi ve ölçülen nitelik sayısı gibi çeřitli kořullar altında deęişimlenerek model performansı deęerlendirilmiştir. Bu nedenle bu çalışmadan elde edilen sonuçlar faktörlerde deęişimlenen kořullarla sınırlı olduğundan, SISM modeli kullanılarak doğru tahmin ve sınıflamalar için gerekli olan teorik ve pratik kořullar hakkında genellemeler yapılamamıştır. Bu bağlamda modelin performansına ilişkin çalışmada kullanılan faktörlere ait farklı kořulları ele alarak ya da farklı faktörlere ait kořulları deęişimleyerek model geçerliğine ilişkin kanıtlar toplama noktasında çeřitli çalışmalar arařtırmacılar tarafından tasarlanabilir. Örneęin arařtırmacılar bu çalışmaya kıyasla daha fazla sayıda beceri ve kavram yanılıęını ölçmeyi hedefliyor olabilir. Bu nedenle çalışma kapsamında ölçülen nitelik sayıları artırılarak SISM'in performansı incelenebilir. Bu aşamada unutulmaması gereken nokta nitelik sayısındaki artışla birlikte modelin çalışma sürelerinin artması ve arařtırmacılar tarafından Q-matris tanımlamalarının zorlařarak, bu tanımlamaların hata yapmaya açık bir hale gelmiş olmasıdır.
- Çalışmada SISM'in sınıflama doğruluęu performansı BUG-DINO ve DINA modellerinin sınıflama doğruluęu performanslarıyla karşılaştırılmıştır. Modelin performansının farklı beceri ve kavram yanılıęı modelleriyle karşılaştırabileceęi çalışmalar arařtırmacılar tarafından tasarlanabilir.
- Beceriler ve kavram yanılıęlarına ilişkin biliřsel tanı literatürü incelendięinde beceri sayısının kaç olması gerektięine ilişkin çok sayıda çalışma olduğu görülse de her bir kavram yanılıęı için istikrarlı birey sınıflamalarının elde edilebilmesi için kavram yanılıęı sayısının kaç olması gerektięiyle ilgili yeterli sayıda çalışmanın olmadığı görölmektedir. Arařtırmacılar literatürde yer alan bu sınırlıęa yönelik çalışmalar tasarlayabilirler.

- Çalışma bir simülasyon çalışmasıdır. SISM'in başarısı gerçek veri uygulamalarında araştırmacılar tarafından değerlendirilebilir. Bu sayede deneysel verilerin kullanımıyla elde edilen bulguların genelleştirilebilirliğine olan güven artacaktır. Bir modele ait psikometrik çerçevede oluşturulan değerlendirmeler, bilişsel tanı değerlendirmeleri kapsamında modellendiğinde bazı sınırlamalar ortaya çıkabilir. Bu sınırlamalar sebebiyle bireylerin kavram yanılgılarının var olup olmadığına ilişkin bilgiler sağlamak amacıyla uyarlama çalışmalarından ziyade, beceri ve kavram yanılgılarının ölçülmesi amacıyla bilişsel tanı modelleri kapsamında tasarlanmış testler kullanılarak pratik ortamlarda SISM'in kullanılabilirliğini üst düzeye çıkarmak için araştırmacılar tarafından çalışmalar tasarlanabilir. Böylelikle gerçek uygulamalardan elde edilen verilerin kullanılmasıyla sadece becerileri ölçmekten ziyade kavram yanılgılarının da ölçülmesi için geliştirilmiş bilişsel tanı modellerinin kullanımı araştırmacılar sayesinde yaygınlaştırılabilir. Ayrıca bilişsel tanı değerlendirmeleri kapsamında herhangi bir ölçme aracı geliştirmenin gerekli bir parçası etkili bir puan raporlamadır. Bu nedenle araştırmacılar test sonuçlarını paydaşlara raporlamak noktasında ölçme araçları için bir yöntem geliştirebilirler.
- Uygulamalarda çoğu durumda alan uzmanları tarafından Q-matrisi hatalı bir şekilde tanımlanmakta ve bilişsel tanı modellerinden elde edilen çıkarımların geçerliği araştırmacılar arasında tartışmaları gündeme getirmektedir. Q-matriste yer alan nitelik tanımlamalarının geçerliği sağlanmadan gerçekleştirilen herhangi bir model analizi güvenilir olmayabilir. Literatür incelendiğinde Q-matris tasarımlarının alan uzmanları tarafından öznel olarak oluşturulmasından sonra genelde bu tasarımların doğru olduğu varsayılmaktadır. Bu varsayım Q-matris geçerliği için iyi tanımlanmış yöntemlerin bulunmaması nedeniyle yapılmaktadır (Terzi, 2017). Bu durumun getirdiği öznellik hem modelde yer alan parametreler üzerinde hem de bireylerin doğru sınıflama oranı değerlerinde belirsizliklere sebep olabilmektedir (De Carlo, 2012). Alan uzmanlarının kapsamlı bir literatür taraması yoluyla gerekli nitelikleri belirlemesi, tanımlanmış niteliklere dayalı test maddeleri yazmaları ve nitelikleri öznellikten uzak bir şekilde doğrulayabilmeleri adına Q-matris geçerliğini sağlama yöntemleri geliştirilmeye devam etmektedir (Terzi, 2017). Becerileri ve kavram yanılgılarını eş zamanlı olarak ölçen modellere (Örneğin Elbulok, 2021; Kuo vd., 2018) ilişkin literatürde herhangi bir geçerlik yöntemi bulunmamaktadır. Başarılı bir bilişsel tanı modeli büyük ölçüde Q-matris geçerliği üzerine kuruludur. Bu nedenle

alandaki bu eksiliğe katkılar sunmak amacıyla beceri ve kavram yanılgılarını eş zamanlı olarak ölçen modellerden biri olan SISM'de kullanılan matrislerin geçerliğine ilişkin yöntem çalışmaları araştırmacılar tarafından tasarlanabilir.

- Çalışma kapsamında Q-matrisin hatalı belirlendiği durumlarla çalışılmamıştır. Uygulamalar sonucu alınacak kararların geçerliği yönünden Q-matrise ilişkin olası hatalı belirleme oranlarının madde parametrelerine ve sınıflama doğruluğu değerlerine etkisinin incelenmesi önemlidir (Uyumaz, 2016). Bu nedenle Q-matrise ilişkin hatasız tanımlanan matrislerle olası hatalı belirleme oranlarının dikkate alındığı matrisler araştırmacılar tarafından karşılaştırılarak modelin performansı incelenebilir.
- Çalışma SISM'e ilişkin Q-matrislerinin oluşturulması sürecinde becerilerin ve kavram yanılgılarının kendi içinde (küme-içi) bağımlılıklarının olmadığı yani sadece beceriler ve kavram yanılgıları arasında dış bağımlılık ilişkisinin kurulduğu durumlarla sınırlıdır. Araştırmacılar küme-içi bağımlılık ya da hem küme-içi bağımlılık hem de dış bağımlılık ilişkileri kurarak SISM'in performansını inceleyebilirler.

KAYNAKLAR

- Ackerman, T. (1996). Graphical representation of multidimensional item response theory analyses. *Applied Psychological Measurement*, 2, 311-329. <https://doi.org/10.1177/014662169602000402>
- Akbay, L. (2016). *Identification, estimation, and Q-matrix validation of hierarchically structured attributes in cognitive diagnosis*. (Doctoral dissertation). Retrieved from <https://www.proquest.com/docview/1877640767?pq-origsite=gscholar&fromopenview=true>
- Akbay, L., & de la Torre, J. (2020). Estimation approaches in cognitive diagnosis modeling when attributes are hierarchically structured. *Psicothema*, 32(1), 122-129. <https://doi.org/10.7334/psicothema2019.182>
- Akbay L., Terzi R., Kaplan M. & Karaaslan K. G. (2018). Expert-based attribute identification and validation: An application of cognitively diagnostic assessment. *Journal on Mathematics Education*, 9, 103-120. <https://doi.org/10.22342/jme.9.1.4341.103-120>
- Alwan, A. A. (2011). Misconception of heat and temperature among physics students. *Procedia-Social and Behavioral Sciences*, 12, 600-614. <https://doi.org/10.1016/j.sbspro.2011.02.074>
- Anıl, Ö., & Küçüközer, H. (2010). Ortaöğretim 9. sınıf öğrencilerinin düzlem ayna konusunda sahip oldukları ön bilgi ve kavram yanlışlarının belirlenmesi. *Türk Fen Eğitimi Dergisi*, 7(3), 104-122. <https://dspace.balikesir.edu.tr/xmlui/handle/20.500.12462/4843> sayfasından erişilmiştir.

- Ardıç, E. Ö. (2020). *Bilişsel tanı ve çok boyutlu madde tepki modellerinin sınıflama doğruluğu ve parametrelerinin karşılaştırılması*. (Doktora tezi). [https://tez.yok.gov.tr/UlusalTezMerkezi/](https://tez.yok.gov.tr/UlusalTezMerkezi/sayfasından) sayfasından erişilmiştir.
- Ayan, C. (2018). *Bilişsel tanı modelinde geleneksel ve bilgisayarlı sınıflamalı test uygulamalarının psikometrik özelliklerinin karşılaştırılması*. (Doktora tezi). [https://tez.yok.gov.tr/UlusalTezMerkezi/](https://tez.yok.gov.tr/UlusalTezMerkezi/sayfasından) sayfasından erişilmiştir.
- Baker, F. (1993) Sensitivity of the linear logistic test model to misspecification of the weight matrix. *Applied Psychological Measurement*, 17(3), 201-210. <https://doi.org/10.1177/014662169301700301>
- Baki, A. (2006). *Kuramdan uygulamaya matematik öğretimi*. İstanbul: Bilge Matbaacılık.
- Başokçu, T.O. (2011). *Bağıl ve mutlak değerlendirme ile DINA modele göre yapılan sınıflamaların geçerliğinin karşılaştırılması*. (Doktora tezi). [https://tez.yok.gov.tr/UlusalTezMerkezi/](https://tez.yok.gov.tr/UlusalTezMerkezi/sayfasından) sayfasından erişilmiştir.
- Bingölbali, E., & Özmantar, M.F. (2009). *Matematiksel kavram yanılgıları: sebepleri ve çözüm arayışları*. Ankara: Pegem Akademik Yayıncılık.
- Bradshaw, L., & Templin, J. (2014). Combining item response theory and diagnostic classification models: A psychometric model for scaling ability and diagnosing misconceptions. *Psychometrika*, 79(3), 403–425. <https://doi.org/10.1007/s11336-013-9350-4>
- Bradshaw, L. P. (2011). *Combining scaling and classification: A psychometric model for scaling ability and diagnosing misconceptions*. (Doctoral dissertation). Retrieved from <https://www.proquest.com/docview/2572548457?pq-origsite=gscholar&fromopenview=true>
- Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, 2, 155-192. [https://doi.org/10.1016/S0364-0213\(78\)80004-4](https://doi.org/10.1016/S0364-0213(78)80004-4)
- Chen, J., & de la Torre, J. (2018). Introducing the general polytomous diagnosis modeling framework. *Frontiers in Psychology*, 9, 1–9. <https://doi.org/10.3389/fpsyg.2018.01474>

- Chen, J., & de la Torre, J. (2013). A general cognitive diagnosis model for expert-defined polytomous attributes. *Applied Psychological Measurement, 37*(6), 419-437. <https://doi.org/10.1177/0146621613479818>
- Chen, J., de la Torre, J., & Zhang, Z. (2013). Relative and absolute fit evaluation in cognitive diagnosis modeling. *Journal of Educational Measurement, 50*(2), 123–140. <https://doi.org/10.1111/j.1745-3984.2012.00185.x>
- Chen, P., Xin, T., Wang, C., & Chang, H. H. (2012). Online calibration methods for the DINA model with independent attributes in CD-CAT. *Psychometrika, 77*(2), 201–222. <https://doi.org/10.1007/S11336-012-9255-7>
- Chen, W. J., Lin, H. M., & Nien, S. F. (2014). The learning effectiveness of the concept map approach of e-learning applied to a math class of special educational students in a vocational school. *International Journal of Information and Education Technology, 4*(5), 388-393. <https://doi.org/10.7763/IJiet.2014.V4.436>
- Cheng, Y. (2010). Improving cognitive diagnostic computerized adaptive testing by balancing attribute coverage: The modified maximum global discrimination index Method. *Educational and Psychological Measurement, 70*(6), 902-913. <https://doi.org/10.1177/0013164410366693>
- Cui, Y., Gierl, M. J., & Chang, H. H. (2012). Estimating classification consistency and accuracy for cognitive diagnostic assessment. *Journal of Educational Measurement, 49*, 19-38. <https://doi.org/10.1111/j.1745-3984.2011.00158.x>
- Çakır, S. Ö., & Yürük, N. (1999). *Oksijenli ve oksijensiz solunum konusunda kavram yanlışları teşhis testinin geliştirilmesi ve uygulanması*. III. Fen Bilimleri Eğitimi Sempozyumu'nda sunulmuş bildiri, M.E.B. ÖYGM, Ankara.
- De Carlo, L. T. (2012). Recognizing uncertainty in the Q-matrix via a bayesian extension of the DINA model. *Applied Psychological Measurement, 36*(6), 447–468. <https://doi.org/10.1177/0146621612449069>
- de la Torre, J. (2009b). A cognitive diagnosis model for cognitively based multiple-choice options. *Applied Psychological Measurement, 33*(3), 163-183. <https://doi.org/10.1177/0146621608320523>

- de la Torre, J. (2008). An empirically-based method of Q-matrix validation for the DINA model: development and applications. *Journal of Educational Measurement*, 45, 343–362. <https://doi.org/10.1111/j.1745-3984.2008.00069.x>
- de la Torre, J., & Douglas, J. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika*, 69(3), 333-353. <https://doi.org/10.1007/BF02295640>
- de la Torre, J., Hong, Y., & Deng, W. (2010). Factors affecting the item parameter estimation and classification accuracy of the DINA model. *Applied Psychological Measurement*, 47, 227-249. <https://doi.org/10.1111/j.1745-3984.2010.00110.x>
- de la Torre, J., & Huo, Y. (2014). Estimating a cognitive diagnostic model for multiple strategies via the EM algorithm. *Applied Psychological Measurement*, 38, 464-485. <https://doi.org/10.1177/0146621614533986>
- de la Torre, J., & Karelitz, T. M. (2009). Impact of diagnosticity on the adequacy of models for cognitive diagnosis under a linear attribute structure: A simulation study. *Journal of Educational Measurement*, 46, 450-469. <https://doi.org/10.1111/j.1745-3984.2009.00092.x>
- de la Torre, J., & Lee, Y. S. (2010). A note on the invariance of the DINA model parameters. *Journal of Educational Measurement*, 47(1), 115-127. <https://doi.org/10.1111/j.1745-3984.2009.00102.x>
- de la Torre, J., & Minchen, N. (2014). Cognitively diagnostic assessments and the cognitive diagnosis model framework. *Psicología Educativa*, 20(2), 89-97. <https://doi.org/10.1016/j.pse.2014.11.001>
- DeMars, C. E. (2003). Sample size and the recovery of nominal response model item parameters. *Applied Psychological Measurement*, 27(4), 275–288. <https://doi.org/10.1177/0146621603027004003>
- DiBello, L. V, Henson, R. A., & Stout, W. F. (2015) A family of generalized diagnostic classification models for multiple choice option-based scoring. *Applied Psychological Measurement*, 39(1), 62–79. <https://doi.org/10.1177/0146621614561315>
- Dogan, E., & Tatsuoka, K. (2008). An international comparison using a diagnostic testing model: Turkish students' profile of mathematical skills on TIMSS-R. *Educational Studies in Mathematics*, 68(3), 263-272. <https://doi.org/10.1007/s10649-007-9099-8>

- Driver, R., & G. Erickson. (1983). Theories-in-action: some theoretical and empirical issues in the study of students' conceptual frameworks in science. *Studies in Science Education*, 10(1), 37–60. <https://doi.org/10.1080/03057268308559904>
- Elbulok, M. (2021). *A cognitively diagnostic modeling approach to diagnosing misconceptions and subskills*. (Doctoral dissertation). Retrieved from <https://www.proquest.com/docview/2492657094>
- Fischer, G. H. (1973). The linear logistic model as an instrument in educational research. *Acta Psychologica*, 37, 359–374. [https://doi.org/10.1016/0001-6918\(73\)90003-6](https://doi.org/10.1016/0001-6918(73)90003-6)
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education*. New York, NY: McGraw-Hill.
- Gao, M., Miller, M., & Liu, R. (2017). The impact of Q-matrix misspecification and model misuse on classification accuracy in the Generalized DINA Model. *Journal of Measurement and Evaluation in Education and Psychology*, 8(4), 391-403 <https://doi.org/10.21031/epod.332712>
- Gierl, M. J., Alves, C., & Majeau, R. T. (2010). Using the attribute hierarchy method to make diagnostic inferences about examinees' knowledge and skills in mathematics: An operational implementation of cognitive diagnostic assessment. *International Journal of Testing*, 1, 318-341. <https://doi.org/10.1080/15305058.2010.509554>
- Haberman, S. J., Sinharay, S., & Puhan, G. (2009). Reporting subscores for institutions. *British Journal of Mathematical and Statistical Psychology*, 62, 79-95. <https://doi.org/10.1348/000711007X248875>
- Haertel, E. H. (1989). Using restricted latent class models to map the skill structure of achievement items. *Journal of Educational Measurement*, 26, 301–321. <http://dx.doi.org/10.1111/j.1745-3984.1989.tb00336.x>
- Harwell, M. (2018). A strategy for using bias and RMSE as outcomes in monte carlo studies in statistics. *Journal Modern Applied Statistical Methods*, 17(2), Article 5. <https://doi.org/10.22237/jmasm/1551907966>
- Henson, R. A., Templin, J. L., & Willse, J. T. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika*, 74, 191–210. <https://doi.org/10.1007/s11336-008-9089-5>

- Huang, H.-Y. (2017). Multilevel cognitive diagnosis models for assessing changes in latent attributes. *Journal of Educational Measurement*, 54(4), 440–480
<https://doi.org/10.1111/jedm.12156>
- Junker, B.W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25(3), 258-272. <https://doi.org/10.1177/01466210122032064>
- Kalkan, Ö. K. (2016). *Bilişsel tanı modellerinin değişen koşullar altında karşılaştırılması: DINA, RDINA, HODINA ve HORDINA Modelleri*. (Doktora tezi). [https://tez.yok.gov.tr/UlusalTezMerkezi/ sayfasından](https://tez.yok.gov.tr/UlusalTezMerkezi/sayfasından) erişilmiştir.
- Khazanov, L. (2009). *A diagnostic assessment for misconceptions in probability*. Paper presented at the annual Georgia Perimeter College Mathematics Conference in Clarkston, GA.
- Koçak, B. (2020). *Matematik öğretmenliği lisans öğrencilerinin üstel belirsizlikler ve diferansiyel konularındaki kavram yanlışlarının incelenmesi*. (Yüksek lisans tezi). [https://tez.yok.gov.tr/UlusalTezMerkezi/ sayfasından](https://tez.yok.gov.tr/UlusalTezMerkezi/sayfasından) erişilmiştir.
- Koyuncu, M. S. (2020). *Bilişsel tanı modellerinde yapısal eşitlik modeli ile Q-matris doğruluğunun belirlenmesi*. (Doktora tezi). [https://tez.yok.gov.tr/UlusalTezMerkezi/ sayfasından](https://tez.yok.gov.tr/UlusalTezMerkezi/sayfasından) erişilmiştir.
- Kunina-Habenicht, O., Rupp, A. A., & Wilhelm, O. (2012). The impact of model misspecification on parameter estimation and item-fit assessment in log-linear diagnostic classification models. *Journal of Educational Measurement*, 49(1), 59-81. <https://doi.org/10.1111/j.1745-3984.2011.00160.x>
- Kuo, B. C., Chen, C. H., & de la Torre, J. (2018). A cognitive diagnosis model for identifying coexisting skills and misconceptions. *Applied Psychological Measurement*, 42(3) 179–191. <https://doi.org/10.1177/0146621617722791>
- Kuo, B. C., Chen, C. H., Yang, C. W., & Mok, M. M. C. (2016). Cognitive diagnostic models for tests with multiple-choice and constructed-response items. *Educational Psychology*, 36(6), 1115–1133. <https://doi.org/10.1080/01443410.2016.1166176>
- Lee, J., & Corter, J. E. (2011). Diagnosis of subtraction bugs using Bayesian networks. *Applied Psychological Measurement*, 35(1), 27-47. <https://doi.org/10.1177/0146621610377079>

- Lee, Y. S., Park, Y. S., & Taylan, D. (2011). A cognitive diagnostic modeling of attribute mastery in Massachusetts, Minnesota, and the US national sample using the TIMSS 2007. *International Journal of Testing*, *11*(2), 144-177. <https://doi.org/10.1080/15305058.2010.534571>
- Lee, Y., & Sawaki, Y. (2009). Application of three cognitive diagnosis models to ESL reading and listening assessments. *Language Assessment Quarterly*, *6*, 239–263. <https://doi.org/10.1080/15434300903079562>
- Lei, P. W., & Li, H. (2016). Performance of Fit Indices in Choosing Correct Cognitive Diagnostic Models and Q-Matrices. *Applied Psychological Measurement*, *40*(6) 405–417. <https://doi.org/10.1177/0146621616647954>
- Leighton, J. P., & Gierl, M. J. (2007). Why cognitive diagnostic assessment. In J. P. Leighton, & M. J. Gierl (Eds.), *Cognitive Diagnostic Assessment for Education: Theory and Applications* (pp. 3-18). Cambridge University Press.
- Leighton, J. P., Gierl, M. J., & Hunka, S. M. (2004). The attribute hierarchy method for cognitive assessment: A variation on Tatsuoka's Rule-Space approach. *Journal of Educational Measurement*, *41*(3), 205-237. <https://doi.org/10.1111/j.1745-3984.2004.tb01163.x>
- Li, L., Zhou, X., Huang, J., Tuc, D., Gaod, X., Yanga, Z., & Lia, M. (2020). Assessing kindergarteners' mathematics problem solving: The development of a cognitive diagnostic test. *Studies in Educational Evaluation*, *66*, 1-10. <https://doi.org/10.1016/j.stueduc.2020.100879>
- Liang, L., Lu, J., Zhang, J., & Shi, N. (2022). Modeling not-reached items in cognitive diagnostic assessments. *Frontiers Psychology*, *13*. <https://doi.org/10.3389/fpsyg.2022.889673>
- Liu, C., & Cheng, Y. (2018). An application of the support vector machine for attribute-by-attribute classification in cognitive diagnosis. *Applied Psychological Measurement*, *42*(1), 58-72. <https://doi.org/10.1177/0146621617712246>
- Liu, R., Huggins-Manley, A. C., & Bradshaw, L. (2017). The impact of Q-matrix designs on diagnostic classification accuracy in the presence of attribute hierarchies. *Educational and Psychological Measurement*, *77*(2), 220-240. <https://doi.org/10.1186/1471-2148-10-242>

- Liu, R., Huggins-Manley, A. C., & Bulut, O. (2018). Retrofitting diagnostic classification models to responses from IRT-based assessment forms. *Educational and Psychological Measurement*, 78(3), 357–383. <https://doi.org/10.1177/0013164416685599>
- Liu, R., Quian, H., Luo, X., & Woo, A. (2018). Relative Diagnostic Profile: A subscore reporting framework. *Educational and Psychological Measurement*, 78(6), 1072–1088. <https://doi.org/10.1177/0013164417740170>
- Ma, W., & de la Torre. (2020). Package “GDINA”. Retrieved from <https://cran.r-project.org/web/packages/GDINA/GDINA.pdf>
- Ma, W., & de la Torre, J. (2019a). Digital Module 05: Diagnostic Measurement-The GDINA Framework. *Educational Measurement: Issues and Practice*, 38(2), 114-115. <https://doi.org/10.1111/emip.12262>
- MacDonald, G. T. (2013). *The performance of the linear logistic test model when the Q-matrix is misspecified: a simulation study*. (Doctoral dissertation). Retrieved from <https://www.proquest.com/docview/1530419261?pqorigsite=gscholar&fromopenview=true>
- Macready, G. B., & Dayton, C. M. (1977). The use of probabilistic models in the assessment of mastery. *Journal of Educational Statistics*, 2(2), 99-120. <https://www.jstor.org/stable/1164802> sayfasından erişilmiştir.
- Madison, M. J., & Bradshaw, L. P. (2015). The effects of Q-matrix design on classification accuracy in the log-linear cognitive diagnosis model. *Educational and Psychological Measurement*, 75(3), 491-511. <https://doi.org/10.1177/0013164414539162>
- McIntosh, M. E., & Draper, R. J. (2001). Using learning logs in mathematics: Writing to learn. *The Mathematics Teacher*, 94(7), 554-557. <https://doi.org/10.5951/MT.94.7.0554>
- Mintzes, J. J. (1984). Naive theories in biology: children's concepts of the human body. *School Science and Mathematics*, 84(7), 548-555. <https://doi.org/10.1111/j.1949-8594.1984.tb10179.x>
- Niss, M. (1999). Aspects of the nature and state of research in mathematics education. *Educational Studies in Mathematics*. 40(1), 1–24. <https://doi.org/10.1023/A:1003715913784>

No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107–11, 115 Stat/ 1449–1452 (2002).

Norris, M. A. (2021). *Rethinking the force concept inventory: Developing a cognitive diagnostic assessment to measure misconceptions in Newton's laws*. (Doctoral dissertation). Retrieved from https://vtechworks.lib.vt.edu/bitstream/handle/10919/105262/Norris_MA_D_2021.pdf?sequence=1&isAllowed=y

Odabaş, M. (2016). *Değişen madde fonksiyonunu belirlemede DINA modelde işaretli alan indeksi, standardizasyon ve lojistik regresyon tekniklerinin karşılaştırılması*. (Doktora tezi). <https://tez.yok.gov.tr/UlusalTezMerkezi/sayfasından> erişilmiştir.

Ozaki, K. (2015). DINA models for multiple-choice items with few parameters: considering incorrect answers. *Applied Psychological Measurement*, 39(6), 431–447. <https://doi.org/10.1177/0146621615574693>

Ozaki, K., Sugawara, S. & Arai, N. (2020). Cognitive diagnosis models for estimation of misconceptions analyzing multiple-choice data. *Behaviormetrika*, 47, 19–41. <https://doi.org/10.1007/s41237-019-00100-9>

Ömür-Sünbül, S. (2013). *Bilişsel tanı modellerinde parametre kestirimini ve sınıflama tutarlılığını etkileyen faktörlerin incelenmesi*. (Doktora tezi). <https://tez.yok.gov.tr/UlusalTezMerkezi/sayfasından> erişilmiştir.

Ömür-Sünbül, S., & Kan, A. (2015). Bilişsel tanı modellerinde parametre kestirimini ve sınıflama tutarlılığını etkileyen faktörlerin incelenmesi. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 31(4), 778-795. <https://doi.org/10.16986/HUJE.2015014663>

Parlak, Ö. İ. (2019). *9. sınıf öğrencilerinin üçgenlerde temel kavramlara ilişkin kavram yanılgılarının incelenmesi*. (Yüksek lisans tezi). <https://tez.yok.gov.tr/UlusalTezMerkezi/sayfasından> erişilmiştir.

Pines, A. L., & West, L. H. T. (1986), Conceptual understanding and science learning: an interpretation of research within a sources-of-knowledge framework. *Science Education*, 70(5), 583-604. <https://doi.org/10.1002/sce.3730700510>

- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of a conceptual change. *Science Education*, 66(2), 211-227. Retrieved from <https://staff.science.uva.nl/e.joling/vakdidactiek/documenten/posner.pdf>
- Qin, C., Zhang, L., Qiu, D., Huang, L., Geng, T., Jiang, H., Ren, Q., & Zhou, J. (2015). Model identification and Q-matrix incremental inference in cognitive diagnosis. *Knowledge-Based Systems*, 86, 66-76. <https://doi.org/10.1016/j.knosys.2015.05.024>
- Ravand, H., & Baghaei, P. (2020) Diagnostic Classification Models: Recent Developments, Practical Issues, and Prospects. *International Journal of Testing*, 20(1), 24-56. <https://doi.org/10.1080/15305058.2019.1588278>
- Robitzsch, A., Kiefer, T., George, A. C., & Uenlue, A. (2020). *Package 'CDM'*. Retrieved from <https://cran.r-project.org/web/packages/CDM/CDM.pdf>
- Romero, S. J., Ordoñez, X. G., Ponsoda, V., & Revuelta, J. (2014). Detection of Q-matrix misspecification using two criteria for validation of cognitive structures under the Least Squares Distance Model. *Psicológica*, 35(1), 149-169. Retrieved from <https://www.redalyc.org/pdf/169/16930557010.pdf>
- Rupp, A. A. (2007). The answer is in the question: a guide for describing and investigating the conceptual foundations and statistical properties of cognitive psychometric models. *International Journal of Testing*, 7(2), 95-125. <https://doi.org/10.1080/15305050701193454>
- Rupp, A., & Templin, J. (2008). Unique characteristics of diagnostic models: A review of the current state-of-the-art. *Measurement*, 6(4), 219–262. <https://doi.org/10.1080/15366360802490866>
- Rupp, A., Templin, J., & Henson, R. (2010). *Diagnostic measurement: Theory, methods, and applications*. New York: Guilford Press.
- Sadler, P. M. (1998). Psychometric models of student conceptions in science: Reconciling qualitative studies and distractor-driven assessment instruments. *Journal of Research in Science Teaching*, 35, 265. [https://doi.org/10.1002/\(SICI\)1098-2736\(199803\)35:3<265::AID-TEA3>3.0.CO;2-P](https://doi.org/10.1002/(SICI)1098-2736(199803)35:3<265::AID-TEA3>3.0.CO;2-P)

- Schellman, M. A. (2021). *Diagnostic concept inventories for misconception classification accuracy and reliability*. (Master's thesis). Retrieved from <https://www.proquest.com/docview/2572548457?pqorigsite=gscholar&fromopenview=true>
- Schulte, P. L. (2001). *Pre service primary teachers' alternative conceptions in science and attitudes towards teaching science*. (Doctoral dissertation). Retrieved from <https://www.proquest.com/docview/304703069>
- Sen, S., & Bradshaw, L. (2017). Comparison of relative fit indices for diagnostic model selection. *Applied Psychological Measurement*, 41(6), 422-438. <https://doi.org/10.1177/0146621617695521>
- Sen, S., & Cohen, A. S. (2021). Sample size requirements for applying diagnostic classification models. *Frontiers in Psychology*, 11, 621-251. <https://doi.org/10.3389/fpsyg.2020.621251>
- Sever, R., Budak, M. F., & Yalçınkaya, E. (2009). Coğrafya eğitiminde kavram haritalarının önemi. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 13(2), 19–32. <https://dergipark.org.tr/en/pub/ataunisobil/issue/2824/38125> sayfasından erişilmiştir.
- Shear, B. R., & Roussos, L. A. (2017). Validating a distractor-driven geometry test using a generalized diagnostic classification model. In B. D. Zumbo & A. M. Hubley (Eds.), *Understanding and investigating response processes in validation research* (pp. 277-304). Cham, Switzerland: Springer.
- Smith, E. L., Blakeslee, T. D., & Anderson, C. W. (1993). Teaching strategies associated with conceptual change learning in science. *Journal of Research in Science Teaching*, 3, 111-126. <https://doi.org/10.1002/tea.3660300202>
- Sun, J., Xin, T., Zhang, S., & de la Torre, J. (2013). A polytomous extension of the generalized distance discriminating method. *Applied Psychological Measurement*, 37(7), 503-521. <https://doi.org/10.1177/0146621613487254>
- Şenyiğit, Ç. (2020). *Sorgulama temelli öğrenmenin sınıf öğretmeni adaylarının bilimsel süreç becerilerine ve kavramsal anlamaların etkisi*. (Doktora tezi), <https://tez.yok.gov.tr/UlusalTezMerkezi/> sayfasından erişilmiştir.

- Tall, D. O., & Razali, M. R. (1993) Diagnosing students' difficulties in learning mathematics. *International Journal of Mathematical Education in Science and Technology*, 24(2), 209–222. <https://doi.org/10.1080/0020739930240206>
- Tatsuoka, K. K., Corter, J. E., & Tatsuoka, C. (2004). Patterns of diagnosed mathematical content and process skills in TIMSS-R across a sample of 20 countries. *American Educational Research Journal*, 41 (4), 901–926. <https://doi.org/10.3102/00028312041004901>
- Templin, J., & Henson, R. (2006). Measurement of psychological disorders using cognitive diagnosis models. *Psychological Methods*, 11(3), 287-305. <https://doi.org/10.1037/1082-989X.11.3.287>
- Templin, J., & Hoffman, L. (2013). Obtaining diagnostic classification model estimates using Mplus. *Educational Measurement: Issues and Practice*, 32(2), 37–50. <https://doi.org/10.1111/emip.12010>
- Terzi, R. (2017). *New Q-matrix validation procedures*. (Doctoral dissertation). Retrieved from <https://rucore.libraries.rutgers.edu/rutgers-lib/55750/>
- Tu, D., Wang, S., Cai, Y., Douglas, J., & Chang, H. H. (2019). Cognitive diagnostic models with attribute hierarchies: model estimation with a restricted Q-matrix design. *Applied Psychological Measurement*, 43(4), 255-271. <https://doi.org/10.1177/0146621618765721>
- Uyumaz, G. (2016). *DINA modelde Q matrisinin hatalı belirlenmesinin farklı örneklem büyüklüklerinde parametre kestirimine ve bireylerin sınıflandırılmasına etkisi*. (Doktora tezi). https://tez.yok.gov.tr/UlusalTezMerkezi/_sayfasından erişilmiştir.
- Ülgen, G. (2004). *Kavram geliştirme, kuramlar ve uygulamalar*. Ankara: Nobel.
- Wang, W., Song, L., Chen, P., & Ding, S. (2019). An item-level expected classification accuracy and its applications in cognitive diagnostic assessment. *Journal of Educational Measurement*, 56(1), 51–75. <https://doi.org/10.1111/jedem.12200>
- Wang, W., Song, L., Chen, P., Meng, Y., & Ding, S. (2015). Attribute-level and pattern-level classification consistency and accuracy indices for cognitive diagnostic assessment. *Journal of Educational Measurement*, 52, 457-476. <https://doi.org/10.1111/jedem.12096>

- Wu, X., Wu, R., Chang, H.-H., Kong, Q., & Zhang, Y. (2020). International comparative study on PISA mathematics achievement test based on cognitive diagnostic models. *Frontiers in Psychology, 11*, 2230. <https://doi.org/10.3389/fpsyg.2020.02230>
- Xi, C., Tu, D., & Cai, Y. (2022). Dual-objective item selection methods in computerized adaptive test using the higher-order cognitive diagnostic models. *Applied Psychological Measurement, 46*(5), 422-438. <https://doi.org/10.1177/01466216221089342>
- Yakar, L. (2017). *Bilişsel tanı ve çok boyutlu madde tepki kuramı modellerinin karşılıklı uyumlarının incelenmesi* (Doktora tezi). <https://tez.yok.gov.tr/UlusalTezMerkezi/sayfasından erişilmiştir>.
- Yamaguchi K (2020) Variational bayesian inference for the multiple-choice DINA model. *Behaviormetrika, 47*(1), 159-187. <https://doi.org/10.1007/s41237-020-00104-w>
- Yavuz, E. (2021). *Parametrik olmayan bilişsel tanılama, yapay sinir ağı ve DINO modellerinin sınıflandırma performanslarının karşılaştırılması*. (Doktora tezi). <https://tez.yok.gov.tr/UlusalTezMerkezi/sayfasından erişilmiştir>.
- Yenilmez, K., & Yaşa, E. (2008). İlköğretim öğrencilerinin geometrideki kavram yanlışları. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi, 21*(2), 461-483. <https://dergipark.org.tr/en/pub/uefad/issue/16688/173427> sayfasından erişilmiştir.
- Zhan, P., Jiao H., & Liao, D. (2018). Cognitive diagnosis modeling incorporating item response times. *British Journal of Mathematical and Statistical Psychology, 71*, 262–286. <https://doi.org/10.1111/bmsp.12114>
- Zhan, P., Man, K., Wind, S. A., & Malone, J. (2022). Cognitive diagnosis modeling incorporating response times and fixation counts: Providing comprehensive feedback and accurate diagnosis. *Journal of Educational and Behavioral Statistics, 20*(10), 1-41. <https://doi.org/10.3102/10769986221111085>

EKLER



EK 1. 6 ve 10 Nitelik için Olası Nitelik Sınıfları

3 Beceri ve 3 Kavram Yanılgısı için Olası Nitelik Sınıfları

Sınıf Grubu	Nitelikler						Olası Bilgi Durumları
	B_1	B_2	B_3	K_1	K_2	K_3	
1	1	1	1	0	*	*	$2^2 = 4$
2	1	1	0	0	*	*	$2^2 = 4$
3	1	0	1	0	*	*	$2^2 = 4$
4	0	1	1	0	*	*	$2^2 = 4$
5	1	0	0	0	*	*	$2^2 = 4$
6	0	1	0	0	*	*	$2^2 = 4$
7	0	0	1	*	*	*	$2^3 = 8$
8	0	0	0	*	*	*	$2^3 = 8$

Tabloda yer alan B_1 , B_2 ve B_3 sırasıyla 1., 2. ve 3. becerileri; K_1 , K_2 ve K_3 sırasıyla 1., 2. ve 3. kavram yanılgılarını temsil etmektedir.

5 Beceri ve 5 Kavram Yanılgısı için Olası Nitelik Sınıfları

Sınıf Grubu	Nitelikler										Olası Bilgi Durumları
	B_1	B_2	B_3	B_4	B_5	K_1	K_2	K_3	K_4	K_5	
1	1	1	1	1	1	0	0	0	*	*	$2^2 = 4$
2	1	1	1	1	0	0	0	0	*	*	$2^2 = 4$
3	1	1	1	0	1	0	0	0	*	*	$2^2 = 4$
4	1	1	0	1	1	0	0	0	*	*	$2^2 = 4$
5	1	0	1	1	1	0	*	0	*	*	$2^3 = 8$
6	0	0	1	1	1	*	0	0	*	*	$2^3 = 8$
7	1	1	1	0	0	0	0	0	*	*	$2^2 = 4$
8	1	1	0	1	0	0	0	0	*	*	$2^2 = 4$
9	1	1	0	0	1	0	0	*	*	*	$2^3 = 8$
10	0	1	1	1	0	*	0	0	*	*	$2^3 = 8$
11	0	1	1	0	1	*	0	0	*	*	$2^3 = 8$
12	1	0	1	0	1	0	*	0	*	*	$2^3 = 8$
13	1	0	1	1	0	0	*	0	*	*	$2^3 = 8$
14	0	1	0	1	1	*	0	0	*	*	$2^3 = 8$
15	0	0	1	1	1	*	*	0	*	*	$2^4 = 16$
16	1	0	0	1	1	0	*	*	*	*	$2^4 = 16$
17	1	1	0	0	0	0	0	0	*	*	$2^2 = 4$
18	1	0	1	0	0	0	*	0	*	*	$2^3 = 8$
19	0	1	1	0	0	*	0	0	*	*	$2^3 = 8$
20	1	0	0	1	0	0	*	*	*	*	$2^4 = 16$
21	1	0	0	0	1	0	*	*	*	*	$2^4 = 16$
22	0	1	0	1	0	*	0	0	*	*	$2^3 = 8$
23	0	1	0	0	1	*	0	0	*	*	$2^3 = 8$
24	0	0	1	1	0	*	*	0	*	*	$2^4 = 16$
25	0	0	1	0	1	*	*	0	*	*	$2^4 = 16$
26	0	0	0	1	1	*	*	*	*	*	$2^5 = 32$
27	1	0	0	0	0	0	*	*	*	*	$2^4 = 16$
28	0	1	0	0	0	*	0	0	*	*	$2^3 = 8$
29	0	0	1	0	0	*	*	0	*	*	$2^4 = 16$
30	0	0	0	1	0	*	*	*	*	*	$2^5 = 32$
31	0	0	0	0	1	*	*	*	*	*	$2^5 = 32$
32	0	0	0	0	0	*	*	*	*	*	$2^5 = 32$

Tabloda yer alan B_1 , B_2 , B_3 , B_4 ve B_5 sırasıyla 1., 2., 3., 4. ve 5. becerileri; K_1 , K_2 , K_3 , K_4 ve K_5 sırasıyla 1., 2., 3., 4. ve 5. kavram yanılgılarını temsil etmektedir.

EK 2. 6 ve 10 Nitelik için 20 Maddelik Q-matris

6 Nitelik için 20 Maddelik Q-matris

Maddeler	Nitelikler					
	Beceriler			Kavram Yanılgıları		
	B_1	B_2	B_3	K_1	K_2	K_3
1	1	0	0	0	0	0
2	0	1	0	0	0	0
3	0	0	1	0	0	0
4	0	0	0	1	0	0
5	0	0	0	0	1	0
6	0	0	0	0	0	1
7	1	0	0	1	0	0
8	0	1	0	1	0	0
9	1	1	0	1	0	0
10	0	1	1	1	0	0
11	1	1	1	1	0	0
12	1	1	1	1	0	0
13	1	0	1	1	0	0
14	1	0	1	0	1	1
15	1	0	0	0	0	1
16	0	0	1	0	0	1
17	0	1	0	0	1	1
18	0	0	1	0	1	1
19	0	1	0	0	1	1
20	0	0	1	0	1	0

Tabloda yer alan B_1, B_2, B_3 sırasıyla 1., 2. ve 3. becerileri; K_1, K_2, K_3 sırasıyla 1., 2. ve 3. kavram yanılgılarını temsil etmektedir.

10 Nitelik için 20 Maddelik Q-matris

Maddeler	Nitelikler									
	Beceriler					Kavram Yanılgıları				
	B_1	B_2	B_3	B_4	B_5	K_1	K_2	K_3	K_4	K_5
1	1	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0
6	0	0	0	0	0	1	0	0	0	0
7	0	0	0	0	0	0	1	0	0	0
8	0	0	0	0	0	0	0	1	0	0
9	0	0	0	0	0	0	0	0	1	0
10	0	0	0	0	0	0	0	0	0	1
11	1	1	0	0	0	1	1	0	0	0
12	0	0	1	1	0	0	0	1	0	0
13	1	1	0	0	1	1	1	0	1	1
14	1	1	1	1	0	1	1	1	0	0
15	1	1	0	1	1	1	1	1	0	0
16	0	0	1	0	1	0	0	0	1	1
17	1	1	0	1	0	1	1	0	1	1
18	0	0	0	1	1	0	0	0	1	1
19	0	0	1	0	1	0	0	1	1	1
20	1	1	1	0	0	1	1	1	0	0

Tabloda yer alan B_1, B_2, B_3, B_4, B_5 sırasıyla 1., 2., 3., 4. ve 5. becerileri; K_1, K_2, K_3, K_4, K_5 sırasıyla 1., 2., 3., 4. ve 5. kavram yanılgılarını temsil etmektedir.

EK 3. 250 Kişilik Veri Seti için Örnek Olarak Oluşturulan Kod

```
##### Bilişsel Tanı Modeli için R paketinin yüklenmesi #####
install.packages("CDM")
library(CDM)

install.packages("GDINA")
library(GDINA)

##### Q-Matrislerinin Yüklenmesi #####
setwd("C:/Users/Hp/Desktop/busra")

#3+3 Q matris
q20_6 <- read.csv2("q20_6.csv", header = TRUE, sep = ";")
rownames(q20_6) <- q20_6[,1]
q20_6 <- q20_6[,-1]

q40_6 <- read.csv2("q40_6.csv", header = TRUE, sep = ";")
rownames(q40_6) <- q40_6[,1]
q40_6 <- q40_6[,-1]

q60_6 <- read.csv2("q60_6.csv", header = TRUE, sep = ";")
rownames(q60_6) <- q60_6[,1]
q60_6 <- q60_6[,-1]

#4+4 Q matris
q20_8 <- read.csv2("q20_8.csv", header = TRUE, sep = ";")
rownames(q20_8) <- q20_8[,1]
q20_8 <- q20_8[,-1]

q40_8 <- read.csv2("q40_8.csv", header = TRUE, sep = ";")
rownames(q40_8) <- q40_8[,1]
q40_8 <- q40_8[,-1]

q60_8 <- read.csv2("q60_8.csv", header = TRUE, sep = ";")
rownames(q60_8) <- q60_8[,1]
q60_8 <- q60_8[,-1]

#5+5 Q matris
q20_10 <- read.csv2("q20_10.csv", header = TRUE, sep = ";")
rownames(q20_10) <- q20_10[,1]
q20_10 <- q20_10[,-1]

q40_10 <- read.csv2("q40_10.csv", header = TRUE, sep = ";")
rownames(q40_10) <- q40_10[,1]
q40_10 <- q40_10[,-1]
```

```

q60_10 <- read.csv2("q60_10.csv", header = TRUE, sep = ";")
rownames(q60_10) <- q60_10[,1]
q60_10 <- q60_10[,-1]

```

```

#### Simülasyon verilerinin oluşturulması ####

```

```

N <- 250

```

```

Q <- q20_8

```

```

itemparm.list <- list(item1=c(0.35, 0.95),
  item2=c(0.35, 0.95),
  item3=c(0.35, 0.95),
  item4=c(0.35, 0.95),
  item5=c(0.95, 0.15),
  item6=c(0.95, 0.15),
  item7=c(0.95, 0.15),
  item8=c(0.95, 0.15),
  item9=c(0.35, 0.95, 0.05, 0.00),
  item10=c(0.35, 0.95, 0.05, 0.00),
  item11=c(0.35, 0.00, 0.35, 0.35, 0.05, 0.95, 0.00, 0.00),
  item12=c(0.35, 0.35, 0.35, 0.35, 0.05, 0.05, 0.35, 0.35, 0.00, 0.05, 0.35, 0.05,
    0.00, 0.05, 0.00, 0.05, 0.95, 0.00, 0.00, 0.00, 0.00, 0.00, 0.05, 0.00,
    0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00),
  item13=c(0.35, 0.35, 0.35, 0.35, 0.05, 0.05, 0.05, 0.35, 0.35, 0.35, 0.05, 0.05,
    0.05, 0.00, 0.05, 0.05, 0.05, 0.00, 0.05, 0.05, 0.05, 0.05, 0.05, 0.95,
    0.00, 0.00, 0.05, 0.00, 0.05, 0.00, 0.05, 0.05, 0.05, 0.00, 0.00, 0.00,
    0.00, 0.05, 0.05, 0.05, 0.00, 0.05, 0.00, 0.00, 0.15, 0.05, 0.00, 0.00,
    0.00, 0.00, 0.00, 0.00, 0.05, 0.00, 0.00, 0.00, 0.05, 0.00, 0.00, 0.00,
    0.00, 0.00, 0.00, 0.00),
  item14=c(0.35, 0.95, 0.05, 0.00),
  item15=c(0.35, 0.95, 0.05, 0.15, 0.15, 0.05, 0.15, 0.05),
  item16=c(0.35, 0.05, 0.35, 0.35, 0.05, 0.95, 0.05, 0.05,
    0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.15),
  item17=c(0.35, 0.35, 0.35, 0.05, 0.05, 0.05, 0.95, 0.00,
    0.05, 0.05, 0.00, 0.05, 0.05, 0.05, 0.05, 0.05,
    0.00, 0.15, 0.15, 0.00, 0.00, 0.05, 0.05, 0.05,
    0.05, 0.05, 0.00, 0.00, 0.15, 0.00, 0.05, 0.00),
  item18=c(0.35, 0.35, 0.35, 0.05, 0.05, 0.05, 0.95, 0.00,
    0.05, 0.05, 0.05, 0.00, 0.05, 0.05, 0.05, 0.05,
    0.00, 0.15, 0.15, 0.00, 0.00, 0.05, 0.05, 0.05,
    0.05, 0.05, 0.00, 0.00, 0.15, 0.00, 0.05, 0.00),
  item19=c(0.35, 0.15, 0.95, 0.05, 0.05, 0.15, 0.15, 0.05),
  item20=c(0.35, 0.05, 0.35, 0.35, 0.05, 0.05, 0.35, 0.35,
    0.05, 0.00, 0.05, 0.05, 0.00, 0.05, 0.05, 0.05,
    0.95, 0.05, 0.00, 0.05, 0.00, 0.00, 0.05, 0.00,
    0.00, 0.05, 0.15, 0.00, 0.00, 0.00, 0.00, 0.00))

```

```

sim <- simGDINA(N,Q,catprob.parm = itemparm.list)

```

```

# True delta parameters
extract(sim,what = "delta.parm")

# simulated data
dat <- extract(sim,what = "dat")

# simulated attributes
extract(sim,what = "attribute")

#SISM fitting
est <- GDINA(dat=dat,Q=Q,model="SISM",no.bugs=4)
coef(est)
data <- coef(est, withSE = TRUE) #estimations and standard errors
summary(est)
?GDINA

item_1 <- unlist(data[[1]])
est_item_1 <- item_1[1,]
est_item_1[3:4] <- NA
se_item_1 <- item_1[2,]
se_item_1[3:4] <- NA

item_2 <- unlist(data[[2]])
est_item_2 <- item_2[1,]
est_item_2[3:4] <- NA
se_item_2 <- item_2[2,]
se_item_2[3:4] <- NA

item_3 <- unlist(data[[3]])
est_item_3 <- item_3[1,]
est_item_3[3:4] <- NA
se_item_3 <- item_3[2,]
se_item_3[3:4] <- NA

item_4 <- unlist(data[[4]])
est_item_4 <- item_4[1,]
est_item_4[3:4] <- NA
se_item_4 <- item_4[2,]
se_item_4[3:4] <- NA

item_5 <- unlist(data[[5]])
est_item_5 <- item_5[1,]
est_item_5 <- c(NA, est_item_5[1], est_item_5[2], NA)
se_item_5 <- item_5[2,]
se_item_5 <- c(NA, se_item_5[1], se_item_5[2], NA)

```

```

item_6 <- unlist(data[[6]])
est_item_6 <- item_6[1,]
est_item_6 <- c(NA, est_item_6[1], est_item_6[2], NA)
se_item_6 <- item_6[2,]
se_item_6 <- c(NA, se_item_6[1], se_item_6[2], NA)

item_7 <- unlist(data[[7]])
est_item_7 <- item_7[1,]
est_item_7 <- c(NA, est_item_7[1], est_item_7[2], NA)
se_item_7 <- item_7[2,]
se_item_7 <- c(NA, se_item_7[1], se_item_7[2], NA)

item_8 <- unlist(data[[8]])
est_item_8 <- item_8[1,]
est_item_8 <- c(NA, est_item_8[1], est_item_8[2], NA)
se_item_8 <- item_8[2,]
se_item_8 <- c(NA, se_item_8[1], se_item_8[2], NA)

item_9 <- unlist(data[[9]])
est_item_9 <- item_9[1,]
est_item_9 <- c(est_item_9[1], est_item_9[2], NA, est_item_9[3])
se_item_9 <- item_9[2,]
se_item_9 <- c(se_item_9[1], se_item_9[2], NA, se_item_9[3])

item_10 <- unlist(data[[10]])
est_item_10 <- item_10[1,]
est_item_10 <- c(est_item_10[1], est_item_10[2], NA, est_item_10[3])
se_item_10 <- item_10[2,]
se_item_10 <- c(se_item_10[1], se_item_10[2], NA, se_item_10[3])

item_11 <- unlist(data[[11]])
est_item_11 <- item_11[1,]
est_item_11 <- c(est_item_11[1], est_item_11[6], NA, est_item_11[5])
se_item_11 <- item_11[2,]
se_item_11 <- c(se_item_11[1], se_item_11[6], NA, se_item_11[5])

item_12 <- unlist(data[[12]])
est_item_12 <- item_12[1,]
est_item_12 <- c(est_item_12[1], est_item_12[17], NA, est_item_12[5])
se_item_12 <- item_12[2,]
se_item_12 <- c(se_item_12[1], se_item_12[17], NA, se_item_12[5])

item_13 <- unlist(data[[13]])
est_item_13 <- item_13[1,]
est_item_13 <- c(est_item_13[1], est_item_13[24], NA, est_item_13[5])
se_item_13 <- item_13[2,]
se_item_13 <- c(se_item_13[1], se_item_13[24], NA, se_item_13[5])

```

```

item_14 <- unlist(data[[14]])
est_item_14 <- item_14[1,]
est_item_14 <- c(est_item_14[1], est_item_14[2], NA, est_item_14[3])
se_item_14 <- item_14[2,]
se_item_14 <- c(se_item_14[1], se_item_14[2], NA, se_item_14[3])

item_15 <- unlist(data[[15]])
est_item_15 <- item_15[1,]
est_item_15 <- c(est_item_15[1], est_item_15[2], est_item_15[4], est_item_15[3])
se_item_15 <- item_15[2,]
se_item_15 <- c(se_item_15[1], se_item_15[2], se_item_15[4], se_item_15[3])

item_16 <- unlist(data[[16]])
est_item_16 <- item_16[1,]
est_item_16 <- c(est_item_16[1], est_item_16[6], est_item_16[16], est_item_16[2])
se_item_16 <- item_16[2,]
se_item_16 <- c(se_item_16[1], se_item_16[6], se_item_16[16], se_item_16[2])

item_17 <- unlist(data[[17]])
est_item_17 <- item_17[1,]
est_item_17 <- c(est_item_17[1], est_item_17[7], est_item_17[18], est_item_17[4])
se_item_17 <- item_17[2,]
se_item_17 <- c(se_item_17[1], se_item_17[7], se_item_17[18], se_item_17[4])

item_18 <- unlist(data[[18]])
est_item_18 <- item_18[1,]
est_item_18 <- c(est_item_18[1], est_item_18[7], est_item_18[18], est_item_18[4])
se_item_18 <- item_18[2,]
se_item_18 <- c(se_item_18[1], se_item_18[7], se_item_18[18], se_item_18[4])

item_19 <- unlist(data[[19]])
est_item_19 <- item_19[1,]
est_item_19 <- c(est_item_19[1], est_item_19[3], est_item_19[2], est_item_19[4])
se_item_19 <- item_19[2,]
se_item_19 <- c(se_item_19[1], se_item_19[3], se_item_19[2], se_item_19[4])

item_20 <- unlist(data[[20]])
est_item_20 <- item_20[1,]
est_item_20 <- c(est_item_20[1], est_item_20[17], est_item_20[27], est_item_20[5])
se_item_20 <- item_20[2,]
se_item_20 <- c(se_item_20[1], se_item_20[17], se_item_20[27], se_item_20[5])

est_data <- rbind(est_item_1, est_item_2, est_item_3, est_item_4, est_item_5, est_item_6,
est_item_7, est_item_8, est_item_9, est_item_10, est_item_11,
est_item_12, est_item_13, est_item_14, est_item_15, est_item_16, est_item_17,
est_item_18, est_item_19, est_item_20)

```

```
se_data <- rbind(se_item_1, se_item_2, se_item_3, se_item_4, se_item_5, se_item_6,
se_item_7, se_item_8, se_item_9, se_item_10, se_item_11,
                se_item_12, se_item_13, se_item_14, se_item_15, se_item_16, se_item_17,
se_item_18, se_item_19, se_item_20)
```

```
colnames(est_data) <- c("g", "h", "w", "e")
est_data <- est_data[, c("h", "g", "w", "e")]
colnames(se_data) <- c("g", "h", "w", "e")
se_data <- se_data[, c("h", "g", "w", "e")]
```

```
true_data <- matrix(rep(c(0.95, 0.35, 0.15, 0.05), 20), ncol = 4, byrow = TRUE)
colnames(true_data) <- c("h", "g", "w", "e")
```

Absolute Deviation (AD) and Mean Absolute Deviation (MAD)

```
abs_dev <- abs(est_data-true_data)
mad_h <- mean(abs_dev[, "h"], na.rm = TRUE)
mad_g <- mean(abs_dev[, "g"], na.rm = TRUE)
mad_w <- mean(abs_dev[, "w"], na.rm = TRUE)
mad_e <- mean(abs_dev[, "e"], na.rm = TRUE)
```

Standart Error (SE)

```
average_se_h <- mean(se_data[, "h"], na.rm = TRUE)
average_se_g <- mean(se_data[, "g"], na.rm = TRUE)
average_se_w <- mean(se_data[, "w"], na.rm = TRUE)
average_se_e <- mean(se_data[, "e"], na.rm = TRUE)
```

#Root Mean Square Error (RMSE)

```
rmse_h <- sqrt(mean((abs_dev[, "h"])^2, na.rm = TRUE))
rmse_g <- sqrt(mean((abs_dev[, "g"])^2, na.rm = TRUE))
rmse_w <- sqrt(mean((abs_dev[, "w"])^2, na.rm = TRUE))
rmse_e <- sqrt(mean((abs_dev[, "e"])^2, na.rm = TRUE))
```

Simülasyon ve İterasyon Bölümü

Bu senaryo 250 kişilik örnekleme ait veri için 8 nitelik, 20 madde ve yüksek ayırt ediciliğe sahip maddelerden oluşan Q-matris içindir

N <- 250

Q <- q20_8

```
itemparm.list <- list(item1=c(0.35, 0.95),
                    item2=c(0.35, 0.95),
                    item3=c(0.35, 0.95),
                    item4=c(0.35, 0.95),
                    item5=c(0.95, 0.15),
                    item6=c(0.95, 0.15),
                    item7=c(0.95, 0.15),
                    item8=c(0.95, 0.15),
                    item9=c(0.35, 0.95, 0.05, 0.00),
```

```

item10=c(0.35, 0.95, 0.05, 0.00),
item11=c(0.35, 0.00, 0.35, 0.35, 0.05, 0.95, 0.00, 0.00),
item12=c(0.35, 0.35, 0.35, 0.35, 0.05, 0.05, 0.35, 0.35, 0.00, 0.05, 0.35, 0.05,
0.00, 0.05, 0.00, 0.05, 0.95, 0.00, 0.00, 0.00, 0.00, 0.00, 0.05, 0.00,
0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00),
item13=c(0.35, 0.35, 0.35, 0.35, 0.05, 0.05, 0.05, 0.35, 0.35, 0.35, 0.05, 0.05,
0.05, 0.00, 0.05, 0.05, 0.05, 0.00, 0.05, 0.05, 0.05, 0.05, 0.05, 0.95,
0.00, 0.00, 0.05, 0.00, 0.05, 0.00, 0.05, 0.05, 0.05, 0.00, 0.00, 0.00,
0.00, 0.05, 0.05, 0.05, 0.00, 0.05, 0.00, 0.00, 0.15, 0.05, 0.00, 0.00,
0.00, 0.00, 0.00, 0.00, 0.05, 0.00, 0.00, 0.00, 0.05, 0.00, 0.00, 0.00,
0.00, 0.00, 0.00, 0.00),
item14=c(0.35, 0.95, 0.05, 0.00),
item15=c(0.35, 0.95, 0.05, 0.15, 0.15, 0.05, 0.15, 0.05),
item16=c(0.35, 0.05, 0.35, 0.35, 0.05, 0.95, 0.05, 0.05,
0.05, 0.05, 0.05, 0.05, 0.05, 0.05, 0.15),
item17=c(0.35, 0.35, 0.35, 0.05, 0.05, 0.05, 0.95, 0.00,
0.05, 0.05, 0.00, 0.05, 0.05, 0.05, 0.05, 0.05,
0.00, 0.15, 0.15, 0.00, 0.00, 0.05, 0.05, 0.05,
0.05, 0.05, 0.00, 0.00, 0.15, 0.00, 0.05, 0.00),
item18=c(0.35, 0.35, 0.35, 0.05, 0.05, 0.05, 0.95, 0.00,
0.05, 0.05, 0.05, 0.00, 0.05, 0.05, 0.05, 0.05,
0.00, 0.15, 0.15, 0.00, 0.00, 0.05, 0.05, 0.05,
0.05, 0.05, 0.00, 0.00, 0.15, 0.00, 0.05, 0.00),
item19=c(0.35, 0.15, 0.95, 0.05, 0.05, 0.15, 0.15, 0.05),
item20=c(0.35, 0.05, 0.35, 0.35, 0.05, 0.05, 0.35, 0.35,
0.05, 0.00, 0.05, 0.05, 0.00, 0.05, 0.05, 0.05,
0.95, 0.05, 0.00, 0.05, 0.00, 0.00, 0.05, 0.00,
0.00, 0.05, 0.15, 0.00, 0.00, 0.00, 0.00, 0.00))

```

```

abs_dev_h <- {}
abs_dev_g <- {}
abs_dev_w <- {}
abs_dev_e <- {}

```

```

mad_dev_h <- {}
mad_dev_g <- {}
mad_dev_w <- {}
mad_dev_e <- {}

```

```

ave_se_h <- {}
ave_se_g <- {}
ave_se_w <- {}
ave_se_e <- {}

```

```

ave_rmse_h <- {}
ave_rmse_g <- {}
ave_rmse_w <- {}
ave_rmse_e <- {}

```

```

rep<-100
K<-ncol(Q)
class.accuracy<-matrix(NA, ncol=2,nrow=rep, byrow = T)
colnames(class.accuracy)<-c("Average_ARR", "PRR")

att.recovery<-matrix(NA, ncol=K,nrow=rep, byrow = T)

att.rate <- function(A,B) {
  att.accuracy<-matrix(NA, nrow=nrow(A), ncol=ncol(A), byrow = T)
  for (k in 1:nrow(A))
  { for (m in 1:ncol(A))
    { if (A[k,m]==B[k,m])
      { att.accuracy[k,m]<-1
        } else {
          att.accuracy[k,m]<-0
        } } }
  apply(att.accuracy, 2, mean)
}

for (i in 1:rep){
  sim <- simGDINA(N,Q,catprob.parm = itemparm.list)

  dat <- extract(sim,what = "dat")
  est <- GDINA(dat=dat,Q=Q,model="SISM",no.bugs=4)
  data <- coef(est, withSE = TRUE)

  obs.pattern<-sim$attribute
  est.pattern<-personparm(est)

  cr<-ClassRate(obs.pattern,est.pattern)
  class.accuracy[i,]<-c(cr$PCA,cr$PCV[[K]])

  att.recovery[i,]<-att.rate(obs.pattern,est.pattern)

  item_1 <- unlist(data[[1]])
  est_item_1 <- item_1[1,]
  est_item_1[3:4] <- NA
  se_item_1 <- item_1[2,]
  se_item_1[3:4] <- NA

  item_2 <- unlist(data[[2]])
  est_item_2 <- item_2[1,]
  est_item_2[3:4] <- NA
  se_item_2 <- item_2[2,]
  se_item_2[3:4] <- NA

```

```

item_3 <- unlist(data[[3]])
est_item_3 <- item_3[1,]
est_item_3[3:4] <- NA
se_item_3 <- item_3[2,]
se_item_3[3:4] <- NA

item_4 <- unlist(data[[4]])
est_item_4 <- item_4[1,]
est_item_4[3:4] <- NA
se_item_4 <- item_4[2,]
se_item_4[3:4] <- NA

item_5 <- unlist(data[[5]])
est_item_5 <- item_5[1,]
est_item_5 <- c(NA, est_item_5[1], est_item_5[2], NA)
se_item_5 <- item_5[2,]
se_item_5 <- c(NA, se_item_5[1], se_item_5[2], NA)

item_6 <- unlist(data[[6]])
est_item_6 <- item_6[1,]
est_item_6 <- c(NA, est_item_6[1], est_item_6[2], NA)
se_item_6 <- item_6[2,]
se_item_6 <- c(NA, se_item_6[1], se_item_6[2], NA)

item_7 <- unlist(data[[7]])
est_item_7 <- item_7[1,]
est_item_7 <- c(NA, est_item_7[1], est_item_7[2], NA)
se_item_7 <- item_7[2,]
se_item_7 <- c(NA, se_item_7[1], se_item_7[2], NA)

item_8 <- unlist(data[[8]])
est_item_8 <- item_8[1,]
est_item_8 <- c(NA, est_item_8[1], est_item_8[2], NA)
se_item_8 <- item_8[2,]
se_item_8 <- c(NA, se_item_8[1], se_item_8[2], NA)

item_9 <- unlist(data[[9]])
est_item_9 <- item_9[1,]
est_item_9 <- c(est_item_9[1], est_item_9[2], NA, est_item_9[3])
se_item_9 <- item_9[2,]
se_item_9 <- c(se_item_9[1], se_item_9[2], NA, se_item_9[3])

item_10 <- unlist(data[[10]])
est_item_10 <- item_10[1,]
est_item_10 <- c(est_item_10[1], est_item_10[2], NA, est_item_10[3])
se_item_10 <- item_10[2,]
se_item_10 <- c(se_item_10[1], se_item_10[2], NA, se_item_10[3])

```

```

item_11 <- unlist(data[[11]])
est_item_11 <- item_11[1,]
est_item_11 <- c(est_item_11[1], est_item_11[6], NA, est_item_11[5])
se_item_11 <- item_11[2,]
se_item_11 <- c(se_item_11[1], se_item_11[6], NA, se_item_11[5])

item_12 <- unlist(data[[12]])
est_item_12 <- item_12[1,]
est_item_12 <- c(est_item_12[1], est_item_12[17], NA, est_item_12[5])
se_item_12 <- item_12[2,]
se_item_12 <- c(se_item_12[1], se_item_12[17], NA, se_item_12[5])

item_13 <- unlist(data[[13]])
est_item_13 <- item_13[1,]
est_item_13 <- c(est_item_13[1], est_item_13[24], NA, est_item_13[5])
se_item_13 <- item_13[2,]
se_item_13 <- c(se_item_13[1], se_item_13[24], NA, se_item_13[5])

item_14 <- unlist(data[[14]])
est_item_14 <- item_14[1,]
est_item_14 <- c(est_item_14[1], est_item_14[2], NA, est_item_14[3])
se_item_14 <- item_14[2,]
se_item_14 <- c(se_item_14[1], se_item_14[2], NA, se_item_14[3])

item_15 <- unlist(data[[15]])
est_item_15 <- item_15[1,]
est_item_15 <- c(est_item_15[1], est_item_15[2], est_item_15[4], est_item_15[3])
se_item_15 <- item_15[2,]
se_item_15 <- c(se_item_15[1], se_item_15[2], se_item_15[4], se_item_15[3])

item_16 <- unlist(data[[16]])
est_item_16 <- item_16[1,]
est_item_16 <- c(est_item_16[1], est_item_16[6], est_item_16[16], est_item_16[2])
se_item_16 <- item_16[2,]
se_item_16 <- c(se_item_16[1], se_item_16[6], se_item_16[16], se_item_16[2])

item_17 <- unlist(data[[17]])
est_item_17 <- item_17[1,]
est_item_17 <- c(est_item_17[1], est_item_17[7], est_item_17[18], est_item_17[4])
se_item_17 <- item_17[2,]
se_item_17 <- c(se_item_17[1], se_item_17[7], se_item_17[18], se_item_17[4])

item_18 <- unlist(data[[18]])
est_item_18 <- item_18[1,]
est_item_18 <- c(est_item_18[1], est_item_18[7], est_item_18[18], est_item_18[4])
se_item_18 <- item_18[2,]
se_item_18 <- c(se_item_18[1], se_item_18[7], se_item_18[18], se_item_18[4])

```

```

item_19 <- unlist(data[[19]])
est_item_19 <- item_19[1,]
est_item_19 <- c(est_item_19[1], est_item_19[3], est_item_19[2], est_item_19[4])
se_item_19 <- item_19[2,]
se_item_19 <- c(se_item_19[1], se_item_19[3], se_item_19[2], se_item_19[4])

item_20 <- unlist(data[[20]])
est_item_20 <- item_20[1,]
est_item_20 <- c(est_item_20[1], est_item_20[17], est_item_20[27], est_item_20[5])
se_item_20 <- item_20[2,]
se_item_20 <- c(se_item_20[1], se_item_20[17], se_item_20[27], se_item_20[5])

est_data <- rbind(est_item_1, est_item_2, est_item_3, est_item_4, est_item_5, est_item_6,
est_item_7, est_item_8, est_item_9, est_item_10, est_item_11,
est_item_12, est_item_13, est_item_14, est_item_15, est_item_16,
est_item_17, est_item_18, est_item_19, est_item_20)

se_data <- rbind(se_item_1, se_item_2, se_item_3, se_item_4, se_item_5, se_item_6,
se_item_7, se_item_8, se_item_9, se_item_10, se_item_11,
se_item_12, se_item_13, se_item_14, se_item_15, se_item_16, se_item_17,
se_item_18, se_item_19, se_item_20)

colnames(est_data) <- c("g", "h", "w", "e")
est_data <- est_data[, c("h", "g", "w", "e")]
colnames(se_data) <- c("g", "h", "w", "e")
se_data <- se_data[, c("h", "g", "w", "e")]

true_data <- matrix(rep(c(0.95, 0.35, 0.15, 0.05), 20), ncol = 4, byrow = TRUE)
colnames(true_data) <- c("h", "g", "w", "e")

abs_dev <- abs(est_data-true_data)

mad_h <- mean(abs_dev[,"h"], na.rm = TRUE)
mad_g <- mean(abs_dev[,"g"], na.rm = TRUE)
mad_w <- mean(abs_dev[,"w"], na.rm = TRUE)
mad_e <- mean(abs_dev[,"e"], na.rm = TRUE)

average_se_h <- mean(se_data[,"h"], na.rm = TRUE)
average_se_g <- mean(se_data[,"g"], na.rm = TRUE)
average_se_w <- mean(se_data[,"w"], na.rm = TRUE)
average_se_e <- mean(se_data[,"e"], na.rm = TRUE)

rmse_h <- sqrt(mean((abs_dev[,"h"])^2, na.rm = TRUE))
rmse_g <- sqrt(mean((abs_dev[,"g"])^2, na.rm = TRUE))
rmse_w <- sqrt(mean((abs_dev[,"w"])^2, na.rm = TRUE))
rmse_e <- sqrt(mean((abs_dev[,"e"])^2, na.rm = TRUE))

```

```

abs_dev_h <- cbind(abs_dev_h,abs_dev[, 1])
abs_dev_g <- cbind(abs_dev_g,abs_dev[, 2])
abs_dev_w <- cbind(abs_dev_w,abs_dev[, 3])
abs_dev_e <- cbind(abs_dev_e,abs_dev[, 4])

mad_dev_h <- rbind(mad_dev_h, mad_h)
mad_dev_g <- rbind(mad_dev_g, mad_g)
mad_dev_w <- rbind(mad_dev_w, mad_w)
mad_dev_e <- rbind(mad_dev_e, mad_e)

ave_se_h <- rbind(ave_se_h, average_se_h)
ave_se_g <- rbind(ave_se_g, average_se_g)
ave_se_w <- rbind(ave_se_w, average_se_w)
ave_se_e <- rbind(ave_se_e, average_se_e)

ave_rmse_h <- rbind(ave_rmse_h, rmse_h)
ave_rmse_g <- rbind(ave_rmse_g, rmse_g)
ave_rmse_w <- rbind(ave_rmse_w, rmse_w)
ave_rmse_e <- rbind(ave_rmse_e, rmse_e)

}

average.att.accuracy<-apply(att.recovery, 2, mean)
average.class.accuracy<-apply(class.accuracy, 2, mean)

abs_dev_all_h <- rowMeans(abs_dev_h, na.rm = TRUE)
abs_dev_all_g <- rowMeans(abs_dev_g, na.rm = TRUE)
abs_dev_all_w <- rowMeans(abs_dev_w, na.rm = TRUE)
abs_dev_all_e <- rowMeans(abs_dev_e, na.rm = TRUE)

mad_dev_all_h <- mean(mad_dev_h, na.rm = TRUE)
mad_dev_all_g <- mean(mad_dev_g, na.rm = TRUE)
mad_dev_all_w <- mean(mad_dev_w, na.rm = TRUE)
mad_dev_all_e <- mean(mad_dev_e, na.rm = TRUE)

mean_ave_se_h <- mean(ave_se_h, na.rm = TRUE)
mean_ave_se_g <- mean(ave_se_g, na.rm = TRUE)
mean_ave_se_w <- mean(ave_se_w, na.rm = TRUE)
mean_ave_se_e <- mean(ave_se_e, na.rm = TRUE)

mean_ave_rmse_h <- mean(ave_rmse_h, na.rm = TRUE)
mean_ave_rmse_g <- mean(ave_rmse_g, na.rm = TRUE)
mean_ave_rmse_w <- mean(ave_rmse_w, na.rm = TRUE)
mean_ave_rmse_e <- mean(ave_rmse_e, na.rm = TRUE)

table_1 <- round(cbind(abs_dev_all_h, abs_dev_all_g, abs_dev_all_w, abs_dev_all_e),2)
colnames(table_1) <- c("h", "g", "w", "e")

mad_table_1 <- round(rbind(mad_dev_all_h, mad_dev_all_g, mad_dev_all_w,
mad_dev_all_e),2)

```

```
se_table_3 <- round(rbind(mean_ave_se_h, mean_ave_se_g, mean_ave_se_w,  
mean_ave_se_e),2)
```

```
rmse_table <- round(rbind(mean_ave_rmse_h, mean_ave_rmse_g, mean_ave_rmse_w,  
mean_ave_rmse_e),2)
```



EK 4. 8 ve 10 Nitelikli Q-matrisler için “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerlerine İlişkin Tablolar

8 Niteliği Ölçen Yüksek Ayırt Ediciliğe Sahip 20 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,15	,14	,13	,09	,34	,32	,23	,13	-	-	-	-	-	-	-	-
2	,13	,12	,11	,16	,24	,18	,07	,05	-	-	-	-	-	-	-	-
3	,14	,14	,14	,11	,32	,29	,18	,05	-	-	-	-	-	-	-	-
4	,15	,15	,12	,08	,37	,34	,26	,14	-	-	-	-	-	-	-	-
5	,04	,03	,03	,03	-	-	-	-	,07	,05	,03	,02	-	-	-	-
6	,04	,03	,01	,01	-	-	-	-	,06	,03	,02	,02	-	-	-	-
7	,09	,05	,04	,04	-	-	-	-	,13	,11	,10	,10	-	-	-	-
8	,10	,06	,03	,02	-	-	-	-	,15	,13	,08	,06	-	-	-	-
9	,12	,12	,11	,08	,13	,12	,09	,09	-	-	-	-	,07	,05	,04	,04
10	,10	,10	,10	,12	,13	,09	,09	,09	-	-	-	-	,07	,04	,04	,04
11	,48	,45	,38	,35	,14	,14	,15	,18	-	-	-	-	,07	,07	,06	,08
12	,24	,22	,18	,15	,10	,07	,04	,03	-	-	-	-	,07	,07	,08	,08
13	,77	,78	,78	,78	,15	,10	,07	,08	-	-	-	-	,03	,03	,03	,03
14	,12	,12	,09	,05	,11	,10	,08	,07	-	-	-	-	,05	,04	,02	,02
15	,18	,14	,14	,09	,13	,12	,09	,09	,05	,05	,05	,04	,08	,07	,07	,07
16	,27	,24	,21	,20	,11	,09	,11	,15	,09	,10	,10	,11	,26	,25	,20	,15
17	,21	,20	,17	,11	,17	,13	,07	,06	,12	,13	,13	,14	,03	,03	,03	,03
18	,24	,19	,17	,11	,16	,11	,06	,05	,12	,13	,14	,14	,03	,03	,03	,03
19	,50	,49	,32	,18	,14	,11	,13	,12	,12	,08	,07	,07	,25	,27	,27	,77
20	,28	,24	,21	,20	,09	,06	,06	,06	,15	,16	,19	,19	,05	,05	,05	,05
<i>Ortalama Mutlak Sapma</i>	,18	,16	,13	,13	,13	,10	,07	,07	,08	,08	,07	,06	,07	,07	,07	,07

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı.

8 Niteliği Ölçen Düşük Ayırt Ediciliğe Sahip 20 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,05	,05	,04	,03	,22	,21	,09	,07	-	-	-	-	-	-	-	-
2	,05	,08	,17	,19	,13	,07	,05	,04	-	-	-	-	-	-	-	-
3	,05	,04	,05	,05	,21	,18	,08	,03	-	-	-	-	-	-	-	-
4	,05	,05	,04	,03	,28	,20	,08	,05	-	-	-	-	-	-	-	-
5	,03	,02	,01	,01	-	-	-	-	,06	,04	,03	,03	-	-	-	-
6	,02	,02	,01	,01	-	-	-	-	,04	,03	,02	,01	-	-	-	-
7	,04	,03	,02	,02	-	-	-	-	,08	,10	,07	,07	-	-	-	-
8	,03	,02	,02	,01	-	-	-	-	,12	,09	,06	,04	-	-	-	-
9	,09	,09	,04	,03	,08	,06	,05	,04	-	-	-	-	,03	,03	,02	,02
10	,06	,05	,10	,12	,08	,07	,09	,11	-	-	-	-	,03	,03	,02	,02
11	,53	,50	,43	,43	,12	,12	,14	,14	-	-	-	-	,04	,04	,05	,06
12	,15	,14	,11	,12	,07	,05	,03	,01	-	-	-	-	,04	,04	,04	,04
13	,90	,90	,90	,91	,11	,07	,06	,06	-	-	-	-	,01	,01	,01	,01
14	,10	,09	,06	,02	,08	,07	,05	,05	-	-	-	-	,04	,02	,01	,01
15	,12	,07	,03	,03	,12	,07	,04	,04	,04	,04	,03	,03	,07	,07	,07	,07
16	,17	,12	,10	,13	,11	,10	,14	,15	,09	,08	,10	,10	,21	,21	,16	,15
17	,14	,11	,06	,04	,14	,08	,05	,04	,10	,09	,09	,09	,02	,02	,01	,02
18	,17	,14	,08	,04	,10	,08	,05	,04	,10	,09	,09	,09	,08	,02	,01	,01
19	,58	,51	,28	,21	,11	,09	,08	,09	,08	,07	,05	,06	,32	,39	,65	,74
20	,21	,20	,14	,13	,09	,08	,08	,09	,11	,12	,12	,13	,02	,02	,02	,02
Ortalama Mutlak Sapma	,22	,20	,17	,15	,18	,15	,11	,09	,10	,10	,09	,09	,09	,08	,08	,08

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı.

8 Niteliği Ölçen Yüksek Ayırt Ediciliğe Sahip 40 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,03	,03	,01	,01	,06	,04	,02	,01	-	-	-	-	-	-	-	-
2	,05	,09	,12	,14	,06	,05	,04	,04	-	-	-	-	-	-	-	-
3	,03	,02	,01	,01	,05	,03	,02	,01	-	-	-	-	-	-	-	-
4	,04	,03	,02	,01	,07	,05	,02	,02	-	-	-	-	-	-	-	-
5	,02	,01	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
6	,02	,01	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
7	,03	,02	,01	,01	-	-	-	-	,04	,03	,01	,01	-	-	-	-
8	,03	,02	,01	,01	-	-	-	-	,04	,03	,01	,01	-	-	-	-
9	,04	,03	,01	,01	,06	,04	,02	,02	-	-	-	-	,03	,02	,01	,01
10	,04	,04	,05	,06	,06	,05	,05	,05	-	-	-	-	,03	,03	,02	,01
11	,48	,42	,41	,39	,12	,12	,12	,12	-	-	-	-	,04	,03	,02	,02
12	,09	,07	,05	,03	,05	,03	,02	,02	-	-	-	-	,03	,03	,03	,03
13	,90	,90	,90	,90	,08	,07	,04	,04	-	-	-	-	,01	,01	,01	,00
14	,04	,02	,01	,01	,05	,04	,03	,02	-	-	-	-	,02	,02	,01	,01
15	,05	,03	,02	,01	,08	,06	,03	,02	,05	,04	,04	,04	,07	,07	,06	,06
16	,06	,06	,04	,05	,11	,12	,12	,12	,09	,08	,09	,09	,19	,18	,18	,18
17	,09	,06	,03	,02	,09	,06	,04	,02	,09	,08	,09	,08	,02	,01	,01	,01
18	,08	,06	,04	,02	,09	,06	,04	,02	,08	,09	,08	,08	,02	,01	,01	,01
19	,51	,44	,39	,36	,10	,08	,04	,03	,07	,05	,04	,04	,39	,46	,51	,54
20	,10	,05	,04	,03	,10	,09	,09	,09	,11	,12	,12	,12	,02	,02	,02	,02
21	,04	,03	,01	,01	,05	,04	,02	,01	-	-	-	-	-	-	-	-
22	,05	,09	,12	,14	,05	,05	,04	,04	-	-	-	-	-	-	-	-
23	,03	,02	,01	,01	,05	,03	,02	,01	-	-	-	-	-	-	-	-
24	,05	,03	,02	,01	,06	,04	,02	,02	-	-	-	-	-	-	-	-
25	,02	,01	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
26	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
27	,03	,02	,01	,01	-	-	-	-	,04	,03	,01	,01	-	-	-	-
28	,03	,02	,01	,01	-	-	-	-	,04	,03	,01	,01	-	-	-	-
29	,03	,03	,01	,01	,06	,05	,02	,02	-	-	-	-	,03	,02	,01	-
30	,04	,04	,05	,05	,06	,05	,05	,05	-	-	-	-	,03	,03	,02	-
31	,48	,42	,41	,39	,12	,12	,12	,12	-	-	-	-	,04	,03	,02	,02
32	,09	,06	,04	,03	,06	,04	,02	,02	-	-	-	-	,03	,03	,03	,03
33	,90	,90	,90	,90	,09	,06	,04	,04	-	-	-	-	,01	,01	,01	,00
34	,03	,03	,01	,01	,07	,04	,02	,02	-	-	-	-	,02	,02	,01	,01
35	,05	,03	,02	,01	,09	,06	,04	,02	,04	,04	,03	,03	,07	,07	,07	,07
36	,06	,06	,05	,05	,12	,12	,12	,13	,08	,09	,09	,09	,18	,18	,18	,17
37	,07	,06	,03	,02	,10	,07	,03	,02	,09	,08	,09	,08	,02	,02	,01	,01
38	,07	,06	,04	,02	,09	,06	,04	,02	,09	,09	,08	,08	,02	,01	,01	,01
39	,50	,44	,39	,36	,09	,07	,04	,03	,07	,05	,03	,04	,40	,46	,51	,54
40	,09	,07	,04	,03	,10	,09	,09	,09	,11	,11	,12	,12	,02	,03	,02	,02
Ortalama Mutlak Sapma	,13	,12	,11	,10	,08	,06	,05	,04	,06	,06	,05	,05	,07	,07	,06	,06

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.

8 Niteliği Ölçen Düşük Ayırt Ediciliğe Sahip 40 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,05	,05	,03	,02	,06	,06	,03	,02	-	-	-	-	-	-	-	-
2	,07	,12	,16	,16	,05	,06	,06	,05	-	-	-	-	-	-	-	-
3	,06	,03	,02	,01	,09	,04	,02	,02	-	-	-	-	-	-	-	-
4	,08	,06	,03	,03	,07	,06	,04	,03	-	-	-	-	-	-	-	-
5	,03	,03	,02	,01	-	-	-	-	,05	,03	,02	,01	-	-	-	-
6	,03	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
7	,05	,04	,02	,02	-	-	-	-	,06	,04	,02	,02	-	-	-	-
8	,07	,03	,02	,02	-	-	-	-	,05	,04	,02	,02	-	-	-	-
9	,07	,05	,03	,02	,06	,06	,04	,04	-	-	-	-	,06	,04	,02	,02
10	,07	,07	,12	,13	,14	,08	,07	,07	-	-	-	-	,05	,06	,04	,03
11	,33	,27	,24	,24	,12	,13	,15	,16	-	-	-	-	,07	,05	,07	,06
12	,16	,13	,11	,10	,07	,04	,03	,02	-	-	-	-	,08	,07	,07	,07
13	,78	,78	,78	,78	,07	,07	,05	,05	-	-	-	-	,04	,03	,03	,03
14	,09	,05	,03	,02	,08	,07	,03	,02	-	-	-	-	,06	,03	,01	,01
15	,09	,07	,03	,02	,13	,08	,05	,04	,09	,05	,04	,04	,06	,08	,06	,06
16	,12	,11	,09	,10	,11	,13	,15	,16	,11	,10	,10	,10	,22	,18	,15	,14
17	,13	,12	,07	,04	,13	,07	,04	,03	,13	,11	,13	,13	,03	,03	,03	,02
18	,17	,11	,08	,07	,10	,08	,04	,03	,16	,12	,12	,13	,03	,03	,03	,03
19	,41	,26	,08	,05	,19	,11	,06	,05	,07	,07	,07	,07	,35	,50	,71	,75
20	,16	,14	,10	,09	,10	,09	,09	,10	,17	,17	,18	,18	,05	,05	,04	,04
21	,07	,06	,03	,02	,06	,06	,03	,02	-	-	-	-	-	-	-	-
22	,08	,12	,16	,16	,08	,06	,05	,05	-	-	-	-	-	-	-	-
23	,04	,03	,02	,01	,11	,04	,02	,01	-	-	-	-	-	-	-	-
24	,08	,06	,03	,03	,09	,07	,04	,03	-	-	-	-	-	-	-	-
25	,04	,02	,02	,01	-	-	-	-	,05	,02	,02	,01	-	-	-	-
26	,02	,02	,01	,01	-	-	-	-	,02	,02	,01	,01	-	-	-	-
27	,06	,04	,02	,02	-	-	-	-	,06	,04	,02	,02	-	-	-	-
28	,06	,04	,02	,02	-	-	-	-	,08	,03	,02	,02	-	-	-	-
29	,07	,05	,03	,02	,07	,07	,04	,03	-	-	-	-	,06	,04	,02	,02
30	,04	,07	,12	,13	,09	,07	,07	,07	-	-	-	-	,06	,06	,04	,03
31	,34	,27	,24	,24	,11	,14	,16	,16	-	-	-	-	,07	,04	,06	,07
32	,14	,14	,10	,10	,06	,05	,03	,02	-	-	-	-	,08	,07	,07	,07
33	,78	,78	,78	,78	,11	,07	,05	,05	-	-	-	-	,03	,04	,03	,03
34	,11	,06	,03	,02	,11	,07	,03	,02	-	-	-	-	,04	,03	,02	,01
35	,09	,06	,03	,02	,06	,09	,05	,04	,07	,05	,04	,04	,07	,07	,07	,06
36	,17	,11	,09	,11	,12	,13	,15	,16	,09	,09	,10	,10	,21	,17	,15	,14
37	,12	,11	,06	,05	,11	,07	,04	,03	,11	,12	,12	,13	,03	,03	,03	,03
38	,13	,10	,09	,07	,12	,07	,04	,03	,13	,12	,12	,12	,03	,03	,03	,03
39	,43	,25	,09	,04	,12	,13	,07	,05	,10	,08	,07	,07	,32	,51	,72	,75
40	,12	,14	,10	,09	,14	,10	,09	,09	,15	,17	,17	,18	,05	,05	,05	,04
Ortalama Mutlak Sapma	,15	,13	,10	,10	,10	,08	,06	,05	,09	,08	,07	,07	,09	,09	,08	,08

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı.

8 Niteliği Ölçen Yüksek Ayırt Ediciliğe Sahip 60 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,02	,02	,01	,01	,04	,03	,01	,01	-	-	-	-	-	-	-	-
2	,05	,08	,10	,10	,04	,04	,03	,03	-	-	-	-	-	-	-	-
3	,02	,01	,01	,01	,04	,02	,01	,01	-	-	-	-	-	-	-	-
4	,03	,02	,01	,01	,04	,03	,02	,01	-	-	-	-	-	-	-	-
5	,02	,01	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
6	,02	,01	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
7	,02	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
8	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
9	,02	,02	,01	,01	,05	,03	,02	,01	-	-	-	-	,03	,02	,01	,01
10	,03	,03	,02	,02	,05	,04	,03	,03	-	-	-	-	,03	,02	,01	,01
11	,47	,43	,41	,41	,12	,12	,12	,11	-	-	-	-	,03	,03	,01	,01
12	,08	,05	,03	,02	,05	,04	,02	,02	-	-	-	-	,03	,03	,03	,03
13	,90	,90	,90	,90	,09	,07	,05	,04	-	-	-	-	,02	,01	,01	,00
14	,03	,02	,01	,01	,06	,03	,02	,02	-	-	-	-	,02	,01	,01	,01
15	,04	,02	,01	,01	,07	,05	,03	,02	,04	,04	,03	,03	,07	,06	,07	,07
16	,05	,05	,04	,03	,11	,11	,11	,11	,08	,08	,08	,08	,19	,19	,19	,19
17	,07	,05	,03	,02	,08	,06	,03	,03	,09	,09	,08	,09	,02	,01	,01	,01
18	,07	,05	,03	,02	,08	,06	,03	,02	,08	,08	,08	,08	,02	,01	,01	,01
19	,50	,46	,44	,45	,08	,06	,03	,03	,06	,04	,03	,03	,40	,44	,47	,45
20	,07	,06	,03	,02	,10	,09	,09	,09	,11	,12	,11	,11	,02	,02	,02	,02
21	,02	,02	,01	,01	,04	,03	,02	,01	-	-	-	-	-	-	-	-
22	,05	,08	,10	,10	,05	,04	,03	,03	-	-	-	-	-	-	-	-
23	,02	,01	,01	,01	,04	,02	,01	,01	-	-	-	-	-	-	-	-
24	,03	,02	,01	,01	,04	,04	,02	,01	-	-	-	-	-	-	-	-
25	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
26	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
27	,02	,01	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
28	,02	,01	,01	,01	-	-	-	-	,04	,02	,01	,01	-	-	-	-
29	,03	,02	,01	,01	,05	,04	,02	,01	-	-	-	-	,02	,02	,01	,01
30	,03	,03	,02	,02	,04	,04	,03	,02	-	-	-	-	,03	,03	,02	,01
31	,46	,43	,41	,41	,11	,11	,12	,12	-	-	-	-	,03	,02	,02	,01
32	,06	,06	,03	,02	,05	,03	,02	,01	-	-	-	-	,03	,03	,03	,03
33	,90	,90	,90	,90	,08	,06	,04	,04	-	-	-	-	,01	,01	,01	,00
34	,03	,02	,01	,01	,05	,04	,02	,02	-	-	-	-	,02	,02	,01	,01
35	,03	,03	,02	,01	,07	,05	,03	,02	,05	,04	,03	,03	,06	,07	,07	,07
36	,07	,05	,03	,03	,11	,12	,11	,11	,09	,07	,08	,08	,19	,18	,19	,19
37	,07	,05	,03	,02	,09	,05	,04	,02	,08	,09	,08	,08	,02	,01	,01	,01
38	,08	,05	,03	,02	,10	,06	,04	,02	,09	,08	,08	,08	,02	,01	,01	,01
39	,51	,46	,43	,45	,09	,06	,03	,03	,05	,04	,03	,03	,39	,44	,47	,45
40	,06	,06	,03	,02	,10	,09	,09	,09	,10	,12	,11	,11	,03	,02	,02	,02
41	,02	,02	,01	,01	,04	,03	,02	,01	-	-	-	-	-	-	-	-
42	,05	,08	,10	,10	,05	,04	,04	,03	-	-	-	-	-	-	-	-
43	,02	,01	,01	,01	,04	,03	,01	,01	-	-	-	-	-	-	-	-
44	,03	,02	,01	,01	,05	,03	,02	,01	-	-	-	-	-	-	-	-
45	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
46	,02	,01	,01	,00	-	-	-	-	,02	,02	,01	,01	-	-	-	-
47	,02	,01	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
48	,02	,01	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
49	,02	,02	,01	,01	,05	,04	,02	,02	-	-	-	-	,03	,02	,01	,01
50	,03	,03	,03	,02	,05	,04	,03	,03	-	-	-	-	,03	,02	,02	,01
51	,46	,42	,41	,41	,12	,12	,12	,12	-	-	-	-	,03	,03	,01	,01
52	,07	,05	,03	,02	,06	,04	,02	,02	-	-	-	-	,03	,03	,03	,03
53	,90	,90	,90	,90	,09	,06	,05	,04	-	-	-	-	,01	,01	,01	,00
54	,03	,02	,01	,01	,05	,04	,02	,02	-	-	-	-	,02	,02	,01	,01
55	,03	,03	,01	,01	,08	,05	,03	,02	,04	,03	,03	,03	,07	,07	,07	,07
56	,06	,05	,03	,03	,11	,12	,11	,11	,08	,08	,08	,08	,19	,18	,19	,19
57	,07	,05	,03	,02	,08	,05	,03	,02	,09	,08	,08	,09	,02	,01	,01	,01
58	,08	,05	,03	,02	,08	,06	,03	,02	,09	,09	,08	,08	,02	,02	,01	,01
59	,50	,47	,44	,45	,09	,06	,03	,02	,05	,04	,03	,03	,40	,43	,46	,45
60	,07	,05	,03	,02	,10	,09	,09	,09	,11	,12	,12	,12	,03	,03	,02	,02
Ortalama Mutlak Sapma	,13	,12	,11	,10	,07	,06	,04	,04	,06	,05	,05	,05	,07	,07	,07	,07

8 Niteliği Ölçen Düşük Ayırt Ediciliğe Sahip 60 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,05	,03	,02	,01	,06	,03	,03	,02	-	-	-	-	-	-	-	-
2	,09	,12	,15	,16	,07	,06	,04	,04	-	-	-	-	-	-	-	-
3	,04	,02	,01	,01	,05	,03	,02	,01	-	-	-	-	-	-	-	-
4	,05	,03	,02	,01	,06	,04	,03	,02	-	-	-	-	-	-	-	-
5	,03	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
6	,03	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
7	,04	,02	,01	,01	-	-	-	-	,04	,03	,02	,01	-	-	-	-
8	,03	,02	,01	,01	-	-	-	-	,04	,03	,02	,01	-	-	-	-
9	,06	,04	,02	,02	,07	,05	,03	,02	-	-	-	-	,04	,03	,02	,01
10	,06	,07	,11	,13	,07	,07	,05	,05	-	-	-	-	,06	,05	,03	,02
11	,31	,26	,25	,26	,13	,13	,15	,15	-	-	-	-	,05	,04	,05	,06
12	,14	,12	,08	,08	,06	,04	,02	,02	-	-	-	-	,07	,07	,07	,07
13	,78	,78	,78	,78	,09	,07	,05	,05	-	-	-	-	,04	,03	,03	,03
14	,05	,04	,02	,01	,08	,04	,02	,02	-	-	-	-	,04	,03	,01	,01
15	,07	,05	,03	,02	,10	,07	,03	,03	,07	,04	,04	,04	,08	,07	,06	,06
16	,11	,09	,08	,08	,13	,12	,14	,14	,10	,09	,10	,10	,19	,18	,16	,16
17	,12	,09	,05	,04	,10	,06	,04	,02	,12	,12	,13	,12	,03	,03	,02	,02
18	,14	,10	,05	,04	,10	,07	,04	,03	,12	,12	,12	,11	,03	,03	,03	,02
19	,34	,24	,11	,04	,11	,09	,05	,04	,10	,06	,06	,07	,41	,52	,65	,73
20	,16	,11	,07	,05	,09	,10	,09	,09	,15	,16	,17	,17	,05	,04	,04	,04
21	,04	,04	,02	,01	,05	,04	,02	,02	-	-	-	-	-	-	-	-
22	,09	,12	,15	,16	,07	,05	,05	,04	-	-	-	-	-	-	-	-
23	,04	,02	,01	,01	,05	,03	,02	,01	-	-	-	-	-	-	-	-
24	,05	,03	,02	,02	,06	,04	,03	,02	-	-	-	-	-	-	-	-
25	,03	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
26	,03	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
27	,03	,02	,01	,01	-	-	-	-	,04	,03	,02	,01	-	-	-	-
28	,04	,03	,02	,01	-	-	-	-	,04	,03	,02	,01	-	-	-	-
29	,05	,03	,02	,02	,07	,04	,03	,02	-	-	-	-	,05	,03	,02	,01
30	,06	,07	,11	,13	,07	,06	,06	,05	-	-	-	-	,06	,05	,03	,03
31	,30	,26	,25	,25	,13	,13	,15	,15	-	-	-	-	,05	,04	,05	,07
32	,15	,12	,09	,08	,06	,04	,03	,02	-	-	-	-	,07	,07	,07	,07
33	,78	,78	,78	,78	,09	,08	,05	,05	-	-	-	-	,03	,03	,03	,03
34	,06	,04	,02	,01	,07	,05	,03	,02	-	-	-	-	,04	,02	,01	,01
35	,07	,04	,03	,02	,09	,07	,04	,03	,05	,04	,04	,03	,07	,07	,06	,07
36	,13	,09	,08	,08	,11	,11	,14	,14	,09	,10	,10	,10	,21	,19	,16	,16
37	,14	,09	,05	,03	,10	,07	,04	,03	,12	,12	,13	,12	,03	,03	,02	,02
38	,13	,12	,06	,05	,08	,07	,04	,03	,12	,12	,12	,12	,03	,03	,03	,03
39	,34	,25	,12	,04	,10	,09	,05	,04	,09	,07	,06	,06	,41	,51	,65	,73
40	,14	,10	,07	,05	,10	,10	,09	,09	,15	,16	,17	,17	,05	,05	,04	,04
41	,04	,03	,02	,01	,05	,04	,02	,02	-	-	-	-	-	-	-	-
42	,09	,12	,15	,16	,06	,05	,04	,05	-	-	-	-	-	-	-	-
43	,04	,02	,01	,01	,05	,03	,02	,01	-	-	-	-	-	-	-	-
44	,05	,03	,02	,02	,07	,04	,03	,02	-	-	-	-	-	-	-	-
45	,03	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
46	,03	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
47	,03	,02	,02	,01	-	-	-	-	,04	,03	,01	,01	-	-	-	-
48	,03	,02	,01	,01	-	-	-	-	,04	,03	,02	,01	-	-	-	-
49	,05	,04	,02	,01	,08	,05	,03	,02	-	-	-	-	,04	,03	,02	,01
50	,06	,07	,11	,13	,07	,06	,05	,05	-	-	-	-	,06	,05	,03	,03
51	,31	,26	,25	,25	,12	,14	,15	,15	-	-	-	-	,06	,04	,05	,07
52	,14	,12	,08	,09	,06	,04	,03	,02	-	-	-	-	,07	,07	,07	,07
53	,78	,78	,78	,78	,10	,07	,04	,05	-	-	-	-	,03	,03	,03	,03
54	,06	,04	,02	,02	,07	,05	,02	,02	-	-	-	-	,04	,03	,01	,01
55	,07	,05	,03	,02	,10	,06	,04	,02	,05	,04	,03	,04	,08	,07	,07	,06
56	,13	,10	,08	,08	,12	,12	,14	,14	,09	,09	,09	,10	,19	,18	,16	,16
57	,15	,11	,06	,04	,09	,06	,04	,03	,12	,12	,13	,13	,03	,03	,02	,02
58	,14	,09	,06	,05	,09	,08	,03	,03	,12	,12	,12	,11	,03	,03	,03	,03
59	,35	,25	,11	,04	,13	,09	,05	,04	,09	,06	,06	,06	,40	,52	,65	,73
60	,14	,11	,06	,05	,10	,09	,09	,09	,16	,17	,17	,17	,05	,04	,04	,04
Ortalama Mutlak Sapma	,14	,12	,10	,09	,08	,07	,05	,05	,08	,07	,07	,06	,09	,09	,09	,09

10 Niteliği Ölçen Yüksek Ayırt Ediciliğe Sahip 20 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,15	,15	,15	,15	,37	,39	,40	,39	-	-	-	-	-	-	-	-
2	,15	,15	,15	,15	,39	,40	,40	,40	-	-	-	-	-	-	-	-
3	,14	,15	,15	,15	,25	,33	,38	,40	-	-	-	-	-	-	-	-
4	,15	,15	,15	,15	,40	,40	,40	,40	-	-	-	-	-	-	-	-
5	,15	,15	,15	,15	,36	,37	,39	,38	-	-	-	-	-	-	-	-
6	,15	,15	,12	,04	-	-	-	-	,18	,19	,19	,16	-	-	-	-
7	,15	,15	,15	,08	-	-	-	-	,20	,20	,20	,17	-	-	-	-
8	,15	,15	,15	,10	-	-	-	-	,20	,20	,20	,19	-	-	-	-
9	,14	,13	,07	,02	-	-	-	-	,17	,15	,10	,07	-	-	-	-
10	,14	,14	,10	,05	-	-	-	-	,16	,15	,10	,09	-	-	-	-
11	,60	,67	,67	,64	,18	,18	,16	,12	-	-	-	-	,12	,12	,14	,18
12	,48	,49	,49	,48	,13	,12	,12	,12	-	-	-	-	,18	,18	,18	,18
13	,77	,77	,77	,78	,21	,19	,14	,10	,12	,12	,12	,13	,02	,02	,02	,03
14	,79	,78	,78	,79	,24	,24	,23	,21	-	-	-	-	,04	,03	,03	,04
15	,76	,76	,77	,77	,22	,23	,22	,19	-	-	-	-	,02	,02	,02	,02
16	,25	,26	,28	,28	,13	,09	,05	,04	,06	,04	,05	,06	,20	,22	,27	,28
17	,77	,77	,77	,78	,22	,17	,14	,09	,12	,12	,12	,13	,03	,03	,02	,03
18	,24	,20	,16	,13	,10	,09	,08	,09	,05	,03	,03	,03	,09	,08	,07	,07
19	,17	,19	,15	,14	,15	,10	,08	,09	,06	,03	,03	,04	,09	,09	,07	,06
20	,29	,42	,41	,34	,16	,12	,09	,07	,15	,14	,14	,15	,02	,02	,03	,04
Ortalama Mutlak Sapma	,30	,31	,31	,30	,20	,18	,18	,16	,10	,09	,09	,08	,07	,08	,08	,08

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı.

10 Niteliği Ölçen Düşük Ayırt Ediciliğe Sahip 20 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,05	,05	,05	,05	,33	,33	,35	,34	-	-	-	-	-	-	-	-
2	,05	,05	,05	,05	,35	,35	,35	,35	-	-	-	-	-	-	-	-
3	,05	,05	,05	,04	,22	,26	,28	,20	-	-	-	-	-	-	-	-
4	,05	,05	,05	,05	,35	,35	,35	,34	-	-	-	-	-	-	-	-
5	,05	,05	,05	,05	,25	,26	,22	,11	-	-	-	-	-	-	-	-
6	,05	,05	,04	,03	-	-	-	-	,13	,14	,15	,13	-	-	-	-
7	,05	,05	,05	,03	-	-	-	-	,15	,15	,14	,13	-	-	-	-
8	,05	,05	,05	,04	-	-	-	-	,15	,15	,15	,14	-	-	-	-
9	,04	,03	,02	,01	-	-	-	-	,11	,09	,08	,06	-	-	-	-
10	,04	,04	,03	,02	-	-	-	-	,09	,09	,07	,05	-	-	-	-
11	,76	,78	,80	,78	,13	,13	,13	,11	-	-	-	-	,17	,17	,17	,19
12	,53	,54	,55	,53	,09	,08	,09	,09	-	-	-	-	,21	,22	,21	,21
13	,90	,90	,90	,90	,14	,12	,12	,10	,10	,10	,10	,10	,01	,01	,01	,01
14	,90	,90	,90	,91	,20	,20	,21	,19	-	-	-	-	,01	,01	,01	,01
15	,89	,89	,89	,89	,19	,19	,19	,17	-	-	-	-	,02	,01	,01	,01
16	,16	,18	,18	,12	,11	,09	,08	,09	,06	,05	,05	,05	,20	,22	,22	,21
17	,90	,90	,90	,90	,14	,13	,11	,09	,10	,10	,10	,10	,01	,01	,01	,01
18	,09	,11	,09	,04	,10	,07	,04	,03	,04	,04	,03	,03	,08	,07	,07	,07
19	,09	,10	,12	,09	,11	,08	,04	,05	,04	,03	,03	,03	,07	,07	,07	,07
20	,24	,33	,42	,38	,12	,10	,07	,07	,12	,12	,12	,13	,02	,02	,01	,02
<i>Ortalama Mutlak Sapma</i>	,33	,34	,33	,33	,23	,23	,22	,21	,13	,13	,13	,11	,08	,08	,09	,09

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.

10 Niteliği Ölçen Yüksek Ayırt Ediciliğe Sahip 40 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,13	,13	,11	,04	,17	,17	,14	,05	-	-	-	-	-	-	-	-
2	,14	,14	,15	,08	,25	,33	,33	,18	-	-	-	-	-	-	-	-
3	,09	,08	,08	,11	,08	,09	,07	,08	-	-	-	-	-	-	-	-
4	,14	,13	,04	,03	,29	,25	,04	,03	-	-	-	-	-	-	-	-
5	,11	,11	,07	,03	,14	,11	,07	,02	-	-	-	-	-	-	-	-
6	,09	,06	,03	,03	-	-	-	-	,07	,06	,03	,02	-	-	-	-
7	,10	,07	,03	,03	-	-	-	-	,09	,08	,04	,03	-	-	-	-
8	,09	,06	,03	,03	-	-	-	-	,10	,06	,04	,02	-	-	-	-
9	,06	,04	,02	,02	-	-	-	-	,07	,05	,02	,02	-	-	-	-
10	,08	,05	,02	,01	-	-	-	-	,06	,05	,02	,02	-	-	-	-
11	,48	,57	,62	,67	,14	,11	,08	,07	-	-	-	-	,18	,21	,22	,23
12	,44	,40	,25	,17	,15	,16	,14	,13	-	-	-	-	,15	,14	,16	,17
13	,78	,78	,78	,77	,16	,12	,06	,07	,13	,13	,13	,12	,03	,03	,03	,02
14	,80	,79	,80	,79	,16	,15	,13	,13	-	-	-	-	,05	,04	,05	,04
15	,78	,78	,78	,78	,16	,13	,12	,11	-	-	-	-	,03	,03	,03	,03
16	,11	,11	,07	,06	,11	,10	,12	,16	,08	,06	,07	,06	,21	,20	,18	,14
17	,78	,78	,78	,77	,17	,12	,06	,05	,13	,13	,13	,12	,03	,03	,03	,02
18	,09	,07	,03	,03	,11	,07	,04	,03	,06	,05	,03	,03	,07	,06	,07	,07
19	,11	,07	,06	,10	,12	,10	,09	,13	,07	,05	,04	,03	,07	,07	,08	,08
20	,18	,13	,12	,11	,13	,10	,05	,08	,17	,17	,18	,19	,04	,04	,04	,03
21	,13	,13	,11	,04	,18	,21	,17	,06	-	-	-	-	-	-	-	-
22	,14	,14	,14	,09	,11	,11	,14	,15	-	-	-	-	-	-	-	-
23	,09	,08	,07	,11	,14	,11	,07	,09	-	-	-	-	-	-	-	-
24	,13	,12	,04	,03	,11	,09	,05	,03	-	-	-	-	-	-	-	-
25	,11	,10	,07	,02	,12	,10	,06	,03	-	-	-	-	-	-	-	-
26	,08	,06	,03	,03	-	-	-	-	,09	,08	,03	,03	-	-	-	-
27	,09	,08	,03	,03	-	-	-	-	,09	,08	,03	,03	-	-	-	-
28	,08	,06	,03	,03	-	-	-	-	,08	,07	,03	,02	-	-	-	-
29	,06	,04	,02	,02	-	-	-	-	,07	,04	,02	,02	-	-	-	-
30	,07	,05	,02	,02	-	-	-	-	,09	,05	,02	,01	-	-	-	-
31	,53	,57	,61	,69	,12	,10	,08	,05	-	-	-	-	,20	,20	,22	,25
32	,44	,39	,25	,17	,16	,16	,13	,12	-	-	-	-	,15	,14	,17	,18
33	,78	,78	,78	,77	,16	,11	,08	,06	,13	,13	,13	,12	,03	,03	,03	,02
34	,80	,80	,80	,80	,16	,15	,14	,12	-	-	-	-	,05	,05	,05	,05
35	,78	,78	,78	,78	,15	,13	,12	,11	-	-	-	-	,03	,03	,03	,03
36	,13	,10	,06	,07	,12	,10	,13	,15	,10	,07	,07	,06	,19	,20	,17	,15
37	,78	,78	,78	,77	,15	,11	,06	,05	,13	,13	,13	,12	,03	,03	,03	,02
38	,09	,06	,04	,02	,11	,07	,05	,03	,05	,05	,04	,03	,07	,06	,06	,07
39	,09	,07	,07	,10	,12	,08	,08	,12	,06	,05	,04	,03	,07	,07	,08	,09
40	,19	,14	,13	,10	,15	,10	,06	,08	,16	,17	,18	,18	,04	,04	,04	,03
Ortalama Mutlak Sapma	,28	,27	,24	,23	,10	,09	,08	,08	,06	,06	,05	,05	,08	,08	,08	,08

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı.

10 Niteliği Ölçen Düşük Ayırt Ediciliğe Sahip 40 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,05	,05	,06	,05	,07	,08	,08	,07	-	-	-	-	-	-	-	-
2	,05	,05	,08	,09	,07	,10	,20	,17	-	-	-	-	-	-	-	-
3	,04	,04	,06	,09	,06	,06	,06	,08	-	-	-	-	-	-	-	-
4	,05	,05	,04	,02	,15	,09	,04	,02	-	-	-	-	-	-	-	-
5	,04	,04	,03	,02	,05	,04	,03	,02	-	-	-	-	-	-	-	-
6	,04	,03	,02	,02	-	-	-	-	,03	,03	,02	,02	-	-	-	-
7	,04	,04	,03	,02	-	-	-	-	,04	,04	,03	,01	-	-	-	-
8	,04	,04	,02	,02	-	-	-	-	,06	,04	,02	,02	-	-	-	-
9	,02	,02	,01	,01	-	-	-	-	,03	,02	,02	,01	-	-	-	-
10	,03	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
11	,73	,78	,80	,83	,11	,10	,08	,06	-	-	-	-	,20	,20	,22	,24
12	,52	,49	,43	,38	,12	,12	,11	,11	-	-	-	-	,19	,18	,19	,19
13	,91	,91	,91	,90	,15	,10	,08	,06	,11	,11	,11	,10	,01	,01	,01	,01
14	,91	,91	,91	,91	,16	,17	,15	,14	-	-	-	-	,02	,01	,01	,01
15	,90	,89	,90	,89	,16	,15	,14	,13	-	-	-	-	,01	,01	,01	,01
16	,06	,05	,03	,03	,11	,11	,11	,12	,06	,05	,05	,05	,19	,19	,19	,18
17	,91	,90	,90	,90	,13	,09	,07	,04	,11	,10	,10	,10	,01	,01	,01	,00
18	,04	,03	,02	,01	,09	,06	,04	,02	,04	,04	,03	,03	,07	,07	,07	,07
19	,04	,03	,03	,04	,08	,06	,05	,05	,04	,04	,03	,03	,07	,07	,07	,08
20	,09	,08	,07	,05	,10	,07	,04	,05	,13	,13	,14	,13	,02	,02	,02	,02
21	,05	,06	,06	,05	,08	,05	,08	,07	-	-	-	-	-	-	-	-
22	,06	,07	,11	,10	,06	,05	,11	,16	-	-	-	-	-	-	-	-
23	,04	,04	,05	,09	,07	,06	,06	,07	-	-	-	-	-	-	-	-
24	,08	,06	,04	,02	,07	,03	,04	,02	-	-	-	-	-	-	-	-
25	,04	,04	,03	,01	,05	,04	,03	,02	-	-	-	-	-	-	-	-
26	,03	,03	,02	,02	-	-	-	-	,06	,03	,03	,02	-	-	-	-
27	,04	,04	,03	,02	-	-	-	-	,04	,04	,03	,02	-	-	-	-
28	,05	,04	,02	,02	-	-	-	-	,05	,04	,02	,02	-	-	-	-
29	,02	,02	,01	,01	-	-	-	-	,04	,03	,02	,01	-	-	-	-
30	,03	,02	,01	,01	-	-	-	-	,04	,03	,02	,01	-	-	-	-
31	,76	,79	,81	,82	,10	,09	,08	,06	-	-	-	-	,21	,21	,22	,24
32	,52	,49	,43	,38	,11	,11	,11	,11	-	-	-	-	,19	,19	,19	,19
33	,91	,90	,90	,90	,14	,10	,08	,06	,11	,10	,10	,10	,01	,01	,01	,00
34	,91	,91	,91	,91	,16	,17	,15	,14	-	-	-	-	,02	,01	,01	,01
35	,90	,90	,90	,90	,16	,15	,14	,13	-	-	-	-	,01	,01	,01	,01
36	,05	,04	,03	,03	,11	,11	,11	,12	,06	,05	,06	,05	,20	,19	,19	,18
37	,91	,90	,90	,90	,13	,11	,07	,05	,11	,10	,10	,10	,01	,01	,01	,00
38	,04	,03	,02	,01	,08	,06	,03	,02	,04	,04	,04	,03	,07	,07	,06	,07
39	,05	,03	,03	,04	,08	,06	,04	,05	,04	,04	,03	,03	,08	,07	,07	,07
40	,10	,08	,06	,06	,10	,07	,04	,04	,13	,13	,13	,13	,02	,02	,02	,02
Ortalama Mutlak Sapma	,28	,27	,27	,26	,15	,13	,10	,09	,09	,08	,07	,06	,09	,09	,09	,09

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.

10 Niteliği Ölçen Yüksek Ayırt Ediciliğe Sahip 60 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_j (Mutlak Sapma)				g_j (Mutlak Sapma)				ω_j (Mutlak Sapma)				ϵ_j (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,08	,07	,05	,05	,09	,09	,04	,05	-	-	-	-	-	-	-	-
2	,10	,09	,06	,07	,13	,10	,09	,09	-	-	-	-	-	-	-	-
3	,06	,06	,10	,12	,07	,07	,05	,06	-	-	-	-	-	-	-	-
4	,07	,06	,05	,04	,11	,07	,08	,03	-	-	-	-	-	-	-	-
5	,07	,06	,01	,01	,07	,06	,04	,01	-	-	-	-	-	-	-	-
6	,03	,03	,02	,02	-	-	-	-	,05	,03	,02	,01	-	-	-	-
7	,04	,02	,02	,01	-	-	-	-	,05	,03	,01	,02	-	-	-	-
8	,04	,03	,02	,02	-	-	-	-	,05	,03	,02	,02	-	-	-	-
9	,03	,03	,02	,01	-	-	-	-	,05	,02	,02	,01	-	-	-	-
10	,04	,02	,01	,01	-	-	-	-	,04	,02	,02	,01	-	-	-	-
11	,44	,57	,66	,67	,12	,09	,07	,04	-	-	-	-	,19	,21	,23	,26
12	,39	,33	,24	,16	,15	,13	,07	,11	-	-	-	-	,16	,17	,23	,19
13	,78	,78	,77	,77	,15	,09	,06	,04	,13	,13	,12	,12	,03	,03	,02	,02
14	,80	,80	,79	,79	,14	,15	,16	,12	-	-	-	-	,05	,05	,04	,04
15	,79	,79	,78	,77	,14	,12	,12	,11	-	-	-	-	,04	,04	,03	,02
16	,11	,07	,05	,05	,11	,11	,13	,14	,08	,07	,05	,06	,20	,19	,17	,16
17	,78	,78	,76	,77	,13	,10	,06	,06	,13	,13	,11	,12	,04	,03	,01	,02
18	,07	,04	,02	,02	,09	,06	,02	,03	,06	,04	,03	,03	,06	,07	,07	,07
19	,08	,05	,05	,08	,08	,07	,08	,08	,06	,05	,04	,03	,07	,07	,08	,08
20	,14	,12	,11	,10	,10	,07	,04	,04	,17	,18	,18	,18	,04	,03	,03	,03
21	,09	,07	,04	,04	,10	,09	,06	,05	-	-	-	-	-	-	-	-
22	,09	,08	,06	,08	,12	,09	,08	,09	-	-	-	-	-	-	-	-
23	,06	,07	,08	,11	,07	,07	,06	,06	-	-	-	-	-	-	-	-
24	,07	,05	,06	,04	,08	,07	,06	,03	-	-	-	-	-	-	-	-
25	,08	,06	,03	,01	,07	,06	,02	,02	-	-	-	-	-	-	-	-
26	,04	,03	,03	,02	-	-	-	-	,04	,03	,02	,01	-	-	-	-
27	,04	,03	,01	,01	-	-	-	-	,04	,03	,03	,01	-	-	-	-
28	,04	,03	,02	,02	-	-	-	-	,04	,04	,02	,02	-	-	-	-
29	,04	,02	,01	,01	-	-	-	-	,05	,03	,02	,01	-	-	-	-
30	,03	,02	,01	,01	-	-	-	-	,04	,03	,01	,01	-	-	-	-
31	,53	,59	,68	,68	,12	,10	,06	,05	-	-	-	-	,20	,20	,24	,25
32	,39	,34	,25	,17	,15	,13	,10	,11	-	-	-	-	,16	,17	,20	,19
33	,79	,78	,77	,77	,13	,11	,06	,04	,14	,13	,12	,12	,04	,03	,02	,02
34	,80	,80	,79	,79	,14	,14	,13	,13	-	-	-	-	,05	,05	,04	,04
35	,78	,78	,78	,78	,13	,12	,14	,10	-	-	-	-	,03	,03	,03	,03
36	,10	,06	,04	,05	,11	,12	,13	,14	,08	,07	,07	,06	,20	,18	,17	,16
37	,78	,78	,77	,77	,15	,08	,07	,04	,13	,13	,12	,12	,03	,03	,02	,02
38	,07	,05	,03	,02	,09	,05	,03	,03	,06	,04	,03	,03	,07	,07	,07	,07
39	,07	,05	,06	,07	,10	,08	,04	,08	,06	,05	,04	,03	,08	,07	,07	,07
40	,17	,14	,07	,08	,12	,06	,06	,05	,16	,17	,19	,18	,04	,04	,03	,03
41	,08	,06	,04	,04	,10	,09	,05	,06	-	-	-	-	-	-	-	-
42	,08	,07	,05	,07	,10	,12	,08	,09	-	-	-	-	-	-	-	-
43	,06	,06	,09	,11	,08	,06	,05	,06	-	-	-	-	-	-	-	-
44	,06	,06	,07	,05	,10	,06	,04	,03	-	-	-	-	-	-	-	-
45	,06	,05	,02	,02	,06	,05	,03	,01	-	-	-	-	-	-	-	-
46	,04	,03	,02	,02	-	-	-	-	,04	,03	,01	,01	-	-	-	-
47	,04	,03	,02	,02	-	-	-	-	,04	,03	,02	,01	-	-	-	-
48	,04	,03	,02	,02	-	-	-	-	,04	,04	,02	,02	-	-	-	-
49	,03	,03	,01	,01	-	-	-	-	,03	,03	,01	,01	-	-	-	-
50	,03	,03	,02	,01	-	-	-	-	,04	,02	,02	,01	-	-	-	-
51	,49	,56	,68	,68	,12	,10	,06	,05	-	-	-	-	,20	,21	,26	,25
52	,39	,34	,25	,16	,15	,13	,09	,11	-	-	-	-	,16	,17	,21	,19
53	,79	,78	,77	,77	,15	,12	,06	,05	,14	,13	,12	,12	,04	,03	,02	,02
54	,80	,80	,79	,79	,15	,14	,16	,13	-	-	-	-	,05	,05	,04	,04
55	,78	,79	,77	,77	,13	,11	,14	,10	-	-	-	-	,03	,04	,02	,02
56	,10	,08	,05	,06	,12	,12	,13	,14	,08	,07	,06	,06	,19	,18	,17	,16
57	,79	,78	,76	,77	,14	,09	,05	,04	,14	,13	,11	,12	,04	,03	,01	,02
58	,07	,05	,04	,02	,08	,06	,05	,03	,05	,05	,03	,03	,08	,07	,07	,07
59	,07	,06	,06	,08	,09	,06	,05	,07	,06	,05	,04	,03	,08	,06	,07	,08
60	,15	,13	,06	,07	,11	,07	,05	,06	,17	,18	,18	,18	,04	,04	,04	,03
Ortalama Mutlak Sapma	,25	,24	,23	,23	,09	,08	,07	,06	,06	,05	,05	,04	,08	,08	,08	,08

10 Niteliği Ölçen Düşük Ayırt Ediciliğe Sahip 60 Maddede Farklı Örneklem Büyüklüğü Koşullarında Madde Parametrelerine Ait Tahmini “Mutlak Sapma” ve “Ortalama Mutlak Sapma” Değerleri

Madde	h_i (Mutlak Sapma)				g_i (Mutlak Sapma)				ω_i (Mutlak Sapma)				ϵ_i (Mutlak Sapma)			
	Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü			
	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000
1	,04	,04	,02	,01	,06	,05	,03	,02	-	-	-	-	-	-	-	-
2	,05	,04	,03	,03	,08	,06	,05	,03	-	-	-	-	-	-	-	-
3	,03	,03	,04	,07	,04	,04	,04	,05	-	-	-	-	-	-	-	-
4	,04	,03	,02	,01	,07	,04	,02	,03	-	-	-	-	-	-	-	-
5	,04	,03	,01	,01	,05	,03	,02	,01	-	-	-	-	-	-	-	-
6	,02	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
7	,02	,02	,01	,01	-	-	-	-	,03	,03	,01	,01	-	-	-	-
8	,02	,02	,01	,01	-	-	-	-	,03	,02	,01	,00	-	-	-	-
9	,02	,02	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
10	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,00	-	-	-	-
11	,74	,79	,81	,84	,12	,10	,09	,07	-	-	-	-	,19	,20	,21	,23
12	,50	,48	,45	,42	,11	,11	,11	,12	-	-	-	-	,20	,19	,19	,18
13	,91	,90	,90	,90	,15	,12	,07	,07	,11	,10	,10	,10	,01	,01	,01	,00
14	,91	,91	,91	,91	,16	,15	,15	,15	-	-	-	-	,01	,01	,01	,01
15	,90	,90	,90	,90	,14	,15	,14	,13	-	-	-	-	,01	,01	,01	,00
16	,05	,03	,02	,02	,11	,10	,10	,10	,06	,04	,05	,05	,19	,20	,20	,20
17	,91	,90	,90	,90	,14	,11	,08	,03	,11	,10	,10	,10	,01	,01	,01	,00
18	,04	,02	,02	,01	,06	,04	,03	,01	,04	,03	,03	,02	,07	,07	,07	,08
19	,04	,03	,02	,02	,08	,06	,03	,02	,04	,03	,03	,03	,07	,07	,07	,09
20	,08	,06	,04	,04	,08	,07	,03	,01	,13	,13	,13	,14	,02	,02	,02	,01
21	,04	,04	,02	,02	,07	,05	,03	,02	-	-	-	-	-	-	-	-
22	,05	,03	,03	,03	,07	,05	,04	,02	-	-	-	-	-	-	-	-
23	,03	,03	,04	,07	,05	,05	,04	,04	-	-	-	-	-	-	-	-
24	,04	,03	,02	,01	,07	,05	,03	,02	-	-	-	-	-	-	-	-
25	,03	,03	,01	,01	,05	,03	,02	,02	-	-	-	-	-	-	-	-
26	,02	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
27	,02	,02	,01	,00	-	-	-	-	,03	,02	,01	,02	-	-	-	-
28	,02	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
29	,02	,01	,01	,01	-	-	-	-	,03	,02	,01	,00	-	-	-	-
30	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,01	-	-	-	-
31	,77	,79	,81	,80	,10	,09	,08	,08	-	-	-	-	,21	,21	,22	,22
32	,51	,47	,45	,43	,11	,10	,11	,10	-	-	-	-	,19	,20	,19	,20
33	,91	,91	,90	,90	,13	,11	,08	,09	,11	,11	,10	,10	,01	,01	,01	,00
34	,91	,91	,91	,91	,15	,15	,15	,16	-	-	-	-	,01	,01	,01	,01
35	,90	,90	,90	,90	,15	,15	,14	,12	-	-	-	-	,01	,01	,01	,00
36	,05	,04	,02	,03	,10	,10	,10	,10	,06	,05	,05	,06	,20	,20	,20	,20
37	,91	,90	,90	,91	,13	,10	,08	,03	,11	,10	,10	,11	,02	,01	,01	,01
38	,04	,03	,01	,01	,07	,05	,03	,03	,04	,04	,03	,03	,06	,06	,07	,07
39	,04	,03	,02	,02	,08	,06	,03	,04	,04	,03	,03	,02	,07	,07	,07	,08
40	,08	,06	,03	,01	,08	,06	,04	,02	,13	,13	,13	,13	,02	,02	,02	,02
41	,05	,04	,02	,01	,06	,04	,03	,02	-	-	-	-	-	-	-	-
42	,05	,04	,03	,02	,07	,06	,04	,03	-	-	-	-	-	-	-	-
43	,03	,03	,04	,07	,05	,05	,04	,05	-	-	-	-	-	-	-	-
44	,04	,03	,02	,01	,08	,04	,02	,01	-	-	-	-	-	-	-	-
45	,04	,03	,01	,01	,04	,04	,02	,01	-	-	-	-	-	-	-	-
46	,02	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
47	,02	,02	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
48	,02	,02	,01	,01	-	-	-	-	,03	,02	,01	,02	-	-	-	-
49	,02	,01	,01	,01	-	-	-	-	,03	,02	,01	,01	-	-	-	-
50	,02	,01	,01	,00	-	-	-	-	,03	,02	,01	,00	-	-	-	-
51	,74	,80	,82	,82	,11	,10	,08	,08	-	-	-	-	,20	,21	,22	,22
52	,50	,47	,45	,41	,11	,10	,11	,10	-	-	-	-	,19	,20	,19	,20
53	,91	,91	,91	,90	,13	,11	,08	,10	,11	,11	,11	,10	,01	,01	,01	,01
54	,91	,91	,91	,91	,16	,15	,16	,16	-	-	-	-	,01	,01	,01	,01
55	,90	,90	,89	,90	,15	,16	,14	,15	-	-	-	-	,01	,01	,01	,00
56	,05	,04	,02	,02	,12	,10	,11	,11	,06	,05	,05	,06	,19	,20	,19	,19
57	,91	,90	,90	,91	,13	,11	,08	,05	,11	,10	,10	,11	,01	,01	,01	,01
58	,04	,03	,01	,01	,07	,05	,03	,01	,04	,04	,03	,02	,06	,06	,07	,08
59	,04	,03	,02	,01	,07	,05	,03	,04	,04	,04	,03	,03	,07	,07	,07	,07
60	,08	,06	,04	,01	,09	,06	,03	,02	,13	,13	,13	,14	,02	,02	,02	,02
Ortalama Mutlak Sapma	,27	,26	,26	,25	,11	,09	,07	,07	,08	,07	,06	,05	,09	,09	,09	,09

EK 5. 6, 8 ve 10 Nitelikli Q-matrislerin Kullanıldığı Durumlarda Dört Madde Parametresine ait “Hataların Ortalama Karekökü” Değerleri

Tablo 29

Faktörler için Değişimlenen Her Bir Koşulda Dört Madde Parametresine ait “Hataların Ortalama Karekökü” Değerleri

		h_j				g_j				ω_j				ϵ_j				
		Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				Örneklem Büyüklüğü				
		250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	250	500	1500	3000	
Madde Ayırt Edicilik Düzeyi	Nitelik Sayısı																	
20 Madde	Y	6	,22	,20	,18	,18	,22	,20	,18	,18	,21	,20	,18	,18	,22	,20	,18	,18
	D	6	,29	,29	,28	,28	,29	,29	,28	,28	,29	,29	,28	,28	,29	,29	,28	,28
	Y	8	,30	,28	,25	,23	,30	,28	,25	,23	,29	,28	,25	,23	,30	,28	,25	,23
	D	8	,30	,28	,25	,23	,30	,28	,25	,24	,30	,28	,25	,24	,30	,28	,25	,24
	Y	10	,42	,43	,42	,42	,42	,43	,42	,41	,42	,43	,42	,40	,42	,43	,42	,40
	D	10	,47	,47	,47	,47	,47	,46	,47	,46	,47	,46	,47	,46	,47	,46	,47	,46
40 Madde	Y	6	,20	,18	,17	,17	,20	,18	,17	,17	,20	,18	,17	,17	,20	,18	,17	,17
	D	6	,25	,27	,27	,26	,25	,26	,26	,26	,24	,26	,26	,26	,25	,26	,26	,26
	Y	8	,24	,22	,20	,19	,23	,22	,20	,19	,23	,22	,20	,19	,23	,22	,20	,20
	D	8	,26	,25	,24	,24	,26	,25	,24	,24	,26	,25	,24	,24	,26	,25	,24	,24
	Y	10	,40	,40	,39	,39	,40	,39	,38	,38	,40	,39	,38	,38	,40	,39	,39	,38
	D	10	,46	,46	,46	,45	,45	,45	,45	,45	,45	,45	,45	,45	,45	,45	,45	,45
60 Madde	Y	6	,19	,18	,17	,17	,19	,18	,17	,17	,19	,18	,17	,17	,19	,18	,17	,17
	D	6	,22	,17	,15	,14	,22	,17	,15	,14	,22	,17	,15	,14	,22	,17	,15	,14
	Y	8	,23	,21	,20	,19	,23	,21	,20	,19	,23	,21	,20	,19	,23	,21	,20	,19
	D	8	,26	,25	,24	,24	,26	,25	,24	,24	,26	,25	,24	,24	,26	,25	,24	,24
	Y	10	,39	,39	,39	,38	,39	,38	,36	,37	,39	,38	,36	,37	,39	,38	,36	,37
	D	10	,46	,46	,45	,45	,45	,45	,45	,36	,45	,45	,45	,36	,45	,45	,45	,37

Y: Madde ayırt edicilik düzeyi yüksek, D: Madde ayırt edicilik düzeyi düşük, h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.

EK 6. 6, 8 ve 10 Nitelikli Q-matrisler için Faktörlerde Değişimlenen Her Bir Koşulda Dört Parametreye Ait Tüm Maddeler Arasındaki Ortalama Standart Hata Değerleri

Tablo 30

Dört Parametreye Ait Tüm Maddeler Arasındaki Ortalama Standart Hata Değerleri

	Örneklem Büyüklüğü	Madde Sayısı	Madde Ayırt Edicilik Düzeyi	h_j	g_j	ω_j	ϵ_j
6 Nitelik	250	20	<i>Yüksek</i>	,09	,11	,06	,04
			<i>Düşük</i>	,11	,12	,07	,04
	500	20	<i>Yüksek</i>	,05	,06	,03	,03
			<i>Düşük</i>	,05	,07	,04	,03
	1500	20	<i>Yüksek</i>	,02	,04	,02	,01
			<i>Düşük</i>	,03	,04	,02	,02
	3000	20	<i>Yüksek</i>	,02	,03	,01	,01
			<i>Düşük</i>	,02	,03	,01	,01
	250	40	<i>Yüksek</i>	,23	,13	,07	,05
			<i>Düşük</i>	,23	,17	,06	,05
	500	40	<i>Yüksek</i>	,08	,06	,03	,03
			<i>Düşük</i>	,07	,07	,03	,03
	1500	40	<i>Yüksek</i>	,02	,03	,01	,01
			<i>Düşük</i>	,02	,04	,02	,02
	3000	40	<i>Yüksek</i>	,01	,02	,01	,01
			<i>Düşük</i>	,02	,03	,01	,01
	250	60	<i>Yüksek</i>	,34	,21	,09	,06
			<i>Düşük</i>	,42	,25	,13	,07
	500	60	<i>Yüksek</i>	,06	,07	,03	,03
			<i>Düşük</i>	,14	,07	,03	,03
1500	60	<i>Yüksek</i>	,02	,03	,01	,01	
		<i>Düşük</i>	,02	,04	,02	,02	
3000	60	<i>Yüksek</i>	,01	,02	,01	,01	
		<i>Düşük</i>	,02	,03	,02	,01	

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.

Tablo 31

Dört Parametreye Ait Tüm Maddeler Arasındaki Ortalama Standart Hata Değerleri

	Örneklem Büyüklüğü	Madde Sayısı	Madde Ayırt Edicilik Düzeyi	h_j	g_j	ω_j	ϵ_j
8 Nitelik	250	20	<i>Yüksek</i>	,09	,08	,05	,10
			<i>Düşük</i>	,10	,09	,06	,10
	500	20	<i>Yüksek</i>	,05	,05	,03	,06
			<i>Düşük</i>	,06	,06	,04	,07
	1500	20	<i>Yüksek</i>	,02	,03	,02	,04
			<i>Düşük</i>	,04	,04	,03	,05
	3000	20	<i>Yüksek</i>	,02	,02	,01	,02
			<i>Düşük</i>	,03	,03	,02	,04
	250	40	<i>Yüksek</i>	,15	,09	,07	,20
			<i>Düşük</i>	,50	,08	,06	,24
	500	40	<i>Yüksek</i>	,06	,05	,03	,06
			<i>Düşük</i>	,06	,05	,04	,10
	1500	40	<i>Yüksek</i>	,02	,02	,02	,02
			<i>Düşük</i>	,03	,03	,03	,04
	3000	40	<i>Yüksek</i>	,01	,02	,01	,01
			<i>Düşük</i>	,02	,02	,02	,02
250	60	<i>Yüksek</i>	,27	,12	,09	,34	
		<i>Düşük</i>	,62	,14	,12	,35	
500	60	<i>Yüksek</i>	,06	,05	,03	,08	
		<i>Düşük</i>	,16	,06	,05	,19	
1500	60	<i>Yüksek</i>	,02	,02	,02	,02	
		<i>Düşük</i>	,03	,03	,02	,07	
3000	60	<i>Yüksek</i>	,01	,02	,01	,01	
		<i>Düşük</i>	,02	,02	,02	,03	

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanılgıları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanılgıları var olan bireylerin başarı olasılığı.

Tablo 32

Dört Parametreye Ait Tüm Maddeler Arasındaki Ortalama Standart Hata Değerleri

	Örnekleme Büyüklüğü	Madde Sayısı	Madde Ayırt Edicilik Düzeyi	h_j	g_j	ω_j	ϵ_j
10 Nitelik	250	20	<i>Yüksek</i>	,11	,09	,06	,05
			<i>Düşük</i>	,12	,11	,06	,05
	500	20	<i>Yüksek</i>	,06	,06	,04	,03
			<i>Düşük</i>	,07	,07	,04	,03
	1500	20	<i>Yüksek</i>	,03	,04	,02	,02
			<i>Düşük</i>	,04	,04	,03	,02
	3000	20	<i>Yüksek</i>	,02	,03	,02	,01
			<i>Düşük</i>	,03	,03	,02	,02
	250	40	<i>Yüksek</i>	,57	,11	,09	,13
			<i>Düşük</i>	,57	,11	,08	,07
	500	40	<i>Yüksek</i>	,13	,06	,04	,03
			<i>Düşük</i>	,07	,07	,04	,04
	1500	40	<i>Yüksek</i>	,03	,03	,02	,01
			<i>Düşük</i>	,03	,04	,02	,02
	3000	40	<i>Yüksek</i>	,02	,02	,01	,01
			<i>Düşük</i>	,02	,03	,01	,02
	250	60	<i>Yüksek</i>	,68	,16	,18	,20
			<i>Düşük</i>	,85	,18	,23	,11
	500	60	<i>Yüksek</i>	,09	,06	,05	,03
			<i>Düşük</i>	,33	,07	,04	,04
1500	60	<i>Yüksek</i>	,03	,03	,02	,01	
		<i>Düşük</i>	,03	,04	,02	,02	
3000	60	<i>Yüksek</i>	,01	,02	,01	,01	
		<i>Düşük</i>	,02	,02	,01	,01	

h_j : Ölçülen tüm becerilere sahip olan ve kavram yanlışları olmayan bireylerin başarı olasılığı, g_j : Ölçülen tüm becerilere sahip olmayan ve kavram yanlışları olmayan bireylerin başarı olasılığı, ω_j : Ölçülen tüm becerilere sahip olan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı, ϵ_j : Ölçülen tüm becerilere sahip olmayan ve bazı kavram yanlışları var olan bireylerin başarı olasılığı.



GAZİLİ OLMAK AYRICALIKTIR..