

**ÇOKTAN SEÇMELİ TESTLERDE SEÇENEK SAYISININ MADDE
PARAMETRELERİ VE FARKLI YETENEK DÜZEYLERİNDE BİLGİ
FONKSİYONLARI İLE MADDE UYUMLULUKLARI ÜZERİNDE
ETKİSİNİN İNCELENMESİ**

Gholamreza GOLMOHAMMAD NAZHAD BAHRAMI

Hacettepe Üniversitesi
Sosyal Bilimler Enstitüsü

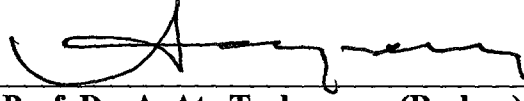
Eğitim Bilimler Anabilim Dalı Eğitimde Ölçme ve Değerlendirme Bilim
Dalı

DOKTORA TEZİ

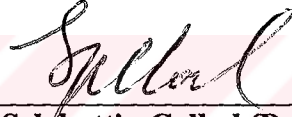
Ankara
Ekim, 2006

KABUL VE ONAY

Gholamreza Gol. Mohammed Nazhad Bahrami tarafından hazırlanan Çoktan Seçmeli Tetslerde Seçenek Sayısının Madde Parametreleri ve Farklı Yetenek Düzeylerinde Bilgi Fonksyonları İle Madde Uyumlulukları Üzerinde Etkisinin İncelenmesi” başlıklı bu çalışma, 13.10.2006 tarihinde yapılan savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından Doktora tezi olarak kabul edilmiştir.



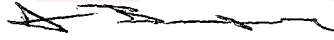
Prof. Dr. A. Ata Tezbaşaran (Başkan)



Doç. Dr. Selahattin Gelbal (Danışman)



Doç. Dr. Aysun Umay



Doç. Dr. Şener Büyüköztürk



Yrd. Doç. Dr. Hülya Kelecioğlu

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylarım.



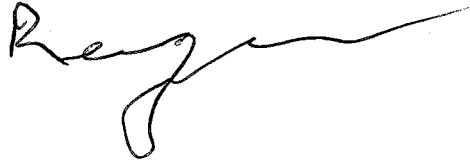
BİLDİRİM

Hazırladığım tezin tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder, tezimin/raporumun kâğıt ve elektronik kopyalarının Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım:

- ☒ Tezimin/Raporumun tamamı her yerden erişime açılabilir.
- ☐ Tezim/Raporum sadece Hacettepe Üniversitesi yerleşkelerinden erişime açılabilir.
- ☐ Tezimin/Raporumun.Yıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığı m takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir.

13.10.2006

Gholamreza GOLMOHAMMAD NAZHAD BAHRAMI



TEŞEKKÜR

Araştırmamın her aşamasında görüş, öneri ve desteklerini esirgemeyen değerli hocam ve danışmanım Doç. Dr. Selahattin GELBAL'a sonsuz teşekkür ederim. Çalışmalarım süresince bana görüş ve önerileri ile destek olan Prof. Dr. Ata TEZBAŞARAN'a ve Yrd. Doç. Dr. Hülya KELECİOĞLU'na, Yrd. Doç. Dr. Mohammad Ali FARNIA'ya teşekkürlerimi sunarım.

Araştırmama katkılarından dolayı sevgili arkadaşlarım Araş. Gör. Neşe GÜLER ve Araş. Gör. Emre ÇETİN'e teşekkür ederim. Bugünlere gelmemi sağlayan bana her zaman güç veren, bana hem anne hem baba olan sevgili anneciğime ve kalbimdeki yerini hiç kaybetmeyen rahmetli babama teşekkürlerimi bir borç bilirim.

Gholamraza GOLMOHAMMAD NAZHAD BARAMI

Ekim, 2006

ÖZET

GOLMOHAMMAD NAZHAD BAHRAMI, Gholamreza. Çoktan Seçmeli Testlerde Seçenek Sayısının Madde Parametreleri ve Farklı Yetenek Düzeylerinde Bilgi Fonksiyonları ile Madde Uyumlulukları Üzerinde Etkisinin incelenmesi, Doktora Tezi, Ankara, 2006

Eğitimde ve psikolojide öğretim görevlileri, danışmanlar ve uzmanlar her zaman öğrencilere eğitim ve öğretim vermekte, onlara rehberlik etmekte, bazen öğrencileri sınıflandırmakta ve öğrenciler arasında seçim yapmakta veya çeşitli araştırmalar yaparak önemli kararlar almaktadırlar. Bu faaliyetlerle ilgili değişik testlerden yararlanarak bireyler hakkında bilgi toplamaktadırlar. Örneğin; zeka testleri, yetenek ve başarı testleri vb.

Güvenilir kararlar almak için doğru bilgilere ulaşmak şarttır ve bu nedenle doğru ve güvenilir testler geliştirmek son derece önem taşır.

Bu çalışmada, başarı testlerinde seçenek sayının madde parametreleri üzerindeki etkisi en doğru ve güvenilir test formuna ulaşmak amacıyla incelenmektedir. Madde ayırt edicilik gücü, madde güçlük ve şans parametreleri dört seçenekli ve altı seçenekli testler arasında karşılaştırılmıştır. Ayrıca iki test arasında madde bilgi fonksiyonlarının değişik başarı düzeylerindeki değişimleri göz önüne alınarak incelenmiştir.

Bu araştırmada, 985 öğrenci İRAN'ın Tabriz Azad Üniversitesi'nden örneklem alınarak R. B. Cattell'in zeka testinin üçüncü ölçeğinin A formu altı seçenekli test halinde ve B formu dört seçenekli test halinde (A ve B formları, paralel formlardır) uygulanmış ve bireylerin zeka ölçüleri ve düzeyleri ile madde parametreleri, bilgi fonksiyonları ve madde uyumluluğu, IRT'de üç ve iki parametrelili modeller kullanılarak elde edilmiştir. Bu test, eğitilmiş insanlara ve üniversite öğrencilerine özeldir. A ve B formu halinde olan testlerin her birisi üç alt testten oluşmuştur ve her testin toplam 36'sar maddesi vardır.

Birinci alt problem olan “seçenek sayısı arttığında ayırtedicilik gücünde (a_j) ne tür değişiklikler ortaya çıkmaktadır?” sorusunun incelenmesinde ortaya çıkan ilgili sonuç, iki testin ayırt edicilik güçlerinin değişmediği yönündedir.

İkinci alt problemde, iki testin güçlük derecesi karşılaştırılmış ve elde edilen sonuçlar A testinin güçlük derecesinin daha da yüksek olduğunu göstermiştir. Bu sonuç, gayet normal bir

sonuç olarak değerlendirilebilir. Başarı testlerinde seçenek sayısı artığında sorular zorlaşmaktadır.

Üçüncü alt problemde, değişik yetenek düzeylerinde, bireylerin testlerdeki yanlış cevap sayılarının eşit olup olmadığını incelenmiştir. Bu probleme ilişkin elde edilen sonuç, yetenek düzeyleri arasında yanlış cevap sayılarının eşit olmadığı yönündedir.

Dördüncü alt problemin incelenmesi sonucunda; seçenek sayısı değiştiğinde yanlış cevap sayılarının, farklı yetenek düzeylerinde eşit olmadığı görülmektedir.

Beşinci alt problemde, madde bilgi fonksiyonunun seçenek sayısından etkilenip etkilenmediği, ayrıca bu etkinin farklı yetenek düzeylerinde eşit olup olmadığı incelenmiştir. Ortaya çıkan sonuçlara göre, seçenek sayısının fazlasıyla artırılması ve bu artışın bireylerin başarı düzeylerini aşması durumunda, madde bilgi fonksiyonlarının düştüğü gözlenmiştir. Bu alt problemle ilgili diğer bir sonuç ise, maddelerin seçenek sayısı artırıldığında maddelerin zorlaşması nedeniyle daha yukarı düzeylerde, maddelerin ayırt edicilik gücü artmış ve bu da yüksek başarı düzeylerinde madde bilgi fonksiyonlarının yükselişine sebep olmuştur. Bu durum, düşük başarı düzeylerinde tersine işlemektedir. Maddedeki seçenek sayısı azaldığında da, madde bilgi fonksiyonu düşük düzeylerde artmakta ve yüksek düzeylerde düşmektedir.

Altıncı alt problemde, madde sayısındaki değişimin madde uyumluluğu üzerinde ne gibi bir etkisi olduğu incelenmiştir. Buradan elde edilen sonuç; maddedeki seçenek sayısının değiştirilmesinin, madde uyumluluğu üzerinde (IRT 3 ve 2 parametrelili modellerde) anlamlı bir etkisinin olmadığı yönündedir.

Anahtar Sözcükler

Ayırt Edicilik Gücü, Çoktan Seçmeli Test, Madde Bilgi Fonksiyonu, Madde Güçlük Düzeyi, Madde Parametreleri, Madde Uyumluluğu, Örtük Özellikler Kuramı, Seçenek Sayısı, Şans.

ABSTRACT

GOLMOHAMMAD NAZHAD BAHRAMI, Gholamreza. *A Study on The Effect of The Numbers of Options On The Parameters And The Question's Goodness of Fit, And Its Effect on Information Function of The Question in Different Levels of Ability in Multiple Choice Tests*. Ph.D. Dissertation, Ankara, 2006.

In behavioral science experts are always in search of such objectives as classification of individuals, discrimination of interferences, selection and evaluation. Investigators also look for studies which increase the human knowledge about the humanistic problems and make suggestion about the application of these findings. In order to deal with these sort of problems and obtaining the objectives, collection of valid evidence, reliable information and careful measurement of individual's mental and behavioral characteristics is required.

The aim of the present study is to investigate the effect of the number of option in multiple choice items which is used as the measurement and information collection instrument, so that the best form which is more accurate and suitable among the others can be determined.

The present study compares the discrimination parameter (a_j), difficulty parameter (b_j), chance among and the goodness of fit items in four and six options. It also compares and investigates the information function in different ability levels.

The statistical population was 985 Tabriz Azad University students in Iran from different fields of study which have been selected by random sampling. Parallel form (A & B) R.B. Cattell intelligence test was implemented without chronological interval. In the study A from was six option and B from was four option items, where each of them included 36 items.

The research questions and their results are as follow:

The first question deals with the difference of (a_j) between four and six option items, the result showed no difference in discrimination parameter between them, i.e. there is no effect based on the number of options in discrimination parameter item.

The second question explores the difference of the difficulty degrees of items between the two tests. The results show that the level of difficulty in items with six options is higher than that of four options, and it seems quite logical.

The third question, studies the difference of the number of wrong answers between the different levels of ability and the results show that there is significance different between various levels of ability.

The fourth question examines whether the use of the number of wrong answers the different levels of ability, will change equally. The result showed that with the change in the number of item options, there will also be a change in the number of wrong answers, but this change among the different levels of ability is not equal.

The fifth question was to investigate whether the change in the number of options will lead to the change in information function items, and if this change is equal in different levels of ability. The result showed that with the change in the number of options, the information function items will also change and the more the number of the options the more difficult the items will become, consequently in the higher level, the information function will be higher and it will be lower in lower levels of ability.

The sixth question was to investigate the effect of the number of option in the goodness of fit items (i.e. the fit of Item to the measurement model). The result showed no considerable effect of the number of options on the goodness of fit item. This investigation was implemented both in two parametric and three parametric IRT model.

Key words

Chance, Choice, Difficulty Parameter, Discrimination Parameter, Goodness of Fit, Guess, Information Function, Item, Item Response Theory (IRT), Option .

İÇİNDEKİLER

	Sayfa
TEŞEKKÜR.....	i
ÖZET.....	ii
ABSTRACT	iv
İÇİNDEKİLER	vi
TABLolar LİSTESİ	ix
ŞEKİLLER LİSTESİ.....	xi
 BÖLÜM I	
GİRİŞ	1
Klasik Test Kuramı Modeli	2
Klasik Test Kuramının Varsayımları	2
Klasik Test Kuramında Test ve Madde İstatistikleri.....	3
Klasik Test Kuramını Sınırlılıkları.....	5
Klasik Test Kuramının Avantajları	7
Madde Tepki Kuramı Modeli	8
Örtük Özellikler Kuramının Varsayımları	10
Örtük Özellik Kuramında Kullanılan Modeller	13
Örtük Özellikler Kuramının Avantajları	19
Örtük Özellikler Kuramının Sınırlılıkları	20
Çoktan Seçmeli ve Klasik Testler.....	22
Problem cümlesi.....	23
Alt problemler	23
Araştırmanın Önemi ve Amacı	24
Varsayımlar	25

Sınırlılıklar	25
İlgili Araştırmalar	26

BÖLÜM II

YÖNTEM	30
Araştırmanın Türü	30
Evren ve Örneklem	30
Araç ve Verilerin Toplanması	30
Veri Çözümleme Yöntemleri	32
Testlerin Puan, Parametre ve Bilgi Fonksiyonlarının Betimsel Bilgileri.....	34
Ölçütlerin Tek Boyutluluklarıyla İlgili Olan Analizler ve Sonuçları	41
A ve B Test Puanlarının Normal Dağılıp Dağılmadığı Testi	44

BÖLÜM III

BULGULAR VE YORUM.....	46
1. Alt Probleme İlişkin Bulgular ve Yorumlar	46
2. Alt Probleme İlişkin Bulgular ve Yorumlar	47
3. Alt Probleme İlişkin Bulgular ve Yorumlar	49
4. Alt Probleme İlişkin Bulgular ve Yorumlar	53
5. Alt Probleme İlişkin Bulgular ve Yorumlar	57
6. Alt Probleme İlişkin Bulgular ve Yorumlar	66

BÖLÜM IV

SONUÇ VE ÖNERİLER	70
Sonuçlar	70
Öneriler	71

KAYNAKÇA	72
EKLER	78
EK 1 B (4 SEÇENEKLİ) TEST FORMU	79
EK 2 A VE B TESLERİNE AİT MADDE PARAMETRELERİ	84
EK 3 A VE B TESTLERİN MADDE ÖZELLİKLER GRAFİKLERİ	85
EK 4 AVE B TESTLERİN MADDE BİLGİ FONKSİYONLARI	103



TABLOLAR LİSTESİ

	Sayfa
Tablo 1 A ve B Testlerine Ait Olan İstatistiksel Ölçüler	34
Tablo 2 A ve B Testlerinin Madde Parametreleri	37
Tablo 3 A ve B Testlerinin Maddelerine Ait olan Maksimum Bilgi Fonksiyonu, Maksimum Bilgi Fonksiyona Sahip Olan Nokta ve Ortalama Bilgi Fonksiyonu	39
Tablo 4 A ve B Testine Ait Faktör Analiz Sonuçları	41
Tablo 5 A ve B testlerinin puan dağılımı için Kolmogorov-Smirnov testi	44
Tablo 6 A Test (6seçenekli) ve B Testine (4 seçenekli) Ait Olan Ayırt Edici Güçlerinin Betimsel İstatistikleri ve Wilcoxon Testi İle Karşılaştırılma Sonuçları	46
Tablo 7 A Test (6seçenekli) ve B Testine (4 Seçenekli) Ait Olan Madde Güçlük Derecesinin Betimsel İstatistikleri ve Wilcoxon Test İle Karşılaştırılma Sonuçları	47
Tablo 8 A ve B Testine Ait Olan Yetenek Düzeylerinin Ortalamasına Göre 8 Yetenek Grubuna Yerleşen Bireylerin İki Teste Olan Yanlış Cevap Sayılarının Ortalamasının İstatistiği	49
Tablo 9 Bireylerin A Test ve B Testlerindeki Yanlış Cevaplarının Ortalamaları İle İki Testteki Yetenek Düzeylerinin Ortalamaları Arasındaki Korelasyon Katsayısı	50
Tablo 10 A ve B Testlerinde Yanlış Cevap Sayılarının Ortalamaları Arasındaki Farkın, Değişik 8 Yetenek Grupta ANOVA ile İncelenmiş Sonuçları ...	50
Tablo 11 Yetenek Grupları Arasında, A ve B Testlerinde Ortalama Yanlış Cevap Sayılarının Ortalama Farkı	51
Tablo 12 A Test (6seçenekli) ve B Testine (4 seçenekli) Ait Olan Yanlış Cevap Sayısının Farkı ve Yetenek Düzeylerine Göre Bu Farkın Nasıl Olduğu Varyans Analiz İle İncelenmiş Sonuçları	53
Tablo 13 Yetenek Grupları Arasında, A Testindeki (6 seçenekli test) ve B Testindeki Yanlış Cevap Farkının (A Test Yanlışları – B Test Yanlışları) Değişik Yetenek Gruplarında İncelenme Sonuçları	54
Tablo 14 A Ve B Testlerinin Madde Bilgi Fonksiyonlarının Ortalamasının T-Testi İle İncelenmiş Sonuçları	57

Tablo 15	Yetenek Düzeyler Arasında Madde Bilgi Fonksiyonları Farklı Olup Olmamasının Her İki Teste (Ayrı Ayrı) Tek Yönlü Varyans Analiz İle Yapılmış İncelemenin Sonuçları	58
Tablo 16	A Test (6 Seçenekli) ve B Testine (4 Seçenekli) Ait Olan Madde Bilgi Fonksiyon Farkı ve $\theta = 12$ 'ten $\theta = -12$ 'ye Olan Yetenek Düzeylerine Göre Bu Farkın Nasıl Olduğu Varyans Analiz İle İncelenmiş Sonuçları	58
Tablo 17	A Test (6seçenekli) ve B Testine (4 seçenekli) Ait Olan Madde ilgi Fonksiyon Farkı ve $\theta = -3$ 'ten $\theta = -12$ 'ye ve $\theta = 3$ 'ten $\theta = 12$ 'ye Olan Yetenek Düzeylerine Göre Bu Farkın Nasıl Olduğu Varyans Analiz İle İncelenmiş Sonuçları	59
Tablo 18	A ve B Testlerinin Madde Ayırt Edicilik Güçleri İle Güçlük Derecelerinin Arasındaki Korelasyon Katsayısı	63
Tablo 19	A ve B Testlerin Uyumlu Madde Sayısını Üç ve İki Parametrelili IRT Modelde Göstermektedir. A ve B Testleri Toplam 36 Şer Maddeye Sahiptiler	66
Tablo 20	Tekrarlı Datalara Ait Olan İki Yönlü Varyans Analizini Kullanarak A ve B Testlerinin Uyumlu Madde Sayısının Üç ve İki Parametrelili IRT Modelde Hem Testler Hem Modeller Arasında Fark, Hem de Test ve Modellerin Karşılıklı Etkisi İncelenmiş Sonuçları	67
Tablo 21	A (6 seçenekli) Test ve B (4 seçenekli) Testin, 3 ve 2 Parametrelili Modellerde Belirlenen Uyumlu Madde Sayılarının Farkı Ki-kare Testi İle İncelenmiş İstatistiksel Sonuçlarını Göstermektedir	68
Tablo 22	A (6 seçenekli) Testi ve B (4 seçenekli) Testinin 3 ve 2 Parametrelili Modellerde Belirlenen Uyumlu Madde Sayılarının Farkını McNemar Testi İle İncelenmiş İstatistiksel Sonuçlarını Göstermektedir	68

ŞEKİLLER LİSTESİ

		Sayfa
Şekil 1	IRT Modelinde Bir Maddenin Bilgi Fonksyonu, Ayırt Ediciliği, Güçlük Derecesi ve Şans Başarısını Gösteren Bir Grafik	9
Şekil 2	A Testine Ait Puan Grafiği	35
Şekil 3	B Testine Ait Puan Grafiği	36
Şekil 4	A Testine Ait Özdeğer-Bileşen Grafiği	42
Şekil 5	B Testine Ait Özdeğer-Bileşen Grafiği	43
Şekil 6	A Testine Ait Normallik P-P Grafiği	45
Şekil 7	B Testine Ait Normallik P-P Grafiği	46
Şekil 8	Bireylerin İki Testte Olan Yanlış Cevapları Sayısının Yetenek Düzeylerine Göre Yansıtmaktadır	55
Şekil 9	A ve B Testlerinin Madde Bilgi Fonksiyonlarının Ortalamasını Yetenek Düzeylerine Göre Yansıtmaması	60
Şekil 10	A Testine Ait Başarı Düzeyleri Üzerinde Bilgi Fonksiyon Grafiği	61
Şekil 11	B Testine Ait Başarı Düzeyleri Üzerinde Bilgi Fonksiyon Grafiği	61
Şekil 12	Teorik Olarak Eşit Ayırt Edicilik Gücüne Sahip Olan ve Değişik Güçlük Derecelerini Yansıtan Maddeleri Gösterir. Burada Madde Ayırt Edicilik Gücü İle Güçlük Derecesi Arasında Hiç Bir İlişkinin Olmadığı Söz Konusudur	63
Şekil 13	Teorik Olarak Değişik Ayırt Edicilik Gücü ve Değişik Güçlük Derecelerine Sahip Olan Maddeleri Gösterir. Burada Madde Ayırt Edicilik Gücü İle Güçlük Derecesi Arasında Bir İlişkinin Olduğu Söz Konusudur	64

BÖLÜM I

GİRİŞ

Her bilim dalında ve her araştırmada, belli özellikleri ölçmeye yönelik çeşitli araçlar kullanılır. Gerek sosyal bilimlerde gerekse eğitim bilimlerinde, çeşitli ölçme araçları yardımıyla bilgi toplayarak araştırma sorularının çözümünde ve eğitimsel kararların alınmasında önemli bir yere sahiptir. Ölçme araçları 20'nci yüzyılın başında ilk defa Alfred Binet tarafından öğrencilerin yeteneklerini ölçme amacıyla geliştirilmiştir (Sharifi,1991). O zamandan şimdiye kadar farklı bilim dallarında insan özelliklerini ölçmeye yönelik çok değişik testler geliştirilmiştir.

İlk olarak, başarı ve zeka testleri, sonra da kişilik, tutum ve görüş gibi özellikleri ölçme amacıyla araçlar geliştirilmiştir. Günümüzde değişik ölçme araçları eğitim kurumlarında, rehberlik merkezlerinde, hastanelerde vb. birçok birimde değişik amaçlar için kullanılmaktadır.

Güvenirliği ve geçerliği düşük testler ve bu testleri oluşturan maddeler yukarıda belirtilen amaçları gerçekleştirmede yetersiz kalır. İstenilen amaçlara uygun ölçme araçları yaratmak için test geliştirme süreçlerinin uygulanması faydalı olabilir. Bu yüzden test ve maddeler geliştirilme aşamasında güvenilirlik ve geçerlilik gibi özellikleri açısından incelenmelidir.

Bilim adamları ilk testlerin gelişimi ile beraber bu testlerin güvenilirlik ve geçerliklerini inceleme amacıyla çeşitli yöntemler ve teknikler bulup ortaya koymuşlardır. Genelde bu yöntemler ve teknikler iki yaklaşımda yer alırlar. Bu yaklaşımlardan birisi daha eski yıllardan gelen “Klasik Test Kuramı” ve diğeri ise daha sonra geliştirilen “Örtük Özellikler Kuramı”dır. Bu kuramlar aşağıda sırasıyla özetlenerek verilmiştir.

Klasik Test Kuramı Modeli

Bu modele göre elde edilen ölçüm veya gözlenen puan; bir gerçek ve bir hata puanlarından oluşmuştur;

$$X_o = X_t + X_e$$

Burada X_o gözlenen puan, X_t gerçek puan ve X_e ise hata puanıdır. Hata puanı ne kadar düşürülürse ölçme güvenirliği o derecede artar. Bu denklem, klasik test kuramının temel denklemi kabul edilir. Bu denklemden yararlanarak, bazı varsayımlarla ulaşılabılır (Gulliksen 1950; Ramsay 1971).

Klasik Test Kuramının Varsayımları

Klasik test kuramı bazı varsayımlara dayandırılmıştır. Bunlar;

- 1- Hata puanlarının beklenen değerinin sıfır olması,
- 2- Gerçek puanlarla hata puanları arasındaki korelasyonun sıfır olması,
- 3- İki ayrı ölçmeye ait hata puanları arasındaki korelasyonun sıfır olması,
- 4- İki ayrı ölçmeye ait hata ve gerçek puanlar arasındaki korelasyonun sıfır olması şeklinde özetlenebilir (Gulliksen 1950; Lord ve Novic 1968:36; Thorndike 1982:144-145; Baykul 2000:114; Owens 2000).

Klasik test kuramında, yukarıdakilere ek olarak paralellik (eşdeğerlik) kavramı ve bileşik ölçmeler için de bazı varsayımlar bulunmaktadır (Gulliksen 1950; Baykul 2000).

Klasik Test Kuramında Test ve Madde İstatistikleri

Klasik test kuramına göre, test istatistikleri; testin güçlüğü, ortalaması, varyansı, güvenilirliği ve geçerliği, madde istatistikleri kullanılarak kestirilebilir. Testin ortalaması ya da testin ortalama güçlüğü madde güçlük indekslerine göre seçilecek maddelerle istenilen bir değer olacak şekilde ayarlanabilir. Yine madde ayırt edicilik güçleri ve varyansları yüksek maddeleri seçmek yoluyla testin güvenilirliği ve geçerliği yüksek tutulabilir. Ayrıca madde sayısını belli oranda fazlalaştırmak yoluyla da güvenilirliği artırmak mümkündür (Gulliksen 1950; Crocker ve Algina 1986; Baykul 2000).

Klasik test kuramında test ve madde istatistiklerinin hareket noktası madde güçlük düzeyi, madde ayırt edicilik toplam test puanlarıdır. Bu istatistikler yardımıyla teste ait özellikler kestirilebilir.

Madde güçlük düzeyi, bir maddeye doğru cevap verenlerin o maddeyi cevaplayanların tümüne (toplam cevaplayıcı sayısına) oranı olarak tanımlanır ve genel olarak aşağıdaki formülle ifade edilebilir.

$$p_j = \frac{n_d}{N}$$

Bu denklemde p_j madde güçlük düzeyini n_d maddeye doğru cevap verenlerin sayısını, N ise j maddesine cevap verenlerin tümünü ifade etmektedir. Madde güçlük düzeyi 0 ile 1 arasında değerler alır. Bu değer 0,50 civarında ise madde orta güçlükte, büyüdükçe maddenin kolaylaştığı ve küçüldükçe maddenin zorlaştığı yorumu yapılabilir. Yukarıdaki denklem madde güçlük düzeyi için temel denklem olup, maddeye erişemeyenler, maddeyi atlayanlar veya cevap veremeyenler ile şans başarısını dikkate alarak da madde güçlük indeksini hesaplamak mümkündür. Ayrıca alt üst grup seçmek yoluyla madde güçlük indeksini hesaplama formülünün yaygın bir şekilde kullandığından söz edilebilir (Henrysson; 1971; Turgut 1971; Baykul 1979; Turgut

1982; Cunningham 1986; Osterlid 1989; Özçelik 1989, 1992; Tekin 1993; Aiken 2000; Baykul ve diğerleri 2001a).

Henrysson (1971), madde güçlük indeksi yerine bu değerlerin standart normal dağılım değerlerinin kullanılabileceğini belirtir. Bu dönüştürme de bir maddeye verilen cevapların yetenekle ilişkili olduğu, madde güçlük boyutunda maddeye verilen cevapların normal dağıldığı varsayılır. Böylece standart normal dağılım yardımıyla bir maddeyi doğru cevaplamak için gerekli yetenek düzeyi hakkında bilgi edinilebilir.

Madde ayırt edicilik gücü, kabaca bir maddenin o maddeyle ölçülmek istenen özelliğe sahip olanla olmayanı ayırabilmesi olarak tanımlanabilir. Madde puanları ile test puanlarının korelasyonu olarak ifade edilir. Bu değer, değişkenlerin ikisi de sürekli kabul edildiğinde Pearson Moment Çarpımları korelasyon (PMÇK) katsayısı ile; değişkenler süreksiz ve yapay süreksiz kabul edildiğinde PMÇK'nın özel halleri olan çift serili korelasyon katsayısı (r_{bis}) ve nokta-çift serili korelasyon katsayısı (r_{pbis}) denklemlerinden birisiyle elde edilmektedir. Madde ayırt edicilik gücü - 1 ile 1 arasında değerler alabilir.

Bu çalışmada, madde ayırt edicilik gücünün göstergesi olarak çift serili korelasyon katsayısı dikkate alınmıştır. Çift serili korelasyon katsayısı aşağıdaki denklemlerle gösterilebilir (Ancı 1990; Akhun 1994).

$$r_{bis} = \frac{\bar{x}_p - \bar{x}_q}{s_t} \cdot \frac{pq}{y}$$

$$r_{bis} = \frac{\bar{x}_p - \bar{x}_t}{s_t} \cdot \frac{p}{y}$$

r_{bis} : Çift serili korelasyon katsayısı

$\bar{x}_p$: Maddeye doğru cevap verenlerin test puanları ortalaması

$\bar{x}_q$: Maddeye yanlış cevap verenlerin test puanları ortalaması

$\bar{x}_t$: Testin ortalaması

S_t : Testin standart sapması

p : Maddeye doğru cevap verenlerin tüm cevaplayıcılara oranı

$$q = 1 - p$$

y : p oranına karşı gelen birim normal dağılımdaki ordinat değeri

Klasik Test Kuramını Sınırlılıkları

Yıllardır test geliştirmede, geliştirilmiş test ve testlerin analizinde ve bunların uygulanması sonucunda elde edilen puanların yorumlanmasında kullanılıyor olmasına rağmen, klasik test kuramının bir takım sınırlılıklara sahip olduğuna ilişkin bilgilere sık sık rastlanmaktadır.

Klasik test kuramının literatürde rastlanan sınırlılıkları maddeler halinde şöyle sıralanabilir.

1 - Klasik test kuramında madde ve test istatistikleri elde edildiği örnekleme bağımlıdır (sample dependency). Bir örneklemdaki yetenek düzeyi ve dağılımı, klasik test kuramının madde güçlük indeksi ve madde ayırt edicilik indeksi gibi madde istatistiklerini ve test istatistiklerini önemli ölçüde etkilemektedir. Bu nedenle bu istatistiklerin genellenmesi sıkıntılar yaratır; hesaplandıkları örnekleme benzeyen örneklemelere genellemede kullanışlı olabilirler.

2- Bireylerin yetenekleri kendilerine uygulanan maddelere bağımlıdır (test dependency). Bir birey, aynı özelliği ölçen farklı güçlükteki testlerden farklı puanlar

alır. Bireylerin karşılaştırılması aynı ya da paralel testlere tabi tutulduğunda mümkündür. Farklı güçlükteki testlere tabi tutulan bireyleri klasik test kuramıyla karşılaştırmak mümkün olmamaktadır.

3- Klasik kuramla geliştirilen testler genellikle orta yetenek düzeyine hitap eder. Bu da, özellikle alt ve üst düzeydeki yeteneklerin güvenilir şekilde belirlenmesine imkan vermemektedir. Alt ve üst yetenek gruplarına gidildikçe sağlanan bilginin yetersiz olması söz konusudur. Belirli bir yetenek düzeyine uygun güçlükteki maddelerin seçilmesi ile test güçlüğü'nün ayarlanması, o yetenek grubundaki bireyler için uygulanan testin geçerliğini artırsa da geniş bir yetenek dağılımına sahip grupta testin geçerliği düşmektedir.

4- Ölçme hatası varyansının tüm cevaplayıcılar için eşit olduğu varsayılır. Ölçme hatasının varyansının tüm cevaplayıcılar için eşit olması da bireyler hakkında yanlış kararlar almaya götürmektedir. Çünkü; hata yetenek düzeyine bağlı olarak artar veya azalabilir.

5- Klasik test kuramının çok fazla varsayımı vardır. Hepsinin karşılanması her zaman mümkün olmayabilir.

6- Testin güvenilirliği de örnekleme bağımlıdır. Güvenirlik kavramında sıkıntılar ve çelişkiler bulunmaktadır. Güvenirliğin bir testin tekrarlı uygulamalarından ya da paralel formunun uygulanmasından elde edilmesi sorun oluşturur. Bir testin tekrar uygulanması durumunda bireylerin ölçülen özelliklerinin değişmediği varsayılmaktadır. Gerçekte ise bireylerin ölçülen psikolojik özellikleri iki uygulama arasında yeni öğrenmeler, unutmalar, kaygı, motivasyon, hatırlama gibi bazı etkilerle değişir. En azından değişip değişmediği hakkında bir yorum yapılması güçtür. Dolayısıyla böyle elde edilen güvenilirlik, olsa olsa güvenilirlik için bir alt sınırdır. Daha iyi bir deyişle güvenilirlik değerinin hatalar içeren bir tahminle elde edilmesi söz konusudur.

7. Bir cevaplayıcının, bir test maddesine doğru cevap verme olasılığını

belirlemek zordur. Bir cevaplayıcının bir test maddesine doğru cevabı verme olasılığını bilememek, testin, kişinin yetenek düzeyine uygunluğu hakkında karar vermeyi zorlaştırmaktadır.

8. Bireyler arasındaki farklar sifıra yaklaştığında klasik kurama göre elde edilen istatistikler anlamlarını kaybederler. Çünkü karşılaştırmalar bireylerin birbirine göre konumlarının varyansına dayalı olarak yapılır (Baykal 1980b; Hambleton ve Swaminathan 1985; Somer 1998; Yu 1999; Harvey ve Hammer, 1999; Bayku! 2000).

Yukarıda belirtilmiş olan sınırlılıkların yanı sıra test içindeki yanlış maddelerin belirlenmesi, test puanlarının eşitlenmesi gibi sorunlara çözüm bulmanın klasik test kuramında istenilir düzeyde olmadığı belirtilmektedir. Madde istatistikleri, test puanları ölçeğinde, maddelerin maksimum ayırt edicilik gücü ve ayırdıkları yetenek ranjı hakkında bilgi vermemekte, madde yanlışlığını belirleme çalışmalarında ilgili gruplar arasındaki gerçek yetenek farklılıklarını uygun biçimde ele almada yetersiz kalmaktadır (Hambleton ve Swaminathan, 1985; Hambleton, Swaminathan ve Rogers, 1991; Fan 1998; Harvey ve Hammer, 1999).

Klasik Test Kuramının Avantajları

Açıklanan sınırlılıklarına rağmen, klasik kuram testi uygulama ve parametreleri kestirme kolaylığı yanında, üzerinde çok çalışılması nedeniyle diğer kuramlardan daha yaygın kullanılmaktadır. Klasik test kuramı, geliştirilen ilk modern ölçme kuramı olma özelliği taşımaktadır.

Klasik kuramın en büyük avantajının, varsayımlarının çeşitli test uygulamalarına uyarlanabilmesi olduğu belirtilmektedir (Hambleton ve Jones 1993).

Klasik test kuramının sınırlılıkları alternatif kuramların geliştirilmesi çalışmalarını hızlandırmıştır. Klasik test kuramının sınırlılıklarını aşma çalışmaları örtük özellikler kuramının gelişmesiyle sonuçlanmıştır.

Madde Tepki Kuramı Modeli

Psikometri uzmanları tarafından üretilen diğer model ise Madde Tepki Kuramı (IRT) modelidir. Bu modelde, klasik test kuramı modelindeki sınırlılıkların bulunmadığı iddiası vardır. IRT modeli, Örtük Özellikler Kuramına dayanmaktadır. Örtük özellik, tek boyutlu farzedilmiş bir temel özellik olarak, test tarafından ölçülmesi varsayılır. Örtük Özellikler Kuramı , temel bir özelliğin çeşitli düzeyleri ile bir madde arasındaki ilişkiyi açıklamaktadır. Başka bir deyişle bu kuram, bir bireyin bir maddeyi doğru cevaplama olasılığını, o bireyin temel özelliğindeki düzeyine bağlayan bir denklemi oluşturur. Bu denklem şöyledir:

$$P_j = f(\theta_1, \theta_2, \theta_3, \dots, \theta_k)$$

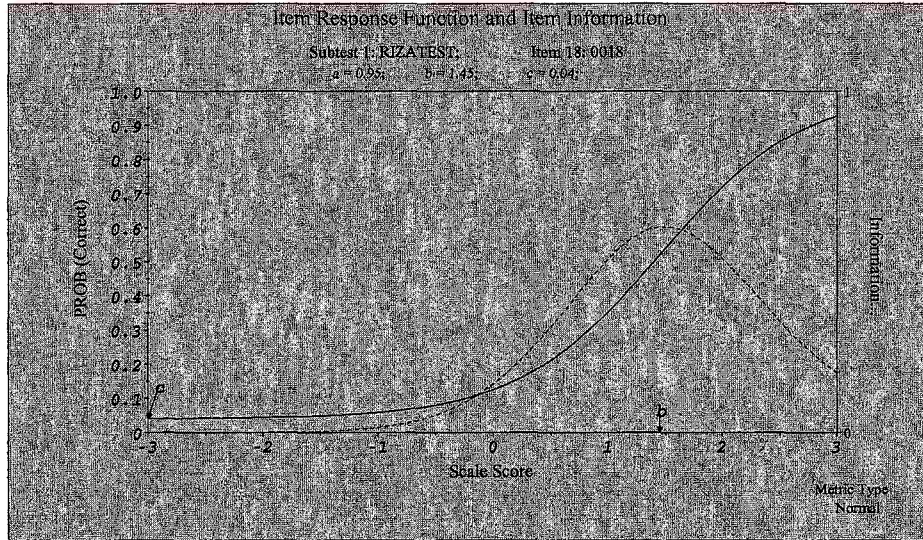
Bu denklemde; “ P_j ” j. maddedeki başarı olasılığını ve “ θ ” madde tarafından ölçülen temel özelliği göstermektedir. Bu nedenle, her bir maddenin bu temel özellikle olan bağlantısının incelenmesi önem taşır ve çoğunlukla bu inceleme Madde Karakteristik Eğrisi ile (Item Characteristic Curve- ICC) yapılır. Bu eğri, maddeyi doğru cevaplama olasılığını, test tarafından ölçülen örtük özelliğin düzeyine bağlı yani maddeyi doğru cevaplama olasılığını örtük özelliğin düzeyinin bir fonksiyonu olarak gösterir. Bu eğrinin oluşumunda ise bir kaç faktör rol oynar. Bu faktörlerden birisi maddenin ayırt etme gücüdür.

Ayrırt edicilik faktörü; bir maddenin, farklı düzeylerde olan bireyleri birbirinden ayırıp teşhis etme gücünü gösterir. Eğer bir maddenin doğru cevaplanma olasılığı, başarı düzeyi ile birlikte artıyorsa o madde ayırtedicilik gücüne sahiptir.Yani, alt yetenek düzeylerinde olan bireylerin o maddeyi doğru cevaplama olasılıkları yüksek düzeylerde olan bireylerden daha azdır. Eğer, bir maddenin doğru cevaplanma olasılığı başarı düzeyiyle birlikte hızla artıyorsa, o maddenin o oranda ayırtedicilik gücü yüksektir. Bir madde eğrisinin yükseliş veya düşüşü o maddenin ayırtedicilik gücünü göstermektedir ve bu yükseliş veya düşüş her ne kadar 45 dereceye yakın olursa o maddenin

ayırteıcılık gücü o derecede yüksektir. Maddenin bu özelliğine ayırt edicilik parametresi denir ki “ a_j ”le gösterilir.

Bir başka faktörse, madde güçlüğüdür. Bu faktör, maddenin cevaplanmasındaki zorluğu yansıtır. IRT modelinde bir madde güçlüğü bireylerin o maddeyi doğru cevaplama olasılığının 0.5 olduğu örtük özelliğın düzeyine (θ) eşittir. Madde güçlüğü de güçlük parametresi olarak tanımlanır ve “ b_j ” ile gösterilir. Diğer faktör ise şans faktörüdür ve maddenin cevabını bilmeden, şans kullanarak maddeyi doğru cevaplama olasılığı olarak tanımlanır. Başka bir deyişle en alt başarı düzeyinde bulunan bireylerin maddeyi doğru cevaplama olasılığıdır. Bu faktör şans parametresi olarak adlandırılır ve “ c_j ”ile gösterilir.

Aşağıda madde karakteristik eğrisine (ICC) bir örnek gösterilmiştir. Bu örnekte gösterilen maddenin madde ayırteıcılık parametresi 0.95, güçlük parametresi 1.45 ve şans parametresi ise 0.04 dır.



Şekil 1: IRT Modelinde Bir Maddenin Bilgi Fonksyonu, Ayırt Ediciliği, Güçlük Derecesi ve Şans Başarısını Gösteren Bir Grafik

Klasik test kuramı ve IRT modelinin arasındaki bir başka farklılık da güvenirlik konusudur. Klasik test kuramında bir madde güvenirliği; o maddenin tüm testle olan homojenliğinde bulunur ve tüm test güvenirliği de, o testin paralel testlerle olan korelasyonu anlamındadır. Bu modelde, test güvenirliği tüm bireyler için eşittir ve bundan dolayı ölçme hatası da tüm bireyler için hesaplanır ve bu da ölçme sonuçlarında hataya sebep olabilir.

IRT modellerinde madde güvenirliği, o maddenin bilgi fonksiyonuna bağlıdır ve farklı yetenek düzeylerinde (örtük özellik düzeylerinde) farklı olabilir. Eğer tüm maddelerin bilgi fonksiyonları bir yetenek düzeyi üzerinde toplanırsa o yetenek düzeyinde tüm test güvenirliği elde edilir ve bunun için ölçme hatası her yetenek düzeyinde farklı olabilir. Dolayısıyla ölçme sonuçları klasik test kuramına göre daha güvenilir olmaktadır. Yukarıdaki şekilde madde bilgi fonksiyonu eğrisi kesik noktalarla gösterilmiştir.

IRT modellerinin maddelere uyumlu olup olmamasına göre karar vermek, o maddenin o modele uymasına bağlıdır.

Örtük Özellikler Kuramının Varsayımları

Tek boyut/ufuk: Kuramın temel varsayımlardan birisi, belirli bir maddeye doğru cevap verme olasılığının, kişinin tek bir karakteristiği veya yetenek düzeyiyle belirlendiğidir. Geleneksel olarak örtük özellik modelleri, son yıllardaki çok boyutlu model geliştirme çalışmaları bir yana bırakılırsa tek boyutluluk varsayımına dayanır. Tek boyutluluk, maddelerin tek bir boyutu, bir değişkeni ölçmesi anlamındadır. Örtük özellik modelleri tipik olarak, belli bir boyut yapısına sahip olan araçlara uygulanır. Ancak tek boyutluluk pratikte doğrulanması zor bir varsayımdır. Bu nedenle, alt test içeren testlerin hazırlanması için, her bir alt test, tek boyutluluk açısından ayrı ayrı analiz edilebilir. Uygulamada tek boyutluluğu sağlamak mükemmel bir iş olacaktır. Ancak, tek boyutluluğu sağlamak gerçekten güçtür. Ayrıca, tek boyutluluk çalışmalarında, tek bir boyutun bulunması durumunda, bulunan boyutun ne olduğunu kestirmek, boyutu

tanımlamak her zaman kolay olmayabilir (Crocker ve Algina 1986).

Tek boyutluluk varsayımının bozulmasının etkilerini değerlendiren araştırmaların sonuçlarına göre, ölçme araçlarının tek boyutlu çıkmamasını varsayım bozulumu olarak kabul etmek gerekmez. Önemli olan, madde havuzunun tek örtük özellik (dominant bir faktör) tarafından belirlenmesidir. Genellikle tek boyutluluk çalışması faktör analizi ile yapılmaktadır. Bununla beraber yerel bağımsızlık çalışmasında kullanılan korelasyonlardan da faydalanmak mümkündür (Lord 1980; Hambleton ve Ravinelli 1986; Crocker ve Algina 1986; Harvey ve Hammer 1999).

Hambleton ve Swaminathan (1985) tek boyutluluğu şöyle açıklamaktadır; Örtük özellikler teorisinde, maddelerin “k” tane örtük özelliği ölçtüğü, “k” tane örtük özelliğin “k” boyutlu bir özellik uzayını tanımladığı ve genellikle kişilerin testteki performansını tanımlayan bir tek özellik boyutunun olduğu varsayılır. Her ne kadar test maddeleri ile ölçülmek istenen özellik bir tek özellik ve de tek bir boyut olsa da testteki maddelere yanıt verme davranışları ölçülmek istenilen özelliklerle sınırlı kalmayabilir. Testteki maddelere yanıt vermekteki isteklilik, endişe, sınav kaygısı, ölçülen özellik dışında kalan duygusal ya da bilişsel etkiler gibi değişkenlerde başka boyutlar olarak ortaya çıkabilir. Ancak, baskın olan faktör testle ölçülmek istenen özellik olarak kabul edilebilir.

Yerel Bağımsızlık: Bu varsayım, örtük özellik uzayının tek boyutlu olmasının bir sonucudur. Bu nedenle kimi kuramcılar (Lord 1980; Hambleton ve Swaminathan 1985; Algina ve Crocker 1986) tarafından iki varsayım tek başlık altında incelenmektedir.

Yerel bağımsızlık, belirli bir yetenek düzeyinde olan alt grupların, bir maddeye verdikleri cevapların bağımsızlığına işaret etmektedir. Örtük özellikler teorisinin ikinci genel varsayımı olan yerel bağımsızlık, bir bireyin değişik test maddelerine verdiği yanıtların istatistiksel olarak bağımsız olduğudur. Maddelerin birbirinden bağımsız olarak cevaplandırılması, maddelerden birine verilen cevabın başka maddeye verilecek cevabı etkilememesi, her bir maddenin tek başına cevaplanabilir olmasıdır. Başka bir ifadeyle, bir maddeye verilen doğru ya da yanlış yanıtın, başka bir maddeye verilecek

doğru ya da yanlış yanıtı bağımsız olmasıdır. Bu da maddelerin olasılık yoğunluk fonksiyonlarının çarpımlarının, maddelerin bileşik olasılık yoğunluk fonksiyonlarına eşit olmasıdır. Yani maddelerin birlikte doğru cevaplandırılma olasılıkları, maddelerin ayrı ayrı doğru cevaplandırılma olasılıklarının çarpımına eşit ise yerel bağımsızlık varsayımının sağlandığı söylenebilir. Bireylerin performanslarını etkileyen yetenek düzeyleri sabit tutulduğunda, maddelere verilen tepkilerin istatistiksel olarak birbirinden bağımsız olması yerel bağımsızlık göstergesidir. Farklı yetenek düzeylerini içeren bir grubun, iki maddeye gösterdikleri tepkilerin korelasyon göstermesi, bu korelasyonun o yetenek boyutundan kaynaklandığına işaret ederken; yetenek bakımından aynı düzeyde olan kişilerin iki maddeye gösterdikleri tepkiler arasında ilişki olması, ölçülen yetenekten başka özelliklerin de maddedeki performansı etkilediğine işaret etmektedir.

θ , yetenek göstergesi tek boyutlu olduğundan, yerel bağımsızlık varsayımı tek boyutluluk varsayımına eşdeğer olmaktadır. Tek boyutluluk ve yerel bağımsızlık varsayımlarının bu eşitliğinden yola çıkarak, yerel bağımsızlık için de faktör analizi kullanılabilir. Bu durumun test edilmesi için farklı teknikler de önerilmiştir. Madde puanları arasındaki tetrakorik korelasyonlar veya ki-kare analizi kullanılabilecek teknikler olarak görünmektedir. Korelasyon tekniğinde, belirli bir yetenek düzeyindeki bireyler için maddeler arasında korelasyonların düşük çıkması yerel bağımsızlığın bir göstergesi olarak düşünülür. Yüksek korelasyon çıkarsa, belirli yetenek düzeyi için doğru cevap verme olasılıklarına başka değişkenler karışıyor demektir. Kİ-kare analizinde, iki maddeye verilen doğru ve yanlış yanıtlar arasında uyum bulunması, maddelere verilen tepkilerin birbirini etkilediği şeklinde yorumlanır (Lord 1980; Hambleton ve Swaminathan 1985; Algina ve Crocker 1986; Gelbal1994; Hambleton ve diğerleri 1991).

Normal Dağılım: Örtük özellikler teorisinin bütün modelleri için verilerin normal dağılıma sahip olması önemli bir gerekliliktir. Verilerin normal dağılıp dağılmadıkları Ki-kare testi, Kolmogorov-Simimov uyum iyiliği testi, betimsel istatistikler aracılığıyla

ya da grafiksel yöntemle test edilebilir. Ancak, gerçek uygulamalarda bu koşulun sağlanması çoğunlukla mümkün olmamaktadır. Koşulun sağlanmamasının nedenlerinden biri de kullanılan tekniklerin yapısından kaynaklanmaktadır. Kİ-kare ve Kolmogorov-Simimov testi beklenen ve gözlenen olasılıkların farkının toplamına dayandığından, uygulama evreni büyüdükçe normallik varsayımını sağlamak güçleşmektedir. Yapılan araştırmaların birçoğunda, diğer iki varsayım sağlanması koşuluyla, puan dağılımı normal olmadığı halde çalışmalara devam edildiği görülmektedir (Lord 1980; Hambleton ve Swaminathan1985; Cunningham 1986; Algina ve Crocker 1986; Gelbal 1994; Stage 1998a; Kelecioğlu 2001).

Örtük Özellik Kuramında Kullanılan Modeller

Ölçme sonuçlarında elde edilen veriler üç sınıfa ayrılabilir. Bunlar; iki kategorili veriler (dichotomous), çok kategorili veriler (polychotomous) ve sürekli verilerdir (continuous). Örtük özellik modelleri genel olarak verilerin doğasına göre sınıflanabilir.

Bireyin bir yanıtını doğru ya da yanlış olarak iki kategoriye ayıran ve 1 ya da 0 olarak puanlanabilen veriler için 1,2 ve 3 parametrelili "normal ogive" ve "lojistik" modeller vardır. Örtük özellikler kuramı, gelişmeye başladığı 1950'li yıllarda, daha çok 1-0 şeklinde iki kategorili puanlamaya elverişli çoktan seçmeli başarı ve yetenek testlerinin geliştirilmesi, puanlanması ve değerlendirilmesinde kullanılmıştır. Bunların dışında sınıflı cevap modeli, 4 parametrelili model ve ölçekli cevap modelinden de bahsedilmektedir.

Uygulanan testlerde, cevaplayıcıların yanıtlarını iki kategoriye ayırmanın dışında da tepki kategorileri olması nedeniyle başka modellere gereksinim duyulmuştur. Çok kategorili veriler için ilk model önerisi Samejima tarafından 1969 yılında ileri sürülmüştür. Bu modellerden en çok bilinenleri şunlardır; Samenjima'nın (1969) ağırlandırılmış tepkiler (Graded Response) modeli, Bock'un (1972) sınıflamalı tepkiler (Nominal Response) modeli ve Masters'ın (1982) kısmi puanlama (Partial Credit) modelidir. Bunların dışında çok boyutlu ölçekler için geliştirilen modellerden de söz

edilebilir. Üzerinde çalışılan çok kategorili ve çok boyutlu model fazladır (Hambleton ve Swaminathan, 1985; De Ayala, 1993; Park 2000).

Modellerdeki farklılık, madde karakteristik eğrisinin formülasyonundan veya puanlamadan ileri gelmektedir. Madde karakteristik eğrisi yardımıyla yetenek boyutunda birbirinden farklı konumlarda bulunan deneklerin maddeye doğru cevap verme olasılıkları elde edilmektedir. Örtük özelliklerde kullanılan modeller için, “lojistik” ve “ogive” modellerdeki madde karakteristik fonksiyonlarının içerdiği madde parametrelerinin sayısına göre tanımlar yapılmıştır. Madde parametreleri kestirilirken maksimum olabilirlik (Maximum Likelihood) veya Bayes tekniklerinden biri kullanılır. Bayes tahminlerinin standart hatalarının küçük olmasına rağmen, 8 parametresine ilişkin varsayımlarının fazlalığı sıkça kullanımını önler. Yaygın bir istatistiksel teknik olan maksimum olabilirlik tekniğinde, gözlenen verilerin olasılıklarını maksimum yapan model parametreleri, parametre değeri olarak alınmaktadır. Bu teknikle elde edilen parametre tahminleri, maksimum olabilirlik tahminleri olarak bilinmektedir. Tahmin etme süreci ile maddeler ve denekler, gerçek performans olasılıkları ile beklenen olasılık parametreleri arasındaki ilişkinin en yakın olmasını sağlayacak biçimde yetenek ölçeği üzerine yerleştirilirler. Önce maddeye verilen cevaplar yardımıyla seçilen modele bağlı olarak, a, b ve c parametreleri kestirilir. Bu parametrelerin yardımıyla, dağılımın normal ve özelliğin bir boyutlu olduğu varsayımı altında bireylerin yetenekleri kestirilir. Madde parametre tahminleri ve deneklerin yetenek tahminleri, gerçek test verileri ile yetenek ve madde tahminleri arasında en yüksek uyum sağlanana kadar sürdürülür. Bu olasılığın elde edilmesinde bir, iki ve üç parametre kullanan farklı modeller bulunmaktadır. Seçilen modele göre madde parametrelerinin sayısı da değişmektedir. Bir parametrelili modelde b; 2 parametrelili modelde a ve b; 3 parametrelili modelde a, b, ve c parametreleri vardır.

Bu modellerin işe koşulabilmesi için, öncelikle örtük özellikler teorisinin yukarıda anlatılan normallik, tek boyutluluk ve yerel bağımsızlık varsayımlarının karşılanması ve bunlar sağlandıktan sonra da seçilen model ile verilerin uyumluluğunun test edilmesi

gerekir. Model-veri uyumu için; verilerin model varsayımlarını ne derece karşıladığına, modeli kullanırken beklenen avantajlardan ne kadarının sağlandığına ve modelden kestirilen değerlerle gerçek değerlerin uyumuna bakılması gerekir. Kaç parametrelili modelin kullanılacağına karar vermek için maddelerin modele uyumuna bakılırken, Ki-kare veya $-2 \log \lambda$ istatistiklerinden yararlanılabilir. Ki-kare istatistiğini kullanmak daha yaygındır. Ki-kare uyum testinde büyük değerler uyumsuzluğa işaret eder. Madde karakteristik eğrisinin teorik konumu ile gözlenen (deneysel) eğri arasındaki fark arttıkça, Ki-kare değeri büyür ve maddenin modele uyum göstermediği anlaşılır. Bu durumu grafiksel olarak gösteren EO-FIT gibi programlar da mevcuttur ve model-veri uyum testinde grafiksel teknik olarak da bilinir. Elde edilen istatistiklere bağlı olarak hangi maddenin hangi modele uyumlu olduğuna karar verilir (Hambleton & Swaminathan 1985; Mislevy ve Bock 1990; Hambleton ve diğerleri 1991; Baker 2001; Pere ve Verbano 2001).

Modelin beklentilerine uymayan maddeler, testle uyum sağlamadığı düşünülerek testten çıkarılır veya üzerinde düzeltme çalışmaları yapılır. Uyum analizleri sonucunda çıkarılan bir maddenin 1- Yazımında teknik hata olduğu, 2- Ayırt edicilik düşük olduğu, 3- Testteki diğer maddelerden farklı olduğu, 4- Test ile ölçülen yetenekten farklı bir değişkeni ölçtüğü, 5- Şansla cevaplanabildiği, 6- Testteki diğer maddelerden güç olduğu 7- Yetenek dağılımı çok geniş bir kitleye uygulandığı düşünülmektedir (Özçelik ve Berberoğlu 1989).

Yukarıda değinildiği üzere, örtük özellikler kuramında, bir boyutlu ve 1-0 puanlamaya sahip testlere uygulanabilen modeller parametre sayısına göre üç başlık altında incelenebilir. Bu modeller ve model parametreleri aşağıda kısaca açıklanmıştır.

Bir Parametrelili Lojistik Model: En basit örtük özellik modellerindendir. Rasch modeli olarak da bilinir. Madde tepki işleyişini betimlemek için bir parametrenin yeterli olduğunu varsayar. Bu da madde güçlük parametresini karşılayan “b” parametresidir. Madde karakteristik fonksiyonu, “lojistik” modele bağlı olarak aşağıdaki gibi formüle edilmiştir.

$$P_j(\theta) = \frac{e^{D(\theta - b_j)}}{1 + e^{D(\theta - b_j)}}$$

(Hambleton ve Swaminathan 1985:47)

Madde karakteristik eğrisinin maksimum eğime sahip olduğu noktada maddenin b parametresi bulunur. Eğrinin büküm noktasının izdüşümündeki yetenek değeridir. Maddenin en iyi ölçtüğü yetenek düzeyini de gösterir. Maddenin örtük özellik boyutu üzerindeki konumunu (yerini) vermektedir. Bir parametrelili modelde şans parametresinin minimum veya 0 olduğu ve ayırt edicilik gücünün bütün maddeler için eşit olduğu düşünülür (bütün maddeler için 1 değeri veya hesaplanan ortalama madde ayırt edicilik değeri). Verilen cevapların sadece b parametresinden etkilendiği düşünülür. Bir parametrelili modelin en önemli sınırlılığı bütün maddelerin eşit ayırt edicilik olduğu varsayımdır. Bu durumla uygulamada hemen hemen hiç karşılaşmaz. Madde güçlük parametresi b , θ yetenek parametresi ile aynı ölçektedir ve yetenek parametresine göre tanımlanır. Örtük özelliğin düzeyini gösterir ve $-\infty$ ile $+\infty$ aralığında değerler alabilir, ancak uygulamada -3 ile +3 arasında değerler aldığı gözlenmektedir (Bayku! 1979; Lord 1980). Negatif b değeri maddenin kolay, pozitif b değeri maddenin zor ve 0'a yakın b değeri maddenin orta güçlükte olduğunun işareti kabul edilebilir. Bu parametre, maddenin doğru cevaplandırılma olasılığının 0,50 olduğu noktadaki yetenek düzeyini (örtük özellik puanını) belirtir. Bu parametrenin değeri yükseldikçe maddeyi doğru cevaplandırmak için gerekli yetenek düzeyi de artmaktadır. Bir parametrelili modelin en önemli avantajı yetenek kestiriminde getirdiği kolaylıktır. Ancak model-veri uyumunu sağlamanın zorluğu çok büyük madde havuzunu gerektirir.

İki Parametrelili Lojistik Model: İki parametrelili model bir parametrelili modelin sadece b parametresini yeterli sayması karşısında, tek parametrenin birçok durumda yetersiz olduğu görüşünden doğmuştur ve bir parametrelili modele ikinci parametreyi eklemiştir. Bu da “ a ” parametresidir. İki parametrelili model için madde karakteristik fonksiyonu, “lojistik” modele bağlı olarak aşağıdaki gibi formüle edilmiştir.

$$P_j(\theta) = \frac{e^{Da_j(\theta - b_j)}}{1 + e^{Da_j(\theta - b_j)}}$$

(Hambleton ve Swaminathan 1985:36)

Bu modele eklenen ikinci parametre, madde ayırt edicilik parametresi olarak da bilinir. Kuramsal olarak a parametresi $-\infty$ ile $+\infty$ aralığında değerler alabilir. Uygulamada ise genellikle 0 ile 2 veya -2,80 ile +2,80 arasında değerler aldığı bildirilen farklı görüşler bulunmaktadır (Baker 2001). Değer yükseldikçe maddenin ayırt ediciliği de artar. Maddenin kalitesini gösteren bir tür madde geçerlik ölçüsüdür. Bu anlamda elde edilen θ yeteneğinin, gerçek θ düzeyi hakkında ne derecede bilgi verdiğini gösterir. İdeal değer olarak, 1 kabul edilir. Diğer özelliklerin sabit tutulması şartıyla, parametre değerleri $b=0$ ve $a=1$ olan maddelerin iyi olduğu söylenmektedir. Ayrıca a değeri 0,80 altında olan maddenin kullanılmamasını önerenler de vardır (Baker 2001). Ayırt edicilik parametresi, bir madde karakteristik eğrisinin eğimine ya da dikliğine (steepness; Thomdike 1982) karşılık gelmektedir. Madde karakteristik eğrisi çok dikse madde ayırtıcılığı, çok dar bir yetenek grubunda seçim yapmak için özel testler hazırlamada kullanılabilir. Bu bakımdan, dik eğriye sahip maddeler fazla kullanışlı olmayabilir. Farklı yetenek seviyelerini ayırt edemeyen maddeler görece düz bir eğri vermekte ve düşük a değerine sahip olmaktadır. Ayırt edicilik 0'a yakınsa eğri düzleşir ve böyle maddeler ölçülen özellik hakkında ve dolayısıyla da 8 yetenek düzeyi hakkında bilgi vermezler. Böyle maddeler teste alınmamalıdır. Ayırt edicilik değeri eksi ise negatif Ayırt edicilikten (Baker 2001:33) söz edilir. Böyle maddelerde, doğru cevaplandırılma olasılığı alt yetenek düzeyinden üst yetenek düzeylerine gidildikçe düşmektedir. Bu tür maddelerin ölçülen değişken dışında bir değişkeni ölçtüğü veya hatalı olduğu kabul edilir. En iyi maddeler yetenek düzeyi arttıkça doğru cevaplandırılma olasılığı da artan maddelerdir. Madde Ayırt ediciliği önemli bir faktör olduğu durumlarda, 2 parametreliliğin kullanılması önerilmektedir. İki parametreliliğin en önemli sınırlılığı şansla doğru cevap verme olasılığını dikkate almamasıdır.

Üç Parametrelili Lojistik Model: İki parametrelili model, bir parametrelili modele önemli bazı eleştiriler getirmiştir. Ancak önemli bir diğer faktör olan şanstaki söz etmemiştir. Üç parametrelili model, bir ve iki parametrelili modelleri eleştirdikten sonra bu parametrelere bir parametre daha eklemiştir. Bu parametre şans parametresi de denilen “c” parametresidir. Üç parametrelili model için madde karakteristik fonksiyonu, "lojistik" modele bağlı olarak aşağıdaki gibi formüle edilmiştir (Hambleton ve Swaminathan 1985:37).

$$P_j(\theta) = c_j + (1 - c_j) \frac{e^{D_{jq}(\theta - b_j)}}{1 + e^{D_{jq}(\theta - b_j)}}$$

Şansın etkili olduğu testlerin analiz edilmesinde üç parametrelili modelin kullanılması önerilmektedir. Kuramsal olarak, üç parametrelili modelin daha az hatalı sonuçlar verdiği, diğer bir deyişle daha fazla bilgi verdiği ya da daha uygun bir model olduğu bildirilmektedir. Model a ve b'ye ek olarak maddeyi şansla doğru cevaplandırma olasılığın da ele alır. Madde karakteristik eğrisinin düşük asimptotuna karşılık gelmektedir. Bir anlamda yetenek düzeyi en düşük kişilerin maddeye doğru cevap verme olasılığıdır. Bu parametre şansla doğru cevabı bulma olanağı veren, özellikle çoktan seçmeli maddeler için önem taşımaktadır. 0 ile 1 arasında değerler alabilir. Doğru cevabı bilmeyen cevaplayıcı doğru cevabı c olasılığı ile tahmin edecektir. Yani c maddenin tahminle cevaplandırılma olasılığıdır. Tartışmalara neden olan c parametresi, bazı araştırmacılara göre maddeyi cevaplandırmak için gerekli en düşük yetenek düzeyi iken; bazı araştırmacılara göre maddeyi şansla cevaplandırma olasılığıdır. (Lord ve Novic 1968; Baykul 1979; Lord 1980; Thorndike 1982; Hambleton ve Swaminathan 1985; Cunningham 1986; Osterlid 1989; Hambleton ve diğerleri 1991 ; Van der Linden ve Hambleton 1997; Harvey ve Hamner 1999; Fan 1998; Baker 2001). Örneğin, Lord'a göre “sözde şans”tır. Lord (1980), c parametresinin, maddenin tahmin parametresi olarak alınması yerine madde karakteristik eğrisinin bir alt sınırı olarak düşünülmesini; parametre olarak alındığında yeteneğin bir parçası olması gerektiğini, ancak yetenekten bağımsız ve madde seçenek sayısına bağımlı olduğunu söylemektedir.

Örtük Özellikler Kuramının Avantajları

Genel varsayımlar ve model-data uyumu sağlandığı takdirde örtük özellikler teorisini klasik test teorisinden üstün kılan bazı avantajların elde edilebileceği belirtilmektedir.

Örtük özellikler kuramının avantajları içinde parametrelerin değişmezliği (invariant) önemli yer tutar. Kuramın diğer avantajlarının değişmezlik avantajına bağlı olduğu söylenebilir. Değişmezlik, test ve madde parametrelerinin uygulanan gruba bağımlılığı veya bireyler için hesaplanan yetenek ölçümlerinin uygulanan teste bağımlılığı gibi sınırlılıkları ortadan kaldırmaktadır. Değişmezlik iki açıdan ele alınabilir. Birincisi; aynı özelliği ölçmeye yönelik olarak hazırlanmış olan farklı maddelere verilen tepkilere dayalı olarak kestirilen yetenek parametreleri değişmezdir. İkincisi; aynı maddelerin farklı bireylere uygulanmasıyla elde edilen madde parametreleri değişmezdir (Hambelton ve Swaminathan 1985; Lynch 1997). Değişmezlik özelliğine göre, bireylerin yetenek kestirimleri testten, madde parametreleri de testin uygulandığı gruptan bağımsızdır. Farklı maddelerden kestirilen yetenek parametreleri ve farklı gruplardan kestirilen madde parametreleri aynı veya benzer sırayı almaktadır. Hambleton ve Swaminathan (1985) bu iki temel avantaja, “yetenek tahminlerinin farklı yetenek düzeyleri için bilinebilir” olmasını eklemektedir.

Bu avantajların kuramın yapısından geldiği iddia edilmektedir. Kurama göre yetenek (θ), hem seçilen test maddelerinden hem de ölçeği cevaplayan kişilerin performansından bağımsızdır. Böylece maddelerin ve ölçeğin uygulandığı çeşitli grupların karşılaştırılması mümkün olmaktadır. Bir kişinin soruya doğru cevap verme olasılığı, madde karakteristik eğrisinin şekline bağlı olduğundan, kişinin yeteneği grubun yetenek dağılımından bağımsız olacaktır. O halde bir kişinin bir soruya doğru cevap verme olasılığı, maddeye cevap veren gruba ya da kişiyle aynı yetenek düzeyindeki kişilerin sayısına bağlı olmayacaktır. Van de Rijt ve diğerleri (1999), kuramın yetenek parametrelerini belirlemedeki tutarlılığından dolayı, özellikle erken dönemlerde (ilköğretim çağında), yetenekleri belirlemede kullanılacak testlerin

geliştirilmesinde, bireysel farklılıkları görme açısından klasik kuramdan daha avantajlı olduğunu belirtmektedir.

Kuram, madde parametrelerinin de gruptan bağımsız olarak kestirilebileceği iddiasındadır. En azından, madde parametrelerinin bağımsız olmaması durumunun (örneğin madde ve test yanlılığının tespit edilmesi) belirlenebileceği iddia edilmektedir. Madde parametrelerinin farklı gruplara uygulandığında büyüklüklerinin aynı sırada olması beklenmektedir. Burada dikkat çeken nokta madde veya test yanlılığının gruba bağımlılıkla aynı anlamda kullanılmamasıdır. Oysa madde ve test yanlılığı, parametre değerlerinin farklı gruplar için değişmesi anlamına gelebilir.

Ayrıca, klasik kuramdaki güvenilirliğe karşılık gelen madde ve test bilgi fonksiyon değerlerinin de değişmediği iddia edilmektedir.

Örtük özellikler kuramının, bahsedilen avantajların dışında, test yapılandırma (test construction), soru bankası oluşturmada (item banking), bireysel test oluşturmada (tailored-adaptive test construction), test maddelerinin belirli bir grup için yanlılık (differential item functioning-DIF) gösterip göstermediğini belirlemede, madde ve seçenekleri ağırlıklandırmada ve testlerin eşitlenmesinde (test equating) karşılaşılan sorunlara rahatlıkla yanıt verdiği iddia edilmektedir. Özellikle bilgisayarda ölçme uygulamalarıyla bireysel testler için en uygun kuram, örtük özellikler kuramı olarak görülmektedir (Wright ve Stone 1979; Hambleton ve Swaminathan 1985; Keesee 1988; Osterlind 1989; Wainer 1990; Hambleton ve diğerleri 1991; Stage 1998e).

Örtük Özellikler Kuramının Sınırlılıkları

Örtük özellikler kuramının da belirli sınırlılıkları bulunmaktadır. Bu sınırlılıklardan en önemlisi, örtük özellikler kuramının klasik kuramın bir uzantısı olmasıdır. Kuramın varsayımlarını karşılamak özellikle gerçek verilerde çok güçtür. Bu sınırlılık nedeniyle yaygın olarak kullanılamamaktadır.

Örtük özellikler kuramını uygulamak teknolojik desteği gerektirmektedir. Madde

karakteristik fonksiyonun tanımlanmasındaki karmaşıklık nedeniyle parametrelerin hesaplanması çok güçtür. Bu nedenle, özellikle çok maddeli ölçeklerde ve çok cevaplayıcı bulunduğu, bilgisayar programlarından yararlanma zorunluluk gibi görünmektedir. Bu amaçla hazırlanmış çok sayıda özel yazılım bulunmaktadır. Bu yazılımlara ek olarak çeşitli istatistik paket programları da menülerine örtük özellikler kuramı analizlerini kısmen de olsa eklemeye başlamışlardır (Systat 8.0 ve üstü ile NCSS2000-P ASS200 bilinen paket programlardır) (Lord ve Novic 1968; Algina ve Crocker 1986; Somer 1998).



Çoktan Seçmeli ve Klasik Testler

Eğitim ve psikoloji alanlarında kullanılan araçlar, cevap formatına göre iki türe ayrılır; Birincisi seçme gerektiren, diğeri de açık uçlu (klasik) araçlardır. Bu araçların her birbirine göre birçok avantaj ve dezavantajları vardır. Çoktan seçmeli testlerde madde sayısının daha fazla olması ve bu yüzden daha çok sayıda özelliği kapsaması bu testlere bir avantaj sayılır. Bu nedenle çoğu ölçmelerde çoktan seçmeli testler kullanılır.

İstenen özellikler açısından yeterli olabilecek çoktan seçmeli test üretmek zordur. Çeldiricilerin, o soruyla ölçülmek istenen bilgiye sahip olmayan bireyleri yanıltırken sahip olanları da yanıltmamaları gerekir. Aksi taktirde madde ve test ayıricılık gücünü yitirecektir. Genelde daha çok seçeneğe sahip olan testler daha güvenilir ve daha uygun kabul edilir. Bu düşüncenin nedeni ise daha çok seçeneğe sahip olan testlerde bilgi sahibi olmayan bireylerin, soruyu rasgele olarak doğru cevaplama olasılıklarının azalmasıdır. Bazı araştırmacılar derslerinde çoktan seçmeli testlerde ideal seçenek sayısını ortaya koymaya çalışmışlardır. Bu çalışmaların çoğu başarı testleri üzerinde ve Klasik Test kuramına dayalı olarak yapılmıştır. Bu çalışmaların sonuçları “İlgili Araştırmalar” kısmında bulunmaktadır.

Bu çalışmanın amacı, çoktan seçmeli testlerde seçenek sayısı değiştiğinde, madde parametreleri, farklı yetenek düzeylerinde madde bilgi fonksiyonları ile madde uyumluluğu üzerinde ne gibi değişiklikler meydana getirdiğini incelemektir. Bu incelemede birisi 4 seçenekli diğeri ise 6 seçenekli iki paralel yetenek testi, Örtük Özellikler Kuramında maddelerin ayırt edicilik güçleri, güçlük dereceleri, şans parametreleri, bilgi fonksiyonları ve uyumlulukları karşılaştırılmıştır.

Problem cümlesi:

Çoktan seçmeli testlerde seçenek sayısı madde parametrelerini, madde uyumluluğunu ve madde bilgi fonksiyonlarını nasıl etkilemektedir ve ne tür değişikliklere yol açmaktadır?

Alt problemler:

Bu çalışmanın alt problemleri şöyledir:

- 1) Seçenek sayısı arttığında ayırt edicilik gücünde (a_j) ne tür değişiklikler ortaya çıkmaktadır?
- 2) Seçenek sayısı arttığında güçlük derecesinde (b_j) ne tür değişiklikler ortaya çıkmaktadır?
- 3) Yanlış cevap sayıları farklı yetenek düzeyleri üzerinde eşit midir?
- 4) Seçenek sayısı arttığında farklı yetenek düzeyleri üzerinde yanlış cevap sayıları değişimleri eşit midir?
- 5) Seçenek sayısı değiştiğinde madde bilgi fonksyonu değişir mi? Bu etki farklı yetenek düzeyleri üzerinde değişir mi?
- 6) Seçenek sayısı değiştiğinde madde uyumluluğunda ne gibi değişiklikler ortaya çıkmaktadır?

Araştırmanın Önemi ve Amacı:

Psikoloji ve eğitimde doğru kararlar verebilmek için, hatasız puanlar ve ölçümler gerekmektedir. Bu hatasız puanlar ve ölçümler; güvenilir ve geçerli testler, yöntemler ve modellerden elde edilir.

Genelde her model ve her yöntem bir çok sayıtlı ve kurallara dayanır. Bu modeller ve yöntemlerin güvenilirliği, bu kurallar ve sayıtların doğru olup olmamasına bağlıdır. Günümüzde, psikolojide ve eğitimde en çok kullanılan ölçme ve değerlendirme modellerinden birisi de IRT modelidir. IRT modelinde de başka modellerde olduğu gibi bir çok sayıtlı ve kural yer alır. IRT modeli kullanıldığı zaman bu sayıtların ve kuralların doğru olup olmaması incelenmelidir. İncelenmesi gereken kurallardan bir de şans parametresi ve miktarının farklı düzeylerde eşit olmasıdır. Ayırt edicilik ve güçlük parametrelerinin veya madde bilgi fonksiyonunun da farklı koşullarda doğru olup olmadığı gerekli olan incelemelerin içinde yer alır.

Hem yeni bir kural ve prensip oluşturmak amacıyla, hem de eski kurallarda gereken değişiklikleri yapmak için, çeşitli çalışmaların ve incelemelerin yapılmasına gerek vardır. Esasında her bilim dalında sayıtların doğru olmasına yönelik deneyler ve incelemeler yapılır ve gereken değişiklikler yapılarak yeni sayıtlar , prensipler ve kurallar ortaya konulur.

Bu çalışmada, ölçme ve değerlendirmede elde edilen sonuçların güvenilirliğini ve geçerliğini arttırmak, sonunda da psikolojide ve eğitimde sağlam kararlara varmak, ayrıca yukarıda bahsettiğimiz IRT modelindeki sayıtlar üzerinde inceleme yapmak ve bu sayıtların sağlanması için önerilerde bulunmak amaçlanmaktadır. Bu önerilerden örneğin, daha çok seçeneğe sahip olan testin varyansının, daha az seçeneğe sahip olan testin varyansına göre daha yüksek olması durumunda testin güvenilirlik ölçüsü olarak önerilebilir.

Varsayımlar:

Bu çalışmanın varsayımı şöyledir :

- Bireylerin testteki yanlış cevap sayısı şanstın yararlanma miktarını yansıtır. Başka bir ifadeyle yanlış cevap sayısı fazla olduğunda o kişinin daha çok şansı kullandığını ve yanlış cevap sayısı az olduğunda da şansı daha az kullandığını gösterir.

Sınırlılıklar

Bu çalışmada aşağıdaki sınırlılıklarla karşılaşmıştır:

- 1- Bu çalışmada iki testin seçenek sayılarının farkı sadece ikidir (birisi 4 ve diğeri 6 seçenekli). Seçenek sayısının farkı arttığında sonuçlar değişebilir.
- 2- Bu çalışma, yetenek testleri üzerinde yapılmış ve okuldaki derslerde kullanılan başarı testlerini (sınavlardaki testleri) içermemektedir.

İlgili Araştırmalar:

Haladyna (1992) yaptığı çalışmada, Karmaşık çoktan seçmeli (complex multiple choice), İki seçenekli (alternate choice), Doğru – Yanlış (true-false), çoklu doğru-yanlış (multiple true-false) ve metne bağlı cevaplanan maddelerin (context dependent items) formunda maddeler uyguladıktan sonra değişiklikleri incelenmiştir.

Mehrens ve Maihoff (1985) yaptığı çalışmada, doğal bilimler bölümünde seçmeli testlerin iki farklı formu karşılaştırılmıştır. İki seçeneğe dönüştürülmüş bir çoktan seçmeli test türü olan “iki seçenekli” ve “doğru-yanlış” testleri uygulanarak başarı testlerinde madde güçlüğü, ayırt edicilik gücü, güvenilirlik ve geçerlikleri birbirleriyle karşılaştırılmıştır. Ayrıca bu çalışmada, bu iki testte (doğru-yanlış (true-false) ve doğru – çeldirici (alternate-choice test)) testlerin seçeneklerinin yerleri değiştiğinde ne tür değişiklikler ortaya çıktığı incelenmiş ve karşılaştırılmıştır. Elde edilen sonuçlara göre doğru-yanlış testinin güçlüğü doğru – çeldirici teste karşı düşük çıkmış fakat iki testin ayırt edicilik gücü, güvenilirlik ve geçerliklerinde hiç bir değişiklik bulunmamıştır.

Bielinski ve diğerlerinin (2000) yaptığı araştırmada düzey dışında, ölçmenin özellikleri tartışılmıştır. Şans faktörü de tartışılan konular arasında yer alır. Maalesef şu ana kadar bu iddiaya yanıt olan hiç bir çalışma yapılmamıştır. Yüksek yetenek seviyesindeki bireylerde mi, yoksa düşük yetenek düzeyindeki bireylerde mi şans başarısına daha çok yönelme olduğu, şans başarısı için düzeltme formülü kullanılarak belirlenmiştir. Düzey dışında, ölçmede düşük düzeylerde olan bireylerin, yüksek düzeylerdeki bireylere göre şans faktörünü daha fazla kullandıkları belirtilmiştir.

1972’de Illinois üniversitesinde Frank Costin, final sınavlarında kullanılan üç seçenekli ve dört seçenekli testlerin güvenilirliği ve iç tutarlıklarını Urbana-Champaign

Kampüsü’nde psikoloji dersi üzerinde karşılaştırmış ve incelemiştir. Bu çalışmada ders konularını içeren 100 maddeden oluşan dört seçenekli bir test ele alınmış, daha sonra bu maddelerden 50 madde rasgele alınmış ve her maddenin de çeldirici seçeneklerinden

birisi rasgele olarak seçilerek atılmıştır. Bu test üç seçenekli hale dönüştürülmüştür. Böylece bu çalışmada iki tane 50'şer maddeli test kullanılmış ve bunlardan birisi üç, diğeri ise dört seçenekten oluşmuştur. Bu testler final sınavı olarak 1566 öğrenci üzerinde uygulanmıştır ve istatistiksel ölçüleri örneğin ortalama, standart sapma, her maddenin doğru cevap sayısı ve diğer ölçüler hesaplanarak KR-20 (Kuder-Richardson 20 formülü) ile her iki testin ayrı ayrı güvenilirliği ve iç tutarlığı ile ayrıca her maddenin tüm testle olan ilişkisi çift serili korelasyon katsayısıyla hesaplanmıştır. Elde edilen sonuçlara göre iki testin güvenilirlik ve iç tutarlık ile maddelerin çift serili korelasyon katsayısı ortalamaları arasında hiç bir anlamlı fark söz konusu olmamaktadır. Bu çalışmada araştırmacı okuldaki sınavlarda üç seçenekli testlerin, nitelik açısından dört seçenekli testlerle aynı seviyede olmasına karşın; zaman tasarrufu ve daha kolay uygulanması açısından tercih edilebilir olduklarını belirtmiştir.

Çoktan seçmeli testlerde üç seçenekli testlerin yanlış anlaşılma sebebi nedir? 1987'de Steven V. Owen (Connecticut Üniversitesi) ve Roben D. Froman (Hartford Üniversitesi) bu soru üzerinde bir çalışma yapmışlardır. Bu çalışmada, sınavlarda kullanan testlerde üç seçenekli ve beş seçenekli testleri iç tutarlık, güvenilirlik, soruları cevaplama zamanı, ayırt edicilik gücü ve güçlük derecelerini karşılaştırarak üç seçenekli testlerin daha uygun olduklarını incelemişlerdir. Çalışmada, iki paralel testten birisi beş seçenekli diğeri üç seçenekli olarak 114 öğrenci üzerinde uygulamış ve ortaya çıkan sonuçlara göre iki testin iç tutarlık, güvenilirlik, ayırt edicilik gücü ve güçlük dereceleri ve puanları arasında hiç bir anlamlı fark bulunamamış, fakat üç seçenekli testteki geçirilen zaman beş seçenekli teste göre daha düşük görülmüş ve $p < 0.01$ düzeyine anlamlı bulunmuştur. Ayrıca öğrencilerin yüzde 97.4'ü üç seçenekli testten memnun kaldıklarını ve beş seçenekli teste tercih ettiklerini söylemişlerdir. Bu çalışmanın araştırmacıları elde edilen sonuçlara göre dört ve beş seçenekli test üretenlere bu işten vazgeçmelerine dair öneride bulunmuşlardır.

Bir başka çalışma Michael S. Trevisan ve Gilbert Sax (Washington Üniversitesi) ve William B. Michael (Southern Üniversitesi) tarafından yapılmıştır. Bu çalışmada testlerdeki seçenek sayısının ve yetenek düzeylerinin etkisi test geçerlik ve güvenilirliği üzerinde incelenmiş ve 1991’de Educational and Psychological Measurement dergisinde yayınlanmıştır. 435 lise öğrencisi üzerinde üç, dört ve beş seçenekli testler uygulanmış olan bu çalışmada, yüksek başarı düzeylerinde üç seçenekli ve düşük düzeylerde beş seçenekli testlerin daha yüksek güvenilirlik ve geçerlik gücüne sahip olduklarına dair bir hipotez kurulmuştur. Elde edilen sonuçlar bu hipotezin onaylanmaması yönündedir. Yani, seçenek sayısı değiştiğinde güvenilirlik ve geçerlikte hiç bir anlamlı değişiklik söz konusu değildir. Ayrıca, sonuçlar yetenek düzeyleri ve seçenek sayısının test güvenilirlik ve geçerliği üzerinde, karşılıklı etkiye sahip olmaması yönündedir. Bu sonuçlara rağmen araştırmacılar önceki diğer araştırmacılar gibi üç seçenekli testlerin zaman tasarrufu ve kolay hazırlanması açısından dört ve beş seçenekli testlere tercih edilebileceğini belirtmişlerdir.

Zimmerman ve Humphries (1953), 700 hava kuvveti askerleri üzerinde beş seçenekli ve 100 maddeden oluşan bir test uygulamışlar, daha sonra da bu testin çeldirici seçeneklerinin arasından en kötüsünü atıp ve böylece iki, üç ve dört seçenekli testler elde edip ve tekrar uygulamışlardır. Elde edilen sonuçlara göre daha az seçeneğe sahip olan testlerde güvenilirlik yükseliş göstermiş ve zaman tasarrufu sağlanmıştır.

Williams ve Ebel (1957), dört seçenekli testleri 700 öğrenci üzerinde uygulayarak, üç ve iki seçenekli testlerle karşılaştırmışlardır. Çalışmadan elde edilen sonuçlara göre; iki ve üç seçenekli testlerde zorluk derecesi düşme (kolaylaşmışlar) ve ayırt edicilik gücü yükselme eğiliminde olup, testlerin iç tutarlıkları arasında hiç bir anlamlı fark olmadığı gözlenmiştir.

Costin (1970), “Psikolojiye Giriş dersinde dört seçenekli başarı testi uygulamış daha sonra bu testi üç seçenekli hale dönüştürerek tekrar uygulamıştır. Elde ettiği sonuçlar üç

seenekli testin hem ayırt edicilik gcnn hem de gvenirliėinin drt seenekli teste gre daha fazla olduėunu gstermiřtir.

Ramos ve Stern (1973) 2400 ėrencinin İřpanyolca ve Fransızca okuma bařarılarını drt ve beř seenekli testlerle lmřler ve iki testin puan ortalamaları, zorluk dereceleri ve madde korelasyonları arasında hi bir anlamlı farka rastlamamıřlardır. Bu arařtırmacılar, beřinci seenek atılırsa, bu durumun madde ve teste hi bir zararı dokunmayacaėı ynnde bir sonuca ulařmıřlardır.

Son olarak Hodson (1983), lisede kimya dersinde uyguladıėı testin madde seenekleri arasından bir seeneėi atarak drt seenekli bir test elde etmiřtir. Bu alıřmada, i tutarlık ve ayırt edicilik gcnn deėiřmemesi ynnde bir sonuca ulařmıř ve drt seenekli testleri, hem ekonomik hem de zaman tasarrufu aısından uygun bulmuř ve beř seenekli testlere gre tecih edilebileceėi ynnde neride bulunmuřtur.

Bu alıřmalar ve diėer benzeri alıřmalar, daha az seenekli testleri, fazla seeneėe sahip olan testler karřısında desteklemektedir. Daha az seeneėe sahip bu testlerin en az, fazla seenekli testler kadar gvenirlik ve geerlik nitelliėini tařıdıkları gsterilmiř ve hem ekonomik hem de zaman tasarrufu ve daha kolay hazırlanabilmeleri aısından daha fazla seeneėe sahip olan testlere gre tercih edilebilir oldukları belirtilmiřtir.

nceki alıřmaların bu alıřmayla iki farkı bulunmaktadır. Bunlardan birisi, nceki alıřmaların geleneksel kuramlara dayalı olan metodlarla yapılmasıdır. rneėin, doėru cevap veren kiřilerin oranının, zorluk derecesinin indeksi olarak alınması ve gvenirlik iin Kuder-Richardson'ın formlnn kullanılmasıdır. Bu alıřma ise, rtk zellikler kuramına dayalı ve madde tepki kuramı modelinde (IRT) yapılmıřtır. Diėer bir fark ise, nceki alıřmaların derslerde uygulanan bařarı testleriyle ilgili olmaları ancak bu alıřmanın ise zeka testleri zerinde yapılmıř olmasıdır.

BÖLÜM II

YÖNTEM

Araştırmanın Türü:

Bu çalışmanın yöntemi, iki testin (dört seçenekli ve altı seçenekli) özellikleri ve parametrelerini karşılaştırmak suretiyle yöntem araştırması sayılır.

Evren ve Örneklem:

İran-Tabriz Üniversitesi'nin öğrencileri bu araştırmanın evrenini oluşturmaktadır ve bu evrenden yaklaşık 985 öğrenci farklı bölümlerden örneklem olarak alınmıştır. Bu örneklem; Fars ve Arap Edebiyatı, İlahiyat, Eğitim Bilimleri, Psikoloji, Matematik, Elektronik, Bilgisayar Mühendisliği, İnşaat Mühendisliği, Kimya, Fizik, İşletme, Sosyoloji, İktisat, Grafik, Spor ve Beden Eğitimi, Tıp ve Sağlık Bilimleri, Yabancı Diller, Hukuk, Veterinerlik ve Ziraat Mühendisliği Bölümlerinden oluşmuştur. Bu bölümler 5 fakültenin listesinden rastgele olarak seçilmiştir. Bu fakülteler de şöyledir: Tıp, İktisat, Fen, Ziraat, Mimarlık ve Sosyal Bilimler Fakülteleri. Bölümler seçildikten sonra, her bölüm içindeki sınıflardan biri rastgele seçilerek ve her sınıftan da öğrenciler rastgele seçilerek oluşturulan örnekleme alınan kişiler üzerinde testler uygulanmıştır.

Araç ve Verilerin Toplanması:

Bu çalışmada, Cattell'in zeka testinin 3. ölçeği A formu altı seçenekli test halinde ve B formu dört seçenekli test halinde (testler paralel sayılırlar) uygulanmışlar. R. B. Cattell tarafından üretilen bu test, "G" faktörünü veya genel yeteneği ölçmektedir. Bu test, gruplara uygulanır ve kültürden bağımsız bir test sayılır. Cattell'in bu yetenek testi üç ölçekten oluşmaktadır. Birinci ölçek 4-8 yaş arasında, ikinci ölçek 8-14 yaş arasında ve diğeri ise 18 yaşın üzerinde olan bireylere aittir. Bu çalışmada üçüncü ölçeğe ait ve

paralel olan A ve B formları uygulanmıştır. Her bir form dört alt testten oluşmaktadır ve toplam 50 madde içerir. Bu alt testlerden bazıları beş ve bazıları altı seçeneğe sahiptir. Bu çalışmada ise, A ve B testlerinin üç alt test ve 36'şar maddeden oluşan formları uygulanmıştır. B formunun alt testlerinin maddelerinin çeldirici seçeneklerinden bir veya ikisi rasgele atılmış ve test maddeleri dört seçeneğe indirilmiştir. A formunun maddeleri ise benzer seçenekler eklenerek altı seçenekli hale dönüştürülmüştür. Bu üç alt testin seçilme nedeni, seçenek sayılarının azaltılıp çoğaltılabilmek kolaylığının olmasındandır. (A ve B test formlarının örneği EK 1'de verilmiştir).

Cattel'in yetenek testinin bu çalışmada kullanılma nedenlerinden birisi, grup halinde ve daha kolay uygulanabilmesi, diğeri ise kültürden bağımsız olup alt yapıyı ölçebilmesi ve tüm evrenlerde genellenebilmesidir.

Testler seçilen örneklem üzerinde uygulandıktan sonra bireylerin zeka ölçüleri ve düzeyleri ile madde parametreleri, bilgi fonksiyonları ve madde uyumluluğu, IRT'de üç parametrelili model kullanılarak elde edilmiştir.

Veri Çözümleme Yöntemleri:

Dört ve altı seçenekli testlerden elde edilen verilerin istatistiksel ölçüleri ve grafikleri elde edilmiş ve testin tek boyutluluğunu tespit etmek için maddeler üzerinde faktör analizi yapılmış ve IRT üç parametrelili modelde madde parametreleri ve madde bilgi fonksiyonları hesaplanmıştır.

Birinci ve ikinci alt problemlerle ilgili olarak seçenek sayısının madde ayırt edicilik gücü (a_j) ve güçlük (b_j) derecesinin üzerindeki etkisinin olup olmadığı, iki testin bu madde özelliklerinin dağılımının normal olduğuna yönelik bir işaret olmaması nedeniyle iki testin bu özelliklerinin farkı parametrik olmayan Wilcoxon'un işaretli sıralar testi ile incelenmiştir.

Üçüncü problem cümlesi ile ilgili olarak önce bireylerin yetenek düzeyleri (Teta düzeyleri), Bilog programı kullanılarak IRT üç parametrelili modelde her teste ayrı ayrı olarak elde edilmiş ve sekiz yetenek grubu oluşturulmuştur. Bu yetenek gruplarının sayısının sekiz olma nedeni ise, normal dağılıma uyumlu olması ve her grubun içindeki bireylerin sayısının normal dağılıma uymasıdır. Bireylerin daha çok alt gruplara ayrılmamış olmasının nedeni; bu gruplarla beklenen sonuçların elde edilmesi ve daha fazla gruplara bölünmesine gerek kalmamış olmasıdır.

Bu alt problemle ilgili yapılmış diğer bir ön işlem ise, bireylerin her iki testte olan yanlış cevaplarını hesaplamak ve bu yanlış cevaplar sayısının şanslı kullanma kanıtı olarak ele alınmasıdır.

Üçüncü alt problem cümlesini inceleme doğrultusunda, bireylerin yetenek düzeyleri ve her iki testteki yanlış cevaplarının arasındaki ilişki, Pearson korelasyon katsayısı yöntemi ile incelenmiştir. Diğer bir analizde, bir kere B testine göre farklı yetenek düzeyinde olan bireyler çeşitli gruplara ayrılarak (sekiz gruba) tek yönlü varyans analizi ile A testinde olan yanlış cevap sayılarının ortalamalarının farkı incelenmiş ve bir kere

de A testine göre farklı yetenek düzeyinde olan bireyler çeşitli gruplara ayrılarak ve tek yönlü varyans analizi ile B testinde olan yanlış cevap sayılarının ortalamalarının farkı incelenmiştir.

Dördüncü alt problem cümlesi için, ilk önce bireylerin iki testte olan yetenek düzeylerinin ortalaması hesaplanmış ve bu ortalamalara göre bireyler üçüncü alt problemde olduğu gibi sekiz yetenek grubuna bölünmüş ve her bireyin her iki testte olan yanlış cevapları da (şans kanıtı olarak) ele alarak tekrarlı verilere ait olan iki yönlü varyans analizi kullanılmıştır.

Beşinci alt problem cümlesini incelemek için, yani seçenek sayısının madde bilgi fonksiyonu üzerinde etkisi ve bu etkinin farklı yetenek düzeylerinde değişik olup olmadığını incelemek için, $\theta = -12, \theta = -11.5, \dots, \theta = 0.00, \dots, \theta = 11.5, \theta = 12$ 'de, iki testte 49 grupta bilgi fonksiyonları hesaplandıktan sonra tekrarlı ölçülere dayalı olan iki yönlü varyans analizi ile incelenmiştir.

Altıncı alt problem cümlesini incelemek için, hem üç parametrelili hemde iki parametrelili IRT modellerinde iki testin uyumlu madde sayısının farkı hem iki yönlü varyans analizi ile hemde Ki-Kare ve McNemar testi ile incelenmiştir.

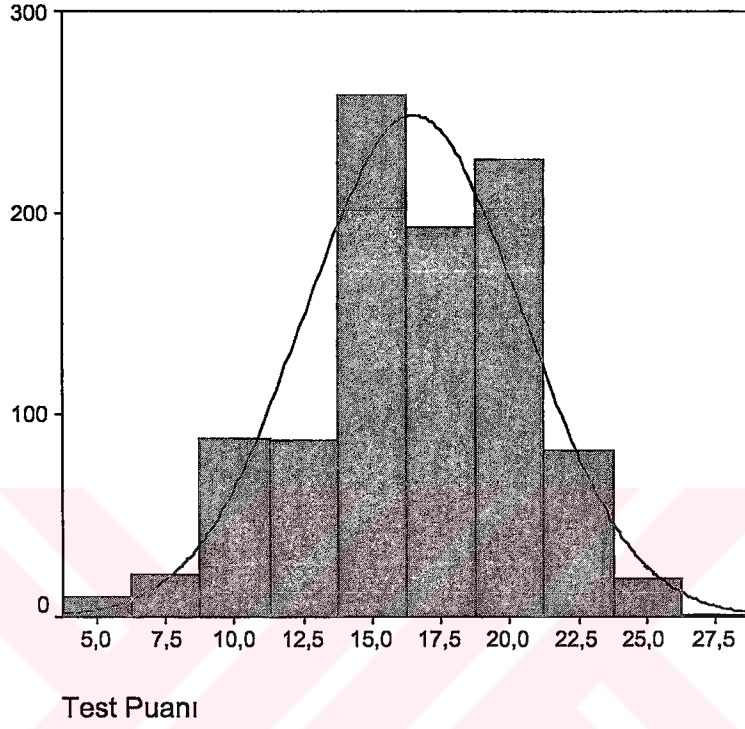
Testlerin Puan, Parametre ve Bilgi Fonksiyonlarının Betimsel Bilgileri:

İki test formu da puanlandıktan sonra, her iki testin madde parametreleri ve yetenek düzeyleri üzerinde madde bilgi fonksiyonları elde edilmiştir. Aşağıdaki tablolarda ve grafiklerde iki test puanları ile ilgili olan istatistiksel bilgiler hem de madde parametreleri ve madde bilgi fonksiyonları görülmektedir:

Tablo 1: A ve B Testlerine Ait Olan İstatistiksel Ölçüler

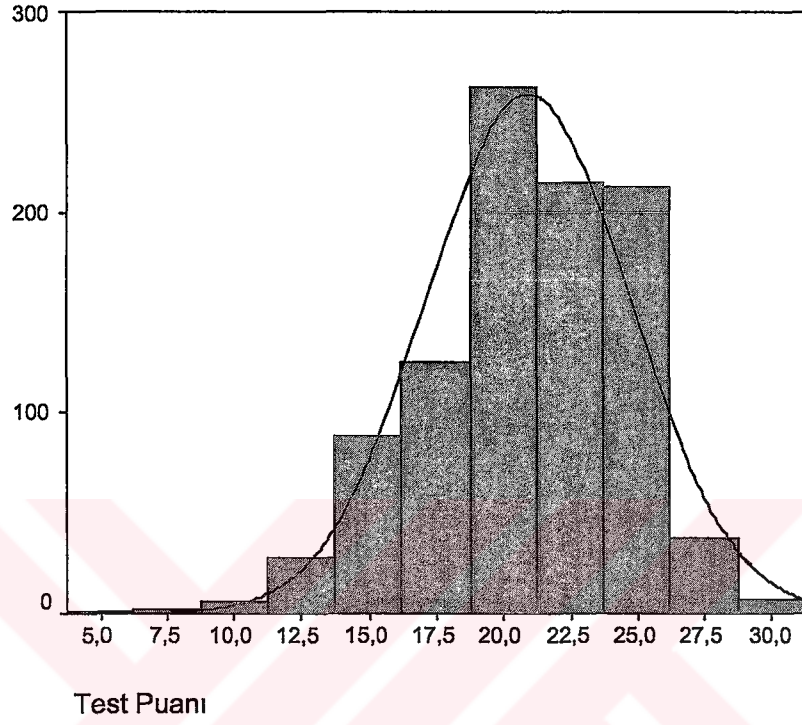
Test	n	Minimum	Maksimum	Ortalama	Std. kayma
Test A (6 Seçenekli)	985	4.00	27.00	16.52	3.95
Test B (4 Seçenekli)	985	6.00	31.00	20.94	3.79

Tablo 1’de görüldüğü gibi A testin ortalaması 16.52 olurken B testin ortalaması 20.94 olmuş, fakat standart kaymalarda pek fazla değişiklik gözlenmemiştir.



Şekil 2: A Testine Ait Puan Grafiği

Şekil 2, A testine ait olan puanların dağılımını göstermektedir. Görüldüğü gibi, en fazla frekanslar yaklaşık 13'le 21 arasındaki puanlara aittir.



Şekil 3: B Testine Ait Puan Grafiği

Şekil 3, B testine ait olan puanların dağılımını göstermektedir. Görüldüğü gibi en fazla frekanslar yaklaşık 18'le 26 arasındaki puanlara aittir. 3. şekil, 2. şekil ile karşılaştırdığında; B testinden bireylerin daha yüksek puanlar elde ettiklerini görebilmek mümkündür.

Tablo 2: A ve B Testlerinin Madde Parametreleri**(a = Ayırt Edicilik , b = Güçlük ve c = Şans Parametresi)**

	Test-A			Test-B		
	1. Alt test			1. Alt test		
NO	aj	bj	cj	aj	bj	cj
1	0.28	-4.45	0.03	0.46	-3.88	0.10
2	0.71	-3.09	0.03	0.37	-2.82	0.10
3	0.54	-2.88	0.03	0.31	-2.53	0.11
4	0.80	-2.12	0.02	0.60	-2.78	0.10
5	0.74	-2.39	0.03	0.53	-4.03	0.10
6	0.59	-0.96	0.03	0.32	-2.08	0.10
7	0.14	3.21	0.04	0.32	6.14	0.03
8	0.52	1.14	0.03	0.47	0.05	0.09
9	0.29	6.15	0.01	1.06	0.86	0.11
10	0.35	1.60	0.03	1.61	0.84	0.35
11	0.20	6.99	0.02	1.09	1.97	0.08
12	0.35	6.62	0.00	0.84	2.74	0.03
13	0.28	6.79	0.01	0.69	4.21	0.02
	2. Alt test			2. Alt test		
1	0.40	2.14	0.04	0.30	-3.88	0.10
2	0.51	-1.63	0.03	0.49	-4.00	0.10
3	0.36	-1.21	0.03	0.62	-2.69	0.10
4	0.52	-1.09	0.03	0.23	-1.08	0.11
5	0.49	1.57	0.01	0.34	-3.25	0.10
6	0.25	-2.20	0.04	0.21	-2.86	0.12
7	0.48	-1.37	0.04	0.18	0.48	0.13
8	0.40	1.47	0.03	0.22	-1.34	0.13
9	0.32	6.87	0.01	0.16	9.38	0.02
10	0.42	4.16	0.01	0.30	5.98	0.06
11	0.18	11.18	0.00	0.50	4.01	0.06
12	0.28	7.80	0.01	0.22	10.14	0.01
13	0.30	8.77	0.00	1.17	2.72	0.03
	3. Alt test			3. Alt test		
1	0.41	-3.12	0.03	1.13	-1.79	0.09
2	0.89	-1.95	0.04	0.96	-2.48	0.09
3	1.07	-0.7	0.03	0.86	-1.61	0.08
4	0.54	-1.26	0.04	0.82	-0.47	0.05
5	0.29	0.81	0.02	0.46	-1.1	0.10
6	1.01	-0.89	0.06	1.22	-0.89	0.10
7	1.19	-0.31	0.05	0.82	-0.6	0.08
8	0.82	1.06	0.08	0.56	-0.43	0.12
9	0.68	2.65	0.02	1.18	-1.14	0.10
10	0.12	14.59	0.00	0.28	1.95	0.11

Tablo 2, A ve B testlerine ait olan madde parametrelerini (a; ayırt edicilik, b; güçlük ve c; şans parametreleri olarak tanımlanır) ayrı ayrı göstermektedir. Bu tablodaki verilere göre; A testinde (a) parametresi en az 0.12 ve en fazla 1.19 olurken, B testinde en az 0.16 ve en fazla 1.61'dir. (b) parametresi de A testinde en az -4.45 ve en fazla 14.49 olurken, B testinde en az -4.03 en fazla 10.14'tür. A testinde (c) parametresi en az 0.00 ve en fazla 0.08 iken, B testinde en az 0.01 ve en fazla 0.35'tir. Görüldüğü gibi her iki testte ve her alt testte, birinci maddeden son maddeye doğru gidildikçe maddelerin güçlük dereceleri artış göstermektedir.



Tablo 3: A ve B Testlerinin Maddelerine Ait olan Maksimum Bilgi Fonksiyonu , Maksimum Bilgi Fonksiyona Sahip Olan Nokta ve Ortalama Bilgi Fonksiyonu

	Madde	Test-A			Test-B		
		Mak.Bilgi Fonksiyonu	Mak. Bilgi Fonk. Sahip Olan Nokta	Ortalama Bilgi Fonksiyonu	Mak.Bilgi Fonksiyonu	Mak. Bilgi Fonk. Sahip Olan Nokta	Ortalama Bilgi Fonksiyonu
1. Alt test	1	0.05	-4.32	0.02	0.12	-3.68	0.03
	2	0.34	-3.05	0.05	0.08	-2.57	0.05
	3	0.20	-2.82	0.06	0.06	-2.21	0.04
	4	0.44	-2.08	0.13	0.22	-2.63	0.06
	5	0.37	-2.35	0.10	0.17	-3.86	0.02
	6	0.24	-0.92	0.17	0.06	-1.79	0.05
	7	0.01	3.50	0.01	0.07	6.25	0.01
	8	0.18	1.20	0.13	0.13	0.24	0.11
	9	0.06	6.20	0.01	0.66	0.96	0.29
	10	0.08	1.69	0.06	0.95	0.98	0.27
	11	0.03	7.08	0.01	0.73	2.04	0.11
	12	0.09	6.63	0.01	0.47	2.78	0.05
	13	0.06	6.82	0.01	0.33	4.24	0.01
2. Alt test	1	0.11	2.24	0.06	0.05	-3.57	0.03
	2	0.18	-1.58	0.11	0.14	-3.82	0.03
	3	0.09	-1.12	0.07	0.23	-2.54	0.07
	4	0.18	-1.02	0.13	0.03	-0.63	0.03
	5	0.17	1.60	0.11	0.07	-2.98	0.04
	6	0.04	-2.05	0.03	0.02	-2.34	0.02
	7	0.16	-1.29	0.11	0.02	1.10	0.02
	8	0.11	1.54	0.08	0.03	-0.83	0.02
	9	0.07	6.90	0.01	0.02	9.54	0.00
	10	0.13	4.18	0.02	0.06	6.18	0.01
	11	0.02	11.20	0.00	0.16	4.13	0.01
	12	0.05	7.83	0.00	0.03	10.18	0.00
	13	0.07	8.78	0.00	0.94	2.74	0.05
3. Alt test	1	0.12	-3.03	0.05	0.78	-1.71	0.20
	2	0.53	-1.90	0.16	0.55	-2.39	0.09
	3	0.77	-0.67	0.43	0.46	-1.52	0.19
	4	0.20	-1.19	0.14	0.44	-0.40	0.30
	5	0.06	0.91	0.05	0.13	-0.91	0.10
	6	0.66	-0.82	0.35	0.90	-0.81	0.42
	7	0.93	-0.26	0.53	0.42	-0.50	0.28
	8	0.41	1.16	0.20	0.18	-0.24	0.15
	9	0.32	2.68	0.06	0.83	-1.06	0.35
	10	0.01	14.61	0.00	0.05	2.32	0.03

Tablo 3, A ve B testlerinin maddelerine ait olan maksimum bilgi fonksiyonu, maksimum bilgi fonksiyonuna sahip olan nokta ve ortalama bilgi fonksiyonunu her test ve alt testlerine g re ayrı ayrı g stermektedir. “Maksimum bilgi fonksiyonu” her maddenin en fazla bilgi fonksiyon miktarını g sterir ve “maksimum bilgi fonksiyonuna sahip olan nokta” bařlıđı ise en fazla bilgi fonksiyonuna sahip olan yetenek d zeyini ifade etmektedir ve “ortalama bilgi fonksiyonu” her maddenin  eřitli yetenek d zeylerinde olan bilgi fonksiyonunun ortalamasını yansıtmaktadır.

Ayrıntılı olarak her yetenek d zeyi  zerinde (“teta”lar  zerinde) hesaplanmış olan madde bilgi fonksiyonları Ek 4’de sunulmuřtur.



Ölçütlerin Tek Boyutluluklarıyla İlgili Olan Analizler ve Sonuçları:

Testlerin tek boyutlu olup olmamasını denemek için faktör analizi sonuçlarına bakılmıştır. Bu sonuçlar aşağıdaki tabloda görülmektedir:

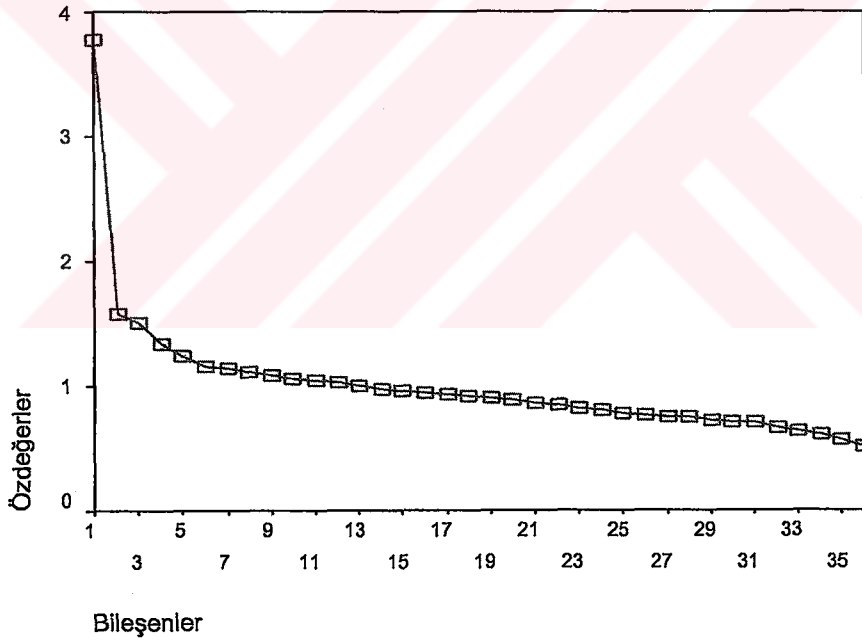
Tablo 4: A ve B Testine Ait Faktör Analiz Sonuçları:

Bileşenler	Öz-değerler					
	A Test			B Test		
	Toplam	Varyans Yüzdesi	Yığılmalı Yüzde	Toplam	Varyans Yüzdesi	Yığılmalı Yüzde
1	3.78	10.49	10.49	3.62	10.05	10.05
2	1.58	4.409	14.89	2.00	5.57	15.61
3	1.51	4.20	19.08	1.71	4.76	20.37
4	1.33	3.70	22.78	1.32	3.67	24.05
5	1.24	3.44	26.23	1.27	3.51	27.56
6	1.16	3.21	29.44	1.23	3.43	30.99
7	1.14	3.18	32.62	1.16	3.23	34.21
8	1.12	3.10	35.72	1.12	3.12	37.33
9	1.09	3.02	38.73	1.10	3.06	40.39
10	1.06	2.94	41.67	1.08	3.00	43.39
11	1.04	2.90	44.58	1.03	2.87	46.27
12	1.03	2.85	47.43	1.02	2.84	49.10

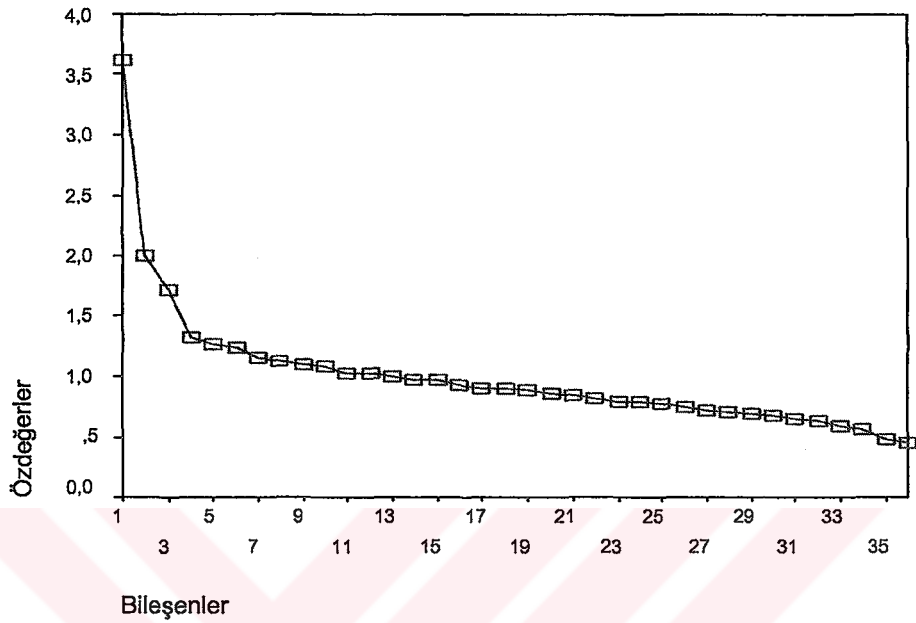
Faktör analizi sonuçlarını yorumlarken iki ölçütün dikkate alındığı gözlenmektedir. Bu ölçütler; özdeğerler ve açıklanan varyanstır. Genellikle ölçeklerin boyut sayısı, birinci özdeğer ile ikinci özdeğerin arasındaki farka veya bileşenlerin açıkladığı varyansa bakılarak kararlaştırılmaktadır. Eğer birinci bileşen özdeğer ile ikinci bileşen özdeğeri arasındaki fark 3 veya 3.5 kat ise ve ikinci bileşen özdeğeri ile sonraki bileşenlerin özdeğerleri birbirine yakınsa, ölçekler tek boyutlu olarak yorumlanmaktadır. Birçok durumda bu işlem, özdeğerlerin azalan eğimini gösteren grafikte, eğimin kaybolduğu ya da küçük olduğu yere kadar olan özdeğerler alınarak boyut sayısını belirleme şeklinde yapılmaktadır.

Tablo 4'teki sonuçlara göre, A testinde birinci bileşen özdeğeri ile ikinci bileşen özdeğeri arasındaki fark yaklaşık 2.4 kat ve B testinde ise yaklaşık 1.8 kattır ve ikinci

bileşen özdeğeri ile sonraki bileşenlerin özdeğerleri birbirine yakındılar. Bu sonuçlara göre, biraz daha riskli olarak her iki testin tek boyutlu olduğunu söylemek mümkündür. Genelde tüm testlere, (özellikle yetenek ve başarı testlerinde) sınav kaygısı, yorgunluk, heyecan gibi çok sayıda değişken karışmaktadır. Bu değişkenlere çoktan seçmeli testlerde şans faktörü de ilave olmaktadır. A ve B testlerinde birinci bileşen özdeğerinin baskın olmaması, karışan değişkenlerin etkisinden kaynaklanabilir. Testlerin tek boyutlarıyla ilgili sonuçlar Şekil 4 ve 5'te özdeğer-bileşen grafiğinde görülmektedir.



Şekil 4: A Testine Ait Özdeğer-Bileşen Grafiği



Şekil 5: B Testine Ait Özdeğer-Bileşen Grafiği

Testin tek boyutlu çıkması, örtük özellikler kuramının varsayımlarından tek boyutluluk ve yerel bağımsızlık varsayımlarının karşılandığını kabul etmek için yeterli bir sonuçtur. Yerel bağımsızlık, tek boyutluluğun bir sonucu olarak görülmektedir. Bu nedenle tek boyutluluk sağlandığında yerel bağımsızlığın da sağlandığı kabul edilmektedir (Lord 1980:19, Hambleton ve Swaminathan 1985:25).

Birçok araştırmacı, tek boyutluluğun sağlandığı durumda yerel bağımsızlığı tespit etmek amacıyla yapılması gereken işlemleri yapmamakta, tek boyutluluğun sağlanmasıyla yerel bağımsızlığın sağlandığını kabul etmektedirler.

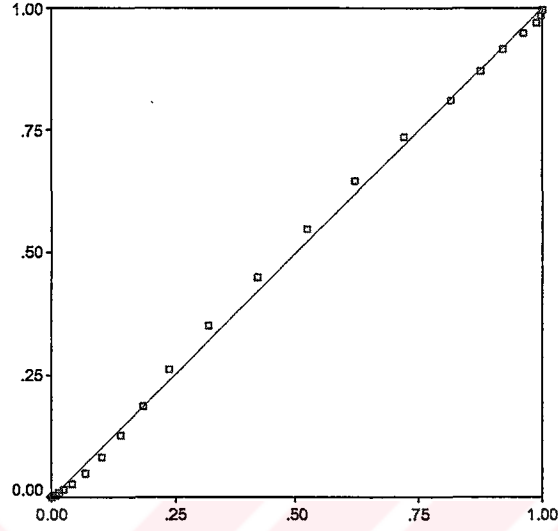
A ve B Test Puanlarının Normal Dağılıp Dağılmadığı Testi

A ve B test puanlarının normal dağılıp dağılmadığı Kolmogorov-Smirnov testi ile test edilmiştir. Tablo 5'te sonuçlar görülmektedir.

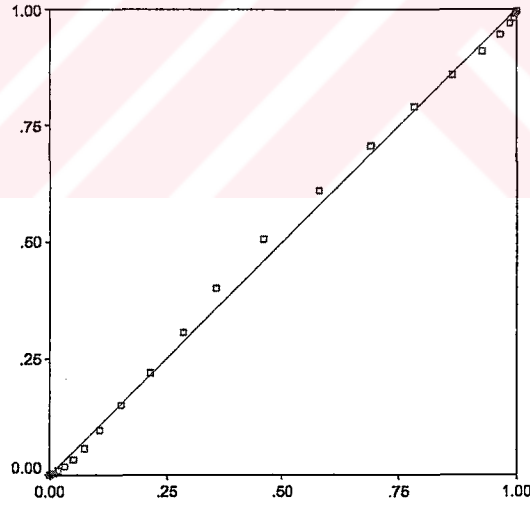
Tablo 5: A ve B testlerinin puan dağılımı için Kolmogorov-Smirnov testi

	D (En Yüksek Fark Değeri)			N	p
	Pozitif	Negatif	Mutlak Değer		
A Test	0.043	-0.082	0.082	985	0.00
B Test	0.045	-0.105	0.105	985	0.00

Kolmogorov-Smirnov testi sonuçları, dağılımın normal olmadığı yönünde yorumlanabilir. Ancak, hem Ki-kare hem de Kolmogorov-Smirnov teknikleriyle normallik testi sonuçları, örneklem büyük olduğunda genellikle manidar sonuçlanmaktadır. Bu testler, beklenen ve gözlenen frekansların farklarını toplama teknikleri olduğundan, N büyüdükçe küçük farkların manidar çıkması mümkündür. Böyle durumlarda grafiksel tekniklerden (P-P veya Q-Q grafikleri) ve betimsel istatistiklerden yararlanılabilir. P-P ve Q-Q grafikleri betimsel olarak verilerin normalliğinin incelenmesinde kullanılan grafiklerdir. P-P grafiği gözlenen ve beklenen verilerin olasılıklarına göre şekillenirken, Q-Q grafiği bu verilen asıl değerlerine göre şekillenmektedir. Aşağıdaki şekiller, A ve B testlerine ait olan normallik P-P grafiklerini yansıtmaktadır.



Şekil 6: A Testine Ait Normallik P-P Grafiği



Şekil 7: B Testine Ait Normallik P-P Grafiği

P-P normallik grafiğinde köşegenden geçen eğri normalliğin göstergesi olmaktadır. Grafikte görülen noktalar ise, çalışmadan elde edilen puanların dağılımının normalliğini göstermektedir. Buna göre, Şekil 6 ve 7 incelendiğinde, her iki testin puanlarının dağılımının normale yakın olduğu görülmektedir. Bu nedenle verilerin dağılımının normal olduğu varsayıp, araştırmaya devam edilmiştir.

BÖLÜM III

BULGULAR VE YORUM

1. Alt Probleme İlişkin Bulgular ve Yorumlar :

Seçenek sayısı arttığında ayırt edicilik gücünde (α_j) ne tür değişiklikler ortaya çıkmaktadır?

İki testin ayırt edicilik gücü dağılımının normal olduğuna yönelik bir işaret olmadığından, bu iki testin (4 ve 6 seçenekli testler) ayırt edicilik güçlerinin karşılaştırılmasında Wilcoxon'un işaretli sıralar testi kullanılmıştır. Bu karşılaştırmaya ilişkin sonuçlar Tablo 6'da verilmiştir.

Tablo 6: A Test (6 seçenekli) ve B Testine (4 seçenekli) Ait Olan Ayırt Edici Güçlerinin Betimsel İstatistikleri ve Wilcoxon Testi İle Karşılaştırılma Sonuçları

Ayırt edici Güçleri	K	Ortalama Sıra	Wilcoxon Z	Anlam düzeyi
A Test	36	14.11	-1.02	0.30
B Test	36	23.41		

Tablo 6'da görüldüğü gibi B test (4 seçenekli test) ve A test (6 seçenekli test) ayırt edicilik gücü arasında anlamlı bir fark yoktur.

Burada elde edilen sonuç daha önce söz ettiğimiz birçok araştırmaların sonuçlarıyla aynıdır. Daha önceki araştırmaların çoğunda da, az seçeneği olan testlerle daha çok seçeneği olan testlerin maddelerinin ayırt edicilik güçlerinin arasında fark bulunamamıştır (Hodson,1983; Steven V. Owen ve Roben D. Froman,1987).

2.Alt Probleme İlişkin Bulgular ve Yorumlar:

Seçenek sayısı arttığında güçlük parametresinde (b_j) ne tür değişiklikler ortaya çıkmaktadır?

Bu alt problemle ilgili olarak, altı seçenekten oluşan test A ve dört seçenekten oluşan test B'nin madde güçlük derecelerinin farkı, birinci alt problemde olduğu gibi Wilcoxon'un işaretli sıralar testi ile incelenmiştir. Burada da iki testin madde güçlük dereceleri dağılımının normal olduğuna yönelik bir işaret olmaması nedeniyle parametrik olmayan Wilcoxon'un işaretli sıralar testi kullanılmıştır. Bununla ilgili sonuçlar Tablo 7'de verilmiştir.

Tablo 7: A Test (6 seçenekli) ve B Testine (4 Seçenekli) Ait Olan Madde Güçlük Derecesinin Betimsel İstatistikleri ve Wilcoxon Test İle Karşılaştırılma Sonuçları

Test	K	Ortalama Sıra	Wilcoxon Z	Anlam düzeyi
A Test	36	20.25	-2.80	0.00
B Test	36	13.09		

Ele gelen Wilcoxon testinin analiz sonuçlarına göre B testinde madde gücü A teste göre ($p= 0.00$) daha düşük bulunmuştur. Bu bulgu ise beklenen bir sonuç olarak değerlendirilebilir. Genelde çoktan seçmeli sorularda madde güçlülüğü iki faktörden etkilenmektedir. Bunlardan birisi maddeyi oluşturan konunun yapısı, diğeri ise maddenin seçenek sayısıdır. Seçenek sayısı arttığında güçlük derecesi de artar. Bunun için B testinde madde güçlük derecesi A testine göre daha da düşüktür.

Bu alt problemde elde edilen sonuçlar, Williams ve Ebel'in (1957) tarafından yapılan araştırmaların sonuçları ile aynı yöndedir. Bu araştırmalarda, az seçeneğe sahip olan

testlerin güçlük derecelerinin, çok seçeneğe sahip olan testlerin güçlük derecelerine göre daha düşük olduğu rapor edilmiştir.



3. Alt Probleme İlişkin Bulgular ve Yorumlar :

Yanlış cevap sayıları farklı yetenek düzeyleri üzerinde eşit midir?

Üçüncü alt problem cümlesi ile ilgili olarak önce bireylerin yetenek düzeyleri (teta düzeyleri), Bilog programını kullanılarak IRT üç parametrelili modelde her testte ayrı ayrı olarak elde edildi ve her bireyin iki testteki yetenek düzeyinin ortalamasına göre sekiz başarı grubu oluşturulmuştur. Bu yetenek grupların sayısının sekiz olmasının nedeni ise normal dağılıma uyumlu olması ve her grubun içindeki bireylerin sayısının normal dağılıma uymasıdır. Böylece bireyler daha çok alt gruplara ayrılmamışlardır. Bunun da nedeni bu gruplarla beklenen sonuçların elde edilmesi ve daha fazla gruba bölünmeye gerek kalmamasıdır. Bu yetenek gruplarıyla ilgili ayrıntılar ve betimsel istatistikler tablo 8’de görülmektedir.

Bu alt problemle ilgili yapılmış diğer bir ön işlem ise, bireylerin her iki testte olan yanlış cevaplarının hesaplanmasıdır.

Tablo 8: İki Teste Olan Yetenek Düzeylerinin Ortalamasına Göre 8 Yetenek Grubuna Yerleşen Bireylerin İki Teste Olan Yanlış Cevap Sayılarının Ortalamasının İstatistiği

Başarı grupları	n	A ve B testin yanlış cevaplarının ortalamasının istatistiği			
		Minimum	Maksimum	Ortalama	Std. Kayma
1 $(-2 > \theta)$	39	8.50	24.00	15.90	4.76
2 $(-1 > \theta \geq -2)$	125	3.00	22.50	12.54	4.34
3 $(-0.5 > \theta \geq -1)$	123	1.00	22.00	10.34	4.01
4 $(-0.0 \geq \theta \geq -0.5)$	169	1.50	16.50	8.44	2.91
5 $(+0.5 \geq \theta > 0.0)$	195	1.00	16.50	7.68	3.07
6 $(+1 \geq \theta > +0.5)$	159	1.00	15.50	7.04	3.10
7 $(+2 \geq \theta > +1)$	148	1.00	14.00	6.52	2.80
8 $(\theta > +2)$	27	1.50	12.50	6.20	2.81

Üçüncü alt problem cümlesine ilişkin olarak, bireylerin iki testteki ortalama yetenek düzeyleri ve iki testteki yanlış cevap ortalamalarının arasındaki ilişki Pearson korelasyon katsayısı yöntemi ile incelenmiştir. Elde edilen korelasyon katsayısı Tablo 9'da görülmektedir:

Tablo 9: Bireylerin A Test ve B Testlerindeki Yanlış Cevaplarının Ortalamaları İle İki Testteki Yetenek Düzeylerinin Ortalamaları Arasındaki Korelasyon Katsayısı

Korelasyon katsayısı	-0.555
P	0.00
n	985

Tablo 9'da gözlenen istatistiksel sonuçlara göre, bireylerin A ve B testlerindeki yanlış cevaplarının ortalamaları ile iki testteki yetenek düzeylerinin ortalamalarının arasındaki korelasyon katsayısı -0.555 olmak üzere, sıfır düzeyinde anlamlıdır ($p= 0.00$). Bu bulgu, yetenek düzeyleri ve yanlış cevap ortalamaları arasında orta seviyede negatif bir ilişki olduğunu göstermektedir.

Bu problem cümlesi ile ilgili bir başka analiz sonucuna göre de, bireyleri A ve B testlerindeki yetenek düzeylerinin ortalamasına göre çeşitli gruplara ayırarak (8 grup) tek yönlü varyans analizi ile her iki testte olan ortalama yanlış cevap sayılarının bu gruplar arasında ortalamalarının farkı incelenmiş ve gruplar arasındaki bu fark anlamlı çıkarak düşük düzeydeki grupların yanlış ortalamaları üst düzeydeki gruplara göre daha yüksek bulunmuştur. Sonuçlar Tablo 10 ve 11'de görülmektedir.

Tablo 10: A ve B Testlerinde Yanlış Cevap Sayılarının Ortalamaları Arasındaki Farkın, Değişik 8 Yetenek Grupta ANOVA ile İncelenmiş Sonuçları

	Kareler Toplamı	sd	Kareler Ortalaması	F	P
Gruplar-arası	5704.81	7	814.97	70.54	0.00
Grup-İçi	11286.95	977	11.55		
Toplam	16991.76	984			

Tablo 11: Yetenek Grupları Arasında, A ve B Testlerinde Ortalama Yanlış Cevap Sayılarının Ortalama Farkı

(I) Yetenek Grubu (J)Yetenek Grubu		Ortalamalar Farkı	P
1.0	2	3.36	0.00
	3	5.56	0.00
	4	7.46	0.00
	5	8.21	0.00
	6	8.85	0.00
	7	9.38	0.00
	8	9.69	0.00
2.0	3	2.20	0.00
	4	4.10	0.00
	5	4.85	0.00
	6	5.49	0.00
	7	6.02	0.00
	8	6.33	0.00
3.0	4	1.90	0.05
	5	2.65	0.00
	6	3.29	0.00
	7	3.82	0.00
	8	4.13	0.71
4.0	5	0.76	0.41
	6	1.40	0.01
	7	1.92	0.00
	8	2.24	0.03
5.0	6	0.64	0.65
	7	1.16	0.04
	8	1.48	0.40
6.0	7	0.52	0.88
	8	0.84	0.94
7.0	8	0.32	1.00

Eğer yanlış cevap sayısını şanstın yararlanma kanıtı olarak varsayarsak, elde edilen sonuçlara göre her ne kadar yetenek düzeyi düşük olursa şanstın yararlanmak o kadar artar ve her ne kadar yetenek düzeyi yüksek olursa şanstın yararlanmak o kadar azalır görülmektedir. Bu yüzden IRT üç parametrelili modelde tüm yetenek düzeylerinde şans parametresinin (C_j) eşit tutulması pek mantıklı görülmemektedir.

Elde edilen sonuçlardan çıkabilecek bir diğer yorum ise “Şans” (Chance) ve “Tahmin” (Guess) terimlerine aittir. Daha üst düzeyde olan bireyler doğru cevaptan bilgisiz oldukları zaman bir kaç seçenek arasında tahminde bulunurlar, fakat aşağı düzeydeki bireyler bu durumda tüm seçenekler arasından rasgele olarak bir seçeneği seçerler ve bu nedenle yüksek düzeyde olan bireylerin yanlış cevap sayıları düşük düzeydeki bireylere göre daha düşük olmaktadır. Bu yüzden “Şans” (Chance) teriminin başarı düzeyine göre daha düşük olan bireyler ve “Tahmin” (Guess) teriminin ise daha yüksek de olan bireyler için kullanılması daha mantıklıdır.



4.Alt Probleme İlişkin Bulgular ve Yorumlar:

Seçenek sayısı arttığında farklı yetenek düzeyleri üzerinde yanlış cevap sayıları değişimleri eşit midir?

Dördüncü alt problem cümlesi için ilk önce bireylerin iki testteki yetenek düzeylerinin ortalaması hesaplanmış ve bu ortalamalara göre bireyler üçüncü alt problemde olduğu gibi sekiz yetenek grubuna bölünerek, her bireyin her iki testte olan yanlış cevaplarını da ele alarak, tekrarlı verilere ait olan iki yönlü varyans analizi kullanılmıştır. Yapılan analizlerle ilgili sonuçlar Tablo 12’de görülmektedir.

TABLO 12: A Test (6 seçenekli) ve B Testine (4 seçenekli) Ait Olan Yanlış Cevap Sayısının Farkı ve Yetenek Düzeylerine Göre Bu Farkın Nasıl Olduğu Varyans Analiz İle İncelenmiş Sonuçları

	Etki	Değer	F	P
Testlerin Yanlış Cevap Sayıları (6 ve 4 Seçenekli Test’lerin Aralarındaki Yanlış Cevap Sayı Farkı)	Pillai’s Trace	0.30	427.30	0.00
	Wilks’ Lambda	0.70	424.30	0.00
	Hotelling’s Trace	0.44	427.30	0.00
	Roy’s Largest Root	0.44	427.30	0.00
Testlerin Yanlış Cevap Sayılarının Farkı *Yetenek Düzey’leri	Pillai’s Trace	0.02	2.25	0.03
	Wilks’ Lambda	0.98	2.25	0.03
	Hotelling’s Trace	0.02	2.25	0.03
	Roy’s Largest Root	0.02	2.25	0.03

Tablo 12’de elde edilen sonuçlara göre iki testin yanlış cevap sayısının farkı $p=0.00$ düzeyinde anlamlıdır. Bu da A testinin daha fazla seçeneğe sahip olmasından ve böylece daha zor olmasından kaynaklanmaktadır.

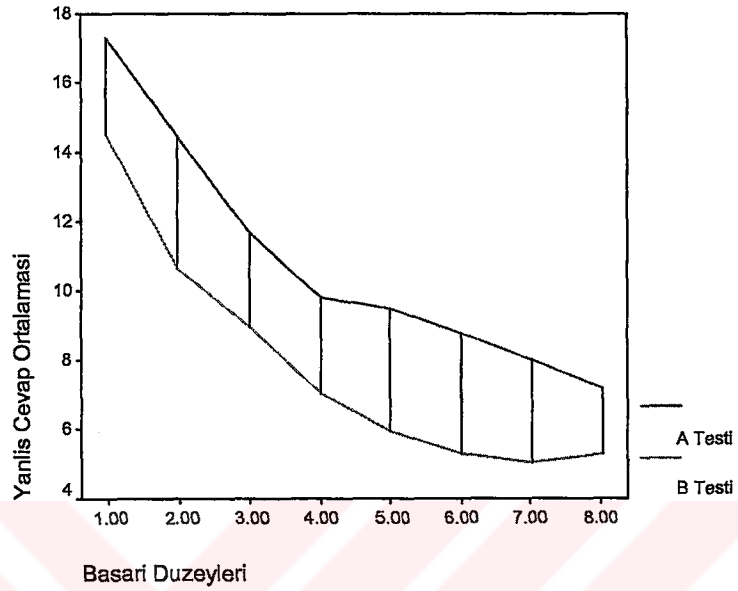
Yukarıdaki analizlerden elde edilen diğer bir sonuç ise, yetenek düzeylerinin, iki testin yanlış cevaplarının arasındaki farka etkisidir. Bu etki ($p=0.03$) düzeyinde anlamlı bulunmuştur.

Tablo 13’te, A ve B testlerinde olan yanlış cevap sayısı farkları ortalamasının yetenek gruplarına göre farklılığın anlamlı olup olmadığına ilişkin sonuçları göstermektedir. Görüldüğü gibi, gruplar tek tek karşılaştırdığında aralarındaki fark anlamlı

çıkılmamaktadır. Fakat Tablo 12’de verilen sonuçlara göre tüm gruplar bir arada incelendiğinde genel olarak gruplar arasında farkın olduğu görülmektedir.

Tablo 13: Yetenek Grupları Arasında, A Testindeki (6 seçenekli test) ve B Testindeki Yanlış Cevap Farkının (A Test Yanlışları – B Test Yanlışları) Değişik Yetenek Gruplarında İncelenme Sonuçları

(I) Yetenek Grubu	(J)Yetenek Grubu	Ortalamalar Farkı (I-J)	P
1.00	2.00	-0.99	0.81
	3.00	0.15	1.00
	4.00	0.03	1.00
	5.00	-0.72	0.95
	6.00	-0.66	0.97
	7.00	-0.13	1.00
	8.00	0.93	0.97
2.0	3.00	1.13	0.21
	4.00	1.02	0.24
	5.00	0.26	1.00
	6.00	0.32	1.00
	7.00	0.86	0.51
	8.00	1.92	0.20
3.0	4.00	-0.11	1.00
	5.00	-0.87	0.42
	6.00	-0.81	0.58
	7.00	-0.27	1.00
	8.00	0.79	0.98
4.0	5.00	-0.76	0.49
	6.00	-0.70	0.66
	7.00	-0.16	1.00
	8.00	0.90	0.93
5.0	6.00	0.06	1.00
	7.00	0.60	0.80
	8.00	1.65	0.33
6.0	7.00	0.54	0.90
	8.00	1.60	0.40
7.0	8.00	1.06	0.86



Şekil 8: Bireylerin İki Testte Olan Yanlış Cevapları Sayısının Yetenek Düzeylerine Göre Yansıtmaktadır

Bu analizlerin sonuçlarına göre, seçenek sayısı arttığında yüksek düzeydeki bireyler düşük düzeydeki bireylere göre daha fazla şanstı yararlanmaya başlıyorlar. Bunun sebebi de muhtemelen testin seçenek sayısının artmasıyla testin daha da zorlaşarak, bireylerin yetenek düzeylerini aşmasıdır. Düşük düzeydeki bireylerin yanlış cevap sayıları seçenek sayısı arttığında eşit seviyede artış gösterir ki bu da bireylerin başarı düzeylerine göre her iki testin madde zorluk derecesinin artması anlamına gelir; ancak yüksek düzeydeki bireyler için az seçeneğe sahip olan test kolaydır ve bu testlerin madde seçeneklerinin sayısı arttığında madde zorluk derecesi bu bireylerde yetenek düzeylerini aşar ve bunun için bunların yanlış cevap sayılarının artışı aşağı düzeydeki bireylere göre daha fazla ortaya çıkar.

Şekil 8'de görüldüğü gibi, bireylerin her iki testte olan yanlış cevapları ve yetenek düzeyleri arasında negatif bir korelasyon vardır. Ayrıca A testinin (6 seçenekli test)

yanlıř cevapları, B testine göre (4 seenekli test) tüm düzeylerde daha fazladır, fakat yüksek düzeylerde bu artış biraz daha fazla görölmektedir.

Bu sonuçlarla ilgili şöyle bir yorum ortaya çıkar:

Seenek sayısı arttığında tüm düzeylerde bireylerin řanstan yararlanmaları artar, fakat yüksek düzeylerde olan bireyler daha fazla řanstan yararlanmaktadırlar. Bunun nedeni ise düşük seviyedeki bireylere hem önceki test (B testi) hem de sonraki test (A testi) zor gelirken, üst düzeydeki bireylere göre B testinde kolay olan soruların A testinde daha da zorlaşmasıdır .

Bu durum (maddelerin zorlaşması) aynı zamanda A testinde madde ayırt edicilik gücünün düşmesine yol açar ve bu da madde bilgi fonksiyonunun azalmasına sebep olur.

5. Alt Probleme İlişkin Bulgular ve Yorumlar:

Seçenek sayısı değiştiğinde madde bilgi fonksiyonu değişir mi? Bu etki farklı yetenek düzeyleri üzerinde değişir mi?

Beşinci alt problem cümlesi ile ilgili yapılan analizler doğrultusunda ilk önce her iki testin üzerinde madde bilgi fonksiyonları , $\theta = -12$ 'den $\theta = 0.00$ 'a ve $\theta = 0.00$ 'dan $\theta = 12$ 'ye yani 49 yetenek düzeyi veya 49 teta (θ) üzerinde hesaplanmıştır. Yetenek düzey sayısının fazla olmasının nedeni, seçenek sayısının yetenek düzeyleri arasında madde bilgi fonksiyonları üzerinde olan etkisinin farkının az olmasının beklentisidir.

Böylesine her iki testte 1764 kere ($36 \text{ madde} \times 49 \text{ yetenek düzeyi} = 1764$) madde bilgi fonksiyonu hesaplanmış ve iki testin bu madde bilgi fonksiyonları ortalamalarının farkı t-testi ile incelenmiş ve anlamlı bir fark ortaya çıkarak B testinin madde bilgi fonksiyonu A testine göre daha yüksek bulunmuştur. Bu sonuçlarla ilgili istatistiksel bilgiler Tablo 14'te görülmektedir.

Tablo 14: A Ve B Testlerinin Madde Bilgi Fonksiyonlarının Ortalamasının T-Testi İle İncelenmiş Sonuçları

Madde bilgi fonksiyonu	Ortalama	n	Std.Kayma	t	sd	Anlam düzeyi
A Test	0.0189	1764	0.0724	-2.92	1763	0.004
B Test	0.0257	1764	0.0987			

Yukarıdaki analizden sonra 4 ve 6 seçenekli testlerde yetenek düzeyleri arasında bilgi fonksiyonun farkı incelenmiş ve ortaya çıkan sonuçlara göre başarı düzeyleri (θ 'lar düzeyleri) arasında da bilgi fonksiyonu sıfır düzeyinde ($p = 0.00$) farklı bulunmuştur.

Tablo 15: Yetenek Düzeyler Arasında Madde Bilgi Fonksiyonları Farklı Olup Olmamasının Her İki Teste (Ayrı Ayrı) Tek Yönlü Varyans Analiz İle Yapılmış İncelemenin Sonuçları

Madde Bilgi Fonksiyonu		Kareler Toplamı	sd	Kareler Ortalaması	F	p
A-Test (6 seçenekli)	Gruplar-arası	1.475	48	.031	6.79	.000
		7.755	1715	.005		
	Grup-İçi Toplam	9.230	1763			
B-Test (4 seçenekli)	Gruplar-arası	2.178	48	.045	5.18	.000
		15.003	1715	.009		
	Grup- İçi Toplam	17.180	1763			

Tablo 15'ten gelen sonuçlara göre her iki teste yetenek düzeyler arasında madde bilgi fonksiyonun farkı sıfır düzeyde ($p=0.00$) anlamlıdır. Genelde B testinin bilgi fonksiyonu ortalaması A testine göre daha yüksek bulunmuş ($p= 0.00$), fakat bu fark yetenek düzeylerinden (θ) etkilenmemiştir yani seçenek sayısının madde bilgi fonksiyonu üzerinde olan etkisi yetenek düzeylerine göre değişmemiştir. Tablo 16'da bu sonuçlar verilmiştir.

Tablo 16:A Test (6 Seçenekli) ve B Testine (4 Seçenekli) Ait Olan Madde Bilgi Fonksiyon Farkı ve $\theta = 12$ 'den $\theta = -12$ 'ye Olan Yetenek Düzeylerine Göre Bu Farkın Nasıl Olduğu Varyans Analiz İle İncelenmiş Sonuçları

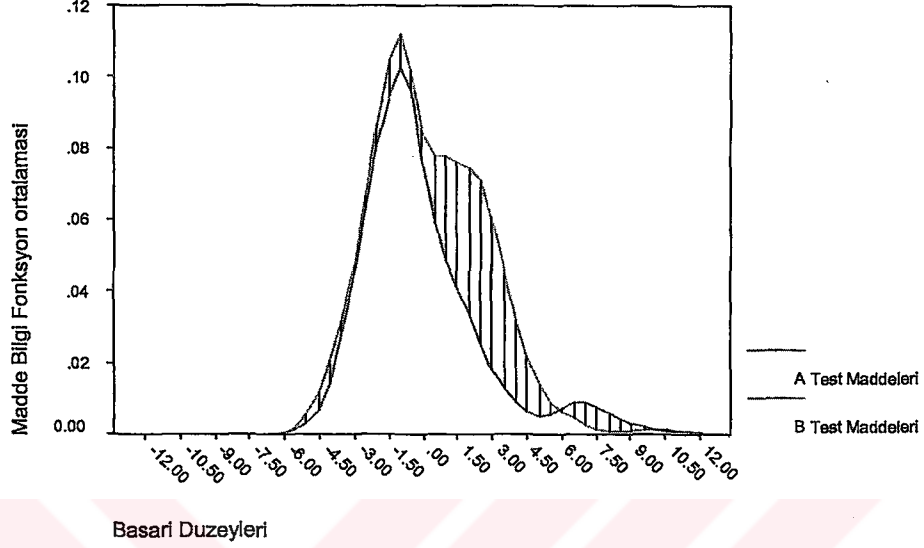
	Etki	Değer	F	P
Test (6 ve 4 seçenekli test'lerin aralarındaki bilgi fonksiyon Farkı)	Pillai's Trace	0.01	8.42	0.00
	Wilks' Lambda	0.99	8.42	0.00
	Hotelling's Trace	0.01	8.42	0.00
	Roy's Largest Root	0.01	8.42	0.00
Test*Yetenek Düzey'leri (6 ve 4 seçenekli test'lerin aralarındaki bilgi fonksiyon Farkı *Başarı Düzeyleri)	Pillai's Trace	0.02	0.65	0.97
	Wilks' Lambda	0.98	0.65	0.97
	Hotelling's Trace	0.02	0.65	0.97
	Roy's Largest Root	0.02	0.65	0.97

Diğer bir analizde sadece $\theta = -3$ ’ den $\theta = -12$ ‘e ve $\theta = 3$ ’ten $\theta = 12$ ’ye olan yetenek düzeylerini aldığımızda, seçenek sayısının madde bilgi fonksiyonu üzerinde olan etkisi yetenek düzeylerine göre farklı bulunmuştur.

Tablo 17: A Test (6 seçenekli) ve B Testine (4 seçenekli) Ait Olan Madde Bilgi Fonksiyon Farkı ve $\theta = -3$ ’ten $\theta = -12$ ’ye ve $\theta = 3$ ’ten $\theta = 12$ ’ye Olan Yetenek Düzeylerine Göre Bu Farkın Nasıl Olduğu Varyans Analiz İle İncelenmiş Sonuçları

	Etki	Değer	F	P
Test (6 ve 4 seçenekli test’lerin aralarındaki bilgi fonksiyon Farkı)	Pillai’s Trace	0.01	6.04	0.01
	Wilks’ Lambda	0.99	6.04	0.01
	Hotelling’s Trace	0.01	6.04	0.01
	Roy’s Largest Root	0.01	6.04	0.01
Test*Başarı Düzey’leri (6 ve 4 seçenekli test’lerin aralarındaki bilgi fonksiyon Farkı *Başarı Düzeyleri)	Pillai’s Trace	0.04	1.57	0.02
	Wilks’ Lambda	0.96	1.57	0.02
	Hotelling’s	0.04	1.57	0.02
	Trace	0.04	1.57	0.02
	Roy’s Largest Root			

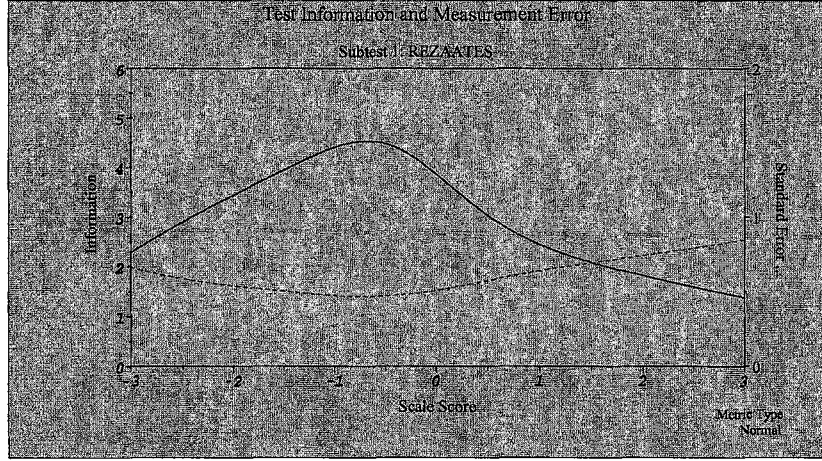
Yukarıdaki tabloda bulunan sonuçlarla ilgili, “seçenek sayısının madde bilgi fonksiyonları üzerinde etkisi, yetenek düzeylerine göre değişir” ifadesini kullanmak mümkündür. Fakat bu etki daha çok $\theta \geq 0.00$ düzeylerde bulunur yani bu çalışmada iki testin madde seçenek sayılarının farkı daha çok yüksek başarı düzeylerinde madde bilgi fonksiyonlarını etkiler. Aşağıdaki grafik bu ifadeyi yansıtmaktadır.



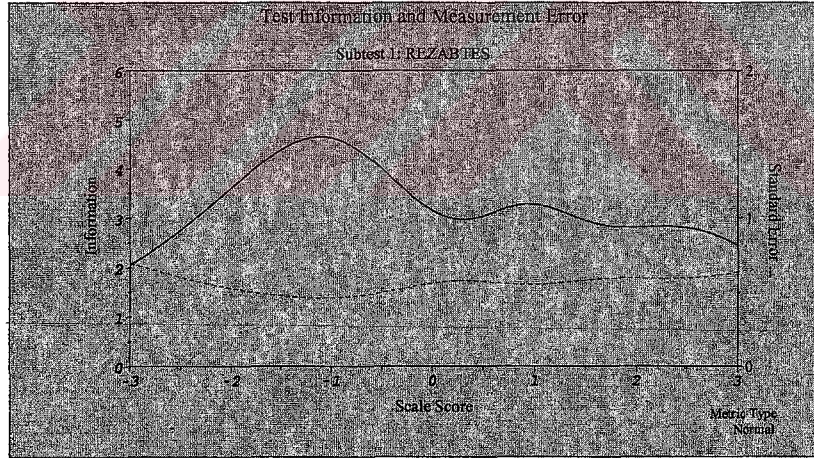
Şekil 9: A ve B Testlerinin Madde Bilgi Fonksiyonlarının Ortalamasını Yetenek Düzeylerine Göre Yansıtması

Şekil 9’da görüldüğü gibi, her iki testin madde bilgi fonksiyon ortalamalarının en yüksek olduğu nokta -1.5 düzeylere ait olurken, -6.00 ve 10.5 ’lerde sıfıra yaklaşmıştır. Bu şekilde görüldüğü gibi B testinin A testine göre $\theta = -6.00$ ’ dan $\theta = 0.5$ ’e kadar fazla üstünlüğü yok ve $\theta > 0.5$ ’ten $\theta = 6.00$ ’ya bu üstünlük daha fazladır ve daha sonraki tetalar düzeyinde ($\theta \geq 6.00$ ’larda) A testi üstün görünmektedir. Bunun nedeni ise A testinin güçlük derecesinin daha da yükselmesidir.

Demek ki, madde seçenek sayısı artarken (yani madde zorlaşırken) test bilgi fonksiyonu yüksek tetalarda artar ve düşük düzeylerde düşer. Ayrıca seçenek sayısı azalırken (yani madde kolaylaşırken) test bilgi fonksiyonu düşük tetalarda artar ve yüksek düzeylerde de düşer.



Şekil 10: A Testine Ait Başarı Düzeyleri Üzerinde Bilgi Fonksiyon Grafiği



Şekil 11: B Testine Ait Başarı Düzeyleri Üzerinde Bilgi Fonksiyon Grafiği

Şekil 10 ve 11, sırayla A testinin (6 seçenekli) ve B testinin (4 seçenekli) yetenek düzeyleri üzerinde bilgi fonksiyon değerlerini göstermektedir. Şekil 9 için yapılan yorumlar burada da yapılabilir. Bu grafiklerde, B testi bilgi fonksiyonunun, A testine göre bir çok düzeyde üstünlüğü görülür fakat bu üstünlük üst düzeylerde daha da fazla görülmektedir. Ölçme teorilerinde ayırt edicilik gücü ile madde bilgi fonksiyonunun arasında bir korelasyon varsayılır. Aşağıdaki formül bu korelasyonu göstermektedir:

$$I(\theta) = \frac{P_j^*(\theta)^2}{P_j(\theta)Q_j(\theta)}$$

(Embretson .Susan E & Reise Steven P 2000: 183) burada $I(\theta)$; Belirli bir teta üzerinde madde bilgi fonksiyonu, $P_j^*(\theta)$ madde ayırt edicilik gücü ve $P_j(\theta)Q_j(\theta)$ ise madde varyansı'dır. Ayrıca bu teorilere göre madde ayırt edici gücü ile madde güçlük derecesi arasında da bir orta seviyede korelasyon vardır ve buda aynı zamanda madde bilgi fonksiyonunu etkilemektedir.

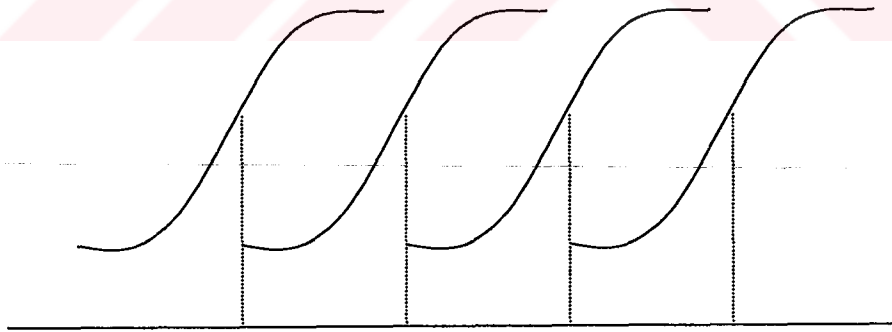
Madde güçlük derecesi her hangi bir teta üzerinde orta seviyede olursa, yani $(b_j - \theta) = 0$, bu teta üzerinde en fazla ayırt edicilik gücü ve nihayet en fazla bilgi fonksiyonuna sahip olacaktır. Bir başka deyişle, madde güçlük derecesi fazla basit veya fazla zor olursa, o maddenin ayırt edicilik gücü düşer ve bu da madde bilgi fonksiyonunu düşürür.

Bu çalışmada, hem B testinin A testine göre ayırt edicilik gücünün fazla olması hem de B testinin madde bilgi fonksiyon ortalamasının A testine göre $\theta = 6.00$ 'ya kadar üstün olması ve A testinin $\theta > 6.00$ 'larda daha fazla olması, yukarıda sözü edilen teorilerle oldukça uyumlu sonuçlardır. Bunun nedenini şöyle söylemek mümkündür; madde seçenek sayısı arttığında güçlük derecesi artar ve bu da orta ve düşük düzeydeki bireylerin soruları rasgele cevaplamalarına sebep olur ve madde ayırt edicilik gücünün düşmesine yol açar. Dolayısıyla madde bilgi fonksiyonu düşük düzeylerde düşmeye başlar, ancak üst düzeylerde madde güçlük derecesi başarı düzeyleri ile uyum sağladığı için ayırt edicilik gücü artar ve bu da madde bilgi fonksiyonunun yükselişine sebep olur.

Tablo 18: A ve B Testlerinin Madde Ayırt Edicilik Güçleri İle Güçlük Derecelerinin Arasındaki Korelasyon Katsayısı

	Ayırt edici gücü (Test A)	Ayırt edici gücü (Test B)
Zorluluk derecesi (Test A)	-0.537	--
Zorluluk derecesi (Test B)	--	-0.09

Bu korelasyonlara bakarken A testinde madde ayırt edicilik gücü ile madde güçlük derecesi arasında orta seviyede negatif bir korelasyon görülmektedir. Ancak bu korelasyon, B testinde sıfır seviyesindedir ve bunun nedeni ise A testinde madde güçlük derecesinin bireylerin başarı düzeylerini aşmış, bunun için madde güçlük derecesi arttığı halde madde ayırt edicilik gücünün düşmüş ve orta seviyede negatif korelasyon ortaya çıkmış olmasıdır. Fakat B testinde, madde güçlük derecesi bireylerin başarı düzeyleri ile uyumlu olduğu için bu iki parametre arasında sıfır korelasyon görülmektedir, yani madde ayırt edicilik gücü madde güçlük derecesinden bağımsızdır ve bu da yukarıdaki teorilerle uyumludur.

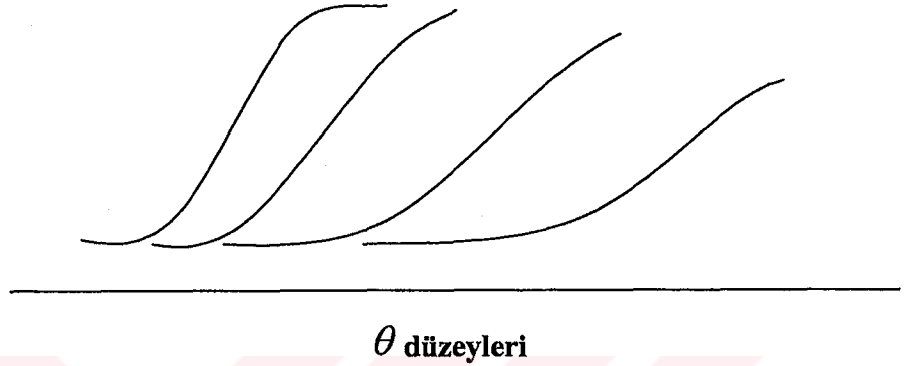


θ düzeyleri

Şekil 12: Teorik Olarak Eşit Ayırt Edicilik Gücüne Sahip Olan ve Değişik Güçlük Derecelerini Yansıtan Maddeleri Gösterir. Burada Madde Ayırt Edicilik Gücü İle Güçlük Derecesi Arasında Hiç Bir İlişkinin Olmadığı Söz Konusudur

B testinin madde zorluklarıyla ayırt edicilik güçleri arasında hiç bir ilişki söz konusu değildir. Üstteki grafik bu ifadeyi anlatmaktadır. Bir başka deyişle, madde ayırt edicilik

gücünün madde güçlük derecesinden bağımsız olduğunu göstermektedir. (Madde ayırt edicilik gücü eşit, ancak güçlük derecesi değişik bulunur)



Şekil 13: Teorik Olarak Değişik Ayırt Edicilik Gücü ve Değişik Güçlük Derecelerine Sahip Olan Maddeleri Gösterir. Burada Madde Ayırt Edicilik Gücü İle Güçlük Derecesi Arasında Bir İlişkinin Olduğu Söz Konusudur

A testinin madde zorluklarıyla ayırt edicilik güçleri arasında -0.537 miktarında bir negatif korelasyon söz konusudur. Üstteki grafik bu ifadeyi anlatmaktadır. Bir başka deyişle, madde zorluk derecesi arttığında madde ayırt edicilik gücün düştüğünü göstermektedir.

Klasik Test Kuramına dayalı ilgili araştırmaların bazılarında az seçeneğe sahip olan testlerin güvenilirliği ve iç tutarlılığı çok seçeneğe sahip olan testlere göre daha da fazla bulunmuştur (Costin,1970; Zimmerman ve Humphries,1953; Williams ve Ebel,1957). Bu demektir ki, az seçeneğe sahip olan testlerin ölçme hatası çok seçeneğe sahip olan testlere göre daha da düşüktür.

Bu alt problemle ilgili çıkan sonuç, normal düzeylerde dört seçenekli testin madde bilgi fonksiyonunun altı seçeneğe sahip olan teste göre daha fazla olduğunu göstermektedir. Bu nedenle dört seçenekli testlerin normal düzeylerde altı seçenekli testlere göre daha da az ölçme hatasına sahip olduğu söylenebilir. Bu yüzden bu sonucun daha önceki

Klasik Test Kuram'ına dayalı olarak yapılan arařtırmaların sonuçlarıyla aynı yönde olduđunu söylemek mümkündür.



6. Alt Probleme İlişkin Bulgular ve Yorumlar:

Seçenek sayısı değiştiğinde madde uyumluluğunda ne gibi değişiklikler ortaya çıkmaktadır?

Altıncı Alt problem cümlesini incelemek için her iki testin maddelerinin uyumluluğu, üç parametrelili IRT modelinde incelenmiş ve uyumlu ve uyumsuz olan maddeler uyum düzeyleri ile birlikte belirlenmiştir. Daha sonra her uyumlu maddeye 1 puan verilerek tekrarlı verilere ait olan iki yönlü varyans analizi ile incelenmiş ve betimsel bilgiler Tablo 19 ve analizlerin sonuçları Tablo 20’de görülmektedir.

Tablo 19: A ve B Testlerin Uyumlu Madde Sayısını Üç ve İki Parametrelili IRT Modelde Göstermektedir. A ve B Testleri Toplam 36 Şer Maddeye Sahiptirler

Test	A Test		B Test	
IRT Modelleri	Uyumlu	Uyumsuz	Uyumlu	Uyumsuz
Üç Parametrelili	29	7	25	11
İki Parametrelili	26	10	24	12

Tablo 19’de görüldüğü gibi, A testinin uyumlu madde sayısı B testine göre daha fazladır. Fakat analiz sonuçları bu iki testin uyumlu madde sayısının farkının anlamlı olmadığını gösterir.

Tablo 20: Tekrarlı Datalara Ait Olan İki Yönlü Varyans Analizini Kullanarak A ve B Testlerinin Uyumlu Madde Sayısının Üç ve İki Parametrelili IRT Modelde Hem Testler Hem Modeller Arasında Fark, Hem de Test ve Modellerin Karşılıklı Etkisi İncelenmiş Sonuçları

Madde Uyumluluk	Etki	Değer	F	P
A ve B Testleri arasında	Pillai's Trace	0.02	0.78	0.38
	Wilks' Lambda	0.98	0.78	0.38
	Hotelling's Trace	0.02	0.78	0.38
	Roy's Largest Root	0.02	0.78	0.38
3 ve 2 Parametrelili arasında	Pillai's Trace	0.04	1.63	0.21
	Wilks' Lambda	0.96	1.63	0.21
	Hotelling's Trace	0.05	1.63	0.21
	Roy's Largest Root	0.05	0.63	0.21
A ve B Testleri* 3 ve 2 Parametrelili	Pillai's Trace	0.01	0.28	0.60
	Wilks' Lambda	0.99	0.28	0.60
	Hotelling's Trace	0.01	0.28	0.60
	Roy's Largest Root	0.01	0.28	0.60

Tablo 20'den elde edilen sonuçlara göre A (6 seçenekli) ve B (4 seçenekli) testlerin uyumlu madde sayılarında anlamlı bir fark yoktur. Ayrıca iki model (3 ve 2 parametrelili modeller) arasında da uyumlu maddeler açısından anlamlı bir değişiklik bulunmamıştır. Bir başka sonuç, modeller değiştiğinde iki testin uyumlu madde sayılarında eşit değişiklik bulunmasıdır.

Diğer bir analizde iki testin uyumlu madde sayısının farkı, her modelde (3 ve 2 parametrelili modellerde) ayrı ayrı Ki-kare ve McNemar testi ile incelenmiş ve Tablo 21 ve 22'deki sonuçlar ortaya çıkmıştır.

Tablo 21: A (6 seçenekli) Test ve B (4 seçenekli) Testin, 3 ve 2 Parametrelili Modellerde Belirlenen Uyumlu Madde Sayılarının Farkı Ki-kare Testi İle İncelenmiş İstatistiksel Sonuçlarını Göstermektedir

	3 parametrelili model			2 parametrelili model		
	Değer	Sd	P	Değer	sd	P
Pearson Chi-Square	1.19	1	0.28	0.262	1	0.609
Continuity correction	0.67	1	0.41	0.065	1	0.798
Likelihood Ratio	1.20	1	0.28	0.262	1	0.609
Fisher's Exact Test						
Linear-by-Linear						
Association	1.17	1	0.28	0.258	1	0.611
N of valid Cases	72			72		

Yukarıda da sözü edildiği gibi, A testinin uyumlu madde sayısı her iki modelde B testine göre daha fazladır. Ancak bu fark, Ki-kare testi ile incelendiğinde anlamlı olmadığı ortaya çıkmıştır. Tablo 21, Ki-kare test sonuçlarını göstermektedir.

Bir başka analizde uyumlu maddelere 1 puan ve uyumsuz maddelere 0 verilerek parametrik olmayan McNemar testi uygulandı . Tablo 22, bu analiz sonuçları göstermektedir.

Tablo 22: A (6 seçenekli) Testi ve B (4 seçenekli) Testinin 3 ve 2 Parametrelili Modellerde Belirlenen Uyumlu Madde Sayılarının Farkını McNemar Testi İle İncelenmiş İstatistiksel Sonuçlarını Göstermektedir

IRT Modelleri	Test	Uyumlu madde sayısı	Uyumsuz madde sayısı	N	Anlam düzeyi
3 Parametrelili model	A Test	29	7	36	0.45
	B Test	25	11	36	
2 Parametrelili model	A Test	26	10	36	0.79
	B Test	24	12	36	

Elde edilen sonuçlara göre iki testin uyumlu madde sayılarında anlamlı bir fark bulunmadı.Yani seçenek sayısının madde uyumluluğu üzerinde etkisi yok.Fakat şunu söylemek mümkündür ki, bu çalışmada iki testin seçenek sayılarının arasında sadece 2 fark bulunmaktadır. Yani, eğer seçenek sayısının farkı daha da fazla olursa başka sonuçların ortaya çıkması mümkündür.

Klasik Test Kuramı'na göre medde uyumluluđu; o maddenin tüm testle olan korelasyonuna bakılarak belirlenir. Frank Costin (1972), Steven V. Owen, Roben D. Froman (1987), Ramos ve stern (1973) tarafından yapılan arařtırmaların sonucuna göre az seçeneđe sahip olan testlerle daha çok seçeneđe sahip olan testlerin maddeleri ile tüm testle olan korelasyon katsayılarının ortalamalarının arasında anlamlı fark bulunamamıştır. Bu nedenle altıncı alt problemin incelenmesiyle elde edilen sonuçla, daha önceki arařtırmaların sonuçları arasında çok fark yoktur.



BÖLÜM IV

SONUÇ VE ÖNERİLER

Sonuçlar

Birinci alt probleme ilişkin elde edilen sonuç, A testi (altı seçenekli) ile B testinin (dört seçenekli) ayırt edicilik güçleri arasında farkın olmadığını göstermiştir.

İkinci alt problem cümlesinin incelemesinden ele gelen sonuç şöyledir: yetenek testlerinde seçenek sayısı arttığında test daha da zorlaşır

Üçüncü alt problemle ilgili sonuç; yetenekler arasında yanlış cevap sayılarının eşit olmaması yönündedir. Bir başka değişle, şans parametresi farklı yetenek düzeylerinde değişmektedir.

Dördüncü alt problemin incelenmesi sonucunda; seçenek sayısı değiştiğinde yanlış cevap sayısı, farklı yetenek düzeylerinde eşit olmadığı görülmektedir. Bu sonuca göre; az seçenekli testte, diğer yetenek düzeyindeki bireylere göre daha az yanlış cevap sayısına sahip olan yukarı düzeydeki bireyler seçenek sayısı arttığında yanlış cevap sayısı, diğer bireylere göre daha fazla artış göstermektedirler.

Beşinci alt problemle ilgili sonuç; seçenek sayısının fazlasıyla artırılması ve bu artışın bireylerin yetenek düzeylerini aşması durumunun, maddelerin bilgi fonksiyonlarının düşmesine sebep olduğu yönündedir. Bu alt probleme ilişkin diğer bir sonuç ise, madde seçenek sayısı değiştiğinde, madde bilgi fonksiyonları farklı yetenek grupları arasında eşit olarak değişmemektedir.

Altıncı alt problemin incelenmesinden elde edilen sonuç; maddedeki seçenek sayısının değiştirilmesinin, madde uyumluluğu üzerinde (IRT 3 ve 2 parametrelili modellerde) anlamlı bir etkisinin olmadığı yönündedir.

Öneriler

Bu çalışmayla ilgili olan bazı öneriler şöyledir:

- 1- Normal yetenek düzeylerinde 4 seçenekli testlerin, 6 veya daha fazla seçeneğe sahip olan testlere göre tercih edilmesi önerilebilir.
- 2- En yüksek yetenek düzeyindeki bireyleri ayırt etmek için daha fazla seçeneğe sahip olan testler önerilebilir.
- 3- Seçenek sayısındaki artışın, testin bilgi fonksiyonuna veya maddelerin normal düzeyde ayırt edicilik gücüne katkısı yoktur. Bunun için seçenek sayısının arttırılması yerine daha yüksek bilgi fonksiyonu ve ayırt edicilik gücüne sahip olan maddelerin teste ilave edilmesi önerilebilir.
- 4- Benzer çalışmalar daha farklı sayıda seçeneğe sahip olan testlerle yapılabilir. Örneğin 3 ve 6 veya 2 ve 5 seçenekli vs... testlerle.
- 5- IRT üç parametrelili modellerde, şans parametresinin yetenek düzeylerine göre ayrı ayrı hesaplanıp ele alınması önerilebilir.

KAYNAKÇA

- AIKEN, Lewis R. Psychological Testing and Assessment. Tenth edition, *Allyn and Bacon, Inc.*, Boston, 2000.
- BAKER, Frank. The Basic of Item Response Theory. *ERIC Clearinghouse on Assessment and Evaluation*, 2001.
- BAYKUL, Yaşar. Örtük Özellikler ve Klasik Test Kuramları Üzerine Bir Karşılaştırma. (Yayımlanmamış Doktora Tezi), Ankara: Hacettepe Üniversitesi, 1979.
- BAYKUL, Yaşar. Eğitimde ve Psikolojide ölçme : Kalsik test teorisi ve uygulaması. *ÖSYM Yayınları*, Ankara 2000.
- BIELINSKI, John; Martha Thurlow, Jane Minnema, Jim Scott. "How Out-of-Level Testing Affects the Psychometric Quality of Test Scores." 2000.
- COSTIN, Frank. Three –Choice Versus Four-Choice Items: Implications for Reliability and Validity of Objective Achievement Test. *Educational and Psychological Measurement*, 1972, 32, 1035-1038.
- CROCKER, Linda. ; Algina, James. Introduction to classical & modern test theory. *Holt, Rinehart and Winston, Inc.* University of Florida 1986.
- EMBRETSON, Susan E.; Riese. Steven P. Item response theory for psychologists. *Lawrence erlbaum associates, publishers, mahwah*, New jersey, London 2000.
- FAN, Xitao. Item Response Theory and Classical Test Theory: An Empirical Comparison of Their Item-Person Statistics. *Educational and Psychological Measurement*, 1998, 58,3,357-381.
- GELBAL. Selahettin. P Madde Güçlük İndeksi ile Rash Modelinin Bir Parametresi ve Bunlarla Dayalı Yetenek Ölçüleri Bir Karşılaştırma. (Yayımlanmamış Doktora Tezi), Ankara: Hacettepe Üniversitesi, 1994.

- GREEN, R., Yen, W. M., & Burket, G. R. Experiences in the applications of item response theory in test construction. *Applied Measurement in Education*, 2, 297-312. 1989.
- GULLIKSEN, Harold . Theory of mental tests. *John Wiley & Sons, Inc.* New York 1967.
- HALADYNA, Thomas. The Effectiveness of Several Multiple-Choice Formats. 1992.
- HAMBLETON, R. K. Principles and selected applications of item response theory. In R. L. Linn (Ed.), *Educational measurement (3rd ed., pp. 147-200)*. New York: American Council on Education, Macmillan 1989.
- HAMBLETON, R. K. and R. W. JONES. Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. *Educational measurement: Issues and Practices*, 1993, 12, 3, 38-47.
- HAMBLETON, R. K. and H. SWAMINATHAN and J. H. ROGER. Fundamentals of Item Response Theory. *Sage Publishing*, Boston, 1991.
- HAMBLETON, R. K. and H. SWAMINATHAN. Item Response Theory: Principles and Application. *Kluwer-Nijhoff Publishing*, Boston, 1985.
- HARVEY, R. J. and A. L. HAMMER. Item Response Theory ; MYERS-Briggs Type Indicator. *Counseling Psychologist*, 1999, 27, 3, 353-374.
- HENRYSSON, Sten. "Gathering, Analyzing and Using Data on Test Items." R. L. Thorndike (Ed), *Educational Measurement*, Second Edition, New York, 1971.
- HOOMAN, Haidar Ali. "*Educational and psychological Measurements.*" *Parsa publishing*, Tehran 1994.
- KELECİOĞLU, Hülya. Örtük Özellikler Teorisindeki b ve a Parametreleri ile Klasik Test Teorisindeki p ve r İstatistikleri Arasındaki ilişki. H. Ü. Eğitim Fakültesi Dergisi, 20, 104-110.

- LORD, Fredric M. Application of item response theory to practical testing problems. *Lawrence Erlbaum associates publishers*, Hillsdale, New jersey 1980.
- LORD, Fredric M.; Novick, Melvin R. Statistical theories of Mental test scores. *Addison-Wesley publishing company*, Inc, New York University 1968.
- LUDLOW, H. Larry. The Structure of The Job Responsibilities Scale: A Multimethod Analysis. *Educational and Psychological Measurement*, 1999, 59, 6, 962-973.
- LYNCH, A, Marie. The Application of Item Response Theory to Employee Attitude Survey Data Using Samejima's Graded Response Theory. (Yayınlanmamış Doktora Tezi), Connecticut Üniversitesi, 1997.
- MEHRENS, Wm; Maihoff, N.A. A Comparison of Alternate-Choice and True-False Item Forms Used in Classroom Examinations. 1985.
- MISILEVY, R. J. and R. D. BOCK. Bilog 3: Item Analysis and Test Scoring with Binary Logistic Models. Second Edition, *Scientific Software, Inc.*, Chicago, 1990.
- NARTGÜN, Zekeriya. Aynı Tutumu Ölçmeye Yönelik Likert Tipi Ölçek İle Metrik Ölçeğin Madde ve Ölçek Özelliklerinin Klasik Test ve Örtük Özellikler Kuramına Göre İncelenmesi. (Yayımlanmamış Doktora Tezi), Ankara:Hacettepe Üniversitesi, 2002.
- OWENS, J. Anne. Using Brand Loyalty Scale to Compare Item Selection Criteria and Scale Evaluation Techniques of Classical Test Theory and Item Response Theory. (Yayımlanmamış Doktora Tezi), Illinois: Southern Illinois University, 2000.
- OWEN, V. Stiven. and R. D. FROMAN. What's Wrong With Three-Option Multiple Choice Item?. *Educational And Psychological Measurement*, 1987, 47.
- ÖZDEMİR, Devrim. Çoktan Seçmeli Testlerin Klasik Test Teorisi ve Örtük Özellikler Teorisine Göre Hesaplanan Psikometrik Özelliklerinin İki Kategorili ve Ağırlıklandırılmış Puanlanması Yönünden Karşılaştırılması. (Yayımlanmamış Doktora Tezi), Ankara: Hacettepe Üniversitesi, 2002.

- PARK, Hyunjeong. Comparison of IRT Models for Ordered Polytomous Response Data. (Yayımlanmamış Doktora Tezi), Minnesota: Minnesota University, 2000.
- SHARIFI, Hassan Pasha. The Principle of Psychometrics & Mental Testing. *Roshd Publishing*. Tehran 1995.
- SOMER, Oya. Kişilik Testlerinde Klasik ve Modern Test Kuramları İke Madde Analizi. *Türk Psikoloji Dergisi*, 1998, 13, 41, 1-15.
- STAGE, Christina. The Applicability of Item Response Models to The SweSAT: A Study of DTM Subtest. *Report*, Sweden: Umea University, Department of Educational Measurement, 1997a.
- STAGE, Christina. The Applicability of Item Response Models to The SweSAT: A Study of ERC Subtest. *Report*, Sweden: Umea University, Department of Educational Measurement, 1997B.
- STAGE, Christina. The Applicability of Item Response Models to The SweSAT: A Study of READ Subtest. *Report*, Sweden: Umea University, Department of Educational Measurement, 1997c.
- STAGE, Christina. The Applicability of Item Response Models to The SweSAT: A Study of WORD Subtest. *Report*, Sweden: Umea University, Department of Educational Measurement, 1997d.
- STAGE, Christina. Item Analysis Based on Item Response Theory and on Classical Test Theory: a Comparison. *The International SweSAT Conference*, San Diego, 1998a.
- STAGE, Christina. A Comparison Between Item Analysis Based on Item Response Theory and Classical Test Theory: A Study of ERC Subtest. *Report*, Sweden: Umea University, Department of Educational Measurement, 1998b.

- STAGE, Christina. A Comparison Between Item Analysis Based on Item Response Theory and Classical Test Theory: A Study of WORD Subtest. *Report*, Sweden: Umea University, Department of Educational Measurement, 1998c.
- STAGE, Christina. A Comparison Between Item Analysis Based on Item Response Theory and Classical Test Theory: A Study of READ Subtest. *Report*, Sweden: Umea University, Department of Educational Measurement, 1998d.
- STAGE, Christina. Predicting Gender Differences in WORD Item: A Comparison of Item Response Theory and on Classical Test Theory. *Report*, Sweden: Umea University, Department of Educational Measurement, 1998e.
- TATLIDİL, Hüseyin. Uygulamalı Çok Değişkenli İstatistiksel Analiz. *Akademi Matbaası*, Ankara, 1996.
- TEKİN, Halil. Eğitimde Ölçme ve Değerlendirme. *Yargı Yay*, Ankara, 1993.
- TEZBAŞARAN, Ata. Likert Tipi Ölçek Geliştirme Kılavuzu. *Psikologlar Derneği Yayınları*, Ankara, 1996.
- THORNDIKE, Robert L. Applied Psychometrics. *Houghton Mifflin Company*, Boston, 1982.
- TREVISAN, S. Michael. and G. SAX. and W. B. MICHAEL. The Effects of The Number of Options Per Item and Student Ability on Test Validity and Reliability. *Educational and Psychological Measurement*, 1991, 51.
- TURGUT, M. F. Eğitimde Ölçme ve Değerlendirme Metodlar. *Saydam Matbaacılık*, Ankara, 1992.
- TURGUT, M. F. ve Y. BAYKUL. Ölçekleme Teknikleri. *ÖSYM Yay*, Ankara, 1992.
- WILCOX, R. R. True-Score Models. H. J. Walberg ve G. D. Haertel (Ed.), *Pergamon Pres*, Oxford, 1990.
- WRIGHT, B. D. AND M. H. STONE. Best Test Design. *Mesa Pres*, Chicago, 1979.

VAN DER LINDEN, Wim J.;Hambleton, Ronald K. Handbook of modern item response theory. *Springer-Verlag*, New York Inc 1997.



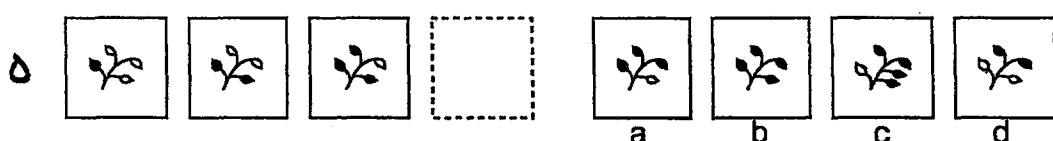
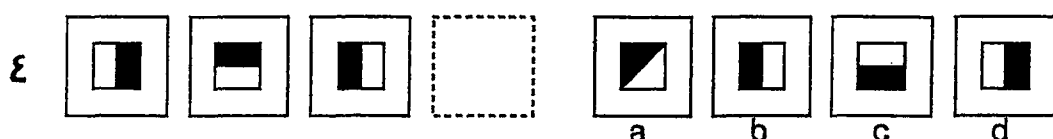
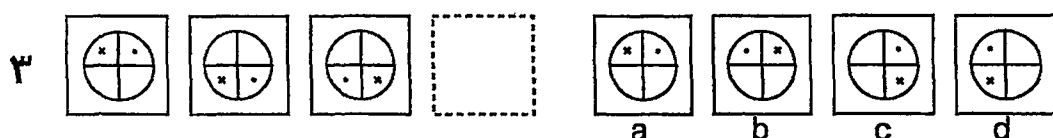
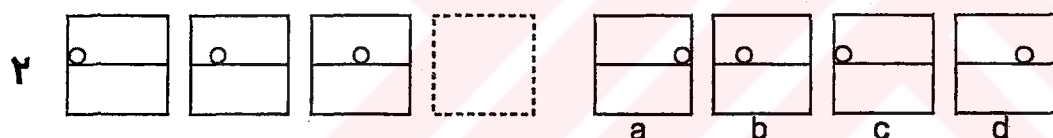
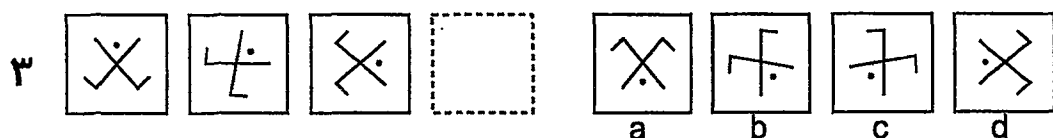
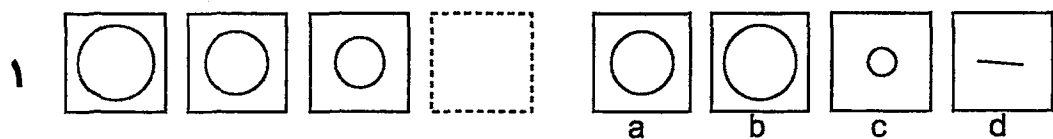
EKLER

- EK 1 B (4 SEÇENEKLİ) TEST FORMU**
EK 2 A VE B TESLERİNE AİT MADDE PARAMETRELERİ
EK 3 A VE B TESTLERİN MADDE ÖZELLİKLER GRAFİKLERİ
EK4 AVE B TESTLERİN MADDE BİLGİ FONKSYONLARI

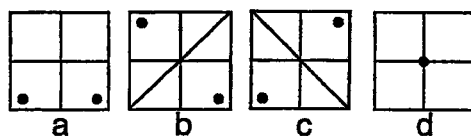
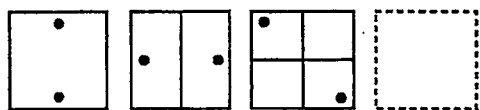


آزمون ۱

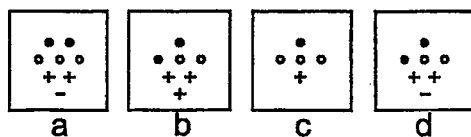
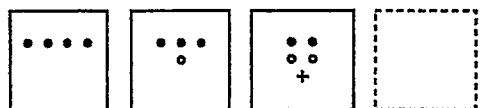
سؤالات نمونه



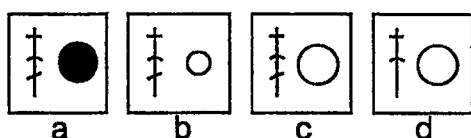
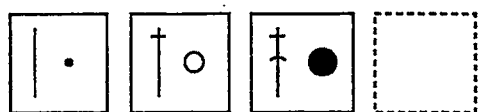
۶



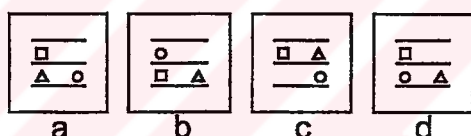
۷



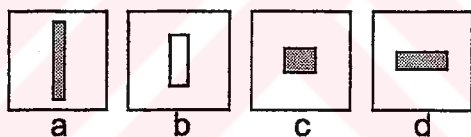
۸



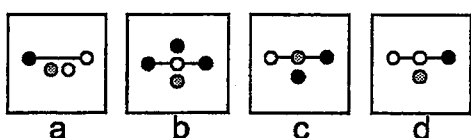
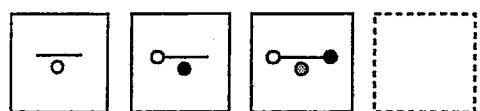
۹



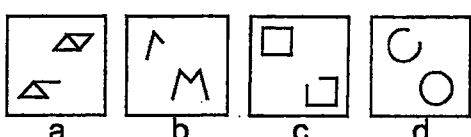
۱۰



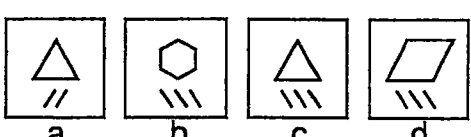
۱۱



۱۲



۱۳

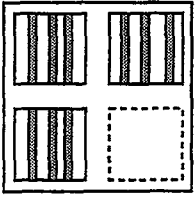


لطفاً قبل از آنکه به شما اجازه بدهند صفحه را برگردانید

آزمون ۲

سؤالات نمونه

۱



a



b

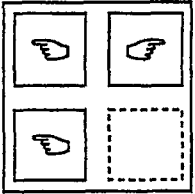


c



d

۲



a



b

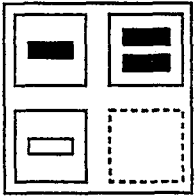


c



d

۳



a



b

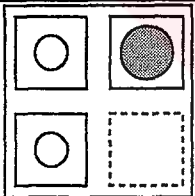


c



d

۱



a



b

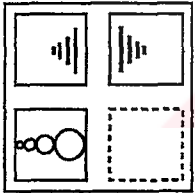


c



d

۲



a



b

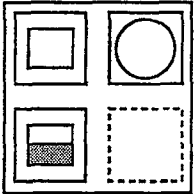


c



d

۳



a



b

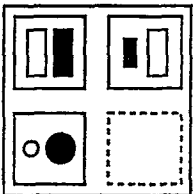


c



d

۴



a



b

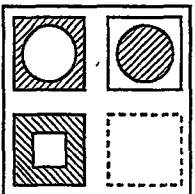


c



d

۵



a



b



c



d

۶

●

●

○

■

■

□

▲

▲

△

▲

△

▲

۷

○●

○●

●○

●○

○●

○●

●○

○●

○●

○●

○●

○●

۸

☁

☁

☁

□

□

□

○

○

○

□

☁

○

۹

●●●

●●●

●●●

○●

○●

○●

○●

○●

●●●

○●

○●

●●●

۱۰

||

||

||

||

||

=

—

||

||

۱۱

⌒

⌒

⌒

□

□

□

○

○

□

□

□

□

۱۲

／

／

／

○

○

○

◇

◇

◇

◇

◇

◇

۱۳

○

○

○

○

○

○

○

○

○

○

○

○

لطفاً قبل از آنکه به شما اجازه بدهند صفحه را برگردانید

آزمون ۳

سؤالات نمونه

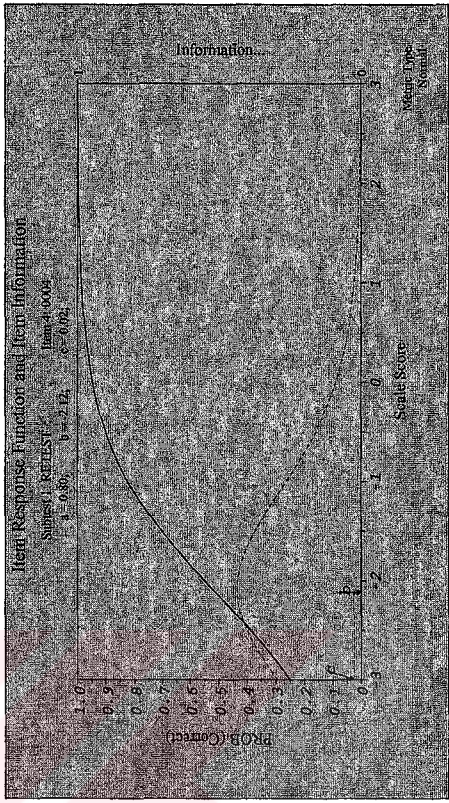
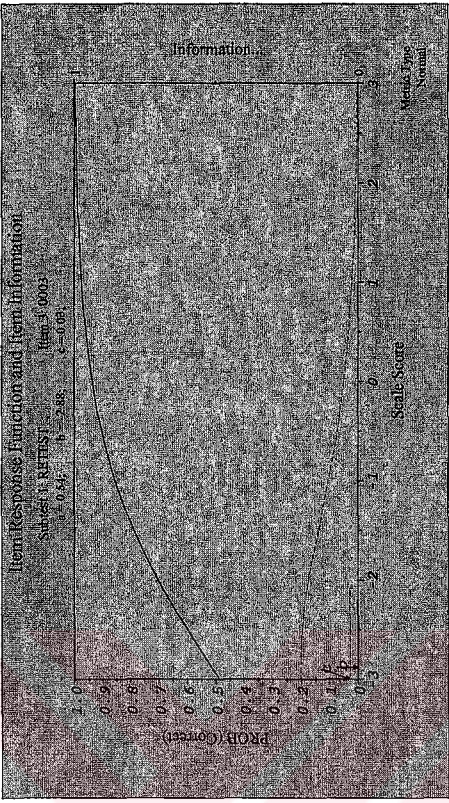
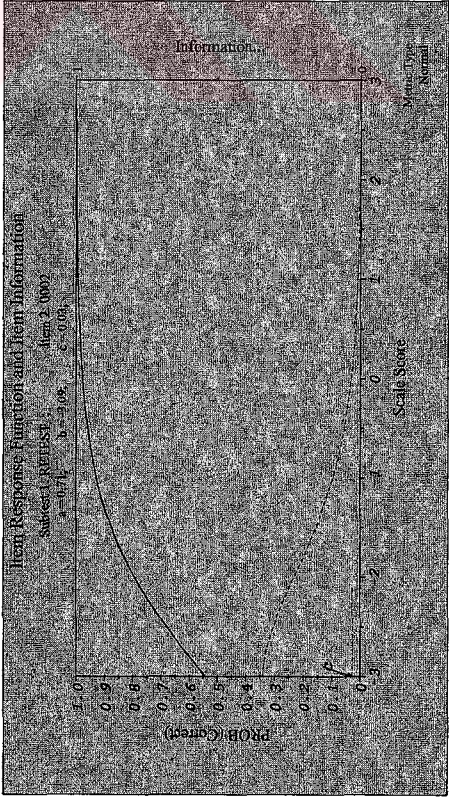
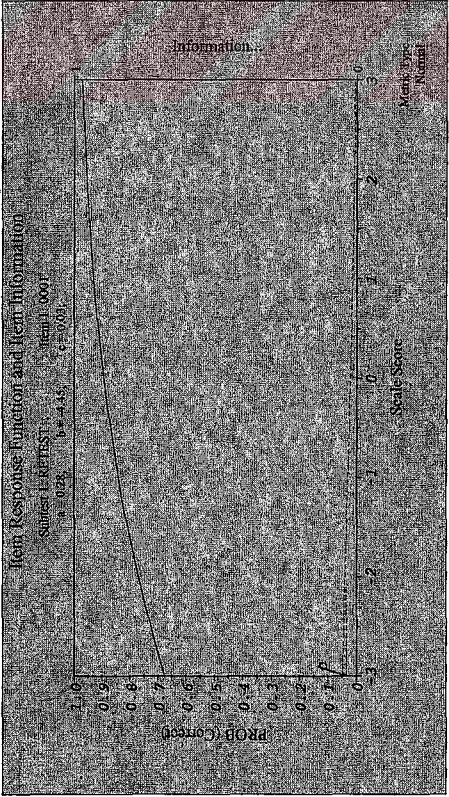
۱					
۲					
۳					
۱					
۲					
۳					
۴					
۵					
۶					
۷					
۸					
۹					
۱۰					

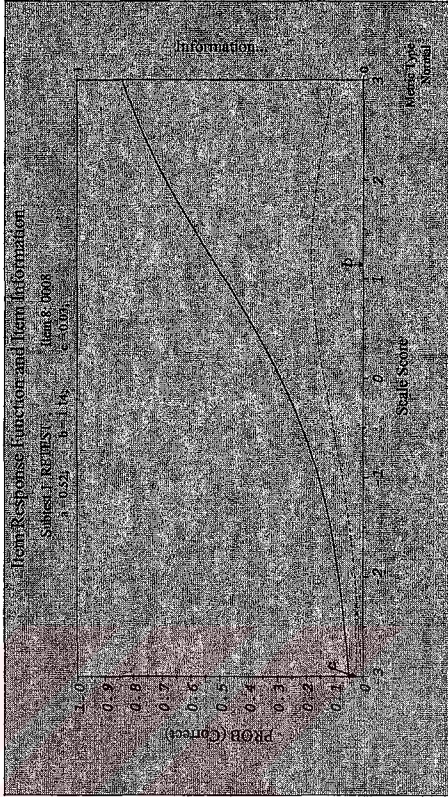
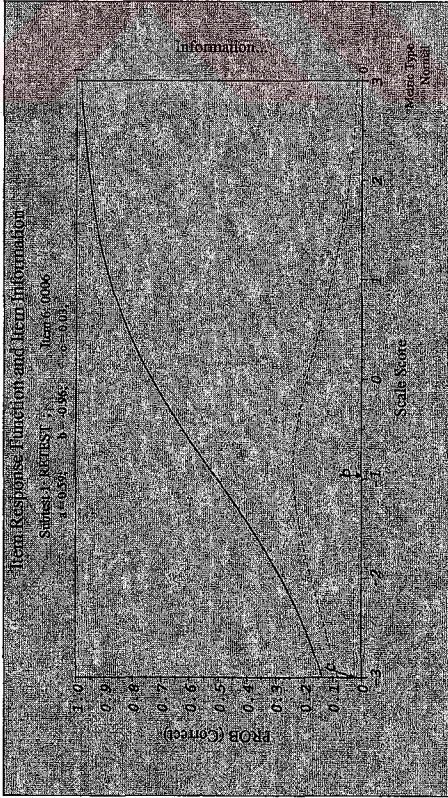
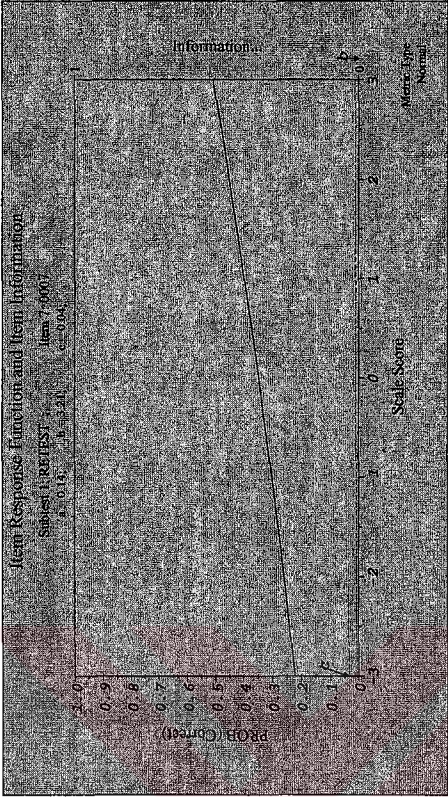
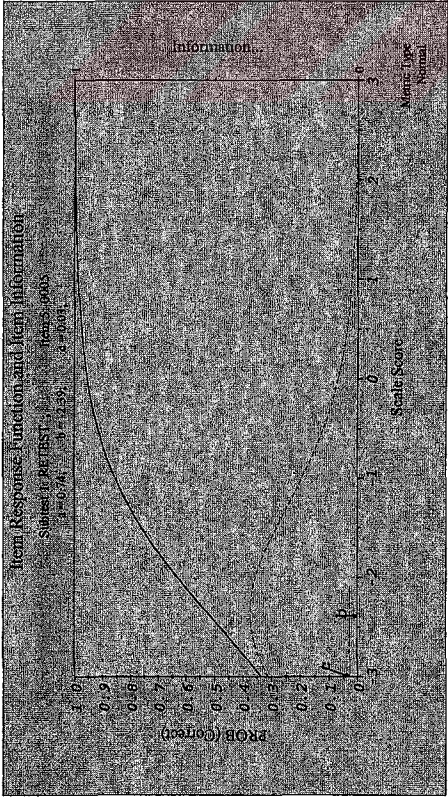
EK 2**TEST PARAMETRELERİ****A= 6 SEÇENEKLİ****B= 4 SEÇENEKLİ**

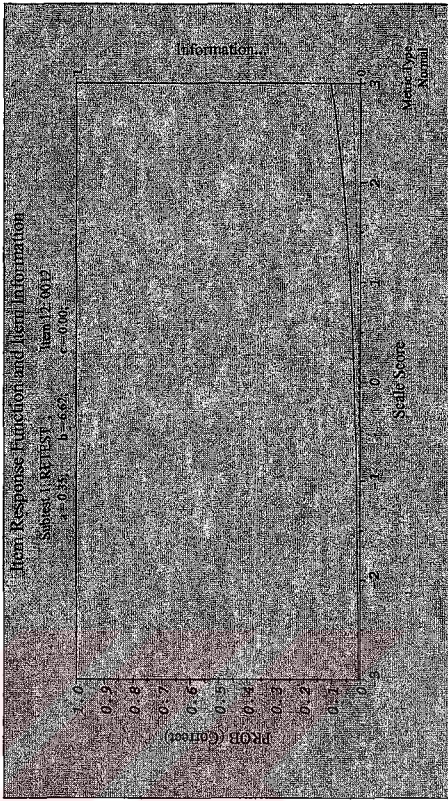
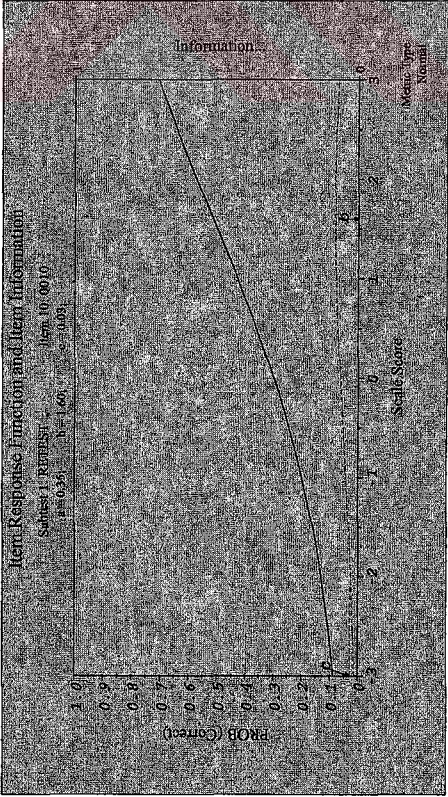
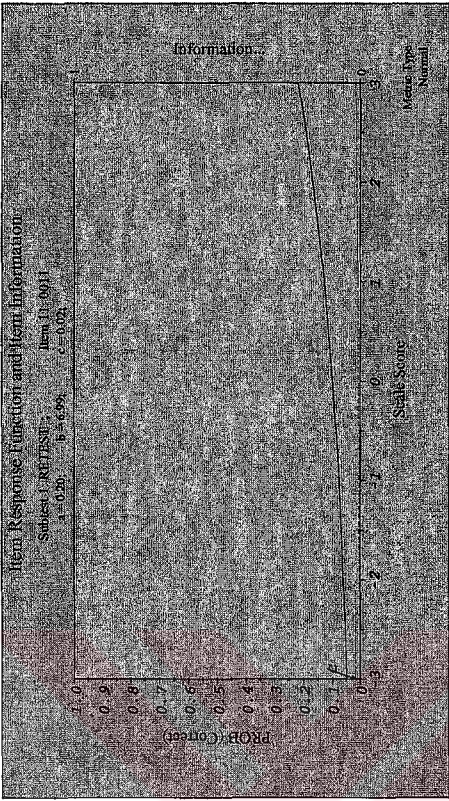
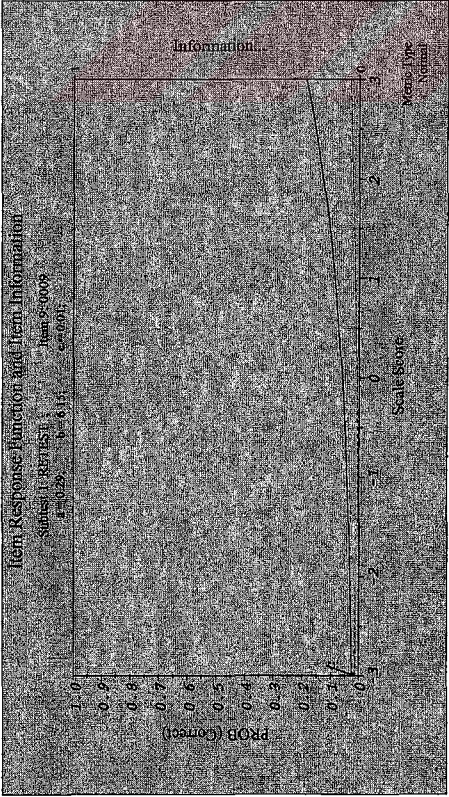
	Test-A			Test-B		
	1. Alt test			1. Alt test		
NO	aj	bj	cj	aj	bj	cj
1	0.28	-4.45	0.03	0.46	-3.88	0.10
2	0.71	-3.09	0.03	0.37	-2.82	0.10
3	0.54	-2.88	0.03	0.31	-2.53	0.11
4	0.80	-2.12	0.02	0.60	-2.78	0.10
5	0.74	-2.39	0.03	0.53	-4.03	0.10
6	0.59	-0.96	0.03	0.32	-2.08	0.10
7	0.14	3.21	0.04	0.32	6.14	0.03
8	0.52	1.14	0.03	0.47	0.05	0.09
9	0.29	6.15	0.01	1.06	0.86	0.11
10	0.35	1.60	0.03	1.61	0.84	0.35
11	0.20	6.99	0.02	1.09	1.97	0.08
12	0.35	6.62	0.00	0.84	2.74	0.03
13	0.28	6.79	0.01	0.69	4.21	0.02
	2. Alt test			2. Alt test		
1	0.40	2.14	0.04	0.30	-3.88	0.10
2	0.51	-1.63	0.03	0.49	-4.00	0.10
3	0.36	-1.21	0.03	0.62	-2.69	0.10
4	0.52	-1.09	0.03	0.23	-1.08	0.11
5	0.49	1.57	0.01	0.34	-3.25	0.10
6	0.25	-2.20	0.04	0.21	-2.86	0.12
7	0.48	-1.37	0.04	0.18	0.48	0.13
8	0.40	1.47	0.03	0.22	-1.34	0.13
9	0.32	6.87	0.01	0.16	9.38	0.02
10	0.42	4.16	0.01	0.30	5.98	0.06
11	0.18	11.18	0.00	0.50	4.01	0.06
12	0.28	7.80	0.01	0.22	10.14	0.01
13	0.30	8.77	0.00	1.17	2.72	0.03
	3. Alt test			3. Alt test		
1	0.41	-3.12	0.03	1.13	-1.79	0.09
2	0.89	-1.95	0.04	0.96	-2.48	0.09
3	1.07	-0.7	0.03	0.86	-1.61	0.08
4	0.54	-1.26	0.04	0.82	-0.47	0.05
5	0.29	0.81	0.02	0.46	-1.1	0.10
6	1.01	-0.89	0.06	1.22	-0.89	0.10
7	1.19	-0.31	0.05	0.82	-0.6	0.08
8	0.82	1.06	0.08	0.56	-0.43	0.12
9	0.68	2.65	0.02	1.18	-1.14	0.10
10	0.12	14.59	0.00	0.28	1.95	0.11

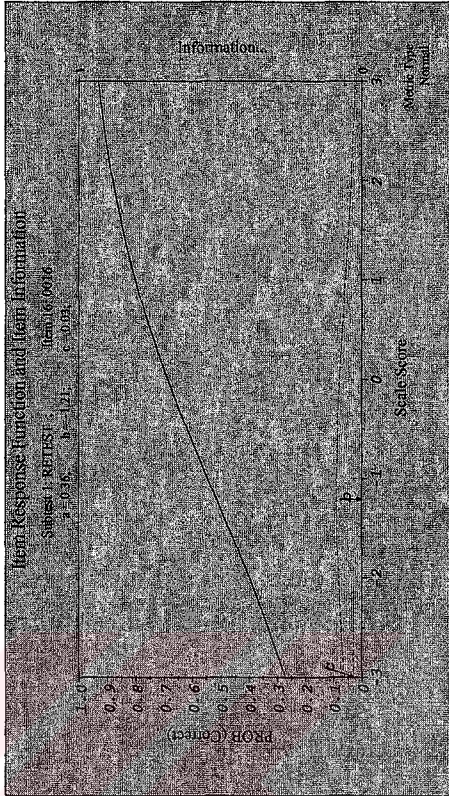
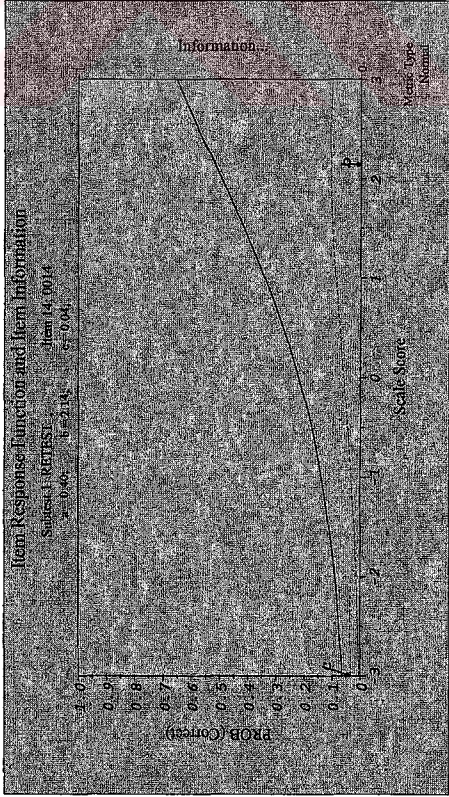
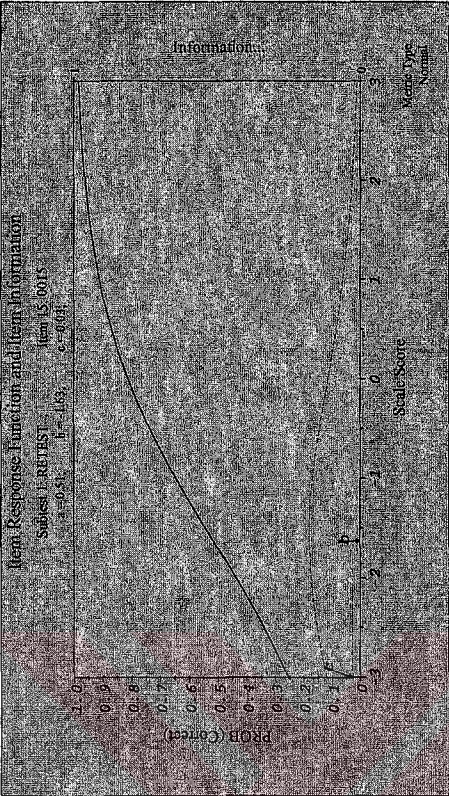
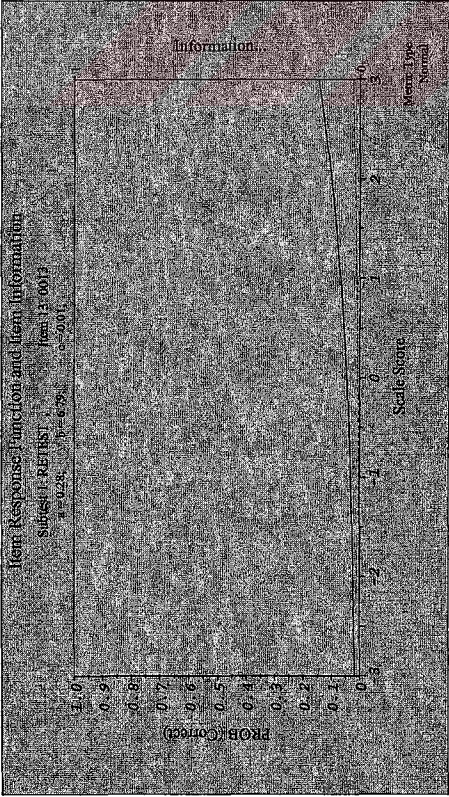
**EK 3 A TESTE AIT OLAN MADDELERİN
ÖZELLİKLERİNİN GRAFİKLERİ**

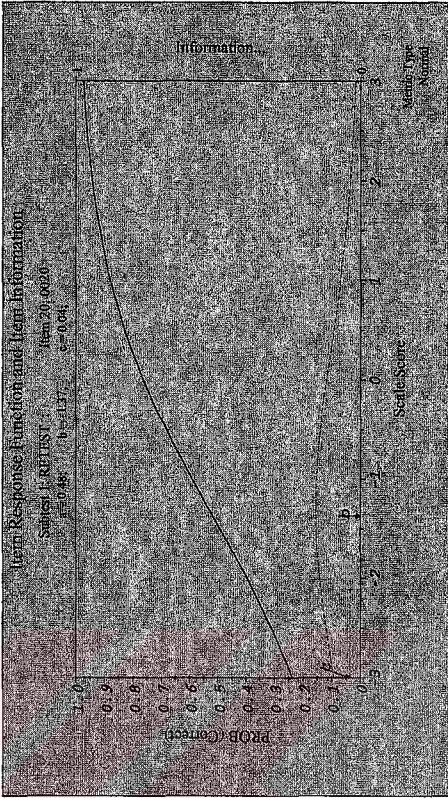
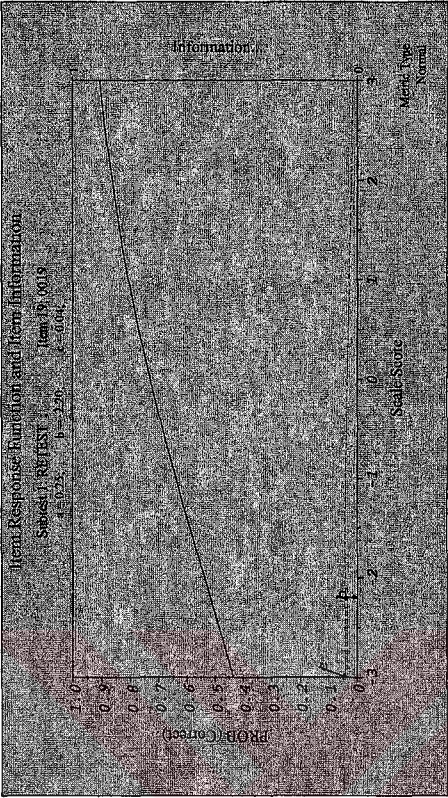
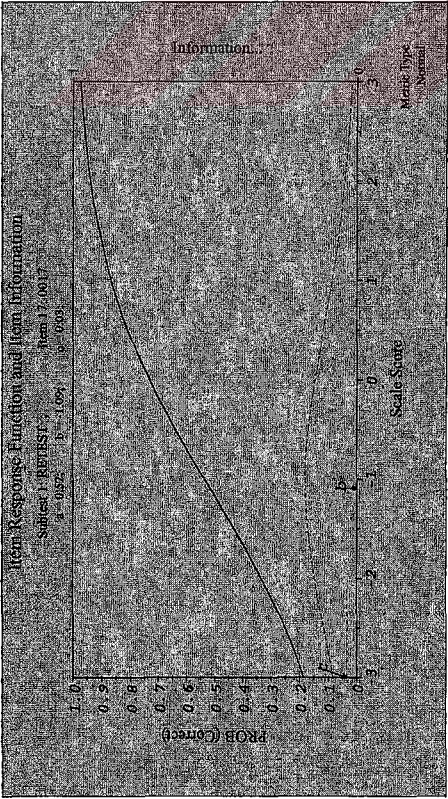
1. ALT TEST= 1-13 ve 2. ALT TEST = 14-26 ve
3. ALT TEST= 27-36 MADDERELEİ İÇERİR

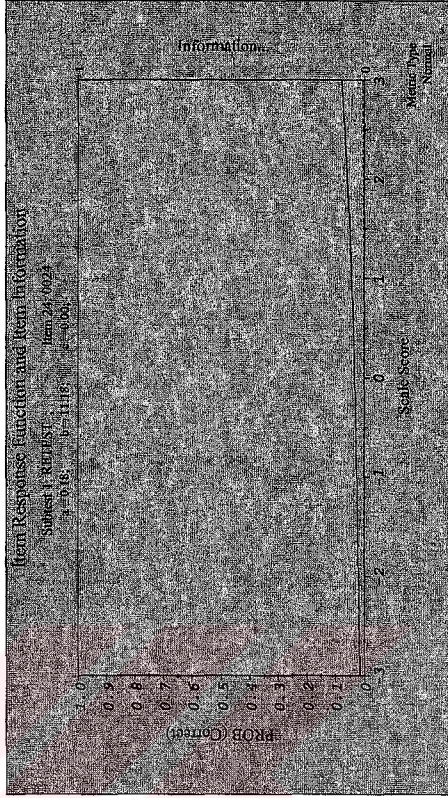
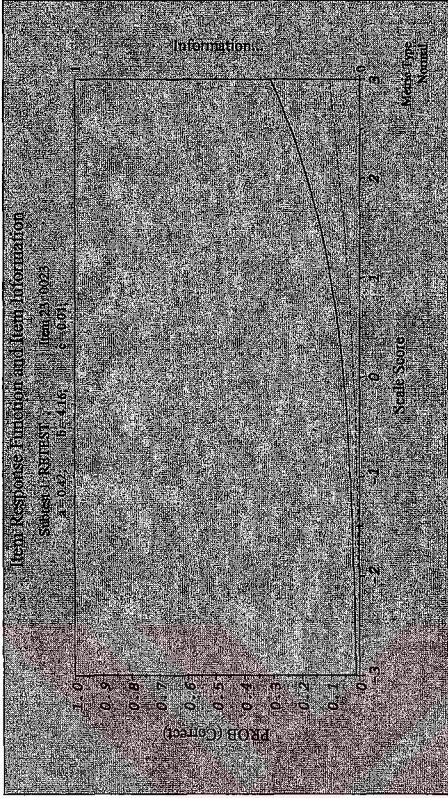
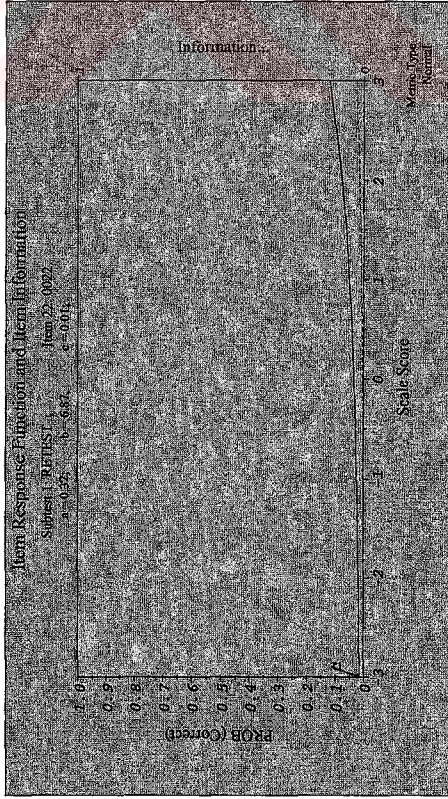
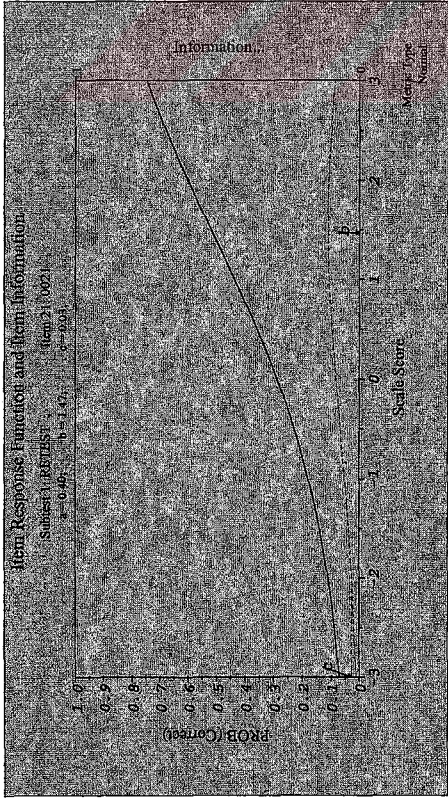


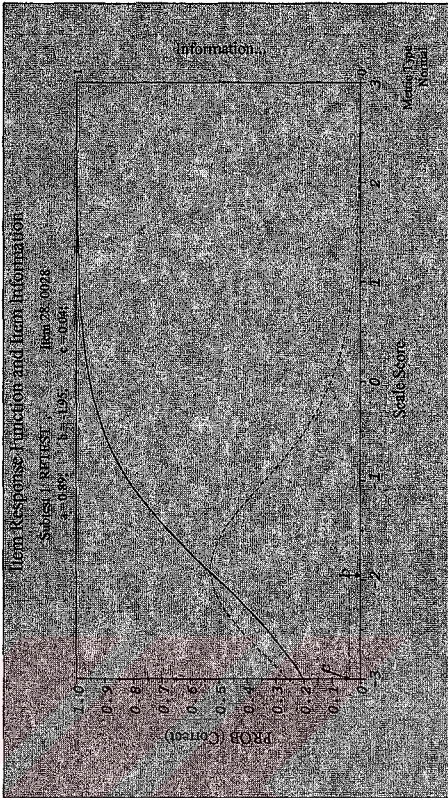
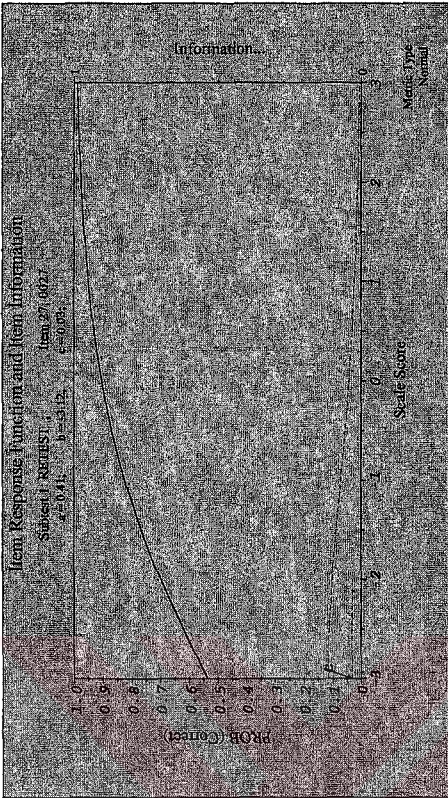
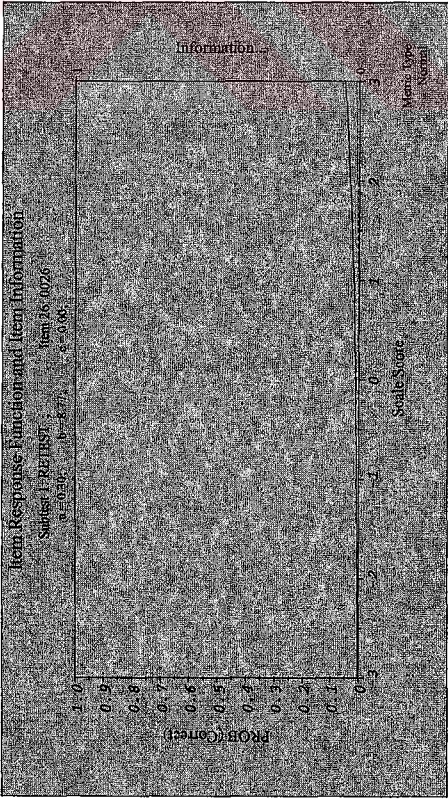
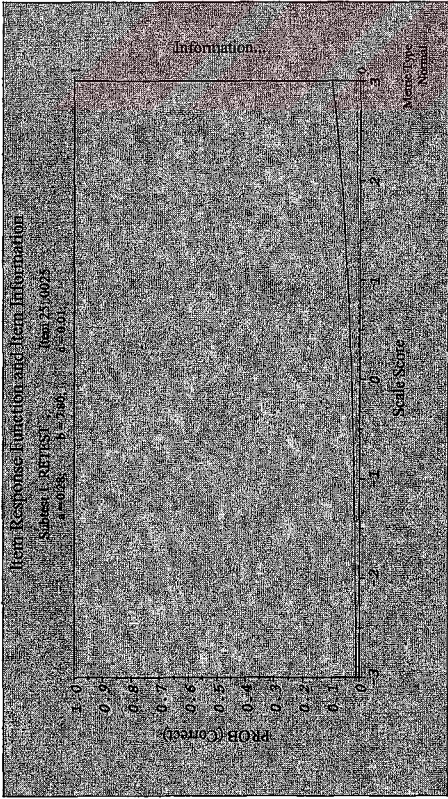


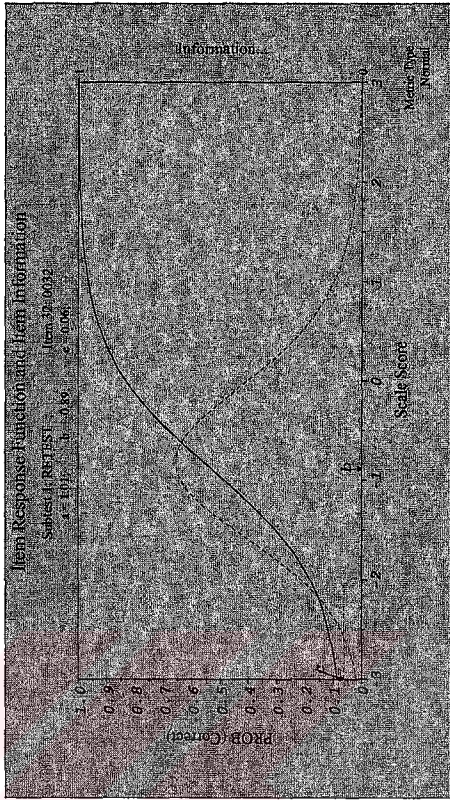
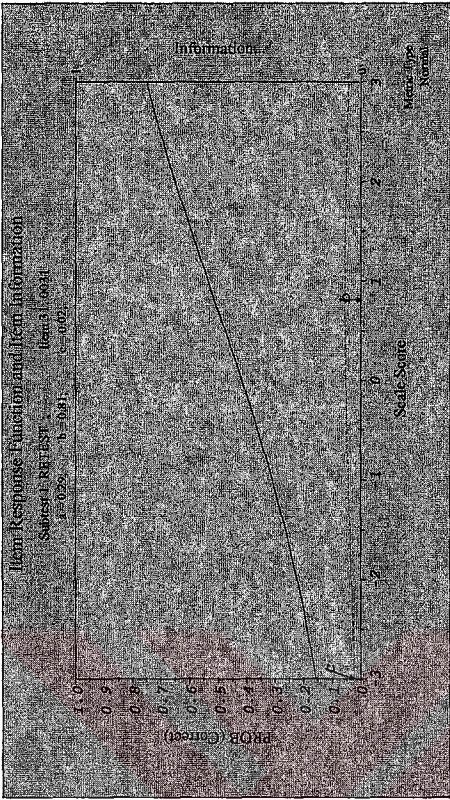


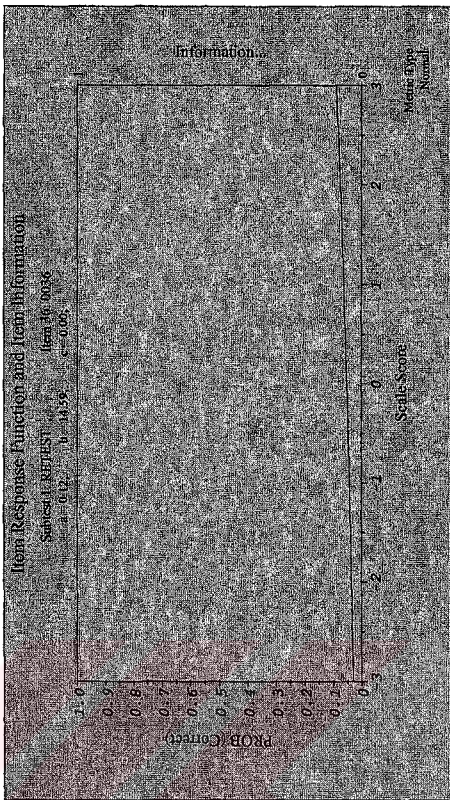
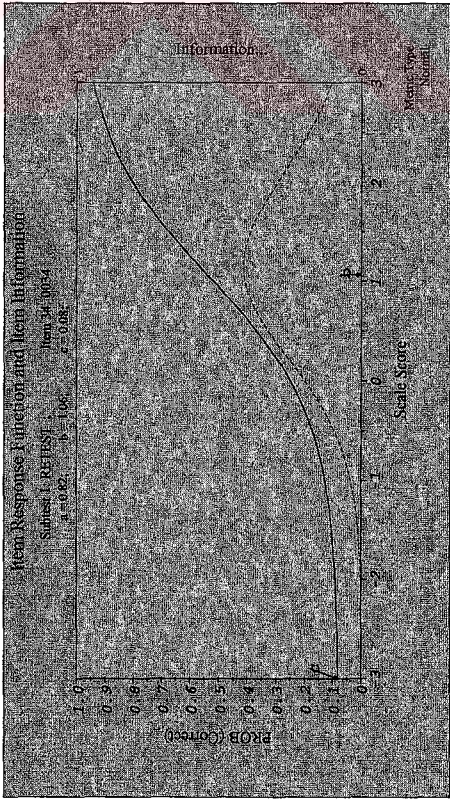
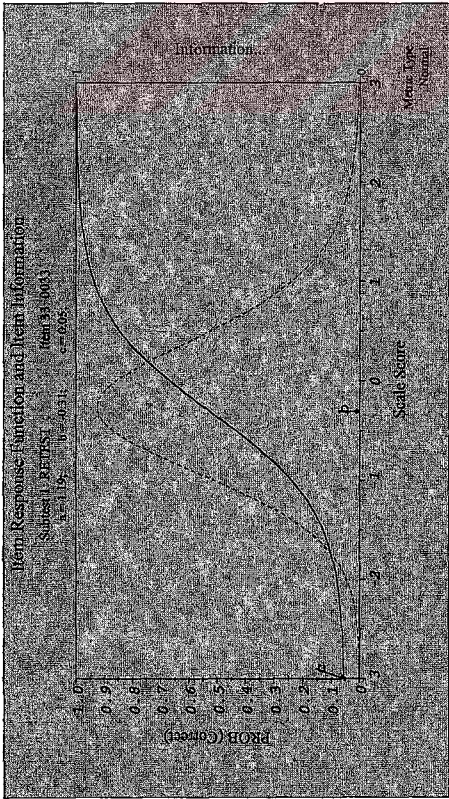






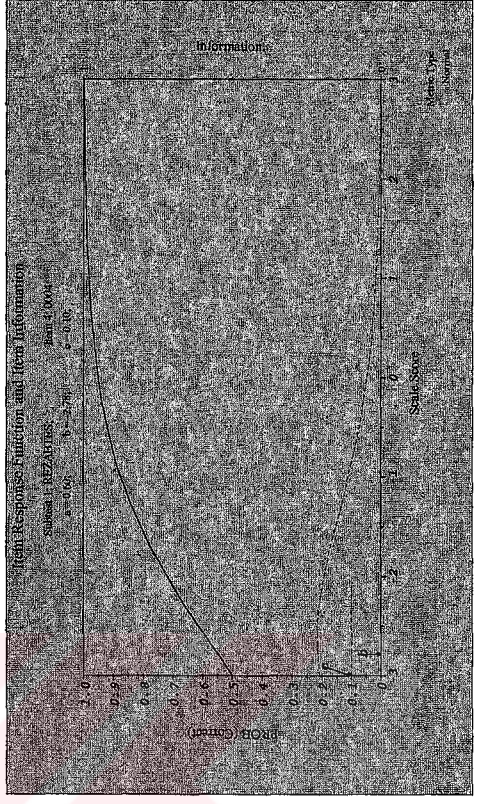
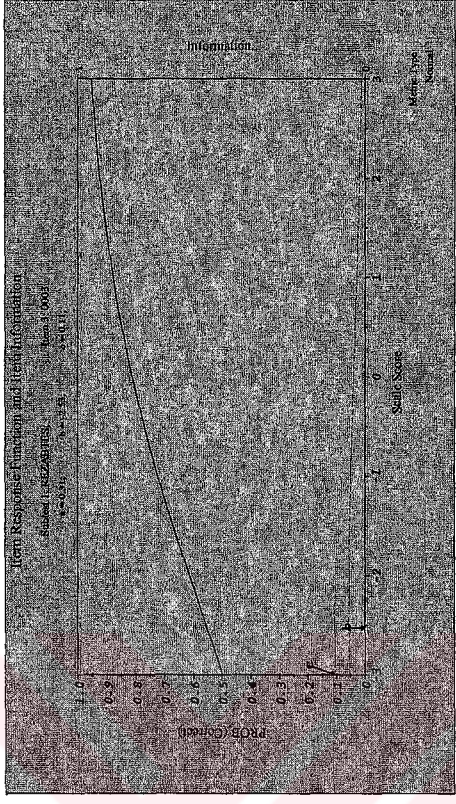
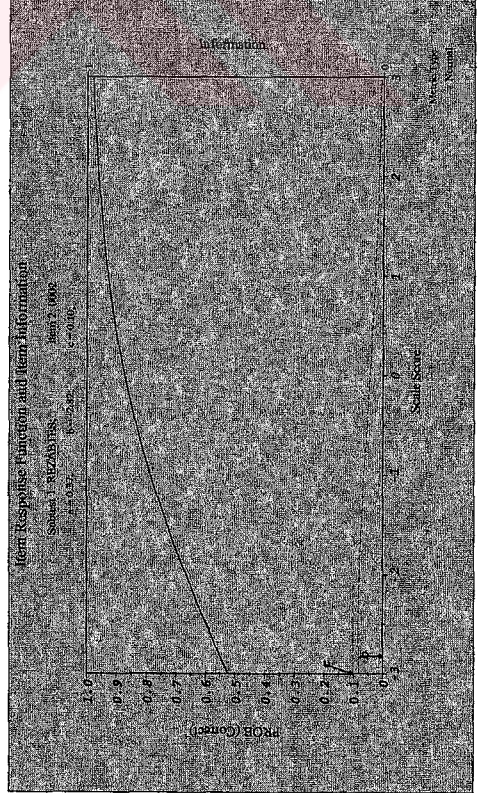
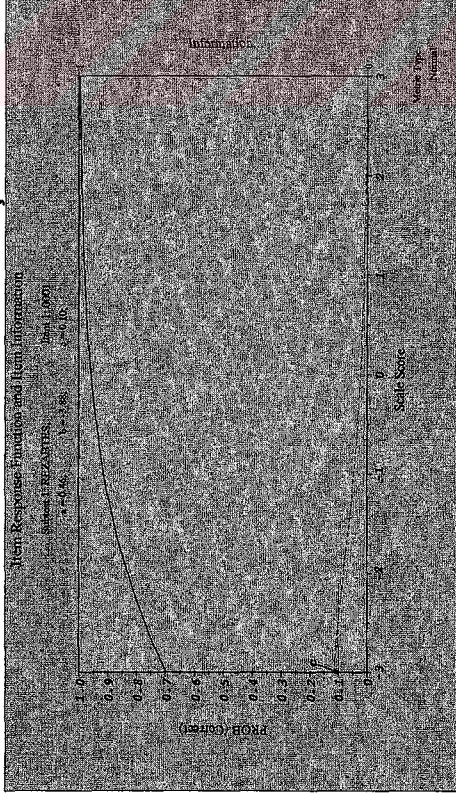


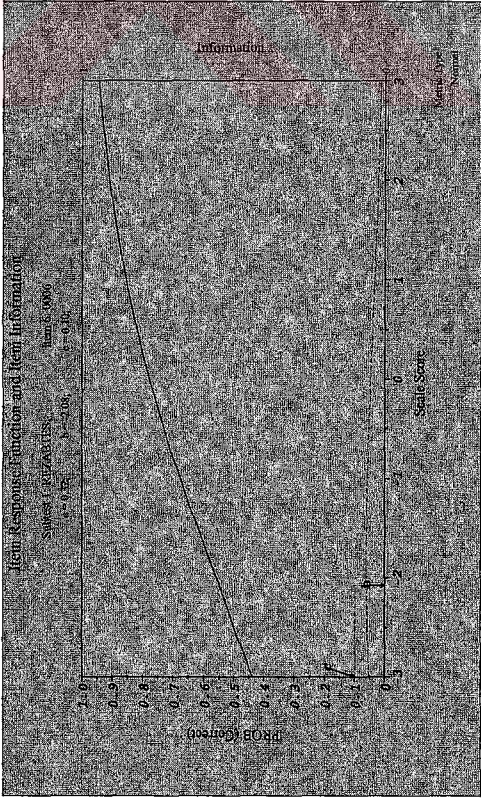
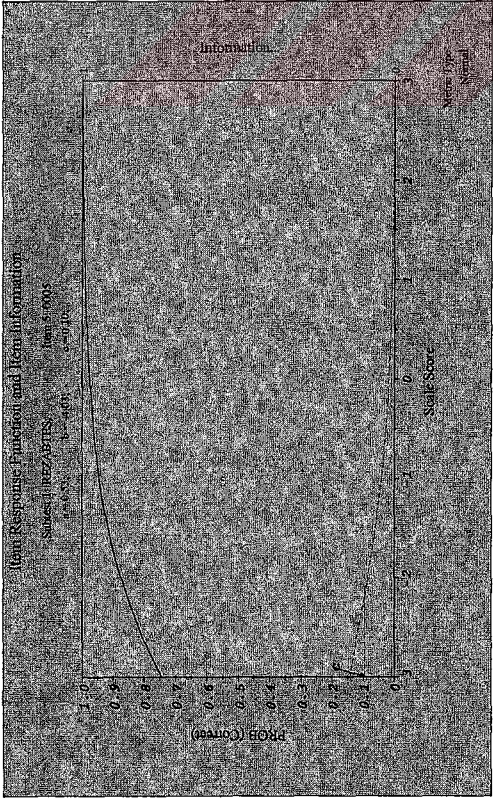


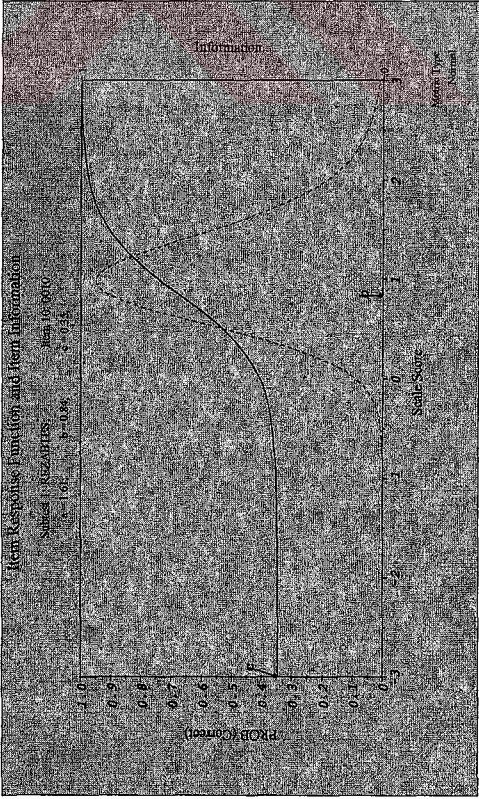


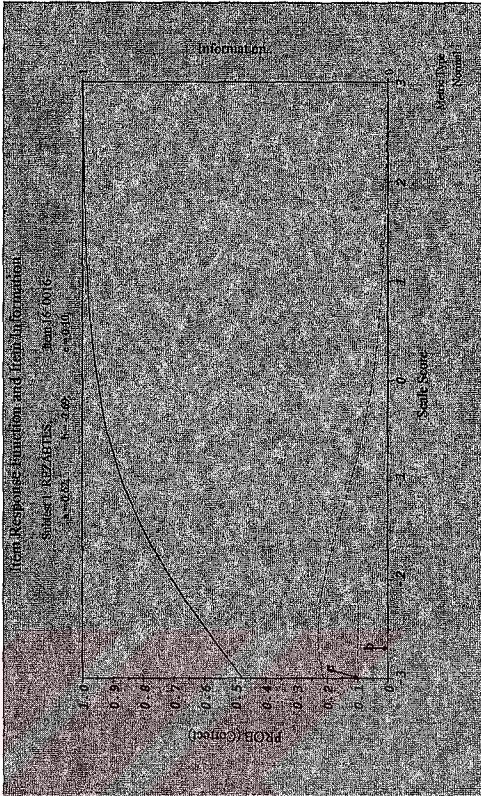
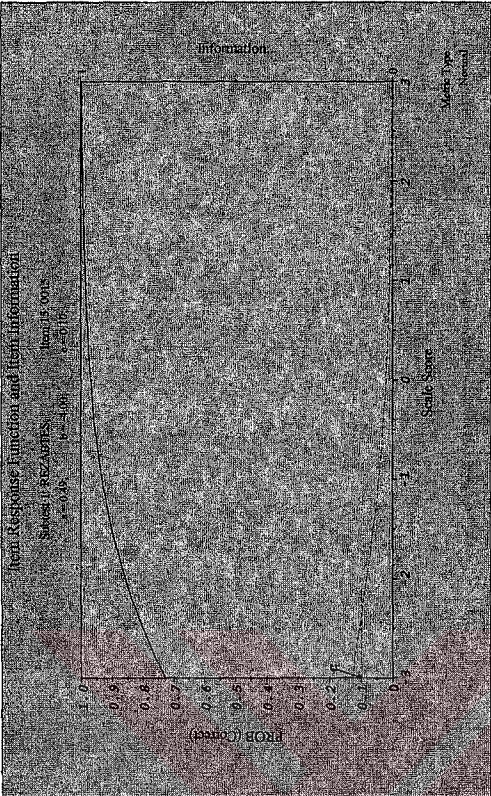
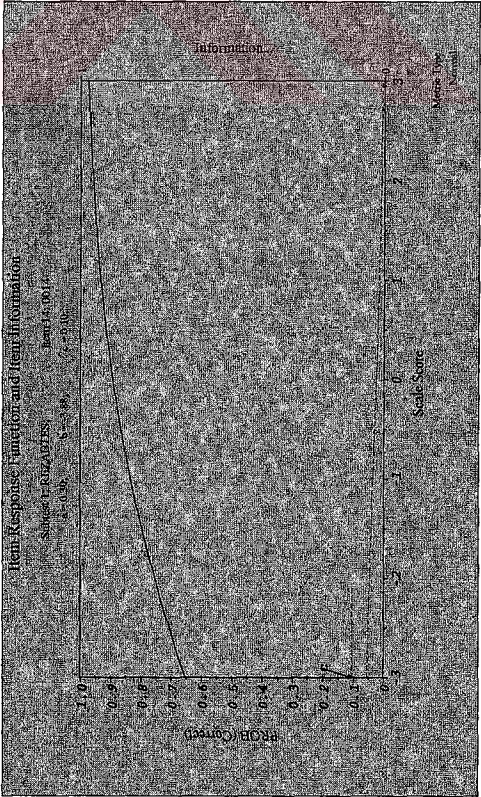
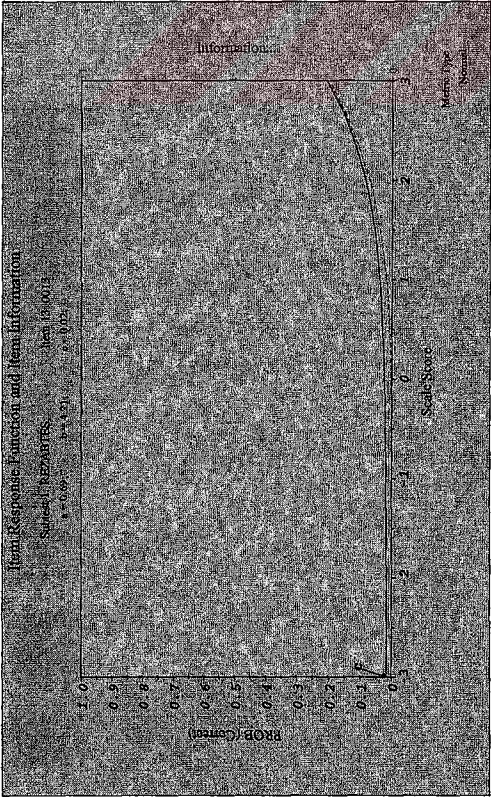
**B TESTE AİT OLAN MADDELERİN
ÖZELLİKLERİNİN GRAFİKLERİ**

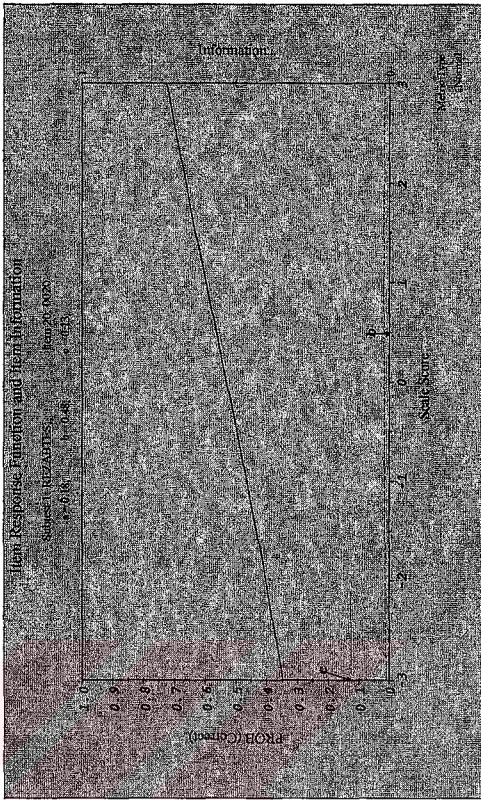
- 1. ALT TEST= 1-13 ve 2. ALT TEST = 14-26 ve
- 3. ALT TEST= 27-36 MADDERELEİ İÇERİR

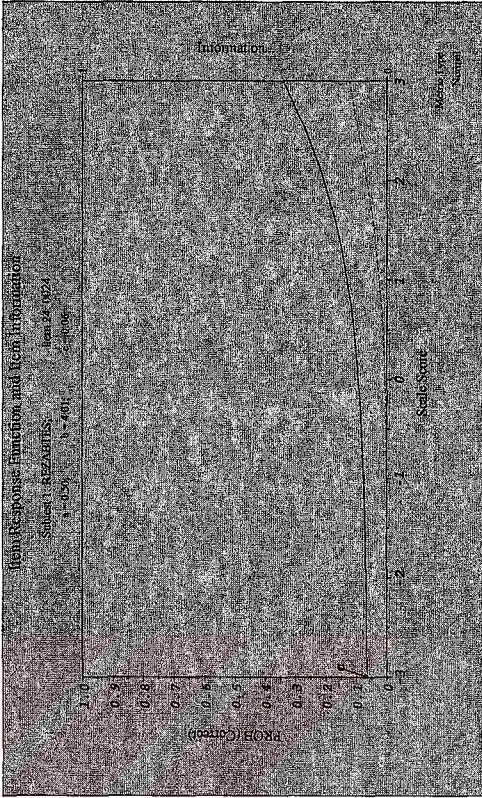
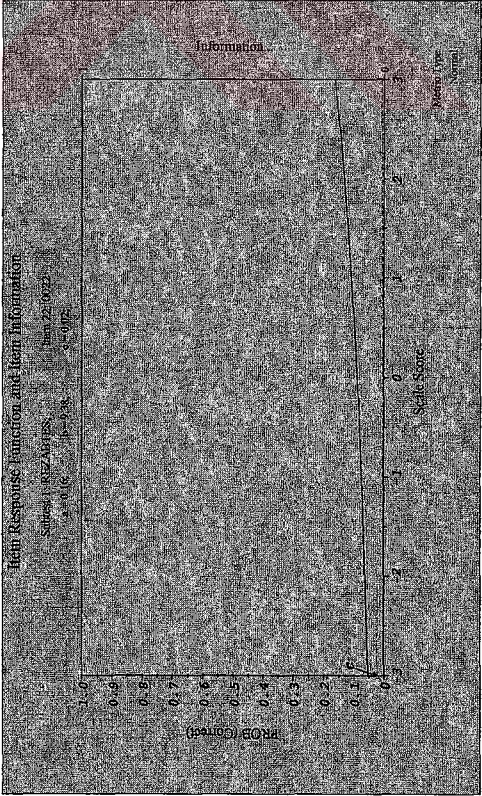
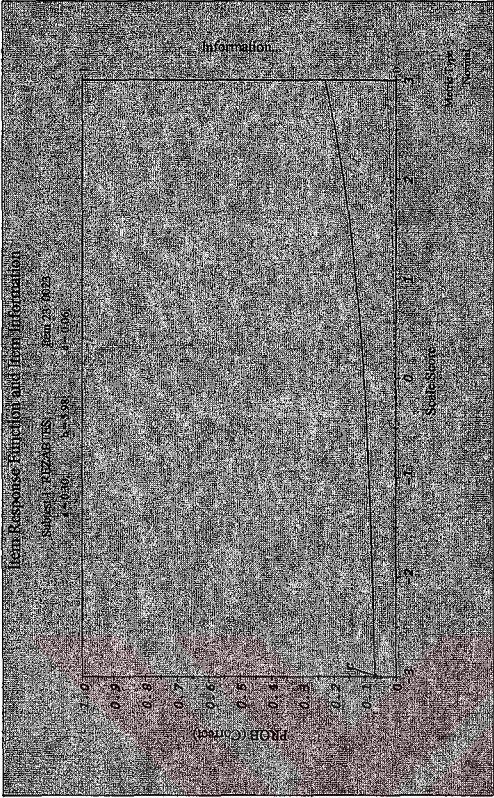
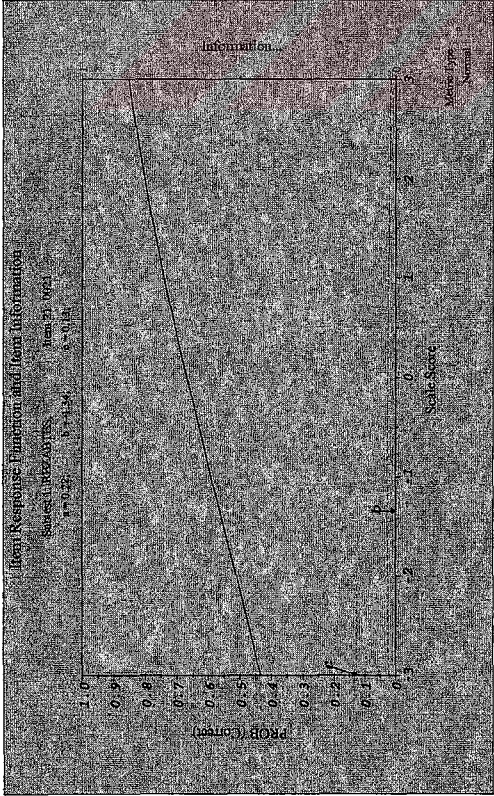


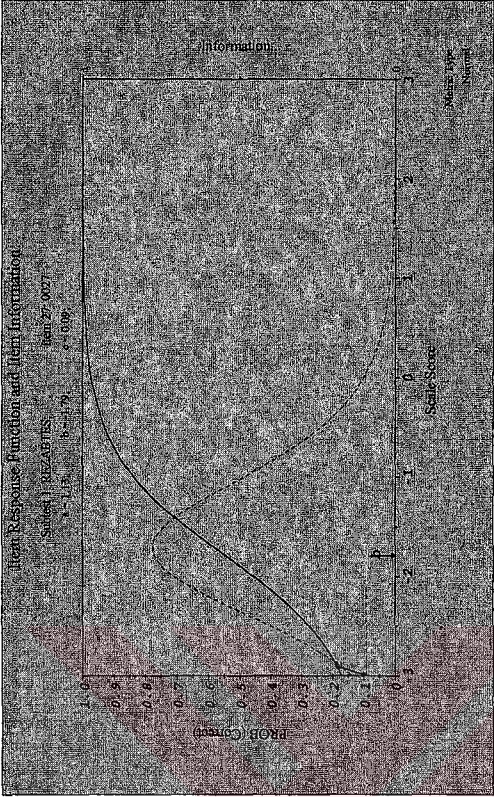


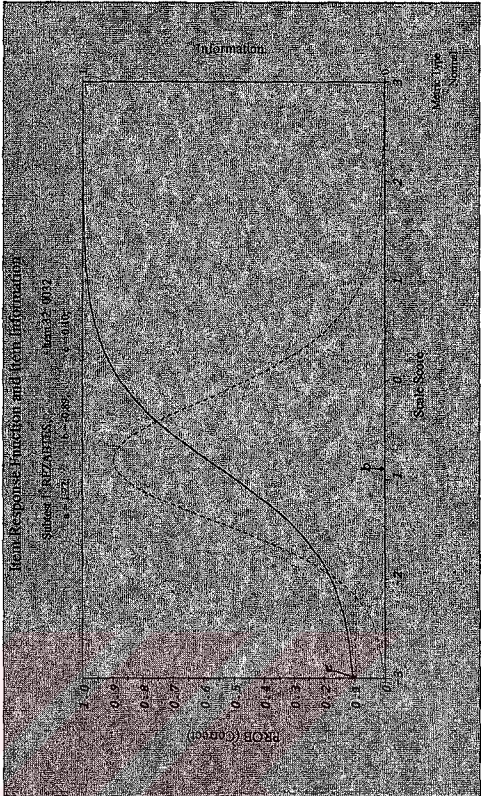
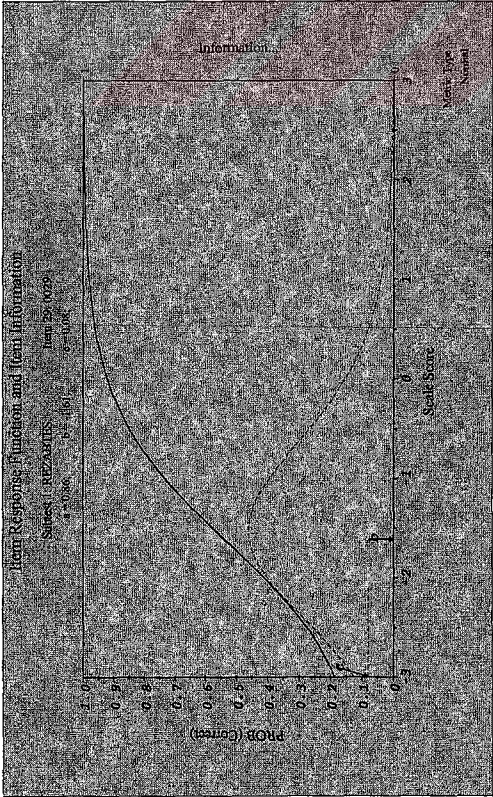


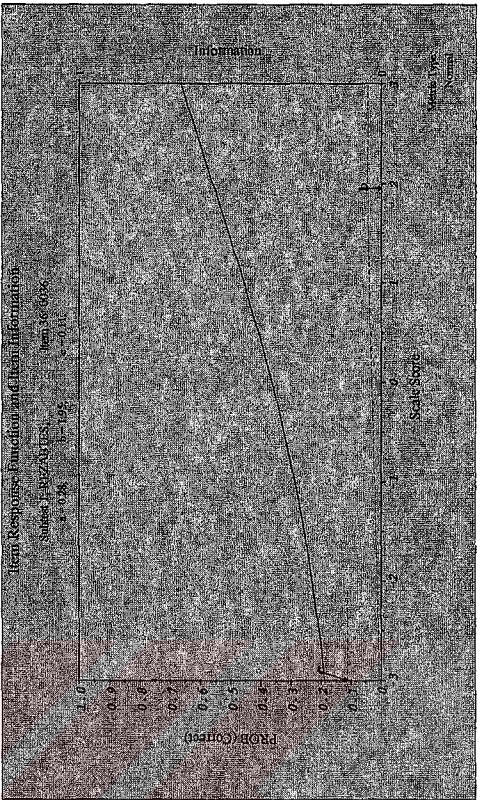
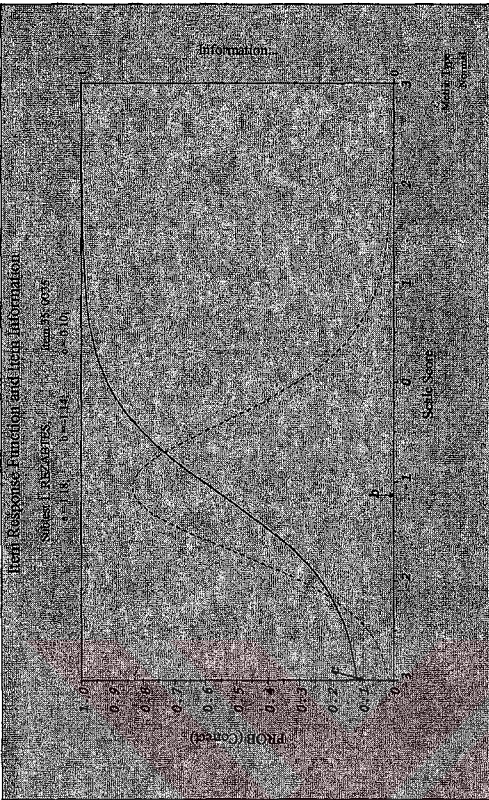
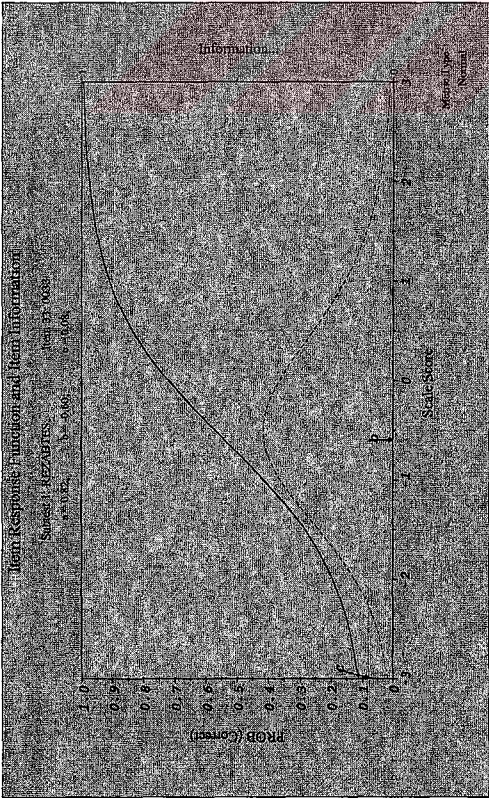












EK 4

A TEST MADEE BİLGİ FONKSYONLARI

ITEM	$\theta=0$	$\theta=-0.5$	$\theta=-1$	$\theta=-1.5$	$\theta=-2$	$\theta=-2.5$	$\theta=-3$
1	0.000113781	0.000265834	0.000619917	0.001439301	0.003307881	0.007427075	0.015829639
2	0.007315124	0.016875175	0.038214629	0.082999806	0.164780516	0.274434934	0.342411385
3	0.006019635	0.013803851	0.030836931	0.065002016	0.121411721	0.18158412	0.195072514
4	0.046723778	0.101878177	0.203935368	0.344993562	0.441273032	0.393946218	0.248815894
5	0.0255005	0.057009708	0.120370761	0.225572508	0.339260838	0.36715652	0.273476975
6	0.132719522	0.207311424	0.236094168	0.185096197	0.103357231	0.043814772	0.014582084
7	2.21684E-05	4.30721E-06	8.08678E-07	1.49498E-07	2.74508E-08	5.02577E-09	9.18978E-10
8	0.069019462	0.027831303	0.008843719	0.002271769	0.00049809	9.93458E-05	1.88841E-05
9	1.9902E-08	3.64183E-09	6.65774E-10	1.21662E-10	2.22283E-11	4.06092E-12	7.41873E-13
10	0.013680995	0.004428091	0.00115512	0.000255865	5.13174E-05	9.78069E-06	1.82033E-06
11	2.70076E-10	4.93481E-11	9.01577E-12	1.64708E-12	3.00895E-13	5.49683E-14	1.00417E-14
12	4.58606E-06	1.96017E-06	8.37806E-07	3.5809E-07	1.53052E-07	6.54165E-08	2.79598E-08
13	2.10975E-09	3.8563E-10	7.04644E-11	1.28739E-11	2.35192E-12	4.29661E-13	7.84917E-14
14	0.004397923	0.00107076	0.000226588	4.43442E-05	8.35351E-06	1.54658E-06	2.84165E-07
15	0.04034771	0.080891695	0.13718339	0.175847418	0.156569561	0.097379818	0.045080504
16	0.036372341	0.063814474	0.086184041	0.081520097	0.053493379	0.025856123	0.009689602
17	0.088388517	0.147207313	0.183669983	0.158672996	0.096144409	0.043578042	0.015418947
18	0.036395353	0.014354008	0.004739983	0.001291713	0.000297193	6.09973E-05	1.17617E-05
19	0.003926211	0.008629527	0.017565151	0.030639476	0.0409549	0.03811037	0.024381981
20	0.051494532	0.09573821	0.141958339	0.150367631	0.108801688	0.055947668	0.02162208
21	0.023052784	0.007928296	0.002179259	0.000500511	0.000102439	1.97179E-05	3.68629E-06
22	2.09961E-09	3.8375E-10	7.01189E-11	1.28106E-11	2.34035E-12	4.27545E-13	7.8105E-14
23	3.34367E-05	6.40085E-06	1.19372E-06	2.20033E-07	4.0351E-08	7.38356E-09	1.34979E-09
24	5.21419E-10	2.22861E-10	9.52536E-11	4.07126E-11	1.74011E-11	7.43743E-12	3.17885E-12
25	6.81091E-11	1.24435E-11	2.2733E-12	4.15298E-13	7.58678E-14	1.38597E-14	2.53191E-15
26	8.71365E-08	3.72433E-08	1.59182E-08	6.80365E-09	2.90797E-09	1.2429E-09	5.31232E-10
27	0.002319272	0.005354026	0.012143762	0.026469058	0.052939689	0.089380706	0.113807274
28	0.074237559	0.158148973	0.301689246	0.467552584	0.525523516	0.403738981	0.218342833
29	0.568822365	0.763325836	0.716581048	0.467034539	0.224511657	0.083731007	0.024755978
30	0.075722548	0.135291195	0.187858498	0.183053666	0.122221244	0.058784194	0.021442172
31	0.035508069	0.017656196	0.006997334	0.002239687	0.000584414	0.000129813	2.60906E-05
32	0.404557444	0.603264012	0.640450582	0.457961704	0.226842829	0.082173568	0.022929747
33	0.881473732	0.885876711	0.604040152	0.291055687	0.104476065	0.029176051	0.00673285
34	0.146342912	0.051529706	0.013892308	0.003112771	0.000626789	0.000119635	2.22784E-05
35	0.005040424	0.001171113	0.000241522	4.66744E-05	8.74204E-06	1.61442E-06	2.96305E-07
36	7.03691E-13	3.00767E-13	1.28551E-13	5.49445E-14	2.3484E-14	1.00373E-14	4.29008E-15

ITEM	$\theta=-3.5$	$\theta=-4$	$\theta=-4.5$	$\theta=-5$	$\theta=-5.5$	$\theta=-6$	$\theta=-6.5$
1	0.030231115	0.046973783	0.053110872	0.041332783	0.022940098	0.00967567	0.003205201
2	0.295810123	0.179239633	0.081241461	0.02874516	0.008102956	0.001895278	0.000392026
3	0.144251966	0.076781426	0.031269767	0.010028072	0.002595781	0.000572017	0.000114404
4	0.119574081	0.045987832	0.014288025	0.00363277	0.000792839	0.000157815	2.9972E-05
5	0.146425021	0.059929899	0.019308014	0.005017224	0.00110846	0.000222003	4.22831E-05
6	0.00389523	0.000876852	0.00017744	3.39651E-05	6.33375E-06	1.16741E-06	2.14083E-07
7	1.67948E-10	3.06862E-11	5.60621E-12	1.02418E-12	1.87102E-13	3.41802E-14	6.24411E-15
8	3.51039E-06	6.46129E-07	1.18418E-07	2.16627E-08	3.95973E-09	7.23552E-10	1.32194E-10
9	1.35528E-13	2.47585E-14	4.52293E-15	8.26257E-16	1.50942E-16	2.75743E-17	5.03731E-18
10	3.35228E-07	6.14522E-08	1.12428E-08	2.05516E-09	3.75541E-10	6.86124E-11	1.25348E-11
11	1.83444E-15	3.35118E-16	6.12198E-17	1.11837E-17	2.04306E-18	3.73229E-19	6.8182E-20
12	1.19504E-08	5.10774E-09	2.18311E-09	9.3309E-10	3.98814E-10	1.70458E-10	7.2856E-11
13	1.4339E-14	2.61948E-15	4.7853E-16	8.74186E-17	1.59698E-17	2.91738E-18	5.32951E-19
14	5.20402E-08	9.51685E-09	1.73934E-09	3.17807E-10	5.80622E-11	1.06073E-11	1.93778E-12
15	0.016257368	0.004661997	0.001104378	0.000230154	4.47033E-05	8.39211E-06	1.55136E-06
16	0.002877526	0.000700079	0.000148272	2.9038E-05	5.47222E-06	1.01332E-06	1.86199E-07
17	0.004346438	0.00101663	0.000210283	4.06898E-05	7.62538E-06	1.40854E-06	2.58545E-07
18	2.20077E-06	4.06251E-07	7.45481E-08	1.36448E-08	2.4947E-09	4.55895E-10	8.32961E-11
19	0.011319448	0.004003098	0.001114727	0.000257516	5.28403E-05	1.01823E-05	1.90449E-06
20	0.006502053	0.001588637	0.000336896	6.60042E-05	1.24401E-05	2.3037E-06	4.23319E-07
21	6.802E-07	1.24797E-07	2.28402E-08	4.17578E-09	7.63096E-10	1.39424E-10	2.54717E-11
22	1.42684E-14	2.60657E-15	4.76173E-16	8.69879E-17	1.58911E-17	2.90301E-18	5.30325E-19
23	2.46657E-10	4.50654E-11	8.23306E-12	1.50406E-12	2.74767E-13	5.01951E-14	9.16973E-15
24	1.35868E-12	5.80717E-13	2.48206E-13	1.06086E-13	4.53426E-14	1.938E-14	8.28325E-15
25	4.62534E-16	8.44963E-17	1.54359E-17	2.81986E-18	5.15135E-19	9.41057E-20	1.71914E-20
26	2.27055E-10	9.70462E-11	4.14787E-11	1.77285E-11	7.5774E-12	3.23867E-12	1.38425E-12
27	0.100568373	0.062139705	0.02861432	0.010270042	0.002932399	0.000692424	0.000144027
28	0.088088466	0.027513159	0.006921148	0.001494419	0.000295532	5.59447E-05	1.03802E-05
29	0.006002343	0.001268587	0.000248173	4.67447E-05	8.65398E-06	1.59003E-06	2.91186E-07
30	0.006131244	0.001443337	0.000299389	5.8003E-05	1.08756E-05	2.00938E-06	3.68868E-07
31	4.97825E-06	9.27014E-07	1.70758E-07	3.13056E-08	5.72768E-09	1.04702E-09	1.91325E-10
32	0.00527455	0.001078287	0.000207351	3.87443E-05	7.14741E-06	1.3112E-06	2.39963E-07
33	0.001379724	0.000265662	4.96703E-05	9.16551E-06	1.68162E-06	3.0777E-07	5.62685E-08
34	4.1037E-06	7.52342E-07	1.37648E-07	2.51622E-08	4.59795E-09	8.4006E-10	1.53471E-10
35	5.4238E-08	9.91676E-09	1.81227E-09	3.3112E-10	6.04936E-11	1.10514E-11	2.01891E-12
36	1.83363E-15	7.83718E-16	3.34971E-16	1.43171E-16	6.1193E-17	2.61547E-17	1.11788E-17

ITEM	$\theta=-7$	$\theta=-7.5$	$\theta=-8$	$\theta=-8.5$	$\theta=-9$	$\theta=-9.5$	$\theta=-10$
1	0.000852764	0.000191436	3.86797E-05	7.39842E-06	1.37918E-06	2.54167E-07	4.66067E-08
2	7.5857E-05	1.42158E-05	2.62591E-06	4.81999E-07	8.82326E-08	1.61326E-08	2.94823E-09
3	2.17749E-05	4.05016E-06	7.45672E-07	1.36677E-07	2.50041E-08	4.57058E-09	8.35179E-10
4	5.56943E-06	1.02495E-06	1.87833E-07	3.43602E-08	6.2806E-09	1.14763E-09	2.09674E-10
5	7.86707E-06	1.44859E-06	2.65533E-07	4.85786E-08	8.87994E-09	1.62263E-09	2.96459E-10
6	3.91732E-08	7.16124E-09	1.30862E-09	2.39091E-10	4.36799E-11	7.97969E-12	1.45776E-12
7	1.14068E-15	2.08382E-16	3.80675E-17	6.95423E-18	1.27041E-18	2.3208E-19	4.23967E-20
8	2.41505E-11	4.41193E-12	8.05985E-13	1.47239E-13	2.68979E-14	4.91374E-15	8.9765E-16
9	9.20223E-19	1.68108E-19	3.07101E-20	5.61017E-21	1.02487E-21	1.87226E-22	3.42026E-23
10	2.28993E-12	4.18331E-13	7.64217E-14	1.39608E-14	2.55039E-15	4.65909E-16	8.51129E-17
11	1.24556E-20	2.27541E-21	4.15675E-22	7.5936E-23	1.38721E-23	2.53418E-24	4.62947E-25
12	3.11395E-11	1.33094E-11	5.68862E-12	2.43139E-12	1.0392E-12	4.44169E-13	1.89843E-13
13	9.73603E-20	1.77859E-20	3.24916E-21	5.93561E-22	1.08433E-22	1.98086E-23	3.61867E-24
14	3.53999E-13	6.46691E-14	1.18139E-14	2.15818E-15	3.94259E-16	7.20237E-17	1.31574E-17
15	2.84856E-07	5.2152E-08	9.53613E-09	1.74277E-09	3.18427E-10	5.81749E-11	1.06278E-11
16	3.41005E-08	6.23622E-09	1.13976E-09	2.08255E-10	3.80475E-11	6.95081E-12	1.2698E-12
17	4.73281E-08	8.65354E-09	1.58143E-09	2.88945E-10	5.27885E-11	9.64375E-12	1.76176E-12
18	1.52176E-11	2.78006E-12	5.07871E-13	9.2779E-14	1.6949E-14	3.09628E-15	5.65633E-16
19	3.51492E-07	6.4494E-08	1.18041E-08	2.15813E-09	3.94386E-10	7.20577E-11	1.31644E-11
20	7.75272E-08	1.4178E-08	2.59126E-09	4.73468E-10	8.65011E-11	1.58027E-11	2.88691E-12
21	4.65333E-12	8.50086E-13	1.55296E-13	2.83697E-14	5.18262E-15	9.4677E-16	1.72957E-16
22	9.68806E-20	1.76983E-20	3.23315E-21	5.90636E-22	1.07898E-22	1.9711E-23	3.60084E-24
23	1.67514E-15	3.06017E-16	5.59037E-17	1.02126E-17	1.86565E-18	3.40819E-19	6.22613E-20
24	3.54036E-15	1.5132E-15	6.46759E-16	2.76433E-16	1.18151E-16	5.04991E-17	2.1584E-17
25	3.14054E-21	5.73719E-22	1.04808E-22	1.91464E-23	3.4977E-24	6.38965E-25	1.16727E-25
26	5.91646E-13	2.52877E-13	1.08083E-13	4.61959E-14	1.97447E-14	8.43913E-15	3.60699E-15
27	2.79476E-05	5.24425E-06	9.6926E-07	1.77957E-07	3.25795E-08	5.95715E-09	1.08869E-09
28	1.90905E-06	3.49753E-07	6.39723E-08	1.16927E-08	2.13652E-09	3.90341E-10	7.1311E-11
29	5.32505E-08	9.73227E-09	1.77825E-09	3.2488E-10	5.93516E-11	1.08426E-11	1.98075E-12
30	6.7526E-08	1.23468E-08	2.25639E-09	4.12268E-10	7.53189E-11	1.37598E-11	2.51369E-12
31	3.49557E-11	6.38609E-12	1.16665E-12	2.13127E-13	3.89344E-14	7.11261E-15	1.29934E-15
32	4.38706E-08	8.01698E-09	1.46476E-09	2.67601E-10	4.88869E-11	8.93083E-12	1.63151E-12
33	1.02827E-08	1.87873E-09	3.4323E-10	6.27035E-11	1.14549E-11	2.09261E-12	3.82281E-13
34	2.8037E-11	5.12188E-12	9.35676E-13	1.70931E-13	3.12259E-14	5.7044E-15	1.04209E-15
35	3.68819E-13	6.73765E-14	1.23084E-14	2.24853E-15	4.10764E-16	7.50389E-17	1.37082E-17
36	4.77797E-18	2.04216E-18	8.72846E-19	3.73065E-19	1.59453E-19	6.81521E-20	2.91291E-20

ITEM	$\theta=-10.5$	$\theta=-11$	$\theta=-11.5$	$\theta=-12$	$\theta=0.5$	$\theta=1$	$\theta=1.5$
1	8.52793E-09	1.55897E-09	2.84879E-10	5.20487E-11	4.86606E-05	2.08035E-05	8.89266E-06
2	5.38673E-10	9.84122E-11	1.79786E-11	3.2844E-12	0.003145576	0.001347952	0.000576771
3	1.52589E-10	2.78765E-11	5.09262E-12	9.30336E-13	0.002595205	0.001113341	0.000476611
4	3.83052E-11	6.99779E-12	1.27838E-12	2.33537E-13	0.020596695	0.008921541	0.003835091
5	5.41602E-11	9.89426E-12	1.80751E-12	3.30201E-13	0.011116634	0.004791945	0.002055613
6	2.66306E-13	4.86493E-14	8.88732E-15	1.62355E-15	0.069285508	0.032459675	0.014447147
7	7.7451E-21	1.41489E-21	2.58473E-22	4.72183E-23	0.000106261	0.000446988	0.001542219
8	1.63984E-16	2.99568E-17	5.47256E-18	9.99735E-19	0.131196774	0.179993819	0.170043941
9	6.24819E-24	1.14143E-24	2.08518E-25	3.80923E-26	1.08518E-07	5.88642E-07	3.15526E-06
10	1.55486E-17	2.84043E-18	5.18894E-19	9.47924E-20	0.033260222	0.061753132	0.082306986
11	8.45718E-26	1.54497E-26	2.82237E-27	5.15595E-28	1.47769E-09	8.0798E-09	4.41129E-08
12	8.11415E-14	3.46809E-14	1.4823E-14	6.33555E-15	1.07295E-05	2.51012E-05	5.87172E-05
13	6.61063E-25	1.20764E-25	2.20613E-26	4.0302E-27	1.15335E-08	6.29396E-08	3.42062E-07
14	2.40361E-18	4.39096E-19	8.02146E-20	1.46537E-20	0.014693373	0.038222088	0.07479234
15	1.94153E-12	3.54683E-13	6.47942E-14	1.18367E-14	0.018489994	0.008147811	0.003528701
16	2.31971E-13	4.23769E-14	7.74149E-15	1.41423E-15	0.017817783	0.008095881	0.003553937
17	3.21842E-13	5.87947E-14	1.07407E-14	1.96213E-15	0.044521221	0.020498385	0.009051869
18	1.03331E-16	1.88766E-17	3.4484E-18	6.29959E-19	0.077696699	0.133467658	0.169224858
19	2.40496E-12	4.39347E-13	8.02609E-14	1.46622E-14	0.001724542	0.000745822	0.000320388
20	5.27388E-13	9.63442E-14	1.76003E-14	3.21525E-15	0.024483511	0.01097086	0.004786328
21	3.15961E-17	5.77202E-18	1.05444E-18	1.92627E-19	0.052471968	0.089979139	0.108953045
22	6.57806E-25	1.20169E-25	2.19526E-26	4.01034E-27	1.148E-08	6.26725E-08	3.40919E-07
23	1.1374E-20	2.07782E-21	3.79579E-22	6.9342E-23	0.000165298	0.000736607	0.002795113
24	9.22526E-18	3.94299E-18	1.68528E-18	7.20311E-19	1.21994E-09	2.85425E-09	6.67798E-09
25	2.13239E-26	3.89547E-27	7.11631E-28	1.30002E-28	3.72741E-10	2.03925E-09	1.11484E-08
26	1.54167E-15	6.5893E-16	2.81635E-16	1.20374E-16	2.0387E-07	4.76985E-07	1.11598E-06
27	1.98917E-10	3.6341E-11	6.63904E-12	1.21284E-12	0.000997011	0.000427187	0.000182778
28	1.30274E-11	2.37989E-12	4.34762E-13	7.94231E-14	0.033070891	0.014390845	0.006198477
29	3.61848E-13	6.6103E-14	1.20758E-14	2.20602E-15	0.325538284	0.159814399	0.072688337
30	4.59207E-13	8.38887E-14	1.53249E-14	2.79958E-15	0.036739688	0.016618019	0.007280263
31	2.37366E-16	4.33623E-17	7.92149E-18	1.44711E-18	0.053774853	0.057192363	0.042721047
32	2.98046E-13	5.44476E-14	9.94657E-15	1.81705E-15	0.216688132	0.102792393	0.046010849
33	6.98358E-14	1.27577E-14	2.3306E-15	4.25757E-16	0.617737267	0.339811142	0.163395123
34	1.9037E-16	3.47771E-17	6.35313E-18	1.1606E-18	0.297544545	0.40950539	0.375603784
35	2.50424E-18	4.57478E-19	8.35727E-20	1.52672E-20	0.018147161	0.052951853	0.124651807
36	1.24501E-20	5.32134E-21	2.27441E-21	9.72111E-22	1.6464E-12	3.85201E-12	9.0124E-12

ITEM	$\theta=2$	$\theta=2.5$	$\theta=3$	$\theta=3.5$	$\theta=4$	$\theta=4.5$	$\theta=5$
1	3.80101E-06	1.62463E-06	6.94395E-07	2.96794E-07	1.26854E-07	5.42189E-08	2.31738E-08
2	2.46636E-04	1.05437E-04	4.50689E-05	1.92637E-05	8.23369E-06	3.51921E-06	1.50416E-06
3	2.03848E-04	8.71524E-05	3.72546E-05	1.59239E-05	6.80624E-06	2.9091E-06	1.24339E-06
4	1.64319E-03	7.03059E-04	0.000300631	0.000128518	5.49348E-05	2.34806E-05	1.00361E-05
5	8.79966E-04	3.76359E-04	0.000160907	6.87818E-05	2.93997E-05	1.25661E-05	5.37096E-06
6	6.28427E-03	2.70633E-03	0.001160468	0.000496686	0.000212415	9.0812E-05	3.88184E-05
7	4.16247E-03	8.50034E-03	0.012458302	0.012495572	0.008834489	0.004888725	0.002358112
8	1.14991E-01	6.19795E-02	0.029502428	0.013227001	0.005772038	0.0024892	0.001068002
9	1.64699E-05	8.12961E-05	0.000361306	0.001365973	0.004264003	0.011108896	0.024413759
10	7.54967E-02	4.99195E-02	0.026551669	0.012554913	0.005611638	0.002445521	0.001054018
11	2.40002E-07	1.29527E-06	6.8638E-06	3.49491E-05	0.000164221	0.000671408	0.00226352
12	1.37317E-04	3.20942E-04	0.000749073	0.001742662	0.004023728	0.009130705	0.01992342
13	1.84155E-06	9.70544E-06	4.88619E-05	0.000224957	0.000894221	0.002946076	0.008056452
14	1.04135E-01	9.90519E-02	0.067186084	0.036260361	0.017269786	0.00774454	0.003379935
15	1.51677E-03	6.49862E-04	0.000278047	0.000118894	5.08263E-05	2.17255E-05	9.28609E-06
16	1.53660E-03	6.60017E-04	0.000282697	0.000120938	5.17103E-05	2.21053E-05	9.44875E-06
17	3.92384E-03	1.68729E-03	0.000723043	0.000309381	0.000132296	5.65565E-05	2.4175E-05
18	1.49971E-01	9.70245E-02	0.050997709	0.023976227	0.010688796	0.004652816	0.002004374
19	1.37235E-04	5.87101E-05	2.51034E-05	1.07313E-05	4.58703E-06	1.96061E-06	8.38002E-07
20	2.06389E-03	8.85481E-04	0.00037908	0.000162136	6.93194E-05	2.96317E-05	1.26657E-05
21	9.10244E-02	5.62393E-02	0.028777145	0.013350525	0.005915693	0.002568211	0.001105072
22	1.83922E-06	9.73816E-06	4.95051E-05	0.000232067	0.000947891	0.003221312	0.009063055
23	8.75860E-03	2.28966E-02	0.050498758	0.090734352	0.122264404	0.115548082	0.078545545
24	1.56242E-08	3.65553E-08	8.55269E-08	2.00103E-07	4.68171E-07	1.09535E-06	2.56266E-06
25	6.08411E-08	3.30697E-07	1.78085E-06	9.3913E-06	4.73406E-05	0.000218463	0.000871419
26	2.61098E-06	6.10863E-06	1.42912E-05	3.34317E-05	7.81919E-05	0.000182795	0.000426874
27	7.81568E-05	3.34116E-05	1.42817E-05	6.1044E-06	2.60914E-06	1.11518E-06	4.76645E-07
28	2.65808E-03	1.13771E-03	0.000486564	0.000208018	8.89191E-05	3.80069E-05	1.6245E-05
29	3.19232E-02	1.38052E-02	0.005930249	0.002540126	0.001086682	0.000464644	0.000198628
30	3.14496E-03	1.35035E-03	0.000578284	0.000247372	0.000105768	4.52134E-05	1.9326E-05
31	2.45075E-02	1.20484E-02	0.005483839	0.002409155	0.001041985	0.00044763	0.00019174
32	2.00638E-02	8.64981E-03	0.00371073	0.001588524	0.000679415	0.000290474	0.000124168
33	7.35962E-02	3.21818E-02	0.013890734	0.005962146	0.002552897	0.001091982	0.000466881
34	2.46212E-01	1.30055E-01	0.061262919	0.027332967	0.011901898	0.005127897	0.002199257
35	2.30827E-01	3.14032E-01	0.296778804	0.201225173	0.108685042	0.051793942	0.023233847
36	2.10860E-11	4.93340E-11	1.15425E-10	2.70055E-10	6.31836E-10	1.47828E-09	3.45868E-09

ITEM	$\theta=5.5$	$\theta=6$	$\theta=6.5$	$\theta=7$	$\theta=7.5$	$\theta=8$	$\theta=8.5$
1	9.90479E-09	4.23343E-09	1.80942E-09	7.73369E-10	3.30548E-10	1.4128E-10	6.03849E-11
2	6.42897E-07	2.74782E-07	1.17445E-07	5.01976E-08	2.14551E-08	9.17018E-09	3.91945E-09
3	5.31442E-07	2.27145E-07	9.70847E-08	4.14952E-08	1.77356E-08	7.58041E-09	3.23996E-09
4	4.28957E-06	1.83342E-06	7.83627E-07	3.34932E-07	1.43154E-07	6.11859E-08	2.61516E-08
5	2.29562E-06	9.8118E-07	4.19369E-07	1.79244E-07	7.6611E-08	3.27445E-08	1.39954E-08
6	1.65922E-05	7.09187E-06	3.03118E-06	1.29557E-06	5.53743E-07	2.36677E-07	1.01159E-07
7	0.001063713	0.000465443	0.000200959	8.62657E-05	3.69396E-05	1.5801E-05	6.75583E-06
8	0.000457227	0.000195562	8.36106E-05	3.57408E-05	1.52769E-05	6.52969E-06	2.7909E-06
9	0.043653587	0.058433168	0.054827181	0.037058669	0.019989135	0.009520356	0.004269611
10	0.000452117	0.000193537	8.27744E-05	3.53887E-05	1.51274E-05	6.46597E-06	2.7637E-06
11	0.006168557	0.013515496	0.022873564	0.02777384	0.023405053	0.014570251	0.007489194
12	0.039949799	0.067851256	0.087591768	0.079878601	0.052935429	0.028232503	0.013371522
13	0.01858993	0.03546433	0.051952953	0.053950529	0.039634752	0.022514033	0.011012914
14	0.001457667	0.00062543	0.000267758	0.000114524	4.89636E-05	2.09304E-05	8.94639E-06
15	3.96905E-06	1.69643E-06	7.25078E-07	3.09908E-07	1.32459E-07	5.66145E-08	2.41977E-08
16	4.03863E-06	1.72619E-06	7.37797E-07	3.15344E-07	1.34782E-07	5.76077E-08	2.46222E-08
17	1.03331E-05	4.41657E-06	1.88771E-06	8.06833E-07	3.44851E-07	1.47394E-07	6.29979E-08
18	0.000859587	0.000367929	0.000157354	6.7273E-05	2.87565E-05	1.22915E-05	5.25365E-06
19	3.58174E-07	1.53089E-07	6.5432E-08	2.79665E-08	1.19532E-08	5.10896E-09	2.18363E-09
20	5.41358E-06	2.31386E-06	9.88975E-07	4.22701E-07	1.80668E-07	7.72198E-08	3.30047E-08
21	0.000473681	0.000202706	8.66846E-05	3.70584E-05	1.58408E-05	6.77082E-06	2.89398E-06
22	0.021487323	0.042418386	0.065228214	0.071781565	0.055498428	0.032598334	0.016232485
23	0.042511852	0.020282475	0.009103397	0.003974533	0.001714393	0.000735636	0.000314949
24	5.99531E-06	1.40246E-05	3.27996E-05	7.66679E-05	0.000178984	0.00041662	0.000963182
25	0.002882103	0.007909312	0.018312752	0.035088949	0.051722184	0.054102533	0.039995863
26	0.000994355	0.002302714	0.005260899	0.01165486	0.024132414	0.04354895	0.061716841
27	2.03724E-07	8.70744E-08	3.72167E-08	1.59069E-08	6.7988E-09	2.90589E-09	1.24201E-09
28	6.94336E-06	2.96769E-06	1.26843E-06	5.42143E-07	2.31719E-07	9.90396E-08	4.23308E-08
29	8.49023E-05	3.62894E-05	1.55108E-05	6.62953E-06	2.83355E-06	1.2111E-06	5.17638E-07
30	8.26042E-06	3.53065E-06	1.50905E-06	6.44989E-07	2.75676E-07	1.17828E-07	5.0361E-08
31	8.20285E-05	3.5074E-05	1.49936E-05	6.40892E-06	2.73934E-06	1.17084E-06	5.00436E-07
32	5.30736E-05	2.26848E-05	9.69588E-06	4.14416E-06	1.77127E-06	7.57063E-07	3.23578E-07
33	0.000199579	8.53076E-05	3.64625E-05	1.55847E-05	6.66112E-06	2.84705E-06	1.21687E-06
34	0.000941371	0.000402606	0.000172125	7.35769E-05	3.14492E-05	1.34421E-05	5.74536E-06
35	0.010141351	0.004373943	0.001876746	0.000803478	0.00034366	0.000146929	6.28076E-05
36	8.09213E-09	1.89328E-08	4.42963E-08	1.03638E-07	2.42476E-07	5.67303E-07	1.32725E-06

ITEM	$\theta=9$	$\theta=9.5$	$\theta=10$	$\theta=10.5$	$\theta=11$	$\theta=11.5$	$\theta=12$
1	2.58093E-11	1.10312E-11	4.71488E-12	2.0152E-12	8.61321E-13	3.68139E-13	1.57347E-13
2	1.67522E-09	7.1601E-10	3.06032E-10	1.30802E-10	5.59063E-11	2.38951E-11	1.0213E-11
3	1.3848E-09	5.91881E-10	2.52977E-10	1.08126E-10	4.62143E-11	1.97526E-11	8.44249E-12
4	1.11775E-08	4.77742E-09	2.04193E-09	8.72746E-10	3.73022E-10	1.59434E-10	6.81443E-11
5	5.98181E-09	2.5567E-09	1.09277E-09	4.67062E-10	1.99628E-10	8.53235E-11	3.64683E-11
6	4.32365E-08	1.84798E-08	7.8985E-09	3.37592E-09	1.44291E-09	6.16718E-10	2.63593E-10
7	2.88795E-06	1.23442E-06	5.27621E-07	2.25514E-07	9.63881E-08	4.11976E-08	1.76084E-08
8	1.19287E-06	5.09849E-07	2.17916E-07	9.314E-08	3.98092E-08	1.7015E-08	7.2724E-09
9	0.001863448	0.000803665	0.000344825	0.000147626	6.31419E-05	2.69958E-05	1.15398E-05
10	1.18125E-06	5.04883E-07	2.15794E-07	9.22329E-08	3.94215E-08	1.68493E-08	7.20158E-09
11	0.003482283	0.001544611	0.000670875	0.000288727	0.000123771	5.29682E-05	2.26515E-05
12	0.005981429	0.00260762	0.001124062	0.000482195	0.000206419	8.82848E-05	3.77448E-05
13	0.005000822	0.002194679	0.000948794	0.000407517	0.000174543	7.46688E-05	3.19266E-05
14	3.82389E-06	1.63439E-06	6.98564E-07	2.98575E-07	1.27615E-07	5.45442E-08	2.33129E-08
15	1.03424E-08	4.42048E-09	1.88937E-09	8.07539E-10	3.45152E-10	1.47522E-10	6.30529E-11
16	1.05239E-08	4.49803E-09	1.92251E-09	8.21706E-10	3.51207E-10	1.5011E-10	6.41591E-11
17	2.69261E-08	1.15085E-08	4.9189E-09	2.1024E-09	8.98591E-10	3.84069E-10	1.64156E-10
18	2.24549E-06	9.59756E-07	4.10212E-07	1.7533E-07	7.49383E-08	3.20296E-08	1.36898E-08
19	9.33312E-10	3.98909E-10	1.70499E-10	7.28733E-11	3.1147E-11	1.33126E-11	5.68997E-12
20	1.41066E-08	6.02935E-09	2.57702E-09	1.10145E-09	4.70773E-10	2.01214E-10	8.60016E-11
21	1.23693E-06	5.28682E-07	2.25966E-07	9.65807E-08	4.12798E-08	1.76435E-08	7.54105E-09
22	0.007433479	0.003274638	0.001418013	0.000609484	0.000261127	0.000111724	4.77731E-05
23	0.00013471	5.75944E-05	2.46198E-05	1.05234E-05	4.49794E-06	1.9225E-06	8.21703E-07
24	0.002192044	0.004814277	0.009786221	0.017049639	0.022869451	0.021759061	0.014912428
25	0.0228108	0.011181757	0.005082565	0.002231547	0.000964921	0.000414478	0.000177531
26	0.062601677	0.04524901	0.02545837	0.012392822	0.005614755	0.002461651	0.001063749
27	5.30853E-10	2.26893E-10	9.69769E-11	4.14492E-11	1.77159E-11	7.57199E-12	3.23636E-12
28	1.80927E-08	7.73304E-09	3.3052E-09	1.41268E-09	6.03799E-10	2.58071E-10	1.10303E-10
29	2.21245E-07	9.45628E-08	4.04173E-08	1.72749E-08	7.3835E-09	3.1558E-09	1.34883E-09
30	2.15249E-08	9.20002E-09	3.9322E-09	1.68067E-09	7.18341E-10	3.07028E-10	1.31228E-10
31	2.13893E-07	9.14207E-08	3.90743E-08	1.67009E-08	7.13816E-09	3.05094E-09	1.30401E-09
32	1.38301E-07	5.91117E-08	2.52651E-08	1.07986E-08	4.61546E-09	1.97271E-09	8.4316E-10
33	5.20104E-07	2.22299E-07	9.50134E-08	4.06099E-08	1.73572E-08	7.41868E-09	3.17084E-09
34	2.45565E-06	1.04958E-06	4.48602E-07	1.91738E-07	8.19514E-08	3.5027E-08	1.4971E-08
35	2.68462E-05	1.14747E-05	4.90448E-06	2.09624E-06	8.95962E-07	3.82946E-07	1.63676E-07
36	3.10504E-06	7.26329E-06	1.69857E-05	3.96973E-05	9.2641E-05	0.000215459	0.000497144

B TEST MADEE BİLGİ FONKSİYONLARI

ITEM	$\theta=0$	$\theta=-0.5$	$\theta=-1$	$\theta=-1.5$	$\theta=-2$	$\theta=-2.5$	$\theta=-3$
1	0.000749524	0.001746922	0.00405081	0.009282566	0.020702807	0.043474329	0.080490891
2	0.002897352	0.006624769	0.014701645	0.030539679	0.055335886	0.078205662	0.076569427
3	0.003256543	0.007338725	0.015765392	0.030570505	0.048602316	0.055910209	0.042499887
4	0.008145272	0.018594407	0.041116261	0.084749203	0.151228853	0.208369934	0.197398453
5	0.000771536	0.001799773	0.004181741	0.009627122	0.021698952	0.046634292	0.090515697
6	0.00730402	0.015801628	0.03109528	0.050808315	0.060957983	0.048804775	0.025794115
7	8.20364E-09	1.49954E-09	2.74008E-10	5.00616E-11	9.14574E-12	1.67079E-12	3.05225E-13
8	0.132036649	0.095671816	0.046524898	0.016086788	0.004259111	0.000942336	0.000188436
9	0.29937866	0.112302834	0.03130184	0.007128226	0.001446088	0.000276928	5.16438E-05
10	0.308831857	0.092993003	0.021776079	0.004460478	0.000857193	0.000160075	2.95209E-05
11	0.031585571	0.007271071	0.001485447	0.000285492	5.33297E-05	9.83674E-06	1.80445E-06
12	0.004422628	0.000946946	0.000186505	3.52405E-05	6.53337E-06	1.20114E-06	2.20025E-07
13	3.93533E-05	7.35344E-06	1.35658E-06	2.48869E-07	4.5546E-08	8.32685E-09	1.52166E-09
14	0.000318796	0.00074302	0.001722934	0.003948161	0.008805542	0.018490972	0.034235256
15	0.000693897	0.001618414	0.003758996	0.008646647	0.019451839	0.041628821	0.080099394
16	0.010104387	0.022974722	0.050347424	0.101799222	0.175021047	0.22710643	0.199569621
17	0.01586049	0.025862999	0.030893088	0.024516719	0.012764524	0.004599295	0.001244832
18	0.001188439	0.002749735	0.006269273	0.013823194	0.028318175	0.049930734	0.067463588
19	0.000853434	0.00195388	0.004348615	0.009089656	0.016668757	0.024006753	0.023992665
20	0.013384354	0.006452529	0.002159094	0.000551509	0.000119015	2.3479E-05	4.43772E-06
21	0.010125994	0.018410321	0.02611265	0.025513626	0.016115712	0.006747994	0.002024516
22	5.11351E-14	9.34146E-15	1.70651E-15	3.11749E-16	5.69507E-17	1.04038E-17	1.90059E-18
23	6.02103E-09	1.10038E-09	2.01055E-10	3.67317E-11	6.71042E-12	1.22589E-12	2.23948E-13
24	1.32984E-05	2.45764E-06	4.51205E-07	8.26028E-08	1.51038E-08	2.76026E-09	5.04332E-10
25	1.47415E-14	2.693E-15	4.91961E-16	8.98721E-17	1.6418E-17	2.99926E-18	5.47909E-19
26	0.009102284	0.001957784	0.000386514	7.31144E-05	1.35618E-05	2.49383E-06	4.56864E-07
27	0.145279049	0.300234848	0.538614561	0.749323801	0.720652601	0.457480151	0.198051198
28	0.03468832	0.077948159	0.166405731	0.318605848	0.495562683	0.554155041	0.411725135
29	0.111753285	0.222016244	0.36949062	0.457013848	0.382005875	0.2133667	0.083591531
30	0.386353721	0.438164062	0.337642616	0.180647715	0.070917048	0.021346101	0.005194954
31	0.062715415	0.103474171	0.125975453	0.102587455	0.055085626	0.020510537	0.005710157
32	0.560288779	0.826230893	0.856194726	0.583462149	0.266499732	0.086926055	0.021997301
33	0.338472154	0.415675335	0.344597087	0.191003608	0.074346954	0.021651213	0.005099517
34	0.165399505	0.17502543	0.120828538	0.055042253	0.017703714	0.004413638	0.000940122
35	0.392810368	0.660403763	0.82748411	0.697099131	0.386475931	0.147802678	0.041981855
36	0.001693832	0.000375644	7.51653E-05	1.43016E-05	2.6594E-06	4.89554E-07	8.97264E-08

ITEM	$\theta=-3.5$	$\theta=-4$	$\theta=-4.5$	$\theta=-5$	$\theta=-5.5$	$\theta=-6$	$\theta=-6.5$
1	0.117995674	0.121296632	0.081965476	0.037161058	0.012049193	0.003036573	0.000651783
2	0.049203072	0.021346579	0.006683529	0.001644421	0.000348156	6.80703E-05	1.28143E-05
3	0.021262743	0.007415311	0.001960576	0.00043241	8.62801E-05	1.63948E-05	3.04686E-06
4	0.12269612	0.051719403	0.01583072	0.003836078	0.000805238	0.000156775	2.94562E-05
5	0.144221368	0.166695766	0.127907827	0.06499986	0.023121684	0.006224043	0.001389267
6	0.009478223	0.002613006	0.000591869	0.000119776	2.29121E-05	4.27081E-06	7.87016E-07
7	5.57591E-14	1.01862E-14	1.86083E-15	3.39939E-16	6.21005E-17	1.13446E-17	2.07245E-18
8	3.58466E-05	6.66532E-06	1.22695E-06	2.24876E-07	4.11383E-08	7.51971E-09	1.37406E-09
9	9.51876E-06	1.74556E-06	3.19404E-07	5.839E-08	1.067E-08	1.94945E-09	3.56149E-10
10	5.41484E-06	9.90909E-07	1.81155E-07	3.31042E-08	6.04834E-09	1.10498E-09	2.01865E-10
11	3.30225E-07	6.03718E-08	1.10324E-08	2.01569E-09	3.68251E-10	6.72744E-11	1.22899E-11
12	4.02415E-08	7.35506E-09	1.34392E-09	2.45532E-10	4.48558E-11	8.19446E-12	1.49699E-12
13	2.78019E-10	5.0792E-11	9.279E-12	1.69512E-12	3.09668E-13	5.65708E-14	1.03344E-14
14	0.050187196	0.051591195	0.034862442	0.015805743	0.005124893	0.001291548	0.000277223
15	0.125633132	0.141931841	0.106164066	0.052709455	0.018394909	0.004883488	0.001080921
16	0.115206636	0.045588104	0.013288462	0.003118195	0.000643057	0.000124126	2.32311E-05
17	0.000278431	5.59583E-05	1.0669E-05	1.98572E-06	3.65686E-07	6.70351E-08	1.22642E-08
18	0.062338827	0.037798773	0.015596994	0.004695421	0.001125292	0.000234766	4.55704E-05
19	0.01558797	0.006730557	0.002076453	0.000503573	0.000105628	2.05524E-05	3.8603E-06
20	8.22761E-07	1.51262E-07	2.77083E-08	5.0677E-09	9.26238E-10	1.69243E-10	3.09203E-11
21	0.000481342	9.98343E-05	1.93178E-05	3.61925E-06	6.68428E-07	1.22684E-07	2.24571E-08
22	3.47202E-19	6.34274E-20	1.1587E-20	2.11673E-21	3.86688E-22	7.06407E-23	1.29047E-23
23	4.09113E-14	7.47374E-15	1.36531E-15	2.49417E-16	4.5564E-17	8.32369E-18	1.52058E-18
24	9.21387E-11	1.68325E-11	3.07503E-12	5.61754E-13	1.02622E-13	1.87472E-14	3.42477E-15
25	1.00093E-19	1.82851E-20	3.34035E-21	6.1022E-22	1.11476E-22	2.03646E-23	3.72023E-24
26	8.35616E-08	1.5273E-08	2.79072E-09	5.09861E-10	9.31459E-11	1.70163E-11	3.10859E-12
27	0.062316146	0.015427697	0.003280454	0.000642867	0.000121153	2.24336E-05	4.12213E-06
28	0.204776713	0.072130049	0.019358791	0.004319021	0.000867415	0.000165347	3.07726E-05
29	0.024467769	0.005782553	0.001198538	0.000231968	4.34691E-05	8.0291E-06	1.47374E-06
30	0.001097784	0.000214636	4.0413E-05	7.48043E-06	1.3743E-06	2.5167E-07	4.60232E-08
31	0.001301327	0.000264197	5.06161E-05	9.44129E-06	1.74034E-06	3.19158E-07	5.84008E-08
32	0.004732893	0.000932813	0.000176251	3.26733E-05	6.0066E-06	1.10027E-06	2.01231E-07
33	0.001054969	0.000203993	3.82108E-05	7.05655E-06	1.29513E-06	2.37069E-07	4.33451E-08
34	0.000184302	3.47355E-05	6.432E-06	1.18187E-06	2.16445E-07	3.95827E-08	7.23432E-09
35	0.009689167	0.001980379	0.000380627	7.10993E-05	1.31142E-05	2.40564E-06	4.40245E-07
36	1.64144E-08	3.00041E-09	5.4826E-10	1.00168E-10	1.82997E-11	3.34308E-12	6.10723E-13

ITEM	$\theta=-7$	$\theta=-7.5$	$\theta=-8$	$\theta=-8.5$	$\theta=-9$	$\theta=-9.5$	$\theta=-10$
1	0.000128307	2.42297E-05	4.49059E-06	8.25454E-07	1.51197E-07	2.76524E-08	5.05403E-09
2	2.37165E-06	4.35692E-07	7.97844E-08	1.45901E-08	2.66652E-09	4.87215E-10	8.90123E-11
3	5.60736E-07	1.02762E-07	1.87982E-08	3.43607E-09	6.27862E-10	1.14711E-10	2.09565E-11
4	5.44708E-06	1.00031E-06	1.83149E-07	3.349E-08	6.12052E-09	1.1183E-09	2.04308E-10
5	0.000278996	5.31771E-05	9.8962E-06	1.82237E-06	3.34057E-07	6.11157E-08	1.11717E-08
6	1.44311E-07	2.64052E-08	4.82704E-09	8.82068E-10	1.61158E-10	2.94421E-11	5.37865E-12
7	3.78598E-19	6.91628E-20	1.26348E-20	2.30814E-21	4.21654E-22	7.70283E-23	1.40716E-23
8	2.51043E-10	4.58631E-11	8.3785E-12	1.53061E-12	2.79615E-13	5.10806E-14	9.33148E-15
9	6.50633E-11	1.1886E-11	2.17136E-12	3.96667E-13	7.24638E-14	1.32378E-14	2.4183E-15
10	3.68774E-11	6.73685E-12	1.2307E-12	2.24826E-13	4.10716E-14	7.50302E-15	1.37066E-15
11	2.24515E-12	4.10148E-13	7.49265E-14	1.36877E-14	2.50049E-15	4.56793E-16	8.34476E-17
12	2.73473E-13	4.99584E-14	9.12648E-15	1.66724E-15	3.04574E-16	5.564E-17	1.01644E-17
13	1.88791E-15	3.44886E-16	6.30043E-17	1.15097E-17	2.10261E-18	3.84108E-19	7.01695E-20
14	5.45729E-05	1.03057E-05	1.90999E-06	3.51091E-07	6.43087E-08	1.17614E-08	2.14964E-09
15	0.000216132	4.11113E-05	7.64383E-06	1.40704E-06	2.5788E-07	4.71757E-08	8.62327E-09
16	4.28857E-06	7.86976E-07	1.44043E-07	2.63357E-08	4.81275E-09	8.79333E-10	1.60648E-10
17	2.24186E-09	4.09657E-10	7.48453E-11	1.36735E-11	2.49795E-12	4.56334E-13	8.3364E-14
18	8.55051E-06	1.58022E-06	2.90119E-07	5.31125E-08	9.71154E-09	1.77481E-09	3.24279E-10
19	7.1374E-07	1.31063E-07	2.39959E-08	4.38777E-09	8.01889E-10	1.46516E-10	2.67677E-11
20	5.64878E-12	1.03195E-12	1.88519E-13	3.4439E-14	6.29137E-15	1.14932E-15	2.09959E-16
21	4.10602E-09	7.5037E-10	1.371E-10	2.50474E-11	4.57582E-12	8.35927E-13	1.52709E-13
22	2.35746E-24	4.30664E-25	7.86743E-26	1.43723E-26	2.62556E-27	4.79641E-28	8.76215E-29
23	2.77782E-19	5.07457E-20	9.2703E-21	1.69351E-21	3.09373E-22	5.65167E-23	1.03246E-23
24	6.25642E-16	1.14293E-16	2.08793E-17	3.81425E-18	6.96793E-19	1.27291E-19	2.32537E-20
25	6.79617E-25	1.24153E-25	2.26805E-26	4.14331E-27	7.56906E-28	1.38273E-28	2.52598E-29
26	5.67883E-13	1.03742E-13	1.89517E-14	3.46213E-15	6.32467E-16	1.1554E-16	2.1107E-17
27	7.54918E-07	1.38057E-07	2.52319E-08	4.61031E-09	8.42288E-10	1.53876E-10	2.81107E-11
28	5.66686E-06	1.0388E-06	1.9005E-07	3.47405E-08	6.34815E-09	1.15982E-09	2.11889E-10
29	2.69773E-07	4.93254E-08	9.0142E-09	1.64699E-09	3.00895E-10	5.49695E-11	1.0042E-11
30	8.41132E-09	1.53689E-09	2.80783E-10	5.12957E-11	9.37091E-12	1.7119E-12	3.12733E-13
31	1.06763E-08	1.95095E-09	3.56449E-10	6.51202E-11	1.18965E-11	2.1733E-12	3.97023E-13
32	3.67795E-08	6.72035E-09	1.22779E-09	2.24304E-10	4.09768E-11	7.48576E-12	1.36751E-12
33	7.92125E-09	1.44729E-09	2.64411E-10	4.83044E-11	8.82442E-12	1.61207E-12	2.94495E-13
34	1.32183E-09	2.41495E-10	4.41182E-11	8.0597E-12	1.47237E-12	2.68975E-13	4.91368E-14
35	8.04855E-08	1.4708E-08	2.68725E-09	4.90939E-10	8.96877E-11	1.63845E-11	2.99315E-12
36	1.11568E-13	2.03815E-14	3.72332E-15	6.80182E-16	1.24257E-16	2.26994E-17	4.14676E-18

ITEM	$\theta=-10.5$	$\theta=-11$	$\theta=-11.5$	$\theta=-12$	$\theta=0.5$	$\theta=1$	$\theta=1.5$
1	9.2347E-10	1.68716E-10	3.08224E-11	5.63078E-12	0.000320882	0.000137245	5.86779E-05
2	1.62615E-11	2.97071E-12	5.42697E-13	9.91409E-14	0.001250684	0.000536832	0.000229866
3	3.82844E-12	6.99391E-13	1.27766E-13	2.33405E-14	0.001414657	0.000608872	0.000261018
4	3.73246E-11	6.8186E-12	1.24564E-12	2.27556E-13	0.00351846	0.00151068	0.00064694
5	2.04141E-09	3.72971E-10	6.81383E-11	1.24479E-11	0.000330184	0.000141201	6.03653E-05
6	9.82588E-13	1.79501E-13	3.27916E-14	5.99043E-15	0.003230906	0.001401582	0.000602885
7	2.57063E-24	4.69606E-25	8.57883E-26	1.56719E-26	4.48447E-08	2.44688E-07	1.32938E-06
8	1.70469E-15	3.11415E-16	5.68898E-17	1.03927E-17	0.120839248	0.078973606	0.041630964
9	4.41779E-16	8.07048E-17	1.47433E-17	2.69332E-18	0.547746883	0.655531241	0.525280955
10	2.50395E-16	4.57424E-17	8.3563E-18	1.52654E-18	0.690356324	0.943492969	0.809965561
11	1.52443E-17	2.78486E-18	5.08741E-19	9.29376E-20	0.112298414	0.301680144	0.573901274
12	1.85685E-18	3.39211E-19	6.19676E-20	1.13203E-20	0.017862851	0.058824663	0.153213047
13	1.28187E-20	2.34173E-21	4.27791E-22	7.81495E-23	0.000204687	0.00100277	0.004392656
14	3.9278E-10	7.17601E-11	1.31097E-11	2.39494E-12	0.000136481	5.83746E-05	2.49575E-05
15	1.57571E-09	2.87885E-10	5.25937E-11	9.60809E-12	0.000296977	0.000127005	5.42966E-05
16	2.93482E-11	5.36144E-12	9.79441E-13	1.78926E-13	0.004372322	0.001878699	0.000804801
17	1.52291E-14	2.78207E-15	5.08233E-16	9.28448E-17	0.008066984	0.00373026	0.001650341
18	5.9244E-11	1.08231E-11	1.97721E-12	3.61201E-13	0.000510387	0.000218593	9.35111E-05
19	4.89011E-12	8.93345E-13	1.63198E-13	2.98134E-14	0.000368191	0.000158001	6.76474E-05
20	3.83556E-17	7.00686E-18	1.28002E-18	2.33836E-19	0.01808721	0.015994208	0.010157693
21	2.78972E-14	5.09631E-15	9.31001E-16	1.70077E-16	0.004864922	0.002190079	0.000957415
22	1.60068E-29	2.92415E-30	5.34188E-31	9.75862E-32	2.79912E-13	1.53221E-12	8.38695E-12
23	1.88611E-24	3.44556E-25	6.29441E-26	1.14987E-26	3.29275E-08	1.79841E-07	9.79313E-07
24	4.24803E-21	7.76036E-22	1.41767E-22	2.58983E-23	7.08809E-05	0.000365319	0.001754716
25	4.6145E-30	8.42984E-31	1.53998E-31	2.81325E-32	8.06945E-14	4.41718E-13	2.41791E-12
26	3.85586E-18	7.04394E-19	1.2868E-19	2.35074E-20	0.036500956	0.119125999	0.307264184
27	5.13533E-12	9.38133E-13	1.7138E-13	3.13079E-14	0.065622609	0.028730864	0.012407862
28	3.87089E-11	7.07147E-12	1.29183E-12	2.35994E-13	0.015087711	0.006497325	0.002785995
29	1.8345E-12	3.3513E-13	6.12221E-14	1.11842E-14	0.05142578	0.022702738	0.009840009
30	5.71307E-14	1.04367E-14	1.9066E-15	3.483E-16	0.247292493	0.128994699	0.060401122
31	7.25289E-14	1.32497E-14	2.42047E-15	4.42176E-16	0.031703286	0.014617596	0.00645876
32	2.4982E-13	4.56375E-14	8.33712E-15	1.52304E-15	0.301580745	0.143369191	0.064232302
33	5.37989E-14	9.82806E-15	1.7954E-15	3.27987E-16	0.20433936	0.10311704	0.04750915
34	8.97638E-15	1.63982E-15	2.99564E-16	5.47249E-17	0.110157029	0.058641413	0.027721763
35	5.46794E-13	9.98892E-14	1.82479E-14	3.33356E-15	0.196581869	0.090206397	0.03977213
36	7.57535E-19	1.38388E-19	2.52809E-20	4.61834E-21	0.006347478	0.017966521	0.03534589

ITEM	$\theta=2$	$\theta=2.5$	$\theta=3$	$\theta=3.5$	$\theta=4$	$\theta=4.5$	$\theta=5$
1	2.50829E-05	1.07213E-05	4.58254E-06	1.95865E-06	8.37157E-07	3.57812E-07	1.52933E-07
2	9.83239E-05	4.20388E-05	1.79705E-05	7.68127E-06	3.28316E-06	1.40328E-06	5.99781E-07
3	0.000111705	4.77702E-05	2.04224E-05	8.72964E-06	3.73132E-06	1.59484E-06	6.8166E-07
4	0.00027674	0.000118324	5.0581E-05	2.16204E-05	9.24107E-06	3.9498E-06	1.6882E-06
5	2.58034E-05	1.10292E-05	4.71409E-06	2.01488E-06	8.61187E-07	3.68083E-07	1.57323E-07
6	0.000258385	0.000110566	4.72809E-05	2.02128E-05	8.63998E-06	3.69298E-06	1.57845E-06
7	7.15139E-06	3.76221E-05	0.000188626	0.000860315	0.003349101	0.010539665	0.026137548
8	0.019589459	0.008735643	0.003803015	0.00163836	0.000702632	0.00030075	0.000128624
9	0.313012823	0.156773665	0.071964552	0.031734035	0.013747597	0.005909995	0.002532273
10	0.497955383	0.252776291	0.116699131	0.051586197	0.022371083	0.009621446	0.004123312
11	0.73345765	0.62846292	0.392694786	0.201925963	0.093877173	0.041635197	0.018082315
12	0.308336011	0.453998585	0.462202629	0.331682695	0.185355136	0.089872037	0.040639083
13	0.016197212	0.048489535	0.117180058	0.22399324	0.317511006	0.313590527	0.220205567
14	1.06685E-05	4.56011E-06	1.9491E-06	8.33076E-07	3.56069E-07	1.52188E-07	6.50473E-08
15	2.32095E-05	9.92046E-06	4.24021E-06	1.81233E-06	7.74617E-07	3.31081E-07	1.41508E-07
16	0.000344315	0.000147226	6.29372E-05	2.69022E-05	1.14987E-05	4.91476E-06	2.10064E-06
17	0.000715976	0.000307984	0.000131998	5.64837E-05	2.4154E-05	1.03259E-05	4.41383E-06
18	3.99828E-05	1.70919E-05	7.30578E-06	3.12267E-06	1.33469E-06	5.70464E-07	2.43824E-07
19	2.89345E-05	1.23709E-05	5.28817E-06	2.26036E-06	9.66129E-07	4.1294E-07	1.76496E-07
20	0.005263816	0.002455794	0.00109084	0.000474071	0.00020408	8.7494E-05	3.74451E-05
21	0.000413203	0.000177345	7.59344E-05	3.248E-05	1.38869E-05	5.93626E-06	2.53738E-06
22	4.59054E-11	2.51224E-10	1.37441E-09	7.51333E-09	4.09981E-08	2.22777E-07	1.19888E-06
23	5.29594E-06	2.81893E-05	0.000144858	0.000691665	0.002880981	0.009702091	0.024810048
24	0.007384854	0.025238839	0.065709296	0.123811492	0.160010025	0.139786766	0.0887555
25	1.32349E-11	7.24383E-11	3.96404E-10	2.16833E-09	1.18492E-08	6.46049E-08	3.50388E-07
26	0.611461015	0.887861653	0.890409522	0.631202612	0.350073371	0.169076059	0.076315057
27	0.0053269	0.00228112	0.000975773	0.000417203	0.000178344	7.62313E-05	3.25831E-05
28	0.001192411	0.000509952	0.000218015	9.31922E-05	3.98333E-05	1.70256E-05	7.27702E-06
29	0.00423106	0.001813062	0.000775778	0.000331733	0.000141815	6.06189E-05	2.59102E-05
30	0.026876198	0.011689279	0.005033741	0.002158407	0.000923799	0.000395075	0.000168903
31	0.00280046	0.001204351	0.000516115	0.000220843	9.44368E-05	4.03718E-05	1.72569E-05
32	0.028020569	0.012082125	0.005183554	0.002219092	0.000949122	0.000405786	0.00017346
33	0.020985095	0.009097696	0.003912277	0.001676533	0.000717373	0.000306761	0.00013114
34	0.012387778	0.005397832	0.002326323	0.000997841	0.000427139	0.000182683	7.81029E-05
35	0.017228698	0.007406292	0.003173358	0.001357765	0.000580587	0.000248198	0.000106092
36	0.045704406	0.039132035	0.024361662	0.012492714	0.005799501	0.002570357	0.001115974

ITEM	$\theta=5.5$	$\theta=6$	$\theta=6.5$	$\theta=7$	$\theta=7.5$	$\theta=8$	$\theta=8.5$
1	6.53657E-08	2.79381E-08	1.19411E-08	5.10378E-09	2.18142E-09	9.32365E-10	3.98505E-10
2	2.56354E-07	1.09569E-07	4.68313E-08	2.00163E-08	8.55521E-09	3.6566E-09	1.56288E-09
3	2.91351E-07	1.24527E-07	5.32245E-08	2.27488E-08	9.72314E-09	4.15579E-09	1.77624E-09
4	7.2156E-07	3.08404E-07	1.31816E-07	5.63397E-08	2.40803E-08	1.02922E-08	4.39902E-09
5	6.72419E-08	2.874E-08	1.22839E-08	5.25027E-09	2.24403E-09	9.59127E-10	4.09943E-10
6	6.74655E-07	2.88357E-07	1.23248E-07	5.26776E-08	2.25151E-08	9.62322E-09	4.11308E-09
7	0.049683985	0.06816334	0.064395339	0.043547003	0.023471526	0.011172517	0.005009042
8	5.499E-05	2.35061E-05	1.00473E-05	4.29442E-06	1.83551E-06	7.84521E-07	3.35315E-07
9	0.001083473	0.0004633	0.000198058	8.46597E-05	3.61859E-05	1.54665E-05	6.61063E-06
10	0.001764367	0.000754481	0.000322542	0.000137871	5.89299E-05	2.51878E-05	1.07656E-05
11	0.007781905	0.00333589	0.001427598	0.000610502	0.000260996	0.000111564	4.76859E-05
12	0.017801574	0.007689605	0.003301596	0.00141389	0.000604817	0.000258598	0.000110545
13	0.121495027	0.058529525	0.026387512	0.01154355	0.004983519	0.002139188	0.000915998
14	2.7802E-08	1.18829E-08	5.07892E-09	2.17079E-09	9.27824E-10	3.96564E-10	1.69496E-10
15	6.04825E-08	2.5851E-08	1.1049E-08	4.72249E-09	2.01845E-09	8.62712E-10	3.68734E-10
16	8.97843E-07	3.8375E-07	1.6402E-07	7.0104E-08	2.99633E-08	1.28067E-08	5.47375E-09
17	1.8866E-06	8.06371E-07	3.44655E-07	1.4731E-07	6.29624E-08	2.69109E-08	1.15021E-08
18	1.04214E-07	4.45422E-08	1.90379E-08	8.13704E-09	3.47787E-09	1.48649E-09	6.35343E-10
19	7.54369E-08	3.22427E-08	1.37809E-08	5.89015E-09	2.51752E-09	1.07602E-09	4.59905E-10
20	1.60135E-05	6.84601E-06	2.92637E-06	1.25082E-06	5.34627E-07	2.28508E-07	9.76676E-08
21	1.08454E-06	4.6355E-07	1.98128E-07	8.46825E-08	3.61944E-08	1.54699E-08	6.61204E-09
22	6.31261E-06	3.17136E-05	0.000145311	0.000571521	0.001839513	0.004782963	0.009952636
23	0.045769404	0.057756303	0.04935766	0.030858915	0.015881256	0.007387129	0.00327707
24	0.046072917	0.021520849	0.009565236	0.004158158	0.00179024	0.000767562	0.000328504
25	1.8775E-06	9.79273E-06	4.82621E-05	0.000213914	0.000805755	0.002505652	0.006505694
26	0.033402061	0.01442335	0.006191847	0.002651452	0.001134175	0.000484927	0.000207294
27	1.39266E-05	5.95243E-06	2.54415E-06	1.0874E-06	4.64769E-07	1.98648E-07	8.49047E-08
28	3.1103E-06	1.32938E-06	5.68195E-07	2.42854E-07	1.03799E-07	4.43649E-08	1.89621E-08
29	1.10745E-05	4.73342E-06	2.02313E-06	8.64712E-07	3.69589E-07	1.57967E-07	6.7517E-08
30	7.21988E-05	3.08601E-05	1.31903E-05	5.63773E-06	2.40964E-06	1.02991E-06	4.40198E-07
31	7.37611E-06	3.15269E-06	1.34751E-06	5.75944E-07	2.46166E-07	1.05214E-07	4.497E-08
32	7.4143E-05	3.16904E-05	1.3545E-05	5.78932E-06	2.47443E-06	1.0576E-06	4.52034E-07
33	5.60559E-05	2.39599E-05	1.02409E-05	4.37712E-06	1.87084E-06	7.99623E-07	3.41769E-07
34	3.33861E-05	1.42704E-05	6.09947E-06	2.60701E-06	1.11427E-06	4.76256E-07	2.03558E-07
35	4.53465E-05	1.9382E-05	8.28415E-06	3.54076E-06	1.51337E-06	6.46832E-07	2.76464E-07
36	0.000480207	0.00020584	8.80873E-05	3.76695E-05	1.6104E-05	6.88373E-06	2.94231E-06

ITEM	$\theta=9$	$\theta=9.5$	$\theta=10$	$\theta=10.5$	$\theta=11$	$\theta=11.5$	$\theta=12$
1	1.70326E-10	7.27994E-11	3.11154E-11	1.32991E-11	5.6842E-12	2.4295E-12	1.0384E-12
2	6.67993E-10	2.85509E-10	1.2203E-10	5.21571E-11	2.22926E-11	9.52814E-12	4.07245E-12
3	7.59186E-10	3.24486E-10	1.38689E-10	5.92775E-11	2.53359E-11	1.08289E-11	4.62841E-12
4	1.8802E-09	8.0362E-10	3.43477E-10	1.46806E-10	6.27469E-11	2.68188E-11	1.14627E-11
5	1.75215E-10	7.4889E-11	3.20085E-11	1.36808E-11	5.84736E-12	2.49923E-12	1.0682E-12
6	1.75798E-09	7.51384E-10	3.21151E-10	1.37264E-10	5.86683E-11	2.50756E-11	1.07176E-11
7	0.00218586	0.000942656	0.00040445	0.000173151	7.40589E-05	3.16632E-05	1.3535E-05
8	1.43318E-07	6.12558E-08	2.61815E-08	1.11903E-08	4.78288E-09	2.04426E-09	8.73743E-10
9	2.82547E-06	1.20764E-06	5.16163E-07	2.20614E-07	9.42934E-08	4.03022E-08	1.72257E-08
10	4.60139E-06	1.96669E-06	8.40589E-07	3.59278E-07	1.5356E-07	6.56335E-08	2.80526E-08
11	2.03819E-05	8.71155E-06	3.72344E-06	1.59145E-06	6.80204E-07	2.90728E-07	1.24261E-07
12	4.72514E-05	2.01964E-05	8.63229E-06	3.68957E-06	1.57697E-06	6.74018E-07	2.88084E-07
13	0.000391817	0.000167524	7.16121E-05	3.06098E-05	1.30833E-05	5.59205E-06	2.39012E-06
14	7.24448E-11	3.09638E-11	1.32343E-11	5.65652E-12	2.41767E-12	1.03334E-12	4.41663E-13
15	1.57601E-10	6.73608E-11	2.87909E-11	1.23056E-11	5.25956E-12	2.248E-12	9.60824E-13
16	2.33955E-09	9.99952E-10	4.27392E-10	1.82673E-10	7.80766E-11	3.33709E-11	1.42631E-11
17	4.91613E-09	2.10122E-09	8.98086E-10	3.83853E-10	1.64064E-10	7.01229E-11	2.99714E-11
18	2.71553E-10	1.16065E-10	4.96078E-11	2.1203E-11	9.06242E-12	3.87339E-12	1.65554E-12
19	1.96569E-10	8.40161E-11	3.59095E-11	1.53482E-11	6.56E-12	2.80383E-12	1.19839E-12
20	4.17444E-08	1.78421E-08	7.62594E-09	3.25942E-09	1.39312E-09	5.95436E-10	2.54497E-10
21	2.82607E-09	1.2079E-09	5.16271E-10	2.20661E-10	9.43131E-11	4.03106E-11	1.72292E-11
22	0.015757849	0.017650921	0.013799742	0.008157415	0.004075127	0.001868951	0.000823868
23	0.001423405	0.000612607	0.000262614	0.000112387	4.80616E-05	2.05469E-05	8.78287E-06
24	0.000140487	6.00605E-05	2.56733E-05	1.09736E-05	4.69033E-06	2.00472E-06	8.56847E-07
25	0.014246414	0.025349511	0.033706019	0.031399508	0.021104301	0.011345116	0.005394204
26	8.86058E-05	3.78722E-05	1.61873E-05	6.91867E-06	2.95713E-06	1.26392E-06	5.40214E-07
27	3.62893E-08	1.55105E-08	6.62939E-09	2.83348E-09	1.21107E-09	5.17625E-10	2.21239E-10
28	8.10464E-09	3.46403E-09	1.48057E-09	6.32813E-10	2.70472E-10	1.15603E-10	4.94103E-11
29	2.88576E-08	1.23341E-08	5.27175E-09	2.25321E-09	9.63052E-10	4.1162E-10	1.75932E-10
30	1.88146E-07	8.0416E-08	3.43708E-08	1.46905E-08	6.27891E-09	2.68369E-09	1.14704E-09
31	1.92207E-08	8.21518E-09	3.51127E-09	1.50076E-09	6.41444E-10	2.74161E-10	1.1718E-10
32	1.93205E-07	8.25782E-08	3.52949E-08	1.50855E-08	6.44773E-09	2.75584E-09	1.17788E-09
33	1.46076E-07	6.24349E-08	2.66854E-08	1.14057E-08	4.87494E-09	2.08361E-09	8.9056E-10
34	8.70032E-08	3.71862E-08	1.58939E-08	6.79324E-09	2.90352E-09	1.241E-09	5.30418E-10
35	1.18164E-07	5.05049E-08	2.15864E-08	9.2263E-09	3.94344E-09	1.68547E-09	7.20393E-10
36	1.2576E-06	5.3752E-07	2.29743E-07	9.81954E-08	4.19699E-08	1.79385E-08	7.66713E-09