

BİREYSELLEŐTİRİLMİŐ TESTLER ÜZERİNE BİR ÇALIŐMA

81928

Metin YAŐAR

Hacettepe Üniversitesi
Sosyal Bilimler Enstitüsü

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin
Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı
İçin Öngördüğü
DOKTORA TEZİ
Olarak Hazırlanmıştır.

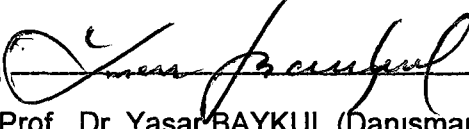
T.C. YÜKSEKÖĞRETİM KURULU
DOKÜMANTASYON MERKEZİ

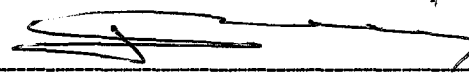
781928

Ankara
Ocak, 1999

Sosyal Bilimler Enstitüsü M¼d¼rl¼ę¼'ne

Bu alıřma, j¼rimiz tarafından Eęitim Bilimleri Anabilim Dalı, Eęitimde ¼lme ve Deęerlendirme Bilim Dalı'nda DOKTORA TEZİ olarak kabul edilmiřtir.

Başkan 
Prof. Dr. Yařar BAYKUL (Danıřman)

¼ye 
Prof. Dr. Nizamettin KO

¼ye 
Do. Dr. A. Ata TEZBAřARAN

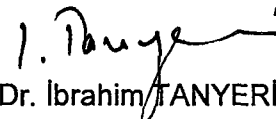
¼ye 
Do. Dr. Aysun UMay

¼ye 
Yrd. Do. Dr. Selahattin GELBAL

Onay

Yukarıdaki imzaların, adı geen ¼ęretim ¼yelerine ait olduęunu onaylarım.

...../...../.....


Prof. Dr. Ibrahim TANYERI
Enstit¼ M¼d¼r¼

ÖZET

Bu çalışmada, Türkçe okuduğunu anlama alanında Klasik Test Teorisi ve İki-parametrelili lojistik model kullanılarak geliştirilen bireyselleştirilmiş iki testin KR-20 güvenirlikleri ve bir grup öğrencinin bu iki yöntemle elde edilen yetenek ölçüleri arasında ilişki düzeyi konusu incelenmektedir.

Bu çalışma için gerekli olan veriler Denizli Ticaret Meslek Lisesi, lise ikinci sınıf düzeyindeki 69 öğrenciye uygulanarak elde edilmiştir.

İki-parametrelili lojistik modele göre geliştirilen bireyselleştirilmiş test, 13 adımda cevaplandırılabilen 61 maddelik bir havuz oluşturulmuştur. Bilgisayar uygulaması BASIC bilgisayar programı vasıtasıyla gerçekleştirilmiştir.

Klasik Test Teorisine göre 47 maddelik kâğıt kalem testi oluşturularak aynı 69 öğrenciye uygulanmıştır.

Çalışmada elde edilen bulgular aşağıda verilmiştir:

Her iki yöneme göre geliştirilen testlere ait KR-20 güvenirlik katsayıları sırasıyla; bireyselleştirilmiş testin güvenirlik katsayısı 0.67 ve kâğıt kalem testine ait güvenirlik katsayısı 0.75 olarak hesaplanmıştır. Her iki güvenirlik katsayısı Fisher'in Z istatistiğine dönüştürülerek karşılaştırılmış; aradaki fark 0.05 düzeyinde manidar bulunmamıştır.

Bireyselleştirilmiş test puanları ile kâğıt kalem testinden elde edilen puanlar arasında Pearson Momentler Çarpımı Korelasyon Katsayısı hesaplanmış, korelasyon katsayısı 0.36 olarak hesaplanmıştır.

Sonuç olarak, her iki y nteme dayalı olarak geliştirilen testlerin KR-20 g venirlik katsayıları arasında manidar bir farklılıđın olmadığı ve uygulamaya katılan  đrencilerin her iki testten elde edilen yetenek  l  leri arasındaki ili ki d zeyinin oldukça d   k olduđu g r lm  t r.



SUMMARY

This study , searches KR-20 reliability of two individualised tests which developed by using Two-parameter logistic model and Classical Test Theory on Turkish reading comprehension and investigates if there is significant difference between ability scores of one group of students obtained these two methods mentioned above. .

The data used in this work were obtained by applying the test to 69 student of the Denizli Vocational High school of trade.

Within the scope of this individualised test which is developed according to Two Parameter logistic model a pool of 61 classes to be answered in 13 step was established. A new computer program was develop in BASIC for the calculations.

According to Classic Test Theory, a paper-pen test of 0.47 clauses was developed and applied to same students.

The findings reached by this study are as follows:

The KR-20 reliability coefficients of individualised and paper-pen tests developed depending on those methods have been calculated as 0.67 and 0.75 consecutively. Both reliability coefficients have been compared by transforming to Fisher's Z statistic and no significant difference found at 0.05 level.

The Pearson Moments Multiplication Correlation Coefficient was calculated and found to be 0.36 for the points between the Converted Individualised test and paper-pen test.

In conclusion there is no significant difference between the develop KR-20 Coefficients based on two methods.

The correlation of the ability measures obtained from two test, of the tested students were found very low.



İÇİNDEKİLER

<u>BÖLÜM</u>	<u>SAYFA</u>
ÖZET (Türkçe ve İngilizce)	I
İÇİNDEKİLER	V
TABLolar	VIII
ŞEKİLLER	IX
ÖNSÖZ	X

BÖLÜM I

GİRİŞ	1
Problem Durumu	1
Tek Boyutluluk	13
Yerel Bağımsızlık	13
Örtük Özellikler Teorisinde Parametrelerin ve Yeteneğin Kestirilmesi	14
Olasılık Fonksiyonu.....	14
Klasik Test Teorisinde Güvenirlik	19
Örtük Özellikler Teorisinde Güvenirlik	21
Bireyselleştirilmiş (Adaptive, Tailored) Test	23
Teste Başlama	27
Testin Devam Ettirilmesi	27
Testin Bitirilmesi	27
Uygulama Stratejileri	30
İki Aşamalı Test	30
Çok Düzeyli Format	30
Kendinden Seçmeli Test	31
Değişken Test	32

Piramitsel Test	33
Tabaka Uyumlu Testler	37
Bireyselleştirilmiş Testlerde Karşılaşılan Bazı Güçlükler	42
Bilgi Yerleştirimi	42
Çapraz Bilgi	42
Dengelenmemiş İçerik	44
Güç (Kuvvet) Eksikliği	45
Test Maddelerinin Güçlük Derecelerine Göre Sıralanmaları .	45
Yüksek Teknolojinin Etkisi	46
İçerikten Kaynaklanan Problemlere Çözüm	47
Güç (Kuvvet) Eksikliğinden Doğan Problemlere Çözüm..	48
Test Maddelerinin Güçlük Derecelerine Göre	
Sıralanmalarıyla İle İlgili Çözüm	49
Yüksek Teknoloji Kullanımından Doğan Problemlere	
Çözüm	50
Bireyselleştirilmiş Testlerin Güvenirliği	50
Problem Cümlesi	51
Alt Problemler	52
Araştırmanın Önemi	52
Sayıtlılar	53
Sınırlamalar	53
Tanımlar ve Kısaltmalar	54
İlgili Araştırmalar	54

BÖLÜM II

YÖNTEM	59
Veri Toplama Aracı	59
Verilerin Elde Edildiği Grup	63
Deneme Uygulaması İçin Gerekli Olan Madde Havuzu	65

Klasik Test Teorisine Göre Grup Testinin Hazırlanması	67
Verilerin Elde Edilmesi	67
Verilerin Çözümlemesi	68
BÖLÜM III	
BULGULAR VE YORUM	69
BÖLÜM IV	
SONUÇ VE ÖNERİLER	77
Sonuç	77
Öneriler	78
KAYNAKÇA	81
EKLER	85

TABLOLAR

<u>Tablo</u>	<u>Sayfa</u>
1.1 Belirli Uzunluktaki Testler İçin Ölçmenin Standart Hata (Ö.S.H) Değerleri	37
1.2 Her Tabakada Yer Alan Maddelere Ait Farklı a_j Değerleri İçin Ölçmenin Hata Varyans Değerleri	40
2.1 69 Öğrencinin Madde Havuzundan Cevaplandıkları Madde Numaraları	63
3.1 Bireyselleştirilmiş ve Kağıt-Kalem Testlerine Ait Puanlar	70
3.2 Bireyselleştirilmiş Test Uygulamasındaki Madde Parametrelerinin Klasik Test Teorisindeki Madde İstatistikleri Karşılıkları	73

ŞEKİLLER

<u>Şekil</u>	<u>Sayfa</u>
1.1 Madde Karakteristik Eğrisi	11
1.2 Normal Ogive (ve lojistik ogive) Modellerine Ait Bazı Hipotetik Madde Karakteristik Eğrileri	12
1.3 Newton - Raphson Yönteminin Gösterimi	18
1.4 Bireyselleştirilmiş Test Algoritması	29
1.5 Değişken Testte Madde Sıralanış Şekli	32
1.6 Piramitsel Testte Madde Sıralanış Şekli	34
2.1 Bireyselleştirilmiş Teste Ait Cevaplandırma Algoritması	61

ÖNSÖZ

Bu araştırma, Türkçe okuduğunu anlama alanında Klasik Test Teorisine ve İki-parametrelili lojistik modele göre geliştirilen iki testin güvenilirlik katsayıları manidar bir farklılığını olup olmadığı ve bir grup öğrencinin bu testlerden elde edilen yetenek ölçüleri arasındaki ilişki düzeyinin belirlenmesi amacıyla yapılmış bir çalışmadır.

Bu araştırma konusunun planlanmasında, yürütülmesi aşamasında her türlü (verilerin çözümlenmesi de dahil olmak üzere) katkıda bulunan, araştırmanın her kademesinde hiç bir zaman desteğini esirgemeyen tez danışmanım sayın Prof. Dr. Yaşar BAYKUL'a teşekkürü bir borç olarak bilirim.

Eldeki çalışma için gerekli olan verilerin elde edilebilmesi (bireyselleştirilmiş testin uygulanması) için gerekli olan BASIC bilgisayar programını yazan ve yine çalışmanın bir çok aşamasında yardımlarını esirgemeyen H.Ü. Eğt. Fak. Eğitim Bilimleri Bölümü öğretim üyesi sayın Yrd.Doç.Dr. Hülya KELECİOĞLU'na çok teşekkür ederim.

Yine çalışmanın her aşamasında daima manevi desteğini gördüğüm eşim Sevilay YAŞAR'a teşekkür ederim.

Metin YAŞAR

BÖLÜM I

GİRİŞ

Bu bölümde problem durumu, problem cümlesi, alt problemler, araştırmanın amacı, araştırmanın önemi, sayıtlar, sınırlamalar, tanımlar ve kısaltmalar ile ilgili araştırmalar üzerinde durulmuştur.

Problem Durumu

Bütün toplumlarda eğitimin temel işlevi, toplumun bütün bireylerini belli niteliklerle donanık kılmak ve onları içinde yaşadığı topluma daha yararlı hale getirmektir. Bunun için de bireylerin ilgi ve yeteneklerinin bilinmesi gerekmektedir. Toplumdaki bireylerin anılan ilgi ve yetenekleri arasında sözel yetenekleri de bulunmaktadır. Bireyler bu özelliklerini kullanarak, toplumla sağlıklı bir iletişim kurmaları mümkün olabilir. Bundan dolayı bireylerin sözel yeteneklerinin bilinmesine gerek vardır. Baykul'un (1979) da belirttiği gibi; bireylerin ilgi ve yetenekleri doğrultusunda yönlendirilmeleri, eğitim sürecine girmeden yeteneklerinin belirlenmesi, özelliklerine en uygun alanlara yönlendirilmeleri ve bir alanda eğitime en uygun kişilerin seçilmesi; hem bireyin hem de toplumun birçok sorununun çözülmesiyle yakından ilgilidir.

Eldeki çalışma, bir grup lise ikinci sınıf öğrencisinin sözel yetenek düzeylerinin belirlenmesi amacına yönelik olup; elde edilen verilerin çözümlenmesinde kullanılan Klasik Test Teorisi ve Örtük Özellikler Teorisi modellerinden İki-parametrelili Lojistik model kullanılarak sonuçlarının karşılaştırılmasını içermektedir.

Eğitimde gözlenmesi amaçlanan nitelikler arasında başarı, yetenekler, ilgi ve tutumlar önemli bir yer tutar. Bu nitelikler, eğitimdeki diğer pek çok değişkende olduğu gibi, doğrudan gözlenemeyen değişkenlerdir. Böyle durumlarda ölçme daha da önem kazanır. Çünkü, doğrudan gözlenemeyen değişkenleri ölçmede, ölçmeye başka değişkenlerin karışması, tesadüfi hataların artması, duyarlı sonuçların elde edilmemesi ve hatta bu değişkenlerin yeterli şekilde tanımlanmasında zorluklarla karşılaşılır. Bu araştırmanın konusu sözel yeteneğin ölçülmesiyle ilgilidir.

Bütün bilim dallarında yapılan ölçmelerde olduğu gibi eğitimdeki ölçmelerde de araç kullanılmaktadır. Ölçmelerde araç kullanma, bir yönden ölçme sonuçlarının duyarlılığını artırır, diğer yandan standartlaştırmayı kolaylaştırır; hatta doğrudan gözlenmesi mümkün olmayan bazı nitelikleri dolaylı olarak gözleme imkanı verir.

Eğitimdeki ölçmelerde, çoktan seçmeli testler, özellikle geniş grupları ilgilendiren uygulamalarda en çok kullanılan ölçme aracı durumundadır. Diğer taraftan istatistiksel olarak çoktan seçmeli testler üzerinde çokça çalışılmakta olup bunlar için temeli istatistiğe dayanan teoriler geliştirilmiştir. Bu teorilerden en çok bilinen ve en çok sözü edilen ikisi klasik test teorisi ve örtük özellikler (latent traits) teorisidir. Örtük Özellikler teorisine bazen madde-cevap teorisi (item response theory) de denilmektedir.

Bu iki teorinin dışında, bir yönden test geliştirme, diğer yandan da bir test uygulaması gibi algılanabilecek olan diğer bir yaklaşımdan daha söz edilmektedir. Bu da, bireyselleştirilmiş (adaptive veya tailored) testtir. Bireyselleştirilmiş test psikometrik özellikleri önceden hesaplanmış çok sayıda test maddesi (bu test maddelerinin tümüne, madde havuzu veya madde bankası adı verilmektedir) arasından uygun olanların alınması suretiyle her bireyin ölçülen özelliğine uygun maddelerden oluşturulan bir testtir.

Yukarıda bahsedilenlerden Klasik Test Teorisi, fazla varsayımları olmaması nedeniyle, test geliřtirmede pek çok ölçme probleminin çözümünde rahatça kullanılmıřtır. Günümüzde bu teoriye göre geliřtirilmiř çok sayıda başarı, yetenek, kiřilik ve diğerk testler vardır.

Klasik Test Teorisi günümüzde sıkça kullanılmasına rağmen birtakım zayıf varsayımlara da sahiptir. Bu nedenle eğitimde ve psikolojide bu teoriye dayalı olarak kullanılan testlerin geliřtirilmesi, uygulanması ve değerkendirilmesinde bir çok eleřtirilere uğramaktadır. Bunlardan birincisi, sıkça kullanılan madde istatistiklerinin seçilen örnekleme oldukça bağımlı olması ve örneklemden etkilenmesidir (Erkkin, 1994, s. 51).

Madde ve dolayısıyla test istatistiklerinin örnekleme bağılı olması, özellikle örneklemin çok geniş olmadığı durumlarda, test puanlarının yorumlanmasını ve testin kullanışlılığını büyük ölçüde sınırlamaktadır; hatta, bazı durumlarda bir örneklem için geçerli olan bir test başka bir örneklem için geçerli olmayabilmektedir.

Halihazırda kullanılan test teknikleri ve yöntemlerinde de bazı problemlere rastlanılmaktadır. Bir örnek vermek gerekirse test maddesinin güçlük ve ayırt edicilik derecelerinin bireylerin düzeylerine göre farklılık göstermesi gerekirken bir çok test tekniğinde bu farklılığı görmek mümkün değildir. Doğal olarak yetenek düzeyi yüksek olan bir grubu teste tabi tutarken test maddelerinin güçlük derecesinin de yüksek olması gözönünde bulundurulması gerekmektedir. Lord ve Novick (1968)'e göre, grup farklılıkları dikkate alınarak test maddelerinin hazırlanması da günümüzde çok önemlidir. Bu yöntemlerin uygulanmasının sonucunda testlerde daha yüksek performans oranı gözlenebilmektedir (Hambleton ve Swaminathan, 1984, s.2).

Sonuçta testler, geliştirilme safhasında uygulanan gruba aynı düzeydeki bir gruba uygulanırken ortaya çok büyük problemler çıkmaz; fakat içinde farklı düzeyde bireylerin olduğu gruplarda problemlerin ortaya çıkması çoğu zaman kaçınılmaz olmaktadır. Bu problemlerden başta gelenlerden biri, esas uygulamada testin nitelikleri değişeceğinden yorumlama ve karar vermede güçlükler, hatta bazı durumlarda hatalarla karşılaşılmasıdır. Diğer bir problem testin grubun tamamına hitap etmesi endişesiyle maddelerin güçlük indekslerinin geniş tutulması zorunluluğudur. Bu durum testin uzunluğunu gereksiz yere artırmaktadır. Diğer taraftan, testlerin değerlendirilmesi ise grup içi bireysel farklılıkları dikkate almaksızın yapıldığından test güvenilirliği ve geçerliğinin de olumsuz yönde etkilenmesidir.

Weiss göre (1983), Klasik Test Teorisinde görülen bir başka problem ise testin desenlenmesi ve analizi konusunda ortaya çıkmaktadır. Bireylerin birbirleriyle karşılaştırılmaları kısıtlı (yetersiz) uzunluktaki testlerle tam anlamıyla gerçekleştirilememektedir. Orta düzeydeki bireyler için kolaylıkla düzey testleri bulunabilirken, üst ve alt düzeydeki bireylerin düzeylerini ölçmek için test bulmak pek o kadar kolay olmamaktadır. Testlerin güvenilirliği, bireylerin ölçülen özellikleri yönünden ayarlanması durumunda daha da yükselecektir (Hambleton ve Swaminathan, 1984, s. 4-8).

Böyle yüksek bir güvenilirliği sağlamak için de bazen test uzunluğunda değişikliğe gidilmekte, fakat ortaya çıkan test tipiyle arzu edilen güvenilirlik derecesinin elde edilmesi güçleşmektedir.

Klasik Test Teorisinin bir başka problemi güvenilirliğin tanımı ve hesaplanmasıyla ilgilidir. Bu teoride güvenilirlik paralel ölçmelere dayandırılmıştır ve paralel ölçmelerde bireyin ölçülen özelliğinin değişmediği varsayılmaktadır. Halbuki bireyler bir testi ikinci defa aldıklarında bile bazı farklılıkların ortaya çıkması kaçınılmaz olmaktadır. Bireyler birşeyleri öğrenirken aynı zamanda da

yine birşeyleri de unutmaktadırlar (Tekin,1984, s.60). Bu durum güvenilirlik katsayısının yorumlanmasında problemlere sebep olmaktadır.

Yukarıda belirtilen problemlerin bir kısmı Klasik Test Teorisinin kendisinden diğer bir kısmı da bu teoriyle bağlı uygulamalardan kaynaklanmaktadır. Klasik Test Teorisinden kaynaklanan problemleri ortadan kaldırmak amacıyla Örtük Özellikler Teorisi (Latent Traits Theory) veya Madde Tepki Teorisi'nin (Item Response Theory) kullanılması son zamanlarda gittikçe önem kazanmaktadır.

Klasik Test Teorisine dayalı olarak geliştirilen bir test, grubun tamamına aynı anda uygulanır. Bu uygulama örtük özellikler teorisine dayalı olarak geliştirilen testler için de geçerlidir. Ancak, örtük özellikler teorisine dayalı testler aynı zamanda bireylere uygun bir test olarak da geliştirilip uygulanabilir. Böyle testlere “tailored test” veya “adaptive test” gibi adlar verilmektedir. Bu araştırma, sözel yeteneğin örtük özellikler teorisine dayalı olarak adaptive testle ölçülmesi üzerinedir.

Hambleton ve Swaminathan'ın (1984) belirttiğine göre, örtük özellikler teorisinin, Richardson (1936), Lawly (1943-44) ve Tucker'ın (1946), çalışmalarıyla ana hatları çizilmiştir. Yukarıda sözü edilen psikometristler arasında ise Tucker “madde karakteristik eğrisi”ni (item characteristic curve) ilk kullanan kişidir.

Lawley (1943,1944,) örtük özellikler test tekniğinde birçok teorik gelişmeye öncü olmuştur. Yine Lord, Lawley'in bu çalışmalarından etkilenmiştir. Lawley klasik test teorisinin parametrelerini, örtük özellikler teorisine uygulayarak bu modeli ilerletmeye çalışmıştır. Fakat Lawleyin geliştirdiği modellerin kısıtlı ve sınırlı olduğu yerler de mevcuttur. Örneğin onun çalışmalarında (1) maddeler arasındaki bağıntılar eşit değildi, (2) performanslarda

tahmine pek yer yoktur. Richardson (1936) ve Tucker (1946) Klasik Test Teorisi modeli parametreleri ile örtük özellikler teorisi parametreleri arasında eğrileri kullanarak bağıntılar oluşturmuşlardır (Hambleton ve Swaminathan, 1984,s. 4-8).

Lazersfield (1950), araştırmalarının çoğunu davranış değerlendirmeleri üzerine yoğunlaştırmıştır. Yine Lazersfield “örtük özellikler” kavramını da ilk kez ortaya çıkaran (kullanan) araştırmacıdır. Lord’un yaptığı çalışmalar örtük özellikler teorisinin temelini oluşturmaktadır. Belki bu nedenle böyle bir test modelini hazırlayıp uygulayan ilk kişi o olmuştur. Samejima’nın yaptığı çalışmalar da madde-cevap teorisinin gelişmesi konusunda çok faydalı olmuştur (Hambleton ve Swaminahan, 1984,s.4-8; Lord ve Novick, 1968).

Birnbaum (1957, 1958a, 1958b) yaptığı çalışmalarla Lord’un elde ettiği eğrileri ve kullanma yöntemlerini daha da geliştirmiştir. Ayrıca bu modelin kullanımını daha da kolaylaştıracak istatistikleri de hazırlamıştır. Buna rağmen 1950 ve 1960’larda bu alanda yeterli bir ilerleme görülememiştir. Buna neden olarak bilgisayar destekli testlerin hazırlanamaması gösterilmiştir. Bu alandaki önemli sayılabilecek gelişmelerden test puanlama eşitliği ve uyumlu testler Lord tarafından gerçekleştirilmiştir (Thissen ve Mislevy, 1990, s.120).

Örtük Özellikler Teorisinin gelişmesinde Rasch’ın Danimarka’da yaptığı çalışmaların da etkisi büyük olmuştur. Kendi ismiyle anılan Rasch modeli daha sonraki yıllarda Wright ve arkadaşlarının çalışmalarında temel kaynak olarak gösterilmektedir. Wright ve arkadaşları bu alandaki uygun bilgisayar programları ile başarılı testler geliştirmişlerdir (Hambleton ve Swaminahan; 1984: s.4-8).

“n” tane maddeden oluşan bir test ve “k” tane de özellik olduğunu kabul edelim; her özelliğin test içerisindeki bir madde ile bireyin performansını etkilediğini varsayalım. Bu özellikler bireyin durumunu açıklamak için

kullanılması gerekli olan psikolojik boyutlar olarak düşünülebilir (Baykul, 1979, s. 32).

Yukarıda değinilen psikolojik boyutlar (özellikler) örtük özellikler veya örtük değişkenler olarak adlandırılmaktadır. Bir örtük değişken teorik olarak

$$\theta = (\theta_1, \theta_2, \theta_3, \dots, \theta_K) \quad (1.1)$$

vektörüyle gösterilmektedir. Bu gösterimde K değişkenin boyutlarının sayısını, θ , K boyutlu olan yeteneği, $\theta_1, \theta_2, \theta_3, \dots, \theta_K$ 'da θ yeteneğinin boyutlarını (boyutlarına ait koordinatları) temsil eder.

Şüphesiz bireylerin θ yetenekleri doğrudan gözlenememekte; dolayısıyla (1.1) vektörünün koordinatları bilinmemektedir. Bu sebeple bireylere ait (1.1) vektörü teoriktir.

Örtük Özellikler Teorisinde, θ yeteneğini ölçme amacıyla bir test hazırlanır. Bireylerin bu testten aldıkları puanlar geometrik bir uzayın noktası olarak düşünülmektedir. Bu ikinci uzay (θ') , θ uzayının temsilcisi olarak alınmakta ve bireyler bu uzayda

$$\theta' = (\theta'_1, \theta'_2, \theta'_3, \dots, \theta'_K) \quad (1.2)$$

vektörüyle temsil edilmektedir.

n maddelik bir testin, tüm evrenler için hazırlandığını ve her bir maddenin de her bir bireye sadece bir defa uygulandığını düşünelim. Böyle bir uygulamada, testin bir g maddesine ait madde puanının herhangi bir θ için şartlı frekans dağılımı, testi alan bütün bireyler grupları için aynı değilse (1.2)

bağıntısıyla verilen θ vektörü, dikkate alınan grubun örtük özellikler uzayı değildir (Baykul, 1979, s.32).

Bu düşünceden hareket ederek, test içerisindeki herhangi bir maddeye ait madde puanını herhangi bir sabit θ değeri için teste tabii tutulan gruba ait bütün özellikler uzayının bütün boyutlarını kapsamıyor demektir. Böyle bir örtük uzayının tamam olabilmesi için örtük özellikler uzayına yeni psikolojik boyutların eklenmesi gerekmektedir . Aynı zamanda Lord'a göre, K boyutlu ve tamam olan bir örtük uzay şu özelliği göstermelidir: Tamam olan bir örtük özellikler uzayında, sabit bir θ için, madde puanlarının dağılımı bütün evrenler için aynı olmalıdır (Lord, 1980, s. 359).

Örtük özellikler kuramında, madde puanının θ vektörü üzerindeki regresyonu “madde karakteristik fonksiyonu” adını almakta ve $P_g(\theta)$ ile gösterilmektedir. $P_g(\theta)$ fonksiyonu tanımı itibarıyla bir örneklemden diğerine değişmez. $P_g(\theta)$ ’ yı belirleyen parametreler de sabit madde parametreleri haline gelir.

Cevaplayıcıların kendilerine yöneltilen her hangi bir maddeyi doğru cevaplandırmaları durumunda 1 ve cevapsız bırakılması veya yanlış cevaplandırmaları durumunda 0 (sıfır) şeklinde puanlanan maddeler için $P_g(\theta)$, g maddesini cevaplandıran örneklemin gözlenebilen özellikleri ile örtük özellikleri arasındaki bağıntıyı oluşturmaktadır (Baykul, 1979, s.33).

Bu özelliğinden dolayı $P_g(\theta)$ fonksiyonu, gözlenebilen madde cevaplarından hareketle gözlenemeyen özelliklerin kestirilmesinde kullanılmaktadır. Bu kullanılış amacı, zihinsel yeteneklerin ölçülmesinde temel yaklaşımı göstermektedir.

Test puanlarının örtük özellikler $\theta = (\theta_1, \theta_2, \theta_3, \dots, \theta_k)$ üzerindeki dağılımı ile madde karakteristik fonksiyonlarının olduğu yukarıda açıklanmıştır. Bu teoride θ yeteneğinin kestirilmesinde, yeteneğin evrendeki dağılımının şekline bağlı olmakta ve kabul edilen dağılımın biçimi yani madde karakteristik fonksiyonu esas rolü oynamaktadır.

Yukarıda genel yapısı kısaca açıklanan madde karakteristik fonksiyonunun Örtük Özellikler Teorisinde matematiksel olarak saptanmış şekilleri vardır. Her bir şekil bu teori içinde ayrı birer model oluşturmaktadır. Örtük Özellikler Teorisi modellerinin başlıcaları (1.3), (1.4) ve (1.5) numaralı eşitliklerle tanımlanan sırayla bir, iki ve üç parametrelili lojistik modellerdir.

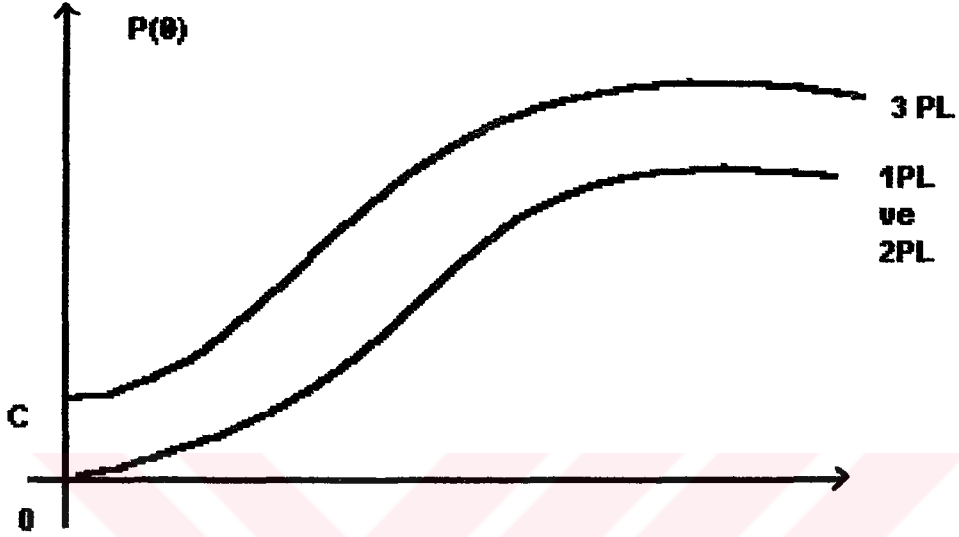
$$P_j(\theta) = \frac{1}{1 + e^{-A(\theta - b_j)}} \quad (1.3)$$

$$P_j(\theta) = \frac{1}{1 + e^{-a_j(\theta - b_j)}} \quad (1.4)$$

$$P_j(\theta) = c_i + (1 - c_i) \frac{1}{1 + e^{-a_j(\theta - b_j)}} \quad (1.5)$$

Yukarıdaki eşitlikler bir j maddesine ait dağılım fonksiyonlarıdır. Bu fonksiyonlardaki $P_j(\theta)$, j maddesinin belli bir θ yeteneğindeki bireyler tarafından doğru cevaplandırılma olasılığını; b_j , j maddesinin en iyi ölçtüğü yetenek düzeyini; a_j , j maddesinin kalitesini (bir yönüyle ayırtıcılığını); c_j , j maddesinin şansla doğru cevaplandırılma olasılığını veya bu maddenin doğru

cevaplandırılması için gerekli minimum yetenek düzeyini gösterir. Bu modellere ait dağılım fonksiyonlarının eğrileri aşağıdaki şekilde olduğu gibi gösterilebilir.

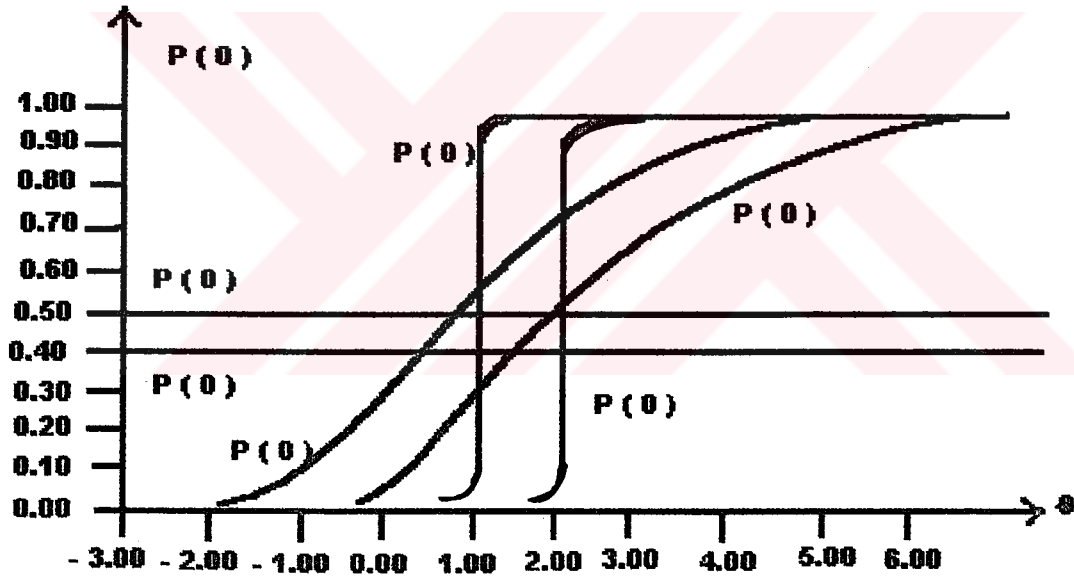


Şekil 1.1: Madde Karakteristik Eğrisi

Bu fonksiyonlara madde karakteristik fonksiyonları ve eğrilerine de madde karakteristik eğrileri adı verilmektedir.

(1.3) eşitliğindeki model, bir parametrelili lojistik model olarak adlandırıldığı gibi buna Rasch Modeli de denilmektedir. Bu modele ait eşitlikteki A bir parametre olmayıp testi oluşturan maddelerin a_j parametrelerinin ortalaması olarak alınmakta olup bütün maddeler için ortaktır. Bu modelde parametre olarak sadece b_j kalmaktadır. Her üç modelde b_j , bir yandan j maddesinin en iyi ölçtüğü yetenek düzeyini; diğer yandan bu maddeyi doğru cevaplayan bir bireyin en az bu yetenek düzeyinde olduğunu gösterir. Teorik olarak b_j $(-\infty; +\infty)$ aralığında değerler alabilir.

(1.4) ve (1.5) eşitliklerindeki a_j yukarıda da belirtildiği gibi j maddesinin kalitesini belirten bir parametre olup Klasik Test Teorisindeki madde ayıricılık gücü indeksine karşılık gelmektedir. En kaliteli maddeler a_j değeri 1.00 civarında olanlardır. a_j parametresinin Klasik Test Teorisindeki değerine bakıldığında; $a_j = 0.5$ olduğunda Klasik Test Teorisinde $r_{bis} = 0.44$ ve yine $a_j = 1.00$ olduğunda ise Klasik Test Teorisinde $r_{bis} = 0.70$ 'e eş değerdedir (Thorndike, 1982, s.83). a_j değeri 0.00' a yaklaştıkça madde karakteristik eğrisi yatay eksene paralel bir doğru halini almakta, 1.00 değerine geldikçe normal ogive halini gelmekte ve 1.00'den büyük değerlere ulaştıkça da buna ait eğri dik bir ogive olmaktadır.



Şekil 1.2 : Normal Ogive (ve lojistik ogive) modellerine ait bazı
hipotetik madde karakteristik eğrileri

Örtük Özellikler Teorisine göre test hazırlamada, istenilen b_j parametresine sahip maddeler seçilmek suretiyle gereksiz yere uzatılmadan belli bir yetenek aralığında istenen etkililikte testler hazırlanabilir. Bu durum, Örtük

Özellikler Teorisinin Klasik Test Teorisine olan üstünlüklerinden birini oluşturur.

Gerçekten, bu teoriye dayalı modellerden birine göre geliştirilen testlerden bireylerin yeteneklerinin ölçülmesi Klasik Teoriye dayalı bir teste göre daha kısa ve belli bir yetenek aralığında daha etkili olmaktadır.

Örtük Özellikler Teorisinin yukarıda belirtilen özelliklerine ulaşılabilmesi için bazı sayıtların yerine getirilmesi gerekir. Bunlar, tek boyutluluk ve yerel bağımsızlıktır.

Tek Boyutluluk

Örtük Özellikler Teorisinde tek boyutluluk, ölçmeye konu edilen özelliğin tek boyutlu olmasını gerektirmektedir. Eğitim gibi, değişkenlerin karmaşık olduğu bir alanda bu şartın yerine getirilmesi oldukça zordur. Ancak, günümüzde karmaşık değişkenler istatistiksel yöntemlerle boyutlarına ayrıştırılarak ve ayrı boyutları ölçen alt testler birleştirilerek bu problem aşılmaya çalışılmaktadır.

Tek boyutluluğun sağlanmasında, en çok başvurulan yöntemlerden biri faktör analizidir. Bu yöntemde madde puanları faktör analizine tabi tutulur, bir faktör altında büyük yük veren maddelerin tek (aynı) boyutta olduğu kabul edilmektedir. Bu yüzden aynı faktörde yığılan maddelerden elde edilen test tek boyutlu olur.

Yerel Bağımsızlık

Yerel bağımsızlık, cevaplayıcıların aynı $\theta_1, \theta_2, \theta_3, \dots, \theta_k$ değerleri ile tanımlandığı her hangi bir örnekleme madde puanlarının (şartlı) dağılımı birbirlerinden bağımsız ise bu özelliğe yerel bağımsızlık adı verilmektedir

(Lord: 1968; s. 361). Diğer bir ifadeyle, yerel bağımsızlık, bir maddeye verilen cevapların, belli bir yetenek düzeyinde, başka maddelerden bağımsız olması anlamına gelir. Bu duruma göre yerel bağımsızlık, tüm gruplar için maddeler arasında korelasyon bulunmadığı anlamına gelmemekte; fakat, maddeye verilen cevapların farklı yetenek düzeylerinde bağımsızlığı anlamına gelmektedir.

Lord ve Novick'e (1968) göre yerel bağımsızlık yaklaşımına göre bir grup test maddesinin bağımsız olacağını düşünmek de yanlış olabilir. Bireylerin yetenekleri arasında farklar görüldüğünde test maddeleri arasında da pozitif ilişkiler görülebilir. Sabit yetenek düzeyinde test puanlarını bu ilişkiler etkilememelidir.

Yerel bağımsızlık ve tek boyutluluk modelin öngörülleri birbirlerine yakın olduğundan bu iki modelin ortak teknikler kullanılabileceğini düşünmek belki mümkün olabilir.

Yerel bağımsızlığın sağlanmasında faktör analizinden ve maddeler arası korelasyonlardan yararlanılmaktadır.

Örtük Özellikler Teorisinde Parametrelerin ve Yeteneğin Kestirilmesi

Örtük Özellikler Teorisinde testin maddelerine verilen cevaplardan yararlanılarak önce seçilen modelin parametreleri ve cevaplayıcıların yetenekleri kestirilir. Bu kestirmelerde olasılık fonksiyonlarından yararlanılır. Kestirmelerde izlenen yolun esası aşağıda kısaca açıklanmaktadır.

Olasılık Fonksiyonu: θ yetenek seviyesindeki bir bireyin i no'lu maddeyi cevaplandırması sonucu oluşan U_i fonksiyonu,

$$U_i = \begin{cases} 1 & \text{doğru cevap için} \\ 0 & \text{yanlış cevap için} \end{cases}$$

biçiminde tanımlanmaktadır.

Bir maddenin doğru cevaplandırılması durumunda olasılık fonksiyonu $P(U_i/\theta)$ şeklinde olup, genellikle $P(U_i = 1/\theta)$ ya da daha kısa bir şekilde P_i olarak gösterilmektedir. Daha önce de belirtildiği gibi, madde cevap fonksiyonu, bir bireyin yetenek seviyesi θ 'nın, bir fonksiyonu olup parametrelerle belirlenmektedir. U_i bir binom değişkeni olduğundan bunun olasılık fonksiyonu,

$$\begin{aligned} P(U_i/\theta) &= P(U_i = 1/\theta) P(U_i = 0/\theta) \\ &= P_i^{U_i} (1 - P_i)^{1-U_i} \end{aligned} \quad (1.6)$$

eşitliği ile tanımlanmıştır.

Burada $Q_i = 1 - P_i$ olmaktadır. Eğer θ yetenek düzeyindeki bir birey n maddeye cevap verirse, oluşan fonksiyonlar, $(U_1, U_2, U_3, \dots, U_n)$ cevaplama vektörü olmak üzere $P(U_1, U_2, U_3, \dots, U_n/\theta)$ şeklinde gösterilmektedir. Eğer maddeler tek boyutlu ve yerel bağımsız iseler θ yetenek düzeyi için n maddeye verilen cevaplar

$$\begin{aligned} P(U_1, U_2, \dots, U_n/\theta) &= P(U_1/\theta) P(U_2/\theta) P(U_3/\theta) \dots P(U_n/\theta) \\ &= \prod_{i=1}^n P_i^{U_i} Q_i^{(1-U_i)} \end{aligned} \quad (1.7)$$

$$L(U_1, U_2, \dots, U_n/\theta) = \prod_{i=1}^n P_i^{U_i} Q_i^{1-U_i} \quad (1.8)$$

şeklinde de ifade edilebilmektedir.

Yukarıda (1.7) eşitliğinde verilen olasılık fonksiyonu bir ölçüt fonksiyon olarak görülebilir. Olasılık fonksiyonu en yüksek yapan θ değeri, θ tahmin değeri olarak kullanılabilir. Bu tahmin değerinin en yüksek olasılık fonksiyonu olarak görülebilir. Tahmin değeri bireylerin yeteneklerini belirleyen bir değer olarak görülebilir ve bu şekilde de bireylerin belirlenen madde-cevap modeli için en yüksek olasılık fonksiyonu hesaplanabilmektedir.

Yukarıda (1.7) eşitliğinde verilen likelihood fonksiyonunun grafiği çizilebilir ve grafikten en yüksek olasılık tahmin değeri (θ) belirlenebilir. Bu grafiği çizmek yerine, $L(U/\theta)$, $\ln L(U/\theta)$ fonksiyonu çizilebilir. $\ln$ burada doğal logaritma fonksiyonu olarak belirtilmektedir. $\ln$ ve L fonksiyonları birbirleriyle ilişkili olduğu için, $L(U/\theta)$ fonksiyonunu en yüksek yapan θ değeri aynı zamanda $\ln L(U/\theta)$ fonksiyonunu da en yüksek yapmaktadır. (1.8) eşitliğindeki $\ln L(U/\theta)$ fonksiyonu Likelihood fonksiyonunun logaritması alınarak bulunmuş olup,

$$\ln L(U/\theta) = \sum_{i=1}^n (u_i \ln P_i + (1 - u_i) \ln Q_i) \quad (1.9)$$

ifadesine sahiptir.

(1.9) eşitliği kullanılarak en yüksek olasılık kestirici (maximum likelihood estimation) metodu yardımıyla maddelere verilen cevaplardan çözülerek madde parametreleri hesaplanmaktadır. Bu hesaplamada, sayısal yaklaşık yöntemlerden Newton-Raphson yöntemi kullanılmaktadır.

Newton-Raphson yöntemi şu şekilde açıklanabilir; çözümü bulunacak olan eşitlik $F(x) = 0$ olsun, yaklaşık bir cevap olarak da X_0 olsun. daha net bir sonuç bulabilmek için

$$X_1 = X_0 - h \quad (1.10)$$

yazılabilir.

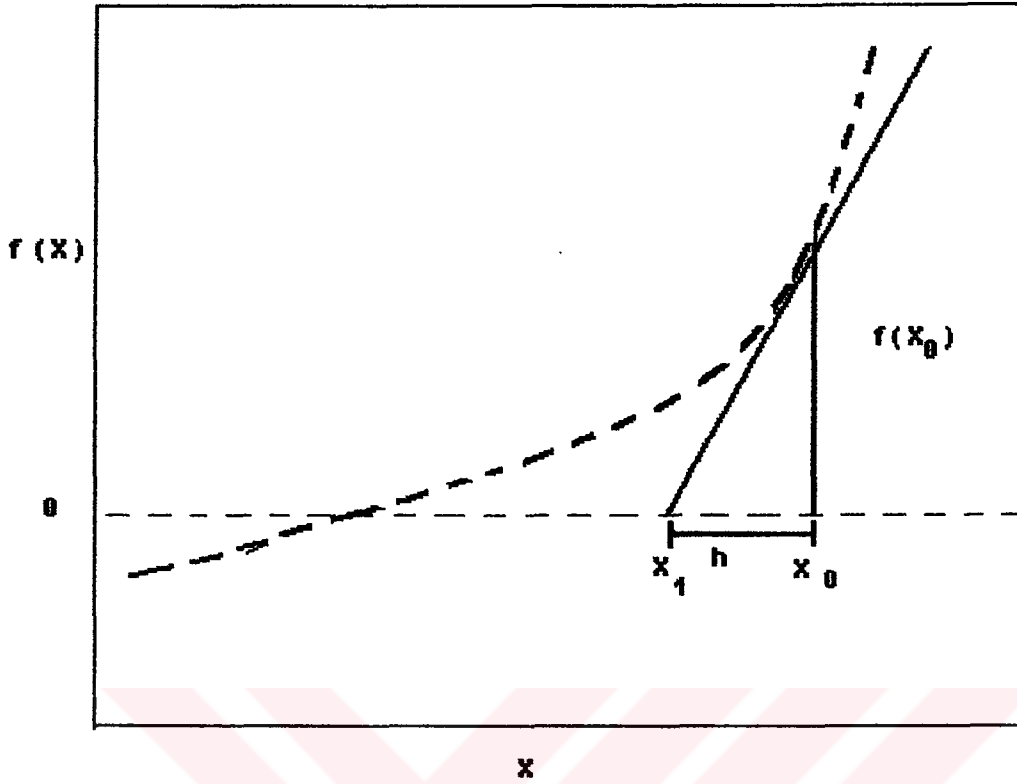
Şekil 1.3'deki eğride, $F(X_0)$, X_0 apsisli noktadaki ordinatıdır. (1.9) eşitliğindeki fonksiyonunun X_0 apsisli noktadaki türevi $F'(X_0)$ ' dir. Bu şekilde üçgenden,

$$h = \frac{F(X_0)}{F'(X_0)} \quad (1.11)$$

yazılabilir ve bu eşitlik (1.10) eşitliğinde kullanılarak,

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} \quad (1.12)$$

elde edilir.



Şekil 1.3: Newton-Raphson Yönteminin Gösterimi.
(Hambleton ve Swaminathan'dan alınmıştır. 1984,S. 80)

Böylelikle x_1 bulunur ve daha da yaklaşık bir sonuç bulunması gerekiyorsa, x_0 üzerinde uygulanan aynı yöntemi x_1 'de de tekrar etmek gerekiyor. Bu işlem periyodik olarak devam ederse daha genel bir eşitlik olan,

$$x_{m+1} = x_m - \frac{f(x_m)}{f'(x_m)} \quad \text{çıkarak.} \quad (1.13)$$

bu tekrarlama $(m+1)$. yaklaşımla, m . yaklaşım arasındaki fark belirli bir değerin altına düşürüldüğünde ise $f(x)$ fonksiyonu x_m 'e yaklaştığı kabul edilir ve $f(x)=0$ için x_m çözüm olur. $(m+1)$ 'nci iterasyonda θ 'nın değeri

$$\theta_{m+1} \equiv \theta_m - \left(\frac{d}{d\theta} \ln L(u/\theta)_m / \left(\frac{d^2}{d\theta^2} \ln L(u/\theta) \right)_m \right) \quad (1.14)$$

elde edilir. Bu yöntem belirli bir sayıya yaklaşıncaya kadar tekrar edilir ya da yaklaşılan değer en yüksek olasılık değeri θ olarak alınır ve θ ile gösterilir.

Bireyselleştirilmiş testler hakkındaki açıklamalara geçmeden önce, ilgisi dolayısıyla, güvenilirlik konusuyla ilgili olarak kısa bir özet sunmakta fayda vardır.

Klasik Test Teorisinde Güvenirlik

Herhangi bir ölçme aracından elde edilen ölçme sonuçlarına kaynağı belli olmayan bir takım hatalar karışmaktadır. Ölçmede kaynağı belli olmayan hatalara tesadüfi hata adı verilmektedir. Ölçme işiyle uğraşanların bütün gayretleri kaynağı belli olmayan bu tesadüfi hataları en aza indirme yönündedir. Bu bağlamda, ölçme sonuçları tesadüfi hatalardan arınık olduğu oranda güvenilirlik özelliğini göstermektedir. Bir başka ifadeyle güvenilirlik, ölçme sonuçlarının tesadüfi (random) hatalardan arınık olma derecesini ifade eden bir kavramdır (Turgut, 1986, s.31).

Yukarıdaki tanımlamaya eş değer olarak güvenilirlik, ölçme aracının duyarlılığı, tutarlılığı ve kararlılığı olarak da ele alınmaktadır. Bu ifadelerin tamamı ölçme aracının (testin) tesadüfi hatalara karşı kapalı olma derecesinin değişik görünümleri olarak ifade edilmektedir (Baykul, 1979, s.15).

Ölçme sonuçlarının güvenilirliği, sonuçların elde edildiği ölçme araçlarının güvenilirliği ile yakından ilgilidir. Bir başka ifadeyle, ölçme sonuçlarının güvenilirliği, onların elde edildiği ölçme araçlarının güvenilirliğine bağlıdır. Bir ölçme aracı yukarıda değinildiği gibi tesadüfi hatalara duyarlı,

tutarlı ve kararlı olduğu oranda ve dolayısıyla bu tür hatalara kapalı olma özelliğini gösterdiği oranda, ölçme sonuçlarının güvenilirliği de yüksek olacaktır.

Klasik Test Teorisinde, güvenilirlik katsayısı; “gerçek puan varyansının gözlenen puan varyansına oranı” olarak tanımlanmakta olup aşağıda verilen eşitlikte olduğu gibi gösterilmektedir (Lord, 1968, s.368).

$$\rho_x = \rho^2_{XT} = \frac{\sigma_T^2}{\sigma_X^2} \quad (1.15)$$

Eşitlik (1.15)’te verilen ve evren parametresi olan ρ_x güvenilirlik katsayısı, pratik uygulamalarda mevcut özelliğinden dolayı güvenilirlik hesaplamalarında kullanılamamaktadır. Bunun nedeni eşitlik (1.15)’te yer alan ifadelerden yalnız gözlenen puanlar varyansının σ_x^2 hesaplanmasının mümkün olmasına karşılık, gerçek puan varyansı σ_T^2 ve hata varyansının hesaplanamamasıdır. Söz konusu hata varyansının σ_E^2 ve gerçek puan varyansının σ_T^2 hesaplanamamasından dolayı, ölçme sonuçlarının ve dolayısıyla da ölçme araçlarının güvenilirlik katsayılarının hesaplanması için farklı değişik yöntemlerin kullanılmasını zorunlu hale getirmiştir. Bu cümleden hareketle, güvenilirlik katsayısının kestirilmesi için aşağıda kısaca değinilen yöntemlerin geliştirilmesine neden olmuştur. Bu yöntemler, bir testin iki eşdeğer formunu esas alan “paralel formlar yöntemi”, bir testin iki defa aynı gruba farklı zamanlarda uygulanması esasına dayanan “test-tekrar test yöntemi”, bir testin uygulanmasından elde edilen cevapların iki eş değer yarıya ayrılması esasına dayanan “yarılama yöntemi” ve bir testin uygulanması sonucu elde edilen madde puanlarının arasındaki ilişkileri inceleyen ve Kuder-Richardson KR-20 ve KR-21 numaralı eşitlikleri olarak bilinen güvenilirlik katsayıları yöntemleri olarak gösterilebilir .

Örtük Özellikler Teorisinde Güvenirlilik

Örtük Özellikler Teorisinde, Klasik Test Teorisindeki güvenilirlik kavramlarının yerini madde ve test bilgi fonksiyonları kavramları almaktadır.

Madde bilgi fonksiyonu,

$$I(\theta, u_i) = \frac{[P_i'(\theta)]^2}{P_i(\theta) \cdot Q_i(\theta)} \quad (1.16)$$

ve test bilgi fonksiyonu da

$$I(\theta) = \sum_{i=1}^n \frac{[P_i'(\theta)]^2}{P_i(\theta) \cdot Q_i(\theta)} \quad (1.17)$$

eşitliği ile tanımlanır. Bu eşitliklerdeki $P_i(\theta)$ seçilen Örtük Özellikler Modelinin olasılık fonksiyonunu, $P_i'(\theta)$ bunun türevini göstermekte olup $Q_i(\theta) = 1 - P_i(\theta)$ 'dır.

(1.16) eşitliği ile tanımlanan madde bilgi fonksiyonunun ifadesinde, $P_i(\theta)$ olasılık fonksiyonu, bunun varyansı ve türevi yer almaktadır. Türev değeri büyüdükçe ve varyans küçüldükçe kesir büyüyeceğinden maddelerden elde edilecek bilgi artmaktadır.

(1.17) eşitliğine göre, test bilgi fonksiyonunun madde bilgi fonksiyonlarının toplamı olduğu görülür. Bu durum, madde bilgi fonksiyonunun teste olan katkısının madde dışındaki faktörlere bağlı olmadan katkısını ifade etmektedir. Bu özellik, Örtük Özellikler Teorisinin, Klasik Test Teorisine olan bir başka üstünlüğüdür.

Örtük Özellikler Teorisinde ölçmenin standart hatası bilgi fonksiyonuna bağlı olarak tanımlanmakta olup madde ve test için aşağıdaki ifadelerle sahiptir.

$$SE_i(\theta) = \frac{1}{\sqrt{I(\theta, u_i)}} \quad (1.18)$$

$$SE_T(\theta) = \frac{1}{\sqrt{\sum I(\theta, u_i)}} \quad (1.19)$$

(1.18) ve (1.19) eşitliklerinden görüldüğü gibi, standart hata bilgi fonksiyonlarıyla ters orantılıdır. Maddenin veya testin verdiği bilgi arttıkça standart hata azalmakta, bilgi azaldıkça standart hata artmaktadır. Test bilgi fonksiyonunu, buna bağlı olarak test standart hatasının hesaplanabilir özellikte olması yeteneğin kestirilmesinde, kestirme işinin, belli bir hata sınırına ininceye kadar devam etmesine imkan sağlamaktadır. Bu husus, bireyselleştirilmiş testlerde kestirmenin hatasının sabit tutulması bakımından önemlidir.

Klasik Test Teorisindeki en önemli düşüncelerden bir tanesi de güvenilirlik katsayısının belirlenmesidir. Büyük örneklerde veya çok sayıda ölçümlerin yanında, güvenilirlik katsayısının dikkate alınması gereken dezavantajlarından bir tanesi, gruplara bağımlı olması ve buna dayalı olarak da genellenebilirlik özelliğinin sınırlanmış olmasıdır. İkinci bir eleştiri ise,

güvenirlilik katsayısının test maddelerini seçme konusunda belirli bir kurala bağlı kalmasıdır. Bu özelliği geçerlilik katsayısında görmek mümkün değildir. Buna ek olarak kestirimin (tahmin) standart hatası, σ_e , (Lord ve Novick, 1968,s.66-68) tarafından, aşağıdaki eşitlikte verilmiştir.

$$\sigma_e = \sigma_x [\rho_{xx}(1 - \rho_{xx})]^{1/2} \quad (1.20)$$

ve ölçmenin standart hatası ise,

$$\sigma_e = \sigma_x [1 - \rho_{xx}]^{1/2} \quad (1.21)$$

eşitliği ile hesaplanmaktadır. Bu iki fonksiyonda güvenirlilik katsayısına bağlı olup, yukarıda söz edilen aynı dezavantajlara sahiptir.

Madde ve test bilgi fonksiyonları, klasik test teorisindeki güvenirlilik katsayısı düşüncesine ve ölçmenin standart hatasına alternatifler sağlamaktadır. Bu araştırma, Örtük Özellikler Teorisinde KR-20 güvenirliliğinin hesaplanmasını da içermektedir.

Bireyselleştirilmiş (Adaptive) Test

Başlangıçtan bu noktaya kadar, eğitimde değerlendirmeye esas olma özelliği gösteren ölçme işleminin yapılması sonucunda elde edilen ölçme sonuçlarının analiz edilmesi için kullanılan birbirinden farklı özellikleri olan iki teori ve bu teorilerin bir takım varsayımları dolayısıyla kullanımdaki amaçlarına ilişkin bir takım bilgiler verildi. Bu iki teorinin (Klasik Test ve Örtük Özellikler Teorisi) dışında, bir yönden test geliştirme, diğer yönden bir test uygulaması gibi algılanabilecek olan diğer bir yöntemden daha söz edilebilir. Bu da bireyselleştirilmiş (adaptive veya tailored) testtir. Bireyselleştirilmiş test

psikometrik özellikleri önceden hesaplanmış çok sayıda madde arasından yapılacak uygun seçimlerle cevaplayıcıların özelliklerine uygun maddelerin seçilerek uygulanması suretiyle her birey için bütün maddelerin aynı olması beklenmeyen testtir.

Bireyselleştirilmiş testler, farklı bireylerin farklı düzeylerine göre her adaya farklı test maddelerinin uygulanması ile oluşturulan testler olduğuna yukarıda değinilmişti; bu yönü ile geleneksel testlerden ayrılmaktadırlar. Bireyselleştirilmiş testlere bu özelliğinden dolayı cevaba bağlı (response-contingent), otomatikleştirilmiş (automated), programlanmış (programmed), branşlaşmış (branched), bireyselleştirilmiş (individualized) ve bireye uydurulmuş (tailored) testler gibi isimler de verilmektedir.

Bir bireyselleştirilmiş testte, maddeler bireye uygulanır ve bireyin almış olduğu test maddesine verdiği cevabın doğru veya yanlış olmasına bakılarak, madde havuzundan değişik güçlük düzeydeki (farklı b parametrelerine sahip) maddeler bireye verilerek ölçme işlemi devam ettirilir. Bu yöntemle belirlenen maddeler, bireyin yetenek düzeyini ölçme doğrultusunda en verimli bir şekilde kullanılmış olur. Seçilen maddeler bireylerin önceki maddelere vermiş oldukları cevaplara bakılarak kestirilen yeteneklerine en uygun maddelerdir.

Bireyselleştirilmiş testlerin uygulanmasında bilgisayarlardan yararlanılmaktadır. Yani bir anlamda bireyselleştirilmiş testler bilgisayarlar için uyumlu bir hale getirilmiştir. Bilgisayarlar ile hazırlanan bireyselleştirilmiş testler bu test tekniğinin tüm avantajlarını çok iyi bir şekilde kullanmaktadır. Günümüzde hazırlanan birçok bireyselleştirilmiş test bireylerin yetenek değişikliklerini ölçmede kullanılırken, yine bu testler kişilik ve davranış (tutum) farklılıklarını ölçmede de başarılı bir şekilde kullanılmaktadırlar. Son zamanlarda başarı testlerinde kullanıma gayretleri de gözlenmektedir.

Bireyselleştirilmiş testlerin geleneksel kağıt-kalem testlerine göre birçok avantajları bulunmaktadır. Bu avantajlarına bakıldığında, daha verimli olmasıdır. Ölçme sırasında bireylerin düzeylerini geleneksel testlere göre daha kısa zamanda belirlenmesi ve tüm yetenek düzeyinde bireylerin ölçülen özellikleri daha az madde ile gözlenerek madde sayısında ekonomiklik sağlanmaktadır.

Geleneksel testler, bireyselleştirilmiş testlerle karşılaştırıldığında, bireyselleştirilmiş testlerin daha etkili ölçümler yaptığı görülmektedir.

Bir geleneksel test iki farklı şekilde hazırlanabilir. İlk olarak “sabit” test hazırlarken aynı zamanda “değişken” testler de ortaya çıkarılabilir. Sabit bir geleneksel testteki maddelerin güçlük dereceleri hep aynı düzeydedir. Bu özelliğe sahip bir test ile belirli düzeydeki bireylerin yetenek düzeyleri ölçülebilir. Fakat bu seviyenin altında veya üstünde aynı hata ölçüsünde ölçülemez. Üst düzeydeki bireyler için bu testler çok kolay olurken, alt düzeydeki bireyler için bu özellikteki test çok zor olabilir. Kısaca, sabit bir yetenek düzeyindeki testi farklı yetenek düzeyindeki bireylere uygulamak çok zor olmaktadır.

Değişken geleneksel testler ise birbirinden çok farklı güçlük derecelerine sahip test maddelerinin bir araya getirilmesiyle oluşturulan testlerdir. Bu özelliğe sahip testler düşük düzeylerdeki bireyler için zor maddeler, yüksek düzeye sahip bireyler için çok kolay maddeler olabileceği gibi normal düzeydeki bireyler için orta güçlükteki maddeler de bulunabilmektedir. Bu tip testlerde bütün bireyler için çok kolay maddeler olabileceği gibi çok zor maddeler de bulunmakta bu özelliğinden dolayı, değişken testler, değişik yetenek düzeylerinin ölçümünde eşit imkanlar sağladığı halde genel ölçüm özelliği oldukça düşüktür.

Yukarıda geleneksel testlerin yapısına bağlı olarak ortaya çıkan sözkonusu iki problem durumu “bireyselleştirilmiş testler” ile çözülebilmektedir.

Bu yöntemde ölçüme tabi tutulan bütün bireylerin kendi yetenek düzeylerine göre hazırlanan test maddeleri, daha önceden oluşturulan madde havuzundan seçilerek bireylere verilmekte ve bu problem durumu giderilmektedir. Bireyler için en uygun madde ise, yukarıda da belirtildiği gibi, bireyin mevcut durumdaki yeteneğine en uygun olan maddedir; bireyin mevcut durumdaki yeteneği de o ana kadar uygulanmış olan maddelerle kabul edilebilir bir hata sınırında ölçülmüş olan yeteneğidir. Bu test sayesinde aralarında çok az yetenek farklılığı bulunan bireylerin yetenek düzeylerindeki ayırım bile yapılabilmektedir.

Bireyselleştirilmiş testler ilk olarak 1900'lü yılların başlarında Fransa'da Alfred Binet tarafından (zeka testleri olarak) kullanılmaya başlanılmış ve bu testlerin uygulanmasına ilişkin bazı ilkeler onun tarafından oluşturulmuştur (Weiss:1983, s.5).

Binet'in bireyselleştirilmiş testlerinde başlangıç noktası bireyin doğru olarak cevaplandığı maddenin güçlük düzeyidir. Birey, maddeleri doğru olarak cevaplandığı sürece biraz daha zor maddeyi cevaplandırmak durumunda kalmaktadır. Bu cevaplandırma işlemi bireyin bir düzeydeki tüm maddeleri yanlış cevaplandırması aşamasına kadar devam eder. Böylelikle birey için hem taban, hem de tavan düzey belirlenmiş olur. Teste tabi tutulan birey taban ve tavan arasında, kendisine verilen maddelerin bazılarını yanlış bazılarını da doğru cevap verebilir.

Herhangi bir niteliği gözlemek (ölçmek) amacıyla kullanılan bütün testler bir takım test algoritmaları izlenerek uygulanmaktadır. Test algoritmaları sorulacak test maddeleri ve bu maddelerin sıralarını belirleyen bir kurallar bütünüdür. Bir testin uygulanmasında, test algoritması genellikle üç aşamadan oluşmaktadır. Bu aşamalar; a) teste başlama, b) testin devam ettirilmesi ve c) test uygulamasının sonuçlandırılmasıdır.

Teste başlama, cevaplayıcıya, testin başlangıcında hangi güçlükteki maddelerin sorulması; teste devam etme, cevaplayıcının doğru veya yanlış olarak cevapladığı her maddeden sonraki sorunun seçilmesi; testin bitirilmesi ise, soruların sorulmasına son verilmesi işlemine ait kuralların konulması veya kararların verilmesi işleridir. Bu sürece bireyselleştirilmiş test algoritması da denilmektedir.

Bu algoritma geleneksel test uygulamalarında oldukça basittir. Teste başlama, testteki ilk maddenin, devam etme onu izleyen maddelerin ve bitirme de sonuncu maddenin cevaplandırılmasıyla olur.

Bireyselleştirilmiş test uygulamalarında bu algoritma daha karmaşık işlemlere dayalıdır. Bu algoritma aşağıda ayrıntılı olarak açıklanmaktadır.

Teste Başlama: Bir cevaplayıcının testi almaya başlaması için ölçülen özelliğe göre bir başlangıç noktasının tayin edilmesi gerekir.

Başlangıç noktasının tayininde başvuru başlıca yöntemler şunlardır.

Benzer bir grubun ortalama yetenek düzeyi: Bu düzey, bir gruba önceden, ölçülen özelliğe göre, benzer bir grup testi uygulanarak saptanabileceği gibi, daha önceden uygulanmış bir testin ortalaması da olabilir. Uygulama kağıt-kalem testi yoluyla veya bilgisayarla yapılabilir.

Benzer yaş grubunun ortalaması: Testi cevaplayacak grubun yaşında olan bir gruba uygulanmış benzer bir test varsa, bu testin ortalaması başlangıç noktası olarak alınabilir.

Cevaplayıcı hakkında ölçülen özellikteki başarısı: Cevaplayıcının, ölçülen özelliğe ait buna benzer bir ölçme aracından aldığı bir puan varsa, bu puan başlangıç noktası olarak alınabilir.

Cevaplayıcının birkaç denemeden sonra kendisinin bulması: Bu halde cevaplayıcı serbest bırakılır, deneme yapmasına izin verilir. Bu deneme sonucuna göre bir başlangıç düzeyi seçmesi istenir, daha sonra test uygulamasına başlaması sağlanır.

$p=0,50$ ($b=0,00$) olan bir madde ile başlanması: Bu durumda bütün cevaplayıcılar güçlük düzeyi $b=0,00$ olan bir maddeden cevaplamaya başlatılır.

Ortak bir test verilip adayın bundan alacağı puan esas alınması: Bu halde cevaplayıcılara topluca bir test verilir, bu testten alacakları puanın standartlaştırılmış değerine en yakın güçlükteki maddeden başlaması sağlanır.

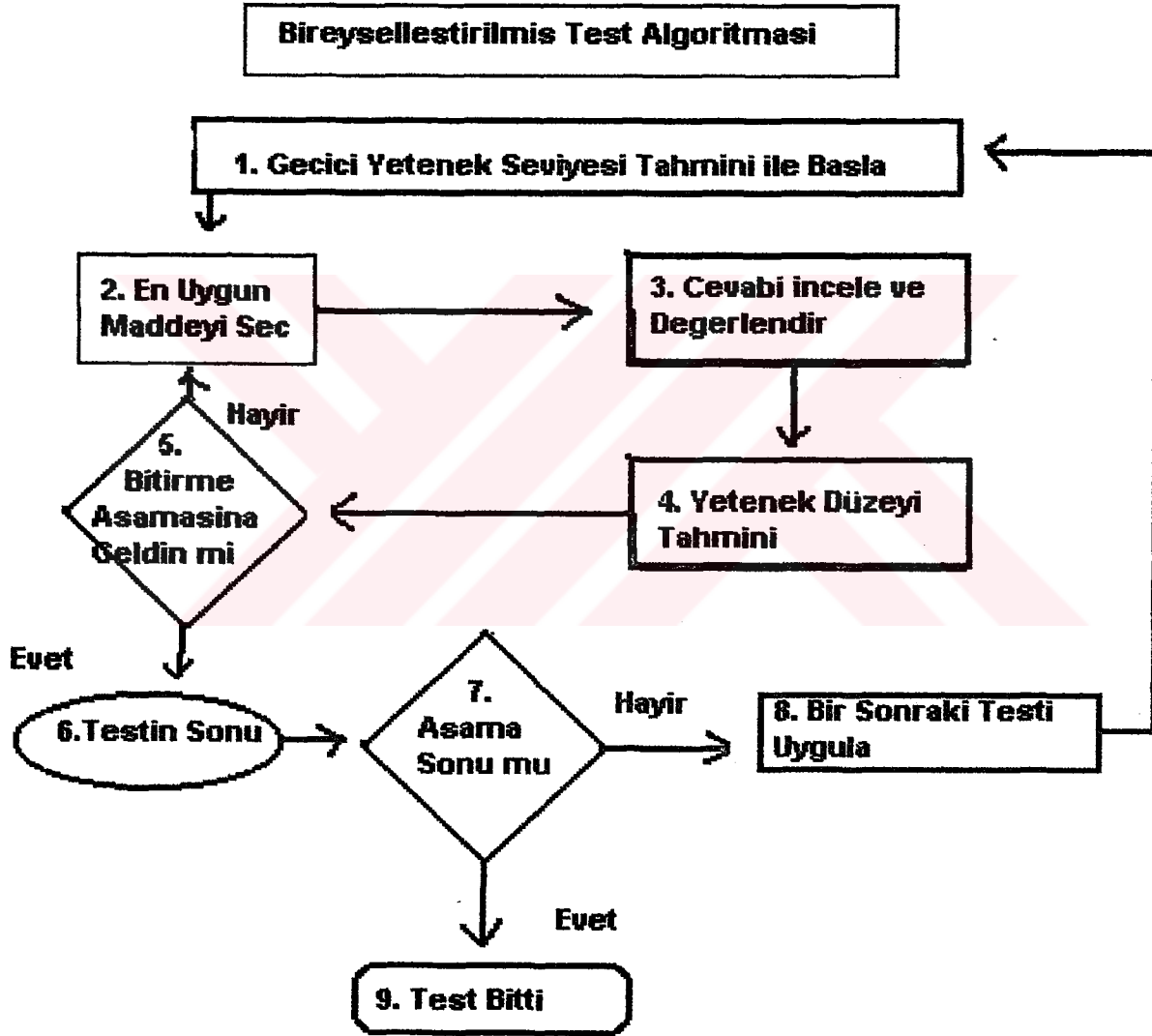
Testin Devam Ettirilmesi: Teste başlamada ilk madde yukarıda açıklanan yöntemlerden birisi yardımıyla gerçekleştirildikten sonra, havuzdaki maddelerden seçilecek olanların uygulanmasına sıra gelir. Bu uygulama için çok çeşitli stratejiler geliştirilmiştir. Bunlar aşağıda özetlenmektedir.

Testin Bitirilmesi: Bireyselleştirilmiş testlerin uygulamalarında test belli bir hedef noktaya ulaşıldığında test sonuçlandırılmaktadır. Bu belli bir madde sayısına ulaşılması veya daha önceden belirlenen bir zaman dilimiyle sınırlama veya belli bir hata düzeyine inme olabilir. Söz konusu yukarıdaki kurallar tek başına olabileceği gibi birlikte de uygulanarak test sonuçlandırılabilir.

Her cevaplayıcıya aynı eşit sayıda maddeyi yönelterek uygulama kolaylığı sağlar. Her cevaplayıcıyı daha önceden belirlenen bir kesinlik

seviyesine kadar test etmek, iyi bir yaklaşım ve test eden kişinin istediği kesinlikte bir sonuç verecektir. Ölçüt, maksimum bilgi veren madde seçimindeki hedef cevaba göre veya maksimum kesinlik veren ilk madde seçimine göre alınır (Thissen ve Mislevy:1990, s.114).

Yukarıda özetlenen bireyselleştirilmiş test uygulama algoritması aşağıdaki şekilde olduğu gibi gösterilebilir.



Şekil 1.4 : Bireyselleştirilmiş Test Algoritması
(Thissen ve Mislevy'den alınmıştır. S.108)

Uygulama Stratejileri

İki Aşamalı Test: Test maddeleri bir testteki yönlendirilebilecek en küçük birimdir. Teoride, her test maddesi cevaplayıcıya sorulduğunda testi yetenek seviyesine göre ayarlamak en iyi yoldur. Daha az uyum sağlayan test algoritmaları bizi iki aşamalı test algoritmalarına götürmektedir.

İki aşamalı testte, cevaplayıcı belirli bir testteki bütün maddeleri cevaplandırır. Bu geleneksel kağıt-kalem testi veya bireyselleştirilmiş test şeklinde olabilir. Bu, testin ilk aşamasını oluşturur. İlk aşamaya ilişkin puanlar belirlenir ve ikinci aşamaya geçilir. İkinci aşama testlerinin daha kolay veya tam tersine daha zor olan formları vardır. İkinci aşamadaki yer alan formlardaki maddelerin güçlük dereceleri cevaplayıcıların ilk aşamadaki test maddelerine vermiş oldukları cevaplara bağlı olarak belirlenir. İlk aşamada başarılı olan cevaplayıcılara, birinci aşamadaki maddelerden daha zor maddelerin oluşturduğu ikinci aşama testi; birinci aşamada beklenen düzeyde başarılı olamayanlara ise daha kolay maddelerden oluşan ikinci aşama testi uygulanır. Daha önce yukarıda değinilen test algoritmalarının üç aşamalı uygulama özellikleri bu testler için de geçerlidir

Çok-Düzeyle Format: Bu yöntemde yeni seçilecek madde tamamen cevaplayıcı hakkında bilinen bilgilere dayanmaktadır. Bu yöntemle hazırlanan test içindeki modüller güçlük derecelerine göre sıralanmakta ve cevaplayıcı en basit modülden başlayarak kendi seviyesine uygun olan modüle kadar devam etmektedir. Her düzeyin bir başlangıç ve bir de bitiş modülü bulunmaktadır. Testler genellikle sekiz-on modülden oluşmakta ve güçlük derecelerine göre sıralanmakta ve her hangi bir cevaplayıcı için düzey yaşı ve önceki test sonuçlarına bakılarak bulunmaktadır.

Kendinden-Seçmeli Test: İki aşamalı testte, puanlama aşamasında bazı problemlerin ortaya çıkmasından dolayı; cevaplayıcıdan cevaplandıracağı maddelerin düzeylerinin kendileri tarafından istenmektedir. Bu durumda cevaplayıcının yaşına ve düzeyine uygun olan test bulunup cevaplayıcıya uygulanmakta ve bu test cevaplayıcı için kolay ise bir sonraki birinci testten biraz daha zor olan düzeye geçmekte; bir sonraki düzey cevaplayıcı için zor geldiği takdirde, cevaplayıcı tekrar bir önceki düzeye geçebilmektedir.

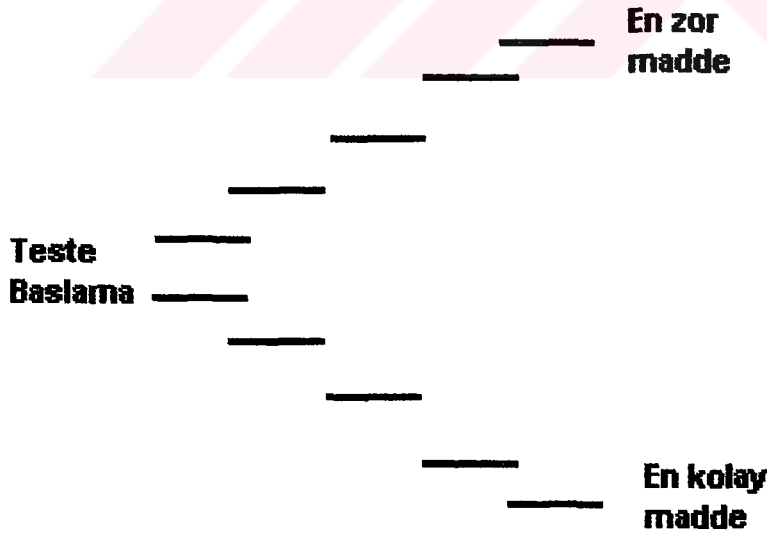
Prestwood ve Weiss (1977) cevaplayıcıların kendileri için kolay maddeleri seçtikleri ve bunun sonucunda bilgi fonksiyonunun verimli bir şekilde kullanılamadığını iddia etmişlerdir (Thorndike,1982, s.315). Bu durumun gerekçesi ise, cevaplayıcıların maddeleri kendi yetenek düzeylerine uygun bir şekilde seçerken bir de o anki durumları etkili olmaktadır. Testin uygulanması aşamasında çok sinirli ve huzursuz olan bir cevaplayıcı kendi yetenek düzeyinin altında maddeleri seçmeyi tercih edebilmekte; çünkü, cevaplarından emin olmayı tercih etmektedirler. Böylesine kendi düzeylerinden daha alt düzeye yönelik maddeleri seçen cevaplayıcıların yetenek düzeylerinin belirlenmesi durumunda onların gerçek yetenekleri hakkında gerçekçi bir bilgi elde edilemeyeceğinden belirlenen yetenek düzeyi çoğu zaman anlamsız olmaktadır. Yine bu tip uygulamalarda doğru olan cevaplayıcıların cevaplandırmaları gereken maddelerin kendileri tarafından seçilmeleri, ancak bu yöntemin her zaman etkili olduğunu söylemek çok zordur.

Bu yöntemin daha verimli olabilmesi için, çok düzeyli testte olduğu gibi modüllere ihtiyaç vardır. Modüller güçlük derecelerine göre sıralandırılmalı ve her modüldeki maddelerin güçlük derecelerinin birbiriyle çok yakın veya birbiriyle aynı olmalıdır. Cevaplayıcılar bu modülden bir kaç maddeyi cevaplandırarak kendilerine en uygun olan modülü seçerek maddeleri cevaplandırmalı ve her modüldeki maddeler cevaplandırıldıktan sonra bir sonraki

modüle geçmeli ve her cevaplayıcının aynı sayıda modülü cevaplandırmalarına özen gösterilmelidir.

Sonuçta her bir cevaplayıcı birden fazla modül içerisindeki maddeleri cevaplandıracak ve bu modülleri cevaplandırdıktan sonra elde edecekleri puanlara bağlı olarak değerlendirilecekler ve bu değerlendirmelerin de doğru olma olasılığı 0.50 olacaktır (Thorndike 1982; s. 316).

Değişken Test: Bu yöntemde de test maddeleri güçlük derecelerine göre sıralanmakta ancak başlangıç maddesi, maddelerin güçlük dereceleri dikkate alındığı zaman ortanca konumunda olan maddedir. Bu başlangıç noktasındaki maddeye göre diğer maddeler “ V “ harfinin kolları gibi iki düzeye ayrılmakta, bu kollardan biri gittikçe zorlaşan maddeleri içerirken, diğer kol bu durumun tam tersine gittikçe daha da kolaylaşan maddelerden oluşmaktadır.



Şekil 1.5 : Değişken Testte Madde Sıralanış Şekli

Cevaplayıcı ilk maddeyi doğru cevaplandırıdığı takdirde, güçlü ve düzeyi daha yüksek bir maddeyi cevaplandırması gerekiyor bu maddeyi de doğru cevaplandırıdığı takdirde güçlük düzeyi daha zor olan bir sonraki maddeye geçmekte ve test uygulaması bu şekilde devam etmektedir. Bu durum cevaplayıcının herhangi bir maddeyi yanlış cevaplandırıncaya kadar devam eder. Eğer cevaplayıcı bu noktada bir maddeyi yanlış cevaplandırmışsa, daha kolay maddelerin yer aldığı kola atlamakta ve bu düzeyin ilk maddesinden cevaplandırmaya devam eder. Cevaplayıcının teste devam etme işleminde cevaplayıcının belirli sayıdaki maddeyi cevaplandırıncaya kadar devam eder.

Bu yöntemde cevaplayıcı, bir maddeyi yanlış cevaplandırıdığı takdirde “V” ‘nin alt kolunu geçer maddeleri cevaplandırmaya devam eder. Cevaplayıcı maddeleri cevaplandırırken sırada hangi maddenin olduğunu ve doğru veya yanlış bir cevaplama sonra hangi maddeyi cevaplandırması gerektiğini bilmelidir. Bu uygulamada cevaplayıcı için belirlenecek puan cevaplayıcının doğru cevap sayısıdır. Bu özellik, bu yöntemin dezavantajlarından biridir.

Bu yöntemde bazı maddelerin bütün cevaplayıcılar için çok zor olması gerekir. Testin ilerleyen aşamalarına doğru maddeler çok geniş bir aralığa yayılması gerekiyor. Orta düzeydeki bir cevaplayıcı bile çok zor düzeydeki bir madde ile çok kolay güçlük derecesine sahip bir maddeyi cevaplandırmalıdır. Bu yöntemde cevaplayıcılara yöneltilen maddeler cevaplayıcının yetenek düzeyinden uzaklaştığı oranda cevaplayıcıya ilişkin elde edilen bilgi fonksiyonuna eklenen değer gittikçe azalmaktadır.

Piramitsel Test: Piramitsel test modeli, daha önceden güçlük derecesine göre ve aşama sırasına göre düzenlenmiş maddelerin oluşturduğu modeldir. Bu modelin gösterimini kolaylaştıracak bir şekil aşağıda gösterilmiştir.

A S A M A L A R					
	1	2	3	4	5
					11 Zor
				7	12
			4	8	13
Teste Başlama	1	2	5	9	14
		3	6	10	15 Kolay

Şekil 1.6 : Piramitsel Testte Madde sıralanış Şekli

Bu modele göre bütün cevaplayıcılar ilk maddeyi cevaplandırarak başlamaktadırlar ve ilk maddeyi doğru cevaplandıranlar ikinci maddeyi cevaplandırırken, ikinci maddeyi doğru cevaplandıramayanlar ise üçüncü maddeyi cevaplandırmaktadır. Her aşamadaki maddeleri doğru cevaplandıranlar daha zor bir düzeydeki maddeler ile karşılaşmakta, maddeleri doğru cevaplandıramayanlar ise daha kolay düzeydeki maddeleri cevaplandırmak yoluna gitmektedir. Piramitsel test modelinde aşama sayısı arttıkça madde havuzunda daha fazla maddeye ihtiyaç duyulmaktadır.

Bu yöntem, cevaplayıcının bir önceki durumuna göre o andaki sorulacak maddeye karar verilmesini gerektirmektedir. Maddeler piramide daha önceden yerleştirildiğinden, her madde bir defa kullanılabilir. Bu özellik test hazırlayıcısına çok fazla bir yük getirmektedir. Bunun anlamı test hazırlayıcısı çok fazla sayıda test maddesi hazırlamak zorunda kalacaktır.

Piramit içerisinde oluşturulan her yeni aşamada bilgi fonksiyonu daha fazla büyüdüğü için birbirini izleyen maddeler arasındaki ilişki gittikçe azalmaktadır (Thorndike, 1982, s. 317).

Piramitsel testlerde bir çok aşamadan oluşturulmuş ise ve test bilgisayar ile uygulanmıyorsa cevaplarda bazı karışıklıklar ortaya çıkabilir ve bu da piramitsel testlerin dezavantajlarından biri olarak kabul edilmektedir (Thorndike, 1982, s.317).

Piramitsel bir test yöntemine göre test hazırlamak isteyen birisi şu üç soruyu cevaplandırmak durumundadır: (1) testin içerisinde kaç aşama olmalıdır? (2) her aşamada madde sayısı ne olmalıdır? (3) ve puanlama nasıl yapılmalıdır? Şimdi bu noktalar üzerinde kısaca durulmaktadır.

Testin içerisinde kaç aşama olmalıdır? Testin içerisindeki aşama sayısı, testten beklenen kesinliğe göre ve madde havuzunda yer alan madde sayısına göre değişmektedir. Tablo 1.1'e bakarak aşama sayısına göre gerekli olan madde sayısı bulunabilir.

Aşağıda verilen tablo 1.1'e göre piramitsel testteki aşama sayısı 15'den fazla değil ise α_j değerleri için kabul edilebilir kesinlik değerleri elde edilmektedir. Bir grup için 4/10 standart hata değeri kabul edilebilir bir hata değeri olarak görüldüğü takdirde, madde ayırdedicilik $\alpha_j = 0.50$ olduğunda, piramitsel test için aşama sayısı 21 olmalı ve bu durumda cevaplayıcıların yetenekleri arzu edildiği şekilde kestirilebilmektedir.

Tablo 1.1
Belirli Uzunluktaki Testler İçin Ölçmenin
Standart Hata (Ö.S.H) Değerleri*

Aşama Sayısı	Madde Sayısı	a_j d e ğ e r l e r i					
		0.30	0.50	0.70	1.00	1.500	2.00
		ÖSH	ÖSH	ÖSH	ÖSH	ÖSH	ÖSH
3	6	.926	.833	.742	.631	.508	.433
6	21	.866	.727	.611	.487	.364	.293
9	45	.816	.654	.530	.409	.295	.232
12	78	.774	.598	.474	.358	.254	.197
15	120	.737	.554	.432	.322	.226	.174
18	171	.706	.518	.399	.295	.205	.158
21	231	.678	.488	.373	.274	.189	.145

* Burada $C = 0$ alınmış ve her maddenin cevaplayıcıların o anki düzeyi için uygun olduğu kabul edilmiştir. Bu tablo Thorndike (1982, s.320)'den alınmıştır

Yine $a_j = 0.70$ olduğunda testin içerisinde olması gereken aşama sayısı 18 ve test maddesi sayısı ise 171 olmalıdır; $a_j = 1.00$ olduğunda ise aşama sayısı 10 olmalıdır (Thorndike, 1982, s. 320).

Eğer test maddelerinin a_j değerleri 0.70'e eşit veya daha büyük değil ise piramitsel test oluşturmak ve bundan elde edilecek kestirimlerin çok etkili ve verimli olacağını söylemek pek mümkün değildir. a_j değerleri 1.00 olan ve 10 maddelik bir test homojen güçlükte olup, yüksek bilgi fonksiyonuna sahip olup ve çok dar bir aralıkta çok yüksek değerlerde kestirim değerleri elde edilmektedir.

Her aşamada yer alan madde sayısı ne olmalıdır? Piramitsel testin her aşamasındaki maddelere bakılarak, bu aralıktaki farklı kesinlik değerleri karşılaştırılabilir. Bireyselleştirilmiş testlerde hedef çok geniş bir aralıkta

homojen kesinlik deęerleri elde etmek olduęuna gre, piramitsel testlerde de bu zellik dikkate alınmalıdır. Byle bir durum saęlandığı takdirde, piramitsel test nce maksimum geniřlięe daha sonra minimum dzeye ulařır.

Her ařamadaki madde sayısı yukarıda verilen tablo 1.1’de gsterilen a_j deęerlerine gre ka maddeyi iermesi gerektirdięine bakılarak karar verilebilir.

Piramitsel testlerde puanlama nasıl olmalıdır? Bu yntemlerle hazırlanan testlerde verilen doęru (D) veya yanlış (Y) cevaplara gre deęiřik puanlamalar yapılmaktadır. Bu yntemde puanlama amacıyla en yksek olasılık (maximum likelihood) yntemi kullanılmaktadır.

Tabaka Uyumlu Testler (The Stradaptive Test): Weiss (1973) tarafından ortaya atılan bu test modelinde maddeler tabaka tabaka glk derecelerine ayrılmıř olup her tabakadaki maddelerin glk dereceleri aynı dzeyde yani homojen bir yapıdadır. Cevaplayıcının yařına, derecesine (sınıf dzeyi) ve daha nceki test sonularına bakılarak uygun bir tabaka seilmekte ve cevaplayıcı bu tabakanın ilk maddesini cevaplandırarak teste bařlamaktadır (Thronkike, 1982, s. 302).

Eęer cevaplayıcı ilk maddeyi doęru cevaplandırırrsa bu tabaka ierisindeki daha zor maddeye gemektedir. Cevaplayıcı maddeyi yanlış cevaplandırdığı takdirde daha kolay tabakaya gemektedir. Sonu olarak, test cevaplayıcının dzeyine yakın olan iki- tabakadaki maddelerin cevaplandırılması ile sonulandırılmaktadır.

Bu yntem biraz daha detaylandırıldığında, glk derecelerine gre maddeler tabakalara ayrılmıřtır. Cevaplandırıcı her bir tabaka ierisinde yer alan maddelere vermiř olduęu cevapların doęru veya yanlış olması durumuna gre bir tabakadan dięer bir tabakaya gemektedir. Bu test ynteminin, piramitsel test

yönteminden farkı ise bir tabakadaki maddeler aynı anda cevaplandırılmamasıdır (Thorndike, 1982, s. 326).

Tabaka uyumlu bir test hazırlayan bir kişi temel olarak aşağıda belirtilen problemlerle karşılaşmaktadır. Bu problemler sıralandığında;

Test kaç tabakadan oluşmalıdır?

Her tabakada yer alacak madde sayısı ne olmalıdır?

Tabakaların içerisinde yer alan maddelerin sıralanması nasıl belirlenmelidir?

Bir cevaplayıcının bir tabakadan bir başka tabakaya geçmesinde etken ne olmalıdır?

Cevaplayıcının testin sonuna geldiğine nasıl karar verilecektir?

Cevaplayıcıya verilecek puanlar nasıl belirlenecek?

Şimdi bu aşamada yukarıda başlıklar şeklinde verilen soruların cevapları çok kısaca verilmektedir.

Tabaka sayısı ne olmalıdır? Bir tabakanın genişliği maksimum bilgi fonksiyonu sağlayan madde sayısı ile doğru orantılıdır. Bu özellikte madde ayıricılık parametresi olan a_j değerlerinin gruplar arası dağılımına bağlıdır. Bir test içerisindeki her hangi bir maddenin madde ayıricılık parametre (a_j) değeri arttıkça, o maddenin bilgi fonksiyon aralığı daralmakta ve maksimum bilgi fonksiyonu en üst düzeye ulaşmaya başlamaktadır. Maksimum bilgi sağlayan (% 90 dolaylarında) en az iki tabakaya ihtiyaç vardır (Thorndike 1982; s. 327).

Aşağıda tablo 1.2’de verilen θ yetenek düzeyinde farklı a_j değerlerinin (%90 ve %80) etkililik düzeylerini göstermektedir. İki tabaka kullanılarak minimum %90 verimlilik elde etmek istendiğinde, a_j değerlerinin 1.00 civarında

olan maddelerin kullanılması gerekir. Ayrıca her tabakanın genişliği 0.50 ve 0.60 standart sapma aralığında olmalıdır.

10 tabakadan oluşmuş bir test ile istenilen düzeyde bilgi fonksiyon değerleri elde edilirken (maksimum % 90 etkililikle), 5 standart sapma aralığı oluşturulmalıdır (θ değerinin -2.5 den + 2.5 ' e kadar).

Madde ayırtıcılık parametre değerleri ($a_j = 1.00$) kullanıldığı bir tabakadan sonra gelen bir başka tabaka ile arasındaki etkililik farkı hiçbir zaman %20'yi aşamaz (0.50 ve 0.30 veya 0.60 ve 0.40 gibi). Tabakalar daraldıkça cevaplayıcılar ilerleyen tabakalarda daha zor maddeler ile karşılaşacaklardır. Böylece cevaplayıcıların performansları bütün tabakalara yayılacak ve böylelikle farklı tabakalar için bulunan kesinlik değerleri, genel kesinlik değerini olumsuz yönde etkileyeceği söz konusu olmaktadır. Bu yüzden daha az sayıda tabaka ve daha geniş tabakalar tercih edilmelidir. Eğer madde etkililik değerini minimum %80 olarak belirlenmişse, $a_j = 1.00$ değeri ile elde edilecek tabaka genişliği -2.8 ile + 2.8 aralığında 7 tabakadan oluşturulmalıdır. Böyle bir durumda ortaya çıkacak madde etkililik değeri ise %90'ların üstüne çıkacaktır (Thorndike 1982; s.328).

Tabakaların sayısı, cevaplayıcıların yaşına, derecesine (sınıf düzeyine) ve sınavın gerekçesine bağlıdır. 6 standart sapma birimlik bir aralıkta 10 yaşındaki bir grubu değerlendirmek mümkündür. Böyle bir testin 7-18 yaş aralığına uygulanabilmesi için de 10 standart sapma birimine ihtiyaç vardır. Böyle bir durumda ise test içerisinde yer alacak tabaka sayısının neredeyse iki katına çıkacaktır.

Tablo 1.2
Her Tabakada Yer Alan Maddelere Ait Farklı a_j
Değerleri İçin Ölçmenin Hata Varyans Değerleri

Her Tabaka İçin Madde Sayısı	a_j d e ğ e r l e r i				
	0.50	0.70	1.00	1.20	1.50
2	.628	.476	.326	.262	.198
3	.528	.373	.237	.183	.133
4	.455	.306	.185	.141	.098
5	.394	.259	.152	.114	.078
6	.351	.224	.128	.097	.064
7	.318	.198	.111	.083	.055
8	.288	.177	.098	.072	.048
9	.265	.160	.087	.064	.042
10	.245	.146	.079	.058	.038

Her tabakada yer alacak madde sayısı ne olmalıdır? Her tabakada olması gereken madde sayısı cevaplayıcıya verilmesi düşünülen madde sayısının yarısı kadar olmalıdır. Maddelerin maksimum etkililik düzeyinde çalıştığı düşünülürse, farklı a_j değerleri ve farklı tabaka sayıları için ölçmenin standart hata değerinin hesaplanan değerleri yukarıda tablo 1.2’de gösterilmiştir.

Tablo 1.2’de gösterilen ölçmenin hata varyans değerleri, yeteneğin belirlenmesinde kullanılan a_j değerlerinin 1.00 olmasına bağlı olarak güvenirlik katsayılarından elde edilmiştir.

Burada test maddelerine verilen cevapların tahmine (şansa) dayalı olmadığı kabul edilmiştir. Yukarıda verilen tablo 1.2’ye bakıldığında her bir tabakada kaç maddenin yer alması gerektiğini görmek mümkündür.

Her tabaka içerisinde yer alan maddelerin sıralanması nasıl olmalıdır? Sıralamada hangi maddenin daha önce geleceği ve bu maddeyi takip edecek maddelerin hangileri olacağının belirlenmesinde Weiss (1973) şunu önermiştir: Maddeler tabaka içerisinde sıralanırken madde ayırıcılık parametresi dikkate alınmalı ve madde ayırıcılık parametresi en yüksek olan maddelerden başlanmalı ve sıralama bu özelliğe doğru yapılmalıdır. Bunun gerekçesi ise madde ayırıcılık gücü yüksek tüm maddeler bütün cevaplayıcılar için kullanılacağından ve böylece testin söz konusu yeteneğe ilişkin performans değerlerini daha etkili bir şekilde kestirilmesini sağlayacaktır. Yine buna bağlı olarak testten elde edilen sonuçların doğruluk değerleri yüksek olacaktır. Madde ayırıcılık parametre değerlerinin bir birlerine çok yakın değerlerde olması durumunda cevaplayıcıların ilgi düzeyleri, kültürel uyumlulukları veya cevaplayıcıların ortak yanlarını içeren diğer özellikler dikkate alınarak maddelerin sıralanması yapılmalıdır (Thorndike, 1982, s. 330).

Bir cevaplayıcının bir tabakadan başka bir tabakaya geçişi nasıl olmalıdır? Bu yöntemde en çok kullanılan tabaka değiştirme kuralı olarak şu görülmektedir: Cevaplayıcı kendisine yöneltilen bir maddeyi doğru cevaplandırdığı takdirde bir üst tabakaya geçmekte ve yine maddeye yanlış cevap vermesi durumunda ise alt tabakaya geçmektedir.

Bir cevaplayıcının testi bitirmesine nasıl karar verilir? Ölçmede hata sabit fakat buna karşılık madde sayısı değişken olduğu zaman, bir testin uzunluğu bütün adaylar için aynı olamaz. Çünkü, cevaplayıcılara ait bilgi fonksiyon değerleri bir birine eşit olmayabilir veya cevaplayıcıların maddelere vermiş oldukları cevaplara karşılık geçmiş oldukları tabaka sayılarını bir birlerinden farklı olabilir, bundan dolayı her cevaplayıcı farklı sayıdaki maddelerden oluşan testi alması sonrasında test bitirilir. Bu yöntemle ilgili bir test uygulayıcısının dikkate alması gereken bir özellik, tabakadaki bütün maddeler bitinceye kadar teste devam edilmesi ve eğer tabakada k tane madde var ise her cevaplayıcı en az $2k$

- 1 tane maddeyi cevaplandıracaktır. Bu özellik yerine getirildiği zaman teste son verilebilir.

Cevaplayıcının testten aldığı puan nasıl belirlenecektir? Eğer kullanılan test bilgisayar kullanımını gerektiren türde hazırlanmış ise cevaplayıcının yeteneğine ilişkin hesaplamalar Maksimum Likelihood yöntemi kullanılarak cevaplayıcının yetenek değeri kestirilir.

Bireyselleştirilmiş Testlerde Karşılaşılan Bazı Güçlükler

Bireyselleştirilmiş testlerde karşılaşılan güçlükler bakıldığında şu sıkıntıların ortaya çıktığı görülmektedir: (1) içerikten dolayı kaynaklanan problemler, (2) güç eksikliği, (3) test maddelerinin güçlük derecelerine göre sıralanmasıyla ilgili problemler ve (4) yüksek teknolojinin etkisine ilişkin problemler.

(1) İçerikten dolayı kaynaklanan problemler denilince genellikle, aralarında bağlantı bulunan test maddelerinin yanlış yorumlanmasından kaynaklanmaktadır. Geleneksel testlerde yani aynı testi alan bütün bireylerin testin içerisinde yer alan aynı maddeleri cevaplandırmaya çalıştığı bir testte böyle bir problem olmamaktadır (Wainer ve Kiely, 1986).

Bireyselleştirilmiş testlerde ise testi alan bireylere göre daha önceden düzenlenen ve bireylerin farklı testlere tabi tutulması sırasında ortaya çıkmaktadır. Bu içeriğe bağlı olarak (a) bilgi gerektiren test maddelerinin test içerisinde yanlış yerlere konulması (bilgi yerleştirimi), (b) çapraz bilgi ve c) belirli bir konunun ve buna bağlı olarak da testte, konuyla ilgili test maddelerinin gereğinden fazla tekrar edilmesidir (dengelenmemiş içerik).

İçerikten kaynaklanan problemleri biraz daha geniş bir şekilde iyi anlaşılabilmesi için bu problem durumlarının açıklanması gerekmektedir. Bu problemler küçük başlıklar altında verilmektedir.

Bilgi Yerleştirimi:

Testlerde yapılan önemli hatalardan biri de bilgi isteyen maddelerin yanlış yerlere yerleştirilmesidir. Maddelerin test içerisinde yanlış yerlere yerleştirilmesi bireylerin performansını dolayısıyla da madde parametrelerini etkilemektedir.

Wainer ve Kiely (1986), Whitely ve Dawes'e (1976) atıfta bulunarak şunu söylemektedirler; Whitely ve Dawes'in yaptıkları bir çalışmada 60 maddelik bir testte 15 maddenin yerlerini değiştirerek bir gruba uyguladıktan sonra sözkonusu 15 maddeden 9'nun parametrelerinde farklılık tespit etmişlerdir. Aynı şekilde Yen'de (1980), Wihtely ve Dawes'in bulgularını destekler nitelikte bulgular elde etmiştir.

Çapraz Bilgi

İçerikten kaynaklanan bir başka problem durumu da belirli bir bilginin maddenin cevabına yönelik cevaplayıcılara kolaylık sağlaması olarak adlandırılan “çapraz bilgi”dir. Bu problem durumuna şu örnek verilebilir;

Madde 1- Karbondioksit (CO_2) suya eklenince ne oluşur?

a) Karbonik asit b) Maden Sodası c) Hidrojenkarbonat

Madde 2- CO_2 ne olarak bilinir?

- a) Karbondioksit b) Karbonmonoksit
c) Karbonatlı su d) Ozon

Yukarıdaki örneklerde, ilk madde konuyla ilgili olarak bireyin düzeyi hakkında ayrıntılı bilgi istemekte ve seçici bir test maddesi olarak kabul edilebilir. Ancak hemen sonra ilk maddeyi izleyen ikinci maddenin cevabı birinci maddede yer almaktadır. Bu durum bir probleme işaret etmektedir. Aslında birinci maddeyi yine ayırdedici özelliği yüksek ve güçlük derecesi yüksek başka bir madde izlemeliydi. Çoğu zaman bu özelliğe bireyselleştirilmiş testlerde dikkat edilmemekte ve bu durum da problem olma özelliği göstermektedir.

Dengelenmemiş İçerik

İçerikten kaynaklanan bir başka problem de belirli bir konunun ve buna bağlı olarak da maddenin test içerisinde çok tekrar edilmesidir. Bu problem durumunu ise dengelenmemiş içerik olarak adlandırmak mümkündür.

Dengelenmemiş içerikle ilgili bir örnek olarak şu özellik örnek gösterilebilir; çoğu zaman bireyselleştirilmiş testlerde cinsiyet, ırk, farklı kültürel yaşama şekli pek fazla dikkate alınmadan hazırlanabilmekte (özellikle A.B.D'lerinde). Bu özelliklerin önemli olduğu testlerde bireyler arasında adaletsizlikler ortaya çıkma olasılığı çok yüksektir (cinsiyete dayalı rollerin ne olduğunun belirlendiği bir sınavda cinsiyetle ilgili test maddeleri dengeli bir şekilde hazırlanmamışsa).

Geleneksel (klasik) testlerde bu dengesizlik test hazırlayıcıları veya puanlayıcılar tarafından belli bir oranda ortadan kaldırılabilir. Fakat bireyselleştirilmiş test tekniklerinde bu pek mümkün gözükmemektedir. Çünkü, puanlayıcı durumunda olan bilgisayarlar insanlar gibi çalışmadıklarından bu

problem durumu ortadan kalkmaz. Böyle bir durumda teknoloji bu problemi bireyselleştirilmiş testlerde dengelenmemiş içeriğe ait probleme yok edebilecek özellikten yoksun gözükmemektedir.

Güç (kuvvet) Eksikliği

Bireyselleştirilmiş test tekniğinin önemli bir avantajı kısa bir zaman dilimi içerisinde doğru ve güvenilir test sonuçlarının elde edilmesidir. Ancak, bireyselleştirilmiş testlerin bu özellikleri bazı durumlarda problem doğurmaktadır.

Geleneksel klasik testlerde belli bir konuda teste tabi tutulan bireylerin, sözkonusu konuya tam anlamıyla hakim olup olmadığını kontrol etmek mümkün olabilmektedir. Özellikle de sayısal yeteneği ölçmek amacıyla kompozisyon tipindeki maddelerin yer almasıyla hazırlanmış bir testi örnek göstermek mümkündür.

Buna karşılık bireyselleştirilmiş test tekniklerine göre hazırlanmış testlerle bu pek mümkün gözükmemektedir. Özellikle de konunun detayları ölçme yapan kişi için çok önemli olduğu durumlarda bu problem kendini göstermektedir.

Test Maddelerinin Güçlük Derecelerine Göre Sıralanmaları:

Burada test oluşturulurken testin en küçük birimi olarak adlandırılabilen test maddelerinin güçlük derecelerinin dikkate alınmadan testin içerisinde yerleştirilmelere bir sıkıntıya neden olmaktadır (Wainer ve Kiely, 1986).

Geleneksel testlerde test maddeleri testin içerisine yerleştirilirken şu özellik test hazırlayıcıları tarafından dikkate alınmaktadır; genellikle testi alan

bütün bireyler tarafından doğru olarak cevaplandırılması mümkün olabilecek güçlüğüne sahip test maddeleri testin hemen başına alınır; daha sonra test maddelerinin güçlük dereceleri kademeli bir şekilde arttırılarak testin içerisinde sıralanmasıyla seçicilik özelliği sağlanır. Basit test maddelerinin testin başına yerleştirilmesinin nedeni, düşük yetenek düzeyine sahip bireylerin motivasyonlarını sağlayarak onların kendilerine olan güvenlerini sağlamaktır. Böyle bir yaklaşım ayrıca bireylerin kendilerine olan güvenlerinin sağlanması yoluyla diğer test maddelerinin doğru olarak cevaplandırmalarını teşvik etmektedir.

Geleneksel testlerdeki bu özelliğe karşı bireyselleştirilmiş test tekniklerinde yukarıda vurgulanan özelliklerin görülmesi pek mümkün değildir (Wainer ve Kiely, 1986).

Bireyselleştirilmiş testlerde, üst düzey yeteneğe sahip olan bireyler test maddelerini kolayca cevaplandırırken alt düzey yeteneğe sahip olan bireylerde tam bunun tersi bir durum ortaya çıkmakta ve daha sonraki test maddelerini cevaplandırmaktan kaçınmaktalar .

Yüksek Teknolojinin Etkisi

Çok büyük paraların harcanarak bu şekilde hazırlanan testlerden elde edilen verimliliği arttırmak amaç edinilmiş olsa da bu tip testlerde kullanılan bazı madde tipleri bu verimliliği engelleyici nitelikte olduğu görülmektedir. Buna örnek olarak özellikle paragraflara bağlı olarak geliştirilen test maddelerini göstermek mümkündür. Özellikle seçilen paragrafların çoğu zaman bilgisayar ekranına sığdırılması pek mümkün olmayabilir; bu paragrafa bağlı olarak test maddelerini cevaplandıracak bireyler bu paragrafların tamamını ekranda görememesi onun motivasyonunu olumsuz yönde etkileyebilir. Aynı zamanda aynı paragrafla ilgili olarak çok sayıda test maddesi sorulmuşsa bireyin bunları cevaplandırmak için tekrar geri dönerek paragrafa ulaşması bir problem olarak görülebilir. Böyle bir yöntem ne kadar etkili ve verimli olabilir?

Yukarıda kısaca değinilen bu problemlerin giderilmesi için belli bir takım geleneksel anlamda çözümler getirmek için şu özellikler dikkate alınmalıdır.

İçerikten Kaynaklanan Problemlere Çözüm

Bir ölçme aracı olan testlerin bireylerin herhangi bir alana ilişkin yeteneklerinin ve yetenek düzeylerinin belirlenmesi amacıyla bireyselleştirilmiş testlerin kullanıldığı daha önce belirtilmiş ancak uygulamada ortaya içerikten dolayı bir takım problemlerin çıktığı ifade edilmişti.

Yukarıda sözkonusu olan problem durumunun giderilmesi için nasıl bir yol izlenmelidir ki bu problem ortadan kaldırılabilsin veya en aza indirmek mümkün olabilsin?

İçerikten kaynaklanan problemlere mantıklı bir çözüm, kullanılmakta olan madde parametrelerine yeni bir takım bilgilerinin eklenerek madde daha anlaşılır hale getirmekle mümkün olabilir. Test maddesine eklenen yeni bilgiler açıklayıcı veya kodlama şeklinde olabilir (Wainer ve Kiely, 1986).

Bu kodlama özellikle bireylerin demografik özelliklerini içeren maddeler (bu tip maddelere aynı zamanda demografik madde adı verilmektedir) için geçerli olmaktadır. Bu kodlamalarda cinsiyete, ırka, bağlı bulunduğu gruba göre sınıflandırarak hangi maddeleri kimlerin cevaplandırması gerektiği belirtilirse “çapraz bilgi” problemi önlenmiş olur. Ve böylelikle bireyler arasında dangesizlik ya da adaletsizlik önlenmiş olur. Hemen şunu da belirtmekte yarar vardır; bu eklemeler test maddesinin kolay anlaşılmasını sağlamaz ancak maddenin kimlere yönelik olduğunu belirler. Kısacası o maddeyi kimlerin cevaplandırması gerektiğini belirler.

İçerikten kaynaklanan problemlere bir başka çözüm yolu ise bireylerin yetenek farklılıklarını dikkate alarak başka bir yöntem aramaktır. Bu yöntemle bireyler yetenekleri doğrultusunda karşılaştırılarak seçim yapılabilir. Ancak bu yöntemin geçerliği henüz tam anlamıyla ısıpatlanmış değildir (Wainer, 1987).

Güç (kuvvet) Eksikliğinden Doğan Problemlere Çözümler:

Bireyselleştirilmiş test tekniğinin bir başka amacı da bireylerin sahip oldukları yetenekleri belirli bir düzen içerisinde ortaya çıkarıp geliştirmektir. Fakat testlerin verimliliği arttıkça kuvvet veya doğruluk buna ters orantılı olarak azalmakta ve sonuçta testlerin doğruluğu ve geçerliği yeterli bir düzeye ulaşamamaktadır.

Doğruluk ya da kuvvet eksikliğine karşı uygulanabilecek bir yol ise, bireysel değerlendirmelerde doğruluğa aşırı bir derecede önem vermektir. Örneğin bir testte yeterli sayıda doğru cevap veremeyen bireyler yeterli derecede puan almış olsalar bile başarısız olabilmelidir.

Başka bir yol bireyselleştirilmiş test tekniği ile yoklamada bir model baz alınabilir. Böylelikle bireyler önlerinde bulunan modeli gözönünde bulundurarak yeterli başarı düzeyine ulaşmaları mümkün olabilir.

Böyle bir yöntemde, bireylerin cevapları ile model uyuşmuyor ise birey devre dışı bırakılabilir ve doğruluk ve kuvvet bu şekilde oluşturulabilir. Bireyselleştirilmiş test tekniğinde bireyin yanlış cevap verdiğini belirten ve bundan sonrasında bireyin maddeleri cevaplandırmasını durduran bir kural kullanmak tercih edilen bir seçenek olarak kullanılabilir (Wainer ve Kiely, 1987).

Test Maddelerinin Sıralanmasıyla İlgili Problemlere Çözüm

Test maddelerinin güçlük derecelerine göre sıralanması aşamasında bireylere ait daha önceki performanslarının bilinmesi gerekmektedir. Eğer bireylerin daha önceki performansları biliniyorsa hazırlanacak test maddeleri bireylerin düzeylerine uygun olarak hazırlanmalıdır. Böyle bir durumda bireyler pek fazla zaman kaybetmeden mevcut performanslarının üzerine çıkmak için daha fazla efor harcayacaktır.

Yüksek Teknoloji Kullanımından Doğan Problemler

Yüksek teknolojinin test tekniğinde kullanılmasıyla test sonuçlarında belirli bir oranda geçerlilik ve doğruluk derecesinin arttığı görülmektedir. Ancak teknolojiden kaynaklanan problemlerin giderilmesi için de bir takım önlemlerin test hazırlayıcıların dikkate almasıyla giderilmesi olasıdır.

Bireyselleştirilmiş testlerin uygulanması için kullanılan bilgisayarların özellikleri dikkate alındığında özellikle ekrana ilişkin özellik, pasaja dayalı test maddeleri kullanılacağı zaman pasajların uzunlukları önem teşkil etmektedir. Ekrana sığabilecek boyuttaki pasajlar kullanım için tercih edilerek uygun bir formatta test maddeleri oluşturulduğu vakit söz konusu problemlerin ortadan kalkması mümkün olabilecektir.

Bireyselleştirilmiş testlerde karşılaşılan diğer önemli bir problem testin kapsam geçerliliğinin düşmesidir. Bu problem bireyselleştirilmiş tek boyutlu testlerde sıkça rastlanan bir durumdur. Bunun ortadan kaldırılması, bireyselleştirilmiş testleri konularına veya boyutlarına göre alt testlerden oluşturmakla gerçekleştirilebilir. Ancak böyle bir durumda testin uygulanmasında kullanılacak bilgisayar programının hazırlanması önemli ve zor olmakta; ayrıca madde havuzunun da bu boyutlara göre oluşturulması gerekmektedir.

Anılan problemleri açık bir şekilde ortaya koyan ve yine bu problemlere ilişkin bir takım çözümler öneren Wainer ve Kiely bunun da ötesine giderek yeni bir model önermektedirler; bu modelin adını da testlet olarak adlandırmaktadırlar.

Bireyselleştirilmiş Testlerin Güvenirliği

Bireyselleştirilmiş testlerde güvenirliliğin hesaplanmasına dair bir çalışma Thissen, Steinberg ve Moony tarafından (1989) testletler üzerinde yapılmıştır. Bu araştırmacılar, bir geleneksel okuma parçasına dayalı olarak geliştirdikleri testte testlet yaklaşımını kullanmışlardır. Verdikleri örnekte, dört okuma parçası ve bu okuma parçalarına bağlı olarak sırasıyla her test 7, 4, 3, ve 8'er maddeden oluşturulmuştur. Örtük Özellikler Teorisine dayalı bir test ile karakterize edildiğinde (tanımlandığında), bilgi fonksiyonu ile tanımlanmasından daha büyük bir doğruluk değeri ortaya çıkmaktadır.

Klasik Test Teorisinde güvenirlilik indeksi, X gözlenen, T gerçek puanı, ρ_{XT}^2 gözlenen ve gerçek puanlar arasındaki korelasyonu, σ_X^2 ve σ_E^2 de gözlenen hata puanlarının varyanslarını göstermek üzere,

$$\rho_{XT}^2 = \frac{\sigma_T^2}{\sigma_X^2} \quad (1.21)$$

veya

$$\rho_{XT}^2 = \frac{\sigma_X^2 - \sigma_E^2}{\sigma_X^2} \quad (1.22)$$

eşitlikleriyle tanımlanır. Testletlerin güvenirliliklerinin hesaplanmasında Klasik Test Teorisinin bu formüllerinden yararlanılabileceği gibi Örtük Özellikler Teorisinin modellerinden de yararlanılabilir.

Örtük Özellikler Teorisi güvenilirlik indeksini (ρ) hesaplamak için, Thissen 1989 ve Bock 1972 modelleri incelenebilir. Bu modellere bakıldığında J tane testlet, $J= 1,2, \dots, j$ şeklinde numaralandırılmıştır. Her testlette, m_j tane madde bulunmaktadır. Bu nedenle j'inci testlet için, $X_j = 0,1,2, \dots, m_j$ tane madde farklı cevap elde edilmektedir.

$X = 0,1,2, \dots, m_j$ değerleri J testlet için;

$$T_{jx}(\theta) = \frac{\exp^{a_{jk}\theta + C_{jk}}}{\sum_{k=0}^m \exp^{a_{jk}\theta + C_{jk}}} \quad (1.23)$$

fonksiyonu ile gösterilmektedir. Burada (a_{jk}, c_{jk}) j, madde-cevap fonksiyonunun şeklini belirleyip, “madde-kategori parametreleri” adını almaktadır. Eğer bu model bilgi ile uyum gösterirse, bilgi fonksiyonun değeri hata varyansının göz önünde bulundurulmasıyla hesaplanabilir.

Bu çalışmada aynı özelliği ölçen Klasik Test Teorisine göre geliştirilmiş bir testle aynı özelliği ölçmeye yönelik bireyselleştirilmiş test arasındaki KR-20 güvenilirliğinin karşılaştırılması çalışılmaktadır.

Problem cümlesi aşağıdaki gibi ifade edilmiştir.

Problem Cümlesi

Türkçe okuduğunu anlama alanında klasik test teorisi ve İki-parametrelili lojistik model kullanılarak geliştirilen bireyselleştirilmiş testin KR-20 güvenilirlikleri ve bir grup öğrencinin bu iki yöntemle elde edilen yetenek ölçüleri arasında manidar bir farklılık var mıdır?

Bu problemdeki soruyu cevaplandırabilmek için aşağıdaki alt problemlerde yer alan sorulara cevap aranmıştır.

Alt Problemler

1- Türkçe okuduğunu anlama alanında Klasik Test Teorisi ve İki-parametrelili Lojistik model kullanılarak geliştirilen bireyselleştirilmiş testin KR-20 güvenirlikleri arasında manidar bir farklılık var mıdır?

2- Türkçe okuduğunu anlama alanında bir grup öğrencinin Klasik Test Teorisi ve İki-parametrelili Lojistik model kullanılarak geliştirilen bireyselleştirilmiş teste göre elde edilen yetenek ölçüleri arasındaki ilişki düzeyi nedir?

Araştırmanın Önemi

İçinde bulunduğumuz yüzyılın son çeyreğinde öğretimi bireyselleştirme çalışmaları hızlı bir artış göstermiştir. Bu artışa paralel olarak ölçme alanında da grup testlerinin yanında bireyselleştirilmiş testlere önem de artmış; bu alanda yapılan araştırmalarda yoğun bir şekilde artışlar olmuştur. Özellikle 1970’li yıllardan sonra bireyselleştirilmiş testler üzerinde yoğun çalışmalar gözlenmiştir.

Bireyselleştirilmiş testler de bu çalışmaların bir grubunu oluşturmaktadır. Ayrıca teknolojik gelişmeler de bu artışı hızlandırmıştır. Bireyselleştirilmiş testler, çok sayıda öğrenciyi aynı anda salonlarda teste tabi tutulma zorunluluğunu ortadan kaldırdığından uygulamada büyük kolaylıklar getirmektedir.

Eldeki çalışmanın, ölçme alanında, bu tip çalışmalara bir katkı getirebilmesi hedeflenmiştir.

Bu araştırmanın bir başka önemi ise yapılan literatür taramasında Türkiye’de bu yaklaşımla ilgili arzu edilen sayıda araştırma bulunmamaktadır. Bu çalışma, konuyla ilgili az sayıdaki araştırmalardan biri olmaktadır.

Bu araştırmada aşağıdaki sayıtların varlığı kabul edilmiştir.

Sayıtlar

1- Araştırmada kullanılan ölçme aracının geliştirilmesi sırasında alan (konu) uzmanı olarak kabul edilen öğretmenlerle bilgisine başvuru alan diğer uzmanların cevaplarında düşünce ve duyguları aynen olduğu gibi yansıtılmıştır.

2- Bu çalışmada kullanılan bireyselleştirilmiş testin tekboyutluluk ve yerel bağımsızlık özelliğine sahip olduğu varsayılmaktadır.

Araştırma, aşağıda belirtilen sınırlamalar dahilinde yürütülmüştür.

Sınırlamalar

1- Araştırma, lise düzeyinde “okuduğunu anlama yeteneği” ile sınırlı tutulmuştur.

2- Araştırma, uygulamanın yapıldığı 1996-97 öğretim yılında öğrenimlerini sürdüren Denizli il merkezi ve Tavas ilçe sınırları içerisindeki lise öğrencileri ile sınırlı tutulmuştur.

Araştırma içerisinde geçen bazı kavramlar aşağıdaki anlamlarda kullanılmıştır.

Tanımlar ve Kısaltmalar

Güvenirlilik: Bir ölçme aracının ölçmek istediği niteliği olabildiği kadar az tesadüfi hata ile ölçebilme derecesidir.

Adaptive Test: Bilgisayar uygulamasına dayalı bireyselleştirilmiş test.

Bireyselleştirilmiş Test: Psikometrik özellikleri bilinen çok sayıda maddenin bir araya getirilmesiyle, belirli bir niteliği ölçmek amacıyla (bireye göre ayrı ayrı olarak) uygulanan test.

Testlet: Aynı düzeyde aynı yeteneği ölçen ve bir birim olarak kabul edilen alt test.

İlgili Araştırmalar

Baykul (1979) yaptığı çalışmada, Klasik Test Kuramına ve Örtük Özellikler Kuramına (Birnbaum'un üç parametrelî modeline göre) dayalı olarak geliştirmiş olduğu matematik testlerin geçerlik ve güvenirliklerini karşılaştırmıştır.

Baykul'un çalışmasının evrenini Ankara'daki genel liselerin birinci ve ikinci sınıf öğrencileri alınmış olup, örneklemini ise tesadüfi olarak Fen, Deneme, Halide Edip, Çankaya ve Aktepe liselerinin birinci ve ikinci sınıflarındaki 1064 öğrenci oluşturmuştur.

Baykul araştırmasında elde ettiği sonuçlar şu şekilde özetlenebilir:

1 - Üç parametrelî lojistik modele göre oluşturulan testin şans başarılarından arındırılarak ve arındırılmadan hesaplanan KR-20, eşit iki yarı güvenirlikleri Klasik Test Teorisine göre geliştirilen testin güvenirliğinden manidar şekilde büyük bulunmuş; KR-21 güvenirlikleri arasında fark manidar bulunmamıştır.

- 2- Her iki testin aynı öğrenci grubuna uygulanması sonucu elde edilen yetenek ölçüleri arasında yüksek düzeyde ilişki bulunmuştur ($r = 0,85$)
- 3- Her iki test de şans başarısından manidar düzeyde etkilenmektedir ($\alpha = 0,01$)
- 4 - a_g ve b_g madde parametreleri örneklemden çok az etkilendiklerinden daha kararlı testlerin geliştirilebileceği ifade edilmektedir.

Az sayıda çalışılan bireyselleştirilmiş testler konusunda yapılan çalışmalardan biri de Kaptan'a aittir. Kaptan (1994) yaptığı çalışmada, yetenek kestiriminde bireysel olarak bilgisayarla uygulanan bireyselleştirilmiş test yöntemi ile grup testini (geleneksel test) karşılaştırmaktadır. Karşılaştırmaları, test süreleri, madde sayıları, sayısal yetenek düzeyleri ve söz konusu düzeylere ilişkin dış ölçüt olarak alınan ÖYS Matematik ağırlıklı puanları arasındaki ilişkileri incelemiştir.

Kaptan araştırmasında aşağıda belirlenen sonuçlara ulaşmıştır:

- 1- Bireyselleştirilmiş test ve grup testinde kullanılan sürelerin ortalamaları arasındaki fark bireyselleştirilmiş testin lehine anlamlı bulunmuştur.
- 2- Bireyselleştirilmiş test uygulama zamanı açısından tasarruf sağlamaktadır.
- 3- Bireyselleştirilmiş test yöntemi madde sayısında da tasarruf sağlayarak, grup testi ile aralarında anlamlı bir fark bulunmayacak şekilde, yeteneği daha az sayıdaki madde ile kestirmek mümkün olmaktadır.
- 4- Her iki test yöntemi ile kestirilen sayısal yetenek düzeyleri arasında anlamlı bir fark olmadığını gözlemiştir.

5- Bireyselleştirilmiş test ve grup testinden elde edilen yetenek düzeyleri, dış ölçüt olarak alınan ÖYS matematik ağırlıklı puan ile anlamlı düzeyde ilişki gösterdiğini gözlenmiştir.

6- Bireyselleştirilmiş test ve grup testinden elde edilen yetenek ölçüleri ile ÖYS matematik ağırlıklı puan arasında ayrı ayrı elde edilen ilişkilerin birbirinden anlamlı düzeyde farklı olmadığını belirlemiştir.

7- Bireyselleştirilmiş test sonucunda bireylerin cevap dağılımlarına bakarak, doğru cevapladıkları maddelerin güçlük ortalamalarına göre sıralanmış, bu sıranın maximum likelihood (en çok olabilirlik) kestirim yöntemini kullanan bilgisayar programından elde edilen yetenek düzeyleri sırası ile aynı sayılabilecek bir ilişkinin olduğu gözlenmiştir.

Mc Kinley ve Reckase (1980), tek parametrelili lojistik ogive (1 PL) ile üç parametrelili lojistik ogive (3 PL) modellerini sözel başarının kestirilmesinde, hem kendi aralarında hem de grup testleri denilen kağıt-kalem testleri ile karşılaştırarak, 1 PL ve 3 PL modellerinin test-tekrar test güvenilirliklerini incelemişlerdir. İnceleme sonucunda 3 PL için $r = 0,77$, 1 PL için $r = 0,61$ olarak bulmuşlardır. Ayrıca 1 PL ve 3 PL test güvenilirliklerinin kağıt-kalem testi güvenilirliklerinden daha yüksek olduğunu göstermişlerdir.

Yine, Mc Kinley ve Reckase'nin (1983) başka bir çalışmasında, yetenek kestiriminde bireyselleştirilmiş test kullanılarak, bir parametrelili lojistik model ile üç parametrelili lojistik modellerinin karşılaştırılması yapılmıştır. Çalışmanın ilk aşamasında gerçek verilen üzerinde, ikinci aşama da ise benzetilmiş veriler üzerinde çalışılmıştır. İlk aşamada Assesment Mathematics Usage alt testindeki 40 madde için 3000 cevaplayıcıdan elde edilen veriler kullanılmıştır. İlk 2000 cevaplayıcı ile ilgili veriler her iki model için madde parametrelerinin elde edilmesinde kullanılmıştır. Elde edilen sonuçlar diğer 1000 cevaplayıcının

verilerinden 1 PL ve 3 PL bireyselleştirilmiş testler geliştirilmiştir. Başka bir ifadeyle ilk 2000 cevaplayıcının verileri ön deneme uygulaması yerine, 1000 cevaplayıcının verileri ise esas uygulama yerine kullanılmıştır. Her iki bireyselleştirilmiş test yönteminde de en çok olabilirlik kestirim yöntemi (maximum likelihood estimation) ve maximum information madde seçimi işlemleri uygulanmıştır. Daha sonra bu iki kümenin yetenek kestirim değerleri karşılaştırılmıştır. Yetenekler, standart normal dağılımdan seçilmiştir. Her iki aşamanın sonuçlarına göre tek parametrelili lojistik model ve üç parametrelili model ile uygulanan bireyselleştirilmiş testleri, yetenek kestiriminde birbirine göre daha avantajlı bulunmamıştır (Aktaran Kaptan, 1994, s.29)

Bütün bunların yanında, bireyselleştirilmiş test uygulamasının, grup testlerine göre üstün olan yönlerinden birinin de daha az sayıdaki madde ile de testin tamamlanabileceğini belirtmektedir.

Thissen, Steinberg ve Mooney (1991) okuduğunu anlama yeteneğine ilişkin teslet şeklinde geliştirilen testin güvenilirliğini, klasik test teorisinin yöntemi ile madde-cevap teorisinin üç-parametrelili lojistik modelini karşılaştıran bir çalışma yapmışlar; çalışmalarından elde ettikleri sonuçlar şöyledir. Dört testlete ait geleneksel α katsayısını 0.62 ve ρ değerini ise 0.66 olarak hesaplanmış; r uyuma değeri ise 0.66 olarak bulunmuştur. Yine bu testletlerin içerisinde yer alan maddeleri testlet şeklinde değil de 22 maddelik bir test olarak kabul ederek her iki yöntemi uyguladıklarında α değerini 0.70 ve ρ değerini ise 0.74 olarak hesaplamışlardır. Yine bu 22 maddenin r uyuma değerini ise 0.65 olarak hesaplamışlardır.

Wainer, Sireci ve Thissen (1991), A.B.D’inde kullanılmakta olan ve SAT-V (Sözel Akademik Yetenek Testi) diye adlandırılan sözel yetenek testini biraz değiştirerek 12,13,10 ve 10’ar maddelik dört ayrı pasajı 45 maddelik bir test haline dönüştürerek, testletlerle karşılaştırdıklarında, 45 maddelik testten

elde edilen gvenirlik katsayısının testletlerden elde edilen gvenirlik katsayısından daha dřk olduėunu grmřlerdir. Yine gvenirlik katsayılarını cinsiyete gre karřılařtırdıklarında, gvenirlik katsayısının cinsiyete gre de farklılık gsterdiėi ortaya çıkmıřtır. Bu arařtırmada 45 maddeye ait gvenirlik katsayıları sırasıyla $\alpha = 0,86$ ve $\rho = 0,88$ olarak hesaplanmıřtır. Aynı maddeleri testlet řeklinde dřnerek hesapladıkları gvenirlik katsayıları yine sırasıyla $\alpha = 0,76$ ve $\rho = 0,80$ olarak hesaplanmıřtır.

BÖLÜM II

YÖNTEM

Bu çalışma temel araştırma niteliğinde bir çalışma olduğundan bir evren örneklem ilişkisine bakılmamıştır. Temel araştırmalar, sadece varolan bilgiye yenilerini katmak amacıyla olan araştırmalardır. Bu tip araştırmalar, araştırma kavramının en yalın bir biçimde temsil edildiği çabalardır. Bu nedenle bu tip araştırmalara temel (basic) araştırma denmektedir (Karasar, 1995, s.24). Eldeki araştırma temel araştırma olarak düşünülmüş ve Denizli merkezindeki liselerden, Ticaret Meslek Lisesi, lise ikinci sınıfı 69 öğrenci seçilerek araştırmada yer alan alt problemlere cevap bulabilmek amacıyla uygulamaya alınmıştır. Ticaret Meslek lisesinin seçilmesindeki neden ise söz konusu lisedeki bilgisayar laboratuvarının uygulama için uygun özelliklere sahip olmasıydı.

Veri Toplama Aracı

Veri toplama aracının hazırlanmasında, Türkçe okuduğunu anlama alanında cümle tamamlama, eş anlamlı kelime, deyim ve atasözü tipindeki maddelerden yararlanılmıştır.

Ölçme aracının geliştirilmesi için yukarıda anılan alana ilişkin test maddelerinin yazılması aşamasında, bu alandaki iki uzmandan yararlanılmıştır. Alan uzmanlarının hazırladığı test maddeleri Denizli ili Tavas Lisesi'nde denemeye tabi tutulmuştur. Deneme için 134 madde hazırlanmıştır. Hazırlanan 134 madde Tavas lisesi, ikinci sınıfındaki 132 öğrenciye uygulanmıştır. Denemeden elde edilen ham verilerin bireyselleştirilmiş test için madde havuzu oluşturabilmek amacıyla BILOG bilgisayar paket programı kullanılarak analiz edilmiştir. Ancak, kullanılan bilgisayarın kapasitesi verilerin analizi için yeterli olmadığından dolayı, madde sayısı 100'e indirilerek analiz

gerçekleştirilebilmiştir. Geriye kalan 34 madde ve bu maddelerden elde edilen ham veriler çalışmanın her hangi bir safhasında kullanılmamıştır (Ölçme aracından çıkarılan maddeler, madde güçlük indeksleri düşük olan maddelerdir).

Söz konusu 100 maddeden yararlanılarak bir madde havuzu oluşturulmuş olup bu madde havuzundan 61 madde seçilerek bireyselleştirilmiş test temeline dayalı 13 adımlık bir uygulama gerçekleştirilmiştir.

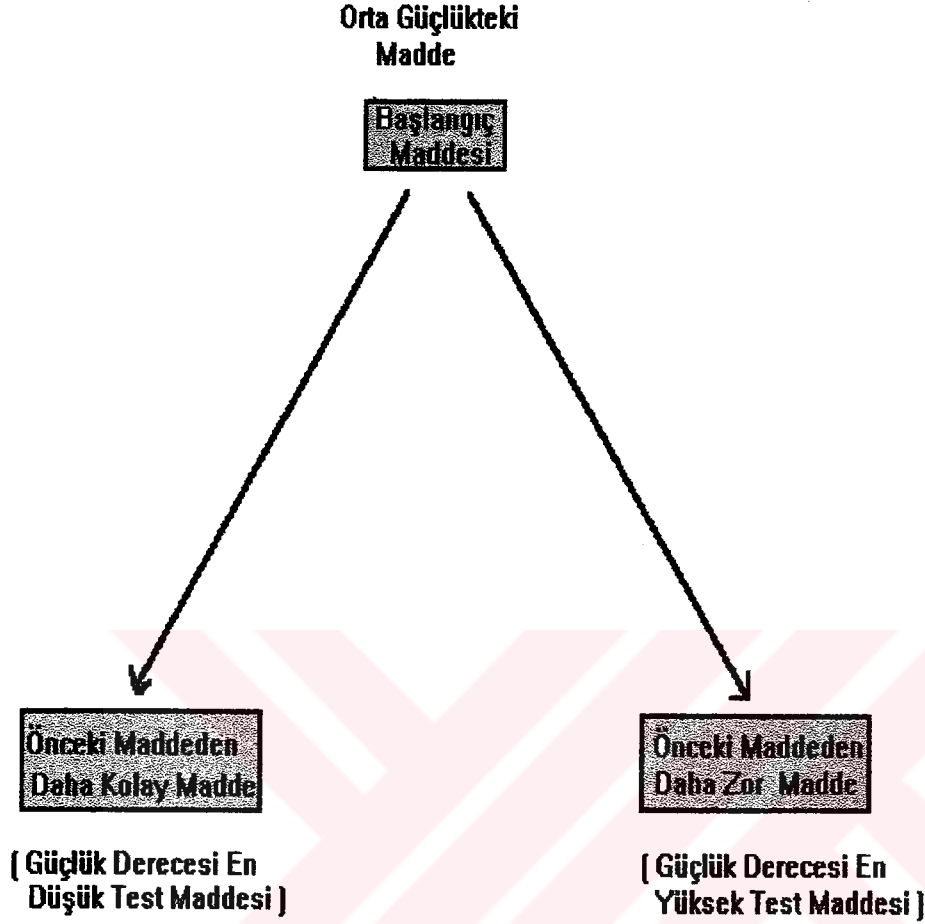
Her bir cevaplayıcının 13 adımda tamamlayabildiği bireyselleştirilmiş testin maddelerinin seçiminde ve testin içerisinde sıralanmasında İki-parametrelili lojistik modele göre maddelerin güçlük ve ayıricılık parametreleri dikkate alınmıştır.

Bireyselleştirilmiş test maddelerinin sıralanması biçimsel olarak üçgene benzeyen bir şekilde hazırlanmış olup, üçgenin tepesinde tüm cevaplayıcıların ortak olarak cevaplandıkları başlangıç maddesi yer almaktadır. Üçgenin başlangıç noktasında bulunan madde, deneme sonuçlarına göre madde güçlük parametresi dikkate alındığında tam orta noktasında (ortancası) bulunan madde olduğu görülmektedir.

Bu araştırmada kullanılan bireyselleştirilmiş test yönteminin algoritması, her cevaplayıcının başlangıç noktasında yer alan maddeyi doğru ya da yanlış cevaplandırmasına bağlı olarak bir sonraki adımda daha zor ya da daha kolay maddenin cevaplandırılması esasına dayandırılmıştır. Bu durum aşağıda şekil 2.1’de gösterilmektedir.

Cevaplayıcının kendilerine yöneltilen soruları cevaplandırabilmeleri için, madde havuzunda yer alan soruları şekil 2.1’de gösterilen strateji doğrultusunda

maddeyi seçip cevaplayıcının çalıştığı bilgisayarın ekranına getiren , BASIC bilgisayar program diliyle yazılmış bir bilgisayar programı kullanılmıştır.



Şekil 2.1 : Bireyselleştirilmiş Teste Ait Cevaplandırma Algoritması

Bu uygulamada BASIC bilgisayar diliyle yazılan program test maddesini ve bu test maddesine alternatif cevap olarak verilen seçenekleri ekrana getirmekte; cevaplayıcı maddeyi okuduktan sonra, A, B, C, D ve E seçeneklerinden kendisine göre birini doğru cevap olarak kabul ederek, o seçeneği klavyedeki tuşa basarak cevabını vermektedir. Cevaplayıcı anahtar (doğru) cevabın tuşuna basmış ise daha zor bir soru, aksi bir durumda ise daha kolay bir soru ekrana gelmektedir. Bu araştırmaya katılan cevaplayıcılar 13 adımda (13 soruyu) uygulamayı (bireyselleştirilmiş testi) bitirmişlerdir. Bu

uygulamada cevaplayıcılar, daha programın başında cevaplama işlemine başlayabilmesi için öğrenci numarasına girmek zorundadır. Öğrenci numarasını girdikten sonra program başlangıç noktasında yer alan soruyu ekrana getirmektedir. Uygulama bittikten sonra ise program hangi cevaplayıcının hangi maddeyi doğru ya da hangi maddeyi yanlış cevaplandırıldığını bir kütükte saklamaktadır. Aşağıda bu durumu gösteren ve bir cevaplayıcıya ait sonuç verilmiştir.

Öğrenci Numarası : 2649

Doğru Cevap Sayısı : 10

Yanlış Cevap Sayısı : 3

<u>Sorular</u>	<u>Cevap</u>	<u>Puanı:</u> 98.29
8	Doğru	
9	Doğru	
13	Doğru	
15	Doğru	
23	Doğru	
24	Doğru	
28	Doğru	
30	Yanlış	
46	Doğru	
50	Doğru	
52	Doğru	
53	Yanlış	
54	Yanlış	

Verilerin Elde Edildiği Grup

Bu araştırma; “Türkçe okuduğunu anlama alanında Klasik Test Teorisi ve İki-parametrelili lojistik model kullanılarak geliştirilen iki testin KR-20 güvenirlikleri ve bir grup öğrencinin bu iki yöntemle elde edilen yetenek ölçüleri arasında manidar bir farklılığın olup olmadığını” anlamak amacıyla yapılmış bir çalışmadır. Bu çalışma aynı zamanda temel araştırma kabul edildiğinden evren örneklem ilişkisine bakılmamış bu nedenle evren örneklem tayinine gidilmemiştir.

Denemeye 132 öğrenci katılmıştır. Denemeden elde edilen ham veriler BILOG bilgisayar paket programıyla analiz edildikten sonra, toplam 61 maddeden ve 13 adımda cevaplandırılabilen 61 maddelik bir havuz oluşturulmuştur. Oluşturulan bu havuz Denizli ili merkezindeki Ticaret Meslek Lisesi, ikinci sınıftaki 69 öğrenciye uygulanmıştır. Araştırmada cevaplandırılması gereken alt problemlere çözüm bulabilmek amacıyla gerekli olan sonuçlar söz konusu bu 69 öğrenciye ait uygulama (ölçme) sonuçlarından elde edilmiştir.

Bireyselleştirilmiş test uygulamasına katılan 69 öğrencinin madde havuzundan cevaplandırdıkları madde numaraları aşağıda tablo 2.1’de verilmiştir.

Tablo 2.1
69 Öğrencinin Madde Havuzundan
Cevaplandıkları Madde Numaraları

Öğrenci No	Öğrencilerin Cevaplandıkları Madde Numaraları												
2499	8	9	13	14	23	24	25	26	34	36	37	38	46
2649	8	9	13	15	23	24	28	30	46	50	52	53	54
2689	1	3	7	8	19	20	22	23	34	36	37	38	46
2763	2	3	7	8	23	24	25	26	34	36	37	38	46
2784	8	9	13	14	16	18	22	23	34	35	37	38	46
2785	8	9	10	12	23	24	28	29	46	54	55	59	61
2800	8	9	10	12	23	24	28	30	34	36	37	38	46
2808	8	9	10	11	23	24	25	26	46	50	51	53	54
2812	8	9	13	14	23	24	28	30	34	36	37	38	46
2813	1	3	7	8	16	18	22	23	46	48	49	53	54
2847	8	9	10	12	23	24	28	29	46	48	49	53	54
2879	1	3	7	8	16	18	22	23	46	48	49	53	54
2945	8	9	10	11	23	24	25	26	32	33	37	38	46
2962	8	9	13	14	23	24	28	29	46	48	49	53	54
2967	8	9	13	14	16	18	22	23	46	54	55	56	57
3019	8	9	13	15	23	24	28	30	33	36	37	38	46
3027	8	9	13	14	23	24	25	26	46	54	55	56	57
3039	8	9	13	14	23	24	28	30	46	54	55	59	61
3074	8	9	13	15	16	17	23	24	46	54	55	56	57
3085	8	9	13	14	23	24	25	26	46	54	55	56	57
3141	8	9	13	14	23	24	28	29	46	50	52	53	54
3251	8	9	13	14	23	24	25	26	31	33	37	38	46
3284	8	9	13	14	16	18	22	23	34	35	37	38	46
3299	8	9	13	15	23	24	28	30	38	39	43	44	46
3306	8	9	13	14	23	24	28	29	34	35	37	38	46
3330	8	9	13	15	17	18	22	23	46	50	52	53	54
3434	8	9	13	14	23	24	28	29	34	35	37	38	46
3540	2	3	7	8	17	18	22	23	34	36	37	38	46
3574	8	9	13	14	16	18	22	23	38	39	43	45	46
3676	8	9	13	15	23	24	25	26	34	35	37	38	46
3685	8	9	13	14	17	18	22	23	46	54	55	56	58
3748	8	9	13	14	23	24	28	29	32	33	37	38	46
3753	8	9	10	11	23	24	25	28	46	54	55	56	57
3754	8	9	13	14	23	24	28	29	34	36	37	38	46
3756	2	3	7	8	19	20	22	23	46	51	52	53	54
3758	8	9	13	14	23	24	25	26	46	50	52	53	54
3762	2	3	7	8	23	24	28	29	46	54	55	56	57
3773	8	9	13	14	23	24	25	26	46	54	55	59	60
3795	8	9	13	15	19	20	22	23	38	39	44	45	46
3911	8	9	13	15	23	24	28	29	46	48	49	53	54
3912	8	9	13	14	23	24	25	26	34	35	37	38	46

3972	8	9	13	14	23	24	28	29	46	47	49	53	54
4049	8	9	13	14	17	18	22	23	38	39	40	42	46
4070	8	9	13	14	16	18	22	23	46	54	55	59	60
4085	8	9	10	11	19	20	22	23	31	33	37	38	46
4089	8	9	13	14	23	24	25	26	34	36	37	38	46
4103	8	9	13	15	16	17	22	23	37	38	43	45	46
4111	8	9	13	14	19	20	22	23	34	36	37	38	46
4162	8	9	10	11	23	24	25	26	46	50	52	53	54
4166	8	9	13	14	16	18	22	23	34	35	37	38	46
4205	8	9	13	14	17	18	22	23	34	35	37	38	46
4216	8	9	13	14	23	24	28	30	46	54	55	56	58
4224	8	9	13	14	17	18	22	23	46	48	49	54	55
4226	8	9	13	14	23	24	25	27	34	35	37	38	46
4246	8	9	10	11	16	18	22	23	46	7	49	53	54
4249	8	9	13	14	17	18	22	23	34	36	37	38	46
4250	8	9	13	15	19	21	22	23	34	35	37	38	46
4257	8	9	13	15	23	24	28	29	34	36	37	38	46
4260	8	9	13	14	19	20	22	23	34	35	37	38	46
4275	8	9	13	14	16	17	22	23	46	48	49	53	54
4275	8	9	13	14	19	20	22	23	34	35	37	38	46
4282	8	9	10	12	16	18	22	23	38	39	43	45	46
4283	8	9	13	14	17	18	22	23	32	33	37	38	46
4289	8	9	13	14	23	24	25	26	34	36	37	38	46
4295	8	9	10	11	16	18	22	23	31	33	37	38	46
4324	2	3	7	8	23	24	28	29	34	36	37	38	46
4361	8	9	10	12	23	24	28	30	31	33	37	38	46
4368	8	9	13	14	19	20	22	23	38	39	43	44	46
4379	8	9	10	12	23	24	28	30	46	50	52	53	54

Bu arařtırmada yer alan alt problemlere özüm getirebilmek amacıyla kullanılan ölme araçlarının hazırlanması ve verilerin elde edilmesiyle ilgili izlenen aşamalar ařağıda açıklanmaktadır.

Deneme Uygulaması İçin Gerekli Olan Madde Havuzunun Hazırlanması

Bu alıřmanın konusu olan “Türke okuduğunu anlama alanındaki Klasik Test Teorisine ve İki-parametrelili lojistik model kullanılarak bireyselleřtirilmiř test geliřtirilen iki testin KR-20 güvenirlikleri ve bir grup öęrencinin bu iki yöntemle elde edilen yetenek ölçüleri arasında manidar bir farklılık var mıdır?” ifadesine iliřkin karşılařtırmaların yapılabilmesi amacıyla öncelikle iře test

maddelerinin yazımıyla (konu alanı uzmanı iki öğretmenin katılmasıyla) başlanılmıştır. İlk aşamada 134 madde hazırlanmıştır. Hazırlanan test maddeleri ayrıca Ölçme ve Değerlendirme alanında bir uzmana kontrol ettirilmiş olup, test maddelerinin, ölçülecek özellikler için uygun olduğuna ilişkin görüş alınmıştır.

Hazırlanan maddelerin denemesi Denizli iline bağlı Tavas ilçesi, Tavas Lisesi'ndeki lise ikinci sınıf öğrencilerinden 132 kişiye uygulanmıştır. Denemeden elde edilen verilerin analizi için kullanılan bilgisayarın kapasitesi yeterli olmadığından madde sayısı 100'e indirilerek asıl uygulama için gerekli olan maddeleri belirlemek yönünde hareket edilmiştir.

Nihai uygulama için seçilen madde tipleri sınıflandırıldığında;

- cümle tamamlama
- deyimler
- sözcük anlamları olduğu görülmektedir.

Denemeden elde edilen sonuçların analizleri için BILOG bilgisayar paket programı kullanılmıştır. Analiz için yöntem olarak da, bireyselleştirilmiş test için İki-parametrelili lojistik model; en çok olabilirlik kestirimi (maximum likelihood estimation) ve Klasik Test Teorisine dayalı olarak madde istatistikleri hesaplanmıştır. Söz konusu deneme sonuçlarına ait analiz sonuçları Ek 1'de verilmiştir. Denemeden elde edilen sonuçlar faktör analizine tabi tutulmuştur. Bireyselleştirilmiş test için madde havuzunun oluşturulmasında faktör analizinden yararlanılmış; olabildiği kadar tek boyutlu bir havuz elde edilmeye çalışılmıştır.

Deneme sonuçlarının analizine göre nihai uygulama için maddelerin seçilmesinde; madde güçlük parametrelerinin geniş olması ve orta noktada

(başlangıç) bulunan maddenin sağında ve solunda yer alan madde sayısının bir birine eşit veya çok yakın olmasına özen gösterilmiştir.

Klasik Test Teorisine Dayalı Grup Testinin Hazırlanması

Deneme sonuçlarına göre yapılan analizler sonrasında, bireyselleştirilmiş test için gerekli olan maddeler kolaydan zora doğru sıraladığı ve bu sıralamanın ortasında bulunan maddenin cevaplayıcılar için başlangıç maddesi olduğu ve cevaplayıcıların bu bireyselleştirilmiş testi 13 adımda tamamladığı daha önce belirtilmişti.

Grup (kağıt-kalem) testinin hazırlanmasında izlenen yöntem ise bireyselleştirilmiş testin uygulaması sonucunda cevaplayıcılar tarafından kullanılmayan veya en az kullanılan (cevaplandırma stratejisini bağlı olarak) maddelerden 47 tanesi belirlenerek grup testi haline dönüştürülmüş, bireyselleştirilmiş testin uygulamasından yaklaşık olarak 27 gün sonra aynı gruba uygulanarak iki yöntemle ilişkin karşılaştırmaların yapılabilmesi için gerekli olan veriler elde edilmiştir.

Hem bireyselleştirilmiş testin hem de grup testinin uygulanması aşamasında cevaplayıcılara herhangi bir sür sınırlandırılması getirilmemiştir.

Verilerin Elde Edilmesi

Bu araştırmada problem ve alt problemlere cevap bulabilmek amacıyla; problem cümlesinde ifade edilen, “Türkçe okuduğunu anlama alanında Klasik Test Teorisi ve İki-parametrelili lojistik model kullanılarak geliştirilen bireyselleştirilmiş iki testin KR-20 güvenirlikleri ve bir grup öğrencinin bu iki yöntemden elde edilen yetenek ölçüleri arasında manidar bir farklılık var mıdır?” sorusun ilişkin karşılaştırmalar yapılmıştır. Bu karşılaştırmaları yapabilmek amacıyla gerekli olan ölçme araçları geliştirilmiştir.

Anılan karşılaştırmaların yapılabilmesi için her iki yönteme göre geliştirilen ölçme araçları, Denizli il merkezinde yer alan Ticaret Meslek Lisesi, lise ikinci sınıf öğrencilerin uygulanmak suretiyle gerekli olan veriler elde edilmiştir.

Verilerin Çözümlemesi

Bu araştırmanın alt problemlerine cevap bulabilmek amacıyla elde edilen verilerin çözümlemesinde Klasik Test Teorisinin madde parametreleri için madde güçlük indeksi ve madde ayırtıcılık indeksi, z istatistiği, farkları test etmek için t testi kullanılmıştır. İki-parametrelili lojistik modele göre ise en yüksek olasılık kestirim (maximum likelihood estimation) yöntemi ve madde bilgi fonksiyonu kullanılmıştır. Madde faktör analizi yapılmıştır.

BÖLÜM III

BULGULAR VE YORUMLAR

Bu bölümde araştırmanın bulguları alt problemlerdeki sıraya göre açıklanmakta ve elde edilen sonuçlar yorumlanmaktadır.

Örtük Özellikler Teorisinin İki-parametrelili lojistik modeline göre madde analizi yapıldıktan sonra 61 madde seçilmiştir. Maddelerin seçilmesinde a_j (madde ayırıcılık parametresi) değerlerinin 1,00 civarında olanlarının alınmasına ve b_j (madde güçlük parametresi) değerlerinin aralığının (ranjının) mümkün olduğu kadar geniş olmasına özen gösterilmiştir. Çalışmada kullanılan a_j ve b_j değerleri Ek 1’de görülmektedir.

Araştırmaya katılan 69 öğrenciye bu maddeler bireyselleştirilmiş test programı yardımıyla bilgisayarda uygulanmıştır. Bu 69 öğrenciye ayrıca test maddelerinden oluşturulan 47 maddelik bir test, klasik kağıt kalem testi olarak uygulanmıştır. Bireyselleştirilmiş testten elde edilen puanlar

$$X_{\theta} = 9,1x\theta + 100 \quad (3.1)$$

eşitliği yardımıyla lineer dönüşüme tabi tutulmuştur.

Bireyselleştirilmiş testten elde edilen puanların lineer dönüştürmeye tabi tutulmasının nedeni; 69 öğrencinin büyük kısmının bireyselleştirilmiş test puanlarının işaretleri eksi (-) olduğundan ve bireyselleştirilmiş test puanlarıyla kağıt-kalem testi puanları arasındaki ilişki düzeyinin (Pearson momentler çarpımı korelasyon katsayısının) hesaplanabilmesi için eksi (-) işaretinin hesaplamadaki

sıkıntısından kurtulmak amacıyla, bireyselleştirilmiş test puanları dönüştürme işlemine tabi tutulmuştur.

Bireyselleştirilmiş teste ait dönüştürülmüş puanlar ile klasik kâğıt kalem testine ait ham puanlar Tablo 3.1’de görüldüğü gibi bulunmuştur.

Tablo 3.1
Bireyselleştirilmiş Ve Kâğıt Kalem Testlerine
Ait Puanlar

Denek no	Bireyselleştirilmiş test puanı (dönüştürülmemiş)	Bireyselleştiril- miş test puanı (dönüştürülmüş)	Kâğıt kalem testi puanı (Ham puan)
1	-0.19	98.29	27.48
2	-3.01	72.58	22.67
3	-0.21	98.06	31.71
4	-0.32	97.07	32.17
5	-0.36	96.76	32.32
6	1.50	113.66	24.41
7	0.43	103.90	28.97
8	-0.62	94.32	33.46
9	-2.57	76.57	26.57
10	1.35	112.24	36.49
11	1.61	114.64	23.95
12	2.02	118.42	42.24
13	-1.24	88.69	36.10
14	-2.15	80.47	26.48
15	0.05	100.42	32.81
16	-2.40	78.20	24.72
17	0.47	104.28	34.45
18	-1.40	87.29	24.65
19	0.01	100.05	30.77
20	0.22	101.99	29.87
21	0.94	108.58	33.74
22	-1.30	88.17	30.93
23	-0.50	95.50	32.91
24	0.94	108.58	31.17
25	0.25	102.23	32.34
26	0.50	104.52	31.29
27	-0.76	93.05	27.62
28	-0.26	97.64	31.90

29	0.32	102.91	37.23
30	-0.62	94.37	33.44
31	0.08	100.69	30.48
32	-0.38	96.53	32.42
33	0.36	103.29	29.26
34	1.00	109.10	42.21
35	-0.17	98.44	31.53
36	0.48	104.34	28.77
37	0.03	100.28	30.67
38	0.70	106.40	29.40
39	1.54	114.00	36.64
40	1.22	111.12	25.59
41	-1.32	87.96	30.27
42	-1.40	87.24	29.40
43	-0.14	98.77	31.38
44	0.16	101.47	30.11
45	-0.83	92.41	34.35
46	-1.93	82.44	29.42
47	-2.17	80.22	26.68
48	0.17	101.58	30.06
49	0.13	101.18	30.25
50	1.75	115.97	34.03
51	0.37	103.36	31.27
52	-0.10	99.13	23.92
53	0.28	102.53	29.62
54	0.54	104.92	28.50
55	1.80	116.42	34.74
56	0.70	106.34	31.45
57	0.15	101.37	30.16
58	-0.81	92.65	21.87
59	1.15	110.48	30.80
60	-0.24	97.78	20.27
61	-0.22	97.99	31.74
62	-1.95	82.27	19.23
63	-0.37	96.66	27.12
64	-0.87	92.12	29.10
65	0.87	107.89	38.94
66	-0.57	94.80	30.28
67	0.53	104.82	31.42
68	2.20	120.00	31.82
69	-2.42	87.02	19.49
$\bar{X}$	-0.10	99.08	31.30
S_x	1.15	10.52	10.24

Alt Problem 1: Türkçe okuduğunu anlama alanında Klasik Test Teorisi ve İki-parametrelili lojistik model kullanılarak geliştirilen bireyselleştirilmiş testin KR-20 güvenirlikleri arasında manidar bir fark var mıdır?

Bu alt probleme cevap bulabilmek için önce araştırmada kullanılan 61 maddenin Örtük Özellikler Teorisindeki İki-parametrelili lojistik modele ait a_j ve b_j parametre değerleri,

$$r_{jx} = \frac{a_j}{\sqrt{1 + a_j^2}} \quad (3.2)$$

ve

$$p_j = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{a_j \cdot b_j}{\sqrt{1 + a_j^2}}} e^{-\frac{1}{2}z^2} \cdot dz \quad (3.3)$$

$$\frac{a_j \cdot b_j}{\sqrt{1 + a_j^2}}$$

(Thorndike, 1980, s. 33-34)

p_j : j maddesinin klasik test teorisindeki madde güçlük indeksi

r_{jx} : j maddesinin klasik test teorisindeki madde ayıricılık indeksi

a_j : j maddesinin klasik test teorisindeki madde ayıricılık gücü parametresi

dönüştürme formülleri yardımıyla klasik test teorisinin p_j madde güçlük indeksi ve r_{jx} çift serili korelasyon katsayısı olarak madde ayıricılık gücü katsayısına dönüştürülmüştür; daha sonra, bu değerlerden

$$r_j = r_{jk} \sqrt{p_j \cdot q_j} \quad (3.4)$$

r_j : Klasik test teorisindeki madde güvenirlik katsayısı
 q_j : $1 - p_j$

eşitliği yardımıyla Klasik Test Teorisindeki madde güvenirlik katsayıları hesaplanmıştır. Hesaplanan bu katsayılar Tablo 3.2'deki gibi bulunmuştur.

Tablo 3.2
 Bireyselleştirilmiş Test Uygulamasındaki Maddelerin
 Parametrelerinin Klasik Test Teorisindeki Madde
 İstatistikleri Karşılıkları

Madde no	a_j	b_j	p_j	r_{jk}
1	0.41	0.11	0.52	0.38
2	0.47	0.13	0.52	0.43
3	0.65	0.12	0.53	0.55
4	0.52	-0.20	0.54	0.46
5	0.72	-0.32	0.58	0.58
6	0.49	0.12	0.52	0.44
7	0.69	0.58	0.63	0.57
8	0.82	0.31	0.58	0.63
9	0.74	-0.13	0.53	0.60
10	1.03	0.46	0.62	0.72
11	0.93	-0.75	0.70	0.68
12	0.83	-0.27	0.57	0.64
13	0.67	1.046	0.72	0.56
14	1.12	0.25	0.57	0.74
15	0.78	1.58	0.83	0.62
16	0.48	-0.53	0.59	0.43
17	0.52	-0.68	0.62	0.46
18	0.55	1.72	0.80	0.48
19	0.66	1.24	0.75	0.55
20	0.59	0.39	0.58	0.51
21	0.47	-0.86	0.64	0.43
22	0.86	0.22	0.55	0.65
23	0.49	0.95	0.63	0.44
24	0.73	0.19	0.54	0.59
25	1.21	-0.43	0.66	0.77
26	0.94	-1.27	0.51	0.69
27	0.51	0.19	0.54	0.45
28	0.73	0.15	0.60	0.59
29	0.59	0.83	0.81	0.51

30	0.66	0.17	0.53	0.55
31	0.51	0.76	0.54	0.45
32	0.73	-0.58	0.66	0.59
33	0.82	-0.69	0.53	0.63
34	0.69	-1.27	0.61	0.57
35	0.78	0.56	0.63	0.62
36	0.57	0.79	0.67	0.50
37	0.94	0.22	0.57	0.69
38	0.93	0.20	0.76	0.68
39	0.86	-0.19	0.63	0.65
40	0.72	0.18	0.65	0.58
41	1.06	0.28	0.45	0.73
42	0.94	0.22	0.55	0.68
43	1.15	-0.16	0.43	0.76
44	1.07	0.86	0.41	0.73
45	0.41	0.99	0.57	0.38
46	0.77	0.43	0.55	0.61
47	0.92	-0.12	0.42	0.68
48	1.07	-0.78	0.72	0.73
49	0.86	0.82	0.70	0.65
50	0.53	1.68	0.78	0.47
51	0.61	0.06	0.41	0.52
52	0.47	2.10	0.81	0.42
53	1.00	0.66	0.57	0.71
54	0.79	-0.74	0.68	0.62
55	0.98	1.03	0.67	0.70
56	0.67	-0.86	0.76	0.56
57	1.09	0.27	0.68	0.74
58	0.94	0.12	0.53	0.69
59	0.81	0.09	0.52	0.06
60	1.06	0.11	0.53	0.08
61	0.80	-0.57	0.63	0.36

Bundan sonra, bireyselleştirilmiş test uygulamasında her öğrenciye sorulan soruların a_j ve b_j değerleri yerine bunların dönüştürülmüş değerleri kullanılarak klasik test teorisindeki test standart kayması

$$S_x = \sum_{k=1}^K r_j \quad (3.5)$$

eşitliği yardımıyla kestirilmiştir. Elde edilen dönüştürülmüş madde istatistikleri bireyselleştirilmiş test uygulamasında her öğrenciye uygulanan maddelerin klasik

test teorisine ait dönüştürülmüş madde istatistikleri ve test standart kayması kullanılarak her öğrenci için bireyselleştirilmiş teste ait KR-20 güvenirlik katsayısı hesaplanmıştır; daha sonra bu güvenirlik katsayılarının aritmetik ortalaması alınarak bireyselleştirilmiş testin tamamına ait KR-20 güvenirlik katsayısı elde edilmiştir. Hesaplamalar sonunda kağıt kalem testine ait KR-20 güvenirlik katsayısı 0,75 ve bireyselleştirilmiş teste ait KR-20 güvenirlik katsayısı da 0,67 bulunmuştur. Bu katsayılar (3.6) ve (3.7) eşitlikleri yardımıyla Fisher'in Z istatistiğine dönüştürülerek karşılaştırılmış; aradaki fark 0,05 düzeyinde manidar bulunmamıştır.

$$Z' = \frac{1}{2} \ln \left(\frac{1+r}{1-r} \right) \quad (3.6)$$

$$t = \frac{Z'_1 - Z'_2}{\sqrt{\frac{2}{N-3}}} \quad (3.7)$$

t = İstatistik

r = Korelasyon katsayısı

Z = r'nin standart normal değeri

N = Örneklemden eleman sayısı

Bu bulgulara göre, bireyselleştirilmiş test ile kağıt kalem testinin KR-20 güvenirliklerinin aynı olduğu; bireyselleştirilmiş test uygulamasının KR-20 anlamındaki güvenirliğine bir katkısının olmadığı söylenebilir. Ancak; bireyselleştirilmiş testte her öğrenciye uygulanan madde sayısının 13 ve kağıt kalem testinde 47 olduğu dikkate alınırsa, bireyselleştirilmiş test uygulamasında madde sayısı artırıldığında kağıt kalem testinden daha yüksek bir güvenirlik elde edileceği sonucuna ulaşılabilir.

Alt Problem 2: Türkçe okuduğunu anlama alanında bir grup öğrencinin Klasik Test Teorisi ve İki-parametrelili Lojistik model kullanılarak geliştirilen bireyselleştirilmiş teste göre elde edilen yetenek ölçüleri arasındaki ilişki düzeyi nedir?

Bu alt probleme cevap bulabilmek için dönüştürülmüş bireyselleştirilmiş test puanları ile kâğıt kalem testinden elde edilen ham puanlar arasındaki Pearson momentler çarpımı korelasyon katsayısı hesaplanmış ve bu katsayı 0,36 bulunmuştur.

Elde edilen korelasyon katsayısı düşük bir değerdir. Bunun sebebi, araştırmacının imkanlarının darlığı sebebiyle bireyselleştirilmiş testin madde havuzunun yeteri kadar geniş tutulamayışı madde sayısındaki gereğinden fazla sınırlama olabilir. Korelasyonun böyle düşük oluşunun diğer bir sebebi, bireyselleştirilmiş testin uygulamasında eldeki madde sayısının azlığı nedeniyle soru sayısı sabit tutulmuştur; bunun yerine ölçmenin hatası sabit tutulsaydı cevaplayıcıların bireyselleştirilmiş testten elde edilen puanları her cevaplayıcıda eşit miktarda ve daha az hatalı olacağı için korelasyon katsayısı daha yüksek çıkabilirdi.

BÖLÜM IV

SONUÇ VE ÖNERİLER

Bu bölümde, araştırmada elde edilen bulgulara dayalı olarak, alt problemlerle ilgili ulaşılan sonuçlar üzerinde durulmakta ve bununla birlikte bazı önerilerde bulunmaktadır.

Sonuçlar

Araştırmada, Türkçe okuduğunu anlama alanında Klasik Test Teorisi ve iki-parametrelili lojistik model kullanılarak geliştirilen iki testin bir grup öğrenciye uygulanması sonucunda her iki yöntemle göre geliştirilen iki testin KR-20 güvenirlikleri karşılaştırılması yapılmıştır. Ayrıca bir grup öğrencinin her iki testten elde edilen yetenek ölçüleri arasındaki ilişki düzeyine bakılmıştır. Karşılaştırmalarda lise ikinci sınıf düzeyi için geliştirilen sözel yetenek testi kullanılmıştır.

Araştırmanın birinci alt problemi, Türkçe okuduğunu anlama alanında Klasik Test Teorisi ve iki-parametrelili lojistik model kullanılarak geliştirilen bireyselleştirilmiş testin KR-20 güvenirlikleri arasında manidar bir farklılığın olup olmadığı yönünde incelenmiştir. Klasik Test Teorisine göre geliştirilen 47 maddelik testten elde edilen puanların KR-20 güvenirlik katsayısı 0,75 ve İki-parametrelili lojistik model kullanılarak geliştirilen bireyselleştirilmiş testten elde edilen puanlara ait KR-20 güvenirlik katsayısı 0,67 olarak hesaplanmıştır. Hesaplanan bu iki güvenirlik katsayısı Fisher'in Z istatistiğine dönüştürülerek 0,05 düzeyinde test edilmiş ve manidar düzeyde bir farklılığın olmadığı görülmüştür.

Araştırmanın ikinci alt problemine bakıldığında; Türkçe okuduğunu anlama alanında Klasik Test Teorisine ve İki-parametrelili lojistik model kullanılarak geliştirilen iki testten elde edilen bir grup öğrencinin yetenek ölçüleri arasındaki ilişki düzeyine yönelik olup; birinci alt problemde olduğu gibi her iki yöntemle ilgili olarak geliştirilen testlerin uygulanması sonucunda elde edilen yetenek ölçüleri arasında Pearson momentler çarpım korelasyon tekniği kullanılarak korelasyon katsayısı hesaplanmıştır. Korelasyon katsayısı hesaplanmadan önce uygulamaya katılan öğrencilerin bireyselleştirilmiş teste ait yetenek ölçüleri dönüştürmeye tabi tutulmuştur. Dönüştürme işleminden sonra korelasyon katsayısı 0,36 olarak bulunmuştur.

İki testten elde edilen korelasyon katsayısının oldukça düşük olduğu görülmektedir. Korelasyon katsayısının böylesine düşük düzeyde çıkmasının nedeni olarak şu gösterilebilir; bireyselleştirilmiş test oluşturulduğunda madde havuzunda yeterli sayıda sorunun olmaması bu katsayının küçük çıkmasını nedeni olabilir.

Sonuç olarak İki-parametrelili lojistik model kullanılarak geliştirilen bireyselleştirilmiş test ile Klasik Test Teorisi kullanılarak geliştirilen kağıt-kalem testinin KR-20 güvenirlik katsayıları arasında anlamlı bir farklılığın olmadığı ve uygulamaya katılan öğrencilerin ait yetenek düzeyleri arasında düşük bir ilişkinin olduğu görülmektedir.

Öneriler

Bu çalışmada iki yönlü öneri getirilmek istenmiştir. Bunlardan birincisi yukarıdaki sonuçlara bağlı olarak getirilmek istenen öneriler ikincisi ise bu konunun herhangi bir boyutuyla ilgili olarak araştırma yapmak isteyen araştırmacılara yönelik öneriler şeklinde olmuştur.

Bu araştırmanın içerdiği konuyla ilgili öneri; eğitimdeki ölçmelerin temel amaçlarından birisi bireylerin her alandaki var olan gerçek özelliklerini en iyi, tutarlı ve geçerli bir şekilde ortaya çıkarmak ve bu özelliklere bakarak bireyler hakkında karar verebilmek esas amaç edinilmiş durumdadır. Bunların gerçekleştirilebilmesinin ön koşulu bu özellikleri ortaya çıkarabilecek ölçme araçlarının oluşturulabilmesi olmakla birlikte bu özelliklerin ortaya çıkarılması aşaması ve sonrasında elde edilen ölçme sonuçlarının analizinde kullanılacak teorilerin genel özellikleri önem kazanmaktadır. Günümüzde sıkça kullanılan teorilerden birisi Klasik Test Teorisidir. İkinci teori ise son zamanlarda gittikçe önem kazanan ve kullanımı sıklaşan Örtük Özellikler Teorisi diye bilinen teori olup ve bu teorinin varsayımlarına bağlı olarak ortaya çıkan bireyselleştirilmiş testlerdir. Bireyselleştirilmiş testlerin kullanımı ise bilgisayar teknolojisine bağlı olup bilgisayar kullanımının yaygınlaşması ve eğitim kurumlarında kullanılması oranında bireyselleştirilmiş testlerin kullanım sıklığı ve alanı gittikçe artmaktadır.

Klasik Test Teorisine göre geliştirilen kağıt-kalem testlerinin problemlerinden biri olan örneklemlerden etkilenme durumu bireyselleştirilmiş testlerde söz konusu değildir. Sadece bu özellik dikkate alındığında bile bireyselleştirilmiş testler bireylerin her hangi bir alana ilişkin bir takım özelliklerinin ortaya çıkarılması, grup testlerine göre daha tutarlı sonuçlar vereceğini söylemek mümkündür. Bu çalışmada da bu güvenilirlik katsayısı anlamında ortaya çıkmıştır. Bireyselleştirilmiş testlerin bu olumlu niteliklerinin yanında bazı problemlerin de olduğunu kabul etmek gerekir. Bu problemlerden birisi madde havuzunun oluşturulması aşamasında a_g ve b_g gibi madde parametrelerinin gerektirdiği maddeleri oluşturmaktır. Uygun özellikte a_g ve b_g gibi kararlı madde parametre değerlerini içeren maddelerin geliştirilmesi durumunda amaca yönelik olarak çok kullanışlı olacağı söylenebilir.

Bu arařtırmada, Trke okuduęunu anlama alanında Klasik Test Teorisi ve İki-parametrelili lojistik model kullanılarak bireyselleřtirilmiř test řeklinde geliřtirilen iki testin KR-20 gvenirlikleri ve bir grup oęrencinin bu iki yntemle elde edilen yetenek lleri arasında manidar bir farklılık var mıdır? Sorusu zerinde durulmuřtur. Bu konuda arařtırma yapmak isteyen arařtırmacılar, bu arařtırmada yer alan konunun bir bařka boyutu olarak, Klasik Test teorisine baęlı olarak geliřtirilecek kaęıt-kalem testleri ile iki parametrelili ve -parametrelili lojistik modele gre geliřtirilecek testletlerin gvenirlik ve geerlilikleri arasında farkın olup olmadıęını veya bireylerin her hangi bir alana ait bařarılarının llmesinde her iki yntemi gvenirlik ve geerlilik zelliklerini karřılařtırılabileceęi bir alıřma yapılabilir. Ve yine iki-parametrelili lojistik model ile -parametrelili lojistik modele gre geliřtirilen testletlerin lmeye alıřtıkları zellikleri dikkate alındıęında psikometrik zellikleri itibarıyla bir farklılıęın olup olmadıęı arařtırılabilir.

İkinci neri bu konuda arařtırma yapacak olan arařtırmacılaraya ynelik neridir. Arařtırmacıların elde edecekleri verilerin zmlenebilmesi iin gerekli olan bilgisayar zel yazılım programlarını veya hazır paket programlarını temin edip edememe konusunu dřnmeleri gerekmektedir. Bu konu lkemiz iin ok yeni bir konu olduęundan sz konusu bilgisayar programlarının temininde bazı sıkıntıların ortaya ıkması muhtemeldir. Belki belli bir sre sonrasında byle bir sıkıntı olmayacaktır. O zaman bu alanda mutlaka arařtırmaların yapılması gerekir.

K A Y N A K Ç A

BAYKUL, Yaşar.

- 1979 "Örtük Özellikler ve Klasik Test Kuramları
Üzerine Bir Karşılaştırma"
(Yayımlanmamış Doktora Tezi)
Ankara: Hacettepe Üniversitesi. 1979
- 1989 **Test Geliştirme.** (Çoğaltma:Ders Notu)
- 1997 **İstatistik: Metodlar ve Uygulamalar.** Anı Yayınları. Ankara.

CREHAN, Kevin D; ve diğerleri.

- 1993 A Comprasion of Testlet Reliability for Polytomous Scoring
Methods. **Paper Presented at the Annual Meeting of the
American Educational Research Associaton.** Atlanta, GA,
April 12-16

CROCER, L., ALGINA, J. **Introduction To Classical And Modern Test**

- 1986 **Theory.** CBS College Publishing. 383 Madison Avenue New
York. N.Y. 10017

ERKTİN, Emine.

- 1994 Ölçmede Yeni Yaklaşımlar: Örtük Özellikler Kuramı.
Türk Psikoloji Dergisi. Cilt 9, Sayı.32,s.51-60.(Özel Sayı).
Ankara.

HAMBLETON, Ronald K.,SWAMINATHAN, Hariharan.

- 1985 **Item Response Theory: Principles and Applications.**
Kluwer-Nijhoff Publishing. Boston

KAPTAN, Fitnat.

- 1994 "Yetenek Kestiriminde Adaptive (Bireyselleştirilmiş) Test
Uygulaması İle Geleneksel Kağıt-Kalem Testi Uygulamasının
Karşılaştırılması"
(Yayımlanmamış Doktora Tezi)
Ankara: Hacettepe Üniversitesi. 1994

KARASAR, Niyazi.

- 1995 **Bilimsel Araştırma Yöntemi.**
3A Araştırma Eğitim Danışmanlık Ltd. Ankara

- 1995 **Araştırmalarda Rapor Hazırlama.**
3A Araştırma Eğitim Danışmanlık Ltd. Ankara

LINN, Robert L. GRONLUND, Norman E.

- 1995 **Measurement and Assessment in Teaching.**
Prentice-Hall, Inc. 7'th Edition. N.J.

LORD, Frederich. M. and NOVICK M.R.

- 1968 **Statistical Theories of Mental Test Scores.**
Reading Mass: Addison-Wesly Publishing Company.

LORD, Frederich. M.

- 1980 **Applications of Item Response Theory to Practical Testing Problems.** Hillsdale, NJ: Erlbaum.

ÖNCÜ, Hüseyin

- 1995 **Eğitimde Ölçme Ve Değerlendirme.** TDFO (Tasarım-Dizgi-Fotokopi-Ofset-Bilgisayar yayıncılık San.Tic.Ltd.Şti). Ankara

Scientific Software, Inc. **DOS-BILOG V3.04.** 1369 Neitzel Road
Mooreville, Indiana 46158.

TEKİN, Halil,

- 1984 **Eğitimde Ölçme Ve Değerlendirme.**
Yargı Yayınları. Yedinci Baskı. Ankara.

TEKİNDAL, Satılmış.

- 1992 "Madde Çeşidinin Test Güvenirliğine Ve Geçerliğine Etkisi" (Yayımlanmamış Doktora Tezi)
Ankara : Hacettepe Üniversitesi.

THISSEN, D., Steinberg L. ve J. Ann Moony.

- 1989 Trace Lines of Testlets: A Use of Multiple-Categorical-Response Models. **Journal of Educational Measurement.**
Fall 1989, Vol. 26, No. 3, pp, 247-260

THISSEN,D.,Mislevy Robert J.

- 1990 **Testing Algorithms. Computerized Adaptive Testing. A Primer.** Lawrance Erlbaum Associates, Publisher Hillsdale, New Jersy.

THORNDIKE, Robert, L.

- 1982 **Applied Pyschometrics.** Houghton Mifflin Company. Boston.

TURGUT, M. Fuat.

- 1977 **Ölçmede Hata ve Güvenirlik.**
(Çoğaltma)
Ankara: H.Ü. eğitim Bilimleri
- 1992 **Eğitimde Ölçme ve Değerlendirme Teknikleri.**
Ankara: Saydam Matbaacılık.

TURGUT, M.Fuat., BAYKUL, Y.

- 1992 **Ölçekleme Teknikleri.** ÖSYM Yayınları.Yayın No 1. Ankara

WAINER,Howard.,Sireci, Stephen G., David Thissen

- 1986 **On The Reability of Testlet Based Tests.**
Journal of Educational Measurement. Vol.28, No.3,
pp.237-247
- 1991 **Differential Testlet Functioning: Definations and Detection.**
Journal of Educational Measurement.
Fall1991,Vol.28,No.3, pp,197-219

WAINER, Howard: ve diğerleri.

- 1991 **Some Empirical Guidelines for Building Testlets.**
Program Statistics Research Technical Report No: 91-14.
Educational Testing Service, Princeton,N.J. October.

WAINER, Howard. Kiely, Gerard. L,

- 1987 **CAT's Testlets and Test Constraction: A Rationale for Putting Test Developers Back Into CAT. Program Statistics Research Project No: 86-33.** Educational Testing Service, Princeton,NJ.

WAINER, Howard.,Kiely, Gerard,L.

- 1987 Item Clusters and Computerized Adaptive Testing:
A Case for Testlets. **Journal of Educational Measurement.**
1987,24.185-201.

WAINER, Howard., C.Lewis.

- 1990 Toward A Psychometrics for Testlets.
Journal of Educational Measurement.1990,27.1-14.

WAINER,H., Lewis,C., Kaplan,B.,J. Braswell.

- 1991 Building Algebra Testlets: A Comprasion of Hierarchial And
Linear Structures. **Journal of Educational Measurement.**
1991,28.No:4.pp.311-323.

WEISS, David. J. **New horizons In Testing:** Latent Trait test Theory and

- 1983 Computerized Adaptive Testing. Academic Press, Inc.
New York

ZARA, R. Anthony. Administering and Scoring Adaptive Test.

- 1997 **National Council of Boards of Nursing,Inc.**March. 1997

Ek 1: Deneme Sonuçlarına Ait İstatistikler

Madde No	a_g	b_g	Ki-Kare
1	0.054	0.141	0.3074
2	0.411	0.109	0.6452
3	0.473	0.128	0.3698
4	0.274	0.122	0.7935
5	0.275	0.146	0.6119
6	0.313	0.124	0.9722
7	0.371	0.124	0.4648
8	0.218	0.143	0.9944
9	0.293	0.105	0.5966
10	0.256	0.13	0.8212
11	0.256	0.139	0.7847
12	0.256	0.121	0.549
13	0.218	0.104	0.3562
14	0.293	0.111	0.4237
15	0.294	0.122	0.6588
16	0.353	0.171	0.6353
17	0.313	0.123	0.9922
18	0.275	0.135	0.9631
19	0.182	0.146	0.7314
20	0.293	0.133	0.2597
21	0.145	0.145	0.637
22	0.332	0.125	0.9249
23	0.108	0.122	0.739
24	0.018	0.123	0.622
25	0.181	0.137	0.796
26	0.108	0.141	0.9237
27	0.313	0.127	0.6669
28	0.036	0.174	0.2616
29	0.126	0.108	0.0071
30	0.351	0.119	0.3053
31	0.255	0.119	0.2807
32	0.313	0.141	0.7616
33	0.339	0.245	0.3938
34	0.392	0.135	0.6495
35	0.353	0.15	0.3299
36	0.255	0.117	0.7596
37	0.145	0.125	0.2402
38	0.255	0.109	0.0376
39	0.09	0.149	0.7492
40	0.182	0.159	0.3624
41	0.332	0.123	0.6323
42	0.072	0.18	0.7851
43	0.182	0.149	0.5797
44	0.275	0.146	0.8270
45	0.218	0.135	0.7062
46	0.278	0.206	0.0822
47	0.432	0.139	0.6395

EK 1: Devamı

Madde No	a_g	b_g	Ki-Kare
48	0.372	0.130	0.3070
49	0.054	0.144	0.4161
50	0.256	0.124	0.9239
51	0.181	0.124	0.2970
52	0.298	0.221	0.4241
53	0.373	0.164	0.5636
54	0.610	0.187	0.4396
55	0.392	0.136	0.7273
56	0.352	0.134	0.9593
57	0.335	0.185	0.7799
58	0.474	0.151	0.8361
59	0.653	0.154	0.7114
60	0.475	0.168	0.4690
61	0.237	0.112	0.3772
62	0.520	0.185	0.6702
63	0.494	0.116	0.5910
64	0.473	0.127	0.5856
65	0.495	0.139	0.4732
66	0.497	0.166	0.6734
67	0.662	0.220	0.6913
68	0.478	0.195	0.8453
69	0.519	0.186	0.8528
70	0.564	0.183	0.6570
71	0.302	0.278	0.1442
72	0.565	0.215	0.0955
73	0.539	0.159	0.9984
74	0.696	0.269	0.2590
75	0.522	0.362	0.1125
76	0.163	0.143	0.1930
77	0.236	0.105	0.2180
78	0.378	0.233	0.1893
79	0.273	0.126	0.1327
80	0.219	0.149	0.9963
81	0.313	0.118	0.2810
82	0.274	0.114	0.1621
83	0.352	0.150	0.5550
84	0.218	0.121	0.3051
85	0.393	0.151	0.9754
86	0.293	0.103	0.3964
87	0.433	0.159	0.9962
88	0.411	0.120	0.9651
89	0.293	0.118	0.1268
90	0.332	0.114	0.4556
91	0.274	0.117	0.1456
92	0.351	0.118	0.6268
93	0.182	0.177	0.6029
94	0.391	0.137	0.1950

EK 1: Devamı

Madde No	a_g	b_g	Ki-Kare
95	0.473	0.115	0.0884
96	0.294	0.129	0.8973
97	0.351	0.110	0.2772
98	0.370	0.091	0.1470
99	0.333	0.153	0.3308
100	0.236	0.112	0.8541



1 - Arkadaşın ne denli iyi olursa olsun, ona yine de fazla güvenme. Çünkü

Bu öğütlemeye boş bırakılan yere anlamca en uygun hangi yargı getirilmelidir?

- A) Bir ölçüp bir biçmek gerekir
- B) Onunla içli dışlı olmamak gerekir
- C) Ay ışığında ceviz silkilmez
- D) İnsanoğlu çiğ süt emmiş
- E) Bal alacak çiçeği bilmelisin

2- Yapılacak bir işin bütün yönleri önceden iyi düşünülmeli, sonucu iyi hesaplanmalı, ondan sonra işe girilmelidir; yani

Bu tümcede anlamın akışına göre en uygun hangi atasözü getirilmelidir?

- A) Bin bilsen de bir bilene danışmalı
- B) Çürük tahta çivi tutmaz
- C) Kanatsız kuş uçmaz
- D) Kişi ne yaparsa kendine yapar
- E) Bin ölçüp bir biçmeli

3- " Hiç bir destek sürekli olmaz. bunun için insan yapacağı işte başkalarının desteğine pek güvenmemeli, yalnız kendi öz gücüne inanmalıdır" **açıklamasına anlamca denk düşen atasözü, aşağıdakilerden hangisidir?**

- A) Ağaç kökünden yıkılır
- B) Ağacın kurdu içinde olur
- C) Ağaca dayanma kurur, adama dayanma ölür
- D) Ağacı kurt, insanı dert öldürür
- E) Ağaçtan maşa olmaz

4 - "Yararlanacağı bir iş peşinde koşan kişi, yararlanmanın gerektirdiği araçları hazır bulundurmaz" **açıklamasını karşılamayan atasözü aşağıdakilerden hangisidir?**

- A) Arı bal alacak çiçeği bilir
 - B) Camı kaymak isteyen, mandayı yanında taşır
 - C) Zemheride yoğurt isteyen cebinde inek taşır
 - D) Aşure yemeye giden kaşığı taşır
 - E) Pilav yiyen, kaşığı yanında taşır
- 5) **Aşağıdakilerin hangisinde deyim yoktur?**

- A) Bu günlerde işi ağırdan almaya başladım
- B) Çobanlar, suyun gözünde buluştu
- C) Ondan söz ederken hiç bozuntuya vermedi
- D) Sonunda o da boyunun ölçüsünü aldı
- E) Öğrenciler, sınıfta hiç çıt çıkarmıyorlardı

6 - **Aşağıdaki tümcelerin hangisinde bir atasözü açıklamasıyla birlikte verilmemiştir?**

- A) Hep çıkar, peşinde koşan kişi tehlikelerden uzak kalmaz; gözü tanede olan kuşun ayağı tuzaktan kurtulmaz
- B) Her kesin bildiği gerçekler ört bas edilemez; güneş balçıkla sıvanmaz
- C) çok sinirlenmek başımızı derde sokabilir; çünkü öfke baldan tatlıdır
- D) Mert olan kişi, sözünde durur. verdiği sözü yerine getirir; yani söz ağızdan çıkar.

E) İş yapacak alanda yeteri kadar. güç bulunmadıkça, bu iş başkalarının küçük yardımıyla yürümez; kısacası taşıma su ile değirmen dönmez.

7 - "Her bir tel çekseler, köseye sakal olur" **atasözünün en uygun açıklaması aşağıdakilerden hangisidir ?**

- A) Çok, azdan oluşur
- B) Domuzdan bir kıl kopsa da domuza bir şey olmaz
- C) Damlaya damlaya göl olur
- D) Herkes biraz yardım etse bir yoksul, perişanlıktan kurtulur
- E) İki gönül bir olunca samanlık seyrân olur

8 - **Aşağıdakilerden hangisinde bir deyim açıklamasıyla birlikte verilmemiştir ?**

- A) verimsiz işlerle uğraşır; kalburla su taşıyorsun
- B) O, görünüşünden beklenen işi yapamadı; kalıbının adamı değilmiş
- C) Ona karşı içimde bir sevgi yeşerdi; kanım kaynadı
- D) Sınavı kazandığından tatilin keyfini çıkardı
- E) Yenilikleri kabullenmeyen, eskiye bağlanıp kalmış, yani örümcek kafalı birisiydi

9 - "Canla başla çalışarak onunla uyumlu bir ortaklık kurduk ..."

Bu tümceyi "sonunda tüm ağır işleri bırakarak beni terk etti" sözleriyle sürdürmek istersek, bu durumu en iyi aşağıdaki deyimlerden hangisi karşılar?

- A) Üstüne yıkmak B) Üstüne yatmak
C) Başına ekşimək D) Üstüne almak
E) Başının etini yemek

10 - Herkesin talihi, kazancı bir değildir aynı çabayı göstermelerine karşın kimisi daha çok, kimisi daha az kazanır; kısacası ...

Yukarıdaki açıklamada boş bırakılan yere anlamın bütünlüğü açısından en uygun hangi atasözü getirilmelidir?

- A) Her kaşığın kısmeti bir olmaz
B) Herkes evinde ağıdır
C) Herkesinaklı bir olsa, koyuna çoban bulunmaz
D) Herkes kaşık yapar ama sapını ortaya getiremez
E) Herkesin tenceresi kapalı kaynar

11 - Aşağıdakilerin hangisinde bir deyim, açıklamasıyla birlikte verilmemiştir?

- A) İmrendiğim, istediğim şeyi elde edemedim; hevesim kursağımda kaldı
B) Çok sinirlendi bağırdı, çağırdı; heyheyleri tuttu
C) Onunla karşılaşacağı için iki dirhem bir çekirdek gibiydi sanki
D) Çok geçim sıkıntısı çekiyor: iki yakası bir araya gelmiyordu
E) çok güzel ve etkili yazılar yazıyordu; adeta kaleminden kan damlıyordu

12 - Aşağıdakilerin hangisinde bir deyim yanlış anlamında kullanılmıştır?

- A) Onun seni görmek istemediğine kalıbımı basarım
B) Gömleklerin fiyatları farklıydı; ama gömlek arasında uzun uzadıya bir fark yoktu
C) Dedem koltuğun üzerinde uyku kestiriyordu
D) Çocuğunu çok seviyor; üzerine titriyordu
E) Gece gündüz bu işin üstüne düşeceksin

13 - Hiç bitmeyecek gibi görünen o işi **yaradana sığınıp** zar zor bitirdi

Bu tümcede altı çizili deyimın yerine anlam gereği en uygun hangi söz getirilmelidir?

- A) Dua ederek B) İnancına sarılarak
C) Düşe kalka D) Kılı kırk yarararak
E) Bütün gücünü kullanarak

- 14 - I. Herkes ne ederse kendine eder
II. Herkes kendi ölüsü için ağlar
III. Herkes kendi çukurunu doldurur
IV. Herkesin ettiği yoluna gelir
V. Herkes bildiğini okur

Yukarıdaki atasözlerinin hangileri "iyilik eden iyilik bulur" anlamını karşılar?

- A) I. ile III. B) II. ile IV. C) IV. ile V.
D) I. ile IV. E) III. ile IV.

- 15 - I. Ağaç yaprağıyla gürler
II. Yalnız taş duvar olmaz
III. Bir elin nesi var; iki elin sesi var
IV. Taşıma su ile değirmen dönmez
V. Yalnız kalamı kurt yer

Yukarıdaki atasözlerinin hangisi, konuya bakış açısı yönünden ötekilerden farklıdır?

- A) I. B) II. C) III.
D) IV. E) V.

- 16 - I. Aş taşınca kepçeye paha olmaz
 II. Aşı pişiren yağ olur, gelinin yüzü ağ olur
 III. Aşını, eşini, işini bil
 IV. Kötü söyleme eşine, zehir katar aşına
 V. Alet işler el öğünür

Yukarıdaki atasözlerinin hangi ikisi anlamca birbirine yakındır?

- A) I. ile V. B) II. ile V. C) I. ile III.
 D) III. ile IV. E) IV. ile V.

- 17 - I. Ne yılandır o, her kötülüğü yapar
 II. Orada tatlı bir kızla tanıştım
 III. "Koçum benim !" haydi at golünü
 IV. Acı soğuklar bir türlü geçmedi
 V. O dar odada at gibi koşturuyorlardı

Yukarıdaki tümcelerde altı çizili hangi sözcük temel (gerçek) anlamıyla kullanılmıştır?

- A) V. B) IV. C) III. D) II.
 E) I.

18 - " Davranışları birbirine ters düşen iki tarafın her birini kendi tutumunda yürüklendirmek " anlamına denk düşen deyim aşağıdakilerden hangisidir?

- A) Acem kılıcı her iki tarafı kesmek
 B) Tavşana kaç, taziya tut demek
 C) Ters yüz etmek
 D) Kargayı bülbül diye satmak
 E) Yaraya tuz biber ekmek

19 - Onu büyük bir dikkatle, iyice kavrayarak dinle ki görenler, "....." desinler.

Bu tümcede boş bırakılan yere anlamca en uygun hangi deyim getirilmelidir?

- A) Can korkusuyla dinliyor
 B) Can havliyle dinliyor
 C) Can kulağıyla dinliyor
 D) Kulak kabartarak dinliyor
 E) Kulak misafiri olarak dinliyor

20 - "**Pireyi deve yapmak**" deyiminin anlamı aşağıdakilerin hangisinde geçmektedir?

- A) Sen çok önemli bir olayı her zaman görmezlikten gelirsın
 B) Sen çok önemsiz bir olayı herkesten farklı yorumlarsın
 C) Sen önemli önemsiz bir olayı büyük bir heyecan içinde anlatırsın
 D) Sen çok önemsiz bir olayı büyütürük aktarırsın
 E) Sen, bazılarının önem vermediği bir olayı çok önemli bulursun

21 - I. Sınavı kazanamayınca aklı başına geldi
 II. İlacı içince yavaş yavaş aklı başına geldi

III. Bir gün gözlerin bozulacak, aklım başına gelecek

IV. Kurallara uymuyordu; yiyince aklı başına geldi

"Aklı başına gelmek" sözü yukarıdaki tümcelerin hangisinde ötekilerden farklı anlamda kullanılmıştır?

- A) I. B) II. C) III. D) IV. E) V.

22 - Karşı takıma gol attıklarında o kadar sevinmişti ki sanki

Bu tümcede boş bırakılan yere, anlamın bütünlüğü açısından en uygun hangi deyim getirilmelidir?

- A) etekleri zil çalıyordu
 B) eli ayağı dolaşıyordu
 C) etekleri tutuşmuştu
 D) yüreği hoplamıştı
 E) gözü parlamıştı

23 - Sınavı kazandığımı öğrenince.....

Bu tümce anlamca en uygun aşağıdaki hangi deyimle tamamlanabilir?

- A) yeniden dünyaya gelmiş gibi oldum
- B) yüreğim ağzıma gelmişti sanki
- C) çok memnun olmuştum
- D) dünyalar benim olmuştum sanki
- E) aklım başımdan gitti

24 - "**Dürbünün tersiyle bakmak**" deyiminin anlamı aşağıdakilerin hangisinde vardır?

- A) Görünüş, insanı yanıltabilir
- B) O, her şeye iyimser gözüyle bakar
- C) Nedense sen herşeyi olduğundan daha küçük görürsün
- D) Arabamızla ters yola girmiştik
- E) Dünya yıkılsa umrunda değildir onun

25 - "Ne yap yap, kendini ona sevdirmeye bak"

Açıklamasını özetleyen aşağıdakilerden hangisidir?

- A) Gözünü dikmek
- B) Gözünü ayırmamak
- C) Gözünün içine bakmak
- D) Gözünü açmak
- E) Gözüne girmek

26 - Zorluklara alışmış bir insan, zorluklardan pek etkilenmez; yani

Yukarıdaki tümcenin anlam bütünlüğü en uygun aşağıdakilerin hangisiyle tamamlanır?

- A) Acı acıyı keser
- B) Acı patlıcanı kırağı çalmaz
- C) Açık yaraya tuz ekilmez
- D) Güneş balçıkla sıvanmaz
- E) Hatasız kul olmaz

27 - "**Damlaya damlaya göl olur**" atasözüyle doğrudan ilgili olan söz, aşağıdakilerden hangisidir?

- A) Dayanışma
- B) Deneyim
- C) Birikim
- D) Yararlanma
- E) Başarı

28 - "**Damar**" sözcüğü aşağıdakilerin hangisinde başka bir sözcükle deyim oluşturmamıştır?

- A) Başkalarının damarına basmaktan zevk alırdı
- B) Onun damarına girdi, isteklerini ona kabul ettirdi
- C) Çocuğun yine damarı tuttu, annesini üzdü
- D) Ne yapalım, küsmek onun damarına işlemiş
- E) Hastanın beyin damarları çatlamış

29 - Ad aktarması; bir adın, benzetme ilgisi olmaksızın kendisiyle ilgili bir kavramın yerine kullanılmasıdır.

Aşağıdaki altı çizili sözcüklerin hangisi böyle bir aktarmaya örnek oluşturmaz?

- A) Ben gelene dek masanı toplamalısın
- B) Okulumuz, yarışmada birinci oldu
- C) Bütün köy, kente göçtü
- D) Sokağın dirseğinde kaza yapmış
- E) Orhan Kemal'i okudunuz mu hiç?

30 - Bir tek sözcükle anlatılabilecek bir kavramı birden çok sözcükle anlatmaya "**dolaylama**" denir.

Aşağıdakilerin hangisinde dolaylama yoktur?

- A) Büyük kurtarıcı bağımsızlık ataşını Samsun'da yaktı
- B) Değerli bir yazarımız daha aramızdan ayrıldı
- C) Beyaz perdenin ünlüleri bir araya geldi
- D) Aktörün rolünü çok beyendim
- E) Kara elmes Zonguldak'ın yaşam kaynağıdır

31 - "On sekizinde var ya da yoktu; genç yaşta anne olmuştu" **tümcesinde** "On sekizinde" sözü "On sekiz yaşında" anlamında kullanılmıştır.

Aşağıdaki tümcelerin hangisinde buna benzer bir kullanım yoktur?

- A) Ahmet'ler altta oturuyorlar
- B) Dikili bir ağacı bile yoktu
- C) Dünyayı buharıyla dolaştı
- D) Otobüse ön kapıdan binip arkadan ininiz
- E) Ticaret beşe alıp ona satmaktır, diyordu

32 - "Yelkenli hızla uzaklaşıyordu sahilden" **tümcesinde** "yelkenli" sözüyle "yelkenli gemi" anlatılmak istenmiş ve böylece bir "düşük anlatım" yapılmıştır.

Aşağıdakilerin hangisinde böyle bir anlatıma baş vurulmamıştır?

- A) Bu akşam balığa gideceğiz
- B) Bu yıl, yavru vatanda bir hafta kaldık
- C) Elinde testi ile suya giden gelin
- D) Kuyruklu, gökyüzünde ne güzel parlıyordu
- E) Onlar, her yaz Söke ovasına pamuğa giderler

33 - Aşağıdaki tümcelerin hangisinde "fazla" sözcüğü tümceye "üstünlük" anlamı katmıştır?

- A) Bu gün de fazla çalıştım
- B) Sana iki araba fazladır
- C) Onun zekası senden fazladır
- D) Fazla paramı nereye koyuyorsun?
- E) İşime fazla karışıyorsun

34 - Aşağıdakilerin hangisinde "gizli" sözcüğü ötekilerden farklı anlamdadır?

- A) Ceketinin cebime gizli bir cep yaptırdı
- B) Arabasında gizli bir alarm düğmesi vardı
- C) Antik kentin altında gizli bir yol bulunuyordu
- D) Gizli bir güç beni ondan soğuttu
- E) Sakladıklarını mahsenin gizli bölümünde bulduk

35 - "Ne" sözcüğü, aşağıdakilerin hangisinde "ne güzel gözlerin var öyle !" tümcesindeki anlamında kullanılmıştır?

- A) Bana, öyle ne diye bakıyorsun?
- B) O kadar ekmeği ne yapacaksın?
- C) Gözlerin ne söylüyor, bir bilsem !
- D) Ne inatçı insandır o, biliyorsun değil mi?
- E) Ne saçlarına bakar, ne de güzel giyinir

36 - Aşağıdakilerin hangisinde "nasıl" sözcüğü, "zorunluluk" anlamında kullanılmıştır?

- A) yemeğini yesin, bak nasıl çalışacak
- B) Bunu bana nasıl yaparsın?
- C) Ne anlatsam, bilmem ki !
- D) Onu doktora nasıl götürmezsin?
- E) yaptığın iyilikleri nasıl unuturum?

37 - Kitapları karıştırırken bakarım, dün içinde türlü güzellikler bulduğum, okudukça coştığım bir yer, bu gün birşey demez olmuş bana. O sayfalar boşalmış artık benim için. Kendi yazılarımda bile her zaman ilk duyduğum, düşündüğüm şeyleri bulamam.

Bu parçanın anlatmak istediği ana düşünce aşağıdakilerden hangisidir?

- A) Güzellik kavramı kişiden kişiye değişiklik gösterir
- B) Sanat yapıtlarını yorumlamada öznel yaklaşım kişilik gösterir
- C) Geçen zaman, beğeni ve sanat anlayışını etkiler
- D) Yaratıcılıktan ve güzellikten yoksun eserler kalıcı olmaz
- E) kitaplar, ancak sürekli okunursa güzelliklerini koruyabilir

38 - Günlerdir işsizdim, iş için nereye başvursam kapılar yüzüme kapanıyordu. Artık son umut olarak küçük bir işletmesi olan baba dostuna gittim. Biraz umutsuz, biraz sıkılgan bir tavırla kendisinden iş istedim. Bana hemen işe başlayabileceğimi söylediğinde **dünyalar benim olmuştu**.

Aşağıdakilerden tümcelerden hangisi, bu parçanın tümcesinin (altı çizili) yerine kullanılamaz?

- A) ne kadar sevindim, bilimezsiz.

- B) sevinçten uçacak gibi oldum
 C) nasıl sevindiğimi anlatamam
 D) ne kadar sevindiğimi söylemeye gerek var mı?
 E) bir kat daha artmıştı sevincim

39 - "Bundan sonra şiir söylemeyeceğim. Ben bir şeyi yaşatlarımdan daha iyi yapabildiğime inanmazsam iç rahatlığı taşımam. Benden çok daha iyi şair Ahmet Haşim'dir. Onun için şiiri ona bırakıyordum. Sözün kısası"

Bu paragraf, dil ve düşünce kurgusu yönünden aşağıdakilerin hangisiyle sürdürülebilir?

- A) Söz bilen, kılıç kuşananın olmalıdır.
 B) bu işi yalnızca ünlüler yapmalıdır
 C) bence Ahmet Haşim iyi bir şairdir
 D) ben şairlik yeteneğimi yitirmiş bulunuyorum
 E) yaşitlarıma şans tanımak da isterdim.

40 - Bir konu üzerinde söz söyleyen çok olursa sonuca varmak gecikir; demem şu ki

Bu tümcede boş bırakılan yere anlamın bütünlüğü açısından aşağıdaki atasözlerinden hangisi getirilmelidir?

- A) El kazam ile aş kaynamaz
 B) akıllı düşman akılsız dosttan akıllıdır
 C) horozu çok olan köyün sabahı geç olur
 D) bir elin nesi var, iki elin sesi var
 E) her kuşun eti yenmez

41 - Aşağıdaki tümcelerin hangisinde istek söz konusu **değildir**?

- A) Bu ödevi haftaya hazırlayıp geleceksin
 B) Çiçekleri gün aşırı sularsın, onların canlandığını göreceksin
 C) Sınavdan çıkar çıkmaz eve geleceksin
 D) Bana iki güne bir telefonla bilgi vereceksin
 E) Büroya gelip gidenlere nazik davranacaksın

42 - "Nasıl olsa yine bir gün
 Döneriz bu yollardan geri"

Aşağıdakilerden hangisi, bu dizelerdeki "nasıl olsa" sözünün anlamına en uzaktır?

- A) Mutlaka
 B) Eninde sonunda
 C) Elbet
 D) Ne olursa olsun
 E) Belki

43 - " Onun geleceğini sanıyordum doğrusu."

Bu tümcede altı çizili sözün bildirdiği anlam, aşağıdakilerin hangisinde vardır?

- A) Duyduğuma göre, o bizimle olmayacakmış
 B) Onun doğruyu söylediğine inanmıyorum
 C) Onun doğruyu söylediğine inanıyorum
 D) Gerçeği söylemek gerekirse onu hiç görmedim
 E) Babanız şu anda belki yola çıkmıştır

44 - " Ne romancı, ne şair olmak elimizdedir. İnsan ya romancı doğmuş ya doğmamıştır. Şair için de öbür sanatçılar için de böyle. Sanatçı, içinden geçen yolun yolcusudur."

Yukarıdaki parçanın ana düşüncesi aşağıdakilerden hangisidir?

- A) Her sanatçı sanat yolunu kendi seçer
 B) Sanatçıyı yaratan, toplumsal koşullardır
 C) Sanat, iç dünyamızın anlatımıdır
 D) Sanatçı olmanın yolu yaratılıştan geçer
 E) Sanatçı olmanın sırrı, zengin bilgi birikiminde olmalıdır

45 - Orhan Veli, yoksul insanların haline dokunur. Küçük insanların dünyasını, beğenilerini yansıtmaya çalışır. Gündelik yaşamı, iç güdülere dayalı küçük şeylerle avunan bir yaşayış biçimini şiirleştirir.

Pu paragrafın konusu aşağıdakilerden hangisidir?

- A) Gündelik yaşamın anlatımı
- B) Orhan Veli'nin şiir anlayışı
- C) Sıradan insanların mutluluğu
- D) Orhan Veli'nin şiirlerinin konusunu
- E) Yaşama sevincinin şiire yansması

46 - Pencerisinin önüne oturmuş, karşıda yükselen dağlara bakıyordu. Kekik ve sümbül kokularını duyuverdi birden. Güneş, morlu ve kıvrıllı renge bürünmüş dağların tepesinde bir noktadaydı. Güzelliklerle dolu bir gün daha yitiyordu. Şairin o ünlü şiirini düşündü: "Ne iyiolurdu penceremden gün eksilmese !" diye içini çekti.

Bu parçanın anlatımında aşağıdakilerden hangisi yoktur?

- A) Duygulanırma öğelerinden yararlanma
- B) Betimleyici bir anlatım kullanma
- C) Öyküleyici bir anlatım biçimine başvurma
- D) Nesnel yargılarla anlatımı geliştirme
- E) Çağrışımlarla, anlatıma güzellik katma

47 - Sinemaya bir ömür verdi. Güzel filmler yaptı; bir de ödül aldı.

Bu tümcede boş bırakılan yere aşağıdaki sözcüklerden hangisi getirilmez?

- A) Hatta B) üstelik C) ayrıca
- D) bundan böyle E) bununla birlikte

48 - Aşağıdakilerin hangisinde, "özgü" sözcüğü, "sadece bir varlıkta, bir şeyde bulunup başkalarında bulunmayan" anlamının dışında kullanılmıştır?

- A) Ozanlara özgü bir duyarlılık var onda
- B) Bu tür organizma bozuklukları yaşlılara özgüdür
- C) Düşünmek, kuşkusuz insanlara özgüdür
- D) Üzerinde zeybeklere özgü bir giysi vardı
- E) Bu yüzme havuzu sadece çocuklara özgüdür

49 - "Büyük" sözcüğü, aşağıdaki tümcelerin hangisinde sayısal anlam ıçermez?

- A) Başarı için büyük düşünmek gerek
- B) Ağacın gölgesi oldukça büyüktü
- C) Evimizin önünü büyük binalar kapatmıştı
- D) Van gölü, büyük göllerimizdendir
- E) Onun boyu benimkinden büyüktü

50 - "Vurmak" sözcüğü, aşağıdakilerin hangisinde "sapmak" anlamında kullanılmıştır?

- A) Yağmur camlara vuruyordu
- B) Boyayı duvara iyi vurmamışlar
- C) Şişeyi yirmi metreden vuramadı
- D) Bu sabah balık karaya vurdu
- E) O sahtekar, yine birinin beş milyonunu vurmuş

51- Aşağıdakilerden hangisinde bir koşula bağlı olma niteliği vardır?

- A) Onu görünce şaşırdım kaldım.
- B) Onu gördükten sonra buraya geldim
- C) Onu gördüğümde henüz bebekti.
- D) Onu görelî rahatım kaçtı
- E) Onu gördükçe seni hatırlıyorum

52- Eleştirmen kendine güvenecek; bilgiye, sezgiye, anlayışa dayanan bir yüreklilikle ortaya çıkacak “şu şöyledir, bu böyledir” diyecektir. Tartışmalar yaratacaktır. Bu onun varlık nedenidir. Eleştiri sanatının gereğidir.

Parçaya göre aşağıdakilerden hangisi, eleştirmenin varlık nedeni olamaz?

- A) Belli bir anlayışı yerleştirme
- B) Sanat ürünleri üzerinde yorum yapmak
- C) Sanat ürünleri üzerinde tartışma açmak
- D) Sanat ortamına canlılık getirmek
- E) Okuru değişik açılardan düşünmeye yöneltmek

53- “Denilebilir ki, buraların en çalışkan, en becerikli, en yırtıcı adamıdır. Ekmeğini taştan çıkarır.”

Bu söz, aşağıdakilerden hangisiyle sürdürülürse, kendisinden söz edilen kişinin olumsuz bir yönü belirtilmiş olur?

- A) Ne kendi ekmeği ile oynatır ne de başkalarınınkiniyle
- B) Ekmek elden su gölden demez
- C) Ekmek aslanın ağzında olsa alır
- D) Ne yapar eder bir ekmek kapısı bulur
- E) başkalarını ekmeğinden etse bile

54- Şiir dünyası, düşünce dünyasıyla o kadar bağdaşmaz ki gerçekten şair olanlar, her şeyden önce kelimeleri, günlük anlamlarından uzaklaştırmakla işe başlarlar. Okuyun büyük şairleri, hepsininin davranışı budur.

Parçadaki “kelimeleri günlük anlamlarından uzaklaştırmak” sözüyle ne anlatılmak isteniyor?

- A) Kelime seçiminde aşırı özen gösterme
- B) Kelimeleri duygu ve düşünce yönünden dengeleme
- C) Dili somutlaştırarak kullanma
- D) Söyleşiye açıklık kazandırma
- E) Kelimelere yeni anlamlar yükleme

55- Sözü uzatmak, büyütme, dallandırma, gereksiz kelimelerle doldurmak yoktur onda. Ne söyleyecek se en açık, en doğru biçimde söyler. Ama bu sözler bir araya geldi mi bir derinlik, bir anlamçoğalması, üzerinde uzun uzun durmak, incelemek, düşünmek gerekliliği yaratır.

Bu paragrafı geliştiren ve sürdüren bir yazı için aşağıdakilerden hangisi en uygun başlık olur?

- A) Kısa yazmanın yararları
- B) Özlü anlatım
- C) Düşüncede açıklık
- D) Yalınlığa giden yol
- E) Anlam bütünlüğü

56- Aşağıdaki cümlelerden hangisinde “boş” kelimesi, ‘elimde mutlaka bir iş olmalı, boş oturmasını sevmem.’ Cümlesindekiyle aynı anlamda kullanılmıştır?

- A) Onun çalıştığı şirkette galiba boş bir yer varmış
- B) Onun bu boş sözlerine artık kimse inanmıyor
- C) Satın aldıkları daire üç yıldan beri boş duruyormuş
- D) Arkadaşımın biyle boş gezmesi ailesini üzüyor
- E) Onun böylesine boş bir insan olduğunu bilmiyordum

57- Atlar, ulu bir sel gibi dağlardan boşanmış. Nalları, toprağı sarsıp şimşekleniyor. Geçtikleri yerlerde süt beyaz yollar bırakıyorlar. Üzerlerindeki binicileri toprağa fırlatıyorlar. Ovanın yüzü köpürmüş bir denizi andırıyor.

Parçanın yazarı, anlatım yönünden aşağıdakilerden hangisiyle nitelendirilebilir?

- A) İnce eleyip sık dokuyan
- B) Bin dereden su getiren
- C) Bire bin katan
- D) Kılı kırk yaran
- E) ağzına geleni söyleyen

58- Kasabaya yeni gelen kaymakam, kısa sürede öyle olaylara tanık olmuştu ki: "Bir ay içerisinde, bir ömürde 'öğrenebileceklerimin hepsini öğrenmiş gibiyim.' diyordu"

Bu sözleri söyleyen kişinin durumu aşağıdakilerden hangisiyle adlandırılabilir?

- A) Düş kırıklığına uğrama
- B) Bilgilerinin yetersizliğinden utanma
- C) Yaşamın gerçeklerinden korkma
- D) Yaşam deneyimi kazanma
- E) Bilgi kazanmakla 'övünme'

59- Kimi yargılar genel bir nitelik taşır. Bu örnekte olduğu gibi: "Dağlık bölgelerin insanları sert yaradılışlı olur." Bu tür yargılar zihnin genel düşünceler üretmesinin ya da özelden genele gidişin sonucudur.

Aşağıdakilerden hangisi bu parçada belirtilen yargı türüne örnek olamaz?

- A) Dünyanın bütün ülkelerinde polisiye romanlar çok satar
- B) Yazarlar yaşlılık dönemlerinde üne kavuşurlar
- C) Üslup, kelimeleri seçme sanattır
- D) Davranışları etkilemede iletişim araçlarının büyük payı vardır
- E) Sıcak ülkelerin insanları heyecanlı olur

60- "Bu yazar hakkında yüzlerce araştırma kitabı yazılmış; bunların bir kaçını son aylarda karıştırdım. Bir kez daha şu yargıya vardım. Bir yazarı, o yazarın kendi kelimeleri ile okumak gerekir. O kelimeler her şeyi içeriyor."

Burada söylenmek isteneni en iyi özetleyen cümle aşağıdakilerden hangisidir?

- A) Her yazar kendini tanıtırken özel kelimeler kullanır
- B) Bir yazarı, en iyi ancak kendi yazdıklarından tanıyabilirsiniz
- C) Kimi kelimeler yazarı tanıtmak için daha uygundur.

- D) Yazar kendini tanıtırken tarafsız davranır
- E) Yazar, öz geçmişini anlatırken öteki eserlerinden daha başarılı olurlar

61- Bir ülkenin sanatı, kültürü, o ülkeyi en az denizi, kumsalları, güneşi kadar tanıtır. Sanatın ve kültürün yaratıcısı insanoğlu olduğu için sanat ve kültür alanlarındaki gelişmeler, eserler, daha anlamlı, daha derin, daha insancıldır.

Bu paragrafı sürdürebilecek en uygun cümle aşağıdakilerden hangisidir?

- A) Sanata ve sanatçıya verilen önem bu açıdan ele alınmalıdır
- B) Bu nedenle sanat değeri, toplumun ihtiyaçlarını karşılamasında aranmalıdır
- C) Kültürle sanat arasındaki ilişkinin sınırı bu olmalıdır
- D) Toplumun gelişmesi bu sanat eserlerine bağlıdır
- E) Bunun için sanat, insanı geliştiren tek öge olmaktadır

**T.C. YÜKSEKÖĞRETİM KURULU
DOKÜMANTASYON MERKEZİ**