

ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ FAKÜLTESİ
ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI

ÇOKTAN SEÇMELİ, AÇIK UÇLU VE KARMA TESTLERİN PSİKOMETRİK
ÖZELLİKLERİNİN MADDE TEPKİ KURAMINA DAYALI
OLARAK DEĞERLENDİRİLMESİ

DOKTORA TEZİ

Selda GÜLTEKİN

Ankara
Haziran, 2011

ANKARA ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ FAKÜLTESİ
ÖLÇME VE DEĞERLENDİRME ANABİLİM DALI

ÇOKTAN SEÇMELİ, AÇIK UÇLU VE KARMA TESTLERİN PSİKOMETRİK
ÖZELLİKLERİNİN MADDE TEPKİ KURAMINA DAYALI
OLARAK DEĞERLENDİRİLMESİ

DOKTORA TEZİ

Selda GÜLTEKİN

Danışman: Doç. Dr. R. Nükhet ÇIKRIKÇI DEMİRTAŞLI

Ankara
Haziran, 2011

Eğitim Bilimleri Enstitüsü Müdürlüğü'ne

Bu çalışma jürimiz tarafından Ölçme ve Değerlendirme Anabilim Dalında DOKTORA TEZ ÇALIŞMASI RAPORU olarak kabul edilmiştir.

Başkan
Prof. Dr. Ezel TAVŞANCIL

Üye
Prof. Dr. Aysun UMay

Üye
Prof. Dr. Şener BÜYÜKÖZTÜRK

Üye
Doç. Dr. Nükhet ÇIKRIKÇI DEMİRTAŞLI (Danışman)

Üye
Doç. Dr. Hülya KELECİOĞLU

Onay

Yukarıdaki imzaların, adı geçen öğretim üyelerine ait olduğunu onaylarım.

07 / 06 / 2011

Prof. Dr. Necla KURUL
Enstitü Müdürü

ÖNSÖZ

Öğretim programlarının öğrenci merkezli hale gelmesi, sınıf içi değerlendirme etkinliklerinin de çeşitlenmesini gerektirmiştir. Ancak, ülkemizde öğrencilerin geleceğini etkileyen SBS, ÖSS gibi geniş ölçekli testler sadece çoktan seçmeli maddelerden oluşması, öğrencilerde tek yönlü bir çalışma alışkanlığı geliştirmektedir. Çoktan seçmeli madde formatı, o maddeyle ölçülen özelliğe sahip olmayan veya kısmen sahip olan bireylerin de şansa doğru yanıtlamalarına olanak verir. Bu durum, test puanlarına ilişkin kararların geçerliğini ve güvenilirliğini önemli ölçüde etkilemektedir. Bu doğrultuda, son yıllarda, farklı madde tiplerinin birbirlerine olan üstünlüğünü avantaja dönüştürmek için geniş ölçekli test uygulamalarında (TIMSS, PISA, SAT, CITO-ÖİS, vb.) çoktan seçmeli maddelerin yanı sıra, açık uçlu ve diğer (doğru-yanlış ve açık uçlu bir arada) madde türlerinin kullanılması ve Madde Tepki Kuramına (MTK) dayalı analizler yapılması yaygınlık kazanmaktadır. Bu noktadan hareketle benzer yapıları ölçen çoktan seçmeli ve yanıtı sınırlı açık uçlu madde formatlarının bir arada kullanıldığı durumlarda testlerin psikometrik niteliklerinin nasıl farklılaştığının ortaya konması ve ölçme uygulamalarında giderek yaygınlaşan MTK'ye dayalı ölçme model ve yöntemlerinin anlaşılması önemli hale gelmiştir.

Bu çalışmanın yapılmasında katkısı ve desteği bulunan insanlara teşekkürü borç bilirim. Öncelikle, tezimin her aşamasında bana yardımcı olan ve yol gösteren tez danışmanım Doç. Dr. Nükhet Çıkrıkçı Demirtaşlı'ya, önemli katkılarından dolayı Prof. Dr. Ezel Tavşancıl, Prof. Dr. Şener Büyüköztürk'e, Prof. Dr. Aysun Umay'a ve Doç. Dr. Hülya Kelecioğlu'na, görüşlerini ve yardımlarını esirgemeyen sevgili arkadaşlarım Yrd. Doç. Dr. Sevilay Kilmen, Yrd. Doç. Dr. İ. Alper Köse ve Araş. Gör. D. Tuğçe Coşkun'er'e, anlayışından dolayı TED Ankara Koleji Vakfı Özel Lisesi Müdürü Aydın Ünal'a teşekkür ederim. Son olarak, teşvikleri ve sabırlarından dolayı annem, babam, eşim ve biricik oğluma sonsuz teşekkürler...

Selda Gültekin

ÖZET

ÇOKTAN SEÇMELİ, AÇIK UÇLU VE KARMA TESTLERİN PSİKOMETRİK ÖZELLİKLERİNİN MADDE TEPKİ KURAMINA DAYALI OLARAK DEĞERLENDİRİLMESİ

Gültekin, Selda

Doktora, Ölçme ve Değerlendirme Anabilim Dalı

Tez danışmanı: Doç. Dr. Nükhet Çıkrıkçı Demirtaşlı

Haziran, 2011, 99 sayfa

Bu çalışmanın amacı, Matematik başarısına ilişkin olarak benzer yapıları ölçen iki madde tipinin (çoktan seçmeli ve yanıtı sınırlı açık uçlu) tek başına ve birlikte kullanıldıkları koşullar ile iki madde tipinin test içinde yer alma yüzdelerinin test formlarında farklı olduğu durumlarda MTK'ye göre kestirilen madde ve yetenek parametrelerinin psikometrik niteliklerinin, madde ve test bilgi fonksiyonlarının (item and test information function), görelî etkinlik (relative efficiency) düzeylerinin ve ölçüt dayanaklı geçerlik kanıtlarının farklı olup olmadığını araştırmaktır. Bu amaç doğrultusunda, TIMSS 2007 sekizinci sınıf matematik testi ikinci kitapçığındaki maddelerden test uzunluğu (madde sayısı: 15 ve 25) ve açık uçlu madde yüzdesi (%20 ve %40) farklı olarak oluşturulan karma testlerle, çoktan seçmeli ve açık uçlu testlere ilişkin, madde ve yetenek parametreleri, madde ve test bilgi fonksiyonları kestirilmiş, görelî etkinlik indeksleri hesaplanmıştır. Parametre kestirimleri, BILOG-MG programı kullanılarak yapılmıştır. Araştırma verileri, TIMSS 2007 matematik uygulamasına Türkiye'den katılan ve matematik testinin 2. kitapçığını alan 320 8.sınıf öğrencisinden elde edilmiştir.

Araştırma sonucunda, uzun karma testlerin, sadece çoktan seçmeli ve sadece açık uçlu maddelerden oluşan testlere göre orta güçlükte ve daha ayırıcı testler olduğu belirlenmiştir. Ayrıca, orta yetenek düzeyinde uzun karma testlerin daha az standart hatayla yetenek kestirimi yaptığı ortaya konmuştur. Buna ek olarak, testlerde, yanıtı sınırlı açık uçlu maddelerin çoktan seçmeli maddelerden daha fazla bilgi verdiği sonucuna ulaşılmıştır. Görelî etkinlik değerleri, açık uçlu testin diğer testlerden etkili olduğunu ortaya koymuştur. Bu doğrultuda, geniş ölçekli test uygulamalarında, çoktan seçmeli

maddelerin yanı sıra, yanıtı sınırlı açık uçlu maddelere de yer verilmesi, hem çoktan seçmeli maddelerden gelen şans başarısı kaynaklı hataların puanlara karışmamasını önleyebilecek, nitelikli sorular olmak koşuluyla ölçülen zihinsel becerilerin düzeyi, yükselebilecek ve bu testlerin sonuçlarına göre verilen seçme veya yeterlik kararlarının daha geçerli ve güvenilir olması sağlanabilecektir.

ABSTARCT

THE EVALUATION BASED ON ITEM RESPONSE THEORY OF THE PSYCHOMETRIC CHARACTERISTICS IN MULTIPLE CHOICE, CONSTRUCTED RESPONSE AND MIXED FORMAT TESTS

Gültekin, Selda

PhD, Measurement and Evaluation

Advisor: Assoc. Prof. N khet  ıkırık ı Demirtaşı

June, 2011, 99 pages

The aim of this study is to find out if there are differences in item and test parameters estimated in item response theory, test information function, reliability and proofs for criterion referenced validity in the conditions of two item types (multiple choice and constructed response) measuring common constructs used separately (alone) and together, and two item types take place in different proportions in tests. In line with this aim, the tests in which the length(15 and 25 items) and the proportion of open ended questions (20% and 40%) were different were combined from eighth grade Maths questions in the second testlet of TIMSS 2007. Through these tests, item and test parametres, item and test information functions, relative efficiency indices were estimated. BILOG-MG was used for estimating parametres. The data was collected from 320 eighth grade students who had taken the second Maths testlet in TIMSS 2007 in Turkey.

At the end of the research it is concluded that long mixed format tests are more intermediate level in difficulty and more discriminating than multiple choice and constructed response tests. In addition, it can also be said that long mixed tests prove to make less standart errors in ability estimation. It can also be determined that constructed response items provide more information function than multiple choice items. Relative efficiency values revealed that constructed response test is effective in other tests. In this regard, large-scale test applications, as well as multiple-choice items, be

included in a constructed response items, both multiple-choice scores to distinguish it from substances that can prevent the errors from the success of chance, provided that the qualified questions measured the level of mental skills, according to the results of these tests to rise and the select or qualification to be valid and reliable than the decisions will be provided.

İÇİNDEKİLER

	Sayfa
JÜRİ ÜYELERİNİN İMZA SAYFASI	i
ÖNSÖZ	ii
ÖZET	iii
ABSTRACT	v
ÇİZELGELER LİSTESİ	ix
ŞEKİLLER LİSTESİ	x
BÖLÜM 1	
GİRİŞ	1
Problem	1
Madde Tepki Kuramının Genel Yaklaşımı	4
Madde Tepki Kuramının Varsayımları	5
Madde Karakteristik Eğrisi ve Değişmezlik	5
Tek Boyutluluk	7
Yerel Bağımsızlık	9
Model Veri Uyumu	9
Madde Tepki Kuramında Parametreler	11
İlgili Araştırmalar	15
Klasik Test Kuramına Dayalı Çalışmalar	16
Madde Tepki Kuramına Dayalı Çalışmalar	19
Amaç	25
Önem	26
Sınırlılıklar	28
Tanımlar	28
Kısaltmalar	28
BÖLÜM 2	
YÖNTEM	30
Araştırma Modeli	30
Araştırma Grubu	30

Veri Toplama Araç ve Teknikleri	33
Verilerin Analizi	38
1. Ön Analizler – MTK Varsayımları.....	39
2. Araştırma Sorularına İlişkin Analizler	49
 BÖLÜM 3	
BULGULAR VE YORUMLAR.....	50
1. KT1, KT2, KT3, KT4, AUT ve ÇST için ayrı ayrı kestirilen madde parametreleri (b, a, c) hangi aralıkta değişmektedir?	50
2. KT1, KT2, KT3, KT4, AUT ve ÇST için ayrı ayrı kestirilen yetenek parametreleri ve bunlara ait standart hata değerleri hangi aralıkta değişmektedir?	53
3. Testlere ait madde ve test bilgi fonksiyonları, görelî etkinlik indeksleri hangi düzeydedir?.....	56
4. KT1, KT2, KT3, KT4, AUT ve ÇST için ayrı ayrı kestirilen yetenek parametreleri arasındaki Pearson Momentler Çarpım Korelasyon Katsayıları ne düzeydedir?	61
5. KT1, KT2, KT3, KT4, AUT ve ÇST için ayrı ayrı kestirilen yetenek parametreleri ile TIMSS 2007 Genel Matematik Başarı Puanları arasındaki Pearson Momentler Çarpımı Korelasyon Katsayıları ne düzeydedir?.....	62
 BÖLÜM 4	
SONUÇ VE ÖNERİLER.....	64
Sonuçlar	64
Öneriler	66
 KAYNAKLAR	68
 EKLER	
EK 1: Araştırma Kapsamında Seçilen Maddelere İlişkin Bilgiler	76
EK 2: Araştırma Kapsamında Ele Alınan Testler (AUT, ÇST, KT1, KT2, KT3, KT4).....	77

ÇİZELGELER LİSTESİ

	Sayfa
Çizelge 1: TIMSS 2007 8. Sınıflar Matematik Testi Katılımcı Ülkeler.....	31
Çizelge 2: TIMSS 2007 8. Sınıflar Matematik Testindeki 14 Kitapçıkta Yer Alan Maddelerin Madde Tipine Göre Dağılımları	35
Çizelge 3: TIMSS 2007 8. Sınıflar Matematik Testi 2. Kitapçıkta Yer Alan Maddelerin Madde Tipi ve Bilişsel Alanlara Göre Dağılımları	36
Çizelge 4: Karma Testler, Çoktan Seçmeli Test ve Açık Uçlu Testteki Madde Sayılarının Bilişsel Alan ve Madde Tipine Göre Dağılımları	37
Çizelge 5: KT1 (k=25, AU Yüzdesi=%40) İçin Faktör Analizi Sonuçları	40
Çizelge 6: KT2 (k=25, AU Yüzdesi =%20) İçin Faktör Analizi Sonuçları	41
Çizelge 7: KT3 (k=15, AU Yüzdesi =%40) İçin Faktör Analizi Sonuçları	42
Çizelge 8: KT4 (k=15, AU Yüzdesi =%20) İçin Faktör Analizi Sonuçları	44
Çizelge 9: AUT (k=10) İçin Faktör Analizi Sonuçları	45
Çizelge 10: ÇST (k=20) İçin Faktör Analizi Sonuçları	46
Çizelge 11: Açık Uçlu ve Çoktan Seçmeli Test ile Karma Testlerin Test Düzeyinde Model Veri Uyumu Analiz Sonuçları	48
Çizelge 12: Karma Testler, Çoktan Seçmeli Test ve Açık Uçlu Teste İlişkin Madde Parametreleri.....	51
Çizelge 13: Yetenek Parametrelerine İlişkin Betimsel İstatistikler	53
Çizelge 14: Yetenek Parametrelerine İlişkin Standart Hata Değerlerine ilişkin betimsel istatistikler	54
Çizelge 15: Karma Testler, Çoktan Seçmeli Test ve Açık Uçlu Teste İlişkin Madde Bilgi Fonksiyonları	57
Çizelge 16: Dört Karma Testte, ÇST ve AUT'ta Madde Formatına Göre Ortalama Test Bilgi Ortalamaları	58
Çizelge 17: Testlere İlişkin Görelî Etkinlik (Relative Efficiency) İndeksi Değerleri	58
Çizelge 18: Yetenek Düzeylerine Göre Test Bilgi Fonksiyonları Dağılımı	59
Çizelge 19: Yetenek Parametreleri Arasındaki Pearson Momentler Çarpımı Korelasyon Katsayıları	61
Çizelge 20: Yetenek Parametreleri İle TIMSS 2007 Genel Matematik Başarı Puanları Arasındaki Pearson Korelasyon Katsayıları.....	62

ŞEKİLLER LİSTESİ

	Sayfa
Şekil 1. Madde Karakteristik Eğrisi.....	6
Şekil 2: KT1 (k=25, AU Yüzdesi =%40) İçin Özdeğerler Grafiği.....	40
Şekil 3: KT2 (k=25, AU Yüzdesi =%20) İçin Özdeğerler Grafiği.....	42
Şekil 4: KT3 (k=15, AU Yüzdesi =%40) İçin Özdeğerler Grafiği.....	43
Şekil 5: KT4 (k=15, AU Yüzdesi =%20) İçin Özdeğerler Grafiği.....	44
Şekil 6: AUT (k=10) İçin Özdeğerler Grafiği	45
Şekil 7: ÇST (k=20) İçin Özdeğerler Grafiği	47
Şekil 8: KT1'i Alan Bireylerin Yetenek Parametrelerine İlişkin Standart Hata Dağılımları	54
Şekil 9: KT2'yi Alan Bireylerin Yetenek Parametrelerine İlişkin Standart Hata Dağılımları	54
Şekil 10: KT3'ü Alan Bireylerin Yetenek Parametrelerine İlişkin Standart Hata Dağılımları	55
Şekil 11: KT4'ü Alan Bireylerin Yetenek Parametrelerine İlişkin Standart Hata Dağılımları	55
Şekil 12: ÇST İçin Kestirilen Yetenek Parametrelerine İlişkin Standart Hata Dağılımları	55
Şekil 13: AUT İçin Kestirilen Yetenek Parametrelerine İlişkin Standart Hata Dağılımları	55
Şekil 14: Yetenek Düzeylerine Göre Test Bilgi Fonksiyonları Dağılımı	60

BÖLÜM 1

GİRİŞ

Bu bölümde, araştırmanın problemi, amacı ve alt amaçları, önemi, varsayımları ve sınırlılıklarına yer verilmiştir.

Problem

Ölçme ve değerlendirme işlemleri, eğitim öğretim sürecinin niteliğini sağlamada en önemli bileşenlerden biridir. Eğitim sisteminin başarılı olması yürütülen öğretim etkinliklerinin niteliği yanında ölçme ve değerlendirme araç ve yaklaşımlarının niteliğinden de etkilenir.

Günümüzde uygulanan öğretim programları, bireysel farklılıkları dikkate alan, öğrenci merkezli öğretme ve öğrenme stratejilerini benimseyerek geliştirildiği için, ölçme ve değerlendirmede de öğrencilerin bilgi, beceri ve tutumlarını sergilemelerine dayalı çoklu değerlendirmeler gerekmektedir (MEB, 2005). Öğrenci merkezli öğrenme yaklaşımlarının benimsendiği değişen öğretim programlarında, sınıf içi değerlendirme etkinliklerinin doğası değişmiş; öğretmenin hem öğretim hem de değerlendirme etkinliklerini çeşitlendirmesi ve birleştirmesini gerektirmiştir. Tek bir soru formatı ya da değerlendirme şeklini kullanmak öğrencilerde tek yönlü bir çalışma alışkanlığı geliştirerek; öğrenme çıktıları açısından yine tek yönlü özellikler ağırlık kazanmaktadır (Berberoğlu, 2006).

Eğitimde ölçme araçlarında kullanılan farklı madde türlerinin zayıf ve güçlü yönleri bulunmaktadır. Çoktan seçmeli maddelerden oluşan testler ile daha kısa sürede daha fazla sayıda beceri ölçülebilmesi, puanlanma kolaylığı ve objektifliği çoktan seçmeli testlerin bir üstünlüğüdür (Haladyna, 1997). Çoktan seçmeli madde türüyle, bilişsel alanın bazı düzeyindeki davranışları (bilgi, kavrama, uygulama, analiz) ölçmek de mümkündür (Tekin, 1991). Bununla birlikte, son yıllarda eğitim-öğretimde, okuduğunu anlama, kritik

etme, yorumlama, bilgiyi toplayıp analiz edebilme, bir sonuca ulaşma, grafik ya da tablo halinde verilen bilgiden sonuç çıkarma, uzaysal muhakeme, gözlem yapma, gözlemlerden sonuca ulaşma, günlük hayatta sıkça karşılaşılan problemleri çözebilme, araştırma yapma gibi öğrencilerin kendilerini sosyal yaşama daha çok hazırlayan becerilerin ağırlık kazandığı görülmektedir (Berberoğlu, 2006). Bu türden üst düzey düşünme becerilerinin sadece çoktan seçmeli maddelerden oluşan testlerle yoklanması zordur.

Çoktan seçmeli maddelerin önemli diğer bir sınırlılığı da bireylerin tahminle puan kazanma olanağının bulunmasıdır. Çoktan seçmeli madde formatının yapısı gereği seçenekler içermesi, o maddeyle ölçülen özelliğe sahip olmayan veya kısmen sahip olan yanıtlayıcıyı şansını kullanarak doğru yanıtı bulmaya yönltebilir. Şans başarısı olarak tanımlanan bu durum testin geçerlik ve güvenirliğini olumsuz etkilemektedir (Tekin, 1991). Öte yandan, uzun yanıtli maddelerde de geçerliğin sağlanmasında sıkıntılar yaşanmaktadır. İfade gücü, yazı güzelliği gibi asıl değişkenle ilgili olmayan bazı faktörlerin, soruların yazılarak yanıtlandığı bir sınavda puanlayıcıyı etkileyerek puanlamanın varyansını artırabilecektir. (Umay, 1997).

Farklı madde formatlarıyla ilgili süregelen tartışmalar, çoktan seçmeli maddelerin yanı sıra farklı madde türlerini de kullanma eğilimi ortaya çıkarmıştır. Farklı madde tiplerinin birbirlerine olan üstünlüğünü avantaja dönüştürmek için ölçme araçlarında farklı madde tiplerinin bir arada kullanılması yaygınlık kazanmaktadır. Son yıllarda geniş ölçekli test uygulamalarında çoktan seçmeli maddelerin yanı sıra, açık uçlu ve diğer (doğru-yanlış ve açık uçlu bir arada) madde türlerinin kullanıldığı görülmektedir. Dünyada “National Assessment of Educational Progress (NAEP), Massachusetts Comprehensive Assessment System (MCAS), Test of English as a Foreign Language (TOEFL), Programme for International Student Assessment (PISA), Advanced Placement Test (AP), CITO Türkiye Öğrenci izleme Sistemi-ÖİS” gibi farklı madde türlerinin bir arada kullanıldığı, akademik başarı ve yeterlik belirlemeye yönelik ölçme uygulamaları bulunmaktadır. Uygulandıkları ülkelerde yüzbinlerce bireyin katılımıyla gerçekleşen bu tür uygulamalar, karma test maddelerinin puanlanmasına yönelik geliştirilmiş özel yazılımlarla (Örn: IntelliMetric, E-rater, Intelligent

Essay Assessor, Project Essay Grade vb.) puanlanmaktadır. (Wainer&Thissen, 1993; Bastari, 2000; Shermis&Burstein, 2003; MEB, 2007; CITO Türkiye, 2009, Demirtaşlı, 2010).

Bu uygulamalarda kullanılmaya başlanan açık uçlu maddeler, kendi içerisinde yanıtı sınırlandırılmış ve yanıtı serbest bırakılmış sorular olmak üzere iki grupta ele alınmaktadır. Yanıtı sınırlandırılmış açık uçlu maddede öğrenciden yanıtın niteliğine, uzunluğuna ya da organizasyonuna yönelik sınırlamalar yapması istenir. Yanıtı sınırlandırılmış açık uçlu maddeler, hem daha çok soru sormaya olanak sağlaması hem de puanlamanın kolay olması nedeniyle daha çok tercih edilmesi gereken bir formattır (Kubizh ve Borich, 2003). Matematikte problem çözme becerisi yoklanırken kısa kısa, ama çok sayıda problem sorulması şans faktörünü büyük ölçüde azaltarak güvenilirliğin artmasına neden olur. “Açık Uçlu” olarak tanımlanan bu tip sorularda yazma ve ifade becerileri puanlamada etkili olmayacak ve yine çoktan seçmeli maddeler gibi 0 ve 1 şeklinde puanlanabilecektir (Umay, 1997).

Test geliştirme ve diğer ölçekleme uygulamaları belli kuramsal temellere dayalı olarak yürütülür. Bu kapsamda test kuramı ölçmelerin geçerlik ve güvenilirliğini sağlamak üzere çeşitli yöntem ve modeller kullanır (Crocker ve Algina, 1986). Psikometride ölçme araçlarının geliştirilmesi, puanlanması ve psikometrik nitelikleriyle ilgili sorunların ele alınmasında yaygın olarak iki test kuramından yararlanılır. Biri, klasik test kuramı (KTK) diğeri de kullanımı giderek yaygınlaşan Madde Tepki Kuramı (MTK)’dır. MTK, olasılıklı ölçme modelleri ile ölçme uygulamalarına bazı avantajlar sağlamıştır. Bunlar arasında; farklı amaçlar için test geliştirme, puanların eşitlenmesi, madde yanlılığının belirlenmesi ve bilgisayarlı bireye uyarlanmış testlerde yetenek puanlarının kestirimi gibi çeşitli ölçme problemlerinin çözümü için kullanışlı bir çerçeve sağlar (Hambleton, Swaminathan ve Rogers, 1991, Baker&Kim, 2004, Zhao, 2008).

Pek çok araştırmacı, farklı madde tiplerinin bir arada kullanılmasının madde tepki kuramındaki uygulamasını ele almıştır. Bu çalışmalarda, kullanılan farklı madde formatlarına ilişkin puanlama süreçleri ve bunların ölçme sonuçlarının geçerlik ve güvenilirliğini artırmaya etkisi olduğu üzerinde

durulmuştur (Wainer&Thissen, 1993, Bastari, 2000; Schaeffer ve diğerleri, 2002; Baker&Kim, 2004; Shin, 2007).

Test alanların geleceğini etkileyen SBS, ÖSS gibi geniş ölçekli testler sadece çoktan seçmeli maddelerden oluşmaktadır. Bu testler de yukarıda özetlendiği üzere çoktan seçmeli testlerin sınırlılıklarını taşımaktadır. Bu noktadan hareketle benzer yapıları ölçen çoktan seçmeli ve yanıtı sınırlı açık uçlu madde formatlarının bir arada kullanıldığı durumlarda testlerin psikometrik niteliklerinin nasıl farklılaştığının ortaya konması önem kazanmıştır. Bu kapsamda ölçme uygulamalarında giderek yaygınlaşan MTK'ye dayalı ölçme model ve yöntemlerinin anlaşılması önemli hale gelmiştir.

MTK'ye dayalı olarak farklı madde formatlarının bir arada kullanıldığı karma testlerin, ayrı ayrı kullanıldıkları testlerden psikometrik nitelikler bakımından nasıl farklılaştığının araştırıldığı bu çalışmada MTK'nin daha iyi anlaşılması için, MTK'nin genel yaklaşımı, varsayımları, modelleri, parametreleri aşağıda geniş biçimde açıklanmıştır.

Madde Tepki Kuramının Genel Yaklaşımı

Yirminci yüzyılın ilk 40 yıllık diliminde zeka, akademik başarı, yeterlik gibi özelliklerin ölçülmesine yönelik girişimler ve sonrasında gelen tartışmalar, Klasik Test Kuramına (KTK) dayalı olarak yürütülmüştür. Klasik Test Kuramına dayalı test geliştirme ve ölçme araçlarının uygulamasında elde edilen madde ve test istatistiklerinin, testin uygulandığı yetenek düzeyine göre değişmesi, buna bağlı olarak bireylerin ve grupların karşılaştırılmasındaki güçlükler alternatif kuram arayışlarına yol açan önemli bir sınırlılık olmuştur (Hambleton, Swaminathan ve Rogers, 1991). Klasik Test Kuramının uygulamasında karşılaşılan bu sınırlılıklar neticesinde Lord ve Novick (1968) modern matematiksel istatistik yaklaşımlarının kullanıldığı kuramın temel yapılarını yeniden formüle etmişler ve günümüzde bilinen adıyla Madde Tepki Kuramının temel kavramlarını oluşturmuşlardır (Baker&Kim, 2004). Madde Tepki Kuramı, kişinin ölçülen özellikteki (yeterlik)

düzeyi ile verdiği yanıtlar arasındaki ilişki olduğunu kabul eder, bu ilişkiyi matematiksel bir fonksiyon ile açıklayan olasılıklı bir model önerir. (Embretson ve Reise, 2000). Madde Tepki Kuramı iki temel kabule dayanır: (a) doğrudan gözlenemeyen örtük özellik ya da yetenek/yeterlik olarak adlandırılan psikolojik yapı, bireylerin test maddelerindeki gözlenen performansından kestirilebilir, (b) bireylerin maddelerdeki performansı ile maddeki performansından sorumlu olan özellik arasındaki ilişki, madde karakteristik fonksiyonu/eğrisi olarak adlandırılan doğrusal olmayan bir fonksiyonla açıklanabilir (Hambleton, Swaminathan ve Rogers, 1991). Kuram, belli varsayımlar altında madde setinden bağımsız yetenek parametreleri ve yanıtlayıcı örneklemeden bağımsız madde parametreleri kestirebileceğini iddia eder.

Madde Tepki Kuramının Varsayımları

Madde Tepki Kuramının avantajlarından (test eşitleme, bilgisayarlı bireye uyarlanmış test, amaca uygun test geliştirme vb.) yararlanabilmek için güçlü varsayımların karşılanması gerekmektedir. Her modelin gerektirdiği ek varsayımlar olsa da tüm tek boyutlu MTK modelleri için iki temel varsayım; (a) madde karakteristik eğrilerinin değişmezliği ve (b) yerel bağımsızlık ve tek boyutluluğun sağlanması gerekmektedir (Embretson ve Reise, 2000).

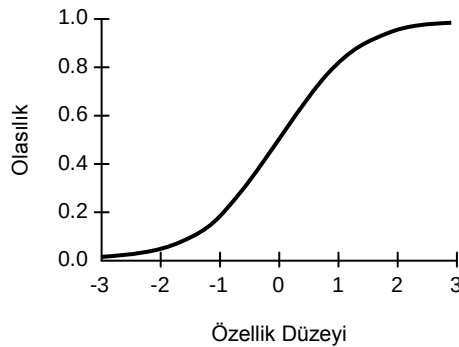
Madde Karakteristik Eğrisi ve Değişmezlik

Klasik Test Kuramında test ve madde istatistikleri testin uygulandığı gruba bağlı olarak hesaplanır. Testin uygulandığı grubun yetenek düzeyi test ve madde istatistiklerini değiştirmektedir. Bu durumda bireyin ilgili testten aldığı puan, maddelerin zor ya da kolay olmasına bağlıdır. Test zor olduğunda, bireyin yeterlik düzeyi düşük gibi görünürken; test kolay olduğunda, bireyin yeterlik düzeyi yüksek gibi görünmektedir. Ayrıca, testin zor ya da kolay olması da testin uygulandığı örnekleme bağlıdır. Test maddelerinin güçlüğü ilgili maddenin doğru yanıtlanma yüzdesidir. Maddenin

kolay ya da zor olması bireylerin ölçülen yeterlik düzeylerine bağlıyken, bireylerin yeterlik düzeyleri test maddelerinin kolay ya da zor olmasına bağlıdır.

Madde ayıricılığı, test puanlarının güvenilirliği ve geçerliği de belirli bir gruba dayalı olarak belirlenmektedir. Bireylerin testle ölçülen özellikleri değiştikçe test ve madde istatistikleri değişmekte, test ve madde özellikleri değiştikçe bireylerin testle ölçülen özellikteki yerleri de değişmektedir. Bu durum, farklı testler alan bireylerin karşılaştırılmasını ve farklı gruplardan elde edilen madde karakteristiklerinin karşılaştırılmasını güçleştirmektedir (Hambleton, Swaminathan ve Rogers, 1991). Ancak, MTK'nin değişmezlik özelliği bu sorunu ortadan kaldırmaktadır. Test ve madde parametreleri testin uygulandığı örneklemden bağımsız (sample-free), bireyin yetenek düzeyi madde parametrelerinden bağımsız (item - free) kestirilmektedir.

Bir maddenin madde karakteristik eğrisi, model veri uyumu sağlanan verilerde aynı evrenden gelen farklı alt grupları için ya da farklı dağılımlara sahip gruplar için değişmezdir (invariant). Bir başka deyişle, madde karakteristik eğrisinin biçimi aynı evrenden gelmiş alt örneklemeler için farklılık göstermez. Bir madde karakteristik eğrisinin şekli (Şekil 1), yeterlik düzeyi değişimine bağlı olarak tepkilerin verilme olasılığındaki değişimin nasıl olacağını ifade etmektedir. Madde karakteristik eğrisinin şeklini üç parametre belirler. Bunlar güçlük, ayıricılık ve şans parametreleridir. Madde güçlüğü'nün artması, eğrinin sağa doğru yatıklaşmasına, madde ayıricılık parametresinin artması eğrinin eğiminin artmasına, şans parametresinin artması ise eğrinin olasılık ekseninde sıfırdan başlamamasına neden olur. (Embretson ve Reise, 2000, Baker ve Kim, 2004).



Şekil 1. Madde Karakteristik Eğrisi

Tek Boyutluluk

MTK modellerinde yaygın olarak kullanılan bir varsayım, uygulanan testin maddelerinin sadece tek bir örtük özelliği ölçmesidir. Bu varsayım tek boyutluluk olarak adlandırılır. Test performansının genellikle çeşitli bilişsel, kişisel faktörlerden etkilenmesi, bu varsayımın kesin olarak karşılanmasını engelleyebilmektedir. Bu faktörler, motivasyon, kaygı, testi yanıtlama hızı, bilişsel beceri düzeyi ve doğru yanıtın tahminle bulunabilme düzeyi, test maddelerinin ölçtüğü yapıyı belirlemede rol oynamaktadır. Bu durumda tek boyutluluk varsayımın karşılanması, test performansını etkileyen baskın (başat) bir faktör ya da bileşen olmasıyla elde edilir. Bu baskın faktör ya da bileşen testle ölçülen özellik – yetenek olarak adlandırılır (Hambleton, Swaminathan ve Rogers, 1991).

Crocker ve Algina'ya (1986) göre, bir testin tek boyutlu olması, maddeler arasındaki istatistiksel bağımsızlığa dayanmaktadır. Maddeler arasındaki istatistiksel bağımlılık tek bir örtük özellik tarafından açıklanabiliyorsa tek boyutluluk söz konusudur. Tek boyutluluğun olduğu durumda, bireyler örtük özellik bakımından homojen bir yapı gösterirken; maddeler istatistiksel olarak birbirinden bağımsızdır. Madde Tepki Kuramında istatistiksel bağımlılık ve istatistiksel bağımsızlık kavramları, yaygın olarak değişkenler arası ilişkileri açıklamada kullanılır. İki kategorili puanlanan iki madde için bu kavramlar şu şekilde ifade edilebilir: $P_i(+)$ i . maddenin doğru yanıtlanma olasılığını, $P_i(-)$, i . maddenin yanlış yanıtlanma olasılığını gösterebilir. $P_j(+)$ ve $P_j(-)$ 'de j . madde için aynı şekilde tanımlanmış olsun. $P(+,+)$, $P(-,-)$, $P(+,-)$ ve $P(-,+)$ bu iki maddenin tepki örüntülerine ilişkin olasılıkları gösterebilir. Örneğin, $P(+,+)$ i . ve j . maddeye doğru yanıt verme olasılığıdır. Bu tanımlamalarla birlikte, eğer iki madde istatistiksel olarak bağımsızsa;

$$P(+,+) = P_i(+). P_j(+)$$

$$P(+,-) = P_i(+). P_j(-)$$

$$P(-,+) = P_i(-). P_j(+)$$

$$P(-,-) = P_i(-). P_j(-) \quad \text{dır.}$$

Eğer iki maddenin istatistiksel olarak bağımlıysa, bu eşitlik de sağlanamaz.

Tek boyutluluğun sınanmasında çeşitli yöntemler kullanılır. Bunlardan biri, tetrakorik korelasyon matrisi üzerinde yapılan faktör analizidir. Bu yöntemde, (i) birinci faktör yüküne ait özdeğerin diğerlerinden oldukça büyük olmasının ve (ii) ikinci faktör yüküne ait özdeğerlerin geriye kalan faktörlerin özdeğerlerinden çok farklı olmamasının yaklaşık bir ölçü olarak alınabileceği belirtilmiştir (Lord 1983). Diğer bir yöntem ise tek boyutluluk araştırmasında yaygın olarak kullanılan temel bileşenler analizidir. Swaminathan ve diğerleri (2006), doğrusal faktör analizinde, maddeler arası korelasyon matrisi ve saçılma diyagramı oluşturulduğunu, en büyük özdeğerin açıkladığı varyans yüzdesi bulunduğunu ve bunlara bağlı olarak boyutluluk hakkında yargıya varıldığını belirtmişlerdir.

Horn (1965) tarafından önerilen paralel analiz ve Drasgow ve Lissak (1983) tarafından geliştirilen değişimlenmiş (modify) paralel analiz, klasik p değeri gibi madde istatistiklerini kontrol ederek, üretilmiş (simülatif) madde tepkilerinin rasgele normal sapmaları için basit bir yol göstermişlerdir. Bu yol, suni / üretilmiş veriye dayalı maddeler arası korelasyon matrislerini verir ve en büyük özdeğerin açıkladığı varyans yüzdesini hesaplar. Yani, gerçek madde yanıtları analizinden elde edilen özdeğerle, üretilmiş veri karşılaştırılabilmektedir. Alternatif olarak, üretilmiş veri ile gerçek veriyi karşılaştırmak, üretilmiş veri için artıkların (residuals) dağılımının analiziyle de yapılabilmektedir. Ancak, tetrakorik korelasyonlar matrisi pozitif sürekli değilse, en çok olabilirlik işlemi başarısız olmaktadır. Bu karşılaştırmaların neticesinde, üretilmiş veri ile gerçek veri setlerinin dağılımları benzer bulunduğu tek boyutluluk kararına varılabilmektedir (Swaminathan, Hambleton & Rogers, 2006).

Yerel Bağımsızlık

Yerel bağımsızlık, bireyin herhangi bir madde çiftine verdiği yanıtların istatistiksel olarak bağımsız olmasıdır. Diğer bir deyişle, bireyin testle ölçülen özelliği sabit tutulduğunda bir maddeye doğru yanıt verme olasılığının, diğer maddelere verdiği yanıtlardan etkilenmeyeceği anlamına gelmektedir. Birey parametreleri (θ_s) ve madde parametreleri (ξ_i) kontrol edildiğinde, herhangi bir maddeye doğru yanıt verme olasılığının ($P(X_{is}=1)$), başka bir maddeye doğru yanıt verme olasılığından ($P(X_{i's}=1)$) bağımsız olması; tepki örüntüsü verilen bir birey için, bir madde setinde tepki örüntüsünün olasılığı, her bir madde tepkilerinin olasılıklarının çarpımına eşit olmasıdır (Embretson&Reise, 2000). Yerel bağımsızlık, madde performansı altında yatan farklı örtük özelliklerin sayısı ile de ilişkilidir. Dolayısıyla tek boyutluluk varsayımı karşılandığında, yerel bağımsızlık da sağlanmış olur (Hambleton, Swaminathan ve Rogers, 1991).

Model Veri Uyumu

Farklı MTK modellerinden hangisinin çalışılan veriye en iyi uyum gösterdiğine karar vermek için uygulanabilecek çok sayıda kriter vardır. Hambleton ve diğerleri (1991), MTK model seçiminin anlamlılığını test etmek için madde ve yetenek parametresi kestirimlerinin değişmezliğinin kontrol edilmesini, gerçek ve üretilmiş madde tepkileri üzerinden gözlenen ve beklenen test puan dağılımlarının karşılaştırılmasını önermişlerdir.

İstatistiksel hipotez testlerinin kullanıldığı uyum iyiliği testleri, iki örneklemin aynı dağılımdan gelip gelmediğini ya da frekansların belirli bir dağılıma uyup uymadığını (Kaykare- χ^2 testi gibi) test eder. Model veri uyumu değerlendirmeleri test düzeyinde ve madde düzeyinde yapılabilir (Embretson ve Reise, 2000).

Madde Düzeyinde Uyum

Model uyum hipotezinde, Pearson istatistiği ve olabilirlik oran (likelihood ratio) istatistiği yaygın olarak kullanılan uyum iyiliği istatistikleridir. Bu istatistiklerin ikisi de kaykare analizlerine dayalıdır. MTK modellerinde, Pearson uyum istatistiği Bock'un kaykare (BCHI) ve Yen'in kaykare (Q_1) istatistiklerini içerir. Olabilirlik oran kaykare istatistiği McKinley ve Mills (G^2) tarafından önerilmiştir. Üç yöntemde de bireyler, kestirilen yetenek düzeylerine göre yetenek ölçeğinde farklı gruplara ayrılırlar ve her bir alt yetenek grubundaki belirli madde puanlarının beklenen ve gözlenen oranları hesaplanır. Bu istatistiğin serbestlik derecesine göre anlamlı olmasıyla model veri uyumuna karar verilir. BCHI, Q_1 ve G^2 istatistikleri BILOG ve PARSCALE programlarında hesaplanabilmektedir (Zhao, 2008). Q_1 istatistiği örneklem büyüklüğünden etkilenmektedir. Küçük örneklerde daha az madde uyumsuz görünürken; büyük örneklerde uyum sağlamayan madde sayısı artmaktadır (Hambleton, Swaminathan ve Rogers, 1991).

Test Düzeyinde Uyum

Bir maddenin, bir modele uyum göstermediği belirlendiğinde, bu maddenin kullanmaya devam edip edilmeyeceği kararına test düzeyinde model veri uyumunun araştırılması ile varılır. Hambleton ve Hen (2005), madde ve yetenek parametreleri kestirimlerine ait beklenen ve gözlenen dağılımlarının karşılaştırılmasını önermişlerdir (Zhao, 2008).

En çok olabilirlik kestiriminde “-2loglikelihood” değeri model veri uyumu indeksi olarak kullanılır. Farklı modeller için kestirilen bu değerler arasındaki fark kaykare istatistiğiyle değerlendirilir. İki model için kestirilen -2loglikelihood değerleri arasındaki fark kaykare tablo değeriyle karşılaştırılır. Fark tablo değerinden büyükse fark anlamlıdır. Fark anlamlı ise daha kompleks model, anlamlı değil ise daha basit model seçilir (Embretson ve Reise, 2000, s.74).

Madde Tepki Kuramında Parametreler

Madde Tepki Kuramında, kestirilen madde ve yetenek parametreleri, madde-test bilgi fonksiyonu ve görelî etkinlik (relative efficiency) indeksinin kestirimi ve yorumuna ilişkin bilgiler şöyle özetlenebilir.

Madde Parametreleri

Madde Tepki Kuramında madde güçlük (b), madde ayıricılık (a) ve madde şans (c) parametreleri kullanıldıkları modellerde açıklanmalıdır. Madde tepki kuramında modeller “boyutluluk, puanlama, doğrusallık yapısı” gibi özelliklere göre sınıflanmaktadır. Tek boyutlu, ikili puanlanan (binary/dichotomous) modellere ve ihtiyaç duyulan parametrelere ilişkin sınıflama ve açıklamalar aşağıda belirtilmiştir (Hambleton, Swaminathan ve Rogers, 1991, Embretson&Reise, 2000, Baker ve Kim, 2004).

İkili Puanlanan Maddeler İçin Modeller

Bu modeller maddelerin 1-0 şeklinde puanlandığı durumlar için uygundur. Bu sınıflamada modeller içerdikleri parametre sayısına göre isimlendirilmektedir.

- **1 Parametrelî Lojistik Model**

Madde Tepki Modellerinin en yalın olanıdır. Bireyin bir maddeye doğru yanıt verme olasılığını kişinin yetenek düzeyine (θ) ve o maddenin güçlüğüne (b) bağılı olarak açıklar. Maddenin ayıricılığı (a) kullanılan programa göre ortalama ayıricılık ya da sabit değer olarak hesaplar. Bu model, ayıricılığın 1 olduğu koşul için Rasch olarak adlandırılmaktadır. Bu modelde, s. bireyin i. maddeyi doğru yanıtlama olasılığı aşağıdaki denklemle ifade edilir:

$$P(X_{is} = 1 | \theta_s, b_i) = \frac{\exp(\theta_s - b_i)}{1 + \exp(\theta_s - b_i)}$$

- **2 Parametrelili Lojistik Model**

1 parametrelili lojistik modele ayırıcılık parametresinin katılmasıyla elde edilir. Bu modelde bireyin bir maddeye doğru yanıt verme olasılığı, kişinin yetenek düzeyine (θ), o maddenin güçlüğüne (b) ve ayırıcılığına (a) bağlıdır. Bu modelde, s. bireyin i. maddeyi doğru yanıtlama olasılığı aşağıdaki denklemle ifade edilir:

$$P(X_{is} = 1 | \theta_s, b_i, a_i) = \frac{\exp[a_i (\theta_s - b_i)]}{1 + \exp[a_i (\theta_s - b_i)]}$$

- **3 Parametrelili Lojistik Model**

Bir madde tahminle de doğru yanıtlanabiliyorsa, örneğin çoktan seçmeli bir madde, madde karakteristik eğrisi sıfırdan başlamaz. Bu durumda şans parametresi (c) devreye girer. Bu modelde bireyin bir maddeye doğru yanıt verme olasılığı, kişinin yetenek düzeyine (θ), o maddenin güçlüğüne (b), ayırıcılığına (a) ve şansla doğru yanıtlama olasılığına (c) bağlıdır. Bu modelde, s. bireyin i. maddeyi doğru yanıtlama olasılığı aşağıdaki denklemle ifade edilir:

$$P(X_{is} = 1 | \theta_s, b_i, a_i, c_i) = c_i + (1 - c_i) \frac{\exp[a_i (\theta_s - b_i)]}{1 + \exp[a_i (\theta_s - b_i)]}$$

Yetenek Parametresi

Yetenek aralığı teorik olarak $(-\infty, +\infty)$ aralığında olsa da, uygulamada genellikle $(-3, +3)$ aralığında sınırlandırılmaktadır. Her bir yetenek düzeyinde, o yetenekteki bireylerin ilgili maddeye doğru yanıt verme olasılıkları tanımlanmaktadır. Yetenek düzeyi (θ) ve buna karşılık gelen doğru yanıtlanma olasılığı $P(\theta)$ 'nin grafiği Madde Karakteristik Eğrisi (MKE) olarak adlandırılmaktadır. MKE, maddeye ilişkin üç teknik özelliği açıklamada kullanıldığından MTK modellerinin temelini oluşturmaktadır. Bunlardan ilki,

madde güçlük (b) parametresidir. Madde güçlüğü, bir maddeye .50 olasılıkla doğru yanıt veren bireylerin yetenek düzeyini göstermektedir. θ ile aynı ölçektedir. -3'e yaklaşması o maddenin kolay, +3'e yaklaşması zor olduğunu gösterir. İkincisi, MKE'nin eğimi olan ayıricılık (a) parametresidir. Ayıricılık parametresi ne kadar büyükse o madde o kadar ayırdedici demektir. Üçüncü özellik ise şans parametresidir. $\theta = -3$ noktasında bireyin o maddeye doğru yanıt verme olasılığını, diğer bir deyişle o maddenin şansla doğru yanıtlanma olasılığını göstermektedir. (Embretson ve Reise, 2000, Baker, 2001).

Ölçmenin Standart Hatası

Klasik Test Kuramında, ölçmenin standart hatası tüm puan düzeylerinde eşit değer alırken; Madde Tepki Kuramında her bir birey için kestirilen yetenek parametresinin standart hatası da ayrı olarak kestirilmektedir. Standart hatalar, maddeler belirli yetenek düzeylerine en iyi uyumu sağladıklarında ve madde ayıricılıkları yüksek olduğunda en küçük değerlerini alırlar. Her bir yetenek düzeyinde daha küçük standart hatalar, güvenilirliği de yükseltmektedir (Embretson ve Reise, 2000).

Madde ve Test Bilgi Fonksiyonu

Madde Tepki Kuramında madde ve testleri tanımlamada, test maddelerini seçmede ve testleri karşılaştırmada kullanılan güçlü bir yöntem "Madde Bilgi Fonksiyonu" ve "Test Bilgi Fonksiyonu"dur.

$$I_i(\theta) = \frac{P_i'(\theta)^2}{P_i(\theta) Q_i(\theta)}$$

Bu eşitlikte; $I_i(\theta)$: i. maddenin madde bilgi fonksiyonunu, $P_i(\theta)$: θ 'ya bağlı madde tepki fonksiyonunu, $Q_i(\theta)$: $1 - P_i(\theta)$ 'yı ve $P_i'(\theta)$: $P_i(\theta)$ 'nın birinci türevini göstermektedir. İkili puanlanan maddelerde üç parametrelili lojistik model için $P_i(\theta)$ 'nın açılımı bu eşitlikte yerine konduğunda aşağıdaki eşitlik elde edilmektedir:

$$I(\theta) = \frac{2,89 a_i^2 (1 - c_i)}{1 + \exp(1,7 a_i (\theta - b_i)) + \exp(-1,7 a_i (\theta - b_i))^2}$$

Bu eşitlikte, madde bilgi fonksiyonunun madde parametreleriyle ilişkisi açıkça görülmektedir. Bir madde, madde güçlük parametresi (b) hitap ettiği yetenek düzeyi θ 'ya yaklaştıkça, madde ayıricılığı (a) arttıkça ve şans parametresi (c) sıfıra yaklaştıkça daha fazla bilgi vermektedir.

θ yetenek düzeyinde bir testin sağladığı bilgi, o yetenek düzeyinde madde bilgi fonksiyonlarının toplamından elde edilir.

$$I(\theta) = \sum_{i=1}^n I_i(\theta)$$

Bu durum, her bir test maddesinin katkısının, testteki diğer maddeler bilinmediğinde de tanımlanabildiğini ortaya koymaktadır. Test bilgi fonksiyonu, Klasik Test Kuramındaki güvenirlik katsayısından farklı olarak, testin uygulandığı örneklemden tamamen bağımsızdır.

Bir test bilgi fonksiyonu, yetenek düzeyi ölçeğinde bazı noktalarda en yüksek değerini alırken; tüm yetenek düzeylerinde eşit olmayan değerler alabilir. Böyle bir testin, yetenek düzeyleri test bilgi fonksiyonunun en yüksek değere ulaştığı noktanın yakınlarındaki bireylerin yetenek kestiminde iyi olduğu söylenmektedir. Bazı testlerde ise test bilgi fonksiyonu yetenek düzeyinin bazı noktalarında daha düz-yassı bir dağılım gösterebilir. Test bilgi fonksiyonu bu testlerin, o yetenek düzeyi aralığındaki bireylere hitap ettiğini gösterir. Test bilgi fonksiyonu yorumlanırken; yetenek düzeyi ve test bilgi fonksiyonu arasındaki ilişki göz önünde tutulmalıdır. (Hambleton, Swaminathan ve Rogers, 1991, Embretson ve Reise, 2000, Baker, 2001).

Görelî Etkinlik (Relative Efficiency) İndeksi

Test bilgi fonksiyonu, test geliştirme ve seçme açısından önem taşımaktadır. Ancak; test ve madde bilgi fonksiyonu, kesin olarak yorumlanamadığında, puan formülünün ya da testlerin değerinin görelî olarak karşılaştırılması mümkündür. Genel olarak, $I_A(\theta)$ ve $I_B(\theta)$ iki teste ilişkin aynı θ düzeyinde test bilgi fonksiyonlarını gösterdiğinde; görelî etkinlik (relative efficiency):

$$RE(\theta) = I_A(\theta) / I_B(\theta)$$

formülüyle hesaplanmaktadır. Örneğin; $I_A(\theta)=25$ ve $I_B(\theta)=20$ olduğunda, $RE(\theta)=1.25$ olarak bulunmaktadır. Bu durumda, “Belirli θ düzeyinde, Test A, Test B’den %25 daha uzunmuş gibi çalışmaktadır.” yorumu yapılabilmektedir. Bu kavram, test modeli ve puanlama formülünün seçiminde önemli rol oynamaktadır. (Lord ve Novick, 1968, Hambleton & Swaminathan, 1985, Hambleton, Swaminathan ve Rogers, 1991).

Bu kuramsal çerçeve kapsamında araştırmanın problemi ile ilgili olarak, farklı madde formatlarının ayrı ayrı ve birarada kullanıldığı durumlarda testlerin psikometrik niteliklerinin nasıl değiştiğini gösteren araştırmaların bulgularına yer verilmiştir.

İlgili Araştırmalar

Bu kısımda, ilgili alanyazında farklı madde tiplerinin birlikte kullanıldığı ölçme durumlarında madde ve testlerin psikometrik niteliklerinin araştırıldığı çalışmaların bulguları özetlenmiştir. Bu çalışmalarda, Klasik Test Kuramı ve Madde Tepki Kuramına dayalı analizler yapılmıştır.

Klasik Test Kuramına Dayalı Çalışmalar

Uğurlu (1996), aynı davranışı ölçmeye yönelik kısa yanıtlı ve çoktan seçmeli maddelerden oluşan testlerin madde ve test istatistikleri arasında farklılık olup olmadığını araştırdığı çalışmasında, ortalama güçlük indeksleri, KR 20 güvenirlik katsayıları ve ölçüt dayanaklı geçerlik katsayıları arasındaki farkların manidarlığını test etmiştir. Araştırmacı tarafından geliştirilen 5. sınıf Fen Bilgisi Vücudumuzu Tanıyalım ünitesine ilişkin başarı testi, 1994-1995 öğretim yılında 10 okuldan 293 öğrenciye uygulanmıştır. Geçerlik ölçütü olarak 1.dönem karne notları alınmıştır. Sonuç olarak, kısa yanıtlı ve çoktan seçmeli maddelerden oluşan testlerin güçlük, güvenirlik ve geçerlik bakımından manidar farklılık göstermediği belirtilmiştir.

Özdemir (1997), aynı davranışı ölçen kompozisyon tipi, kısa yanıtlı ve çoktan seçmeli maddelerden oluşan testlerin psikometrik özellikleri arasındaki ilişkiyi araştırdığı çalışmasında, üç test maddesinin güçlük indeksleri, ayıricılık indeksleri, varyansları ve güvenirlik indeksleri ile test puanlarının ortalama güçlük düzeyleri ve güvenirlik katsayıları değişkenleri ele alınmıştır. Araştırmacı tarafından geliştirilen testler, 1996-1997 öğretim yılında Hacettepe Üniversitesi Eğitim Bilimleri bölümünde okuyan ve İstatistik I dersini alan 173 öğrenciye uygulanmıştır. Değişkenler arasındaki ilişki Pearson Momentler Çarpım Korelasyon Katsayısı hesaplanarak araştırılmıştır. Sonuç olarak, çoktan seçmeli testin daha kolay; ancak daha ayırt edici olduğu, varyansı en büyük olan testin kompozisyon tipi test olduğu, üç testin de madde güvenirlik indekslerinin düşük değerler verdiği, test güvenirliği bakımından en güvenilir testin çoktan seçmeli test olduğu belirlenmiştir.

Araz (2001), aynı davranışı ölçen kısa yanıtlı testlerin madde ve test özellikleri ile şans başarısından arıtılmış ve şans başarısından arıtılmamış üç ve beş seçenekli çoktan seçmeli testlerin; madde güçlük indeksleri, madde ayıricılık gücü indeksleri, test puanları, varyansları, güvenirlikleri ve geçerlikleri bakımından farklılık gösterip göstermediğini araştırmıştır. Araştırmacı tarafından geliştirilen sözel yetenek testi 2000-2001 öğretim

yılında ODTÜ Geliştirme Vakfı Özel İlköğretim Okulunda öğrenim gören 169 6.sınıf öğrencisine uygulanmıştır. Sonuç olarak, seçenek sayısının azalmasının madde güçlük indekslerini artırdığı, en düşük ayıricılığın ve güvenilirliğin kısa yanıtli testten elde edildiği, şans başarısından artıldığına üç ve beş seçenekli testlerin en yüksek ayıricılıklara ulaştığı, şans başarısından artıldığına testin varyansının arttığı belirlenmiştir.

Gözen (2002), çoktan seçmeli ve kısa cevaplı test maddelerinin ağırlıklı puanlanmasının öğrencilerin test puanlarına ve testin geçerliğine etkisini belirlemeyi amaçlamıştır. Araştırmanın verileri 2001-2002 öğretim yılında iki üniversitenin Sınıf Öğretmenliği ve Rehberlik ve Psikolojik Danışma bölümlerinde okuyan 316 birinci sınıf öğrencisinden elde edilmiştir. Uygulamada biri çoktan seçmeli, diğeri kısa cevaplı iki sözel yetenek testi kullanılmıştır. Uygulanan test maddeleri her bir test türü için “1- 0” puanlama yöntemiyle, madde güçlük indekslerine bağı ağırlıklı puanlama yöntemiyle (1-pj) ve madde güçlük indeksleriyle madde ayıricılık gücü indekslerinin birlikte kullanıldığı ağırlıklı puanlama yöntemiyle [(1-pj)rjx] puanlanmıştır. Farklı puanlama yöntemlerinden elde edilen test puanları arasındaki ilişkiler ve testlerin geçerlik katsayıları arasında fark olup olmadığı incelenmiştir. Araştırmanın bulguları, farklı puanlama yöntemlerinin kullanılmasının test puanları üzerinde aynı etkiyi yaptığını ve test geçerlikleri arasında da manidar bir farka neden olmadığını göstermiştir.

Özkul (2002), aynı davranışı ölçen ağırlıklandırılmış puanlamalı kısa yanıtli, “1-0” yöntemiyle puanlanmış kısa yanıtli ve çoktan seçmeli testlerin madde ve test özellikleri arasındaki farklılıkları ortaya çıkarmak amacıyla bir çalışma yapmıştır. Ağırlıklandırılmış puanlamalı kısa yanıtli test, görgül ağırlıklandırma yöntemiyle anahtarlı ve anahtarsız olarak puanlanmıştır. Bu çalışmada, karşılaştırılan test ve madde özellikleri madde ayıricılık gücü indeksleri, test güvenilirlikleri ve test geçerlikleri olarak belirlenmiştir. Araştırmacı tarafından geliştirilen sözel yetenek testleri 2000-2001 öğretim yılında bir Özel İlköğretim Okulunda öğrenim gören 175 7.sınıf öğrencisine uygulanmıştır. Araştırma sonucunda, kısa yanıtli testlerde ağırlıklı puanlama yönteminin kullanılmasının, puanlama cetveli ve anahtarı kullanıldığı

durumlarda madde ayırıcılık gücü indekslerini, test geçerliğini ve güvenilirliğini yükselttiği ortaya konmuştur.

Tosyalı (2002), aynı davranışı ölçen çoktan seçmeli, kısa yanıtlı, 0-1 puanlamalı ve 0-1-2 puanlamalı gerekçeli doğru-yanlış testlerin; madde ayırıcılık gücü indeksleri, test güvenilirlikleri ve test geçerlikleri arasındaki farklılıkları araştırmayı amaçlamıştır. Bu amaç doğrultusunda, araştırmacı tarafından geliştirilen sayısal yetenek testi, 2001-2002 öğretim yılında Lise1. ve 2. sınıf öğrencisi olan 151 öğrenciye uygulanmıştır. Sonuç olarak, en düşük madde ayırıcılık gücü indekslerinin 0-1 puanlamalı doğru-yanlış testinden elde edildiği, verilen önermenin doğruluğunun gerekçesi sorulduğunda ayırıcılığın arttığı, en yüksek güvenilirlik ve geçerliğin 0-1-2 puanlamalı gerekçeli doğru-yanlış testinden elde edildiği belirlenmiştir.

Akdeniz (2003), aynı davranışı ölçmeye yönelik olarak kullanılan kısa yanıtlı ve çoktan seçmeli testlerin madde ve test istatistikleri arasındaki farklılıkları araştırdığı çalışmasında, madde güçlük indeksleri, madde ayırıcılık gücü indeksleri, test güvenilirlik ve geçerlik katsayıları ele alınmıştır. Araştırma verileri, 10. sınıf Fizik dersi Vektörler ünitesine ilişkin başarı testinin Ankara'daki dört lisede toplam 427 öğrenciye uygulanmasıyla elde edilmiştir. Sonuçta, her iki testin güçlük indeksleri arasında manidar fark olmadığı, kısa yanıtlı testin ayırıcılığının daha yüksek olduğu, ölçüt dayanaklı geçerlik katsayılarının ve güvenilirlik katsayılarının birbirine yakın olduğu ortaya konmuştur.

Karaca (2004), seçme gerektiren, kısa cevaplı ve doğru-yanlış türü test maddelerinin aynı davranışları ölçme yönünden etkililiklerinin ortaya konulabilmesi için madde güçlükleri, ayırıcılık gücü indeksleri, test varyansları ve KR-20 güvenilirlikleri arasında anlamlı farklılık olup olmadığını araştırmıştır. Bu farklılıkların ortaya konulabilmesi için, seçme gerektiren, kısa cevaplı ve doğru-yanlış tipi maddelerden oluşturulan üç test hazırlanmış ve 74 yedinci ve sekizinci sınıf öğrencisine uygulanmıştır. Verilerin analizinde varyans analizi, Scheffé testinden yararlanılmıştır. Verilerin analizi sonunda; kısa cevaplı teste ait ortalama madde güçlüğü, doğru-yanlış ve seçme gerektiren teste ait ortalama madde güçlüklerinden daha küçük olduğu;

doğru-yanlış testinin seçme gerektiren ve kısa cevaplı testten daha çok ayırıcı madde içerdiği; kısa cevaplı testin seçme gerektiren ve doğru-yanlış testten daha çok güvenilir olduğu görülmüştür.

Madde Tepki Kuramına Dayalı Çalışmalar

Traub ve Fisher (1977), aynı içerikte fakat farklı formatlardaki testlerin aynı özelliği ölçüp ölçmediğini araştırdıkları çalışmada, matematiksel sonuç çıkarma (mathematical reasoning) ve sözel kavrama (verbal comprehension) testlerinde yanıtı sınırlı açık uçlu, çoktan seçmeli ve “Coombs çoktan seçmeli” formatlarda maddelere yer vermiştir. Coombs çoktan seçmeli formatı, yanlış seçeneklerin elenme gerekçelerini belirtilerek elenmesi ve doğru seçeneğe ulaşılması süreci için dereceli puanlama gerektiren bir madde formatıdır. Araştırma verileri, testlerin 199 8. sınıf öğrencisine uygulanmasıyla elde edilmiştir. Sonuç olarak, farklı madde formatlarının aynı özelliği ölçmede kullanılabilirliği hipotezi, sadece sözel kavrama testinin açık uçlu maddeleri için reddedilmiştir. Sözel kavrama testinin çoktan seçmeli ve Coombs çoktan seçmeli maddelerinin aynı özelliği ölçtüğü, matematiksel sonuç çıkarma testinin üç madde formatı için de aynı özelliği ölçtüğü belirlenmiştir.

Donoghue (1993), çoklu puanlanan maddelerden sağlanan bilginin iki kategorili puanlanan maddelere göre nasıl değiştiğini incelemek amacıyla bir çalışma yapmıştır. Bu amaç doğrultusunda, araştırmada çoktan seçmeli, kısa yanıtlı ve uzun yanıtlı maddelerin bir arada kullanıldığı Ulusal Eğitim Gelişimini Değerlendirme (National Assessment of Educational Progress-NAEP) durum belirleme uygulamasının Okuma (Reading) bölümü 4., 8. ve 12. sınıf verileri kullanılmıştır. Sonuç olarak, uzun yanıtlı (dört kategorili çoklu puanlanan) maddelerde, ikili puanlanan maddelere göre daha fazla bilgi sağlandığı; üst yeterlik düzeyinde yer alan bireylerde, çoklu puanlanan maddelerden sağlanan bilginin daha da arttığı ortaya konmuştur. Uzun yanıtlı maddelerin iki kategorili puanlanması durumunda, elde edilen bilgi düzeyinde ciddi düşüş olduğu belirtilmiştir.

Lukhele ve diğerleri (1994), Kolejlere giriş için İleri Yerleştirme (Advanced Placement) sınavının çoktan seçmeli ve açık uçlu maddeler içeren Kimya ve Amerikan Tarihi sorularını ele almışlardır. Bu testler, bireylere açık uçlu maddelerin alt testlerini seçme imkanı sağlamaktadır. Araştırma bulguları, 1989'da İleri Yerleştirme Kimya sınavını alan 18.462 ve Amerikan Tarihi sınavını alan 82.842 öğrenciden elde edilmiştir. Sonuç olarak, testteki açık uçlu maddelerden, çoktan seçmeli bölümden elde edilen bilgiden az da olsa daha fazla bilgi elde edilmiştir.

Ercikan ve diğerleri (1998), farklı iki madde tipinin (açık uçlu ve çoktan seçmeli) birlikte ölçeklenip ölçeklenemeyeceği ve bu durumun bilgi kaybına yol açıp açmadığı sorularına yanıt aranmıştır. Üç sınıf düzeyinde (3, 5 ve 8) ve dört alanda (okuma, dil, matematik ve fen) alt testler hazırlanmıştır. Bu çalışmayla, karma madde tiplerinin kullanıldığı testlerin puanların güvenirliğini ve geçerliğini artırdığını ortaya koymuşlardır. Öte yandan, iki madde tipi birlikte kullanıldığı karma testlerde test uzunluğu arttığında genel ölçme duyarlılığının da (overall measurement accuracy) arttığı bulunmuştur. Bunun yanı sıra, açık uçlu maddeler şans başarısı içermediğinden, çok düşük yeterlik grubundaki öğrencilerden de yüksek yeterlik grubundan olduğu kadar bilgi sağlanabildiğini vurgulamışlardır. Ayrıca, açık uçlu ve çoktan seçmeli maddelerin birarada ölçeklendiği durumdaki ölçek puanları ile ayrı ayrı ölçeklendiğindeki ölçek puanları arasındaki korelasyonların tümü yüksek bulunmuştur.

Wester ve Henriksson (2000), çoktan seçmeli ve açık uçlu madde formatlarının, matematik başarısının cinsiyete göre farklılaşmasına etkili olup olmadığını araştırmışlardır. Araştırma kapsamında, TIMSS 1995 Matematik testinde yer alan 10 çoktan seçmeli ve 10 açık uçlu madde seçilmiş; bu maddelere ilişkin madde güçlük değerleri kız ve erkek öğrenciler için ayrı ayrı hesaplanmıştır. Problem çözme sıklığı (problem solving frequency) için bir gösterge olarak ele alınan madde güçlük değerleri arasındaki farklar madde bazında incelenmiştir. Sonuçta, kız öğrencilerin her iki madde formatında erkek öğrencilere göre az da olsa başarılı oldukları; çoktan seçmeli maddelere ilişkin madde güçlük değeri varyansının, açık uçlu maddelere göre daha fazla olduğu belirlenmiştir.

Garavaglia (2001), çalışmasında bireylerin çoktan seçmeli ve açık uçlu maddeleri yanıtlarken kullandıkları bilişsel süreçlerin farklılaşp farklılaşmadığını ortaya koymayı amaçlamıştır. Çoktan seçmeli ve açık uçlu formatta maddeler içeren 8. sınıf Matematik-Cebir testi, 34 öğrenciye uygulanmıştır. Öğrencilerin maddeleri yanıtlarken kullandıkları bilişsel süreçler sesli düşünme (think-aloud) yöntemiyle tespit edilmiştir. Sonuçta, öğrencilerin çoktan seçmeli ve açık uçlu maddeleri yanıtlarken kullandıkları bilişsel süreçlerin benzer olduğu ortaya konmuştur.

Sykes ve diğerleri (2001), karma testlerde açık uçlu maddelerin puanlamasına ilişkin MTK'ye dayalı üç yöntemin, test puanlarının güvenilirliğini artırmaya etkisi olup olmadığı araştırmak üzere bir çalışma yapmışlardır. Bu doğrultuda, "ağırlıklandırılmamış (baseline)", "yanıtı sınırlı açık uçlu (constructed response-CR) maddeleri iki kat ağırlıklandırma (CRx2)" ve "uzun yanıtli (extended response-ER) maddeleri iki kat ağırlıklandırma (ERx2)" yöntemlerine göre test formları oluşturulmuştur. Karma test formatında geliştirilen 45 maddelik 5. sınıf Matematik testi 2385, 45 maddelik 8. sınıf Matematik testi 2748, 34 maddelik 3. sınıf Yazma testi 2466, 33 maddelik 8. sınıf Yazma testi 3288 öğrenciye uygulanmıştır. Sonuç olarak, Baseline formların güvenirligi (stratified alfa), CRx2 formlarından az da olsa yüksek bulunmuştur. bu farkın en yüksek olduğu test 5.sınıf Matematik testidir. Genel olarak Baseline formları ile ERx2 formlarının güvenirlikleri hemen hemen aynı düzeyde bulunmuş; ancak 8. sınıf yazma testinde ERx2 formunun güvenirligi Baselin formdan yüksek bulunmuştur. Ayrıca, CRx2 ve ERx2 formlarında kestirilen yetenek parametrelerine ilişkin standart hata miktarlarının Baseline formdan yüksek olduğu; en yüksek bilginin ise yanıtı sınırlı açık uçlu Matematik maddelerinden elde edildiği belirtilmiştir.

Arthur ve diğerleri (2002), geleneksel çoktan seçmeli teste alternatif olarak yazma/işaretleme kağıt-kalem testinde yenilikçi açık uçlu test formatına (innovative constructed response) ilişkin alt gruplardaki performans farklılıklarını incelemişlerdir. Araştırma verisi, Afrikalı Amerikalı ve Beyaz itfaiyeci örnekleme uygulanan uygulamalı terfi sınavından (operational promotional exam) elde edilmiştir. Sonuç olarak, alt gruplardaki performans

farklılığının açık uçlu maddelerde çoktan seçmeli maddelere göre daha az olduğu ortaya konmuştur.

Kinsey (2003), test uzunluğu (10, 20 ve 30 madde), puanlama türü (iki kategorili puanlama, çoklu puanlama) değişkenlerinin yetenek kestirimi, test bilgi düzeyi, karma maddeli testlerinin optimizasyon kriteri üzerine etkisini araştırmıştır. Araştırmada ele alınan değişkenlerin farklı kombinasyonlarında 33 karma test için 1000 kişilik simülatif data kullanılmıştır. Sonuç olarak, test uzunluğu ve çoklu puanlanan madde oranlarının her ikisinin de karma testten elde edilen bilgi üzerinde anlamlı etkiye sahip olduğu belirtilmiştir. Genel olarak maddelerin tamamı çoklu puanlandığında en yüksek test bilgi düzeyi elde edilmiştir. Bu durum özellikle düşük yetenek düzeyindeki gruplar için gözlenmiştir.

Rodriguez (2003), çoktan seçmeli ve açık uçlu madde formatlarının yapı eşdeğerliğini (construct equivalence) araştıran deneysel çalışmaları sentezlemek amacıyla bir meta analizi çalışması yapmıştır. İlgili alan yazın taramasıyla çoktan seçmeli ve açık uçlu maddelerin “yapı eşdeğerliği” şu üç kavram çerçevesinde ele alınmıştır. İlki, açık uçlu maddelerle çoktan seçmeli maddelerin aynı soru kökünü kullanması olarak tanımlanan madde kökü eşdeğerliği (stem-equivalent), ikincisi iki formattaki maddelerin aynı kapsam ve bilişsel alanla ilişkin olması olarak tanımlanan kapsam eşdeğerliği (content-equivalent), üçüncüsü ise, açık uçlu maddelerin çoktan seçmeli maddelerden nitelik bakımından farklılaşması olarak tanımlanan açık uçlu maddelerin kapsam ya da bilişsel alanın farklı bir durumuna ilişkin yazılmış olması durumu (not content-equivalent). Bu doğrultuda, 67 deneysel çalışma incelenmiş ve 29 korelasyonel çalışma araştırma kapsamına alınmıştır. Sonuç olarak, maddeler her iki formatta da aynı kök kullanılarak oluşturulduğunda, iki formatın kullanıldığı testlerden elde edilen puanlar arasındaki ortalama korelasyonlar birbirine yakın ve farklı madde kökü kullanıldığı (non-stem-equivalent) duruma göre daha yüksek bulunmuştur.

Özdemir (2004), çoktan seçmeli test maddelerini iki kategorili (1,0) ve ağırlıklı (1,2,3,4) puanlama yöntemlerinin testin güvenirlik ve geçerliğine etkisi, klasik test teorisi ve madde tepki kuramına göre incelemiştir. Araştırma

verileri, her biri dört seçenekli 20 maddelik çoktan seçmeli bir testin 1608 (4.-7. sınıf) öğrenciye uygulanmasından elde edilmiştir. Sonuç olarak, test geliştirme çalışmalarında madde tepki kuramından ve (1,0) puanlamadan yararlanılmasının daha uygun olacağı, klasik test kuramına göre yapılacak uygulamalarda ise ağırlıklı puanlamadan yararlanılmasının uygun olacağı belirtilmiştir.

Uyeno (2004), bireylerin yeterlik düzeylerine göre sınıflanmasında yanıtı sınırlı açık uçlu (open ended), uzun yanıtı (extended response) ve çoklu yanıt gerektiren (multiple response) maddelerden sağlanan bilginin tekliğini (uniqueness of information) araştırmıştır. Araştırmanın verileri, 2002 yılında 10620 8. sınıf öğrencisinin ve 9068 10.sınıf öğrencisine uygulanan Hawai Durum Belirleme (Hawai'i State Assessment) Testi'nin Okuma (Reading) bölümüne verilen yanıtlardan elde edilmiştir. Test alan bireylerin sınıflandığı yeterlik düzeyleri “iyi altı (well below proficiency)”, “yaklaşık yeterli (approaches proficiency)”, “yeterli (meets proficiency)” ve “üst düzeyde yeterli (exceeds proficiency)” olarak tanımlanmıştır. Sonuç olarak, her iki sınıf düzeyinde de çoktan seçmeli, yanıtı sınırlı açık uçlu ve uzun yanıtı maddelerden, öğrencileri dört yeterlik düzeyine yerleştirmede sağlanan bilgi yeterli bulunmuştur. Buna karşın, çoktan seçmeli maddelerden sağlanan bilgi ile her iki açık uçlu formattan sağlanan bilgi arasında düşük korelasyon bulunmuştur.

In'nami ve Koizumi (2009), okuma ve dinleme testlerinde, test formatının test başarısına etkisini araştıran çalışmaları meta analitik yaklaşımla taramışlardır. Test formatının başarı üzerindeki etkisinde, ortalama etki büyüklüğü değerlerini ölçüt alarak yürüttükleri literatür taraması sonucunda 56 araştırmaya ulaşılmıştır. Sonuçta, çoktan seçmeli maddelerden oluşan testlerin, açık uçlu testlere göre daha kolay olduğu belirlenmiştir. Güçlük düzeyleri arasındaki bu farkın dinleme testlerinde daha fazla olduğu belirtilmiştir.

Kim (2009), bireylerin yeteneklerini kestirmede, çoktan seçmeli ve açık uçlu maddelerin puanlarının Genişletilmiş Hiyerarşik Değerlendirici Modeli (Extended Hierarchical Rater Model-HRM) ile birleştirilme prosedürünü

incelemiştir. Genişletilmiş HRM, değerlendirici davranışının (rater behaviour) kullanımına uygun olan Örtük Sınıf İşaret Tarama Teorisi (Latent Class Signal Detection Theory-SDT) modeli ile bireyin yeteneklerinin kestirildiği MTK modellerini içine alan bir yaklaşımdır. Bu araştırmada, HRM_{SDT} yaklaşımının yetenek kestirim performansı hem simülatif veri hem de gerçek veri üzerinde sınanmıştır. Karma test ve sadece açık uçlu maddelerden oluşan test için 2970 adet veri üretilmiştir. Gerçek veri ise, 38 çoktan seçmeli ve bir tane 30 dakikalık denemeden (essay) oluşan okuma, matematik ve yazma temel beceri ölçeğinin 9336 öğrenciye uygulanmasıyla elde edilmiştir. Sonuç olarak, açık uçlu ve çoktan seçmeli maddeler birlikte kullanıldığında, sadece açık uçlu maddelerden oluşan teste göre daha iyi yetenek kestirimi yapıldığı belirtilmiştir. Ayrıca, HRM yaklaşımının, yetenek kestirimi için çoktan seçmeli ve açık uçlu maddelerin puanlarının birleştirilmesinde, yaygın olarak kullanılan ağırlıklandırma şemaları (weighting schemes) yaklaşımlarından daha başarılı olduğu ortaya konmuştur.

Yukarıda iki başlık altında verilen araştırmalar incelendiğinde, Klasik Test Kuramına dayalı olarak yapılan çalışmalarda, genel olarak çoktan seçmeli ve kısa yanıtlı testlerin ve farklı ağırlıklandırma yöntemlerinin test ve madde istatistiklerine etkisi araştırıldığı; buna karşın Madde Tepki Kuramına dayalı çalışmalarda, farklı madde tiplerinin aynı testte kullanıldığı karma test yapısının da ele alındığı dikkat çekmektedir. MTK analizlerinde ayrıca, test ve madde parametrelerinin yanı sıra testten elde edilen bilgi düzeyine ilişkin kanıtlar aranmıştır.

Klasik Test Kuramına dayalı olarak yapılan araştırmaların sonuçları bir arada ele alındığında; çoktan seçmeli testin daha kolay; ancak daha ayırt edici olduğu, en güvenilir testin çoktan seçmeli test olduğu, en düşük ayıricılığın ve güvenilirliğin kısa yanıtlı testten elde edildiği, kısa yanıtlı testlerde ağırlıklı puanlama yönteminin kullanılmasının, puanlama cetveli ve anahtarı kullanıldığı durumlarda madde ayıricılık gücü indekslerini, test geçerliğini ve güvenilirliğini yükselttiği bulgularının yanı sıra farklı madde tiplerinin ve ağırlıklandırma yöntemlerinin test ve madde istatistikleri ile güvenilirlik ve geçerlik indekslerinde farklılığa yol açmadığının ortaya koyan çalışmalar yapılmıştır.

Öte yandan, Madde Tepki Kuramına dayalı olarak yapılan çalışmalarda, açık uçlu maddelerden sağlanan bilginin çoktan seçmeli maddelere göre daha yüksek olduğu; açık uçlu maddeler şans başarısı içermediğinden düşük yetenek düzeyindeki bireylerde de ölçülen özelliğe ilişkin yüksek bilgi sağlanabildiği belirlenmiştir. Madde sayısı ve açık uçlu madde oranının artmasının testten sağlanan bilgi, güvenirlik kestirimi ve ölçme duyarlılığını artırdığı bulgusuna ulaşılmıştır.

Yukarıda değinilen ilgili alanyazında farklı madde formatlarının birbirine üstünlüğünün araştırıldığı ve son yıllarda farklı madde formatlarının bir arada kullanıldığı karma test uygulamalarında testlerin psikometrik niteliklerinin farklılaşıp farklılaşmadığı araştırmalara konu olmaktadır. Farklı madde formatlarının birbirlerine olan üstünlüğüne dair pek çok araştırma olmasına karşın; özellikle son yıllarda dünyada sıklıkla geniş ölçekli test uygulamalarında yer bulmaya başlayan karma testlerin sadece tek madde formatından oluşan testlere karşı üstün olup olmadığına ilişkin yeterli bulgu bulunmaması tek madde formatının (çoktan seçmeli) eğitsel ve ölçmelerin geçerlik ve güvenirliği boyutunda sahip olduğu sınırlılıklar bu araştırmanın yapılmasına yol açan bir problem olarak görülmektedir. Karma testlerin madde ve yetenek parametrelerinin, madde ve test bilgi fonksiyonlarının tek soru tipi olarak çoktan seçmeli veya açık uçlu maddelerden oluşan testlerden farklılık gösterip göstermediğinin araştırılmasına ihtiyaç duyulmaktadır.

Amaç

Bu çalışmanın amacı, Matematik başarısına ilişkin olarak benzer yapıları ölçen iki madde tipinin (çoktan seçmeli ve yanıtı sınırlı açık uçlu) ayrı (tek başına) ve birlikte kullanıldıkları koşullar ile iki madde tipinin test içinde yer alma yüzdelerinin test formlarında farklı olduğu durumlarda MTK'ye göre kestirilen madde ve yetenek parametrelerinin psikometrik niteliklerinin, madde ve test bilgi fonksiyonlarının (test information function), göreceli etkinlik (relative efficiency) düzeylerinin ve ölçüt dayanaklı geçerlik kanıtlarının farklı olup olmadığını araştırmaktır. Bu amaç çerçevesinde;

Madde sayısının 25 ve 15; açık uçlu madde yüzdesinin %40 ve %20 olduğu dört karma testten (KT1:25,%40; KT2:25,%20; KT3:15,%40; KT4:15,%20), açık uçlu (AUT) ve çoktan seçmeli (ÇST) testlerden kestirilen;

1. madde parametreleri (b, a, c) hangi aralıkta değişmektedir?
2. yetenek parametreleri ve standart hata değerleri hangi aralıkta değişmektedir?
3. madde ve test bilgi fonksiyonları ile görelî etkinlik (relative efficiency) indeksleri hangi düzeydedir?
4. yetenek parametreleri arasındaki ilişki miktarı nedir?
5. yetenek parametreleri ile TIMSS 2007 genel matematik başarı puanları (overall mathematics achievement score) arasındaki ilişki miktarı nedir?

sorularına yanıt aranmıştır.

Önem

Dünyada geniş ölçekli testlerde (TOEFL, AP, PISA, TIMSS, SAT vb.) de soru türlerinde görülen çeşitlilik ulusal sınavlarda tercih edilmemektedir. Bu araştırmayla çoktan seçmeli ve yanıtı sınırlı açık uçlu madde tiplerinin birlikte kullanılmasının test ve madde parametreleri ile yetenek kestiriminde farklılıklara yol açıp açmadığı ortaya konmaya çalışılmıştır. Araştırmanın bulgularının ulusal düzeydeki merkezi sistem sınavlarının uygulamasıyla ilgili gelişmelere ışık tutacağı düşünülmektedir. Nitekim, ölçme araçlarında sadece çoktan seçmeli soru formatının kullanılmasının eğitim sistemi üzerinde yarattığı olumsuz etkiler sadece eğitimciler arasında değil karar vericiler düzeyinde de tartışılmaya başlanmıştır. Berberoğlu (2009), Öğrenci İzleme Sistemi (ÖİS) kapsamında öğrencilerin akademik başarısını etkileyebilecek diğer duyuşsal ve öğretimsel özelliklerin yoklandığı Öğrenci

Sosyal Gelişim Programı çerçevesinde uygulanan anketlerle, öğrencilerin sınıf düzeyi arttıkça daha çok ezberleme stratejilerini kullandıklarını ortaya koymuştur. Bu durum, öğretmenlerin sınıfta “test çözme”yi bir öğretim etkinliği gibi kullanma ve bu türden testlerin ev ödevi gibi verilmesi davranışının bir sonucudur. Öğrenciler, konuları kavramak yerine, soruların çözüm algoritmalarını ezberlemektedirler. Soru üzerinden eğitimin yanlış bir süreç olduğu tartışmaları son dönemde oldukça artmıştır. Karma test uygulamalarının yaygınlaşmasının, modern çağa uygun eğitim-öğretim ve ölçme-değerlendirme sistemlerine geçiş açısından önemli olacağı düşünülmektedir.

Bu türden geniş ölçekli test uygulamalarında maddelerin MTK'ye göre ölçeklenmesinin çeşitli avantajları bulunmaktadır. Klasik Test Kuramında, farklı madde tiplerinin bir arada kullanımı, toplam puan üzerinde maddelerin denk olmayan ağırlıklandırmalarını gerektirmektedir. Bir diğer yaklaşım ise her bir madde için z puanı hesaplamak ve z puanlarının toplamından toplam puan elde etme yaklaşımıdır. Bu yaklaşımın dezavantajı, belirli bir örnekleme ait madde ortalama ve standart sapmasına bağlı olmasıdır. Farklı bir örneklemede başka sonuçlar vermektedir. Bu durum, modele uyum gösteren test verileri elde edildiğinde Madde Tepki Kuramının değişmezlik (invariant) özelliğiyle ortadan kalkmaktadır. Günümüzde, karışık madde formatlarının toplam test puanı üzerinde dengesiz etkilere sahip olma durumu; karma maddelerin optimal test puanını sağladığı kuralıyla değişmiştir (Embretson ve Reise, 2000). MTK modellerinin varsayımları karşılandığı takdirde, testlerin yetenek ve madde parametrelerini gruptan bağımsız olarak kestirilmesi; paralel testlerin geliştirilmesi, test formları veya bir testin farklı dönemlerdeki uygulamaları arasında puanların eşitlenmesi ve bu yolla puanların ortak bir ölçekte karşılaştırılması olanaklı hale gelebilmektedir.

Sınırlılıklar

Bu araştırma, TIMSS 2007 8. sınıflar Matematik Testi, 2. kitapçığında yer alan maddelerle sınırlı tutulmuştur.

TIMSS uygulamasında kullanılan soruların tamamı yayınlanmamaktadır. TIMSS resmi web sitesi (www.timss.bc.edu) üzerinden yayınlanan sorular (released items) incelendiğinde, TIMSS 2007 8. sınıflar Matematik testi 2.kitapçığındaki soruların tamamının yayınlandığı belirlenmiş ve bu nedenle 2. kitapçıktaki maddeler ve bu maddelere verilen yanıtlar araştırma kapsamına alınmıştır.

Araştırma kapsamında TIMSS 2007 Matematik Testi 2 nolu kitapçığındaki iki kategorili puanlanan açık uçlu maddeler ele alınmıştır. Testin 2. kitapçığında yalnız bir madde çoklu puanlanmaktadır. Kısmi puanlamanın testin psikometrik niteliklerine etkisini ortaya çıkarma için yalnızca bir madde yeterli olmayacağından kısmi puanlama araştırma kapsamı dışında tutulmuştur.

Tanımlar

Karma Test: Farklı yüzdelerde açık uçlu ve çoktan seçmeli maddeler içeren test

Kısaltmalar

MTK: Madde Tepki Kuramı

KTK: Klasik Test Kuramı

KT1: Madde sayısı 25, açık uçlu madde yüzdesi %40 olan test

KT2: Madde sayısı 25, açık uçlu madde yüzdesi %20 olan test

KT3: Madde sayısı 15, açık uçlu madde yüzdesi %40 olan test

KT4: Madde sayısı 15, açık uçlu madde yüzdesi %20 olan test

AUT: Tamamı açık uçlu maddelerden oluşan test

ÇST: Tamamı çoktan seçmeli maddelerden oluşan test

1PL: 1 parametrelili lojistik model

2PL: 2 parametrelili lojistik model

3PL: 3 parametrelili lojistik model

AU: Açık uçlu

BÖLÜM 2

YÖNTEM

Araştırma Modeli

Bu araştırmada, çoktan seçmeli ve açık uçlu madde formatlarının Madde Tepki Kuramına dayalı olarak birlikte ölçeklenmesinin madde ve yetenek parametrelerinin psikometrik özelliklerinde fark yaratıp yaratmadığı araştırılmıştır. Bu yönüyle çalışma, kuramsal ve temel araştırma niteliğindedir.

Araştırma Grubu

Araştırmanın verileri, 2007 yılında uygulanan Uluslararası Fen ve Matematik Eğilimleri Araştırması (Trends in International Mathematics and Science Study – TIMSS 2007) Türkiye’den katılan ve matematik testinin 2. kitapçığını alan 320, 8.sınıf öğrencisinden elde edilmiştir.

Uluslararası Eğitim Başarısını Değerlendirme Kurumu (IEA-International Association for the Evaluation of Educational Achievement), eğitim alanında ölçme ve değerlendirme çalışmaları yapan bir kuruluştur. Bu kuruluş tarafından dünya çapında uygulanan sınavlardan biri de TIMSS (Trends in International Mathematics and Science Study) dir. TIMSS, dört yılda bir uygulanan, 4. ve 8. sınıf düzeylerinde öğrencilerin matematik ve fen başarılarını ölçen uluslararası bir karşılaştırma çalışmasıdır.

TIMSS 2007 uygulamasında yer alan ülkelerdeki katılımcılar, tek biçimli (uniform) örnekleme yaklaşımıyla örnekleme minimum sapmayla dahil olmuştur. Bu yöntem bir kalite standardı sağlayarak, araştırma sonuçlarındaki ülkeler arası farklılıkların örnekleme yönteminden kaynaklanması ihtimalini ortadan kaldırmaktadır. Katılımcı her ülkenin TIMSS 2007 Ulusal Araştırma Koordinatörü (National Research Coordinator), TIMSS & PIRLS Uluslararası Çalışma Merkezi (International Study Center) tarafından onaylanan

örnekleme prosedürünün her adımını uygulama ve rapor etme sorumluluğunu üstlenmiştir. Ulusal Araştırma Koordinatörleri *Windows Within-school Sampling Software* yazılımını kullanarak örnekleme dahil olacak öğrencileri belirlemektedir.

TIMSS 2007 uygulaması, dünya çapında 49 ülke ve 7 kıyaslama (benchmarking) katılımcısı ile gerçekleşmiştir. Bu ülkeler ve öğrenci sayıları Çizelge 1’de verilmiştir.

Çizelge 1: TIMSS 2007 8. Sınıflar Matematik Testi Katılımcı Ülkeler

Ülkeler	Populasyon		Örneklem	
	Okullar	Öğrenciler	Okullar	Öğrenciler
Avustralya	2417	270116	228	4069
Bahreyn	74	11667	74	4230
Birleşik Devletler	46112	4219262	239	7377
Bosna Hersek	569	45579	150	4220
Botsvana Cum.	214	40115	150	4208
Bulgaristan	2309	78729	163	4019
Cezayir	3891	624353	149	5447
Çek Cumhuriyeti	2669	124325	147	4845
El Salvador	2626	109671	145	4063
Endonezya	29701	2799024	149	4203
Ermenistan	1332	55289	148	4689
Fas	1636	368656	131	3060
Filistin	1130	94376	148	4378
Gana	7589	346289	163	5294
Gürcistan	2059	46061	135	4178
Hong Kong	455	83267	120	3470
İngiltere	3886	636732	137	4025
İran	29956	1475368	208	3981
İskoçya	418	64812	129	4070
İsrail	805	97132	146	3294
İsveç	1531	125478	159	5215
İtalya	5824	602185	170	4408
Japonya	10708	1201082	146	4312
Katar	67	7332	66	7184
Kıbrıs	67	9500	67	4399
Kolombiya	10034	648634	148	4873
Kore	2727	696156	150	4240
Kuveyt	163	23827	158	4091
Litvanya	1021	49887	142	3991
Lübnan	1574	63755	136	3786
Macaristan	2968	118049	144	4111
Malezya	1930	429048	150	4466

Çizelge 1'in devamı...

Ülkeler	Populasyon		Örneklem	
	Okullar	Öğrenciler	Okullar	Öğrenciler
Malta	60	5260	59	4670
Mısır	8179	1342127	233	6582
Norveç	1070	62348	139	4627
Romanya	6099	251054	149	4198
Rusya Federasyonu	42188	2140032	210	4472
Sırbistan	1310	81275	147	4045
Singapur	164	50904	164	4599
Sloveenya	428	19138	148	4043
Suriye	3756	270389	150	4650
Suudi Arabistan	6271	332479	165	4243
Tayland	9481	844336	150	5412
Tayvan	888	316997	150	4046
Tunus	804	176555	150	4080
Türkiye	16112	1163830	146	4498
Ukrayna	12184	479467	146	4424
Umman	722	56569	146	4752
Ürdün	1691	108856	200	5251
Kıyaslama (Benchmarking) Katılımcıları				
British Kolombiya Kanada	433	51804	150	4256
Dubai UAE	116	11178	88	3195
İspanya Bask Bölgesi	330	17203	130	2296
Massachusetts US	468	75805	48	1897
Minnesota US	656	64566	49	1777
Ontario Kanada	2854	159230	176	3448
Quebec Kanada	605	102112	170	3956

Örnekleme, her ülkede olduğu gibi Türkiye’de de coğrafi bölge ve okul türü örnekleme tabakası olarak alınmıştır. Türkiye uygulamasında örneklem yedi coğrafi bölgede ve iki okul türüne (devlet ve özel) göre tabakalı olarak her bölgeden oranı ölçüsünde okul; her okuldan da seçkisiz (random) bir sınıf seçilecek şekilde belirlenmiştir. Bu doğrultuda Türkiye’de uygulamaya 146 okuldan bu yolla seçilmiş 4498, 8. sınıf öğrencisi katılmıştır (IEA 2008-1).

Araştırma kapsamında yayımlanan maddelerin bulunduğu 2. kitapçık ele alındığından dolayı, araştırmanın amacına yönelik yapılan analizlerde 2. kitapçığı yanıtlayan 320 öğrencinin yanıtlarından elde edilen veriler kullanılmıştır.

Veri Toplama Araç ve Teknikleri

TIMSS verileri ile öğrencilerin matematik ve fen bilimleri alanındaki başarı durumları değerlendirilip; öğretim programları, öğretmen ve okulların özellikleri, öğrenci özellikleri ve eğitim sistemi hakkında da bilgi sahibi olunmaktadır. Bu sayede ülkeler eğitim sistemlerini sorgulayabilir ve mevcut eksiklikleri belirleyerek; gereken düzeltme ve yenilikler yapılabilir (MEB, 2011). Öte yandan TIMSS çalışmaları, örneklemin genişliği, öğrencilerin başarı düzeyleri ve öğrenci özellikleri açısından pek çok değişkeni içermekte ve araştırmacılara büyük bir veri tabanı oluşturmaktadır. Bu nedenlerle araştırma verileri TIMSS 2007 8. sınıflar Matematik testinden elde edilmiştir.

TIMSS, Uluslararası Eğitim Başarısı Değerlendirme Kuruluşu (International Association for the Evaluation of Educational Achievement – IEA) tarafından, katılımcı OECD ülkelerinin dört yılda bir 4. ve 8. sınıf düzeylerinde uygulanan, öğrencilerin matematik ve fen başarılarını ölçmeyi amaçlayan uluslararası bir karşılaştırma sınavıdır. Bu uygulamaların dördüncü ve sonuncusu olan TIMSS 2007, katılımcı ülkelerin matematik ve fen öğretim programlarının içeriği, nitelik ve niceliğiyle ilgili olarak katılımcı ülkelerin eğitim politikalarına bilgi sağlama hedefindedir. 49 ülkenin katıldığı uygulamaya Türkiye sadece 8. sınıf düzeyinde 146 okulda toplam 4498 öğrenciyle katılmıştır. 4. sınıf düzeyinde katılım olmamıştır (IEA, 2008-1).

TIMSS 2007’de öğrenciler, öğretmenler, okul yöneticileri ve Ulusal Araştırma Koordinatörü tarafından yanıtlanan anketler yer almaktadır. Öğrenci anketinde, yaş, cinsiyet, ev ve okul ortamı, evdeki kitap sayısı, anne-babanın eğitim düzeyi, matematik ve fen derslerindeki özgüven, ev ödevi yapma sıklığı ve süresi; Öğretmen anketinde, yaş, cinsiyet, eğitim düzeyi, kıdem, derse hazırlık, mesleki gelişim, ev ödevi, öğretim etkinliklerinde ve değerlendirmede etkililiğine ilişkin algılar; Okul anketinde, okul mevcudu, organizasyon, personel, okul ortamı değişkenlerine ilişkin görüşler ölçülmektedir. Ulusal Araştırma Koordinatörü tarafından yanıtlanan müfredat anketinde ise katılımcı ülkelerin matematik ve fen programlarına ilişkin görüşlerinin yoklandığı sorular yer almaktadır (IEA, 2007).

TIMSS matematik değerlendirmeleri iki boyutta tasarlanmıştır: (i) öğrencilerin öğrenmesi beklenen konu ya da kapsam ve (ii) öğrencilerin göstermesi beklenen bilişsel beceriler. Bu doğrultuda matematik testi, 8. sınıf düzeyinde Sayılar, Cebir, Geometri, Veri ve Olasılık konuları kapsamında (content domain); *bilgi, uygulama ve akıl yürütme* bilişsel alanlarını (cognitive domain) içermektedir. *Bilgi*, öğrencilerin matematik olguları, kavramları, araçları ve yöntemlerine dayalı bilgisini tanımlamaktadır (Ek1: M042003, M042301A, ...). *Uygulama*, öğrencinin problem durumundaki kavramsal algılama ve bilgiyi uygulamadaki yeteneğine odaklanır (Ek1: M042055, M042018, ...). *Akıl yürütme* ise, benzer olmayan durumları içeren ve çok aşamalı problemler gibi rutin problem çözmenin ötesine geçmek olarak tanımlanmaktadır (Ek1: M042265, M042301B, ...). TIMSS uluslararası karşılaştırma ölçütleri (benchmark) kullanarak öğrencilerin yeterliklerini yorumlayabilmektedir. Yeterlik düzeyleri “İleri 625, Yüksek 550, Orta 475, Düşük 400” olarak tanımlanmıştır. Ülkelerin yeterlik düzeyleri, matematik ve fen alanında hem bilişsel alanda hem de konu kapsamında ayrı ayrı rapor edilmiştir. TIMSS puanları ortalaması 500, standart sapması 100 olacak şekilde, 0-1000 arasında standardize edilmiştir. Türkiye 8. sınıf düzeyinde 49 ülke arasında 432 ortalamayla 30. sırada yer almıştır. Ayrıca, öğrencilerin % 5'i ileri düzey, %15'i üst düzey, % 33'ü orta düzey, %59'u düşük düzeye erişebilmiştir. (IEA, 2005 ve 2008-2, Gonzales, 2008).

MTK'ye dayalı olarak geliştirilen ve puanlanan TIMSS 2007 8. sınıflar matematik testinde toplamda 214 maddeden oluşan bir soru havuzu bulunmakla birlikte, uygulamada her öğrenci her birinin güçlük düzeyi denk 28-33 sayıda madde bulunan 14 farklı kitapçıktan birini yanıtlamıştır. Her iki kitapçıkta sayısı 12-19 arasında değişen ortak madde (anchor item) bulunmaktadır. Ortak maddeler 14 farklı formun Madde Tepki Kuramına dayalı olarak diğer kitapçıkları almamış öğrencileri o maddeleri de yanıtlamaları durumunda yeterlik kestirimi yapmak amacıyla kullanılmıştır. Madde Tepki Kuramına dayalı olarak eşitleme (equating) yoluyla test kitapçıklarının güçlük açısından denkliği sağlanmış ve bu yolla öğrencinin almadığı testteki yetenek düzeyinin kestirilebilmesi mümkün olmuştur. TIMSS 2007 8. sınıflar Matematik Testi bu çalışmada kullanılan 2. Numaralı

kitapçıkta ve geri kalan 13 kitapçıkta yer alan maddelerin madde tipine göre dağılımları Çizelge 2'de verilmiştir.

Çizelge 2: TIMSS 2007 8. Sınıflar Matematik Testindeki 14 Kitapçıkta Yer Alan Maddelerin Madde Tipine Göre Dağılımları

Kitapçık	Çoktan Seçmeli		Açık Uçlu				Toplam Madde Sayısı
	Madde Sayısı	Yüzde	1-0 Puanlanan	Kısmi Puanlanan	Toplam		
					Madde Sayısı	Yüzde	
1	16	55,17	9	4	13	44,83	29
2	20	64,52	10	1	11	35,48	31
3	21	65,63	8	3	11	34,38	32
4	14	48,28	8	7	15	51,72	29
5	12	41,38	11	6	17	58,62	29
6	17	53,13	13	2	15	46,88	32
7	16	48,48	17	0	17	51,52	33
8	17	53,13	14	1	15	46,88	32
9	18	58,06	12	1	13	41,94	31
10	18	56,25	13	1	14	43,75	32
11	19	59,38	12	1	13	40,63	32
12	14	50,00	10	4	14	50,00	28
13	14	50,00	7	7	14	50,00	28
14	16	53,33	8	6	14	46,67	30
TESTLERİN TAMAMI	116	54,21	76	22	98	45,79	214

Çizelge 2 incelendiğinde, TIMSS 2007 8. sınıf matematik uygulamasındaki 14 kitapçığın 28-32 maddeden oluştuğu ve açık uçlu maddelerin kitapçıklarda yer alma yüzdelerinin yaklaşık %35 - %60 arasında değiştiği belirlenmiştir. Alt amaçlara yanıt aramak için oluşturulan çoktan seçmeli, açık uçlu ve farklı madde sayısı ve açık uçlu madde yüzdelerinde karma testler için seçilen 2. kitapçıkta çoktan seçmeli 20 (%64,52), açık uçlu 11 (35,48) madde olmak üzere toplam 31 madde bulunmaktadır.

TIMSS 2007 8. sınıf matematik uygulamasındaki 2. kitapçıkta yer alan maddelerin madde tipi ve bilişsel alanlarına göre dağılımları Çizelge 3'te verilmiştir.

Çizelge 3: TIMSS 2007 8. Sınıflar Matematik Testi 2. Kitapçıkta Yer Alan Maddelerin Madde Tipi ve Bilişsel Alanlara Göre Dağılımları

Madde Tipi	Açık Uçlu		Çoktan Seçmeli		Toplam	
	Madde Sayısı	Yüzde	Madde Sayısı	Yüzde	Madde Sayısı	Yüzde
Bilişsel Alan						
Bilgi	3	27,27	9	45,00	12	38,71
Uygulama	5	45,45	10	50,00	15	48,39
Akıl yürütme	3	27,27	1	5,00	4	12,90
Toplam	11	100,00	20	100,00	31	100,00

Çizelge 3 incelendiğinde, 320 öğrenci tarafından alınan TIMSS 2007 8. Sınıflar Matematik Testi 2. kitapçığında yer alan 31 maddenin 12'sinin (%38,71) Bilgi, 15'inin (%48,39) Uygulama, 4'ünün (%12,90) Akıl yürütme bilişsel alanında olduğu görülmektedir. İkinci kitapçıkta yer alan çoktan seçmeli maddeler iki kategorili (0-1) olarak puanlanmıştır. Öte yandan, açık uçlu maddelerden sadece biri (M042220) kısmi puanlamış (0-1-2), diğer 10 madde iki kategorili (0-1) olarak puanlanmıştır. İki uçlu puanlanan maddeleri araştırma kapsamında ele alma sınırlılığından dolayı kısmi puanlanan madde (M042220) iki kategorili puanlanarak teste dahil edilmiştir.

Araştırma kapsamında, TIMSS 2007 Matematik Testi 2. kitapçığındaki maddelerden oluşturulan karma testlerde test uzunluğu 15 ve 25 madde, açık uçlu madde yüzdesi %20 ve %40 ile sınırlı tutulmuştur. İlgili alanyazında, çok farklı sayıda test uzunluğu (10-20-30, 15-25-50, vb.) ve açık uçlu madde yüzdesi (%15-%25, %25-%50, vb.) ile araştırmalar yapılmıştır. Bu araştırmada karma testleri oluşturmada yararlanılan test uzunlukları ve açık uçlu madde yüzdeleri TIMSS 2007 Matematik Testi 2. kitapçığından, bilişsel düzeylere göre dağılımı denk olacak şekilde dört karma test elde etmeye uygun olacak şekilde belirlenmiştir.

Bu kitapçıkta yer alan 31 madde arasından, farklı madde sayısı ve açık uçlu madde yüzdelerinde oluşturulan dört karma teste, açık uçlu ve çoktan seçmeli testlere ilişkin bilgiler Çizelge 4'te verilmiştir. Seçilen maddelerle ilgili ayrıntılı bilgi Ek 1'de verilmiştir.

Çizelge 4: Karma Testler, Çoktan Seçmeli Test ve Açık Uçlu Testteki Madde Sayılarının Bilişsel Alan ve Madde Tipine Göre Dağılımları

Karma Test 1 (KT1) (k=25 ; AU Yüzdesi: %40)				
Bilişsel Alan	Açık Uçlu	Çoktan Seçmeli	Toplam	Bilişsel Alandaki Oranı
Bilgi	3	8	11	44
Uygulama	5	6	11	44
Akıl yürütme	2	1	3	12
Toplam madde sayısı	10	15	25	100
Karma Test 2 (KT2) (k=25 ; AU Yüzdesi: %20)				
Bilişsel Alan	Açık Uçlu	Çoktan Seçmeli	Toplam	Bilişsel Alandaki Oranı
Bilgi	2	9	11	44
Uygulama	1	10	11	44
Akıl yürütme	2	1	3	12
Toplam madde sayısı	5	20	25	100
Karma Test 3 (KT3) (k=15 ; AU Yüzdesi: %40)				
Bilişsel Alan	Açık Uçlu	Çoktan Seçmeli	Toplam	Bilişsel Alandaki Oranı
Bilgi	2	4	6	40
Uygulama	3	4	7	47
Akıl yürütme	1	1	2	13
Toplam madde sayısı	6	9	15	100
Karma Test 4 (KT4) (k=15 ; AU Yüzdesi: %20)				
Bilişsel Alan	Açık Uçlu	Çoktan Seçmeli	Toplam	Bilişsel Alandaki Oranı
Bilgi	1	5	6	40
Uygulama	1	6	7	47
Akıl yürütme	1	1	2	13
Toplam madde sayısı	3	12	15	100

k: madde sayısı , AU: açık uçlu madde

Çizelge 4 incelendiğinde, madde sayısı (k); 15 ve 25, açık uçlu madde yüzdesi; %20 ve %40 olacak şekilde seçilerek oluşturulan dört farklı karma test (KT1, KT2, KT3, KT4) ile Açık Uçlu ve Çoktan Seçmeli Testlerin içerdiği maddelerin bilgi, uygulama ve akıl yürütme bilişsel alanlarında dağılımının benzer olduğu belirlenmiştir.

Verilerin Analizi

Bu bölümde, MTK varsayımlarına ilişkin olarak yapılan analizler ve araştırma sorularına ilişkin analizler olmak üzere iki aşamada açıklanmıştır. Öncelikle MTK varsayımlarının karşılanıp karşılanmadığını test etmek amacıyla; tek boyutluluğu sınamada Faktör Analizi, model veri uyumunu sınamada -2loglikelihood uyum istatistiğinden yararlanılmıştır. Araştırmada tek boyutluluğun sağlanması ilgili kanıt, alanyazında sıkça rastlandığı gibi yerel bağımsızlığın da bir kanıtı olarak değerlendirilmiştir (Hambleton, Swaminathan ve Rogers, 1991, Embretson ve Reise, 2000).

İkinci aşamada ise araştırma sorularına ilişkin analizlere yer verilmiştir. Madde ve yetenek kestirimleri, yetenek kestirimine ilişkin standart hata, madde ve test bilgi fonksiyonu BILOG MG programında beklenen a posteriori (expected a posteriori) yöntemiyle elde edilmiştir. Beklenen a posteriori yöntemi, diğer kestirim yöntemlerinden (ençok olabilirlik, en yüksek posteriori) ayrı olarak tüm tepki örüntüleri için (tümü doğru - tümü yanlış) sonlu bir yetenek kestirimi ($\theta = -3.0 - 3.0$) sağladığı için tercih edilmiştir. Yetenek kestirimleri arasındaki ilişki, Pearson Momentler Çarpımı Korelasyon Katsayısı ile belirlenmiştir. Karma testlerden, açık uçlu ve çoktan seçmeli, testlerden elde edilen yetenek kestirimleri ile TIMSS 2007 testinden aldıkları Genel Matematik Başarı Puanları (Overall Math Achievement Score) arasındaki ilişki, Pearson Momentler Çarpımı Korelasyon Katsayısı ile belirlenmiştir.

Analizlere kaynaklık eden veriler, TIMSS & PIRLS Uluslararası Çalışma Merkezine ait internet sitesindeki uluslararası veri tabanından SPSS dosyaları şeklinde alınmıştır (http://timss.bc.edu/TIMSS2007/idb_ug.html). 8. sınıf verilerinin bulunduğu dosyadan Türkiye katılımcılarının matematik testine verdikleri yanıtlar alınarak araştırmada kullanılan veriler elde edilmiştir.

1. Ön Analizler – MTK Varsayımları

Tek Boyutluluk Testi:

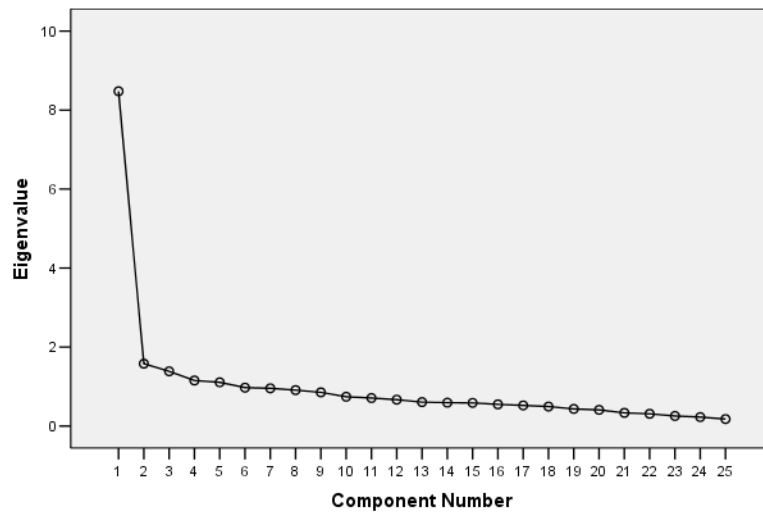
Oluşturulan farklı formatlardaki testlerin maddelerine verilen yanıtlardan oluşan verilerin öncelikle tek boyutluluk varsayımını karşılayıp karşılamadığına bakılmıştır. Tek boyutluluk, SPSS programında açımlayıcı faktör analizi yapılarak araştırılmıştır. Faktör Analizi, tek boyutluluk araştırmalarında yaygın olarak kullanılmaktadır. Bu yöntemde, maddelerin faktör yük değerleri, en büyük özdeğer tarafından açıklanan varyans yüzdesi incelenmiş ve özdeğer grafiğine bakılmıştır. Tek boyutluluğa karar vermede; faktör yük değerlerinin 0,30'dan yüksek olması, başat faktör tarafından açıklanan varyansın %30'a yakın ve/veya yüksek olması, öz değer grafiğinin incelenmesinde yüksek ivmeli, hızlı düşüşün yaşandığı faktörün bulunması ölçütlerine göre karar verilmiştir (Büyüköztürk, 2004). Karma Test 1-4, Açık Uçlu Test ve Çoktan Seçmeli Test için faktör analizi sonuçları Çizelge 5-10'da özdeğerler grafiği Şekil 2-7'de verilmiştir.

Çizelge 5: KT1 (k=25, AU Yüzdesi=%40) İçin Faktör Analizi Sonuçları

Madde No	Faktör Yük Değeri	Madde No	Faktör Yük Değeri
M042003	0,50	M042220	0,69
M042079	0,50	M022097	0,61
M042018	0,67	M022101	0,40
M042055	0,40	M022104	0,62
M042199	0,67	M022106	0,58
M042301A	0,48	M022110	0,57
M042301B	0,68	M022181	0,54
M042301C	0,73	M032307	0,64
M042265	0,43	M032701	0,51
M042137	0,60	M032525	0,63
M042148	0,53	M032579	0,66
M042254	0,38	M032691	0,70
M042250	0,60		

Açıklanan varyans: %33,91

Çizelge 5 incelendiğinde, KT1’de yer alan 25 maddenin birinci faktördeki faktör yükleri 0,38’in üzerindedir. Ayrıca birinci faktörün açıkladığı varyans %33,91 olarak bulunmuştur. Bu değerler, tek boyutluluğun sağlandığına kanıt olarak ele alınmakla birlikte, öz değer grafiği de incelenmiştir.



Şekil 2: KT1 (k=25, AU Yüzdesi =%40) İçin Özdeğerler Grafiği

KT1'e ilişkin özdeğerler grafiği incelendiğinde, birinci faktöre ilişkin özdeğerle ikinci faktöre ilişkin özdeğer arasında büyük farklılık olduğu, ikinci faktörün özdeğerinden sonra bu farkların azaldığı belirlenmiştir.

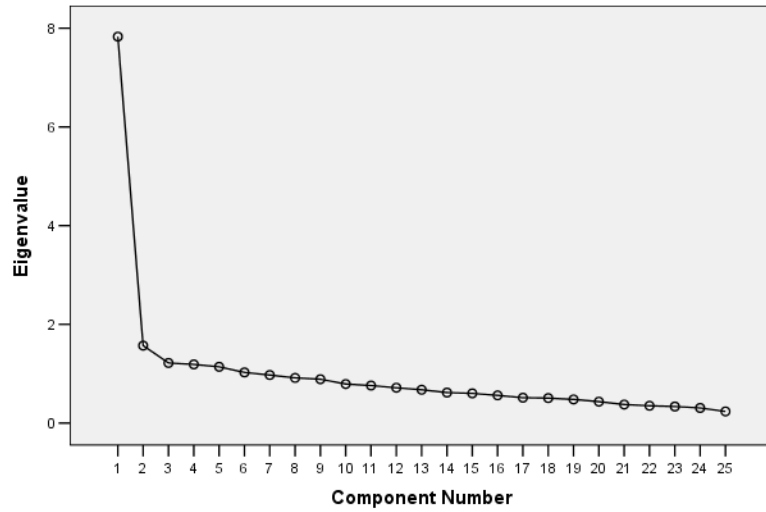
Tüm bu sonuçlar bir arada ele alındığında, KT1'in tek boyutlu olduğuna kanıttır.

Çizelge 6: KT2 (k=25, AU Yüzdesi =%20) İçin Faktör Analizi Sonuçları

Madde No	Faktör Yük Değeri	Madde No	Faktör Yük Değeri
M042003	0,52	M022097	0,60
M042079	0,52	M022101	0,47
M042055	0,45	M022104	0,59
M042039	0,54	M022105	0,15
M042199	0,67	M022108	0,59
M042301A	0,51	M022181	0,58
M042301B	0,65	M032523	0,49
M042301C	0,71	M032701	0,47
M042265	0,41	M032704	0,65
M042137	0,56	M032525	0,64
M042148	0,55	M032579	0,65
M042254	0,43	M032691	0,69
M042250	0,61		

Açıklanan varyans: %31,32

Çizelge 6 incelendiğinde, KT2'de yer alan 25 maddenin 24'ünün birinci faktördeki faktör yükleri 0,41'in üzerindedir. Ayrıca birinci faktörün açıkladığı varyans %31,32 olarak bulunmuştur. Bu değerler, tek boyutluluğun sağlandığına kanıt olarak ele alınmakla birlikte, öz değer grafiği de incelenmiştir.



Şekil 3: KT2 (k=25, AU Yüzdesi =%20) İçin Özdeğerler Grafiği

KT2'ye ilişkin özdeğerler grafiği incelendiğinde, birinci faktöre ilişkin özdeğerle ikinci faktöre ilişkin özdeğer arasında büyük farklılık olduğu, ikinci faktörün özdeğerinden sonra bu farkların azaldığı belirlenmiştir.

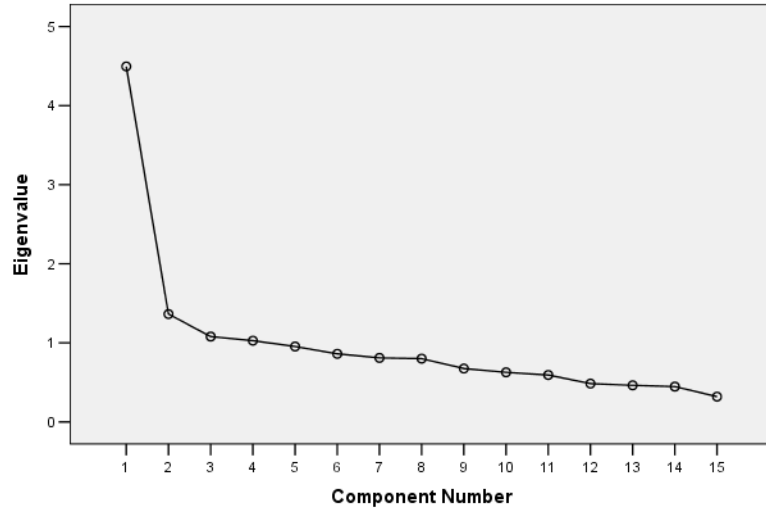
Tüm bu sonuçlar bir arada ele alındığında, KT2'nin tek boyutlu olduğu sonucuna ulaşılmaktadır.

Çizelge 7: KT3 (k=15, AU Yüzdesi =%40) İçin Faktör Analizi Sonuçları

Madde No	Faktör Yük Değeri	Madde No	Faktör Yük Değeri
M042003	0,47	M042254	0,39
M042055	0,37	M042250	0,60
M042199	0,62	M042220	0,67
M042301A	0,50	M022101	0,43
M042301B	0,66	M022106	0,54
M042265	0,43	M032701	0,53
M042137	0,60	M032691	0,69
M042148	0,57		

Açıklanan varyans: %29,97

Çizelge 7 incelendiğinde, KT3'te yer alan 15 maddenin birinci faktördeki faktör yükleri 0,37'nin üzerindedir. Ayrıca birinci faktörün açıkladığı varyans %29,97 olarak bulunmuştur. Bu değerler, tek boyutluluğun sağlandığına kanıt olarak ele alınmakla birlikte, öz değer grafiği de incelenmiştir.



Şekil 4: KT3 (k=15, AU Yüzdesi =%40) İçin Özdeğerler Grafiği

KT3'e ilişkin özdeğerler grafiği incelendiğinde, birinci faktöre ilişkin özdeğerle ikinci faktöre ilişkin özdeğer arasında büyük farklılık olduğu, ikinci faktörün özdeğerinden sonra bu farkların azaldığı belirlenmiştir.

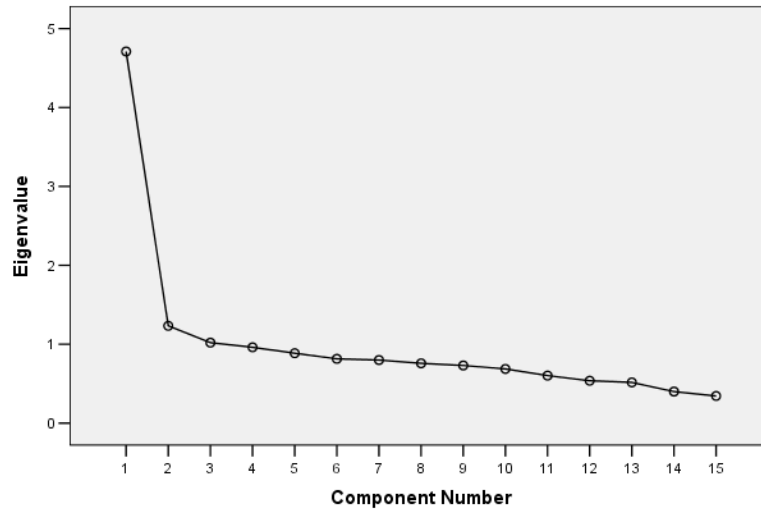
Tüm bu sonuçlar bir arada ele alındığında, KT3'ün tek boyutlu olduğu bulunmuştur.

Çizelge 8: KT4 (k=15, AU Yüzdesi =%20) İçin Faktör Analizi Sonuçları

Madde No	Faktör Yük Değeri	Madde No	Faktör Yük Değeri
M042003	0,55	M022101	0,46
M042055	0,38	M022110	0,49
M042199	0,69	M022181	0,61
M042301B	0,59	M032701	0,48
M042265	0,41	M032525	0,60
M042137	0,60	M032579	0,67
M042148	0,59	M032691	0,68
M042254	0,49		

Açıklanan varyans: %31,40

Çizelge 8 incelendiğinde, KT4'te yer alan 15 maddenin birinci faktördeki faktör yükleri 0,38'in üzerindedir. Ayrıca birinci faktörün açıkladığı varyans %31,40 olarak bulunmuştur. Bu değerler, tek boyutluluğun sağlandığına kanıt olarak ele alınmakla birlikte, öz değer grafiği de incelenmiştir.



Şekil 5: KT4 (k=15, AU Yüzdesi =%20) İçin Özdeğerler Grafiği

KT4'e ilişkin özdeğerler grafiği incelendiğinde, birinci faktöre ilişkin özdeğerle ikinci faktöre ilişkin özdeğer arasında büyük farklılık olduğu, ikinci faktörün özdeğerinden sonra bu farkların azaldığı belirlenmiştir.

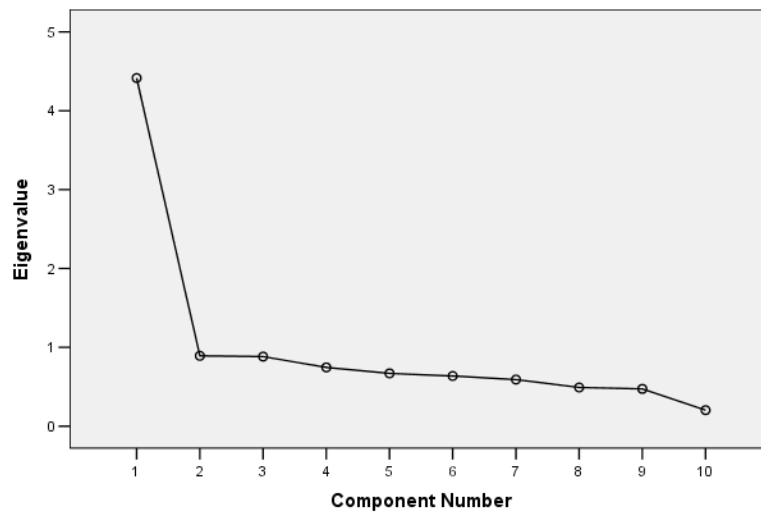
Tüm bu sonuçlar bir arada ele alındığında, KT4'ün tek boyutlu olduğu sonucuna ulaşılmıştır.

Çizelge 9: AUT (k=10) İçin Faktör Analizi Sonuçları

Madde No	Faktör Yük Değeri	Madde No	Faktör Yük Değeri
M042018	0,70	M042220	0,70
M042301A	0,52	M022106	0,61
M042301B	0,75	M022110	0,57
M042301C	0,81	M032307	0,70
M042250	0,53	M032691	0,68

Açıklanan varyans: %44,15

Çizelge 9 incelendiğinde, AUT'ta yer alan 10 maddenin birinci faktördeki faktör yükleri 0,52'nin üzerindedir. Ayrıca birinci faktörün açıkladığı varyans %44,15 olarak bulunmuştur. Bu değerler, tek boyutluluğun sağlandığına kanıt olarak ele alınmakla birlikte, öz değer grafiği de incelenmiştir.



Şekil 6: AUT (k=10) İçin Özdeğerler Grafiği

AUT'a ilişkin özdeğerler grafiği incelendiğinde, birinci faktöre ilişkin özdeğerle ikinci faktöre ilişkin özdeğer arasında büyük farklılık olduğu, ikinci faktörün özdeğerinden sonra bu farkların azaldığı belirlenmiştir.

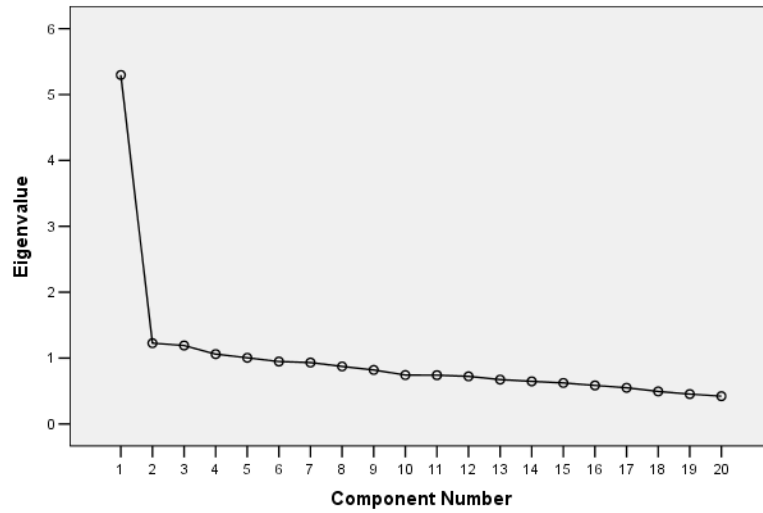
Tüm bu sonuçlar bir arada ele alındığında, AUT'un tek boyutlu olduğu belirlenmiştir.

Çizelge 10: ÇST (k=20) İçin Faktör Analizi Sonuçları

Madde No	Faktör Yük Değeri	Madde No	Faktör Yük Değeri
M042003	0,50	M022101	0,47
M042079	0,47	M022104	0,57
M042055	0,31	M022105	0,22
M042039	0,53	M022108	0,55
M042199	0,62	M022181	0,58
M042265	0,44	M032523	0,47
M042137	0,49	M032701	0,50
M042148	0,52	M032704	0,63
M042254	0,46	M032525	0,59
M022097	0,52	M032579	0,65

Açıklanan varyans: %26,46

Çizelge 10 incelendiğinde, ÇST'de yer alan 20 maddenin birinci faktördeki faktör yükleri 0,22'nin üzerindedir. Ayrıca birinci faktörün açıkladığı varyans %26,46 olarak bulunmuştur. Bu değerler, tek boyutluluğun sağlandığına kanıt olarak ele alınmakla birlikte, öz değer grafiği de incelenmiştir.



Şekil 7: ÇST (k=20) İçin Özdeğerler Grafiği

ÇST'ye ilişkin özdeğerler grafiği incelendiğinde, birinci faktöre ilişkin özdeğerle ikinci faktöre ilişkin özdeğer arasında büyük farklılık olduğu, ikinci faktörün özdeğerinden sonra bu farkların azaldığı belirlenmiştir.

Tüm bu sonuçlar bir arada ele alındığında, ÇST'nin tek boyutlu olduğu belirlenmiştir.

Araştırmanın alt amaçlarına yanıt aramak için oluşturulan karma testler ile çoktan seçmeli ve açık uçlu testlerin tek boyutlu olması ve bilişsel alanlara göre dağılımlarının benzer olması, bu testlerin aynı yapıyı ölçtüğünü göstermektedir.

Model Veri Uyumu Testi:

MTK varsayımlarından model-veri uyumunun üzerinde çalışılacak olan altı test için sağlanıp sağlanmadığı -2loglikelihood istatistiği ölçüt alınarak sınanmıştır. En çok olabilirlik kestiriminde “-2loglikelihood” değeri model veri uyumu indeksi olarak kullanılır. Farklı modeller için kestirilen bu değerler arasındaki fark kaykare istatistiğiyle aşağıdaki gibi değerlendirilir:

$$\chi^2 = -2\loglikelihood_{1PL} - (-2\loglikelihood_{2PL})$$

Bu fark, serbestlik derecesinde göre kaykare tablosundan elde edilen değerden büyükse fark anlamlıdır. Bu durumda, daha kompleks model, fark anlamlı değil ise daha basit model seçilir. Farkın anlamlı olmaması durumunda, bu işlem 2PL ve 3PL modelin -2loglikelihood değerleri için tekrarlanır (Embretson ve Reise, 2000).

Veriler BILOG-MG programında analiz edilmiştir. Test düzeyinde uyumu gösteren bulgular Çizelge 11’de verilmiştir.

Çizelge 11: Açık Uçlu ve Çoktan Seçmeli Test ile Karma Testlerin Test Düzeyinde Model Veri Uyumu Analiz Sonuçları

Test	Test Uzunluğu	Açık Uçlu Madde Yüzdesi (%)	Model	-2Loglikelihood	Fark	Serbestlik Derecesi	p
KT1	25	40	3PL*	8085,237			
			2PL	8125,514	40,276	25	0,027
			1PL	8262,088	136,574	25	0,000
KT2	25	20	3PL*	8511,439			
			2PL	8587,201	75,762	25	0,000
			1PL	8764,478	177,277	25	0,000
KT3	15	40	3PL	5058,286			
			2PL*	5062,535	4,249	15	0,997
			1PL	5123,893	61,358	15	0,000
KT4	15	20	3PL*	5238,325			
			2PL	5275,378	37,053	15	0,001
			1PL	5330,900	55,522	15	0,000
ÇST	20	0	3PL*	7246,961			
			2PL	7335,490	88,529	20	0,000
			1PL	7443,431	107,940	20	0,000
AUT	10	100	2PL*	2885,088			
			1PL	2952,750	67,662	10	0,000

* Test düzeyinde verilerle uyum gösteren modeller

Çizelge 11 incelendiğinde, AUT ve KT3'ün 2 PL; ÇST, KT1, KT2 ve KT4'ün 3 PL modele daha iyi uyum sağladığı görülmektedir. 3 PL model, şans parametresinin eklendiği model olduğu için Açık Uçlu Testin bu modele uyum sağlamaması, 2 PL modele uyum sağlaması öngörülmekteydi. Bununla birlikte KT3'ün de şans parametresinin devre dışı bırakıldığı 2 PL modele uyum göstermesi, madde sayısının az olması ve açık uçlu madde yüzdesinin yüksek olması (%40) nedeniyle AUT gibi çalıştığını düşündürmektedir. Model veri uyumu testinin sonuçlarına bağlı olarak KT1, KT2, KT4 ve ÇST için kestirimler 3PL modele göre; KT3 ve AUT için kestirimler 2PL modele göre yapılmıştır.

2. Araştırma Sorularına İlişkin Analizler

Araştırmanın “madde parametreleri b, a, c hangi aralıkta değişmektedir?”, “yetenek parametreleri ve standart hata değerleri hangi aralıkta değişmektedir?”, “madde ve test bilgi fonksiyonları ile görelî etkinlik indeksleri hangi düzeydedir” biçimindeki 1, 2 ve 3 numaralı sorularına ilişkin madde ve yetenek kestirimleri, yetenek kestirimine ilişkin standart hata, madde ve test bilgi fonksiyonu, görelî etkinlik indeksleri BILOG MG programında beklenen a posteriori (expected a posteriori) yöntemiyle elde edilmiştir. Beklenen a posteriori (expected a posteriori) yöntemi, diğer kestirim yöntemlerinden (ençok olabilirlik, en yüksek posteriori) ayrı olarak tüm tepki örüntüleri için (tümü doğru - tümü yanlış) sonlu bir yetenek kestirimi ($\theta = -3.0 - 3.0$) sağlamaktadır. Ayrıca iteratif bir prosedür değildir (Embretson ve Reise, 2000.)

Araştırmanın “yetenek parametreleri arasındaki ilişki miktarı nedir?” ve “yetenek parametreleri ile TIMSS 2007 genel matematik başarı puanları (overall mathematics achievement score) arasındaki ilişki miktarı nedir?” biçiminde ifade edilen 4 ve 5 numaralı sorularda yer alan yetenek kestirimleri arasındaki ilişkinin büyüklüğü ve yönünü belirlemek amacıyla sürekli iki değişken arasındaki doğrusal ilişkinin derecesini veren Pearson Momentler Çarpımı Korelasyon Katsayıları hesaplanmıştır.

BÖLÜM 3

BULGULAR VE YORUMLAR

Bu bölümde, araştırmanın alt amaçları doğrultusunda yapılan analizlerden elde edilen bulgular ve yorumları yer almaktadır. Madde sayısının 25 ve açık uçlu madde yüzdesinin %40 ve %20 olduğu sırasıyla KT1 ve KT2 karma testleri, madde sayısının 15 ve açık uçlu madde yüzdesinin %40 ve %20 olduğu sırasıyla KT3 ve KT4 karma testleri, açık uçlu test ve çoktan seçmeli test için kestirilen madde ve yetenek parametreleri, yetenek parametrelerine ilişkin standart hata düzeyleri, bu testlerden kestirilen yetenek parametreleri arasındaki Pearson korelasyon katsayıları, yetenek parametreleri ile TIMSS 2007 toplam matematik başarı puanları arasındaki Pearson korelasyon katsayıları, madde ve test bilgi fonksiyonları ile görelî etkinlik indeksleri incelenmiştir.

1. KT1, KT2, KT3, KT4, AUT ve ÇST için ayrı ayrı kestirilen madde parametreleri (b, a, c) hangi aralıkta değişmektedir?

KT1, KT2, KT3, KT4, AUT ve ÇST için Madde Tepki Kuramına göre madde güçlük (b), madde ayırt edicilik (a) ve şans (c) parametreleri kestirilmiş; sonuçlar Çizelge 12’de verilmiştir.

Çizelge 12: Karma Testler, Çoktan Seçmeli Test ve Açık Uçlu Teste İlişkin Madde Parametreleri

Test Adı			KT1				KT2				KT3*			KT4				ÇST				AUT*		
Madde Sayısı			25				25				15			15				20				10		
AU Madde Oranı			%40				%20				%40			%20										
Madde Kodu	Format	Düz.	Madde No	b	a	c	Madde No	b	a	c	Madde No	b	a	Madde No	b	a	c	Madde No	b	a	c	Madde No	b	a
M042003	ÇS	B	1	-0,1	1,5	0,2	1	-0,1	1,4	0,2	1	-0,6	1,1	1	-0,1	1,5	0,2	1	-0,1	1,4	0,2	--	--	--
M042079	ÇS	B	2	0,3	1,9	0,3	2	0,4	2,2	0,3	--	--	--	--	--	--	--	2	0,2	1,9	0,3	--	--	--
M042018	AU	U	3	0,9	3,3	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	1	0,8	2,0
M042055	ÇS	U	4	1,4	2,9	0,2	3	1,4	4,8	0,2	2	1,5	0,7	2	1,4	3,5	0,2	3	1,5	3,1	0,2	--	--	--
M042039	ÇS	U	--	--	--	--	4	0,9	1,8	0,1	--	--	--	--	--	--	--	4	0,9	1,8	0,1	--	--	--
M042199	ÇS	B	5	0,0	2,5	0,2	5	0,0	2,3	0,2	3	-0,3	1,5	3	0,0	2,3	0,2	5	0,0	2,4	0,2	--	--	--
M042301A	AU	B	6	0,7	1,6	--	6	0,6	1,5	--	4	0,4	1,3	--	--	--	--	--	--	--	--	2	0,4	1,3
M042301B	AU	AY	7	1,0	2,7	--	7	1,0	2,5	--	5	0,9	2,1	4	1,1	2,2	--	--	--	--	--	3	0,8	2,5
M042301C	AU	AY	8	1,1	4,2	--	8	1,1	3,6	--	--	--	--	--	--	--	--	--	--	--	--	4	1,0	4,6
M042265	ÇS	AY	9	1,4	1,3	0,1	9	1,4	1,3	0,1	6	1,1	1,0	5	1,3	1,3	0,1	6	1,3	1,3	0,1	--	--	--
M042137	ÇS	U	10	0,8	2,2	0,2	10	0,8	2,3	0,3	7	0,2	1,1	6	0,8	2,2	0,2	7	0,8	2,1	0,2	--	--	--
M042148	ÇS	B	11	-0,2	1,3	0,2	11	-0,2	1,4	0,2	8	-0,6	1,2	7	-0,2	1,5	0,2	8	-0,2	1,5	0,2	--	--	--
M042254	ÇS	U	12	-0,9	1,6	0,2	12	-0,9	1,5	0,2	9	-1,1	1,6	8	-0,9	1,5	0,2	9	-0,9	1,5	0,2	--	--	--
M042250	AU	B	13	-0,5	2,4	--	13	-0,5	2,5	--	10	-0,7	2,0	--	--	--	--	--	--	--	--	5	-0,7	2,1
M042220	AU	U	14	1,0	2,6	--	--	--	--	--	11	1,0	2,0	--	--	--	--	--	--	--	--	6	1,0	2,0
M022097	ÇS	B	15	0,6	1,8	0,2	14	0,6	1,8	0,2	--	--	--	--	--	--	--	10	0,6	1,8	0,2	--	--	--
M022101	ÇS	B	16	-0,4	1,2	0,2	15	-0,4	1,3	0,2	12	-0,9	1,1	9	-0,4	1,3	0,2	11	-0,4	1,4	0,2	--	--	--
M022104	ÇS	B	17	0,5	1,9	0,1	16	0,5	1,8	0,1	--	--	--	--	--	--	--	12	0,5	1,9	0,2	--	--	--
M022105	ÇS	B	--	--	--	--	17	2,4	0,8	0,2	--	--	--	--	--	--	--	13	2,4	0,9	0,2	--	--	--
M022106	AU	U	18	1,1	2,0	--	--	--	--	--	13	0,9	1,4	--	--	--	--	--	--	--	--	7	1,0	1,3
M022108	ÇS	U	--	--	--	--	18	0,7	4,3	0,2	--	--	--	--	--	--	--	14	0,7	4,7	0,3	--	--	--
M022110	AU	B	19	0,8	1,6	--	--	--	--	--	--	--	--	10	0,8	1,5	--	--	--	--	--	8	0,4	1,0
M022181	ÇS	U	20	-0,1	1,9	0,2	19	-0,1	2,0	0,2	--	--	--	--	--	--	--	11	-0,1	2,0	0,2	--	--	--
M032307	AU	U	21	1,3	2,9	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	9	1,3	2,5
M032523	ÇS	U	--	--	--	--	20	1,1	2,5	0,2	--	--	--	--	--	--	--	16	1,2	2,5	0,2	--	--	--
M032701	ÇS	U	22	-0,7	1,9	0,2	21	-0,7	2,0	0,2	14	-1,1	1,4	12	-0,6	2,1	0,2	17	-0,6	2,2	0,2	--	--	--
M032704	ÇS	U	--	--	--	--	22	0,0	2,7	0,2	--	--	--	--	--	--	--	18	0,0	3,1	0,2	--	--	--
M032525	ÇS	B	23	0,1	2,5	0,2	23	0,1	2,6	0,2	--	--	--	13	0,2	2,7	0,2	19	0,1	2,7	0,2	--	--	--
M032579	ÇS	U	24	-0,1	2,6	0,1	24	-0,1	2,8	0,1	--	--	--	14	-0,1	2,5	0,2	20	-0,1	2,6	0,1	--	--	--
M032691	AU	U	25	1,2	2,5	--	25	1,2	2,5	--	15	1,1	2,2	15	1,1	2,7	--	--	--	--	--	10	1,2	1,9
Ortalama				0,44	2,18	0,19		0,45	2,24	0,19		0,12	1,43		0,29	2,04	0,19		0,39	2,13	0,19		0,72	2,13
Ortanca				0,57	1,98	0,18		0,51	2,15	0,18		0,23	1,37		0,03	2,06	0,20		0,16	1,99	0,19		0,91	2,01
Std. Sapma				0,68	0,70	0,05		0,77	0,92	0,05		0,93	0,46		0,74	0,63	0,04		0,79	0,85	0,04		0,57	1,02
En Düşük				-0,87	1,21	0,11		-0,89	0,84	0,10		-1,12	0,73		-0,91	1,27	0,10		-0,89	0,85	0,10		-0,69	1,01
En Yüksek				1,39	4,19	0,31		2,44	4,81	0,33		1,47	2,24		1,43	3,49	0,23		2,41	4,68	0,28		1,27	4,63
Genişlik				2,27	2,98	0,21		3,33	3,97	0,23		2,59	1,52		2,33	2,22	0,13		3,29	3,83	0,18		1,96	3,62

* KT3 ve AUT için parametreler 2 Parametrelili Liojistik modele göre kestirilmiştir.

ÇS: Çoktan Seçmeli, AU: Açık Uçlu, Düz.: Maddenin ölçtüğü bilişsel düzey, B: Bilgi, U: Uygulama, AY: Akıl Yürütme

Çizelge 12 incelendiğinde, en kolay testin KT3, en zor testin AU test olduğu görülmektedir. Bu bulgu, Lukhele (1994), Özdemir (1997), Karaca (2004) ve In'nami&Koizumi'nin (2009) de ortaya koyduğu; açık uçlu testlerin çoktan seçmeli testlere göre daha güç olduğu bulgusuyla da paraleldir. En güç test AU test olmasına rağmen, madde sayısı 25 olan KT1 ve KT2'nin ortalama güçlük değerlerinin birbirlerine çok yakın olması, karma testteki açık uçlu madde yüzdesinin testin güçlüğünde farklılık yaratmadığını göstermiştir. Diğer bir deyişle, açık uçlu madde yüzdesinin farklılaşması (%20, %40) uzun testin ($k=25$) güçlüğüne değiştirmemiştir.

Madde güçlük standart sapmasının en düşük olduğu test AUT'tur. Bu bulgu, Wester ve Henriksson (2000) tarafından da ortaya konmuştur. ÇST ile madde sayısı 25 ve AU yüzdesi %20 olan KT2'nin güçlük değerlerinin hemen hemen aynı aralıkta değiştiği belirlenmiştir. Bu durum, KT2'nin AU yüzdesi düşük olması nedeniyle ÇST'ye benzer sonuçlar verdiği şeklinde yorumlanabilir. AU testin madde güçlük parametrelerinin standart sapması diğer üç teste göre daha küçük bulunmuştur. Bu durumda AUT'un madde güçlüğü bakımından daha homojen bir dağılım gösterdiği söylenebilir. AUT'un madde güçlük ranjının diğer üç testekine göre daha dar olması, madde sayısının az olmasıyla da açıklanabilir.

Açık uçlu ve çoktan seçmeli testin ortalama ayırıcılık değerlerinin aynı olmasına rağmen, madde sayısı 25 ve AU yüzdesi sırasıyla %40 ve %20 olan karma testler, KT1 ve KT2'den elde edilen ortalama ayırıcılık değerleri daha yüksek bulunmuştur. Bu durumda, karma testlerin sadece çoktan seçmeli ve sadece açık uçlu maddelerden oluşan testlere göre daha ayırt edici olduğu sonucuna ulaşılabilir.

Karma testler içinde, ayırıcılığı en yüksek olan KT1 ve KT2'dir. Ayrıca, karma testler içinde madde sayısı 25 olan açık uçlu madde yüzdesi %20 olan KT2'nin ortalama ayırıcılık değeri, açık uçlu madde yüzdesi %40 olan KT1'in ortalama ayırıcılık değerinden yüksek bulunmuştur. Bununla birlikte, madde sayısı 15 olan karma testlerden açık uçlu madde yüzdesi %20 olan KT4'ün ortalama ayırıcılık değeri, açık uçlu madde yüzdesi %40 olan KT3'ün ortalama ayırıcılık değerinden yüksek bulunmuştur. Bu durumda, uzun ve açık uçlu

madde yüzdesi düşük olan karma testlerin daha ayırıcı olduğu yorumu yapılabilir.

TIMSS maddeleri, deneme uygulamaları yapılmış, nitelikli maddeler olduğundan, şans parametresi belli bir değerin üstüne çıkmayan maddelerin uygulamada bulunması beklenmektedir. Madde şans parametreleri kestirimleri incelendiğinde, KT1, KT2, KT4 ve ÇST’de ortalama şans parametresi değerlerinin aynı ve oldukça düşük olduğu görülmektedir.

2. KT1, KT2, KT3, KT4, AUT ve ÇST için ayrı ayrı kestirilen yetenek parametreleri ve bunlara ait standart hata değerleri hangi aralıkta değişmektedir?

Testi alan bireylerin dört karma test, çoktan seçmeli ve açık uçlu testler için ayrı ayrı kestirilen yetenek parametrelerine ilişkin betimsel istatistikler Çizelge 13’te verilmiştir.

Çizelge 13: Yetenek Parametrelerine İlişkin Betimsel İstatistikler

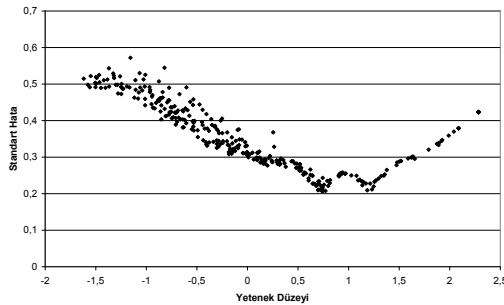
Test Adı	KT1	KT2	KT3	KT4	ÇST	AUT
Madde sayısı	25	25	15	15	20	10
AU Yüzdesi	%40	%20	%40	%20		
Ortalama	0,01	0,01	0,00	0,00	0,01	0,00
Ortanca	-0,12	-0,08	-0,13	-0,08	-0,12	-0,23
Standart Sapma	0,94	0,94	0,91	0,90	0,92	0,90
Çarpıklık	0,48	0,48	0,53	0,41	0,54	0,67
Basıklık	-0,41	-0,45	-0,19	-0,66	-0,49	-0,17
En Düşük	-1,62	-1,69	-1,88	-1,55	-1,58	-1,20
En Yüksek	2,29	2,23	2,24	2,04	2,14	2,16
Genişlik	3,91	3,92	4,12	3,59	3,72	3,36

Çizelge 13 incelendiğinde, dört karma test ile çoktan seçmeli ve açık uçlu testten kestirilen yetenek parametresi dağılımlarının benzer olduğu görülmektedir. Farklı testlerin yetenek kestirimlerini standart hata düzeyleriyle

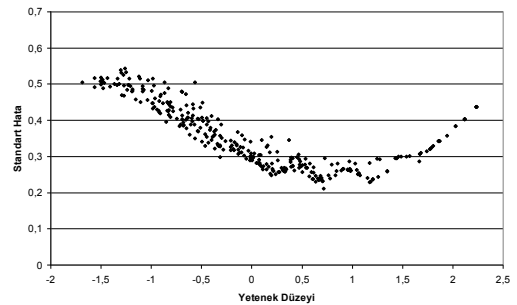
birlikte ele almak yorumların daha sağlıklı yapılmasını sağlamaktadır. Standart hata değerlerinin, her bir yetenek düzeyinde sıfıra yakın olması, yetenek kestiriminin güvenilir olduğunun bir göstergesi olmaktadır. Bu doğrultuda, bireylerin yetenek kestirimlerinin bireysel standart hata değerlerine ilişkin betimsel istatistikler Çizelge 14'te, standart hata değerlerinin yetenek düzeyine göre dağılımları Şekil 8-13'te verilmiştir.

Çizelge 14: Yetenek Parametrelerine İlişkin Standart Hata Değerlerine ilişkin betimsel istatistikler

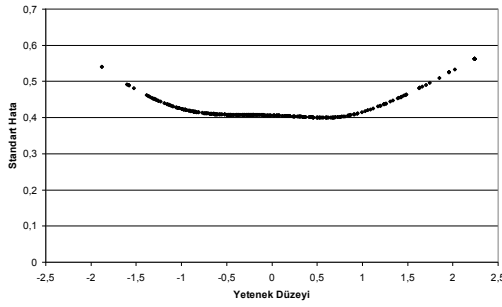
Test Adı	KT1	KT2	KT3*	KT4	ÇST	AUT
Madde sayısı	25	25	15	15	20	10
AU Yüzdesi	%40	%20	%40	%20		
Ortalama	0,36	0,36	0,42	0,44	0,40	0,46
Ortanca	0,34	0,33	0,41	0,43	0,39	0,46
Standart Sapma	0,09	0,09	0,04	0,07	0,08	0,10
Çarpıklık	0,24	0,42	2,46	0,36	0,22	0,17
Basıklık	-1,11	-1,08	5,68	-1,01	-1,15	-0,97
En Düşük	0,21	0,21	0,40	0,34	0,26	0,31
En Yüksek	0,57	0,54	0,56	0,61	0,56	0,62
Genişlik	0,36	0,33	0,16	0,27	0,30	0,31



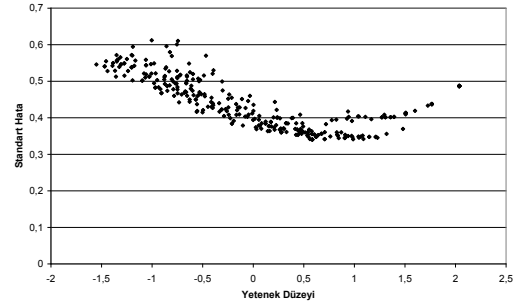
Şekil 8: KT1'i Alan Bireylerin Yetenek Parametrelerine İlişkin Standart Hata Dağılımları



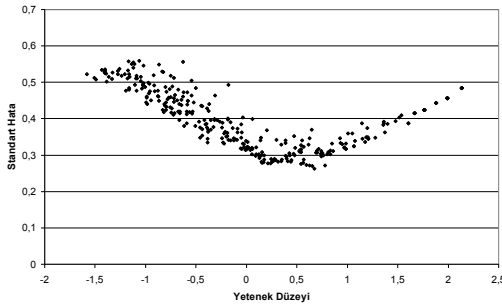
Şekil 9: KT2'yi Alan Bireylerin Yetenek Parametrelerine İlişkin Standart Hata Dağılımları



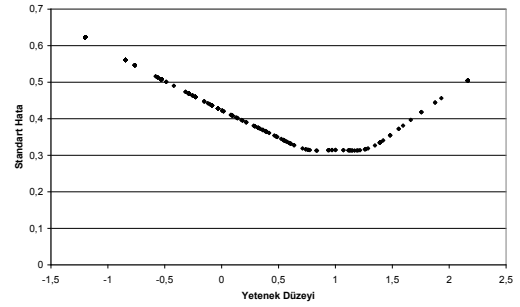
Şekil 10: KT3'ü Alan Bireylerin
Yetenek Parametrelerine İlişkin
Standart Hata Dağılımları



Şekil 11: KT4'ü Alan Bireylerin
Yetenek Parametrelerine İlişkin
Standart Hata Dağılımları



Şekil 12: ÇST İçin Kestirilen
Yetenek Parametrelerine İlişkin
Standart Hata Dağılımları



Şekil 13: AUT İçin Kestirilen
Yetenek Parametrelerine İlişkin
Standart Hata Dağılımları

Çizelge 14 ve şekil 8-13 incelendiğinde, testlerin tümünün düşük standart hatayla yetenek kestirimi yaptıkları görülmektedir. En düşük değer KT1 ve KT2'de $\theta=0,7$ düzeyinde, en yüksek değer AUT'ta $\theta=-1,2$ düzeyinde elde edilmiştir.

Standart hata ortalaması en düşük olan testlerin KT1 ve KT2 olduğu ve bu testlerin özellikle (0,5 – 1,5) yetenek aralığında en düşük standart hatayla yetenek kestirimi yaptığı görülmektedir. Açık uçlu madde yüzdesine bakılmaksızın uzun karma testlerin orta yetenek aralığında daha iyi yetenek kestirimi yaptığı belirlenmiştir.

Genel olarak tüm testlerin yetenek kestirimlerine ilişkin en yüksek standart hatanın (-1,5 – +2.0) gibi yetenek ölçeğinin uç değerlerinde ortaya çıktığı belirlenmiştir. Bu durumda, en düşük ve en yüksek yetenek düzeylerinde testlerin daha büyük standart hatayla kestirim yaptığı söylenebilir.

3. Testlere ait madde ve test bilgi fonksiyonları, görelî etkinlik indeksleri hangi düzeydedir?

Test ve madde bilgi fonksiyonları ile test bilgi fonksiyonlarının oranlanmasıyla elde edilen görelî etkinlik değerleri, test geliştiricilere madde ve test seçme açısından kaynaklık etmektedir. Bu doğrultuda, dört karma test, çoktan seçmeli ve açık uçlu testlerden hangilerinin bireylerin yeteneklerini kestirmede daha fazla bilgi sağladığını ve daha etkin olarak kullanılabileceğini ortaya koymak amacıyla kestirilen madde ve test bilgi fonksiyonları, görelî etkinlik indeksleri bu bölümde tartışılmıştır. Dört karma test, çoktan seçmeli ve açık uçlu testlerden kestirilen ortalama madde bilgi fonksiyonları Çizelge 15'te gösterilmiştir.

Çizelge 15'te testlerin ortalama bilgi düzeyleri incelendiğinde, genel olarak birbirlerine yakın oldukları görülmektedir. Testlere ilişkin ortalama bilgi düzeyleri genel olarak yakın olmakla beraber; en fazla bilginin AUT'tan sağlandığı görülmektedir. Sağladığı bilgi bakımından AUT'u KT1 ve KT2 izlemiştir.

Bu noktada, maddelerden sağlanan bilginin madde tiplerine göre farklılaşıp farklılaşmadığını ortaya koyma ihtiyacı doğmuştur. Bu nedenle, ortalama bilgi değerleri madde tipine göre ayrı ayrı değerlendirilmiş ve Çizelge 16'da verilmiştir.

Çizelge 15: Karma Testler, Çoktan Seçmeli Test ve Açık Uçlu Teste İlişkin Madde Bilgi Fonksiyonları

Test Adı	KT1		KT2		KT3*		KT4		ÇST		AUT*	
Madde Sayısı	25		25		15		15		20		10	
AU Madde Oranı	%40		%20		%40		%20					
Madde Kodu	Format	Düz.	Madde No	Ort. Bilgi	Madde No	Ort. Bilgi	Madde No	Ort. Bilgi	Madde No	Ort. Bilgi	Madde No	Ort. Bilgi
M042003	ÇS	B	1	0,24	1	0,23	1	0,23	1	0,23	--	--
M042079	ÇS	B	2	0,24	2	0,27	--	--	2	0,26	--	--
M042018	AU	U	3	0,57	--	--	--	--	--	--	1	0,49
M042055	ÇS	U	4	0,23	3	0,35	2	0,09	3	0,21	--	--
M042039	ÇS	U	--	--	4	0,26	--	--	4	0,25	--	--
M042199	ÇS	B	5	0,50	5	0,47	3	0,38	5	0,46	--	--
M042301A	AU	B	6	0,25	6	0,25	4	0,31	--	--	2	0,29
M042301B	AU	AY	7	0,48	7	0,42	5	0,47	4	0,39	3	0,62
M042301C	AU	AY	8	0,80	8	0,64	--	--	--	--	4	1,08
M042265	ÇS	AY	9	0,13	9	0,13	6	0,15	6	0,13	--	--
M042137	ÇS	U	10	0,27	10	0,27	7	0,22	7	0,24	--	--
M042148	ÇS	B	11	0,20	11	0,22	8	0,24	8	0,23	--	--
M042254	ÇS	U	12	0,25	12	0,23	9	0,31	9	0,22	--	--
M042250	AU	B	13	0,51	13	0,53	10	0,52	--	--	5	0,55
M042220	AU	U	14	0,46	--	--	11	0,42	--	--	6	0,44
M022097	ÇS	B	15	0,28	14	0,28	--	--	10	0,27	--	--
M022101	ÇS	B	16	0,17	15	0,19	12	0,19	11	0,20	--	--
M022104	ÇS	B	17	0,33	16	0,32	--	--	12	0,33	--	--
M022105	ÇS	B	--	--	17	0,03	--	--	13	0,03	--	--
M022106	AU	U	18	0,27	--	--	13	0,28	--	--	7	0,25
M022108	ÇS	U	--	--	18	0,64	--	--	14	0,68	--	--
M022110	AU	B	19	0,22	--	--	--	--	10	0,16	8	0,20
M022181	ÇS	U	20	0,38	19	0,38	--	--	11	0,41	--	--
M032307	AU	U	21	0,42	--	--	--	--	--	--	9	0,45
M032523	ÇS	U	--	--	20	0,27	--	--	--	--	--	--
M032701	ÇS	U	22	0,31	21	0,33	14	0,26	12	0,39	--	--
M032704	ÇS	U	--	--	22	0,57	--	--	--	--	--	--
M032525	ÇS	B	23	0,49	23	0,52	--	--	13	0,34	--	--
M032579	ÇS	U	24	0,59	24	0,63	--	--	14	0,56	--	--
M032691	AU	U	25	0,40	25	0,40	15	0,46	15	0,53	10	0,37
Ortalama				0,36		0,35		0,30		0,33		0,47
Std. Sapma				0,16		0,16		0,12		0,18		0,25
En Düşük				0,13		0,03		0,09		0,03		0,20
En Yüksek				0,80		0,64		0,52		0,68		1,08
Genişlik				0,67		0,62		0,43		0,65		0,89

* KT3 ve AUT için parametreler 2 Parametrelili Lojistik modele göre kestirilmiştir.

ÇS: Çoktan Seçmeli, **AU:** Açık Uçlu, **Düz.:** Maddenin ölçüldüğü bilişsel düzey, **B:** Bilgi, **U:** Uygulama, **AY:** Akıl Yürütme

Çizelge 16: Dört Karma Testte, ÇST ve AUT'ta Madde Formatına Göre Ortalama Test Bilgi Ortalamaları

		Madde Tipi	
		Açık Uçlu	Çoktan Seçmeli
Ortalama Bilgi	KT1	0,44	0,31
	KT2	0,45	0,33
	KT3	0,41	0,23
	KT4	0,36	0,30
	ÇST	--	0,33
	AUT	0,47	--

Çizelge 16 incelendiğinde, farklı testlerdeki açık uçlu maddelerden sağlanan ortalama bilginin çoktan seçmeli maddelerden sağlanan bilgiden daha fazla olduğu görülmektedir. Bu bulgu, Lukhele ve diğerleri (1994) ve Sykes ve diğerleri (2001) yanıtı sınırlı açık uçlu maddelerin çoktan seçmeli maddelerden daha fazla bilgi verdiği bulgusuyla örtüşmektedir.

Farklı testlerin birbirine tercih edilmesi sözkonusu olduğunda kıyaslanması amacıyla kullanılan göreceli etkinlik (relative efficiency) değerleri, testlere ait ortalama test bilgi fonksiyonlarının birbirlerine oranlanmasıyla hesaplanmış ve Çizelge 17'de verilmiştir.

Çizelge 17: Testlere İlişkin Göreceli Etkinlik (Relative Efficiency) İndeksi Değerleri

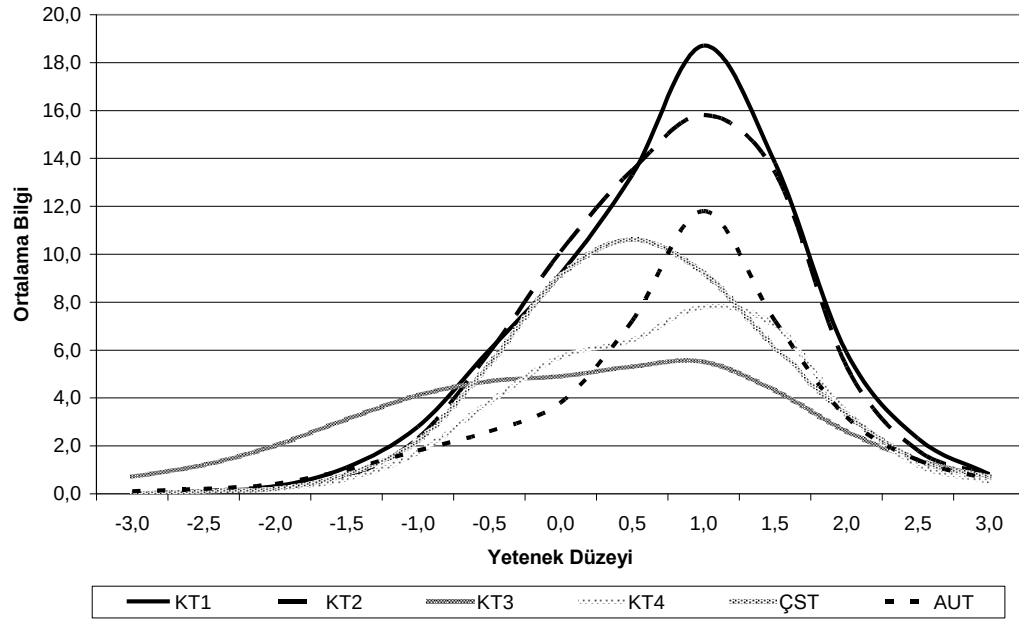
Test Adı	KT1	KT2	KT3	KT4	ÇST	AUT
KT1	1	1,02	1,19	1,16	1,09	0,76
KT2	0,98	1	1,17	1,13	1,07	0,74
KT3	0,84	0,85	1	0,97	0,92	0,63
KT4	0,86	0,88	1,03	1	0,94	0,65
ÇST	0,91	0,93	1,09	1,06	1	0,69
AUT	1,32	1,34	1,57	1,53	1,44	1

Görelî etkinlik indeksi değerlerine göre, KT1; KT2'den %2, KT3'ten %19, KT4'ten %16, ÇST'den %9 daha uzun bir test gibi ölçülen özelliğe ilişkin bireylerarası farklılıkları daha fazla ortaya koyduğu görülmüştür. Öte yandan, AUT; KT1'den %32, KT2'den %34, KT3'ten %57, KT4'ten %53, ÇST'den %44 daha uzun bir test gibi işlem görmüştür. Bu durum, AUT'un madde sayısı az olmasına rağmen, daha uzun testlere göre testle ölçülen özellik bakımından bireylerarası farkları göstermede daha etkili olduğunu göstermiştir. Ayrıca, AU maddelerin karma testler arasında madde sayısı 25 ve açık uçlu madde yüzdesi %40 olan KT1'in, AUT haricindeki diğer testlerden daha etkili olduğu söylenebilir.

Test bilgi fonksiyonu yorumlanırken; yetenek düzeyi ve test bilgi fonksiyonu arasındaki ilişki göz önünde tutulmalıdır (Baker, 2001). Dolayısıyla testten elde edilen bilgi, hangi yetenek düzeylerinde en yüksek değerlerini alıyorsa, o test o yetenek düzeyindeki bireylere hitap ediyor demektir. Bu çıkarımdan hareketle, dört farklı karma testin, çoktan seçmeli ve açık uçlu testten elde edilen bilgi miktarının yetenek düzeyi aralıklarına göre dağılımı Çizelge 18 ve Şekil 14'te verilmiştir.

Çizelge 18: Yetenek Düzeylerine Göre Test Bilgi Fonksiyonları Dağılımı

Testler	YETENEK DÜZEYİ												
	-3,0	-2,5	-2,0	-1,5	-1,0	-0,5	0,0	0,5	1,0	1,5	2,0	2,5	3,0
KT1	0,0	0,1	0,3	1,1	2,8	6,0	9,2	13,3	18,7	13,8	5,9	2,3	0,8
KT2	0,0	0,1	0,3	0,7	2,3	5,9	10,1	13,5	15,8	13,4	5,3	1,8	0,8
KT3	0,7	1,2	2,0	3,1	4,1	4,7	4,9	5,3	5,5	4,3	2,6	1,4	0,7
KT4	0,0	0,1	0,2	0,6	1,7	3,8	5,7	6,4	7,8	7,0	3,4	1,2	0,5
ÇST	0,0	0,1	0,2	0,8	2,2	5,4	9,1	10,6	9,2	6,1	3,3	1,4	0,7
AUT	0,1	0,2	0,4	1,0	1,8	2,6	3,8	7,2	11,8	7,2	3,2	1,4	0,6



Şekil 14: Yetenek Düzeylerine Göre Test Bilgi Fonksiyonları Dağılımı

Genel olarak, araştırma kapsamında düzenlenen testlerin tümünün farklı yetenek düzeylerinde bilgi verdiği görülmektedir. Bununla birlikte, Çizelge 18 ve Şekil 14 incelendiğinde, KT3'ün ($k=15$, AU yüzdesi: %40), hemen her yetenek düzeyinde bilgi verdiği, özellikle alt yetenek düzeyinde (-3,-1) diğer testlere göre daha fazla bilgi verdiği görülmektedir. KT3'ün, b değerleri bakımından ortalama olarak en düşük olan test, diğer bir deyişle diğerlerine göre kolay test olduğu önceki kısımlarda ortaya konmuştu. Test bilgi fonksiyonu, madde gücü ve ayıricılık ilişkili olduğundan, alt yetenek düzeylerine hitap eden KT3, alt yetenek düzeylerinde daha fazla bilgi vermiştir.

Diğer testlerin benzer yetenek aralığında bilgi verdiği, sadece KT4'ün ($k=15$, AU yüzdesi: %20) diğerlerine göre biraz daha dar bir yetenek aralığında bilgi verdiği belirlenmiştir.

Test bilgi fonksiyonu, madde bilgi fonksiyonlarının toplanmasıyla elde edilmektedir (Embretson ve Reise, 2000). Bu durumda madde sayısı fazla olan testten daha fazla bilgi elde edilmesi beklenir. Buna rağmen; en az madde sayısına sahip olan AUT, $\theta=1$ düzeyinde madde sayısı daha fazla

olan ÇST, KT3 ve KT4'ten daha fazla bilgi sağlamıştır. Öte yandan, AUT en yüksek bilgiyi (0.5 – 1.5) yetenek aralığında vermiştir. AUT, en zor test olduğundan orta ve üst yetenek düzeylerinde daha fazla bilgi vermiştir.

Test bilgi fonksiyonu, madde ayıricılığı ve şans parametresiyle de ilişkilidir. Ancak, araştırma kapsamında ele alınan testlerin ayıricılık düzeyleri ve şans parametreleri (3PL için) birbirine yakın bulunduğundan (Çizelge 12), bu noktada test bilgi fonksiyonları üzerindeki farklı bir etkisi görülememiştir.

4. KT1, KT2, KT3, KT4, AUT ve ÇST için ayrı ayrı kestirilen yetenek parametreleri arasındaki Pearson Momentler Çarpım Korelasyon Katsayıları ne düzeydedir?

Testi alan bireyler için dört farklı karma teste, çoktan seçmeli ve açık uçlu teste göre ayrı ayrı kestirilen yetenek parametreleri arasındaki Pearson Momentler Çarpım Korelasyon katsayıları Çizelge 19'de verilmiştir.

Çizelge 19: Yetenek Parametreleri Arasındaki Pearson Momentler Çarpımı Korelasyon Katsayıları

	KT2	KT3	KT4	ÇST	AUT
KT1	0,98	0,96	0,97	0,96	0,92
KT2	--	0,94	0,97	0,98	0,88
KT3	--	--	0,93	0,90	0,91
KT4	--	--	--	0,96	0,83
ÇST	--	--	--	--	0,81

Çizelge 19 incelendiğinde, kestirilen yetenek parametreleri arasında hesaplanan Pearson Momentler Çarpımı Korelasyon katsayılarının tamamının pozitif yönde ve oldukça yüksek olduğu görülmektedir. Bu bulgu, Ercikan ve diğerlerinin (1998) açık uçlu, çoktan seçmeli ve karma testlerin

ölçek puanları arasındaki korelasyonların yüksek olduğu bulgusuyla da paraleldir. En yüksek korelasyonlar, KT1 ile KT2 ve ÇST ile KT2 yetenek parametreleri arasında bulunmuştur (0,98). Buna karşın ÇST ile AUT ve KT4 ile AUT yetenek parametreleri arasında sırasıyla 0,81 ve 0,83 değerlerinde, diğerlerine göre düşük, tek başına değerlendirildiğinde yine yüksek kabul edilebilen bir korelasyon bulunmuştur. Bu durumda, test uzunluğu ve açık uçlu madde yüzdesi birbirine yakın olan testlerden kestirilen yetenek parametreleri arasındaki korelasyonun daha yüksek olduğu sonucuna ulaşılabilir.

5. KT1, KT2, KT3, KT4, AUT ve ÇST için ayrı ayrı kestirilen yetenek parametreleri ile TIMSS 2007 Genel Matematik Başarı Puanları arasındaki Pearson Momentler Çarpımı Korelasyon Katsayıları ne düzeydedir?

KT1, KT2, KT3, KT4, AUT ve ÇST için ayrı ayrı kestirilen yetenek parametreleri ile öğrencilerin TIMSS 2007 testinden aldıkları Genel Matematik Başarı Puanları (Overall Math Achievement Score) arasındaki Pearson Momentler Çarpımı Korelasyon katsayıları hesaplanmış ve Çizelge 20’de verilmiştir.

Çizelge 20: Yetenek Parametreleri İle TIMSS 2007 Genel Matematik Başarı Puanları Arasındaki Pearson Korelasyon Katsayıları

Test	Test Uzunluğu	AU Madde Yüzdesi	TIMSS Puanı İle Korelasyon
KT1	25	%40	0,96
KT2	25	%20	0,96
KT3	15	%40	0,93
KT4	15	%20	0,93
ÇST	20	%0	0,94
AUT	10	%100	0,89

Testi alan öğrencilerin TIMSS 2007 testinin tamamından test eşitleme yoluyla kestirilen genel matematik başarı puanları ile KT1, KT2, KT3, KT4, ÇST ve AUT için kestirilen yetenek parametreleri arasındaki Pearson Momentler Çarpımı Korelasyon katsayıları pozitif yönde oldukça yüksek bulunmuştur. En yüksek korelasyonlar KT1 ve KT2'den; bunlardan daha düşük korelasyon AUT'tan elde edilmiştir. Diğer bir deyişle, TIMSS 2007 testinin tamamından kestirilen yetenek parametreleri ile açık uçlu madde yüzdesine bakılmaksızın uzun testler arasındaki ilişkinin daha yüksek olduğuna ilişkin ipuçları bulunmuştur.

BÖLÜM 4

SONUÇ VE ÖNERİLER

Bu bölümde, araştırmadan elde edilen bulgulara dayalı sonuçlar tartışılmış ve öneriler sunulmuştur.

Sonuçlar

Araştırma, benzer yapıları ölçen iki madde tipinin (çoktan seçmeli ve yanıt sınırlı açık uçlu) ayrı ayrı (tek başına) ve birlikte kullanıldıkları koşullarda oluşturulan testlerde, madde ve yetenek parametreleri; madde ve test bilgi fonksiyonları ile görelî etkinlik indeksleri bakımından farklılıklar olup olmadığı ortaya konmaya çalışılmıştır. Bu amaç doğrultusunda, TIMSS 2007 matematik testi ikinci kitapçığındaki maddelerden test uzunluğu ($k=15$ ve 25) ve açık uçlu madde yüzdesi farklı (%20 ve %40) olan testlere ilişkin, madde ve yetenek parametreleri, madde ve test bilgi fonksiyonları, görelî etkinlik indeksleri kestirilmiştir. Araştırma sonunda elde edilen bulgulardan aşağıdaki sonuçlara ulaşılmıştır:

- 1- Madde güçlük parametrelerinin dağılımına göre, en güç test açık uçlu maddelerden oluşan testtir. Öte yandan, karma testlerde açık uçlu madde yüzdesinin farklılaşması karma testlerin ortalama madde güçlüğüne değiştirmemiştir. Madde ayıricılık parametreleri bakımından; açık uçlu madde yüzdesi %20 ve %40 olan uzun karma testlerden daha yüksek ayıricılık parametreleri elde edilmiştir. Sonuçta, uzun karma testlerin, sadece çoktan seçmeli ve sadece açık uçlu maddelerden oluşan testlere göre orta güçlükte ve daha ayırıcı testler olduğu sonucuna ulaşılmıştır.
- 2- Maddelerin düşük yeterlik düzeyindeki bireyler tarafından doğru yanıtlanma olasılıklarını gösteren madde şans parametreleri (c) karma testler için birbirine benzer sonuçlar vermiştir. Aynı özelliği

öçtüğü kanıtlanmış TIMSS 2007 Matematik testi maddelerinden, farklı test uzunluklarında ($k=15$ ve 25) ve farklı yüzdelerde açık uçlu maddelerden (%20 ve %40) oluşturulan dört karma test ile sadece çoktan seçmeli ve sadece açık uçlu maddelerden oluşturulan testlerden kestirilen yetenek parametreleri dağılımları benzer olmakla birlikte; yetenek parametrelerine ilişkin standart hata düzeyleri de göz önünde bulundurulduğunda, orta yetenek aralığında uzun karma testlerin daha az standart hatayla yetenek kestirimi yaptığı sonucuna ulaşılmıştır.

- 3- Test maddelerinden elde edilen bilgi düzeyini gösteren madde bilgi fonksiyonları, en fazla bilginin açık uçlu test maddelerinden sağlandığını ortaya koymuştur. Farklı testlerdeki açık uçlu maddelerden sağlanan ortalama bilginin, çoktan seçmeli maddelerden sağlanan bilgiden daha fazla olduğu sonucuna ulaşılmıştır.
- 4- Test bilgi fonksiyonlarının birbirilerine oranlanmasıyla elde edilen göreceli etkinlik (relative efficiency) değerleri, AUT'un madde sayısı az olmasına rağmen, daha uzun testlere göre daha etkili olduğunu göstermiştir. Ayrıca, açık uçlu madde yüzdesinin testin %40'ını oluşturduğu ve uzun karma testin açık uçlu test haricindeki diğer testlerden ölçülen özellik bakımından bireylerarası farklılıkları daha etkili ortaya koyduğu sonucuna ulaşılmıştır.
- 5- Genel olarak araştırma kapsamında düzenlenen tüm testlerin yetenek düzeyi ölçeğinin $(-3, +3)$ çoğunu kapsayacak şekilde bilgi verdiği belirlenmiştir. Öte yandan, açık uçlu madde yüzdesi yüksek kısa karma testin, alt yetenek düzeylerinde daha yüksek bilgi verdiği; açık uçlu testin ise üst yetenek düzeylerinde daha fazla bilgi verdiği sonucuna ulaşılmıştır.
- 6- Kestirilen yetenek parametreleri arasında hesaplanan Pearson Momentler Çarpımı Korelasyon katsayılarının tamamı pozitif yönde ve oldukça yüksek bulunmakla birlikte; açık uçlu madde yüzdesi

birbirine yakın olan testlerden kestirilen yetenek parametreleri arasındaki korelasyonun daha yüksek olduğu sonucuna ulaşılmıştır.

- 7- TIMSS 2007 8. sınıf Matematik Testi 2. Kitapçığını alan öğrencilerin bu testten hesaplanmış genel matematik puanları ile, 2. kitapçıktan bu araştırmanın amacı doğrultusunda düzenlenen altı testten kestirilen yetenek parametreleri arasındaki Pearson Momentler Çarpımı Korelasyon katsayıları pozitif yönde oldukça yüksek bulunmakla birlikte; TIMSS 2007 testinden kestirilen yetenek parametreleri ile açık uçlu madde yüzdesine bakılmaksızın uzun testler arasındaki ilişki daha yüksektir.

Öneriler

Bu araştırmadan elde edilen temel bulgu, ,test uzunluğu nispeten daha uzun olan karma testlerin sadece çoktan seçmeli maddelerden oluşan testlere göre daha ayırıcı olduğu, testle ölçülen özellik bakımından birey hakkında daha fazla bilgi verdiği ve daha düşük standart hatayla yetenek kestirimi yaptığıdır. Bu araştırmanın bulgularından hareketle sırasıyla uygulama alanına ve benzer konular kapsamında araştırma yapacaklara şu önerilerde bulunulabilir:

- 1- Geniş ölçekli test uygulamalarında kullanılacak testlerde, testin tamamının yanıtı sınırlı açık uçlu olmasının puanlamada yaratacağı güçlükler düşünüldüğünde, açık uçlu madde yüzdesi yarıya yakın ve test uzunluğu yüksek karma testlerin kullanılmasının testten elde edilen bilgi bakımından yarar sağlayacağı düşünülmektedir. Bu araştırmada kullanılan testlerde, yanıtı sınırlı açık uçlu madde formatı kullanılmıştı. Bu maddelerin yanıtları iki kategorili (0-1) olarak puanlanan türdendi. Araştırma sonuçları, YGS, LYS veya yine çoktan seçmeli maddelerden oluşan ÜDS, ALES gibi testlerde iki formatın

birlikte kullanılabileceği yönünde ipuçları vermiştir. Buna göre, bu tür testlerde iki formatın, birbirlerine yakın yüzdelerde kullanılması, hem çoktan seçmeli maddelerden gelen şans başarısı kaynaklı hataların puanlara karışmamasını önleyebilecek, nitelikli sorular olmak koşuluyla ölçülen zihinsel becerilerin düzeyi de yükseldebilir. Bu durumda bu testlerin sonuçlarına göre verilen seçme veya yeterli kararlarının daha geçerli ve güvenilir olması sağlanabilecektir.

Yapılacak diğer araştırmalar için de şu önerilerde bulunulabilir:

- 1- Bu çalışma 0-1 ile puanlanan açık uçlu maddelerin yer aldığı testler için yapılmıştır. Benzer çalışmalar kısmi puanlanan açık uçlu maddelerin bulunduğu karma testler için yapılabilir.
- 2- Benzer bir çalışma Matematik kadar önemli bir ders alanı olan Fen ve Teknoloji ders alanı için de gerçekleştirilebilir. Çünkü Fen ve teknoloji dersinde geliştirilmesi amaçlanan bilimsel süreç becerileri sadece çoktan seçmeli maddelerle ölçülememektedir. Bu türden beceriler açık uçlu maddelerle ölçülebilecek türdendir.
- 3- Benzer çalışmalar simülatif veri setlerinde yapılabilir. Buna göre test uzunluklarında, kısmi ve 0-1 ile puanlanan açık uçlu maddelerin farklı yüzdelerde yer aldığı testler farklı örneklem büyüklüklerinde çalışılabilir. Bu çalışmalardan elde edilecek sonuçlar kuramsal bilgilere ve uygulamalara da katkı getirebilecektir.
- 4- Yanıtı sınırlı açık uçlu ve çoktan seçmeli maddelerin ayrı ayrı ve birarada kullanıldığı test durumlarında, testlerin psikometrik niteliklerinin nasıl farklılaştığı bu çalışmayla ortaya konmaya çalışılmıştır. İleriki araştırmalarda karma testlerin bireylerin kaygı düzeyleri, kullandıkları öğrenme stratejileri, çoktan seçmeli ve açık uçlu maddelerin test içerisindeki sırasının farklılaşması gibi boyutlara ilişkin olarak testlerin psikometrik özelliklerinin ele alınması faydalı olacaktır.

KAYNAKLAR

Akdeniz, M. (2003), *Kısa Cevaplı ve Çoktan Seçmeli Testlerin Madde ve Test İstatistiklerinin Karşılaştırılması*. Yayımlanmamış yüksek lisans tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara

Araz, G. (2001), *Aynı Davranışı Ölçmeye Yönelik Kısa Cevaplı, Üç ve Beş Seçenekli Çoktan Seçmeli Testlerin Madde ve Test Özelliklerinin Şans Başarısı İle Birlikte İncelenmesi*. Yayımlanmamış yüksek lisans tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara

Arthur, Jr. W., Edwards, B. D., Barrett, G. V., (2002), Multiple-Choice and Constructed Response Tests of Ability: Race-Based Subgroup Performance Differences On Alternative Paper-And-Pencil Test Format. *Personnel Psychology*, 55, 985-1008.

Baker, F. B. (2001). *The Basis of Item Response Theory*. USA: ERIC Clearinghouse on Assessment and Evaluation.

Baker, F. B., Kim, S., (2004), *Item Response Theory Parameter Estimation Techniques*. New York:Marcel Dekker, Inc.

Bastari, B., (2000), *Linking Multiple-Choice and Constructed-Response Items to a Common Proficiency Scale*. Doctoral Dissertation. University of Massachusetts Amherst.

Berberoğlu, G., (2006), *Sınıf İçi Ölçme ve Değerlendirme Teknikleri*. İstanbul: Morpa Kültür Yayınları.

Berberoğlu, G., (2009), CİTO Türkiye Öğrenci İzleme Sistemi (ÖİS) Öğrenci Sosyal Gelişim Programı'na (ÖSGP) İlişkin Ön Bulgular, *CİTO Eğitim: Kuram ve Uygulama Dergisi*, Kasım-Aralık Sayısı, 32-42.

Büyüköztürk, Ş. (2004), *Sosyal Bilimler İçin Veri Analizi El Kitabı*, Ankara: PEGEM Yayıncılık.

Ceylan, E., Berberoğlu, G. (2007), Öğrencilerin Fen Başarısını Açıklayan Etmenler: Bir Modelleme Çalışması. *Eğitim ve Bilim*, Cilt 32, Sayı 144, 36-48.

Crocker, L. ve Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. N.Y.: CBS College Publishing Company.

Demirtaşlı, N. (2010). Açık uçlu soru formatı ve öğrenci izleme sistemi (ÖİS) akademik gelişimi izleme ve değerlendirme (AGİD) modülündeki kullanımı. *Cito Eğitim: Kuram ve Uygulama*. Nisan-Haziran, 21-30.

Embretson, S. E., Reise, S. P., (2000), *Item Response Theory For Psychologists*. New Jersey: Lawrence Erlbaum Associates, Publishers.

Ercikan, K., Schwarz, R.D., Julian, M.W., Burket, G.R., Weber, M.M., Link, V., (1998), Calibration and Scoring of Tests With Multiple-Choice and Constructed-Response Item Types. *Journal of Educational Measurement*, Vol. 35, No. 2, 137-154.

Garavaglia, D. R. (2001), *The Relationship Between Item Format And Cognitive Processes in Wide-Scale Assessment of Mathematics*. Doctoral Dissertation, Michigan State University.

Gonzales, P. (2008), *Highlights From TIMSS 2007, Mathematics and Science Achievement of U.S. Fourth and Eighth Grade Students in an International Context*, National Center for Education Statistics, USA. (<http://nces.ed.gov/pubs2009/2009001.pdf> adresinden 20.05.2009 tarihinde alınmıştır.)

Gözen, G. (2002), *Kısa Cevaplı ve Çoktan Seçmeli Maddelerin "1-0" ve Ağırlıklı Puanlama Yöntemleriyle Puanlanmasının Testin Psikometrik Özellikleri Açısından İncelenmesi*. Yayımlanmamış yüksek lisans tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara

Haladyna, T. M. (1997). *Writing Test Item to Evaluate Higher Order Thinking*. USA: Allyn & Bacon.

Hambleton, R. K., Swaminathan, H. (1985), *Item Response Theory. Principles and Applications*. Boston: Kluwer Academic Publishers.

Hambleton, R. K., Swaminathan, H., Rogers, H. (1991), *Fundamentals of Item Response Theory*. Newbury Park CA: Sage Publications.

IEA, (2005), TIMSS 2007 Assessment Frameworks, International Study Center, Lynch School of Education, Boston College: USA. (<http://timss.bc.edu/> adresinden 03.04.2009 tarihinde alınmıştır.)

IEA, (2007), TIMSS 2007 Questionnaires. (<http://timss.bc.edu/> adresinden 03.04.2009 tarihinde alınmıştır.)

IEA, (2008-1), TIMSS 2007 Technical Report, International Study Center, Lynch School of Education, Boston College: USA. (<http://timss.bc.edu/> adresinden 03.04.2009 tarihinde alınmıştır.)

IEA, (2008-2), TIMSS 2007 International Mathematics Report, International Study Center, Lynch School of Education, Boston College: USA. (<http://timss.bc.edu/> adresinden 03.04.2009 tarihinde alınmıştır.)

In'nami, Y., Koizumi, R. (2009), A Meta-Analysis Of Test Format Effects On Reading And Listening Test Performance: Focus On Multiple-Choice And Open-Ended Formats. *Language Testing*, 26 (2), 219-244.

Karaca, E. (2004), Seçme Gerektiren, Kısa Cevaplı ve Doğru - Yanlış Testlerinin Madde ve Test Özelliklerinin Karşılaştırılması. *Dumlupınar Üniversitesi Sosyal Bilimler Dergisi*, Sayı:10, 161-174.

Kim, Y. K. (2009), *Combining Constructed Response Items and Multiple Choice Items Using a Hierarchical Rater Model*. Columbia University. Doctoral Dissertation.

Kinsey, T. L. (2003), *A Comparison of IRT and Rasch Procedures in a Mixed-Item Format Test*. University of North Texas. Doctoral Dissertation.

Lord, F. M., Novick, M. R. (1968), *Statistical Theories of Mental Test Scores*. Reading, MA: Addison-Wesley.

Lord, F. M. (1983). Small n Justifies Rach Model in D. J. Weis (Ed.). *New Horizons in Testing: Latent Trait Test Theory And Computerized Adaptive Testing*. NY : Academic Pres.

Lukhele, R., Thissen, D., Wainer, H. (1994), On The Relative Value Of Multiple Choice, Contructed Response, And Examinee Selected İtems On Two Achievement Tests. *Journal Of Educational Measurement*, 31, 231-250.

Milli Eğitim Bakanlığı Talim Terbiye Kurulu Başkanlığı, (2005), *İlköğretim Matematik Dersi Öğretim Programı ve Kılavuzu 1-5. Sınıflar*. Ankara: Devlet Kitapları Müdürlüğü.

Milli Eğitim Bakanlığı Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı, (2007), *PISA 2006 Uluslararası Öğrenci Başarılarını Değerlendirme Programı Ulusal Ön Rapor*. (28.01.2009 tarihinde http://earged.meb.gov.tr/pisa/dokuman/2006/rapor/Pisa_2006_Ulusal_On_Rapor.pdf adresinden alınmıştır.)

Milli Eğitim Bakanlığı Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı, (2011), *TIMSS 2011 Tanıtım Kitapçığı*. (18.02.2011 tarihinde http://earged.meb.gov.tr/pdf/TIMSS_2011_kitapcigi.pdf adresinden alınmıştır.)

Özdemir, D., (1997), *Kısa Cevaplı, Çoktan Seçmeli ve Kompozisyon Tipi Maddelerden Oluşan Testlerin Psikometrik Özelliklerinin Karşılaştırmalı Olarak İncelenmesi*. Yayımlanmamış yüksek lisans tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara

Özdemir, D., (2004), *Çoktan Seçmeli Testlerin Klasik Test Teorisi Ve Örtük Özellikler Teorisine Göre Hesaplanan Psikometrik Özelliklerinin İki Kategorili ve Ağırlıklandırılmış Puanlanması Yönünden Karşılaştırılması*. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 26, 117-123.

Özkul, F., (2002), *Ağırlıklandırılmış Puanlamalı Kısa Cevaplı, “1-0” Puanlamalı Kısa Cevaplı ve Çoktan Seçmeli Testlerin Madde ve Test Özelliklerinin Karşılaştırılması*. Yayımlanmamış yüksek lisans tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara

Rodriguez, M. C. (2003), Construct Equivalence of Multiple-Choice and Constructed-Response Items: A Random Effects Synthesis of Correlations. *Journal Educational Measurement*, 40 (2), 163-184.

Schaeffer, G. A., Montero, D. H., Julian, M., Bené, N. H., (2002), A Comparison of Three Scoring Methods for Tests With Selected-Response and Constructed-Response Items. *Educational Assessment*, 8(4), 317–340

Shermis, M. D., Burstein, J. C., (2003), *Automated Essay Scoring: A Cross-Disciplinary Perspective*. Mahwah, NJ: Lawrence Erlbaum Associates.

Shin, D., (2007), *A Comparison of Method of Estimating Subscale Scores for Mixed-Format Tests*. Pearson Educational Measurement Research Reports. (24.12.2008 tarihinde www.pearsonedmeasurement.com/research/research.htm adresinden alınmıştır.)

Swaminathan, H., Hambleton, R. K., & Rogers, H. J. (2006). *Assessing the Fit of Item Response Theory Models*. (Center for Educational Assessment Research Report No. 541). Amherst, MA: University of Massachusetts, Center for Educational Assessment.

Sykes, R. C., Truskosky, D., White, H. (11-12 April 2001), *Determining The Representation of Constructed Response Items in Mixed-Item-Format Exams*. Paper presented at Annual Meeting of the National Council on Measurement in Education, Seattle

Tekin, H., (1991), *Eğitimde Ölçme ve Değerlendirme*. Ankara: Yargı Yayınları.

Tosyalı, S. (2002), *Yetenek Testlerinde Çoktan Seçmeli, Kısa Cevaplı ve Doğru-Yanlış Testlerin Madde ve Test Özelliklerinin Karşılaştırılması*. Yayımlanmamış yüksek lisans tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara

Traub, R. E., Fisher, C. W. (1977), On the Equivalence of Constructed-Response and Multiple-Choice Tests. *Applied Psychological Measurement*, 1(3), 355-369.

Uğurlu, N. (1996), *Aynı Davranışı Ölçmeye Yönelik Kısa Cevaplı ve Çoktan Seçmeli Maddelerden Oluşan Testlerin Psşkometrik Özelliklerinin İncelenmesi*. Yayınlanmamış yüksek lisans tezi. Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü, Ankara

Umay, A. (1997), Yanıtlayıcı Davranışların Analizi Yolu İle Matematikte Problem Çözümleri İçin Bir Güvenirlik ve Geçerlik Araştırması, *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*. 13, 47-56.

Uyeno, R. K. (2004), Assessing The Content Standarts of a Large-Scale, Standards-Based Test: A Psychometric Validty Study of The 2002 Hawai'i State Assessment Grade 8 and Grade 10 Reading Tests. University of Hawai'i. Doctoral Thesis.

Wainer, H., Thissen, D. (1993), Combining Multiple-Choice and Constructed-Response Test Scores: Toward a Marxist Theory of Test Construction. *Applied Measurement in Education*, 6(2), 103-118.

Wester, A., Henriksson, W. (2000), The Interaction Between Item Format and Gender Differences in Mathematics Performance Based on TIMSS Data. *Studies in Educational Evaluation*, 26, 79-90.

Zhao, Y. (2008). Approaches For Addressing The Fit Of Item Response Theory Models To Educational Test Data. University of Massachusetts: Doctoral Thesis.

EK 1: Araştırma Kapsamında Seçilen Maddelere İlişkin Bilgiler

MADDE KODU	ADI (ETİKETİ)	MADDE TÜRÜ	ÖĞRENME ALANI	BİLİŞSEL ALAN	KT1	KT2	KT3	KT4	AU	ÇS
M042003	Sayıların sıralanması (Ordering of number)	ÇS	SAYILAR	Bilgi	X	X	X	X		X
M042079	Basit sayılar (Simplify the number)	ÇS	SAYILAR	Bilgi	X	X				X
M042018	İfadeyi (+/-) olarak işaretleme (Place + or - sign in expression)	AU	SAYILAR	Uygulama	X				X	
M042055	Sınıftaki erkek çocuk sayısı (Number of boys in the class)	ÇS	SAYILAR	Uygulama	X	X	X	X		X
M042039	Ahmet'in ne kadar kar ettiği (How much alan save with discount)	ÇS	SAYILAR	Uygulama		X				X
M042199	Denk ifadeler (Equivalent expression)	ÇS	CEBİR	Bilgi	X	X	X	X		X
M042301A	Tabloyu doldur (Complete the table)	AU	CEBİR	Bilgi	X	X	X		X	
M042301B	On kenarlı poligonun iç açıları (Interior angles of 10-sided polygon)	AU	CEBİR	Akıl Yürütme	X	X	X	X	X	
M042301C	n kenarlı poligonun iç açıları (Interior angles of n-sided polygon)	AU	CEBİR	Akıl Yürütme	X	X			X	
M042265	Katlandığında bir küp oluşturan açık şekil (Which net when folded forms cube)	ÇS	GEOMETRİ	Akıl Yürütme	X	X	X	X		X
M042137	x+y'nin değeri kaçtır (What is the value of x + y)	ÇS	GEOMETRİ	Uygulama	X	X	X	X		X
M042148	Grafikteki noktalar (Point on the graph)	ÇS	GEOMETRİ	Bilgi	X	X	X	X		X
M042254	Grafikte verileri göstermek (Representing data on the graph)	ÇS	VERİ VE OLASILIK	Uygulama	X	X	X	X		X
M042250	Satılan bilet sayısı (Number of the tickets sold)	AU	VERİ VE OLASILIK	Bilgi	X	X	X		X	
M042220	Sütun grafiği çizme (Make a bar chart)	AU	VERİ VE OLASILIK	Uygulama	X		X		X	
M022097	Metrede voltajı gösterme (Voltage shown on meter)	ÇS	SAYILAR	Bilgi	X	X				X
M022101	En şanslı kutuyu seçme (Box with greater chance to pick 17)	ÇS	VERİ VE OLASILIK	Bilgi	X	X	X	X		X
M022104	Hangi parça en küçük (Which fraction is smallest)	ÇS	SAYILAR	Bilgi	X	X				X
M022105	p noktası etrafında döndürülmüş üçgen (Triangle rotated about point p)	ÇS	GEOMETRİ	Bilgi		X				X
M022106	Otobüsteki çocuk/yetişkin oranı (Ratio of children to adults on bus)	AU	SAYILAR	Uygulama	X		X		X	
M022108	Bütünler açının derecesi (Degree of angle from supplements)	ÇS	GEOMETRİ	Uygulama		X				X
M022110	İki ondalık sayıyı çarpma (Multiply two decimals)	AU	SAYILAR	Bilgi	X			X	X	
M022181	Tabloda gösterilen bilginin grafiği (Graph showing information in table)	ÇS	VERİ VE OLASILIK	Uygulama	X	X		X		X
M032307	10 zede satılan bilet payı (Fraction of tickets sold for 10 zeds)	AU	SAYILAR	Uygulama	X				X	
M032523	Şekeri tarifi 1 1/2'si kadar artırmak (Cups of sugar for 1 1/2 times recipe)	ÇS	SAYILAR	Uygulama		X				X
M032701	12 öğrenciye 1 öğretmen (1 teacher for every 12 students)	ÇS	SAYILAR	Uygulama	X	X	X	X		X
M032704	Bir otobüs sabit hızla gitmekte (A bus travels at a constant speed)	ÇS	SAYILAR	Uygulama		X				X
M032525	-6'ya bölündüğünde 12 veren sayı (The number divided by -6 gives 12)	ÇS	SAYILAR	Bilgi	X	X		X		X
M032579	Üçgende c açısının ölçümü (The measure of angle c in triangle)	ÇS	GEOMETRİ	Uygulama	X	X		X		X
M032691	Bütünler açı çizme (Draw angles that are supplementary)	AU	GEOMETRİ	Uygulama	X	X	X	X	X	

ÇS: 20 , AU: 11 ; Sayılar : 14 , Cebir: 5 , Geometri: 7 , Veri ve olasılık: 4 ; Bilgi: 12 , Uygulama: 15 , Akıl yürütme: 4

EK 2

Arařtırma Kapsamında Ele Alınan Testler

(AUT, ST, KT1, KT2, KT3, KT4)

AÇIK UÇLU TEST

1. Aşağıdaki karelerden her birinin içine, bu ifadenin mümkün olan en yüksek değerde olmasını sağlayacak şekilde + ya da - işareti koyunuz.

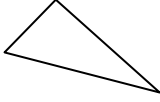
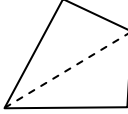
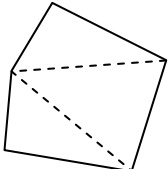
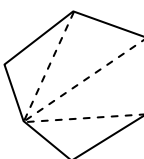
$$-5 \square -6 \square 3 \square -9$$

2-4. maddeler aşağıdaki metne bağlı olarak yanıtlanacaktır.

İç Açılar

Kağan, çokgenlerin özelliklerini inceliyordu. Çokgenlerde kenarlar ve açılar arasında bir bağıntı olup olmadığını nortaya çıkarmak için aşağıdaki tabloyu hazırladı.

2. Tablodaki boş yerleri tamamlayınız.

Çokgen	Kenar Sayısı	Üçgen Sayısı	İç Açılar Top.
	3	1	$1 \times 180^\circ$
	—	—	$__ \times 180^\circ$
	—	—	$__ \times 180^\circ$
	—	—	$__ \times 180^\circ$

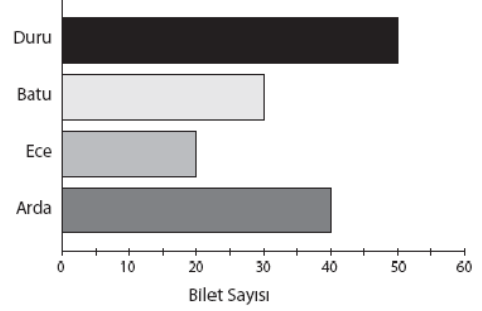
3. Aşağıdaki eşitlikte verilen karenin içine doğru sayıyı yazınız.

10 kenarlı bir çokgenin iç açıları toplamı = $\square \times 180^\circ$

4. Kağan bir örüntü bulabildi ve n simgesini kullanarak her çokgen için doğru olan bir bağıntı (formül) yazdı. Kağan'ın yazdığı aşağıdaki formülü tamamlayınız.

n kenarlı bir çokgenin iç açıları toplamı = $__ \times 180^\circ$

- 5.

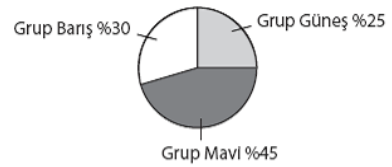


Duru, Batu, Ece ve Arda okul konserleri için bilet sattılar. Bunlardan her birinin satmış olduğu biletlerin sayısı yukarıdaki grafikte görülmektedir. Bu dört kişiden iksinin sattığı biletlerin sayısı Duru'nun tek başına sattığı bilet sayısı kadardır.

Bu iki kişi hangileridir?

Yanıt: _____ ve _____

6. Aşağıdaki daire grafiğinde, 200 öğrenci üzerinde yapılan bir araştırmanın sonuçları görülmektedir.



Aşağıda ayrılan yere, daire grafiğinin her bölümündeki öğrenci sayısını gösteren bir sütun grafiği çiziniz.



7. Bir otobüste 36 yolcu var. otobüsteki çocukların yetişkinlere oranı 5 çocuğa 4 yetişkindir. Bu otobüste kaç çocuk vardır?

Yanıt: _____

8. $0,402 \times 0,53 =$

Yanıt: _____

9. Bir konserin biletleri 10 zed, 15 zed ve 30 zed fiyata satılmaktadır. Satılan 900 biletten $\frac{1}{5}$ 'i 30 zed'lik, $\frac{2}{3}$ 'ü 15 zed'lik biletlerdir.

Buna göre satılan biletlerin KESİR olarak ne kadarı 10 zed'dir?

Yanıt: _____

10. Aşağıdaki AO doğru parçasını kullanarak AOB dar, AOC geniş açı olacak şekilde O'dan geçen düz bir BC doğrusu çizin. B ve C noktalarını işaretleyiniz.



YANIT ANAHTARI

MADDE	DOĞRU YANIT
1	-, +, -
2	4 2 2 5 3 3 6 4 4
3	8
4	n-2 köşeli (parantezli veya parantezsiz)
5	Ece ve Batu
6	Üçü de doğru – (50,90,60) 50'nin doğru çizgide olması gerekir. 90, 100'den küçük fakat 80'den büyük olmalıdır. 60, 70'ten küçük fakat 50 den büyük olmalıdır.
7	20
8	0,21306 0,21306 öğrencinin bir parçası olarak gösterilmiş ve daha sonra yanlış veya doğru olarak yuvarlanmış.
9	$\frac{2}{5}$ veya dengi
10	O dan geçen doğru çizilmiş; dar açı ve geniş açı adlandırılmış.

ÇOKTAN SEÇMELİ TEST

1. Aşağıdaki seçeneklerin hangisinde verilen sayılar, BÜYÜKTEN KÜÇÜĞE doğru sıralanmıştır?

A) 10,001; 10,110; 11,001; 11,100
 B) 10,110; 10,011; 11,100; 11,011
 C) 11,001; 11,100; 10,110; 10,011
 D) 11,100; 11,001; 10,110; 10,011

2. $3,4 \times 10^2$ işleminin sonucu kaçtır?

A) 3,4 B) 34 C) 340 D) 3400

3. Bir sınıftaki öğrenci sayısı 30'dur. Bu sınıftaki erkek öğrencilerin sayısını kız öğrencilerin sayısına oranı $\frac{2}{3}$ 'tür. Sınıftaki erkek öğrenci sayısı kaçtır?

A) 6 B) 12 C) 18 D) 20

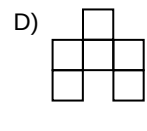
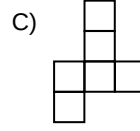
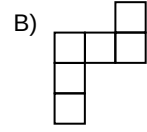
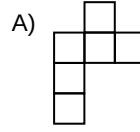
4. Bir paltonun fiyatı 60 zed'dir. Ahmet bu paltıyı satış fiyatı üzerinden %30 indirim yapıldığı bir zamanda almıştır. Ahmet bu paltıyı kaç zed ucuza almıştır?

A) 18 zed B) 24 zed
 C) 30 zed D) 42 zed

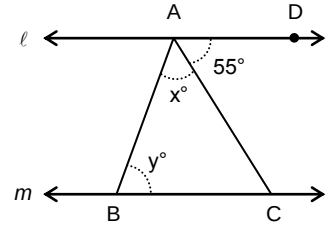
5. $4x - x + 7y - 2y$ ifadesine aşağıdakilerden hangisi eşittir?

A) 9 B) $9xy$
 C) $4 + 5y$ D) $3x + 5y$

6. Aşağıda verilen açık şekillerden hangisi katlandığında bir küp oluşturur?



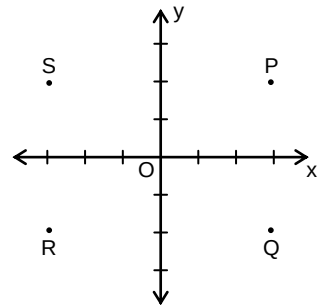
7.



Bu şekilde, ℓ doğrusu m doğrusuna paraleldir. DAC açısının ölçüsü 55° dir. $x + y$ kaç derecedir?

A) 55 B) 110 C) 125 D) 135

8.



Grafikteki hangi nokta $(3, -2)$ noktasıdır?

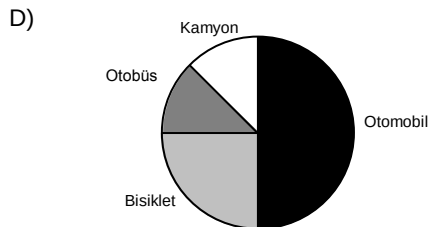
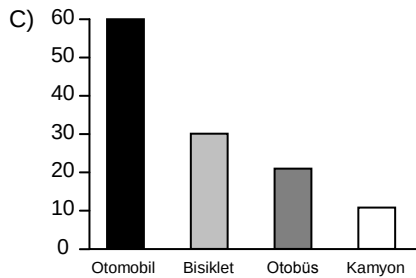
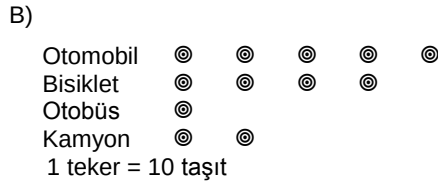
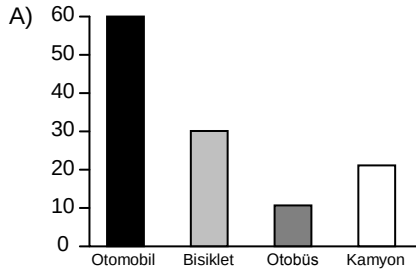
A) P B) Q C) R D) D

9. 4 öğrenci, 1 saat boyunca okullarının önündeki yoldan geçen taşıtları incelemişlerdir.

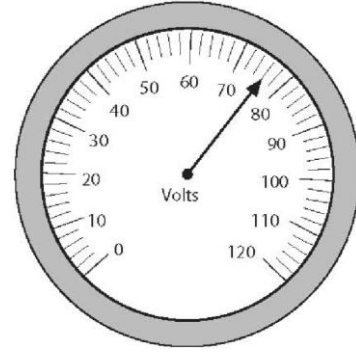
Tablo onların gördükleri taşıtları göstermektedir.

Taşıt Cinsi	Sayı
Otomobil	60
Bisiklet	30
Otobüs	10
Kamyon	20

Tablodaki sonuçları göstermek için her öğrenci bir grafik çizmiştir. Aşağıda verilen bu grafiklerin hangisinde sonuçlar doğru olarak gösterilmiştir?

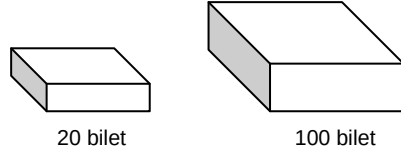


10. Aşağıdaki voltmetrede okunan değer kaç voltur?



- A) 73 B) 74 C) 76 D) 78

11. Küçük kutuda 1'den 20'ye kadar numaralanmış bilet; büyük kutuda 1'den 100'e kadar numaralanmış 100 bilet vardır.



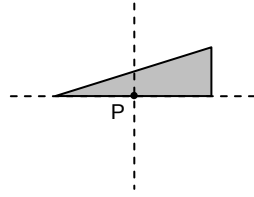
Kutuların içindeki biletlere bakmadan her kutudan bir bilet çekiliyor. Hangi kutudan 17 numaralı biletin çekilme olasılığı daha fazladır?

- A) İçinde 20 bilet olan kutu
 B) İçinde 100 bilet olan kutu
 C) İki kutu için de olasılık aynıdır.
 D) Bu soruya cevap verilemez.

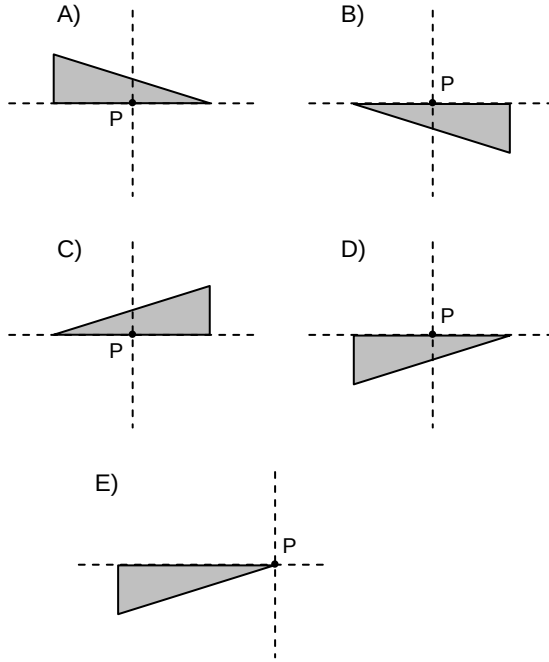
12. Aşağıdaki sayıların hangisi EN KÜÇÜKTÜR?

- A) $\frac{1}{2}$ B) $\frac{5}{8}$ C) $\frac{5}{6}$ D) $\frac{5}{12}$

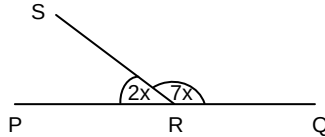
13. Taralı şekle, kağıt düzlemindeki P noktası etrafında yarım dönüş yaptırılıyor.



Bu dönüşle elde edilecek sonuç aşağıdakilerden hangisidir?



14. Aşağıdaki şekilde PQ bir doğru parçasını göstermektedir.



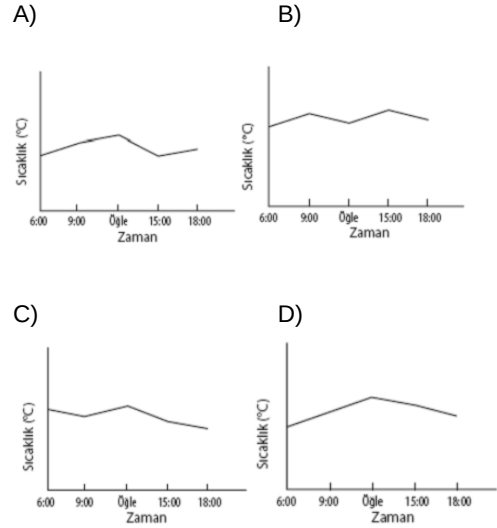
PRS açısının ölçüsü kaç derecedir?

- A) 10° B) 20° C) 40°
D) 70° E) 140°

15. Tabloda bir gün içindeki değişik zamanlarda ölçülen hava sıcaklıkları görülmektedir.

Zaman	6:00	9:00	Öğle	15:00	18:00
Sıcaklık °C	12	17	14	18	15

Bu sonuçlar, sıcaklık ölçeği olmayan bir grafikte gösteriliyor. Bu grafik aşağıdakilerden hangisi olabilir?



16. Deniz, tarifte verilmiş olan miktarın 1,5 katı kadar bir meyveli kek yapacaktır.

Kek tarifinde bir su bardağının $\frac{3}{4}$ ü kadar şeker konacağı söylenmişse Deniz'in yapacağı kek için kaç su bardağı şeker gerekir?

- A) $\frac{3}{8}$ B) $1\frac{1}{8}$ C) $1\frac{1}{4}$ D) $1\frac{3}{8}$

17. Okulun düzenlediği gezilere her 12 öğrenci için 1 öğretmen katılmaktadır. Geziye 108 öğrenci gidiyorsa bu geziye kaç öğretmen katılacaktır?

- A) 7 B) 8 C) 9 D) 10

18. Bir otobüs sabit bir hızla hareket etmektedir. Yolculuk için harcana zaman ile alınan yol doğru orantılıdır. Bu otobüs 5 saatte 120 km yol almışsa 8 saatte kaç km yol alır?

A) 168 B) 192
C) 200 D) 245

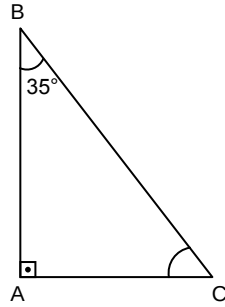
19. Hangi sayı -6 sayısına bölünürse sonuç 12 olur?

A) -72 B) -2 C) 2 D) 72

YANIT ANAHTARI

SORU NO	DOĞRU YANIT
1	D
2	C
3	B
4	A
5	D
6	C
7	C
8	B
9	A
10	C
11	A
12	D
13	D
14	C
15	B
16	B
17	C
18	B
19	A
20	B

20.



Yukarıdaki üçgende C açısı kaç derecedir?

A) 45° B) 55° C) 65° D) 145°

KARMA TEST 1

1. Aşağıdaki seçeneklerin hangisinde verilen sayılar, BÜYÜKTEN KÜÇÜĞE doğru sıralanmıştır?

A) 10,001; 10,110; 11,001; 11,100
 B) 10,110; 10,011; 11,100; 11,011
 C) 11,001; 11,100; 10,110; 10,011
 D) 11,100; 11,001; 10,110; 10,011

2. $3,4 \times 10^2$ işleminin sonucu kaçtır?

A) 3,4 B) 34 C) 340 D) 3400

3. Aşağıdaki karelerden her birinin içine, bu ifadenin mümkün olan en yüksek değerde olmasını sağlayacak şekilde + ya da - işareti koyunuz.

-5 ☐ -6 ☐ 3 ☐ -9

4. Bir sınıftaki öğrenci sayısı 30'dur. Bu sınıftaki erkek öğrencilerin sayısını kız öğrencilerin sayısına oranı $\frac{2}{3}$ 'tür. Sınıftaki erkek öğrenci sayısı kaçtır?

A) 6 B) 12 C) 18 D) 20

5. $4x - x + 7y - 2y$ ifadesine aşağıdakilerden hangisi eşittir?

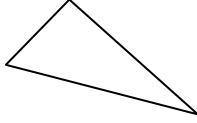
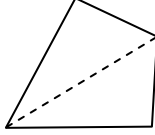
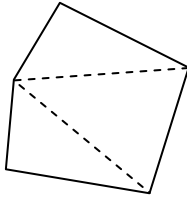
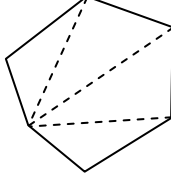
A) 9 B) $9xy$
 C) $4 + 5y$ D) $3x + 5y$

6-8. maddeler aşağıdaki metne bağlı olarak yanıtlanacaktır.

İç Açılar

Kağan, çokgenlerin özelliklerini inceliyordu. Çokgenlerde kenarlar ve açılar arasında bir bağıntı olup olmadığını nortaya çıkarmak için aşağıdaki tabloyu hazırladı.

6. Tablodaki boş yerleri tamamlayınız.

Çokgen	Kenar Sayısı	Üçgen Sayısı	İç Açılar Top.
	3	1	$1 \times 180^\circ$
	—	—	$__ \times 180^\circ$
	—	—	$__ \times 180^\circ$
	—	—	$__ \times 180^\circ$

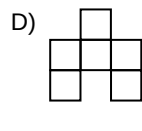
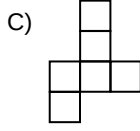
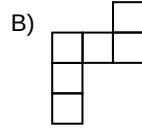
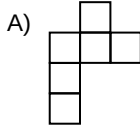
7. Aşağıdaki eşitlikte verilen karenin içine doğru sayıyı yazınız.

10 kenarlı bir çokgenin iç açıları toplamı = $\times 180^\circ$

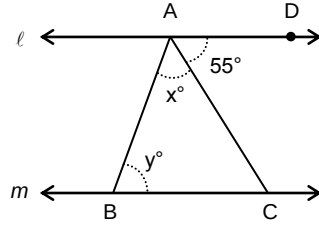
8. Kağan bir örüntü bulabildi ve n simgesini kullanarak her çokgen için doğru olan bir bağıntı (formül) yazdı. Kağan'ın yazdığı aşağıdaki formülü tamamlayınız.

n kenarlı bir çokgenin iç açıları toplamı = $__ \times 180^\circ$

9. Aşağıda verilen açık şekillerden hangisi katlandığında bir küp oluşturur?



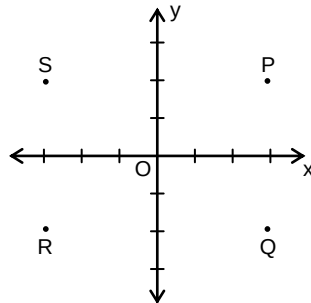
10.



Bu şekilde, ℓ doğrusu m doğrusuna paraleldir. $\angle DAC$ açısının ölçüsü 55° dir. $x + y$ kaç derecedir?

- A) 55 B) 110 C) 125 D) 135

11.



Grafikteki hangi nokta $(3, -2)$ noktasıdır?

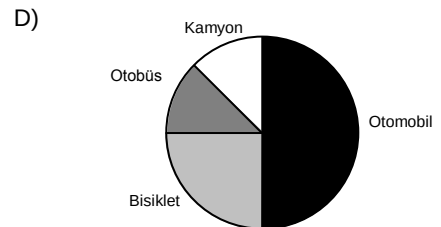
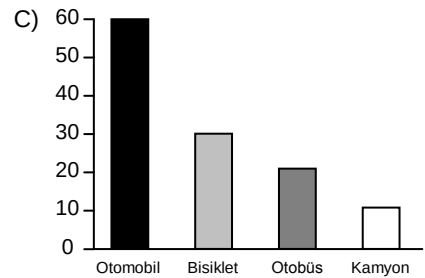
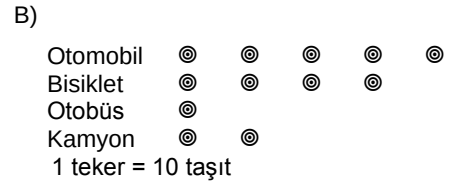
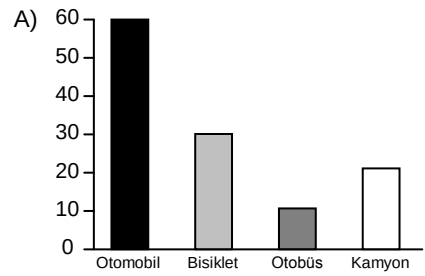
- A) P B) Q C) R D) D

12. 4 öğrenci, 1 saat boyunca okullarının önündeki yoldan geçen taşıtları incelemişlerdir.

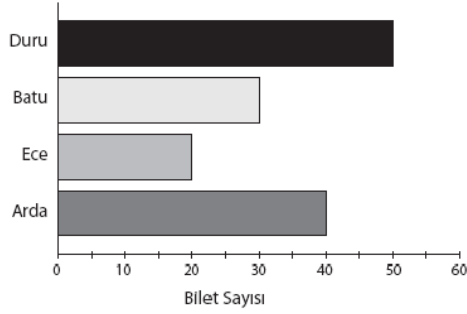
Tablo onların gördükleri taşıtları göstermektedir.

Taşıt Cinsi	Sayı
Otomobil	60
Bisiklet	30
Otobüs	10
Kamyon	20

Tablodaki sonuçları göstermek için her öğrenci bir grafik çizmiştir. Aşağıda verilen bu grafiklerin hangisinde sonuçlar doğru olarak gösterilmiştir?



13.

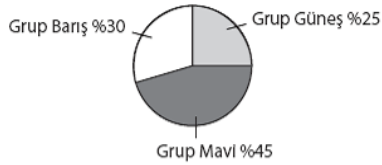


Duru, Batu, Ece ve Arda okul konserleri için bilet sattılar. Bunlardan her birinin satmış olduğu biletlerin sayısı yukarıdaki grafikte görülmektedir. Bu dört kişiden iksinin sattığı biletlerin sayısı Duru'nun tek başına sattığı bilet sayısı kadardır.

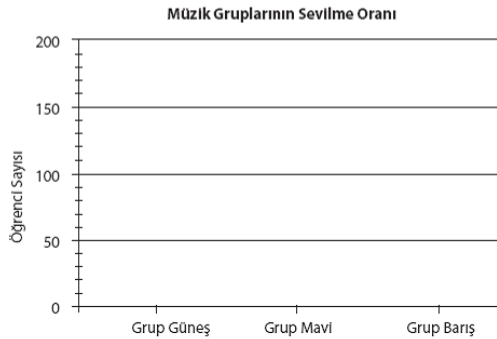
Bu iki kişi hangileridir?

Yanıt: _____ ve _____

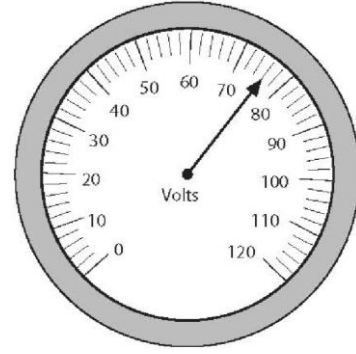
14. Aşağıdaki daire grafiğinde, 200 öğrenci üzerinde yapılan bir araştırmanın sonuçları görülmektedir.



Aşağıda ayrılan yere, daire grafiğinin her bölümündeki öğrenci sayısını gösteren bir sütun grafiği çiziniz.

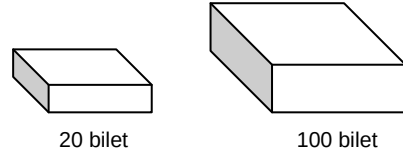


15. Aşağıdaki voltmetrede okunan değer kaç volttur?



A) 73 B) 74 C) 76 D) 78

16. Küçük kutuda 1'den 20'ye kadar numaralanmış bilet; büyük kutuda 1'den 100'e kadar numaralanmış 100 bilet vardır.



Kutuların içindeki biletlere bakmadan her kutudan bir bilet çekiliyor. Hangi kutudan 17 numaralı biletin çekilme olasılığı daha fazladır?

A) İçinde 20 bilet olan kutu
B) İçinde 100 bilet olan kutu
C) İki kutu için de olasılık aynıdır.
D) Bu soruya cevap verilemez.

17. Aşağıdaki sayıların hangisi EN KÜÇÜKTÜR?

A) $\frac{1}{2}$ B) $\frac{5}{8}$ C) $\frac{5}{6}$ D) $\frac{5}{12}$

18. Bir otobüste 36 yolcu var. otobüsteki çocukların yetişkinlere oranı 5 çocuğa 4 yetişkindir. Bu otobüste kaç çocuk vardır?

Yanıt: _____

19. $0,402 \times 0,53 =$

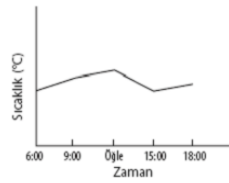
Yanıt: _____

20. Tabloda bir gün içindeki değişik zamanlarda ölçülen hava sıcaklıkları görülmektedir.

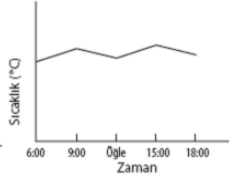
Zaman	6:00	9:00	Öğle	15:00	18:00
Sıcaklık°C	12	17	14	18	15

Bu sonuçlar, sıcaklık ölçeği olmayan bir grafikte gösteriliyor. Bu grafik aşağıdakilerden hangisi olabilir?

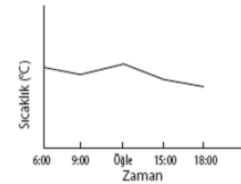
A)



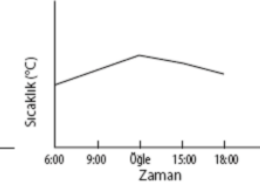
B)



C)



D)



21. Bir konserin biletleri 10 zed, 15 zed ve 30 zed fiyata satılmaktadır. Satılan 900 biletten $\frac{1}{5}$ 'i 30 zed'lik, $\frac{2}{3}$ 'ü 15 zed'lik biletlerdir.

Buna göre satılan biletlerin KESİR olarak ne kadarı 10 zed'dir?

Yanıt: _____

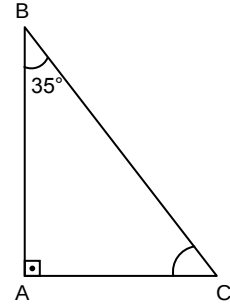
22. Okulun düzenlediği gezilere her 12 öğrenci için 1 öğretmen katılmaktadır. Geziye 108 öğrenci gidiyorsa bu geziye kaç öğretmen katılacaktır?

A) 7 B) 8 C) 9 D) 10

23. Hangi sayı -6 sayısına bölünürse sonuç 12 olur?

A) -72 B) -2 C) 2 D) 72

24.



Yukarıdaki üçgende C açısı kaç derecedir?

A) 45° B) 55° C) 65° D) 145°

25. Aşağıdaki AO doğru parçasını kullanarak AOB dar, AOC geniş açı olacak şekilde O'dan geçen düz bir BC doğrusu çizin. B ve C noktalarını işaretleyiniz.



YANIT ANAHTARI

SORU NO	DOĞRU YANIT
1	D
2	C
3	-, +, -
4	B
5	D
6	4 2 2 5 3 3 6 4 4
7	8
8	n-2 köşeli (parantezli veya parantezsiz)
9	C
10	C
11	B
12	A
13	Ece ve Batu
14	Üçü de doğru – (50,90,60) 50'nin doğru çizgide olması gerekir. 90, 100'den küçük fakat 80'den büyük olmalıdır. 60, 70'ten küçük fakat 50 den büyük olmalıdır.
15	C
16	A
17	D
18	20
19	0,21306 0,21306 öğrencinin bir parçası olarak gösterilmiş ve daha sonra yanlış veya doğru olarak yuvarlanmış.
20	B
21	$\frac{2}{5}$ veya dengi
22	C
23	A
24	B
25	O dan geçen doğru çizilmiş; dar açı ve geniş açı adlandırılmış.

KARMA TEST 2

1. Aşağıdaki seçeneklerin hangisinde verilen sayılar, BÜYÜKTEN KÜÇÜĞE doğru sıralanmıştır?

A) 10,001; 10,110; 11,001; 11,100
 B) 10,110; 10,011; 11,100; 11,011
 C) 11,001; 11,100; 10,110; 10,011
 D) 11,100; 11,001; 10,110; 10,011

2. $3,4 \times 10^2$ işleminin sonucu kaçtır?

A) 3,4 B) 34 C) 340 D) 3400

3. Bir sınıftaki öğrenci sayısı 30'dur. Bu sınıftaki erkek öğrencilerin sayısını kız öğrencilerin sayısına oranı $\frac{2}{3}$ 'tür. Sınıftaki erkek öğrenci sayısı kaçtır?

A) 6 B) 12 C) 18 D) 20

4. Bir paltonun fiyatı 60 zed'dir. Ahmet bu paltoyu satış fiyatı üzerinden %30 indirim yapıldığı bir zamanda almıştır. Ahmet bu paltıyı kaç zed ucuza almıştır?

A) 18 zed B) 24 zed
 C) 30 zed D) 42 zed

5. $4x - x + 7y - 2y$ ifadesine aşağıdakilerden hangisi eşittir?

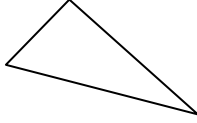
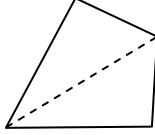
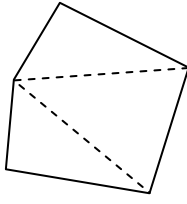
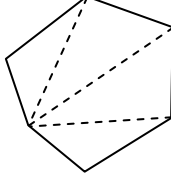
A) 9 B) $9xy$
 C) $4 + 5y$ D) $3x + 5y$

6-8. maddeler aşağıdaki metne bağlı olarak yanıtlanacaktır.

İç Açılar

Kağan, çokgenlerin özelliklerini inceliyordu. Çokgenlerde kenarlar ve açılar arasında bir bağıntı olup olmadığını nortaya çıkarmak için aşağıdaki tabloyu hazırladı.

6. Tablodaki boş yerleri tamamlayınız.

Çokgen	Kenar Sayısı	Üçgen Sayısı	İç Açılar Top.
	3	1	$1 \times 180^\circ$
	—	—	$__ \times 180^\circ$
	—	—	$__ \times 180^\circ$
	—	—	$__ \times 180^\circ$

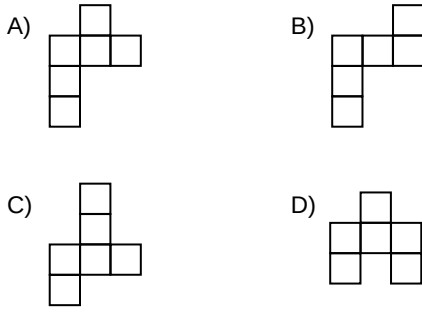
7. Aşağıdaki eşitlikte verilen karenin içine doğru sayıyı yazınız.

10 kenarlı bir çokgenin iç açıları toplamı = $\square \times 180^\circ$

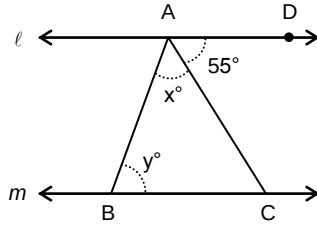
8. Kağan bir örüntü bulabildi ve n simgesini kullanarak her çokgen için doğru olan bir bağıntı (formül) yazdı. Kağan'ın yazdığı aşağıdaki formülü tamamlayınız.

n kenarlı bir çokgenin iç açıları toplamı = $__ \times 180^\circ$

9. Aşağıda verilen açık şekillerden hangisi katlandığında bir küp oluşturur?



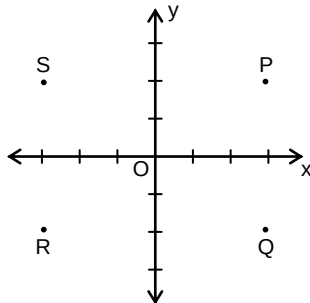
10.



Bu şekilde, ℓ doğrusu m doğrusuna paraleldir. DAC açısının ölçüsü 55° dir. $x + y$ kaç derecedir?

- A) 55 B) 110 C) 125 D) 135

11.



Grafikteki hangi nokta $(3, -2)$ noktasıdır?

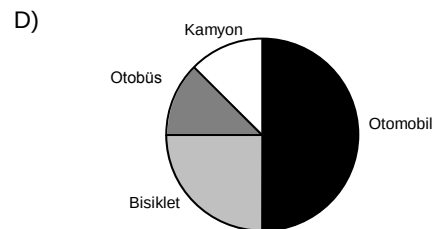
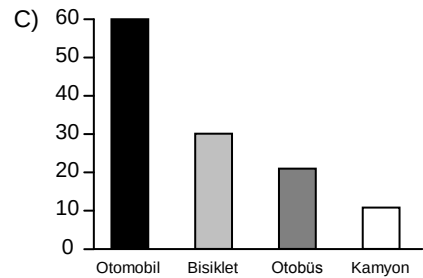
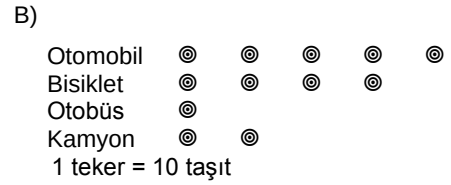
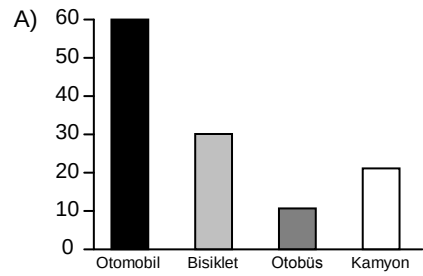
- A) P B) Q C) R D) D

12. 4 öğrenci, 1 saat boyunca okullarının önündeki yoldan geçen taşıtları incelemişlerdir.

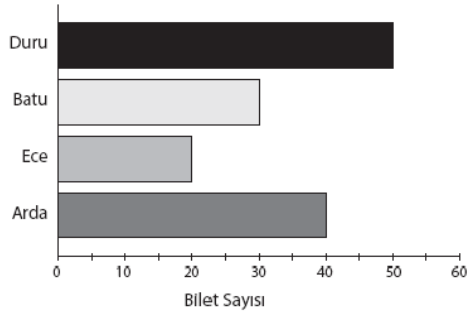
Tablo onların gördükleri taşıtları göstermektedir.

Taşıt Cinsi	Sayı
Otomobil	60
Bisiklet	30
Otobüs	10
Kamyon	20

Tablodaki sonuçları göstermek için her öğrenci bir grafik çizmiştir. Aşağıda verilen bu grafiklerin hangisinde sonuçlar doğru olarak gösterilmiştir?



13.

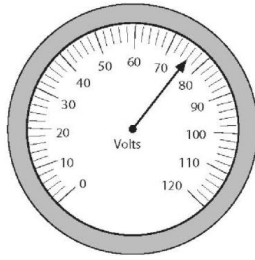


Duru, Batu, Ece ve Arda okul konserleri için bilet sattılar. Bunlardan her birinin satmış olduğu biletlerin sayısı yukarıdaki grafikte görülmektedir. Bu dört kişiden iksinin sattığı biletlerin sayısı Duru'nun tek başına sattığı bilet sayısı kadardır.

Bu iki kişi hangileridir?

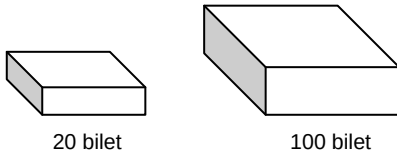
Yanıt: _____ ve _____

14. Aşağıdaki voltmetrede okunan değer kaç voltur?



A) 73 B) 74 C) 76 D) 78

15. Küçük kutuda 1'den 20'ye kadar numaralanmış bilet; büyük kutuda 1'den 100'e kadar numaralanmış 100 bilet vardır.



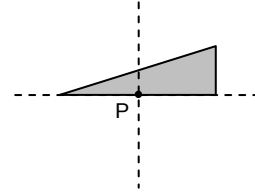
Kutuların içindeki biletlere bakmadan her kutudan bir bilet çekiliyor. Hangi kutudan 17 numaralı biletin çekilme olasılığı daha fazladır?

- A) İçinde 20 bilet olan kutu
 B) İçinde 100 bilet olan kutu
 C) İki kutu için de olasılık aynıdır.
 D) Bu soruya cevap verilemez.

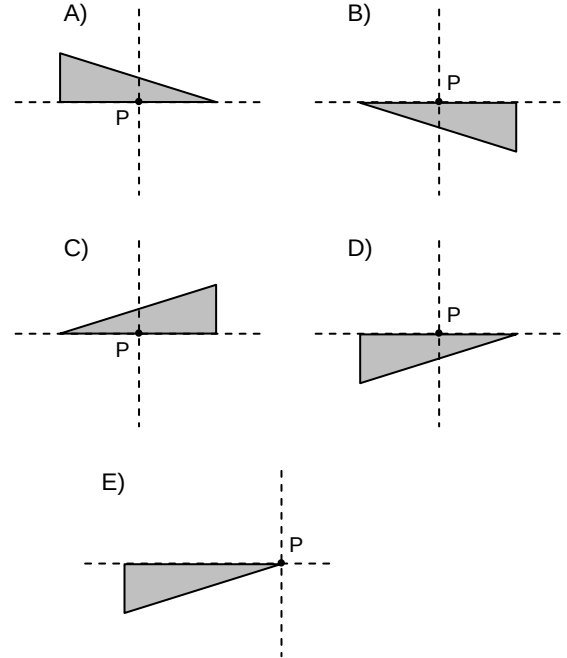
16. Aşağıdaki sayıların hangisi EN KÜÇÜKTÜR?

- A) $\frac{1}{2}$ B) $\frac{5}{8}$ C) $\frac{5}{6}$ D) $\frac{5}{12}$

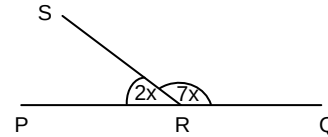
17. Taralı şekle, kağıt düzlemindeki P noktası etrafında yarım dönüş yaptırılıyor.



Bu dönüşle elde edilecek sonuç aşağıdakilerden hangisidir?



18. Aşağıdaki şekilde PQ bir doğru parçasını göstermektedir.



PRS açısının ölçüsü kaç derecedir?

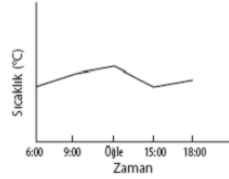
- A) 10° B) 20° C) 40°
 D) 70° E) 140°

19. Tabloda bir gün içindeki değişik zamanlarda ölçülen hava sıcaklıkları görülmektedir.

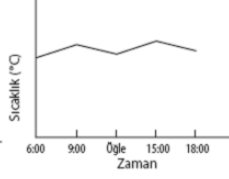
Zaman	6:00	9:00	Öğle	15:00	18:00
Sıcaklık°C	12	17	14	18	15

Bu sonuçlar, sıcaklık ölçeği olmayan bir grafikte gösteriliyor. Bu grafik aşağıdakilerden hangisi olabilir?

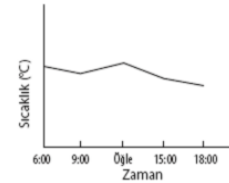
A)



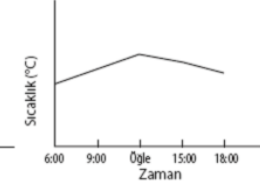
B)



C)



D)



20. Deniz, tarifte verilmiş olan miktarın 1,5 katı kadar bir meyveli kek yapacaktır. Kek tarifinde bir su bardağının $\frac{3}{4}$ ü kadar şeker konacağı söylenmişse Deniz'in yapacağı kek için kaç su bardağı şeker gerekir?

A) $\frac{3}{8}$ B) $1\frac{1}{8}$ C) $1\frac{1}{4}$ D) $1\frac{3}{8}$

21. Okulun düzenlediği gezilere her 12 öğrenci için 1 öğretmen katılmaktadır. Geziye 108 öğrenci gidiyorsa bu geziye kaç öğretmen katılacaktır?

A) 7 B) 8 C) 9 D) 10

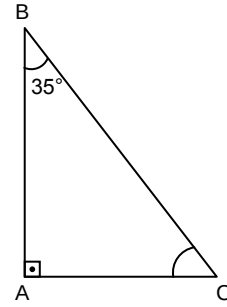
22. Bir otobüs sabit bir hızla hareket etmektedir. Yolculuk için harcama zaman ile alınan yol doğru orantılıdır. Bu otobüs 5 saatte 120 km yol almışsa 8 saatte kaç km yol alır?

A) 168 B) 192
C) 200 D) 245

23. Hangi sayı -6 sayısına bölünürse sonuç 12 olur?

A) -72 B) -2 C) 2 D) 72

- 24.



Yukarıdaki üçgende C açısı kaç derecedir?

A) 45° B) 55° C) 65° D) 145°

25. Aşağıdaki AO doğru parçasını kullanarak AOB dar, AOC geniş açı olacak şekilde O'dan geçen düz bir BC doğrusu çizin. B ve C noktalarını işaretleyiniz.



YANIT ANAHTARI

SORU NO	DOĞRU YANIT
1	D
2	C
3	B
4	A
5	D
6	4 2 2 5 3 3 6 4 4
7	8
8	n-2 köşeli (parantezli veya parantezsiz)
9	C
10	C
11	B
12	A
13	Ece ve Batu
14	C
15	A
16	D
17	D
18	C
19	B
20	B
21	C
22	B
23	A
24	B
25	O dan geçen doğru çizilmiş; dar açı ve geniş açı adlandırılmış.

KARMA TEST 3

1. Aşağıdaki seçeneklerin hangisinde verilen sayılar, BÜYÜKTEN KÜÇÜĞE doğru sıralanmıştır?

A) 10,001; 10,110; 11,001; 11,100
 B) 10,110; 10,011; 11,100; 11,011
 C) 11,001; 11,100; 10,110; 10,011
 D) 11,100; 11,001; 10,110; 10,011

2. Bir sınıftaki öğrenci sayısı 30'dur. Bu sınıftaki erkek öğrencilerin sayısını kız öğrencilerin sayısına oranı $\frac{2}{3}$ 'tür. Sınıftaki erkek öğrenci sayısı kaçtır?

A) 6 B) 12 C) 18 D) 20

3. $4x - x + 7y - 2y$ ifadesine aşağıdakilerden hangisi eşittir?

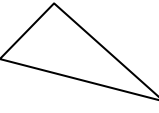
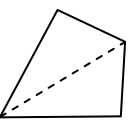
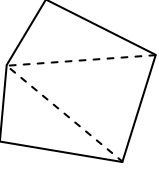
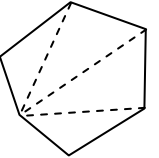
A) 9 B) $9xy$
 C) $4 + 5y$ D) $3x + 5y$

4-5. maddeler aşağıdaki metne bağlı olarak yanıtlanacaktır.

İç Açılar

Kağan, çokgenlerin özelliklerini inceliyordu. Çokgenlerde kenarlar ve açılar arasında bir bağıntı olup olmadığını ortaya çıkarmak için aşağıdaki tabloyu hazırladı.

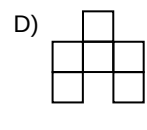
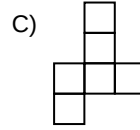
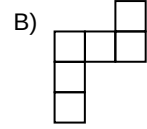
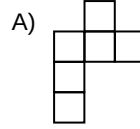
4. Tablodaki boş yerleri tamamlayınız.

Çokgen	Kenar Sayısı	Üçgen Sayısı	İç Açılar Top.
	3	1	$1 \times 180^\circ$
	—	—	$— \times 180^\circ$
	—	—	$— \times 180^\circ$
	—	—	$— \times 180^\circ$

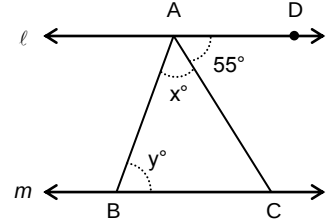
5. Aşağıdaki eşitlikte verilen karenin içine doğru sayıyı yazınız.

10 kenarlı bir çokgenin iç açıları toplamı = $\square \times 180^\circ$

6. Aşağıda verilen açık şekillerden hangisi katlandığında bir küp oluşturur?



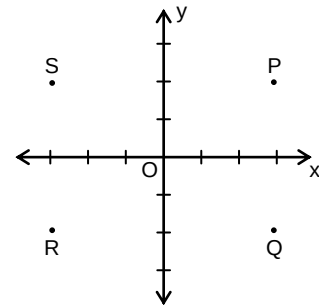
- 7.



Bu şekilde, ℓ doğrusu m doğrusuna paraleldir. DAC açısının ölçüsü 55° dir. $x + y$ kaç derecedir?

A) 55 B) 110 C) 125 D) 135

- 8.



Grafikteki hangi nokta $(3, -2)$ noktasıdır?

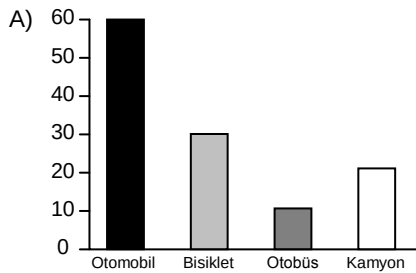
A) P B) Q C) R D) D

9. 4 öğrenci, 1 saat boyunca okullarının önündeki yoldan geçen taşıtları incelemişlerdir.

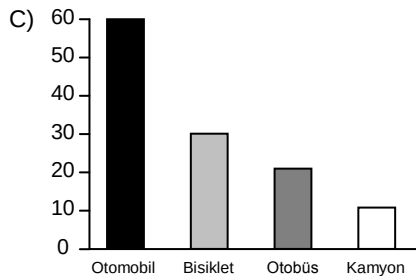
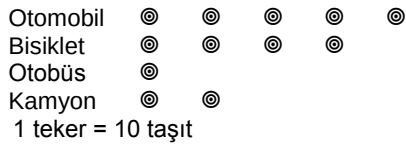
Tablo onların gördükleri taşıtları göstermektedir.

Taşıt Cinsi	Sayı
Otomobil	60
Bisiklet	30
Otobüs	10
Kamyon	20

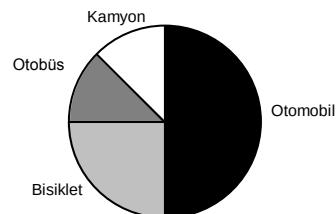
Tablodaki sonuçları göstermek için her öğrenci bir grafik çizmiştir. Aşağıda verilen bu grafiklerin hangisinde sonuçlar doğru olarak gösterilmiştir?



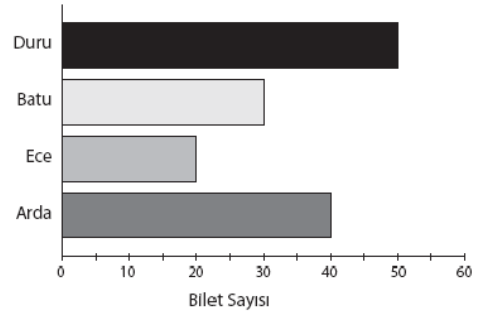
B)



D)



10.

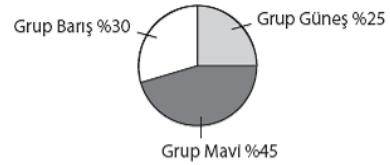


Duru, Batu, Ece ve Arda okul konserleri için bilet sattılar. Bunlardan her birinin satmış olduğu biletlerin sayısı yukarıdaki grafikte görülmektedir. Bu dört kişiden iksinin sattığı biletlerin sayısı Duru'nun tek başına sattığı bilet sayısı kadardır.

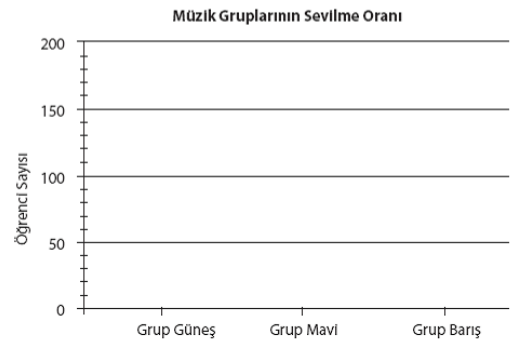
Bu iki kişi hangileridir?

Yanıt: _____ ve _____

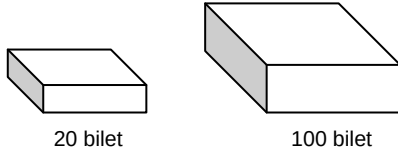
11. Aşağıdaki daire grafiğinde, 200 öğrenci üzerinde yapılan bir araştırmanın sonuçları görülmektedir.



Aşağıda ayrılan yere, daire grafiğinin her bölümündeki öğrenci sayısını gösteren bir sütun grafiği çiziniz.



12. Küçük kutuda 1'den 20'ye kadar numaralanmış bilet; büyük kutuda 1'den 100'e kadar numaralanmış 100 bilet vardır.



Kutuların içindeki biletlere bakmadan her kutudan bir bilet çekiliyor. Hangi kutudan 17 numaralı biletin çekilme olasılığı daha fazladır?

- A) İçinde 20 bilet olan kutu
 B) İçinde 100 bilet olan kutu
 C) İki kutu için de olasılık aynıdır.
 D) Bu soruya cevap verilemez.
13. Bir otobüste 36 yolcu var. otobüsteki çocukların yetişkinlere oranı 5 çocuğa 4 yetişkindir. Bu otobüste kaç çocuk vardır?

Yanıt: _____

14. Okulun düzenlediği gezilere her 12 öğrenci için 1 öğretmen katılmaktadır. Geziye 108 öğrenci gidiyorsa bu geziye kaç öğretmen katılacaktır?
- A) 7 B) 8 C) 9 D) 10

15. Aşağıdaki AO doğru parçasını kullanarak AOB dar, AOC geniş açı olacak şekilde O'dan geçen düz bir BC doğrusu çizin. B ve C noktalarını işaretleyiniz.



YANIT ANAHTARI

SORU NO	DOĞRU YANIT
1	D
2	B
3	D
4	4 2 2 5 3 3 6 4 4
5	8
6	C
7	C
8	B
9	A
10	Ece ve Batu
11	Üçü de doğru – (50,90,60) 50'nin doğru çizgide olması gerekir. 90, 100'den küçük fakat 80'den büyük olmalıdır. 60, 70'ten küçük fakat 50 den büyük olmalıdır.
12	A
13	20
14	C
15	O dan geçen doğru çizilmiş; dar açı ve geniş açı adlandırılmış.

KARMA TEST 4

1. Aşağıdaki seçeneklerin hangisinde verilen sayılar, BÜYÜKTEN KÜÇÜĞE doğru sıralanmıştır?

- A) 10,001; 10,110; 11,001; 11,100
 B) 10,110; 10,011; 11,100; 11,011
 C) 11,001; 11,100; 10,110; 10,011
 D) 11,100; 11,001; 10,110; 10,011

2. Bir sınıftaki öğrenci sayısı 30'dur. Bu sınıftaki erkek öğrencilerin sayısını kız öğrencilerin sayısına oranı $\frac{2}{3}$ 'tür. Sınıftaki erkek öğrenci sayısı kaçtır?

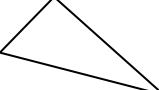
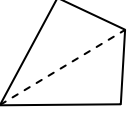
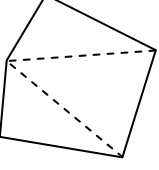
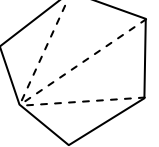
- A) 6 B) 12 C) 18 D) 20

3. $4x - x + 7y - 2y$ ifadesine aşağıdakilerden hangisi eşittir?

- A) 9 B) $9xy$
 C) $4 + 5y$ D) $3x + 5y$

4. İç Açılar

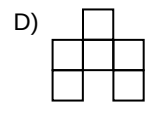
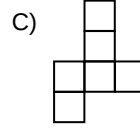
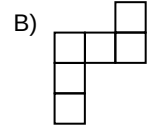
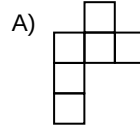
Kağan, çokgenlerin özelliklerini inceliyordu. Çokgenlerde kenarlar ve açılar arasında bir bağıntı olup olmadığını nortaya çıkarmak için aşağıdaki tabloyu hazırladı.

Çokgen	Kenar Sayısı	Üçgen Sayısı	İç Açılar Top.
	3	1	$1 \times 180^\circ$
	—	—	$__ \times 180^\circ$
	—	—	$__ \times 180^\circ$
	—	—	$__ \times 180^\circ$

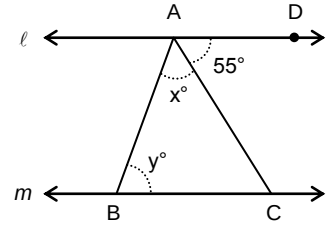
Aşağıdaki eşitlikte verilen karenin içine doğru sayıyı yazınız.

10 kenarlı bir çokgenin iç açıları toplamı = $\square \times 180^\circ$

5. Aşağıda verilen açık şekillerden hangisi katlandığında bir küp oluşturur?



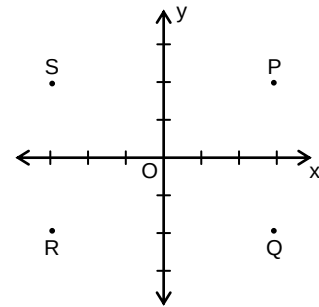
- 6.



Bu şekilde, ℓ doğrusu m doğrusuna paraleldir. DAC açısının ölçüsü 55° dir. $x + y$ kaç derecedir?

- A) 55 B) 110 C) 125 D) 135

- 7.



Grafikteki hangi nokta (3, -2) noktasıdır?

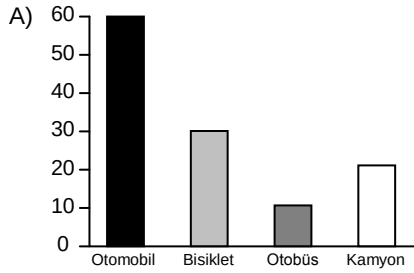
- A) P B) Q C) R D) D

8. 4 öğrenci, 1 saat boyunca okullarının önündeki yoldan geçen taşıtları incelemiştir.

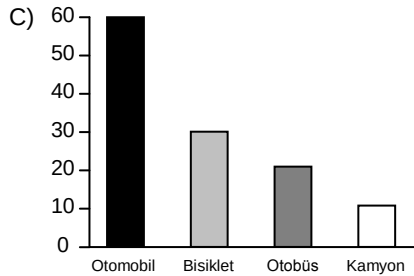
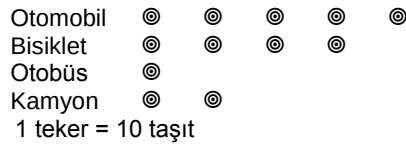
Tablo onların gördükleri taşıtları göstermektedir.

Taşıt Cinsi	Sayı
Otomobil	60
Bisiklet	30
Otobüs	10
Kamyon	20

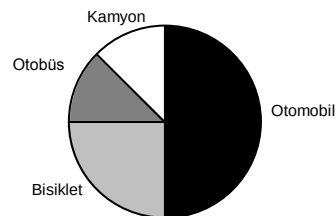
Tablodaki sonuçları göstermek için her öğrenci bir grafik çizmiştir. Aşağıda verilen bu grafiklerin hangisinde sonuçlar doğru olarak gösterilmiştir?



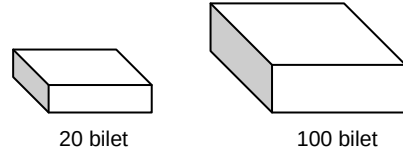
B)



D)



9. Küçük kutuda 1'den 20'ye kadar numaralanmış bilet; büyük kutuda 1'den 100'e kadar numaralanmış 100 bilet vardır.



Kutuların içindeki biletlere bakmadan her kutudan bir bilet çekiliyor. Hangi kutudan 17 numaralı biletin çekilme olasılığı daha fazladır?

- A) İçinde 20 bilet olan kutu
 B) İçinde 100 bilet olan kutu
 C) İki kutu için de olasılık aynıdır.
 D) Bu soruya cevap verilemez.

10. $0,402 \times 0,53 =$

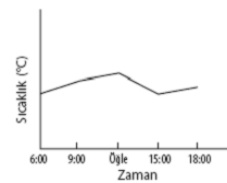
Yanıt: _____

11. Tabloda bir gün içindeki değişik zamanlarda ölçülen hava sıcaklıkları görülmektedir.

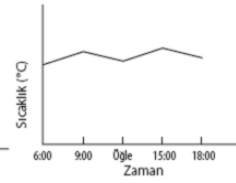
Zaman	6:00	9:00	Öğle	15:00	18:00
Sıcaklık °C	12	17	14	18	15

Bu sonuçlar, sıcaklık ölçeği olmayan bir grafikte gösteriliyor. Bu grafik aşağıdakilerden hangisi olabilir?

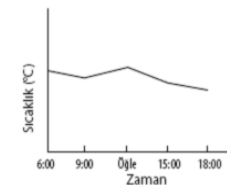
A)



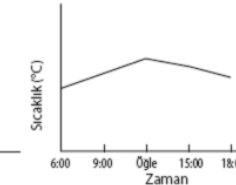
B)



C)



D)



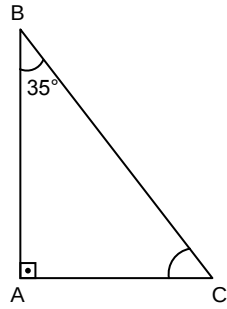
12. Okulun düzenlediği gezilere her 12 öğrenci için 1 öğretmen katılmaktadır. Geziye 108 öğrenci gidiyorsa bu geziye kaç öğretmen katılacaktır?

A) 7 B) 8 C) 9 D) 10

13. Hangi sayı -6 sayısına bölünürse sonuç 12 olur?

A) -72 B) -2 C) 2 D) 72

14.



Yukarıdaki üçgende C açısı kaç derecedir?

A) 45° B) 55° C) 65° D) 145°

15. Aşağıdaki AO doğru parçasını kullanarak AOB dar, AOC geniş açı olacak şekilde O'dan geçen düz bir BC doğrusu çizin. B ve C noktalarını işaretleyiniz.



YANIT ANAHTARI

SORU NO	DOĞRU YANIT
1	D
2	B
3	D
4	8
5	C
6	C
7	B
8	A
9	A
10	0,21306 0,21306 öğrencinin bir parçası olarak gösterilmiş ve daha sonra yanlış veya doğru olarak yuvarlanmış.
11	B
12	C
13	A
14	B
15	O dan geçen doğru çizilmiş; dar açı ve geniş açı adlandırılmış.

