

**FEN BAŞARISINDAKİ ARTIŞIN BELİRLENMESİNDE
MADDE TEPKİ KURAMINA DAYALI DİKEY ÖLÇEKLEME
YÖNTEMLERİNİN KARŞILAŞTIRILMASI**

**A COMPARISON OF IRT VERTICAL SCALING METHODS
IN DETERMINING OF THE INCREASE IN ACHIEVEMENT
OF SCIENCE EDUCATION**

Aylin ALBAYRAK SARI

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

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

Öngördüğü

Doktora Tezi

olarak hazırlanmıştır.

2014

Eđitim Bilimleri Enstitüsü M¼d¼rl¼đ¼'ne,

Aylin ALBAYRAK SARI'nın hazırladıđı "Fen Bařarısındaki Artıřın Belirlenmesinde Madde Tepki Kuramına Dayalı Dikey ¼lçekleme Y¼ntemlerinin Karřılařtırılması" bařlıklı bu çalıřma j¼rimiz tarafından **Eđitim Bilimleri Anabilim Dalı, Eđitimde ¼lçme ve Deđerlendirme Bilim Dalı'nda Doktora Tezi** olarak kabul edilmiřtir.

Bařkan Prof. Dr. Selahattin GELBAL

¼ye (Danıřman) Prof. Dr. H¼lya KELECIOđLU

¼ye Doç. Dr. Duygu ANIL

¼ye Doç. Dr. Neře G¼LER

¼ye Yrd. Doç. Dr. ¼mer KUTLU

ONAY

Bu tez Hacettepe ¼niversitesi Lisans¼st¼ Eđitim-¼đretim ve Sınav Y¼netmeliđi'nin ilgili maddeleri uyarınca yukarıdaki j¼ri ¼yeleri tarafından 22/10/2014 tarihinde uygun g¼r¼lm¼ř ve Enstit¼ Y¼netim Kurulunca/..../..... tarihinde kabul edilmiřtir.

Prof. Dr. Berrin AKMAN
Eđitim Bilimleri Enstit¼s¼ M¼d¼r¼

FEN BAŞARISINDAKİ ARTIŞIN BELİRLENMESİNDE MADDE TEPKİ KURAMINA DAYALI DİKEY ÖLÇEKLEME YÖNTEMLERİNİN KARŞILAŞTIRILMASI

Aylin ALBAYRAK SARI

ÖZ

Bu araştırmada iki farklı veri seti kullanılarak Madde Tepki Kuramına dayalı dikey ölçekleme çalışması yürütülmüş, farklı kalibrasyon yöntemleri (ayrı ve eş zamanlı kalibrasyon) ve farklı yetenek kestirim yöntemleri (maksimum olabilirlik, beklenen önsel dağılım, maksimum önsel dağılım) kullanarak elde edilen dikey ölçekleme sonuçları karşılaştırılmıştır. Elde edilen dikey ölçekler, bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme (grade-to-grade growth), sınıf düzeyleri arasındaki çeşitlilik (grade-to-grade variability) ve düzey dağılımlarının ayrımı (the separation of grade distributions) kriterlerine göre değerlendirilmiştir. Bir sınıf düzeyinden diğer sınıf düzeyine olan büyümeyi değerlendirebilmek için ortalamalar ve ortalama farkları, sınıf düzeyleri arasındaki çeşitliliği değerlendirebilmek için her sınıf düzeyi için standart sapma değerleri ve düzey dağılımlarının ayrımını değerlendirebilmek için de etki büyüklüğü değerleri incelenmiş ve karşılaştırılmıştır. Altıncı, yedinci ve sekizinci sınıf fen bilgisi başarı testleri 2008-2012 yılları arasında uygulanan SBS, OKS ve PYBS testlerinden seçilerek oluşturulmuştur. Veri toplama araçları Ankara ilinden farklı sosyoekonomik kültüre sahip on iki ilköğretim okulundan toplam 1500 öğrenciye uygulanmış ve gerçek veri seti oluşturulmuştur. Yapay veri seti oluşturulurken alanyazın incelenmiş ve alanyazındaki çalışmalara göre belirlenen madde ve yetenek parametreleri kullanılmıştır. İki farklı veri seti kullanılarak dikey ölçekleme yapılmıştır. Elde edilen bulgular değerlendirme kriterlerine göre karşılaştırılmıştır. Uygulanan fen bilgisi başarı testi verileri ile yapılan eş zamanlı kalibrasyon yöntemi ile hesaplanan ortalamalar incelendiğinde, öğrenci başarısının 6. sınıftan 8. sınıfa doğru arttığı görülmektedir. Her üç sınıf düzeyi için ortalama farkları incelendiğinde, en yüksek ortalama farklarının EAP yöntemi ile elde edildiği görülmektedir. Standart sapma değerleri incelendiğinde, 6. sınıftan 8. sınıfa doğru standart sapma değerlerinin azaldığı ve en düşük standart sapma değerinin ML yöntemi ile elde edildiği

görülmektedir. Etki büyüklüğü kriterleri incelendiğinde, etki büyüklüğünün 6. sınıftan 8. sınıfa doğru arttığı ve en yüksek etki büyüklüğü değerlerinin EAP yöntemi ile elde edildiği görülmektedir. Fen bilgisi başarı verileri ile yapılan ayrı kalibrasyon yöntemi ile hesaplanan ortalamalar incelendiğinde, öğrenci başarısının 6. sınıftan 8. sınıfa doğru arttığı, ortalamalar arası farkın ise 8. sınıfa doğru azaldığı görülmektedir. Ayrı kalibrasyonda en yüksek ortalama fark EAP yöntemi ile ölçülmüştür. Ayrı kalibrasyon ile elde edilen ortalama farklarının, eş zamanlı kalibrasyon ile elde edilen ortalama farklarından daha düşük olduğu görülmektedir. Standart sapma değerleri incelendiğinde, genel olarak 6. sınıftan 8. sınıfa doğru standart sapma değerlerinin azaldığı, en düşük standart sapma değerinin ML yöntemi ile elde edildiği ve ayrı kalibrasyon ile elde edilen standart sapma değerlerinin genel olarak eş zamanlı kalibrasyon ile elde edilen değerlere göre daha düşük olduğu görülmektedir. Etki büyüklüğü kriterleri incelendiğinde, etki büyüklüklerinin 8. sınıfa doğru azaldığı, en yüksek etki büyüklüğü değerlerinin ML yöntemi ile elde edildiği ve ayrı kalibrasyon ile elde edilen etki büyüklüğü değerlerinin eş zamanlı kalibrasyon ile elde edilen değerlere göre daha düşük olduğu görülmektedir. Yapay veri seti ile yapılan dikey ölçekleme sonuçlarına göre, eş zamanlı ve ayrı kalibrasyon yönteminin her ikisinde de ortalama farkların 8. sınıf düzeyine göre arttığı ve her iki yöntemde de en yüksek değerlerin EAP yöntemi ile elde edildiği bulunmuştur. Ayrı kalibrasyon ile elde edilen ortalama farkları eş zamanlı kalibrasyon ile elde edilen ortalama farklarına göre daha düşüktür. Standart sapma değerleri karşılaştırıldığında, her iki kalibrasyon yönteminde de en yüksek değerlerin EAP ile elde edildiği ve her iki yöntemle de genel olarak birbirine yakın standart sapma değerleri elde edildiği görülmektedir. Etki büyüklüğü değerleri incelendiğinde her iki kalibrasyon yönteminde de 8. sınıfa doğru etki büyüklüğü değerlerinin arttığı görülmektedir. Ayrı kalibrasyon ile elde edilen etki büyüklüğü değerleri eş zamanlı kalibrasyona göre daha düşüktür. Her üç yetenek kestirim yöntemi birbirine yakın sonuçlar üretmiştir.

Anahtar sözcükler: Madde Tepki Kuramı, Dikey Ölçekleme, Kalibrasyon Yöntemleri, Yetenek Kestirim Yöntemleri, Simülasyon, Yapay Veri

Danışman: Prof. Dr. Hülya KELECİOĞLU, Hacettepe Üniversitesi, Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı

A COMPARISON OF IRT VERTICAL SCALING METHODS IN DETERMINING OF THE INCREASE IN ACHIEVEMENT OF SCIENCE EDUCATION

Aylin ALBAYRAK SARI

ABSTRACT

In this research, a vertical scaling study based on Item Response Theory was conducted by using two different datasets and vertical scaling results obtained by different calibration methods and different proficiency estimation methods (maximum likelihood, expected prior distribution, maximum prior distribution) are compared. The resulting vertical scales were evaluated according to the criteria of the growth from one grade level to another (grade-to-grade growth), variability between grade levels (grade-to-grade variability) and the separation of grade distributions. Averages and average differences, standard deviation values and effect size values were examined and compared in order to evaluate and compared the growth from one grade level to another, the variability between grade levels and the separation of grade distributions, respectively. Sixth, seventh and eighth grade science education achievement tests were created from questions of SBS, OKS and PYBS exams had applied between 2008-2012. Data collection tools were applied to 1500 students of twelve different primary schools in Ankara consist of different socio-economic culture and real data set was created. The literature was examined while artificial data set was formed and items and proficiency parameters those had determined according to the studies in literature were used. The vertical scaling was performed using two different datasets. The findings were compared according to the evaluation criteria. Averages calculated by application of concurrent calibration method with applied science education achievement test data were analyzed and it was observed that student achievement from sixth to eighth grade were increasing. Average differences for each grade levels were analyzed and it was observed that the highest average differences were obtained by Expected Apriori (EAP) method. Standard deviation values were analyzed and it was observed that standard deviation values were decreasing from sixth to eighth grade and the lowest standard deviation were obtained by Maximum Likelihood (ML) method. Effect size criteria were analyzed and it was observed that effect size were increasing from sixth to eighth grade and the highest effect size values were obtained by EAP

method. Averages calculated by application of separate calibration method with applied science education achievement test data were analyzed and it was observed that student achievement from sixth to eighth grade were increasing and differences between averages were decreasing from sixth to eighth grade. The highest difference between averages were measured in the separate calibration method by EAP method. The differences between averages obtained by separate calibration method were lower than the ones obtained by concurrent calibration method. Standard deviation values were analyzed, it was observed that the values were usually decreasing from sixth to eighth grade, the lowest standard deviation value was obtained by ML method and standard deviation values obtained by separate calibration method were lower than the ones obtained by concurrent calibration method. Effect size criteria were analyzed and it was observed that effect size values were decreasing from sixth to eighth grade, the highest effect size value was obtained by ML method and effect size values obtained by separate calibration method were lower than the ones obtained by concurrent calibration method. According to the vertical scaling results applied with artificial dataset, average differences were increasing from sixth to eighth grade in both concurrent and separate calibration methods and the highest values were obtained by EAP method in both calibration methods. Average differences obtained by separate calibration method were lower than the values obtained by concurrent calibration method. Standard deviation values were compared and it was observed that the highest values were obtained by EAP method in both calibration methods and generally, standard deviation values obtained by two calibration methods were close to each other. Effect size criteria were analyzed and it was observed that effect size values were increasing from sixth to eighth grade in both calibration methods. Effect size values obtained by separate calibration method were lower than the values obtained by concurrent calibration method. The results produced by all three proficiency estimation methods were close to each other.

Keywords: Item Response Theory, Vertical Scaling, Calibration Methods, Proficiency Estimation Methods, Simulation, Simulated Data

Advisor: Prof. Dr. Hülya KELECİOĞLU, Hacettepe University, Department of Educational Sciences, Division of Educational Measurement and Evaluation

ETİK BEYANNAMESİ

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

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

beyan ederim.

Aylin ALBAYRAK SARI

TEŞEKKÜR

Yüksek lisans ve doktora eğitimim boyunca yakın ilgi ve desteğini esirgemeyen, beni yüreklendiren, tezimin yapılanmasında büyük emeği olan hocam Prof. Dr. Hülya Kelecioğlu'na;

Yüksek lisansa başladığım zamandan bu yana kendimi geliştirmemde büyük emeği ve katkısı olan, tecrübelerini ve akademik birikimini cömertçe paylaşan, sağlam karakteriyle bizlere her zaman örnek olan hocam Prof. Dr. Selahattin Gelbal'a;

Tezimi okuyarak değerli görüşleriyle katkı sağlayan, tanıdığım için kendimi çok şanslı hissettiğim hocalarım Doç. Dr. Nuri Doğan, Doç. Dr. Duygu Anıl, Doç. Dr. Neşe Güler ve Yrd. Doç. Dr. Ömer Kutlu hocalarıma;

Bugünlere gelmemdeki sonsuz emekleri, destekleri için varını yoğuna katarak beni okutan kariyerimin mimarı Canım AİLEM'e;

Her zaman yanımda olan, sabrı ve anlayışıyla tez yazma sürecimde beni cesaretlendiren, destekleyen, beni her yorulduğumda yüreklendiren ve motive eden EŞİM'e;

Sonsuz teşekkür ederim.

**Bu tez Hacettepe Üniversitesi Bilimsel Araştırmal Birimi tarafından desteklenmiştir (Proje no: 1591).*

***Doktora eğitimim süresince beni, 2211 Yurtiçi Doktora Bursu ile destekleyen TÜBİTAK'a teşekkür ederim.*

Canım aileme ve Biricik eşime;

İÇİNDEKİLER

ÖZ	i
ABSTRACT	iii
ETİK BEYANNAMESİ	v
TEŞEKKÜR	vi
İÇİNDEKİLER.....	viii
TABLolar DİZİNİ	ix
SİMGELER VE KISALTMALAR DİZİNİ	xi
1. GİRİŞ.....	1
1.1. Problem Durumu.....	1
1.2. Araştırmanın Amacı ve Önemi	3
1.3. Problem Cümlesi.....	4
1.3.1. Alt Problemler.....	4
1.4. Sınırlılıklar	5
1.5. Araştırmanın Kuramsal Temeli.....	5
1.5.1 Test Eşitleme	5
1.5.2 Eşitleme Desenleri	9
1.5.3 Eşitleme Yöntemleri	13
1.5.4 Değerlendirme Ölçütleri.....	27
2. İLGİLİ ARAŞTIRMALAR.....	29
3. YÖNTEM	38
3.1. Araştırmanın Türü	38
3.2. Çalışma Grubu.....	38
3.3. Araştırma Deseni	38
3.4. Verilerin Analizi	39
3.4.1. Gerçek Veri Seti	39
3.4.2 Simülasyon ile Yapay Verilerin Türetilmesi.....	43
4. BULGULAR VE TARTIŞMA	46
4.1 Birinci Alt Probleme Ait Bulgular ve Yorumlar.....	46
4.2 İkinci Alt Probleme Ait Bulgular ve Yorumlar	52
4.3 Üçüncü Alt Probleme Ait Bulgular ve Yorumlar	59
4.4 Dördüncü Alt Probleme Ait Bulgular ve Yorumlar	64
5. SONUÇ ve ÖNERİLER	71
5.1 Sonuçlar.....	71
5.2 Öneriler	73
5.2.1 Uygulamaya Dönük Öneriler	73
5.2.2 Araştırmacılara Dönük Öneriler	74
KAYNAKÇA.....	76
EKLER DİZİNİ	83
EK-1.....	85
EK-2.....	97
EK-3.....	109
EK-5.....	120
ÖZGEÇMİŞ	121

TABLolar DİZİNİ

Tablo 1.1 Tek Grup ve Dengelenmiş Tek Grup Deseni.....	10
Tablo 1.2 Random Grup Deseni.....	11
Tablo 1.3 Denk Olmayan Gruplarda Ortak Madde Deseni	11
Tablo 3.1 Ortak Maddelerin Test Düzeylerine Göre Dağılımı	39
Tablo 3.2 Fen Bilgisi Başarı Testi Doğrulayıcı Faktör Analizi Sonucu Elde Edilen Uyum İyiliği İndeksleri.....	41
Tablo 3.3 Ayrı Kalibrasyona Ait Sınıflara Göre Puanların İki Parametrelili Lojistik Model Ki-Kare Uyum İstatistikleri.....	42
Tablo 3.4 Madde Tepki Kuramına Göre İki Parametrelili Lojistik Model İle Ölçekleme Sonucu Elde Edilen Madde Parametrelerinin Betimsel İstatistikleri....	43
Tablo 4.1 Fen Bilgisi Başarı Testi Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen ML, EAP ve MAP Yetenek Kestirimine Ait Sonuçlar.....	47
Tablo 4.2 Stocking Lord Dönüşümü İçin Elde Edilen A ve B Sabitleri	52
Tablo 4.3 Fen Bilgisi Başarı Testi Ayrı Kalibrasyon Yöntemi İle Elde Edilen ML, EAP ve MAP Yetenek Kestirimine Ait Sonuçlar.....	53
Tablo 4.4 Yapay Veri Seti Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen ML, EAP ve MAP Yetenek Kestirimine Ait Sonuçlar.....	60
Tablo 4.5 Stocking-Lord Dönüşümü İçin Elde Edilen A ve B Sabitleri	65
Tablo 4.6 Ayrı Kalibrasyon Yöntemi İle Elde Edilen ML, EAP ve MAP Yetenek Kestirimine Ait Sonuçlar	65

GRAFİKLER DİZİNİ

Grafik 4.1 Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen Ortalama ve Ortalama Farkları.....	49
Grafik 4.2 Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen Standart Sapma Değerleri.....	50
Grafik 4.3 Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen Etki Büyüklüğü Değerleri.....	51
Grafik 4.4 Ayrı Kalibrasyon Yöntemi İle Elde Edilen Ortalama ve Ortalama Farkları.....	56
Grafik 4.5 Ayrı Kalibrasyon Yöntemi İle Elde Edilen Standart Sapma Değerleri....	57
Grafik 4.6 Ayrı Kalibrasyon Yöntemi İle Elde Edilen Etki Büyüklüğü Değerleri.....	59
Grafik 4.7 Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen Ortalama ve Ortalama Farkları.....	62
Grafik 4.8 Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen Standart Sapma Değerleri.....	63
Grafik 4.9 Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen Etki Büyüklüğü Değerleri.....	64
Grafik 4.10 Ayrı Kalibrasyon Yöntemi İle Elde Edilen Ortalama ve Ortalama Farkları....	67
Grafik 4.11 Ayrı Kalibrasyon Yöntemi İle Elde Edilen Standart Sapma Değerleri.....	69
Grafik 4.12 Ayrı Kalibrasyon Yöntemi İle Elde Edilen Etki Büyüklüğü Değerleri.....	70

SİMGELER VE KISALTMALAR DİZİNİ

MTK: Madde Tepki Kuramı

IRT: Item Response Theory

EAP: Beklenen Önsel Bilgi (Epected a Posteriori)

MAP: Maksimum Önsel Bilgi (Maximum a Posteriori)

ML: Maksimum Olabilirlik (Maximum Likelihood Estimation)

MEB: Milli Eğitim Bakanlığı

SBS: Seviye Belirleme Sınavı

OKS: Ortaöğretim Kurumları Seçme ve Yerleştirme Sınavı

PYBS: Parasız Yatılılık ve Bursluluk Sınavı

TEOG: Temel Eğitimden Ortaöğretime Geçiş

1. GİRİŞ

1.1. Problem Durumu

Bilgi çağı olarak nitelendirilen günümüzde sadece bilgiye kolay ulaşabilen bireyler değil, aynı zamanda bilgi üretebilen ve bilgiyi kullanabilen bireylere ihtiyaç duyulmaktadır. İstenilen özelliklere sahip nitelikli gelecek kuşaklar yetiştirmenin en temel yolunun iyi bir eğitimden geçtiği bilinmektedir. Çağın ihtiyacı olan bu bireyleri yetiştirebilmek için iyi bir fen eğitimine ihtiyaç duyulmakta, bu nedenle tüm ülkelerde olduğu gibi ülkemizde de fen eğitimine önem verilmektedir.

Fen eğitimi geleceğin yetişkinleri olan miniklerin dünyayı anlamasına yardımcı olur. Fen eğitimi ile öğrenciler bilimsel süreç becerilerini, eleştirel düşünebilmeyi, karşılaştığı problemleri çözebilmeyi, gelişen teknolojiyi takip edebilmeyi ve çağa ayak uydurabilmeyi öğrenir. Fen Bilgisi dersi öğretim programı, öğrencilerin bilim ve bilginin doğasını, temel fen kavramlarını anlamalarını, günlük yaşamlarında bunların önemini fark etmelerini ve ilişkilendirmelerini, okulda ve okul dışında istekli bir şekilde çalışmalarını sürdürmelerini hedeflemektedir (MEB 2005).

Öğrenciler, fen öğretimindeki öğrenme yaşantıları ile bağımsız şekilde kendilerini yöneten insanlar olarak yetişirler. Öğrencileri problemleri tanıma, çözüm için plan yapma, veri toplama, karar verme, bu karara göre eyleme geçme ve en sonunda ürünleri değerlendirme becerileri ile donatmak fen eğitimiyle olanaklıdır.

Fen eğitimi ile öğrenciler, problem çözme, eleştirel düşünme, bilgi edinme ve elde edilen bilgiyi güncel sorunlara uygulayabilme becerisi kazanmaktadır (Fidan ve Baykul, 1993). Fen eğitimi, bilimin günlük hayattaki önemini kavranmasına yardımcı olmakta, bireylerin bilimsel bilgiyi, bilimsel süreç becerilerini ve bilime karşı olumlu tutum kazanmalarını sağlamaktadır. Bu nedenle, fen eğitimi bireyin gelecekteki yaşamını yönlendirme açısından oldukça önemli bir yere sahiptir. Araştıran, tartışan, bilimsel süreç becerilerini kullanabilen, bilime karşı olumlu tutumlar sergileyen fen okuyuları bireylerin yetiştirilmesinde fen eğitimi önem kazanır (Çepni, Ayvaci ve Bacanak, 2004).

Fen eğitimi tüm dünyada üzerinde sıklıkla durulan bir konu olduğu için, eğitime yapılan yatırım gittikçe artmaktadır. Bu yatırımların sonuçları yani yapılan eğitim-öğretim faaliyetlerinin etkililiği merak edilmektedir. Eğitim her anlamda bireyin davranışlarını etkilese de, eğitimin etkililiğinin değerlendirilmesinde, öğrenci başarısının belirlenmesi önemli bir boyuttur. Ülkemizde öğrenci başarısının belirlenmesi amacıyla yapılan sınavlar genel olarak bir üst eğitim basamağına öğrenci seçmek ve yerleştirmek için yapılmaktadır. Öğrenci başarısının izlenebilmesi için yapılan çalışmalar azınlıktadır.

Sınavlar çok çeşitli nedenlerle yapılmakta ve sınav sonuçlarına göre bazı kararlar verilmektedir. Bir öğrencinin hangi okula devam edeceğine karar verilirken, bir üniversite alacağı öğrencinin sahip olması gereken test puanını belirlerken, eğitim sistemini geliştirmek için neler yapılması gerektiği hakkında kararlar alınırken ve eğitim uygulamalarındaki değişiklikler değerlendirilirken uygulanan sınavlardan elde edilen bu bilgiler kullanılır (Kolen & Brennan, 2004). Test uygulayıcıları da farklı formlardan ve farklı gruplardan elde edilen test puanlarını karşılaştırabilmeyi ve ihtiyaç duyulduğunda testleri birbirinin yerine kullanabilmeyi isterler. Fakat karşılaştırmanın yapılabilmesi için ham puanların standart puanlara dönüştürülmesi gerekmektedir (AERA, APA & NCME, 1999). Benzer özellikleri ölçmesi amacıyla hazırlanan farklı testlerden elde edilen puanların ortak bir ölçeğe dönüştürülmesi ile karşılaştırma yapılabilmekte ve bu tür puanlara karşılaştırılabilir puanlar denilmektedir (Angoff, 1971).

Karşılaştırılabilir puanların en çok kullanıldığı alanlardan biri de gelişimsel ölçek puanlarıdır (developmental score scale). Akademik gelişimin yıldan yıla ne düzeyde olduğunun belirlenmesi için, farklı sınıf düzeylerindeki öğrencilerden elde edilen puanların ortak bir ölçeğe dönüştürülmesiyle elde edilen gelişimsel ölçek puanları kullanılır (Kolen & Brennan, 2004). Test bataryalarında ölçekleme işlemlerinin yapılmasının temel nedeni, test geliştiricilere öğrenci başarısındaki ilerlemeyi izleyebilecekleri gelişimsel ölçek puanı sağlamaktır (Loyd ve Hoover, 1980). Akademik gelişimin yıldan yıla ne düzeyde olduğunun belirlenmesinde en temel sorun; uygulanan grubun, testlerin güçlüklerinin ve test içeriklerinin farklı olmasıdır. Bu sorunu aşabilmek için, ardışık sınıf seviyelerindeki öğrencilere ortak sorular sorularak iki farklı yetenek düzeyindeki öğrencilerin puanları ortak bir

ölçeğe dönüştürülür. Bu durumda dikey ölçekleme olarak adlandırılan bağlama süreçlerine başvurulur (Kolen & Brennan, 2004).

1.2. Araştırmanın Amacı ve Önemi

Bu çalışmanın amacı, Madde Tepki Kuramı (MTK)'ya dayalı dikey ölçekleme çalışması yürüterek, farklı kalibrasyon yöntemleri (ayrı ve eş zamanlı kalibrasyon) ve farklı yeterlik tahmini yöntemleri (maksimum olabilirlik, beklenen önsel dağılım, maksimum önsel dağılım) kullanarak oluşturulan farklı dikey ölçekleme sonuçlarının ölçeklerin ortalama, standart sapma ve etki büyüklüğü değerlerine göre karşılaştırılmasıdır. Bu çalışmada çeşitli dikey ölçeklerin, bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme (grade-to-grade growth), sınıf düzeyleri arasındaki çeşitlilik (grade-to-grade variability) ve düzey dağılımlarının ayrımı (the separation of grade distributions) özellikleri üzerinde durulmuştur. Bir sınıf düzeyinden diğer sınıf düzeyine olan büyümeyi değerlendirebilmek için ortalamalar ve ortalama farkları, sınıf düzeyleri arasındaki çeşitliliği değerlendirebilmek için her sınıf düzeyi için standart sapma değerleri ve düzey dağılımlarının ayrımını değerlendirebilmek için de etki büyüklüğü değerleri incelenmiştir.

Alanyazında hangi ölçekleme yöntemlerinin öğrenci başarısındaki artışı en iyi ve doğru ortaya koyduğu konusunda ortak bir görüş bulunmamaktadır. Buna rağmen dikey ölçekleme birçok test geliştirici tarafından kullanılmaktadır. Ancak her test geliştirici geliştirdiği ölçek için dikey ölçek geliştirme süreçlerini kendisi belirlemektedir (Tong & Kolen, 2007).

Öğrencilerin sınıf düzeyleri arttıkça başarılarındaki gelişimlerinin ortaya konulmasına olanak sağlayan dikey ölçekleme konusu giderek önem kazanmış ve bu konuda daha fazla çalışma yapılmaya başlanmıştır. Bu çalışma öğrenci başarısındaki gelişimin izlenmesi konusunda örnek oluşturabilecektir.

Alanyazın incelendiğinde ülkemizde yapılan çalışmaların çoğunun gerçek verilerle yapıldığı, simülasyon çalışmalarının oldukça az olduğu görülmektedir. Bu çalışmada ise hem gerçek veri hem de simülasyon ile türetilen yapay veriler kullanılarak sonuçların farklılaşıp farklılaşmadığı kontrol edilmiştir. Böylece bu çalışmanın alanyazına katkı sağlayacağı düşünülmektedir.

1.3. Problem Cümlesi

Madde Tepki Kuramı ile dikey ölçekte farklı kalibrasyon yöntemleri ve farklı yetenek kestirim yöntemleri kullanıldığında, elde edilen değerlendirme ölçütleri fen bilgisi başarı testinde ve simülasyon ile türetilen yapay veri setinde nasıl değişmektedir?

1.3.1. Alt Problemler

1. Fen Bilgisi Başarı Testinde eş zamanlı kalibrasyon yöntemi ile elde edilen maksimum olabilirlik, beklenen önsel bilgi ve maksimum önsel bilgi yetenek kestirimlerindeki;
 - a. bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme,
 - b. sınıf düzeyleri arasındaki çeşitlilik,
 - c. düzey dağılımlarının ayrımınasıl değişmektedir?
2. Fen Bilgisi Başarı Testinde ayrı kalibrasyon yöntemi ile elde edilen maksimum olabilirlik, beklenen önsel bilgi ve maksimum önsel bilgi yetenek kestirimlerindeki;
 - a. bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme,
 - b. sınıf düzeyleri arasındaki çeşitlilik,
 - c. düzey dağılımlarının ayrımınasıl değişmektedir?
3. Yapay veri setinde eş zamanlı kalibrasyon yöntemi ile elde edilen maksimum olabilirlik, beklenen önsel bilgi ve maksimum önsel bilgi yetenek kestirimlerindeki;
 - a. bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme,
 - b. sınıf düzeyleri arasındaki çeşitlilik,
 - c. düzey dağılımlarının ayrımınasıl değişmektedir?

4. Yapay veri setinde ayrı kalibrasyon yöntemi ile elde edilen maksimum olabilirlik, beklenen önsel bilgi ve maksimum önsel bilgi yetenek kestirimlerindeki;
 - a. bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme,
 - b. sınıf düzeyleri arasındaki çeşitlilik,
 - c. düzey dağılımlarının ayrımı
nasıl değişmektedir?

1.4. Sınırlılıklar

1. Araştırma 1-0 şeklinde puanlanan iki kategorili verilerle yürütülmüştür.
2. Araştırmada madde tepki kuramı modellerinden iki parametrelili model ele alınmıştır.
3. Test uzunluğu 10'u ortak madde olmak üzere toplam 40 madde ile sınırlıdır.
4. Örneklem büyüklüğü yapay veride de gerçek veride ulaşılabilen 500 kişi ile sınırlıdır.

1.5. Araştırmanın Kuramsal Temeli

1.5.1 Test Eşitleme

Test eşitleme aynı özelliği ölçmek amacıyla hazırlanmış testler arasındaki güçlük farklarını gidermek amacıyla kullanılan yöntemlere verilen addır (Kolen, 2004). Test eşitleme çalışması yürütebilmek için bazı sayıtların karşılanmış olması gerekir. Karşılanması gereken sayıtlar şunlardır (Hambleton, R. K. & Swaminathan, 1985; Kolen & Brennan, 2004):

- 1) **Eşitlik:** Eşitlik denk olan X ve Y testleri için, aynı θ yetenek düzeyindeki cevaplayıcıların hangi test formunu aldığı sonucunu etkilememesidir. Eşitlik varsayımına göre; eşit güvenilirliğe sahip olmayan testler, paralel olmayan testler, farklı yetenek ya da becerileri ölçen testler ve farklı güçlük düzeyine sahip testlere ilişkin ham puanlar eşitlenemez.
- 2) **Simetri:** Bir testin X formundan Y formuna yapılan eşitlemenin, benzer şekilde Y formundan X formuna da yapılabilmesi ve eşitlenmiş puanların aynı olması gerekmektedir. Buna eşitlemenin simetrikliği denir. Bu özellik dönüşümün tersine çevrilebilir olması olarak da ifade edilebilir.

- 3) **Gruplar Arası Değişmezlik:** Eşitleme sürecinin cevaplayıcılardan bağımsız olmasıdır. Eşitleme sonucu elde edilen denklem sadece o cevaplayıcı grubuna bağlı olmamalı, benzer tüm koşullara uygulanabilir olmalıdır.
- 4) **Aynı Yapıyı Ölçme:** Eşitleme yapılacak test formlarının her ikisi de aynı özelliği (örtük özellik, yetenek, beceri vb.) ölçmelidir. Farklı kapsamları olan ve farklı özelliği ölçen testler eşitlenemezler. Yatay eşitlemede farklı güçlükte olan testlere ait ham puanlar eşitlenemezken, dikey ölçeklemede testlerin güçlükleri farklıdır. Fakat her iki eşitleme yönteminde de eşitlenecek/ölçeklenecek test formlarının aynı yapıyı ölçmesi beklenir. Yatay ölçeklemenin amacı formların içerikleri arasındaki farkları gidermek değil, güçlükleri arasındaki farkları gidermektir. Örneğin fen ve teknoloji dersinde ışık ve ses konusuna ait bir başarı testinden alınan puanlar ile kuvvet ve hareket konusuna ait bir başarı testinden alınan puanlar birbirine eşitlenemez.
- 5) **Tek Boyutluluk:** Eşitleme yapılacak test formlarından her birinin tek boyutluluk varsayımını karşılaması gerekmektedir. Tek boyutluluk testte yer alan maddelerin aynı yapıyı ölçmesi anlamına gelmektedir.
- 6) **Yerel Bağımsızlık:** Eşitlemenin gruptan bağımsız olması gerekmektedir. Eşitleme olası tüm gruplarda aynı sonucu vermelidir. Nandakumar (1994) yerel bağımsızlık varsayımını karşılamak üzere maddelere verilen cevaplar için sadece bir yeteneğe gereksinim duyulursa, bu maddelerin tek boyutlu olarak ifade edildiğini ileri sürer. Bir başka deyişle, test tek boyutluluk varsayımını karşıladığında yerel bağımsızlık varsayımını da sağlamış olur.

Uygulamada, yukarıda açıklanan durumların birlikte karşılanması pek mümkün olmamaktadır. Dorans (1990) bir testi diğerine eşitlemek için sadece aynı yapıyı ölçme ve tek boyutluluk varsayımını sağlamasının yeterli olduğunu belirtmiştir.

Test eşitlemeye genel olarak; eşitlenecek formlar, benzer güçlük düzeyinde ve uygulandığı grupların yetenek dağılımları birbirine benzer olduğu zaman ya da eşitlenecek formlar, farklı güçlük düzeyinde ve uygulandığı grupların yetenek dağılımları birbirinden farklı olduğu zaman ihtiyaç duyulmaktadır (Hambleton & Swaminathan, 1985). Bu nedenle uygulanacak yöntem, eşitlenecek test formlarının güçlüklerine ve uygulandığı grubun özelliklerine göre, yatay (horizontal)

eşitleme ve dikey (vertical) ölçekleme olmak üzere ikiye ayrılmaktadır (Cook & Eignor, 1983).

1.5.1.1 Yatay Eşitleme

Yatay eşitlemeye, bir testin güvenlik nedeniyle farklı formlarının kullanıldığı durumlarda başvurulmaktadır (Hambleton ve Swaminathan, 1985). Bu eşitleme türünde amaç farklı test formlarından doğrudan karşılaştırılabilen test puanları elde etmektir (Skaggs & Lissitz, 1982, 1986). Yatay eşitlemeyi uygulayabilmek için test formlarının güçlük, kapsam ve psikometrik özellikler bakımından birbirine paralel olması gerekir. Eğer formların içeriği, güçlük düzeyi ve güvenilirliği birbirinden çok farklı ise yatay eşitleme uygun olmayacaktır (Kolen, 1988). Ayrıca, yatay eşitlemede formların uygulandığı gruptaki bireylerin yetenek dağılımının yaklaşık olarak eşit olması beklenmektedir.

Aynı sınıf düzeyinde farklı yıllarda uygulanan fen ve teknoloji testlerinden elde edilen puanların eşitlenmesi, güz ve bahar döneminde uygulanan benzer güçlükte ve benzer kapsamda olan Akademik Personel ve Lisansüstü Eğitime Giriş Sınavı (ALES)'den elde edilen puanların eşitlenmesi vb. durumlarda yatay eşitleme kullanılabilir.

1.5.1.2 Dikey Ölçekleme

Ülkemizde de uygulanan ve öğrencilerin başarısını hem ulusal hem uluslar arası platformda karşılaştırabilmeye olanak sağlayan PISA, PIRLS, TIMSS gibi uluslararası sınavların sayısı arttıkça öğrencilerin akademik başarısının belirlenebilmesi için yapılan çalışmaların sayısında ve bu başarının yıldan yıla nasıl değiştiğinin belirlenebilmesi için yapılan çalışmaların sayısında son yıllarda artış görülmektedir. Amerika Birleşik Devletleri'nde 2001 yılında alınan No Child Left Behind -Hiçbir Çocuk Geride Kalmasın- (NCLB; Public Law 107-110) yasaları gereği ülkedeki tüm çocukların akademik başarı gelişimleri izlenmeye bu konuda çalışmalar yapılmaya başlanmıştır. NCLB uygulamaları ile çoğu ülke öğrencilerinin başarı gelişimlerini izleyebilmek için yapılan çalışmalara ağırlık vermektedir. Her ne kadar alanyazında bir sınıftan daha üst sınıfa başarı gelişimlerini ölçmek için üzerinde hem fikir olunan belli bir yöntem olmasa da dikey ölçekleme çalışmaları ile öğrencilerin gelişimleri belirlenmektedir.

Dikey ölçeklemede, aynı bilgi veya becerileri ölçen fakat güçlükleri farklı olan iki testten elde edilen puanlar eşitlenir. Testi alan bireylerin yetenek düzeyleri ve testlerin güçlük düzeyleri farklıdır (Felan, 2002). Dikey ölçekleme, daha çok ilköğretimde uygulanan başarı testlerinin eşitlenmesinde kullanılmaktadır. Okulöncesi eğitimden on ikinci sınıfa kadar olan büyük çaplı değerlendirmelerde, öğrencilerin akademik gelişimlerini belirlemek için birçok çalışma yapılmaktadır. Belirlenen yıllar arasında karşılaştırma yapabilmek veya sınıf düzeyine bakılmaksızın tüm test puanlarını aynı ölçekte gösterebilmek için verilen sınıf düzeyleri arasında, geniş bir aralıktaki tüm öğrenci performansları için tek bir ölçek puanı elde etmek gerekmektedir. Böyle ölçüklere dikey ölçek, böyle bir ölçek geliştirme sürecine ve bütün sınıf düzeylerindeki değerlendirme puanlarını böyle bir ölçeğe yerleştirmeye dikey ölçekleme denmektedir (McBridge & Wise, 2001). Öğrencilere, sınıf düzeylerine uygun olarak hazırlanan testler uygulanarak farklı sınıf düzeylerinden elde edilen puanlar aynı puan ölçeğinde ifade edilmektedir. Bu işlemler sayesinde, farklı sınıf düzeylerinden elde edilen puanlar karşılaştırılabilir ve bireylerin gelişimleri hakkında bilgi edinilmektedir (Kolen, 1988).

Dikey ölçekleme süreci, farklı düzeylerden elde edilen test puanlarını ortak bir ölçekte ilişkilendirmek için kullanılmaktadır. Aynı sınıf düzeyinde yapılan ölçeklemelerden farklıdır. Ölçekleme sürecinde çeşitli kararların verilmesi gerekmektedir. Bu kararların dikey ölçeklemeyi dolayısıyla da öğrenci başarısındaki gelişmeyi gösteren örüntüleri etkilediği görülmüştür (Tong & Kolen, 2007).

Dikey ölçekleme, karmaşık bir süreç gerektirir. Dikey ölçeklemenin amacı, iki farklı cevaplayıcı grubunu ve iki farklı yetenek düzeyini karşılaştırmaktır (Skaggs & Lissitz, 1986). Dikey ölçeklemede, hem grupların başarı düzeyi hem de grupların aldıkları test içerikleri birbirinden farklı olduğundan, kontrol edilecek değişken sayısı çok fazladır. Dikey ölçeklemede, testi alan grupların düzeyi farklı olduğundan, test üst sınıfa çok kolay gelebilir; bu da tavan etkisi (ceiling effect) probleminin doğmasına ya da tersine alt gruba çok zor gelen bir testte taban etkisi (floor effect) probleminin neden olabilir.

Dikey ölçekleme çalışmaları için öncelikle veri toplama deseninin belirlenmesi, uygun istatistiksel yöntemin ve sonuçları değerlendirecek ölçütlerin seçilmesi gerekir.

1.5.2 Eşitleme Desenleri

Belirlenen amaca göre eşitleme türüne karar verildikten sonra yatay eşitleme ya da dikey ölçekleme yapabilmek için formlar oluşturulur ve veri toplama desenine karar verilir. Eşitlemede veri toplama deseni “eşitleme deseni” olarak ifade edilir (von Davier & Wilson, 2008). Uygun bir eşitleme deseninin kullanılması eşitlemenin önemli basamaklarından biridir. Alanyazında geçen eşitleme desenleri şunlardır:

- Tek grup deseni (Single group design)
- Random grup deseni (Random group design)
- Denk olmayan gruplarda ortak madde deseni (Non-equivalent groups anchor test design - NEAT).
- Ölçekli test deseni (Scaling design)
- Eşit grup deseni (Equal to group design)

Bu araştırmada dikey ölçekleme sürecine uygun bir desen olan “denk olmayan gruplarda ortak madde/test deseni” kullanılmıştır.

1.5.2.1 Tek Grup Deseni

Uygulanabilirlik anlamında basit bir bağlama deseni olan bu desende bağlanacak olan test formlarının ikisi de (X formu ve Y formu) rastgele olarak seçilen bir gruba sırasıyla uygulanır. Bu iki uygulama sonucunda elde edilen puanlar üzerinden bağlama işlemi yapılır. Tek grup desenin örtük sayılıtsı, seçilen gruptan elde edilen eşitleme denkleminin, hedef gruba genellenebilir olmasıdır. Bu yöntem pratikte pek sık kullanılmaz. Tek grup deseninin avantajı, tek grup kullanıldığı için farklı grupların yetenek farklılıklarından doğan hatayı ortadan kaldırmasıdır. Bunun yanı sıra, ölçeklemede tek grup deseni bazı dezavantajlara sahiptir. Bunlardan biri sıra etkisidir, testlerin hangi sırayla verildikleri bir hata kaynağı olabilir. Önce ve sonra verilen testteki öğrenci performansları değişebilir. Tek grup desenin diğer önemli bir dezavantajı da sıra etkisiyle paralel olarak gelişen yorgunluk etkisidir. Bireyler ilk testi aldıktan sonra ikinci teste yorulup daha düşük performans gösterebilirler.

Sıralamadan kaynaklanan hatanın etkisini ortadan kaldırabilmek için grup random olarak ikiye ayrılır. Grubun ilk yarısına önce X sonra Y formu, ikinci yarısına ise önce Y sonra X formu verilir. Bu desene alanyazında **dengelenmiş tek grup deseni** denilmektedir (Kolen & Brennan, 2004). Yorgunluk ve sıra etkisi kontrol edildiğinde, bu desen random grup desenine göre daha küçük örneklem gerektirdiği için tercih edilebilir bir yöntemdir (Kolen & Brennan, 2004). Bu desen daha çok benzer yetenek düzeylerinin başarısının eşitlendiği yatay eşitleme çalışmalarında kullanılır. Tablo 1.1’de tek grup deseni ve dengelenmiş tek grup deseni gösterilmiştir.

Tablo 1.1: Tek Grup ve Dengelenmiş Tek Grup Deseni

	Form X	Form Y
Tek grup deseni		
Grup A	✓	✓
Dengelenmiş Tek Grup deseni		
Grup A	✓*	✓
Grup B	✓	✓*

*Önce uygulanan test

1.5.2.2 Random Grup Deseni

Random grup deseninde, benzer özelliklere sahip iki grup random olarak seçilir. Seçilen bu gruplar da random olarak ikiye ayrılır ve testlerin her bir yarısı bir gruba uygulanır. Grup A'ya, X_1 testini aldıktan sonra Y_2 testi, grup B'ye Y_1 testini aldıktan sonra X_2 testi uygulanır. Bu sayede iki gruptan gelen farklılık azaltılarak eşitleme hatası azaltılır, eşdeğer grupların karşılaştırılmasına olanak sağlanmış olur. Formlardan elde edilen puanlar üzerinden bağlama çalışması yapılır. Random grup deseninde gruplar birer form aldığı için bu desenin uygulama zamanı, bireylerin birden fazla formu aldığı tek grup deseninden ve dengelenmiş tek grup deseninden daha kısadır. Böylece aynı gruba üst üste iki formun uygulandığında oluşan yorgunluktan kaynaklanan hata etkisi ortadan kaldırılmış olur. Bu desenin tek grup desenine ve dengelenmiş tek grup desenine göre dezavantajı random olarak seçilen eşdeğer grupların yetenek dağılımlarındaki farklılıktır. Bağlanacak iki formdaki performans düzeyleri arasındaki fark, iki formun güçlük düzeyi arasındaki farkı göstermektedir (Kolen ve Brennan, 2004). Bu desen Tablo 1.2’de gösterilmiştir.

Tablo 1.2: Random Grup Deseni

	Form X		Form Y	
	X ₁	X ₂	Y ₁	Y ₂
Grup A	✓			✓
Grup B		✓	✓	

1.5.2.3 Ortak Madde Deseni

Denk olmayan gruplarda ortak madde deseni, ilköğretim başarı ve yetenek test bataryalarının örtüşen yapısından yararlanarak, ortak maddeler üzerinden grupların performanslarının karşılaştırılmasına olanak sağlar. Cevaplayıcılar eşitlenecek formlardan sadece birini cevaplar. Her sınıf düzeyi için o sınıf düzeyine uygun test geliştirilir ve her test ait olduğu sınıf düzeyine uygulanır. Bireylerin bitişik test düzeylerinde ortak maddelerindeki başarıları, bir sınıftan diğerine oluşan büyümenin miktarını belirlemek için kullanılmaktadır (Kolen & Brennan, 2004). Bu desen denk olmayan iki gruba uygulandığı için denk olmayan gruplarda ortak madde (ya da ortak test) deseni (NEAT) olarak adlandırılır (von Davier, Holland & Thayer, 2004). Bu desen Tablo 1.3'te gösterilmiştir.

Tablo 1.3: Denk Olmayan Gruplarda Ortak Madde Deseni

	Form X	Ortak Test/Maddeler	Form Y
Grup A	✓	✓	
Grup B		✓	✓

Ortak madde deseni iki şekilde uygulanabilir. Eğer ortak maddeler test formlarının içerisinde tek bir form olarak tüm gruplara uygulanırsa bu maddelere **iç ortak madde**, grupların kendi düzeylerine uygun olarak aldıkları test formlarının yanı sıra verilen ikinci bir form olarak uygulanırsa **dış ortak madde** olarak adlandırılır. İç ortak maddeler cevaplayıcının test puanının bir parçası olarak hesaplanır, fakat dış ortak maddeler puanlamaya katılmaz ve bağlama için kullanılır (Kolen & Brennan, 2004; Crocker & Algina, 1986).

Ortak maddeler iki cevaplayıcı grup arasındaki farklılıkları değerlendirmede kullanılacağı için dikkatli seçilmelidir. İçerik, ayırıcılık gücü, madde güçlüğü gibi özellikleri bakımından tüm test maddeleri ile benzer özellikler içermesi gerekmektedir (Lord, 1980). Ortak maddeler doğru bir şekilde seçilirse, bu desen

tek grup desenindeki veya dengelenmiş tek grup desenindeki bağlama hatalarını azaltır (Holland & Dorans, 2006; Hambleton, Swaminathan & Rogers, 1991). Ortak maddeler tüm testlerde aynı yerde yer almalı ve birbiriyle eşdeğer olduğu düşünülen maddeler ortak madde olarak kullanılmamalı, birebir aynı olmalıdır (Kolen & Brennan, 2004).

Dikey ölçekleme çalışmalarında ortak maddelerin hangi seviyeden olacağı konusu da tartışmalı bir konudur. Maddeerin düşük sınıf düzeyine göre mi, yüksek sınıf düzeyine göre mi, yoksa her iki sınıf düzeyine göre mi belirleneceği önemli bir konudur. Alanyazında, ortak maddelerde üst sınıf düzeyinden maddeler yer aldığına, alt sınıf düzeyindeki öğrencilerin performansının öğrenmedikleri konularla belirlenmesinin dezavantajları olduğu vurgulanarak eleştirilmektedir (Tong & Kolen, 2007).

Ortak madde deseninde dikkat edilmesi gereken bir diğer konu da ortak madde sayısıdır. Ortak madde sayısı eşitleme sonuçlarını ve eşitleme hastasını etkilemektedir. Hambleton, Swaminathan ve Rogers (1991), ortak madde sayısının tüm testteki madde sayısının %20 - %25'i kadar olması gerektiğini belirtirken, Angoff (1971), ortak madde sayısının tüm testteki madde sayısının %20'si kadar olmasının yeterli olduğunu belirtmiştir. Smith ve Kramer (1992), bir tek madde bile olsa ortak madde için yeterli olacağını savunurken, Kolen ve Brennan (2004) ortak madde sayısının mümkün olduğunca fazla olması gerektiğini, ortak madde sayısı arttıkça hatanın arttığını belirtmişlerdir.

Dikey ölçekleme çalışmalarında bir sınıf düzeyinden diğer sınıf düzeyine öğrenci performansındaki değişimin belirlenmesinde, ortak maddeler farklı sınıf düzeylerinin puanlarının birbiriyle ilişkilendirilmesinde kullanılır. Aşağıda Şekil 1.1'de dikey ölçekleme şematize edilmiştir.

Sınıf Seviyesi	Test Maddeleri
3	a
4	a b
5	a b c
6	a b c d
7	a b c d e
8	a b c d e f g

(Kaynak: Kolen&Brennan, 2004)

Şekil 1.1: Ortak Madde Deseninin Gösterilişi

Bir sınıf temel düzey olarak seçilir ve bu sınıf düzeyi ile diğer sınıf düzeylerine ortak maddeler sorulur, bu temel sınıf sayesinde sınıflar Şekil 1.1’de görüldüğü gibi zincirleme olarak birbirine bağlanır.

1.5.2.4 Eşdeğer Gruplar Deseni

Her bir öğrenci rastgele olarak kendi seviyesine uygun testi yada bitişik sınıf seviyesine uygun olan testini, yanı sıra kendi seviyelerine uygun olan testi de alır.

Eşdeğer grup deseninde her sınıf seviyesindeki öğrenciler kendi sınıf seviyelerine uygun veya ardışık sınıf seviyesindeki testleri alırlar. Hangi öğrencinin hangi sınıf düzeyindeki testi alacağı rastgele belirlenir. Rastgele eşitlenen gruplar boyunca ard arda gelen sınıf seviyelerinde öğrencilerin gösterdikleri performans bir sınıf düzeyinden diğer sınıf düzeyine ne kadar gelişme gösterdiğinin belirlenmesine yardımcı olur. Bu desende de bir grubun temel grup olarak belirlenmesi gerekmektedir. Sınıf düzeyleri arasında zincirleme yapılarak bu uygulamadan elde edilen veri tüm test düzeylerindeki puanların aynı ölçeğe yerleştirilmesinde kullanılır. Bu desen art arda gelen sınıf düzeyleri için ortak madde kullanımı gerektirmez. Ortak madde yerine ortak kişileri kullanarak bağlantıyı kurar. Bu desende öğrenciler kendi sınıf düzeyleri dışında testler aldıkları için pratikte sıkça kullanılan bir yöntem değildir (Tong & Kolen, 2007).

1.5.2.5 Ölçekleme Test Deseni (Scaling Design)

Ölçekleme test deseninde tüm sınıf düzeylerine ikişer test uygulanır. Öncelikle her sınıf kendi seviyesine uygun bir test alır. Buna ek olarak, içeriği tüm sınıf düzeylerini kapsayan ortak bir test hazırlanır ve bu teste ölçek testi denir. Her sınıf kendi seviyesine uygun olan testin yanı sıra bu tüm sınıf düzeylerine uygun olarak geliştirilmiş ölçek testini de alır. Her sınıf düzeyindeki puanlar ölçek testine bağlanır ve ortak bir ölçeğe dönüştürülmüş olur. Bu desenin en önemli dezavantajı, en düşük sınıf düzeyindeki öğrencilere de en yüksek sınıf düzeyindeki öğrencilere de aynı maddelerin uygulanmasıdır. Bu desen uygulama ve testlerin geliştirilmesi aşamasındaki zorluklar ve maliyetin fazlalığı sebebiyle yaygın olarak tercih edilmez (Kolen & Brennan, 2004; Tong & Kolen, 2007).

1.5.3 Eşitleme Yöntemleri

Eşitleme yöntemlerinin her biri farklı kuram ve varsayımlara dayanmaktadır. Eşitleme yöntemleri dayandığı kurama göre; Klasik Test Kuramı (KTK)’ya göre

eşitleme yöntemleri ve Madde Tepki Kuramı (MTK)'ya göre eşitleme yöntemleri olarak sınıflandırılmaktadır. KTK'ya göre gözlenen puan, gerçek puanın ve hatanın toplamından oluşmaktadır. KTK'ya dayalı eşitleme yöntemleri ise eşit yüzdelikli eşitleme, doğrusal eşitleme ve ortalama eşitleme olarak üçe ayrılmaktadır. KTK'ya dayalı eşitleme yöntemleri MTK'ya dayalı eşitleme yöntemlerine göre daha az varsayım gerektirdiği için daha pratiktir (Kolen & Brennan, 2004).

Alanyazında, MTK ve KTK'ya dayalı eşitleme yöntemlerinin karşılaştırıldığı çalışmalar bulunmaktadır. Bu çalışmaların bulguları incelendiğinde, MTK ve KTK'nın benzer sonuçlar verdiği bulgusuna ulaşılan çalışmaların yanı sıra; MTK'ya dayalı eşitleme yöntemlerinin daha kararlı ve daha az hatalı sonuçlar verdiğini belirten çalışmalar da yer almaktadır.

1.5.3.1 Klasik Test Kuramına Göre Eşitleme Yöntemleri

1.5.3.1.1 Doğrusal Eşitleme

Doğrusal eşitleme farklı iki test formu olan X ve Y formlarının, ortalama ve standart sapması dışındaki özelliklerin eşit olduğu varsayımına dayanmaktadır. Eşitleme için iki test formunun standart puanları kullanılır. Angoff (1971) doğrusal eşitlemenin tanımını yaparken “eğer iki test eşit standart sapmayı sağlıyorsa, testlerden alınan puanlar da eşittir” demiştir. X ve Y formlarının standart puanları kullanılarak yapılan eşitleme;

μ_X = Birinci test formunun ortalaması

μ_Y = İkinci test formunun ortalaması

σ_X = Birinci test formunun standart sapması

σ_Y = İkinci test formunun standart sapması

A = Eşitleme denkleminin eğim katsayısı

B = Eşitleme denkleminin sabiti

olmak üzere;

$$X = Y - m_Y + m_X$$

$$\frac{X - m_X}{S_X} = \frac{Y - m_Y}{S_Y}$$

$$X = \frac{S_Y}{S_X}(Y - m_Y) + m_X$$

$$X = \frac{S_Y}{S_X} Y - \frac{S_Y}{S_X} m_Y + m_X$$

Son denklem düzenlendiğinde,

$$X = AY + B \quad (\text{Eşitlik 1.1})$$

doğru denklemi elde edilir. Eşitlik 1.1'de görüldüğü gibi eşitleme denkleminin eğim katsayısı A , $\frac{S_Y}{S_X}$ oranına; eşitleme denkleminin sabiti B ise $m_X - \frac{S_Y}{S_X} m_Y$ 'ye eşittir.

Bu dönüşüm eşitliği X formundan Y formuna, Y formundan X formuna puan dönüştürmede kullanılabilir, bir başka deyişle simetriktir (Angoff, 1971).

Benzer yetenek düzeyine sahip gruplara farklı test formları uygulandığında, doğrusal eşitleme yönteminin uygulanması uygundur (Donlon, 1984). Doğrusal eşitlemede eşitleme sonuçlarının gruba bağımlı olması doğrusal eşitlemenin bir sınırlılığıdır (Livingston, 2004). Test sonuçları eşitlenecek olan öğrenci gruplarının yetenek düzeyleri eşit olmadığı zaman, doğrusal eşitlemenin diğer bir türü olan Levine Doğrusal Eşitleme yönteminin daha uygun olabileceğini belirtmiştir (Donlon, 1984).

1.5.3.1.2 Ortalama Eşitleme

Ortalama eşitleme yönteminde, eşitlenecek olan testlerin güçlükleri arasında fark olduğu fakat bu farkın puan ölçeği boyunca sabit olduğu kabul edilmektedir. Bir başka deyişle, X ve Y formlarının ortalamaları arasındaki farkın tüm ölçek boyunca sabit olduğu kabul edilir. Bu yöntem eşitlenecek testlerdeki puanların ortalamaya uzaklıklarının eşit olduğu varsayımına dayanmaktadır (Kolen & Brennan, 1995). X ve Y formları için ortalama eşitleme,

μ_X = Birinci test formunun ortalaması

μ_Y = İkinci test formunun ortalaması

olmak üzere;

$$X - m_x = Y - m_y$$

$$Y = X - \mu_x + \mu_y \quad (\text{Eşitlik 1.2})$$

şeklinde ifade edilir (Kolen, 2007). Ortalama eşitlemede eşitleme yapabilmek için eşitlenecek formların ortalama değerleri kullanılırken, doğrusal eşitlemede test formlarının ortalamaları ve standart sapmaları kullanılır. Eşitlenecek formların standart sapmaları eşitse, ortalama eşitleme ve doğrusal eşitleme ile yapılan dönüşümler aynı sonucu verecektir (Kolen, 1988).

1.5.3.1.3 Eşit Yüzdellikli Eşitleme

Eşit yüzdellikli eşitleme, iki testin puan dağılımlarında aynı yüzdellik sıraya denk gelen puanların eşitlenmesine dayanmaktadır. Angoff (1981), eşit güvenilirliğe sahip aynı özelliği ölçen iki farklı X ve Y formlarından elde edilen puanların, aynı yüzdellik sıraya denk gelen puanların eşdeğer olarak kabul edilebileceğini belirtmiştir.

Eşit yüzdellikli eşitleme yapabilmek için öncelikle her bir formdaki puanlar yüzdellik sıraya çevrilir. Sonra ilk formda p. yüzdellik sıraya denk gelen puan ile ikinci formda p. yüzdellik sıraya denk gelen puanlar eşitlenir (Aiken, 2000).

Eşit yüzdellikli eşitleme yöntemi doğrusal eşitlemeye göre daha karmaşık bir işlem sürecini içermektedir. Çünkü X formundaki bir ham puan ile Y formundaki bir ham puan nadiren aynı yüzdellik sıraya karşılık gelmektedir. Puan dağılımları küçük gruplarda düzensizlik gösterebilir. Aynı yüzdellik sıraya karşılık gelen ham puanları bulmak için puan dağılımı sürekli hale getirilmektedir. Bu doğrultuda, çeşitli düzgünleştirme (smoothing) ve öteleme (interpolation) yöntemleri kullanılmaktadır (Cook ve Petersen, 1987; Akt. Kelecioğlu, 1994). Eğer testlere ait puan dağılımları benzerse, doğrusal eşitleme ve eşit yüzdellikli eşitleme aynı sonucu verecektir (Kolen, 1988).

1.5.3.2 Madde Tepki Kuramına Dayalı Eşitlemede Kullanılan Modeller

KTK'ya dayalı istatistiklerinin bazı zayıf noktaları, madde tepki kuramının ortaya çıkmasına zemin hazırlamıştır. Bu kuram, bireyin yetenek düzeyinin, belirli bir madde grubundan bağımsız olarak kestirilebileceği sayılısı üzerine kurulmuştur (Hambleton, 1985). Klasik test teorisine dayalı eşitlemede eşitlik, tek boyutluluk, simetriklik ve gruptan bağımsızlık varsayımlarının karşılanmasıyla yaşanan

sıkıntılar sonucu elde edilen bulgular anlamlı olmamakta ve analizlerde bazı problemler yaşanmaktadır. Model-veri uyumu sağlandığında MTK ile eşitlemede bu problemlerin üstesinden gelinebilmektedir (Kolen, 1981). MTK'da, ölçülen bir özellik için farklı yetenek düzeyindeki bireylerin bir maddeye verebilecekleri cevabın matematiksel bir model ile açıklaması sunulmaktadır (Cook & Eignor, 1991). KTK'da madde ve test parametreleri örnekleme bağılı iken, MTK'da madde ve test parametreleri testin uygulandığı gruptan bağımsız olarak elde edilmektedir. Bireylerin yetenek düzeyi de maddelere verdikleri cevaplardan bağımsızdır. Bu yüzden bir bireyin, aynı yeteneği ölçen iki test formundan herhangi birini alması, testlerin güçlüğü farklı olsa bile yetenek kestirimini değiştirmeyecektir (Hambleton & Swaminathan, 1985). MTK yöntemlerini anlamak, uygulamak ve açıklamak KTK yöntemlerinden daha zordur fakat MTK yöntemleri daha esnektir (Harris, 2003).

Test formlarının eşitlenmesinde MTK yaygın olarak kullanılan bir yöntemdir. MTK'ya dayalı eşitlemede, iki test formundaki puanlar arasında matematiksel bir ilişki modellenir (Dongyang, 2009). MTK'ya dayalı eşitleme yöntemleri, cevaplayıcıların θ yetenek düzeyi ile cevaplayıcının maddeyi doğru cevaplama olasılığı arasındaki ilişkiyi tanımlayan matematiksel bir fonksiyon olduğu varsayımını temel alır (Kolen & Brennan, 2004).

MTK özellikle dikey ölçekleme çalışmalarında sıkça kullanılmaktadır. MTK'ya dayalı ölçekleme yöntemleri, diğer ölçekleme yöntemlerine göre daha güçlü sayılıtlara sahiptir. Alanyazında tek boyutlu ve çok boyutlu, iki kategorili ve çok kategorili, parametrik ve parametrik olmayan olmak üzere çok çeşitli modeller bulunmaktadır.

Çok boyutlu MTK modelleri kestirimlerin karmaşıklığından dolayı yaygın olarak kullanılmaz. Tek boyutluluk ve yerel bağımsızlık varsayımları olan tek boyutlu modeller dikey ölçeklemede yaygın olarak kullanılan MTK modelleridir (Çetin, 2009). MTK'ya dayalı eşitleme yöntemleri iki kategorili yada çok kategorili puanlanan maddeler için geliştirilmiştir. İki kategorili MTK madde 1-0 olarak puanlandığı durumlarda, çok kategorili MTK, maddeler holistik olarak puanlandığı zaman kullanılırlar (Kolen & Brennan, 2004). Parametrik ve parametrik olmayan yöntemleri karşılaştıran bir çok çalışma bulunmaktadır. Bu çalışmalar genel olarak, daha güçlü varsayımlara sahip olan parametrik modellerin parametrik olmayan modellerden daha kararlı sonuçlar verdiği yönündedir (Lei, Dunbar & Kolen, 2004;

Ansley & Forsyth, 1985). Bu çalışmada tek boyutlu, iki kategorili ve parametrik modeller kullanıldığı için bu yöntemler daha ayrıntılı olarak ele alınmıştır.

İki kategorili puanlanan maddelerde kullanılan MTK yöntemleri bir parametrelili lojistik model, iki parametrelili lojistik model ve üç parametrelili lojistik model olarak geliştirilmiştir.

1.5.3.2.1 Bir Parametrelili Lojistik Model ile Eşitleme (Rasch Modeli)

Sadece bir madde parametresi barındıran bu model en basit MTK modeli olarak bilinen bir parametrelili lojistik model (1PLM)'dir. Her bir maddenin ayırt ediciliği a_i eşittir ve ortalama bir değerdir. Şans başarısı c_i ise 0 olarak kabul edilir. Maddeler sadece madde gücü bakımından farklılaştığı için kestirilecek madde parametresi bir tanedir. Bu model Rasch modeli olarak da adlandırılrsa da 1 parametrelili model ile rasch modelin farkı, rasch modelinde a parametresi 1 olarak alınır fakat 1 parametrelili model de a parametresi ortalama bir değerdir ve 1'den farklı değerler alabilmektedir.

Bir parametrelili lojistik model ile θ yetenek düzeyindeki bir bireyin i maddesini doğru cevaplama olasılığı:

$P_i\theta$: θ yetenek düzeyindeki bir bireyin i maddesini doğru cevaplama olasılığı

b_i : i maddesine ait madde güçlük indeksi

D : Ölçek sabiti (1,7)

olmak üzere;

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

(Eşitlik 1.3)

şeklinde ifade edilebilir.

θ yetenek parametresi kuramsal olarak $-\infty$ ile $+\infty$ arasında değerler alabilir. Fakat pratikte -4 ile +4 arasında değerler almaktadır (Hambleton & Swaminathan, 1985). Slinde ve Linn (1978) öğrenci yeteneğinin heterojen ve şans başarısının minimum olduğu durumlarda Rasch modelinin dikey ölçekleme için uygun olduğunu ortaya koymuştur. Loyd ve Hoover (1980) da bu sonuçları destekler bulgulara ulaşmıştır.

1.5.3.2.2 İki Parametrelili Lojistik Model ile Eşitleme

Bir parametrelili lojistik modelde tüm maddelerin a parametresi olan ayırt edicilik gücü eşit ve 1 olarak varsayılırken, iki parametrelili lojistik modelde (2 PLM) ise modele ayırtedicilik parametresi de dahil edilir. Bireyin maddeyi doğru cevaplama olasılığını maddenin ayırtıcılık gücü indeksi etkiler, şans başarısı yine 0 olarak kabul edilir. İki parametrelili lojistik modelde θ yetenek düzeyindeki bir bireyin i maddesini doğru cevaplama olasılığı:

$P_i\theta$: θ yetenek düzeyindeki bir bireyin i maddesini doğru cevaplama olasılığı

a_i : i maddesine ait madde ayırtıcılık gücü indeksi

b_i : i maddesine ait madde güçlük indeksi

D: Ölçek sabiti (1,7)

olmak üzere;

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

(Eşitlik 1.4)

şeklinde ifade edilebilir.

Kuramsal olarak a parametresi $-\infty$ ile $+\infty$ arasında değerler olsa da pratikte bu pek mümkün olmaz (D.N.M. & van der Kamp, L. J.Th., 2005). Genel olarak 0 ile 2 arasında değerler alabilir (Hambleton, Swaminathan & Rogers, 1991).

1.5.3.2.3 Üç Parametrelili Lojistik Model ile Eşitleme

Bir ve iki parametrelili lojistik model bireylerin şans başarısını dikkate almaz ve 0 olarak kabul ederler. Üç parametrelili lojistik modelde (3PLM) modele cevaplayıcıların doğru cevabı şansa bulma olasılığı da dahil edilir. Üç parametrelili lojistik model ile θ yetenek düzeyindeki bir bireyin i maddesini doğru cevaplama olasılığı:

$P_i\theta$: θ yetenek düzeyindeki bir bireyin i maddesini doğru cevaplama olasılığı

a_i : i maddesine ait madde ayırtıcılık gücü indeksi

b_i : i maddesine ait madde güçlük indeksi

c_i : i maddesine ait şans parametresi

D: Ölçek sabiti (1,7)

olmak üzere;

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

(Eşitlik 1.5)

şeklinde ifade edilebilir.

Baker (2001) çalışmasında, bu modelin özellikle şans başarısının etkili olduğu testlerde kullanılmasının uygun olacağını belirtmiştir.

Madde tepki kuramına göre ölçekleme yapılırken parametre sayısı (1, 2 ya da 3 parametrelili lojistik model), kalibrasyon yöntemi (ortak ya da ayrı), yetenek kestirim yöntemi (EAP, MAP ya da ML) gibi çok sayıda yöntem karar vermek gerekmektedir. MTK'ya dayalı ölçeklemenin karmaşık ve tutarsız sonuçlar vermesi, bu kararların birbirleriyle olan ilişkileri ile ilgilidir (Tong, 2005: Kolen & Brennan, 2004).

1.5.3.3 Madde Tepki Kuramına Dayalı Ölçek Dönüştürme Yöntemleri

Uygun olan ölçekleme yöntemi seçildikten sonra, madde ve yetenek kestirimleri farklı kalibrasyon yolları ile elde edilir (Brennan, 2004). Farklı gruplardan elde edilen verileri ortak bir ölçeğe yerleştirmek için çeşitli kalibrasyon yöntemleri kullanılmaktadır. Dikey ölçeklemede ortak madde deseni kullanıldığında ortak olan maddeler bir sınıf seviyesinden diğer sınıf seviyesine test seviyelerinin ortak bir ölçekte bağlanmasını sağlar. Ortak madde test deseninde, MTK parametreleri her test düzeyi için kullanılan program ayrı ayrı çalıştırılarak kestirilmekte ya da program bir defa çalıştırılarak parametreler aynı anda kestirilmektedir (Kolen & Brennan, 2004). Bu kalibrasyon yöntemleri eş zamanlı (concurrent) kalibrasyon ve ayrı (separate) kalibrasyon yöntemleri olarak adlandırılmaktadır (Meng, 2007).

1.5.3.3.1 Eş Zamanlı (Ortak) Kalibrasyon

Eş zamanlı kalibrasyonda bütün sınıf düzeylerindeki verinin aynı anda kalibre edilmesiyle dikey ölçek oluşturulmaktadır. Ortak maddelerin ardışık sınıf düzeylerinde aynı madde parametrelerine sahip olduğu varsayılarak formların madde parametreleri birlikte kestirilir (Meng, 2007). Ortak kalibrasyon ile her bir seviyeye ait madde parametreleri ve yetenek kestirimleri dönüşümü tek bir ölçekte elde edilmektedir. Böylece tekrar yetenek kestirimleri için ölçek dönüşümü

yapılmasına gerek kalmaz. Tek bir dosyada birleştirilen ardışık sınıf seviyelerine ait veriler tek bir analizle ölçeklenmiş olur. Bunun için öncelikle herhangi bir sınıf referans sınıf olarak belirlenmekte ve ardışık sınıf seviyelerine ait ölçeklenmiş yetenek kestirimleri ortalaması 0 ve standart sapması 1 olan bir ölçeğe dönüştürülmektedir (Çetin, 2009). Hedef testteki ortak maddeler için madde parametreleri referans testin değerlerine ayarlanıp tekrar tahmin edilmektedir. Ortak maddeler için madde parametreleri bilinmekte ve MTK kalibrasyonları hedef testte ortak olmayan maddeleri referans test ölçeğine yerleştirmektedir (Meng, 2007).

1.5.3.3.2 Ayrı Kalibrasyon

Ayrı kalibrasyon kullanıldığında MTK parametreleri her sınıf düzeyi için ayrı ayrı hesaplanmaktadır. İki farklı test formu için ayrı ayrı kestirilen madde ve yetenek parametreleri farklı ölçekte oldukları için karşılaştırılmaları uygun olmamaktadır. Karşılaştırmanın yapılabilmesi amacıyla, bir sınıf düzeyi referans sınıf düzeyi olarak seçilmekte ve bir sınıf düzeyi için θ ölçeği başlangıç ölçeği olarak belirlenmektedir. Denk olmayan gruplarda ortak madde deseni ile ortak maddeler kullanılarak, madde ve yetenek parametreleri kestirimler bir dizi doğrusal dönüşüm kullanarak başlangıç ölçeğine yerleştirmek için kullanılmaktadır (Kolen & Brennan, 2004). Ortak maddelerin ayırıcılık gücü indeksi ve madde güçlük indeksi kullanılarak elde edilen A ve B sabitleri kullanılarak bir formdaki θ yetenek düzeyini diğer formdaki θ yetenek düzeyine dönüştürülmekte ve iki form arasında dönüşüm sağlanmaktadır. Dikey ölçeklemede ardışık sınıf seviyeleri için zincirleme olarak tekrar edilerek ölçek dönüşümü sağlanmaktadır. Test I'daki i kişisi için θ değerinin test J'deki karşılığı:

θ_{ji} = i kişisinin J testine ait yetenek düzeyi

θ_{ii} = i kişisinin I testine ait yetenek düzeyi

A= Eşitleme denkleminin eğim katsayısı

B= Eşitleme denkleminin sabiti

olmak üzere,

$$\theta_{ji} = A\theta_{ii} + B \quad (\text{Eşitlik 1.6})$$

şeklinde ifade edilebilir.

İki test formuna ait madde parametreleri de birbirine dönüştürülür. I ve J testlerindeki j maddesine ait madde parametrelerinin dönüşümleri;

a_{Jj} , b_{Jj} ve c_{Jj} = j maddesi için J testinden elde edilen madde parametreleri

a_{Ij} , b_{Ij} ve c_{Ij} = j maddesi için I testinden elde edilen madde parametreleri

A = Eşitleme denkleminin eğim katsayısı

B = Eşitleme denkleminin sabiti

olmak üzere;

$$a_{Jj} = \frac{a_{Ij}}{A}$$

$$b_{Jj} = Ab_{Ij} + B$$

$$c_{Jj} = c_{Ij} \quad (\text{Eşitlik 1.7})$$

şeklinde ifade edilebilir. Yukarıdaki eşitliklerde de görüldüğü gibi şans parametresi (c) ölçek dönüşümünden bağımsızdır (Kolen & Brennan, 2004).

Ayrı kalibrasyonda sonuçlarını ortak bir ölçeğe yerleştirmek için birçok bağlama prosedürü geliştirilmiştir. Bu prosedürlerde doğrusal ve doğrusal olmayan dönüşüm yöntemleri izlenmektedir fakat doğrusal olmayan dönüşüm yöntemleri seviyeler arasındaki ortalama farklarını değiştireceği için kullanılmamaktadır.

İki kategorili MTK modelleri ile dikey ölçekleme yapılırken denk olmayan gruplar için ortak madde deseni kullanıldığında ayrı kalibrasyonda kullanılan dört temel doğrusal dönüşüm yöntemi; Moment yöntemleri, Karakteristik eğri yöntemleri, minimum ki kare yöntemi ve en küçük kareler yöntemidir (Kim, 2007). Alanyazında ayrı kalibrasyon için yaygın olarak kullanılan yöntemler moment yöntemleri ve karakteristik eğri yöntemleridir.

1.5.3.3.2.1 Moment Yöntemleri

Ortalama-ortalama (mean/mean-MM) (Loyd & Hover, 1980) ve ortalama-standart sapma (mean/sigma-MS) (Macro, 1977) yöntemleri moment yöntemleri olarak adlandırılır ve her iki yöntem de ölçekleme sabitlerini hesaplamada kullanılır (Meng, 2007). Moment yöntemlerinde ortak maddelerin a ve b madde

parametrelerinin ortalaması ve standart sapması kullanılır. Bu bağlama yöntemleri Rasch modelde daha yaygın olarak kullanılmaktadırlar.

Ortalama-ortalama yönteminde ölçekleme sabitleri, ortak maddelerden elde edilen madde ayırıcılık parametrelerinin ortalaması kullanılarak kestirilir. Ortalama-ortalama yöntemi ile katsayılar:

$\mu(a_I)$ = I testindeki ortak maddelerden elde edilen a parametrelerinin ortalaması

$\mu(a_J)$ = J testindeki ortak maddelerden elde edilen a parametrelerinin ortalaması

olmak üzere;

$$A = \frac{\mu(a_I)}{\mu(a_J)}$$

(Eşitlik 1.8)

şeklinde hesaplanır.

Ortalama-standart sapma yönteminde ölçekleme sabitleri, ortak maddelerden elde edilen madde güçlük parametrelerinin ortalaması ve standart sapması kullanılarak kestirilir. Ortalama-ortalama yöntemi ile katsayılar:

$\sigma(b_I)$ = I testindeki ortak maddelerden elde edilen b parametrelerinin standart sapması

$\sigma(b_J)$ = J testindeki ortak maddelerden elde edilen b parametrelerinin standart sapması

$\mu(a_I)$ = I testindeki ortak maddelerden elde edilen a parametrelerinin ortalaması

$\mu(a_J)$ = J testindeki ortak maddelerden elde edilen a parametrelerinin ortalaması

olmak üzere;

$$A = \frac{\sigma(b_J)}{\sigma(b_I)} \quad \text{ve} \quad B = \mu(b_J) - A\mu(b_I)$$

(Eşitlik 1.9)

şeklinde hesaplanır.

Bu iki moment yöntemi kestirim için farklı madde parametrelerini kullanmakta ve bu nedenle hesaplamalar ile farklı sonuçlar elde edilmektedir. Madde güçlük parametresi kestirimi, madde ayırıcılık parametresi kestirimine göre daha kararlı sonuçlar verdiği için bazı çalışmalarda ortalama-standart sapma yöntemi, ortalama-ortalama yöntemine tercih edilir (Gök, 2012).

1.5.3.3.2 Karakteristik Eğri Dönüştürme Yöntemleri

Stocking-Lord metodu (SL) (Stocking & Lord,1983) ve Haebara metodu (Haebara, 1980) karakteristik eğri dönüştürme yöntemleri olarak adlandırılmaktadırlar (Meng, 2007).

Norman-Dvorak (2009) ortalama-ortalama ve ortalama-standart sapma yöntemlerinde madde parametrelerinin hepsi aynı anda hesaplamalara katılmadığı için hatalı sonuçlara yol açabildiğini belirtmiş ve Haebara (1980) bu hataya çözüm olarak madde parametrelerinin tümünü dikkate alan Haebara yöntemini geliştirmiştir. Stocking ve Lord (1983) ise sonraki yıllarda aynı soruna çözüm olarak Haebara'nın yöntemine benzer farklı bir yöntem olarak Stocking Lord yöntemini geliştirmişlerdir. Bu yöntemler ortak maddelerin karakteristik eğrileri arasındaki farkların karelerini kullanarak, madde ya da test karakteristik eğrileri arasındaki farkı azaltır.

Haebara yönteminde;

P_{ij} = j. cevaplayıcının i. maddeyi doğru cevaplama olasılığı (i maddesinin j. cevaplayıcı için madde karakteristik fonksiyonu:

m = ortak madde sayısı

a_{Jj}, b_{Jj}, c_{Jj} = j. ortak madde için J testine ait madde parametreleri

a_{Ij}, b_{Ij}, c_{Ij} = j. ortak madde için I testine ait madde parametreleri

olmak üzere;

$$H(\theta_i) = \sum_{j=1}^m \left[p_{ij}(\theta_i, a_{Jj}, b_{Jj}, c_{Jj}) - p_{ij}(\theta_i, a_{Ij}, b_{Ij}, c_{Ij}) \right]^2 \quad (\text{Eşitlik 1.10})$$

şeklinde ifade edilir. Formüldeki ikinci forma ait madde parametreleri, ayrı kalibrasyon başlığında anlatılan eşitlikler (Eşitlik 1.7) yardımı ile yeniden ölçeklenerek Haebara fonksiyonuna yerleştirildiğinde fonksiyon:

$$H(\theta_i) = \sum_{j=1}^m \left[p_{ij} \left(\theta_i, a_{j_j}, b_{j_j}, c_{j_j} \right) - p_{ij} \left(\theta_i, \frac{a_{I_j}}{A}, Ab_{I_j} + B, c_{I_j} \right) \right]^2 \quad (\text{Eşitlik 1.11})$$

şeklinde yeniden ölçeklenir. Haebara fonksiyonu tüm bireyler için hesaplanır (Kolen & Brennan, 2004).

Stocking Lord yönteminde, Haebara yönteminin tersine, kareler arası farkın toplamı kullanılır. Yine yukarıdaki ifadeler kullanılarak Stocking Lord yöntemi:

$$SL(\theta_i) = \left[\sum_{j=1}^m p_{ij} \left(\theta_i, a_{j_j}, b_{j_j}, c_{j_j} \right) - \sum_{j=1}^m p_{ij} \left(\theta_i, a_{I_j}, b_{I_j}, c_{I_j} \right) \right]^2 \quad (\text{Eşitlik 1.12})$$

şeklinde ifade edilir. Formüldeki ikinci forma ait madde parametreleri, ayrı kalibrasyon başlığında anlatılan Eşitlik 1.7 yardımı ile yeniden ölçeklenerek Stocking-Lord fonksiyonuna yerleştirildiğinde fonksiyon:

$$SL(\theta_i) = \left[\sum_{j=1}^m p_{ij} \left(\theta_i, a_{j_j}, b_{j_j}, c_{j_j} \right) - \sum_{j=1}^m p_{ij} \left(\theta_i, \frac{a_{I_j}}{A}, Ab_{I_j} + B, c_{I_j} \right) \right]^2 \quad (\text{Eşitlik 1.13})$$

şeklinde yeniden ölçeklenir. Stocking-Lord fonksiyonu tüm bireyler için hesaplanır (Kolen & Brennan, 2004).

Stocking-Lord (SL) yöntemi üç parametrelili model için daha sık kullanılmaktadır. SL metodunda ortak maddelerden elde edilen test karakteristik eğrileri arasındaki farkların karesi kullanılarak verilen θ hesaplamaları yapılmaktadır (Kim, Lee, Kim & Kelly, 2009).

1.5.3.4 Madde Tepki Kuramına Göre Yetenek Kestirim Yöntemleri

Madde parametreleri uygun bir kalibrasyon yolu ile ortak bir ölçeğe dönüştürüldükten sonra yetenek düzeyi kestiriminin yapılacağı yöntemin belirlenmesi gerekmektedir. Madde tepki kuramına göre θ yetenek düzeyi kestirimi yapılırken, toplam puan ya da örüntü puanlama kullanılabilir. Daha pratik ve kolay

bir yöntem olan toplam puan yöntemi örüntü puanlama yöntemine göre daha çok kullanılır. Fakat örüntü puanlamaya göre daha çok hata payı içerir ve daha az bilgi verir (Tong & Kolen, 2010).

MTK'da 1-0 şeklinde kodlanan iki kategorili maddelerin yetenek kestiriminde sıklıkla üç farklı yetenek kestirim yöntemi kullanılır. Bunlar; maksimum olabilirlik kestirimi (Maximum Likelihood Estimation – ML), maksimum önsel dağılım (Maximum A Posteriori – MAP) ve beklenen önsel dağılım (Expected A Posteriori – EAP) kestirim yöntemleridir.

En yaygın kullanılan yeterlilik kestirimi olan ML, asimptotik olarak yansızdır. Buna rağmen, “0” ve tam puana karşılık gelen bir ML bulunmamaktadır. Bir başka deyişle, birey tamamen yanlış ya da tamamen doğru yanıt verdiğinde ML yöntemi ile kestirim yapılamamaktadır. Bayesian EAP ve MAP kestirim metotları her türlü yanıt deseni için kullanılabilirler ve tipik olarak ML kestirimine göre daha küçük MSE'ye sahiplerdir. Ancak bu metotlar kendi ortalamalarına yakınsayacak şekilde yanlıdır (Lord, 1986). EAP ve MAP kestirim yöntemleri tipik olarak ortalama yeterliliğe ML yöntemine göre daha yakındır. Yeterlilik kestirimi üzerine geliştirilen Bayesian temelli yaklaşımlar yeterlilik kestirimine yetenek dağılımını katarak ML'den daha farklı bir yöntem ortaya koyarlar. Bunun sonucunda, bu prosedürler, araştırma evreninin yeterlilik dağılımı üzerindeki varsayımlarıyla, maddelerin doğru ve yanlış olarak elde edilmesi olabilirliği hakkındaki bilgileri birleştiren yeterlilik kestirimleri üretirler. Bu varsayılan yeterlilik dağılımı önsel dağılım olarak adlandırılır. Yanıt desenleri ile ilgili olabilirlik, sonsal dağılım olarak adlandırılan düzeltilmiş bir dağılım üretmek üzere önsel dağılım hakkındaki bilgi ile birleştirilir. Bayesian tabanlı kestirimler, EAP ve MAP, EAP'ın sonsal dağılımın ortalamasını alması ve MAP'ın sonsal dağılımın modunu alması noktasında ayrılırlar.

Thissen ve Orlando (2001) ve Yen (1984), toplam puanlar üzerinden puanlama (summed scoring) ile pseudo-MLE ya da pseudo-EAP kestirimlerinin nasıl hesaplanacağını ortaya koymuşlardır. Bunu anlamak için, hem üçüncü sınıf öğrencisinin hem de dördüncü sınıf öğrencisinin aynı test düzeyini aldığını ve aynı test maddelerini doğru bir şekilde yanıtladığını varsayalım (Kolen & Brennan, 2004). Üçüncü sınıf öğrencisinin EAP sonucu üçüncü sınıf yeterlilik dağılımı ile, dördüncü sınıf öğrencisinin EAP sonucu dördüncü sınıf yeterlilik dağılımı ile

hesaplanır. Üçüncü sınıf öğrencileri için ortalama yeterlilik dördüncü sınıf öğrencileri için olandan daha düşük olduğundan, üçüncü sınıf öğrencilerinin EAP kestirimi, dördüncü sınıf öğrencileri için olan kestirimden daha düşük olacaktır. Bu örnek, EAP kestirimi kullanımının, dikey ölçekleme durumunda önemli kullanım endişeleri yaratabileceğini ortaya koymaktadır. EAP'tan farklı olarak, ML hem üçüncü hem de dördüncü sınıf öğrencileri için aynı olacaktır.

MAP kestirimi, EAP kestirimi ile aynıdır ancak MAP kestirimi ortalama yerine sonsal dağılımın modunu kullanır. MAP kestirimi, tıpkı EAP gibi, tüm yanıt desenleri için kullanılmaktadır ve genel olarak araştırma evreni ortalamasına doğru yanlıdır, çünkü θ (theta) kestirimleri öncellerin kullanımıyla otomatik olarak bir aralığa sınırlandırılmıştır (Lord, 1986). Buna rağmen, sonsal dağılımın modu ve ortalaması aynı olmadıkça, MAP kestirimi, EAP kestiriminden bir şekilde daha büyük bir ortalama standart sapmaya sahiptir (Lord, 1986).

1.5.4 Değerlendirme Ölçütleri

Ölçekleme çalışmasının son aşaması elde edilen sonuçların karşılaştırma aşamasıdır. Gelişimsel ölçek puanlarının normatif özellikleri bir çok tartışmanın ve çalışmanın konusu olmuştur. Dikey ölçekleme çalışması sonuçlarını karşılaştırabilmek için ölçek puanlarının özellikleri karşılaştırılmaktadır. Bu özellikler bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme (grade-to-grade growth), sınıf düzeyleri arasındaki çeşitlilik (grade-to-grade variability) ve düzey dağılımları arasındaki ayrımı (separation of grade distributions)'dir.

Bir sınıf düzeyinden diğer sınıf düzeyine olan büyümeyi belirlemek için ard arda olan sınıf düzeylerinin ortalamaları arası farklar kullanılmaktadır. Alternatif olarak medyanlar ya da seçilen yüzdeler puanlar kullanılabilir (Kolen & Brennan, 2004). Ölçekleri karşılaştırmak için ardışık sınıf düzeylerinin ortalamalarının testlerin içeriğine, analizde kullanılan kalibrasyon yöntemi ya da yetenek kestirimi yöntemlerine bakılmaksızın artması beklenmektedir (Kim, 2007).

Sınıf düzeyleri arasındaki farklılığı belirlemek için ardışık sınıflarda sınıfların standart sapmaları kullanılır. Her sınıf düzeyinin kendi içerisindeki değişkenliğine bakılmaktadır. Alternatif olarak diğer değişkenlik ölçüleri kullanılabilir (Kolen & Brennan, 2004).

Düzey dağılımlarının ayırımı Holland (2002) dağılımlar arasındaki boşluklar olarak tanımlarken, Hoover (1984a, 1984b) ve Petersen vd. (1989) dağılımların örtüşme derecesi olarak tanımlamıştır. Düzey dağılımları arasındaki ayırımı yorumlamak için kullanılan istatistiksel yöntemlerden biri Yen (1986) tarafından tavsiye edilen etki büyüklüğü indeksidir. Etki büyüklüğü indeksi,

$$\mu(Y)_{üst} = \text{Üst sınıfın ortalaması}$$

$$\mu(Y)_{alt} = \text{Alt sınıfın ortalaması}$$

$$\sigma^2(Y)_{üst} = \text{Üst sınıfın varyansı}$$

$$\sigma^2(Y)_{alt} = \text{Alt sınıfın varyansı}$$

olmak üzere;

$$\text{Etki büyüklüğü} = \frac{\mu(Y)_{üst} - \mu(Y)_{alt}}{\sqrt{\frac{(\sigma^2(Y)_{üst} + \sigma^2(Y)_{alt})}{2}}}$$

(Eşitlik 1.14)

şeklinde hesaplanır (Kolen & Brennan, 2004). Etki büyüklüğü standardize edilmiş ortalamalar farkıdır ve sınıf varyanslarının ortalamasının karekökü kullanılarak hesaplanır. Etki büyüklüğünün küçük olması ard arda gelen sınıf düzeylerinin yetenek düzeyleri arasındaki farkın az olduğunun göstergesi olarak kullanılmaktadır. Etki büyüklüğü 0-0.20 arasında değer alıyorsa zayıf etki, 0.21-0.50 arasındaki değerler için küçük etki, 0.51-1 orta etki ve 1'den büyük değerler için ise güçlü etki olarak tanımlanmaktadır (Cohen, Manion & Morrison, 2007).

2. İLGİLİ ARAŞTIRMALAR

Wang (2013) araştırmasında dikey ve yatay ölçekleme yaklaşımlarını kullanarak yapılan yetenek kestirimlerinin sonuçlarını karşılaştırmıştır. Bu karşılaştırma, form içi sınıflar arası, sınıf içi formlar arası, sınıflar arası ve formlar arası ölçek kararlılığı izlenerek yapılmış ve böylece sınıf içi değişkenlik desenlerinin ve katılımcı örnek niteliklerinin hangi koşulları altında, ölçekleme yaklaşımlarından birinin diğerine üstünlük kurduğu araştırılmıştır. Bu çalışmayla elde edilen bulgular ışığında uygulayıcılara dikey bir ölçek oluşturulduktan sonra, bu ölçeğin nasıl sürdürüleceğine dair önemli deneysel bir rehberlik sağladığı ifade edilmiştir. Eğer ölçeğin amacı, sınıflar arası ayırımı, sınıf içi değişkenliği ve gelişimi yakalamak ise, dikey ölçekleme yaklaşımı bu hedeflere ulaşmakta marjinal en iyi yaklaşım olarak görülmektedir.

Altun (2013) araştırmasında gerçek veri seti kullanarak ortak madde deseni ile ayrı ve eş zamanlı kalibrasyon yöntemleri ve Beklenen Önsel Bilgi (EAP), Maksimum Önsel Bilgi (MAP) ve Maksimum Olabilirlik (ML) yetenek kestirim yöntemleri kombinasyonu ile elde edilen dikey ölçekleme sonuçlarını karşılaştırmıştır. Değerlendirme kriterleri olarak, ortalamalar, ortalamalar arası farklar, etki büyüklüğü ve yatay uzaklıklar hesaplanmış ve yapılan seçimlerin dikey ölçeklemeyi nasıl etkilediği belirlenmeye çalışılmıştır. Araştırma bulgularına göre kalibrasyon yöntemlerinin benzer sonuçlar verdiği ancak etki büyüklüğü ve yatay uzaklık değerlerinin ayrı kalibrasyonda daha düşük olduğu ifade edilmiştir. Yetenek kestirim yöntemleri kıyaslandığında ortalamalar açısından MAP'ın çoğunlukla daha düşük değerler ürettiğini, etki büyüklüğü ve yatay uzaklıklar incelendiğinde kalibrasyon yöntemlerinin benzer örüntüler sergilemediği ve örüntünün farklılaştığını ifade etmiştir.

Wysel ve Reckase (2011)'in çalışmasında, dikey ölçeklemede farklı kalibrasyon yöntemlerinin karşılaştırılması amacı ile gerçek veri seti ve simülasyon veri seti kullanılmıştır. Elde edilen bulgulara göre, sabit kalibrasyon ve ayrı kalibrasyonun (SL metodu ile) en iyi sonucu verdiği ifade edilmiştir. Eş zamanlı kalibrasyonda kullanılan veri sayısı ayrı kalibrasyona göre daha fazla olduğu için genellikle daha iyi sonuç verdiği belirtilmiştir. Ancak eş zamanlı kalibrasyon yönteminde test

formları arasında ölçülen yapının eşit olmaması veya tek boyutluluk varsayımının ihlali nedeniyle hatalı sonuçlar elde edilebildiği, bu kalibrasyon yönteminde yanlılık görülebildiği vurgulanmıştır.

Lee ve Ban (2010) tarafından yapılan çalışmada gerçek veri seti parametrelerine benzer parametrelerle simule edilen yapay veri seti kullanılmıştır. Test karakteristik eğrisi yöntemleri ve beklenen gözlenen puan dağılımı ile bağlama yöntemlerinin karşılaştırılması amaçlanmıştır. Elde edilen bulgulara göre, ayrı kalibrasyonda Heabera ile SL yöntemi kıyaslandığında, Heabera'nın daha düşük bağlama hatası ürettiği ifade edilmiştir.

Kim, Lee, Kim ve Kelley (2009) tarafından yapılan çalışmada simülasyon ile türetilmiş yapay veri seti kullanılmıştır. Çalışmada farklı koşullar altında (kalibrasyon yöntemleri, bağlama yöntemleri, yetenek kestirimleri, ortak madde sayısı, her sınıf düzeyi için örneklem büyüklüğü ve farklı sınıfların temel sınıf olarak alınmasıyla) oluşturulan dikey ölçekler karşılaştırılmıştır. Elde edilen bulgulara göre, kullanılan iki kalibrasyon yönteminde sonuçların benzer olduğu ifade edilmiştir. Yetenek kestirimleri için genel olarak test karakteristik eğrisi ve ortalama kriteri her iki kalibrasyon metodunda benzer sonuçlar vermektedir. Örneklem büyüklüğü ve ortak madde sayısı arttığı zaman hatanın azaldığı bulunmuş ve örneklem büyüklüğünün 250'den daha fazla ve ortak madde sayısının tüm testteki madde sayısının % 25'ten daha fazla alarak dikey ölçekleme yapılması önerilmiştir. Standart sapma kriteri ise EAP da en düşük sonucu verdiği için EAP kullanımı önerilmiştir. Bunun dışında Rasch modelle dikey ölçekleme yapılırken herhangi bir yetenek kestirim yönteminin kullanılabileceği ifade edilmiştir.

Kang ve Petersen (2009) tarafından yapılan çalışmada gerçek veri seti parametrelerine benzer parametrelerle simule edilen yapay veri seti kullanılmıştır. Farklı kalibrasyon yöntemleri, örneklem büyüklüğü, ortak madde sayısı ve yetenek dağılımı koşulları altında oluşturulan dikey ölçekler karşılaştırılmıştır. Elde edilen bulgular doğrultusunda, örneklem sayısı arttıkça standart hatanın azaldığı, ortak madde sayısının toplam madde sayısının %20'si kadarının yeterli olduğu ifade edilmiştir. Kalibrasyon yöntemleri hakkında elde edilen sonuçlar doğrultusunda, karşılaştırılan ayrı, eş zamanlı ve sabit kalibrasyon yöntemlerinin doğru ve yansız sonuçlar ürettiği ifade edilmiştir.

Çetin (2009) araştırmasında gerçek veri seti kullanarak Klasik Test Teorisi (Thurstone) ve Madde Yanıt Kuramı (EAP) yöntemlerini kullanarak sonuçlarını karşılaştırmıştır. 6ncı, 7nci ve 8inci sınıf düzeylerine Matematik ve Türkçe testi uygulanmış ve örneklem büyüklüğünün analiz sonuçlarına etkisi araştırılmıştır. Elde edilen bulguların karşılaştırılabilmesi için değerlendirme ölçütü olarak ortalama, standart sapma ve etki büyüklüğü değerleri kullanılmıştır. Araştırma sonunda, KTT'ye dayalı Thurstone ölçeklemede Matematik ve Türkçe testlerinde sınıf seviyesi ile birlikte standart sapmaların artış gösterdiği ifade edilmiştir. MTK'ya dayalı EAP yetenek kestirimi ile yürütülen ölçekleme sonunda, matematik dersinde, sınıf seviyesi ile ortalamalar tutarlı bir artma ya da azalma göstermez iken, standart sapmalar sınıf seviyelerine bağlı olarak artış göstermiş fakat Türkçe testine ait, MTK ölçeklemesinde gerek ortalama gerekse standart sapmalarda düzenli bir örüntü meydana gelmemiştir. Örneklem sayısı arttıkça standart hatanın azaldığı ve 1000 kişilik örneklemin, hatayı minimize ettiği ifade edilmiştir. Araştırmacı elde edilen bulgular sonucunda, hem KTK hem de MTK uygulamalarında, ortalamaların seyrinden farklı olarak standart sapmaların arttığını ifade belirtmiştir.

Tong ve Kolen (2008) iki yılı aşkın bir süre uygulanan ve dikey ölçekleme yapılabilmesi için bağlama maddeleri içeren büyük ölçekli bir uygulamadan elde edilen verileri analiz etmişlerdir. Çoktan seçmeli maddelere Rasch Model ve yapılandırılmış maddelere ise Kısmi Kredi Modeli uygulanmıştır. Bu iki yöntemin sonuç ölçeği üzerindeki etkisini ve gelişimin yorumlanmasını gözlemeyi amaçlamışlardır. Elde edilen bulgular doğrultusunda, sınıf düzeyi arttıkça ortalamaların arttığı, standart sapmaların sınıf seviyesi ile birlikte artıp azaldığı, etki büyüklüğünün ise sınıf düzeyi arttıkça azalma eğilimi gösterdiği ifade edilmiştir. Bu bulgular, büyüme oranının başarılı ve başarısız öğrenciler arasında sınıf seviyesi arttıkça benzer olduğu şeklinde ifade edilmiştir. Tong ve Kolen (2009) dikey ölçeklerin yürütülmesi üzerine daha derin bir çalışma gerçekleştirmiş, 2008 yılında yürüttükleri bir önceki çalışmalarıyla aynı verileri kullanarak, 3PLM ve genel kısmi kredi modeliyle iki konu alanı (Matematik ve İngilizce) incelenmiştir. İki yıl üst üste 2005 ve 2006 yılında testler uygulanmış, hem iki yılda aynı seviyeye uygulanan testlere yatay eşitleme, hem de aynı yıl farklı seviyelere uygulanan testlerle dikey ölçekleme çalışması yürütülmüştür.

Ito, Sykes ve Yao (2008) gerçek veri seti kullanarak yürüttüğü çalışmalarında ayrı ve eş zamanlı kalibrasyon yöntemlerini karşılaştırmıştır. Anasınıfından 9. sınıf düzeyine kadar olan öğrencilere Matematik ve Okuma testleri uygulanmış ve MTK ile madde parametreleri elde edilmiştir. Ayrı kalibrasyonda SL yöntemi ile test puanları bağlanmış, yetenek kestirim yöntemi olarak da ML yöntemi kullanılmıştır. Elde edilen bulgulara göre, ayrı ve eş zamanlı kalibrasyon benzer sonuçlar üretmişlerdir. Her iki öğrenme alanında da ölçek ortalamaları benzer olduğu ifade edilmiştir.

Briggs, Weeks ve Wiley'in (2008) gerçek veri seti ile yürüttüğü çalışmasında farklı modeller kurarak dikey ölçekleme yapmış ve büyüme örüntüsünün belirlenebilmesi için dikey ölçekleme çalışması yürütmüştür. Çoktan seçmeli testler için 1PLM ve 3PLM, yapılandırılmış maddeler için ise genelleştirilmiş kısmi puan modeli kullanılmış; yetenek kestirim yöntemi olarak EAP ve ML; değerlendirme ölçütü olarak ise ortalamalar ve standart sapma değerleri kullanılmıştır. Elde edilen bulgulara göre, 3PLM kullanıldığında elde edilen ortalamalar arası fark 1PLM ile elde edilen farklara göre daha büyüktür. Standart sapmalar arasında fark ise daha küçüktür. Bu nedenle ortalama ve standart sapmalara bağlı olarak elde edilen etki büyüklüğü değerleri incelendiğinde bu farkın biraz daha azaldığı ifade edilmiştir. EAP ve ML kestirim yöntemleri ve 1 PLM kullanıldığında korelasyon büyük ve negatiftir. Fakat EAP ve ML kestirim yöntemleri ile 3PLM kullanıldığında korelasyon daha düşüktür. Çalışma bulgularına göre araştırmacılar yöntemlerin birbiriyle etkileşiminin sonuçları etkilediğini ifade etmişlerdir.

Kim (2007) çalışmasında farklı içerik alanlarında ve kalibrasyon ile yetenek kestirim yöntemlerinin farklı kombinasyonlarını temel alan farklı dikey ölçekler geliştirmek ve böylece, dikey ölçeklerin özellikleri üzerinde yapılacak farklı seçimlerin yaratacağı etkiyi araştırmayı amaçlamıştır. Bu farklı dikey ölçeklerin değerlendirilmesi için, bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme, sınıf düzeyleri arasındaki çeşitlilik ve düzey dağılımlarının ayrımı incelenmiştir. İncelenen kalibrasyon metotları eşzamanlı kalibrasyon, ayrı kalibrasyon ve ortak maddeler için sabit a, b, ve c madde parametreleri ile basit öncel güncellemeleridir (FSPU). İncelenen yetenek kestirim yöntemleri örüntü/desen skorları ile Maksimum Olabilirlik Kestirimi (ML), örüntü/desen skorları ile Beklenen Önsel Dağılım (EAP), toplam skorlar ile pseudo-ML, toplam skorlar ile pseudo-EAP, ve

Quadrature Dağılımı (QD) olarak belirlenmiştir. Her araştırma sorusu için bu çalışmadan şöyle sonuçlar çıkarılmıştır: Üç kalibrasyon yöntemi karşılaştırıldığında okuma ve fen testleri için eşzamanlı kalibrasyon, FSPU ve ayrı kalibrasyona göre daha az gelişim, düşük düzeylerde daha yavaş azalan bir gelişim, düzeyler arası çeşitlilikte daha az azalma ve yatay aralıklar açısından düşük düzeylerde daha az ayırım göstermiştir. Kelime ve matematik testi için, bir sınıf düzeyinden diğer sınıf düzeyine olan büyümedeki ve düzey dağılımları ayırımındaki farklılıkları önemsizdir. Beş yetenek kestirim yöntemi karşılaştırıldığında, tüm içerik alanları için, düzey içi standart sapmalarda pseudo-ML \geq ML $>$ QD $>$ EAP \geq pseudo-EAP, etki boyutlarında pseudo-EAP \geq EAP $>$ QD $>$ ML \geq pseudo-ML yönelimi elde edilmiştir. Buna rağmen, düzeyler üzeri çeşitlilikteki düşüş derecesi tüm yeterlilik kestirim metotları için benzerdir. Dört içerik alanı karşılaştırıldığında, kelime ve matematik testlerinde, okuma becerisi ve fen testlerine göre gelişimin daha az, ancak bir şekilde daha istikrarlı olduğu gözlenmiş, düzeyler arası çeşitlilikteki düşüşün de fazla olmadığı ortaya konmuştur.

Cao, Li ve Hendrickson (2007) tarafından yürütülen bu çalışmada, ölçek yürütme için üç yöntem üzerinde 3-6 sınıfları arası simülasyon ile türetilmiş veri kullanılarak araştırma yapılmıştır. Bu üç yöntem (1) ayrı kalibrasyon + yatay eşitleme, (2) ayrı kalibrasyon + yatay eşitleme + dikey ölçekleme ve (3) eş zamanlı kalibrasyon + yatay eşitleme olarak belirlenmiştir. Bu üç yöntem üzerine yapılan araştırmaya ek olarak örneklem büyüklüğü de incelenmiştir. İki farklı örneklem büyüklüğünden ilki, toplam 4000 katılımcıya uygulanmak üzere her formun her sınıfı için 500, ikincisi ise toplam 8000 katılımcıya uygulanmak üzere her formun her sınıfı için 1000 olarak seçilmiştir. Buna rağmen, bu çalışmanın sonuçları, iki farklı örneklem büyüklüğünün, üç farklı ölçek yürütme yönteminin performansını etkilemediğini göstermiştir.

Harris (2007) araştırmasında, ölçek uygulamada nelerin ele alınması gerektiği üzerine çeşitli soruların cevabını araştırmıştır. Yeni 3. sınıf formları, orijinal 3. sınıf formlarına eşitlendiği durumlar ile K. sınıftan 8inci sınıfa kadar tüm aralıktaki formları ölçeğin kurulduğu orijinal form kümesine bağlandığı durumları araştırmıştır.

Meng, Kolen ve Lohman (2006), kalibrasyon yöntemlerini (ayrı kalibrasyon, eşzamanlı kalibrasyon ve yarı eşzamanlı kalibrasyon), içerik alanlarını (sözel,

sayısal, sözel olmayan/beden dili (nonverbal)), IRT modellerini (1PL ve 2PL), ve yeterlilik kestirim yöntemlerini (ML, EAP ve QD) farklı koşullar altında karşılaştırmıştır. Bu çalışma, Bilişsel Yetenekler Testleri (CogAT) Form 6'nın (Lohman & Hagen, 2003) sözel, sayısal ve sözel olmayan (nonverbal) testleri 7 seviye kullanılarak ortak madde deseni ile yürütülmüştür. Yarı eşzamanlı kalibrasyonda, dört seviye eşzamanlı olarak kalibre edilmiştir: A, B, C ve D seviyeleri D ölçeği üzerinde eş zamanlı olarak kalibre edilmiş ve aynı şekilde D, E, F ve G seviyeleri de D ölçeği üzerinde eş zamanlı olarak kalibre edilmiştir. Çalışmada, ortak maddelerden elde edilen bilgiler kullanılarak yarı eşzamanlı kalibrasyon metodunun çok boyutluluk problemlerine çözüm olarak sunulabileceği ifade edilmiştir. Bunun yanı sıra çalışmada, üç farklı içerik alanı kullanılmasına rağmen, CogAT setindeki gibi yetenek testlerinin, düzeyler ve test seviyeleri üzerinde başarı testlerine göre daha tek boyutlu olarak ele alındığını ifade etmişlerdir. Sınıflar arası değişkenlik ile ilgili olarak, sonuçlar göstermiştir ki, genelde eş zamanlı ve yarı eş zamanlı kalibrasyon azalan bir desen üretirken, ayrı kalibrasyon tutarlı bir değişkenlik üretmiştir. Ortalama farklarına ve etki büyüklüğü değerleri incelendiğinde, yarı eşzamanlı kalibrasyonun eşzamanlı ve ayrı kalibrasyon sonuçlarının bir birleşimi olduğu gözlenmiştir. Çalışmanın sonuçlarına göre, ML'nin en büyük seviyeler arası standart sapmayı ürettiğini göstermiş, QD ve EAP'ın da sırasıyla ML'yi takip ettiğini ortaya koymuştur. Etki büyüklüğü değerlerine bakıldığında, EAP en büyük kestirimleri üretmiş, QD ve ML de sırasıyla EAP'ı takip etmiştir. Farklı yetenek kestirim yöntemleri kullanmanın, ortalama farklılıkları ve yatay uzaklıklar üzerindeki etkisi önemsizdir. Ortaya konan bu sonuçlar genel olarak Tong'un (2005) elde ettiği sonuçlarla tutarlıdır.

Tong (2005), QD, MLE, EAP, MAP ve pseudo-EAP yetenek kestirim yöntemlerini ve ortak madde deseni ve ölçekleme test deseni kullanılarak farklı koşullar oluşturmuş; bu koşullar altında dikey ölçekleme sonuçlarını karşılaştırmıştır. Tong ITBS'i matematik, dil, okuma ve kelime testlerini kullanarak araştırmasını yürütmüştür. Bu çalışmada ayrıca üç farklı yeterlilik dağılımı (artan, azalan ve düzey artarken sabit kalan standart sapma) ile veri kümeleri simüle edilmiştir. Gerçek ve simüle edilen veri kümelerinden elde edilen sonuçlar, düzey arttıkça tüm ortalamaların arttığını ve öğrencilerin tüm dört alanda da gelişim gösterdiğini ortaya koymuştur. Gerçek veriler kullanılarak yapılan analiz sonuçları sınıf

düzeyleri arası çeşitlilik ölçütü açısından değerlendirildiğinde, ML en büyük düzey içi standart sapmayı üretmiş, ML yöntemini QD, EAP, pseudo-EAP ve MAP izlemiştir. Tüm alanlarda, en büyük etki büyüklüğü MAP kestirimi ile elde edilmiştir. Pseudo-EAP, EAP, QD ve ML de sırasıyla MAP'ı izlemiştir. Simülasyon verileri ile elde edilen analiz sonuçlarında, düzeyler arası çeşitliliği kestirmede $ML > QD > EAP > pseudo-EAP > MAP$ ve etki büyüklüğü kestirmede $ML < QD < EAP < pseudo-EAP < MAP$ eğilimleri ortaya konulmuştur. Tong (2005)'un çalışma bulgularında en küçük standart sapma arzulandığında EAP'ın seçilecek en iyi yöntem olduğu ifade edilmiştir.

Kolen ve Brennan (2004) yaptıkları çalışmalarında ise kalibrasyon yöntemlerini karşılaştırmış ve test karakteristik eğrisi yöntemlerini kullanan ayrı kalibrasyon ile kestirimin en güvenli yöntem olarak görüldüğünü ve eşzamanlı kalibrasyonun, ayrı kalibrasyona bir yardımcı/ek olarak kullanılabileceğini önermişlerdir.

Harris, Hendrickson, Tong, Shin ve Shyu (2004) çalışmalarında, zamanla daha kararlı bir ölçüğe varacak yöntemlerin tanımlanması ve ölçüklerin uzun süreli kullanımların değerlendirmesini sağlayacak kriterlerin geliştirilmesi gerekliliğini önermişlerdir. Dikey ölçeklemedeki yöntem bilimlerin ve sonuçların karşılaştırılması üzerine yapılan bir çalışmada dikey ölçüklerin yıllar üzerinde yürütülmesi konusunda daha yeni araştırmaların yapılmasının gerekliliğine değinmişlerdir.

Hoskens, Lewis ve Patz (2003) yıllar arası temel bir dikey ölçük kullanarak, 4. sınıf düzeyinden 7nci sınıf düzeyine kadar olan düzeylerde dikey bir ölçük yürütme konusu üzerinde durmuşlardır. Her düzey (1) önceden uygulanan, temel dikey ölçük parametrelerine sahip değerlendirme formları ve (2) henüz parametre kestirimleri bulunmayan uygulamanın yeni formunun komşu düzeyleri ortak maddelerinden oluşmaktadır. Yazarlar bu çalışmalarında çoktan seçmeli madde analizleri için Üç Parametrelili Lojistik Model ve cevapları yapılandırılmış maddelerin analizleri için İki Parametrelili Kısmi Kredi Modeli yöntemlerini kullanmışlardır. Ayrı kalibrasyon ve Stocking-Lord yöntemi eşitleme ve bağlama için kullanılmıştır. Hoskens, Lewis ve Patz (2003) yıldan yıla ya da formdan forma dikey ölçük eşitleme tasarımlarının, ölçüğün karşılaştırılabilirliğini üç farklı yönde geliştirmesi gerektiğini belirtmişlerdir. İlk olarak, oluşturulan dikey ölçüğün yürütülme işleminin sınıf içi formlar arası karşılaştırılabilirliği geliştirilmelidir. Birinci yol, gelişimin ölçüsü dikey bir ölçüğe gerek duymaz ve her sınıf içi performans

karşılaştırmaları iki form arasında yapılabilir. İkinci yol olarak, formdan forma dikey ölçek eşitleme tasarımları form içi sınıflar arası karşılaştırılabilirliği desteklemelidir. Yıl içi sınıflar arası ölçek kararlılığı, sınıflar arası ortalama notlar arasında anlamlı karşılaştırmalar yapabilmek için hayati önem taşır. Üçüncü yol olarak ise, formdan forma dikey ölçek eşitleme, formlar ve sınıflar arası karşılaştırılabilirliğe de imkân vermelidir. Bu karşılaştırılabilirlik, her bir öğrenci gelişimini, öğrencilerin notlarındaki değişimi bir uygulamadan diğerine izleyerek ölçmeye olanak vermesi gerektiğini ifade etmişlerdir.

Hanson ve Beguin (2002) ayrı ve eşzamanlı madde parametre kestirimleri üzerine bir simülasyon çalışması yürütmüşlerdir. Elde edilen bulgular, IRT model varsayımlarına uyulduğunda, eşzamanlı kestirimin genel olarak daha az hata içerdiği ve böylece ayrı kalibrasyondan daha isabetli sonuçlar üretildiğini ortaya koymuştur. Buna rağmen, Hanson ve Beguin (2002) bu çalışmanın sonuçlarının, ayrı kalibrasyonun tamamen terkedilip yerine eşzamanlı kalibrasyonun tercih edilmesi için yeterli olmadığını belirtmişlerdir.

Beguin, Hanson ve Glas (2000) ile Beguin ve Hanson'ın (2001) çalışmalarına ek olarak, bu çalışmada, ayrı kalibrasyon ile kestirimin, eşzamanlı kalibrasyon ile kestirime nazaran, tek boyutluluk varsayımının ihlaline karşı daha dirençli (robust) olduğu belirtilmiştir. Ayrı kalibrasyonun bir avantajı da, ortak maddelerin uç değerlerini grafikler üzerinden belirlemek için, madde parametreleri kestirimlerinin incelenmesine olanak tanınması olduğu ifade edilmiştir.

Kim ve Cohen (1998) üç yöntemi IRT altında genel bir ölçü geliştirmek için karşılaştırmışlardır: Stocking-Lord yöntemi ile ayrı kalibrasyon, marjinal maksimum önsel bilgi kestirimi tabanlı eşzamanlı kalibrasyon (MULTILOG kullanılarak) ve marjinal maksimum olabilirlik kestirimi tabanlı eşzamanlı kalibrasyon (BILOG kullanılarak). Tüm simülasyonlar 3PLM'ye fit edilen veriye dayalı olarak gerçekleştirilmiştir. Elde edilen bulgular, az sayıdaki ortak maddeler için, ayrı kalibrasyon üzerinden bağlamanın madde ayırıcılık ve güçlük parametrelerinde daha küçük ortalama hatalar ürettiğini, çok sayıdaki ortak maddeler için ise üç yöntemin de benzer sonuçlar ürettiğini ortaya koymuşlardır.

Skaggs ve Lissitz (1988), Rasch Modeli ile 3PLM eşitleme yöntemlerini tek boyutlu üç parametrelilik lojistik modeli için üretilen veriler kullanılarak karşılaştırmışlardır.

Madde güçlüğü, ayırıcılık gücü ve kestirme derecesi manipüle edilerek Rasch modelin varsayımları sistematik bir biçimde ihlal edilmiştir. Elde edilen bulgular, iki testte de aynı derecede kestirim bulunduğu zamanda Rasch modeli eşitlemesinin iyi çalıştığını göstermektedir. Puanlamanın şans derecesi iki test için eşit olmadığına Rasch modeli eşitlemesi uygun görülmemiştir.

Kolen (1981), iki modeli de inceleyerek Marco'nun (1977) çalışmasını genişletmiştir. Kolen yeni bir versiyonu, Iowa Tests of Educational Development'ın (ITED) bir önceki versiyonunu, Kelime ve Nicel Düşünme başlıklı iki alt kümesi için eşitlemiştir. Bu eşitleme, kestirilmiş gerçek puan eşitleme ve kestirilmiş gözlenen puan eşitleme kullanılarak gerçekleştirilmiştir. Buna ek olarak, Rasch modeli ortak ayırımın (eğim) birbiriyle eşitlenen iki test arasında değişebilmesine izin verecek şekilde düzeltilmiştir. Bulgular, madde içerikleri, zorluk düzeyi ve modeller arasında karmaşık bir etkileşim bulunduğunu göstermiştir.

Marco, Peterson ve Stewart (1979), Rasch modeli ile 3PLM yöntemini, rassal ve farklı örnekleri, iç ve dış ortak madde testleri ve farklı türdeki kriter puanları içeren çeşitli koşullar altında incelemiştir. Araştırmacılar, bu incelemelerde verilen bazı sonuçlara ulaşmışlardır: bir ortak madde testi dışsal olduğunda ve iki test arasında zorluk derecesi bağlamında eşit olduğunda, hem Rasch hem de Üç Parametrelili Lojistik Modelin iyi performans gösterdiği bulgusuna ulaşmışlardır. Dışsal bir ortak madde testi kullanıldığında, Rasch modelinin, 3PLM yönteminden nispeten daha iyi sonuçlara sahip olduğunu ifade etmişlerdir.

Alanyazındaki çalışmalar incelendiğinde, hangi yöntemlerin öğrencilerin başarılarındaki artışı en iyi ve doğru ortaya koyduğu konusunda ortak bir görüş olmadığı ve her araştırmacının dikey ölçek geliştirme süreçlerini kendisi belirlediği görülmektedir. Alanyazın incelendiğinde gerçek veri ile yapılan çalışmaların oldukça az olduğu ve simülasyon verileri ile yapılan çalışmalara ağırlık verildiği görüldüğü için bu çalışmada gerçek veri seti ve yapay veri seti kullanılarak sonuçların karşılaştırılması amaçlanmıştır. Böylece bu çalışmanın alanyazına katkı sağlayacağı düşünülmektedir.

3. YÖNTEM

Bu bölümde araştırmanın türü, evren ve örneklem, veri toplama araçları, verilerin toplanması ve analizine yer verilmiştir.

3.1. Araştırmanın Türü

Bu araştırmanın amacı, dikey ölçekleme işlemi sonucunda sınıf düzeyleri arttıkça oluşan başarı gelişiminin örüntüsünü çıkarmak ve dikey ölçekleme sonuçlarının kullanılan farklı yöntemlere göre nasıl değiştiğini belirlemektir. Araştırmada var olan yöntem ve teknikler gerçek veri ve simülasyon veri üzerinden sınındığı ve en az hatalı yöntemler belirlenerek kuramsal çalışmalara katkı sağlaması amacı taşıdığı için araştırma temel araştırma niteliğinde, değişkenler üzerinde herhangi bir manipülasyon yapılmadan var olan durumu var olduğu şekliyle açıklamaya çalışıldığı için ise betimsel araştırma türündedir.

3.2. Çalışma Grubu

Araştırmada dikey ölçekleme uygulaması için çalışma grubu; 6ncı, 7nci ve 8inci sınıf öğrencilerinden oluşmaktadır. Çalışma, Ankara ili Altındağ, Çankaya, Gölbaşı, Keçiören ve Mamak ilçelerinden ikişer okul olmak üzere 12 farklı okuldaki toplam 1500 öğrenci ile yürütülmüştür. Okullar farklı sosyoekonomik düzeylerdeki ilçelerden seçilerek heterojen bir dağılım sağlanmaya çalışılmıştır. Çalışmada kullanılan yapay veriler ise gerçek verilerden farklı olarak belirlenen madde ve yetenek parametrelerine göre türetilmiştir. Her bir sınıf düzeyinden 500 olmak üzere toplam 1500 öğrenci yanıtı simüle edilmiştir.

3.3. Araştırma Deseni

Bu çalışmada denk olmayan gruplarda ortak madde deseni kullanılmıştır. Bu desen uygulamada yaygın olarak kullanılan desenlerden biri olmakla birlikte, en esnek ve en karmaşık desenlerden biridir (Sinharay & Holland, 2008). Pratiklik açısından tercih edilen bir yöntem olmakla birlikte, diğer desenlere göre de daha az sınırlayıcıdır (Zhu, 1998). Bu desen hem uygulamada, hem de simülasyon verilerinde kullanılmıştır. Belirlenen yetenek ve madde parametreleri doğrultusunda yapay veri seti türetilmiş ve gerçek veride kontrol edilen faktörlerin her biri yapay veriler için de kontrol edilmiştir.

3.4. Verilerin Analizi

Araştırma iki farklı veri seti üzerinden yürütüldüğü için bu bölümde gerçek veri seti ve yapay veri seti betimlenmiş ve gerçek veri seti için MTK varsayımlarının sağlanıp sağlanmadığı kontrol edilmiştir.

3.4.1. Gerçek Veri Seti

Üç sınıf düzeyi için hazırlanan testlerde hem öğrencilerin düzeyine uygun maddeler, hem de eşitlemenin yapılabilmesi için ortak maddelere yer verilmiştir. 7nci sınıf referans sınıf seçildiği için 6ncı ve 7nci sınıfta ve 7nci ve 8inci sınıfta ortak maddeler yer almaktadır. 6ncı ve 7nci sınıftaki ortak maddeler 6ncı ve 7nci, 7nci ve 8inci sınıftaki ortak maddeler ise 7nci ve 8inci sınıf düzeyleri arasındaki bağlantıyı sağlamaktadır. Bu yöntem ile üç sınıf düzeyindeki puanlar zincirleme ölçekleme işlemi ile birbirine bağlanmıştır.

Ortak maddeler kullanılarak öğrencilerin bir sınıf düzeyinden diğer sınıf düzeyine ne kadar gelişim gösterdiği belirlenmektedir. Hambleton vd. (1991), ortak maddelerin sayısının testin tamamının %20'si kadar olmasının uygun olduğunu belirtirken, birçok araştırmada ortak madde sayısındaki artışın testteki ölçmenin standart hatasını azalttığını belirtilmektedir (Boughton, Lorie & Yao, 2005, Kim, Lee, Kim & Kelley, 2009). Bu nedenle bu çalışmada toplam madde sayısının %25'i kadar ortak madde kullanılmıştır. Ortak maddelerin sınıf düzeylerine göre dağılımı aşağıda Tablo 3.1'de verilmiştir.

Tablo 3.1: Ortak Maddelerin Test Düzeylerine Göre Dağılımı

	6ncı Sınıf Düzeyine Uygun Maddeler	6-7 Ortak Maddeler	7nci Sınıf Düzeyine Uygun Maddeler	7-8 Ortak Maddeler	8inci Sınıf Düzeyine Uygun Maddeler
6ncı Sınıf	a	b			
7nci Sınıf		b	c	d	
8inci Sınıf				d	e

Tablo 3.1'deki ortak madde deseni kullanılarak 6ncı sınıf, 7nci sınıf ve 8inci sınıf düzeyinin her birine 40'er maddelik testler hazırlanmış ve 10'er madde ortak madde olarak kullanılmıştır.

3.4.1.1. Veri Toplama Aracının Hazırlanması ve Verilerin Toplanması

Araştırmada 6, 7, ve 8inci sınıflara 40'er maddelik birer başarı testi uygulanmıştır. Testin maddeleri 2008-2012 yılları arasında uygulanan SBS (Seviye Belirleme

Sınavı), OKS (Ortaöğretim Kurumları Seçme ve Yerleştirme Sınavı) ve PYBS (Parasız Yatılılık ve Bursluluk Sınavı) testlerinden seçilerek oluşturulmuştur. Maddelerin ayırt edicilik düzeyleri ve madde güçlükleri kontrol edilmiş, ayırt edicilik düzeyi yüksek ve orta güçlükte olan maddeler seçilmiştir. Fen programı sarmal bir yaklaşım ile hazırlandığı için öğrenme alanları 6ncı sınıftan 8inci sınıfa doğru sarmallık ilkesine dayanarak verilmektedir. Uygulanacak test formları denk olmayan gruplarda ortak madde deseni kullanılarak analiz edileceği için testlerin %25'i ortak maddelerdir. 6ncı sınıfta 30 madde, 7nci sınıfta 20 madde ve 8inci sınıfta 30 madde sınıf seviyesine uygun olarak hazırlanan sadece bu sınıf düzeyindeki öğrencilerin cevapladığı maddelerdir. 6ncı ve 7nci sınıfa ortak olarak hazırlanan 10 madde, 7nci ve 8inci sınıfa ortak olarak hazırlanan 10 madde bulunmaktadır. 6ncı sınıf testinde 6ncı sınıf öğrencilerinin düzeyine uygun olarak SBS-OKS-PYBS maddelerinden oluşan 30 madde ve 10 tane de 7nci sınıf testinde de yer alan ortak madde yer almaktadır. 6ncı sınıf başarı testi Ek-1'de verilmiştir. 7nci sınıf testinde 7nci sınıf öğrencilerinin düzeyine uygun olarak SBS-OKS-PYBS maddelerinden oluşan 20 madde ve 10 tane 6ncı sınıf ile ortak madde 10 tane de 8inci sınıf ile ortak madde yer almaktadır. 7nci sınıf başarı testi Ek-2'de verilmiştir. 8inci sınıf testinde 8inci sınıf öğrencilerinin düzeyine uygun olarak SBS-OKS-PYBS maddelerinden oluşan 30 madde ve 10 tane de 7nci sınıf testinde de yer alan ortak madde yer almaktadır. 8inci sınıf başarı testi Ek-3'te verilmiştir.

3.4.1.2. Gerçek Veriler İçin Madde Tepki Kuramı Varsayımlarının Test Edilmesi

MTK'ya göre eşitleme yaparken dikey ölçeklemenin bazı varsayımlarının test edilmesi gerekmektedir.

3.4.1.2.1 Tek Boyutluluk

Eşitlemede tek boyutlu MTK kullanıldığında, eşitlenecek testlerin tek boyutlu olması varsayımının mutlaka test edilmesi gerekmektedir (Hambleton & Swaminathan, 1985).

MTK bireylerin ölçmeye konu olan özelliklerinin hazırlanan test ile ölçülebildiğini varsaymaktadır. Bireylerin k tane örtük özelliği, k boyutlu bir örtük özellik uzayı ile tanımlanır. Her bir özelliği ölçen k maddelik bir test ile bireylerin k boyuttaki özelliği

ölçülebilir. Bir testteki tüm maddeler aynı özelliği ölçüyorsa bireyin testten aldığı puan, onun bu boyuttaki yerini gösterir. MTK'nın bütün modelleri testle ölçülen yeteneğin tek boyutlu olduğunu varsaymaktadır (Hambleton & Swaminathan,1985)

MTK'nın tek boyutluluk varsayımını test etmek amacıyla, öğrencilere uygulanan fen başarı testlerinin her üç sınıf düzeyine DFA uygulanmıştır. Belirlenen model 0.05 anlamlılık düzeyinde test edilmiştir. Model veri uyumunu değerlendirmek amacıyla pek çok uyum iyiliği indeksleri kullanılmaktadır. Bunlar içinde görece en sık kullanılanlardan, Ki-kare Uyum Testi (χ^2 / sd), Yaklaşık Hataların Ortalama Karekökü (RMSEA), İyilik Uyum İndeksi (GFI), Düzeltilmiş İyilik Uyum İndeksi (AGFI), Karşılaştırmalı Uyum İndeksi (CFI), Normlaştırılmış Uyum İndeksi (NFI) kontrol edilmiş ve bu değerler Tablo 3.2'de gösterilmiştir.

Tablo 3.2: Fen Bilgisi Başarı Testi Doğrulayıcı Faktör Analizi Sonucu Elde Edilen Uyum İyiliği İndeksleri

Uyum Ölçüsü	Mükemmel Uyum Değerleri	Kabul Edilebilir Uyum Değerleri	Model Değerleri		
			6nci sınıf	7nci sınıf	8inci sınıf
χ^2 / sd	$0 < \chi^2 / sd \leq 2$	$2 < \chi^2 / sd \leq 5$	1.76	2.35	1.98
RMSEA	$0 < RMSEA < 0.05$	$0.05 \leq RMSEA \leq 0.10$	0.05	0.08	0.05
GFI	$0.95 \leq GFI \leq 1$	$0.90 \leq GFI \leq 0.95$	0.93	0.93	0.94
AGFI	$0.90 \leq AGFI \leq 1$	$0.85 \leq AGFI \leq 0.90$	0.92	0.90	0.95
CFI	$0.97 \leq CFI \leq 1$	$0.95 \leq CFI \leq 0.97$	0.97	0.97	0.98
NFI	$0.95 \leq NFI \leq 1$	$0.90 \leq NFI \leq 0.95$	0.97	0.94	0.95

(Kaynak: Schermelleh-Engel, Moosbrugger & Müller, 2003)

Tablo 3.2'deki model için elde edilen uyum indeksleri değerleri incelendiğinde ki-kare uyum testi değerinin 6ncı sınıf düzeyinde 1.76, 7nci sınıf düzeyinde 2.35, 8inci sınıf düzeyinde 1.98 olduğu görülmektedir. Ki-kare uyum iyiliği indeksleri sıfıra yaklaştıkça model mükemmelliğe yaklaşmaktadır. 6ncı ve 8inci sınıfa ait ki-kare değerleri modelin mükemmel uyuma sahip olduğunu, 7nci sınıfın ise kabul edilebilir uyuma sahip olduğu görülmektedir. RMSEA değerinin 0 ile 0.05 arasında olması modelin mükemmel uyuma, 0.05 ile 0.10 arasında olması modelin kabul edilebilir uyuma sahip olduğunu göstermektedir. RMSEA değeri 6ncı ve 8inci sınıfta 0.05, 7nci sınıfta ise 0.08 olarak bulunmuştur. Bu değerler modelin kabul edilebilir uyuma sahip olduğunu ifade etmektedir. Diğer uyum indekslerinin (GFI, AGFI, CFI, NFI) 1'e yaklaşması daha iyi bir uyumun göstergesidir. Bu indeksler de

tek tek incelendiğinde, her üç sınıf düzeyi için modelin uyumunun yüksek olduğu görülmektedir. Bütün indekslerin sonucunda verinin tek boyutluluk sayılığını karşıladığı ve fen başarı testinin tüm sınıflarda tek bir örtük özelliği ölçtüğü gözlenmektedir. Her bir sınıf seviyesine ait yapısal eşitlik modeli grafiği Ek-4'te verilmiştir.

2.4.1.2.3 Yerel Bağımsızlık

Yerel bağımsızlık maddelerin birbirinden bağımsız olarak cevaplandırılması, bir maddenin doğru cevaplanma olasılığının diğer maddelerden etkilenmemesi demektir. Yetenek düzeyi sabit tutulduğunda, maddeler arası korelasyonun sıfıra yaklaşması beklenmektedir.

Yerel bağımsızlık varsayımını karşılamak üzere, maddelere verilen cevaplar için sadece bir yeteneğe gereksinim duyulursa, bu maddelerin tek boyutlu olarak ifade edilir (Nandakumar, 1994). Tek boyutluluk varsayımı karşılanması yerel bağımsızlık varsayımına bir kanıt olarak kullanılmaktadır (Hambleton, 1991; Lord & Novick, 1968). Bu çalışmada da tek boyutluluk varsayımı karşılandığı için yerel bağımsızlık varsayımının da karşılandığı kabul edilmiştir.

3.4.1.3 Model Veri Uyumunun Kontrol Edilmesi

Madde Tepki Kuramına göre varsayımlar test edildikten sonra, hangi modelin veri setine daha iyi uyum sağladığını belirleyebilmek için model-veri uyumu kontrol edilmiştir. Her sınıf düzeyi için ayrı kalibrasyon yapılarak uyum istatistikleri hesaplanmıştır. Model veri uyumunu değerlendirmek amacıyla ki-kare, ki-kare/sd değerleri kontrol edilmiştir. 1 PLM ve 2 PLM için model veri uyumu sağlanmış, 3 PLM için model veri uyumu sağlanamamıştır. Her üç sınıf seviyesi için χ^2/sd incelenmiş ve verilerin iki parametrelili lojistik modele uyum sağladığı görülmektedir. İki parametrelili lojistik modelin model veri uyumuna ilişkin sonuçlar aşağıda Tablo 3.3'te verilmiştir.

Tablo 3.3: Ayrı Kalibrasyona Ait Sınıflara Göre Puanların İki Parametrelili Lojistik Model Ki-Kare Uyum İstatistikleri

	χ^2	χ^2/sd	p
6ncı sınıf	512.5	1.82	0.000
7nci sınıf	774.7	2.83	0.000
8inci sınıf	443.5	1.51	0.000

Her üç sınıf düzeyindeki test model veri uyumunun sağlandığı iki parametrelili lojistik modele göre ölçeklenmiş ve hesaplanan a ve b parametre değerlerinin betimsel istatistikleri aşağıda Tablo 3.4'te gösterilmiştir.

Tablo 3.4: Madde Tepki Kuramına Göre İki Parametrelili Lojistik Model İle Ölçekleme Sonucu Elde Edilen Madde Parametrelerinin Betimsel İstatistikleri

	6ncı sınıf		7nci sınıf		8inci sınıf	
	a	b	a	b	a	b
Ortalama	0.481	0.355	0.639	0.487	0.754	0.778
Std. Sapma	0.254	1.584	0.468	0.804	0.183	0.718
Min.	0.137	-2.353	0.169	-0.916	0.142	-2.345
Max.	1.131	4.070	2.266	2.833	0.823	2.285

Tablo 3.4'teki ayırt edicilik gücünü incelendiğinde, 6ncı sınıf a parametresinin 0.48, 7nci sınıf a parametresinin 0.63, 8inci sınıfın a parametresinin 0.77 olduğu görülmektedir. Tablo 3.4'teki maddelerin ortalama güçlükleri incelendiğinde her üç sınıf düzeyinde b parametreleri pozitif olduğu görülmektedir. Altıncı sınıf düzeyinden sekizinci sınıf düzeyine doğru b parametrelerinin arttığı görülmektedir. Her bir maddenin model veri uyumu incelenmiş ve model veri uyumunun sağlanmadığı madde olmadığı gözlenmiştir.

3.4.2 Simülasyon ile Yapay Verilerin Türetilmesi

Araştırmada, gerçek veriler ile yapay verilerin sonuçlarını karşılaştırabilmek için gerçek verilere benzer olarak iki parametrelili veriler türetilmiş, şans başarısı sıfır olarak belirlenmiştir. Yapay veri türetilmesi için R programı kullanılmıştır. İstatistiksel veri işleme amacıyla oluşturulan R programlama dilinde geliştirilen kodlarla yapay veri türetilmiştir. Veride olması istenen her bir özellik R'da kod yazılarak belirtilmiştir. Bu çalışmanın temel amacı, kalibrasyon yöntemleri ve yetenek kestirim yöntemlerinin çeşitli kombinasyonlarının dikey ölçekleme sonucuna olan etkisini araştırmaktır. Çalışmanın bu bölümünde de simülasyon çalışması ile araştırmacılar tarafından belirlenen yetenek ve madde parametrelerine göre veri türetilmiş ve gerçek veride karşılaştırılan tüm yöntemler simülasyon ile türetilen yapay veriler için de test edilmiştir. Böylece farklı yetenek dağılımlarında ve farklı madde parametreleriyle nasıl sonuçlar elde edileceği sınıanmıştır.

3.4.2.1. Yetenek Parametrelerinin Türetilmesi

Yetenek dağılımları her bir grup için standart normal dağılım kullanılarak türetilmiştir. θ yetenek düzeyleri altıncı sınıf düzeyi için $[-1,0]$ aralığında ortalaması -0.5 standart sapması 0.5 olan standart normal dağılımla türetilmiştir. Yedinci sınıf düzeyine ait θ yetenek parametreleri $[0,1]$ aralığında ortalaması 0.5 standart sapması 0.5 olan standart normal dağılımla türetilmiştir. Sekizinci sınıf düzeyine ait θ yetenek parametreleri ise $[1,2]$ aralığında ortalaması 1.5 standart sapması 0.5 olan standart normal dağılımla türetilmiştir. Sınıf seviyesi arttıkça cevaplayıcıların yetenek dağılımları artırılmış, standart sapmaları ise sabit tutulmuştur. Bu çalışmada pek çok simülasyon çalışmasında (Kim & Kolen, 2006; Kim & Lee, 2006; Keller III, 2007; Cao, 2008; Zhao, 2008) olduğu gibi sadece θ yetenek düzeyi ortalamaları arasındaki farklılıklar değerlendirilmiş ve standart sapmaları sabit tutulmuştur.

3.4.2.2. Madde Parametrelerinin Türetilmesi

Araştırmada teorik olarak madde parametreleri belirlenmiş ve tüm düzeyler için veriler normal dağılım kullanılarak türetilmiştir. Her bir sınıf düzeyi için üç veri seti oluşturulmuştur. Yapay veriler ile gerçek veri setinin paralel olması için her bir formun %25'i ortak madde olarak belirlenmiştir. Verilerin gerçeği yansıtabilmesi için sınıf düzeylerine göre a ve b parametreleri belirlenmiştir.

6ncı sınıfta 30 madde, 7nci sınıfta 20 madde ve 8inci sınıfta 30 maddenin yetenek parametreleri ve madde parametreleri seviyelerine uygun olarak belirlenmiştir. 6ncı ve 7nci sınıfa ortak olan 10 madde 6ncı ve 7nci sınıf için farklı parametrelerle, 7nci ve 8inci sınıfa ortak olan 10 madde de 7nci ve 8inci sınıf için farklı parametrelerle türetilmiştir.

Altıncı sınıf düzeyindeki ilk 30 maddenin a parametreleri ortalaması 1.0 ve standart sapması 0.5 olan standart normal dağılımla türetilmiştir. Altıncı sınıf b parametreleri ortalaması 0 ve standart sapması 3.0 olan standart normal dağılımla türetilmiştir. Altıncı sınıf, yedinci sınıf ve sekizinci sınıf c parametreleri, 2 parametrelili lojistik model kullanılacağı için sıfır olarak belirlenmiştir. 7nci sınıf ile ortak olan son 10 maddenin ise a parametreleri ortalaması 1 , standart sapması 0.5 ; b parametreleri ortalaması -0.5 , standart sapması 0.5 olarak belirlenmiştir.

Yedinci sınıf düzeyindeki maddelerden ilk 10 tanesi altıncı sınıf ile ortak, 20 tanesi yedinci sınıf düzeyine uygun, 10 tanesi de sekizinci sınıf ile ortak maddelerdir. Yedinci sınıfın ait ilk 10 madde için a parametrelerinin ortalaması 1.5 standart sapması 0.5; b parametrelerinin ortalaması 0, standart sapması 1 olarak belirlenmiştir. İkinci 20 madde için a parametrelerinin ortalaması 1.5, standart sapması 0.5; b parametrelerinin ortalaması 0.5, standart sapması 2.5 olarak belirlenmiştir. Son 10 madde için a parametrelerinin ortalaması 1.5, standart sapması 0.5; b parametrelerinin ortalaması 0.5, standart sapması 0.5 olarak belirlenmiştir.

Sekizinci sınıfın yedinci sınıf ile ortak olan ilk on maddesi için a parametreleri ortalaması 1.75, standart sapması 0.25; b parametreleri ortalaması 1.5, standart sapması 0.5 olarak belirlenmiştir. Kalan son 30 maddenin ise a parametrelerinin ortalaması 1.75, standart sapması 0.25; b parametrelerinin ortalaması 1, standart sapması 2 olarak belirlenmiştir.

3.4.2.3. Madde Cevaplarının Türetilmesi

Her bir sınıf düzeyi için yetenek ve madde parametreleri belirlendikten sonra, 500'er kişilik madde cevapları türetilmesi için R programında yazılan kodlar kullanılmıştır. Random olarak türetilen madde cevapları araştırmanın yapay verilerini oluşturmaktadır. Madde cevapları hem gerçek veri ile uyumlu olması açısından hem de c parametresini sifıra sabitleyerek daha iyi güçlük ve ayırt edicilik parametreleri elde edebilmek için araştırma verileri 2 parametrelili lojistik modele göre türetilmiştir.

Araştırmada belirlenen her bir koşul için 100 tekrar yapılmıştır.

4. BULGULAR VE TARTIŞMA

Bu bölümde, araştırma sonunda elde edilen bulgulara ve bulgulara dayalı yorumlara yer verilmiştir.

Araştırma bulguları, sınıf seviyelerine, kullanılan kalibrasyon yöntemlerine ve yetenek kestirim yöntemlerine göre elde edilen sonuçlar ortalama, standart sapma ve etki değeri ölçütlerine göre incelenmiştir.

Araştırmanın amacı doğrultusunda kurulan problem cümlesini araştırılırken öncelikle uygulamadan elde edilen veriler için MTK varsayımları kontrol edilmiş ve bu varsayımlara ilişkin bulgular, yöntem bölümünde ayrıntılı olarak verilmiştir. Yapılan analiz sonucunda MTK varsayımlarının sağlandığı görülmüştür. Hangi modelin veriye daha uyumlu olduğunu belirlemek için her sınıf düzeyine ait testlerin madde ve yetenek parametreleri 1 parametrelili, 2 parametrelili ve 3 parametrelili modele göre kestirilerek, her bir modelin uyum iyiliği katsayıları kontrol edilmiştir. Model veri uyumu 2 parametrelili lojistik modelde sağlanmıştır.

Araştırmanın alt problemlerine ilişkin bulgular aşağıda verilmiştir.

4.1 Birinci Alt Probleme Ait Bulgular ve Yorumlar

Fen Bilgisi Başarı Testinde eş zamanlı kalibrasyon yöntemi ile elde edilen maksimum olabilirlik, beklenen önsel bilgi ve maksimum önsel bilgi yetenek kestirimlerinde; bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme (ortalama ve ortalama farkları), sınıf düzeyleri arasındaki çeşitlilik (standart sapma) ve düzey dağılımlarının ayrımı (etki büyüklüğü) değerleri sınıf düzeyi ile birlikte nasıl değişmektedir?

Bu alt probleme cevap aramak için tüm sınıf düzeylerine ait veriler tek bir dosyada birleştirilmiş ve tüm veriler BILOG-MG 3 programı kullanılarak aynı anda kalibre edilmiştir. Eş zamanlı kalibrasyon yöntemi ile her bir seviyeye ait madde ve yetenek parametreleri kestirilmiştir. ML, EAP ve MAP yetenek kestirim yöntemlerine göre elde edilen θ yetenek düzeyi ortalamaları, ortalamalar arası farklar, standart sapmalar ve etki büyüklüğü değerleri hesaplanmış ve bu değerler aşağıda Tablo 4.1'de gösterilmiştir.

Tablo 4.1: Fen Bilgisi Başarı Testi Eş Zamanı Kalibrasyon Yöntemi İle Elde Edilen ML, EAP ve MAP Yetenek Kestirimine Ait Sonuçlar

	Sınıf Düzeyi	ML	EAP	MAP
Ortalamalar	6	-0.084	-0.379	-0.318
	7	0.000	0.000	0.000
	8	0.558	0.633	0.585
Ortalama farkları	7-6	0.084	0.379	0.318
	8-7	0.558	0.633	0.585
Standart sapmaları	6	0.643	0.930	0.785
	7	1.000	1.000	1.000
	8	0.415	0.336	0.346
Etki büyüklükleri	7-6	0.0709	0.2777	0.2505
	8-7	0.5154	0.6000	0.5530

Tablo 4.1’de her bir sınıf düzeyine ilişkin ML, EAP ve MAP yetenek kestirimi sonucunda elde edilen ortalama, ortalamalar farkı ve standart sapma değerleri görülmektedir.

Eş zamanlı kalibrasyon ile MTK parametreleri kestirilirken her üç yetenek kestiriminde de referans sınıf olarak 7nci sınıf seçildiğinden bu sınıfın θ yetenek düzeyi ortalaması 0 (sıfır), standart sapması ise 1 (bir)’dir.

ML yetenek kestirimi sonucuna göre; θ yetenek düzeyi ortalamaları 6ncı sınıf düzeyinden 7nci sınıf düzeyine ve 7nci sınıf düzeyinden 8inci sınıf düzeyine artış göstermektedir. Büyüme örüntüsünü belirlemek için ardışık sınıf düzeyleri arasındaki ortalama fark değerleri incelendiğinde, 7nci sınıf ile 6ncı sınıf arasındaki ortalama farkı 0.084, 8inci sınıf ile 7nci sınıf arasındaki ortalama farkı değerinin ise 0.558 olduğu görülmektedir.

ML yetenek kestirim yöntemine göre hesaplanan standart sapma değerleri incelendiğinde en yüksek standart sapma değerinin referans sınıf olan 7nci sınıfa ait olduğu, en düşük standart sapma değerinin ise 8inci sınıfa (0.415) ait olduğu görülmektedir. Bu bulgu 8inci sınıf düzeyindeki grubun diğer sınıf düzeylerindeki gruplardan ölçülen özellikler bakımından daha homojen olduğu şeklinde yorumlanabilir.

ML kestirimine göre elde edilen ortalama ve standart sapma deęerleri kullanılarak elde edilen etki byklklerinin mutlak deęerleri incelendięinde; 6ncı sınıf ile 7nci sınıf arasındaki etki byklęu deęeri 0.071, 7nci sınıf ile 8inci sınıf arasındaki etki byklęu deęeri 0.515 olarak hesaplanmıřtır. 6ncı sınıf ile 7nci sınıf arasındaki etki deęeri kk etki, 7nci sınıf ile 8inci sınıf arasındaki etki deęeri ise orta etki olarak yorumlanabilir.

EAP yetenek kestirimi sonucuna gre; θ yetenek dzeyi ortalamaları 6ncı sınıf dzeyinden 7nci sınıf dzeyine ve 7nci sınıf dzeyinden 8inci sınıf dzeyine artıř gsterdięi grlmektedir. Byme rntsn belirlemek iin ardıřık sınıf dzeyleri arasındaki ortalama fark deęerleri incelendięinde, 7nci sınıf ile 6ncı sınıf arasındaki ortalama farkı 0.379, 8inci sınıf ile 7nci sınıf arasındaki ortalama farkı deęerinin ise 0.663 olduęu grlmektedir.

EAP yetenek kestirim yntemine gre hesaplanan standart sapma deęerleri incelendięinde en yksek standart sapma deęerinin referans sınıf olan 7nci sınıfa ait olduęu, en dřk standart sapma deęerinin ise 8inci sınıfa (0.336) ait olduęu grlmektedir. Bu bulgu 8inci sınıf dzeyindeki grubun dięer sınıf dzeylerindeki gruplardan daha homojen bir grup olduęu řeklinde yorumlanabilir.

EAP kestirimine gre elde edilen ortalama ve standart sapma deęerleri kullanılarak elde edilen etki byklklerinin mutlak deęerleri incelendięinde; 6ncı sınıf ile 7nci sınıf arasındaki etki byklęu deęeri 0.278, 7nci sınıf ile 8inci sınıf arasındaki etki byklęu deęeri 0.600 olarak hesaplanmıřtır. 6ncı sınıf ile 7nci sınıf arasındaki etki deęeri kk etki, 7nci sınıf ile 8inci sınıf arasındaki etki deęeri ise orta etki olarak yorumlanabilir.

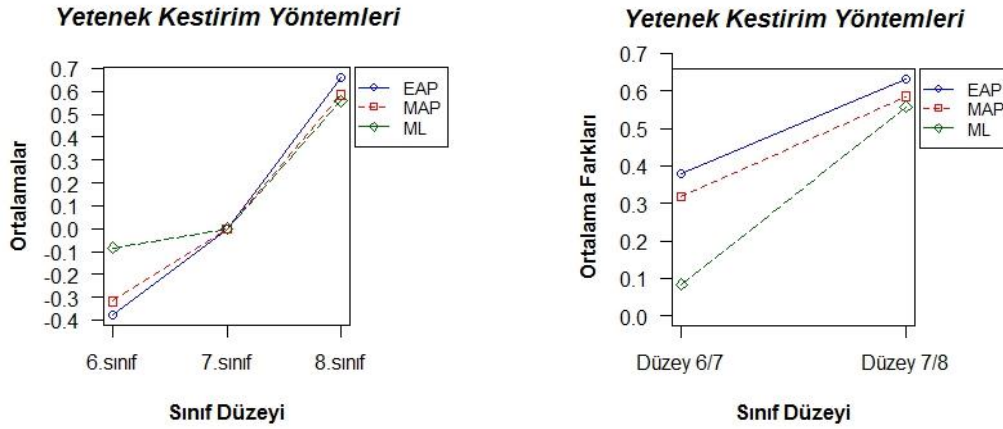
MAP yetenek kestirimi sonucuna gre; θ yetenek dzeyi ortalamaları 6ncı sınıf dzeyinden 7nci sınıf dzeyine ve 7nci sınıf dzeyinden 8inci sınıf dzeyine artıř gsterdięi grlmektedir. Byme rntsn belirlemek iin ardıřık sınıf dzeyleri arasındaki ortalama fark deęerleri incelendięinde, 7nci sınıf ile 6ncı sınıf arasındaki ortalama farkı 0.318, 8inci sınıf ile 7nci sınıf arasındaki ortalama farkı deęerinin ise 0.585 olduęu grlmektedir.

MAP yetenek kestirim yntemine gre hesaplanan standart sapma deęerleri incelendięinde en yksek standart sapma deęerinin referans sınıf olan 7nci sınıfa ait olduęu, en dřk standart sapma deęerinin ise 8inci sınıfa (0.346) ait olduęu

görülmektedir. Bu bulgu 8inci sınıf düzeyindeki grubun diğer sınıf düzeylerindeki gruplardan daha homojen bir grup olduğu şeklinde yorumlanabilir.

MAP kestirimine göre elde edilen ortalama ve standart sapma değerleri kullanılarak elde edilen etki büyüklüklerinin mutlak değerleri incelendiğinde; 6ncı sınıf ile 7nci sınıf arasındaki etki büyüklüğü değeri 0.251, 7nci sınıf ile 8inci sınıf arasındaki etki büyüklüğü değeri 0.553 olarak hesaplanmıştır. 6ncı sınıf ile 7nci sınıf arasındaki etki değeri küçük etki, 7nci sınıf ile 8inci sınıf arasındaki etki değeri ise orta etki olarak yorumlanabilir.

Bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme kriterini değerlendirmek için hesaplanan ortalama ve ortalama farkların yetenek kestirim yöntemlerine göre karşılaştırılması aşağıda Grafik 4.1’de verilmiştir.



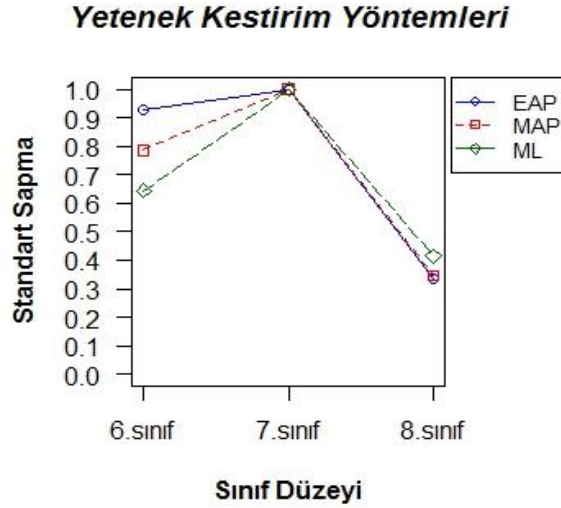
Grafik 4.1: Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen Ortalama ve Ortalama Farkları

Grafik 4.1’deki ortalamalara ait grafik incelendiğinde, 6ncı sınıf düzeyinde en düşük ortalama θ değeri MAP yetenek kestirim yöntemi ile hesaplanırken, en yüksek ortalama θ değerinin ML kestirim yöntemi ile hesaplandığı görülmektedir. 8inci sınıf düzeyinde en düşük ortalama θ değeri ML ile hesaplanırken, en yüksek ortalama θ değeri EAP ile hesaplanmıştır.

Grafik 4.1’deki ortalama farklarına ait grafik incelendiğinde 6ncı sınıf ile 7nci sınıf arasındaki θ ortalama farkları $ML < MAP < EAP$ şeklinde sıralanırken, 7nci sınıf ile 8inci sınıf arasındaki θ ortalama farklarının da yine aynı şekilde $ML < MAP < EAP$

şeklinde sıralandığı görülmektedir. Bu değerler incelendiğinde 6ncı sınıftan 8inci sınıfa doğru artan bir başarı örüntüsü olduğu görülmektedir.

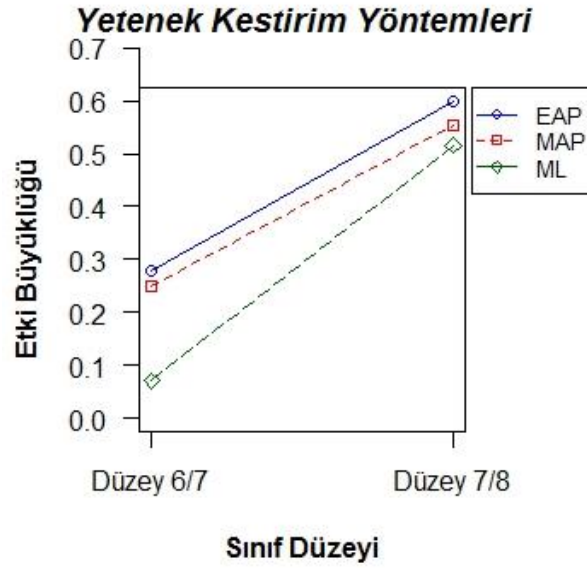
Sınıf düzeyleri arasındaki çeşitlilik kriterini değerlendirmek için; 3 yetenek kestirim yöntemi ile elde edilen standart sapma değerlerinin karşılaştırılması aşağıda Grafik 4.2'de verilmiştir.



Grafik 4.2: Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen Standart Sapma Değerleri

Grafik 4.2'deki standart sapmalara ait grafik incelendiğinde; 6ncı sınıf için en yüksek standart sapma değeri EAP ile hesaplanırken, en düşük standart sapma değeri ML yöntemi ile hesaplanmıştır. 6ncı sınıf düzeyi için standart sapma değerlerini sıralamak gerekirse, standart sapma değerleri $ML < MAP < EAP$ şeklinde sıralandığı görülmektedir. 8inci sınıf düzeyinde en yüksek standart sapma ML ile hesaplanırken, en düşük standart sapma değeri EAP ile hesaplanmıştır. 8inci sınıf düzeyi için standart sapma değerlerini sıralamak gerekirse, standart sapma değerleri $EAP < MAP < ML$ şeklinde sıralandığı görülmektedir.

Düzye dağılımları arasındaki ayırım kriterini değerlendirmek için; yetenek kestirim yöntemleri ile elde edilen ortalamalar ve standart sapmalar kullanılarak etki büyüklüğü kriterleri hesaplanmıştır. Etki büyüklüğü değerlerinin yetenek kestirim yöntemlerine göre karşılaştırılması aşağıda Grafik 4.3'te verilmiştir.



Grafik 4.3 Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen Etki Büyüklüğü Değerleri

Etki büyüklüğü, artarda gelen sınıflar arasındaki ortalama farklarını sınıf düzeyleri içindeki varyansların karekökünü kullanarak standardize etmektedir. Etki büyüklüğü 0-0.20 arasında değer alıyorsa zayıf etki, 0.21-0.50 arasındaki değerler için küçük etki, 0.51-1 orta etki ve 1'den büyük değerler için ise güçlü etki olarak tanımlanmaktadır (Cohen, Manion ve Morrison, 2007). Etki büyüklüğünün zayıf olması, art arda gelen sınıf düzeyleri arasındaki farkın az olduğu, güçlü olması ise art arda gelen sınıf düzeyleri arasındaki farkın fazla olduğu şeklinde yorumlanmaktadır.

Grafik 4.3 incelendiğinde her üç yetenek kestirim yönteminde de benzer sonuçlar alan etki büyüklüğü indeksinin, hem 6ncı ve 7nci sınıflar arasında hem de 7nci ve 8inci sınıflar arasında pozitif değer aldığı görülmektedir. Tüm kestirim yöntemlerinde, 6ncı ve 7nci sınıflar arasındaki etki büyüklüğü değerleri, 7nci ve 8inci sınıf değerlerine göre daha küçüktür. Grafik 4.3'teki grafik incelendiğinde, etki büyüklüğü indeksinin $ML < MAP < EAP$ olarak sıralandığı görülmektedir. EAP ile kestirilen etki büyüklüklerinin diğer yetenek kestirimlerine göre daha büyük olduğu bulunmuştur. Bu bulgu Tong ve Kolen (2010) ve Kim (2007)'in eş zamanlı kalibrasyon analizi yaptığı çalışmasındaki bulgular ile paralellik göstermektedir. Aynı zamanda Meng vd. (2006) ve Tong (2005) da çalışmasında bu bulgu ile

paralel olarak en küçük etki büyüklüğü değerinin ML kestirimi ile elde edildiğini ifade etmişlerdir.

4.2 İkinci Alt Probleme Ait Bulgular ve Yorumlar

Fen Bilgisi Başarı Testinde ayrı kalibrasyon yöntemi ile elde edilen maksimum olabilirlik, beklenen önsel bilgi ve maksimum önsel bilgi yetenek kestirimlerinde; ortalama, standart sapma ve etki büyüklüğü değerleri sınıf düzeyi ile birlikte nasıl değişmektedir?

Her bir sınıf düzeyi için veriler 2PLM ile ayrı ayrı kalibre edilmiş ve parametreler BILOG-MG 3 programı ile hesaplanmıştır. Her sınıf düzeyinin parametre tahminlerinin temel olarak kabul edilen 7nci sınıf ölçeğine yerleştirilmesi için C programlama dilinde yazılan ve MTK ölçekleme sabitlerini hesaplayan bir program olan ST (Hanson, Zeng & Chien, 2004) programı kullanılmış karakteristik eğri yöntemlerinden Stocking Lord yöntemi ile eğim ve keşişim değerleri hesaplanmıştır. Alanyazında dört dönüşüm yönteminin karşılaştırıldığı ve kullanıldığı çalışmalarda madde parametrelerinin kullanıldığı moment yöntemleri yerine, madde ve test karakteristik eğrilerinin kullanıldığı Haebara ve Stocking Lord yöntemlerinin kullanılması gerektiği tavsiye edilmiştir (Kolen & Brennan, 2004; Hanson & Béguin, 2002; Kim & Kolen, 2006). Araştırmalarda SL yönteminin diğer yöntemlere göre daha az hata ürettiği ifade edildiği için (Hanson & Béguin, 2002; Karkee & Wright, 2004; Kim, 2007) bu çalışmada da SL dönüşüm yöntemi ile elde edilen değerler kullanılmıştır.

Stocking Lord yöntemi için quadrature noktaları kullanılması gerekmektedir. Bu nedenle analizler icl_win programı kullanılarak yapılmıştır. Analiz sonucunda elde edilen quadrature noktaları kodlara eklenerek SL dönüşümü yapılmıştır.

Test karakteristik eğrileri kullanılarak SL yöntemi uygulanmış ve elde edilen eğim ve keşişim değerleri aşağıda Tablo 4.2'de gösterilmiştir.

Tablo 4.2: Stocking Lord Dönüşümü İçin Elde Edilen A ve B Sabitleri

Sınıf Düzeyi	A (Slope)	B (Intercept)
6-7	1.121	0.767
7-8	1.574	-0.962

Tablo 4.2'deki SL dönüşümünden elde edilen A ve B sabitleri kullanılarak dönüşümler yapılır. 7. sınıf referans sınıf olduğu için 6. sınıfı 7. sınıf ölçeğine dönüştürürken, yetenek kestirimleri $\theta_{\text{yeni}} = \theta_{\text{eski}} \times 1.121 + (0.767)$ eşitliğiyle yapılırken, 8. sınıfı 7. sınıfa dönüştürürken yetenek kestirimi $\theta_{\text{yeni}} = \theta_{\text{eski}} \times 1.574 + (-0.962)$ eşitliği yardımı ile hesaplanır. 8. sınıfı 6. sınıfa dönüştürmek için ise iki dönüşüme ihtiyaç vardır. 8. sınıfı dönüştürmek için, $\theta_{\text{yeni}} = (\theta_{\text{eski}} \times 1.121 + (0.767)) \times 1.574 + (-0.962)$ eşitliği kullanılmıştır. Kesişim değerleri 6. ve 7. sınıf arasında pozitif iken 7. ve 8. sınıfları arasında negatiftir.

BILOG-MG 3 programını çalıştıracak kodlar yazılarak her bir sınıf düzeyi için ayrı kalibrasyon yöntemi kullanılarak ayrı ayrı kestirimler yapılmıştır. ML, EAP ve MAP yetenek kestirim yöntemlerine göre elde edilen θ yetenek düzeyi ortalamaları, ortalamalar arası farklar, standart sapmalar ve etki büyüklüğü değerleri hesaplanmış ve aşağıda Tablo 4.3'te gösterilmiştir.

Tablo 4.3: Fen Bilgisi Başarı Testi Ayrı Kalibrasyon Yöntemi İle Elde Edilen ML, EAP ve MAP Yetenek Kestirimine Ait Sonuçlar

	Sınıf Düzeyi	ML	EAP	MAP
Ortalamalar	6	-0.317	-0.058	-0.147
	7	0.007	0.002	-0.021
	8	-0.009	-0.005	-0.025
Ortalama farkları	6-7	0.320	0.060	0.126
	7-8	-0.016	-0.007	-0.004
Standart sapmaları	6	0.608	0.822	0.354
	7	1.091	0.874	0.814
	8	1.091	0.586	0.575
Etki büyüklükleri	6-7	0.2558	0.0498	0.1420
	7-8	-0.0104	-0.0030	-0.0040

Tablo 4.3'te her bir sınıf düzeyine ilişkin ML, EAP ve MAP yetenek kestirimi sonucunda elde edilen ortalama, ortalamalar farkı ve standart sapma değerleri görülmektedir.

ML yetenek kestirimi sonucuna göre; θ yetenek düzeyi ortalamaları 6ncı sınıf düzeyinden 7nci sınıf düzeyine yükselirken, 7nci sınıf düzeyinden 8inci sınıf düzeyine tekrar düştüğü görülmektedir. Büyüme örüntüsünü belirlemek için

hesaplanan ardışık sınıf düzeyleri arasındaki ortalama farkı değerleri de sınıf düzeyleri arasındaki bu farklılaşmayı bize göstermektedir. 6ncı sınıf ile 7nci sınıf arasındaki ortalama farkı 0.320, 7nci sınıf ile 8inci sınıf arasındaki ortalama fark ise -0.016 olarak hesaplanmıştır.

ML kestirim yöntemine göre hesaplanan standart sapma değerleri ise benzer şekilde 6ncı sınıf düzeyinden 7nci sınıf düzeyine yükselmekte ve 7nci sınıf düzeyinden 8inci sınıf düzeyine tekrar düşmektedir. Standart sapma değerleri incelendiğinde 6ncı sınıfın en düşük standart sapmaya (0.608) sahip olduğu görülmektedir. Bu bulgu, 6ncı sınıf düzeyindeki grubun, 7nci sınıf (1.091) ve 8inci sınıf (1.091) düzeyindeki gruba göre daha homojen bir grup olduğu şeklinde yorumlanabilir.

ML kestirimine göre elde edilen ortalama ve standart sapma değerleri kullanılarak elde edilen etki büyüklüklerinin mutlak değerleri incelendiğinde; 6ncı sınıf ile 7nci sınıf arasındaki etki büyüklüğü değeri 0.256, 7nci sınıf ile 8inci sınıf arasındaki etki büyüklüğü değeri 0.010 olarak hesaplanmıştır. 6ncı sınıf ile 7nci sınıf arasındaki etki değeri küçük etki, 7nci sınıf ile 8inci sınıf arasındaki etki değeri ise zayıf etki olarak yorumlanabilir.

EAP yetenek kestirimi sonucuna göre; θ yetenek düzeyi ortalama değerlerinin ML yetenek kestirimine benzer olarak 6ncı sınıf düzeyinden 7nci sınıf düzeyine yükseldiği ve 7nci sınıf düzeyinden 8inci sınıf düzeyine tekrar düştüğü gözlenmektedir. Büyüme örüntüsünü belirlemek için hesaplanan ardışık sınıf düzeyleri arasındaki ortalama farkı değerleri Tablo 4.4'ten incelendiğinde 6ncı sınıf ile 7nci sınıf arasında 0.060, 7nci sınıf ile 8inci sınıf arasında -0.007 olduğu görülmektedir.

EAP yetenek kestirimi yöntemi ile hesaplanan standart sapma değerlerinin üç sınıf düzeyi için birbirine yakın değerler aldığı görülmektedir. 6ncı sınıfın standart sapması 0.822, 7nci sınıfın standart sapması 0.874 ve 8inci sınıfın standart sapması da 0.586 olarak hesaplanmıştır. En düşük standart sapma değerine sahip olan 8inci sınıfın diğer sınıflara göre daha homojen bir grup olduğu söylenebilir.

EAP kestirimine göre elde edilen etki büyüklüklerinin mutlak değerleri incelendiğinde; 6ncı sınıf ile 7nci sınıf arasındaki etki büyüklüğü değeri 0.050, 7nci sınıf ile 8inci sınıf arasındaki etki büyüklüğü değeri ise 0.003 olarak

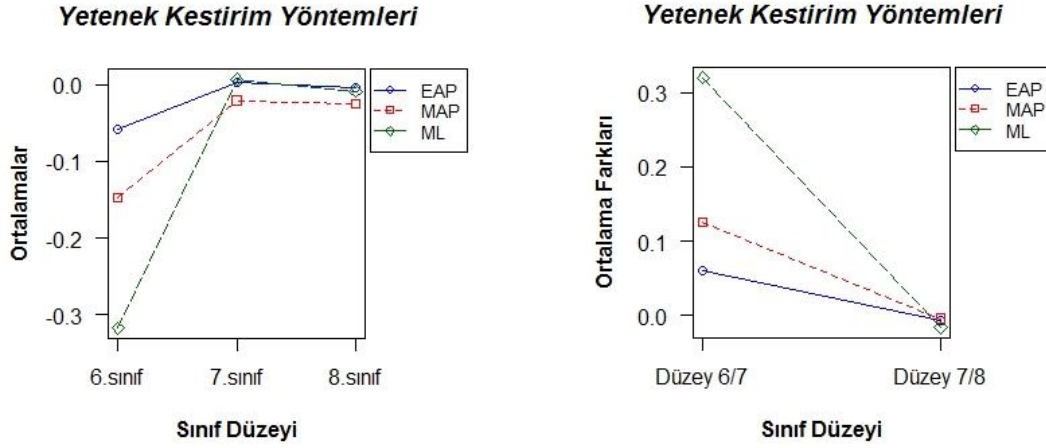
hesaplanmıştır. 6ncı sınıf ile 7nci sınıf arasındaki etki değeri ve 7nci sınıf ile 8inci sınıf arasındaki etki değeri zayıf etki olarak yorumlanabilir.

MAP yetenek kestirimi sonucuna göre; θ yetenek düzeyi ortalama değerinin ML ve EAP yetenek kestirimi yöntemlerine benzer olarak 6ncı sınıf düzeyinden 7nci sınıf düzeyine yükseldiği ve 7nci sınıf düzeyinden 8inci sınıf düzeyine tekrar düştüğü gözlenmektedir. Büyüme örüntüsünü belirlemek için hesaplanan ardışık sınıf düzeyleri arasındaki ortalama farkı değerleri Tablo 4.4'ten incelendiğinde 6ncı sınıf ile 7nci sınıf arasındaki ortalama farkı 0.126, 7nci sınıf ile 8inci sınıf arasındaki ortalama farkı değeri -0.004 olduğu görülmektedir.

MAP yetenek kestirimi yöntemi ile hesaplanan standart sapma değerleri incelendiğinde, 6ncı sınıfın en düşük, 7nci sınıfın ise en büyük standart sapma değerine sahip olduğu görülmektedir. 6ncı sınıfın standart sapması 0.354, 7nci sınıfın standart sapması 0.814 ve 8inci sınıfın standart sapma değeri 0.575 olarak hesaplanmıştır. En düşük standart sapma değerine sahip olan 6ncı sınıfın diğer sınıflara göre daha homojen bir grup olduğu söylenebilir.

MAP kestirimine göre elde edilen ortalama ve standart sapma değerleri kullanılarak elde edilen etki büyüklüklerinin mutlak değerleri incelendiğinde; 6ncı sınıf ile 7nci sınıf arasındaki etki büyüklüğü değeri 0.142, 7nci sınıf ile 8inci sınıf arasındaki etki büyüklüğü değeri ise 0.004 olarak hesaplanmıştır. 6ncı sınıf ile 7nci sınıf arasındaki etki değeri ve 7nci sınıf ile 8inci sınıf arasındaki etki değeri zayıf etki olarak yorumlanabilir.

Bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme kriterini değerlendirmek için yetenek kestirim yöntemleri ile elde edilen ortalama ve ortalama farkların karşılaştırılması aşağıda Grafik 4.4'te verilmiştir.



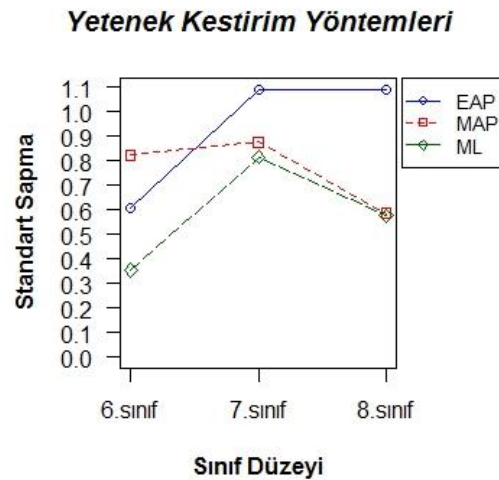
Grafik 4.4 Ayrı Kalibrasyon Yöntemi İle Elde Edilen Ortalama ve Ortalama Farkları

Grafik 4.4'teki ortalamalara ait grafik incelendiğinde, 6ncı sınıf düzeyinde en düşük ortalama θ değeri ML yetenek kestirim yöntemi ile hesaplanırken, en yüksek ortalama θ değerinin EAP kestirim yöntemi ile hesaplandığı görülmektedir. ML 7nci sınıf düzeyinde en düşük ortalama θ değeri MAP ile hesaplanırken ML ve EAP birbirine benzer sonuçlar vermiştir. 8inci sınıf düzeyinde en düşük ortalama θ değeri MAP ile hesaplanırken, en yüksek ortalama θ değeri EAP ile hesaplanmıştır. Ayrı kalibrasyon yöntemi ile yürütülen alanyazındaki çalışmaların ortalama kriteri incelendiğinde ML, EAP ve MAP yetenek kestirim yöntemlerine göre kestirilen yetenek ortalamalarda yakın sonuçlar elde edildiği belirtilmektedir (Meng vd., 2006; Kolen & Tong, 2007).

Grafik 4.4'teki ortalama farklarına ait grafik incelendiğinde, her üç kestirim yönteminde de benzer olarak 6ncı sınıf ile 7nci sınıf arasındaki ortalama farkı pozitif olarak hesaplanmış ve $EAP < MAP < ML$ şeklinde sıralanmaktadır. Bu bulgu 7nci sınıftaki öğrencilerin 6ncı sınıftaki öğrencilerden daha başarılı olduğu ve bir sınıf düzeyinden diğer sınıf düzeyine büyümenin sağlandığı şeklinde yorumlanabilir. 7nci sınıf ile 8inci sınıf arasındaki ortalama farkı ise negatif olarak bulunmuş ve ortalamalar $ML < EAP < MAP$ şeklinde sıralanmıştır. Bu bulgu da 7nci sınıf öğrencilerinin 8inci sınıf öğrencilerinden daha başarılı olduğu, bir sınıf düzeyinden diğer sınıf düzeyine istenen ve beklenen büyümenin sağlanmadığı şeklinde yorumlanabilir. 8inci sınıf düzeyindeki öğrencilerin 7nci sınıfa göre daha az başarılı olmalarının nedeni, uygulamaya yaklaşan TEOG (Temel Eğitimden Ortaöğretime Geçiş) sınavı olabilir. Öğrencilerin kaygı düzeylerindeki artış

başarılarını olumsuz etkilemiş olabilir. Ayrıca lise eğitimine başlayacak olan 8inci sınıf öğrencilerinin ergenlik dönemine girmiş olmaları da psikolojileri ile birlikte başarılarını olumsuz yönde etkilemiş olabilir. Briggs ve Weeks (2009)'in çalışmasında da bu bulguya paralel olarak büyüme örüntüleri bir yıldan diğer yıla lineer olarak bir artış göstermediği belirtilmiştir. Alanyazında bu bulguları destekler çalışmaların olduğu görülmektedir (Tong & Kolen, 2008; Çetin, 2009; Wysel & Reckase, 2011; Altun, 2013). Kolen ve Tong (2010) de çalışmasında düşük sınıf düzeylerinde ortalama farkın fazla olduğunu, sınıf seviyesi arttıkça ortalama farkın azaldığı sonucuna ulaşmıştır. Ito, Skykes ve Yao (2008)'nin ve Tong ve Kolen (2007)'nin dikey ölçekleme yöntemlerinin karşılaştırıldığı çalışmalarında da Kolen ve Tong (2010)'nun çalışmasına benzer olarak küçük sınıf düzeylerinde, büyük sınıf düzeylerine göre daha fazla puan artışı olduğu sonucuna ulaşılmıştır. MTK analizleri sonucunda öğrencilerin puanları sınıf seviyelerine göre artıp azalmaktadır. Yani başarısız öğrencilerin başarı düzeyleri 6ncı sınıftan 7nci sınıfa geçişte, 7nci sınıftan 8inci sınıfa geçişte kıyaslandığı zaman daha fazla artmaktadır.

Diğer bir değerlendirme kriteri olan sınıf düzeyleri arasındaki çeşitlilik kriterini değerlendirmek için; yetenek kestirim yöntemleri ile elde edilen standart sapma değerlerinin karşılaştırılması aşağıda Grafik 4.5'te verilmiştir.



Grafik 4.5 Ayrı Kalibrasyon Yöntemi ile Elde Edilen Standart Sapma Değerleri

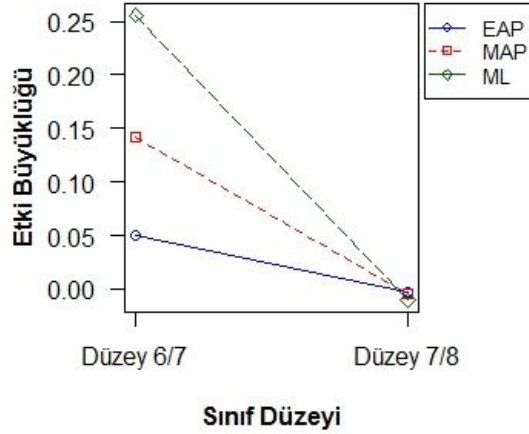
Grafik 4.5'deki standart sapmalara ait grafik incelendiğinde; hemen hemen her üç sınıf düzeyi için en yüksek standart sapma değerlerinin ML yetenek kestirim

yöntemi ile elde edildiği görülmektedir. ML ve MAP yetenek kestirim yöntemlerinde en düşük standart sapma değeri 6ncı sınıf için hesaplanmış, EAP yönteminde en düşük standart sapma değeri 8inci sınıf için hesaplanmıştır. Her bir sınıf düzeyi için standart sapma değerlerini sıralamak gerekirse 6ncı sınıf düzeyi için standart sapma değerleri $MAP < ML < EAP$, 7nci sınıf düzeyi için $MAP < EAP < ML$, 8inci sınıf düzeyi için $MAP < EAP < ML$ olarak sıralanmaktadır.

Alanyazın incelendiğinde elde edilen bu bulguları destekleyen birçok çalışma olduğu görülmektedir. Tong ve Kolen (2007) bu bulgulara benzer olarak standart sapmaların $MAP < EAP < ML$ şeklinde sıralandığı sonucuna ulaşmışlardır. Kim (2007) çalışmasında yetenek kestirim yöntemlerinin standart sapma değerlerini etkilediği ve sınıf düzeyi arttıkça standart sapmaların azaldığı sonucuna ulaşmış ancak bu azalma miktarının kullanılan yetenek kestirim yönteminden etkilenmediğini vurgulamıştır. Tong (2005) ayrı kalibrasyon ile sınıf düzeyleri arasındaki çeşitlilik kriterini değerlendirdiğinde ML yönteminin en büyük standart sapmayı ürettiği sonucuna ulaşmıştır. Tong'un çalışmasındaki standart sapma değerleri $MAP < pseudo-EAP < EAP < QD < ML$ şeklinde sıralanmıştır. Yetenek kestirim yöntemlerinin karşılaştırıldığı çalışmalarda bu bulgularla paralel olarak en yüksek standart sapma değerinin ML ile elde edildiği, EAP kestiriminin ML yöntemine göre daha küçük standart sapmalar verdiği bulgusuna ulaşılmıştır (Beard, 2008 ; Tong & Kolen, 2008; Meng vd., 2006).

Düzye dağılımları arasındaki ayırım kriterini değerlendirmek için; yetenek kestirim yöntemleri ile elde edilen ortalamalar ve standart sapmalar kullanılarak hesaplanan etki büyüklüğü kriterleri karşılaştırılması aşağıda Grafik 4.6'da verilmiştir.

Yetenek Kestirim Yöntemleri



Grafik 4.6 Ayrı Kalibrasyon Yöntemi İle Elde Edilen Etki Büyüklüğü Değerleri

Grafik 4.6 incelendiğinde her üç yetenek kestirim yönteminde de benzer sonuçlar alan etki büyüklüğü indeksi, 6ncı ve 7nci sınıflar arasında pozitif değer alırken, 7nci ve 8inci sınıflar arasında negatif değer aldığı görülmektedir. 6ncı ve 7nci sınıflar arasındaki etki büyüklüğü değerleri tüm kestirim yöntemlerinde, 7nci ve 8inci sınıf arasındaki etki büyüklüğü değerlerine göre daha küçük değerler almıştır. Tablo 4.2'deki etki büyüklüğü indekslerinin mutlak değerleri incelendiğinde değerlerin $EAP < MAP < ML$ olarak sıralandığı görülmektedir. Bu bulgu Kolen ve Tong (2007)'un çalışmasında ulaştığı bulgu ile paraleldir. Meng vd. (2006) ve Tong (2005) ise bu bulgunun aksine en küçük kestirimi ML kestirimi ile elde edildiği bulgusuna ulaşmışlardır.

4.3 Üçüncü Alt Probleme Ait Bulgular ve Yorumlar

Yapay veri setinde eş zamanlı kalibrasyon yöntemi ile elde edilen maksimum olabilirlik kestirimi, beklenen önsel bilgi ve maksimum önsel bilgi yetenek kestirimlerinde; bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme (ortalama ve ortalama farkları), sınıf düzeyleri arasındaki çeşitlilik (standart sapma) ve düzey dağılımlarının ayrımı (etki büyüklüğü) değerleri sınıf düzeyi ile birlikte nasıl değişmektedir?

Yapay veri analizinde belirtilen koşullar altında her bir analiz 100 defa tekrar edilmiş ve çıkan değerlerin ortalaması alınmıştır. Eş zamanlı kalibrasyon analizi

için R programında kod yazılmış ve BILOG-MG 3 programı kullanılarak eş zamanlı kalibrasyon yöntemi ile her bir seviyeye ait madde ve yetenek parametreleri kestirilmiştir. ML, EAP ve MAP yetenek kestirim yöntemlerine göre elde edilen θ yetenek düzeyi ortalamaları, ortalamalar arası farklar, standart sapmalar ve etki büyüklüğü değerleri hesaplanmış ve bu değerler aşağıda Tablo 4.3'te gösterilmiştir.

Tablo 4.4: Yapay Veri Seti Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen ML, EAP ve MAP Yetenek Kestirimine Ait Sonuçlar

	Sınıf Düzeyi	ML	EAP	MAP
Ortalamalar	6	-0.624	-0.492	-0.488
	7	0.000	0.000	0.000
	8	1.136	1.013	0.997
Ortalama farkları	6-7	0.624	0.492	0.488
	7-8	1.136	1.013	0.997
Standart sapmaları	6	0.783	0.610	0.601
	7	1.000	1.000	1.000
	8	0.785	0.659	0.632
Etki büyüklükleri	6-7	0.732	0.719	0.722
	7-8	1.332	1.433	1.445

Tablo 4.4'te her bir sınıf düzeyine ilişkin ML, EAP ve MAP yetenek kestirimi sonucunda elde edilen ortalama, ortalamalar farkı ve standart sapma değerleri görülmektedir. Eş zamanlı kalibrasyon ile MTK parametreleri kestirilirken her üç yetenek kestiriminde de referans sınıf olarak seçilen 7nci sınıfın θ yetenek düzeyi ortalamasının sıfır, standart sapmasının ise bir olduğu görülmektedir.

ML yetenek kestirimi sonucuna göre; θ yetenek düzeyi ortalamaları 6ncı sınıf düzeyinden 7nci sınıf düzeyine ve 7nci sınıf düzeyinden 8inci sınıf düzeyine artış gösterdiği görülmektedir. Büyüme örüntüsünü belirlemek için ardışık sınıf düzeyleri arasındaki ortalama fark değerleri incelendiğinde, 6ncı sınıf ile 7nci sınıf arasındaki ortalama farkı 0.624, 7nci sınıf ile 8inci sınıf arasındaki ortalama farkı değerinin ise 1.136 olduğu görülmektedir.

ML yetenek kestirim yöntemine göre hesaplanan standart sapma değerleri incelendiğinde en yüksek standart sapma değerinin referans sınıf olan 7nci sınıfa

ait olduđu, 6ncı sınıfa (0.783) ve 8inci sınıfa (0.785) ait standart sapma deęerlerinin ise benzer olduđu grlmektedir.

ML kestirimine gre elde edilen ortalama ve standart sapma deęerleri kullanılarak elde edilen etki byklklerinin mutlak deęerleri incelendięinde; 6ncı sınıf ile 7nci sınıf arasındaki etki byklę deęeri 0.732, 7nci sınıf ile 8inci sınıf arasındaki etki byklę deęeri 1.332 olarak hesaplanmıřtır. 6ncı sınıf ile 7nci sınıf arasındaki etki deęeri orta etki, 7nci sınıf ile 8inci sınıf arasındaki etki deęeri ise gçl etki olarak yorumlanabilir.

EAP yetenek kestirimi sonucuna gre; θ yetenek dzeyi ortalamaları 6ncı sınıf dzeyinden 7nci sınıf dzeyine ve 7nci sınıf dzeyinden 8inci sınıf dzeyine artıř gsterdięi grlmektedir. Byme rntsn belirlemek iin ardıřık sınıf dzeyleri arasındaki ortalama fark deęerleri incelendięinde, 6ncı sınıf ile 7nci sınıf arasındaki ortalama farkı 0.492, 7nci sınıf ile 8inci sınıf arasındaki ortalama farkı deęerinin ise 1.013 olduđu grlmektedir.

EAP yetenek kestirim yntemine gre hesaplanan standart sapma deęerleri incelendięinde en yksek standart sapma deęerinin referans sınıf olan 7nci sınıfa ait olduđu, 6ncı sınıfa ait standart sapma deęerinin 0.610, 7nci sınıfa ait standart sapma deęerinin ise 0.659 olduđu grlmektedir.

EAP kestirimine gre elde edilen ortalama ve standart sapma deęerleri kullanılarak elde edilen etki byklkleri incelendięinde; 6ncı sınıf ile 7nci sınıf arasındaki etki byklę deęeri 0.732, 7nci sınıf ile 8inci sınıf arasındaki etki byklę deęeri 1.332 olarak hesaplandığı grlmektedir. 6ncı sınıf ile 7nci sınıf arasındaki etki deęeri orta etki, 7nci sınıf ile 8inci sınıf arasındaki etki deęeri ise gçl etki olarak yorumlanabilir.

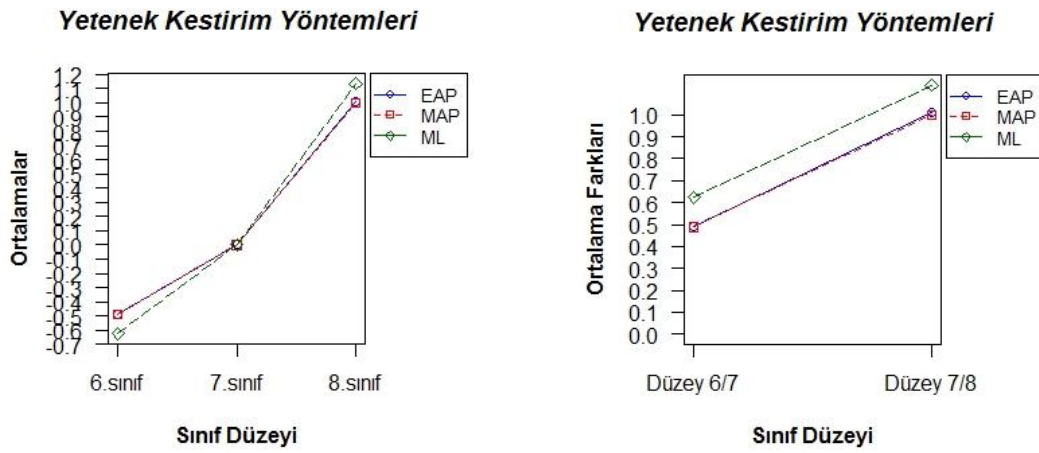
MAP yetenek kestirimi sonucuna gre; θ yetenek dzeyi ortalamaları 6ncı sınıf dzeyinden 7nci sınıf dzeyine ve 7nci sınıf dzeyinden 8inci sınıf dzeyine artıř gsterdięi grlmektedir. Byme rntsn belirlemek iin ardıřık sınıf dzeyleri arasındaki ortalama fark deęerleri incelendięinde, 6ncı sınıf ile 7nci sınıf arasındaki ortalama farkı 0.488, 7nci sınıf ile 8inci sınıf arasındaki ortalama farkı deęerinin ise 0.997 olduđu grlmektedir.

MAP yetenek kestirim yntemine gre hesaplanan standart sapma deęerleri incelendięinde en yksek standart sapma deęerinin referans sınıf olan 7nci sınıfa

ait olduğu, 6ncı sınıfa ait standart sapma değerinin 0.601, 7nci sınıfa ait standart sapma değerinin ise 0.659 olduğu görülmektedir.

MAP kestirimine göre elde edilen ortalama ve standart sapma değerleri kullanılarak elde edilen etki büyüklüklerinin mutlak değerleri incelendiğinde; 6ncı sınıf ile 7nci sınıf arasındaki etki büyüklüğü değeri 0.722, 7nci sınıf ile 8inci sınıf arasındaki etki büyüklüğü değeri 1.445 olarak hesaplanmıştır. 6ncı sınıf ile 7nci sınıf arasındaki etki değeri orta etki, 7nci sınıf ile 8inci sınıf arasındaki etki değeri ise güçlü etki olarak yorumlanabilir.

Bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme kriterini değerlendirmek için yetenek kestirim yöntemleri ile elde edilen ortalama ve ortalama farkları karşılaştırılması aşağıda Grafik 4.7’de verilmiştir.

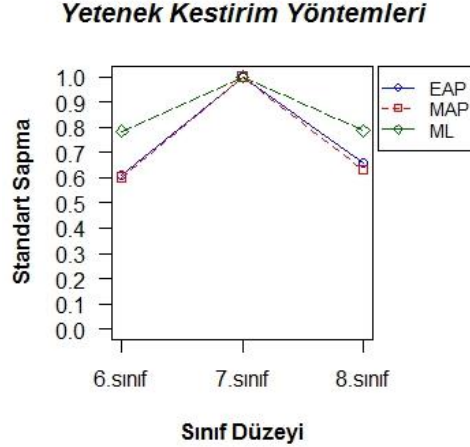


Grafik 4.7 Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen Ortalama ve Ortalama Farkları

Grafik 4.7’deki ortalamalara ait grafik incelendiğinde, 6ncı sınıf düzeyinde en düşük ortalama θ değeri ML yetenek kestirim yöntemi ile hesaplanırken, ML ve EAP kestirim yöntemi ile hesaplanan ortalama θ değerlerinin birbirine yakın olduğu görülmektedir. 7nci sınıf düzeyi referans sınıf seçildiğinden tüm ortalamalar sıfır olarak bulunmuştur. 8inci sınıf düzeyinde en düşük ortalama θ değeri MAP ile hesaplanırken, en yüksek ortalama θ değeri ML ile hesaplanmıştır. Ortalama farklar incelendiğinde 6ncı sınıf ile 7nci sınıf arasındaki θ ortalama farkları MAP < EAP < ML şeklinde sıralanırken, 7nci sınıf ile 8inci sınıf arasındaki θ ortalama farklarının da yine aynı şekilde MAP < EAP < ML şeklinde sıralandığı

görülmektedir. Bu değerler incelendiğinde 6ncı sınıftan 8inci sınıfa doğru artan bir başarı örüntüsü olduğu görülmektedir.

Sınıf düzeyleri arasındaki çeşitlilik kriterini değerlendirmek için; üç yetenek kestirim yöntemi ile elde edilen standart sapma değerleri karşılaştırılması aşağıda Grafik 4.8'de verilmiştir.

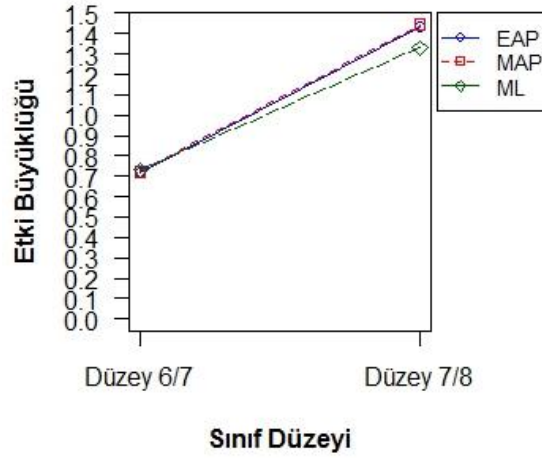


Grafik 4.8 Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen Standart Sapma Değerleri

Grafik 4.8 incelendiğinde, 6ncı sınıf için en yüksek standart sapma değeri ML ile hesaplanırken en düşük standart sapma değerinin ise MAP yöntemi ile hesaplandığı görülmektedir. 6ncı sınıf düzeyi için standart sapma değerlerini sıralamak gerekirse, standart sapma değerleri MAP < EAP < ML şeklinde sıralanabilir. 7nci sınıf referans sınıf olarak seçildiği için tüm yöntemlerde standart sapma değeri 1 olarak hesaplanmıştır. 8inci sınıf düzeyinde en yüksek standart sapma ML ile hesaplanırken, en düşük standart sapma değeri MAP ile hesaplanmıştır. 8inci sınıf düzeyi için standart sapma değerlerini sıralamak gerekirse, standart sapma değerleri 6ncı sınıf ile benzer olarak MAP < EAP < ML şeklinde sıralanmıştır.

Düzye dağılımları arasındaki ayırım kriterini değerlendirmek için; yetenek kestirim yöntemleri ile elde edilen ortalamalar ve standart sapmalar kullanılarak hesaplanan etki büyüklüğü kriterlerinin karşılaştırılması Grafik 4.9'da verilmiştir.

Yetenek Kestirim Yöntemleri



Grafik 4.9 Eş Zamanlı Kalibrasyon Yöntemi İle Elde Edilen Etki Büyüklüğü Değerleri

Grafik 4.9 incelendiğinde, etki büyüklüğü değerlerinin her üç yetenek kestirim yönteminde de benzer sonuçlar aldığı görülmektedir. Grafik 4.9'daki etki büyüklüğü indeksleri incelendiğinde etki büyüklüğü indeksinin hem 6ncı ve 7nci sınıf arasında hem de 7nci ve 8inci sınıf arasında $EAP < MAP < ML$ şeklinde sıralandığı görülmektedir.

4.4 Dördüncü Alt Probleme Ait Bulgular ve Yorumlar

Yapay veri setinde ayrı kalibrasyon yöntemi ile elde edilen maksimum olabilirlik kestirimi, beklenen önsel bilgi ve maksimum önsel bilgi yetenek kestirimlerinde; bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme (ortalama ve ortalama farkları), sınıf düzeyleri arasındaki çeşitlilik (standart sapma) ve düzey dağılımlarının ayrımı (etki büyüklüğü) değerleri sınıf düzeyi ile birlikte nasıl değişmektedir?

Yapay veri analizinde belirtilen koşullar altında her bir analiz 100 defa tekrar edilmiş ve çıkan değerlerin ortalaması alınmıştır. Ayrı kalibrasyon analizi için R programında kod yazılmış ve BILOG-MG 3 programı kullanılarak ayrı kalibrasyon yöntemi ile her bir seviyeye ait madde ve yetenek parametreleri ayrı ayrı kestirilmiştir. Bağlama yöntemi olarak SL kullanılmıştır. Test karakteristik eğrilerini kullanan SL yöntemi sonucu elde edilen değerler aşağıda Tablo 4.5'te verilmiştir.

Tablo 4.5: Stocking-Lord Dönüşümü İçin Elde Edilen A ve B Sabitleri

Sınıf Düzeyi	A (Slope)	B (Intercept)
6-7	1.417	0.357
7-8	1.250	0.812

Tablo4.5'teki değerler kullanılarak dönüşüm yapılır. 7nci sınıf referans sınıf olduğu için 6ncı sınıfı 7nci sınıf ölçeğine dönüştürürken, yetenek kestirimleri $\theta_{yeni} = \theta_{eski} \times 1.417 + (0.357)$ eşitliğiyle yapılırken, 8inci sınıfı 7nci sınıfa dönüştürürken yetenek kestirimi $\theta_{yeni} = \theta_{eski} \times 1.250 + (0.812)$ eşitliği yardımı ile hesaplanır. 8inci sınıfı 6ncı sınıfa dönüştürmek için ise iki dönüşüme ihtiyaç vardır. 8inci sınıfı dönüştürmek için, $\theta_{yeni} = (\theta_{eski} \times 1.417 + (0.357)) \times 1.250 + (0.812)$ eşitliği kullanılmıştır. Kesişim değerleri tüm sınıflar için pozitifdir.

ML, EAP ve MAP yetenek kestirim yöntemlerine göre elde edilen θ yetenek düzeyi ortalamaları, ortalamalar arası farklar, standart sapmalar ve etki büyüklüğü değerleri hesaplanmış ve bu değerler aşağıda Tablo 4.6'da gösterilmiştir.

Tablo 4.6: Ayrı Kalibrasyon Yöntemi İle Elde Edilen ML, EAP ve MAP Yetenek Kestirimine Ait Sonuçlar

	Sınıf Düzeyi	ML	EAP	MAP
Ortalamalar	6	-0.015	-0.002	0.004
	7	-0.007	0.001	0.005
	8	0.013	0.003	0.008
Ortalama farkları	6-7	0.008	0.003	0.001
	7-8	0.020	0.002	0.003
Standart sapmaları	6	1.159	0.773	0.764
	7	1.130	0.884	0.868
	8	1.155	0.863	0.856
Etki büyüklükleri	6-7	0.007	0.004	0.001
	7-8	0.018	0.002	0.003

Tablo 4.6 incelendiğinde her bir sınıf düzeyine ilişkin ML, EAP ve MAP yetenek kestirimi sonucunda elde edilen ortalama, ortalamalar farkı ve standart sapma değerleri görülmektedir.

ML yetenek kestirimi sonucuna göre; θ yetenek düzeyi ortalamaları 6ncı sınıf düzeyinden 7nci sınıf düzeyine ve 7nci sınıf düzeyinden 8inci sınıf düzeyine

yükseldiği görülmektedir. Büyüme örüntüsünü belirlemek için hesaplanan ardışık sınıf düzeyleri arasındaki ortalama farkı değerleri de sınıf düzeyleri arasındaki bu farklılaşmayı bize göstermektedir. 6ncı sınıf ile 7nci sınıf arasındaki ortalama fark 0.008, 7nci sınıf ile 8inci sınıf arasındaki ortalama fark ise 0.020 olarak hesaplanmıştır.

ML kestirim yöntemine göre hesaplanan standart sapma değerleri ise birbirine yakın değerler almıştır. Standart sapma değerleri incelendiğinde 7nci sınıfın en düşük standart sapmaya (1.130) sahip olduğu görülmektedir. Bu bulgu, 7nci sınıf düzeyindeki grubun, 6ncı sınıf (1.159) ve 8inci sınıf (1.155) düzeyindeki gruba göre daha homojen bir grup olduğu şeklinde yorumlanabilir.

ML kestirimine göre elde edilen ortalama ve standart sapma değerleri kullanılarak elde edilen etki büyüklükleri incelendiğinde; 6ncı sınıf ile 7nci sınıf arasındaki etki büyüklüğü değeri 0.007, 7nci sınıf ile 8inci sınıf arasındaki etki büyüklüğü değeri 0.018 olarak hesaplanmıştır. Bu etki değerleri zayıf etki olarak yorumlanabilir.

EAP yetenek kestirimi sonucuna göre; θ yetenek düzeyi ortalama değerlerinin ML yetenek kestirimine benzer olarak 6ncı sınıf düzeyinden 7nci sınıf düzeyine, 7nci sınıf düzeyinden 8inci sınıf düzeyine doğru doğrusal olarak yükseldiği gözlenmektedir. Tablo 4.5'teki büyüme örüntüsünü belirlemek için hesaplanan ardışık sınıf düzeyleri arasındaki ortalama farkı değerleri incelendiğinde, 6ncı sınıf ile 7nci sınıf arasında 0.003, 7nci sınıf ile 8inci sınıf arasında 0.002 olduğu görülmektedir.

EAP yetenek kestirimi yöntemi ile hesaplanan standart sapma değerleri incelendiğinde, standart sapma değerlerinin üç sınıf düzeyi için birbirine yakın değerler aldığı görülmektedir. 6ncı sınıfın standart sapması 0.773, 7nci sınıfın standart sapması 0.884 ve 8inci sınıfın standart sapması da 0.863 olarak hesaplanmıştır. En düşük standart sapma değerine sahip olan 6ncı sınıfın diğer sınıflara göre daha homojen bir grup olduğu söylenebilir.

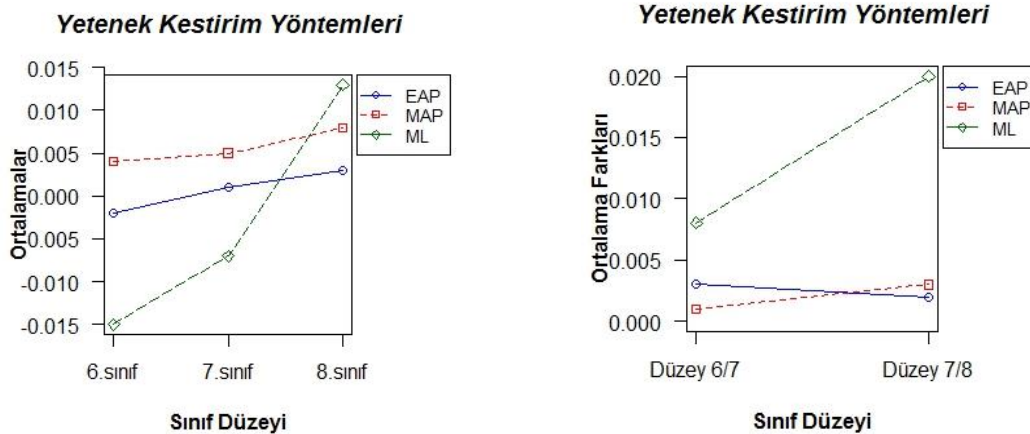
EAP kestirimine göre elde edilen ortalama ve standart sapma değerleri kullanılarak elde edilen etki büyüklüklerinin mutlak değerleri incelendiğinde; 6ncı sınıf ile 7nci sınıf arasındaki etki büyüklüğü değeri 0.004, 7nci sınıf ile 8inci sınıf arasındaki etki büyüklüğü değeri ise 0.002 olarak hesaplanmıştır. Bu etki değerleri zayıf etki olarak yorumlanabilir.

MAP yetenek kestirimi sonucuna göre; θ yetenek düzeyi ortalama değerinin ML ve EAP yetenek kestirimi yöntemlerine benzer olarak 6ncı sınıf düzeyinden 7nci sınıf düzeyine ve 7nci sınıf düzeyinden 8inci sınıf düzeyine yükseldiği gözlenmektedir. Büyüme örüntüsünü belirlemek için hesaplanan ardışık sınıf düzeyleri arasındaki ortalama farkı değerleri Tablo 4.5'ten incelendiğinde 6ncı sınıf ile 7nci sınıf arasındaki ortalama farkı 0.001, 7nci sınıf ile 8inci sınıf arasındaki ortalama farkı değeri 0.003 olduğu görülmektedir.

MAP yetenek kestirimi yöntemi ile hesaplanan standart sapma değerleri incelendiğinde, standart sapma değerlerinin diğer yöntemlere benzer şekilde üç sınıf düzeyi için birbirine yakın değerler aldığı görülmektedir. 6ncı sınıfın standart sapması 0.764, 7nci sınıfın standart sapması 0.868 ve 8inci sınıfın standart sapma değeri 0.856 olarak hesaplanmıştır. En düşük standart sapma değerine sahip olan 6ncı sınıfın diğer sınıflara göre daha homojen bir grup olduğu söylenebilir.

MAP kestirimine göre elde edilen ortalama ve standart sapma değerleri kullanılarak elde edilen etki büyüklükler incelendiğinde; 6ncı sınıf ile 7nci sınıf arasındaki etki büyüklüğü değeri 0.001, 7nci sınıf ile 8inci sınıf arasındaki etki büyüklüğü değeri ise 0.003 olarak hesaplanmıştır. 6ncı sınıf ile 7nci sınıf arasındaki etki değeri ve 7nci sınıf ile 8inci sınıf arasındaki etki değeri zayıf etki olarak yorumlanabilir.

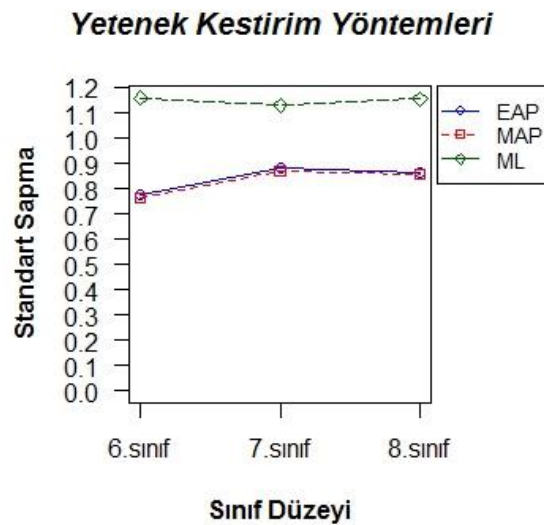
Bir sınıf düzeyinden diğer sınıf düzeyine olan büyüme kriterini değerlendirmek için yetenek kestirim yöntemleri ile elde edilen ortalama ve ortalama farkları karşılaştırılması aşağıda Grafik 4.10' da verilmiştir.



Grafik 4.10 Ayrı Kalibrasyon Yöntemi İle Elde Edilen Ortalama ve Ortalama Farkları

Grafik 4.10'daki ortalamalara ait grafik incelendiğinde, 6ncı sınıf düzeyinde en düşük ortalama θ değeri ML yetenek kestirim yöntemi ile hesaplanırken, en yüksek ortalama θ değerinin MAP kestirim yöntemi ile hesaplandığı görülmektedir. Bu sonuç gerçek veri seti ile benzerlik göstermektedir Gerçek veri setinde de en düşük ortalama θ değeri ML yetenek kestirim yöntemi ile hesaplanmıştır. 7nci sınıf düzeyinde en düşük ortalama θ değeri yine ML ile hesaplanırken ML ve EAP birbirine benzer sonuçlar vermiştir. 8inci sınıf düzeyinde en düşük ortalama θ değeri EAP ile hesaplanırken, en yüksek ortalama θ değeri ML ile hesaplanmıştır. Ayrı kalibrasyon yöntemi ile yürütülen alanyazın çalışmalarının ortalama kriteri incelendiğinde ML, EAP ve MAP yetenek kestirim yöntemlerine göre kestirilen yetenek ortalamalarda yakın sonuçlar elde edildiği belirtilmektedir (Meng vd., 2006; Kolen ve Tong, 2007). Grafik 4.10'daki Ortalama farkları incelendiğinde, 6ncı sınıf ile 7nci sınıf arasındaki ortalama farkı $MAP < EAP < ML$ şeklinde sıralanmaktadır. Gerçek veri seti ile karşılaştırıldığında gerçek veri setinde de en yüksek ortalama fark ML yöntemi ile hesaplanmıştır. 7nci sınıf ile 8inci sınıf arasındaki ortalama farkı ise $EAP < MAP < ML$ şeklinde sıralanmaktadır. Yapay veri setinde en yüksek ortalama ML ile elde edilirken, gerçek veri setinde ise en yüksek ortalama MAP ile hesaplanmış, $ML < EAP < MAP$ sıralaması elde edilmiştir.

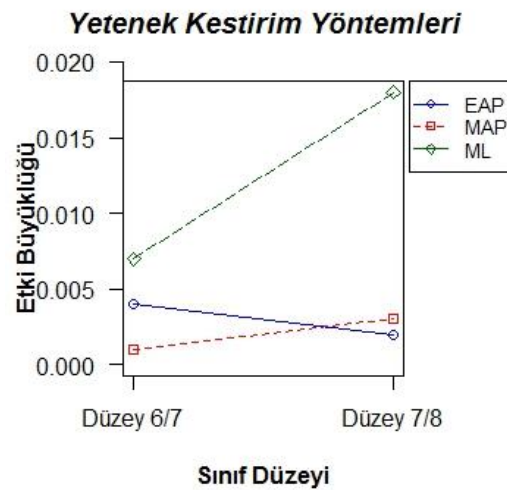
Sınıf düzeyleri arasındaki çeşitlilik kriterini değerlendirmek için; yetenek kestirim yöntemleri ile elde edilen standart sapma değerleri karşılaştırılması Grafik 4.11'de verilmiştir.



Grafik 4.11 Ayrı Kalibrasyon Yöntemi İle Elde Edilen Standart Sapma Değerleri

Grafik 4.11 incelendiğinde, her üç sınıf düzeyi için en yüksek standart sapma değerlerinin, gerçek veri seti ile paralel olarak, ML yetenek kestirim yöntemi ile elde edildiği görülmektedir. EAP ve MAP yetenek kestirim yöntemlerinde en düşük standart sapma değeri 6ncı sınıf için hesaplanmış, ML yönteminde en düşük standart sapma değeri 7nci sınıf için hesaplanmıştır. Standart sapma değerlerini sıralamak gerekirse tüm sınıf düzeyleri için standart sapma değerleri MAP < EAP < ML olarak sıralanmaktadır. Alan yazın incelendiğinde elde edilen bu bulguları destekleyen birçok çalışma olduğu görülmektedir. Tong ve Kolen (2007) bu bulgulara benzer olarak standart sapmaların MAP < EAP < ML şeklinde sıralandığı sonucuna ulaşmışlardır. Tong (2005) ayrı kalibrasyon ile sınıf düzeyleri arasındaki çeşitlilik kriterini değerlendirdiğinde ML yönteminin en büyük standart sapmayı ürettiği sonucuna ulaşmıştır. Tong'un çalışmasındaki standart sapma değerleri MAP < pseudo-EAP < EAP < QD < ML şeklinde sıralanmıştır. Yetenek kestirim yöntemlerinin karşılaştırıldığı çalışmalarda bu bulgularla paralel olarak en yüksek standart sapma değerinin ML ile elde edildiği, EAP kestiriminin ML yöntemine göre daha küçük standart sapmalar verdiği bulgusuna ulaşılmıştır (Beard, 2008; Tong & Kolen, 2008; Meng vd., 2006).

Düzye dağılımları arasındaki ayırım kriterini değerlendirmek için; yetenek kestirim yöntemleri ile elde edilen ortalamalar ve standart sapmalar kullanılarak hesaplanan etki büyüklüğü kriterleri karşılaştırılması Grafik 4.12'de verilmiştir.



Grafik 4.12 Ayrı Kalibrasyon Yöntemi İle Elde Edilen Etki Büyüklüğü Değerleri

Grafik 4.12 incelendiğinde, etki büyüklüğü değerlerinin her üç yetenek kestirim yönteminde de benzer sonuçlar aldığı görülmektedir. Grafik 4.12'de 6ncı ve 7nci sınıf etki büyüklüğü değerlerinin $MAP < EAP < ML$ olarak sıralandığı, 7nci ve 8inci sınıf etki büyüklüğü değerlerinin ise $EAP < MAP < ML$ olarak sıralandığı görülmektedir. Bu bulgu Kolen ve Tong (2007)'un çalışmasında ulaştığı bulgu ile paraleldir.

5. SONUÇ ve ÖNERİLER

Bu bölümde araştırmanın bulgu ve yorumlarına dayalı olarak ulaşılan sonuçların özetine ve bu sonuçlardan yola çıkarak geliştirilen önerilere yer verilmiştir.

5.1 Sonuçlar

Bu araştırmada dikey ölçekleme işlemi sonucunda sınıf düzeyleri arttıkça oluşan başarı gelişiminin örüntüsünü çıkarmak ve dikey ölçekleme sonuçlarının kullanılan farklı yöntemlere göre nasıl değiştiğini belirlemek amaçlanmıştır. Gerçek veriler ve yapay veriler kullanılarak farklı kalibrasyon yöntemleri (ayrı ve eş zamanlı kalibrasyon) ve farklı yeterlik tahmini yöntemleri (maksimum olabilirlik, beklenen önsel dağılım, maksimum önsel dağılım) kullanılarak oluşturulan farklı dikey ölçekleme sonuçlarının ölçeklerin ortalama, standart sapma ve etki büyüklüğü değerlerine göre karşılaştırılması yapılmıştır.

Araştırmada sonucunda ulaşılan sonuçlar şöyle özetlenebilir;

1. Gerçek veriler kullanıldığında eş zamanlı kalibrasyonda bir sınıf düzeyinden diğer sınıf düzeyine büyümeyi ifade eden ortalamalar arası fark; yetenek kestirim yöntemlerine göre ortalamalar arası fark büyüklükleri $ML < MAP < EAP$ şeklinde sıralanmaktadır.
2. Gerçek veriler kullanıldığında eş zamanlı kalibrasyonda sınıf düzeyleri arasındaki çeşitliliği ifade eden standart sapma değerleri; 6ncı sınıf düzeyinde $ML < MAP < EAP$, 8inci sınıf düzeyinde $EAP < MAP < ML$ şeklinde sıralanmaktadır.
3. Gerçek veriler kullanıldığında eş zamanlı kalibrasyonda düzey dağılımları arasındaki ayrımı ifade eden etki büyüklüğü değerleri; $ML < MAP < EAP$ şeklinde sıralanmaktadır.
4. Gerçek veriler kullanıldığında ayrı kalibrasyonda bir sınıf düzeyinden diğer sınıf düzeyine büyümeyi ifade eden ortalamalar arası fark; yetenek kestirim yöntemlerine göre ortalamalar arası fark büyüklükleri 6ncı Sınıf ile 7nci Sınıflar arasında $EAP < MAP < ML$ şeklinde sıralanırken, 7nci Sınıf ve 8inci Sınıflar arasında $ML < EAP < MAP$ şeklinde sıralanmaktadır.
5. Gerçek veriler kullanıldığında ayrı kalibrasyonda sınıf düzeyleri arasındaki çeşitliliği ifade eden standart sapma değerleri; 6ncı sınıf düzeyi için standart

sapma deęerleri $MAP < ML < EAP$, 7nci sınıf d¼zeyi iin $MAP < EAP < ML$, 8inci sınıf d¼zeyi iin $MAP < EAP < ML$ Őeklinde sıralanmaktadır.

6. Gerek veriler kullanıldıęında ayrı kalibrasyonda d¼zey daęılımları arasındaki ayrımı ifade eden etki b¼y¼kl¼ę¼ deęerleri; $EAP < MAP < ML$ Őeklinde sıralanmaktadır.
7. Gerek veriler ile yapılan dikey ¼leklemede farklı kalibrasyon y¼ntemlerinden elde edilen sonular karŐılaŐtırıldıęında; ayrı kalibrasyon ile elde edilen ortalama farkların eŐ zamanlı kalibrasyon ile elde edilen ortalama farklarından daha d¼Ő¼k olduęu, ayrı kalibrasyon ile elde edilen standart sapma deęerlerinin genel olarak eŐ zamanlı kalibrasyon ile elde edilen deęerlere g¼re daha d¼Ő¼k olduęu ve ayrı kalibrasyon ile elde edilen etki b¼y¼kl¼ę¼ deęerlerinin eŐ zamanlı kalibrasyon ile elde edilen deęerlere g¼re daha d¼Ő¼k olduęu g¼r¼lmektedir.
8. Yapay veri seti kullanıldıęında eŐ zamanlı kalibrasyonda bir sınıf d¼zeyinden dięer sınıf d¼zeyine b¼y¼meyi ifade eden ortalamalar arası fark; yetenek kestirim y¼ntemlerine g¼re ortalamalar arası fark b¼y¼kl¼kleri $MAP < EAP < ML$ Őeklinde sıralanmaktadır.
9. Yapay veri seti kullanıldıęında eŐ zamanlı kalibrasyonda sınıf d¼zeyleri arasındaki eŐitlilięi ifade eden standart sapma deęerleri $MAP < EAP < ML$ Őeklinde sıralanmaktadır.
10. Yapay veri seti kullanıldıęında eŐ zamanlı kalibrasyonda d¼zey daęılımları arasındaki ayrımı ifade eden etki b¼y¼kl¼ę¼ deęerleri; $ML < MAP < EAP$ Őeklinde sıralanmaktadır.
11. Yapay veri seti kullanıldıęında ayrı kalibrasyonda bir sınıf d¼zeyinden dięer sınıf d¼zeyine b¼y¼meyi ifade eden ortalamalar arası fark b¼y¼kl¼kleri; 6ncı sınıf ile 7nci sınıf ortalamalar arası fark deęerleri $MAP < EAP < ML$ Őeklinde sıralanırken, 7nci sınıf ile 8inci sınıf ortalamalar arası fark deęerleri $EAP < MAP < ML$ Őeklinde sıralanmaktadır.
12. Yapay veri seti kullanıldıęında ayrı kalibrasyonda sınıf d¼zeyleri arasındaki eŐitlilięi ifade eden standart sapma deęerleri $MAP < EAP < ML$ Őeklinde sıralanmaktadır.

- 13.Yapay veri seti kullanıldığında ayrı kalibrasyonda düzey dağılımları arasındaki ayrımı ifade eden etki büyüklüğü değerleri; 6ncı ve 7nci sınıf etki büyüklüğü değerleri $MAP < EAP < ML$ şeklinde, 7nci ve 8inci sınıf etki büyüklüğü değerlerinin ise $EAP < MAP < ML$ şeklinde sıralanmaktadır.
- 14.Yapay veriler ile yapılan dikey ölçeklemede farklı kalibrasyon yöntemlerinden elde edilen sonuçlar karşılaştırıldığında; ayrı kalibrasyon ile elde edilen ortalama farkların eş zamanlı kalibrasyon ile elde edilen ortalama farklarından daha düşük olduğu, ayrı kalibrasyon ile elde edilen standart sapma değerleri ile eş zamanlı kalibrasyon ile elde edilen standart sapma değerlerinin genel olarak benzer olduğu ve ayrı kalibrasyon ile elde edilen etki büyüklüğü değerlerinin eş zamanlı kalibrasyon ile elde edilen değerlere göre daha düşük olduğu görülmektedir.
- 15.Yapay veri seti kullanılarak yapılan analizler sonucunda ölçekleme kombinasyonlarının benzer sonuçlar ürettiği görülmüştür. Fakat gerçek veri seti kullanıldığında ölçekleme kombinasyonlarının sonuçlarında önemli farklılıklar olduğu görülmektedir. Gerçek veriler kullanılarak yetenek kestirim yöntemleri ve kalibrasyon yöntemleri ile farklı büyüme modelleri elde edilmiştir.

5.2 Öneriler

Araştırma bulguları doğrultusunda araştırmacılara bazı öneriler getirilmiştir. Bu öneriler araştırma sonuçlarından çıkarılan öneriler ve araştırmacılara yönelik öneriler olarak iki başlıkta incelenmiştir.

5.2.1 Uygulamaya Dönük Öneriler

1. Bu araştırmadan elde edilen bulgular sadece bu çalışmada ele alınan koşullarla sınırlıdır. Dikey ölçekleme süreci karmaşık bir süreçtir. Bu süreçte ele alınan koşulların birbiriyle etkileşimi dikey ölçekleme sonucunu dolayısıyla öğrenci başarısının gelişimine yönelik yapılacak yorumları etkileyecektir. Öğrenci başarıları hakkında karar verirken farklı yöntemlerin de kullanılarak karşılaştırma yapılması önerilebilir.
2. Alanyazında kesin doğru bir dikey ölçekleme yöntemi olmadığı sıklıkla ifade edilmektedir. Dikey ölçekleme kararları ölçeğin doğasını etkileyeceği için öğrenci başarısının belirlenmesi ile ilgilenen test geliştirme uzmanları ve

test uygulayıcılarının dikey ölçekleme yaparken dikkatli olması önerilmektedir.

3. Gerçek veri seti kullanıldığında ölçekleme kombinasyonlarının sonuçlarında önemsiz olmayan farklılıklar olduğu görülmektedir. Bunun nedeni, olası varsayım ihlalleri olabileceği için, gerçek veri seti kullanıldığında mutlaka varsayımların tamamen karşılandığının kontrolü yapılmalıdır.
4. Ülkemizde öğrencilerin yıldan yıla başarılarındaki değişimin belirlenebilmesi için dikey ölçekleme uygulamaları oldukça önemlidir. Öğrencilerin K-12 seviyesinde başarılarının takibi için dikey ölçekleme çalışmaları yürütülmesi önerilebilir.

5.2.2 Araştırmacılara Dönük Öneriler

1. Bu araştırma tüm dikey ölçekleme sürecinin bir parçasıdır. Test geliştirme uzmanları ve uygulayıcılarına dikey ölçekleme sürecinin son aşamasında olan gözlenen-gerçek puan eşitleme süreci ile ilgili çalışmaları, gözlenen puanları etkileyen faktörleri incelemeleri önerilebilir.
2. Bu çalışmada test uzunluğu (40ar madde), ortak madde sayısı (10ar madde), örneklem sayısı (500er kişi), uygulanan model (2PLM) gibi faktörler koşul olarak belirlenmemiş ve sabit tutulmuştur. Araştırmacılar bu farklı koşulların dikey ölçekleme sonuçlarına etkisini araştırabilirler.
3. Bu çalışmada dikey eşitlemede kalibrasyon yöntemlerinden eş zamanlı ve ayrı kalibrasyon, yetenek kestirim yöntemlerinden ise EAP, MAP ve ML kullanılmıştır. Araştırma bulgularından yöntemlerin farklı koşullarla farklı sonuçlar verdiği görülmüştür. Araştırmacılar sabit kalibrasyon, yarı eş zamanlı kalibrasyon gibi kalibrasyon yöntemleri veya pseudo-EAP, pseudo-MAP gibi farklı yetenek kestirim yöntemleri kullanarak karşılaştırma yapabilirler.
4. Bu çalışma ilköğretim 6ncı, 7nci ve 8inci sınıf düzeyinde yürütülmüştür. Çalışma genişletilerek K-12 seviyesine yapılabilir.
5. Bu çalışma sadece fen bilgisi dersi üzerinden yürütülmüştür. Matematik, Türkçe gibi farklı alanlarda dikey ölçekleme çalışmaları yapılabilir.

6. Arařtırmacılar, boylamsal bir arařtırma yrtp uzun yıllar aynı đrencilerin bařarısını inceleyebilir ve analizlerini aynı đrenciler zerinden yapabilirler.
7. Dikey lklemede kullanılan bu yntemlerin dođruluđunu deđerlendirmek iin tek ve kesin bir lt olmadığı iin arařtırmacılara likleme sonularını karřılařtırırken birden fazla deđerlendirme lt (ortalama, ortalama farkları, standart sapma, etki byklđ, yatay uzaklık, eřitleme hatası (RMSE) ve yanlılık (bias) deđerleri) kullanmaları nerilebilir.
8. Bu alıřmada gerek veri ve simlasyon verisi kullanılarak farklı trde iki veri setinden elde edilen sonular karřılařtırılmıřtır. Simle verilerde 6ncı, 7nci ve 8inci sınıf dzeyleri iin θ yetenek parametreleri [-1, 2] aralıđında, a parametresi [0.5, 2] aralıđında, b parametreleri ise [-3,+3] aralıđında retilmiřtir. Farklı aralıklarda madde ve yetenek parametreleri belirlenerek yetenek dzeyinin alıřma sonularını nasıl etkilediđi arařtırılabilir.

KAYNAKÇA

- Aiken, L. R. (2000). *Psychological Testing and Assessment*. Boston: Allyn and Bacon.
- Altun, A. (2013). *Dikey ölçeklemede madde tepki kuramına dayalı farklı kalibrasyon ve yetenek kestirim yöntemlerinin karşılaştırılması*. (Yayınlanmamış Doktora Tezi), Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.
- Angoff, W. H. (1971). Scales, Norms, and Equivalent Scores. In Thorndike, R.L. (ed.) *Educational Measurement*, 508-600. American Council on Education: Washington D.C.
- Ansley, N. A., & Forsyth, R. A. (1985). An examination of the characteristics of unidimensional IRT parameter estimates derived from two-dimensional data. *Applied Psychological Measurement*, 9, 37-48.
- Baker, F. (2001). *The Basics of Item Response Theory*. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, MD.
- Beard, J. J. (2008). *An Investigation of vertical scaling with item response theory using a multistage testing framework*. Unpublished Doctoral Thesis. University of Iowa, Iowa.
- Béguin, A. A., & Hanson, B. A. (2001). *Effect of noncompensatory multidimensionality on separate and concurrent estimation in IRT observed score equating*. Paper presented at the annual meeting of the National Council on Measurement in Education, Seattle, WA.
- Béguin, A. A., Hanson, B. A., & Glas, C. A. W. (2000). *Effect of multidimensionality on separate and concurrent estimation in IRT equating*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.
- Boughton, K. A., Lorie, W. & Yao, L. (2005). A multidimensional multigrup irt models for vertical scales with complex test structure: An emprical evaluation of student growth using real data. National Council on Measurement in Education: 2005. Monreal/ Quebec/ Canada.
- Briggs, D. C., Weeks, J. P. & Wiley, E. (2008). *Vertical scaling in value-added models for student learning*. Paper presented at the National Conference on Value- Added Modeling, April 22-24, 2008, Madison, WI.
- Cao, L. (2008). *Mixed format test equating: Effects of test dimensionality and common-item sets*. Unpublished dissertasion, University of Maryland.
- Cao, Y., Li, D.Y., & Hendrickson, A.B. (2007). *Maintaining vertical scale under the common-item design: A simulation Study*. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, IL.

- Cohen, L., Manion, L. & Morrison, K. (2007). *Research methods in education* (6th Edition). Routledge, Canada.
- Cook, L. L. & Eignor, D. R. (1983). Practical considerations regarding the use of item response theory to equate tests. In Hambleton, R. K. (Ed.) *Applications of item response theory* (pp.175–195). Vancouver, British Columbia: Educational Research Institute of British Columbia.
- Cook, L. L. & Eignor, D. R. (1991). An NCME instructional module on IRT equating methods. *Educational measurement: Issues and Practice*. 10 (3), 37-45.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart & Winston.
- Çepni, S., Ayvaci, Ş. H., ve Bacanak, A. (2004). *Fen Teknoloji Toplum*. Top-Kar Matbaacılık, s:40, Trabzon.
- Çetin, E. (2009). Dikey ölçekteleme klasik test ve madde tepki kuramına dayalı yöntemlerin karşılaştırılması. (Yayınlanmamış Doktora Tezi), Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Çilenti, K. (1979). *Eğitim Teknolojisi: Kuramlar-Amaç ve Yöntemler-Merkezler*. Kadioğlu Matbaası, Ankara.
- De Gruijter, D. N. M., & Van der Kamp, L. J. Th. (2005). *Statistical Test Theory for Education and Psychology*.
- Donlon, T. (1984). *The college board technical handbook for scholastic Aptitude Test and Achievement Tests*. New York: College Entrance Examination Board.
- Dongyang, L. (2009). *Developing a common scale for testlet model parameter estimates under the common-item nonequivalent groups design*. Unpublished doctorate thesis, University of Maryland.
- Dorans, N. J. (1990). Equating methods and sampling designs. *Applied Measurement in Education*, 3(1), 3-17.
- Felan, G. D. (2002). *Test equating: mean, linear, equapercentile and item response theory*. Paper presented at the annual meeting of the Southwest Educational Research Association, Austin.
- Fidan, N. ve Baykul, Y. (1993). *İlkokul ve ilköğretim okullarında temel öğrenme ihtiyaçlarının karşılanması*. MEB Yayınları, s: 32-33, Ankara.
- Gelbal, Selahattin (1994). p madde güçlük indeksi ile rash modelinin b parametresi ve bunlara dayalı yetenek ölçüleri üzerine bir karşılaştırma (Yayınlanmamış Doktora Tezi), Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Gök, B. (2012). Denk olmayan gruplarda ortak madde deseni kullanılarak madde tepki kuramına dayalı eşitleme yöntemlerinin karşılaştırılması (Yayınlanmamış Doktora Tezi), Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Haebara, T. (1980). Equating logistic ability scales by a weighted least squares method. *Japanese Psychological Research*, 22, 144-149.

- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: principles and applications*. Boston: Kluwer.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Hanson, B. A., & Béguin, A. A. (2002). Obtaining a common scale for item response theory item parameters using separate versus concurrent estimation in the common-item equating design. *Applied Psychological Measurement*, 26, 3-24.
- Hanson, B. A., Zeng, L., & Chien, Y. (2004). *ST: A Computer Program for IRT Scale Transformation* [Computer software]. Retrieved January 24, 2005, from <http://www.education.uiowa.edu/casma>
- Harris, D. J. (2003). Equating the multistate bar examination. *The Bar Examiner*, 72(3), 12-18.
- Harris, D. J. (2007). Practical issues in vertical scaling. In N.J. Dorans, M. Pommerich, & P. M. Holland (Eds.), *Linking and aligning scores and scales*. (pp. 233-251) New York, NY: Springer Science+Business Media, LLC.
- Harris, D. J., Hendrickson, A. B., Tong, Y., Shin, S. H., & Shyu, C. Y. (2004). *Vertical scales and the measurement of growth*. Paper presented at the National Council of Measurement in Education, San Diego, CA
- Holland, P. W. (2002). Two measures of changes in the gaps between the CDFs of test-score distributions. *Journal of Educational and Behavioral Statistics*, 27, 3-17.
- Holland, P. W. & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), *Educational measurement* (pp. 187–220). Westport, CT: Praeger Publishers
- Hoover, H. D. (1984a). The most appropriate scores for measuring educational development in the elementary schools: GE's. *Educational Measurement: Issues and Practice*, 3(4), 8-14.
- Hoover, H. D. (1984b). Rejoinder to Burket. *Educational Measurement: Issues and Practice*, 3(4), 16-18.
- Hoskens, M., D. M. Lewis, and R. J. Patz. April 2003. *Maintaining vertical scales using a common item design*. Paper in D. Lewis (organizer): *Issues in Vertical Scaling under Common Item Designs*, a symposium at the 2003 annual meeting of the National Council on Measurement in Education, Chicago.
- Ito, K., Sykes, R. C. & Yao, L. (2008). Concurrent and separate grade-groups linking procedures for vertical scaling. *Applied Measurement in Education*, 21, 187-206.
- Kang, T. & Petersen, N.S. (2009). *Linking item parameters to a base scale*. ATC Research Report Series 2009-2.
- Karkee, T. B. & Wright K. R. (2004). *Evaluation of linking methods for placing three-parameter logistic item parameter estimates onto a one-parameter scale*. Paper presented at the Annual Meeting of the American Educational Research Association in San Diego, California, April 16, 2004.

- Kelecioğlu, Hülya. (1994). Öğrenci seçme sınavı puanlarının eşitlenmesi üzerine bir çalışma (Yayınlanmamış Doktora Tezi), Ankara: Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü.
- Keller III, R. R. (2007). *A comparison of item response theory true score equating and item response theory-based local equating*. Unpublished dissertation, University of Massachusetts.
- Kim, J. (2007). *A comparison of calibration methods and proficiency estimators for creating IRT vertical scales*. Unpublished Doctoral Thesis, University of Iowa.
- Kim, S. H., & Cohen, A. S. (1998). A comparison of linking and concurrent calibration under item response theory. *Applied Psychological Measurement, 22*, 131-143.
- Kim, S., & Kolen, M. J. (2006). Robustness to format effects of IRT linking methods for mixed-format tests. *Applied Measurement in Education, 19*(4), 357-381.
- Kim, S., & Lee, W. (2006). An extension of four IRT linking methods for mixed-format tests. *Journal of Educational Measurement, 43*(1), 53-76.
- Kim, J., Lee, W.C., Kim, D. & Kelley, K. (2009). Investigation of Vertical Scaling Using the Rasch Model. National Council on Measurement in Education: April 2009.
- Kolen, M. J. (1981). Comparison of traditional and item response theory methods for equating tests. *Journal of Educational Measurement, 18*, 1-11.
- Kolen, M. J. (1988). Traditional equating methodology. *Educational Measurement: Issues and Practice, 7*(4), 29-36.
- Kolen, M. J. (2007). *Data collection designs and linking procedures*. In Dorans, N. J.; Pommerich, M. Holland, P. W. (Eds.), *Linking and Aligning Scores and Scales*. USA: Springer.
- Kolen, M. J., & Brennan, R. L. (1995). *Test Equating: methods and practices*. New York: Springer-Verlag.
- Kolen, M. J & Brennan, R. L. (2004). *Test equating, scaling, and linking: methods and practices (2nd edn)* (New York, Springer Verlag).
- Lee, W. C. & Ban, J. C. (2010). A Comparison of IRT linking procedures. *Applied Measurement in Education, 23*, 23-48
- Lei, P. W., Dunbar, S. B., & Kolen, M. J. (2004). A comparison of parametric and nonparametric approaches to item analysis for multiple-choice tests. *Educational and Psychological Measurement, 64*, 565-587.
- Livingston, S. (2004) *Equating Test Scores (without IRT)*. New Jersey: Educational Testing Service.
- Lohman, D. F. & Hagen, E. (2003). *Interpretive guide for teachers and counselors: Cognitive Abilities Test Form- 6 All Levels*. Itasca, Illinois: Riverside Publishing.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, N. J.: Lawrence Erlbaum.

- Lord, F. M. (1986). Maximum likelihood and Bayesian parameter estimation in item response theory. *Journal of Educational Measurement*, 23, 157-162.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison Wesley.
- Loyd, B. H., & Hoover, H. D. (1980). Vertical equating using the Rasch model. *Journal of Educational Measurement*, 17, 179-193.
- Macro, G. L. (1977). Item characteristic curve solutions to three intractable testing problems. *Journal of Educational Measurement*, 14(2), 139- 160
- Marco, G. L., Petersen, N.S. & Stewart, E.E. (1979). *A test of adequacy of curvilinear score equating models*. Paper presented at the 1979 Computerized Adaptive Testing Conference, Minneapolis.
- McBridge, J., & Wise, L. (2001) Developing the Vertical Scale for the Florida Comprehensive Assessment Test (FCAT). A Harcourt Educational Measurement, San Antonio, Texas.
- MEB (2005). İlköğretim Fen ve Teknoloji Dersi Öğretim Programı ve Kılavuzu (4.-5. Sınıflar) Ankara.
- Meng, H (2007). A comparison study of IRT calibration methods for mixed-format tests in vertical scaling. Unpublished Ph.D. Thesis, University of Iowa, Iowa.
- Meng, H., Kolen, M. J. & Lohman, D. (2006). *An empirical investigation of IRT scaling methods: How different IRT models, parameter estimation procedures, proficiency estimation methods, and estimation programs affect the results of vertical scaling for the Cognitive Abilities Test*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.
- Nandakumar, R. (1994). Assessing dimensionality of a set of item responses: Comparison of different approaches. *Journal of Educational Measurement*, 31(1), 17-35.
- Norman-Dvorak, R. L. (2009). *A comparison of kernel equating to the test characteristic curve method*. Unpublished Doctoral Thesis, University of Nebraska, Lincoln, Nebraska.
- Petersen, N.S., Kolen, M.J., & Hoover, H.D. (1989). Scaling, norming, and equating. In R.L. Linn (Ed.), *Educational Measurement* (3rd Ed., pp. 221-261). New York: American Council on Educational and Macmillan.
- Sinharay, S. & Holland, P. W. (2007). Is it necessary to make anchor tests mini versions of the tests being equated or can some restrictions be relaxed? *Journal of Educational Measurement*, 44(3), 249-275.
- Skaggs, G. & Lissitz, R. W. (1982). *Test equating, relevant issues and a review of recent research*. Paper presented at the annual meeting of the American Educational Research Association. Los Angeles, CA. Eric document ED 241 540.
- Skaggs, G. & Lissitz, R. W. (1986). IRT Test equating: relevant issues and a review of recent research. *Review of Educational Research*, 56(4), 495-529.

- Skaggs, G., & Lissitz, R. W. (1988). Effect of examinee ability on test equating invariance. *Applied Psychological Measurement*, 12(1), 69-82
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23-74.
- Slinde, J.A & Linn, R.L. (1978). An exploration of the adequacy of the Rasch Model for the problem of vertical equating. *Journal of Educational Measurement*, 5, 245-262.
- Smith, R. M. & Kramer G. A. (1992). A comparison of two methods of test Equating in the Rasch model. *Educational and Psychological Measurement*, 52(4), 835-846.
- Stocking, M. L. & Lord, F. M. (1983). Developing a common metric in item response theory. *Applied Psychological Measurement*, 7, 201-210.
- Thissen, D., & Orlando, M. (2001). Item response theory for items scored in two categories. In D. Thissen & W. Howard (Ed.), *Test Scoring* (pp.73-140). Mahwah, NJ: Lawrence Erlbaum Associates.
- Tong, T (2005). Comparison of methodologies and results in vertical scaling for educational achievements tests. Unpublished Doctoral Thesis, University of Iowa, Iowa.
- Tong, Y. & Kolen M. (2007). Comparison of methodologies and results in vertical scaling for educational achievement tests. *Applied Measurement in Education*, 20(2), 227-253
- Tong, Y. & Kolen, M. (2008). Maintenance of Vertical Scales. National Council on Measurement in Education: March 2008. New York City.
- Tong, Y. & Kolen, M. (2010) Scaling: An ITEMS Module. *Educational Measurement: Issues and Practice*, 29(4), 39-48
- von Davier, A. A., Holland, P. W. & Thayer, D. T. (2004). *The kernel method of test equating*. New York: Springer.
- von Davier, A. A. & Wilson, C. (2008). Investigating the population sensitivity assumption of item response theory true-score equating across two subgroups of examinees and two test formats. *Applied Psychological Measurement*, 32(1), 11-26.
- Wang, X. (2013). *Linking across forms in vertical scaling under the common-item nonequivalent groups design*. Unpublished Doctoral Dissertation, University of Iowa.
- Wright, B.D. & M.H. Stone (1979). Best test design. Chicago, IL: MESA Press.
- Wyse1, A. E.& Reckase, M. D. (2011). A graphical approach to evaluating equating using test characteristic curves. *Applied Psychological Measurement*, 35(3) 217–234.
- Yen, W. M. (1984). Obtaining maximum likelihood trait estimates from number-correct scores for the three-parameter logistic model. *Journal of Educational Measurement*, 21, 93-111.

- Zhao, Y. (2008). *Approaches for addressing the fit of item response theory models to educational test data*. Unpublished Doctoral Dissertasion, University of Massachusett.
- Zhu, W. (1998). Test equating: What, why, who?. *Research Quarterly for Exercise and Sport*, 69(1), 11–23.

EKLER DİZİNİ

EK-1

ETİK KURUL İZİN MUAFİYET FORMU

6. SINIF FEN BİLİMLERİ DERSİ SEVİYE BELİRLEME TESTİ

(1)

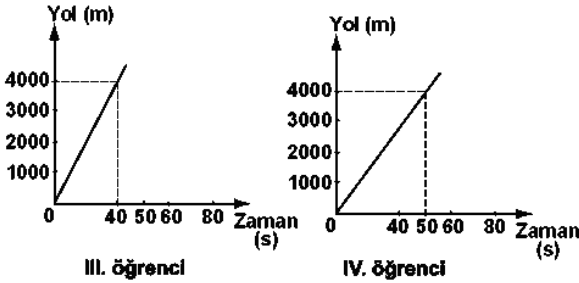
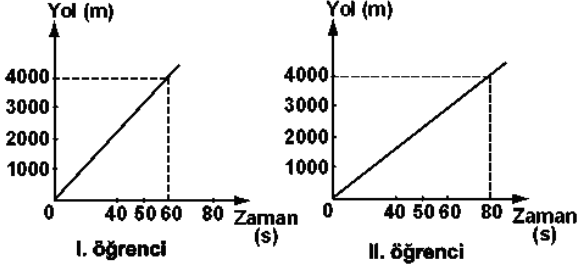


Emel, şekildeki 75 m'lik yolu izleyerek 15 saniye-
de K noktasından L noktasına sabit süratle
geliyor. Buna göre Emel'in sürati kaç m/s'dir?

- A) 3 B) 5 C) 8 D) 10

(2)

Hafta sonu İstanbul Park'ta yapılan Formu-
la 1 yarışlarını izlemeye giden dört arkada-
ş, ellerindeki kronometrelerle favori pilotları-
nın pistin bir turunu kaç saniyede tamamladıklarını ölçüyor. Pist mesafesinin
4000 m olduğunu bilen öğrenciler, favori pi-
lotları için aşağıdaki yol-zaman grafiklerini
çiziyorlar.



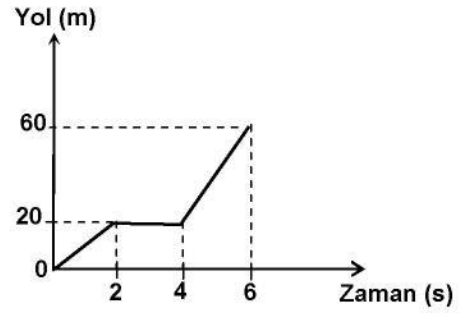
Bu grafiklere göre, hangi öğrencinin favori
pilotu en süratlidir?

- A) I B) II C) III D) IV

(3)



Bir çitanın hareketini inceleyen gözlemci, çitaya
ait aşağıdaki yol - zaman grafiğini elde ediyor.



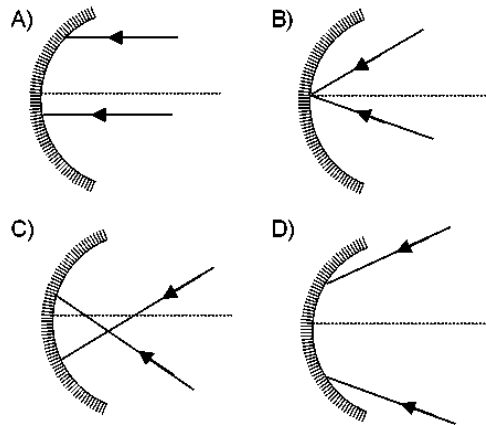
Bu grafiğe göre, aşağıdaki ifadelerden hangileri
doğrudur?

- I. 0 - 2 saniye arasında sürati 10 m/s'dir.
II. 2 - 4 saniye arasında hareket enerjisi vardır.
III. 4 - 6 saniye arasında sürati 30 m/s'dir.

- A) Yalnız I B) I - III
C) II - III D) I - II - III

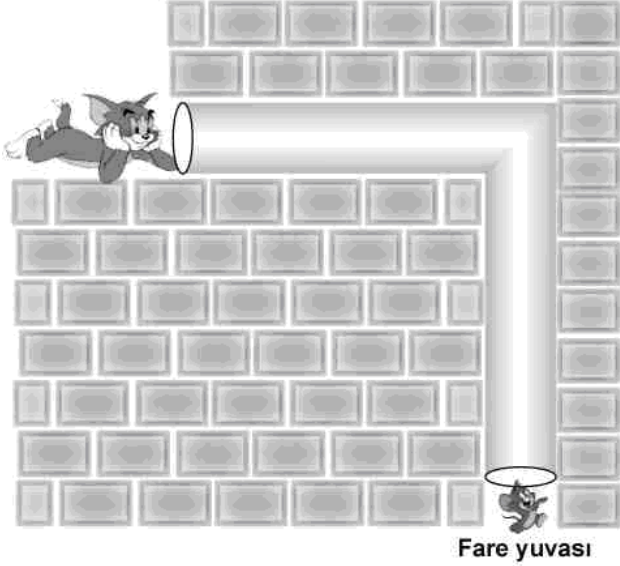
(4)

Arda, bir çukur aynanın odak noktasını bulmak
istiyor. Bunun için çukur aynaya iki ışık ışını
gönderiyor. Bu ışınları aşağıdakilerden hangi-
sindeki gibi gönderirse amacına ulaşabilir?

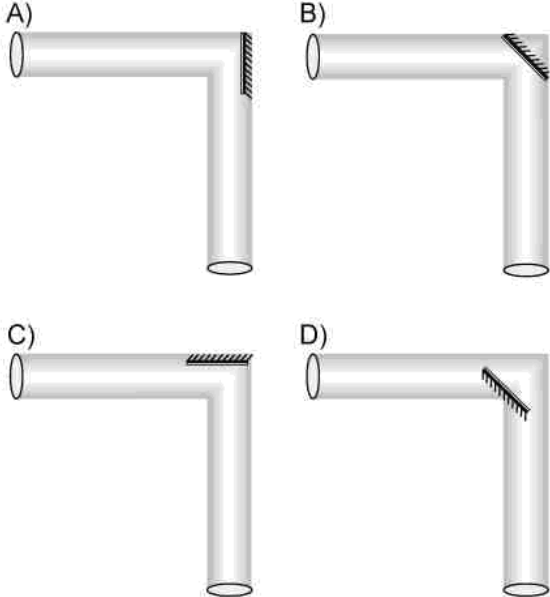


(5)

Şeffaf olmayan bir borunun içine şekildeki gibi bakan kedi, yuvasındaki fareyi görememektedir.

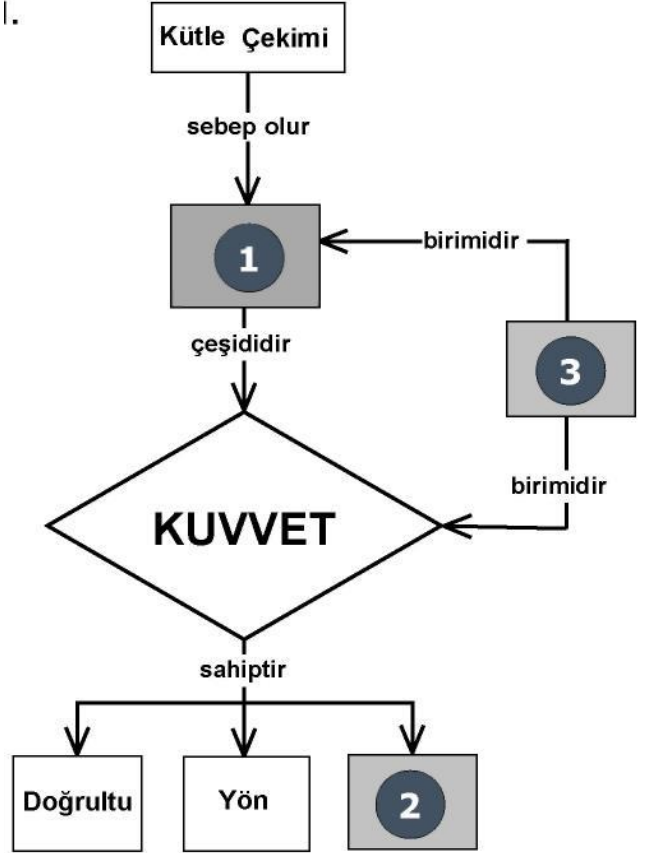


Borunun içine düzlem ayna aşağıdakilerden hangisindeki gibi yerleştirilirse kedi fareyi görebilir?



(6)

1.



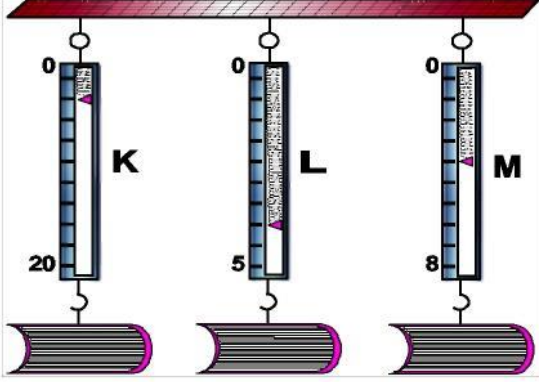
Öğretmen, “kuvvet” konusu ile ilgili yukarıdaki kavram haritasını, numaralandırılmış kısımlarını boş bırakarak tahtaya çiziyor. 1, 2 ve 3 kutucuklarına yazılacak kavramlar için söz alan dört öğrenci aşağıdaki seçenekleri oluşturuyor.

Buna göre, hangi seçenekte bu kavramlar doğru olarak verilmiştir?

	1	2	3
A)	Yerçekimi	Kütle	Kilogram
B)	Kütle	Büyüklik	Dinamometre
C)	Ağırlık	Dinamometre	Newton
D)	Ağırlık	Büyüklik	Newton

(7)

Zeynep, Fen ve Teknoloji ders kitabının ağırlığını, K, L ve M dinamometreleri ile ölçtüğünde dinamometreleri şekildeki gibi görüyor.

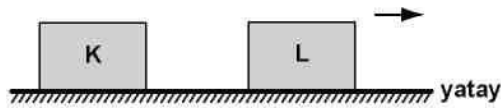


Bu etkinliğe göre, Zeynep'in aşağıda çıkardığı sonuçlardan hangileri doğrudur?

- I. K dinamometresi ile daha ağır kitapları ölçebilirim.
- II. Daha hassas ölçüm yapmak istediğimde L dinamometresini kullanmalıyım.
- III. Üç tane Fen ve Teknoloji ders kitabının ağırlığını M dinamometresi ile ölçebilirim.

- A) Yalnız I B) I - II
C) II - III D) I - II - III

(8)



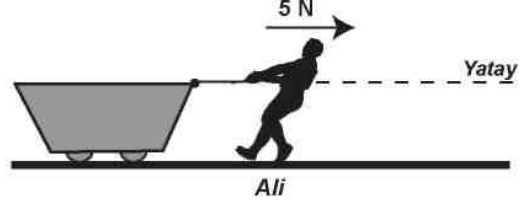
Şekildeki K cismi durmaktadır. L cismi ise belirtilen yönde sabit süratle hareket etmektedir.

Buna göre, K ve L'ye etki eden kuvvetlerle ilgili olarak aşağıdakilerden hangisi doğrudur?

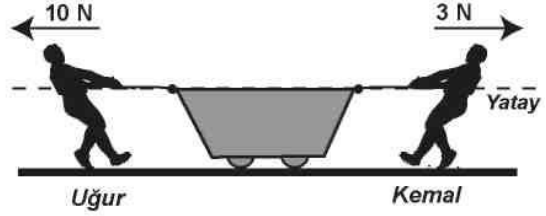
- | K | L |
|------------------|---------------|
| A) Dengelenmiş | Dengelenmiş |
| B) Dengelenmiş | Dengelenmemiş |
| C) Dengelenmemiş | Dengelenmiş |
| D) Dengelenmemiş | Dengelenmemiş |

(9)

Ali'nin yatay yolda 5 N'luk kuvvetle çektiği vagon sabit süratle ilerleyebilmektedir.



Vagon aynı yolda durmaktayken Uğur ve Kemal tarafından aynı anda şekildeki gibi çekiliyor.



Vagonun bundan sonraki hareketi için aşağıdakilerden hangisi doğru olur?

- A) Hareket etmez.
B) Kemal'in çektiği yönde hızlanır.
C) Uğur'un çektiği yönde hızlanır.
D) Uğur'un çektiği yönde sabit süratle hareket eder.

(10)

Sesin yansımaları ve yankı olayları, yaşamımızda önemli bir yer tutar. Bunlara birçok örnek verilebilir.

- I- Doktorların ultrason cihazı kullanarak hastalık teşhisi yapmaları,
- II- Gemilerin sonar cihazı kullanarak yön ve derinlik tespiti yapmaları,
- III- Tiyatro salonlarının duvar ve tavanlarının yumuşak ve pürüzlü yüzey ile kaplanması.

Yukarıda verilen örneklerden hangileri, bilim ve teknolojiye sesin yansımalarından yararlandığını gösterir?

- A) Yalnız I B) I ve II
C) II ve III D) I, II ve III

(11)

Bir öğretmen, Fen ve Teknoloji dersinde, öğrencilerden sesin yansımaya örnekler vermelerini istiyor. Öğrenciler aşağıdaki örnekleri veriyorlar:

Kemal: "Doktorların, ultrason cihazı kullanarak iç organlarımızı görmeleri"

Ahmet: "Balıkçıların, sonar cihazı kullanarak balık avlamaları"

Sema: "Kar yağdığında, sokağımızın diğer günlerden daha sessiz olması"

Özlem: "Derin bir vadi kenarında bağırduğumuzda, sesimizin yankı yapması"

Buna göre, hangi öğrencilerin verdiği örnekler sesin yansımaya ile ilgili olabilir?

- A) Kemal, Ahmet ve Sema
- B) Ahmet, Sema ve Özlem
- C) Kemal, Sema ve Özlem
- D) Kemal, Ahmet ve Özlem

(12)

Sesin yansımaya ve yankı olayları, yaşamımızda önemli bir yer tutar. Bunlara birçok örnek verilebilir.

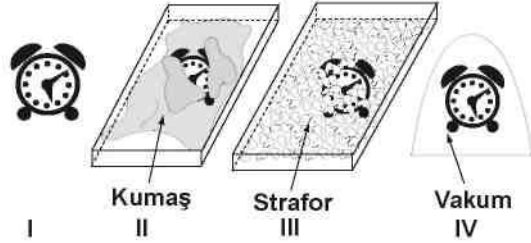
- I- Doktorların ultrason cihazı kullanarak hastalık teşhisi yapmaları,
- II- Gemilerin sonar cihazı kullanarak yön ve derinlik tespiti yapmaları,
- III- Tiyatro salonlarının duvar ve tavanlarının yumuşak ve pürüzlü yüzey ile kaplanması.

Yukarıda verilen örneklerden hangileri, bilim ve teknolojiye sesin yansımayandan yararlandığını gösterir?

- A) Yalnız I
- B) I ve II
- C) II ve III
- D) I, II ve III

(13)

Öğretmeni, Neşe'den "sesin şiddetinin soğurulma ile azaldığını" gösteren bir deney tasarlamasını istiyor.

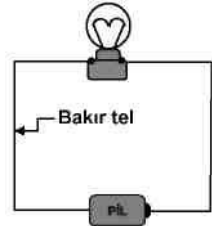


Neşe, çalar saat kullanarak hazırladığı düzeneklerden hangi ikisiyle amacına ulaşmış olur?

- A) I - III
- B) I - IV
- C) II - III
- D) II - IV

(14)

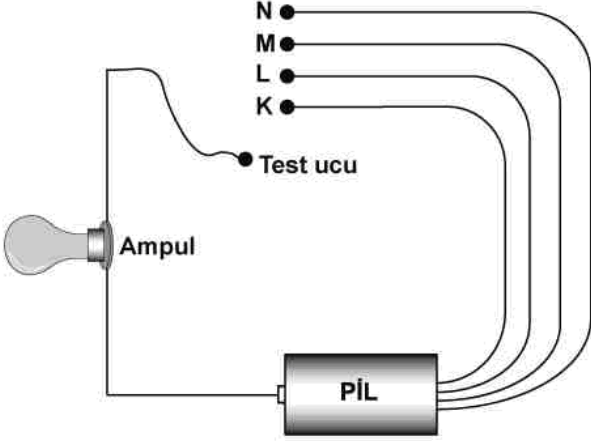
Ali, şekildeki elektrik devresinde ampulün parlaklığını artırmak istiyor. Bunu gerçekleştirebilmek için aşağıdakilerden hangisini yapmalıdır?



- A) Ampulü pilin hemen yanına bağlamalıdır.
- B) Bakır tel yerine aynı uzunluk ve kalınlıkta gümüş tel kullanmalıdır.
- C) Uzunluğu aynı, daha ince bakır tel kullanmalıdır.
- D) Kalınlığı aynı, daha uzun bakır tel kullanmalıdır.

(15)

Arif şekildeki devreyi kuruyor ve boşta kalan test ucunu sırasıyla K, L, M ve N maddelerinden yapılmış kablolara dokunduruyor.



Sonra ampulün ışık verip vermediğini aşağıdaki tabloya kaydediyor.

	Ampul ışık veriyor	Ampul ışık vermiyor
K	√	
L	√	
M		√
N		√

Buna göre, K, L, M ve N maddeleri aşağıdakilerden hangisi olabilir?

- | K | L | M | N |
|--------------|-----------|-----------|---------|
| A) Bakır | Alüminyum | Plastik | Cam |
| B) Alüminyum | Plastik | Cam | Bakır |
| C) Cam | Alüminyum | Bakır | Plastik |
| D) Plastik | Cam | Alüminyum | Bakır |

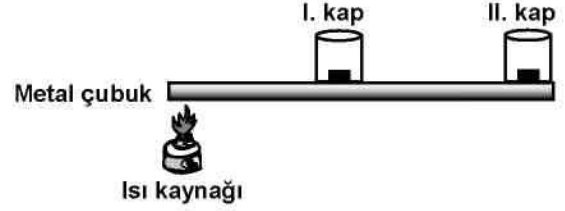
(16)

Mete'nin bugün doğum günü. O, bir yaşında olduğu için ancak elinden tutulunca yürüyebiliyor. Yedinci dişi yeni çıktı.

Sağlıklı ve normal bir gelişim gösteren Mete, iki yıl sonraki doğum gününde aşağıdakilerden hangisini yapabilir?

- A) Merdivenleri kendisi rahat inip çıkabilir.
- B) Sadece sıvı besinler yiyebilir.
- C) Sadece baba, anne, mama sözcüklerini söyleyebilir.
- D) Ayakkabısının bağcıklarını kolayca bağlayabilir.

(17)



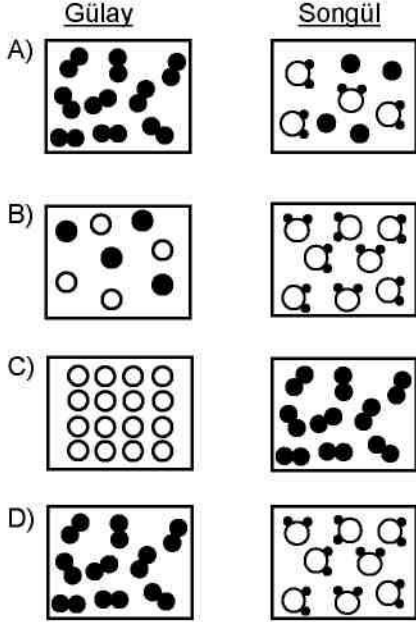
İçinde eşit miktarda aynı katı yağ bulunan özdeş kaplar şekildeki gibi metal çubuk üzerine yerleştiriliyor. Metal çubuk bir ucundan ısıtılmaya başladığında önce I. kaptaki, sonra II. kaptaki yağ eriyor. Bu deney aşağıdakilerden hangisini ispatlamak için yapılmıştır?

- A) Erime süresinin madde miktarına bağlı olduğunu
- B) Katı maddelerde ısının iletim yoluyla yayıldığını
- C) Farklı maddelerin ısıyı farklı hızlarda iletmediğini
- D) Kaplardan hangisinin ısıyı daha iyi iletmediğini

(18)



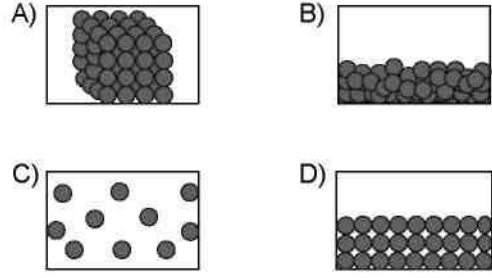
Gülay ve Songül'ün tanecik modelleri aşağıdakilerden hangisi olabilir?



(19)

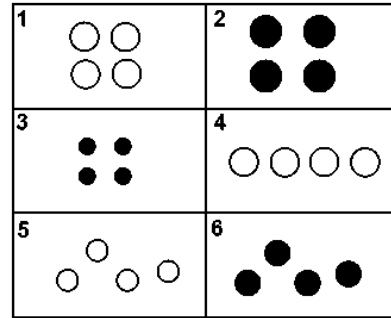
Saf bir maddenin;

- Akma özelliği gösterdiği,
- Taneciklerinin bağımsız hareket ettiği bilindiğine göre, maddenin hâline ait tanecik modeli aşağıdakilerden hangisidir?



(20)

Aşağıda numaralandırılmış kutucuklarda bazı elementlerin tanecik modelleri verilmiştir. Hangi kutucuklardaki tanecik modelleri aynı elementi temsil eder?



- A) 1 ve 4
B) 1, 4 ve 5
C) 5 ve 6
D) 2, 3 ve 6

(21)

Aynı miktarda alınan X ve Y maddeleri özdeş pistonlu kaplara konuluyor. Bu maddelere eşit kuvvet uygulandığında Y maddesinin sıkıştırılabildiği, X'in ise sıkıştırılmadığı gözleniyor.

Bu gözleme bağlı olarak X ve Y maddelerini oluşturan tanecikler ile ilgili aşağıdakilerden hangisi söylenebilir?

- A) X maddesi bütünsel, Y maddesi ise tanecikli yapıya sahiptir.
- B) Y maddesini oluşturan tanecikler arasındaki boşluk daha fazladır.
- C) X maddesini oluşturan taneciklerin arasında boşluk yoktur.
- D) X ve Y maddelerini oluşturan tanecikler arasındaki boşluk aynıdır.

(22)

Öğrenciler, yaptıkları etkinliklerin sonuçlarını aşağıdaki gibi ifade ediyor.

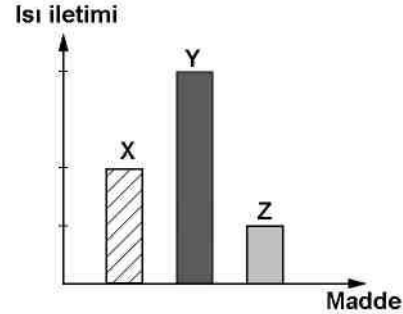


Hangi öğrencinin yaptığı etkinlik sonucuna göre, "Sıvıları oluşturan moleküller birbirine çok yakındır." bilgisine ulaşılabilir?

- A) Oğuz
- B) Şeyda
- C) Ferda
- D) Cemil

(23)

X, Y, Z katı maddelerinin ısı iletimi, grafikteki gibidir:



Bu maddelerin günlük yaşamdaki kullanımlarıyla ilgili olarak aşağıdakilerden hangisi yanlıştır?

- A) Binaların dış cephe kaplamasında Z maddesi tercih edilmelidir.
- B) Ütü tabanı yapılırken Y maddesi kullanılmıdır.
- C) Yemek pişirirken karıştırmada kullanılacak kaşık yapımı için en uygun, X maddesidir.
- D) Kışlık ayakkabı yapımında Z maddesi tercih edilmelidir.

(24)

Fiziksel değişim	Kimyasal değişim
① Cevizleri kırdım.	② Peyniri dillimledim.
③ Ekmeği kızarttım.	④ Yumurtayı pişirdim.

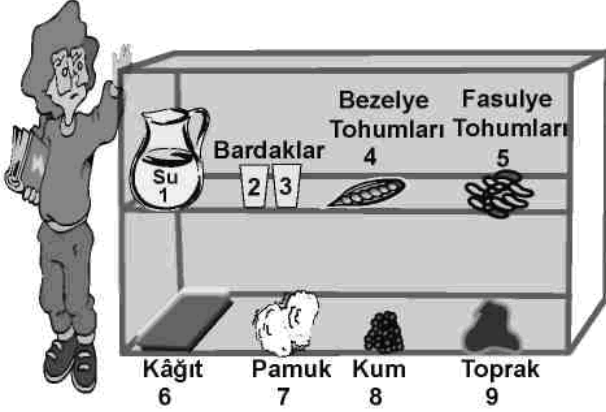
Kahvaltayı hazırlarken yaptığı işlemleri çizelgeye kaydeden Zeynep, hangi işlemleri yanlış yere yazmıştır?

- A) 1 ve 4
- B) 2 ve 3
- C) 3 ve 4
- D) 1 ve 2

(25)

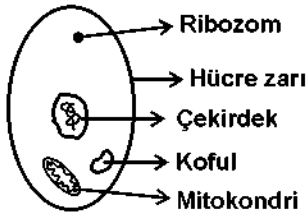
Bir öğrenci nemli topraktaki fasulye tohumlarının, nemli pamuk arasındaki fasulye tohumlarından daha önce çimleneceğini düşünüyor.

Bu durumu deneyle gözlemek için aşağıdaki dolaptan kaç numaralı araç ve gereçleri seçmelidir?



- A) 2, 4, 8
B) 3, 5, 6, 9
C) 1, 4, 5, 6, 7
D) 1, 2, 3, 5, 7, 9

(26)







Yukarıda verilen hücrenin kısımlarını inceleyen Mehmet, bu hücreye sahip olan canlı ile ilgili olarak hangi yorumu yapamaz?

- A) Kendi besinini yaptığı
B) Solunum yaptığı
C) Kalıtsal maddeye sahip olduğu
D) Dış ortamla madde alışverişinin olduğu

(27)

Öğretmen Müge'den, kendi görevlerini anlatan aşağıdaki hücre organellerini hayvan hücresi şeklinin içine çizmesini istiyor. Müge hangi organeli çizerse hata yapmış olur?

- A)  : Fotosentezle besin üretim.
B)  : Hücre içi sindirimi yaparım.
C)  : Hücre için önemliyim, bulunduğum hücreye enerji üretirim.
D)  : Küçük bir organelim ancak hücrede protein sentezi yaparım.

(28)

Mikrop kaptdığın için kanında akyuvar sayısı 16.000'e çıkmış. Sağlıklı bir insanın 1 mm³ kanındaki akyuvar sayısı genellikle 10.000'den fazla olmaz.

Yazdığım ilaçları kullandıktan sonra yine kan tahlili yaptırıp sonucu getir.



Adı Soyadı:
Emel Kaya
Akyuvar Sayısı:
16.000

İlaç kullanmadan önceki kan tahlili

Adı Soyadı:
Emel Kaya
Akyuvar Sayısı:
8.000

İlaç kullandıktan sonraki kan tahlili

İlaçları kullandıktan sonraki kan tahlili sonucunda da görüldüğü gibi akyuvar sayısı 8.000'e inen Emel'e doktor ne söylemiştir?

- A) Hastalığın devam ediyor.
B) Bu ilaçları sürekli kullanmalısın.
C) Başka ilaçlar kullanmalısın.
D) İlaçlar etkili olmuş, iyileşmişsin.

(29)

Mine, Sinan ve Hülya, su çiçeği hastalığı geçirmekte olan arkadaşlarını ziyaret etmek istiyor.

Öğretmen: Bulaşıcı olan bu hastalığa karşı bağışıklık kazanmışsanız gidebilirsiniz. Tablodaki soruları cevaplayarak bağışıklık kazanıp kazanmadığınızı öğrenin.

Sorular İsimler	Su çiçeği hastalığı aşısı oldunuz mu?	Su çiçeği hastalığı geçirdiniz mi?
Mine	Hayır	Hayır
Sinan	Hayır	Evet
Hülya	Evet	Hayır





Tablodaki cevaplara göre öğretmen, kimlerin hastalanan arkadaşlarını ziyaret etmesine izin vermiştir?

- A) Mine ve Sinan B) Sinan ve Hülya
C) Yalnızca Mine D) Yalnızca Hülya

(30)

Sorular Hayvan	Suda yaşar mı?	Başkalaşım geçirir mi?	Embriyo gelişimi anne karnında mı olur?
?	Evet	Hayır	Hayır

Selma evinde beslediği hayvanla ilgili olarak tablodaki sorulara cevap veriyor. Buna göre, Selma'nın beslediği hayvan hangisidir?

- A)  B) 
C)  D) 

(31)

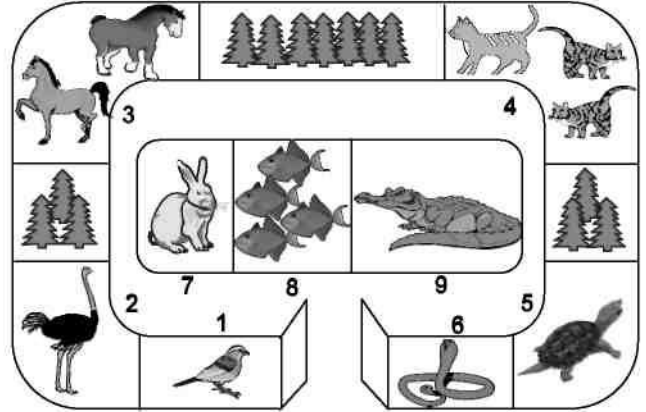


Öğrenci, bir kurbağa türüne ait resmi şekildeki gibi kartona yapıştırıp yanına bazı özelliklerini yazmıştır.

Öğrenci, kurbağayı çıkarıp hamsi balığı resmini yapıştırırsa kaç numaralı özelliği değiştirmesi gerekir?

- A) 1 B) 2 C) 3 D) 4

(32)

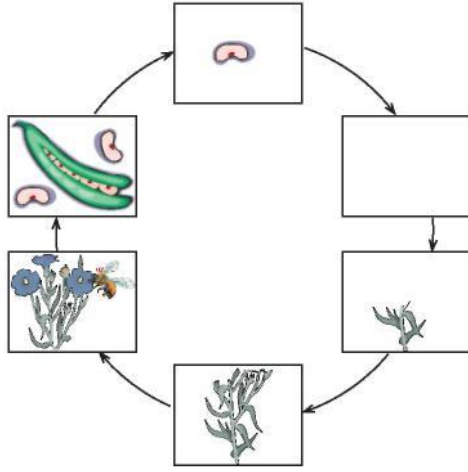


Şekildeki gibi bir hayvanat bahçesine gezi yapan öğrenciler, sadece "yavrusunu sütle besleyen hayvanlar"ı görmek istiyor.

Buna göre öğrenciler kaç numaralı bölümleri ziyaret etmelidir?

- A) 2, 7, 9 B) 3, 4, 7
C) 1, 5, 6, 8 D) 2, 4, 5, 9

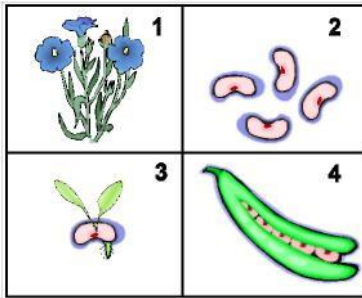
(33)



Fasulyenin hayat döngüsü

Emel, fasulyenin hayat döngüsünün aşamalarını gösteren resimleri kâğıtlara yapıştırdığında, bir resmin eksik kaldığını görüyor.

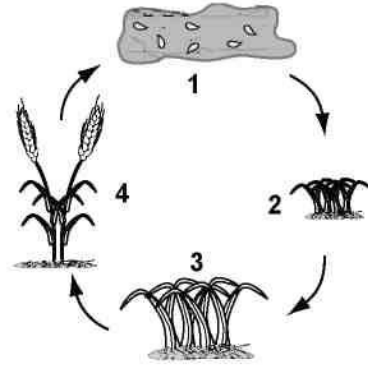
Buna göre, Emel eksik kalan resmin yerine aşağıdaki tablodan kaç numaralı resmi seçmelidir?



Tablo

- A) 1 B) 2 C) 3 D) 4

(34)



Buğdayın Hayat Döngüsü

Ahmet : Dedeciğim, tarladaki bu küçük yapraklı bitkilerin adı ne?

Dede : Bunlar buğday bitkileri... Tohumlarını sonbaharda ektik. Yaz tatilinde köye geldiğinde bunları büyümüş ve üzerinde tohumları oluşmuş göreceksin.

Ahmet, yazın köye gidince, buğdayın yukarıdaki hayat döngüsünde kaç numaralı basamağı görecektir?

- A) 1 B) 2 C) 3 D) 4

(35)

Fosillerden yararlanılarak;

- I- Kayaçların hangi zamanda oluştuğu
- II- Geçmişteki coğrafi ve iklim koşulları
- III- Soyları tükenmiş bitki ve hayvanlar

ile ilgili verilenlerden hangileri hakkında bilgi edinilebilir?

- A) I ve II B) I ve III
C) II ve III D) I, II ve III

(36)

Sivilce çıkması	Sperm oluşması	Kasların gelişmesi
Bıyık çıkması	Kilo artması	Adet görme
Ter salgısının artması	Yumurta oluşması	

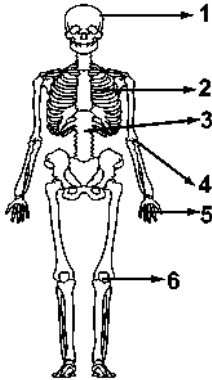
Şemada, ergenlik dönemindeki kızlarda ve erkeklerde görülen özellikler ile her ikisinde görülen ortak özellikler farklı taranmıştır.

Tarama şekline de dikkat edilerek boş bırakılan yere aşağıdakilerden hangisi yazılabilir?

- A) Boy uzaması
- B) Sakal çıkması
- C) Ses kalınlaşması
- D) Göğüslerin büyümesi

(37)

İnsan iskeletinde yer alan eklemler oynamaz eklemler, yarı oynar eklemler ve oynar eklemler şeklinde gruplandırılır.



Şekilde numaralandırılmış kısımların hangilerinde oynar eklem bulunur?

- A) 1 ve 2
- B) 3 ve 4
- C) 2, 3 ve 5
- D) 4, 5 ve 6

(38)

Burun, solunum sisteminde yer alan organlardan biridir. Burunda gerçekleşen olaylardan hangisi, akciğerlere ulaşan havanın tozlardan ve mikroplardan temizlenmesinde etkilidir?

- A) Havanın nemlendirilmesi
- B) Havanın ısıtılması
- C) Mukus üretilmesi
- D) Koku alınması

(39)

İnsanda;

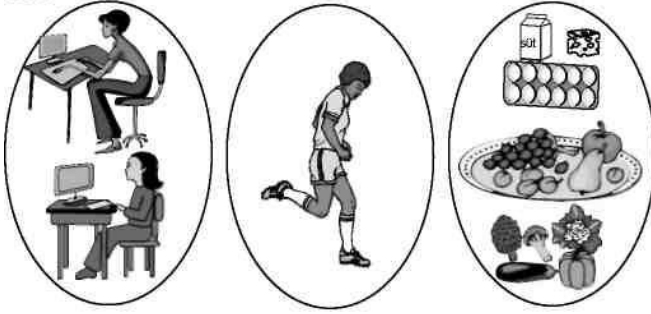
- D) Diyafram, soluk alıp vermede görevlidir.
- Y) Akciğer, kanı kalbe gönderir.
- Y) Kaburga kasları, soluk alıp verirken kasılıp gevşer.

Bir öğrenci çalışma kitabında verilen yukarıdaki bilgilerin yanındaki kutucuklara D (doğru) ve Y (yanlış) yazmıştır.

Bu çalışmayı yapan öğrenci ile ilgili olarak aşağıdakilerden hangisi söylenebilir?

- A) Sadece diyaframın görevini doğru biliyor.
- B) Sadece kaburga kaslarının görevini doğru biliyor.
- C) Solunum sistemindeki tüm organların görevini doğru biliyor.
- D) Solunum sistemindeki organlardan hiçbirinin görevini doğru bilmiyor.

(40)



Bir öğrenci sınıfta anlatacağı konu için yukarıdaki üç resmi hazırlamıştır.

Bu öğrencinin anlatacağı konu aşağıdakilerden hangisidir?

- A) Yaşımıza uygun sporlar
- B) Vitamin değeri yüksek olan bitkisel besinler
- C) Destek ve hareket sisteminin sağlığı
- D) Dolaşım sisteminde görülen hastalıklar

TEST BİTTİ.

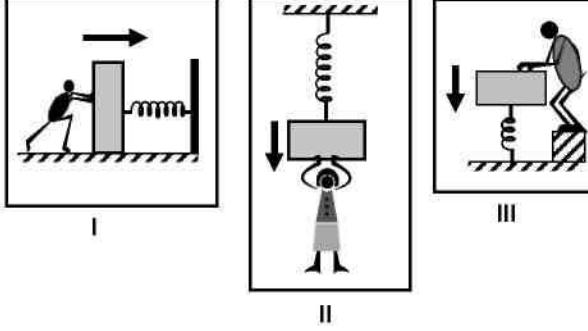
LÜTFEN CEVAPLARINIZI

KONTROL EDİNİZ.

7. SINIF FEN BİLİMLERİ DERSİ SEVİYE BELİRLEME TESTİ

(1)

Üç öğrenci I, II, III'teki yaylara oklarla gösterilen yönlerdeki kuvvetleri uyguluyorlar.



Yayların bu kişilere uyguladıkları kuvvetlerin yönleri hangi seçenekte doğru olarak verilmiştir?

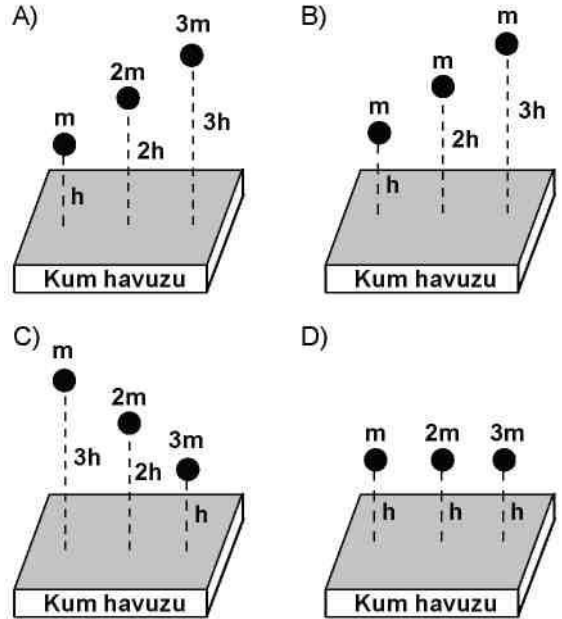
	I	II	III
A)	→	↓	↓
B)	←	↑	↓
C)	←	↑	↑
D)	→	↓	↑

(2)

Bir öğretmen öğrencilerinden "Kinetik enerji kütle ile doğru orantılıdır." ifadesini doğrulayan bir deney düzeneği hazırlamalarını istiyor.

Öğrencilerin hazırladığı aşağıdaki düzeneklerde kütleleri verilmiş eşit hacimli küresel cisimler, belirtilen yüksekliklerden serbest bırakılıyor ve bu cisimlerin kum havuzunda oluşturdukları çukurların derinlikleri not ediliyor.

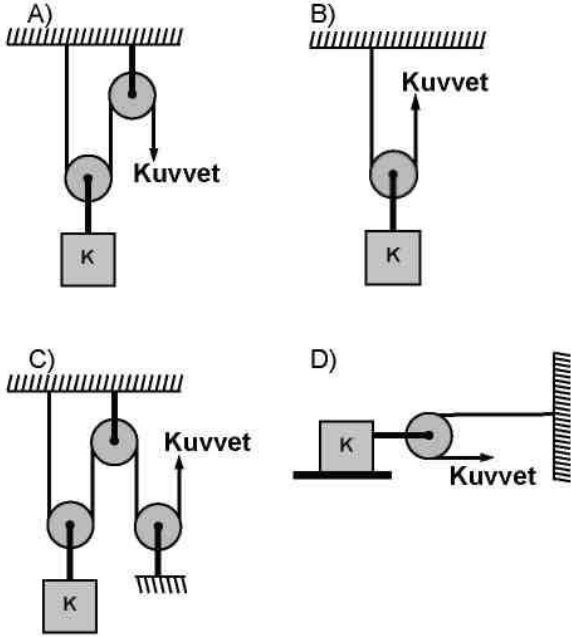
Bunlardan hangisi öğretmenin istediği düzenektir?



(3)

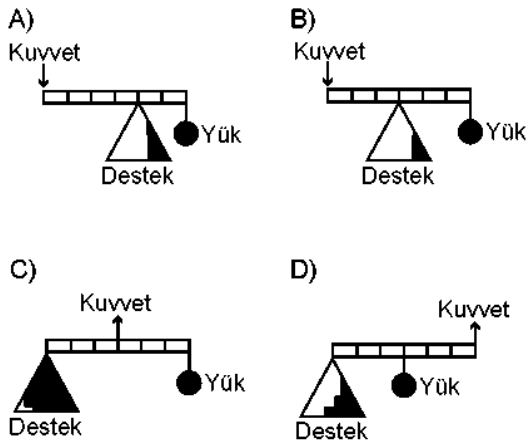
Öğretmen öğrencilerine, “Bana öyle bir makara sistemi hazırlayın ki bu sistem, uyguladığım kuvveti K cismine zıt yönde iletсин.” diyor. Öğrenciler de aşağıdaki düzenekleri hazırlıyorlar.

Hangisi öğretmenin istediği düzenektir?

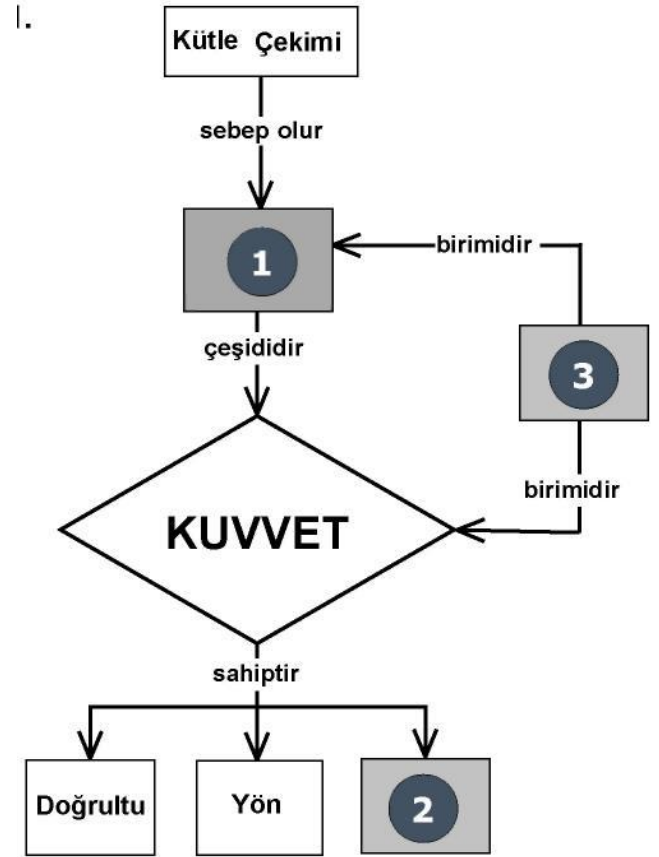


(4)

Aşağıdaki kaldıraçlardan hangisi, yükü yukarı kaldırmak için uygulanması gereken kuvvetin hem yönünü hem de büyüklüğünü değiştirmiştir? (Kaldıraçın ağırlığını ihmal ediniz.)



(5)



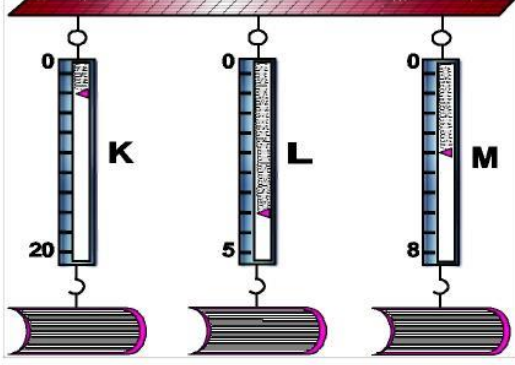
Öğretmen, “kuvvet” konusu ile ilgili yukarıdaki kavram haritasını, numaralandırılmış kısımlarını boş bırakarak tahtaya çiziyor. 1, 2 ve 3 kutucuklarına yazılacak kavramlar için söz alan dört öğrenci aşağıdaki seçenekleri oluşturuyor.

Buna göre, hangi seçenekte bu kavramlar doğru olarak verilmiştir?

	1	2	3
A)	Yerçekimi	Kütle	Kilogram
B)	Kütle	Büyükölük	Dinamometre
C)	Ağırlık	Dinamometre	Newton
D)	Ağırlık	Büyükölük	Newton

(6)

Zeynep, Fen ve Teknoloji ders kitabının ağırlığını, K, L ve M dinamometreleri ile ölçtüğünde dinamometreleri şekildeki gibi görüyor.



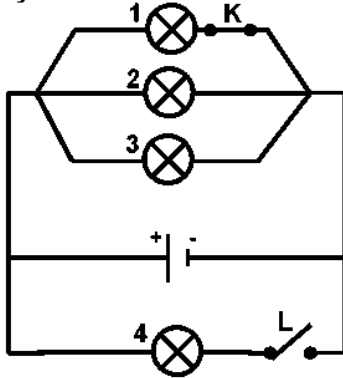
Bu etkinliğe göre, Zeynep'in aşağıda çıkardığı sonuçlardan hangileri doğrudur?

- I. K dinamometresi ile daha ağır kitapları ölçebilirim.
- II. Daha hassas ölçüm yapmak istediğimde L dinamometresini kullanmalıyım.
- III. Üç tane Fen ve Teknoloji ders kitabının ağırlığını M dinamometresi ile ölçebilirim.

- A) Yalnız I B) I - II
C) II - III D) I - II - III

(7)

Numaralandırılmış özdeş ampullerle kurulu şekildeki devrede K anahtarı açılıp L anahtarı kapatıldığında aşağıdaki durumlardan hangisi gerçekleşir?



- A) 2 ve 3 nolu ampullerin parlaklığı aynı kalır.
B) Ana koldan geçen akım artar.
C) 1 nolu ampulün parlaklığı artar.
D) Devrenin eşdeğer direnci artar.

(8)

K, L ve M cisimlerinden M'nin nötr olduğu biliniyor, K ve L'nin yük durumları ise bilinmiyor. K, L'ye dokundurulup ayrıldığında K ve L'nin birbirine itme-çekme kuvveti uygulamadıkları görülüyor. K, L'ye dokundurulmadan önce M'ye dokundurulup ayrıldığında ise K ve M'nin birbirini ittikleri görülüyor.

Buna göre, K ve L'nin ilk yük durumlarıyla ilgili aşağıda verilenlerden hangileri doğru olabilir?

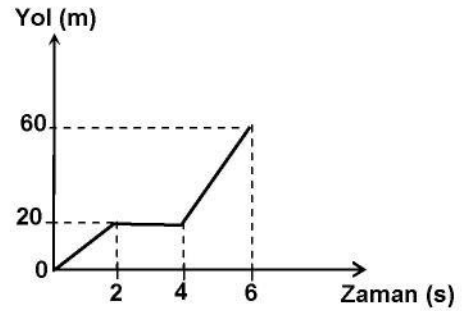
- I- K ve L nötrdür.
- II- K ve L pozitif yüklüdür.
- III- K pozitif, L negatif yüklüdür.

- A) Yalnız I B) Yalnız III
C) I - III D) II - III

(9)



Bir çitanın hareketini inceleyen gözlemci, çitaya ait aşağıdaki yol – zaman grafiğini elde ediyor.



Bu grafiğe göre, aşağıdaki ifadelerden hangileri doğrudur?

- I. 0 - 2 saniye arasında sürati 10 m/s'dir.
- II. 2 - 4 saniye arasında hareket enerjisi vardır.
- III. 4 - 6 saniye arasında sürati 30 m/s'dir.

- A) Yalnız I B) I - III
C) II - III D) I - II - III

(10)

Sesin yansımaları ve yankı olayları, yaşamımızda önemli bir yer tutar. Bunlara birçok örnek verilebilir.

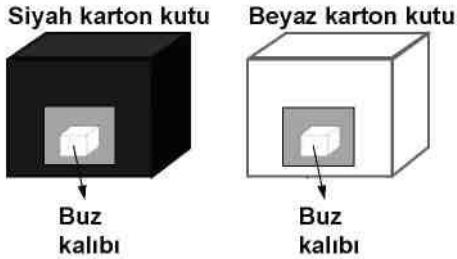
- I- Doktorların ultrason cihazı kullanarak hastalık teşhisi yapmaları,
- II- Gemilerin sonar cihazı kullanarak yön ve derinlik tespiti yapmaları,
- III- Tiyatro salonlarının duvar ve tavanlarının yumuşak ve pürüzlü yüzey ile kaplanması.

Yukarıda verilen örneklerden hangileri, bilim ve teknolojide sesin yansımalarından yararlandığını gösterir?

- A) Yalnız I
- B) I ve II
- C) II ve III
- D) I, II ve III

(11)

Ali buzdolabından aldığı iki özdeş buz kalıbını şekildeki gibi karton kutuların içerisine koyarak, bu kutuları Güneş altına bırakıyor.



Belli bir süre sonra kutuların içindeki buz kalıplarını kontrol ediyor. En fazla eriyen buz kalıbının siyah kutudaki olduğunu gözlemliyor.

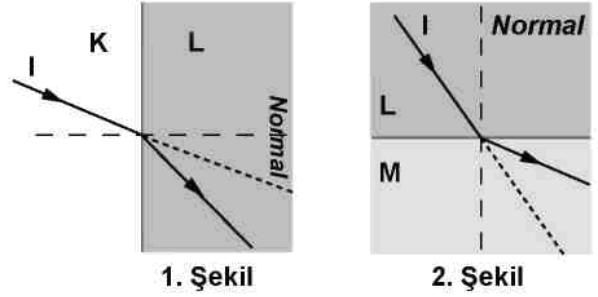
Bunun nedeni aşağıdakilerden hangileri ile açıklanabilir?

- I- Koyu renklerin ışığı daha fazla soğurması
- II- Koyu renklerin ışığı daha az yansıtması
- III- Koyu renklerin ışığı daha fazla kırması

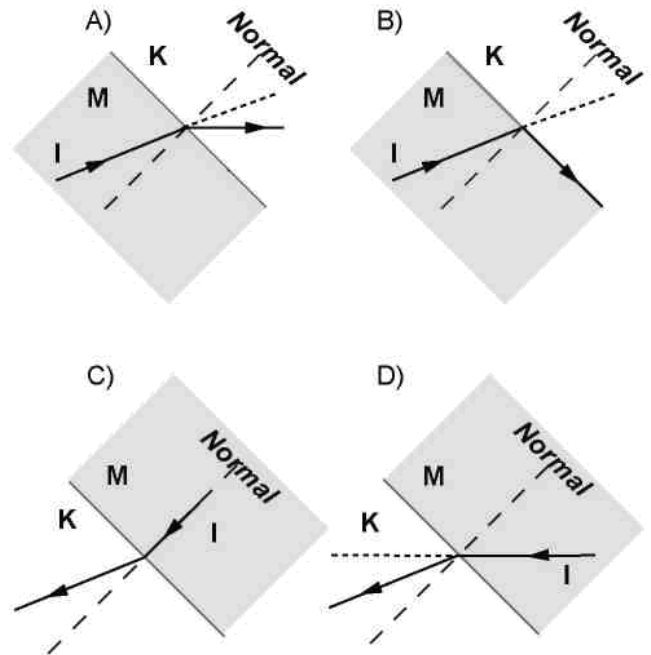
- A) Yalnız I
- B) Yalnız III
- C) I - II
- D) II - III

(12)

I ışık ışını K saydam ortamından L saydam ortamına 1. Şekildeki gibi, L saydam ortamından M saydam ortamına ise 2. Şekildeki gibi kırılarak geçiyor.



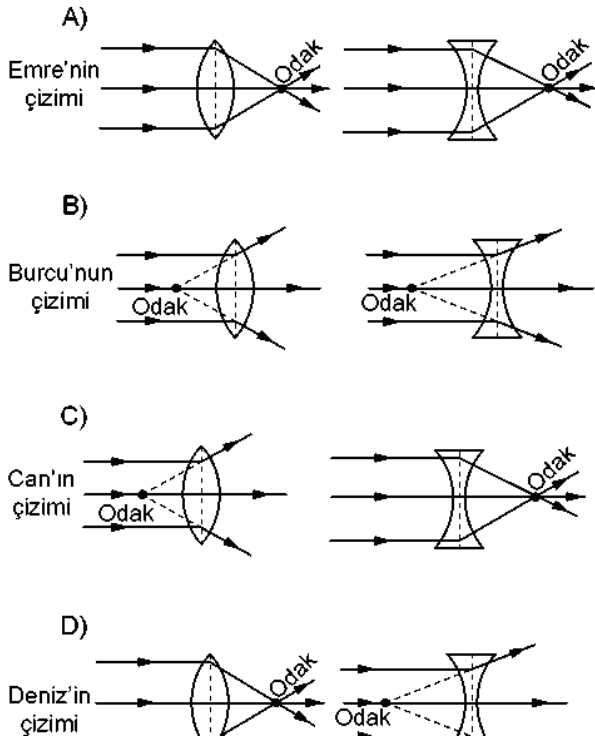
Buna göre, ışık ışınının M'den K'ye geçişini gösteren çizim aşağıdakilerden hangisi olabilir?



(13)

Öğretmen sınıfa biri ince kenarlı, diğeri de kalın kenarlı iki mercek ile geldi. Öğrencilerden etkinlikler yapıp bu merceklerin odak noktalarını çizerek göstermelerini istedi. Bunun üzerine öğrencilerden bazıları tahtaya aşağıdaki çizimleri yaptılar.

Buna göre, hangi öğrenci doğru çizim yapmıştır?



(14)

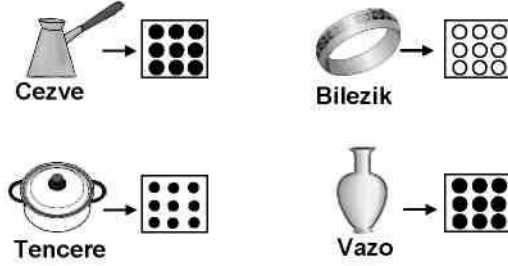
Sürtünme kuvvetinin etkili olduğu bazı eylemler aşağıda verilmiştir:

- I- Bir atletin yarışa hızlı başlayabilmesi
- II- Bir dağcının dağa tırmanabilmesi
- III- Bir sporcunun buz pistinde kayabilmesi
- IV- Bir bisikletlinin hareket hâlinde iken durabilmesi

Bu eylemlerin hangisinde sürtünme kuvvetinin az olması, eylemin daha iyi gerçekleşmesini sağlar?

- A) I'de B) II'de C) III'te D) IV'te

(15)



Şekildeki eşyaları ve bu eşyaların yapıldığı maddelerin tanecik modellerini gösteren kartları sınıfa getiren öğretmen, öğrencilerine;

“Tanecik modellerine göre hangi eşyalar aynı maddeden yapılmış olabilir?” sorusunu yöneltir.

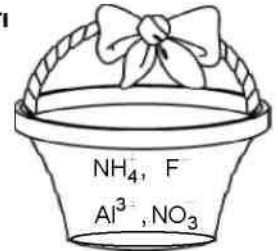
Öğrencilerin verdikleri aşağıdaki cevaplardan hangisi doğrudur?

- A) Cezve – Vazo
B) Cezve – Tencere
C) Cezve – Tencere – Vazo
D) Cezve – Bilezik – Vazo

(16)

Özlem, sepetteki iyonları anyon ve katyon olarak ayırmak istiyor.

Bu iyonları anyon ve katyon kutularına aşağıdakilerden hangisindeki gibi yerleştirmelidir?



- A) Anyon: Al^{3+}, F^{-} Katyon: NH_4^{+}, NO_3^{-}
- B) Anyon: Al^{3+}, NH_4^{+} Katyon: F^{-}, NO_3^{-}
- C) Anyon: Al^{3+}, NO_3^{-} Katyon: NH_4^{+}, F^{-}
- D) Anyon: NO_3^{-}, F^{-} Katyon: NH_4^{+}, Al^{3+}

(17)

Emel, sınıfta yapılacak bir etkinlik için üzerlerinde Güneş sistemindeki üç gezegene ait bilgilerin olduğu aşağıdaki I, II, III kartlarını hazırlıyor.

<p>◆ Güneş sisteminin en büyük gezegenidir.</p> <p>◆ Güneş'e en yakın beşinci gezegendir.</p>	<p>◆ Güneş sisteminin en küçük gezegenidir.</p> <p>◆ Güneş'e en yakın gezegendir.</p>	<p>◆ Güneş sisteminin ikinci en küçük gezegenidir.</p> <p>◆ Güneş'e en yakın dördüncü gezegendir.</p>
---	---	---

I

II

III

Emel'in hazırladığı bu kartların arka yüzüne yazması gereken gezegen isimleri hangi seçenekte doğru olarak verilmiştir?

I	II	III
A) Jüpiter	Merkür	Mars
B) Neptün	Merkür	Dünya
C) Satürn	Venüs	Mars
D) Jüpiter	Venüs	Merkür

(18)

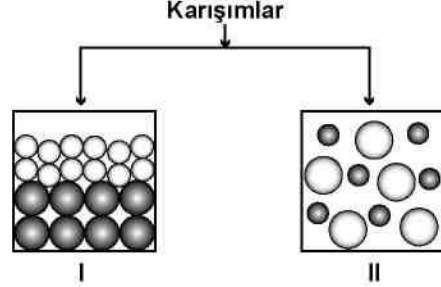
Dil	Elementin adı	Sembölü
İtalyanca	Azoto	N
Almanca	Stickstoff	
Türkçe	Azot	
Latince	Nitrum	

Çizelgeye göre, aşağıdaki yorumlardan hangisi yapılabilir?

- A) Bir elementin adı her dilde farklı olsa da sembolü aynıdır.
B) Elementler sembollerle, bileşikler formüllerle gösterilir.
C) Elementlerin adları eski dönemlerde işaretlerle gösterildi.
D) Bilimsel çalışmalarda elementleri adlarıyla kullanmak iletişimi kolaylaştırır.

(19)

Karışımların sınıflandırılması, şemada tanecik modeli ile gösterilmiştir:



Buna göre, I ve II ile gösterilen tanecik modelleri hangi tür karışımı temsil eder? Neden?

- A) I. model, homojen karışımdır. Çünkü maddelerden biri dibe çökmüştür.
B) I. model, heterojen karışımdır. Çünkü karışımdaki maddelerin tanecik sayıları farklıdır.
C) II. model, heterojen karışımdır. Çünkü karışım iki tür madde içerir.
D) II. model, homojen karışımdır. Çünkü tanecikler karışımın her tarafına eşit dağılmıştır.

(20)

Aşağıda numaralandırılmış kutucuklarda bazı elementlerin tanecik modelleri verilmiştir. Hangi kutucuklardaki tanecik modelleri aynı elementi temsil eder?

1 	2
3 	4
5 	6

- A) 1 ve 4
B) 1, 4 ve 5
C) 5 ve 6
D) 2, 3 ve 6

(21)

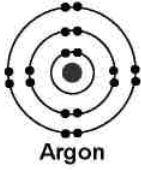
Aynı miktarda alınan X ve Y maddeleri özdeş pistonlu kaplara konuluyor. Bu maddelere eşit kuvvet uygulandığında Y maddesinin sıkıştırılabildiği, X'in ise sıkıştırılamadığı gözleniyor.

Bu gözleme bağlı olarak X ve Y maddelerini oluşturan tanecikler ile ilgili aşağıdakilerden hangisi söylenebilir?

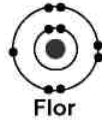
- A) X maddesi bütünsel, Y maddesi ise tanecikli yapıya sahiptir.
- B) Y maddesini oluşturan tanecikler arasındaki boşluk daha fazladır.
- C) X maddesini oluşturan taneciklerin arasındaki boşluk yoktur.
- D) X ve Y maddelerini oluşturan tanecikler arasındaki boşluk aynıdır.

(22)

Merhaba, benim adım oksijen. Oyun oynamayı seviyorum. Benimle iyonik bağ yapabilecek arkadaşlarla oynamak istiyorum. Aşağıdaki hangi arkadaşlarla oynayabilirim?



Argon



Flor

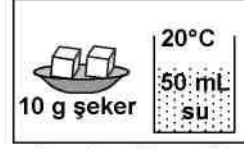


Magnezyum

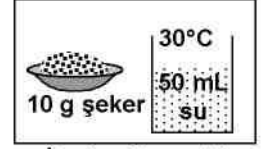
- A) Flor
- B) Magnezyum
- C) Argon
- D) Argon ve flor

(23)

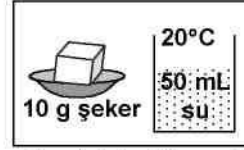
Bazı öğrencilerin aynı miktardaki şekerleri suda çözmek için hazırladıkları düzenekler aşağıda verilmiştir.



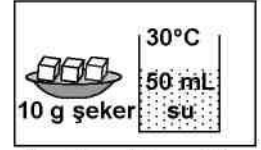
Eren'in düzenegi



İlay'ın düzenegi



Sevde'nin düzenegi



Ece'nin düzenegi

Hangi öğrencinin hazırladığı düzenekte şeker daha kısa sürede çözünür?

- A) Eren'in
- B) Sevde'nin
- C) Ece'nin
- D) İlay'ın

(24)

Korku, heyecan, mutluluk ve öfke gibi durumlarda vücutta adrenal hormonu seviyesi artar. Buna göre, aşağıdaki durumların hangisinde Hülya'nın adrenal hormonu seviyesinde artma beklenir?

- A) Yemek yerken su içtiğinde
- B) Ders çalıştıktan sonra uyuduğunda
- C) Her gün, ev işlerinde annesine yardım ettiğinde
- D) Sınavda başarılı olunca aşırı sevindiğinde

(25)

Aşağıdaki tabloyu hazırlayan öğretmen, öğrencilerine tahtadaki yazıları yakından ve uzaktan okumalarını söylüyor.

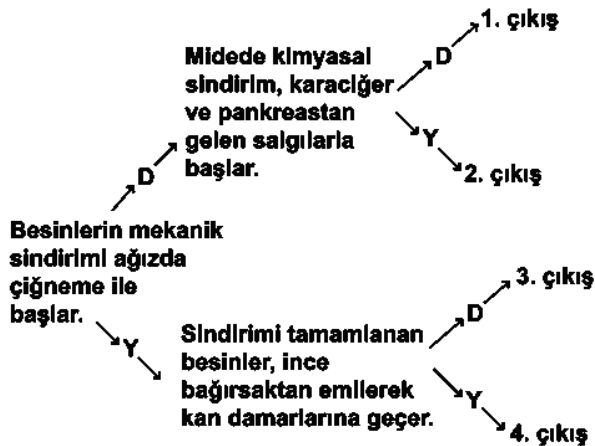
Sorular İsimler	Tahtadaki yazıyı yakından okuyor mu?	Tahtadaki yazıyı uzaktan okuyor mu?
Ayşe	Evet	Hayır
Ahmet	Evet	Evet
Mehmet	Hayır	Evet
Seda	Hayır	Evet
Ali	Evet	Evet

Sonuçları tabloya yazan öğretmen öğrencilerinde hangi göz kusurlarını belirlemek istiyor?

- A) Miyopluk ve hipermetropluk
- B) Şaşılık ve renk körlüğü
- C) Astigmatlık ve şaşılık
- D) Renk körlüğü ve miyopluk

(26)

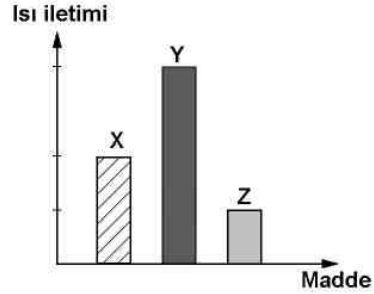
Aşağıda verilen birbiriyle bağlantılı cümleler arasındaki D "Doğru", Y "Yanlış" anlamındadır. Sindirim sistemi hakkında yeterli bilgiye sahip olan bir öğrenci kaçınıcı çıkışa ulaşır?



- A) 1.
- B) 2.
- C) 3.
- D) 4.

(27)

X, Y, Z katı maddelerinin ısı iletimi, grafikteki gibidir:



Bu maddelerin günlük yaşamdaki kullanımlarıyla ilgili olarak aşağıdakilerden hangisi yanlıştır?

- A) Binaların dış cephe kaplamasında Z maddesi tercih edilmelidir.
- B) Ütü tabanı yapılırken Y maddesi kullanılmalıdır.
- C) Yemek pişirirken karıştırmada kullanılacak kaşık yapımı için en uygun, X maddesidir.
- D) Kışlık ayakkabı yapımında Z maddesi tercih edilmelidir.

(28)

Sinir sistemimizde bulunan aşağıdaki merkezlerden hangisi omurilikte yer alır?

- A) Konuşma merkezi
- B) Açlık ve tokluğu algılama merkezi
- C) Vücut sıcaklığını düzenleme merkezi
- D) Refleksleri kontrol etme merkezi

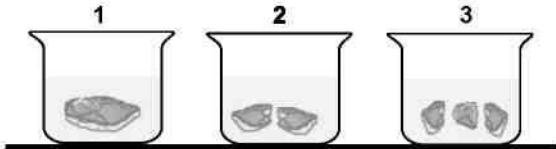
(29)

Burun, solunum sisteminde yer alan organlardan biridir. Burunda gerçekleşen olaylardan hangisi, akciğerlere ulaşan havanın tozlardan ve mikroplardan temizlenmesinde etkilidir?

- A) Havanın nemlendirilmesi
- B) Havanın ısıtılması
- C) Mukus üretilmesi
- D) Koku alınması

(30)

Şekildeki özdeş cam kaplarda bulunan et parçaları üzerine, aynı sindirim enziminin eşit miktarda konuluyor. Sonra kaplara eşit miktarda su ekleniyor.



Kaplarda aynı miktarda et bulunmaktadır.

Aynı süre içinde et parçalarının sindiriminin önce 3., sonra 2., en son olarak ta 1. kaptaki tamamlandığı görülüyor.

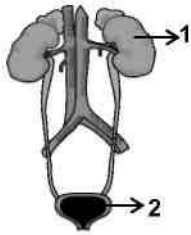
Bu deney, aşağıdaki yorumlardan hangisini destekler?

- A) Sindirimi kolaylaştırmak için yiyecekleri iyice çiğnemeliyiz.
- B) Et türü yiyecekleri diğer besinlerden daha önce tüketmeliyiz.
- C) Yemek sırasında sık sık su içmeliyiz.
- D) Dengeli ve yeterli beslenmeliyiz.

(31)

Öğretmen:

Şekildeki boşaltım sisteminde verilen 1 ve 2 numaralı organların isim ve görevlerini söyler misin?



Öğrenci:

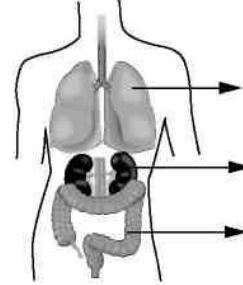
1 numaralı organ böbrektir, idrarı depo eder.
2 numaralı organ idrar kesesidir, kanı süzer.

Bu açıklamalara göre öğrenci ile ilgili olarak aşağıdakilerden hangisi söylenebilir?

- A) Boşaltım sistemi organlarını bilmiyor.
- B) Boşaltım sistemi organları ile diğer sistemlerin organlarını ayırt edemiyor.
- C) Boşaltım sistemi organlarının şeklini biliyor, ancak görevlerini birbirine karıştırıyor.
- D) Boşaltım sistemi organlarını ve görevlerini çok iyi biliyor.

(32)

Öğretmenin sorduğu bir soruya Ahmet, şekil üzerindeki işaretlenmiş üç organı göstererek doğru cevap vermiştir.

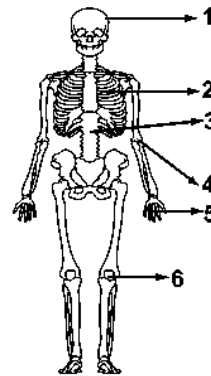


Buna göre, öğretmenin sorduğu soru aşağıdakilerden hangisidir?

- A) Hangi organlar sindirime yardımcı olur?
- B) Hangi organlar solunum sisteminde yer alır?
- C) Kandaki zararlı atıklar hangi organlarda süzülür?
- D) Atık maddelerin vücuttan atılmasında hangi organlar görevlidir?

(33)

İnsan iskeletinde yer alan eklemler oynamaz eklemler, yarı oynar eklemler ve oynar eklemler şeklinde gruplandırılır.



Şekilde numaralandırılmış kısımların hangilerinde oynar eklemler bulunur?

- A) 1 ve 2
- B) 3 ve 4
- C) 2, 3 ve 5
- D) 4, 5 ve 6

(34)

İnsanda;
 D) Diyafram, soluk alıp vermede görevlidir.
 Y) Akciğer, kanı kalbe gönderir.
 Y) Kaburga kasları, soluk alıp verirken kasılıp gevşer.

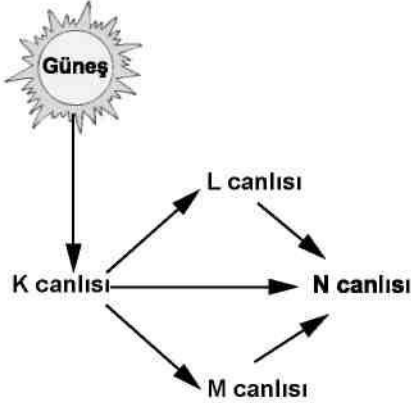
Bir öğrenci çalışma kitabında verilen yukarıdaki bilgilerin yanındaki kutucuklara D (doğru) ve Y (yanlış) yazmıştır.

Bu çalışmayı yapan öğrenci ile ilgili olarak aşağıdakilerden hangisi söylenebilir?

- A) Sadece diyaframın görevini doğru biliyor.
- B) Sadece kaburga kaslarının görevini doğru biliyor.
- C) Solunum sistemindeki tüm organların görevini doğru biliyor.
- D) Solunum sistemindeki organlardan hiçbirinin görevini doğru bilmiyor.

(35)

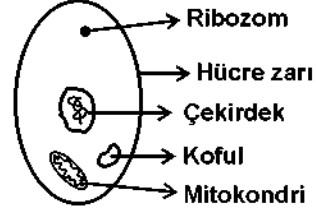
Şemada, bazı canlılardaki beslenme ilişkileri gösterilmiştir.



Buna göre, aşağıdakilerden hangisi yanlıştır?

- A) K canlısı, fotosentez yapabilmektedir.
- B) L canlısı, sadece bitkilerle beslenmektedir.
- C) M canlısı, hem bitkilerle hem de etle beslenmektedir.
- D) N canlısı, farklı türden besin kaynaklarını kullanabilmektedir.

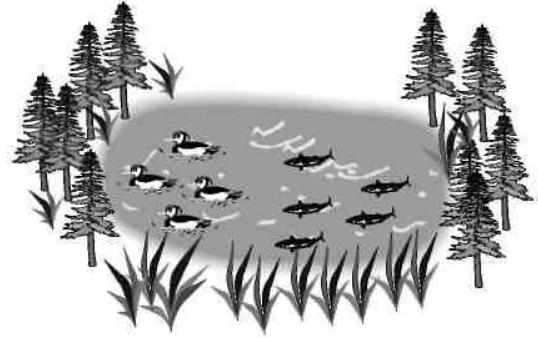
(36)



Yukarıda verilen hücrenin kısımlarını inceleyen Mehmet, bu hücreye sahip olan canlı ile ilgili olarak hangi yorumu yapamaz?

- A) Kendi besinini yaptığı
- B) Solunum yaptığı
- C) Kalıtsal maddeye sahip olduğu
- D) Dış ortamla madde alışverişinin olduğu

(37)



Öğretmen: Yukarıda verilen resimdeki popülasyonlardan birini söyleyiniz.

Aydın : Göldeki ördekler

Neşe : Göl kenarındaki tüm bitkiler

Öğretmen: Aydın'ın cevabı doğru, Neşe'nin cevabı yanlış. Çünkü popülasyon ...

Aşağıdakilerden hangisi öğretmenin açıklamasını tamamlar?

- A) sadece hayvanlardan oluşur.
- B) aynı tür canlılardan oluşur.
- C) gölde yaşayan canlılardan oluşur.
- D) hem karada hem suda yaşayan canlılardan oluşur.

(38)

Öğrenciler, sınıftaki bir etkinlik için şekildeki fotoğrafları kullanacaklardır.

Önce



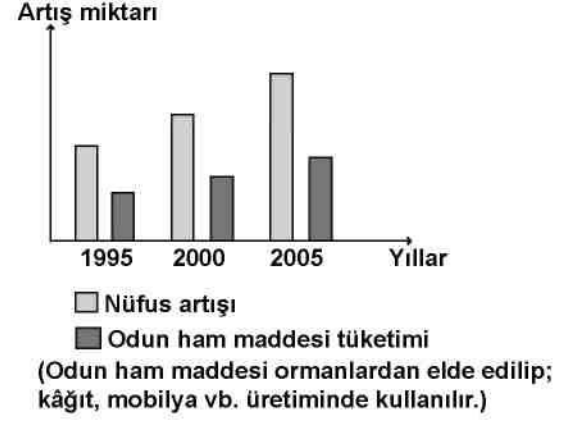
Sonra

Kuzey Kutbu'ndaki bir bölgenin önceki ve sonraki durumunu anlatan bu fotoğraflar, **daha çok aşağıdaki konulardan hangisiyle ilgilidir?**

- A) Küresel ısınma ve sonuçları
- B) Biyolojik çeşitliliğin azalması
- C) Ozon tabakasının zarar görmesi
- D) Su kirliliğinin canlılar üzerine etkileri

(39)

İlköğretim 7. sınıf öğrencisi olan Hülya, sınıfındaki bir etkinlik için araştırma yaparken "Çevre" dergisinden aşağıdaki grafiği bulmuştur.

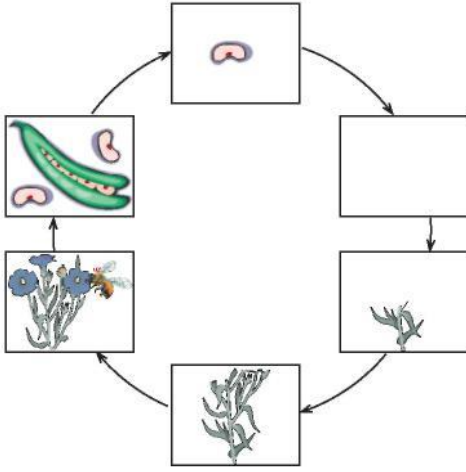


Hülya ve arkadaşları bu grafiği inceledikten sonra, odun ham maddesi tüketimini azaltmak için kendilerinin hem **çabuk** hem de **kolay** uygulayacakları bir karar alırlar.

Bu karar aşağıdakilerden hangisi olabilir?

- A) Isınmak için evimizde odun yerine güneş enerjisi ve rüzgâr enerjisini kullanalım.
- B) Mobilya üretiminde başka maddeler kullanalım.
- C) Sınıfımızda kâğıtların gereksiz kullanımını engelleyelim.
- D) Yangına karşı dayanıklı bitki türleri elde edelim.


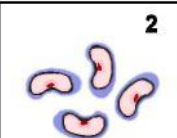

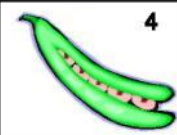
(40)



Fasulyenin hayat döngüsü

Emel, fasulyenin hayat döngüsünün aşamalarını gösteren resimleri kâğıtlara yapıştırdığında, bir resmin eksik kaldığını görüyor.

Buna göre, Emel eksik kalan resmin yerine aşağıdaki tablodan kaç numaralı resmi seçmelidir?

 <p>1</p>	 <p>2</p>
 <p>3</p>	 <p>4</p>

Tablo

- A) 1 B) 2 C) 3 D) 4

TEST BİTTİ.

LÜTFEN CEVAPLARINIZI

KONTROL EDİNİZ.

8. SINIF FEN BİLİMLERİ DERSİ SEVİYE BELİRLEME TESTİ

(1)

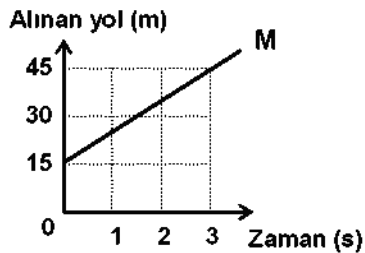
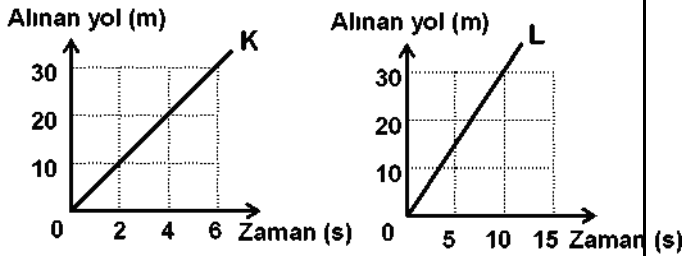
Deniz, bir kutuyu yatay doğrultuda 100 m ileriye götürmek istiyor.

Kutuyu aşağıdakilerden hangisindeki gibi götürürse, fiziksel anlamda en az iş yapmış olur?

- A) Yerden kucağına alarak
- B) Yerde iterek
- C) El arabası ile taşıyarak
- D) İple çekerek

(2)

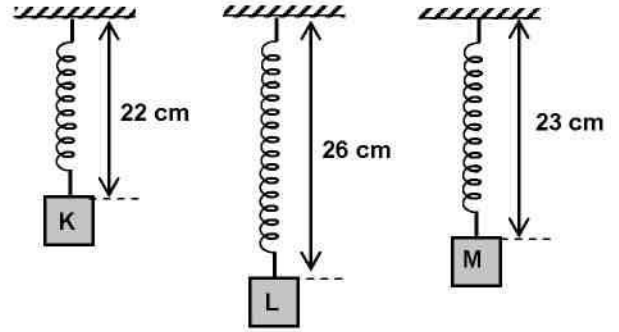
K, L ve M araçlarına ait alınan yol-zaman grafikleri aşağıda verilmiştir.



K, L ve M araçlarının ortalama süratleri, aşağıdakilerin hangisinde doğru olarak verilmiştir?

	K (m/s)	L (m/s)	M (m/s)
A)	10	10	15
B)	10	5	10
C)	5	2	15
D)	5	3	10

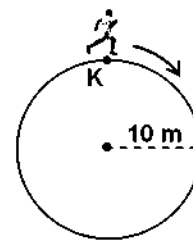
(3)



Cem, tavana astığı 20 cm uzunluğundaki yayın ucuna K, L, M cisimlerini astığında yay uzunlukları şekildeki gibi oluyor. K'nın ağırlığı 22 N olduğuna göre, L'nin ve M'nin ağırlıkları hangisinde doğru olarak verilmiştir?

	L (N)	M (N)
A)	66	33
B)	26	23
C)	60	33
D)	66	30

(4)



Bir çocuk 10 m yarıçaplı çembersel pistte, büyüklüğü sabit ve 5 m/s olan bir hızla K noktasından koşmaya başlıyor.

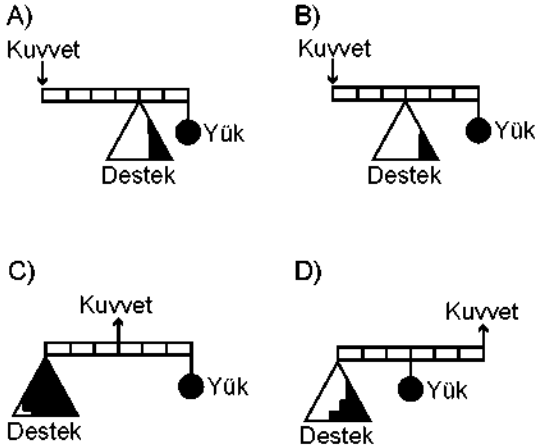
Buna göre, aşağıdakilerden hangileri doğrudur? ($\pi = 3$ alınız.)

- I. 3 saniye sonundaki yer değiştirme ile 9 saniye sonundaki yer değiştirme eşit büyüklüktedir.
- II. 6 saniye sonundaki yer değiştirmenin büyüklüğü 20 m dir.
- III. 6 saniye sonundaki yer değiştirmenin büyüklüğü, 8 saniyede aldığı yolun yarısıdır.

- A) Yalnız I
- B) Yalnız II
- C) II ve III
- D) I, II ve III

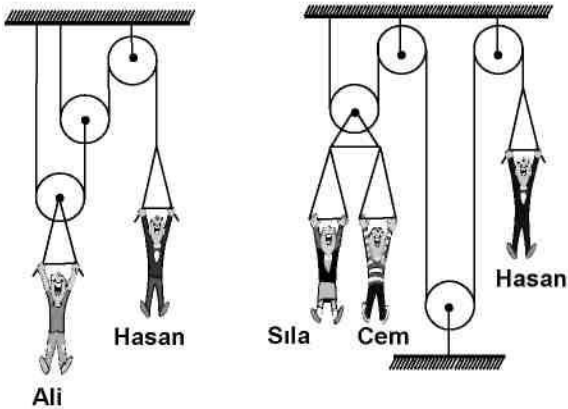
(5)

Aşağıdaki kaldıraçlardan hangisi, yükü yukarı kaldırmak için uygulanması gereken kuvvetin hem yönünü hem de büyüklüğünü değiştirmiştir? (Kaldıraçın ağırlığını ihmal ediniz.)



(6)

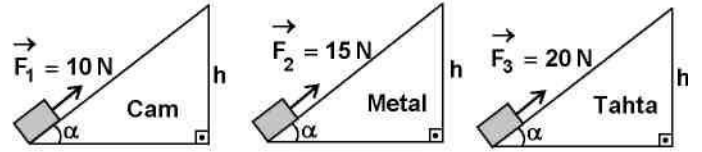
Aşağıda verilen ağırlıksız makara sistemlerinde Ali, Hasan ile, Hasan da Sıla ve Cem ile dengededirler.



Ali'nin kütesinin 50 kg olduğu bilindiğine göre Sıla ve Cem'in kütleleri hangisindeki gibi olamaz? (Sürtünmeler ihmal edilecektir.)

	Sıla (kg)	Cem (kg)
A)	15	12
B)	12	13
C)	11	14
D)	10	15

(7)



Şekildeki cam, metal ve tahtadan oluşan eğik düzlemlerde minimum \vec{F}_1 , \vec{F}_2 ve \vec{F}_3 kuvvetleri uygulanarak eşit kütleli cisimler h yüksekliğine çıkartılıyor.

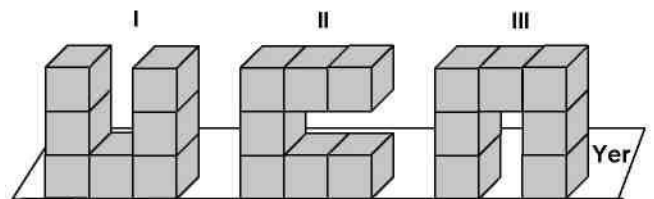
Buna göre, aşağıdakilerden hangileri doğrudur?

- Eğik düzlemlerde yapılan net işler birbirine eşittir.
- En fazla enerji tahta düzlemde harcanmıştır.
- En fazla kuvvet kazancı cam düzlemde sağlanmıştır.

- A) Yalnız I B) Yalnız III
C) I ve II D) I, II ve III

(8)

Birbirine yapışık, özdeş ve türdeş 7 küpten oluşan bir cisim aşağıdaki gibi üç farklı konumda yerleştiriliyor.



Cismin I, II ve III konumlarında yere göre potansiyel enerjileri sırası ile E_I , E_{II} ve E_{III} 'tür.

Buna göre, potansiyel enerjiler arasındaki ilişki hangisinde doğru olarak verilmiştir?

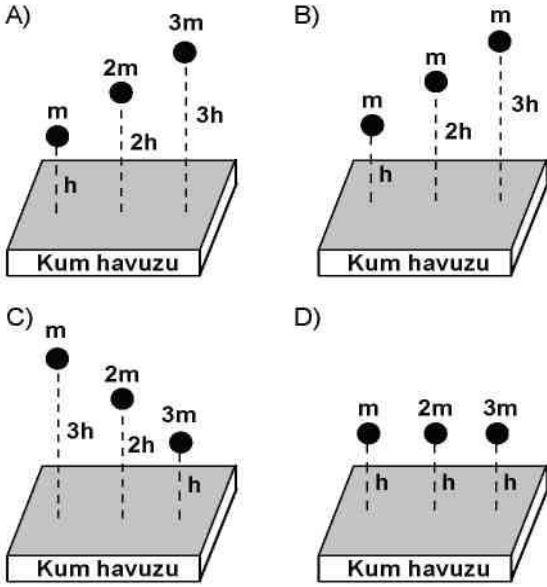
- A) $E_I = E_{II} = E_{III}$ B) $E_{III} > E_I = E_{II}$
C) $E_{III} > E_{II} > E_I$ D) $E_I = E_{III} > E_{II}$

(9)

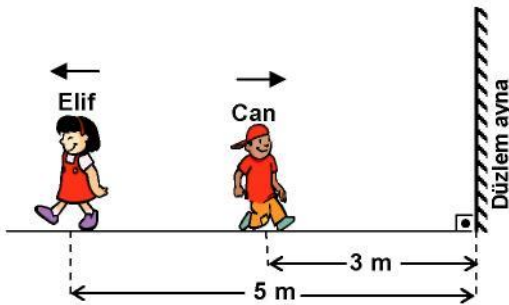
Bir öğretmen öğrencilerinden "Kinetik enerji kütle ile doğru orantılıdır." ifadesini doğrulayan bir deney düzeneği hazırlamalarını istiyor.

Öğrencilerin hazırladığı aşağıdaki düzeneklerde kütleleri verilmiş eşit hacimli küresel cisimler, belirtilen yüksekliklerden serbest bırakılıyor ve bu cisimlerin kum havuzunda oluşturdukları çukurların derinlikleri not ediliyor.

Bunlardan hangisi öğretmenin istediği düzenektir?



(10)

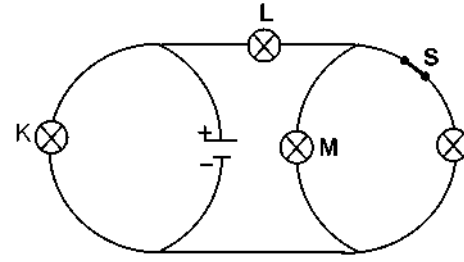


Can ve Elif şekildeki konumlarından aynı anda eşit büyüklükte ve zıt yönlü hızlarla yürümeye başlıyorlar.

Can, aynanın 1 m yakınına geldiğinde Elif'in aynadaki görüntüsünü kendisinden kaç metre uzakta görür?

- A) 6 B) 7 C) 8 D) 10

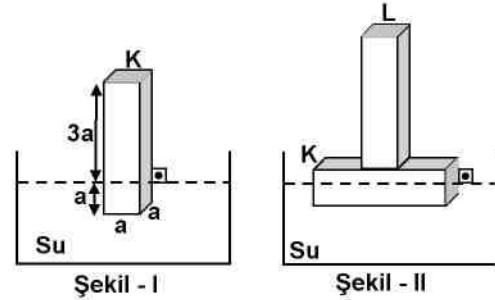
(11)



Özdeş ampullerle kurulu devrede S anahtarı kapalıdır. Anahtar açıldığında K, L ve M ampullerinin parlaklıklarının ilk durumlarına göre değişimi hangisinde doğru olarak verilmiştir?

	<u>K</u>	<u>L</u>	<u>M</u>
A) Değişmez	Artar	Artar	Artar
B) Değişmez	Azalı	Artar	Artar
C) Artar	Artar	Artar	Artar
D) Artar	Azalı	Değişmez	Değişmez

(12)

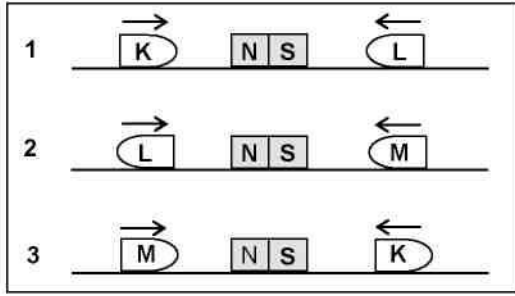


Suda çözünmeyen, homojen K cismi su içinde Şekil - I'deki gibi dengededir. K ile özdeş bir L cismi K'ya orta noktasından dik olarak perçinlenip su içine konulduğunda Şekil - II'deki gibi dengeye geliyor.

Bu durumda, KL cisminin kaç a'lık kısmı suya batar? (Cisimler dikdörtgenler prizması şeklindedir.)

- A) $\frac{5}{4}$ B) 1 C) $\frac{1}{2}$ D) $\frac{1}{4}$

(13)



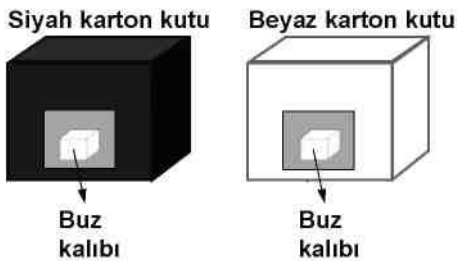
Yatay zemin üzerine sabitlenmiş çubuk mıknatısların yanlarına K, L ve M cisimleri 1, 2 ve 3'teki gibi konulduklarında oklarla belirtilen yönlere hareket ediyorlar. Buna göre, aşağıdakilerden hangileri kesinlikle doğrudur?

- I. K mıknatıstır.
- II. L mıknatıstır.
- III. M mıknatıs değildir.

- A) Yalnız III
- B) I ve II
- C) II ve III
- D) I, II ve III

(14)

Ali buzdolabından aldığı iki özdeş buz kalıbını şekildeki gibi karton kutuların içerisine koyarak, bu kutuları Güneş altına bırakıyor.



Belli bir süre sonra kutuların içindeki buz kalıplarını kontrol ediyor. En fazla eriyen buz kalıbının siyah kutudaki olduğunu gözlemliyor.

Bunun nedeni aşağıdakilerden hangileri ile açıklanabilir?

- I- Koyu renklerin ışığı daha fazla soğurması
- II- Koyu renklerin ışığı daha az yansıtması
- III- Koyu renklerin ışığı daha fazla kırması

- A) Yalnız I
- B) Yalnız III
- C) I - II
- D) II - III

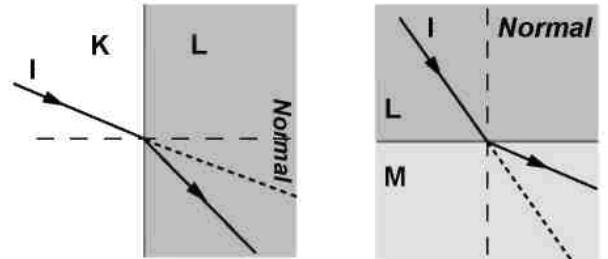
(15)

Aşağıdakilerden hangisi basıncı arttırmak için yapılmıştır?

- A) Rayların şekil bozukluğuna uğramaması için trenlerde tekerlek sayısının artırılması
- B) Meyvenin daha rahat kesilmesi için bıçağın keskinleştirilmesi
- C) Karda daha rahat yürümek için kar ayakkabısı kullanılması
- D) Traktörlerin toprağa saplanmaması için geniş tekerlekli yapılması

(16)

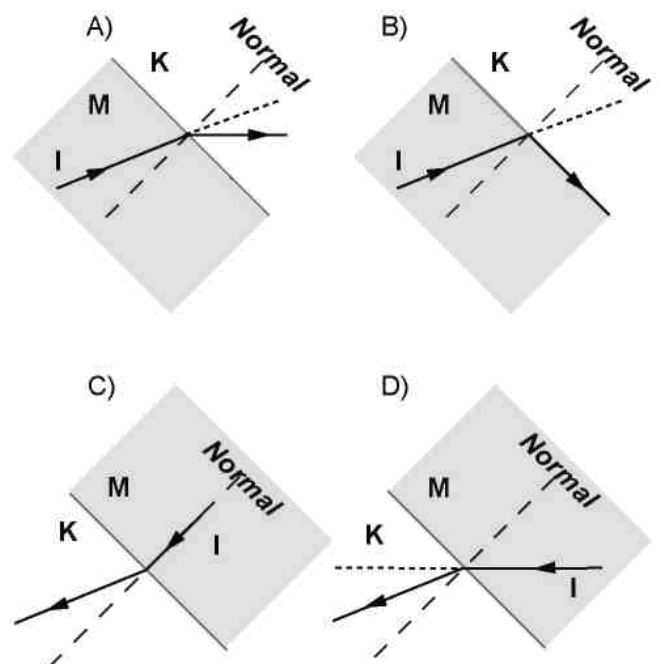
I ışık ışını K saydam ortamından L saydam ortamına 1. Şekildeki gibi, L saydam ortamından M saydam ortamına ise 2. Şekildeki gibi kırılarak geçiyor.



1. Şekil

2. Şekil

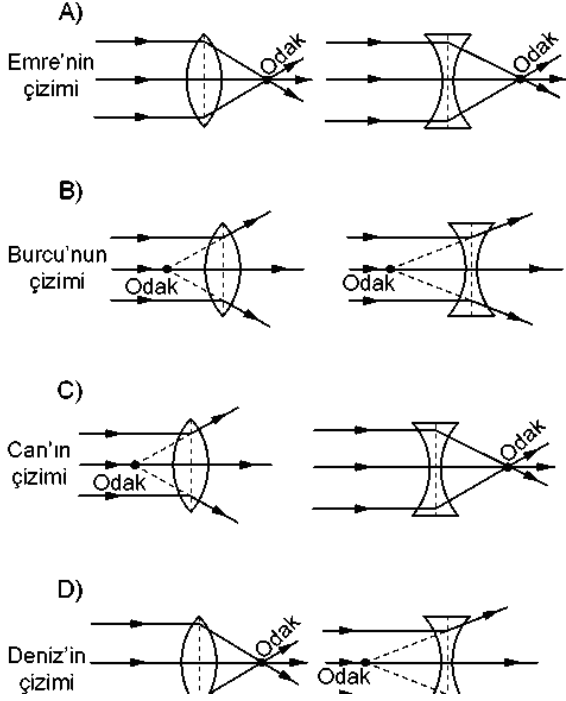
Buna göre, ışık ışınının M'den K'ye geçişini gösteren çizim aşağıdakilerden hangisi olabilir?



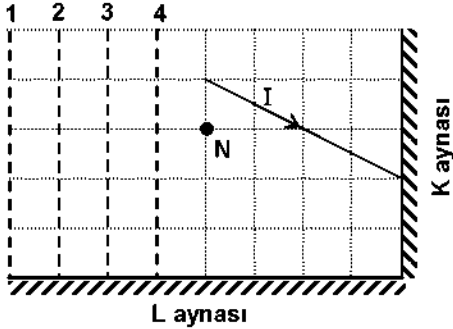
(17)

Öğretmen sınıfa biri ince kenarlı, diğeri de kalın kenarlı iki mercekle geldi. Öğrencilerden etkinlikler yapıp bu merceklerin odak noktalarını çizerek göstermelerini istedi. Bunun üzerine öğrencilerden bazıları tahtaya aşağıdaki çizimleri yaptılar.

Buna göre, hangi öğrenci doğru çizim yapmıştır?



(18)



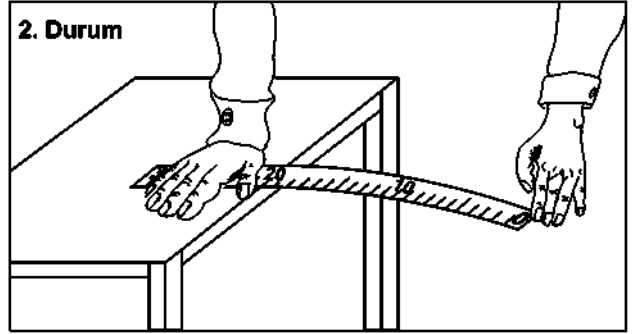
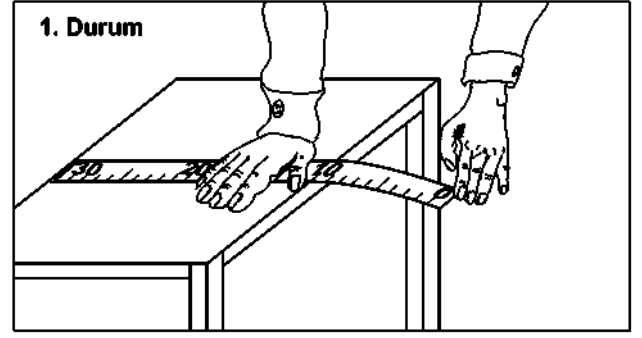
Bir optik düzende K ve L düzlem aynalarının konumu şekildeki gibidir. K düzlem aynasına gelen I ışık ışını K, L ve M aynalarından yansıdıktan sonra N noktasından geçiyor.

Buna göre, M düzlem aynası 1, 2, 3 ve 4 konumlarından hangisindedir?

(Şekildeki bölmeler eşit karelerden oluşmuştur.)

- A) 1 B) 2 C) 3 D) 4

(19)



Her iki durumda da cetveli bir eliyle sehpa üstüne bastırarak Efe, diğer eliyle de cetvelin boştaki ucunu aşağı doğru esnetip serbest bırakıyor.

Bu işlem sonunda 2. Durumda çıkan sesin 1. Durumda çıkan sestene daha kalın olduğunu fark ediyor.

Buna göre Efe, sesteki kalınlaşmanın nedenini aşağıdakilerden hangisi ile açıklar?

- A) Sesin genliğinin artmasıyla.
B) Sesin frekansının artmasıyla.
C) Sesin genliğinin azalmasıyla.
D) Sesin frekansının azalmasıyla.

(20)

Özellikleri Madde- nin hâli	Kütle	Hacim	Şekil
Katı	○	○	○
Sıvı	○	○	▲
Gaz	○	▲	▲

○ → Belirlidir.

▲ → Belirsizdir.

Çizelgede maddenin katı, sıvı ve gaz hâlinin özellikleri verilmiştir.

Buna göre;

Katı → Sıvı

Sıvı → Gaz

Gaz → Sıvı

Sıvı → Katı

dönüşümlerinde maddenin hâllerine ait özelliklerden hangisi kesinlikle değişmez?

A) Kütle

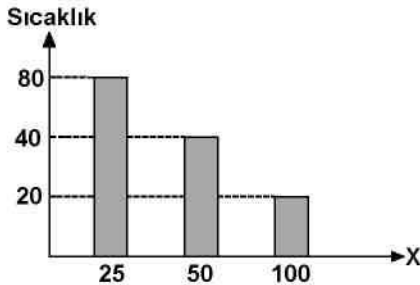
B) Hacim

C) Şekil

D) Kütle - Hacim

(21)

Bir öğrenci başlangıç sıcaklıkları eşit olan aynı sıvının farklı miktarlarını özdeş ısıtıcılarla eşit süre ısıtıktan sonra sıcaklıklarını ölçüyor ve şekildeki grafiği elde ediyor.



Buna göre, öğrencinin X ile belirttiği değişken aşağıdakilerden hangisidir?

A) Isı miktarı

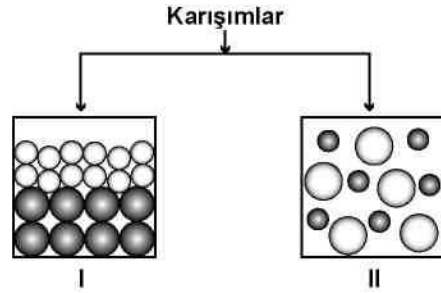
B) Zaman

C) Madde miktarı

D) Özkütle (Yoğunluk)

(22)

Karışımların sınıflandırılması, şemada tanecik modeli ile gösterilmiştir:

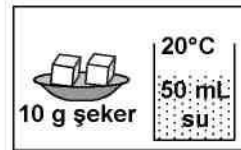


Buna göre, I ve II ile gösterilen tanecik modelleri hangi tür karışımı temsil eder? Neden?

- A) I. model, homojen karışımdır. Çünkü maddelerden biri dibe çökmüştür.
 B) I. model, heterojen karışımdır. Çünkü karışımındaki maddelerin tanecik sayıları farklıdır.
 C) II. model, heterojen karışımdır. Çünkü karışım iki tür madde içerir.
 D) II. model, homojen karışımdır. Çünkü tanecikler karışımın her tarafına eşit dağılmıştır.

(23)

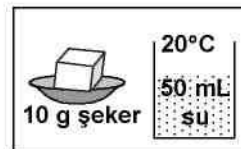
Bazı öğrencilerin aynı miktardaki şekerleri suda çözmek için hazırladıkları düzenekler aşağıda verilmiştir.



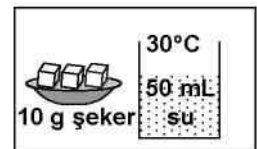
Eren'in düzenegi



İlay'ın düzenegi



Sevde'nin düzenegi



Ece'nin düzenegi

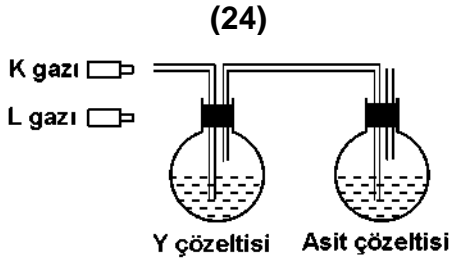
Hangi öğrencinin hazırladığı düzenekte şeker daha kısa sürede çözünür?

A) Eren'in

B) Sevde'nin

C) Ece'nin

D) İlay'ın

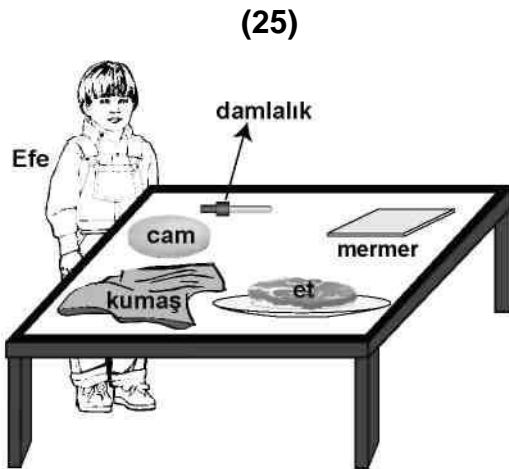


Şekildeki düzeneğe önce K gazı sonra L gazı gönderildiğinde;

- K gazı asit çözeltisi ile,
- L gazı Y çözeltisi ile nötrleşme (nötrleşme) tepkimesi veriyor.

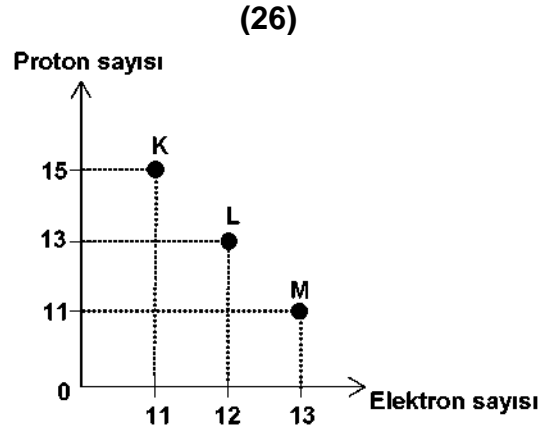
Buna göre aşağıdakilerden hangisi söylenebilir?

- A) Y çözeltisi asidik özelliktedir.
 B) K gazı bazik, L gazı asidik özelliktedir.
 C) K ve L gazı bazik özelliktedir.
 D) L gazı bazik, Y çözeltisi asidik özelliktedir.



Efe, masadaki maddelerin üzerine asit damlattıktan sonra etin, mermerin ve kumaşın tahriş olduğunu, camın ise tahriş olmadığını gözlemliyor. Efe'nin etkinlik sonucu edindiği bilgiye göre, aşağıdakilerden hangisini yapması uygun olmaz?

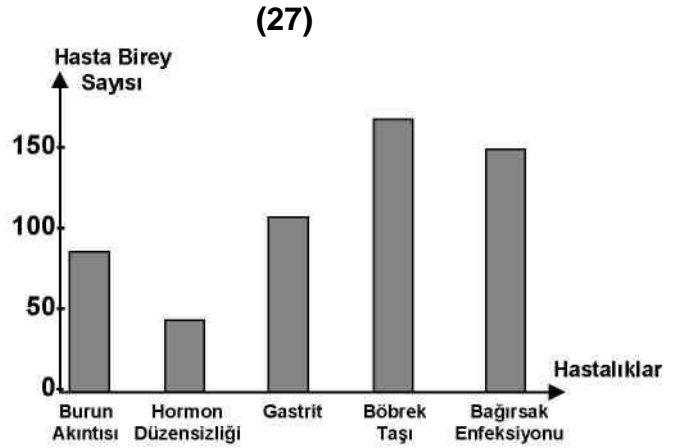
- A) Sirkeyi cam şişede saklaması
 B) Mermer tezgahın üzerinde limon kesmesi
 C) Laboratuvarda çalışırken koruyucu kıyafet kullanması
 D) Tuz ruhu ile banyoyu temizlerken koruyucu eldiven kullanması



Bir atomdan oluşan iyonun yükü, proton sayısı ile elektron sayısının farkına eşittir.

K, L ve M iyonlarının proton ile elektron sayısı arasındaki ilişki grafikteki gibi olduğuna göre, bu iyonların yükleri aşağıdakilerden hangisidir?

	K	L	M
A)	+4	+1	-2
B)	-4	-1	+2
C)	-4	-1	-2
D)	+4	+1	+2



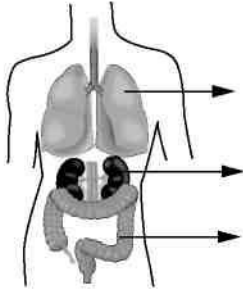
Yukarıdaki grafik bir kasabada bir ay içinde hastaneye giden bireylerdeki hastalıkları ve hasta sayısını göstermektedir.

Buna göre bireylerde aşağıdakilerden hangisiyle ilgili hastalıklar daha fazla görülmektedir?

- A) Denetleyici ve düzenleyici sistem
 B) Solunum sistemi
 C) Sindirim sistemi
 D) Boşaltım sistemi

(28)

Öğretmenin sorduğu bir soruya Ahmet, şekil üzerindeki işaretlenmiş üç organı göstererek doğru cevap vermiştir.

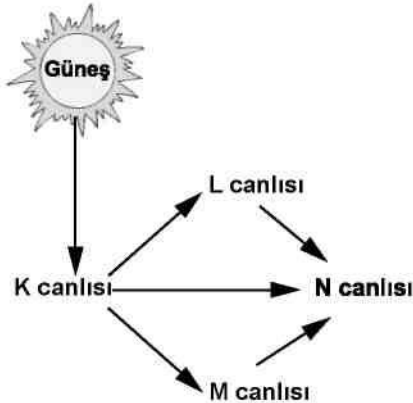


Buna göre, öğretmenin sorduğu soru aşağıdakilerden hangisidir?

- A) Hangi organlar sindirime yardımcı olur?
- B) Hangi organlar solunum sisteminde yer alır?
- C) Kandaki zararlı atıklar hangi organlarda süzülür?
- D) Atık maddelerin vücuttan atılmasında hangi organlar görevlidir?

(29)

Şemada, bazı canlılardaki beslenme ilişkileri gösterilmiştir.



Buna göre, aşağıdakilerden hangisi yanlıştır?

- A) K canlısı, fotosentez yapabilmektedir.
- B) L canlısı, sadece bitkilerle beslenmektedir.
- C) M canlısı, hem bitkilerle hem de etle beslenmektedir.
- D) N canlısı, farklı türden besin kaynaklarını kullanabilmektedir.

(30)



Serpil, birinde nemli, diğerinde kurutulmuş ekmek bulunan iki naylon torbayı gevşek bir şekilde bağlıyor. Beş gün sonra sadece nemli ekmek üzerinde küf mantarının oluştuğunu gözlüyor.

Serpil, bu gözleme göre aşağıdaki sonuçlardan hangisine ulaşır?

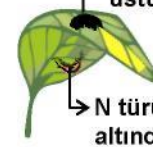
- A) Hava alan tüm yiyecekler çabuk bozulur.
- B) Kuru ortamlar, mantar türleri için uygundur.
- C) Kurutulmuş ekmek daha uzun süre dayanır.
- D) Küf mantarı bölünerek çoğalır.

(31)

K türü, yaprakta delikler açarak beslenir ve yaşar.



M türü, yaprağın üstünde yaşar.



N türü, yaprağın altında yaşar.

L türü, yaprak kenarlarını yiyerek beslenir ve yaşar.

Yukarıdaki şekiller incelendiğinde bazı böcek türlerinin yaşam alanları ile ilgili bilgilere ulaşılmaktadır.

Buna göre aşağıdaki yorumlardan hangisi yapılabilir?

- A) Bitkilerle beslenen ve bir arada yaşayan canlılar arasında iş bölümü vardır.
- B) Aynı ortamda yaşayan iki canlıdan çevre şartlarını en iyi kullanan hayatta kalır.
- C) Aynı ortamda yaşayan canlılar, buldukları ortamı farklı şekillerde kullanır.
- D) Bir arada yaşayan canlılar, birbirine fayda sağlar.

(32)

Sıcakkanlı canlılar, çeşitli davranış şekilleriyle vücut sıcaklıklarını dengede tutmaya çalışır. Sıcak havada dışarıya fazla ısı verecekleri ortamı ararlar. Soğuk havada ise vücut yüzeyini azaltıp hareketlerini artırırlar.

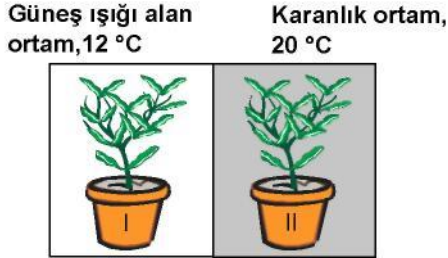
Buna göre sıcakkanlı canlılar, sıcak ve soğuk havada aşağıdakilerin hangisinde verilen davranışları gösterirler?

Sıcak havada

Soğuk havada

- | | |
|--------------------------|-----------------------|
| A) Sıcak yer arama | Hareketlerini azaltma |
| B) Hareketlerini azaltma | Serin yer arama |
| C) Hareketlerini artırma | Kıvrılıp büzülme |
| D) Serin yer arama | Kıvrılıp büzülme |

(33)



Mehmet, güneş ışığının besin yapımına etkisini gözlemek istiyor. Eşit miktarda su verip, aynı tür toprağa ektiği özdeş saksı bitkilerini farklı ortamlara koyarak yukarıdaki gibi deney düzeneği hazırlıyor.

Buna göre Mehmet, aşağıdakilerden hangisini yaparsa deneyden sonuç alır?

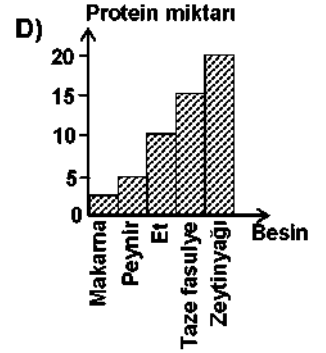
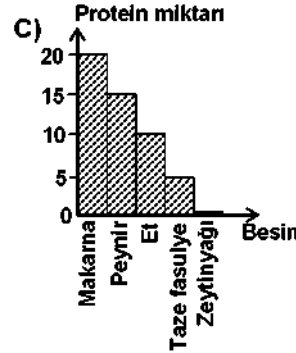
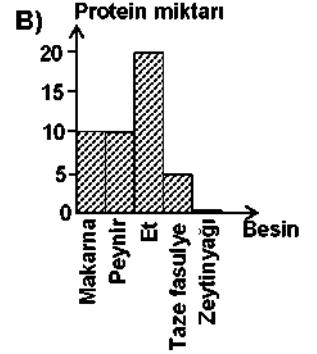
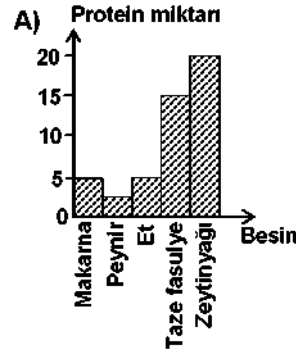
- I. saksının bulunduğu ortamın sıcaklığını 20 °C'a çıkarmalı.
- II. saksıyı sıcaklığı 12 °C olan ışıklı ortama koymalı.
- Bitkilerden birinin yapraklarını azaltmalı.
- Bitkilerden birine daha az su vermeli.

(34)

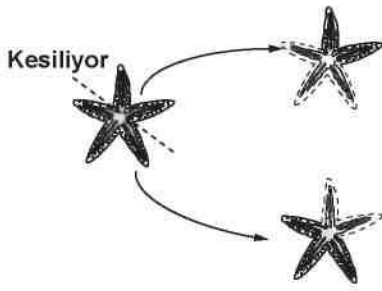
Yemek adı	Protein miktarı
Haşlanmış makarna	* *
Peynirli makarna	* * * *
Etili taze fasulye	* * * * *
Zeytinyağlı taze fasulye	*
Zeytinyağlı makarna	* *

Bazı besinlerden eşit miktarda alınarak yapılan tablodaki yemeklerin protein miktarı "*" ile gösterilmiştir.

Buna göre, yemeklerde kullanılan besinlerin protein miktarı hangi grafikteki gibi olabilir?



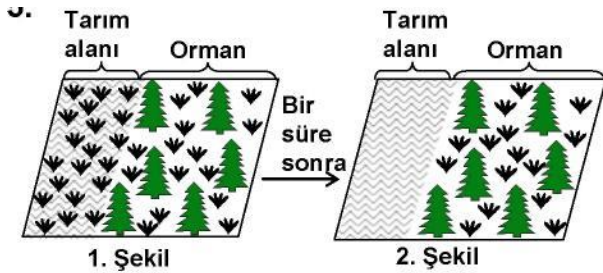
(35)



Yukarıda verilen deniz yıldızındaki yenilenme olayı ile ilgili olarak aşağıdakilerden hangisi söylenebilir?

- A) Eşeyli üremedir.
- B) Mitoz bölünme ile gerçekleşir.
- C) Yavru canlıların genotipi ana canlıdan farklıdır.
- D) Yavru canlılar ana canlıdan daha gelişmiş yapıdadır.

(36)



↙: Yabani ot

Bir bölgede yabani ot, 1. şekildeki gibi yayılma göstermiştir.

Bunun üzerine çiftçiler bu bölgede yabani otları beslenen böcek türünü çoğaltmışlardır.

Bir süre sonra bölgede 2. şekildeki durum gözlemlendiğine göre aşağıdakilerden hangisi söylenemez?

- A) Böcek, yabani otun yayılmasını sınırlamıştır.
- B) Böceğin ormanda yaşamasını engelleyen koşullar vardır.
- C) Bölgenin iklim koşulları, yabani otun yayılmasını engellemiştir.
- D) Yabani ot, böceğin yaşadığı alandan daha geniş alanlara yayılmıştır.

(37)

Baba	■	●
Anne	■	●

Yukarıdaki tabloda ■ ve ● ile gösterilen genlere (alel) sahip anne babanın çocuklarının saç fenotipleri verilmiştir.

Bu verilere göre aşağıdakilerden hangisi doğrudur?

- A) ● geni, her durumda düz saç özelliği kazandırmaktadır.
- B) ■ geni, her durumda kıvrıkcık saç özelliği kazandırmaktadır.
- C) Baba, yalnızca kıvrıkcık saç genine sahiptir.
- D) Anne, yalnızca düz saç genine sahiptir.

(38)



Şekilde verilen azot döngüsünde aşağıdaki olaylardan hangisi gözlenir?

- A) Hayvanların solunumuyla havadan azot aldığı
- B) Hayvan atıklarındaki azotun havaya karıştığı
- C) Bitkilerin yapraklarıyla havaya azot verdiği
- D) Bitkilerin topraktan kökleriyle azot aldığı

(39)

Öğrenciler, sınıftaki bir etkinlik için şekildeki fotoğrafları kullanacaklardır.

Önce



Sonra

Kuzey Kutbu'ndaki bir bölgenin önceki ve sonraki durumunu anlatan bu fotoğraflar, **daha çok** aşağıdaki konulardan hangisiyle ilgilidir?

- A) Küresel ısınma ve sonuçları
- B) Biyolojik çeşitliliğin azalması
- C) Ozon tabakasının zarar görmesi
- D) Su kirliliğinin canlılar üzerine etkileri

(40)

Emel, sınıfta yapılacak bir etkinlik için üzerlerinde Güneş sistemindeki üç gezegene ait bilgilerin olduğu aşağıdaki I, II, III kartlarını hazırlıyor.

<p>◆ Güneş sisteminin en büyük gezegenidir.</p> <p>◆ Güneş'e en yakın beşinci gezegendir.</p>	<p>◆ Güneş sisteminin en küçük gezegenidir.</p> <p>◆ Güneş'e en yakın gezegendir.</p>	<p>◆ Güneş sisteminin ikinci en küçük gezegenidir.</p> <p>◆ Güneş'e en yakın dördüncü gezegendir.</p>
---	---	---

I

II

III

Emel'in hazırladığı bu kartların arka yüzüne yazması gereken gezegen isimleri hangi seçenekte doğru olarak verilmiştir?

- | | | |
|---------------|----------------|-----------------|
| <u> </u> I | <u> </u> II | <u> </u> III |
| A) Jüpiter | Merkür | Mars |
| B) Neptün | Merkür | Dünya |
| C) Satürn | Venüs | Mars |
| D) Jüpiter | Venüs | Merkür |

TEST BİTTİ.

LÜTFEN CEVAPLARINIZI

KONTROL EDİNİZ.

EK-5
ORJİNALLİK RAPORU*

30-Nov-2014 10:35PM 23773 words • 59 matches • 30 sources FAQ

iThenticate® TEZ_AAS
BY AYLIN ALBAYRAK SARI

Quotes Included
Bibliography Excluded **3%**
SİMLAR

Match Overview

1	Internet 124 words crawled on 20-Jun-2014 www.egitimbilimlerinstitusu.hacettepe.edu.tr	1%
2	Internet 98 words crawled on 28-May-2013 acikarsiv.ankara.edu.tr	<1%
3	Internet 81 words crawled on 19-Oct-2011 fenbilimleri.uludag.edu.tr	<1%
4	Internet 50 words crawled on 25-Jul-2012 mersin.mitosweb.com	<1%
5	Publications 39 words DOĞAN, Nuri and ÖĞRETMEN, Tuncay. "Madde Yanlılı ... nı Belirleme Teknikleri Arasında Bir Karşılaştırma", Anı Y	<1%
6	Internet 37 words crawled on 13-Feb-2014 egitimvebilim.ted.org.tr	<1%
7	Publications 32 words Üstün, Emre and Özçimen, Aycan. "TÜRKİYEDE MÜZİK	<1%

FEN BAŞARISINDAKİ ARTIŞIN BELİRLENMESİNDE
MADDE TEPKİ KURAMINA DAYALI DİKEY ÖLÇEKLEME
YÖNTEMLERİNİN KARŞILAŞTIRILMASI

A COMPARISON OF IRT VERTICAL SCALING METHODS
IN DETERMINING OF THE INCREASE IN ACHIEVEMENT
OF SCIENCE EDUCATION

Aylin ALBAYRAK SARI

Hacettepe Üniversitesi

4 Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

PAGE: 1 OF 140

Text-Only Report

*Bu tez orijinalliğinin kontrolü için iThenticate programı ile taranmıştır. Program bu tezin sadece %3 oranında yapılan yayınlarla benzerlik barındırdığını rapor etmiştir.

ÖZGEÇMİŞ

Kişisel Bilgiler

<i>Adı Soyadı</i>	Aylin ALBAYRAK SARI
<i>Doğum Yeri</i>	ANKARA
<i>Doğum Tarihi</i>	1984

Eğitim Durumu

<i>Lise</i>	Ankara Atatürk Anadolu Lisesi	2002
<i>Lisans</i>	Hacettepe Üniversitesi, Eğitim Fakültesi, Fen Bilgisi Öğretmenliği Anabilim Dalı	2006
<i>Yüksek Lisans</i>	Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı	2009
<i>Doktora</i>	Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Eğitim Bilimleri Anabilim Dalı, Eğitimde Ölçme ve Değerlendirme Bilim Dalı	2014
<i>Yabancı Dil</i>	İngilizce, Fransızca	

İş Deneyimi

<i>Çalıştığı Kurumlar</i>	Hacettepe Üniversitesi, Eğitim Fakültesi, Fen Bilgisi Öğretmenliği ABD.	2007-Halen
---------------------------	---	------------

İletişim

<i>e-Posta Adresi</i>	aylinalbayrak@gmail.com
-----------------------	-------------------------

<i>Jüri Tarihi</i>	22.10.2014
--------------------	------------