

**ÇOKLU CEVAPLI MADDE TÜRÜNÜN PUANLANMASINDA
KULLANILAN YÖNTEMLERİN TEST VE MADDE
İSTATİSTİKLERİ AÇISINDAN KARŞILAŞTIRILMASI**

Ömer Doğan

**YÜKSEK LİSANS TEZİ
EĞİTİM BİLİMLERİ ANA BİLİM DALI**

**GAZİ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ**

HAZİRAN, 2020

TELİF HAKKI VE TEZ FOTOKOPİ İZİN FORMU

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

YAZARIN

Adı: Ömer

Soyadı: Doğan

Bölümü: Eğitimde Ölçme ve Değerlendirme

İmza:



Teslim Tarihi: 30/06/2020

TEZİN

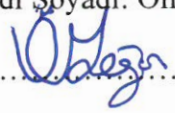
Türkçe Adı: Çoklu Cevaplı Madde Türünün Puanlanmasında Kullanılan Yöntemlerin Test ve Madde İstatistikleri Açısından Karşılaştırılması

İngilizce Adı: Comparison of the Methods Used in the Scoring of the Multiple Response Item Type in terms of Test and Item Statistics

ETİK İLKELERE UYGUNLUK BEYANI

Tez yazma sürecinde bilimsel ve etik deęerlere uyduęumu, yararlanduęım tüm kaynakları kaynak gösterme ilkelerine uygun olarak kaynakçada belirttięimi ve bu bölümler dışındaki tüm ifadelerin şahsıma ait olduęunu beyan ederim.

Yazar Adı Soyadı: Ömer DOĞAN

İmza:..........

JÜRİ ONAY SAYFASI

Ömer DOĞAN tarafından hazırlanan “Çoklu Cevaplı Madde Türünün Puanlanmasında Kullanılan Yöntemlerin Test ve Madde İstatistikleri Açısından Karşılaştırılması” adlı tez çalışması aşağıdaki jüri tarafından oy birliği / oy çokluğu ile Gazi Üniversitesi Eğitim Bilimleri Anabilim Dalı’nda Yüksek Lisans tezi olarak kabul edilmiştir.

Danışman: Prof. Dr. İsmail KARAKAYA

(Eğitim Bilimleri Anabilim Dalı, Gazi Üniversitesi)

Başkan: Prof. Dr. Hakan ATILGAN

(Eğitim Bilimleri Anabilim Dalı, Ege Üniversitesi)

Üye: Doç. Dr. Dilara BAKAN KALAYCIOĞLU

(Eğitim Bilimleri Anabilim Dalı, Gazi Üniversitesi)

Tez Savunma Tarihi: 30/06/2020

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

Prof. Dr. Selma YEL

Eğitim Bilimleri Enstitüsü Müdürü



Çok Değerli Aileme

TEŞEKKÜR

Bu tezin yazılmasında emeđi geen, beni destekleyip teze ve şahsıma birok katkı sunan, bildiklerini aktarıp bizi bilinmeyenleri arařtırmaya teřvik eden danıřmanım Prof. Dr. İsmail KARAKAYA'ya teřekkürü bir bor bilirim.

Lisansüstü eğitime bařladığım ilk günden beri her türlü desteđi veren ve benim için rehber olan ilk tezimin danıřmanı Do. Dr. Aynur BOZKURT BOSTANCI'ya bana kattıkları ve ilgisi için ok teřekkür ederim.

Her iki yüksek lisans sürecinde de akademik olan veya olmayan her konuda yanımda olan kıymetli arkadaşlarım Ali TOSUN ve Yusuf GİDİř ve Dr. Öğr. Üyesi Ümit KAHRAMAN'a ok teřekkür ederim.

Tez alışması sürecinde başarı testlerini dikkatli bir şekilde cevaplayan tüm öğrencilere ve bana öğrencilere ulaşma fırsatı veren tüm öğretmenlere teřekkür ederim.

Lisansüstü eğitim sürecimin dayanakları olan ve bu süreçte kendilerine yeterince vakit ayıramadığım, hayattaki yegane varlıklarım olan ok kıymetli aileme sabırları, destekleri, ilgileri ve anlayışları için ok ok teřekkür ederim.

**ÇOKLU CEVAPLI MADDE TÜRÜNÜN PUANLANMASINDA
KULLANILAN YÖNTEMLERİN TEST VE MADDE
İSTATİSTİKLERİ AÇISINDAN KARŞILAŞTIRILMASI
(Yüksek Lisans Tezi)**

**Ömer Doğan
GAZİ ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
Haziran 2020**

ÖZ

Bu araştırmanın amacı çoklu cevaplı madde türünün puanlanma yöntemlerini incelemektir. Bu amaç doğrultusunda çoklu cevaplı maddelerden oluşan başarı testlerini puanlamada kullanılan sekiz puanlama yöntemi arasında; ortalama test güclüğü, madde güclüğü, madde ayırt ediciliği ve güvenilirliklerine ait farklılıklar incelenmiştir. Bu sebeple araştırmanın modeli temel araştırmadır. Araştırmanın çalışma grubunu olasılığa dayalı olmayan örnekleme yöntemlerinden amaçlı örnekleme yoluyla seçilen 2019-2020 eğitim öğretim yılı ikinci döneminde Uşak'ta öğrenim gören ve Liselere Giriş Sınavı (LGS)'na hazırlanan ortaokul sekizinci sınıf öğrencileri oluşturmaktadır. Araştırmada verilerin toplanması amacıyla; araştırmacı tarafından geliştirilen Türkçe ve matematik başarı testleri kullanılmıştır. Her bir başarı testi beş seçenekli yirmi maddeden oluşmaktadır. Bu testlerden Türkçe testi 277 ve matematik testi ise 263 sekizinci sınıf öğrencisine araştırmacı tarafından uygulanmıştır. Test uygulamaları sonrasında 259 öğrenciye beş maddeden oluşan anket uygulaması yapılmıştır. Toplanan verilerin analizinde betimsel istatistikler, z testi, ANOVA ve test ile madde istatistiklerini hesaplamak için uygun istatistiksel yöntemler kullanılmıştır. Araştırma sonucunda puanlama yöntemleri arasında güvenilirlik bağlamında anlamlı bir fark olmadığı, test ortalama güçlükleri bakımından düzeltme uygulamasının yapıldığı çoklu puanlama yöntemleri ile İkili Puanlama (Dichotomous-İP) yönteminin, düzeltme uygulaması yapılmayan puanlama yöntemlerine göre güçlük değerlerinin daha düşük olduğu bulunmuştur. Madde güçlük indeksleri incelendiğinde indeks değerleri en az olan puanlama yöntemlerinin İkili Puanlama ve Morgan Algoritması olduğu en yüksek ise Çoklu

Doğru/Yanlış (Multiple True/False-ÇDY) ve Pozitif Sayım (Positive Count-PS) puanlama yöntemleri olduğu görülmüştür. Madde ayırt edicilik indekslerine bakıldığında matematik testi için en yüksek ayırt edicilik indeksleri Çoklu Doğru/Yanlış, Tuzaklı Çoklu Puanlama (Polytomous Trapdoor-TÇC) ve İkili Puanlama Yöntemleri iken Türkçe testi için Çoklu Doğru/Yanlış ve İkili Puanlama Yöntemi şeklindedir. Test uygulamalarına katılan sekizinci sınıf öğrencilerinden gönüllü olanlara uygulanan çoklu cevaplı ve çoktan seçmeli maddelerin karşılaştırılmasının yapıldığı anketten elde edilen sonuçlara göre öğrenciler çoklu cevaplı maddelerin yer aldığı testte çoktan seçmeli maddelere göre daha yüksek puan alacaklarını ve kişiler arası ayrımın daha iyi sağlanacağını düşünmektedirler. Ayrıca bildiklerini daha iyi yansıtabileceklerini düşündükleri ve cevaplaması daha kolay olan madde türünün çoktan seçmeli madde türü olduğunu belirtip LGS’de çoktan seçmeli maddelerin yer almasını tercih ettikleri sonuçlarına ulaşmıştır.



Anahtar Kelimeler: Çoklu cevaplı madde türü, kısmi puanlama, en az bir doğru cevap, puanlama yöntemleri.

Sayfa Adedi: 126

Danışman: Prof. Dr. İsmail KARAKAYA

COMPARISON OF THE METHODS USED IN THE SCORING OF THE MULTIPLE RESPONSE ITEM TYPE IN TERMS OF TEST AND ITEM STATISTICS

(M.S. Thesis)

Ömer Doğan

GAZI UNIVERSITY

GRADUATE SCHOOL OF EDUCATIONAL SCIENCES

June 2020

ABSTRACT

The purpose of this study is to examine the scoring methods of the multiple response item type. For this purpose, the differences between average test difficulty, item difficulty, item discrimination and reliability among the eight scoring methods used in scoring achievement tests consisting of multiple answers items were examined. For this reason, the model of the research is basic research. The study group of the research is the secondary school eighth grade students who have been studying in Uşak in the second semester of the 2019-2020 academic year and prepared for the High School Entrance Exam (LGS) and the group was formed by purposive sampling from non-probability sampling methods. In order to collect data in the research; Turkish and mathematics achievement tests developed by the researcher were used. Each achievement test consists of twenty items with five options. Among these tests, Turkish test 277 and mathematics test were applied to 263 eighth grade students by the researcher. After the study group had the test, a questionnaire with five items was applied to 259 students. In the analysis of the collected data, descriptive statistics, z test, ANOVA and appropriate statistical methods were used to calculate the test and item statistics. As a result of the research, it was found that there is no significant difference between the scoring methods in terms of reliability. In terms of test average difficulties, multiple scoring methods with correction application and Dichotomous scoring method were found to have lower difficulty values than scoring methods without correction. When item difficulty indexes are analyzed, it is seen that the scoring methods with the lowest index values are Dichotomous Scoring and Morgan Algorithm, while the Multiple True/False and Positive Count Scoring

methods are the highest. Considering item discrimination indices, the highest discrimination indices for mathematics test are Multiple True/False, Polytomous Trapdoor and Dichotomous Scoring Methods and the highest discrimination indices for the Turkish test are Multiple True/False and Dichotomous Scoring Method. A questionnaire was applied to compare the multiple response and multiple-choice items from the eighth grade students who participated in the tests to the volunteers. According to the results obtained from the questionnaire, the students think that they will score higher than the multiple choice items by doing the test with multiple response items. Students think that interpersonal discrimination with multiple response items will be better provided. In addition, they stated that the type of item that they think could reflect what they know better and which is easier to answer is a multiple choice item type and it was concluded that they preferred multiple choice items in LGS.

Key Words: Multiple response item type, partial scoring, at least one correct answer, scoring methods.

Page Number:126

Supervisor: Prof. Dr. İsmail KARAKAYA

İÇİNDEKİLER

TELİF HAKKI VE TEZ FOTOKOPİ İZİN FORMU	i
ETİK İLKELERE UYGUNLUK BEYANI	ii
JÜRİ ONAY SAYFASI.....	iii
TEŞEKKÜR.....	v
ÖZ	vi
ABSTRACT	viii
TABLolar LİSTESİ	xiii
SİMGELER VE KISALTMALAR LİSTESİ.....	xvii
BÖLÜM I	1
GİRİŞ.....	1
1.1. Problem Durumu.....	1
1.2. Kuramsal Çerçeve	8
1.2.1. Çoktan Seçmeli Madde Türü.....	8
1.2.2. Çoklu Cevaplı Madde Türü.....	11
1.3. Araştırmanın Önemi	25
1.4. Araştırmanın Amacı	28
1.5. Araştırma Soruları.....	28
1.6. Araştırmanın Sınırlılıkları	29
1.7. Varsayımlar	29

BÖLÜM II.....	30
İLGİLİ ARAŞTIRMALAR.....	30
2.1. Çoktan Seçmeli Madde Türü İle İlgili Araştırmalar	30
2.2. Çoklu Cevaplı Madde Türü İle İlgili Araştırmalar	33
BÖLÜM III	37
YÖNTEM.....	37
3.1. Araştırma Modeli	37
3.2. Çalışma Grubu	37
3.3. Veri Toplama Araçları.....	38
3.4. Verilerin Toplanması	56
3.5. Veri Analizi	57
BÖLÜM IV	60
BULGULAR	60
4.1. Birinci Alt Probleme İlişkin Bulgular	60
4.2. İkinci Alt Probleme İlişkin Bulgular	62
4.3. Üçüncü Alt Probleme İlişkin Bulgular	64
4.4. Dördüncü Alt Probleme İlişkin Bulgular	71
4.5. Beşinci Alt Probleme İlişkin Bulgular	77
BÖLÜM V	83
SONUÇ, TARTIŞMA ve ÖNERİLER	83
5.1. Sonuçlar ve Tartışma	83
5.1.1 Birinci Alt Probleme İlişkin Sonuçlar ve Tartışma	83
5.1.2. İkinci Alt Probleme İlişkin Sonuçlar ve Tartışma.....	84
5.1.3. Üçüncü Alt Probleme İlişkin Sonuçlar ve Tartışma.....	85
5.1.4. Dördüncü Alt Probleme İlişkin Sonuçlar ve Tartışma	86
5.1.5. Beşinci Alt Probleme İlişkin Sonuçlar ve Tartışma	87

5.1.6. Genel Sonular ve Tartışma	88
5.2. Öneriler	89
KAYNAKLAR.....	91
EKLER.....	99
EK 1. Araştırma İzni Belgesi	100
EK 2. Çoklu Cevaplı Madde Türü ile Hazırlanan Matematik Testi Maddeleri.....	101
EK 3. Çoklu Cevaplı Madde Türü ile Hazırlanan Türke Testi Maddeleri	113
EK 4. Anket Formu.....	124



TABLÖLAR LİSTESİ

Tablo 1. <i>Dengeli Puanlama Yönteminde Kullanılan Kısaltmalar ve Anlamları</i>	22
Tablo 2. <i>Örnek Uygulamadaki Beş Maddeye Öğrencinin İşaretlediği Seçenekler ve Doğru Seçenekler</i>	23
Tablo 3. <i>Çoklu Cevaplı Madde Türüne Ait Puanlama Yöntemleri İçin Genel Bilgiler ve Örnek Uygulamada Öğrencinin Alacağı Puanlar</i>	24
Tablo 4. <i>Çoklu Cevaplı Madde Türü ile Oluşturulan Matematik Testinin Puanlama Yöntemlerine Göre Betimsel İstatistikleri</i>	25
Tablo 5. <i>Çoklu Cevaplı Madde Türü ile Oluşturulan Matematik Testine Ait Maddelerin Güçlük İndekslerinin Puanlama Yöntemlerine Göre Gösterimi</i>	26
Tablo 6. <i>Çoklu Cevaplı Madde Türü ile Oluşturulan Matematik Testinin Puanlama Yöntemlerine Göre Test Güvenirlikleri</i>	27
Tablo 7. <i>Çoklu Cevaplı Madde Türü ile Oluşturulan Matematik Testine Ait Maddelerin Ayırt Edicilik İndekslerinin Puanlama Yöntemlerine Göre Gösterimi</i>	28
Tablo 8. <i>Çoklu Cevaplı Madde Türü ile Oluşturulan Türkçe Testinin Puanlama Yöntemlerine Göre Betimsel İstatistikleri</i>	29
Tablo 9. <i>Çoklu Cevaplı Madde Türü ile Oluşturulan Türkçe Testine Ait Maddelerin Güçlük İndekslerinin Puanlama Yöntemlerine Göre Gösterimi</i>	30
Tablo 10. <i>Çoklu Cevaplı Madde Türü ile Oluşturulan Türkçe Testinin Puanlama Yöntemlerine Göre Test Güvenirlikleri</i>	32
Tablo 11. <i>Çoklu Cevaplı Madde Türü ile Oluşturulan Türkçe Testine Ait Maddelerin Ayırt Edicilik İndekslerinin Puanlama Yöntemlerine Göre Gösterimi</i>	33

Tablo 12. Çoklu Cevaplı Madde Türü ile Oluşturulan II. Uygulamaya Ait Matematik Testinin Puanlama Yöntemlerine Göre Betimsel İstatistikleri.....	34
Tablo 13. Çoklu Cevaplı Madde Türü ile Oluşturulan II. Uygulamaya Ait Matematik Testi Madde Güçlük İndekslerinin Puanlama Yöntemlerine Göre Gösterimi.....	35
Tablo 14. Çoklu Cevaplı Madde Türü ile Oluşturulan II. Uygulamaya Ait Matematik Testinin Puanlama Yöntemlerine Göre Test Güvenirlikleri.....	36
Tablo 15. Çoklu Cevaplı Madde Türü ile Oluşturulan II. Uygulamaya Ait Matematik Testine Ait Maddelerin Ayırt Edicilik İndekslerinin Puanlama Yöntemlerine Göre Gösterimi.....	37
Tablo 16. Çoklu Cevaplı Madde Türü ile Oluşturulan II. Uygulamaya Ait Türkçe Testinin Puanlama Yöntemlerine Göre Betimsel İstatistikleri.....	38
Tablo 17. Çoklu Cevaplı Madde Türü ile Oluşturulan II. Uygulamaya Ait Türkçe Testi Madde Güçlük İndekslerinin Puanlama Yöntemlerine Göre Gösterimi.....	39
Tablo 18. Çoklu Cevaplı Madde Türü ile Oluşturulan II. Uygulamaya Ait Türkçe Testinin Puanlama Yöntemlerine Göre Test Güvenirlikleri.....	40
Tablo 19. Çoklu Cevaplı Madde Türü ile Oluşturulan II. Uygulamaya Ait Türkçe Testine Ait Maddelerin Ayırt Edicilik İndekslerinin Puanlama Yöntemlerine Göre Gösterimi.....	41
Tablo 20. Çoklu Cevaplı Madde Türü ile Oluşturulan Esas Uygulamaya Ait Matematik Testinin Puanlama Yöntemlerine Göre Betimsel İstatistikleri.....	42
Tablo 21. Çoklu Cevaplı Madde Türü ile Oluşturulan Esas Uygulamaya Ait Türkçe Testinin Puanlama Yöntemlerine Göre Betimsel İstatistikleri.....	43
Tablo 22. Deneme Uygulamalarına Ait Frekans Değerleri.....	44
Tablo 23. Esas Uygulamaya Ait Frekans Değerleri.....	45
Tablo 24. Farklı Puanlama Yöntemleriyle Elde Edilen Test Puanlarına Ait Ortalama Güçlük Değerleri.....	46
Tablo 25. Farklı Puanlama Yöntemleriyle Elde Edilen Matematik Testi Puanlarına Ait Varyans Analizi Sonuçları.....	47
Tablo 26. Farklı Puanlama Yöntemleriyle Elde Edilen Türkçe Testi Puanlarına Ait Varyans Analizi Sonuçları.....	48

Tablo 27. <i>Farklı Puanlama Yöntemleriyle Elde Edilen Test Puanlarının Güvenirlik Katsayıları</i>	49
Tablo 28. <i>Farklı Puanlama Yöntemleriyle Elde Edilen Test Puanlarının En Yüksek ve En Düşük Güvenirlik Katsayılarına Ait Z İstatistiği Sonuçları</i>	50
Tablo 29. <i>Farklı Puanlama Yöntemlerinden Hesaplanan Matematik Testine Ait Madde Güçlük İndeksleri</i>	51
Tablo 30.1 <i>Farklı Puanlama Yöntemleriyle Hesaplanan Matematik Testine Ait Madde Güçlük İndeksleri İçin Varyans Analizi Sonuçları (1-10. Maddeler)</i>	52
Tablo 30.2 <i>Farklı Puanlama Yöntemleriyle Hesaplanan Matematik Testine Ait Madde Güçlük İndeksleri İçin Varyans Analizi Sonuçları (11-20. Maddeler)</i>	53
Tablo 31. <i>Farklı Puanlama Yöntemlerinden Hesaplanan Türkçe Testine Ait Madde Güçlük İndeksleri</i>	54
Tablo 32.1 <i>Farklı Puanlama Yöntemlerinden Hesaplanan Türkçe Testine Ait Madde Güçlük İndeksleri İçin Varyans Analizi Sonuçları (1-10. Maddeler)</i>	55
Tablo 32.2 <i>Farklı Puanlama Yöntemlerinden Hesaplanan Türkçe Testine Ait Madde Güçlük İndeksleri İçin Varyans Analizi Sonuçları (11-20. Maddeler)</i>	56
Tablo 33. <i>Farklı Puanlama Yöntemlerinden Hesaplanan Matematik Testine Ait Madde Ayırt Edicilik İndeksleri</i>	57
Tablo 34. <i>Farklı Puanlama Yöntemlerinden Hesaplanan Matematik Testine Ait Madde Ayırt Edicilik İndekslerinin Karşılaştırılmasına Ait Sonuçlar</i>	58
Tablo 35. <i>Farklı Puanlama Yöntemlerinden Hesaplanan Türkçe Testine Ait Madde Ayırt Edicilik İndeksleri</i>	59
Tablo 36. <i>Farklı Puanlama Yöntemlerinden Hesaplanan Türkçe Testine Ait Madde Ayırt Edicilik İndekslerinin Karşılaştırılmasına Ait Sonuçlar</i>	60
Tablo 37. <i>Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anket Maddeleri ve Bu Maddelere Verilen Yanıtlara Göre Oluşan Frekans ve Yüzde Dağılımı</i>	61
Tablo 38. <i>Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 1. Maddesine Ait Gözlenen ve Beklenen Dağılım Tablosu</i>	63

Tablo 39. Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 1. Maddesine Ait Kay Kare Testinin Sonucu.....	64
Tablo 40. Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 2. Maddesine Ait Gözlenen ve Beklenen Dağılım Tablosu.....	65
Tablo 41. Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 2. Maddesine Ait Kay Kare Testinin Sonucu.....	67
Tablo 42. Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 3. Maddesine Ait Gözlenen ve Beklenen Dağılım Tablosu.....	68
Tablo 43. Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 3. Maddesine Ait Kay Kare Testinin Sonucu.....	69
Tablo 44. Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 4. Maddesine Ait Gözlenen ve Beklenen Dağılım Tablosu.....	71
Tablo 45. Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 4. Maddesine Ait Kay Kare Testinin Sonucu.....	72
Tablo 46. Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 5. Maddesine Ait Gözlenen ve Beklenen Dağılım Tablosu.....	73
Tablo 47. Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 5. Maddesine Ait Kay Kare Testinin Sonucu.....	74

SİMGELER VE KISALTMALAR LİSTESİ

LGS	Liselere Giriş Sınavı
K-R	Kuder-Richardson
KTK	Klasik Test Kuramı
MTK	Madde Tepki Kuramı
MEB	Milli Eğitim Bakanlığı
ÖSYM	Ölçme Seçme ve Yerleştirme Merkezi Başkanlığı
YGS	Yüksek Öğretime Geçiş Sınavı
LYS	Lisans Yerleştirme Sınavı
SBS	Seviye Belirleme Sınavı
MCSE	Microsoft Sertifikalı Sistem Mühendisliği Sertifikasyon Programı
GPA	Lisans Not Ortalaması
ÇDY	Çoklu Doğru/Yanlış
İP	İkili Puanlama
PS	Pozitif Sayım
TÇC	Tuzaklı Çoklu Cevaplı
NP	Negatif Puanlama
RA	Ripkey Algoritması
MA	Morgan Algoritması
DP	Dengeli Puanlama

BÖLÜM I

GİRİŞ

Giriş kısmında problem durumu, kuramsal çerçeve, araştırmanın önemi, araştırmanın amacı, araştırma soruları, araştırmanın sınırlılıkları, araştırmanın varsayımları ve araştırmayla bağlantılı çalışmalara yer verilmiştir.

1.1. Problem Durumu

Modern çağın gelişmiş ve gelişmekte olan pek çok ülkesinde sosyal, kitlesel, iktisadi, teknolojik ve daha başka birçok alanda önemli dönüşümler yaşanmaktadır. Dönüşüm ve değişimle yeni bilgiler ortaya çıkarken var olanlar güncellenmekte, olaylar ve buna bağlı sonuçlar birden çok boyutla açıklanmaktadır. Bu değişime paralel olarak eğitimin ve eğitimin verildiği kurumların amacı da değişmektedir. Günümüzde bu amaç, bireyleri sürekli bir değişim içinde olan dünyanın gereksinimlerine yanıt verebilecek, değişime uyum sağlayabilecek ve alternatif çözümler yapabilecek şekilde yetiştirmektir.

Eğitim kavramına yönelik farklı tanımlar ve bakış açısına göre farklı yorumlamalar yapılmaktadır. Eğitim, kişinin davranışlarında kendi yaşantıları yoluyla kasıtlı olarak istedik yönde meydana gelen kalıcı izli değişiklik oluşturma süreci olarak tanımlanır. Bunun yanında eğitim, toplumun refahı ve ilerlemesi için yapılan hizmettir ve bu hizmet, bir bütünlük ve bir sistem işidir. Bir sistem demekle belli amaçlara ulaşabilmek için öğelerin işbirliği içinde çalışarak, bir arada, bir bütünü oluşturacak şekilde organize edilmeleri kastedilmektedir. Her sistem, gerçekleştirdiği etkinliklerin amaca ne derece hizmet edip etmediğini saptayabilmek için öncelikle gerçekleştirdiği etkinlikler hakkında bilgiler toplar yani ölçme (gözlem) yapar ve topladığı gözlem sonuçlarından yararlanarak bazı çıkarımlara,

kararlara ulaşmaya çalışır. Bu sebeple gözlemler yapma ve gözlem sonuçlarına dayalı doğru kararlar verebilmek için, ölçme ve değerlendirmenin bilinmesi, ilkelerinin ve yaklaşımlarının tanınması ve bu ilkelerin doğru biçimde uygulanması gerekir (Ertürk, 1972, s. 12; Kutlu, 2016; Tan, 2010, s. 5).

Ölçme bir varlığın ilgilenilen niteliğinin belirli kurallara göre niteliğin doğasına uygun araçlarla gözlemlenerek (sayarak, tartarak, gruplayarak vb.) elde edilen gözlem sonuçlarının sayı, sıfat veya sembollerle eşleştirilmesi (senkronize etme) işidir. Ölçme için belli bir nesnenin ya da nesnelerin belli bir özelliğe sahip olup olmadığının, sahipse sahip oluş derecesinin gözlenip gözlem sonuçlarının sembollerle ve özellikle sayı sembolleriyle ifade edilmesi, ayrıca herhangi bir özelliği (nesne, olay, durum), o özelliğin niteliğine uygun bir araçla karşılaştırarak, sonucu aracın birimi cinsin ifade etme işi şeklinde tanımlar da mevcuttur. Ölçme işlemi sonunda nesnelerin, olayların, durumların belli bir nitel ve nicel özelliğe sahip oluş dereceleri belirlenmeye çalışılır. Ölçme bize bir şeyin niceliğine ilişkin bilgi verir; bu bilginin yeterli olup olmadığı ve istenilen türden olup olmadığı ise bir değerlendirme sorunudur (Kutlu, 2016; Tekin, 2014, s. 26; Turgut, 1988; Turgut & Baykul, 2013, s. 102; Yıldırım, 1983).

Eğitimde ölçme ve değerlendirme gereksinimi kavramın kendisi kadar eskidir. İnsanoğlunun insanı tanıma, eğitme ve bu amaçla ölçme ve değerlendirme gayretleri milattan önceki yıllara kadar gider (Baykul, 2011). Bu süreçte her zaman eğitimin ürünü, bireylerin davranışları olmuştur. Bireylerde oluşması beklenen davranış değişikliği, uygulanan eğitimin başarısını ya da başarısızlığını gösterir. Bu sebeple davranış değişikliklerinin ölçülmesi ve değerlendirilmesi gerekir. Ayrıca bireyin başarısının, öğrenmedeki güçlü ve zayıf yönlerinin tanımlanmasında ölçme ve değerlendirmeden faydalanılmaktadır.

Eğitim bir süreç işidir ve eğitim süreci ile bu süreci etkileyen faktörleri doğru bir şekilde ölçmek ve değerlendirmek, beklenen bir hedefe ulaşıldığına ya da bu hedefe ulaşabilmek için daha ne kadar çalışılması gerektiğine karar vermek ancak geçerli ve güvenilir ölçme sonuçları ile olanaklıdır. Böyle ölçme sonuçları elde etmek ve eğitimle kazandırılmak istenen yeni davranışların beklenen düzeyde öğrenilmiş olup olmadığına karar verebilmek için bazı koşulların yerine getirilmesi gerekir. Bu sebeple öncelikle ölçülecek davranışların geçerliği ve güvenilirliği yeterli derecede yüksek olan ölçme araçlarıyla ölçülmesine ihtiyaç vardır. Kullanılacak ölçme araçlarının belirlenmesinde; öğrencilerin yetenekleri, gelişim düzeyleri ve gelişim potansiyelleri ile öğretim hedefleri göz önünde bulundurulmalı, bu

amaçlara en iyi hizmet edecek ölçme araçları seçilmelidir. Bireysel farklılıklar gerçeğinden dolayı bütün öğrencileri kapsayan, bütün öğrenciler için genel geçer, tek tip bir ölçme ve değerlendirme yönteminden söz edilmesi uygun olmaz. Bu sebeple öğrencinin akademik gelişimi tek bir yöntemle veya teknikle ölçülüp değerlendirilmemelidir (Kulm, 2013; Milli Eğitim Bakanlığı (MEB), 2018; Özçelik, 1998; Umay, 1997).

Eğitim ve psikolojideki ölçmelerde kullanılan soru, özellikle bilişsel davranışların ölçülmesinde, doğrudan gözlenemeyen davranışların ortaya çıkarılması için cevaplayıcılara verilen uyarıcılardır. Bu uyarıcıları kullanarak herhangi bir davranışı yoklamak için sadece belli bir türü sürekli kullanmak yerine davranışı ortaya çıkarmaya en uygun türü seçme yoluna gidilmelidir (Baykul, 2015, s. 289). Ölçme aracı doğru seçilmezse elde edilen bilgilerle yapılan değerlendirme hatalı olabilir ve testin geçerliği de bu durumdan olumsuz etkilenir. MEB'in (2018) farklı madde türleri kullanılması konusundaki görüşü şöyledir:

...Hiçbir insan bir başkasının birebir aynısı değildir. Bu sebeple öğretim programlarının ve buna bağlı olarak ölçme ve değerlendirme sürecinin “herkese uygun”, “herkes için geçerli ve standart olması” insanın doğasına terstir. Bu sebeple ölçme ve değerlendirme sürecinde azami çeşitlilik ve esneklik anlayışıyla hareket edilmesi şarttır. Öğretim programı, ölçme sürecinde kullanılabilecek ölçme araç ve yöntemleri açısından uygulayıcılara kesin sınırlar çizmez, sadece yol gösterir. Ancak tercih edilen ölçme ve değerlendirme araç ve yönteminde, gereken teknik ve akademik standartlara uyulmalıdır.

Bir eğitim sisteminde öğrenciler hakkında verilen kararların çokluğuna ve çeşitliliğine paralel olarak, çok çeşitli bilgi edinme yollarından yararlanılmaktadır. Bir ders bazında öğrencilere kazandırılmak istenen davranışlar ve öğrencilerin bunları kazanma düzeyleri ise başarı testleriyle ölçülür (Koç, 2009; Tan, 2010, s. 12). Bir ölçme aracının amaca hizmet edebilmesi için bazı yeterliklere sahip olması gerekir. Bunlar güvenilirlik, geçerlik ve kullanışlılıktır.

İyi bir ölçme aracı, özelliği ölçülen varlık veya olayların bu özelliğinde bir değişme olmadıkça onları hep aynı sıraya koyan ölçüler vermelidir. Ölçme aracı neyi ölçmeyi amaçlıyorsa onu kararlı bir şekilde ölçmelidir. Bu durum ölçme aracının güvenilirliğini gösterir (Özçelik, 2013, s. 123). Ölçme aracının güvenilirliği testten, cevaplayıcıdan ve bu ikisinin dışındaki bazı faktörlerden etkilenir. Bunlardan bazıları testin uygulandığı grubun homojenliği, testin uzunluğu, verilen cevaplama süresi, öğrencilerin teste olan tutumları, sınav şartları vb. hususlardır (Baykul, 2015, s. 185). Ayrıca objektif puanlama ve duyarlı bir ölçme aracı oluşturma testin güvenilirliğini etkileyen önemli etkenlerdendir. Bu sebeple hem kısmi puanlama imkanı sağlayıp hem de objektif olarak puanlanabilen bir madde türünden oluşturulan başarı testinin, kısmi puanlama yapılmayan bir testle güvenilirliklerinin

karşılaştırılıp aralarındaki farkın incelenmesi güvenilir test oluşturma bağlamında faydalı çıkarımlar sağlayabilir.

Bir ölçme aracının ölçmeyi amaçladığı özelliği başka faktörlerle karıştırmadan ölçme derecesi o ölçme aracının geçerliğini gösterir. Geçerlik, test puanlarına dayanarak yapılan çıkarımların uygunluğuna, anlamlılığına ve faydalılığına işaret eder ve test geçerliği bu tür çıkarımları desteklemek için kanıt toplama sürecidir. Ölçme aracından elde edilen puanlara aracın ölçmek istediği niteliğin dışındaki niteliklerin karışması (dili kullanma becerisi, öğrencinin kişiliği, kültür, cinsiyet vb.) o ölçme aracından elde edilecek ölçümlerin geçerliğini düşürür (Başol, 2012, s. 136; Özbek, 2010, s. 67; Tan, 2010, s. 34). Bu istenmeyen niteliklere seçilen ölçme aracının türüne göre yazı güzelliği, yazma becerisi, okuma hızı, test kurnazlığı ve şans başarısı gibi etkenler de eklenebilir. Bu etkenlerden şans başarısı, kişinin bir maddenin doğru cevabını tamamen tahminle veya seçenek eleyerek şans yardımıyla bulmasıyla o sorudan puan almasına denir. Şans başarısı sebebiyle testten elde edilen puanlara hata karışmakta ve bu hata öncelikle puanların geçerliğini düşürmektedir.

Baykul'a (2015, s.341) göre şans başarısı, maddeyi kolaylaştırıp madde güçlük indeksini düşürür. Bu düşüş farklı seçenek sayılarına göre farklı değerlerde olmaktadır. Şans başarısı maddeler arası kovaryansı küçültür ve bunun yanında testin güvenilirliğini (iki yarı güvenilirliği, eşdeğer formlar ve KR (Kuder-Richardson) güvenilirliklerinin hepsinde) düşürür. Ayrıca şans başarısı test geçerliğini yordayıcı ve ölçütün ortalama ve varyanslarının şans başarısından etkilenme dereceleri ile maddelerin şansla doğru cevaplandırılma olasılıklarına göre küçültür. Turgut (1971, s. 48), şans başarısının geçerliğe ve güvenilirliğe olan etkisini düşürmek için üç önlemden söz edilebileceğini belirtir. Bunlar sırasıyla puanları şans başarısından arıtma (düzeltme yapma), yönergede açıklama yapma ve madde yapısını değiştirme olarak ifade edilebilir.

Tek doğru cevaplı çoktan seçmeli maddelerde düzeltme formülü, yanlış cevap sayısının seçenek sayısının bir eksiğine bölünmesi ile elde edilen sonucun doğru cevap sayısından çıkarılması şeklinde işler. Fakat bazı durumlarda şans başarısından elde edilen puanı azaltmak için uygulanan bu formül aşırı düzeltme yaparak bireyin test puanının sıfırın altına düşmesine sebep olabilir. Bu durum önemli bir eksiklik oluşturur. Literatürde ceza puanı (Auer & Tarasowa, 2013; Ma, 2004; Siddiqui, Bhavsar, V., Bhavsar, A., & Bose, 2016) olarak da isimlendirilen düzeltmeler için çoktan seçmeli testlerde farklı metotlar önerilse de hiçbirisi şans başarısının etkisini ortadan kaldıramamıştır. Klasik tek doğru cevaplı çoktan

seçmeli maddelerden farklı bir yapıda hazırlanmış maddelerle oluşturulan bir testte farklı formüllerle yapılan düzeltme işlemlerinin şans başarısına etkisinin araştırılması faydalı sonuçlar ortaya çıkarabilir. Araştırma kapsamında çoktan seçmeli maddelerin bir türü olan çoklu cevaplı madde türü (multiple response item type) incelenecek ve puanlama yöntemleri üzerine karşılaştırmalar yapılacaktır. Çoklu cevaplı maddeler, maddenin doğru cevabının birden fazla olabileceği ve buna göre birden fazla işaretleme yapılabilmesine imkan sağlayan maddelerdir. Kısmi puanlamaya imkan sağlayan bu madde türünde tam puan almak için bütün doğru seçeneklerin işaretlenip bütün yanlış seçeneklerin boş bırakılması gerekir.

Test uygulamasında yaygın olarak iki tür çoklu cevaplı madde kullanılmaktadır. Birincisi kişilere maddelerdeki doğru cevap sayısının önceden verildiği tür iken; diğer daha açık uçlu tipte ise kişilere doğru cevapların tam sayısından bahsedilmeden kişiden ona göre doğru olan tüm seçenekleri işaretlemelerinin istendiği türdür. Çoklu cevaplı madde formatının kullanımı, bu tür bir maddeyi çoktan seçmelilerin çoğul cevaplı sorusu olarak tanımlayan ve tek yanıtıtan oluşan çoktan seçmeli madde türünden ayırt etmek için Orleans ve Sealy (1928) tarafından yapılan test geliştirmenin etkili bir şekli olarak önerilmiştir. Orleans ve Sealy, çoklu cevaplı maddelerin, bilginin geri çağırılması gibi ezberci bilgidен, ortak uyaranlara bağlı muhakeme gibi daha karmaşık bilişsel süreçlere kadar kullanılabileceğini göstermiştir. Çoklu cevaplı maddeler, bireylerin çok yönlülüğünün farklı seviyelerinde cevap vermelerine izin vererek ve bilişsel sürecin farklı aşamalarının sistematik olarak daha anlamlı bir şekilde gözlemlenmesini mümkün kılarak, kişilerin gerçek bilgi düzeylerini daha geçerli bir şekilde ölçmelerini sağlar (Ma, 2004, s. 20).

Resnick (1991), üst düzey düşünmenin, genellikle çoklu çözümlere götürdüğünü, düşünme sürecini kontrol etmeyi gerektirdiğini ve üst düzeyde çaba gerektirdiğini belirtmiştir (aktaran Doğanay, 2007, s. 306). Üst düzey düşünme becerileri içinde yer alan esnek düşünme ile de durumlara ya da sorunlara tek yönden değil, çok yönlü bakış açısı getirilebilir. Esnek düşünebilen bireyler çok sayıda alternatif fikir üretebilir, açık zihinliği sürdürebilir, farklı seçeneklere sahip olduğunun farkında olup alternatifleri görebilir (Duman, 2018). Bu sebeple değişimin hızlı olduğu bilgi çağında, gelişmeleri takip edip bilgiyi bilerek seçen, esnek düşünebilen ve alternatif çözümler üreten bireylere sahip olmak için esnek düşünebilme yetisini geliştirecek uygulamalara yönelim sağlanmalıdır. Çoklu cevaplı madde türünün bu amacı gerçekleştirmek için kullanılabilecek yapıda olduğu alanyazında (Hsu,

Moss & Khampalikit, 1984; Ma, 2004; Verbic, 2012) yer alan birçok araştırmacı tarafından ifade edilmektedir.

Çoklu cevaplı madde formatının testte kullanılmasını başlatan itici prensip, farklı bilgi düzeylerine sahip kişiler arasında ayırım yapmadaki etkinliği olduğundan, çoklu cevaplı maddeler için uygun puanlama kuralı çok net bir yapıya sahip olmalı (yani, farklı kısmi bilgiye sahip kişiler arasında ayırım yapmalı) ve kesin olmalıdır. Çoklu cevaplı maddelerin çoklu puanlanması için altta yatan varsayım, çeldiriciler ile doğru cevaplar arasında sistematik bir ilişkinin bulunmasıdır (Levine & Drasgow, 1983) ve puanlama algoritmaları, ölçüm hassasiyetini artırmak için farklı yanıt alternatiflerinde temsil edilen bu farklılaşan bilgileri kullanmalıdır. Her ne kadar çoktan seçmeli madde formatı ile ilgili olsa da, bu varsayım çoklu cevaplı madde puanlaması için de geçerlidir. Bu iki format birçok ortak özelliği paylaşmaktadır (Ma, 2004, s. 10).

Şans başarısını önlemek için, uygulanacak teste ait bir yönerge oluşturularak kişilere sadece doğru cevabını bildikleri maddeleri işaretlemelerini, cevabını bilmedikleri maddeleri ise boş bırakmaları gerektiği uyarısı test öncesinde yapılır. Ayrıca yönergeye yukarıda açıklandığı gibi yapılan yanlışlardan dolayı düzeltme işlemi uygulanacağını uyarısı eklenebilir. Bu sayede şans başarısı azaltılmaya çalışılır. Baykul (2015, s. 367), böyle bir tedbirin cevaplayıcıların yönergeye uymaları ölçüsünde işleyeceğini ifade etmiştir. Çelen (2002) çoktan seçmeli testlerle ilgili çalışmasında düzeltme yönergesi verilse de verilme de, test puanlarında şans başarısı için yapılan düzeltmenin, testi alanların sıralamasını neredeyse hiç değiştirmedeği sonucuna ulaşmıştır. Koçak (2013) araştırmasında, öğrencilerin yanıtlama davranışlarının farklı yönergelerle uygulanan testlerde farklı düzeylerde gösterilmesinin, yönergede yer alan uyarıların öğrencilerin gösterdikleri yanıtlama davranışlarında belirleyici olduğunu belirtmiştir. Çoban (2015) yaptığı çalışmada, farklı yönergelerle verilen testlerin psikometrik özelliklerini karşılaştırmış ve farklı yönergelerle uygulanan testlerin ham puan ortalamaları, düzeltilmiş puan ortalamaları ve düzeltilmiş puanlara göre hesaplanan güvenilirlik katsayıları ile madde güçlük indeksleri arasında fark olduğunu tespit etmiştir. Klasik tek doğrulu çoktan seçmeli madde türünden farklı olarak doğru cevap sayısı birden fazla olabilen maddeler kullanılarak hazırlanmış bir testte yönerge kullanılması veya kullanılmamasının test sonuçlarına etkisinin incelenmesi bu konuya katkı sağlayabilir.

Çoklu cevaplı maddeler, çoktan seçmeli maddelere kıyasla daha fazla cevap alternatifi sağlar ve bu sayede şans başarısını azaltır. Bunun yanında test yönergesi testi alanlara tam ve doğru olarak aktarıldığında, kişilerin tahmin etme eğilimleri azalır.

Kullanışlılık, geçerlik ve güvenilirlik gibi psikometrik özelliklerin yanında, ölçme aracında bulunması gereken bir diğer özelliktir. Kullanışlı bir ölçme aracı zaman, para ve emek açısından ekonomik olmalıdır. Ölçme aracı, hem uygulayıcı hem de uygulanan kişi açısından uygulama kolaylığına sahip olmalıdır. Yani ölçme aracını uygulayan kişi uygulamasını gereksiz çaba sarf etmeden yapabilmelidir ve uygulama yapılan kişiler de ölçme aracını açık, anlaşılır ve kolay bir şekilde algılayabilmelidir. Ayrıca ölçme aracı kolay, doğru ve hızlı puanlanabilmelidir (Kan, 2009, s. 70). Çoklu cevaplı maddeler çoktan seçmeli maddelerin kullanışlı olan bütün özelliklerine sahip olabilir ve bu sebeple eğitimde ölçme ve değerlendirmeye önemli katkılar sağlayabilir.

Bir testin; puanlanma tekniği (ikili veya çoklu), puanlayıcısı (makinelere veya insanlar), puanlanma süresi, puanlama güvenilirliği (objektif veya subjektif) gibi durumlar o testin geçerliğini, güvenilirliğini ve kullanışlılığını etkiler. Çoktan seçmeli maddelerden oluşan testler kullanıcılarına puanlama bakımından hem nesnellik hem de kolaylık sağlamaktadır. Bu testler makine yardımıyla da kolaylıkla ve çok kısa sürede puanlanabilir. Çoktan seçmeli testlerin en basit ve en çok kullanılan puanlama tekniği 0-1 puanlamasıdır (N. Doğan, 2009, s. 257). Bu puanlama tekniğinde doğru cevap bir puan diğer bütün alternatifler (yanlış seçenek işaretleme, boş bırakma vb.) sıfır puan değerindedir. Bu puanlama tekniğinin olumsuz tarafı, kişilere kısmi puan tayin edememesidir. Bu sebeple kişinin sadece tam doğru cevabı puan kazandırırken, onun konu hakkındaki kısmi bilgisini ortaya koyabilecek ve bu bilgiyi ödüllendirecek bir puanlama sisteminden söz edilememektedir. Waters (1977) cevaplayıcının kısmi bilgisinin değerlendirilmesi ile testin geçerlik ve güvenilirliğinin artabileceğini belirtmiştir. Anastasia (1982), 0-1 puanlama tekniğiyle elde edilen puanların, kişilerin bulundukları grup içindeki sıralarıyla ilgili bilgi verdiğini ama bireyler arasındaki farklılıkları ortaya çıkarmada yetersiz kaldığını belirtmiştir (aktaran N. Doğan, 2009, s. 258). Bu durum çoktan seçmeli maddelerden oluşan testler için önemli bir sınırlılık olarak kabul edilmiştir. Bu sınırlılığı ortadan kaldırmak için farklı puanlama yöntemleri önerilmiştir. Ancak bu yöntemlerin de ortaya çıkardığı dezavantajlardan dolayı kullanımı yaygınlaşmamıştır. Hem objektif, seri ve ekonomik puanlama yapıp hem de kısmi puan bilgisi sunan çoklu cevaplı maddeler birçok farklı puanlama yöntemine sahiptir (Domnich

vd. 2015; Verbic, 2012) . Puanlama yöntemlerinin farklılaşma sebebi, kısmi puanın nasıl dağıtılacağı, düzeltme yapılıp yapılmayacağı, eğer düzeltme yapılacaksa nasıl yapılacağı ve sıfır puanın hangi şartlarda verileceği gibi çeşitli olası senaryoların olmasıdır. Bazı puanlama yöntemleri düzeltme yapmazken bazıları fazla düzeltme yapabilmekte, bazılarında tek yanlış işaretleme sıfır puan verilmesine sebep olurken bazılarında sebep olmamaktadır. Örneğin Kanada gibi bazı ülkelerde yapılan tıp yeterlik sınavlarında çoklu cevaplı maddeler kullanılmaktadır ve bir yanlışın önemli sorunlara yol açabileceği durumlar göz önünde bulundurularak çok uygunsuz bir seçenek işaretlendiği için o madde için kişiye sıfır puan verilmektedir. Bu sebeple testin kullanım amacına göre farklı puanlama yöntemlerinin kullanılabileceği söylenebilir. Ancak çoklu cevaplı maddelerden oluşturulan bir başarı testi oluşturduğumuzda puanlama için hangi yöntemin diğerlerine oranla daha güvenilir, geçerli ve kullanışlı olduğunun belirlenmesi önem arz etmektedir.

Bu amaçla araştırmada farklı puanlama yöntemleri kullanılarak çoklu cevaplı maddelerle hazırlanan bir testin test ve madde istatistiklerinin ne derece değiştiğini belirlemektir. Dolayısıyla bu değişimin testin geçerlik ve güvenilirliğine yönelik etkilerini ortaya çıkarmaktır.

1.2. Kuramsal Çerçeve

Bu başlık altında çoktan seçmeli madde türü ve çoklu cevaplı madde türü hakkında yapılan incelemeler ele alınmıştır.

1.2.1. Çoktan Seçmeli Madde Türü

Çoktan seçmeli maddeler, sınıf içi başarı ölçümlerinde ve geniş ölçekli çalışmaların genelinde ölçüm aracı olarak oldukça yaygın bir şekilde kullanılmaktadır. Bu popülerliğin arkasındaki ana neden, çoktan seçmeli maddelerin çok fazla avantajının olmasıdır. Çoktan seçmeli maddeler teşhis ve biçimlendirici amaçlar için kullanılabilirler ve geniş bir bilgi yelpazesini değerlendirebilirler. Ayrıca, insanlar veya puanlama makineleri ile kolayca, hızlı, objektif ve ekonomik olarak puanlanabilirler. Bu ve buna benzer birçok avantaj, çoktan seçmeli testleri sınıf içi başarı testinden geniş ölçekli standart testlere kadar kapsamlı bir amaç için uygun hale getirmektedir ve alanda dominant test tekniği haline gelmesini sağlamıştır (Auer & Tarasowa, 2013, s. 1; Baghaei & Amrahi, 2011, s.192; Ben-Simon,

Budescu & Nevo, 1997, s. 65; Frisbie & Sweeney, 1982, s. 29; Jodoin, 2003, s. 1; Ma, 2004, s. 1; Wan & Henly, 2012, s. 59).

Çoktan seçmeli maddelerin ilk örneklerinin kullanıldığı testler 20. yüzyılın başlarında Edward Thorndike tarafından geliştirilmiştir. Frederick J. Kelly, 1914'te Kansas Devlet Okulu'nda (günümüzde Emporia Eyalet Üniversitesi) geniş çaplı bir değerlendirme için çoktan seçmeli soruları ilk kullanan kişi olmuştur. Testler, Birinci Dünya Savaşı'nda askere alımda askerlerin zekalarını değerlendirmek için kullanılmıştır. Çoktan seçmeli maddeler değerlendirme için iyi bir nesnel yönteme hizmet eder. Günümüzde çoktan seçmeli maddeler, öğrencilerin bilgisini ölçmek, düşünme ve okuduğunu anlama becerilerini ölçtüğünü iddia eden çeşitli rekabetçi sınavlardaki performanslarını değerlendirmek için yaygın olarak kullanılmaktadır. (Siddiqui vd., 2016).

Çoktan seçmeli madde formatının yapısı aşağıdaki gibidir (Haladyna & Rodriguez, 2013, s.62):

Kök: Bir soru veya kısmi bir cümle şeklinde olabilir.

Seçenekler: Sorunun cevabı ya da cümlenin tamamlayıcısı şeklinde olabilir.

Doğru Seçenek: Bazen doğru cevap veya anahtar cevap şeklinde isimlendirilebilir.

Yanlış Seçenek: Ayrıca çeldirici, dikkat dağıtıcı veya yanlış yönlendirici olarak da adlandırılır.

Doğru seçenek, sadece tek bir cevaptır. Soru biçimine göre doğru cevap, bir kelime, cümle parçası veya cümle olabilir. Çeldiriciler, test maddesinin yazılması en zor kısmıdır. Çeldiriciler tartışmasız yanlış cevaplardır ve her bir çeldirici, test maddesinin ölçmesi gereken bilgi ya da beceriyi henüz öğrenmemiş olan sınav katılımcılarına uygun olmalıdır. Maddede istenen bilgilere sahip olanlar için, çeldiricilerin açıkça yanlış seçimler olduğu söylenebilir. Çeldiriciler dil bilgisi biçimiyle, cümle tarzıyla ve uzunluğuyla doğru cevaba benzemelidir. Doğru cevaba götüren ince veya keskin ipuçlarından daima kaçınılmalıdır. İyi çeldiriciler, öğrencilerin ortak hatalarına dayanmalıdır. Çeldiriciler asla aldatıcı veya hileli bir şekilde doğru olmamalıdır.

Çoktan seçmeli madde formatları için iki ana içerik geçerlidir. Bunlardan ilki, bir çoktan seçmeli testin amacının, öğrencinin öğrenmesinin etkin bir şekilde ölçülmesini sağlamak olduğu sınıf içi test etmedir. Bu ölçme, öğretmenin değerlendirme sürecinin sonunda bir not vereceği durumlarda fayda sağlar. Ayrıca öğretmenlere ve öğrencilere gelecekteki

öğrenmelerde geri bildirim ve yardım sağlamaları ya da öğrenilmemiş içeriği yeniden öğretme ve yeniden öğrenme konusunda değerlidir. İkinci bağlam, geniş ölçekli bir test programıdır. Bu geniş ölçekli test programının amacı mezuniyet, terfi, sertifikalandırma, ruhsatlandırma, değerlendirme, yerleştirme veya kabul olabilir. Bu ikinci bağlamda da, çoktan seçmeli testler seçilmiştir çünkü etkili bir araç olmakla birlikte öğrencinin bilgi ve bilişsel becerileri öğrenmesi için faydalı bir özet sunar (Haladyna, 2004, s. 67).

Geleneksel çoktan seçmeli testler açıkça seçilen yanıt kategorisine girer ve öğrencilerin başarı düzeyini ölçmek için yarım yüzyıldan fazla bir süredir geniş ölçekte kullanılmaktadır. Bu testlerle genellikle bilgi veya bilişsel beceri ölçülür. Puanlama nesneldir. Test puanından yapılan çıkarım genellikle bir bilgi, beceri alanı ya da her ikisi içindir. Bununla birlikte, bu yaklaşım yapay ve geçersiz olarak eleştirilmiştir çünkü testler hataların yapılmasının asıl nedenini tespit edememektedir (Cheng, 2004; Haladyna, 2004, s. 134).

N. Doğan'a (2009, s. 235) göre çoktan seçmeli madde türü ile ilgili bazı özellikler şöyle sıralanabilir:

- Çoktan seçmeli testlerde yazma süresi çok az olduğundan çok sayıda madde testte kullanılabilir. Bu sebeple testin kapsam geçerliği ve güvenilirliği artar. Öğrencilerin doğru cevap için yalnızca işaretleme yapmalarının yeterli olması onlara zaman tasarrufu bakımından kullanışlılık sağlar.
- Çoktan seçmeli testlere güzel yazı, kompozisyon yazma yeteneği gibi değişkenler karışmamaktadır. Bu nedenle ölçülmek istenen değişkenler başka değişkenler ile karışmadan ölçülebilmektedir. Bu durum testin geçerliğini yükseltir. Ancak bu testlerde okuma hızı ve okuduğunu anlama gibi değişkenlerin sonuçlara karıştığı görülmektedir. Bu durum ise sınavın geçerliğini olumsuz etkiler.
- Çoktan seçmeli testler objektif puanlamayı sağlar. Bu nedenle puanlama güvenilirliği oldukça yüksektir. Çok kişiye aynı anda uygulanabilmesi, okuma kolaylığı da bu testin kullanışlılığının yüksek olmasına neden olur.
- Çoktan seçmeli testler istatistiksel temellere dayanmaktadır. Özellikle verilen cevapların kesinlik taşıması, maddelerin güçlük düzeylerinin istatistiksel olarak belirlenmesini sağlamaktadır. Bu nedenle test maddeleri ile ilgili testin ortalaması, değişkenliği, güvenilirliği, geçerliği hesaplanabilir.

Çoktan seçmeli madde türünün yukarıdaki verilen özellikleri hem baskın tür olmasını sağlayan hem de eleştirilmesine sebep olan bazı yönlerini göstermiştir. Tan (2010, s. 286)'a göre çoktan seçmeli maddeler bilgi, kavrama ve uygulama düzeyindeki davranışların ölçümünde kullanılabilir sınav türü olmakla birlikte, cevaplayıcının yaratıcılığını ölçmede kullanımı pek uygun değildir. Umay (1997) ise çoktan seçmeli testlerin, yanıtlayıcıların düşünme sürecini ve yanıtlama davranışını izleme olanağını vermemesini ise en büyük eksiklik olarak görür ancak aynı eksikliğin özelliği nedeniyle puanlamada objektiflik sağlanabildiğini belirtir. Bu da çoktan seçmeli testlerin ikilemidir.

Çoktan seçmeli testlerde test kurnazlığı yapılabilmektedir. Siddiqui ve diğerlerine (2016) göre, çoktan seçmeli testlerin çözümünde, kişi seçenekleri gözden geçirerek doğru seçeneği arıyor olabilir. Önce yanlış seçenekleri eleyebilir ve sonra doğru olanı seçmeye karar vermek için daha az sayıda seçeneğe yoğunlaşmış olabilir. İlk bakışta doğru cevabı bulması durumunda, kalan seçenekleri tam olarak okumayabilir. Çoktan seçmeli testler alternatif cevaplama ve puanlama yöntemlerine göre sınıflandırılmıştır. Lord ve Novick (1968, s. 302), çoktan seçmeli testlerin puanlanması için üç alternatif tür belirlemişlerdir. Bunlar yeni bir madde yapısı ya da cevaplama yöntemi, farklılaşan madde puanlama yöntemleri (seçenek ağırlıklandırma) ve farklılaşan test puanlama yöntemleridir (madde ağırlıklandırmadır).

Ayrıca çoktan seçmeli testlerin kullanılan maddelerin yapı bakımından birbirinden farklı bir çok çeşidi vardır. Bu çeşitlilik sebebiyle çoktan seçmeli testler, doğru cevaba göre, madde köküne göre ve maddelerin gruplanışına göre olmak üzere üç sınıfa ayrılmıştır. Doğru cevaba göre sınıflandırmada doğru cevabın niteliği ve sayısına göre maddeler beşe ayrılır (N. Doğan, 2009, s. 227; Turgut & Baykul, 2013, s. 188). Çoklu cevaplı maddeler bu sınıflamada yer alan beş türden birisidir.

1.2.2. Çoklu Cevaplı Madde Türü

Çoklu cevaplı madde türü, şüphesiz, bugün mevcut olan en kullanışlı alternatif madde türlerinden biridir ve yeni bir madde türü değildir. Çoklu cevaplı maddeler uzun yıllardır çoklu seçim, çoklu listeler, çoktan seçmeli vb. dahil olmak üzere çeşitli isimler altında kağıt tabanlı değerlendirmelerde kullanılmıştır (Parshall, 2015). Çoklu cevaplı test maddeleri bir madde kökü, bir yönerge ve sayısı testi oluşturan kişiye bağlı olmak üzere doğru ve yanlış seçeneklerden oluşur. Bireyler doğru seçenekleri seçtikleri ve yanlış seçenekleri işaretler

bıraktıkları takdirde çoklu cevaplı test maddesine tam olarak doğru yanıt vermiş olurlar (Verbic, 2012).

Çoklu cevaplı test maddeleri, birden fazla alternatifin doğru olabileceği durumlarda bazı önemli şekillerde esneklik kazanır. Alışılmış bir cevabı ve sadece bir cevabı olan (bir anahtarlı alternatif) test maddesinin genelleştirilmesi, çoklu cevaplı maddeler olarak adlandırılır. Çoklu cevaplı maddelerin kullanımı, Orleans ve Sealy (1928) tarafından etkili bir test yapım şekli olarak önerilmektedir. Çoklu maddelerin formatı çoğu zaman soruların ve alternatiflerin daha rahat ve doğal ifadesine izin verir. Ayrıca, test yapımını basitleştirir ve bazı durumlarda daha az makul çeldiriciye ihtiyaç duyulmasına olanak sağlar (Duncan & Milton, 1978).

Çoklu cevaplı maddeler, her bir soru için birden fazla doğru cevabı olan ve her doğru cevabı seçmek için seçenekleri incelemeyi gerektiren alışılmış bir doğru cevabı olan ve bir yanıt seçmeli (tek doğru cevabı olan) çoktan seçmeli maddelerin bir uzantısıdır. Bu tip bir madde bir gövde ve çeşitli cevap alternatiflerinden oluşur (tipik olarak maksimum sayı altıdır, ancak bazı durumlarda yirmiye veya daha fazlasına kadar çıkabilir), bunların herhangi bir kısmı doğru cevaplar olabilir. (Multiple Multiple-Choice) Çoklu Çoktan Seçmeli (Cronbach, 1941), (Multiple Answer) Çoklu Cevap (Dressel & Schmid, 1953), (Multiple-Mark Items) Çoklu İşaretlemeli (Pomplun & Omar, 1997; Scalise & Gifford, 2006), (Selected Response Items) Seçilen Yanıt Maddeleri (Sireci & Zenisky, 2006, s. 331) dahil olmak üzere bu tür maddelere çeşitli isimler verilmiştir (Pomplun & Omar, 1997; Verbic, 2012).

Bir maddede birden fazla doğru seçenek olması mümkün olduğunda, kişinin yaklaşımı farklı olacaktır. Burada, kişi seçenekleri elemek isteyebilir, ancak her seçeneğin doğruluğu olasılığını ayrı ayrı arayacaktır. Bir seçeneğin doğruluğu, kişiye bir sonraki seçeneğe dikkat etmemesini gerektirecek rahatlığı sağlamayacaktır. Ayrıca bu durum, tek bir doğru cevabı olan çoktan seçmeli maddelerin eksikliklerinin çoklu cevaplı madde tipi ile kapatılabileceğini dikkate almak için önemli bir göstergedir. Çoklu cevaplı madde tipinde, bir konunun birden fazla kavramı / uygulaması / yönü test edilebilir ve çeşitli hiyerarşi seviyesindeki bilişsel beceriler test edilebilir. Bu tür maddeleri cevaplarken öğrencilerin stratejileri bile farklıdır. Bu sebeple çoklu cevaplı maddeler ne çerçeve bakımından kullanışsızdır ne de herhangi bir karmaşık kod ile ilişkilidir (Siddiqui vd. 2016). Çoklu cevaplı maddelerdeki daha geniş çözüm alanı kişilerin tahmin etme davranışından uzaklaşmalarını sağlar ve öğrencilerin öğrendiklerinin ayrımını daha kolay yapmasına katkı

sunar. Ayrıca çoklu cevaplı maddelerden oluşan testler geniş bir konu alanı hakkında etkili geri bildirim yapılmasını sağlayacak şekilde tasarlanabilir (Peterson, Craig & Danny, 2016). Çoklu cevaplı maddelerdeki çeldiriciler sadece doğru cevabı belirlemeyi zorlaştırmakla kalmaz, aynı zamanda doğru cevabın neden uygun olduğunu ve doğru cevap ile karşılaştırıldığında çeldiricideki yanlışlıkları gösterir (Matsuda, Ogawa, Hirashima & Taki, 2015).

Çoklu cevaplı madde türü çok geniş bir kullanım alanına sahiptir. Aslında, çoğu sınavın, birden fazla anahtar içeren maddelerin alakalı olduğu konu alanlarını içerdiğini söylemek muhtemelen doğrudur. Bunlar, doğal olarak kümelenen unsurları olan konu alanlarıdır. Örneğin tıbbi değerlendirmelerde, çoklu cevaplı madde türü, teşhisle ilgili olarak kullanılabilir (Parshall, 2015). Tıbbi testlerde Anahtar Özellik (Key Feature) problemleri olarak geçen madde türü çoklu cevaplı madde formatıdır. Page, Bordage ve Allen (1995), Kanada'daki Kanadalı ve Yabancı Tıp Fakültelerinin mezunları tarafından, alınan bir lisans sınavı olan Kanada'daki Tıp Yeterlik Sınavında kullanılmak üzere Anahtar Özellik problemlerinin gelişimini tanımlamışlardır. Anahtar Özellik bir klinik problemin çözümünde kritik bir adım olarak tanımlanmaktadır ve bir ana problem, klinik durum senaryosundan oluşmakta ve sadece bu kritik aşamalara odaklanan sorular gelmektedir. Anahtar Özellik problemine uygun yanıtlar bir veya daha fazla olabilir. Anahtar-Özellik problemleri için kullanılan formatlardan biri, cevap seçeneklerinin bir listesinin öğrencilere sunulduğu çoklu cevaplı madde formatıdır. Page vd. Anahtar Özellik problemlerinin psikometrik özelliklerini değerlendirmiş ve klinik problem çözme becerilerinin geçerli ve güvenilir ölçütleri olduğu ve tıbbi test uzmanları tarafından dikkate değer olduğu sonucuna varmışlardır. Ayrıca Amerikan Hekimler Koleji ve diğer tıp okulları, daha sonra test programlarında kullanılmak üzere Anahtar Özellik problemlerinin benimsendiğini belirtmişlerdir.

Çoklu cevaplı maddeler eğitim alanında farklı ülkelerde kullanılmıştır. ABD'de geniş ölçekli devlet değerlendirme programlarında çoklu cevaplı maddelerin uygulamaları yapılmıştır. Örneğin, Kansas Eyaleti Değerleme Programı (Kansas State Assessment Programme) 3, 7 ve 10. sınıflarda okuma testlerinde ve 4, 7 ve 10. sınıflarda matematik testlerinde çoklu cevaplı maddeleri kullanmıştır. Çalışmalarında çoklu cevaplı test maddeleri için tehdit olabilecek dört özellik olarak; öğrencilerin yönergeyi anlama durumları, öğrencilerin maddeyi boş bırakmaları, seçenekler arası bağımlılık ve öğrencilerin tahmin etme davranışı

sıralanmıştır. Çalışmada bu muhtemel tehditlerden biri olan yönergenin anlaşılma ihtimali sorunu öğrencilerin %95'inden fazlasının doğru anlamasıyla ortadan kalkmıştır. Öğrencilerin %1'inden azı boş madde bırakmıştır ve öğrencilerin maddeyi tahmin edip işaretlemektense boş bırakmayı tercih ettikleri görülmüştür. Güvenirlik sonuçlarına göre ise bazı seçenekler arasında bağımlılık olduğu görülmüştür. Çoklu cevaplı maddelerin psikometrik özellikleriyle ilgili araştırma sonucunda ise çoklu cevaplı madde formatının kullanımını desteklemek için yeterli güvenirlik ve geçerlik kanıtı olduğu sonucuna varılmıştır ve istenen özellikleri nedeniyle, devlet değerlendirme programlarında kullanım için umut vaat eden bir madde formatı olduğu belirtilmiştir (Pomplun & Omar, 1997). Ayrıca Hsu, Moss ve Khampalikit (1984), ABD dışında bazı yabancı ülkelerde de çoklu cevaplı maddelerin kullanıldığını belirtmişlerdir ve Tayvan Kolejlere Giriş Sınavı'nı örnek vermişlerdir. Dünya çapında 70 farklı ülkede faaliyet gösteren bir eğitim şirketi olan Pearson, içerisinde çoklu cevaplı maddelerin de bulunduğu PTE (Pearson Test of English) adında uluslararası geçerliliği olan bilgisayar tabanlı bir dil sınavı uygulamaktadır. Ayrıca Microsoft, çeşitli sertifika programları için uyguladığı bilgisayar tabanlı sınavlarda çoklu cevaplı madde türünü kullanmaktadır. Türkiye'de çoklu cevaplı madde türü pek uygulanmayan bir madde türüdür. Türkçe literatürde çoklu cevaplı maddelerle ilgili bazı yazarların kısa açıklamaları mevcuttur. Turgut (1988), anahtarlanmış cevabı birden fazla olan maddeler şeklinde isimlendirip madde türü hakkında kısa bir bilgi vermiştir. Birden fazla doğru cevabın şansa doğru cevabın bulunmasından doğan hatayı önemli derecede azaltacağını ve bazı yoklanmak istenen kazanımların tespitinde etkili olabileceğini belirten Turgut (1988) , bu madde türünü puanlamanın güçlüğü ve madde analizinin zorluğunu da dile getirmiştir. Turgut ve Baykul (2013, s. 191), anahtarlanmış cevapları birden fazla olan maddeler, Özdemir (2003), çoklu doğru cevap yöntemi, N. Doğan (2009, s.230) ve Çetin (2016, s. 135), birden fazla doğru yanıtı olan maddeler gibi isimler altında çoklu cevaplı madde türünün kısa tanımlarını yapmışlardır.

Dressel ve Schmid (1953), bir tanesi çoklu cevaplı madde formatı olan çoktan seçmeli maddelerin çeşitli tiplerini ampirik olarak araştırmıştır. Çoklu cevaplı maddelerin çoktan seçmeli madde formatından ve diğer çoktan seçmeli madde formatlarından (özgür seçim ve kesinlik dereceleri) biraz daha yüksek bir güvenirliliğe sahip oldukları sonucuna varmışlardır. Dahası, Dressel ve Schmid, çoklu cevaplı maddelerin, çeşitli kısmi bilgi seviyelerini ölçme ve kişiler arasında daha iyi bir ayırtıcılık sağlama potansiyeline sahip olduğunu öne sürmüştür.

Cronbach (1941), Dressel ve Schmid'in (1953) çoklu cevaplı maddelere ilişkin ufuk açıcı çalışmasından bu yana, bu tür bir madde üzerinde yapılan araştırmalar ve testteki kullanımı, muhtemelen çoktan seçmeli madde formatının baskınlığı, iyi tasarlanmış ve yazılmış çoklu cevaplı maddelerin hazırlanmasında ve puanlanmasında karşılaşılan güçlükler nedeniyle az olmuştur. Çoklu cevaplı maddeler, çoktan seçmeli maddelerin yararlı yönlerinin birçoğunu koruyabilir ve aynı zamanda, performans değerlendirme görevlerinde daha düşük maliyetle daha yüksek mertebe düşünme becerilerini ölçebilir (Ma, 2004).

Geleneksel kağıt-kalem testlerinden bilgisayar tabanlı testlere son zamanlarda yapılan geçişler, bilgisayar teknolojisi ile sağlanan grafik ve zamanlama özelliklerinden yararlanmak için bilgisayarlı test yöntemine uyarlanabilen yenilikçi ürün türlerine olan ilgiyi de ortaya çıkardı. Yenilikçi madde türlerinden biri çoklu cevaplı madde formatıdır (Parshall, Davey & Pashley, 2000, s. 131). Bilgisayarlı testlerde iki tip çoklu cevaplı maddelerin kullanımı ile ilgili raporlar vardır (Jodoin, 2003; O'Neil & Folk, 1996). Örneğin Parshall, Stewart ve Ritter (1996), bilgisayarlı testlerde grafik, ses ve alternatif cevap modları gibi yeniliklerin kullanımının uygunluğunu araştırmak için bir çalışma yürütmüşler ve çalışmalarının bir bölümünü çoklu cevaplı madde formatının değerlendirmesine ayırmışlardır. Burada çoklu cevaplı maddelerden, daha açık uçlu olan formatı (Doğru Olanların Tümünü Seç) kullanılmış ve ikili olarak puanlanmıştır (aktaran Ma, 2004).

Jodoin (2003), bilgisayar tabanlı olarak uygulanan Microsoft sertifikalı sistem mühendisliği sertifikasyon programı (MCSE) için kullanılan çoklu cevaplı maddelerin iki türü (doğru cevabı önceden verilen ve verilmeyen çoklu cevaplı maddeler) ile çoktan seçmeli testleri Madde Tepki Kuramı'na (MTK) göre karşılaştırmıştır. MCSE sertifikası, teknoloji endüstrisindeki birçok çalışan için bir iş gereksinimidir ve Microsoft Windows yazılımının uygulanmasına odaklanan bir dizi incelemeden oluşur. Bu uygulama ile bilgisayar tabanlı uygulamalarda çoklu cevaplı maddelerin kullanıldığını görülür. Araştırmada yenilikçi madde türü olarak belirtilen çoklu cevaplı maddelerin çoktan seçmeli maddelere oranla bütün yetenek seviyelerinde daha fazla bilgi sağladığını belirtmiştir.

Eggen ve Lampe'ye (2011) göre de çoklu cevaplı maddeler, eşleşen maddeler ve dizilim maddeleri yenilikçi madde türleri olarak adlandırılır, çünkü bu madde türleri genellikle bilgisayar tabanlı değerlendirme sistemlerinde bulunur. Bu madde türleri kağıt ve kalem testlerinde de kullanılabilmesine rağmen, bu madde tiplerinin puanlanması karmaşık bir hesaplama gerektirebilir. Bilgisayar tabanlı değerlendirme yazılımı, bu madde türlerinin

puanlarını otomatik olarak hesaplamayı sağlar. Bu madde türleri, çok amaçlı puanlama (çoklu) veya kısmi puanlama olanağı sunar ve bu nedenle geleneksel çoktan seçmeli maddelere göre daha fazla bilgi sağlayabilir.

Çoklu cevaplı madde türünün başka test türleriyle karşılaştırıldığı çalışmalar incelenmiştir. Qian, Woo ve Kim (2018), çoklu cevaplı maddelerden oluşan testler ile çoktan seçmeli maddelerden oluşan testleri karşılaştırdıklarında çoklu cevaplı testlerin madde gücünün daha yüksek olduğu, madde ayırt ediciliğinin daha yüksek olduğu, tahminle doğru cevabı bulma ihtimalinin daha düşük olduğu sonuçlarına ulaşmışlardır. Ayrıca çoklu cevaplı testlerin çoktan seçmeli testlere göre cevaplama süresinin daha uzun olduğunu belirtmişlerdir. Albanese (1993), farklı çoktan seçmeli madde türlerini karşılaştırdığı çalışmasında çoklu cevaplı madde türünün kompleks çoktan seçmeli madde türüne göre daha yüksek güvenilirlik gösterme eğiliminde olduğunu belirtmiştir. Albanese, Kent ve Whitney (1977), sağlık bilimleri öğrencileri için çoklu cevaplı ve K tipi test (birleşik yanıt vermeyi gerektiren madde türü) puanlarının lisans not ortalamaları (GPA) ile korelasyonunu incelemiştir ve çoklu cevaplı maddelerin GPA ile 0,41 ve K Tipi maddelerin ise 0,39 korelasyona sahip olduğunu ifade ederlerken, Frisbie ve Sweeney (1982) ise çalışmalarında çoklu cevaplı maddenin cevaplama süresinin A tipi (en iyi tek cevaplı) çoktan seçmeli maddeye göre yaklaşık % 15 daha fazla zaman gerektirdiğini belirtmişlerdir (aktaran Albanese, 1993). Çoklu cevaplı madde türünü başka madde türleriyle karşılaştıran çalışmalardan birisi de Hochstetter, Page, Chang ve Cappaert'in (2016) çalışmalarıdır. Bu çalışmada çoklu cevaplı madde türü ile evet/ hayır türü madde ve hotspot türü madde karşılaştırılmıştır ve değişen madde fonksiyonu açısından anlamlı bir fark bulunamamıştır. Ayrıca, çoklu cevaplı maddeler daha yüksek gücüne sahip olma eğiliminde ve evet/ hayır veya hotspot maddelerden biraz daha iyi ayırt edici olduğu bulunmuştur. Çoklu cevaplı maddeler ayrıca Madde Tepki Kuramı modelinde evet/ hayır türü maddelerden daha fazla madde bilgisi sağlamıştır. Haladyna (2004, s. 84), yapılan araştırmalarla, çoklu cevaplı maddeleri keşfetmeye devam ederken, bu formatın hem sınıf testlerinde hem de resmi, standartlaştırılmış test programlarında daha yaygın olarak kullanıldığını göreceğimizi belirtir. Ayrıca kısa sürede çok sayıda maddeyi gösterme ekonomikliği, bu formatın esas öne çıkan yönüdür.

Çoklu cevaplı madde formatı, birçok test kuruluşu tarafından ölçüm kalitesini iyileştirmek için yararlı olarak algılanmıştır. Çoklu cevaplı maddenin çoktan seçmeli maddeye olan yapı

benzerliđi ve çoklu cevaplı maddenin bu benzerlikten kaynaklı olarak sorularının yazılmasıyla ilgili göreceli kolaylık, çoklu cevaplı madde formatını test kuruluşları için uygun bir alternatif haline getirmektedir. Çoklu cevaplı maddeler, yapı kapsamını genişletmek, tahminin etkisini ve kişiler için problem olacađı öngörülen ve birçok madde yazma kılavuzunda yasaklanan tuzak anlamında “negatif gövdeler” olarak adlandırılan (örn, olumsuzla biten ve haricidir gibi sözcük seçimleri) yapıları azaltmak için umut vaat etmektedir. Çoklu cevaplı maddeler sıklıkla daha yüksek güçlük ve ayırıcılık göstermiş olup, istenirse bir sınav veya soru bankasının güçlük ve güvenirliğine ince ayar yapmasını sađlayan bir test programına sahiptir (Muckle, Becker & Wu 2011).

1.2.2.1. Çoklu Cevaplı Madde Türünün Puanlanması

Çoklu cevaplı maddeler hem çoklu hem de ikili puanlama yöntemine göre puanlanır. Bu madde türünde çoktan seçmeli madde türündeki gibi bir madde için öğrencinin alabileceđi puan sadece “0” yada “1” olmak zorunda değildir. Çünkü testi alan kişi işaretlediđi dođru ve yanlış seçeneklere göre farklı kısmi puanlar alabilir. Haladyna ve Rodriguez’e (2013, s. 121) göre kısmi puanlama uygulaması güvenirliği artırma eğilimindedir. Aynı zamanda kavram ilke ve yöntemler için öğrenmenin derecesi hakkında bilgi verir. Pomplun ve Omar (1997) da çalışmalarında bu madde türünün kullanımına dair olumlu söylemlerde bulunmuşlardır ancak madde türünün sınavı yapacakların alışkanlıkları dolayısıyla bazı problemlerin yaşanacağını belirtmişlerdir. Haladyna ve Rodriguez (2013, s. 121) bu problemin her yeni uygulamada olduđu gibi pratik ve aşinalıkla ortadan kaldırılabileceđini ve revize edilmiş Yüksek Lisans Kayıt Sınavı (GRE) genel testlerinde de artık çoklu cevaplı maddelerin kullanılmakta olduğunu belirtmişlerdir.

Literatürde anlatılan tüm iyi özelliklerine rağmen çoklu cevaplı maddelerin kullanımı diđer türler kadar yaygın değildir. Siddiqui ve diđerlerine (2016) göre bunun sebebi olarak puanlama sistemi önemli bir etkidir. Deđerlendirme için pratik uygulanabilir yöntem eksikliği, çoklu cevaplı maddelerin sıklıkla kullanılmamasından sorumlu faktördür. Çoklu cevaplı maddelerin puanlanmasına yönelik ortak bir görüş ve buna bađlı olarak ortak bir puanlama modeli yoktur. Bu sebeple birçok araştırmacı farklı puanlama modelleri önermişlerdir. Puanlama modellerinin farklılıklarına sebep olan genel durum, kısmi dođru cevapların nasıl puanlanacađı ve tahmin etme davranışını azaltmak için uygulanan düzeltme

puanının (ceza puanı) nasıl hesaplanacağıdır. Buna rağmen bütün puanlama modelleri testin kullanım amacına göre seçilip uygulanabilir.

Muckle vd.'nin (2011) araştırmalarında kullandıkları puanlama modelleri şunlardır:

1. Puanlama Modeli - İkili Puanlama (Dichotomous-İP): Tam puan almak için tüm doğru seçeneklerin işaretlenmesi ve hiçbir yanlış seçeneğin işaretlenmemesi gerekir. Diğer tüm durumlarda sıfır puan verilir. Bu puanlama yöntemi Frisbie ve Druva (1986) tarafından Küme Puanlama (Cluster Scoring) olarak isimlendirilmiştir.
2. Puanlama Modeli – Tuzaklı Çoklu Cevaplı (Polytomous Trapdoor-TÇC): İşaretlenen her doğru seçenek için 1 puan verilir, ancak herhangi bir yanlış işaretleme varsa sıfır puan verilir.
3. Puanlama Modeli – Tuzaklı İki Puanlı (Two-Point Trapdoor): İşaretlenen tüm seçenekler doğru ise 2 puan verilirken, bazı doğru seçenekler işaretlenmişse ve hiçbir yanlış seçenek işaretlenmemişse 1 puan verilir. Herhangi bir yanlış seçenek işaretlenmişse ya da hiçbir doğru seçenek işaretlenmemişse sıfır puan verilir.
4. Puanlama Modeli – Negatif Puanlama (Negative Scoring-NP): İşaretlenen her bir doğru seçenek için 1 puan verilirken, işaretlenen her bir yanlış seçenek için -1 puan verilir. Taban etkisi uygulanırsa minimum puan sıfırdır. Uygulanmazsa kişi negatif puanları da alabilir.
5. Puanlama Modeli – Çoklu Doğru/Yanlış (Multiple True/False-ÇDY): İşaretlenen her bir doğru seçenek için 1 puan verilirken işaretlenmeyen her yanlış seçenek için 1 puan verilir.
6. Puanlama Modeli – Pozitif Sayım (Positive Count-PS): İşaretlenen her bir doğru seçenek için 1 puan verilir.

Sayısal olarak en basit puanlama algoritması ikili puanlamadır. Bu puanlamada katılımcı tüm seçenekleri doğru işaretlenmişse tam puan alır, ancak diğer bütün seçimlerde hiç puan alamaz. Bununla birlikte, ikili puanlamanın önemli bir dezavantajı, kısmi bilgi için puan verememesidir; biri hariç tümünü doğru bilenle herhangi bir doğru cevabı bilmeyen ya da hepsini yanlış işaretleyen kişiler aynı puanı alırlar (Domnich vd. 2015). Bu puanlama yöntemi, genellikle testin kalitesinin iyi olduğu ve sonuçları puanlarken öğretmenin daha katı olmasını sağlayan kağıt kaleme dayalı testlerde kullanılır. Yanlış bir seçenek seçilmesi durumunda, öğrencinin sorunun ölçmek istediği kazanımı iyi bilmediğinden doğru cevabı tahmin etmeyi umduğunu gösterir. Testlerin amacı sadece sonuçları puanlamak değil, aynı zamanda bilgi boşluklarını yakalamak ve kısmen doğru cevapların puanlanmasıyla öğrencinin gerçek bilgisini daha iyi göstermesini sağlamaktır. Ayrıca ikili puanlama

öğrencideki gelişimi doğru bir şekilde göstermez (Auer & Tarasowa, 2013). Jiao, Liu, Haynie, Woo ve Gorham (2012), her bir çoklu puanlamanın sonucunun ikili puanlamanın sonucu ile karşılaştıran simülasyon çalışmasında, çoklu puanlamadan kaynaklanan yetenek tahminlerinin, ikili puanlamadan elde edilenlerden biraz daha yüksek ölçüm hassasiyetine sahip olduğunu bulmuşlardır.

Hsu, Moss ve Khampalikit'in (1984), çalışmaları, kısmi puan verildiğinde çoklu cevaplı maddelerin önemli ölçüde daha ayırt edici ve güvenilir olduğunu göstermektedir. Çalışmada çoklu cevaplı maddelerin puanlanması için kullanılan 6 formülden herhangi birinin diğerlerine göre belirgin bir üstünlük sergilemediği çıkarımı yapılmıştır. Ayrıca tahmin için düzeltme yapılmayan formüllerin, en az düzeltme yapılan formüller kadar etkili olduğu belirtilmiştir.

Page vd. (1995), özellikle tıpta kullanılan Anahtar Özellik isimli çoklu cevaplı madde türü için bir puanlama sistemi geliştirmişlerdir. Buna göre kişiye doğru işaretlenen seçenek sayısının maddedeki doğru cevap sayısına oranı kadar o madde için puan verilir. Ancak bu puanlamayı ayıran nokta çok uygunsuz bir cevabın seçilmesi durumunda kişinin o maddeden puan alamaması durumudur. Bu ciddi cezadaki prensip, bu cevapların ifade ettiği eylemlerin ciddi sonuçlar doğurabileceğidir (örneğin, yaşamı tehdit edebilecek seçimler). Dolayısıyla bu yanıtları işaretleyen bir kişiye, bu kişinin diğer yanıtlarının doğruluğuna bakılmaksızın, hiçbir puan verilmez.

Çoklu cevaplı maddeleri puanlamak için tarihsel olarak ilk yöntemlerden biri 1979'da Morgan tarafından tanımlanmıştır . Yönteme göre, puanlar aşağıdaki algoritma ile belirlenir (Morgan, 1979):

1. Doğru olarak seçilen her seçenek için öğrenci +1 puan alır.
2. Yanlış olarak seçilen her seçenek için öğrenci -1 puan alır.
3. Seçili olmayan her seçenek için, seçeneğin doğru mu ya da yanlış mı olduğuna bakılmaksızın, pozitif veya negatif hiçbir puan verilmez.

Test sırasında, pozitif ve negatif puanlar ayrı ayrı toplanır ve testin sonunda birikmiş pozitif ve negatif puanlar, testteki toplam doğru ve yanlış alternatif sayısının yüzdesi olarak ifade edilir; kişinin puanı, bu iki yüzde arasındaki farktır.

Morgan Algoritması (MA), 1 olan sabit sayıyı dinamik olarak belirlenen bir puanla değiştirerek Auer ve Tarasowa (2013) tarafından geliştirilmiştir. Buna göre:

1. Doğru seçilen her seçenek için, öğrencinin alacağı puan (p_{max} / n). (p_{max} : maddeden alınabilecek en yüksek puan, n : doğru seçeneklerin sayısı)
2. Yanlış seçilen her seçenek için öğrencinin kaybedeceği puan (p_{max} / k) (k : çeldirici sayısı)

Ayrıca Auer ve Tarasowa (2013), Morgan Algoritması'nda doğru ve yanlış seçenek sayıları ile ceza puanı arasında büyük bir bağımlılık olduğunu gösteren çalışmalar olduğunu ve bunun da yöntemdeki çarpıklığı gösterdiğini ifade etmişlerdir.

Çoklu Doğru/Yanlış maddeleri için geliştirilen yaklaşımlarla çoklu cevaplı maddeler puanlanabilir. Çoklu cevaplı madde türü çoklu doğru/yanlış madde türünden daha farklıdır. Bu fark kişinin çoklu cevaplı maddelerde sadece doğru olan seçenekleri işaretlemesinin istenmesidir ve bu durum sanıldığından daha önemli farklılığa sebep olur. Emmerich (1991), kişinin çoklu doğru/yanlış maddelerini cevaplarken yanıltan daha çok doğruyu işaretleme eğiliminde olduğunu ifade etmiştir. Çoklu doğru/yanlış maddelerini puanlamak için temel yaklaşım, öğrencinin doğru cevaba ne kadar yakın olduğunu tespit etmektir. Tsai ve Suen (1993), çoklu cevaplı madde türünün altı farklı puanlama uygulamasını değerlendirmiştir. Araştırmada, kısmi puanlamanın ikili puanlamadan üstünlüğü bulunsa da, algoritmaların birinin diğerine göre tercih edilebilir olarak değerlendirmesi yapılmamıştır. Bunun yanında araştırma sonucunda çoklu cevaplı olarak puanlanan (ikili puanlama yöntemine göre) çoklu doğru/yanlış maddelerinin en yüksek madde gücüne sahip olduğu görülmüştür.

Tüm çoklu doğru/yanlış puanlama yöntemi algoritmalarında, herhangi bir maddenin n seçeneğe sahip olduğu ve tüm doğru seçeneklerin işaretlenmesi sonucu tam puan verildiği belirtilir. Testi alan kişi doğru seçeneği işaretlememişse veya bir çeldiriciyi işaretlemişse, ($s = \max \text{ puan} / n$) kadar puan düzeltme yapılarak düşülür. Dolayısıyla kişi bir çeldiriciyi seçmediği için ve doğru bir seçenek seçtiği için puan alır. Bu noktada, iki madde tipi arasındaki farklılıklar nedeniyle çoklu cevaplı maddelere mükemmel uyum olmadığı belirlenmiştir (Cronbach, 1941; Pomplun & Omar, 1997).

Çoklu Doğru/Yanlış Puanlama Yöntemi'nin asıl sorunu, çoklu cevaplı maddelere uygulandığında, herhangi bir seçenek işaretlememiş olsa bile, kişinin puan almasıdır. Her ne kadar ilave bir kural oluşturularak problem çözülebilse de, deneyler puanlama uygulamasının çoklu cevaplı maddeler için kullanıldığında algoritmanın diğer problemlerini de göstermektedir (Auer & Tarasowa, 2013). Siddiqui vd. (2016) Çoklu Doğru/Yanlış Puanlama Yöntemi ile ilgili bu tür sorunları önlemek için, (a) hiçbir seçenek

işaretlenmemişse veya (b) tüm seçenekler işaretlenmişse, madde “girişimde bulunulmamış(çözülmemiş)” olarak değerlendirilebileceğini belirtir. Bu durumda kişiye hiçbir not verilmeyeceği ve böylece, Tarasowa ve Auer tarafından tanımlanan Çoklu Doğru/Yanlış Puanlama Yöntemi' nin dezavantajının ortadan kaldırılabilceğini ifade eder. Ancak, sınava giren kişiye önceden ilgili talimatlar verilmesi gerektiğini de belirtir.

Ripkey, Case ve Swanson (1996), yazarın adını alan karmaşık olmayan bir puanlama algoritması önermişlerdir. Ripkey Algoritması (RA) yaklaşımında, işaretlenen her doğru seçenek için toplam doğru seçenek sayısına bağlı olarak tam puanın bir kısmı verilir. Yaklaşım yanlış işaretlemeler için puan kesintisi yapmaz, ancak doğru seçenek sayısından daha fazla seçenek işaretlendiğinde bu maddeye sıfır puan verilir.

Ripkey'in araştırması, gerçek yaşam değerlendirmesinde umut verici sonuçlar vermiştir. Bununla birlikte, daha sonra yapılan araştırmalarda, Ripkey'in çalışmasının sınırları fark edilmiştir. Ripkey algoritmasındaki ana sorun, dengeli olmayan cezadır. Yapılan deneyler, çoğu durumda algoritmanın ciddi bir şekilde cezalandırdığını göstermektedir, öyle ki öğrenciler sanki ikili puanlamadaki gibi değerlendirilmiş olmaktadır (Auer & Tarasowa, 2013).

Auer ve Tarasowa (2013), mevcut yaklaşımların çoklu cevaplı maddelerin puanlama sorununu mükemmel şekilde çözmediğini iddia ederek Dengeli Puanlama (Balanced Scoring-DP) adında bir puanlama algoritması tasarlamışlardır. Konseptleri, puanlamanın maddenin tahmin edilme seviyesini temel alan bir varsayıma dayanmaktadır. Bu nedenle, bir öğrenci olası tüm seçenekleri işaretlediğinde, bu durum tahmin edilme seviyesini 1'e kadar yükseltir. Eğer sorunun tüm seçenekleri doğru iken öğrenci tüm seçenekleri işaretlemiş ise bu durumda, öğrencinin tam puan alması gerekir ama bu soruda en az bir çeldiricisi varsa ve yine öğrenci tüm seçenekleri işaretlemişse bu durumda öğrenci sıfır puan alır . Bununla birlikte, eğer bir öğrenci herhangi bir seçeneği işaretlememişse, tüm soruların en az bir doğru seçeneği olduğunu varsayıldığından, alacağı puan daima sıfır olmalıdır. Bu nedenle algoritmanın görevi, yanıtın doğruluk yüzdesini bulmak ve eğer tahmin etme seviyesi yapay olarak doğru seçenek sayısından fazla seçenek işaretlenerek arttırıldıysa da bu durumu bir ceza puanı ile azaltmaktır.

Auer ve Tarasowa'ya (2013) göre Dengeli Puanlama Algoritması'nın mantığında , soruların tahmin düzeyinin doğal seviyesine sahip olduğu vardır ve yalnızca öğrencinin yanıtından sonra tahmin düzeyi artarsa yani öğrenci doğru seçenek sayısından daha fazla seçenek

işaretlerse ceza belirlenir ve uygulanır. Yaklaşımda yalnızca değerlendirme maddeleri için aşağıdaki şartlara uyan sistemlerde kullanımı önerilir:

- Maddenin seçeneklerinin tümü aynı ağırlığa sahip olmalı;
- En az bir doğru seçenek olmalı;
- Diğerlerini dışarda tutan hiçbir seçenek olmamalıdır. (örneğin, yukarıdakilerin hepsi doğrudur gibi)

Algoritmanın pozitif puanlama kısmı Ripkey ve diğerlerinin (1996) puanlama önerisinin aynısıdır. Dengeli puanlama modelini ayıran kısım ceza puanı verme ve bunu dengeli olarak verme kısmıdır. Tablo 1’de puanlamada kullanılan kısaltmalar ve anlamları yer almaktadır.

Tablo 1

Dengeli Puanlama Yönteminde Kullanılan Kısaltmalar ve Anlamları

a max: sorudaki seçenek sayısı	c max: sorudaki doğru seçenek sayısı
ch: İşaretlenen seçenek sayısı	c ch: sorudaki doğru işaretlenen seçenek sayısı
p: bir sorudan alınabilecek temel puan [(p c) . (c ch)]	p max: sorudan alınabilecek maksimum puan
p c: doğru işaretlenen seçenek için alınacak puan (p max/ c max)	b: soru için temel tahmin etme seviyesi (c max / a max)
n: çok fazla seçenek işaretleyerek doğru cevabı tahmin etmeye çalışanların ihtimalini gösteren ölçme. Bu formül için “n” sayısı “b” den küçük ya da eşitse bu durum değerlendirmeye alınmaz (ch / a max)	s: tahmin etme davranışı için ceza (n-b)
s k: maksimal olası puanlar için ceza [(n-b). (p max / (1-b))]	T: nihai puan (p- s k)
$T \% = (p - s k / p \text{ max }) . 100$	

Çalışmanın sonunda akla gelen sorulardan biri olan puanlama uygulamasının öğrencilerden tarafından anlaşılıp anlaşılmadığı konusu öğrencilere sorulmuş ve öğrencilerin çoğunluğunun açık bir şekilde anladığı dönütü alınmıştır (Auer & Tarasowa, 2013).

Morgan Algoritması ve Dengeli Puanlama yöntemine göre bir öğrenci işaretlediği seçeneklere göre negatif puan da alabilir. Araştırmada kullanılacak puanlama türlerinin puanlanması, avantaj ve dezavantajları ile örnek bir puanlama uygulaması tablolar halinde gösterimi aşağıdaki gibidir.

Örnek uygulamada 5 madde vardır. Her maddenin 5 seçeneği vardır ve öğrencilere maddedeki doğru seçenek sayısı verilmemiştir.

Tablo 2

Örnek Uygulamadaki Beş Maddede Öğrencinin İşaretlediği Seçenekler ve Doğru Seçenekler

Soru Numarası	Öğrenci Cevabı	Doğru Cevap
1	A-B-C	A-B
2	D-E	B-D
3	C-D-E	A-B-C
4	A-E	A-E
5	D	D-E

Tablo 3

Çoklu Cevaplı Madde Türüne Ait Puanlama Yöntemleri İçin Genel Bilgiler ve Örnek Uygulamada Öğrencinin Alacağı Puanlar

Puanlama Yöntemi	Puanlama Kuralı	Olumlu Tarafları	Olumsuz Tarafları	Örnek uygulama puanı
İkili Puanlama	İşaretlenen tüm seçenekler doğru ve işaretlenmeyen tüm seçenekler yanlış ise 1 puan verilirken, diğer tüm durumlarda sıfır puan verilir.	Sayısal olarak basit puanlama yapılabilen algoritmalarından birisidir.	Kısmi puanlama yapılamaz.	1) 0 2) 0 3) 0 4) 1 5) 0
Tuzaklı Çoklu Cevaplı	İşaretlenen her doğru seçenek için 1 puan verilir, ancak herhangi bir yanlış işaretleme varsa sıfır puan verilir.	Kısmi puanlama yapılabilir. Tahmin etme davranışını azaltır.	Ceza puanı çok yüksektir.	1) 0 2) 0 3) 0 4) 2 5) 1
Tuzaklı İki Puanlı	İşaretlenen tüm seçenekler doğru ve işaretlenmeyen tüm seçenekler yanlış ise 2 puan verilirken, bazı doğru seçenekler işaretlenmişse ve hiçbir yanlış seçenek işaretlenmemişse 1 puan verilir. Herhangi bir yanlış seçenek işaretlenmişse ya da hiçbir doğru seçenek işaretlenmemişse sıfır puan verilir.	Kısmi puanlama yapılabilir. Tahmin etme davranışını azaltır.	Puanlamada bir netlik yoktur. “Bazı doğru seçenekler” ifadesi belirsizlik oluşturur. Ayrıca ceza puanı çok yüksektir.	1) 0 2) 0 3) 0 4) 2 5) 1

Negatif Puanlama	İşaretlenen her bir doğru seçenek için 1 puan verilirken, işaretlenen her bir yanlış seçenek için -1 puan verilir. Taban etkisi uygulanırsa minimum puan sıfırdır. Uygulanmazsa kişi negatif puanları da alabilir.	Kısmi puanlama yapılabilir. Tahmin etme davranışını azaltır.	Ceza puanı yüksektir. Taban etkisi uygulanmazsa öğrenci negatif puan alabilir.	1) 1 2) 0 3) 0 ya da -1 4) 2 5) 1
Anahtar Özellik	İşaretlenen doğru seçenek sayısının maddedeki doğru seçenek sayısına oranı kadar puan verilir. Çok uygunsuz bir cevabın seçilmesi durumunda sıfır puan verilir	Kısmi puanlama yapılabilir. Tıp fakültelerinde kritik öneme sahip konularla ilgili sınavlar için daha uygundur.	Puanlamada bir netlik yoktur. “Çok uygunsuz cevap” ifadesi kişiden kişiye değişebileceği için tartışmalı bir durum oluşturabilir.	1) 1 ya da 0 2) 0,5 ya da 0 3) 0,33 ya da 0 4) 1 5) 0,5
Çoklu Doğru/Yanlış	İşaretlenen her doğru seçenek için bir puan verilirken işaretlenmeyen her yanlış seçenek için bir puan verilir.	Kısmi puanlama yapılabilir. Öğrencinin cevaba ne kadar yaklaştığı hakkında bilgi sağlayabilir.	Herhangi bir seçenek işaretlenmeden ya da tüm seçenekler işaretlenerek puan alınabilir. Tahmin etme davranışını azaltmak için ceza puanı uygulanmamaktadır.	1) 4 2) 3 3) 1 4) 5 5) 4
Pozitif Sayım	İşaretlenen her bir doğru seçenek için bir puan verilir.	Kısmi puanlama yapılabilir. Puanlanması kolaydır.	Tüm seçenekler işaretlendiğinde tam puan alınabilir. Tahmin etme davranışını azaltmak için ceza puanı uygulanmamaktadır.	1) 2 2) 1 3) 1 4) 2 5) 1
Morgan Algoritması	İşaretlenen doğru seçenek sayısının maddedeki toplam doğru seçenek sayısına oranı kadar puan verilir. İşaretlenen yanlış seçenek sayısının maddedeki toplam yanlış seçenek sayısına oranı kadar ceza puanı verilir.	Kısmi puanlama yapılabilir. Tahmin etme davranışını azaltır.	Ceza puanı yüksektir. Öğrencinin negatif puan alma olasılığı vardır.	1) 0,67 2) 0,17 3) -0,67 4) 1 5) 0,5
Ripkey Algoritması	İşaretlenen doğru seçenek sayısının toplam doğru seçenek sayısına oranı kadar puan verilir. Toplam doğru seçenek sayısından daha fazla yapılan işaretlemelerde sıfır puan verilir.	Kısmi puanlama yapılabilir. Puanlamanın pozitif kısmı ayıricıdır. Tahmin etme davranışını söndürmek için ceza puanı uygulanır.	Ceza puanı dengeli değildir.	1) 0 2) 0,5 3) 0,33 4) 1 5) 0,5
Dengeli Puanlama	İşaretlenen doğru seçenek sayısının toplam doğru seçenek sayısına oranı kadar puan verilir.	Algoritma öğrencinin işaretlememe veya tamamını işaretleme	Öğrencinin negatif puan alma ihtimali vardır. Algoritma	1) 0,67 2) 0,5 3) 0,33

Toplam doğru seçenek sayısından daha fazla yapılan işaretlemelerde ceza puanı uygulanır. Ceza puanı ise işaretlenen seçenek sayısına bağlı olarak farklılaşır.	durumlarında puan verecek tasarlanmıştır. puanı dengelidir. Kısmi puanlama yapılabilir.	sıfır şekilde Ceza karmaşık gelebilir.	alışık kişilere	4) 1 5) 0,5
--	---	--	-----------------	----------------

Çoklu Doğru/Yanlış Puanlama Yöntemi'nde kişinin işaretlediği doğru seçeneklerin ve işaretlemediği yanlış seçeneklerin toplamının, toplam seçenek sayısına oranı ile elde edilecek puan 0-1 aralığında olacaktır. Bu sayede puanlamaların karşılaştırılmasında kolaylık sağlanacaktır. Auer ve Tarasowa (2013), Siddiqui vd. (2016) yaptıkları çalışmalarda puanlamaları kıyaslarken bu şekilde dönüşüm yapmışlardır. Pozitif Sayım Puanlaması da benzer şekilde 0-1 aralığında dağılacak şekilde uyarlanırsa kişinin doğru işaretlediği seçenek sayısını toplam doğru seçenek sayısına oranlamak yeterlidir. Negatif Puanlama yöntemi Domnich ve diğerlerinin (2015) yaptığı gibi 0-1 arası standart puan aralığına dönüştürülebilir. Bu uygulama için Muckle ve diğerlerinin (2011) belirttiği şekilde taban uygulaması yapılarak minimum alınacak puan sıfır olacak şekilde işaretlenen doğru seçenek sayısından işaretlenen yanlış seçenek sayısı çıkarılır ve bu sonuç toplam seçenek sayısına bölünür. Tuzaklı Çoklu Cevaplı puanlama yöntemi için ise kişinin doğru işaretlediği seçenek sayısını toplam doğru seçenek sayısına oranlanır. Ancak yanlış bir seçenek işaretlenmişse kişiye o madde için sıfır puan verilir. Anahtar Özellik ve Tuzaklı İki Puanlı puanlama yöntemlerinde netlik olmadığı için karşılaştırılacak puanlamalar arasına dahil edilmemişlerdir. Çünkü ilk puanlamada alakasız seçenek ifadesi testi yapan kişiden kişiye değişebileceği için belirsizlik oluştururken, ikinci puanlamada ise bazı doğru işaretlenen cevaplar ibaresi doğru sayısının her maddede farklılaşabileceğinden ötürü karmaşaya yol açacaktır.

1.3. Araştırmanın Önemi

Bilimsel ve teknolojik gelişmelerin bir sonucu olarak, çağdaş toplumlarda bireylerin sahip olması gereken özellikler farklılaşmakta, bu durum eğitim sistemlerine, öğretim programlarının büyük değişimler geçirmesi olarak yansımaktadır. Bu ilerlemeler ile ölçme ve değerlendirme alanında paralellik kaydedilememiştir. Ülkemiz öğrenci seçme ve yerleştirme sınavlarının yapısında bir değişikliğe gidilmemesi ve 1974'ten beri çoktan seçmeli soruların kullanımının sürdürülmesi bu durumu örneklemektedir (Yavuz & Bilgeç,

2016). Ölçme, Seçme ve Yerleştirme Merkezi Başkanlığı (ÖSYM) 2015 tarihinde bu konu hakkında duyuru niteliğinde şöyle bir açıklaması vardır:

“Zaman içerisinde bu uygulama, eğitim camiası içinde eleştirilmeye başlanmış, *çoktan seçmeli sorulardan* oluşan sınavlara karşılık yurt dışındaki uygulamalar, yurt dışındaki uygulamaların ülkemiz gerçeklerine göre alternatif çözüm önerileri kamuoyunda sıklıkla tartışılmıştır. Bu eleştirilerde de görüleceği üzere, çoktan seçmeli sorularla yapılan sınavların eğitim sistemimize olumsuz etkileri, sistemin tüm paydaşları tarafından tartışmasız kabul edilen bir gerçektir. Eleştirilerin odak noktasını; çoktan seçmeli sorularda adayların, cevabını bilmedikleri bir soruyu, % 20 oranında doğru cevaplama olasılıklarının olduğu ... oluşturmaktadır.”

Çoktan seçmeli maddeler ile oluşturulan testlere yapılan eleştiriler neticesinde MEB ve ÖSYM tarafından açık uçlu maddelerden oluşan testler yapılması için çalışmalar başlatılmış ancak sadece 2017 yılı Yüksek Öğretime Geçiş Sınavı (YGS) ve Lisans Yerleştirme Sınavı'nda (LYS) yer alan madde türü daha sonra sınavlardan çıkarılmıştır. Bunun sebebi olarak da toplumda açık uçlu maddelerin puanlanmasına yönelik eleştiriler ile öğrenci ve veli kaygıları en önemli etkenler olarak öne çıkmıştır. Geniş ölçekli sınavlarda, üst düzey düşünmeyi sağlayıp, kısmi puanlama imkanı veren ayrıca objektif ve seri olarak puanlanabilen çoklu cevaplı madde türünün kullanılabilecek uygun bir alternatif olduğu düşünülmektedir.

Geniş ölçekli testlerden sınıf içi başarı ölçümüne uzandığımızda ise öğretmenlerin yaklaşımları incelenmelidir. Karakaya ve N. Doğan (2008) yaptıkları çalışmada, öğretmenlerin ölçme araçlarını tercih ederken aradıkları özelliklerin başında öğrenci gelişimine uygunluk, ölçme aracının güvenilir olması, eğitsel yararın görülmesi ve ölçme aracının geçerli olmasının gelmekte olduğu ayrıca öğretmenlerin ölçme araçlarından en fazla çoktan seçmeli maddelerden oluşan testleri kullandıkları sonuçlarına ulaşmışlardır. Öğretmenlerin önerilen yeni ölçme araçlarını ise kullanmaya eğilimli olmadıkları araştırmada ulaşılan başka bir bulgudur. Bu bilgilerden hareketle araştırma konusu olan çoklu cevaplı madde türünün öğrenci gelişimine hitap eden, güvenilir ve çoktan seçmeli madde türüne benzerliği sebebiyle okullarda öğretmenlerin kullanım açısından zorluk çekmeyecekleri, kullanmaya istekli olacakları ve yeni tanıtılan ölçme araçlarındaki gibi kullanmaya mesafeli olmayacakları bir madde türü olacağı tahmin edilmektedir. Ayrıca çoktan seçmeli madde türüne göre şans başarısının daha az olması beklenmekte, tahmin etme davranışını söndüreceği umulmakta ve açık uçlu maddelerin dezavantajı olan objektif ve hızlı puanlanamama gibi olumsuzluklardan da uzak olacağı beklenmektedir. Üst düzey zihinsel becerilerin ölçülmesinde en etkili ölçme aracı olarak açık uçlu maddeler görülmektedir. Bunun en önemli sebeplerinden iki tanesi alternatif düşünmeyi ve

cevaplamayı sağlaması ile kısmi puan verebilme imkanının olmasıdır. Çoklu cevaplı madde türü bu iki fark oluşturan durumu karşılayabileceğinin düşünülmesinin yanında açık uçlu soruların en çok eleştirilen kısmı olan puanlamanın da objektif ve hızlı yapılmasını sağlamaktadır.

Çoklu cevaplı maddelerden oluşan testler hem kağıt tabanlı hem de bilgisayar tabanlı olarak düzey belirleyici testlerin haricinde de kullanılmaya uygun yapıya sahiptir. Örneğin biçimlendirici testlerde, tıpkı standart belirleme veya tanılama testleri için yapılan testler gibi, bireylerin ne kadar bildiklerinden çok tam olarak ne yaptıklarını ya da yapmadıklarını bulmak daha önemlidir. Verbic'e (2012) göre bu amaçla, yorumlama kolaylığı nedeniyle, test yapıcılar genellikle yalnızca iki madde türünü kullanırlar. Bunlardan birisi çoktan seçmeli maddeler iken diğeri yapılandırılmış yanıt maddeleridir. Ancak, özellikle bilgisayar tabanlı testlerde, tamamen kapalı ve tamamen açık maddeler arasındaki boşluğu dolduracak başka madde türlerine ihtiyaç duyulmaktadır.

Çoklu cevaplı maddeler, son yıllarda, daha geleneksel çoktan seçmeli madde türüyle erişilemeyen madde sunumundaki ve cevaplamadaki esnekliği nedeniyle dikkat çekmiştir. Dahası, çoklu cevaplı maddelerin, karmaşık yetenek, bilgi ve becerileri çoktan seçmeli maddelerden daha kolay ölçebildiği kabul edilmiştir. Çoklu cevaplı madde türü ayrıca performans değerlendirme görevlerine göre puanlamanın doğruluğu ve ekonomisi bakımından üstünlük göstermektedir (Ma, 2004). Esnek, eleştirel ve alternatifli düşünmenin önem kazandığı, 21. yüzyıl becerilerinin kazandırılması amaçlanan günümüz sistemine çoklu cevaplı madde türünün önemli katkılarının olacağı tahmin edilmektedir.

Üst düzey düşünme becerilerinin ölçülmesine yönelik girişimlere önemli katkı sunması beklenen çoklu cevaplı madde türünün, araştırmadan elde edilebilecek olumlu sonuçlar ışığında geniş ölçekli ve ayırt edici sınavlarda kullanımının sağlanmasına öncülük etmesi önemli bir durumdur . Bu sayede geniş ölçekli sınavlarda daha geçerli, şans başarısından daha az etkilenen ve yakın yetenek düzeyindeki kişiler arası farkı güvenilir biçimde ortaya çıkarabilmesi beklenen maddelerin kullanımının önünü açarak açık uçlu madde kullanımına gerek kalmadan alternatifli bakış açısıyla maddeleri çok yönlü inceleyip tek doğruya takılmadan kişileri esnek düşünmeye yönlendirebilmesi de araştırmanın önemli noktalarından bir tanesidir.

Çoklu cevaplı madde türü ölçülmek istenen becerinin özelliğine göre sınıf içi veya geniş ölçekli testlerde kullanılabilecek objektif, tek bir doğrunun peşinde koşturmayan, seri

puanlanabilen ve kısmi puan vermeye imkan sađlayan bir madde türüdür. Çoklu cevaplı maddelerin alternatif puanlama yöntemlerinin bilinmesi, test ve madde istatistiklerine etkilerinin incelenip karşılaştırılması bu test türünü kullanacaklar için yol gösterici olacaktır. Bu sayede ölçme ve değerlendirme uzmanı hangi amaçla test oluşturacaksa ona uygun puanlama yöntemini seçip uygulayabilecektir. Araştırmanın ölçme araçlarının kullanımındaki çeşitliliğe katkı sağlayacağı ve alanyazını zenginleştireceği düşünülmektedir. Ayrıca eğitimcilere öğrenmeleri değerlendirmede hangi madde türünden nasıl ve hangi durumlarda en uygun biçimde fayda sağlayacaklarına dair katkı sağlayacağı umulmaktadır. Bu alanda yapılan çalışmalarda çoklu cevaplı maddeler ya başka madde türleriyle karşılaştırılmış, ya birkaç puanlama yöntemi kendi içinde karşılaştırılmış ya da yazarların kendi önerdikleri puanlama yöntemleri anlatılmıştır. Bu çalışmanın farklılığı ve önemi ise çoklu cevaplı maddelerin en çok eleştirilen yönü olan puanlamadaki tutarsızlığı ve istatistiksel analizlerin zorluğunu ortadan kaldıracak şekilde puanlama yöntemlerinin standartlaştırılması ve bu standarda bağılı olarak rasyonel bir şekilde karşılaştırılması ve ayrıca standart hale getirilebilen sekiz farklı puan türünün aynı anda madde ve test özelliklerinin incelenmesidir.

1.4. Araştırmanın Amacı

Bu araştırmanın amacı başarı testlerinde ve sınıf-içi ölçme değerlendirme uygulamalarında sıklıkla kullanılan çoktan seçmeli maddelerin bir türü olan ve çoktan seçmeli maddelerin önemli eksik ve sınırlılıklarına sahip olmayan çoklu cevaplı madde türünün puanlanması sürecinde kullanılan puanlama yöntemlerini test ve madde istatistikleri açısından karşılaştırmak, bu madde türünü tanıtmak, bu madde türünün özelliklerini açıklamak ve çoklu cevaplı maddelerin ölçme değerlendirme uygulamalarında kullanımını yaygınlaştırmak için bir adım atmaktır. Çoklu cevaplı madde türünün puanlanmasında yaşanan karmaşayı ortadan kaldırıp ölçme amacına uygun puanlama yöntemlerinin önerilmesini sağlamak araştırmanın başka bir amacıdır.

1.5. Araştırma Soruları

Çalışma kapsamında “Çoklu cevaplı madde türünün puanlanmasında kullanılan yöntemlerin (Dengeli Puanlama, İkili Puanlama, Morgan Algoritması, Negatif Puanlama, Çoklu Doğru/Yanlış Puanlama, Pozitif Sayım, Tuzaklı Çoklu Cevaplı Puanlama, Ripkey Algoritması) test ve madde istatistikleri açısından aralarında anlamlı bir farklılık var mıdır?”

sorusu cevaplanacaktır.

Bu kapsamda aşağıdaki alt sorulara yanıtlar aranmıştır.

1. Çoklu cevaplı madde türü kullanılarak hazırlanan bir testin farklı puanlama yöntemleriyle elde edilen test puanlarına ait ortalama güçlükler arasında istatistiksel olarak anlamlı bir farklılık var mıdır?
2. Çoklu cevaplı madde türü kullanılarak hazırlanan bir testin farklı puanlama yöntemleriyle elde edilen test puanlarının güvenilirlik katsayıları arasında anlamlı bir farklılık var mıdır?
3. Çoklu cevaplı madde türünün puanlanmasında kullanılan yöntemlerden hesaplanan madde güçlük indeksleri arasında istatistiksel olarak anlamlı bir farklılık var mıdır?
4. Çoklu cevaplı madde türünün puanlanmasında kullanılan yöntemlerden hesaplanan madde ayırt edicilik indeksleri arasında istatistiksel olarak anlamlı bir farklılık var mıdır?
5. Çoklu cevaplı maddelerden oluşan başarı testlerine katılan öğrencilerin çoklu cevaplı maddeler ile çoktan seçmeli maddelerin karşılaştırılmasına ilişkin görüşleri nelerdir?

1.6. Araştırmanın Sınırlılıkları

1. Hazırlanan başarı testi 8. sınıf öğrencilerine uygulanmıştır.
2. Hazırlanan başarı testi Türkçe ve matematik dersleri ile sınırlıdır.
3. Hazırlanan başarı testinde sadece doğru cevabın sayısının öğrenciye verilmediği türe göre hazırlanan yönergeler ile testler uygulanmıştır.

1.7. Varsayımlar

1. Başarı testi uygulamasında güvenilirliği düşürecek herhangi olumsuz bir koşul olmamıştır.
2. Başarı testi uygulaması öncesi yönerge tüm öğrenciler tarafından tam olarak anlaşılmıştır.

BÖLÜM II

İLGİLİ ARAŞTIRMALAR

Bu bölümde, çoktan seçmeli madde türü ve çoklu cevaplı madde türü ile ilgili çalışmalara yer verilmiştir.

2.1. Çoktan Seçmeli Madde Türü İle İlgili Araştırmalar

Özdemir (2003) “Çoktan Seçmeli Testlerde İki Kategorili ve Önsel Ağırlıklı Puanlamanın Diferansiyel Madde Fonksiyonuna Etkisi İle İlgili Bir Araştırma” isimli çalışmasında, iki kategorili ve önsel ağırlıklı puanlama yöntemleri kullanılarak puanlanan bir testin, cinsiyet grupları ve okul türüne göre diferansiyel madde fonksiyonunu ortaya koymayı amaçlamıştır. Çalışma Türkçe okuduğunu anlama yeteneğini ölçme amacı ile araştırmacı tarafından geliştirilen 20 maddelik çoktan seçmeli maddelerle oluşturulan test yardımıyla özel okul ve devlet okullarında 4. ,5. ,6. ve 7. sınıflarda öğrenim gören 1478 öğrenciye uygulanmıştır. Çalışma sonunda iki kategorili ve önsel ağırlıklı puanlama bakımından testte yer alan maddelerin büyük çoğunluğunun okul türüne göre de cinsiyet faktörüne göre de önemli düzeyde yanlılık göstermediği bulunmuştur. Sadece bir maddenin her iki alt grup için de orta düzey (B düzeyi) yanlılık gösterdiği ve bu yanlılığın okul türü bakımından özel okullarda okuyan öğrenciler lehine ve cinsiyet bakımından erkek öğrenciler lehine olduğu belirlenmiştir.

Kan ve Kayapınar (2006) “Yabancı Dil Eğitiminde Aynı Davranışları Yoklayan Çoktan Seçmeli ve Kısa Cevaplı İki Testin Madde ve Test Özelliklerinin Karşılaştırılması” adlı çalışmalarında, aynı hedef davranışları ölçmeyi amaçlayan çoktan seçmeli ve kısa cevaplı maddelerin ve bu maddelerden oluşan iki İngilizce testinin madde güçlük indeksleri, madde ayırt edicilik gücü indeksleri, test güvenirlikleri, bir ölçüte dayalı geçerlik katsayıları ve test

puanları ortalamaları arasında farklılıklar olup olmadığını araştırmışlardır. 139 öğrenci üzerinde yapılan çalışmada her biri 4 seçenekli 19 çoktan seçmeli maddeden oluşan test ile yine 19 sorudan oluşan kısa cevaplı test uygulanmıştır. Çalışmanın sonuçlarına göre bu iki testin madde güçlük indeksleri, madde ayırt edicilik gücü indeksleri ve test güvenirlikleri arasında belirgin farklılıklar bulunmasına rağmen test ortalamaları olarak belirgin bir farklılık bulunamamıştır. Ayrıca kısa cevaplı testlerin, çoktan seçmeli testlere göre öğrencilerin final sınavı puanlarını daha iyi yordadığı belirlenmiştir.

Gültekin (2011), “Çoktan Seçmeli, Açık Uçlu ve Karma Testlerin Psikometrik Özelliklerinin Madde Tepki Kuramına Dayalı Olarak Değerlendirilmesi” adlı çalışmasında matematik başarısına ilişkin olarak benzer yapıları ölçen iki madde tipinin (çoktan seçmeli ve yanıt sınırlı açık uçlu) tek başına ve birlikte kullanıldıkları koşullar ile iki madde tipinin test içinde yer alma yüzdelerinin test formlarında farklı olduğu durumlarda MTK’ya göre kestirilen madde ve yetenek parametrelerinin psikometrik niteliklerinin, madde ve test bilgi fonksiyonlarının, göreceli etkililik düzeylerinin ve ölçüt dayanaklı geçerlik kanıtlarının farklı olup olmadığını araştırmıştır. Çalışma sonucunda uzun karma testlerin, sadece çoktan seçmeli ve sadece açık uçlu maddelerden oluşan testlere göre orta güçlükte ve daha ayırıcı testler olduğu belirlenmiştir. Ayrıca, orta yetenek düzeyinde uzun karma testlerin daha az standart hatayla yetenek kestirimi yaptığı ortaya konmuştur. Buna ek olarak da, testlerde, yanıt sınırlı açık uçlu maddelerin çoktan seçmeli maddelerden daha fazla bilgi sağladığı çıkarımı yapılmıştır.

Pehlivan (2011) “Çoktan Seçmeli Türkçe Test Maddelerinde Yanıtlama Davranışlarının İncelenmesi” isimli çalışmasında, ilköğretim 8. sınıf öğrencilerinin çoktan seçmeli Türkçe test maddelerinde gösterdikleri yanıtlama davranışlarının maddelerin ölçtüğü boyuta, maddelerin psikometrik özelliklerine, alt puan, orta puan, üst puan grubundaki öğrencilere ve cinsiyete göre nasıl bir değişim gösterdiğini belirlemeyi amaçlamıştır. Bu kapsamda yirmi maddelik bir test oluşturulmuş ve öğrencilere uygulanmıştır. Çalışmanın sonuçlarına göre öğrencilerin en çok gösterdikleri tepki, soruyu ve seçenekleri okuyup ona göre doğru yanıt olarak “düşündüğüm seçeneği işaretledim” şeklindedir ve bu tepki madde güçlüğü azaldıkça daha çok görülmektedir. Ayrıca çoktan seçmeli testlerin en çok eleştirilen yönü olan şans başarısının hem yanıtlayıcı puanlarında hem de madde ve test özelliklerinde oluşturduğu olumsuz etkilerin azaltılabilmesi için test sorularını yanıtlama sürecinin ve bu süreç içerisinde gerçekleşen yanıtlama davranışlarının dikkate alınması gerektiği ileri sürülmüştür.

Baghaei ve Amrahi (2011) “Çoktan Seçmeli Maddelerin Seçenek Sayısının Madde Psikometrik Özellikleri Üzerindeki Etkisi” adlı çalışmalarında çoktan seçmeli maddeler için ideal seçenek sayısını belirlemeyi amaçlamışlardır. Madde istatistiklerini, çeldiricilerin performans indekslerini, kişi istatistiklerini ve test güvenilirliklerini karşılaştırmak için, sadece madde başına düşen seçenek sayısı farklı olan üç benzer kelime test formu (A, B ve C) kullanılarak ortak bir eşitlik tasarımı kullanılmıştır. Madde başına düşen seçenek sayısından ters olarak etkilendiğini ortaya koyan çeldiricilerin ayırıcılık gücü hariç, üç test formu için madde güçlüklerinde ($p < 0.05$), madde uyum istatistiklerinde ve kişi yanıt davranışlarında ve güvenilirlikte önemli bir değişiklik gözlenmemiştir. Ayrıca daha fazla çeldiricili çoktan seçmeli testler hazırlamak için gereken zaman ve enerji miktarı göz önüne alındığında, madde başına üç seçeneğin optimal olduğu sonucuna varılmıştır.

S. Doğan (2012) “Kavram Haritası Ve Yapılandırılmış Grid Tekniğinin Çoktan Seçmeli Testlerle Karşılaştırılması” isimli çalışmada aynı kazanımları ölçmek üzere geliştirilmiş çoktan seçmeli test, yapılandırılmış grid ve kavram haritası tekniklerinin öğrencilerin karne puanını ve Seviye Belirleme Sınavı (SBS) düzeltilmiş puanlarını yordama düzeylerini incelemiştir. Çalışmada Fen ve Teknoloji Dersi ‘Maddenin İç Yapısına Yolculuk’ ünitesi ele alınmıştır. 15 maddeden oluşan yapılandırılmış grid, 15 maddeden oluşan çoktan seçmeli test ve öğrencilere hazırlatılan kavram haritaları kullanılmıştır. Yordama ilişkisinde öğrencilerin karne notu ve SBS düzeltilmiş puanları esas alınmıştır. Araştırma sonunda karne puanlarının, SBS düzeltilmiş puanlarına göre daha iyi yordandığı belirlenmiştir ve karne puanlarının yordanmasında da en etkili ölçeğin çoktan seçmeli test olduğu görülmüştür.

Koçak (2013), “Farklı Yönergelerle Verilen Çoktan Seçmeli Testlerde Yanıtlama Davranışlarının İncelenmesi” adlı çalışmada çoktan seçmeli testlerde gösterilen yanıtlama davranışlarının farklı yönergelere, cinsiyete, başarı düzeyine, bir dershaneye gidip gitmeme durumuna, okul türüne, farklı güçlük ve ayırt edicilik indekslerine sahip maddelere göre nasıl bir dağılım gösterdiğini saptamayı amaçlamıştır. Araştırmada 3 farklı yönerge ile sunulan 8. sınıf öğrencileri için hazırlanmış çoktan seçmeli matematik testi ve yanıtlama davranışları formu kullanılmıştır. Araştırmadan elde edilen bulgular, farklı okul türlerinde ve farklı yönergeler ile uygulanan testlerin bütününde en çok gösterilen yanıtlama davranışının “Soruyu çözdüm, bulduğum sonucu seçenekler arasında bulup işaretledim.” olduğu belirlenmiştir. Farklı yönergeler ile uygulanan testlerin bütününde gösterilen

yanıtlama davranışlarının cinsiyetlere göre dağılımı incelendiğinde, kız ve erkek öğrencilerin gösterdikleri yanıtlama davranışlarının benzer olduğu sonucuna ulaşılmıştır. Farklı yönergeler ile uygulanan testlerin bütününde gösterilen yanıtlama davranışlarının, öğrencilerin bir önceki dönem matematik notlarına ve madde güçlük indekslerine (kolay-orta-zor) göre dağılımları incelendiğinde, tüm yönerge türlerinde gösterilen yanıtlama davranışlarının farklı dağılım gösterdiği belirlenmiştir. Yanıtlama davranışlarının madde ayırt edicilik indekslerine (düşük-yüksek) göre farklı dağılım gösterdiği sonucuna ulaşılmıştır.

Eren (2015), “Çoktan Seçmeli ve Karma Test Uygulamalarına İlişkin Öğrenci Başarıları İle Öğrenci ve Öğretmen Görüşlerinin Karşılaştırılması” isimli çalışmasında, matematik başarısına ilişkin olarak aynı öğrenme hedeflerini ölçen, yalnızca çoktan seçmeli maddelerden oluşan başarı testi ile çoktan seçmeli ve yanıtı sınırlı açık uçlu maddelerin yer aldığı karma başarı testinden alınan puanlar arasındaki ilişkilerin incelenmesi ile öğretmen ve öğrencilerin farklı test formatlarındaki testlere ilişkin görüşlerini belirlemeyi amaçlamıştır. Çalışmada iki farklı formattaki matematik testi öğrencilere uygulanmış, bu farklı testlere ilişkin öğrenci ve öğretmen görüş anketi uygulanmış ve ayrıca geçmiş yıllara ait öğrencilerin matematik notları ve genel notları çalışmaya eklenmiştir. Araştırma sonucunda, iki farklı formattaki testten alınan toplam puanlar arasında pozitif yönlü yüksek ve manidar bir ilişki olduğu belirlenmiştir. Karma testin yalnız çoktan seçmeli maddelerden oluşan testten manidar düzeyde daha güvenilir olduğu gözlenmiştir. Ayrıca öğrencilerin test formatları ile ilgili görüşleri incelendiğinde öğrencilerin çoğunlukla yalnız çoktan seçmeli maddelerden oluşan testi tercih ettikleri ve bunun nedenlerinin farklılık gösterdiği ortaya çıkmıştır. Öbür taraftan, öğretmenlerin karma testleri öğrencilerin öğrenme süreçlerini ve başarılarını daha isabetli belirleyebilme ve öğrencilere dönüt verebilme açısından tercih ettikleri, fakat çoktan seçmeli maddeleri daha kullanışlı buldukları ortaya çıkmıştır.

2.2. Çoklu Cevaplı Madde Türü İle İlgili Araştırmalar

Hsu vd. (1984) çalışmalarında, madde güçlüğü ve ayırt edicilik indeksleri ve altı farklı puanlama formülünü kullanarak tek doğru cevaplı ve çoklu cevaplı maddeler için ortalama madde bilgisi hesaplamış ve karşılaştırmışlardır. Çalışmada kullanılan formüllerden ikisi hem tek doğru cevaplı hem de çoklu cevaplı maddeler için kullanılabilirken diğer dört tanesi çoklu cevaplı maddeler için özel olarak tasarlanmıştır. Bu çalışmada yer alan enstrümanlar,

Çin Halk Cumhuriyeti'ne bağlı Tayvan'daki 1979 Ortak Kolej Giriş Sınavının üç alt testi idi. Bu testler, Çince, San-Min-Chu I (Yat-sen'in politik teori dersi) ve tarihtir. Çalışmaya göre çoklu cevaplı maddeler tek doğru cevaplı maddelere göre sürekli olarak daha zor ve daha ayırt edicidir. Madde güvenirliğinin verimliliği için bir madde türünün diğerine göre göreceli olarak önemli bir üstünlüğü bulunmamıştır. Yalnızca Çince alt testinde çoklu cevaplı maddelerin tek doğru cevaplı maddelere göre güvenirliği biraz daha yüksek bulunmuştur. Araştırmaya göre ortalama madde bilgi fonksiyonu olarak tanımlanan göreceli verimlilik için, tek doğru cevaplı maddeler ortalamanın altındaki kişiler için daha fazla bilgi sağlarken, ortalama ve ortalamanın üstünde olan kişiler için çoklu cevaplı maddeler, tek doğru cevaplı maddelere göre çok daha üstündür. Bu eğilim hem Çince hem de tarih alt testleri için tutarlı bulunmuştur. Çoklu cevaplı maddelerin yüksek yetenekli kişiler için daha fazla bilgi vermesinin olası bir nedeni olarak da, çoklu cevaplı maddelerin yüksek yetenekli kişiler arasında daha iyi ayırım yapabilmesi gösterilmiştir. Araştırmadan elde edilen sonuçlar, çoklu cevaplı maddelerin kullanımında, özellikle de ortalama ve üstünde olan kişilerin test edilmesi sırasında belirli yararların olduğunu göstermiştir. Çoklu cevaplı maddelerin puanlanması için kullanılan formüllerin hiçbirisi diğerlerine göre belirgin bir üstünlük sergilememiştir. Ayrıca tahmin için düzeltmenin uygulanması gereksiz görünmektedir, ancak kısmi doğru cevaplar için kısmi puan verilmesinin cazip bir durum olduğu belirtilmiştir.

Ma (2004) çalışmasında, çoklu cevaplı test maddelerinin alternatif puanlanması üzerine incelemelerde bulunmuştur. Çoklu cevaplı madde türlerinin ikili ve çoklu puanlandığı ancak farklı seçenek ağırlıklandırma yöntemlerinin güvenirliğini ve geçerliğini inceleyen çalışmaların Klasik Test Kuramı'na (KTK) göre yapıldığı belirtilen çalışmada araştırmacı MTK'ya dayalı çoklu cevaplı puanlama yöntemlerini araştırmıştır. Bu kapsamda yaklaşık 3000 kişiye 70 maddeden oluşan içerisinde 1/7 ile 1/3 oranında çoklu cevaplı test maddeleri içeren ve 6 paralel testten oluşan bir sertifika sınavı uygulanmıştır. Bu çalışmadan elde edilen bulgular, çoklu cevaplı test maddelerinin çoklu puanlanmasının ölçüm verimliliğini ve sınıflandırma doğruluğunu artırma potansiyeline sahip olduğunu göstermektedir. Çoklu cevaplı maddelerin ortalama olarak daha zor, daha ayırıcı ve çözmek için daha fazla zaman gerektirdiği tespit edilmiştir. Ayrıca MTK modelleri kullanılarak seçenek ağırlıklandırma ve çoklu puanlama, belirli durumlarda test puanlarının psikometrik özelliklerini geliştirdiği çıkarımı yapılmıştır.

Eggen ve Lampe (2011) çalışmalarında, çoklu cevaplı test maddeleri ile eşleştirme maddeleri ve dizilim maddeleri için farklı puanlama yöntemlerini karşılaştırmışlardır. Bu çalışma, farklı bilgi düzeylerini tutarlı bir şekilde temsil etmenin uygunluğunu belirlemek için tek bir madde düzeyinde ikili ve çoklu puanlama yöntemlerini karşılaştırmaya yönelik ampirik olmayan bir girişimdir. Çalışmada bu maddelerin hem kişilerin farklı bilgi seviyelerini temsil etme yöntemleri hem de puanlamanın güvenilirliğine dair imkanları incelenmiştir. Çalışma sonucunda üç madde türünün her bir bilgi seviyesi için puanlama yöntemlerinin sonuçlarının bir olasılık dağılımı türetilmiş ve hesaplanmıştır ve farklı bilgi düzeylerine sahip kişi grupları için genelleştirilebilir olacağı belirtilmiştir.

Muckle vd. (2011) yaptıkları çalışmada, çoklu cevaplı madde türünü ve bu madde türünü kapsayan puanlama konularını incelemişlerdir. Bu kapsamda maddeye doğru cevap sayısını bildirecek ek ipucu dahil etmenin veya etmemenin etkilerini, ikili ve çoklu puanlamanın neler oldukları ile bunların ölçme sonuçlarına olan etkilerini ve kişi yeterlik düzeyi tahminine olan etkilerini ve son olarak da madde güçlüğü, madde ayırt ediciliği gibi madde özelliklerine olan etkilerini araştırmışlardır. Araştırmacılar çoklu cevaplı ve çoktan seçmeli türündeki maddeleri içeren dört test programından topladıkları verileri kullanmışlardır. Topladıkları verilerdeki bütün çoktan seçmeliler ikili puanlanırken, çoklu cevaplılar altı farklı puanlama yöntemiyle puanlanmıştır. Araştırma sonunda sınavı alan kişiye cevap sayısı hakkında ipucu vermekle 3 tane çoklu puanlamanın (Tuzaklı Çoklu Cevaplı, Tuzaklı İki Puanlı ve Negatif Puanlama) yapılamayacağı, ipucu verilmeyip tüm doğruları seçmesi istenirse de Pozitif Sayım puanlama yöntemini tahmin etme davranışını artıracığı belirtilmiştir. Bu sebeple puanlamanın nasıl yapılacağına karar verme işini soru hazırlamadan önce yapmak gerektiği belirtilmiştir. Puanlama modellerinin ortalama madde güçlükleri ve bilgi grafikleri ölçekte farklı noktalarda bilgi sağlamıştır. Bu sebeple puanlama modelinin seçiminin değerlendirilecek örnekleme bağılı olarak belirlenmesi gerektiği vurgulanmıştır. “Tuzaklı” puanlamaların diğer puanlama türlerinden daha fazla bilgi sağladığı bulunmuş ve bunun nedenlerinin ayrıntılı araştırılması gerektiği belirtilmiştir. Ayrıca tahmin etme davranışının bazı puanlama türleri için ciddi bir sorun kaynağı olabileceği belirtilmiştir.

Verbic (2012) çalışmasında, test oluşturma sürecinin erken dönemlerinde madde psikometrik özelliklerinin iyi bilinmesinin, testi oluşturanlar için önemli olduğunu ve bu sebeple çoklu cevaplı maddelerden elde edilen bilgi değerini artırabilecek birkaç puanlama

yöntemi önermiş ve değerlendirmiştir. Çalışma 2009 yılında Eğitim Kalitesi ve Değerlendirilmesi Enstitüsünde gerçekleştirilmiş ve 4. sınıf öğrencilerine bir kurs kapsamında uygulanan test ile veriler elde edilmiştir. Çalışma sonuçlarına göre çoklu cevaplı maddelere verilen cevaplar genellikle diğer türlerle yapılan çalışmalardan daha fazla bilgi taşır. Ayrıca çoklu cevaplı maddeler iki şekilde puanlanabilir. Bunlar, maddelere puan vermek veya maddedeki tüm seçeneklere puan vermek şeklindedir ki bu şekilde, puanlanan maddelerinin sayısını artırılır ve sonuç olarak testin güvenilirliğini artırılmış olur.

Auer ve Tarasowa (2013), yaptıkları çalışmada çoktan seçmeli maddelerin popülerliğinin sebeplerinden birinin otomatik puanlama ve değerlendirme olduğunu belirtmişlerdir ve çoklu cevaplı maddelerin puanlanmasında matematiksel temele dayalı, tahmin davranışını algılayan ve çoktan seçmeli maddelerle çoklu cevaplı maddelerin karışık olduğu bir testte de uygulanabilecek puanlama modeli önermişlerdir. Puanlama yaklaşımlarının mantığını, puanlama işleminin cevabın tahmin edilme seviyesine dayanması gerektiği fikri oluşturmaktadır. Çalışma içerisinde verilen başka puanlama yaklaşımlarına yaptıkları tahmin etme davranışını azaltmak için ceza puanı verilmemesi ya da yüksek ceza puanı verilmesini eleştiren araştırmacılar ürettikleri puanlama sisteminin uygulamalarının da mevcut puanlama yöntemlerinden daha iyi sonuçlar verdiğini yaptıkları uygulamalarla kanıtladıklarını belirtmişlerdir.

BÖLÜM III

YÖNTEM

Bu bölümde araştırmanın modeli, çalışma grubu, veri toplama araçları hakkında bilgiler verilmiş; araştırma için kullanılan verilerin toplanması ve verilerin hangi analizler ile yapıldığı belirtilmiştir.

3.1. Araştırma Modeli

Araştırmada ortaokul sekizinci sınıf öğrencilerine uygulanacak olan başarı testlerinde kullanılacak çoklu cevaplı maddelerden oluşan testlerin puanlanmasında kullanılacak yöntemlerin test ve madde istatistikleri açısından karşılaştırılması ve incelenmesi amaçlanmıştır. Bu sebeple araştırmanın modeli temel araştırmadır. Temel araştırma, odaklandığı problemin çözümü için ihtiyaç duyulan bilginin üretimi amacı ile yapılan araştırmadır. Pratik ihtiyaçlardan kaynaklandığı gibi, “bilgi, bilgi içindir” kabullü tamamen kuramsal ihtiyaçlara dönük olarak, var olan bilgiye yenilerini katmak için de yapılabilir (Karasar, 2016, s. 46). Ayrıca bu araştırma bir örneklemden elde edilen bulguların örneklemin temsil ettiği evrene genellenmesi gibi bir amaç taşımadığı ve çoklu cevaplı maddeleri puanlama yöntemlerinin karşılaştırılması yoluyla alanyazına yeni bilgiler katmayı amaçladığı için temel araştırma niteliği taşır.

3.2. Çalışma Grubu

Araştırmanın çalışma grubunu 2019-2020 eğitim öğretim yılında Uşak İli ve ilçelerinde eğitim görmekte olan sekizinci sınıf öğrencilerinden araştırma amacı doğrultusunda sınavla

öğrenci alan liselere gitme amacıyla olup LGS'ye girecek öğrenciler arasından başarı testlerini cevaplamaya gönüllü olanları oluşturmuştur. Bu sebeple çalışma grubu olasılığa dayalı olmayan örnekleme yöntemlerinden amaçlı örnekleme yöntemi ile oluşturulmuştur.

3.3. Veri Toplama Araçları

Araştırmacı tarafından matematik ve Türkçe derslerinin kazanımları doğrultusunda 8. sınıf seviyesine göre geliştirilmiş olan çoklu cevaplı madde türünden oluşan başarı testleri araştırmanın veri toplama araçlarını oluşturmaktadır. Başarı testleri hazırlanırken öğrencinin öğrenim gördüğü sınıf seviyesi ve ilgili dersin öğretim programında yer alan kazanımlara uygunluğu dikkate alınmıştır. Oluşturulan maddeler 11 kişiden oluşan uzman grubundan alınan görüşler çerçevesinde düzenlenmiştir. Uzman grubunu; ders kitaplarının yazım ve imla kuralları bakımından incelenmesi için oluşturulan milli eğitim müdürlüğü ders kitapları yazma komisyonunda görevli bir uzman, ölçme ve değerlendirme alanında bir Dr. ve iki Doç. Dr. olmak üzere üç öğretim üyesi, öğretmenlik deneyimi olan ve matematik eğitimi alanında Dr. olan bir uzman ve halen MEB'de görev yapan üç ilköğretim matematik öğretmeni ile üç Türkçe öğretmeni oluşturmaktadır. Testlerin deneme formlarının geliştirilmesi aşamasında öncelikle 8. sınıf Türkçe ve matematik derslerinin kazanımları incelenmiştir. Kazanımlara uygun olarak çoklu cevaplı madde formatına uygun madde yazma çalışmaları yapılmış ve her iki ders için de madde havuzu oluşturulmuştur. Oluşturulan madde havuzundaki maddeler önce uzman grubundaki alana hakim ve tecrübeli Türkçe ve matematik öğretmenlerinin görüş ve önerilerine sunulmuştur. Elde edilen dönütler doğrultusunda bazı maddeler tekrar düzenlenmiş ve matematik eğitimi alanında Dr. olan uzmanın görüşlerine sunulmuştur. Sonrasında maddeler uzman grubunda yer alan Eğitimde Ölçme ve Değerlendirme Anabilim Dalı öğretim üyelerinin incelemesinden geçmiştir. İnceleme sonrası yazım, imla ve düzen yönlerinden kitap yazma komisyonundaki uzmandan görüş alınmıştır. Son halini alan maddelerden 20'ser tanesi kazanımların ağırlığına göre havuzdan çekilmiş ve Türkçe ile matematik derslerine ait çoklu cevaplı madde türünden oluşan iki test oluşturularak deneme uygulaması için hazır hale getirilmiştir. Öğrenciler aşına olmadıkları bir madde türü ile karşılaşacakları için uygulanacak olan testlerin yönergelerine çoklu cevaplı madde türünü açıklayıcı temel bilgiler konulmuştur. Yönergede madde türü tanıtılmış, her bir maddenin en az bir doğru cevabının olacağı, düzeltme işleminin puanlama türüne göre uygulanacağı belirtilip şansla tahminde bulunulmaması gerektiği belirtilmiştir.

Türkçe ve matematik testlerinin deneme uygulaması farklı tarihlerde Uşak İli ve ilçelerinde bizzat tez yazarı ya da tez yazarının bilgilendirdiği öğretmenler tarafından yapılmıştır. Uygulamaların farklı zamanlarda yapılması sebebiyle iki teste katılan öğrenci sayıları farklılaşmış ve her uygulama öncesinde madde formatı üzerinde durularak yönergeler öğrencilere okunmuştur. Uygulama sonrasında elde edilen veriler istatistik paket programları ile analiz edilmiştir. Matematik sınavına ait verilerin analizi aşağıdaki gibidir. Yapılan analiz ile sekiz farklı puan türüne ait betimsel istatistikler, madde güçlük indeksleri, testlerin ortalama güçlükleri, madde ayırt edicilik indeksleri ve test güvenirlikleri hesaplanmıştır.

Tablo 4

Çoklu Cevaplı Madde Türü İle Oluşturulan Matematik Testinin Puanlama Yöntemlerine Göre Betimsel İstatistikleri

Betimsel İstatistikler	Puanlama Yöntemlerine Göre Betimsel İstatistik Değerleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
A.O.	35,70	21,27	26,44	29,99	43,6	37,7	25,8	30,4
Ortanca	35,45	20	27,15	28,40	44	38,35	24,15	27,50
Std. Sapma	16,81	17,33	21,38	17,83	16,30	16,57	17,57	16,35
Ranj	81,70	75	82,50	83,35	84	83,35	83,4	81,15
Çarpıklık	0,263	0,618	0,080	0,354	0,074	0,296	0,470	0,486
Basıklık	-0,428	-0,464	-0,839	-0,645	-0,003	-0,069	-0,405	-0,321

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 4 incelendiğinde puanlama yöntemlerine ait aritmetik ortalama ve ortanca değerlerinin birbirine çok yakın değerler aldıkları; çarpıklık değerlerinin pozitif ve sıfıra yakın (hafif sağa kayışlı), basıklık değerlerinin negatif ve sıfıra yakın (hafif basık) değerler aldıkları görülmektedir. Ranj değerlerine bakıldığında en düşük ve en yüksek puanlar arasındaki farkın oldukça fazla olduğu dikkati çekmektedir.

Tablo 5

Çoklu Cevaplı Madde Türü İle Oluşturulan Matematik Testine Ait Maddelerin Güçlük İndekslerinin Puanlama Yöntemlerine Göre Gösterimi

Madde No	Puanlama Yöntemlerine Göre Testteki Maddelerin Güçlük İndeksleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
1	0,79	0,66	0,74	0,74	0,82	0,78	0,68	0,73
2	0,57	0,39	0,46	0,50	0,67	0,65	0,39	0,43
3	0,36	0,16	0,21	0,28	0,52	0,41	0,19	0,26
4	0,13	0,01	0,09	0,01	0,26	0,23	0,02	0,01
5	0,37	0,15	0,18	0,28	0,54	0,45	0,20	0,26
6	0,43	0,26	0,35	0,40	0,56	0,44	0,38	0,41
7	0,31	0,11	0,19	0,21	0,42	0,33	0,18	0,27
8	0,47	0,38	0,35	0,43	0,55	0,51	0,38	0,41
9	0,50	0,40	0,37	0,46	0,52	0,46	0,44	0,49
10	0,11	0,01	-0,03	0,04	0,21	0,14	0,03	0,08
11	0,43	0,25	0,35	0,37	0,53	0,41	0,36	0,41
12	0,46	0,28	0,38	0,39	0,48	0,42	0,37	0,44
13	0,60	0,33	0,51	0,49	0,61	0,65	0,37	0,43
14	0,32	0,19	0,26	0,30	0,32	0,30	0,26	0,32
15	0,06	0,06	0,01	0,08	0,19	0,11	0,06	0,06
16	0,21	0,04	0,13	0,20	0,30	0,23	0,15	0,18
17	0,09	0,01	0,00	0,09	0,10	0,09	0,04	0,09
18	0,31	0,21	0,22	0,25	0,33	0,29	0,22	0,28
19	0,36	0,17	0,30	0,30	0,43	0,34	0,25	0,33
20	0,27	0,18	0,22	0,18	0,36	0,30	0,18	0,18
Testin								
Ortalama	0,36	0,21	0,26	0,30	0,44	0,38	0,26	0,30
Güçlüğü								

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 5 incelendiğinde madde güçlüklerinin ve testlerin ortalama güçlüğüne genel olarak sıfıra yakın değerler aldıkları görülmektedir. Bu da maddelerin ve testlerin zor olduğu çıkarımında bulunmamıza neden olur.

Tablo 6

Çoklu Cevaplı Madde Türü İle Oluşturulan Matematik Testinin Puanlama Yöntemlerine Göre Test Güvenirlikleri

Puanlama Yöntemleri	Güvenirlik Türü	Test Güvenirliği
Dengeli Puanlama	Cronbach Alpha	0,79
İkili Puanlama	KR-20	0,80
Morgan Algoritması	Cronbach Alpha	0,81
Negatif Puanlama	Cronbach Alpha	0,83
Çoklu Doğru/Yanlış Puanlama	Cronbach Alpha	0,78
Pozitif Sayım Puanlama	Cronbach Alpha	0,77
Tuzaklı Çoklu Cevaplı Puanlama	Cronbach Alpha	0,82
Ripkey Algoritması	Cronbach Alpha	0,78

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 6 incelendiğinde sekiz puanlama türü ile elde edilen güvenilirliklerin yeterince yüksek oldukları söylenebilir.

Tablo 7

Çoklu Cevaplı Madde Türü İle Oluşturulan Matematik Testine Ait Maddelerin Ayırt Edicilik İndekslerinin Puanlama Yöntemlerine Göre Gösterimi

Madde No	Puanlama Yöntemlerine Göre Testteki Maddelerin Ayırt Edicilik İndeksleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
1	0,46	0,53	0,56	0,55	0,33	0,39	0,52	0,45
2	0,62	0,69	0,70	0,71	0,47	0,45	0,71	0,70
3	0,43	0,42	0,45	0,49	0,37	0,45	0,42	0,38
4	0,18	0,20	0,26	0,18	0,08	0,11	0,11	0,19
5	0,50	0,48	0,48	0,55	0,41	0,39	0,47	0,39
6	0,35	0,33	0,34	0,38	0,43	0,35	0,41	0,37
7	0,32	0,26	0,33	0,31	0,32	0,31	0,31	0,30
8	0,54	0,59	0,57	0,58	0,41	0,45	0,55	0,52
9	0,62	0,62	0,60	0,63	0,55	0,59	0,62	0,60
10	0,12	-0,06	0	0,06	0,34	0,28	0,07	0
11	0,65	0,58	0,64	0,66	0,59	0,61	0,63	0,64
12	0,59	0,56	0,59	0,65	0,65	0,58	0,65	0,58
13	0,53	0,52	0,55	0,55	0,44	0,51	0,44	0,40
14	0,54	0,56	0,55	0,57	0,51	0,50	0,52	0,55
15	0,24	0,31	0,26	0,30	0,39	0,29	0,32	0,33
16	0,35	0,08	0,41	0,35	0,52	0,40	0,45	0,37
17	-0,03	-0,03	0,12	-0,05	0,11	0,10	-0,07	-0,05
18	0,54	0,59	0,58	0,60	0,53	0,55	0,57	0,53
19	0,55	0,56	0,50	0,53	0,57	0,56	0,57	0,55
20	0,53	0,48	0,43	0,51	0,63	0,54	0,52	0,52

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 7 incelendiğinde maddelerin büyük çoğunluğunun yeterince ayırt edici olduğu söylenebilir. Matematik başarı testine ait tüm veriler incelendiğinde 4-10 ve 17. maddelerin madde güçlüklerinin ve madde ayırt ediciliklerinin çok düşük olduğu görülmüştür. Bu

sebeple bu üç madde testten çıkarılmış ve yerlerine üç yeni madde eklenmiştir. Ayrıca 15. madde düzenlenmiştir.

Matematik testinin en az bir hafta sonrasında uygulanan Türkçe testinin analizleri aşağıdaki gibidir. Yapılan analiz ile sekiz farklı puan türüne ait betimsel istatistikler, madde güçlük indeksleri, testlerin ortalama güçlükleri, madde ayırt edicilik indeksleri ve test güvenilirlikleri hesaplanmıştır.

Tablo 8

Çoklu Cevaplı Madde Türü İle Oluşturulan Türkçe Testinin Puanlama Yöntemlerine Göre Betimsel İstatistikleri

Betimsel İstatistikler	Puanlama Yöntemlerine Göre Betimsel İstatistik Değerleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
A.O.	57,4	35,2	43,5	45,16	68,2	64,4	40,4	47,6
Ortanca	57,9	35	44,4	45	70	65,4	39,6	46,7
Std. Sapma	17,30	17,18	22,7	18,15	14,95	17,17	17,88	16,99
Ranj	80,1	75	90,5	86,70	80	85	86	81,7
Çarpıklık	-0,351	0,276	-0,332	-0,049	-0,75	-0,467	0,126	-0,022
Basıklık	-0,308	-0,477	-0,149	-0,338	0,713	-0,270	-0,483	-0,366

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 8 incelendiğinde puanlama yöntemlerine ait aritmetik ortalama ve ortanca değerlerinin birbirine çok yakın değerler aldıkları; çarpıklık değerlerinin altı tanesinin negatif ve sıfıra yakınken (hafif sola kayışlı) iki tanesinin pozitif ve sıfıra yakın (hafif sağa kayışlı) olduğu, basıklık değerlerinin yedi tanesinin negatif ve sıfıra yakınken (hafif basık) bir tanesinin pozitif ve sıfıra yakın (hafif sivri) değerler aldıkları görülmektedir. Ranj değerlerine bakıldığında en düşük ve en yüksek puanlar arasındaki farkın oldukça fazla olduğu dikkati çekmektedir.

Tablo 9

Çoklu Cevaplı Madde Türü İle Oluşturulan Türkçe Testine Ait Maddelerin Güçlük İndekslerinin Puanlama Yöntemlerine Göre Gösterimi

Madde No	Puanlama Yöntemlerine Göre Testteki Maddelerin Güçlük İndeksleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
1	0,79	0,61	0,74	0,74	0,85	0,86	0,65	0,66
2	0,72	0,48	0,70	0,47	0,81	0,85	0,47	0,47
3	0,82	0,68	0,75	0,75	0,88	0,83	0,72	0,78
4	0,54	0,37	0,48	0,51	0,65	0,55	0,50	0,53
5	0,48	0,41	0,23	0,46	0,61	0,53	0,43	0,45
6	0,42	0,24	0,33	0,24	0,66	0,59	0,24	0,24
7	0,57	0,19	0,39	0,32	0,69	0,63	0,24	0,46
8	0,47	0,16	0,31	0,34	0,66	0,50	0,31	0,41
9	0,47	0,47	0,25	0,52	0,61	0,58	0,51	0,53
10	0,77	0,65	0,69	0,73	0,83	0,83	0,66	0,69
11	0,62	0,39	0,50	0,51	0,74	0,67	0,47	0,54
12	0,71	0,56	0,64	0,65	0,79	0,78	0,59	0,62
13	0,40	0,14	0,17	0,24	0,56	0,44	0,22	0,38
14	0,67	0,39	0,60	0,56	0,74	0,78	0,40	0,44
15	0,55	0,26	0,41	0,43	0,67	0,66	0,31	0,37
16	0,63	0,34	0,60	0,34	0,73	0,77	0,34	0,34
17	0,40	0,06	0,06	0,24	0,42	0,42	0,16	0,40
18	0,38	0,18	0,31	0,18	0,59	0,55	0,18	0,18
19	0,43	0,11	0,19	0,31	0,50	0,43	0,25	0,43
20	0,63	0,35	0,41	0,50	0,64	0,63	0,43	0,62
Testin								
Ortalama	0,57	0,35	0,43	0,45	0,68	0,64	0,40	0,48
Güçlüğü								

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 9 incelendiğinde madde güçlüklerinin ve testlerin ortalama güçlüğüne genel olarak

0,5 değerine yakın değerler aldıkları görülmektedir. Bu da maddelerin ve testlerin ortalama zorlukta olduğu çıkarımında bulunmamıza neden olur.

Tablo 10

Çoklu Cevaplı Madde Türü İle Oluşturulan Türkçe Testinin Puanlama Yöntemlerine Göre Test Güvenirlikleri

Puanlama Yöntemleri	Güvenirlik Türü	Test Güvenirliği
Dengeli Puanlama	Cronbach Alpha	0,80
İkili Puanlama	KR-20	0,71
Morgan Algoritması	Cronbach Alpha	0,77
Negatif Puanlama	Cronbach Alpha	0,78
Çoklu Doğru/Yanlış Puanlama	Cronbach Alpha	0,83
Pozitif Sayım Puanlama	Cronbach Alpha	0,80
Tuzaklı Çoklu Cevaplı Puanlama	Cronbach Alpha	0,75
Ripkey Algoritması	Cronbach Alpha	0,74

Tablo 10 incelendiğinde sekiz puanlama türü ile elde edilen güvenilirliklerin yeterince yüksek oldukları söylenebilir.

Tablo 11

Çoklu Cevaplı Madde Türü İle Oluşturulan Türkçe Testine Ait Maddelerin Ayırt Edicilik İndekslerinin Puanlama Yöntemlerine Göre Gösterimi

Madde No	Puanlama Yöntemlerine Göre Testteki Maddelerin Ayırt Edicilik İndeksleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
1	0,31	0,22	0,28	0,27	0,37	0,33	0,26	0,26
2	0,43	0,33	0,39	0,35	0,48	0,40	0,37	0,36
3	0,43	0,45	0,45	0,41	0,37	0,36	0,43	0,46
4	0,54	0,54	0,56	0,57	0,53	0,53	0,55	0,53
5	0,45	0,41	0,45	0,44	0,45	0,47	0,41	0,39
6	0,39	0,30	0,38	0,34	0,44	0,37	0,33	0,31
7	0,23	0,14	0,17	0,22	0,32	0,26	0,20	0,21
8	0,23	0,18	0,24	0,26	0,35	0,28	0,19	0,21
9	0,30	0,35	0,30	0,38	0,33	0,33	0,33	0,29
10	0,53	0,49	0,51	0,51	0,54	0,52	0,47	0,48
11	0,55	0,48	0,55	0,53	0,50	0,51	0,50	0,53
12	0,54	0,55	0,56	0,54	0,49	0,50	0,53	0,54
13	0,38	0,28	0,39	0,42	0,48	0,38	0,42	0,39
14	0,62	0,52	0,62	0,63	0,64	0,60	0,54	0,47
15	0,60	0,53	0,61	0,63	0,64	0,61	0,57	0,51
16	0,61	0,42	0,57	0,46	0,63	0,59	0,45	0,45
17	0,52	0,23	0,28	0,39	0,48	0,48	0,28	0,48
18	0,47	0,33	0,46	0,36	0,57	0,48	0,37	0,36
19	0,54	0,36	0,51	0,54	0,57	0,54	0,51	0,53
20	0,59	0,49	0,55	0,59	0,57	0,59	0,50	0,56

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 11 incelendiğinde maddelerin çok büyük çoğunluğunun yeterince ayırt edici olduğu söylenebilir. Türkçe başarı testine ait veriler incelendiğinde 1-7 ve 8. maddelerin madde ayırt

ediciliklerinin düşük olduğu görülmüştür. Bu yüzden bu üç madde testten çıkarılmış ve yerlerine üç yeni madde eklenmiştir.

Yapılan düzenlemelerle oluşturulan yeni test formları ikinci kez uygulanmıştır. Uygulama Uşak İli ve ilçelerinde yapılmış ve uygulama grubu ilk uygulamaya katılmayan öğrencilerden oluşturulmuştur. Uygulamaya katılacak öğrencilere çoklu cevaplı madde türü tanıtımı yapıp şans başarısına karşı düzeltme uygulamasının yapılacağı ve düzeltme işleminin puanlama türüne göre farklılaşacağı belirtilip şansla tahminde bulunulmaması gerektiği belirtilmiştir. İkinci uygulama yapıldıktan sonra elde edilen verilen istatistik paket programları ile analiz edilmiştir. Matematik sınavına ait verilerinin analizi aşağıdaki gibidir. Yapılan analiz ile sekiz farklı puan türüne ait betimsel istatistikler, madde güçlük indeksleri, testlerin ortalama güçlükleri, madde ayırt edicilik indeksleri ve test güvenilirlikleri hesaplanmıştır.

Tablo 12

Çoklu Cevaplı Madde Türü İle Oluşturulan II. Uygulamaya Ait Matematik Testinin Puanlama Yöntemlerine Göre Betimsel İstatistikleri

Betimsel İstatistikler	Puanlama Yöntemlerine Göre Betimsel İstatistik Değerleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
A.O.	49,61	34,83	43,10	45,15	60,08	53,39	42,38	44,52
Ortanca	45,42	30	40	40	58,50	49,18	37,93	40
Std. Sapma	21,51	23,94	24,76	22,35	18,92	20,69	23,24	22,79
Ranj	84,15	90	96,25	86,65	79	82,50	92,50	92,50
Çarpıklık	0,906	1,256	0,741	1,058	0,547	0,673	1,105	1,076
Basıklık	0,199	0,767	0,045	0,467	-0,157	0,114	0,592	0,542

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 12 incelendiğinde puanlama yöntemlerine ait aritmetik ortalama ve ortanca değerlerinin birbirine yakın değerler aldıkları; çarpıklık değerlerinin pozitif ve sıfıra yakın (hafif sağa kayışlı), basıklık değerlerinin pozitif (ÇDY hariç) ve sıfıra yakın (hafif sivri) değerler aldıkları görülmektedir. Ranj değerlerine bakıldığında en düşük ve en yüksek puanlar arasındaki farkın oldukça fazla olduğu dikkati çekmektedir. I. ve II. Uygulama

arasındaki aritmetik ortalama artışı çıkarılan maddelerin yerine eklenen maddelerin madde güçlük indeks değerlerini artırması sebebiyle oluşmuştur.

Tablo 13

Çoklu Cevaplı Madde Türü İle Oluşturulan II. Uygulamaya Ait Matematik Testi Madde Güçlük İndekslerinin Puanlama Yöntemlerine Göre Gösterimi

Madde No	Puanlama Yöntemlerine Göre Testteki Maddelerin Güçlük İndeksleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
1	0,87	0,75	0,85	0,84	0,93	0,88	0,83	0,84
2	0,49	0,34	0,36	0,45	0,67	0,55	0,36	0,39
3	0,42	0,25	0,29	0,35	0,57	0,47	0,28	0,34
4	0,41	0,18	0,36	0,18	0,58	0,67	0,18	0,18
5	0,46	0,25	0,34	0,39	0,60	0,53	0,30	0,33
6	0,54	0,44	0,49	0,52	0,63	0,54	0,53	0,53
7	0,51	0,36	0,41	0,44	0,63	0,52	0,44	0,50
8	0,72	0,52	0,67	0,66	0,80	0,78	0,57	0,59
9	0,71	0,60	0,66	0,71	0,75	0,71	0,70	0,70
10	0,29	0,23	0,19	0,27	0,38	0,31	0,26	0,28
11	0,55	0,36	0,49	0,51	0,65	0,58	0,49	0,50
12	0,44	0,22	0,43	0,44	0,59	0,45	0,44	0,44
13	0,75	0,59	0,73	0,70	0,82	0,81	0,64	0,66
14	0,50	0,32	0,46	0,48	0,57	0,50	0,47	0,50
15	0,27	0,25	0,22	0,25	0,36	0,27	0,25	0,25
16	0,39	0,17	0,33	0,35	0,54	0,41	0,32	0,35
17	0,38	0,19	0,29	0,35	0,45	0,38	0,35	0,38
18	0,40	0,33	0,35	0,38	0,45	0,40	0,37	0,39
19	0,46	0,27	0,39	0,43	0,58	0,52	0,38	0,41
20	0,36	0,33	0,31	0,33	0,47	0,38	0,33	0,33
Testin								
Ortalama	0,50	0,35	0,43	0,45	0,60	0,53	0,42	0,45
Güçlüğü								

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 13 incelendiğinde madde güçlüklerinin ve testlerin ortalama güçlüğüne genel olarak orta güçlüğü yakın değerler aldıkları görülmektedir. Bu da yeni eklenen maddelerle birlikte testin ilk haline göre daha kolaylaştığı çıkarımında bulunmamıza neden olur.

Tablo 14

Çoklu Cevaplı Madde Türü İle Oluşturulan II. Uygulamaya Ait Matematik Testinin Puanlama Yöntemlerine Göre Test Güvenirlikleri

Puanlama Yöntemleri	Güvenirlik Türü	Test Güvenirliği
Dengeli Puanlama	Cronbach Alpha	0,86
İkili Puanlama	KR-20	0,86
Morgan Algoritması	Cronbach Alpha	0,85
Negatif Puanlama	Cronbach Alpha	0,87
Çoklu Doğru/Yanlış Puanlama	Cronbach Alpha	0,84
Pozitif Sayım Puanlama	Cronbach Alpha	0,84
Tuzaklı Çoklu Cevaplı Puanlama	Cronbach Alpha	0,87
Ripkey Algoritması	Cronbach Alpha	0,87

Tablo 14 incelendiğinde sekiz puanlama türü ile elde edilen test güvenilirlik indekslerinin oldukça yüksek oldukları söylenebilir. I. Uygulama sonrasında değiştirilen maddeler ayırt ediciliği artırdığı için testlerin güvenilirlikleri de artmıştır.

Tablo 15

Çoklu Cevaplı Madde Türü İle Oluşturulan II. Uygulamaya Ait Matematik Testine Ait Maddelerin Ayırt Edicilik İndekslerinin Puanlama Yöntemlerine Göre Gösterimi

Madde No	Puanlama Yöntemlerine Göre Testteki Maddelerin Ayırt Edicilik İndeksleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
1	0,29	0,33	0,26	0,29	0,29	0,33	0,32	0,31
2	0,59	0,70	0,56	0,66	0,66	0,53	0,68	0,64
3	0,69	0,75	0,68	0,71	0,71	0,65	0,72	0,68
4	0,58	0,76	0,54	0,73	0,73	0,47	0,74	0,74
5	0,53	0,50	0,52	0,48	0,48	0,48	0,49	0,56
6	0,46	0,45	0,47	0,48	0,48	0,44	0,49	0,47
7	0,48	0,43	0,49	0,46	0,46	0,47	0,47	0,51
8	0,35	0,33	0,37	0,35	0,35	0,31	0,37	0,36
9	0,42	0,34	0,45	0,43	0,43	0,47	0,39	0,32
10	0,62	0,68	0,65	0,67	0,67	0,58	0,68	0,68
11	0,48	0,52	0,46	0,45	0,45	0,42	0,48	0,47
12	0,61	0,69	0,59	0,61	0,61	0,58	0,61	0,61
13	0,36	0,37	0,37	0,36	0,36	0,29	0,39	0,38
14	0,53	0,42	0,46	0,48	0,48	0,53	0,41	0,50
15	0,65	0,67	0,63	0,69	0,69	0,66	0,69	0,69
16	0,67	0,68	0,66	0,69	0,69	0,61	0,72	0,72
17	0,60	0,59	0,57	0,61	0,61	0,59	0,60	0,59
18	0,50	0,41	0,50	0,52	0,52	0,49	0,53	0,51
19	0,47	0,53	0,37	0,50	0,50	0,43	0,55	0,59
20	0,58	0,55	0,55	0,58	0,58	0,57	0,56	0,56

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 15 incelendiğinde maddelerin hemen hepsinin iyi ayırt edici maddeler oldukları söylenebilir. Özellikle değiştirilen üç madde ve düzenlenen bir maddenin ayırt ediciliklerinde oldukça büyük farklar gözlenmiştir. Bazı maddelerde ilk uygulamaya göre

madde ayırt ediciliklerinde azalma görülmüştür. Bunun sebebi her iki uygulamanın farklı öğrenci gruplarında yapılması sebebiyledir.

II. Uygulama matematik testinin en az bir hafta sonrasında uygulanan II. Türkçe testinin analizleri aşağıdaki gibidir. Yapılan analiz ile sekiz farklı puan türüne ait betimsel istatistikler, madde güçlük indeksleri, testlerin ortalama güçlükleri, madde ayırt edicilik indeksleri ve test güvenilirlikleri hesaplanmıştır.

Tablo 16

Çoklu Cevaplı Madde Türü İle Oluşturulan II. Uygulamaya Ait Türkçe Testinin Puanlama Yöntemlerine Göre Betimsel İstatistikleri

Betimsel İstatistikler	Puanlama Yöntemlerine Göre Betimsel İstatistik Değerleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
A.O.	59,90	38,60	49,90	49,53	70,90	66,90	44,90	51,20
Ortanca	61,70	40	51,10	50	74	70,40	45	51,70
Std. Sapma	16,82	18	21,81	18,46	15,62	17,95	18,32	16,55
Ranj	87,60	85	94,2	85,85	86	85,90	88,40	88,40
Çarpıklık	-0,852	0,285	-0,544	-0,274	-1,475	-1,053	0,074	0,091
Basıklık	1,058	0,176	-0,050	0,002	3,193	0,929	0,059	0,631

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 16 incelendiğinde puanlama yöntemlerine ait aritmetik ortalama ve ortanca değerlerinin birbirine yakın değerler aldıkları; çarpıklık değerlerinden beş tanesinin negatif ve sıfıra yakın (hafif sola kayışlı) iken, üç tanesinin pozitif ve sıfıra yakın (hafif sağa kayışlı), basıklık değerlerinin pozitif (MA hariç) ve sıfıra yakın (hafif sivri) değerler aldıkları görülmektedir. Ranj değerlerine bakıldığında en düşük ve en yüksek puanlar arasındaki farkın oldukça fazla olduğu dikkati çekmektedir.

Tablo 17

Çoklu Cevaplı Madde Türü İle Oluşturulan II. Uygulamaya Ait Türkçe Testi Madde Güçlük İndekslerinin Puanlama Yöntemlerine Göre Gösterimi

Madde No	Puanlama Yöntemlerine Göre Testteki Maddelerin Güçlük İndeksleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
1	0,65	0,33	0,56	0,52	0,75	0,75	0,38	0,46
2	0,43	0,30	0,35	0,30	0,67	0,59	0,30	0,30
3	0,44	0,15	0,26	0,27	0,56	0,51	0,22	0,35
4	0,65	0,48	0,57	0,58	0,74	0,66	0,58	0,64
5	0,62	0,52	0,50	0,59	0,69	0,67	0,54	0,57
6	0,51	0,37	0,44	0,37	0,72	0,62	0,37	0,37
7	0,61	0,28	0,50	0,49	0,72	0,75	0,33	0,37
8	0,65	0,35	0,55	0,53	0,79	0,69	0,50	0,60
9	0,50	0,41	0,30	0,51	0,64	0,58	0,48	0,53
10	0,86	0,75	0,79	0,81	0,90	0,89	0,76	0,79
11	0,68	0,40	0,60	0,58	0,79	0,72	0,53	0,60
12	0,77	0,70	0,69	0,73	0,83	0,78	0,72	0,75
13	0,37	0,25	0,36	0,39	0,64	0,51	0,38	0,48
14	0,64	0,38	0,56	0,54	0,72	0,75	0,40	0,45
15	0,60	0,39	0,50	0,50	0,69	0,65	0,44	0,52
16	0,63	0,29	0,60	0,29	0,71	0,78	0,29	0,29
17	0,55	0,22	0,30	0,41	0,56	0,56	0,31	0,55
18	0,51	0,30	0,45	0,30	0,69	0,63	0,30	0,30
19	0,60	0,30	0,49	0,53	0,65	0,60	0,49	0,60
20	0,71	0,54	0,62	0,67	0,72	0,71	0,67	0,72
Testin								
Ortalama	0,60	0,39	0,50	0,50	0,71	0,67	0,45	0,51
Güçlüğü								

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 17 incelendiğinde madde güçlüklerinin ve testlerin ortalama güçlüğüne genel olarak orta güçlüğü yakın değerler aldıkları görülmektedir. Yeni eklenen maddelerle birlikte testin ilk hali benzer zorlukta olduğu çıkarımında bulunmamızı sağlar.

Tablo 18

Çoklu Cevaplı Madde Türü İle Oluşturulan II. Uygulamaya Ait Türkçe Testinin Puanlama Yöntemlerine Göre Test Güvenirlikleri

Puanlama Yöntemleri	Güvenirlik Türü	Test Güvenirliği
Dengeli Puanlama	Cronbach Alpha	0,76
İkili Puanlama	KR-20	0,70
Morgan Algoritması	Cronbach Alpha	0,76
Negatif Puanlama	Cronbach Alpha	0,78
Çoklu Doğru/Yanlış Puanlama	Cronbach Alpha	0,84
Pozitif Sayım Puanlama	Cronbach Alpha	0,81
Tuzaklı Çoklu Cevaplı Puanlama	Cronbach Alpha	0,74
Ripkey Algoritması	Cronbach Alpha	0,72

Tablo 18 incelendiğinde sekiz puanlama türü ile elde edilen test güvenirlilik değerlerinin bazılarının orta bazılarının ise yüksek oldukları söylenebilir.

Tablo 19

Çoklu Cevaplı Madde Türü İle Oluşturulan II. Uygulamaya Ait Türkçe Testi Maddelerinin Ayırt Edicilik İndekslerinin Puanlama Yöntemlerine Göre Gösterimi

Madde No	Puanlama Yöntemlerine Göre Testteki Maddelerin Ayırt Edicilik İndeksleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
1	0,29	0,37	0,26	0,41	0,29	0,33	0,32	0,31
2	0,59	0,25	0,56	0,26	0,66	0,53	0,68	0,64
3	0,69	0,33	0,68	0,35	0,71	0,65	0,72	0,68
4	0,58	0,28	0,54	0,38	0,73	0,47	0,74	0,74
5	0,53	0,41	0,52	0,38	0,48	0,48	0,49	0,56
6	0,46	0,43	0,47	0,41	0,48	0,44	0,49	0,47
7	0,48	0,36	0,49	0,45	0,46	0,47	0,47	0,51
8	0,35	0,33	0,37	0,47	0,35	0,31	0,37	0,36
9	0,42	0,49	0,45	0,51	0,43	0,47	0,39	0,32
10	0,62	0,49	0,65	0,57	0,67	0,58	0,68	0,68
11	0,48	0,40	0,46	0,45	0,45	0,42	0,48	0,47
12	0,61	0,51	0,59	0,59	0,61	0,58	0,61	0,61
13	0,36	0,36	0,37	0,43	0,36	0,29	0,39	0,38
14	0,53	0,43	0,46	0,59	0,48	0,53	0,41	0,50
15	0,65	0,32	0,63	0,43	0,69	0,66	0,69	0,69
16	0,67	0,53	0,66	0,50	0,69	0,61	0,72	0,72
17	0,60	0,38	0,57	0,49	0,61	0,59	0,60	0,59
18	0,50	0,39	0,50	0,41	0,52	0,49	0,53	0,51
19	0,47	0,31	0,37	0,32	0,50	0,43	0,55	0,59
20	0,58	0,31	0,55	0,45	0,58	0,57	0,56	0,56

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 19 incelendiğinde maddelerin hemen hepsinin iyi ayırt edici maddeler oldukları söylenebilir. Özellikle değiştirilen üç maddenin ayırt ediciliklerinde önceki maddelere göre ayırt ediciliklerinde artış gözlenmiştir.

İki uygulama sonrasında elde edilen veriler ışığında araştırmanın veri toplama araçları oluşturulmuştur. Oluşturulan esas uygulamaya ait test formları Uşak İli ve ilçelerinde uygulanmış ve uygulama grubu deneme uygulamalarına katılmayan öğrencilerden oluşturulmuştur. Uygulamaya katılacak öğrencilere çoklu cevaplı madde türü tanıtımı yapıp şans başarısına karşı düzeltme uygulamasının yapılacağı ve düzeltme işleminin puanlama türüne göre farklılaşacağı belirtilip şansla tahminde bulunulmaması gerektiği belirtilmiştir. Esas uygulama yapıldıktan sonra elde edilen verilerin istatistik paket programları ile analiz edilmiştir. Matematik ve Türkçe başarı testleri için sekiz farklı puan türüne ait betimsel istatistikler Tablo 20 ve Tablo 21’de verilmiştir.

Tablo 20

Çoklu Cevaplı Madde Türü İle Oluşturulan Esas Uygulamaya Ait Matematik Testinin Puanlama Yöntemlerine Göre Betimsel İstatistikleri

Betimsel İstatistikler	Puanlama Yöntemlerine Göre Betimsel İstatistik Değerleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
A.O.	51,3	33,2	42,1	44,6	63,8	55,6	41,3	45,4
Ortanca	48,4	30	39,6	63	40,9	52,5	38,4	43,4
Std. Sapma	19,25	21,45	23,40	16,97	20,58	19,49	20,85	19,21
Ranj	87,6	95	98,4	77	93,2	89,2	91,8	88,3
Çarpıklık	,393	0,753	0,366	-0,044	0,542	0,220	0,561	0,543
Basıklık	-0,574	-0,012	-0,654	-0,596	-0,484	-0,678	-0,359	-0,255

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Farklı puanlama yöntemlerden elde edilen verilerin betimsel istatistik sonuçlarına göre puanlama yöntemlerine ait aritmetik ortalama ve ortanca değerlerinin birbirine yakın değerler aldıkları; çarpıklık değerlerinden yedi tanesinin pozitif ve sıfıra yakın (hafif sağa kayışlı) iken, Negatif Puanlama yönteminin negatif ve sıfıra yakın (hafif sola kayışlı), basıklık değerlerinin negatif ve sıfıra yakın (hafif basık) değerler aldıkları görülmektedir. Ranj değerlerine bakıldığında en düşük ve en yüksek puanlar arasındaki farkın oldukça fazla olduğu dikkati çekmektedir.

Tablo 21

Çoklu Cevaplı Madde Türü İle Oluşturulan Esas Uygulamaya Ait Türkçe Testinin Puanlama Yöntemlerine Göre Betimsel İstatistikleri

Betimsel İstatistikler	Puanlama Yöntemlerine Göre Betimsel İstatistik Değerleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
A.O.	61,1	37,2	49	49,3	72,6	68,2	44,2	51,1
Ortanca	62,1	35	50	74	50	70	44,1	51,7
Std. Sapma	14,17	16,05	19,08	11,12	15,48	14,28	16,08	14,62
Ranj	70,5	85	90	64	80,9	70	85,1	85,6
Çarpıklık	-0,196	0,373	-0,156	-0,507	0,059	-0,430	0,198	0,092
Basıklık	-0,320	0,082	-0,415	0,191	-0,302	-0,233	-0,167	-0,085

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Tablo 21 incelendiğinde puanlama yöntemlerine ait aritmetik ortalama ve ortanca değerlerinin birbirine yakın değerler aldıkları; çarpıklık değerlerinden dört tanesinin negatif ve sıfıra yakın (hafif sola kayışlı) iken, diğer dört tanesinin pozitif ve sıfıra yakın (hafif sağa kayışlı), basıklık değerlerinin iki tanesinin pozitif ve sıfıra yakın (hafif sivri), beş tanesinin negatif ve sıfıra yakın (hafif basık) değerler aldıkları görülmektedir. Ranj değerlerine bakıldığında en düşük ve en yüksek puanlar arasındaki farkın oldukça fazla olduğu dikkati çekmektedir.

3.4. Verilerin Toplanması

Araştırma, 2019-2020 eğitim-öğretim yılının ikinci döneminde sekizinci sınıf Türkçe ve matematik derslerinin kazanımlarına uygun olarak çoklu cevaplı maddelerden oluşturulmuş başarı testleri kullanılarak yapılmıştır. Yapılan deneme uygulamalarına ait frekans değerleri Tablo 22'deki gibidir.

Tablo 22

Deneme Uygulamalarına Ait Frekans Değerleri

Uygulama	Testlerin Uygulandığı Öğrenci Sayısı	
	Matematik	Türkçe
I. Uygulama	158	134
II. Uygulama	88	92

Yapılan uygulamalar sonrasında son halini alan Türkçe ve matematik başarı testlerinin esas uygulaması birer hafta aralıklarla sekizinci sınıf öğrencilerine yapılmıştır. Uygulama bizzat araştırmacı tarafından yürütülmüş ve deneme uygulamalarından elde edilen tecrübe ile ilk olarak Türkçe testi uygulanmış, uygulamadan sonraki hafta ise matematik testi uygulanmıştır. Ayrıca öğrencilerin test motivasyonunu sağlamak için çoklu cevaplı madde türünün yurt dışında kullanıldığı alanlar ve sınavların da içerisinde yer aldığı tanıtıcı bir sunu hazırlanarak uygulama öncesi sunum yapılmıştır. Yapılan esas uygulamaya ait frekans değerleri Tablo 23'teki gibidir.

Tablo 23
Esas Uygulamaya Ait Frekans Değerleri

Uygulama	Testlerin Uygulandığı Öğrenci Sayısı	
	Matematik	Türkçe
Esas Uygulama	263	277

Esas uygulamanın ikinci kısmı yapılmadan önce öğrencilere ilk uygulanan testin sonuçları puan türlerine göre açıklanmıştır ve ikinci uygulamaya başlamadan önce öğrencilerin çoklu cevaplı maddeler ile çoktan seçmeli maddelerin karşılaştırılmasına ilişkin görüşlerini almak için beş sorudan oluşan anket uygulanmıştır. Anket uygulamasına 259 öğrenci gönüllü olarak katılmıştır.

3.5. Veri Analizi

Araştırmanın birinci aşamasında veriler, betimsel tarama ve anlam çıkartıcı istatistik yöntemleri kullanılarak değerlendirilmiştir. Araştırmanın ikinci aşamasındaki veriler sekizinci sınıf öğrencilerinin matematik ve Türkçe dersleri kazanımlarına uygun olarak hazırlanmış başarı testlerine verdikleri yanıtlardan oluşmaktadır. Uygulanan başarı testleri Zipgrade programıyla taranmış ve bilgisayar ortamına aktarılmıştır.

Araştırmanın birinci sorusuna ait istatistiksel işlemler için testlerin ortalama güçlükleri ve güvenilirlik indekslerinin belirlenmesi gerekmektedir. Testin ortalama güçlüğü testte bulunan maddelerin güçlük indeksleri toplamının testteki madde sayısına bölümü ile elde edilir. Farklı puanlama yöntemleriyle elde edilen test puanlarına ait ortalama güçlükler arasındaki ilişki varyans analizi ile tespit edilmiştir. Testin güvenilirliği testteki maddelerin birbiriyle tutarlı olup olmadığı gösterir. Araştırma verilerinin güvenilirliği puanlama metodunun yapısına göre KR-20 ya da Cronbach Alpha iç tutarlılık yöntemlerinden uygun olanı

seçilerek belirlenmiştir. Farklı puanlama yöntemleriyle elde edilen test puanlarının güvenilirlikleri arasındaki ilişkiyi belirlemek için önce güvenilirlik katsayıları Fisher's Z puanlarına aşağıdaki gibi dönüştürülmüş (Tan, 2016, s. 302):

$$Z_r = \frac{1}{2} \ln \frac{1+r_{xy}}{1-r_{xy}}$$

ardından elde edilen katsayılar arasındaki en büyük farklara sahip olan değerlerden başlayarak ikiyeşerli olarak manidarlık aşağıdaki eşitlikten hesaplanmıştır (Akhun, 1995):

$$Z = \frac{Z_{r1} - Z_{r2}}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

Araştırmanın ikinci sorusuna ait istatistiksel işlemler için test maddelerinin güçlük indeksleri ve maddelerin ayırt edicilik indekslerinin belirlenmesi gerekmektedir. Madde güçlük indeksi bir maddeden alınan puanların ortalamasının madde puan ranjına oranı ile elde edilir. Bu hesaplama ile hem ikili hem de çoklu puanlamada kullanılacak indeks hesaplanabilir. Madde güçlük indeksi “0” ile “1” arasında değerler alır ve değer “0” a yaklaştıkça madde zorlaşırken “1” e yaklaştıkça madde kolaylaşır. Farklı puanlama yöntemleri kullanılarak elde edilen madde güçlükleri arasındaki farkın anlamlılığı varyans analizi yöntemi kullanılarak incelenmiştir. Madde ayırt edicilik indeksi, maddenin ölçmesi beklenen özelliğe sahip oluş derecesidir. Bu sebeple madde ayırt ediciliğine madde geçerliği de denir. Madde ayırt ediciliğinin farklı puanlama türlerine göre farklı formülleri bulunmaktadır. Bunlar Pearson momentler çarpımı korelasyonu, nokta çift serili korelasyon, çift serili korelasyon, Phi ve Tetrachoric korelasyonlar ile grup farklılıklarına dayalı olan tekniktir. Grup farklarına göre de farklı yüzdeler alınmakla birlikte en sık %27' lik gruplar seçilerek uygulama yapılır. Kısmi puanlanan maddelerin madde ayırt edicilik indeksleri hesaplanırken Pearson korelasyonu ve üst-alt grup yöntemi önerilirken, ikili puanlanan maddeler için bazı yazarlar (Kilmen, 2012, s. 364; Turgut & Baykul, 2013, s.226) nokta çift serili korelasyonu bazıları (Erkuş, 2016, s. 141 ; Özçelik,2013, s. 134; Tan, 2016, s. 163) ise (ikili puanlamanın yapay süreksiz olduğu gerekçesiyle) çift serili korelasyonu önerir. Ayrıca Erkuş (2016, s. 145), üst-alt grup yönteminin 0-1 şeklinde puanlanan maddeler için kullanışlı olmadığını belirtmiştir. Tüm bu bilgilerden hareketle çoklu puanlamada ayırt ediciliği hesaplamak için Pearson korelasyonun kullanımına karar verilirken ikili puanlanan maddeler için dağılımın normal olmaması ihtimalinin bulunduğu durumlarda nokta çift serili ve çift serili korelasyon katsayıları arasında Tan(2016, s.168), nokta çift serili korelasyonun tercih edilmesinin daha

uygun olacağını ifade ettiği için nokta çift serili korelasyonun kullanılmasına karar verilmiştir. Farklı puanlama yöntemleri kullanılarak elde edilen madde ayırt edicilik indeksleri arasındaki farkın anlamlılığı belirlemek için önce ayırt edicilik puanları Fisher's Z puanlarına aşağıdaki gibi dönüştürülmüş (Tan, 2016, s. 302):

$$Z_r = \frac{1}{2} \ln \frac{1+r_{xy}}{1-r_{xy}}$$

ardından elde edilen katsayılar arasındaki en büyük farklara sahip olan değerlerden başlayarak ikiyeşerli olarak manidarlık aşağıdaki eşitlikten hesaplanmıştır (Akhun, 1995):

$$Z = \frac{Z_{r1} - Z_{r2}}{\sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}}$$

Öğrencilerin çoklu cevaplı maddeler ile çoktan seçmeli maddelerin karşılaştırılmasına ilişkin görüşleri arasında anlamlı farklılık olup olmadığı tek örneklem kay kare testi ile belirlenmeye çalışılmıştır.

Araştırmanın verilerinin analizi için Excel ve istatistik paket programlarından yararlanılmıştır.

BÖLÜM IV

BULGULAR

Bu bölümde, araştırma sürecinde toplanan verilerin analizleri sonucu elde edilen bulgular yer almaktadır.

4.1. Birinci Alt Probleme İlişkin Bulgular

Çoklu cevaplı madde türünün puanlanmasında kullanılan yöntemlerin test istatistiklerine etkilerinin incelendiği ilk alt problemde, çoklu cevaplı madde türü kullanılarak hazırlanan bir testin farklı puanlama yöntemleriyle elde edilen test puanlarına ait ortalama güçlükler arasında istatistiksel olarak anlamlı bir farklılık olup olmadığını incelemek için varyans analizi yapılmıştır. Testlerin ortalama güçlükleri Tablo 24’te gösterilirken, matematik testi varyans analizine ait sonuçlar Tablo 25’te, Türkçe testi varyans analizine ait sonuçlar Tablo 26’da verilmiştir.

Tablo 24

Farklı Puanlama Yöntemleriyle Elde Edilen Test Puanlarına Ait Ortalama Güçlük Değerleri

Puanlama Yöntemleri	Testlere Ait Ortalama Güçlük Değerleri	
	Matematik	Türkçe
Dengeli Puanlama (1)	0,51	0,61
İkili Puanlama (2)	0,33	0,37
Morgan Algoritması (3)	0,42	0,49
Çoklu Doğru/Yanlış Puanlama (4)	0,64	0,73
Negatif Puanlama (5)	0,45	0,49
Pozitif Sayım Puanlama (6)	0,56	0,68
Tuzaklı Çoklu Cevaplı Puanlama (7)	0,41	0,44
Ripkey Algoritması (8)	0,45	0,51

Sekiz farklı puanlama yöntemiyle elde edilen test puanlarına ait ortalama güçlük değerleri incelendiğinde matematik ve Türkçe testlerinin genel olarak orta güçlükte olduğu görülmektedir.

Tablo 25

Farklı Puanlama Yöntemleriyle Elde Edilen Matematik Testi Puanlarına Ait Varyans Analizi Sonuçları

Varyans Kaynağı	Kareler Toplamı	Serbestlik Derecesi	Kareler Ortalaması	F	p	Fark
Gruplar arası	6613,352	7	944,765	57,708	,000	(1-2),(1-3),(1-5),(1-7),(1-8) (3-2),(4-1),(4-2),(4-3),(4-5) (4-6),(4-7),(4-8),(5-2),(6-2)
Grup içi	34314,318	2096	16,371			(6-3),(6-5),(6-7),(6-8),(7-2) (8-2)
Toplam	40927,670	2103				

Dengeli Puanlama (1), İkili Puanlama (2), Morgan Algoritması (3), Çoklu Doğru/Yanlış Puanlama (4), Negatif Puanlama (5), Pozitif Sayım (6), Tuzaklı Çoklu Cevaplı (7), Ripkey Algoritması (8)

Analiz sonuçlarına göre farklılaşan puanlama yöntemleriyle elde edilen test ortalama güçlük değerleri arasında anlamlı farklılıklar görülmüştür. Özellikle ÇDY puanlama yöntemiyle elde edilen ortalama güçlük değerinin diğer puanlama yöntemlerinin tamamıyla elde edilen ortalama güçlük değerleriyle anlamlı olarak farklılaşmaktadır.

Tablo 26

Farklı Puanlama Yöntemleriyle Elde Edilen Türkçe Testi Puanlarına Ait Varyans Analizi Sonuçları

Varyans Kaynağı	Kareler Toplamı	Serbestlik Derecesi	Kareler Ortalaması	F	p	Fark
Gruplar arası	11451,471	7	1635,924	175,561	,000	(1-2),(1-3),(1-5),(1-7),(1-8) (3-2),(3-7),(4-1),(4-2),(4-3), (4-5),(4-6),(4-7),(4-8),(5-2) (6-1),(6-2),(6-3),(6-5),(6-7) (6-8),(7-2),(8-2),(8-7)
Grup içi	20574,754	2208	9,318			
Toplam	32026,225	2215				

Dengeli Puanlama (1), İkili Puanlama (2), Morgan Algoritması (3), Çoklu Doğru/Yanlış Puanlama (4), Negatif Puanlama (5), Pozitif Sayım (6), Tuzaklı Çoklu Cevaplı (7), Ripkey Algoritması (8)

Türkçe testine ait analiz sonuçlarına göre farklılaşan puanlama yöntemleriyle elde edilen test ortalama güçlük değerleri arasında anlamlı farklılıklar görülmüştür. Yine özellikle ÇDY ve PS puanlama yöntemleriyle elde edilen ortalama güçlük değerlerinin diğer puanlama yöntemleriyle elde edilen ortalama güçlük değerleriyle anlamlı olarak farklılaşmıştır.

4.2. İkinci Alt Probleme İlişkin Bulgular

Çoklu cevaplı madde türünün puanlanmasında kullanılan yöntemlerin test istatistiklerine etkilerinin incelendiği ikinci alt problemde, çoklu cevaplı madde türü kullanılarak hazırlanan bir testin farklı puanlama yöntemleriyle elde edilen test puanlarının güvenirlik katsayıları arasında anlamlı bir farklılık olup olmadığını incelemek için yedi yöntem ile elde edilen Cronbach Alpha değerleri ile İkili Puanlama Yöntemi ile elde edilen KR-20 değeri bulunmuş ve elde edilen değerler arasında anlamlılığa bakılmıştır. Farklı puanlama yöntemleriyle elde edilen test puanlarının güvenirlik değerleri Tablo 27’de ve analiz sonuçları Tablo 28’de gösterilmiştir.

Tablo 27

Farklı Puanlama Yöntemleriyle Elde Edilen Test Puanlarının Güvenirlik Katsayıları

Puanlama Yöntemleri	Güvenirlik Türü	Testlere Ait Güvenirlik Katsayıları	
		Matematik	Türkçe
Dengeli Puanlama	Cronbach Alpha	0,83	0,76
İkili Puanlama	KR-20	0,80	0,70
Morgan Algoritması	Cronbach Alpha	0,81	0,73
Negatif Puanlama	Cronbach Alpha	0,83	0,78
Çoklu Doğru/Yanlış Puanlama	Cronbach Alpha	0,84	0,74
Pozitif Sayım Puanlama	Cronbach Alpha	0,82	0,75
Tuzaklı Çoklu Cevaplı Puanlama	Cronbach Alpha	0,83	0,73
Ripkey Algoritması	Cronbach Alpha	0,81	0,72

Sekiz puanlama yöntemiyle elde edilen test puanlarının güvenirlik katsayıları incelendiğinde matematik testine ait güvenirlik katsayılarının oldukça yüksek olduğu görülmektedir. Türkçe testine ait güvenirlik katsayılarının ise matematik testine ait güvenirlik katsayılarının altında yer almasına rağmen orta düzeyde oldukları söylenebilir. Tablo 28’de matematik ve Türkçe testleri için en yüksek ve en düşük güvenirlik katsayılarına sahip puanlama yöntemlerinin karşılaştırılması verilmiştir.

Tablo 28

Farklı Puanlama Yöntemleriyle Elde Edilen Test Puanlarının En Yüksek ve En Düşük Güvenirlik Katsayılarına Ait Z İstatistiği Sonuçları

Test	Yöntem	r	Z _r	z
Matematik	ÇDY	0,84	1,2211	1,3974
	İP	0,80	1,0986	
Türkçe	NP	0,78	1,0445	1,6179
	İP	0,70	0,9076	

İkili Puanlama (İP), Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY)

Matematik ve Türkçe testlerine ait güvenirlik katsayılarının kendi içlerinde yapılan karşılaştırılmasında Fisher’in Z istatistiği kullanılmıştır ve test güvenirlikleri arasında

anlamalı bir farklılık bulunmamıştır.

4.3. Üçüncü Alt Probleme İlişkin Bulgular

Çoklu cevaplı madde türünün puanlanmasında kullanılan sekiz farklı yöntemin madde istatistiklerine etkilerinin incelendiği üçüncü alt problemde, çoklu cevaplı madde türünün puanlanmasında kullanılan yöntemlerden hesaplanan madde güçlük indeksleri arasında istatistiksel olarak anlamalı bir farklılık olup olmadığını incelemek için varyans analizi yapılmıştır. Testlere ait madde güçlük indeksleri Tablo 29 ve Tablo 31’de gösterilirken, varyans analizine ait sonuçlar Tablo 30 ve Tablo 32’da verilmiştir.



Tablo 29

Farklı Puanlama Yöntemlerinden Hesaplanan Matematik Testine Ait Madde Güçlük İndeksleri

Madde No	Puanlama Yöntemlerine Göre Matematik Testine Ait Maddelerin Güçlük İndeksleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
1	0,83	0,66	0,82	0,80	0,92	0,86	0,77	0,78
2	0,57	0,39	0,45	0,51	0,72	0,63	0,43	0,46
3	0,53	0,37	0,41	0,46	0,66	0,54	0,45	0,51
4	0,33	0,18	0,25	0,18	0,58	0,51	0,18	0,18
5	0,48	0,23	0,34	0,34	0,62	0,56	0,28	0,32
6	0,49	0,35	0,40	0,46	0,62	0,50	0,45	0,48
7	0,52	0,26	0,42	0,42	0,66	0,56	0,39	0,47
8	0,64	0,48	0,55	0,60	0,75	0,70	0,53	0,54
9	0,68	0,57	0,59	0,66	0,71	0,68	0,65	0,68
10	0,31	0,18	0,19	0,27	0,44	0,32	0,26	0,30
11	0,59	0,39	0,51	0,54	0,71	0,62	0,52	0,56
12	0,55	0,30	0,47	0,48	0,69	0,56	0,47	0,53
13	0,71	0,43	0,65	0,61	0,81	0,79	0,50	0,56
14	0,53	0,34	0,45	0,50	0,61	0,54	0,47	0,54
15	0,29	0,27	0,23	0,27	0,45	0,34	0,27	0,27
16	0,51	0,29	0,45	0,43	0,62	0,56	0,38	0,43
17	0,34	0,11	0,18	0,26	0,42	0,34	0,24	0,34
18	0,49	0,33	0,40	0,43	0,59	0,52	0,41	0,46
19	0,47	0,25	0,33	0,38	0,61	0,52	0,33	0,40
20	0,41	0,29	0,34	0,29	0,59	0,48	0,29	0,29

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Farklı puanlama yöntemleri ile elde edilen matematik testine ait madde güçlük indekslerine göre testin genelinde orta güçlükte maddeler yer alırken İP ve düzeltme işlemi uygulayan

çoklu puanlama yöntemleri ile elde edilen madde güçlük indekslerinin diğerlerine düşük oldukları görülmektedir.

Tablo 30.1

Farklı Puanlama Yöntemleriyle Hesaplanan Matematik Testine Ait Madde Güçlük İndeksleri İçin Varyans Analizi Sonuçları (1-10. Maddeler)

Test Madde No	Madde İçin Varyans Kaynağı	Kareler Toplamı	Sd	Kareler Ortalaması	F	p	Fark
1	Gruplar arası	10,180	7	1,454	14,954	,000	(1-2),(3-2),(4-2),(4-3)
	Grup içi	203,842	2096	,097			(4-5),(4-6),(4-7),(4-8)
	Toplam	214,022	2103				(5-2),(6-2),(6-7)
2	Gruplar arası	22,985	7	3,284	15,810	,000	(1-2),(1-7),(4-1),(4-2)
	Grup içi	435,297	2096	,208			(4-3),(4-5) (4-7),(4-8)
	Toplam	458,282	2103				(6-2),(6-3),(6-7),(6-8)
3	Gruplar arası	15,049	7	2,150	10,901	,000	(1-2),(1-3),(4-1),(4-2)
	Grup içi	413,369	2096	,197			(4-3),(4-5) (4-7),(4-8)
	Toplam	428,418	2103				(6-2),(6-3),(8-2)
4	Gruplar arası	48,777	7	6,968	39,046	,000	(1-2),(4-1),(4-2),(4-3)
	Grup içi	374,050	2096	,178			(4-5),(4-7) (4-8),(6-1)
	Toplam	422,827	2103				(6-2),(6-3),(6-5)
5	Gruplar arası	35,814	7	5,116	27,961	,000	(1-2),(1-3),(1-7),(1-8)
	Grup içi	383,523	2096	,183			(4-1),(4-2) (4-3),(4-5)
	Toplam	419,337	2103				(4-7),(4-8),(5-2),(6-2)
6	Gruplar arası	11,843	7	1,692	8,547	,000	(1-2),(4-1),(4-2),(4-3)
	Grup içi	414,899	2096	,198			(4-5),(4-6) (4-7),(4-8)
	Toplam	426,742	2103				(6-2),(8-2)
7	Gruplar arası	26,295	7	3,756	24,032	,000	(1-2),(1-3),(1-7),(4-1)
	Grup içi	327,614	2096	,156			(4-2),(4-3) (4-5),(4-7)
	Toplam	353,909	2103				(4-8),(5-2) (6-2),(6-3)
8	Gruplar arası	15,329	7	2,190	10,868	,000	(1-2),(4-2),(4-3),(4-5)
	Grup içi	422,321	2096	,201			(4-7),(4-8) (6-2),(6-3)
	Toplam	437,650	2103				(6-7),(6-8)
9	Gruplar arası	4,237	7	,605	2,985	,004	(4-2),(4-3)
	Grup içi	424,922	2096	,203			
	Toplam	429,159	2103				
10	Gruplar arası	12,281	7	1,754	11,034	,000	(1-2),(1-3),(4-1),(4-2)
	Grup içi	333,268	2096	,159			(4-3) (4-5),(4-6)(4-7)
	Toplam	345,549	2103				(4-8),(6-2),(6-3)

Dengeli Puanlama (1), İkili Puanlama (2), Morgan Algoritması (3), Çoklu Doğru/Yanlış Puanlama (4), Negatif Puanlama (5), Pozitif Sayım (6), Tuzaklı Çoklu Cevaplı (7), Ripkey Algoritması (8)

Tablo 30.2

Farklı Puanlama Yöntemleriyle Hesaplanan Matematik Testine Ait Madde Güçlük İndeksleri İçin Varyans Analizi Sonuçları (11-20. Maddeler)

Test Madde No	Madde İçin Varyans Kaynağı	Kareler Toplamı	Sd	Kareler Ortalaması	F	p	Fark
11	Gruplar arası	15,940	7	2,277	13,075	,000	(1-2),(3-2),(4-1),(4-2)
	Grup içi	365,029	2096	,174			(4-3) (4-5),(4-7)(4-8)
	Toplam	380,969	2103				(5-2)(6-2) (7-2),(8-2)
12	Gruplar arası	21,378	7	3,054	19,737	,000	(1-2),(3-2),(4-1)(4-2)
	Grup içi	324,328	2096	,155			(4-3) (4-5),(4-6),(4-7)
	Toplam	345,706	2103				(4-8)(5-2),(6-2),(7-2)
13	Gruplar arası	33,462	7	4,780	32,307	,000	(8-,2)
	Grup içi	310,130	2096	,148			(1-2),(1-7),(1-8),(3-2)
	Toplam	343,592	2103				(4-2),(4-3),(4-5),(4-7)
14	Gruplar arası	12,319	7	1,760	9,955	,000	(4-8),(5-2) (5-7),(6-2)
	Grup içi	370,543	2096	,177			(6-3),(6-5),(6-7),(6-8)
	Toplam	382,863	2103				(1-2),(4-2),(4-3),(4-5)
15	Gruplar arası	8,757	7	1,251	5,902	,000	(4-7) (5-2),(6-2),(7-2)
	Grup içi	444,299	2096	,212			(8-2)
	Toplam	453,057	2103				(4-1),(4-2),(4-3),(4-5)
16	Gruplar arası	19,891	7	2,842	16,183	,000	(4-7),(4-8)
	Grup içi	368,030	2096	,176			(1-2),(1-7),(3-2),(4-2)
	Toplam	387,921	2103				(4-3) (4-5),(4-7),(4-8)
17	Gruplar arası	19,007	7	2,715	22,269	,000	(5-2),(6-2) (6-5),(6-7)
	Grup içi	255,561	2096	,122			(6-8)
	Toplam	274,568	2103				(1-2),(1-3),(1,7),(4-2)
18	Gruplar arası	12,075	7	1,725	8,775	,000	(4-3),(4-5),(4-7),(5-2)
	Grup içi	412,033	2096	,197			(6-2)(6-3) (6-7),(7-2)
	Toplam	424,108	2103				(8-2),(8-3),(8-7)
19	Gruplar arası	25,175	7	3,596	20,535	,000	(1-2),(4-2),(4-3),(4-5)
	Grup içi	367,097	2096	,175			(4-7) (4-8),(6-2),(6-3)
	Toplam	392,273	2103				(8-2)
20	Gruplar arası	23,433	7	3,348	15,395	,000	(1-2),(1-3),(1-7),(4-1)
	Grup içi	455,767	2096	,217			(4-2),(4-3),(4-5),(4-7)
	Toplam	479,200	2103				(4-8)(5-2) (6-2),(6-3)

Dengeli Puanlama (1), İkili Puanlama (2), Morgan Algoritması (3), Çoklu Doğru/Yanlış Puanlama (4), Negatif Puanlama (5), Pozitif Sayım (6), Tuzaklı Çoklu Cevaplı (7), Ripkey Algoritması (8)

Farklı puanlama yöntemleriyle hesaplanan matematik testine ait madde güçlük indeksleri için varyans analizi sonuçları incelendiğinde bütün maddelerde bazı puanlama yöntemleri arasında anlamlı farklılık olduğu görülmüştür.

Tablo 31

Farklı Puanlama Yöntemlerinden Hesaplanan Türkçe Testine Ait Madde Güçlük İndeksleri

Madde No	Puanlama Yöntemlerine Göre Türkçe Testine Ait Maddelerin Güçlük İndeksleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
1	0,61	0,25	0,44	0,39	0,73	0,66	0,33	0,52
2	0,53	0,31	0,48	0,31	0,73	0,69	0,31	0,31
3	0,47	0,16	0,28	0,31	0,63	0,49	0,28	0,42
4	0,71	0,56	0,64	0,65	0,80	0,73	0,64	0,68
5	0,58	0,51	0,37	0,56	0,66	0,62	0,53	0,54
6	0,56	0,39	0,49	0,39	0,75	0,69	0,39	0,39
7	0,58	0,27	0,47	0,46	0,72	0,73	0,30	0,35
8	0,62	0,32	0,52	0,52	0,77	0,66	0,48	0,56
9	0,36	0,38	0,06	0,45	0,53	0,49	0,44	0,45
10	0,85	0,72	0,81	0,81	0,89	0,89	0,76	0,78
11	0,65	0,35	0,56	0,53	0,78	0,70	0,46	0,54
12	0,78	0,68	0,72	0,75	0,84	0,81	0,71	0,73
13	0,47	0,16	0,29	0,31	0,64	0,51	0,27	0,40
14	0,74	0,44	0,69	0,64	0,82	0,86	0,47	0,51
15	0,65	0,39	0,57	0,55	0,75	0,74	0,44	0,49
16	0,74	0,43	0,72	0,43	0,81	0,87	0,43	0,43
17	0,56	0,19	0,28	0,40	0,60	0,57	0,32	0,56
18	0,48	0,27	0,41	0,29	0,69	0,64	0,27	0,27
19	0,53	0,16	0,34	0,42	0,61	0,52	0,37	0,52
20	0,77	0,51	0,66	0,70	0,80	0,77	0,66	0,77

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Farklı puanlama yöntemleri ile elde edilen Türkçe testine ait madde güçlük indekslerine göre testin genelinde genel olarak orta güçlükte ve kolay maddelerden oluştuğu söylenebilir. Türkçe testinde de İP ve düzeltme işlemi uygulayan çoklu puanlama yöntemleri ile elde edilen madde güçlük indekslerinin diğerlerine düşük oldukları görülmektedir.

Tablo 32.1

Farklı Puanlama Yöntemlerinden Hesaplanan Türkçe Testine Ait Madde Güçlük İndeksleri İçin Varyans Analizi Sonuçları (1-10. Maddeler)

Test Madde No	Madde İçin Varyans Kaynağı	Kareler Toplamı	Sd	Kareler Ortalaması	F	p	Fark
1	Gruplar arası	54,034	7	7,719	59,547	,000	(1-2),(1-3),(1-5),(1-7)
	Grup içi	286,223	2208	,130			(3-2) (3-7),(4-1),(4-2)
	Toplam	340,257	2215				(4-3)(4-5) (4-7),(4-8)
2	Gruplar arası	61,983	7	8,855	41,601	,000	(5-2),(6-2)(6-3)(6-5)
	Grup içi	469,975	2208	,213			(6-7),(6-8),(8-2)
	Toplam	531,958	2215				(1-2),(1-5),(1-7),(1-8)
3	Gruplar arası	43,586	7	6,227	46,688	,000	(3-2) (3-5),(3-7),(3-8)
	Grup içi	294,470	2208	,133			(4-1)(4-2) (4-3),(4-5)
	Toplam	338,056	2215				(4-7),(4-8)(6-1) (6-2)
4	Gruplar arası	9,607	7	1,372	8,032	,000	(6-3),(6-5)(6-7),(6-8)
	Grup içi	377,285	2208	,171			(1-2),(1-3),(1-5),(1-7)
	Toplam	386,892	2215				(3-2) (4-1),(4-2),(4-3)
5	Gruplar arası	14,472	7	2,067	7,906	,000	(4-5)(4-7) (4-8),(5-2)
	Grup içi	577,406	2208	,262			(6-2),(6-3),(6-5) (6-7)
	Toplam	591,878	2215				(7-2),(8-2)
6	Gruplar arası	42,748	7	6,107	26,594	,000	(1-2),(4-2),(4-3),(4-5)
	Grup içi	507,032	2208	,230			(4-7) (4-8),(6-2)
	Toplam	549,780	2215				
7	Gruplar arası	63,466	7	9,067	57,436	,000	(1-3),(2-3),(4-2),(4-3)
	Grup içi	348,544	2208	,158			(5-3) (6-3),(7-3),(8-3)
	Toplam	412,010	2215				
8	Gruplar arası	34,314	7	4,902	35,035	,000	(1-2),(1-5),(1-7),(1-8)
	Grup içi	308,931	2208	,140			(3-2) (3-7),(4-1),(4-2),
	Toplam	343,245	2215				(4-3),(4-5) (4-7),(4-8)
9	Gruplar arası	40,961	7	5,852	19,243	,000	(6-1),(6-2),(6-3) (6-5)
	Grup içi	671,419	2208	,304			(6-7),(6-8),(8-2)
	Toplam	712,380	2215				
10	Gruplar arası	7,532	7	1,076	8,349	,000	(1-3),(2-3),(4-1),(4-2)
	Grup içi	284,578	2208	,129			(4-3),(5-3) (6-3),(7-3)
	Toplam	292,111	2215				(8-3)

Dengeli Puanlama (1), İkili Puanlama (2), Morgan Algoritması (3), Çoklu Doğru/Yanlış Puanlama (4), Negatif Puanlama (5), Pozitif Sayım (6), Tuzaklı Çoklu Cevaplı (7), Ripkey Algoritması (8)

Tablo 32.2

Farklı Puanlama Yöntemlerinden Hesaplanan Türkçe Testine Ait Madde Güçlük İndeksleri İçin Varyans Analizi Sonuçları (11-20. Maddeler)

Test Madde No	Madde İçin Varyans Kaynağı	Kareler Toplamı	Sd	Kareler Ortalaması	F	p	Fark
11	Gruplar arası	35,726	7	5,104	34,520	,000	(1-2),(1-5),(1-7),(1-8)
	Grup içi	326,448	2208	,148			(3-2) (4-1) (4-2),(4-3)
	Toplam	362,174	2215				(4-5)(4-7) (4-8),(5-2)
12	Gruplar arası	5,601	7	,800	4,890	,000	(6-2),(6-3)(6-5) (6-7)
	Grup içi	361,283	2208	,164			(6-8),(7-2) (8-2)
	Toplam	366,885	2215				(4-2),(4-3),(4-7),(4-8)
13	Gruplar arası	45,411	7	6,487	49,412	,000	(6-2)
	Grup içi	289,888	2208	,131			(1-2),(1-3),(1-5),(1-7)
	Toplam	335,299	2215				(3-2) (4-1) (4-2),(4-3)
14	Gruplar arası	48,381	7	6,912	45,217	,000	(4-5)(4-6) (4-7),(4-8)
	Grup içi	337,496	2208	,153			(5-2),(6-2)(6-3) (6-5)
	Toplam	385,877	2215				(6-7),(6-8) (7-2)(8-2)
15	Gruplar arası	34,450	7	4,921	28,177	,000	(8-3),(8-7)
	Grup içi	385,658	2208	,175			(1-2),(1-5),(1-7),(1-8)
	Toplam	420,108	2215				(3-2),(3-7) (3-8),(4-2)
16	Gruplar arası	72,040	7	10,291	56,721	,000	(4-3)(4-5) (4-7),(4-8)
	Grup içi	400,617	2208	,181			(5-2),(5-7)(5-8) (6-1)
	Toplam	472,657	2215				(6-2),(6-3) (6-5)(6-7)
17	Gruplar arası	48,664	7	6,952	52,325	,000	(6-8)
	Grup içi	293,364	2208	,133			(1-2),(1-7),(1-8),(3-2)
	Toplam	342,028	2215				(3-7) (4-2) (4-3),(4-5)
18	Gruplar arası	57,009	7	8,144	38,612	,000	(4-7)(4-8) (5-2),(6-2)
	Grup içi	465,721	2208	,211			(6-3),(6-5)(6-7) (6-8)
	Toplam	522,730	2215				(1-2),(1-5),(1-7),(1-8)
19	Gruplar arası	39,554	7	5,651	47,855	,000	(3-2) (3-5) (3-7),(3-8)
	Grup içi	260,714	2208	,118			(4-1)(4-2) (4-3),(4-5)
	Toplam	300,268	2215				(4-7),(4-8)(6-1) (6-2)
20	Gruplar arası	17,746	7	2,535	19,232	,000	(6-3),(6-5) (6-7)(6-8)
	Grup içi	291,046	2208	,132			(1-2),(1-3),(1-5),(1-7)
	Toplam	308,792	2215				(3-2) (4-2) (4-3),(4-5)

Dengeli Puanlama (1), İkili Puanlama (2), Morgan Algoritması (3), Çoklu Doğru/Yanlış Puanlama (4), Negatif Puanlama (5), Pozitif Sayım (6), Tuzaklı Çoklu Cevaplı (7), Ripkey Algoritması (8)

Farklı puanlama yöntemlerinden hesaplanan Türkçe testine ait madde güçlük indeksleri için varyans analizi sonuçları incelendiğinde testin tüm maddelerinde bazı puanlama yöntemleri arasında anlamlı farklılık olduğu görülmüştür.

4.4. Dördüncü Alt Probleme İlişkin Bulgular

Çoklu cevaplı madde türünün puanlanmasında kullanılan sekiz farklı yöntemin madde istatistiklerine etkilerinin incelendiği dördüncü alt problemde, çoklu cevaplı madde türünün puanlanmasında kullanılan yöntemlerden hesaplanan madde ayırt edicilik indeksleri arasında istatistiksel olarak anlamlı bir farklılık olup olmadığını belirlemek için önce ayırt edicilik indeksleri Fisher's Z puanlarına dönüştürülmüş ve ardından manidarlıklarına dair inceleme yapılmıştır. Testlerden elde edilen madde ayırt edicilik indeksleri Tablo 33 ve Tablo 35'te gösterilirken, ayırt edicilik indeksleri arasındaki ilişkiye ait sonuçlar Tablo 34 ve Tablo 36'da gösterilmiştir.

Tablo 33

Farklı Puanlama Yöntemlerinden Hesaplanan Matematik Testine Ait Madde Ayırt Edicilik İndeksleri

Madde No	Puanlama Yöntemlerine Göre Matematik Testine Ait Maddelerin Ayırt Edicilik İndeksleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
1	0,30	0,36	0,31	0,33	0,26	0,26	0,34	0,31
2	0,52	0,55	0,51	0,53	0,48	0,47	0,54	0,53
3	0,47	0,55	0,45	0,50	0,42	0,48	0,49	0,46
4	0,41	0,36	0,39	0,36	0,44	0,46	0,38	0,40
5	0,48	0,37	0,45	0,43	0,46	0,49	0,33	0,33
6	0,40	0,38	0,38	0,42	0,43	0,39	0,43	0,40
7	0,38	0,30	0,40	0,41	0,43	0,37	0,40	0,36
8	0,43	0,46	0,44	0,46	0,41	0,41	0,45	0,44
9	0,54	0,50	0,49	0,54	0,53	0,55	0,50	0,49
10	0,52	0,49	0,47	0,53	0,59	0,51	0,54	0,49
11	0,43	0,50	0,46	0,46	0,36	0,39	0,47	0,45
12	0,54	0,50	0,58	0,59	0,56	0,53	0,59	0,55
13	0,39	0,41	0,38	0,41	0,35	0,43	0,38	0,32
14	0,47	0,44	0,41	0,45	0,48	0,46	0,43	0,47
15	0,59	0,60	0,55	0,61	0,59	0,57	0,60	0,60
16	0,55	0,42	0,53	0,55	0,61	0,56	0,50	0,49
17	0,56	0,61	0,52	0,59	0,54	0,57	0,61	0,54
18	0,59	0,59	0,59	0,60	0,59	0,59	0,60	0,58
19	0,65	0,63	0,64	0,65	0,60	0,63	0,63	0,60
20	0,54	0,53	0,50	0,57	0,58	0,51	0,58	0,57

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Sekiz puanlama yöntemiyle hesaplanan matematik testine ait madde ayırt edicilik indeksleri sonuçlarına göre maddelerin tamamına yakınının oldukça iyi ayırt edici oldukları çıkarımı yapılabilir.

Tablo 34

Farklı Puanlama Yöntemlerinden Hesaplanan Matematik Testine Ait Madde Ayırt Edicilik İndekslerinin Karşılaştırılmasına Ait Sonuçlar

Madde No	Yöntem	r_{jx}	z_r	z
1	İP	0,36	0,376	1,2635
	ÇDY	0,26	0,266	
2	İP	0,55	0,618	1,2349
	PS	0,47	0,510	
3	İP	0,55	0,618	1,9461
	ÇDY	0,42	0,447	
4	PS	0,46	0,497	1,3730
	NP	0,36	0,376	
5	PS	0,49	0,536	2,2131*
	TÇC	0,33	0,342	
	PS	0,49	0,536	2,2131*
	RA	0,33	0,342	
	DP	0,48	0,522	2,0540*
	TÇC	0,33	0,342	
	DP	0,48	0,522	2,0540*
	RA	0,33	0,342	
6	ÇDY	0,43	0,459	0,6822
	MA	0,38	0,400	
7	RA	0,46	0,497	2,1411*
	İP	0,3	0,309	
8	NP	0,46	0,497	0,7034
	PS	0,41	0,435	
9	PS	0,55	0,618	0,9386
	MA	0,49	0,536	
10	ÇDY	0,59	0,677	1,9108
	MA	0,47	0,510	
11	İP	0,50	0,549	1,9658*
	ÇDY	0,36	0,376	
12	TÇC	0,59	0,677	1,4635
	İP	0,5	0,549	
13	PS	0,43	0,459	1,4622
	RA	0,32	0,331	
14	ÇDY	0,48	0,522	0,9962
	MA	0,41	0,435	
15	NP	0,61	0,708	1,0323
	MA	0,55	0,618	
16	ÇDY	0,61	0,708	2,9784**
	İP	0,42	0,447	
	PS	0,56	0,632	2,1109*
	İP	0,42	0,447	
	ÇDY	0,61	0,708	1,9709*

	RA	0,49	0,536	
17	TÇC	0,61	0,708	1,5116
	MA	0,52	0,576	
18	NP	0,60	0,693	0,3498
	RA	0,58	0,662	
19	DP	0,65	0,775	0,9366
	ÇDY	0,6	0,693	
20	TÇC	0,58	0,662	1,2901
	Ma	0,50	0,549	

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Matematik testine ait analiz sonuçlarına göre 5-7-11 ve 16. maddelerde farklı puanlama yöntemleriyle elde edilen madde ayırt edicilik indeksleri arasında anlamlı farklılıklar bulunmuştur.

Tablo 35

Farklı Puanlama Yöntemlerinden Hesaplanan Türkçe Testine Ait Madde Ayırt Edicilik İndeksleri

Madde No	Puanlama Yöntemlerine Göre Türkçe Testine Ait Maddelerin Ayırt Edicilik İndeksleri							
	DP	İP	MA	NP	ÇDY	PS	TÇC	RA
1	0,33	0,33	0,31	0,31	0,33	0,38	0,32	0,31
2	0,37	0,36	0,35	0,37	0,32	0,37	0,35	0,36
3	0,38	0,25	0,41	0,35	0,46	0,40	0,35	0,41
4	0,51	0,46	0,51	0,51	0,47	0,50	0,50	0,49
5	0,39	0,44	0,43	0,41	0,36	0,37	0,43	0,42
6	0,40	0,28	0,38	0,29	0,37	0,38	0,29	0,32
7	0,34	0,25	0,30	0,29	0,29	0,41	0,29	0,27
8	0,36	0,37	0,39	0,39	0,44	0,41	0,40	0,36
9	0,27	0,29	0,29	0,29	0,26	0,21	0,30	0,31
10	0,33	0,33	0,35	0,34	0,35	0,32	0,34	0,33
11	0,34	0,44	0,33	0,39	0,34	0,27	0,42	0,40
12	0,50	0,46	0,46	0,50	0,50	0,52	0,45	0,45
13	0,36	0,30	0,38	0,35	0,37	0,34	0,34	0,31
14	0,45	0,32	0,40	0,40	0,52	0,51	0,32	0,30
15	0,45	0,38	0,39	0,45	0,52	0,45	0,42	0,43
16	0,39	0,39	0,34	0,41	0,40	0,40	0,41	0,41
17	0,47	0,40	0,44	0,45	0,54	0,50	0,44	0,42
18	0,44	0,29	0,45	0,36	0,39	0,45	0,36	0,36
19	0,43	0,36	0,39	0,44	0,50	0,43	0,41	0,42
20	0,44	0,42	0,39	0,40	0,42	0,46	0,35	0,37

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Sekiz puanlama yöntemiyle hesaplanan Türkçe testine ait madde ayırt edicilik indeksleri sonuçlarına göre genel olarak maddelerin iki tanesi haricinde diğerlerinin oldukça iyi ayırt edici oldukları çıkarımı yapılabilir.

Tablo 36

Farklı Puanlama Yöntemlerinden Hesaplanan Türkçe Testine Ait Madde Ayırt Edicilik İndekslerinin Karşılaştırılmasına Ait Sonuçlar

Madde No	Yöntem	r_{jx}	z_r	Z
1	PS	0,38	0,400	0,9340
	RA	0,31	0,320	
2	NP	0,37	0,388	0,6669
	ÇDY	0,32	0,331	
3	ÇDY	0,32	0,331	2,8416**
	İP	0,36	0,376	
	MA	0,35	0,365	2,6935**
	İP	0,36	0,376	
	RA	0,41	0,435	2,1168*
	İP	0,36	0,376	
	PS	0,4	0,423	1,9763*
	İP	0,36	0,376	
4	DP	0,51	0,562	0,3132
	RA	0,49	0,536	
5	İP	0,44	0,472	1,1200
	ÇDY	0,36	0,376	
6	DP	0,4	0,423	1,5972
	İP	0,28	0,287	
7	PS	0,41	0,435	2,1168*
	İP	0,25	0,255	
8	ÇDY	0,44	0,472	1,1200
	RA	0,36	0,376	
9	RA	0,31	0,320	1,2613
	PS	0,21	0,213	
10	ÇDY	0,35	0,365	0,3970
	PS	0,32	0,331	
11	İP	0,44	0,472	2,2950
	PS	0,27	0,276	
	TÇC	0,42	0,447	2,0067
	PS	0,27	0,276	
12	PS	0,52	0,576	1,0765
	RA	0,45	0,484	
13	MA	0,38	0,400	1,0636
	İP	0,30	0,309	
14	ÇDY	0,52	0,576	3,1344**
	RA	0,30	0,309	
	ÇDY	0,52	0,576	2,8744**
	TÇC	0,32	0,331	
	ÇDY	0,52	0,576	2,8744**
	İP	0,32	0,331	
	PS	0,51	0,562	2,9745**
	RA	0,30	0,309	

	PS	0,51	0,562	2,7146**
	TÇC	0,32	0,331	
	PS	0,51	0,562	2,7146**
	İP	0,32	0,331	
15	ÇDY	0,52	0,576	2,0708*
	İP	0,38	0,400	
16	RA	0,41	0,435	
	MA	0,34	0,354	0,9576
17	ÇDY	0,54	0,604	2,1204*
	TÇC	0,40	0,423	
	ÇDY	0,54	0,604	2,1204*
	İP	0,40	0,423	
18	MA	0,45	0,484	2,1865*
	İP	0,29	0,298	
	PS	0,45	0,484	2,1865*
	İP	0,29	0,298	
	DP	0,44	0,472	2,0400*
	İP	0,29	0,298	
19	NP	0,44	0,472	1,1200
	İP	0,36	0,376	
20	ÇDY	0,52	0,576	2,4774**
	TÇC	0,35	0,365	
	ÇDY	0,52	0,576	2,2075*
	RA	0,37	0,388	

Dengeli Puanlama (DP), İkili Puanlama (İP), Morgan Algoritması (MA) Negatif Puanlama (NP), Çoklu Doğru/Yanlış Puanlama (ÇDY), Pozitif Sayım (PS), Tuzaklı Çoklu Cevaplı (TÇC), Ripkey Algoritması (RA)

Türkçe testine ait analiz sonuçlarına göre 3-7-14-15-17-18 ve 20. maddelerde farklı puanlama yöntemleriyle elde edilen madde ayırt edicilik indeksleri arasında anlamlı farklılıklar bulunmuştur.

4.5. Beşinci Alt Probleme İlişkin Bulgular

Çoklu cevaplı maddelerden oluşan başarı testlerine katılan öğrencilerin çoklu cevaplı maddeler ile çoktan seçmeli maddelerin karşılaştırılmasına ilişkin görüşlerini almak için Tablo 37’de yer alan beş ifadeden oluşan anket öğrencilere uygulanmıştır. Çoktan seçmeli maddelere aşina olan ve çoklu cevaplı maddelerden oluşan başarı testlerini cevaplayan öğrencilerin verilen ifadelerle göre hangi madde türünün daha uygun olduğunu seçmeleri istenmiş ve 259 öğrenciden elde edilen sonuçların frekans ve yüzde dağılımları verilmiştir.

Tablo 37

Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anket Maddeleri ve Bu Maddelere Verilen Yanıtlara Göre Oluşan Frekans ve Yüzde Dağılımı

İfadeler	Çoklu Cevaplı		Çoktan Seçmeli	
	f	%	f	%
1) Daha yüksek puan alabileceğinizi düşündüğünüz madde türü	235	90,73	24	9,27
2) Bildiklerinizi daha iyi yansıtabileceğiniz madde türü	119	45,95	140	54,05
3) Cevaplaması daha kolay olan madde türü	13	5,02	246	94,98
4) Kişiler arası ayrımı daha iyi yapabilen madde türü	194	74,90	65	25,10
5) LGS’de kullanılmasını isteyeceğiniz madde türü	56	21,62	203	78,38

Anketten elde edilen verilerle Tek Örneklem Kay Kare Testi yapılmıştır. Tablo 38 ve Tablo 39’da 1. maddeye ait Kay Kare Testi sonuçları verilmiştir.

Tablo 38

Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 1. Maddesine Ait Gözlenen ve Beklenen Dağılım Tablosu

	Gözlenen N	Beklenen N	Kalan
Çoklu Cevaplı	235	129,5	105,5
Çoktan Seçmeli	24	129,5	-105,5
Toplam	259		

Tablo 39

Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 1. Maddesine Ait Kay Kare Testinin Sonucu

	Madde Türü
Chi-Square	171,896
df	1
Asymp. Sig	,000

Çoklu cevaplı maddelerden oluşan başarı testlerine katılan öğrencilerin çoklu cevaplı maddeler ile çoktan seçmeli maddelerin karşılaştırılmasına ilişkin görüşlerini almak için uygulanan ankette yer alan "Daha yüksek puan alabileceğinizi düşündüğünüz madde türü " maddesine öğrencilerin 235 tanesi çoklu cevaplı madde türünü 24 tanesi çoktan seçmeli madde türünü seçmiştir. Öğrencilerin tercihleri arasında anlamlı bir farkın olup olmadığını belirlemek için yapılan Tek Örneklem Kay Kare Testi sonucuna göre öğrencilerin tercihleri arasında anlamlı bir farklılık bulunmuştur [$\chi^2_{(1)} = 171,896$, $p < 0,005$]. Bu fark anlamlılık sınırları dahilinde oldukça yüksektir [Etki Büyüklüğü= 0,664].

Anket uygulamasının 2. maddesine ait Kay Kare Testi sonuçları Tablo 40 ve Tablo 41’de verilmiştir.

Tablo 40

Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 2. Maddesine Ait Gözlenen ve Beklenen Dağılım Tablosu

	Gözlenen N	Beklenen N	Kalan
Çoklu Cevaplı	119	129,5	-10,5
Çoktan Seçmeli	140	129,5	10,5
Toplam	259		

Tablo 41

Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 2. Maddesine Ait Kay Kare Testinin Sonucu

	Madde Türü
Chi-Square	1,703
df	1
Asymp. Sig	,192

Çoklu cevaplı maddelerden oluşan başarı testlerine katılan öğrencilerin çoklu cevaplı maddeler ile çoktan seçmeli maddelerin karşılaştırılmasına ilişkin görüşlerini almak için uygulanan ankette yer alan "Bildiklerinizi daha iyi yansıtabileceğiniz madde türü "

maddesine öğrencilerin 119 tanesi çoklu cevaplı madde türünü 140 tanesi çoktan seçmeli madde türünü seçmiştir. Öğrencilerin tercihleri arasında anlamlı bir farkın olup olmadığını belirlemek için yapılan Tek Örneklem Kay Kare Testi sonucuna göre öğrencilerin tercihleri arasında anlamlı bir farklılık bulunmamıştır [$\chi^2_{(1)} = 1,703$, $p > 0,005$].

Anket uygulamasının 3. maddesine ait Kay Kare Testi sonuçları Tablo 42 ve Tablo 43'te verilmiştir.

Tablo 42

Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 3. Maddesine Ait Gözlenen ve Beklenen Dağılım Tablosu

	Gözlenen N	Beklenen N	Kalan
Çoklu Cevaplı	13	129,5	-116,5
Çoktan Seçmeli	246	129,5	116,5
Toplam	259		

Tablo 43

Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 3. Maddesine Ait Kay Kare Testinin Sonucu

	Madde Türü
Chi-Square	209,610
df	1
Asymp. Sig	,000

Çoklu cevaplı maddelerden oluşan başarı testlerine katılan öğrencilerin çoklu cevaplı maddeler ile çoktan seçmeli maddelerin karşılaştırılmasına ilişkin görüşlerini almak için uygulanan ankette yer alan "Cevaplaması daha kolay olan madde türü " maddesine öğrencilerin 13 tanesi çoklu cevaplı madde türünü 246 tanesi çoktan seçmeli madde türünü seçmiştir. Öğrencilerin tercihleri arasında anlamlı bir farkın olup olmadığını belirlemek için yapılan Tek Örneklem Kay Kare Testi sonucuna göre öğrencilerin tercihleri arasında anlamlı

bir farklılık bulunmuştur [$\chi^2_{(1)} = 209,610$, $p < 0,005$]. Bu fark anlamlılık sınırları dahilinde oldukça yüksektir [Etki Büyüklüğü= 0,809].

Anket uygulamasının 4. maddesine ait Kay Kare Testi sonuçları Tablo 44 ve Tablo 45'te verilmiştir.

Tablo 44

Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 4. Maddesine Ait Gözlenen ve Beklenen Dağılım Tablosu

	Gözlenen N	Beklenen N	Kalan
Çoklu Cevaplı	194	129,5	64,5
Çoktan Seçmeli	65	129,5	-64,5
Toplam	259		

Tablo 45

Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 4. Maddesine Ait Kay Kare Testinin Sonucu

	Madde Türü
Chi-Square	64,251
df	1
Asymp. Sig	,000

Çoklu cevaplı maddelerden oluşan başarı testlerine katılan öğrencilerin çoklu cevaplı maddeler ile çoktan seçmeli maddelerin karşılaştırılmasına ilişkin görüşlerini almak için uygulanan ankette yer alan "Kişiler arası ayrımı daha iyi yapabilen madde türü " maddesine öğrencilerin 194 tanesi çoklu cevaplı madde türünü 65 tanesi çoktan seçmeli madde türünü seçmiştir. Öğrencilerin tercihleri arasında anlamlı bir farkın olup olmadığını belirlemek için yapılan Tek Örneklem Kay Kare Testi sonucuna göre öğrencilerin tercihleri arasında anlamlı bir farklılık bulunmuştur [$\chi^2_{(1)} = 64,251$, $p < 0,005$]. Bu fark anlamlılık sınırları dahilinde yüksektir [Etki Büyüklüğü= 0,248].

Anket uygulamasının 5. maddesine ait Kay Kare Testi sonuçları Tablo 46 ve Tablo 47'de verilmiştir.

Tablo 46

Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 5. Maddesine Ait Gözlenen ve Beklenen Dağılım Tablosu

	Gözlenen N	Beklenen N	Kalan
Çoklu Cevaplı	56	129,5	-73,5
Çoktan Seçmeli	203	129,5	73,5
Toplam	259		

Tablo 47

Çoklu Cevaplı Maddelerden Oluşan Başarı Testlerine Katılan Öğrencilere Yöneltilen Anketin 5. Maddesine Ait Kay Kare Testinin Sonucu

	Madde Türü
Chi-Square	83,432
df	1
Asymp. Sig	,000

Çoklu cevaplı maddelerden oluşan başarı testlerine katılan öğrencilerin çoklu cevaplı maddeler ile çoktan seçmeli maddelerin karşılaştırılmasına ilişkin görüşlerini almak için uygulanan ankette yer alan "LGS'de kullanılmasını isteyeceğiniz madde türü" maddesine öğrencilerin 56 tanesi çoklu cevaplı madde türünü 203 tanesi çoktan seçmeli madde türünü seçmiştir. Öğrencilerin tercihleri arasında anlamlı bir farkın olup olmadığını belirlemek için yapılan Tek Örneklem Kay Kare Testi sonucuna göre öğrencilerin tercihleri arasında anlamlı bir farklılık bulunmuştur [$\chi^2_{(1)} = 83,432$, $p < 0,005$]. Bu fark anlamlılık sınırları dahilinde yüksektir [Etki Büyüklüğü= 0,322].

BÖLÜM V

SONUÇ, TARTIŞMA ve ÖNERİLER

Bu bölümde, araştırmanın bulgularına dayalı olarak bulunan sonuçlar, bulguların önemi ve etkileri üzerine tartışma ve çalışmanın sonuçlara dayalı gelecekte yapılabilecek çalışmalar için önerilere yer verilmiştir.

5.1. Sonuçlar ve Tartışma

Alt problemlere ait sonuç ve tartışmalar aşağıda sıralanmıştır.

5.1.1 Birinci Alt Probleme İlişkin Sonuçlar ve Tartışma

Farklı puanlama yöntemleriyle elde edilen testlerin ortalama güçlük değerlerine ait sonuçlar incelendiğinde her iki başarı testinde de Çoklu Doğru/Yanlış ve Pozitif Sayım Puanlama Yöntemleri ile elde edilen ortalamaların diğerlerinden anlamlı şekilde yüksek olduğu görülmektedir. Bu yüzden öğrenme eksikliklerinin tespit edilmesinin amaçlandığı, ders kazanımlarına ne derece ulaşıldığının belirlenmesi için yapılan sınıf içi uygulamalarda bu puanlama yöntemlerine yer verilmesinin daha uygun olacağı düşünülmektedir. Tuzaklı Çoklu Cevaplı, İkili Puanlama ve Morgan Algoritması Puanlama Yöntemleri ile elde edilen ortalama güçlük indekslerinin diğerlerinden oldukça düşük olduğu görülmektedir. Bu yüzden seçme amaçlı ya da maksimum puana dayalı değerlendirmenin yapıldığı geniş ölçekli sınavlarda yukarıdaki puanlama yöntemlerinden uygun olanı kullanılabilir. Muckle, Becker ve Wu (2011), araştırmalarında çoklu cevaplı maddelerin altı farklı puanlama yöntemini kıyaslamışlardır. Bu yöntemler; Tuzaklı İki Puanlı, İkili, Negatif Puanlama,

Tuzaklı Çoklu Cevaplı, Pozitif Sayım ve Çoklu Doğru/Yanlış Puanlama Yöntemi şeklindedir. Standardize hale getirilmeden karşılaştırılan altı yöntemin analizi Rasch modeli ile yapılmıştır. Araştırmanın sonuçlarına göre ortalama güçlük değerleri ve bilgi grafikleri farklı puan yöntemlerinin farklı puan aralıklarında bilgi sağladığını göstermiştir. Araştırmaları neticesinde, üst kısımda ifade edilen sonuçlarla aynı doğrultuda bir çıkarımla giriş seviyesi bilgi gerektiren işler için Çoklu Doğru/Yanlış Puanlama Yöntemi'nin kullanılabileceği ve en iyilerin arandığı durumlarda "Tuzaklı" yöntemleri kullanılabileceği belirtilmiştir. Domnich vd. (2015)'nin çoklu cevaplı maddelerin puanlama yöntemlerini karşılaştırdıkları çalışmalarında da İkili Puanlama Yöntemi'nin diğer yöntemlerine göre göreceli olarak daha zor olduğu, Çoklu Doğru/Yanlış Puanlama yönteminin ise diğerlerine göre madde gücünün daha yüksek olduğu sonuçlarına ulaşılmıştır. Ripkey vd. (1996), çalışmalarında çoktan seçmeli maddeleri, eşleştirmeli maddeleri, çoklu cevaplı maddeleri kullanarak oluşturulan testlerden elde edilen puanları karşılaştırmak için çoktan seçmeli maddeler için İkili, eşleştirmeli maddeler için puanlama ve çoklu cevaplı puanlama yöntemlerinden İkili ve Ripkey Algoritması yöntemlerini karşılaştırmışlardır. Elde ettikleri sonuçlara göre madde gücü bakımından çoklu cevaplı İkili Puanlama Yöntemi diğer yöntemlere göre anlamlı şekilde düşük çıktığı görülmüştür. Bauer, Holzen, Kopp ve Fischer (2011)'in tıp fakültesi öğrencilerinin dönem sonu sınavlarında kullanılan doğru cevap sayısının bilindiği çoklu cevaplı maddeleri karşılaştırdıkları çalışmalarında İkili Puanlama Yöntemi ile iki farklı kısmi puanlama yönteminin (kısmi (%50) puanlama ve kısmi (1/n) puanlaması) karşılaştırılması yapmışlar ve araştırmada kullandıkları diğer iki yönteme kıyasla İkili Puanlama Yöntemi'nin madde gücünün daha düşük olduğu sonucuna ulaşmışlardır.

5.1.2. İkinci Alt Probleme İlişkin Sonuçlar ve Tartışma

Çoklu cevaplı maddelerin puanlanmasında kullanılan yöntemler ile elde edilen test güvenilirlik katsayıları arasında anlamlı bir farklılık olmaması bu madde türüne ait puanlama yöntemleri arasında güvenilirlik anlamında iyi veya kötü bir yöntemin olmadığını göstermektedir. Ayrıca çoklu puanlama yöntemleriyle elde edilen güvenilirlik katsayılarının İkili Puanlama Yöntemi ile elde edilen güvenilirlik katsayılarına göre daha yüksek oldukları görülmektedir. Bauer vd. (2011) araştırmalarında kullandıkları kısmi puanlamaların (kısmi (%50) puanlama ve kısmi (1/n) puanlaması) psikometrik özelliklerinin İkili Puanlama

Yöntemi'nden daha üstün olduklarını belirtmişlerdir. Çalışmalarına ait test güvenirlik ve madde güçlük indekslerinde, bu araştırmada olduğu gibi kısmi puanlamalar İkili puanlamaya göre daha yüksektir. Hsu vd. (1984) Tayvan Üniversiteye Giriş Ortak Sınavı'nda yer alan tarih, Çince ve San Min Chu-I (Doktor Sun Yat Sen'in politik teori dersi) alt testlerini kullanarak çoklu cevaplı maddelerin puanlama yöntemleri ve çoktan seçmeli maddeleri karşılaştırmışlardır. Elde ettikleri sonuçlara göre tarih ve San Min Chu-I testlerinde puanlama yöntemlerine ait güvenirlikler arasında anlamlı bir farklılık bulunmamıştır. Sadece Çince dersinde puanlama yöntemlerine ait güvenirliklerde İkili puanlama ile S1 olarak isimlendirilen puanlama yönteminin güvenirlik değerinin diğerlerinden anlamlı olarak düşük olduğu tespit edilmiştir. Domnich vd. (2015)'nin yaptıkları çalışmada Ripkey ve Dengeli Puanlama Yöntemleri'nin diğerlerinden daha yüksek güvenirlik değerlerine sahip oldukları sonucuna ulaşılmıştır. Verbic (2012), içerisinde çoktan seçmeli ve çoklu cevaplı maddelerin bulunduğu bir sınav uygulamış ve çoklu cevaplı maddelerin analizini Küme (clustering) Puanlama Yöntemi ve Madde (item) Puanlama Yöntemleri adını verdiği yöntemlerle yapmıştır. Küme Puanlama, İkili Puanlama Yöntemi ile aynı kurallara sahiptir. Madde puanlama ise Pozitif Puanlama Yöntemi ile aynı kurallara sahiptir. Araştırma sonuçlarına göre puanlama yöntemlerinin güvenirlik değerleri arasında anlamlı bir farklılık bulunmamıştır.

5.1.3. Üçüncü Alt Probleme İlişkin Sonuçlar ve Tartışma

Çoklu cevaplı madde türünün puanlanmasında kullanılan sekiz puanlama yöntemi ile elde edilen madde güçlük indeksleri incelendiğinde hem Türkçe hem de matematik testlerinde en düşük değerlerin İkili Puanlama Yöntemi ile daha sonra ise ceza puanı uygulayan Morgan, Ripkey ve Tuzaklı Çoklu Cevaplı puanlama yöntemleri ile elde edildiği görülmüştür. Madde güçlüğü en yüksek olan puanlama yöntemleri düzeltme işleminin uygulanmadığı Pozitif Sayım ile işaretlenmeyen yanlış seçenekler için de puan alınan Çoklu Doğru/Yanlış Yöntemleridir. Testlerin ortalama güçlük değerlerinde olduğu gibi bütün test maddelerinde yüksek madde güçlüğüne sahip bu iki madde türü öğrenmenin izleneceği, etkili dönüt verilmesi amaçlanan, biçimlendirme ve yetiştirmeye yönelik formatif değerlendirme yapılacak testlerde kullanılabilir. Düzeltme formüllerinin uygulandığı çoklu puanlanan Morgan, Ripkey ve Tuzaklı Çoklu Cevaplı Puanlama Yöntemleri ise daha çok düzey belirleme ve değer biçmeye yönelik summatif değerlendirme yapılacak testlerde

kullanılabilir.

5.1.4. Dördüncü Alt Probleme İlişkin Sonuçlar ve Tartışma

Standart hale getirilmiş sekiz puanlama yöntemi ile elde edilen madde ayırt edicilik indeksleri incelendiğinde matematik testinde dört, Türkçe testinde yedi maddenin ayırt edicilik indekslerinde anlamlı farklılık bulunmuştur. Matematik testindeki dört madde incelendiğinde iki maddede Pozitif Sayım Yöntemi'nin anlamlı olarak farklı olduğu görülürken, Türkçe testindeki yedi maddenin beş tanesinde Çoklu Doğru/Yanlış Puanlama Yöntemi'nin anlamlı olarak farklı olduğu görülmektedir. Ayrıca İkili Puanlama Yöntemi'nin özellikle Türkçe testinde anlamlı olarak düşük olduğu görülmektedir. Puanlama yöntemlerinin madde madde en yüksek ayırt edicilik indekslerine bakıldığında Türkçe testi için Çoklu Doğru/Yanlış Puanlama Yöntemi ile Pozitif Sayım Puanlama Yöntemi en fazla ayırt edicilik indeksleri yüksek madde sayılarına sahiplerken matematik testi için Tuzaklı Çoklu Cevaplı, Çoklu Doğru/Yanlış ve İkili Puanlama Yöntemleri en fazla ayırt edicilik indeksleri yüksek madde sayılarına sahiptir. Verbic (2012)'in, araştırma sonucuna göre küme puanlamamanın madde puanlama yönteminden daha sağlam olduğu ortaya çıkmış ve çoklu puanlamaların İkili puanlamalara göre ayırt edicilik bağlamında daha iyi sonuçlara sahip oldukları görülmüştür. Hsu vd. (1984), yaptıkları çalışmada çoklu puanlama yöntemleri ikili puanlama yöntemine göre daha ayırıcı ve güvenilir bulmuşlardır. Türkçe ve matematik testlerindeki maddeleri puanlama yöntemlerine göre elde edilen madde güçlük indeksi sonuçları ile kolay, orta güçlükte ve zor olarak sınıflandırılıp maddelerin ayırt edicilikleri ile karşılaştırıldığında ortaya tutarlı bir sonuç çıkmaktadır. Bu çıkarıma göre kolay maddeler için İkili Puanlama Yöntemi ile elde edilen ayırt edicilik indekslerinin diğer puanlama yöntemleri ile elde edilenlerden biraz yüksek olduğu görülmüştür. Orta güçlükte maddeler için Çoklu Doğru/Yanlış Puanlama Yöntemi ile elde edilen ayırt edicilik indekslerinin diğer puanlama yöntemleri ile elde edilenlerden biraz yüksek olduğu dikkati çekmiştir. Madde güçlük indeksleri düşük olup zor olarak isimlendirilen maddeler için Tuzaklı Çoklu Cevaplı Puanlama Yöntemi ile elde edilen ayırt edicilik indekslerinin diğer puanlama yöntemleri ile elde edilenlerden biraz yüksek olduğu belirlenmiştir. Muckle vd. (2011) çalışmalarında “Tuzaklı” puanlama yönteminin farklı yetenek seviyelerinde kişiler arasında en yüksek ayrımı sağlayabileceğini ifade etmişlerdir. Domnich vd. (2015)'nin

araştırma sonuçlarına göre çoklu cevaplı madde türünün puanlama yöntemlerinin ayırt edicilikleri arasında anlamlı bir farklılık bulunmamaktadır.

5.1.5. Beşinci Alt Probleme İlişkin Sonuçlar ve Tartışma

Öğrencilere uygulanan anketlerden elde edilen sonuçlara uygulanan Kay Kare Testi sonuçlarına göre beş anket maddesinin birinci, üçüncü, dördüncü ve beşinci maddelerinde öğrencilerin tercihlerinde anlamlı farklılık bulunurken ikinci maddede öğrenci tercihleri arasında anlamlı bir farklılık bulunmamıştır. Anket uygulaması öncesinde öğrencilerin ilk test uygulamasından elde ettikleri puanları öğrenmeleri sonucunda aldıkları kısmi puanlar ile çoklu puanlama yöntemleri kullanıldığında ikili puanlamaya göre daha iyi sonuçlar alacaklarını fark etmişlerdir. Bu sebeple çoklu cevaplı maddelerde çoktan seçmeli maddelere göre daha yüksek puan alacağını düşünen öğrenciler tüm öğrencilerin %90'ıdır. Öğrencilerin bildiklerini daha iyi yansıtabildikleri madde türünde Kay Kare Testi sonucuna göre anlamlı bir farklılık bulunmamıştır. Bu durum ilk defa karşılaştıkları çoklu cevaplı madde türünün okul hayatları boyunca karşılarına çıkan çoktan seçmeli madde türü kadar öğrencilerde bildiklerini yansıtabileceklerini ortaya çıkarmıştır. Öğrencilere göre çoklu cevaplı ve çoktan seçmeli maddeler arasında cevaplama kolaylığı bakımından çoktan seçmeliler %95'lik bariz bir üstünlüğe sahiptir. Alanyazında da ifade edildiği gibi çoklu cevaplı maddelerin daha üst düzey düşünmeyi sağlayabilecek maddeler olduğu düşünülmektedir. Bu sebeple çoktan seçmeli maddelere göre daha zor olarak algılanmaları alanyazını destekleyen bir sonuçtur. Öğrencilere göre kişiler arası ayrımı sağlamada daha etkili madde türü yaklaşık %75'lik oranla çoklu cevaplı maddelerdir. İlk uygulamalar sonrasında elde edilen puanlar ve kıyaslamalarla çoklu cevaplı maddelerin kısmi puanlama imkanı sağlamasının öğrencilerin birbiriyle arasındaki farkın daha kolay ortaya çıkmasını sağlaması öğrencilerin bu seçimi yapmasında oldukça etkili olduğu düşünülmektedir. Öğrenciler LGS'de kullanılmasını isteyecekleri madde türünün %78 oranla çoktan seçmeli madde olmasını istemişlerdir. Çoktan seçmeli maddelerle yapılacak bir sınava hazırlanan öğrencilerin bu yönde düşünceleri ve seçim yapmaları beklenen bir sonucu ortaya koymuştur. Çoklu cevaplı maddelerin kullanılmasını isteyen yaklaşık %22'lik öğrenci grubunun madde türüne adapte oldukları ve uygulanması halinde şu anki durumlarından daha üst seviyeye çıkacakları ya da sınava katılacak diğer öğrenciler ile aralarındaki farkın bu madde türü ile açılacağı fikrinde oldukları tahmin edilmektedir.

5.1.6. Genel Sonuçlar ve Tartışma

Çoklu cevaplı madde türü ile oluşturulan testlerden elde edilen verilere göre bu madde türünün hem sayısal hem sözel dersler için uygulanabilir olduğu test ve madde istatistikleri ile testlerin güvenilirlik değerlerinden anlaşılmaktadır. Bu madde türünün uygulanabilirliğine ve puanlama yöntemlerine dair olumsuz bakışa sahip Turgut ve Baykul (2013, s. 191) çoklu cevaplı maddeleri yazmanın kolay olacağını ancak puanlama ve madde analizinde önemli güçlükler yaşanacağını belirtmişler ve zorunluluk olmadıkça bu madde türünün kullanılmaması gerektiğini tavsiye etmişlerdir. Benzer şekilde N. Doğan (2009, s.230), bu madde türünün puanlama ve madde istatistiklerini hesaplama zorluğundan bahsedip madde türünü başka madde türüne çevirmenin daha uygun olacağını ifade etmiştir. Yapılan bir araştırmada Karakaya ve Ö. Doğan (2020) katılımcıları ÖDM (Ölçme Değerlendirme Merkezi) çalışanı, ölçme ve değerlendirme alanında lisansüstü eğitim gören kişiler ve eğitimde ölçme değerlendirme bölümünde görev yapan akademisyenler olan kişilerle görüşmeler yapmışlardır. Araştırmaya göre katılımcıların çoklu cevaplı maddelerin puanlama yöntemlerine dair kafa karışıklığı bulunduğu fark edilmiş ve bu sebeple katılımcıların puanlamaya dair zıt yönde fikir beyan ettikleri durumlar ortaya çıkmıştır. Özellikle puanlama yöntemlerinin çeşitliliğinin güvenilirliği olumsuz etkileyeceğine dair katılımcı ifadeleri dikkat çekmiştir. Ancak bu tez çalışması sonucunda çoklu cevaplı madde türünün puanlama yöntemlerinin standart hale getirilebileceği ve madde analizinin yapılabileceği görülmüştür. Ayrıca standart hale getirilen puanlama yöntemlerinin tek bir kurala bağlı kalmadan, yapılan ölçme ve değerlendirmenin amacına göre alternatifler sunup farklı puanlama yöntemlerinin kullanılmasına olanak tanıyacak yapıda olduğu bulunmuştur. Bu sayede sınıf içi değerlendirmelerde öğrenme eksikliklerinin fark edilmesinin ve etkili dönüt vermenin amaçlandığı ölçme değerlendirme faaliyetlerinde Çoklu Doğru/Yanlış ve Pozitif Sayım Puanlama Yöntemleri önerilebilirken, seçme ve kişiler arası ayrımın sağlanması gereken geniş ölçekli uygulamalarda Tuzaklı Çoklu Cevaplı, Morgan Algoritması ya da İkili Puanlama Yöntemleri kullanılabileceği görülmüştür. Puanlama yöntemlerinin güvenilirlikleri arasında anlamlı bir farklılık olmaması ve puanlama yöntemlerinin yeterince güvenilir olması kişilerin bu madde türünün güvenilirliğine ait çekincelerini giderecektir. Öğrencilere uygulanan anketin sonuçlarına göre öğrencilerin büyük çoğunluğunun çoklu cevaplı madde türünün kişiler arası ayrımı sağlamada çoktan seçmeli maddelere göre daha iyi olduğu belirttikleri anlaşılmıştır. Ayrıca öğrenciler çoklu

cevaplı madde türü ile çoktan seçmeli maddelere oranla daha yüksek puan alacaklarını düşünmektedirler. Yıllardır kendilerine uygulanan madde türünün farklı bir formatı olan çoklu cevaplı maddelere karşı öğrencilerin tutumlarının genel olarak olumlu olduğu söylenebilir. Çoklu cevaplı madde türü ve puanlama yöntemlerine yönelik yapılan bu ve buna benzer çalışmalar ile madde türünün yaygınlaşacağı ve hem kağıt kalem testlerinde hem de ülke dışında geniş ölçekli sınavlarda kullanıldığı gibi bilgisayar tabanlı ölçme ve değerlendirme uygulamalarında kullanımın yaygınlaşacağı düşünülmektedir.

5.2. Öneriler

1. Çoklu cevaplı maddeler kullanılarak oluşturulan testlerin ölçme ve değerlendirme hedeflerine göre uygun puanlama yönteminin seçilmesinin test geliştiricilere fayda sağlayacağı düşünülmektedir. Bu sebeple Türkçe ve matematik dersleri haricinde çoklu cevaplı maddelerden oluşan testler yardımıyla çoklu cevaplı maddelerin puanlanmasında kullanılan yöntemlerin test ve madde istatistikleri karşılaştırılabilir.
2. Çoklu cevaplı maddeler hakkında öğrencilerden alınan görüşlerin genel olarak olumlu olduğu anlaşılmaktadır. Bunun yanında çoklu cevaplı madde türünün hazırlanması ve puanlanması hakkında gereken bilgilerin verildiği öğretmenlerin de madde türü hakkında düşünceleri incelenebilir.
3. Çoklu cevaplı maddeler kullanılarak oluşturulan testlerde doğru cevap sayısının bilinmediği türde uygulama yapılmıştır. Test yönergesinde doğru cevabın sayısının verildiği ve verilmediği durumların test sonuçlarına etkisi incelenebilir.
4. Çoklu cevaplı maddeler kullanılarak oluşturulan başarı testleri sınava hazırlanan sekizinci sınıf öğrencilerine uygulandı. Farklı sınıf seviyelerine göre hazırlanan çoklu cevaplı maddelerden oluşan testler yardımıyla çoklu cevaplı maddelerin puanlanmasında kullanılan yöntemlerin test ve madde istatistikleri karşılaştırılabilir.
5. Çoklu cevaplı maddeler kullanılarak oluşturulan ölçme araçlarının geniş ölçekli uygulamalarda kullanılıp test ve madde istatistiklerinin incelenmesi bu sınavı yapan kurumlara önerilebilir. Bunun yanında sınıf içi ölçme değerlendirme faaliyetlerinde kullanımının olumlu ve olumsuz tarafları da araştırmacılar tarafından incelenebilir.
6. Çoklu cevaplı maddeler kullanılarak oluşturulan ölçme araçları eğitim kurumlarında uygulanmıştır. Bu ölçme araçları yardımıyla tıp ve mühendislik gibi farklı uygulama

alanlarında yapılan ölçmelerin geçerlik ve güvenirliği incelenebilir.

7. Çoklu cevaplı madde türü kullanılarak yapılan sınıf içi değerlendirmelerde öğrenme eksikliklerinin fark edilmesinin ve etkili dönüt vermenin amaçlandığı ölçme değerlendirme faaliyetlerinde Çoklu Doğru/Yanlış ve Pozitif Sayım Puanlama Yöntemleri önerilebilirken, seçme ve kişiler arası ayrımın sağlanması gereken geniş ölçekli uygulamalarda Tuzaklı Çoklu Cevaplı, Morgan Algoritması ya da İkili Puanlama Yöntemleri kullanılabilir.



KAYNAKLAR

- Akhun, İ. (1995). *İstatistiklerin manidarlığı ve örneklem*. Ankara: Özel Basım.
- Albanese, M. A. (1993). Type K and other complex multiple choice items: An analysis of research and item properties. *Educational Measurement: Issues and Practice* 12(1), 28-33.
- Auer, S. & Tarasowa, D. (2013). Balanced scoring method for multiple-mark questions. https://www.researchgate.net/publication/288682442_Balanced_scoring_method_for_multiple-mark_questions sayfasından erişilmiştir.
- Baghaei, P. & Amrahi, N. (2011). The effects of the number of options on the psychometric characteristics of multiple choice items. *Psychological Test and Assessment Modeling*, 53(2), 192-211.
- Başol, G. (2012). *Eğitimde ölçme ve değerlendirme*. Ankara: Atalay.
- Bauer, D., Holzer, M., Kopp, V. & Fischer, M. (2011). Pick-N multiple choice-exams: a comparison of scoring algorithms. *Advances in Health Sciences Education*, 16(2), 211-221.
- Baykul, Y. (2011) Türklerde eğitimde ölçme ve değerlendirme. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 2(0).
- Baykul, Y. (2015). *Eğitimde ve psikolojide ölçme: klasik test teorisi ve uygulaması*. Ankara: Pegem Akademi.
- Ben-Simon, A., Budescu, D. V. & Nevo, B. (1997). A comparative study of measures of partial knowledge in multiple-choice tests. *Applied Psychological Measurement* 21(1), 65-88.

- Cheng, H. (2004). A comparison of multiple choice and open ended response formats for the assessment of listening proficiency in English. *Foreign Language Annals*, 37(4).
- Cronbach, L. J. (1941). An experimental comparison of the multiple true-false and multiple-choice test. *Journal of Educational Psychology*, 32, 533-543.
- Çelen, Ü. (2002). *Şans başarısı için düzeltme formülü kullanılacağına ilişkin yönergenin testin psikometrik özelliklerine etkisinin araştırılması*. Yüksek Lisans Tezi, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Çetin, (2016). Bilişsel alan davranışlarının ölçülmesi. M. Gömleksiz, S. Erkan (Ed.). *Eğitimde ölçme ve değerlendirme*. (s. 102-142). Ankara: Nobel.
- Çoban, E. (2015). *Test yönergesi ve risk alma eğiliminin seçmeli testlerin psikometrik özelliklerine ve tahminle yanıtlama davranışına etkisi*. Yüksek Lisans Tezi, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Doğan, N. (2009). Çoktan seçmeli testler. H. Atılğan, (Ed.), *Eğitimde ölçme ve değerlendirme* (223-268). Ankara: Anı.
- Doğan, S. (2012). *Kavram haritası ve yapılandırılmış grid tekniğinin çoktan seçmeli testlerle karşılaştırılması*. Yüksek Lisans Tezi, Mersin Üniversitesi, Eğitim Bilimleri Enstitüsü, Mersin.
- Doğanay, A. (2007). Üst düzey düşünme becerilerinin öğretimi. A. Doğanay (Ed.), *Öğretim ilke ve yöntemleri* (s.280-327). Ankara: Pegem Akademi.
- Domnich, A., Panatto, D., Arata, L. Bevilacqua, I., Apprato, L., Gasparini, R. & Amicizia, D. (2015). Impact of different scoring algorithms applied to multiple-mark survey items on outcome assessment: An in-field study on health-related knowledge. https://www.researchgate.net/publication/291827497_Impact_of_different_scoring_algorithms_applied_to_multiple-mark_survey_items_on_outcome_assessment_An_in-field_study_on_health-related_knowledge sayfasından erişilmiştir.
- Dressel, P. L., & Schmid, J. (1953). Some modifications of multiple-choice items. *Educational and Psychological Measurement*, 13, 574-595.
- Duman, E. Z. (2018). Bir düşünme türü olarak esnek düşünme. *Turkish Studies Social Sciences*. 13(26). 547-561.

- Duncan, G. T. & Milton, E. O. (1978). Multiple answer multiple choice test items: Responding and scoring through bayes and minimax strategies. *Psychometrica*, 43(1).
- Eggen, T. & Lampe, T. (2011). Comparison of the reliabitiy of scoring methods of multiple response items, matching items and sequencing items. https://www.researchgate.net/publication/281912731_Comparison_of_the_reliabilit_r_of_scoring_methods_of_multiple-response_items_matching_items_and_sequencing_items_sayfasından erişilmiştir.
- Emmerich, W. (1991). *The development, investigation, and evaluation of new item types for the GRE analytical measure*. Princeton, New Jersey: Educational Testing Service.
- Eren, B. (2015). *Çoktan seçmeli ve karma test uygulamalarına ilişkin öğrenci başarıları ile öğrenci ve öğretmen görüşlerinin karşılaştırılması*. Yüksek Lisans Tezi. Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Erkuş, A. (2016). *Psikolojide ölçme ve ölçek geliştirme-I*. Ankara: Pegem Akademi.
- Ertürk, S. (1972). *Eğitimde program geliştirme*. Ankara: Meteksan.
- Frisbie, D. A., & Druva, C. A. (1986). Estimating the reliability of multiple true-false tests. *Journal of Educational Measurement*, 23(2), 99–105.
- Frisbie, D. A. & Sweeney, D. C. (1982). The relative merits of multiple true-false achievement tests. *Journal of Educational Measurement* 19(1).
- Gültekin, S. (2011). *Çoktan seçmeli, açık uçlu ve karma testlerin psikometrik özelliklerinin madde tepki kuramına dayalı olarak değerlendirilmesi*. Doktora Tezi. Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Haladyna, T. M. (2004). *Developing and validating multiple choice test items*. New Jersey: Lawrence Erlbaum Associates.
- Haladyna, T. M. & Rodriguez, M. C. (2013). *Developing and validating test items*. New York: Routledge.
- Hochstetter, Page, Chang & Cappaert (2018). *Validity inferences for different types of technology-enhanced items*. National Council On Measurement In Education 2018 Annual Meeting New York.

- Hsu TC, Moss PA, Khampalikit C. (1984). The merits of multiple-answer items as evaluated by using six scoring formulas. *The Journal of Experimental Education*. 52(3), 152-158.
- Jiao, H., Liu, J., Haynie, K., Woo, A. & Gorham, J. (2012). Comparison between dichotomous and polytomous scoring of innovative items in a large scale computerized adaptive testing. *Educational and Psychological Measurement* 72(3) 493–509.
- Jodoin, M. G. (2003). Measurement efficiency of innovative item formats in computer-based testing. *Journal of Measurement* 40(1) 1-15.
- Kan, A. & Kayapınar, U. (2006). Yabancı dil eğitiminde aynı davranışları yoklayan çoktan seçmeli ve kısa cevaplı iki testin madde ve test özelliklerinin karşılaştırılması. *Eğitim ve Bilim*,31(142), 65-71.
- Kan, A. (2009). Ölçme araçlarında bulunması gereken nitelikler. H. Atılğan, (Ed.), *Eğitimde ölçme ve değerlendirme* (23-80). Ankara: Anı.
- Karakaya, İ. & Doğan, N. (2008, Mayıs). İlköğretim öğretmenlerinin ölçme araçlarını tercih etmede aradıkları özellikler ve kullanım sıklığına ilişkin görüşleri. *I. Ulusal Eğitimde ve Psikolojide Ölçme Ve Değerlendirme Kongresi bildirileri 1*, 241-252.
- Karakaya, İ. & Doğan, Ö. (2020. Mart). Çoklu cevaplı madde türünün psikometrik özellikleri hakkında ödm çalışanı, alan uzmanı ve akademisyenlerin görüşlerinin incelenmesi. *Ankara II. Uluslararası Bilimsel Araştırmalar Kongresi bildirileri 1*, 271-279.
- Karasar, N. (2016). *Bilimsel araştırma yöntemi: kavramlar, ilkeler, teknikler*. Ankara: Nobel.
- Kilmen, S. (2012). Madde analizi, madde seçimi ve yorumlanması. N. Çıkrıkçı-Demirtaşlı, (Ed.), *Eğitimde ölçme ve değerlendirme* (363-385). Ankara: Elhan.
- Koç, N. (2009), *Eğitimimizde psikolojik testlerin (standart testlerin) kullanım durumu, sorunlar ve öneriler*. I. Ulusal Eğitimde Ve Psikolojide Ölçme Ve Değerlendirme Kongresi bildirileri 1, 305-316.

- Koçak, D. (2013). *Farklı yönergelerle verilen çoktan seçmeli testlerde yanıtlama davranışlarının incelenmesi*. Yüksek Lisans Tezi, Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Kulm, G. (2013). Back To The Future; Reclaiming Effective Mathematics Assessment Strategies. *Middle Grades Research Journal*, 8(2), 1–10.
- Kutlu, Ö. (2016). <http://www.kurumsalegitimdergisi.org/olcme-ve-degerlendirme/> sayfasından erişilmiştir.
- Levine, M. V. & Drasgow, F. (1983). The relation between incorrect option choice and estimated ability. *Educational and Psychological Measurement*. 43, 675-685.
- Lord, F. M., & Novick, M.R. (1968). *Statistical theories of mental test scores*. New Jersey: Information Age
- Ma, X. (2004). *An investigation of alternative approaches to scoring multiple response items on a certification exam*. Doktora Tezi, University of Massachusetts Amherst. Boston.
- Matsuda, N., Ogawa, H., Hirashima, T. & Taki, H. (2015). A generating technique and knowledge representation of multiple-answer problems for learning with solving knowledge. *Research and Practice in Technology*, 10(6), 1-21.
- MEB Talim Terbiye Kurulu Başkanlığı (2018). Matematik dersi öğretim programı (İlkokul ve Ortaokul 1,2,3,4,5,6,7 ve 8. Sınıflar). <http://mufredat.meb.gov.tr/Dosyalar/201813017165445-MATEMATİK%20ÖĞRETİM%20PROGRAMI%202018v.pdf> sayfasından ulaşılmıştır.
- Morgan, M. (1979). MCQ: An interactive computer program for multiple-choice self-testing. *Biochemical Education*, 7(3), 67–69.
- Muckle, T., Becker, K. A. & Wu, B. (2011). *Investigating the multiple answer multiple choice item format. Presentation and scoring consideration*. 2011 annual meeting of the National Council on Measurement in Education, New Orleans, LA.
- O'Neill, K., & Folk, V. (1996). Innovative CBT item formats in a teacher licensing program. Paper presented at the annual meeting of the National Council on Measurement in Education, New York, NY.

- ÖSYM, (2015). <http://www.osym.gov.tr/TR,721/yazili-sinav-acik-uclu-sorularla-sinav-04022015.html> sayfasından erişilmiştir.
- Özbek, Ö. Y. (2010). Ölçme araçlarında bulunması gereken nitelikler. S. Tekindal (Ed.), *Eğitimde ölçme ve değerlendirme* (s.43-89). Ankara: Pegem Akademi.
- Özçelik, D. A. (1998). *Ölçme ve değerlendirme*. Ankara: ÖSYM.
- Özçelik, D. A. (2013). *Test hazırlama kılavuzu*. Ankara: Pegem Akademi.
- Özdemir, D. (2003). Çoktan seçmeli testleri puanlama yöntemlerine bir bakış. apbs.mersin.edu.tr/files/devrimozdemir/Publications_026.pdf sayfasından erişilmiştir.
- Özdemir, D. (2003). Çoktan seçmeli testlerde iki kategorili ve önsel ağırlıklı puanlamanın diferansiyel madde fonksiyonuna etkisi ile ilgili bir araştırma. *Eğitim ve Bilim*, 28(129), 37-43.
- Page, G., Bordage, G., & Allen, T. (1995). Developing key-feature problems and examinations to assess clinical decision-making skills. *Academic Medicine*. 70(3), 194-201.
- Parshall, C. G. (2015). The multiple response – why, when, and how to use this item type. <http://www.proftesting.com/blog/2015/06/17/2015617the-multiple-response-why-when-and-how-to-use-this-item-type/> sayfasından erişilmiştir .
- Parshall, C. G., Davey, T., & Pashley, P. (2000). Innovative item types for computerized testing. In W. J. van der Linden & C. A. W. Glas (Eds), *Computerized adaptive testing: Theory and practice* (pp. 129-148). Boston, MA: Kluwer Academic Publishers.
- Pehlivan, E. B. (2011). *Çoktan seçmeli türkçe test maddelerinde yanıtlama davranışlarının incelenmesi*. Yüksek Lisans Tezi. Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Peterson, A. , Craig M. & Danny, P. (2016). Employing multiple answer multiple choice questions. <https://dl.acm.org/citation.cfm?id=2925503> sayfasından erişilmiştir.
- Pomplun, M., & Omar, MD. H. (1997). Multiple-mark items: An alternative objective item format? *Educational and Psychological Measurement*, 57(6), 949-962.

- Ripkey, D. , Case, S. & Swanson, D. (1996). A “new” item format for assessing aspects of clinical competence. *Academic Medicine*, 71(10), 34–36.
- Scalise, K., & Gifford, B. (2006). Computer-Based Assessment in E-Learning: A Framework for Constructing "Intermediate Constraint" Questions and Tasks for Technology Platforms. *The Journal of Technology, Learning and Assessment*, 4(6).
- Siddiqui, N. I., Bhavsar, V. H., Bhavsar, A. V., & Bose, S. (2016). Contemplation on marking scheme for Type X multiple choice questions, and an illustration of a practically applicable scheme. *Indian journal of pharmacology*, 48(2), 114-21.
- Sireci, S. G., & Zenisky, A. L . (2006). Innovative item formats in computer-based testing: In pursuit of improved construct representation. *Handbook of test development*, 329–347.
- Tan, Ş. (2010). *Öğretimde ölçme ve değerlendirme*. Ankara: Pegem Akademi.
- Tan, Ş. (2016). *Spss ve excel uygulamalı temel istatistik-1*. Ankara: Pegem Akademi.
- Tekin, H. (2014). *Eğitimde ölçme ve değerlendirme*. Ankara: Yargı.
- Tsai, F. & Suen, H. (1993). A brief report on a comparison of six scoring methods for multiple true-false items. *Educational and psychological measurement*, 53(2), 399–404.
- Turgut, M. F. (1971). *Şans başarısının test puanlarına etkisi*. Ankara: Orta Doğu Teknik Üniversitesi.
- Turgut, M. F. (1988). *Eğitimde ölçme ve değerlendirme metotları*. Ankara: Saydam.
- Turgut, M. F. & Baykul, Y. (2013). *Eğitimde ölçme ve değerlendirme*. Ankara: Pegem Akademi.
- Umay, A. (1997). Yanıtlayıcı davranışlarının analizi yolu ile matematikte problem çözümleri için bir güvenirlik ve geçerlik araştırması. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 13, 47-56.
- Verbic, S. (2012). Information value of multiple response questions. *Psihologija*, 45(4), 467–485.
- Wan, L. & Henly, G. A. (2012). Measurement properties of two innovative item formats in

a computer-based test. *Applied Measurement In Education*, 25, 58-78.

Waters, B. (1977). The measurement of partial knowledge: A comporison between two empirical option-weighting methods and rights-only scoring. *The Journal of Educational Research*, 70(6), 256-260.

Qian, H., Woo, A. & Kim, D. (2018). Exploring the psychometric properties of innovative items in computerized adaptive testing. Jiao & Lissitz (Ed.), *Tecnology enhanced innovative assessment: development, modelling and scoring from an interdisciplinary perspective*. (s.65-116). New York: Information Age.

Yavuz, İ. & Bilgeç, İ. (2016). Açık uçlu sorularla yapılan matematik sınavlarının ölçme ve değerlendirilmesinin incelenmesi. *Eğitim ve Öğretim Araştırmaları Dergisi*, 5(3).

Yıldırım, C. (1983). *Eğitimde ölçme ve değerlendirme (öğretmenler için el kitabı)*, Ankara: ÖSYM.

EKLER



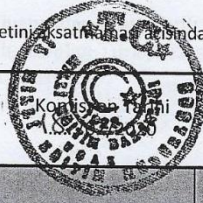
EK 1. Araştırma İzni Belgesi

EK-2

ANKET VE ARAŞTIRMA İZİN KOMİSYONU ARAŞTIRMA ÖN İNCELEME FORMU

Adı Soyadı	: Ömer DOĞAN
Kurumu/Üniversitesi	: Gazi Üniversitesi
İletişim Bilgisi	:
Konu	: Çoklu Cevaplı Madde Türünün Puanlamasında Kullanılan Yöntemlerin Test ve Madde İstatistikleri Açısından Karşılaştırılması
Başvuru Tarihi	: 14.02.2020
Veri Toplama Araçları	: var

MEB 07/03/2012 tarih ve 3616 sayılı 2012/13 Nolu Genelge Kapsamında Araştırma, Yarışma ve Sosyal Etkinlik İznilerinde Dikkat Edilecek Hususlar	Uygun	Uygun Değil	Açıklama
Anayasa, Millî Eğitim Temel Kanunu ve Türk Millî Eğitiminin Genel Amaçlarına uygunluğu açısından;	✓		
Millî ve manevi değerler açısından;	✓		
Kişilik hakları açısından (kişisel bilgiler istenilmemeli, ad-soyad vb.);	✓		
Cinsiyet, din, dil ve ırk gibi farklılıkları istismar etmeme açısından;	✓		
İnsan Hakları Evrensel Beyannamesi ve uluslararası bağlayıcılığı olan diğer belgelerce suç kabul edilen hususları içermeme açısından;	✓		
Kişisel ve ailevi mahremiyetini ifşa eden sorular, ifadeler, resimler ve simgeler yer almaması açısından;	✓		
Veri toplama araçlarında kişi, kurum ve kuruluşlara yönelik reklâm veya tanıtım gibi ifade ve öğeler yer almaması açısından;	✓		
Araştırma önerisi ile veri toplama araçlarının tamamının idareye sunulması açısından;	✓		
Uygulama, okul ve kurumların eğitim-öğretim faaliyetini kısıtlaması açısından;	✓		



Komisyon Üyeleri		Uygun	Uygun Değil	İmza
Başkan	Ahmet Turan YÜZÜK	✓		
Üye	Zeki ÇAKIR	✓		
Üye	Şirin YÖRÜK	✓		
Üye	Serap DEĞİRMENCİ ARIKAN	✓		

**EK 2. Çoklu Cevaplı Madde Türü ile Hazırlanan Matematik Testi
Maddeleri**



**ÇOKLU CEVAPLI MADDE TÜRÜ İLE HAZIRLANAN
8. SINIF MATEMATİK SINAVI SORULARI**

AÇIKLAMALAR

- Bu sınav 20 sorudan oluşmaktadır ve her sorunun beş seçeneği vardır.
- Bu sınavda kullanılan sorular çoklu cevaplı madde türüne göre oluşturulmuştur. Çoklu cevaplı madde türünde her sorunun **en az bir doğru cevabı** vardır. Yani sorudaki doğru cevap sayısına bağlı olarak öğrenci iki veya daha fazla seçenek işaretleyebilir.
- Doğru olan tüm seçenekleri işaretleyen kişi **tam puan** alırken, bazı doğru seçenekleri işaretleyen kişi **kısmi puan** alır. Tüm seçenekleri işaretleyen ya da sadece yanlış seçenekleri işaretleyen kişi puanlama yöntemine göre **sıfır puan alır ya da eksi puanlar alabilir**.
- Yanlış işaretlenen seçenekler sorudan kazanılacak puanı azaltacaktır. (Sınavda düzeltme işlemi uygulanacaktır) Bu sebeple **sadece doğruluğuna emin olduğunuz seçenekleri işaretleyiniz**.
- İşaretlemelerinizi kitapçığın arkasındaki cevap formuna doğru ve eksiksiz bir şekilde kodlayınız.
- Bu sınav için verilen süre 60 dakikadır.



1) Asal çarpanlarına ayrıldığında her bir asal çarpanının kuvveti 1 olan pozitif tam sayıya “karesiz sayı” denir.

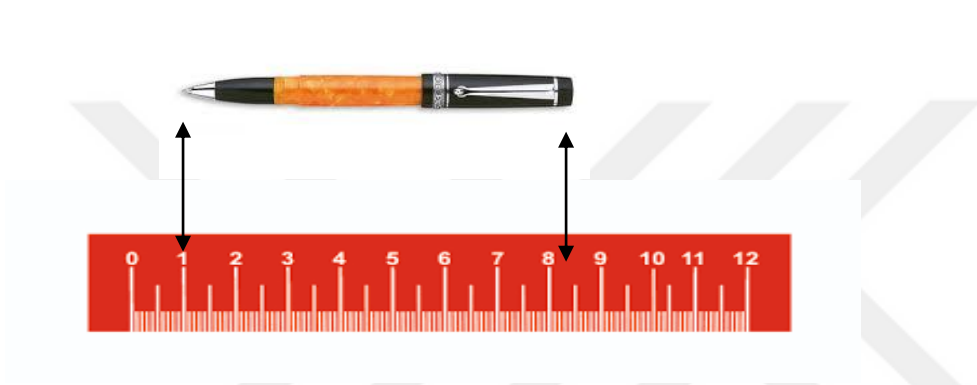
Örneğin 70 sayısı asal çarpanlarına ayrıldığında $70 = 2^1 \cdot 5^1 \cdot 7^1$ olacağından 70 karesiz bir sayıdır.

Buna göre aşağıdaki sayılardan hangisi ya da hangileri karesiz sayıdır?

- A)90 B)98 C)104 D) 110 E)111

2) a ve b birer doğal sayı olmak üzere $a\sqrt{b} = \sqrt{a^2 b}$ dir.

Ali kaleminin boyunu kırmızı cetveli ile aşağıdaki gibi ölçüyor.



Buna göre kalemin boyu santimetre cinsinden aşağıdakilerden hangisi ya da hangileri olabilir?

- A) $4\sqrt{3}$ B) $7\sqrt{2}$ C) $2\sqrt{13}$ D) $3\sqrt{6}$ E) $6\sqrt{2}$

3) $|a|$, 1 veya 1’den büyük, 10’dan küçük bir gerçekte sayı ve n bir tam sayı olmak üzere $a \cdot 10^n$ gösterimi bilimsel gösterimdir.

Ay’ da atmosfer bulunmadığından Ay’a ayak basan astronotların ayak izlerinin ay üzerinde 80 ile 100 milyon yıl kadar değişmeden kalacağı tahmin edilmektedir.



Verilen bilgiye göre, Ay’a ayak basan ilk insan olduğu iddia edilen Neil Armstrong’a ait ayak izinin ay üzerinde tahmini kaç gün kalacağının bilimsel gösterimi aşağıdakilerden hangisi ya da hangileri olabilir?

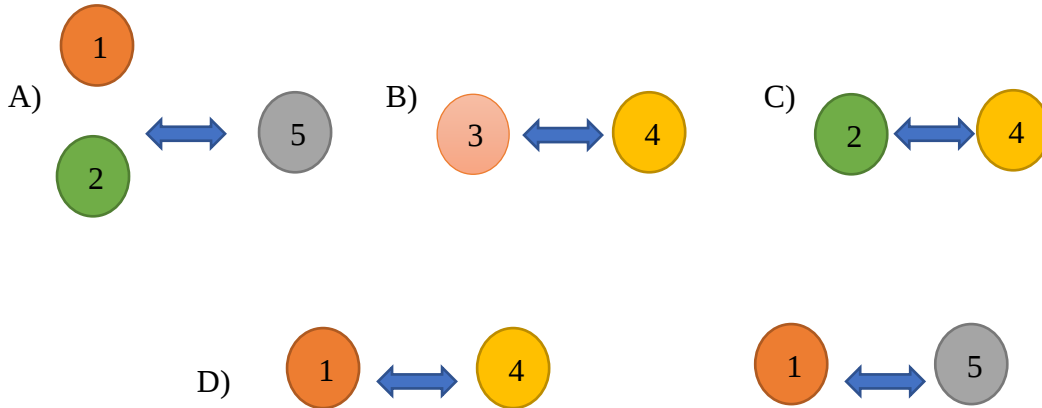
- A) $2,92 \cdot 10^{10}$ B) $3 \cdot 10^8$ C) $3,2 \cdot 10^{10}$ D) $4,2 \cdot 10^{10}$ E) $5,6 \cdot 10^9$

4)



1	$\sqrt{50}$
2	$\sqrt{18}$
3	$6\sqrt{2}$
4	$2\sqrt{8}$
5	$\sqrt{162}$

Yukarıdaki tahterevallideki karşılıklı yükler arasındaki fark 2,5 kilogram altında ise tahterevalli dengede durmaktadır. Şekilde denge durumundaki tahterevallide aşağıda verilen değişikliklerden hangisi ya da hangileri uygulanırsa tahterevalli dengede kalmaya devam eder? ($\sqrt{2}$ 'nin yaklaşık değeri 1,4 'tür.)



5) Tamamı dolu olmayan 58 kişi kapasiteli bir otobüsten 8 evli çift indiğinde otobüsteki kadın yolcu sayısı erkek yolcu sayısının 6 katı oluyor. Buna göre başlangıçta otobüste kaç yolcu olabilir?

- A) 22 B) 30 C) 38 D) 44 E) 58

6)



Renkli Çark yarışmasına katılan Ali, Betül ve Osman’a yarışma sunucusu kuralları şöyle anlatıyor.

“Merhaba değerli yarışmacılarımız, Renkli Çark yarışmamızın kuralları şunlardır.

- Renkli çarka ok atan yarışmacı okun geldiği rengin harf sayısına göre puan kazanacaktır.
Harf sayısı 5 ve üzerinde olan renkler için rengin harf sayısı 2’nin kuvveti olarak yazılır.
(Örneğin pembe rengi için 2^5 puan)
Harf sayısı 5’ in altında olan renkler için rengin harf sayısı 5’in kuvveti olarak yazılır.
(Örneğin gri rengi için 5^3 puan)
- Her yarışmacı iki kez ok atacaktır ve her isabette kazandığı puanlar çarpılacaktır.”

Bu yarışmada Ali 1000 puandan daha düşük bir puan kazanmıştır.

Betül oku ilk attığında ok beyaz renge denk gelmiş ve yarışma sonunda 20.000 puan kazanmıştır.

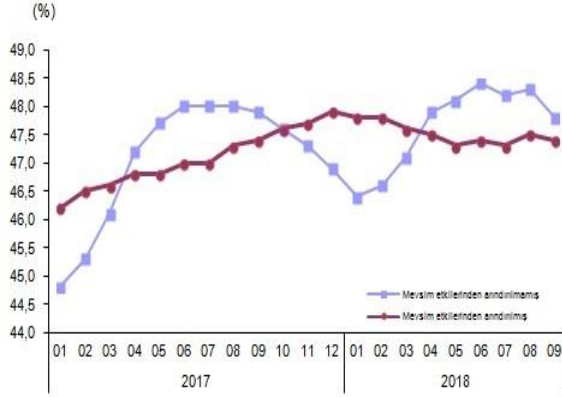
Osman her iki atışında da oku aynı rengi denk getirmiş ve dört basamaklı bir puan kazanmıştır.

Bu verilen bilgiler ışığında Ali, Betül ve Osman’ın yarışmada oklarıyla hangi renklere isabet ettirmiş olabileceklerini dair aşağıdaki ifadelerden hangisi ya da hangileri doğrudur? (Bütün yarışmacıların oku çarka isabet etmiştir.)

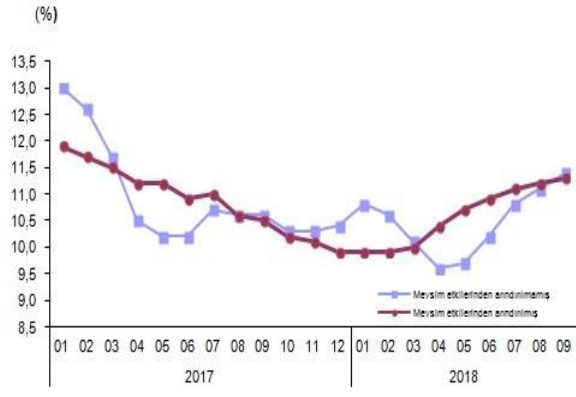
	Ali	Betül	Osman
A)	Su-Beyaz	Sarı-Beyaz	Mürdüm-Mürdüm
B)	Beyaz-Beyaz	Beyaz-Kırmızı	Sarı-Sarı
C)	Sarı-Mürdüm	Turkuaz-Beyaz	Su-Su
D)	Mavi-Turkuaz	Beyaz-Su	Turkuaz-Turkuaz
E)	Su-Su	Beyaz-Mavi	Beyaz-Beyaz

7)

İstihdam oranı, Eylül 2018



İşsizlik oranı, Eylül 2018



(Kareli şekilde olan çizgi grafikleri mevsim etkisinden arındırılmamış verileri, daireli şekilde olan çizgi grafikleri mevsim etkisinden arındırılmış verileri temsil etmektedir.)

TÜİK verilerine göre 2018 yılı Eylül ayı istihdam ve işsizlik oranları yukarıdaki çizgi grafiğinde gösterilmiştir. Bu grafiğe göre aşağıdaki ifadelerden hangisi ya da hangileri doğrudur?

- A) Yaz ve kış aylarındaki istihdam oranında mevsim etkisiyle beraber istihdam oranlarında büyük farklılaşmalar olmuştur.
- B) Mevsim etkisinden arındırılmamış verilere göre 2017 yılı işsizlik oranı 2018 yılı verilerine göre oldukça düşüktür.
- C) Mevsim etkisinden arındırılmış istihdam oranı verileri 2018 yılı boyunca sürekli azalmıştır.
- D) 2018 yılı mevsim etkisinden arındırılmamış işsizlik oranı yılın son beş ayında sürekli bir artış göstermiştir.
- E) 2017 yılı istihdam oranlarının mevsim etkisinden arındırılmamış verileri incelendiğinde sürekli bir artış olduğu görülmektedir.

8) Matematiği çok seven bir müteahhit, yeni inşa ettiği bir binanın zilinin düğmelerinde aşağıdaki asal sayıları kullanmıştır.

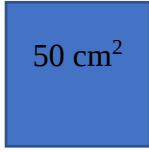
2	13	*2
3	17	*3
5	19	*4
7	23	*5
11	29	*6

Bu düğmelerde bulunmayan bir dairenin zilini çalmak isteyen kişi o daire numarasının asal çarpanlarını sırasıyla basmalıdır. Örneğin 20 numaralı dairenin zilini çalmak isteyen biri $20 = 2^2 \cdot 5$ olduğundan sırasıyla önce 2'ye ardından *2'ye ve en son 5'e basmalıdır.

Bu sistemden memnun olmayan ve müteahhite sistemin yetersiz olduğunu söyleyen ev sahipleri hangi katlarda oturuyor olabilir?

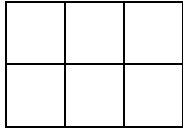
- A) 78
- B) 74
- C) 69
- D) 62
- E) 51

9)

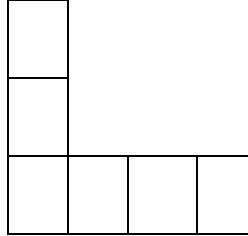


Yanda alanı verilen kare şeklindeki levhalardan 6 tanesinin bir araya gelmesiyle oluşan şekillerden hangisi ya da hangilerinin çevre uzunlukları en az $30\sqrt{10}$ 'dir.

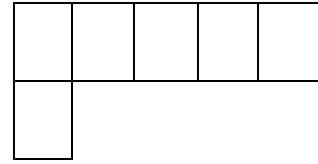
A)



B)



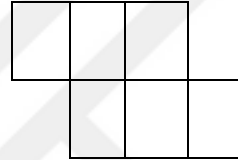
C)



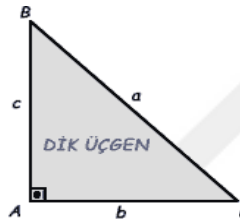
D)



E)

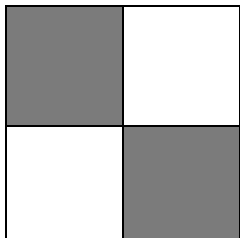


10) Bir dik üçgenin alanı dik kenarların çarpımının yarısına eşittir.

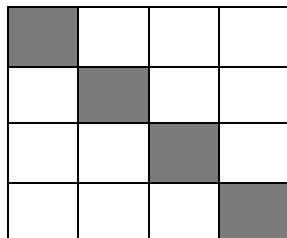


$$A(ABC) = \frac{b \cdot c}{2}$$

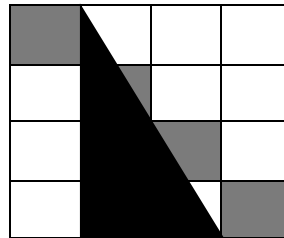
Resfebe; harf, resim ve sayıların bir arada kullanılması ile bir kelimeyi bulmaya dayanan zeka oyunudur. Örneğin; $x \square \downarrow 5 = x^2 - 5$ $y \square \uparrow zzz = y^2 + 3z$ ya da $aaa \square \downarrow bb = 3a^2 - 2b$ şeklindedir. Aşağıdaki kağıt önce dört eş kareye sonra 16 eş kareye bölünmüştür. Ardından III. şekildeki gibi bir dik üçgen çizilmiştir. Birinci durumdaki eş karelerden birinin resfebe ile gösterimi $aaaa \square \uparrow aaaaaa \uparrow 4$ ise III. şekildeki dik üçgenin alanı aşağıdakilerden hangisi ya da hangilerine eşittir?



I



II



III

A) $4 \cdot (a+1)^2$

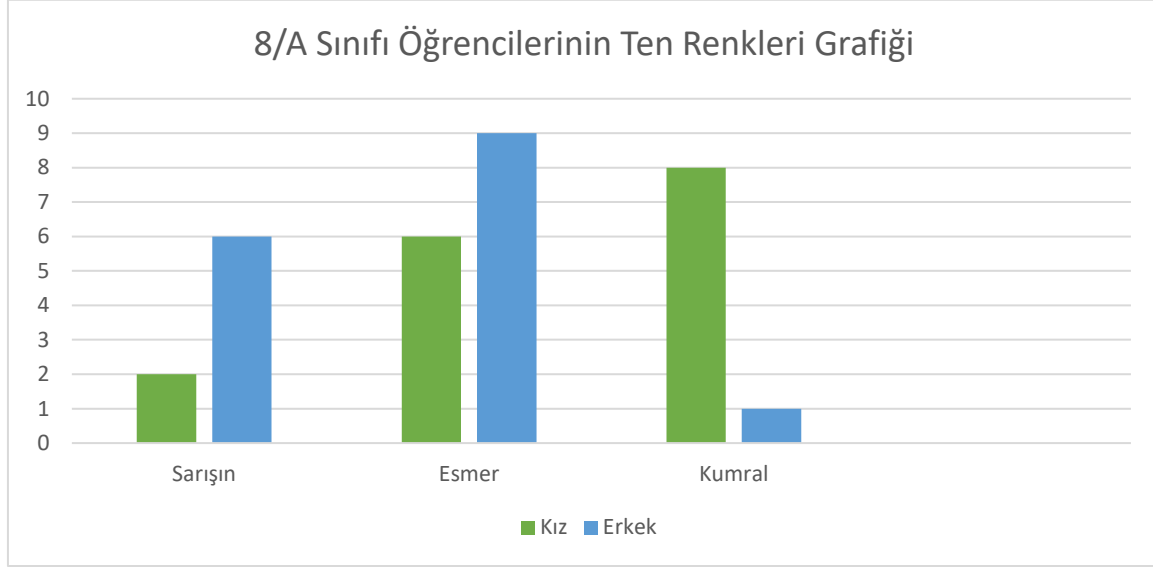
B) $8 \cdot (a-1) \cdot (a+1)$

C) $(4a+4) \cdot (a+1)$

D) $4 \cdot (2a+1)^2$

E) $8 \cdot (a+1) \cdot (a+1)$

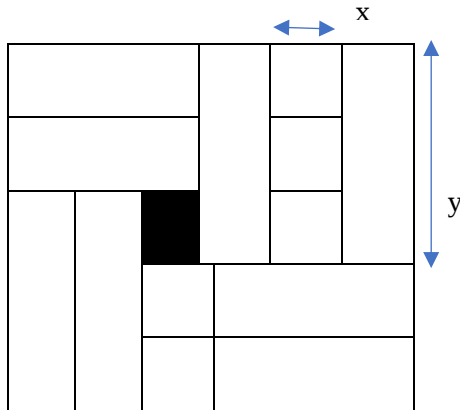
11) İncelenen bir olayın gerçekleşme olasılığının kendi dışında kalan diğer olayların gerçekleşme olasılığına oranına “odds değeri” denir. Örneğin üzerinde 1’ den 10’ a kadar sayıların yazılıp torbaya atıldığı bir deneyde çekilen topun 4 ve 4’ den küçük olma olasılığı $\frac{4}{10}$ iken geriye kalan olasılık $\frac{6}{10}$ dur ve odds değeri $\frac{4/10}{6/10} = \frac{2}{3}$ tür.



8/A sınıfı öğrencilerinin ten rengine göre gruplara ayrıldığı grafikteki bilgilere göre yapılan bir deneyin odds değeri 1 ise yapılan deneye ait olay aşağıdakilerden hangisi ya da hangileri olabilir?

- A) Sınıftan rastgele seçilen birinin erkek olması
- B) Esmer öğrenciler arasından seçilen birinin kız olması
- C) Erkek öğrenciler arasından seçilen birinin esmer olması
- D) Sarışın öğrenciler arasından seçilen birinin erkek olması
- E) Kızlar arasından seçilen birinin kumral olması

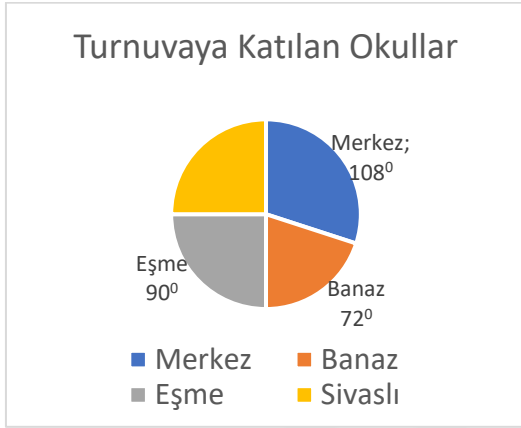
12) Birbirine eş 8 dikdörtgen ve ortalarında yer alan bir kareden oluşan şekilde eş dikdörtgenlerden her birinin uzun kenarı y birim ve kısa kenarı x birimdir. Ortadaki siyah bölgeyi gösteren cebirsel ifade aşağıdakilerden hangisi ya da hangileri olabilir?



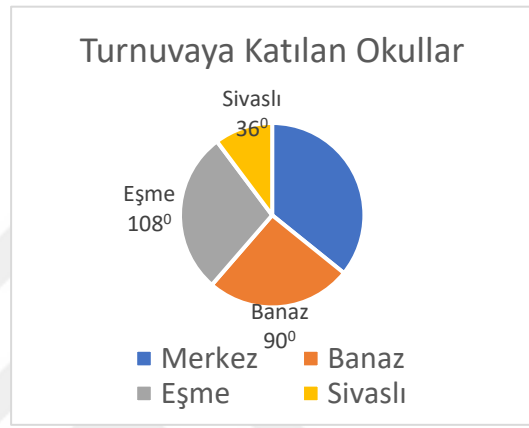
- A) $(y - 2x)^2$
- B) $(y + 2x)^2 - 8xy$
- C) $(2y - x)^2$
- D) $(2x + 2y)^2$
- E) $(x + 2y)^2$

13) Uşak İli genelinde yapılacak olan 3. Geleneksel Bocce Turnuvası'na merkez ilçe ve diğer üç ilçeden (Eşme, Banaz, Sivaslı) toplam 100 okul katılmıştır. Turnuva sonunda okullardan herhangi birine rastgele dağıtılacak olan Fair Play ödülünü alma ihtimali en yüksek olan ilçe merkez ilçe iken en az olan ilçe Sivaslı'dır. Bu verilere göre turnuvaya katılan okulların dağılımı daire grafiğiyle gösterilirse aşağıdaki grafiklerden hangisi ya da hangileri doğru olur?

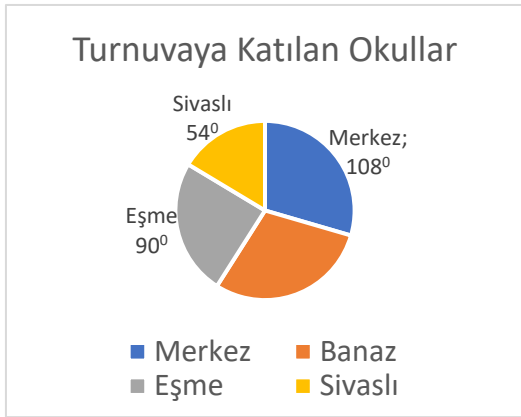
A)



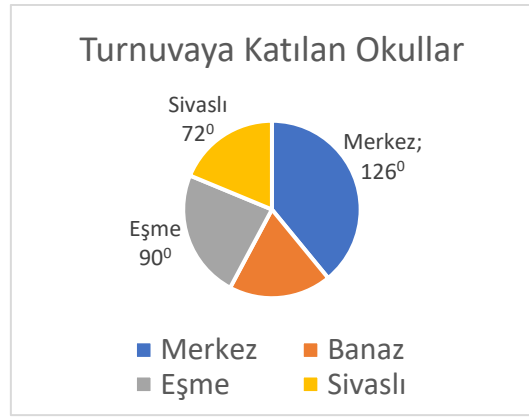
B)



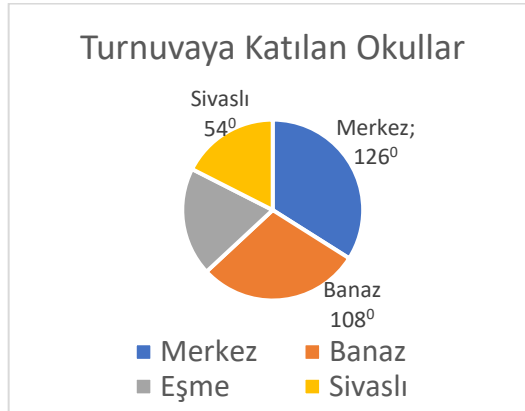
C)



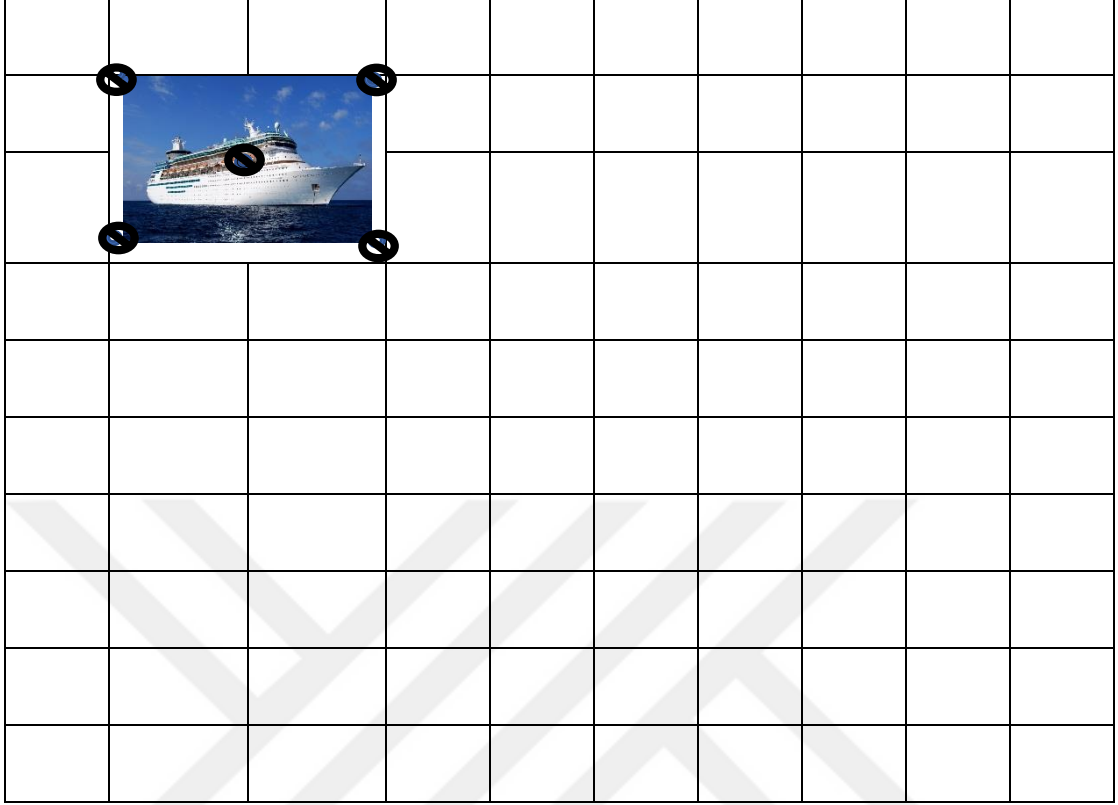
D)



E)



14)

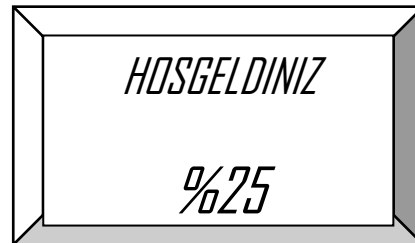


Amiral Battı oyunu şansa ve tahmine bağlı bir strateji oyunudur. Oyun başlangıcında koordinat düzlemi hazırlanır ve gemiler yerleştirilir. Daha sonra sıralı ikililer söylenerek rakibin yerleştirdiği gemiler vurulmaya çalışılır. Rakibinin bütün gemileri önce batıran oyuncu oyunun galibi olur.

Firdevs ve Büşra'nın oynadıkları oyunda Büşra'nın son gemisi şekildeki gibidir ve Firdevs $(-1,1)$ ve $(-2,2)$ sıralı ikililerini söyleyerek Büşra'nın gemisinin iki noktasından vurmaya başarmıştır. Buna göre Firdevs aşağıdaki sıralı ikililerden hangisi veya hangilerini söylerse Büşra'nın gemisine isabet ettirebilir? (Şekilde verilen 5 noktanın hepsi geminin noktalarıdır.)

- A) $(-3,3)$ B) $(-3,-3)$ C) $(-3,1)$ D) $(-1,3)$ E) $(3,-1)$

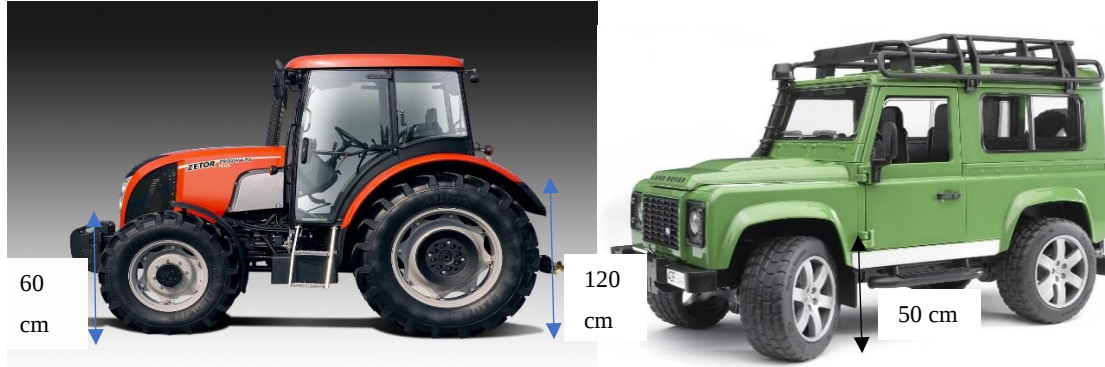
15) İzmir - Gaziantep arası uçak seferleri düzenleyen bir havayolu firmasının başlattığı kampanya ile firmayı ilk defa tercih eden yolculara özel %25 “HOŞGELDİNİZ” indirimi yapılmaktadır.



30/04/2019 saat 16.30 seferinde 150 yolcu kapasiteli bir uçakta bazı boş yerlerin kaldığı ve “HOŞGELDİNİZ” biletli yolcuların, diğer yolcu sayısının 2 katından 6 fazla olduğu biliniyor. Bu seferden havayolu firması 16.055 TL gelir elde ettiğine göre kaç koltuk boş kalmış olabilir?

- A) 8 B) 18 C) 28 D) 38 E) 48

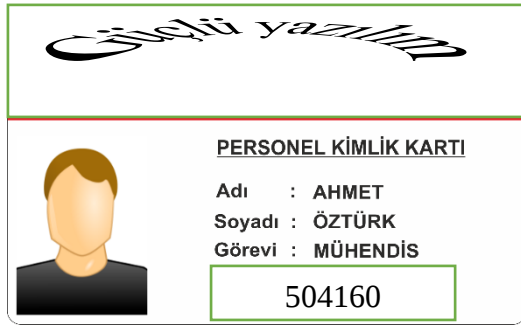
16) Yarıçapının uzunluğu r olan bir çemberin çevresi $2\pi r$ dir



Şekildeki traktörün ön teker çapı 72 cm ve arka teker çapı 120 cm, arazi aracının ise tekerlek çapı 50 cm dir. Tarlasındaki işleri halledip evine dönmek için yola çıkan iki kardeşten biri traktörü diğeri arazi aracını kullanıyor. Kardeşler eve ulaştıklarında hem traktörün ön ve arka tekerleklerin ikisinin de tam tur döndüğünü hem de arazi aracının tekerlerinin tam tur döndüğünü görüyorlar. Kardeşlerin tarlaları ile evleri arası 7 km den az olduğuna göre bu uzaklık metre cinsinden aşağıdakilerden hangisi ya da hangileri olabilir? ($\pi=3$ alınız.)

- A) 1500 B) 3600 C) 5400 D) 6600 E) 7200

17) Bir yazılım firması, çalışanlarına işe girişte okutmalar için çipli personel kartları dağıtmıştır. Bu kartların üzerinde altı haneli bir seri numaraları ve o seri numaralarından üretilen iki haneli kart kodları bulunmaktadır.



Seri No: 504160

Kart Kodu : $5 \cdot 2^3 + 0 \cdot 2^2 + 4 \cdot 2^1 + 1 \cdot 2^0 + 6 \cdot 2^{-1} + 0 \cdot 2^{-2}$

Buna göre Kart Kodu 39 olan bir çalışanın personel kartının seri numarası aşağıdakilerden hangisi ya da hangileri olabilir?

- A) 320428 B) 411004 C) 234124 D) 153224 E) 312544

18) Bir olayın olma olasılığı $= \frac{\text{İstenilen olası durumların sayısı}}{\text{Tüm olası durumların sayısı}}$

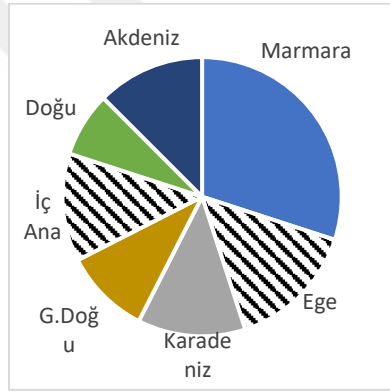
Türkiye'nin nüfusunun bölgelere göre dağılımının yaklaşık değerleri aşağıdaki tabloda verilmiştir.

Tablo: Türkiye'nin Bölgelere Göre Nüfus Dağılımı

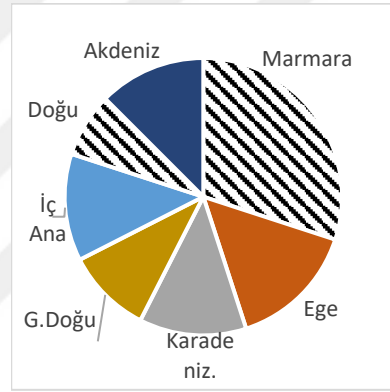
Bölge Adı	Nüfusu(Milyon)
Marmara	24
Ege	12
İç Anadolu	10
Güneydoğu Anadolu	8
Karadeniz	10
Doğu Anadolu	6
Akdeniz	10

Tablo daire grafiğine dönüştürülüp dart hedef tahtası haline getirilmiştir. Hedef tahtasına atış yapan birinin okunun hedef tahtasındaki hangi bölgelere isabet etme olasılığı toplamı $\frac{1}{4}$ tür.(Taralı bölgeler okun isabet etme ihtimali olan bölgelerdir.)

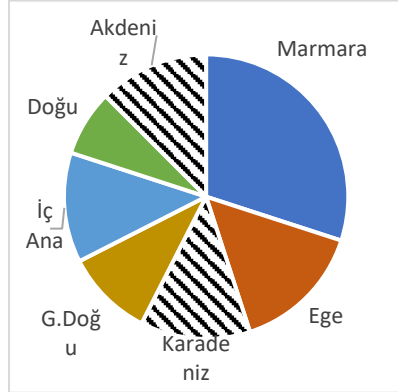
A)



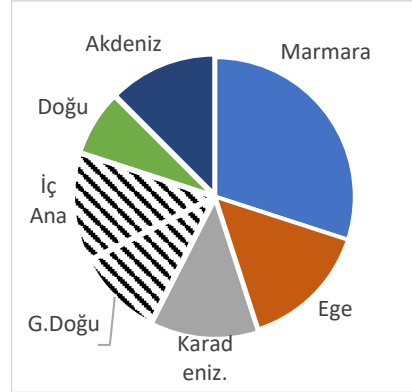
B)



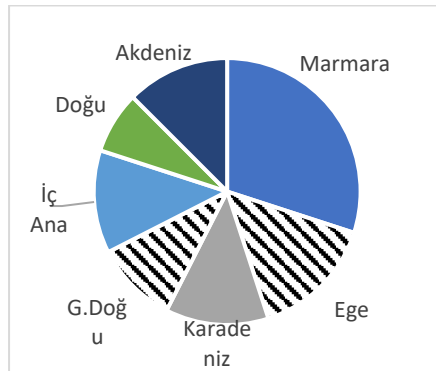
C)



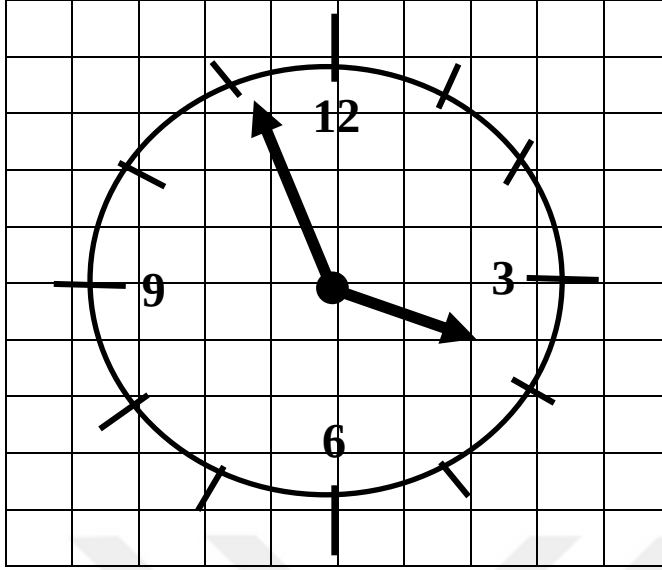
D)



E)



19)



Kareli zemin üzerine yerleştirilen saatin akrep ile yelkovanının bağlı olduğu merkez (0,0) noktasıdır.

Saat 03.56 de uyuyan bir kişi 5 saat sonra uyandığında akrep aşağıda verilen hangisi ya da hangilerinden geçmiştir?(Akrep boyu kısa olan göstergenin adıdır.)

- A) (-2,-1) B) (2,0)
C) (1,2) D) (-2,1)
E) (1,-2)

20) Bir ayrıtının uzunluğu x cm olan küp şeklindeki bir hediye paketinin 4 yan yüzünden geçen ve eni y cm olan kurdele haricindeki kısımlar pembe renge boyanacaktır.



Buna göre boyanacak bölgenin alanı gösteren cebirsel ifade aşağıdakilerden hangisi ya da hangileri olabilir.

- A) $x^2 + 2xy + y^2$ B) $8x^2 - y^2$
C) $2x(3x - 2y)$ D) $6x^2 + 4xy$
E) $4x^2 - 4xy$

Adı Soyadı

Sınıfı

S. No

ZIPGRADE.COM

ÇokluMatematik2 (0906)

- | | | | |
|----|---------------------|----|---------------------|
| 1 | (A) (B) (C) (D) (E) | 12 | (A) (B) (C) (D) (E) |
| 2 | (A) (B) (C) (D) (E) | 13 | (A) (B) (C) (D) (E) |
| 3 | (A) (B) (C) (D) (E) | 14 | (A) (B) (C) (D) (E) |
| 4 | (A) (B) (C) (D) (E) | 15 | (A) (B) (C) (D) (E) |
| 5 | (A) (B) (C) (D) (E) | 16 | (A) (B) (C) (D) (E) |
| 6 | (A) (B) (C) (D) (E) | 17 | (A) (B) (C) (D) (E) |
| 7 | (A) (B) (C) (D) (E) | 18 | (A) (B) (C) (D) (E) |
| 8 | (A) (B) (C) (D) (E) | 19 | (A) (B) (C) (D) (E) |
| 9 | (A) (B) (C) (D) (E) | 20 | (A) (B) (C) (D) (E) |
| 10 | (A) (B) (C) (D) (E) | | |
| 11 | (A) (B) (C) (D) (E) | | |

Test bitti, yanıtlarınızı kontrol ediniz.

EK 3. Çoklu Cevaplı Madde Türü ile Hazırlanan Türkçe Testi Maddeleri



ÇOKLU CEVAPLI MADDE TÜRÜ İLE HAZIRLANAN

8. SINIF TÜRKÇE SINAVI SORULARI

AÇIKLAMALAR

- Bu sınav 20 sorudan oluşmaktadır ve her sorunun beş seçeneği vardır.
- Bu sınavda kullanılan sorular çoklu cevaplı madde türüne göre oluşturulmuştur. Çoklu cevaplı madde türünde her sorunun **en az bir doğru cevabı** vardır. Yani sorudaki doğru cevap sayısına bağlı olarak öğrenci iki veya daha fazla seçenek işaretleyebilir.
- Doğru olan tüm seçenekleri işaretleyen kişi **tam puan** alırken, bazı doğru seçenekleri işaretleyen kişi **kısmi puan** alır. Tüm seçenekleri işaretleyen ya da sadece yanlış seçenekleri işaretleyen kişi puanlama yöntemine göre **sıfır puan alır ya da eksi puanlar alabilir**.
- Yanlış işaretlenen seçenekler sorudan kazanılacak puanı azaltacaktır. (Sınavda düzeltme işlemi uygulanacaktır) Bu sebeple **sadece doğruluğuna emin olduğunuz seçenekleri işaretleyiniz**.
- İşaretlemelerinizi kitapçığın arkasındaki cevap formuna doğru ve eksiksiz bir şekilde kodlayınız.
- Bu sınav için verilen süre 50 dakikadır.



Aşağıdaki metni kullanarak 1. 2. ve 3. soruları cevaplayınız.

İnsanların tek başına bırakıldığı bencil ve hoyrat bir dünyada yaşıyoruz. Dostluk, komşuluk, ilgi, sevecenlik, vefa, yardım duygusu kalmadı. Kendileri için yaşayan, boş zamanları bile olsa bunu başkaları için harcamayı israf sayan bir çıkarıcılar yarışındayız. Kimsenin kimseyi göreceği hali yok.

Almanya’da oturan işçilerimizin en çok yakındığı şeylerin başında komşusuzluk geliyor. Komşusuzluk onların buradaki sıkıntılı durumunu daha da vurguluyor, büsbütün öksüz yapıyor. Oysa alışmış köyünde komşusuna başvurmaya, ona dert yanmaya, içini boşaltıp rahatlamaya, komşusunun ona ihtiyacı olunca sıcak bir çorba yapıp duasını almaya. Burada yok böyle şeyler... Sadece ona değil Alman komşulara da kapalı Alman’ın kapısı. Oturduğumuz apartmanda tam on iki daire var. Yedi aydır bu güne dek hiçbirini görmedik. Kapıda bazı gölgelerle karşılaşılıyor. Birbirimize kötü insan olmadığımızı belgeler gibi adet yerini bulsun diye merhabalaşıyoruz.(...)

Diyeceğim o ki sadece Almanya’da değil, dünyanın her yerinde dostluk, komşuluk azaldıkça azalıyor. Sanayileşme humması içinde, tüketim furyası içinde iç dünyanın en değerli unsurlarını yitiren insanoğlu doğayı unutup çevreyi kirlettiği gibi içini ısıtan insancıl değerleri de kırıp geçiriyor. Bu arada bencilliği aşmanın en yakın aşaması olan komşuluk müessesesini de miadı geçmiş romantik bir duygu olarak müzeliğe bir eşya gibi geçmişe bırakıyor.

Haldun TANER

1) Verilen metinle ilgili olarak aşağıdaki yargılardan hangisi ya da hangilerine ulaşılabilir?

- A) Günümüzde yardımlaşma, komşuluk ve vefa gibi değerler geçmişe kıyasla gitgide yok olmaktadır.
- B) Almanya’da yabancılara karşı düşmanlık beslenmekte ve yabancılara mesafe konulmaktadır.
- C) İnsanların tüketme çılgınlığı, insani değerlerin kaybına ve doğanın tahribatına sebep oluyor.
- D) Avrupa özelinde dostluk ve komşuluk bağları gittikçe azalmış, herkes bencilleşmiştir.
- E) İnsanlar sanayileşme çağında çevreyi önemsememiş ve bu durum küresel ısınmaya sebep olmuştur.

2) Metnin dil ve anlatım özellikleriyle ilgili aşağıdakilerden hangisi ya da hangileri söylenemez?

- A) Metin içerisinde deyimlerden yararlanılmıştır.
- B) Yazar eleştirel bir dil kullanmış ve günümüz sorunlarına değinmiştir.
- C) Metinde birden fazla düşünceyi tartışma ve geliştirme yoluna gidilmiştir.
- D) Metinde gerçek anlamı dışında kullanılan sözcükler bulunmaktadır.
- E) Metin içerisinde hayali unsurlara yer verilmiştir.

3) Metinde altı çizili sözcüklerin cümlelere kattığı anlam aşağıdaki seçeneklerin hangisi ya da hangilerinde bulunmaktadır?

- A) Arkadaşıma zor zamanında yaptığım iyiliğin mükafatını hiç aramadım.
- B) Eczacılar odasından gelen uyarıyla son kullanım tarihi geçen ilaçlar toplandı ve imha edildi.
- C) Başboş gezen sokak köpekleri özellikle kış aylarında çocuklarımız için çok tehlikeli oluyorlar.
- D) Öğretmenler en çok ilgisiz ve disiplinsiz öğrencilerden şikayet ediyorlar.
- E) Dostun evi gönüllerdir, gönüller yapmaya geldim.

4) Arkadaşlarıyla birlikte çıktığı turu anlatan Sümeyye'nin kurduğu aşağıdaki cümlelerden hangisi ya da hangilerinin öznesi söz öbeği şeklindedir?

- A) İstanbul turumuzun ilk gününde coşkuyla turistik mekanları gezmeye gittik.
- B) Bütün bir günümüzü İstanbul'un güzellikleri karşısında mest olarak geçirdik.
- C) Peri masallarını kışkırtacak güzelliklerle bezenmiş mekanlar çok etkileyiciydi.
- D) Bol bol resim çekme şansı yakaladığımız gezimizi anlatmak oldukça güç olacak.
- E) Turun sonunda hepimiz birer İstanbul aşığı olarak evlerimize döndük.

5) Aşağıdaki bilgilere göre, hangi teknede olduğu kesin olarak bilinmeyen kişiler kim veya kimlerdir?

Ahmet, Bünyamin, Can, Dünya, Ezel, Ferhat, Güner ve Halise adlı kişiler mavi beyaz renklerde tekneler ile Kızılırmak Nehri'nde gezintiye çıkmışlardır. Bu gezi ile ilgili olarak şunlar bilinmektedir:

- ✓Mavi tekneye beş, beyaz tekneye üç kişi binmiştir.
- ✓Bünyamin ile Dünya mavi teknededir.
- ✓Ahmet, Ferhat'ın bulunduğu teknede değildir.
- ✓Halise ve Güner aynı teknededir.
- ✓Güner ve Bünyamin farklı teknelerdedir.

- A) Ferhat
- B) Ahmet
- C) Halise
- D) Can
- E) Ezel

6)

Kişiler	Olaylar	Mimikler	Eşyalar
Sarışın Delikanlı	Buluşmaya geç kalmış biri	Telaşlı	Tarak, Ayakkabı
Ayakkabıcı	İflas etmiş biri	Öfkeli	Defter, Kalem
Misafir Komşu	Suratı kızarmış biri	Utanmış	Vazo, Top
Anne	Gözleri yaşarmış biri	Mutlu	Diploma, Mendil
Mühendis	Kaza yapmış biri	Korkmuş	Tutanak, Cep Telefonu

Bir ortaokulda oluşturulan tiyatro ekibi yukarıdaki durumlara uygun drama çalışması yapacaklardır. Bu tablodan hareketle aşağıda verilenlerden hangisi ya da hangileri yanlıştır?

- A) Oğlunun mezuniyet törenine giden annesi bir elinde mendiliyle gözlerindeki yaşları silerken diğer elinde tuttuğu oğlunun diplomasına gururla bakıyordu.
- B) Elinde defter kalemiyle hesap işlerine dalan ayakkabıcı, iflas etme sebebinin çok fazla veresiye ürün dağıtması olduğunu fark edince deliye döndü.
- C) Bir yandan ayakkabılarını giyerken diğer yandan saçlarını taramaya çalışan sarışın delikanlı kız arkadaşını bekletmek istemiyordu.
- D) Oğlunun, top ile ev sahibinin vazosunu kırması misafir komşunun yüzünü mahcubiyetten kızartmıştı.
- E) İşe yetişirken önündeki araca çarpan mühendis, tutanak tutup kaza mahallinin resmini cep telefonuyla çekerken neden dikkatli araç kullanmadığının iç hesabını yapıyordu.

7) Firdevs:

Furkan: Senin söylediklerine katılmıyorum. Bir filmin çok izlenmesinde ölçüt sadece bu değildir. Filmin senaryosu, akıcılığı, etkileyici bir son ile bitmesi de çok izlenmesini sağlayan etkenlerdendir.

Yukarıdaki karşılıklı konuşmada Firdevs'in savunduğu Furkan'ın ise karşı olduğu düşünce aşağıdakilerden hangisi ya da hangilerinde ifade edilmiştir?

- A) Konusuyla birçok kişiyi etkileyen yönetmenin son filmi herkes tarafından beğenilir bence.
- B) Özgün bir içeriği olan ve beklenmeyen bir sonla biten filmlerin de izleyicisi beklenenden fazla olabilir.
- C) Yönetmenin filmin oyuncu kadrosunu ünlülerden oluşturması gişede büyük başarı getirecektir.
- D) Film yapımcısı en son filmi için gala düzenleyip reklamını yapmış, filmin epey izleyicisi çıkar.
- E) Bu film kurgusuyla hem çocuklara hem de büyüklere hitap ediyor, diğerlerinden daha fazla seyredilir.

8)



Aşağıdaki seçeneklerden hangisi ya da hangileri yukarıdaki görseli doğru biçimde ifade edebilir?

- A) Çevre kirliliği doğaya zarar vererek doğrudan ya da dolaylı şekilde doğadaki canlıları olumsuz etkiler.
- B) Yeniden değerlendirilebilme imkanı olan atıkların çeşitli işlemlerle üretim sürecine tekrar katılması kaynakların tükenmesini önler.
- C) Fabrikalardan çıkan zararlı gazlar atmosferdeki ozon tabakasını etkiler ve bu durum dünyamızın zararlı ışınlarla maruz kalmasına sebep olur.
- D) Yetişkin bir gürgen ağacı günde yaklaşık bin kişinin çıkardığı karbondioksiti emebilir ve yılda yaklaşık 300 kg zehri süzebilir.
- E) Geri dönüşüm sisteminin yaygınlaşmaması atık çöplerin artmasına ve çevre kirliliğine sebep olur.

9) Gökova Koyu, körfezin kuzeybatısında küçük bir girinti şeklindedir. Beton rıhtım üzerindeki iskele koyun belirlenmesini sağlar. Doğusunda denize dökülen ırmak, yürürken ayaklarınızın altını yakan kum bir sıgılık oluşturur. Irmağın kuzeyindeki köyün meydanında küçük bir lokanta vardır. Çınar ağaçlarının altındaki bu yerde kuş cıvıltılarının eşliğinde sipariş verilen ev yemeklerinin gelmesini beklerken insanın iştahı daha da açılır.

Bu parçanın anlatımında aşağıdaki duyuların hangisi ya da hangilerinden yararlanılmamıştır?

- A) Tatma B) İşitme C) Görme D) Hissetme E) Koklama

Verilen bilgilere göre 10 ve 11. Soruları cevaplayınız.

Bir koşu yarışmasına; Akif, Burak, Ceren, Davut, Enes, Faruk ve Gamze adlı sporcular katılmıştır. Bu koşu yarışması sonunda elde edilen derecelerle ilgili aşağıdakiler bilinmektedir.

- Yarışmada Gamze birinci olamamıştır.
- Yarışmayı Ceren'den hemen sonra bitiren kişi Akif'tir.
- Akif yarışmayı 5. sırada tamamlamıştır.
- Enes ile Akif arasında yarışmayı 2 kişi tamamlamıştır.
- Gamze yarışmayı Davut'tan önce bitirmiştir.

10) Yukarıda verilen bilgilere göre yarışmayı kimlerin birinci bitirebilme ihtimali vardır?

- A) Faruk B) Burak C) Ceren D) Enes E) Davut

11) Yukarıda verilen bilgilere göre Gamze kaçınıcı olmuş olabilir?

- A) 2 B) 3 C) 4 D) 6 E) 7

12) Edilgen fiiller, yüklemde belirtilen işi gerçekleştiren öznenin belli olmadığı, işin kim tarafından yapıldığının tahmin edilemediği cümlelerin çatı özelliğini ifade etmektedir.

Etken fiiller, yüklemde belirtilen işi gerçekleştiren öznenin apaçık ortada olduğu cümlelerin çatı özelliğini ifade etmektedir.

Aşağıdaki cümlelerin hangisi ya da hangilerinde edilgen çatılı bir fiil vardır?

- A) 19 Mayıs için hazırlanan resimler okulun giriş kısmındaki panoda sergilendi.
- B) Hızla yağan yağmurun altında arabasına koşar adımlarla gidiyordu kırmızı ceketli kişi.
- C) Gelecek yıl uygulanacak olan yeni sınav sistemiyle ilgili düzenlemeler tamamlandı.
- D) Dere kenarından hızla geçen atlıların ardından bakıp kaldılar.
- E) Tembelliği ve iş beğenmezliğiyle herkesi kendinden soğutmayı başarmıştı.

13) Düşündükleri gerçekleşmeyince planlarını askıya aldı.

Yukarıdaki cümle ile ilgili aşağıda verilen bilgilerden hangisi ya da hangileri yanlıştır?

- A) Adlaşmış sıfat fiil kullanılmıştır. B) Zarf fiil kullanılmıştır.
C) Vurgu yer tamlayıcısı üzerindedir. D) Nesne yüklem ilişkisine göre geçişlidir.
E) Cümle yapısı bakımından basittir.

Aşağıdaki metni kullanarak 14, 15 ve 16. soruları cevaplayınız.

Cep telefonlarının yaygınlaşmasıyla birlikte insanlar kendilerini telefon ekranlarına bakmaktan alıkoyamaz oldu. Çocuklar hafta içi ve hafta sonu saatlerce dijital araçları kullanarak zaman harcamaktadırlar. Özellikle hafta sonları genel olarak öğrencilerin ortalama 6- 7 saat bilgisayar başında zaman harcadıkları bilinmektedir. Birçok sorunu da beraberinde getiren bu durumun önlenmesi için önlemler alınmaya çalışılmaktadır. Kullanıcılar en fazla neye vakit geçiriyor diye bakıldığında sosyal medya, oyun, video ve sohbet gibi başlıklar öne çıkmaktadır. Kitap okuma, araştırma yapma gibi yararlı amaçlar yerine daha çok oyun ve “Youtuber”lerin videolarının seyredilmesi beraberinde olumsuz sonuçları da getirmektedir.

Bu kadar yoğun telefonla zaman geçirmek bağımlılığa neden olmaktadır. Bu bağımlılıktan kurtulmak için son günlerde *dijital detoks* akımı başladı. Dijital detoks, kişinin bir süre teknolojik cihazlardan uzak kalması olarak tanımlanmaktadır. İşiniz gereği telefonsuz kalmanız mümkün değilse bile sadece arama yapabilen ve mesaj atabilen eski tip bir telefon kullanmanız gerekiyor.

Dijital oyun bağımlılığı ve oyunların olumsuz etkisi tüm dünyada ölümle ve yaralanmayla sonuçlanan facialara yol açabiliyor. Geçtiğimiz günlerde Florida’da düzenlenen bir sanal oyun turnuvasında oyunculardan biri çevresine ateş açarak üç kişinin ölümüne, on bir kişinin yaralanmasına neden oldu. Sanal oyunlara ayrılan sürenin fazla olması, kimi oyunların bağımlılık yapması, çocukların şiddet eğiliminin artması, kimi oyunların kumara özendirilmesi gibi nedenlerle, Çin hükümeti yeni oyun satışlarını onaylamamaktadır. Dünyanın en büyük bilgisayar oyun pazarına sahip Çin’de geçtiğimiz 2018 Mart ayından bu yana hiçbir oyunun satılmasına izin vermemektedir. Sadece önceden onay alınmış oyunlar satılabilmektedir. Üstelik oyunlara yönelik bağımlılık sadece çocuklarda olan bir durum değildir. İngiltere’de 2018 yılında 200’den fazla çift sadece bilgisayar oyunları nedeniyle boşanmışlardır. Ayrıca Dünya Sağlık Örgütü oyun bağımlılığının toplum sağlığını oldukça olumsuz etkilediğini açıklamışlardır. Bütün bu çabaların ne kadar işe yaracağı tam olarak bilinmemektedir. Ancak sağlıklı bir dijital yaşam için başta Facebook, YouTube gibi teknoloji platformlarının da bu konuda hükümetlerle birlikte çalışması gerekmektedir.

14) Metnin yazarı siz olsaydınız ve yetkili bir konumda bulunsanız halka sağlıklı dijital yaşam sunmak için alacağınız tedbirler hangisi ya da hangileri olacaktır?

- A) Teknolojik araçlar yaşam için vazgeçilmez araçlar olduğundan kullanımına asla kısıtlama getirmezdim.
- B) Sosyal medya kullanımını artırmak için çalışmalar yapar ve bu konuda ülkemizin dünya lideri olmasını isterdim.
- C) Teknoloji bağımlılığıyla mücadele için devlet bünyesinde sağlık kuruluşları açılması için çalışmalar yapardım.
- D) Sağlıklı teknoloji kullanımı için ebeveynleri bilinçlendirecek kampanyalar yapardım.
- E) Çocukların faydalanmaları için derslere ait ücretsiz videolar yayınlanmasını sağladım.

15) Parçada verilen bilgilere göre dijital oyunlarla ilgili aşağıdakilerden hangisi ya da hangileri çıkarılabilir?

- A) Dijital oyunlar sadece çocuklara yönelik eğlenceli ancak bağımlılık yapabilecek teknolojik uğraşlardır.
- B) Çin’de birçok ailenin boşanmasında dijital oyun bağımlılığı etkili olmuştur.
- C) İngiltere’de dijital oyun bağımlılığıyla mücadele için Dünya Sağlık Örgütü’nden yardım istenmiştir.
- D) Dijital oyun bağımlılığı tüm dünyada ölümle veya yaralanmayla sonuçlanan facialara yol açabilmektedir.
- E) Dijital oyunlar çocukların şiddet eğilimini artırmasının yanında onları kumara da özendirilmektedir.

16) Metinde koyu yazılmış cümlede yapılan anlatım bozukluğu aşağıdakilerden hangisi ya da hangileri yapılarak giderilebilir?

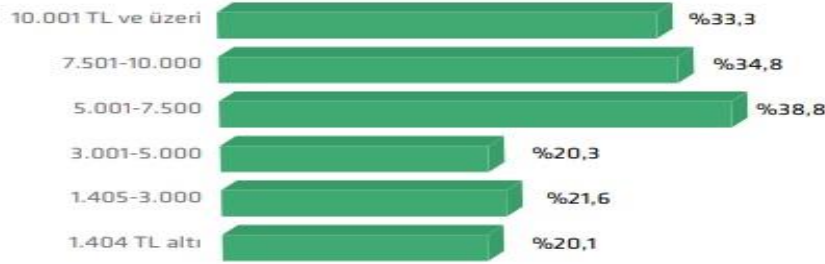
- A) “Telefonla” sözcüğü başa getirilerek.
- B) “Telefonla” ve “zaman” sözcüklerinin yerleri değiştirilerek.
- C) “Bu kadar yoğun” sözcük grubu “geçirmek” sözcüğünden sonra getirilerek.
- D) “Bağımlılığa” sözcüğü başa getirilerek.
- E) “Neden olmuştur” sözcük grubu yerine “sebebi olmuştur” sözcük grubu getirilerek.

17)

Grafik 99. Bireylerin Satın Alınan Gıdaları Tüketmeden Çöpe Atma Nedenleri



Grafik 98. Bireylerin Aylık Gelir Düzeyine Göre Satın Alınan Gıdaları Tüketmeden Çöpe Atma Durumu



Ticaret Bakanlığı'nın 2018 yılı aralık ayında yayınladığı "Türkiye İsrar Raporu" adlı çalışmasından elde edilen sonuçlardan iki tanesi yukarıdaki grafiklerde verilmiştir.

Buna göre aşağıdakilerden hangisi ya da hangilerine ulaşılabilir?

- A) Bireylerin gelir düzeyi ile tüketilmeden çöpe atılan gıda miktarları arasında doğrudan bir bağ kurulamaz.
- B) Bireylerin satın aldıkları gıdaları tüketmeden çöpe atmalarındaki en önemli etkenlerden biri gıdaları gereğinden fazla satın almalarıdır.
- C) Gıdaları tüketmeden çöpe atma durumunda en bilinçli kesimin gelir düzeyi en düşük kesim olduğu söylenebilir.
- D) Bireylerin gıdaları alırken saklama kapasitelerine göre hareket etmeleri israfı çok önemli düzeyde önleyecektir.
- E) Gıdaları çöpe atma nedenleri değerlendirildiğinde; gerek bozulması gerekse tüketilememesi, gıda alışverişinde bilinçsiz davranıldığını düşündürmektedir.

18) Aşağıdakilerden hangisi ya da hangilerinde birinci dize ikinci dizenin gerekçesidir?

- A) Fani ömür biter, bir uzun sonbahar olur
çiçek, kuşlar ve dağlar tarumar olur.
- B) Dağlar gibi yığılmış tarlada ekinler Yaprak,
Sağa sola düştükçe sarı başaklar inler.
- C) Her şey güzel; ufuklar, deniz, akşam, su, rüzgar
Her yerde seni gördüm, gezerken diyar diyar
- D) Birçok gidenin her biri memnun ki yerinden
Birçok seneler geçti dönen yok seferinden.
- E) Yıllar yılı saatim işlemiş, ben durmuşum
Gökyüzünden habersiz uçurtma uçurmuşum.

19) Tablo: Küresel akıllı telefon pazarındaki 2018 yılı 3. Çeyreği ile 2017 yılı 3. Çeyreği karşılaştırması

Şirket	Satış Adedi (2018 yılı 3. Çeyrek)	Pazar Payı (2018 yılı 3. Çeyrek)	Satış Adedi (2017 yılı 3. Çeyrek)	Pazar Payı (2017 yılı 3. Çeyrek)	Yıllık Değişim
Samsung	72,2	%20,30	83,3	%22,10	-%13,40
Huawei	52	%14,60	39,1	%10,40	%32,90
Apple	46,9	%13,20	46,7	%12,40	%0,50
Xiaomi	34,3	%9,70	28,3	%7,50	%21,20
OPPO	29,9	%8,40	30,6	%8,10	-%2,10
Diğerleri	119,9	%33,80	149,8	%39,60	-%19,90
Toplam	355,2	%100,00	377,8	%100,00	-%6,00

Yukarıdaki tablo küresel bazda akıllı telefon üreten firmaların 2018 ve 2017 yılları 3. çeyreğine ait pazar paylarını ve satış adetlerini göstermektedir. Bu tabloya göre aşağıdakilerden hangisi ya da hangilerine ulaşılabilir?

- A) Xiaomi marka telefonların geçen bir yıllık süreçte hem pazar payları hem de satış adetleri artmıştır.
- B) Toplam akıllı telefon satışları 2018 yılı 3. çeyreği itibariyle geçen yılın aynı çeyreğine göre %6 azalmıştır.
- C) Samsung marka telefonlar geçen bir yıllık süreçte satışlarında büyük azalma yaşamıştır ve pazardaki liderliğini kaybetmiştir.
- D) Akıllı telefon pazarının önde gelen beş firmasından üç tanesinin 2018 yılında bir önceki yıla göre satışlarında azalma olmuştur.
- E) Diğer akıllı telefon üreticileri bir yılda satışlarında yaklaşık %20' lik kayıp yaşamışlardır.

20)

I

Kişi adlarından
sonra gelen saygı
sözlerine getirilen
ekleri ayırmak
için kullanılır.

II

Kısaltmalara
getirilen ekleri
ayırmak için
kullanılır.

III

Sayılarla getirilen
ekleri ayırmak için
kullanılır.

IV

Şiirde seslerin ölçü
dolayısıyla düştüğünü
göstermek için
kullanılır

V

Bir ek veya harften
sonra gelen ekleri
ayırmak için kullanılır.

TBMM' nin 99. kuruluş yıl dönümü olan 2018 yılı Nisanının 23'üncü günü coşkuyla güne uyanmış ve hazırlıklarımıza başlamıştık. Bizim tatlı telaşımızı gören Doğan Bey'in de heyecanı artmış ve torununu yanına alıp tören alanına doğru gitmişti. Tören alanı hıncahınç doluydu ve herkes ellerindeki bayrakları sallıyordu. Meydana kendini atan bir ozan Karacaoğlu'nun bir dördlüğünü sazının eşliğinde söylüyordu. "Türk oğlu Türk'üm ben" diye haykırıyordu gaziler.

Bu metinde kesme işareti numaralandırılmış görevlerinden hangisi veya hangileriyle kullanılmıştır?

A) I

B) II

C) III

D) IV

E) V

Adı Soyadı

Sınıfı

S. No

ZİP GRADE.COM

- | | | | |
|----|---------------------|----|---------------------|
| 1 | (A) (B) (C) (D) (E) | 12 | (A) (B) (C) (D) (E) |
| 2 | (A) (B) (C) (D) (E) | 13 | (A) (B) (C) (D) (E) |
| 3 | (A) (B) (C) (D) (E) | 14 | (A) (B) (C) (D) (E) |
| 4 | (A) (B) (C) (D) (E) | 15 | (A) (B) (C) (D) (E) |
| 5 | (A) (B) (C) (D) (E) | 16 | (A) (B) (C) (D) (E) |
| 6 | (A) (B) (C) (D) (E) | 17 | (A) (B) (C) (D) (E) |
| 7 | (A) (B) (C) (D) (E) | 18 | (A) (B) (C) (D) (E) |
| 8 | (A) (B) (C) (D) (E) | 19 | (A) (B) (C) (D) (E) |
| 9 | (A) (B) (C) (D) (E) | 20 | (A) (B) (C) (D) (E) |
| 10 | (A) (B) (C) (D) (E) | | |
| 11 | (A) (B) (C) (D) (E) | | |

Çoklu Türkçe (9140)

Test bitti, yanıtlarınızı kontrol ediniz.

EK 4. Anket Formu

Değerli Öğrenci;

Bu çalışmada sizden uygulamasına katılmış olduğunuz çoklu cevaplı madde türü ile çoktan seçmeli madde türünün karşılaştırılmasına dair görüşlerinizi almak amaçlanmaktadır. Elde edilen veriler akademik çalışmalar için kullanılacak ve isimsiz olarak saklanacaktır. Aşağıdaki maddelere vereceğiniz yanıtlar çalışmaya katkı sağlayacaktır bu sebeple çalışmaya vereceğiniz katkı için teşekkür ederiz.

Ömer DOĞAN

Gazi Üniversitesi Eğitimde Ölçme ve Değerlendirme Anabilim

Dalı Tezli Yüksek Lisans Öğrencisi

Prof. Dr. İsmail KARAKAYA

Gazi Üniversitesi Eğitimde Ölçme ve Değerlendirme

Anabilim Dalı Öğretim Üyesi

İfadeler	Çoklu Cevaplı	Çoktan Seçmeli
1) Daha iyi puan alabileceğinizi düşündüğünüz madde türü		
2) Bildiklerinizi daha iyi yansıtabileceğiniz madde türü		
3) Cevaplaması daha kolay olan madde türü		
4) Kişiler arası ayrımı daha iyi yapabilen madde türü		
5) LGS’de kullanılmasını isteyeceğiniz madde türü		





GAZİLİ OLMAK AYRICALIKTIR..