

**T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İLKÖĞRETİM BÖLÜMÜ**

**İLKÖĞRETİM ÖĞRETMEN ADAYLARININ
ÖLÇME-DEĞERLENDİRME OKURYAZARLIĞI VE
ÖLÇME-DEĞERLENDİRMEYE İLİŞKİN
TUTUMLARININ BELİRLENMESİ**

Emrah GÜL

Yüksek Lisans Tezi

İlköğretim Anabilim Dalı

Danışman: Yrd. Doç. Dr. İsmail TÜRKOĞLU

HAZİRAN-2011

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
İLKÖĞRETİM BÖLÜMÜ

İLKÖĞRETİM ÖĞRETMEN ADAYLARININ ÖLÇME-DEĞERLENDİRME
OKURYAZARLIĞI VE ÖLÇME-DEĞERLENDİRMEYE İLİŞKİN
TUTUMLARININ BELİRLENMESİ

YÜKSEK LİSANS TEZİ

Emrah GÜL
091140105

Tezin Enstitüye Verildiği Tarih: 16.05.2011

Tezin Savunulduğu Tarih: 03.06.2011

Tez Danışmanı: Yrd. Doç. Dr. İsmail TÜRKÖĞLU (F.Ü)
Diğer Jüri Üyeleri: Doç. Dr. Erdal CANPOLAT (F.Ü)
Yrd. Doç. Dr. İ. Bakır ARABACI (F.Ü)

HAZİRAN – 2011

ÖNSÖZ

En zor anımda yanımda olan ve manevi desteğini esirgemeyen değerli hocam Yrd. Doç. Dr. İsmail TÜRKOĞLU'na sonsuz teşekkürlerimi sunuyorum. Görüş ve önerilerini hiçbir zaman esirgemeyen Doç. Dr. Çetin SEMERCİ ve Doç. Dr. Burhan AKPINAR'a ayrıca akademik yazışmalarıyla beni yönlendiren Prof. Dr. Plake Barbara'ya teşekkür ederim.

Emrah GÜL
ELAZIĞ-2011

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ	II
ÖZET	VII
SUMMARY	VIII
TABLolar LİSTESİ	IX
GRAFİKLER LİSTESİ	XI
1. GİRİŞ	1
1.1. Problem Durumu	1
1.2. Araştırmanın Amacı	4
1.3. Araştırmanın Önemi	6
1.4. Araştırma Soruları	7
1.5. Sayıtlar	8
1.6. Sınırlılıklar	8
1.7. Tanımlar	8
2. LİTERATÜR VE KAVRAMSAL ÇERÇEVE	10
2.1. LİTERATÜR	10
2.1.1. Ülkemizde Yapılan Araştırmalar	10
2.1.2. Yurtdışında Yapılan Araştırmalar	16
2.2. KAVRAMSAL ÇERÇEVE	19
2.2.1. Eğitimde Ölçme Ve Değerlendirme	19
2.2.1.1. Eğitimde Ölçme	19
2.2.1.1.1. Ölçme Araçlarının Özellikleri	20
2.2.1.1.2. Eğitimde Değerlendirme	23
2.2.1.1.2.1. Tanıma ve Yerleştirmeye Yönelik Değerlendirme	24
2.2.1.1.2.2. Biçimlendirme ve Yetiştirmeye Yönelik Değerlendirme	25
2.2.1.1.2.3. Değer Biçmeye Yönelik Değerlendirme	26
2.2.2. Eğitimde Ölçme ve Değerlendirme Yaklaşımları	26
2.2.1.3. Geleneksel Ölçme ve Değerlendirme Yaklaşımları	27
2.2.1.3.1. Yazılı Yoklamalar	29
2.2.1.3.2. Kısa Cevaplı Testler	30
2.2.1.3.3. Çoktan Seçmeli Testler	32

2.2.1.3.4.	Dođru-Yanlıř Testleri	34
2.2.1.3.5.	Sözlü sınavlar	36
2.2.1.3.6.	Eřleřtirmeli Sorular	37
2.2.1.3.7.	Kısa Cevaplı Sorular.....	38
2.2.1.4.	Alternatif Ölçme ve Deđerlendirme Yaklařımları.....	38
2.2.1.4.1.	Puanlama Ölçeđi (Rubrik).....	41
2.2.1.4.2.	Öđrenci Ürün Dosyası (Portfolyo)	42
2.2.1.4.3.	Performans Ödevi	43
2.2.1.4.4.	Akran Deđerlendirmesi.....	44
2.2.1.4.5.	Grup Deđerlendirmesi	44
2.2.1.4.6.	Öz Deđerlendirme.....	45
2.2.1.4.7.	Kavram Haritası.....	45
2.2.1.4.8.	Kontrol listesi	47
2.2.1.4.9.	Gözlem Formu.....	47
2.2.1.4.10.	Yapılandırılmıř Grid.....	48
2.2.1.4.11.	Tutum Ölçeđi.....	50
2.2.1.4.12.	Tanılayıcı Dallonmıř Ađaç	50
2.2.1.4.13.	Poster Deđerlendirme	52
2.2.1.4.14.	Mülakat (Görüşme)	52
3.	MATERYAL VE METOT	54
3.1.	Arařtırmanın Modeli	54
3.2.	Evren ve Örneklem.....	54
3.2.1.	Pilot Çalıřma	54
3.2.2.	Ana Çalıřma	55
3.3.	Veri Toplama Teknikleri	56
3.3.1.	Sınıfta Ölçme ve Deđerlendirme Okuryazarlık Envanteri (SDOE)	56
3.3.1.1.	Envanterin Türkçeye Çevrilmesi, Uyarlanması ve Geliřtirilmesi	57
3.3.1.2.	Pilot Çalıřma	57
3.3.1.3.	Envanterin Düzenlenmesi.....	58
3.3.1.4.	Ana Çalıřma	58
3.3.2.	Ölçme ve Deđerlendirmeye İliřkin Tutum Ölçeđi	59
3.3.2.1.	Ölçek Maddelerinin Oluřturulması	59
3.4.	Verilerin Toplanması.....	60

3.5.	Verilerin Analizi	60
4.	BULGULAR	61
4.1.	Araştırma Sorularına Ait İstatistikî Veriler	61
4.1.1.	Birinci Araştırma Sorusu	64
4.1.1.1.	Fen Bilgisi Öğretmen Adaylarının (N=54) Ölçme ve Değerlendirme Okuryazarlık Seviyeleri	64
4.1.1.2.	Sosyal Bilgiler Öğretmen Adaylarının (N=55) Ölçme ve Değerlendirme Okuryazarlık Seviyeleri	66
4.1.1.3.	Matematik Öğretmen Adaylarının (N=28) Ölçme ve Değerlendirme Okuryazarlık Seviyeleri	68
4.1.1.4.	Türkçe Öğretmen Adaylarının (N=43) Ölçme ve Değerlendirme Okuryazarlık Seviyeleri	70
4.1.2.	İkinci Araştırma Sorusu	72
4.1.2.1.	Yeterlilik Alanı 1	74
4.1.2.2.	Yeterlilik Alanı 2	75
4.1.2.3.	Yeterlilik Alanı 3	77
4.1.2.4.	Yeterlilik Alanı 4	79
4.1.2.5.	Yeterlilik Alanı 5	81
4.1.2.6.	Yeterlilik Alanı 6	83
4.1.2.7.	Yeterlilik Alanı 7	85
4.1.3.	Üçüncü Araştırma Sorusu	87
4.1.3.1.	Öğretmen Adaylarının Ölçme ve Değerlendirmeye Yönelik Tutum Ölçeğine İlişkin Sonuçları	89
4.1.4.	Dördüncü Araştırma Sorusu	100
5.	SONUÇLAR VE TARTIŞMA.....	101
5.1.	Sonuçlar	101
5.1.1.	Ölçme ve Değerlendirme Okuryazarlığı Envanteri ile İlgili Sonuçlar	102
5.1.1.1.	Yeterlilik Bazında Sonuçlar	102
5.1.2.	Ölçme ve Değerlendirmeye İlişkin Tutum Sonuçları	105
5.1.3.	Ölçme ve Değerlendirme Okuryazarlık Seviyesi ile Ölçme ve Değerlendirmeye İlişkin Tutumlar Arasındaki İlişki	105
6.	ÖNERİLER.....	106

7.	KAYNAKLAR	108
8.	EKLER	119
	ÖZGEÇMİŞ	139

ÖZET

Bu araştırmanın amacı ilköğretim öğretmen adaylarının (Fen Bilgisi, Sosyal Bilgiler, Türkçe ve Matematik) ölçme-değerlendirme okuryazarlığı ve ölçme-değerlendirmeye ilişkin tutumlarını belirlemektir. Araştırmaya, 2010-2011 eğitim öğretim yılı bahar döneminde Fırat Üniversitesi Eğitim Fakültesi İlköğretim Bölümünde öğrenim gören toplam 180 öğretmen adayı katılmıştır. Öğretmen adaylarının ölçme-değerlendirme okuryazarlıklarını belirlemek için, Plake ve Impara tarafından hazırlanan *Teacher Assessment Literacy Questionnaire* (1993) Ölçme ve Değerlendirme Okuryazarlığı Envanteri kullanılmıştır. Envanter literatüre uygun bir şekilde Türkçeye uyarlanmış ve geçerlilik sağlanmıştır. Sonuçları analiz etmek için, TAP (Test Analysis Programme) programı kullanılmıştır. Veriler KR-20 güvenilirlik analizi ve test maddeleri güçlük indeksleri kullanılarak analiz edilmiştir. Ayrıca SPSS 16,0 programıyla Tek Yönlü Varyans (ANOVA) analizi yapılmış, gruplar arasındaki korelasyona ise “Pearson Korelasyon Analizi” ile bakılmıştır. Öğretmen adaylarının ölçme-değerlendirmeye ilişkin tutumlarını belirlemek için, araştırmacı tarafından öğretmen adayları ile mülakatlar sonucunda oluşturulmuş “Ölçme ve Değerlendirmeye İlişkin Tutum Ölçeği” kullanılmıştır. Bu veriler ise, SPSS 16,0 programı ile analiz edilmiştir.

Araştırma sonuçlarına göre envanterin KR-20 güvenilirlik katsayısı 0,718, maddelerin ortalama güçlük değeri 0,556, ortalama ayırt edicilik değeri ise 0,374 olarak bulunmuştur. Bu değerler maddelerin psikometrik açıdan kullanılabilmesini göstermektedir. Bulgulara göre öğretmen adaylarının ölçme-değerlendirme okuryazarlık seviyeleri oldukça düşük olduğu görülmüştür. Ölçme-değerlendirme okuryazarlık seviyesinde en başarısız oldukları alan “ Ölçme ve Değerlendirme Sonuçlarını Paylaşma” yeterlilik alanı olmuştur. Öğretmen adayları arasında standartlar açısından önemli farklılıklar bulunmuştur. Öğretmen adaylarının tutumları dört faktör altında incelenmiş ve adayların tutumlarının yüksek olduğu görülmüştür. Başarı ile tutum arasında yapılan korelasyon ise tutum ve başarı arasında zayıf derecesinde de olsa doğrusal bir ilişkinin olduğunu göstermektedir.

Anahtar kelimeler: Ölçme ve Değerlendirme Okuryazarlığı, Ölçme ve Değerlendirmeye İlişkin Tutum, İlköğretim Öğretmen Adayları, Yeterlilik

SUMMARY

PRESERVICE ELEMENTARY TEACHERS ASSESSMENT LITERACY AND ATTITUDE TOWARDS ASSESSEMENT

The aim of this study is to determine the assessment literacy of preservice elementary school teachers (Science Teachers, Social Sciences Teacher, Turkish Teachers And Mathematics Teachers) and their attitude towards assessment. The research was participated in the academic year of 2010-2011, spring semester through 180 preservice elementary school teachers studing at the department of Elementary Education at the Faculty of Education at Fırat University. To determine the assessment literacy of teacher candidates, *Assessment Literacy Inventory* (1993), Measurement And Evaluation Literacy Inventory, prepared by Barbara S. Plake and James C. Impara, was used. The Inventory has been adapted to Turkish literature in a suitable manner. To analyze the results, the TAP (Test Analysis Program) program was used. In addition, the SPSS 16.0 program variance (ANOVA) analysis has been prepared. The correlation between the groups is examined with the "Pearson Correlation Analysis". To determine the attitudes of teacher candidates about measurement and evaluation, "Attitude Scale about Measurement and Evaluation which was prepared by the researcher as an overall of the interviews made with the teacher candidates is used. These data were analyzed with the SPSS 16.0 program.

According to the item analysis results, the instrument reliability (KR-20) was 0,718, mean item difficulty was 0,556 and item discrimination was 0,374. These values show that substances of the scale can be used in a psychometric aspect. According to the findings, teacher candidates were found to be considerably at a low level of assessment literacy. The level of assessment literacy that the teacher candidates fail most has proved to be the "Sharing of the Assessment and Evaluation Results". In terms of standards, there were significant differences between preservice elementary school teachers. Attitudes of the teacher candidates were examined under four factors and they were found to be high. The correlation between success and attitude shows that, though at a low level, there is a linear relationship between success and attitude.

Keywords: Assessment Literacy, Attitude towards Assessment, Preservice Elementary Teachers, Competency

TABLÖLAR LİSTESİ

	<u>Sayfa No</u>
Tablo 1.1. Ölçme ve Değerlendirme Okuryazarlığı	6
Tablo 2.1. Günümüzdeki ve önceki ölçme değerlendirme yaklaşımları (McMillian,1997).....	39
Tablo 2.2. Alternatif Değerlendirme Türleri (Reeves, 1996)	40
Tablo 2.3. Geleneksel ve Alternatif Değerlendirme Yöntemlerinin Karşılaştırılması.....	41
Tablo 2.4. Yapılandırılmış Grid.....	49
Tablo 3.1. Pilot Çalışma Örneklem	55
Tablo 3.2. Ana Çalışma Örneklem	55
Tablo 3.3. Ölçme ve Değerlendirme Okuryazarlığı Envanter İçeriği.....	56
Tablo 3.4. Araştırma Zaman Çizelgesi	58
Tablo 3.5. Ölçme ve Değerlendirmeye İlişkin Tutum Anketi İçeriği.....	59
Tablo 4.1. Envanter Maddelerinin Psikometrik Özellikleri.....	62
Tablo 4.2. Fen Bilgisi Öğretmen Adaylarının Değerlendirme Okuryazarlık Seviyeleri	64
Tablo 4.3. Sosyal Bilgiler Öğretmen Adaylarının Değerlendirme Okuryazarlık Seviyeleri	66
Tablo 4.4. Matematik Öğretmen Adaylarının Değerlendirme Okuryazarlık Seviyeleri	68
Tablo 4.5. Türkçe Öğretmen Adaylarının Değerlendirme Okuryazarlık Seviyeleri	70
Tablo 4.6. Tüm Grupların Ölçme ve Değerlendirme Başarı Puanları	73
Tablo 4.7. Yeterlik Alanı 1 (ANOVA).....	74
Tablo 4.8. Yeterlik Alanı 1 Çoklu Regresyon Analizi	75
Tablo 4.9. Yeterlik Alanı 2 (ANOVA).....	76
Tablo 4.10. Yeterlik Alanı 2 Çoklu Regresyon Analizi.....	77
Tablo 4.11. Yeterlik Alanı 3 (ANOVA).....	78
Tablo 4.12. Yeterlik Alanı 3 Çoklu Regresyon Analizi.....	79
Tablo 4.13. Yeterlik Alanı 4 (ANOVA).....	80
Tablo 4.14. Yeterlik Alanı 4 Çoklu Regresyon Analizi.....	81

Tablo 4.15. Yeterlik Alanı 5 (ANOVA)	82
Tablo 4.16. Yeterlik Alanı 5 Çoklu Regresyon Analizi.....	83
Tablo 4.17. Yeterlik Alanı 6 (ANOVA)	84
Tablo 4.18. Yeterlik Alanı 6 Çoklu Regresyon Analizi.....	85
Tablo 4.19. Yeterlik Alanı 7 (ANOVA)	86
Tablo 4.20. Yeterlik Alanı 7 Çoklu Regresyon Analizi.....	87
Tablo 4.21. Tutum Anketi Tüm Faktörler ve Çıkarılan Faktörler	88
Tablo 4.22. Özel İlgi Faktörü Toplam Grup Sonuçları.....	89
Tablo 4.23. Özel İlgi Faktörü (ANOVA)	91
Tablo 4.24. Özel İlgi Faktörü Tukey Testi Sonucu	92
Tablo 4.25. Endişe-Kaygı Faktörü Toplam Grup Sonuçları.....	92
Tablo 4.26. Endişe-Kaygı Faktörü (ANOVA).....	94
Tablo 4.27. Endişe-Kaygı Faktörü Tukey Testi Sonuçları	94
Tablo 4.28. Önem Faktörü Toplam Grup Sonuçları	95
Tablo 4.29. Önem Faktörü (ANOVA).....	96
Tablo 4.30. Önem Faktörü Tukey Testi Sonuçları	97
Tablo 4.31. Yeterlilik Faktörü Toplam Grup Sonuçları	97
Tablo 4.32. Yeterlilik Faktörü (ANOVA)	99
Tablo 4.33. Öğretmen Adaylarının Tüm Faktörler Bazında Toplam Sonuçları.....	99
Tablo 4.34. Pearson Korelasyon İlişkisi	100
Tablo 4.35. Tutum ve Başarı Puanları Pearson Korelasyonu	100
Tablo 5.1. Ölçme ve Değerlendirme Okuryazarlığı Toplam Sonuçlar.....	101

GRAFİKLER LİSTESİ

	<u>Sayfa No</u>
Grafik 4.1. Fen Bilgisi Öğretmen Adayları Başarı Grafiği	65
Grafik 4.2. Sosyal Bilgiler Öğretmen Adayları Başarı Grafiği	67
Grafik 4.3. Matematik Öğretmen Adayları Başarı Grafiği.....	69
Grafik 4.4. Türkçe Öğretmen Adayları Başarı Grafiği.....	71
Grafik 4.5. Yeterlilik alanı 1	74
Grafik 4.6. Yeterlilik alanı 2	76
Grafik 4.7. Yeterlilik alanı 3	78
Grafik 4.8. Yeterlilik alanı 4	80
Grafik 4.9. Yeterlilik alanı 5	82
Grafik 4.10. Yeterlilik alanı 6.....	84
Grafik 4.11. Yeterlilik alanı 7.....	86
Grafik 4.12. Ölçme ve Değerlendirmeye İlişkin Özel İlgi Faktörü Toplam Sonuçları Grafiği	90
Grafik 4.13. Özel İlgi Scree Plot Grafiği.....	91
Grafik 4.14. Ölçme ve Değerlendirmeye İlişkin Endişe-Kaygı Faktörü Toplam Sonuçları Grafiği	93
Grafik 4.15. Endişe-Kaygı Scree Plot Grafiği.....	93
Grafik 4.16. Ölçme ve Değerlendirmeye İlişkin Önem Faktörü Toplam Sonuçları Grafiği	95
Grafik 4.17. Önem Scree Plot Grafiği	96
Grafik 4.18. Ölçme ve Değerlendirmeye İlişkin Yeterlilik Faktörü Toplam Sonuçları Grafiği	98
Grafik 4.19. Yeterlilik Scree Plot Grafiği	98

1. GİRİŞ

Bu bölümde araştırmaya ait problem durumu, araştırmanın amacı, araştırmanın önemi, araştırma soruları, sayıtlılar, sınırlılıklar ve tanımlar yer almıştır.

1.1. Problem Durumu

Son yıllarda eğitim alanında yapılan bilimsel araştırmalar, öğrenme-öğretim modelleri, öğrenci-öğretmen ilişkileri ve ölçme-değerlendirme yöntemlerinin, öğrencilerin fen bilimlerinde, sosyal bilimlerde ve diğer alanlardaki başarısını etkileyen faktörlerin başında geldiğini göstermektedir. Bu çalışmalarda ölçme ve değerlendirme uygulamalarının öğretimin ayrılmaz bir parçası olduğu önemle vurgulanmıştır. Buna rağmen, yükseköğretimde geleneksel öğretim ve değerlendirme yöntemleri hala kullanılmakta, ölçme ve değerlendirmeye yeterince önem verilmemektedir (Mertler, 2003). Öğretmenlerin en çok sorun yaşadıkları aynı zamanda da yeni yaklaşımlardan habersiz oldukları alan ölçme-değerlendirmedir.

Yurtdışında yapılan araştırmaların çoğunda ölçme değerlendirmenin öğretim kadar önemli olduğu ve bu alanla ilgili çalışmaların yapılması gerektiği vurgulanmaktadır. Ölçme ve değerlendirme, öğrencilerin neleri bildiği, bildiklerini nasıl yorumladığı ve elde ettiği değerlendirme sonuçlarıyla öğrencide ne tür eksiklikler olduğunu anlamaya ve kavramaya çalışan uygulamalardır (Black, 1998). Birçok bilimsel dergi, sadece ölçme değerlendirme üzerine yayınları kabul etmektedirler. Bu nedenle ölçme-değerlendirme son yıllarda önemli bir çalışma alanı olmuştur.

Eğitimde ve nitelikli öğretmen yetiştirmede değerlendirme bilgisi son yıllarda üzerinde çalışılan en önemli konulardan biridir. Uluslararası eğitim literatürü değerlendirme bilgisinin öğretmen yetiştirmede en önemli bilgilerden biri olduğunu, ayrıca öğretme bilgisi kadar önemli olduğunu göstermektedir. Ölçme ve değerlendirme uygulamaları uygun bir şekilde kullanıldığında, öğrencinin öğrenme performansı da önemli biçimde artacaktır (Mertler ve Campbell, 2005). Yapılan araştırmalar da öğretmenlerin bu konuda yeterli olmadıklarını ve birçok eksikliğin olduğunu göstermektedir. Hargreaves (2009) ölçme ve değerlendirme yöntem tekniklerinin öğrencilerin öğrenme yaklaşımlarının belirlenmesinde önemli bir rol oynadığını ve öğretmenlerin bu yöntemleri uygun bir

şekilde kullandığında derin öğrenme yaklaşımlarının kullanılmasına olanak tanıyacağını belirtmiştir. Bu nedenle öğrenmenin niteliğinin artırılması ve öğrencilerin günümüzün gereği olan bilgi ve becerilerin geliştirilmesinde öğretmenlerin ölçme ve değerlendirme okuryazarlık seviyeleri ve ölçme değerlendirmeyi uygun bir şekilde yapabilmelerinin önemi ortaya çıkmaktadır (Daniel ve King, 1998).

Eğitim literatürü son 30 yılda öğrenci merkezli öğrenme ve öğrenenin bağımsız olduğu öğretim modelleriyle ilgili konulara büyük bir ilgi duymaya başlamıştır (Nunan, 1988). Geleneksel eğitim anlayışının yerine yeni arayışlar başlamıştır ve son gelişmeler eğitimde ve öğretme-öğrenme sürecinde öğretmenin tek otorite olduğu geleneksel anlayışın yerine, öğrencinin ihtiyaçlarına önem veren yeni bir anlayışı desteklemektedir. Eğitimde öğretmenin bakış açısı yerine öğrencinin bakış açısının ne olduğuna odaklanılmaya başlanmıştır. Öğrenci merkezli yapılandırmacı yaklaşıma göre, öğretme ve öğrenmenin etkililiği öğrencilerin öğrenme sürecine, bu süreç içerisinde kendilerine olan bakış açısına bağlıdır (Haris, 1997; McNamara ve Deane, 1995; Nunan, 1988). Öğrenci merkezli yaklaşım yalnızca öğrencinin ihtiyaçlarını göz önünde bulundurmak değil, aynı zamanda öğrencilere ihtiyaçlarının, amaçlarının ve ilerlemelerinin farkında olabilmelerini sağlayacak öğrenme ortamlarını da sağlamaktır.

Son yıllarda ölçme-değerlendirmede tamamlayıcı yöntemler önem kazanmaya başlamış ve eğitimin her seviyesinde otantik ölçme değerlendirme yöntemleri kullanılmaya başlamıştır. Araştırmalar, ölçme değerlendirme yöntemlerinin öğrencilerin öğrenme yaklaşımlarını ve yöntemlerini etkilediğini ortaya koymuştur (Beckwith, 1991). Sınav kültürü barındıran, öğrenme ve öğretmeden bağımsız olan, daha çok ezbere dayalı bilgileri ölçen, öğrencinin pasif olduğu, belirli zaman aralıklarında yapılan, öğrenciler arası etkileşimin nerdeyse sıfır olduğu geleneksel değerlendirmeler çağdaş öğrenme yöntemlerinin uygulanmasına olanak tanımamaktadır. Boud (1995) eğitimde gereksinim duyulan yeni ölçme değerlendirmenin özelliklerini aşağıdaki gibi belirtmektedir.

Ölçme değerlendirme etkinlikleri;

- Bireysel performansla sınırlı olmamalıdır.
- Programla ilintili olmalı; fakat programda öğretildiği şekilde sınırlı olmamalıdır.
- Problemler birden fazla çözüm yolu içermeli ve bir sorunun birden fazla kabul edilebilir cevabı olmalıdır.
- Öğrencilerin hayatta karşılaşılabilecekleri durumları içermeli sadece okullarla sınırlı olmamalıdır.

Öğretmenlik mesleği, öğretmenin önemi, öğretmen yetiştirme ve öğretmenlerin rolleri hemen hemen tüm toplumlarda güncelliğini koruyan kavramlardır. Birçok gelişmiş ve gelişmekte olan ülke, eğitim sistemlerini ve bu sistem içinde öğretmenlik ve öğretmen yetiştirme konusunu gündemlerinde ön sıralara çıkarmışlardır. Bu değişikliğin temel gerekçesi, 21. yüzyılda eğitimin ve öğretmenliğin toplum yaşamı üzerinde etkisinin giderek artma beklentisidir (Atanur, 2001).

Millî Eğitim Bakanlığının 2002 yılında öğretmen yeterlikleri ile ilgili yaptığı çalışmada, Millî Eğitim Bakanlığınca saptanan öğretmen yeterlikleri “eğitme-öğretme yeterlikleri”, “genel kültür bilgi ve becerileri” ve “özel alan bilgi ve becerileri” alt başlıklarından oluşmaktadır. Bu yeterlik alanlarından “eğitme ve öğretme yeterlikleri” 14 alt bölüm ve 206 yeterlik maddesinden oluşmaktadır. Öğrenci başarısını ölçme ve değerlendirmeye yönelik öğretmenlerden beklenen yeterlikler aşağıdaki gibidir.

(Millî Eğitim Bakanlığı [MEB], 2002):

- Amaca uygun ölçme ve değerlendirme yöntemini belirleme,
- Öğrencinin performans ölçütlerini belirleme,
- Yazılı yoklama yapma,
- Sözlü yoklama yapma,
- Doğru-yanlış türü test yapma,
- Eşleştirme türü test yapma,
- Tamamlama türü test yapma,
- Çoktan seçmeli test yapma,
- Hazırladığı ölçme araçlarının geçerlik ve güvenirliklerini saptama,
- Ödev ve proje değerlendirme,
- Öğrenci çalışma dosyasını değerlendirme,
- Değerlendirme sonuçlarını okul yönetimine bildirme,
- Öğrenci velisine geribildirim sağlama,
- Mutlak değerlendirme yapma,
- Bağıl değerlendirme yapma,

Bu bağlamda öğretmenlerin değerlendirme ile ilgili bilgileri, uygulamaları, değerlendirmeyi teknoloji destekli yapmaları, uygun değerlendirme metodu seçmeleri ve değerlendirme yolları geliştirmeleri, eğitimsel kararlar almaları, değerlendirmede güvenirlik ve geçerlik ve değerlendirmede etik kurallar ve değerlendirme ile ilgili temel bilgilere sahip olmaları “Değerlendirme Okuryazarlığı” adı altında toplanmıştır.

Değerlendirme okuryazarlığı, eğitimsel kararlarda, uygun değerlendirme sonuçlarını kullanmak, seçmek, yorumlanmak, tasarlamak için bir öğretmenin sahip olması gereken yetenek olarak tanımlanabilir. Stiggins (1995) ölçme ve değerlendirme okuryazarı olan öğretmenlerin neyi değerlendireceklerini, öğrencinin neyi öğrendiğini en iyi şekilde nasıl değerlendireceğini, ölçme değerlendirmeyi olumsuz etkileyecek faktörlerin neler olduğunu, bunların önüne nasıl geçileceği konularında yeterli bilgi, beceri ve deneyime sahip olmaları gerektiğini belirtmiştir. Ülkemizde ve diğer ülkelerdeki öğretmen yetiştiren kurumlarda, birçok öğretmen ve öğretmen adayı ölçme ve değerlendirme derslerini ve çeşitli kurslarını alırlar. Fakat çok az sayıda öğretmen veya öğretmen adayı, çeşitli amaçlar için değerlendirme sonuçlarını kullanma, iyi bir şekilde değerlendirme yapma, değerlendirme sonuçlarını yorumlamada gerçek sınıflarda başarılı olurlar. Schfer ve Lissitz (1987) yaptıkları araştırmada genelde, öğretmenlerin öğrenci çıktılarını değerlendirmede iyi bir şekilde hazırlık yapmadıklarını ortaya koymuştur. Eğitimsel ölçme değerlendirmelerde, yapılan birçok araştırmada öğretmenlerin eksik oldukları ve öğrenci başarısını gerçek anlamda değerlendirmede farklı yolları kullanmada eksikliklerinin olduğu ortaya çıkmıştır.

1.2. Araştırmanın Amacı

Bu çalışmanın amacı ilköğretim öğretmen adaylarının ölçme değerlendirme okuryazarlık seviyelerini ve ölçme değerlendirmeye ilişkin tutumlarını belirlemektir.

Eğitim literatüründe değerlendirme okuryazarlığı, Amerikalı araştırmacılar tarafından 6 yıl süren ve önemli kuruluşların katılımıyla 7 yeterlilik alanı belirlenmiştir.

- Uygun ölçme ve değerlendirme yöntemlerini seçme
- Uygun ölçme ve değerlendirme yöntemlerini geliştirme
- Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama
- Eğitimsel kararlar alırken ölçme ve değerlendirme sonuçlarını kullanma
- Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma
- Ölçme ve değerlendirme sonuçlarını paylaşabilme
- Ölçme ve değerlendirmede kanun dışı ve ahlaki olmayan durumları bilme

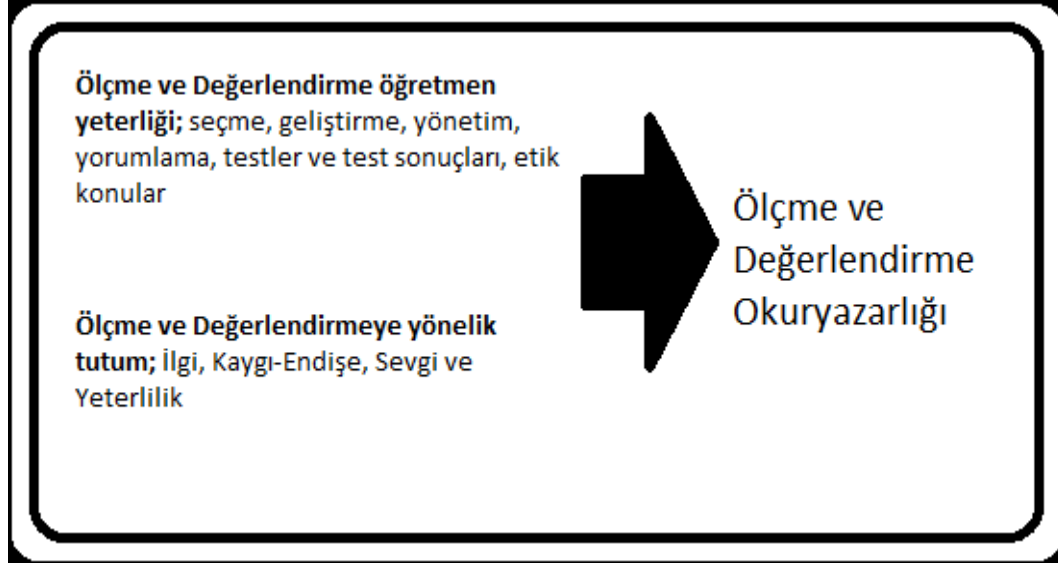
Yukarıda belirtilen yeterlilik alanları göz önünde bulundurularak öğretmenlerin okuryazarlık seviyeleri belirlendiği gibi bunlar dikkate alınmadan da öğretmen

yeterlilikleri belirlenmiştir. Bu çalışmaların çoğunda öğretmen adaylarının görüş, düşünce ve tutumları göz önünde bulundurularak belirlenmeye çalışılmıştır. Fakat elde edilen verilerin öğretmen veya öğretmen adaylarının yeterliliklerini yansıtmamakta veya yüzeysel olarak yansıtmaktadır. Tutum ve düşüncenin dışında, ölçme-değerlendirme okuryazarlığı konu alanında belli bir seviyede bilgi ve beceriye sahip olmak anlamına gelir. Yeterlilik alanlarının belirlenmesinde kullanılan ölçme araçları (anket, mülakat vb.) öğretmen yeterliliğini belirlemede eksik ve sınırlı kalmaktadır. Bu nedenle ölçme ve değerlendirme ile ilgili konu alanını kapsayan ölçme araçları kullanılmalı ve geliştirilmelidir.

Bu amaçla çalışmamızda Barbara S. Plake ve James C. Impara tarafından hazırlanan *Teacher Assessment Literacy Questionnaire* (1993) ölçme ve değerlendirme okuryazarlığı envanteri kullanılmıştır. Çalışmaya farklı bölümlerden 180 son sınıf öğretmen adayı katılmıştır.

Çalışmanın bir diğer amacı da öğretmen adaylarının ölçme değerlendirmeye ilişkin tutumlarının belirlenmesidir. Öğretmen adaylarının ölçme ve değerlendirme yeterlilik alanları ile ölçme değerlendirmeye ilişkin tutumları karşılaştırılacaktır. Öğretmen adaylarının ölçme değerlendirmeye ilişkin algıları araştırmacı tarafından hazırlanan "*Ölçme- Değerlendirmeye İlişkin Tutum Ölçeği*" ölçeğiyle araştırılacaktır. Bu ölçekte, öğretmen adaylarının ölçme değerlendirmeye ilişkin tutumları; özel ilgi, sevgi, endişe-kaygı ve yeterlilik alt boyutlarında ele alınacaktır.

Tablo 1.1. Ölçme ve Değerlendirme Okuryazarlığı



Değerlendirme okuryazarlığı kavramıyla ilgili yapılan çalışmalarda, yalnızca ölçme ve değerlendirmede öğretmen yeterliliklerine bakılmış fakat, yapılan çalışmaların çok azında öğretmen adaylarının ölçme ve değerlendirmeye ilişkin tutumları belirlendiği gözlemlenmiştir. Ölçme ve değerlendirmeyi iyi bir şekilde yapabilmek için öğretmen adaylarının ölçme ve değerlendirme konusunda algılarının yüksek olması gerekmektedir. Yapılan birçok araştırmada öğretmen adaylarının ölçme ve değerlendirmeye yönelik tutumları, ölçme ve değerlendirme yöntemlerini seçmelerine veya tercihlerine göre yapılmıştır. Oysaki tutum bir insanın herhangi bir konuya karşı göstermiş olduğu bakış açısidir. Örneğin ölçme ve değerlendirmede alternatif yöntemleri seçen bir öğretmen için bu durum tutum değil bir davranış veya algıdır. Tutum bir olay veya olguyu sevip sevmemek, o olguya ilgili endişe veya kaygı duymamak olarak tanımlanmaktadır. Bu araştırmada da öğretmen adaylarının ölçme ve değerlendirme konusundaki tutumları yukarıda belirtilen faktörler ışığında belirlenecektir.

1.3. Araştırmanın Önemi

Bir eğitim durumu ve sistemi hakkındaki gerçeklerin belirlenebilmesi için, öncelikle eğitim sisteminin, ölçme ve değerlendirme süreçlerine bakılması gerekmektedir (Rowntree, 1987). Yapılan birçok araştırmada, ölçme ve değerlendirme uygulamalarının, öğrencilerin

eksikliklerinin belirlenmesi ve eksiklikler göz önünde bulundurularak öğretimin yeniden planlanması amacıyla yapıldığını vurgulamaktadır.

Değerlendirme okuryazarlığı uzun yıllar Amerikalı araştırmacılar tarafından çalışılmıştır. Fakat ülkemizde bu konuyla ilgili herhangi bir çalışmaya rastlanmamıştır. Bu anlamda araştırmamız ülkemiz eğitim literatürüne bu kavramın girmesi ve ileride bu kavramın geliştirilmesi aynı zamanda Türk eğitim sistemine değerlendirme okuryazarlık standartlarının getirilmesi konusunda önemli bir araştırma olacaktır.

Değerlendirme okuryazarlığı, 1990 yılında Amerikan Ölçme ve Değerlendirme Birliği (NCME) tarafından geliştirilen bir kavram olmasına rağmen, literatürde ilköğretim öğretmen adaylarıyla yapılmış bir çalışmaya rastlanmamıştır. Son yıllarda ülkemizde ilköğretim müfredatında değerlendirme üzerine birçok değişiklik olmasına rağmen ne değerlendirme okuryazarlığının belirlenmesi ne de geliştirilmesi üzerine çalışmaya rastlanmıştır. Bu anlamda ülkemiz eğitim literatürüne de çok büyük katkı sağlayacaktır.

1.4. Araştırma Soruları

Değerlendirme okuryazarlığı ve ölçme değerlendirmeye ilişkin tutum arasındaki ilişki birçok araştırmada ele alınmış, araştırılmıştır. Fakat ölçme değerlendirmeye ilişkin öğretmen adaylarının tutumu, onların ölçme değerlendirmeye olan bakış açıları yapılandırmacı ölçme değerlendirme metodu kullanma veya geleneksel ölçme değerlendirme metotlarını kullanma olarak tanımlanmıştır. Fakat tutum faktörü bu araştırmada öğretmen adaylarının sınıf içerisinde ölçme değerlendirmeyi yapmadaki davranışları olarak ele alınmıştır. Öğretmen adayının ölçme değerlendirmeye ne kadar önem verdiği, ölçme-değerlendirmeyi ne kadar yapmak istediği, ölçme değerlendirmeyi yapmada kendini ne kadar hazır ve yeterli hissettiği ve bu tutumlarının ölçme değerlendirme okur-yazarlığı kavramıyla ne derecede ilişkili olduğu araştırılmıştır.

Araştırma soruları şunlardır;

1. İlköğretim öğretmen adaylarının (fen bilgisi, sınıf, sosyal bilgiler ve matematik öğretmen adayları) değerlendirme okur-yazarlık seviyeleri nedir?
2. İlköğretim öğretmen adaylarının (fen bilgisi, sınıf, sosyal bilgiler ve matematik öğretmen adayları) değerlendirme okur-yazarlık seviyeleri arasında anlamlı bir farklılık var mıdır?

3. İlköğretim öğretmen adaylarının (fen bilgisi, sınıf, sosyal bilgiler ve matematik öğretmen adayları) ölçme-değerlendirmeye ilişkin tutumları nedir?
4. İlköğretim öğretmen adaylarının ölçme-değerlendirme okur-yazarlığı ve ölçme değerlendirmeyle ilişkin tutumları arasında bir ilişki var mıdır?

1.5. Sayıtlar

- Araştırmanın örnekleme, evreni temsil eder niteliktedir. Anket ve döküm yoluyla toplanan veriler Türkiye genelinde üniversitelerdeki ölçme ve değerlendirme uygulamalarını yansıtabilecek niteliktedir.
- İngilizceden çevrilen psikometrik ölçme aracı gerekli prosedür göz önünde bulundurularak çevrilmiş ve ülkemiz kültürel öğeleri dikkate alınmıştır.
- Öğretmen adayları anket ve döküm sorularını içtenlikle cevaplamışlardır.
- Öğretmen adayları anket ve döküm sorularının uygulanması sırasında aynı derecede güdülenmişlerdir.

1.6. Sınırlılıklar

- Araştırma 2010-2011 eğitim-öğretim yılı ile sınırlıdır.
- Araştırma yükseköğretim kurumuna bağlı üniversiteler ile sınırlıdır.
- Araştırma Fırat Üniversitesi ile sınırlıdır.
- Araştırma üniversitelerin ilköğretim bölümlerinde öğrenim gören öğretmen adaylarıyla sınırlıdır.
- Kullanılan ölçme aracı Amerika'da yapıldığından ülkemiz kültürel öğelerinin ve eğitim sisteminin farklılığı nedeniyle sınırlılık sahibidir.

1.7. Tanımlar

Eğitim: Bireyde öğrenme yoluyla kasıtlı ve istendik davranış oluşturma sürecidir (Çelik, 2000, s.137).

Öğretim: Yaşantı ürünü kalıcı izli davranış değişikliklerini kılavuzlama faaliyetleridir (Çelik, 2000, s.137).

Eđitim Programı: Okul ierisinde veya dıřında mili eđitimin ve okulun amalarını etkili olarak gerekleřtirmek amacı ile hazırlanan ierik ve etkinliklerin uygun yntem ve tekniklerle geliřtirilmesine ynelik koordine abaların tmdr (Varıř, 1996).

lme: Geerli grgl yollarla test edilebilecek kurallar erevesinde nesnelere belirli zelliklere sahip oluř derecelerine gre sayılar ve semboller vermektir (Karip ve diđerleri, 2007, s.3).

Deđerlendirme: Deđerlendirme, elde edilen bilgilere bir anlam verme, onları belli amalara elveriřlilik, belli kořullar karřılama, belli anlamlarda olup olmama vb. bakımdan yorumlama iřlemidir (zelik, 1992, s.229).

Yapılandırıcılık: Yapılandırıcılık bina etme ve oluřturmanın bir sreci olarak, bilginin kazanılmasına dayanan bir anlayıř olarak tanımlanabilir (Fox, 2001).

z Yeterlik: Kiřinin bir iři yapmak iin gerekli becerilere sahip olduđu konusundaki inancıdır (Bandura, 1997).

lme ve Deđerlendirme Okuryazarlıđı: lme ve deđerlendirme kavram ve ilkelerinin anlaşılması olarak tanımlanmaktadır (Popham, 2004; Stiggins, 2002; Akt: Volante ve Fazio, 2007).

Deđerlendirme Okuryazarlıđı: Eđitimde deđerlendirme standartlarına dayanır. Deđerlendirme okuryazarlıđı řyle tanımlanır; “deđerlendirme uygulamalarının basit ilkelerine sahip olma”, “deđerlendirme tekniklerini ve metotlarını kullanma ve geliřtirme”, “deđerlendirmenin kalitesini standartlarla iliřkilendirme” (Paterno, 2001) olarak tanımlanmaktadır.

Gvenirlik: Herhangi bir lme ara veya ynteminin ne derece tutarlı lm yapabildiđidir.

Geerlik: Bir lme aracının lmeyi amaladıđı zelliđi, bařka herhangi bir zellikle karıřtırmadan, dođru olarak lebilme derecesidir (Tekin, 1996).

2. LİTERATÜR VE KAVRAMSAL ÇERÇEVE

Bu bölümde ölçme ve değerlendirme okuryazarlık ve yeterlilikleri ile ilgili ülkemizde ve yurt dışında yapılan araştırmalar ve kavramsal çerçeve yer almaktadır.

2.1. LİTERATÜR

2.1.1. Ülkemizde Yapılan Araştırmalar

Ülkemizde değerlendirme okuryazarlığı ile ilgili çok fazla çalışma yoktur. Bu nedenle öğretmen yeterliliği ve ölçme değerlendirme okuryazarlığına daha yakın olan literatürlerden bahsedeceğiz. Ölçme ve değerlendirme okuryazarlığı ile ilgili ülkemizde tek bir çalışma vardır. Bütüner ve diğ. (2010) yaptıkları araştırmada ölçme ve değerlendirme okuryazarlığı envanterini Türkçeye uyarlamış ve güvenilirlik ve geçerlilikleriyle ilgili çeşitli analizler yapmıştır. Envanterin ilk uygulamasına 80 ikinci uygulamasında ise 180 öğretmen adayına katılmıştır. Araştırmada madde analizi kullanarak ölçeğin, KR-20 güvenilirlik değerine bakılmıştır. Sonuç olarak KR-20 güvenilirlik değerini 0,74 olarak bulmuştur. Envanterdeki 7 yeterlilik alanından 1 tanesi çıkarılmıştır.

Türkiye’de lise ve dengi okullarda öğrenci gelişiminin değerlendirilmesinde Milli Eğitim Bakanlığınca öngörülen esaslarla, uygulamadaki durumu, değerlendirmede kabul edilen temel felsefe ve ilkeler bakımından karşılaştırmalı olarak inceleyen ilk kapsamlı araştırma, 1961 yılında Uysal tarafından yapılmıştır. Bu araştırmada şu bulgular elde edilmiştir: Öğretmenler dersin amaçlarını ölçülebilir ve gözlenebilir davranışlara dönüştürememektedirler. Öğrenci başarısının değerlendirilmesinde yazılı ve sözlü sınavlar kullanılmakta, öğrencinin değerlendirilmesinde bireysel farklılıklar dikkate alınmamakta, öğrencilere değerlendirme sonrasında iyi bir geri bildirim verilmemektedir.

Ortaöğretimdeki alan öğretmenlerinin öğrenci başarısını ölçme ve değerlendirme ile ilgili görüşlerinin değerlendirildiği bir araştırmada, öğretmenlerin, öğrenci başarısını değerlendirmede ölçme araçlarından en çok yazılı sınavları kullandıkları belirlenmiştir. Öğretmenler, kısa cevap gerektiren testleri, çoktan seçmeli testleri, sözlü yoklamaları ve doğru yanlış testlerini ara sıra kullandıklarını ifade etmişlerdir. Yabancı Dil Bölümünden mezun olan öğretmenler diğer bölümlere göre kısa cevap gerektiren testleri daha sık

kullanmaktadırlar. Çoktan seçmeli testleri ise erkek öğretmenler, bayan öğretmenlere göre daha sık kullanmaktadırlar (Kaynak, 2000).

Milli Eğitim Bakanlığı Eğitimi Araştırma ve Geliştirme Dairesi tarafından yapılan bir çalışmada (1999), “Öğretmenler değerlendirme sonuçlarını öğrencilerin başarısını arttırmada kullanabiliyor mu?” sorusuna, öğretmenlerin %79,6’sı kendilerinin değerlendirme sonuçlarını öğrenci başarısını arttırmada kullandıklarını belirtmişlerdir. Yöneticilerin ise % 41,1’i orta, % 14’ü az derecede, müfettişlerin % 51,1’i orta, % 40,9’u da yetersiz derecede kullandıklarını ifade etmişlerdir. Öğretmenler değerlendirme sonuçlarının kullanımı ile ilgili değerlendirildiklerinde genellikle yetersiz bulunmakta ancak öğretmenler kendilerini değerlendirdiklerinde yeterli olduklarını düşünmektedir.

Öztürk (1988), yaptığı çalışmada, lise sosyal bilimler dersleri öğretmenlerinin başarı testi hazırlamadaki yeterliklerini incelemiştir. Çalışmada, öğretmenlerin eğitim hedeflerini istenilen niteliklere uygun olarak hazırlamada ve ifade etmede, öğrenci başarısının ölçülmesinde kullandıkları başarı testlerinin geçerliğini ve güvenilirliğini sağlamada ve başarı testlerinin puanlamasında kullandıkları yanıt anahtarını istenilen niteliklere uygun olarak hazırlama boyutlarında yeterliklerini belirlemeye çalışmıştır. Elde edilen veriler sonucunda öğretmenlerin, eğitim hedeflerini istenilen niteliklere uygun olarak hazırlamada ve ifade etmede yetersiz oldukları ortaya çıkmıştır. Öztürk, öğrenci başarısının ölçülmesinde kullandıkları başarı testlerinin geçerliğini sağlamada “az yeterlikte”, güvenilirliğini sağlamada ise “orta yeterlikte” beceriye sahip olduklarını, ancak öğretmenlerin sahip oldukları bu beceri seviyelerinin, çalışmada yeterlik ölçüsü olarak kabul edilen puanın altında olması nedeni ile öğretmenlerin yetersiz olduklarını belirtmiştir. Aynı şekilde başarı testlerinin puanlamasında kullandıkları yanıt anahtarının istenilen niteliklere uygun olarak hazırlanmasında “az yeterli” bir beceriye sahip oldukları, bu beceri seviyesinin çalışmada yeterlik ölçüsü kabul edilen puanın altında olması nedeniyle “yetersiz” olduklarının görüldüğünü belirtmiştir. (95)

Çakan (2004) tarafından, öğretmenlerin ölçme aracı kullanımlarına ilişkin olarak yapılan bir başka çalışmada ise, öğretmenlerin % 56, 2’sinin çoktan seçmeli maddeleri, %48’inin yazılı yoklamaları ve % 46,1’inin kısa cevaplı maddeleri kullandıkları belirtmiştir. Veriler öğretmenlerin çoğunluğunun, ölçme ve değerlendirme konusunda kendilerini yetersiz ya da eksik bulduklarını göstermiştir.

Gökçe (1999), ilköğretim öğretmenlerinin kendi algılarına dayalı olarak yeterliklerini araştırdığı çalışmasında, öğretmenlerin yeterliklerini 10 boyutta incelemiş ve bu

boyutlardan birini de ölçme ve değerlendirme yeterlikleri olarak belirlemiştir. Bununla ilgili bulgularında göze çarpan, öğretmenlerin öğrenci başarısını değerlendirme alanı ve bu alanda öngörülen yeterliğe yeterince sahip olmadıklarını göstermesidir. En düşük düzeyde sahip olunan ya da yeterince sahip olunmayan yeterlik, “geçerli, güvenilir ve anlaşılır ölçme ve değerlendirme araçları geliştirme”; en yüksek düzeyde sahip olunan yeterlik ise, “öğrenci başarısını değerlendirmede objektif olma” olarak bulunmuştur. Bu araştırmada yeterlik düzeyinde görev ve cinsiyet bakımından farklılaşma olmadığı buna karşılık görev yapılan okul türü, öğretmenlik sertifikasına sahip olma gibi etkenler bakımından farklılık olduğu bulunmuştur. Ayrıca, özel okullarda görev yapan öğretmenlerin, resmi okullarda görev yapan öğretmenlerden; sertifikası olan öğretmenlerin, sertifikası olmayan öğretmenlerden daha yeterli olduğu ortaya çıkmıştır (Aktaran Arık, 2006).

2005'ten sonra değişen Hayat Bilgisi ve Sosyal Bilgiler Programı'na ilişkin yapılan araştırmalarda ise, öğretmenlerin uygulamada en zorlandıkları kısmın ölçme ve değerlendirme olduğu, değerlendirme formlarının mevcut koşullara uygun olmaması ya da mevcut koşulların daha önceden düzeltilmesi gerektiği ve bu ölçeklerin değerlendirilmesinin çok uzun zaman almasının, öğretmenleri olumsuz etkilediği belirtilmektedir. Değerlendirme sürecinde, birçok değerlendirme etkinliğinin bir arada kullanılmasının olumlu gibi görünmesine karşın, öğretmenlerin bu teknikler konusunda yeterli donanımına sahip olmaması ve sınıfların kalabalık olması, özellikle bireysel değerlendirme tekniklerinin kullanılmasını engellediği saptanmıştır. Ayrıca ilköğretim Hayat Bilgisi Programı'nın en zayıf halkasının, ölçme ve değerlendirme uygulamaları olduğu saptanmıştır (Aykaç ve Başar, 2005; Küçükahmet, 2005; Özden, 2005).

Yeni ilköğretim programlarının uygulanmasına ilişkin sınıf öğretmenlerinin hazır bulunuşluk düzeylerinin ve eğitim gereksinimlerinin belirlenmesine yönelik araştırmalarda, öğretmenlerin, yenilenen programın ölçme ve değerlendirme boyutuna ilişkin olarak hizmet içi eğitime ihtiyaçları olduğu belirlenmiştir. Gözlem, çalışma dosyası, tartışma, deneyler, projeler, çalışma kâğıtları, öğrenci ürün dosyası ve performans değerlendirme konularında yetersiz oldukları ve eğitim almaları gerektiği vurgulanmaktadır. Öğretmenlerin kendilerini ölçme ve değerlendirme alanında yetersiz gördükleri saptanmıştır (Gözütok, Ergün ve Karacaoğlu, 2005; Yaşar, Gültekin, Türkan, Yıldız ve Girmen, 2005).

Gümüş (1981), “Ölçme ve Değerlendirme Dersi Öğretmenlerinin Öğrenci Başarısını Değerlendirme Davranışına Katkısı” adlı araştırmasında meslek öncesinde ölçme ve

değerlendirme bilgi ve becerilerini kazanmış ve kazanmamış öğretmenlerin, öğrenci başarısını değerlendirmelerine ilişkin uygulamalarını karşılaştırarak, öğretmen yetiştiren kurumların bir kısmında verilmekte olan ölçme ve değerlendirme dersinin işlevselliğini saptamayı amaçlamıştır. Araştırma sonucunda, öğretmenlerin öğrenci başarısını değerlendirmelerindeki davranışları arasında, ölçme ve değerlendirme dersi görenlerin daha olumlu tutum sergilediklerini belirlemiştir. Hizmet süresi arttıkça iki grubun öğrenci başarısının ölçülmesi ve değerlendirilmesine ilişkin uygulamaları arasındaki fark azalmakta, ancak yine de ortadan kalkmamaktadır.

Ulutaş (2003) tarafından öğretmenlerin ölçme ve değerlendirme alanı ile ilgili yeterliklerinin, bir alan bilgisi testi ile saptandığı araştırmada, öğretmenlerin testteki başarı ortalamasının yüz puan üzerinden 46 olduğu bulunmuştur. Bu araştırmanın bulgularına göre, öğretmenler ölçme ve değerlendirme ilkelerini büyük ölçüde uygulamakta olduklarını ifade etmektedirler. Araştırmanın temel bulgusu, öğretmenlerin algılarına dayalı yeterliklerle, alan bilgisi testi ile saptanan yeterliklerin arasında düşük düzeyde bir ilişkinin olmasıdır.

Doğan (2005), yaptığı araştırmayla fen bilgisi öğretmenlerinin derslerinde ölçme değerlendirmeyi nasıl yaptıklarını, sadece bilgi düzeyini ölçen soruları değil, duyuşsal ve devinişsel alanları da ölçen soruları sorup sormadıklarını, uygulamaya yer verip vermediklerini, bu konudaki eksikliklerini ve yanlışlıklarını tespit etmeyi amaçlamaktadır. Araştırmanın örneklemini Van ili Merkez ilçesindeki 41 ilköğretim okulunda görev yapan, 50 fen bilgisi öğretmeni oluşturmuştur. Öğretmenlere anket dağıtılarak, öğretmenlerin değerlendirme etkinlikleri üzerine görüşleri alınmıştır. Araştırmadan elde edilen sonuçlara göre; alternatif değerlendirme yaklaşımları hakkında öğretmenlerin % 20' sinin “hiç” bilgileri olmadığı, % 50 sinin “biraz” bilgisi olduğu, “tamamen” bilgisi olan öğretmenlerin olmadığı sonucu ortaya çıkmıştır. Aynı araştırmada öğretmenlere geleneksel değerlendirme yaklaşımlarının yanında bu yaklaşımları uygulayıp uygulamadıkları sorulduğunda % 44 ünün “az” uyguladığı görülmüştür.

Aydın (2001) ölçme ve değerlendirme bilgi, becerileri ve deneyimlerini almış ve almamış öğretmenlerin, öğrenci başarısını ölçme ve değerlendirmeye ilişkin uygulamalarının karşılaştırıldığı araştırmasında da benzer bir sonuç elde etmiştir. Öğretmenlerin ölçme ve değerlendirme davranışları arasında, değerlendirme dersi alanların lehine bir farklılık olduğu saptanmıştır.

Eyitmiş (2007) liselerde görev yapan alan öğretmenlerinin ölçme değerlendirme tekniklerini ne derece etkin kullandıklarını ve uygulama konusundaki görüşlerini belirlemek amacı ile bir araştırma yapmıştır. Araştırma sonucunda, üniversitede ve öğretmen yetiştiren kurumlarda öğretmen adaylarına yeterli derecede ölçme ve değerlendirme bilgisinin verilmediği, ölçme ve değerlendirme ile ilgili derslerin teorik ve uygulamalı olarak verilmesi istendiği, ölçme araçlarından yazılı sınavların her zaman, sözlü sınavların, çoktan seçmeli, doğru yanlış ve kısa cevap gerektiren testlerin ise bazen kullandıkları görüldü. Bu sonuçlar doğrultusunda, öğretmen yetiştiren öğretim kurumlarında ölçme değerlendirme alanındaki derslerin ağırlıklı olarak okutulması, öğretmenlik uygulama çalışmalarına ağırlık verilmesi, öğretmenlerin bu konuda eksikliklerinin giderilmesi, gelişmeleri takip etmelerine yönelik olarak ölçme ve değerlendirme konusunda MEB tarafından zorunlu kurslar açılması ve öğretmenlerin ölçme değerlendirme birimlerinden yardım almaları konularında önerilerde bulunmuştur.

Okur (2008) araştırmasında fen ve teknoloji dersini veren öğretmenlerin hem alternatif ölçme ve değerlendirme teknikleri ile ilgili görüşleri hem de bu tekniklerin ne boyutta kullanıldığının ortaya çıkarılmasını amaçlamıştır. Yapılan çalışmalarda elde edilen bulgulardan ilköğretim 4. ve 5. sınıf Fen ve Teknoloji dersi öğretmenlerinin alternatif ölçme ve değerlendirme tekniklerini kullanmaları konusundaki görüşlerinin cinsiyet, hizmet yılı değişkenlerine göre farklılık gösterdiği, eğitim durumlarına göre ise farklılık göstermediği tespit edilmiştir. Öğretmenlerce en çok tercih edilen tekniklerin genelde klasik ölçme ve değerlendirme teknikleri olduğu; alternatif teknikler arasında proje, performans değerlendirme, portfolyo, kavram haritası ve posterleri sıklıkla kullanıldığı anlaşılmıştır. Alternatif ölçme-değerlendirme tekniklerinin kullanımında en büyük problemin zaman yetersizliği ve sınıf mevcudunun fazlalığı olduğunu belirtmişlerdir. Çalışmadan elde edilen sonuçlara dayanarak; okullarda sınıf mevcutlarının azaltılması, ders saatlerinde artırım yapılması, öğretmenlere alternatif ölçme ve değerlendirme teknikleriyle ilgili hizmet içi eğitim verilmesinin bu tekniklerin kullanılmasında katkı sağlayacağı görüşleri belirtilmiştir.

Gözütok ve arkadaşları (2005), öğretmenlerin mesleki yeterliklerini araştırmayı ve araştırma sonuçlarına bakılarak öğretmen yetiştirmeye ilgili sorunlara çözüm önerileri getirmeyi uygun bulmuştur. Bulgular madde bazında incelendiğinde en yüksek puanı alan; 85,6 ile “Öz değerlendirme yapıyor musunuz?” ve “Mesleki gelişmeleri izliyor musunuz?” maddeleridir. En düşük puanın alındığı maddeler ise “Öğrenciyi değerlendirme

konusundaki yeterlik düzeyiniz”, “Öğrencilere portfolyo hazırlatma (süreç değerlendirme) konusundaki yeterlik düzeyiniz” ve “Ölçme değerlendirmeye ilişkin yeterlik düzeyinizdir. Bu maddelerin tümünün ölçme değerlendirme alanı ile ilgili olduğu, dolayısıyla öğretmenlerin kendilerini en düşük düzeyde algıladıkları konunun ölçme değerlendirme olduğu ortaya çıkmıştır.

Erdemir (2007) araştırmasında ilköğretim okullarının ikinci kademesinde görev yapan alan öğretmenlerinin ölçme değerlendirme tekniklerini etkin bir şekilde kullanabilme yeterliklerini belirleyebilmeyi ve öğretmenlerin ölçme değerlendirmeye ilişkin yaklaşımlarını tespit etmeyi amaçlamıştır. Araştırmaya Kahramanmaraş il merkezinde bulunan Türkçe, Matematik, Fen Bilgisi, Sosyal Bilgiler, İngilizce, Din Kültürü ve Ahlak Bilgisi branşlarında görev yapan 568 öğretmen katılmıştır. Öğretmenlere anket dağıtılarak araştırmanın bulguları elde edilmiştir. Araştırmanın bulgularına dayanılarak; öğretmenlerin mezun oldukları öğretim kurumlarından ölçme değerlendirme ile ilgili yeterli bilgi almadıkları ve ölçme değerlendirme teknikleri uygulamalarından yeterli bilgi düzeyine sahip olmadıkları sonuçlarına ulaşılmıştır.

Gelbal ve Kelecioğlu (2007), “Öğretmenlerin Ölçme ve Değerlendirme Yöntemleri Hakkındaki Yeterlik Algıları ve Karşılaştıkları Sorunlar” çalışması ile yapılandırmacı öğrenme yaklaşımına göre eğitim yapılan sınıflarda öğretmenlerin kullandıkları ölçme ve değerlendirme yöntemlerine yönelik görüşleri betimlenmeye çalışılmıştır. Araştırmanın bulgularına göre, öğretmenlerin öğrenci başarısının belirlenmesinde, kendilerini daha yeterli olarak gördükleri geleneksel ölçme yöntemlerini tercih ettikleri görülmüştür. Ayrıca, öğretmenlerin hiç kullanmadıkları teknikler arasında ilk sırayı öğrencilerin kendilerini değerlendirmesi almaktadır. Ölçme araçlarını kullanmada karşılaştıkları sorunların başında sınıfların kalabalık oluşu ve zaman yetersizliği gelmektedir. Öğretmenlerin büyük bir bölümünün yeni karşılaştıkları değerlendirme yaklaşımlarının olumlu ve olumsuz yanlarına ilişkin görüşlerinin, programda belirtilen özelliklerle paralellik gösterdiği ortaya çıkmıştır. Diğer sonuçlar dikkate alındığında, öğretmenlerin ölçme tekniklerinin kullanımı ve hazırlanması konusunda eğitime ihtiyaçları olduğu anlaşılmaktadır. Öğretmenlerin ölçme yöntemlerini kullanmada karşılaştıkları sorunları azaltabilmek için buldukları bölgede veya okullarda ölçme ve değerlendirme uzmanının bulunması gerektiğini belirtmektedir.

2.1.2. Yurtdışında Yapılan Araştırmalar

Bir sınıfı yönetmek ve o sınıfta eğitimsel kararları almak için ölçme ve değerlendirmenin kullanılması yeni bir kavram değildir. Robert Davis (1940) öğretmenlerin sınıfta yaşadıkları en büyük sorunların başında öğrencileri değerlendirmek olduğundan bahsetti. Özellikle öğretmen adaylarının üniversitede aldıkları ölçme ve değerlendirme dersleriyle ilgili araştırma yapanlar çok azdı (Noll, 1955; Mayo, 1957). Son yıllarda ise özellikle değerlendirme okuryazarlığıyla ilgili araştırmalar hizmet içinde çalışan öğretmenler üzerine yapıldığı görülmektedir.

Değerlendirme okuryazarlığı, 1990 yılında Amerikan Ölçme ve Değerlendirme Birliği tarafından geliştirilen bir kavramdır. Eğitimde Değerlendirme birliği (AFT, NCME&NEA, 1990) “Eğitimde Değerlendirmede Öğretmen Yeterliliği Standartları” belirlemiştir. S. Plake ve James Impara (1993) tarafından bu standartların her birini ölçecek çoktan seçmeli bir anket ”Sınıfta Değerlendirme Okuryazarlığı Anketi” (CALİ) geliştirilmiştir. Bu anket 7 değerlendirme standardını, her bir standardı ölçen 5 madde ile toplamda 35 maddelik çoktan seçmeli bir ölçektir. Anket ilk kez 6 farklı yapıdaki (ekonomik, sosyal farklılıklar vb.) okulda, 555 öğretmene uygulanmıştır. Sonuçlar öğretmenlerin özellikle değerlendirme sonuçlarını yorumlamada, not verme standardında en yüksek puanlara sahip olduklarını fakat uygun değerlendirme metodu seçmede ve geliştirmede çok kötü sonuçların olduğunu belirtmiştir (Plake, 1993).

Popham’a (2004) göre eğitimcilerin büyük bir çoğunluğu “değerlendirme okuryazarı değildir”. Çünkü eğitimciler, hizmet içinde veya üniversite eğitiminde ölçme ve değerlendirmeyi önemsemezler. Stiggins (2004) öğretmenlerin öğrencileri değerlendirmelerinde sürekli olarak aynı yolları kullandığını, çünkü öğretmenlerin ve öğretmen adaylarının ölçme değerlendirme ile ilgili yeni yollar öğrenmelerine fırsat tanınmadığını belirtmiştir. Buna ek olarak da öğretmenlerin çoğu ölçme ve değerlendirmeyi anlamamanın, istatistiksel ve nicel veriler nedeniyle çok karmaşık ve zor olduğunu düşünmektedirler (Popham, 2004).

Son 20 yıldan önceki çalışmalarda, öğrenci başarısını değerlendirme ve eğitimsel kararlar almadaki öğretmenlerin ölçme ve değerlendirme ile ilgili bilgi ve becerilerinin çok az bir kısmını formal olarak aldığını göstermektedir. Öğrencinin nasıl değerlendirileceğini anlamadan, değerlendirme sonuçlarını kullanmak öğretmenler ve öğrencileri için bir hayal kırıklığı olacaktır.

1980 yıllarının başlarında, birkaç profesyonel eğitim derneği bir öğretmenin öğrencilerinin başarısını en iyi şekilde değerlendirebilmesi için öğretmenin sahip olması gereken yeterlilikler için işbirliği içine girdiler. Amerikan Öğretmenler Federasyonu, Uluslararası Eğitimde Ölçme Derneği bu uygulamaların başında yer almışlardır.

Son 20 yıl içerisinde öğretmenlerin değerlendirme okuryazarlığı standartlara dayalı olarak birkaç çalışma yapılmıştır (Plake ve Impara, 1993; Campbell ve diğ., 2002; Mertler, 2003). Plake ve Impara (1993) bu standartlara dayalı olarak “*Teacher Assessment Literacy Questionnaire*” anketini geliştirmişlerdir. Bu çalışmada da bu anket örneği Türkçeye uyarlanarak yapılmıştır. Yapılan bu araştırma göstermiştir ki, sorulan 35 sorunun 23 (% 66) tanesine ortalama olarak doğru cevaplar verilmiştir. Bu durum da öğretmenlerin büyük çoğunluğunun öğrencileri değerlendirmede yetersiz olduğunu göstermektedir. Bu çalışmada, geçmişte ölçme değerlendirme eğitimi almış öğretmenlerin skorlarının eğitim almayan öğretmenlerin skorlarından daha yüksek olduğu da bulunmuştur.

Campbell ve diğ. (2002) aynı ölçeği öğretmen adaylarına kullanmış ve öğretmen adayları ortalama olarak değerlendirme okuryazarlığı envanterinde ki 35 sorudan 21 tanesine doğru cevap vermişlerdir. Plake ve Impara (1993) araştırmasından 2 tane daha az soruya doğru cevap vermişlerdir.

Mertler (2003) Plake ve Impara (1993) tarafından geliştirilen anketi uyarlayarak, hizmet içinde çalışan öğretmenler ile öğretmen adaylarının değerlendirme okuryazarlıklarını karşılaştırmış ve diğer çalışmalarda olduğu gibi benzer sonuçlara varmıştır. Bu çalışmanın amacı öğretmen adaylarını ve hizmet içinde çalışan öğretmen adaylarını karşılaştırmaktır. Bu amaçla yine Plake tarafından geliştirilen “Sınıfta Değerlendirme Okuryazarlığı Anketi” CALİ’yi kendi araştırma amacına göre düzenleyerek uygulamıştır. 2002 yılında 67 öğretme adayı ve 197 hizmet içi öğretmenle araştırma yapmıştır. Araştırma istatistiksel analizlerle çözümlenmiştir. Araştırma sonucunda, öğretmen adaylarının değerlendirme okuryazarlıkları; uygun değerlendirme metodu seçmede en yüksek puanı almıştır. Her iki grup karşılaştırıldığında özellikle öğretmen adaylarının değerlendirme okuryazarlıklarında farklı standartlar açısından farklı sonuçların olduğu görülmüştür.

Stiggins (2000) yaptığı çalışmada öğretmenler için 6 tane sınıfta değerlendirme yeterliliği belirlemiştir: 1) Eğitimciler değerlendirmenin kullanımını ve kullanıcının tüm sırasını anlamalıdır. 2) Öğrenciden istediği başarı konusunda net olmalıdır. 3) Öğrenci başarısı için kullanacağı tüm ölçme ve değerlendirme metotlarını kullanmaya hazır

olmalıdır. 4) Öğrenci performansını etkili ve etkin biçimde ortaya koyacak yöntemleri seçmelidir. 5) Özel durumlardan ortaya çıkacak ön yargıları anlamalı ve sonuçları olumsuz etkileyecek durumları engelleyebilmelidir. 6) öğrenci motivasyonu ve değerlendirme arasındaki ilişkiyi anlamalıdır.

McTighe ve O'Connor (2005) ise yaptıkları araştırma sonucunda öğrencilerin etkili öğrenmeleri için bir öğretilerde olması gereken 7 tane özel değerlendirme uygulamalarını belirlemişlerdir: 1) Summative (ürün odaklı) değerlendirme kullanılmalı, 2) gelişmiş model ve kriterleri göstermeli, 3) öğretimden önce değerlendirmeli, 4) uygun seçimle yapmalı, 5) sık sık ve erken bir şekilde dönüt sağlamalı, 6) amaca uygun öz değerlendirme için öğrencilerini cesaretlendirmeli, 7) eskiyi ortadan kaldırıp yeni gelişmelere açık olmalıdır.

Ölçme ve değerlendirme okuryazarı öğretmenler, öğretim için değerlendirmeye nasıl odaklanacağını bilmelidirler. Çünkü bu öğretmenlerin öğrencilerin yeteneklerini ortaya çıkarmak ve onlar için uygun öğretim yöntemlerini seçmelerine ve planlamalarına kayda değer bir şekilde faydalı olacaktır (Leahy diğ., 2005; McNamee ve Chen, 2005; McTighe ve O'Connor, 2005).

En iyi öğretmen, hem kendisi hem de öğrencisinin performansının en üst düzeye çıkaracak ve sürekliliğini sağlayacak şeyin ölçme ve değerlendirme olduğunun farkına varan kişidir (McTighe ve O'Connor, 2005).

Öğretmen yetiştiren kurumların ve eğitim politikacılarının öğrenci başarısını geliştirmede sonuç, uygulama ve eğitimsel kararların gerçekleşmesinde etkili bir yol olması için en büyük görevlerinden biri öğretmenlerin değerlendirme okuryazarlıklarını geliştirmek olmalıdır. Özellikle son yıllarda öğrenci başarısını değerlendirme üzerine yapılan araştırmalar, sınıfta değerlendirme ve öğrenci performansı arasındaki ilgidir (Stiggins, 1999).

Öğretmenler için önemli olan yalnızca değerlendirme okuryazarı olmak, bu değerlendirme tekniklerini bilmek ve anlamak değildir. Aynı zamanda bildiklerini sınıflarında uygulamalıdır. Popham (2006) şöyle açıklıyor: “Bugün ve bugünden sonra eğitimde ölçme ve değerlendirme öğrencinin eğitiminde temel bir rol oynar. Bu yüzden eğitimciler ve eğitimle ilgilenen herkes değerlendirme okuryazarlığına bir parça ihtiyaç duyacaktır.” Sözleriyle bu kavramın önemini net bir şekilde ortaya koymaktadır.

Sınıf amaçları için, test sonuçlarının ve öğrenci performansının kullanımı öğrenci başarısını yansıtmada en geçerli olacaktır (Popham, 2005). Bu bir değerlendirme ölçeceği değildir, daha ziyade uygun eğitimsel kararlar için sonuçların etkili bir şekilde

yorumlanmasıdır. Sonuç olarak öğretmenlerin yaptıkları değerlendirmelerde bir teste dayalı olmaktan daha çok sonuca odaklanması öğrenci hakkında alınan kararı daha çok güvenli kılacaktır (Popham, 2003). Çünkü insan kararıyla verilmiş değerlendirmenin geçerliliği daha güvenilirdir. Bazen yanlış yapılabilir, kararlar doğru veya geçersiz olabilir ve bunlardan yola çıkarak alınmış eğitimsel hatalar yapılabilir (Popham, 2003).

Son 20 yılda yapılan araştırmalar öğretmenlerin, etkili öğretimi sağlayacak ve öğrenci başarısını uygun bir şekilde değerlendirecek gerekli bilgi becerilerde okuryazar olmadıklarını gösteriyor (Plake ve Impara, 1993; Mertler ve Campbell, 2005). Şüphesiz son yıllarda eğitimde yapılan araştırmalar ölçme ve değerlendirmenin önemini açık bir şekilde dile getirmektedir. Ölçme ve değerlendirme okuryazarlığı yurt dışında çok çalışılmış olmasına, bu konuyla ilgili yeterli araştırma olmasına rağmen çağdaş ölçme ve değerlendirme yöntemleriyle harmanlanarak yapılması gerekmektedir. Ülkemizde bu konuyla ilgili yalnızca bir araştırmaya rastlanmıştır. Bütüner ve diğ. (2010) yaptıkları araştırmada ölçme ve değerlendirme okuryazarlığı envanterini Türkçeye uyarlamış ve güvenilirlik ve geçerlilikleriyle ilgili çeşitli analizler yapmıştır. Bu çalışmada tüm yeterlilik alanları ele alınmamıştır. Etik kurallar ile ilgili yeterliliklerin ülkemiz standartlarında olmaması gerekçesiyle göz önünde bulundurulmamıştır. Yaptığımız araştırmada ise tüm yeterlilik alanları ele alınmış ve ülkemizde de bu yeterlilik alanıyla ilgili yapılacak çeşitli çalışmalara da yol göstereceği düşünülmektedir.

2.2. KAVRAMSAL ÇERÇEVE

Bu bölümde eğitimde ölçme ve değerlendirme kavramları, alternatif ölçme ve değerlendirme yaklaşımları ve teknikleri ayrıntılı bir şekilde açıklanmış, bu yaklaşımların kullanımına yönelik öz yeterlik ve okuryazarlık düzeyi ile ilgili bilgilere yer verilmiştir.

2.2.1. Eğitimde Ölçme Ve Değerlendirme

2.2.1.1. Eğitimde Ölçme

Bir özelliğin gözlenip, gözlem sonucunun sayı ya da semboller ile ifade edilmesi olarak tanımlanabilir (Turgut, 1997). Eğitim, içinde girdi, çıktı ve kontrol öğelerini

barındıran bir süreçtir. Bu süreçte bireyde kazanımların gerçekleşip gerçekleşmediğini belirlemek için yapılan değerlendirmelerin gerçekçi olması ancak geçerli ve güvenilir ölçümler yoluyla meydana gelebilir (Atılğan, Kan ve Doğan, 2009). Ölçme karar verme işi değildir, karar vermek için birçok faktör hakkında bilgiler sağlamaktır (Thorndike, 1997). Eğitimde ölçmenin önemi, verilecek kararlara bir destek sağlamasından gelir. Hassas ölçme araçları geliştirildikçe eğitimle ilgili yapılacak ölçmeler daha objektif hale getirilebilir (Turgut, 1997). Böylelikle eğitimle ilgili ölçme sonuçlarına dayanılarak alınan kararlar daha objektif olacaktır.

Eğitimde ölçme ve değerlendirme, insanların zihninde meydana gelen değişiklikleri ve öğrenme çıktılarını ölçmeye çalışmaktadır. Bu yüzden de ölçülmek istenen değerler soyuttur. Bu durum ise insanların üzerinde durduğu en önemli noktalardan biridir. Somut şeyleri ölçmek kolay ve sonuçlar daha nettir (Tekin,2000).

Ölçmeyi, Weltzenhoffer (1951) “bir gözlemci tarafından fiziksel dünya hakkında bir karar oluşturma” , Stevens (1946) “kurallara göre olay ve objelere sayılar tayin etme” ve Turgut (1986) “bir niteliğin gözlenip, gözlem sonucunun sayı veya sembollerle gösterilmesi” olarak tanımlar (Tekindal, 2002).

Ölçme işleminin aşamaları şunlardır (Tan ve Erdoğan, 2004):

- Ölçülmek istenen özelliklerin neler olduğunun tespit edilmesi
- Ölçülecek özelliklerin hangi sayı veya sembollerle ifade edileceğine karar verilmesi
- Belirli kurallara göre niteliklerle nicelikleri eşleştirme işleminin yapılması

2.2.1.1.1. Ölçme Araçlarının Özellikleri

Güvenirlilik

Güvenirlilik, ölçme sonuçlarının tesadüfi hatalardan arınlık derecesini ifade eden bir kavramdır (Turgut, 1997). Bir ölçme sonucu, içindeki tesadüfi hataların azlığı oranında güvenilir sayılır. Güvenirlilik, herhangi bir ölçme aracının ölçtüğü özellikleri ne derece duyarlılıkla ölçebildiği, yani ölçme sonuçlarının hatadan ne derece arınmış olduğudur. Güvenirlilik, ölçü aracının tutarlılığıdır. Eğer iki ölçüm arasında fark yok ya da az ise o testin güvenirliliği yüksek, bu fark çok ise güvenirliliği düşüktür. Ölçme aracının duyarlı ve tutarlı ölçme yapması aracın güvenirliliğini belirler (Küçükahmet, 1999).

Eđitimde kullanılan ölçme araç ve yöntemlerinin güvenilirliğini artırmak için alınabilecek tedbirler şunlardır (Turgut, 1997):

- Bir sınavda kullanılan soru sayısı arttıkça, çođunlukla o sınavdan elde edilen toplam puanın güvenilirliđi de artar.
- Bir sınavda kullanılacak soruların açık, anlaşılır ve kesinlikle cevaplanabilir düzeyde olması, o sınavdan elde edilen puanın güvenilirliğini artırır.
- Bir sınavda öğrencilerin, her soruyu dikkatli ve hızlı cevaplandırmaya teşvik etmek güvenilirliği artıracaktır.
- Sınav süresi öğrencilerin hemen hepsinin bütün soruları cevaplandırmalarına yetecek uzunlukta olmalıdır.
- Bir sınavda kullanılacak sorular, o sınavı cevaplayacak bireylerin yaklaşık olarak yarısı tarafından dođru cevaplandırılabilir güçlükte olmalıdır.
- Her sınav objektif yollarla puanlanabilmelidir.
- Ölçmenin yapılmasında gösterilecek titizlik ve dikkat hemen her zaman güvenilirliği artırır. Ölçme işleminin herhangi bir basamađında yapılan dikkatsizlik hataları, puanlama hatalarında olduđu gibi çođu halde tesadüfi hata türündedir ve ölçme sonuçlarının güvenilirliğini düşürür.
- Güvenirliđi artırmak için duyarlıđı yeterince yüksek araç veya yöntemler kullanılmalıdır. Duyarlıđı yüksek bir araç veya ölçme yöntemi daima güvenilirliđi yüksek sonuçlar verir.
- Ölçme sonuçlarının ölçme işleminin elverdiđi duyarlılıkta kaydedilmesi güvenilirliđi gerçek deđerine yaklaştırır.

Geçerlik

Geçerlik, genel anlamda bir ölçme aracının ya da ölçme yönteminin ölçmeyi amaçladığı özelliđi başka herhangi bir özellikle karıştırmadan dođru olarak ölçebilme derecesi olarak tanımlanabilir (Demirel, 2002). Eđitimde kullanılan araç ve yöntemlerle ölçülmek istenen deđişkenler bireylerin bazı özellikleridir. Bu özellikler çođu halde öğrencinin başarısı, kişinin yetenekleri gibi deđişkenlerdir. Aracın veya yöntemin ölçülmek istenen deđişkenin ölçüsü olabilecek bir puan vermesi, bu ölçüyü başka

değişkenlerle karıştırmaması beklenir. Bir araç veya yöntem bu niteliği sağlayabildiği derecede geçerli puanlar verir (Turgut, 1997). Bir aracın geçerli olması için onun sadece ölçmek amacıyla düzenlendiği özelliği ölçmesi, yani onu başka bir özellikle karıştırmadan ölçmesi gerekir. Eğitimde ölçmek istediğimiz bilgi ve becerilerin tümünü ölçmemiz zor olduğu gibi, sınavda elde edilen puana ölçmek istediğimiz özelliğin dışındaki etkenlerin de etkisi fazla olmaktadır. Bu etkenlerden dolayı elde edilen puanlara hata karışmaktadır (Küçükahmet, 1999).

Ölçme sonuçlarının geçerliğini sağlamak için gerek araçların yapımında, gerekse yöntemlerin uygulanmasında alınabilecek önlemlerin başlıca olanları şunlardır (Turgut,1997):

- Kapsam geçerliği sadece konular açısından değil, testteki soruları doğru cevaplandırabilmek için gerekli öğrenci davranışları açısından da anlam kazanır. Bir araç, kapsadığı soruların geçerliği oranında geçerli olabilir.
- Her soru, o soruyla ölçülmek istenen bilgi ve becerilere sahip öğrencilerin tereddütsüz doğru cevaplandırabileceği; bu bilgi, beceri ve yeteneklere sahip olmayanların doğru cevaplandıramayacağı nitelikte olmalıdır.
- Sınav belli bir sürede okutulan konular ve o konulara dayanılarak geliştirilmesi beklenen bilgi, beceri ve yetenekler bakımından hem kapsayıcı, hem de dengeli olmalıdır.
- Herhangi bir puanın geçerliği için ön şart, onun güvenilirliğidir. Güvenirlikten yoksun bir puan sadece tesadüfî hataları gösterir; hata da geçerli bir ölçü olamaz. Her ölçme sonucu ancak güvenilirliği oranında geçerlik kazanır.
- Sınavın ve ölçme araçlarının güçlüğü de geçerliğe etki eder. Güçlük derecesi o sınavdan elde edilecek puanların kullanılacağı amaca göre ayarlanmalıdır.
- Aynı soruların yıldan yıla hiç değiştirilmeksizin kullanılması, soruların sınavdan önce açıklanması, derslerin sınav sorularını çözmeye ayrılması, kopya çekilmesi, sınav süresinin yetersizliğinden dolayı bir kısım soruların cevapsız bırakılması, öğrencilerin çeşitli nedenlerle sınavı yarıda bırakması, öğrencinin gürültü ve hastalık gibi nedenlerle soruları tüm yetenekleriyle cevaplandıramaması gibi etkenler geçerliği düşürür.

2.2.1.2. Eğitimde Değerlendirme

Değerlendirme; ölçme sonuçlarını bir ölçüte ilişkilendirerek bir değer yargısına ulaşma süreci olarak tanımlanabilir (Turgut, 1997). Değerlendirmede iki temel öge vardır. Bunlardan birisi ölçüm, diğeri ise ölçümleri karşılaştırmak üzere kullanılan ölçüttür. Değerlendirme işlemi bu iki ögenin karşılaştırılması ile gerçekleşir. Harlen'e (1998) göre ise değerlendirme, ölçümler dâhil her türlü bilgi toplama ve bu bilgilerden bir değer yargısına ulaşmaktır. Değerlendirme, ölçme sonucunu bir ölçüt ile karşılaştırma ve bu yolla ölçme sonucuyla belirlenmiş olan özellik hakkında bir karara varma işlemidir (Özçelik,1981).

Yapılan değerlendirmenin doğruluğu hem bu ögelere hem de bu ögelerin nasıl karşılaştırıldıklarına bağlıdır (Atılgan ve diğ.,2009). Değerlendirme bir veya daha fazla gözlem verisini bir ölçüte vurma işlemidir (Yıldırım, 1999). Eğitimde değerlendirme, belirlenen hedeflere ulaşıp ulaşılmadığını ya da ne derece ulaşıldığını gösterir. Ölçülen nesnenin ya da niteliğin hangi özellikte olduğu konusunda karar verme ya da yargılama yapma işlemidir (İşman ve Eskicumalı,2003). Değerlendirme, bir ölçme işlemi sonunda elde edilen ölçme sonuçlarının bir ölçütle karşılaştırılması ve buna bağlı olarak bir değer yargısına ulaşma işidir (Demirel, 2002).

Ölçme ve değerlendirme birbirinden farklı kavramlardır. Bu farklılık şöyle açıklanabilir (Küçükahmet, 1999):

- Değerlendirme ölçmeyi de içine alan geniş bir kavramdır.
- Ölçme özelliğin miktarını gösterir; değerlendirme ise bu miktarın yeterli olup olmadığını ya da amaca uygun olup olmadığını belirtir.
- Ölçme daha çok gözleme; değerlendirme ise karşılaştırma, yorum ve yargıya dayanır.
- Ölçme, değerlendirmeden önce gelen bir işlemdir. Ölçmeden sonra değerlendirmeye gidilir.
- Ölçme daha objektiftir; değerlendirme ise daha çok kişisel kanılara dayanır.

Öğretim sürecinin farklı aşamalarında, farklı amaçlarla değerlendirmeler yapılır. Buna göre çeşitli değerlendirme türleri vardır.

Literatürde değerlendirme türleri sıklıkla şu şekilde sınıflandırılmaktadır (Demirel, 2006; Özçelik, 1992; Öztürk, 2005; Semerci, 2008; Sönmez, 2007).

- Tanıma ve yerleştirmeye yönelik değerlendirme
- Yetiştirme ve biçimlendirmeye yönelik değerlendirme
- Sonuca dayalı değerlendirme

Değerlendirme yapılırken göz önünde bulunması gereken unsurlar şunlardır (Moon, 2002)

- **Üst düzey düşünme:** Üst düzey düşünme stratejileri; yaratıcılığı, uzmanlık gelişimini ve kritik düşünmeyi kolaylaştırır.
- **Öğrenme sürecinin doğası:** Öğrenme, bilgi ve yaşantıdan gelen bir keşif ve yapılandırma sürecidir.
- **Öğrenmedeki bireysel farklılıklar:** Tüm öğrencilere, öğrenme, motivasyon ve etkili öğretimin bütün temel prensipleri uygulanır. Ancak buna rağmen öğrenciler, farklı öğrenme kapasitelerine sahiptirler. Her öğrencinin farklı bir öğrenme biçimi vardır, buna bağlı olarak da öğrencilerin öğretim stratejisi tercihleri farklıdır (akt. Ören, 2005).
- **Öğrenme sürecinin amaçları:** Öğrenen, tutarlı ve anlamlı bir yaratıcılık için araştırır.
- **Bilginin yapılandırılması:** Öğrenen, yeni bilgisini önceki bilgileri ile birleştirir.
- **Motivasyon artırıcı öğrenme görevlerinin özellikleri:** Merak, yaratıcılık ve üst düzey düşünme; öğrenci için orta zorluktaki gerçek öğrenme görevlerinin yenileriyle ilişkilendirilmesiyle uyandırılır.

2.2.1.2.1. Tanıma ve Yerleştirmeye Yönelik Değerlendirme

Bu tür değerlendirmede, öğrencilerin öğrenme sürecine girmeden önce belirlenen konular ile ilgili getirdikleri ön bilgileri ve hazır bulunuşluk düzeylerinin ölçülmesi esasına dayanır. Tanıma ve yerleştirmeye yönelik değerlendirmelerde öğrencilerin farklı yönlerdeki gelişimleri, farklı etkinliklere hazır oluş düzeyleri, ilgileri, eğilimleri ve yatkınlıkları belirlenerek onların en iyi şekilde gelişerek mutlu olabilecekleri etkinliklere yönlendirilmeleri amaçlanır. Bu yönlendirmeler okul ve program türü gibi öğrencilerin bütün okul ve okul dışı yaşamlarını etkileyecek düzeylerde olabileceği gibi, sadece bir dersteki yaşantılarını etkileyecek düzeylerde de olabilir (Özçelik, 1981). Öğrencilerin, belirlenen derste öğretilecek konular ile ilgili daha önceden öğrendikleri bilgilerin belirlenmesi ve bu bilgiler kullanılarak öğrenci başarısının değerlendirilmesidir (İşman ve

Eskicumalı, 2003). Tanıma ve yerleştirmeye yönelik değerlendirme genellikle öğrencinin, hem belirli bir program için gerekli ön davranışlara hem de o programın uygulanışı sırasında kazandırılacak davranışlara ne derecede sahip olduğunu anlamak amacıyla, programı uygulamaya başlamadan önce yapılır (Tan, Kayabaşı ve Erdoğan, 2003).

Bu değerlendirme biçimi öğrencinin giriş davranışları ile ilgilidir ve şu sorulara cevap arar:

- Eğitim şeklinin gerektirdiği ilgi, iş alışkanlığı ve kişilik özelliklerinin öğrencilerdeki derecesi nedir? (Tekindal, 2002)
- Planlanan eğitimin amaçladığı anlama ve beceriye öğrencinin sahip olma derecesi nedir?
- Öğrenci planlanan eğitimin başında ihtiyaç duyulan bilgi, ilgi ve beceri gibi özelliklere sahip midir?

Bu değerlendirme şekli ile öğrenciler bilgi seviyelerine göre çeşitli gruplara ayrılabilir. Öğretmen bireysel farklılıklardan kaynaklanan sorunları ortadan kaldırmak için öğretimi nitelikli bir hale getirir. Ders veya kursun başlangıcında rahatlıkla uygulanabilir (İşman ve Eskicumalı, 2003).

2.2.1.2.2. Biçimlendirme ve Yetiştirmeye Yönelik Değerlendirme

Bu değerlendirmelerde, bir yandan öğrencinin öğrenme hızı veya güçlüklerini gösteren işaretler, öte yandan da bu zorlukları ortadan kaldırmak için eğitim sistemindeki yetersizlikler ve hatalar eğitim-öğretim süreci içinde keşfedilir (Ertürk, 1982). Yetiştirme ve biçimlendirmeye yönelik değerlendirme üç aşamalı olarak ele alınabilir. İlk aşamada öğrencilerin belirlenen hedeflere ulaşip ulaşamadıklarına bakılır. Eğer öğrenciler hedeflere tam olarak ulaşamadıysa, hangi seviyede olduklarını belirlemek ikinci aşamadır. Son aşamada ise, öğrencilerin ulaşmada başarısız oldukları hedefler belirlenerek yeniden öğretim faaliyetine geçilir (İşman ve Eskicumalı, 2006).

Sönmez'e (2007) göre, yetiştirme ve biçimlendirmeye dönük değerlendirme; öğrencinin öğrenmede güçlük çektiği davranışları, öğrenemediklerini, yani öğrenme güçlük ve eksiklerini belirlemek, daha sonra ona gerekli yardımı yapmak, bu eksikleri tamamlamak, güçlükleri gidermek için yapılabilir. Öğrencilerin bir programa girdikten sonra süreç içinde sürekli değerlendirilmeleri önemli görülmektedir. Bu süreç içinde öğrencilerin öğrenme güçlüklerini ortaya çıkarmak ve gerekli düzeltmeleri gerçekleştirmek

için yapılan değerlendirmedir. Bu değerlendirme, programa sürekli dönüt sağlamak ve iyileştirici önlemlerin alınması için bir kontrol sistemi oluşturur (Demirel, 2002).

2.2.1.2.3. Değer Biçmeye Yönelik Değerlendirme

Eğitim-öğretim dönemi içerisinde veya sonunda, programın öngördüğü hedeflere ulaşıp ulaşılmadığına bakılarak öğrenci, öğretmen ve programla ilgili yargılarda bulunan değerlendirmelerdir.

Değerlendirme yapmanın amaçları şunlardır (Harlen, 1998):

- Başarıyı belirlenen zamanlara göre özetlemek (kayıt tutarak velilere, diğer öğretmenlere ve öğrencilere bilgi vermek)
- Öğrenmeye yardımcı olmak (gelişim düzeyi, karşılaşılan güçlükler vb.)
- Öğretim etkinliklerinin değerlendirilmesi (ortaya çıkan sonuçlara göre düzenlemelerin yapılması)
- Yeni materyallerin ve eğitim reformlarının incelenmesine yardımcı olmak
- Öğrencileri seçmek (daha sonraki seviyeler için alan belirlemede rehberlik etmesi)
- Öğrencilerin performanslarının okul, bölge veya ülke çapında izlenmesi

Tekin (1996) bu değerlendirme türünü, genellikle öğretim devresi sonunda, ara sıra öğretim devresi içinde, programın öngördüğü hedeflere ulaşıp ulaşılmadığına bakılarak öğrenci, öğretmen ve programa ilişkin yargılarda bulunmak, şeklinde açıklamıştır. Bu değerlendirme türünde öğretim programının amaçlarına ulaşılma derecesine bakılarak öğrenci, öğretmen ve programa ilişkin kararlar verilir (Öztürk, 2005).

Eğitim-öğretim sonunda öğrencilerin kazanmış oldukları bilişsel davranış, duyuşsal özellik ve devinişsel becerilerini ölçmeye yarayan değerlendirme türüdür. Bu değerlendirmelerle eğitim programının öğrencilere istenilen davranışları kazandırma açısından yeterli olup olmadığı hakkında bir yargıya varılabilir (Demirel, 2002).

2.2.2. Eğitimde Ölçme ve Değerlendirme Yaklaşımları

Kazanımların bireyde gerçekleşme düzeyini belirlemek için birçok ölçme ve değerlendirme yöntem ve tekniğı kullanılmaktadır. Bu yöntem ve teknikler özelliklerine göre iki başlık altında toplanabilir. Bunlardan ilki geçmişten günümüze kadar eğitim

ortamlarında sıklıkla kullanılan geleneksel ölçme ve değerlendirme yaklaşımları, diğeri ise son yıllarda kullanılmaya başlayan alternatif ölçme ve değerlendirme yaklaşımlarıdır. Bu bölümde geleneksel ve alternatif ölçme-değerlendirme yaklaşımları hakkında bilgiler, literatür ile birlikte açıklanmıştır.

2.2.1.3. Geleneksel Ölçme ve Değerlendirme Yaklaşımları

Eğitim-öğretim sürecinin içinde bulunan öğretmenlerin çoğu tarafından bilinen ve eğitimin her kademesinde kullanılan teknikler olarak tanımlanan (Bahar ve diğ., 2008) geleneksel ölçme ve değerlendirme yaklaşımları, genellikle kâğıt-kaleme dayalı testler olarak bilinmektedir. Kısa bir süre içerisinde geniş bir konu alanını değerlendirmek için kullanılırlar.

Bu yaklaşımları kullanmanın çeşitli avantajları vardır. Enger ve Yager (1998) geleneksel ölçme ve değerlendirmenin avantajlarını şu şekilde sıralamışlardır:

- Ekonomiktir ve birden çok öğrenciye kolaylıkla uygulanabilir.
- Standartlaştırılmış testlerin puanları rahatlıkla yerleşim birimlerine, şehirlere ve ülkelere göre karşılaştırmalarda kullanılabilir.
- Öğrencilerin özel konulardaki bilgilerini hızlı bir şekilde belirleyebilir.

Geleneksel değerlendirme; süreç talimatları, ölçme aracı ve puanlaması hazırlayanlar tarafından sabitlenmiş ve bu suretle sınavı uygulayan farklı kişilerce ve farklı oturularda sınava girenlerin tümünde kıyaslanabilir sonuçlara ulaşmak için hep aynı şekilde uygulanan ve puanlanan (özdeş) testlerdir (Enger ve diğ., 1998). Geleneksel değerlendirme bir öğrencinin gelişim süreci hakkında bilgi vermez (Cubertson ve diğ., 1999; Curtis ve diğ., 2002). Geleneksel ölçme araçlarıyla yapılan değerlendirme, öğrencinin bilgisini belirli bir zaman diliminde ölçmeye çalışmakta, öğrencinin başarısını ya da başarısızlığını ayrıntılı olarak yansıtmamakta ve öğrencinin zihinsel semasını ortaya koyamamaktadır (Manning ve diğ., 1995; Romberg, 1993; Shepard, 2000). Geleneksel değerlendirmeler, performansın açık bir şekilde sergilenebilmesini engeller (Chen ve diğ., 2000). Öğrencilerin kendi gelişimlerini süreç boyunca takip etmelerine izin vermez. Öğrenciler sadece aldıkları notlara odaklanırlar (Enger ve diğ., 1998). Geleneksel değerlendirme araçları üst seviyedeki becerileri değerlendirmede yetersiz görülmektedir (Yaman ve diğ., 2005; Enger ve diğ., 1998).

Enger ve diğ., (1998) geleneksel ölçme ve değerlendirme yöntemlerinin dezavantajlarını şu şekilde sıralamıştır:

- Genellikle bilginin parçalarının bir araya toplanması ile temsil edilebileceği ve bir doğru cevabın var olduğu varsayımı üzerine kurulmuştur. Oysa bu, yanlış ve sakıncalıdır.
- Öğretime yönelik sakıncaları vardır. Öğretmenler testlere yönelik öğretirler ve okullar test sonucunda elde edilen bilgilere göre değer sırasına konurlar. Üst düzey becerileri yansıtmada bu sıralama doğru değildir.
- Öğrencilerin sınıfta karşılaştıkları basit tartışmaları yansıtabilir. Fakat değişen ve kabul gören öğrenme teorilerini ve bilişsel teorileri yansıtmaz. Öğrencilerin gelecekteki başarısı için gerçekten ihtiyaç duyacağı yetenekleri temel almaz.
- Öğrenci gelişimini sürekli olarak yansıtamaz. Yakın olmayan aralıklarda uygulanır ve bağımsız olarak yorumlanır.

Geleneksel ölçme ve değerlendirme yaklaşımlarının yukarıda belirtilen avantajlarının dışında bazı dezavantajları da bulunmaktadır. Korkmaz'a (2004) göre öğrencilerin ne öğrendiğinin bir kanıtı olan bu testlerle gerçek öğrenme yaşantılarının değerlendirilmesi zor görünmektedir. Hein ve Price'da (1994) geleneksel testlerin modern ve geniş kapsamlı fen eğitimi programlarının tüm öğelerini değerlendirmede yetersiz kaldığını ifade etmişlerdir. Geleneksel testlerle yapılan değerlendirmede öğrencinin bilgisi belirli bir zaman dilimiyle ölçülmeye çalışılmaktadır. Bu nedenle geleneksel yaklaşımlar öğrencinin başarısını ya da başarısızlığını ayrıntılı olarak yansıtmamakta ve öğrencinin zihinsel şemasını ortaya koyamamaktadır (Manning ve diğ., 1995; Romberg, 1993; Shepard, 2000; Akt: Orhan, 2007; Haney ve Madaus, 1989; Neillve Medina, 1989; O'Neill, 1992).

Wiggins (1989) araştırmalarında çoktan seçmeli testlerin açık uçlu problem çözme gibi üst düzey becerileri ölçmede yetersiz olduğunu rapor etmişlerdir (Akt: Enger ve Yager, 1998). Bu sınırlılıklar göz önüne alındığında öğrencilerin üst düzey düşünme becerilerini bilişsel, duyuşsal ve psikomotor alanlardaki kazanımlarını değerlendirebilmek için geleneksel ölçme ve değerlendirmenin yanı sıra farklı ölçme ve değerlendirme tekniklerinin kullanımına ihtiyaç duyulduğu ortaya çıkmaktadır.

Literatürde sıklıkla rastlanılan geleneksel ölçme ve değerlendirme teknikleri şunlardır:

- Yazılı yoklamalar
- Kısa cevaplı testler

- Çoktan seçmeli testler
- Doğru-yanlış testleri
- Sözlü sınavlar
- Eşleştirme soruları
- Boşluk doldurmalı sorular
- Kısa cevaplı sorular

2.2.1.3.1. Yazılı Yoklamalar

Öğrencilere birkaç soru yazdırılıp veya yazılı verilip bunlara belli bir sürede yazılı cevap istenmesi, bir sınav türü olarak her öğretmenin başvurduğu bir yöntemdir. Böyle bir sınavda öğrenciden istenen işler cevapları düşünmek, tasarlamak, düzenlemek, yazmak ve gerekirse yazdıklarını kontrol etmektir. Öğretmene düşen işler ise, sınavdan önce soruları hazırlamak, sınavı yönetmek, sınavdan sonra cevapları okuyup puanlamaktır (Turgut, 1997). Yazılı sınavlar; öğrencilerin problemleri formüle etme, organize etme, bazı fikirleri değerlendirme, bilgilerini değişik durumlara uygulama olanağı sağlama, orijinal görüş veya ürünler ortaya koyabilme davranışlarını ölçmede gayet uygun bir sınav türüdür (Tan ve Erdoğan, 2004). Yazılı sınavlar eğitim sistemimizde çok kullanılan bir sınav tipidir. Bunun nedenlerinden bazıları soruların hazırlanmasının görünüşte kolay oluşu, öğretmenlerin geleneksel olarak yazılı sınav tipine alışık olması, diğer sınav türlerini iyi tanımamaları ve bazı eğitimcilerin uygulama, organize etme, analiz, değerlendirme gibi yüksek seviyedeki davranışların diğer sınav türleri ile ölçülemeyeceği düşüncesinde olmaları olabilir (Tekindal, 2002). Cevapların yazılı verilmesi zorunluluğu, cevapların uzunluğu, cevaplayıcı bağımsızlığı, puanlamanın öğretmen veya puanlayıcı kanısına dayanması gibi özellikler yazılı sınavların niteliklerini etkileyen önemli etkenlerdir. Bunlar çoğu halde yazılı sınavların bir ölçme aracı olarak yeterince güvenilir ve geçerli sonuçlar vermesini engelleyen etkenlerdir (Turgut, 1997). Yazılı sınavların temel özelliklerinden biri öğrencilerin sorulara serbest olarak cevap verebilmeleridir. Yazılı sınavlarda cevaplayıcı bağımsızlığı vardır. Öğrenciler kendi düşünce ve fikirlerini serbest olarak kendi kelimeleriyle ifade edebilmektedir (Tan ve Erdoğan, 2004).

Yazılı sınavların başlıca özellikleri şöyle sıralanabilir (Turgut, 1997):

- Yazılı sınavda cevaplayıcı, cevabını kendi düşünüp bulmak zorundadır.
- Yazılı sınavda cevaplayıcı cevabını yazılı olarak ifade etmek zorundadır.

- Yazılı sınavda cevaplar çoğu halde sınırlı değildir. Her cevaplayıcı cevaplarını dilediği gibi ifade etme bağımsızlığına sahiptir.
- Genellikle bir yazılı sınav sorusuna verilen cevapları tamamen yanlış veya tamamen doğru olarak iki sınıfa ayırmak mümkün değildir.
- Yazılı sınavlarda cevapların doğruluk derecesini tayin etmek puanlayıcıya düşer.
- Cevaplama işlemi cevaplayıcının çok zamanını alır. Yazılı sınavlarda sınav süresinin çok büyük bir kısmı cevapların düzenlenip yazılmasına harcanır.
- Yazılı sınavların puanlanması güçtür.
- Bir yazılı sınavın hazırlanması için gerekli zaman genellikle diğer ölçme araçlarına göre kısa, uygulamadan sonra yapılacak puanlama işlemleri için gerekli zaman ise oldukça uzundur.

Yazılı sınav sorularını hazırlamak ve uygulamak için aşağıda bazı öneriler sıralanmıştır (Kemertaş, 2003):

- Mutlaka bir sınav planı hazırlanmalıdır. Kritik davranışlar ile sınav kapsamına giren konular listelenmelidir.
- Soru kâğıdının başına açıklama yapılmalıdır.
- Soruların öğrencilerin seviyesine uygun ve ayırt edici olmasına özen gösterilmelidir.
- Sorular önceden hazırlanıp çoğaltılarak öğrencilere dağıtılmalıdır.
- Sınavda sorulacak soru sayısından daha fazla soru hazırlanmalıdır. Bunlar arasından sorular seçilmelidir. Böylece nitelikli, zorluk dereceleri ayarlanmış, konuları örnekleyebilen sorular seçilmiş olur.

2.2.1.3.2. Kısa Cevaplı Testler

Cevaplayıcının bir kelime, bir rakam, bir ibare veya en çok bir cümle ile cevaplandırabileceği maddelerden meydana gelmiş test tipine kısa cevaplı test denir. Bu test çeşidi ayrıntılı cevap veya kompozisyon gerektiren yazılı sınavlardan cevapların kısalığı ile seçme veya sınıflama gerektiren testlerden ise cevapların verilenler arasından seçilmeyip cevaplayıcı tarafından bulunması ve yazılması özelliği ile ayrılır (Turgut, 1997). Kısa cevaplı bir test maddesi soru cümleleri olarak sorulabileceği gibi, doldurulmalı test yapısında da sorulabilir. Kısa cevaplı testler diğer objektif testlerde olduğu gibi, yazılı sınavla kıyasla daha alt düzeyde zihinsel etkinlikleri ölçmede kullanılır. Kısa cevaplı test

maddeleri genelde bilgi ve kavrama düzeyindeki davranışları ölçmede uygundur (Tan ve Erdoğan, 2004).

Kısa cevaplı testlerin başlıca özellikleri şöyle sıralanabilir (Turgut, 1997):

- Cevaplayıcı, cevapları düşünüp bulmak ve yazmak zorundadır. Bu yapıdaki test maddeleri cevaplayıcının sorulan bilgiyi hatırlama ve bulma yeteneğini yoklar.
- Cevaplayıcı, istediği cevabı verme bağımsızlığına sahiptir. Bu özelliği ile kısa cevaplı testler yazılı sınavlara yaklaşıp.
- Cevaplar kısadır. Bu özelliğinden dolayı kısa cevaplı testler seçmeli testlere yaklaşıp.
- Kısa cevap gerektiren maddeler çok çeşitli bilgi ve yetenek alanlarına uygulanabilir.
- Kısa cevaplı testler her eğitim düzeyine uygun düşer.
- Puanlama işlemi nispeten kolaydır; fakat tamamen objektif değildir. Puanlamanın kolaylığı cevapların kompozisyon gerektiren sorulara kıyasla sınırlı, kısa ve kesin oluşundan ileri gelir.
- Kısa cevaplı testlerin hazırlanması, diğer yapıdaki testlere kıyasla daha kolaydır.

Kısa cevap gerektiren sınavların sorularını geliştirmek için öneriler şu şekilde sıralanabilir (Tekindal, 2002):

- Ders kitapları ve benzeri materyaldeki ifadeler olduğu gibi alınarak soru şekline dönüştürülmemelidir.
- Sorular açık olarak ifade edilmeli ve anlaşılması güç sorular testten çıkarılmalıdır.
- Her soru bir tek doğru cevabı olacak şekilde oluşturulmalıdır.
- Her soru önemli bir bilgiyi yoklamalıdır. Eğitimsel değeri olan ve üst düzeyde düşünme gerektiren sorular tercih edilmelidir.
- Bir soru, o soru ile yoklanmak istenen bilgi, beceri ve davranışa sahip olan bireylerce doğru, sahip olmayanlarca yanlış cevaplandırılmalıdır.
- Soruların yazılmasında şekil, harita, grafik ve metinlerden yararlanılmalıdır.

2.2.1.3.3. Çoktan Seçmeli Testler

Çoktan seçmeli bir test maddesinde soruyla birlikte, sorunun cevabı olabilecek seçenekler de sunulur. Öğrencilerin sunulan seçeneklerden doğru olanı bulup işaretlemeleri beklenir. Çoktan seçmeli testler; sorulan bir sorunun doğru cevabını verilen seçeneklerden bulduran sınav türüdür (Tan ve Erdoğan, 2004).

Seçmeli test maddeleriyle ilgili kavramlar aşağıda açıklanmıştır (Turgut, 1997):

Madde Kökü: Madde ile ne sorulduğunu ifade eden veya ima eden ve sorunun çözümü için gerekli bilgileri veren kısımdır.

Cevap Şıkları: Madde kökündeki soruya muhtemel doğru cevap olarak verilen ifadelerdir. Maddede verilen seçeneklerden genellikle yalnız bir tanesi doğru veya en doğru cevap olur. Madde yazarının doğru cevap olarak belirttiği bu seçeneğe anahtara göre doğru cevap veya kısaca anahtarlanmış cevap denir.

Çeldiriciler: Bir maddenin anahtara göre yanlış seçenekleridir. Çeldiriciler bilgisi yetersiz, bilgisiz veya yanlış bilgili cevaplayıcıları yanıltmak için verilir.

Madde Yapısı: Bir test maddesinin cevaplama yönergesi de dâhil olmak üzere yazıldığı zaman aldığı şekle madde yapısı veya madde formu denir. Maddeyi oluşturan sembol, kelime ve ifadelerin düzeni, maddenin nasıl cevaplanacağını belirten açıklamalar, cevapların kaydediliş şekli madde yapısının temel öğeleridir.

Seçmeli testlerin özellikleri şunlardır (Turgut, 1997):

- Seçme gerektiren maddelerde, sorunun doğru cevabı kendi içinde verilmiştir. Cevaplayıcı, cevabı verilen seçenekler arasından seçmek zorundadır.
- Test süresinin büyük bir kısmı maddeleri okumaya veya doğru cevabı bulmaya harcanır. Seçilen cevabın kaydı pek az zaman alır. Bunun için sınırlı bir test süresinde yazılı sınavlara ve kısa cevaplı testlere kıyasla daha çok sayıda soru sorulabilir.
- Seçmeli testler cevaplayıcıda testle yoklanmak istenen bilgi ve yeteneklerin yanında okuma yeteneği gerektirir.
- Verilen cevaplar kesinlikle doğru veya kesinlikle yanlış şeklinde sınıflandırılabilir. Bu sebeple, seçmeli testler objektif olarak puanlanabilir.
- Seçmeli testlerde kullanılacak çok çeşitli madde formu vardır. Bunun için seçmeli testler çok çeşitli bilgi, beceri ve yeteneklerin ölçülmesine uygun düşer.

- Seçmeli testler basit madde formları kullanılmak şartıyla, aşağı eğitim basamaklarında da uygulanabilir.
- Seçmeli maddelerde doğru cevabın şansla bulunabilme ihtimali vardır.
- Seçmeli testlerde maddeler yazılırken test güçlüğü ayarlamak veya test denendikten sonra istenilen güçlükte madde seçmek mümkündür.
- İyi niteliklerde seçmeli testlerin hazırlanması, bu konuda tecrübeli madde yazarlarını gerektirir.
- Seçmeli testlerin hazırlanması uzun zaman alır. Buna karşılık sınavdan sonraki puanlama işlemi az zaman alır. Bu yüzden seçmeli sorularla yapılacak bir sınav, önceden önemli bir hazırlık süresi gerektirir; fakat sınavın sonuçları oldukça çabuk alınır.

Çoktan seçmeli test maddeleri yazarken dikkat edilmesi gereken genel hususlar şunlardır (Demirel, 2002):

- Test maddesi gayet açık bir dille yazılmalı, ifade kesinlik taşımalı, birbirinden farklı yorumlara imkân verilmemelidir.
- Her maddede ifade açıklığına, kesinliğine ve kısalığına dikkat edilmelidir.
- Test maddelerinde cinaslı, aldatıcı, bilmece gibi ifadeler kullanılmamalıdır.
- Test maddelerinin dil bakımından güçlük seviyesi, o testi cevaplandıracak grubun ortalama okuma yeteneğinin oldukça altında tutulmalıdır.
- Test maddeleri ve testlerde kullanılacak diğer yazılı malzeme, ders kitabı gibi öğrencilerin kullandığı yazılı kaynaklardan aynen alınmamalıdır.
- Test maddesi önemsiz ayrıntıları sormamalıdır.
- Her test maddesi diğer maddelerden bağımsız olarak cevaplandırılabilir.
- Bir maddede verilen bilgi, diğer bir maddede sorulan sorunun cevabını açıklamamalıdır.

Çoktan seçmeli test sorularının seçeneklerini yazarken göz önünde bulundurulması gereken noktalar şunlardır (Tekindal, 2002):

- Seçeneklerin her biri kökle tutarlı, ifade tarzı, kapsam ve uzunluk bakımından birbirine benzer olmalıdır.
- Seçenekler mümkün olduğunca kısa yazılmalı, gerekli bilgiler kökte toplanmalıdır.
- Seçeneklerde farklı anlamlara gelebilecek daima, asla gibi kelimeler kullanılmamalıdır.

- Seçenekler birbirinden bağımsız olmalıdır.
- Seçenekler birbirinin anlamını içermemelidir.
- Birbirine zıt düşen seçenekler yazmak özel dikkat ister.
- Bir testteki maddelerin hepsinin seçenekleri birbirine eşit sayıda olmalıdır.
- Seçenekler mantıksal düzenleme, alfabetik sıralama gibi bir durumda olursa, bunlar sıra ile verilmelidir.
- Çeldiriciler, soruyla yoklanan davranışı öğrenmemiş olanlara doğru cevapmış gibi gözükecek nitelikte olmalıdır.
- Doğru cevap bir kitap veya başka materyalden değiştirmeden alınmamalı ve kalıplaşmamış bir ifadeyle verilmemelidir.
- Seçeneklerde doğru cevabın veriliş şekli bir sistematiklik göstermemeli, rastgele dağıtılmalıdır.
- Yukarıdakilerin hepsi ve yukarıdakilerin hiçbiri ifadelerinin seçenek olarak kullanılması özel bir dikkat gerektirir.
- Anlam bakımından ele alındığında, çeldiricilerin doğru cevaba yakınlık derecesi yoklanan davranışla, erişilmesi beklenen düzeye uygun olmalıdır.

2.2.1.3.4. Doğru-Yanlış Testleri

Doğru-yanlış testleri, öğrencinin test maddelerini doğruluk-yanlışlık ölçütüne göre sınıflamaları şeklinde uygulanır. Doğru-Yanlış testlerindeki maddeler bazıları doğru, bazıları yanlış önermelerden oluşur (Tan ve Erdoğan, 2004). Prensiplerin ve genellemelerin iyi anlaşılıp anlaşılmadığının yoklanacağı durumlarda ve belli bir nokta ile ilgili yalnız iki ihtimal olduğu durumlarda doğru-yanlış maddelerinin kullanılabilceği ileri sürülmektedir (Yılmaz, 1998).

Doğru-yanlış testlerinin genel özellikleri şunlardır (Turgut,1997):

- Test maddeleri bir kısmı yanlış, bir kısmı doğru önermeler halinde verilmiştir. Cevaplayıcıdan her cümlenin doğru veya yanlış sembolleriyle kaydetmesi istenir. Cevaplayıcı, maddede verilen önermenin doğruluğunu ya da yanlışlığını kendi bilgilerine göre kendi yargısını kullanarak kararlaştırır.
- Cevaplayıcı, cevaplarını yalnız bir harf yazarak veya bir işaret koyarak kaydeder.

- Cevapların bütün cevaplayıcılar için ortak olan iki işaretle verilmiş olması, bu yapıdaki maddelerin puanlanmasını hem kolaylaştırır, hem de tamamen objektifleştirir.
- Madde yapısı sabittir; fakat basitliğine rağmen, oldukça ileri basamaktaki bilgileri yoklamaya da elverişlidir.
- Cevaplayıcıya verilecek test yönergesi kısa ve basittir. Cevaplayıcıdan istenen cevaplama işlemi, ölçülen bilgi veya yeteneğin dışında bir zorluk göstermez. Bu sebeple, doğru-yanlış testleri eğitimin aşağı basamaklarında bile kolaylıkla uygulanabilir.
- Sadece iki cevap seçeneği bulunduğu için, önermenin doğru olup olmadığını bilmeden işaretleyen bir cevaplayıcının istenilen cevabı tutturabilme şansı oldukça yüksektir.

Her test çeşidinde olduğu gibi, bu test çeşidinde de maddelerin yapısından gelen bazı teknik zorlukların farkına varmak ve bunlardan doğacak sakıncaları önlemek, madde yazmada yeterliğe sahip olmakla mümkündür.

Doğru-Yanlış testleri için madde yazmada dikkat edilecek esaslar aşağıdaki dokuz noktada toplanabilir (Turgut, 1997):

- Her maddede sadece bir tane ana fikir bulunmalıdır.
- Her madde kesinlikle doğru ya da yanlış olmalıdır.
- Bir maddenin yanlışlığı önemsiz ve atlatıcı bir noktada olmamalıdır.
- Her madde kısa ve açık yazılmalıdır.
- Olumsuz, özellikle çift olumsuz ifadeler kullanılmamalıdır.
- Maddelerin uzunlukları birbirine yakın olmalıdır.
- Bir maddede birden fazla yanlış fikir bulunmamalıdır.
- Bir maddenin yanlışlığını veya doğruluğunu apaçık belirtecek ipuçları bulunmamalıdır.
- Test maddesiyle önemsiz ayrıntılar hemen her kaynakta bulunabilecek ezberlenebilir bilgiler yoklanmamalıdır.

2.2.1.3.5. Sözlü sınavlar

Soruların genellikle sözlü sorulduğu ve cevapların sözlü verildiği sınav çeşidine sözlü sınav denir. Birkaç üyeden oluşan bir komisyonun bir cevaplayıcıya sorular sorması ve sözlü cevaplara hemen not takdir etmesi suretiyle yapılan sınav, sözlü sınavların en çok başvurulan türlerinden biridir. Öğretmenin sınıftaki öğrencilerden birini yazı tahtası önüne çağırıp, birkaç soru sorması ve genellikle sözlü olarak verilen cevaplara not takdir etmesi de okullarımızda çok kullanılan bir sözlü sınav türüdür (Turgut, 1997). Sözlü sınavlar öğrencilerin sözlü ifade becerilerinin, yabancı dildeki kelimelerin telaffuzunun ölçülmesinde, okul öncesi eğitiminde, üniversite düzeyinde yüksek lisansa girişte, tez savunmasında ve bazı mesleki yerleştirmelerde geleneksel hali ile kullanılmaktadır. Sözlü sınavlar diğer sınav türlerine göre az kullanılır, bunun nedeni onun yapısındaki zayıflıklardır (Tekindal, 2002).

Sözlü sınavı diğer sınav çeşitlerinden ayıran özelliklerin başlıca olanları şöyle sıralanabilir (Turgut, 1997):

- Cevap sözlü olarak verilir. Böyle olunca sözlü ifade yeteneği, konuşmanın etkililiği, ses tonu ve benzeri değişkenler de puana karışabilir. Bu yeteneklerin ölçülmesini amaçlamayan bir sınavda puanların sözlü ifade yeteneklerinden etkilenmesi, sınavın geçerliğini düşürür.
- Sınavı yapan kişi ile sınava tabi tutulan kişi arasında karşılıklı ve devamlı bir etkileşim vardır.
- Sözlü sınavlarda bir defada ancak bir tek cevaplayıcı dinlenebilir. Bu sebeple, sözlü sınavların bireysel olarak uygulanması zorunludur.
- Sınav bireysel olarak uygulandığından ve sınavı tamamlamış bireylerle, henüz sınava alınmamış bireyleri birbirinden ayrı tutmak zor olduğundan her bireye ayrı sorular sorma zorunluluğu doğar.
- Sorular genellikle sözlü sorulur. Böyle olunca cevaplayıcının soruları tekrar gözden geçirmesine imkân kalmaz.
- Cevaplama çoğu halde, üzerinde iyice düşünülüp tasarlanmadan yapılır.
- Cevaplar kaydedilmediği için puanlamaya cevaplar verilirken ya da cevaplama tamamlandıktan hemen sonra yapılır. Bu durum, sözlü sınavlarda puanlamanın sadece genel izlenimle yapılmasına yol açmıştır.

Sözlü sınavların daha etkili kullanımına yönelik dikkat edilmesi gereken bazı hususlar şunlardır (Tan ve Erdoğan, 2004):

- Sorular önceden hazırlanmalı ve cevabı fazla uzun olmamalıdır.
- Sorular açık ve anlaşılır olmalıdır.
- Önceden puanlama anahtarı hazırlanmalıdır.
- Öğrenciye sınav davranışlarının neler olacağı söylenmelidir.
- Öğretmen soru sormadan önce öğrenciyi rahatlatmalıdır.

2.2.1.3.6. Eşleştirmeli Sorular

Eşleştirme tipi sorular; öğrencilere birbirine benzer fikirler, olaylar, kurallar, maddeler, sebep ve sonuçlar arasındaki ilişkileri buldurmak için kullanılan soru çeşididir (Kemertaş, 2003). Eşleştirme maddeleri, iki grup halinde verilen ve birbirleriyle ilgili olan bilgi öğelerinin belirli bir açıklamaya göre eşleştirilmesini gerektirir (Tekin, 1993). Sorular ve cevaplar iki ayrı bölümde verilir. Soruların sıra numarası rakamla, cevapların sırası harfle gösterilir. Öğrenciden her sorunun doğru cevabını bularak, o cevabın sırasını gösteren harfi ilgili sorunun baş tarafındaki yere yazması istenir. Seçmeli soru tipinden sonra en çok kullanılan eşleştirmeli soru tipidir. Bu soru tipinin hazırlanması kolaydır. Şansa az yer verir. Öğrencileri ezberlemeye zorlamaz. Objektif olup, her konuya uygulama olanağı vardır (Kemertaş, 2003).

Eşleştirmeli testlerin etkin kullanımı için şunlar önerilebilir (Tan ve Erdoğan, 2004):

- Eşleştirmeli testlerde cevap seçenekleri soru sayısından çok olmalıdır. Seçenek sayısının daha çok olması, sorunun şansla cevaplanma olasılığını düşürür.
- Her bir eşleştirmeli madde setinin ortak bir boyutu olmalıdır.
- Sorular ve seçenekler iki kolonda sunulmalıdır. Genelde, madde kökleri soldaki kolonda ve seçenekler sağdaki kolonda olmalıdır.
- Madde köklerinin ve seçeneklerinin uzunlukları benzer olmalı ve gramer bakımından eşdeğer olmalıdır. Eşleştirmeli sorularda özellikle seçenekler mümkün olduğunca kısa yazılmalıdır.
- Seçenekler belli bir sıraya ve düzene göre verilmelidir.
- Öğrencilerin madde köklerinden bir ok çıkartarak cevaplamalarına izin verilmemelidir.

2.2.1.3.7. Kısa Cevaplı Sorular

Kısa cevaplı sorulara verilecek cevaplar çok kısadır ve cevaplama işi çok az zaman almaktadır. Kısa cevaplı testlerin puanlanması oldukça kolay, çabuk ve nesneldir. Kısa cevaplı testler hemen her eğitim düzeyindeki öğrencilere uygulanabilir niteliktedir. Kısa cevap maddeleri, özellikle kavramsal bilgileri ölçmede etkilidir (Tekin, 2000).

2.2.1.4. Alternatif Ölçme ve Değerlendirme Yaklaşımları

Alternatif ölçme ve değerlendirme yaklaşımları; öğrencilerin ne bildikleri ve ne yapabildikleri hakkında bilgi veren, konu ile ilgili neler anladıklarını ortaya koyan, performanslarını değerlendiren, geleneksel olmayan yaklaşımlar olarak ifade edilmektedir (Gummer ve Shepardson, 2001; Pierce ve O'Malley, 1992). Alternatif ölçme ve değerlendirme yaklaşımları güvenilir, performans temelli, gerçekçi, yapılandırmacı ve uygulanabilir özelliklere sahiptir. Beceri ve bilginin kullanımı birbiriyle bağlantılı olan, tamamlanmış ve ortamlara uyum sağlanmış etkinlikleri ölçer. Amaca yönelik ölçme yapan standardize edilmiş, yönergeli etkinliklerden ayrıştırılabilir. Başka bir deyişle alternatif ölçme ve değerlendirme anlamlı öğrenme deneyimleri demektir. Farklı sunuş şekillerine ve öğrencilerin tercih ettikleri öğrenme biçimine olanak sağlar. Bu da esnek olduğunu göstermektedir. Öğrencilerin kendi öğrenme biçimlerini ve düşüncelerini analiz etmelerini sağlamak yoluyla kendine dönük düşünmeyi ve kendini irdelemeyi öğretir (Korkmaz, 2004).

Eğitimde ölçme ve değerlendirme, öğretim programlarının dört önemli ögesinden biridir. Öğretim programları hedefler (öğrenci kazanımları), öğrenme durumları, içerik ve ölçme değerlendirme öğelerinden oluşur. Ölçme değerlendirme diğer öğeleri de yakından etkiler. Değişen eğitim anlayışı ve yenilenen öğretim programları beraberinde ölçme ve değerlendirme yöntemlerini de etkilemiştir.

Günümüzdeki ve daha önceki ölçme değerlendirme yaklaşımlarını McMillian (1997) Tablo 2.1. 'deki gibi özetlemektedir.

Tablo 2.1. Günümüzdeki ve önceki ölçme değerlendirme yaklaşımları (McMillian,1997)

Önceleri	Günümüzde
Sonuca kıymet verme	Sürecin ölçülmesi
Birbirinden ayrılmış becerilerin ölçümü	Birbirini tamamlayan becerilerin ölçümü
Bilginin hatırlanması	Bilginin uygulanması
Yazıya dayalı görevler	Uygulamalı görevler
Tek bir doğru cevap	Birden fazla doğru cevap
Gizli veya belirsiz ölçütler	Açık ve belirli ölçütler
Öğretimin sonunda	Öğretimin tamamında
Çok az dönüt	Yeterli ve zamanında dönüt
Klasik sınavlar	Performansa dayalı ölçümler
Tek bir yöntemle ölçüm	Çoklu yöntemlerle ölçüm
Aralıklarla yapılan ölçümler	Sürekli yapılan ölçümler

Günümüzde giderek önem kazanan geleneksel değerlendirmelerden işlev ve yapı olarak farklı olan değerlendirme, genel bir kabul ile alternatif değerlendirme olarak adlandırılmaktadır. Pierce ve O'Malley'e (1992) göre alternatif değerlendirme; öğrencinin ne bildiğini gösteren ve ne yapabileceğini ortaya koyabilen bir değerlendirmedir. Amacı, öğrencinin ilerlemesini göstermek ve öğretimi bilgilendirmektir. Standart veya geleneksel bir değerlendirme değildir. Yapılandırıcı öğrenme içerisinde yer alan alternatif değerlendirme, öğrenim ortamının kendisi kadar çeşitli olmalıdır.

Reeves ve diğ. (1996) göre yapılandırıcı öğrenme ortamındaki alternatif değerlendirme türlerini Tablo 2.2. 'de görüldüğü gibi ifade etmişlerdir.

Tablo 2.2. Alternatif Değerlendirme Türleri (Reeves, 1996)

Alternatif Değerlendirme Türü	Tanımı
Otantik	<ul style="list-style-type: none">➤ Öğrencinin yaptıkları ile oluşan sahiplenme duygusu ve artan öğrenci motivasyonu söz konusudur.➤ Değerlendirme güvenilirliği: Öğrenmenin hızlı karşılığı. Şartlar/deneyim. Değerlendirmeye karşı olumlu öğrenci tutumu. Artan öğrenci motivasyonu.
Performans	<ul style="list-style-type: none">➤ Öğrenciler bir aktivite ile meşgul olarak kavrama gösterirler. Örneğin karmaşık öğrenmeye odaklanma, daha yüksek düzeyde düşünme ve problem çözme becerileri ile ilgilenme, daha geniş çapta aktiflik gösterme söz konusudur.➤ Çoklu adım gerektiren çalışmalara ihtiyaç duyar.➤ Öğrenci zaman ve çabasını gerektirir.
Öğrenci Ürün Dosyası (Portfolyo)	<ul style="list-style-type: none">➤ Ürünlerin ve öğrencinin kendi kendine değerlendirmesi için bir depolama yeridir.➤ Öğrenci ürün dosyası öğrencinin çalışma örneklerini, çalışmanın aşamalarında oluşturulan taslakları, günlük ya da yazılarını içerebilir.
Cevaplayıcı	<ul style="list-style-type: none">➤ Değerlendirme yönteminden ziyade değerlendirme aracındaki bir değişiklik olarak göz önünde bulundurulur.➤ Değerlendirmenin bu perspektifi; öğrencilerin gücünü vurgular, öğrencilere kendi hedef ve aktivitelerini oluşturabilinler diye doğrudan bilgi sağlar, öğrencilere değerlendirme sürecinde en temel söz sahibi olarak yer verir.
Bilgisayar Tabanlı Simülasyonlar (Sanal Gerçeklik: VR)	<ul style="list-style-type: none">➤ Değerlendirmenin biçimlerini değiştirme, performansı yargılama basamaklarını oluşturma ve değerlendirme için ölçüt ve standartları paylaşma yeterliliğine sahip olan yöntem olarak tanımlanır.➤ Teknolojinin fonksiyonları değerlendirmeyi etkiler; otantiklik, kabul edilebilirlik, yeniden oynanabilirlik, zengin kaynak, iletişim, uygulamalı, işbirlikçi.

Tablo 2.3. Geleneksel ve Alternatif Değerlendirme Yöntemlerinin Karşılaştırılması

Geleneksel Ölçme ve Değerlendirme Yaklaşımları	Alternatif Ölçme ve Değerlendirme Yaklaşımları
Standardize edilmiş testlere dayalıdır.	Performans temelli, gerçekçi, yapılandırıcı ve uygulanabilir testlere dayalıdır. Standardize edilmiş testleri de önemser.
Bilişsel alanla ilgili kazanılan hedef davranışları değerlendirmeye yönelik açıklamalar getirir.	Öğrenenin yeterliliklerini daha geniş bir çevrede (bilişsel, duyuşsal ve psikomotor) tanımlar.
Öğrenenlerin bilgi ve özümstediklerinden çok testlerde belirlenen hedeflere yönelik gösterdikleri performans önemlidir.	Öğrenme bireyseldir.
İlk yapılan etkinlikler öğrencilerin öğrenmeleri hakkında bilgi verir.	Öğrenme süreklidir.
Başarı gelişmeden bağımsız ele alınır.	Başarı gelişimle birlikte ele alınır.

2.2.1.4.1. Puanlama Ölçeği (Rubrik)

Eğitim-öğretim aşamasının daha en başında öğrencinin başvuracağı ölçütler geliştirilmelidir. Ölçütlerin yer aldığı puanlama ölçekleri sayesinde öğrenci öğretmenin ne istediğini anlarken öğretmen de öğrenciyi daha iyi değerlendirebilir (Enger vd., 1998). Puanlama ölçeği, performans düzeylerinin hızlı ve net değerlendirilmesini sağlar. Bu yüzden öğretmenler ve öğrenciler için çok değerli bir uygulamadır (Goodrich 1997, Popham 1997). Puanlama ölçekleri hedeflenen amaca uygun olarak bütüncül, analitik veya her ikisinin birleşimi olarak hazırlanabilirler. Bütüncül puanlama ölçekleri, bir görevin niceliğini, niteliğini veya hem nicelik hem de niteliğini tanımlayan farklı seviyelerdeki performansları içeren yapılardır. Böyle bir puanlama ölçeği için en uygun seviye belirlenmelidir. Analitik puanlama ölçekleri, farklı performans seviyelerinin kendi içinde bölümlere ayrılması ve ölçütlerin belirlenmesi ile oluşur. Analitik puanlama ölçeklerinde değerlendirilecek nitelikler sütunların ilk sırasında yer alır. Her sıra ise değerlendirilecek olan ölçütün yer aldığı bir hücreyle başlar. Her bir bitişik sıra içerisinde performansın farklı seviyelerini tanımlayan ölçütler yer alır. Bütüncül puanlama ölçeği, genellikle performansın seviyeleri hakkında daha geniş tanımlamalar içerirken, analitik puanlama ölçeği daha spesifik performans durumlarıyla alakalı daha özel ve ayrıntılı tanımlamalar içerir (Korkmaz, 2004).

Puanlama ölçeklerinin geliştirilmesinde dikkat edilmesi gereken noktalar şunlardır:

- Belirlenen ölçütler amaçlanan hedeflere uygun ve tutarlı olmalıdır.
- Ölçütler, üründe bulunması gereken özellikleri veya davranışları ifade etmelidir.
- Puanlama ölçekleri, öğrenci seviyesine uygun olarak açık ve anlaşılır bir şekilde yazılmalıdır.
- Puanlama sistemi anlamlı ve açık olmalıdır.
- Performans düzeylerindeki farklılık net bir şekilde verilmelidir (Moskal, 2000).

Puanlama ölçeklerinin avantajları şu şekilde ifade edilebilir:

- Öğretim ile değerlendirmeyi birleştirir.
- Ölçütler sayesinde en iyi durum belirlenir.
- Öğrenci çalışmalarının değerlendirilmesi ve notlandırılması işlemi aynı anda gerçekleşir.
- Öğrenme hedeflerinin açık bir şekilde belirlenmesini sağlar.
- Öğrencilere hedefleri açık bir şekilde göstererek yardım eder.
- Beklentilere uygun ürünlerin ortaya çıkmasına yardım eder.
- Geri bildirimlerde hız kazandırır. Ayrıca ayrıntılı bir dönüt sağlar (Yıldız, 2005; Whittaker vd., 2001).

2.2.1.4.2. Öğrenci Ürün Dosyası (Portfolyo)

Öğrenci ürün dosyası, öğrenci çalışmasının koleksiyonudur (Collins, 1992). Öğrenci çalışmalarının sistematik ve anlamlı olarak öğrenmeyi değerlendirme amacıyla kullanılmasına denir (Thompson, 2001). Öğrenci ürün dosyalarının oluşturulması ve değerlendirilmesi hem öğrenciler hem de öğretmen açısından üzerinde düşünme ve zaman gerektiren bir olaydır. Öğrenci ürün dosyaları net bir düşünce ve amaç için hazırlanmalıdır.

Öğrenci ürün dosyasını bir değerlendirme metodu olarak kullanmaya karar vermek için cevaplandırılması gereken sorular şunlardır (Howe vd., 1998);

- Öğrenci ürün dosyasının kullanımındaki gayesi nedir? Öğrencileri motive etmek, çalışmalarını organize etmek, yardım etmek, başarıyı değerlendirmek veya ilerlemeyi göstermek için mi kullanılacak?
- Öğrenci ürün dosyaları, puanlama sisteminin her bir parçası olarak mı kullanılacak yoksa öğrenci çalışmasının gösterilmesine vitrin mi olacak? Her

öğrenci ürün dosyası bir bütün olarak mı değerlendirilecek yoksa öğrenci ürün dosyasına giren her çalışma ayrı ayrı mı değerlendirilecek?

- Öğrenci ürün dosyasına neler dâhil edilecek? Bu kararın sorumluluğu paylaşılmalıdır; fakat kesin hatlar belirtilmezse öğrenci ürün dosyası karışık kâğıt yığınlarına dönüşebilir

Öğrenci ürün dosyası kullanılmasının avantajları şunlardır:

- Öğrencilerin, tüm alanlar için yararlı olan kendini değerlendirme stratejileri geliştirmesine yardım eder.
- Öğrenci gelişimini izler ve gelişimle ilgili kanıtları gösterir.
- Öğrencilerin sınıf içinde olduğu kadar sınıf dışında da araştırmasını ve çaba harcamasını sağlar.
- Farklı yaş seviyeleri ve farklı fen alanları dikkate alındığında öğrenci ürün dosyasına nelerin gireceği konusunda farklı seçenekler sunar.
- Öğrencilerin eleştirel düşünme becerilerini, bilgilerini, kendilerini ifade edebilme yeteneklerini ve katılımlarını artırır (Thompson, 2001; Enger vd., 1998; Oosterhof, 1999; Niguidula, 1993; Hamilton, 1994; Zollman, 1994; Gilman vd., 1995; Valencia, 1990).

McMillian'a (1997) göre portfolyo kullanımında çeşitli dezavantajlar da söz konusudur. Öncelikle, öğretmenin öğrenci ürün dosyasına konulan çalışmalara göre ölçütler geliştirmesi ve çalışmalarını bu ölçütlere göre değerlendirmesi için zaman harcaması gerekir. Buna dikkat edilmezse öğrenci ürün dosyasındaki çalışmaların puanlandırılması düşük güvenilirliğe sebep olabilir. Ayrıca, öğrencilerin portfolyolarına koyacakları çalışmaların seçimine dikkat edilmezse performansların değerlendirilmesinde sıkıntılar yaşanabilir (Thompson, 2001).

2.2.1.4.3. Performans Ödevi

Performans ödevleri, öğrencinin neyi bildiğinden çok neyi yapabileceğini değerlendirme amacını taşır. Buradaki fark, sürücü belgesi almak için yazılı sınav ile direksiyon sınavı arasındaki fark gibidir. Yazılı sınav sürücü adayının kurallar hakkındaki bilgisini gösterir ve yol testi otomobilin direksiyonda neyi yapabildiğini gösterir.

Performans ödevleri hakkında genel olarak şunları ifade etmek mümkündür (Pierce vd., 1992):

- Alternatif değerlendirme olarak dikkate alınmalıdır.
- Genellikle öğrencinin özel becerilerini ve yeterliliğini gösterebileceği performans çalışmalarına yöneliktir.
- Öğretim çalışmalarında öğrenci performansını yansıtır ve ölçütlerinin belirlenmesi gerekir. Çalışmalarının sunumunda, değerlendirmeyi yapan öğretmen geri iletimde bulunabilir ve öğrencinin ihtiyaç duyduğu durumlarda onu destekleyebilir.

Performans değerlendirmesi 3 veya 4 kişilik gruplara uygulanabilir. Çünkü gerçek hayatla ilgili olan bu çalışmalarda, muhtemelen tek bir kişinin bu işi yapma becerisi, zamanı ve motivasyonu bu iş için yetmeyecektir (Enger vd., 1998).

2.2.1.4.4. Akran Değerlendirmesi

Akran değerlendirmede bir öğrenci başka bir öğrencinin çalışmasını değerlendirir. Akran değerlendirmesi öğrencilerin her çeşit performansları için kullanılabilir. Akran değerlendirmede değerlendirme yapan öğrenci hangi arkadaşını değerlendirdiğini bilmezse değerlendirme daha objektif olur (Bahar vd., 2006). Akran değerlendirmenin birçok avantajı vardır; derse katılımı artırma, öğrenciler arasındaki etkileşimi artırma, iletişim becerilerini artırma, değerlendirmedeki gayeyi keşfetme, öz değerlendirme becerilerini geliştirme bunlardan bazılarıdır.

2.2.1.4.5. Grup Değerlendirmesi

İşbirliğine dayalı öğrenme faaliyetlerinin birçoğunda grup çalışması yapılır. İşbirliğine dayalı öğrenme, öğrencilerin oluşturduğu küçük grupların işbirliği içerisinde ortak öğrenme amaçlarını en üst düzeyde gerçekleştirmek amacıyla kullanılan bir öğrenme ve öğretme tekniğidir (Bahar vd., 2006). Grup değerlendirmesi hem öğrenciler hem de öğretmen tarafından gerçekleştirilir. Grup değerlendirmesi öğrenci etkileşimini zenginleştirir ve birbirlerinden öğrenmeyi sağlar. Aynı zamanda birlikte çalışma alışkanlığı kazandırır. Bu tür değerlendirmelerde dönüt vermek önemlidir. Çünkü bu sayede öğrencilerin sosyal sorumlulukları ve başarılarında gelişme sağlanır (Ornstein vd., 2004).

2.2.1.4.6. Öz Değerlendirme

Kendi öğrenme süreçlerini, özellikle başarı düzeylerini ve öğrenme sonuçlarını değerlendirmeleri olarak açıklanabilir. Kısacası öğrencinin kendisini değerlendirmesi olarak tanımlanabilir. Öğrencilerin neleri öğrendiğini veya hangi alanlarda problemi olduğunu belirlemesi ve gelişimlerine paralel olarak sorumluluğu üstlenmesi öz değerlendirmenin amacıdır (Bahar vd., 2006). Etkin bir öz değerlendirme öğrencilerin özgüvenlerini geliştirir ve öğrendiklerinin farkında olmayı sağlar. Öğrencinin kendi güçleri ve becerilerinin farkında olmasının yanında öğrenme yaklaşımı hakkında değerlendirme yapmasını sağlar.

2.2.1.4.7. Kavram Haritası

Kavram haritaları; öğrenme, öğretme etkinliklerinde geniş kullanım alanına sahip olan görsel bir yöntemdir. Bu önemli yöntemi Ausubel'in bilişsel öğrenme kuramı ışığında ilk defa Joseph Novak tanıtmıştır. İnsanlar iletişim için çeşitli yollar kullanmaktadırlar; doğal diller (konuşma-yazma), müzik, resimler- şemalar gibi. Kavram haritaları ise bilgi iletişiminin grafiksel bir yoludur. Öğrenme-öğretme yöntemi olarak geniş kullanım alanı bulan kavram haritaları, kavram yanılıklarını belirlemede de kullanılabilir (Baki, Şahin 2004). Kavram haritaları, bilgiyi organize etmek ve sunmak için yapılmış grafiksel araçlardır. Bu araçlar daire ya da bir çeşit kutu içine yazılmış olan kavramları içerir. Kavram haritalarında iki kavram arasındaki ilişki, üzerine ilişkiyi belirleyen ifadelerin yazıldığı doğrularla gösterilir. İlişkiyi belirleyen bağlantı ifadeleri ile iki kavram tamamlanarak anlamlı bir cümle oluşturur (Novak, 1998). 1990'dan beri kavram haritası farklı konu alanlarında yapılan araştırmalarda kullanılmıştır. Barenholz ve Tamir (1992), Trowbridge ve Wandersee (1994) kavram haritasını fen öğretiminin değerlendirmesinde kullanmışlardır. Hegarty-Hazel ve Prosser (1991), McClure, Sonak ve Suen (1999)'in çalışmaları ise daha çok kavramsal öğrenmeler ile çalışma yöntemleri arasındaki ilişkileri ortaya koymaya yöneliktir. Bütün bu araştırmalarda bir öğretim metodu olarak kavram haritasının geçerlik, güvenilirlik ve uygulanabilirliği kanıtlanmıştır (Duru, Gürdal; 2002).

Kavram haritaları (KH) kavram ağlarına benzer grafik araçlardır; ancak, onlardan farklı olarak kavram haritalarında kavramlar arası ilişkiler önermeler veya ilkeler olarak yer alır.

Kavram haritalarının yapımında izlenmesi önerilen genel kurallar aşağıdaki gibi sıralanır (Martin, ve diğ., 1994, s. 89-91; BSCS, 1994, s. CM1-CM14).

- Öğretilecek konunun kavramları listelenir. Kavramlarla ilgili açıklama gerekmez. Eşya ve olayların tekil örnekleri, özel adlar kavram olmadıkları için bu listeye alınmaz. İlkeler ve kavramlar arası ilişkiler de bu listeye dâhil değildir.
- Kavramlar listesinden en genel veya en üst düzeyde olan sözcük ayrı bir sayfanın başına yazılır. Bu bir kavram olabileceği gibi bir tema da olabilir. Bundan sonra öğretilmek istenen ilişkili kavramlar aşamalı bir düzende sayfaya yerleştirilir. Düşey düzenlemede en genel kavram en üstte, eşit genellikteki kavramlar aynı satırda, diğerleri genellik derecelerine göre azalan sırada sayfanın altına doğru sıralanır. KH aşamalılığı öğreteceği için bu sıralama önemlidir. Her kavram haritada yalnız bir kez yer almalıdır.
- Kavramlar haritadaki diğer sözcüklerden kolayca ayırt edilebilmelidir; bunun için kavramlar kutu veya yuvarlak içine alınır.
- Öğretilmek istenilen kavramlar arası ilişkiler, genelleme ve ilkeler ayrıca listelenir.
- Kavram haritasında iki kavram arasındaki ilişkiyi göstermek üzere iki kutu bir çizgi ile bağlanır. İlişki bu çizginin üzerine birkaç kelimelik bir ibareyle yazılır. Bu ilişki haritadaki kavramlardan en az birini ilgilendiren bir önermedir. İlişkiler ve ilkeler kutulanmaz. Bazı hallerde ilişkinin yönü önemli olduğu için belirtilecek ilişki yönü ok ile gösterilir. İlişkileri içermeyen bir kavram haritası daha ziyade bir akış diyagramına benzer, öğretimde yeterince etkili olmaz.
- Kavram haritası gereğinden fazla şişirilmemelidir. Harita başlangıçta basit tutulmalıdır. Harita çok sayıda kavramı, ilişkiyi ve ilkeyi içeriyorsa önce en önemli elemanları topluca gösteren bir genel harita, sonra genel haritanın bölümlerini ayrı ayrı gösteren ayrıntılı haritalar yapılmalıdır.

Kavram haritaları bir olayı veya konuyu topluca gösteren, kavramları, kavramlar arası ilişkileri ve ilkeleri kısaca belirten araçlardır. Doğru yapılmaları halinde öğretimin her basamağında kullanılabilir. Haritalar tüm sınıf etkinliğinde veya küçük grup etkinliklerinde

öğrencilerin katılımlarıyla geliştirilebilir. Ayrıca, KH hazırlandığı seviyeye göre kelimeler içermelidir. Örneğin su için hazırlanan ilköğretim seviyesindeki bir haritada suyun nerelerde kullanıldığını, yapı birimlerinin ne olduğunu ve halleri ana gruplar olarak seçilebilir. Fakat aynı harita ortaöğretim seviyesi için yapıldığında suyun KH çok farklı olur. Çünkü bu düzeyde vurgu çözücülüğe, iyonlaşmaya ve polarizasyona yapılabilir.

Novak ve Gowin (1984) KH'nin aşağıdaki durumlarda kullanılabilceğini belirtmektedirler.

- Bilgileri organize hale getirmede,
- Öğrencilerle kavramların anlamlılığını tartışmada,
- Yanlış anlamaları gidermede ve
- Yüksek seviyeli düşünme yeteneği geliştirmede

2.2.1.4.8. Kontrol listesi

Kontrol listesi, öğrencilerin öğrenme aktiviteleri sırasındaki davranışların gözlenmesi için kullanılır. Hangi davranışlara dikkat edileceğinin önceden planlanması onları bir kâğıt üzerinde listelemenin iki fonksiyonu vardır (Ornstein vd., 2004):

- Gözleme odaklanmak
- Kaydetmeye kolaylık sağlamak

Davranış kontrol listeleri herhangi bir amaç alanında kullanılabilir olsalar da özellikle bilişsel olmayan alanlarda kullanılırlar. Çünkü bu amaçların doğrudan değerlendirilmesi daha zordur. Kontrol listesi hazırlanırken derecelendirme çizelgesinden daha çok olumlu ya da olumsuz düşüncenin ifadesi olarak “+” ve “-” işareti kullanılmalıdır. Böylelikle kontrol listesinin kullanımı kolaylaşır. Dikkat, listeye kaymadan kullanılabilir ve öğrencilerin davranışlarının kaydedilmesi kolay olur (Ornstein vd., 2004).

2.2.1.4.9. Gözlem Formu

Gözlem, öğrencinin bir konudaki yeteneğinin değerlendirilmesinin bir yoludur; fakat çok nadir olarak bir gözlem eylemi bir değerlendirme biçimi olarak görülür. Eğer değerlendirme, öğrencinin sınıftaki performansının bütünü üzerine odaklanırsa laboratuvarında veya sınıfta öğrenci performansının gözlemsel kayıtlarını tutmak suretiyle öğrencinin çalışmasını değerlendirebilirse çok kıymetli olur (Enger vd., 1998).

Öğretmen, öğrencilerin farklı ödevlerini tek başlarına ya da grup halinde yaparlarken onları izleme sansına sahiptir. Öğretmen sınıfta öğrenciler ile az ya da çok göz teması halindedir. Öğretmenin bu gözlemi unutmaması ve değerlendirmesi açısından birtakım ölçütleri kayıt altına alması gerekir. Öğretmen hazırladığı gözlem formu ile bütün öğrencileri değerlendirmeyi amaçlasa da öğrenciler arasında farklı davranış sergileyen ve özel çalışma gerektiren öğrencilere yoğunlaşır (Ornstein vd., 2004). Gözlem aynı zamanda öğretmeni hangi öğrencinin laboratuvarı kullanmayı, not almayı, bu notlar için nereden bilgi edinmesi gerektiğini, bilgisayarı kullanmayı, interneti kullanmayı vb. bilip bilmediği hakkında bilgilendirir. Değerlendirme, tüm öğrencilerin değerlendirilmesini dikkate alıyorsa gözlem değerlendirmede anahtar bir rol oynayabilir ve gözlem, öğretmenin öğrencilere maksimum başarıya ulaşmalarında yardım etmesi için sürekli rehberlik sağlar (Enger vd., 1998). Gözleme bağlı değerlendirme önerileri, öğrencinin sınıftaki faaliyetlere katılımının daha anlamlı bir resmini sunabilir.

Bir gözlem formu gözlemlenen davranışların dokümantasyonu için yararlı olabilir. Gözlem formu içerisinde yer alabilecek bazı ifadeler şunlardır (Enger vd., 1998):

- Meraklı, öğrenmeye karşı olumlu tutuma sahip, araştırmacı ruha sahiptir.
- Bireysel olarak çalışmalarını sürdürür.
- Üretken alternatifler ve / veya ilaveler sunar.
- Şahsi sorumluluk gösterir.
- Zamanı iyi kullanmaya örnek oluşturur.
- Laboratuvar araç ve gereçlerini iyi kullandığını gösterir.
- Sınıf aktivitelerine hazırdır.
- Grup sorumluluğu alır.
- Grup anlayışına katkıda bulunur.
- Fikirlere, görüşlere karşı açıktır ve isteklidir.
- Grup çalışmasının sürdürülmesine yardımcı olur.

2.2.1.4.10. Yapılandırılmış Grid

Yapılandırılmış grid alternatif ölçme değerlendirme tekniklerinden biridir. Bu teknik uygulanırken yaşa ve seviyeye bağlı olarak dokuz ya da on iki kutucuktan oluşan bir şekil hazırlanır. Tablo 2.4 'de buna bir örnek verilmiştir.

Tablo 2.4. Yapılandırılmış Grid

1	2	3
4	5	6
7	8	9

Yapılandırılmış Gridin Genel Yapısı

Konu ile ilgili kavramlar, resimler, sayılar, eşitlikler, tanımlar veya formüller gelişigüzel kutucuklara yerleştirilir. Kutucukların içeriğinin değiştirilebilmesi hem görsel hem de analitik düşünebilme olanağı sağlar. Öğrencilere konuyla ilgili değişik sorular verilir.

Öğrencilerden,

- Her sorunun cevabı için uygun kutucukları bulmaları,
- Bu kutucuk numaralarını mantıksal veya işlevsel sıraya göre dizmeleri istenir.

Gridi hazırlamak üzere öğretmen kendisine bir soru sorar ve bu sorunun cevabını gelişigüzel kutucuklardan birine veya birkaçına yerleştirir. Sonra 2. soruyu sorar ve cevabını gene kutucuklara yerleştirir; fakat 2. sorunun cevabını teşkil eden kutucuklardan bir kısmı birinci sorunun da cevapları arasında olabilir. Diğer bir deyişle 2. sorunun cevabının bir kısmı 1. sorunun da cevabı olabilir. Bu şekilde kutucukların tamamı doluncaya kadar soru hazırlanarak cevaplar kutucuklara dağıtılır. Sonuçta öğrencilerden; her sorunun cevabı için uygun kutucukların bulunması ve bu kutucuk numaralarının mantıksal veya işlevsel sıraya göre dizilmesi istenir. Öğrencilerin verdiği cevap o konudaki bilgi seviyesini, kavramsal bağları, yanlış kavramları ve bilgi eksikliğini gösterir (Bahar, 2002).

Her sorunun cevabı için uygun kutucukların bulunması aşamasında aşağıdaki formül uygulanır:

$$\frac{C1}{C2} - \frac{C3}{C4}$$

C1= Doğru seçilen kutucuk sayısı

C2= Toplam doğru

C3= Yanlış seçilen kutucuk sayısı

C4= Toplam yanlış kutucuk sayısı

Bu formüle göre öğrencilerin puanları -1, 0 ve +1 arasında değişir. Bu puanı on üzerinden değerlendirmek için önce negatifliği ortadan kaldırmak amacı ile 1 ile toplanır ve elde edilen sayı 5 ile çarpılır (Bahar, 2002). Kutucuk numaralarının mantıksal veya işlevsel sıraya dizilmesinin değerlendirilmesinde aşağıdaki puanlama sistemi kullanılır. Öncelikle doğru sıralanmış kutucuk numaraları için (Ek-5 deki 1. sorunun cevabı 8, 6 ve 2 dir.) (Atıcı ve Karahan 2006). Soru 1: 8 numara 6'dan önce mi? Cevap evet ise art arda mı geliyor? Soru 2: 6 numara 2'den önce mi? Cevap evet ise art arda mı geliyor? Bu 2 soruda 4 tane "evet" veya "hayır" cevabı alınır. Her "evet" cevabı "2,5" ve her "hayır" cevabı "0" ile değerlendirilirse bir öğrenci 4 "evet" için 10 puan alır.

2.2.1.4.11. Tutum Ölçeği

Tutum, bireylerin belli bir kişiyi, grubu, kurumu veya bir düşünceyi kabul ya da reddetme şeklinde gözlenen, duygusal bir hazır oluşluk hali veya eğilimdir. Tutumların ölçülmesinde gözlem, soru listeleri, tamamlanmamış cümleler ve hikâyeler anlatma gibi çeşitli yöntemler ile yanlışı seçme tekniği, içerik analizi gibi çeşitli teknikler de kullanılmaktadır (MEB, 2005). Tutum ölçekleri, tutum ölçme yöntemleri içerisinde en önde gelen ve yaygın olarak kullanılanıdır. Tutum ölçeklerinden en çok kullanılan yöntem de likert ölçeğidir. Likert tipi ölçeklerle, ölçülmek istenen tutumla ilgili çok sayıda olumlu ve olumsuz ifade yazılır. Bu ifadeler için, "Tamamen Katılıyorum", "Katılıyorum", "Kararsızım", "Katılmıyorum" ve "Kesinlikle Katılmıyorum" biçiminde tepkide bulunulur. Böylece her cevaplayıcı, ölçekteki her ifadenin kapsadığı tutum objesine katılma derecesini bildirmiş olur (MEB, 2005).

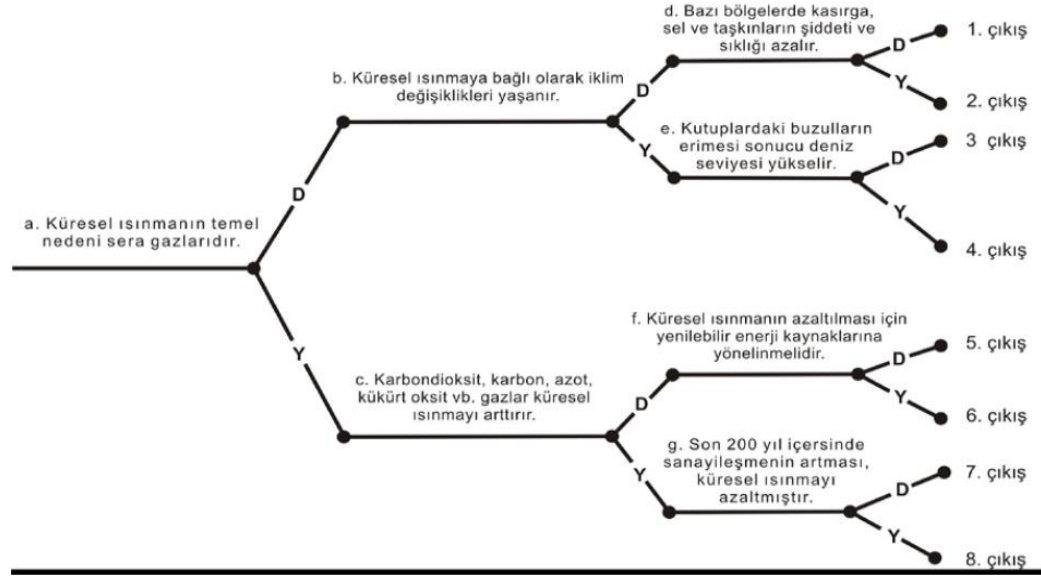
2.2.1.4.12. Tanılayıcı Dallanmış Ağaç

Belli bir konuda öğrencinin neleri öğrendiğini ve neleri öğrenemediğini belirlemek için kullanılabilir değerlendirme araçlarından biridir. Bu teknikte, temelden ayrıntıya giden bir sıraya göre doğru ve yanlış ifadeler seçilerek öğrenciden doğru seçimi yapması istenir. Böylece, 8 veya 16 seçimlik bir ifadeler listesi ile sonlanan bir dallanmış ağaç oluşturulur.

Bu yöntemin sağlayacağı avantajlar şu şekilde sıralanabilir;

- Öğrencilerin hangi önermelerde yanlışlıklar yaptığını belirlemede
- Öğrencide var olan kavram yanlışlıklarını ortaya çıkarmada

- Öğrencilerin eksik olduğu ya da yanlış öğrendiği konuları belirlemede
- Öğrencinin ön bilgilerini belirlemede
- Öğrencilerin önermelerden yararlanarak öğrenmenin gerçekleştirilmesinde



Şekil 2.1. Tanılayıcı Dallanmış Ağaç (<http://ttkb.meb.gov.tr/>)

Bu yöntemin değerlendirilmesinde ise;

- Öğrenci 1. çıkışa ulaştı ise; a ifadesine “D” diyerek doğru yanıt vermiş ve b ifadesine ulaşmıştır. “b” ifadesine “D” diyerek doğru yanıt vermiş ve d ifadesine ulaşmıştır. “d” ifadesine “Y” diyerek doğru yanıt vermiştir. Bu durumda öğrencinin 3 doğru yanıtı vardır ve 3 puan almıştır.
- Öğrenci 3. çıkışa ulaştı ise; a ifadesine “D” diyerek doğru yanıt vermiş ve b ifadesine ulaşmıştır. “b” ifadesine “Y” diyerek yanlış yanıt vermiş ve e ifadesine ulaşmıştır. e ifadesine “D” diyerek doğru yanıt vermiştir. Bu durumda öğrencinin 2 doğru yanıtı vardır ve 2 puan almıştır.
- Öğrenci 7. çıkışa ulaştı ise; a ifadesine “Y” diyerek yanlış yanıt vermiş ve c ifadesine ulaşmıştır. “c” ifadesine “D” diyerek yanlış yanıt vermiş ve g ifadesine ulaşmıştır. g ifadesine “D” diyerek doğru yanıt vermiştir. Bu durumda öğrencinin 1 doğru yanıtı vardır ve 1 puan almıştır.
- Öğrenci 8. çıkışa ulaştı ise; a ifadesine “Y” diyerek yanlış yanıt vermiştir ve c ifadesine ulaşmıştır. “c” ifadesine “D” diyerek yanlış yanıt vermiş ve g ifadesine

ulaşmıştır. g ifadesine “Y” diyerek yanlış yanıt vermiştir. Bu durumda öğrencinin doğru yanıtı yoktur ve 0 puan almıştır.

2.2.1.4.13. Poster Değerlendirme

Posterler bir projeyi daha önceden proje ve sistem hakkında hiçbir bilgisi olmayan izleyicilere ana hatları ile tanıtıcı nitelikte hazırlanan iki boyutlu grafiklerdir (Aydogdu vd., 2005).

Posterin hazırlanmasında dikkat edilecek noktalar şunlardır (Yaman vd., 2005):

- Posterin genişliği 60–70 cm, yüksekliği ise 90–100 cm aralığında olmalıdır.
- Posterler en az bir metre uzaklıktan okunabilecek şekilde düzenlenmelidir.
- Poster; giriş, amaçlar, bulgular ve sonuçlar şeklinde bölümlere ayrılmalıdır. Bu bölümler içerisindeki geçişlere ve vurgulara özen gösterilmelidir.
- Posterdeki yazılar ve konular, bütün bir ifade oluşturmalıdır.
- Posterde konu ile ilgili resimler, şekiller, tablolar, yazılar ve önemli bilgiler arasında bir bütünlük oluşturulmalıdır.

2.2.1.4.14. Mülakat (Görüşme)

Görüşme, öğretmen ve öğrenci arasında geçer. Buradaki amaç, öğrencinin bir kavram, durum veya olay hakkındaki anlayışını derinlemesine incelemek ve analiz etmektir (Bahar vd., 2006). Öğretmen basit bir yol izleyerek görüşmenin güçlü bir değerlendirme aracı olmasını sağlayabilir.

Enger vd. (1998), öğretmenin görüşmede izleyeceği yolu şu şekilde ifade etmişlerdir:

- Görüşmeyi planla. Mümkünse somut objeler, durumlar ve resimler kullan.
- Öğrenciye soracağın uygun soruları seç. Evet-hayır sorularından kaçın. Doğru-yanlış korkusu olmadan konuşmalarına izin ver. Öğrenciden tanım, tarif, tahmin ya da açıklama iste.
- Soruları kolaydan zora doğru sırala ve öğrencinin güvenini sağla. Bu sayede öğrenci görüşme sırasında rahat bir tavır sergiler.
- Öğrenciye soruları yanıtlaması için yeterli zaman verilmelidir. Öğrenci cevap vermezse veya soruyu anlamadığını söylerse soru başka bir şekilde bir kez daha sorulmalıdır.

- Öğretmen ve öğrenci için rahat olan bir ortam seçilmelidir.
- Görüşmede bir önceki görüşmeden bahsedip öğrenci ile yakınlık sağlanmalıdır.
- Sorular anlaşılmadığı takdirde farklı ifadeler kullanılarak sorulmalı veya daha fazla bilgi almak üzere öğrencinin dili kullanılmalıdır.
- Son olarak, görüşme pozitif bir izlenimle sonuçlanmalıdır. Öğrencinin işbirliği, davranış tarzı ve sorulara verdiği cevaplar not edilebilir, öğrencinin aklına gelen sorulara cevap verilebilir.

3. MATERYAL VE METOT

Araştırmanın bu bölümünde araştırma modeli, evren, örneklem, veriler, verilerin toplanması, verilerin çözümlenmesi ve verilerin yorumu konularına yer verilmiştir.

3.1. Araştırmanın Modeli

İlköğretim öğretmen adaylarının ölçme ve değerlendirme okuryazarlıkları ve ölçme-değerlendirmeye ilişkin tutumlarını araştırmak için çeşitli değişkenlere göre yapılan bu araştırma, tarama türü bir araştırmadır. Tarama modelleri, geçmişte ya da halen var olan bir durumu var olduğu sekliyle betimlemeyi amaçlayan araştırma yaklaşımıdır (Karasar, 2003). Araştırmada, cinsiyet, öğrenim gördükleri program çeşitli bağımlı değişkenlere göre incelenmiştir. Ayrıca bu değişkenlerin birbirleri arasındaki ilişkiler çeşitli istatistikî yöntemlerle ortaya konmuştur.

3.2. Evren ve Örneklem

3.2.1. Pilot Çalışma

Bu araştırmanın ilk kısmının evrenini Fırat Üniversitesi eğitim fakültesinde öğrenim gören öğrencilerden oluşmaktadır. Örnekleme ise fen bilgisi öğretmenliği 3. sınıf öğrencileri oluşturmaktadır. Pilot çalışma verileri bu gruptan toplanmıştır. Çalışmanın örnekleme ise; I. ve II. öğretimde toplam 89 3. sınıf öğrencisi oluşturmuştur Bu öğrenci grubu bu dönem içerisinde ölçme ve değerlendirme dersi almaktadırlar. Pilot çalışma ile ilgili istatistikî veriler sonuçlar kısmında detaylı bir şekilde açıklanacaktır.

Tablo 3.1. Pilot Çalışma Örneklem

Sınıf	Kız		Erkek		Toplam	
	N	%	N	%	N	%
Fen Bilgisi Öğretmenliği 3. Sınıf Öğrencileri	37	41,5	52	58,5	89	100

3.2.2. Ana Çalışma

Araştırmanın ana çalışması, yani ikinci kısmının evreni Fırat Üniversitesi İlköğretim Bölümü programlarının oluşturduğu öğretmen adaylarıdır. Ana çalışmanın örneklemini ise, fen bilgisi öğretmenliği, sosyal bilgiler öğretmenliği, Türkçe öğretmenliği ve matematik öğretmenliği programlarında öğrenim gören 180 öğretmen adayı oluşturmaktadır.

Tablo 3.2. Ana Çalışma Örneklem

Sınıflar	Cinsiyet				Ara Toplam	
	N	Erkek %	N	Kız %	N	%
Fen Bilgisi Öğretmenliği	28	51,8	26	48,2	54	100
Sosyal Bilgiler Öğretmenliği	29	52,7	26	47,3	55	100
Türkçe Öğretmenliği	20	46,5	23	53,5	43	100
Matematik Öğretmenliği	14	50	14	50	28	100
Ana Σ Toplam	91	50,5	89	49,5	180	100

3.3. Veri Toplama Teknikleri

Bu bölümde veri toplama araçları ve bunlara ilişkin detaylı bilgiler verilmiştir.

3.3.1. Sınıfta Ölçme ve Değerlendirme Okuryazarlık Envanteri (SDOE)

Bu araştırmada, Amerikan Eğitimde Ölçme Birliği ve W.K. Kellogg Derneği işbirliğiyle Nebraska-Lincoln Üniversitesinde, Barbara S. Plake & James C. Impara tarafından hazırlanan” *Teacher Assessment Literacy Questionnaire*” (1993) “Değerlendirme Okuryazarlığı Envanteri” Türkçeye uyarlanarak uygulanmıştır. Bunun yanında, öğretmen adaylarının ölçme değerlendirmeye ilişkin tutumlarının belirlenmesi için araştırmacı tarafından öğretmen adaylarıyla yapılan mülakatlar sonucunda “ ölçme ve değerlendirmeye ilişkin tutum anketi” oluşturulmuş ve uygulanmıştır.

Değerlendirme Okuryazarlığı Envanteri, kendi içerisinde yedi temel öğretmen yeterliliğini kapsamaktadır. Her bir yeterlilik alanı ile ilgili 5 soru vardır. Aşağıdaki Tablo 3.3’te bu yeterlilik alanlarının hangi soru numaralarıyla olduğu verilmiştir.

Tablo 3.3. Ölçme ve Değerlendirme Okuryazarlığı Envanter İçeriği

<i>Ölçme ve Değerlendirme Standartları</i>	<i>Soru numaraları</i>
1. Uygun ölçme ve değerlendirme yöntemlerini seçme	1, 2, 3, 4 ve 5
2. Uygun ölçme ve değerlendirme yöntemlerini geliştirme	6, 7, 8, 9 ve 10
3. Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama	11, 12, 13, 14 ve 15
4. Eğitimsel kararlar alırken ölçme ve değerlendirme sonuçlarını kullanma	16, 17, 18,19 ve 20
5. Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma	21, 22, 23, 24 ve 25
6. Ölçme ve değerlendirme sonuçlarını paylaşabilme	26, 27, 28, 29 ve 30
7. Ölçme ve değerlendirmede kanun dışı ve ahlaki olmayan durumları bilme	31, 32, 33,34 ve 35

3.3.1.1. Envanterin Türkçeye Çevrilmesi, Uyarlanması ve Geliştirilmesi

Araştırma amacına uygun hale getirilmesi için envanter öncelikle Türkçeye çevrilmiştir. Çeviri yapılırken, beş araştırmacı ve biri İngilizce alan uzmanı olmak üzere 6 çeviri araştırmacılar tarafından birbirlerinden bağımsız bir şekilde yapılmış ve daha sonra her bir çeviri diğerleriyle karşılaştırılmıştır. Envantere ilgili tüm uzmanlar ortak bir kararla son şeklini vermişlerdir. Türkçeye uyarlanmış envanter, orijinal test formatına göre biçimlendirilmiştir. Yani yönergelerin şekli, soruların ve işaretleme yapılacak ölçek kısmının biçimi orijinaline uygun bir şekilde yapılmıştır. Burada amaç dil geçerliği için korelasyon analizlerinin yapılabilmesidir. Sadece ölçek sorularının değil, ayrıca ölçeklendirmelerinde tutarlı olması amacıyla format uygunluğunun aranması gerekmektedir. Ölçeğin dil geçerliliği uzmanlar tarafından sağlanmıştır.

Envantere son şekli verildikten sonra 20 öğretmen adayına (araştırmaya katılmayacak olan) envanter dağıtılmış, envanteri anlama ve okuma yönünden incelemeleri istenmiştir. Öğretmen adayları anketlerle ilgili fikirlerini sunduktan sonra verdikleri dönütler ışığında envanter yeniden düzenlenmiştir. Envanterde birçok kültürel öge ve eğitim sistemi farklılığından kaynaklanan maddeler belirlenmiş ve ülkemiz milli eğitim sistemine uygun hale getirilmiştir. Araştırmanın bu kısmındaki son çalışma ise Türkçe açısından dil geçerliliğinin olup olmadığı hakkındadır. İki Türkçe uzmanı envanteri okumuş ve dil geçerliliğiyle ilgili eleştirilerini yapmışlardır. Bu eleştiriler doğrultusunda envanterden Türkçeye uygun olmayan cümle ve kelime yapıları çıkarılmış ve gerekli eklemeler yapılmıştır. Görünüş geçerliliği sağlanan anket, pilot çalışma için uygun hale getirilmiştir.

3.3.1.2. Pilot Çalışma

Bu araştırmanın pilot çalışması şubat ayı içerisinde, 130 öğretmen adayına uygulanmıştır. Uygulama yapılırken öğrencilerin envanterle ilgili eleştirileri de alınmıştır. İlk uygulama olduğu için öğrenciler envanterle ilgili fikirlerini açık bir şekilde sunmuşlardır. Bu uygulama sonucunda envanterin güvenilirlik, geçerlik araştırmaları yapılmış, maddelerin tek tek ayırt edicilik ve madde güçlük analizleri ortaya konmuştur. Sorun olan maddeler üzerinde değişiklikler yapılmıştır. Bu konuyla ilgili daha ayrıntılı bilgi sonuçlar bölümünde açıklanacaktır.

3.3.1.3. Envanterin Düzenlenmesi

Envanter ile ilgili güvenilirlik ve geçerlilik analizleri yapılmış, ve tekrar envanter üzerinde sorun olduğu varsayılan maddeler ve maddelerin cevap seçenekleri üzerinde çeşitli çalışmalar ve tartışmalar yapılmıştır. Envanterin ilk ve son halleri eklerde verilecektir.

3.3.1.4. Ana Çalışma

Envanter ana çalışmaya uygun hale geldikten sonra, 180 öğretmen adayına uygulanmıştır. Envanterde maddelerin güçlük indeksleri dikkate alınarak bazı maddeler üzerinde çeşitli tartışmalar yapılarak düzeltilmiştir. Envanter ile ilgili ikinci bir dil geçerliliği yapıldıktan sonra envanter uygulamaya hazır hale gelmiştir. Aşağıdaki Tablo 3.4. envanterin hazırlanış takvimini göstermektedir.

Tablo 3.4. Araştırma Zaman Çizelgesi

<i>Aralık</i>	<i>Ocak</i>	<i>Şubat</i>	<i>Mart</i>	<i>Nisan</i>
Envanterin Türkçeye çevrilmesi				
	Envanterin Türkçeye Uyarlanması ve Geliştirilmesi			
		Pilot Çalışma		
			Envanterin Tekrar Düzenlenmesi	
				Ana Çalışma

3.3.2. Ölçme ve Değerlendirmeye İlişkin Tutum Ölçeği

3.3.2.1. Ölçek Maddelerinin Oluşturulması

Literatürde ölçme değerlendirmeye ilişkin birçok tutum ölçeği bulunmaktadır. Bu ölçekler öğretmenlerin ve öğretmen adaylarının ölçme ve değerlendirmeye ilişkin tutumlarını belirlemek için oluşturulmuşlardır. Bu tür ölçeklerdeki temel problem ölçülen fikrin tutum mu yoksa algı mı olduğudur. Bir öğretmen adayının ölçme ve değerlendirmede alternatif ölçme ve değerlendirme yöntemlerini kullanması onun algısıdır. Başka bir deyişle bu durum öğretmen adayının tercihini gösterir. Oysaki tutum, psikolojik bir objeye yönelen olumlu veya olumsuz bir yoğunluk sıralaması ve derecelemesidir (Thurstone, 1931). Bu nedenle bir tutum ölçeği oluşturulurken öğrencilerin ölçülecek konuyla ilgili olumlu veya olumsuz fikirleri sorulmalıdır. Bu amaçla 20 öğretmen adayıyla yapılan mülakatlar sonucunda onların ölçme ve değerlendirmeye ilişkin tutumlarını ortaya koyacak sorular sorarak anket maddeleri oluşturulmuştur. Örneğin ölçme ve değerlendirmeye ilişkin sevgi boyutunda çeşitli cevaplar doğrultusunda anket maddeleri öğrencilerin kendi ağızlarından ortaya çıkmıştır. Anket 4 faktörden oluşmaktadır. Bunlar, özel ilgi, endişe-kaygı, önem ve yeterlilik alt boyutlarıdır. Ankette olumlu veya olumsuz ifadelerin eşitliği göz önünde bulundurulmuştur. Tablo 3.5.'te faktörler, soru numaralarıyla verilmiştir. Yapılan güvenilirlik analizleri sonucunda bazı maddelerin ankette çıkarılmasına karar verilmiştir. Çıkarılan maddeler koyu harfler belirtilmiştir.

Tablo 3.5. Ölçme ve Değerlendirmeye İlişkin Tutum Anketi İçeriği

<i>Ölçme ve Değerlendirmeye ilişkin tutum Faktörler;</i>	<i>Faktör# Soru numaraları</i>
Ölçme ve değerlendirmeyi yapmadaki ilgi seviyesi (specific interest)	1, 4* , 10, 13, 15, 20, 21, 23, 24, 28, 31, 33, 36, 39, 41 ve 45*
Ölçme ve değerlendirmeyi yaparken duyulan endişe ve korku	2, 5, 8, 11* , 16, 18* , 22, 26, 34, 37 ve 44
Ölçme ve değerlendirmeyi yapmanın önemi	6, 9, 14* , 17, 29, 35, 40, 42 ve 46
Ölçme ve değerlendirmeyi yapmadaki yeterlilik	3, 7, 12, 19, 25, 27, 30, 32, 38, 43 ve 47

*Güvenilirlik analizi sonucunda ölçekten çıkarılması uygun görülen maddeler

3.4. Verilerin Toplanması

Veriler, 2010-2011 eğitim öğretim yılı bahar döneminde Fırat Üniversitesi İlköğretim Bölümü öğrencilerinden alınmıştır. Envanter ve tutum ölçeği ile ilgili verilerin alınması için, dekanlık kanalıyla gerekli izinler alınmıştır. Bölümdeki öğretim üyelerinden birer ders saati alınarak verilerin uygun bir şekilde toplanması sağlanmıştır. Envanterin uzun olması nedeniyle öğrencilerin motivasyonu yüksek seviyede tutulmaya çalışılmıştır. Envanterin ölçme ve değerlendirme ile ilgili olduğu ve gerekli açıklamalar yapıldıktan sonra envanter ve tutum anketi uygulanmıştır.

3.5. Verilerin Analizi

Bu araştırmadan elde edilen veriler SPSS 16,0 (Statistic Package for Social Sciencies) paket programı ve Test Analiz Programı (TAP) kullanılarak yapılmıştır (Brooks, 2003). TAP, Gordon Brooks tarafından geliştirilmiş olup, programda veri girişinin yapıldığı veri editörü penceresi 10 bölümden oluşmaktadır. Bu bölümlere, testin başlığı, başlığa dayalı yorumlar, testin cevap anahtarı, testin seçenek sayısı, testte analizi yapılacak maddeler, örneklem sayısı, testi oluşturan madde sayısı, örneklemeleri ifade eden sözcüklerin maksimum uzunluğu, kayıp veri sembolü ile örneklem testteki maddelere verdikleri cevaplar yazılmıştır (TAP: Test analysis program user guide). Test analiz programıyla yapılan analiz sonucunda madde ayırıcılık değeri 0,20' den küçük olan maddeler ile madde güçlük değerleri 0,1' den küçük olan maddeler envanterden çıkarılmıştır. SPSS programı ile ilköğretim öğretmen adaylarının ölçme ve değerlendirme okuryazarlıkları ve tutumları t-test, tek yönlü varyans analizi (ANOVA) ile analiz edilmiştir.

4. BULGULAR

Bu çalışmanın amacı ilköğretim öğretmen adaylarının ölçme değerlendirme okuryazarlık seviyelerini ve ölçme değerlendirmeye ilişkin tutumlarını belirlemektir. Araştırmanın bu bölümünde araştırma sorularına ilişkin istatistiksel sonuçlar sunulacaktır.

4.1. Araştırma Sorularına Ait İstatistikî Veriler

Araştırmanın pilot çalışmasına fen bilgisi öğretmenliği 3. sınıfta öğrenim gören 89 öğrenci katılmıştır. Pilot çalışmanın amacı, envanterin eksiklerini görüp tekrar düzenlemek üzere incelenmesini sağlamaktır. Envanter uygulandıktan sonra veriler (TAP) Test Analiz Programı ile analiz edilmiş ve madde ayıricılık değeri 0,20' den küçük olan maddeler ile madde güçlük değerleri 0,1' den küçük olan maddeler envanterden çıkarılmış, envanter ana çalışmaya hazır hale getirilmiştir. İlk uygulama sonucunda KR-20 güvenilirlik katsayısı 0,56, ortalama güçlük değeri 0,48, madde ayırt ediciliği 0,31 bulunmuştur. Analizden çıkarılan maddeler 13, 19, 27 ve 29 numaralı maddelerdir. İkinci uygulama sonucunda ise, KR-20 güvenilirlik katsayısı 0,714, ortalama güçlük değeri 0,54, madde ayırt ediciliği ise 0,35 bulunmuştur. Bu bölüm Tablo 4.1. de daha kapsamlı açıklanmaktadır.

Tablo 4.1. Envanter Maddelerinin Psikometrik Özellikleri

Maddeler	Doğru Cevaplayanların Sayısı	Madde Ayırt Edicilik Değeri	Madde Güçlük Değeri
1.	169	0,34	0,94
2.	110	0,37	0,61
3.	122	0,39	0,68
4.	101	0,44	0,56
5.	90	0,32	0,50
6.	121	0,35	0,67
7.	86	0,26	0,48
8.	102	0,47	0,57
9.	143	0,44	0,79
10.	70	0,29	0,39
11.	133	0,39	0,79
12.	106	0,56	0,59
13.	*Envanterden çıkarılmıştır		
14.	123	0,44	0,68
15.	53	0,31	0,30
16.	143	0,48	0,79
17.	115	0,39	0,64
18.	80	0,33	0,44
19.	*Envanterden çıkarılmıştır		
20.	104	0,49	0,58
21.	84	0,29	0,47
22.	*Envanterden çıkarılmıştır		
23.	106	0,35	0,59
24.	79	0,40	0,44
25.	80	0,38	0,44
26.	102	0,63	0,57
27.	*Envanterden çıkarılmıştır		
28.	65	0,32	0,36
29.	*Envanterden çıkarılmıştır		
30.	37	0,45	0,21
31.	90	0,52	0,50
32.	99	0,43	0,55
33.	99	0,57	0,55
34.	88	0,33	0,49
KR-20 Güvenilirlik katsayısı = 0,718			
Ortalama Güçlük Değeri = 0,556			
Ortalama Ayırt edicilik Değeri = 0,374			

*p<0.2 veya p>0.9, D<0.2

Chase (1999) ve Laatsch ve Choca (1991), güçlük değerleri 0,20 ile 0,85 arasında olan maddelerin psikometrik açıdan iyi olarak değerlendirilebileceğini belirtmişlerdir. Envanterdeki maddelerin ortalama güçlük değeri 0,556 olarak bulunmuştur. Bayrakçeken (2008) ve Tekin (1991) ortalama güçlük değerinin 0,50 olması gerektiğini vurgulamışlardır. Bütüner ve diğ. (2010) yaptıkları araştırmada maddelerinin ortalama güçlük değerlerini 0,644 olarak bulmuşlardır. Envanterdeki tüm maddelerin güçlük değerleri 0,2 ile 0,79 arasında değişmektedir. Envanterde sadece bir sorunun güçlük indeksi 0,96 olarak bulunmuştur.

Nunnally ve Berstein (1994) ; Chase (1999), ayırt edicilik değeri 0,30 ve daha fazla olan maddeleri psikometrik açıdan iyi olarak ifade etmişlerdir. Kryspin ve Feldhusen (1974), ayırt edicilik değeri 0,40 ve üstü olan maddelerin çok iyi, 0,20 ile 0,40 arasındaki maddelerin iyi, 0,0 ile 0,20 arasındaki maddelerin zayıf, 0,0"dan az olan maddelerin atılması veya gözden geçirilmesi gerektiğini belirtmişlerdir. Envanterdeki 25 maddenin ayırt edicilik değerleri ise 0,30"dan büyük, 4"ü 0,20 ile 0,30 arasında olduğu bulunmuştur. Ayrıca 12 maddenin ayırt ediciliği 0,4 ten büyük olduğu görülmüş ve maddelerin literatüre oranla daha iyi olduğu tespit edilmiştir.

Tabloda da görüldüğü gibi envanter maddelerinin psikometrik açıdan iyi olduğu görülmektedir. Chase (1999) ölçme değerlendirme okuryazarlığı envanteri tipindeki envanterler için güvenilirlik katsayısının en az 0,65 olması gerektiğini ifade etmiştir. Kehoe (1995), kısa envanterler için KR-20 güvenilirlik değerinin en az 0,50, 50 maddeden fazla olan envanterler için ise bu değer en az 0,80, Nitko (2001), bir envanterin güvenilir olabilmesi için güvenilirlik katsayı değerlerinin 0,70 ile 1.00 arasında olması gerektiğini belirtmişlerdir. Plake (1993) geliştirdiği ölçme değerlendirme okuryazarlığı anketinin güvenilirlik değerini 0,54, Campbell ve diğ. (2002) ve Mertler (2004) lise öğretmen ve öğretmen adaylarının ölçme değerlendirme bilgi ve beceri düzeylerini belirlemek için kullandıkları envanterlerin güvenilirlik değerini ise 0,74 olarak bulmuşlardır. Ölçme ve değerlendirme okuryazarlığı envanterinin KR-20 güvenilirlik katsayısı 0,716 olarak bulunmuştur. Sonuçlara bakılarak envanterin uygun bir şekilde Türkçeye uyarlandığı ve verdiği sonuçların birçok araştırmacının yaptığı envanterlerden daha güvenilir olduğu söylenebilir. Envanterin güvenilirliğinin sağlanması, yapılan araştırmanın da güvenilirliğini ve geçerliliğini sağlayacaktır.

4.1.1. Birinci Araştırma Sorusu

İlköğretim öğretmen adaylarının (fen bilgisi, sosyal bilgiler, matematik, Türkçe öğretmen adayları) değerlendirme okur-yazarlık seviyeleri nedir?

Bu araştırma sorusuna cevap bulmak için ölçme ve değerlendirme okuryazarlığı envanteri 180 öğretmen adayına uygulanmış ve bu verilere ait istatistikî bilgiler Tablo 4.2. de verilmiştir.

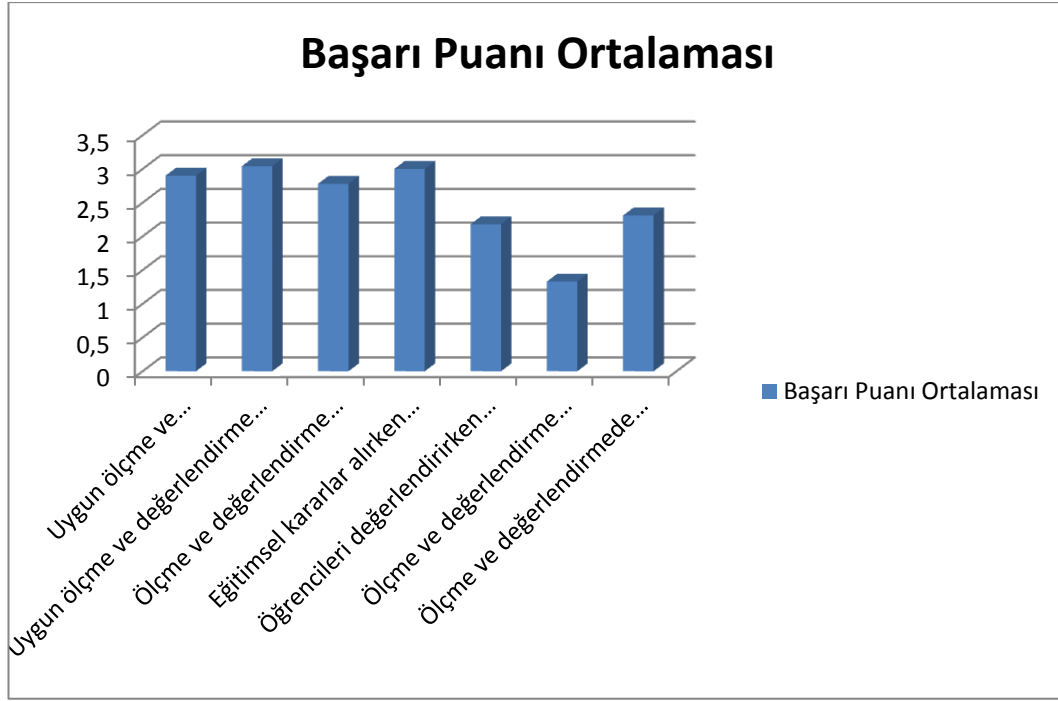
4.1.1.1. Fen Bilgisi Öğretmen Adaylarının (N=54) Ölçme ve Değerlendirme Okuryazarlık Seviyeleri

Tablo 4.2. Fen Bilgisi Öğretmen Adaylarının Değerlendirme Okuryazarlık Seviyeleri

<i>Standart</i>	^a <i>Ortalama</i>	<i>S.Sapma</i>
1. Uygun ölçme ve değerlendirme yöntemlerini seçme	2,90	1,1
2. Uygun ölçme ve değerlendirme yöntemlerini geliştirme	3,04	1,02
3. Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama	2,78	0,92
4. Eğitimsel kararlar alırken ölçme ve değerlendirme sonuçlarını kullanma	3,00	1,08
5. Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma	2,18	1,04
6. Ölçme ve değerlendirme sonuçlarını paylaşabilme	1,33	0,83
7. Ölçme ve değerlendirmede kanun dışı ve ahlaki olmayan durumları bilme	2,31	1,05
Toplam	17,57	4,01

^a Her bir standart için ortalama sonuçlar, en yüksek 5 puan, en düşük 0 puandır.

Grafik 4.1. Fen Bilgisi Öğretmen Adayları Başarı Grafiği



Fen bilgisi öğretmen adaylarından (N=54) toplanan verilere göre iç tutarlılık güvenilirlik katsayısı Cronbach's Alpha = 0,639 olarak bulunmuştur.

Güvenilirlik Analizi

Cronbach's Alpha	Maddelerin Sayısı	Toplam N
0,639	7 (ST)	54

Ortalama olarak fen bilgisi öğretmen adayları 34 sorudan, yaklaşık olarak 18 tanesine ($X=17,57/^{34}$) doğru cevap vermişlerdir. Yedi yeterlilik alanından en yüksek puanı 2. Ölçme ve değerlendirme standardı olan Uygun ölçme ve değerlendirme yöntemlerini geliştirmede ($X=3,04/^{5}$) en düşük puanı ise, 6. Ölçme ve değerlendirme standardı olan Ölçme ve Değerlendirme sonuçlarını paylaşabilme ($X=1,33/^{5}$) standardından almışlardır. Tablo 4.2. de 7 standart alanıyla ilgili ortalamalar ve standart sapmalar verilmiştir. Yurt dışında yapılan benzer araştırmalarla karşılaştırıldığında genel olarak alınan puanların çok düşük olduğu görülmektedir. Mertler (2005) yaptığı araştırmada öğretmen adaylarının 35 maddeden en az 4 tanesine % 90 üzerinde doğru cevap verildiğini belirtiyor. Fen bilgisi öğretmen adaylarının maalesef hiçbir standart açısından % 90 ve üzerinde doğru cevap

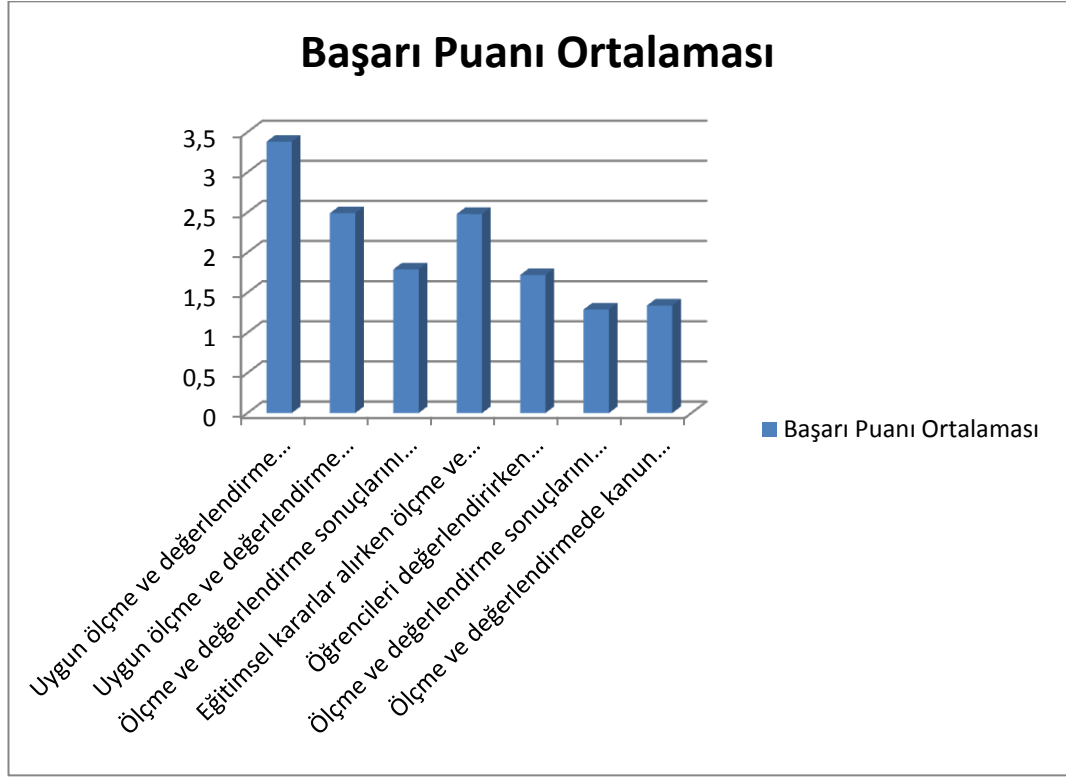
vermediği gözleniyor. Sadece 3 yeterlilik alanında % 60 ve üzerinde doğru cevap verildiği görülmektedir. Dikkat çeken bir başka nokta ise 6. yeterlilik alanı olan “Ölçme ve değerlendirme sonuçlarını paylaşabilme” alanında fen bilgisi öğretmen adaylarının çok düşük bir puan ($X=1,33/5$) almasıdır.

4.1.1.2. Sosyal Bilgiler Öğretmen Adaylarının (N=55) Ölçme ve Değerlendirme Okuryazarlık Seviyeleri

Tablo 4.3. Sosyal Bilgiler Öğretmen Adaylarının Okuryazarlık Seviyeleri

<i>Standart</i>	^a <i>Ortalama</i>	<i>S.sapma</i>
1. Uygun ölçme ve değerlendirme yöntemlerini seçme	3,38	1,38
2. Uygun ölçme ve değerlendirme yöntemlerini geliştirme	2,49	1,31
3. Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama	1,79	1,12
4. Eğitimsel kararlar alırken ölçme ve değerlendirme sonuçlarını kullanma	2,48	1,18
5. Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma	1,72	1,03
6. Ölçme ve değerlendirme sonuçlarını paylaşabilme	1,29	0,90
7. Ölçme ve değerlendirmede kanun dışı ve ahlaki olmayan durumları bilme	1,34	1,00
Toplam	14,51	4,26
^a Her bir standart için ortalama sonuçlar, en yüksek 5 puan, en düşük 0 puandır.		

Grafik 4.2. Sosyal Bilgiler Öğretmen Adayları Başarı Grafiği



Sosyal bilgiler öğretmen adaylarından (N=55) toplanan verilere göre iç tutarlılık güvenilirlik katsayısı Cronbach's Alpha = 0,577 olarak bulunmuştur.

Güvenilirlik Analizi

Cronbach's Alpha	Maddelerin Sayısı	Toplam N
0,577	7 (ST)	55

Sosyal bilgiler öğretmen adayları ölçme ve değerlendirme okuryazarlık seviyesinde en düşük sonuçları alan gruptur. Tablo 4.3.'te görüldüğü gibi, 34 maddeden ortalama olarak ($X=14,51/^{34}$) 15 doğru cevap vermişlerdir. Yedi yeterlilik alanında en yüksek puanı 1.yeterlilik alanı olan “Uygun ölçme ve değerlendirme yöntemlerini seçme” yeterlilik alanından ($X=3,38/^5$) almışlardır. Sosyal bilgiler öğretmen adayları, fen bilgisi öğretmen adaylarından ortalama olarak düşük olsalar da 2. yeterlilik alanında fen bilgisi öğretmenlerinden daha başarılı oldukları görülüyor (fen bilgisi; $X=3,04$ sosyal $X=3,38$).

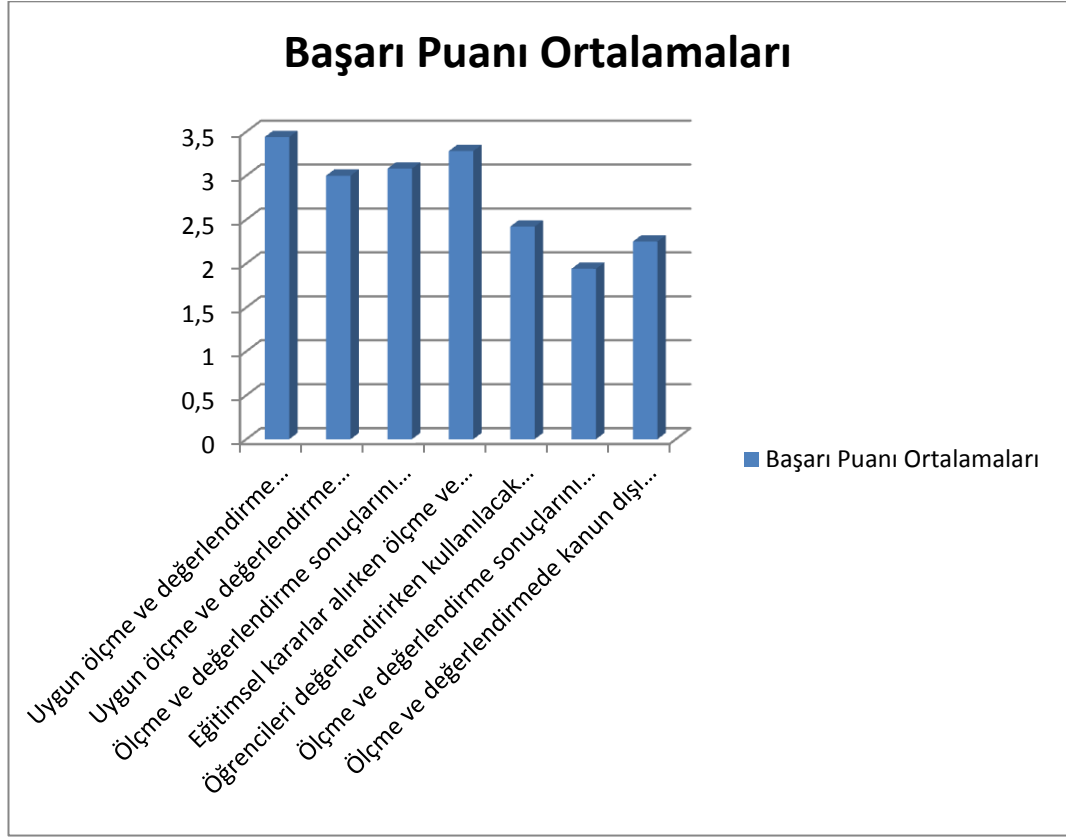
Sosyal bilgiler öğretmen adaylarının en başarısız oldukları alan 6. ölçme ve değerlendirme yeterliliği olan “Ölçme ve değerlendirme sonuçlarını paylaşabilme” yeterlilik alanı olmuştur ($X=1,29/5$). Bunun yanında göze çarpan diğer bir kayda değer sonuç ise, 3, 5, 6 ve 7 yeterlilik alanlarında sosyal bilgiler öğretmen adaylarının soruların % 40’ına bile doğru cevap verememiş olmalarıdır. Sadece üç (1, 2 ve 4) yeterlilik alanlarında ortalama seviyesinde başarılı oldukları görülmüştür.

4.1.1.3. Matematik Öğretmen Adaylarının (N=28) Ölçme ve Değerlendirme Okuryazarlık Seviyeleri

Tablo 4.4. Matematik Öğretmen Adaylarının Değerlendirme Okuryazarlık Seviyeleri

<i>Standart</i>	^a <i>Ortalama</i>	<i>S.sapma</i>
1. Uygun ölçme ve değerlendirme yöntemlerini seçme	3,44	1,04
2. Uygun ölçme ve değerlendirme yöntemlerini geliştirme	3,00	1,37
3. Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama	3,08	0,60
4. Eğitimsel kararlar alırken ölçme ve değerlendirme sonuçlarını kullanma	3,28	1,04
5. Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma	2,42	1,25
6. Ölçme ve değerlendirme sonuçlarını paylaşabilme	1,94	0,92
7. Ölçme ve değerlendirmede kanun dışı ve ahlaki olmayan durumları bilme	2,25	1,07
Toplam	20,21	2,95
^a Her bir standart için ortalama sonuçlar, en yüksek 5 puan, en düşük 0 puandır.		

Grafik 4.3. Matematik Öğretmen Adayları Başarı grafiği



Matematik öğretmen adaylarından (N=28) toplanan verilere göre iç tutarlılık güvenilirlik katsayısı Cronbach's Alpha = 0,44 olarak bulunmuştur.

Güvenilirlik Analizi

Cronbach's Alpha	Maddelerin Sayısı	Toplam N
0,44	7 (ST)	28

Matematik öğretmen adayları toplam üzerinden ölçme ve değerlendirme okuryazarlığı en yüksek ($X=20,21/34$) gruptur. Ortalama olarak 34 sorudan 21 tanesine doğru cevap vermişlerdir. Yeterlilik alanlarında en yüksek 1. yeterlilik alanı olan “Uygun ölçme ve değerlendirme yöntemlerini seçme” alanından almışlardır ($X=3,44/5$). En başarısız oldukları yeterlilik alanı ise 6. yeterlilik alanı olan “Ölçme ve değerlendirme sonuçlarını paylaşabilme” alanıdır ($X=1,94/5$). Matematik öğretmen adayları tüm gruplar

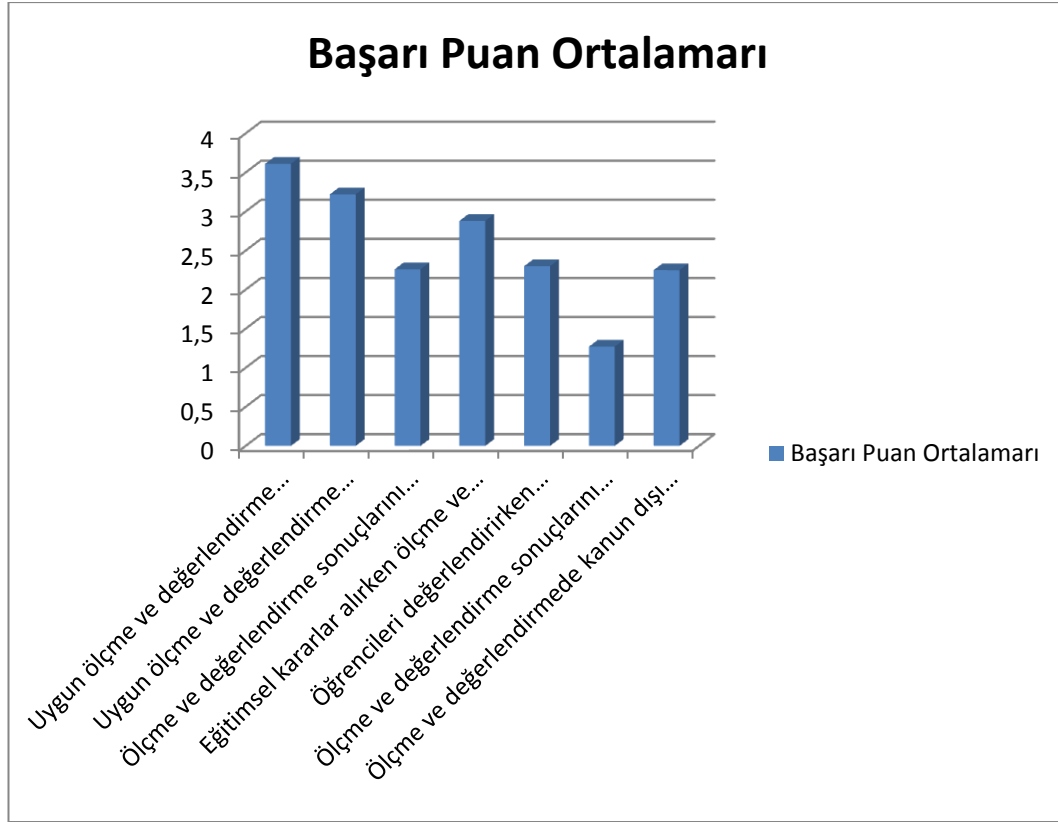
içersinde en başarılı olmasına rağmen, literatürde yapılan araştırmalara öğretmen adaylarının bazı yeterliliklerde % 60 ve üzerinde başarı gösterdiği görülmüş maalesef matematik öğretmen adaylarında bu başarı düzeyi görülmemiştir. Dört yeterlilik alanında (1, 2, 3 ve 4) % 60 ve üzerinde başarı gösterdikleri, geriye kalan yeterlilik alanlarından sadece 1 (6 yeterlilik) tanesinde % 40'ın altına düştükleri görülmüştür. Fen bilgisi, sosyal ve matematik öğretmen adaylarında göze çarpan en önemli sonuç kuşkusuz 6. yeterlilik alanı olan ölçme sonuçlarını paylaşabilmedeki önemli derecedeki sıkıntıdır.

4.1.1.4. Türkçe Öğretmen Adaylarının (N=43) Ölçme ve Değerlendirme Okuryazarlık Seviyeleri

Tablo 4.5. Türkçe Öğretmen Adaylarının Değerlendirme Okuryazarlık Seviyeleri

<i>Standart</i>	^a <i>Ortalama</i>	<i>S.sapma</i>
1. Uygun ölçme ve değerlendirme yöntemlerini seçme	3,61	1,17
2. Uygun ölçme ve değerlendirme yöntemlerini geliştirme	3,22	1,01
3. Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama	2,26	1,04
4. Eğitimsel kararlar alırken ölçme ve değerlendirme sonuçlarını kullanma	2,88	1,27
5. Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma	2,30	1,16
6. Ölçme ve değerlendirme sonuçlarını paylaşabilme	1,27	0,85
7. Ölçme ve değerlendirmede kanun dışı ve ahlaki olmayan durumları bilme	2,25	1,07
Toplam	17,83	4,32
^a Her bir standart için ortalama sonuçlar, en yüksek 5 puan, en düşük 0 puandır.		

Grafik 4.4. Türkçe Öğretmen Adayları Başarı Grafiği



Türkçe öğretmen adaylarından (N=43) toplanan verilere göre iç tutarlılık güvenilirlik katsayısı Cronbach's Alpha = 0,64 olarak bulunmuştur.

Güvenilirlik Analizi

Cronbach's Alpha	Maddelerin Sayısı	Toplam N
0,64	7 (ST)	43

Türkçe öğretmen adayları ile fen bilgisi öğretmen adaylarının ölçme ve değerlendirme okuryazarlık başarıları aynı düzeye yakındır. İkinci araştırma sorumuzda bu yakınlığın istatistikî olarak anlamlı olup olmadığına bakılacaktır. Türkçe öğretmen adayları 34 sorudan yaklaşık olarak 18 tanesine doğru cevap vermişlerdir ($X=17,83/^{34}$). Yeterlilik alanı en düşük olan, diğer üç öğretmen adayı grubunda olduğu gibi “Ölçme ve değerlendirme sonuçlarını paylaşabilme” yeterlilik alanıdır ($X=1,27/^{5}$). En başarılı oldukları yeterlilik alanı ise, 1. yeterlilik alanı olan “Uygun ölçme ve değerlendirme

yöntemlerini seçme” yeterlilik alanıdır ($X=3,61/5$). Türkçe öğretmen adayları yalnızca 2 yeterlilik alanında (1 ve 2) % 60 ve üzerinde başarı gösterebilmiştir. Geriye kalan 4 yeterlilik alanında % 40 ile % 60 arasında, 1 yeterlilik alanında ise % 40’ın altında başarı göstermiştir.

4.1.2. İkinci Araştırma Sorusu

İlköğretim öğretmen adaylarının (fen bilgisi, sosyal bilgiler, matematik ve Türkçe öğretmen adayları) değerlendirme okur-yazarlık seviyeleri arasında anlamlı bir farklılık var mıdır?

Bir önceki araştırma sorumuzda bazı gruplar açısından ortalamalar üzerinden ne tür farklılıkların olduğunu görmüştük. Bu farklılıkların istatistikî olarak anlamlı olup olmadığını görmek için SPSS 16,00 programından yararlanılmıştır. Gruplar bağımsız değişken olarak alınıp Tek Yönlü Varyans (One Way Anova) analiziyle bu farklılıklar ortaya çıkarılmaya çalışılmıştır. Tablo 4.6.’da ortalamalar karşılaştırmalı olarak verilmiştir.

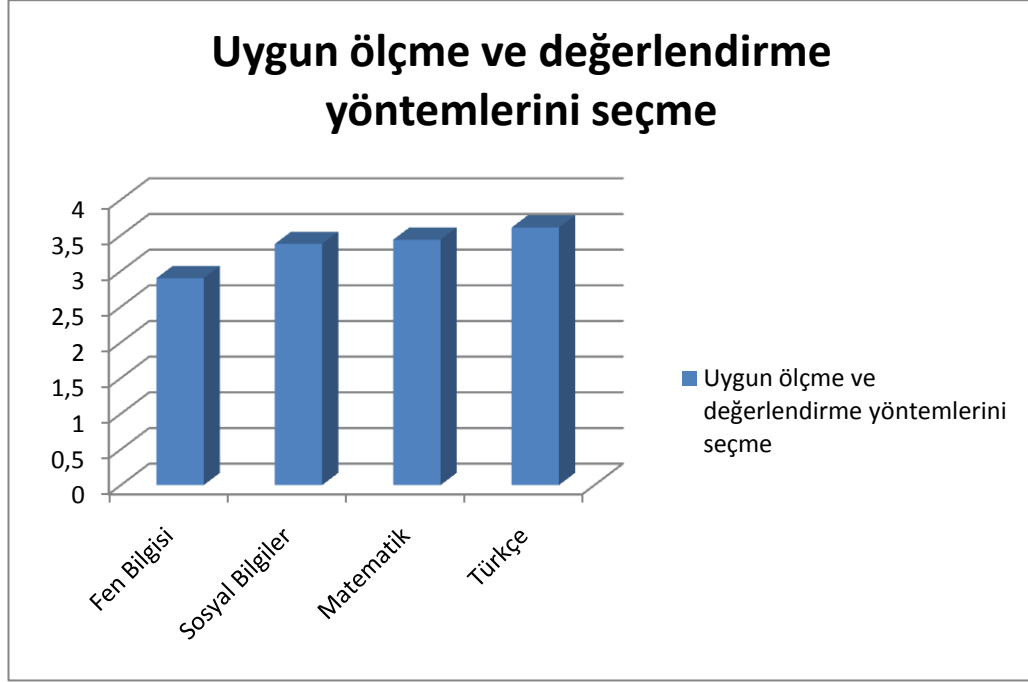
Tablo 4.6. Tüm Grupların Ölçme ve Değerlendirme Başarı Puanları

		N	Ortalama *X	S. Sapma
Uygun ölçme ve değerlendirme yöntemlerini seçme	Fen Bilgisi	54	2,90	1,13
	Sosyal Bilgiler	55	3,38	1,38
	Matematik	28	3,44	1,04
	Türkçe	43	3,61	1,17
	Toplam	180	3,30	1,23
Uygun ölçme ve değerlendirme yöntemlerini geliştirme	Fen Bilgisi	54	3,04	1,02
	Sosyal Bilgiler	55	2,49	1,31
	Matematik	28	3,00	1,37
	Türkçe	43	3,22	1,01
	Toplam	180	2,91	1,20
Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama	Fen Bilgisi	54	2,78	0,92
	Sosyal Bilgiler	55	1,79	1,12
	Matematik	28	3,08	0,60
	Türkçe	43	2,26	1,04
	Toplam	180	2,40	1,09
Eğitimsel kararlar alırken ölçme ve değerlendirme sonuçlarını kullanma	Fen Bilgisi	54	3,00	1,08
	Sosyal Bilgiler	55	2,48	1,18
	Matematik	28	3,28	1,04
	Türkçe	43	2,88	1,27
	Toplam	180	2,85	1,18
Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma	Fen Bilgisi	54	2,18	1,04
	Sosyal Bilgiler	55	1,72	1,03
	Matematik	28	2,42	1,25
	Türkçe	43	2,30	1,16
	Toplam	180	2,10	1,13
Ölçme ve değerlendirme sonuçlarını paylaşabilme	Fen Bilgisi	54	1,33	0,83
	Sosyal Bilgiler	55	1,29	0,90
	Matematik	28	1,94	0,92
	Türkçe	43	1,27	0,85
	Toplam	180	1,40	0,89
Ölçme ve değerlendirmede kanun dışı ve ahlaki olmayan durumları bilme	Fen Bilgisi	54	2,31	1,05
	Sosyal Bilgiler	55	1,34	1,00
	Matematik	28	2,25	1,05
	Türkçe	43	2,25	1,07

4.1.2.1. Yeterlilik Alanı 1

Uygun ölçme ve değerlendirme yöntemlerini seçme yeterlilik alanında anlamlı bir farklılık var mıdır?

Grafik 4.5. Yeterlilik alanı 1



Tablo 4.7. Yeterlilik Alanı 1 (ANOVA)

		Kareler Toplamı	<i>df</i>	Kareler Ortalaması	<i>F</i>	<i>p</i>
Uygun ölçme ve değerlendirme yöntemlerini seçme	Gruplar arası	13,90	3	4,63	3,14	0,026*
	Grup içi	259,13	176	1,47		
	Toplam	273,03	179			

* $p < 0,05$

Tablo 4.7.'de görüldüğü gibi, 1. yeterlilik alanının olan “Uygun ölçme ve değerlendirme yöntemlerini seçme”, ortalamalarının öğretmen adayları gruplarına göre farklılaşıp farklılaşmadığını belirlemek üzere yapılan tek yönlü varyans analizi sonucunda gruplar arası istatistiki olarak anlamlı bir farklılığın olduğu ortaya çıkmaktadır ($df (3, 176)=3,14$ $p=0,026$). Yani öğretmen adayları arasında bu standart açısından önemli

farklılıklar olduğu görülmüştür. Hangi gruplar arasında farklılığın olduğu ve ne derecede olduğu Çoklu Regresyon Analizi ile belirlenmiştir.

Tablo 4.8. Yeterlik Alanı 1 Çoklu Regresyon Analizi

Bağımlı Değişken	(A) Sınıflar	(B) Sınıflar	Ortalamalar farkı (A-B)	Standart hata	<i>p</i>
<i>Yeterlilik 1</i> <i>Uygun ölçme ve değerlendirme yöntemlerini seçme</i>	Fen Bilgisi	Sosyal Bilgiler	-0,48	0,23	0,167
		Matematik	-0,54	0,28	0,218
		Türkçe	-0,71	0,24	0,022*
	Sosyal Bilgiler	Fen Bilgisi	0,48	0,23	0,167
		Matematik	-0,06	0,28	0,996
		Türkçe	-0,23	0,24	0,778
	Matematik	Fen Bilgisi	0,54	0,28	0,218
		Sosyal Bilgiler	0,06	0,28	0,996
		Türkçe	-0,16	0,29	0,939
	Türkçe	Fen Bilgisi	0,71	0,24	0,022*
		Sosyal Bilgiler	0,23	0,24	0,778
		Matematik	0,16	0,29	0,939

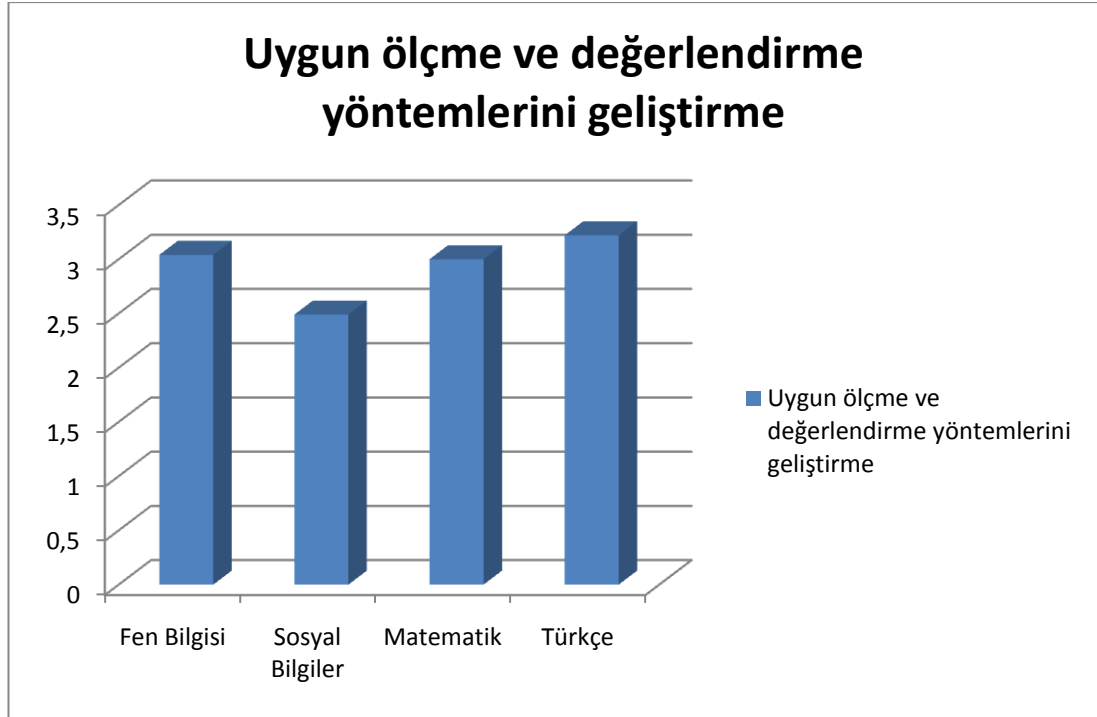
*P<0,05

Tablo 4.8.'de görüldüğü gibi Fen bilgisi öğretmen adayları ile Türkçe öğretmen adayları arasında istatistikî olarak anlamlı bir farklılık görülmektedir. Bu standart açısından, Türkçe öğretmen adaylarının, fen bilgisi öğretmen adaylarından daha başarılı olduğu görülmektedir. Diğer gruplar arasında ise bu standart açısından anlamlı bir farklılık görülmemektedir.

4.1.2.2. Yeterlilik Alanı 2

Uygun ölçme ve değerlendirme yöntemlerini geliştirme yeterlilik alanında anlamlı bir farklılık var mıdır?

Grafik 4.6. Yeterlilik alanı 2



Tablo 4.9. Yeterlik Alanı 2 (ANOVA)

		Kareler Toplamı	<i>df</i>	Kareler Ortalaması	<i>F</i>	<i>p</i>
Uygun ölçme ve değerlendirme yöntemlerini geliştirme	Gruplar arası	15,29	3	5,09	3,67	0,013*
	Grup içi	244,366	176	1,388		
	Toplam	259,663	179			

* $p < 0,05$

Yeterlilik alanı olan “Uygun ölçme ve değerlendirme yöntemlerini geliştirme” alanında, gruplar arasında anlamlı bir farklılığın olduğu görülmektedir ($df (3, 176) = 3,67$ $p=0,013$). Bu standart açısından, hangi gruplar arasında farklılığın olduğu ve ne derecede olduğu “Post-hoc Scheffe” testi ile belirlenmiştir.

Tablo 4.10. Yeterlik Alanı 2 Çoklu Regresyon Analizi

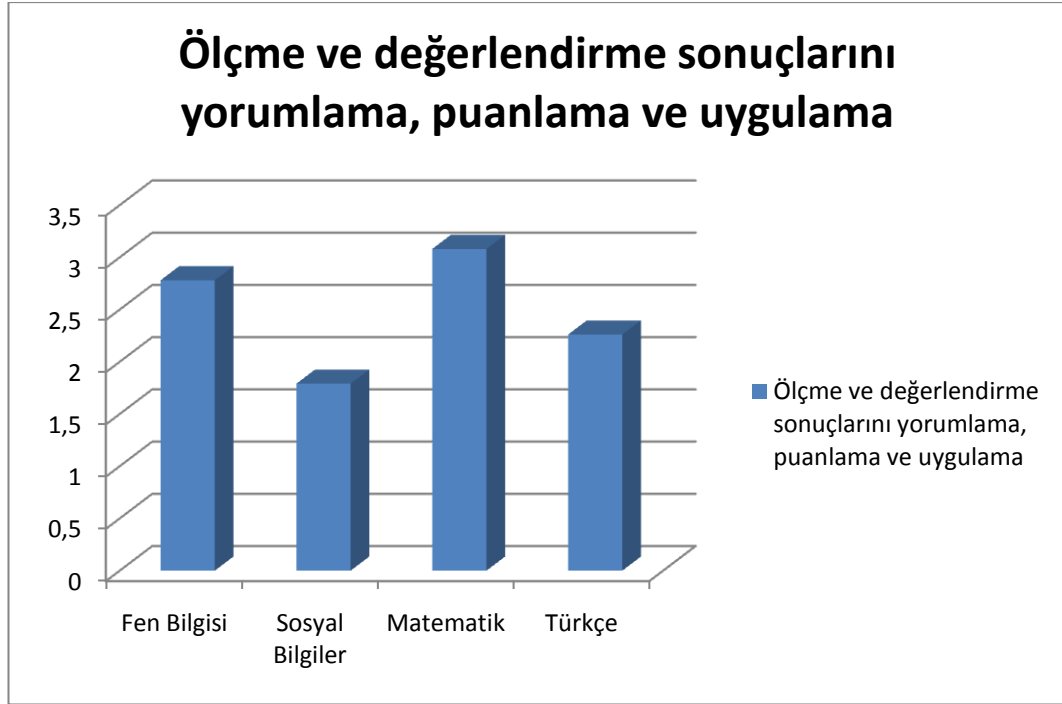
Bağımlı Değişken	(A) Sınıflar	(B) Sınıflar	Ortalamalar farkı (A-B)	Standart hata	<i>p</i>
<i>Yeterlilik 2</i> <i>Uygun ölçme ve değerlendirme yöntemlerini geliştirme</i>	Fen Bilgisi	Sosyal Bilgiler	0,55	,225	0,071
		Matematik	0,03	0,274	0,999
		Türkçe	-0,18	0,240	0,868
	Sosyal Bilgiler	Fen Bilgisi	-0,55	0,225	0,071
		Matematik	-0,51	0,273	0,235
		Türkçe	-0,73	0,239	0,013*
	Matematik	Fen Bilgisi	-0,03	0,274	0,999
		Sosyal Bilgiler	0,51	0,273	0,235
		Türkçe	-0,22	0,286	0,867
	Türkçe	Fen Bilgisi	0,18	0,240	0,868
		Sosyal Bilgiler	0,73	0,239	0,013*
		Matematik	0,22	0,286	0,867

* $p < 0,05$

4.1.2.3. Yeterlilik Alanı 3

Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama yeterlilik alanında anlamlı bir farklılık var mıdır?

Grafik 4.7. Yeterlilik alanı 3



Tablo 4.11. Yeterlilik Alanı 3 (ANOVA)

		Kareler Toplamı	<i>df</i>	Kareler Ortalaması ₁	<i>F</i>	<i>p</i>
Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama	Gruplar arası	42,27	3	14,092	14,53	0,000**
	Grup içi	170,59	176	,969		
	Toplam	212,87	179			

** $p < 0,005$

Yeterlilik alanı olan “Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama” alanında, gruplar arasındaki farklılığın diğer standartlara nazaran daha anlamlı bir farklılıkların olduğu görülmektedir ($df(3, 176) = 3,67$ $p = 0,000$). Bu standart açısından, hangi gruplar arasında farklılığın olduğu ve ne derecede olduğu “Post-hoc Scheffe” testi ile belirlenmiştir.

Tablo 4.12. Yeterlik Alanı 3 Çoklu Regresyon Analizi

Bağımlı Değişken	(A) Sınıflar	(B) Sınıflar	Ortalamalar farkı (A-B)	Standart hata	<i>p</i>
Yeterlilik 3 Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama	Fen Bilgisi	Sosyal Bilgiler	0,99	0,188	0,000**
		Matematik	-0,29	0,229	0,577
		Türkçe	0,52	0,201	0,048*
	Sosyal Bilgiler	Fen Bilgisi	-0,99	0,188	0,000**
		Matematik	-1,28	0,228	0,000**
		Türkçe	-0,47	0,200	0,091
	Matematik	Fen Bilgisi	0,29	0,229	0,577
		Sosyal Bilgiler	1,28	0,228	0,000**
		Türkçe	0,81	0,239	0,004**
	Türkçe	Fen Bilgisi	-0,52	0,201	0,048*
	Sosyal Bilgiler	0,47	0,200	0,091	
	Matematik	-0,81	0,239	0,004*	

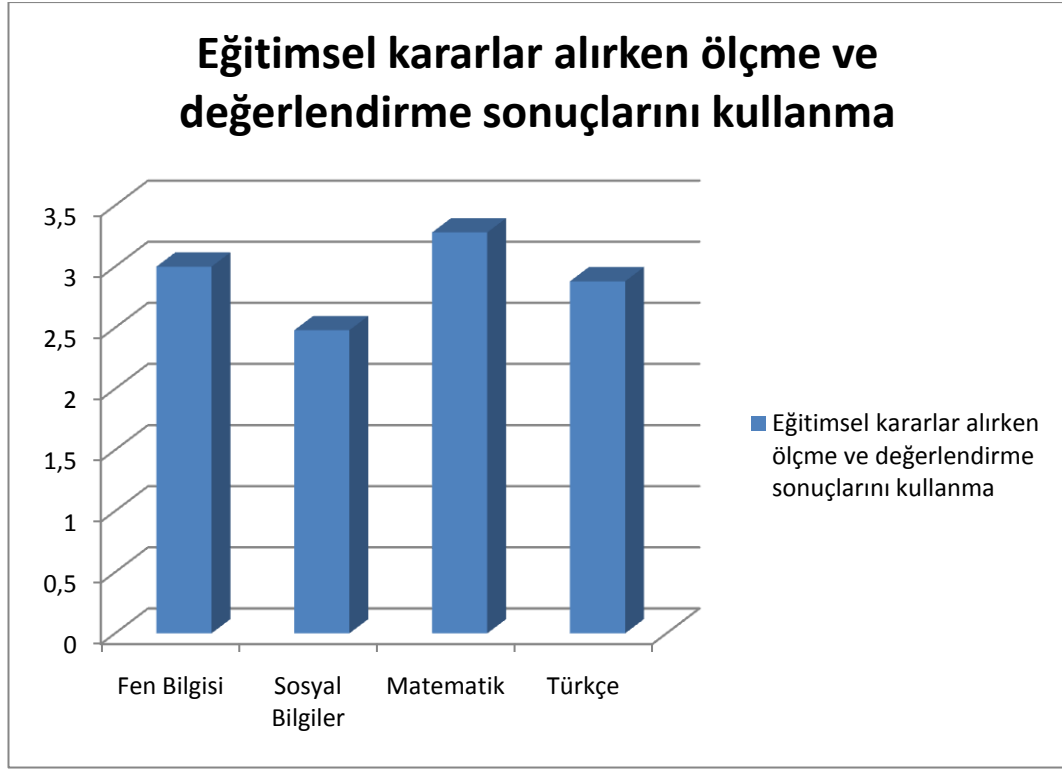
* $p < 0,05$; ** $p < 0,005$

3.yeterlilik alanında fen bilgisi öğretmen adayları ile sosyal bilgiler öğretmen adayları arasında istatistikî olarak büyük bir farklılığın olduğu görülmektedir ($p=0,000$). Bunun yanında fen bilgisi öğretmen adayları ile Türkçe öğretmen adayları arasında da anlamlı farklılığın olduğu belirlenmiştir. Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama yeterlilik alanında matematik öğretmen adaylarının fen bilgisi öğretmenliği haricindeki diğer tüm gruplarla arasında önemli bir fark bulunmaktadır. Matematik öğretmen adayları bu standartta en başarılı gruptur.

4.1.2.4. Yeterlilik Alanı 4

Eğitimsel kararlar alırken ölçme ve değerlendirme sonuçlarını kullanma yeterlilik alanında anlamlı bir farklılık var mıdır?

Grafik 4.8. Yeterlilik alanı 4



Tablo 4.13. Yeterlik Alanı 4 (ANOVA)

		Kareler Toplamı	<i>df</i>	Kareler Ortalaması	<i>F</i>	<i>p</i>
Eğitimsel kararlar alırken ölçme ve değerlendirme sonuçlarını kullanma	Gruplar arası	14,132	3	4,711	3,51	0,016*
	Grup içi	235,676	176	1,339		
	Toplam	249,808	179			

* $p < 0,05$

“Eğitimsel kararlar alırken ölçme ve değerlendirme sonuçlarını kullanma” yeterlilik alanında, gruplar arasında anlamlı farklılıkların olduğu görülmektedir ($df (3, 176) = 3,51$ $p=0,016$). Bu standart açısından, hangi gruplar arasında farklılığın olduğu ve ne derecede olduğu Çoklu Regresyon Analizi ile belirlenmiştir.

Tablo 4.14. Yeterlik Alanı 4 Çoklu Regresyon Analizi

Bağımlı Değişken	(A) Sınıflar	(B) Sınıflar	Ortalamalar farkı (A-B)	Standart hata	<i>p</i>
Yeterlilik 4 Eğitimsel kararlar alınırken ölçme ve değerlendirme sonuçlarını kullanma	Fen Bilgisi	Sosyal Bilgiler	0,52	0,221	0,088
		Matematik	-0,28	0,269	0,724
		Türkçe	0,12	0,236	0,956
	Sosyal Bilgiler	Fen Bilgisi	-0,52	0,221	0,088
		Matematik	-0,80	0,268	0,016*
		Türkçe	-0,40	0,235	0,322
	Matematik	Fen Bilgisi	0,28	0,269	0,724
		Sosyal Bilgiler	0,80	0,268	0,016*
		Türkçe	0,40	0,281	0,482
	Türkçe	Fen Bilgisi	-0,12	0,236	0,956
		Sosyal Bilgiler	0,40	0,235	0,322
		Matematik	-0,401	0,281	0,482

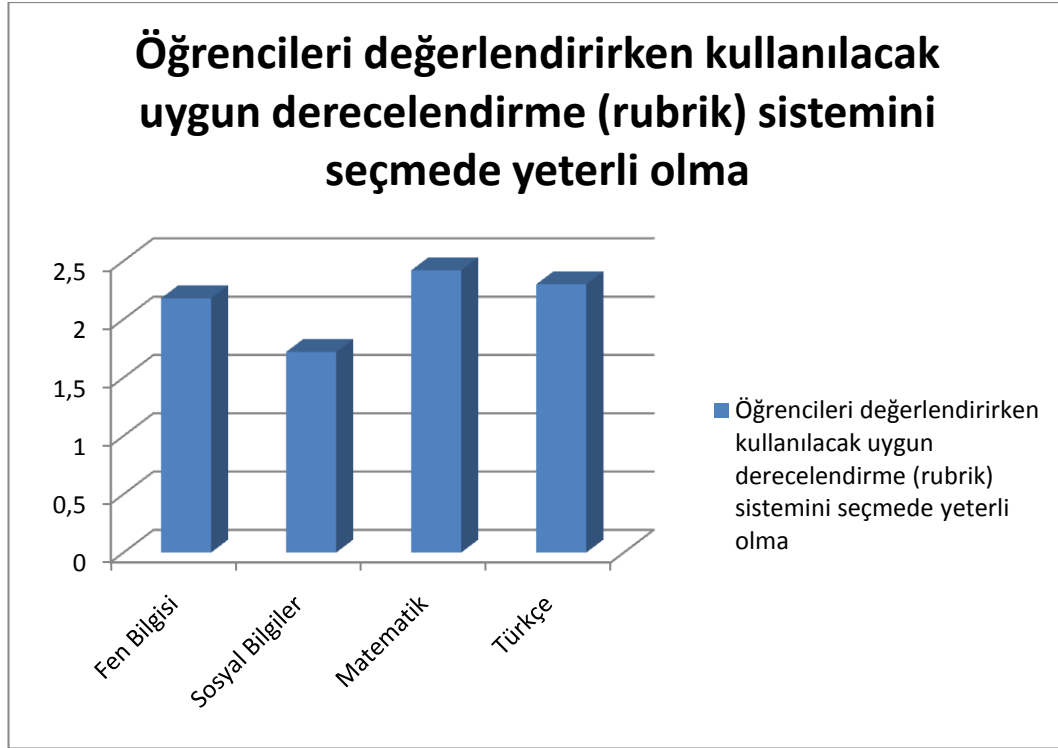
* $p < 0,05$

Bu standart açısından yalnızca matematik ve sosyal bilgiler öğretmenliği gruplarında istatistiki olarak anlamlı bir farklılık bulunmuştur ($p=0,016$). Diğer guruplar arasında anlamlı bir farklılık yoktur.

4.1.2.5. Yeterlilik Alanı 5

Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma yeterlilik alanında anlamlı bir farklılık var mıdır?

Grafik 4.9. Yeterlilik alanı 5



Tablo 4.15. Yeterlilik Alanı 5 (ANOVA)

	Kareler Toplamı	<i>df</i>	Kareler Ortalaması	<i>F</i>	<i>p</i>
Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma	12,979	3	4,326	3,53	0,016
Gruplar arası	215,717	176	1,226		
Grup içi	228,696	179			
Toplam					

* $p < 0,05$

“Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma” yeterlilik alanında, gruplar arasında anlamlı farklılıkların olduğu görülmektedir ($df(3, 176) = 3,53$ $p = 0,016$). Bu standart açısından, hangi gruplar arasında farklılığın olduğu ve ne derecede olduğu “Post-hoc Scheffe” testi ile belirlenmiştir.

Tablo 4.16. Yeterlik Alanı 5 Çoklu Regresyon Analizi

Bağımlı Değişken	(A) Sınıflar	(B) Sınıflar	Ortalamalar farkı (A-B)	Standart hata	<i>p</i>
Yeterlilik 5 Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma	Fen Bilgisi	Sosyal Bilgiler	0,46	0,212	0,133
		Matematik	-0,24	0,257	0,781
		Türkçe	-0,11	0,226	0,955
	Sosyal Bilgiler	Fen Bilgisi	-0,46	0,212	0,133
		Matematik	-0,70	0,257	0,033*
		Türkçe	-0,57	0,225	0,053
	Matematik	Fen Bilgisi	0,24	0,257	0,781
		Sosyal Bilgiler	0,70	0,257	0,033*
		Türkçe	0,12	0,268	0,966
	Türkçe	Fen Bilgisi	0,11	0,226	0,955
		Sosyal Bilgiler	0,57	0,225	0,053
		Matematik	-0,12	0,268	0,966

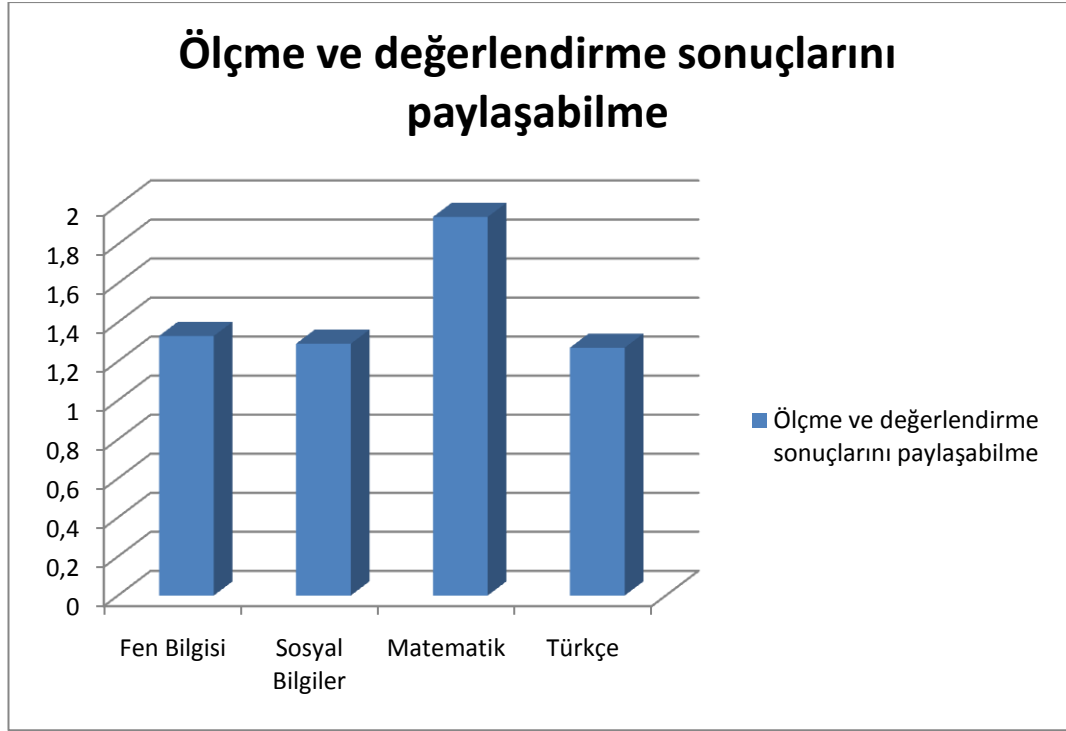
* $p < 0,05$

Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma, bu standart açısından yalnızca Matematik ve sosyal bilgiler öğretmenliği gruplarında istatistikî olarak anlamlı bir farklılık bulunmuştur ($p=0,033$). Diğer guruplar arasında anlamlı bir farklılık yoktur.

4.1.2.6. Yeterlilik Alanı 6

Ölçme ve değerlendirme sonuçlarını paylaşabilme yeterlilik alanında anlamlı bir farklılık var mıdır?

Grafik 4.10. Yeterlilik alanı 6



Tablo 4.17. Yeterlik Alanı 6 (ANOVA)

		Kareler Toplamı	<i>df</i>	Kareler Ortalaması	<i>F</i>	<i>p</i>
Ölçme ve değerlendirme sonuçlarını paylaşabilme	Gruplar arası	9,678	3	3,226	4,21	0,007
	Grup içi	134,787	176	0,766		
	Toplam	144,465	179			

* $p < 0,05$

“Ölçme ve değerlendirme sonuçlarını paylaşabilme” yeterlilik alanında ise, gruplar arasında anlamlı farklılıkların olduğu görülmektedir ($df (3, 176) = 4,21$ $p=0,007$). Bu standart açısından, hangi gruplar arasında farklılığın olduğu ve ne derecede olduğu “Post-hoc Scheffe” testi ile belirlenmiştir.

Tablo 4.18. Yeterlik Alanı 6 Çoklu Regresyon Analizi

Bağımlı Değişken	(A) Sınıflar	(B) Sınıflar	Ortalamalar farkı (A-B)	Standart hata	<i>p</i>
Yeterlilik 6 Ölçme ve değerlendirme sonuçlarını paylaşabilme	Fen Bilgisi	Sosyal Bilgiler	0,03	0,167	0,997
		Matematik	-0,61	0,203	0,016*
		Türkçe	0,05	0,178	0,991
	Sosyal Bilgiler	Fen Bilgisi	-0,03	0,167	0,997
		Matematik	-0,64	0,203	0,010*
		Türkçe	0,01	0,178	1,000
	Matematik	Fen Bilgisi	0,61	0,203	0,016
		Sosyal Bilgiler	0,64	0,203	0,010**
		Türkçe	0,66	0,212	0,011**
	Türkçe	Fen Bilgisi	-0,05	0,178	0,991
		Sosyal Bilgiler	-0,01	0,178	1,000
		Matematik	-0,66	0,212	0,011**

* $p < 0,05$ ** $p < 0,01$

Ölçme ve değerlendirme sonuçlarını paylaşabilme standardı, en çok göze çarpan standart olmuştur. Çünkü bu standart açısından öğretmen adaylarının tümü çok başarısız oldukları belirlenmişti. Çoklu regresyon analizi sonuçlarına göre, fen bilgisi ve matematik öğretmenliği öğretmen adayları arasında bu standart açısından anlamlı farklılıklar tespit edilmiştir ($p=0,016$). Matematik öğretmen adaylarının fen bilgisi öğretmen adaylarından anlamlı bir şekilde daha başarılı oldukları ortaya çıkmıştır. Bir diğer anlamlı farklılık ise matematik öğretmen adayları ile sosyal ve Türkçe öğretmen adayları arasındadır. Bu anlamlı farklılık fen bilgisi öğretmen adaylarında olduğundan daha büyüktür ($p=0,01$).

4.1.2.7. Yeterlilik Alanı 7

Ölçme ve değerlendirmede kanun dışı ve ahlaki olmayan durumları bilme yeterlilik alanında anlamlı bir farklılık var mıdır?

Grafik 4.11. Yeterlilik alanı 7



Tablo 4.19. Yeterlilik Alanı 7 (ANOVA)

		Kareler Toplamı	<i>df</i>	Kareler Ortalaması	<i>F</i>	<i>p</i>
Ölçme ve değerlendirilmede kanun dışı ve ahlaki olmayan durumları bilme	Gruplar arası	58,291	3	19,43	17,819	0,000**
	Grup içi	191,909	176	1,090		
	Toplam	250,199	179			

** $p < 0,005$

“Ölçme ve değerlendirilmede kanun dışı ve ahlaki olmayan durumları bilme” yeterlilik alanında, gruplar arasında anlamlı farklılıkların olduğu görülmektedir ($df(3, 176) = 17,81$ $p = 0,000$). Bu standart açısından, hangi gruplar arasında farklılığın olduğu ve ne derecede olduğu “Post-hoc Scheffe” testi ile belirlenmiştir.

Tablo 4.20. Yeterlik Alanı 7 Çoklu Regresyon Analizi

Bağımlı Değişken	(A) Sınıflar	(B) Sınıflar	Ortalamalar farkı (A-B)	Standart hata	<i>p</i>
Yeterlilik 7 Ölçme ve değerlendirmede kanun dışı ve ahlaki olmayan durumları bilme	Fen Bilgisi	Sosyal Bilgiler	0,97	0,200	0,000**
		Matematik	-0,69	0,243	0,023*
		Türkçe	0,06	0,213	0,991
	Sosyal Bilgiler	Fen Bilgisi	-0,97	0,200	0,000**
		Matematik	-1,67	0,242	0,000**
		Türkçe	-0,90	0,212	0,000**
	Matematik	Fen Bilgisi	0,69	0,243	0,023*
		Sosyal Bilgiler	1,67	0,242	0,000**
		Türkçe	0,76	0,253	0,016*
	Türkçe	Fen Bilgisi	-0,06	0,213	0,991
		Sosyal Bilgiler	0,90	0,212	0,000**
		Matematik	-0,76	0,253	0,016*

* $p < 0,05$ ** $p < 0,005$

Ölçme ve değerlendirmede kanun dışı ve ahlaki olmayan durumları bilme, bu standart açısından gruplar arasında birçok anlamlı farklılıklar bulunduğu görülmektedir. Öncelikle fen bilgisi ve matematik öğretmenliği arasındaki farklılığın anlamlı olduğu görülmektedir ($p=0,023$). Aynı zamanda soysa bilgiler ile de anlamlı farklılığın bulunduğu fakat bu farklılığın matematik ile olan farklılıktan daha anlamlı olduğu görülmektedir ($p=0,000$). Türkçe öğretmen adaylarının ise, sosyal bilgiler ve matematik öğretmen adaylarıyla arasında anlamlı bir şekilde farklılıkların olduğu görülmektedir.

4.1.3. Üçüncü Araştırma Sorusu

İlköğretim öğretmen adaylarının (fen bilgisi, sosyal bilgiler, matematik ve Türkçe öğretmen adayları) ölçme-değerlendirmeye ilişkin tutumlar nedir?

İlköğretim öğretmen adaylarının Ölçme ve Değerlendirmeye ilişkin tutumlarının belirlenmesi için “Ölçme ve Değerlendirmeye İlişkin Tutum Ölçeği” geliştirilip 180 öğretmen adayına uygulanmıştır. Bu anket ile ilgili güvenilirlik analiz sonuçları aşağıdaki tabloda verilmiştir.

Güvenilirlik Analizi

Cronbach's Alpha	Maddelerin sayısı
0,887	41

“Ölçme ve Değerlendirmeye İlişkin Tutum Ölçeği”

Yukarıda görüldüğü gibi Cronbach's Alpha=0,887 olarak bulunmuştur. Anket başlangıçta 47 maddeden oluşmaktadır. Yapılan güvenilirlik analizleri sonucunda bazı maddelerin anketten çıkarılması uygun görülmüştür. “*” işaretli maddeler anketten çıkarılmıştır.

Tablo 4.21. Tutum Anketi Tüm Faktörler ve Çıkarılan Faktörler

<i>Ölçme ve Değerlendirmeye ilişkin tutum faktörler;</i>	<i>Faktör# Soru numaraları</i>
Ölçme ve değerlendirmeyi yapmadaki ilgi seviyesi	1, 4* , 10, 13, 15, 20, 21, 23, 24, 28, 31, 33, 36, 39, 41 ve 45*
Ölçme ve değerlendirmeyi yaparken duyulan endişe ve korku	2, 5, 8, 11* , 16, 18* , 22, 26, 34, 37 ve 44
Ölçme ve değerlendirmeyi yapmanın önemi	6, 9, 14* , 17, 29, 35, 40, 42 ve 46
Ölçme ve değerlendirmeyi yapmadaki yeterlilik	3, 7, 12, 19, 25, 27, 30, 32, 38, 43 ve 47

Anket önermelerinde

Hiç katılmıyorum	1 puan
Kararsızım	3 puan
Katılmıyorum	2 puan
Katılıyorum	4 puan
Tamamen Katılıyorum	5 puan

olarak değerlendirilmiştir.

4.1.3.1. Öğretmen Adaylarının Ölçme ve Değerlendirmeye Yönelik Tutum Ölçeğine İlişkin Sonuçları

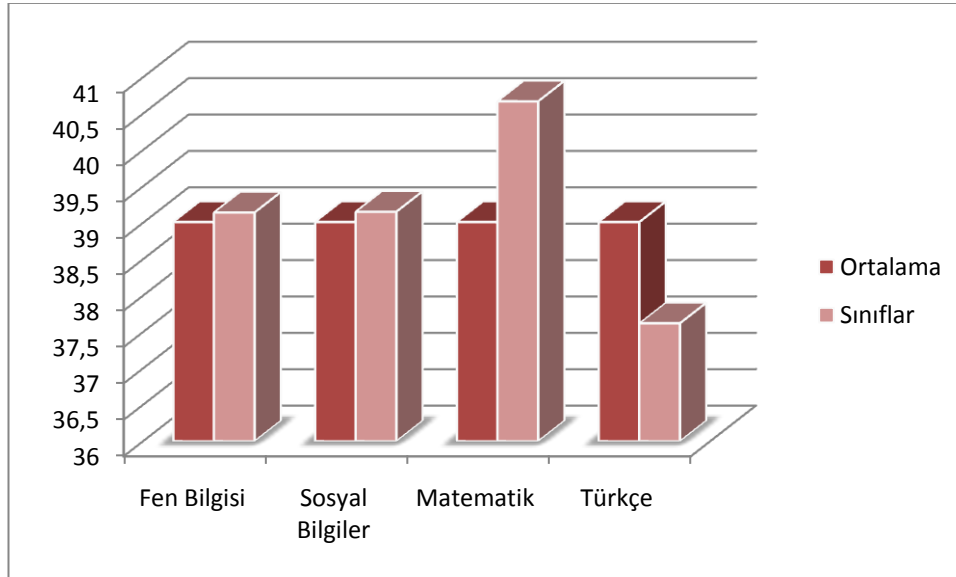
Tablo 4.22. Özel İlgi Faktörü Toplam Grup Sonuçları

Sınıflar		1. Madde	10. Madde	13. Madde	15. Madde	20. Madde	21. Madde	23. Madde	24. Madde
Fen Bilgisi Öğrt.	Ortalama	4,24	3,50	3,24	3,20	3,96	3,24	3,94	3,55
	N	54	54	54	54	54	54	54	54
	<i>S.Sapma</i>	,642	,837	,867	,939	,823	1,026	,833	1,223
Sosyal Bilgiler Öğrt.	Ortalama	4,10	2,99	3,46	3,19	3,31	3,23	3,90	3,26
	N	55	55	55	55	55	55	55	55
	<i>S.Sapma</i>	,978	1,067	1,081	1,036	1,282	1,033	,878	1,216
Matematik Öğrt.	Ortalama	4,25	3,10	3,42	2,96	3,77	3,35	3,96	3,96
	N	28	28	28	28	28	28	28	28
	<i>S.Sapma</i>	,700	,994	1,069	,838	,786	1,026	,922	1,070
Türkçe Öğrt.	Ortalama	4,02	3,19	3,07	2,74	3,36	3,26	3,83	2,97
	N	43	43	43	43	43	43	43	43
	<i>S.Sapma</i>	,859	1,159	,986	1,049	1,109	1,235	1,044	1,311
Toplam	Ortalama	4,14	3,21	3,29	3,05	3,59	3,26	3,91	3,39
	N	180	180	180	180	180	180	180	180
	<i>S.Sapma</i>	,816	1,028	1,000	,993	1,078	1,073	,908	1,254

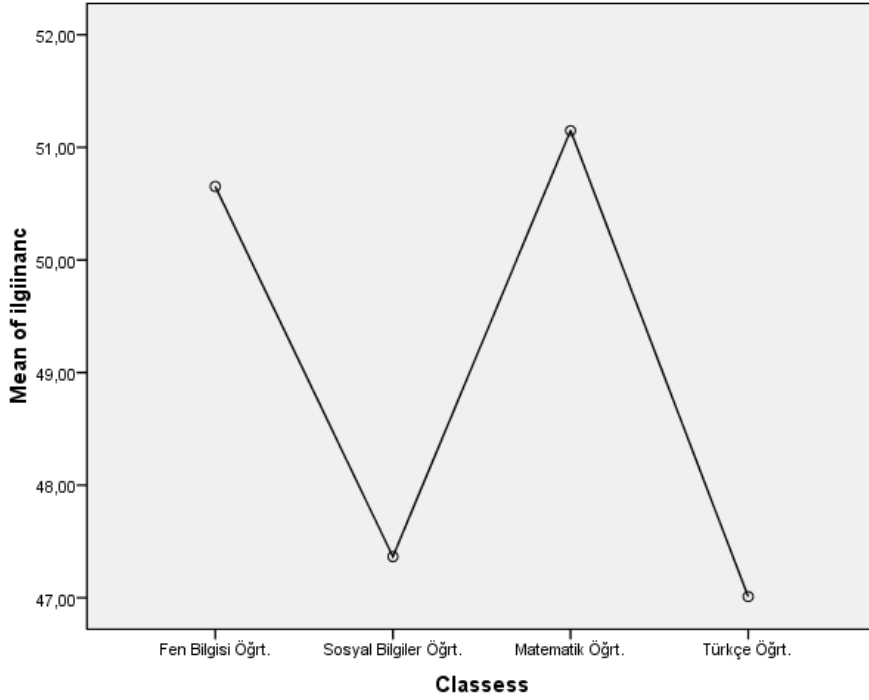
Özel İlgi Faktörü Toplam Grup Sonuçları (Devamı)

Sınıflar		28. Madde	31. Madde	33. Madde	36. Madde	39. Madde	41. Madde
Fen Bilgisi Öğrt.	Ortalama	3,44	3,47	3,79	3,42	3,64	3,96
	N	54	54	54	54	54	54
	S.Sapma	1,021	1,140	,786	,881	,704	,867
Sosyal Bilgiler Öğrt.	Ortalama	2,91	3,09	3,34	3,43	3,71	3,37
	N	55	55	55	55	55	55
	S.Sapma	1,081	1,101	,995	,989	,800	1,107
Matematik Öğrt.	Ortalama	3,50	3,65	3,78	3,57	3,89	3,92
	N	28	28	28	28	28	28
	S.Sapma	1,102	,862	,917	,920	,628	1,086
Türkçe Öğrt.	Ortalama	3,02	3,48	3,68	2,96	3,63	3,73
	N	43	43	43	43	43	43
	S.Sapma	1,101	1,026	,830	1,069	,945	1,176
Toplam	Ortalama	3,19	3,38	3,63	3,34	3,70	3,72
	N	180	180	180	180	180	180
	S.Sapma	1,092	1,074	,898	,984	,785	1,076

Grafik 4.12. Ölçme ve Değerlendirmeye İlişkin Özel İlgi Faktörü Toplam Sonuçları Grafiği



Grafik 4.13. Özel İlgi Scree Plot Grafiği



Öğretmen adaylarının ölçme ve değerlendirmeyi yapmaya karşı özel ilgileri faktörü açısından fen bilgisi ve matematik öğretmen adaylarının diğer öğretmen adaylarından daha yüksek olduğu görülmektedir. Genel sonuçlara bakıldığında öğretmen adaylarının ölçme ve değerlendirmeye ilişkin özel ilgilerinin “kararsızım” ($X=3,19$) önermesine daha yakın olduğu görülmektedir. Bu farklılığın istatistikî olarak anlamlı olup olmadığına “Tek Yönlü Varyans Analizi” ile bakılacaktır. Aşağıda ANOVA sonuçları verilmiştir. Tablo 4.23’te görüldüğü gibi Özel ilgi faktörü açısından gruplar arasında anlamlı farklılıklar olduğu görülmektedir ($df (3, 176) = 4,184$ $p=0,007$). Tablo 4.24.’te görüldüğü gibi sadece fen bilgisi ve Türkçe öğretmen adayları arasında anlamlı farklılığın olduğu görülmektedir ($p=0,049$).

Tablo 4.23. Özel İlgi Faktörü (ANOVA)

Özel İlgi	Kareler Toplamı	<i>df</i>	Kareler Ortalaması	F	<i>p</i>
Gruplar arası	590,203	3	196,734	4,184	0,007*
Grup içi	8275,088	176	47,018		
Toplam	8865,291	179			

* $p<0,05$

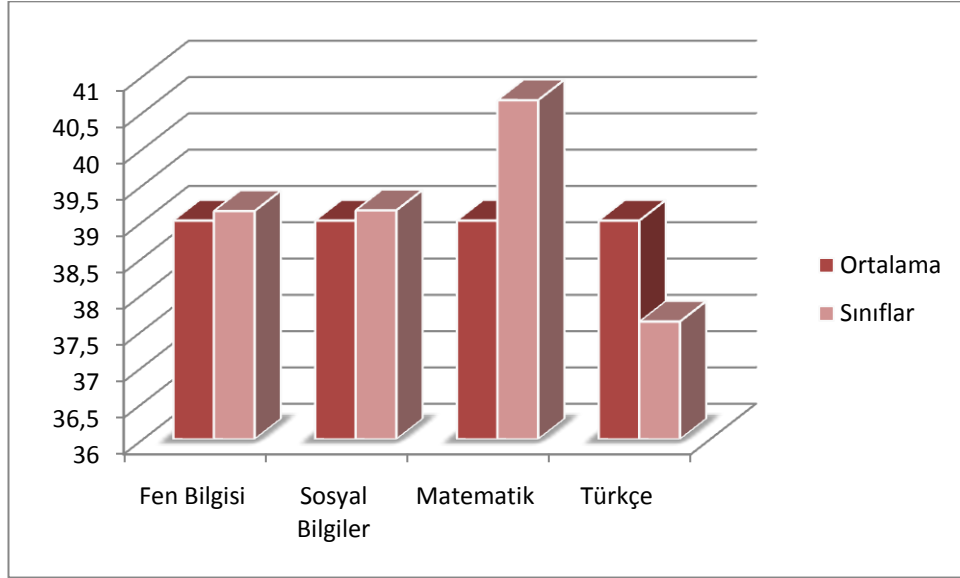
Tablo 4.24. Özel İlgi Faktörü Tukey Testi Sonucu

	(A) Sınıflar	(B) Sınıflar	Ortalamalar Farkı	
			(A-B)	<i>p.</i>
Ölçme ve Değerlendirme Özel İlgi	Fen Bilgisi	Sosyal Bilgiler	3,28	0,063
		Matematik	-0,496	0,990
		Türkçe	3,64	0,049*
	Sosyal Bilgiler	Fen Bilgisi	-3,28	0,063
		Matematik	-3,78	0,085
		Türkçe	0,355	0,994
	Matematik	Fen Bilgisi	0,496	0,990
		Sosyal Bilgiler	3,783	0,085
		Türkçe	4,138	0,066
	Türkçe	Fen Bilgisi	-3,642	0,049
		Sosyal Bilgiler	-0,355	0,994
		Matematik	-4,138	0,066

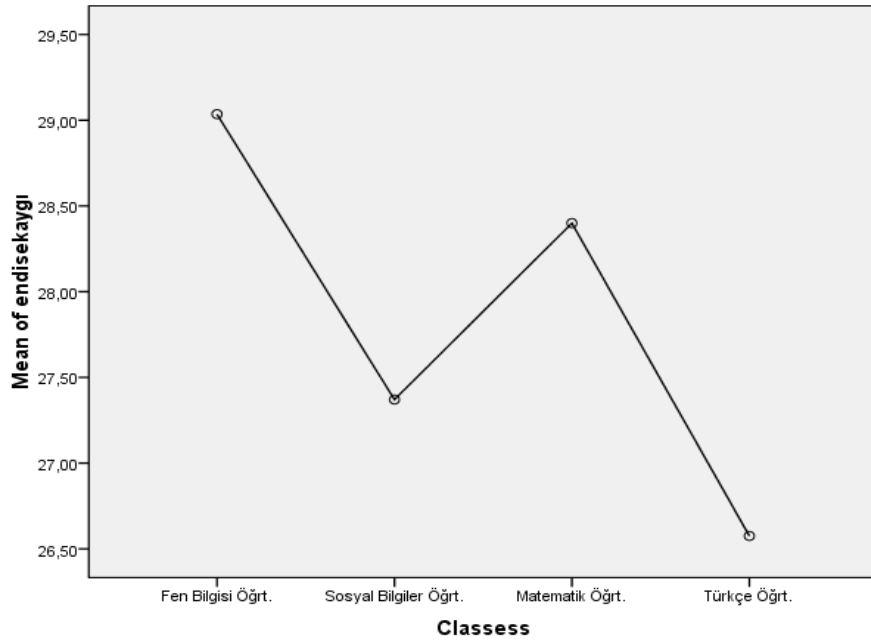
p*<0,05Tablo 4.25.** Endişe-Kaygı Faktörü Toplam Grup Sonuçları

Sınıflar		2. Madde	5. Madde	8. Madde	22. Madde	26. Madde	34. Madde	37. Madde	44. Madde
Fen Bilgisi öğrt.	Ortalama	4,01	3,22	3,92	3,77	3,59	3,94	3,22	3,33
	N	54	54	54	54	54	54	54	54
	<i>S.Sapma</i>	,713	1,110	1,00	,634	1,037	,737	1,021	1,098
Sosyal Bilgiler Öğrt.	Ortalama	4,05	2,86	3,59	3,97	3,11	3,79	3,05	2,91
	N	55	55	55	55	55	55	55	55
	<i>S.Sapma</i>	,692	1,018	1,127	,862	,924	,953	,922	1,023
Matematik Öğrt.	Ortalama	3,85	2,67	4,42	3,78	3,51	3,85	3,17	3,10
	N	28	28	28	28	28	28	28	28
	<i>S.Sapma</i>	1,044	1,278	0,572	0,738	0,833	1,043	1,020	1,100
Türkçe Öğrt.	Ortalama	3,67	3,02	3,31	3,92	3,24	3,73	2,95	2,69
	N	43	43	43	43	43	43	43	43
	<i>S.Sapma</i>	1,016	0,858	1,163	0,910	1,041	0,873	1,068	0,964
Toplam	Ortalama	3,92	2,98	3,75	3,87	3,35	3,83	3,10	3,01
	N	180	180	180	180	180	180	180	180
	<i>S.Sapma</i>	0,852	1,064	1,088	0,792	0,988	0,886	1,000	1,064

Grafik 4.14. Ölçme ve Değerlendirmeye İlişkin Endişe-Kaygı Faktörü Toplam Sonuçları Grafiği



Grafik 4.15. Endişe-Kaygı Scree Plot Grafiği



Genel olarak bakıldığında, tüm öğretmen adaylarının kaygı düzeylerinin yüksek olmadığı görülmektedir ($X=3,92$). İlköğretim öğretmen adaylarının ölçme ve değerlendirmeye ilişkin endişe-kaygı seviyeleri bakımından fen bilgisi ve matematiğin daha az kaygı düzeyine sahip oldukları görülmektedir. Bu farklılığın anlamlı olup

olmadığına eğer varsa hangi gruplara ait olduğuna bakılacaktır. Tablo 4.26.'da gruplar arasında anlamlı farklılıkların olduğu görülmektedir ($df(3, 176) = 3,377$ $p=0,020$). Sadece fen bilgisi ve Türkçe öğretmen adayları arasında anlamlı farklılıklar bulunmuştur ($p=0,018$).

Tablo 4.26. Endişe-Kaygı Faktörü (ANOVA)

Endişe-Kaygı	Kareler Toplamı	df	Kareler Ortalaması	F	p
Gruplar arası	166,874	3	55,625	3,377	,020*
Grup içi	2899,220	176	16,473		
Toplam	3066,094	179			

* $p<0,05$

Tablo 4.27. Endişe-Kaygı Faktörü Tukey Testi Sonuçları

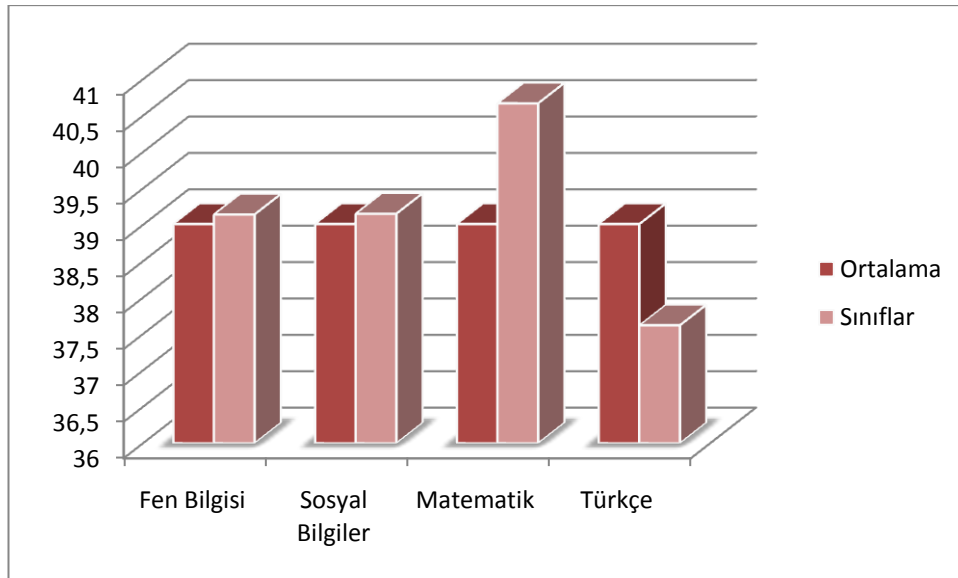
	(A) Sınıflar	(B) Sınıflar	Ortalamalar	
			Farkı (A-B)	p
Ölçme ve Değerlendirme <i>Endişe-Kaygı</i>	Fen Bilgisi	Sosyal Bilgiler	1,66	0,144
		Matematik	0,63	0,907
		Türkçe	2,46	0,018*
	Sosyal Bilgiler	Fen Bilgisi	-1,66	0,144
		Matematik	-1,02	0,695
		Türkçe	0,79	0,771
	Matematik	Fen Bilgisi	-0,63	0,907
		Sosyal Bilgiler	1,02	0,695
		Türkçe	1,82	0,253
	Türkçe	Fen Bilgisi	-2,46	0,018*
		Sosyal Bilgiler	-0,79	0,771
		Matematik	-1,82	0,253

* $p<0,05$

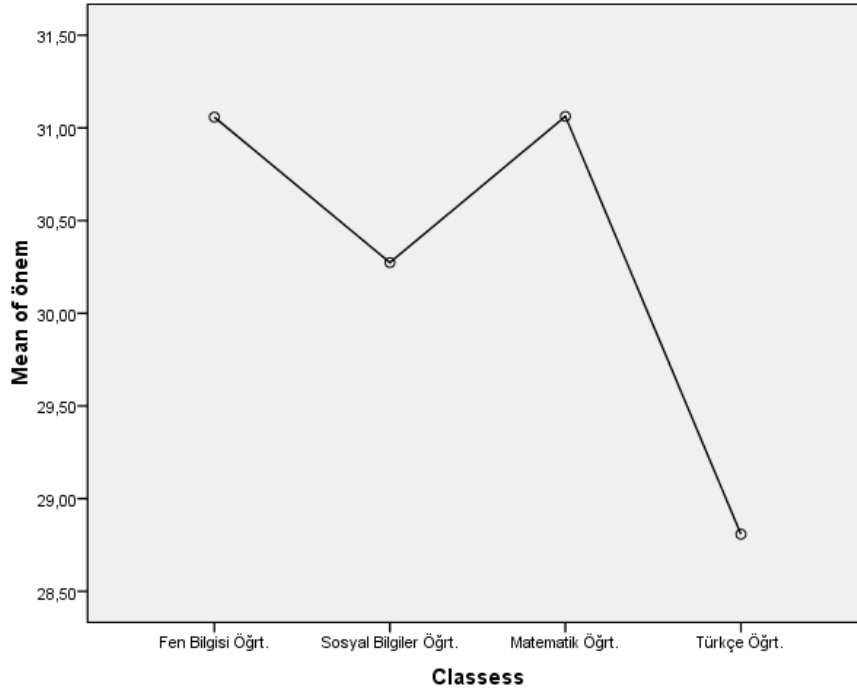
Tablo 4.28. Önem Faktörü Toplam Grup Sonuçları

Sınıflar		6. Madde	9. Madde	17. Madde	29. Madde	35. Madde	40. Madde	42. Madde	46. Madde
Fen Bilgisi Öğrt.	Ortalama	4,16	4,13	3,62	3,49	3,96	4,01	3,72	3,94
	N	54	54	54	54	54	54	54	54
	<i>S.Sapma</i>	0,693	0,753	0,895	0,838	1,045	0,980	0,877	0,711
Sosyal Bilgiler Öğrt.	Ortalama	4,16	4,11	3,74	3,61	3,50	3,51	3,87	3,73
	N	55	55	55	55	55	55	55	55
	<i>S.Sapma</i>	0,908	1,029	0,994	0,855	1,181	1,086	0,859	0,853
Matematik Öğrt.	Ortalama	4,03	4,10	3,51	3,44	3,96	3,92	3,99	4,07
	N	28	28	28	28	28	28	28	28
	<i>S.Sapma</i>	1,035	1,196	1,101	0,874	1,035	1,274	0,721	0,716
Türkçe Öğrt.	Ortalama	3,93	3,79	3,05	3,34	3,52	3,82	3,43	3,88
	N	43	43	43	43	43	43	43	43
	<i>S.Sapma</i>	0,986	1,125	1,070	1,167	1,180	0,973	1,027	0,792
Toplam	Ortalama	4,09	4,04	3,50	3,48	3,71	3,80	3,74	3,88
	N	180	180	180	180	180	180	180	180
	<i>S.Sapma</i>	0,889	1,010	1,028	0,934	1,132	1,072	0,902	0,779

Grafik 4.16. Ölçme ve Değerlendirmeye İlişkin Önem Faktörü Toplam Sonuçları Grafiği



Grafik 4.17. Önem Scree Plot Grafiği



Ortalama olarak bakıldığında önem faktörü açısından, öğretmen adaylarının yüksek derecede tutuma sahip oldukları görülmüştür ($X=4,09$). Bu faktör açısından öğretmen adaylarının ortalamalarının birbirine çok yakın olduğu görülmektedir. ANOVA tablo 4.29.'a bakıldığında gruplar arasında anlamlı bir farklılık görülmemektedir ($df (3, 176)=2,852$ $p=0,056$).

Tablo 4.29. Önem Faktörü (ANOVA)

Önem	Kareler Toplamı	df	Kareler Ortalaması	F	p
Gruplar arası	142,960	3	47,653	2,852	0,056*
Grup içi	2940,715	176	16,709		
Toplam	3083,675	179			

* $p>0,05$

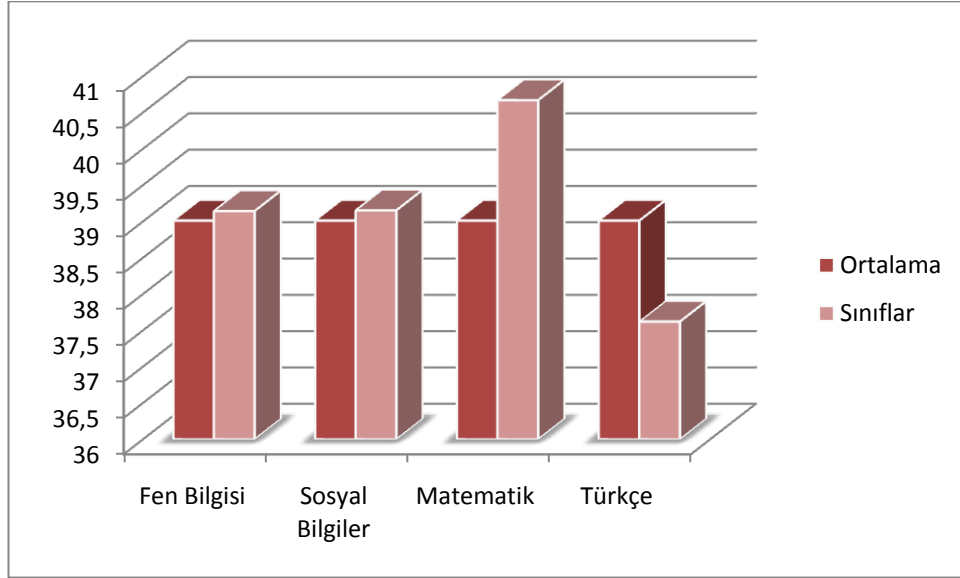
Tablo 4.30. Önem Faktörü Tukey Testi Sonuçları

	(A) Sınıflar	(B) Sınıflar	Ortalamalar Farkı	
			(A-B)	* <i>p</i>
Ölçme ve Değerlendirme Önem	Fen Bilgisi	Sosyal Bilgiler	0,78	0,748
		Matematik	-0,00	1,000
		Türkçe	2,24	0,053
	Sosyal Bilgiler	Fen Bilgisi	-0,78	0,748
		Matematik	-0,78	0,840
		Türkçe	1,46	0,296
	Matematik	Fen Bilgisi	0,00	1,000
		Sosyal Bilgiler	0,78	0,840
		Türkçe	2,25	0,109
	Türkçe	Fen Bilgisi	-2,24	0,053
		Sosyal Bilgiler	-1,46	0,296
		Matematik	-2,25	0,109

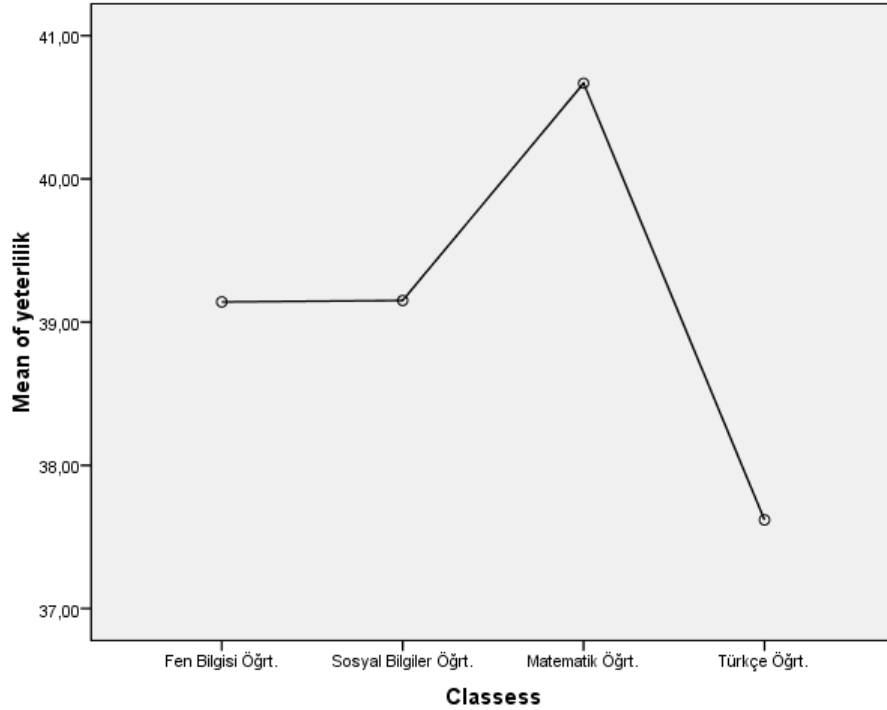
p*>0,05Tablo 4.31.** Yeterlilik Faktörü Toplam Grup Sonuçları

Sınıflar		3. Madde	7. Madde	12. Madde	19. Madde	25. Madde	27. Madde	30. Madde	32. Madde	38. Madde	43. Madde	47. Madde
Fen Bilgisi Öğrt.	Ortalama	3,85	3,64	3,68	3,62	3,42	3,24	3,37	3,53	3,38	3,74	3,62
	N	54	54	54	54	54	54	54	54	54	54	54
	<i>S.Sapma</i>	0,810	0,675	0,842	0,734	1,020	0,822	0,830	0,745	0,855	0,756	0,734
Sosyal Bilgiler Öğrt.	Ortalama	3,31	3,61	3,68	3,68	3,14	3,43	3,62	3,92	3,55	3,71	3,44
	N	55	55	55	55	55	55	55	55	55	55	55
	<i>S.Sapma</i>	1,095	0,980	0,970	1,008	1,190	0,968	0,896	0,706	0,908	0,907	1,094
Matematik Öğrt.	Ortalama	3,92	3,87	3,89	3,39	3,21	3,67	3,87	3,85	3,42	3,78	3,73
	N	28	28	28	28	28	28	28	28	28	28	28
	<i>S.Sapma</i>	0,899	0,833	0,916	1,030	1,066	0,669	0,834	0,650	0,959	0,786	0,798
Türkçe Öğrt.	Ortalama	3,57	3,39	3,54	3,05	3,17	3,33	3,38	3,47	3,45	3,75	3,47
	N	43	43	43	43	43	43	43	43	43	43	43
	<i>S.Sapma</i>	1,003	0,838	0,876	1,155	1,067	0,942	0,898	0,929	1,028	0,893	1,052
Toplam	Ortalama	3,63	3,61	3,68	3,47	3,24	3,38	3,53	3,69	3,46	3,74	3,55
	N	180	180	180	180	180	180	180	180	180	180	180
	<i>S.Sapma</i>	0,987	0,847	0,901	1,005	1,090	0,883	0,880	0,788	0,926	0,836	0,942

Grafik 4.18. Ölçme ve Değerlendirmeye İlişkin Yeterlilik Faktörü Toplam Sonuçları Grafiği



Grafik 4.19. Yeterlilik Scree Plot Grafiği



Yeterlilik faktörü açısından bakıldığında Türkçe öğretmenliğinin diğer sınıflara oranla daha düşük puanlar aldığı görülüyor ($X=37,6188$). Fakat bu farklılığın ne derede önemli olduğunu analiz sonuçları gösterecektir. Genel olarak bu faktör bazında

bakıldığında öğretmen adaylarının ölçme-değerlendirmeyi yapmada kendilerini yeterli hissettikleri görülmektedir ($X=3,63$). ANOVA tablo 4.32.'de görüldüğü gibi gruplar arasında istatistiki olarak anlamlı bir farklılık yoktur ($df(3, 176) = 2,058$ $p=0,108$).

Tablo 4.32. Yeterlilik Faktörü (ANOVA)

Yeterlilik	Kareler Toplamı	<i>df</i>	Kareler Ortalaması	F	<i>p</i>
Gruplar arası	162,181	3	54,060	2,058	0,108*
Grup içi	4623,114	176	26,268		
Toplam	4785,296	179			

* $p>0,05$

Tablo 4.33. Öğretmen Adaylarının Tüm Faktörler Bazında Toplam Sonuçları

		N	Ortalama (X)	S. Sapma	Minimum	Maksimum
Özel İlgi	Fen Bilgisi	54	50,65	6,610	36,39	66,00
	Sosyal Bilgiler	55	47,36	6,862	32,00	66,00
	Matematik	28	51,14	5,883	39,00	63,00
	Türkçe	43	47,01	7,687	33,00	66,00
	Toplam	180	48,85	7,037	32,00	66,00
Endişe-Kaygı	Fen Bilgisi	54	29,03	4,291	19,00	39,00
	Sosyal Bilgiler	55	27,37	3,888	18,00	35,00
	Matematik	28	28,39	3,814	20,00	36,00
	Türkçe	43	26,57	4,123	17,00	36,00
	Toplam	180	27,84	4,138	17,00	39,00
Önem	Fen Bilgisi	54	31,05	3,599	19,00	38,00
	Sosyal Bilgiler	55	30,27	3,985	18,00	38,00
	Matematik	28	31,06	4,468	19,00	37,00
	Türkçe	43	28,80	4,518	18,00	39,00
	Toplam	180	30,28	4,150	18,00	39,00
Yeterlilik	Fen Bilgisi	54	39,14	4,999	24,00	50,00
	Sosyal Bilgiler	55	39,15	4,618	29,00	51,00
	Matematik	28	40,66	4,994	29,00	48,00
	Türkçe	43	37,61	5,922	22,00	50,00
	Toplam	180	39,01	5,170	22,00	51,00

4.1.4. Dördüncü Araştırma Sorusu

İlköğretim öğretmen adaylarının ölçme-değerlendirme okuryazarlığı ve ölçme değerlendirmeye ilişkin tutumları arasında bir ilişki var mıdır?

Öğretmen adaylarının toplam başarı puanları ile toplam tutum puanları arasında anlamlı bir ilişkinin olup olmadığı “Pearson Korelasyon” katsayısına bakılarak analiz edilmiştir. Normal dağılımlar incelendiğinde, tutum ile başarı arasında ortalamalar üzerinde bir ilişki olmadığı gözlenmiştir. Pearson Korelasyon katsayısı = 0,295 değeri bulunmuştur. SPSS verilerine göre korelasyonlar, +1 ile -1 arasında değişmektedir. +1 yüksek oranda doğrusal olarak yüksek bir ilişkinin olduğunu gösterirken, -1 olması ters orantılı bir şekilde mükemmel bir ilişki olduğunu gösterir (Pallant J., 2005).

Tablo 4.34. Pearson Korelasyon İlişkisi

Pearson Korelasyon	r =0,10 ile 0,30 arasında ise ilişki	Çok Zayıf
Pearson Korelasyon	r =0,30 ile 0,50 arasında ise ilişki	Orta
Pearson Korelasyon	r =0,50 ile 1,00 arasında ise ilişki	Çok İyi

Pearson Korelasyon katsayısı = 0,295 değeri yorumlandığında ölçme ve değerlendirmeye ilişkin, tutum ve başarı arasında “Çok Zayıf” oranda ilişki bulunmaktadır.

Tablo 4.35. Tutum ve Başarı Puanları Pearson Korelasyonu

		Tutum	Başarı
Tutum Toplam	Pearson Korelasyon	1	0,295**
	<i>p</i>		0,000
	N	180	180
Başarı Toplam	Pearson Korelasyon	0,295**	1
	<i>p</i>	0,000	
	N	180	180

**Korelasyon 0,01 seviyesinde anlamlandırılmıştır.

5. SONUÇLAR VE TARTIŞMA

Araştırmanın bu bölümünde, çalışmadan elde edilen bulgulara bağlı olarak ulaşılan sonuçlar ve öneriler yer almaktadır.

5.1. Sonuçlar

Araştırmada ilköğretim öğretmen adaylarının (Fen Bilgisi, Sosyal Bilgiler, Matematik ve Türkçe) ölçme ve değerlendirme okuryazarlıkları ve ölçme-değerlendirmeye ilişkin tutumları belirlenmiştir. Ölçme ve değerlendirme kavramları son yıllarda önemli araştırma konusu olmasına ve eğitim fakültelerinde birçok araştırmacının ölçme-değerlendirme üzerine deneysel çalışma yapmasına rağmen, araştırmadan elde edilen istatistikî sonuçlar, ölçme ve değerlendirme okuryazarlık (yeterlilik) seviyesinde öğretmen adaylarının birçok eksiği olduğunu göstermektedir.

Tablo 5.1. Ölçme ve Değerlendirme Okuryazarlığı Toplam Sonuçlar

Sınıflar	Ortalama (X)	N	<i>S.Sapma</i>
Fen Bilgisi	17,57	54	4,015
Sosyal Bilgiler	14,51	55	4,268
Matematik	20,21	28	2,958
Türkçe	17,83	43	4,324
Toplam	17,11	180	4,443

Mertler (2005) yaptığı araştırmada öğretmen adaylarının, 35 madden 19 tanesini doğru cevapladığını söylüyor. Tablo 5.1.'de görüldüğü gibi bu değere en yakın matematik öğretmen adaylarının olduğu görülüyor. Anket ilk kez 6 farklı yapıdaki (ekonomik sosyal farklılıklar vb.) okulda, 555 öğretmene uygulanmıştır. Sonuçlar öğretmenlerin özellikle değerlendirme sonuçlarını yorumlamada, not verme standardında en yüksek puanlara sahip olduklarını fakat uygun değerlendirme metodu seçmede ve geliştirmede çok kötü sonuçların olduğunu belirtmiştir (Plake, 1993). Ülkemiz öğretmen adaylarının ise değerlendirme sonuçlarını paylaşabilme standardından en düşük puanlara sahip oldukları görülmektedir. Campbell ve diğ. (2002) aynı ölçeği öğretmen adaylarına kullanmış ve

öğretmen adayları ortalama olarak değerlendirme okuryazarlığı envanterindeki 35 sorudan 21 tanesine doğru cevap vermişlerdir.

5.1.1. Ölçme ve Değerlendirme Okuryazarlığı Envanteri ile İlgili Sonuçlar

Yapılan envanterin bulguları incelendiğinde, öğretmen adaylarının ölçme ve değerlendirme okuryazarlık seviyesinin yurt dışında yapılan diğer araştırmaların bulgularından daha düşük olduğu görülmüştür. Öğretmen adayları 34 envanter maddesinin ortalama olarak 18 tanesine (%52) doğru cevap vermişlerdir. Bu durumda ölçme ve değerlendirme konusunda öğretmen adaylarının büyük eksikliklerinin olduğu tespit edilmiştir. Özellikle sosyal bilgiler öğretmen adaylarının değerlendirme konusunda çok kötü sonuçlar aldığı görülmektedir ($X=14,51$). En yüksek ortalama matematik öğretmen adaylarında görülmüştür ($X=20,21$). Yapılan ilk çalışmalarda, en yüksek yeterlilik alanı 3. standartta; en düşük yeterlilik alanı ise, 6. standart olmuştur. Yapılan bu çalışmada ise en yüksek puanı 1. yeterlilik alanı en düşüğünü ise 6. yeterlilik alanı almıştır ($KR-20=0,718$). Mertler (2005) yaptığı çalışmada en yüksek performansı 3. standartta; en düşük performansı ise 5. standartta almıştır ($KR-20=0,54$).

5.1.1.1. Yeterlilik Bazında Sonuçlar

Envanter yeterlilik bazında incelendiğinde ise sonuçlar çarpıcıdır.

- 1. Yeterlilik alanı** olan “Uygun ölçme ve değerlendirme yöntemlerini seçme” yeterliliğinde öğretmen adaylarının aynı envanter ile yurt dışında yapılan araştırmalar ile paralel olduğu görülmektedir. Yapılan çalışmada ortalama (X) 3,30 iken; yurt dışında öğretmen adayları ile yapılan araştırmalarda bu oran yaklaşık $X=3,25$ civarındadır. İlk yeterlilik alanında en başarılı grup, $X=3,61$ ortalama ile Türkçe öğretmenliği grubudur. Bu ortalama ile yapılan diğer çalışmalardaki ortalamalar yakındır. En başarısız grup ise, fen bilgisi öğretmenliği grubu olmuştur ($X=2,90$). Yapılan istatistiksel araştırmalar sonucunda, fen bilgisi ile Türkçe öğretmenliği grupları arasında istatistiksel olarak da anlamlı farklılıkların olduğu görülmüştür ($p=0,026$; $p<0,05$).
- 2. Yeterlilik alanı** olan “Uygun ölçme ve değerlendirme yöntemlerini geliştirme” alanında ise ortalamaları $X=2,91$ ’dir. Diğer araştırmalar ile

karşılaştırıldığında bu değer sonuçlara paraleldir. Bu yeterlilik alanında en yüksek ortalama $X=3,22$ ile Türkçe öğretmen adaylarıdır. En başarısız grup ise, $X=2,49$ ile sosyal bilgiler öğretmen adaylarıdır. Gruplar arası ANOVA sonuçlarına göre anlamlı farklılık Türkçe öğretmenliği ve sosyal bilgiler öğretmenliği arasında anlamlıdır (0,013).

- 3. Yeterlilik alanı** “Ölçme ve değerlendirme sonuçlarını yorumlama, puanlama ve uygulama” alanında, öğretmen adaylarının diğer yapılan araştırmalara göre çok düşük puanlar aldığı görülmüştür. En yüksek sonuçları fen bilgisi öğretmen adayları ($X=2,78$), en düşük puanı ise sosyal bilgiler öğretmen adayları almıştır ($X=1,79$). Bu standart açısından öğretmen adayları arasında çok ciddi farkların olduğu görülmektedir. Özellikle fen bilgisi ve matematik öğretmen adaylarına kıyasla sosyal bilgiler öğretmen adaylarının çok başarısız oldukları görülmektedir ($p=0,000$ $p<0,005$). Tüm öğretmen adaylarının bu standart açısından aldıkları puanların ortalaması (X) 2,40 tır. Toplam 5 puan üzerinden düşünüldüğünde, bu oran % 50’ye bile denk gelmemektedir. Öğretmen adaylarının ölçme ve değerlendirme sonuçlarını yorumlama, puanlama konusunda birçok eksiklikleri olduğu görülmektedir. Bu oran yurt dışında yapılan araştırmalarda % 60’ın üzerine çıkmaktadır ($X=3,24$). Bu durum hala yüksek öğretimde geleneksel yöntemlerle ölçme ve değerlendirme uygulamalarının sürdüğünü ve çok acil bir müdahalenin yapılması gerektiğini vurgulamaktadır.
- 4. Yeterlilik alanı** “Eğitimsel kararlar alırken ölçme ve değerlendirme sonuçlarını kullanma” yeterliliği yapılan diğer çalışmalarla paralellik göstermektedir ($X=2,87$). Bu yeterlilik alanında öğretmen adaylarının başarı orana % 60 civarındadır. En yüksek başarıyı, $X=3,00$ ile fen bilgisi öğretmenliği; en düşük başarı oranını ise $X=2,48$ ile sosyal bilgiler öğretmen adayları oluşturmaktadır. İstatistikî olarak matematik öğretmen adayları ile sosyal bilgiler öğretmen adayları arasında anlamlı farklılıklar tespit edilmiştir ($p=0,016$ $p<0,05$).
- 5. Yeterlilik alanı** “Öğrencileri değerlendirirken kullanılacak uygun derecelendirme (rubrik) sistemini seçmede yeterli olma” yeterliliğinde ise, öğretmen adaylarının çok düşük bir puan ortalaması aldığı görülmektedir ($X=2,10$). Öğretmen adayları maalesef yaklaşık % 40 civarında doğru

cevaplar vermişlerdir. Yurt dışında yapılan arařtırmalara paralel olan bu sonuç öğretmen adaylarının bu yeterlilik alanında başarısız olduklarını göstermektedir. Bu yeterlilik alanında en yüksek puanı $X=2,42$ ile matematik öğretmen adaylarının aldığı görülmektedir. Bu standart açısından en başarısız grup ise sosyal bilgiler öğretmen adaylarıdır. Yalnızca sosyal bilgiler öğretmen adayları ile matematik öğretmen adayları arasında istatistikî olarak anlamlı farklılıkların olduğu ortaya çıkmaktadır ($p=0,033$; $p<0,05$).

6. **Yeterlik alanı** “Ölçme ve değerlendirme sonuçlarını paylaşabilme” yeterliliği en çarpıcı sonuçların olduğu alandır. Bu alanda öğretmen adaylarının tümü çok başarısızdır. Ortalama olarak 5 puan üzerinden $X=1,40$ puan almışlardır. Bu değerlendirme öğretmen adaylarının % 35 civarında doğru yanıt verdiğini göstermektedir. Mertler (2005) yaptığı arařtırmada bu değeri $X=2,26$ olarak bulmuştur. İki değer arasında anlamlı farklılıklar bulunmaktadır. Plake (1993) yaptığı benzer arařtırmada öğretmenlerin en düşük olduğu alanı 6. yeterlilik alanı olarak rapor etmiştir. Bu anlamda bir paralellik söz konusudur. Bu alanda en yüksek puanı, matematik öğretmen adayları en düşük puanı ise Türkçe öğretmen adayları almıştır. İstatistikî olarak gruplar arasında matematik öğretmenliği ile Türkçe ve sosyal bilgiler öğretmenliği arasında anlamlı farklılıklar olduğu görülmüştür ($p=0,011$; $p<0,05$).
7. **Yeterlik alanı** “Ölçme ve değerlendirmede kanun dışı ve ahlaki olmayan durumları bilme” yeterliliği ülkemizde üzerinde durulmayan ve ölçme-değerlendirmede çok önemli olan bir yeterlilik alanıdır. Bu alanda ülkemiz öğretmen adaylarının yurt dışına oranla daha düşük puanlar aldıkları görülmüştür. Öğretmen adaylarının ortalama puanları $X=2,11$ 'dir. En başarılı grup, fen bilgisi ve Türkçe öğretmen adaylarıdır. Ortalama olarak puanlarının birbirine yakın oldukları görülmektedir.

5.1.2. Ölçme ve Değerlendirmeye İlişkin Tutum Sonuçları

Yapılan anket sonucunda öğretmen adaylarının tutumları 4 faktör altında toplanmıştır.

Özel ilgi faktörü açısından öğretmen adaylarının tutumlarının genel olarak yaklaşık 3 ile 4 puan arasında ortalamalar aldığı ve tutumlarının iyi olduğu görülmüştür. Yalnızca sosyal bilgiler öğretmen adaylarının tutum seviyesinin ortalamanın aşağısında olduğu görülmektedir. Tutumu en yüksek grup matematik öğretmen adaylarıdır.

Endişe-kaygı faktörü açısından, gruplar arasında çok büyük farklılıkların olmadığı gözlenmiştir. Yalnızca fen bilgisi ve Türkçe öğretmenliği gruplarında anlamlı farklılıklar görülmüştür ($p=0,018$; $p<0,05$). Tutumu en yüksek grup fen bilgisi öğretmenliğidir.

Önem faktörü açısından, öğretmen adaylarının tutumlarının yüksek olduğu ve tutum puanlarının 3 ile 4 arasında değiştiği görülmüştür. Hiçbir grup arasında anlamlı farklılıklar görülmemiştir.

Yeterlilik faktörü açısından, matematik öğretmen adaylarının kendini en çok yeterli bulan grup olduğu anlaşılmaktadır. Bu faktör açısından gruplar arasında anlamlı bir farklılık yoktur.

5.1.3. Ölçme ve Değerlendirme Okuryazarlık Seviyesi ile Ölçme ve Değerlendirmeye İlişkin Tutumlar Arasındaki İlişki

Tutum ve başarı arasında önemli bir korelasyon vardır (Mertler, 2005) yapılan birçok araştırmada bu durum açıkça belirtilmiştir. Fakat ölçme ve değerlendirme üzerine yapılan birçok araştırmada tutum faktörü öğretmenlerin veya öğretmen adaylarının ölçme ve değerlendirme yöntemlerini seçmedeki algıları tutum adı altında araştırılmıştır. Oysaki tutum psikolojik bir objeye yönelen olumlu veya olumsuz bir yoğunluk sıralaması ve derecelemesidir (Thurstone, 1931). Literatür araştırıldığında ölçme ve değerlendirmeye ilişkin bu bakış açısında olan çok az araştırma yapılmıştır.

Başarı puanları ve tutum toplam puanları karşılaştırıldığında öğretmen adaylarının başarı durumu arttıkça tutumlarının da yükseldiği görülmektedir.

6. ÖNERİLER

Günden güne eğitim alanında ölçme ve değerlendirmeye ilişkin araştırmalar artmaktadır. Bunun nedeni ise ölçme ve değerlendirme alanının eğitim sistemlerinde öneminin artmasından kaynaklanmaktadır. İlköğretim öğretmenlerinin işi sınıfta yaptıkları etkinlik ve yapılandırmacı yaklaşımın getirdikleri ile günden güne artmaktadır. Öğretmenlerin en büyük sorumluluğu şüphesiz ki ölçme ve değerlendirme uygulamalarıdır (Mertler, 2003). Yapılan araştırmalar öğretmenlerin, öğrencilerini değerlendirmede hazırlıksız ve yetersiz olduklarını rapor etmişlerdir (Mertler, 1999; Plake, 1993). Yapılan araştırmalara göre, öğretmenler genelde üniversite düzeyinde eğitim fakültelerinde yeterince ölçme ve değerlendirme eğitimi almadıklarından şikâyet ederler. Yaptığımız araştırmada da görüldüğü gibi öğretmen adaylarının bu konuda ne derece eksik olduğu göz önünde bulundurulursa, eğitim fakültelerindeki ölçme ve değerlendirme dersleri ile ilgili büyük değişikliklerin yapılması gerekmektedir. Öğretmen adaylarının, özellikler sosyal ve radikal yapılandırmacılığın etkin olduğu, öğretmen adaylarına özellikle alanın ölçme ve değerlendirmesini kapsayan eğitimler verilmelidir. Okul deneyimi ve özel öğretim yöntemleri dersleri kapsamında alana özgü ölçme ve değerlendirme uygulamalarına yer verilmelidir. Bu araştırmanın devamında ilköğretim öğretmenleri üzerine araştırmalar yapılmalı ve öğretmen adaylarının ölçme ve değerlendirme konusundaki eksikleri belirlenmelidir.

Bu çalışmadan çıkarılan sonuçlar göz önüne alındığında, öğretmen adaylarının ölçme-değerlendirme standartları açısından birçok eksiği olduğu görülmektedir. Bu standartlar göz önünde bulundurularak düzenlenecek bir ders programı öğretmen adaylarının ölçme ve değerlendirme okuryazarlık seviyelerini yükseltecektir. Tüm yeterlilik alanları, alternatif ve otantik ölçme ve değerlendirme yöntemleri ile ilişkilendirilerek ölçme ve değerlendirme dersleri yeniden tasarlanmalıdır. Yapılacak olan deneysel çalışmalarla hem standartların etkililiği hem de ölçme ve değerlendirme seviyeleri ortaya konacaktır. Bu araştırmanın sonuçları gösteriyor ki hem hizmet içi hem de eğitim fakültesi seviyesinde ölçme ve değerlendirme dersleri hayati önem taşımaktadır (Mertler, 2005). Bu eğitimler bu derece önemli olmasına rağmen öğretim üyeleri geleneksel ders yöntemlerinden ve uygulamalarından vazgeçmemektedirler. Öğrencinin hiçbir uygulamayı

yapmadığı gerçek sınıf ortamlarında olmayan, öğrenenden bağımsız işlenen dersler ve ders içerikleri öğretmen adaylarına hiçbir katkı sağlamamaktadır.

Gelecekte yapılacak arařtırmalar özellikle de ilköğretim seviyesinde hizmet içinde çalışan öğretmenler ve öğretmen adayları arasındaki farklılıkları ortaya koymalıdır. Bu çalışmada özellikle öğretmen adaylarının ölçme ve değerlendirme sonuçlarını paylaşmada, geri dönüt vermede çok başarısız oldukları göz önünde bulundurulmalıdır. Öğretmenlere ve öğretmen adaylarına hangi standart açısından müdahale edilmesi gerektiği arařtırılmalıdır.

Sonuç olarak, tüm eğitim bilimciler ve öğretmenler bu konu ile ilgili sorumluluk almalıdır. Bu standartların ortaya çıkmasını sağlayan Amerikan eğitimcileri ve dernekleri gibi ülkemiz milli eğitim kültürüne uygun bir şekilde ölçme ve değerlendirme standartları oluşturulmalıdır. Bu sorumlulukları; okul yönetim, aile, öğretmenler, bilimsel kuruluşlar yüklenmelidir. Öğrencileri değerlendirmek için en uygun, en geçerli, en güvenilir yollar bulunmalıdır.

7. KAYNAKLAR

- American Federation of Teachers, National Council on Measurement in Education, National Education Association. (1990). Standards for teacher competence in educational assessment of students. *Educational measurement: issues and practice*, 9(4), 30-32.
- Arık, S. R. (2006). İlköğretim Öğretmenlerinin Ölçme ve Değerlendirme Alanındaki Kavram Yanılgılarının Belirlenmesi. Yüksek Lisans Tezi, Ankara Üniversitesi Eğitim Bilimleri Enstitüsü.
- Atanur-Başkan, G. (2001). Öğretmenlik Mesleği ve Öğretmen Yetiştirmede Yeniden Yapılanma. Ankara: Denge Matbaacılık
- Atıcı, T; Karahan, U (2006). Osmos ve Bitkisel Hücrelerdeki Plastidler Konularındaki Öğrenci Anlama Düzeylerinin Yapılandırılmış Grid Metodu ve Kısa Cevaplı Sorularla Değerlendirilmesi. VII. Fen Bilimleri ve Matematik Eğitimi Kongresi
- Atılğan, H., Kan, A. ve Doğan, N. (2009). *Eğitimde Ölçme ve Değerlendirme*. Ankara: Anı Yayıncılık.
- Aydın, A. (2001). Eğitim Fakültesi Mezunu Olan ve Olmayan Öğretmenlerin Ölçme ve Değerlendirme Yeterliklerinin Karşılaştırılmasına Yönelik Bir Çalışma. Yayımlanmamış yüksek lisans tezi. Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Aykaç, N. ve Başar, E. (2005). İlköğretim Sosyal Bilgiler Dersi Eğitim Programının Değerlendirilmesi. Eğitimde Yeni Yansımalar VIII, Yeni İlköğretim Programlarını Değerlendirme Sempozyumu. Ankara: Sim Matbaası, ss. 343-361.
- Bahar, M. Öztürk, E. ve Ateş, S. (2002). “Yapılandırılmış Grid Metodu İle Lise Öğrencilerinin Newton’un Hareket Yasası, İş, Güç Ve Enerji Konusundaki Anlama Düzeyleri ve Hatalı Kavramlarının Tespiti”, V. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi’nde Sözlü Bildiri, ODTÜ, Ankara, 2002.
- Bahar, M., Nartgün, Z., Durmuş S. ve Bıçak, B. (2008). Geleneksel- Alternatif Ölçme ve Değerlendirme Öğretmen El Kitabı. Ankara: Pegem A Yayıncılık.
- Bahar, M., Nartgün. Z., Durmuş. S. ve Bıçak, B. (2006) Geleneksel-Alternatif Ölçme ve Değerlendirme Öğretmen El Kitabı. Ankara: Pegem A Yayıncılık.

- Baki, A. & Birgin, O. (2004). Reflections of Using Computer-Based Portfolios as an Alternative Assessment Tools: A Case Study. *The Turkish Online Journal of Educational Technology*, 3 (3), Article 11, <http://www.tojet.net/articles/3311.htm>
- Bandura, A. (1997). *Self-Efficacy: The Exercise of Control*. New York: W.H. Freeman Company.
- Barenholz, H., & Tamir, P. (1992) A comprehensive use of concept mapping in desing instruction and assessment. *Research in science & Technological Education*, 10, 37-52
- Bayrakçeken, S., (2008). Test geliştirme. İçinde, E. Karip (Ed.). Ölçme ve değerlendirme, ss:244–274, Ankara: Pegem Akademi Yayıncılık.
- Beckwith, J. B. (1991) Approaches to learning, their context and relationship to assessment performance. *Higher Education* 22 , pp. 17-30
- Boud, D.(1995) *Enhancing Learning Through Self Assessment*. London and Philadelphia, Kogan Page
- Brooks, G.P. and Johanson, G.A., (2003). Test analysis program. *Applied Psychological Measurement*, 27, pp:305-306.
- Campbell, C., Murphy, J., & Holt, J. (October 2002). Psychometric analysis of an assessment literacy instrument: Applicability to preservice teachers. Paper presented at the annual meeting of the Mid- Western educational Research Association. Columbus, OH.
- Chase, C.I., (1999). *Contemporary assessment for educators*, New York: Longman.
- Chen, Y., Martin, M.A. (2000). Using Performance Assessment and Portfolio Assessment Together in Elementary Classroom. *Reading Improvement*. 37(1). 32-38.
- Collins, A. (1992) Portfolios for Science Education. *Science Education*. 76. 451-463.
- Cubertson, L.D., Laongo, M.R.(1999). But What's With Letter Grades?. *Childhood Education*. 75(3). 130-135.
- Curtis, M.J., Hunley, S.A., Chesno Grier, J.E. (2002). Relationships Among the Professional Practices and Demographic Characteristics of School Psychologists. *School Psychology Review*. 31. 30–42.
- Çakan, M. (2004). Öğretmenlerin Ölçme ve Değerlendirme Uygulamaları ve Yeterlik Düzeyleri: İlk Ve Ortaöğretim. *Ankara Üniversitesi Eğitim Bilimleri Dergisi*, 37 (2), 99-114.

- Çelik, D. (2000). Okullarda Ölçme ve Değerlendirme Nasıl Olmalı? İstanbul: *Milli Eğitim Bakanlığı Yayınları*
- Daniel, L.G. and D.King., (1998). A knowledge and use of testing and measurement literacy of elementary and secondary teachers. *Journal of Educational Research*. 91(6), pp:96-100.
- Demirel, Ö. (2002). Planlamadan Değerlendirmeye Öğretme Sanatı. Ankara: Pegem A.
- Demirel, Ö. (2006). Eğitimde Program Geliştirme. Ankara: Pegem A Yayıncılık
- Doğan, B.A. (2005). Fen Öğretiminde Değerlendirme Etkinlikleri Üzerine Öğretmen Görüşleri (Van İli Örneği). Yüksek Lisans Tezi, Van Yüzüncü Yıl Üniversitesi, Fen Bilimleri Enstitüsü.
- Duru, M. K. ve Gürdal, A. (2002). İlköğretim fen bilgisi dersinde kavram haritasıyla ve gruplara kavram haritası çizdirilerek öğretimin öğrenci başarısına etkisi, V. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi, 16-18 Eylül, ODTÜ, Ankara.
- Enger, S.K. ve Yager, R.E. (1998). The Iowa Assessment Handbook. (ERIC Dökümanı Servis Numarası: ED424286).
- Erdemir, A. Z. (2007). İlköğretim İkinci Kademe Öğretmenlerinin Ölçme Değerlendirme Tekniklerini Etkin Kullanabilme Yeterliklerinin Araştırılması. Yayınlanmamış Yüksek Lisans Tezi, Kahramanmaraş Sütçü İmam Üniversitesi, Sosyal Bilimler Enstitüsü, Kahramanmaraş.
- Ertürk, S. (1982). Eğitimde Program Geliştirme. Ankara: Meteksan Ltd. Şti.
- Eyitmiş, A.N. (2007). Ortaöğretim Öğretmenlerinin Ölçme Değerlendirme Tekniklerini Etkin Kullanabilme Yeterliklerinin Araştırılması (Kahramanmaraş Örneği). Yüksek Lisans Tezi, Kahramanmaraş Sütçü İmam Üniversitesi. Sosyal Bilimler Enstitüsü.
- Fox, R. (2001). Constructivism Examined, Oxford Review of Education, 27,(1).
- Gelbal, S. ve Kelecioğlu, H. (2007). Öğretmenlerin Ölçme ve Değerlendirme Yöntemleri Hakkındaki Yeterlik Algıları ve Karşılaştıkları Sorunlar, *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33, 135-145.
- Gilman, D.A., Hasset, M.(1995). More than Work Folder: Using Portfolios for Educational Assessment, Clearing House, 68(5), 310-312.
- Goodrich, H. (1997). Understanding Rubrics. *Educational Leadership*. 54 (4).14-17

- Gözütok, D., Ergün, E.Ö. ve Karacaoğlu, C. Ö., (2005). İlköğretim Programlarının Öğretmen Yeterlikleri Açısından Değerlendirilmesi, Eğitimde Yeni Yansımalar VIII, Yeni İlköğretim Programlarını Değerlendirme Sempozyumu. Ankara: Sim Matbaası, ss:17-40.
- Gözütok, D., Ergün, E.Ö. ve Karacaoğlu, C. Ö., (2005). İlköğretim Programlarının Öğretmen Yeterlikleri Açısından Değerlendirilmesi, Eğitimde Yeni Yansımalar VIII, Yeni İlköğretim Programlarını Değerlendirme Sempozyumu. Ankara: Sim Matbaası, ss:17-40.
- Gummer, S. ve Shepardson, D. P. (2001). The Nrc Standarts as a Tool in the Professional Development of Science Teachers Assessment Knowledge and Practice. Eds. Shepardson, D.P. Assessment in science a guide to Professional development and classroom practice. 39–51. Dordrecht: Kluwer Academic Publishers.
- Gümüş, B. (1981). Ölçme ve Değerlendirme Dersinin Öğretmenlerin Öğrenci Başarısını Değerlendirme Davranışına Katkısı. Yayımlanmamış yüksek lisans tezi. Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Hair, J. F. Anderson, R. E., Tahtam, R. L. ve Black, W. C. (1998). Multivariate Data Analysis. Fifth Edition, Pearson Education.
- Hamilton, L.S. (1994). Validating Hands-on Science Assessments Through an Investigation of Response Process. American Educational Research Association'ın New Orleans, LA. ERIC Document Reproduction Service No: ED 376 202.
- Hargreaves, J. (n.d). Engendering student learning by use of specific assessment strategies, http://www.tedi.uq.edu.au/Conferences/A_conf/papers/Hargreaves.html
- Harlen, W. (1998). The Teaching of Science in Primary Schools. (Second Edition). Great Britain: The Cromwell Press, Trowbridge.
- Harris, M (1997) Self-assessment of language learning in formal settings. *ELT Journal* cilt:51, sayı:1, ss. 12-20.
- Hegarty-Hazel, E. and Prosser, M. (1991b). Relationship between students' conceptual knowledge and study strategies. Part II: Student learning in biology. *International Journal of Science Education*, 13, 421-429.
- Hein, G.E. ve Price, S. (1994). Active Assessment for Active Science A Guide for Elementary School Teachers. Portsmouth: Heinemann.

- Howe, C. A. and Jones, L. (1998). Engaging Children in Science. (Second Edition). New Jersey: Macmillan College Publishing Company, Prentice- Hall Inc. http://eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01
- İşman, A. ve Eskicumalı, A. (2003). Eğitimde Planlama ve Değerlendirme. İstanbul: Değişim Yayınları
- Karip, E. Ve Diğerleri (2007). Ölçme ve Değerlendirme (Ed. Emin Karip). Ankara: PegemA Yayıncılık
- Kaynak, S. (2000). Ortaöğretimdeki Branş Öğretmenlerinin Öğrenci Başarısını Ölçme ve Değerlendirme ile İlgili Görüşlerinin Değerlendirilmesi. Yayımlanmamış yüksek lisans tezi, Mustafa Kemal Üniversitesi, Sosyal Bilimler Enstitüsü, Hatay.
- Kehoe, J., (1995). Basic item analysis for multiple-choice tests. *Practical Assessment*.
- Kemertaş, İ. (2003). Öğretimde Planlama ve Değerlendirme. İstanbul.
- Koç, N. (1981). Liselerde Öğrencilerin Akademik Başarılarının Değerlendirilmesi Uygulamalarının Etkililiğine İlişkin Bir Araştırma. Ankara: *Ankara Üniversitesi Eğitim Fakültesi Yayınları*.
- Korkmaz, H. (2004). Fen ve Teknoloji Eğitiminde Alternatif Değerlendirme Yaklaşımları. Ankara: Yeryüzü Yayınevi
- Küçükahmet, L. (1999). Öğretimde Planlama ve Değerlendirme. Ankara: Alkım Yayıncılık.
- Küçükahmet, L. (2005). 2004 Hayat Bilgisi Programının Değerlendirilmesi. Eğitimde Yeni Yansımalar VIII, Yeni İlköğretim Programlarını Değerlendirme Sempozyumu. Ankara: Sim Matbaası, ss: 373-381.
- Leahy, S., Lyon, C., Thompson, M., & William, D. (2005). Classroom assessment: minute by minute, day by day. *Educational Leadership*, 63(3), 18-24.
- Manning, M., & Gary (1995). Portfolios in Reading and Writing. *Teaching Pre K-8*. 25(5). 94-95.
- Martin, S. (1994). The Mentoring Process in Preservice Teacher Education. School
- Mayo, S. (1967). Preservice preparation a teachers in educational measurement (Report No. 5-0807). Washingtoni, DC: Office of Education.
- McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, 36(4), 475-492.

- McMillian, J.H. (1997). Classroom assessment: Principles and practice for effective instruction. Needham Heights. MA: Allyn and Bacon.
- McNamara, M. J. and Deane, D. (1995). Self-assessment activities: Toward autonomy in language learning, *TESOL Journal*, sayı:5, ss17-21
- McNamee, G., & Chen, J. (2005). Dissolving the line between assessment and teaching. *Educational Leadership*, 63(3), 72-76.
- McTighe, J., & O'Connor, K. (2005). Seven practices for effective learning. *Educational Leadership*, 63(3), 10- 17.
- MEB. (1999). 2000 Yılında Milli Eğitim. Ankara: Milli Eğitim Basımevi Gökçe E. (1999). İlköğretim Öğretmenlerinin Yeterlikleri, A.Ü. Sosyal Bilimler Enstitüsü, Ankara: Yayınlanmamış Doktora Tezi, Ankara.
- MEB. (2002). Okulda performans yönetimi (taslak). Ankara: Milli Eğitim Basımevi.
- Mertler, A.C. and Campbell, C., (2005). Measuring teachers knowlede& application of classroom assessment concepts: development of the assessment literacy inventory. Annual Meetin of American Educational Research Association, Quebec, Canada, 11-15, April.
- Mertler, A.C., (2004). Secondary teachers assessment literacy: does classroom experience make a difference. *American Secondary Education*, 33(1), pp:49–64.
- Mertler, C. (2003). Preservice versus inservice teachers' assessment literacy: does classroom experience make a difference? Paper presented at the annual meeting of the Mid-Western Educational Research Association. Columbus, Ohio.
- Mertler, C. A. (2003). Classroom assessment: A practical guide for educators. Los Angeles, CA: Pyczak.
- Milli Eğitim Bakanlığı (MEB) (2005). İlköğretim Fen ve Teknoloji Dersi: Öğretim Programı ve Kılavuzu (Taslak Basım). Ankara: Devlet Kitapları
- Moon, Tonya R., (2002). Development of Differentiated Performance Assessment Tasks for Middle School Classrooms. Office of Educational Research and Improvement, Washington, DC. ERIC Document Reproduction Service No: ED 476 371.
- Moskal, B.M (2000). Scoring Rubrics: What, When and How? *Practical Assesment, Research & Evaluation*. 7 (3).
<http://pareonline.net/getvn.asp?v=7&n=3>. 25.01.2011

- Nıgıudula, D. (1993). The Digital Portfolio.
- Nitko, A.J., (2001). Educational assessment of students (3th edition). New Jersey-USA: Merrill Prentice Hall
- Novak, J. , Gowin, B. , (1984), “Learning How to Learn” , Cambridge University Press, USA.
- Nunan, D. (1988) The learner centered curriculum. Cambridge:CUP
- Nunnally, J.C. and Bernstein, I.H., (1994). Psychometric theory (3th edition). New York: Mc Graw-Hill.
- Okur, M. (2008). 4. ve 5. Sınıf Öğretmenlerinin Fen ve Teknoloji Dersinde Kullanılan Alternatif Ölçme ve Değerlendirme Tekniklerine İlişkin Görüşlerinin Belirlenmesi. Yayınlanmamış Yüksek Lisans Tezi, Zonguldak Karaelmas Üniversitesi, Sosyal Bilimler Enstitüsü, Zonguldak.
- Oosterhof, A. (1999). Developing and Using Classroom Assessments. Upper Saddle River. NJ: Prentice Hall
- Orhan, T. A. (2007). Fen Eğitiminde Alternatif Ölçme ve Değerlendirme Yöntemlerinin İlköğretim Öğretmen Adayı, Öğretmen ve Örgenci Boyutu Dikkate Alınarak İncelenmesi. Yayınlanmamış Doktora Tezi, Gazi Üniversitesi Eğitim Bilimleri Enstitüsü, Ankara.
- Ornstein, A. C. ve Thomas J.L. (2004). Strategies for Effective Teaching. (Fourth Edition). New York: The McGraw-Hill Companies Inc.
- Ören, F.S. (2005). İlköğretim 7. Sınıf Fen Bilgisi Dersinde Öğrenme Halkası Yaklaşımının, Öğrencilerin Başarı, Tutum ve Mantıksal Düşünme Yetenekleri Üzerine Etkisi. Yayınlanmamış Doktora Tezi. Gazi Üniversitesi, Eğitim Bilimleri Enstitüsü. Ankara.
- Özçelik, D. A. (1981). Okullarda Ölçme ve Değerlendirme, Ankara: ÖSYM Yayınları
- Özçelik, D. A. (1992). Ölçme ve Değerlendirme. Ankara: ÖSYM Yayınları
- Özden, M., (2005), “Kimya Öğretiminde Okul ve Dershane Eğitiminin Karşılaştırılması”. VII. Ulusal Fen Bilimleri ve Matematik Eğitimi Bilimleri Kongresi, Gazi Üniversitesi, Gazi Eğitim Fakültesi, Ankara.
- Öztürk, M. (2005). Değerlendirme. Editör: Öztürk, M. Öğretimde Planlama ve Değerlendirme. 223- 234. İstanbul: Lisans Yayıncılık.
- Pallant, 2005 J. Pallant, SPSS Survival Guide (second ed.), Open University Press Maidenhead (2005).

- Paterno, J. (2001). Measuring success: A glossary of assessment terms. In *Building cathedrals: Compassion for the 21st century*. Retrieved July 24, 2003, from <http://www.angelfire.com/wa2/buildingcathedrals/measuringsuccess.html>
- Pierce, L. V. ve O'Malley. J. M. (1992). Performance and Portfolio Assessment for Language Minority Students. Program Information Guide Series, 9.(ERIC Dökümanı Servis Numarası: ED347747).
- Pierce, L. V. ve O'Malley. J. M. (1992). Performance and Portfolio Assessment for Language Minority Students. Program Information Guide Series, 9.(ERIC Dökümanı Servis Numarası: ED347747).
- Plake, B.S. & Impara, J.C. (1993). Teacher assessment literacy questionnaire. University of Nebraska- Lincoln. In cooperation with the National Council on Measurement in Education and the W.K. Kellogg Foundation.
- Plake, B.S. (1993). Teacher assessment literacy: Teachers' competencies in the educational assessment of students. *Mid-Western Educational Researcher*, 6(1), 21-27.
- Plake, B.S., Impara, J.C., & Fager, J.J. (1993). Assessment competencies of teachers: A national survey. *Educational Measurement: Issues and Practice*, 12(4), 10-12, 39.
- Popham, J.W. (1997). What is Wrong and What is Right with Rubrics. *Educational Leadership*. 55 (2). 72-77
- Popham, W. J. (2003). Test better, teach better – the instructional role of assessment. Association for Supervision and Curriculum Development: Alexandria, VA.
- Popham, W. J. (2004). All about accountability: why assessment illiteracy is professional suicide. *Educational Leadership*, 62(1), 82-83.
- Popham, W. J. (2005). Classroom assessment: what teachers need to know. *Pearson Education, Inc.:* Boston.
- Popham, W. J. (2006). Needed: a dose of assessment literacy. *Educational Leadership*, 63(6), 84-85.
- Reeves, T.C, and Okey, J.R. (1996). Alternate Assessment for Constructivist Learning Environments. Editör: In B. Wilson. *Constructivist Learning Environments: Case Studies in Instructional Design*. 191-202. Englewood Clirffs, NJ: Educational Technology Publications

- Romberg, T.A. (1993). How One Comes to Know Models and Theories of the Learning of Mathematics. Editör: M. Niss. Investigations into Assessment in Mathematics Education. 97-111. Netherlands: Kluwer Academic Publishers.
- Rowntree, D., (1987). Assessing students: how shall we know them?. London: Kogan Page
- Schafer, W. D., & Lissitz, R. W. (1987). Measurement training for school personnel: Reconrmmentations and reality. *Journal of Teacher Education*, 38(3), 57-63.
- Semerci, Ç. (2008). Eğitimde Ölçme ve Değerlendirme. Editör: Karip, E. Ölçme ve Değerlendirme. 1- 15. Ankara: Pegem A Yayıncılık.
- Shepard L. (2000). The Role of Assessment in a Learning Culture. *Educational Researcher*. 20(7). 2-16.
- Sönmez, V. (2007). Program Geliştirmede Öğretmen El Kitabı Ankara: Anı Yayıncılık
- Stiggins, R. (1999). Are you assessment literate? *The High School Journal*, 6(5), 20-23.
- Stiggins, R. (2000). Learning teams for assessment literacy. *Orbit*, 30(4), 5-7.
- Stiggins, R. (2004). New assessment beliefs for a new school mission. *Phi Delta Kappan*, 86(1), 22-27.
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Ph Delta Kappan*, 83(10), 758-765.
- Stiggins, R.J., (1995). Assessment literacy for the 21st century. *Phi Delta Kappan*, 77(3), pp:238-245.
- Tan, Ş. ve Erdoğan, A. (2004). Öğretimi Planlama ve Değerlendirme. Ankara: PegemA Yayıncılık.
- Tan, Ş., Kayabaşı, Y. ve Erdoğan, A. (2003). Öğretimi Planlama ve Değerlendirme. Ankara: Anı Yayıncılık
- Tekin, H. (1996). Eğitimde Ölçme ve Değerlendirme. Ankara: Yargı Yayınları
- Tekin, H. (2000). Eğitimde ölçme ve değerlendirme, Yargı Yayınları, Ankara.
- Tekin, H., (1991), Eğitimde ölçme ve değerlendirme. Ankara: Yargı Yayınevi
- Tekindal, S. (2002). Okullarda Ölçme ve Değerlendirme Yöntemleri. Kocaeli.
- Thompson, S. J., Benson, S. N. K., Pachnowski L. M. ve Salzman, J. A. (2001). Decision-Making in Planning and Teaching. Addison-Wesley Educational Publishers Inc., United States.
- Thorndike, M. R. (1997). Measurement And Evaluation In Psychology and Education. New Jersey: Prentice Hall.

- Trowbridge, J.E., & Wandersee, J.H. (1994). Identifying critical junctures in learning in a college course on evolution. *Journal of Research in Science Teaching*, 31, 459–473.
- Turgut, F. (1997). Eğitimde Ölçme ve Değerlendirme Metotları. Ankara: Nüve Matbaası
- Ulutaş, S. (2003). Genel Liselerdeki Öğretmenlerin Ölçme ve Değerlendirme Alanındaki Yeterlikleri İle Ölçme ve Değerlendirme İlkelerini Uygulama Düzeylerinin Araştırılması. Yayımlanmamış yüksek lisans tezi. Ankara Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara.
- Uysal, Ş. (1961). Evaluation Of Pupil Progress In Secondary Schools Of Turkey. Yayımlanmamış doktora tezi, University of Illinois, Urbana.
- Valencia, S. (1990). A Portfolio Approach to Classroom Assessment: The Whys, Whats And Hows. *The Reading Teacher*. 43 (4). 338-340
- Variş, F. (1996). Eğitimde Program Gelistirme- Teori ve Teknikler. (6. Baskı). Ankara: Alkım Kitapçılık Yayıncılık
- Volante, L. ve Fazio, X. (2007). Exploring Teacher Candidates' Assessment Literacy: Implications For Teacher Education Reform And Professional Development, *Canadian Journal Of Education*, 30, 3, 749–770.
- Whittaker, C.R., Salend, S.J., Duhaney, D. (2001). Creating Instructional Rubrics for Inclusive Classroom. *Teaching Exceptional Children* 34 (2). 8- 13.
- Yaman, S., Karamustafaoğlu, O., ve Karamustafaoğlu, S. (2005). Fen ve Teknoloji Eğitiminde Kavram Öğretimi Editör: Aydoğdu, M., ve Kesercioğlu, T. İlköğretimde Fen ve Teknoloji Öğretimi. 236- 278. Ankara: Anı Yayıncılık.
- Yaşar, Ş., Gültekin, M., Türkan, B., Yıldız, N. ve Girmen, P. (2005). Yeni İlköğretim Programlarının Uygulanmasına İlişkin Sınıf Öğretmenlerinin Hazır Bulunuşluk Düzeylerinin ve Eğitim Gereksinimlerinin Belirlenmesi. Eğitimde Yeni Yansımalar VIII, Yeni İlköğretim Programlarını Değerlendirme Sempozyumu Ankara: Sim Matbaası, ss: 51-63.
- Yıldırım, C. (1999). Eğitimde Ölçme ve Değerlendirme. Ankara: ÖSYM Yayınları
- Yıldız, Nida. (2005). 6. Sınıf Fen Bilgisi Dersinde Rubrik Yöntemiyle Öz Değerlendirme Çalışması. Eğitimde İyi Örnekler Konferansı. 15-16 Ocak 2005. İstanbul: Sabancı Üniversitesi.
<http://www.erg.sabanciuniv.edu/iok2005bildiriler/nidayildiz.doc>. 12.02.2011.

Zollman, A. ve Jones, D. L. (1994). Accommodating Assessment and Learning: Utilizing Portfolios In Teachers Education With Preservice Teachers. (ERIC Dökümanı Servis Numarası: ED 368 551).

8. EKLER

Ek 1. Sınıfta Değerlendirme Okuryazarlığı Envanteri (SDOE)

Amerikan Eğitimde Ölçme Birliği ve W.K. Kellogg Derneği işbirliğiyle Nebraska-Lincoln Üniversitesinde, Barbara S. Plake & James C. Impara tarafından hazırlanan ”*Teacher Assessment Literacy Questionnaire*” (1993) envanterinden Türkçeye uyarlanmıştır”

Bu envantere toplam 35 soru vardır ve bu sorular, “Eğitimde Ölçme ve Değerlendirmeyi Yapmada Öğretmen Yeterlilikleri” ile ilgili 7 standarttan oluşmaktadır. Sorular ölçme ve değerlendirme kavramlarıyla, öğrencilere not verirken neleri göz önünde bulundurduğumuz, ölçme ve değerlendirme aktiviteleri, ölçme ve değerlendirme sonuçlarının hem aile hem de öğrencilerle paylaşımı, bazı maddeler de standart testlerle ve sınıfta değerlendirme ile ilgilidir. İkinci kısımda ise eğitimde ölçme ve değerlendirmeye karşı tutumunuzu (ölçme ve değerlendirmeyi sevip sevmediğiniz, kaygılarınız ve yeterliliğiniz ile ilgili) belirlemek için tutum anketi verilmiştir.

Her soruyu dikkatlice okuyup yalnızca düşündüğünüz tek seçeneği işaretleyiniz. Eğer hiçbir seçeneğin doğru olmadığını düşünüyorsanız size en yakın gelen ve doğru cevap olabileceğini düşündüğünüz seçeneği işaretleyiniz.

Arş. Gör. Emrah GÜL
Fırat Üniversitesi, 2011

1. Öğrencilerinizin, öğrettiğiniz konuyla ilgili başarılarınızı değerlendirmek için uygun bir ölçme ve değerlendirme yöntemi seçerken aşağıdakilerin hangisine daha çok önem verirsiniz?

- A. Kolayca not verebileceğiniz bir yöntem olmasına
- B. Değerlendirme aracının kolayca hazırlanabilmesine
- C. Öğretimsel amacın, gerçekleşip gerçekleşmediğini gösteren bir yöntem olmasına
- D. Okul yönetimi tarafından kabul edilecek bir yöntem olmasına

2. Bir testin “güvenilir” olmasını, aşağıdaki ifadelerden hangisi en iyi şekilde açıklar?

- A. Bu testten alınan notlar tüm eğitimsel kararlarda kullanılabilir.
- B. Herhangi bir öğrenci bu testi tekrar çözerse, benzer sonuçlar alır.
- C. Bu testin sonucu öğretmenin kararından daha doğrudur.
- D. Bu testin sonucu öğretilen konuları tamamen ölçebilecek niteliktedir.

3. Emrah öğretmen, işlediği dersle ilgili olarak öğrencilerinin konuyu öğrenip öğrenmediklerini değerlendirmek istiyor. Buna göre, aşağıdakilerden hangisi bu değerlendirmeyi yapmak için en geçerli yol olur?

- A. Öğretmen rehber kitabından aldığı bir testi kullanması
- B. Sınıfta öğrettiklerini yansıtan bir değerlendirme yöntemi seçmesi ve uygulaması
- C. Öğretmenin kendi hazırladığı bir testi kullanması
- D. Öğretmenin bir tutum anketi yapması

4. Bir öğretmen öğrencilerinin daha başarılı olmalarını sağlamak amacıyla aşağıdaki yollardan hangisini başvurmalıdır?

- A. Bir matematik problemini çözüme ulaştırdıklarında, onlara not vermelidir.
- B. Öğrencilerinin verdiği cevaplara bireysel olarak dönüt sağlamalıdır.
- C. Öğrencilerini, bir problemin çözümüyle ilgili motive etmelidir.
- D. Yukarıdakilerin hiçbiri.

5. Aşağıdaki belgelerden hangisi bir öğretmenin öğrencilerini, dersleri ve akademik başarıları ile ilgili motive edip etmediğini anlamada en geçerli kanıttır?

- A. Öğretmenin kendi ders planı
- B. MEB müfredatı ders planı
- C. Öğretmenin öğrencileri değerlendirmede kullandığı yöntem-teknik veya kullandığı testlerden birer kopya
- D. Öğretmenin öğrencilerine not vermeden, onlarla ilgili bazı bilgileri yazdığı bir belge

6. Öğretmen, öğrettiği konu ile ilgili bir ölçme ve değerlendirme stratejisi belirler; öğrencilerine notlar verir ve bu notların geçerli olup olmadığını belirlemek ister. Aşağıdaki bilgilerden hangisi öğretmenin bu amacı gerçekleştirip gerçekleştirmediğini en iyi şekilde kanıtlar?

- A. Kullanılan ölçme ve değerlendirme stratejisinin, diğer öğretmenler tarafından değerlendirilmesi
- B. Kullanılan ölçme ve değerlendirme stratejisinin öğretim içeriğinin kazanımlarıyla uyumlu olması
- C. Ölçme ve değerlendirmenin geçerli olup olmadığını, öğrencilerden aldığı fikirlerle belirlemesi
- D. Öğrenci velisine, öğrencinin öğretilen konuyla ilgili kazanımlarını sağlayıp sağlamadığının sorulması

7. Bir çoktan seçmeli testin güvenilirliği nasıl arttırabilir?

- A. Testteki soruları geliştirmek için, bir taslak kullanarak
- B. Çoktan seçmeli testin formatını doğru-yanlış formatına çevirerek
- C. Teste ekstra soru ekleyerek
- D. Testin sonuna yazılı bir cevap gerektiren soru ekleyerek

8. Kenan öğretmen, öğrencilerinin bilgileri ezberlemesinden ziyade fikirleri organize edebilme, problem çözebilme, üst düzey düşünme vb durumlarını değerlendirmek için birer paragraf yazdırıp değerlendirmek istiyor. Aşağıdaki kelime gruplarından hangisi bu paragraflarda olması gerekenleri tamamen yansıtır?

- A. Karşılaştırma, zıtlık, eleştirme
- B. Tanımlama, özelleştirme, listeleme
- C. Düzen, uyum, seçme
- D. Tanımlama, hatırlama, yeniden isimlendirme

9. Türkçe öğretmeni Deniz Hanım, öğrencilerinin Orhan Pamuk'un eserlerini ve edebi yapısını anlayıp anlamadıklarını belirlemek için bir test uyguluyor. Bu testindeki hangi soru bu amaca daha uygundur?

- A. “Sessiz Ev” ve “Kar” Orhan Pamuk'un eserleri midir?
- B. Doğru mu yanlış mı? “Orhan Pamuk annesini babasını tanımayan bir yetimdir.”
- C. Doğru mu yanlış mı? “Orhan Pamuk; şiir, roman, hikâye yazmıştır”.
- D. Orhan Pamuk'un edebiyat dünyasına katkısı nedir? Bununla ilgili kısa bir paragraf yazınız.

10. Matematik öğretmeni Ahmet Bey'in sınıfındaki birkaç öğrenci ünite sonundaki dört işlem problemleri testinden düşük notlar almıştır. Ahmet Bey, benzer sorunu yaşayan öğrencileri gruplandırmak için bu öğrencilerin kimler olduğunu bilmek istiyor. Aşağıdaki değerlendirme yollarından hangisi bu öğrencileri gruplamak için en uygundur?

- A. Öğretmen rehber kitabından aldığı çoktan seçmeli bir testi uygulamak
- B. Öğretmenin konuyla ilgili süreci tamamen kapsayan bir testi hazırlayıp uygulaması
- C. Öğrencilerin test, yazılı ve sözlü kayıtlarına bakmak
- D. Öğrencilere dört işlem problemleri verip önce çözmelerini daha sonra da adım adım anlatmalarını istemek

11. Birçok öğretmen genelde notlarını 100 üzerinden değerlendirir. Böyle bir not sisteminden 90 puan alan bir öğrenci için aşağıdakilerden hangisi doğru olur?

- A. Bu öğrenci testteki soruların % 90'ını doğru bir şekilde cevaplamıştır.
- B. Bu öğrenci testi kapsayan ünitenin % 90'ını bilir.
- C. Bu testi çözen tüm öğrencilerin % 90'ından daha yüksek not almıştır.
- D. Bu öğrenci sınıftaki ortalama bir öğrenciden % 90 daha fazla not almıştır.

12. Esmâ öğretmen, Fen Bilgisi dersinde dönem sonu notunu vermek için öğrencilerinden güneş sisteminin yapısı ile ilgili bir model geliştirmelerini istemiştir. Hangi ölçme ve değerlendirme stratejisi bu modelleri değerlendirmede ve not vermede doğru bir seçim olur?

- A. Esmâ öğretmen tüm modelleri topladıktan sonra en ilgi çeken ve en güzel olan modeli belirler ve ona en yüksek notu verir. Daha sonra gelen modellere de bu modelle kıyaslayarak diğer modellerin notlarını belirlemesi
- B. Esmâ öğretmen diğer öğretmenlerden her bir projenin niteliğini gerçek anlamda ortaya çıkaracak 5'li likert bir ölçek hazırlamaları istemesi ve bu ölçeğe göre değerlendirmesi
- C. Projeler teslim edilmeden önce esmâ öğretmen sınıfta performansı yüksek olan çalışkan öğrencilerine bu projeleri değerlendirmek için bir derecelendirme ölçeği hazırlamalarını istemesi ve ölçme ve değerlendirmeyi bu ölçeğe göre yapması
- D. Projeler teslim edilmeden önce Esmâ öğretmen projelerin belirli özelliklerini yansıtacak bir derecelendirme (rubrik) ölçeği geliştirmesi ve bu ölçeğe göre en yüksek notu alan en başarılı proje seçmesi

13. Eğitim-Öğretim yılı bahar döneminin sonlarına doğru Ali öğretmen Sosyal Bilgiler dersi için kitaplardan ve bazı testlerden faydalanarak geliştirdiği bir testi öğrencilerine veriyor. Bu testte bir paragraf ve bu paragraf ile ilgili bazı sorular yer almaktadır. Bu testi uyguladıktan sonra sınıfta performansı yüksek olan, derse aktif katılan başarılı iki öğrencisinin kötü notlar aldıklarını fark ediyor. Bu bilgilere aşağıdakilerden hangisi eklenirse, Ali öğretmen bu sınav sonucunu daha iyi yorumlar?

- A. Öğrencilerin cinsiyeti
- B. Öğrencilerin yaşı
- C. Model olarak kullanılan testin, güvenilirliği ve geçerliliğiyle ilgili bilgiler
- D. Öğrencilerin paragraf not sonuçları

14. Öğretmeni, çok başarılı bir 5. sınıf öğrencisi olan Gizem'e Türkçe dersinde okuduğunu anlama konusunda 8. sınıf düzeyinde sorular soruyor ve Gizem'in tüm soruları doğru cevapladığını görüyor. Bunun anlamı...

- A. Gizem 8. sınıf düzeyinde okuyabilir ve anlayabilir.
- B. Gizem'in okuduğunu anlama notu 8. sınıftaki herhangi bir öğrencinin notu ile aynıdır.
- C. Gizem okuduğunu anlamada 8. sınıf düzeyinde bir başarı gösterir.
- D. Gizem 8. sınıfa geldiğinde okuduğunu anlamada en başarılı öğrenci olacaktır.

15. Bir sınavda hangi bölüme ne kadar zaman ayrılacağı belirtilmiştir. Örneğin Fen Bilgisi soruları için, 30 dk veya Türkçe soruları için 45 dk vb. Bunlara zaman yönergeleri denmektedir. Aşağıdaki öğrenci davranışlarından hangisi zaman yönergeleri açısından en kabul edilebilir davranıştır?

- A. Ali'nin ilk olarak Matematik testini bitirdikten sonra, bu testteki cevaplarını kontrol etmesi
- B. Ahmet'in ilk olarak Matematik testini bitirdikten sonra, önceden çözdüğü bölümleri kontrol etmesi
- C. Deniz'in ilk olarak Matematik testini bitirip, daha sonra diğer testteki sorulara şöyle bir göz atması fakat hiçbir soruyu işaretlememesi
- D. Ulaş'ın Matematik testini bitirmeden diğer testlere geçmemesi ve en sonunda testteki tüm soruları çözmeden zamanını tüketmesi.

16. Ömer öğretmen, dönemin ilk ünitesi olan sayılar konusu ile döneme başlamıştır. Üniteye başlamadan önce çarpma işleminin değişme ve dağılma özelliğiyle ilgili öğrencilerine bir test vermiş ve çözmelerini istemiştir. Aşağıdakilerden hangisi öğretmenin bu davranışının nedeni olabilir?

- A. Okul yöneticisi bu testin sonuçlarını istemesi
- B. Ömer öğretmen öğrencilerine dönem başından itibaren çeşitli ödevler ve uygulamalar vermek istemesi
- C. Ömer öğretmen üniteye başlamadan önce öğrencilerinin önceden öğrendikleri bilgilerini kontrol etmek istemesi
- D. Ömer öğretmen, bu kavramlarla ilgili öğrenci başarısındaki gelişimi ölçmek istemesi

17. Bir matematik öğretmeni 6. sınıf Matematik programının etkililiğini değerlendirmek için öğrencilerine 8. sınıf Matematik testi verip çözmelerini istemiştir. Ve bu testten elde ettiği sonuçları normalde 8. sınıfta olan öğrencilerin notlarıyla karşılaştırmıştır. Bu uygulama neden doğru değildir?

- A. Bu durum 6. sınıf seviyesi için güvenilir değildir.
- B. Bu durum 6. sınıf seviyesi için geçerli değildir.
- C. 8. sınıf matematik konuları 6. sınıf seviyesine göre çok zordur.
- D. 6. sınıf öğrencilerinin test çözme zamanı çok sınırlıdır.

18. Bir matematik öğretmeni kesirler ünitesinde aritmetik işlemler konusunu planlarken, aşağıdaki bilgilerden hangisi kendisi için daha yararlı olacaktır?

- **Norm-dayalı bilgi:** Her öğrencinin performansının, diğer öğrencilerin performansına göre değerlendirilmesi
- **Kritere dayalı bilgi:** Her bir öğrenciyi öğrenme çıktılarına göre ortaya koyduğu bireysel başarıyı referans alarak değerlendirme

- A. Norma-dayalı bilgi
- B. Kritere dayalı bilgi
- C. Öğretimi planlamak için her iki tür bilgiye de ihtiyacı vardır.
- D. Öğretimi planlamak için her iki bilgi türüne de ihtiyacı yoktur.

19. Öğrencinin SBS'deki (Seviye Belirleme Sınavı) başarısı ile sınıf içindeki öğretmenin yaptığı uygulamalar sonucundaki başarısı tutarsız olursa; Aşağıdakilerden hangisi bu durumun nedenini en iyi açıklar?

- A. Bazı öğrenciler SBS'de aşırı heyecan ve stres yaşadıklarından başarılı olamazlar; fakat sınıf içerisinde daha rahattırlar.
- B. Öğrenciler çoğu zaman SBS' deki testleri sınıf içindeki performanslarından daha az ciddiye alırlar.
- C. SBS' deki testler daha çok ezbere dayalı bilgileri ölçer, fakat sınıfta yapılan değerlendirmeler daha üst düzey (mantıksal ve nedensel düşünme, problem çözme, fikir üretme vb...) olduğu düşünülen bilgileri ölçer.
- D. SBS'nin sınıfta yapılan değerlendirmelerden daha az geçerliğe sahip olması

20. Özel bir kolej, tüm öğretmenlerinin işbirliği ile okul adına yeni bir Fen Bilgisi ve Matematik müfredatı geliştirmiştir. Fakat geliştirilen bu müfredat SBS'ye yönelik değildir. Ve o yıl bu okulun SBS'de en başarısız okul olduğu tespit edilmiştir. Aşağıdakilerden hangisi bu kolej ile diğer okullar arasındaki karşılaştırmada geçersiz bir yorum olacaktır?

- A. Diğer okulların müfredatı, SBS 'ye daha yakındır.
- B. Diğer okullar kendi öğretim müfredatlarını SBS' ye daha yakın organize etmişlerdir.
- C. Bu özel kolejin eğitim-öğretim seviyesi oldukça düşüktür.
- D. Diğer okulların müfredat programları bu özel kolejinkinden farklıdır.

21. Aşağıdaki ölçme ve değerlendirme yöntemlerinden hangisi, öğrenci başarı ve performansını en iyi şekilde ortaya koyar ve bir öğretmen için en güvenilir bir değerlendirmeye olanak sağlar?

- A. Ünitenin öğretim amaçlarına uygun 2 ya da 3 soruluk klasik sorulardan oluşan bir sınav yapmak
- B. Ünitenin öğretim amaçlarını kapsayan 20 soruluk çoktan seçmeli bir sınav yapmak
- C. Öğretmenin ünitenin sonunda öğrencilere sorduğu sorular ve yüz yüze aldığı cevaplar ile değerlendirme yapmak
- D. Öğretmenin ünite boyunca günlük olarak öğrencilerine sorduğu sorular ve onların performans göstergeleri

22. Öğretmenler bir ünite boyunca genelde 3 sınav yaparlar ve ağırlıklı ortalama olarak öğrencinin karne notunu belirlerler. Bu uygulamanın amacı öğrencileri başarılarına göre sıralamaktır. Bu amaca ulaşmak için aşağıdaki ifadelerden hangisinin eşit olması gerekir?

- A. Her 3 sınavdaki soruların sayısı
- B. Her sınavda testi yapan öğrencilerin sayısı
- C. Her sınavdaki ortalama notlar
- D. Her sınavda alınan notların çeşitliliği

23. Bir çocuğunun ebeveynleri (anne-babaları) çocuklarının dersleriyle ve notlarıyla ilgili öğretmenden bilgi almak istediğinde, öğretmenin aşağıdaki davranışlardan hangisini yapması en uygun olur?

- A. Öğrencinin sınıftaki performansına ve diğer ilgili faktörlere bağlı olarak verdiğini açıklamalıdır.
- B. Ebeveynlere, çocuklarına not verirken öğretmenlerin nasıl davranması gerektiğini sormalıdır.
- C. Öğretmenler, öğrencinin çalışmalarına, performanslarına ve ödevlerine bakarak onlara nasıl not verdiklerini ebeveynlere tam olarak açıklamalıdır.
- D. Not vermede sadece okul yönetiminin etkisi olduğunu, öğretmenlerin kesinlikle bu konuda söz sahibi olmadıklarını ebeveynlere belirtmelidir.

24. Bir öğretmenin, aşağıdaki not verme stratejilerden hangisini kullanması, öğrenci başarısını en az yansıtır?

- A. Fatma öğretmen çoktan seçmeli 20 soruluk bir test uygular, fakat sadece tek rakamlı olan sorulara not verir.
- B. Ahmet öğretmen, ünite içerisinde 3 büyük sınav ve haftalık küçük sınavlar yaparak öğrenci başarısını değerlendirir.
- C. Efe öğretmen öğrencilerine notlarını yükseltmeleri için devamlı fırsatlar vererek onlar istedikleri zaman not verme şansı verir.
- D. Osman öğretmen öğrencilerinin her kötü davranışı için, onlardan 5 puan kırarak not verir.

25. Çilem öğretmen dönem boyunca ev ödevi vermez, ünite bitince sadece bir sınav yapar ve öğrencilerine bu şekilde karne not verir. Aşağıdakilerden hangisi onun bu şekilde not vermesine en büyük eleştiridir?

- A. Bu şekilde değerlendirme yanlıştır, çünkü bazı öğrenciler tek sınavda başarısız olabilirler
- B. Bu şekilde değerlendirme yanlıştır, çünkü not verirken birçok test sonucundan yola çıkılarak verilmelidir.
- C. Bu şekilde değerlendirme yanlıştır, çünkü öğretmenin yaptığı bu test müfredatın içeriğine göre çok sınırlı olacaktır.
- D. Bu test ünitenin tamamını kapsıyorsa önemli bir eleştiri yapmaya gerek yoktur. Bu uygulama doğrudur.

26. Emrah öğretmen, bir öğrencisinin SBS'deki matematiksel hesaplamalar testi sorularına, matematiksel kavramlarla ilgili test sorularından daha fazla doğru cevap verdiğini fark ediyor. Bu durum şu anlama gelir:

- A. Emel'in matematiksel hesaplamalar testi ile ilgili sorulardaki ortalaması düşüktür.
- B. Emel matematiksel kavramlar konusunda başarılı bir öğrencidir.
- C. Emelin, matematiksel kavramlar testi ile matematiksel hesaplamalar testi sonuçları birbirinden farklıdır.
- D. Matematiksel kavramlar testi oldukça zordur.

27. Birçok ülke SBS (Seviye Belirleme Sınavı) ve STS (Seviye Tespit Sınavı) gibi sınavlarla ilgili olarak okulların başarı farklılıklarını açıklamak için, okul programlarını gözden geçirip incelemeye alırlar. Aşağıdaki bilgilerden hangisi bu farklılığı açıklamada gerekli değildir?

- A. Her okuldaki öğrenci sayısı
- B. Her okulun sosyoekonomik düzeyi
- C. Her okuldaki etnik farklılıklar
- D. Her okulda başarısız olan öğrenci sayısı

28. Aşağıda bir öğrencinin bazı derslerde almış olduğu notlar verilmiştir.

(100 üzerinden değerlendirme)

- ❖ Fen Bilgisi: 75
- ❖ Sosyal Bilgiler: 75
- ❖ Matematik: 75

Aşağıdakilerden hangisi alınan bu notlar için geçerli bir yorumdur?

- A. Selma 3 testin her birine eşit sayıda doğru cevap vermiştir.
- B. Selma her 3 testte de çok başarılıdır.
- C. Selma her 3 testte de aynı yüzdelerle sıradadır.
- D. Selma her 3 testte de ortalamasının üstündedir.

29. Önder öğretmen öğrencilere notlarını genelde ev ödevleri ve yaptığı testlerden aldıkları sonuca göre verirken, Esmâ Öğretmen ise öğrencilerin sınıf içindeki performanslarına bağlı olarak not vermektedir. Bu iki değerlendirme stratejisi aşağıdaki çiftlerden hangisidir?

- A. Formal ve informal değerlendirme
- B. Performans ve uygulamalı değerlendirme
- C. Şekillenmiş ve özel değerlendirme
- D. Toplam ve biçimlendirici değerlendirme

30. Hasan Matematik sınavında 60, Fen Bilgisi sınavında ise 57 almıştır. Hasanın bu notlarını 2. sınavlarda her ikisi için % 5 oranında artırması durumunda Hasan'ın öğretmeni ne yapmalıdır?

- A. Bu farklılığı görmezden gelmelidir.
- B. Fen Bilgisi dersi için ona bireysel yardımda bulunmalıdır.
- C. Okul dışında da Hasan'ı Fen Bilgisi çalışmaya yönlendirmelidir.
- D. Performansı Matematikte daha yüksek olduğu için, onu bu alanda daha fazla çalışmaya yönlendirmelidir.

31. Tülin öğretmen, yaptığı sınavın sonuçlarını öğrencilerine bir an önce duyurmak istemektedir. Bu nedenle öğrencilerine sınav kâğıtlarını ders çıkışında öğretmenler odasından alıp notlarına bakabileceklerini söylüyor. Öğrenciler de diğer arkadaşlarının kâğıtlarını görerek kendi notlarına bakıyorlar. Tülin öğretmenin bu konuda yaptığı yanlış nedir?

- A. Öğrencilerin tümü diğer arkadaşlarının notlarını görür ve diğer öğrencilerin gizlilik hakkı ihlal edilmiş olur.
- B. Ders bitince okuldan ayrılmak zorunda olan öğrencileri bekletmek, doğru bir davranış olmaz.
- C. Tülin öğretmen sınav kâğıtlarını hızlı hızlı okuyacağından özel yardıma ihtiyacı olan öğrencileri dinleyemeyecektir ve bu durum onlar için olumsuz olacaktır.
- D. Okula gelmeyen öğrenciler için bir dezavantaj olacak, çünkü bu öğrencilerin kopya çekmesine olanak sağlayacaktır.

32. Bir ülkenin Mili Eğitim Bakanlığı okullara kaynak dağıtımını, okulların ülke çapında yapılan sınavlardaki başarısını dikkate alarak yapmaktadır. Yani başarılı ve başarısız okulları belirleyerek bu duruma göre kaynakları dağıtmaktadır. Hangi özlü söz bu durumu en iyi tanımlar?

- A. Her bulutun bir gümüş kaplaması vardır.
- B. Hayatta bazen yağmurlu zamanlar vardır.
- C. Yağmurda hep denizin üstüne yağarmış.
- D. Eldeki bir kuş, daldaki iki kuştan daha iyidir.

33. Ünite sonunda çoktan seçmeli bir test yapan bir öğretmen, sınavdan sonra bazı öğrencilerin cevap kâğıdında baskı hatasından dolayı bazı soruların seçeneklerinin eksik olduğunu fark ediyor. Buna göre; hangi öğretmen davranışı daha etik olur?

- A. Mustafa öğretmen, derslerinde çok başarılı olmayan Ali'nin kâğıdındaki eksiklikleri rastgele tamamlar ve bu durumun Ali için adil olacağını düşünür.
- B. Esmâ öğretmen, o gün sınavda kendini iyi hissetmeyen Faruk'un kâğıdını onun sınıf içindeki performansını düşünerek işaretlemeleri yapar (ör.: Faruk bu soruyu doğru yapardı ya da bunu yapamazdı gibi...).
- C. Hasan öğretmen, Eda iyi bir öğrenci olmasına rağmen kâğıdında hiçbir değişiklik yapmaz, yapacağı değişikliğin Eda'ya çok daha iyi bir not aldirabileceğini düşünür.
- D. Çilem öğretmen, çok başarılı bir öğrenci olan Emrah'ın tüm soruları doğru cevaplayacağını düşünerek tüm soruları okur ve doğru cevabı işaretler.

34. Sevda öğretmen, öğrencilerinin SBS (Seviye Belirleme Sınavı)' den yüksek not alamayacaklarından çok endişelidir. Bir gün Sevda öğretmenin eline SBS soruları geçer. Aşağıdaki davranışlardan hangisi Sevda öğretmen için etik değildir?

- A. Öğrencilerine sınavda nasıl davranmaları gerektiğini ve süreyi nasıl kullanmaları gerektiğini açıklamak
- B. Eline geçen testin benzer bir formunu yapıp öğrencilere dağıtmak.
- C. Öğreteceği konuları elindeki testin içeriğine göre planlamak
- D. Yukarıdakilerden hiçbiri etik değildir.

35. Öğrenim? I. Öğretim II. Öğretim

Fen Bilgisi Öğretmenliği		
Sosyal Bilgiler Öğretmenliği		
Türkçe Öğretmenliği		
Matematik Öğretmenliği		

36. Cinsiyetiniz

Bay	
Bayan	

Ek 2. EĞİTİMDE ÖLÇME VE DEĞERLENDİRMEYE İLİŞKİN TUTUM ÖLÇEĞİ

Bu anket bir ilköğretim öğretmeni olarak sizlerin eğitimde ölçme ve değerlendirmeye ilişkin tutumunuzu belirlemek için oluşturulmuştur. Bu amaç ile ilgili 47 cümle verilmiştir. Her bir ifadeyi okuduktan sonra inandığınız ya da düşündüğünüz yalnızca bir cevabı işaretleyiniz. Cevaplarınızda dürüst ve içten olmanız, çalışmanın amacı için çok önemlidir. Lütfen samimiyetle cevap veriniz. Zaman ayırdığınız için teşekkürler.

Arş. Gör. Emrah GÜL

	Tamamen Katılıyorum	Katılıyorum	Kararsızım	Katılmıyorum	Hiç Katılmıyorum
1. Öğrencilerimi ölçme ve değerlendirme uygulamaları içerisinde görmekten mutluluk duyarım.					
2. Sınıfımda ölçme ve değerlendirme uygulamaları sırasında kendimi rahat hissederim.					
3. Çok fazla uğraşsam bile ölçme ve değerlendirmeyi etkin bir şekilde yapamam.					
4. Ölçme ve değerlendirme uygulamaları için harcayacağım zaman ve gayretin amacına ulaşacağına inanıyorum.					
5. Ölçme ve değerlendirme bilgimi etkili bir şekilde kullanamamaktan dolayı endişe duyarım.					
6. Ölçme ve değerlendirme uygulamaları, hangi ders olursa olsun bir konunun öğretimi için çok faydalıdır.					
7. Ölçme ve değerlendirme uygulamaları sırasında yeni bir durumla karşılaştığımda ne yapacağımı bilirim.					
8. Ölçme ve değerlendirme uygulamalarını yapmaktan korkarım.					
9. Ölçme ve değerlendirme bilgisi, tüm öğretmenlerin sahip olması gereken en temel bilgilerden biridir.					
10. Öğretim sürecinde yapmayı en çok istediğim uygulamalar, ölçme ve değerlendirme uygulamalarıdır.					
11. Ölçme ve değerlendirme sürecinin beklediğim gibi ilerlememesi bende gerginlik yaratır.					
12. Ölçme ve değerlendirme yaparken bir sorunla karşılaştığımda çözüm bulamam.					
13. Ölçme ve değerlendirme uygulamalarına bir an önce başlamayı istiyorum.					
14. Bir öğretmen zamanının çoğunu, ölçme ve değerlendirme uygulamalarını geliştirmeye ayırmalıdır.					
15. Sınıflarımda ölçme ve değerlendirme ile ilgili yaptığım her şey kendimi mükemmel hissetmemi sağlar.					
16. Öğrencilerimi ölçme ve değerlendirme uygulamalarına etkin olarak dâhil etmede başarısız olmaktan korkarım.					
17. Bir öğretmenin, en çok ölçme ve değerlendirme ile ilgili bilgi, beceri ve uygulamalarını geliştirmesi gerekir.					

	Tamamen Katılıyorum	Katılıyorum	Kararsızım	Katılmıyorum	Hiç Katılmıyorum
18. Öğrencilerimi ölçme ve değerlendirme uygulamalarında yeterince cesaretlendirememekten korkarım.					
19. Üniversite eğitiminde ölçme ve değerlendirme uygulamalarına etkin olarak katılabilirim.					
20. Ölçme ve değerlendirme uygulamaları ilgimi çekmez.					
21. Bir öğretmen olarak, ölçme ve değerlendirme en önemli ilgi alanlarımdan biridir.					
22. Ölçme ve değerlendirme uygulamalarını planlarken kendimi rahat hissederim.					
23. Ölçme ve değerlendirme ile ilgili gelişmeleri öğrenmek için gelecekte mutlaka bir hizmet içi eğitime katılmayı isterim.					
24. Ölçme ve değerlendirme uygulamalarının, etkili bir öğretim için gerekli olduğunu düşünmüyorum.					
25. Ölçme ve değerlendirme uygulamalarını yapmak için kendimi yeterli görmüyorum.					
26. Ölçme ve değerlendirme sonuçlarını yorumlarken tedirginlik yaşıyorum.					
27. Ölçme ve değerlendirme uygulamalarını kalabalık sınıflarda da yapabilirim.					
28. Ölçme ve değerlendirme ile ilgili gelişmeleri takip etmemi sağlayacak kitap, dergi ve makale okumaktan sıkılırım.					
29. Ölçme ve değerlendirmeyi ne kadar iyi yaparsam, o kadar başarılı bir öğretmen olurum.					
30. Ölçme ve değerlendirme uygulamalarını, yöntemlerini ilgili bilgi ve deneyimlerimi vs... gelecekte meslektaşlarıma öğretebilirim.					
31. Ölçme ve değerlendirmeden kaynaklanan sorunları çözmek en son isteyeceğim şeydir.					
32. Ölçme ve değerlendirme uygulamalarında karşılaşılabilecek problemleri çözebilirim ve farklı çözüm yolları geliştirebilirim.					
33. Kullanabileceğim ölçme ve değerlendirme ile ilgili yeni uygulamaları araştırmaktan ve öğrenmekten hoşlanırım.					
34. Ölçme ve değerlendirme sonuçları hakkında öğrencilerime dönüt verirken kendimi rahat hissederim.					
35. Ölçme ve değerlendirme uygulamalarının, öğrencilerin öğrenecekleri konuyu daha anlamlı ve kalıcı olarak öğrenmelerine katkısı yoktur.					
36. Ölçme ve değerlendirme uygulamalarını yaparken çok eğlenirim.					
37. Ölçme ve değerlendirme uygulamaları için zamanın yetmeyeceği konusunda kaygılıyım.					
38. Ölçme ve değerlendirme yapabilmek için yeterli bilgi ve beceriye sahibim.					
39. Ölçme ve değerlendirme ile ilgili bir sorunla karşılaştığımda bu sorunu çözene kadar vazgeçmem.					
40. Ölçme ve değerlendirmenin gereksiz ve zaman kaybı olduğunu düşünüyorum.					
41. Ölçme değerlendirme uygulamalarından kaynaklanan sorunları çözmek için çaba sarf etmem.					
42. Ölçme ve değerlendirme uygulamalarının öğrencilerin üst düzey düşünme becerileri geliştirmede etkili olacağını düşünüyorum.					
43. Ölçme ve değerlendirme uygulamalarının zorlukları olmasına rağmen başarabilirim.					
44. Ölçme ve değerlendirme uygulamaları sürecinde, öğrencilerin yaşayacakları sorunları çözememekten endişe duyarım.					
45. Meslektaşlarım, okul yönetimi ve velilerle en çok ölçme ve değerlendirme üzerinde konuşmaktan haz duyarım.					
46. Ölçme ve değerlendirme sayesinde bir konunun öğretimiyle ilgili eksik olduğum yerleri görürüm.					
47. Sınıfımda ölçme ve değerlendirme uygulamalarıyla ilgili yeni yöntemler geliştirebilirim.					

Ek 3. İzin Belgesi

Ek-4

Topic: Permission to use and adapt "Teacher Assessment Literacy Questionnaire"
from Dr. Barbara Plake

12.03.2011

Emrah GÜL



Emrah GÜL

scienceeduemrah@hotmail.com

Kime: Barbara S. Plake bplake1@unl.edu

Kimden Emrah GÜL (scienceeduemrah@hotmail.com)

Gönderme tarihi: 12 Mart 2011 Cumartesi 17:28:53

Kime: bplake1@unl.edu

Dear Professor Barbara S. Plake;

My name is Emrah GÜL. I am a graduate student at Firat University, Turkey. I have long been working on the concept of "Assessment Literacy". I want to work with this concept in my thesis. Therefore; You have created "Teacher Assessment Literacy Questionnaire (1993)", I would like to adapt and use this survey

I would like your permission in this regard.

Sincerely

Res.Assist. Emrah GÜL

Hakkari University, Faculty of Education,

Department of Educational Sciences

Assessment & Evaluation Department (A&E)

Topic: Permission to use and adapt “Teacher Assessment Literacy Questionnaire” from Dr. Barbara Plake

12.03.2011

Plake barbara



plake barbara

bplake@unlserve.unl.edu

Kime: Emrah GÜL

Kimden: **Barbara Plake** (bplake@unlserve.unl.edu)

Gönderme tarihi: 12 Mart 2011 Cumartesi 18:23:14

Kime: Emrah GÜL (scienceeduemrah@hotmail.com)

2 ek (toplam 52,3 KB)



Tcher Ass...doc

Karşıdan yükle(52,0 KB)



KEY Liter...doc

Emrah,

I have attached two documents, the Teacher Assessment Literacy Questionnaire (Parts 1 & 2) and the key for Part 1. I do not own copyright for this instrument; the National Council on Measurement in Education does. NCME has released the instrument for research purposes, so you should be able to adapt it for your thesis providing that you cite the original instrument.

Barbara Plake

ÖZGEÇMİŞ

RSQJUISGPw 80x

Özlük Bilgileri

Ad: EMRAH **Soyadı:** GÜL
Cinsiyet: Erkek **Baba Adı:** YAŞAR
Doğum Tarihi: 31.01.1987

RSQJUISGPw 80x

İletişim Bilgileri

Telefon: Cep: 0 (507) 959 98 49
e-posta: gulemrh@gmail.com

RSQJUISGPw 80x

Lise Bilgileri

Lise	Şehir	Giriş Tarihi	Mezuniyet Tarihi	Ortalama
ELAZIĞ LİSESİ	ELAZIĞ	9/2001	6/2004	4.73/5

RSQJUISGPw 80x

Üniversite Bilgileri

Derece	Üniversite	Program	Giriş Tarihi	Mezuniyet Tarihi	Ortalama
Lisans	FIRAT ÜNİVERSİTESİ	FEN BİLGİSİ ÖĞRETMENLİĞİ	9/2005	6/2009	3.36/4

Derece	Üniversite	Program	Giriş Tarihi	Mezuniyet Tarihi	Ortalama
Y.Lisans	FIRAT ÜNİVERSİTESİ	FEN BİLGİSİ EĞİTİMİ A.B.D.	9/2009	5/2011	92,6

RSQJUISGPw 80x

İş Bilgileri

ÜNİVERSİTE	BÖLÜM	PROGRAM	ÜNVANI	BAŞLANGIÇ
HAKKÂRİ ÜNİVERSİTESİ	EĞİTİM BİLİMLERİ	ÖLÇME VE DEĞERLENDİRME A.B.D.	ARAŞTIRMA GÖREVLİSİ	2010-EYLÜL