T. C.

DİCLE ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

EĞİTİM BİLİMLERİ EĞİTİMİ ANABİLİM DALI

EĞİTİM PROGRAMLARI VE ÖĞRETİM BİLİM DALI

ORTAOKULLARDA İNGİLİZCE ÖĞRETİMİNDE ÖLÇME DEĞERLENDİRME SÜRECİNİN İNCELENMESİ

DOKTORA TEZİ

Berivan TOKUR ÜNER

Diyarbakır-2022

T. C.

DİCLE ÜNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ

EĞİTİM BİLİMLERİ EĞİTİMİ ANABİLİM DALI

EĞİTİM PROGRAMLARI VE ÖĞRETİM BİLİM DALI

ORTAOKULLARDA İNGİLİZCE ÖĞRETİMİNDE ÖLÇME DEĞERLENDİRME SÜRECİNİN İNCELENMESİ

HAZIRLAYAN

Berivan TOKUR ÜNER

TEZ DANIŞMANI

Prof. Dr. Bayram AŞILIOĞLU

Diyarbakır-2022

T.C

DİCLE UNİVERSİTESİ

EĞİTİM BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜ

DİYARBAKIR

Berivan TOKUR ÜNER tarafından yapılan "Ortaokullarda İngilizce Öğretiminde Ölçme Değerlendirme Sürecinin İncelenmesi" konulu bu çalışma, jürimiz tarafından Eğitim Bilimleri Eğitimi Anabilim Dalında DOKTORA tezi olarak kabul edilmiştir

Jüri Üyesinin

Ünvanı	Adı Soyadı
Başkan: Doç. Dr.	Hilal KAZU
Üye : Doç. Dr.	İbrahim YAŞAR KAZU
Üye : Prof. Dr.	Bayram AŞILIOĞLU
Üye : Prof. Dr.	Şengül KOCAMAN
Üye : Doç. Dr.	Taha YAZAR

Tez Savunma Sınavı Tarihi: 29/06/2022

Yukarıdaki bilgilerin doğruluğunu onaylarım.

Prof. Dr. Bahar BURTAN DOĞAN

ENSTİTÜ MÜDÜRÜ

(MÜHÜR)

BİLDİRİM

Tezimin içerdiği yenilik ve sonuçları başka bir yerden almadığımı ve bu tezi Dicle Üniversitesi Eğitim Bilimleri Enstitüsünden başka bir bilim kuruluşuna akademik gaye ve unvan almak amacıyla vermediğimi; tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada kullanılan her türlü kaynağa eksiksiz atıf yapıldığını, aksinin ortaya çıkması durumunda her türlü yasal sonucu kabul ettiğimi beyan ediyorum.

29/06/2022

Berivan TOKUR ÜNER

ÖNSÖZ

Akademik hayata adım attığımdan beri her zaman hoşgörüsü ve sabrıyla yanımda olan ve bu süreçte yetişmemde bana inanan ve desteğini esirgemeyen çok değerli hocam ve tez danışmanım Prof. Dr. Bayram AŞILIOĞLU'na sonsuz teşekkürlerimi ve en içten saygılarımı sunarım.

Çalışmanın veri toplama araçlarının geliştirilmesi ve pek çok aşamasında sunduğu değerli görüş ve katkılarından dolayı Prof. Dr. Behçet ORAL'a, Prof. Dr. Şengül KOCAMAN'a, Prof. Dr. Nilüfer BEKLEYEN'e, Doç. Dr. Taha YAZAR'a, Doç. Dr. İsmail Kınay'a, Doç. Dr. Osman SOLMAZ'a, Doç. Dr. Süleyman BAŞARAN'a, Doç. Dr. Mustafa İLHAN'A ve Arş. Gör. Serpil ÇELİKTEN'e en içten teşekkürlerimi sunarım.

Doktora tez çalışmamda sunduğu katkılarla destek olan arkadaşımların Dr. Filiz KOYUNCU ve Dr. Fikriye YAMAN'a ve veri toplama sürecinde yardımcı olan okul idarecilerine ve öğretmenlere teşekkür ederim. Ayrıca doktora öğrenimim boyunca 2211-A Genel Yurt İçi Doktora Bursu kapsamında sunduğu maddi ve manevi katkılarından dolayı Türkiye Bilimsel ve Teknolojik Araştırma Kurumu'na (TÜBİTAK) teşekkür ederim.

Eğitim hayatım boyunca her zaman fedakarlıkları ve bana olan inançlarıyla yanımda olan aileme, yaşadığım zorlukları kolaylaştıran eşime, desteklerinden dolayı arkadaşlarıma teşekkür ederim.

Son olarak, doktora sürecimin tamamlandığını gururla izlemesini istediğim canım babamı bu süreçte kaybetmenin derin üzüntüsünü yaşıyor ve onu sonsuz sevgi ve rahmetle anıyorum.

İÇİNDEKİLER

BİLDİRİM……..……**ii**
ÖNSÖZ ..... **iii**
İÇİNDEKİLER.....**iv**
ÖZET ..... **vii**
ABSTRACT .....**ix**
TABLOLAR LİSTESİ .....**xi**
ŞEKİLLER LİSTESİ ..... **xiii**
KISALTMALAR LİSTESİ .....**xiv**
1. GİRİŞ.....**1**
1.1 Problem Durumu.....1
1.2 Araştırmanın Amacı.....8
1.3 Araştırmanın Önemi 9
1.4 Sayıltılar.....11
1.5 Sınırlılıklar.....11
1.6 Tanımlar.....11
2. KURAMSAL ÇERÇEVE **12**
2.1 Ölçme ve Değerlendirme 12
2.1.1 Eğitimde Ölçme Değerlendirmenin Yeri ve Önemi.....15
2.2 Ölçme ve Değerlendirme Yaklaşımları 18
2.2.1 Geleneksel ve Alternatif Değerlendirme 18
2.2.2 Biçimlendirici ve Özetleyici Değerlendirme.....22
2.2.3 Formal ve İnformal Değerlendirme.....25
2.2.4 Doğrudan ve Dolaylı Değerlendirme 25
2.2.5. Nesnel ve Öznel Değerlendirme……26
2.2.6 Norma ve Ölçüte Bağlı Değerlendirme.....27
2.3 Yabancı Dil Öğretiminde Ölçme ve Değerlendirme 29
2.3.1 Bilimsel Olmayan Dönem 30
2.3.2 Psikometrik/Yapısalcı Dönem.....31
2.3.3 Bütünlükçü/Sosyo-Dilbilimsel Dönem 34
2.3.4 İletişimsel Yaklaşımlar Doğrultusunda Ölçme Değerlendirme.....36
2.4 Yabancı Dil Becerilerinin Değerlendirilmesi 39

2.4.1 Dinleme Becerilerinin Değerlendirilmesi 39
2.4.2 Konuşma Becerilerinin Değerlendirilmesi 46
2.4.3 Okuma Becerilerinin Değerlendirilmesi 55
2.4.4 Yazma Becerilerinin Değerlendirilmesi 62
2.5 İlgili Araştırmalar 67
2.5.1 Yurt İçinde Yapılan İlgili Çalışmalar 67
2.5.2 Yurt Dışında Yapılan İlgili Çalışmalar 78
3. YÖNTEM 88
3.1 Araştırmanın Modeli 88
3.2 Araştırmanın Deseni 88
3.3 Evren Örneklem 92
3.4 Veri Toplama Araçları 96
3.4.1 İngilizce Öğretiminde Ölçme Değerlendirme Sürecinin İncelenmesine İlişkin Görüşme Formu 96
3.4.2 Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketi 97
3.5 Verilerin Toplanması 100
3.5.1 Nitel Verilerin Toplanması 100
3.5.2 Nicel Verilerin Toplanması 101
3.6 Verilerin Analizi 101
3.6.1 Nitel Verilerin Analizi 102
3.6.2 Nicel Verilerin Analizi 106
4. BULGULAR 107
4.1 İngilizce Öğretmenlerinin Yabancı Dil Becerilerini Ölçme ve Değerlendirme Düzeylerine İlişkin Bulgular 119
5. TARTIŞMA 150
6. SONUÇ VE ÖNERİLER 178
6.1 Sonuçlar 178
6.2 Öneriler 180
KAYNAKÇA 182
EKLER 219
EK-1 Uzman Değerlendirme Formu-1 219
EK-2 Katılımcı Görüşme Formu 221
EK-3 Uzman Değerlendirme Formu-2 223

EK-4 Veri Toplama Aracı 227
EK-5 Etik Kurul İzni-1 231
EK-6 Uygulama İzni-2 232

ÖZET

Ortaokullarda İngilizce Öğretiminde Ölçme Değerlendirme Sürecinin İncelenmesi

Bu araştırmanın amacı ortaokullarda görev yapan İngilizce öğretmenlerinin ölçme değerlendirme sürecine ilişkin görüşlerini incelemektir. Karma araştırma yöntemi çerçevesinde yürütülen bu çalışma, betimsel tarama modelindedir. Araştırmanın nitel veri toplama sürecine dahil edilen katılımcıları, 2019-2020 eğitim öğretim yılı Diyarbakır ili merkez ilçelerine bağlı ortaokullarda görev yapan ve amaçlı örnekleme alma yöntemiyle belirlenen 45 öğretmenden oluşmaktadır. Araştırmanın nicel verileri ise, 2020-2021 eğitim öğretim yılında Diyarbakır ili merkez ilçelerinde görev yapan 543 İngilizce öğretmeninden tabakalı örnekleme yöntemi ile seçilen 306 öğretmenden elde edilmiştir. Araştırmada veri toplama araçları olarak araştırmacı tarafından geliştirilen yarı yapılandırılmış görüşme formu ve "Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketi" kullanılmıştır. Nitel verilerin analizinde içerik analiz yöntemi kullanılırken; nicel verilerin analizinde betimsel istatistiklerden yararlanılmıştır.

Araştırma sonuçlarına göre İngilizce öğretmenlerinin ölçme değerlendirme sürecine başlamadan önce "Hazırlık" aşamasında planlama yaptıkları ve testin geçerlik, güvenirlik ve kullanışlık özelliklerine çoğunlukla dikkat ettikleri belirlenmiştir. Bununla birlikte, öğretmenlerin çoğu hazırladıkları ölçme araçlarının İngilizce öğretim programıyla uyumlu olmasına kısmen dikkat ettiklerini ifade etmişlerdir. "Uygulama" ana teması altında ise, öğretmenlerin İngilizce dil becerilerinin ölçülmesinde klasik ölçme araçları arasında en fazla çoktan seçmeli testleri; tamamlayıcı ölçme değerlendirme araçları arasında ise en fazla proje ve performans görevleri ile ürün dosyasını (portfolyo) tercih ettikleri belirlenmiştir. Araştırmanın yabancı dil beceri alanlarının değerlendirilmesiyle ilgili nitel ve nicel bulguları incelendiğinde; öğretmenlerin dinleme ve konuşma alanlarında yeterli düzeyde değerlendirme yapamadıkları kaydedilmiştir. Diğer taraftan öğretmenlerin okuma becerilerinin değerlendirilmesine yönelik görüşleri çoğunlukla olumludur. Yazma becerilerinin ise ne dinleme ve konuşma alanları kadar ihmal edildiği ne de okuma becerileri kadar önemsendiği belirlenmiştir. Öğretmenler yabancı dil yapıları arasında en fazla sözcük bilgisinin değerlendirilmesine yönelik uygulamalar yapmaktadır. Ayrıca öğretmenlerin büyük bir çoğunluğunun duyuşsal becerilerin ölçülmesiyle ilgili olumlu görüş bildirdikleri kaydedilmiştir.

Araştırmanın “Değerlendirme” ana temasında öğretmenlerin tümünün bir tercih olmaktan ziyade zorunlu olduğu için düzey belirleyici değerlendirme yöntemini kullandıkları tespit edilmiştir. Araştırmanın “Geribildirim” ana temasına ilişkin bulgular incelendiğinde ise, çoğu öğretmenin öğrencilere ölçme sonuçları hakkında geribildirimde bulunmayı oldukça önemsediği belirlenmiştir. Ayrıca öğretmenler nitelikli bir değerlendirme süreci için, ölçme değerlendirme araç ve yöntemleri, öğrenme ortamı, öğretim programı ve öğretim materyallerinin geliştirilmesine yönelik çeşitli öneriler sunmuşlardır.

Anahtar Kelimeler: Ölçme, Değerlendirme, İngilizce Öğretimi, İngilizce Öğretmenleri

ABSTRACT

The Investigation of Testing and Assessment Process in English LanguageTeaching at Secondary School Level

The aim of this study is to identify teachers' opinions about testing and assessment process in English language teaching. The study which was conducted within the scope of mixed research method, is a descriptive survey. The participants included in qualitative data collection process consist of 45 teachers selected by purposive sampling method and working in different secondary schools in the central districts of Diyarbakır in 2019-2020 academic year. As for quantitative data collection, the sample of the study consists of 306 teachers determined by stratified sampling method method among 543 English language teachers working in the central districts of Diyarbakır in 2020-2021 academic year. A semi-structured interview form and "The Questionnaire of Testing and Assessing Foreign Language Skills" developed by the researcher were used to collect data. The qualitative data was analyzed through content analysis method while descriptive analysis techniques were used to analyse quantitative data.

The results of the study reveal that English language teachers make some arrangements to design a test and mostly pay attention to a test's basic principles regarding validity, reliability and practicality before beginning testing and assesment process at "Preparation" stage. Moreover, most of the teachers state that they are partly careful about designing measurement tests which are truly in accord with English language teaching curriculum. Considering the findings regarding "Implementation" main theme, it was determined that teachers usually prefer using multiple choice tests among traditional assessment techniques; and projects, performance tasks and porfolios among alternative assessment tasks. When it comes to assessing language skills, qualitative and quantitative data reveal that teachers can't assess listening and speaking skills at a satisfactory level. On the other hand, teachers' opinions about assesing reading skills are mostly positive. Writing skills are neither as ignored as listening and speaking skills, nor they are considered as important as reading skills. Most of the teachers make assessment practices to test learners' vocabulary knowledge. Also, it was revealed that a clear majority of teachers view measuring affective skills significant.

The results regarding "Assessment" main theme show that all of the teachers use summative assessment as it is stated compulsory in the curriculum to carry out this kind of

assessment. When it comes to findings regarding “Feedback” main theme, teachers mostly care about informing learners about their testing scores. In addition, improving testing instruments, learning environment, foreign language curriculum and teaching materials are suggestions provided by teachers to experience a more qualified assessment process.

Key Words: Testing, Assessment, English Language Teahing, English Language Teachers

TABLOLAR LİSTESİ

Tablo 1 Araştırmanın nitel boyutuna dahil edilen öğretmenlere ait özellikler.....93
Tablo 2 Araştırmanın nicel boyutuna dahil edilen öğretmenlere ait özellikler.....95
Tablo 3 Veri toplama aracındaki maddelerin uyuşma düzeyleri.....98
Tablo 4 Veri toplama aracında kullanılan değer aralıkları.....106
Tablo 5 "Bir İngilizce testi geliştirirken planlama yapıyor musunuz? Planlama yapıyorsanız nelere dikkat ettiğinizi açıklayabilir misiniz" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları.....107
Tablo 6 "Ölçme aracı geliştirirken veya bir ölçme aracını kullanmaya karar verirken geçerlik, güvenirlik ve kullanışlık özelliklerinin sağlanması için nelere dikkat ettiğinizi açıklar mısınız?" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları.....109
Tablo 7 "Hazırladığınız ölçme araçlarının İngilizce öğretim programının kazanımlarına uygun olması için nelere ve ne düzeyde dikkat ettiğinizi açıklar mısınız? sorusuna verilen cevaplardan elde edilen görüşler ve frekansları.....112
Tablo 8 "İngilizce öğretim programında öngörülen tüm dil beceri alanlarının (dinleme, konuşma, okuma ve yazma) değerlendirilmesi ile ilgili görüşleriniz nelerdir?" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları.....115
Tablo 9 Öğretmenlerin dinleme becerilerini ölçme ve değerlendirme süreçlerine ilişkin ortalamaları, standart sapmaları ve düzeyleri.....119
Tablo 10 Öğretmenlerin konuşma becerilerini ölçme ve değerlendirme süreçlerine ilişkin İlişkin ortalamaları, standart sapmaları ve düzeyleri.....120
Tablo 11 Öğretmenlerin okuma becerilerini ölçme ve değerlendirme süreçlerine ilişkin ortalamaları, standart sapmaları ve düzeyleri.....121
Tablo 12 Öğretmenlerin yazma becerilerini ölçme ve değerlendirme süreçlerine ilişkin ortalamaları, standart sapmaları ve düzeyleri.....122
Tablo 13 "Klasik ölçme araçlarından hangi/hangilerini kullanıyorsunuz? Kullanma nedenleriniz ile birlikte açıklar mısınız?" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları.....124

Tablo 14 “Tamamlayıcı ölçme değerlendirme araç ve tekniklerinden hangi/ hangilerini kullanıyorsunuz? Kullanma nedenleriniz ile birlikte açıklar mısınız?” sorusuna verilen cevaplardan elde edilen görüşler ve frekansları 127
Tablo 15 Öğretmenlerin yabancı dil beceri alanlarına göre tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini kullanma ortalamaları, standart sapmaları ve düzeyleri 130
Tablo 16 “İngilizce öğretiminde dilbilgisi, sözcük dağarcığı ve telaffuz gibi yapıların ölçülmesi ile ilgili görüşleriniz nelerdir?” sorusuna verilen cevaplardan elde edilen görüşler ve frekansları 131
Tablo 17 Öğretmenlerin dilbilgisi, sözcük yapıları ve telaffuzla ilgili davranışları değerlendirme ortalamaları, standart sapmaları ve düzeyleri 135
Tablo 18 “İngilizce öğretiminde duyuşsal becerilerin ölçülmesi ile ilgili görüşleriniz nelerdir?” sorusuna verilen cevaplardan elde edilen görüşler ve frekansları 136
Tablo 19 “İngilizce öğretiminde tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirmeyi hangi düzeyde ve niçin kullandığınızı açıklar mısınız?” sorusuna verilen cevaplardan elde edilen görüşler ve frekansları 139
Tablo 20 Öğretmenlerin Yabancı Dil Beceri Alanlarına Göre Tanılayıcı, Biçimlendirici ve Düzey Belirleyici Değerlendirme Tekniklerini Kullanma Ortalamaları, Standart Sapmaları ve Düzeyleri 143
Tablo 21 “Ölçme sonuçları hakkında öğrencilere geribildirimde bulunma konusundaki görüşleriniz nelerdir?” sorusuna verilen cevaplardan elde edilen görüşler ve frekansları 143
Tablo 22 “İngilizce derslerinde nitelikli bir ölçme değerlendirme süreci için önerileriniz nelerdir? Açıklayınız” sorusuna verilen cevaplardan elde edilen görüşler ve frekansları 146

ŞEKİLLER LİSTESİ

Şekil 1 Ölçme, değerlendirme ve öğretim arasındaki ilişkiyi gösteren Venn şeması 15
Şekil 2 Keşfedici sıralı desen akış şeması 91
Şekil 3 Anket Geliştirme Süreci 97

KISALTMALAR LİSTESİ

IELTS: International English Language Testing System (Uluslararası İngilizce Dil Sınavı Sistemi)

MEB: Milli Eğitim Bakanlığı

OECD: Organisation for Economic Co-operation and Development (Avrupa Ekonomik İşbirliği Örgütü)

1. GİRİŞ

1.1. Problem Durumu

Eğitim pek çok öğeyi içinde barındıran ve bu öğelerin karşılıklı etkileşimi ile gerçekleşen karmaşık bir süreçtir. Bu süreç öğretimin planlı olarak yürütüldüğü formal uygulamaların yanı sıra plansız ve kendiliğinden gerçekleşen informal öğrenmeleri kapsamaktadır. Eğitimin formal yapısı incelendiğinde; öğretmenler, öğrenciler, öğrenme konuları, öğretim materyalleri ve yöntemleri, programlar, araç gereçler gibi pek çok nitel ve nicel değişkenin planlı ve amaçlı bir şekilde bir araya gelerek oluşturduğu bir yapıdan söz edilebilir. Söz konusu bu öğeler önceden belirlenen hedefleri kazandırmak üzere bir bütün olarak eğitim sistemini oluşturmaktadırlar. Bu sistem içerisinde yer alan öğelerin birbirleriyle uyum içinde işlemesinin, eğitimin kazandırmak istediği davranışları gerçekleştirmek için oldukça önemli olduğu söylenebilir. Bu anlamda eğitim kurumlarında yürütülen faaliyetlerin nasıl işlediğini belirlemenin eğitim kalitesini arttırmak için fayda sağlayacağı düşünülebilir.

Alanyazında eğitim kurumlarının nasıl işlediğiyle ilgili farklı bakış açıları sunan çeşitli örgütsel yaklaşımlar sınıflandırılmıştır. Aydın'a (2010) göre bu yaklaşımlar tarihsel süreç dikkate alınarak incelendiğinde şu üç temel kuramdan bahsedilebilir: Klasik, neoklasik ve sistem yaklaşımı. Klasik kurama göre formal örgütler, verimliliğini arttırabilmek için tüm çabalarını biçimsel yapının düzenlenmesi ve adeta bir makine gibi işlemesi üzerine kurgulamışlardır (Çetin, 2003). Diğer taraftan neoklasik kuramda insan temele alınmış ve eğitim kurumları gibi örgütlerin de içinde bulunduğu formal örgütlerin her şeyden önce bir sosyal yapı olduğu vurgulanmıştır (Topaloğlu, 2011). Çağdaş bir yaklaşım olarak kabul edilen sistem yaklaşımı ise bireylere organizasyonun bütün yönlerini inceleme fırsatı sunan, verilen bir dizi kararın diğer öğeleri nasıl etkilediğini açıklayan ve problemlerin çözümüne yönelik eldeki bütün kaynakların en ideal düzeyde kullanımını savunan bir yönetim aracıdır (Gupta & Gupta, 2013). Klasik yaklaşımda biçimsel çatı, neoklasik yaklaşımda ise beşeri ilişkilerin ön plana çıkmasına karşın, sistem yaklaşımı diğer kuramlardan daha kapsamlı ve çok yönlü çözümlemelere dayanmaktadır (Aydın, 2010). Bu geniş çerçeveden bakıldığında, Özerbaş (2015) eğitimi sistem yaklaşımıyla incelemenin, eğitim sisteminin temelini oluşturan öğelerin nasıl işlediğini kavrama ve sistemin parçaları arasındaki ilişkileri ortaya koymak için daha net ve bütünsel bir kavrayış sunduğunu ifade etmiştir.

Eğitim bir sistem olarak incelendiğinde, sistemin girdi, süreç, çıktı ve kontrol olmak üzere dört temel öğeden oluştuğu kabul edilmektedir (Odabaşı, 1997). Araç ve gereçler, para, eğitime alınacak bireylerin sayısı ve özellikleri, kazandırılması planlanan davranışların tasarımı, eğitimle ilgili yasalar, tüzükler, yönetmelikler, toplumumuzun örf ve adetleri, öğretmen ve yöneticilerin özellikleri, kullanılan eğitim yöntemleri, insangücü planlaması, eğitim felsefesi ve diğer pek çok eğitim faaliyetlerini etkileyen, bazen davranışların değişmesine yardımcı olan, kolaylaştırıcı; bazen de değişmeye direnmede rol oynayan zorlaştırıcı öğeler eğitim sisteminin girdileri arasındadır (Baykul, 1992). Eğitimin ikinci öğesi olarak belirtilen, öğrencilere istendik ve planlanan davranışların kazandırılması için yürütülen tüm öğrenme-öğretme etkinlikleri sistemin sürecini oluşturmaktadır (Tan, 2013). Sistemin üçüncü öğesi olan çıktı, eğitim süreci sonunda bireyin davranışlarında meydana gelen değişikliklerdir. Dördüncü öğesi olan kontrol öğesi ise sistemin işleyip işlemediğini, işlemiyorsa neden işlemediğini ya da hangi açılardan işlemediğini ve bunun nedenlerini belirlemek amacıyla yürütülen bir değerlendirme sürecini kapsamaktadır (Aydın, 2010).

Eğitim sürecinin sonunda bireylerin istendik davranışları beklenilen düzeyde gerçekleştirip gerçekleştirmediğinin sınanması oldukça önemlidir. Bu sınama durumları sonucunda hedeflerin bazıları istenilen düzeyde kazanılmış; bazıları hiç kazanılmamış ya da bazı beklenmedik yeni davranışlar ortaya çıkmış olabilir. Bu durumlar eğitimde bir kontrol mekanizmasının gerekliliğine işaret etmektedir. Söz konusu programın istendik ürünler çıkarma anlamında yüksek bir performans sergileyeceğine dair beklentiler davranışların kazanıldığını garantilememektedir. Bu durumda programın yeterlilik ve amaca hizmet etme derecesinin sürekli araştırılması gerekmektedir. Ertürk'ün (2013: 116) de belirttiği gibi "eğitime kendi kendini onarıcı olanağını veren değerlendirme" eğitim faaliyetlerinin belirtilen maksada hizmet edip etmediğinin veya olumsuz olarak bakıldığında istenmedik sonuçlara sebep olup olmadığının araştırılması sürecidir.

Eğitimde öğrenciler tarafından edinilen kazanımların nitelik ve niceliği, başarının ölçülmesi ve öğrenme sürecinde yaşanılan güçlüklerin belirlenmesi için ölçme değerlendirme faaliyetlerine geniş ölçekte yer verilmektedir (Altmışdört, 2010). Özçelik'e (2010) göre, ilk bakışta "Beklenen davranış değişiklikleri gerçekleşmiş midir?" sorusuna kolayca cevap verilebileceği düşünülebilir. Ancak, eğitimle kazandırılmak istenilen davranış değişiklikleri niteliksel olarak ayrıntılı bir şekilde incelendiğinde bu sorunun

sanıldığı kadar kolay bir yanıtının olmadığı ortaya çıkar. Çünkü eğitimde istenilen davranış değişikliklerinin bir kısmı doğrudan gözlenebilirken (devinişsel davranışlar), bir kısmı ise doğrudan gözlenemeyen (bilişsel ve duyuşsal davranışlar) davranışlardan oluşur. Bu durumda istendik davranışların tespit edilmesi ancak doğru ölçümler ve amacına uygun şekilde hazırlanan ölçeklerle yapılabilir. Son aşamada da beklenen ve gerçekleşen kazanımlar dikkate alınarak bir karar verme yani değerlendirme işlemi gerçekleştirilir (Özçelik, 2010).

Ölçme ve değerlendirme öğretim sürecinin vazgeçilmez bir parçasıdır. Tekin'e (1982) göre, öğrenciye kazandırılması hedeflenen davranış değişikliklerinin ne derece gerçekleştiğine dair geçerli ve güvenilir bir yanıt bulunmaksızın sürdürülen eğitim çabaları oldukça yetersiz kalacaktır. Öyle ki, ölçme değerlendirme olmadan bir öğretim sürecini düşünmek bile mümkün değildir. Yaşar (2011), eğitimde ölçme ve değerlendirmenin önemi ve gerekliliğinin eğitim uygulamalarında verilecek olan kararlara güçlü bir dayanak olma noktasında kendini gösterdiğini ifade etmektedir. Ona göre bu alanda gün geçtikçe artan geçerli ve güvenilir ölçme araçları geliştirerek uygulamaya koyma çabaları, ölçme ve değerlendirmeye olan ilgiyi de arttırmıştır. Çoğu eğitim uzmanı da eğitim kalitesinin geliştirilmesi için ölçme ve değerlendirmenin etkili ve önemli bir dinamik olduğu noktasında uzlaşmaktadırlar (Crooks, 1988; Black & Wiliam, 1998; Brown, 2004; Bachman, Palmer & Palmer, 2010).

Ölçme ve değerlendirmenin eğitim sürecine hizmet eden pek çok işlevi bulunmaktadır. En genel anlamda bu işlevler; 1) öğrencileri tanıma, hazırbulunuşluk düzeylerini belirleme ve uygun bir eğitim programına yerleştirme (Güler, 2011); 2) öğretimin ve öğretim materyallerinin işlevselliklerini belirleme ve öğretmenlere kendi öğretim faaliyetlerinin etkililiğini sorgulayabilmeleri için araç olma (Çelikkaya, 2008); 3) öğrenci kazanımlarını teşhis etme, onlara kendi öğrenme düzeyleri ile güçlü ve zayıf yönlerini kontrol etme fırsatı sunma (Kelleghan & Greaney, 2001) ve 4) eğitim problemlerinin tespiti ile çözümleri araştırma ve program geliştirme sürecine katkı sağlama (Yılmaz, 2002) olarak özetlenebilir.

Ölçme ve değerlendirme süreci yabancı dil öğretiminin de ayrılmaz bir parçasıdır. Yabancı dil öğrenimi ve öğretiminin istenilen şekilde işleyip işlemediğine dair sunduğu geribildirimlerle bu sürecin verimli bir şekilde yürütülmesine oldukça önemli katkılar sunmaktadır. Doğru ve etkili bir yabancı dil öğretiminde öğrenci, eğitim yaşantıları,

materyaller ve öğretmenler (Lado, 1964) ile çeşitli bilişsel ve duyuşsal faktörler (Krashen, 1981) gibi pek çok değişkenin rol oynamasının yanı sıra ölçme ve değerlendirme faaliyetlerinin de bu süreçte önemli bir yeri olduğunu söylemek mümkündür (Brown, 2004).

Yabancı dil öğretiminin temel amacı; bireylere kendi anadilleri dışındaki dil ya da dillerin, dinleme, konuşma, okuma ve yazma gibi dört temel beceriyi ve telaffuz, sözcük dağarcığı ve dilbilgisi gibi yapıları kazandıracak şekilde öğretmektir (Ergüder, 2005). Söz konusu bilgi ve becerileri bireylere daha kısa sürede ve daha etkili bir şekilde kazandırmak için geçmişten günümüze farklı yöntem ve teknikler kullanılagelmiştir. Uygulanan yöntemlerin farklılığı var olan ölçme değerlendirme yöntemlerinin de değişimini zorunlu hale getirmiştir. Şahin'e (2018) göre çağdaş yabancı dil öğretiminde bütünsel bir kavrayış sunan, dört temel dil becerisine aynı ölçüde önem veren yeni yöntem ve yaklaşımların geliştirilmesiyle, ölçme ve değerlendirme araçlarının niteliğinde de önemli değişiklikler gözlenmektedir. Bu durum, Yavuz-Kırık'ın (2008) dil eğitimi sürecinin unsurlarından olan İngilizce öğretimi yaklaşım ve uygulamalarının ölçme ve değerlendirme uygulamalarıyla ne denli iç içe ve birbiriyle karşılıklı etkileşim içindeki bağımlı kavramlar olduğu yönündeki tezini doğrular niteliktedir.

Görüldüğü üzere, yabancı dil öğretiminde işe koşulan ölçme değerlendirme araçlarının öğretim yöntem ve uygulamalarına göre şekillendiğini söylemek mümkündür. Bu bağlamda tarihsel süreç içinde yabancı dil öğretim yaklaşımlarına dair yaşanan gelişmelerin ölçme ve değerlendirme faaliyetlerini etkilediği öne sürülebilir. Örneğin, davranışçılık yaklaşımının etkili olduğu dönemlerde dil öğretiminde Dilbilgisi-Çeviri Yöntemi (Grammar-Translation Method) ya da daha önceleri bilenen adıyla Geleneksel Yöntem (Traditional Method) kullanılmaktaydı (Mitchell & Vidal, 2001). Dilin mantıksal kurallar üzerine kurulduğu ve hedef dil ile anadil arasında çeviri yapabilmenin dilin öğrenildiğine dair en büyük kanıt olması bu yöntemin temel sayıltılarındandır (Stern, 1983). Dolayısıyla bu öğretim yönteminin gerektirdiği ölçme ve değerlendirme araçları öğretmenler tarafından hazırlanan ve öncelikli olarak öğrencilerin dilin kurallarını ne kadar öğrendiğini ölçen ve değerlendiren çeviri odaklı sınavlardı (Şenel, 2015). Diğer taraftan 1950'lerde davranışçılık ve karşılaştırmalı analiz yaklaşımlarının dikkat çektiği dönemlerde ana dil ile hedef dil arasındaki dilbilimsel, sesbilimsel ve sözcüksel karşılaştırmalar gibi dilin özel öğelerine odaklanan ölçme teknikleri kullanılırken, 1970 ve

1980'lerde bütünün parçalarından daha önemli olduğu görüşüyle iletişimsel dil teorilerinin daha etkileşimli test tekniklerini ön plana çıkardığı görülmektedir (Brown, 2004). Bugüne geldiğimizde ise yaşama daha yakın, güncel, geçerli ve öğrenenlerin ihtiyaçlarına yanıt olabilecek ölçme değerlendirme araçlarının kullanılması gerektiğine dikkat çekildiği söylenebilir.

Ülkemizdeki yabancı dil öğretim faaliyetleri incelendiğinde, evrensel boyutta geçerliği kabul edilen eğitim felsefelerinin öğretim programlarına yansıdığı söylenebilir. Nitekim MEB tarafından 2004 yılında geliştirilen ve 2005-2006 eğitim öğretim yılıyla uygulanmaya konulan eğitim programlarında, yapılandırmacı yaklaşımın benimsendiği ve ölçme değerlendirme uygulamalarının da bu yaklaşımın temel ilkeleri çerçevesinde tasarlandığı açıklanmıştır. Yayınlanan programda değerlendirme kavramının uygulama alanı genişletilerek sadece ürünün değil, öğrenme süreçlerinin de değerlendirildiği özellikle vurgulanmıştır. Ayrıca öğretmenlere halihazırda kullandıkları geleneksel ölçme araçlarının yanında; performans, portfolyo değerlendirme ve öğrenci ürün dosyasının hazırlanması ile öğrencilerin duyuşsal ve kişisel gelişimleri hakkında bilgi edinmek için ölçekler (gözlem, görüşme) kullanmaları da önerilmiştir (MEB, 2005). 2018 İngilizce dersi öğretim programı incelendiğinde ise ölçme ve değerlendirme süreçlerinin teorik temelinin asıl olarak içinde farklı ölçme ve değerlendirme teknikleri barındıran CEFR (Avrupa Dilleri Ortak Çerçeve Programı) uygulamalarına dayandığı belirtilmektedir. Ayrıca bu uygulamaların çoğunun alternatif ve süreç temelli değerlendirmeyi kapsadığı da belirtilmiştir. Buna ek olarak, öğrencilerin kendi becerilerinin gelişim adımlarını izleyebildikleri öz-değerlendirme sürecinin destekleyici olduğu vurgulanmıştır (MEB, 2018).

Alanyazında ölçme ve değerlendirmeyle ilgili çalışmalar incelendiğinde, araştırmacıların farklı bağlamlardan yola çıkarak analiz yaptıkları söylenebilir. Örneğin, Tarhan (1992) tarafından yapılan tez çalışmasında, öğretim sürecinde kullanılan testlerin sıklık derecelerinin öğrencilerin başarı durumlarına etkisi incelenmiştir. Bilkent Üniversitesi Yabancı Diller Bölümünde gerçekleştirilen bu deneysel çalışma, ön-test ve son-testlerin sıklıkla uygulanmasının öğrenci başarıları üzerinde belirgin bir etki yaratmadığını otaya koyarken; belli aralıklarla yapılan testlerin öğrenci başarı puanlarını arttırdığı gözlemlenmiştir.

Araştırmacıların yöneldiği alanlardan biri yabancı dil öğretiminde kullanılan ölçme ve değerlendirme araçlarına yönelik tutumları incelemek olmuştur. Bu amaçla Yavuz-Kırık (2008) tarafından gerçekleştirilen tez çalışmasında, yabancı dil olarak İngilizce öğretmenlerinin ölçme ve değerlendirme bağlamında tutum ve yaklaşımları değerlendirilmiştir. Anadolu liselerinde yapılan bu araştırmada öğretmenlerden iletişimsel dil öğretim yöntemi kapsamında sunulan ölçme ve değerlendirme tekniklerini kullanma ve sorunlarla karşılaşma sıklıkları analiz edilmiştir. Sonuç olarak ise alternatif ölçme ve değerlendirme yaklaşımlarının kullanılması önerilmiştir.

Yabancı dil becerilerinin test edilmesi ve değerlendirilmesine yönelik gerçekleştirilen uygulamalar alanyazındaki tezlerin ve makalelerin bir diğer konusunu oluşturmaktadır. Bu beceri alanlarından özellikle konuşma becerisinin test edilmesi sürecine odaklanan Güllüoğlu (2004), Gazi Üniversitesi İngilizce hazırlık okulunda konuşma becerilerine ne kadar yer verildiğini inceleyip bu alanda yapılan sınavları ele almıştır. Araştırmanın sonuçlarına göre dilbilgisi, okuma ve kelime becerilerinin ölçülmesi sürecinin daha fazla önemsendiği ve az da olsa yazma ve dinleme becerilerine yönelik ölçme ve değerlendirme etkinliklerinin gerçekleştiği görülmektedir. Diğer taraftan konuşma becerilerinin değerlendirilmesine ilişkin herhangi bir uygulamanın olmadığı da kaydedilmiştir.

Araştırma sonuçları göz önünde bulundurulduğunda, İngilizce derslerinde uygulanan ölçme değerlendirme faaliyetlerinin öğretim hedeflerini istenilen düzeyde gerçekleştirdiğini söylemek güçtür. Nitekim ülkemizde ilkokul 2. sınıftan başlanarak üniversiteye kadar örgün eğitimin her kademesinde yabancı dil olarak İngilizce öğretimi yapılmakta olup, beklenilen verimin sağlanamaması araştırma sonuçlarını doğrular niteliktedir. Altan (2017) bu durumun en önemli sebeplerinden birinin, öğretim yöntemiyle neredeyse hiç ilgisi olmayan bir şekilde bilgileri anlamaya ve hatırlamaya yönelik ölçme araçlarıyla ölçüm yapılması olduğunu belirtmektedir. Ona göre bu şekilde alan içerikli ve tamamen bilgi odaklı testlerle yapılan ölçümler bireysel farklılıkları göz ardı etmekte ve öğrenenlerin kazanmaları gereken becerileri gerçekleştirememelerine sebep olmaktadır. Özmen (2012) bir öğretim programının üst boyutlarından birini oluşturan ölçme ve değerlendirmenin kesinlikle öğretim programının doğasıyla örtüşmesi gerektiğini, aksi bir durumun ise öğrenilen yabancı dilin bütün dilbilimsel ve pedagojik alt yapısı ile kazanımlarını riske sokabileceğini belirtmektedir.

21. yüzyılda eğitim ihtiyaçları ve öğrenme ve öğretim yaklaşımları üzerine değişen anlayışın, öğrenciler ve belki de onlar için çoğu zaman kaygı dolu bir deneyimi hatırlatan değerlendirmeyi bir araya getirdiği görülmektedir. Özellikle değerlendirmenin öğrenme için yapılması gerektiğiyle ilgili görüşlerin ön plana çıkmasıyla birlikte değerlendirme kavramına atfedilen önemin daha kapsayıcı bir nitelik kazandığı söylenebilir (Brown, 2005). Öz'e (2014) göre geleneksel eylem tarzında meydana gelen bu paradigma değişikliği, değerlendirmenin öğrenme üzerindeki desteğini ön plana çıkarmış ve bu iki ayrı kavramı birbirinden ayrılamaz bir bütünlüğe ulaştırmıştır. Ancak eğitim öğretim ortamında değerlendirmelerin öğrenme sürecini kolaylaştırmaktan ziyade daha çok not vermek için yapıldığı ve öğrenenlerinin ne yapabildiklerinden çok öğrenme çıktılarına odaklanıldığı görülmektedir (Kuyumcu-Vardar ve Acar, 2018). Nitekim bu alanda yapılan çalışmalar, öğretim programında farklı ölçme değerlendirme uygulamaları önerilmesine rağmen öğretmenlerin hala geleneksel kalem kağıt testlerini kullandıklarını ve öz değerlendirme, portfolyo değerlendirme gibi uygulamaların ihmal edildiğini göstermektedir (Paker, 2006; Kırkgöz ve Ağçam, 2012; Öz, 2014; Tatlı, 2014). Ayrıca hazırlanan testlerin içeriğinin dilbilgisi ve sözcük yapılarına dayandığı ve çoğu zaman dinleme ve konuşma becerilerinin değerlendirilmesine yönelik uygulamaların yapılmadığı görülmektedir (Çakır ve Genç, 2021). Bu durum hedeflenen davranışlar ile uygulamalar arasındaki farklılıkların sonuçlarını gösterir niteliktedir.

Türkiye'de yabancı dil öğretiminde ölçme değerlendirme alanında yaşanan farklılıklar bu araştırmanın temel çıkış noktası olmuştur. Yabancı dil becerilerinin ne kadarının test edildiği ve programın benimsediği eğitim felsefisine uygun ölçme araçlarıyla doğru ölçümler yapılıp yapılmadığıyla ilgili sorulara uygun yanıtlar bulmak için, konunun bütün boyutlarıyla araştırılması gerektiği söylenebilir. Ayrıca eğitim programında benimsenen yapılandırmacı yaklaşım ve İletişimsel Dil Öğretim Yönteminin ilkelerine uygun olarak geliştirilen ölçme değerlendirme araçlarının uygulayıcılar tarafından nasıl algılandığını ortaya koymak önemlidir. Çünkü öğretmenlerin yabancı dil programında benimsenen anlayışa uygun bilgi, beceri ve tutumlara sahip olması istenilen kazanımların gerçekleşmesi için oldukça gereklidir. Bu noktada ölçme değerlendirme sürecinde işlemeyen, yanlış işleyen ya da zorlayıcı olarak algılanan unsurların tespit edilmesinin ölçme değerlendirme alanında yaşanan uyuşmazlıkların nedenlerine ilişkin bilgi sunacağı söylenebilir.

Araştırmanın Amacı

Bu araştırmanın amacı ortaokullarda görev yapan İngilizce öğretmenlerinin ölçme değerlendirme sürecine ilişkin görüşlerini incelemektir. Bu genel amaç doğrultusunda aşağıdaki sorulara cevap aranmıştır:

1. İngilizce öğretmenlerinin bir ölçme aracı hazırlarken planlama yapılmasıyla ilgili görüşleri nelerdir?
2. İngilizce öğretmenlerinin ölçme aracı geliştirirken veya bir ölçme aracını kullanmaya karar verirken geçerlik, güvenirlik ve kullanışlık özelliklerinin sağlanmış olması hakkındaki görüşleri nelerdir?
3. İngilizce öğretmenleri hazırladıkları ölçme araçlarının İngilizce öğretim programının kazanımlarına uygun olması için nelere dikkat etmektedirler?
4. İngilizce öğretmenlerinin tüm dil beceri alanlarının (dinleme, konuşma, okuma ve yazma) değerlendirilmesine ilişkin görüşleri nelerdir?
5. İngilizce öğretmenlerinin tüm dil beceri alanlarındaki (dinleme, konuşma, okuma ve yazma) ölçme ve değerlendirme uygulamalarını gerçekleştirme düzeyleri nelerdir?
6. İngilizce öğretmenlerinin İngilizce öğretiminde kullanılan klasik ölçme değerlendirme araçlarıyla ilgili görüşleri nelerdir?
7. İngilizce öğretmenlerinin İngilizce öğretiminde kullanılan tamamlayıcı değerlendirme araç ve teknikleriyle ilgili görüşleri nelerdir?
8. İngilizce öğretmenlerinin dilbilgisi, sözcük dağarcığı ve telaffuz gibi dil yapılarının değerlendirilmesine ilişkin görüşleri nelerdir?
9. İngilizce öğretmenlerinin duyuşsal becerilerin ölçülmesi ile ilgili görüşleri nelerdir?
10. İngilizce öğretmenlerinin tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirmeye ilişin görüşleri nelerdir?
11. İngilizce öğretmenlerinin ölçme sonuçları hakkında öğrencilere geri bildirimde bulunma konusundaki görüşleri nelerdir?
12. İngilizce öğretmenlerinin nitelikli bir ölçme değerlendirme süreci için önerileri nelerdir?

1.2. Araştırmanın Önemi

Eğitim ve öğretim hedeflerinin ne düzeyde gerçekleştirildiğiyle ilgili önemli geribildirimler sunan ölçme ve değerlendirme; öğretmenler, öğrenciler, veliler, program geliştirme uzmanları gibi pek çok paydaş için oldukça önemli bir yere sahiptir. Yabancı dil öğretimi açısından incelendiğinde ise, ölçme ve değerlendirme öğrencilerin hedef dildeki performanslarını ortaya çıkaran ve elde ettiği bu bilgiyle öğrenme hakkında bir karara ulaşılmasını sağlayan bir mekanizmadır. Öğrenciler hakkında doğru kararlara ulaşmak, onların ihtiyaçlarına yanıt bulabilmek ve öğretim etkinliklerini düzenlemek gibi pek çok eylem etkili bir değerlendirme sürecinin işe koşulmasını gerektirmektedir. Malone'a (2013) göre öğretim ve değerlendirme birbirini etkileyen kavramlar olmakla birlikte birbirlerini geliştirmektedirler. Bu iki kavram arasında önemli bir köprü kuran ise öğretmenlerdir. Değerlendirmenin gücünün öğretmenlerin bilgi ve uygulamalarına bağlı olduğu söylenebilir.

Türkiye'de yabancı dil öğretmenleri çeşitli standart testler, yazılı yoklamalar ve diğer ölçme araçlarıyla öğrencilerin beklenen davranışları ne düzeyde kazandıklarına yönelik bir yargıya ulaşmaktadırlar. Öğretmenlerden beklenen, öğretim programıyla uyumlu olacak şekilde ölçme araçları geliştirmeleri, puanlamaları ve sonuçlarla ilgili geribildirimde bulunmalarıdır. Bu yönüyle değerlendirme basit bir işlemmiş gibi görünse de aslında içerisinde pek çok öğeyi barındıran oldukça karmaşık bir süreçtir. Bu araştırmayla yabancı dil öğretiminde ölçme değerlendirmeye geniş bir perspektiften bakılarak, İngilizce öğretmenlerinin bu görevi nasıl yerine getirdikleri, zorlandıkları alanlar, karşılaştıkları sorunlar, nedenleri ve çözüm önerileri gibi konular tartışılmıştır. Ortaokullarda İngilizce öğretiminde ölçme ve değerlendirme sürecinin incelenmesi konusunda yapılan bu araştırmanın, eğitim sürecinin mevcut durumu hakkında geçerli ve güvenilir bilgiler sunarak uygulayıcılara ve program geliştirme uzmanlarına katkı sağlayacağı düşünülmektedir.

İngilizce öğretiminde ölçme değerlendirme sürecine ilişkin birçok bilimsel çalışmanın yürütüldüğü görülmektedir (Jannati, 2012; Ayas, 2014; Çiçen, 2014; Gökçe, 2014; Kayacan ve Eskicumalı, 2014; Öz, 2014; Wewer, 2014; Bayram, 2015; Britton, 2015; Chan, 2016; Kırkgöz, Babanoğlu ve Ağçam, 2017; Mede ve Atay, 2017; Sellan, 2017; Tsagari & Vogt, 2017; Ölmezer-Öztürk, 2018; Saeed, Tahir & Latif, 2018; Önalan ve Gürsoy, 2020; Ballıdağ ve İnan-Karagül; 2021; Çakır ve Genç, 2021; Tokur-Üner ve

Aşılıoğlu, 2022). Söz konusu çalışmalarda İngilizce öğretmenlerinin ölçme değerlendirme sürecine yönelik algıları, hangi ölçme ve değerlendirme yöntemlerini tercih ettikleri, test tekniklerini ne sıklıkla kullandıkları ve ölçme araçlarının uygulanması sürecinde yaşadıkları güçlükler kaydedilmiştir. Yapılan çalışmalar incelendiğinde, araştırmaların ölçme değerlendirmenin belirli alanlarına odaklanılarak gerçekleştirildiği söylenebilir. Ancak ölçme değerlendirmenin içerisinde pek çok öğe barındırmasından dolayı oldukça kapsayıcı bir alan olması, bu mekanizmanın detaylı bir şekilde incelenmesini gerektirmektedir. Dolayısıyla bu araştırmada ölçme değerlendirme süreciyle ilgili kuramsal bilgiler incelendikten sonra "Hazırlık", Uygulama", "Değerlendirme" ve "Geribildirim" olmak üzere ana bölümler belirlenmiş ve her bölümle ilgili aydınlatıcı verilere ulaşmak için araştırma soruları özenle hazırlanmıştır. Böylece ölçme değerlendirme sistemine dahil olan tüm boyutların ayrıntılı bir şekilde incelenmesi hedeflenmiştir. Örneğin, öğretmenlerin planlama yapıp yapmadıkları, öğretim programına bağlılıkları, tüm yabancı dil becerileri ve dil yapılarının değerlendirilmesi, duyuşsal becerilerin ölçülmesi, öğrencilere geribildirim sunulması gibi alan taramalarında daha az rastlanılan ama ölçme değerlendirmenin önemli boyutlarından olan bu konuların tümü araştırmaya dahil edilmiştir. Ayrıca bu çalışma karma araştırma yöntemi ilkeleri çerçevesinde sistematik bir şekilde yürütülmüştür. Böylece nitel veriler ışığında araştırmacı tarafından geliştirilen bir anketle, yabancı dilin temel yapıtaşlarından olan dinleme, konuşma, okuma ve yazma becerilerinin değerlendirilmesi sürecinin öğretmenler tarafından nasıl gerçekleştirildiği geniş ölçekte analiz edilmiştir.

Teorik olarak programlarda yer bulan yaklaşım ve tekniklerin uygulamadaki ölçme ve değerlendirme sürecine nasıl yansıdığını incelemek, teori ve uygulama arasındaki farklılıkları ortaya çıkarmak adına önemlidir. Bu noktada Diyarbakır ili ölçeğinde böylesi bir çalışmanın yürütülmesinin bu alanda önemli bir boşluğu dolduracağı düşünülmektedir.

1.3. Sayıltılar

1. Araştırmaya katılan öğretmenler görüşme sorularına ve "Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketi"nde yer alan davranışlara gönüllülük ilkesi çerçevesinde içtenlikle yanıt vermişlerdir.

2. Araştırmaya katılan öğretmenler görüşme sorularını ve "Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketi"nde yer alan maddeleri tam ve doğru anlamışlardır.

1.5. Sınırlılıklar

1. Bu araştırma 2019-2020 ve 2020-2021 eğitim öğretim yılları arasında Diyarbakır ili merkez ilçelerine bağlı ortaokullarda görev yapan İngilizce öğretmenleri ile sınırlıdır.

2. Bu araştırma görüşme soruları ve "Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketi"nde yer alan ifadeler ile sınırlıdır.

1.6. Tanımlar

Ölçme: Bir niteliğin gözlenerek gözlem sonuçlarının sayı ya da başka sembollerle gösterilmesidir (Demirel, 2012).

Değerlendirme: Bireylerin yetenekleri ya da performanslarıyla ilgili elde edilen ölçme sonuçlarının kendi bağlamına ait ölçütlerle karşılaştırılarak bir değer yargısına ulaşma sürecidir (Yılmaz, 2002).

2. KURAMSAL ÇERÇEVE

Eğitim ve öğretim hedeflerine ne düzeyde ulaşıldığıyla ilgili önemli veriler sunan ölçme değerlendirme süreci, eğitim sistemindeki öğretmenler, öğrenciler ve program hazırlayıcılar gibi tüm paydaşlar için vazgeçilmez bir öğedir. Etkili bir ölçme değerlendirme yapmak için ise öncelikle bu süreçle ilgili kavramları doğru bir şekilde tanımlamanın oldukça önemli olduğu söylenebilir. Dolayısıyla araştırmanın bu bölümünde ilk olarak ölçme ve değerlendirme sürecine dair kuramsal temeller ilgili araştırmalar ışığında sunulmuştur.

2.1. Ölçme ve Değerlendirme

Eğitimle kazandırılmak istenen davranışların hedeflenen düzeyde gerçekleşip gerçekleşmediğine karar verebilmek için öncelikle bu davranışların ölçülmesi gerekmektedir. Özçelik'e (2010) göre, yapılacak ilk iş öğrencilerin hedef davranışları kazanıp kazanmadıklarının ve bu davranışlarda eriştikleri yetkinlik ve kararlılık derecelerinin geçerli ve güvenilir ölçme araçlarıyla nesnel bir biçimde ortaya konmasıdır. Sonraki aşamada yapılacak iş ise, gerçekleştiği görülen öğrenme düzeyi ile gerçekleşmesi beklenen öğrenme düzeyi arasında fark olup olmadığına dair doğru bir yargıya ulaşmaktır. Dolayısıyla önce öğrenci öğrenmelerinin ne düzeyde gerçekleştiği ile ilgili bir ölçme ve ardından bu öğrenme düzeyinin hedeflenen öğrenme düzeyi ile karşılaştırıldığı bir değerlendirme işleminin yapıldığı söylenebilir.

Doğru bir ölçme yapabilmek için öncelikle ölçme kavramının ne olduğu ve onu oluşturan unsurların özelliklerini bilmek oldukça önemlidir. İngilizce öğretiminde ölçme ve değerlendirmeye yönelik yapılan yurt dışı alan taramalarında, "testing" ve "measurement" terimlerinin ölçme kavramı için kullanıldığı göze çarpmaktadır. Bu kavramlarla ölçme süreci içindeki aynı etkinliklere göndermeler yapıldığı için anlam olarak birbirleriyle örtüştükleri söylenebilir. İlk olarak ölçme kavramının sözlük anlamları incelendiğinde, ölçme "Bir niteliğin gözlenerek gözlem sonuçlarının sayı ya da başka sembollerle gösterilmesi" olarak belirtilirken (Demirel, 2012:106); başka bir tanımlamada "Birinin bilgisini ve yeteneğini sınayan bir takım sorular" şeklinde ifade edilmiştir (Cambridge Learner's Dictionary, 2004: 669). Brown (2004) ölçmeyi, öğrenenlerin belirli bir alandaki yetenek, bilgi ya da performansını sınama yöntemi olarak tanımlamaktadır. Ona göre, bu işlem farklı teknik, metot ve öğeleri içeren bir yöntemdir. Daha geniş bir perspektiften bakıldığında ise ölçme; birinin yeterlik ya da yetersizliğini, yeteneğini ya da

kabiliyeti olmadığı alanları gösteren ve bunu yaparken de onu başarılı, başarısız, ortalama, yeterli, iyi ve mükemmel şeklinde konumlandıran bir sınama sürecidir (Rahman, 2017). Kan'a göre (2018) ölçme bireyler ve nesnelerin belli bir özelliğe sahip olma dereceleri arasındaki farklılığı saptamak için ortaya çıkmıştır. Madsen (1983) ise ölçme için bir sınır çizmeden, ölçmenin öğretim ve öğrenme deneyimlerinin her aşamasında öne çıkan önemli bir mekanizma olduğuna dikkat çekmiştir.

Ölçmenin kullanılma sebebinin nesnelerin karakteristik özelliklerini niceliksel olarak belirlemek olduğunu ifade eden Nunnaly'e (1970) göre, bu tür nicelleştirmeler çeşitli avantajlar sunmaktadır. Bu avantajlar şöyle özetlenebilir (Nunnaly, 1970):

- Ölçme nesnelerin mevcut durumlarına ait niteliklerini ortaya çıkarır.
- Ölçme doğru bilgilere ulaşmayı kolaylaştırır.
- Ölçme grupların sınıflandırılmasını sağlar
- Ölçme grupların nesnel bir şekilde karşılaştırılmasını sağlar.
- Ölçme sonuçları daha fazla analiz yapılması için veri sağlar.

Bireylerin yetenekleri ya da performanslarıyla ilgili elde edilen ölçme sonuçlarının kendi bağlamına ait ölçütlerle karşılaştırılarak bir değer yargısına ulaşma süreci olarak adlandırılan değerlendirme (Yılmaz, 2002), karar verici bir rolünün olmasından dolayı ölçmeden ayrılır (Özçelik, 2010). Aslında öğretim sürecinin sonunda başvurulan bir mekanizma gibi görünse de White'a (2009) göre değerlendirme öğrenme sürecini başlatmaktan sorumlu olan bir motor olarak kabul edilmelidir. Magno ve Ouano'ya (2010) göre ise değerlendirme, öğrenme ve öğretim sürecinin tüm aşamalarıyla bütünleşmiş bir yapıdadır. Yani bu öğretimden önce, öğretim sırasında ve öğretimden sonra da gerçekleşebileceği anlamını taşır.

Değerlendirme kavramına ilişkin yapılan yurt dışı alan taramalarında, değerlendirme kavramı için "assessment" ve "evaluation" terimlerinin kullanıldığı göze çarpmaktadır. Kimi araştırmacıların değerlendirmeyi "evaluation" başlığı altında tartıştığı ve bu kavramı karar vermek amacıyla sistematik bir şekilde bilgi toplamak olarak tanımladıkları görülmektedir (Bachman, 1990; Harmer, 2001; Lynch, 2001). Kimi araştırmacıların ise "assessment" ve "evaluation" kavramlarını ayrı ayrı kullanmayı tercih ettikleri belirlenmiştir (Nunan, 1999; Dickins and Germanie, 1995; Murphy, 2000). Nunan (1999), "evaluation" kavramını öğretim süreciyle ilgili eksiklikleri saptayıp öğretim programına karar vermek için bilgi toplama ve yorumlama olarak tanımlarken,

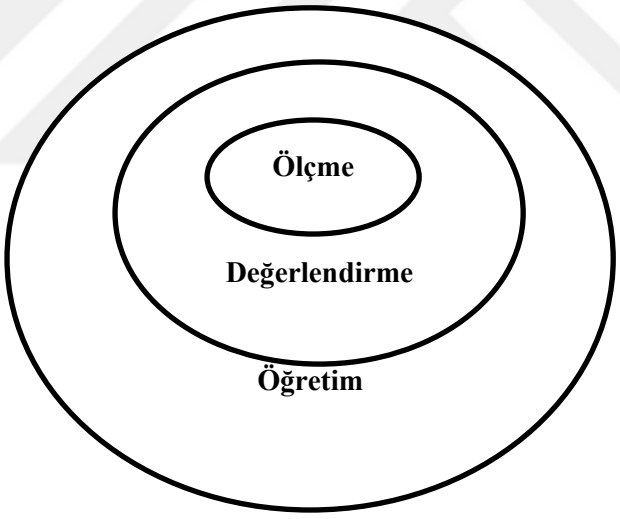
“assessment” kavramının ise “evaluation” kavramının bir alt başlığı olduğunu ve öğrencilerle ilgili bilgi toplama ve yorumlama sürecini ifade ettiğini belirtmiştir. Benzer şekilde Murphy (2000), “evaluation” kavramının bir programın hedeflerinin ne düzeyde gerçekleştiğine karar vermek için kullanıldığını ifade ederken, “assessment” kavramının ise öğrenmeyle ilgili ölçümler yapma ve oradan da bir yargıya ulaşma olduğunu belirtmiş ve bu iki kavram arasındaki farklılığa dikkat çekmiştir.

Son zamanlarda değerlendirme kavramının ölçmeyle eş anlamlı olduğu ve bu yüzden de ölçmenin yerine kullanılabileceğine dair bir algının yerleştiği görülebilir (Inbar-Lourie, 2008). Brown (2004) ölçme ve değerlendirmenin aynı anlama geldiğine dair yaygın olan kanıya dikkat çekmiş ve bu iki kavramı birbirinden ayırmıştır. Ona göre, ölçme yönetimsel süreçlerle ilgilidir ve öğrenenler bu süreçte gösterdikleri performanslarının aldıkları puanlarla belli bir karşılığının olduğunu bilirler. Diğer taraftan değerlendirme daha geniş bir alanı kapsayan ve sürekli devam eden bir süreçtir. Benzer şekilde Clapham (2000), ölçme ve değerlendirme kavramlarının birbirinden ayrı terimler olduğunu vurgulamış ve ölçmenin geniş gruplar için tasarlandığını, değerlendirmenin ise asıl amacının puan elde etmek olmadığı aksine öğrenenlerin problemlerinin tespiti için bireysel olarak yürütülen bir uygulama türü olduğunu belirtmiştir. Valette (1994) ise farklı bir bakış açısı sunarak, ölçmede geniş çaplı standart yeterlik testlerinin kullanıldığını belirtirken, değerlendirmede okul bazlı testlerin kullanıldığını ifade etmiştir.

Alanyazında ölçme ve değerlendirme kavramlarının farklı uygulamalar içerdiğini kabul eden ancak değerlendirmenin ölçmeyi kapsadığını belirten kullanımlar göze çarpmaktadır (Kunnan, 1998). Nitekim Clapham (2000) değerlendirme kavramını ölçme ve sınama faaliyetlerinin tümünü kapsayan bir şemsiye terim olarak görmektedir. Diğer taraftan Kizlik (2012) her ölçmenin bir değerlendirme olduğu ama her değerlendirmenin bir ölçme olmadığını ifade etmiştir. Bu düşünceyi destekleyen Dickins ve Germanie’ye (1995) göre öğrenenlerin performanslarını görmek için yapılan ölçme, değerlendirme sürecinin bir parçasıdır. Değerlendirme ise sadece öğrencilerin bilgi ve yeteneklerini test edilmesini değil aynı zamanda öğretmenlerin de kendi deneyimlerini gözden geçirmelerini sağlayan bir süreçtir. Dolayısıyla değerlendirme kavramının daha geniş bir anlamı karşılayacak biçimde kullanıldığı söylenebilir. Bu iki kavram arasındaki ilişkiyi belki de en iyi şekilde özetleyen Green’e (1970) göre ise, ölçme ve değerlendirme bir anlamda madeni

bir paranın ayrılamayan iki yüzü gibidir ve işbirliği içinde çalışmaktadırlar (Akt. Cesur, 2008).

Ölçmeyi değerlendirmenin bir alt kümesi olarak tanımlayan Hughes'a (2003) göre ölçme, öğrencilerin bazı becerilere ait özel amaçları belirli bir zaman diliminde gösterme yeteneklerini test etmek için yürütülen bir süreçtir. Değerlendirme ise ölçme işlemini de kapsayarak öğretim ile birlikte ilerler. Öğretim süreci başladığında, hemen hemen tüm öğretmenler öğrencilerin ilgilerini ve performanslarını bilinçli ya da bilinçsiz bir şekilde gözlemleyip bir yargıya ulaşırlar. Hatta öğrenciler bile kendi gelişimlerini akranlarıyla karşılaştırıp yorum yaparlar. Daha iyi bir öğrenme için öğretimi şekillendirmeyi amaçlayan bütün bu formal ve informal ölçümler değerlendirme olarak kabul edilir (Sah, 2012). Dolayısıyla nasıl ki ölçme değerlendirmenin bir parçasıysa, değerlendirmenin de öğretimin etkili bir şekilde işlemesi için çalışan ayrılmaz bir unsuru olduğu söylenebilir. Aşağıda Şekil 1'de ölçme, değerlendirme ve öğretim arasındaki ilişki gösterilmiştir.



Şekil 1. Ölçme, değerlendirme ve öğretim arasındaki ilişkiyi gösteren Venn şeması (Brown & Abeywickrama, 2010)

2.1.1. Eğitimde Ölçme Değerlendirmenin Yeri ve Önemi

Eğitim sürecinde ölçme ve değerlendirmenin öğrenciler, öğretmenler, veliler ve eğitim uzmanları gibi pek çok paydaş için oldukça önemli olduğu söylenebilir. Özellikle öğrenme, öğretim, öğretim programı gibi temel kavramların ne düzeyde kazandırıldığı veya işlediğiyle ilgili sağladığı dönütlerle çeşitli kararların alınmasına dayanak

oluşturmaktadır. Yaşar'a (2011) göre, ölçme ve değerlendirme eğitim sisteminde yer alan kişilere karar verme noktasında şu konularda fayda sağlamaktadır:

- Öğrenci başarı veya başarısızlıklarıyla ilgili kararlar
- Öğrencilerin ihtiyaç ve ilgileri doğrultusunda rehberlik sunulması ile ilgili kararlar
- Seçme ve yerleştirmeye dönük kararlar
- Öğretimin etkililiği ile ilgili kararlar
- Öğretmenlerin kendi öğretim faaliyetlerini değerlendirmelerine yönelik kararlar
- Öğrencilerin öğretim sürecine ve öğretmenlere ilişkin algılarına yönelik kararlar
- Mevcut programın nasıl işlediğine yönelik kararlar
- Öğretim sürecinin sonunda ulaşılan çıktılara yönelik kararlar
- Değerlendirme öğesinin nasıl uygulandığına ilişkin kararlar.

Görüldüğü üzere, ölçme ve değerlendirmenin öğrenme ve öğretim deneyiminin her aşamasında başvurulan kararlara cevap verici nitelikte olduğu söylenebilir. Öğrenenler açısından incelendiğinde, ölçme ve değerlendirme öğrenenlerin öğretim materyalleriyle anlamlı bir bağ kurmasını teşvik etmekte, onlara derste öğrendiklerini gözden geçirme şansı tanımakta ve öğrenenlerin kendi performanslarına yönelik dönüt sağlamaktadır (Cohen, 1994). Scanlan'a (2012) göre sevsek de sevmesek de pek çok öğrenci için ölçme bir motivasyon kaynağıdır ve çoğu öğrenci enerjisini sınavları en iyi şekilde geçmek için odaklamaya meyillidir. Angelo ve Cross (1993) öğrencilerin oldukça fazla anlam atfettikleri ölçmenin, doğru stratejilerle birlikte kullanılmasıyla öğrenme üzerinde etkili sonuçlar yaratabileceğini ileri sürmüştür. Örneğin, eğer öğretmenler bilgiyi hatırlamaya yönelik ölçmeye odaklanırsa yüzeysel bir öğrenmenin gerçekleşmesine sebep olacaklardır. Diğer taraftan problem çözmeyi ya da eleştirel düşünmeyi gerektiren değerlendirme stratejilerini kullanırlarsa, büyük ihtimalle öğrenci performanslarının daha yüksek bir seviyede gerçekleştiğini fark edeceklerdir (Angelo & Cross, 1993). Dolayısıyla doğru ve etkili bir değerlendirme deneyiminin öğrencilerin üstbilişsel düşünme becerilerini geliştirebileceği savunulabilir.

Eğitim sürecinde ölçme ve değerlendirmenin pek çok önemli fonksiyonunun olduğunu vurgulayan Turgut ve Baykul'a (2013) göre, bu dinamiğin öğrencilere sağladığı en önemli avantajlar; onlara davranışlarını nasıl geliştirecekleri hakkında bilgi sunma ve başarılı olan öğrencileri daha fazla motive etme olarak özetlenebilir. Diğer taraftan

Brissenden ve Slater (2002) farklı bir çerçeve sunarak, ölçme değerlendirme uygulamalarının öğrencilerin kendileriyle ilgili yanıt aradıkları birtakım sorulara cevap olacağını ifade etmişlerdir. Daha nitelikli bir öğrenme için cevaplanması gereken sorular ise şu şekildedir (Brissenden & Slater, 2002):

- Bu dersin sonunda kaç puan alacağım?
- Bu derse çalışma yöntemimi nasıl geliştirebilirim?
- Ders içeriğine yeteri kadar hakim miyim?

Eğitim sürecindeki öğretenler açısından bakıldığında ise, ölçme ve değerlendirme öğesinin onlara (1) öğretim hedeflerini belirlemede yardımcı olma, (2) daha sonraki değerlendirme faaliyetlerini geliştirmeleri için dönüt sağlama ve (3) öğrencilerin güçlü ve zayıf yönlerini saptama noktasında son derece önemli faydalar sağladığı söylenebilir (Cohen, 1994). Öğrenciler hakkında sunduğu verilere ek olarak, ölçme değerlendirme süreci öğretmenlerin kendi öğretim faaliyetlerini gözden geçirmelerine fırsat tanır ve öğretimin ne düzeyde etkili olup olmadığıyla ilgili sorulara yanıt bulmada oldukça işlevseldir (Madsen, 1983). Brissenden ve Slater'a (2002) göre sınıf içi yapılan değerlendirmeler öğretmenlere şu özel soruları cevaplamaları için yardımcı olabilecektir:

- Öğrencilerim önceden belirlenen amaçları ne ölçüde gerçekleştiriyorlar?
- Nasıl bir puanlama yapacağım?
- Bu konuyu daha etkili bir şekilde öğretebilir miyim?
- Öğrencilerim için dersin/ünitenin hangi bölümleri daha önemli?
- Bu dersi/üniteyi bir dahaki sefere nasıl anlatacağım?
- Şu anki ders konusu için ne kadar süre ayıracağım?

Yukarıdaki sorular incelendiğinde, ölçme değerlendirmenin öğrencilere sadece not vermenin ötesinde bir öneme sahip olduğu söylenebilir. Yılmaz'a (2002) göre öğretmenler ölçme ve değerlendirmenin öğretimin her aşamasında sunduğu sağlıklı bilgilerle, öğrenme sürecini kontrol altında tutabilir ve kendi başarılarını değerlendirebilirler.

Ölçme ve değerlendirmenin eğitim problemlerinin tespiti ve program geliştirme açısından da incelendiğinde pek çok işlevinin olduğu söylenebilir. Nitekim Magno ve Ouano'ya (2010) göre, değerlendirme sonuçları program uygulayıcılarına ve diğer ilgili paydaşlara sunduğu geribildirimlerle programın ne düzeyde işlediğini ortaya koymaktadır. Diğer taraftan bakıldığında, eğer öğretim programları ölçme ve değerlendirme temeline dayalı olmayan tahminlerle, öznel görüşlerle ya da öğrenci ve öğretmenlerden elde

edilecek verilerden yoksun bir anlayışla geliştirilirse, eğitim sistemi ve ülke adına büyük kayıplara yol açacaktır (Yılmaz, 2002).

2.2. Ölçme Değerlendirme Yaklaşımları

Ölçme ve değerlendirmenin eğitim sürecinde sunduğu geribildirimler göz önünde bulundurulduğunda, bu aracı işlevsel bir şekilde öğretim süreciyle bütünleştirmenin öğretimin her kademesinde yer alan kişiler için oldukça fayda saylayacağı söylenebilir. Bu noktada önemli olan değerlendirmelerin eğitsel amaçlara hizmet edecek şekilde doğru olarak kullanılmasıdır. Dolayısıyla araştırmanın bu başlığı altında değerlendirme amacına bağlı olarak alanyazında yer alan ölçme değerlendirme yaklaşımları sunulacaktır.

2.2.1. Geleneksel ve Alternatif Değerlendirme

Dünyanın pek çok yerinde öğrenme ve öğretim yöntemleri üzerine değişen paradigmalara rağmen hala kullanılan geleneksel yaklaşımın temelinin 1800'lü yıllarda ortaya çıkan ve öğretim faaliyetlerinin gözlenmesi için kullanılan testlere dayandığı söylenebilir (Atta-Alla, 2012). Ghaicha ve Omarkaly'ye (2018) göre, o zamandan beri süregelen geleneksel yaklaşım pek çok farklı eğitim sisteminde en fazla tercih edilen değerlendirme yöntemlerinin başında gelmektedir. Nitekim çeşitli disiplinlerde yapılan çalışmaların ortaya konan öğretmenlerin büyük bir çoğunluğunun geleneksel değerlendirme tekniklerine sıklıkla başvurdukları yönündeki bulgular bu tezi doğrular niteliktedir (Band; 2005; Watt, 2005; Akbulut ve Akbulut, 2011; Majid, 2011). Brown'a (2004) göre bu tür geleneksel yöntemlerin daha fazla tercih edilmesinin sebebi geçerli ve güvenilir ölçme sonuçları sağlaması ve öğretmenler için kullanımının daha kolay olmasıdır. Bu görüşü destekleyen Bailey (1998) geleneksel testlerin doğası gereği çoğunlukla standart olduklarını ifade etmiştir. Geleneksel değerlendirme uygulamalarını savunan kimi araştırmacılara göre, bu tür standart testler öğretmenlerin öğretim programında kazandırılması öngörülen bilgilerin öğrencilere hızlıca transferini sağlayabilmektedir (Serafini, 2001).

Geleneksel değerlendirmede çoğunlukla çoktan seçmeli testler, doğru-yanlış ifadeleri, boşluk doldurma, eşleştirme ve kısa cevap gibi öğrencilerin belirli bir alandaki bilgi ve yeterliklerini sınayan teknikler yer almaktadır (Nasab, 2015). Kwako'ya (2003) göre, bu tür geleneksel teknikler öğrenci gelişimini izlemenin sadece kolay ve hızlı yoludur. Ancak hızlı uygulama ve puanlama öğretmenlere kolaylık sunsa da öğrencilerin eğitsel ihtiyaçları, ilgileri ve gelişimleriyle ilgili karar vermek için tüm gerekli bilgileri

sağlamamaktadır (Letina, 2014). Geleneksel değerlendirme tekniklerinin dezavantajlarına dikkat çeken Nasab'a (2015) göre iki seçenekten birinin doğru olduğunu gösteren doğru-yanlış testleri öğrencilere doğru cevabı bulmak için %50 şans sunmaktadır. Eşleştirme alıştırmaları ya da çoktan seçmeli testlere bakıldığında, doğru cevabı bulmak için yine belli düzeyde şans faktörünün etkili olduğu görülmektedir. Dolayısıyla geleneksel değerlendirme tekniklerinin güvenilir ölçme sonuçları sunduğuna dair görüşlerin eleştirildiği söylenebilir.

Öğrenmenin ne düzeyde gerçekleştiğini ölçmek için temel ve tek kaynak olarak geleneksel değerlendirme tekniklerine aşırı bağlanmak pek çok sorun yaratabilir (Torrance & Pryor, 1998). Shepard'a (1991) göre, geleneksel değerlendirmede kullanılan standart testler öğretmen ve öğrencilerin dikkatini sadece sınavlarda önemsenen konulara yöneltmektedir. Bu yüzden standart testler tüm öğretim programına yön vermekte ve bir risk oluşturmaktadır. Diğer taraftan Wall (1996) geleneksel araçların öğretim yöntemini de etkilediğini ifade etmiş ve öğretmenlerin diğer aktiviteleri bu testler uğruna harcayarak çeşitli sınava hazırlık alıştırmaları yapmayı tercih ettiklerini öne sürmüştür. Geleneksel değerlendirmeye psikolojik bir bakış açısı sunan Broadfoot'a (2003) göre ise, standart testler öğrencilerin motivasyonları, ilgileri ve çabalarına odaklanmadan onları pasif bir alıcı konumuna getirdiği için psikolojik olarak da onları olumsuz bir şekilde etkilemektedir.

Zaman içerisinde öğrenme ve öğretim yöntemlerinin doğasına ilişkin gelişen anlayışın değerlendirme yöntemlerinin değişimini de zorunlu kıldığı söylenebilir. Özellikle bir şeyler öğrenmenin niçin önemli olduğu ve bilginin karmaşık ve gerçek yaşam durumlarında nasıl başarılı bir şekilde kullanılabileceğine ilişkin soruların vurgulanmaya başlamasıyla (Letina, 2014); geleneksel değerlendirmenin sunduğu yüzeysel geribildirimlere yönelik eleştiriler gelmeye başlamıştır (Higuchi, 1993). Alternatif değerlendirme olarak adlandırılan bu yeni anlayışla artık öğrencilerin neyi hatırladıkları veya ezberlediklerinden çok ne ürettiklerine dayalı değerlendirmeler yapılması görüşü ön plana çıkmıştır (Huerta-Macias, 1995). Buhagiar'ın (2007) ifade ettiği gibi, alternatif değerlendirme daha yüksek seviyedeki eğitsel çıktıları kazandırmak için öğretenlerin yönelmesi gereken bir alandır.

Alternatif değerlendirme öğrencilerin düşünme, analiz yapma, bilgiyi yeni durumlarda kullanma ve kavramlar arasında ilişki kurma yeteneklerini tespit etmeye çalışır

(Earl & Cousins, 1995). Bu tür değerlendirmeler öğrenci merkezlidir ve öğrencilerin kendi öğrenme süreçlerini kontrol etmeleri ve sorumluluk almaları için onları motive eder (Stiggins, 1997). Ayrıca öğretim sürecinin sonunda ve kısıtlı koşullar altında yürütülen geleneksel değerlendirmelerin aksine, informal ve süregelen bir süreç içerisinde devam eden alternatif değerlendirme, öğrencilere öğrenmeyi daha çok deneyimleyebilmeleri için bol bol fırsat sunmaktadır (Huerta-Macias, 1995). Brown ve Hudson (1998) alternatif değerlendirmenin sunduğu fırsatları şu şekilde özetlemişlerdir:

- Öğrencilere her gün sınıfta normal bir şekilde yaptıkları görevlerle değerlendirilme olanağı tanır
- Ürünün yanında süreç de değerlendirilir
- Öğrencilerden bir şey yapmaları, yaratmaları, üretmeleri ya da performans göstermeleri beklenir
- Anlamlı öğretim aktivitelerini içeren görevler sunar
- Öğrencilerin problem çözme ve üst düzey düşünme becerilerini geliştirir
- Öğrencilere zayıf ve güçlü yanlarıyla ilgili bilgi sağlar
- Hümanistiktir ve insanların makine olmadığını göz önünde bulundurarak bir puanlama sunar.
- Öğretmenlerin yeni öğretici ve değerlendirici roller üstlenmelerini gerektirir.

Geleneksel değerlendirmede genellikle öğrencilerin alıcı becerilerine yönelik yeterliklerini ölçmek için kullanılan seçmeyi gerektiren testlerin (Anderson, 1998) aksine, alternatif değerlendirme öğrenciye anlamlı ve ilginç gelen ve gerçek yaşamla bağlantılı pek çok farklı yeterlik alanın değerlendirilmesi için kapsayıcı teknikler sunar (Duban ve Küçükyılmaz, 2008). Değerlendirmeyi öğrenmenin bir parçası olarak gören alternatif yaklaşım, öğrenci yeteneklerini ve onların problem çözme becerilerini değerlendirirken öğrenci ürün dosyası, öz-akran değerlendirme, performans değerlendirme, proje, portfolyo gibi pek çok teknik işe koşar (Mintah, 2003). Brown'a (2004) göre, alternatif değerlendirme stratejileriyle öğrenciler eleştirel düşünme becerilerini geliştirerek kendi öğrenmelerinin, ürettiklerinin, güçlü ve zayıf yönlerinin üzerine analiz yapma fırsatı bulurlar. Örneğin, alternatif değerlendirme tekniklerinden olan öz değerlendirmeyle, öğrenciler yansıtıcı bir bakış açısıyla kendi yeterlikleriyle ilgili yargıda bulunabilir ve bu durum onların kendilerine olan güvenleri ve motivasyonlarını arttırabilir (Nurov, 2000).

Geleneksel ve alternatif değerlendirme yaklaşımlarının kuramsal ve pragmatik olarak pek çok farklı yönlerinin olduğu ve içinde çeşitli avantaj ve dezavantajları barındırdıkları söylenebilir. Nasab'a (2015) göre geleneksel değerlendirme araçlarının sunduğu en önemli avantaj, uygulanma ve puanlamasının kolay ve objektif olmasıdır. Ancak geleneksel değerlendirmede kullanılan testler çoğunlukla seçmeyi gerektiren cevaplara dayandığı için öğrencilerin üretimsel becerilerinin kontrolü ihmal edilmekte ve belirli alanlardaki bilgi ve yeterlikleri ölçülmektedir. Ayrıca Bailey (1998) geleneksel değerlendirmelerin tek bir beceriye ve hıza dayalı ve norm-referanslı olduklarını vurgulayarak, bu tür değerlendirmelerin otantik olmadığını eklemiştir. Diğer taraftan, alternatif değerlendirmelerin doğası gereği öğretim süreci boyunca devam ettiği ve öğrenenlerin eğitsel alanlardaki üretimsel beceriler de dahil olmak üzere pek çok yeterlik alanındaki gelişim süreçlerine ayna tuttuğu ifade edilmiştir (Hamayan, 1995). Ayrıca Genesee ve Upshur (1996) öğrencilerin başarısını etkileyen öğrenme stratejileri, davranışlar ve duygular gibi değişkenlerin açığa çıkarılması için alternatif değerlendirmelerin oldukça gerekli olduğunu vurgulamışlardır. Alternatif değerlendirme yaklaşımlarıyla ilgili en büyük eleştirilerden biri ise, bu tür bir değerlendirme yapmanın uygun analizler ve sürekli geribildirim sunmayı gerektirmesi nedeniyle öğretmenler açısından zaman alıcı ve maliyetli olduğudur (Brindley, 2001). Ek olarak, öğretmenlerin alternatif değerlendirme araçlarını hazırlama ve uygulama konusunda yeterli bilgi ve donanıma sahip olmaları gerekmektedir (Clark & Gipps, 2000). Geleneksel ve tamamlayıcı değerlendirme yöntemlerinin dayandıkları temel felsefe ve içerdikleri uygulamalar bakımından karşılaştırmalı analizi aşağıdaki çizelgede sunulmuştur (Anderson, 1988)

Geleneksel Değerlendirme	**Alternatif Değerlendirme**
Bilginin tek bir anlamı olduğunu varsayar	Bilginin birden çok anlamı olduğunu varsayar
Öğrenme pasif bir süreçtir	Öğrenme aktif bir süreçtir
Öğrenme sürecini üründen ayrı tutar. (Sadece ürünü değerlendirir)	Öğrenme süreci ve ürünü birlikte değerlendirir
Bilgi parçalarını kullanmaya odaklıdır	Araştırma ve problem çözme odaklıdır
Değerlendirmenin amacı öğrenmeyi belgelemek ve öğrencileri puanlarına göre sınıflandırmaktır.	Değerlendirmenin amacı öğrenmeyi kolaylaştırmaktır

Öğrencilerin bilişsel yeterliklerine odaklanır	Öğrencilerin bilişsel, duyuşsal ve psikomotor becerileri arasında bir bağlantı olduğunu kabul eder
Öğrenciler karar verme sürecine dahil değildir	Öğrenciler kara verme sürecine dahildir
Öğrenmeyi bireysel bir çaba olarak algılar (Öğrenciler verilen görevleri bağımsız bir şekilde gerçekleştirir)	Öğrenmeyi işbirlikçi bir süreç olarak algılar (Öğretmenleri ve öğrenciler verilen görevleri birlikte yürütür)

Geleneksel ve alternatif değerlendirme yaklaşımlarının karşılaştırıldığı çizelgede, Anderson (1988) geleneksel değerlendirmenin geleneksel öğretim yöntemlerini değerlendirmenin en iyi yolu olduğuna dikkat çekmiştir. Nitekim geleneksel öğretim gerçeklerin sunulmasına ve onların yeniden tekrarlanmasına dayalıdır. Ancak bu değerlendirme yaklaşımı, öğrencilerin bilişsel farklılıkları ve öğrenme stillerinin göz ardı edilmediği proje çalışması, araştırma etkinliği, öğrenmeyi keşfetme gibi uygulamaları içinde barındıran çağdaş öğretim yöntemleri için uygun değildir (Anderson, 1988). Diğer taraftan alternatif değerlendirmenin yapılandırmacı eğitim uygulamalarına uygun strateji ve teknikleri içerdiği söylenebilir (Mintah, 2003).

2.2.2. Biçimlendirici ve Özetleyici Değerlendirme

Biçimlendirici (formative) değerlendirme, öğrencilerin performansını sadece sınamak için değil aynı zamanda onların gelişimini izlemek ve geliştirmek için tasarlanmış bir değerlendirme türüdür (Wiggins, 1998). Black ve William'a (2010) göre, biçimlendirici değerlendirme öğrenim ve öğretim aktivitelerini düzenlemek için geribildirim sunan ve hem öğrencilere hem de öğretmenlere kendilerini değerlendirme fırsatı tanıyan bir değerlendirmedir. Biçimlendirici değerlendirme öğrenmenin değerlendirilmesi anlayışından ziyade değerlendirmenin öğrenme için yapıldığı felsefesine dayanmaktadır (Stiggins, 2002). Yani bu yöntemdeki temel amaç öğrenci performanslarının değerlendirilmesiyle ortaya çıkan verileri onların başarılarını arttırmak için kullanmaktır. Böylece ortaya çıkan geribildirimlerle sadece mevcut durumla ilgili analiz yapılmakla kalmayıp öğrencilerin gelecekteki başarılarını olumlu olarak etkilemek istenmektedir (Nichols, Meyers & Burling, 2009). Dolayısıyla biçimlendirici değerlendirmenin temel özelliklerinden birinin öğrenme eksikliklerinin farkına varmak ve bunları gidermek olduğu söylenebilir. Bu değerlendirme türünde göze çarpan bir diğer özellik ise değerlendirme

uygulamalarının öğretim süreci boyunca devam etmesidir. Böylece sürece yayılan bu değerlendirme yönteminden elde edilen bilgiler ışığında öğretim sürecinin yeniden planlanması yapılabilir (Dick, Carey & Carey, 2001).

Biçimlendirici değerlendirmenin temel olarak öğrencilerin devam eden gelişimine odaklanmasından dolayı, Brown (2004) neredeyse her tür informal değerlendirme türünün biçimlendirici değerlendirme olduğunu ileri sürmüştür. Ona göre öğrencilerin hatalarının düzeltildiği, yorum yapıldığı veya onlara bir öneride bulunulduğu zaman sunulan geribildirimlerin amacı onların becerilerini geliştirmektir. Ya da bir öğretmen biçimlendirici değerlendirmenin sunduğu bu veriler sayesinde öğrenme süreci boyunca kullandığı öğretim tekniklerini değiştirebilir. Bu düşünceyi destekleyen Cook'a (2009) göre biçimlendirici değerlendirmeler kendiliğinden ve planlı olmak üzere iki şekilde gerçekleşebilir. Kendiliğinden gerçekleşen değerlendirmelerin Brown'un da öne sürdüğü gibi öğretim etkinleri boyunca yapılan soru cevap etkinlikleri veya öğrenci hatalarını düzeltme gibi faaliyetlerle uygulandığı belirtilirken; ara sınav, ev ödevleri, projeler gibi aktivitelerle ise planlı bir şekilde yapılabileceği vurgulanmıştır (Akt. Dixson & Worrel, 2016). William ve Thompson'a (2008) göre, biçimlendirici değerlendirmede önemli olan ne şekilde uygulandığından ziyade öğretmenlerin öğrenci gelişimini destekleyici stratejileri planlı ya da plansız mutlaka kullanmaları gerektiğidir.

Bir diğer değerlendirme türlerinden olan özetleyici (summative) değerlendirme ise genel olarak öğrencilerin okulu bitirmeleri, bir derece almaları ya da sertifika verilmesi gibi durumlarda öğrencilerin yeterlik düzeyleri ile ilgili bir karar vermek için kullanılır (OECD, 2013). Öğrencilere ve öğretmenlere geribildirim sunması amacıyla kullanılan biçimlendirici değerlendirmenin aksine özetleyici değerlendirme, katılımcılar için önemli sonuçları olan bir süreçtir ve öğrenmenin ne ölçüde gerçekleştiğiyle ilgili nihai bir yargıya ulaşmak için kullanılır (Gardner, 2010). Final sınavları, ülke genelinde yapılan sınavlar, lise, üniversite giriş sınavları ve dönem sonu ödevleri gibi sonuca ve ürüne dayalı testler özetleyici değerlendirmenin örneklerindendir (Dixson & Worrel, 2016). Bu özetleyici değerlendirme araçları sadece çoktan seçmeli standart testleri değil aynı zamanda performans görevlerini de içermektedir. Böylece öğrencilerin genel olarak başarı ya da başarısızlık durumlarıyla ilgili karar vermek için sonuca bakılsa da aslında problem çözmeleri ve eleştirel düşünmeleri gerekmektedir (Common Core Standards, 2014).

Biçimlendirici ve özetleyici değerlendirme türlerinin temel özelliklerine ilişkin karşılaştırmalı analiz aşağıdaki gösterilmiştir (Dixson & Worrel, 2016).

Karakteristik Özellikler	**Biçimlendirici Değerlendirme**	**Özetleyici Değerlendirme**
Amaç	Öğretimi ve öğrenmeyi geliştirmek, Öğrenci eksikliklerini teşhis etmek	Öğrenme çıktılarının değerlendirilmesi, Yerleştirme, bir üst seviyeye çıkarma kararı
Resmilik	Genellikle informal	Genellikle formal
Uygulama zamanı	Sürekli devam eden, öğretimden önce ya da öğretim sırasında	Öğretimden sonra
Önem derecesi	Ölçme sonuçları ikinci derecede önemlidir	Katılımcılar için sonuçları birinci derecede önemlidir
Zorluk derecesi	Düşük seviyeden yüksek seviyeye	Orta seviyeden yüksek seviyeye
Cevap aranan sorular	Mevcut durum ne? Hangi ihtiyaçlar geliştirilmeli?	Öğrenci konuyu anladı mı? Öğrenci bir sonraki aktivite düzeyi için hazır mı?
Örnekler	Gözlemler, ev ödevleri, soru-cevap etkinlikleri, öz-değerlendirmeler, Öğretim programına dayalı ölçmeler	Sınıf içi sınavlar, ulusal düzeydeki sınavlar, performans değerlendirme, portfolyolar

Biçimlendirici ve özetleyici değerlendirmenin temel özellikleri karşılaştırıldığında, biçimlendirici değerlendirmenin öğrenim sürecine daha fazla katkı sunduğu söylenebilir. Nitekim Harris (2007) ve Taras (2008) yaptıkları çalışmalarla biçimlendirici değerlendirmenin özetleyici değerlendirmeden daha öncelikli ve önemli olduğuna dair bulguları paylaşarak bu düşünceyi desteklemişlerdir. Benzer bir bulguya Kordes, Sicherl & Brunauer'in (2014) araştırmasında ulaşılmış ve öğrencilerin biçimlendirici değerlendirme süreçleri kullanıldığı zaman daha olumlu tutumlar geliştirdikleri saptanmıştır. Diğer taraftan alanyazında biçimlendirici değerlendirmenin öğrenim ve öğretim sürecine sunduğu katkılar bakımından özetleyici değerlendirme karşısındaki üstünlüğünü vurgulayan çalışmalar göze çarpsa da Black ve William (1998) yabancı dil öğrenimiyle ilgili yaptıkları çalışmada, biçimlendirici değerlendirmenin etkililiğiyle ilgili herhangi bir kanıta ulaşmamışlardır. Mahshanian, Shoghi ve Bahrami'ye (2019) göre ise bu değerlendirme

türlerini birbirinden bağımsız bir şekilde kullanmaktansa bütünleştirip kullanmak daha yüksek düzeyde öğrenmelerin gerçekleşmesini sağlayacaktır.

2.2.3. Formal ve İnformal Değerlendirme

Formal değerlendirmeler öğrencilerin belirli bir değerlendirme amacı için çeşitli görevleri yerine getirmeleridir (Scanlan, 2012). Formal değerlendirmeler önceden belirlenmiş, yapılandırılmış ve standart bir hale getirilmiş ölçme araçları kullanılarak yapılır (Wortham & Hardin, 2012). Brown'a (2004) göre, bu tür değerlendirmeler bilgi ve becerilerin kaynağına ulaşmak için özel olarak tasarlanmış süreçlerdir. Ayrıca öğrenci başarılarıyla ilgili bir değer yargısına ulaşmak için sistematik, planlı ve yapılandırılmış bir şekilde yürütülmektedir. Çoğu formal değerlendirmelerin temelinin özetleyici değerlendirmelere dayandığını belirten Scanlan'a (2012) göre, bu tür değerlendirmelerin doğası gereği karar verici bir mekanizma olmalarından dolayı daha yüksek geçerlik ve güvenirlik indekslerine sahip olmaları beklenmektedir.

Diğer taraftan informal değerlendirmeler görüşme, gözlem, sınıf içi etkinlikler gibi standart olmayan ya da önceden yapılandırılmamış ölçme ve değerlendirme tekniklerini kapsamaktadır (McAfee & Leong, 2012). Biçimlendirici değerlendirmede olduğu gibi planlanmamış ve kendiliğinden gelişen değerlendirmeleri ifade etmek için kullanılır (Brown, 2014). Daha çok öğretmen yapımı olan informal değerlendirme araçlarını hazırlamak ve uygulamak standart testlerde olduğu gibi uzmanlık gerektirmeyebilir (Kargın, 2007). Popham'ın (2009) tanımlamasına göre ise formal ve informal değerlendirmelerin ikisi de sınıf içi değerlendirmelerdir ve biçimlendirici değerlendirmenin bir alt grubudur. Popham bu tür değerlendirmelerin öğrencilerin ne bildikleri ve asıl potansiyellerinin ne olduğunu ortaya çıkarmak, programda bir değişiklik yapmak ve kullanılan öğretim yöntemlerinin uygun ve etkili olup olmadığını saptamak için öğretmenler tarafından kullanıldığını belirtmiştir.

2.2.4. Doğrudan ve Dolaylı Değerlendirme

Öğrenci başarıları hakkında verilecek kararlara doğrudan veya dolaylı yapılacak ölçümlerle ulaşılabilmektedir. Merhout vd. (2008) doğrudan ölçmeyi, öğretim programının öngördüğü öğrenme çıktılarını öğrencilerin ne düzeyde kazandığını tespit etmek için, ortaya çıkan öğrenci ürünlerini sistematik ve objektif bir şekilde sınamak olarak tanımlamaktadır. Menhart'a (2011) göre doğrudan değerlendirme ilk olarak öğrenci performansının ölçülmesini ve ardından onların başarı-başarısızlık durumları hakkında bir

sonuca varmayı ifade etmektedir. Bu karar verme sürecinde kullanılabilecek doğrudan ölçme araçları, öğrencilerin performansını göstermelerini sağlayan öğretmen yapımı testler, ödevler, dönem sonu yapılan sınavlar ve projeler gibi değerlendirme araçlarını kapsamaktadır (Rajak, Shrivastava, Tripathi, 2019). Doğrudan ölçmeye farklı bir perspektif sunan Hughes' a (2003) göre, konuşma ve yazma gibi üretimsel becerilerin test edilmesi amaçlandığı zaman doğrudan ölçümler yapmak daha kolaydır. Öğrenenler konuşma ve yazma becerilerine yönelik davranışları doğrudan bir şekilde gösterebildikleri için onların bu alanlardaki yeterlikleriyle ilgili bir sonuca varmak daha kolay olacaktır. Ancak okuma ve dinleme becerileri ölçülmek istendiğinde, öğrenenlerin önce dinleme ve okuma yapmaları ve ardından bunu başarılı bir şekilde yaptıklarını göstermeleri gerekmektedir.

Doğrudan yapılan ölçümlerin aksine dolaylı değerlendirmeler, öğrenci performansının doğrudan ölçülmesini amaçlamaz ancak bu konuda yararlı bilgiler sunabilir (Yamayee & Albright, 2008). Dolaylı değerlendirme öğrenme çıktılarıyla ilgili becerileri öğrencilerin ne düzeye algıladıklarını belirlemek için kullanılır (Merhout vd., 2008). Hughes'ın (2003) tanımlamasına göre ise, öğrenci başarısını dolaylı olarak sınamak demek ölçmek istediğimiz temel becerilerin alt becerilerini öğrencilerin ne düzeyde gerçekleştirdiğini ortaya koymak demektir. Örneğin, TOEFL sınavında yazma becerisini ölçmek için öğrencilerden cümle içindeki uygun olmayan ya da hatalı yapıları belirlemesini istemek gibi (Hughes, 2003).

Diğer taraftan Bailey (1998), dolaylı ölçme ve değerlendirme sürecinin dil becerilerinin nasıl tanımlandığına bağlı olarak anlam kazandığını ifade etmiştir. Ona göre eğer yazma becerisinin daha çok dilbilgisiyle ilgili olduğunu düşünüyorsanız, o zaman bu beceriyi ölçmek için dilbilgisi temelli ölçümler yaparsınız. Sesletim bilgisini konuşma becerisinin önemli bir alt becerisi olarak kabul ediyorsanız o zaman da sesbirim bilgisine yönelik sorularla bu beceriyi değerlendirirsiniz (Akt. Yavuz-Kırık, 2008). Dolayısıyla ölçmek istediğiniz beceriyi çoktan seçmeli sorular, doğru-yanlış testleri gibi daha yapay uygulamalar kullanarak öğrencilerin performanslarına dair çıkarımlarda bulunabilirsiniz (Davies vd., 1999).

2.2.5. Nesnel ve Öznel Değerlendirme

Öğrencilerin sınav sonuçlarının değerlendirilmesi sürecinde yapılan puanlamalar nesnel veya öznel ölçütlere göre gerçekleşebilir. Bachman'a (1990) göre eğer sınavı alan

kişilerin sorulara verdikleri yanıtlar önceden kararlaştırılan kriterlere uygun puanlanıyorsa nesnel değerlendirme; ancak puanlayıcının doğru cevaplara karar vermesi bir yorumlama gerektiriyorsa bu bir öznel değerlendirmedir (Akt. Ölmezer-Öztürk, 2018). Çoktan seçmeli, kısa cevaplı, eşleştirmeli soruların içinde bulunduğu seçmeyi gerektiren testlerde doğru cevap önceden belirlendiği için nesnel bir değerlendirme; diğer taraftan açık uçlu sorular ya da proje görevlerinin puanlanması işlemi ise öznel bir değerlendirme gerektirmektedir (Baran, 2020). Dolayısıyla nesnel değerlendirme okuduğunu anlama, sözcük, dil yapıları gibi yeterlikleri ölçerken, öznel değerlendirmenin ise fikirleri, üretkenliği ve yaratıcılığı ölçtüğü söylenebilir.

Nesnel ve öznel değerlendirmeleri karşılaştıran Shaban'a (2014) göre, nesnel değerlendirmeler öğrencilerin kısa bir sürede cevaplaması gereken soruları içerdiğinden daha geçerli ve güvenilir sonuçlar elde edilmektedir. Ama öznel değerlendirme öğrencilerin deneyimlerine dayanmaktadır. Öğrencinin sunduğu yanıtlar farklı puanlayıcılar tarafından aynı yorumlanmayabilir. Bu yüzden geçerlik ve güvenirlik indeksleri daha düşüktür.

2.2.6. Norma ve Ölçüte Dayalı Değerlendirme

Norma ve ölçüte dayalı değerlendirme ölçme sonuçlarının yorumlanmasıyla ilgili bilgi sağlayan değerlendirme türleridir (Hughes, 2003). Yurt içi kaynaklarda norma ve ölçüte dayalı değerlendirme yerine bağıl ve mutlak değerlendirme kavramlarının kullanıldığı görülmektedir (Tekin, 2004; Nartgün, 2007; Atılgan, 2018). Bu kavramların ölçme değerlendirme sürecinde aynı uygulamaları içermelerinden dolayı anlam olarak örtüştükleri söylenebilir. İlk olarak Glaser (1963) tarafından kavramsallaştırılan ölçüte bağlılık kavramı özellikle testlerin kapsam geçerliğini ve kullanışlılığını arttırmak amacıyla ortaya çıkmıştır (Shepard, 1980). Bond'a (1996) göre bu iki değerlendirme türü amaçları, içerik seçim yöntemleri ve ölçme sonuçlarının puanlamasının nasıl yapılacağı gibi pek çok yönden farklılık göstermektedir.

Norma dayalı değerlendirme öğrencilerin test puanlarının belirli bir normla karşılaştırıldıktan sonra başarı ya da başarısızlık durumlarına karar verildiği bir değerlendirme türüdür (Shepard, 1980). Norma dayalı değerlendirmede öğrencilerin performansı gruptaki akranlarıyla karşılaştırılır ve aldıkları puanlara göre sıralanır (Burton, 2006). Öğrencileri aldıkları test puanlarına göre bu şekilde sıralamak genellikle "çan eğrisi" olarak tanımlanır (Burton, 2006). Yani önceden kararlaştırılan bir başarı ölçütü

yerine sınava giren öğrencilerin aldıkları puanların ne düzeyde yeterli olduğu diğer sınava girenlerin puanları referans alınarak yorumlanır (Hussain, Tadesse & Sajid, 2015). Örneğin bir öğrencinin aldığı puanın yüzdelik dilimi 34 ise, bu öğrenci norm gruptaki öğrencilerin % 34'ünden daha iyi bir performans göstermiş demektir (Bond, 1996). Norma bağlı değerlendirmelerde gruptaki bireylerin puanları yüzdelikler, Z, T gibi standart puanlar, normal dağılım eğrisi altında kalanlar, standart sapma ve sınıfın başarı ortalaması gibi grup puanları ölçüt alınarak değerlendirilir (Atılgan, 2018). TOEFL gibi standart testlerden alınan puanların yorumlanması norma bağlı değerlendirmeye örnek gösterilebilir (Brown, 2004)

Bir ölçüte dayalı yapılan değerlendirme ise norma dayalı değerlendirmeden farklı olarak öğrenci başarısının gruptaki diğer öğrencilerden bağımsız olarak önceden belirlenen ölçütlerle karşılaştırılarak bir karara varmaya dayanır (Cox & Vargas, 1966). Scarino'ya (2005) göre bu ölçütler şu amaçlara hizmet etmektedir: Değerlendirme için savunulabilir bir çerçeve sunmak, ihtiyaçları tanımlamak ve açıklamak, bir bağlam oluşturmak, alınan kararlarını ispatını kolaylaştırmak, öznelliğe ve önyargılara karşı önlem almak ve adil değerlendirmeyi garantilemek. Shepard (1980) bu ölçütlerin iyi bir şekilde önceden belirlenmesiyle, öğrenci performanslarının doğru olarak puanlanabileceğini belirtmiştir.

Alan taramalarında mutlak değerlendirme olarak yer bulan bu değerlendirme türünü Erkuş (2006), öğrenci performansının yeterli olarak kabul edilmesi için çeşitli yollarla önceden belirlenmiş olan minimum standart bir değerin referans alınması olarak tanımlamıştır. Örneğin, MEB'in puan ortalaması 45 puan ve üstü olanlar başarılı sayılır ölçütü buna örnek verilebilir. Nartgün (2007) mutlak değerlendirmenin genel olarak işlem adımlarını şu şekilde özetlemiştir: İlk olarak değerlendirme ölçütleri belirlenir. Sonra ilgili derse ilişkin öğretim çalışmaları yürütülür. Ardından öğrencilerin erişmeye çalıştıkları kazanımlarla ilgili sınav yapılır. Son olarak ise her bir öğrencinin elde ettiği puanlar önceden belirlenen ölçütle karşılaştırılarak ne düzeyde başarı sağladığıyla ilgili yargıya ulaşılır. Shepard'a (2008) göre değerlendirme öncelikleriyle ilgili bir başkalaşmanın yaşandığına dair kanıtlardan biri norma bağlı değerlendirmeden ölçüte bağlı değerlendirmeye yavaş ama istikrarlı bir geçişin olduğudur. Bu durum bizim değerlendirme amacıyla ilgili kolektif düşüncemizin de paralel bir şekilde değişmesini sağlamış ve öğrencileri sadece başarılarına göre kıyaslamaktansa kimin hangi standardı karşılayıp karşılayamadığını sorgulamamıza katkı sunmuştur.

Sonuç olarak şimdiye kadar incelenen ölçme değerlendirme yaklaşımlarının temel anlayış ve uygulamaları göz önünde bulundurulduğunda, değerlendirmenin hangi amaçla, ne zaman, nasıl ve hangi ölçütlerle yapıldığına dair sorulara verilecek yanıtların ölçme değerlendirme ilkelerine uygun bir çerçevede temellendirilmesinin, öğrenci performanslarına yönelik doğru karar verilmesini kolaylaştıracağı söylenebilir. Değerlendirmeyi öğrenmenin önemli bir partneri olarak gören Brown (2004), uygulayıcılara rehberlik yapması için iyi bir değerlendirmenin genel özelliklerini şu şekilde özetlemiştir:

1. Hem formal hem de informal periyodik değerlendirmeler öğrenci gelişiminin yapıtaşları olarak işe koşulduğunda öğrencilerin motivasyonu artacaktır.
2. Doğru değerlendirmeler bilginin akılda tutulması ve sağlamlaştırılmasını kolaylaştırır.
3. Değerlendirme daha sonraki çalışmalar için gereken güçlükleri ve önemli noktaları saptamaya yardımcı olur.
4. Değerlendirme öğrencilerin öz değerlendirmelerini ve gelişimlerini destekleyerek öğrenci özerkliğini teşvik eder.
5. Değerlendirme öğrenenlere kendileri için hedef koymaya yönlendirir.
6. Değerlendirme öğretimin etkililiğini sınamaya yardım eder.

2.3. YABANCI DİL ÖĞRETİMİNDE ÖLÇME VE DEĞERLENDİRME

Yabancı dil öğretiminde ölçme ve değerlendirme süreciyle ilgili yapılan araştırmalar, bu alanda kullanılan test araçlarının daha işlevsel, güvenilir ve geçerli olması için oldukça fazla katkı sunmuştur. Dil öğretiminde kullanılan ölçme ve değerlendirme araçlarının şu anki seviyede olmalarının yabancı dil öğretiminde yaşanan yeniliklerin ve sürekli devam eden kuramsal ve deneysel çalışmaların sonucu olduğu söylenebilir.

Tarihsel süreç içinde yaşanan değişimler incelendiğinde, dil öğretim yöntembilimindeki gelişmelerin ölçme değerlendirme anlayışına ve uygulamalarına yön verdiği görülmektedir. Örneğin; 1950'lerde özellikle davranışçılık kuramının ve karşılaştırmalı analizlerin ilgi odağı olduğu dönemde, ölçme ve değerlendirme araçları ana dil ile hedef dil arasındaki sesbilim, dilbilgisi ve kelime bilgisi gibi dil yapılarına odaklanmaktaydı (Brown, 2004). 1970 ve 1980'lerde ise iletişimsel dil teorileri test sürecine daha etkileşimsel bir boyut kazandırmış ve "bütün bir iletişimsel olayın onu

oluşturan dil öğelerinin toplamından çok daha önemli olduğu" görüşü test süreçlerine yansımıştır (Clark, 1983: 432). Dolayısıyla yabancı dil öğretim kuramlarına dair yaşanan gelişmelerin ölçme değerlendirme faaliyetlerini etkilediği söylenebilir.

Spolsky (1978), ikinci dil öğretiminde test süreçlerinin gelişimini genel olarak üç döneme ayırmıştır: bilimsel olmayan dönem (1950'lerden önce), psikometrik-yapısalcı dönem (1950'lerden 1960'ların sonu) ve psikolinguistik-sosyolinguistik dönem (1960'ların sonu ve takiben). Diğer taraftan Morrow (1981) bu dönemleri Cennet Bahçesi (The Garden of Eden), Fani Dünya (The Vale of Tears) ve Vadedilmiş Topraklar (The Promised Land) olarak adlandırmıştır. Brown (1996) ise "term (dönem)" yerine "movement (hareket)" kavramını kullanmayı tercih ederek, bu hareketlerin kronolojik olarak birbirini kapsayabileceğine ve bugün bile farklı yerlerde aynı anda bulunabileceklerine dikkat çekmiştir. Dolayısıyla tarihsel olarak keskin bir ayrım yapmak yerine dil testlerinin hazırlanma ve uygulama süreçlerini göz önünde bulundurarak bir sınıflama yapmıştır. Bu test hareketleri; Discrete-Point Testing (Parça-Birimli Sınav), Integrative Testing (Bütünleştirici Sınav) ve Communicative Language Testing (İletişimsel Dil Sınavı) şeklindedir.

2.3.1. Bilimsel Olmayan Dönem

1900'lü yılların başından ikinci dünya savaşına kadar geçen bu dönemde İngilizce öğretiminde ön planda olan Dilbilgisi-Çeviri (Grammar-Translation) yönteminin test süreçlerini yakından etkilediği görülmektedir. Bu dönem öğretim ve test süreçlerinin öğretmenler ve test uygulayıcılarının kişisel yargılarına bağlı olmasından dolayı Madsen (1983) tarafından "sezgisel dönem" olarak da adlandırılmıştır. Madsen'e (1983) göre bilimsel olmayan dönemde dil öğrenimi, bir takım dilbilgisi ve sözcük bilgisi kurallarının öğrenildiği ve bu kuralların yazma ve çeviri aktivitelerinde kullanıldığı bir süreçti. Yabancı dildeki testler ise, öğrenenlerin hedef dildeki dilbilgisi ve kelime bilgilerini yazılı metinlerle ne derece kullanabildiklerini ölçmekteydi.

Bilimsel olmayan dönemde kullanılan tüm testler genellikle herhangi bir ilkeye bağlı olmaksızın sezgisel bir şekilde hazırlanmış ve uygulanmıştır (Alduais, 2012). Öğrenenlere tümdengelimsel bir şekilde öğretilen dilbilgisi kurallarını ve kelimelerin ezberlenmesini gerektiren testler sunulmuştur (Giri, 2003). Dilbilgisi-Çeviri (Grammar-Translation) yöntemiyle uyumlu olarak okuma temelli öğretim yöntemleri kullanılmıştır. Bu yöntemler temel olarak öğrencilere çevirmeleri için verilen oldukça uzun metinlerin ya

da seçilen dilbilgisi konuları ve çeşitli kültürel öğeler üzerine yapılan alıştırmaların öğrencilere hedef dili kazandırmada yeterli ve geçerli olduğunu kabul etmektedir (Clark, 1983).

Bilimsel olmayan dönemdeki dil testleri, öğrenenlerin sözdizimsel olarak kabul edilir cümleler kurmaları ve kelime bilgisi ile dil kurallarını sentezlemelerini gerektiren kompozisyon ve yazma aktiviteleriyle hedef dildeki yeteneklerini genel olarak değerlendirmeyi amaçlamaktadır (Ingram, 1985). Dilbilgisi yeterliğini arttırmak için okuma ve anlama temelli çeviri, makale, mektup ve açık uçlu sorular gibi yazma etkinlikleri oldukça yoğun bir şekilde organize edilmiştir. Test uygulayıcıları öğrenenlerin bu yazma etkinliklerindeki dilbilgisi ve kelime bilgisine dayanarak yabancı dildeki yeteneklerine dair bir yargıda bulunmuşlardır (Madsen, 1983).

Bu dönemin en önemli özelliklerinden biri de test hazırlamada herhangi özel bir becerinin veya uzmanlığın gerekmediği görüşüdür (Clark, 1983). Öğretmenler kendi testlerini hazırlama ve değerlendirme süreçlerinde bağımsız hareket etmişlerdir. Aslında dil öğretmek için yeterli olan herhangi bir birey, ölçme ve değerlendirme sürecini kontrol etmek için de tam olarak yeterli görülmüştür (Clark, 1983). Sujana'a (2016) göre, bu tür testlerle ilgili en önemli problem, objektif bir şekilde puanlanmasının zor olmasıdır. Dolayısıyla bu dönemde hazırlanan testlerin, öğrencilerin hedef dili gerçek hayatta işlevsel bir şekilde kullanmaları için gereken dil becerilerini kazandırmaktan yoksun olduğu kabul edilmiştir (Madsen, 1983; Brown, 1996; McNamara, 2000 and Weir, 2005). Ayrıca bu dönemde dil öğretiminde kullanılan ölçme araçlarının geçerlik, güvenirlik ve diğer önemli psikometrik özelliklerine genel olarak çok az önem verilmiştir (Howard, 1980).

2.3.2. Psikometrik-Yapısalcı Dönem

Bilimsel olmayan dönemde kullanılan testlerin güvenirlik, geçerlik ve objektiflik gibi ilkeler açısından hem öğrenenler hem de öğretenler için tatmin edici olmaması psikometrik-yapısalcı akımın ortaya çıkmasına temel olmuştur. 1950'lerde başlayan bu yeni dönemde dil testleri psikoloji ve yapısalcı dilbilim alanındaki çalışmalardan oldukça etkilenmiştir. Madsen'ın (1983) "bilimsel test süreçlerinin" başladığı dönem olarak kabul ettiği bu akımla; dilbilimciler ile psikoloji ve eğitim alanındaki uzmanların ortak çalışmaları sonucu bilimsel ilkelere daha uygun test teknikleri kullanılmaya başlanmıştır. Bu dönemin en önemli karakteristik özelliği, psikometrik ölçümlerin ve yapısalcı dilbilim

ilkelerinin bilimsel olmayan döneme kıyasla uygulamada daha fazla yer bulmasıdır (Lado, 1961).

1960'lar boyunca özellikle Skinner'ın çalışmalarıyla ortaya çıkan davranışçı dil öğrenme teorileri ile Fries ve Bloomfield gibi yapısalcı dilbilimcilerin çalışmaları bu dönemin temel çıkış noktaları olmuştur. Skinner'ın ortaya koyduğu bu yaklaşıma göre dil, uyarıcı-tepki bağıyla alışkanlık oluşturma sürecidir (Ingram, 1985). Etki-tepki ilişkisinin kurulması için ise pekiştireçlerin kullanılması gerekmektedir. Olumlu pekiştireçlerin verilmesi ile birlikte davranışlar tekrar edilir ve alışkanlığa dönüşür (Demirel, 2004). Yapısalcı dilbilimcilere göre ise dil sözel olarak gerçekleştirilen bir davranıştır; küçük parçalara ayrıştırılabilir, karşılaştırılabilir ve tekrar bir bütün haline dönüştürülebilir (Brown, 2007; Chen, 2005; Richards & Rodgers, 2001). Dil öğrenimi bu bağlamda değerlendirildiğinde; dilin temel yapılarından olan sesletim, dilbilgisi ve sözcük yapılarını bilinçli bir farkındalıkla edinme süreci olarak tanımlanır. Bu yapılar üzerinde üst düzeyde yeterli olmanın yolu da tekrar ve uygulama aşamalarının sürekliliğidir (Carroll, 1983).

Psikometrik-yapısalcı dönemde dil öğretiminde yeni yaklaşımlar benimsenmiştir. Bu yaklaşımlardan en çok vurgulananlar İşitsel-Dilsel (Audolingual), Sözel-İşitsel (Oral-Aural) ve Sözel-Yapısal (Oral-Structural) yöntemlerdir. Bu yapısalcı yöntemler dil öğretimini bir takım dil yapılarının edinim süreci olarak kabul etmektedir. Dil yeteneği ise, dilin birbirinden bağımsız öğeleri üzerinde yetkin olma durumudur (Giri, 2003). Dönemin benimsenen yaklaşımları doğrultusunda, test edilecek yabancı dil bilgisinin parçalara ayrılması ve bağlamından ayrı bir şekilde incelenmesi yönünde bir eğilim vardı (McNamara, 2000). Bu fikrin öncülerinden olan Lado (1961), bir dilin karmaşık yapısının, o dilin birbirinden bağımsız birimlere ayrılarak öğrenilebileceğini savunmuştur. Ancak bu şekilde, neyin test edileceği ve bu test süreçlerinin nasıl gerçekleştirilmesi gerektiğine dair bir yol haritası oluşturulabilir. Bu düşünceye göre, dinleme, konuşma, okuma ve yazma gibi her dil becerisi ve sesletim, dilbilgisi ve sözcük bilgisi gibi her dil alt beceri ya da parçası birbirinden bağımsız bir şekilde test edilmelidir. Böylece tüm dil becerileri ve parçaları uygun bir şekilde ölçülebilir olacaktır (McNamara, 2000; Weir, 2005).

Psikometrik-yapısalcı dönem Brown (2005) tarafından Ayrık-nitelikli (Discrete-Point Testing) test hareketleri olarak adlandırılmıştır. Ona göre bu test yaklaşımları bugün bile varlığını sürdürmektedir. Çoktan seçmeli testler bu dönemin en çok kullanılan testleri olmuştur. Öğrenenlerin sadece herhangi bir beceriye ait kısmi bilgisi ölçülmüştür. Yani her test sadece bir dilbilgisi kuralını, sözcük bilgisini ya da diğer becerilerin herhangi birini

ayrı ayrı bağlamlarda ölçmeyle sınırlandırılmışlardı (Lado, 1961). Weir'e (1988) göre dil bağlamından kopuk, soyut bir şekilde algılanmamalıdır. Bu tür bir anlayış dil testlerini oldukça yapay, gerçek hayat durumlarıyla ilgisi olmayan ve sadece kuramsal bilgileri ölçen bir seviyeye indirgemektedir. Ancak sadece kuramsal olarak yeterli olmak, tek başına iletişimsel becerilerin bir göstergesi değildir. Örneğin, bir ehliyet ya da müzik sınavına girecek biri için önemli olan, yeteneğini bir bütün olarak göstermesidir. Sadece motor ya da org bilgisine sahip olmak, iyi bir sürücü ya da sanatçı olmak için yeterli değildir (Weir, 1988).

Bilimsel olmayan dönemde kullanılan testlerin aksine, bu dönemde kullanılan testlerin en önemli özelliği güvenirlik ve geçerlik kavramlarına daha fazla ilgi gösterilmesidir. Lado (1961) tarafından öne sürülen "objektiflik" kavramı psikometrik analizlerin temelini oluşturmuştur. Lado'ya (1961) göre bu dönemde yapılan testlerin hem kapalı-uçlu ve tek doğru yanıtın arandığı testler olması hem de ölçütlerin çok net olması sebebiyle, yapılan değerlendirmeler oldukça nesnel ve güvenilirdir. Ayrıca bu dönemde psikometrik ölçümler ve yapısal dil ilkelerinin uygulamaları göze çarpmaktadır. İstatistiksel analizlerle tanışılmış ve TOEFL (Test of English as a Foreign Language) (1963) gibi ilk defa dikkatli bir şekilde tasarlanmış standart testler yayımlanmıştır. Bu tür testler objektif bir şekilde tasarlanmış ve güvenirlik, geçerlik özelliklerinin sağlanması için içerik, puanlama ve uygulama süreçlerine dikkat edilmiştir (Sujana, 2016). Bilimsel test süreçlerini yeteri kadar objektif bulmayan Spolsky'a (2000) göre ise, aslında objektiflik kavramı ortaya atılsa da bunun nasıl sağlanacağı yıllarca problem oluşturmuştur. Bu düşünceyi destekleyen Pilliner (1968), sadece niceliksel olarak ortaya çıkan sonuçlar ve belki de en önemlisi testin kendisinin nasıl oluşturulduğu düşünüldüğünde, objektif olmayan faktörlerin değerlendirme sürecine büyük ölçüde karıştığını ileri sürmüştür. Diğer taraftan geçerlik faktörü açısından incelendiğinde ise, bir testin taşıması gereken görünüş, yapı ve kapsam geçerliği gibi özellikler Davies (1968) tarafından bu dönemde tanımlanmıştır.

Psikometrik-yapısalcı döneme ait yaklaşımlar ve testler savundukları ilkeler itibariyle oldukça fazla eleştiri almıştır. Dil öğreniminin en önemli parçası olan iletişimi yeteri kadar sağlayamadığı görüşü, bu eleştirilerin en büyük kaynağı olmuştur. Alderson ve Hughes'a (1981) göre eğer öğrenenler dil öğelerine dair sahip oldukları bilgileri yeni durumlarla karşılaştıklarında uygun bir şekilde kullanamıyor ve istenilen performansı sergileyemiyorlarsa öğrendiklerinin işlevsel olduğu söylenemez. Bütünün onu oluşturan

parçalardan daha büyük ve anlamlı olduğu noktasında uzlaşan Davies (1968) ve Ingram (1985), bu dönemin temel hipotezi olan dilin öğelere ayrılıp bağlamından ayrı değerlendirilmesine karşı çıkmışlardır. Oller (1979) benzer bir yaklaşım sergileyerek, dilin onu oluşturan elementlerine ayrıldığında önemli özelliklerinin kaybolduğuna ve bu elementlerin daha farklı iletişimsel bir ortamda birlikte kullanımına ya hiç ya da çok az önem verildiğine dikkat çekmiştir.

2.3.3. Bütünlükçü/ Sosyo-Dilbilimsel Dönem

Dil öğretimi ve test süreçlerine etki eden bütünlükçü yaklaşımın ortaya çıkması iki temel dilbilim akımının çağdaş dilbilimciler tarafından vurgulanmasıyla başlamıştır. Bunlardan ilki, dile dair bilginin onu oluşturan parçaların toplamından daha fazla olduğu fikridir (Spolsky, 1975; Oller, 1979). İkincisi ise modern sosyo-dilbilimsel görüşlerle ilgili olan iletişimsel dil öğretim yaklaşımlarıdır. Hymes'a (1967) göre, dil öğrenenler sadece dilbilgisel açıdan doğru cümleler üretebilmekle kalmamalı aynı zamanda farklı sosyal durumlara uygun yapılar kullanabilmelilerdir. Bu görüşler göz önünde bulundurulduğunda bütünlükçü düşüncenin psikometrik-yapısalcı döneme dair yapılan eleştirilerden ortaya çıktığı savunulabilir. Yani bilimsel dönemin başladığı kabul edilen psikometrik-yapısalcı dönemde dil becerilerinin bağlamından kopuk ve ayrı ayrı değerlendirilmesine yönelik bir anlayış hakimken, bütünlükçü dönemde bu becerilerin parçalara ayrılmak yerine bir bütün halinde değerlendirilmesi söz konusu olmuştur.

Holistik (Bütünsel) yaklaşım olarak da adlandırılan bu dönemde, dil yeteneğinin birbirinden bağımsız bir şekilde test edilebilir parçalara ayrılıp ayrılamayacağı üzerine yapılan tartışmalar Oller'ın (1979) "*Birleşik Özellik Hipotezi (Unitary Trait Hypothesis)"* nin ortaya çıkmasına öncülük etmiştir. Bu yaklaşıma göre, dil yeterliği ayrık nitelikli testlerin öne sürdüğünden daha bütünseldir. Yani, sözcük bilgisi, dilbilgisi, sesletim, dört temel dil becerisi (dinleme, konuşma, okuma ve yazma) ve dilin diğer ayrı parçaları birbirinden ayrı düşünülemez. Bu beceriler birbiriyle iç içedir ve onlar arasında bir ayrım yapmak yerine genel dil yeterliğini göz önünde bulundurmak daha doğru olacaktır. Bu yaklaşım, dili öğelere indirgemeden önce bir bütün olarak incelenmesi gerektiği savunmaktadır. Aksi takdirde, dil öğrenenlerin dili tam olarak anlayamayacakları savunulmuştur (Brockman, 1994; Rupp, 1986).

Dil öğrenimini çevresel faktörler ile şekillenen dışsal bir mekanik süreç olarak algılayan davranışçı ve yapısalcı yaklaşımların aksine içsel ve soyut zihinsel süreçlerin bir

yansıması olarak gören bilişsel psikoloji ve bilişsel öğrenme teorileri, bu dönemde kullanılan dil öğretim yöntemlerinin ve testlerin temel felsefesini oluşturmuştur. Davranışçı kuramcılar dil performansının yalnızca gözlemlenebilir kısmına odaklanırken, Chomsky dil öğrenmenin bir uyarıcıya tepki vermenin ve davranışın taklit edilmesinin ötesinde daha karmaşık bir süreç olduğuna dikkat çekmiştir. Ona göre, dil birimlerindeki yüzey yapı ile derin yapının birbirine bağlanmasıyla bellekte anlam oluşur. Derin yapının yüzey yapıya bağlanması çeşitli zihinsel işlemlerle gerçekleşir. Söz konusu zihinsel işlemleri gerçekleştirmek için de tüm insanlar zihinlerinde bir dil öğrenme aygıtına sahip olarak dünyaya gelirler. Bu aygıtı, özellikle dil öğrenmek için kurulmuş zihinsel bir yetenek gibi algılamak gerekmektedir (Chomsky, 2001). Chomsky'nin bu teorisi farklı dil öğretim şemalarının geliştirilmesine ve dil öğrenmenin bilinçli bir içselleştirme süreci olduğuna dair fikirlere öncülük etmiştir. Bu bilişsel teoriler ışığında öğretim yöntemlerinin değişmesi test süreçlerinin de değişimini zorunlu kılmıştır. Nitekim yabancı dilde ölçme vurgusu dilbilimsel doğruluktan işlevsel yeteneği kullanmaya dönük bir değişim yaşamıştır. Bu doğrultuda dil testleri problem çözme yaklaşımlarına odaklanmış ve öğrenenlerden içselleştirdikleri kuralları açığa vurmaları beklenmiştir (Corder, 1967).

Bu dönemi yapılandıran bir diğer önemli olgu da sosyo-dilbilimsel ilkelerin sorgulanmaya başlamasıdır. Sosyo-dilbilimsel oluşum Chomsky'nin bilişsel teorisini genişleterek, öğrenenlerin dil becerilerini iletişimsel bir şekilde ve farklı bağlamlara uygun olarak kullanma yeteneklerinin test edilmesi gerektiğine dayanır. Bu anlayışın savunucularından olan Hymes (1972), bir mesajın kültür ve sosyal kurallar gibi bağlamının dilbilimsel bağlamı kadar önemli olduğuna ve mesajın sosyo-kültürel işaretlerinin gözden kaçırılmasının yanlış anlaşılmalara yol açabileceğine dikkat çekmiştir. Giri'ye (2003) göre bu doğrultuda hazırlanan dil testleri sadece öğrenenlerin dil becerilerine yönelik bilgisini değil aynı zamanda durumsal ve bağlamsal uygunluk göz önünde bulundurularak cümleleri anlama ve üretme yeteneklerini ölçmelidir.

Tarihsel olarak bütüncül döneme ait olduğu kabul edilen iki test türü vardır: Çıkartmalı (cloze) ve dikte (dictation). Çıkartmalı test içinde belirli aralıklarla boşluklar olan ve uygulayıcıların bu boşluklara uygun olan kelimeleri yerleştirmelerini gerektiren okuma parçalarıdır. Dikte ise öğrenenlerin test uygulayıcıları tarafından sesli bir şekilde okunan metni yazım kurallarına uygun bir şekilde yazma etkinliğidir. Sujana'ya (2016) göre, bu iki test tekniğinin tercih edilmesinin en önemli sebepleri, öğrencilerin dil yeterliğini değerlendirirken hem dil becerilerini birleştirme olanağı sunması hem de belirli

bir bağlam içinde dilin kullanımını gerektirmesidir. Böylece psikometrik-yapısalcı dönemin tek bir beceriyle sınırlandırdığı ölçme kavramının ötesine geçilmiştir.

Oller (1979) çıkartmalı ve dikte testlerinin birkaç beceriyi aynı andan ölçebildiğine dikkat çekerek, öğrenenlerin genel dil yeterliğine dair değerlendirme yapmada uygun sonuçlar sağladığını ileri sürmüştür. Weir'e (1988) göre ise, bu tür test teknikleri iletişimsel yeterliği test etmek için yeterli olmasa da öğrenenlerin kabiliyetlerini ve iletişimsel yeteneklerine dair sahip oldukları potansiyeli ortaya çıkarmak için iyi bir kılavuzdur. Ayrıca bu testlerin hazırlanmasının ve uygulanmasının ekonomikliğine ve kayda değer bir derecede güvenilir olduklarına dikkat çekmiştir.

Çıkartmalı ve dikte gibi test tekniklerini dil becerilerini ölçmede yeterli gören bütüncül yaklaşım bazı karşıt görüşlerin de odağı olmuştur. Davies (1978) çıkartmalı testlerin genel dil yeterliğini ölçmekten çok sadece anlama becerilerini ölçmeye dayalı olduğunu ve öğrenenlerin yabancı dil yeterliğiyle ilgili bir yargıya varmak için güverlik ve geçerliğinin düşük olduğunu savunmuştur. Bu görüşü destekleyen Alderson (1979), bu tür testlerin işlevsel bilgiler yerine kuramsal bilgileri test eden bir araç olduğunu eklemiştir.

Bütünlükçü ya da sosyo-dilbilimsel dönemde gelinen noktayı Davies (1978) şu şekilde özetlemiştir: Birincisi, dilin özü itibariyle birbirinden bağımsız parçalar olmadığıdır. Bu yüzden mutlaka bütünsel bir şekilde öğretilmeli ve test edilmelidir. İkincisi ise, dil kullanımı amaçlı olmalıdır. Bu amaç da her zaman iletişimsel olmalıdır. Dolayısıyla test edilmesi gereken şey kavramsal bilgiler değil iletişimsel yeterliktir.

2.3.4. İletişimsel Yaklaşımlar Doğrultusunda Ölçme Değerlendirme

1980'li yıllardan itibaren dil öğretiminde bilişsel ve yapılandırmacı öğrenim kuramlarına dayanan *İletişimsel Dil Öğretim Yöntemi* (Communicative Language Teaching Method) benimsenmiş ve yeni bir dönem açılmıştır. Şimdiye değin dil öğrenim sürecinde dilbilgisi yapılarının kusursuz bir şekilde kazanılmasını hedefleyen yapısalcı yaklaşımların yerini iletişim odaklı etkinliklerle dil öğreniminin gerçekleşmesini amaçlayan bir yaklaşım almıştır (Paker, 2015).

İletişimsel Dil Öğretim Yönteminin ortaya çıkışı sosyal dilbilimci Hymes (1971) tarafından ortaya atılan "iletişimsel yeti" kavramına dayanmaktadır. Bu kavram daha sonra Savignon (1987) tarafından iletişimsel süreçlerde sosyal dilbilimsel ve söylemsel yetiler ile dilbilgisi kombinasyonun başarılı bir şekilde kullanımı olarak yeniden tanımlanmıştır. Littlewood (1983) "iletişimsel" kavramına dikkat çekerek, bu kavramın dil öğretim

sürecinde kazandırılması beklenen hedefleri işaret ettiğini belirtmiştir. Ona göre iletişimsel olarak nitelendirilebilecek ve tüm sınırların kesin bir şekilde belirlendiği bir öğretim sistemi yoktur. İletişimsel dil öğretimi terimi, iletişimsel becerilerin geliştirilmesi için belirlenen hedeflere ulaştıran bir yöntem ya da yöntemler karmasıdır. Richards ve Rodgers'a (2002) göre ise iletişimsel dil öğretimi; iletişimsel becerileri dil öğretiminin hedef noktası haline getirmeye çalışan, dilin ve iletişimin birbiriyle ayrılmaz ve ilişkili olduğu ilkesiyle dört temel dil becerisinin (dinleme, okuma, konuşma, yazma) öğretimini aktif süreçlerle geliştirmeyi amaç edinen bir yaklaşımdır.

İletişimsel Dil Öğretimi daha çok temel ilkeleri ve özellikleri ile tanımlanmaktadır. Nunan (1991) bu öğretim yönteminin beş temel özelliğini aşağıdaki şekilde özetlemiştir. Ona göre İletişimsel Dil Öğretimi:

- hedef dilde etkileşim yoluyla iletişim kurmayı öğrenmeye vurgu yapılması,
- otantik metinlerin öğrenme durumlarında kullanılması,
- öğrenenlerin kendi dil öğrenme süreçlerine odaklanma fırsatı sunması,
- öğrenenlerin kendi kişisel deneyimlerinin sınıftaki öğrenmelere katkı sunan unsurlar olması,
- sınıf ve sınıf dışı dil öğrenme faaliyetleri arasında bağlantı kurma çabasıdır.

Nunan'ın (1991) sıraladığı bu özellikler, İletişimsel Dil Öğretim Yönetiminin öğrenenlerin istek ve ihtiyaçlarını göz önünde bulundurmasının yanı sıra sınıfta öğretilen dil ile dışarıda kullanılan dil arasında bağlantı kurulmasının önemini vurguladığını göstermektedir. Bu geniş tanımlamaya göre öğrencilerin iletişimsel yetilerini otantik süreçlerle geliştirmelerine yardım eden her türlü eğitsel faaliyetin geçerli ve yararlı olarak kabul edildiği düşünülebilir. Oiler'a (1979) göre bu yöntemde ölçme ve değerlendirme adına ortaya atılan en büyük söylem ise dil yeteneğinin birbiriyle bağlantılı etkileşimsel süreçlerden oluşan bir yapı olduğu ve bu yapıların ayrı ayrı test edilemeyeceğidir.

Dil öğretimindeki bu yenilik, dil testlerinin de yeniden gözden geçirilmesine öncülük etmiştir. Uzmanlar iletişimsel dil yaklaşımının öğretim ilkelerini gerçekleştirecek en etkili ve en uygun ölçme değerlendirme yöntem ve araçlarını bulmaya çalışmışlardır. Brindley'e (1986) göre, iletişimsel dil yaklaşımının benimsenmesiyle öğrenenlerin belirli dil öğelerine dair bilgilerini ölçen geleneksel yöntemlerin yerini, dili gerçek durumlarda etkili bir şekilde kullanabilme yeteneğini test eden ölçme ve değerlendirme yöntemleri almıştır.

İletişimsel anlamda dil bilmenin ne olduğuna dair pek çok model öne sürülmüştür. Bu modellerin en önemli özelliği, iletişimsel dil testlerinin ne olduğuna ve nasıl olmasına gerektiğine dair ilkeler ortaya koyarak referans noktası olmalarıdır. Rea'ya (1985) göre, alanyazında iletişimsel yeteneğin ölçülebilir bir süreç haline nasıl gelebileceği ile ilgili pek çok farklı görüş bulunmaktadır. Bu karmaşıklığın üstesinden gelmeyi amaçlayan Canale ve Swain (1980), Hymes'ın (1972) "iletişimsel yeti" kavramına daha geniş bir perspektif sunmuşlardır. Onlara göre iletişimsel beceri dilbilgisi yeteneği, sosyo-linguistik yetenek, stratejik yetenek ve söylem yeteneği olmak üzere dört farklı bilgi alanını kapsamalıdır. Bunlardan ilki olan dilbilgisi yeteneği; sözcük, anlam, yapı ve sesletim bilgisi ile sözdizimi kurallarını içermektedir. İkinci olarak sosyo-linguistik yetenek, sosyo-kültürel kuralların kullanım bilgisinden oluşmaktadır. Stratejik yetenek ise iletişimsel stratejileri kullanabilme, iletişim sürecinde meydana gelebilecek hatalarla baş edebilme yeterliğini kapsamaktadır. Son olarak 1983 yılında yapılan eklemeyle kendine yer edinen söylem yeteneği; dilbilgisi formlarının ve anlamın farklı yazın türlerinde birleştirilerek tutarlı bir içeriğin nasıl oluşturulabileceğine odaklanmaktadır.

1990'lı yılların başında Morrow (1979) ve Bachman (1990) tarafından öne sürülen yabancı dil testi modelleri, test sürecinin hem yetenek hem de performansa dayalı olması gerektiği fikrini ortaya koymuştur. Onlara göre iletişimsel bir dil testi, hem bir öğrenenin yabancı dili nasıl öğrendiğine dair bilgisini hem de dile dair bu bilgisini ne derece yerine getirebildiğini ölçmelidir. Yetenek ve performans arasındaki teorik ilişkiye dair daha belirgin bir tablo oluşturan Canale ve Swain (1980) ise, kullanılan test tekniklerinin öğrenenlerin sadece dilin yapısına dair bilgilerine değil, aynı zamanda kişinin bu bilgiyi anlamlı ve en önemlisi iletişimsel bir şekilde ne derece gösterebildiğine odaklanması gerektiğini öne sürmüşlerdir.

Yabancı dil bilgisinin nasıl ölçülmesi gerektiğine dair önemli bulgular sunan Bachman ve Palmer'a (1996) göre, iletişimsel bir dil testinin dört önemli boyutu vardır. Bunlar: içeriğin özgün olması, materyallerin ile test araçlarının otantikliği ve gerçek yaşam durumlarının test süreçlerine yansıtılmasıdır. Ayrıca bu testler tüm dilbilimsel, sosyo-dilbilimsel, söylem ve stratejik yetenekleri birlikte harekete geçirebilmelidir. Katsumasa (1997), iletişimsel dil testlerinin içeriği ile ilgili benzer ilkeleri tartışmış ve bu boyutlara yeni maddeler eklemiştir. Ona göre yabancı dilde etkileşimsel bir ölçme ve değerlendirme süreci oluşturmak için; neyin ölçülüyor olduğunun bilinmesi, iletişimsel süreçlere odaklanılması, yapılan önemsiz dilbilgisi hatalarının iletişimi çoğu zaman kısıtlamadığının

farkında olunması ve mükemmel bir dilbilgisi yetisinin başarılı bir iletişim için tek geçerli araç olmadığının bilinmesi gerekmektedir.

İletişimsel Dil Öğretim Yönteminde ara sınav, son sınav, kısa sınav (quiz) gibi geleneksel yöntemlerin yanında ödev, proje, sunum gibi sınıf içi performansa dayalı alternatif ölçme ve değerlendirme teknikleri kullanılmaktadır (Paker, 2013). Brown (2004), performansa dayalı değerlendirmenin öğrencilerin kağıt kalemle birbirinden bağımsız sorulara cevap vermek yerine; sözlü ve yazılı üretimlerinin, açık uçlu soruların, grup performansının ve diğer etkileşimsel görevlerin ölçülmesine olanak sağlandığına dikkat çekmiştir. Ona göre böylesi bir değerlendirme uygulaması pahalı ve zaman alıcı gibi düşünülebilir ancak öğrenciler mevcut ya da temsili gerçek hayat görevlerini yerine getirdiklerinden, bu çabalar diğer doğrudan ölçme tekniklerinden daha fazla sonuç verecektir. Daha teknik anlamda söylemek gerekirse, öğrenenler hedeflenen dilbilimsel becerileri üretme sürecinde test edileceklerinden daha yüksek kapsam geçerliği sağlanacaktır.

2.4. YABANCI DİL BECERİLERİNİN DEĞERLENDİRİLMESİ

Yabancı dil öğretim yaklaşımlarının geleneksel bakış açısından iletişimsel bir sürece doğru evrilmesinin dil testlerini de oldukça yakından etkilediği söylenebilir. Araştırmanın bu bölümünde yabancı dil becerilerine ayrı ayrı odaklanılarak bu alanların değerlendirilmesi sürecinde ele alınan yaklaşımlar ve yöntemler tarihsel süreç göz önünde bulundurularak incelenmiştir.

2.4.1. Dinleme Becerilerinin Değerlendirilmesi

2.4.1.1. Dinleme

Bireysel ve toplumsal etkileşimin en önemli parçalarından biri olan dinleme, pasif bir eylemmiş gibi görünse de aslında bireylerin anlama, yorumlama ve değerlendirme gibi zihinsel mekanizmalarını oldukça aktif bir şekilde işlemelerini gerektiren bir süreçtir. Flowerdew ve Miller'a (2005) göre, tüm çocuklar duyabilme yetisiyle doğarlar. İlk önce dinlerler ve sonra konuşmaya başlarlar. Ardından okuma ve yazma becerilerini edinirler. Yani tüm dil becerileri arasında dinleme ilk ortaya çıkan beceri alanıdır. Dolayısıyla dinleme, konuşma da dahil olmak üzere diğer tüm dil becerilerinin kazanılmasının ön koşulu niteliğindedir (Nunan, 1998).

Duymayı seslerin algılanması, dinlemeyi ise seslerle anlam arasında bağlantı kurmak olarak tanımlayarak Kline (1996), bu iki kavram arasındaki farklılığa dikkat çekmiştir. Benzer farklılığı vurgulayan Rost'a (2002) göre, duyma bir algılama biçimidir. Dinleme ise sadece duymayı değil, aynı zamanda bellekteki bilginin yapılandırılması ve yorumlanması süreçlerini de içeren oldukça karmaşık, aktif ve kasıtlı bir süreçtir. Dinleme sürecinin ne olduğuna dair doğru tanımlamalar yapılması, ikinci dil olarak bu becerinin kazandırılması ve değerlendirilmesi için birbiriyle tutarlı yöntemler ve araçlar sunulması bakımından oldukça önemlidir.

Günümüzde dinleme becerilerinin sadece günlük hayatta değil, öğrenme ortamlarında da kritik bir role sahip olduğu büyük ölçüde kabul edilmekte ve bu becerinin gelişimini sağlamak için çağdaş yöntem ve araçların kullanılması gerektiğine dikkat çekilmektedir (Rubin, 1994). Şimdilerde yaygın olan bu inanışın aksine, 1960'lı yılların sonuna kadar hem araştırmacılar hem de yabancı dil öğretmenleri okuma ve dilbilgisi aktivitelerine daha fazla önem vermekteydiler. Dil becerilerinin geliştirilmesi ve ölçülmesi ise, dil öğretimi için çok gerekli bir alan olarak görülmemekteydi (Richards ve Rodgers, 2001). İngilizce öğretiminin ilk yıllarında öğrencilerin yapmak zorunda oldukları tek dinleme etkinliği, hedef dilin kurallarını anlatan açıklamaları dinlemekti (Flowerdew ve Miller, 2005).

Dil öğretim yöntemleri açısından tarihsel süreç incelendiğinde, her yaklaşımın dil öğretim süreçleriyle farklı şekillerde ilgilendiği ve dolayısıyla dinleme alanın da her yaklaşımda farklı bir yer edindiği söylenebilir. Örneğin, dil öğretimini belirli kuralların öğretilmesi olarak gören ve öğrencilerin edebi metinleri okumaları ve anlamalarını amaçlayan Dilbilgisi-Çeviri Yönteminde, dinleme becerisinin öğretimi hiçbir zaman gerekli görülen bir konu olmamış ve aslında öğretmenler de bu becerinin öğretilmesi için gerekli donanıma sahip değillermiş (Larsen-Freeman, 2011). Field'a (2009) göre bu dönemde dinleme, temel olarak örnek diyaloglar aracılığıyla yeni dilbilgisi yapılarının öğretilmesini kolaylaştırmak için kullanılmaktaydı.

Dilbilgisi-Çeviri Yönteminden sonra benimsenen Düzevarım (Direct Method) ve İşitsel-Dilsel Yöntemler (Audio-Lingual Method), ikinci dilin doğal bir sıra izlenilerek geliştirilebileceğini öne sürmüşlerdir. Yani öğrenenler anadilde olduğu gibi, sırasıyla dinleme, konuşma, okuma ve son olarak yazma becerilerini öğrenirler (Larsen-Freeman & Anderson, 2011). Şimdiye dek kabul edilen inanışın aksine, bu yöntemlerin savunduğu

temel ilke, diğer tüm becerilerin öğretilmesinden önce dinleme alanının geliştirilmesine odaklanmak gerektiğiydi. Richards ve Rodgers'a (2001) göre hedef dilin kullanımı ön plana çıksa da, dinleme stratejilerini geliştirmek veya bu beceriyi öğretmek ve değerlendirmek için hiç çaba harcanmamıştır. Öğrencilere sadece tekrar etmeleri ve telaffuzlarını geliştirmeleri için dinleme aktiviteleri sunulmuştur. Dinleme öğretimine bu kadar sınırlı yer verilmesi ve sadece diğer beceri alanlarını destekleyen bir araç olarak sunulması, dinleme becerilerinin ölçülmesini de zorlaştırmıştır.

Dil öğretiminde değişen eğilimlere yön veren 2. Uluslararası Uygulamalı Dil Konferansı Birliği'nin (II. International Association of Applied Linguistics Conference) özellikle öğrenenlerin bireyselliğini vurgulaması ve dinlemeyi pasif olmayan oldukça karmaşık bir süreç olarak kabul etmesi yabancı dil eğitimi açısından bir dönem noktası olmuştur (Morley, 2001). Bu süreci takip eden dönemde yabancı dil öğretiminde yaşanan değişikliklerden en çok etkilenen alanlardan biri dinleme olmuştur. Özellikle 1970'lerde İletişimsel Dil Yönteminin benimsenmesiyle, dinleme-anlama alanının iletişim için oldukça kritik olduğuna dair anlayış güç kazanmış ve dinleme, öğretilmesi ve ölçülmesi gereken bir beceri alanı olarak dil öğretim programlarında yer almıştır (Morley, 2001). Rost'a (2002) göre, dinleme alanın bu denli öne çıkmasının sebebi, öğrenenler için girdi (input) sağlaması ve öğrenme için de önemli bir yapı oluşturmasıdır. Çünkü bu girdileri dinleme aracılığıyla doğru algılamak, başta konuşma olmak üzere diğer becerilerin de gelişimini etkilemektedir. O zamandan beri değişen anlayış ile birlikte araştırmacılar arasında dinleme becerilerinin geliştirilmesi ve değerlendirilmesine yönelik bir ilgi olmuştur (Goh, 2002; Vandergrift, 2007).

2.4.1.2. Dinleme Türleri

Yabancı dil öğretiminde, dinleme hedefleri ve ölçütleri belirlenerek öğrenenlere farklı dinleme deneyimlerinin kazandırılması oldukça önemlidir. Bütün etkili testlerde olduğu gibi, dinleme alanında da uygun ölçme araçlarının hazırlanması ve uygulanması için öncelikle amaçların belirlenmesi gerekir. Brown'a (2004) göre dinleme hedeflerinin gerçekleştirmeye yönelik dört temel dinleme türü vardır. Bunlar:

1. *Yoğun Dinleme:* Konuşma dilindeki dil yapılarını ve belirli vurguları saptamak için yapılan dinleme.
2. *Yanıtlayıcı Dinleme:* Konuşulan dilin kısa bir taslağını uygun şekilde yanıtlamak için yapılan dinleme.

3. *Seçici Dinleme:* Yalnızca istenilen bilgiye ulaşmak için yapılan dinleme. Böyle bir dinleme performansının genel amacı aslında dinlenen metnin tüm ve genel anlamını bulmak değil, sadece istenilen bilgiyi aramaktır.
4. *Yaygın Dinleme:* Daha uzun zaman alan dinleme metinlerinin dinlenerek, bütünden bir anlam oluşturmak için yapılan dinleme.

Konuşma dilini algılamak ve karmaşık bir yapıda olan dinlemenin nasıl işlediğini anlamak için Rumelhart ve Ortony (1997) tarafından bir başka sınıflandırma yapılmış ve iki dinleme modeli tanımlanmıştır. Bunlardan ilki aşağıdan-yukarıya (bottom-up) dinleme modeli iken, diğeri yukarıdan-aşağıya dinleme modelidir.

Aşağıdan-yukarıya dinleme modelinde, dinleyiciden dil girdisinin detaylarından başlayarak anlam oluşturması beklenirken; yukarıdan-aşağıya modelde ise ses, sözcük, dilbilgisi gibi öğelere takılmadan mesajın kendisi algılanmaya ve yorumlanmaya çalışılır (Morley, 2001). Nunan'a (1998) göre, aşağıdan-yukarıya işleme modeli, dinleme metninin en küçük anlamlı biriminden tümüne ulaşıncaya kadar doğrusal düzeyde duyduğu sesleri çözümlemesidir. Bu görüşe göre seslerin kelimelerle, kelimelerin cümlelerle, cümlelerin de birbirleriyle doğrusal bir bağ kurmasıyla anlama ulaşılır. Anlam sürecin son aşamasında elde edilir (Nunan, 1998). Diğer taraftan yukarıdan-aşağıya modelde ise; dinleyici gelen sesleri ipucu olarak kullanarak aktif bir şekilde anlam oluşturur. Bunu yaparken de geçmiş bilgi birikiminden oluşan şemalarını kullanır (Hinkel, 2006). Dinleyicinin gelen mesajı anlamlandırırken hem geçmiş hem de dilsel bilgi birikiminden yararlanması, bu iki dinleme modelinin işbirliği içinde çalıştığını göstermektedir (Nunan, 1998).

Dil öğrenimi için özel hedeflerin belirlenmesi, bu alanda ölçülecek kritik davranışların seçilmesi, bu davranışlara uygun ölçütlerin geliştirilmesi ve tutarlı ölçme değerlendirme araçlarının tasarlaması için dinleme türlerinin ve modellerinin bilinmesinin, test uygulayıcılarına fayda sağlayacağı düşünülmektedir.

2.4.1.3. Dinleme Performansının Mikro ve Makro Becerileri

Mikro beceriler aşağıdan-yukarıya dinleme modelinde olduğu gibi dilin daha küçük parçalarına odaklanırken; makro beceriler ise yukardan-aşağıya dinleme modelindeki gibi dilin daha büyük bileşenleriyle ilgilenmektedir (Brown, 2004). Dinleme-anlama becerilerinin geliştirilmesi için belirlenen mikro ve makro beceriler listesi, bir yandan öğrenenlere etkili bir dinleyici olabilmeleri için stratejiler sunarken, diğer taraftan dinleme

becerilerinin ölçülmesi için öğretmenlere çeşitli ölçütler sunan bir rehber niteliğindedir. Dinleme alanı için Richards (1989) tarafından belirlen mikro ve makro beceriler şunlardır:

- ***Mikro Beceriler***

1. İngilizcenin kendine özgü seslerini ayırt etme
2. Dilin farklı uzunluktaki parçalarını kısa süreli hafızada tutma
3. İngilizcedeki vurgu kalıplarının ve tonlamaların farkına varma
4. Kelimelerin kısaltılmış biçimlerinin farkına varma
5. Sözcük sınırlarını ayırt etme, sözdizimi kalıplarını ve önemini yorumlama
6. Farklı tempolardaki konuşmaları ilerletme
7. Konuşmayı duraksamalar, hatalar, düzeltmeler ve diğer performans değişkenleriyle birlikte sürdürme
8. Dilbilgisel kelime sınıfları (isim, fiil, vb.), sistemler (zamanlar, çoğul yapma vb.), kalıplar ve kuralların farkına varma
9. Cümlenin öğelerini tespit etme
10. Özel bir anlamın farklı biçimlerde ifade edilebileceğini anlama
11. Konuşma dilinde kullanılan bağlaçların farkına varma

- ***Makro Beceriler***

12. Durumlara, katılımcılara, amaçlara göre söylemlerin iletişimsel işlevlerinin farkına varma
13. Arka plan bilgisini (schemata) kullanarak durumlardan anlam çıkarma
14. Olaylar arasında bağlantı kurma, sebeplerini ve etkilerini belirleme
15. Gerçek anlam ile mecaz anlamı ayırt etme
16. Anlam oluşturmak için mimik, vücut dilini ve diğer sözel olmayan ipuçlarını kullanma
17. Dinleme stratejilerini kullanma ve geliştirme

2.4.1.4. Dinleme Becerilerini Ölçme Değerlendirme Yöntem ve Teknikleri

Dinleme-anlama alanında amaçları ve ölçülecek davranışları belirleyip, ölçütler geliştirdikten sonra diğer adım öğrencilerin performansını ölçme ve karar verme sürecidir. Bu aşamada hem dil öğrenenlerin öğretim hedeflerini ne derece gerçekleştirdiklerini görmek hem de öğretim sürecinin etkililiğini ortaya çıkarmak için ölçme ve değerlendirme faaliyetlerine ihtiyaç duyulmaktadır. Yaman'a (2015) göre, dinleme becerisinin ölçülmesi kimi zaman zor ve karmaşık bir süreç olarak görülebilir. Ayrıca dinlemenin soyut bir süreç

ve ölçülemez yapıda olduğuna dair yargılar da vardır. Oysa dinleme beceri öğrenenlerden tek başına yanıtlaması beklenen tekniklerle ölçülebileceği gibi, grup içinde veya diğer becerilerle birleştirilmiş şekilde de ölçülebilir.

Dil testlerindeki geleneklerin değişmesiyle birlikte, dinleme becerilerinin ölçülmesine yönelik yöntem ve teknikler de değişmeye başlamıştır. Sujana'ya (2000) göre, ayrık-nitekli (discrete-point) ve dolaylı ölçme gibi dilin belirli dilsel öğelerine odaklanan test süreçleri, öğrenenlerin yabancı dildeki becerilerini ölçmede daha az tercih edilmeye başlanmıştır. Dil öğretiminde değişen bu paradigmalara dikkat çeken Madsen (1983), daha bütüncül ve doğrudan test süreçlerinin uygulanmaya başlamasıyla, dinleme testlerinin daha kapsamlı iletişim becerilerini ölçmeyi amaçladığını belirtmektedir. Bu durumda kapsamlı iletişim becerileri dilin küçük parçalarını anlamaktan daha fazlasını içermekte; konuşmacıların niyetini yorumlamanın yanı sıra karşılıklı fikir alışverişi yapmayı gerektirmektedir (Madsen, 1983).

İletişimsel performansını değerlendirmenin en uygun yolu, öğrencilerin dili gerçek yaşam durumlarına en yakın olacak şekilde kullanabilecekleri öğrenme görevleri yaratmaktır (Brindley, 2001). Bu anlamda seçilecek dinleme metinlerinin öğrenenlerin gerçek yaşamda karşılaşabilecekleri durumları yansıtmalarına dikkat edilmelidir. Buck'a (1988) göre öğrenenlerden bağlamından kopmuş kısa cümleleri çözmelerini istemek ve böyle doğal olmayan aktivitelerle onları meşgul etmek asla gerçek bir ölçme değildir. Eğer öğrenenlerin dilin öğelerine dair bilgi ve yeteneklerini ne derece kullandıklarının yanı sıra bağlamın bütününü nasıl anlamlandıkları ve arka alan bilgileri ile bağlam arasında nasıl bir bağlantı kurduklarına odaklanırsak, o zaman gerçek bir ölçme sürecinden bahsedebiliriz (Lewkowicz, 1991).

Gerçek bir ölçme sürecinden bahsedebilmek için, dinleme amaçlarına hizmet edecek ölçme araçlarının hazırlanması bir diğer önemli konudur. Test uygulayıcıları dinleme becerilerini ölçmek için bir test geliştirirken, ölçme sürecine dahil olan unsurların çok iyi farkında olmaları gerekir. Başka bir deyişle, test geliştirenler konuşulan dilin yapısının farkında olmalılar ki otantik ve iletişimsel dinleme testleri oluşturulabilsin. Brown (2004), bir dinleme testi hazırlarken ilk göz önüne alınması gereken düşüncenin, eğitsel hedefleri tanımlamak olduğunu ifade etmektedir. Ona göre dinleme testlerinin de amaçlarından biri öğrencilerin ya da grupların hedefleri nasıl gerçekleştirdiklerini ölçmektir. Bu hedefler birer rehber gibidir. Mead ve Rubin'e (1985) göre uygun bir ölçme aracı seçme, değerlendirmenin amacına ve hangi ölçme araçlarının mevcut olduğuna

bağlıdır. Eğer amaç bir eğitsel hedefin yeterliğine dair özel becerileri ölçmekse, test de bu becerileri ölçmeye hizmet etmelidir.

Bir ölçme aracı seçerken veya geliştirirken ikinci önemli nokta, eğitsel hedeflerle uyumlu olan muhtemel öğrenme çıktılarını düşünmektir. Yani dinleme testi geliştirenlerin, dinleyiciler tarafından gerçekleştirmeleri beklenen öğrenme çıktılarını tanımlamaları gerekmektedir (Sujana, 2016). Bir sonraki adım konuşma ve metin türü örneklerinin seçimidir. Mead ve Rubin'e (1985) göre, işitsel uyarıcıların tipik bir konuşma dilini temsil etmesi gerekmektedir. Sunulan materyal, öğrencilerin sınıfta, çeşitli kitle iletişim araçlarında ya da konuşmalarda duymaları beklenilen dili örneklemelidir. Dördüncü adım ise, istenilen öğrenme çıktıları için en uygun görevleri ve soru tiplerini seçmektir. Sujana'ya (2016) göre bu noktada önemli olan kriter, eğitsel hedeflere uygun dikkatlice seçilmiş ölçme tekniklerinin kullanılmasıdır.

Dinleme becerilerini ölçmek için kullanılabilecek pek çok teknik vardır. En çok kullanılan tekniklerden biri, öğrencilerin bir ses kaydını dinledikten sonra dinlenilen metin ile ilgili ana fikri bulma ya da detaylı bilgi gerektiren sorulara cevap vermelerini istemektir. Bir başka teknik performans testidir. Performans testiyle öğrencilerden beklenilen yine kayıtlı bir metni dinlemeleri ve duyduklarıyla ilişkili olan resimleri seçmeleri ya da bir rolü canlandırmalarıdır (Decker, 2004). Yaylı'ya (2015) göre, bir başka teknik de öğrencilerin kayıtlı metni dinledikten sonra konuşarak veya yazılı bir şekilde karşılık vermeleri ya da grup içinde tartışmalarının istenmesidir. Böylece birkaç teknik bir arada kullanılarak dil becerilerinin de bütünleştirilmesi sağlanmış olur. Ayrıca öz-değerlendirme, akran değerlendirme, dereceli puanlama araçları ve kontrol listeleri gibi pek çok tamamlayıcı ölçme ve değerlendirme aracı da dinleme becerilerinin ölçülmesi için kullanılabilir. Örneğin; öz-değerlendirme ile öğrenciler kendi dinleme becerilerini değerlendirebilir, öğrenmeye nerden başlayacakları ve nasıl ilerleyeceklerine dair bir plan yapabilirler (Baleghizadeh ve Rahimi, 2011).

Dinleme becerilerinin ölçülmesi için farklı sınıf seviyelerinde kullanılabilecek en yaygın teknikler şunlardır (Thompson, 1995):

Dinleme Türü	Ölçme ve değerlendirme teknikleri	Örnekler
Yoğun Dinleme	➢ Sesbirimsel çiftleri ayırt etme ➢ Biçimsel çiftleri ayırt etme ➢ Vurgu kalıplarını ayırt etme ➢ Tekrarlama	➢ Glass- Grass ➢ Miss- Missed ➢ I can go- I can't go ➢ Kelime tekrarlama
Yanıtlayıcı Dinleme	➢ Sözel sorulara uygun yanıt bulma ➢ Açık uçlu soruları yanıtlama	➢ Çoktan seçmeli testte uygun yanıtı bulma ➢ Sorulara öğrencinin kendi yanıtı beklenir
Seçici Dinleme	➢ Çıkartmalı metni (cloze) dinleme ➢ Resimli bilgi aktarımı ➢ Tablo tamamlama ➢ Cümle tekrarlama	➢ Boşlukları doldurma ➢ Uygun resmi seçme ➢ Haftalık programdaki saatleri yazma ➢ Uyarıcı cümleyi tekrarlama
Yaygın Dinleme	➢ Dikte etme ➢ Diyalog ➢ Konferans dinleme ➢ Yorumlama gerektiren görevler ➢ Hikaye dinleme	➢ Metni dinleyip paragraf yazma ➢ Metni dinledikten sonra çoktan seçmeli ya da açık uçlu soruları yanıtlama ➢ Not alma, özetleme, temel noktaları Listeleme ➢ Şiir dinleme ve anlamını yorumlama ➢ Hikayeyi yeniden anlatma

2.4.2. Konuşma Becerilerinin Değerlendirilmesi

2.4.2.1. Konuşma

Bir beceri olarak konuşma, bireylerin duygu, düşünce ve inançlarını ifade etmelerini sağlayan en önemli araçtır. Burns ve Jayce (1997) konuşmayı, bilginin üretilip işlenerek anlamın oluşturulduğu etkileşimsel bir süreç olarak tanımlanmaktadır. Bu süreç muhatapların sırasıyla konuşmacı ve dinleyici rollerini üstlendikleri ve ayrıca sözel veya sözel olmayan öğelerin işe koşulduğu iletişimsel bir işlemdir. Dolayısıyla dinleme becerisi ile de doğal bir etkileşim içerisindedir. Bütün bu özellikler konuşmayı yabancı dil

eğitiminin de kaçınılmaz bir öğesi yapmaktadır. Ur'a (1992) göre, yabancı dil öğretiminin en önemli amacı öğrencilerin kendilerini hedef dilde sözel olarak ifade edebilmeleridir. Bir dilde yeterli olarak kabul edilmenin en önemli ölçütü de o dili ne kadar iyi konuşabildiğimizdir.

Tarihsel olarak bakıldığında, konuşma becerilerinin öğretilmesi ve test edilmesi ile yabancı dil öğretimi yöntembilimindeki yaklaşımlar arasında çok da sıkı olmayan bir ilişki olduğunu söyleyebiliriz. Örneğin, konuşma Dilbilgisi-Çeviri Yönteminde dikkate değer bir beceri alanı olarak kabul görmemiştir. Kelly'e (1969) göre, bu metodun 1940'lara kadar entelektüel alanlardaki gelişime yönelik yarattığı baskın dayatma, konuşma becerilerinin göz ardı edilmesine sebep olmuştur. Diğer taraftan İşitsel-Dilsel yaklaşımla sözel akıcılık ve ses bilimsel (fonetik) doğruluk vurgulansa da konuşma yeteneğinin yazma becerileri ile test edilmesinden öteye geçilememiştir (Bygate, 2009). Ancak İkinci Dünya Savaşından sonra hedef dilde konuşma becerilerinin değerlendirilmesine yönelik güçlü bir anlayış oluşmuştur (Fulcher & Reiter, 2003). Özellikle de Kaulfers (1944) tarafından askeri amaçlar için geliştirilen sözel testler, anlaşılır bir şekilde iletişim kurmanın gerekliliğini vurgulamaktaydı. Kaulfers'ın askeri durumlar için ikinci dil hazırbulunuşluğunu ölçen sözel akıcılık testleri, istenilen bilgileri sorma ya da yanıtlama, yön belirtme, komut verme gibi geniş kategoriler altında, anadilden hedef dile anahtar cümlelerin çevrilmesini gerektirmekteydi. Daha sonra teste girenlerin, konuşma performansları, sözel ve dil performanslarının niteliğine bakılarak bir derecelendirme ölçeğiyle değerlendirilmekteydi (Kaulfers, 1944)

İkinci dil çalışmalarındaki iletişimsel yaklaşımın etkisiyle, iletişimsel dil testleri odak noktası haline gelmiştir. İletişimsel dil testleri, dil öğrenenlerin gerçek yaşam durumlarında dili nasıl kullandıklarına dair yeteneklerini ölçmeyi amaçlamaktadır. Dolayısıyla bu tür testlerin iletişimsel yeterlilik temelinde dinleme, konuşma, okuma ve yazma olmak üzere dört beceriyi de kapsadığı söylenebilir. Bu bağlamda iletişimsel konuşma becerilerinin ölçülmesi ele alındığında ise, öğrenenlere hedef dilde belirli içeriğe sahip görevler sunularak, konuşma performanslarının test edilmesi amaçlanmaktadır (Roever & McNamara, 2006). Brown (2004), otantik durumların kullanılmasının öğrenenlere hedef dili üretken bir şekilde kullanabilmeleri için fırsat sunduğunu ve iletişimsel yeterliliklerinin daha anlamlı ve verimli bir şekilde ölçmeyi sağladığını belirtmektedir.

2.4.2.2. Konuşma Türleri

Konuşma performansını ölçmeden önce, bu beceri alanına uygun ölçme araçlarının hazırlanması ve uygulanması için dinleme alanında olduğu gibi, öncelikle amaçların belirlenmesi gerekir. Bu amaçları gerçekleştirmek için kullanılabilecek 5 temel dinleme türü vardır (Brown and Abeywickrama, 2010) :

1) *Taklitçi Konuşma:* Bu konuşma türünde öğrenenlerden bir kelime ya da cümleyi aynı şekilde söylemeleri beklenir. Telaffuz değerlendirmenin asıl parçasıdır ancak dil bilgisi de bir değerlendirme ölçütü olarak kabul görür. Öğrenciler ekstra açıklama eklemeksizin verilen bilgiyi sözel olarak tekrar üretirler.
2) *Yoğun Konuşma:* Taklitçi konuşmanın aksine, yoğun konuşma dilin ses-bilimsel öğelerine odaklanmaz. Bu konuşma türü sesli okuma, cümle veya diyalog tamamlama gibi belirli görevleri yerine getirmek için gereklidir.
3) *Yanıtlayıcı Konuşma:* Bir sohbette otantiklik oldukça önemlidir. Bu yüzden konuşmacı hemen cevap vermesi için teşvik edilir. Kısa bir konuşmayı yanıtlamak, basit bir ricada bulunmak, yorum yapmak bu konuşma türüne ait olan aktivitelerdir.
4) *Etkileşimsel Konuşma:* Bu konuşma türünü diğerlerinden ayıran cümlelerin yoğunluğu ve karmaşıklığıdır. Bir sohbette bazen ikiden fazla kişi gerektiği için konuşmacıların sayısı da önemli bir konudur.
5) *Extensive (Kapsamlı) Konuşma:* Dil öğelerine dair oldukça güçlü bilgi ve becerilere sahip olmayı gerektiren, ulaşılabilecek en iyi konuşma türüdür. Kapsamlı konuşma oldukça geniş konuşma üretimini gerektirmektedir.

2.4.2.3. Konuşma Performansının Mikro ve Makro Becerileri

Dinleme alanında olduğu gibi konuşma becerilerinin etkili bir şekilde değerlendirilmesi için mikro ve makro becerilerin bilinmesi, test uygulayıcılarına ölçülecek davranışları belirlemede kolaylık sağlamaktadır. Brown (2004) tarafından belirlenen konuşma sürecinin mikro ve makro becerileri şunlardır:

- ***Mikro Beceriler***

1. İngilizce fonemler ile sesbirim değişkenleri arasındaki farklılıkları üretme.
2. Farklı uzunluklarda cümleler üretme.
3. İngilizce vurgu kalıplarını, vurgulu veya vurgusuz durumdaki kelime durumlarını, ritmik yapıyı ve tonlamayı üretme

4. Kelimelerin ve söz öbeklerinin kısaltılmış biçimlerini üretme
5. Pragmatik amaçları gerçekleştirmek için yeterli sayıda anlamlı kelimeler kullanabilme
6. Farklı hızlarda akıcı konuşmayı üretme
7. Kendi sözel üretimini izleme ve mesajın açıklığını arttırmak için farklı stratejiler kullanma
8. Dil bilgisel kelime sınıflarını (isim, fiil vb.), sistemleri, sözcük düzenini, kalıpları ve kuralları kullanma
9. Doğal yapıda konuşma üretme: duraklama, nefesi ayarlama
10. Belirli bir anlamı farklı dil bilgisel biçimlerde açıklama
11. Sözel söyleme uygun bağlaçlar kullanma

- ***Makro Beceriler***

12. İletişimsel işlevleri durumlara, katılımcılara ve hedeflere göre uygun şekilde tamamlama.
13. Yüz yüze konuşmaya uygun stilleri, sezdirmeleri, pragmatik uygulamaları, sohbet kurallarını ve diğer sosyo-lingustik özellikleri kullanma
14. Olaylar arasında bağlantı ve ilişki kurma
15. Jest, mimik, vücut dili ve diğer sözel olmayan ipuçlarını sözel dille birlikte kullanma
16. Konuşma stratejilerini kullanma ve geliştirme.

2.4.2.4. Konuşma Becerilerini Ölçme Değerlendirme Yöntem ve Teknikleri

Yabancı dil becerilerinin geliştirilmesinde ölçme ve değerlendirme süreci oldukça gereklidir. Öyle ki, bu süreç olmaksızın gösterilecek çabaların eksik kalacağı söylenebilir. Konuşma becerilerinin değerlendirilmesi de diğer beceriler gibi kritik öneme sahiptir. Hughes'a (2003) göre eğer öğretmenler öğrenenlerin sözel yeteneklerini geliştirmek istiyorlarsa o zaman bu yeteneği test etmeleri gerekmektedir. Yani öğrenenler yabancı dil başarılarının göstermiş oldukları sözel performansa bağlı olduğunu bilirlerse, konuşma becerilerini geliştirmek için yeteri kadar ciddi olabileceklerdir (Matin, 2013).

Konuşma becerilerinin değerlendirilmesi, öğrencilerin daha önce sınıf ortamında yaptıkları konuşmaların benzerlerini bir sınav ortamında göstermeleri ve bu konuşmaların önceden belirlenen ölçme-değerlendirme ölçütlerine göre teste edilmesi sürecini kapsar. Daha önce Brown (2004) tarafından belirtilen konuşma yeteneğine ait mikro ve makro

becerilerin birer değerlendirme ölçütü olarak kullanılabileceğinden bahsetmiştik. Nunan'a (1999) göre ise, iletişimsel becerilerin gerektirdiği dilsel, işlevsel, stratejik ve sosyo-kültürel yetenekler konuşma sınavları için ölçüt olarak geliştirilebilir. Bu bağlamda dilbilgisel yetenek, sesleri doğru telaffuz etme, yeterli sözcük dağarcığına sahip olma ile yapısal ve dilbilgisel öğelere hakim olmaktır. Ayrıca konuşma sorulara mantıklı ve tam olarak cevap vermek anlamına gelen işlevsel yeteneği gerektirir. Sohbet anında meydana gelebilecek hataları düzeltebilme stratejik bir yetenektir. Son olarak konuşmacının içinde bulunduğu bağlama uygun şekilde dili kullanması sosyo-kültürel yeteneğin bir gereğidir (Nunan, 1999). Ölçütler geliştirildikten sonraki adımda ise dikkat edilmesi gereken hususları Dağtan (2015) şu şekilde özetlemiştir:

- Hangi konuşma türünün (yoğun, kapsamlı, etkileşimsel vb.) ölçüleceğine karar verilmesi
- Hangi temel öğelere (mekanik öğeler, edimsel beceriler ve sosyo-kültürel kurallar) ağırlık verileceğinin belirlenmesi.
- Sorulacak soruların öğrenilmiş olan konuşma stratejilerini ölçmeye yönelik olması

Yukarıda da belirtildiği gibi konuşma sınavları ölçülecek olan konuşma türüne, hangi sözel öğelerin önem kazandığına bağlı olarak farklı amaç ve ölçütlerle tasarlanabilir. Bu noktada bir diğer önemli konu nasıl test ettiğimiz kadar neyin test edileceğidir. Viyoga (2000) konuşma sınavlarıyla ölçülmesi gereken önemli unsurları aşağıdaki gibi özetlemiştir:

<table>
<tr><td>Doğruluk:
Dili Bir Sistem Olarak İyice Kavrayabilme Yeteneği</td><td>Akıcılık:
Dili İletişim İçin Kullanabilme Yeteneği</td></tr>
<tr><td>Telaffuz:
Doğru Sesleri Üretebile, Vurgu Ve Tonlamaları Takip Etme Yeteneği</td><td>Mekanik Öğeler:
Cümle Uzunlukları, Duraklamalar, Ritim Ve Hızı Kullanabilme Yeteneği</td></tr>
<tr><td>Kelime Bilgisi:
Gerekli Sözcükleri Kullanabilme Ve Sorulardaki Kelimeleri Anlayabilme Yeteneği</td><td>Dil Kullanımı:
Dilin Anlamsal Ve İşlevsel Öğelerini Kullanabilme Yeteneği</td></tr>
<tr><td>Dilbilgisi:
Verilen bir konuşmadaki uygun söz-dizimsel ve biçimsel kalıpları üretebilme yeteneği</td><td>Muhakeme Becerileri:
Çeşitli konular üzerine konuşabilme yeteneği</td></tr>
<tr><td colspan="2">Anlama:
Bilgiyi akıcı ve doğru bir tavırla müzakere etme, aktarma ve paylaşma yeteneği</td></tr>
</table>

Konuşma becerileri test edilirken, öğrenenlerin dilin kurallarını uygun ve doğru bir şekilde takip edip etmediklerinin ölçülmesi gerekmektedir (Susanti, 2019). Konuşma becerilerinin test edilmesi sürecinde ölçülmesi gereken bir diğer unsur da kelime dağarcığıdır. Mertova'ya (2013) göre kelimeler anlamlı cümleler kurma, soruları yanıtlama ve fikirleri açıklamak için gerekli bir bileşendir.

Konuşmacıların kendilerini doğru olduğu kadar akıcı bir şekilde ifade etmeleri beklenmektedir. Fillmore (1979), akıcılık kavramını uzun sohbetleri birkaç duraklamayla konuşabilme, zamanı konuşarak doldurabilme yeteneği olarak tanımlarken; Lennon'a (2000) göre bu kavram en dar anlamıyla; düşüncelerin hızlı, düz, doğru, anlaşılır ve etkili bir şekilde iletilmesidir. Akıcılık, konuşma becerilerinin değerlendirilmesine yönelik bir anlayış geliştiğinden beri önemli bir ölçme kriteri olmuştur (Fulcher, 2003). Makul bir hızda konuşma ve az sayıdaki küçük duraklamalar akıcılığın en önemli göstergeleridir (Mertova, 2013). Bu göstergeler aslında, konuşmacıların hedef dilde konuştukları zaman mesajı iletmek için gereken kelime ve dil bilgisi yapıları üzerinde düşünmek için fazla zaman harcamadıklarının işaretidir (Agung, 2012). Böylelikle hedef dilde akıcı konuşabilmenin konuşma becerilerinin ölçülmesinde önemli bir kriter olduğunu söyleyebiliriz.

Konuşma becerilerinin test edilmesinde önemli bir ölçüt olan telaffuz, konuşulan dilin hem seslerini hem de tonlama, vurgu ve ritim gibi özelliklerini ifade eder (Isaacs, 2014). Konuşma yeterliliğinin değerlendirilmesi için bir ölçüt olmasına karşın telaffuz, yabancı dil araştırmalarında, uygulama ve değerlendirmede bu zamana dek ihmal edilen bir alan olmuştur (Baker 2013; Derwing & Munro, 2005). Yapılan araştırmalar bu yapının değerlendirilmesini hem teste uygulayıcıları hem de öğrenenler açısından zorlayıcı ve kaygılı bir süreç olduğunu göstermektedir (MacDonald, 2002; Levis, 2006). Telaffuz yeterliliğinin ölçülmesi gerektiğini savunan Pramujunarta'ya (2007) göre, dil sınıflarında telaffuzun öğretilmesi ve ölçülmesi, hedef dili anadili ya da ikinci dil olarak konuşan kişilerin, dili öğrenenleri anladığından emin olmak için önemlidir.

Konuşma becerilerinin ölçülmesinde iki önemli yaklaşım vardır. Bunlardan ilki olan, doğalcı yaklaşım; öğrencilerin sınıf gibi normal ortamlarda kendiliğinden meydana gelen iletişimsel davranışlarının dolaylı olarak ölçülmesini önerir. Bir diğer değerlendirme yaklaşımı olan yapısalcı yaklaşım ise, öğrencilerin diğer öğrenciler ya da test uygulayıcılarıyla etkileşimini içeren ya da içermeyen konuşma etkinlikleri kullanılarak

öğrencilerin iletişimsel davranışlarının daha doğrudan ve açık bir şekilde ölçülmesini hedefler (Mead, 1980). Doğalcı ölçme yaklaşımı bir öğrencinin günlük konuşma ve aktivitelerini ölçerken, yapısal ölçme yaklaşımı ise bir öğrencinin verilen konuşma görevlerini tamamlamak için en iyi çabasını ölçer (Mead, 1980).

Dil yeteneğinin ölçülmesi için Clark (1983) tarafından önerilen dolaylı, doğrudan, yarı-doğrudan ölçme yöntemleri öğrenenlerin konuşma performanslarını anlamak için fayda sağlamaktadır. Dolaylı ölçme, sözel yeterlilik gibi becerilerin temelinde yatan diğer yetenekleri ölçmek için kullanılır (Hughes, 2003). Yani sözel becerileri doğrudan ölçmek yerine, bu tür testler konuşma yeteneğini dolaylı bir şekilde ölçen parçaları içerir. Örneğin; Lado (1964) kitabında telaffuzun konuşmaya gerek olmadan kalem-kağıt testleriyle ölçülebileceğini öne sürmüştür. Ancak bu yaklaşım bilimsel olmayan dönemde ortaya atılmış ve günümüzde geçerliliğini yitirmiştir (O'loughlin, 2001). Qian'a (2009) göre, dolaylı testler öğrenenlerin gerçek performanslarını doğru ve güvenilir bir şekilde ölçemediğinden, ölçme sonuçlarının geçerli olduğunu söylemek oldukça güçtür.

Bir diğer sözel test türü doğrudan ölçmedir. Hughes'a (2003) göre, bir test ölçmek istediğimiz beceriyi adayın tam olarak yerine getirmesini gerektiriyorsa, o zaman ölçmenin doğrudan olduğunu söyleyebiliriz. Sözel performansın doğrudan bir şekilde ölçülmesi, öğrenenlerin konuşma yeterliliklerini gösterebilecek sözel görevleri ne derece gerçekleştirdikleri ile ilgilidir. Yan ve Ginther'a (2017) göre, bu test yöntemi ölçülmek istenen beceriyi gerçek performansa en yakın şekilde ölçtüğü için daha geçerli sonuçlar ortaya çıkarmaktadır.

Yarı-doğrudan ölçme yöntemi, önceden kaydedilmiş kayıtlarla sözel performansı değerlendirmeyi amaçlamaktadır (Clark, 1983). Günümüzde, bu tanımın alanı genişletilerek, bilgisayar ve internet gibi son teknolojik araçlarla yapılan test süreçlerinin de yarı-doğrudan ölçme kapsamında olduğu belirtilmektedir. Yarı-doğrudan bir testte, öğrenenler kaset, video, bilgisayar gibi pek çok araçtan önceden kaydedilmiş sorular veya görevler almakta ve yanıtlar sonraki değerlendirme için kaydedilmektedir (Qian, 2009). Bu yöntemin en belirgin avantajı güvenilir olmaları, zaman ve maliyet tasarrufu sağlamalarıdır. Ayrıca, öğrenenlerin konuşma performansları test uygulayıcılarının teknikleri ya da davranışlarından dolayı olumsuz bir şekilde etkilenebileceği için, yarı-doğrudan ölçme bu etkileri azaltabilir (Yan & Ginther, 2017).

Konuşma becerilerinin değerlendirilmesine yönelik diğer yaklaşımlar ise puanlama yöntemleri arasındaki farklılığa dayanır. Bu yaklaşımlardan olan holistik ve analitik puanlama sözel yeterliği değerlendirmek için yaygın şekilde kullanılır (Xi, 2007; Al-Amri, 2010). Holistik yaklaşım konuşma performansının geneline bakılarak puanlama yapılmasıyla ilgiliyken; analitik yaklaşım testteki farklı özelliklerin birbirinden bağımsız tanımlanması ve puanlamasına odaklanır (Richards & Schmidt, 2013).

Her iki yaklaşımın da puanlama sürecinde çeşitli avantajları ve ya dezavantajlarının olduğu savunulabilir. Kondo-Brown'a (2002) göre, holistik puanlama, analitik puanlamaya göre, daha az zaman alır ve kısmen daha kullanışlıdır ve hazırlaması kolaydır. Diğer taraftan sağladıkları yarar göz önünde bulundurulduğunda ise analitik yaklaşım öğrenenlerin dil yeteneklerine dair daha fazla bilgi sunmaktadır. Tuan (2012) analitik yaklaşımın üstünlüklerine dikkat çekerek, bu tür puanlamanın öğrenenlerin konuşma performanslarının değerlendirilmesine yönelik daha tanımlayıcı bilgiler ortaya çıkardığını belirtmektedir. Bununla birlikte, analitik puanlama öğrenenlerin performansını objektif olarak değerlendirme fırsatı sunduğundan dolayı ölçme ve değerlendirme sürecinin güvenirliği de artmaktadır (Doğan ve Uluman, 2017).

2003 yılında Cambrigde Üniversitesi'nin İngilizce Yeterlilik Sertifikasını (Certificate of Proficiency of English) almak için yapılan konuşma sınavlarını ineceleyen Weir, aslında tarihsel süreç içerisinde konuşma becerilerinin ölçülmesinde yaşanan yöntem değişikliklerini de ortaya çıkarmıştır (Akt. O'Sullivan, 2011). O'Sullivan (2011), Weir'in bu çalışmasına dayanarak 1913 ve 2003 yılları arasında yaşanan değişimi aşağıdaki gibi özetlemiştir:

Yıl	**Görev**	**Görev**	**Görev**	**Görev**	**Değerlendirme**
1913	Yazım çalışmaları (dikte)	Sesli Okuma	Konuşma		Belirsiz
1934 - 1966	Yazım çalışmaları (dikte)	Sesli Okuma	Konuşma	*Hikaye yazma (İsteğe bağlı)	Belirsiz

1975	Resimlere dayalı görüşme soruları	Uzun konuşma (Sınavdan 15 dk önce konu verilir-not almaya izin var)	Sesli Okuma (Sınavdan 10 dk önce metin verilir. Aday ve değerlendirici tarafından birlikte okunur)	Değerlendirici tarafından okunan metin dinlenir ve uygun cevaplar verilir	Değerlendirme ölçeği: Sözcük, dilbilgisi ve yapı, tonlama, ritim, vurgu, sesletim, genel iletişim
** 1984	Resimlere dayalı görüşme soruları	Uzun konuşma ve dinlenen metinlere uygun karşılık verme	Sesli okuma (Sadece aday kısa metinler okur)	Rol canlandırma	Akıcılık, Dilbilgisel Doğruluk, Sesletim ve Vurgu, İletişim Becerisi, Sözcük Bilgisi
2003	Bire bir görüşmeler	Resim temelli ikili/küçük grup tartışmaları	Bireysel uzun konuşmalar	Grup tartışmaları	Dilbilgisi ve Sözcük varlığı, Söylem Yönetimi, Sesletim

*Özellikle 1938-1945 yılları arasında kullanılmıştır

**Seçilen ölçütlere uygun bireysel ve grup görevleri verilmiştir

Tabloda görüldüğü gibi, 1975 yılına kadar konuşma becerilerinin ölçülmesi için kesin bir ölçme değerlendirme anlayışı yoktur. Örneğin; sesli okuma görevleri alt düzey okuma becerileri ile birlikte telaffuz ve tonlamayı gerektirirken; yazım görevleri ise konuşma yerine dinleme ve yazma becerilerinin bir birleşimi gibidir (O'Sullivan, 2011). İletişimsel test süreçlerinin gelişimine paralel olarak, sözlü becerilerin ölçülmesi önem kazanmıştır. Özellikle ikili ya da grup çalışması, görev-temelli öğrenme, dil farkındalığı gibi kavramların ön plana çıkmasıyla birlikte konuşma becerilerinin ölçülmesinde kullanılan yöntemler de iletişimsel yeteneği daha çok vurgulamaya başlamıştır (Hughes, 2003).

Günümüzde, konuşma becerisinin ölçmek için bire bir görüşmelerden, rol canlandırmaya, sözlü sunumlardan, resim betimlemeye, işbirlikçi görev ve tartışmalara kadar birçok yöntem kullanılmaktadır. Bu tekniklerin her birinin kendi başına etkili olduğu

söylenebilir. Gampper'a (2013) göre, bireysel farklılıklar göz önünde bulundurularak pek çok tekniğin bir arada kullanılması, öğrenenler açısından daha faydalı olacaktır.

2.4.3. Okuma Becerilerinin Değerlendirilmesi

2.4.3.1. Okuma

Yabancı öğretiminin en gerekli becerilerinden biri olan okuma; yazılı materyallerin anlamlandırılması, değerlendirilmesi ve etkili iletişim kurulabilmesi için önde gelen yapılardandır. Krashen'a (2009) göre, okumanın sadece okuyucuya yazılı metinlerdeki belli bir konuyu kazandırma süreci olarak tanımlanması oldukça yanıltıcıdır. Aksine okuma, yazılı metinlerdeki ipuçlarından anlam oluşturma süreci olduğu için aktif bir yapıdadır (Anderson & Pearson, 1988). Dubin ve Bycina (1991) okumayı; metin ile okuyucu arasında meydana gelen ve okuyucunun metni anlamlandırmak için ön bilgileri ile dilsel öğelere dair diğer bilgilerini kullanmasını gerektiren seçici bir süreç olarak açıklamaktadırlar. Grabe ve Stoller'a (2002) göre ise okuma, bir metindeki bilgileri anlama ve bunu göre uygun şekilde yorumlama yeteneğidir (Akt. Antoni, 2010)

Tarihsel süreç içerisinde okuma becerilerinin tanımlanması, bu sürece etki eden unsurların belirlenmesi, okuma türlerinin saptanması ve ölçme değerlendirme süreçlerine yönelik anlayışın, dil öğreniminde değişen yaklaşım ve yöntemlere göre değiştiğini söyleyebiliriz. Maarof'a (2016) göre, okuma sürecinde metinlerin ön planda olduğuna dair görüşlerin yerini okuyucu ile metin arasındaki etkileşimin en önemli süreç olduğuna dair bir anlayış almıştır. Bu değişimi sağlayan en önemli aşamalardan biri, okuma sürecinde okuyucunun aktif bir rol aldığına dair görüşlerin ön plana çıkmasıdır (Samuels & Kamil, 1984). Yapılandırıcı yaklaşımla değişen bu paradigma, okuyucuyu önceden hafızada muhafaza edilen bilgiler ile yeni bilgileri ilişkilendiren (Bernhardt & James, 1987) ve aktif bir şekilde zihinsel bir şema oluşturan biri olarak tasvir etmektedir (Spivey & King, 1989).

Okuduğunu, anlamaya dair şimdiki görüşlerin yapılandırmacı bakış açısından büyük ölçüde etkilendiğini söylemek mümkündür. Bernhardt'ın (1993) yapılandırmacı modeline göre, yeni bilgiler ile eski bilgilerin ilişkilendirilmesi metnin temasına, okuyucunun ön bilgilerine ve dilsel yeteneğine göre farklı şekillerde oluşur (Akt. Grabe, 1991). Bunun gibi pek çok okuyucu ve metin değişkenlerinin sonucu olarak da sadece tek bir gerçek anlamın oluştuğunu söylemek doğru değildir. Dolayısıyla okumanın farklı okuyucular için farklı yorumlamalar gerektiren bireysel bir süreç olduğunu söylemek

mümkündür. Anlamın yeniden yapılandırılmasında oluşan bireysellik büyük ölçüde her okuyucunun sahip ön bilgilerin farklılığından kaynaklanmaktadır (Grabe, 1991).

Okuma becerilene dair değişen modeller ve teoriler bu süreci anlamlandırmak için temel oluşturmaktadır. Okuma sürecinin ne olduğuna dair bir farkındalık oluştuktan sonra, ölçme ve değerlendirme faaliyetlerinin değişen anlayışa uygun olarak düzenlenmesi ile ideal ve istenilen ölçmenin gerçekleştirilebileceği söylenebilir.

2.4.3.2. Okuma Türleri

Okumanın; pek çok sürecin, bilgi ve yeteneklerin etkileşimini içerdiğine dair genel bir görüş birliği vardır (Garcia & Pearson, 1991; Grabe, 1991; Hedge, 2000). Bu süreçler; ön bilgilerin kullanımı, stratejileri işleme ve metnin yapısını anlama gibi daha yüksek düzey bilişsel süreçlerin yanında yazı birimlerinin tanınması, sözcük erişimi, sesbilimsel temsiliyet ve dilsel yapıyı işleme gibi temel çözümleme işlemlerini de içermektedir (Hudson, Torgesen, Lane & Turner, 2012). Okuma aynı zamanda değerlendirme becerileri gibi üst bilişsel bilginin kullanımını ve öz-izleme gibi etkileşimli süreçleri de içermektedir. Bu noktada önemli olan okuma sürecinde yer alan bu bileşenlerin birbirleriyle olan ilişkilerini incelemektir. Bu parçaların birbirleriyle olan ilişkisi çeşitli okuma modellerini belirlemekte ve okuma alanındaki başarı veya başarısızlıkların açıklanmasını kolaylaştırmaktadır (Hudson, Torgesen, Lane & Turner, 2012).

Bir okuma modeli olarak aşağıdan-yukarıya okuma süreci, okuyucunun, bir metni en küçük biriminden başlayarak en büyüğe doğru anlamlandırması olarak tanımlanır (La Berge & Samuels, 1974). Özellikle geleneksel yaklaşımlarda yer bulan bu model, okuyucunun sırasıyla harf, kelime ve cümlelerden anlam oluşturduğunu savunur (Gough, 1972). Böylece parçadan bütüne gidilerek anlama ulaşılmaya çalışılır (Bekleyen, 2015). Yukarıdan- aşağıya okuma modelinde ise; okuyucunun dünya ve dile dair arka plan bilgisini kullanarak anlam oluşturduğu vurgulanır. Aşağıdan-yukarıya okuma modelinin aksine, bu modele göre anlam; metnin başlığı, her paragrafın ana fikri gibi yönlere bakılarak daha genel bir bakış açısıyla oluşur ve daha sonra daha küçük dilsel öğelere odaklanılır (Smith, 1971). Çakır'a (2015) göre, okumanın anlamaya dayandığı bu süreçte temel amaç, metnin bütününe bakıldıktan sonra ayrıntılara gidilmesi ve daha bütünsel bir anlamın oluşmasıdır. Bu noktada önemli olan dilsel yapıdan çok anlama odaklanmaktır.

Aşağıdan-yukarıya ile yukarıdan-aşağıya okuma modellerinin taban tabana zıt noktaları vurguladıkları görülmektedir. Ancak çoğu araştırmacı okumayı bu iki modelin de öğelerini içeren etkileşimi bir süreç olarak görmektedirler. Etkileşimli okuma süreci diye adlandırılan bu model, okumayı; yukarıdan-aşağıya ve aşağıdan-yukarıya süreçlerin unsurlarını eş zamanlı bir şekilde kullanarak anlama ulaşmak olarak değerlendirir (Rumelhart, 1991). Okuyucular okudukları bir metni yorumladıkları zaman hem dil yapılarına dair bilgilerini hem de arka plan bilgisini (schemata) kullanırlar. Dolayısıyla okuyucu okuma sürecinde elde ettiği bilgileri, daha önce sahip olduğu bilgilerle ilişkilendirerek metindeki dilsel öğeleri yorumlamaya çalışır (Carrell & Eisterhold, 1983).

Yabancı dil öğrenenlerin okuma süreçlerinden birine aşırı bağlanıp diğerini ihmal etmesi okuma güçlüklerinin oluşmasına neden olmaktadır (Carrell, Devine & Eskey, 1988). Doğru ve akıcı bir okuma için, okuyucuların hem aşağıdan-yukarıya algılama becerilerini hem de yukarıdan-aşağıya yorumlama stratejilerini geliştirmeleri gerekmektedir. Eskey'e (1988) göre akıcı okuma hem ustaca çözümlemeyi hem de yeni bilgiler ile önceki bilgiler arasında bağlantı kurmayı gerektirmektedir. Bu iki model arasında kurulması gereken denge ölçme-değerlendirme faaliyetlerini de etkilemektedir. Okuma becerilerinin test edilmesi için kullanılacak ölçme araçlarının ve seçilecek metinlerin tek bir okuma modeline bağlı kalmaktansa, her iki modelin öğeleri arasındaki etkileşimi ve dengeyi gözeterek hazırlanması gerekmektedir.

Dinleme ve konuşma becerilerinde olduğu gibi, okuma becerilerinin de test edilmesinden önce amaçların belirlenmesi, uygun ölçme ve değerlendirme araçlarının seçilmesi ve uygulanması gereklidir. Bu amaca hizmet eden çeşitli okuma performanslarının tanımlanması, yabancı dilde okuma becerilerinin ölçülmesi ve değerlendirmesi için önemli ölçütler sunmaktadır. Brown (2004) tarafından bu amaçları gerçekleştirme yönelik 4 farklı okuma performansı tanımlanmıştır.

1) Algısal Okuma: Bu okuma becerisi, harfler, kelimeler, noktalama ve diğer yazıbilimi parçalarına dikkat etmeyi gerektirir. Aşağıdan-yukarıya okuma süreci uygulanmaktadır.

2) Seçici Okuma: Bu okuma türünde, okuyucunun kısa bir yazılı metinde sözcüksel, dilsel ve söylem özelliklerini ne kadar tanıdığını ortaya çıkarmak amaçlanmaktadır. Eşleştirme, doğru-yanlış, çoktan seçmeli gibi belirli görevler bu beceriyi ölçmek için

kullanılabilir. Kısa yanıtlarda istenilebilir. Aşağıdan-yukarıya ve yukarıdan-aşağıya okuma süreçlerinin bir birleşimdir.

3) Etkileşimli (Interactive): Okuma, bir anlam oluşturma sürecidir. Okuyucu metni anlamlandırmak için bir takım arka alan bilgisini kullanır. Etkileşimli bir etkinliğin odak noktası, metnin içindeki bağlantılı özellikleri (sözcüksel, dilsel, sembolik ve söylemsel) belirlemektir. Bazı aşağıdan-yukarıya işleme performansının örnekleri gerekli olsa da genel olarak yukarıdan-aşağıya işleme modelinin tipik görevlerini yansıtır. Anekdotlar, uzun metinlerden alıntılar, anketler, ilanlar, yönler, yemek tarifleri bu okuma türüne örnek verilebilir.

4) Kapsamlı (Extensive Reading): Bu okuma türüyle ölçülmek istenen küçük detaylara odaklanmaktansa okuyucuların metinle ilgili genel algısını belirlemektedir. Bir sayfadan fazla profesyonel makaleler, denemeler, teknik raporlar, kısa hikayeler ve kitaplar bu okuma türünün örneklerindendir.

2.4.3.3. Okuma Performansının Mikro ve Makro Becerileri

Okuduğunu anlama becerilenin etkili bir şekilde değerlendirilebilmesi için, ölçülecek davranışların belirlenmesinde mikro ve makro becerilerin bilinmesi oldukça önemlidir. Brown, (2004) tarafından belirlenen okuma performansına ait mikro ve makro beceriler şunlardır:

- **Mikro Beceriler:**

1. İngilizcenin belirli yazımsal kalıplarını ayırt etme
2. Kısa süreli hafızada farklı uzunluktaki dil parçalarını muhafaza etme
3. Amaca uygun şekilde ve yeterli bir hızda yazıyı işleme
4. Kelimeleri tanıma ve sözdizim kalıplarının ve onların önemini yorumlama
5. Dilbilgisel sözcük sınıflarını (isimler, fiiller vb.), sistemleri (zaman uyumu, çoğul yapma vb.), kalıpları ve kuralları tanıma
6. Belirli bir anlamın farklı dilbilgisel biçimlerde ifade edilebileceğinin farkına varma
7. Yazılı metindeki bağlaçların ve onların cümleler arasındaki bağlantıyı kurmadaki rollerinin farkına varma

- **Makro Beceriler:**

8. Yazılı metinlerin sözbilimsel biçimlerinin ve onları yorumlamadaki öneminin farkına varma
9. Amaca ve biçime göre, yazılı metinlerin iletişimsel işlevlerini tanıma

10. Arka alan bilgisini kullanarak bağlamdan sonuç çıkarma
11. Olaylar arasında bağlantı kurma, nedenlerini ve etkilerini ortaya çıkarma, ana fikir, destekleyici fikirler, yeni bilgiler, genelleme ve örnekleme gibi ilişkileri belirleme
12. Gerçek anlam ile mecaz anlamı ayırt etme
13. Kültürel örnekleri tespit etme ve yorumlama
14. Tarama, hızlıca göz atma, bağlamdan bilinmeyen kelimelerin anlamını tahmin etme ve metinleri yorumlamak için arka alan bilgisini aktifleştirme gibi bir takım okuma stratejilerini kullanma ve geliştirme.

2.4.3.4. Okuma Becerilerini Ölçme Değerlendirme Yöntem ve Teknikleri

Algısal bir beceri olan okuma performansının doğrudan gözlemlenebilmesi mümkün değildir. Sadece gerçek okumayı yansıttığı düşünülen davranış örneklerine bakılarak öğrenenlerin okuma performanslarına dair sonuçlar çıkarabiliriz. Cain ve Oakhill'e (2006) göre, araştırmacıların ve test uygulayıcıların öğrenenlerin okuma yeteneklerini ölçmeye neden ihtiyaç duyduklarına dair pek çok farklı sebep vardır. Örneğin; okuma güçlüklerini belirlemek ve teşhis etmek veya öğrenci gelişimini takip etmek bu sebepler arasındadır. Her bir amaç için okuma yeteneğinin doğru şekilde değerlendirilmesine ihtiyaç vardır. Eğer okuyucuların performansına yönelik bir ölçme yapmak gerekiyorsa, bu noktada önemli olan uygulamadaki değerlendirme yaklaşımlarının okuma süreci ile ilgili teori ve araştırmalarla uyumlu olmasıdır (Bernhardt & James, 1987).

Okuduğunu anlama yeteneğinin tanımlanması ve sürecin nasıl işlediğini saptamaya yönelik yapılan araştırmalar, bu becerilerin nasıl değerlendirilmesi gerektiğini açıklayan görüşleri de yakından etkilemiştir. Grabe and Jiang'a (2013) göre 20.yüzyıldan önce, anlamdan çok kültürel ve edebi öğelerin ölçülmesine ağırlık verilmiş ve daha öznel değerlendirmeler yapılmıştır. Ayrıca çoktan seçmeli, doğru-yanlış testleri ve eşleştirme gibi daha sınırlı ölçme ve değerlendirme araçları kullanılmıştır. Okuyucular davranışçılığın önemli metaforlarından olan "tabula rasa" (boş levha) olarak görülmüş ve bilginin mutlak alıcısı olarak kabul edilmişlerdir (Maarof, 2016)

20. yüzyıldan sonra ise özellikle de yapılandırmacı anlayışın etkisiyle, okuyucuların pasif bir alıcı olduğu yönündeki görüşler değişmiştir. Aksine, anlamın aktif bir şekilde okuyucu tarafından metinle etkileşime geçerek yapılandırıldığına dair genel bir kavrayış oluşmuştur (Carrell, Devine & Eskey, 1988). Böylelikle yazılı metinlere gösterilen ilgi

yerini, okuyucu ve metin arasındaki etkileşime bırakmıştır. Yapılandırmacı yaklaşıma göre, okuma becerilerini değerlendirirken, okuyucunun anlamı nasıl oluşturduğuna odaklanan ölçümler yapılmaktadır (Bernhardt, 1983). Yani, ürün yerine artık süreç vurgulanmaya başlamıştır. Maarof'a (2016) göre, bireylerin anlamı nasıl oluşturduğuna dair yaşanan bu değişim, pek çok eğitmenin ve araştırmacının performans ölçümlerindeki yetersizlikleri fark etmelerini sağlamıştır. Sonuç olarak, okuyucuların yazılı metni nasıl çözümlediklerini oraya çıkaran ölçme ve değerlendirme araçlarının gerekliliğine dikkat çekilmiştir.

Okuma performansının hangi araçlarla ve tekniklerle ölçüleceği nasıl bir değerlendirme yapılacağı gibi süreçler belirlenmeden önce asıl yanıtlanması gereken soru, bu performansın hangi amaçlar için test edileceğidir. Cross ve Paris'e (1987) göre, okuduğunu anlama becerilerinin değerlendirilmesi üç özel amaca hizmet etmelidir. Bunlarda birincisi; öğrenenlerin akademik başarılarını tahmin etmek ya da öğretim programının yeterliliğini görmek için. İkincisi, öğrenenlerin stratejileri ve süreçleri hakkında bilgi toplamak için. Bir tür teşhis koyma aşamasıdır. Böylece öğretmen öğretim sürecinin nasıl işlediğiyle ilgili karar verir. Üçüncü amaç eğitim programını değerlendirmek; yani programın belirli bir toplum üzerindeki etkilerini saptamak. Grabe ve Stoller (2002) ise, değerlendirme uygulamalarından önce göz önünde bulundurulması gereken en önemli kriterin, şu becerileri birleştirmek olduğunu öne sürmüştür: akıcılık ve okuma hızı, otomatiklik ve kelimeleri tanıma hızı, araştırma süreçleri, kelime bilgisi, biçim bilgisi, sözdizimi hatırlama, sonuç çıkarma, değerlendirme ve son olarak eleştirel okuma. Onlara göre, ölçme uygulamaları dış dünyadaki okuma ihtiyaçları ve aktivitelerine dayanılarak hazırlanmalıdır. Hubley (2012), okuma performansının etkili bir şekilde değerlendirmesi için şu davranışların dikkate alınması gerektiğini belirtmektedir.

- Programınızdaki hangi becerilerin öğrencileriniz açısından önemli olduğuna karar verin
- Programınızda her seviyeye uygun test kriteri geliştirin ve bunları takip edin
- Uygun metinler seçin, bunları uzunluk ve okunabilirlik açısından kontrol edin ve eğer gerekliyse öğrenci seviyelerine göre üzerinde değişiklik yapın
- Öğrencilerin daha önce karşılaştıkları metinleri kullanmayın ama benzer konular seçebilirsiniz
- Metindeki bilinmeyen kelimelerin sayısını 5-10 olacak şekilde sınırlayın

- Soruların seviyesini okuma metninden daha az zor yapın
- Sadece kolay olan neyse onu ölçmektense, bir takım alt becerileri ölçün
- Testi uyguladıktan sonra, analiz yapın ve öğretmenler ile öğrencilere yapıcı geribildirimler sunun
- Okuma performansını, okuma stratejilerini vurgulayarak sınıfta test edin.

Öğretmenler yukarıda belirtilen okuma davranışlarını değerlendirmek için pek çok ölçme aracı kullanırlar. Aebersold ve Firel (1997) bu ölçme değerlendirme araçlarından bazılarını informal, alternatif, gelişimsel, diğerlerini de formal, öğretmen kontrollü, geleneksel ve standart araçlar olarak tanımlamışlardır. Frodden, Restrepo ve Maturana (2004) tarafından yapılan başka bir sınıflamada ise ölçme araçları "sıkı" veya "yumuşak" olarak gruplandırılmıştır. Sıkı ölçme araçları süreçten çok ürüne odaklanan, nesnelliği ve güvenirliği vurgulayan geleneksel araçlarken, yumuşak ölçme araçları alternatif, otantik araçlardır (Akt. Munoz Marin, 2009).

Okuma performansının değerlendirilmesine yönelik doğru-yanlış cümleleri tanıma, çoktan seçmeli test, açık uçlu sorular, boşluk doldurma, metnin parçalarını sıralama veya kısa yanıtlı eşleştirme gibi pek çok teknik kullanılabilir. Ayrınca informal ve alternatif değerlendirme seçenekleri de etkili bir değerlendirme için oldukça kullanışlıdır. Grabe (2009), öğretmenlerin sıklıkla başvurduğu 25 informal değerlendirme aktivitesi tanımlamıştır. Bunlardan bazıları şunlardır: Öğrencilere sınıfta sesli okuma yaptırma ve değerlendirme, okumadan sonra sorulara verilen cevapların kaydını tutma, öğrencilerin metin üzerinde ne kadar zaman harcadıklarını gözlemleme, öğrencileri bir ses bandındaki okuma metnini dinlerken gözlemleme, okumadan sonra öğrenciden öğrenmek istedikleri kelimeleri kaydetmelerini isteme, öğrencilere basit kitap raporları yazdırma, öğrencilerin okuma hızlarının gelişimini takip etme, öğrencilerin gelişim dosyalarını hazırlama. Grabe ve Jiang'a (2013) göre bütün bu informal değerlendirme aktivitelerinin arasında, alanyazında en çok dikkat çeken ve anlama becerisinin güçlü bir yardımcısı olarak kabul edilen sesli okuma akıcılığı olmuştur.

Okuma becerilerinin değerlendirilmesi için üründen çok sürece odaklanan ve öğrenci gelişimini destekleyen alternatif ölçme-değerlendirme araçları da kullanılabilir. Aebersold ve Field (1997) öğrencilerin öğrenme sürecini, onların derse katılımını vurgulayan ve bu süreçleri gözlenebilir hale getiren alternatif araçları şu şekilde sıralamıştır:

1) Öğrencilerin bilgiyi kullanmalarını, dil yeterliliği kazanmalarını sağlayan dergiler (sesli ya da yazılı),

2) Öğrencilerin çalışmalarının değerlendirilmesi için kullanılan portfolyolar

3) Öğrencilere bilmedikleri öğrenme veya ne hakkında soru sormaya ihtiyaç duyduklarını gösteren ev ödevleri

4) Öğrencilerin grup ya da akran çalışmalarını izleme, öğrenci katılımını gözlemleme gibi öğretmen değerlendirmeler

5) Öğrencilerin kendi başarılarını ne kadar yansıttıklarını soran öz-değerlendirmeler

6) Öğrencilere kendi sınıf arkadaşlarının görüşlerini sunan akran değerlendirme.

Okuma becerilerinin ölçmeye yönelik pek çok yöntem ve araç olsa da hiçbir araç ya da teknik tek başına yeterli bir şekilde okuyucuların testi ne kadar algıladıklarını ortaya koymaz (Maarof, 2016). Bu yüzden pek çok araştırmacı hem süreci hem de ürünü yansıtan geleneksel ve tamamlayıcı çeşitli araçların birlikte kullanılmasını tavsiye etmektedirler (Bernhardt, 1991; Hancock, 1994; Grabe and Jiang, 2013).

2.4.4. Yazma Becerilerinin Değerlendirilmesi

2.4.4.1. Yazma

İkinci dilde yazma, modern okuryazar toplumlar için artık daha fazla vurgulanan ve kazanılması zaruri görülen bir beceri olarak karşımıza çıkmaktadır. Özellikle diğer becerilere kıyasla daha karmaşık bir yapıda olan yazma; eğitim, alıştırma, uygulama ve deneyim gerektirmektedir (Grabe & Kaplan, 2014). Yani, yüzmede olduğu gibi olgunlaşmayla birlikte otomatik olarak gerçekleştiren doğal bir yetenek değildir (Liberman & Liberman, 1990). Diğer dil becerileri ile güçlü bir ilişkisi olan yazma, bu beceri alanlarının gelişimine paralel olarak ilerleme gösterecektir (Maden, Dinçel ve Maden, 2015). Dolayısıyla yazma becerisinin belirli bir deneyim ve bilgi birikimi sonucu elde edildiği söylenebilir.

Yazmanın ne olduğu ve hangi süreçleri kapsadığına dair yapılacak tanımlamaların, bu karmaşık yapıyı anlamada ve etkili bir değerlendirme sunmada fayda sağlayacağı düşünülmektedir. Wheeler ve Mcdonald'a (2000) göre yazma; beyin fırtınası yapma, planlama, uygulama, kontrol etme ve düzeltme süreçlerini içeren bir eylemdir. White ve Arndt'a (1991) göre ise yazma, dili yazılı sembollere dönüştürmek gibi basit bir

tanımlamanın ötesinde; düşünsel ve zihinsel bir süreçtir. Demirel'in (2003) eleştirel bir düşünme süreci olarak gördüğü yazma, dört temel yabancı dil beceri zincirinin son halkasıdır ve üretkenliği, aktif katılımı gerektirmektedir. Ayrıca, yazma bir yaratma sürecidir. Aslında bu süreç bir taraftan mekanik ve formal bir şekilde işleyen aktiviteleri kapsarken, diğer taraftan daha karmaşık olan yaratıcı davranışları da içermektedir (Hadley & Reiken, 1993). Hatta bu süreç öğrenenlerin bir takım problemlere çözüm aradıkları, bir problem çözme aktivitesi olarak da görülebilir (Hyland, 2008). Problem çözümleri ve kavramları birlikte ortaya koymaya çalışan öğrenenler, sürekli gelişen bilgi ve metin arasında iki yönlü bir etkileşime girerler (Bereiter & Scardamalia, 1987). Anlam oluşturmak için bilgiyi keşfetmeyi ve yeniden yapılandırmayı gerektiren bu zihinsel süreç, özellikle ikinci dil öğrenenler için zorlayıcı olabilmektedir (Nunan, 1999). Dolayısıyla öğrenenlerin yabancı dil yazma becerilerini ölçmeden önce, bu sürecin neler gerektirdiğini bilmek ve buna uygun ölçümler yapmak test uygulayıcıları için oldukça önemlidir.

Tarihsel süreç içerisinde, ikinci cilde yazma özellikle son kırk yılda (1980'lerin sonundan itibaren) katlanarak devam eden bir önem kazanmış ve kendine ait özellikleriyle anadilde yazmadan ayrı bir disiplin olarak ele alınmaya başlamıştır (Matsuda vd., 2003). Geleneksel olarak bakıldığında yazma başarısı, doğru bir yazılı metnin oluşturulması için gereken dil yapılarını anlamlı bir şekilde kullanmak demekti (Kepner, 1991). Bu tanımlamadan anlaşılacağı üzere öğretmenlerin yazma görevlerini değerlendirirken göz önünde bulundurdukları ölçüt, öğrencilerin dilbilgisi, kelime söz dizim kurallarını ne kadar doğru uygulayabildikleriydi. Bernhardt vd.'ne (2015) göre yabancı dilde yazma, dil öğelerine odaklanarak daha çok dil öğrenmeyi destekleyen bir mekanizma olarak görülmekteydi. Özellikle de 1970'lerden önce yazma öğretimi öğrencilerin hatalarını görüp, düzeltmeye odaklanmaktaydı. 1980'lerden sonra ise, araştırmacılar yazma sürecine dair daha kapsamlı ve teorik bir çerçeve oluşturmaya çalışmışlardır (Becker, 2006).

Yabancı dil eğitiminde yapısalcılıktan iletişimsel yaklaşımlara doğru ilerleyen süreç, yazım becerisini öğretme ve ölçme değerlendirme anlayışını da değiştirmiştir. Öğrenmenin, alışkanlık oluşturma olarak kanıksandığı yapısalcı dönemde "kontrollü kompozisyon" adı verilen yazım modeli ön plandaydı. Bu modelde yazma diğer dil becerilerini (dinleme, konuşma ve okuma) geliştirmek için kullanılan bir unsurdur (Randaccio, 2013). Öğrenenlere verilen metinlerde sadece dilsel öğeler ve bunların yarattığı anlam ilişkileri ön plana çıkmaktadır (Yaylı, 2015). Örneğin; öğrencilere

çalışmaları için bir metin verilir ve onlardan beklenilen metni tekrar yazmaları, birkaç özel dilbilgisel ya da yapısal değişiklik yapmalarıdır. Öğrenciler metnin içeriği, organizasyonu ya da cümleleri biçimlendirmek gibi süreçlerle ilgilenmek zorunda değillerdir (Raimes, 1983). Yazma görevleriyle ölçülmek istenen davranışlar ise, yazma sürecinin kendi bileşenleriyle ilgili olmayıp, dilbilgisel veya sözdizimsel öğelerin doğruluğunu tespit etmektir

Yapısalcı yaklaşımlardan sonra yazma sürecini açıklamak için bilişsel süreçlerin önemine vurgu yapan teoriler güç kazanmıştır. Bu teorinin öncülerinden olan Flower ve Hayes (1981), yazma sürecinin doğrusal bir yapıda olmadığını ve öğrenenlerin süreçteki farklı aşamaları sürekli gözden geçirebileceklerini öne sürmüşlerdi. Onlara göre, yazma eylemi öncelikle bir beyin fırtınası yapmayı, ön yazımı, taslak oluşturmayı, yeniden gözden geçirmeyi ve düzenlemeyi gerektiren bir süreçtir. Hodges'e (2017) göre yazmaya sadece yaratıcı bir çaba olarak odaklanmak yerine, bilişsel süreçlerin birlikte çalıştıkları karmaşık bir yapı olarak bakmak gerekir.

Yazma sürecinin basit ve doğrusal olmadığına dair değişen anlayış, yazma için nasıl bir ölçme ve değerlendirmeye gerek olduğuna dair fikirleri de etkilemiştir. Bu dönemde üründen ziyade süreci öncelikli hale getiren bir yaklaşım benimsenmiştir. Badger ve White'a (2000) göre, süreç temelli yaklaşımla dilbilgisi ve metin yapısı gibi dilsel bilgilere yapılan vurgu değişmiş ve yazma süreci planlama, tasarlama gibi zihinsel yapılandırmalarla ilişkilendirilmiştir. Boscolo (2009) özellikle sosyokültürel yaklaşımın da etkisiyle yazmanın sosyal, işbirlikçi bir aktivite olarak kabul edildiğini ve öğretmen ve öğrenci rollerinin de farklılaştığını vurgulamıştır. Süreç temelli yazma yaklaşımıyla, yazma motivasyonel ve yansıtıcı bir aktivite olarak kabul edilmiş ve öğretmenler de öğrencilerin öğrenmelerini kolaylaştıran, onlara yapıcı dönütler veren ve yazmaları için en uygun koşulları sağlayan birer rehber rolüne sahip olmuşlardır. Ayrıca öğretmenler öğrencileri değerlendiren bir pozisyonda değil, aksine onların sorularını yanıtlayarak fikirlerini yeniden düzenlemelerini sağlayan bir gözlemcidirler (Boscolo, 2009).

İletişimsel dil öğretim yönteminin benimsenmesiyle birlikte eleştirel ve yaratıcı düşüncenin önem kazandığı, öğrenmenin sosyal ve etkileşimli bir yapıda olduğu ve tamamlayıcı değerlendirme yöntemlerinin kullanıldığı bir döneme geçilmiştir (Richard, 2006). Bu anlayışa paralel olarak, ikinci dilde yazma süreci sosyo-kültürel bağlamın önem kazandığı, öğrenme ürünlerinin ve sürecin birlikte değerlendirildiği bir aşamaya geçilmiştir

(Cumming, 1998; Akt. Larios, Murphy & Marin, 2002). Connor ve Mbaye (2002) Cande ve Swaine'ın (1980) dilbilgisel, söylem, sosyo-linguistik ve stratejik yeteneklerden oluşan iletişimsel modelini yazma becerisi için yeniden uyarlamışlardır. Dilbilgisel yeteneği dilbilgisi, sözcük ve noktalama işaretleri bilgisi olarak kabul etmişlerdir. Söylem yeteneğiyle kastettikleri metinlerin nasıl bağlantılı ve tutarlı bir şekilde yapılandırıldığıdır. Sosyo-linguistik yetenek kavramı, yazım türünün ve niteliğinin uygunluğuyla ilgilidir. Son olarak stratejik yetenek okuyucu ile yazar arasındaki ilişkiyle alakalıdır. (Okuyucuya uygun dille hitap etme ve güçlü argümanlar sunmak gibi). De Haan ve Van Esch'e (2004) göre, eğer öğrenenlerin yazma yeteneklerini değerlendirmeyi istiyorsak bu yeteneklerin her birini birbiriyle ilişkilendirerek değerlendirmeliyiz.

2.4.4.2. Yazma Türleri

Diğer kinci dil beceri alanlarında olduğu gibi yazma becerilerinin ölçülmesinde yazma türlerinin bilinmesi, hem ölçülecek davranışların belirlenmesinde hem de ölçüt geliştirme sürecinde fayda sağlayabilecektir. Bu anlamda Brown (2004) tarafından dört farklı yazma performansı tanımlanmıştır. Bunlar:

1. *Taklitçi Yazma:* Yazılı dili üretmek için öğrenenlerin sahip olmaları gereken en temel ve basit yazım harfleri, kelimeler, noktalama işaretleri ve çok kısa cümle görevleridir.
2. *Yoğun (Kontrollü) Yazma:* İçeriğe uygun kelimeleri, deyimleri ve doğru dilbilgisel özellikleri kullanmayı gerektiren yazma görevleridir.
3. *Yanıtlayıcı Yazma:* Öğrenenlerin sınırlandırılmış ve ana hatları belirlenmiş bir seviyede, mantıksal bir sıralamayla bağlanmış iki ya da üç paragraf yaratmaları istenilen ölçme görevleridir
4. *Kapsamlı Yazma:* Yazmanın tüm strateji ve süreçlerinin başarılı bir şekilde uygulanmasını vurgular.

2.4.4.3. Yazma Performansının Mikro ve Makro Becerileri

Yazma becerilerine ait mikro ve makro becerilerin tanımlanması, ölçülecek davranışların belirlenmesinde ve değerlendirilmesinde kolaylık sağlayabilecektir. Brown (2004) tarafından belirlenen yazma performansına ait mikro ve makro beceriler şunlardır:

- **Mikro Beceriler**

1. İngilizce yazıbirim ve yazımsal sözcükleri üretme

2. Yeterli bir hızda yazma
3. Makul kelimeleri ve uygun sözcük dizilimini kullanma
4. Uygun dilbilgisel sistemleri (zaman uyumu, çoğullaştırma gibi), kalıpları ve kuralları kullanma
5. Özel bir anlamı farklı dilbilgisel biçimlerde açıklama
6. Yazılı anlatımda bağlaçları kullanma

- **Makro Beceriler**

7. Yazılı anlatımın kurallarını ve söz-bilimsel biçimlerini kullanma
8. Amaca ve biçime uygun şekilde yazılı metinlerin iletişimsel işlevlerini uygun şekilde tamamlama
9. Olaylar arasında bağlantı ve ilişki kurma
10. Yazarken gerçek ile mecaz anlamı birbirinden ayırt etme
11. Yazılı metnin içeriğindeki kültürel öğeleri doğru bir şekilde iletme
12. Bir takım yazma stratejilerini kullanma ve geliştirme

2.4.4.4. Yazma Becerilerini Ölçme Değerlendirme Yöntem ve Teknikleri

İkinci dilde yazmayı değerlendirirken akla gelen temel soru yazmanın ne olduğu bunun nasıl anlaşılacağıdır (Coombe, 2010). Ölçülmek istenen yazma davranışları belirlendikten sonra öğrenenlerin yazma performansı doğrudan ve dolaylı olarak ölçülebilir. Dolaylı ölçme cümle kurma, dilbilgisi, heceleme ve noktalama işaretleri gibi yazma alt becerilerine odaklanan ve iletişimden ziyade daha çok dilbilgisel doğruluğun ön planda olduğu bir ölçme türüdür. Doğrudan ölçme ise bir performans testidir (Hamp-Lyons, 2003). Bu tür bir ölçme öğrencilerin içerik bulmalarını, fikirlerini organize etmelerini, uygun kelime, cümle ve dil yapılarını kullanmalarını gerektirdiği için dilin tüm öğelerini bütünleştirmektedir (Coombe, 2010). Stiggins'e (1982) göre doğrudan ve dolaylı ölçme biçim olarak birbirinden farklı olsa da her iki yöntem de öğrencinin dile dair bilgisini az ya da çok yansıtabilmektedir. Öğrencilerin yazma sürecinde yaşadıkları güçlükleri ve zayıflıkları tanımlamada her iki yöntem de kullanılabilir ancak doğrudan ölçümler gerçek bir yazma performansını yansıttığı için daha detaylı ve geçerlidirler (Stiggins, 1982).

Yazma becerilerini doğrudan ölçmek için çeşitli puanlama türleri kullanılabilir. Temel özellik puanlaması (Primary trait scoring) özel bir yazma göreviyle alakalı bir ya da daha fazla özelliğin tanımlanmasını gerektirir (Huot, 1990). Örneğin; bir görevin nasıl

tamamlandığına dair bir takım açıklamaların yazılması isteniliyorsa, yazma performansıyla ilgili temel özelliklerden biri kullanılan adımların mantıklı ve doğru bir sırayla sunulması olacaktır (Lloyd-Jones, 1977).

Başka bir puanlama yöntemi olarak analitik puanlama yönteminde ise iyi bir yazıda olması gereken pek çok özellik önceden tanımlanır ve bir derecelendirme ölçeği oluşturulur. Daha sonra öğrenenlerin ortaya koydukları yazılı metinler, iyi bir yazıda olması gereken özelliklerin ne kadarını yansıtıyorsa ona göre bir puanlama yapılır (Wiseman, 2012). Son olarak bütüncül (holistic) puanlama ise puanlayıcının yazılı metnin niteliğine genel olarak bakarak bir değerlendirme yapmasıdır. Yazmaya bütüncül bir bakış açısıyla yaklaşan bu metot, öğrenenlerin zayıf oldukları alanlara odaklanmaktansa neyi iyi yaptıklarını göz önünde bulundurmaktadır (Hyland, 2002). Puanlaması daha az zaman alan ve ekonomik olan bu yöntem daha çok geniş ölçekli sınavlarda kullanılmaktadır (Bauer, 1981). Barkaoui'ye (2010) göre, bütüncül puanlama yöntemi daha kullanışlı olsa da kelime yeterliği, organizasyonel kontrol gibi yazma puanını etkileyebilecek değişkenleri ayırt etmede pek yeterli değildir. Bu yüzden analitik puanlama kadar güvenilir olmasa da geçerli bir alternatif sayılabilir (Bauer, 1981).

2.5. İlgili Çalışmalar

Araştırmanın bu bölümünde İngilizce öğretiminde ölçme değerlendirmeye ilişkin yurt içi ve yurt dışında yapılan çalışmaların özeti sunulmuştur.

2.5.1. Yurt İçinde Yapılan İlgili Çalışmalar

Paker ve Höl (2012) tarafından yapılan araştırmada hazırlık sınıfı üniversite öğrencilerinin ve okutmanların "İletişimsel Konuşma Sınavına" ilişkin tutum ve algılarını belirlemek hedeflenmiştir. Özellikle genel hedef öğrencilerin sınav öncesi, sınav süresince ve sonrasında yaşadıkları tutum ve algılar ile okutmaların ölçme aracı olarak kullandıkları sınav ve materyaller hakkındaki deneyim ve düşüncelerini ortaya çıkarmaktır. Veri toplama aracı olarak Güllüoğlu (2004) tarafından geliştirilen anket uyarlanmıştır. Araştırmanın örneklemini, orta-alt ve orta düzey olmak üzere 2 seviye grubunda öğrenim gören toplam 210 hazırlık sınıf öğrencisi ve 32 okutman oluşturmaktadır. Veriler sınav sonrası hem öğrencilere hem de okutmanlara uygulanan anketler aracılığıyla elde edilmiş ve betimsel istatistik analizleri yapılmıştır. Araştırmanın bulgularına göre, öğrencilerin daha önceki yıllarda herhangi bir konuşma sınavına girmedikleri tespit edilmiştir. Bu tür sınavlarla ilgili bir deneyimlerinin olmamasının onların kaygı düzeyini arttırdığı bilgisine

ulaşılmıştır. Sınav içeriği hakkında önceden sunulan yönergeye rağmen bu durum onların kaygı seviyelerini düşürmek için bir fayda sağlamamıştır. Öğrenciler konuşma sınavlarını dil beceri alanındaki en zor sınav olarak algılamaktadır. Diğer taraftan, okutmanlar konuşma sınavlarının uygulanması en zor sınav olduğunu ve sınav esnasında kullanılan ölçme aracının iletişimsel konuşma becerilerini ölçmede yeterli olduğunu ifade etmişlerdir.

Şahinkarakaş (2012) yaptığı araştırmada İngilizce öğretmenlerinin mesleki kıdemlerinin onların değerlendirme algıları üzerindeki etkilerini incelemiştir. Araştırmaya İngilizce Öğretmenliği Bölümünde öğrenim gören 53 öğretmen adayı ve mesleki kıdemi 5 yıldan az olan 23 ve 5 yıldan daha fazla olan 24 olmak üzere toplam 47 İngilizce öğretmeni katılmıştır. Öğretmenlere değerlendirmeye dair metaforik görüşlerini gösteren bir ifade sunulmuş ve ardından onlardan bu metaforu neden seçtiklerini belirtmeleri istenmiştir. Araştırma sonuçlarına göre, öğretmenlerin ürettikleri 73 metafor "bir biçimlendirme aracı olarak değerlendirme", "düzey belirleyici bir araç olarak değerlendirme", "kaygı yaratan bir şey olarak değerlendirme" ve "öz-yeterliliğin işareti olarak değerlendirme" olarak 4 temada kategorize edilmiştir. Biçimlendirici bir araç olarak değerlendirme temasını altındaki metaforların devam eden bir süreçle ilişkili olduğu, geribildirimlerle öğrenmeyi desteklediği ve öğrencilerin güçlü ve zayıf yanlarını teşhis ettiği belirlenmiştir. Düzey belirleyici bir araç olarak değerlendirme teması içerisindeki metaforların ise, ortaya çıkan ürün ve öğrencileri derecelendirmekle ilişkili olduğu saptanmıştır. Diğer taraftan kaygı yaratan bir şey olarak değerlendirme temasındaki metaforların korku ve zorlukları yansıttığı belirtilirken, öz-yeterliliğin işareti olarak değerlendirme temasındaki metaforların ise bireylerin kendileri ve kişisel doyumları ile ilişkilendirildiği kaydedilmiştir. Ayrıca deneyimli öğretmenlerin kendi tecrübelerini çok değerli buldukları ve öğretim uygulamalarının etkililiğini kanıtlamak için çaba harcamaya gerek duymadıkları yönündeki ifadeler araştırmanın önemli bulgularındandır.

Yıldırım'ın (2012) yüksek lisans tez çalışmasında İngilizce dil becerilerinin öğretiminde kullanılan etkinliklerin ve ölçme değerlendirme uygulamalarının sıklığını belirlemek ve dil becerilerinin öğretimi ile değerlendirilme durumları arasındaki tutarlılığı tespit etmek amaçlanmıştır. Bu amaçla, Ankara ilinde görev alan ve öğrenim gören 100 İlköğretim İngilizce öğretmeni ile 400 8. sınıf öğrencisinin görüşlerine başvurulmuştur. Araştırma verileri araştırmacı tarafından geliştirilen ve hem öğretmenler hem de öğrenciler için hazırlanan anketlerle elde edilmiştir. İki ankette birbiriyle benzerlik göstermekte ve

toplam 40 maddenin 20 maddesini dil becerilerinin öğretimi için kullanılan etkinlikler oluştururken, 20 maddesini de ölçme değerlendirme uygulamaları oluşturmaktadır. Araştırmanın verilerine göre, İngilizce öğretiminde ve ölçme değerlendirme sürecinde okuma becerilerine diğer becerilere oranla daha fazla odaklanıldığı saptanmıştır. Ayrıca dil becerilerinin değerlendirilmesi için kullanılan teknikler ile öğretmenlerin cinsiyeti, mezun oldukları fakülteler ve mesleki kıdemleri arasında anlamlı ilişki olduğu belirlenmiştir.

Ayas (2014) tarafından yapılan çalışmada, İngilizce öğretmenlerinin ölçme değerlendirme süreciyle ilgili kavrayışları ve ölçme değerlendirme yöntem ve tekniklerine dair tercihleri araştırılmıştır. Osmaniye'deki ilkokul ve ortaokullarda görev alan 43 İngilizce öğretmeninden veriler elde edilmiştir. Araştırmada veri toplama aracı olarak Brown (2008) tarafından geliştirilen "Öğretmenlerin Ölçme ve Değerlendirme Sürecine İlişkin Kavrayışları" ölçeğinin 27 maddelik kısaltılmış formu ve bir açık uçlu sorunun yer aldığı anket kullanılmıştır. Araştırma sonuçlarına göre, öğretmenlerin değerlendirmeyi öğrencilerini başarı düzeylerini belirlemek için kullanma ve değerlendirmenin okul başarısını gösterdiğine dair görüşleri destekledikleri belirtilmiştir. Diğer taraftan değerlendirmenin okul başarısını gösterdiği fikrine de karşı çıkmışlardır. Ölçme değerlendirme araçlarına ilişkin veriler incelendiğinde ise, öğretmenlerin en çok kullandıkları ölçme araçları şu şekilde listelenmiştir: Sözel soru ve cevaplar, standart testler, yazılı yoklamalar, portfolyolar, öz ve akran değerlendirmeler ve kontrol listeleri. Ayrıca araştırmanın önemli bulgularından biri de öğretmenlerin alternatif ve biçimlendirici değerlendirmeyi daha fazla kullanmak istemelerine rağmen uygulamada geleneksel ölçme ve değerlendirme araçlarını kullanmak zorunda kaldıklarına dair bildirilen görüşlerdir.

Çiçen'in (2014) değerlendirme yöntemlerine dair yürüttüğü çalışmada, özel ilkokullarda çalışan İngiliz öğretmenlerinin kullandığı değerlendirme yöntemlerini ve neden bu yöntemleri tercih ettiklerini belirlemeyi amaçlamıştır. Karma araştırma desenine uygun olarak yürütülen bu çalışmaya Gaziantep'teki özel okullarda görev alan 42 İngilizce öğretmeni katılmıştır. Araştırmada bilgi toplamak için kullanılan ve araştırmacı tarafından geliştirilen ankette hem çoktan seçmeli sorulara ve açık uçlu sorulara yer yerilmiştir. Araştırma bulgularına göre, öğretmenlerin daha çok öğrenci gelişimini takip etmek ve onların motivasyonu arttırmak amacıyla değerlendirme yaptıkları ve en fazla tercih ettikleri değerlendirme yönteminin de portfolyo olduğu ortaya çıkmıştır.

Gökçe (2014) tarafından yapılan araştırmada, İngilizce öğretmenlerinin biçimlendirici değerlendirmeye yönelik algılarının okul türüne göre karşılaştırılması amaçlanmıştır. Araştırmanın örneklemini ilkokul, ortaokul ve lise olmak üzere toplam 39 kamuya bağlı okul ve özel okullarda görev alan 100 öğretmen oluşturmaktadır. Araştırmada nitel ve nicel araştırma desenleri birlikte kullanılmıştır. Araştırmanın bulgularına göre, İngilizce öğretmenlerinin çoğu biçimlendirici değerlendirmeyi öğrencilerin öğrenme sürecine yaptığı katkılardan dolayı önemli görmektedir. Bununla birlikte, özel okullarda çalışan öğretmenlerin biçimlendirici değerlendirmeyi daha etkili kullandıkları sonucuna ulaşılmıştır. Özal okullarda görev alan öğretmenlerin yarısından fazlası biçimlendirici değerlendirmenin öğrenmeyi teşvik ettiğini vurgulamış ve daha çok kendi öğretim uygulamalarını geliştirmek ve yeniden tasarlamak için bu değerlendirme yöntemine başvurduklarını belirtmişlerdir.

Kayacan ve Eskicumalı (2014) tarafından yapılan araştırmada, ortaöğretim kurumlarında görev yapan İngilizce öğretmenlerinin test tekniklerini kullanma sıklıkları ile bu araçların hazırlanması ve sağladığı faydalara ilişkin görüşlerini belirlemek amaçlanmıştır. Ayrıca öğretmenlerin test tekniklerini kullanma sıklığı ile faydalılığına ilişkin görüşleri arasında anlamlı bir ilişki olup olmadığını saptamak hedeflenmiştir. İlgili alanyazı ışığında hazırlanan anket çevrimiçi (online) olarak uygulanmış ve 92 İngilizce öğretmenine ulaşılmıştır. Yapılan betimsel istatistiklere göre, öğretmenlerin en sık soru-cevap tekniğini kullandıkları, dikte tekniğini ise an kullandıkları ortaya çıkmıştır. Ayrıca elde edilen bulgulara göre, hazırlanması en kolay tekniğin çeviri olduğu, çoktan seçmeli testlerin ise hazırlanması en zor olan teknik olduğu belirlenmiştir. Bununla birlikte yapılan T-testi ikili karşılaştırma istatistikleri ile her tekniğin kullanım sıklığı ile faydalılık derecesi karşılaştırılmış ve elde edilen korelasyonların ölçme teknikleri ile faydalılığı arasında anlamlı bir farkın olduğunu göstermiştir. Sonuç olarak bu çalışmanın yabancı dilde kullanılan test teknikleri ve faydalılığına ilişkin sunduğu bulguların, öğretmenlerin ölçme tekniklerinin önemine dair algı ve tutumlarını geliştirmesi için katkı sunacağı düşünülmüştür.

Öz'ün (2014) yaptığı çalışmada İngilizce öğretmenlerinin ölçme değerlendirme tercihleri cinsiyet, mesleki kıdem, okul türü değişkenlerine göre incelenmiştir. Çalışmaya 120 öğretmen çevrimiçi (online) bir anket doldurarak katılmıştır. Elde edilen bulgulara göre İngilizce öğretmenlerinin geleneksel ölçme ve değerlendirme araçlarını alternatif

değerlendirmeye göre daha fazla kullanmaya meyilli oldukları sonucuna ulaşılmıştır. Geleneksel ölçme değerlendirme araçları arasında ise en sık başvurdukları araçlar boşluk doldurma, çoktan seçmeli testler, doğru-yanlış testleri, eşleştirmeler ve kısa cevaplı sorular olmuştur. Öğretmenlerin öğrenci öğrenmelerini takip etmek amacıyla yaptıkları değerlendirme uygulamalarının cinsiyet, mesleki kıdem ve okul türü değişkenlerine göre anlamlı olarak farklılaştığı tespit edilmiştir.

Bayram (2015) çalışmasında İngilizce öğretmen adaylarının İngilizce dersleri için ölçme ve değerlendirme tercihlerini ve bu tercihleri yapmaya yönelten faktörleri araştırmıştır. Karma araştırma yöntemine uygun olarak yürütülen bu çalışmada; nicel veriler için "Ölçme Teknikleri Farkındalık Anketi", "Ölçme Teknikleri Kullanım Anketi" ve "Ölçme Teknikleri Tercihi Anketi" kullanılırken, nitel veriler odak grup görüşmesi tekniğiyle elde edilmiştir. Araştırmanın örneklemini Karadeniz Teknik Üniversitesinde öğrenim gören 326 İngilizce öğretmeni adayı ve bu katılımcılardan odak grup görüşmesi için rastgele seçilen 38 kişi oluşturmaktadır. Araştırma verilerine göre, İngilizce öğretmeni adaylarının sınıf seviyesi, cinsiyet ve kişilik özelliklerine göre derslerde farklı ölçme teknikleriyle değerlendirilmeyi talep ettikleri ortaya çıkmıştır. Bu tercihlere etki eden faktörlerin ise sınıfların kalabalık oluşu, ulusal sınavlar, sınav kaygısı, bireysel farklılıklar gibi sebeplere dayandığı belirlenmiştir. Araştırmanın nitel verileri incelendiğinde ise, katılımcıların büyük bir kısmı dil performansına yönelik gerçek sonuçlar ortaya çıkardığına inandıkları ürün dosyası, çeviri, sunum ve çoktan seçmeli testlerle değerlendirme yapılmasını istediklerini ifade etmişlerdir. Ayrıca öğretmen adayları öğretim elemanlarından ölçme değerlendirmeyle ilgili daha fazla bilgi ve detaylı geribildirim almak istediklerini belirtmişlerdir.

Toprak (2015) tarafından yapılan bir araştırmada bilişsel tanılayıcı bir test geliştirmek ve "Bilişsel Sınıflandırma Modeli" ve "Logaritmik-doğrusal Bilişsel Tanı Modelini" kullanarak analiz edilen verileri test etmek amaçlanmıştır. Araştırmaya İngilizce Öğretmenliği Bölümünde öğrenim gören 1058 öğrenci katılmıştır. Araştırmacı karma araştırma yöntemini kullanmış ve veri toplama aracı olarak "İkinci Dilde Okuma Becerisinin Bilişsel Tanılayıcı Değerlendirilmesi" testini kullanmıştır. Bu test 5 okuma parçasından ve 27 çoktan seçmeli testten oluşmuştur. Testin pilot uygulaması yapılmış ve katılımcılara öz değerlendirme yaptıkları bir anket verilmiştir. Bununla birlikte, İkinci Dilde Okuma Becerisinin Bilişsel Tanılayıcı Değerlendirilmesi testinden daha iyi

performans gösteren ve daha az performans gösteren gruplardan 15 öğrenciyle sesli okuma protokolü gerçekleştirilmiştir. Logaritmik-doğrusal Bilişsel Tanı Modeli çerçevesinde yürütülen araştırmada hesaplamalar M Plus kullanılarak yapılmıştır. Ayrıca sözdizimsel analizler SAS 9.4. kullanılarak gerçekleştirilmiştir. Araştırma sonuçları bu testin amacına uygun ve test maddelerinin kalitesinin makul düzeyde kabul edilebilir olduğu sonucuna ulaşılmıştır. Bunun yanında testin yetenekli ve daha az yetenekli okuyucuları ayırt ettiği belirlenmiştir. Testin yapı geçerliği ve tanılayıcı dönüt sunması bakımından da iyi düzeyde olduğu saptanmıştır. Sonuç olarak bu araştırmayla öğrenenlerin okuma yeterlikleri ve güçsüzlüklerini tespit etmek için uygun ve geçerli bir test sunulmuştur.

Kırkgöz, Babanoğlu ve Ağçam (2017) tarafından yapılan çalışmada ilkokul ve ortaokullarda görev alan İngilizce öğretmenlerinin değerlendirme algıları ve sınıflarda kullandıkları ölçme değerlendirme yöntemleri araştırılmıştır. Çalışmaya 42 devlet okulundan 4. 5. ve 6. sınıfta öğretmenlik yapan 56 İngilizce öğretmeni katılmıştır. Veri toplama aracı olarak anket kullanılmıştır. Araştırmanın sonuçlarına göre, hemen hemen bütün öğretmenler yabancı dil sınıflarında değerlendirme yapmanın önemini vurgulamışlardır. Öğretmenlerin değerlendirme yapma amaçları şu şekilde sıralanmıştır: öğrencilerin yabancı dil gelişimini takip etmek, öğretim programındaki kazanımlara ulaşılıp ulaşılmadığını belirlemek, hem öğretmen hem de öğrencilerin performanslarını değerlendirmek ve öğretimin kalitesini arttırmak. Öğretmenlerin en fazla kullandıkları değerlendirme araçları ise, kalem-kağıt testleri, sınıf gözlemleri ve performans görevleridir. Araştırmanın önemli bulgularından biri de İngilizce öğretmenlerinin öğrenci başarısını değerlendirmek için performansa ve iletişime dayanan değerlendirme türlerini daha fazla ettikleridir. Diğer taraftan sözlü sınavlar, sunum, portfolyo, proje ve ara sınav gibi araçların ise daha az sıklıkla tercih edildiği kaydedilmiştir.

Mede ve Atay (2017) tarafından dil değerlendirme okuryazarlığı üzerine yapılan bir çalışmada, Türk yabancı dil öğretmenlerinin ölçme değerlendirme yeterlik ve uygulamalarını belirlemek amaçlanmıştır. Çalışmaya 4 devlet ve 7 özel olmak üzere toplam 11 üniversitenin hazırlık programlarında görev alan 350 İngilizce öğretmeni katılmıştır. Veri toplamak için Vogt ve Tsagari (2014) tarafından hazırlanan anket uyarlanmış ve her grupta 10-12 gönüllü katılımcının olduğu odak grup görüşmeleri yapılmıştır. Elde edilen bulgulara göre, öğretmenlerin dil değerlendirme okuryazarlığının sınırlı olduğu ve ölçme değerlendirmenin özellikle sınıf temelli değerlendirme olmak üzere

birçok alanında eğitime ihtiyaç duydukları saptanmıştır. Ayrıca öğretmenlerin dil beceri alanlarının değerlendirmesinde kendilerini yeterli görmedikleri; dilbilgisi ve sözcük bilgisini test etmede ise daha yeterli oldukları kaydedilmiştir.

Ölmezer-Öztürk'ün (2018) doktora tezinde öğretmenlerin yabancı dilde ölçme değerlendirme bilgisini ölçmek için "Dilde Ölçme Değerlendirme Bilgisi Ölçeğini"ni geliştirmek ve yükseköğretim kurumlarında görev alan İngilizce öğretmenlerinin dilde ölçme ve değerlendirme bilgilerinin ne düzeyde olduğunu belirlemek amaçlanmıştır. Yapılan geçerlik ve güvenirlik analizlerinden sonra ortaya çıkan 4 boyutlu ve 60 maddeli ölçeğin mükemmel model veri uyumuna sahip ve güvenirliğin yüksek olduğu saptanmıştır. Araştırmanın verileri için yükseköğretimde çalışan 542 katılımcının görüşüne başvurulmuştur. Elde edilen bulgulara göre, katılımcıların ölçeğin genelinden 60 üzerinden 25 aldığı ve bu ortalamanın toplam puanın yarısından anlamlı olarak düşük olduğu verisine ulaşılmıştır. Ayrıca ölçme değerlendirme ofisinin çalışanı olup olmama değişkeni dışında hiçbir değişkeninin öğretmenlerin dilde ölçme değerlendirme bilgileri üzerinde bir etkisinin olmadığı ortaya çıkan diğer bulgular arasındadır.

İleri (2019) tarafından yapılan çalışmada, devlet okullarında görev alan İngilizce öğretmenlerinin konuşma becerilerinin değerlendirilmesi için yaptıkları uygulamalar ve yeterlik düzeyleri belirlenerek, ihtiyaç duyulan alanların dahil edildiği bir hizmet içi eğitim taslağı oluşturulması hedeflenmiştir. Bu hedef doğrultusunda veri toplamak için çevrimiçi (online) bir anket hazırlanmış ve öğretmenlerle yüz yüze görüşmeler yapılmıştır. Araştırmaya 69 farklı şehirde görev yapan 519 İngilizce öğretmeni katılmıştır. Örneklemi oluşturan katılımcılardan 505'i anketi çevrimiçi olarak yanıtlamış ve 14 katılımcıyla da yüz yüze görüşmeler yapılmıştır. Elde edilen bulgulara göre, İngilizce öğretmenlerinin mesleğe başladıkları ilk yıllarda özellikle konuşma becerisinin değerlendirilmesi olmak üzere ölçme ve değerlendirme alanında kendilerini yetersiz gördükleri ancak mesleki deneyimleri arttıkça bu alanda gelişim gösterdikleri sonucuna ulaşılmıştır. Ayrıca üniversitelerde İngilizce Öğretmenliği Bölümünde ders olarak okutulan "Ölçme ve Değerlendirme" dersinin istenilen düzeyde verim sağlayamadığı ve öğretmen adaylarının "Öğretmenlik Uygulaması" dersinde ölçme değerlendirmeyle ilgili uygulamalara dahil edilmediği araştırmanın diğer bulguları arasındadır. Sonuç olarak, konuşma becerilerinin değerlendirilmesiyle ilgili sunulacak hizmet içi eğitimlerin öğretmelerin bu alandaki eksikliklerini gidermeleri için faydalı olacağı önerilmiştir.

Çakır ve Genç'in (2021) karşılaştırmalı bir analiz yaptıkları çalışmada, Türkiye, İtalya ve Finlandiya'da çalışan yabancı dil öğretmenlerinin 4. sınıf öğrencileri için kullandıkları etkinlikler ve ölçme değerlendirme araç ve uygulamaları ile ilgili görüşlerini tespit etmek amaçlanmıştır. Araştırmaya Türkiye'de 89, İtalya'dan 20 ve Finlandiya'dan 41 olmak üzere toplam 170 İngilizce öğretmeni katılmıştır. Öğretmenlere sosyal ağlar ve öğretmen forumları aracılığıyla ulaşılmıştır. Veri toplamak için, Brumen vd. (2009) tarafından geliştirilen ölçek araştırmanın amacı doğrultusunda uyarlanarak kullanılmıştır. Frekans, yüzdelik, ortalama, ve standart sapma gibi betimsel istatistik teknikleri ve tek-yönlü ANOVA kullanılarak verilerin analizi yapılmıştır. Elde edilen bulgulara göre, Türk, İtalyan ve Finlandiyalı öğretmenler İngilizceyi yabancı dil olarak öğrenen 4. sınıf öğrencileri için daha öğrenci merkezli ve biçimlendirici değerlendirmeye dayalı tekniklerin kullanılması gerektiğine inanmaktadırlar. Ayrıca araştırma bulguları öğretmenlerin öğrencileri sınamak için farklı tür alıştırma yapma ihtiyacı, becerilere odaklanma ve değerlendirme yöntemlerine ilişkin görüşlerinin ülkelere göre değişmediğini göstermektedir. Ancak öğretmenlerin genel değerlendirme algıları ülkelere göre anlamlı olarak farklılaşmaktadır. İtalyan öğretmenlerin değerlendirme sürecine ilişkin genel algılarının Finlandiyalı ve Türk öğretmenlere kıyasla daha yüksek olduğu ortaya çıkmıştır.

Kalay ve Büyükkarcı (2020) tarafından yapılan araştırmada, ortaokul düzeyindeki öğrencilerin yazma becerilerini geliştirme ve değerlendirmek için İngilizce öğretmenleri tarafından kullanılan teknikler araştırılmıştır. Hem nitel hem de nicel verilerin toplandığı bu çalışmaya, özel ve kamuya bağlı ortaokullarda görev alan 97 İngilizce öğretmeni katılmıştır. İlk olarak yazma tekniklerinin nasıl öğretildiğiyle ilgili bir değerlendirme yapılmış ve ardından öğretmenlerin yazma becerilerini değerlendirme yöntemleri incelenmiştir. Ayrıca rastgele seçilen 6 İngilizce öğretmeniyle daha detaylı bilgi toplamak istenmiştir. Nicel araştırma bulgularına göre, öğretmenlerin öğrencilerin yazma becerilerini geliştirmek için daha çok öğretmen-merkezli ve doğrudan öğretim yöntemlerini kullandıkları belirlenmiştir. Diğer taraftan yazma becerilerinin nasıl değerlendirildiği incelendiğinde, öğretmenlerin daha çok yazılıya dayalı değerlendirme tekniklerini kullandığı ve verdikleri puanların daha mekanik ve sınava dayalı olduğu kaydedilmiştir. Araştırmanın nitel bulguları göz önünde bulundurulduğunda ise, öğretmenlerin yazma becerilerini öğretme ve değerlendirmek için kullandıkları tekniklerin zaman, motivasyon eksikliği, ders kitaplarının içeriği ve hizmet içi eğitim faaliyetlerinin yetersiz oluşundan etkilendiği ortaya çıkmıştır.

Önalan ve Gürsoy (2020) tarafından yapılan bir çalışmada, özel ya da kamuya bağlı ortaokullarda görev alan İngilizce öğretmenlerinin kendi öğretim süreçlerinde değerlendirme ve değerlendirme sonuçlarını nasıl kullandıklarını belirlemek amaçlanmıştır. Araştırmanın örneklemini oluşturan 348 İngilizce öğretmeninden çevrimiçi (online) olarak veriler elde edilmiştir. Veri toplama aracı olarak araştırmacılar tarafından geliştirilen ve geçerlik ve güvenirlik analizleri istatistiksel olarak yapılan "Dil Öğretmenlerinin Değerlendirme ve Değerlendirme Sonuçlarına İlişkin Anlayışları" ölçeğinden yararlanılmıştır. Ölçeğin ilk bölümünde katılımcıların demografik bilgilerine ilişkin sorular yer alırken; ikinci bölümde onların değerlendirme sürecine ilişkin algılarını sorgulayan 5'li Likert tipi bir ölçek yer almıştır. Veriler ölçeğin öğretmenlerin farklı değerlendirme yöntemlerini kullanmaları, öğrencilerin değerlendirme sürecine katılımı, dil becerilerini değerlendirme bilgisi ve öğretmenlerin değerlendirme sonuçlarını kullanması olmak üzere ölçeğin dört temel boyutuna uygun olgun olarak analiz edilmiştir. Araştırma sonuçlarına göre, özel okullarda görev alan öğretmenlerin kamu okullarında çalışan öğretmenlere göre değerlendirme yöntemlerinin kullanılmasına ilişkin daha fazla olumlu görüş bildirdikleri saptanmıştır. Bu durum özel okullarda yabancı dil derslerinin daha fazla olması ile açıklanmıştır. Ayrıca ölçeğin tümünde değerlendirme sürecine yönelik olumlu algılar saptansa da öğrenci katılımı ve değerlendirme yöntemleri boyutlarında orta düzeyde bir görüş birliği olduğu ve mesleki tecrübe olarak deneyimli öğretmenlerin daha olumlu tutumlar geliştirdikleri belirlenmiştir. Sonuç olarak alternatif değerlendirme yöntemleri önerilmiştir.

Ballıdağ ve İnan-Karagül (2021) tarafından yapılan bir başka çalışmada Türk İngilizce öğretmenlerinin dil değerlendirme okuryazarlığı araştırılmıştır. Çalışmaya Türkiye'nin farklı şehirlerindeki ilkokul, ortaokul ve lise düzeyindeki devlet okullarında çalışan İngilizce öğretmenleri katılmıştır. Karma yöntem araştırmasına uygun olarak yürütülen bu çalışmada nicel veriler anket aracılığıyla, nitel veriler de görüşmeler de toplanmıştır. Araştırmada ölçme aracı olarak Vogt ve Tsagari (2014) tarafından hazırlanan ve dilde ölçme ve değerlendirme, sınıf odaklı ölçme ve değerlendirme ve ölçmenin amaçları ve dilde ölçme ve değerlendirmenin içerik ve kavramları olmak üzere üç bölümden oluşan "Öğretmen Anketi" kullanılmıştır. Nitel veriler için ise, 18 gönüllü öğretmenle yarı yapılandırılmış görüşme yapılmıştır. Araştırmanın nicel bulgularına göre, yabancı dilde ölçme ve değerlendirmenin üç boyutunda da katılımcıların kendi eğitim seviyelerini yetersiz gördükleri ve bu alanların geliştirilmesi içi hizmet içi eğitime ihtiyaç

duydukları belirlenmiştir. Ayrıca nitel bulgular da katılımcıların dil değerlendirme okuryazarlığı konusundaki yetersizliklerine ilişkin nedenleri sunarak, nicel araştırma bulgularını desteklemektedir.

Son olarak Tokur-Üner ve Aşılıoğlu (2022) tarafından yapılan araştırmada İngilizce öğretiminde ölçme ve değerlendirme sürecine ilişkin öğretmen görüşlerini belirlemek amaçlanmıştır. Araştırmaya Diyarbakır il merkezindeki ortaokullarda görev alan 45 öğretmen katılmış ve veriler yarı yapılandırılmış görüşme formuyla elde edilmiştir. Betimsel analiz süreçleri sonunda elde edilen verilere göre, katılımcıların geleneksel ölçme değerlendirme uygulamalarını daha sık kullandıkları ve özellikle konuşma ve dinleme becerilerinin değerlendirilmesine yönelik yürütülen faaliyetlerin yetersiz kaldığı kaydedilmiştir. Bununla birlikte, ölçme değerlendirme sürecinde yaşanan sorunların özellikle ders kitaplarının yetersizliği, sınıf mevcutlarının kalabalıklığı ve yetersiz fiziki ve teknolojik donanımdan kaynaklandığı tespit edilmiştir. Bu bakımdan öğretim materyallerinin görsel ve işitsel olarak zenginleştirilmesi ve teknolojiyle bütünleştirilerek kullanılması önerilmiştir.

Dil öğrenimi ve öğretiminin önemli bileşenlerinden olan ölçme değerlendirme, içerisinde pek çok öğeyi barındıran oldukça kapsamlı bir süreçtir. Yabancı dilde ölçme değerlendirmeye ilişkin yurt içinde yapılan çalışmalar incelendiğinde; araştırmaların İngilizce öğretmenlerinin ölçme değerlendirme algılarını saptama, yabancı dilde kullanılan ölçme değerlendirme yöntem ve araçlarını belirleme, dil beceri alanlarını değerlendirme, ölçme aracı geliştirme ve dil değerlendirme okuryazarlığına ilişkin ölçme değerlendirmenin farklı boyutlarına vurgu yaptığı söylenebilir. Ortaya konulan bu çalışmalarla genel olarak Türkiye bağlamında yürütülen dilde ölçme değerlendirme faaliyetlerinin dikkat çeken yönleri, var olan eksiklerin belirlenmesi ve çözüm önerileri bildirilerek bu alanın geliştirilmesine katkı sunmak amaçlanmıştır. Ayrıca özellikle 2012-2022 yılları arasındaki araştırmalar sunularak, geçirilen bu 10 yıllık süreçte ölçme değerlendirme anlayışında değişen trendlerin öğrenenler ve uygulayıcılar tarafından nasıl deneyimlendiğini sunmak amaçlanmıştır.

Alanyazındaki araştırmalar incelendiğinde, araştırmacıların yöneldiği alanlardan biri, ölçme değerlendirme sürecinin İngilizce öğretmenleri tarafından nasıl algılandığını ortaya koymak olmuştur. Bu konuda yapılan pek çok araştırmayla (Şahinkarakaş, 2012; Ayas, 2014; Kırkgöz, Babanoğlu ve Ağçam, 2017; Çakır ve Genç, 2021) İngilizce

öğretmenlerinin ölçme değerlendirmeye yönelik tutumları ve algıları cinsiyet, mesleki kıdem, okul türü gibi çeşitli değişkenler göz önünde bulundurularak incelenmiştir.

Araştırmaların yoğunlaştığı önemli alanlardan biri de İngilizce öğretmenlerinin ölçme değerlendirme uygulamalarını nasıl gerçekleştirdiği, hangi yöntem ve tekniklere başvurdukları ve onları bu tercihlere yönelten nedenlerin neler olduğudur. Bu sorulara yanıt aramak için yapılan araştırmalar incelendiğinde (Ayas, 2014; Çiçen, 2014; Gökçe, 2014; Kayacan ve Eskicumalı, 2014; Öz, 2014; Bayram, 2015; Kırkgöz, Babanoğlu ve Ağçam, 2017; Önalan ve Gürsoy, 2020; Çakır ve Genç, 2021; Tokur-Üner ve Aşılıoğlu, 2022); ilkokul, ortaokul, lise ve üniversite düzeyindeki okullarda görev alan İngilizce öğretmenlerinin hangi ölçme ve değerlendirme yöntemlerini tercih ettikleri, test tekniklerini ne sıklıkla kullandıkları, ölçme araçlarının hazırlanması ve uygulanması süreçlerinde yaşadıkları güçlükler kaydedilmiştir. Araştırmalarda cinsiyet, mesleki kıdem, okul türü gibi değişkenler göz önünde bulundurularak yapılan analizlerin yanı sıra; Çakır ve Genç'in (2021) araştırmasında Türkiye, İtalya ve Finlandiya'dan İngilizce öğretmenlerinin görüşlerine başvurularak karşılaştırmalı analizler yapılmıştır. Ayrıca Önalan ve Gürsoy'un (2020) çalışmasında öğretmenlerin değerlendirme tercihlerinin yanı sıra değerlendirme sonuçlarını da nasıl kullandıkları araştırılarak çalışmaya farklı bir bağlam kazandırılmıştır.

Dil beceri alanlarının değerlendirilmesi ile ilgili ulaşılan çalışmaların diğer araştırmalara göre nispeten daha sınırlı olduğu söylenebilir. Bu konuda yapılan çalışmalar incelendiğinde, Yıldırım'ın (2012) dört temel dil beceri alanındaki öğretim uygulamaları ile ölçme değerlendirme faaliyetleri arasındaki ilişkiyi irdelediği görülmektedir. Dil becerilerine bütünsel bir yaklaşım sunan bu araştırmanın yanı sıra bazı araştırmacıların özellikle konuşma (Paker ve Höl, 2012; İleri, 2019) ve yazma (Kalay ve Büyükkarcı, 2020) alanlarında nasıl bir ölçme ve değerlendirme gerçekleştiğine odaklandıkları söylenebilir.

İngilizce öğretiminde ölçme değerlendirme sürecinin nasıl gerçekleştiğine yönelik veri toplamak için araştırmacıların yöneldiği alanlardan biri de ölçme araçları geliştirmek olmuştur. Toprak (2015) tarafından öğrenenlerin okuma yeterlikleri ve güçsüzlüklerini tespit etmek için bir okuma testi geliştirilirken; Ölmezer-Öztürk'ün (2018) çalışmasında ise "Dilde Ölçme Değerlendirme Bilgisi Ölçeği" ile yükseköğretim kurumlarında görev alan İngilizce öğretmenlerinin dilde ölçme ve değerlendirme bilgilerinin ne düzeyde olduğunu

belirlemek amaçlanmıştır. Araştırmacılar tarafından yapılan analizler ölçme araçlarının geçerli ve güvenilir ve test edilmek istenen olgular için uygun olduğunu göstermiştir.

Son zamanlarda yapılan araştırmalarda göze çarpan konulardan biri de öğretmenlerin yabancı dilde ölçme değerlendirme bilgisi ve yeterliğini belirlemek olmuştur. İlk olarak Stiggins (1991) tarafından ortaya konulan "Değerlendirme Okuryazarlığı" kavramı, etkili bir ölçme değerlendirme sürecinin yaşanabilmesi için öncelikle öğretmenlerin bu konudaki yeterliklerinin test edilmesi gerektiğini ifade etmektedir. Zamanla öğretmenlerin ölçme ve değerlendirme alanındaki profesyonelliklerine yönelik yapılan vurguların artmasıyla ise "Dilde Değerlendirme Okuryazarlığı" kavramı ortaya çıkmıştır (Farhady, 2019). O'Loughlin (2013) tarafından test puanlarının yorumlanması ve sonuçların değerlendirilmesi olarak tanımlanan bu kavram; Pill ve Harding'e (2013) göre, bazen dil testleri oluşturma ve analiz yapma gibi pek çok farklı beceriyi de barındırır. Dilde Değerlendirme Okuryazarlığına dair araştırmaların ilgili alanayazında son zamanlarda öne çıkan konulardan biri olduğu görülebilir (Mede ve Atay, 2017; Ölçmezer-Öztürk, 2018; Ballıdağ ve İnan-Karagül; 2021).

2.5.2. Yurt Dışında Yapılan İlgili Çalışmalar

Fulcher (2012) yaptığı araştırmada dil öğretmenlerinin değerlendirme sürecine ilişkin eğitim ihtiyaçlarını tespit etmeyi amaçlamıştır. Öncelikle yapılan pilot çalışma uluslararası alanlarda çalışan 24 dil öğretmeni ile yürütülmüştür. Ardından 278 öğretmen araştırmaya katılmıştır. Araştırma sonuçlarını elde etmek için hem nitel hem de nicel analizler uygulanmıştır. Elde edilen sonuçlara göre, katılımcıların farklı değerlendirme ihtiyaçlarının gerçekten farkında oldukları belirlenmiştir. Ayrıca değerledirme ilkeleri ve uygulamalarının tarihsel ve sosyal bağlamla uyumlu olacak şekilde yürütülmesi gerektiği ifade edilmiştir. Bununla birlikte, sınıf içi değerlendirmeler ve büyük ölçekli sınavların dengeli bir şekilde kullanılması gerektiği araştırmanın diğer bulgularındandır.

Jannati (2012) tarafından yapılan çalışmada İranlı İngilizce öğretmenlerinin dil sınıflarındaki ölçme değerlendirme sürecine yönelik algıları ve uygulamaları araştırılmıştır. Çalışmada İngilizce dil kurumlarında çalışan toplam 18 İngilizce öğretmeni mesleki kıdemleri bakımından düşük, orta ve yüksek olarak kategorize edilmiştir. Araştırmanın amacını gerçekleştirmek için öğretmenlerle görüşmeler yapılmıştır. Yapılan içerik analizi sonunda tüm öğretmenlerin dil öğretimindeki değerlendirme süreciyle ilişkili temel

kavramlar ve teknolojinin farkında oldukları gözlenmiştir. Öğretmenlerin ölçme değerlendirme algılarının mesleki kıdemlerine göre değişmediği kaydedilmiştir. Diğer taraftan farklı mesleki tecrübeye sahip öğretmenler iyi bir değerlendirme yapmanın karakteristik özelliklerini farklı yollara dayandırmışlardır. Bunlar arasında en sık dile getirilen özellik, değerlendirmenin öğrenci başarısını ölçmesi ve kursun amaçlarına dayanarak tasarlanması gerekliliğidir. Öğretmenlerin en yaygın kullandıkları değerlendirme uygulamaları ise soru cevap yapma, dikte etkinlikleri, rol oynama, grup ve akran çalışmalarıdır.

Kirkova-Naskova vd. (2013) tarafından yapılan araştırmada öğretmenlerin İngilizce telaffuz becerilerinin değerlendirilmesine yönelik uygulamaları ve profesyonel eğitimlere dair görüşleri incelenmiştir. Araştırma için Avrupa'da İngilizce Telaffuz Öğretimi Merkeziyle ortak bir çalışma yürütülmüş ve bu merkezde çalışanlarla birlikte toplam 10 araştırmacı geniş kapsamlı çevrimiçi bir anket hazırlamıştır. Anketi yedi Avrupa ülkesinden 630 öğretmen yanıtlamıştır. Araştırma sonuçlarına göre, öğretmenler aldıkları eğitimleri genel olarak eleştirmiş ve bu eğitimlerin onlaın telaffuz becerilerini geliştirdiğini ancak teleaffuz becerilerini öğretmek için eksik kaldığını ifade etmişlerdir. Telaffuz becerilerinin değerlendirilmesiyle ilgili verilere göre ise, öğretmenler sınıftaki uygulamaların öğretim materyallerinde ya da ilgili alanyazında önerildiği gibi gerçekleşmediğini belirtmişlerdir. Ayrıca araştırma sonuçları öğretmenlerin tanılayıcı değerlendirmeyi çok sık yapmadıklarını göstermektedir. Bu durum öğretmenlerin öğrenenlerin telaffuz becerilerinin analiziyle ilgili detaylı geribildirim sunmak için hazırlıklı olmamaları ile açıklanmıştır.

Matin (2013) yaptığı araştırmada orta ve daha ileri düzeydeki öğrencilerin İngilizce konuşma becerilerini değerlendirmeyi amaçlamıştır. Bu genel amaç doğrultusunda öğrencilerin İngilizce konuşma becerilerini geliştirmek için neden isteksiz olduklarına da cevap aranmıştır. Çalışmaya Dakka şehrindeki üç özel üniversiteden toplam 100 öğrenci rastgele seçilmiştir. Araştımada veri toplamak için tarama modelini betimsel açıklayıcı bir araştırma yöntemi izlemiş ve birincil ve ikincil veriler çalışmayı desteklemek için toplanmıştır. İlk verileri elde etmek için açık ve kapalı uçlu sorularla yapılandırılmış bir anket tasarlanmıştır. İkincil veriler yayımlar, araştırma çalışmaları, dergiler ve çevrimiçi kaynaklardan toplanmıştır. Araştırma bulgularına göre, öğrencilerin ailelerindeki bireylerin eğitimli olması onların konuşma alanındaki gelişimlerini etkilemede önemli bir role

sahiptir. Ancak çoğu öğrencinin onları İngilizce konuşmaya teşvik edecek bir çevrelerinin olmadığı kaydedilmiştir. Ayrıca öğretim programının hedefleri ile uygulamaları arasında bir boşluk olduğu belirtilmiştir. Öğrencilerin İngilizce konuşma isteksizlikleri ise şu üç temel nedenle açıklanmıştır: Birinci neden öğretmenlerin ana dilde konuşmalarından dolayı öğrencilerin sınıfta hiç motive olamamalarıdır. İkincisi, konuşma becerisi değerlendirilmediği ve bir puanlama yapılmadığı için öğrenciler bu beceri alanını geliştirmek için yeteri kadar hevesli değillerdir. Üçüncü olarak ise, öğrenciler konuşma becerilerini geliştirmek için başka kaynaklara ulaşabildiklerinden okullara ya da üniversitelere olan bağlılıkları ve beklentileri azalmıştır.

O'Loughlin (2013) üniversiteye giriş sınavlarından sorumlu yöneticilerin değerlendirme ihtiyaçlarını tespit etmeyi amaçlamıştır. Bu yöneticiler öğrencilerinin %25'ni uluslararası ülkelerden gelenlerin oluşturduğu Avustralya'nın iki büyük üniversitesinde çalışmaktadır. Öğrenciler bu üniversitelere girebilmek için Uluslararası Dil Yeterlik Sınavını (IELTS) geçmek zorundadırlar. Bu araştırmada yöntecilere IELTS kullanımı, değerlendirmesi ve diğer özellikleriyle ilgili sorular yönlendirilerek görüşme yapılmıştır. Araştırma bulgularına göre, yöneticilerin daha fazla değerlendirme yeterliğine sahip olmaları gerektiği ve test puanlarının geçerli ve güvenilir bir şekilde yorumlanması için eğitime ihtiyaç duydukları sonucuna ulaşılmıştır. Bu sonuçlar doğrultusunda, yöneticilerin öğrencilerin üniversiteye giriş kararlarından sorumlu olmalarından dolayı, bu kararları daha iyi bir değerlendirme anlayışıyla yerine getirmeleri önerilmiştir.

Safa ve Goodarzi (2014) tarafından yapılan çalışmada görev temelli dil değerlendirme yöntemiyle uygulanan sınavların, İngilizceyi yabancı dil olarak öğrenenlerin dilbilgisi gelişimi üzerindeki etkilerini incelemek amaçlanmıştır. Araştırmanın örneklemini alt-orta seviyedeki bir dil kurumunda rastgele seçilen 74 öğrenen oluşturmaktadır. Katılımcılar kontrol ve deney gruplarına ayrılmışlardır. Aynı dilbilgisi yeterliğinde olduklarını garantilemek için iki gruba da araştırmaya başlamadan ön test uygulanmıştır. 10 seanslık eğitimler boyunca gruplara aynı dilbilgisi öğretimi uygulanmıştır. Ancak deney grubuna her üç seans da bir araştırmacı tarafından hazırlanan görev temelli dilbilgisi ara sınavı uygulanırken, kontrol grubu geleneksel dilbilgisi ara sınavlarına girmişlerdir. Tüm eğitimler bittikten sonra iki gruba da dilbilgisi testi uygulanmıştır. Analiz sonuçları görev temelli anlayışa uygun yapılan dilbilgisi sınavlarının etkisinin (washback) olumlu olduğunu göstermektedir. Elde edilen bulgulara göre, ölçme

ve değerlendirme süreçlerinin planlı bir şekilde yürütülmesi gibi bütün eğitsel çabaları içinde barındıran görev temelli dil değerlendirmenin klasik değerlendirme süreçlerinin yerine kullanılabilecek pedagojik bir ölçme aracı olduğu sonucuna ulaşılmıştır.

Tao'ya (2014) ait bir çalışmada, öğretmenlerin değerlendirme yeterliklerini tespit etmek amaçlanmıştır. Kamboçya'da görev alan 108 İngilizce öğretmeni araştırmanın örneklemini oluşturmaktadır. Veri toplamak için araştırmanın amacı doğrultusunda geliştirilen 4 farklı ölçek kullanılmıştır. Bu ölçekler sınıf içi değerlendirme bilgisi, yenilikçi yöntemler, öyargıları tespit etme ve sürecin kalitesini ölçmek için tasarlanmış araçlardır. İlk araç öğretmenlerin değerlendirme bilgisini ölçmeyi amaçlayan bir çoktan seçmeli testtir. Geriye kalan son üç araç ise katılımcıların değerlendirme sürecine dair inançlarını tespit etmeyi hedeflemektedir. Bütün ölçeklerin yeterli ve kabul edilebilir ölçme özelliklerine sahip olduğu analiz edilmiştir. Elde edilen nicel verilerle birlikte, nitel veriler de toplanmış ve 6 öğretmenle yarı-yapılandıırlmış görüşmeler gerçekleştirilmiştir. Araştırma bulgularına göre, öğretmenlerin değerlendirme bilgisinin sınırlı olduğu ve bunun da değerlendirme uygulamalarını olumsuz bir şekilde etkilediği sonucuna ulaşılmıştır.

Wewer (2014) yaptığı araştırmada, Finlandiya'daki ilkokullarda ikinci dil sınıflarında öğretmenlerin yürüttüğü değerlendirme uygulamalarını incelemeyi amaçlamıştır. Ayrıca öğrencilerin, anne ve babaların değerlendirmeye yönelik algıları ve bir bilgisayar simulasyon programı (LangPerfom) da incelenmiştir. Karma araştırma deseni çerçevesinde gerçekleştirilen bu çalışmaya çiftdilli öğretim yapan toplam 42 öğretmen, 3.,4. ve 5. sınıf öğrencileri ve 99 anne baba katılmıştır. Veriler deneysel çalışma, öğenciler ve anne babalarla yapılan anketler ve videolu görüşmelerle elde edilmiştir. Araştırmanın bulgularına göre, dil sınıflarında en yaygın şekilde kullanılan değerlendirme aracı öğretmen gözlemleri, diyaloğa dayalı etkileşim ve çiftdilli testlerdir. Diğer taraftan en az kullanılan değerlendirme araçları ise, akran değerlendirme, simulasyonlar ve portfolyolardır. Öğretmenlerin bir kısmı değerlendirmeyi öğrenci erişimleri hakkında bilgi toplamak için kullandığını belirtirken; sadece üçte biri değerlendirmenin önemli olduğunu düşünmelerine rağmen neredeyse hiçbir zaman öğrencilere değerlendirme sonuçları hakkında dönüt sağlamamaktadır.

Britton (2015) yaptığı araştırmada yabancı dil öğretmenlerinin ilkokullardaki değerlendirme uygulamalarını ve bu uygulamaların sınıftaki etkileşim üzerindeki etkisini incelemiştir. Karma araştırma yöntemine uygun olarak yürütülen çalışmada gözlem,

bireysel ve odak grup görüşmeleri ve okul dokümanları gibi farklı veri kaynakları kullanılmıştır. Yapılan içerik analizi sonunda, öğretmenlerin 7-9 yaşındaki öğrenciler ile 10-11 yaşındaki öğrencilerin olduğu sınıflarlarda yürüttükleri değerlendirme faaliyetlerinin amaçlar, beklentiler, öğrenci performansını takip ve başarılarını kontrol etme bakımından farklılaştığı saptanmıştır. Ayrıca öğretmenlerin değerlendirmeyi öğrenme sürecinin bir parçası olarak kullanmalarının işbirlikli bir sınıf ortamının yaratılmasında etkili olabileceği belirtilmiştir. Katılımcıların çocuklara İngilizce öğretiminin temel ilkeleri ve öğrenmenin desteklenmesi için değerlendirme yapılmasına ilişkin görüşlerinin alanyazında yer alan kuramsal çerçeveyle aynı çizgide olduğu tespit edilmiştir.

Chan (2016) yaptığı araştırmada İngilizce öğretmenlerinin değerlendirme uygulamaları, inanç ve bilgileri arasındaki ilişkileri incelemiştir. Ayrıca katılımcıların çoklu değerlendirmeyi kullanmalarının mesleki deneyime göre değişip değişmediğini belirlemeyi amaçlamıştır. Çalışmaya Kuzey Tayvan'dan 520 ilkokul öğretmeni katılmıştır. Veriler anket, çoktan seçmeli test ve açık uçlu sorularla toplanmıştır. Araştırma sonuçları katılımcıların çoklu değerlendirmenin ne olduğu ve hangi süreçleri kapsadığına dair bilgilere açık bir şekilde sahip olduklarını göstermiştir. Elde edilen önemli bulgulardan biri de katılımcıların özellikle portfolyolar olmak üzere çoklu değerlendirmenin etkili olduğuna inanmalarıdır. Bununla birlikte, çoğu öğretmenin geleneksel kalem kağıt testlerindense biçimlendirici değerlendirmeyi kullanmaya daha fazla meyilli oldukları ortaya çıkmıştır. Katılmcılar ya biçimlendirici değerlendirmeyi ya da biçimlendirici ve özetleyici değerlendirmeyi birlikte kullanmak istediklerini ifade etmişlerdir. Ancak onların hiçbiri geleneksel değerlendirmeyi temel değerlendirme türü olarak görmemektedir. Bunun yanında deneyimli katılımcılar ile onların inaçları arasında anlamlı bir ilişkinin olduğu sonucuna ulaşılmıştır.

Halali, Singh ve Making (2017) tarafından yapılan çalışmada ikinci dilde okuma dersi bağlamında öğretmenlerin kullandıkları sınıf içi değerlendirmeye yönelik algılarını belirlemek amaçlanmıştır. Çalışmaya ilkokullarda görev alan ve en az bir yıllık mesleki deneyimi olan 28 İngilizce öğretmeni katılmıştır. Veri toplama aracı olarak araştırmacı tarafından geliştirilen anket kullanılmıştır. Araştırmanın bulgularına göre, öğretmenlerin sınıf içi değerlendirme uygulamalarına dair agılarını büyük ölçüde etkileyen bazı iç ve dış faktörler mevcuttur. Bu faktörlerden teknolojik ve fiziki olanaklar, motivasyon, zaman yönetimi, uygun ders kitapları ve bakanlık tarafından sağlanan destek daha gözlenebilir

olanlardır. Öğretmenler sınıf içi değerlendirme faaliyetlerinin onların resmi öğretim uygulamalarını geliştirmeye yardım ettiğini ifade etmişlerdir. Okuma alanında yürütülen değerlendirme sürecinde olumlu iç ve dış faktörlerin öğretmenlerin motivasyonunu arttırdığı kaydedilmiştir.

Sellan (2017) tarafından yapılan çalışmada, öğretmenlerin değerlendirme uygulamalarına ilişkin anlayışlarını derin bir şekilde incelemek amaçlanmıştır. Araştırmanın örneklemini Sinagpur'da görev yapan ve on yıllık deneyime sahip 8 İngilizce öğretmeni oluşturmaktadır. Bu çalışmadaki öğretmenlerin en önemli özelliği sınavlara önem vermeyen ve öğretmen merkezli değerlendirmeyi destekleyen özel bir Bütünleştirilmiş Program (Integrated Programmes) uygulamalarıdır. Veri toplamak için görüşme, gözlem ve doküman analizi yapılmıştır. Araştırmanın bulgularına göre, katılımcıların kültere dikkat ederek, içerik bilgisine oldukça önem vererek ve üst bilişsel becerilere daha fazla odaklanarak kendi değerlendirme yeterliklerini geliştirdiği belirlenmiştir. Ayrıca Bütünleştirilmiş Program uygulamalarının öğrenenlerin değerlendirme ihtiyaçlarının daha fazla farkına varılması için desteklediği ve öğretmenleri değerlendirme sürecinde daha yetkin kıldığı sonucuna ulaşılmıştır.

Tsagari ve Vogt (2017) öğretmenlerin değerlendirme okuryazarlığı algıları ve dilde ölçme değerlendirme sürecine ilişkin bireysel ihtiyaçlarını belirlemeyi amaçlamıştır. Araştırmanın örneklemini, Kıbrıs'tan 16, Yunanistan'dan 22 ve Almanya'dan 25 olmak üzere 63 normal öğretmen oluşturmaktadır. Bu araştırmada normal öğretmen; standart bir eğitim alan, İngilizceyi devlet kurumları, kolejler ve okullarda öğreten ve ek olarak farklı bir değerlendirici rolüne sahip olmayan öğretmen olarak tanımlanmıştır. Karma araştırma çerçesinde yürütülen bu çalışmada nitel araştırmayla elde edilen veri setinin sonuçlarına ağırlık verilmiştir. Yani yarı yapılandırılmış görüşme tekniği araştırmanın büyük bir bölümünü oluşturmaktadır. Araştırma sonuçları öğretmenlerin algıladıkları değerlendirme okuryazarlığı seviyelerinin yeterli olmadığı ve değerlendirme odaklı aktivitelerin etkili bir şekilde uygulanması için kendilerini hazırlıklı görmediklerini göstermektedir. Ek olarak, öğretmen yetiştirme programlarının öğretmen adaylarına verimli bir ölçme ve değerlendirme eğitimi sağlayamamasından dolayı bu programların öğretmen adaylarını gelecekteki kariyerleri için yeterli düzeyde hazırlayamadıkları ifade edilmiştir. Son olarak çoğu öğretmenin yenilikçi değerlendirme uygulamalarındansa test kullanımına daha fazla meyilli oldukları kaydedilmiştir.

Xu ve Brown (2017) yaptıkları araştırmada öğretmenlerin değerlendirme okuryazarlığını belirlemeyi amaçlamışlardır. Bu amaç doğrultusunda Çin'de görev alan 891 İngilizce öğretmeninin görüşlerine başvurulmuştur. Verileri toplamak için "Öğretmenlerin Değerlendirme Okuryazarlığı Anketinin" uyarlanmış ve 24 sorudan oluşan formu kullanılmıştır. Araştırma sonuçları katılımcıların ya temel ya da minimal düzeyde değerlendirme okuryazarlığına sahip olduğunu göstermiştir. Ayrıca psikometrik analizler yapıldıktan sonra zor olarak atfedilen test maddelerinin çoğu katılımcı tarafından doğru bir şekilde yanıtlanamadığı bulgusuna ulaşılmıştır. Bununla birlikte yaş, mesleki deneyim, yeterlik, yaşanılan bölge, değerlendirme eğitimi alma gibi demografik değişkenlerin öğretmenlerin değerlendirme okuryazarlığı üzerinde anlamlı bir etkisinin olmadığı belirlenmiştir.

Joyce (2018) tarafından yapılan araştırmada ikinci dilde kelime öğrenme ve değerlendirmede kelimeleri birinci dile çevirmenin etkisi araştırılmıştır. Çalışma için 48 Japon üniversite öğrencisi akademik kelime listesinden 200 İngilizce kelimeyi 10 haftalık bir periyotta çalışmıştır. Öncelikle katılımcılar A grubu ve B grubu olmak üzere dengeli bir şekilde iki gruba ayrılmıştır. A grubu ilk listedeki 100 kelimeyi İngilizce anlamıyla, ikinci listedeki 100 kelimeyi de Japonca karşılığıyla çalışmıştır. B grubuna da aynı görevler verilmiş ama ilk listeyi Japonca karşılığıyla, ikinci listeyi ise İngilizce tanımıyla çalışmışlardır. Katılımcıların sözcük bilgisine dair yeterlikleri çoktan seçmeli ön-test ve son-test uygulamalarıyla değerlendirilmiştir. Araştırma sonuçları öğrencilerin ikinci dildeki kelimeleri anadildeki karşılıklarıyla eşleştirdikleri zaman anlamlı olarak daha yüksek bir oranla tanıdıklarını göstermektedir. Bununla birlikte dili çalışma ile dili test etme arasında anlamlı bir ilişkinin olduğu saptanmıştır. Yani katılımcılar dili çalışıp hemen test edildikleri zaman anlamlı olarak daha yüksek puanlar elde etmişlerdir.

Saeed, Tahir ve Latif (2018) tarafından yapılan araştırmada öğretmenlerin ilkokul ve ortaokullarda kullandıkları sınıf içi değerlendirme tekniklerini incelemek amaçlanmıştır. 300 ilkokul ve 200 ortaokul olmak üzere rastgele seçilen toplam 500 öğretmen araştırmanın örneklemini oluşturmaktadır. Veri toplama aracı olarak ise 55 maddeden oluşan 5'li Likert tipinde bir anket kullanılmıştır. Anketin iç tutarlılığı .93 olarak hesaplandığından güvenilir bir araç olarak kabul edilmiştir. Araştırmanın sonuçları özel ve devlet okullarında görev alan çoğu öğretmenin biçimlendirici ve düzey belirleyici değerlendirmenin sınıfta öğrenmeyi desteklediğini belirtmelerine rağmen, çoğunlukla

düzey belirleyici değerlendirmeler yaptıklarını göstermektedir. Ayrıca öğretmenler akran ve portfolyo değerlendirme gibi diğer alternatif değerlendirme tekniklerini kullandıklarını ifade etmişlerdir. Sonuç olarak, değerlendirme yöntemlerinin etkili bir şekilde kullanılmasının hem özel hem de devlet okullarındaki öğrencilerin gelişimlerini destekleyeceği belirtilmiştir.

Kim (2019) yaptığı araştırmada, öğretmen ve öğrencilerin okuma alanında kullanılan değerlendirme tekniklerine ilişkin algılarını incelemeyi amaçlamıştır. Araştırmanın örneklemini farklı yeterlik seviyelerine göre sınıflara yerleştirilen 232 üniversite öğrencisi ve 13 profesör oluşturmaktadır. Veriler araştırmacı tarafından hazırlanan bir anket ve gönüllü öğrenci ve profesörlerin katıldıkları görüşmeler yoluyla elde edilmiştir. T-test kullanılarak veriler analiz edilmiş ve iki farklı seviyedeki öğrenci grubu karşılaştırılmıştır. Araştırma sonuçları temel seviyedeki öğrenciler için kısa cevap yazma ve akranlarla ve profesörle tartışmayı içeren değerlendirme tekniklerinin daha fazla kullanıldığı ve öğrencilerin özellikle tartışma tekniğini en etkili değerlendirme tekniği olarak gördüklerini göstermektedir. Diğer taraftan orta seviyedeki öğrenciler için daha çok doğru-yanlış ifadeler ve çoktan seçmeli testlerin değerlendirme için kullanıldığı; ancak öğrencilerin en etkili olarak algıladıkları tekniğin ise bir tema hakkında görüşlerini yazma olduğu belirlenmiştir. Ayrıca profesörlerin kısa cevap gerektiren sorular ve sınıf içi tartışma tekniklerini derecelendirilmiş okuyucular için daha fazla tercih ettikleri ve bir görev yazmanın en yeterli değerlendirme tekniği olduğunu belirtmeleri araştırmanın diğer bulguları arasındadır. Ek olarak, profesörler öğrencileri yeterliklerine göre sınıflandırmanın tüm dil becerilerinin gelişimi için daha faydalı olduğunu ifade etmiş ve dil programlarının bu şekilde yapılandırılmasını desteklediklerini dile getirmişlerdir.

Morales ve Fernandez (2019) tarafından yapılan çalışmada dinleme alanında alternatif değerlendirme yapmanın sağladığı fayda ve çıkarımlar incelenmiştir. Karma araştırma yöntemine uygun olarak yürütülen bu çalışmaya 38 üniversite öğrencisi katılmıştır. Veri toplama araçları olarak gelişim planları, portfolyolar, öz-değerlendirme formları, öğretmen ve öğrenci konferansları kullanılmıştır. Verilerin üçgenleme gibi çeşitli süreçlerle geçerliği sağlanmış ve toplanan bilgilerden elde edilen temalara göre analizi yapılmıştır. Araştırma sonuçlarına göre alternatif değerlendirmeler dinleme alanında öğretimi ve değerlendirmeyi bir araya getirdiği ve diğer becerilerle bütünleşme olanağı sunduğu için daha fazla kullanılmalıdır. Bu araştırma teori ve uygulamalar için de çeşitli

çıkarımlar sağlamış ve gelenekselden süreç değerlendirmeye ve norm-bazlı değerlendirmeden daha ölçüt bazlı değerlendirmeye doğru gelişen bir anlayışı önermiştir. Ancak amaç radikal bir şekilde kalem kağıt testleri gibi geleneksel araçları tamamıyla reddetmek değil, aslında diğer stratejilerle dengeli bir şekilde birleştirilip kullanılmasını sağlamaktır. Böylece değerlendirme süreci yabancı dil sürecindeki tüm aktörler için daha geçerli, güvenilir, adil ve otantik bir yapıya dönüşebilir.

Yurt dışında ölçme ve değerlendirme alanıyla ilgili yapılan çalışmalar incelendiğinde, araştırmanların öğretmen ve yöneticilerin ölçme değerlendirme algıları, inançları ve uygulamalarını tespit etme, öğretmenlerin ölçme değerlendirme araçları konusundaki tercihleri ve yeterliklerini belirleme, dil beceri alanlarının değerlendirilmesi, sınav etkisi ve dilde değerlendirme okuryazarlığı konularına yoğunlaştığı söylenebilir. Öğretmenlerin değerlendirme algılarına yönelik yapılan çalışmalar (Jannati, 2012; Chan, 2016; Sellan, 2017) göz önünde bulundurulduğunda, araştırmacıların öğretmenlerin değerlendirme kavramına yükledikleri anlamın ne olduğunu açığa çıkarmaya çalıştıkları söylenebilir. Pek çok araştırmanın temel konularından biri öğretmenlerin sınıf içi ölçme değerlendirme uygulamalarını tespit etmek olmuştur (Jannati, 2012; Wewer, 2014; Britton, 2015; Chan, 2016; Sellan, 2017; Saeed, Tahir & Latif, 2018). Bu alanda yapılan çalışmalarla öğretmenlerin ölçme değerlendirme uygulamaları mesleki kıdem, cinsiyet, aile gelir, değerlendirme eğitimi alıp almama gibi değişkenler açısından incelenmiştir. Ek olarak Wewer (2014) tarafından yapılan araştırmada, öğretmenlerin yanı sıra öğrenciler ve anne babaların da değerlendirme uygulamalarına yönelik algı ve tutumları incelenmiştir.

Araştırmacıların yöneldiği alanlardan biri de öğretmenlerin değerlendirme okuryazarlıklarını belirlemek ve bu doğrultuda eğitim ihtiyaçlarını saptamaktır (Fulcher, 2012; O'Loughlin, 2013; Xu & Brown, 2017; Tsagari & Vogt, 2017). Ayrıca temel dil beceri alanları ile dilbilgisi, sözcük bilgisi ve telaffuz gibi yapıların değerlendirilmesi ile ilgili çalışmaların yoğun bir şekilde alanyazında yer bulduğu ve her bir becerinin ayrı ayrı incelendiği söylenebilir (Kirkova-Naskova vd., 2013; Safa & Goodarzi, 2014; Halali, Singh & Making, 2017; Joyce, 2018; Morales & Fernandez, 2019).

Yurt içi ve yurt dışında ölçme ve değerlendirme alanında yapılan çalışmalar birlikte incelendiğinde, genel olarak benzer konuların alanyazında tartışıldığı görülebilir. Ancak özellikle dil becerileri ve dil yapılarının değerlendirilmesi konusunda yurt dışında daha geniş kapsamlı çalışmaların yürütüldüğü ve bu alanların her birine özel olarak

odaklanılarak çalışmaların gerçekleştiğini söylemek mümkündür. Yapılan araştırmalar örneklem grupları olarak karşılaştırıldığında, özellikle yurt dışı alanyazındaki bazı araştırmalarda katılımcıların uluslararası alanlardaki farklı öğretim kurumlarında görev alan kişilerden oluştuğu görülmektedir (Kirkova-Naskova vd., 2013; Tsagari & Vogt, 2017). Ayrıca yurt içinde yapılan çalışmalarda öğretmenler değerlendirici rolüyle ön plandayken; yurt dışında yapılan kimi çalışmada ise (O'Loughlin, 2013; Wewer, 2014) öğrenciler, ebeveynler, yöneticiler gibi diğer paydaşların da algı ve tutumlarının değerlendirme sürecinde göz önünde bulundurulduğu söylenebilir.

3.YÖNTEM

3.1. Araştırmanın Modeli

Ortaokullarda İngilizce öğretiminde öğretmenlerin ölçme ve değerlendirme olgusunu nasıl deneyimlediklerini keşfetmeyi amaçlayan bu araştırmada betimsel tarama modeli kullanılmıştır. Bu modelde geniş bir örneklemle yürütülen çalışmalarda gruplardaki bireylerin bir olay ya da olguyla ilgili tutumları, görüşleri veya başarıları betimlenmeye çalışılır (Karakaya, 2012). Karasar'a (2013) göre, bu araştırma modelinde önemli olan araştırma konusu dahilindeki olayları, bireyleri veya nesneleri kendi ortamında olduğu gibi yansıtmaktır.

3.2. Araştırma Deseni

Araştırmada karma araştırma deseni tercih edilmiştir. Bir araştırma paradigması olarak karma yöntem, nitel ve nicel araştırmalar arasında bir köprü kurulmasına yardımcı olmaktadır (Onwuegbuzie & Leech, 2004). Tashakkori ve Teddlie (1998) karma yöntemi, bir çalışma yönteminin nitel ve nicel yaklaşımlarının kombinasyonu olarak tanımlamışlardır. Karma yöntemi "görmenin birden fazla yolu" olarak tanımlayan Creswell ve Clark'a (2015: 5) göre ise bu yöntem, sadece bir araştırma modeli olarak kullanılmanın ötesinde geniş bir uygulama alanı ortaya çıkarmaktadır. Bu desenin kullanılma amacı nitel ve nicel desenlerin avantajlarını kullanarak bir olgunun daha kapsamlı ve detaylı bir şekilde anlaşılmasını sağlamaktır (Mills & Gay, 2016).

Nitel ve nicel desenlerin birlikte kullanıldığı karma yöntemlerde her iki tekniğin de sunduğu avantajların birleştirilmesiyle, araştırma sorularının çoklu yaklaşımlarla incelenmesi sağlanabilir. Byrne ve Humble'a (2007) göre tüm veri toplama yöntemlerinin sınırlılıkları olduğu için çoklu yöntemlerin bir arada kullanılması belirli yöntemlerin dezavantajlarını etkisiz hale getirebilir. Ayrıca her yaklaşımın güçlü tarafları birbirini tamamlayabilir. Karma yöntem araştırmalarının sunduğu pek çok avantaj şu şekilde özetlenebilir (Wisdom & Creswell, 2013):

- Nitel ve nicel veriler karşılaştırılabilir. Özellikle bu iki veri seti arasındaki benzerlik ve farklılıkların ortaya çıkarılmasında kullanışlıdır.
- Katılımcıların fikirlerini ve kendi deneyimlerini yansıtır.
- Yöntembilimsel olarak esneklik sağlar. Kapsamlı araştırma sorularına cevap bulmak için uygundur.

- Nitel ve nicel verilerin birlikte kullanılması daha geniş ve kapsamlı bilgiler üretir.

Karma yöntem desenine uygun şekilde yürütülen bu çalışma yapılandırılırken nitel ve nicel desenler belirli bir amaç doğrultusunda kasıtlı bir şekilde bir araya getirilmiştir. Bryman'a (2006) göre, karma yöntem desenini tercih eden araştırmacıların nicel ve nitel yaklaşımların bütünleştirilmesi veya birleştirilmesine karar verirken hangi gerekçeyle bunu yaptıklarını açıklamaları gerekmektedir. Alanyazında karma yöntem kullanma gerekçeleri hakkında araştırmacılara rehber olabilecek çeşitli tartışmalar mevcuttur. Bunlardan biri Greene, Caracelli ve Graham'ın (1989) çalışmalarında karma yöntem araştırmaları için tanımladıkları beş kapsamlı gerekçeyi içeren listedir. Bu gerekçeler: Üçgenleme, tamamlayıcılık, geliştirme, başlatma ve genişletmedir.

Çeşitleme (Triangulation): Aynı olayı değerlendirmek için nitel ve nicel yöntemlerin birbirinden bağımsız bir şekilde kullanılmasıdır. Böylece farklı yöntem sonuçlarının birbirini doğrulaması veya birbirine yakınlığı araştırılır.

Tamamlayıcılık (Complementarity) : Bir yöntemden elde edilen sonuçlarla diğer yöntemin sonuçlarını detaylandırma, genişletme, örnekleme ve açıklamayı amaçlar.

Geliştirme (Development): Bir yöntemin uygulanmasının ardından ortaya çıkan sonuçların diğer yöntemin örnekleminin seçiminde ya da ölçme aracını geliştirmede yardımcı olmasıdır.

Başlatma (Initiation): Paradoks ve çelişkileri ortaya çıkarmak için nitel ve nicel yöntemlerin birlikte kullanılmasıdır. Bir yöntemin uygulanmasının ardından soruların başka biçimde ifade edilmesinin veya sunduğu bakış açılarının keşfedilmesini ve diğer yöntemde değişiklikler yapılabilmesini amaçlar.

Genişletme (Expansion): Birbirinden farklı olgular için farklı yöntemlerin kullanılmasını amaçlar. Örneğin, bir programın sonuçlarını değerlendirirken nicel araştırma yöntemleri; uygulamaları değerlendirirken ise nitel yöntemler kullanılabilir (Greene, Caracelli & Graham, 1989).

Araştırmacıların belirtilen bu gerekçeleri dikkate alarak karma yöntem desenini yapılandırmaları beklenmektedir. Böylesi bir yapılandırmanın araştırmacılara planlama sürecinde kolaylık sağlayabileceği de söylenebilir (Alkan, Şimşek ve Erbil, 2019). Bu araştırma "Geliştirme" amacına uygun olarak yürütülmüştür. Bu amaçla öncelikle nitel araştırma desenine uygun olarak veriler elde edilmiştir. Yıldırım ve Şimşek'e (2013) göre nitel araştırmalar insan davranışlarına dair deneyimleri ortaya koyarak daha esnek ve

bütüncül bir yaklaşım sunmaktadır. Bu bakış açısından hareketle, yabancı dil öğretiminde ölçme değerlendirme olgusunun öğretmenler tarafından nasıl deneyimlendiğini derin bir şekilde analiz etmek için nitel bir araştırmayla çalışmaya başlanmıştır. Ardından İngilizce öğretmenlerinin yabancı dil beceri alanlarına dair davranışları ne düzeyde gerçekleştirdiklerine dair ayrıntılı bir analiz yapılmak istenmiştir. Bu bağlamda, nitel araştırma verilerine dayanarak uygulanacak anket için maddeler hazırlanmış ve geniş bir örneklem seçilerek nitel verilerin yaygınlığını keşfetmek amaçlanmıştır.

Bu çalışmada olduğu gibi bir olgunun nicel araştırmadan önce nitel olarak keşfedilmeye çalışılması karma yöntem desenlerinden keşfedici sıralı desen olarak adlandırılır. Bu zincirleme desenin birçok uygulamasında, araştırmacı nitel verilere dayalı olarak oluşturulan aşamalar arasında orta adım olarak bir araç geliştirir ve bu aracı nicel verileri elde etmek için kullanır (Creswell, Fetters & Ivankova, 2004). Bu desende araştırmacı ortaya çıkan nitel sonuçları test eder ve nicel verilerin nitel veriler üzerine nasıl inşa edildiğini açıklar (Creswell, 2003).

Keşfedici sıralı desenin birincil amacı, ilk aşamada sınırlı sayıda katılımcıdan elde edilen nitel bulguları ikinci aşamada daha geniş bir örneklem kullanarak araştırma sonuçlarını genellemektir (Greene, Caracelli & Graham, 1989). Bu desen ayrıca araştırmacının istediği bir olguyu derinlemesine analiz etmek ve tekrarlanma sıklığını ölçmek istediği durumlar için de uygundur (Morse, 1991). Creswell ve Clark (2015) keşif deseninin ek olarak aşağıdaki durumlarla da ilişkili olduğunu açıklamaktadırlar:

- Araştırmacı ve araştırma problemi nitel yönelimlidir.
- Araştırmacının çalışmayı iki aşamalı yürütmek için yeteri kadar zamanı vardır.
- Araştırmacı hangi yapıların çalışmak için önemli olduğunu bilmemektedir ve ilgili nicel araçların kullanımı elverişli değildir.
- Araştırmacı, nitel sonuçlara dayanılarak cevaplanamayan ve yeni ortaya çıkan araştırma sorularını tanımlamaktadır.

Keşfedici sıralı desen bir olguyu keşfetmek amacıyla nitel verilerin toplanması ve çözümlenmesi ile başlar. Bir sonraki adımda araştırmacılar nitel aşamadan elde ettikleri sonuçlar üzerinden bir araç geliştirerek, biçimleri tanımlayarak veya ortaya çıkan kuram veya çerçeveye dayalı testler için önermeler saptayarak devam ederler. Böylece nitel bulgular nicel aşama ile ilişkilendirilir. Üçüncü adımda ise araştırmacı geliştirilen aracı kullanarak yeni bir örneklem ile çalışmanın nicel aşamasını uygular. Son olarak, nicel

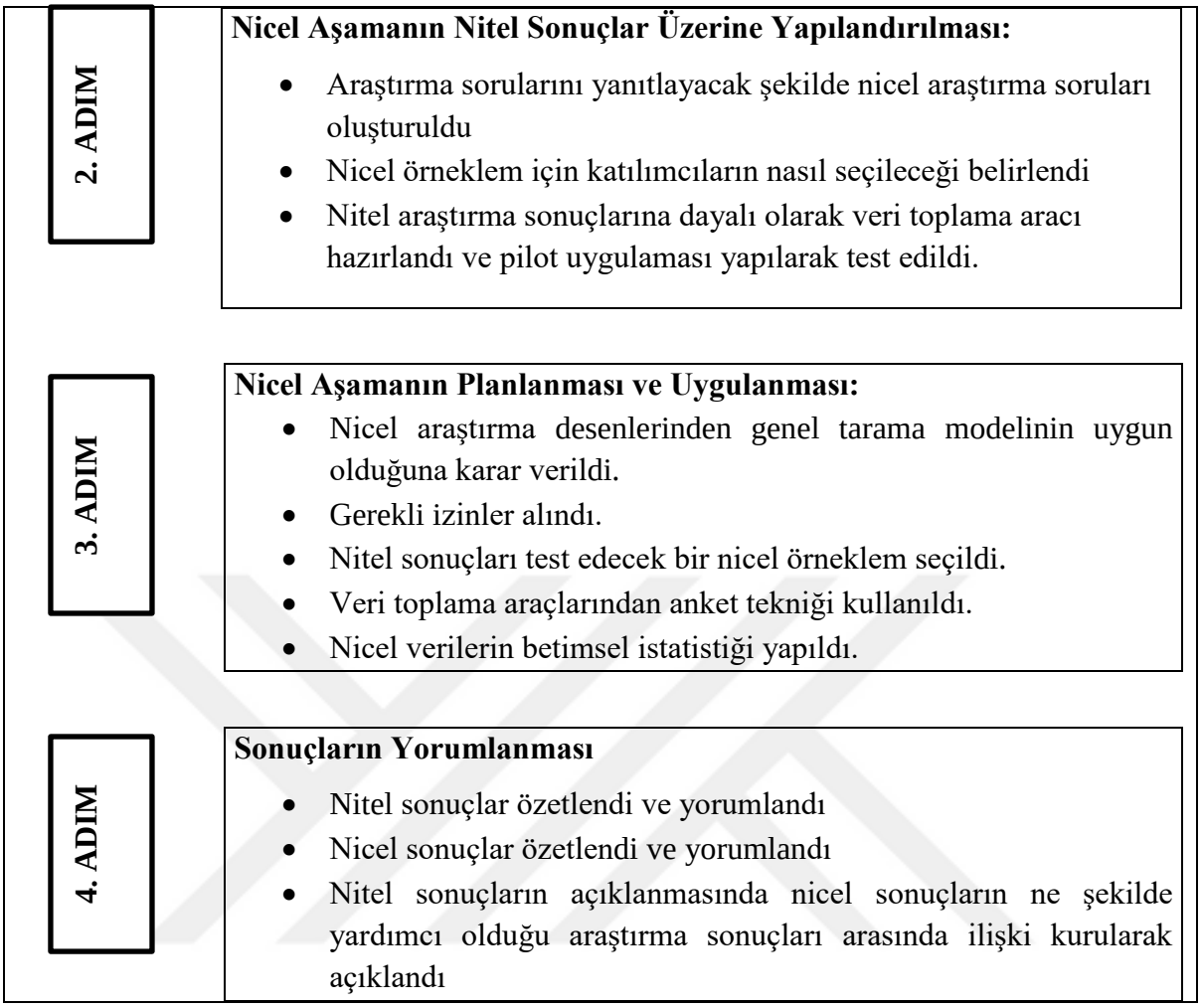
sonuçların nitel bulguları ne şekilde ve düzeyde genelleştirebildiğini veya genişletebildiğini yorumlar (Creswell ve Clark, 2015).

Ortaokullarda İngilizce öğretiminde ölçme değerlendirme sürecinin incelendiği bu çalışma, Creswell ve Clark (2015) tarafından tanımlanan keşfedici sıralı desen akışına uygun olarak yürütülmüştür. Öncelikle nitel araştırma yöntemlerinden durum çalışması yöntemi kullanılarak araştırma gerçekleştirilmiştir. Özellikle dil öğrenimi, anadil ve ikinci dil edinimi ve öğretmen araştırmalarında sık sık kullanılan durum çalışması (Paker, 2015); bir kişi, olay veya olguyu niçin ve nasıl sorularını temele alarak derinlemesine inceleyen bir yöntemdir (Yıldırım ve Şimsek, 2013). Daha sonra yarı yapılandırılmış görüşme tekniğiyle veriler toplanmış ve kullanılacak bilgiler analiz edilmiştir. Ardından nitel araştırma sonuçlarına dayalı olarak anket hazırlanmış ve nicel araştırma desenlerinden genel tarama modeline uygun olarak araştırma yapılmıştır. Genel tarama modelleri, çok sayıda elemandan oluşan bir evrende, evren hakkında genel bir yargıya varmak amacıyla evrenin tümü ya da ondan alınacak bir grup, örnek ya da örneklem üzerinde yapılan tarama düzenlemeleridir (Karasar, 2013). Bu tür araştırmalar olayların, objelerin, varlıkların, kurumların ve çeşitli alanların ”ne” olduğunu açıklamaya çalışır (Kaptan, 1998). Nicel araştırma verileri elde edildikten sonra analizler yapılmış ve son aşamada araştırma sonuçları arasında ilişki kurularak, nitel verilerin nicel sonuçları açıklamada ne ölçüde yardımcı olduğu tartışılmıştır. Aşağıda Şekil 2’de araştırmanın uygulanmasına ilişkin temel akış şeması şu şekilde özetlenmiştir:

1. ADIM

Nitel Aşamanın Planlanması ve Uygulanması:

- Araştırma yöntemi olarak durum çalışması tercih edildi.
- Araştırma konusuna uygun olarak nitel araştırma soruları ifade edildi
- Nitel araştırma desenlerinden yarı yapılandırılmış görüşme tekniği seçildi
- Gerekli izinler alındı
- Nitel örneklem tanımlandı
- Nitel veriler toplandı
- İçerik analizi yapılarak temalar ve alt temalar belirlendi.
- İkinci aşama için gerekli bilgiler analiz edildi



Şekil 2: Creswell ve Clark (2015) tarafından oluşturulan keşfedici sıralı akış şemasına göre uyarlanmıştır.

3.3. Evren ve Örneklem

Ortaokullarda İngilizce öğretiminde ölçme değerlendirme sürecinin incelendiği bu araştırmanın evrenini Diyarbakır ili merkez ilçelerindeki ortaokullarda görev yapan toplam 543 İngilizce öğretmeni oluşturmaktadır. Araştırmanın nitel verileri için amaçlı örnekleme alma yöntemiyle örneklem belirlenmiş ve 45 öğretmenle görüşme yapılarak veriler toplanmıştır. Zaman ve maliyet açısından tasarruf sağlamak için nitel araştırmalarda küçük örneklem grupları tercih edilir ve nicel araştırmalardan farklı olarak süreç içerisinde örneklem büyüklüğüne karar verilir (Mason, 2002). Dolayısıyla nitel araştırmalarda örneklem büyüklüğünü hesaplamak için kesin bir kural yoktur (Gurbetoğlu, 2018). Nitel çalışmalar araştırma konusuyla ilgili amaçları ne ölçüde gerçekleştirebildikleri ve örneklemdeki kişi, olay ya da nesnelerin uygunluğuyla değerlendirilir (Luborsky & Rubinstein, 1995). Yağar ve Dökme'ye (2018) göre bu amaçlar sistematik bir süreçle

seçilen küçük örneklem gruplarının sunduğu yoğun araştırma fırsatıyla en iyi şekilde gerçekleştirilmektedir. Örneklem grubundan yeterli veri elde edildikten sonra yani veri doygunluğuna ulaşıldığında ise veri toplama süreci sona erdirilir (Mason, 2002).

Nitel çalışmalarda sınırlı sayıdaki kaynaklardan etkin ve derinlemesine bilgi elde edebilmek için amaçlı örnekleme yöntemlerinin kullanılması önerilmektedir (Marshall, 1996). Amaçsal örnekleme, çalışmanın amacına bağlı olarak bilgi açısından zengin durumların seçilerek derinlemesine araştırma yapılmasına olanak tanır (Büyüköztürk, Çakmak, Akgün, Karadeniz ve Demirel, 2013). Bu çalışmada katılımcıların seçiminde amaçsal örnekleme yöntemlerinden maksimum çeşitlilik örneklemesi tercih edilmiştir. Maksimum çeşitlilik yönteminde araştırılan olay ya da olguyla alakalı farklılıkları içeren ana temaları keşfetmek ve bunları yansıtmak amaçlanmaktadır (Neuman, 2014). Farklı durumların ortaya çıkarılma olasılığı ve durumların çoklu bakış açısıyla bütünüyle betimlenmesi bu örnekleme yönteminde daha fazladır (Creswell, 2015). Bu amaçla araştırmada farklı sosyo-ekonomik özelliklere sahip bölgelerdeki okullardan örneklem seçilerek durumlarda çeşitlilik sağlanacağı düşünülmektedir. Çalışma grubundaki katılımcıların görev aldığı okulların 8'ini kamuya bağlı ortaokullar, 6'sını İmam Hatip Ortaokulu ve 4'ünü özel ortaokullar oluşturmaktadır. Öğretmenlerin seçiminde ise okul yöneticilerinin görüş ve önerileri dikkate alınarak gönüllü olan İngilizce öğretmenleriyle görüşmeler yapılmış ve veri doygunluğuna ulaşıldığında süreç sonlandırılmıştır. Araştırmaya dahil edilen katılımcılara ait özellikler Tablo 1'de verilmiştir.

Tablo 1. Araştırmanın nitel boyutuna dahil edilen öğretmenlere ait özellikler

Kişisel Özellikler	**Kategori**	**Öğretmen Sayısı (N)**	**Yüzde (%)**
Cinsiyet	Kadın	32	71,1
	Erkek	13	28,9
Mesleki Kıdem	1-5 yıl	10	22,2
	6-10 yıl	17	37,8
	11-15 yıl	10	22,2
	16-20 yıl	3	6,7
	21+ yıl	5	11,1
Öğretim Kurumunun Statüsü	Ortaokul	25	55,6
	İmam Hatip Ortaokulu	10	22,2
	Özel Ortaokul	10	22,2
Toplam		**45**	**100**

Araştırmada örneklem olarak seçilen öğretmenlerin % 71,1'i kadın, % 28,9'u erkektir. Öğretmenlerin mesleki kıdem olarak örneklem içindeki oranlarına bakıldığında; % 22,2'si 1 ile 5 yıl, %37,8'i 6 ile 10 yıl, % 22,2'si 11 ile 15 yıl, % 6,7'si 16 ile 20 yıl ve %11,1'i 20 ve daha fazla yıl aralığında görev yapmışlardır. Görev yapılan öğretim kurumunun statüsüne göre incelendiğinde ise; öğretmenlerin % 55,6'sı kamuya ait ortaokullarda, %22,2'si kamuya ait İmam Hatip Ortaokullarında ve %22,5'i ise özel ortaokullarda görev almaktadır. Araştırmaya dahil edilen öğretmenlerin dengeli bir dağılım gösterdiği söylenebilir.

Araştırmanın nicel verileri ise 2020-2021 eğitim öğretim yılında Diyarbakır ili merkez ilçelerinde görev yapan 543 İngilizce öğretmeninden tabakalı örnekleme yöntemi ile seçilen 306 öğretmenden elde edilmiştir. Tabakalı örnekleme yönteminin temelinde evreni çeşitli alt gruplara bölmek ve ardından örnekleri bu gruplardan seçmek yatmaktadır (Onwuegbuzie & Collins, 2007). Alt tabakalar belirlenirken ise kullanılan ölçütlerin araştırmanın amaçları ve değişkenleri ile ilgili olması oldukça önemlidir (Vogt, Gardner & Haeffele, 2012). Ardından her tabakadan basit seçkisiz veya tesadüfi örnekleme yöntemleri ile öğeler seçilip örneklem oluşturulabilir (Neuman & Robson, 2014). Bu çalışmada farklı okul türlerinin araştırmanın alt problemleri üzerinde etkili olduğuna karar verilmiş ve uzman görüşleri doğrultusunda bu okullar kamuya bağlı ortaokullar, İmam Hatip ortaokulları ve özel ortaokullar olmak üzere üç tabakaya ayrılmıştır. Daha sonra basit seçkisiz örnekleme yöntemi kullanılarak her katmandan seçilen öğeler bir araya getirilmiştir. Bu yöntemde evrendeki tüm birimlerin örnekleme seçilme şansları eşit ve birbirinden bağımsızdır (Baştürk ve Taştepe, 2013). Bryman'a (2007) göre evren coğrafi olarak belirli bir alanda yoğunlaştığı ve iyi bir örneklem çerçevesi mevcut olduğunda basit seçkisiz örnekleme yöntemi oldukça kullanışlıdır.

Örneklem büyüklüğünün belirlenmesinde araştırmanın deseni, yöntemi, evrenin niteliği, kabul edilen güven düzeyi, değişken sayısı, araştırma imkanları, veri analiz yöntemleri ve hata payı gibi faktörler göz önünde bulundurulur (Baştürk ve Taştepe, 2013). En önemli nokta ise evreni temsil etme gücünün yüksek olmasıdır (Balcı, 2015). Araştırmanın örneklem büyüklüğü belirlenirken şu formülden yararlanılmıştır (Büyüköztürk vd., 2013):

$$n = \frac{\mathrm{N t^2 pq}}{\mathrm{d^2 (N-1) + t^2 pq}}$$

N: Evren büyüklüğü

n: Örneklem büyüklüğü

t: Belirli bir anlamlılık düzeyi, t tablosunda bulunan teorik değer

p: İncelenen olayın gerçekleşme olasılığı

q: İncelenen olayın gerçekleşmeme olasılığı

d: Duyarlılık (kabul edilen +/- örnekleme hatası)

Örneklem büyüklüğü belirlenirken evren için p tahmini yoksa p=q=0,5 alınabilir ve %95 güven aralığında d=0,05 örneklem hatasına karşılık gelen t-tablo değeri 1,96'ya eşittir (Büyüköztürk vd., 2013). Bu formüle göre araştırmanın örneklem büyüklüğü n=225 çıkmaktadır. Araştırma kapsamında 306 İngilizce öğretmenine ulaşılmıştır. Araştırmaya dahil edilen katılımcılara ait özellikler Tablo 2'de verilmiştir.

Tablo 2. Araştırmanın nicel boyutuna dahil edilen öğretmenlere ait özellikler

Kişisel Özellikler	**Kategori**	**Öğretmen Sayısı (N)**	**Yüzde (%)**
Cinsiyet	Kadın	217	71
	Erkek	89	29
Mesleki Kıdem	1-5 Yıl	55	18
	6-10 Yıl	112	37
	11-15 Yıl	77	25
	16-20 Yıl	45	15
	21+ Yıl	17	5
Öğrenim Durumu	Lisans	283	92
	Lisansüstü	23	8
Okul Türü	Genel Ortaokul	246	80
	İmam Hatip Ortaokulu	33	11
	Özel Ortaokul	27	9
Ölçme Değerlendirme İle İlgili Kursa/Seminere Katılma Durumu	Evet	68	22
	Hayır	238	78
Toplam		**306**	**100**

Tablo 2'ye bakıldığında, araştırmaya katılan öğretmenlerin %71'i kadın, %29'u ise erkektir. Katılımcıların mesleki kıdem dağılımlarına bakıldığında % 37'lik bir oranla öğretmenlerin çoğunun 6-10 yıl arası bir mesleki kıdeme sahip oldukları görülmektedir. Bu oranı %25 ile 11-15 yıl ve %18 ile 1-5 yıl arası mesleki tecrübesi olanlar takip etmektedir. Katılımcıların büyük bir kısmının (% 92) lisans mezunu oldukları ve kamuya bağlı genel

ortaokullarda görev aldıkları (% 80) belirlenmiştir. Ayrıca katılımcıların %78'nin daha önce ölçme değerlendirmeyle ilgili herhangi bir kursa ya da seminere katılmadığı kaydedilmiştir.

3.4. Veri Toplama Araçları

Araştırmada veri toplama aracı olarak araştırmacı tarafından geliştirilen yarı yapılandırılmış görüşme formu ve anket kullanılmıştır.

3.4.1. İngilizce Öğretiminde Ölçme ve Değerlendirme Sürecinin İncelenmesine İlişkin Görüşme Formu

Araştırma sorularına yanıt aramak için yarı yapılandırılmış görüşme tekniğiyle veriler toplanmıştır. Yarı yapılandırılmış görüşme, yapılacak görüşmenin araştırmacı tarafından önceden planlanıp sorulacak soruların önceden hazırlandığı bir veri toplama tekniğidir (Ekiz, 2009). Yarı yapılandırılmış görüşme tekniğinde görüşmeciler konu listesini hazırlar ve bunlara yanıt aramak isterler. Ancak soruların sıralanmasında, kesin ifadelerde ve farklı konulara verdikleri önem ve zaman miktarında önemli bir özgürlüğe sahiptirler (Robson, 2015).

İngilizce öğretiminde ölçme ve değerlendirme süreciyle ilgili veri toplamadan önce konuyla ilgili kaynak taraması yapılmış ve görüşme soruları hazırlanmıştır. Daha sonra Adıgüzel (2016) tarafından geliştirilen *"Uzman Görüşme Formu"* göz önünde bulundurularak araştırmanın amaç ve sorularına uygun olacak şekilde bir uzman değerlendirme formu hazırlanmış ve Yabancı Dil ile Eğitim Bilimleri alanında uzman 7 öğretim üyesinin görüşlerine başvurulmuştur. Uzmanların soruları değerlendirmesine yardımcı olmak amacıyla öncelikle araştırmanın amacı ve kuramsal dayanaklarını özetleyen bilgiler formun birinci kısmında sunulmuştur. İkinci kısımda ise tüm sorular tablo halinde sunulmuş ve uzmanlardan görüşme formundaki her bir sorunun karşısında bulunan *"Uygundur", "Uygun değildir", "Uygun ancak düzeltilmeli"* ve *"Öneri/Açıklama"* başlıklarının altına görüşlerini yazıyla aktarmaları beklenmiştir. Uzman değerlendirmeleri alındıktan sonra, uzmanlarla tekrar görüşülmüş ve geri bildirim yapılarak görüşlerin veri toplama aracına nasıl yansıtıldığı belirtilmiştir.

Veri toplama aracı öğretim üyelerinin görüş ve önerileri doğrultusunda yeniden düzenlenmiş ve ön uygulamaya hazır hale getirilmiştir. Ön uygulama için ortaokullarda görev yapan 5 İngilizce öğretmeniyle görüşülmüştür. Veri toplama sürecinde katılımcıların isteği doğrultusunda ses kaydı alınmamıştır. Ön görüşme yapılan katılımcılara görüşme formundaki sorular hakkında düşünceleri sorulmuş ve katılımcılardan kendi el yazılarıyla

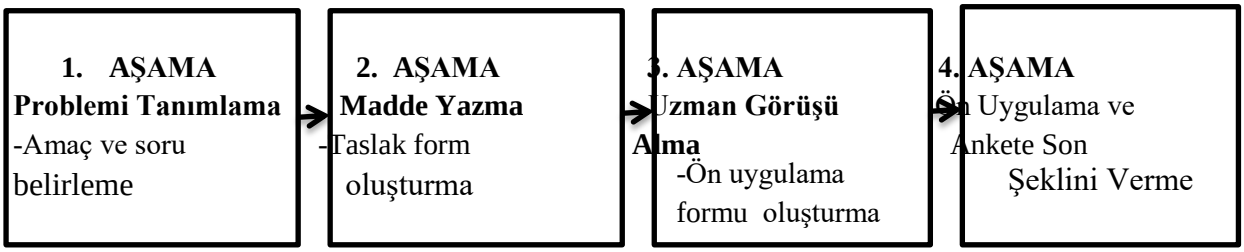
düşüncelerini aktarmaları istenmiştir. Yapılan bu görüşmeler sonucunda soruların yeterince açık ve anlaşılır olduğuna karar verilmiş ve veri toplama sürecine başlanmıştır. Ön görüşme yapılan katılımcılar araştırmada esas çalışma grubuna dahil edilmemiştir. Böylece 45 öğretmenden alınan görüşler incelemeye alınmıştır.

3.4.2. Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketi

Araştırmanın birinci aşamasında elde edilen nitel araştırma verileri ışığında yabancı dil beceri alanlarının değerlendirilmesi ile ilgili ortaya çıkan görüşlerin yaygınlığıyla ilgili bir çalışma yapılmak istenmiştir. Ayrıca nitel verileri daha iyi açıklamak ve daha geniş çaptaki bir örneklemin özelliklerini saptamak amaçlanmıştır. Evrenin temsiliyetini arttırmak için ise geniş bir veri kaynağına ulaşmak hedeflenmiş ve daha hızlı ve kolay ulaşılabilir veri toplama tekniklerinden olan anket tercih edilmiştir.

Doğrudan gözlenemeyen davranışlar, fikirler, tutumlar ya da inançlar hakkında bilgi toplamamızı sağlayan anketler (Lee, 2006); ihtiyaç analizlerinde, eğitim programlarını değerlendirmede ve diğer pek çok alanda tanılayıcı bir araç olarak sık sık kullanılmaktadırlar (Hayes, 1992). Ayrıca anketlerin diğer veri toplama tekniklerine göre daha geniş örneklemlere uygulanabilmesi, maliyetinin daha düşük olması ve kısa sürede zengin bir veri kaynağı sunması gibi avantajları vardır (Acharya, 2010). Codo' ya (2005) göre eğer bir araştırmanın amacı diller hakkında geniş bir örneklemden ölçülebilir bilgi toplamaksa, anketler bu amaç için kullanılabilecek en uygun araçlardandır. Anketler aracılığıyla elde edilen bilgiler, dil kullanımıyla ilgili genel bir çerçeve çizmede, araştırılan sosyal içeriklerin kavramsallaştırılmasında ya da hedef kişi ya da konuları tespit etmede kullanılabilir (Codo, 2005).

Anket kullanımına karar verilmeden önce cevaplanması gereken en önemli konu, bu aracın araştırma sorularına yanıt aramak için yeterli olup olmadığıdır. Bir diğer önemli konu ise anketleri önceden belirlenen amaçları gerçekleştirmek üzere iyi bir şekilde tasarlamaktır (Lee, 2006). Bu doğrultuda anket geliştirme süreci planlı bir şekilde önceden tasarlanmış ve aşağıda Büyüköztürk (2005) tarafından tanımlanan *"Problemi tanımlama", "Madde (soru) yazma", "Uzman görüşü alma"* ve *"Ön uygulama yapma"* aşamalarına uygun olarak çalışma yürütülmüştür (Şekil 3).



Şekil 3. Anket Geliştirme Süreci (Büyüköztürk, 2005).

Şekil 3'te belirtildiği gibi 1. aşamada öncelikle problemi daha iyi keşfetmek için araştırmanın amacına uygun olarak incelenecek konular ve anketin sınırları belirlenmiştir. 2. aşamada nitel verilerin analizi ile elde edilen tema ve alt temalar ve ilgili alanyazı taranarak (Richards, 1990; Brown, 2002; Richards and Rodgers, 2002; Brown, 2004; Brown ve Abeywickrama, 2009) anket maddeleri yazılmış ve taslak form oluşturulmuştur. Anket taslak formunda 4 soruluk kişisel bilgi formu ve 48 maddelik bir madde havuzu oluşturulmuştur. 3. aşamada madde havuzunda yer alan soruların ne derece yeterli olduğuyla ilgili uzman görüşüne başvurulmuştur. Bu süreçte Büyüköztürk, vd. (2013) tarafından geliştirilen *"Uzman Görüşme Formu"*na madde havuzundaki sorular aktarılmış ve Yabancı Dil ve Eğitim Bilimleri alanında uzman 9 öğretim üyesinin incelemesine sunulmuştur. Uzmanlardan soruların uygunluğuna dair görüşlerini *"Uygun"* ve "*Uygun Değil"* şeklinde belirtmeleri beklenmiştir. Veri toplama aracı öğretim üyelerinin görüşleri doğrultusunda incelenmiş ve her bir madde için uyuşma düzeyi hesaplanmıştır. Uzmanların her bir sorunun geçerliğiyle ilgili uyuşma düzeyleri Tablo 3'te gösterilmiştir.

Tablo 3. Veri toplama aracındaki maddelerin uyuşma düzeyleri

Maddeler	**Uygun (%)**	**Uygun Değil (%)**	**Maddeler**	**Uygun (%)**	**Uygun Değil (%)**
m1	100	0	m25	77,8	22,2
m2	66,7	33,3	m26	66,7	33,3
m3	100	0	m27	100	0
m4	88,9	11,1	m28	100	0
m5	77,8	22,2	m29	88,9	11,1
m6	77,8	22,2	m30	88,9	11,1
m7	88,9	11,1	m31	55,6	44,4
m8	100	0	m32	100	0
m9	66,7	33,3	m33	100	0
m10	88,9	11,1	m34	88,9	11,1
m11	100	0	m35	88,9	11,1
m12	77,8	22,2	m36	100	0
m13	88,9	11,1	m37	100	0
m14	100	0	m38	77,8	22,2

m15	100	0	m39	88,9	11,1
m16	88,9	11,1	m40	100	0
m17	88,9	11,1	m41	77,8	22,2
m18	77,8	22,2	m42	88,9	11,1
m19	88,9	11,1	m43	100	0
m20	100	0	m44	66,7	33,3
m21	100	0	m45	88,9	11,1
m22	100	0	m46	88,9	11,1
m23	66,7	33,3	m47	100	0
m24	100	0	m48	100	0

Tablo 3'te anket maddelerinin uygunluğunun belirlenmesinde uzmanların görüşlerinin uyuşma düzeylerinin yüzdelik puanları gösterilmiştir. Büyüköztürk'e (2005) göre, veri toplama aracındaki maddelerin geçerli sayılabilmesi için uzmanların uyuşma düzeylerinin %90-100 olması beklenilmektedir. Uzmanların %70-80 düzeyinde uyuşma gösterdikleri maddelerin ise gerekli düzeltmeler yapılarak ölçekte tutulabileceği belirtilmiştir. Uzmanların görüşleri incelendiğinde %70'in altında uyuşma gösteren 6 madde (m2, m9, m23, m26, m31, m44) araştırmanın amacı ve kapsamı göz önünde bulundurularak anketten çıkarılmıştır. Uygun uyuşma düzeyine sahip olup geçerli sayılan 21 madde üzerinde ise öneriler doğrultusunda düzeltmeler yapılmış ve dil, anlatım ve yazım kuralları göz önünde bulundurularak uygun şekilde tekrar yapılandırılmıştır. Ayrıca 1 soru da kişisel bilgi formuna eklenmiştir. Böylece ön uygulamaya hazır bir anket oluşturulmuştur. Ön uygulama için ortaokullarda görev alan 15 İngilizce öğretmenine veri toplama aracı olarak anket uygulanmıştır. Anket maddelerinin anlamsal açıdan belirsizlik ifade etmediği katılımcılar tarafından belirtildikten sonra çalışma uygulamaya hazır hale getirilmiştir.

3.4.2.1. Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketinin Geçerlik Çalışması

Anketler bilimsel ölçme araçları oldukları için yeterli düzeyde geçerli ve güvenilir sonuçlar ortaya koymaları oldukça önemlidir. Standart ölçme araçlarının kullanımı oldukça sıkı geçerlik ve güvenirlik hesaplamalarını gerektirirken, belirli araştırma amaçları hakkında bilgi toplamak için geliştirilen anketlerin ise geçerlik ve güvenirliğin her

boyutuyla ilgili indeksleri sağlaması her zaman makul olmayabilir (Dörnyei & Taguchi, 2010). Şimşek'e (2014) göre anketlerin geçerliğini belirlemek için sadece yargısal yaklaşımlar kullanırken, standart araçlar içinse hem yargısal hem de istatiksel yaklaşımlar kullanılabilir. Yargısal yaklaşımla geçerliği sağlamak için alanyazı taramaları ve uzman görüşlerinin değerlendirilmesi yeterli kabul edilebilir (Taherdoost, 2016).

Bu araştırmada anketin geçerliğini sağlamak için görünüş ve kapsam geçerlikleri incelenmiştir. Araştırma kapsamı ve soruları göz önünde bulundurularak öncelikle nitel veriler ışığında ortaya çıkan temalar ayrıntılı şekilde incelenmiş, daha sonra ilgili alanyazı taraması yapılmış ve ardından anket maddeleri hazırlanmıştır. Anket maddelerinin geçerliğini sağlamak için ise uzman görüşlerine başvurulmuş ve Büyüköztürk, vd. (2013) tarafından geliştirilen *"Uzman Görüşme Formu"*na verilen yanıtlar neticesinde maddelerin uyuşma düzeyleri hesaplanmıştır. Daha sonra uyuşma düzeyi düşük olan 6 madde anketten çıkarılırken, 21 madde ise uzmanların önerileri doğrultusunda yeniden yapılandırılmıştır.

3.5. Verilerin Toplanması

Keşfedici karma yöntem desenine uygun şekilde yürütülen bu araştırmada ilk önce nitel veriler toplanmış, analiz edilmiş ve daha sonra bu bilgiler veri toplama sürecinin nicel aşamasını geliştirmek için kullanılmıştır.

3.5.1. Nitel Verilerin Toplanması

Nitel araştırmaya ait veriler Dicle Üniversitesi Eğitim Bilimleri Enstitüsü aracılığıyla Diyarbakır İl Milli Eğitim Müdürlüğünden gerekli izinler alındıktan sonra, 2019-2020 eğitim öğretim yılı bahar döneminde Diyarbakır ili merkez ilçelerine (Bağlar, Kayapınar, Sur ve Yenişehir) bağlı ortaokullardan toplanmıştır. Çalışma grubundaki katılımcıların görev aldıkları okulların 8'ini kamuya bağlı ortaokullar, 6'sını İmam Hatip Ortaokulları ve 4'ünü özel ortaokul olmak üzere toplam 18 ortaokul oluşturmaktadır. Araştırmada veri toplama aracı olarak araştırmacı tarafından geliştirilen ve uzman görüşlerine göre yapılandırılan görüşme formu kullanılmıştır.

Görüşmelerin daha etkili gerçekleştirilebilmesi için okullar ile önceden görüşülerek gerekli randevular alınmış ve görüşme yeri ve saatine katılımcılara uygun olacak şekilde karar verilmiştir. Ayrıca katılımcıların iletişim bilgileri ve görüşme yeri, saati, görüşme talebinin tarihi ile ilgili detayları içeren bir form hazırlanmıştır. Böylece planlı bir şekilde yüz yüze görüşmelerin gerçekleştirilmesi sağlanmıştır.

Görüşme formunda katılımcılara gönüllülüğü teyit eden açıklamalara, görüşmenin hangi amaçla yapıldığına, nasıl kayıt altına alınacağına ve kimlerle paylaşılacağına dair gerekli bilgilendirmelere yer verilmiştir. Yüz yüze görüşmeler gerçekleştikten sonra ise veri ve analizlerin katılımcılara sunularak, araştırmacının elde ettiği verilerin doğru anlaşılıp yorumlandığının teyit edilmesi gerekmektedir. Eğer verilere dair herhangi bir yanlış anlama veya yorumlama mevcut ise, bunun katılımcıların söyledikleri doğrultusunda tekrar düzeltilmesi oldukça önemlidir (Ekiz, 2009). Bu amaçla, yarı yapılandırılmış görüşme tekniğiyle elde edilen veriler tekrar katılımcıların bilgisine sunulmuş ve belirtilen görüşlerin katılımcılara ait olduğuna dair imza alınmıştır.

3.5.2. Nicel Verilerin Toplanması

Nicel araştırmaya ait veriler Dicle Üniversitesi Eğitim Bilimleri Enstitüsü aracılığıyla Diyarbakır İl Milli Eğitim Müdürlüğünden gerekli izinler alındıktan sonra, 2020-2021 eğitim öğretim yılı bahar döneminde Diyarbakır ili merkez ilçelerine (Bağlar, Kayapınar, Sur ve Yenişehir) bağlı ortaokullardan toplanmıştır. Küresel çapta yaşanan pandemi salgını araştırmacıyı verilerin elde edilmesi surecinde kısıtladığından dolayı, verilerin bir kısmı yüz yüze, bir kısmı da çevrimiçi (online) olarak toplanmıştır. Veri toplama sürecinde araştırmacı tarafından geliştirilen Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketi Google Forms'a aktarılmış ve çevrimiçi olarak katılımcılara sunulmuştur. Araştırma örneklemini oluşturan toplam 306 katılımcıdan 50'si formu çevrimiçi olarak yanıtlarken, 256'sı ile yüz yüze görüşülerek veriler toplanmıştır.

3.6. Verilerin Analizi

Karma yöntem araştırmalarında veri analizi, nitel ve nicel veri analiz stratejilerinin araştırma çalışmalarında bir araya getirildiği, bağlandığı ve birleştirildiği süreçleri içermektedir (Peker ve Güngör, 2015). Verilerin analizleri tamamlandıktan sonra ise karma yöntem araştırmasının doğru bir şekilde yorumlanması için nicel ve nitel bulguları birlikte incelemek ve ortaya çıkan sonuçların karma yöntem araştırmasındaki sorulara nasıl cevap verdiğine ilişkin değerlendirmeler yapmak gerekir (Güzel ve Kula, 2015). Alanyazında karma yöntem veri analizi stratejilerinin nasıl yapılabileceğine ilişkin farklı sınıflandırmalar mevcuttur (Caracelli & Grene, 1993; Onwuebuzie & Teddlie, 1998; Creswell & Clark, 2015). Bukova-Güzel ve Kula'ye göre karma yöntem veri analizine ilişkin tartışmalar bu sürece dair genel bir yaklaşım sunarken, Carecelli ve Greene'nin 1993'de ortaya koyduğu analitik stratejiler, veri analiz sürecine dair belirli tasarımlar

sunmaktadır. Greene (2007) karma yöntem analiz stratejilerini şu şekilde özetlemiştir (Akt. Peker ve Güngör, 2015) :

1. Paralel Karma Veri Analizi
2. Dönüştürücü Karma Veri Analizi
   a. Nitel verileri nicelleştirme
   b. Profil bilgisi içeren sayısal veriyi nitelleştirme
   c. Doğası gereği karma veri analizi teknikleri
3. Sıralı Karma Veri Analizi
   a. Tipoloji gelişimini içeren sıralı nitel → nicel analizler
   b. Tipoloji gelişimini içeren sıralı nicel → nitel analizler
   c. Tekrarlı sıralı karma analiz
4. Çok aşamalı Karma Veri Analizi
5. Tamamen Bütünleşik Karma Veri Analizi
6. Analitik Tekniklerin Bir Gelenekten Diğer Bir Geleneğe Uygulanması

Bu araştırmanın verileri keşfedici karma yöntem desenine uygun olarak sıralı karma veri analizi teknikleri kullanılarak analiz edilmiştir. Sıralı karma veri analizinde araştırmanın nitel veya nicel aşamalarından biri, bir önceki aşamaya bağlı olarak ortaya çıkar ve bu aşamalılık kronolojik bir sırada gerçekleşir (Tashakkori & Teddlie, 2009). Caracelli ve Greene'ne (1993) göre sıralı veri analizinde tipoloji oluşturulması için bir veri türü analiz edilir ve daha sonra bu tipoloji diğer veri türlerinin analizinde bir çerçeve olarak kullanılır (Akt. Güzel ve Kula, 2015). Dolayısıyla bu çalışmada öncelikle nitel ön veriler elde edilmiş ve bu aşamayı takiben nicel veriler toplanmıştır. Daha sonra ilk nitel keşfedici bulguları keşfetmek için nicel analizlerin nasıl katkıda bulunulduğuna değinilmiş ve iki veri tabanı birbiriyle ilişkilendirilerek değerlendirilmiştir.

3.6.1. Nitel Veri Analizi

Araştırmanın ilk aşamasında elde edilen nitel verilerin çözümlenmesinde nitel veri analizi yöntemlerinden içerik analizi kullanılmıştır. Nitel içerik analizi, metin verilerinin içeriğindeki temaların ve örüntülerin tanımlanmasını ve kodlamanın sistematik bir sınıflandırmayla yorumlanmasını ifade eden bir analiz türüdür (Hsieh & Shannon, 2005). İçerik analizinin sadece kelimeleri saymanın ötesinde olduğunu belirten Yıldırım ve Şimşek'e (2013) göre, içerik analizi yoluyla verilerin içindeki açık ve gizli gerçekler ortaya çıkarılmaya çalışılır. Nitekim bu araştırmayla da çalışmaya katılan katılımcıların

araştırmanın konusuyla doğrudan ya da dolaylı olarak ilgili gördükleri kelimeleri, kalıpları ya da temaları belirlemek ve bu kavramları ne sıklıkla ve hangi bağlamda kullandıklarını sistematik bir şekilde sınıflandırmak amaçlanmıştır. Böylece hem verilere ait frekanslar sunularak görüşlerin sıkılığı vurgulanmış hem de içerikteki anlama odaklanılarak yorumlamacı ve bilimsel bir paradigmayla analiz yapılmıştır.

Nitel veri analizi için veri toplama sürecinin başından itibaren veri ile ilgili olabilecek olası örüntüler ve temalar araştırılmıştır. Böylece verilerin sürekli düzenlenmesi ve metindeki açık veya saklı kalan örüntülerin keşfedilmesi amaçlanmıştır. Nitel verilerin analizi için araştırmaya katılan her bir katılımcının verdiği yanıtlar öncelikle bilgisayar ortamına aktarılmıştır. Bilgisayar ortamında her soru için ayrı dosya açılmış ve tüm katılımcılara ilişkin görüşler Ö1, Ö2, Ö3,...…..Ö45 olarak kodlanıp düzenlenmiştir. Ayrıca farklı okul türleri olduğu için İmam Hatip ortaokullarındaki öğretmenlerden elde edilen görüşler İHO1, İHO2…..İHO10 ve özel okullardan elde edilen veriler ise ÖO1, ÖO2,….ÖO10 olarak tanımlanmıştır.

Tüm katılımcılara ait görüşler aktarıldıktan sonra veri analiz süreci Yıldırım ve Şimsek'in (2013) dört aşamalı içerik analizine uygun olarak yürütülmüştür. Birinci aşamada, veriler anlamlı bölümlere ayrılmış ve her bölümü kavramsal olarak ifade eden kodlar elde edilmiştir. İkinci aşamada, kodlar arasındaki ortak özellikler incelenmiş ve kodları belirli kategoriler altında toplayan temalar tanımlanmıştır. Üçüncü aşamada veriler tanımlanan kodlara ve temalara göre sunulmuştur. Son aşamada ise toplanan bulgular arasındaki ilişkiler açıklanmış ve sonuçlara dair birtakım çıkarımlar yapılarak analiz süreci tamamlanmıştır.

3.6.1.1. Nitel Veri Analizine Dair Geçerlik ve Güvenirlik Çalışmaları

3.6.1.1.1. Geçerlik

Nitel araştırmalarda geçerliği sağlama yolları nicel araştırmalardan farklı olsa da asıl amaç verilerin doğruluğu yansıtma düzeyini incelemektir (Silverman, 2000). En genel haliyle nitel geçerliği kontrol etme, toplanan verilerin, bulguların ve yorumların doğruluğunu değerlendirmektir (Güzel ve Kula, 2015). Lincoln ve Guba'ya (1985) göre, nitel araştırmalarda araştırmacı ve katılımcıların doğru, inandırıcı ve güvenilir görüşler bildirip bildirmediklerini belirlemek amaçlandığından, güvenirlikten daha çok geçerliğe odaklanılır. Alanyazında nitel araştırmalarda geçerliğin nasıl sağlanabileceğine ilişkin çeşitli yaklaşımlardan bahsedilmiştir (Lincoln & Guba, 1985; Kirk & Miller; 1986;

Mishler, 1990; Miles & Huberman, 1994; Creswell & Miller, 2000; Maxwell, 2005). Bu araştırmanın geçerlik analizleri Maxwell'in (1992; 2005) çalışmalarında ortaya koyduğu nitel araştırmalarda kullanılan geçerlik stratejilerine dayalı olarak yürütülmüştür. Bu bağlamda araştırmanın verileri betimsel geçerlik, yorumlama geçerliği, iç geçerlik ve dış geçerlik boyutlarında incelenmiştir.

Betimsel geçerlik ortaya konulan verilerin olduğu gibi doğrudan yansıtılması ile ilgilidir (Wolcott, 1990). Betimsel geçerliği sağlamanın yollarından biri yalnızca bir araştırmacının bakış açısını yansıtma olasılığını düşüren araştırmacı çeşitlemesidir (Maxwell, 1992). Bu araştırmanın verilerinin toplanması ve yorumlanması süreçlerinde uzman görüşlerine başvurulmuş ve araştırmada vurgulanan betimsel detaylarla ilgili birden çok araştırmacıyla görüşülerek analiz süreçleri eleştirel bir bakış açısıyla sorgulanmıştır.

Yorumlama geçerliği betimsel geçerlikle oldukça bağlantılıdır ve katılımcıların araştırma konusuna dair perspektiflerinin ne derece doğru yansıtıldığına odaklanmaktadır (Sandelowski, 2000). Yorumlama geçerliği sağlamanın yolları katılımcı geribildirimi almak ve çok sayıda alıntı yapmaktır (Maxwell, 2005). Bu araştırmada hem katılımcı hem de araştırmacı nezdinde oluşabilecek olası ön yargıları veya yanlış anlamaları düzeltmek ve katılımcıların düşüncelerini en doğru şekilde ifade ettiklerinden emin olmak için yüz yüze görüşmeler gerçekleştikten sonra elde edilen veriler tekrar katılımcıların bilgisine sunulmuş ve belirtilen görüşlerin katılımcılarına ait olduğunu teyit eden imzalar alınmıştır. Ayrıca araştırma bulgularında çok sayıda doğrudan alıntılara yer verilmiştir.

İç geçerlik, gözlenen ilişkinin nedensel olduğuna dair araştırmacının yaptığı çıkarımların ne derece doğru olup olmadığıyla ilgilidir (Christensen, Johnson & Turner, 2015). Lincoln ve Guba (1985), nitel araştırmanın doğasına daha uygun olduğu gerekçesiyle iç geçerlik yerine "inandırıcılık", dış geçerlik yerine ise "aktarılabilirlik" kavramlarını kullanmayı tercih etmişlerdir. Bu araştırmada katılımcılarla olan görüşmeler geniş bir zamana yayılarak okul ortamında gerçekleştirilmiş ve yeterli sayıda katılımcıya ulaşılarak veri doygunluğuna ulaşıldığında araştırma sona erdirilmiştir. Araştırmada derinlemesine veri toplamak amaçlanmış ve elde edilen bulguların araştırmanın sorularına ne derece cevap verdiğiyle ilgili sürekli sorgulamalar yapılmıştır. Ayrıca araştırmanın deseninden görüşme formu hazırlanmasına, toplanan verilere, bunların analizine ve raporlaştırılmasına kadar olan tüm süreçlerde uzman görüşüne başvurulmuş ve yapılan geri bildirimlerle gerekli düzeltmeler yapılmıştır. Bunun yanı sıra araştırma karma yöntem desenine uygun olarak yürütüldüğü için, nitel araştırmadan sonra ikinci aşamada nicel bir

çalışma yapılmış ve elde edilen verilerin birbiriyle ne düzeyde örtüştüğü tartışılmıştır. Bu şekilde bir yöntem çeşitlemesi ile araştırmanın geçerliğini arttırmak hedeflenmiştir.

Dış geçerlik, araştırma sonuçlarının diğer insanlara, durumlara, zamana ya da ortamlara olan benzerliği üzerine yapılan genellemelerdir (Maxwell, 1992). Nicel araştırmalardan farklı olarak nitel araştırmalarda genelleme yerine "aktarılabilirlik" kavramının benimsenmesi, araştırma sonuçlarının benzer ortamlara transfer edilebileceğini vurgulamaktadır (Becker, 1991). Erlandson, Harris, Skipper ve Allen (1993) araştırma sonuçlarının aktarılabilirliğini arttırmak için ayrıntılı betimleme ve amaçlı örnekleme yöntemlerinin kullanılmasını önermiştir. Nitekim bu araştırmada, doğrudan alıntılar sıkça kullanılmış ve katılımcıların seçiminde farklı olguların değişkenlik gösteren özelliklerini ortaya koymak amaçlandığından amaçlı örnekleme yöntemlerinden maksimum çeşitlilik örneklemesi tercih edilmiştir. Ayrıca araştırmanın ikinci aşamasında yürütülen nicel çalışmayla nitel verileri daha iyi keşfetmek ve daha geniş çaptaki bir örneklemin özelliklerini saptamak amaçlanmıştır. Böylece nitel araştırma sonuçlarının dış geçerliğinin arttığını söylemek mümkündür.

3.6.1.1.2. Güvenirlik

Güvenirlik, farklı zamanlarda yapılan araştırmanın aynı sonucu vermesi olarak anlaşılsa da Kirk ve Miller (1986) bu tür bir yönlendirmenin nitel araştırmalar için yanlış olduğunu ifade etmiş ve bu araştırmalardaki görüşlerin nadiren sabit kaldığını savunmuşlardır. Nitekim insan doğasının sürekli değişen dinamizmi gereği sosyal araştırmaların tekrarlanması mümkün olmadığından (Baltacı, 2019), nitel araştırmalarda güvenirlik kavramı yerine "inandırıcılık" olması gerektiğine dikkat çekilmiştir (Lincoln ve Guba, 1985). Merriam'a (1998) göre, elde edilen bulguların sürekli kontrol edilmesi ve denenmesi araştırmanın inandırıcılığı açısından oldukça önemlidir. Poggenpoel ve Myburgh (2003) kodlar, kategoriler, temalar ve diğer veri işleme süreçleri kontrol edildikten sonra elde edilen sonuçların, analiz edilen veri setini ne düzeyde temsil ettiğinin belirlenebileceğine dikkat çekmişlerdir. Nitel araştırmada kodlayıcılar arası uyum olarak adlandırılan bu süreç, verilerin farklı araştırmacılar tarafından kodlanmasını ve daha sonra bu araştırmacıların kodlamalarını karşılaştırmayı içerir (Miles & Huberman, 1994). Miles ve Hubermann (1994) tarafından "kodlayıcılar arası görüş birliği" olarak kavramsallaştırılan bu kavram, ortaya koyduğu benzerlik oranıyla iç tutarlılığın ne düzeyde sağlandığına ilişkin araştırmacılara temel bir kavrayış sunmaktadır. Bu modele göre nitel

araştırmalar için güvenirlik oranı, üzerinde uzlaşı sağlanan görüş sayısının toplam uzlaşılan ve uzlaşılamayan görüş sayısına oranlanmasıyla saptanabilir.

Bu araştırmada katılımcıların görüşme formundaki sorulara verdikleri cevaplar, hem araştırmacı hem de alanında uzman bir öğretim üyesi tarafından ayrı ayrı incelenmiştir. Elde edilen veriler sonucunda ortaya çıkan tema ve alt temalar, Miles ve Huberman (1994) tarafından geliştirilen formüle uygun olarak "görüş birliği" ve "görüş ayrılığı" başlıkları altında toplanmıştır. Kodlayıcılar arasında görüş birliği sağlanan tema ve alt tema sayısı 214, görüş birliği sağlanamayan tema ve alt tema sayısı ise 18 olarak belirlenmiştir. Bu sonuçlar Güvenirlik = Görüş Birliği / (Görüş Birliği + Görüş Ayrılığı) formülüne göre hesaplandığında araştırmanın güvenirlik oranının % 92 olduğu görülmektedir. Kodlayıcılar arası görüş birliğinin % 80'e yakın olması önerildiğinden (Miles & Huberman, 1994), ortaya çıkan bu sonuç araştırma için güvenilir kabul edilmiştir.

3.6.2. Nicel Verilerin Analizi

Araştırmada nicel verilerin elde edilmesinde araştırmacı tarafından geliştirilen "Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketi" kullanılmıştır. Anketlerden elde edilen verilerin analizinde SPSS paket programları kullanılmıştır. Verilerin analizinde madde bazında betimsel inceleme yapılmış ve betimsel istatistik tekniklerinden aritmetik ortalama ve standart sapma puanları hesaplanmıştır. Beşli Likert tipi 42 maddeden oluşan anketten elde edilen puan ortalamalarının yorumlanmasında Tablo 5'te belirtilen değer aralıkları dikkate alınmıştır.

Tablo 4. Veri toplama aracında kullanılan değer aralıkları

Seçenekler	Veri Ağırlığı	Değer Ağırlığı
Her zaman	5	4.21-5.00
Çoğu zaman	4	3.41-4.20
Ara sıra	3	2.61-3.40
Nadiren	2	1.81-2.60
Hiçbir zaman	1	1.00-1.80

4. BULGULAR

Araştırmanın nitel verileri inceledikten sonra kodlamalar yapılmış ve temalar ile alt temalar belirlenmiştir. Elde edilen veriler doğrultusunda İngilizce öğretiminde ölçme ve değerlendirme sürecinin incelenmesi "Hazırlık", "Uygulama", "Değerlendirme" ve "Geribildirim" olmak üzere dört ana/üst temaya ayrılmıştır. Daha sonra nitel veriler ve ilgili alanyazı göz önünde bulundurularak hazırlanan anket ile nicel veriler toplanmıştır. Nicel veriler nitel araştırmadan elde edilen temalar altında tablolar halinde sunulmuştur.

Araştırmada "Hazırlık" ana temasındaki *"Bir İngilizce testi geliştirirken planlama yapıyor musunuz? Planlama yapıyorsanız nelere dikkat ettiğinizi açıklayabilir misiniz"* sorusuna verilen yanıtlar doğrultusunda elde edilen tema ve alt temalar Tablo 5'te gösterilmiştir.

Tablo 5. "Bir İngilizce testi geliştirirken planlama yapıyor musunuz? Planlama yapıyorsanız nelere dikkat ettiğinizi açıklayabilir misiniz" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları

Temalar ve Alt temalar	**Frekans**
A. Ölçülecek Özellikleri Planlama	
A.1. Ünite kazanımlarına uygun olan davranışları belirlerim	11
A.2. Sınıf düzeyine uygun olan davranışları belirlerim	20
A.3. Kazanılması zorunlu olan davranışları seçerim	2
A.4. Hedef davranışların tümünü belirlerim	4
B. Ölçme Aracını Planlama	
B.1. Testin amacını belirlerim	1
B.2. Testin amacına uygun test türü belirlerim	4
C. Soruların Niteliğini Planlama	
C.1. Soru sayısını belirlerim	19
C.2.Madde türleri ya da çeşitlerini belirlerim.	3
C.3. Soruların güçlük düzeyini belirlerim	20
C.4. Sınıf düzeyine uygun sorular seçerim	21
C.5.Öğrencilerin hazırbulunuşluklarına dikkat ederim	4
D. Uygulama Süresini Planlama	5
E. Test Puanlarının Dağılımını Planlama	11

Tablo 5'te İngilizce öğretmenlerinin bir ölçme aracı kullanırken veya geliştirirken göz önünde bulundurdukları durumlar "Ölçülecek özellikleri planlama", "Ölçme aracını planlama", "Soruların niteliğini planlama", "Uygulama süresini planlama" ve "Test puanlarının dağılımının planlanması" temaları altında sunulmuştur. Öğretmenlerin bir testi planlarken dikkat ettikleri en önemli alt temanın soruların sınıf düzeyine uygunluğunun belirlenmesi ile ilgili olduğu görülmektedir. İngilizce öğretmenleriyle yapılan görüşmelerden elde edilen verilere göre, öğretmenlerin ölçülecek özellikleri planlarken en gerekli gördükleri ölçüt sınıf düzeyine uygun olan kazanımları belirlemektir. Bu temayı takip eden diğer alt tema ünitelerdeki kazanımlara uygun olan davranışların seçilmesidir. Soruların niteliğini planlama temasına dair görüşleri incelendiğinde ise, soru sayını belirleme ve sınıf düzeyine uygun sorular seçmenin en yüksek frekansa sahip diğer alt temalar olduğu görülmektedir. İngilizce öğretmenlerinin önemsedikleri bir diğer süreç de test puanlarının uygun şekilde dağılımının yapılması için yapılan planlamalardır. Ölçülecek özelliklerin ve soruların niteliğinin planlanması temalarına ait bazı öğretmen görüşleri şu şekildedir:

> *"Dört temel beceriyi ölçecek şekilde bir planlama yapmamız gerektiği halde sınıf mevcutları çok olduğu için ancak okuduğunu anlaması ve kelime ile dilbilgisini ölçecek şekilde bir planlamaya gitmek zorunda kalıyoruz." (Ö3)*

> *"Ünite kazanımlarına göre bir planlama yapıyorum. Değerlendirme sorularının bir kısmı kolay; herkesin yapabileceği şekilde, bir kısmı orta düzey ve bir kısmı da zor olacak şekilde bir planlama yapıyorum." (Ö5)*

> *"İşlenen konulara ve önem derecesine göre soru dağılımını gerçekleştiriyorum. Sınıf seviyesi ve öğrencilerin hazırbulunuşluk durumlarına göre planlama yapıyorum" (Ö6)*

> *"Ölçmede MEB'in belirlediği müfredata uygun kazanımları ölçmek hedefimdir. Bu yüzden bu kazanımlara göre soru hazırlarım. Soru sayısı genelde çoktur; çünkü kısa cevaplı sorular hazırlarım. İyi öğrenciler belirlemek için çok zor sorulara %20-25, seviyesi düşük öğrencilerin derse motivasyonunu artırmak için çok kolay sorulara %30-35 yer veririm. Geriye kalan sorular orta zorluktadır. Puanlama yönergesi sınav kâğıdında her zaman bellidir."(Ö11)*

> *"Planlama yapıyorum. Öncelikle sınıf düzeyini gözeterek bir ölçme aracı geliştirmeye dikkat ederim. Örneğin, girdiğim tüm 5. sınıflarda aynı konuları işlememe rağmen, tüm sınıflara aynı ölçme aracını kullanmıyorum. Aynı şekilde, bir sınıfta tüm öğrenciler aynı seviyelerde olmadığı için*

sınıftaki tüm öğrencileri gözeterek bir planlama yapıyorum. Güçlük düzeyini orta seviyede tutmaya çalışırım, fakat kim zaman düşük seviyedeki öğrenciler için çok kolay soruları barındıran bölümleri de kullanırım." (Ö12)

"Planlama yaparken soru sayısı ve güçlük düzeyini dikkate alıyorum. Ders saatinde çözülebilecek yoğunlukta ve tabi ki sınıf ortalamasına denk gelebilecek zorluktaki soruları hazırlıyorum. Bazı soruları kolay sormaya özen gösterirken diğer taraftan da konuların her yönüyle kavrandığını (bazı öğrenciler tarafından) anlamak için zor sorular kullanıyorum. Puanlama ve yönerge hazırlıyorum." (Ö29; ÖO4)

"Elbette planlama yaparım. Yaptığım her ölçmenin dört dil becerisini de kapsamasına çok dikkat ederim ama bazı sınavlarda dinleme becerilerini ölçememekteyim. Bu yüzden daha sonra bu edinimi ölçecek sınavlar hazırlıyorum. Sınavlarda en kolay sorudan en eleyici soruya kadar her türlü soruyu sorarım." (Ö31;ÖO6).

Araştırmaya katılan öğretmenlerin tümünün İngilizce dersi için bir ölçme aracı geliştirirken planlama yaptıkları görülmektedir. Planlama yapılmadığına dair bir tema yoktur. Çoğu öğretmenin planlama yaparken izledikleri ilk adımın sınıf düzeyi ve öğretim programındaki kazanımlara uygun olan davranışları belirlemek olduğu söylenebilir. Ölçme aracında yer alan sorular hazırlanırken ise, öğretmenlerin göz önünde bulundurduğu en önemli ölçüt öğrencilerin bilişsel olarak hazır olmalarıdır. Ayrıca ölçme sorularının farklı seviyelerdeki öğrencilere uygun olacak şekilde kolay, orta ve zor güçlükte hazırlanması çalışmaya katılan öğretmenlerin vurguladıkları ortak görüşlerdendir.

Araştırmada "Hazırlık" ana temasında yer alan bir diğer soru İngilizce öğretmenlerinin bir ölçme aracı geliştirirken veya kullanmaya karar verirken geçerlik, güvenirlik ve kullanışlık özelliklerinin sağlanması için nelere dikkat ettikleriyle ilgilidir. Bu soruya verilen yanıtlar doğrultusunda belirlenen tema ve alt temalar Tablo 6'da gösterilmiştir.

Tablo 6. "Ölçme aracı geliştirirken veya bir ölçme aracını kullanmaya karar verirken geçerlik, güvenirlik ve kullanışlık özelliklerinin sağlanması için nelere dikkat ettiğinizi açıklar mısınız?" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları

Tema ve Alt Temalar	Frekans
A. Geçerlik	
A.1. Amaca uygun test türü belirlenmesi	3

A.2. Kazanımlara uygun olması	11
A.3. Farklı sınıf düzeylerine uygun maddeler hazırlanması	5
A.4. Öğretim programına uygun ölçütler geliştirme	1
A.5. Tüm yabancı dil becerileri alanlarının ölçülmesi	2
A.6. Testin bilen ile bilmeyeni ayırt etmesi	1
A.7. Kapsam geçerliğinin sağlanması	6
A.8. Derste işlenenlerle testin örtüşmesi	5
B. Güvenirlik	
B.1. Soru sayısını arttırma	1
B.2. Soruların açık ve anlaşılır olması	3
B.3. Sınav süresinin yeterli olması	5
B.4. Testin orta güçlükte olması	4
B.5. Açıklayıcı yönergeler bulunması	2
B.6. Puanlamanın objektif olması	2
B.7. Cevap anahtarlarının test öncesi hazırlanması	2
B.8. Şans faktörünün azaltılması	1
C. Kullanışlık	
C.1. Ekonomik olması	3
C.2. Hazırlanmasının kolaylığı	4
C.3. Uygulama süresinin uygunluğu	2
C.4. Puanlamanın kolaylığı	9
D. Ölçme aracında olması gereken özelliklere uygun test hazırlamak benim için zordur.	3

Araştırmaya katılan öğretmenlerin ölçme aracında bulunması gereken özellikleri nasıl sağladıklarına dair görüşleri incelendiğinde, "Geçerlik" temasında yer alan "Kazanımlara uygun olması" ve "Kapsam geçerliğinin sağlanması" alt temalarının en yüksek frekansa sahip olduğu görülmektedir. Ayrıca "Derste işlenenlerle testin örtüşmesi" alt teması da ölçme aracının geçerliğini sağlamak için öğretmenlerin başvurduğu yollardan biridir. "Güvenirlik" temasına dair görüşler incelendiğinde ise, "Sınav süresinin yeterli olması" alt temasının öğretmenlerin en çok vurguladıkları güvenirlik ölçütü olduğu söylenebilir. Testin zorluğunun sınıf düzeyi göz önünde bulundurularak "Orta güçlükte olması" güvenirliği sağlamanın bir başka yoludur. Ölçme aracında olması gereken "Kullanışlık" temasına dair görüşler analiz edildiğinde, "Puanlamanın kolaylığı" alt temasının en yüksek frekansa sahip görüş olduğu görülmektedir. Bunun yanında ölçme aracındaki özelliklerin sağlanmasının zorluğuna dikkat çeken görüşler de bildirilmiştir. Bir

ölçme aracının sahip olması gereken geçerlik, güvenirlik ve kullanışlık özelliklerinin nasıl sağlandığıyla ilgili öğretmen görüşleri şu şekildedir:

> *"Bu özellikleri dikkate alarak sınavı yapmak temel amacımız. Asıl sıkıntı sınıflarda 15 iyi öğrenci, 15 orta öğrenci ve 20 çok zayıf öğrencinin olması. Dolayısıyla öğrenci sayıları ve seviyeleri arasındaki farklılıktan dolayı ölçme aracında olması gereken bu özelliklere uyabildiğim söylenemez." (Ö2)*

> *"Geçerlik konusunda istenen kazanımı ölçmesine dikkat ederim. Güvenirlik konusunda aynı seviyedeki sınıflarda aynı sonuçları alıyorsam ölçme aracını kullanmaya devam ederim. Kullanışlık konusunda cevaplandırılmasının kolay olmasına ve ölçme aracının ekonomik olmasına dikkat ederim." (Ö4)*

> *"Ölçme aracının güvenirliğinin sağlanması için puan dağılımına önceden karar veririm, zorluk derecesi öğrenciye uygun olan sorular hazırlarım ve soruların açık ve anlaşılır olmasına dikkat ederim. Ayrıca kapsam geçerliği göz önüne alıp, sorulacak soru sayısını belirlerim" (Ö22)*

> *"Ölçme aracında hazırladığım sorular işlediğim konularla bütünlük sağladığı için geçerliği yüksektir. Genel olarak bileni ve bilmeyeni de ayırt ettiği için güvenilirdir. Sınav süresi ve soru sayısı standartlara uygun olduğu için kullanışlık özelliği de yüksektir denilebilir." (Ö23)*

> *"Güvenirliği arttırmak için soruların yalın ve açık olmasına ve soru sayısını arttırmaya dikkat ederim. Sınav süresini uygun ayarlamaya çalışırım. Geçerliği arttırmak için derste öğrenilen konulara yönelik sorular sorarım. Hazırladığım testlerin çoğaltılmasının ekonomik olması ve analizinin pratik olması da kullanışlılığı arttırır." (Ö25)*

> *"Kullanacağım testin amacına uygun şekilde hazırlanması, soruların derste işlenen konulara uygunluğu, öğrenci seviyesine uygun olması ve yeterli sürede çözülebilmesi geçerlik ve güvenirliğini arttırır. Hazırlama ve puanlaması bakımından da tasarruf sağlaması kullanışlı olduğunu gösterir." (Ö27; ÖO2)*

> *"Kesinlikle her hazırladığım testin geçerlik ve güvenirliğinin tam olması gerekmektedir. İlk olarak Neyi ölçeceğim? ve Nasıl ölçeceğim? demeliyim. Örneğin, bir testte çocukların okuma ve anlama becerilerini ölçmek istiyorum. Mutlaka bu edimlere göre sorular hazırlarım. Diyelim ki bu ölçme 5 aşamadan oluşacak; her aşamayı kolaydan zora doğru sorular hazırlayarak oluştururum ve puanlamaya öyle karar veririm. Daha sonra aynı sınavı bir süre sonra tekrar uygularım. Sonuçlar tutarlılık gösterirse testin güvenirliği de söz konusudur." (Ö31;ÖO6)*

"Ölçme aracı hazırlarken geçerlik, güvenirlik ve kullanışlılığa dikkat etmeye çalışırım. Soruların açık ve anlaşılır olmasına, derste işlenen konulara uygunluğuna dikkat ederim. Sınıf mevcutları kalabalık olduğu için puanlaması kolay olabilecek sorular hazırlarım." (Ö34;ÖO9).

Araştırmaya katılan öğretmenlerin büyük bir çoğunluğunun İngilizce dersi için bir ölçme aracı geliştirirken veya bir ölçme aracını kullanmaya karar verirken geçerlik, güvenirlik ve kullanışlık özelliklerinin sağlanmasına dikkat ettikleri söylenebilir. Testin geçerliğini sağlamak için öğretmenlerin ünite kazanımlarına ve testin amacına uygun olan davranışları seçtiği ve farklı sınıf düzeylerine uygun olacak şekilde test araçlarını belirledikleri görülmektedir. Öğretmenlerin testin güvenirliğini sağlamak için göz önünde bulundurdukları ölçütler ise soruların açık ve anlaşılır olması, soru sayısının ve test süresinin uygun olmasıdır. Ayrıca testin orta güçlükte hazırlanmasının güvenirliği arttırdığına dair görüşler bildirilmiştir. Araştırmaya katılan öğretmenler testin kullanışlığını sağlamak için puanlama ve ölçme aracını hazırlamada zaman açısından kolaylık sağlayacak sorular belirlediklerini ifade etmişlerdir. Diğer taraftan öğretmenler bir ölçme aracı hazırlarken sınıf mevcutlarının kalabalık olması ve öğrenci seviyeleri arasındaki farklılıktan dolayı geçerlik, güvenirlik ve kullanışlık özelliklerine dikkat edemediklerini de belirtmişlerdir.

Araştırmada "Hazırlık" ana temasındaki bir diğer konu İngilizce öğretmenlerinin hazırladıkları ölçme araçlarının öğretim programındaki kazanımlara uygun olması için neleri ne düzeyde yaptıklarıyla ilgilidir. Bu soruya verilen yanıtlar doğrultusunda belirlenen temalar ve alt temalar Tablo 7'de gösterilmiştir.

Tablo 7. "Hazırladığınız ölçme araçlarının İngilizce öğretim programının kazanımlarına uygun olması için nelere ve ne düzeyde dikkat ettiğinizi açıklar mısınız? sorusuna verilen cevaplardan elde edilen görüşler ve frekansları

Tema ve Alt Temalar	Frekans
A. Yeteri Kadar Dikkat Ediyorum	
A.1. Yıllık planları göz önünde bulundururum.	5
A.2. Kazanımlara uygun farklı ölçme araçları geliştiririm.	8
A.3. MEB'in hazırladığı İngilizce ders kitaplarına uygun ölçme araçları belirlerim.	1
A.4. Programda önerilen ölçme ve değerlendirme tekniklerini kullanırım.	1
A.5. Klasik ve tamamlayıcı ölçme değerlendirme araçlarını	4

birlikte kullanırım.	
A.6. Süreç ve ürün değerlendirmesini birlikte yaparım	3
A.7. Geçerlik, güvenirlik ve kullanışlık özelliklerinin sağlanmış olmasına dikkat ederim.	2
B) Kısmen Dikkat Ediyorum	
B.1. Sadece temel kazanımlara uygun olan bir ölçme aracı hazırlarım	3
B.2. Öğrenci seviyesine uygun olan kazanımları ölçerim.	12
B.3. Okulun fiziki şartları tüm beceri alanların ölçülmesini kısıtlamaktadır.	4
B.4. İngilizce ders saatlerinin sınırlı olması tüm kazanımların ölçülmesini aksatmaktadır	3
B.5. Sadece klasik ölçme araçlarını kullanıyorum	2
B.6. Ürün değerlendirmesine ağırlık veriyorum.	3
B.7. Sınav odaklı eğitim anlayışı ölçme ve değerlendirme sürecini sınırlamaktadır	1

Tablo 7'de araştırmaya katılan İngilizce öğretmenlerinin hazırladıkları ölçme araçlarının İngilizce öğretim programının kazanımlarına uygun olması için nelere ve ne düzeyde önem verdikleri "Yeteri kadar dikkat ediyorum" ve "Kısmen dikkat ediyorum" temaları altında sunulmuştur. Araştırmanın verilerine göre, ölçme araçlarının kazanımlara uygunluğuna yeterli ölçüde dikkat eden İngilizce öğretmenlerinin vurguladıkları en sık uygulama "Kazanımlara uygun farklı ölçme araçları geliştiririm" alt temasıdır. Öğretmenlerin bu tema altında en sık dile getirdikleri bir diğer görüş ise "Yıllık planları göz önünde bulundururum" alt temasıdır. İngilizce öğretim programındaki kazanımlara uygun ölçme araçları hazırlamaya kısmen dikkat eden öğretmenlerin görüşleri incelendiğinde ise çeşitli faktörlerin onları sınırlandırdığı görülmektedir. Bu noktada öğretmenlerin ölçme araçlarını hazırlarken kazanımlara son derece uygun olmasından ziyade daha fazla dikkat ettikleri durum "Öğrenci seviyesine uygun olan kazanımları ölçerim" alt teması olmuştur. Öğretmenleri en çok sınırlandıran bir diğer faktör ise "Okulun fiziki şartları tüm beceri alanların ölçülmesini kısıtlamaktadır" alt temasıyla belirtilmiştir. İngilizce öğretmenlerinin hazırladıkları ölçme araçlarının İngilizce öğretim programına uygun olması için nelere dikkat ettiklerini açıklayan öğretmen görüşleri şu şekildedir:

"Programda öngörülen kazanımlar çok iyi temellere dayansa da maalesef bazen öğrencilerimizin hazırbulunuşluk düzeylerine göre yeniden oluşturulabiliyor." (Ö6)

"Ölçme araçlarının öğretim programının kazanımlarına uygun olması için epey bir çaba harcıyorum. Fakat ne yazık ki, programda istenilen kazanımlar yüzde yüz uygulamaya yansıtılamıyor. Örneğin, ben İngilizce hazırlık programın olduğu bir okulda görev yapıyorum. Fakat haftada 13 saat programdaki kazanımların gerçekleştirilmesine yetmiyor. Çünkü her öğrenci aynı düzeyde değil ve hepsinin aynı kazanımları aynı seviyede edinmeleri pek mümkün değil. Buna rağmen elimden geldiği kadarıyla kazanımları göz önünde bulundurmaya çalışıyorum."(Ö12).

"Programda öngörülen kazanımlar kağıt üzerinde yeterli ama uygulamada yetersiz kalıyor. Öğrencilerin hazırbulunuşluk düzeyleri, eğitimde fırsat eşitliğinin farklı seviyelerde olması ve ekonomik koşullardan dolayı programdaki kazanımlara uygun ölçme araçları hazırlamada aksaklıklar yaşıyorum. Genelde ürün değerlendirilmesine ağırlık veriyorum."(Ö13)

"İngilizce öğretim sürecinde dört becerinin olması gerekiyor. Bunların hepsinin bir arada öğretilmesi ve değerlendirilmesi için gerekli donanımın sağlanması gerekmektedir. Fakat okulun fiziki şartlarından dolayı bu mümkün olmamaktadır. Bu süreçte bazı noktalarda aksaklık olması mümkündür. Değerlendirme yaparken bu gibi durumları da göz önünde bulundurarak ölçme araçları hazırlıyorum." (Ö18)

"Ders saati açısından ve fiziksel yetersizlikten kaynaklı maalesef öngörülen kazanımları her zaman tam tamına gerçekleştiremiyorum." (Ö21)

"Öncelikle hazırladığım ölçme araçlarının objektif bir ölçme aracı olmasına dikkat ederim. Farklı ölçme araçlarıyla öğrencileri test etmeye çalışırım. Süreçleri adım adım takip eder ve ortaya çıkan ürünün kazanımlarla ne kadar uyumlu olduğunu değerlendiririm." (Ö26;ÖO1).

"Kazanımlara uygun olmasına yeteri kadar dikkat ederim. Farklı soru tarzlarıyla kazanımları gerçekleştirmeye çalışıyorum." (Ö29;ÖO4).

Araştırmaya katılan öğretmenlerin görüşleri incelendiğinde, İngilizce öğretim programındaki kazanımları tümüyle gerçekleştirebilecek ölçme araçları hazırlamanın öğretmenler açısından güç olduğu görülmektedir. Öğretmenlerin çoğu hazırladıkları ölçme araçlarının İngilizce öğretim programıyla uyumlu olmasına kısmen dikkat ettiklerini ifade etmişlerdir. Okulun fiziki koşulları, öğrenci seviyeleri arasındaki farklılıklar, ders saatlerinin kısıtlı olması gibi çeşitli faktörlerin, öğretmenlerin öğretim programındaki kazanımların gerçekleşme düzeylerini ölçebilecek araçlar hazırlamalarını engellediği

belirtilmiştir. Bütün bu durumlar göz önünde bulundurulduğunda, eldeki kaynak ve öğrenci hazırbulunuşluklarına uygun olacak şekilde ölçülecek davranışların belirlendiği ve bu doğrultuda bir ölçme değerlendirme sürecinin yaşandığı söylenebilir.

Araştırmada ölçme değerlendirme sürecinin planlanmasından sonraki aşama olan "Uygulama" ana teması altında öğretmenlere yöneltilen soru yabancı dil beceri alanlarının (dinleme, konuşma, okuma ve yazma) değerlendirilmesi ile ilgili görüşleri olmuştur. Çalışmaya katılan öğretmenlerin verdikleri yanıtlar doğrultusunda belirlenen tema ve alt temalar Tablo 8'de gösterilmiştir.

Tablo 8. "İngilizce öğretim programında öngörülen tüm dil beceri alanlarının (dinleme, konuşma, okuma ve yazma) değerlendirilmesi ile ilgili görüşleriniz nelerdir?" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları

Tema ve Alt Temalar	Frekans
A. Dinleme Becerilerinin Değerlendirilmesi	
A.1. Dinleme becerisini ölçmeye dayalı teknikler kullanırım	10
A.2. Dinleme becerilerini ölçmek benim için zordur	20
A.3. Ders kitabındaki dinleme metinleri çok uzun	1
A.4. Okulun fiziki şartları yetersiz	8
A.5. Sınıf mevcutları kalabalık olduğundan ölçülmesi zor	6
A.6. Sınav sisteminden dolayı göz ardı ediliyor	5
A.7. Öğrenciler için zorlayıcı	3
A.8. Öğrencilerin motivasyonunu düşürüyor	1
A.9. Ders saatleri yeterli değil	4
B. Konuşma Becerilerinin Değerlendirilmesi	
B.1. Konuşma becerisini ölçmeye dayalı teknikler kullanırım	10
B.2. Konuşma becerilerini ölçmek benim için zordur.	22
B.3. Öğrenciler için zorlayıcı	3
B.4. Konuşma kaygısı yaşanıyor	3
B.5. Sınav sisteminden dolayı göz ardı ediliyor	5
B.6. Sınıflar kalabalık olduğundan ölçülmesi zor	6
B.7. Ders saatleri yeterli değil	4
C. Okuma Becerilerinin Değerlendirilmesi	
C.1. Okuma becerisini ölçmeye dayalı teknikler kullanırım.	22
C.2. Herhangi bir sıkıntı yaşamıyorum	15
C.3. En sık ölçtüğüm beceri alanıdır	5
C.4. Sınav sisteminden dolayı önemli görülen bir alandır	5
D. Yazma Becerilerinin Değerlendirilmesi	

D.1. Daha çok yazılı yoklamalarla ölçme yapıyorum	4
D.2. Tamamlayıcı ölçme ve değerlendirme tekniklerini kullanıyorum	3
D.3. Herhangi bir sıkıntı yaşamıyorum	5
D.4. Ders kitabındaki uygulamaları takip ediyorum	3
D.5. Yazma becerilerini ölçmeye daha az zaman ayırırım	3
D.6. Yazma becerilerini ölçmede zorlanırım	3
D.7. Güvenilir ölçme sonuçlarına ulaşmak zordur.	1

Tablo 8'de araştırmaya katılan öğretmenlerin İngilizce öğretim programında öngörülen tüm dil beceri alanlarının ölçülmesi ile ilgili görüşleri "Dinleme becerilerinin değerlendirilmesi", "Konuşma becerilerinin değerlendirilmesi", "Okuma becerilerinin değerlendirilmesi" ve "Yazma becerilerinin değerlendirilmesi" temaları altında sunulmuştur. Bu temalar içerisinde en fazla üzerinde durulan temanın dinleme becerileri ile ilgili olduğu anlaşılmaktadır. Dinleme becerilerinin ölçülmesinde öğretmenler tarafından en çok "Dinleme becerilerini ölçmek benim için zordur" alt temasının ifade edildiği görülmektedir. Bu temaya ilişkin bir diğer önemli alt tema ise dinleme becerilerinin ölçülmesinin önemini vurgulayan "Dinleme becerisini ölçmeye dayalı teknikler kullanırım" görüşüdür. Konuşma becerilerinin ölçülmesi temasına dair görüşler incelendiğinde ise, dinleme alanında olduğu gibi bu alanın da değerlendirilmesi yönelik zorluğa dikkat çeken "Konuşma becerilerini ölçmek benim için zordur" alt temasının öne çıktığı görülmektedir. Herhangi bir olumsuz görüşün kaydedilmediği okuma becerilerinin ölçülmesi temasında, öğretmenlerin en çok ifade ettikleri alt tema "Okuma becerisini ölçmeye dayalı teknikler kullanırım" olmuştur. Bu temanın büyük çoğunluğunu oluşturan diğer görüş ise okuma becerilerini ölçme ve değerlendirme sürecinde "Herhangi bir sıkıntı yaşamıyorum" alt temasıdır. Yazma becerilerinin ölçülmesi temasına ilişkin de pek çok farklı görüşün olduğu ve bunların içerisinde en çok vurgulanan ifadenin "Herhangi bir sıkıntı yaşamıyorum" alt teması olduğu anlaşılmaktadır. Araştırmaya katılan öğretmenlerin yabancı dil beceri alanlarının ölçülmesi ile ilgili görüşleri ise şu şekildedir:

> *"Küçük yaş gruplarında dinleme becerilerini geliştirmeye çalışırım ve gerekli görüyorum. Diğer öğrenciler için becerilerin hepsinin eşit ağırlıkta kullanılması ve ölçülmesi gerektiğini düşünüyorum. Ancak konuşma ve dinleme becerilerini ölçme uygulamaları pratikte ne yazık ki pek yer bulamıyor" (Ö4).*

> *"Aslında dil ediniminde ve öğretiminde tüm beceri alanlarının ölçülmesi esas alınmalıdır. Fakat bazı alt yapı eksiklikleri nedeni ile bu tüm*

okullarımızda çok sağlıklı bir şekilde ilerletilememektedir. Konuşma ve dinleme dil becerilerinin ölçülmesi kalabalık sınıflarımızda rahat ve anlaşılır bir şekilde yapılamamaktadır" (Ö6).

"Dört dil becerisini de kullanıyoruz ama öğrenciler özellikle konuşma ve yazmada zorlanıyor. Dinleme etkinliğinde paragrafı veya diyaloğu birkaç defa dinletmem gerekiyor. Okuma ve telaffuza önem veriyorum ve sık sık tekrar yaptırıyorum. Yazma çok sıkıntılı bir süreç onlar için. Türkçe düşünüp Türkçe dil kurallarına göre yazmaya çalışıyorlar. Bizim okul mevcudumuz kalabalık ve 50 kişilik sınıflarda sınırlı bir zaman diliminde özellikle konuşma ve dinleme becerilerini ölçmeyi tam olarak yetiştiremiyorum. Bu da bizi ve öğrencileri zorluyor." (Ö9)

"Faydalı olduğuna inanıyorum ama öğrenciler dinlemede bazen zorlanıyor. Birkaç dinleterek üstesinden gelmeye çalışıyorum. Bazen ders kitaplarındaki dinleme metinleri seviyelerinin üzerinde (çok uzun) veya zor anlaşılır (telaffuz) olduğu için öğrencinin motivasyonu bozuluyor ve dinlemeyi bırakıyor. Yazma becerilerini ölçmek benim için zor oluyor. Çünkü evde yapsalar da başkalarından yardım alarak yaptıkları için seviye üstü oluyor. Ya da çok karmaşık cümleler yazıyorlar. Türkçeden İngilizceye bire bir çevirdikleri için çok hata yapıyorlar. Bunları 40 kişilik sınıflarda kontrol etmek hem de öğrencilerle birlikte çok zor. Bu yüzden yazma ve konuşma becerilerini ders sırasında daha kısa ve öz cümleler kullanarak hep beraber yaptırıyorum" (Ö11)

"Her becerinin mutlak suretle ölçülmesi gerekir. Ben haftalık olarak her bir beceri için bir değerlendirme hazırlar ve uygularım. Bu beceriler ölçülmezse öğrencilerin İngilizceyi gerçek anlamda öğrenmeleri imkansız. Fakat sıkıntımız şu ki, 13 saatlik İngilizce derslerinde deneme sınavlarına yakın zamanlarda öğrencilere bolca test çözme etkinliği yaptırıyoruz. Bunu çok yanlış görsem de deneme sınavındaki 10 soruluk İngilizce testleri, çocukların İngilizceyi konuşmasından daha fazla önemseniyor. Böyle bir sistem olmasaydı belki çok daha verimli İngilizce dersleri geçirebilirdik." (Ö12)

"Ülkemizde maalesef dinleme becerileri çok fazla ölçülemiyor. Bu alana daha fazla ağırlık verilerek bu sorunun da çözüleceğine inanıyorum. Konuşma ve okuma becerilerini sözlü, yazma becerisini ise yazılı olarak ölçüyorum. Okulumuz proje okulu olduğu için 5. sınıflarda İngilizce 7 saat ana ders, 4 saat okuma ve yazma 2 saat de seçmeli İngilizce şeklinde bölüştürülmüş. Bu nedenle bu becerilerin elde edilmesinde ve ölçülmesinde zaman yönünden bir sıkıntı yok." (Ö13).

"Okuma ve yazma konusunda çok sıkıntı yaşamıyorum. Öğrencilerime maksimum seviyede faydalı olmaya çalışıyorum. Dinleme fiziki şartlarından

dolayı bazen aksayabiliyor. Konuşma konusunda çok verimli olamıyoruz. Çocukların en çekimser kaldıkları konu konuşma. Bu alanlarda etkinlikler sırasında gözlem yaparak değerlendirme yapıyorum. Ama ayrı bir sınav yapamıyorum" (Ö17).

"Tüm dil beceri alanları bana göre yetersiz kalmaktadır. Özellikle dinleme kısmında materyal eksikliği, okul durumu göz önünde bulundurulursa çok büyük eksiklik var. Dinleme olmayınca konuşma ve telaffuz alanlarında da çok eksik kalmaktayız" (Ö18)

"Şu an içerisinde bulunduğumuz sistemde öne çıkan dil becerisi ve neredeyse sadece ona odaklanılan yeterlilik okumadır. Okullarda pek çok sebepten ötürü dinleme ve yazma becerileri ihmal edilmekte. Konuşma ise çok cüzi bir şekilde yürütülmektedir. Bu yüzden yapılan ölçümler buna yönelik olup diğer dil becerileri ihmal edilmektedir" (Ö22).

"Tüm dil becerilerin ölçülmesinden yanayım. Ölçülmeyen becerinin öğrenilmesi ve önemsenmesi de zorlaşır. Özellikle konuşma ve dinleme becerilerini ölçmeye dayalı teknikler kullanırım. Çünkü bu iki beceri iletişimde en önemli alanlardır." (Ö24).

"İngilizce öğretim programında öngörülen bu dil becerilerinin mutlaka ölçülmesi gerekmektedir. Bu edinimler birbirinden ayrı düşünülemez. Her kazanılan edinim bir sonrakinin yardımcısıdır. Örneğin; dinleme becerisi yani kulak dolgunluğu kazanamayan bir öğrencinin diğer beceri alanlarında başarılı olması oldukça güçtür. Dolayısıyla tüm beceri alanlarında değerlendirme yapıyorum" (Ö31; ÖO6).

"Tüm becerilerin geliştirilmesi ve ölçülmesi gerekir. Ancak dinleme becerisi imkansızlıktan dolayı ölçülemiyor. Konuşma becerisi de sınıflar kalabalık olduğu için. Ama bunlar da ders içinde gözlem yoluyla ölçülebilir." (Ö34; ÖO9).

"Ders saatleri dinleme ve konuşma becerilerinin ölçülmesi için yeterli değil. Daha çok okuma ve yazma becerilerinin değerlendirilmesine ağırlık veriyorum." (Ö36; İHO;1)

Araştırmaya katılan öğretmenlerin görüşleri incelendiğinde, yabancı dil beceri alanlarının İngilizce öğretim programında belirtildiği şekilde değerlendirilmesi gerektiğine ilişkin bir görüş birliği olduğu söylenebilir. Ancak uygulama aşamasında öğretmenlerin çoğu çeşitli sebeplerden dolayı bazı becerilerin eksik bırakıldığını veya ihmal edildiğini ifade etmişlerdir. Özellikle dinleme ve konuşma becerilerinin ölçülmesinin hem öğretmenler hem de öğrenciler için oldukça zorlayıcı olduğu anlaşılmaktadır. Bu alanlarda ne düzeyde öğrenme sağlandığına dair yeterli değerlendirmelerin yapılamadığı

görülmektedir. Öğretmenlerin görüşlerine göre bu durumun en önemli sebepleri; sınıf mevcutlarının kalabalıklığı, materyal eksiklikleri ve ders saatlerinin az olması gibi fiziki şartların yetersizliği ve var olan sınav sisteminden dolayı diğer beceri alanlarında yaşanan kaygıların iletişim becerilerine olan ilgiyi geride bırakmasıdır. Yaşanan sınav kaygılarının okuma becerilerini ise olumlu yönde etkilediği görülmektedir. Nitekim, sınav odaklı bir yaklaşımın okuduğunu anlamaya yönelik ölçümleri gerektirdiği ve daha çok bu alanın değerlendirilmesine yönelik uygulamalar yapıldığı ifade edilmektedir. Yazma becerilerinin değerlendirilmesine ilişkin çeşitli görüşler bildirilmekle birlikte, bu alanın ne konuşma ve dinleme becerileri kadar ihmal edildiği ne de okuma becerileri kadar önemsendiği söylenebilir.

Bu tema altında katılımcıların yabancı dil beceri alanlarının değerlendirilmesine yönelik görüşleri göz önünde bulundurularak hazırlanan Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketinin verilerine yer verilmiştir.

4.1. İngilizce Öğretmenlerinin Yabancı Dil Becerilerini Ölçme ve Değerlendirme Düzeylerine İlişkin Bulgular

Araştırmanın nitel araştırma sonuçları ve ilgili alanyazı göz önünde bulundurularak oluşturulan Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketinden elde edilen verilerin ortalama, standart sapma ve değer aralığını gösteren düzeyleri Tablo 9, Tablo 10, Tablo 11, ve Tablo 12'de gösterilmiştir.

Tablo 9. Öğretmenlerin dinleme becerilerini ölçme ve değerlendirme süreçlerine ilişkin ortalamaları, standart sapmaları ve düzeyleri

Maddeler	$\bar{X}$	Ss	Düzey
1.İngilizce dersi öğretim programındaki kazanımlara uygun metinler seçerim.	4,23	,87	Her zaman
2.Öğrencilerin dikkat sürelerini aşan metinler seçmemeye dikkat ederim.	4,27	,90	Her zaman
3.Sınıf içi etkinlikler sırasında öğrencilerin dinleme becerileriyle ilgili değerlendirmeler yaparım.	3,69	,96	Çoğu zaman
4. Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım	2,83	1,20	Ara sıra
5. Öğrencilerin dilin ses, sözcük gibi öğelerine ne ölçüde hakim olduklarını değerlendiririm.	3,84	,87	Çoğu zaman
6.Öğrencilerin dilbilgisi kuralları ve kalıplarını ne ölçüde anladıklarını değerlendiririm	3,92	,89	Çoğu zaman

7.Öğrencilerin dinleme metinlerini ne kadar anladıklarıyla ilgili değerlendirmeler yaparım	3,90	,94	Çoğu zaman
8.İngilizce dersi öğretim programında önerilen test tekniklerini kullanırım	3,90	1,09	Çoğu zaman
9. Tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) kullanırım	3,37	,97	Çoğu zaman
10. Sınıf içi etkinlikler yapılırken veya sınav esnasında uygun bir dinleme ortamı oluşturmada güçlük yaşarım	2,63	1,10	Ara sıra

Tablo 9'da İngilizce öğretmenlerinin dinleme becerilerini ölçme ve değerlendirme süreçlerine dair görüşleri incelendiğinde, "Öğrencilerin dikkat sürelerini aşan metinler seçmemeye dikkat ederim ($\bar{X}$=4,27)" ve "İngilizce dersi öğretim programındaki kazanımlara uygun metinler seçerim ($\bar{X}$=4,23) maddelerinin en yüksek ortalamaya sahip olduğu ve "Her zaman" düzeyine denk geldiği saptanmıştır. "Öğrencilerin dilbilgisi kuralları ve kalıplarını ne ölçüde anladıklarını değerlendiririm ($\bar{X}$= 3,92)*"* görüşünün ise "Çoğu zaman" düzeyine denk gelen en yüksek ortalamaya sahip madde olduğu belirlenmiştir. Diğer taraftan "Sınıf içi etkinlikler yapılırken veya sınav esnasında uygun bir dinleme ortamı oluşturmada güçlük yaşarım ($\bar{X}$= 2,63)" ve "Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım ($\bar{X}$= 2,83)*"* maddelerinin en düşük ortalamaya sahip olduğu ve katılımcıların bu maddelere "Ara sıra" düzeyinde katıldıkları görülmektedir.

Tablo 10. Öğretmenlerin konuşma becerilerini ölçme ve değerlendirme süreçlerine ilişkin ortalamaları, standart sapmaları ve düzeyleri

Maddeler	**$\bar{X}$**	**Ss**	**Düzey**
1. İngilizce dersi öğretim programındaki kazanımlara uygun bir ölçme aracı hazırlarım.	3,19	1,52	Ara sıra
2. Sınıf içi etkinlikler sırasında öğrencilerin konuşma becerileriyle ilgili değerlendirmeler yaparım	3,85	,85	Çoğu zaman
3. Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım	3,17	1,16	Ara sıra
4. İngilizce dersi öğretim programında önerilen test tekniklerini kullanırım	3,12	1,32	Ara sıra
5. Öğrencilerin konuşmalarının doğru ve anlaşılır olup olmadığını değerlendiririm.	3,99	,85	Çoğu zaman
6. Öğrencilerin bilgiyi akıcı bir şekilde aktarıp aktaramadıklarını değerlendiririm	3,92	,94	Çoğu zaman

7. Öğrencilerin gerekli sözcükleri kullanabilme becerilerini değerlendiririm	4,03	,88	Çoğu zaman
8. Sözcüklerin doğru telaffuz edilmesine dikkat ederim	4,14	,93	Çoğu zaman
9. Öğrencilerin dilbilgisi kuralları ve kalıplarını ne düzeyde kullanabildiklerini değerlendiririm	3,97	,92	Çoğu zaman
10. Tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) kullanırım	3,34	1,06	Çoğu zaman
11. Sınıf içi etkinlikler yapılırken veya sınav esnasında uygun bir konuşma ortamı oluşturmada güçlük yaşarım.	2,86	1,20	Ara sıra

Tablo 10'da İngilizce öğretmenlerinin konuşma becerilerini ölçme ve değerlendirme süreçlerine dair görüşleri incelendiğinde, "Sözcüklerin doğru telaffuz edilmesine dikkat ederim ($\bar{X}$=4,14)", "Öğrencilerin gerekli sözcükleri kullanabilme becerilerini değerlendiririm ($\bar{X}$=4,03)" ve "Öğrencilerin konuşmalarının doğru ve anlaşılır olup olmadığını değerlendiririm ($\bar{X}$=3,99)" maddelerinin en yüksek ortalamaya sahip olduğu ve "Çoğu zaman" düzeyine denk geldiği saptanmıştır. Diğer taraftan en düşük ortalamaya sahip maddelerin "Sınıf içi etkinlikler yapılırken veya sınav esnasında uygun bir konuşma ortamı oluşturmada güçlük yaşarım ($\bar{X}$=2,86)" ve "İngilizce dersi öğretim programında önerilen test tekniklerini kullanırım ($\bar{X}$=3,12)" olduğu ve katılımcıların bu davranışlara ilişkin "Ara sıra" düzeyinde görüş bildirdikleri görülmektedir.

Tablo 11. Öğretmenlerin okuma becerilerini ölçme ve değerlendirme süreçlerine ilişkin ortalamaları, standart sapmaları ve düzeyleri

Maddeler	**$\bar{X}$**	**Ss**	**Düzey**
1.İngilizce dersi öğretim programındaki kazanımlara uygun olan metinleri seçerim	4,56	,59	Her zaman
2. Sınıf içi etkinlikler sırasında öğrencilerin okuma becerileriyle ilgili değerlendirmeler yaparım	4,38	,73	Her zaman
3. Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım	4,08	,85	Çoğu zaman
4. İngilizce dersi öğretim programında önerilen ölçme ve değerlendirme araçlarını kullanırım.	4,12	,72	Çoğu zaman
5. Öğrencilere farklı yazın türlerine dair okuma metinleri sunarım	3,80	,90	Çoğu zaman
6. Öğrencilerin kelimeleri ve söz kalıplarını ne ölçüde tanıdıklarını değerlendiririm	4,19	,71	Çoğu zaman
7. Öğrencilerin dilbilgisi kuralları ve kalıplarını uygun bir şekilde kullanıp kullanmadıklarını değerlendiririm.	4,14	,75	Çoğu zaman

8. Öğrencilerin okuma metnindeki ana fikri belirleme, neden-sonuç bulma gibi ilişkileri nasıl yorumladıklarını değerlendiririm.	4,14	,77	Çoğu zaman
9. Tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) kullanırım	3,46	1,01	Çoğu zaman
10. Sınıf içi etkinlikler yapılırken veya sınav esnasında uygun bir okuma ortamı oluşturmada güçlük yaşarım.	2,53	1,17	Nadiren

Tablo 11'de katılımcıların İngilizce okuma becerilerini ölçme ve değerlendirme süreçlerine dair görüşleri incelendiğinde, en yüksek ortalamaya sahip görüşlerin "İngilizce dersi öğretim programındaki kazanımlara uygun olan metinleri seçerim ($\overline{X}$=4,56)" ve **"**Sınıf içi etkinlikler sırasında öğrencilerin okuma becerileriyle ilgili değerlendirmeler yaparım ($\overline{X}$=4,38)" şeklinde olduğu ve "Her zaman" düzeyine denk geldiği saptanmıştır. Diğer taraftan "Sınıf içi etkinlikler yapılırken veya sınav esnasında uygun bir okuma ortamı oluşturmada güçlük yaşarım ($\overline{X}$=2,53)" maddesi ise hem okuma hem de anketin tüm bölümleri içerisinde en düşük ortalamaya sahip olan ve "Nadiren" düzeyindeki tek maddedir.

Tablo 12. Öğretmenlerin yazma becerilerini ölçme ve değerlendirme süreçlerine ilişkin ortalamaları, standart sapmaları ve düzeyleri

Maddeler	**$\overline{X}$**	**Ss**	**Düzey**
1.İngilizce dersi öğretim programındaki kazanımlara uygun bir ölçme aracı hazırlarım.	3,87	1,22	Çoğu zaman
2. Sınıf içi etkinlikler sırasında öğrencilerin yazma becerileriyle ilgili değerlendirmeler yaparım.	3,85	,95	Çoğu zaman
3. Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım	3,62	1,11	Çoğu zaman
4.Öğrencilere yazma türü, konusu, metin uzunluğu vb. ile ilgili bilgileri içeren detaylı bir yönerge sunarım	3,51	1,10	Çoğu zaman
5. Öğrencilerin farklı yazın türlerine dair metinler oluşturmalarını isterim	3,33	1,06	Ara sıra
6. Öğrencilerin yazma amacına uygun olan kelimeleri ne ölçüde kullandıklarını değerlendiririm	3,82	,99	Çoğu zaman
7.Öğrencilerin yeterli bir hızda yazıp yazmadıklarını değerlendiririm	3,31	1,07	Ara sıra
8. Öğrencilerin dilbilgisi kuralları ve kalıplarını uygun bir şekilde kullanıp kullanmadıklarını değerlendiririm	3,91	,95	Çoğu zaman
9. Yazım kurallarına dikkat ederim	3,75	1,03	Çoğu zaman
10. Tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı,	3,49	1,02	Çoğu zaman

gözlem vb.) kullanırım			
11. Sınıf içi etkinlikler yapılırken veya sınav esnasında uygun bir yazma ortamı oluşturmada güçlük yaşarım.	2,61	1,15	Ara sıra

Tablo 12'de İngilizce öğretmenlerinin yazma becerilerinin değerlendirilmesine yönelik görüşleri incelendiğinde, "Öğrencilerin dilbilgisi kuralları ve kalıplarını uygun bir şekilde kullanıp kullanmadıklarını değerlendiririm ($\bar{X}$=3,91)", "İngilizce dersi öğretim programındaki kazanımlara uygun bir ölçme aracı hazırlarım ($\bar{X}$=3,87)" ve "Sınıf içi etkinlikler sırasında öğrencilerin yazma becerileriyle ilgili değerlendirmeler yaparım ($\bar{X}$= 3,85)" maddelerinin en yüksek ortalamaya sahip olan görüşler olduğu ve "Çoğu zaman" düzeyine denk geldiği belirlenmiştir. "Sınıf içi etkinlikler yapılırken veya sınav esnasında uygun bir yazma ortamı oluşturmada güçlük yaşarım ($\bar{X}$=2,61)" ve "Öğrencilerin yeterli bir hızda yazıp yazmadıklarını değerlendiririm ($\bar{X}$= 3,31)" görüşlerinin en düşük ortalamaya sahip olduğu ve katılımcıların bu davranışları "Ara sıra" düzeyinde sergiledikleri görülmektedir.

Araştırmanın yabancı dil beceri alanlarında yapılan ölçme ve değerlendirme süreçlerine yönelik nitel verileri incelendiğinde, öğretmenlerin konuşma ve dinleme alanlarının ölçülmesinde yaşanan eksiklere dikkat çektikleri ve bu alanların sınav sisteminin iletişim becerilerini göz ardı etmesi, okulların fiziki alt yapısının yetersizliği, ders saatlerinin az olması gibi öne çıkan sebeplerden dolayı öğretim programında öngörülen şekilde değerlendirilemediği şeklinde görüşler bildirilmiştir. Özellikle bu alanlarda yapılan değerlendirmelerin sınıf içi etkinliklerle sınırlı kaldığına dair ifadeler yer bulmuştur. Dinleme ve konuşma alanlarının değerlendirilmesine yönelik bildirilen olumsuz görüşlerin olumlu görüşlerden daha fazla olduğu görülebilir. Araştırmanın dinleme ve konuşma alanlarının ölçülmesine dair nicel bulguları incelediğinde ise, öğretmenlerin bu becerilerin sınıf içi etkinlikler sırasında değerlendirildiğinde dair görüşlere "Çoğu zaman" düzeyinde katıldıkları görülmektedir. Bu alanlarda detaylı bir değerlendirme sürecinin yaşanıp yaşanmadığını sorgulayan anketin "İngilizce dersi öğretim programındaki kazanımlara uygun bir ölçme aracı hazırlarım" ve "Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım" maddelerine ise "Ara sıra" düzeyinde katıldıkları saptanmıştır. Dolayısıyla araştırmanın nitel ve nicel verilerinin birbirini desteklediği, çoğu öğretmenin ders sırasında dinleme ve konuşma etkinliklerine odaklandığı ancak bu becerilerin bir ölçme aracıyla test edilmesine yönelik faaliyetlerinin ise daha sınırlı kaldığı anlaşılmaktadır.

Araştırmanın nitel aşamasında ortaya çıkan bir diğer önemli bulgu, okuma becerilerini değerlendirme sürecinin daha fazla önemsendiği ve herhangi bir sıkıntı yaşanmadığına dair bildirilen olumlu görüşlerin dinleme ve konuşma becerilerinin aksine daha fazla olduğudur. Diğer taraftan nicel araştırmanın verileri incelendiğinde, öğretmenlerin okuma becerilerinin değerlendirilmesine yönelik sergiledikleri davranışların "Her zaman" ya da "Çoğu zaman" düzeyinde olduğu görülmektedir. Dolayısıyla hem nitel hem de nicel verilerle benzer sonuçlara ulaşıldığı ve öğretmenler için okuma becerilerini değerlendirme sürecinin diğer beceri alanlarına göre nispeten daha kolay olduğu söylenebilir.

Araştırmada "Uygulama" ana teması altında İngilizce öğretmenlerinin ölçme değerlendirme sürecinde hangi araç ve yöntemleri kullandıklarına dair görüşlerine başvurulmuştur. Tablo 13'te öğretmenlerin hangi klasik ölçme araçlarını kullandıklarına dair verdikleri yanıtlar doğrultusunda oluşturulan tema ve alt temalar sunulmuştur.

Tablo 13. "Klasik ölçme araçlarından hangi/hangilerini kullanıyorsunuz? Kullanma nedenleriniz ile birlikte açıklar mısınız?" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları

Temalar	Alt temalar	Frekans
Yazılı ifadeyi gerektiren testler (Yazılı Yoklamalar/Kısa Cevaplı Testler)	Yazma becerilerini arttırır	1
	Kazanımların gerçekleşme düzeyini en iyi gösteren araçtır	4
	Bilen ile bilmeyeni ayırt eder	1
	Üst düzey bilişsel becerileri ölçer	1
Seçmeyi Gerektiren Testler (Doğru-Yanlış Testleri/ Eşleştirme Soruları)	Hazırlanması kolaydır	5
	Puanlaması kolaydır	3
	Her seviyeye uygun hazırlanabilir	2
	Öğrencilerin en alışık olduğu yöntemdir	3
	Puanlaması objektiftir	1

İşaretlemeyi Gerektiren Testler (Çoktan Seçmeli Testler)	Kapsam geçerliği yüksektir	1
	Puanlaması kolaydır	6
	Puanlaması objektiftir	2
	Kullanımı hızlıdır	5
	Konuları pekiştirmek için daha avantajlıdır	1
	Sınav odaklı eğitim anlayışına daha uygundur	2
	Kaygı seviyesini düşürür	1
	Daha çok soru sorma imkanı sağlar	2
	Kısa sürede daha fazla beceri ölçülebilir	3
Sözlü İfadeyi Gerektiren Testler (Sözlü Yoklamalar)	Anında geribildirim olanağı sağlar	1
	Kelime bilgisini ölçmek için kullanışlıdır	2
	Öğretimin her aşamasında yapılabilir	1
	Yabancı dilde kendini ifade etme becerileri için kullanılır	2

Tablo 13'te araştırmaya katılan öğretmenlerin ölçme ve değerlendirme sürecinde en çok kullandıkları klasik ölçme araçları "Yazılı ifadeyi gerektiren testler", "Seçmeyi gerektiren testler", "İşaretlemeyi gerektiren testler" ve "Sözlü ifadeyi gerektiren testler" temaları altında sunulmuştur. Öğretmenlerin görüşlerine göre yazılı ifade gerektiren testleri kullanmalarının en önemli sebebi bu araçların "Kazanımların gerçekleşme düzeyini en iyi gösteren araç" olmalarıdır. Seçmeyi gerektiren testleri kullanmalarının ise pek çok sebebi olduğu görülmektedir. Bu tür testlerin "Hazırlanmasının kolay", "Puanlamasının kolay" ve "Öğrencilerin en alışık oldukları yöntem" olması en sık vurgulanan alt temalardır. Çoktan seçmeli testler gibi işaretlemeyi gerektiren testlerin öğretmenler tarafından daha fazla tercih edildiği söylenebilir. "Puanlamasının kolaylığı" ve "Kullanımının hızlı" olması gibi alt temalarda belirtildiği üzere bu tür testler uygulayıcılara kolaylık sağladığı için tercih edilmektedir. Sözlü yoklamalar gibi sözel ifadeyi gerektiren testlerin ise daha az kullanıldığı görülmektedir. Öğretmenlerin sözlü yoklamalarla öğrencilerin daha çok kelime dağarcıklarının zenginliğine yönelik değerlendirmeler yaptıkları "Kelime bilgisini ölçmek için kullanışlıdır" alt temasıyla ifade edilmiştir. Araştırmaya katılan İngilizce öğretmenlerinin klasik ölçme araçlarından en çok hangilerini tercih ettikleri ve bu araçları kullanma sebeplerinin neler olduğuna dair görüşler şu şekildedir:

> *"Kısa cevaplı testleri kullanırım. Hazırlanışının ve puanlamasının kullanışlı olmasına dikkat ederim." (Ö9)*

"Klasik ölçme araçlarını ve çoktan seçmeli testleri kullanırım. Çünkü bu şekilde sınavlara daha iyi hazırlanabileceklerini düşünüyorum. Klasik yazılı sınavları da dil becerilerini arttırmak için kullanıyorum." (Ö10)

"Klasik ölçme araçlarını kullanıyorum. Programdaki kazanımların kazanılıp kazanılmadığını en iyi klasik ölçme araçları kullanarak anlayabiliriz. Ayrıca zaman, mekan ve sınıf mevcudu bakımından en uygun araçlar çoktan seçmeli testlerdir." (Ö13)

"Klasik ölçme araçlarından yazılı yoklama ve sözlü yoklamalar en çok kullandığım ölçme araçları arasındadır. Öğrencilerin analiz, sentez, değerlendirme becerilerini ölçmede yazılı yoklamalar; öğrencilerin sözel anlatım, kendini ifade etme becerilerini ölçmede ise sözlü yoklamalar kullanılabilir." (Ö14)

"Daha çok yazılı yoklama tarzında olmasına dikkat ediyorum. Çoktan seçmeli olanları tercih etmiyorum. Öğrenciyi ezbere yönlendirdiğini düşünüyorum." (Ö19)

"İstemeyerek de olsa kısa cevaplı testler ile çoktan seçmeli testleri kullanıyorum. Sınav odaklı bir eğitim anlayışı geçerli olduğu ve fiziki şartlar uygun olmadığı için." (Ö20)

"Yazılı testler, çoktan seçmeli sorular, eşleştirmeli soruları genelde harmanlayıp sınav hazırlıyorum. Çünkü zorluk seviyesini ancak bu şekilde ayarlayabiliyorum. Örneğin, çoktan seçmeliler açık uçlu sorulara nazaran daha kolay olabiliyor. Ayrıca bu şekilde pek çok beceri de kısa sürede ölçülebiliyor" (Ö29;ÖO4).

"Daha çok çoktan seçmeli testleri kullanırım. Sınav odaklı bir eğitim anlayışı olduğu için öğrencilerin de beklentilerini karşılayan bir ölçme aracıdır. Ayrıca çoktan seçmeli testlerin öğrencileri daha fazla motive ettiği ve kaygı seviyelerini düşürdüğünü söyleyebilirim." (Ö32;ÖO7).

Araştırmaya katılan öğretmenlerin görüşleri incelendiğinde, İngilizce becerilerinin ölçülmesinde klasik ölçme araçları arasında en çok çoktan seçmeli testlerin tercih edildiği söylenebilir. Özellikle sınav odaklı eğitim anlayışının hem öğrencileri hem de öğretmenleri bu tür testler kullanmaya yönelttiği ifade edilmektedir. Ayrıca sınıfların kalabalık olması gibi fiziki şartlar göz önünde bulundurulduğunda, test uygulayıcıları için puanlaması ve uygulaması hızlı olan testlerin kullanımının öncelikli olduğu belirtilmiştir. Diğer taraftan doğru-yanlış testleri, kısa cevaplı sorular gibi seçmeyi gerektiren testlerin tercih edilme sebepleri incelendiğinde de aynı şekilde uygulayıcılara hazırlanması ve puanlanması bakımından kolaylık sunduğu için daha fazla kullanıldığı görülmektedir. Dolayısıyla

öğretmenlerin bir İngilizce dersi için bir ölçme aracı hazırlarken testin kullanışlı olmasına daha fazla önem verdikleri söylenebilir.

Araştırmada "Uygulama" ana teması altında İngilizce öğretmenlerine klasik ölçme araçlarının yanı sıra hangi tamamlayıcı ölçme değerlendirme araç ve tekniklerini kullandıkları sorulmuştur. Çalışmaya katılan öğretmenlerin verdikleri yanıtlar doğrultusunda belirlenen tema ve alt temalar Tablo 14'te gösterilmiştir.

Tablo 14. "Tamamlayıcı ölçme değerlendirme araç ve tekniklerinden hangi/hangilerini kullanıyorsunuz? Kullanma nedenleriniz ile birlikte açıklar mısınız?" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları

Tema	Alt Temalar	Frekans
Proje/Performans Değerlendirme	Gerçek hayatla ilişkilidir	1
	Öğrencinin daha aktif olmasını sağlar	7
	Ürün ve süreç birlikte değerlendirilir	2
	Yaratıcılığı geliştirir	1
	Öğrenme sürecini destekler	7
Öğrenci Ürün Dosyası (Portfolyo)	Bir geribildirim aracıdır	2
	Öğrencilerin gelişim sürecinin takibini sağlar	6
	Öğrenciyi daha iyi tanıma fırsatı sunar	4
	Yabancı dilde yazma becerisini geliştirir	1
	Üst düzey zihinsel becerilerin ölçülmesini sağlar	1
Kontrol Listesi / Dereceli Puanlama Anahtarı	Objektif puanlama imkanı sunar	7
	Etkili bir planlama için yol göstericidir	1
	Öğretimin etkililiği ile ilgili dönüt sağlar	1
Öz Değerlendirme	Öğrenciler güçlü ve zayıf yönlerini fark eder	2
	Öğrenciler kendi öğrenmelerinden sorumludur	1
Akran Değerlendirme	İşbirliğini arttırıyor	1
	Öğrenci merkezlidir	2
Gözlem	Diğer tekniklere göre daha kullanışlıdır	2
	Doğru ve hızlı bilgilere ulaşmayı sağlar	2
Kullanmıyorum	Öğrenciler için zorlayıcıdır	1
	Sınıf mevcutları oldukça kalabalık	3

Tablo 14'te araştırmaya katılan İngilizce öğretmenlerinin kullandıkları tamamlayıcı ölçme değerlendirme araç ve yöntemleri "Proje/Performans değerlendirme", "Öğrenci ürün

dosyası", "Kontrol listesi/Dereceli puanlama anahtarı", "Öz değerlendirme", "Akran değerlendirme", "Gözlem" ve "Tamamlayıcı ölçme ve değerlendirme araçlarını kullanmıyorum" temaları altında sunulmuştur. Öğretmenlerin proje ya da performans görevlerini tercih etmelerinin en önemli sebepleri "Öğrencinin daha aktif olmasını sağlar" ve "Öğrenme sürecini destekler" alt temalarıyla belirtilmiştir. Tamamlayıcı ölçme ve değerlendirme tekniklerinden en yüksek frekansa sahip olan öğrenci ürün dosyasının daha fazla tercih edilme nedenleri ise "Öğrencilerin gelişim sürecinin takibini sağlar" ve "Öğrencileri daha iyi tanıma fırsatı sunar" alt temalarıyla sunulmuştur. Kontrol listeleri ve dereceli puanlama anahtarlarının öğretmenlere en çok sağladığı kolaylık "Objektif puanlama imkanı sunar" alt temasıyla belirtilmiştir. Bir diğer tamamlayıcı ölçme değerlendirme yöntemlerinden olan Öz değerlendirmenin öğretmenler tarafından tercih edilmesinin en önemli sebebi "Öğrenciler güçlü ve zayıf yönlerini fark eder" alt temasıyla sunulmuştur. Diğer taraftan akran değerlendirmenin öğretmenler tarafından kullanılmasının en önemli sebebi ise "Öğrenci merkezlidir" alt temasıyla belirtilmiştir. Öğretmenlerin görüşlerine göre Gözlem yönteminin tercih edilmesinin nedenleri "Diğer tekniklere göre daha kullanışlıdır" ve "Doğru ve hızlı bilgilere ulaşmayı sağlar" alt temalarıdır. Öğretmenlerin büyük çoğunluğunun tamamlayıcı ölçme değerlendirme yöntem ve tekniklerini kullandığı görülmektedir. Ancak bu araç ve yöntemlerin kullanılmadığını ifade eden öğretmen görüşleri de mevcuttur. Bu durumun en önemli nedeni ise "Sınıf mevcutları oldukça kalabalık" alt temasıyla belirtilmiştir. Araştırmaya katılan İngilizce öğretmenlerinin tamamlayıcı ölçme değerlendirme araç ve tekniklerinden en çok hangilerini tercih ettikleri ve bu araçları kullanma sebeplerinin neler olduğuna dair görüşler şu şekildedir:

> *"Performans ve portfolyo ödevleri verip bunları değerlendiriyorum. Sadece eğlenceli ve yaratıcılıklarını ortaya çıkarabilecekleri ödevleri olsun diye. Ders kitabında her ünite sonunda proje ödevleri var, onları kullanıyorum. Öğrenciler için ünite değerlendirmesi oluyor." (Ö1)*
>
> *"En çok kullandığım performans, proje ve portfolyodur. Çünkü bunlar sürece odaklı olduğu için daha uzun vadede öğrencinin performansını daha iyi kullanmasını sağlıyor. Öğretmen tarafından yapılan geri dönütler de süreç içinde daha sağlıklı bir öğrenme ortamı geliştiriyor. Öğrenci daha çok keşfediyor ve daha fazla kendine güveniyor." (Ö6)*
>
> *"Mümkün olduğunca tamamlayıcı araç ve tekniklerini kullanıyorum. Çünkü öğrencileri en iyi tanımamıza yardımcı olan teknikler olduğunu*

düşünüyorum. Bu teknikleri kullanıp sonuçları aldıktan sonra, bir sonraki aşamada neler yapabileceğini, eksiklerin nerede olduğunu, planlamada hata yapıp yapmadığını görebiliyorum. Geçen yıl bir sınıfımda özellikler akran değerlendirme üzerinde durdum. Sonrasında akran yardımı planlaması yaparak öğrencilerden çok güzel geri dönütler aldım." (Ö12)

"Performans, gözlem, öz değerlendirme tekniklerini kullanıyorum. Öğrencilerin öğrenmelerinde oluşan gelişimleri görmeleri ve anlayabilmeleri açısından sıklıkla bu yöntemleri kullanıyorum. İyi bir geri dönüt sağlıyor." (Ö13).

"Performans, öz değerlendirme, dereceli puanlama anahtarı ve gözlem kullandığım araç ve teknikler arasındadır. Öğrencileri daha iyi tanıyabilmek ve onların eksik oldukları noktaları belirleyerek tam bir öğrenme sağlamak için kullanıyorum." (Ö18).

"Gözlem ve performans değerlendirmeyi daha çok kullanıyorum. Öğrenci sayısı fazla olduğu için bu tekniklerin uygulanmasını diğer tekniklere göre daha kullanışlı buluyorum. Öğrenci sayısı daha az olsaydı diğer teknikleri kullanırdım." (Ö19).

"Hâlihazırdaki sistemde öğretmen başına düşen öğrenci sayısı yüksek olduğundan ve yine bu sistem içerisindeki değerlendirme ölçütlerinin öğretmenleri sadece bilgi parçalarını ölçmeye zorladığından ve öğrenci tutum ve davranışlarının önemsenmediğinden değerlendirmeler salt klasik yöntemlerle yapılmaktadır." (Ö22).

"Tamamlayıcı ölçme araç ve tekniklerini kullanırken 2 temel amacım var; birincisi öğrencinin düzeyini görmek; ikincisi gelişiminin önündeki engelleri tespit ederek gerekli müdahalelerde bulunmak ve gelişimine katkı sunmak. Öğrencinin sahip olduğu potansiyeli ortaya çıkarmak diyebiliriz." (Ö26; ÖO1).

"Performans değerlendirmenin daha objektif olduğunu düşünüyorum. Aynı zamanda öğrenci merkezli olduğu için daha kıymetli buluyorum. Gözlem, bir öğretmenin her zaman yapabileceği en pratik yöntemdir. Ben de çoğunlukla kullanırım" (Ö28; ÖO3)

"Bunların çoğunu kullanmak zorundayız çünkü bu kadar öğrencinin zaten ne yaptığını aklımızda tutmamız imkansız. O yüzden performans değerlendirme, kontrol listeleri ve dereceli puanlama anahtarını mutlaka kullanmamız lazım. Böylelikle öğrencinin bir gerileme veya ilerleme durumunda olup olmadığını da daha iyi anlarız." (Ö35; ÖO10)

Araştırmaya katılan öğretmenlerin görüşleri incelendiğinde, tamamlayıcı ölçme değerlendirme araç ve tekniklerinin öğretmenler tarafından sıklıkla kullanıldığı

söylenebilir. Öğrencilerin dil becerilerini değerlendirme sürecinde kullanılan bu araçların çeşitli öğretim amaçlarını gerçekleştirmek için tercih edildiği ifade edilmiştir. Öğretmenlere göre tamamlayıcı araç ve tekniklerin en önemli özelliği öğrenciyi daha fazla tanıma imkanı sunmasıdır. Daha kapsayıcı bir ölçme değerlendirme süreci sunan bu tamamlayıcı araçların öğretmenlere öğrenme süreci hakkında dönüt sağlaması, öğrencileri çok yönlü tanıma fırsatı sunması ve öğrencilerin aktif olması gibi özellikleriyle tam öğrenmeyi de desteklediği dile getirilmiştir. Ayrıca öğretmenler bu araçları öğrenciler için de oldukça faydalı bulmaktadırlar. Öğrencilerin kendi gelişimlerini takip etmelerini, öğrenme sorumluluklarını üstlenmelerini, yeterliklerinin ve zayıf yönlerinin farkına varmalarını sağlayan motive edici araçlar olarak kabul etmektedirler. Tüm bu görüşlerin aksine tamamlayıcı ölçme değerlendirme araç ve yöntemlerini okulların fiziki şartlarından dolayı çok sınırlı bir şekilde kullanan ya da hiç kullanmayan görüşler de bildirilmiştir.

Araştırmanın nicel verilerine göre, Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketinin 9., 20., 30. ve 41. maddeleri öğretmenlerin tamamlayıcı ölçme ve değerlendirme araçlarını kullanma düzeyleriyle ilgilidir. Tablo 15'te araştırmaya katılan İngilizce öğretmenlerinin yabancı dil beceri alanlarına göre tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini kullanma ortalamaları, standart sapmaları ve düzeyleri gösterilmiştir.

Tablo 15. Öğretmenlerin yabancı dil beceri alanlarına göre tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini kullanma ortalamaları, standart sapmaları ve düzeyleri

Bölüm	Madde no	Maddeler	$\bar{X}$	Ss	Düzey
Dinleme	(9) 9	Tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) kullanırım	3,37	,97	Çoğu zaman
Konuşma	(10) 20	Tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) kullanırım	3,34	1,06	Çoğu zaman
Okuma	(9) 30	Tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) kullanırım	3,46	1,01	Çoğu zaman

Yazma	(10) 41	Tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) kullanırım	3,49	1,02	Çoğu zaman

Tablo 15'te araştırmaya katılan öğretmenlerin tüm yabancı dil beceri alanlarına göre tamamlayıcı ölçme ve değerlendirme araçlarını "Çoğu zaman" düzeyinde kullandıkları saptanmıştır. Öğretmenlerin en çok yazma becerilerini değerlendirmek için tamamlayıcı araçları kullandıkları görülürken, konuşma alanında ise bu tür araçların daha az tercih edildiği belirlenmiştir.

Araştırmanın tamamlayıcı ölçme ve değerlendirme araç ve tekniklerinin incelenmesine dair nitel ve nicel verileri incelendiğinde, katılımcıların hem görüşme formuna hem de ankete verdikleri yanıtların tutarlı olduğu söylenebilir. Nitekim, nitel araştırma verilerine göre katılımcıların çoğu bu tür araçları öğrenci başarısını ve yeteneklerini değerlendirmede kullandıklarını belirtmişlerdir (f=41). Diğer taraftan nicel araştırma verileri incelendiğinde, katılımcıların tüm yabancı dil beceri alanlarını ölçme ve değerlendirme sürecinde tamamlayıcı araçları "Çoğu zaman" düzeyinde kullandıkları saptanmıştır. Dolayısıyla araştırmanın iki aşamasında da ortaya çıkan bulguların birbiriyle örtüştüğü ifade edilebilir.

Araştırmada "Uygulama" ana temasındaki bir diğer soru İngilizce öğretiminde dilbilgisi, sözcük dağarcığı ve telaffuz gibi dil yapılarının nasıl ölçüldüğü ile ilgilidir. Çalışmaya katılan öğretmenlerin verdikleri yanıtlar doğrultusunda bu soruyu açıklamaya yönelik belirlenen temalar ve alt temalar Tablo 16'da gösterilmiştir.

Tablo 16. "İngilizce öğretiminde dilbilgisi, sözcük dağarcığı ve telaffuz gibi yapıların ölçülmesi ile ilgili görüşleriniz nelerdir?" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları

Tema ve Alt Temalar	Frekans
A. Dilbilgisi	
A.1.Genellikle dilbilgisi yapılarını ölçecek testler hazırlarım	14
A.2.Klasik ölçme araçlarıyla değerlendirme yaparım	4
A.3.Dilbilgisi yapılarını ölçmek oldukça gereklidir.	5
A.4. Eskisi kadar önemsenen bir alan değildir	8
A.5. İletişim becerilerinin önüne geçmemelidir.	3

B. Sözcük Dağarcığı	
B.1. Sözcük bilgisini ölçmek oldukça gereklidir	16
B.2. Genellikle sözcük bilgisi ağırlıklı ölçümler yaparım	15
B.3. Klasik ölçme araçlarını kullanırım	4
B.4. Sınav sisteminden dolayı önem kazanmıştır.	2
C. Telaffuz	
C.1. Telaffuz becerilerini ölçmek önemlidir	6
C.2.Telaffuz ile ilgili yapıların ölçülmesi gerekli görülmemektedir	2
C.3. Gözlem yoluyla ölçme yaparım	3
C.4. Sınav sisteminden dolayı göz ardı edilmektedir.	5
C.5. Herhangi bir ölçme yapmıyorum	25
D. Bütün beceriler aynı derecede önemlidir ve mutlaka ölçülmelidir	4

Tablo 16'da araştırmaya katılan öğretmenlerin görüşleri "Dilbilgisi", "Sözcük dağarcığı", "Telaffuz" ve "Bütün becerileri aynı derecede önemlidir ve mutlaka ölçülmelidir" temaları altında sunulmuştur. Dilbilgisi teması içerisinde en fazla öne çıkan düşünce "Genellikle dilbilgisi yapılarını ölçecek testler hazırlarım" alt teması olmuştur. Bu durumun aksine dikkat çeken "Eskisi kadar önemsenen bir alan değildir" alt teması ise ikinci sırada gelmektedir. Sözcük dağarcığı teması altındaki görüşler incelendiğinde ise öğretmenlerin bu alandaki değerlendirmeleri önemsedikleri "Sözcük bilgisini ölçmek oldukça gereklidir" alt temasıyla anlaşılmaktadır. Bu düşünceyi "Genellikle sözcük bilgisi ağırlıklı ölçümler yaparım" alt teması takip etmektedir. Telaffuz temasıyla ilişkili çeşitli görüşlerin olduğu ve bunların içerisinde en çok dile getirilen düşüncenin "Herhangi bir ölçme yapmıyorum" alt teması olduğu görülmektedir. Ayrıca tüm dil yapılarının en az yabancı dil becerileri (dinleme, konuşma, okuma ve yazma) kadar önemli olduğuna dikkat çeken ve mutlaka değerlendirilmesi gerektiğini savunan görüşler de mevcuttur. Bu görüşler "Bütün beceriler aynı derecede önemlidir ve mutlaka ölçülmelidir" teması altında birleştirilmiştir. Dilbilgisi, sözcük dağarcığı ve telaffuz gibi dil yapılarının ölçülmesi ile ilgili doğrudan alıntılar şu şekildedir:

> *"Genelde dilbilgisi ve sözcük dağarcığıyla ilgili yapıları ölçecek testler hazırlıyorum. Ders etkinlikleri yapılırken telaffuza dikkat ediyorum. Ancak bunu herhangi bir ölçme aracıyla ölçmüyorum." (Ö4)*

"Her kademeye göre kazanımlar değişiklik gösterebilir. Ancak, ortaokul düzeyindeki öğrencilerin kelimelerin yazımında diğer kazanımlara göre daha iyi durumda olmaları gerekir. Ünitelerde verilen dilbilgisi ve sözcük bilgisini sınavlar ve sözlü yoklamalarla mutlaka ölçerim." (Ö5).

"İngilizce öğretiminde sözcük dağarcığı benim için büyük önem taşımaktadır. Öğrenci ne kadar çok kelime bilirse kendini o dilde o kadar çok rahat ifade etmektedir. Ama son zamanlarda tecrübelerime dayanarak edindiğim izlenimde dilbilgisi ve telaffuz yapıları öğrencinin kendini ifade etmesinde çok büyük önem taşımamaktadır. Bir veya iki kelime ile öğrencinin ne demek istediğini çok rahat anlayabiliyoruz. Dilbilgisinin bu açıdan pek önemi kalmamaktadır. Örneğin liseye geçiş sınavlarında bile dilbilgisi değil de anlama ve kavrama becerileri daha çok ölçülmektedir." (Ö6).

"Dilbilgisi ve sözcük dağarcığı mevcut sistemde ölçülen şeylerdir. Telaffuz ise asla ölçülmüyor. Biz İngilizce öğretmenleri olarak sınav odaklı öğrenci yetiştirmek zorunda kalıyoruz. Bu durumda telaffuz önemini yitiriyor" (Ö7).

"Genelde dilbilgisi ve sözcük dağarcığı ağırlıklı ölçümler yapılıyor. Telaffuz ise ölçülemiyor." (Ö8).

"Dilbilgisi ve sözcük dağarcığını genellikle klasik ölçme araçları kullanarak ölçerim. Telaffuz ve bazı zamanlarda sözcük dağarcığını sözlü sınavlarla ölçüyorum" (Ö13).

"Ortaokulda görev yaptığım için sınav odaklı çalışmalar yürütüyorum. Dolayısıyla dilbilgisi son yıllarda yerini okuyup anlamaya ve sözcük öğrenmeye bırakmış durumda. Eskisi gibi dilbilgisi ağırlıklı etkinlikler ve ölçümler yapmıyorum." (Ö17).

"Ara sınavlar sözcük bilgisini kısmen ölçebilmekte. Çünkü günlük ezberletilen kelimeler kısa bir zaman sonra unutulmaktadır. Telaffuzla ilgili yeterince geri dönüt verilmemekle birlikte dilbilgisi yeterliliği yazılı sınav kağıtlarındaki başarıyla ölçülebilmektedir." (Ö22).

"Kelime bilgisi ve dilbilgisi için yazılı ve sözlü sınavlar yaparken telaffuz için herhangi bir ölçme aracı kullanmıyorum." (Ö23).

"Hiçbir becerinin diğerlerinden ayırt edilip daha az önemsenmesinin söz konusu olmaması gerekir. Mutlaka tüm becerilerin bir şekilde ölçülmesi ve üzerinde durulması lazım." (Ö24).

"İlk olarak telaffuz olmazsa olmazımızdır. Dil öğreniminde en önemli nokta budur. İlk yapmamız gereken öğrencilerimizin doğru telaffuz edebilmelerini sağlamaktır. Sözcük dağarcığı da oldukça önemlidir. Kelime hazinelerinin

güçlü olması o dile hakim olmalarını sağlamada önemli bir yoldur. Dilbilgisi en sevmediğim ancak öğrenilmesi gereken önemli bir konudur. Ama bana göre ilk 8 yıl çocuklara hissettirilmeden ve kurallar verilmeden bir yol çizilmelidir. Çünkü amacımız bu dili konuşturmak olmalıdır. Daha sonra çocuğun yaptığı dilbilgisi hataları onu rahatsız edecektir ve yaşları ilerledikçe kendileri farkına varacaklardır. Bu yüzden bu yapılar direkt ölçmek yerine gözlem ve etkinlikler yoluyla dolaylı ölçmeler yapılmalıdır." (Ö31; ÖO6).

"Hepsinin önemli olduğunu düşünüyorum. Son yıllarda dilbilgisine çok önem verilmiyor ama ben bir çocuğun İngilizce öğrenip başarılı olması için dilbilgisini de çözmesi gerektiğine inanıyorum. O yüzden konularda veya ünitelerde geçen bütün dilbilgisi yapılarını ve konularını anlatır, bunları yazılı olarak da ölçerim. Ama şu an öncelik öğrencilerin kelime dağarcığını geliştirmek. Bunu da sözlü ve yazılı ölçme araçlarıyla yapıyorum." (Ö35; ÖO10)

"Dilbilgisi ve sözcük dağarcığının değerlendirilmesinin dil öğrenimi için hayati önem taşıdığını düşünüyorum ve bu alanlarla ilgili daha fazla değerlendirme yapıyorum" (Ö41; İHO6).

Araştırmaya katılan öğretmenlerin görüşleri incelendiğinde, öğretmenlerin çoğunun genellikle sözcük ve dilbilgisi yapılarının değerlendirilmesine yönelik ifadelerinin olduğu görülmektedir. Hatta değişen ölçme ve değerlendirme yaklaşımının daha çok okuma ve anlama becerisini ön plana çıkardığı ve bu yüzden öğretmenlerin sözcük dağarcığının geliştirilmesi için daha fazla hassasiyet yaşadıkları söylenebilir. Dil öğreniminde dilbilgisi yapılarının mutlaka ölçülmesi gerektiğine dair fikirlerin de eskisi kadar dile getirilmediği söylenebilir. Ancak bu konuda değişen anlayışa rağmen ölçme ve değerlendirme uygulamalarında daha çok dilbilgisi öğelerinin test edildiği ifade edilmiştir. Dolayısıyla dil öğretim yaklaşımları açısından bir paradigma değişikliği yaşansa da uygulama alanında köklü bir anlayışın yerleşemediği ve geleneksel yöntemlerin yabancı dil öğretim felsefesinin temelini oluşturmayı sürdürdüğü söylenebilir. Telaffuz becerilerinin ölçülmesinin ise daha fazla ihmal edildiği anlaşılmaktadır. Nitekim öğretmenlerin çoğunun telaffuz becerilerinin ölçülmesine dair herhangi bir değerlendirme yapmadıkları kaydedilmiştir.

Bu tema altında Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketinin verileri sunulmuştur. Araştırmanın nicel verilerine göre anketin 5., 6., 17., 18., 19., 27., 28., 37. ve 39. maddeleri öğretmenlerin ölçme ve değerlendirme sürecinde dilbilgisi, sözcük

bilgisi ve telaffuzla ilgili davranışları ne sıklıkta gerçekleştirdiklerini göstermektedir. Tablo 17'de araştırmaya katılan İngilizce öğretmenlerinin bu alanların ölçülmesine dair sergiledikleri davranışların ortalamaları, standart sapmaları ve düzeyleri gösterilmiştir.

Tablo 17. Öğretmenlerin dilbilgisi, sözcük yapıları ve telaffuzla ilgili davranışları değerlendirme ortalamaları, standart sapmaları ve düzeyleri

Bölüm	Madde no	Maddeler	$\bar{X}$	Ss	Düzey
Dinleme	5 (5)	Öğrencilerin dilin ses, sözcük gibi öğelerine ne ölçüde hakim olduklarını değerlendiririm.	3,84	,87	Çoğu zaman
	6 (6)	Öğrencilerin dilbilgisi kuralları ve kalıplarını ne ölçüde anladıklarını değerlendiririm	3,92	,89	Çoğu zaman
Konuşma	7 (17)	Öğrencilerin gerekli sözcükleri kullanabilme becerilerini değerlendiririm	4,03	,88	Çoğu zaman
	8 (18)	Sözcüklerin doğru telaffuz edilmesine dikkat ederim	4,14	,93	Çoğu zaman
	9 (19)	Öğrencilerin dilbilgisi kuralları ve kalıplarını ne düzeyde kullanabildiklerini değerlendiririm	3,97	,92	Çoğu zaman
Okuma	6 (27)	Öğrencilerin kelimeleri ve söz kalıplarını ne ölçüde tanıdıklarını değerlendiririm	4,19	,71	Çoğu zaman
	7 (28)	Öğrencilerin dilbilgisi kuralları ve kalıplarını uygun bir şekilde kullanıp kullanmadıklarını değerlendiririm.	4,14	,75	Çoğu zaman
Yazma	6 (37)	Öğrencilerin yazma amacına uygun olan kelimeleri ne ölçüde kullandıklarını değerlendiririm	3,82	,99	Çoğu zaman
	8 (39)	Öğrencilerin dilbilgisi kuralları ve kalıplarını uygun bir şekilde kullanıp kullanmadıklarını değerlendiririm	3,91	,95	Çoğu zaman

Tablo 17 incelendiğinde, öğretmenlerin anketin tüm boyutlarındaki dilbilgisi, sözcük bilgisi ve telaffuzla ilgili ölçme değerlendirme davranışlarını "Çoğu zaman" düzeyinde gösterdikleri görülmektedir. Araştırmanın nitel ve nicel verileri incelendiğinde, öğretmenlerin dilbilgisi ve sözcük yapılarının ölçülmesine ilişkin görüşlerinin benzerlik gösterdiği saptanırken, telaffuz becerisinin ölçülmesiyle ilgili bulguların örtüşmediği söylenebilir. Nitel araştırmaya katılmaya öğretmenlerin büyük bir kısmının (f=32) telaffuz yapılarının ölçülmesiyle ilgili olumsuz görüş bildirdiği kaydedilirken, nicel araştırma örneklemindeki öğretmenlerin "Sözcüklerin doğru telaffuz edilmesine dikkat ederim" davranışını "Çoğu zaman" düzeyinde sergiledikleri ortaya çıkmıştır. Telaffuz yapılarının bu düzeyde ölçüldüğüyle ilgili ulaşılan bulgular araştırmanın nitel aşamasında ortaya çıkan

"Gözlem yoluyla ölçme yaparım" alt temasıyla açıklanabilir. Yani nicel araştırmaya katılan öğretmenlerin sınıf içi etkinlikler sırasında gözlem yoluyla bu becerileri değerlendirdikleri söylenebilir.

İngilizce öğretiminde öğrencilerin bilişsel kazanımları ne düzeyde gerçekleştirdiğinin ölçülmesi kadar önemli olan bir diğer boyut da duyuşsal becerilerin ne kadar önemsendiği ve ölçme değerlendirme sürecinin bir parçası olup olmadığıdır. Bu yüzden araştırmada "Uygulama" ana teması altındaki bir diğer soru duyuşsal becerilerin ölçülmesi ile ilgili olmuştur. Tablo 18'de araştırmaya katılan öğretmenlerin bu konudaki görüşlerini belirten temalar ve alt temalar sunulmuştur.

Tablo 18. "İngilizce öğretiminde duyuşsal becerilerin ölçülmesi ile ilgili görüşleriniz nelerdir?" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları

Tema ve Alt Temalar	Frekans
A. Duyuşsal becerilerin ölçülmesi gereklidir	
A.1. İlgi ve tutumları ortaya çıkarır	12
A.2. Bireysel farklılıkları ortaya koyar	1
A.3. Önyargıları ortaya çıkarır	5
A.4. Öğrenmeye isteksiz oluşun sebeplerini ortaya koyar	5
A.5. Öğrenme güçlüklerini belirlemede etkilidir	2
A.6. Başarı için gerekli bir ölçüttür	2
A.7. Dil becerilerinin öğrenimini kolaylaştırır	3
A.8. Öğretim yöntemiyle ilgili ipucu sağlar	2
A.9. Bilişsel öğrenmelerin hazırlayıcısıdır	1
A.10. Gözlem yoluyla ölçme yaparım.	16
A.11. İlgi envanterleri veya tutum ölçekleri kullanırım	2
B. Duyuşsal becerilerin ölçülmesi gerekli değildir	3
B.1. Değerlendirme yapmak zordur	1
B.2. Uzun zaman alır	2
C. Gereklidir ancak herhangi bir ölçme yapmıyorum.	3

Tablo 18'de araştırmaya katılan öğretmenlerin duyuşsal becerilerin ölçülmesiyle ilgili görüşleri doğrultusunda "Duyuşsal becerilerin ölçülmesi gereklidir", "Duyuşsal becerilerin ölçülmesi gerekli değildir" ve "Gereklidir ancak herhangi bir ölçme yapmıyorum" temaları belirlenmiştir. Öğretmenlerin çoğunun duyuşsal becerilerin

ölçülmesinin gerekliliğine dair görüş bildirdiği söylenebilir. Bu tema içerisinde pek çok farklı ifadenin yer aldığı ve bunların içerisinde en fazla vurgulan görüşün "Gözlem yoluyla ölçme yaparım" alt teması olduğu görülmektedir. Öğrencilerin duyuşsal becerilerinin ölçülmesinin sunduğu katkılara değinen öğretmenlerin en çok dile getirdikleri ifade "İlgi ve tutumları ortaya çıkarır" alt teması olmuştur. Duyuşsal becerilerin ölçülmesinin gerekli olmadığını savunan öğretmenler için ise bu durumun en önemli gerekçesi "Uzun zaman alır" alt teması olmuştur. Ayrıca duyuşsal becerilerin gerekli olduğu ancak herhangi bir ölçme yapılmadığına dair görüşler de kaydedilmiştir. Duyuşsal becerilerin ölçülmesi ile ilgili öğretmen görüşleri şu şekildedir:

"Herhangi bir ilgi envanteri uygulamıyorum ancak gözlem yoluyla, öğrencilerin derse katılımlarıyla bu konuda bilgi sahibi olabiliyorum." (Ö4)

"Öğrenci sevse de sevmese de bu dersi almak zorunda. Ayrıca ders saatleri kısıtlı olduğundan dolayı bu tür ölçümler için gerekli zamanı bulamıyorum. Çünkü uygulama ve değerlendirme süreçleri uzun zaman da alacaktır. O yüzden ölçülmesinin bir mantığı yok." (Ö7).

"Dersi sevdirmek ya da öğrencinin istekli olması tabi ki de başarılı olmanın ilk ölçütüdür. Ancak bu becerileri direkt olarak ölçmek için herhangi bir ölçme aracı kullanmıyorum." (Ö11)

"İlgi envanterleri, gözlüm formu kullanarak duyuşsal becerilerin harekete geçmesini sağlıyorum." (Ö13).

"Öğrencilerin derse ilgisinin çok fazla olmayışı, dersi önemsememeleri gibi olumsuzluklar duyuşsal becerilerin ölçülmesi gerektiğinin en büyük kanıtıdır. Gözlem yoluyla değerlendirmeler yaparım" (Ö14)

"Genellikle gözlem formunu yaygın olarak kullanıyorum. İngilizce dersini sevme veya sevmeme gibi durumları belirlemek için kişisel anket düzenliyorum." (Ö18)

"Öğretmeni sevmek, dersi sevmek, dersin öneminin farkında olma öğretimin de verimini arttırır ve karşılıklı olarak derse olan tutum değişir. Bu yüzden bu beceriler mutlaka ölçülmelidir." (Ö21).

"Öğrencilerin pek çoğu Türkçeyi dahi etkili bir şekilde kullanamamakta ve kendini ifade edememektedir. Haliyle bu yeni bir dil öğrenme konusunda hem motivasyonlarını hem de bu konudaki yeterliklerini olumsuz bir şekilde etkilemektedir. Ayrıca öğrendikleri dili de hâlihazırda kullanabilecekleri pek sosyal ortam olmadığı için öğrenme konusundaki istekleri de

kırılabilmektedir. Öğrencilere uygulanacak bir anketle bütün bu durumların kaynağı, öğrenme süreçlerini nasıl etkilediği, bu konuda hissettikleri ve duyguları açığa çıkarılabilir ve buna göre öğrenme ortamları düzenlenebilir. Ancak ders saatleri kısıtlı olduğu için gözlem yoluyla değerlendirme yapıyorum" (Ö22).

"Öğrenci isteksizse İngilizce dersi için ilgi duyduğu alanların öğrenilmesi gerekir. Bireysel veya grup çalışmasına uygun olup olmadığının gözlemlenmesi gerekir. Yani bireysel farklılıklar öğretmen tarından ortaya çıkarılmalı ki verimli bir öğrenme çevresi oluşabilsin." (Ö25).

"Öğrencilerin seviyesini ve derse olan ilgilerini dikkate alarak gözlem formu tutuyorum". (Ö29; ÖO4).

"Dil öğreniminde öğrencilerin duyuşsal becerileri önemlidir. İngilizce dersini ve öğretmenini sevmesi, derse istekli katılması ve önemsemesi başarılı olmasını sağlar." (Ö30; ÖO5).

"Duyuşsal becerilerin ölçülmesinin oldukça zor olduğuna inanıyorum. Zaten duyuşsal beceriler ders esnasında öğretmenler tarafından algılanabiliyor. Çok zaman da aldığı için bu tür değerlendirmeleri gerekli görmüyorum." (Ö42; İHO7).

Araştırmaya katılan öğretmenlerin görüşleri incelendiğinde, öğretmenlerin çoğunun İngilizce öğretiminde duyuşsal becerilerin ölçülmesi gerektiğine dair görüş bildirdikleri anlaşılmaktadır. Duyuşsal becerilerin ölçülmesi gerektiğine dair görüş bildiren pek çok öğretmen; bu süreç için herhangi bir ölçme aracı kullanmaktansa öğrencilerin derse olan ilgi ve tutumlarını gözlemleyerek bir yargıya ulaşabildiklerini ifade etmişlerdir. İngilizce dersinde başarılı olmanın ön koşulu olarak öğrencilerin bu dersi ve öğretmeni sevmesi ve önemsemesi gerektiği belirtilmiştir. Diğer taraftan duyuşsal becerilerin ölçülmesi için kullanılacak araçlarının seçimi, uygulanması ve değerlendirilmesi gibi aşamalardan dolayı uzun zaman gerektirdiği ve İngilizce ders saatlerinin kısıtlı olmasından dolayı bu tür çalışmalar yapılamadığı da kaydedilmiştir. Ayrıca duyuşsal becerilerin ölçülmesini son derece önemli gören ama herhangi bir değerlendirme yapmayan ifadeler de bulunmaktadır.

Araştırmada "Uygulama" ana temasında bilişsel ve duyuşsal süreçlerin nasıl ve ne düzeyde ölçüldüğü ile ilgili sorular yer almıştır. Bir sonraki süreç olan "Değerlendirme" ana temasında ise ölçme uygulamaları ile elde edilen sonuçların belirli ölçütlerle karşılaştırılarak bir sonuca ya da değer yargısına ulaşılması işi gerçekleşmektedir. Araştırmanın bu bölümünde öğretmenlere belirli bir amaca göre yapılan değerlendirme

türlerinden hangilerini kullandıklarına dair bir soru yer almaktadır. Bu doğrultuda Tablo 19'da İngilizce öğretmenlerinin değerlendirme türlerine yönelik görüşlerini belirten temalar ve alt temalar sunulmuştur.

Tablo 19. "İngilizce öğretiminde tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirmeyi hangi düzeyde ve niçin kullandığınızı açıklar mısınız?" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları

Tema	Alt Temalar	Frekans
Tanılayıcı Değerlendirme	Dönem başında yaparım	12
	Hazırbulunuşluğu ölçmek için	8
	Öğrenci seviyesini belirlemek için	10
	Planlama yapmak için	7
	Tamamlayıcı çalışmalar yapmak için	2
	Sınıf seviyesini belirlemek için	4
	Tanılayıcı değerlendirme yapmam	6
Biçimlendirici Değerlendirme	Her hafta yaparım	2
	Her üniteden sonra yaparım	6
	Ayda bir kez yaparım	1
	İhtiyaç halinde yaparım	15
	Bir üst seviyeye geçmek için	2
	Öğretim etkinliklerini düzenlemek için	3
	Öğrenme eksiklerini belirlemek için	10
	Öğrenme düzeyini belirlemek için	5
	Öğrencilerin gelişim süreçlerini izlemek için	1
	Öğrenciler için bir geribildirim aracı olduğu için	1
	Biçimlendirici değerlendirme yapmam	3
Düzey Belirleyici Değerlendirme	Her dönemin ortasında ve sonunda yapılır	45
	Uygulanması zorunludur	12
	Kazanımların gerçekleşme düzeyini belirmek için	3
	Başarı-başarısızlık durumlarını belirlemek için	15

Tablo 19'da araştırmaya katılan öğretmenlerin görüşleri doğrultusunda "Tanılayıcı değerlendirme", "Biçimlendirici değerlendirme" ve "Düzey belirleyici değerlendirme" temaları belirlenmiştir. Tanılayıcı değerlendirme temasında görüş bildiren öğretmenlerin en çok dile getirdikleri alt tema "Dönem başında yaparım" ifadesi olmuştur. Öğretmenlerin tanılayıcı değerlendirme yapmalarındaki en önemli amaçları "Öğrenci seviyesini belirlemek" ve "Hazırbulunuşluğu ölçmek" alt temaları olmuştur. Biçimlendirici değerlendirme teması incelendiğinde, öğretmenlerin en çok vurguladıkları ifade hangi

sıklıkla değerlendirme yaptıklarını yansıtan "İhtiyaç halinde yaparım" alt teması olmuştur. Bu görüşü takip eden diğer alt tema ise bu değerlendirme türünü neden tercih ettiklerini açıklayan "Öğrenci eksikliklerini belirlemek için" alt teması olmuştur. Düzey belirleyici değerlendirme teması içerisindeki görüşlere göre tüm öğretmenler "Her dönemin ortasında ve sonunda yapılır" alt teması üzerinde görüş birliğine varmışlardır. Bu değerlendirme türünün en çok kullanılma nedeni ise "Başarı-başarısızlık durumlarını belirlemek için" alt teması olmuştur. Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme türleri ile ilgili öğretmen görüşleri şu şekildedir:

> *"Sene başında sözlü olarak tanılayıcı değerlendirme yaparım. Yaz tatilinden dönüşte öğrencinin aklında kalan bilgileri yoklamak için. Her üniteden sonra da yazılı veya sözlü olarak biçimlendirici değerlendirme yapıyorum. Düzey belirleyici değerlendirmeyi zorunlu olarak dönemin ortasında ve sonunda yapmak zorundayız zaten." (Ö1)*
>
> *"Hazırbulunuşluğu ölçmek için tanılayıcı değerlendirmeyi kullanırım. Çocuğun bulunduğu seviyeyi belirleyip bir üst düzeye çıkarmak için de biçimlendirici değerlendirme yaparım. Düzey belirleyici değerlendirmeyi de karne notlarını vermek için yapıyorum" (Ö4).*
>
> *"Tanılayıcı değerlendirmeyi genellikle dönem başlarında çocukların hazırbulunuşluk düzeylerini ölçmek amacı ile yaparım. Böylece çocuğun hazır olmadığı durumlarda öncelikle tamamlayıcı çalışmalar yaparım. Biçimlendirici değerlendirmeyi öğrencilerin öğrenme eksikliklerini gidermek ve geri dönüt almak için yaparım. Ünite sonlarında genel bir test yaparım. Yine kelime sınavları en çok başvurduğum biçimlendirici değerlendirme yöntemimdir. Tanılayıcı ve biçimlendirici değerlendirmelerin katkısı ile dönem sonlarında düzey belirleyici sınavları ve değerlendirmeleri yaparım. Bunları kullanma amacım öğrenme sürecini daha sağlıklı ilerletebilmek ve öğrencinin öğrenme seviyesiyle ilgili bilgi sahibi olmaktır." (Ö6).*
>
> *"Tanılayıcı değerlendirme öğretmen için önemlidir. Sınıf düzeyini belirlemeyi ve dersi bu çerçevede işlemeyi sağlar. Biçimlendirici değerlendirme ise öğrenci gelişimini takip etmekte çok önemlidir. Öğrenci konuyu anlayabilmiş mi?, Doğru yöntemle mi öğretim yapılıyor? gibi sorulara cevap verir niteliktedir. Düzey belirleyici değerlendirme zorunludur. Öğrenci başarıları hakkında bir değerlendirme yapmamızı sağlar." (Ö7)*
>
> *"Tanılayıcı değerlendirmeyi öğrencilerin hazırbulunuşluğunu belirlemek için dönemin başında yaparım. Biçimlendirici değerlendirmede öğrenme*

eksikliklerini belirlemek için ilave testler, ünite testleri ve ara sınavlar yaparım. Düzey belirleyici değerlendirme dönem sonunda öğrenci başarısı hakkında karar vermek için yapılır." (Ö8).

"Tanılayıcı değerlendirme ilk derslerde öğrenciyi tanımak için uygun sorular sorarak uygularım. Biçimlendirici değerlendirme için ara sınavlar, sözlüler ve çoktan seçmeli testleri kullanırım. Düzey belirleyici değerlendirme zorunlu olarak her dönem iki kez uygulanır." (Ö9).

"Tanılayıcı değerlendirmeyi ilk derslerde öğrenciyi tanımak için, seviyesine uygun sorular sorarak uyguluyorum. Biçimlendirici değerlendirmeyi; yazılı yoklamalarla çok sık olmayan ara sınav veya testlerle ihtiyaç duyunca yapıyorum. Bu değerlendirme türü üzerine daha çok uygulama yapmak isterdim ancak müfredatın yoğunluğu nedeniyle süre sıkıntısı yaşıyorum. Düzey belirleyici değerlendirme zaten karne notları girilerek belirlenmiş olur" (Ö11).

"Tanılayıcı değerlendirmeyi, konulara başlamadan önce çocukların ne düzeyde olduğunu anlamak ve ona göre bir planlama yapmak için kullanıyorum. Biçimlendirici değerlendirmeyi, biten bir ünitenin ardından konunun ne düzeyde anlaşıldığını, nerelerde eksiklikler olduğunu, bir sonraki üniteye geçmeden önce son bir tekrar yapıp yapmamam gerektiğini anlamak için kullanıyorum. Düzey belirleyici değerlendirmeyi de öğrencilerin son geldikleri seviyeyi belirlemek için kullanırım. İstediğim sonucu alamadıysam, bir sonraki aşama için daha dikkatli planlama yaparım." (Ö12).

"Tanılayıcı değerlendirmeler, sözlü yoklama şeklinde sınıfın düzeyini belirlemek için yapılmaktadır. Biçimlendirici değerlendirme; öğrencilerin eksiklerini belirlemek amacıyla ayda bir ara sınav şeklinde yapılır. Düzey belirleyici değerlendirme ise, her dönemin sonunda zorunlu olarak yapılan yazılı yoklamalardır" (Ö14).

"Her dönem iki kez olmak üzere sadece düzey belirleyici değerlendirme yapıyorum. Bu değerlendirmeyle öğrenci başarısı hakkında karar verilir." (Ö17).

"Gerçek anlamda hiçbirini net, etkili bir şekilde kullanmıyorum. Sadece zorunlu olduğu için düzey belirleyici değerlendirmeler yapıyorum." (Ö20).

"Biçimlendirici değerlendirmeyi her hafta kullanıyorum. Ara sınavlarla ve testlerle mutlaka ölçüm yapıyorum. Düzey belirleyici değerlendirmeyi öğretim programının öngördüğü şekilde yapmak zorundayız zaten." (Ö21).

"Tanılayıcı değerlendirme benim için çok önemli. Çünkü öğrencilerin bildiklerinden yola çıkarak bilmediklerini ulaşmalarını sağlamak daha

kolay ve etkili oluyor. Tabi ki biçimlendirici değerlendirme de çok önemli. Çünkü bu değerlendirmeyle öğrencilerin eksiklerini daha hızlı bir şekilde görebiliyorum. Ayrıca rakamsal sonuçlar öğrencilerin kendilerini toparlamalarına da çok yardımcı oluyor. Dönem ortasında veya sonunda da genel başarılarını ölçer ve değerlendiririm." (Ö31; ÖO6).

"Öğrencilerin sene başında ne seviyede olduklarını incelemek için tanılayıcı ölçme; dönem içerisinde özellikle kelime bilgisini ölçmek için biçimlendirici değerlendirme ve yazılı sınavlarla da düzey belirleyici değerlendirmeyi kullanıyorum" (Ö38; İHO3).

Araştırmaya katılan öğretmenlerin görüşleri incelendiğinde, öğretmenlerin özellikle düzey belirleyici değerlendirme türünü kullandıkları görülmektedir. Bu değerlendirme türünün bir tercih olarak kullanılmasından ziyade aslında zorunlu olduğu için uygulandığı anlaşılmaktadır. Daha çok öğrencilerin başarıları hakkında bir karara varmak için kullanıldığı ifade edilmiştir. Tanılayıcı değerlendirmelerin yapılmasının en önemli sebebinin ise öğrenci eksiklerini belirlemek olduğu belirtilmiştir. Öğretmenlerin görüşlerine göre bu tür değerlendirmeler yapılırken öğretmenlerin en çok tercih ettiği yöntemlerden biri sözlü yoklamalardır. Dolayısıyla bu değerlendirme türüyle ilgili çok geniş çaplı çalışmaların yapıldığı veya farklı araç ve tekniklerin kullanıldığı söylenemez. Biçimlendirici değerlendirmelerin ise tanılayıcı değerlendirmelerden daha çok tercih edildiği söylenebilir. Öğretmenlerin bu değerlendirme türünü tercih etmelerinin en önemli sebebi ise öğrenme eksikliklerini ortaya çıkarmak ve kelime bilgisini pekiştirmektir. Ayrıca sadece zorunlu olduğu için düzey belirleyici değerlendirmeler yapan ve diğer değerlendirme türlerini kullanmadıklarını ifade eden öğretmen görüşleri de mevcuttur.

Bu tema altında Yabancı Dil Becerilerini Ölçme ve Değerlendirme Anketinin verileri sunulmuştur. Anketin 4., 13., 24. ve 34. maddeleri öğretmenlerin yabancı dil beceri alanlarının değerlendirilmesi sürecinde tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerini ne düzeyde kullandıklarıyla ilgilidir. Tablo 20'de araştırmaya katılan İngilizce öğretmenlerinin bu değerlendirme tekniklerini kullanma ortalamaları, standart sapmaları ve düzeyleri gösterilmiştir.

Tablo 20. Öğretmenlerin Yabancı Dil Beceri Alanlarına Göre Tanılayıcı, Biçimlendirici ve Düzey Belirleyici Değerlendirme Tekniklerini Kullanma Ortalamaları, Standart Sapmaları ve Düzeyleri

Bölüm	Madde no	Maddeler	$\bar{X}$	Ss	Düzey
Dinleme	4 (4)	Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım	2,83	1,20	Ara sıra
Konuşma	3 (13)	Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım	3,17	1,16	Ara sıra
Okuma	3 (24)	Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım	4,08	,85	Çoğu zaman
Yazma	3 (34)	Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım	3,62	1,11	Çoğu zaman

Tablo 20 incelendiğinde, araştırmaya katılan öğretmenlerin dinleme ve konuşma alanlarında "Ara sıra" düzeyinde değerlendirme tekniklerini kullandıkları görülürken, okuma ve yazma alanlarında ise "Çoğu zaman" düzeyinde kullandıkları belirlenmiştir. Öğretmenlerin en fazla "Okuma" boyutunda tanılayıcı, biçimlendirici ya da düzey belirleyici değerlendirmeler yaptığı görülürken ($\bar{X}$=4,08), en az "Dinleme" becerilerinin değerlendirilmesiyle ilgili uygulamaları bulunmaktadır ($\bar{X}$=2,83). Araştırmanın yabancı dil beceri alanlarıyla ilgili nitel ve nicel bulguları göz önünde bulundurulduğunda, sonuçların tutarlı olduğu ve dinleme ve konuşma becerilerinin bir ölçme aracıyla test edilip değerlendirilmesi işleminin daha sınırlı bir şekilde yürütüldüğü söylenebilir.

İngilizce öğretiminde etkili bir ölçme değerlendirme için planlama, uygulama ve değerlendirme süreçleri tamamlandıktan sonra öğrencilere geribildirim sunmak oldukça önemlidir. Araştırmada "Geribildirim" ana teması olarak belirlenen bu aşamada, öğretmenlere ölçme sonuçları hakkında geribildirimde bulunma konusundaki görüşleri sorulmuştur. Tablo 21'de araştırmaya katılan öğretmenlerin bu konudaki görüşleri doğrultusunda belirlenen temalar ve alt temalar sunulmuştur.

Tablo 21. "Ölçme sonuçları hakkında öğrencilere geribildirimde bulunma konusundaki görüşleriniz nelerdir?" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları

Temalar	Frekans
Öğrencilere her sınav sonrası sınav evrakını dağıtırım	30
Öğrencilerin hatalarını görmelerini sağlar	27
Öğrenciler hata yapmaktan korkmazlar	3
Öğrenciler soru sormaktan çekinmezler	1
Öğrenmeler daha kalıcı hale gelir	4
Anında dönüt verme imkanı sağlar	6
Puanlama hataları giderilir	3
Ölçme sonuçlarının güvenirliği artar	2
Dikkati arttırır	1
Motivasyonu arttırır	2
Ara sıra sınav evrakını dağıtırım	10
Geribildirim sağlamak benim için zordur.	2

Tablo 21'de araştırmaya katılan öğretmenlerin görüşleri çeşitli temalar altında sunulmuştur. Araştırmaya katılan öğretmenlerin çoğunun ölçme sonuçlarıyla ilgili öğrencilere geribildirimde bulunma konusunu oldukça önemsedikleri söylenebilir. Ölçme sonuçlarıyla ilgili geribildirimde bulunmanın sağladığı faydalara dikkat çeken öğretmenlerin en sık vurguladıkları alt tema "Öğrencilerin hatalarını görmelerini sağlar" olmuştur. Bu alt temayı takip eden ve geribildirimin önemini anlatan diğer görüş ise "Anında dönüt verme imkanı sağlar" alt temasıdır. Ölçme sonuçları hakkında geribildirimde bulunma sıklığıyla ilgili görüşler incelendiğinde ise, öğretmenlerin çoğunun sınav evrakını dağıtmaya özen gösterdikleri "Öğrencilere her sınav sonrası sınav evrakını dağıtırım" alt temasıyla anlaşılmaktadır. Bu alt temayı takip eden diğer görüş ise "Ara sıra sınav evrakını dağıtırım" alt temasıdır. Geribildirimde bulunma konusundaki bazı öğretmen görüşleri şu şekildedir:

> *"Sınav kağıtlarını öğrencilere hatalarını görmeleri için geri veririm. Ara sınavlar yaparım. Öğrencilerin birbirlerinin kağıtlarını benim gözetimimde değerlendirmesini sağlarım." (Ö4).*

> *"Yaptığım her sınav sonrası mutlaka sınav kağıtlarını okuyup öğrencilere hatalarını görmeleri için dağıtırım. Böylece diğer sınavlarda daha dikkatli olmaları gerektiğini anlamış olurlar." (Ö5)*

"Öğrencilerime anında geri dönüt vermek benim için çok önemlidir. Çünkü geri dönüt çocuğun hatasını görüp daha doğruya ulaşmasını ve bilgiyi içselleştirmesini sağlayacaktır. Hatasını görmeyen öğrenci zihninde karışıklık yaşayacaktır ve bu sağlıklı öğrenme sürecini kötü etkileyecektir. Sınav kağıdı dışında ders anlatırken de soru cevaplarda anında yapılan geri dönütler işimizi kolaylaştıracaktır" (Ö7).

"Ölçme sonuçları sonunda öğrencilere kesinlikle geri bildirimde bulunması gerektiğini düşünüyorum. Bu nedenle her ölçme değerlendirme sonunda öğrencilerle tek tek kağıtlarını incelerim. Sadece sınavlarda sonra değil, haftalık yaptığım test yoklamalarında da öğrencilerle tek tek kağıtlarını inceler yaptıkları hataları görmelerini sağlarım. Bunun birçok faydasını görüyorum. Çocuklarda hem özgüven gelişimi oluyor, hem de soru sormaktan, hata yapmaktan korkmadan devam ediliyor öğrenme sürecine." (Ö12)

"Cevap kağıtlarının değerlendirildikten sonra kesinlikle öğrencilere dağıtılması taraftarıyım. Hatalarını görüp, doğruyu öğrenmeleri açısından çok yararlı olduğunu düşünüyorum. 19 yıldır hiç aksatmadan yaptığım bir uygulama." (Ö17).

"Sözlü sınavlarda yaklaşık 2-3 dakikalık kısa bir sınav yaptıktan sonra öğrenciye artı ve eksilerini puanlayarak detaylı anlatabiliyorum. Ancak, yazılı sınavların sınıf yönetimini zorlaştırması hem de sınıf mevcutlarının kalabalık oluşundan dolayı pek dağıtma taraftarı değilim." (Ö20).

"Sonuçların bir geribildirim aracı olarak kullanılması çok önemlidir. Bana göre geri bildirimler ne kadar hızlı yapılırsa o kadar yararlı olacaktır. Nasıl ki üzerinden zaman geçen bir olayın üzerimizde etkisi azsa bu durum ölçme için de geçerlidir. Her dönüt anında veya en kısa süre içerisinde gerçekleşmelidir. Ayrıca sınavdan sonra öğrencilere kağıtlarını dağıtır ve yanlışları hakkında tartışır, bir sonuca varırız. Ders içi ölçme değerlendirmem ise anında olur çünkü bunu anında yapmazsam çocuğa yanlış deneyimler kazandırmış olurum ve ona daha sonra da doğruyu öğretmem çok zor olur ya da imkansızlaşır." (Ö31; ÖO6).

"Her ünite sonunda yapılan sınavları tekrar dağıtıp, her soruyu sınıfta tekrar çözüyoruz. Yine ana sınavların soruları da sınıfta çözülür. Böylece nerede hata yaptıklarını görebilirler. Yine aynı şekilde öğretmenin yanlış hesaplama veya yanlış değerlendirmeleri ortaya çıkar. Böylece ölçme sonucu daha güvenilir olur." (Ö34; ÖO9)

"Ara sıra yaptığım testleri ve ara sınavları dağıtırım. Çözemediklerini mutlaka açıklar ve çözdürürüm. Sözlü olarak da onları sınıfta konuşturmaya çalışarak özgüvenlerini sağlamaya çalışırım. Böylece kendilerini ifade

edebilmek için İngilizce kelime bulma araştırmasına girerler" (Ö35; ÖO10).

"Öğrencilerin yanlışlarını görmelerini sağlamak ve puanlamada yanlışlık olup olmadığını kontrol etmek amacıyla sınav kağıdını muhakkak her sınav sonrası öğrencilere dağıtırım." (Ö42; İHO7)

Araştırmaya katılan öğretmenlerin görüşlerine göre hem öğretim süreci esnasında hem de bu sürecin başında, ortasında veya sonunda gerçekleştirilen sınavlarda öğrencilere geribildirim sağlamak oldukça önemlidir. Öğretmenlerin çoğunun bu görüşü desteklediği ve yapılan yazılı sınavlar sonrası öğrencilere sınav evrakını dağıtmaya özen gösterdikleri görülmektedir. Öğretmenler anında dönüt sağlamanın önemine oldukça fazla dikkat çekmiş ve sağladığı pek çok yararı sıralamışlardır. Öğrencilerin hatalarını görmelerini sağlaması, onları hata yapma korkusundan alıkoyarak özgüvenlerini ve motivasyonlarını arttırması en çok dile getirilen faydalar arasındadır. Tüm bu görüşlerin aksine, sınıfların kalabalık olmasından dolayı yaşanan kargaşa ve zaman kaybı gibi endişelerin geribildirim sürecini kimi öğretmenler için güç kıldığı anlaşılmaktadır.

Araştırmada öğretmenler ölçme ve değerlendirme uygulamalarıyla ilgili genel bir tablo sunarak yaşadıkları tecrübeleri aktarmışlardır. Bu bölümde ise ölçme ve değerlendirme öğesinin daha etkili bir şekilde kullanılması için neler yapılabileceği araştırılmıştır. "Geribildirim" ana teması altında sunulan bu aşamada, öğretmenlere yabancı dilde nitelikli bir ölçme ve değerlendirme sürecinin gerçekleşmesi için önerilerinin neler olduğu sorulmuştur. Tablo 22'de araştırmaya katılan öğretmenlerin bu konudaki görüşleri doğrultusunda belirlenen temalar ve alt temalar sunulmuştur.

Tablo 22. "İngilizce derslerinde nitelikli bir ölçme değerlendirme süreci için önerileriniz nelerdir? Açıklayınız" sorusuna verilen cevaplardan elde edilen görüşler ve frekansları

Tema ve Alt Temalar	Frekans
A. Ölçme değerlendirme araç ve yöntemleri ile ilgili öneriler	
A.1. Ölçme aracı günlük yaşamdan kopuk olmamalı	2
A.2. Tüm yabancı dil beceri alanları ölçülmeli	8
A.3. Farklı ölçme ve değerlendirme teknikleri kullanılmalı	5
A.4. Tamamlayıcı ölçme araçları kullanılmalı	4
A.5. Üst düzey düşünme becerileri ölçülmeli	1
A.6. Öğrencilerin hazırbulunuşluğu ölçülmeli	3
A.7. Biçimlendirici değerlendirmeye ağırlık verilmeli	2

A.8. Duyuşsal beceriler ölçülmeli	1
A.9. Öğretim etkinliklerine uygun ölçüm yapılmalı	2
B. Öğrenme ortamı ile ilgili öneriler	
B.1. Dil sınıfları oluşturulmalı	3
B.2. Sınıf mevcutları azaltılmalı	13
B.3. Yeterli fiziki donanım sağlanmalı	7
C. Öğretim programıyla ilgili öneriler	
C.1. Kazanımlar sadeleştirilmeli	3
C.2. İletişime dayalı test örnekleri sunulmalı	3
C.3. Ders saati arttırılmalı	6
C.4. Ünite sayısı azaltılmalı	4
D. Öğretim materyalleri ile ilgili öneriler	
D.1. Ders kitapları iletişimi desteklemeli	3
D.2. İlgi çekici görsel-işitsel materyaller sunulmalı	2
D.3. Dil becerileri için uygun içerik sunulmalı	3
D.4. Akıllı tahtalar yaygınlaştırılmalı	5
D.5. Ders kitapları akıllı tahta uygulamalarıyla zenginleştirilmeli	3
D.6. Günlük hayata dair video ve animasyonlar sunulmalı	2

Tablo 22 incelendiğinde, araştırmaya katılan İngilizce öğretmenlerinin nitelikli bir ölçme ve değerlendirme süreci için önerileri "Ölçme ve değerlendirme araç ve yöntemleri ile ilgili öneriler", "Öğrenme ortamı ile ilgili öneriler", "Öğretim programıyla ilgili öneriler" ve "Öğretim materyalleri ile ilgili öneriler" temaları altında sunulmuştur. Öğretmenlerin ölçme ve değerlendirme yöntemleri ile ilgili sundukları öneriler arasında en fazla "Tüm yabancı dil beceri alanları ölçülmeli" ve "Farklı ölçme ve değerlendirme teknikleri kullanılmalı" alt temalarının vurgulandığı görülmektedir. Öğrenme ortamının iyileştirilmesi konusunda ise "Sınıf mevcutları azaltılmalı" alt temasının en sık dile getirildiği görülmektedir. Öğretim programıyla ilgili öneriler incelendiğinde, "Ders saatleri artırılmalı" ve "Ünite sayısı azaltılmalı" alt temaları ön plana çıkmaktadır. Öğretmenlerin öğretim materyalleri ile ilgili görüşleri arasında ise "Akıllı tahtalar yaygınlaştırılmalı" önerisi en sık vurgulanan alt tema olmuştur. Yabancı dil öğretiminde nitelikli bir ölçme ve değerlendirme sürecinin gerçekleşmesi için öğretmenlerin sunduğu bazı öneriler şu şekildedir:

> *"Nitelikli bir ölçme için öğrenci sayısının az olması, ders saatinin daha fazla olması, ünite sayısının daha az ve bol etkinlikle olması gerektiğini*

düşünüyorum. Öğrencilerin konuşma ve dinleme becerilerinin de mutlaka ölçülmesi gerekmektedir." (Ö9)

"Tüm beceri alanlarının ayrı ayrı mutlaka ölçümü zorunlu olmalı. Çoktan seçmeli sorular azaltılmalı. Değişik soru türlerine yer verilmeli." (Ö10)

"Nitelikli bir ölçme için öğrenci sayısının az olması, ünite sayısının azaltılmış olması, daha çok ölçme yapabilmek için de ders saatinin attırılmış olması gerekir." (Ö11)

"Öğrencilerin sınav kaygısı yaşamayacak şekilde, dört temel beceriyi de ölçecek nitelikte bir uygulamaya geçilmesi gerekiyor." (Ö14)

"Tamamen iletişime dayalı sözel ölçme öneriyorum. Öğrencilerin ne kadar konuşup ne ölçüde dinlediklerini anlayabildikleri ölçülürse o kadar sağlıklı bir ölçüm yapılmış olur. Bu şekilde öğrenme de kendiliğinden gerçekleşir. Yazılılar, ara sınavlar ve testler sadece bilgi ölçer. Bilgiyi nasıl kullandığının ölçülmesi daha uygun olurdu." (Ö16)

"Dil laboratuvarının olması gerektiğini düşünüyorum. Dönem başında seviyelini görmek için mutlaka sınavlar yapılmalı ve ders planı ile kaynak seçimi buna göre ayarlanmalıdır." (Ö17)

"Nitelikli bir ölçme öncelikle öğrenciyi iyi tanıma ile başlar. Dolayısıyla öncelikle öğrencilerin hazırbulunuşlukları, yetenekleri ve ilgileri belirlenmelidir. Bunun için gözlem ve derecelendirme ölçekleri tüm yıl kullanılmalıdır. Ayrıca dört temel beceriyi ölçecek şekilde hazırlanacak ölçme araçları hem İngilizce öğrenimini kolaylaştıracak hem de sağlıklı bir değerlendirme yapmamıza imkan sağlayacaktır." (Ö18)

"Nitelikli bir ölçme; dinleme, konuşma ve yazma gibi önemli dil unsurlarını içermelidir. Bir tanesinin eksik ölçülmesi diğerlerini de olumsuz etkileyecektir. Bu yüzden bu becerileri kapsayan ölçme araçlarının kullanılması gerekir." (Ö19)

"Konuşma, dinleme ve yazma becerilerine yönelik öğrenme sürecinde yeterli çalışmalar yapılıp, bu becerilerin ölçme araçlarına mutlaka eklenmesi sağlanmalıdır." (Ö28; ÖO3)

"Tamamlayıcı değerlendirme ile öğretime başlanılması gerekir. Öğrencinin Hazırbulunuşluğu belirlenmeden öğrenme ve öğretme gerçekleşmez. Biçimlendirici değerlendirmeler de önemsenmelidir." (Ö33; ÖO8)

"Okumuzun fiziki alt yapısı ve teknolojik donanımı verimli bir öğretim ve değerlendirme sürecinin yaşanmasını kısıtlıyor. Öncelikle akıllı tahtaların kullanımın yaygınlaştırılması gerekiyor. Ayrıca ders kitapları da yeteri kadar ilgi çekici değil. Ancak devlet okullarında ek kaynak kitapların

kullanımına da izin verilmiyor. Bu yüzden ders kitaplarının öğrenci ilgisini ve motivasyonunu arttıracak şekilde düzenlenmesi ve akıllı tahtalar ile uyumlu hale getirilmesi sağlanabilir." (Ö40; İHO5).

Araştırmaya katılan öğretmenlerin görüşleri incelendiğinde, öğretmenlerin genel olarak ölçme değerlendirme yöntem ve teknikleriyle ilgili öneriler sunduğu görülmektedir (f=28). Öğretmenlerin dikkat çektikleri en önemli görüş, tüm dil beceri alanlarının ayrı ayrı ölçülmesi gerektiğidir. Özellikle konuşma, dinleme ve yazma becerilerinin ayrı ayrı değerlendirilmesi gerektiği vurgulanmıştır. Çoktan seçmeli testler yerine her bir beceri alanı için farklı ölçme ve değerlendirme araçlarının kullanılması önerilmiştir. Öğretmenlerin sıklıkla dile getirdikleri bir diğer önemli görüş de okulların fiziki ve teknolojik yapısındaki yetersizlikler nedeniyle yaşadıkları güçlükler ve bu konudaki önerileridir. Akıllı tahta kullanımının yaygınlaştırılması gerektiği belirtilmiş ve dil laboratuvarlarının oluşturulabileceği ifade edilmiştir. Ayrıca nitelikli bir ölçme ve değerlendirme için ünite yoğunluğunun azaltılması, ders saatlerinin arttırılması ve ders kitaplarının teknolojiyle uyumlu hale getirilerek öğrenci motivasyonunu arttıracak şekilde düzenlenmesi gerektiği sunulan öneriler arasındadır.

5. TARTIŞMA

Araştırmada "Hazırlık" ana teması altında sunulan ilk nitel veriler, öğretmenlerin bir İngilizce testi hazırlarken nasıl bir planlama yaptıkları ve hangi ölçütlere dikkat ettikleri ile ilgilidir. Araştırmaya katılan tüm İngilizce öğretmenlerinin bir ölçme aracı hazırlarken planlama yapmaya özen gösterdikleri kaydedilmiştir. Dolayısıyla katılımcıların planlamayı ölçme sürecinden önce önemli bir adım olarak gördükleri söylenebilir. Bachman ve Palmer'a (1996) göre, dil testlerini geliştirme sürecinde dikkatli bir planlama yapmanın üç gerekçesi vardır. Birincisi ve de en önemlisi, testin beklenilen hedefleri gerçekleştirmesi için etkili olabilecek en iyi araçları sunmayı garantilemesidir. İkincisi iyi bir planlamanın nelerin niçin yapıldığıyla ilgili hesap verilebilirliği artırmasıdır. Yani testin dikkatli bir şekilde hazırlandığına dair kanıt sunmasıdır. Üçüncü gerekçe ise, bir işe değer verip tamamladıktan sonra yarattığı ödüllendirilmiş hissiyle, yaşadığımız deneyimle ilgili memnuniyetimizi arttırır. Ayrıca ölçme değerlendirme uygulamalarından önce yapılacak iyi bir planlama ile öğrenci başarıları hakkında ulaşılacak değer yargılarının hatasız ve etkili olması sağlanabilir (Gözütok, 2006).

Araştırma bulgularına göre planlama süreciyle ilgili ortaya çıkan alt temalar arasında, öğretmenlerin göz önünde bulundurdukları en önemli ölçüt hazırlanan ölçme aracının sınıf düzeyine uygun olmasıdır. *"Sınıf seviyesi ve öğrencilerin hazırbulunuşluk durumlarına göre planlama yapıyorum (Ö6)"* ve "*Bir sınıfta tüm öğrenciler aynı seviyelerde olmadığı için sınıftaki tüm öğrencileri gözeterek bir planlama yapıyorum (Ö12)"* şeklinde dile getirilen görüşler doğrultusunda, öğretmenlerin sınıf düzeylerine uygun olan kazanımlara odaklanarak planlama yaptıkları söylenebilir. Arı'nın (2014) altıncı sınıf İngilizce öğretim programının değerlendirilmesiyle ilgili ulaştığı kazanımların ulaşılabilirliğinin sınıf düzeyine göre farklılaştığı ve ulaşılabilirliğinin mümkün olmadığı yönündeki sonuçlar bu araştırmanın bulgularıyla örtüşmektedir. Berkant, Özaslan ve Doğan (2019) yabancı dil ağırlıklı eğitimle ilgili öğretmen görüşlerini inceledikleri çalışmada, öğretmenlerin pedagojik açıdan karşılaştıkları en önemli sorunun yabancı dil ağırlıklı eğitim uygulamalarının öğrenci seviye ve hazırbulunuşluklarına uygun olmamasından dolayı kazanımlarla uyumlu sınavların yapılamaması olduğunu kaydetmişlerdir. Öztürk (2003) öğrencilerin giriş davranışlarının programın ölçme ve değerlendirme öğesini büyük ölçüde etkilediğini belirtmiş ve öğretmenlerin sınıf seviyesini göz önünde bulundurarak ulaşılabilir kazanımlara odaklanmak zorunda kaldıklarını ifade

etmiştir. Ortaya çıkan bu sonuçların araştırmanın *"Ünite kazanımlarına uygun olan davranışları belirlerim"* ve *"Kazanılması zorunlu olan davranışları seçerim"* alt temalarıyla örtüştüğü söylenebilir.

Araştırmada planlama sürecinde sınıf düzeyine uygun olan davranışların seçilmesinden sonra öğretmenler tarafından en sık dile getirilen bir diğer görüş soruların güçlük düzeyini belirlemektir. Öğretmenler genel olarak soru sayısına ve testin kolay ve zor soruları kapsayacak şekilde dengeli bir şekilde dağılım göstermesine dikkat ettiklerini ifade etmişlerdir. Ayrıca sınıf düzeyini gözeterek sınavları orta düzeyde hazırladıklarını belirtmişlerdir. Bu görüşleri destekleyen Gronlund ve Linn'e (1990) göre, öğretmen yapımı testlerin kabul edilebilir bir güvenirlik düzeyi sağlaması için yaklaşık %50 gibi bir zorluk seviyesinde hazırlanması tavsiye edilir. Testlerin güçlük düzeyinin ayarlanmasıyla ilgili yapılan çalışmalar, toplam test güçlüğü seviyesinin öğrencilerin sınava hazırlanma sürecini ve başarılarını etkileyebileceğini göstermektedir (Marso, 1969). Marso'ya (1969) göre zorluk derecesi makul bir şekilde belirlenmiş testler, öğrencilerin çalışma performansını ve sınav başarılarını artıracaktır.

Araştırmada öğretmenlerin planlama sürecinde vurguladıkları bir diğer alt temalardan biri testin amacına uygun bir test türü belirlemektir. Bu görüşü destekleyen Ivanova'ya (2011) göre, test hazırlayanlar ölçülecek özellikleri tüm boyutlarıyla detaylı bir şekilde düşündükten sonra amaçlarına en uygun şekilde hizmet edecek test türünü seçmelidirler. Böylece testin geçerliğini de yüksek düzeyde sağlamış olacaklardır. Ayrıca öğretmenlerin görüşlerine göre, test formatının önceden öğrencilere bildirilmesi, puan dağılımının ve zamanın makul bir düzeyde ayarlanması planlama sürecinde karar verilmesi gereken durumlardır. Saigh'e (1984) göre, eğer test uygulamasından önce bu durumlar titizlikle yürütülmezse, öğrenciler testin adil olmadığı gibi bir algıya kapılabilir veya sınav kaygısı yaşayabilirler.

Araştırmada planlama aşamasından sonra öğretmenlere ölçme aracında bulunması gereken geçerlik, güvenirlik ve kullanışlık özelliklerini nasıl sağladıklarına yönelik bir soru sorulmuştur. Öğretmenlerin çoğu bu özelliklerin sağlanmış olmasına dikkat ettiklerini belirtirken, sınıfların kalabalıklığı ve sınıf düzeyindeki farklılıklardan dolayı bu özelliklere uygun ölçme aracı hazırlayamadıklarını belirten görüşler de bildirilmiştir. Öğretmenlerin hazırladıkları ölçme aracının geçerlik şartını sağlayabilmesi için kazanımlara ve testin amacına uygun olan davranışları seçtikleri ve kapsam geçerliğinin sağlanmasına dikkat

ettikleri belirtilmiştir. Öğretmenlerin testin güvenirliğini sağlamak için göz önünde bulundurdukları en önemli ölçütler ise sınav süresinin yeterli olması, orta güçlükte bir sınav hazırlama ve soruların açık ve anlaşılır olmasını sağlamaktır. Ölçme aracında bulunması gereken bir diğer özellik olan kullanışlılığı sağlamak için test aracının hazırlama ve puanlamasının kolay olması öğretmenlerin en sık başvurdukları yöntemlerdir.

Çalışmaya katılan öğretmenlerin ölçme aracının geçerlik, güvenirlik ve kullanışlılık özelliklerinin sağlanması için başvurduğu yöntemlerin alanyazında da sıkça yer bulan ölçütler olduğu söylenebilir (Bachman, 1996; Bachman & Palmer; 1996; Brown, 1996; Hughes, 2003; Brown & Abeywickrama; 2010). Weir'e (2005) göre doğru bir ölçme testte olması gereken geçerlik, güvenirlik ve kullanışlık gibi genel ilkelerin ne düzeyde sağlandığıyla doğrudan ilişkilidir. Micheels ve Karnes'e (1950) göre ise, iyi bir ölçme aracı ölçülmek istenen davranışları doğru ve objektif bir biçimde ölçmeli, uygulanması ve puanlaması kolay olmalı, bilen ile bilmeyeni ayırt edebilmelidir.

Ölçme aracının ölçülmek istenilen özellikleri ne derece doğru ölçütüyle ilgili olan geçerlik, test puanlarının öğrenci başarısını başka bir değişkenle karıştırılmadan doğru bir biçimde yansıtmasıdır. (Hughes, 2003). Henning'e (1992) göre test geçerliğine karar vermenin iki yolu vardır. Birincisi verilerin toplanmasını ve geçerlik katsayısını hesaplamak için çeşitli istatistiksel formüllerin kullanımını gerektiren deneysel bir metottur. İkincisi ise denetlemeye ve genel bilgilere dayanan deneysel olmayan yöntemdir. Sireci'ye (1998) göre, birinci yöntem olan geçerlik çalışmalarının sadece istatistiksel yöntemlerle hesaplanması oldukça kısıtlayıcıdır. Diğer taraftan ikinci yöntem ise deneysel metot kadar bilgi ve beceri gerektirmediğinden öğretmenlerin kolaylıkla kavrayabileceği bir yöntemdir (Henning, 1988). Bu yönteme göre öğretmenlerin öncelikle öğrenci seviyelerini göz önünde bulundurarak test etmek istedikleri yeteneği ya da yapıyı doğru bir şekilde tanımlayıp işe başlamaları gerekir (Rapi & Miconi, 2014). Örneğin; eğer öğretmenler ilkokul seviyesindeki öğrencilerin dilbilgisi yeterliğini ölçmek istiyorlarsa, bu seviyede öğretilen dilbilgisi yapılarını test ettikleri konusunda açık olmaları gerekmektedir (Hughes, 2003). Dolayısıyla araştırmaya katılan öğretmenlerin kendi ölçme araçlarını hazırlarken dikkat ettiklerini belirttikleri *"Farklı sınıf düzeylerine uygun maddeler hazırlanması"* ve *"Derste işlenenlerle testin örtüşmesi"* alt temalarının geçerliği sağlamada kullanılabilecek doğru yöntemlerden olduğu söylenebilir. Bu görüşü destekleyen Altmışdört (2010), yabancı dilde ölçme değerlendirme araçlarının geçerliğinin

sağlanması için derste işlenenlerle testteki görevlerin ya da metinlerin benzerlik göstermesinin gerekliliğine dikkat çekmiştir.

Geçerliği sağlamanın bir diğer yolu testlerde ölçülmek istenen konular ile bu konuların hangi bilişsel düzeyde olduğunu gösteren kapsam geçerliğine (Kutlu ve Altıntaş, 2021) bakmaktır. Büyüköztürk'e (2012) göre kapsam geçerliği, testteki maddelerin tanımlanmış davranışlar kümesini ne düzeyde örneklediğini ifade eder. Özçelik (2010), testin kapsam geçerliğinin sağlamak için hedeflenen kazanım ve bilişsel düzeylerinin yer aldığı bir belirtke tablosu oluşturmanın oldukça önemli olduğuna dikkat çekmiştir. Belirtke tabloları planlama, öğretim ve değerlendirme süreçleri arasındaki bağlantıyı açık bir şekilde yansıttığından öğretmenlerin test hazırlamadan önce bu tabloları işlevsel olarak kullanmaları önerilmiştir (Fives & DiDonato-Barnes, 2013).

Araştırmanın bulgularına göre öğretmenler hazırladıkları ölçme araçlarının geçerliğini sağlamak için kapsam geçerliğine dikkat ederek soru sayısını belirlediklerini ifade etmişlerdir. Ancak bu ölçüte göre testin geçerliğini sağladıklarını ifade eden öğretmen sayısının oldukça az olduğu görülmektedir (f=5). Newell (2002), öğretmen yapımı testleri sadece sınırlı bir konu alanını ölçtüğü ve hatırlama düzeyindeki sorulara yoğun bir şekilde odaklandığı konusunda eleştirmiştir. Bu bağlamda kapsam geçerliğinin göz ardı edilmesi kazanılması gerekli olan davranışların kapsam dışında kalmasına sebep olabilir. Çağlar ve Kılıç'ın (2019) merkezi sınavlar ile öğretmen yapımı sınavları bazı değişkenler açısından inceledikleri çalışmada benzer bulgulara ulaşılmış ve katılımcıların hazırladıkları sınavlarda kapsam geçerliğinin %30'un altına düşmediği, genel ortalamanın ise %72 olduğu kaydedilmiştir. Araştırmanın sonuçları ölçülmesi beklenilen kazanımların yaklaşık %30'unun ölçülmediği ve kapsama alınmadığını göstermiş ve bu oranın küçümsenmeyecek bir miktar olduğu ifade edilmiştir. Çepni vd. (2009) bu konunun öğretmenler açısından oldukça önemli olduğuna dikkat çekmiş ve kritik davranışların kapsama dahil edildiğini garantilemek için kapsam geçerliğine mutlaka bakılması gerektiğini ifade etmişlerdir.

Belirli bir özelliği test etmek amacıyla yapılan ölçümlerin benzer şartlar altında tekrar edildiğinde tutarlı sonuçlar sunması olarak adlandırılan güvenirlik (Furwana, 2019), test puanlarının doğruluğunu ve kararlılığını açık bir şekilde gösteren bir niteliktir (Bachman & Palmer, 1996). Rapi ve Miconi'ye (2014) göre eğitim ve psikolojide yapılan ölçümler bir uzunluğu, genişliği ya da uzaklığı ölçmekten farklıdır ve bütünüyle duyarlı

sonuçlar beklemek oldukça güçtür. Yabancı dil becerilerinin ölçülmesinde daha da zor hale gelen bu süreç, ölçme araçlarında hem aynı davranışı ölçecek nitelikte çok sayıda soru kullanılmasını hem de boşluk tamamlama, eşleştirme gibi farklı türde soru tiplerinin kullanılmasını gerektirmektedir (Thorndike & Thorndike-Christ, 2010). Henning (1988) test güvenirliğini nesnel bir şekilde sağlamanın istatistiksel yollarını sıralasa da öğretmenlere kolay bir şekilde uygulayabilecekleri yöntemleri kullanmalarını önermektedir. Güvenirliği sağlayan bu yöntemler; test yönergelerinin açık olması, test süresinin yeterli bir şekilde ayarlanması, madde sayılarının makul bir uzunlukta ve güçlükte olması ile ilgilidir (Henning, 1988). Bu doğrultuda araştırmada katılımcıların testin güvenirliği sağlamak için ifade ettikleri *"Sınav süresinin yeterli olması"*, *"Testin orta güçlükte olması"*, *"Soruların açık ve anlaşılır olması"* ve "*Soru sayısını arttırma"* alt temalarının hazırladıkları ölçme aracının güvenirliğini arttırmada etkili olduğu söylenebilir. Diğer taraftan soru sayısının arttırılmasının yüksek güvenirlikle doğrudan ilişkili olduğu yönündeki anlayışı eleştiren Kutlu ve Altıntaş (2021), madde sayısının sadece güvenirliği sağlamak için arttırılmasının, öğretmenleri çoğunlukla bilgi, hatırlama ve uygulama düzeyindeki bilişsel hedeflere yönelttiğini ifade etmiştir. Dolayısıyla soru sayısının dil becerilerini kapsamlı olarak ölçecek nitelikte düzenlenmesi oldukça önemlidir

Ölçme aracında bulunması gereken niteliklerden bir diğeri de kullanışlılıktır. Brown ve Abeywickrama'a (2010) göre, bir testin kullanışlı olması demek zaman, uygulama, puanlama ve maliyet açısından test uygulayıcılarına kolaylık sağlaması demektir. Bu tanım doğrultusunda, araştırmaya katılan öğretmenlerin belirttikleri testin puanlama kolaylığı sağlaması, ekonomik olması ve hazırlanmasının kolaylığı gibi görüşlerin kullanışlılığı sağlama konusunda yeterli olduğu söylenebilir. Bir testin kullanışlı olması için puanlamanın kolay olması gerektiği bu tema altında en sık dile getirilen görüş olmuştur (f=9). Araştırma bulguları arasında öğretmenlerin bir ölçme aracının özellikle kullanışlı olmasına dikkat ettiklerini gösteren "*Sınıf mevcutları kalabalık olduğu için puanlaması kolay olabilecek sorular hazırlarım (Ö34;ÖO9)"* şeklinde ifadeler kaydedilmiştir.

Araştırmada "Hazırlık" ana teması altında öğretmenlere yöneltilen bir diğer soru hazırladıkları ölçme aracının İngilizce öğretim programına uygun olması için nelere dikkat ettikleridir. Araştırma bulgularına göre öğretmenlerin çoğu hazırladıkları ölçme araçlarının İngilizce öğretim programıyla uyumlu olmasına kısmen dikkat ettiklerini ifade etmişlerdir.

Çeşitli bilişsel, duyuşsal, fiziksel ve teknolojik kısıtlamaların öğretmenleri ölçme aracı hazırlama konusunda sınırlandırdığını belirten görüşler bildirilmiştir. Öğretim programının tüm boyutlarıyla tamamen uyumlu olacak şekilde ölçme aracı geliştirdiklerini belirten herhangi bir görüş olmamıştır.

Araştırma bulgularına göre öğretmenlerin öğretim programına kısmen bağlı kalmalarının en önemli sebebi öğrenci seviyelerinde yaşanan farklılıklardır. Öğrencilerin programa giriş davranışlarındaki yetersizliklerin, öğretmenleri öğretim programında sunulan standart kazanımları ölçecek araçlar hazırlama konusunda zorladığı belirtilmiştir. Dilekli (2018) ortaokul İngilizce hazırlık sınıfı programını öğretmen görüşlerine göre incelediği çalışmada benzer sonuçlara ulaşmış ve öğretmenlerin büyük bir kısmının programın benimsediği ölçme değerlendirme anlayışından uzak faaliyetlere yöneldiklerini ifade etmiştir. Dinçer (2018) 7. sınıf İngilizce öğretim programını değerlendirdiği çalışmada, öğrencilerin hazırbulunuşluk düzeylerinin öğretim programında yer alan kazanımları gerçekleştirecek düzeyde olmadığı sonucuna ulaşmıştır. Bir başka çalışmada Gök-Çatal'ın (2015) ulaştığı öğrencilerin yetenek ve kapasitelerinin üzerindeki kazanımların öğretim programında yer aldığı ve bu durumun dil öğretiminde hedeflenen başarıyı kazandıramadığı şeklindeki sonuçlar bu araştırmanın bulgularıyla örtüşmektedir. Yavuz'a (2005) göre öğretim programının içerik olarak yoğunluğuna rağmen esnek olmaması öğretmenleri kısıtlamaktadır. Bu durumun öğretmenleri belirli kazanımların gerçekleşme düzeyini ölçmeye yönlendirdiği söylenebilir.

Araştırma bulgularına göre, öğretmenlerin öğretim programına kısmen uygun olan araçlar hazırlamak zorunda kalmalarının diğer sebepleri, okulların fiziki koşullarının yetersizliği ve ders saatlerinin az olmasıdır. Öğretmenlerin eldeki kaynak ve öğrenci hazırbulunuşluklarına uygun olacak şekilde ölçülecek davranışları belirlediği ve bu doğrultuda ölçme araçları hazırladıkları belirtilmiştir. Alanyazında ders saatleri ve fiziki donanım gibi yetersizliklerin İngilizce öğretiminin program hedeflerine uygun şekilde yürütülmesini engellediğine dair çeşitli araştırmalar mevcuttur (Yılmaz, 2005; Aküzel, 2006; Gelbal ve Kelecioğlu, 2007; Doğan, 2009; Başaran, Özdemir ve Can, 2020; Çelebi ve Narinalp, 2020). Büyükduman'ın (2005) çalışmasında İngilizce öğretmenlerinin yarısı her ünite için ayrılan zamanın hem ünitenin kazanımları ile hem de güçlük dereceleri ile tutarlı olmadığını dile getirmişlerdir. Ersen-Yanık (2007) sınıfların kalabalık olmasının öğretmenlerin sınavları ve ödevleri sağlıklı bir şekilde kontrol etmesini ve öğrencilere

geribildirim sunmasını engellediğini belirtmiştir. Bu araştırmalardaki öğretmen görüşlerinin aksine, Gök-Çatal, Şahin ve Çelik (2018) İngilizce öğrenimine öğrencilerin bakış açısıyla farklı bir perspektif sunmuş ve öğretmenlerin ifade ettikleriyle benzer sonuçlara ulaşmışlardır. Onların çalışmasında öğrencilerin yaşadıkları program kaynaklı sorunlar arasında en fazla vurgulan görüşlerin ders saatlerinin az olması ve konuların çok ağır olmasına dair ortaya çıkan sonuçlar bu araştırmanın bulgularıyla örtüşmektedir. Dolayısıyla araştırma sonuçları göz önünde bulundurulduğunda, programının öğrenci kapasitesini aşan bir içerik sunduğuna dair görüşlerin hem öğretmen hem de öğrenciler nezdinde en fazla dile getirilen sorunlardan olduğu söylenebilir.

Araştırmaya katılan öğretmenlerin ölçme değerlendirme faaliyetleri ile ilgili yaptıkları hazırlık aşamasından sonra "Uygulama" boyutunda öğretmenlerin yabancı dil beceri alanlarında (dinleme, konuşma, okuma ve yazma) ölçme ve değerlendirme faaliyetlerine yönelik görüşleri incelenmiştir. Araştırmanın nitel aşamasında ortaya çıkan verilere göre, öğretmenlerin dinleme ve konuşma becerilerinin değerlendirilmesine yönelik olumsuz görüşleri olumlu görüşlerden daha fazladır. Bu beceri alanlarının değerlendirilmesine yönelik faaliyetlerin okulların fiziki ve teknolojik yapısındaki yetersizlikler ve ders saatlerinin az olması gibi öne çıkan nedenlerden dolayı ihmal edildiği ve ölçülemediği öğretmenlerin vurguladıkları görüşlerdendir. Bu alanlarda yapılan ölçümlerin daha çok ders esnasındaki uygulamalarla sınırlı kaldığı ifade edilmiştir.

Araştırmanın ikinci aşamasında yürütülen nicel çalışmanın verilerine göre ise, öğretmenlerin dinleme ve konuşma becerilerini sınıf içi etkinlikler sırasında değerlendirdiklerine dair görüşlerin ortalamasının *"Çoğu zaman"* düzeyinde olduğu görülmektedir. Öğretmenlerin dinleme ve konuşma becerilerinin ölçme ve değerlendirme araçlarıyla test edilip edilmediğini sorgulayan anketin *"İngilizce dersi öğretim programındaki kazanımlara uygun bir ölçme aracı hazırlarım"* ve *"Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım"* maddelerini ise "*Ara sıra"* düzeyinde gerçekleştirdikleri kaydedilmiştir. Dolayısıyla nitel ve nicel veriler göz önünde bulundurulduğuna sonuçların tutarlı olduğu ve dinleme ve konuşma becerilerinin değerlendirilmesine yönelik faaliyetlerin öğretim programında öngörülen şekilde çoğu zaman gerçekleşmediği ve eksik bırakıldığı söylenebilir.

Araştırmanın nitel aşamasında ortaya çıkan bir diğer bulgu, öğretmenlerin okuma becerilerinin değerlendirilmesine yönelik görüşlerinin dinleme ve konuşma alanlarında

kaydedilen olumsuzlukların aksine çoğunlukla olumlu olduğudur. Benzer şekilde nicel araştırma örneklemindeki katılımcıların ankete verdikleri yanıtlara göre, okuma becerilerinin değerlendirilmesine yönelik sergiledikleri davranışların "Her zaman" ya da "Çoğu zaman" düzeyinde olduğu görülmektedir. Her iki araştırma sonuçları göz önünde bulundurulduğunda okuma becerilerinin değerlendirilmesinin öğretmenler için diğer alanlar kadar zorlayıcı olmadığı anlaşılmaktadır. Yazma becerilerinin değerlendirilmesiyle ilgili nitel veriler incelendiğinde, pek çok farklı görüşün olduğu ve bunların içerisinde daha fazla vurgulanan ifadenin *"Herhangi bir sıkıntı yaşamıyorum"* alt teması olduğu anlaşılmaktadır. Ayrıca öğretmenlerin yazılı yoklamalarla, tamamlayıcı araçlarla ya da ders kitabındaki uygulamalarla yazma becerilerini değerlendirdiklerine dair görüşler saptanmıştır. Yazma becerilerinin ne düzeyde değerlendirildiğine ilişkin nicel veriler incelendiğinde ise, öğretmenlerin bu alanın değerlendirilmesine yönelik sergiledikleri davranışların çoğunlukla *"Çoğu zaman"* düzeyinde olduğu belirlenmiştir. Araştırmanın nitel ve nicel bulguları birlikte analiz edildiğinde, yazma alanında dinleme ve konuşma becerilerinden daha fazla ölçme ve değerlendirme faaliyetlerinin gerçekleştirildiği söylenebilir.

Etkili bir iletişimin en önemli öğelerinden olan yabancı dilde dinleme ve konuşma becerilerinin değerlendirilmesiyle ilgili yaşanılan zorluklar alanyazında sıkça vurgulanan konulardan olmuştur (Demirtaş ve Erdem, 2015). Arı'nın (2013) 6. sınıf İngilizce dersi öğretim programına yönelik öğretmen görüşlerini incelediği çalışmasında, katılımcıların dinleme ve konuşma alanlarında öğrencilerin isteksiz oluşları ve materyal eksikliklerinden dolayı bu alanlarda yeterli değerlendirmeler yapamadıklarına dair ulaşılan sonuçlar bu araştırmanın bulgularıyla örtüşmektedir. Bir başka çalışmada Çelebi ve Narinalp (2020), ortaokullarda İngilizce öğretiminde karşılaşılan en önemli sorunlardan birinin öğretmenlerin dilin dört temel beceri alanlarıyla ilgili değerlendirme uygulamalarına yeterli bir düzeyde odaklanamamaları olduğunu belirtmiştir. Bu durumun ise sınav sisteminin öğretimi mekanikleştirmesi ve okulların fiziki altyapısının öğretim programındaki kazanımları ölçecek imkanı sunmaması olarak gösterilmiştir. Sınıf mevcutlarının kalabalık olmasının ölçme değerlendirme sürecini olumsuz yönde etkilediğine dair benzer sonuçlara ulaşan Ersen-Yanık (2007), bu durumun her öğrenciye yeteri kadar zaman ayrılması ve sınavların programın öngördüğü niteliklerde gerçekleştirilmesini kısıtladığını belirtmiştir. Sonuç olarak yabancı dilde öğretim sürecini etkileyen bu olumsuzlukların iletişimin en önemli unsurlarından olan dinleme ve konuşma

becerilerini istenilen düzeyde değerlendirmeyi güç kıldığı kaydedilmiştir. Ortaya çıkan bu sonuçlar ile araştırmaya katılan öğretmenlerin ifade ettiği sorunların benzer ve tutarlı olduğu söylenebilir.

Ortaokul düzeyi için İngilizce dersi öğretim programı incelendiğinde, ölçme ve değerlendirme bağlamında uygulanacak yazılı veya sözlü sınavlar gibi performansa dayalı araçlar ya da projeler gibi alternatif araçlarla öğrenci başarısı hakkında objektif kayıtlar sunmak amaçlanmıştır (MEB, 2018). Ayrıca uygulanacak sınavların dilin dört temel beceri alanını kapsaması ve dilbilgisi, sözcük bilgisi gibi mikro düzeydeki yapıların ise örtük bir şekilde ölçülmesi gerektiği ifade edilmiştir (MEB, 2018). Yaman'a (2018) göre, İngilizce derslerinde tüm beceri alanlarıyla ilgili etkinlikler yapıp dinleme ve konuşma becerilerini sınavlarda ölçmemek, öğrencilere bu becerilerin diğer beceriler kadar önemli ya da gerekli olmadığına dair yanlış izlenimler kazandırabilir. Bu bağlamda öğrencilere yabancı dil başarısının kendini o dilde iyi bir şekilde ifade etmek gibi sözel bir performansa bağlı olduğu ve öğretmenlerin de bu performansı test etmesi gerektiğine dair olumlu bir mesaj verilirse, öğrenciler konuşma becerilerini geliştirmek için daha fazla çaba göstereceklerdir (Matin, 2012). Bu düşünceyi destekleyen bulgulara ulaşan Duran (2011), çalışmasında konuşma sınavlarının öğrencilerin ne düzeyde iletişim kurabildiğiyle ilgili sunduğu geribildirimlerin hem öğrenciler hem de öğretmenler için oldukça kritik öneme sahip olduğunu belirtmiştir. Araştırmada öğrencilerin konuşma becerilerinin bir ölçme aracıyla test edilip edilmediğini sorgulayan *"İngilizce dersi öğretim programındaki kazanımlara uygun bir ölçme aracı hazırlarım"* maddesine verilen yanıtların ortalaması incelendiğinde ($\bar{X}$=3,19), öğretmenlerin bu davranışı *"Ara sıra"* düzeyinde sergiledikleri görülmüştür. Diğer taraftan konuşma becerisinin bir ölçme aracıyla test edilmesine ilişkin nitel veriler analiz edildiğinde, bu beceri alanını değerlendirmeye yönelik araçlar kullandığını belirten katılımcı sayısının kullanmayanlardan daha az olduğu kaydedilmiştir (f=10). Dolayısıyla öğrencilerin sözel performanslarını değerlendirmeye yönelik yürütülen faaliyetlerin yeterli düzeyde olmadığı savunulabilir.

Dinleme ve konuşma becerilerinin neden yeterli düzeyde ölçülemediğiyle ilgili araştırma verileri incelendiğinde, mevcut sınav sisteminin iletişim becerilerini göz ardı ettiği, sınıf mevcutlarının kalabalıklığı ve ders saatlerinin kısıtlı olması gibi etkenlerin ön plana çıktığı görülmektedir. Günday (2007) yaptığı araştırmada benzer sonuçlara ulaşmış ve özellikle sınıf mevcutlarının kalabalık olmasının konuşma ve dinleme gibi birebir

iletişimi gerektiren alanlarda öğretmenleri zorladığı ve bu becerileri geliştirme ve değerlendirme süreçlerini olumsuz bir şekilde etkilediğini ifade etmiştir. Aküzel'in (2006) yabancı dil öğretiminde yaşanan başarısızlıkların nedenleri üzerine yürüttüğü çalışmada; okullarda yeterli fiziki ve teknolojik donanımın olmadığı, ders saatlerinin yetersizliği ve sınıf mevcutlarının yabancı dil öğrenimine uygun olarak düzenlenmediği yönünde ulaşılan sonuçların bu araştırmanın bulgularıyla tutarlı olduğu söylenebilir. Araştırmada öğrencilerin konuşma kaygısı yaşadıkları ve çekimser davrandıklarına dair görüşler de belirtilmiştir. Paker ve Höl'e (2012) göre öğrenciler uzun yıllar yabancı dil derslerine katılsalar da çoğu öğrenci için konuşma sınavları en zor sınav olarak algılanmaktadır. Bu yüzden öğrencilerin kendilerini konuşma sınavlarına daha hazır hissetmeleri için öğretmenlere ders sırasında daha fazla çift ya da grup çalışmaları ve tartışma etkinlikleri yaptırmaları önerilmiştir.

Araştırmaya katılan öğretmenlerin görüşleri incelendiğinde, tüm yabancı dil becerilerinin öğretim programının öngördüğü gibi eşit bir şekilde önemsenmesi ve test edilmesi gerektiğine dair ifadeleri göze çarpmaktadır. Ancak uygulamada özellikle iletişim becerilerini değerlendirme faaliyetlerinin eksik bırakıldığı veya ihmal ediği söylenebilir. Öğretmenlerin bir ölçme aracıyla değerlendirme yapmaktan ziyade konuşma ve dinleme alanlarında çoğu zaman sınıf içi etkinlikler sırasında öğrenci yeteneklerine ilişkin değerlendirme yaptıkları araştırmanın nicel bulguları arasındadır. Diğer taraftan nitel araştırma verileri incelendiğinde, öğretmenlerin gözlem yoluyla öğrenci başarılarına dair değerlendirmeler yaptıklarına dair *"...Dinleme becerisi imkansızlıktan dolayı ölçülemiyor. Konuşma becerisi de sınıflar kalabalık olduğu için. Ama bunlar da ders içinde gözlem yoluyla ölçülebilir"* şeklinde ifade edilen görüşler kaydedilmiştir. Cooper'a (1984) göre bu şekilde dolaylı değerlendirmeler yapılıyorsa, mutlaka sistematik bir yaklaşım sergilenmeli ve hem öğretmenlere hem de öğrencilere etkili geribildirimler sunmak amaçlanmalıdır. Chaudron (1988) gözlem tekniğin dil öğretiminde objektif bir şekilde kullanılması için hangi ölçütlerle değerlendirme yapılacağını gösteren değerlendirme ölçeklerinin önemine dikkat çekmiştir. Ona göre bu şekilde yapılacak ölçümlerle geçerli ve güvenir sonuçlar elde edilebilir.

Araştırmada yabancı dilde okuma becerilerinin diğer alanlardan daha sık değerlendirildiği yönünde bulgulara ulaşılmıştır. Öğretmenlerin okuma alanının değerlendirilmesine yönelik gösterdikleri davranışlarların "Çoğu zaman" ve "Her zaman"

düzeyinde olduğu saptanmıştır. Ortaya çıkan bu sonuçlar çeşitli araştırma sonuçlarıyla tutarlılık göstermektedir. Büyükduman (2005), dil becerileriyle ilgili öğretmen görüşlerini araştırdığı çalışmada, okuma alanı ile ilgili kazanımların diğer alanlara göre daha ulaşılabilir olduğu sonucuna ulaşmıştır. Yıldırım (2012) dil becerilerinin öğretimi ile değerlendirme süreçleri arasındaki tutarlılığı araştırmış ve hem öğretmenlerin çoğunlukla okuma becerilerinin geliştirilmesine yönelik etkinlikler yaptığına hem de okuma alanının ölçülmesinin diğer alanlardan daha fazla önemsendiğine dair benzer bulgular elde etmiştir. Öğrencilerin okuma alanına yönelik becerilerini değerlendirmenin bu denli önemsenmesinin temel nedeni araştırmanın nitel aşamasında ortaya çıkan *"Sınav sisteminden dolayı önemli görülen bir alandır"* alt temasıyla açıklanabilir. Bu bulguları destekleyen Başaran, Özdemir ve Can (2020), yaptıkları çalışmada öğretmenlerin öğrencileri merkezi sınavlara hazırlamak için okuma etkinliklerine daha fazla yöneldiğini ve dinleme ve konuşma gibi iletişim becerilerinin değerlendirilmesine ise yeteri kadar zaman ayırmadıklarını kaydetmişlerdir. Dolayısıyla sınav sisteminin gerektirdiği okuduğunu anlama becerilerinin bu alanın öncelikli olarak değerlendirilmesi gerektiğine dair bir anlayış kazandırdığı söylenebilir.

Araştırmanın önemli bulgularından biri de öğretmenlerin anketin tüm bölümleri içerisinde sadece okuma alanında "Nadiren" düzeyinde güçlük yaşadıklarını belirtmeleridir. Ayrıca nitel araştırmada okuma becerilerinin değerlendirilmesinde herhangi bir olumsuz yargıya rastlanılmaması bu bulguyu doğrular niteliktedir. Öztürk ve Aydın'ın (2019) öğretmenlerin okuma becerilerini değerlendirmede daha bilgili olduklarına dair ulaştıkları sonuçlar araştırma sonuçlarıyla örtüşmektedir. Nitekim çalışmaya katılan öğretmenlerin büyük bir kısmının okuma alanında herhangi bir güçlük yaşamadıklarını belirtmeleri, bu alanda kendilerini daha yeterli hissetmeleri ile açıklanabilir. Kim (2019) yaptığı çalışmada İngilizcenin temel ve orta seviyesinde öğrenim gören öğrenci ve öğretmelerin okuma alanındaki değerlendirme uygulamalarına karşı daha olumlu tutum sergilediklerini saptamıştır. Öğretmenlerin olumlu tutum geliştirmelerindeki en önemli sebeplerden biri ise okuma becerilerini ölçmek için pek çok hazır materyale hemen ulaşabilmeleridir (Backlund, Brown, Gurry, & Jandt, 1980). Dolayısıyla araştırmalarda ortaya çıkan bu benzer sonuçlar göz önünde bulundurulduğunda, öğretmenler için okuma becerilerini değerlendirme sürecinin diğer beceri alanlarına göre nispeten daha kolay olduğu savunulabilir.

Yazma becerilerinin değerlendirilmesiyle ilgili bulgular incelendiğinde ise, bu alanın ne dinleme ve konuşma becerileri kadar ihmal edildiği ne de okuma becerileri kadar çok önemsendiği söylenebilir. Cihan ve Gürlen'in (2009) 5. sınıf İngilizce öğretim programına ilişkin öğretmen görüşlerini sundukları çalışmada benzer sonuçlara ulaşılmış ve öğretmenlerin dört temel yabancı dil becerisini önemsedikleri ancak daha çok okuma ve yazma alanlarıyla ilgili ölçme ve değerlendirme tekniklerini kullandıkları kaydedilmiştir. Diğer taraftan Han ve Kaya'nın (2014) çalışmasında yabancı dil öğretmenlerinin ölçme ve değerlendirme uygulamaları ile tercihleri araştırılmış ve yazma ile dinleme alanlarının öğretmenler tarafından daha az önemsendiği sonucuna ulaşılmıştır. Aslında öğretim programındaki yazma ve dinleme ile ilgili kazanımların ulaşılabilir olduğu bulgusuna ulaşan Büyükduman (2005), sınıfların kalabalık olmasından dolayı bu alanlarda yeterli düzeyde ölçme ve değerlendirme uygulamalarının yürütülemediğini ifade etmiştir. Demirtaş ve Erdem (2015) benzer şekilde yazma becerilerinin değerlendirilmesinin öğretmenler için zorlayıcı olduğu sonucuna ulaşmıştır. Bu araştırmada ise nitel ve nicel veriler göz önünde bulundurulduğunda öğretmenlerin yazma becerilerinin değerlendirilmesine yönelik olumlu görüşlerinin daha fazla olduğu kaydedilmiştir. Dolayısıyla alanyazında bu çalışmaların aksine, araştırmada yazma alanı öğretmenlerin önemsedikleri bir beceri olarak karşımıza çıkmaktadır.

Yazma becerilerini geleneksel veya alternatif araçlarla değerlendirmenin pek çok yolu vardır (Brown & Abeywickrama, 2010). Bu araştırmanın nitel verileri göz önünde bulundurulduğunda öğretmenlerin daha çok yazılı yoklamalar, ders kitabındaki uygulamalar ve tamamlayıcı araçlarla öğrencilerin yazma becerilerini değerlendirdikleri kaydedilmiştir. Nitekim nicel araştırmada öğretmenlerin tamamlayıcı araç ve teknikleri en fazla yazma becerilerini değerlendirirken kullandıklarına dair ulaşılan sonuçlar nitel araştırma verileriyle örtüşmektedir. Kırkgöz ve Ağçam'ım (2012) çalışmasında öğretmenlerin yazılı sınavlar, ev ödevleri ve projelerle öğrencilerin yazma becerilerini değerlendirdiklerine dair ulaştıkları sonuçlar bu araştırmanın bulgularıyla tutarlılık göstermektedir.

Araştırmada "Uygulama" ana teması altında öğretmenlerin hangi klasik ölçme araçlarını ne amaçla kullandıklarına dair bulgular elde edilmiştir. Elde edilen bulgulara göre öğretmenlerin İngilizce becerilerinin ölçülmesinde klasik ölçme araçları arasında en çok çoktan seçmeli testleri tercih ettikleri görülmektedir (f=23). Daha sonra doğru-yanlış

testleri, eşleştirme soruları gibi işaretlemeyi gerektiren testlerin kullanıldığı ortaya çıkan sonuçlar arasındadır. Bu tür testlerin daha fazla tercih edilmesinin sebepleri *"Puanlaması kolay"* ve *"Hazırlaması kolay"* alt temalarıyla dile getirilmiştir. Ayrıca sınıfların kalabalık olması gibi fiziki şartlar göz önünde bulundurulduğunda, test uygulayıcıları için puanlanması ve uygulaması hızlı olan testlerin kullanımının öncelikli olduğu belirtilmiştir. Büyükduman (2005) yaptığı çalışmada benzer sonuçlara ulaşmış ve sınıf mevcutlarının kalabalık olmasından dolayı öğretmenlerin hazırlaması ve puanlaması zaman alacak soru tiplerini kullanmaktan kaçındıklarını belirtmiştir.

Dil becerilerini ölçmede testleri hem öğretmenler hem de öğrenciler için önemli bir geribildirim aracı olarak gören Davies (1990), öğretmenlerin ölçmenin öğrenme üzerindeki etkisini göz ardı edip ölçmeyi sadece öğrenci performansını puanlama olarak algıladıkları ve ölçme değerlendirme uygulamalarına dikkat etmediklerini ileri sürmüştür. Araştırmada yazılı ve sözlü cevap gerektiren testlerin diğer test türlerine göre daha az tercih edildiğine dair ortaya çıkan sonuçlar Davies'in bu tezini doğrular niteliktedir. Nitekim öğretmenlerin görüşlerine göre çoktan seçmeli araçların tercih edilmesinin en önemli sebebi puanlamasının kolay olmasıdır. Bible, Simkin ve Kuechler (2008), ölçme aracı seçme konusunda öğretmenlerin neden daha fazla ilgili olması gerektiğini şu sebeplerle açıklamışlardır: Birincisi, öğrencilerin bilgisini doğru ve açık bir şekilde ölçecek sınavlar hazırlamak için. İkincisi, öğrencilerin konuları anladıklarını yeterli bir detaylandırmayla test etmek ve objektif bir puanlama yapmak için. Üçüncüsü ise, öğretim yöntemlerinin etkililiğiyle ilgili dönüt elde etmek için. Dolayısıyla araştırmacıların öğretmenlere ne düzeyde doğru ölçümler yaptıklarına dair yeterli veri sunmayan çoktan seçmeli testlerin yerine öğrencilerin cevapları yapılandırmaları gereken soru türlerini kullanmalarını önerdikleri söylenebilir.

Öğrencilerin cevapları sözlü ya da yazılı bir şekilde yapılandırmaları gereken soru tipleri çoktan seçmeli testlere göre gerçek yaşam durumlarına daha yakın bulunmaktadır (Hancock, 1994; Bacon, 2003). Ayrıca bu tür testler şans faktörünü ortadan kaldırdığı için öğrencilerin sorulara bilinçli bir şekilde cevap verip vermediklerine dair şüpheyi en aza indirmektedir (Tsui, Lau & Font, 1995). Güneş (2012) çoktan seçmeli testleri öğrencileri bilgiyi sadece hatırlama ve anlamaya yönelttiği için eleştirmiştir. Araştırmanın bulguları arasında yer alan "*Çoktan seçmeli testleri tercih etmiyorum. Öğrenciyi ezbere yönlendirdiğini düşünüyorum (Ö19)"* ifadesi bu eleştiriyi doğrular niteliktedir. LeClercq'a

(1986) göre soru tiplerinin ve cevaplarının öğrenciler tarafından ezberlenmesi zamanla öğrencilerin zihinsel süreçlerini ve yeteneklerini sınırlandırarak dünyayla ilgili algılarını da basitleştirecektir.

Çoktan seçmeli testlerin çeşitli zayıflıklarına rağmen yabancı dilde dinleme, okuma, dilbilgisi, sesletim ve kelime bilgisini ölçmede oldukça kullanışlı araçlar olduğu ifade edilmiştir (Brown & Hudson, 1998). Dursun'un çalışmasında (2014) öğretmenlerin özellikle okuma ve dinleme becerilerini ölçmek için en fazla tercih ettikleri araçların çoktan seçmeli test formatı olduğu kaydedilmiştir. Ancak Tekin'e (2004) göre, konuşma ve yazma gibi üretimsel becerilerin test edilmesinde, ilgi ve tutumlarının belirlenmesinde yazılı yoklamalar daha kullanışlıdır. Ayrıca bu tür sınavlar sistemli ve kompleks düşünme gibi üst düzey düşünme becerileri gerektirdiğinden diğer sınav türlerinden oldukça farklıdır (Reiner, Bothell, Sudweeks & Wood, 2002). Araştırmada yazılı veya sözel ifadeyi gerektiren testleri tercih eden katılımcıların sayısının daha az olduğu saptanmıştır. Bu durum yazılı ifade gerektiren testlerin puanlamasının seçmeyi ve işaretlemeyi gerektiren testlere göre daha zor olması ile açıklanabilir.

Araştırmada ölçme araçları olarak çoktan seçmeli testlerin tercih edilmesinin sebeplerinden biri olarak sınav odaklı eğitim anlayışı gösterilmiştir. Bu anlayışın hem öğrencileri hem de öğretmenleri bu tür testleri kullanmaya yönelttiği ifade edilmiştir. Sınav sisteminin ölçme araçlarının seçimini etkilediğine dair ortaya çıkan bu bulgular Çelebi ve Narinalp'in (2020) araştırma sonuçlarıyla benzerlik göstermektedir. Nitekim bu çalışmada sınav sisteminin İngilizce öğretimini olumsuz etkilediği ve dil becerilerinin testlerle ölçülmesinin öğrenmeyi mekanikleştirdiği yönünde bulgulara ulaşılmıştır.

Araştırmada ortaya çıkan bir diğer önemli görüş çoktan seçmeli testlerin hem öğretmenler hem de öğrencilerin en alışık oldukları yöntem olmasından dolayı tercih edilmesidir. Fulcher'a (2012) göre öğretmenlerin birer öğrenen ve aday öğretmen olarak hizmet öncesi ölçme değerlendirmeyle ilgili yaşadıkları deneyimler onların bu alandaki görüşlerini ve seçimlerini etkilemektedir. Dolayısıyla öğretmenlerin çoktan seçmeli testleri kullanmayı tercih etmeleri mesleki hayatlarından önce bu tür ölçme araçlarına daha fazla maruz kalmaları ile açıklanabilir. Benzer şekilde Gelbal ve Kelecioğlu (2007) ve Arı (2013) tarafından yapılan araştırmalarda hem öğretmenler hem de öğrencilerin geleneksel olarak adlandırılan bu sınav türlerinde kendileri daha yeterli gördüklerine dair ortaya çıkan sonuçlar araştırmanın bulgularını desteklemektir.

Bir başka çalışmada Peterson ve Reider (2002) test formatının öğrencilerin testle ilgili algılarını ve performanslarını etkilediğini ileri sürmüştür. Bu iddiayı yaptığı araştırmayla destekleyen Kobayashi (2002), ölçme aracındaki soru tiplerinin yabancı dil öğrenenlerin performansını etkilediğine dair anlamlı bulgular elde etmiştir. Shohamy (1984) yabancı dil öğrencilerinin okuma becerilerini çoktan seçmeli test ve açık uçlu sorularla değerlendirmiş ve öğrencilerin çoktan seçmeli testte daha başarılı oldukları sonucuna ulaşmıştır. Soru tiplerinin öğrenci performansını etkilediğine dair sonuçlar bu araştırmanın bulgularını da desteklemektedir. Araştırmada ortaya çıkan "…*Çoktan seçmeli testlerin öğrencileri daha fazla motive ettiği ve kaygı seviyelerini düşürdüğünü söyleyebilirim (Ö32;ÖO7)"* şeklindeki görüşler bu bulguları doğrulamaktadır. Bu tür ölçme araçlarının sınav sisteminden dolayı öğrencilerin beklentilerini karşıladığı ve motivasyonlarını arttırdığı söylenebilir.

Araştırmada "Uygulama" ana teması altında İngilizce öğretmenlerine klasik ölçme araçlarının yanı sıra hangi tamamlayıcı ölçme değerlendirme araç ve tekniklerini kullandıkları sorulmuştur. Çalışmaya katılan öğretmenlerin görüşleri incelendiğinde, tamamlayıcı ölçme ve değerlendirme araçları arasında en fazla proje ve performans görevleriyle birlikte öğrenci ürün dosyasının (portfolyo) tercih edildiği sonucuna ulaşılmıştır. Bu sıralamayı kontrol listeleri ve dereceli puanlama anahtarlarının kullanımı takip etmektedir. Öz değerlendirme ve akran değerlendirme tekniklerinin ise daha az kullanıldığı kaydedilmiştir. Ayrıca bu tür araçları sınırlı bir şekilde ya da hiç kullanmadığını belirten görüşler de bildirilmiştir. Fidan ve Sak'ın (2012) araştırmasında öğretmenlerin tamamlayıcı ölçme ve değerlendirme tekniklerini kullanma sıklıkları incelenmiş ve performans ve proje değerlendirme, portfolyo, puanlama ölçeklerinin daha fazla kullanıldığı belirtilmiştir. Benzer şekilde Karamustafaoğlu, Çağlak ve Meşeci (2012) yaptıkları çalışmada, öğretmenlerin performans ve proje değerlendirme ile değerlendirme ölçeklerini diğer araç ve tekniklerden daha sık kullandıklarını kaydetmişlerdir. Ortaya çıkan araştırma bulgularının bu çalışmaya katılan öğretmenlerin tamamlayıcı ölçme ve değerlendirme tercihleriyle örtüştüğü söylenebilir.

Araştırmaya katılan öğretmenler proje ve performans görevlerini çoğunlukla öğrencilerin aktif olmasını sağladığı ve öğrenme sürecini desteklediği için kullandıklarını belirtmişlerdir. Portfolyoları ise daha çok öğrenci gelişim sürecinin takibini sağladığı için tercih ettiklerini ifade etmişlerdir. Çetin'in (2011) araştırmasında öğretmenlerin

tamamlayıcı araçlar arasında en fazla gözlem ve portfolyoyu kullandıkları ve bu araçların öğrencilerin kendi öğrenmelerinden sorumlu olmalarını sağlayarak özerkliklerini desteklediği belirtilmiştir. Bu durumun öğrencileri daha iyi tanıma fırsatı sunduğu ve öğrenme süreciyle ilgili dönüt sağladığına ilişkin ortaya çıkan sonuçlar bu araştırmanın bulgularıyla tutarlılık göstermektedir. Bu görüşleri destekleyen Hamayan'a (1995) göre, alternatif değerlendirmeler süreç içerisinde sürekli devam ettiği için öğrenenlerin dil gelişim süreçlerine dair bir ayna tutmaktadır. Genesee ve Upshur (1996) bu tür değerlendirmelerin, öğrenme stratejileri ve stilleri, öğrenci davranışları ve tutumları gibi öğrencilerin başarılarını etkileyen faktörlerle ilgili bilgi sunduğunu belirterek alternatif değerlendirmenin önemini vurgulamıştır. Hatta Alderson ve Banerjee'ye (2001) göre alternatif araçların öğrenciler hakkındaki sundukları bilgiler geleneksel test puanlarından daha kapsamlı ve bilgilendiricidir. Araştırmada tamamlayıcı araçların iyi bir dönüt sağladığı şeklindeki görüşleri destekleyen Genesee ve Hamayan (1994), alternatif değerlendirmenin öğretim ile değerlendirme arasında bir geribildirim döngüsü oluşturarak uygulayıcılara öğretimin izlenmesi ve düzenlenmesi ile ilgili sürekli bilgi sunduğunu ifade etmiştir.

Hem nitel hem de nicel araştırma bulgularına göre, araştırmaya katılan öğretmenlerin çoğu tamamlayıcı ölçme değerlendirme araç ve tekniklerini değerlendirme sürecinde kullandıklarını belirtmişlerdir. Çelebi ve Kuşuçuran (2019) İngilizce öğretmenlerinin alternatif değerlendirme yöntemlerini kullanma yeterliklerini inceledikleri çalışmada benzer bulgulara ulaşmış ve öğretmenlerin genelinin alternatif değerlendirme araçları konusunda bilgi sahibi ve yeterli olduğunu kaydetmişlerdir. Ayrıca onların araştırmasında öğretmenlerin geleneksel yöntemlerle birlikte bu tür araçları kullandıkları görüşü ön plana çıkmıştır. Geleneksel ve alternatif değerlendirme araçlarının birlikte kullanıldığına dair ulaşılan sonuçların bu araştırmanın bulgularıyla örtüştüğü söylenebilir. Araştırmaya katılan öğretmenlerin bir önceki soruda yanıtladıkları klasik ölçme ve değerlendirme araçlarına dair görüşleri ile bu soruda yanıtladıkları tamamlayıcı ölçme değerlendirme araç ve tekniklerine dair görüşleri birlikte incelendiğinde; öğretmenlerin çoktan seçmeli testler gibi ölçme araçlarını sıklıkla kullandıkları ve aynı zamanda tamamlayıcı ölçme değerlendirme araçlarının kullanımına özen gösterdiklerini belirtmeleri bu tezi doğrular niteliktedir.

Araştırmanın önemli bulgularından biri de öğretmenlerin sınıf mevcutlarının kalabalıklığı ve öğrenciler için zorlayıcı olması nedenleriyle tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini kullanamadıklarına dair görüşlerdir. Gelbal ve Kelecioğlu'nun (2007) çalışmasında öğrencileri tanıma ve başarı düzeylerini değerlendirmede daha çok geleneksel yöntemlerin tercih edildiği ve yeni değerlendirme yaklaşımlarının ise hazırlanmasının zorluğu, sınıfların kalabalık oluşu ve öğretmenlerin kendilerini yeterli görmedikleri bir alan olması sebebiyle kullanılmadığı ifade edilmiştir. Öğretmenlerin büyük bir kısmı tamamlayıcı yöntemleri kullanmadıklarına dair benzer sebepler sunsa da bu araştırmanın bulgularıyla bütünüyle örtüştüğü söylemez. Çünkü bu araştırmada tamamlayıcı araçları kullanmadıklarını vurgulayan katılımcı sayısının oldukça az olduğu görülmektedir (f=3). Demirtaş ve Erdem (2015) tarafından yapılan başka bir çalışmada, öğretmenlerin büyük bir çoğunluğunun (% 84.2) öğretim programında öngörülen tamamlayıcı ölçme değerlendirme araçlarını kullanmadıklarına dair bulgular bu araştırmanın sonuçlarıyla örtüşmemektedir. Onların araştırmasında katılımcıların bu tür araçları kullanmama nedenlerine dair en çok vurgulanan görüşler, öğrenci seviyesine uygun olmadığı ve öğrencilerin objektif bir şekilde değerlendirilemedikleri yönündedir. Merter, Kartal ve Çağlar (2012) İngilizce dersi öğretim programının öğretmen görüşlerine göre incelemiş ve performans ve proje görevlerinin öğrencilerin kendi başlarına somut ve yaratıcı ürünler oluşturabilecekleri görevler olmadığı konusunda öğretmenlerin çoğunlukla olumsuz görüş bildirdiklerini belirtmişlerdir.

Alanyazında tamamlayıcı ölçme ve değerlendirme araçlarının çeşitli nitelikleriyle ilgili benzer kaygıların olduğu ifade edilmiştir (Nasab, 2015). Bu kaygılardan biri alternatif araçların dikkatli ve özenli bir analiz gerektirmesi sebebiyle çok zaman alıcı ve maliyetli olmasıdır (Brindley, 2001). Bir diğer önemli husus öğretmenlerin alternatif araçları başarılı bir şekilde uygulayabilmeleri için yeteri kadar bilgi ve donanım sahibi olmaları gerektiğidir (Clark & Gipps, 2000). Ayrıca büyük ölçekteki sınıflarda alternatif araçların doğru bir şekilde kullanılmasının mümkün olamayacağı savunulmaktadır (Worthen, 1993). Worthen'in belirttiği bu durum araştırmada kimi öğretmenlerin sınıf mevcutlarının kalabalıklığı nedeniyle tamamlayıcı araçları kullanmadıklarına dair görüşleri destekler niteliktedir.

Araştırmada yabancı dilde ölçme ve değerlendirme uygulamalarıyla ilgili bir diğer soru dilbilgisi, sözcük bilgisi ve telaffuz gibi dil yapılarının nasıl ölçüldüğüyle ilgilidir.

Araştırmanın nitel bulgularına göre öğretmenler en fazla sözcük bilgisinin değerlendirilmesine yönelik uygulamalar yapmaktadır. Eskisi kadar önemsenen bir alan olmadığı belirtilmekle birlikte dilbilgisi yapılarının gerekliliğini savunan ve test ettiğini ifade eden öğretmen sayısı ise azımsanamayacak kadar fazladır (f=23). Telaffuz becerisinin ölçülmesine ilişkin görüşler arasında ise herhangi bir ölçme yapmadığını belirten görüşlerin ön plana çıktığı tespit edilmiştir (f=25). Diğer taraftan araştırmanın ikinci aşamasında yürütülen nicel araştırma sonuçlarının nitel verilerle tutarlı olduğu görülmüştür. Nitekim öğretmenlerin sözcük bilgisi ve dil yapılarının ölçülmesiyle ilgili davranışları *"Çoğu zaman"* düzeyinde sergiledikleri belirlenmiştir. Ancak telaffuz yapılarının da çoğu zaman ölçüldüğüyle ilgili ulaşılan veriler nitel araştırmanın bulgularıyla örtüşmemektedir. Bu durum öğretmenlerin telaffuz yapılarının ölçülmesiyle ilgili yürüttükleri faaliyetlerin çoğunlukla gözlem yoluyla gerçekleştiğini ifade eden ve araştırmanın nitel aşamasında ortaya çıkan *"Gözlem yoluyla ölçme yaparım"* alt temasıyla açıklanabilir. Dolayısıyla nicel araştırmaya katılan öğretmenlerin sınıf içi etkinlikler sırasında gözlem yoluyla bu becerileri değerlendirdikleri söylenebilir. Balım'ın (2020) çalışmasında öğretmenlerin öğrencilerin konuşma performanslarına yönelik becerilerini ölçerken bir ölçme aracı kullanmak yerine süreç içerisinde gözlem yaptıklarına dair ortaya çıkan bulgular bu tezi destekler niteliktedir.

Dil öğretiminin en önemli iki yapı taşı olarak kabul edilen dilbilgisi ve sözcük bilgisi, geçmişte dilin değişmeyen kurallarını yansıtan ve durağan bir yapıda olan bilgi yığınlarını temsil etmekteydi (Yang, 2020). Şimdilerde ise dil öğretiminde geleneksel anlayışın yerini alan ve iletişimin gücünü savunan paradigmalarla birlikte öğrenenlerin dili iletişimsel bir şekilde kullanmaları için dilbilgisi ve sözcük bilgisi birer kaynak olarak görülmeye başlanmış ve farklı bir perspektif kazanmıştır (Larsen-Freeman, 2011). Bu kaynak özellikle dil yeterlik testlerinde dilin makro becerileri olan dinleme, konuşma, okuma ve yazmayı içeren görevleri tamamlamak için kullanılmaktadır (Jamieson vd., 2000). Böylece dilin dört dil beceri alanını bir araya getiren bir yapıda oldukları ve dil öğretiminde ve değerlendirme sürecinde eski zamanlardan beri süregelen önemli mikro beceriler oldukları söylenebilir (Purpura, 2004).

Dil yapılarının makro becerilerle birlikte öğrenilmesi ve test edilmesi dilin bağlam içinde kullanılmasını da kolaylaştırmakta ve dile iletişimsel bir boyut kazandırmaktadır (Yang, 2020). Nitekim yeni İngilizce öğretim programı incelendiğinde, dil sınavlarının

temel olarak dört beceriyi kapsaması; dilbilgisi, sözcük bilgisi gibi dil yapılarının ise örtük bir şekilde ve bir bağlam içinde ölçülmesi gerektiği vurgulanmaktadır (MEB, 2018). Araştırmanın nitel verileri ve dil becerilerinin mikro yapılarına ilişkin bu kuramsal temeller göz önünde bulundurularak hazırlanan anketin maddeleri incelendiğinde, öğretmenlerin dinleme, konuşma, okuma ve yazma alanlarında dilin sözcük ve yapı bilgisine dair ölçümleri çoğu zaman yerine getirdikleri belirlenmiştir. Ancak araştırmada öğretmenlerin çoğu zaman düzeyinde sergiledikleri dil yapılarına dair bu ölçümlerin belirli bir bağlam gözetilerek yapılıp yapılmadığına dair verilere rastlanılmamaktadır. Benzer bir vurguya değine Dursun (2014) araştırmasında tüm dil becerilerinin birbiriyle bağlantılı olduğu ve alt beceriyle birlikte ölçülmesi gerektiğini ifade etmiştir.

Araştırmanın önemli bulgularından biri ise sınav sisteminden dolayı okuma ve anlama becerilerinin ön plana çıktığı ve bu durumun da sözcük bilgisinin geliştirilmesine yönelik bir hassasiyet oluşturduğu yönündeki görüşlerdir. Hatta sözcük bilgisinin dilbilgisinden daha önemli olduğu ve daha fazla test edildiğine yönelik pek çok ifade kaydedilmiştir. Çakır (2017) yaptığı araştırmada benzer bulgulara ulaşmış ve İngilizce derslerinde öğretmenlerin çoğunun (%80) sınav sisteminden dolayı öğrencilerin test başarıları üzerinde durdukları ve kelime öğretimine sınıf içi ve sınıf dışında daha fazla yer verdiklerini ifade etmiştir. Kelime bilgisini takiben en fazla dilbilgisi yapılarıyla ilgili ölçümler yapıldığına dair ulaşılan diğer bulgular da bu araştırmanın nitel verileriyle örtüşmektedir. İngilizce öğretmenlerinin çoğunlukla dilbilgisi temelli ölçümler yaptıklarını belirten Dilekli (2018) ve Merter, Kartal ve Çağlar (2012), öğretmenlerin öğretim programının yoğun olmasından dolayı konuşma ve yazma alanlarına yeteri kadar zaman ayıramadıklarını ve daha kısa sürede bu tür dil yapılarını ölçmeye yöneldiklerini ifade etmişlerdir.

Birbirini destekleyen bu araştırma bulgularının aksine İleri (2019), İngilizce öğretmenlerinin konuşma becerisinin ölçülmesine ilişkin gereksinimlerini belirlemeyi hedeflediği çalışmada öğretmenlerin konuşma becerisini değerlendirmede en fazla önem verdikleri ölçütlerin telaffuz ve akıcılık daha sonra ise dilbilgisi ve sözcük bilgisi olduğu sonucuna ulaşmıştır. Araştırmanın nicel verileri incelendiğinde, öğretmenlerin anketin konuşma bölümünde "*Sözcüklerin doğru telaffuz edilmesine dikkat ederim"* maddesine verdikleri yanıtların ortalamasının ($\bar{X}$=4,14), sözcük ($\bar{X}$=4,03) ve dilbilgisi ($\bar{X}$=3,97) yapılarına verilen yanıtların ortalamasından daha yüksek olduğu görülmektedir.

Dolayısıyla her iki araştırma sonucunun birbiriyle tutarlılık gösterdiği ve telaffuzla ilgili yapıların öğrencilerin konuşma performansını ölçmek için gereken önemli bir ölçüt olarak algılandığı söylenebilir. Diğer taraftan Büyükdumanın'ın (2005) çalışmasında öğretmenlerin büyük bir çoğunluğunun sınıf mevcutlarının kalabalık olmasından dolayı öğrencilerin telaffuzlarını düzeltme ve onlarla tek tek ilgilenme fırsatı bulamadıkları belirtilmiştir. Bu araştırmanın nitel verilerine göre ise öğretmenler çoğunlukla sınav sisteminden dolayı telaffuzla ilgili ölçümler yapamadıklarını ifade etmişlerdir.

İngilizce öğretiminde bilişsel kazanımların ne düzeyde gerçekleştirildiğinin ölçülmesi kadar önemli olan bir diğer boyut da duyuşsal becerilerin ne kadar önemsendiği ve ölçme ve değerlendirme sürecinin bir parçası olup olmadığıdır. Bu yüzden araştırmanın "Uygulama" ana teması altındaki bir diğer sorusu duyuşsal becerilerin ölçülmesi ile ilgili olmuştur. Araştırmaya katılan öğretmenlerin görüşleri incelendiğinde, öğretmenlerin büyük bir çoğunluğunun duyuşsal becerilerin ölçülmesiyle ilgili olumlu görüş bildirdikleri kaydedilmiştir. Öğretmenlere göre duyuşsal becerilerin dikkate alınmasının en önemli gerekçesi öğrencilerin ilgi ve tutumlarını ortaya çıkarmasıdır. Etkili bir dil öğretimi için öncelikle öğrencilerin derse ve öğretmene karşı ilgisinin olumlu yönde olması gerektiğine dikkat çekilmiştir.

20. yüzyılın ortalarına kadar araştırmacılar sadece bilişsel yaklaşımlarla öğrenmenin nasıl gerçekleştiğini açıklamaya çalışmışlardır (Djigunovic, 2006). 1970'lerde ise bilişsel yaklaşımların zihnin nasıl çalıştığına dair sunduğu bilgilerin sınırlı olduğuna ve eğitim ve psikolojide öne çıkan hümanistik yaklaşımların etkisiyle duyguların mantık ve düşünmeyle olan ilişkisine dikkat çekilmiştir (Arnold, 2011). Bu bağlamda zihnin duygular olmadan aslında gerçek bir zihin olamayacağı düşüncesi hızla önem kazanmış (Keblowska, 2012) ve duygular öğrenenlerin algı ve başarılarını etkileyen faktörlerden biri olarak kabul edilmiştir (Brown, 2001). Artık öğrenmenin bilişsel ve duyuşsal alanlarının birbiriyle bağlantılı olduğu ve birbirlerini tamamladıkları vurgulanarak her iki alanın birlikte düşünülmesi gerektiği savunulmuştur (Schmumann, 1994) Bununla birlikte duyguların sadece bilişsel değil psiko-motor alan becerilerini de etkilediği ileri sürülmüştür (Harrow, 1977). Hareket etme isteğinin tutumlarla komuta edildiği ve eğer öğrenenlerin yeteri kadar hareket etme istekleri olmazsa hiçbir şey elde edemeyecekleri ifade edilmiştir (Harrow, 1977).

Temel olarak davranışları şekillendiren duygu, his, durum ve tutum olarak tanımlanan duyuşsal alan (Arnold & Brown, 1999); dil öğreniminde öğrenenlerin iletişim becerilerini pozitif ya da negatif bir şekilde etkileyebilmektedir (Piasecka, 2013). Arnold'a (2011) göre pozitif etki öğrenme için oldukça değerli bir destek sunarken, negatif etki tam tersi zihni kapatmakta ve öğrenmeyi engellemektedir. Bu etkinin uygun girdiler sunulduğunda bile dil öğrenenlerin performansını etkileyen metaforik bir bariyer gibi çalıştığı söylenebilir (Lightbown & Spada, 2006). Krashen'ın (1987) "Duyuşsal Filtre Hipotezi" olarak adlandırdığı bu metaforik bariyer; bir filtre gibi öğrenenlerin kaygılı ya da motivasyonlarının düşük olması durumunda yükselirken, özgüvenli ve rahat hissettiklerinde ise düşmektedir. Bu filtre hipotezine göre olumsuz duygular dil başarısını etkileyen bir bariyer gibidir ve istenilen davranışların gerçekleşmesini kısıtlamaktadır.

Görüldüğü üzere zaman içinde duygulara atfedilen değer oldukça önem kazanmış ve bilişsel ve psiko-motor alanlarla ilişkisine dikkat çekilerek öğrenme üzerindeki etkileri araştırmacıların ilgisini çekmiştir. Dolayısıyla bilişsel alan davranışları kadar duyuşsal becerilerin test edilmesi gerektiği öne sürülmüştür (Worrell, Evans-Fletcher & Kovar, 2002). Ancak Popham'a (1995) göre, öğretmenler duyuşsal alan davranışlarını sistematik bir şekilde takip edip değerlendirme faaliyetlerini nadiren gerçekleştirmektedirler. Özellikle daha büyük yaş gruplarıyla çalışan öğretmenler için ise bu durum daha az rastlanılır bir hal almaktadır (Popham, 1995). Araştırmaya katılan öğretmenlerin görüşleri incelendiğinde, duyuşsal alan davranışlarının değerlendirilmesiyle ilgili olumlu görüşlerin sayısı fazla olsa da bu işi tutum ölçekleri, ilgi envanterleri veya gözlem formu ile sistematik bir şekilde yürütenlerin sayısının oldukça az olduğu söylenebilir. Duyuşsal becerileri değerlendirdiklerine dair görüş bildiren öğretmenlerin çoğu bunu öğrencilerin sınıf içindeki davranışlarını doğal bir şekilde izleyerek yaptıklarını belirtmişlerdir. Adeyato'ya (2014) göre bu tür informal gözlemler olayların gerçek sebeplerini ortaya çıkarmada yetersiz kalabileceğinden derecelendirme ölçekleri gibi araçların gözlem sırasında kullanımı güvenirliği arttırabilecektir.

Araştırmada duyuşsal becerilerin ölçülmesi için kullanılacak araçlarının seçimi, uygulanması ve değerlendirilmesi gibi aşamaların uzun zaman gerektirdiği ve İngilizce ders saatlerinin kısıtlı olmasından dolayı bu tür çalışmalar yapılamadığı ortaya çıkan bulgular arasındadır. Bu düşünceyi savunan Martin ve Reigeluth'a (1992) göre, duyuşsal becerileri değerlendirmek bir ölçme aracı hazırlamak kadar zor ve karmaşık bir süreçtir.

Ayrıca geçerli, güvenilir ve objektif sonuçlar elde etmek güç olabilir. Diğer taraftan duyuşsal alan davranışlarına dair değerlendirmelerin sadece testle sınırlı kalamayacağı belirtilerek (Arnold, 2011), öğretmenlerin sınıfta destekleyici bir atmosfer yaratmaları, test formatı ve öğretim materyallerini öğrencilerin ilgileri doğrultusunda düzenleyerek test süreçlerini daha az stresli bir hale getirmelerinin de öğrencilerin duyuşsal giriş davranışları üzerinde olumlu bir etki yaratacağı ifade edilmiştir (Chastain, 1988).

Bilişsel ve duyuşsal davranışların ne düzeyde ve nasıl ölçüldüğüyle ilgili sorulara aranan yanıtlardan sonra, araştırmaya katılan öğretmenlerin ölçme sonuçlarını kullanarak belirli bir karara ulaştıkları süreç olan değerlendirme faaliyetlerine yönelik uygulamaları araştırılmıştır. "Değerlendirme" ana teması altında öğretmenlere belirli bir amaca yönelik yapılan tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme türlerinden hangilerini kullandıklarına dair görüşleri sorulmuştur. Elde edilen bulgulara göre, öğretmenlerin tümü bir tercih olmaktan ziyade zorunlu olduğu için düzey belirleyici değerlendirme yöntemini kullandıklarını ifade etmişlerdir. Düzey belirleyici değerlendirmeden sonra en fazla tercih edilen yöntemin biçimlendirici değerlendirme olduğu göze çarpmaktadır. Katılımcıların çoğu bu tür değerlendirmeyi öğrenme eksikliklerini belirlemek için ihtiyaç halinde kullandıklarını belirtmişlerdir. Tanılayıcı değerlendirmenin ise diğer değerlendirme türlerine göre nispeten daha az tercih edildiği ve çoğunlukla öğrenci seviyesi ve hazırbulunuşluklarını belirlemek amacıyla yapıldığı saptanmıştır. Diğer taraftan araştırmanın nicel bulguları incelendiğinde, öğretmenlerin okuma ve yazma alanlarında *"Çoğu zaman"* düzeyinde; dinleme ve konuşma alanlarında ise *"Ara sıra"* düzeyinde tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirmeler yaptıkları kaydedilmiştir. Bu sonuçlar katılımcıların dinleme ve konuşma alanlarında yeterli düzeyde ölçme faaliyetleri yürütmedikleri ancak okuma ve yazma alanlarına daha fazla yoğunlaştıkları yönünde ortaya çıkan önceki bulguları desteklemektedir. Nitekim uygun ölçme araçlarıyla yeterli düzeyde test edilmeyen dinleme ve konuşma beceri alanlarının yine yeterli düzeyde değerlendirilmesi beklenemez. Dolayısıyla araştırma sonuçlarının birbiriyle tutarlı olduğu söylenebilir.

Değerlendirme türleri öğretime girişte, süreçte ve çıkışta yapılandırılarak birbirini tamamlamakta ve öğrenmeye farklı açılardan katkı sunmaktadır (Demirel, 1992). Keeley'e (2008) göre tanılayıcı değerlendirme dönemin başında öğrencilerin zayıflıkları ve güçlü yönlerini tespit etmek için yapılır. Öğrenme faaliyetleriyle ilgili bilgi sunmak için

kullanılan biçimlendirici değerlendirme öğretim süreci boyunca kullanılabilir (Black & William, 2010). Düzey belirleyici değerlendirme ise genellikle öğretim sürecinden sonra tercih edilen bir yöntemdir (Boud, 2000). Araştırma bulgularına göre, öğretmenlerin tanılayıcı değerlendirmeyi dönem başında, biçimlendirici değerlendirmeyi süreç içerisinde ihtiyaç halinde ve düzey belirleyici değerlendirmeyi ise her dönemin ortasında ve sonunda tercih ettikleri sonucuna ulaşılmıştır. Elde edilen bulgular göz önünde bulundurulduğunda katılımcıların değerlendirme türlerini ne amaçla ve nerede kullanacaklarına dair geçerli algılar geliştirdikleri söylenebilir.

Öğretmenler çoğunlukla soru sorarak, gözlem yaparak ya da sınavlarla öğrencilerin ne bilip bilmediklerini tespit etmeye çalışırlar (Rea-Dickins, 2001). Ancak çok az öğretmen tanılayıcı değerlendirme yapmak için bu tür araçları kullanmayı tercih etmektedir (Edelenbos & Kubanek-German, 2004). Huff ve Goodman'ın (2007) araştırmasında öğretmenlerin çoğu kendileri tarafından geliştirilen ölçme araçlarının öğrenciler hakkında tanılayıcı bilgi toplamada kullanılabilecek en iyi araçlar olduğunu belirtmişlerdir. Ayrıca devlet denetiminde gerçekleştirilen sınavlar ya da ticari amaç taşıyan geniş kapsamlı testlerin tanılayıcı bilgi toplamak için oldukça uygun araçlar olarak kabul edildiği ortaya çıkan diğer bulgulardandır. Bu araştırmaya katılan öğretmenlerin görüşleri incelendiğinde ise, daha çok sözlü yoklamalarla tanılayıcı değerlendirmeler yapıldığına dair görüşler kaydedilmiştir. Dolayısıyla bu değerlendirme türüyle ilgili geniş çaplı çalışmaların yapıldığı ve tanılayıcı bilgi elde etmek için farklı ölçme araçlarının işe koşulduğunu söylemek güçtür.

Araştırmada biçimlendirici değerlendirmenin öğrenme eksiklerini veya düzeyini belirlemek, öğretim etkinliklerini düzenlemek gibi çeşitli amaçlarla tercih edildiği ifade edilmiştir. Buldur ve Hasbek (2020), öğretmen adaylarının biçimlendirici değerlendirmeye karşı algılarını metaforlar aracılığıyla araştırmış ve öğrenme düzeyini artırma ile öğrenme eksiklerini belirleme kategorilerinde en fazla metaforun geliştirildiğini saptamışlardır. Elde edilen bulgular bu araştırmaya katılan öğretmenlerin biçimlendirici değerlendirmeyi kullanma nedenlerine dair görüşlerini destekler niteliktedir. Dolayısıyla öğretmenlerin biçimlendirici değerlendirmenin karakteristik özelliklerine dair geçerli bilgilere sahip oldukları söylenebilir. Ayrıca araştırmanın bulguları arasında öğretmenlerin kelime bilgisini pekiştirmek için biçimlendirici değerlendirmeye ihtiyaç duyduklarına dair görüşler mevcuttur. Bu bulgular Körhasanoğulları'nın (2020) çalışmasında ortaya çıkan ve

özellikle 8. sınıfların kelime bilgisini test etmek için izleme testlerinin kullanıldığına dair sonuçlarla benzerlik göstermektedir.

Biçimlendirici değerlendirmenin öğrenme standartlarını arttırdığına dair bir yığın bilgiye rastlanabilir ancak pragmatik olarak bakıldığında çoğu sınıfta etkili ve planlı bir şekilde yürütülmemektedir (Black ve William, 2010). Araştırmaya katılan öğretmenlerin görüşleri incelendiğinde, biçimlendirici değerlendirmeyi ihtiyaç halinde tercih ettiklerine dair görüşün ön plana çıktığı görülmektedir. Dolayısıyla öğretmenlerin bu değerlendirme türüne yönelik yürüttükleri uygulamaların çoğunlukla planlı olduğu söylenemez. Bu bağlamda biçimlendirici değerlendirmelerin planlı ya da kendiliğinden gelişebileceğine dikkat çeken Dixson ve Worrell'a (2016) göre, ders esnasında yapılabilecek soru cevap gibi spontane süreçler ya da ara sınavlar ve ev ödevleri gibi planlı etkinliklerin hepsi öğrenci gelişimini değerlendirebilecek biçimlendirici araçlardır. Bu noktada önemli olan hangi strateji kullanılırsa kullanılsın öğrencilerin öğrenmelerini desteklemesidir (Wiliam & Thompson, 2007).

Öğrenci davranışlarının kapsamlı bir şekilde değerlendirilmesi ve öğrenmelerinin desteklenmesi için kullanılabilecek bir diğer strateji de biçimlendirici ve düzey belirleyici değerlendirmeleri bir arada kullanmaktır (Zeng & Fang, 2017). Hatta alanyazında bu değerlendirme yöntemlerinin birlikte kullanılmasının tek başına tercih edilmelerinden daha fazla yarar sağladığı bulgusuna ulaşılmıştır (Mahshanian, Shoghi & Bahrami, 2019).

İngilizce öğretiminde planlama, uygulama ve değerlendirme uygulamalarından sonra öğrencilere ölçme sonuçları hakkında geribildirim sunmak ölçme değerlendirme sürecini daha dinamik ve verimli kılabilecektir. Bu doğrultuda araştırmanın bu bölümünde katılımcıların ölçme sonuçları hakkında öğrencileri bilgilendirip bilgilendirmedikleri araştırılmıştır. Elde edilen bulgulara göre öğretmenlerin çoğu öğrencilere ölçme sonuçları hakkında geribildirimde bulunmayı oldukça önemsemektedir (f=30). Geribildirimde bulunmanın sağladığı faydalara dikkat çeken öğretmenlerin en fazla vurguladığı alt tema ise *"Öğrencilerin hatalarını görmelerini sağlar"* olmuştur (f=27). Göçer'e (2019) göre, öğretmenler tarafından yapılan testlerin not verme işlemi bittikten sonra sınav evrakının bizzat öğrencilere eksik ve üstün yönlerini görmeleri için dağıtılması önemli geribildirim sağlama yollarındandır. Nitekim bu araştırmada öğretmenlerin çoğu yazılı sınavlar sonrası sınav evrakını dağıttıklarını, sözlü sınavlarda veya sınıf içi etkinlikler sırasında ise anında dönüt verdiklerini belirtmişlerdir. Bu bulguları destekleyen Altunöz ve Bay (2018),

İngilizce öğretiminde kullanılan etkili stratejilerin analizini yaptıkları çalışmada, öğretmenlerin öğrencilere fayda sağladığı için geribildirim stratejisini kullandıklarını saptamışlardır. Ayrıca öğretmenlerin bu stratejiyi daha çok sınıf içi alıştırmalar sırasında ya da sonunda, ev ödevi kontrolünde ve telaffuz hatalarının giderilmesinde tercih ettikleri bulgusuna ulaşmışlardır.

Sınav evrakını öğrencilere geribildirim sağlamak için dağıtmak kadar önemli olan bir diğer konu da kağıtların ne zaman dağıtıldığıdır. Yaman'a (2018) göre sınav kağıtlarının öğrencilere hiç gösterilmemesi veya sınav sonuçlarının çok geç değerlendirilip öğrencilerle paylaşılması, bilgilerin hatırlanmasını zorlaştıracağından İngilizce öğrenmeyi de olumsuz etkilemektedir. Sınav sonrası sürece benzer şekilde vurgu yapan Lee (2017), özellikle kaliteli ve kapsayıcı dönütlerin ancak uygun zamanda verilmesiyle, öğrenci öğrenmeleri üzerinde olumlu etkisinden söz edilebileceğini ifade etmiştir. Geribildirim sunma zamanının önemine değinen görüşlere araştırma bulgularında da rastlanılmaktadır. Katılımcılardan birinin "*Bana göre geri bildirimler ne kadar hızlı yapılırsa o kadar yararlı olacaktır. Nasıl ki üzerinden zaman geçen bir olayın üzerimizde etkisi azsa bu durum ölçme için de geçerlidir. Her dönüt anında veya en kısa süre içerisinde gerçekleşmelidir" (Ö31; ÖO6)* şeklinde belirttiği görüş araştırmacıların açıklamalarını destekler niteliktedir. Bu bulgular göz önünde bulundurulduğunda, etkili bir geribildirim sağlanması için sınav sonuçlarının öğrencilere makul bir zaman sonra iletilmesi gerektiği söylenebilir.

Araştırmada geribildirim sunmanın önemli olduğu ancak sınıfların kalabalık olmasından dolayı yaşanan kargaşa ve zaman kaybı gibi endişelerin geribildirim sürecini kimi öğretmenler için güç kıldığı ulaşılan diğer bulgular arasındadır. Ersen-Yanık'ın (2007) araştırmasında sınıf mevcutlarının fazla olmasından dolayı öğretmenlerin her öğrenciye yeteri kadar zaman ayıramadığı bu yüzden de ev ödevleri ve sınavları sağlıklı bir şekilde inceleyip geribildirim sunamadıklarına dair ulaşılan sonuçlar bu araştırmanın bulgularını desteklemektedir. Bir başka çalışmada Altunöz ve Bay (2018) sınıf mevcutlarının kalabalıklığı ve ders saatlerinin kısıtlı olmasının öğrencilerle yakından ilgilenmeyi zorlaştırdığı ve öğretmenlerin sadece uygun gördükleri zamanda dönüt verdiklerini tespit etmişlerdir. Arı'nın (2014) çalışmasında öğretmenlerin yaşadığı zaman kaygısına değinilmiş ve proje, performans gibi görevlerin değerlendirilmesi ve dönüt sunulmasının ağır iş yükü gerektirdiğine dair görüşler bildirilmiştir. Araştırma örneklemi olarak ortaokullarda görev alan katılımcılarla yürütülen bu çalışmaların aksine, Ballantyne,

Borthwick ve Packer (2000) öğretim elemanlarının ölçme değerlendirme yeterlik algılarını araştırmış ve öğretim elemanlarının nitelikleri boyutunda, katılımcılarının en fazla öğrenmenin değerlendirilmesi ve dönüt sağlanması konularında eksikliklerinin olduğu sonucuna ulaşılmışlardır. Dolayısıyla çalışma grupları farklılaşsa da öğretenlerin geribildirim sunma konusunda benzer güçlükler yaşadıkları öne sürülebilir.

Şimdiye kadar İngilizce öğretmenlerinin ölçme ve değerlendirme uygulamalarıyla ilgili tecrübeleri aktarılarak var olan süreç kapsamlı bir şekilde sunulmak istenmiştir. Araştırmanın bu bölümünde ise, ölçme değerlendirme öğesinin daha etkili bir şekilde kullanılması için neler yapılması gerektiği tartışılmıştır. "Geribildirim" ana teması altında sunulan bu aşamada öğretmenlere yabancı dilde nitelikli bir ölçme değerlendirme sürecinin gerçekleşmesi için çözüm önerilerinin neler olduğu sorulmuştur. Elde edilen bulgular "Ölçme ve değerlendirme araç ve yöntemleri ile ilgili öneriler", "Öğrenme ortamı ile ilgili öneriler", "Öğretim programıyla ilgili öneriler" ve "Öğretim materyalleri ile ilgili öneriler" temaları altında sunulmuştur. Öğretmenlerin ölçme değerlendirme araçlarına yönelik görüşleri içerisinde tüm yabancı dil beceri alanlarının ölçülmesi ve farklı ölçme değerlendirme tekniklerinin kullanılması gerektiği en sık vurgulanan ifadelerdir. Erdem ve Yücel-Toy (2017) ve Mede ve Akyel (2014) yaptıkları araştırmalarda benzer bulgulara ulaşmışlar ve katılımcıların tüm yabancı dil beceri alanlarının eşit bir şekilde önemsenmesi ve ölçülmesi gerektiğine dikkat çekmişlerdir. Ayrıca Ortaokul İngilizce Öğretim Programlarının değerlendirilmesine ilişkin yapılan pek çok makale ve lisansüstü tez çalışmasında öğretmenlerin ölçme değerlendirme sorunlarıyla ilgili benzer önerilerde bulunduğu göze çarpmaktadır (Işık, 2008; Yörü, 2012; Dinçer, 2013). Çalışmaların çoğunda konuşma ile dinleme becerilerinin özellikle ihmal edilen alanlar olduğu ve bu alanların test edilmesine yönelik uygulamaların attırılması gerektiği sıkça dile getirilmiştir.

Araştırmaya katılan öğretmenlerin öğrenme ortamı ile ilgili görüşleri incelendiğinde, sınıf mevcutlarının azaltılması, yeterli fiziki ve teknolojik donanımın sağlanması ve dil sınıflarının oluşturulması ile ilgili önerilerin sıralandığı görülebilir. Özellikli yetersiz materyal ve donanım ile sınıfların kalabalık oluşunun öğretmenleri sadece ölçme değerlendirme sürecinde değil aynı zamanda öğretim yöntemi, materyal seçimi gibi pek çok konuda sınırlandırdığı ifade edilmiştir. Alanyazında okulların fiziki ve teknolojik yapılarının iyileştirilmesi gerektiğini vurgulayan çeşitli çalışmalar mevcuttur (Ergüç, 2004; Peker, 2006; Dilekli, 2008;). Bunun yanında dil sınıflarının oluşturulması

önerisini yaptıkları çalışmayla destekleyen Çelebi ve Narinalp (2020), dört temel dil becerisini geliştirmek için öğrencilerin faydalanabilecekleri yabancı dil kütüphanelerinin ya da dil laboratuvarlarının okullara sağlanmasının önemine vurgu yaptıkları görülmektedir.

Öğretim programıyla ilgili sunulan öneriler incelendiğinde ise, katılımcıların ders saatlerinin arttırılması ve ünite sayısını azaltılması gerektiğine dair görüşleri ön plana çıkmaktadır. Balım'ın (2020) çalışmasında benzer bulgulara ulaşılmış ve ünitelerin içeriğinin öğrenci ilgi ve ihtiyaçlarına uygun olduğu fakat ünite sayısı ve kazanımlarının yoğunluğunun azaltılması gerektiğine dikkat çekilmiştir. Benzer bir şekilde Dilekli (2018), İngilizce Ağırlıklı Öğretim Programıyla ilgili öğretmen görüşlerini araştırdığı çalışmada, katılımcıların kazanımların yoğunluğuna vurgu yaptıkları dile getirilmiştir. Nitekim bu araştırma da öğretmenlerin kazanımların sadeleştirilmesine yönelik görüşlerine rastlanılmaktadır. Parmaksız ve İncirci'nin (2016) çalışmasında ortaya çıkan ve bu araştırmanın da bulgularıyla örtüşen ortak görüş kazanım sayısının fazla ve öğrencilerin bilişsel hazırbulunuşluklarının üzerinde olduğu yönündedir. Diğer taraftan Seçkin'in (2011) yaptığı çalışmada öğretmenlerin büyük bir kısmının kazanımların sadeleştirilmesi yönünde görüş bildirdiği görülürken, konuşma ve dinleme becerileriyle ilgili kazanımların arttırılması gerektiğine dair verilere de rastlanılmaktadır. Güneş (2009) ise konuşma ve dinleme becerileriyle ilgili kazanımların ölçme ve değerlendirme için yol gösterici olmadığı sonucuna ulaşmıştır. Çelebi ve Narinalp'e (2020) göre, İngilizce ders saatleri konusunda düzenlemeler yapılmasıyla dil beceri alanlarındaki kazanımların gerçekleştirilmesi için öğretmenlere yeterli zaman sağlanabilir. Bu araştırmaya katılan öğretmenlerin de ders saatlerinin öğretim faaliyetlerini ve sınama durumlarını gözeterek uygun bir şekilde arttırılması gerektiğine vurgu yaptıkları görülmektedir.

Öğretim programıyla ilgili katılımcıların belirttiği görüşler arasında öne çıkan bulgulardan biri de iletişime dayalı daha fazla test örneklerinin sunulması önerisidir. Tokur-Üner ve Aşılıoğlu'nun (2022) İngilizce öğretiminde ölçme değerlendirme sürecine ilişkin öğretmen görüşlerini inceledikleri çalışmada benzer bir bulguya ulaşılmış ve programda yer alan ölçme değerlendirme araçlarına ilişkin örneklerin yeterince açıklayıcı olmamakla birlikte nicelik olarak da yeterli olmadığı ifade edilmiştir. Bu araştırma sonuçlarıyla tutarlı sonuçlara ulaşan Cihan ve Gürlen (2013), iletişimsel test örneklerinin azlığı sebebiyle öğretmenlerin ölçme aracı hazırlama ve uygulamada zorluk yaşadıklarını

kaydetmişlerdir. Bu yüzden Büyükduman (2005) öğretmenlerin programdan daha fazla yararlanmaları için her ünite sonunda örnek testlerin bulunmasını önermiştir.

Araştırmada katılımcıların öğretim materyalleri ile ilgili sundukları öneriler incelediğinde ise, öğretimin etkileşimsel bir süreç olarak gerçekleşmesine katkı sunacak teknolojik ve yazılı-işitsel materyallerin kullanımının yaygınlaştırılması gerektiği belirtilmiştir. Ayrıca ders kitaplarının öğrencilerin ilgi ve motivasyonunu arttıracak şekilde yeniden düzenlenmesi önerilen görüşler arasındadır. Ayrıca katılımcılardan birinin belirttiği "…*Ancak devlet okullarında ek kaynak kitapların kullanımına izin verilmiyor"(Ö40; İHO5)* şeklinde ortaya çıkan ifadelerin özel okullara ders kitaplarının seçimi ve kullanımı ile ilgili bir üstünlük sağladığı söylenebilir. Diğer taraftan kamuya bağlı ortaokul ve İmam Hatip Ortaokullarında öğretmenlerin Milli Eğitim Bakanlığınca sunulan kitapların dışında farklı kaynaklara başvurmaları tercihleri dışında bırakılmıştır. Kararmaz ve Aslan'ın (2014) çalışmasında özel okullarda dil becerilerini geliştirmeye yönelik etkinlik içeren ders kitaplarının kullanılması ve yeterli fiziki ve teknolojik donanımın sağlanmasından dolayı, özel okullarda çalışan öğretmenlerin devlet okullarında çalışan öğretmenlere göre kendilerini mesleki alan bilgisi konusunda daha yeterli algıladıklarına dair ortaya çıkan sonuçlar bu tezi destekler niteliktedir. Bu durumun yarattığı farklılaşmaya dikkat çeken Yaman (2018), yerli kitap kullanımının oldukça kıymetli olduğunu belirtse de özellikle farklı öğretim kurumlarında görev alan öğretmenler arasında güven ortamının oluşabilmesi için tecrübeli yayınevleriyle işbirliği kurularak Türkiye'ye özgü İngilizce kitapların hazırlanabileceğini önermiştir.

6. SONUÇ VE ÖNERİLER

6.1. Sonuçlar

Araştırmanın bu bölümünde nitel ve nicel veriler göz önünde bulundurularak ulaşılan sonuçlara ve bu sonuçlarla ilgili araştırmacılara, program uygulayıcılarına ve diğer eğitim paydaşlarına öneriler sunulmuştur.

- Araştırmaya katılan tüm İngilizce öğretmenlerinin bir ölçme aracı hazırlarken planlama yapmaya özen gösterdikleri kaydedilmiştir. Planlama süreciyle ilgili ortaya çıkan alt temalar arasında, öğretmenlerin göz önünde bulundurdukları en önemli ölçüt hazırlanan ölçme aracının sınıf düzeyine uygun olmasıdır.
- Öğretmenlerin çoğu bir ölçme aracı hazırlarken geçerlik, güvenirlik ve kullanışlık özelliklerinin sağlanmış olmasına dikkat ettiklerini belirtmişlerdir. Ancak sınıfların kalabalıklığı ve öğrenci seviyelerindeki farklılıklardan dolayı bu özelliklere uygun ölçme aracı hazırlayamadıklarını belirten görüşler de bildirilmiştir (n=3). Geçerlik şartını sağlamak için öğretmenlerin daha çok kazanımlara uygun olan davranışları seçtikleri belirlenirken; güvenirliği sağlamak için ise, sınav süresinin yeterli olması ve testin orta güçlükte olmasına dikkat ettikleri kaydedilmiştir. Öğretmenlerin puanlama kolaylığı sağlayan ölçme araçları hazırlamaları da testin kullanışlı olması için en sık başvurulan yöntemdir.
- Öğretmenlerin çoğu hazırladıkları ölçme araçlarının İngilizce öğretim programıyla uyumlu olmasına kısmen dikkat ettiklerini ifade etmişlerdir. Öğretim programının tüm boyutlarıyla tamamen uyumlu olacak şekilde ölçme aracı hazırladıklarını belirten herhangi bir görüş olmamıştır. Öğretmenlerin öğretim programına kısmen bağlı kalmalarının en önemli sebebi ise öğrenci seviyelerinde yaşanan farklılıklardır. Ayrıca çeşitli fiziksel ve teknolojik kısıtlamaların öğretmenleri ölçme aracı hazırlama konusunda sınırlandırdığı kaydedilmiştir.
- Araştırmanın yabancı dil beceri alanlarının değerlendirilmesiyle ilgili temasında, öğretmenlerin dinleme ve konuşma becerilerinin değerlendirilmesine yönelik olumsuz görüşlerinin olumlu görüşlerden daha fazla olduğu sonucuna ulaşılmıştır. Bu beceri alanlarının değerlendirilmesine yönelik faaliyetlerin okulların fiziki ve teknolojik yapısındaki yetersizlikler ve ders saatlerinin az olması gibi öne çıkan nedenlerden dolayı ihmal edildiği ve ölçülemediği öğretmenlerin vurguladıkları görüşlerdendir. Dinleme ve konuşma alanlarında kaydedilen olumsuzlukların

aksine, öğretmenlerin okuma becerilerinin değerlendirilmesine yönelik görüşleri çoğunlukla olumludur. Yazma alanında ise herhangi bir sıkıntı yaşanmadığına dair görüşün en yüksek frekansa sahip olduğu görülmektedir.

- Yabancı dil beceri alanlarıyla ilgili nicel veriler incelendiğinde ise, öğretmenlerin dinleme ve konuşma becerilerini sınıf içi etkinlikler sırasında değerlendirdiklerine dair görüşlerin ortalamasının "Çoğu zaman" düzeyinde olduğu görülmektedir. Öğretmenlerin dinleme ve konuşma becerilerinin ölçme ve değerlendirme araçlarıyla sınanıp sınanmadığını sorgulayan anketin "İngilizce dersi öğretim programındaki kazanımlara uygun bir ölçme aracı hazırlarım" ve "Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım" maddelerini ise "Ara sıra" düzeyinde gerçekleştirdikleri kaydedilmiştir. Diğer taraftan öğretmenlerin okuma becerilerinin değerlendirilmesine yönelik sergiledikleri davranışların "Her zaman" ya da "Çoğu zaman" düzeyinde olduğu saptanmıştır. Yazma becerilerinin ne düzeyde değerlendirildiğine ilişkin nicel veriler incelendiğinde ise, öğretmenlerin bu alanın değerlendirilmesine yönelik sergiledikleri davranışların çoğunlukla "Çoğu zaman" düzeyinde olduğu belirlenmiştir.
- Öğretmenlerin İngilizce dil becerilerinin ölçülmesinde klasik ölçme araçları arasında en fazla çoktan seçmeli testleri tercih ettikleri saptanmıştır (f=23). Sınıfların kalabalık olması gibi fiziki şartların yetersizliğinden dolayı, test uygulayıcıları öncelikli olarak puanlanması ve uygulaması hızlı olan testlerin kullanımına yönelmişlerdir.
- Öğretmenlerin tamamlayıcı ölçme ve değerlendirme araçları arasında en fazla proje ve performans görevleri ile ürün dosyasını (portfolyo) tercih ettikleri görülmektedir. Öz değerlendirme ve akran değerlendirme tekniklerinin ise daha az kullanıldığı kaydedilmiştir. Ayrıca sınıf mevcutlarının kalabalıklığı ve öğrenciler için zorlayıcı olmasından dolayı tamamlayıcı araçları hiç kullanmadıklarını belirten görüşler de bildirilmiştir (n=4)
- Öğretmenler en fazla sözcük bilgisinin değerlendirilmesine yönelik uygulamalar yapmaktadır. Dilbilgisi yapılarının gerekliliğini savunan ve test ettiğini ifade eden öğretmenlerin sayısı da oldukça fazladır (f=23). Telaffuz becerisinin ölçülmesine ilişkin görüşler arasında ise herhangi bir ölçme yapmadığını belirten görüşlerin ön plana çıktığı tespit edilmiştir (f=25).

- Öğretmenlerin büyük bir çoğunluğunun duyuşsal becerilerin ölçülmesiyle ilgili olumlu görüş bildirdikleri kaydedilmiştir. Öğretmenlere göre duyuşsal becerilerin dikkate alınmasının en önemli gerekçesi öğrencilerin ilgi ve tutumlarını ortaya çıkarmasıdır.
- Öğretmenlerin tümü bir tercih olmaktan ziyade zorunlu olduğu için düzey belirleyici değerlendirme yöntemini kullandıklarını ifade etmişlerdir. Düzey belirleyici değerlendirmeden sonra en fazla tercih edilen yöntemin biçimlendirici değerlendirme olduğu göze çarpmaktadır. Tanılayıcı değerlendirmenin ise diğer değerlendirme türlerine göre nispeten daha az tercih edildiği saptanmıştır.
- Öğretmenlerin çoğu öğrencilere ölçme sonuçları hakkında geribildirimde bulunmayı oldukça önemsemektedir (f=30).
- Öğretmenler nitelikli bir değerlendirme süreci için, ölçme değerlendirme araç ve yöntemleri, öğrenme ortamı, öğretim programı ve öğretim materyalleri ile ilgili çeşitli öneriler sunmuşlardır.

6.2. Öneriler

- Öğretim programında belirtilen ölçme değerlendirme uygulamalarının tam olarak gerçekleştirilebilmesi için okulların fiziki ve teknolojik donanımlarının iyi bir şekilde yapılandırılması gerekmektedir.
- Araştırmaya katılan öğretmenler kazanımların yoğun ve öğrenci seviyesinin üzerinde olduğunu belirtmişlerdir. Bu yüzden kazanımların öğrenci giriş davranışları ve okullar arasındaki sosyo-ekonomik ve kültürel farklılıklar göz önünde bulundurularak sadeleştirilmesi sağlanabilir.
- Ders saatlerinin az olmasının öğretmenleri ölçme değerlendirme sürecinde kısıtladığı ve bazı becerilerin tam olarak değerlendirilmesini engellediği ifade edilmiştir. Bu durumda İngilizce ders saatlerinin arttırılması gerekmektedir.
- Ders kitapları öğrenci ilgisini çekecek şekilde görsel-işitsel materyallerle zenginleştirilmeli ve etkileşimli tahtalarla uyumlu hale getirilmelidir. Ayrıca ders kitaplarındaki her ünitenin video veya sunumlarla desteklenmesi için gerekli materyaller sağlanmalıdır.
- Araştırma sonuçları öğretmenlerin çoğunlukla puanlama kolaylığından dolayı çoktan seçmeli testleri tercih ettiklerini göstermiştir. Bu noktada öğretmenlere

ölçme aracı hazırlama ve uygulama süreçleriyle ilgili konu alanı uzmanları tarafından eğitimler verilebilir.

- Araştırmada konuşma ve dinleme alanlarının en fazla ihmal edilen beceriler olduğu tespit edilmiştir. Sınav yönetmeliğinde sınavların tüm beceri alanlarını dahil edecek şekilde yeniden düzenlemesi ve bu konuda ülke genelinde büyük ölçekli kararların alınması önerilebilir.
- Öğretmenlere ölçme değerlendirmeyle ilgili hem kuramsal hem de uygulamalı faaliyetleri içeren hizmet içi eğitimler sunulmalıdır.
- Öğretmenlerin dil öğretiminde değerlendirme okuryazarlıklarının ne düzeyde olduğu araştırılabilir
- Yabancı dilde ölçme değerlendirme sürecinde öğrenci, ebeveyn ve yöneticiler gibi diğer paydaşların algı ve tutumları incelenebilir.
- Sözcük bilgisi, telaffuz gibi dil yapılarının nasıl değerlendirildiğine ilişkin süreçler çeşitli değişkenlerle geniş ölçekte incelenebilir.
- Yapılan bu çalışma öğretmenlerin ölçme değerlendirmeyle ilgili görüşlerini beyan etmeleri ile gerçekleşmiştir. Öğretmenlerin uygulamada neler yaptıklarını gözlemleyen çalışmalar yapılabilir.

KAYNAKÇA

Acharya, B. (2010, June). Questionnaire design. Workshop in Research Methodology organised by Centre for Post Graduate Studies, University Grant Commission Nepal, Pulchok

Adetayo, J. O. (2014). Assessing the affective behaviors in learners. *Journal of Education and Practice*, 5(4), 8-15

Adıgüzel, O. C. (2016). E*ğitim programlarının geliştirilmesinde ihtiyaç analizi el kitabı*. Ankara: Anı Yayıncılık

Aebersold, J. A. & Field, M. L. (1997). *From reader to reading teacher: Issues and strategies for second language classrooms*. Cambridge: Cambridge University Press.

Agung, H. (2012). The Implementation of Cards in Teaching Speaking at Eighth Grade of MTs Negeri Bandung in Academic Year 2011/2012. http://repo.uinsatu.ac.id/1211/ adresinden 16.05.2021 tarihinde ulaşılmıştır.

Akarsu. B. (2018). *Ölçme ve Değerlendirme*. İstanbul: Cinius Yayınları

Akbulut, E. O. & Akbulut, K. (2011). Science and technology teacher candidates' opinions regarding alternative assessment. *Procedia Social and Behavioral Sciences*, 15, 3531-3535

Akman-Yeşilel, D. A. (2015). İşitsel-Dilsel Yöntem (Audi-Lingual Method). N. Bekleyen (Ed.), *Dil Öğretimi* içinde (s. 60-78). Ankara: Pegem Akademi.

Aküzel, G. (2006). İlköğretim 4-8. sınıflarda yabancı dil öğretimindeki başarısızlık nedenlerinin incelenmesi (Adana Örneği). Yayınlanmamış Yüksek Lisans Tezi, Çukurova Üniversitesi Sosyal Bilimler Enstitüsü, Adana

Al-Amri, M. N. (2010). Direct spoken English testing is still a real challenge to be worth bothering about. *English Language Teaching*, *3*(1), 113-117.

Alderson, J. C. & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language teaching*, 35(2), 79-113.

Alduais, A. M. S. (2012). An account of approaches to language testing. *International journal of Academic research in progressive education and development*, 1(4), 203-208.

Alkan, V., Şimşek, S. & Erbil, B. A. (2019). Karma yöntem deseni: Öyküleyici alanyazın incelemesi. *Eğitimde Nitel Araştırmalar Dergisi*, *7*(2), 559-582.

Altan, M. Z. (2017). *Hayallerindeki Eğitim Faküktesi*. Ankara: Pegem Akademi

Altmışdört, G. (2010). Yabancı dil öğretiminde nasıl bir ölçme-değerlendirmeye gerek vardır. *Kurumsal Eğitimbilim*, 3(2), 175-200.

Altunöz, H. B. & Erdal, B. A. Y. (2018). Yabancı dil olarak ingilizce öğretiminde etkili öğretim stratejilerinin kullanımının analizi. *Uluslararası Türk Kültür Coğrafyasında Sosyal Bilimler Dergisi*, 3(1), 129-149.

Anderson, R. S. (1998). Why talk about different ways to grade? The shift from traditional assessment to alternative assessment. *New directions for Teaching and Learning*, 74, 5-16.

Anderson, R. C. & Pearson, P. D. (1988). A schema-theoretic view of basic processes in reading comprehension. *Interactive approaches to second language reading*, 37-55.

Angelo, T. A. & Cross, K. P. (1993). Minute paper. *Classroom assessment techniques: A handbook for college teachers*, 148-153.

Antoni, N. (2010). Exploring EFL teachers' strategies in teaching reading comprehension. *Jurnal penelitian pendidikan*, 11(2), 39-51.

Arı, A. (2014). İlköğretim altıncı sınıf İngilizce dersi öğretim programına ilişkin öğretmen görüşleri. *Journal of Theoretical Educational Science/Kuramsal Eğitimbilim Dergisi*, 7(1), 172-194.

Arnold, J. (2011). Attention to affect in language learning. *Online Submission*, 22(1), 11-22.

Arnold, J. & Brown, H. (1999). A map of the terrain. J. Arnold (Ed.), In *Affect in language learning* (pp. 1-24). Cambridge: Cambridge University Press.

Atılgan, H. (Ed.). (2018). *Eğitimde ölçme ve değerlendirme*. Ankara: Anı Yayıncılık

Atta-Alla, M. (2012). Investigating alternative assessment practices in pre-service English teacher education programs. *International Journal of English Language Education*, 1(1).

Ayas, N. (2014). An investigation of teachers' conception and practices concerning assessment in English for young learners classroom. Unpublished Master Thesis, Çağ University, Institute of Social Sciences, Mersin.

Aydın, M. (2010). *Eğitim Yönetimi.* Ankara: Hatipoğlu Yayımcılık

Baker, A. (2013). Integrating fluent pronunciation use into content-based ESL instruction: Two case studies. In J. Levis, & K. LeVelle (Eds.), Proceedings of the 4th pronunciation in second language learning and teaching conference (pp. 245–254). Ames, IA: Iowa State University

Balcı, A. (2015). *Sosyal bilimlerde araştırma yöntem teknik ve ilkeler* (11. Baskı). Ankara: Pegem Akademi Yayınları.

Ballıdağ, S. & Karagül, B. I. (2021). Exploring the language assessment literacy of Turkish in-service EFL teachers. *Balıkesir Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 24(45), 73-92.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford: Oxford University Press.

Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice: Designing and developing useful language tests*. Oxford: Oxford University Press.

Bachman, L. F., Palmer, A. S. & Palmer, A. S. (2010). L*anguage assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford: Oxford University Press.

Backlund, P., Brown, K., Gurry, J. & Jandt, F. (1980). Evaluating Speaking and Listening Skill Assessment Instruments: Which One Is Best for You?. *Language Arts*, 57(6), 621-627.

Bacon, D. R. (2003). Assessing learning outcomes: A comparison of multiple-choice and short-answer questions in a marketing context. *Journal of Marketing Education*, 25(1), 31-36.

Badger, R. & White, G. (2000). A process genre approach to teaching writing. *ELT Journal, 54*(2), 153-160.

Bailey, K. M. (1998). *Learning about language assessment: dilemmas, decisions, and directions*. US: Heinle & Heinle

Baleghizadeh, S. & Rahimi, A. H. (2011). The relationship among listening performance, metacognitive strategy use and motivation from a self-determination theory perspective. *Theory and Practice in Language Studies*, 1(1), 61-67.

Balım, D. (2020). Beşinci sınıf İngilizce ağırlıklı öğretim programının aydınlatıcı değerlendirme modeliyle değerlendirilmesi. Yayınlanmamış yüksek lisans tezi, Pamukkale Üniversitesi, Eğitim Bilimleri Enstitüsü, Denizli.

Ballantyne, R., Borthwick, J. & Packer, J. (2000). Beyond student evaluation of teaching: Identifying and addressing academic staff development needs. *Assessment & Evaluation in Higher Education*, 25(3), 221-236.

Baltacı, A. (2019). Nitel araştırma süreci: Nitel bir araştırma nasıl yapılır?. *Ahi Evran Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 5(2), 368-388.

Baran, H. (2020). Açık ve uzaktan eğitimde ölçme ve değerlendirme. *Açıköğretim Uygulamaları ve Araştırmaları Dergisi, AUAd*, 6(1), 28-40.

Barkaoui, K. (2010). Explaining ESL essay holistic scores: A multilevel modeling approach. *Language Testing*, 27(4), 515-535.

Boscolo, P. (2009). Engaging and motivating children to write. R. Beard et al. (Eds.), In *The SAGE handbook of writing development* (pp. 300–312). Los Angeles: SAGE. https://books.google.com.tr/book adresinden 05.04.2021 tarihinde ulaşılmıştır

Başaran, M., Özdemir, O. İ. & Can, M. S. (2020). 8. sınıf İngilizce öğretim programının bağlam, girdi, süreç ve ürün (CIPP) modeline göre değerlendirilmesi. *e-Uluslararası Eğitim Araştırmaları Dergisi*, *11*(3), 154-178.

Baştürk, S. ve Taştepe, M. (2013). *Bilimsel Araştırma Yöntemleri.* Ankara: Vize Yayıncılık.

Bauer, B. A. (1981). *A study of the reliabilities and cost efficiencies of three methods of assessments for writing ability.* Champaign: University of Illinois.

Baykul, Y. (1992). Eğitim sisteminde değerlendirme. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi,* 7, 85-94.

Bayram, E. (2015). An investigation of language assessment preferences of pre-service EFL teachers and underlying factors. Unpublished Master Thesis, Karadeniz Teknik Üniversitesi, Trabzon

Becker, H. S. (1991). Generalizing from case studies. E. W. Eisner & A. Peshkin (Eds.), In *Qualitative inquiry in education: The continuing debate* (pp. 233–242). New York: Teachers College Press.

Becker, A. (2006). A review of writing model research based on cognitive processes. *Revision: History, theory, and practice*, 25-49.

Bekleyen, N. (2015). Dil öğretiminin tarihsel gelişimi. N. Bekleyen (Ed.), *Dil Öğretimi* içinde (s. 4-20). Ankara: Pegem Akademi.

Bereiter, C. & Scardamalia, M. (1987). An attainable version of high literacy: Approaches to teaching higher-order skills in reading and writing. *Curriculum inquiry*, 17(1), 9-30.

Berkant, H.G., Özaslan, D. ve Doğan, E. (2019). Yabancı dil ağırlıklı eğitim uygulamasına ilişkin İngilizce öğretmenlerinin görüşleri. *Milli Eğitim,* 48(1), 553-570.

Bernhardt, E. B. (1993). *Reading development in a second language: Theoretical, empirical, & classroom perspectives*. Norwood, NJ: Ablex Publishing Corporation

Bernhardt, E. B. & James, C. J. (1987). The teaching and testing of comprehension in foreign language learning. D.W. Birckbichler (Ed.), In *Foreign languages: proficiency, policy and professionalism* (pp. 65-81). Lincolnwood, IL: National Textbook.

Bernhardt, E., Molitoris, J., Romeo, K., Lin, N. & Valderrama, P. (2015). Designing and sustaining a foreign language writing proficiency assessment program at the postsecondary level. *Foreign Language Annals*, 48(3), 329-349.

Bible, L., Simkin, M. G. & Kuechler, W. L. (2008). Using multiple-choice tests to evaluate students' understanding of accounting. *Accounting Education: An International Journal*, 17(1), 55-68.

Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment In Education,* 5(1), 7 -74

Black, P. & Wiliam, D. (2010). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 92(1), 81 – 90

Bond, L. A. (1996). Norm-and criterion-referenced testing. *Practical Assessment, Research and Evaluation*, 5(1), 2.

Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies in continuing education*, 22(2), 151-167.

Brindley, G. (2001). Outcomes-based assessment in practice: some examples and emerging insights. *Language Testing*, 18(4), 393-407.

Brissenden, G., Slater, T. F. & Mathieu, R. D. (2002). The role of assessment in the development of the college introductory astronomy course. *Astronomy Education Review*, *1*(1).

Britton, A. M. (2015). Assessment for learning in teaching English to young learners: Teachers' understanding, classroom practice and impact on interaction. Unpublished Doctoral Dissertation, University of Reading, UK

Broadfoot, P. M. (2005). Dark alleys and blind bends: Testing the language of learning. *Language Testing*, *22*(2), 123-141.

Brown, D. C. (2001). Cognitive and Affective Minds: A Necessary Marriage in Criminal Justice Education, *Journal of Criminal Justice Education,* 12(1), 101-115.

Brown, H.D. (2004). *Language assessment: Principles and classroom practices*. The USA: Pearson Education

Brown, H. D. & Abeywickrama, P. (2010). *Language Assessment: Principles and Classroom Practices.* White Plains, NY: Pearson Education.

Brown, J. D. (1996). *Testing in Language Programs.* New Jersey: Prentice Hall Regents

Brown, J. D. & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-675.

Bryman, A. (2006). Integrating quantitative and qualitetive research: How is it done? *Qualitative Research*. 6 (1), 97-113.

Byrne, J. & Humble, Á. M. (2007). An introduction to mixed method research. *Atlantic Research Centre For Family-Work Issues*, 1, 1-4.

Buck, G. (1988). Testing listening comprehension in Japanese university entrance examinations. *JALT Journal*, *10*(1), 15-42.

Buhagiar, M. A. (2007). Classroom assessment within the alternative assessment paradigm: Revisiting the territory. *The Curriculum Journal*, 18(1), 39-56.

Buldur, S. & Hasbek, G. (2020). Öğretmen Adaylarının Biçimlendirici Değerlendirmeye İlişkin Algıları: Metaforik Bir İnceleme. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 21(2), 595-615.

Burns, A. & Joyce, H. (1997). *Focus on Speaking*. Sydney: National Centre for English Language Teaching and Research.

Burton, K. (2006). Designing criterion-referenced assessment. *Journal Of Learning Design*, 1(2), 73-82.

Büyükduman, F. İ. (2005). İlköğretim okulları İngilizce öğretmenlerinin birinci kademe İngilizce öğretim programına ilişkin görüşleri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *28*(28), 55-64.

Büyüköztürk, Ş. (2005). Anket geliştirme. *Türk Eğitim Bilimleri Dergisi, 3*(2), 133-151.

Büyüköztürk, B., Çakmak, E., Akgün, Ö.E., Karadeniz, Ş. & Demirel, F. (2013). *Bilimsel Araştırma Yöntemleri.* Ankara: Pegem Akademi

Bygate, M. (2009). Teaching and testing speaking, M. H. Long & C. J. Doughty (Eds.). In *The handbook of language teaching* (pp. 412-440). UK: Wiley-Blackwell.

Cain, K. & Oakhill, J. (2006). Assessment matters: Issues in the measurement of reading comprehension. *British Journal of Educational Psychology*, 76(4), 697-708.

Caracelli, V. J. & Greene, J. C. (1993). Data analysis strategies for mixed-method evaluation designs. *Educational evaluation and policy analysis*, 15(2), 195-207.

Carrell, P. L., Devine, J. & Eskey, D. E. (1988). *Interactive approaches to second language reading*. New York: Cambridge University Press.

Carrell, P. L. & Eisterhold, J. C. (1983). Schema theory and ESL reading pedagogy. *TESOL Quarterly,* 17(4), 553-573.

Cesur, K. (2008). Students' And Teachers' Perceptions Of The Test Techniques Used To Assess Language Skills At Unıversity Level. Yayınlanmamış Yüksek Lisans Tezi, Çanakkale 18 Mart Üniversitesi, Çanakkale

Chan, Y. (2016). Elementary school EFL teachers' beliefs and practices of multiple assessments. *Reflections on English Language Teaching,* 7(1), 37-62.

Chastain, K. (1988). Affective and ability factors in second langugae learning. *Language Learning,* 25, 153-161

Chaudron, C. (1988). Validation in second language classroom research: The role of observation. *University of Hawaii Working Papers in English as a Second Language* 7(2), 1-15

Chomsky, N. (2001). *Dil ve zihin.* (Çev. A. Kocaman). Ankara: Ayraç Yayınevi

Christensen L. B., Johnson R. B., & Turner, L. A. (2015) *Araştırma Yöntemleri: Desen ve Analiz* (Çev. A. Alpay). Ankara: Anı Yayıncılık.

Cihan, T. ve Gürlen, E. (2013). İlköğretim 5. sınıf İngilizce dersi öğretim programına ilişkin öğretmen görüşleri. *Anadolu University Journal of Social Sciences*, *13*(1), 131-146.

Clapham, C. (2000). Assessment and testing. *Annual Review of Applied Linguistics*, 20, 147-161.

Clark, J. L. (1983). Language testing: Past and current status-directions for the future. *Modern Language Journal*, 431-443.

Clark, S. & Gipps, C. (2000). The role of teachers in teacher assessment in England 1996-1998. *Evaluation and Research in Education, 14*, 38-52.

Codo, E. (2008). Interviews and questionnaires. *The Blackwell guide to research methods in bilingualism and multilingualism*, 158-176.

Cohen, A. D. (1994). A*ssessing language ability in the classroom*. Boston: Heinle & Heinle

Connor, U. & Mbaye, A. (2002). Discourse approaches to writing assessment. *Annual Review of Applied Linguistics,* 22, 263-278

Coombe, C. (2010). Assessing foreign/ second language writing ability. *Education Business and Society: Contemporary Middle Eastern Issues*, 3(3), 178-187.

Cooper, P. L. (1984, May). The assessment of writing ability: A review of research. ETS Research Report Series, Educational Testing Service, Princeton.

Cox, R. C. & Vargas, J. S. A. (1966, April) Comparison of item selection techniques for norm-referenced and criterion-referenced tests. National Council on Measurement in Education, Chicago.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* . Thousand Oaks, CA: Sage

Creswell, J. W. (2015). *Nitel Araştırma Yöntemleri* (2.Baskı) M. Bütün ve S. B. Demir (Çev. Edt.). Ankara: Siyasal Kitabevi

Creswell, J. W., Fetters, M. D. & Ivankova, N. V. (2004). Designing a mixed methods study in primary care. *The Annals of Family Medicine*, *2*(1), 7-12.

Creswell, J. W. & Miller, D. L. (2000). Determining validity in qualitative inquiry. *Theory into practice*, 39(3), 124-130.

Creswell, J.W. & Plano Clark, V. L. (2015). *Karma Araştırma Yöntemleri: Tasarımı ve Yürütülmesi.* Y. Dede ve S. B. Demir (Çev. Edt.). Ankara: Anı Yayıncılık

Crooks, T. J. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research,* 52(4), 438-481.

Cross, D. R. & Paris, S. G. (1987). Assessment of reading comprehension: Matching test purposes and test properties. *Educational Psychologist*, 22(3-4), 313-332.

Çağlar, M. & Kılıç, A. (2019). Merkezi sınav ve öğretmen yapımı sınavların bazı değişkenler açısından incelenmesi: Ortaöğretime geçiş sınavı örneği. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 19(4), 1288-1305.

Çakır, İ. (2015). Okuma becerisinin öğretimi. N. Bekleyen (Ed.), *Dil Öğretimi* içinde (s. 266-283). Ankara: Pegem Akademi.

Çakır, İ. (2017). The washback effects of secondary education placement examination on teachers, school administrators and parents with specific reference to teaching English as a foreign language. *Turkish Journal of Teacher Education*, 6(2), 61-73.

Çakır, N. C. & Genç, Z. S. (2022). A comparative analysis of teachers' beliefs about the assessment of 4th grade-efl students in Turkey, Italy and Finland. *European Journal of Foreign Language Teaching*, 5(6), 67-98

Çelebi, M. & Kuşuçuran, B. N. (2019). İngilizce öğretmenlerinin alternatif değerlendirme yöntemlerini kullanma yeterliliklerinin incelenmesi. *OPUS Uluslararası Toplum Araştırmaları Dergisi*, *13*(19), 983-1009.

Çelebi, M. ve Narinalp, N. Y. (2020). Ortaokullarda İngilizce öğretiminde karşılaşılan sorunlar. *OPUS Uluslararası Toplum Araştırmaları Dergisi*, 15(1), 4975-5005.

Çelikkaya, T. (2008). Yapılandırmacı yaklaşımın sosyal bilgiler öğretiminde başarı, tutum ve kalıcılığa etkisi (5.sınıf örneği). Yayımlanmamış Doktora Tezi, Atatürk Üniversitesi, Sosyal Bilimler Enstitüsü, Erzurum.

Çepni, S., Baki, A., Ayas, A., Demircioğlu, G. ve Akyıldız, S. (2009). *Ölçme ve değerlendirme.* Trabzon: Celepler Matbaacılık.

Çetin, M. (2003). Örgüt kuramları perspektifinden halkla ilişkilerin gelişimi. *İletişim Dergisi*, *18*, 31-58.

Çetin, P. (2011). İlköğretimde hayat bilgisi dersinde probleme dayalı öğrenme yöntemi uygulamalarının öğrencilerin öğrenme ürünlerine etkisi. Yayınlanmamış yüksek lisans tezi, Dokuz Eylül Üniversitesi, Eğitim Bilimleri Enstitüsü, İzmir.

Çiçen, M. (2014). The assessment methods used by EFL teachers at the primary level in private schools . Unpublished Master Thesis, Çağ University, Institute of Social Sciences, Mersin.

Dağtan, E. (2015). Konuşma becerisinin öğretimi. N. Bekleyen (Ed.), *Dil Öğretimi* içinde (s. 334-349). Ankara: Pegem Akademi.

Davies, A. (1968) *Language testing symposium: A psycholinguistic approach*. Oxford University Press

Davies, A. (1978). Language testing. *Language Teaching*, *11*(3), 145-159.

Davies, A. (1990). *Principles of Language Testing.* Oxford: Basil Blackwell

Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T. & McNamara, T. (1999). *Dictionary of Language Testing* (Vol. 7). Cambridge University Press.

De Haan, P. & Van Esch, K. (2004). Towards an instrument for the assessment of the development of writing skills. U. Connor & T. Upton (Eds.). In *Applied Corpus Linguistics* (pp. 267-279). New York: Rodopi

Decker, M. A. (2004). Incorporating guided self-study listening into the language curriculum. *Language Teacher-KYOTO-JALT-*, *28*(6), 5-10.

Delice, A. (2015). Karma yöntem desen seçimi. Y. Dede ve S.B. Demir (Çev. Ed.). *Karma Yöntem Araştırmaları* içinde (s. 61-114). Ankara: Anı Yayıncılık.

Demirel, Ö. (1992). Türkiye'de Program Geliştirme Uygulamaları. *Hacettepe Üniversitesi Eğitim Fakültesi Eğitim Fakültesi Dergisi* (7), 27-43.

Demirel, Ö. (2003). *Yabancı dil öğretimi*. Ankara: Pegem Yayıncılık.

Demirel, Ö. (2012). *Eğitim bilimleri sözlüğü* (5. Baskı). Ankara: Pegem Akademi

Demirtaş, Z., & Erdem, S. (2015). 5. sınıf İngilizce dersi öğretim programı: Güncellenen programın bir önceki programla karşılaştırılması ve programa ilişkin öğretmen görüşleri. *Sakarya University Journal of Education*, 5(2), 55-80.

Derwing, T. M. & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39(3), 379-397.

Dick, W., Carey, L. & Carey, J. O. (2001). *The systematic design of instruction*. New York: Longmann.

Dickins, R. & Germaine, K. (1995). Evaluation. *ELT Journal*. 49(2), 194-195.

Dilekli, Y. (2018). Ortaokul İngilizce hazırlık sınıfı programı pilot uygulamasının öğretmen görüşlerine göre değerlendirilmesi. OPUS *Uluslararası Toplum Araştırmaları Dergisi, 8*(15), 1401-1425.

Dinçer, B. (2013). 7. sınıf ingilizce öğretim programının Stufflebeam'in bağlam-girdi-süreç-ürün (CIPP) modeline göre değerlendirilmesi. Yayımlanmamış Doktora Tezi, Adnan Menderes Üniversitesi, Aydın

Dixson, D. D. & Worrell, F. C. (2016). Formative and summative assessment in the classroom. *Theory Into Practice*, 55(2), 153-159.

Djigunović, J. M. (2006). Language anxiety and language processing. *EUROSLA Yearbook*, 6(1), 191-212.

Doğan, İ. (2009). İngilizce öğretmenlerinin mesleki yeterlilik algıları ve İngilizce öğretiminde karşılaştıkları sorunlar. Yayınlanmamış yüksek lisans tezi, Yeditepe Üniversitesi, Sosyal Bilimler Enstitüsü, İstanbul

Doğan, C. D. & Uluman, M. (2017). A Comparison of Rubrics and Graded Category Rating Scales with Various Methods Regarding Raters' Reliability. *Educational Sciences: Theory and Practice, 17*(2), 631-651.

Duban, N. & Küçükyılmaz, E. A. (2008). Sınıf öğretmeni adaylarının alternatif ölçme-değerlendirme yöntem ve tekniklerinin uygulama okullarında kullanımına ilişkin görüşleri. *İlköğretim Online*, 7(3), 769-784.

Dubin, F. & Bycina, D. (1991). Academic reading and the ESL/EFL teacher. *Teaching English as a Second or Foreign Language*, 2, 195-215.

Duran, Ö. (2011). Teachers' and students' perceptions about classroom-based speaking tests and their washback. Unpublished doctoral dissertation, Bilkent University, Ankara

Dursun, G. (2014). An analysis of assessment and evaluation activities in the schools of foreign languages in Turkey. Unpublished master's thesis, Pamukkale University, Denizli

Earl, L. & Cousins, B. (1996). *Classroom assessment: Changing the face, facing the change.* Toronto: Ontario Teachers' Federation (OPSTF).

Edelenbos, P. & Kubanek-German, A. (2004). Teacher assessment: The concept of 'diagnostic competence'. *Language testing*, 21(3), 259-283.

Ekiz, D. (2009). *Bilimsel Araştırma Yöntemleri*. Ankara: Anı Yayıncılık

Erdem, S. & Yücel-Toy, B. (2017). Yabancı dil ağırlıklı beşinci sınıf ingilizce programına yönelik ihtiyaçların belirlenmesi. *Electronic Turkish Studies*, 12(28), 259-280

Ergüç, B. (2004). İlköğretim 4. ve 5. sınıflarda yabancı dil (İngilizce) öğretimi ile ilgili okul yönetiminin karşılaştığı sorunlar ve çözüm önerileri (Kırıkkale ili örneği). Yayımlanmamış Yüksek Lisans Tezi, Kırıkkale Üniversitesi, Kırıkkale

Ergüder, I. (2005). Anadolu Lisesi ve Süper Lise hazırlık sınıfı öğrencilerinin İngilizce öğreniminde karşılaştıkları sorunlar. Yayınlanmamış doktora tezi, Dokuz Eylül Üniversitesi, İzmir.

Erkuş, A. (2006). *Sınıf Öğretmenleri için Ölçme ve Değerlendirme: Kavramlar ve Uygulamalar*. Ankara: Ekinoks Yayınları.

Erlandson, D. A., Harris, E. L., Skipper, B. L. & Allen, S. D. (1993). *Doing naturalistic inquiry: A guide to methods*. London: Sage.

Ersen-Yanık, A. (2007). A study of english language curriculum implementation in 6th, 7th and 8th grades of public primary schools through teachers' and students' perceptions. Unpublished P.hD thesis, Middle East Technical University, Ankara.

Eskey, D. E. (1988). Holding in the bottom: An interactive approach to the language problems of second language readers. *Interactive approaches to second language reading*, 6, 93-100.

Ertürk, S. (2013). *Eğitimde Program Geliştirme* (6. Baskı). Ankara: Edge Akademi

Farhady, H. (2019). A Cross-contextual perspective on EFL teachers' assessment knowledge. *EDU7*, 8(10), 1-19.

Fidan, U. M. & Sak, I. M. (2012). İlköğretim öğretmenlerinin tamamlayıcı ölçme değerlendirme teknikleri hakkında görüşleri. *Bartın University Journal of Faculty of Education*, 1(1), 174-189.

Fillmore, C. J. (1979). On fluency. C. J. Fillmore, D. Kempler, & W. S. Y. Wang (Eds.), In *Individual differences in language ability and language behavior* (pp. 85–101). New York: Academic Press.

Fives, H. & DiDonato-Barnes, N. (2013). Classroom test construction: The power of a table of specifications. *Practical Assessment, Research, and Evaluation*, 18(1), 3.

Flower, L. & Hayes, J. R. (1981). A cognitive process theory of writing. *College composition and communication*, 32(4), 365-387.

Flowerdew, J. & Miller, L. (2005). *Second Language Listening: Theory and Practice.* New York: Cambridge University Press

Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9 (2), 113-132.

Fulcher, G. & Reiter, R. M. (2003). Task difficulty in speaking tests. *Language Testing*, *20*(3), 321-344.

Furwana, D. (2019). Validity and reliability of teacher-made English summative test at second grade of vocational high school 2 Palopo. *Language Circle: Journal of Language and Literature*, *13*(2), 113-124

Gampper, C. (2013). Improving English test qualities. *Thammasat Review*, *16*(3), 73-83.

Garcia, G. E. & Pearson, P. D. (1991). The role of assessment in a diverse society. E. H. Hiebert (Ed.), In *Literacy for a diverse society: Perspectives, practices, and policies* (pp. 253- 278). New York: Teachers College Press

Gardner, J. (2010). Developing teacher assessments: An introduction. In J. Gardner, W. Harlen, L. Hayward, G. Stobart, & M. Montgomery (Eds.), *Developing Teacher Assessment* (pp. 1 – 11). New York, NY: Open University Press

Gelbal, S. & Kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *33*(33), 135-145.

Genç, A. (2004). Türkiye'de ilk ve ortaöğretim okullarında yabancı dil öğretimi. *Bişkek, Kırgızistan-Manas Sosyal Bilimler Dergisi,* 10, 107-111.

Genesee, F. & Upshur, J. (1996). *Classroom-based Evaluation in Second Language Education*. Cambridge: Cambridge University Press.

Genesee, F. & Hamayan, E.V. (1994). Classroom-based assessment. F. Genesee (Ed.), In *Educating second language children* (pp. 212-239). Cambridge: Cambridge University Press.

Ghaicha, A., & Omarkaly, E. (2018). Alternative assessment in the Moroccan EFL classrooms teachers' conceptions and practices. *Higher Education of Social Science*, 14(1), 56-68.

Giri, R. A. (2010). Language testing: Then and now. *Journal of NELTA*, *8*(1), 49-67.

Goh, C. C. (2002). Exploring listening comprehension tactics and their interaction patterns. *System*, *30*(2), 185-206.

Gough, P. B. (1972). One second of reading. *Visible Language*, *6*(4), 291-320.

Göçer, A. (2019). Türkçe eğitiminde ölçme ve değerlendirmenin işlevselliğine etki eden önemli bir uygulama tekniği olarak geribildirim kullanımı. *Kırıkkale Üniversitesi Sosyal Bilimler Dergisi, 9*(1), 111-126.

Gök-Çatal, Ö. (2015). İlköğretim 6. 7. ve 8. sınıflarda İngilizce öğrenme ve öğretme sürecinde karşılaşılan sorunların incelenmesi: Öğretmen ve öğrenci görüşleri. Yayınlanmamış Yüksek lisans tezi, Mehmet Akif Üniversitesi, Burdur.

Gök-Çatal, Ö., Şahin, H. & Çelik, H. (2018). 6., 7. ve 8. Sınıflarda İngilizce öğrenme sürecinde karşılaşılan sorunların incelenmesi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, (47), 123-136.

Gökçe, Ö. F. (2014). A comparison of EFL teachers' perception of formative assessment in

public and private schools. Unpublished Master Thesis, Çağ University, Institute of Social Sciences, Mersin

Gözütok, D. (2006). *Öğretim İlke ve Yöntemleri.* Ankara: Ekinoks Eğitim Hizmetleri

Grabe, W. (1991). Current developments in second language reading research. *TESOL Quarterly*, *25*(3), 375-406.

Grabe, W. (2009). Teaching and testing reading. M.H. Long & C.J. Doughty (Eds.), In *The handbook of language teaching* (pp. 441-462). MA: Blackwell Publishing LTD.

Grabe, W. & Jiang, X. (2013). Assessing reading. *The companion to language assessment*, 1, 185-200.

Grabe, W. & Kaplan, R. B. (2014). *Theory and practice of writing: An applied linguistic perspective*. Routledge.

Grabe, W.B. & Stoller, F.L. (2002). *Teaching and researching reading.* Harlow, Essex: Pearson Education.

Greene, J., Caracelli, V. & Graham, W. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis*, 11, 255-274.

Gronlund, N. E. & Linn, R. L. (1990). *Measurement and Evaluation in Teaching* (6th Edition). New York: Macmillan Publishing

Gupta, S. & Gupta, A. (2013). The system approach in education. *International Journal of Managemet*, 1(1), 52-55.

Gurbetoğlu, A. 2008. Bilimsel Araştırma Yöntemleri. http://agurbetoglu.com/bilimselarastirma.html adresinden 16.04.2021 tarihinde ulaşılmıştır.

Güler, N. (2011). *Eğitimde Ölçme ve Değerlendirme.* Ankara: Pegem Akademi

Güllüoğlu, Ö. (2004). Attitudes Towards Testing Speaking at Gazi University Preparatory School of English and Suggested Speaking Tests. Yayımlanmamış yüksek lisans tezi, Gazi Üniversitesi, Ankara.

Günday, R. (2007). Yabancı Dil Öğretiminde Başarısızlığa Neden Olan Etmenler Üzerine Bir Araştırma, *Milli Eğitim Dergisi*, 175, 210-229.

Güneş, F. (2009). Türkçe Öğretiminde Günümüz Gelişmeleri ve Yapılandırıcı Yaklaşım/The New Developments In Teaching of Turkish and Constructivist Approach. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, *6*(11), 1-21.

Güneş, F. (2012). Testlerden etkinliklere Türkçe öğretimi. *Dil ve Edebiyat Eğitimi Dergisi*, *1*(1), 31-42.

Güzel, E. B. & Kula, S. (2015). Karma yöntem araştırmalarında verilerin analizi ve yorumlanması. Y. Dede ve S.B. Demir (Ed.), *Karma yöntem araştırmaları: Tasarımı ve yürütülmesi* içinde (s. 217-262). Ankara: Anı Yayıncılık

Hadley, A. O. & Reiken, E. (1993). *Teaching Language in Context, and Teaching Language in Context.* Heinle & Heinle Publishers.

Halali, A. A. S., Singh, H. K. D., Saeed, I. M. & Making Z. (2017). Teachers' perceptions toward the use of classroom-based assessment in English reading. *International Journal of Education and Research,* 5(11), 153-168.

Hamayan, E. V. (1995). Approaches to Alternative Assessment. *Annual Review of Applied Linguistics*, 15, 212-226.

Hamp-Lyons, L. (2003). Writing teachers as assessors. B. Kroll (Ed.), In *Exploring the dynamics of second language writing* (pp. 162–189). Cambridge: Cambridge University Press.

Han, K. & Kaya, H. (2014). Turkish EFL teachers' assessment preferences and practices in the context of constructivist instruction. *Journal of Studies in Education*, 4(1), 77-93.

Hancock, C. R. (1994). Alternative Assessment and Second Language Study: What and Why? https://files.eric.ed.gov/fulltext/ED376695.pdf adresinden 10.03.2022 tarihinde ulaşılmıştır

Harmer, J. (2007). *The Practice Of English Language Teaching*. Pearson Longman.

Harris, L. (2007). Employing formative assessment in the classroom. *Improving Schools*, 10, 249-260

Harrow, A. J. (1977). The behavioral objectives movement: its impact on physical education. *Educational Technology, 17*(6), 31-38.

Hayes, B. E. (1992). *Measuring customer satisfaction: development and use of questionnaires*. Milwaukee, WI: ASQC Quality Press.

Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford, UK: Oxford University Press. http://www.tesl-ej.org/wordpress/issues/volume5 adresinden 04.04.2021 tarihinde ulaşılmıştır.

Henning, G. (1992). Dimensionality and construct validity of language tests. *Language Testing*, *9*(1), 1-11.

Higuchi, C. (1992, September). Performance-based assessments and what teachers need. What Works in Performance Assessment? Conference, University of California, Los Angeles.

Hinkel, E. (2006). Current perspectives on teaching the four skills. *Tesol Quarterly*, *40*(1), 109-131.

Hodges, T. S. (2017). Theoretically speaking: An examination of four theories and how they support writing in the classroom. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 90(4), 139-146.

Howard, F. (1980). Testing communicative proficiency in French as a second language: A search for procedures. *Canadian Modern Language Review*, 36(2), 272-289.

Hsieh, H. F. & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative health research*, 15(9), 1277-1288.

Hubley, N. N. (2012). Assessing reading. C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoynoff. (Eds.). In *The Cambridge guide to second language assessment.* (pp. 211-217). Cambridge University Press, USA. https://books.google.com.tr/books adresinden 04.04.2021 tarihinde ulaşılmıştır.

Hudson, R. F., Torgesen, J. K., Lane, H. B., & Turner, S. J. (2012). Relations among reading skills and sub-skills and text-level reading proficiency in developing readers. *Reading and Writing*, *25*(2), 483-507.

Huerta-Macias, A. (1995). Alternative assessment: responses to commonly asked questions. *TESOL Journal*, 5(1), 8-11.

Huff, K. & Goodman, D. P. (2007). The demand for cognitive diagnostic assessment. *Cognitive diagnostic assessment for education: Theory and applications*, 19-60.

Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge England: Cambridge University Press.

Huot, B. (1990). The literature of direct writing assessment: Major concerns and prevailing trends. *Review of Educational research*, 60(2), 237-263.

Hussain, S., Tadesse, T. & Sajid, S. (2015). Norm-referenced and criterion-referenced test in EFL classroom. *International Journal of Humanities and Social Science Invention*, 4(10), 24-30.

Hyland, K. (2002). *Teaching and researching writing*. Toronto, ON: Longman

Hyland, K. (2008). Genre and academic writing in the disciplines. *Language Teaching*, 41(4), 543-562.

Hymes, D. (1971). Competence and performance in linguistic theory. R. Huxley & E. Ingram (Eds.), In Language Acquisition: Models And Methods içinde (pp. 3-28). London: Academic Press.

İleri, S. A. (2019). İngilizce öğretmenlerinin konuşma becerisi ölçümüne ilişkin gereksinimleri ve bir hizmet içi eğitim önerisi. Yayınlanmamış doktora tezi, Ankara Üniversitesi, Ankara.

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, *25*(3), 385-402.

Ingram, D. E. (1985). Assessing proficiency. *Modelling and Assessing Second Language Acquisition*, 18, 215.

Işık, A. (2008). Yabancı dil eğitimimizdeki yanlışlar nereden kaynaklanıyor? *Journal of Language and Linguistic Studies, 4*(2), 15-26.

Ivanova, V. (2011, May). Construction and evaluation of achievement tests in English. National Conference, Education in The Information Society, University of Plovdiv, Bulgaria

İleri, S. A. (2019). İngilizce öğretmenlerinin konuşma becerisi ölçümüne ilişkin gereksinimleri ve bir hizmet içi eğitim önerisi. Yayımlanmamış doktora tezi, Ankara Üniversitesi, Ankara.

Jamieson, J., Jones, S., Kirsch, I., Mosenthal, P. & Taylor, C. (2000). *TOEFL 2000 Framework.* Princeton, NJ: Educational Testing Service.

Jannati, S. (2015). ELT teachers' language assessment literacy: Perceptions and practices. Educational Research Association *The International Journal of Research in Teacher Education,* 6(2), 26-37.

Joyce, P. (2018). L2 vocabulary learning and testing: The use of L1 translation versus L2 definition. *The Language Learning Journal*, 46(3), 217-227.

Kalay, S. & Büyükkarcı, K. (2020). English language teachers' views on teaching and assessment of writing skills. *SDU International Journal of Educational Studies*, 7(2), 262-286

Kan, A. (2018). Ölçmenin temel kavramları. H. Atılgan (Ed.), *Eğitimde ölçme ve değerlendirme* içinde (s. 19-42). Ankara: Anı Yayıncılık

Kaptan, S. (1998). *Bilimsel Araştırma Teknikleri ve İstatistik Yöntemleri*. Ankara: Tekışık Matbaası.

Karakaya, İ. (2012). *Bilimsel Araştırma Yöntemleri.* Ankara: Anı Yayınları

Karamustafaoğlu, S., Çağlak, A. & Meşeci, B. (2012). Alternatif ölçme değerlendirme araçlarina ilişkin sinif öğretmenlerinin öz yeterlilikleri. *Amasya Üniversitesi Eğitim Fakültesi Dergisi*, *1*(2), 167-179.

Kararmaz, S. & Arslan, A. (2014). İlköğretim İngilizce öğretmenlerinin öğretmenlik mesleği özel alan yeterliklerine ilişkin algılarının belirlenmesi. *Uşak Üniversitesi Sosyal Bilimler Dergisi*, *7*(4), 203-232

Karasar, N. (2013). *Bilimsel Araştırma Yöntemi.* Ankara: Nobel Yayın Dağıtım.

Kargın, T. (2007). Eğitsel değerlendirme ve bireyselleştirilmiş eğitim programı hazırlama süreci. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Özel Eğitim Dergisi*, 8(01), 1-15.

Kaulfers, W. V. (1944). Wartime development in modern-language achievement testing. *The Modern Language Journal*, *28*(2), 136-150.

Kayacan, İ. & Eskicumalı, A. (2014, Haziran). Ortaöğretim Kurumlarında Görev Yapan İngilizce Öğretmenlerinin Derslerinde Kullandıkları Ölçme ve Değerlendirme Tekniklerinin Analizi. III. Sakarya'da Eğitim Araştırmaları Kongresi, Sakarya

Kębłowska, M. (2012). The place of affect in second language acquisition. In *New perspectives on individual differences in language learning and teaching* (pp. 157-167). Berlin, Heidelberg: Springer.

Keeley, P. (2008). *Science formative assessment: 75 practical strategies for linking assessment, instruction, and learning*. Thousand Oaks, CA: Corwin

Kelleghan, T. & Greaney, V. (2001). *Using Assessment To İmprove The Quality Of Education.* Paris: UNESCO

Kelly, L. G. (1976) *25 Centuries of Language Teaching*. Rowley, MA: Newbury House.

Kepner, C. G. (1991). An experiment in the relationship of types of written feedback to the development of second-language writing skills. *The Modern Language Journal*, 75(3), 305-313.

Kırkgöz, Y. & Ağçam, R. (2012). Investigating the written assessment practices of Turkish teachers of English at primary education. *Journal of Language and Linguistic Studies, 8*(2).

Kırkgöz, Y., Babanoğlu, M. P. & Ağçam, R. (2017). Turkish EFL teachers' perceptions and practices of foreign language assessment in primary education. *Journal of Education and e-Learning Research*, 4(4), 163-170.

Kim, M. (2019). The perceptions of students and teachers onthe practice of assessment in extensive reading. *English Teaching, 74*(4), 179-203.

Kirk, J. & Miller, M. L. (1986). *Reliability and validity in qualitative research*. London: Sage.https://books.google.com.tr/books adresinden 10.05. 2021 tarihinde ulaşılmıştır

Kirkova-Naskova, A., Tergujeff, E., Frost, D., Henderson, A., Kautzsch, A., Levey, D., Murphy, D., WaniekKlimczak, E. (2013, August). Teachers' views on their professional training and assessment practices: Selected results from the English Pronunciation Teaching in Europe survey. *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference*, Ames, IA: Iowa State University.

Kizlik, B. (2012). Measurement, assessment, and evaluation in education. *Retrieved October, 10*, 2015.

Kline, J. A. (1996). *Listening Effectively.* Washington: Air University Press

Kobayashi, M. (2002). Method effects on reading comprehension test performance: Text organization and response format. *Language testing*, *19*(2), 193-220.

Kondo-Brown, K. (2002). A FACETS analysis of rater bias in measuring Japanese second language writing performance. *Language Testing*, *19*(1), 3-31.

Kordes, U., Sicherl-Kafol, B. & Brunauer, A. H. (2014). A Model of Formative Assessment in Music Education. *Athens Journal of Education*, 1(4), 295-307.

Köksal, M. S. (2015). Karma yöntem desenleri. Y. Dede ve S.B. Demir (Çev. Ed.). *Karma Yöntem Araştırmalarının Temelleri* içinde (s. 165-202). Ankara: Anı Yayıncılık

Körhasanoğulları, N. (2020). İlköğretim İngilizce Dersi Öğretim Programının Iletişimsel dil Yaklaşımına Uygunluğu Açısından Incelenmesi (Bir Durum Çalışması). Yüksek Lisans Tezi, Anadolu Üniversitesi, Eskişehir

Krashen, S. (1981). Second language acquisition. *Second Language Learning*, *3*(7), 19-39.

Krashen, S. (1987). *Principles and practice in second language acquisition*. UK: Prentice Hall International

Krashen, S. (2009). Anything but reading. *Knowledge Quest*, *37*(5), 18-25

Kunnan, A. J. (1998). An introduction to structural equation modelling for language assessment research. *Language Testing*, *15*(3), 295-332.

Kutlu, Ö. & Altıntaş, Ö. (2021). Psikolojik ölçmelerin kısa tarihi ve 21. yüzyılda sınıf içi durum belirleme anlayışı. *Trakya Eğitim Dergisi*, 11(3), 1599-1620.

Kuyumcu-Vardar, A. & Acar, F. (2018). Öğrenme-öğretme sürecinde kullanılabilecek hızlı ölçme araçları. *Electronic Turkish Studies*, *13*(11).

Kwako, J. (2003). A brief summary of traditional and alternative assessment in the college classroom. www.stat.wisc.edu/~nordheim/Kwako_assessment4.doc adresinden 25.02.2022 tarihinde ulaşılmıştır.

LaBerge, D. & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive psychology*, *6*(2), 293-323.

Lado, R. (1964). *Language Teaching: A Scientific Approach.* New York: McGraw-Hill

Larios, J. R. D., Murphy, L. & Marín, J. (2002). A critical examination of L2 writing process research. In *New directions for research in L2 writing* (pp. 11-47). Springer, Dordrecht.

Larsen-Freeman, D. (2011). A complexity theory approach to second language development/acquisition. In D. Atkinson (Ed.), *Alternative approaches to second language acquisition* (pp. 60-84). London and New York: Routledge.

Larsen-Freeman, D. & Anderson, M. (2013). *Techniques and principles in language teaching: Oxford handbooks for language teachers*. Oxford: Oxford university press.https://books.google.com.tr/books? adresinden 24.03.2021 tarihinde ulaşılmıştır.

LeClercq, A. (1986). The Academic Library/High School Library Connection: Needs Assessment and Proposed Model. *Journal of Academic Librarianship*, *12*(1), 12-18.

Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Singapore: Springer Singapore.

Lee, S. H. (2006). Constructing effective questionnaires. *Handbook of human performance technology, Hoboken, NJ: Pfeiffer Wiley*, 760-779.

Lennon, P. (2000). The lexical element on spoken second language fluency. H. Riggenbach (Ed.), In *Perspectives on fluency* (pp. 25–42). Ann Arbor: University of Michigan Press

Letina, A. (2015). Application of traditional and alternative assessment in science and social studies teaching. *Croatian Journal of Education*, 17(1), 137-152.

Lewkowicz, J. (1991). Testing Listening Comprehension: A New Approach?. *Hong Kong Papers in Linguistics and Language Teaching*, *14*, 25-31.

Liberman, I. Y. & Liberman, A. M. (1990). Whole language vs. code emphasis: Underlying assumptions and their implications for reading instruction. *Annals of Dyslexia*, 40(1), 51-76.

Lightbown, P.M. & Spada, N. (2006). *How languages are learned*. Oxford: Oxford University Press.

Lincoln, Y. S. & Guba, E. G. (1985). *Naturalistic inquiry*. London: Sage Publications https://books.google.com.tr/books adresinden 10.05.2021 tarihinde ulaşılmıştır.

Lloyd-Jones, R. (1977). Primary trait scoring. C. R. Cooper & L. Ode (Eds.), In *Evaluating writing: Describing, measuring, judging* (pp. 3-32). Urbana: NCT

Luborsky, M. R. & Rubinstein, R. L. (1995). Sampling in qualitative research: rationale, issues, and methods. *Journal of Aging Research*, 17 (1), 89-113.

Lynch, B. K. (2001). Rethinking assessment from a critical perspective. *Language testing*, *18*(4), 351-372.

Maarof, N. (2016). *Assessing second language reading*. Singapore: Partridge Publishing https://books.google.com.tr/books adresinden 04.04.2021 tarihinde ulaşılmıştır.

Maden, S., Dinçel, Ö. & Maden, A. (2015). Türkçeyi Yabancı Dil Olarak Öğrenenlerin Yazma Kaygıları. *Uluslararası Türkçe Edebiyat Kültür Eğitim Dergisi*, 4(2), 748-769.

Madsen, H. S. (1983). *Techniques in Testing*. Oxford University Press,

Magno, C. & Ouano, J. (2010). Designing written assessment for student learning. *Philippines: Phoenix Pub*.

Majid, A. F. (2011). School-based assessment in Malaysian schools: The concerns of the English teachers. *Journal of US-China Education*, 8 (5), 1-15.

Mahshanian, A., Shoghi, R. & Bahrami, M. (2019). Investigating the Differential Effects of Formative and Summative Assessment on EFL Learners' End-of-term Achievement. *Journal of Language Teaching and Research*, 10(5), 1055-1066.

Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing, 30* (3), 329-344.

Marshall, M. N. (1996). Sampling for qualitative research. *Family Practice*, 13(6), 522-526.

Marso, R. N. (1969). The influence of test difficulty upon study efforts and achievement. *American Educational Research Journal*, 6(4), 621-632.

Martin, B. L. & Reigeluth, C. M. (1992). Affective education and the affective domain: Implications for instructional-design theories and models. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: A new paradigm of instructional theory* (pp. 485-511). Hillsdale, NJ: Lawrence Erlbaum Associates.

Mason, J. (2002). *Qualitative Researching*. (2nd Edition). London: Sage Pub Inc

Matin, Z. N. (2013). Speaking assessment at secondary and higher secondary levels and students deficiency in speaking skill: A study to find interdependence. *Stamford Journal of English*, 7, 234-251.

Matsuda, P. K., Canagarajah, A. S., Harklau, L., Hyland, K. & Warschauer, M. (2003). Changing currents in second language writing research: A colloquium. *Journal of Second Language Writing*, 12(2), 151-179.

Maxwell, J. (1992). Understanding and validity in qualitative research. *Harvard Educational Review*, *62*(3), 279-301.

Maxwell, J. (2005). *Qualitative research design: An interactive approach*. London: Sage Publications.

McAfee, O. & Leong, D. J. (2012). *Erken çocukluk döneminde gelişim ve öğrenmenin değerlendirilmesi ve desteklenmesi* (Çev. B. Ekinci). Ankara: Nobel Yayıncılık.

McNamara, T. (2000). *Language testing*. Oxford University Press.

Mead, N. A. (1980, November). Assessing speaking skills: Issues of feasibility, reliability, validity and bias. Paper presented at the Speech Communication Association Convention, New York

Mead, N. A. ve Rubin, D. L. (1985). Assessing Listening and Speaking Skills. *ERIC Digest*, 1-3.

MEB. (2005). *İlköğretim 1–5. sınıf programları tanıtım el kitabı.* Ankara: Devlet Kitapları Müdürlüğü Basımevi.

MEB. (2018). *İngilizce dersi öğretim programı (İlkokul ve ortaokul 2, 3, 4, 5, 6, 7 ve 8. Sınıflar).* Ankara: Milli Eğitim Bakanlığı.

Mede, E. & Atay, D. (2017). English language teachers' assessment literacy: The Turkish context. *Dil Dergisi*, 168(1), 43-60

Menhart, S. (2011, September). Direct versus indirect assessment methodologies. *Proceedings of the 2011 Midwest Section Conference of the American Society for Engineering Education*, Russellville, AK.

Merhout, J., Benamati, J., Rajkumar, T. M., Anderson, P. & Marado, D. (2008). Implementing direct and indirect assessment in the MIS curriculum. *Communications of the Association for Information Systems*, 23(1), 24.

Merriam, S.B. (1998). *Qualitative research and case study applications in education*. San Francisco: Jossey-Bass

Merter, F., Kartal, Ş. & Çağlar, İ. (2012). Ortaöğretim İngilizce dersi yeni öğretim programının öğretmen görüşlerine göre değerlendirilmesi. *Mehmet Akif Ersoy Üniversitesi Eğitim Fakültesi Dergisi*, *1*(23), 43-58.

Mertova, K. (2013). Issues of designing effectıve conversational lessons for secondary school students. Unpublished master's thesis, University of West Bohemia, Plzeň, Czech Republic

Micheels, W. J. & Karnes, M. (1950). *Measuring Educational Achievement*. New York: McGraw-HillBook Company.

Miles, M. B. & Huberman, A.M. (1994). *Qualitative data analysis: An expanded sourcebook.* (2nd Edition). California: SAGE Publications.

Mills, G. E. & Gay, L. R. (2016). *Educational research: competencies for analysis and applications*. (11. Baskı) USA: Pearson Education.

Mintah, J. K. (2003). Authentic assessment in physical education: Prevalence of use and perceived impact on students' self-concept, motivation, and skill achievement. *Measurement in physical education and exercise science*, 7(3), 161-174.

Mishler, E. (1990). Validation in inquiry-guided research: The role of exemplars in narrative studies. *Harvard Educational Review*, 60(4), 415-443.

Mitchell, C. B. & Vidal, K. E. (2001). Weighing the ways of the flow: Twentieth century Language instruction. *The Modern Language Journal*, 85(1), 26-38

Morales, H. S. & Fernández, L. C. (2019). Authentic assessment in the listening comprehension classroom: benefits and implications. *Gist: Education and Learning Research Journal*, (19), 6-30.

Morley, J. (2001). Aural comprehension instruction: Principles and practices. *Teaching English as a second or foreign language*, 3, 69-85.

Morrow, K. (1981). Communicative language testing: evolution or revolution? J. C. Alderson & A. Hughes (Eds.). In *Issues in Language Testing.* (pp. 9-25). London: The British Council.

Morse, J. M. (1991). Approaches to qualitative-quantitative methodological triangulation. *Nursing Research*, 40(2), 120-123.

Munoz Marin, J. H. (2009). Exploring teachers practices for assessing reading comprehension abilities in english as a foreign language. *Profile Issues in Teachers' Professional Development*, 11(2), 71-84.

Murphy, D. (2000). Key Concepts in ELT. *ELT Journal, 54*(2): 210-211.

Nartgün, Z. (2007). Aynı puanlar üzerinden yapılan mutlak ve bağıl değerlendirme uygulamalarının notlarda farklılık oluşturup oluşturmadığına ilişkin bir inceleme. *Ege Eğitim Dergisi*, 8(1), 19-40.

Nasab, F. G. (2015). Alternative versus traditional assessment. *Journal of Applied Linguistics and Language Research*, 2(6), 165-178.

Neuman, D. (2014). Qualitative research in educational communications and technology: A brief introduction to principles and procedures. *Journal of Computing in Higher Education*, 26(1), 69-86.

Neuman, W. L. & Robson, K. (2014). *Basics of social research*. Toronto: Pearson Canada.

Newell, R. J. (2002). A different look at accountability: The EdVisions Approach. *Phi Delta Kappan*, 84(3), 208-211.

Nichols, P. D., Meyers, J. L. & Burling, K. S. (2009). A framework for evaluating and planning assessments intended to improve student achievement. *Educational Measurement: Issues and Practice*, 28(3), 14-23.

Nunan, D. (1991). Communicative tasks and the language curriculum. *TESOL Quarterly*, 25, 279-295.

Nunan, D. (1998). Approaches to teaching listening in language classroom. Proceedings of the 1997 Korea TESOL Conference. Taejon, Korea: KOTESOL

Nunan, D. (1999). *Second Language Teaching and Learning*. Heinle & Heinle Publishers,

Nunnaly, J. C. (1970). *Introduction To Psychological Measurement*. New York: McGraw Hill.

Nurov, A. (2000). Self-assessment of foreign language achievement: The relationship between students' self-assessment, teachers' estimates and achievement test, Doctoral dissertation, Bilkent University, Ankara.

Odabaşı, F. (1997). Eğitimde sistem yaklaşımı ve eğitim teknolojisi. *Eğitim ve Bilim*, 21(106), 23-34

OECD Publishing. (2013). *Synergies for better learning: An international perspective on evaluation and assessment*. OECD.

Oiler, J. W. (1979). *Language Tests At School*. London: Pearson-Longman.

O'Loughlin, K. (2001). *The equivalence of direct and semi-direct speaking tests*. Cambridge: Cambridge University Press.

O'Loughlin, K. (2013). Developing the assessment literacy of university proficiency test users. *Language Testing,* 30(3), 363-380.

Onwuegbuzie, A. J. & Leech, N. L. (2004). Enhancing the interpretation of "significant" findings: The role of mixed methods research. *The qualitative report*, 9(4), 770-792.

Onwuegbuzie, A. J. & Collins, K. M. (2007). A typology of mixed methods sampling designs in social science research. *Qualitative Report*, 12(2), 281-316.

O'Sullivan, B. (2011). Language testing. J. Simpson (Ed.), In *Routledge handbook of applied linguistics* (pp. 262–276). Oxford, England: Routledge.

Ölmezer-Öztürk, E. (2018). Developıng and valıdatıng language assessment knowledge scale (LAKS) and explorıng the assessment knowledge of EFL teachers. Doctoral dissertation, Anadolu University, Turkey

Ölmezer-Öztürk, E. & Aydın, B. (2019). Voices of EFL teachers as assessors: Their opinions and needs regarding language assessment. *Eğitimde Nitel Araştırmalar Dergisi*, 7(1), 373-390.

Önalan, O. & Gursoy, E. (2020). EFL teachers' views and needs on in-service training as a part of professional development: A case study in Turkish context. *Bartın University Journal of Faculty of Education*, 9(2), 373-387.

Öz, H. (2014). Turkish teachers' practices of assessment for learning in the English as a foreign language classroom. *Journal of Language Teaching and Research, 5* (4), 775-785.

Özçelik, D. A. (2010). *Ölçme ve Değerlendirme* (3. Baskı) Ankara: Pegem Akademi

Özerbaş, M.A. (2005). Eğitim teknolojisi öğelerinin sistem kuramları açısından değerlendirilmesi. *Kazım Karabekir Eğitim Fakültesi Dergisi,* 11, 252-268.

Özmen, K. S. (2012). A study of the washback effect of exams and associated teacher burnout. *The Teacher Trainer Journal*, 26(3), 13-24.

Öztürk, E. (2003). *An assessment of high school biology curriculum implementation*. Unpublished doctoral dissertation, METU, Ankara.

Öztürk, G. & Aydın, B. (2019). English language teacher education in Turkey: Why do we fail and what policy reforms are needed?. *AJESI - Anadolu Journal of Educational Sciences International,* 9(1), 181-213

Paker, T. (2006, Eylül). Çal bölgesindeki okullarda İngilizce öğretiminin sorunları ve çözüm. Çal Sempozyumu'nda sunulmuş bildiri (s. 684-690). Denizli.

Paker, T. (2015). İletişimsel Dil Öğretim Yöntemi. N. Bekleyen (Ed.), *Dil Öğretimi* içinde (s. 188-208). Ankara: Pegem Akademi.

Paker, T. & Höl, D. (2012). Attitudes and perceptions of the students and instructors towards testing speaking communicatively. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 32(11), 13-24

Parmaksız Ş. R & İncirci, A. (2016). 10. Sınıf ortaöğretim İngilizce öğretim programının Stufflebeam (CIPP) bağlam, girdi, süreç ve ürün modeline göre değerlendirilmesi, *International Journal of Language Academy*, 4(2), 153-173.

Peker, M. & Güngör, F. (2015). Karma yöntem çalışmasına giriş. Y. Dede ve S. B. Demir (Ed.), *Karma yöntem araştırmaları: Tasarımı ve yürütülmesi* içinde (s. 155-182). Ankara: Anı Yayıncılık

Peker, T. (2006, Eylül). Çal bölgesindeki okullarda İngilizce öğretiminin sorunları ve çözüm. Çal Sempozyumu'nda sunulmuş bildiri (s. 684-690). Denizli.

Peterson, B. K. & Reider, B. P. (2002). Perceptions of computer-based testing: a focus on the CFM examination. *Journal of Accounting Education*, 20(4), 265-284.

Piasecka, L. (2013). What does it feel like to use English? Empirical evidence from EFL students. E. Piechurska-Kuciel & E. Szymańska-Czaplak (Eds.), In *Language in Cognition and Affect* (pp. 219-237). Heidelberg, Germany: Springer.

Pill, J. & Harding, L. (2013). Defining the language assessment literacy gap: evidence from a parliamentary inquiry. *Language Testing*, 30(3), 381–402.

Poggenpoel, M. & Myburgh, C. (2003). The researcher as research instrument in educational research: A possible threat to trustworthiness?(A: research_instrument). *Education*, 124(2), 418-423.

Popham J. W. (1995). *Classroom assessment: What teachers need to know*. Boston: Allyn and Bacon

Popham, J. W. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, 48, 4-11.

Qian, D. D. (2009). Comparing direct and semi-direct modes for speaking assessment: Affective effects on test takers. *Language Assessment Quarterly*, *6*(2), 113-125.

Purpura, J. E. (2004). *Assessing grammar*. UK: Cambridge University Press.

Pramujunarta, A. (2007). The Problems of learning speaking through retelling story at mahesa English course of Pare-Kediri. Unpublished master's thesis, IAIN Kediri, Indonesia

Rahman, M.S. (2017). The advantages and disadvantages of using qualitative and quantitative approaches and methods in language "testing and assessment" research: a literature review. *Journal of Education and Learning*, 6(1), 102-112

Raimes, A. (1983). *Techniques in teaching writing*. New York: Oxford University Press

Rajak, A., Shrivastava, A. & Tripathi, A. (2019). An approach to evaluate program outcomes and program educational objectives through direct and indirect assessment tools. *International Journal of Emerging Technologies in Learning (iJET)*, *14*(23), 85-97

Randaccio, M. (2013). Writing skills: theory and practice. *QuaderniCIRD*, 7, 51-74

Rapi, L. & Miconi, H. (2015). Foreign Language in the Albanian Primary School, Benefits and Problems. *Journal of the Association-Institute for English Language and American Studies*, *3*(9), 15-20.

Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language testing*, *18*(4), 429-462.

Reiner, C. M., Bothell, T. W., Sudweeks, R. R. & Wood, B. (2002). Preparing effective essay questions. http://www.ndcl.info/Writing%20Effective%20Essay%20Questions.pdf adresinden 10.02.2022 tarihinde ulaşılmıştır

Richards, J. C. (2006). *Communicative language teaching today.* New York: Cambridge University Press

Richards, J. C. & Rodgers, T. S. (2002). *Approaches and Methods in Language Teaching.* Cambridge, UK: Cambridge University Press.

Richards, J. C. & Schmidt, R. W. (2013). *Longman dictionary of language teaching and applied linguistics*. Routledge.

Robson, C. (2015). *Bilimsel Araştırma Yöntemleri*. Ş. Çınkır ve N. Demirkasımoğlu (Çev. Ed.) Ankara: Anı Yayıncılık

Roever, C. & McNamara, T. (2006). Language testing: The social dimension. *International Journal of Applied Linguistics, 16*(2), 242-258.

Rost, M. (2002). *Teaching and Researching Listening*. London, UK: Longman

Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal*, 78(2), 199-221

Rumelhart, D. E. (1991). Understanding understanding. *Memories, thoughts and emotions: Essays in honor of George Mandler*, 257-275.

Rumelhart, D. E. & A. Ortony. (1997). The Representation of knowledge in memory. R. C. Anderson, R. J. Sprio & W. E. Montague (Eds.), In *Schooling and the Acquisition of Knowledge* (pp. 99-135). Hillsdale, NJ: Lawrence Erlbaum

Saeed, M., Tahir, H. & Latif, I. (2018). Teachers' Perceptions about the Use of Classroom Assessment Techniques in Elementary and Secondary Schools. *Bulletin of Education and Research*, 40(1), 115-130.

Safa, M. A. & Goodarzi, S. (2014). The washback effects of task-based assessment on the Iranian EFL learners' grammar development. *Procedia - Social and Behavioral Sciences*, 98, 90-99.

Sah, K. P. (2012). Assessment and test in teaching and learning. *Academic Voices: A Multidisciplinary Journal*, *2*, 28-32.

Saigh, P. A. (1984). Unscheduled assessment: Test anxiety, academic achievement, and social validity. *Educational Research Quarterly*, *9*(4), 6–11

Samuels, S. J. & Kamil, M. L. (1984). Models of the reading process. *Handbook of reading research*, 1, 185-224.

Sandelowski, M. (2000). Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in Nursing & Health*, *23*(3), 246-255.

Savignon, S. (1987). Communicative language teaching. *Theory Into Practice*, 26(4), 235-242

Scanlan, C. L. (2012). Assessment, Evaluation, Testing and Grading http://www.umdnj.edu/idsweb/idst5350/assess_eval_test_grade.htm adresinden 05.03.2022 tarihinde ulaşılmştır.

Scarino, A. (2005). Going beyond criteria for judging performance. Barbel: *Journal of the Australian Federation of Modern Language Teachers' Association*, 39(3), 8.

Schumann, J. (1994). "Where is cognition? Emotion and cognition in second language acquisition". *Studies in Second Language Acquisition*, 16, 231-242.

Seçkin, H. (2011) İlköğretim 4. sınıf İngilizce öğretim programına ilişkin öğretmen görüşleri, *Uluslararası İnsan Bilimleri Dergisi*, (8), 2, 511-525.

Sellan, R. (2017). Developing assessment literacy in Singapore: How teachers broaden English language learning by expanding assessment constructs. *Papers in Language Testing and Assessment, 6* (1), 64-87.

Serafini, F. (2001). Three Paradigms of Assessment: Measurement, Procedure, and Inquiry. *The Reading Teacher,* 54, 384-393.

Shaban, A. M. S. (2014). A comparison between objective and subjective tests. *Journal of the College of Languages (JCL)*, (30), 44-52.

Shepard, L. (1980). Standard setting issues and methods. *Applied Psychological Measurement*, *4*(4), 447-467.

Shepard, L. A. (1991). Will national tests improve student learning?. *The Phi Delta Kappan*, 73(3), 232-238.

Shepard, L. A. (2008). A brief history of accountability testing, 1965– 2007. K. E. Ryan, & L. A. Shepard (Eds.), In *The future of test-based educational accountability* (pp. 25–46). New York, NY: Routledge.

Shohamy, E. (1984). Does the testing method make a difference? The case of reading comprehension. *Language testing*, 1(2), 147-170.

Silverman, D. (2000). *Doing Qualitative Research. A Practical Handbook*. London: Sage Publications Inc

Sireci, S. G. (1998). The construct of content validity. *Social indicators research*, *45*(1), 83-117.

Smith, H. K. (Ed.). (1971). *Meeting Individual Needs in Reading*. Newark: International Reading Association.

Spivey, N. N. & King, J. R. (1989). Readers as writers composing from sources. *Reading Research Quarterly*, 7-26.

Spolsky, B. (1978). *Approaches to language testing: Advances in Language Testing Series* VA: Center for Applied Linguistics

Stern, H. H. (1983). *Fundemental Concepts of Language Teaching.* Oxford, England: Oxford University Press

Stiggins, R. J. (1982). A comparison of direct and indirect writing assessment methods. *Research in the Teaching of English*, 101-114.

Stiggins, R. J. (1991). Assessment literacy. *Phi Delta Kappan*, 72, 534-539.

Stiggins, R. J. (1997). Dealing with the practical matter of quality performance assessment. *Measurement in Physical Education and Exercise Science*, 1(1), 5-17.

Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758-765.

Sujana, I. M. (2000). Movements In Language Testıng: From Grammar-Based To Communicative Language Testing. *Jurnal Ilmu Pendidikan Fkıp Unram* 13(48), 1-8

Sujana, I. M. (2016): Assessing oral proficiency: problems and suggestions for elicitation techniques.https://d1wqtxts1xzle7.cloudfront.net/52384192/TECHNIQUE-SPEAKING- adresinden 20.03.2021 tarihinde ulaşılmıştır.

Susanti, N. I. (2019). Spoken language features in a project based learning of ESP class activity. *EEAL Journal (English Education and Applied Linguistics Journal)*, 2(2), 24-28.

Şahin, Y. (2018). *Yabancı Dilde Ölçme ve Değerlendirme.* Ankara: Pegem Akademi

Şahinkarakaş, S. (2012). The role of teaching experience on teachers' perceptions of language assessment. *Procedia – Social and Behavioral Sciences,* 47, 1787-1792

Şenel, M. (2015). Dilbilgisi Çeviri Yöntemi (Grammar-Translation Method) N. Bekleyen (Ed.), *Dil Öğretimi* içinde (s. 24-37). Ankara: Pegem Akademi

Taherdoost, H. (2016). Validity and reliability of the research instrument; how to test the validation of a questionnaire/survey in a research. *International Journal of Academic Research in Management (IJARM),* 5(3), 28-36.

Tan, Ş. (2013). *Öğretimde ölçme ve değerlendirme*. Ankara: Pegem Akademi

Tao, N. (2014). Development and validation of classroom assessment literacy scales: English as a Foreign Language (EFL) teachers in a Cambodian Higher Education Setting. Unpublished PhD dissertation, Victoria University, Australia.

Taras, M. (2005). Assessment: Summative and Formative-Some Theoretical Reflections. *British Journal of educational Studies*, 466-478

Tarhan, M. (1992). An Experimental Study on the Effects of Frequent Testing. Yayımlanmamış yüksek lisans tezi, Bilkent Üniversitesi Yabancı Diller Bölümü, Ankara.

Tashakkori, A. & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. CA: Thousand Oaks.

Tashakkori, A. & Teddlie, C. (2009). Integrating qualitative and quantitative approaches to research. *The SAGE handbook of applied social research methods*, *2*, 283-317.

Tatlı, L. (2014). EFL teachers' attitudes toward teaching English to young learners. Unpublished master thesis, Çağ University, Institute of Social Sciences, Mersin.

Tekin, H. (2004). *Eğitimde Ölçme ve Değerlendirme* (17. Baskı). Ankara: Yargı Yayınevi.

Thompson, I. (1995). Testing Listening Comprehension. *AATSEEL Newsletter*, *37*(5), 24-31.

Thorndike, R. M. & Thorndike-Christ, T. (2010). *Measurement and evaluation in psychology and education* (8th Edition). Boston: Prentice Hall.

Tokur-Üner, B. & Aşılıoğlu, B. (2022). İngilizce öğretiminde ölçme ve değerlendirme sürecine ilişkin öğretmen görüşleri. *EKEV Akademi Dergisi*, 26(89), 25-50

Topaloğlu, C. (2011). Yönetim kuramları ve örgütiçi çatışmalar. *Girişimcilik ve Kalkınma Dergisi*, 6(1), 249-265

Toprak, T. E. (2015). Cognitive diagnostic assessment of second language reading comprehension: Application of the Log-Linear cognitive diagnosis modeling to language testing. Unpublished doctoral dissertation, Gazi University, Ankara, Turkey.

Torrance, H. & Pryor, J. (1998). *Investigating Formative Assessment: Teaching, Learning and Assessment in the Classroom*. McGraw-Hill Education (UK).

Tsagari, D. and Vogt, K. (2017). Assessment literacy Of foreign language teachers around Europe: Research, challenges and future prospects. *Papers in Language Testing and Assessment*, 6 (1), 41-64.

Tsui, J. S., Lau, T. S. & Fong, S. C. (1995). Analysis of accounting students' performance on multiple-choice examination questions: a cognitive style perspective. *Accounting Education*, *4*(4), 351-357.

Tuan, L. T. (2012). Teaching and assessing speaking performance through analytic scoring approach. *Theory and Practice in Language Studies*, *2*(4), 673-679

Turgut, M. F. & Baykul, Y. (2013). *Eğitimde Ölçme ve Değerlendirme*. Ankara: Pegem Akademi

Ur, P. (2012). *A course in English language teaching*. UK: Cambridge University Press. https://books.google.com.tr/books adresinden 24.03.2021 tarihinde ulaşılmıştır.

Valette, R. (1977). *Modern Language Testing*. New York: Harcourt Brace.

Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language teaching*, 40(3), 191-210.

Vigoya, F. (2000). Testing Accuracy and Fluency in Speaking Through Communicative Activities. *HOW Journal*, *5*(1), 95-104.

Vogt, W. P., Gardner, D. C. & Haeffele, L. M. (2012). *When to use what research design*.

https://books.google.com.tr/books adresinden 16.04.2021 tarihinde ulaşılmıştır.

Wall, D. (1996). Introducing new tests into traditional systems: insights from general education and from innovation theory. *Language Testing,* 13 (3), 334-354.

Walter, E. (Ed.). (2004). *Cambridge Learner's Dictionary*. Cambridge University Press.

Watt, H. M. G. (2005). Attitudes to the use of alternative assessment methods in mathematics: A study with secondary mathematics teachers in sydney, Australia. *Educational Studies in Mathematics,* 58, 21-44.

Weir, C. J. (2005). Language Testing and Validation. New York: Palgrave Macmillan.

Wewer, T. (2014). Assessment of young learners' English proficiency in bilingual content instruction CLIL. Unpublished Doctoral Dissertation, University of Turku, Finland

Wheeler, E. & McDonald, R. L. (2000). Writing in engineering courses. *Journal of Engineering Education,* 89(4), 481-486.

White, E. (2009). Student perspectives of peer assessment for learning in a public speaking course. *Asian EFL Journal*, *33*(1), 1-36.

White, R. & Arndt, V. (1991). *Process writing*. London: Longman.

Wiggins, G. P. (1998). *Educative assessment: Designing assessments to inform and improve student performance.* San Francisco, CA: Jossey-Bass

Wiliam, D. & Thompson, M. (2008). Integrating assessment with learning: What will it take to make it work? *Routledge,* 53-82

Wisdom, J. & Creswell, J. W. (2013). Integrating quantitative and qualitative data collection and analysis while studying patient-centered medical home models. *Agency for Healthcare Reseach and Quality,* 13, 1-5.

Wiseman, C. S. (2012). A comparison of the performance of analytic vs. holistic scoring rubrics to assess L2 writing. *International Journal of Language Testing*, *2*(1), 59-92.

Wolcott, H. F. (1990). On seeking-and rejecting-validity in qualitative research. *Qualitative inquiry in education: The continuing debate*, 121- 152.

Worrell, V., Evans-Fletcher, C. & Kovar, S. (2002). Assessing the cognitive and affective progress of children. *Journal of Physical Education, Recreation & Dance*, *73*(7), 29-34.

Wortham, S. C. & Hardin, B. J. (2001). *Assessment in early childhood education*. Merrill/Prentice Hall.

Worthen, B. R. (1993). Critical issues that will determine the future of alternative assessment. *Phi Delta Kappan*, *74*(6), 444-453.

Xi, X. (2007). Evaluating analytic scoring for the TOEFL Academic Speaking Test (TAST) for operational use. *Language Testing*, *24*(2), 251-286.

Xu, Y. & Brown, G.T. L. (2016). Teacher assessment literacy in practice: Reconceptualization. *Teaching and Teacher Education*, 58, 149-162.

Yağar, F. & Dökme, S. (2018). Niteliksel Araştırmaların Planlanması: Araştırma Soruları, Örneklem Seçimi, Geçerlik ve Güvenirlik. *Gazi Sağlık Bilimleri Dergisi*, 3(3), 1-9.

Yaman, İ. (2018). Türkiye'de İngilizce öğrenmek: Zorluklar ve fırsatlar. *RumeliDE Dil ve Edebiyat Araştırmaları Dergisi,* (11),161-175

Yaman, Ş. (2015). Dinleme becerisinin öğretimi. N. Bekleyen (Ed.), *Dil Öğretimi* içinde (s. 310-330). Ankara: Pegem Akademi.

Yamayee, Z. A. & Albright, R. J. (2008). Direct and indirect assessment methods: Key ingredients for continuous quality improvement and ABET accreditation. *The International journal of engineering education*, *24*(5), 877-883.

Yan, X. & Ginther, A. (2017). Listeners and raters: Similarities and differences in evaluation of accented speech. *Assessment in second language pronunciation*, 67-88.

Yang, R. (2020). Grammar and vocabulary testing in the senior high school entrance English test in China: A washback study from a learning oriented assessment perspective. Unpublished doctoral dissertation, Queensland University of Technology, Australia

Yaşar, M. (2011). Ölçme ve değerlendirmenin önemi. S. Tekindal (Edt.). *Eğitimde Ölçme ve Değerlendirme* içinde (s. 2-8). Ankara: Pegem Akademi.

Yavuz, Ş. (2005). Socio-demographic and institutional predictors of variations in English as foreign language teachers' efficacy perceptions. Yayınlanmamış yüksek lisans tezi, Marmara Üniversitesi, İstanbul.

Yavuz-Kırık, M. (2008). Yabancı dil olarak İngilizce öğretmenlerinin ölçme değerlendirme bağlamında tutum ve yaklaşımları. Yayınlanmamış doktora tezi, İstanbul Üniversitesi, İstanbul.

Yaylı, D. (2015). Yazma becerisinin öğretimi. N. Bekleyen (Ed.), *Dil Öğretimi* içinde (s. 288-305). Ankara: Pegem Akademi.

Yıldırım, A. (2012). İngilizce dil becerilerinin öğretimi ve değerlendirilme durumları arasındaki tutarlılığın öğretmen ve öğrenci görüşlerine göre belirlenmesi. Yayımlanmamış yüksek lisans tezi, Ankara Üniversitesi, Ankara

Yıldırım, A. ve Şimşek, H. (2013). *Sosyal Bilimlerde Nitel Araştırma Yöntemleri* (9. Baskı) Ankara: Seçkin Yayıncılık.

Yılmaz, C. (2005). İngilizce öğretmenliği öğrencilerinin yabancı dili iletişim aracı olarak kullanabilme yeterlilikleri. Yayınlanmamış Doktora Tezi, Ondokuz Mayıs Üniversitesi, Samsun.

Yılmaz, H. (2002). *Eğitimde ölçme ve değerlendirme*. Konya: Çizgi Kitabevi.

Yörü, B. (2012). İlköğretim sekizinci sınıf İngilizce öğretim programına ilişkin öğretmen görüşleri: Eskişehir örneği. Yayınlanmamış yüksek lisans tezi, Osman Gazi Üniversitesi, Eskişehir.

Zheng, M. & Fang, Q. (2017). An Empirical Study of the Influence of Formative Assessment on Classroom Environment of College English Reading, Writing and Translating. *Advances in Computer Science Research (ACSR),* 73, 171-178

URL-1, Common core state standards initiative: Preparing America's students for college & career. http://www.corestandards.org/ adresinden 05.02.2022 tarihinde ulaşılmıştır.

EK-1

Değerli Uzman,

"Ortaokullarda İngilizce öğretiminde ölçme değerlendirme sürecinin incelenmesi" başlıklı doktora tezi kapsamında İngilizce öğretmenleriyle gerçekleştirilecek görüşmeler için sizlerin değerleri görüşlerinize ihtiyaç duymaktayım. Araştırmada ortaokul İngilizce öğretmenlerinin kullandıkları ölçme değerlendirme yaklaşım, yöntem ve araçlarının İngilizce öğretim programının hedeflerine uygunluğu açısından incelemek amaçlanmaktadır.

Araştırmanın görüşme soruları aşağıdadır. Sizden istenilen her bir görüşme sorusunu, çalışmanın kuramsal dayanakları ve amacını göz önünde bulundurarak değerlendirip görüşlerinizi belirtmenizdir.

Her bir sorunun karşısında *"Uygundur", "Uygun değildir", "Uygun ancak düzeltilmeli*" ve *"Öneri/Açıklama"* şeklinde dört seçenek bulunmaktadır. Sizden soruları inceleyerek uygun kutucuğu işaretlemeniz beklenmektedir.

Değerli katkılarınız için teşekkür ederim.

Berivan TOKUR ÜNER

SORULAR	Uygundur	Uygun değildir	Öneri/Açıklama
1. Bir İngilizce testi geliştirirken planlama yapıyor musunuz? Planlama yapıyorsanız nelere dikkat ettiğinizi açıklayabilir misiniz?			
2. Ölçme aracı geliştirirken veya bir ölçme aracını kullanmaya karar verirken geçerlik, güvenirlik ve kullanışlık özelliklerinin sağlanmış olması hakkındaki görüşleriniz nelerdir?			
3.Hazırladığınız ölçme araçlarının İngilizce öğretim programının kazanımlarına uygun olması için nelere dikkat ettiğinizi açıklar mısınız?			
4. Hangi ölçme ve değerlendirme araçlarını kullanıyorsunuz? Kullanma nedenlerinizle birlikte açıklar mısınız?			

	Uygundur	Uygun değildir	Öneri/Açıklama
5. Tamamlayıcı ölçme değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) ne ölçüde kullanıyorsunuz? Kullanıyorsanız, kullanma nedenleriniz ile birlikte açıklar mısınız?			
6. Ölçme sonuçları hakkında öğrencilere geri bildirimde bulunma konusundaki görüşleriniz nelerdir? (Sınav kağıdını öğrencilere dağıtarak, hatalarını görmelerini sağlamak gibi)			
7. İngilizce öğretim programında öngörülen tüm dil beceri alanlarının (dinleme, konuşma, okuma ve yazma) ölçülmesi ile ilgili görüşleriniz nelerdir?			
8. İngilizce öğretiminde dilbilgisi, telaffuz ve sözcük dağarcığı yapılarının ölçülmesi ile ilgili görüşleriniz nelerdir?			
9. İngilizce öğretiminde duyuşsal becerilerin (İngilizce dersini sevme, İngilizce konuşmaya istekli/isteksiz olma, dersi önemseme vb.) ölçülmesi ile ilgili düşünceleriniz nelerdir?			
10. İngilizce öğretiminde tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirmeyi hangi düzeyde ve niçin kullandığınızı açıklar mısınız?			

EK-2

Değerli Katılımcı,

"Ortaokullarda İngilizce öğretiminde ölçme değerlendirme sürecinin incelenmesi" başlıklı araştırma kapsamında, sizlerin değerli görüşlerine ihtiyaç duymaktayım. Araştırmanın amacı, ortaokul İngilizce öğretmenlerinin kullandıkları ölçme değerlendirme yaklaşım, yöntem ve araçlarını İngilizce öğretim programının hedeflerine uygunluğu açısından incelemektir. Bu amaçla sizlere ölçme ve değerlendirme süreciyle ilgili sorular yöneltilecektir. Görüşleriniz sadece araştırma kapsamında kullanılacak olup, kişisel verileriniz kimseyle paylaşılmayacaktır.

Bu çalışmaya katılımınız isteğinize bağlıdır. Araştırmanın herhangi bir aşamasında araştırmaya katılmadan vazgeçebilirsiniz, verilerinizin yok edilmesini veya size iade edilmesini talep edebilirsiniz. Yapacağımız görüşme yaklaşık olarak 20 dakika sürmektedir.

Değerli görüş ve önerilerinizle araştırmaya sunduğunuz katkılar için teşekkür ederim.

Görüşme yeri:

Tarih:

Görüşmeci: Berivan TOKUR ÜNER

KİŞİSEL BİLGİLER

Cinsiyet: Kadın □ Erkek □

Mesleki Kıdeminiz: 1-5 Yıl □ 6-10 Yıl □ 11-15 Yıl □ 16-20 Yıl □ 20+ Yıl □

Görev yapılan öğretim kurumunun statüsü: Devlet okulu □ İmam Hatip □ Özel okul □

GÖRÜŞME SORULARI

1. İngilizce dersi için bir ölçme aracı geliştirirken planlama yapıyor musunuz? Planlama yapıyorsanız nelere dikkat ettiğinizi açıklayabilir misiniz?
2. Ölçme aracı geliştirirken veya bir ölçme aracını kullanmaya karar verirken geçerlik, güvenirlik ve kullanışlık özelliklerinin sağlanmış olması hakkındaki görüşleriniz nelerdir?
3. Hazırladığınız ölçme araçlarının İngilizce öğretim programının kazanımlarına uygun olması için nelere dikkat ettiğinizi açıklar mısınız?
4. Hangi ölçme ve değerlendirme araçlarını kullanıyorsunuz? Kullanma nedenlerinizle birlikte açıklar mısınız?

5. Tamamlayıcı ölçme değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) ne ölçüde kullanıyorsunuz? Kullanıyorsanız, kullanma nedenleriniz ile birlikte açıklar mısınız?
6. İngilizce öğretim programında öngörülen tüm dil beceri alanlarının (dinleme, konuşma, okuma ve yazma) ölçülmesi ile ilgili görüşleriniz nelerdir?
7. İngilizce öğretiminde dilbilgisi, telaffuz ve sözcük dağarcığı yapılarının ölçülmesi ile ilgili görüşleriniz nelerdir?
8. İngilizce öğretiminde duyuşsal becerilerin (İngilizce dersini sevme, İngilizce konuşmaya istekli/isteksiz olma, dersi önemseme vb.) ölçülmesi ile ilgili düşünceleriniz nelerdir?
9. İngilizce öğretiminde tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirmeyi hangi düzeyde ve niçin kullandığınızı açıklar mısınız?
10. Ölçme sonuçları hakkında öğrencilere geri bildirimde bulunma konusundaki görüşleriniz nelerdir? (Sınav kağıdını öğrencilere dağıtarak, hatalarını görmelerini sağlamak gibi)
11. İngilizce dersinde nitelikli bir ölçme değerlendirme süreci için önerileriniz nelerdir?

EK-3

Değerli Uzman,

"Ortaokullarda İngilizce öğretiminde ölçme değerlendirme sürecinin incelenmesi" başlıklı doktora tezi kapsamında, öğretmenlerin İngilizce becerilerini ölçerken ve değerlendirirken belirlenen maddeleri ne düzeyde gerçekleştirdiklerini öğrenmek için bir ölçek geliştirmek istemekteyim. Bu amaçla sizlerin değerleri görüşlerine ihtiyaç duymaktayım.

Araştırma ölçeği aşağıdadır. Öğretmenlerden belirlenen maddelerin karşısındaki *"Her zaman", "Çoğu zaman", "Ara sıra", "Nadiren"* ve *"Hiçbir zaman"* seçeneklerinden birini seçmeleri beklenecektir. Sizden istenilen de her bir maddeyi, çalışmanın kuramsal dayanakları ve amacını göz önünde bulundurarak değerlendirip görüşlerinizi belirtmenizdir.

Her bir sorunun karşısında *"Uygundur", "Uygun değildir"* ve *"Öneri/Açıklama"* şeklinde üç seçenek bulunmaktadır. Sizden soruları inceleyerek uygun kutucuğu işaretlemeniz beklenmektedir.

Değerli katkılarınız için teşekkür ederim.

Berivan TOKUR ÜNER

MADDELER	Uygundur	Uygun değildir	Öneri/Açıklama
A)İngilizce dinleme becerilerini değerlendirirken;			
1.Programda öngörülen öğrenci seviyesine uygun metinler seçmeye dikkat ederim			
2.Programda dinleme becerilerini ölçmek için önerilen ölçütleri kullanırım.			
3.Uzun dinleme metinlerini seçmekten kaçınırım.			
4.Dinleme becerisinin öğrenme sürecinin bir parçası olmasına dikkat ederim.			
5. Dinleme öncesi ölçme konusu ile ilgili gerekli yönergeyi öğrencilere sunarım.			
6.Dinleme öncesinde/sonrasında öğrencilere sözlü sorular yöneltirim.			
7.Dinleme öncesinde/sonrasında öğrencilere yazılı sorular yöneltirim.			
8.Dinleme sonrası öğrencilerin dilbilgisi yapılarına ne derece hakim olduklarını anlamaya çalışırım			

MADDELER	Uygundur	Uygun değildir	Öneri/Açıklama
9. Dinleme sonrası öğrencilerin sözcük dağarcıklarının zenginliğine odaklanırım.			
10. Öğretim programında kullanılması önerilen test tekniklerini kullanırım.			
11. Uygun bir dinleme ortamı oluşturmada güçlükler yaşarım.			
12. Yeterli bilgi ve beceriye sahip olmadığımı düşünürüm.			
B) İngilizce konuşma becerilerini değerlendirirken;			
1. Öğrencilerin kendilerini rahat hissedecekleri bir ortam hazırlarım.			
2. Programda konuşma becerilerini ölçmek için önerilen ölçütleri kullanırım.			
3.İngilizce dersi öğretim programında önerilen ölçme ve değerlendirme araçlarını kullanmaya özen gösteririm.			
4. Konuşma becerisinin öğrenme sürecinin bir parçası olmasına dikkat ederim.			
5. Gözlem kontrol listesi geliştiririm.			
6. Öğrencilerin anlattıklarının doğru ve anlaşılır olup olmadığına dikkat ederim.			
7. Öğrencilerin konuşmalarının akıcılığına dikkat ederim.			
8. Öğrencilerin konuşma amacına uygun kelimeleri kullanmalarını önemserim.			
9. Öğrencilerin kurdukları cümlelerin kurallı olmasına önem veririm.			
10. Kelimelerin doğru telaffuz edilmesine dikkat ederim.			
11. Her bir öğrenciye benzer güçlükte sorular sormaya dikkat ederim.			
12. Öğrencilere konuşma performanslarına dair dönüt veririm.			
13. Uygun ölçme araçları geliştirmede zorluklar yaşarım.			
14. Puanlamanın objektif olmasına dikkat ederim.			
C) İngilizce okuma becerilerini değerlendirirken;			
1.Metinlerin seçiminde öğrencilerin seviyelerini göz önünde bulundururum.			

MADDELER	Uygundur	Uygun değildir	Öneri/Açıklama
2. Soruların kazanımları ölçecek niteliklere sahip olmasına dikkat ederim.			
3. Otantik okuma metinlerini kullanmaya özen gösteririm.			
4. Okuma becerisinin öğrenme sürecinin bir parçası olmasına dikkat ederim.			
5. İngilizce dersi öğretim programında önerilen ölçme ve değerlendirme araçlarını kullanmaya özen gösteririm.			
6. Tamamlayıcı ölçme araçlarını (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) değerlendirme sürecine dahil ederim.			
7. Okuma öncesi/sonrası öğrencilere sözlü sorular yöneltirim.			
8. Okuma öncesi/sonrası öğrencilere yazılı sorular yöneltirim.			
9. Okuma sonrası öğrencilerin dilbilgisi yapılarına uygun cümleler kurmalarına dikkat ederim.			
10. Okuma sonrası öğrencilerin kelime bilgisini göz önünde bulundururum.			
D) İngilizce yazma becerilerini değerlendirirken;			
1.Ölçme aracının kazanımları ölçebilecek niteliğe sahip olmasına dikkat ederim.			
2. Öğrencilerin farklı yazın türlerine dair metinler oluşturmalarını önemserim.			
3. Yazma becerisinin öğrenme sürecinin bir parçası olmasına dikkat ederim			
4.Öğrencilere yazma türü, konusu, metin uzunluğu vb. ile ilgili detaylı bir yönerge sunmaya özen gösteririm.			
5. Ölçme aracındaki soruların öğrencilerin düzeyine uygun olmasını göz önünde bulundururum.			
6. Öğrencilerin dilbilgisi yapılarını uygulamaya aktarmalarına dikkat ederim.			
7. Noktalama kurallarına dikkat ederim.			
8. Yazım kurallarına dikkat ederim.			
9. Geleneksel ve tamamlayıcı ölçme araçlarını birlikte kullanmaya dikkat ederim.			

MADDELER	Uygundur	Uygun değildir	Öneri/Açıklama
10. Akran değerlendirme yaparım.			
11. Öğrencilerin öz-değerlendirme yapmalarını sağlarım.			
12. Puanlamanın objektif olmasına dikkat ederim.			
13. Öğrencilere ölçme sonuçları hakkında geri bildirim sağlarım.			

EK 4- Ölçme Aracı

Değerli Katılımcı,

"Ortaokullarda İngilizce öğretiminde ölçme değerlendirme sürecinin incelenmesi" başlıklı doktora tezi kapsamında sizlerin değerli görüşlerine ihtiyaç duymaktayım. Araştırmanın amacı, İngilizce öğretmenlerinin öğrencilerin yabancı dilde dinleme, konuşma, okuma ve yazma performanslarını değerlendirirken hangi süreçlere ne ölçüde dikkat ettiklerini incelemektir. Bu amaçla sizlere yabancı dilde ölçme ve değerlendirme süreciyle ilgili sorular yöneltilecektir. Çalışmanın amacına hizmet edebilmesi için aşağıda yer alan sorulara samimi ve eksiksiz bir şekilde yanıt vermeniz beklenmektedir. Görüşleriniz sadece araştırma kapsamında kullanılacak olup, kişisel verileriniz kimseyle paylaşılmayacaktır. Değerli görüşlerinizle araştırmaya yapacağınız katkılar için teşekkür ederim.

KİŞİSEL BİLGİLER

Cinsiyet: Kadın □ Erkek □

Mesleki Kıdeminiz: 1-5 Yıl □ 6-10 Yıl □ 11-15 Yıl □ 16-20 Yıl □ 21 Yıl ve üzeri □

Öğrenim Durumunuz: Lisans □ Yüksek Lisans □ Doktora □

Görev yapılan öğretim kurumunun statüsü: Devlet okulu □ İmam Hatip □ Özel okul □

Ölçme ve değerlendirme ile ilgili bir hizmet içi eğitim seminerine/kursa katıldınız mı?
Evet □ Hayır □

Berivan TOKUR ÜNER

MADDELER	**Her zaman**	**Çoğu zaman**	**Ara sıra**	**Nadiren**	**Hiçbir zaman**
A)İngilizce dinleme becerilerini değerlendirirken;					
1.İngilizce dersi öğretim programındaki kazanımlara uygun metinler seçerim.					
2.Öğrencilerin dikkat sürelerini aşan metinler seçmemeye dikkat ederim.					
3.Sınıf içi etkinlikler sırasında öğrencilerin dinleme becerileriyle ilgili değerlendirmeler yaparım.					
4. Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım					

MADDELER	Her zaman	Çoğu zaman	Ara sıra	Nadiren	Hiçbir zaman
5. Öğrencilerin dilin ses, sözcük gibi öğelerine ne ölçüde hakim olduklarını değerlendiririm.					
6.Öğrencilerin dilbilgisi kuralları ve kalıplarını ne ölçüde anladıklarını değerlendiririm					
7.Öğrencilerin dinleme metinlerini ne kadar anladıklarıyla ilgili değerlendirmeler yaparım					
8.İngilizce dersi öğretim programında önerilen test tekniklerini kullanırım					
9. Tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) kullanırım					
10. Sınıf içi etkinlikler yapılırken veya sınav esnasında uygun bir dinleme ortamı oluşturmada güçlük yaşarım					
B) İngilizce konuşma becerilerini değerlendirirken;					
1. İngilizce dersi öğretim programındaki kazanımlara uygun bir ölçme aracı hazırlarım.					
2. Sınıf içi etkinlikler sırasında öğrencilerin konuşma becerileriyle ilgili değerlendirmeler yaparım					
3. Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım					
4. İngilizce dersi öğretim programında önerilen test tekniklerini kullanırım					
5. Öğrencilerin konuşmalarının doğru ve anlaşılır olup olmadığını değerlendiririm.					
6. Öğrencilerin bilgiyi akıcı bir şekilde aktarıp aktaramadıklarını değerlendiririm					
7. Öğrencilerin gerekli sözcükleri kullanabilme yeteneklerini değerlendiririm					
8. Sözcüklerin doğru telaffuz edilmesine dikkat ederim					
9. Öğrencilerin dilbilgisi kuralları ve kalıplarını ne ölçüde kullanabildiklerini değerlendiririm					
10. Tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) kullanırım					
11. Sınıf içi etkinlikler yapılırken veya sınav esnasında uygun bir konuşma ortamı oluşturmada güçlük yaşarım.					

MADDELER	Her zaman	Çoğu zaman	Ara sıra	Nadiren	Hiçbir zaman
C) İngilizce okuma becerilerini değerlendirirken;					
1.İngilizce dersi öğretim programındaki kazanımlara uygun olan metinleri seçerim					
2. Sınıf içi etkinlikler sırasında öğrencilerin okuma becerileriyle ilgili değerlendirmeler yaparım					
3. Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım					
4. İngilizce dersi öğretim programında önerilen ölçme ve değerlendirme araçlarını kullanırım.					
5. Öğrencilere farklı yazın türlerine dair okuma metinleri sunarım					
6. Öğrencilerin kelimeleri ve söz kalıplarını ne ölçüde tanıdıklarını değerlendiririm					
7. Öğrencilerin dilbilgisi kuralları ve kalıplarını uygun bir şekilde kullanıp kullanmadıklarını değerlendiririm.					
8. Öğrencilerin okuma metnindeki ana fikri belirleme, neden-sonuç bulma gibi ilişkileri nasıl yorumladıklarını değerlendiririm.					
9. Tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) kullanırım					
10. Sınıf içi etkinlikler yapılırken veya sınav esnasında uygun bir okuma ortamı oluşturmada güçlük yaşarım.					
D)İngilizce yazma becerilerini değerlendirirken					
1.İngilizce dersi öğretim programındaki kazanımlara uygun bir ölçme aracı hazırlarım.					
2. Sınıf içi etkinlikler sırasında öğrencilerin yazma becerileriyle ilgili değerlendirmeler yaparım.					
3. Tanılayıcı, biçimlendirici ve düzey belirleyici değerlendirme tekniklerinden yararlanırım					
4.Öğrencilere yazma türü, konusu, metin uzunluğu vb. ile ilgili bilgileri içeren detaylı bir yönerge sunarım					
5. Öğrencilerin farklı yazın türlerine dair metinler oluşturmalarını isterim					
6. Öğrencilerin yazma amacına uygun olan kelimeleri ne ölçüde kullandıklarını değerlendiririm					

MADDELER	Her zaman	Çoğu zaman	Ara sıra	Nadiren	Hiçbir zaman
7.Öğrencilerin yeterli bir hızda yazıp yazmadıklarını değerlendiririm					
8. Öğrencilerin dilbilgisi kuralları ve kalıplarını uygun bir şekilde kullanıp kullanmadıklarını değerlendiririm					
9. Yazım kurallarına dikkat ederim					
10. Tamamlayıcı ölçme ve değerlendirme araç ve tekniklerini (performans, portfolyo değerlendirme, akran değerlendirme, öz değerlendirme, kontrol listeleri, dereceli puanlama anahtarı, gözlem vb.) kullanırım					
11. Sınıf içi etkinlikler yapılırken veya sınav esnasında uygun bir yazma ortamı oluşturmada güçlük yaşarım.					

EK-5 ETİK KURUL İZNİ

Evrak Tarih ve Sayısı: 08/03/2021-37468

T.C.
DİCLE ÜNİVERSİTESİ REKTÖRLÜĞÜ
Hukuk Müşavirliği

K-Q TSE-ISO-EN 9000

Sayı : E-14679147-663.05-37468
Konu : İnceleme (Projenin Değerlendirilmesi)

Sayın Berivan TOKUR ÜNER
DİCLE ÜNİVERSİTESİ EĞİTİM BİLİMLERİ,ENSTİTÜSÜ,EĞİTİM BİLİMLERİ ANABİLİM DALI
EĞİTİM POGRAMLARI VE ÖĞRETİM BİLİM DALI SUR/DİYARBAKIR

"Ortaokullarda İngilizce Öğretiminde Ölçme Değerlendirme Sürecinin İncelenmesi" başlıklı çalışmanız Yükseköğretim Kurumları Bilimsel Çalışma ve Yayın Etiği Yönergesi uyarınca Üniversitemiz Sosyal ve Beşeri Bilimler Etik Kurulu Başkanlığı tarafından değerlendirilmiş olup söz konusu çalışmanın bilimsel etik açısından uygun olduğuna ilişkin Üniversitemiz Sosyal ve Beşeri Bilimler Etik Kurulu Başkanlığı kararı Üniversitemiz Rektörlük Makamının 07.03.2021 tarih ve 36561 sayılı Olur'u ile uygun görülmüştür.

Bilgilerini rica ederim.

Av. Rengin AVCI
Hukuk Müşaviri V.

Ek: İlgili Belgeler(1 sayfa)

Bu belge, güvenli elektronik imza ile imzalanmıştır.
Belge Doğrulama Kodu :BE6E61DVP Pin Kodu :62612 Belge Takip Adresi : https://www.turkiye.gov.tr/dicle-universitesi-ebys
Adres:Dicle Üniversitesi Rektörlüğü, 21280-Diyarbakır
Telefon:+90 412 241 10 00 Faks:+90 412 241 10 56
e-Posta:gensek@dicle.edu.tr Elektronik Ağ:http://www.dicle.edu.tr
Kep Adresi: dicleuniversitesi@hs01.kep.tr
Bilgi için: Erkan Seyrek
Unvanı: Büro Personeli
Tel No: 2237

EK-6 YABANCI DİL BECERİLERİNİ ÖLÇME VE DEĞERLENDİRME ANKETİ UYGULAMA İZNİ

T.C.
DİYARBAKIR VALİLİĞİ
İl Millî Eğitim Müdürlüğü

Sayı : E-30769799-44-22849742
Konu : Araştırma İzni (Berivan TOKUR)

23.03.2021

MÜDÜRLÜK MAKAMINA

İlgi: a) MEB Yenilik ve Eğitim Teknolojileri Genel Müdürlüğünün 24/01/2020 tarih ve 1563890 sayılı 2020/2 Nolu Genelgesi.
b) Dicle Üniversitesi Rektörlüğünün 14/03/2021 tarih ve 41149 sayılı yazısı.

Dicle Üniversitesi Eğitim Bilimleri Enstitüsü Eğitim Programları ve Öğretim Bilim Dalı Doktora Öğrencisi Berivan TOKUR'un **"Ortaokullarda İngilizce Öğretiminde Ölçme Değerlendirme Sürecinin İncelenmesi "** konulu araştırma çalışmasını İlimiz ilçelerine bağlı resmi ve özel ortaokullarda görev yapan ingilizce öğretmenlerine uygulama talebi ilgi (b) yazıda belirtilmiştir.

Söz konusu araştırma çalışmasının **Türkiye Cumhuriyeti Anayasası, Milli Eğitim Temel Kanunu ile Türk Milli Eğitiminin Genel amaçlarına uygun olarak ilgili yasal düzenlemelerde belirtilen ilke,esas ve amaçlara aykırılık teşkil etmeyecek şekilde** hijyen, sosyal mesafe ve maske kuralına uymak kaydıyla araştırma çalışmasının gönüllülük esasına bağlı olarak yapılması uygun görülmektedir.

Makamlarınızca da uygun görülmesi halinde olurlarınıza arz ederim.

Süleyman İLGE
Müdür Yardımcısı

OLUR
Yüksel ARSLAN
İl Milli Eğitim Müdürü

EK:
1- Anket İzin Değerlendirme Formu
2- Anket Belgeleri

Bu belge güvenli elektronik imza ile imzalanmıştır.
Adres : ŞEHİTLİK MAHALLESİ ESKİ EĞİTİM FAKÜLTESİ BİNASI YENİŞEHİR/DİYARBAKIR
Telefon No : 0 (412) 322 22 36
E-Posta: arge21@meb.gov.tr
Kep Adresi : meb@hs01.kep.tr
Belge Doğrulama Adresi : https://www.turkiye.gov.tr/meb-ebys
Bilgi için: HATİP YAVUZ /MEMUR
Unvan : Memur
İnternet Adresi: diyarbakır@meb.gov.tr
Faks:4123222248
Bu evrak güvenli elektronik imza ile imzalanmıştır. https://evraksorgu.meb.gov.tr adresinden a71d-569d-3f98-9ef9-7d9d kodu ile teyit edilebilir.