

ÇOKTAN-SEÇMELİ ÖLÇÜNLENDİRİLMİŞ SINAVLARIN GEÇERLİLİK ÖLÇÜTÜ BAĞLAMINDA DEĞERLENDİRİLMESİ

Muazzez Yavuz KIRIK ^(*)

ÖZET

Yabancı dilde başarı ya da yeterliliğe dair kararların geçerlilik kazanması alana özgü ölçütlerin sağlanmasına bağlıdır. Ölçme değerlendirme etkinlikleri bir ölçütler oluşturma etkinliği olarak öğretim içerik ve uygulamalarının, hedeflere ulaşmayı sağlayan nitelikler taşıyıp taşımadıklarının göstergeleridir. Ölçme değerlendirme süreci aynı zamanda kararlar alma sürecidir. Yabancı dilde yeterlilik düzeyiyle ilgili kararların geçerliliği, ölçme değerlendirme içerik ve tekniklerinin iletişimsellik ve özgünlük düzeyiyle ilgilidir. Bu bağlamda, çoktan-seçmeli soru yapılarını içeren ölçünleştirilmiş sınavların sınava giren adayların yaşamındaki belirleyiciliği göz önünde bulundurulduğunda sorgulanmaları anlam kazanmaktadır. Bu çalışmada, çoktan-seçmeli soru tekniğiyle hazırlanmış bir ölçünleştirilmiş sınav örneği olarak 2009-YDS sınavının geçerlilik ölçütü bağlamında değerlendirilmesi amaçlanmıştır.

Anahtar Sözcükler: Yabancı Dil Öğretiminde Ölçme Değerlendirme Teknikleri, Ölçünleştirilmiş Sınavlar, Çoktan-seçmeli Sorular, Geçerlilik Ölçütü.

ABSTRACT

In the field of foreign language teaching, validating the decisions about achievement or proficiency in language depends on providing criteria that are specific to domain. Assessment tasks, as the process of establishing criteria in language teaching are the indicators to see whether teaching content and practices carry the qualities that help to attain the aims and goals desired. Testing at the same time is a process of decision making. In foreign language testing, validity of the decisions about the level of proficiency is closely related to the level of authenticity of the test

(*) İstanbul Üniversitesi, Hasan Ali Yücel Eğitim Fakültesi, Yabancı Diller Eğitim Bölümü, İngiliz Dili Eğitimi Anabilim Dalı.

techniques and communicative nature of language use in content. In this respect it would be meaningful to question these standardized tests including multiple-choice questions because their results have decisive influence on candidates' future life and education. This study aims to evaluate the 2009-YDS (University entrance exam) as an example of standardized test in which questions are designed in the form of multiple-choice.

Key Words: *Testing Techniques in Foreign Language Teaching, Standardized Testing, Multiple-choice Questions, Validity.*

Yabancı dil olarak İngilizce'de 'başarı' ya da 'yeterlilik' kararları almak üzere uygulanan sınavlar bireyin yaşamını büyük oranda belirleyici özellikler taşımaktadır. Ulusal ya da uluslar arası ölçünlendirilmiş sınavlardan alınan sonuçlar bireylerin kimliklerini tanımlamaktadır. Bu tür sınavlar, 'başarı' ya da 'yeterlilik' kararlarıyla bireye dair bir 'değer biçme' işlevi taşımaktadır. Siyasal, toplumsal değerler ve tutumlar adı geçen sınavların kullanım amaçlarını belirlerken aynı zamanda tüm bu değerler sınavlar yoluyla yeniden doğallaştırılmaktadır. Dünya diline dönüşmekte olan İngilizce'de 'yeterliliğin' siyasal, sosyal ve ekonomik düzlemlerde bireylerin yaşamında ne denli etkin ve belirleyici olduğu basit gözlemlerle bile fark edilebilir bir olgudur.

Yabancı dil olarak İngilizce'de başarı ya da yeterliliği belirlemeye yönelik sınavların bireyin yaşamındaki rolü, etkisi ve sonuçları göz önünde bulundurulduğunda, sınav yoluyla alınan kararların geçerliliği tartışması önem kazanmaktadır. Geçerlilik kavramı, yabancı dil olarak İngilizce öğretim ve öğrenimine yönelik tüm süreçlerdeki belirleyiciliğinin yanı sıra ulusal ve uluslar arası ölçünlendirilmiş sınavların bireylerin yaşamındaki belirleyiciliği bağlamında da tartışılmalıdır. Yani, geçerlilik tartışmaları sınav sonuçlarının da göz önünde bulundurulduğu bir yeni geçerlilik kavramı oluşturma çalışmasına dönüştürülmelidir.

Sınavlarda alınan puanların hangi bağlamda başarı ya da yeterliliğin göstergesi olduğu tartışması bir geçerlilik tartışması olarak kabul edilirken, puanlara göre alınan kararların birey yaşamında yarattığı sonuçlara yönelik tartışma, geçerlilik tartışmalarında yer almamaktadır. Oysa ölçme değerlendirme araçları olarak ulusal/uluslararası ölçünlendirilmiş sınavların yabancı dil olarak İngilizce öğretimi/öğrenimi bağlamında yöntem, içerik ve uygulamalarına olan etkisinin yanı sıra bireyin toplumsal yaşamında bir 'değer biçme' aracı olarak gördüğü işlevi de göz önünde bulundurmak gerekir.

Ölçme değerlendirme etkinliği olarak sınavlarda kullanılan teknikler, adaya ilişkin hangi bilginin elde edilmek istendiği ile ilgili olarak belirlenir.

Yabancı dil olarak İngilizce’de başarı ya da yeterlilik tanımları kurumların gerekliliklerine göre değişebilir.

Ölçülendirilmiş Sınavlar

Dil kullanımı, gerek yazılı gerek sözlü, bir eylemlilik bir etkinliktir. Etkin dil kullanımı gerçek yaşam durumlarında ve bağlamlarında uygunluğu da gözeten birçok stratejiyi içerir. Günümüzde yabancı dilde yeterlilik tanımları iletişimsel dil kullanımı tanımı üzerinden yapılmakta ve bağlam ve uygunluk önemli ölçütler olarak öne çıkmaktadır. Yani, özgün ortamlara uygun etkileşim ve etkili dil kullanımı yabancı dilde yeterliliğin genel ölçütleri olarak sınav içeriklerinde belli tekniklerle yansıtılmakta ve geçerlilik ölçütleri olarak kullanılmaktadır.

“Yabancı dil olarak İngilizce’de yeterlilik kararları iletişim ortamlarında gerçekleşen dil kullanım etkinliklerinin (yazılı, sözlü) gözlemine dayandırılmalıdır” (McNamara, Roever, 2006:25). Gözlenebilir dil kullanım etkinlikleri ancak böylesi doğrudan tekniklerle geçerli bir şekilde ölçülebilir.

“Özellikle ölçülendirilmiş sınavların daraltılmış içeriklerinde olduğu gibi belli bir konu, başlık ya da bağlam üzerinden gerçekleştirilen sözlü/yazılı etkinliklerdeki yeterlilik göstergeleri diğer bağlamlarda yeterlilik göstergeleri olamazlar” (McNamara, Roever, 2006:26). Çünkü bilinmektedir ki ölçme değerlendirme süreçlerinde elde edilen puanlar/sonuçlar sınavın gerçekleştiği ortam, sınav uygulayıcıları, sınav içeriği, kullanılan teknikler, ürünü değerlendirenler, değerlendirme ölçütleri ve öğrencinin art-alan bilgisi gibi birçok etmene göre değişebilir. Bu durumda bireyin performansına dair çıkarsamaları tek bir sınav sonucuna dayandırmak, bireyin birçok farklı durum ve bağlamda göstereceği performansı yansıtamayacağı için geçerli bir yaklaşım olamayacaktır.

Bir diğer geçerlilik sorgulaması, belli zaman, kural, içerik, teknik ve ölçütlere sıkıştırılmış sınav uygulamalarında dil etkinlikleri bağlamında gösterilen performans ile gerçek yaşam durumlarında gösterilen performans arasındaki örtüşmenin düzeyi ile ilgili olmalıdır (McNamara, Roever, 2006:27). Her anlamda tanımlı bir süreç olarak sınavlarda gösterilen başarı/başarısızlık ya da yeterlilik/yetersizlikleri, gerçek yaşam durumlarında ki performansın göstergesi olarak değerlendirmek alınacak kararların geçerliliğini tartışmalı hale getirebilir.

Yabancı dil olarak İngilizce’de yeterliliği ölçme değerlendirmeye yönelik ulusal ya da uluslar arası ölçülendirilmiş sınavların üniversitelere giriş sınavı olarak kullanılması, bireyin/adayın yaşamındaki belirleyiciliği bağlamında bu sınavları çok önemli bir yere taşımaktadır (McNamara, Roever, 2006:28). Bu nedenle, ‘dışarıda bırakan, içeri alan, değer biçen’ işlevleriyle bu sınavların, topluma, bireye/adaya karşı sorumlulukları bağlamında da sorgulanmaları gerekmektedir. ‘İçeri alma, dışarıda bırakma ve değer biçme’ kararlarının geçerliliğini sorgulamak sınavların görünür, resmi işlevlerinin ötesinde ne tür işlevler gerçekleştirdiklerini de görmeyi kolaylaştıracaktır.

Geçerlilik

Geçerlilik, birleşik bir kavram olarak (unitary concept) sınav sonuçlarına dayanan çıkarsamaları/kararları destekleyecek delillerin sağlanmasına dayanır. Örneğin gerçek yaşam durum ve bağlamlarında uygun doğal dil kullanımında yeterlilik bir ölçüt olarak alındığında bu ölçütü değerlendirebilecek ortam/içerik ve tekniklerin sunulmadığı sınav sonuçlarına dayandırılan bireyin performansına dair çıkarsamalar/kararlar geçerli olamazlar.

Birçok bağlamda doğal ve uygun dil kullanarak etkileşimde bulunma becerisi ya da düzeyi, aynı koşulların sağlanması ve buna uygun teknik/araçların sunulması ile geçerli bir şekilde ölçülüp değerlendirilebilir. Sözlü yazılı düzlemlerde olmak üzere anlama ve anlatma yetilerinden oluşan iletişim yetisi de bu etkinliklerin doğasına uygun araçlar/teknikler geliştirilerek ölçülüp değerlendirilebilir.

Geçerlilik ölçüleceği iddia edilen şeyin ölçülme düzeyidir (Spolsky, 1990:4). Ölçülmek istenen şey ölçülüyorsa o sınav geçerlidir (Brown, 2005:220). İstenen bilgi elde edilen bilgiyi ne denli yansıtır sorusu geçerlilik sorusudur (Genesee, Upshur, 1996:62). Ölçme değerlendirme işlemlerinde birçok değişkenin etkisi geçerliliği sekteye uğratabilir (Harrison, 1983:11). Bir sınav içeriğinin geçerliliği, öncesinde yer alan öğretim içerik ve uygulamalarının doğru bir şekilde yansıtılmasıyla sağlanabilir (Brown, Hudson, 2002). Yani, yeterlilik sınavları öncesinde yer alan öğretime denk düşmelidir (Weir, 1993).

Sınava dair kararlar almak her biri farklı işlemler dizgesini gerektirir. Geçerlilik yaklaşımları olarak adlandırılabilir bu işlem dizgeleri üç temel alanda tanımlanabilir; İçerik uygunluğu ve bağlantısı, ölçütü ilgili geçerlilik ve yapı geçerliliği. Geçerlilik tartışması bu üç ölçütün gerekliliklerini tartışmaktan geçmektedir. Dolayısıyla her bir ölçütün tanımlarını netleştirmek sınavların belirlenen bağlam ve koşullarda ne kadar geçerli olduklarını görmeye yardımcı olacaktır.

İçerik uygunluğu ölçütü sınavın amacıyla ilintilidir. Özellikle sınıf içi değerlendirmelerde öğrencinin gelişimini görmeye yönelik tanısıl sınavlarda amaç, müfredatta belirlenen hedefleri öğretim içerik ve uygulamalarıyla örtüşen içerik ve tekniklerle ölçmektir. Bu amacı gütmeyen sınav içerik ve yöntemleri geçerlilik ölçütünün sağlanıp sağlanmadığının delillerini sunamaz.

Bu bağlamda, ölçülendirilmiş sınavlar içerik uygunluğu tartışmasının dışında kalmaktadır (Hasselgreen, 2004; McNamara, 2000). Geçerlilik dizgelerinden bir diğeri ölçütü ilgili geçerlilik ya da ölçüte bağlılık kavramıdır. Geçerlilikte ölçüte bağlılık düzeyi, bir ölçme değerlendirme aracının paralel içerik ve yapıdaki diğer bilinen ölçme aracının ölçütleriyle benzerlik düzeyine göre belirlenir. Yani, ölçüte bağlılık bir sınavın sonuçlarıyla, diğer ölçme araçlarıyla elde edilen sonuçlar arasındaki ilişkiyle ilgilidir (Davies, Brown, Elder, Hill, Lumley ve McNamara, 1999).

Niteliklerle ilgili elde edilecek bilginin türü ve bu bilginin ne türde yöntemlerle değerlendirileceği tartışması ölçüt belirlemeyi gerektirir. Bir sınavın geçerliliğini irdelemek birden fazla etmeni göz önünde bulundurmakla mümkündür. Ölçütle ilgili geçerliliğin sağlanması için ölçüt alınacak sınavların seçimi önemlidir. Bu bağlamda, çok dikkatle incelenmesi gereken nokta öğretim hedefi, içerik ve yöntemlerinin ölçüt alınacak sınav içerik ve yöntemleriyle örtüşme düzeyidir. Bu noktanın göz ardı edilmesi durumunda elde edilen bilgi ve buna dayanılarak alınan kararlar yanıltıcı olabilmektedir.

Yapı geçerliliği tartışması bir yabancı dil sınavının belli bir dil öğretim kuramını yansıtırma düzeyini belirlemeye yöneliktir (Davies, Brown, Elder, Hill, Lumley, McNamara, 1999). Yapı geçerliliği bir yabancı dil sınavının ölçülmesi planlanan yapıyı ölçüp ölçmediği ile ilgilidir (Cumming, 1996; Hasselgreen, 2004; McNamara, 2000; Brown, Hudson, 2002). Sınav sonuçları, yabancı dil becerilerinde ölçülmek istenen başarı düzeyini her bir becerinin ifade araçlarını kullanarak ölçtüğünün delillerini sağlamalıdır. Bu nedenle sınav sonuçların değerlendirilmesi/yorumlanması tutarlı ölçütlere göre yapılmalıdır (Bachman, Palmer, 1996:21). Yapı geçerliliği, ölçülmek istenen becerinin tümüyle ölçüldüğünün göstergesidir. Dolayısıyla her becerinin alt bileşenleri için farklı yapıda sınav hazırlamak gerekir. Örneğin yazma becerisini çoktan-seçmeli bir teknikle ölçmek ya da konuşma becerisini konuşurmayı gerektiren doğrudan, performansla dayalı tekniklerin ötesinde tekniklerle ölçmek o sınavın geçerli olmadığını gösterir (Hughes, 1989:26-7).

Yeterlilik ya da başarıyı belirlemeye yönelik yabancı dil sınavlarının niteliklerini belirlemede bir ölçüt olarak yapı geçerliliği ölçülendirilmiş sınavların niteliklerini yorumlamakta önemli bir ölçüttür. Çünkü adayların içinden geçtikleri dil öğretim içerik ve yöntemleriyle ya da diğer deyişle belli yabancı dil öğretim programlarının içerik ve uygulamalarıyla örtüşmeyen ölçülendirilmiş sınav içerikleri adayın gösterebileceği başarı düzeyini etkilemektedir (Cushingweigle, Lynch, 1996:58-71).

Yabancı dil olarak İngilizce öğretim programları yeterlilik tanımlarını belirli bir dil kuramına bağlı olarak oluştururlar. Örneğin yabancı dil öğretim programında iletişimsel edinç tanımı üzerinden yeterlilik hedefleri belirlenmişse özgün ortamlarda, doğal dil kullanım becerilerini kazandıran öğretim içerikleri oluşturulur. Hedef ve uygulama arasındaki tutarlılık aynı oranda ölçme değerlendirme tekniklerine de yansıtılır. Yabancı dil olarak İngilizce sınavları dil becerilerindeki başarıyı her bir becerinin ifade araçlarını kullanarak yansıttığının delillerini sağlamalıdır. Bu bağlamda, örneğin yazma ve konuşma becerilerinde 'yeterlilik/başarı' düzeyini çoktan seçmeli yapıda hazırlanmış ölçme teknikleriyle değerlendirmek sorgulanması gereken bir uygulamadır.

Ölçülendirilmiş yabancı dil sınavları kesin ve belirli standart hedefleri içerir. Ulusal ve uluslar arası ölçülendirilmiş yabancı dil sınavları belirli bir yabancı dil öğretim programının müfredat içeriğine göre hazırlanmayan, geniş bir tanımlamayla dil yetisini ölçmek için oluşturulan sınavlardır. Ölçülendirilmiş

sınavlarda uygulama ve puanlama için standart işlemler tanımlanır ve sınava giren adayların başarı düzeyleri adayların başarıları karşılaştırılarak belirlenir (Davies, Brown, Elder, Hill, Lumley, McNamara, 1999).

Ölçünleştirilmiş sınavlar genellikle öğrenci kabulü için Üniversitelerde, gerekli alan ve birimlere eleman alımı amacıyla ticari kurumlar ve devlet kuruluşlarında ekonomik ve güvenilir olduğu varsayımıyla kullanılmaktadırlar. “Ölçünleştirilmiş sınavların gerekliliğini savunanlar aynı anda çok sayıda insanın eşit koşullarda test edilebilmesi özelliğini gerekçe olarak göstermektedirler” (Brown, 2004:66). Ölçme değerlendirme işlemlerinin ölçünleştirilmesi bir geçerlilik görüntüsü yaratmaktadır.

Öğrenciye ilişkin edinilmesi zorunlu olan verilerin sağlanabilmesi çok farklı türde ölçme değerlendirme araçlarının geliştirilmesini gerektirir. Sınav içerik ve tekniklerinin oluşturulması sürecinde sadece tek bir tekniğin (çoktan-seçmeli) kullanılması, eğitim alanında ayrıcalıklı adaylar yaratmaktadır. Bazı adayların öğrenme ve bilginin aktarımı konusunda diğer adaylardan daha ayrıcalıklı hale gelmeleri bu süreçle pekiştirilen bir olgudur. Sınavlara göre hazırlanmış müfredat hedefleri, öğretim uygulamaları ve sınıf içi değerlendirmeler, bu tür programlarda yer alan adayları diğerlerine göre daha şanslı kılmaktadır.

Bu olguyu en iyi açıklayan örnek Türkiye’de Üniversite sınavlarına hazırlık programları olarak işlev gören özel dershanelerdir. Öğretim içerik ve uygulamalarını Yabancı Dil Sınavının (YDS) içerik ve soru tekniklerine göre oluşturan özel ‘dershaneler’ den yararlanabilen öğrenciler bu tür programlardan yararlanamayan öğrencilere göre sınavda daha şanslı olabilmektedir. Çoktan-seçmeli soru tekniğiyle oluşturulan sınavların tek tipliliği, nesnelliği ve ‘eşitlikçiliği’ gerçek dünyanın eğitimdeki ‘eşitlikçiliğini’ yansıtmaktadır.

Bilginin üretim ve aktarımında belli guruplara verilmiş ayrıcalıklar ya da toplumsalın önemli addettiği konuları öğrenme zorunluluğu gibi adil olmayan farklılıklar ölçme araçlarının yeniden yapılanmasını, farklı türlerde ölçme araçlarının geliştirilmesini gerekli kılmaktadır (Neill, 2003:1-15). Örneğin meslek okullarında müfredat içeriğinin daraltılması ve sınırlandırılması beklentide düşüklük ve sınırlılığa yol açmakta, çoğunluğu düşük gelir gurubundan gelen öğrencilerin gelişimini engelleyebilmektedir.

Öğretim ve değerlendirme uygulamaları öğrencinin öğrenme sürecini ve tarzını belirlemek için kullanıldığında beklentiler yüksek tutulabilir ve öğrenciye destek daha bilinçli verilebilir. Bu desteği verebilmenin yolu öğrencinin duygusal, zihinsel ve sosyal art-alan bilgisine ulaşmayı sağlayan öğrenciye/bireye dayalı ölçme değerlendirme ile mümkündür.

Sınav teknik ve içeriğinin biçimlendirdiği öğretim, sınava dayalı öğretim olarak tanımlanabilir. Başarı düzeyi, öğrenme isteği, öğrenme sürecine katılımı da içeren performans göstergelerine göre belirlenir. Tek oturumluk, belirli zaman dilimiyle sınırlı testler bu türden göstergeleri değerlendiremeyeceği için eksik araçlardır. Dolayısıyla “...bu uygulamanın geri-etkisi, öğrencinin sınıf içi

performansını etkileyecek ve öğrenme/edinme için gerekli olan düşünme, istek ve katılım gibi çok önemli göstergeler önemini yitirecek ve öğrenme süreci zaafa uğrayacaktır.” (Wall, 2005:26). Sınıf içi uygulamaları ne olursa olsun genelde ölçme ve değerlendirme uygulamaları büyük oranda “okulun/programın dışındaki birimlerin hazırladığı, çoktan-seçmeli, tek oturumluk, kağıt-kalem, norm-dayanlı” yani, ölçülendirilmiş sınavlarla belirlenmektedir (Nunan, 1999:84).

Yapı, içerik ve uygulamalarıyla baskın ve belirleyici olan bu tür sınavların (örneğin ÖSS-YDS) değişmesi zorunluluğunun iki temel nedeni vardır; daha zengin içerikli ve adil bir değerlendirme yapabilmek ve sınıf içi öğretiminin sınavlara olan bağımlılığını ortadan kaldırarak yeniden yapılandırmak.

Ölçülendirilmiş sınavlar genel dil öğretim programlarında belirlenen içeriğe uygun içerikte hazırlanmamaktadırlar. Bu amaç güdülse bile, ‘belirli zaman dilimiyle sınırlı’, ‘tek oturumluk’, testlerle bu içeriklerin ölçülmesi ve değerlendirilmesi mümkün değildir. Ölçülendirilmiş sınavlar nitelikli içerikte programları değerlendirmek için oldukça zayıf araçlardır. Diğer yanda öğretim içerikleri, sınav içeriklerine göre düzenlendiğinde öğrenciler üniversitede ya da yaşamda gerekli olacak nitelik ya da becerileri edinemeyeceklerdir (FairTest, 2007).

Yabancı dil öğretiminde tanıya yönelik elde edilen verilerin ayrıntılı, anlaşılır olmaları gerekmektedir. Ölçülendirilmiş sınavlar belli bir başlıkla ilgili genellikle birkaç soru içerdiklerinden tanısal amaçlı bir ölçme aracı olmaktan uzaktırlar. Ölçülendirilmiş sınavlar özetleyici, düzey belirleyici ya da eleyici bir özellik taşımaktadır ve gereksinim duyulan verileri sağlamaktan uzak bir nitelik taşımaktadır (FairTest, 2007:1-13). Tek bir sınav sonucuna bakarak gelişim ve başarıyı belirlemek zararlı sonuçlar doğurabilir (Bachman, 1990:74). Ölçülendirilmiş sınavların belirlediği ‘başarı/yeterlilik’ ölçütü yabancı dil programlarının müfredat içeriklerini ve sınıf öğretim uygulamalarını belirlemektedir. Bir beceri ya da bilginin ölçülme yöntemi/teknisi o beceri ya da bilginin öğretiminde model olmaktadır (Jones, 2000). Böylece dil öğretim programları, sınava hazırlık programlarına dönüşmektedirler (Turner, 2001).

Yukarıda değinilen nedenlerden ötürü öğretimin kapsamı ve niteliği değişebilmekte bunun sonucu olarak öğretim programları eleştirel düşünme, işbirliğine dayalı öğrenme ve yaratıcı becerileri geliştirmek yerine sınav endüstrisinin belirlenimlerinin öncelendiği birimler haline gelebilmektedir. Sınava dayalı yabancı dil öğretim programlarında öğrenciler bireysel öğrenme biçimlerindeki farklılıklarını keşfetme olasılığından yoksun kalabilmekte dolayısıyla kendiyile ve ötekiyle ilgili farkındalıkları gelişmeyen, özgün kimlikler olmaktan uzak, tek tipleşen bireylere dönüşebilmektedirler. Yaşamı sınavlardan oluşan bir yarış olarak algılayan bireyler böylesi bir süreç sonunda toplumsal yapılanmalara dair katılımcı ve eleştirel duyarlılıklarını kaybedebilmektedirler.

Genellikle okulların başarılarına dair ilk gösterge olarak öğrencilerin ölçülendirilmiş sınavlarda aldıkları puanlar dikkate alınmaktadır. Öğrencilerin

genel not ortalamaları yükseldikçe okulun eğitim kalitesinin yükseldiği varsayıldığından öğretmenler ve yöneticiler sınavlardan elde edilen puanların yüksek olması için yoğun bir baskı hissetmektedirler. Bu anlayışın baskısıyla okullar, sınav içeriğine uygun olarak müfredat içeriklerini daraltıp değiştirebilmektedir. Sınıf içi öğretim içeriği sınav içeriklerine göre değiştirilip düzenlenebilmekte sınavların çoktan seçmeli formatlarına uygun olarak öğretim yöntemleri seçilmekte ve öğretim giderek sınava dönüşmektedir (Bkz. Yavuz Kırık, 2008, Yayınlanmamış Doktora Tezi).

Örneğin test, çoktan seçmeli soru formatıyla oluşturulmuşsa bu tekrarların ve bolca ezberin zorunlu olduğu bir öğretim uygulamaları sürecini başlatmaktadır. Bu tür bir yöntem, sınav geçme becerisi ve sınavda yüksek not almanın dışında gerçek akademik performansı geliştirmekten yoksundur. ‘Test öğretmek’ müfredat içeriğini daralttığı gibi öğretmeni ve öğrenciyi bağlam dışı kavramların ezberlenmesine yönelik çalışmaya odaklar.

Okulları daha çok sınava dayalı çalışmaya yönlendiren etmenlerin başında Üniversite giriş sınavında başarı barajının yükseltilmesi gelmektedir. Bu yönelim, üniversite giriş sınavı içeriğinde yer almayan becerilerin öğretilmediği bir süreci başlatmaktadır. Örneğin İngilizce öğretiminde özgün ve performansa dayalı doğrudan teknikler olarak araştırma ve raporlama (sözlü/yazılı) çalışmaları sınav içeriğine dahil edilemedikleri için giderek uygulamadan kalkmaktadır (Bkz. Yavuz Kırık, 2008, Yayınlanmamış Doktora Tezi).

Öğretim bir tür teste, örneğin okuma-anlama etkinlikleri kısa alıntıların okunması ve ardından konuyla ilgili çoktan seçmeli soruların yanıtlanmasına, yazma etkinlikleri parçada bütünlüğü bozan tümcenin bulunmasına, sözlü ifade ya da iletişim/etkileşim etkinlikleri ise kısa bir diyalogda boş bırakılan ifadenin bulunması çalışmasına dönüşmektedir.

Çoktan-seçmeli ölçünlendirilmiş sınavların nesnel olarak kabul edilmesi tartışmaya açık bir yaklaşımdır. Çoktan-seçmeli testlerin nesnel oldukları iddiası puanlamada insan unsurunun yani, öznel olma potansiyeli taşıyan değerlendirme araçlarının yerine teknolojik araçların kullanılmasıyla ilgilidir. Ancak sınav içeriğine, soruları oluşturmada kullanılan tekniklere ve çeldiricilere insanların karar vermesi nesnellik iddiasının görünür çelişkilerindendir.

Doğrudan, performansa dayalı ölçme değerlendirme teknikleriyle oluşturulan sınavlarda adayın göstereceği başarı ile çoktan-seçmeli sorulardan oluşan sınavlarda elde edeceği başarı farklı olabilmektedir. Çünkü doğası gereği çoktan-seçmeli soruların yanıtını bulmak yani, çeldiricilerin arasından doğruyu ‘tanımak/fark etmek’ ile sorunun yanıtını düşünüp bulmak ya da hatırlamaktan farklı becerileri gerektirir. ‘Test-uzmanı’ öğrenciler sorunun yanıtına çeldiricileri eleyerek nasıl ulaşacaklarını çok iyi bilmektedirler (FairTest, 2007:1-3).

Çoktan-seçmeli sorular eleştirel düşünme yetisini, bilginin yabancı dilde sözlü/yazılı olarak ifade edilmesini, özgün ortamlarda etkili iletişim ve etkileşim becerisini ve doğal dil kullanım becerisini ölçmek için kusurlu ve zayıf araçlardır.

Çoktan-seçmeli ölçülendirilmiş sınavlar genellikle bu tekniğe uygun bilgiyi ölçmek üzere tasarlanmaktadır. Öğrenciler, sınav sorusu olarak karşılına çıkan kavramların önemli olduğu algısı üzerinden sınavda sorulmayan kavramların ya da becerilerin önemli olmadığı kanısıyla kısa sürede yanıtlanan, tek doğru yanıtın bulunup çıkarıldığı soru tiplerine odaklanmakta düşünme, akıl yürütme, çıkarsama yapma gibi yetilerini geliştirememektedirler.

Yabancı dil olarak İngilizce öğretim programları ya da özel dershaneler sınavlarda başarılı olma kaygısıyla test tekniğine ve içeriğine uygun öğretim içeriği belirlemekte ve tanımların hatırlanması ve doğru yanıtın fark edilmesi çalışmasına odaklanmaktadır. Sınırlı içerikteki çoktan-seçmeli ölçülendirilmiş sınavların etkisiyle daraltılan öğretim uygulamaları test benzeri sorular üzerine çalışmaya yönelmekte bütün etkinlikler tekrarlama çalışmasına dönüşmektedir. Bu süreçte, öğrenci kitap okuma, kendi sorularını sorma, tartışma açma-tartışmaya katılma, yazınsal metinlere ilişkin yorum yapabilme, duygu, düşünce ve bilgi birikimini aktaracağı yazılı raporlar hazırlama gibi gerçek yaşam etkinliklerine hazırlanamamaktadır.

2009- Yabancı Dil-İngilizce-Sınavı (YDS)

YDS (Yabancı Dil Sınavı), Öğrenci Seçme Sınavında (ÖSS) yer alan bir bölümdür. 2009 yılı ÖSS sınavının bir bölümü olarak İngilizce sınavı, yabancı dil olarak İngilizce’de iletişimsel dil kullanımı bağlamında yapılan yeterlilik tanımları üzerinden incelenecektir. Yabancı Dil Sınavı bölümünün alt başlığı olarak İngilizce sınavı, 100 sorudan oluşmaktadır. Sorular temel olarak sözcük, dil bilgisi, okuma-anlama, konuşma ve yazma becerisini ölçmeye yönelik beş çeldiricili çoktan-seçmeli yapıda sunulmuştur. İlk on soru tümcede boş bırakılan yere uygun sözcük bulma (Gap-filling) bilgisini ölçmeye yönelik sorulardır.

Yapı odaklı yabancı dil öğretiminin ölçme değerlendirme yaklaşımı kendini sözcük bilgisini ölçme değerlendirme bağlamında çoktan-seçmeli tekniklerle göstermiştir. Sözcük bilgisinin, çoktan-seçmeli sorularla ölçülmesi geleneksel ölçme ve değerlendirme araçları olarak ayrık-nitelikli yaklaşımın dil öğretim ve ölçümü anlayışına denk düşmektedir (Read, 2000). Ayrık nitelikli ölçme değerlendirme yaklaşımına göre tasarlanan boşluk doldurma soruları sözcüğü bağlam içinde uygun ve anlamlı kullanma becerisini geçerli bir şekilde ölçmekten uzak dolaylı araçlardır. Büyük oranda bellek gücünü ölçen çoktan-seçmeli sorularda doğru seçeneğin tahmin ya da rastlantısallıkla bulunma olasılığı yüksektir. Yoruma açık olmayan, tek doğru yanıt ilkesinden dolayı güvenilirlik ve uygulanabilirlik ölçütlerini yüksek düzeyde sağlayan fakat geçerli ve özgün olmayan tekniktir.

Günümüz gereksinimleri doğrultusunda değişen yabancı dil olarak İngilizce’de yeterlilik tanımı gereği sözcük bilgisi, özgün ortamlarda doğal dil kullanım becerisinin ve uygunluk bilgisinin bir bileşeni olarak görülmektedir. Dolayısıyla sözcük bilgisinde yeterliliğe dair kararlar ancak iletişimsel, özgün ve performansla dayalı ölçme değerlendirme teknikleriyle alınabilir.

“1,-22. sorularda cümlede boş bırakılan yerlere uygun düşen sözcük ya da ifadeyi bulunuz.” yönergesi altında sunulan 1. ve 2. sorular isim, 3. ve 4. sorular sıfat, 5. ve 6. sorular zarf/ belirteç, 7. ve 8. soruların fiil ve 9. ve 10. sorular ise takım fiiller (phrasal verb) yani, fiil+edat (verb+proposition) türünden sözcüklerin bilgisini ölçmeye yöneliktir.

Geniş bir sözcük dağarcığı bile gerçek yaşam dil kullanım etkinliklerinde etkili iletişim kurma becerisini tek başına sağlayamaz. Sözcüğün anlamını bilmek, sözcük bilgisinin sadece bir ögesidir. Sözcük bilgisi, sözlü ve yazılı dil kullanım etkinliklerinde sözcüklerin doğru ve uygun kullanım becerisinin ölçülmesiyle belirlenebilir.

Sözcük bilgisini ölçme değerlendirmeye yönelik, bağlamların birer tümceden oluştuğu 10 soruyla doğal dil kullanımı, etkin iletişim ve bağlama uygun sözcük kullanabilme bilgi ve becerisinin geçerli olarak ölçülme düzeyi tartışmalıdır. Sözcüğün yüklendiği anlam(lar) yorumlanırken kullanıldığı bağlam önemli bir açıklayıcıdır. Konuşmanın bağlamı, iletişim ortamı, etkileşime katılanların sayısı, konumu, yaşı, cinsi gibi birçok değişken bilinmeyen sözcüğün anlamını ya da yüklendiği farklı anlamları kestirmekte yardımcı unsurlardır.

Bu bağlamda, sözü edilen sorular ayrıık nitelikli ölçme değerlendirme yaklaşımıyla oluşturulmuş, yeterli bağlamlardan yoksun, gerçek yaşam etkinliklerini yansıtmakta eksik yapılar olarak yabancı dil olarak İngilizce’de yeterlilik kararları için sözcük bilgisi düzeyini doğrudan ölçen geçerli yapılar sayılamazlar.

2009-Yabancı Dil olarak İngilizce Sınavı’nda dilbilgisini ölçme değerlendirmeye yönelik olarak hazırlanan 11,-22. sorularda verilen tümcede ikişer boşluk bırakılarak çeldiricilerdeki ikişerli yapıdan uygun olanın bulunması istenmiştir.

11,-15. sorular dilbilgisinde zaman kullanım bilgisini 16. ve 17. sorular ilgeçlerin (preposition- at, to, from, with, for, through, of, by, into, in, about, on) kullanım bilgisini ölçmektedir. 18. soru belirteçlerin (adv.sen- however, nevertheless, finally), 19. soru sıralama (coordinating conjunction- and, but, yet, or, nor) ve yan tümce bağlaçlarının (subordinating conjunction- though, unless, if since, as if), 20. soru belirleyicilerin (determiner- that, in which), 21. soru belirleyici isimlerin (the noun- the few, the best) 22. soru ise belirteçlerin ve ilgeçlerin kullanım bilgisini ölçmek üzere tümcede tek bir boşluk bırakılarak sunulmuştur.

Dilbilgisi tanımı yabancı dilde iletişimsel yeterlilik tanımıyla birlikte değişmektedir. Örneğin Rea-Dickins (2001) dilbilgisine iletişimsel dilbilgisi (communicative grammar) adını vermektedir. Dil bilgisinin ne tür öğelerden oluştuğunu, ne tür değerlendirme etkinlikleriyle ölçülebileceğini, geçerli ve güvenilir tekniklerin nasıl tasarlanacağına dair soruların yanıtları konusunda bir uzlaşmaya varılamadığını söyleyen Purpura (2004:4-21) bu sorunun dil kuramı,

dilbilimsel kuram ve ikinci dilde dilbilgisi kavramının tanımındaki farklılıklardan kaynaklandığını belirtmektedir.

Yabancı dil öğretim hedef, içerik ve yöntemlerinde iletişimsel yaklaşımlar doğrultusunda gerçekleşen dönüşümler ölçme değerlendirme amaç, içerik ve tekniklerinde de değişimlere yol açmıştır. Bu değişim doğal olarak dilbilgisini ölçme değerlendirme yaklaşımlarında da aynı paralellikte sürmektedir. Son dönem yabancı dil olarak İngilizce öğretiminde ölçme değerlendirme yaklaşımları, çoktan-seçmeli soru yapılarını dolaylı teknikler olarak değerlendirmekte iletişimsel dilbilgisi yeterlilik düzeyini ölçen özgün ve geçerli araçlar tanımının dışında tutmaktadır.

Yabancı dil öğretiminde iletişimsel yaklaşımlar dilbilgisini gerçek yaşam sözlü/yazılı üretim etkinliklerinde amaca/bağlama uygun iletişim kurma becerisi içinde değerlendirilmesi gereken bir dil bileşeni olarak tanımlamaktadır. Bu bağlamda, uygulanabilirlik ve güvenilirlik ölçütleri öncelenerek kısıtlı zaman dilimlerinde, kolay değerlendirilen, kısıtlı bağlamlarla oluşturulan dolaylı ölçme değerlendirme teknikleri iletişimsel dilbilgisi yeterliliğini belirlemenin geçerli araçları olarak kabul görmemektedir.

İnceleme konusu sınavda çoktan-seçmeli 12 soru İngilizce’de ilgeçler, belirteçler, sıralama ve yan tümce bağlaçları, belirleyiciler, yardımcı fiiller gibi işlev sözcüklerini ve zamanları doğru yerde kullanma bilgisini ölçmek üzere oluşturulmuştur.

23.-32. sorular boşluklar bırakılarak verilen iki paragrafa dayanılarak yanıtlanması gereken sorulardır. Dörder tümceden oluşan paragraflarda doldurulması istenen beşer boşluk için istenen sözcük türlerine bakıldığında -zaman, niteleyici sıfatlar, ilgeçler, yan tümce bağlaçları, dönüşlü adılar, belirteçler gibi- bir metni anlamaya yönelik bütün dilbilgisel yapıların bilgisinin ölçüldüğü görülmektedir.

Fark etme/tanıma/hatırlama becerisini ölçen çoktan-seçmeli tekniklerde doğru yanıt bulmada tahmin ve şansın büyük payı olabilir. Bu tür teknikler doğaları gereği ölçülecek alanları daraltmakta ve öğretim teknik ve içeriğini olumsuz olarak etkilemektedirler (Bailey, 1998:129). Bir paragrafta bırakılan boşlukların yerine uygun düşen sözcüğü -içerik ya da işlev- ‘bulma’ becerisi iletişimsel dilbilgisinde uygunluk ve yeterlilik düzeyinin göstergesi sayılamaz. Dolayısıyla *çoktan-seçmeli boşluk doldurma* tekniği uygulanabilir ve güvenilir puanlamaya uygun fakat özgünlük ve geçerlilik ölçütlerini sağlayamayan dolaylı tekniklerden biridir.

33.-42 sorularda “verilen tümceyi uygun şekilde tamamlayan ifadeyi bulunuz” yönergesi ile dilbilgisel yapılar -yan tümce bağlaçları- zamanlar ve tümce anlamında tutarlılık bilgisinin ölçüldüğü sorulardır. Bu bölümde, ölçme değerlendirmede geleneksel teknikler olarak görülen *çoktan-seçmeli tümce tamamlama tekniği* kullanılmıştır. Çoktan-seçmeli tümce tamamlama tekniği (Multiple-Choice Completion) doğru yanıt vermek için bağlamı anlamayı gerektirmeyen, fark etmeyi ya da hatırlama becerisini ölçen, tek doğru yanıt

mantığı üzerinden tasarlanmaktadır (Purpura, 2004; Ellis, 2001). Dilbilgisi ve kullanım bilgisini doğrudan ölçmediği için özgün ve geçerli sayılamazlar. Bu yaklaşımlarla değerlendirilen sınav sonuçları yeterlilik kararlarında yanılsamalara yol açabilir.

İletişimsel dilbilgisinde her ifade ya da öbek, bir düz anlam bir ikincil anlam ve bir de iletişim ortamında yer alan kişilere özgü yorumla ilintili anlamı içerir. İletişime katılanların konum, tutum ve ilişki biçimlerine bağlı olarak değişebilen anlamlar yani, bağlam, gerçekten neyin söylenmek istendiğini belirler. Bu bağlamda, sözlü iletişim ortamlarının kestirimsiz gelişen etkileşimine uygun dil kullanım becerisi daha doğrudan gözleme ve performansa dayalı tekniklerle ölçülüp değerlendirilmelidir.

Yazılı/sözlü bağlamlarda dilbilgisel yapıları uygun kullanabilme düzeyine dair yeterlilik kararlarının çoktan-seçmeli tümcede/paragrafta boşluk doldurma ve tümce tamamlama sorularını doğru yanıtlama başarısına göre alınması geçerlilik ölçütü bağlamında tartışmaya açık bir alan yaratmaktadır.

43.-46 sorular, verilen ifadenin hangi sorunun yanıtı olduğunu çeldiricilerden bulma bilgisini ölçmek için tasarlanmış sorulardır. Konuşma dilinde soru-yanıt biçiminde oluşan etkileşimi andıran ifadeler dilbilgisel yapı ve sözcük bilgisi bağlamında hangi soruların karşılığı olduğuna dair ipuçları taşımaktadır.

Çoktan-seçmeli soru-yanıt tekniği hangi beceri için tasarlanırsa tasarlansın hazırlanması zor, hedeflenen beceriyi ölçmekten uzak, öğrencinin verili seçeneklerde doğru olanı fark etme becerisini ölçen, doğru yanıt bulmada tahmin olasılığı yüksek bir tekniktir (Hughes, 1989; Weir, 1988).

47.-50. sorular “...verilen İngilizce cümleye anlamca en yakın Türkçe cümleyi bulunuz” yönergesi ile öğrencinin çeviri becerisini ölçmektedir. Örneğin 47. Soruda “...have great sense of smellcannot see at all.... ifadelerinin çevirisinde doğru seçenek olarak belirlenen D seçeneği “vardır” ve “hiç” sözcükleriyle diğerlerinden fark edilir bir şekilde ayrılmaktadır. Bu soruda “cannot see at all” yapısının anlamını bilmek doğru seçeneği bulmak için yeterli bir bilgidir.

- A) “...koku duyusu çok gelişmiş.....görme duyusu tam olarak gelişmemiş...”
- B) “...koku duyusu çok gelişmiş...görme duyusu güçlü..”
- C) “...çok iyi koku duyusuna sahip.....görme duyusu gelişmemiş ...”
- D) “...çok iyi bir koku duyusu vardır.....hiç göremeyen...”
- E) “...koku duyuları oldukça gelişmiş...gözleri görmeyen...”

50. soru doğru seçeneğe dair çok daha açık ipuçları barındırmaktadır. Örneğin “...Onur watched , and” ifadelerinin çevirisinde doğru seçenek olarak belirlenen E seçeneği diğer seçeneklerde yer almayan “..seyretti ve...” sözcükleriyle diğerlerinden tümüyle farklıdır. Verilen İngilizce tümcede sadece “and” işlev sözcüğünün Türkçe anlamını bilmek ve çeviri tümcelerinde var olup olmadığını ‘fark etmek’ doğru yanıt bulmak için yeterli görünmektedir.

- A)
- B)
- C)
- D) ...seyrettikçe ...
- E) ...seyretti ve ...

51.-54. sorularda ise "...verilen Türkçe cümleye anlamca en yakın İngilizce cümleyi bulunuz." yönergesi ile sunulan sorularda iletişimsel dil yeterliliği ile ilgili olmayan çeviri teknikleri bilgisini ölçmeye yönelik sorular sorulmuştur.

51. soruda Türkçe verilen tümcede "...bile..." sözcüğü doğru seçenek olan B seçeneği dışında (Even in a small....) hiçbir çeldiriciye yer almamaktadır. 'Test uzmanı' öğrencilerin kolaylıkla fark edebileceği bu yapı çeldirici tümceleri anlamayı gerektirmeyen bir ipucu işlevi görmektedir.

52. soru sadece "...alışıl gelmiş plastik..." ve "...az sayıdaki madde..." sözcüklerinin anlamını bilmeyi gerektirmektedir. B, D ve E çeldiricileri "alışıl gelmiş" sözcüğünün anlamını vermedikleri için, A çeldiricisi ise "az sayıdaki madde" sözcüğünü doğru içermediği için kolaylıkla elenebilecek yapıdadır.

- A) Conventional plastic....the only material...
- B) Plastic, which is traditionally made...one of the few material...
- C) Conventional plastic one of the few material...
- D) Traditioanal plastic..... one of the few material...
- E) Normal plastic..... one of the few material...

Kaynak dilden hedef dile (Türkçe→İngilizce) ve hedef dilden kaynak dile (İngilizce→Türkçe) çeviri tekniği bir gerçek yaşam etkinliğidir. Çeviride sözcük seçimi, anlamı doğru ifade eden dilbilgisel yapıların seçimi değerlendirmede ölçüt olarak kullanılmaktadır. Ancak çeviri becerisine yönelik yeterliliğin belirlenmesi çoktan-seçmeli tekniklerle değil doğrudan ve özgün tekniklerle mümkündür. Çoktan-seçmeli tekniklerin kullanıldığı sınav sonuçlarına göre alınan yeterlilik kararları yanıltıcı ve geçersiz olabilir.

55.-75. maddeler okuma-anlama becerisini çoktan-seçmeli tümce tamamlama tekniğiyle ölçmek üzere tasarlanmış yedi farklı içerikte okuma parçalarına bağlı olarak yanıtlanması gereken 21 sorudan oluşmaktadır. Sorular temel olarak metinde vurgulanan ana fikri çıkarma, ana fikri destekleyen önermeleri çıkarsama ve yorumlama, yazarın konuyla ilgili duygu ve düşüncelerini belirleme ve öneri niteliğindeki ifadeleri fark etme bilgisini ölçmeye odaklanmıştır.

"Okuma becerisinin -anlama, yorumlama, çıkarsama gibi- çoktan-seçmeli sorularla ölçülmesi yaygın bir uygulamadır. Çoktan-seçmeli okuma soruları için hazırlanan çeldiriciler öğrenciyi yanıltmakta tuzağa düşürmekte ve öğrencinin

düşünmesini engellemektedir. Öğrencinin yapacağı tek şey bu tuzaklara düşmeden istenen yanıtı bulmaktır. ... Çoğu araştırmacı çoktan-seçmeli soruları yanıtlama becerisinin, okuma becerisinden farklı olduğunu iddia etmektedir. Öğrenciler ilgisiz çeldiricileri elemeyi ya da sorunun yapısallığının mantıksal çözümlemelerini testlere hazırlanırken öğrenmektedirler” (Alderson, 2000:211).

Okuma-anlama becerisini değerlendirirken özellikle edebi metinlerde çoklu okuma/çoklu anlam kuramının göz önünde bulundurulması gerekir. Metnin anlaşıldığına dair veri, metinle ilgili anlama sorularıyla sağlanamaz (Delonay, 1991:106-112). Metnin ana fikri ya da mesajı okuyucu bağlamında göreceli olduğundan yanıtların bu yaklaşımla değerlendirilmeleri gerekir (Shohamy, 2001:50).

Üniversite giriş sınavlarına yönelik çalışma yapan okul ve özel dershaneler öğrencilerine çoktan-seçmeli sorularda doğru yanıtı ulaşmak ya da çeldiricileri elemeye yönelik stratejiler edindirmek konusunda oldukça yoğun bir çalışma yapmaktadırlar. Yani, zamana karşı yarışan test çözme uzmanı haline dönüşen öğrencilerin okuma-anlama stratejileri geliştirmelerine gerek kalmamaktadır. Dolayısıyla sınavlarda adayların tahmin, şans ya da ‘test çözme uzmanı’ olmaları gibi etmenlerin dışında başarı ya da yeterlilik düzeyleri bu tür dolaylı tekniklerle geçerli bir şekilde ölçülmesi oldukça zor görünmektedir.

76.-80. maddeler, verilen tümcelere “anlamca en yakın cümleyi” bulmaya yönelik *çoktan-seçmeli açıklama tekniği* ile hazırlanmıştır. Konuşma ve yazma becerisinde açıklama etkinliği gerçek yaşam dil kullanım etkinliği olarak günümüz ölçme değerlendirme yaklaşımlarında özgün ve geçerli bir teknik olarak kabul edilmektedir. Okuma-anlama, dinleme-anlama gibi doğrudan gözlenemeyen becerileri sözlü ve yazılı ürünlerle gözlenebilen becerilere dönüştüren açıklama tekniği öznel bir etkinlik olarak tek doğru yanıt ilkesiyle çelişen bir tekniktir. Bu bağlamda, açıklama becerisinde yeterlilik düzeyine dair kararlar çoktan-seçmeli tekniklerle yapılan ölçme değerlendirme sonuçlarına göre alınamazlar.

81.-85. maddelerde “...boş bırakılan yere, parçanın anlam bütünlüğünü sağlamak için getirilebilecek...” tümcenin bulunması istenmiştir. Anlam bütünlüğü olmayan paragraflar daha çok metin organizasyonuna ait bilgiyi ölçmektedir. Metinde konu ya da anlam bütünlüğünü sağlama stratejileri yazma becerisi bağlamında edinilmesi gereken bir niteliktir. Ancak ‘bulma, fark etme’ becerisini geliştirmeye uygun teknikler olarak okuma-anlamada yeterlilik düzeyini ölçme değerlendirmede eksik araçlardır.

“86.-90. sorularda verilen durumda söylenmiş olabilecek sözü...” bulma yönergesi altında konuşma becerisini ölçmek için tasarlanmış sorular bulunmaktadır. Bu sorular oluşturulan bağlamlar açısından olası gerçek yaşam durumlarını yansıtmaktadır. Verili durumlarda söylenebilecek en uygun söz ya da ifadeyi üretme becerisi bağlamı yani, kültürel nüveleri yorumlama ya da çıkarsamalar yapabilme düzeyiyle ilintilidir. Ancak etkileşimsel dil kullanım özelliklerini yansıtan bu tür etkinliklerin verili durumlara uygun söz ya da ifadeyi

'seçeneklerden bulma' etkinliğine dönüştürülmesi geçerlilik ölçütünün göz ardı edildiğinin göstergesidir.

91.-95. maddeler etkileşim halinde kestirimsiz durum ve konularda doğal dile özgü uygunluk ilkesi gözetilerek yanıt verme becerisini ölçmeye yönelik olarak tasarlanan sorulardan oluşmaktadır. Söyleşi/diyalog gibi sözel dil kullanım etkinliklerinde uygunluk bilgisi yabancı dilde iletişimsel edinç tanımının ölçütlerinden biridir. Etkin iletişimde bağlama uygun konuşabilme bilgi ve beceri düzeyi, gerçek yaşam bağlarını yansıtan özgün ve doğrudan yanı gözleme, performansa dayalı tekniklerle ölçülebilir. Kestirimsiz durumlarda birçok değişkene bağlı olarak oluşturulan doğal dile özgü konuşmaların, iki katılımcısı olan söyleşilere bağlı çoktan-seçmeli sorulara indirgenerek ölçülmesi konuşma becerisinde uygunluk bilgisine dair geçerli veri sağlaması oldukça zordur.

96.-100 maddeler yazma becerisini ölçmeye yönelik tasarlanmış ve temel olarak bağdaşıklık ve bağlaşıklık kurma bilgisine odaklanılmıştır. Yabancı dilde yeterlilik bağlamında yazma becerisini ölçmek için farklı başlıklarda ve türde kompozisyonlar yazdırılması önemlidir (Cumming, Mellow, 1996). Yazma becerisi öğrenciye gerçek yaşam koşul ve durumlarına özgü bağlamların verilmesi, amaç ve hedef okuyucunun belirtilmesiyle ölçülebilir. Birçok tipte yazma etkinliğinin sağlanması bu bağlamda önemlidir (Herzog, 1988:174-5).

Yazma becerisinde yeterliliği değerlendirirken, yazılı dilin kurallarını bilme düzeyi, olaylar arasında bağlantı kurma becerisi, ana temanın ne denli güçlü vurgulandığı, savunuyu destekleyici öğelerin etkin şekilde kullanım becerisi, yüzeysel ve derin anlamı oluşturma becerisi, kültüre özgü referansları doğru kullanarak hedeflenen okuyucuya ve amaca uygun yazma becerisi göz önünde bulundurulur.

Yazma, konuşma becerisinde olduğu gibi üretime dönük öznel bir etkinlik olarak birçok stratejiyi kullanma becerisini içerir. Beş tümceden oluşan paragraflarda "...anlam bütünlüğünü bozan..." tümceleri bulma yani, ana örgünün dışında kalan tümceleri bulma becerisi yazma becerisini oluşturan bileşenlerden sadece bir tanesidir. Sözü edilen sınavda beş çoktan-seçmeli soruyla ölçülen amaç, okuyucu ve bağlamın belirlediği, etkili ve yaratıcı yazma etkinliği biçimlendirici, gözleme, performansa dayalı birçok tekniğin kullanıldığı süregelen ölçme değerlendirme yaklaşımlarıyla mümkündür.

Sonuç

Yukarıdaki değerlendirmeler ışığında 2009-Yabancı Dil Sınavının İngilizce bölümü soru içerik ve tekniklerine bakıldığında, uygulanabilirlik ve güvenilirlik ölçütleri öncelenerek ölçülmesi gereken niteliklere değil, daha çok ölçülebilen niteliklere odaklanıldığı görülebilir. Uygulanabilirlik ve güvenilirlik ölçütleri gözetilerek, sorular dolaylı bir teknik olan çoktan-seçmeli yapıda tasarlanmıştır. Bu bağlamda, yabancı dil olarak İngilizce’de yeterliliğin değerlendirilmesi için olmazsa olmaz ölçütler olarak kabul edilen geçerlilik ve özgünlük ölçütleri göz ardı edilmiştir.

KAYNAKÇA

- Alderson, J.C. (2000). *Assessing Reading*. Cambridge University Press
- Bachman, L.F. (1990). *Fundamental Considerations in Language Testing*. Oxford University Press.
- Bachman, L.F. & Palmer A.S. (1996). *Language Testing in Practice: Designing and developing useful Language Tests*. Oxford University Press.
- Bailey, K.M. (1998). *Learning about Language Assessment: Dilemmas, Decisions and Directions*. Heinle & Heinle Publishers
- Brown, J. D. (2005). *Testing in Language Programs*. Mc Graw
- Brown, H.D. (2004). *Language Assessment: Principles and Classroom Practice*. Pearson Education Inc.
- Brown, J.D. & Hudson, T. (2002). *Criterion-Referenced Language Testing*. Cambridge University Press
- Cumming, A. (1996). Introduction: The Concept of Validation (Ed: A. Cumming, R. Berwick). *Validation in Language testing, Modern Language in Practice 2*. (pp.1-14). Multilingual Matters Ltd.
- Cumming, A., & Mellow, D. (1996). An Investigating into the Validity of Written Indicators of Second Language Proficiency (Ed: A. Cumming, R. Berwick). *Validation in Language testing, Modern Language in Practice 2*. (pp.72-93). Multilingual Matters Ltd.
- Cushingweigle, S., & Lynch, B. (1996). Hypothesis Testing in Construct Validation (Ed: A. Cumming, R. Berwick). *Validation in Language testing, Modern Language in Practice 2*. (pp.58-71). Multilingual Matters Ltd.
- Davies, A. & Brown, A. & Elder, C., & Hill, K. & Cumley, T. & McNamara, T. (1999). *Dictionary of language Testing, Studies in Language Testing 7*. Cambridge University Press.
- Delonay, V. (1991). Assessing Textual Understanding and Literature Learning in ESL (Ed: C. Brumfit). *Assessment in Literature Teaching: Development*

- in English Language Teaching*. (pp.106-112). McMillan Publishers Limited.
- Ellis, R. (2001). Some Thoughts on Testing Grammar: An SLA perspectives (Ed: C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara, K. O'Loughlin). *Experimenting with Uncertainty, Studies in Language Testing* 11. (pp. 251-263). Cambridge University Press.
- Genesee, F., Upshur, J. A. (1996). *Classroom- based Evaluation in Second Language Education*. Cambridge University Press.
- Harrison, A. (1983). *A Language Testing Handbook*. MacMillan Publishers Limited, London.
- Hasselgreen, A. (2004). *Testing the Spoken English of Young Norwegians, A study of Test Validity and the Role of 'Smallwords' in Contributing to Pupils Fluency, Studies in Language Testing* 20. Cambridge University Press.
- Herzog, M. (1988). Issues in Writing Proficiency Assessment (Ed: P. Lowe, Jr., C. W. Stansfield). *Language in Education: Theory and Practice Second Language Proficiency Assessment: current issues*. (pp. 149-175). Prentice-Hall, Inc, New Jersey.
- Hughes, A. (1989). *Testing for Language Teachers*. Cambridge University Press
- McNamara, T. (2000). Language Testing (Ed: H.G. Widdowson). *International Journal of AppliedLinguistics*. 16. (3) (pp. 424-434). Oxford University Press.
- McNamara, T. & Carsten, R. (2006). *Language Testing: The Social Dimension*. Blackwell Publishing.
- Nunan, D. (1999). *Second Language Teaching and Learning*. Heinle & Heinle Publishers, USA.
- Purpura, J. E. (2004). *Assessing Grammar*. Cambridge University Press.
- Rea-Dickins, P. (2001). Fossilization or Evolution: The case of grammar testing" (Ed: C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara, K. O'Loughlin). *Experimenting with Uncertainty, Studies in Language Testing* 11. (pp.21-32). Cambridge University Press.
- Read, J. (2000). *Assessing Vocabulary*. Cambridge University Press.
- Shohamy, E. (2001). *The Power of Tests: A critical perspective on the uses of language tests*. Pearson Education Limited.
- Spolsky, B. (1990). Social Aspects of Individual Assessment (Ed: J.H. A.L. De Jong, D.K. Stevenson). *Individualizing the Assessment of Language Abilities*. Multilingual.

- Turner, C. E. (2001). The Need for Impact Studies of L2 Performance Testing and Rating: Identifying areas of potential consequences at all levels of the testing cycle (Ed: C. Elder , A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara, K. O'Loughlin). *Experimenting with Uncertainty, Studies in Language Testing* 11. (pp. 138-149). Cambridge University Press.
- Yavuz Kırık, M. (2008). *Yabancı Dil Olarak İngilizce Öğretmenlerinin Ölçme Değerlendirme Bağlamında Tutum ve Yaklaşımları*. Yayınlanmamış doktora tezi. İstanbul Üniversitesi.
- Wall, D. (2005). *The Impact of High-stakes Examinations on Classroom Teaching. Studies in Language Testing* 22, Cambridge University Press.
- Weir, C.J. (1988). *Communicative Language Testing*. Prentice Hall.
- Weir, C.J. (1993). *Understanding and Developing language Tests*. Prentice Hall
- Neill, M. (2003 Ocak 13). Transforming student Assessment. <http://www.fairtest.org/facts/template.html>' den Kasım 2006'da alınmıştır. (s.1-8).
- FairTest. (2007 Aralık 17). The Dangerous Consequences of High-Stakes Standardized Testing. <http://www.fairtest.org/facts/Dangerous%20Consequences.html>'den 29 Kasım 2006'da alınmıştır. (s. 1-3).
- FairTest. (2007 Aralık 17). Multiple-Choice Tests. <http://www.fairtest.org/facts/mctfcats.html>' den 29 Kasım 2006'da alınmıştır.
- Jones, K. (2000 Ekim). High Stakes vs. Democracy. <http://www.fairtest.org/high-stakes-vs-democracy>' den 31Mart 2009'da alınmıştır.
- ÖSYM. (2009) 2009 Öğrenci Seçme ve Yerleştirme Sistemi (ÖSYS) 2009-ÖSYS: YDS Soruları ve Yanıtları <ftp://dokuman.osym.gov.tr/2009/2009YDS/yds2009ingilizce.pdf>