

EFL ÖĞRETMENLERİ İÇİN KAÇINILMASI GEREKEN YIRMI YAYGIN TEST HATASI

Grant HENNING
Çev.: Zübeyde YILDIZ*

Bir dereceye kadar iyi test metodu, iyi dil kullanımı gibi, hatalardan sakınmayla gerçekleştirilebilir. Hemen hemen her dil öğretim programı test hazırlığı ve idaresini gerektirir, ancak belli yaygın test hatalarından kaçınıldığı derecede böyle testlerin zahmete değer seçmeler, teşhise yarar değerlendirme araçları oldukları söylenebilir. Burada bulunan yaygın test problemleri listesi detaylı değildir, fakat sınıf ve çevresi kullanımı için hazırlanan testlerle ilgili geniş bir deneyimden edinilmiştir, böylece bir temsilci olduğu söylenebilir. Bu listenin, kendi sınavlarını hazırlamakta olan EFL öğretmenlerine kılavuz olarak hizmet etmesi için bir çeşit kontrol listesi olması amaçlandı.

En yaygın hatalar aşağıdaki gibi 4 kategoride gruplanır: Genel sınav nitelikleri, soru nitelikleri, test geçerliği ile ilgili kavramlar, idari ve puanlama konuları. Her kategorinin altında 5 özgün hata tanımlanmıştır. Bazı hatalar ve iyileştirme metodları tekrar edilmiş olmasına rağmen, aşağıdaki 20 hatanın her biri, tahlil edildiğinde, gelişmiş bir test programında ortaya çıkacak gerçek bir problem oluşturabilir.

GENEL SINAV NİTELİKLERİ:

1- Çok Zor ya da Çok Kolay Testler

Testler çok zor ya da çok kolay olduğunda puanlama tablosunun başlarında ya da sonlarında bir yığılma vardır. Bu olaylar genel olarak "sınırsal etkiler" olarak bilinir. Bu etkiler sonucu olarak, öğrencilerin kabiliyetlerini ayırmak için testin kapasitesinde azalma ve bilgi kaybı vardır. Gerçek sonuç, geliştirme amaçları için hem uygun olmayan hem de güvenilmez bir testtir. Birçok nedenle, öğrencinin başarısında yaklaşık % 50 etkisi olan özenin, ihtimamın sınavların ve soruların hazırlanmasında gösterilmesi gereklidir. Bu tip kurallar sınav bilgisini ve güvenilirliği en yüksek dereceye yükseltir. Bu, sınavın öğrenciler ya da program geliştirme amaçları için kullanımından önce erek dili konuşan, seçilmiş kişilerde denenmesi gerektiğini vurgular.

* Araş. Gör.; U.Ü. Eğitim Fakültesi, İngiliz Dili Eğitimi Anabilim Dalı.

2- Yetersiz Sayıda Soru

Sınavın güvenilirliği sınavdaki soruların sayısı ile direkt olarak ilişkilidir. Sınav çok uzun olup böylece gereksiz yere öğrencileri yorarken, diğer bir hata da sınavın çok kısa olup güvenilir olmaz oluşudur. Birçok kağıt-kalem EFL sınavları için 50 sorudan daha az soruyla makul bir güvenilirlik sağlamak zordur. Aynı zamanda, 100 veya daha çok sorulu EFL sınavları, ilave soruların dahil edilmesiyle güvenilirlikte çok az ya da hiç artış meydana getirmeyen bir noktaya ulaşır.

Benzer bir sonuç, soruların kullanımını içermeyen yazılı veya sözlü sınavlar için geçerlidir. Bu sınavlarda da güvenilir ölçüm için konu bakımından yeterince geniş ve çeşitli örnek bir dil kullanımı öğrencilerden istenmelidir.

3- Test Türü Fazlalığı

Genel dil yeterliğini test etmede önemli dil yeteneklerinin bir bütün olarak ele alınmasını sağlamak için bir takım yan testlerin düzenlenmesi yaygın bir uygulamadır. Bu ayrıca, genel yeterliğin ölçümünde geçerliğe sahip sınavların geliştirilmesinde gerekli bir adım da olabilir. Problem, bu tür birleşimler veya dil yeteneği ölçme birimleri gelişigüzel sürdürüldüğünde ortaya çıkar.

Birçok durumda yantestlerin sayısını arttırmamanın test grubuna, iki ya da üç çok iyi, güvenilir yantestten verebileceğinden daha fazla hiçbir önemli değişiklik katmadığı örneklerle açıklanabilir. Bunun uygulamadaki anlamı, yantestleri çoğaltığımızda bir çeşit "Aşırı" ölçüme kapılıp gitmemizdir. Örneğin, hata tanıma, gramer doğruluğu, kelime tanıma ve kompozisyon yazma testlerinin dahil edilmesi işitsel çözümleme yeteneği testi için hiç yer bırakmaz. Baska bir deyişle, testin genel EFL yeterliğini önceden tahmin etme ve açıklama yeteneği adıyla, teste var olan ana parçalarından başka şey eklenmez. Bunun gibi birçok gelişigüzel sürdürülen yerlilik testleri, çok fazla extra yük taşıdıklarından yetersizdirler.

4- Ölçmede Güven Eksikliği

Birçok standardize olmuş testler, bir elkitabı ile birlikte gelir. Elkitabı bize hem ne oldukları hem de nasıl amaçlandıkları konusunda testlerin geçerliği ve güvenilirliği hakkında bilgi sağlar. Bu bilgi, bize değişik durumlara uygulandığında test sonucuna duyacağımız güvenin düzeyini tahmin etmemize olanak verir. Bölgesel olarak hazırlanan testler önemli değerlendirmelerde kullanıldığında, güvenilirlik ve geçerlik tahminleri bu testler için sağlanmalıdır. Uygun tahmin formülleri, ölçme teorisi kitaplarında kolayca bulunabilir.

Değerlendirme aşamasında testin uygulandığı kişilerin, sonunda testin kullanılacağı kişilerle aynı toplumdaki olmasını sağlama gereği bu sorunla yakından ilişkilidir. Geliştirilen testin ait olduğu topluluğun niteliği onu kullanan topluluktan oldukça farklı olduğunda yabancı standarda uygun testlere teminatsız güven duyma alışlagelmiş bir olaydır. Bu çeşit büyük farklılıklar, testin yeni durumlarda tekrar incelenmesi gereğini vurgular.

5- Varolmayan Yapılarla Yanlış Öğretme

Dilin uygun olmayan yapılarını kullanma yoluyla öğrencilere hataları göstermek mümkündür. Aşağıdaki örneği ele alalım:

I _____ here since 5 o'clock.

- a) am being
b) have been
c) will be
d) am be

Açıkça görülüyor ki, "d" seçeneği İngilizce'de hiç bir normal durumda görülmez. Özellikle başlangıç safhasındaki bir öğrencinin bu formu öğrenebilme ve "am" in "be" yardımcı fiili olarak görev yaptığını hatırlama tutabilme olasılığı vardır. Seçeneklerin çeldirici olarak doğru olmayan formlar olmaları gerekli olduğu halde, en iyisi bu formların yukarıdaki "a" ve "c" gibi dilde muhtemel bazı yapılar içermeleridir.

SORU NİTELİKLERİ

6- Yanıltıcı Sorular

Yanıltıcı soruların kullanılmasından kaçınılmalıdır. Bir kural olarak bu tip sorular öğrencinin güdüsünü, öğretmenin inancını ve testin kalitesini eksiltir. Bunların kullanımı, hatalı pedagojik yaklaşımın açık bir işareti. Aşağıdaki örneği düşünün:

I did not observe him not failing to do his work because he was

- a) always working.
b) ever conscientious
c) consistently lazy
d) never irresponsible

Bu soruya göz attığımızda İngilizce'nin normal kullanım sınırlarını aşan çift olumsuzluk yapısını içeren asıl parçayı farkederiz. Bu tip sorular, çoğunlukla olumsuzluk yapısının fark edilebilmesi için bulundurulur... yani daha zayıf öğrenciler belki de "c" seçeneğinin diğer seçeneklerden farklı oluşuna dikkat ederek geçebilirken, oldukça geniş sözlük bilgisine sahip daha iyi birçok öğrenci başarısız olur.

7- Kelime Fazlalığı

Soru yazmada, özellikle çok seçenekli soru tipinde yaygın bir problem, gereksiz tekrarlamalardır. Aşağıdaki örneğe bakalım:

He went to school

- a) because he wanted to learn more.
b) because he wanted to meet new friends.
c) because he wanted to get a better job.
d) because he wanted to please his parents.

Bu tip bir soru aşağıdaki gibi yazılırsa çok daha iyi olur:

He went to school because he wanted to

- a) learn more.
b) meet new friends.
c) get a better job.
d) please his parents.

Kelime fazlalığı olan sorular, belirli eşit zamanda okuyabileceği bilgi miktarını azalttığı için testin verimini düşürür.

8- Ayırıcı İpuçları

Çoktan seçmeli soru tipi için seçenek yazarken, doğru seçeneğin seçimine göre ipuçları hazırlamak önemlidir. Teste yatkın öğrenciler çoğunlukla bu tip soruları konu hakkında hiç bilgisi olmadan doğru olarak cevaplayabilirler. Tipik ayırıcı ipuçları, seçenek özel bilgi ya da daha fazla detay verdiği zaman meydana gelir. Aşağıdaki örneği düşünelim:

In the story we know it was raining because

- a) the sky was dark.
- b) everyone insisted on wearing an overcoat and carrying an umbrella outside.
- c) it was that time of year.
- d) they heard thunder.

Hikayeyi okumadan "b" seçeneğinin doğru yanıt olduğunu sadece daha fazla ayrıntı verildiğinden tahmin edebiliriz. Bu çeşit seçenek ayırcılıktan kaçınmak gerekir.

9- Benzer İpuçları

Ayırıcı ipuçlarının anlaşılmasından daha güç olan, benzer ipuçlarının varlığıdır. Burada teste yatkın öğrenciler, nitelik benzeşmesi nedeniyle doğru seçeneği tanıyabilirler. Sorusu olmayan, seçenekleri olan aşağıdaki örneğe bakalım:

- a) crawl
- b) creep
- c) brawl
- d) trudge

Soru hakkında herhangi bir bilgi ya da soru olmadan tahminen "crawl" in doğru seçenek olduğunu cesaret edip söyleyebiliriz. Bu seçimin ardındaki mantık, "a)", "b" ye "d" seçeneklerinin oldukça yavaş ve basit hareketi belirtmeleridir. "c" seçeneğinin diğer seçeneklerle tek belirli ortak noktası, "a" seçeneği ile kafiyeli olmasıdır. Bu durumda hem semantik hem de fonolojik konuda zihni başka yöne çekmek yapılmış olur ve birbirine yakın olan nokta "a" seçeneğidir. Sınav olan kişiye ne sorulduğuna dikkat etmeden bu yolla doğru yanıtlanan soruların sayısı şaşırtıcı derecede çoktur.

10- Seçenek Sayısı

Yetersiz ya da değişen sayıda seçenekli sorular içeren testler oldukça çoktur. Çoktan seçmeli ya da Doğru - Yanlış sorular gelişigüzel tahminle muhtemel başarılarla olanak verir. Seçenek ne kadar az olursa, kurnazca ve başarılı tahminler nedeniyle yapılan yanlış ölçme ihtimali o derece yüksek olur. Doğru - Yanlış test tipinde öğrencilerin tahmin ile 50 puan almalarını beklemeliyiz. Gerçek ölçme sadece, 51'den 100 puana kadar olan bölümde yer alacaktır. Bu, bu tip testlerde hatasız ölçme için oldukça fazla soruya gerek olduğunu vurgular.

Seçeneklerin yetersizliği ile ilgili bir sorun da, seçenek sayısındaki düzensizliktir. Estetik görünümün dışında bu düzensizlik tahminden meydana gelen hataların düzeltilmesi için çeşitli formüllerin kullanımını imkansız kılmaktadır. Genelde, bir testteki sorular için seçenek sayısında istikrarlı olmak en iyi yoldur.

TEST GEÇERLİĞİ İLE İLGİLİ KAVRAMLAR

11- Karışık Konular

Bir test sadece ölçmek istediği konuya ya da yeteneği doğru bir şekilde ölçebildiği derecede geçerlidir. Bazen testlerin içindeki soruların gerçekte ölçtüğünden daha farklı şeyler ölçtüğü iddia edilir. Aşağıdaki iki soru örnek olarak verilebilir. Birinci sorunun fiil-zaman bilgisini, ikincisinin kelime bilgisini ölçtüğü söyleniyor:

He the man yesterday.

- a) see b) saw c) will see d) is seeing

The lady to many cities in Europe last year.

- a) visited b) traveled c) visits d) climbed

İlk örnek, "zaman"ı test etmek için amaçlandıysa da aslında "a" seçeneğinin "özne-fiil uyumu" bilgisini ölçtüğünü görüyoruz. Benzer şekilde ikinci sorunun kelime bilgisini ölçtüğü farzedilirse de "zaman"ı test eden "c" seçeneğini içeriyor. Bu tip zıt durumlar testi geçersiz kılar.

12- Yanlış Yol

Bazen kişi test edilenden başka bir yanıtlandırma ortamında geniş yetenek gerektiren testlerle karşı karşıya gelir. Pasajı anladığını göstermek için hatasız, yazılı yanıtlar gerektiren "okuyup kavrama" sorularını düşünün. Araştırma, bu tip testlerin, asıl ölçülmek istenenden daha başka şeyler ölçtüğünden geçersiz olduklarını vurgulamaktadır (Henning, 1975). Test edilen yeteneğin yanıtlandırma şeklinin temsilcisi olduğuna dikkat edilmelidir.

13- Yaygın Bilgi

Yaygın bilgi isteyen sorulardan kaçınmak gerekir. Aşağıda "okuyup kavrama" sorusunu örnek olarak düşünün:

According to the story, Napoleon was born in

- a) England b) France c) Germany d) Italy

Bu tip bir soruyu doğru olarak yanıtlama, okuma parçasını anlama yeteneğini gerektirmez ve böylece bu tip soru içeren testlerdeki yüksek başarı, okuduğunu kavrama yeteneğinden daha başka yetenekleri gösterir.

14- Ders Programının Uymaması

Belki de testlerin geçersiz olmalarındaki en yaygın neden, testin gerçekten ya öğretim konuları ya da dersin içeriğini ölçmesindeki başarısızlıktır. Bu meydana geldiğinde testin görünüş ya da içerik geçerliğinden yoksun olduğunu söyleriz. Başarı testi hazırlanırken, örnek ders içeriği için öğretmenin sistematik bir metodu olmalıdır. Kurs, öğretim konularına uymalı ve test de öğretim materyallerini, kelime, yapılar ve gerçekten öğretilen yetenekler ile yansıtmalıdır.

15- İçerik Karşılaştırma

Bir sözcük, içerik karşılaştırma gerektiren anlama testlerinde (ya okuyup ya da dinleyip anlama) düzen içindedir. Kavrama pasajında tam karşılığı olan test so-

rusundaki bir deyim ya da sözcüğü karşılaştırma, mutlaka kavramayı gerektirmez. Bellek ya da tanıma yetenekleri gerektirir ve bunlar da önemlidir, fakat kavramaya eş değerde değildirler. Bu tip içerik karşılaştırma görevleri bulunan testler genellikle kavramayı ölçmede geçersizdirler.

YÖNETİM VE PUANLAMA KONULARI

16- Kopya Kontrollerinin Yetersizliği

Öğrenciler kopya ile yüksek notlar aldıklarında, hiç şüphe yok ki testler hem güvenilirmez hem de geçersizdir. Bu tip bir hareketi önlemek, öğretmenin ya da test yöneticisinin sorumluluğudur. Bazı toplumlarda testte işbirliği ya da kopya daha az kötü olarak nitelendirilir. Öğretmen öğrencileri ayırmayı dikkate almalı ve mümkünse testin alternatif formlarını kullanmalıdır. Bu alternatif formlar değişik sırada aynı soruları içerebilir. Formlar bir sıradaki her öğrenciye değişik form verilecek şekilde dağıtılmalıdır. Bu etkin bir şekilde kopya girişimini azaltacaktır.

17- Yetersiz Açıklamalar

Açıklamalar, hem öğrenciye hem de testi yöneten herhangi bir yetkiliye anlaşılabilir şekilde açık olmalıdır. Öğrenciler ancak kuralları anladıkları takdirde doğru yanıt verebileceklerinden, ne yapacaklarını anlayamazlarsa yanıtları geçersiz olur. Yabancı dilde açıklamaları anlamayı ölçen bir test olmadıkça, açıklamaların ana dilde verilmesi aslında değerlendirme bakımından sakıncalı değildir.

Eğer yetkili asıl kuralları anlamada başarısızlığa uğrarsa, idareciden idareciye ya da gruptan gruba yönetsel eşitsizlikler oluşacaktır. Açıklamalar, test yöneticileri için özel bir eğitim gerektirse bile, dikkatlice standardize edilmelidir.

18- İdari Eşitsizlikler

Sadece idarecilere verilen değişik açıklamalar idari eşitsizliklere yol açmaz, aynı zamanda diğer faktörler de testin güvenilirliğini bozabilir. Aydınlatmanın bir sıfıfta yetersiz, diğerinde iyi olduğunu, ya da testi idare eden kişinin açıklamaları ve ya anlama parçalarını değişik sınıflarda değişik hız ve tonda okuduğunu düşünün. İkinci problem, bazen yüksek kaliteli kayıt araçlarının kullanımıyla çözülebilir.

Bu tür haksızlıkları ve dikkati dağıtan farklı sesler, testin uzunluğu, zamanı, idarecilerin yeterliği gibi diğer eşitsizlikleri önlemek için dikkat edilmelidir.

19- Yetersiz Rehberlik

Genel kullanıma konulmadan önce, testin erek toplumdan; sınırlandırılmış model bir gruba uygulanması önemlidir. Bu uygulama, sınavı yapan kişinin, sürenin ve testin zorluk derecesinin öğrencilere uygun olduğundan ve soruların amaçlandığı gibi görev yaptığından emin olunmasına olanak verecektir. Sıkıntıya sokan, düşüncesizce yapılan birçok işlemlerden bu basit adımla kaçınılmış olur. Şüphesiz rehber model, asıl sınav olan gruptan, uygulama etkilerini ve güvenilirliğin bozulmasını önlemek için ayrı tutulmalıdır.

20- Subjektif Puanlama

Son, yaygın bir problem de öğreticilerin öğrenciler hakkında subjektif, sabit

fikre sahip olduđunda meydana gelir. Örneđin, kompozisyon puanlamada, deđerlendirmenin çok kere elyazısının düzenliliđinden etkilendiđi gözlenmiřtir. Diđer faktörler de dođru deđerlendirmeyi bozabilir. Bazı öđretmenler ya da deđerlendiriciler, kađıtları puanlarken kendilerini çok katı ya da çok iyimser bulurlar. Kısaca subjektif puanlamaya güvenilmesi gerekiyorsa, bunu hafifletici birkaç tane yöntemin kullanılması gerekir. Öncelikle, diđer deđerlendiricilerce verilen notlar, birden fazla deđerlendiricinin danıřmanlıđından geçirilmelidir. Tüm deđerlendiricilerin verdiđi puanların toplamı, öđrencinin notunu belirlemelidir. İkinci olarak, deđerlendiriciler, hassas puanlama programlarından yararlanmalıdır. Belirli bir puan, belirtilen tip ve sayıdaki hatalar için çıkarılmalıdır. Bu řekilde deđerlendiriciler aynı tür performans eřit notlar vermiř olacaklardır. Son olarak , yeterli dil örneklerini öđrencilerden almak gerekir. Yazma ve konuřma testlerinde öđrencilere, çeřitli ortamlarda dillerini kullanarak daha geniř tasvir alınmasını sađlamak için birden fazla konu verilmelidir.

Bu problemler, sınıf EFL testlerinin hazırlıđında da řařırtıcı řekilde benzer durumdadır. Bunlardan kaçınıldıđı ya da düzeltildiđi takdirde EFL testlerinin kalitesi artacaktır. Diđer sorunlar da gösterilebilir fakat burada sıralanmıř olanlar, řüphesiz daha yaygın olanlar arasındadır.