



## DEĞERLENDİRME TERCİHLERİ ÖLÇEĞİNİN TÜRKÇEYE UYARLANMASI ADAPTATION OF ASSESSMENT PREFERENCES INVENTORY TO TURKISH

Yasemin GÜLBAHAR\*, Şener BÜYÜKÖZTÜRK\*\*

**ÖZET:** Öğrencilerin başarısına ilişkin değerlendirme, genellikle öğretme-öğrenme sürecinde ve sonunda gerçekleştirilir ve öğretimden bağımsız bir olay olarak görülür. Oysa ki, öğrencileri değerlendirmek için seçilen yöntemler öğrencilerin nasıl çalıştıklarını, değerlendirme sürecine nasıl hazırlandıkları ve içeriği ne düzeyde öğrendikleri gibi çok farklı değişkenler üzerinde farklı etkiler yaratmaktadır. İçeriğin öğretilmesi ve ne kadar öğrenildiğinin ölçülmesi birbirlerini bütünleyen ve çift yönlü etkileyen süreçlerdir. Ayrıca, seçilen değerlendirme yöntemi öğrencilerin içeriğe yaklaşımlarını ve öğrenme düzeylerini etkilemekte, dolayısı ile üst düzey düşünme becerilerini kazanmalarına yardımcı olabilmektedir. Bu yaklaşımdan hareketle bu çalışma, özgün formu Birenbaum (1994) tarafından üniversite öğrencileri için geliştirilen “Değerlendirme Tercihleri Ölçeğinin” Türkçeye uyarlanması amacıyla gerçekleştirilmiştir. Ölçeğin faktör yapısı açımlayıcı ve doğrulayıcı faktör analizi ile incelenmiştir. Analiz sonuçları, üç temel ölçekten oluşan ve birbirinden bağımsız olarak kullanılabilen ölçeklerin faktör yapısının özgün ölçekle benzer olduğunu göstermiştir. Ölçeklerin alfa iç tutarlılık katsayıları .58 ile .92 arasında değişmekte olup ortalaması .79’dur.

**Anahtar sözcükler:** değerlendirme tercihleri, ölçek uyarlama

**ABSTRACT:** Evaluation of student achievement is done during and after the teaching-learning process, and it is perceived as an event which is independent from teaching. However, the methods selected for evaluating the students have various effects on different variables like how students study, how they prepare themselves for evaluation process and to what extent they learn the content. Teaching the content and assessing the level of success are two processes which complete and affect each other in a two dimensional way. Furthermore, selected evaluation method facilitates students’ approaches to content and affects their level of understanding, thus supports students for gaining higher order thinking skills. From this point of view, the purpose of this study was to adopt “Assessment Preferences Inventory”, which is originally developed by Birenbaum (1994), to Turkish. The factor structure of the inventory was investigated by descriptive and confirmatory factor analysis. The findings of the analysis showed that factor structures of the inventory, which have three independent scales with their sub scales, were similar with the original inventory. The internal consistency values of scales range between .58 and .92 having a mean of .79.

**Keywords:** assessment preferences, inventory adaptation

### 1. GİRİŞ

Bütün yüksek öğretim kurumlarında tüm öğretim elemanları, her yıl verdikleri derslerde farklı biçimlerde değerlendirmeler yaparak, öğrencilere not vererek dersteki başarı durumları hakkında karar vermektedir. Bu değerlendirme sonuçları öğrencilerin doğrudan yaşamlarını ve gelecekteki planlarını etkilemenin yanı sıra, daha pek çok değişken üzerinde rol oynamaktadır. Öğrencilerin çalışma alışkanlıkları, öğrenme yaklaşımları, ne düzeyde öğrendikleri, ne öğrendikleri, bilgiyi kullanma ve uygulamaya dönüştürme becerileri gibi pek çok değişken seçilen değerlendirme yöntemine bağlı olarak değişiklik göstermektedir. Öğrencilerin başarı düzeyleri de seçilen yönteme bağlı olarak, öğrencilerin kişilik özellikleri doğrultusunda değişim gösterebilmektedir.

Değerlendirme, çoğu zaman öğrenme sürecinin sonunda gerçekleştirilen ve öğretimden bağımsız bir olay olarak görülmektedir. Oysa ki, öğrencileri değerlendirmek için seçilen yöntemlerin öğrencilerin içeriği ne düzeyde öğrendiklerinin öncesinde derse nasıl çalıştıkları, öğrenme sürecinde neler yaşandığı, sınava nasıl hazırlandıkları gibi farklı değişkenlere ilişkin betimleme yapma, geri bildirim sunma fırsatı da vermelidir (Birenbaum, 1997; Struyven, Dochy & Janssens, 2005;

\* Yard. Doç. Dr. Başkent Üniversitesi Eğitim Fakültesi Bilgisayar ve Öğretim Teknolojileri Eğitimi Bölümü, gulbahar@baskent.edu.tr

\*\* Doç. Dr. Başkent Üniversitesi Eğitim Fakültesi Ortaöğretim Fen ve Matematik Alanlar Eğitimi Bölümü, senerb@baskent.edu.tr

Birenbaum, 2007). Bu çalışmada değerlendirme (assesment), “çeşitli araç ya da teknikler kullanılarak öğrenci özellikleri hakkında durum belirleme” anlamında kullanılmıştır.

### 1.1 Değerlendirmenin Önemi

İyi veya kaliteli olarak nitelendirilen bir değerlendirme yöntemi, öğrencilerin öğrenme sürecinin bir parçası olmalı, ayrıca öğrencilere neyi nasıl öğrenmeleri gerektiği konusunda yol göstermelidir. Bu yaklaşımın temelinde şu düşünce vardır: Aslında öğretim ve öğrenme yöntemleri ile değerlendirme yöntemleri benzer şekilde tanımlanmaktadır ve değerlendirmenin öğrenmenin gerçekleşmesine katkı sağlaması gerekir (Biggs, 2003).

Kullanılan değerlendirme yöntemlerine paralel olarak öğrenciler içeriği “yüzeysel” olarak ezberlemekte veya “derin” bir şekilde anlamaktadırlar. Gijbels ve Dochy (2006) yaptıkları çalışmada biçimlendirici değerlendirme yöntemini kullanmanın, öğrencilerin değerlendirme tercihleri ve öğrenme konusuna yaklaşımları açısından bir fark yaratıp yaratmadığını sorgulamışlardır. Değerlendirme tercihlerinin öğrencilerin öğrenme kavramına nasıl yaklaştıklarına bağlı olarak değiştiği sonucuna ulaşmışlardır. Ezbere dayalı öğrenme yaklaşımını benimseyenler yüzeysel ve basit değerlendirme yöntemlerini tercih ederken, daha derin öğrenmeler gerçekleştiren öğrencilerin becerilerini ve ne kadar anladıklarını gösterebildikleri karmaşık ve oluşturmacı değerlendirme yöntemlerini tercih ettikleri ortaya çıkmıştır.

Değerlendirme tercihleri, öğrencilerin sadece nasıl öğrendiklerini açıklamakla kalmayıp, nasıl ders çalıştıklarını da etkilemektedir (Struyven, Dochy & Janssens, 2005). Diğer yandan, öğrencilerin çalışma alışkanlıkları da nasıl değerlendirilmek istediklerini ilişkin tercihleri üzerinde rol oynamaktadır. Yapılan çalışmalar daha az heyecanlandıkları ve daha yüksek notlar alabildikleri için öğrencilerin çoktan-seçmeli ve kısa-yanıtlı soru tiplerini içeren yüzeysel değerlendirme yöntemlerini tercih ettiklerini ortaya çıkarmıştır. Ancak, içeriği daha derinlemesine öğrenmek isteyen öğrencilerin klasik yöntemler dışında kalan alternatif değerlendirme yöntemlerini tercih ettiklerini, çünkü bu yöntemlerin öğrenme düzeylerini daha doğru ölçtüğünü düşündükleri de ortaya çıkan başka bir araştırma sonucudur (Sambell, McDowell & Brown, 1997).

Biranbaum ve Rosenau (2006) öğretmen adayları ile öğretmenleri öğretim stratejileri ve değerlendirme tercihleri açısından karşılaştırdıkları çalışmalarında, öğretmenlerin öğrenme ve değerlendirmeye ilişkin daha derin ve detaylı bir yaklaşım içerisinde olduklarını bulmuşlardır. Bu sonucun öğretmenlerin sürekli olarak, anlamlı öğrenme deneyimlerine katıldıkları için ortaya çıktığını belirtmişlerdir. Bu sonuçta öğretim elemanlarının öğrencilere aktif öğrenme stratejileri uygulayarak, oluşturmacı yaklaşımı temel alan otantik ve iş başında öğrenme gibi deneyimler yaşatmaları gerektiğini göstermektedir (Walsh, 2007).

Trotter (2006) öğrencilere sürekli düzey belirleyici değerlendirme yapmanın etkilerini incelemiştir. Öğretim elemanına çok yük getirmesine rağmen bu yaklaşımın öğrencilerin motivasyonlarını arttırmada, öğrenme anlayışlarının ve öğretim ortamlarının geliştirilmesi konularına olumlu yönde katkı sağladığı sonucuna ulaşmıştır. Öğrenmeye katkı sağlayan başka bir yaklaşım ise çalışma soruları yöntemi ile çalışma sürecinin planlanmasıdır. Wilhelm ve Pieters (2007) yaptıkları çalışmada öğrencilerin çalışmalarını planlamak amacıyla çalışma soruları vermişler, özellikle işbirliğine dayalı grup çalışmalarında başarı daha fazla olmakla birlikte öğrencilerin akademik başarılarının arttığı sonucuna ulaşmışlardır.

Gelbal ve Kelecioğlu (2007) tarafından yapılan başka bir çalışma ise, Ankara’da kamu ve özel ilköğretim okullarında görev yapan 242 öğretmen üzerinde gerçekleştirilmiştir. Öğretmenler, öğrencileri tanımada ve başarılarını daha çok ölçmede kendilerini daha çok yeterli gördükleri geleneksel yöntemleri kullanmaktadırlar. Öğretmenler, kendilerini yeterli görmedikleri öğrenci-merkezli değerlendirme yöntemlerini çok az kullanmaktadırlar. Yeni değerlendirme yöntemleriyle ilgili olarak karşılaşılan sorunlardan dikkat çeken noktalar olarak; “sınıfların kalabalık olması”, “zaman yetersizliği”, “değerlendirme zorluğu” ve “uygulamanın zorluğu” gösterilmektedir. Geleneksel yöntemlerin kullanılmasında karşılaşılan en önemli dört sorun ise sırasıyla, “hazırlanmanın

zorluğu”, “sınıfların kalabalık olması”, “zaman yetersizliği” ve “veli desteğinin olmamasıdır” olarak belirtilmiştir.

Literatüre bakıldığında yapılan araştırma sonuçlarını özetlemek gerekirse, içeriğin öğretilmesi ve ne kadar öğrenildiğinin değerlendirilmesi süreçleri birbirlerini bütünlemektedir. Her iki süreçte çift yönlü birbirlerini etkilemektedir. Ayrıca, seçilen değerlendirme yöntemi öğrencilerin içeriğe yaklaşımlarını ve öğrenme düzeylerini etkilemekte, dolayısı ile üst düzey düşünme becerilerini kazanmalarına yardımcı olabilmektedir.

Bu nedenle öğrencilerin değerlendirme tercihlerinin tahmin edilmesi süreci önemli bir sorun olarak görülmektedir. Birenbaum (1997), üniversite öğrencilerinin değerlendirme tercihlerini belirlemeye yönelik modüler yapıda üç ölçekten, bunlara ait yedi temel alt ölçekten oluşan bir araç geliştirmiştir. Alt ölçekler; “Değerlendirme Yöntemlerine İlişkin Boyutlar”, “Öğrencilere İlişkin Boyutlar” ile “Notlandırma ve Raporlaştırma” olmak üzere temel üç başlık altında toplanmıştır.

Yapılan taramada Türkiye’de bu amaçla geliştirilen böyle bir ölçeye rastlanmamıştır. Ölçeğin Türkçe formunun oluşturulmasının öğrencilerin akademik başarılarını geliştirmeye yönelik çalışmalara bir başka boyut daha ekleyebileceği düşünülmektedir. Bu bağlamda çalışmanın amacı, 1994 yılında Birenbaum tarafından geliştirilen ve özgün adı “Assessment Preferences Inventory – API” olan “Değerlendirme Tercihleri Ölçeği”ni Türkçeye uyarlamaktır.

## 2. YÖNTEM

### 2.1. Katılımcılar

Araştırmanın evreni, bir özel üniversitenin Eğitim Fakültesi bünyesindeki lisans programlarına kayıtlı toplam 1230 öğrenciden oluşmaktadır. Katılımcıların belirlenmesinde seçkisiz olmayan örnekleme yöntemlerinden tabakalı amaçsal örnekleme yöntemi kullanılmıştır. Çalışmada lisans programları ve sınıf tabakalamada ölçüt olarak ele alınmıştır. Araştırmacılar tarafından Ekim 2007 tarihinde gerçekleştirilen uygulamaya 476 lisans öğrencisi katılmıştır.

Katılımcılardan 77 kişi (%16,2) erkek, 399 kişi (%83,8) ise bayandır. Katılımcıların sınıflara göre dağılımı ise; birinci sınıftan 137 kişi (%28,8), ikinci sınıftan 101 kişi (%21,2), üçüncü sınıftan 176 kişi (%37,0) ve son sınıftan 62 kişi (%13,0) olmuştur. Katılımcıların bölümlere göre dağılımına bakıldığında, 60 kişi (%12,6) bilgisayar ve öğretim teknolojileri eğitimi, 65 kişi (%13,7) okul öncesi öğretmenliği, 74 kişi (%15,5) sınıf öğretmenliği, 113 kişi (%23,7) ilköğretim ve ortaöğretim matematik öğretmenliği, 77 kişi (%16,2) Türkçe ve Türk dili ve edebiyatı öğretmenliği, 80 kişi (%16,8) İngilizce öğretmenliği programlarına kayıtlıdır. Grupları 7 kişi okuduğu programa ilişkin soruyu cevaplamamıştır.

### 2.2. Ölçme Aracı

Değerlendirme tercihlerini belirlemek amacı ile “Değerlendirme Tercihleri Ölçeği - DTÖ” kullanılmıştır. Birenbaum (1994, 1997, 2007) tarafından geliştirilen Değerlendirme Tercihleri Ölçeği’nin kullanılabilmesi için araştırmacının kendisinden e-posta yöntemi ile izin alınmıştır. Birenbaum tarafından geliştirilen ölçek toplam 67 sorudan oluşmakta ve üç temel başlık altında yedi boyutta tercihleri sorgulamaktadır. Her bir başlığı farklı bir özelliği ölçen aracın birbirini bütünleyen ölçeklerden oluşan modüler bir yapıya sahip olduğu ifade edilebilir. Buna göre üç farklı ölçekten oluşan DTÖ’nün buna dair konu başlıkları ve onların alt ölçeklerine ilişkin açıklayıcı bilgiler aşağıda verilmiştir.

#### I. Değerlendirme Yöntemlerine İlişkin Boyutlar (32 madde)

1. Değerlendirme Türleri (17 madde)
  - a. Testler
  - b. Diğer değerlendirme türleri
2. Madde biçimi/İşlem türü (11 madde)
  - a. Seçenek/oluşturma
  - b. Bağlamsal (otantik) veya bağlam dışı (sentetik)

- c. Basit/Karmaşık
- d. Farklı türler.
- 3. Değerlendirmeye hazırlık (4 madde).
- II. Öğrencilere İlişkin Boyutlar (27 madde)
  - 4. Bilişsel Süreçler (14 madde).
  - 5. Öğrenci rolü/sorumlulukları (11 madde).
  - 6. Gayret göstergeleri (2 madde).
- III. Notlandırma ve Raporlaştırma (8 madde)
  - 7. Notlandırma ve dönüt (8 madde)
    - a. Kритere dayalı, öğrencinin kendi gelişimine paralel.
    - b. Tek bir not/Profil
    - c. Notlandırmada objektif olma
    - d. Dönüt

Ölçeğin Türkçe'ye uyarlanması, Hambleton ve Patsula (1999)'nın kültürler arası ölçek uyarlamaya ilişkin önerileriyle tutarlı olarak üç aşamada tamamlanmıştır:

1. Özgün ölçek Türkçe'ye, her iki dile hakim olan ve aynı zamanda test yapısı hakkında bilgisi olan iki uzman tarafından çevrilmiştir. Uzman çevirileri arasında dikkate değer farklar için tekrar görüş alınmış ve çeviri işlemleri tamamlanmıştır. Daha sonra, bu kez iki uzman tarafından özgün (kaynak) dile çevrilmiş ve özgün madde yapıları ile olan tutarlılıkları incelenmiştir. Yapılan incelemede, özgün ölçekteki maddeler ile Türkçe'den yapılan çeviri ile elde edilen formdaki maddelerin dil denkliğinin olduğu görülmüştür.
2. Ölçekte yer alan maddelerin anlamsal (kelimelerin anlamları), deyimisel (yaşamda kullanılan deyimlerin anlamı), deneyimsel (deneyimlerin varlığı ve anlamları) ve kavramsal (kavramların aynı bağlamda kullanılması) açılarından denkliğine bakılmıştır. Bu amaçla iki alan uzmanından görüş istenmiştir. Uzman görüşleri doğrultusunda bazı maddelerde düzeltmeler yapılmış ve 13 yeni madde yazılarak ölçeğe eklenmiştir. Böylece ölçeğin toplam madde sayısı 80 olmuştur (Bakınız Ek).
3. Ölçeğin faktör yapısı, yapı geçerliği ile ölçek puanlarının güvenilirliği ve maddelerin ayırt ediciliğini değerlendirmek amacıyla ölçek, bir özel üniversitenin eğitim fakültesi öğrencilerine uygulanmıştır. Uygulamadan elde edilen verilere dayalı olarak ölçeğin Türkçe formu için faktör yapıları, aracı oluşturan temel ölçekler ve/veya bu ölçekleri oluşturan alt ölçekler dikkate alınarak incelenmiştir. Analiz sonuçları makalenin bulgular başlığı altında verilmiştir.

### 2.3. Veri Analizi

DTÖ'nün faktör yapısı açımlayıcı faktör analizi (AFA) ve doğrulayıcı faktör analizi (DFA) ile incelenmiştir. Açımlayıcı faktör analizi, değişkenler arasındaki ilişkilere dayalı olarak faktör yapısını keşfetmeyi amaçlar. Model-veri uyumunu inceleyen doğrulayıcı faktör analizinde ise değişkenler arasındaki ilişkiye dair kurulan hipotezler test edilir (Kline, 1994; Tabachnick ve Fidell, 2001). DFA'da modelin geçerliliğini değerlendirmek için çok sayıda uyum indeksi kullanılmaktadır. Bunlar içinde en sık kullanılanları Ki-Kare Uyum Testi, İyilik Uyum İndeksi (GFI), Düzeltilmiş İyilik Uyum İndeksi (AGFI), Ortalama Hataların Karekökü (RMR veya RMS) ve Yaklaşık Hataların Ortalama Karekökü'dür (RMSEA). Alanyazında, DFA ile hesaplanan ( $\chi^2/sd$ ) oranının 5'ten küçük olması, modelin gerçek verilerle iyi uyumun bir göstergesi olarak görülebilmektedir (Sümer, 2000). Model veri uyumu için GFI ve AGFI değerlerinin .90'dan yüksek çıkması, RMS ya da standartlaştırılmış RMS ile RMSEA değerlerinin ise .05'den küçük küçük olması beklenir. Buna karşılık GFI değerinin 0.85'ten, AGFI değerinin 0.80'den yüksek ve RMS değerinin ise 0.10'dan düşük çıkması modelin gerçek verilerle uyumu için birer ölçüt olarak da kabul edilmektedir (Anderson ve Gerbing, 1984; Cole, 1987; Marsh, Balla ve McDonald, 1988). Faktör yapısı belirlenen ölçekler ve alt ölçekler için Cronbach alfa iç tutarlılık katsayıları hesaplanmıştır.

### 3. GEÇERLİK VE GÜVENİRLİK ANALİZLERİNE İLİŞKİN BULGULAR

Modüler bir yapıya sahip olan Değerlendirme Tercihleri Ölçeğini (DTÖ) oluşturan dört farklı ölçek için özgün faktör yapısının Türk kültüründe geçerliliği ilk olarak doğrulayıcı faktör analizi ile incelenmek istenmiştir. Ancak DTÖ'yü oluşturan ilk ölçek olan değerlendirme yöntemleri tercihleri ölçeği için yapılan ilk analizde model-veri uyumu için öngörülen yukarıda belirtilen ölçütlerin alt sınırlarının dahi karşılanmaktan oldukça uzak olduğu görülmüştür. Bu nedenle DTÖ'yü oluşturan üç ölçek için ayrı ayrı olmak üzere ilk aşamada Türk kültüründe geçerli olan faktör yapılarını ortaya çıkartmak amacıyla açımlayıcı (keşfedici) faktör analizi uygulanmıştır. Analiz ile saptanan faktör yapısını tanımlayan ölçme modellerinin verilerle uyumu incelemek için ek olarak doğrulayıcı faktör analizi uygulanmıştır. Yapılan her iki analiz sonuçları sırasıyla aşağıda verilmiştir.

#### 3.1. Değerlendirme Yöntemleri Tercihleri Ölçeği

Değerlendirme yöntemleri tercihleri ölçeği, a) değerlendirme türleri, b) madde biçimi/işlem türü ve c) değerlendirmeye hazırlık boyutlarını ölçmek amacıyla oluşturulmuş olup, bağımsız ölçekler olarak kullanılmaktadır. Ölçeklerin faktör yapıları ve güvenilirlik analizlerine ait sonuçlar aşağıda sırasıyla verilmiştir.

##### 3.1.1. Değerlendirme Türleri Alt Ölçeği (DTAÖ)

**Açımlayıcı Faktör Analizi :** Ölçekte 19 madde yer almaktadır. Analiz için Kaiser-Meyer-Olkin (KMO) değeri 0,84 ve "Bartlett's testi anlamlı ( $p < .01$ ) bulunmuştur. Analiz ölçeğin iki faktörlü olduğunu göstermiştir. Önemli olarak belirlenen faktörlerden birincisi ölçeğe ilişkin toplam varyansın %29,57'sini ve ikinci faktör ise %13,92'sini açıklamaktadır. Bu iki faktörün açıkladığı toplam varyans %43,49'dir. Faktör döndürme sonrasında ölçeğin birinci faktörünün 12 maddeden (8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19), ikinci faktörün ise 5 maddeden (1, 3, 4, 6, 7) oluştuğu belirlenmiştir. Her iki faktör grubuna da girmeyen 2 ve 5 numaralı sorular değerlendirme dışı bırakılmıştır. Birinci faktördeki yük değerleri 0,463-0,777 arasında değişirken, ikinci faktördeki yük değerleri ise 0,445-0,692 değerleri arasındadır (Tablo-1).

**Tablo 1: DTAÖ'nin Faktör Analizi Sonuçları**

	Faktör 1	Faktör 2	Ortak Varyans
<b>Faktör 1: Alternatif değerlendirme</b>			
D13	,777		,651
D11	,757		,587
D14	,754		,585
D10	,707		,557
D15	,676		,459
D17	,627		,418
D19	,598		,422
D12	,583		,340
D18	,573		,439
D9	,560		,345
D16	,519		,312
D8	,463		,273
<b>Faktör 2: Klasik Değerlendirme</b>			
D7		,692	,495
D1		,690	,500
D4		,688	,493
D3		,521	,272
D6		,445	,245
Not: .40'dan düşük yük değerleri tabloda gösterilmemiştir.			

Faktörlerde yer alan maddeler incelendiğinde Birenbaum (1994) tarafından önerilen başlıklara uyduğu gözlenmiştir. İlk faktörde yer alan maddelerin tümünün “Alternatif Değerlendirme Türleri”, ikinci faktörde yer alan maddelerin ise “Klasik Değerlendirme Türleri” başlıkları altında toplandığı gözlenmiştir. Sonuç olarak, ölçeğin bu bölümünün değerlendirme türlerini sorguladığı söylenebilir.

**Doğrulayıcı Faktör Analizi :** DTAÖ'nün iki faktörden oluşan modelinin toplanan verilerle ne derece uyum gösterdiğini incelemek amacıyla yapılan doğrulayıcı faktör analizinde 6. madde, faktörle ilişkisine dair standardize edilmiş katsayısı anlamlı çıkmadığı ( $r=.08$ ,  $p>.05$ ) için modelden çıkartılmış ve analiz tekrarlanmıştır. İkinci analiz ile model-veri uyumu için hesaplanan ki-kare değeri anlamlı bulunmuştur,  $\chi^2(sd=100)=456.85$ ,  $p<.01$ . Aynı analiz ile hesaplanan bazı uyum istatistikleri şöyledir: ( $\chi^2/sd=4.57$ ,  $RMSEA=.08$ ,  $RMR=.12$ ,  $StRMR=.07$ ,  $GFI=.89$ ,  $AGFI=.85$ ). Maddelerin faktörleri ile olan ilişkilerini gösteren standardize edilmiş katsayılar .37 ile .86 arasında değişmekte olup tümü .01 düzeyinde anlamlıdır.

**Güvenirlilik Analizi :** Değerlendirme türlerine ilişkin alt ölçeğinden elde edilen puanlar için hesaplanan alfa iç tutarlılık katsayısı birinci faktör için .86 ve ikinci faktörü için .67'dir.

### 3.1.2. Madde Biçimi-İşlem Türü Alt Ölçeği (MBAÖ)

**Açımlayıcı Faktör Analizi :** Ölçekte 13 madde yer almaktadır. Analizde KMO değeri 0,79 ve Bartlett testi manidar ( $p<.01$ ) bulunmuştur. Analiz, ölçeğin iki faktörlü bir yapıya sahip olduğunu göstermiştir. Önemli olarak belirlenen faktörlerden birincisi ölçeğe ilişkin toplam varyansın %24,57'sini ve ikinci faktör ise %16,29'sini açıklamaktadır. Bu iki faktörün açıkladığı toplam varyans %40,86'dir. Faktör döndürme sonrasında ölçeğin birinci faktörünün 8 maddeden (20, 21, 22, 23, 25, 27, 29, 30), ikinci faktörün ise 5 maddeden (24, 26, 28, 31, 32) oluştuğu belirlenmiştir. Birinci faktördeki yük değerleri 0,440-0,782 arasında değişirken, ikinci faktördeki yük değerleri ise 0,438-0,724 değerleri arasındadır (Tablo-2).

**Tablo 2: MBAÖ'nin Faktör Analizi Sonuçları**

	Faktör 1	Faktör 2	Ortak Varyans
<b>Faktör 1: Basit/Seçmeli</b>			
D22	,782		,638
D23	,723		,550
D21	,687		,513
D20	,591		,432
D25	,545		,308
D27	,498		,334
D30	,465		,243
D29	,440		,349
<b>Faktör 2: Karmaşık/Oluşturmacı</b>			
D31		,724	,542
D26		,618	,382
D28		,588	,501
D24		,467	,322
D32		,438	,199

Not: .40'dan düşük yük değerleri tabloda gösterilmemiştir.

Faktörlerde yer alan maddeler incelendiğinde Birenbaum (1994) tarafından önerilen “Madde Biçimi/İşlem Türü” başlığı ile tutarlılık gösterdiği belirlenmiştir. İlk faktörde yer alan maddelerin tümünün “basit/seçmeli”, ikinci faktörde yer alan maddelerin ise “karmaşık/oluşturmacı” başlıkları altında toplanabildiği düşünülmüştür.

**Doğrulayıcı Faktör Analizi :** MBAÖ'nün iki faktörden oluşan modelinin toplanan verilerle ne derece uyum gösterdiğini incelemek amacıyla yapılan doğrulayıcı faktör analizinde, 29. madde, çok sayıda madde ve ikinci faktörle tanımlanan gizil değişkenle ilişki verdiği görüldüğünden modelden çıkartılarak analiz tekrarlanmıştır. Yapılan ikinci analiz ile model-veri uyumu için hesaplanan ki-kare

değeri anlamlı bulunmuştur,  $\chi^2(sd=50)=211.10$ ,  $p<.01$ . Aynı analiz ile hesaplanan bazı uyum istatistikleri şöyledir: ( $\chi^2/sd$ )=4.22, RMSEA=.08, RMR=.09, StRMR=.07, GFI=.93, AGFI=.89. Maddelerin faktörleri ile olan ilişkilerini gösteren standardize edilmiş katsayılar .28 ile .67 arasında değişmekte olup tümü .01 düzeyinde anlamlıdır.

**Güvenirlilik Analizi** : Madde Biçimi/İşlem Türü alt ölçeğinden elde edilen puanlar için hesaplanan alfa iç tutarlılık katsayısı birinci faktör için .75 ve ikinci faktörü için .58'dir.

### 3.1.3. Değerlendirmeye Hazırlık Alt Ölçeği (DHAÖ)

**Açımlayıcı Faktör Analizi** : Ölçekte yer alan 4 maddenin Birenbaum (1994) tarafından önerildiği şekilde tek faktör altında toplanıp toplanmadığı incelendi. Analizde KMO değeri 0,63 ve Bartlett testi manidar ( $p<.01$ ) bulunmuştur. Analiz sonuçları, maddelerin tek bir faktör altında toplandığını, bu faktörün ölçeğe ilişkin toplam varyansın %58,79'unu açıkladığını göstermiştir. Bu dört maddenin yük değerlerinin 0,655-0,824 değerleri arasındadır (Tablo-3).

**Tablo:3 DHAÖ'nin Faktör Analizi Sonuçları**

	Faktör 1	Ortak Varyans
D48	,824	,678
D49	,799	,639
D47	,779	,606
D50	,655	,429

**Doğrulayıcı Faktör Analizi** : DHAÖ'nün tek faktörlü yapısını, modelinin toplanan verilerle ne derece uyum gösterdiğini incelemek amacıyla yapılan doğrulayıcı faktör analizi ile model-veri uyumu için hesaplanan ki-kare değeri anlamlı değildir,  $\chi^2(sd=1)=0.12$ ,  $p>.05$ . Aynı analiz ile hesaplanan bazı uyum istatistikleri şöyledir: ( $\chi^2/sd$ )=.12, RMSEA=.00, RMR=.00, StRMR=.00, GFI=1.00, AGFI=1.00. Maddelerin faktörleri ile olan ilişkilerini gösteren standardize edilmiş katsayılar .35 ile .82 arasında değişmekte olup tümü .01 düzeyinde anlamlıdır.

**Güvenirlilik Analizi** : Değerlendirmeye hazırlık alt ölçeğinden elde edilen puanların için hesaplanan alfa iç tutarlılık katsayısı .75'dir.

### 3.2. Öğrencilere İlişkin Boyutlar Ölçeği (ÖİBÖ)

**Açımlayıcı Faktör Analizi** : Ölçekte 27 madde yer almaktadır. Analizde KMO değeri 0,90 ve Bartlett testi manidar ( $p<.01$ ) bulunmuştur. Faktör analizi sonuçları, maddelerin iki faktörde toplandığını göstermiştir. Önemli olarak belirlenen faktörlerden birincisi ölçeğe ilişkin toplam varyansın %26,49'sini, ikinci faktör ise %18,55'ini açıklamaktadır. İki faktörün açıkladığı toplam varyans %45,05'dir. Faktör döndürme sonrasında ölçeğin birinci faktörünün 14 maddeden (33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46), ikinci faktörün ise 13 maddeden (51, 52, 53, 55, 56, 57, 60, 61, 62, 66, 71, 72, 73) oluştuğu belirlenmiştir. Birinci faktördeki yük değerleri 0,409-0,768 arasında değişirken, ikinci faktör için ise 0,498-0,728 değerleri arasındadır (Tablo-4).

Faktörlere maddelerin içerikleri dikkate alınarak yapılan isimlendirme Birenbaum (1994) tarafından önerilen yapıya uyum sağlamaktadır. İlk faktörde yer alan maddeler "Bilişsel Süreçler" başlığında toplanırken, ikinci faktörde yer alan maddeler ise öğrencinin değerlendirme sürecine katılımı veya sorumluluk alması anlamında "Öğrenci rolü/sorumlulukları" kapsamındadır.

**Doğrulayıcı Faktör Analizi** : ÖİBÖ'nün iki faktörden oluşan modelinin toplanan verilerle ne derece uyum gösterdiğini incelemek amacıyla yapılan doğrulayıcı faktör analizinde, 71. maddenin çok sayıda madde ve ikinci faktörle tanımlanan gizil değişkenle ilişki verdiği görülmüştür. Bu madde modelden çıkartılarak analiz tekrarlanmıştır. Yapılan ikinci analiz ile model-veri uyumu için

hesaplanan ki-kare değeri anlamlı bulunmuştur,  $\chi^2(sd=294)=1071.60$ ,  $p<.01$ . Aynı analiz ile hesaplanan bazı uyum istatistikleri şöyledir: ( $\chi^2/sd$ )=3.64, RMSEA=.07, RMR=.09, StRMR=.07, GFI=0.85, AGFI=0.82. Maddelerin faktörleri ile olan ilişkilerini gösteren standardize edilmiş katsayılar .26 ile .87 arasında değişmekte olup tümü .01 düzeyinde anlamlıdır.

**Güvenirlilik Analizi :** Öğrencilere İlişkin Boyutlar ölçeğinden elde edilen puanların güvenirliliği ile ilgili olarak alfa iç tutarlılık katsayısı birinci faktör için .92, ikinci faktör için .83 olarak hesaplanmıştır.

**Tablo:4 ÖİBÖ'nin Faktör Analizi Sonuçları**

	Faktör 1	Faktör 2	Ortak Varyans
<b>Faktör 1: Bilişsel Süreçler</b>			
D43	,768		,595
D37	,757		,612
D38	,757		,588
D39	,754		,583
D46	,742		,551
D35	,717		,529
D40	,695		,492
D36	,684		,499
D42	,663		,481
D41	,660		,503
D45	,647		,419
D44	,627		,423
D34	,545		,310
D33	,409		,172
<b>Faktör 2: Öğrenci Rolü/Sorumlulukları</b>			
D57		,728	,535
D53		,714	,515
D71		,681	,466
D56		,670	,493
D55		,630	,506
D73		,619	,474
D61		,604	,456
D66		,588	,478
D60		,568	,404
D62		,511	,294
D72		,498	,256
D52		,491	,271
D51		,455	,257
Not: .40'dan düşük yük değerleri tabloda gösterilmemiştir.			

### 3.3. Sınav Alma, Notlandırma ve Raporlaştırma Ölçeği (SNRÖ)

**Açımlayıcı Faktör Analizi :** Ölçekte 14 madde yer almaktadır. Analizde KMO değeri 0,87 ve Bartlett testi manidar ( $p<.01$ ) bulunmuştur. Faktör analizi ile maddelerin tek faktör altında toplandığı belirlendi. Tek faktör ölçeğe ilişkin toplam varyansın %34,28'ini açıklamaktadır. Bu maddelerin yük değerlerinin 0,279-0,719 değerleri arasında olduğu görüldü (Tablo-5).



**Tablo:5 SNRÖ'nin Faktör Analizi Sonuçları**

	Faktör 1	Ortak Varyans
D54	,583	,340
D58	,600	,360
D59	,608	,370
D63	,461	,212
D65	,556	,309
D69	,426	,182
D70	,611	,373
D74	,713	,508
D75	,543	,295
D76	,684	,468
D77	,279	,078
D78	,577	,333
D79	,719	,516
D80	,674	,454

**Doğrulayıcı Faktör Analizi :** SNRAÖ'nün tek faktörlü yapısının toplanan verilerle ne derece uyum gösterdiğini incelemek amacıyla yapılan doğrulayıcı faktör analizi ile hesaplanan ki-kare değeri anlamlı bulunmuştur,  $\chi^2(sd=75)=287.64$ ,  $p<.01$ . Aynı analiz ile hesaplanan bazı uyum istatistikleri şöyledir: ( $\chi^2/sd$ )=3.84, RMSEA=.07, RMR=.06, StRMR=.05, GFI=0.92, AGFI=0.89. Maddelerin faktörleri ile olan ilişkilerini gösteren standardize edilmiş katsayılar .27 ile .69 arasında değişmekte olup tümü .01 düzeyinde anlamlıdır.

**Güvenirlilik Analizi :** Sınav Alma, Notlandırma ve Raporlaştırma ölçeğinden elde edilen puanlar için hesaplanan alfa iç tutarlılık katsayısı.83'dür.

Sonuç olarak, bu çalışmanın verileri doğrultusunda güncellenen ölçek toplam 72 maddeden oluşmuştur Ölçeğin boyutları ve içerdikleri madde sayıları aşağıda verilmiştir:

### I. Değerlendirme Yöntemlerine İlişkin Boyutlar (32 madde)

1. *Değerlendirme Türleri (16 madde)*
  - a. Alternatif Değerlendirme Türleri (projeler, ürün dosyaları, sınıfcı tartışmalar vb.)
  - b. Klasik Değerlendirme Türleri (yazılı ve sözlü sınavlar)
2. *Madde biçimi/İşlem türü (12 madde)*
  - a. Basit/Seçmeli (çoktan seçmeli, doğru-yanlış vb.)
  - b. Karmaşık/Oluşturmacı (kavram haritaları, performans dayalı beceriler vb.)
3. *Değerlendirmeye hazırlık (4 madde):* Hazırlık sürecinde rehberlik, örnekler.

### II. Öğrencilere İlişkin Boyutlar (26 madde)

1. *Bilişsel Süreçler (14 madde):* Açıklayıcı, işlemsel, kavramsal, sentez, kritik düşünme, değerlendirme ve yaratıcılık.
2. *Öğrenci Rolü/Sorumlulukları (12 madde):* Kriterlerin ve değerlendirme standartlarının belirlenmesine katılım; madde yazımı, değerlendirme (kendini değerlendirme, akran değerlendirme), işbirliği.

### III. Sınav Alma, Notlandırma ve Raporlaştırma (14 madde):

Kritere dayalı, öğrencinin kendi gelişimi, tek bir not, notlandırmada objektif olma, dönüt

#### 4. TARTIŞMA VE SONUÇ

Değerlendirme o kadar geniş bir kavram ki, her yaşta, her öğretim ortamında, her kurumda, her eğitimde, kısaca öğrenme oluşması beklenen her ortamda vazgeçilmez bir olay olarak karşımıza çıkmaktadır. Öğretilecek içerik, öğrenci kitlesi, eğitimci, öğretim ortamı, öğretim süreci gibi değişkenler sabitlenemediği için değerlendirme yöntemleri ve süreçleri sürekli yapılandırılmakta ve değişiklik göstermektedir. Russell, Elton, Swinglehurst ve Greenhalgh (2006) e-öğrenme amaçlı değerlendirme konusunu tartıştıkları makalelerinde, işbirliğine dayalı öğrenme süreçlerini ve değerlendirmeye ilişkin nasıl daha kaliteli ve detaylı dönüt verebilecekleri üzerinde durmuşlardır. Değerlendirme her zaman öğrenme düzeyini ölçmek amacıyla değil, bazen de ortaya çıkan ürün hakkında karara varmak amacıyla yapılmaktadır. Dobson (2006) yaptığı çalışmada sunu hazırlamanın öğrenmeyi destekleyici rolüyle birlikte bir öğrenciyi değerlendirmek amacıyla alternatif bir yaklaşım olarak kullanılabileceğini belirtmiştir. Ancak, bu konuda akademik ortamlarda yapılan araştırmalara ve kuramsal temellere gereksinim olduğunu da vurgulamıştır.

Değerlendirme, yine farklı bir açıdan incelendiğinde bilgi verici ve düzeltici olabileceği gibi, not, diploma ve sertifika gibi daha kapsamlı sonuçlara ulaşmak için de kullanılmaktadır. Diğer bir ifade ile, süreç veya ürün değerlendirme olarak ayrılmaktadır. Yüksek öğretim düzeyinde ise değerlendirme boyutunda çok ciddi sorunlar yaşanmaktadır. Konuya ve uygulama alanlarına göre yapılan çalışmalar, çoğu zaman bağlamdan uzak kuramsal boyuttan oluştuğu için, öğrencilerin mezun olduklarını varolan bilgilerini uygulamaya dönüştürmede sorunlar yaşadıkları bilinmektedir (Knight, 2006). Değerlendirme sürecindeki ve sonucundaki pek çok sorunu gidermek amacıyla çeşitli yeni yaklaşımlar önerilmektedir. Öğrencinin değerlendirme sürecinin bir parçası haline gelmesi (Dancer & Kamvounias, 2005) ve öğretim elemanının verdiği eğitimin kalitesinin değerlendirilmesi (Shoa, Anderson & Newsome, 2007) bu yaklaşımlara birer örnek olarak verilebilir.

Bu makalenin kapsamına alınamayacak kadar çok değerlendirme türleri ve yaklaşımları dışında, farklı bir yaklaşım ise bu çalışma kapsamında irdelenen değerlendirme tercihleridir. Öğrencilerin değerlendirme tercihleri, çalışma alışkanlıkları, öğretim elemanının yaklaşımı ve değerlendirme yöntemleri, içerik, motivasyon ve not kaygısı gibi pek çok değişkene göre değişmektedir. Bu nedenle her öğrenciye hangi yöntemle değerlendirilmek istediğini sormak veya değerlendirme yöntemleri konusunda seçenek sunmak, öğrencilerin motivasyonunun artmasına, daha fazla öğrenme sorumluluğu almasına, daha detaylı ve kapsamlı öğrenmesine yol açabilir. Bu yaklaşımla bu çalışmada değerlendirme tercihlerini ele alan bir ölçek geliştirilmiş, geçerlik ve güvenilirlik çalışması yapılmıştır. Açıklayıcı faktör analizi sonuçları, özgün formu Birenbaum (1994) tarafından geliştirilen değerlendirme tercihleri ölçeğinin Türkçe formunun faktör yapısının özgün ölçek ile genel olarak tutarlı olduğunu göstermiştir. Genel ölçek altında birbirinden bağımsız üç ölçek için ayrı ayrı yapılan analizler ile ortaya konulan faktör yapılarına ilişkin her bir ölçme modeli için model-veri uyumu doğrulayıcı faktör analizi incelenmiştir. Analiz sonuçları, makalenin yöntem kısmında model-veri uyumu için belirtilen ölçütleri karşıladığı ifade edilebilir. Ölçek puanları için hesaplanan güvenilirlik katsayılarının da genel olarak .70 civarında olması puanların güvenilirliği için bir kanıt olarak görülmektedir. Sonuç olarak, bir grup ölçekten oluşan değerlendirme tercihleri ölçeğinin Türk üniversite öğrencilerinin değerlendirme tercihlerinin belirlenmesinde kullanılabilir olduğu söylenebilir. Uygulamanın bir üniversitenin lisans öğrencileri ile sınırlı tutulması, çalışmanın bir sınırlılığı olarak kabul edilmiştir. Ayrıca, gelecekte varyansı daha geniş ve daha büyük örneklem üzerinde yapılacak çalışmalarla ölçme aracının faktör yapısı ile ilgili farklı modellerin farklı analiz yöntemleri kullanılarak yeniden incelenmesinin de yararlı olacağı düşünülmektedir.

Bu çalışma kapsamında uyarlanan, geçerlik ve güvenilirlik analizleri yapılan “Değerlendirme Tercihleri Ölçeği”nin, gelecekte değerlendirme tercihleri doğrultusunda yapılacak çalışmalarda yol gösterici ve belirleyici olacağı düşünülmektedir. Tercihlerin öğrenciye bırakılmasının, öğrenciyi, süreci ve başarıyı nasıl etkilediği konusunda daha fazla araştırmaya gerek vardır. Bu araştırmalar sonucunda “değerlendirmeye dayalı öğrenme” konusunda kuramsal bir çerçeve oluşması beklenmektedir.

**KAYNAKLAR**

- Anderson, J. C. & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, 49, 155-173.
- Birenbaum, M. (1994). Toward adaptive assessment - the student's angle. *Studies in Educational Evaluation*, 20, 239-255.
- Birenbaum, M. (1997). Assessment preferences and their relationship to learning strategies and orientations. *Higher Education*, 33, 71-84.
- Birenbaum, M. (2007). Assessment and instruction preferences and their relationship with test anxiety and learning strategies. *Higher Education*, 53, 749-768.
- Birenbaum, M. & Rosenau, S. (2006). Assessment preferences, learning orientations, and learning strategies of pre-service and in-service teachers. *Journal of Education for Teaching*, 32(2), 213-225.
- Biggs, J. (2003). *Teaching for Quality Learning at University* (2nd edition). United Kingdom: The Society for Research into Higher Education and Open University Pres.
- Cole, D. A. (1987). Utility of confirmatory factor analysis in test validation research. *Journal of Consulting and Clinical Psychology*, 55, 1019-1031.
- Dancer, D. & Kamvounias, P. (2005). Student involvement in assessment: a Project designed to assess class participation fairly and reliably. *Assessment & Evaluation in Higher Education*, 30(4), 445-454.
- Dobson, S. (2006). The assessment of student Powerpoint presentations – attempting the impossible?. *Assessment & Evaluation in Higher Education*, 31(1), 109-119.
- Gelbal, S. Ve kelecioğlu, H. (2007). Öğretmenlerin ölçme ve değerlendirme yöntemleri hakkındaki yeterlik algıları ve karşılaştıkları sorunlar. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, 33, 135-147.
- Gijbels, D. & Dochy, F. (2006). Students' assessment preferences and approaches to learning: can formative assessment make a difference?. *Educational Studies*, 32(4), 399-409.
- Hambleton, R.K. & Patsula, L. (1999). *Increasing the validity of adapted tests: myths to be avoided an guidelines for improving test adaptation practies* 1,2.(online). Web: <http://www.testpublishers.org.journal.html>.
- Kline, P. (2000). *An easy guide to factor analysis*, New York: Routledge.
- Marsh, H. W., Balla, J. R., & McDonald, R. P.(1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin*, 103, 391-410.
- Russell, J., Elton, L., Swinglehurst, D. & Greenhalgh, T. (2006). Using the online environment in assessment for learning: a case-study of a web-based course in primary care. *Assessment & Evaluation in Higher Education*, 31(4), 465-478.
- Sambell, K., McDowell, L. & Brown, S. (1997). 'But is it fair?': an exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation*, 23(4), 349-371.
- Struyven, K., Dochy, F. & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: a review. *Assessment & Evaluation in Higher Education*, 30(4), 325-341.
- Shao, L. P., Anderson, L. P. & Newsome, M. (2007). Evaluating teaching effectiveness: where we are and where we should be. *Assessment & Evaluation in Higher Education*, 32(3), 355-371.
- Sümer, N. (2000). Yapısal Eşitlik Modelleri: Temel Kavramlar ve Örnek Uygulamalar. *Türk Psikoloji Yazıları*, 3(6) 49-74.
- Tabachnick, B. G. & Fidell, L.S. (2001). *Using multivariate statistics* (4th edition). Boston: Allyn and Bacon.
- Walsh, A. (2007). An exploration of Biggs' constructive alignment in the context of work-based learning. *Assessment & Evaluation in Higher Education*, 32(1), 79-87.
- Wilhelm, P. & Pieters, J. M. (2007). Fostering effective studying and study planning with study questions. *Assessment & Evaluation in Higher Education*, 32(3), 373-382.

**EK – DEĞERLENDİRME TERCİHLERİ ÖLÇEĞİ MADDELERİ****Bir dersteki başarınız değerlendirilirken aşağıdaki yöntemlerin her birinin ne düzeyde kullanılmasını istersiniz?**

1. Yazılı test (kitap ve ders notları açık)
2. Yazılı test (kitap ve ders notları kapalı, süre sınırlı)
3. Yazılı test (kitap ve ders notları açık, süre sınırlı)
4. Yazılı test (kitap ve ders notları açık, süre sınırsız)
5. Bireysel sözlü sınav (kitap ve ders notları kapalı)
6. Bireysel sözlü sınav - sorular sınav başlamadan yarım saat önce verilir (kitap ve ders notları kapalı)
7. Bireysel sözlü sınav - sorular sınav başlamadan yarım saat önce verilir (kitap ve ders notları açık)
8. Sözlü sınav – Öğretmenin her bir öğrencinin katkısını gözlemlediği grup tartışmaları şeklinde
9. Ev Ödevleri
10. Makaleler/Raporlar
11. Bireysel Sunumlar (Poster, slayt vb. kullanarak)

12. Grup halinde Sunumlar (Poster, slayt vb. kullanarak)
13. Projeler
14. Ürün dosyaları (portfolyo - bitmiş veya devam eden çalışmaların toplanması ile oluşan dosya)
15. Elektronik Ürün dosyaları (e-portfolyo - çalışmaların elektronik olarak saklanması )
16. Drama/Gösteri
17. Gözlem
18. Kendi kendini değerlendirme
19. Akran/Grup değerlendirme
20. Tamamlama (boşluk doldurma) soruları
21. Eşleştirme soruları
22. Doğru-Yanlış türünde sorular
23. Çoktan seçmeli sorular
24. Kavram haritaları
25. Kısa-yanıtlı açık-uçlu sorular
26. Uzun yanıtlı açık-uçlu sorular
27. Ders süresince yapılan uygulamalara benzer uygulamalar
28. Performansa dayalı işlemler/beceriler (alanında uzman birisi tarafından gösterilmiş örneklere benzer şekilde)
29. Gerçek hayattaki olay ve durumlara benzer işlemler/beceriler
30. Tek bir doğru yanıtı olan basit işlemler/beceriler
31. Birden fazla olası yanıtı olan karmaşık işlemler/beceriler
32. Her bir adımı öğretmen tarafından yönlendirilen detaylı işlemler/beceriler
33. Bilgi düzeyinde sorular (öğretim içeriğine ilişkin)
34. Kavrama düzeyinde sorular (öğretmen tarafından öğretilecek içeriğe ilişkin)
35. Öğrenilen içeriğin yeni durumlara uygulanmasını gerektiren sorular (uygulama düzeyi)
36. Örnekler sunulmasını gerektiren sorular
37. Farklı kavram ve fikirlerin karşılaştırılmasını gerektiren sorular
38. Analiz ve yorumlama gerektiren sorular
39. Sonuç çıkarmayı gerektiren sorular
40. Öğrenilen tüm konular arasındaki ilişkilerin genel bir özeti sunmayı gerektiren sorular
41. Yaratıcılık ve hayal gücü gerektiren sorular
42. Kişisel açıklama ve fikir sunmayı gerektiren sorular
43. Kritik düşünme gerektiren sorular
44. Başkalarının fikir ve çözümlerinin değerlendirilmesi istenen sorular
45. Bilimsel araştırma gerektiren sorular
46. Problem çözme becerisi gerektiren sorular

---

**Öğretmeninizin aşağıdaki olayları ne düzeyde gerçekleştirmesini istersiniz?**

47. dersin başında, başarı düzeyinizin ne şekilde değerlendirileceğine dair detaylı bir açıklama yapması
  48. Testin kapsamını netleştirmesi ve nasıl hazırlanılması gerektiğini açıklaması
  49. Testten önce, testte çıkabilecek sorulara benzer sorular vermesi
  50. Test sorularını içerisinden seçeceği bir liste vermesi
  51. Test sorularını öğrencilere hazırlatması
  52. Her öğrencinin kendi sorusunu hazırlamasına izin vermesi
  53. Değerlendirme gereksinimlerine öğrencilerle birlikte karar vermesi
  54. Sadece ortaya çıkan ürünü değil, sürecide değerlendiriyor olması
  55. Değerlendirme ve öğrenme sürecinin tartışıldığı bilgilendirme toplantıları yapması
  56. Dersteki başarılarını ve kendinizi değerlendirme etkinliklerini de dikkate almasını
  57. Hangi yöntem(ler)le değerlendirileceğiniz konusunda seçme özgürlüğü tanıması
  58. Notunuzun bir parçası olarak sınıf-içi tartışmalara katılımınızı dikkate alması
  59. Notunuzun bir parçası olarak ödev ve uygulamaları dikkate alması
  60. Notunuzun bir parçası olarak derse olan ilginizi dikkate alması
  61. Notunuzun bir parçası olarak öğrenmek için harcadığınız çabayı dikkate alması
  62. Grup olarak proje çalışmaları yürütmenize izin vermesi
  63. Dersin sonunda her bir sınav sorusuna ilişkin istatistiksel bilgi sunması
  64. Testleri bilgisayar üzerinden alma olanağı vermesi (online testler)
  65. Değerlendirme yaparken “Dereceleme Ölçekleri” (rubrik) kullanması
  66. Değerlendirme amaçlı bireysel “Görüşmeler” yapması
-

**Siz, aşağıdaki olayların ne düzeyde gerçekleşmesini istersiniz?**

67. Sınav sorularının basitten zora doğru sıralanmış bir şekilde hazırlanması
68. Sınav sorularının içerikte yer alan konu akışına göre hazırlanması
69. Dönem boyunca birkaç tane quiz (küçük sınav) yapılması
70. Akademik başarınızın farklı değerlendirme yöntemleri ile belirlenmesi
71. Kendi tercih ettiğiniz değerlendirme yöntemini seçebilme hakkının tanınması
72. Değerlendirme etkinliklerinin bir bölümünün akran değerlendirme sonuçlarını kapsamaları
73. Başarı değerlendirme için gerekli standartların/isteklerin oluşturulması sürecine katılım için bir şans verilmesi
74. Çalışma sonuçlarının detaylı ve iyi tanımlanmış kriterlere göre değerlendirilmesi
75. Çalışmaların iki farklı öğretmen tarafından değerlendirilerek ortak bir sonuç belirlenmesi
76. Notun ders süresince gözlenen kişisel ilerlemeye göre verilmesi
77. Notun derse katılan diğer öğrencilerin başarılarına göre verilmesi
78. Notun, konunun ne düzeyde öğrenildiğini mutlak olarak yansıtmaması
79. Ders sonucunda tek bir not değil, her farklı konu için başarı profilini görebilme
80. Çalışma veya sınav sonucuna ilişkin detaylı dönüt alabilme

**EXTENDED ABSTRACT**

Evaluation is a broad term that covers all ages, in all learning environments, in different disciplines at various educational setting. Parallel with the variables like content, target group, learning environment and teaching-learning process where the standardization and control is difficult, the evaluation methods and processes also reshaped by showing diversity.

Evaluation of student achievement is done during and after the teaching-learning process, and it is perceived as an event which is independent from teaching. However, the methods selected for evaluating the students have various effects on different variables like how students study, how they prepare themselves for evaluation process and to what extend they learn the content. Teaching the content and assessing the level of success are twofold processes, which complete and affect each other in a two dimensional way. Furthermore, the selected evaluation method facilitates students' approaches to content and affects their level of understanding, thus supports students for gaining higher order thinking skills.

Evaluation may be formative or summative which focus on either the process or the product. Serious problems faced in higher education showed that there is a gap between the theoretical knowledge and application. Among the possible alternatives for overcoming these obstacles may be evaluation of the process and product from various dimensions and in terms of several variables that affects the decision. Other than the evaluation methods and approaches, where the scope is outside of this manuscript, an alternative approach this study focused is assessment preferences. Assessment preferences of students changes according to different variables like study habits, content, motivation, instructors approach to evaluation and concerns about achievement scores. Therefore, providing students the freedom to choice preferred assessment methods may increase motivation, improve the responsibility and lead deeper understanding of content.

From this point of view, the purpose of this study was to adopt "Assessment Preferences Inventory", which is originally developed by Birenbaum (1994), to Turkish. Thus, for adaptation of the original inventory developed by Birenbaum (1994, 1997, 2007), permission were taken through e-mail by the researcher. The inventory, which is composed of 67 questions and three independent scales, was questioning assessment preferences in seven dimensions. First, the inventory was translated into Turkish by two experts. After aligning the language between these two experts, the inventory was again translated into English and inconsistencies were checked. Following this process, 13 items, which addresses emerging alternative approaches, were added to the inventory leading a total number of 80 items finally.

The final version of the inventory was administered to 476 students, via clustered sampling in all departments of Faculty of Education in a private university. The factor structure of the inventory was analyzed through Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA).

The findings of the analysis showed that factor structures of the inventory, which have three independent scales with their sub scales, were similar with the original inventory. The internal consistency values of scales range between .58 and .92 having a mean of .79. Of 80 items, 8 were eliminated after related analysis. As a result, the final version of the inventory was composed of 72 items, where the subscales and dimensions are described in detail below.

**I. Assessment-form Related Dimensions (32 items)**

1. *Assessment Types (16 items)*
  - a. Alternative Assessment Types (projects, portfolios, discussion etc.)
  - b. Classical Assessment Types (written and oral exams etc.)
2. *Item Format/Task Type (12 items)*
  - c. Simple/Multiple Choice (multiple choice, true-false etc.)
  - d. Complex/Constructivist (concept maps, performance based skills etc.)
3. *Pre-assessment preparation (4 items):* Guidelines for preparation, examples etc.

**II. Examinee's Related Dimensions (26 items)**

1. *Cognitive Processes (14 items):* Declarative, procedural, conceptual, synthesis, critical thinking, evaluation and creativity.
2. *Student's Role/Responsibilities (12 items):* Participation in determining evaluation standards and criteria; item development, evaluation (self-evaluation, peer-evaluation etc.), *collaboration*.

**III. Grading and Reporting (14 items):** Criterion referenced, relative to student's own progress, single score, objectivity in grading and feedback.

As a result, adopted assessment preferences inventory is said to be reliable and can be used to maintain assessment preferences of students in higher education in Turkey. It is supposed that this inventory would facilitate the future research studies on assessment preferences in Turkey. Future studies should focus on achievement when the evaluation is done according to individual preferences. It is expected a framework for "evaluation based learning" will emerge as a result of these future studies.