



BURSA ULUDAĞ ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ FAKÜLTESİ
HEMŞİRELİK BÖLÜMÜ



ÖLÇME VE DEĞERLENDİRME REHBERİ

Ölçme ve Değerlendirme Komisyonu

Kasım - 2023

İÇİNDEKİLER

ÖNSÖZ	3
Ölçme ve Değerlendirme Komisyonu Çalışma Usul ve Esasları	4
1. Eğitimde Ölçme ve Değerlendirmenin Önemi	8
2. Ölçme ve Değerlendirmede Temel Kavramlar	10
2.1. Ölçme Türleri	12
2.2. Ölçek Türleri	12
2.3. Değerlendirme	14
3. Ölçme Araçlarında Bulunması Gereken Özellikler	17
3.1. Güvenilirlik	17
3.1.1. Güvenilirlik Yöntemleri	19
3.2. Geçerlilik	21
3.2.1. Geçerlilik Yöntemleri	22
3.3. Güvenirlilik ile Geçerlilik Arasındaki İlişki	26
4. Eğitimde Kullanılan Ölçme Araçları	26
4.1. Bilişsel Alan Davranışları	28
4.2. Yazılı (Açık Uçlu) Sınavlar	33
4.2.1. Puanlama Anahtarının Oluşturulması	35
4.3. Çoktan Seçmeli Test Sınavları	42
4.3.1. Çeldiriciler	43
4.4. Doğru-Yanlış Soruları	48
4.5. Eşleştirme Soruları	48
4.6. Boşluk Doldurma Soruları	50
4.7. Sözlü Sınavlar	51
5. Test Geliştirme Süreci ve Madde/Soru İstatistikleri	53
5.1. Test Geliştirme Süreci Basamakları	53
5.2. Madde/Soru İstatistikleri	54
5.2.1. Madde Güçlük İndeksi	54
5.2.2. Madde Ayırt Edicilik İndeksi	56
5.2.3. Açık Uçlu Maddelerin Analizleri	58
Önemli Noktalar	62
KAYNAKLAR	64

ÖNSÖZ

Bursa Uludağ Üniversitesi Sağlık Bilimleri Fakültesi Hemşirelik Bölümü Ölçme ve Değerlendirme Komisyonu tarafından hazırlanan bu rehberin amacı, program yeterliliklerine göre belirlenen derslere ilişkin öğrenim kazanımlarının ölçülmesinde oluşturulacak ölçme araçlarının taşınması gereken özellikleri belirlemek ve bu özelliklere göre ölçme aracı hazırlama ilkelerini tüm yönleriyle ele almak, programımızın eğitim-öğretim süreçlerinin niteliğini arttırmak ve ortak bir dil oluşturmaktır. Bu kapsamda rehber; ölçme ve değerlendirmeye ilişkin temel kavramlar, ölçme aracında bulunması gereken özellikler, ölçme ve değerlendirme süreçlerinde kullanılan araçlar, ölçme ve değerlendirmede geçerlilik ve güvenilirlik gibi konuları içermektedir. Derslere ait ölçme araçlarının bu rehberde belirtilen niteliklere uygun olarak hazırlanması, uygulanması ve sonuçlandırılması gerekmektedir. Bu nedenle rehber doğrultusunda komisyon tarafından hazırlanan kurallar, belirtke tabloları ve kontrol listeleri kullanılarak sürecin yürütülmesi esastır.

Programımız ölçme ve değerlendirme süreçlerinin geçerli ve güvenilir şekilde yürütülmesini sağlayacak olan bu rehberin sizlere yol gösterici olmasını diliyoruz.

Ölçme ve Değerlendirme Komisyonu

Başkan: Dr. Öğr. Üyesi Öznur ERBAY DALLI

Başkan Yardımcısı: Dr. Öğr. Üyesi Hava GÖKDERE ÇİNAR

Üye: Dr. Öğr. Üyesi Nevin UTKUALP

Üye: Dr. Öğr. Üyesi Derya AKÇA DOĞAN

Üye: Öğr. Gör. Dr. Aylin PALLOŞ

Raportör Üye: Öğr. Gör. Dr. Eda ÜNAL

Üye: Öğr. Gör. Kübra Bağcı DERİNPINAR

Üye: Arş. Gör. Muhammed GÜLTAŞ

Üye: Öğrenci Faruk ERKUL

BURSA ULUDAĞ ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ FAKÜLTESİ HEMŞİRELİK BÖLÜMÜ
ÖLÇME VE DEĞERLENDİRME KOMİSYONU ÇALIŞMA USUL ve ESASLARI

BİRİNCİ BÖLÜM

Amaç, Kapsam, Dayanak ve Tanımlar

Amaç

MADDE 1 – Bu usul ve esasların amacı, Bursa Uludağ Üniversitesi Sağlık Bilimleri Fakültesi Hemşirelik Bölümü Ölçme ve Değerlendirme Komisyonunun yapısını ve görevlerini tanımlamaktır.

Kapsam

MADDE 2 – Bu usul ve esaslar, Ölçme ve Değerlendirme Komisyonunun oluşumu, işleyişi, görev, yetki ve çalışma esaslarını kapsamaktadır.

Dayanak

MADDE 3- Bu usul ve esaslar, 02.02.2008 tarih ve 26775 sayılı Resmi Gazete’de yayımlanan “Doktorluk, Hemşirelik, Ebelik, Diş Hekimliği, Veterinerlik, Eczacılık ve Mimarlık Eğitim Programlarının Asgari Eğitim Koşullarının Belirlenmesine Dair Yönetmeliğin” 4. Maddesi, 17/06/2021 tarihli Resmi Gazete’de yayımlanarak yürürlüğe giren 31514 sayılı “Yükseköğretimde Uygulamalı Eğitimler Çerçeve Yönetmeliği” maddeleri ve 20.09.2020 tarih ve 31250 sayılı Resmi Gazete’de yayımlanan Bursa Uludağ Üniversitesi Önlisans Lisans Eğitim-Öğretim Yönetmeliği doğrultusunda hazırlanmıştır.

Tanımlar

MADDE 4 – Bu usul ve esaslarda geçen;

- a) **Üniversite:** Bursa Uludağ Üniversitesini,
- b) **Fakülte:** Bursa Uludağ Üniversitesi Sağlık Bilimleri Fakültesini,
- c) **Dekan:** Bursa Uludağ Üniversitesi Sağlık Bilimleri Fakültesi Dekanını,
- d) **Bölüm:** Bursa Uludağ Üniversitesi Sağlık Bilimleri Fakültesi Hemşirelik Bölümünü,
- e) **Bölüm Başkanı:** Bursa Uludağ Üniversitesi Sağlık Bilimleri Fakültesi Hemşirelik Bölümü Başkanını,
- f) **Anabilim Dalı:** Sağlık Bilimleri Fakültesi Hemşirelik Bölümü bünyesinde bulunan anabilim dallarını,
- g) **Ölçme ve Değerlendirme Komisyonu:** Sağlık Bilimleri Fakültesi Hemşirelik Bölümü Ölçme ve Değerlendirme Komisyonunu,
- h) **Bölüm Sekreteri:** Bursa Uludağ Üniversitesi Sağlık Bilimleri Fakültesi Hemşirelik Bölüm Sekreterini,
- i) **Raportör:** Bursa Uludağ Üniversitesi Sağlık Bilimleri Fakültesi Hemşirelik Bölümü Ölçme ve Değerlendirme Komisyonu raportörlüğüne seçilen komisyon üyesini,
- j) **Ölçme ve Değerlendirme:** Bursa Uludağ Üniversitesi Sağlık Bilimleri Fakültesi Hemşirelik Bölümü bünyesinde yapılan başta öğrenciler olmak üzere personel, yöneticiler ve paydaşlar üzerindeki ölçme ve değerlendirme süreçlerini,
- k) **Üye:** Ölçme ve Değerlendirme Komisyonu’nda görev alan öğretim üye ve elemanlarını, idari personel ve öğrencilerini,
- l) **Paydaşlar:** Fakültemizde ders vermek üzere görevlendirilen öğretim elemanlarını ve öğrencilerini ifade eder.

İKİNCİ BÖLÜM

Komisyonun Oluşturulması, Çalışma İlkeleri ve Görevleri

Komisyonunun Oluşturulması

MADDE 4 - Hemşirelik Bölümü Ölçme ve Değerlendirme Komisyonu aşağıda belirtilen esaslar çerçevesinde oluşturulur:

- a) Komisyon, Bölüm Başkanının önerisi ve Bölüm Kurulu kararı ile kurulur.
- b) Bir takvim yılı içinde, mazeretsiz olarak üç toplantıya katılmayan üyenin üyeliği düşer, Bölüm Kurulu kararı ile yeni atama yapılır.
- c) Komisyonun faaliyet süresi 3 yıldır.
- d) Faaliyet süresi biten komisyon üyesi, Bölüm Kurulu kararıyla yeniden görevlendirilebilir.
- e) Ölçme ve Değerlendirme Komisyonu, Fakültenin web sitesinde ilan edilir.

Komisyonun Çalışma İlkeleri

MADDE 5 -

- a) Hemşirelik Bölümü Ölçme ve Değerlendirme Komisyonu, her akademik dönemde en az bir kez önceden belirlenen yer ve zamanda olağan toplantısını yapar. Gerekli hâllerde Başkan, Komisyonu olağanüstü toplantıya çağırabilir.
- b) Komisyon gerektiğinde konu ile ilgili uzmanları toplantılara çağırabilir ve alt komisyonlar oluşturabilir. Alt komisyonların oluşturulması ve çalışmaları için tüm görevlendirmeler, Komisyon Başkanının önerisiyle Bölüm Başkanı tarafından yapılır.
- c) Herhangi bir nedenle komisyon toplantılarına üç kez üst üste katılmayan veya 6 ay ve daha uzun süreli raporlu ya da izinli olan üyenin komisyon üyeliği kendiliğinden sona erer.
- d) Herhangi bir nedenle komisyondan ayrılmak isteyen üye, mazeretini belirten dilekçesini komisyon başkanına ve Bölüm Başkanlığına bildirir. Bölüm Başkanı mazereti göz önünde bulundurarak değerlendirme yapar ve kararını ilgili üyeye yazılı olarak bildirir.
- e) Üyelerin izin, özür ve mazeret sebebiyle toplantıya katılamayacak olmaları halinde, söz konusu durumla ilgili olarak Komisyon Başkanını bilgilendirmeleri zorunludur.

Komisyonun Görevleri

MADDE 6 -

- a) Mezuniyet öncesi eğitim dönemlerinde, ölçme ve değerlendirmede objektif, geçerli ve güvenilir ölçme ve değerlendirme araçlarının uygulanmasına yönelik öneriler geliştirmek, gerekli durumlarda Anabilim dalları ve diğer komisyonlarla iş birliği içinde olmak.
- b) Sınav ve öğrenci değerlendirmesine ilişkin kural ve süreçleri tanımlamak /güncellemek.
- c) Eğitimde objektif ölçme ve değerlendirme yöntemlerini belirlemek ve ilgili kurul/komisyon /anabilim dallarına öneride bulunmak.
- d) Eğitim uygulamalarına ilişkin öğrenci ve eğitici geri bildirimlerini almak, değerlendirmek ve sonuçlarını ilgili kurullarla paylaşmak.
- e) Ölçme ve değerlendirme sürecinde elde edilen verileri değerlendirmek ve rapor haline getirerek Bölüm Başkanlığı'na sunmak.
- f) Ölçme ve değerlendirme faaliyetlerinin bölümün eğitim amaçlarına ve öğrenim hedeflerine uygunluğunu değerlendirmek.
- g) Ölçme ve değerlendirmenin geliştirilmesi amacıyla öğretim elemanlarına yönelik kurs, seminer vb. bilimsel toplantılar düzenlemek ve geri bildirim almak.
- h) Dersin sorumlu öğretim üyeleri/elemanları tarafından hazırlanan sınav sorusu analiz sonuçlarının ilgili komisyona, kurula ve öğretim elemanlarına gönderilmesini sağlamak.
- i) Sınav sonrası yapılan soru analizi sonuçları konusunda, dersin sorumlu öğretim üyelerini bilgilendirerek soruların güncellenmesine yönelik öneriler sunmak.

- j) Ölçme ve değerlendirmeye ilgili sorunları ve olası çözümleri belirleyerek, tüm paydaşlar ile paylaşılması için Bölüm Başkanlığı'na sunmak.

Komisyon Başkanının Görevleri

MADDE 7 -

- a) Ölçme ve Değerlendirme Komisyonunun görevlerini yerine getirmesinde bölüm başkanına karşı sorumlu olmak.
- b) Komisyonu temsil etmek ve komisyon çalışmalarını yönetmek.
- c) Her yarıyıl için toplantı takvimini belirleyerek Komisyonu o takvim doğrultusunda toplantıya davet etmek.
- d) Komisyonun belirlenen amaç ve faaliyet kapsamına uygun olarak verimli işlemesini sağlamak.
- e) Komisyon toplantı çağrılarını yapmak, ihtiyaca göre komisyonu özel gündemli/acil toplantıya çağırarak.
- f) Komisyon kararlarıyla ilgili konuları (görevlendirme, iletişim, yazışma vb.) Bölüm Başkanlığı'na bildirmek.
- g) Her eğitim - öğretim döneminin sonunda Komisyonun faaliyet raporunu Bölüm Başkanlığı'na, ilgili kurul ve komisyonlara, gerektiğinde anabilim dalı başkanlıklarına sunmak.
- h) Ölçme ve değerlendirme sürecinde gerek duyulan raporların son kontrolünü sağlayıp, Bölüm Başkanı onayına sunmak.
- i) Komisyonun toplantı tutanaklarını dosyalayıp arşivlenmesinin kontrolünü sağlamak.

Komisyon Üyelerinin Görev ve Sorumlulukları

MADDE 8 -

- a) Bursa Uludağ Üniversitesi Sağlık Bilimleri Fakültesinin stratejik planı ve hedefleri doğrultusunda, bölümün ölçme değerlendirme faaliyetlerinin değerlendirilmesi ve kalitesinin geliştirilmesi için çalışmalar yapmak.
- b) Her dönem sonunda ders yürütücülerinden gelen ders değerlendirme raporlarını (sınav sorularının analizi, ders başarı durumu, başarısız öğrencilere yönelik planlamalar, ders değerlendirme anket sonuçları, vb. içeren) değerlendirilerek döneme ilişkin raporu oluşturmak.
- c) Komisyon çalışmalarının sonuçlarını içeren yıllık değerlendirme raporunu hazırlamak ve Bölüm Başkanlığına sunmak, onaylanan ölçme değerlendirme faaliyetleri raporunu web sayfasında paylaşmak.
- d) Öğretim üyelerine verilecek eğitim programlarının ölçme değerlendirme ile ilgili kısımları hakkında ilgili komisyonlara önerilerde bulunmak.
- e) Öğrencilerin performanslarının değerlendirilmesi için uygun yöntemlerin kullanılması amacıyla ilgili komisyonlarla iş birliği yapmak.
- f) Her eğitim-öğretim yılı sonunda gelecek eğitim-öğretim yılı için ölçme ve değerlendirme faaliyetleri ile ilgili planları yapmak.
- g) Komisyon Çalışma Usul ve Esaslarının değişen koşullar ve durumlar karşısında sürekli geliştirilmesini ve güncellenmesini sağlamak.

Raportörün Görevleri

MADDE 9 -

- a) Komisyon toplantı kararlarını yazmak, üye imza listelerinden oluşan Komisyon karar dosyasını oluşturmak ve görev süresi bitiminde görevlendirilen yeni raportöre komisyon ile ilgili bütün kayıtları eksiksiz olarak iletmek.
- b) Toplantı günlerini komisyon üyelerine bildirmek; gündem ve gündemle ilgili bilgi ve belgelerin komisyon üyelerine teslim etmek.

- c) Gerekli görülmesi halinde komisyon adına yazılacak yazıları hazırlamak, konu ile ilgili bilgi ve belgeleri sağlamak.
- d) Raportörün bulunmadığı toplantılarda raportörlük görevini başkanın önerisi ile üyelerden biri yürütür.

MADDE 10 - Toplantı tutanaklarında aşağıdaki bilgiler yer alır;

- a) Komisyon adı,
- b) Toplantı sıra numarası,
- c) Toplantı tarihi (günü, saati),
- d) Toplantı yeri,
- e) Toplantıya katılanlar,
- f) Toplantı gündemi ve kararlar.

Komisyon Kararlarının Uygulanması

MADDE 11 -

- a) Komisyon, Başkan tarafından temsil edilir. Alınan kararlarda uygulamadan ve kararların takip edilmesinden sorumlu olan kişiler belirtilir. Kararlar Başkan tarafından ilgili kişi ve birimlere bildirilir. Komisyon adına yazışmalar Başkanın imzasıyla yapılır.
- b) Komisyon, faaliyetleri için harcama gerektiğinde, gerekçesi ve kullanım alanlarını yazılı olarak belirtmek kaydıyla Fakülte Yönetiminden ödenek talep edebilir.

Komisyon Kararlarının Takibi

MADDE 12 - Ölçme ve Değerlendirme süreciyle ilgili çalışmaların ve Komisyon kararlarının takibi Başkan ve Bölüm Başkanı tarafından yapılır.

MADDE 13 - Komisyon kararları katılımcıların oy çokluğu ile alınır, eşitlik halinde Başkanın oyu iki oy sayılır.

MADDE 14 - Bu çalışma usul ve esaslarında hüküm bulunmayan hallerde konuya ilişkin ilgili mevzuat hükümleri uygulanır.

Yürürlük

MADDE 15- Bu çalışma esasları Fakülte Kurulu tarafından onaylandığı tarihten itibaren yürürlüğe girer.

Yürütme

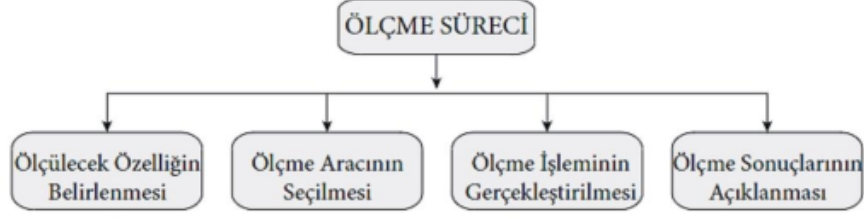
MADDE 16- Bu usul ve esasları hükümlerini Hemşirelik Bölüm Başkanı yürütür.

1. Eğitimde Ölçme ve Değerlendirmenin Önemi

Eğitim sürecinin kalite kontrol niteliğini taşıyan ölçme ve değerlendirme sistemleri, sürecin yalnızca sonunda değil, her aşamasında yer almaktadır. Eğitim etkinliklerinin ne denli başarılı olduğu, dönem ya da yıl içerisinde kaydedilen ilerlemelerin eğitimcinin öğretim hedeflerini kazandırmada ve öğretim sürecinde öğrencilere rehberlik etmede ne derecede başarılı olduğu ölçme ve değerlendirme süreçleri yardımı ile ortaya konmaktadır. Ölçme ve değerlendirme süreçlerinin ne kadar etkili yürütüldüğü, eğitimcinin bazı yeterliliklere sahip olmasını gerektirmektedir. Bunlar;

- Öğrenim sürecinin durumunu eğitim amaçlarına uygun olacak şekilde belirleyebilmesi,
- Uygun ölçüm aracı geliştirilebilecek bilgi düzeyine sahip olması,
- Ölçme ve değerlendirme süreçlerini uygun şekilde sürdürebilmesi ve yorumlayabilmesi,
- Eğitim ile ilgili kararlarda ölçme sonuçlarını kullanabilmesi,
- İlgili kişi ve kurumlara sonuçlara ilişkin geri bildirim verebilmesi,
- Uygun veya etik olmayan ölçme süreçlerini tanıyabilmesidir.

Ölçme, değerlendirme ve durum belirleme kavramları zaman zaman birbiri yerine kullanılsa da, anlamları birbirinden farklıdır. **Ölçme**, bir performansın gerçekleşme düzeyini belirleme veya gözlenen performansa sayı tayin etme süreci olarak tanımlanmaktadır. **Değerlendirme**, öğrenci performansının niteliği hakkında karar verme işlemidir ve bu süreçte ölçme sonuçları ve seçilen ölçüt arasında bir karşılaştırma yapmayı gerektirmektedir. **Durum belirleme** ise, eğitim öğretim süreci devam ederken öğrenci performansı hakkında tüm bilginin toplanması, sentezlenmesi ve yorumlanması işlemidir. Durum belirleme, öğreticinin sınıf hakkında kullandığı tüm bilgi edinme yollarını temsil etmektedir. Durum belirlemeler öğrenci, eğitici, program ve eğitim politikaları hakkında verilere dayalı olarak yapılan, süreklilik arz eden genel yorumlardır. Durum belirleme, değerlendirmeden daha kapsamlıdır. Bu durumda ölçme; durum belirleme ve değerlendirme için bir temel oluşturulmasına yardımcı olmaktadır. Ölçme işlemi, Şekil 1’de gösterildiği üzere ölçülecek özelliğin tanımlanması, ölçmeye konu olan özelliğe karşı gelen değerin saptamasını en iyi şekilde sağlayacak ölçme aracının belirlenmesini, ölçme işleminin gerçekleştirilmesi ve sonuçlandırılması süreçlerini gerektirmektedir.



Şekil 1. Ölçme sürecinin aşamaları

Ölçme, değerlendirme ve durum belirlemeye yönelik örnekler:

1. Cemre'nin boyu 176 cm'dir.
2. Okuldaki öğretmen sayısı 13'tür.
3. Aslı'nın matematik sınavından aldığı puan 98'dir.
4. Mehmet KPSS sınavında başarılı oldu.
5. Hemşirelik son sınıf öğrencilerinin hepsi sınıfı geçerek mezun oldu.
6. Ayşe sene boyunca derslerine konsantre olmakta zorluk çekiyor.

İlk örnekte ölçmeye konu olan özellik, fiziksel bir özellik olan “boy uzunluğu”dur. Bu örnekte Cemre'nin boyu, uzunluk ölçme aracı ile ölçülmüş ve bu ölçme aracında 176 değerine karşılık geldiği bulunmuştur. İkinci örnekte bir sayma işlemi yapılmış ve bunun sonucunda okulda 13 öğretmen olduğu belirlenmiştir. Üçüncü örnekte ise Aslı'nın matematik başarıları, bir matematik testinden aldığı puanla belirlenme çalışılmış ve bu puan başarı göstergesi olarak kullanılmıştır. Dördüncü ve beşinci örneklerde ise “başarılı olmak”, “sınıf geçmek”, “mezun olmak” gibi yorumlar yapılarak ölçme sonuçları bir ölçüte dayandırılmış ve değer yargısına ulaşmaya çalışılarak değerlendirme yapılmıştır. Altıncı ve son örnekte ise, öğrencinin sene boyunca derslerindeki hal ve tavırlarına yönelik bir durum belirleme yapılmıştır. Örneklerde görüldüğü üzere, ölçülen özelliğe göre ölçme ve değerlendirme süreçleri farklı şekillerde gerçekleştirilebilmektedir. Ölçmenin sağlam verilere dayanması, değerlendirme sürecinin geçerli ve güvenilir olmasını sağlamaktadır.

Ölçme ve değerlendirmenin temel işlevi, öğrencilerin eğitim ve öğretimle hedeflenen kazanımlara ne derecede sahip olduklarının ortaya konmasıdır. Ölçme ve değerlendirme olmaksızın bireylerin başarı veya başarısızlığını somut olarak ortaya koymak mümkün değildir. Ölçme ve değerlendirme yoluyla öğrencilerin öğretimle ilgili özel ihtiyaçları ve tercihleri ortaya çıkarılabilir, ölçümlere dayanarak hangi alanlarda başarılı olabileceklerine dair yönlendirmeler yapılabilir. Verilen kararların doğruluğu, ölçme

işleminde kullanılan ölçme araçlarının ve bu ölçme araçları ile elde edilen sonuçların geçerlilik ve güvenilirlik derecesiyle yakından ilişkilidir.

2. Ölçme ve Değerlendirmede Temel Kavramlar

Ölçüm: Ölçme sonucu elde edilen sayı veya semboller, ölçüm olarak isimlendirilir. Sınavdan alınan 60 puan, ağırlık olarak ifade edilen 70 kg, uzunluk olarak ölçülen 150 cm bir ölçümdür.

Ölçme Kuralı: Ölçme işleminin nasıl yapılacağını ifade eder. Yapılacak olan bir sınavda her bir sorunun kaç puana karşılık geldiğinin belirtilmesi ölçme kuralıdır.

Ölçüt: Ölçüt, değerlendirmenin neye göre yapılacağını belirtir. Değerlendirme yaparken ölçülen özelliğin sonucunda elde edilen ölçümün başarılı veya başarısız, yeterli veya yetersiz sayılabilecek bir performansı belirleme durumunda ölçütlerden yararlanılır. Ölçüt; ölçme sonuçları hakkında bir karara ve yargıya varmada kullanılan kriterlerdir. Örneğin; bir öğrencinin dersten başarılı olması için vize ve final değerlendirmesi toplamının 70 puan ve üzeri olması.

Değişken: Bir durumdan diğerine farklı değerler alabilen özelliklerdir. Buna göre insanların zekâ düzeyleri, ekonomik gelirleri, cinsiyetleri, boy uzunlukları, başarıları, tutumları birer değişkendir. Değişken olmayan özelliklere **sabit** denir. Bireyin herhangi bir özelliğinin değişken olması, bireyler arasında farklı değerler alabilmesi demektir. Bir özelliğinin değişken olması o özelliğinin ölçülmesini gerekli kılar. Değişken türleri Şekil 2’de gösterilmiştir.



Şekil 2. Değişkenlerin sınıflandırılması

Değişkenin özelliği miktar ve sayı olarak ifade ediliyorsa buna **nicel değişken** denir. Başarı puanı, uzunluk ve ağırlık ölçüleri nicel değişkenlere örnek olarak verilebilir. Değişkenin özelliği sınıflandırma şeklinde ifade ediliyorsa buna **nitel değişken** denir. Cinsiyet, eğitim durumu, medeni durum nitel değişkenlere örnek olarak verilebilir.

Değişkenler aldıkları değerlere göre **sürekli** veya **süreksiz** olarak sınıflandırılır. **Sürekli değişkenler** iki ölçüm arasında sonsuz sayıda değer alabilirken, **süreksiz değişkenler** ölçülen özelliklerle ilgili sadece sınırlı sayıda değer alır. Örneğin, cinsiyet durumu sadece kadın ve erkek değerleri alabildiğinden süreksiz bir değişkendir. Bir bireyin ağırlığı ise sürekli değişkendir.

Değişkenler sebep sonuç ilişkisi içinde ele alınıyorsa bu durumda **bağımlı** ve **bağımsız değişken** olarak adlandırılır. Araştırmanın amacını tanımlayan değişken **bağımlı değişken** (sonuç değişkeni); bağımlı değişkeni etkileyen, bağımlı değişkendeki değer değişimlerine neden olan değişken/değişkenler ise **bağımsız değişkendir**. Örneğin; “Sigara içmenin astım hastalığının gelişimi üzerine etkisi var mıdır?” hipotezinde bağımlı değişken astım, bağımsız değişken ise sigaradır.

Ölçmede Birim: Ölçülen özelliğin en küçük parçasına **birim** denir. Örneğin; bir evin kaç metrekare olduğunu ölçtüğümüzde birim metrekaredir. Meyve, sebze, bakliyat alırken birim kilo; kahve alırken ise gramdır.

Birimlerin Genelliği: Birimlerin herkes tarafından aynı şekilde anlaşılmasıdır. Örneğin; kilo, metre, cm denildiğinde Türkiye’de herkes aynı şeyi anlar. Çünkü bu birimler geneldir ve herkes tarafından bilinir.

Birimlerin Eşitliği: Birimler arasındaki mesafelerin aynı olmasıdır. Örneğin; farklı renklerde, metal, plastik veya mukavvadan yapılmış cetveller incelendiğinde, hepsinde 1 cm’ye karşılık gelen uzunluk eşit olacaktır. Birimlerin eşitliği ölçme sonuçlarının duyarlılığını artırır. Kulaç, adım ve karış gibi doğal birimler kişiden kişiye çeşitlilik göstereceğinden, bu ölçümler güvenilir değildir. Bu bakımdan başarıyı ölçmede eşit güçlük düzeyinde ve eşit puanda sorulardan oluşan bir sınav farklı puanlar üzerinden değerlendirilen bir sınavından daha duyarlı olacaktır.

Birimlerin Kullanışlılığı: Birimler seçilirken ölçülecek değişkene uygun olmasına dikkat edilmesidir. Örneğin; uzunluk ölçmek için metre, ağırlık ölçmek için terazi kullanmak kullanışlılık için örnek gösterilebilir.

2.1. Ölçme Türleri

Ölçme, **doğrudan** ve **dolaylı** ölçme olmak üzere ikiye ayrılır (Tablo 1). Ölçmeye konu olan özellik ile ölçme aracının özelliği aynıysa yapılan ölçmeye **doğrudan ölçme** denir. Cetvelle uzunluk, çift kollu terazi ile ağırlık ölçmek doğrudan ölçmeye bir örnektir. Bazı durumlarda bir özelliğin doğrudan ölçülmesi mümkün değilken, başka bir özellik yardımıyla veya ilgili özelliğin varlığına ya da yokluğuna gösterge olabilecek bir araçla ölçümler yapılır. Zekâ, yetenek, ilgi, tutum ve başarı gibi soyut özelliklerin doğrudan ölçümleri mümkün değildir. Yazılı sınavlarda ve testlerde öğrencilere yöneltilen sorularla belli bir konuda başarı düzeylerini ortaya koymak **dolaylı ölçmeye** örnek verilebilir.

Tablo 1. Doğrudan ve dolaylı ölçmelerin özellikleri

Doğrudan Ölçme	Dolaylı Ölçme
• Terazi ile ağırlık ölçümü, uzunluk, zaman ve hız gibi fiziksel özelliklerin ölçülmesinde kullanılır. • Aracın özelliği ve ölçülen nesnenin özelliği aynıdır. • Ölçme işleminin güvenilirliği yüksektir. • Ölçme sonucunu yorumlamak kolaydır.	• Eğitim ve psikoloji alanındaki ölçümlerde yaygındır. • Ölçülen özellik ile ölçme aracında kullanılan özellik farklıdır. • Doğrudan ölçümlere göre güvenilirliği düşüktür. • Ölçme sonucunu yorumlamak uzmanlık gerektirir.

2.2. Ölçek Türleri

Ölçek; ölçme işleminde kullanılan araçtır. Ölçme değerlendirmeye konu olan değişkenlerin/özelliklerin ölçülmesinde ölçeklerden yararlanır. Ölçmede farklı özellikler barındıran ölçekler temel olarak dört sınıfa ayrılmaktadır (Şekil 3).



Şekil 3. Ölçek türleri

Sınıflama Ölçekleri: En basit düzeyde ölçme bilgisi sağladığımız ölçek türüdür. Kişileri/nesneleri belli bir özelliğe sahip olup olmamalarına göre kategorize etmeyi amaçlar. Örneğin, öğretim türüne göre

öğrencileri sınıflandırmak istediğimizde, birinci öğretimde öğrenim gören ve ikinci öğretimde öğrenim gören öğrenciler olmak üzere iki kategoride ele alabilir, 1 ve 2 kodlarını verebiliriz. Medeni duruma göre bireyleri evli ve bekar olarak iki kategoride ele alabilir, evlilere E, bekarlara B harfini vererek kodlama yapabiliriz. Bu sınıflamalar, günlük yaşamı kolaylaştırmak için yapılan ölçmelerdir ve aldıkları sayısal değerler hiçbir matematiksel anlam taşımaz. Bu ölçek düzeyinde elde edilen ölçme sonuçları için her bir sınıfta tekrar eden değer sayısı olan frekans, buna bağlı olan yüzdeler ile ifade edilebilir, en çok tekrar eden değer olan mod saptanabilir.

Sıralama Ölçekleri: Bu ölçek düzeyinde yapılan ölçmeler, nesneleri ya da kişileri belli bir özelliğe sahip oluş derecesine göre sıralamayı gerektirir. Öğrencileri boylarına göre, sınavdan aldıkları puanlarına göre sıralamak, bu ölçek türü için örnek olarak verilebilir. Bu ölçek türü için bir öğrencinin, diğer öğrenciye göre daha yüksek veya düşük puan aldığı söylenebilir. Ancak bu düzeyde yapılan ölçmelerde elimizde sadece ölçülen özelliğin miktarından dolayı ölçme nesnesinin yer aldığı sıra bilgisi vardır, aradaki farka ilişkin bir bilgi yoktur. Bu ölçek türünde toplama, çıkarma, çarpma, bölme gibi işlemler yapılamaz. Asimetrik özellik mevcuttur ve $A > B$ ise $B < A$ bağıntısı kurulabilir. Bu ölçek türünde ölçme sonuçları üzerinde Spearman Brown Sıra Farkları Korelasyon Katsayısı hesaplanabilir.

Eşit Aralık Ölçekleri: Eğitim ve psikoloji alanında en çok kullanılan ölçektir. Bu alanlarda kullanılan ölçeklerin sıfırı gerçek sıfır olmadığı için bu ölçekler eşit aralıklı ölçek özelliği taşırlar. Bu ölçek düzeyinde gerçekleştirilen ölçme işlemlerinde eşit birimlerden söz edilebilir. Ancak başlangıç noktası keyfidir. Sıfır noktası, gerçek yokluk anlamı taşımaz. Termometre, takvim ve eğitimde kullanılan ölçme araçlarının verdiği ölçme bilgisi, eşit aralık ölçeği düzeyindedir. Aritmetik ortalama, standart sapma, korelasyon katsayısı gibi hesaplamalar yapılabilir. Eşit aralıklı ölçeklerdeki sıfır noktası, ölçülen özelliğin mutlak yokluğunu göstermediği için, bu ölçeklerle elde edilen ölçümler üzerinde, toplama ve çıkarma işlemleri yapılabildiği halde, oran belirtmek, çarpma ve bölme işlemleri yapmak mümkün değildir. İki ölçüm arasındaki farkın miktarı anlamlıdır. Zekâ puanı 120 olan birinin zekâ puanı 60 olan birinden, 60 puan daha fazla aldığı söylenebildiği halde iki kat daha zeki olduğu söylenemez.

Oranlı Ölçekler: Bu ölçek türünde gerçek sıfır noktası söz konusudur. Diğer bir deyişle bu ölçek düzeyinde yapılan ölçmelerde ölçülen özellik için sıfır, gerçek yokluk anlamını taşır. 0 adet kalem, 0 öğrenci, bahsedilen kişi veya nesnenin olmadığını ifade eder. Bu ölçek türünde, her türlü istatistiksel işlem kullanılabilir. Kilosu 50 olan birisi, 25 kg olan birisinin iki katı ağırlıktadır denilebilir.

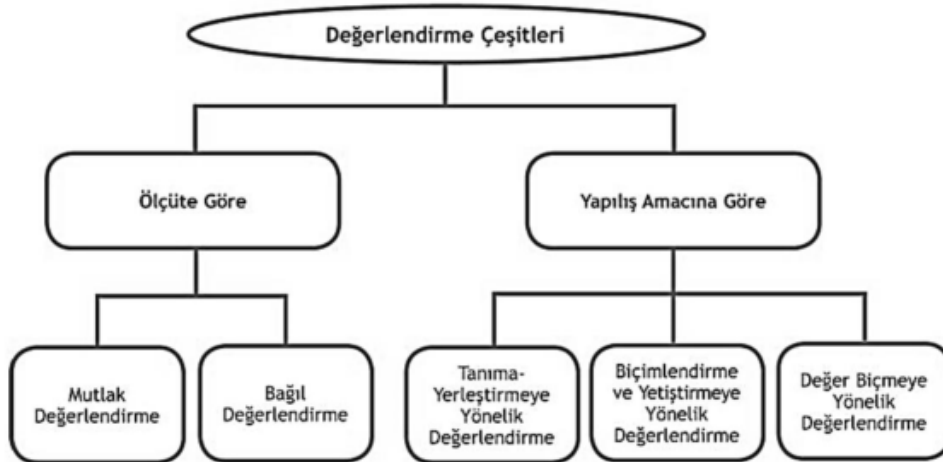
2.3. Değerlendirme

Ölçme sonuçlarının anlamlandırılmasını, bir ölçütle karşılaştırılarak karar verilmesini ve yorum yapılmasını sağlar. Değerlendirme bir yargılama işlemidir. Her değerlendirme, bir ölçme sonucu ile bir ölçütün karşılaştırılmasına dayanır (Şekil 4).



Şekil 4. Değerlendirme süreci örneği

Değerlendirme çeşitleri, kullanılan ölçüte ve kullanılış amacına göre çeşitlilik göstermektedir (Şekil 5).



Şekil 5. Değerlendirme çeşitleri

Tablo 2. Ölçüte göre değerlendirme türleri

Mutlak Değerlendirme	Bağıl Değerlendirme
• Mutlak değerlendirmede, mutlak ölçüt adı verilen önceden belirlenmiş bir kriter kullanılır. • Mutlak ölçütle yapılan değerlendirmede, her bir bireyin (öğrencinin) değerlendirme sonucu, içinde bulunduğu sınıf, topluluk, gruptan bağımsızdır. • “Her koyun kendi bacağından asılır” deyimi bu değerlendirme için uygundur. • Örneğin; bir sınavdaki 50 sorudan 30'unu doğru yanıtlamış olmak, sınavdan 100 üzerinden en az 65 almak ya da 10 hedef davranışın 7'sini başarmış olmak gibi bireyin sadece kendi başarısıyla değerlendirileceği durumlar, mutlak değerlendirme kapsamına girer. Eğer 50 sorudan 30'unu doğru yanıtlamış olmanın başarı sayılacağı bir sınavda bir öğrenci, 35 soruyu doğru yanıtlamışsa, sınavın uygulandığı sınıftaki diğer arkadaşlarının tümü 50 soruyu da doğru cevaplasa ya da tümü hiçbir soruyu da doğru bilmemiş olsa, bu öğrenci bu sınav için belirlenen ölçüte göre daima başarılı sayılacaktır. • Mutlak değerlendirmenin yapıldığı sınıf ortamlarında ders çalışırken öğrenciler arası bilgi alışverişinde bulunmak, yardımlaşmak daha yaygındır (Sınav ortamında değil).	• Bağıl değerlendirmede, bağıl ölçüt adı verilen ölçme sonuçlarına göre sonradan belirlenmiş bir kriter kullanılır. • Bağıl ölçütle yapılan değerlendirmede, her bir bireyin (öğrencinin) değerlendirme sonucu, içinde bulunduğu sınıf, topluluk, gruptan etkilenir. • Her öğrenci, sınıftaki diğer arkadaşlarının da başarılarını önemsemek durumundadır. • Örneğin; bir sınav sonucunda, sınıftaki öğrencilerin aldıkları puanların aritmetik ortalamasının üstünde puan almış olmanın başarının ölçütü olması ya da tüm öğrenciler içinde sınavdan en yüksek puan alan ilk üç öğrencinin arasına girmenin başarı sayılması gibi durumlarda öğrencinin sadece kendi başarısı değil, içinde bulunduğu grubun başarı durumu da başarılı sayılabilmesi için önem taşımaktadır. • Bağıl değerlendirmenin yapıldığı sınıf ortamları öğrencileri daha çok çalışmaya sevk eder. Ancak bu tür değerlendirmenin yapılması, öğrenciler arası rekabeti arttıracak gibi, öğrencilerin arasındaki dayanışmanın azalmasına sebep olacaktır.

Tablo 3. Amaca göre değerlendirme türleri

Tanım-Yerleştirmeye Dönük Değerlendirme	Biçimlendirme ve Yetiştirmeye Dönük Değerlendirme (Formatif Değerlendirme)	Değer Biçmeye Dönük Değerlendirme (Summatif Değerlendirme)
- Eğitim sürecinin başında yapılır. - Amaç, öğrenciyi tanımak ve onu uygun olan eğitim programına yerleştirmektir. - Bir bakıma öğrencinin hazır bulunuşluk düzeyini belirlemeyi amaçlar. - Bu değerlendirmede amaç öğrencinin aldığı nota göre en doğru programa (ya da sınıfa) yerleşmesini sağlamak olduğu için ASIL AMAÇ öğrenciye NOT vermek DEĞİLDİR. - Her öğrencinin kendisine uygun olan programa yerleştirilmesi, zaman, para ve emek açısından ekonomi sağlar. - Bu değerlendirme türüne; seviye tespit sınavları, muafiyet sınavları örnek olarak verilebilir.	- Eğitim süreci devam ederken, belirli periyotlarda yapılır. - Amaç, öğrencinin eksiklerini, hatalarını belirleyerek bunları düzeltmek, öğrenciye geri bildirimde bulunmaktır. - Bir bakıma eğitim süreci bitmeden, belirli aralıklarla eğitimin hedefine ulaşıyor olup olmadığı hakkında fikir sahibi olmak için yapılır. Böylelikle eksiklikler ya da yapılan yanlışların birikerek başa çıkılmaz bir hal alması engellenmiş olur. - Bu değerlendirmede amaç öğrencinin aldığı nota göre geri bildirimde bulunmak olduğu için ASIL AMAÇ öğrenciye NOT vermek DEĞİLDİR. - Eğitimde istenen hedefe doğru daha sağlam adımlarla ilerleme sağlanır. - Bu değerlendirme türüne; quizler, ünite testleri, tarama testleri örnek olarak verilebilir.	- Eğitim sürecinin sonunda yapılır. - Amaç, öğrenciye tamamlanan eğitim sonucundaki seviyesini, ulaştığı noktayı bildirmektir. - Bir bakıma öğrenciye eğitim süreci sonunda aldığı nota göre bir değer biçilmiş olur. - Bu değerlendirmede ASIL AMAÇ öğrenciye NOT VERMEKTİR. Çünkü artık hedeflenen eğitim süreci sona ermiştir. - Bu değerlendirme türüne; bitirme sınavları, uzmanlık sınavları, sertifika sınavları, finaller örnek olarak verilebilir.

3. Ölçme Araçlarında Bulunması Gereken Özellikler

Ölçme sonuçlarının olabildiğince gerçeğe yakın ve objektif olabilmesi, bu sonuçlara göre yapılacak değerlendirmelerin de o kadar yerinde ve isabetli olmasını sağlar. Bu sebeple ölçme araçlarından yararlanmak kaçınılmazdır. Ancak ölçme araçlarının da birtakım niteliklere sahip olması istenir: güvenilirlik, geçerlilik ve kullanılabilirlik. Ölçme aracının güvenilirlik ve geçerliliğe sahip olması mutlaka gerekli olan nitelikleridir. Kullanılabilirlik ise güvenilirlik ve geçerliliğin sağlanması koşuluyla istenen bir diğer niteliktir.

3.1. Güvenilirlik

Güvenilirlik genel tanımıyla “ölçme sonuçlarının tesadüfi hatalardan arınık olma derecesi” olarak ifade edilebilir. Ortaya çıkabilecek tesadüfi hatalar ne kadar azaltılabilirse o kadar güvenilir ölçme sonuçları elde edilmiş olacaktır. Bu sebeple güvenilirlikte “ya hep ya hiç” söz konusu değildir, güvenilirlik bir derece ifadesidir. Güvenilirlik farklı kavramlarla da ifade edilebilir: tutarlılık, kararlılık ve duyarlılık.

Tutarlılık: Ölçme aracı, aynı gruba iki farklı kişi tarafından uygulandığında benzer sonuçlar alınması, aynı sınav kağıtlarını değerlendiren iki eğitimcinin verdikleri puanların birbirine yakın olması ve ölçme aracı içerik bakımından benzer olan başka bir araçla birlikte uygulandığında benzer sonuçlar alınması şeklinde açıklanabilir.

Kararlılık: Ölçme aracının belli bir arayla aynı öğrenci grubuna tekrar uygulandığında benzer sonuçlar vermesidir. Burada kullanılan aracın ve grubun aynı olması önemlidir. Kararlı ölçüm yapma özelliği olan bir ölçme aracının tekrar uygulanması durumunda benzer sonuçlar vereceği kabul edilir. Tutarlı sonuçlar elde etmek için uygulama koşullarını standartlaştırmak önemlidir. Tutarlılık ile kararlılık arasındaki tek fark kararlılığın, elde edilen sonuçların zamana karşı direnç göstermesi ve benzerliğini koruyabilme derecesini de ifade ediyor olmasıdır.

Duyarlılık: Ölçme aracının biriminin olabildiğince hassas, küçük olmasının derecesidir. Ölçme aracında kullanılan birim ne kadar küçükse ölçmenin duyarlılığı o ölçüde artacaktır. Örneğin, uzunluk ölçerken kullanılacak cetvel mm bölmeli ise, cm bölmeli bir cetvele göre daha duyarlı sonuçlar verecektir. Eğitimde en çok kullandığımız ölçme araçlarımız sınavlardır ve sınavları oluşturan her bir birim de sorulardır. Sınavların duyarlılığı da soru sayısının artmasıyla sağlanır. Bir sınavda ne kadar çok soru sorulursa, öğrencilere ilişkin o kadar duyarlı sonuçlar elde edilir.

Ölçme Aracının Güvenilirliğini Arttırmanın Yolları:

1. Soru sayısını arttırmak: Yukarıda da belirtildiği gibi ne kadar çok soru sorarsak o kadar duyarlı yani güvenilir sonuçlar elde edebiliriz. Ancak soru sayısının da öğrencinin seviyesi, yaşı, sınavın ortamı, süresi gibi durumlara bağlı olarak mutlaka makul bir düzeyde tutulması gerekmektedir. Aksi takdirde güvenilir sonuç elde etmek için öğrencilerin cevaplayamayacakları kadar çok soru sormak uygun olmayacaktır.

2. Soruları açık, anlaşılır ve net ifadelerle hazırlamak: Sınavların amacı öğrenciye ilişkin ölçülecek özellik açısından olabildiğince gerçek verilere ulaşmaktır. Bu sebeple ölçülmek istenilen özellik için sorulan sorular net ve açık olmalıdır ki öğrencinin soruyu tam olarak anlayıp, bilip bilmediğini ortaya koyabilmesi mümkün olsun. Şayet öğrenci soruyu tam anlamazsa bizim öğrenmeye çalıştığımız bilgi düzeyini gözlemlememiz de mümkün olmayacaktır. Sınavlar “ölçülmek istenilen özelliğe ilişkin bilgi toplamada ölçme aracı” olmanın dışında başka hiçbir amaçla kullanılmamalıdır.

3. Öğrencilerin sınava olan motivasyonunu arttırmak: Bunun iki yolu vardır;

- *Eğitimcinin ya da sınavı yapan kişi(ler)in tutumları:* “Sınav” kavramı öğrencilerde az ya da çok kaygı yaratır. Sınav sürecinde sınavı yapan kişilerin öğrencilerin bu kaygılarını azaltmaları, onların rahatlamasını sağlamaları çok önemlidir. Çünkü öğrencinin sınav kaygısı ne kadar azalır, gerçek durumunu ortaya koyma olasılığı o kadar artacak, daha az hata yapacak ve sınav sonuçları da böylece daha güvenilir olabilecektir.
- *Sınav sorularının kolaydan zora doğru sıralanması:* Öğrencilerin sınavdaki kaygıları soruları çözebildikleri ölçüde azalmaya başlar. Bu sebeple özellikle ilk sorular kolay olacak şekilde düzenlenebilir. Soruların zorluğu giderek arttırılırsa, öğrencilerin soruları yapabildikleri ölçüde kaygıları azalacak, sınava olan motivasyonları artacak ve öğrencilere ilişkin elde edilen sonuçlar böylece olabildiğince gerçeğe yakın ve güvenilir olacaktır.

4. Sınav süresini az ya da çok fazla olmayacak şekilde ayarlamak: Sınav süresinin öğrencinin seviyesine, sınavın zorluk düzeyine göre olabilecek en makul düzeyde vermek de ölçme sonuçlarının güvenilirliği açısından çok önemlidir. Yetersiz sınav süresi öğrencilerin soruları boş bırakmasına ya da bilmedikleri halde tahminle doğru cevaplamalarına sebep olacaktır ki bu da sonuçların güvenilirliğini düşürür. Benzer şekilde gerekenden çok daha fazla verilen sınav süresi de güvenilirliği olumsuz yönde etkilemektedir. Sınav süresinin çok fazla verilmesi halinde öğrenci yorgunluk vb. sebeplerle doğru verdiği cevaplarda çelişkiye düşmeye başlayarak yanlış cevaplara yönelmekte ya da kopya çekme eğilimi göstermektedir.

5. Tesadüfi hata kaynaklarını olabildiğince azaltmak: Hata kaynakları (öğretmen, öğrenci, sınav ortamı ve sınav araçları) ne kadar engellenirse, ölçme sonuçları da o kadar güvenilir olacaktır. Öğretmenin sınav sonuçlarını okurken göstereceği dikkat, öğrencinin hem fiziksel hem de psikolojik açıdan sınava hazır gelmesi, sınav ortamının ve sınav aracının sahip olması gereken özellikleri olabildiğince taşıması, sınavdan elde edilen sonuçların da o ölçüde güvenilir olmasını sağlayacaktır.

3.1.1. Güvenilirlik Yöntemleri

Güvenilirliğin kestirilmesinde korelasyon katsayısından yararlanılmaktadır. Korelasyon değeri normalde $[-1, +1]$ aralığında bulunurken, güvenilirlik değeri olarak $[0, +1]$ arasında değer alır. Güvenilirlik katsayısının tahmin edilmesinde kullanılan yöntemler iki ana başlık altında ele alınabilir. Bunlar **tek uygulamaya dayalı** yöntemler ve **birden çok uygulamaya dayalı** yöntemlerdir (Şekil 6).



Şekil 6. Güvenilirlik yöntemleri

Birden Çok Uygulamaya Dayalı Yöntemler

Test-Tekrar Test: Bir testi aynı gruba aralıklı olarak iki kez uygulamaya dayalı bir yöntemdir. Burada güvenilirliği hesaplanacak test aynı gruba belli bir aralıkla iki kez uygulanır. İki uygulama arasındaki korelasyon yüksekse, yani 1'e ne kadar yakınsa, güvenilirlik o kadar yüksek olacaktır. Test-tekrar test yönteminde en önemli durum, iki uygulama arasındaki sürenin ne kadar olması gerektiğidir. İki uygulama arasındaki sürenin unutmayı önleyecek kadar kısa, hatırlama etkisini önleyecek kadar uzun olması gerekir. İlk uygulama günün erken saatlerinde, haftanın ilk günü klimalı bir sınıfta yapılmışsa, ikinci uygulama da benzer zamanda ve benzer koşullar sağlanarak tekrar edilmelidir. Test-tekrar test uygulamasında ilgili özelliğin zaman içinde çok fazla değişiklik göstermesi beklenmiyorsa, ikinci

uygulama ilk uygulamadan iki hafta ve üçüncüsü bir ay sonra olmak üzere birden çok olarak yapılabilir. Bu durumda, aralarındaki zaman farkı daha az olduğundan ve ölçümlerin daha kararlı olması bekleneceğinden birinci ve ikinci uygulama arasındaki korelasyonun, birinci ve üçüncü uygulama arasındaki korelasyondan yüksek olması beklenir.

Eşdeğer (Paralel) Testler: Bu yöntemle güvenilirlik tahmini için, bir testin en az iki eş değer (denk) formunun aynı öğrenci grubuna eşzamanlı uygulanması gerekir. İki uygulama arasındaki korelasyona göre testin güvenilirliği yorumlanır. Test-tekrar testte olduğu gibi testin uygulandığı öğrenci grubu aynı olmalıdır. Ancak test-tekrar test yönteminde ikinci uygulamada daha uzun bir ara ile testin tıpatıp aynısı kullanılırken, burada ilk testle içeriği aynı olan, ancak soruları farklı, güçlük bakımından eş değer bir test ikinci sınav olarak verilir. Uygulamada eş değer formlardan ilki önce verilir ve aynı gün, aynı oturumda, iki test arasında öğrencinin temel ihtiyaçlarını karşılayacağı kadar bir ara verildikten sonra, ikinci form uygulanır.

Tek Uygulamaya Dayalı Yöntemler

Test Yarılama: Yarıya bölme yöntemleri geleneksel yöntem olarak da adlandırılır. Uygulama kolaylığından dolayı test güvenilirliğini tahmin etmede en çok tercih edilen yöntemlerden biridir. Testin iki yarıya bölünmesiyle güvenilirliğinin tespit edilmesi yönteminde, uygulamadan sonra test iki eş değer yarıya bölünür. Genellikle konu dağılımını dengede tutmak amacıyla tek numaralı sorular ve çift numaralı sorular olmak üzere aynı teste ait iki puan elde edilir. Öğrencilerin iki yarıdan aldıkları puanlar arasındaki korelasyon katsayısı hesaplanır. Elde edilen bu değer testin her bir yarısının güvenilirlik değerini vermektedir. Ancak bu yöntem testteki tüm sorular homojen (güçlük düzeyleri aynı, aynı kapsamı ölçen sorular) olduğu taktirde kullanılabilir. Aksi durumda bu yöntemle hesaplanan güvenilirlik değeri doğru olmayacaktır.

Kuder Richardson 20 ve 21 (KR-20 ve KR-21): Bu formüller doğru cevabın “1” puan, yanlış veya boş bırakılmış soruya verilen puanın “0” olduğu tüm testlerin güvenilirliğini hesaplamak için kullanılır. KR-20 ve KR-21, ikili cevap sistemli testlerin (çoktan seçmeli testler, doğru/yanlış) iç tutarlılık katsayısıdır. KR-20 ve KR-21 değerleri, [0, +1] arasında değişmekte olup, sonuç 1’e ne kadar yakınsa, testin iç tutarlılığı o kadar yüksek demektir. KR-20, her maddenin birbirine paralel olduğu varsayımına dayalı olarak hesaplanan bir korelasyon katsayısıdır. Maddelerin güçlük düzeyleri biliniyor ya da hesaplanabiliyorsa kullanılır. KR-21 ise maddelerin güçlük düzeyleri bilinmiyor ya da hesaplanamıyorsa kullanılır. Tüm maddelerin güçlüklerinin birbirine eşit olduğu varsayılır. Alt testlerden oluşan bir ölçüm aracının (sayısal yetenek, sözel yetenek) iç tutarlılık katsayısının düşük çıkması aracın güvenilir olmadığı

anlamına gelmez. Farklı ögeler tutarlı sonuçlar vermeyeceğine göre, böyle bir testin iç tutarlılığının yüksek olması beklenemez. Bu nedenle güvenilirlik yöntemine karar verilirken veya sonuçlar değerlendirilirken hangi güvenilirlik hesaplama yönteminin amacımıza uygun olduğu iyi değerlendirilmelidir. Testin iç tutarlılığı hakkında bilgi veren KR-20 ve KR-21 arasında seçim yapmak gerekirse, daima $KR-20 > KR-21$ olduğu hatırlanmalıdır. Madde güçlük indeksleri hesaplanabiliyorsa KR-20'den yana tercih kullanılmalıdır. KR-20 ve KR-21 değerleri yüksek ise, testteki maddelerin homojen bir yapı sergilediği ve iç tutarlılıklarının yüksek olduğu yorumu yapılabilir.

Cronbach Alfa: En çok bilinen güvenilirlik belirleme yöntemidir. Bu yöntem, maddelerin birbiriyle uyumuna bakarak güvenilirliğin kestirilmesini sağlar. KR-20 ve 21 yöntemine benzerdir. Tek farkı anket ve tutum ölçekleri gibi çoklu puanlanan (0-1 puanlamadan farklı puanlama yapılan) maddelerden (sorulardan) oluşan testlerde uygulanır.

KR-20, KR-21 ve Cronbach Alfa iç tutarlılık sonuçlarını yorumlamak için Tablo 4'te verilen yönergelerden yararlanılabilir.

Tablo 4. İç tutarlılık değerlerinin yorumlanması

Güvenilirlik sonucu	Yorum
0.90 ve üzeri	Mükemmel derecede güvenilirlik; en iyi standart testler düzeyinde.
0.80 – 0.90	Çok iyi derecede güvenilirlik; bir ders testi/sınavı için çok iyi düzeyde.
0.70 – 0.80	İyi derecede güvenilirlik; bir ders testi/sınavı için iyi düzeyde ve çoğu kişi için uygun. Muhtemelen geliştirilmesi gereken birkaç madde mevcuttur.
0.60 – 0.70	Düşük derecede güvenilirlik; bu testin notları belirlemek için başka ölçütlerle (Örn; ek testler/sınavlar) desteklenmesi gerekmektedir. Muhtemelen geliştirilmesi gereken bazı maddeler mevcuttur.
0.50 – 0.60	Çok düşük derecede güvenilirlik; çok kısa olmadığı sürece (10 veya daha az madde) sınavın/testin gözden geçirilmesi gerekir. Notlandırma için testin kesinlikle başka ölçütlerle (Örn; ek testler/sınavlar) desteklenmesi gerekir.
0.50 veya altı	Sorgulanabilir derecede güvenilirlik; bu sınavın/testin ders notuna büyük bir katkısı olmaması gerekir ve revizyona ihtiyacı vardır.

3.2. Geçerlilik

Bir ölçme aracının ölçmesi istenilen özelliği ne derece ölçebildiğinin derecesine o ölçme aracının geçerliliği adı verilir. Bir başka deyişle, ölçme aracının ölçtüğü değişkeni (özelliği, niteliği) başka hiçbir değişken karışmaksızın ölçebilmesinin derecesidir. Bir ölçme aracının (test, sınav, anket vb.) hatasız ölçebilmesi güvenilirliğini, ölçülen özelliği başka özellik karışmaksızın ölçebilmesi de geçerliliğini

gösterir. Ölçme aracının geçerli olabilmesi için güvenilir olması şarttır ancak yeterli değildir. Güvenilir olan her ölçme aracı geçerli olmayabilir. Bir ölçme aracının geçerliliği varsa mutlaka güvenilirirdir.

Ölçme Aracının Geçerliliğini Arttırmanın Yolları:

1. Test içeriği ve soru sayısı: Test maddeleri dersin içeriğinin iyi bir örnekleme olmalıdır. Programı dengeli ve yeterli olarak temsil eden bir testin geçerliliği yüksek olacaktır. Soru sayısını arttırmak güvenilirliği, böylece geçerliliği arttırır. Ancak sınava kapsam dışından eklenecek her soru kapsam geçerliliğini tehdit eder. Bu sebeple her sorunun mutlaka kapsam dâhilinde olmasına dikkat edilmelidir. Her yıl aynı sorular kullanılmamalı, sınavdan önce sorular ile ilgili ipucu verilmemeli, sorular belirli bir kaynaktan aynen alınmamalıdır.

2. Madde/soru özellikleri: Madde istatistikleri yoluyla testin geçerliliği artırılabilir. Madde bazında **orta güçlük düzeyinde** ve **ayrıt ediciliği yüksek** sorulardan oluşturulan bir testin geçerliliği yüksek olacaktır.

3. Güvenilirliği arttırmak: Güvenilirliği arttırarak geçerliliğin arttırılması sağlanabilir. Kopyanın önlenmesi güvenilirliği arttıracak tedbirlerin başında gelir. Puanlamada objektiflik, sınav birden fazla kişi tarafından puanlanacaksa puanlayıcılar arası tutarlılık benzer şekilde önemlidir.

4. Geçerlilik $\leq \sqrt{\text{Güvenilirlik}}$: Bir testin geçerliliği, güvenilirlik katsayısının karekökünden daha yüksek değer alamaz. Bu nedenle geçerliliği belli bir dereceye kadar arttırmak mümkündür. Örneğin; güvenilirliği 0.81 olan bir testin geçerliliği 0.90 değerinden daha yüksek olamaz. Dolayısıyla geçerliliği artırmanın en temel yolu güvenilirliği arttırmaktır.

5. Testin yansız olması: Maddenin aynı testin içinde tekrar sorulması güvenilirliği artırsa da yetenek, zeka ve başarı gibi pek çok konuda testler geliştirilirken ölçüm sonuçlarında öğrenciler arası kültürel farklılıkların etkili olmadığından emin olunmalı, testin belli bir gruba avantaj sağlaması önlenmelidir. Kısacası test yanlı ölçüm yapmamalıdır. Örneğin; bir testte soruların mekanik, futbol gibi daha çok erkeklerin dikkatini çeken konulardan çıkması erkeklere bu sorularda avantaj sağlayacaktır. Bu şekilde cinsiyet ayrımcılığı yapan bir test, sınava sadece erkekler girmiyorsa uygulanmamalıdır.

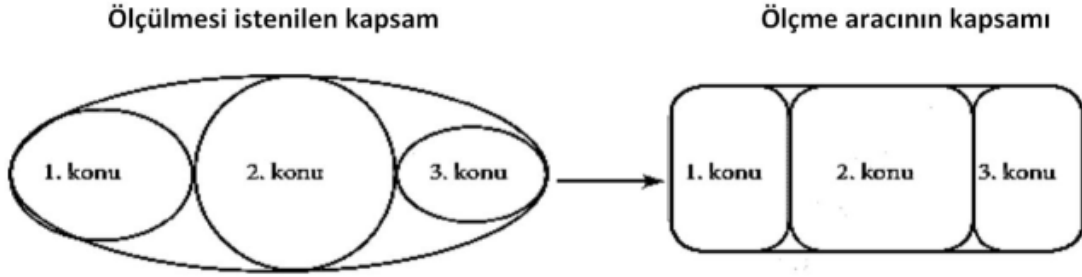
3.2.1. Geçerlilik Yöntemleri

Geçerlilik yöntemleri; kapsam, ölçüte dayalı ve yapı geçerliliği olmak üzere üç ana başlık altında ele alınmaktadır (Şekil 7).



Şekil 7. Geçerlilik yöntemleri

Kapsam Geçerliliği: Ölçme aracının ölçmesi istenilen kapsamı, dengeli bir şekilde ne eksik ne de fazla olmaksızın ölçebilmesinin derecesidir.



Yukarıdaki şekille de açıklanmaya çalışıldığı gibi, ölçülmesi istenilen kapsam farklı içerik ve yoğunluktaki üç konudan oluşuyorsa, bu kapsamı ölçecek ölçme aracı da kapsam dahilindeki üç konunun üçünü de ne eksik ne fazla olmaksızın ve yoğunluklarına uygun şekilde soru sayısına sahip olacak kapsamda oluşturulmalıdır. Kapsam geçerliliğinin belirlenmesinde iki yol izlenebilir:

- **Belirtke tablosu hazırlamak:** Bu yola rasyonel (mantıksal) yaklaşım da denir. Belirtke tablosu bir boyutunda (satır) ölçmenin yapıldığı konuların ana başlıklarının; diğer boyutunda (sütun) ise davranışların hangi bilişsel boyutta yer aldığını göstererek, hücrelerde satır ve sütunda kesişen davranışlardan kaç tane bulunduğunu belirten bir tablodur. Ölçme aracının belirtke tablosundaki davranışları ne derecede ölçtüğü, konu alanında ve ölçme alanında uzman olan kişilerin görüşleri alınarak incelenir. Bunun sonucunda da kapsam geçerliliğine ilişkin bir kanıya varılır. Önemli olan ölçme aracının tablodaki tüm davranışları yine tabloda belirtilen ağırlıkta yansıtır olmasıdır.

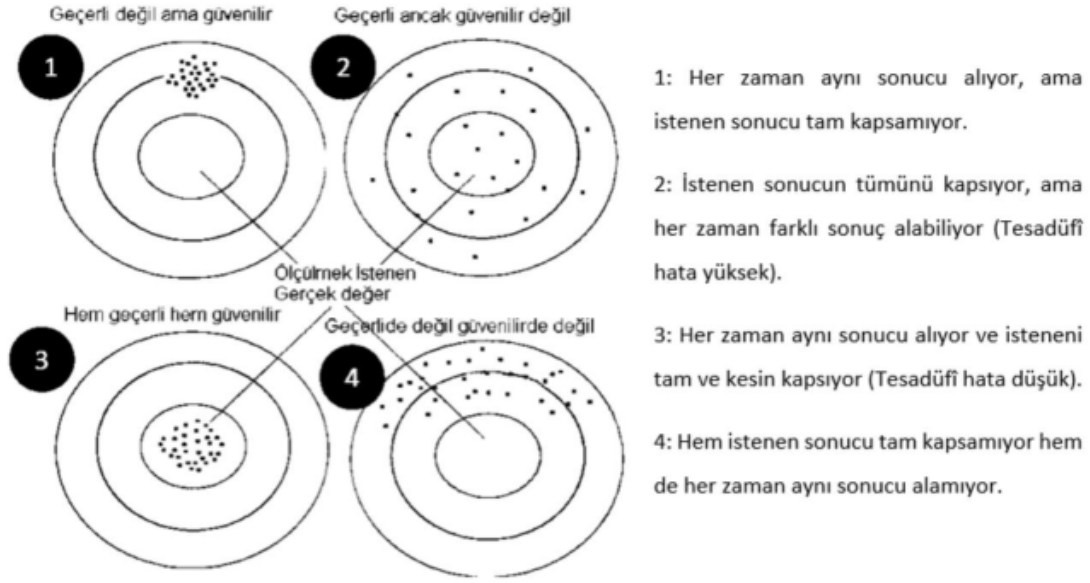
- **İstatistiksel hesaplama:** Ölçmek istenilen özellik ya da özelliklere ilişkin daha önceden hazırlanmış, güvenilirlik ve geçerliliğinin yüksek olduğu belirlenmiş bir ölçme aracı ile hazırlanan (kapsam geçerliliği hesaplanmak istenen) yeni ölçme aracı aynı gruba uygulanarak aralarındaki korelasyon hesaplanır. Hesaplanan bu değer kapsam geçerliliğinin bir göstergesi olarak yorumlanır.

Ölçüt Dayanaklı Geçerlilik (Ölçüt Geçerliliği): Ölçüt geçerliliği, hazırlanan ölçme aracından elde edilen puanlar ile ölçüt (kriter) kabul edilen diğer bir ölçme aracından elde edilen puanlar arasındaki korelasyon katsayısı hesaplanarak bulunur. Burada ölçüt olarak kabul edilen ölçme aracının güvenilir ve geçerli olması gerekmektedir. Ölçüt kabul edilen ölçme aracıyla elde edilen sonuçlar ile hazırlanan ölçme aracından elde edilen sonuçlar arasındaki korelasyon katsayısı ne kadar yüksek çıkarsa hazırlanan ölçme aracının da o derece geçerli ölçme sonuçları verdiği düşünülür. Ölçüt geçerliliği, ölçüt olarak seçilen ölçme aracının hazırlanan ölçme aracından önce ya da sonra uygulanmasına göre ikiye ayrılır;

- **Yordama Geçerliliği:** Ölçülen davranış ya da özelliğın gelecekteki durumu hakkında tahminde bulunmaya **yordama** adı verilir. Yordama geçerliliğinde, hazırlanan ölçme aracıyla elde edilen puanlar gelecekteki bir durumu tahmin etmekte kullanılır. Bu sebeple ölçüt kabul edilecek puanlar henüz mevcut değil, gelecekte var olacak puanlardır. Hazırlanan ölçme aracından elde edilen puanların, gelecekte ortaya çıkacak ve asıl durumu gösterecek puanlar ile arasındaki korelasyon değeri ne kadar yüksek çıkarsa, hazırlanan ölçme aracının yordama geçerliliğinin o derece yüksek olduğu yorumu yapılır. Örneğın; KPSS'ye hazırlık etütleri, asıl sınav öncesi öğrencilerine deneme sınavları (tahminde bulunmak üzere hazırlanan ölçme aracı) yaparlar. Bu sınavların amacı öğrencilerin asıl sınavdaki (gelecekte ortaya çıkacak, puanları ölçüt olarak kullanılacak ölçme aracı) durumları hakkında fikir edinebilmelerini sağlamaktır. Böyle bir deneme sınavından alınan puanlar ile öğrencilerin asıl sınavdan aldıkları puanlar arasındaki korelasyon hesaplanırsa, elde edilecek değeri etütün uyguladığı sınavın yordama (tahmin) geçerliliğinin ölçüsünü verecektir.
- **Zamandaş Geçerliliği:** Hazırlanan ölçme aracının, ölçülen davranış ya da özelliğın var olan durumu uygun bir şekilde ölçüp ölçmediğini belirlemek için **zamandaş** geçerliliğinden yararlanılır. Halihazır geçerlilik, mevcut hal geçerliliği, eşzaman geçerliliği ve uygunluk geçerliliği gibi farklı isimlerde tanımlanabilmektedir. Bu geçerlilikte, hazırlanan ölçme aracıyla elde edilen puanların var olan durumu doğru yansıtır yansıtmadığı tespiti yapılır. Bu sebeple ölçüt kabul edilecek puanlar zaten mevcuttur ve bu puanların elde edildiğı ölçme aracının

güvenilirlik ve geçerliliğinin yüksek olması gerekmektedir. Hazırlanan ölçme aracından elde edilen puanların, var olan durumu gösteren puanlar ile arasındaki korelasyon değeri ne kadar yüksek çıkarsa hazırlanan ölçme aracının zamandaş geçerliliğinin o derece yüksek olduğu yorumu yapılır. Örneğin; orta öğretimdeki bir sınıfın ders yılı ortalarında Türkçe öğretmeni değişmiş olsun. Bu sınıfta derse başlayan öğretmen bir süre sonra uyguladığı Türkçe sınavının zamandaş geçerliliğini belirlemek istesin. Bu durumda Türkçe öğretmeni kendi uyguladığı sınavdan öğrencilerinin aldıkları puanlar ile öğrencilerinin önceki sınıftaki başarı puanları (bu puanların güvenilir ve geçerli olması gerekir) ile korelasyonuna bakarsa, bulacağı değer öğretmenin uygulamış olduğu sınavın zamandaş geçerliliğinin ölçüsünü verecektir.

Yapı Geçerliliği: Bireylerde var olduğu kabul edilen ancak gözle görülemeyen özellikler “yapı” olarak adlandırılır. Örneğin; zeka, tutum, ilgi, kaygı vb. gibi bireysel özellikler birer yapıdır. Bu yapılar gözlenemediği için bunlarla ilişkisi olduğu kabul edilen maddelere (sorulara) verilen cevaplarla, bireylerin bu yapılara sahip olup olmadıkları gözlenmeye çalışılır. Yapı geçerliliği ölçme aracının teoride var olduğu düşünülen, gözle görülemeyen bu yapıları pratikte ölçebilme derecesine denir. Başarı testleri için yapı geçerliliği testin belli konular çerçevesinde hazırlanması ve testi oluşturan maddelerin homojen bir yapı sergilemesine bağlıdır. Maddeleri belli bir konuda olan ve çoğunluğu orta güçlük seviyesinde maddelerden oluşan bir test, bu bakımdan daha ümit verici olacaktır. Örneğin; matematik başarısını ölçmek amacıyla hazırlanan bir testin, matematik dışında (Türkçe, Fizik vb.) başka alanlardaki başarıyı ölçmeden, yalnızca matematik başarısını doğru bir şekilde ölçebilme derecesi o testin yapı geçerliliği ile ilgilidir. Yapı geçerliliği her ölçme aracında olması gereken önemli bir geçerlilik türüdür. Yapı geçerliliğinin incelenmesinde alan uzmanlarının görüşünün alınmasıyla birlikte faktör analizi gibi istatistiksel yöntemler de kullanılmaktadır.



Şekil 8. Geçerlilik ve güvenilirlik arasındaki ilişki

3.3. Güvenirlilik ile Geçerlilik Arasındaki İlişki

- Güvenirlilik, geçerlilik için bir ön koşuldur. Bir sınav geçerli ise, aynı zamanda güveniliridir.
- Bir ölçme aracının güvenilir olması her zaman geçerli olduğu anlamına gelmez.
- Geçerlilik, güvenilirliği kapsayan bir kavramdır.
- Geçerlilik her türlü hatadan etkilenmesine rağmen, güvenilirlik yalnızca tesadüfî hatadan etkilenir.
- Güvenirlilik ve geçerliliğin varlığı bir derece olarak ifade edilebilir, “var” ya da “yok” olarak ifade edilmez.

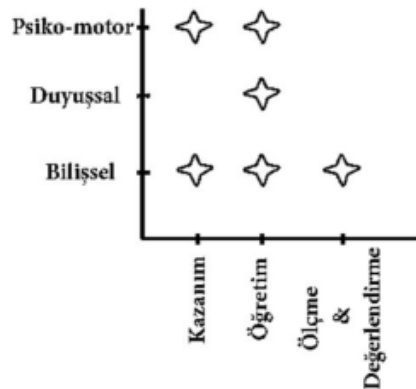
4. Eğitimde Kullanılan Ölçme Araçları

Öğretim yöntem ve teknikleri gibi ölçme ve değerlendirmede başvurulan yöntem ve teknikler de çeşitlenme gösterir. Öyle ki tıp eğitimi, beden eğitimi ve müzik eğitimi gibi birbirinden oldukça farklı alanların, kullandığı ölçme ve değerlendirme yöntemleri de farklıdır. Bazı alanlarda bilişsel alan davranışları ağırlıktayken, bir diğerinde duyuşsal veya psiko-motor alan davranışları önem taşıyabilir. Meslek liselerinde ve meslek yüksek okullarında, güzel sanatlar lisesi ve konservatuarda pek çok dersin uygulaması vardır ve bu derslerde psiko-motor beceriler ölçülür. Diğer yandan, istatistik, fizik ve matematik gibi alanlarda derslerin tamamına yakını bilişsel ve duyuşsal alan kazanımları ile sınırlıdır.

Becerilerin ağırlıklı olduğu her kazanımın bilişsel ve duyuşsal boyutu vardır. Bir makine mühendisi adayından bilişsel düzeyde, motor parçalarını ve işlevlerini bilmesi beklenir. Bunun yanı sıra motoru parçalarına ayırıp takmak psiko-motor düzeyde bir öğrenmedir. Enjeksiyon uygulaması hemşirelerin birinci sınıf konusudur. Basit bir beceri olmakla birlikte, öğrencinin enjeksiyon yaparken dikkatli olmasının önemini öğrenmesi, tedbirli olması ve her türlü önlemi alması gerektiği konusunda farkındalık kazanması duyuşsal düzeyde öğrenmelerdir.

Etkili bir öğretimin amacı, bireyi kısa bir süre sonra unutacağı bilgilerle donatmak değil, duyuşsal anlamda bireyin gelişimine katkıda bulunmak ve uygulama gerektiren konularda yaparak, dokunarak öğrenmesi için bireye fırsat tanımak olmalıdır. Hemşirelik eğitiminde temel hedeflerden biri, öğrencilerin hasta bakımında hedeflenen sonuçlara ulaşabilecek bilişsel, duyuşsal ve psiko-motor becerileri geliştirebilmesidir. Bu hedefin gerçekleştirilebilmesi için, öğrencilerin bakım uygulamalarına ilişkin bilgiyi öğrenmeleri ve bu bilgileri hasta bakım uygulamalarına aktarmaları gerekmektedir. Bu doğrultuda eğitimciler bilişsel süreçleri birleştiren ders konu içeriği, sınav soruları ve öğretim planları geliştirerek, öğrencilere bilgi ve beceri aktarımını kolaylaştırmalıdır. Bu nedenle, birçok alanda olduğu gibi hemşirelikte de eğitim hedeflerinin belirlenmesi, öğretim planlarının oluşturulması ile ölçme ve değerlendirmeye yönelik sınav sorularının hazırlanmasında en yaygın kullanılan modellerden biri Bloom Taksonomisi'dir.

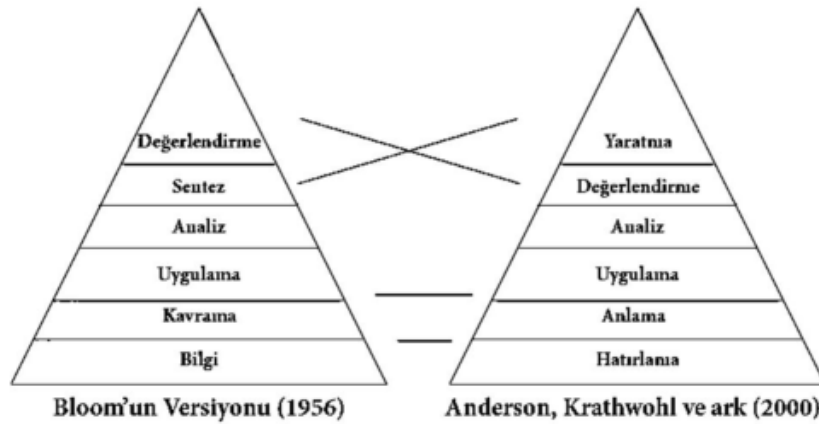
Bloom Taksonomisi, tüm bu hedeflerin gerçekleştirilmesinde eğitimcilere rehberlik etmektedir. Şekil 9'da kazanım ifadesi olarak bilişsel ve psiko-motor alanda, öğretimde bilişsel, psiko-motor ve duyuşsal alanda faaliyetlere başvuru ve ölçümlerin bilişsel alanla sınırlandırıldığı bir öğretim etkinliği görülmektedir. Öğrenim alanları birbiriyle ilişkili olmakla beraber ayrı ayrı ele alınmaları öğretimin planlanmasını, uygulanmasını ve değerlendirilmesini kolaylaştırır.



Şekil 9. Öğrenim alanları arasındaki geçişlilik

4.1. Bilişsel Alan Davranışları

Bilişsel alan sınıflaması **Bloom'un Öğrenme Taksonomisi** olarak adlandırılır. Bilişsel alan öğrenmeleri zihinsel etkinliklere dayanır. Matematik, fizik ve mantık gibi akıl yürütmeye dayalı alanlarda bilişsel alan öğrenmeleri ağırlıklıdır. Benjamin Bloom ve bir grup araştırmacı öğrenmeleri bilişsel, duyuşsal ve psiko-motor alan olmak üzere üç gruba sınıflamıştır. Bu alanlar kendi içinde hiyerarşik bir yapıda birbirinin ön koşulu olacak şekilde gruplanır. 2000 yılında Bloom'un öğrencilerinden Lorin Anderson başkanlığında bir grup araştırmacı, Bloom'un taksonomisini gözden geçirerek bazı değişiklikler yapmışlardır. Şekil 10'da, Tablo 5 ve 6'da Bloom'un Bilişsel Alan Taksonomisi'nin revize edilmiş yeni hali verilmiştir. Revize edilen Bloom taksonomisi, bilgi boyutu ve bilişsel olarak yeniden düzenlenmiştir. Bilgi boyutunda olgusal bilgi, kavramsal bilgi, işlemsel bilgi ve üstbilişsel bilgi basamakları yer alırken, bilişsel süreç boyutunda hatırlama, anlama, uygulama, analiz etme, değerlendirme ve yaratma basamakları bulunmaktadır. Yeni versiyonda "Bilgi" basamağının yerine "Hatırlama"; "Kavrama"nın yerini "Anlama" basamağı almıştır. Anlama basamağında öğrenciden anladığını açıklaması yönünde davranışlar sergilemesi beklenir. Yaratma olarak gördüğümüz son basamakta ise birtakım ürünlerle öğrencinin yaratıcılığını ortaya koyması hedeflenir.



Şekil 10. Bloom'un taksonomisi ve gözden geçirilmiş versiyonu

Yeni taksonomide bilgi boyutu taksonominin "ad" kısmını, bilişsel süreç boyutu ise "fiil" kısmını karşılamaktadır. Ad kısmında öğrencilerin öğrenmeleri gereken bilgi belirtilirken, fiil boyutunda bu bilgiyi öğrenirken kullandıkları bilişsel süreç karşılanmaktadır.

Tablo 5. Yenilenmiş Bloom taksonomisi

Bilgi Boyutu	Bilişsel Süreç Boyutu					
	Hatırlama	Anlama	Uygulama	Analiz	Değerlendirme	Yaratma
A. Olgusal Bilgi						
B. Kavramsal Bilgi						
C. İşlemsel Bilgi						
D. Üstbilişsel Bilgi						

Tablo 6. Yenilenmiş Bloom taksonomisindeki bilgi boyutunun yapısı

Olgusal Bilgi	• Terimlerin bilgisi • Özel detay ve öğeler bilgisi
Kavramsal Bilgi	• Sınıflamalar bilgisi • İlke ve genellemeler bilgisi • Teori, model ve yapılar bilgisi
İşlemsel bilgi	• Konuya özel beceri ve algoritmaların bilgisi • Konuya özel yöntem ve tekniklerin bilgisi • Ölçüt bilgisi
Üstbilişsel bilgi	• Stratejik bilgi • Bilişsel görevler bilgisi • Bireyin kendine dönük biliş ve öğrenmeyle ilgili bilgisi

Olgusal Bilgi: Olgusal bilgi en genel anlamıyla “temel öğelerin” bilgisidir. Kişi olgusal bilgiler sayesinde bir disiplini tanıyabilir ve bir problemi çözebilir. Kişi ilgili olduğu alanla ilgili okurken ve alanları ile ilgili bilgileri anlamaya çalışırken temel terimleri anlamak için olgusal bilgileri kullanır. Örneğin, öğrencinin kurallı cümle yapısı anlayabilmesi için özne, yüklem gibi kavramları bilmesi gerekir. Terimler bilgisi ve özel detay bilgisi olmak üzere iki alt kategorisi mevcuttur.

Kavramsal Bilgi: Bir alanla ilgili gruplar, sınıflamalar ve bu grup ve sınıflamaların arasındaki ilişkiyi içeren bilgidir. Olgusal bilgiye göre daha karmaşıktır. Kavramsal bilginin sınıflamalar bilgisi, ilkeler ve genellemeler bilgisi, teori, model ve yapıların bilgisi olmak üzere üç alt kategorisi bulunmaktadır.

İşlemsel Bilgi: Prosedür bilgisi olarak da tanımlanmaktadır. Bir bilginin nasıl kullanılacağını, bir işin nasıl yapılacağını bilgisidir. Olgusal ve kavramsal bilgide “ne” sorusu ön plana çıkarken, işlemsel bilgide “nasıl” sorusunun cevabı aranmaktadır. Olgusal ve kavramsal bilgide daha çok ürünler ön plandayken, işlemsel bilgide süreç ön plana çıkmaktadır. İşlemsel bilgi daha çok teknik, algoritma becerisi ile ilişkili kullanılmaktadır. Konuya/alana özel beceri ve algoritmaların bilgisi, konuya/alana özel yöntem ve tekniklerin bilgisi ve ölçütlerin bilgisi olmak üzere üç alt kategorisi bulunmaktadır.

Üst Bilişsel Bilgi: Kişinin kendi bilişi hakkındaki bilgisidir. Kişinin bilişsel olarak kullandığı stratejilerin bilgisi, bu stratejileri ne zaman ve nasıl kullanacağını bilgisi üst bilişsel bilgi içinde yer almaktadır. Orijinal taksonomide olmayan bu bilgiye zamanla öğrenme ve gelişme ile ilgili çeşitli kuramların ön plana çıkması ile ihtiyaç duyulmuştur. Stratejik bilgi, bilişsel görevlerin bilgisi ve bireyin kendine dönük bilgisi olmak üzere üç alt kategorisi mevcuttur.

Hatırlama: Bilişsel alanın en temelinde yer alan bu basamak hatırlama düzeyinde bilgileri içerir. Hatırlamada öğrenciden bir kavramı tanımlaması, başka dilde karşılığının ne olduğunu bilmesi, gösterilen bir nesnenin ne olduğunu söylemesi, tanımı verilen bir kavramın adını söylemesi ve resmini çizmesi gibi davranışlar sergilemesi beklenir. Bir konuda hatırlamanın oluşması için öğrenmenin gerçekleşmiş olması şarttır. Burada hatırlama bireyin o ana dek öğrendiklerini hafızasından geri çağırması anlamında kullanılmıştır.

Anlama: Hatırlama düzeyinde öğrenmelerden sonra gelen bu basamakta, öğrenciden verilen konuda açıklama yapması, bir konuyu kendi sözcükleriyle ifade etmesi ve örnekler vermesi beklenir.

Uygulama: Bu basamakta önceki iki basamakta öğrenciden öğrendiği bilgileri yeni bir duruma uygulaması, verilenleri kullanarak problem çözmesi ve bir çemberin çevresini hesaplamak gibi bilişsel bir kazanımı doğru şekilde gerçekleştirmesi beklenir.

Analiz: Analiz basamağında bütünü parçalarına ayırma becerisini göstermek esastır. Bu basamakta öğrenci problemlerin nedenlerini bulur, bir grafiği yorumlar veya bir grup maddeyi kendi içinde bütünlüğünü koruyacak şekilde tekrar sınıflar.

Değerlendirme: Bu aşamada öğrenciden bir ürün ya da materyalin kalitesini, standartlara uygunluğunu belli ölçütlere göre sorgulaması, yargıda bulunması, bir konuda temel prensipleri dikkate alarak olayları değerlendirerek tespitlerde bulunması beklenir.

Yaratma: Yaratma basamağında daha önceki basamaklardaki kazanımları edinmiş olan öğrenci, yeni öğrenmelerin önceki öğrenmelerle bağlantısını kurar, elindekileri kullanarak yeni bir ürün ortaya çıkarır. Örneğin; kompozisyon yazar, resim çizer, beste yapar veya bir şarkıyı kendi tarzında yorumlar.

Tablo 7. Bilişsel alan sınıflamasına göre öğrenme çıktılarında yer alabilecek fiiller

Hatırlama 	Anlama 	Uygulama 	Çözümleme 	Değerlendirme 	Yaratma 
sıralamak, listelemek, tanımlamak, tarif etmek, tekrarlamak, sunmak, söylemek, ana hatlarıyla belirtmek, sınıflandırmak, anlatmak, belirtmek, ifade etmek, saymak, eşleştirmek, adlandırmak	ilişkilendirmek, değiştirmek, ayırt etmek, tartışmak, tahmin etmek, sınıflandırmak, açıklık getirmek, kestirmek, izah etmek, genelleştirmek, yorumlamak, örnekle açıklamak, çıkarımda bulunmak, özetlemek	uygulamak, hesaplamak, değiştirmek, tercih etmek, karar vermek, sergilemek, keşfetmek, çalıştırmak, kullanmak, göstermek, uyarlamak, hazırlamak	parçalara ayırmak, şekil çizmek, ayırım yapmak, göstermek, çıkarımda bulunmak, işaret etmek, çözümlemek, test etmek, eleştirmek, sorgulamak, çıkarsamak	karşılaştırma, sonuca ulaşmak, eşleştirmek, değerlendirmek, yargılamak, hüküm vermek, değer biçmek, düzeltmek, eleştirmek, ispatlamak, doğrulamak, ölçmek, geçerliğini denetlemek, inandırmak	birleştirmek, planlamak, üretmek, değiştirmek, yeniden düzenlemek, yaratmak, tasarlamak, formüle etmek, inşa etmek, organize etmek, geliştirmek, ilişkilendirmek, yazmak

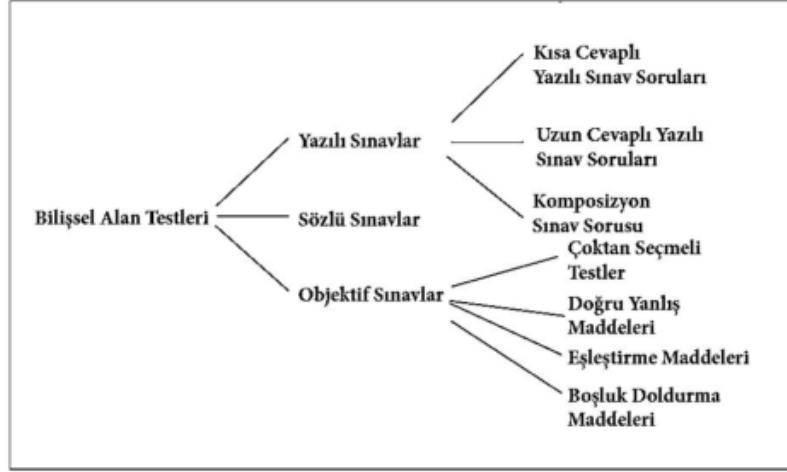
Tablo 8. Bilişsel alan sınıflamasına göre soru kökleri ve ölçme türleri

Hatırlama 	Anlama 	Uygulama 	Çözümleme 	Değerlendirme 	Yaratma 
Kim Ne Ne zaman Nerede Hangisi	Kaç/ Ne kadar Hangi sırada Nasıl Ne açıdan Ne tür farklar var Açıklayınız Sonuç ne olurdu	Nasıl çözersiniz Nasıl kullanılır ...yoluyla gösterin ...deneyini yapınız Neler kullanılır Hangi yöntemi/yaklaşımı kullanırsınız	Neden/Niçin Aralarında ne tür ilişkiler var Nasıl sınıflandırılabilir Çizelgesini oluşturunuz Hangi parçalardan oluşur	Kime/neye göre Neden doğru/yanlış Siz olsaydınız... Ölçütlerimiz ne olmalı Geçerliği nedir Nasıl ispatlarız Niçin böyle	Üretebilir misiniz Nasıl bir plan uygulayabiliriz Tasarım oluşturun Kısa film çekin Web sayfası hazırlayın Makale yazın Resim çizin
Çoktan Seçmeli Boşluk Doldurma Doğru-Yanlış Kısa- cevaplı sorular	Çoktan Seçmeli Boşluk Doldurma Doğru-Yanlış	Açık uçlu sorular Performans görevleri Rapor yazdırma Portfolyo	Açık uçlu sorular Performans görevleri Rapor yazdırma Portfolyo	Açık uçlu sorular Performans görevleri Rapor yazdırma Portfolyo	Açık uçlu sorular Performans görevleri Rapor yazdırma Portfolyo

Tablo 9. Bilişsel alan sınıflamasına göre soru örnekleri

Hatırlama	• Ölçme nedir? • Proper kelimesinin İngilizce karşılığı nedir? • Aritmetik ortalama, standart sapma, varyans ve korelasyonun Yunan alfabesindeki sembollerini yazınız? • Aşağıdaki şekillerden hangisi kolaylaştırılmış difüzyondur? • Duyuşsal alanın ilk basamağında basamağı yer alır. • Hipertansiyon nedir? • Hipertansiyonun belirtileri nelerdir?
Anlama	• Deneme formu sonuçlarına göre sınıfın aritmetik ortalamasının oldukça yüksek olduğunu gördünüz. Maddeleri düzenlemek için neler yaparsınız? • Hipertansiyonun teşhisinde nelere dikkat edilmelidir? • Ölçme ve değerlendirme arasındaki farkı beş cümleyi geçmeyecek şekilde açıklayınız.
Uygulama	• Maksimum olası puanın 20 olduğu bir sınavın aritmetik ortalaması 18, standart sapması 2 olarak bulunmuştur. Buna göre puanların dağılımı hakkında ne söylenebilir? • Hastanın bilincini kaybetmesi durumunda ne yaparsınız?
Analiz	• Eğitim-öğretimde testlerin hangi amaçlarla kullanıldığını örnekler vererek açıklayınız. • Fotosentezin işlevlerini açıklayınız. • Hava kirliliğinin nedenleri neler olabilir açıklayınız. • Çevre kirliliği sorununun kaynakları sizce nelerdir? Kavram haritası oluşturarak gösteriniz. • Hipertansiyonun belirtileri nelerdir? • Aşağıdakilerden hangisi birinci dünya savaşının çıkma nedenleri arasında sayılamaz?
Değerlendirme	• Aşağıda sunulan araştırmanın metodoloji bölümünü araştırma yöntem ve ilkelerine uygunlukları bakımından değerlendiriniz. • Şikayetlerle gelen hastaya tetkikler yapılmış, ... tanısı konularak tedavisi uygulanmıştır. Tanıyı nasıl doğrularsınız?
Yaratma	• Kendi alanınızda bir test sınavı hazırlayınız. • Alternatif ölçme ve değerlendirme yöntemlerinin kullanıldığı üç öğrenme senaryosu geliştiriniz. • Çok su içme, sık idrara çıkma, ağız kuruluğu belirtileri ile seyreden hastalık ne olabilir? Teşhis için neler yapılmalıdır? Neler önerirsiniz?

Bilişsel alan öğrenmelerini ölçmede tercih edilen geleneksel ölçme ve değerlendirme yöntemleri; yazılı sınavlar, sözlüler ve objektif testler olmak üzere üç başlıkta ele alınabilir. Bunlar da kendi içinde farklı başlıklarda incelenir (Şekil 11).



Şekil 11. Geleneksel yaklaşımda tercih edilen bilişsel alanı ölçme araçları

4.2. Yazılı (Açık Uçlu) Sınavlar

Soruların genellikle yazılı olarak verildiği, cevapları öğrencilerin düşünerek yazılı olarak verdikleri sınav türüdür. Cevabı öğrencinin oluşturması en temel özelliği olan açık uçlu sınavlarda cevabı **kısa/sınırlandırılmış** ve cevabı **uzun/kapsamlı** olan sorular olmak üzere iki tip soru yöneltmek mümkündür. Cevabı kapsamlı sorularda öğrencinin bir konuda kapsamlı olarak açıklama, analiz ya da yorum yapması istenir. Cevabı sınırlandırılmış sorularda ise öğrenciden belli bir madde sayısı ya da kelime ile sınırlandırılarak bir konuda en etkili olan faktör veya özellikleri ana hatları ile yazması beklenir. Yazılı sınavların özellikleri Tablo 10'da özetlenmiştir.

Tablo 10. Açık uçlu sınavların olumlu ve olumsuz özellikleri

Olumlu özellikleri	Olumsuz özellikleri
- Hazırlanması kolaydır. - Öğrenci cevaplama özgürlüğüne sahiptir. - Eğitici yorum sorabilir, öğrenci yorum katabilir. - Detaylı açıklama gerektiren konulardan soru sorulabilir. - Konunun gerçekten öğrenilip öğrenilmediği ve öğrenmenin derinliği hakkında bilgi verir. - Üst düzey öğrenmelerin ölçülmesinde etkilidir. - Cevabı seçeneklerden seçmek söz konusu olmadığından, şans başarısı faktörü yoktur.	- Değerlendirmesi zaman alıcıdır. - Değerlendirme subjektifliğe açıktır. - Puanlama güvenilirliği düşüktür. - Soru sayısının sınırlılığından dolayı kapsam geçerliliği düşüktür.

Açık uçlu sınav sorusu hazırlarken, konunun önem derecesi ve programdaki yoğunluğu dikkate alınmalıdır. Cevaplama süresinin yeterli olduğundan emin olunmalı; cevabı düşünmek ve yazmak için ortalama bir öğrencinin ne kadar zaman harcayacağı kestirilmelidir. Soruların ifade edilişi açık ve net olmalı, kullanılan dil ve anlatım öğrencilerin yaş seviyesine uygun olmalı, öğrencinin ne sorulduğunu anlamasını güçleştiren kafa karıştırıcı ifadelerden kaçınılmalıdır. İyi bir açık uçlu sınav sorusunda okuyucu ne sorulduğunu kısa sürede anlamalı, anlam düşüklüğü veya karmaşıklıktan dolayı gereksiz yere düşünerek vakit kaybetmemelidir. Açık uçlu sınavların en olumsuz yönü puanlama güvenilirliğinin düşük olmasıdır. Bu nedenle puanlamaya başlamadan önce cevap anahtarı ve nereye ne kadar puan verileceğinin gösterildiği puanlama yönergesi hazır bulundurulmalıdır.

Açık uçlu maddelerin hazırlanma sürecinde en önemli kısımlardan biri şüphesiz madde kökünün yazılmasıdır. Açıkça tanımlanmış bir madde (madde kökü), bir yönerge fiilinden ve o fiilin nesnesinden oluşur. Madde kökünün oluşturulması sürecinde bazı hususlara dikkat edilmesi gerekmektedir. **Madde kökündeki öğrencilerden istenen eylem (fiil) ile hedef öğrenme çıktısındaki fiil aynı olmalıdır.** Burada öncelikli amaç, hedef öğrenme çıktısındaki (kazanımlar) bilişsel düzey ile madde kökünde yer alan içeriğin bilişsel düzey açısından aynı olmasıdır. Diğer amaç ise hedef öğrenme çıktısındaki fiil ile madde kökünde yer alan fiilin aynı anlamda veya eş/yakın anlamda olmasını sağlamaktır. Tablo 11'deki örnekte görüldüğü üzere, öğrenciden istenen çıktı metinden hareketle sorunu çözmek amacıyla çözüm önerisini oluşturmaktır. Bu amaç doğrultusunda (kazanım) soruda öğrenciden çözüm önerileri istenmektedir. Buradaki temel amaç, hem kazanım hem maddede öğrenciden istenen zihinsel aktivitenin aynı olmasıdır.

Tablo 11. Kazanım ve madde kökündeki fiilin aynı olduğu açık uçlu soru örneği

Ders: Türkçe (8. Sınıf)
Hedeflenen Öğrenme Çıktısı: Metinde ele alınan probleme farklı çözümler üretir.
Lokman ve Hizmetçiler
Lokman efendisinin yanındayken diğer çalışanlar tarafından sürekli kısıklanılmaktaydı. Bir gün efendisi hizmetçilerini bağa meyveleri toplamak amacıyla gönderdi. Dönüşte hizmetçiler topladıkları meyveleri yediler ve efendilerine de "Bütün meyveleri Lokman yedi" dediler.
Efendileri, Lokmana kızdı. Lokman, öncelikle efendisinin kendisine neden kızdığını bilmiyordu. Efendisi sözünü tamamladıktan sonra Lokman;
"Efendim meyveleri ben yemedim. Sorumlu kişiyi bulmak için bir önerim var. İzniniz olursa söylemek isterim."
Efendisi "söyle bakalım dedi"
Lokman sözüne devam etti: "Hepimize öncelikle sıcak su içirin. Sonra siz atla, bizler de yaya olarak koşalım. Böylece meyveleri kimin yediği ortaya çıkacaktır." Efendisi Lokman'ın önerisini kabul etti ve yerine getirdi.
Lokman ile birlikte diğer hizmetçiler koşmaya başladıktan sonra meyveleri yiyen herkes yarı yolda yediklerini istifra etmeye başladılar. Böylece kimin meyveleri yediği ve yalan söylediği ortaya çıkmış oldu.
Soru 1. Siz Lokman'ın yerinde olsaydınız, diğer hizmetçilerinin söyledikleri karşısında <u>çözüm öneriniz</u> ne olurdu? Açıklayınız.

Madde oluşturulurken öğrencinin sadece bir eyleme yönelik zihinsel bir sürece girmesi, odaklanması gerekir. Cümlelerin birden fazla fiil içermesi ya da sıralı cümlelerden oluşması öğrencinin de birden farklı noktaya odaklanması anlamına gelir. Öğretim programlarındaki kazanımlar, içerik olarak öğrencilerden bir öğrenme çıktısına yönelik eylemleri içermesini gerekmektedir. Bazı derslere ait öğretim programlarında ise kazanımların içerik olarak geniş ve kapsamlı olması ve birden fazla eylemi içerdiği görülmektedir. Bir diğer ifadeyle birden fazla öğrenme çıktısı söz konusudur. Bu ise öğrenciden istenen çıktının belirsizleşmesi, daha karmaşık olması anlamına gelmektedir. Dolayısıyla hem hedeflenen öğretim çıktısının öğretiminde hem de ölçmek amacıyla madde oluşturmada problem oluşturabilmektedir. Bu gibi durumlarda ise öğrenciden hangi zihinsel eyleme yönelik bir iş istendiği netleştirilmelidir.

Açık uçlu maddenin problem durumu (bilgi) var ise madde kökü ile bağlantılı olması gerekmektedir. Bazı durumlarda problem durumu ya hiç olmamakta ya da verilen metinde problem verilmemekte, herhangi bir nesne veya kavrama ilişkin betimleyici açıklamalar verilmektedir. Eğer madde, problem durumu okunmadan veya bağlantı kurulmadan yanıtlanabiliyorsa bu verilen açıklamaların (metin) gereksiz olduğunu göstermektedir.

Açık uçlu maddelerin diğer maddelerde olduğu gibi çok genel olmaması gerekir. Çok genel olması, öğrencinin istenilen amaç doğrultusunda performans göstermesini engelleyecektir. Ancak açık uçlu maddeyi çok fazla sınırlandırmamak da gereklidir. Sınırlandırılmamış bir şekilde hazırlamak, öğrencilerin yanıtlarını oluştururken kendi sınırlarını belirleyememesine neden olabilmektedir. Bu ise öğrencilerin, istenen dışında çıkmasına veya görevin sadece bir kısmına odaklanmasıyla sonuçlanabilir. Bu durumda öğrenciden öğrenciye maddeye ilişkin farklı yanıtların gelme olasılığı artmaktadır. Öğrencilerin kendi koydukları sınırlardan hareketle oluşturdukları yanıtları farklılaşacağı için yanıt anahtarı oluşturmada öğreticinin zorlanmasına neden olacaktır. Ayrıca öğrenci yanıtlarının geçerliliği ve güvenilirliği düşük olacaktır.

4.2.1. Puanlama Anahtarının Oluşturulması

Hazırlanan açık uçlu maddelerin uygun bir şekilde puanlanarak değerlendirilmesi gerekmektedir. Puanlama için dereceli puanlama anahtarları; hem objektif olarak değerlendirme, hem de uygulanan öğretimin ne derece başarılı olduğu, öğrencilerin kazanımları ne derece kazandıklarına yönelik ayrıntılı dönüt vermek amacıyla kullanılabilir. Dereceli puanlama anahtarı, öğrenci yanıtlarına ilişkin gözlemleri puanlanmış kategoriler halinde kaydetmeye yardımcı olur. Bu yolla öğrencilerin yanıtlara dağılımları belirlenmekte ve öğrenmelerle ilgili durum net olarak görülebilmektedir.

Açık uçlu maddeler, diğer madde türlerine göre genel olarak yazı güzelliği, yazma hızı, ifade gücü gibi değişkenlerin puanlama sürecine karışacağı, yanıtın gereksiz açıklamalarla daha avantajlı olacağı ve puanlamanın daha öznel olduğu yönünde eleştiriler almaktadır. Bu eleştiriler, iyi bir puanlama anahtarı hazırlanarak ortadan kaldırılabilir ve objektif olarak öğrenme çıktıları puanlanabilir. Puanlama anahtarı ayrıntılı bilgi sunması nedeniyle de öğretim süreçlerine yönelik daha ayrıntılı dönütler sunabilmektedir. Bu nedenle puanlama anahtarları amaç doğrultusunda hem içerik hem de teknik olarak doğru bir şekilde hazırlanması gerekir.

Açık uçlu maddelerin puanlanması için öncelikle **analitik** veya **bütünsel (holistik)** puanlama anahtarlarından hangisinin daha uygun olacağına karar verilmelidir. Öğrencilerin akademik başarılarının belirlenmesi amacıyla okullarda uygulanan sınavlar ile ulusal ve uluslararası sınavlardaki açık uçlu maddelerin puanlanmasında daha çok bütünsel puanlama anahtarı tercih edildiği söylenebilir. Geri bildirimin daha çok önem arz ettiği ve karmaşık yapıların ölçülmeye çalışıldığı maddelerde dikkate alınacak ölçüt sayısı fazla olduğundan analitik puanlama anahtarları tercih edilebilir.

Bütünsel dereceli puanlama anahtarları öğrencilerin her bir görev veya maddede sergiledikleri performansın bir bütün olarak, alt bileşenlerine ayrılmadan puanlanmasında kullanılan bir puanlama anahtarıdır. Bütünsel dereceli puanlama anahtarlarının kullanıldığı puanlamada, her bir görev veya maddenin tamamında gösterilen başarı yüksek ise bazı bileşenlerindeki noksanlar göz ardı edilebilir. Bütünsel dereceli puanlama anahtarlarıyla puanlamada, puanlayıcı her bir görev ya da yanıtı bütün olarak inceler ve bütüne ilişkin bir puan verir. Bu yöntem, puanlayıcının daha subjektif olmasına sebep olabilir ancak gereksiz ayrıntılara takılmadan puanlama yapmasına imkân verir. Bu puanlama yaklaşımı, bütünün parçalarından daha büyük önem taşıdığı durumlarda kullanılmalıdır.

Bütünsel Dereceli Puanlama Anahtarı Örneği

Puanlar	Tanımlar
5 puan	Problemin tümüyle anlaşılmış olduğu gösterilmiştir. Göreve ya da soruya ilişkin tüm gereklilik cevapta yer almıştır.
4 puan	Problemin oldukça anlaşılmış olduğu gösterilmiştir. Göreve ya da soruya ilişkin tüm gereklilik cevapta yer almıştır.
3 puan	Problemin kısmen anlaşılmış olduğu gösterilmiştir. Göreve ya da soruya ilişkin gereklilik cevapta yer almıştır.
2 puan	Problemin çok az bir kısmının anlaşılmış olduğu gösterilmiştir. Göreve ya da soruya ilişkin çoğu gereklilik cevapta yer almamıştır.
1 puan	Problemin anlaşılmamış olduğu gösterilmiştir.
0 puan	Cevap verilmiştir/ görev hiçbir şekilde gerçekleştirilmemiştir.

Analitik dereceli puanlama anahtarlarında, bütünsel olanların tersine, öğrencinin her bir görev veya soruda göstereceği performans, öncelikle alt bileşenlerine ayrılarak puanlanır ve bu puan, alt bileşenlerden alınan puanların toplamıdır. Bir görev veya cevapta, bu bileşenler boyutunda öğrencinin güçlü ve zayıf yönleri belirlenebilir. Bu yöntemde, cevapların alt bileşenlerine ayrılması ve puanlanması ile puanlayıcının daha objektif olması sağlanmaktadır.

Analitik Dereceli Puanlama Anahtarı Örneği (Araştırma Becerisi)

ÖLÇÜTLER	PUAN		
	1	2	3
Kaynakların sayısı	Ulaşılan kaynaklar yetersiz.	Ulaşılan kaynaklar kısmen yeterli.	Ulaşılan kaynaklar yeterli.
Tarihsel Doğruluk	Çok fazla yanlış var.	Çok az yanlış var.	Açık bir yanlış yok.
Organizasyon	Bilgilerin düzenlenmesi, akıcı ve etkili değil.	Bilgilerin düzenlenmesi, kısmen akıcı ve etkili.	Bilgilerin düzenlenmesi, yeterince akıcı ve etkili.
Bibliyografya	Kaynakların çok azı etkili kullanılmış.	Kaynakların çoğu etkili kullanılmış.	Tüm kaynaklar, etkili kullanılmış.

Açık uçlu maddenin puanlanması için puanlama anahtarı türünün seçiminden sonra ise **genel** mi yoksa **maddeye özgü** bir puanlama anahtarının hazırlanacağına karar verilmelidir. Özellikle ulusal ve uluslararası uygulanan sınavlardaki açık uçlu maddelerin puanlanması için bütünsel puanlama anahtarlarının maddeye özgü olanları yaygın olarak kullanıldığı görülmektedir. Bu anahtarlar, her bir madde için özgün hazırlanan puanlama araçlarıdır.

Bir maddenin puanlanması için bütünsel puanlama anahtarı türü tercih edildikten sonra her bir madde için özgü hazırlanması, hem maddenin objektif olarak puanlanması hem de eğitim öğretim sürecinde dönütlerin daha etkili verilmesi açısından önemlidir. Genel ifadeler olmayıp doğrudan maddenin yanıtına ilişkin bilgilerin yer alması ve net açıklamaların olması nedeniyle objektifliği arttırmaktadır. Ayrıca konuya özgü bilgilerin puanlama anahtarlarında yer alması öğretim sürecindeki eksik öğrenmelerin tamamlanmasına, kavram yanlışlarının belirlenmesine, öğrencilerin düşünme stratejilerinin ve biçimlerinin ortaya çıkarılması gibi birçok alanda katkı sağlamaktadır.

Açık uçlu sınav sorusu hazırlarken dikkat edilmesi gerekenler:

- Bloom'un bilişsel alan taksonomisinde "Hatırlama" düzeyinde soru yazmak daha kolay olduğundan çoğu zaman sınavda bu seviyedeki sorular ağırlıklı olabilir. Ancak eğitimciler mümkün olduğunca Anlama, Uygulama, Analiz, Değerlendirme ve Yaratma düzeyinde sorular yazmaya gayret etmelidir.
- Spesifik olarak öğrenciden ne istendiği, soruya ve sorunun alt maddelerine kaçır puan verileceği yazılı kağıdında belirtilmeli, sınavdan sonra kağıtların durumuna göre puanlama takdiri yapılmamalıdır.
- Sınav süresi ortalama düzeyde bir öğrencinin sınavı tamamlamak için ne kadar zamana ihtiyacı olacağı göz önünde bulundurularak belirlenmelidir.
- Sınavlar test veya yazılı sınav olsun, süreç değerlendirme olarak ele alınmalı ve öğrenciye bildiklerini stres altında kalmadan ifade etmeleri için gereken ek süre tanınmalıdır. Ancak, eklenen sürenin kopyaya ortam yaratacak kadar uzun olmamasına dikkat edilmelidir.
- Yazılı sınavlar performans görevleri ve projelerle desteklenerek öğrencinin bu sınavlarla ilgili kaygıları hafifletilebilir.
- Sınav sonrası öğrencilerden sınavdan bekledikleri puanı kağıdın sonuna not almaları istenebilir. Doğru yapılırsa, böyle bir uygulama öğrencilerin başarılarına ilişkin gerçekçi bir algı geliştirmelerine yardımcı olacaktır.

Puanlama anahtarı hazırlarken dikkat edilmesi gerekenler:

- Açık uçlu maddenin puanlanması için en uygun puanlama anahtarının tür olarak doğru seçilmesi gerekmektedir.
- Puanlama anahtarındaki performans düzeyleri "tam doğru", "kısmen doğru" ve "yanlış yanıtlar" gibi doğru ve eksiksiz olarak belirlenmelidir. Örneğin bir maddenin "tam doğru yanıtlar ve yanlış yanıtlar"dan mı yoksa "tam doğru, uzak doğru ve yanlış yanıtlar"dan mı oluşacağına karar verirken maddenin ölçmeye çalıştığı kapsam, içerik ve bilişsel süreçler dikkate alınmalıdır. Açık uçlu maddeler; ikili, üçlü ve dörtlü dereceli bir şekilde olmak üzere farklı düzeylerde tanımlanarak puanlanabilir. Dolayısıyla bu süreçte puanlama anahtarı için performans düzeyleri veya ölçüt sayısı açısından kararını verirken dikkatli olunması gerekmektedir. Her cevaba ve cevapların alt bölümlerine ait puanlamalar verilmelidir.
- Eğer kısmi doğru yanıtlar söz konusu ise, tam doğru yanıtlardan farklı olarak hangi öğelerin kapsamamış olduğundan açıkça ifade edilmelidir.

- Tam doğru, kısmen doğru ve yanlış yanıtlar kısmında yer alacak bütün olası yanıtlar yazılmalıdır. Olası tüm yanıtların yazılması için maddeyi oluşturan eğiticinin alan bilgisinin çok iyi ve puanlama anahtarı hazırlama konusunda deneyimli olması gerekmektedir. Yanıtlar oluşturulurken öğrenci gibi düşünerek olası bütün yanıtlar öğrencilerin bakış açısıyla ele alınmalıdır.
- Yanıtlar oluşturulurken açık ve anlaşılır olmasına dikkat edilmelidir.
- Açık uçlu maddenin bütün bölümlerinin, özellikle puanlama anahtarının etkili bir şekilde hazırlanabilmeleri için mümkünse küçük bir pilot uygulamanın yapılması önem arz etmektedir.

Puanlama anahtarının hazırlanışı:

- İyi hazırlanmış, derecelenmesi ve puanlaması yapılandırılmış cevap anahtarları hazırlanmalıdır.
- Yanıtların açık ve anlaşılır olmasına dikkat edilmelidir.
- Puanlama için “En doğru yanıtlar”, “Uzak doğru yanıtlar”, “Yanlış yanıtlar”, “Boş bırakılan yanıtlar” ve “Diğer (Anlamsız-İlişkisiz) yanıtlar” olarak beş boyutta tanımlanabilir. Bunu yapabilmek için öncelikle sorunun/maddenin olası doğru, yanlış ve diğer yanıtları saptanmalıdır.
- Doğru yanıtlar kendi içerisinde en doğru yanıttan, uzak doğru yanıtı doğru sıralanmalıdır.
- “Yanıtı tanıma kodları” oluşturulmalıdır. Bu kapsamda cevaplayanların yanıtlarını birbirinden ayırmak ve her bir yanıtın hangi koda karşılık geldiğini bulmak, değerlendirme açısından belirleyicidir.

Açık uçlu sınav okurken izlenmesi önerilen adımlar:

- Sınav öncesinde hazırlanan cevap anahtarı ve nereye kaç puan verileceğini gösteren puanlama yönergesi hazır bulundurulmalıdır. Bunlar eşliğinde sınav uygun bir ortamda okunmalıdır.
- Öncelikle sınav kağıtlarının tercihen beşte biri her soru için gözden geçirilerek sınıfın ilgili soruya cevapları incelenmelidir. Yanlılık etkisinden dolayı iyi bir cevaptan sonra okunan orta halli bir cevap, olduğundan daha kötü görünürken, birkaç zayıf kağıttan sonra okunan orta halli bir kağıt, olduğundan daha yüksek bir puanla ödüllendirilebilir. Bu nedenle yazılı kağıtlarını sınıf düzeyini kestirerek okumakta fayda vardır.
- Benzer şekilde kağıtlar tek tek okunduğunda cevabı iyi yazılmış bir sorudan sonra gelen eksik cevaplı bir soruya eğitici normalde vereceğinden daha fazla puan verebilir. Bu nedenledir ki öğrenciler iyi bildikleri soruları başta cevaplarlar.
- Öğrencinin ismini bilmek puanlarına olumlu veya olumsuz şekilde yansıyabileceğinden, sınav kağıtları öğrencilerin isimleri kapatılarak okunmalıdır.

- Kağıtların beşte biri puanlama yapmaksızın gözden geçirildikten sonra, analitik bir yaklaşımla her kağıdın aynı sorusu okunarak puan verilmeye başlanmalıdır.
- Soru temelli okuma işlemi her bir soru için aynı şekilde gerçekleştirilerek kağıtlar okunduktan sonra, sayfa toplamaları alınmalıdır.
- Son olarak sayfa toplamaları toplanarak öğrencinin puanı sınav kağıdının sağ üst köşesine toplam puan üzerinden not düşülmelidir. Örneğin; öğrencinin puan toplamı yüz üzerinden 70 ise 70/100 şeklinde ifade edilir.
- Sayfa toplamaları alınarak kağıt üzerine düşülen notların doğruluğu tekrar gözden geçirilmelidir.
- Soru temelli puanlamada öğrenciler arası temel farklılıklar not alınırsa, eğitici öğrencilerin hangi soruda daha çok sorun yaşadığı hakkında dönüt elde etmiş olacaktır.
- Sınav sonucu öğrencilere duyurulduktan sonra sorular sınıfta birlikte ele alınmalı ve cevaplar elektronik olarak kaydedilmişse öğrencilere yanlışlarını görme fırsatı sunulmalıdır. Bu şekilde öğrenciler hatalarını görme fırsatı yakalarlar.
- Puanlar not platformuna girilirken dikkatli olunmalı, hata yapmamaya özen gösterilmelidir.

Tablo 12. Doğru ve yanlış açık uçlu sınav sorusu örnekleri

Yanlış örnek	Doğru örnek
1) İyi bir test maddesinde bulunması gereken özellikler yazınız.	1) Test maddelerini değerlendirirken dikkat edilmesi gereken hususları aşağıdaki başlıkları temel alarak yazınız. Mantıksal ipuçları (5 puan), Dil ve anlatım (5 puan), Biçim ve düzen (5 puan).
2) Çevre kirliliğinin etkilerini açıklayınız.	2) Çevre kirliliğinin etkilerini bir kavram haritası üzerinde gösteriniz (İlgili her grup 5 puan).
3) Açık uçlu bir objektif test maddesi test maddesi yazınız. (15 puan)	3) Madde yazma kurallarına dikkat ederek bir adet boşluk doldurmalı test maddesi yazınız. Maddenizi oluştururken göz önünde bulundurduğunuz üç özelliği yazınız (Her bir özellik 3 ve maddenin bu özelliklere uygunluğu 2 puan).
4) Aşağıda bir grup öğrencinin sınav puanları verilmiştir. 40, 40, 55, 60, 60, 66, 70, 75, 75, 75, 75, 80, 80, 82, 82, 90, 90, 100, 100 Puanlara ait frekans dağılımını oluşturunuz. (10 puan)	4) Aşağıda bir grup öğrencinin sınav puanları verilmiştir. 5 gruplu bir frekans dağılımı oluşturarak puanlara ait dizi genişliğini, tepe değeri ve tepe değerin altında puan alan öğrencilerin yüzdesini hesaplayınız (her madde 2 puan, frekans dağılımı 4 puan). 40,40,55,60,60,66,70,75,75,75,75,80,80,82,82,90,90,100,100

ÖRNEK PUANLAMA ANAHTARI-1 (DERECELİ-GENEL)

Dereceli Puanlama Anahtarı (Genel)

Puan	Açıklama
2	Yanıt, maddenin gerektirdiği bilimsel içeriğin, kavramların ve prosedürlerin tam olarak anlaşıldığını gösterir. Yanıt, maddenin gerektirdiği şekilde açık, eksiksiz ve doğrudur. Yanıt, tam bir anlayış sergilemekten vazgeçmeyen çalışma veya açıklamadaki küçük bir kusur veya eksiklik içerebilir.
1	Yanıt, maddenin gerektirdiği bilimsel içerik, kavramlar ve prosedürlerin kısmen anlaşıldığını gösterir. Gösterilen ve/veya açıklanan gerekli bilimsel içeriğin, kavramların ve/veya prosedürlerin kısmen anlaşılmasıyla yanıt biraz doğrudur. Yanıt, eksik veya belirsiz bazı çalışmalar içerebilir.
0	Yanıt, o sınıf düzeyindeki maddenin gerektirdiği şekilde bilimsel içeriğin, kavramların ve prosedürlerin anlaşıldığını göstermek için yetersiz kanıt sağlıyor. Yanıt, yalnızca maddeden kopyalanan veya yeniden ifade edilen bilgileri veya 1 puan almak için yetersiz doğru bilgileri gösterebilir.

Sıfır içindeki özel kategoriler ayrıca raporlanır:

Kategoriler	Açıklamalar
Boş	Boş, tamamen silinmiş, tamamen üstü çizili veya tamamen boşluktan oluşuyor
Reddetme	Göreve yanıt vermeyi reddetme
Görev Dışı	Öğeye atıfta bulunmaz ancak kasıtlı değildir/ reddetme
Yabancı Dil	Tamamen İngilizce dışında bir dilde yazılmış
Okunmaz	Okunaksız veya tutarsız

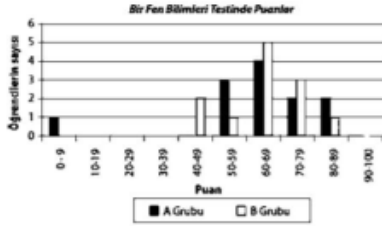
ÖRNEK PUANLAMA ANAHTARI-2 (BÜTÜNSEL-GENEL)

Bütünsel Dereceli Puanlama Anahtarı (Genel)

Puan	Açıklama
3 puan	- Verilen problem durumunu tam olarak anlamıştır. - Öğrenci; sıvıların bir dirence sahip olduğunu ve sıvının cinsine göre değiştiğini bilir. Ayrıca bir cismin farklı sıvı içindeki hareketinin de sıvı direncinden dolayı farklı olduğunu bilir. Sıvı direncinin bir kuvvet olduğunu farkındadır.
2 puan	- Verilen problem durumunu anlamıştır. - Öğrenci; sıvıların bir dirence sahip olduğunu ve sıvının cinsine göre değiştiğini bilir. Ayrıca bir cismin farklı sıvı içindeki hareketinin de sıvı direncinden dolayı farklı olduğunu bilir.
1 puan	- Verilen problem durumunu kısmen anlamıştır. - Öğrenci; sıvı direncini kavramsal olarak bilmemektedir. fakat yaşamdaki etkisini bilmektedir.
0 puan	- Verilen problem durumunu anlamamıştır. - Öğrenci; verilen bilgiyi olduğu gibi aktarmakta, sıvı direncini ve yaşamadaki etkisini bilmemektedir.

ÖRNEK PUANLAMA ANAHTARI-3 (DERECELİ-ÖZEL)

Aşağıdaki grafik, A Grubu ve B Grubu olarak adlandırılan iki grubun bir fen bilimleri testinde aldıkları puanları göstermektedir.



A Grubu için ortalama 62,0 ve B Grubu için ortalama 64,5'tir. Puanları, 50 ya da daha fazla olan öğrenciler, bu testten geçerler. Bir öğretmen, grafiğe bakarak bu testte B Grubunun A Grubundan daha başarılı olduğunu ileri sürmektedir.

A Grubundaki öğrenciler, öğretmenleriyle aynı düşüncede değiller. Onlar, B Grubundaki öğrencilerin, daha başarılı sayılmamaları gerektiği konusunda öğretmenlerini inandırmaya çalışıyorlar.

Grafiği kullanarak A grubundaki öğrencilerin kullanabileceği matematiksel bir dayanak veriniz.

Dereceli Puanlama Anahtarı (Özel)

Kod	Doğru Yanıtlar	Puan
10	Geçen öğrencilerin sayısına, sınırlayıcıların orantısız etkisine ya da en üst düzeyde puan alan öğrencilerin sayısına bağlı olan geçerli kanıtlar.	1
11	• A Grubunda, B grubundan daha fazla öğrenci testten geçmiştir.	
12	• Eğer A Grubunun en zayıf öğrencisini dikkate almazsanız, A Grubundaki öğrenciler B Grubundaki öğrencilerden daha başarılı olmuştur. • B Grubu öğrencilerinden daha çok sayıdaki A Grubu öğrencileri, 80 ve üzeri puan almıştır.	
Yanlış Yanıtlar		0
20	Hiçbir matematiksel nedene dayanmayan ya da yanlış matematiksel nedenlere dayanan ya da basitçe farkları tanımlayan, ama B Grubunun daha iyi yapmamış olabileceğini belirtmeyen geçersiz kanıtlar dahil olmak üzere diğer yanıtlar.	
21	• Fen bilimlerinde A Grubu öğrencileri normal olarak B Grubu öğrencilerinden daha başarılıdır. Bu test puanları sadece bir rastlantıdır.	
22	• Çünkü B Grubu için en yüksek ve en düşük puanlar arasındaki fark A Grubununkinden daha küçüktür.	
23	• A Grubu 80-89 aralığında ve 50-59 aralığında daha iyi puan sonuçlarına sahiptir. • A Grubu, B Grubundan daha geniş çeyrekler-arası aralığa sahiptir.	
Boş ve ilgisiz yanıtlar		0
30		

4.3. Çoktan Seçmeli Test Sınavları

Çoktan seçmeli testler en sık kullanılan objektif sınavlardandır. Öğrenci, cevabı birden çok seçenek arasından seçip işaretlediğinden sınav çoktan seçmeli olarak adlandırılır. Bilişsel alanın “Hatırlama”, “Anlama” ve “Uygulama” gibi alt basamaklarındaki öğrenmelerin yoklanmasında yaygın olarak kullanılan çoktan seçmeli sınavların özellikleri Tablo 13’te sıralanmıştır.

Tablo 13. Çoktan seçmeli test sınavlarının olumlu ve olumsuz özellikleri

Olumlu özellikleri	Olumsuz özellikleri
• Çok sayıda kazanımı ölçmek mümkün, dolayısıyla kapsam geçerliliği yüksektir. • Öğrenci cevaplama özgürlüğüne sahiptir. • Birim hata düşük olduğundan güvenilirliği diğer yöntemlere göre daha yüksektir. • Puanlanması kolay ve objektiftir. • Kalabalık gruplara uygulanabilir. • Değerlendirilmesi çabuk ve kesindir.	• Hazırlanması zordur ve uzmanlık gerektirir. • Öğrencinin kendi yorumunu katabilme özgürlüğü yoktur. • Üst düzey öğrenmeleri ölçen soru sormada sınırlıdır. • Şans başarısına açıktır; bu nedenle öğrencinin bir konuyu gerçekten öğrenip öğrenmediği tam olarak kestirilemez. • Cevap seçenekler arasından seçildiğinden üst düzey düşünme becerileri geliştirmede sınırlıdır. • Öğrenci seçeneklere sınırlandığından sık kullanıldığında sorgulama ve öteleme yeteneklerinde gerilemeye neden olabilir.

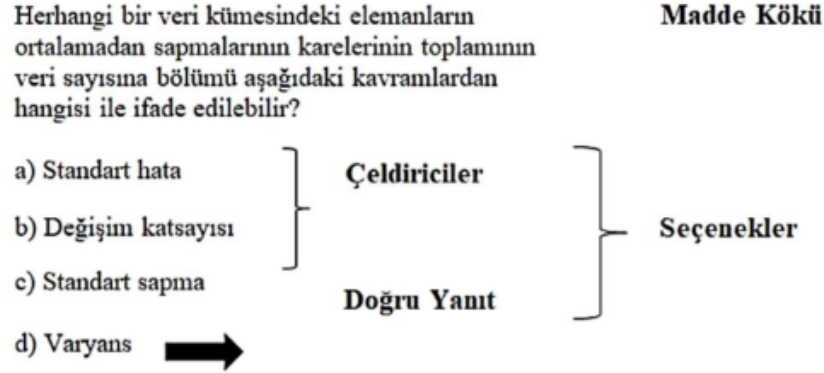
4.3.1. Çeldiriciler

Çeldiriciler, doğru cevap olmayan çoktan seçmeli cevap seçenekleridir (Şekil 12). Bunlar, genellikle öğrencilerin yaygın kavram yanlışlarına veya yanlış çıkarımlarına dayalı olarak geliştirilen ve bunların ötesine geçip geçmediklerini görmek için kullanılan makul ancak yanlış seçeneklerdir. Çeldiriciler akla yatkın seçenekler olmalıdır. Sınav hazırlayan eğiticiler, öğrencilerin kavram yanlışlarını, ödevlerinde yaptıkları hataları veya kaçırdıkları sınav sorularını çeldiricileri hazırlamak için kullanabilir. Çeldirici yazımına bu şekilde yaklaşıldığında, öğrencilerin yanlış cevap seçimleri konu ile ilgili bilgi düzeyleri hakkında fikir verebilir.

Çeldiricileri değerlendirirken göz önünde bulundurulması gereken bazı hususlar vardır;

- Her çeldirici madde yanıtlayıcıların bazıları tarafından mutlaka seçilmiş mi veya işaretlenmiş mi? Her bir madde için 4 olası seçeneğiniz varsa ancak öğrenciler bunlardan sadece bir veya ikisini sadece seçiyorsa, bu diğer çeldiricilerin etkisiz olduğunun bir göstergesidir. Düşük bilgili öğrenciler bile "gerçek" seçenekleri bir veya ikiye indirebilir, bu nedenle doğru seçim yapma olasılıkları yüksektir. Bu durum, bilgi sahibi olmayan öğrencilerin doğru tahmin yapma olasılığını artırmaktadır.
- Yanlış cevaplanan her soruyu ayrıntılı şekilde gözden geçirmek gerekli değildir. Bunun yerine, çeldiricilerin etkisiz gibi görüldüğü yerlerde dikkatli ve duyarlı olmak gereklidir. Tipik olarak bu sorular, hiç kimse tarafından seçilmeyen/işaretlenmeyen iki veya daha fazla çeldiricinin olduğu sorulardır.

- Çeldiriciler aşırı karmaşık, belirsiz şekilde ifade edilmiş veya açıkça yanlış seçenek olduğunu gösteren ipuçları bulundurmamalıdır.
- Teorik olarak, çeldiriciler konuyu iyi anlayan öğrenciler tarafından seçilmemelidir. Bu nedenle, (sınavdaki) en iyi öğrencilerinizin herhangi bir nedenle yanlış bir cevaba yönelip yönelmediğini belirlemek önemlidir.



Şekil 12. Çoktan seçmeli test sorusu düzeni

Çoktan seçmeli test sorusu hazırlarken dikkat edilmesi gerekenler:

Teknik özellikler

- Her soru mutlaka bir kazanımı ölçmelidir. Öğrencinin bir kazanımı iyi kavramamış olması başka bir kazanımdaki bilgisinin ortaya çıkarılmasının önüne geçmemelidir.
- Normal şartlardaki bir başarı testinde soruların çoğunluğu orta güçlükte olmalı, seçenekler yazılırken bu durum göz önünde bulundurulmalıdır.
- Sorular konu, uzunluk, dil ve anlatım bakımından homojen olmalıdır.
- İyi bir çoktan seçmeli test sorusu seçenekler kapatıldığında çözülebilir.
- Soruda kullanılan grafik, şekil veya diyagramlar metinle tutarlı olmalı, yanlış anlamaya neden olmamasına dikkat edilmelidir.
- Çoktan seçmeli test formatı soruyla yoklanan kazanıma uygun olmalıdır.
- Soru kökü ve seçeneklerde yoruma açık, subjektif ifadelerden kaçınılmalıdır.
- Soru kökünde “Hangisi yanlış olamaz?” “Hangisi olumsuz özelliği değildir?” gibi çifte olumsuz ifadeler yer verilmemelidir.
- Çeldiriciler dikkatli seçmeli ve diğerlerinden fazla çalışan çeldirici yazmamaya dikkat edilmelidir.

- Ayrıntı olarak görülebilecek önemsiz hususlarda soru yazmamaya dikkat edilmelidir.
- Öğrencilerin sınava motive olmalarını sağlamak için kolay sorular sınav formunun başında verilmelidir.
- Sorular zorluk dereceleriyle öğrencileri yıldırmamalı, öğrenmeye teşvik eder nitelikte olmalıdır.
- Soru veya seçeneklerde “genellikle”, “çoğu zaman”, “asla” gibi ifadeler sunulan önermenin yanlış olabileceğini akla getirdiği için kullanılmamalıdır.
- Ters de mümkün olabilir anlamına geldiğinden seçeneklerde “bazen” ifadesi kullanılmamalıdır. Aynı nedenden ötürü seçenekler –abilir, -ebilir gibi eklerle sonlandırılmamalıdır.

Mantıksal ipuçları ve test kurnazlığı

- Soru, test kurnazlığı yoluyla çözülmeye açık olmamalıdır.
- Sorular birbirinden bağımsız olmalı ve biri diğerine ipucu olmamalıdır.
- Seçenekler soru kökü ile tutarlı olmalıdır.
- Seçenekler ilgili konuda doğru cevaba yakınlıklarına göre tek bir düzlemde, en yakındaki seçenek doğru cevap, diğerleri çeldirici olacak şekilde sıralanmalıdır. Çeldiricilerin doğru cevaba yakınlığı arttıkça madde güçlüğü de artar. Ancak diğerleri hiç işaretlenmemişken, yani çalışmıyorken, başka bir çeldiricinin çok sayıda cevaplayıcıyı çekmemesine dikkat edilmelidir.
- Cevaplayıcı diğerlerinden farklı olana yönleneceğinden seçenekler olabildiğince aynı yönde ifade edilmelidir (hep olumlu ya da hep olumsuz).
- Seçenekler mantıksal bakımdan soru kökü ile uyumlu olmalıdır. Seçenekler arasında eleme yoluyla sorunun doğru cevabını bulduracak türde net ayrılıklar olmamalıdır.
- “Hepsi” seçeneği diğer seçeneklerden birinin olmadığı anlaşıldığında kolayca eleneceğinden kullanılmamalıdır. Benzer şekilde “Hiçbiri” seçeneği seçeneklerin birinin olmasının mümkün olduğu anlaşılırsa kolayca eleneceğinden tercih edilmemelidir.
- Soruda test çözme stratejilerinde ustalaşmış cevaplayıcılar için ipucu bulunmamalıdır. Örneğin; aşağıda 2010 yılı KPSS Tekrar sınavında sorulan eğitim bilimleri maddelerinden birinin seçenekleri verilmiştir.

a) 1A,2B,3A,4B b) 1A,2A,3B,4B c) 1B,2B,3A,4A d) 1B,2A,3B,4B e) 1A,2B,3A,4A

Burada madde kökünü görmeksizin cevabın A seçeneği olduğu kestirilebilir. Dikkatli bir cevaplayıcı ilk sırada en çok tekrarlayan 1A, ikinci sırada en çok tekrarlayan 2B ve üçüncü sırada en çok tekrarlayan 3A'nın sıralı yer aldığı “A” seçeneğinin doğru cevap olduğunu tahmin edebilir. Benzer şekilde, seçenekler sayı olduğunda cevabı ortadaki seçeneğe saklama (C seçeneğine),

doğru cevabı eksik bilgi kalmasın diye uzun yazma, aynı seçeneği üç defadan çok tekrar ettirmeme gibi durumların doğru cevabı ele vermesi önlenmelidir. Soru yazarı seçenekten yola çıkarak, test kurnazlığı ve mantık hatalarından yararlanarak soruyu çözmenin mümkün olmadığından emin olmalıdır.

Kullanılan dil ve anlatım

- Sorular açık ve anlaşılır biçimde ifade edilmelidir.
- Sorularda kullanılan dil, testin hedef kitlesinin düzeyine uygun olmalıdır.
- Soru kökü ve seçenekler Türkçe dil bilgisi kurallarına uygun olarak yazılmalıdır.
- Gereksiz tekrarlardan kaçınmalı, soruyu okuyup anlamak öğrencinin çok vaktini almamalıdır.
- Seçeneklerde soru kökü ile birlikte okunduğunda yanlış olduğuna işaret edecek uyumsuzluklar olmamalıdır.
- Okuma yükünün azaltılması testin geçerliliğini de arttıracığından, minimum kelime, maksimum anlam yaklaşımı izlenmelidir.
- Seçeneklerde ipucu verilecekse doğru cevaba saklanmalıdır. Çeldiriciler, konuyu öğrenmiş olan öğrencileri değil, bilmeyen ya da eksik/yanlış öğrenenleri yanıltacak şekilde olmalıdır.

Biçim ve düzen

- Soru kökündeki olumsuz ifadelerin altı çizilmelidir.
- Soruların hepsi eşit sayıda seçeneğe (Örn; tüm sorular 5 seçenekli) sahip olmalıdır.
- Seçenekler uzunlukları bakımından birbirine yakın olmalıdır.
- Cevap anahtarında desen oluşmamasına dikkat edilmelidir (AAB BBC CCD DDE ECA vb. gibi).
- Doğru cevapların sayısı eşit olmamalı, beş seçenekli bir test için her bir seçenek beş-altı soruda bir tekrarlanmalıdır (A B C D D E A gibi).
- Soru kökü yeterli uzunlukta olmalı ve soruyu çözmek için gerekli olan bilgiyi içermelidir.
- Biçimsel olarak soru ve seçenekler aynı sayfada yer almalıdır. Test yazı puntosu ve karakteri bakımından homojen olmalıdır.
- Sınavlarda kopya tehdidi olmadığı sürece seçenekler sayısal ifadelerse büyüklük sırasında, tek kelimelikse alfabetik sırada verilmelidir. Literatürde seçeneklerin sayısal veya alfabetik sırada verilmesinin pratikte etkisine dayanak oluşturacak sayıda araştırma mevcut değildir. Ancak alfabetik ve sayısal sıra izlenmesi psikolojik bakımdan mükemmeliyetçi öğrencileri rahatlatılabilir, ayrıca bu şekilde öğrencinin şifre arayışına girerek zaman kaybetmesinin önüne geçilir. Bu nedenle seçeneklerin sıralı verilmesi testlerde daha uygundur.

Tablo 14. Uygun hazırlanmayan çoktan seçmeli soru örnekleri

Örnekler	Sorunlar
1. Elli bir yaşında Tip II DM’u olan erkek hasta sol bacak ön yüzünde kızarıklık ve ağrı yakınması ile başvuruyor. Hastanın ateşi 39°C, TA:150/90 mmHg ve sol bacakta dizden bileğe kadar şişlik, kızarıklık ve ısı artışı dışında patoloji saptanmıyor. Bu hastanın tedavisinde en uygun seçenek aşağıdakilerden hangisidir? a. 21 gün Amikasin 1x1 g b. Aztreonam 3x500 mg-7 gün c. 400 mg Teikoplanin günde 1 kez-5 gün d. Günde 4 kez Ampisilin Sulbaktam 1.5g- 15 gün*	• “En uygun” kalın yazılmamış. • Seçeneklerde verilen ilaçlara yönelik bilgiler birbiri ile uyumsuz. Her seçenekteki ilaç; ilaç ismi, dozu, veriliş süresi gibi bir sıralamada verilmeli. • Doğru cevap seçeneği diğer seçeneklerden daha uzun. Konuya çalışmamış bir öğrenci, bu seçeneği dikkat çekici bulup işaretleyebilir.
2. Astımda hava yolu inflamasyonunun oluşmasında aşağıdaki faktörlerden hangisi rol oynamaz? a. Genetik yatkınlık b. CD8 T lenfositlerin inflamasyonu artırıcı rol oynaması* c. Alerjen teması d. Eozinofillerden inflamatuvar sitokinlerin salgılanması	• Doğru cevap olan “b” seçeneğinde “rol oynaması” cümlesi kullanılmış ve doğru cevap belli edilmiştir. • Seçeneklerde rol alan faktörlerin sadece isimlerinin verilmesi yeterli olacaktır. • “rol oynamaz” olumsuz ifadesinin altı çizilmesi ve kalın yazılması bu örnekte doğru bir uygulamadır.
3. Akut miyokard infarktüsünde hangi zaman aralığı haricinde ışık mikroskopisinde patolojik değişiklikler izlenir? a. 3-4.günler b. 4-12. saatler c. 1-2.günler d. 0-4.saatler*	• Soru kökü yeterince iyi anlaşılmıyor. • Seçeneklerde sayılar küçükten büyüğe sıralama yapılmamış. • Seçeneklerde sayılar gün ve saat şeklinde karışık bir şekilde verilmiş. Sadece gün veya sadece saat şeklinde verilmeli. • “b” ve “e” seçeneğinde 4. saat ortak olarak verilmiş.
4. Kimyasal nörotransmisyonun aşamalarını sıralayınız. I. Nörotransmitterin salınımı II. İkincil mesajların uyarılması III. G proteinlerin bağlanması IV. Birinci mesajın reseptöre bağlanması V. Genom düzeyinde etkinin oluşması a. I, II, IV, II ve IV b. IV, I, II, III ve V c. I, II, III, IV ve V d. I, III, V, IV ve II e. I, IV, III, II ve V*	• Soru kökü uygun şekilde hazırlanmamış. • “a” seçeneğinde “II” iki kez kullanılmış. • a, c, d ve e seçenekleri hep “I” ile başlamış. Çeldiriciler uygun değil. • Birçok şık “V” ile bitmiş. • Aşamaları sıralayın denildiği için virgül yerine tire işareti konulması daha uygun olacaktır.
5. Romatoid artrit tanısı koymada sizce en uygun serum proteini aşağıdakilerden hangisidir? a. Haptogloblin b. Transferrin c. Seruloplazmin d. C-reaktif protein* e. Prealbümin	• Sorudaki “sizce” ifadesi uygun değil. • Sorudaki “en uygun” ifadesi kalın veya altı çizili olarak vurgulanmamış. • Soruda serum proteini soruluyor ve d şıkında “protein” kelimesi var. Doğru cevaba yönlendiriyor.
6. Genital tüberkülozda en sık rastlanan yakınma aşağıdakilerden hangisidir? a. Karın ağrısı b. İnfertilite* c. Öksürük d. Hepsi e. Hiçbiri	• Soruda sorunlar mevcut. En sık rastlanan “yakınma” soruluyor ancak c şıkkı bir yakınma değil. • Seçeneklerde “hepsi” ve “hiçbiri” kullanılmış. • “En sık” ifadesi vurgulanmamış. • a ve c şıkları solunum sistemi yakınmaları.

4.4. Doğru-Yanlış Soruları

Objektif testlerden biri olan doğru-yanlış sorularından öğrenciye bir grup tek cümlelik kısa önerme verilir. Öğrenciden cümlenin sonunda boş bırakılan alana, önermeler doğruysa “D”, yanlışsa “Y” harfi yazması istenir. Doğru-yanlış soruları çoktan seçmeli testler gibi objektif puanlanabilir. Ancak şans başarısı bu testlerde daha yüksektir (1/2). Doğru-yanlış soru formatı kesinlikle doğru ve kesinlikle yanlış önermeler için kullanılmalıdır. Aksi takdirde yanıltıcı olabilir ve yanlış öğrenmelere yol açılır.

Doğru-Yanlış soruları hazırlarken dikkat edilmesi gerekenler:

- Testteki doğru-yanlış madde sayıları birbirine eşit ya da yakın olmalıdır. Yanlışlığından emin olunan yeterince önerme yoksa testte doğru önermeler çoğunlukta olacaktır. Bu durumu fark eden cevaplayıcı bilmediği önermeleri doğru olarak işaretleyerek testi manipüle edebilir.
- Önermelerin uzunlukları birbirine yakın olmalıdır.
- İfadeler kısa, basit ve anlaşılır olmalıdır.
- Önermeler kesinlik taşımamalıdır. -ebilir, olsa da olur, da mümkündür, muhtemeldir gibi kesinlik taşımayan ifadelerden kaçınılmalıdır.
- Olumsuz ifadeler dikkatli kullanılmalı, özellikle çifte olumsuz önermeler kullanılmamalıdır.

Hatalı madde: Maddelerin anlaşılmasının güçlüğü güvenilirliği olumsuz etkiler. ()

Doğru madde: Test maddelerinin anlaşılır olması güvenilirliği olumlu etkiler. ()

- Her maddede yalnızca bir önermenin doğruluğu veya yanlışlığı ölçülmelidir.

Madde sayısını artırmak, maddelerin anlaşılır olması.....etkiler.

(Cevaplayıcının bu soruyu doğru yapması için kapsam geçerliliği ve güvenilirlik terimlerinin her ikisini de bilmesi gerekir).

- Virgülle birleştirilmiş önermeler kullanılmamalıdır.

Hollanda'nın el yapımı çakıları, İsveç'in saatleri meşhurdur. ()

- Yanlış sayısı doğru sayısından düşülerek, puanlar şans başarısından arındırılmalıdır.
- Kafa karıştırıcı önermelerden kaçınılmalıdır.

4.5. Eşleştirme Soruları

Eşleştirme soruları iki grup halinde verilen, birbiriyle ilgili olan bilgilerin belli bir açıklamaya göre eşleştirilmesini gerektirir. Terimlerin tanımlarıyla, sembollerin anlamlarıyla, ilkelerin açıklamalarıyla, tarihi olayların tarihleriyle, yazarların eserleriyle, problemlerin çözümleriyle ve buluşların kaşifleriyle

eşleştirilmesini gerektiren objektif puanlanabilen test sorularıdır. Öncüller ve seçenekler arasındaki ilişkilere göre öğrenciden eşleştirme yapması beklenir. Seçenek sayısı öncüllerden daha fazla tutularak öğrencinin seçenekleri eleyip bazı öncüllere cevap bulması ihtimali engellenir.

Eşleştirme sorularının özellikleri:

- Eşleştirme soruları, seçenek sayısının beşten çok olduğu durumlarda kullanılır.
- Üst düzey bilişsel alan davranışlarını ölçmekten ziyade, olgu bilgisi türünde ya da temel düzeyde soruları cevaplarıyla eşleştirilme şeklindeki sorular için uygundur.
- Eşleştirme sorusu; yönerge, öncüller ve seçeneklerden oluşur.
- Öncüller sayfanın solunda, seçenekler ise sağında yer alır. Öncüllerin sol başında yer alan parantezin içine seçeneklerin başında yer alan harflerden uygun olanı yazılarak eşleştirme maddesi cevaplanır.
- Eşleştirme sorularıyla çok sayıda kazanımı yoklamak mümkündür.

Eşleştirme sorusu hazırlarken dikkat edilmesi gereken hususlar:

- Eşleştirme sorularının yönergeleri cevaplayıcının ne yapacağını anlayacağı açıklıkta olmalıdır.
- Eşleştirme grubundaki öncül ve seçenekler konu birliği bakımından homojen olmalı; birden çok kazanım yoklanmamalıdır.
- Sayıca çoklukları düşünüldüğünde, eşleştirme maddelerinin ifadelerinin kısa olması gereklidir.
- İfadelerden daha uzun olanlar öncül olarak kullanılmalıdır.
- Cevap için ayrılan kutucukların öncüllerin başında olması puanlamayı kolaylaştırır.
- Eşleştirme maddesinde hatırd tutma güçleşeceğinden öncüllerin sayısı en az altı, seçeneklerin sayısı en fazla 15 olmalıdır.
- Seçenekler kelimeyse alfabetik, sayıysa büyüklük sırasında verilmelidir.
- Eşleştirme maddelerinde seçenek sayısı öncül sayısından fazla olmalıdır.

Örnek Eşleştirme Sorusu

Yönerge: Aşağıda I. sütunda verilen bilim adamlarının her biri II. sütunda yer alan tekniklerden birini geliştirmiştir. I. sütunda belirtilen bilim adamlarının II. sütunda verilenlerden hangisini literatüre kazandırdığını I. sütunun başında yer alan boşluğa yazarak cevaplayınız.

I	II
() 1. Spearman	A. ANOVA
() 2. Pearson	B. Çan Eğrisi
() 3. Galton	C. Fan Tablosu
() 4. Gauss	D. İç tutarlılık
() 6. Kuder-Richardson	F. Momentler Çarpımı Korelasyonu
	G. Sıra Farkları Korelasyon Katsayısı
	H. Normallik testi
	I. Tüm gruba ait güvenirlilik
	İ. %27'lik alt grup üst grup kıyaslaması

4.6. Boşluk Doldurma Soruları

Boşluk doldurma soruları eksik soru kipinde çoktan seçmeli sorulara (boşluğu tamamlayan ifade aşağıdakilerden hangisidir) benzer. Ancak burada seçenekler sunulmaksızın cevaplayıcının, önermede boş bırakılan yere cevabı yazması beklenir. Boşluk doldurma soruları Bloom'un Bilişsel Alan Taksonomisi'nde alt basamaklardaki kazanımların ölçülmesi için uygundur. Bir cümlelik önermedeki boşluk, önemli bir tarih, yer ya da kişi olabileceği gibi bir teori, kural ya da özellik de sorulabilir. Soru, önerme formatında ya da soru cümlesi formatında olabilir; ancak cevaplayıcıdan ne yapmasının beklendiği açık olmalıdır. Boşluk doldurma soruları eğitim seviyesi düşük bireylerde uygulandığında güvenilir sonuçlar verdikleri için tercih edilir. Bu tür sorularda şans başarısının olmaması bir avantajdır. Sorular kısa olduğu için çok sayıda soru sormak mümkündür ve bu yönüyle kapsam geçerliliği bakımından avantajlıdır. Ancak bu soru türünün üst düzey madde yazılmasına olanak tanımaması önemli bir sınırlılığıdır.

Boşluk doldurmalı soru hazırlarken dikkat edilmesi gerekenler:

- Mümkün olduğunca boşluk cümlelerin sonuna getirilmelidir.
- Her soruda tek bir boşluk yer almalı, dolayısıyla tek kazanım sorulmalıdır.

Örnek: dünyanın güneşin etrafında döndüğünü ifade eden yüzünden engizisyon mahkemesinde yargılanmıştır (Burada öğrenci ilk boşluktaki ifadenin “Galilei” olduğunu bilebilir ancak “Copernicus Kuramı”nı bilmeyebilir.)

- Boşluktan cevabın tahmin edilmesini önlemek için maddelerin boşluk uzunluğu standart olmalı; her maddede en uzun cevap uzunluğunda boşluk bırakılmalıdır.

- Öğrencilerin boşlukların uzunluğundan cevabı tahmin etmelerine neden olacağından metin, boşluklar eşit tutularak yazılmalıdır.
- Soru kökünde -e,-a,-de,-da, -nin, -nun gibi ipucu olabilecek eklere yer verilmemelidir.

Yanlış: Oltu taşlarıyla ünlü olan ilimiz.....'dur. (Cevap: Erzurum)

Doğru: Oltu taşlarıyla ünlüdür. (Cevap: Erzurum)

- Öğrenciyi ezbere yönelteceğinden ya da ezberleyeni ödüllendireceğinden, sorular ders kitabından cümlelerin kalıp olarak alınması şeklinde oluşturulmamalıdır.

4.7. Sözlü Sınavlar

Sözlü sınavlarda öğrenci, kendisine sözlü olarak yöneltilen soruları yine sözlü olarak cevaplar. Bu tür sınavlarla öğrencinin bilgi, beceri ve tutumunu bir arada ölçmek mümkündür. **Yapılandırılmış ve klasik sözlü** olmak üzere yaygın olarak kullanılan iki türü vardır. Yapılandırılmış sözlü sınavlar bir komisyon tarafından önceden hazırlanmış soruların genellikle 3-5 kişilik gruplarla sınava giren öğrencilere sırasıyla yöneltilmesi şeklinde yürütülür. Klasik sözlüler bir komisyon önüne öğrencilerin tek tek çıkması şeklinde yürütülebileceği gibi sınıf ortamında rastgele adı çağrılan öğrenciye öğreticinin soru yöneltilmesi şeklinde de yapılabilir.

Sözlü anlatım becerisinin yeterli olduğu bu tür sınavlar yazma becerisi iyi olmayan veya okuma-yazma bilmeyen kişilere de uygulanabilir. Bununla birlikte öğrencilerin yabancı dilde konuşma düzeyini ölçmek için de sözlü sınavlara başvurulur. Sözlü sınav öncesi, her konudan düşük, orta ve yüksek güçlük düzeyinde toplam öğrenci sayısının üç katı kadar soru hazırlanmasına dikkat edilmelidir. Sınav, komisyon tarafından yapılacaksa komisyon üyelerinin elinde birbirine alternatif olabilecek çok sayıda farklı güçlükte soru olmasında fayda vardır. Sorular sayıca az ve güçlük düzeyi bakımından dengesiz olursa aynı sorular tekrar edebileceğinden, sınavın sonlarına doğru sözlü sınav olanların lehine bir eşitsizlik doğacaktır. Sözlü sınavlar, öğrencinin sınıfı geçip kalması hakkında kurul kararı vermek amacıyla kullanılacaksa sınav öncesinde eğitimciler bir araya gelerek soruları gözden geçirmelidir.

Diğer sınavlarda olduğu gibi, sözlü sınavlarda da puanlama için cevap anahtarı ve yönerge hazırlanmalıdır. Bu şekilde soruyu yönelten komisyon üyesi, cevap anahtarına göre açıklatıcı soruları yerinde sorabilir. Her sınav aslında öğrenci açısından bir öğrenme fırsatıdır, bu nedenle iyi bir soru öğrenciyi daha fazlasını öğrenmeye sevk etmelidir. Sorular bu tür bir yaklaşımla yöneltilirse, sözlüler tehdit edici olmaktan çıkıp, öğrenciler için kendilerini gösterme şansına, sınavı yapanlar içinse öğrenciyi daha yakından tanıma fırsatına dönüşebilir.

Sözlü sınavların en büyük eksikliği subjektifliğe açık olmasıdır. Bireyin dış görünüşü, giyimi, iletişim becerisi, sınav komisyonu üyelerinin aday hakkındaki ön bilgi ve görüşleri puanlamaya etki edebilir. Bu nedenle sözlü sınavlar mümkünse kamerayla kayıt altına alınmalıdır. Sözlü sınav yaparken dikkat edilmesi gereken bazı hususlar şu şekilde özetlenebilir;

- Sözlü sınavın başında öğrenci kısa bir sohbetle rahatlatılırsa sözlülerin doğasında olan stres hafifletilmiş olur.
- Sözlü sınavlarda pozitif bir yaklaşımla öğrenciye söylediklerini açıklatıcı sorular yöneltilebilir. Adayın ne demek istediğinden emin olunamadığında verilen cevap tekrar soruları ile onaylatılabilir.
- Öğrencinin söylemek istediğinin doğru anlaşıldığından emin olunması, subjektifliğe oldukça açık olan bu sınav türünün güvenilirliğini artırır.
- Açıklatıcı sorular genellikle adayın son söylediği cümlenin ardından sorulduğundan, aday bir konuda önemli maddeleri veya alt başlıkları sıralarken en iyi bildiğini en sona saklayarak bir sonraki soruyu garantileyebilir. Sözlü sınavların bu şekilde manipüle edilmesi mümkün olduğundan sınav komisyonu dikkatli olmalıdır.
- Öğrenciler sözlü sınavlarda konuyu iyi bildikleri bir başlığa veya sınav komisyonu üyelerinin önemsedikleri bir konuya çekmeye çalışabilir. Yanlılık etkisinden dolayı adayın çok iyi açıkladığı bir sorudan sonra, cevabı nispeten yetersiz bir sorunun olumsuz etkisi düşük olacaktır. Haksız avantaj sağlayacağından bu şekilde yönlendirme yapılmasına göz yumulmamalıdır.
- Cevapların yazılı olarak kaydedildiği durumda adayın sınav tutanağını okuyup imzalaması sınavın güvenilirliğini artırmada etkilidir. “Ne?” ve “Nasıl?” soruları detay almak için, “Sen bu konuda ne düşünüyorsun?” şeklindeki sorulara adayın yansıtma becerisini ortaya koymak amacıyla kullanılabilir.
- Yazılı olarak kayıt altına alınan sınavda adayın cevapları tek tek değerlendirilmeli, eksik olduğu noktalar sınav tutanağında belirtilmelidir.

5. Test Geliştirme Süreci ve Madde/Soru İstatistikleri

5.1. Test Geliştirme Süreci Basamakları

Amacın belirlenmesi: Sınavın türü ne olursa olsun öncelikle yapılması gereken, sınavın amacının açık ve net olarak belirlenmesidir.

Kapsamın belirlenmesi: Testin uygulanmasının amacı belirlendikten sonra, bu amaca hizmet edecek konu kapsamının ne olacağının tespit edilmesi gerekir. Testin amacı öğrencilerin belirli periyotlar sonunda öğrenmelerinin izlenmesi, eksiklerinin belirlenerek geri dönüt verilmesi ise testte her hedef-davranışı ölçen **en az üç soru** bulunmalıdır. Eğer testle amaçlanan, öğrenim süreci sonunda öğrencilerin düzeylerini belirlemek ise, bu durumda her bir kritik davranışı ölçen **en az iki sorudan** oluşan bir testin hazırlanması uygun olacaktır. Böylece testin kapsamını geçerli bir şekilde ölçecek soru türü (doğru-yanlış, kısa cevaplı, çoktan seçmeli vb.) ve sayısı da belirlenmiş olacaktır.

Soruların yazılması ve düzeltilmesi: Testin kapsamı belirlenerek bu kapsamı ölçecek uygun soru türü ve yeterli soru sayısı belirlendikten sonra, dersin uzmanları tarafından soru yazımına geçilmelidir. Sorular yazıldıktan sonra gözden geçirilerek yazım hatalarına, cümle düşüklüklerinin olup olmadığına dikkat edilmelidir. Testin görünüş geçerliğinin uygun olup olmadığı kontrol edilmelidir. Ayrıca testi hazırlayan uzman dışında başka bir uzman tarafından soruların incelenmesi tavsiye edilmektedir. Böylece hazırlayan kişinin gözden kaçırdığı hatalar ya da eksiklikler daha iyi belirlenmiş olacaktır. Tüm bu düzeltme işlemleri yapıldıktan sonra ön uygulamanın yapılabilmesi için test çoğaltılarak uygulamaya hazır hale getirilmelidir.

Ön uygulama: İdeal olan, testte bulunması amaçlananın en az iki/üç katı sorunun bulunduğu ön testin, asıl (nihai) uygulamanın yapılacağı gruba benzer bir gruba, soru analizinin yapılabilmesi için uygulanmasıdır. Ancak eğitim öğretim sürecindeki sınav uygulamalarında, ön uygulamanın yapılması çok fazla mümkün olmamaktadır.

Soru analizi ve soru seçimi: Ön uygulamanın yapıldığı öğrenci grubundan alınan cevaplardan yararlanılarak, asıl teste alınabilecek en güvenilir ve uygun soruların seçimine geçilmelidir.

Son testin oluşturulması, uygulanması ve puanlanması: Testin amacına uygun hazırlanan, en az iki/üç katı sorudan oluşan testin ön uygulaması sonucu elde edilen cevapların kullanılarak soru analizi yapıldıktan sonra, son (nihai) testi oluşturacak en güvenilir sorular belirlenmelidir. Bu sorulardan oluşan son test, ölçmenin yapılması istenilen asıl gruba uygulanmalıdır. Uygulama sonucu elde edilen

cevaplar puanlanarak ölçme süreci tamamlanmış olur. Bundan sonraki süreçte elde edilen puanlarla öğrencilerin başarı düzeyleri vb. durumlar hakkında bilgi edinmek mümkün olacaktır.

5.2. Madde/Soru İstatistikleri

Madde/soru istatistiklerini hesaplayabilmek için her öğrencinin soru bazında teste verdiği cevapları bilmek gerekir. Dolayısıyla soru puanları matrisi olmadan soru istatistiklerinin çoğunu hesaplamak mümkün olmayacaktır. Soruların istatistiklerini belirlemede; madde güçlük indeksi, madde ayırt edicilik gücü indeksi, madde varyansı, madde standart sapması ve madde güvenilirlik katsayısından yararlanılabilir. Bu bölümde madde güçlük ve ayırt ediciliği analizleri ele alınacaktır.

5.2.1. Madde Güçlük İndeksi

Testteki sorulara ait madde güçlük derecesinin genel anlamda orta düzeyde olması tavsiye edilir. İyi bir testte zor ve kolay sorular azınlıktayken, orta güçlükte maddeler çoğunlukta olmalıdır. Madde güçlük derecesi tek başına bir soru hakkında karar vermek için yeterli değildir. Bu nedenle madde ayırt edicilik gücü ile birlikte ele alınmalıdır.

Kolay ve zor sorular, sorunun ve testin ayırt ediciliğini düşürür. Kuramsal olarak incelendiğinde, soru güçlük düzeyi orta düzeyde olan soruların ayırt edicilik gücünün daha yüksek olduğu görülecektir. Ancak bu zor ve kolay soruların sorulmaması gerektiği anlamına gelmez. Başarılı öğrencileri teşvik için ayırt edici sorular sormak, zayıf öğrencileri cesaretlendirmek için kolay sorulara yer vermek yerinde bir karar olacaktır. Ancak analiz, değerlendirme ve yaratma gibi üst düzey bilişsel alan davranışlarının test formatında yoklanması güçlüğünden dolayı testlerde çok fazla üst düzeyde soru sorulamadığı da belirtilmelidir.

Madde güçlük indeksi, 0 ve 1 arasında bir değer alır. Bir soruyu ne kadar az kişi doğru cevaplarsa soru o kadar zordur ve madde güçlük derecesi 0'a yakın olacaktır. Soru ne kadar çok kişi tarafından doğru cevaplanırsa o derece kolaydır ve madde güçlük derecesi 1'e yakın olacaktır. Madde güçlük düzeyine göre maddelerin sınıflandırması birçok kaynakta değişebilmektedir. Bu kaynaklara göre en çok kullanılan sınıflama yöntemleri ve yorumlamaları Tablo 15 ve 16'da verilmiştir.

Tablo 15. Madde güçlük indeks aralıkları (5 kategori)

Değer	Kategori	Yorum
0.00 - 0.15	Çok Zor	Soru/madde çıkarılmalı
0.16 - 0.40	Zor	Soru/madde revize edilmeli ^a
0.41 - 0.60	Orta	Soru/madde kabul edilebilir ve idealdir
0.61 - 0.80	Kolay	Soru/madde revize edilmeli ^b
0.81 - 1.00	Çok Kolay	Soru/madde çıkarılmalı

a: Önemli bir konuda ise ve kolaylaştırmak mümkün değilse, başarılı öğrencileri motive etmek amacıyla testte kullanılabilir.

b: İhtiyaca göre revize edilerek zorlaştırılabilir. Bu mümkün değilse ve önemli bir özelliği ölçüyorsa, öğrencilerin kaygı düzeyini azaltmak için testin başında kullanılabilir.

Tablo 16. Madde güçlük indeks aralıkları (3 kategori)

Değer	Kategori	Yorum
< 0.30	Zor	Soru/madde çıkarılmalı veya revize edilmeli ^a
0.30 - 0.70	Orta	Soru/madde kabul edilebilir ve idealdir
> 0.70	Kolay	Soru/madde çıkarılmalı veya revize edilmeli ^b

a: Önemli bir konuda ise ve kolaylaştırmak mümkün değilse, başarılı öğrencileri motive etmek amacıyla testte kullanılabilir.

b: İhtiyaca göre revize edilerek zorlaştırılabilir. Bu mümkün değilse ve önemli bir özelliği ölçüyorsa, öğrencilerin kaygı düzeyini azaltmak için testin başında kullanılabilir.

Literatürde, soru analizlerini hesaplamak için çeşitli yöntemler mevcuttur. Bu yöntemlerden en yaygın kullanılanlar; **Henryson Yöntemi** ve **Basit Yöntem**'dir. Bu iki madde analizi yönteminin birbirlerinden ayrıldığı temel nokta, hesaplamalarda kullanılan ölçme sayıları, yani hesaplama da dahil edilen örnek sayısıdır. Henryson yönteminde, ölçmeye dahil edilen tüm yanıtlayıcılar kullanılırken, basit yöntemde tüm yanıtlayıcıların toplam puanlarından en başarılı %27'lik üst ve en başarısız %27'lik alt grup belirlenerek, kitlenin %54'lük kısmı kullanılır.

Henryson Yöntemine göre, madde güçlük indeksi, soruyu doğru cevaplayanların tüm cevaplayıcı sayısına oranıdır. Oranlama yapıldığında sınıfın yüzde kaçının soruyu doğru yanıtladığı görülebilir. Bu indeks 0 ile 1 arasında değerler alır. Güçlük indeksi 0'a yaklaştıkça sorunun zor bir madde, 1'e yaklaştıkça sorunun kolay bir madde olduğu yorumu yapılabilir. Bu yöntem e göre güçlük indeksi aşağıdaki formülle hesaplanır;

$$P = \frac{N_a}{N}$$

N_a : Doğru yanıtlayan kişi sayısı

N : Tüm cevaplayıcıların sayısı

Basit yöntem e göre güçlük indeksinin hesaplanabilmesi için alt ve üst grupların belirlenmesi gerekir. Bu yöntemde öncelikle, ölçme aracından elde edilen toplam puanlar en yüksekten düşüğe doğru sıralanır.

Başarısı en yüksek olan ve başarısı en düşük olan cevaplayıcıların %27'si alınarak alt ve üst gruplar belirlenir. Buradan hareketle madde güçlük indeksi (p_i) aşağıdaki formülle hesaplanır;

$$p_i = \frac{\frac{\text{Maddeyi üst grupta doğru cevaplayanların sayısı}}{\text{üst gruptaki öğrenci sayısı}} + \frac{\text{Maddeyi alt grupta doğru cevaplayanların sayısı}}{\text{Alt gruptaki öğrenci sayısı}}}{2}$$

İki yöntemden herhangi birinden elde edilen indeks değerine göre maddenin güçlüğü ile ilgili Tablo 15 veya 16 yardımıyla yorum yapılabilir.

5.2.2. Madde Ayırt Edicilik İndeksi

Ayırt edicilik indeksi, bir sorunun başarı düzeyi yüksek cevaplayıcılar ile düşük cevaplayıcıları ayırt etme derecesidir. Özetle, ilgili sorunun bilenle bilmeyeni birbirinden ayırabilme ölçüsüdür. Ayırt edicilik, sorunun/maddenin geçerliliği ile ilgili bir indekstir. Ayırt edicilik indeksi, -1 ile +1 arasında değer alır. İndeksin 0'a yaklaşması durumunda, ilgili sorunun alt ve üst grubu ayırt ediciliği düşük, 1'e yaklaşması durumunda ise sorunun alt ve üst grubu ayırt ediciliği yüksektir olarak yorum yapılabilir. İndeksin pozitif olması, sorunun doğru cevaplanma oranının üst grupta yüksek olduğunu, negatif olması ise sorunun doğru cevaplanma oranının alt grupta yüksek olduğunu gösterir. Ayırt ediciliği negatif olarak belirlenmiş bir soru testten çıkarılmadan önce bu durumun sorudaki bir belirsizlik ya da cevap anahtarındaki yanlışlık gibi nedenlerden kaynaklanıp kaynaklanmadığı kontrol edilmelidir. İlgili yanlışlık bu durumlardan kaynaklıysa ve düzeltilirse indeks pozitifte dönebilir. Madde ayırt edicilik indeks aralıkları Tablo 17'de verilmiştir.

Tablo 17. Madde ayırt edicilik indeks aralıkları

Değer	Kategori	Yorum
> 0.40	Çok iyi	Ayırt ediciliği çok yüksek soru/madde
0.30 - 0.39	Oldukça İyi	Ayırt ediciliği iyi soru/madde
0.20 - 0.29	Orta düzey	Ayırt ediciliği orta düzeyde soru/madde, revize edilmeli
0.10 – 0.19	Düşük	Ayırt ediciliği zayıf/uygun olmayan soru/madde, çıkarılmalı
< 0.10	Çok düşük	Ayırt ediciliği çok zayıf soru/madde, kesinlikle çıkarılmalı

Ayırt edicilik için en çok tercih edilen yöntem, basit yöntemle göre hesaplamadır. Buna göre ayırt edicilik indeksi hesaplanırken, yöntem gereği yanıtlayıcılar alt ve üst grup olarak ikiye ayrılır. Bu yöntemde, öncelikle ölçme aracından elde edilen sonuçlara göre toplam puanlar hesaplanır ve en yüksekten

düşüğe doğru sıralanır. Başarısı en yüksek olan %27'lik grup üst grup, başarısı en düşük olan %27'lik grup alt grup olarak alınır. Bu yöntemle göre ayırt edicilik indeksi (a_i) aşağıdaki formülle hesaplanır;

$$a_i = \frac{\text{üst grupta doğru cevaplayan sayısı} - \text{alt grupta doğru cevaplayan sayısı}}{\text{grupların herhangi birindeki öğrenci sayısı}}$$

Örnek Soru

	A	B	C	D	E *
Tüm grup öğr.s.:113	6	12	4	4	87
Üst grup öğr.s.: 31	0	1	0	0	30
Alt grup öğr.s.: 31	3	8	2	2	16

* Doğru seçenek

Yukarıdaki örneğe göre %27'lik üst grup ve alt grup öğrenci sayısı = 31

Doğru cevap olan "E" şıkkına üst gruptan doğru cevap veren öğrenci sayısı = 30

Doğru cevap olan "E" şıkkına alt gruptan doğru cevap veren öğrenci sayısı = 16

$$\text{Madde güçlüğü} = \frac{30 + 16}{31 + 31} = 0.74 \quad \text{Madde ayırt ediciliği} = \frac{30 - 16}{31} = 0.45$$

Testin güçlüğü **kolay** düzeyde, ayırt ediciliğinin ise **çok iyi** düzeyde olduğu söylenebilir.

5.2.3. Açık Uçlu Maddelerin Analizleri

Açık uçlu maddeler için **madde güçlüğü (p)** ve **ayırt edicilik (D)** hesaplaması aşağıda gösterilmiştir;

$$p = \frac{\text{Maddeden alınan ortalama puan} - \text{Maddeye verilen en düşük puan}}{\text{Madde puan ranjı}}$$

$$D = \frac{\text{Üst grubun madde puan ortalaması} - \text{alt grubun madde puan ortalaması}}{\text{Madde puan ranjı}}$$

Madde puan ranjı: Puanlama anahtarındaki alınabilecek en yüksek ve en düşük puanlar arasındaki farkı temsil eder. "0" ile "5" arası derecelenmiş 6 dereceli bir puanlama anahtarında en yüksek puan "5", en düşük puan ise "0" olduğunda madde puan ranjı 5-0=5 olarak hesaplanır.

Örnek Soru (Açık uçlu madde güçlük derecesi hesaplama)

İç hastalıkları servisinde görev yapan bir hemşire, diyabetli bir hastada hipoglisemi belirtileri fark eder. Hemşirenin bu duruma müdahalesi nasıl olmalıdır? Ayrıca, hastanın yaşamsal bulgularını değerlendirirken hangi belirtilere özellikle dikkat etmesi gerekmektedir?

Yukarıda verilen açık uçlu soru, 10 öğrenci tarafından cevaplanmış ve öğretim elemanı tarafından yedili derecelendirilmiş (0-6) bir dereceli puanlama anahtarı ile puanlanmıştır. Buna göre bu açık uçlu sorunun madde güçlük indeksinin hesaplanması aşağıdaki gibidir;

10 öğrencinin bu maddeden aldığı puanlar:

Öğrenciler	Alınan puan
1. Öğrenci	5
2. Öğrenci	6
3. Öğrenci	2
4. Öğrenci	4
5. Öğrenci	5
6. Öğrenci	3
7. Öğrenci	1
8. Öğrenci	2
9. Öğrenci	6
10. Öğrenci	1

Maddeden alınan ortalama puanı belirlemek için her öğrencinin maddeden aldığı puanlar toplanır ve öğrenci sayısına bölünür;

$$\frac{5 + 6 + 2 + 4 + 5 + 3 + 1 + 2 + 6 + 1}{10} = 3.5$$

Elde ettiğimiz sonuçları yukarıda belirtilen formüle yerleştirdiğimizde sonuç aşağıdaki gibi olacaktır;

$$p = \frac{3.5 - 0}{6} = \mathbf{0.58}$$

Bu sonuca göre ilgili maddenin **ideal (orta) güçlükte** olduğu söylenebilir.

Örnek Soru (Açık uçlu madde ayırt edicilik indeksi hesaplama)

Bir açık uçlu soruya 45 öğrencinin verdiği yanıtlar, öğretim elemanı tarafından kullanılan dereceli puanlama anahtarıyla (0-5) puanlanmıştır. Bu açık uçlu sorunun madde ayırt edicilik indeksinin hesaplanması aşağıdaki gibidir;

İlk olarak öğrenci puanları en yüksekten en düşüğe sıralanır. Grubun en başarılı %27'si ile en başarısız %27'si alınır. Sınava giren 45 öğrenci olduğundan, %27'si 12 öğrenciye denk gelmektedir. Aşağıdaki tabloda açık uçlu maddelerden en yüksek puan alan 12 ve en düşük puan alan 12 öğrenci olmak üzere 24 öğrencinin verisi sunulmuştur.

Grup	Soru 1	Soru 2	Soru 3	Soru 4	Toplam puan
Üst grup	5	5	5	3	18
	5	4	4	5	18
	4	5	5	4	18
	4	4	5	4	17
	3	5	4	5	17
	4	4	5	3	16
	2	5	5	4	16
	5	5	3	3	16
	2	3	5	5	15
	4	5	3	3	15
	4	4	3	4	15
	2	4	5	3	14
Alt grup	3	1	1	2	7
	2	1	1	2	6
	1	1	2	2	6
	0	1	2	2	5
	4	0	0	1	5
	0	0	1	3	4
	1	0	2	1	4
	1	1	1	1	4
	2	1	0	0	3
	0	1	2	0	3
	0	0	2	1	3
	0	1	0	1	2
Üst grup ortalama	3.67	4.42	4.33	3.83	
Alt grup ortalama	1.17	0.67	1.17	1.33	
D	0.50	0.75	0.63	0.50	

Soru 1 için;

Üst grubun madde puan ortalaması:

$$\frac{5 + 5 + 4 + 4 + 3 + 4 + 2 + 5 + 2 + 4 + 4 + 2}{12} = 3.67$$

Alt grubun madde puan ortalaması:

$$\frac{3 + 2 + 1 + 0 + 4 + 0 + 1 + 1 + 2 + 0 + 0 + 0}{12} = 1.17$$

Madde puan ranjı: $5 - 0 = 5$

Elde ettiğimiz sonuçları yukarıda belirtilen formüle yerleştirdiğimizde sonuç aşağıdaki gibi olacaktır;

$$D^* = \frac{3.67 - 1.17}{5} = 0.50$$

Maddenin ayırt edicilik indeksinin **çok iyi** düzeyde olduğu söylenebilir.

Önemli Noktalar

- Hazırlanan sınavdaki/testteki soruların zorluk derecesini incelerken aşağıdaki durumlar göz önünde bulundurulmalıdır;
 - Öğrenciler hangi soruları kolay, hangilerini zor buldu? Bu sorular sizin öğrenciler için kolay/zor olacağını düşündüğünüz sorularla eşleşiyor mu? Örneğin; bazen bir eğitmen bir soruyu sınavdaki en kolay sorulardan biri olduğuna inanarak teste koyabilir ancak öğrenciler bu soruyu zor bulabilir.
 - Çok kolay ve çok zor sorular, konuyu bilen ve bilmeyen öğrenciler arasında ayırım yapma konusunda genellikle iyi bir performans ortaya koymaz. Ancak eğiticinin hazırladığı sınava her iki soru türünü koymak için bazı nedenleri olabilir. Örneğin; bazı eğitmenler sınavlarına kasıtlı olarak çok kolay bir soruyla başlayarak kaygılı sınav katılımcılarını sakinleştirir ya da öğrencilerin sınavda erken bir başarı hissetmelerine yardımcı olur.
- Eğer bir soru neredeyse herkesin doğru bileceği kadar kolay (Örn; güçlük derecesi = 0.98) veya neredeyse herkesin yanlış bileceği kadar zorsa (Örn; güçlük derecesi = 0.12), konuyu gerçekten bilenlerle bilmeyenleri ayırt etmek çok zorlaşır. Ancak bu, tüm çok kolay ve çok zor soruların elenmesi gerektiği anlamına gelmez. İyi bir ayırt edicilik göstermeyeceği bilinen sorular, öğrencileri gerçekten zorlama ya da herkesin belirli bir içeriği bildiğinden emin olma gibi bir niyet ile teste eklenebilir. Bununla birlikte, uygun olmayan biçimde yazılmış bir sorunun ayırt etme yeteneği çok az olacaktır.
- Genellikle madde ayırt ediciliğinin en az 0.30 olması tavsiye edilir. Ayırt ediciliği negatif olan sorular teorik olarak ya testte genel olarak düşük performans gösteren öğrencilerin soruyu doğru bildiğini ya da genel test performansı yüksek olan öğrencilerin soruyu doğru bilemediğini gösterir. Dolayısıyla, bu durum bir dizi soruna işaret edebilir;
 1. Puanlama anahtarında bir hata olabilir.
 2. Sınava yeterince hazırlanmamış öğrenciler doğru tahminde bulunmaktadır.
 3. Sınava iyi hazırlanmış öğrenciler bir şekilde yanlış cevap vererek yanılabilmektedir.Bu durumlarda önlem alınmalıdır. Madde ayırt ediciliği negatif olan sorular mutlaka incelenmelidir. Öğrencilere sadece 1'i anahtar ve diğerleri çeldirici ifadeler içeren toplam 4 veya 5 seçenek sunulduğunda güvenilirlik artacaktır. Sadece tek bir olası cevap olduğundan, sorunun ve seçeneklerin uygun bir şekilde yazıldığından ve cevap anahtarının doğru olduğundan emin olunmalıdır.
- Soruları çıkarmak veya revize etmek için hem güçlük düzeyini hem de ayırt edicilik indeksini birlikte ele almak önemlidir. Güçlük düzeyi 0.30-0.70 olan bir sorunun kabul edilebilir bir ayırt edicilik

oranına (en az 0.30 veya üzeri) sahip olması beklenir. Bir sorunun güçlük düzeyi orta ancak ayırt ediciliği çok düşükse, bu sorunun gözden geçirilmesi veya çıkarılması gerekebilir. Ancak bir soru 0.28 gibi sınırdan bir ayırt edicilik oranına sahip, güçlüğü orta ise bu sorunun mutlaka gözden geçirilmesi gerekmeyebilir.

- Yukarıdaki tablolarda belirtilen güçlük düzeyi ve ayırt edicilik indeksleri için tanımlanan ölçütler dikkate alınarak soruların yorumlanması için Tablo 18’de bir örnek oluşturulmaya çalışılmıştır.

Tablo 18. Madde güçlük ve ayırt edicilik indeksine göre soru yorumlamaları

Güçlük indeksi	Ayırt edicilik indeksi	Yorum
Kolay soru (< 0.30)	• < 0.10 • 0.10-0.19 • 0.20-0.29 • 0.30-0.39 • > 0.40	• Soru çıkarılmalı • Soru çıkarılmalı • Soru mutlaka revize edilmeli • Soru mutlaka revize edilmeli • Soru revize edilebilir
Orta düzey soru (0.30-0.70)	• < 0.10 • 0.10-0.19 • 0.20-0.29 • 0.30-0.39 • > 0.40	• Soru çıkarılmalı • Soru çıkarılmalı • Soru revize edilebilir • Soru kabul edilir • Soru kabul edilir
Zor soru (> 0.70)	• < 0.10 • 0.10-0.19 • 0.20-0.29 • 0.30-0.39 • > 0.40	• Soru çıkarılmalı • Soru çıkarılmalı • Soru revize edilebilir • Soru revize edilebilir • Soru kabul edilir

KAYNAKLAR

Başol, G. (2019). Eğitimde Ölçme ve Değerlendirme (6. baskı). Pegem Akademi.

Belay, L. M., Sendekie, T. Y., & Eyowas, F. A. (2022). Quality of multiple-choice questions in medical internship qualification examination determined by item response theory at Debre Tabor University, Ethiopia. *BMC Medical Education*, 22(1), 635. <https://doi.org/10.1186/s12909-022-03687-y>

Bhat, S. K., & Prasad, K. H. L. (2021). Item analysis and optimizing multiple-choice questions for a viable question bank in ophthalmology: A cross-sectional study. *Indian Journal of Ophthalmology*, 69(2), 343–346. https://doi.org/10.4103/ijo.IJO_1610_20

Bhattacharjee, S., Mukherjee, A., Bhandari, K., & Rout, A. J. (2022). Evaluation of Multiple-Choice Questions by Item Analysis, from an Online Internal Assessment of 6th Semester Medical Students in a Rural Medical College, West Bengal. *Indian Journal of Community Medicine*, 47(1), 92–95. https://doi.org/10.4103/ijcm.ijcm_1156_21

Chae, Y. M., Park, S. G., & Park, I. (2019). The relationship between classical item characteristics and item response time on computer-based testing. *Korean Journal of Medical Education*, 31(1), 1–9. <https://doi.org/10.3946/kjme.2019.113>

Christian, D. S., Prajapati, A. C., Rana, B. M., & Dave, V. R. (2017). Evaluation of multiple choice questions using item analysis tool: a study from a medical institute of Ahmedabad, Gujarat. *International Journal Of Community Medicine And Public Health*, 4(6), 1876–1881. <https://doi.org/10.18203/2394-6040.ijcmph20172004>

Çıkrıkçı, N. (2022). Eğitimde Ölçme ve Değerlendirme (6. baskı). Anı Yayıncılık.

De Champlain, A. F. (2010). A primer on classical test theory and item response theory for assessments in medical education. *Medical Education*, 44(1), 109-17. <https://doi.org/10.1111/j.1365-2923.2009.03425.x>

Ebel, R. L., Frisbie, D. A. (1991). *Essentials of educational measurement* (5th ed). Prentice Hall.

Friatma, A., Anhar, A. (2019). Analysis of validity, reliability, discrimination, difficulty and distraction effectiveness in learning assessment. *Journal of Physics: Conference Series*, 1387, 012063. <https://doi.org/10.1088/1742-6596/1387/1/012063>

Güler, N. (2019). Eğitimde Ölçme ve Değerlendirme (14. baskı). Pegem Akademi.

- Haladyna, T. M., Downing, S. M., Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), 309- 333. https://doi.org/10.1207/S15324818AME1503_5
- Hasançebi, B., Terzi, Y., Küçük, Z. (2020). Madde Güçlük İndeksi ve Madde Ayırt Edicilik İndeksine Dayalı Çeldirici Analizi. *Gümüşhane Üniversitesi Fen Bilimleri Dergisi*, 10 (1), 224-240. <https://doi.org/10.17714/gumusfenbil.615465>
- Hingorjo, M. R., Jaleel, F. (2012). Analysis of one-best MCQs: the difficulty index, discrimination index and distractor efficiency. *J Pak Med Assoc.* 62(2), 142-47.
- Karakaya, İ. (2022). Açık uçlu soruların hazırlanması, uygulanması ve değerlendirilmesi. Pegem Akademi.
- Kaur, M., Singla, S., & Mahajan, R. (2016). Item analysis of in use multiple choice questions in pharmacology. *International journal of Applied & Basic Medical Research*, 6(3), 170–173. <https://doi.org/10.4103/2229-516X.186965>
- Kumar, D., Jaipurkar, R., Shekhar, A., Sikri, G., & Srinivas, V. (2021). Item analysis of multiple choice questions: A quality assurance test for an assessment tool. *Medical journal, Armed Forces India*, 77(Suppl 1), S85–S89. <https://doi.org/10.1016/j.mjafi.2020.11.007>
- McDonald, M. (2018). *The nurse educator's guide to assessing learning outcomes* (4th ed.). Jones & Bartlett Learning.
- Oermann, M. H., Gaberson, K. B. (2021). *Evaluation and testing in nursing education* (6th ed.). Springer Publishing Company.
- Raagas, E. L., Absin, S. B. (2004). *Introduction to Measurement and Evaluation, Concepts and Applications*. Iwag Printing and Publications.
- Sekerci A. R. (2015). Development of Diagnostic Branched Tree Test For High School Chemistry Concepts. *Oxidation Communications*. 38(2A), 1060–1067.
- Sutherland, K., Schwartz, J., Dickison, P. (2012). Best practices for writing test items. *Journal of Nursing Regulation*, 3(2), 35-39. [https://doi.org/10.1016/S2155-8256\(15\)30217-9](https://doi.org/10.1016/S2155-8256(15)30217-9)
- Tekindal, S. (2020). *Eğitimde Ölçme ve Değerlendirme* (6. baskı). Pegem Akademi.