

İÇİNDEKİLER

BÖLÜM 1

ÖLÇME VE DEĞERLENDİRMEDE TEMEL KAVRAMLAR

A. ÖLÇME	1
B. ÖLÇME TÜRLERİ	1
C. ÖLÇMEDE SIFIR	2
D. ÖLÇMEDE BİRİM	2
E. ÖLÇME KURALI	3
F. DEĞİŞKEN	4
G. ÖLÇEK VE TÜRLERİ	5
1. Sınıflama Ölçeği	5
2. Sıralama Ölçeği	5
3. Eşit Aralıklı Ölçekler	5
4. Oranlı Ölçekler	5
H. DEĞERLENDİRME	7
1. Amaca Göre Değerlendirme	9
2. Ölçüte Göre Değerlendirme	10
I. ÖLÇMEDE HATA	11
1. Hata Kaynakları	11
2. Hata Türleri	12
a. Sabit Hata	12
b. Sistemotik Hata	12
c. Tesadüfi Hata	12

BÖLÜM 2

ÖLÇME ARAÇLARINDA BULUNMASI GEREKEN ÖZELLİKLER

A. GEÇERLİK	13
1. Kapsam Geçerliği	14
2. Ölçüt Geçerliği	14
3. Yapı Geçerliği	14
4. Görünüş Geçerliği	14
5. Geçerliği Etkileyen Faktörler	15
B. GÜVENİRLİK	17
1. Test - Tekrar Test	18
2. Paralel (Eşdeğer) Formlar	18
3. İki Yarı (Eşdeğer Yarılar)	18
4. İç Tutarlık KR-20/KR-21	18
5. Ölçmenin Standart Hatası	19
6. Güvenirlilik Artırma Yolları	21
7. Güvenirlilik-Geçerlik Arasındaki İlişki	21
C. KULLANISILILIK	22

BÖLÜM 3

SINAVLAR VE TESTLER

A. TESTLERİN GRUPLANDIRILMASI... 24

1. Uygulama Zamanına Göre Testler... 24

2. Hazırlanma Biçimine Göre Testler... 24

B. EĞİTİMDE KULLANILAN TESTLER 25

1. Yazılı Yoklamalar... 25

2. Kısa Cevaplı (Boşluk Doldurma) Testler 26

3. Doğru-Yanlış Testleri... 26

4. Çoktan Seçmeli Testler... 27

5. Eşleştirmeli Testler... 28

6. Sözlü Yoklamalar... 28

C. YENİ (GAĞDAŞ) ÖLÇME DEĞERLEN-

DİRME ARAÇLARI... 31

1. Portfolyo (Öğrenci Etkinlik Dosyası)... 31

2. Performans Değerlendirme... 33

3. Rubrik Puanlama Anahtarı... 34

4. Kontrol Listeleri... 35

5. Gözlem Formları... 35

6. Derecelendirme Ölçekleri... 36

7. Kelime İlişkilendirme Testleri (KIT)... 36

8. Otomatik Değerlendirme... 37

9. Yapılandırılmış Grid... 37

10. Tanıyıcı Doldurulmuş Ağaç (TDA)... 38

11. Proje Yöntemi... 38

12. Kavram Haritaları... 39

13. Görüşme Tekniği... 39

14. Öz Değerlendirme... 40

15. Akran Değerlendirme... 40

BÖLÜM 4

1. TEST İSTATİSTİKLERİ... 47

A. MERKEZİ EĞİLİM ÖLÇÜLERİ... 49

Aritmetik Ortalama, Ağırlıklı Arit. Ortalama,
Ortanca (Medyan), Tepe Değer (mod)

B. MERKEZİ DAĞILIM ÖLÇÜLERİ... 53

C. STANDART PUANLAR... 58

1. Z Puanı... 58

2. T Puanı... 59

D. DAĞILIM EĞRİLERİ... 59

E. TESTİN AYIRT EDİCİLİĞİ... 61

F. TESTİN ORTALAMA GÜÇLÜĞÜ... 62

G. MUTLAK BAŞARI YÜZDESİ... 62

2. MADDE İSTATİSTİKLERİ... 63

A. MADDE GÜÇLÜK İNDEKSİ... 64

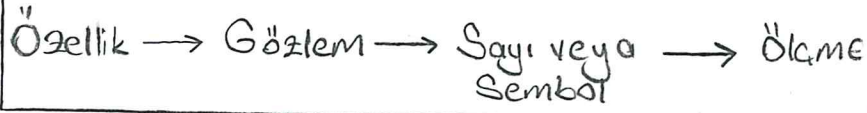
B. MADDE AYIRT EDİCİLİK İNDEKSİ... 65

C. MADDE VARYANSI... 68

D. MADDE STANDART SAPMASI... 68

ÖLÇME

Ölçme; bir betimleme, gözlem işidir. Bir nesnenin belli bir özelliğe sahip olup olmadığını. Sahipse, sahip oluş derecesinin gözlenip, gözlem sonuçlarının Sayı veya Sembollerle ifade edilmesidir. Ölçme işlemi farktan doğmuştur.



ÖRNEK: Masanın boyu 150 cm'dir
Odanın sıcaklığı 20°C'dir
Otomobilin hızı (80 km)
Sınav sonuçları (60 p.)

Kumaş boyunun m ile ölçülmesi
Oda uzunluğunun ip ile ölçülmesi
Bir avval patatesin tartılması

Ölçüm; ölçme işlemi sonunda elde edilen sayı veya semboldür.

ÖRNEK: Hakan 2019 KPSS'de 90 puan aldı. } sonucu ölçmedir fakat
ölçüm } 90 ölçümdür

Ölçme Türleri

Doğrudan Ölçme

Ölçülen nitelik ile ölçmede kullanılan aracın niteliği aynıdır. Ölçülecek değerler doğrudan gözlenir, somuttur.

- Eksi değerler olmaz. Fiziksel özellikler ölçülür.
- Başlangıç noktası gerçek sıfırdır.
- Sonuçları oranlı olarak karşılaştırılabilir.

ÖRNEK: 20 m. 10 m.'nin iki katıdır.

Dolaylı Ölçme

Kendine ait ölçme aracı yoktur. Başka araçlardan yararlanılır. Doğrudan gözlenemeyen soyut durumlardır.

- Eksi değerler alınır. İzafi sıfırla başlanır.
- Kişi beğenisi, zekası, tutumu vs. ölçülür.
- Sonuçları oranlı olarak karşılaştırılmaz.

ÖRNEK: 100 p. alan öğrenci, 50 p. alan öğrencinin iki katı zekidir diyemeyiz.

1. Göstergeyle ölçme 2. Türetilmiş ölçme

Ölçmede Sıfır Noktası

Mutlak / Doğal Sıfır

Ölçme işleminde yokluk, hiclik ifade eder.

ÖRN: Öğretmen sınıfa gelir, sınıfı bastırır
gerçekten öğrenci yoktur

İzâfi / Keyfi Sıfır

Ölçek üzerinde herhangi bir nokta sıfır kabul edilir. Uzmanlar tarafından tanımı yapılır, yokluk-hiclik ifade etmez.

ÖRN: Kopya çeken öğrenciye, hastalığı yarından baş koştur veren öğrenciye verilen sıfır.
(Bu sıfır onların bir şey bilmediğini ifade etmez!)

Not = Eğitimde kullanılan tüm testler

Ölçmede Birim

Bir ölçme aracını oluşturan en küçük parçadır.

Ölçme sonuçlarının sayılara ifade edilmesi için standart birimlere ihtiyaç vardır.

ÖRN: Metrenin \rightarrow cm'leri (100 cm.)
Kilogramın içindeki gr'lar (1000 gr.)
Bir saatte altmış dakika (60 sn.)

Doğal Birim

Tabiatla var olan her değişken doğal birimdir

ÖRN: Sınıfta bulunan her öğrenci

Eşitlik: Aracı oluşturan birimlerin hepsi birbirine eşittir.

Genellik: Herkesin aynı iş için aynı birimi kullanması

Kullanışlılık: Bir birimin ölçülecek niteliğe uygun olması

Tanımlanmış Birim

Uzmanlarca, ne kadarlık miktara birim derecesi sonradan açıklanmış, karşılaştırılmıştır.

ÖRN: 1 m. 100 cm'den oluşur.

Ölame Kuralı

Ölameye konu olan özelliğin hangi miktora ne değeri verileceğinin belirlenmesidir.

ÖRNEK: Öğrencilere yapılan sınavda hangi soruya kaç puan verileceğinin belirlenmesi.

Futbol karşılaşmalarında kendi sahasında kazanılan maç için 3 puan verilmesi

DEĞİŞKEN

En az iki değeri alabilen, gözlemden gözleme değışik değeri alabilen özellik veya durumlara "değişken" denir.

ÖRNEK: Boy uzunluğu, Ağırlık, Hava sıcaklığı, Matematik yeteneği

<u>Değişken Türü</u>	<u>Tanım</u>	<u>Örnek</u>
1. Nicel Değişken	Sayılarla ifade edilebilen özelliklerdir.	Ağırlık, Uzunluk, Başarı puanı, Öğrenci sayısı
2. Nitel Değişken	Sıfat veya Sembolle ifade edilebilen özelliklerdir.	Cinsiyet, Dil, Ayakkabı markası, Saç rengi, Evli, Bekar vs.
3. Sürekli Değişken	Matematiksel olarak iki değeri arasına bir başka değeri alabilen değişkenler.	1 puan 2 puan 1.1 / 1.3 / 1.5
4. Süreksiz Değişken	Sınırlı sayıda değeri alabilen değişkenler	Sıfatlar, Semboller, Doğal Sayılar, Cinsiyet
5. Bağımsız Değişken (Sebebi)	Araştırma sonucuna etkisi incelenen değişken	<u>Deney çalışmasının</u> Bağımsız D.
6. Bağımlı Değişken (Sonuç)	Bağımsız D. etkilenen değişken	<u>başarıya etkisi-</u> Bağımlı D.

Değişken	Nitel	Nicel	Süreklî	Sürekli
Saç rengi	X			X
Uzunluk		X	X	
Zeka Puanı				
Medeni durum				
Sıcaklık				
Zaman				
Plakalar				

Örnek	Bağımsız Değişken	Bağımlı Değişken
Eğitimde bilgisayar tek. kullanımının öğrencilerin derse yönelik ilgi düzeylerinin artması.	Bilgisayar teknolojilerinin kullanımı	Derse yönelik ilgi
Ailenin öğrenciye ödevlerinde yardımcı olma düzeylerinin okuldaki başarıya etkisi		
Öğrencilerin derse devansızlık yapmalarında öğretmenin verdiği cezanın etkisi		
Bir sporcuların ayakta duruşunun kaliteli olmasının ve beslenmesine dikkat etmesinin başarılı olmasına etkisi.		
Kanser hastalığına yakalanma da sigara içmenin ve genetik özelliklerin etkisi		
Çoklu zeka kuramına göre yürütülen derslerde öğrencinin akademik başarı düzeyinin yükseltilmesi		

Not: Nitel değişkenler sınıflama ve sıralama ölçekleri ile
Nicel değişkenler eşit aralıklı ve orantı ölçeklerinde elde edilir.

ÖLÇEK VE TÜRLERİ

Ölçek: Ölçme işleminde kullanılan araçtır. İncelenen özelliğin niteliğini, niceliğini, sınırlarını belirleyen araçlardır.

Sınıflama Ölçekleri

Belli bir gruba ait olup olmayı gösteren "daire"ler veren ölçeklerdir.

Benzer ve farklılıklara göre gruplara ayırır. Sadece nitel ayrımlardır.

Rakamlar matematiksel değeri yoktur. Sadece verileri birbirine benzerlik benzerliğine göre gruplara ayırır.

En basit verileri toplar

Başlangıç noktası ve birimi yoktur.

~~Cinsiyet ayrımı~~ T.C. no'ları

Öğrenci numaraları Sınıf rengi

İl Plakaları Futbolcu no'ları

Yüzde Frekans Mod

Sıralama Ölçekleri

Belirli bir özelliğe sahip olma bakımından nesneleri ve olayları sıraya koyar.

Nesne ve olaylara verilen sayılar incelenen özelliğe ne ölçüde sahip olduklarını gösterir fakat;

Ölçümler arasında ne kadar farklılık olduğunu göstermez.

~~Ölçümler~~ A=95 B=80 C=65

Ağardan hafif = A > B > C

Sadece A en ağır B en hafif

A ile B, B ile C arasında fark eşit diyemeyiz!

Yüzde Medyan (ortanca) Sıra farkları (korelasyon)

Eşit Aralıklı Ölçek

Birimleri arasında farklılıklar eşit olan, negatif değerler alabilen ve başlangıç noktası 0'a eşit olan ölçeklerdir.

~~Ölçek~~ Bir öğrenci matematik dersinden 60 aldığında diğer öğrencinin bu konuyu hakkında hiçbir şey bilmediği söylenemez.

Eğitimde kullanılan ve öğrencilerin zihinsel becerilerini ölçen araçların hepsi eşit aralıklı ölçeklerdir.

Sosyal bilimlerde kullanılır.

~~Not:~~ Sadece aralıklı karşılaştırma yapılabilir.

Zeka, Kızılık, Detenek, ilgi, Başarı testleri Termometre / Takvim

Oranlı Ölçek

Birimler arasında farklılıklar eşit, başlangıç noktası mutlak sıfır ve sıfırdan her zaman pozitifdir. Farklar birbirine her zaman eşittir. Fen bilimlerinde kullanılır.

Bütün ölçeklerin yaptığı işi yapar. En kapsamlı bilgi bundan alınır.

Uzunluk = m/dm

Ağırlık = gr / ton

Zaman = sn / saat

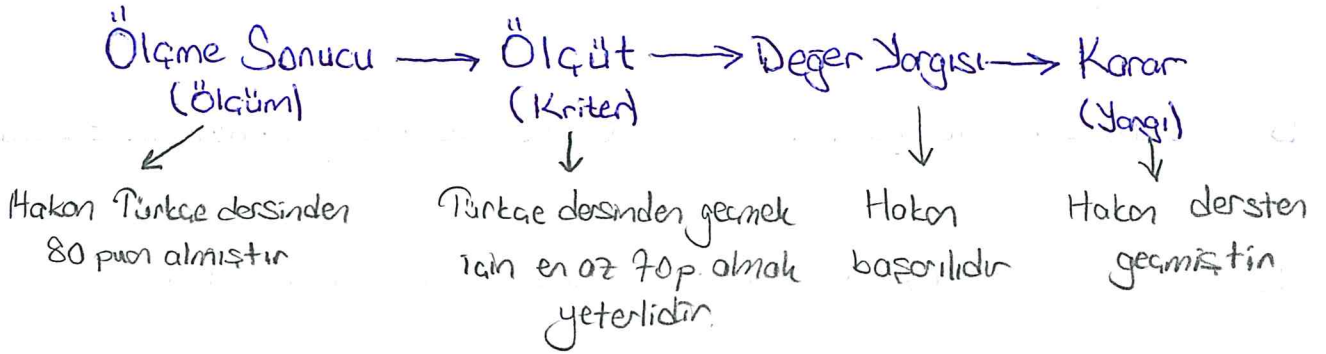
~~Ölçek~~ A=50 cm B=100 cm

AB = 1/2 oranında sığdırılır.

Her türlü istatistiksel hesaplamalar yapılır

ÖLÇME	Belli bir nesnenin ya da nesnelerin belli bir özelliğe sahip olup olmadıklarının gözlenip sonuçlarının sayı veya sembollerle gösterilmesidir.
ÖLÇÜM	Ölçme işleminden sonra elde edilen sayı veya semboldür.
ÖLÇÜT	Ölçme sonucunda konu verilecek değerdir.
ÖLÇME KURALI	Ölçmeye konu olan özelliğin hangi miktara ne kadar verileceğinin belirlenmesidir.
ÖLÇEK	Ölçmeye konu olan özelliklerin sınıflanması, sıralanması, derecelenmesi, miktarların belirlenmesi için uygulanması gereken kurallar ve sınırlara araçlarıdır.
Ölçmede Sıfır	
Ölçmede Birim	
DEĞERLEN DİRME	Ölçme araçlarıyla elde edilen ölçümlerin, bir ölçütle karşılaştırılarak anılan kategorize edilmesidir.

DEĞERLENDİRME



Değerlendirme, ölçme sonuçlarının bir ölçütle karşılaştırarak karar verilmesidir.

Not: Ölçme sonuçlarını kendi başına değerlendiremeyiz!
Ölçüt olmadan değerlendirme olmaz.

ÖLÇÜT (Kriter)

Ölçme sonuçlarının karşılaştırıldığı değerdir.

Sınır, koşul veya kural belirtir.

Mutlak Ölçüt

Ölçüt, değerlendirilecek grup dikkate alınmadan, önceden kesin olarak belirtilir.

- Soru sayısına göre
- Hedef ve davranışlar
- Kesin sınıra göre
- Notla göre,

ÖRNEK: 45 ve üzeri başarılı sayılır
TYT'de başarı için 185 p. gerekir
ALES'ten 70 p. almak gerekir
Mülakatta 60 p. almak gerekir.

Bağıl Ölçüt

Ölçüt, gruba ve ortalama başarıya göre, başarılarının başarısına göre belirtilir.

- Öğrenciler birbirine göre,
- Sınıftaki başarıya göre,
- Standart sapmaya göre,
- Gm eğrisine göre,

ÖRNEK: Gm eğrisiyle sınıf geçmek,
öss'de alımlara yerleşmek
Sınıf ve okul birincisi olmak
Akademik başarı için ilköğ'e girmek,

Ölçme	Değerlendirme
Değerlendirmenin ilk basamağıdır	Ölçmeyi de kapsama alan karar sürecidir.
Nitelikleri nicilemektir ve miktar belirtir	Ölçme sonuçlarını ölçütle karşılaştırarak miktar ve amaca göre karara verir.
Objektif olmalıdır	Subjektif olabilir.

Ölçme (Ölçme Sonucu)	Ölçüt (Kriter)	Değerlendirme (Karar)
Fatma'nın boyu 1,68 cm'dir	Bayanın polis olabilmesi için boyunun 162 cm + olması gerekir.	Fatma polis olabilir.
Emre, Mat. sınavından 78 p. almıştır	Mat. sınavından geçmek için en az 70 p. almak gerekir.	Emre, mat. dersinden geçmiştir.
Suyun sıcaklığı 68 °C'dir.	Suyun kaynaması için sıcaklığın 100 °C olması gerekir.	Su kaynamamıştır.
Arabanın hızı 88 km'dir	Ceza almanın için en fazla 90 km hız yapılabilir	Bu araba ceza almaz.

1. Amaca Göre Değerlendirme (Hedefe Dayalı)

a. Tanıma ve Yerleştirme (Diagnostic)

Dönem Başı

Öğrencilerin; bir derse, kursa veya yeni bir süre başlamadan önce sahip oldukları bilgi seviyesini ölçmek için yapılır.

Öğrencinin hazırbulunusluk, ilgi ve ön bilgi düzeyi belirlenir.

ÖRNEK

: Yabancı dil muafiyet sınavı, TST, LGS müzik, Resim, Beden olarındaki seçme sınavları

b. Yerleştirme ve Geliştirme (Formative)

Dönem İçi

Öğrencilerin; süreci içinde belirli bir zaman kodu işlenen konulardaki bilgi seviyesini test etmek için yapılır.

Sorular, o zamanı kodu olan konuları kapsar.

Eksik ve yanlış öğrenmeleri belirlemek için yapılır.

ÖRNEK

: Vize, Quiz, Ünite Sonu Sınavları,

c. Seviye Belirleme (Summative)

Dönem Sonu

Eğitim - öğretim süreci sonunda istenen hedefe ulaşma derecesini belirlemek için yapılır.

Bütün bir dönem boyunca işlenen konular sorulur.

Üst kademeye geçmek amacıyla yapılır.

ÖRNEK

: Final Sınavları, Yüklek Lisans Testleri
Doktora Sınavları (sunumları)

2. Ölçüte Göre Değerlendirme (Norma Göre)

Mutlak Değerlendirme

Ölçüt, ölçme sonuçları dikkate alınmadan belirlenir.

Ölçme sonuçları birbirinden bağımsız olarak değerlendirilir.

Bireysel değerlendirilir. Gruptaki diğer bireylerin aldığı notların başka öğrenci üzerinde etkisi yoktur.

ÖRNEK: 5. Sınıf'a başvurmak için 25'ten 90 olmak.
Müzikler geçmek için 60 p. almak.

Bağıl Değerlendirme

Ölçüt, ölçme sonuçlarından sonra grubun performansına göre belirlenir, ona göre değerlendirme yapılır.

Öğrenciler rekabet içindedir.

Bir kişinin kazansı diğerinde başkalarının kazansı da etkilidir.

En iyi olanların seçildiği grup.

ÖRNEK: Tarih dersinden başarılı olabilmek için sınıf ort. üstünde olmak.
Turna olmak için ilk 3'e girmek.

Ölçme	Ölçüt	Değerlendirme	Türü
Mehmet sınavdan 76 p. aldı.	Dersten geçmek için 70 almak yeterli	Mehmet dersten geçti.	Mutlak Değerlendirme
Tugba okul birincisidir.	Okul vakti ilk 5 kişiye burs verecektir.	Tugba, bursu kazandı.	
Honde, bilgi yarışmasında 2. oldu	Yarışmada ilk üç öğrenci ödül alacak.	Honde, ödül kazandı.	
Merve, sınavda 18. oldu.	İlk 20'ye giren öğrencilerden derece sınıfı yapılacaktır.	Merve, derece sınıfına girdi.	
Sinem, sınavdan 82 p. aldı.	Sınavda 80 p. üstünde olanların ataması yapılacaktır.	Sinem, atandı.	

ÖLÇMEDE HATA

$$\text{Hata} = \text{Gerçek Değer} - \text{Gözlemlenen Ölçüm}$$

Hata; Ölçülmek istenen özelliğin gerçek değeri ile ölçme yapıldıktan sonra elde edilen değer arasındaki farktır.

Hatayı ne azaltır? Teknolojiden yararlanılarak geliştirilen ölçme araçları
Ölçme işlemi yapacak kişinin donanım ve bilgileri
Uygun ortamda ve zamanda ölçme yapması.

Hata Kaynakları

Ölçmeciden Kaynaklanan Hatalar	Ölçme Aracından Kaynaklanan Hatalar	Ölçülen Özelliğten veya Ortamdan Kaynaklanan Hatalar
Ölçmeyi yapan kişinin eğitimi, psikolojik durumu, duyu organları dikkat düzeyi, motivasyonu. Öğretmenin cevap kağıtlarını okurken dikkatsiz olması, kopya çekmeye izin vermesi, dikkati dağılmasına rağmen okumaya devam etmesi.	Ölçme aracının hatalı üretilmesi, duyarlı olmaması, kullanımının karnosik ve zor olması. Testteki soruların açık olmaması Modellerin öğrenci düzeyine uygun olmaması Farklı alanlarla ilgili bilgi yoldanması.	Ölçme ortamının şartlarının uygun olmaması Sınav esnasında sırtın çok soğuk veya gürültülü olması Öğrencilerin testten önce puanla korkutulması

Hata Türleri

Sabit Hata

Ölçme sonucunda her ölçmeye aynı yönde olan ölçüm hatasıdır.

ÖRN: Öğretmen, öğrencilerin boyunu

ölçmek için metre getirir ve sırayla ölçer. Sonra metrenin 1 cm kırık olduğunu fark eder.

Herkese 1 cm eksik ölçer.

Sistematik Hata

Ölçülen büyüklüğe, ölçmeye ölçme koşullarına ve duruma bağlı olarak miktarı değişen hatalardır.

Hata, bütün ölçme sonuçlarına farklı miktarda yansır

ÖRN: Her öğrenciye aldığı puanın % 10'u kadar puan eklemek.

Tesadüfi Hata

Hatanın kaynağı bilinmeyen ve ölçme sonuçlarına gelişigüzel karışan hatalardır.

ÖRN: Sınav başansı, sınav günü öğrencinin hastalığı
Sınav koşullarının olumsuzluğu
Puanlamada dikkatsizlik

1941-1942

1941-1942 (1941-1942) 1941-1942

1941-1942

1941-1942 (1941-1942) 1941-1942

ÖLÇME ARAÇLARINDA BULUNMASI GEREKEN ÖZELLİKLER

GEÇERLİK

Kapsam Geçerliği
Ölçüt Geçerliği
Yapı Geçerliği
Görünüş Geçerliği

GÜVENİRLİK

Korunluluk
Tutarlılık
Duyarlılık
Objektiflik

KULLANIŞLIK

1. GEÇERLİK

Bir ölçme aracının ölçmeyi amaçladığı özellikleri başka değişkenlerle ilişkilendirilmeden ölçebilmesidir.

Ölçme aracı neyi ölçmek için hazırlanmışsa onu ölçmesidir.

Bir aracın geçerliğini, amacını gerçekleştirme ve işe yarama durumu belirler.

ÖN: Bir eğitim süreci sonunda sınav yapmanın amacı; o süreçte kazandırılmaya çalışılan davranışların kazanılıp kazanılmadığını ortaya koymaktır. Aras sadece o dönem içinde kazandırılan davranışları ölçüyorsa amacını gerçekleştiriyor demektir.

ÖN: Bir çay içmek için elimizde 5 TL var. Ancak paranız sahteyse para çay içme amacına hizmet etmez. Bu durumda elimizdeki para geçerli olmaz.

Geçerliğin Temel Özellikleri

- Bir ölçme aracı geçerli olmak için güvenilir olmak zorundadır.
- Geçerlik daima belli bir amaç için geçerlidir.
- Bir testin geçerliği bir derece sorundur.

Kapsam Geçerliği

En önemli geçerlik türüdür.
 Bir ölçme aracının ölçülme istenen davranışları ne derece kapsadığının belirlenmesidir.
 Testteki toplam modeller ölçülecek davranışları ve kavı içerdiğini ne derece kapsıyor?
 Testteki her bir madde ölçmek istediği davranışı ne derece ölçüyor?
 Uyarı! Kapsam geçerliğini belirleyen temel kriter, ilişkiler, önemli veya önemli olsun o kavı ile ilgili mutlaka kavı olmalıdır.
 Öğretmene göre en az 1 soru sorulmalıdır.

Ölçüt Geçerliği

Ölçüt ve test puanları arasındaki ilişkiyi gösterir.
 Bir testin geçerliği, o testten elde edilen puanların bir ölçüt ya da ölçütlerden elde edilen puanlar ile ilişkisine bakılarak belirlenir.

İşin Geçerliği (Naradana)

Daha önce elde edilmiş puanlarla karşılaştırıp oradaki korelasyona bakılır.
 Örneğin: TYT puanları için LGS sınavı puanları arasındaki korelasyona bakılır.

Naradana Geçerliği

İstatistiksel ilişkilerle gelecekte yapılacak tahminde bulunma işi.
 Örneğin: LGS puanı kullanılarak lise okul başarısının tahmini yapılması.

Yapı Geçerliği

Testteki modellerin birbiriyle ve testi oluşturan modellerin bütünü ile ilişkisinin aynı yapıyı oluşturmamasıdır.
 Testteki modellerin diğerleriyle bir bütün olmasındadır.
 Daha çok psikolojik testler.

Nasıl Anılır?

Test modellerini umutlar halinde daha önce yapılan benzer bir testle karşılaştırılmalıdır.
 Sonuç ortasında korelasyon yapılmalıdır.

Görünüş Geçerliği

Testin, ölçmek istediği özelliği ölçüyor görünmesidir.
 Tüm test ve tek tek sorular için geçerlidir.

Örneğin: Kapsamda kimya testi yapan sorular kimya bilgisini ölçüyorsa görünüş geçerliği vardır.

GecerliĒi Etkileyen Faktörler

- ✓ Güvenirlik, geçerlik için şarttır. Ancak tek başına yeterli değildir.
GüvenirliĒi artıran tüm çalışmalar geçerliĒi de artırır.
- ✓ Sabit ve sistematik hatalar geçerliĒi düşürür.
- ✓ Testin başka özellikleri ölçmesi geçerliĒi düşürür.
- ✓ Davranışların ölçümünde belirtke tabloları kullanılmalıdır.
- ✓ Kapsan geniş olursa geçerliliğe yüksek olur.
- ✓ **Madde Ayırt Etme Gücü:** Testin geçerliĒi orta güçlükte ayırt edici sorularla sağlanır.
- ✓ **Madde Güçlük Düzeyi:** Orta güçlükteki sorular kullanılmalıdır.

- Aynı soruların tekrar sorulması,
- Bilinen kaynaktan hazırlanması,
- Sınavdan önce soru verilmesi,
- Sınavda kopya çekilmesi,
- Cevap şıklarının diziliş,
- Sınav süresinin çok uzun veya kısa olması,

GecerliĒi düşüren durumlardır.

Bu tür durumların önlenmesi de testin geçerliĒini artırır.

✓

✓

✓

✓

✓

✓

✓

GÜVENİRLİK

Ölçme aracının güvenirligi, ölçme aracının ölçmek istediği değişkeni ne derece doğrulukla ölçtüğüyle ya da tesadüfi hatalardan arınıklığıyla ilgilidir.

Ölçme sonucu, tesadüfi hatalardan arınıklığı oranında güvenilirdir.

$$\text{GÜVENİRLİK} = \left[\text{Doğruluk} / \text{Tutarlılık} / \text{Korarılık} / \text{Objektiflik} \right]$$

Doğruluk : Ölçme aracının birimiyle ilgilidir. Birimin küçük ve detaylı olması hatanın da az olması demektir.

Tutarlılık : Ard arda yapılan ölçmelerin aynı veya benzer sonuçlar vermesi. Kısa zamanda benzer sonuçlar.

Korarılık : Yapılan ölçümlerde zaman uzun olur ve sonuçlarda benzer olursa korarılık denir.

Objektiflik : Ölçmeyi kim yapsa yapsın aynı sonuca ulaşılması. En az hata ile puanlama yapılması.

→ Objektiflik Sağlamak İçin

- Cevap anahtarıyla puanlama yapmak
- Bir kaç farklı kişi ile puanlama yapmak
- Ölçme yapan kişinin bir kaç kez yapması
- İsimlerin kapatılarak puanlama yapılması.

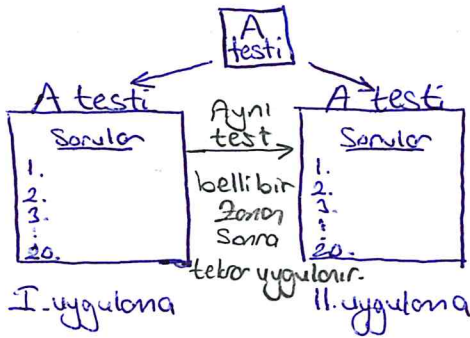
Not : Ölçmenin tutarlılığı 2 şekilde hesaplanır.

"Güvenirlilik Katsayısı" "Ölçmenin Standart Hatası"

"Güvenirlilik Katsayısı Hesaplama"

- | | |
|------------------------------|--------------------------------|
| 1. Test Tekrar Test | 3. İki Yarı (eş değer yarılar) |
| 2. Paralel (eşdeğer) Formlar | 4. İç Tutarlılık (KR-20/21) |

Test-Tekrar Test

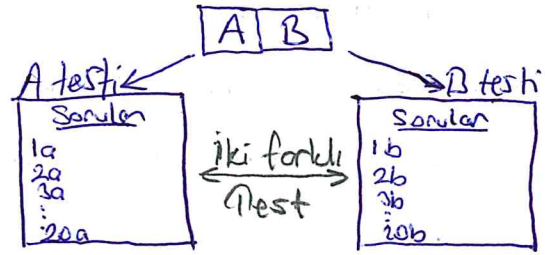


Bir test aynı gruba belli bir zaman aralığıyla iki kez uygulanır.

İki uygulama arasındaki puanların korelasyon katsayıları hesaplanır. (Güvenirlilik Katsayısı)

Not: Aradan uzun zaman geçerek uygulanırsa korelasyon katsayısı olur.

Paralel (Eşdeğer) Formlar



Paralel iki farklı test hesaplanır.

Aynı davranışları ölçen soru sayısı aynı olan benzer iki farklı test öğrencilere aynı anda uygulanır.

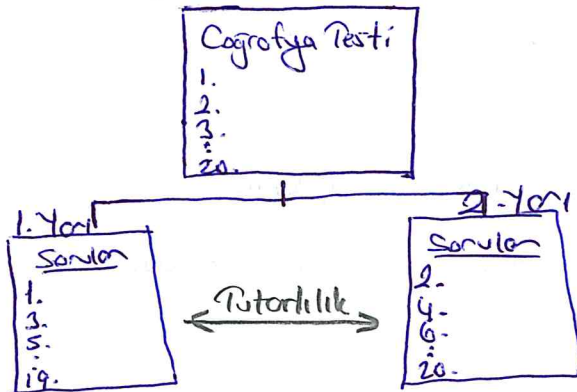
Elde edilen puanlar arasındaki korelasyon katsayısı hesaplanır. (Eşdeğerlik Katsayısı)

İki Yöri (Eşdeğer Yarılar)

Bir test aynı gruba bir kez uygulanır.

Sorular puanlanırken sorular 2 yöriye ayrılır.

Öğrencilerin cevapları iki yöri için ayrı ayrı puanlanır ve hesaplanır.



Not: Testin güvenirlik katsayısı her zaman testin iki yöri katsayısından büyüktür.

Not: Bütünleri için $r = \frac{2 \cdot r_x}{1 + r_x}$ r = testin bütün güvenirliği
 r_x = iki yöri arasındaki korelasyon.

İç Tutorluluk (KR 20/KR 21)

Testin madde güçlük indeksi yapılıyor.
 madde analizi

Testteki her bir maddenin aynı değişkeni ölçtüğü olasılıktan sonra (homajenliği)

Testteki maddelerin birbirleriyle ne derece tutarlı olduğunu hesaplamak için KR 20-21 formülü uygulanır.

Ölçmenin Standart Hatası

Ölçme sonuçlarına değişik kaynaklardan, değişik miktarlarda hata karışır. Aynı kişiye yönelik birden fazla ölçme yapılmalıdır. Bu ölçmeler bir dağılım gösterir. Buna normal dağılım denir. Bu dağılımın standart kayması, ölçmenin standart hatasıdır.

Standart Hata = Standart ve Güvenlik ile
Kayma Katsayısı hesaplanır.

$$\begin{aligned} Se &= \text{Ölçmenin Standart Hatası} \\ S_x &= \text{Standart Sapma} \\ r &= \text{Güvenlik Katsayısı} \end{aligned} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} Se = S_x \sqrt{1-r}$$

ÖRN: Standart sapması 6, güvenlik katsayısı 0,75 olan bir testin standart hatası nedir?

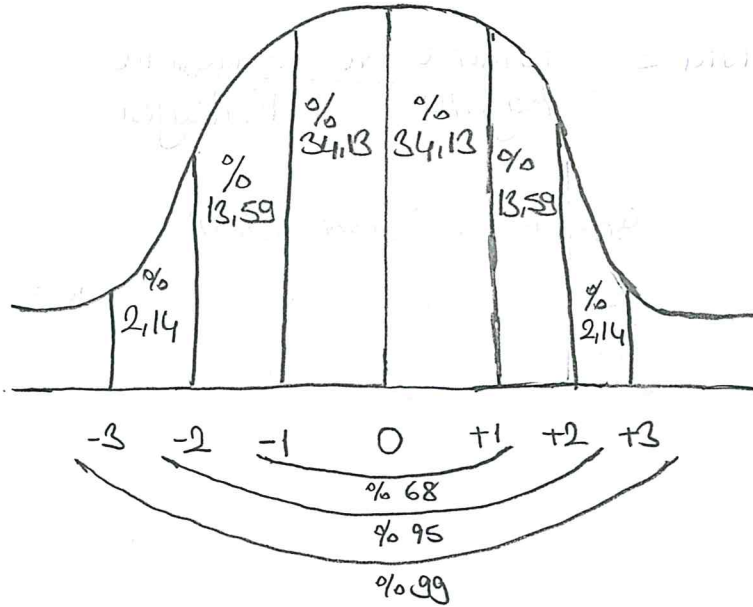
$$Se = S_x \sqrt{1-r} \quad 6\sqrt{1-0,75} \quad 6\sqrt{0,25} \quad 6 \cdot 0,5 \Rightarrow (3)$$

Not: Hatasız ölçme olmaz. Bu nedenle ölçme sonuçları ölçek üzerinde tek bir noktadan ziyade iki nokta arasındaki bir aralık olarak düşünülmelidir.

Güven aralıkları standart hatanın eklenmesi ve çıkarılması, Sonuçta bulunur.

Testlerde Güven Aralıkları

1. % 68 olasılıkla güven aralığı için ; (+1) eklenir ve (-1) çıkarılır
2. % 95 . " güven aralığı için ; (+2) eklenir ve (-2) çıkarılır
3. % 99 " güven aralığı için ; (+3) eklenir ve (-3) çıkarılır



ÖRNEK: Standart sapması 4 güvenlilik katsayısı 0,75 olan bir testten 50 olan bir öğrencinin %68 %95 %99 olasılıkla hata koruma oranı düşünüldüğünde puan aralıkları nelerdir?

Bu soruyu çözmek için öncelikle standart hata bulunur.

$$Se = S \cdot \sqrt{1-r} \quad 4 \cdot \sqrt{1-0,75} \quad 4 \cdot \sqrt{0,25} \quad 4 \cdot (0,5) = (2)$$

50 puan için güven aralıkları

- % 68 olasılıkla (+2/-2) → (52-48)
- % 95 olasılıkla (+4/-4) → (54-46)
- % 99 olasılıkla (+6/-6) → (56-44)

Standart Hata azaldıkça güvenlilik oranı/ Standart sapma küçük ise standart hata da küçük olur

Güvenirlilik Artırma Yolları

1. Soru sayısı arttıkça sınavdan elde edilen toplam puanların güvenirliliği artar. Amaç: Tesadüfi hataların ölçmedeki etkisini azaltmak.
2. Soruların objektif olması gerekir. Sınavda kullanılan soruların öğrenciler tarafından açıkça anlaşılır olması ve tesadüfi cevaplar verilmemesi elde edilen puanların güvenirliliğini artırır.
3. Öğrencilerin sınava motive edilmesi, soruları dikkatlice cevaplamalarını sağlar ve güvenirlilik artırır.
4. Tüm öğrencilerin soruları okuyarak cevaplayacakları kadar süre verilmesi güvenirliliği artırır.
5. Sınav sonuçları objektif olarak puanlanmalı.
6. Sınav gözetmeninin tutumu ortamı etkiler.
7. Ölçme aracının duyarlı olması güvenirliliği artırır.

Güvenirlilik - Geçerlik Arasındaki İlişki

1. Güvenirlilik, geçerlik için ön koşuldur ancak tek başına yeterli değildir.
2. Güvenirlilik düşerse geçerlikte düşer ancak güvenirlilik artarken geçerlik artmayabilir.
3. Güvenirliliği etkileyen durumlar doğrudan veya dolaylı geçerliği de etkiler.
4. Bir testin geçerliği, güvenirlilik kökükünden fazla olamaz.

KULLANILILIK

Bir ölçme aracının; hazırlanmasının
 dağıtılmasının
 cevaplanmasının
 puanlanmasının } Kolay olmasıyla ilgilidir
 "Test, ilgili herkese
 kolaylık sağlanacak"

Ölçme aracının kullanışlı olması "güvenirliği ve geçerliği" yükseltir.

- Yazılı veya sözlü testlerin hazırlanması, dağıtılması kolay ve ekonomik olmasına karşın; cevaplanması ve puanlanması zordur.
- Çoktan seçmeli, Boşluk doldurma, Eşleştirme soruları hazırlanması, dağıtılması zor ama cevaplanması ve puanlanması kolaydır.
- Doğru - yanlış testleri hazırlanması, dağıtılması, cevaplanması ve puanlanması kolaydır. (En kolay kullanışlı testlerdir)
- Kontrol listeleri, Gözlem formları hazırlanması zor puanlama kolaydır.
- Portfolyolar kullanışlığı en düşük araçlardır.

SINAVLAR VE TESTLER

Bir Test Planı Nasıl Olmalıdır?

- Testin amacı belirlenir
- Hedef davranışlar belirlenir
- Belirtke Tablosu hazırlanır
- Kullanılacak soru türü belirlenir
- Sorular yazılır
- Sorular test düzenine sokulur

Test Planlarken Nelere Dikkat Edilmeli?

- Sorular öğrenci seviyesine göre olmalı
- Açık, anlaşılır ve net olmalı
- Yazım ve imla kuralına uygun olmalı
- Ezber bilgi sormaktan kaçınılmalı
- Ayrıntı içeren sorulardan kaçınılmalı
- Sınav hakkında bilgiler içeren sınav yönergesi kullanılmalı
- Bir sorunun cevabı başka soruda olmamalı

A. TESTLERİN GRUPLANDIRILMASI

1. Uygulama Zamanının Uzunluğuna Göre
2. Hazırlanma Biçimine Göre

B. EĞİTİMDE KULLANILAN TESTLER

1. Yazılı Yoklamalar
2. Kısa Cevaplı (Başlık Doldurma) Testleri
3. Doğru-Yanlış Testleri
4. Çoktan Seçmeli Testler
5. Eşleştirmeli Testler
6. Sözlü Yoklamalar

C. YENİ ÖLÇME-DEĞERLENDİRME

YAKLAŞIMLARI

1. Rubrik Puanlama Anahteri
2. Portfolyo Değerlendirme
3. Performans Değerlendirme
4. Gözlem Formları
5. Tonlayıcı Doldurmuş Ağaç (TDA)
6. Yapılandırılmış Grid
7. Kelime İlişkileştirme Testi (KIT)
8. Proje
9. Problem Gözme
10. Kavran Haritası
11. Görüşme Teknikleri
12. Öz Değerlendirme
13. Akron Değerlendirme
14. Grup Değerlendirme

A. TESTLERİN GRUPLANDIRILMASI

1. Uygulama Zamanına Göre Testler

Tür	Özellik
Hız Testleri	<p>Tanınım süresi sınırlıdır. Süre sınırlı tutulmazsa tüm öğrenciler soruları çözebilir.</p> <p>Rahatlıkla cevaplamak kolaylıkta sorular olur.</p> <p>Zamanı iyi kullanan ve test çözmede deneyimli öğrenciler daha başarılı olurlar.</p> <p>ÖRNEK: Özel yetenek, beden, müzik, resim vb. sınavları</p>
Güç Testleri	<p>Başarı düzeyi belirlemek için kullanılır.</p> <p>Verilen süre soruların cevaplanmasına yeter.</p> <p>Çalışan ve çalışmayan öğrencileri ayırt eder.</p> <p>Zorluk düzeyi yüksek testlerdir.</p> <p>ÖRNEK: TYT, AYT, KPSS vb. sınavlar</p>

2. Hazırlama Biçimine Göre Testler

Tür	Özellik
Öğretmen Yapımı Testler vize, final Quiz	<p>Öğretmen kendi sınıfına yönelik hazırlar. Tek bir gruba bir defa kullanılır.</p> <p>Soru sayısı az, grup küçüktür. Her türlü soru tipi bulunur.</p> <p>Geçerlik-Güvenirlilik düşük / Sorularla istatistik-opsiyon yapılmaz.</p>
Standart Testler TYT, KPSS TUS, LGS	<p>Komisyon veya uzmanlar hazırlar. Geniş gruplara yapılır.</p> <p>Uzun zamanla iltiyas vardır.</p> <p>Gökten seçmeli, güvenilir ve geçerli testlerdir.</p> <p>Toplanan verilerle istatistikî işlemler yapılabilir.</p>

B. EĞİTİMDE KULLANILAN TESTLER

1. YAZILI SORULAR

Öğrenciler, soruların cevaplarını belli bir süre içinde düşünür ve yazarlar. Çoktan seçmeli testlerle birlikte en fazla kullanılan türdür.

Genel izlenimle puanlama: Bir öğrencinin kağıdındaki bütün sorular okunur, toplam bir puan verilir.

Sıralama yöntemi: Bütün öğrencilerin kağıtları okunur ve en iyiden en kötüye sıraya koyulur. En yüksekten en düşüğe puan verilir.

Sınıflama yöntemi: Bütün kağıtlar okunur ve sınıflandırılır. "iyi - orta - zayıf" şeklinde. Tekrar her gruptaki bütün kağıtlar kendi içinde iyi, orta, zayıf diye ayrılır.

Daha sonra en iyi kağıtlara 9, 8, 7 diye p. verilir.

Anahtarla puanlama: Test uygulanmadan önce bütün soruların olası cevapları listelenerek puanlama anahtarı oluşturulur.

Ne zaman kullanılır?

- Soru hazırlamak için yeterli zaman yoksa
- Dil ve imla bilgileri ölçülmek istenirse
- Üst düzey becerilerin ölçülmesinde
- Yaratıcı, eleştirel düşünme, problem çözme becerileri ölçülmesinde
- Öğrenci sayısının az olduğu zaman.

Yararı Nelerdir?

- Konuyu kavrayıp kavranmadıkları belirlenir.
- Kavrayışları ve elistik öğrenmeleri belirlenir.
- Soru hazırlamak fazla zaman almaz.
- Özgün ve yaratıcı fikirleri ortaya çıkarır.

Sınırlılıkları Nelerdir?

- Cevaplama ve puanlama uzun sürer.
- Puanlamada objektiflik zordur.
- Soru az, geçerlik - güvenirlik düşüktür.
- Sadece bilgi değil, imla ve yazı güzellikleri de puanlamayı etkiler.
- Cevapların sınırı zordur. Yanlış olasıdır.

2. KISA CEVAPLI (BOŞLUK DOLDURMA) TESTLER

Cümle tamamlanarak, kısa tanım yapmak, sorudaki bir boşluğu doldurmak için sembol, rakam veya kelime yazılması.

Önemli İlkeleri

- Soru sayısı çok olmalıdır.
- Ezber yoklayan sorular sorulmamalı.
- Soruların içinde ipuçları olmalıdır.
- Her soruda tek bilgi sorulmalı.
- Yorumla açık bilgi bulunmalıdır.
- Kitaptaki soruların aynısı sorulmamalı.

Yararları nelerdir?

- Gök soru ver, kapsam geçerliği yüksektir.
- Her eğitim kademesinde kullanılır, objektif olarak puanlanabilir.
- Bilgilerin hatırlanıp hatırlanmadığını öker.
- Cevaplamak ve puanlamak kolaydır.
- Yanlış kavramları ortaya çıkarır.

Sınırlılıkları nelerdir?

- Kitaptan alınan sorular ezbere götürür.
- Bilgi ve kavrama ölçtüğü için geçerliği düşüktür.
- Öğretmen tecrübesiz ise uygulama düzeyinde soru hatırlanmaz.

3. DOĞRU-YANLIŞ TÜRÜ TESTLER

Doğru-yanlış önermelerinin yer olduğu testler. İki seçenek vardır. Soru basur, yüksek (%50) Dışeltme formülü ile geçerlik, güvenilirlik artar.

Önemli İlkeleri

- Yorumla açık sorular yer almalıdır.
- Olumsuz ifadelerin kullanımından kaçınılmalıdır.
- İpuucu olacak vurgulanmalar yapılmalıdır.
- Kitaptan yer alan benzer cümleler alınmalıdır.
- Doğru ve yanlış soruları yakın olmalıdır.
- Ayrıntı ve önemsiz bilgiler yoklanmalıdır.
- Önemli noktaya vurgu yapılmalıdır.

Yararları nelerdir?

- Hatırlamak, puanlamak, cevaplamak kolaydır.
- Eğitimin her aşamasında kullanılır.
- Gök soru olacağı için kapsam geçerliği ve duyarlılık yüksektir.
- Objektif puanlama yapılır.

Sınırlılıkları nelerdir?

- Soru faktörü %50, geçerlik düşüktür.
- Yanlış bilgileri belirlemez.
- Üst düzey becerileri ortaya çıkarmaz.

4. ÇOKTAN SEÇMELİ TESTLER

Bir soru kökü ve alternatif seçeneklerden oluşan testlerdir.

Madde kökü: Problemin bulunduğu bölüm

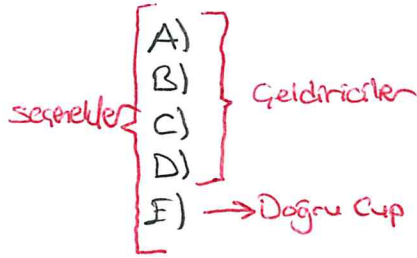
Test maddesi: Testi oluşturan her bir soru

Seçenekler: Sorunun muhtemel cevapları

Doğru cevap: Cevaplanması beklenen seçenek

Geldirici: Doğru cevabın alternatifleri. Konyu çalışmayan öğrencileri eler.

ÖZNE: Soru? (madde kökü)



Önemli İlkeler

- Sıras faktörünü en aza indirmek için seçenek sayısı artırılmalıdır.
- Düzeltme formülü uygulanmak gerekir. (44.10 gösterir)
- Soruların güçlük düzeyleri 0,50 civarında olmalı.
- Olumsuz soru kökü fazla kullanılmamalı
- Sorular ayrıntı değil ana temayı yansıtmalı.
- Hepsisi, Hiçbiri fazla kullanılmamalı
- Aynı köke bağlı çok fazla soru olmamalı.

Yararları nelerdir?

- Soru sayısı fazladır. Kapsamı ve doğruluğu yüksektir.
- Bütün bilgileri ölçer.
- Cevaplama ve puanlama kolaydır.
- Objektif ve kullanışlıdır.
- Her aşamada kullanılabilir.
- Sonuçlarla birçok istatistiksel işlem yapılabilir.

Sınırlılıkları nelerdir?

- Soru okuma ve oluma hızları farklı olduğundan cevaplamada etkilidir.
- Farklı bilgilerin kaynağı belirlenemez.
- Özel bilgi ve beceri gerektirdiğinden hazırlanması uzun sürer.
- Soruların güçlük düzeyini ayarlamak zordur.
- Yaratıcı ve eleştirel düşünme becerilerini ortaya çıkarmaz.
- Kendini ifade etme becerisini geliştirmez.

5. EŞLEŞTİRMELİ TESTLER

Öğrenciler, iki aynı sütunda verilen bilgileri en mantıklı şekilde eşleştirir.

Birinci sütunda soru kökleri, diğerinde ise muhtemel cevaplar vardır.

ÖRNEK:

Coğrafi Bölgeler	Şehirler
(B) Marmara Bölgesi	A. Nevşehir
() Karadeniz Bölgesi	B. Balıkesir
() Akdeniz Bölgesi	C. Mersin
() İç Anad. Bölgesi	D. Muğla
() Doğu Anad. Bölgesi	E. Kars
	F. Batman
	G. Samsun

Önemli İlteler

- Soru kökü ile muhtemel seçeneklerin sayıları farklı olmalı
- Soru kökü en fazla 15 tane olmalı
Soru-cevap oranı 5/7 olmalı
- Soru kökleri cevaplardan ayrı olmalı
- Cevaplar rakamlardan oluşuyorsa büyüklük sırasına konmalı
- Her iki sütunda yer alan bilgiler birbirleriyle tutarlı olmalı

6. SÖZLÜ YOKLAMALAR

Öğrenciye, bir konuda sözel olarak yöneltilen soruların sözlü olarak cevaplandırılması

Önemli İlteler

- Tüm öğrencilere sorulacak soruların güçlük düzeyleri aynı olmalı
- Önceden cevap modelini oluşturmak objektifliği ve güvenilirliği artırır.
- Kısa cevaplı çok soru sorulması geçerlik ve güvenilirliği yükseltir.
- Sınav kaydedilmeli, birden fazla puanlayıcı kullanılmalı.

Yorantı nelerdir?

- En hızlı dönütü sağlar. Üst düzey bilgi, yaratıcılık, hızlı cevaplılık sağlar.
- Kavram yanlışlığı önünde betirler.
- Küçük yaş çocuklar için daha uygundur.
- Sözel ifade yeteneği, konuşma becerisi geliştirir.

Not: Yaratıcılığı ve yanlış kavramları ortaya çıkaran fakat soru sayısı azdır.

Özellikler	Yazılı	Bosluk Doldurma	Doğru-Yanlış	Çoktan Seçmeli	Eşleştirmeli	Sözlü
Yoklama Bilgi Düzeyi	Üst düzey	Alt düzey	Alt düzey	Her düzey	Her düzey	Üst düzey
Kapsamı	Düşük	Yüksek	Yüksek	Yüksek	Yüksek	Düşük
Hazırlama	Kısa zaman alır	Kısa zaman alır	Kısa zaman alır	Uzun zaman alır	Uzun zaman alır	Kısa zaman alır
Puanlama	Uzun zaman alır Subjektiftir	Kısa zaman alır Objektiftir	Kısa zaman alır Objektiftir	Kısa zaman alır Objektiftir	Kısa zaman alır Objektiftir	Hemen yapılır Subjektiftir
Güvenirlilik	Düşük	Yüksek	Yüksek	Yüksek	Yüksek	Düşük
En Önemli Yararı	Yanlış kavramları yaratıcılığı ortaya çıkarır	Yanlış kavramları ortaya çıkarır.	Kullanıcılığı en yüksek testtir.	Her düzeyde kullanılabilir ve kapsamı iyi örnekler	Kapsam geçerliği oldukça yüksek	Sürekli çalışmaya yönlendirir.
En Büyük Sakıncası	Kapsam geçerliği düşük. Puanlama öznelir.	Öğrencileri ezere yönlendirir.	Sıfır boş soruları yüksektir (%50)	Yaratıcı ve eleştirel düşünce yoklamaz.	Görüşme soru sayısı çoktur.	Kapsamı düşük Puanlama öznelir.

1. PORTFOLYO (ÖĞRENCİ ETKİNLİK DOSYALARI)

Portfolyo; öğrenciye ve başkalarına, öğrencinin bir ya da daha fazla alandaki başarılarını sunmak için, amaçlı olarak öğrenci çalışmalarının toplanmasıdır. Öğrencinin nasıl düşündüğünü, nasıl çalıştığını, ne kadar öğrendiğini gösteren dosyalardır.

Amaç; öğrenciyi tüm özellikleriyle bir bütün olarak tanımak ve öğrencinin kendini tanımasını sağlamak.

Öğrenci bu süreçte bizzat sorumluluk alır.

Başarısı diğerleriyle paylaşılır. Çünkü öğrenciye özgü kişisel gelişim araçlarıdır.

Öğrencinin çalışmalarını, çalışmalarını ve gelişimini, velisinin, öğretmenin ve diğer ilgililerin izlenmesini sağlar.

Portfolyoların Kullanım Amacı ve Temel İlkeler

- Öğrencinin gelişim sürecini izlemek
- Öğrencide sorumluluk duygusu geliştirmek
- Öğrencinin toplumla çalışma yeteneğini geliştirmek
- Öğrencinin öğrenmelerini süreç içinde ve düzenli izlemek
- Öğrencileri ilgi alanı ve yetenekleri yönünden tanımak
- Öğrencide öğrenme sorumluluğu ve özdisiplin sağlamak
- Öğretimi, öğrenen merkezi duruma getirmek
- Aile ve öğrenci arasındaki iletişimi güçlendirmek
- Öğrencinin öğretim dışındaki etkinliklerini de kapsamak
- Öğrenci ortadoğasının gelişimini de izlen değerlendirme fırsatı bulur.

Geleneksel Değerlendirme

- Öğrencinin yeteneği tek bir zamanda ölçülür.
- Öğrenci ölçütlerden habersizdir. Değerlendirmeyi öğretmeni yapar.
- Öğrencileri grup içinde değerlendirir.
- Öğrencinin dil yeteneğini geliştirmez.
- Öğrenciye fazla sorumluluk vermez.
- Değerlendirmeler, öğrenme sürecinden bağımsızdır.

Portfolyo Değerlendirme

- Öğrenci yeteneği süreç içinde ölçülür.
- Öğrenci ölçütlerden haberdardır, değerlendirme öğrenci de katılır.
- Öğrenci bireysel değerlendirilir.
- Dil öğrenme performansı yakalanır.
- Öğrenci aktiftir, sorumluluk alır.
- Öğrenme ve değerlendirme sürecini birleştirir.

Portfolyoların Genel Amaçları

1. Öğrencilerin öğrenmelerini; süreç içinde sağlıklı izlemek
2. Öğrencilerde sorumluluk olmayı ve öz disiplini geliştirmek
3. Öğrencinin yaratıcılığını geliştirmek
4. Öğrencileri gerçekçi koşullarda değerlendirmek
5. Öğrenme sürecini ve değerlendirmeyi birleştirmek
6. Öğrencilerin ilgi alanlarını keşfetme ve yeteneklerini sergilemeye yardımcı olmak
7. Sadece sınav kağıdına bağlı kalmadan öğretim süreci içinde öğrenci gelişimini takip etmek.
8. Bir dönem içinde öğrencinin yaşadığı zamanı, eksikliklerini, düzeltmelerini ve performansını tümüyle değerlendirmek.

2. PERFORMANS DEĞERLENDİRME

Performans değerlendirme, öğrencinin sahip olduğu becerilerin gözlenemediği durumlarda; öğrencinin yaptığı görevlerden veya davranışlardan ölçme sonuçlarını değerlendirmeye yarayan süreçtir.

Öğretim sonuçlarını, çalışma veya uygulama sırasında doğrudan doğruya ölçmeye yarayan değerlendirmedir.

Öğrencilerin gerçek yaşam problemlerine akademik bilgilerini uygulayabilme ve bunu problem üzerinde gösterebilmeleri ile ilgilidir.

Eksikleri teşhis eder, tekrar öğretme, yanlışları düzeltme yönünden önemli bir uygulamadır.

Per-formans Değ. Aşamaları

1. Amacın belirlenmesi
2. Performans ölçütlerinin belirlenmesi
3. Performansın gözlenileceği ortamı oluşturma
4. Performansın puanlanması

Per-formans Değerlendirmede Dikkat Edilecek Hususlar

- Ödev, birçok beceriyi kapsamalıdır.
- Ödev hem ürüne hem de sürece odaklanmalıdır.
- Öğrencilere ödevi tamamlamak için verilen süre ödevin niteliğine göre ayarlanmalıdır.
- Ödevlerin bazıları bireysel, bazıları ise grup ödevi şeklinde verilmelidir.

NOT: Performans değerlendirmek için ; performans ödevleri ve projeler kullanılır.
Öğrenci ürün dosyaları

3. RUBRİK PLANLAMA ANAHTARI (PLANLAMA YÖNERGESİ)

Performans değerlendirmede öğrencileri puanlama anahtarı ile kullanılan ölçme aracıdır.

Performansı tanımlayan ve herhangi bir gelişmenin puanlanması için geliştirilmiş ölçütleri içeren araçtır.

Bir ölçme veya etkinliğin tanınmıyada bölümlerinin nasıl puanlanacağını ana hatlarıyla gösterir.

Kabul edilir veya edilenez performans sınırlarının öğretmen ve öğrenciler için açıkça belirtildiği kriterlerin olduğu araçtır.

Not: Bu ölçekler kullanılırken öğrencilerden beklenen performansın iyi tanımlanması ve sınırların iyi çizilmesi gerekir.

Özellikleri

- 1- Değerlendirmeye hem öğretmen hem de öğrenci katılır
- 2- Öğretmen, değerlendirmede kullanacağı ölçütleri belirler
- 3- Öğrenci, değerlendirme ölçütlerini bilir ve etkinliklerini buna göre düzenler
- 4- Öğrenci, nasıl değerlendirildiğini görür. (Değerlendirme şeffaftır)
- 5- Değerlendirme, sürece ve sonuca yayılır.

A. Analitik Rubric (Süreç)

Bir değerlendirmenin farklı bölümlerindeki adı belirli cevapları puanlamada kullanılır.

Değerlendirmenin her bölümündeki puanlar toplanır, öğrencinin toplam puanı belirlenir.

Örn: Doğru cevap 45 gr olan bir deneyde
44-46 gr bulursa 3P.
42-43/47-48 " 2P.
41-42-48+ gr " 1P.

B. Holistik Rubric (Sonuç)

Öğretmen, öğrencilerin cevaplarını toplan olarak değerlendirmek isterse kullanılır.

Bosanak veya bölümlerden çok toplan performansta, sonuçla ilgilidir.

Kriterleri

Üstün - Çok iyi - 3P
Kabul edilebilir - 2P
Geçersiz - 1P
Kabul edilenez - 0P

4. KONTROL LİSTELERİ

Gözlenen performansın, belirtilen performans ölçütlerine ne derece uyumlu olduğunu listeye belirtilmesidir.

Öğrenciden beklenen davranışın özellikleriyle ilgili bilgiler için ve öğrenci performansının eksik noktalarını belirlemek amacıyla kullanılır.

Var/Yok Evet/Hayır 0/1 şeklinde puanlama yapılır.

Örnek: (Bedensel dersinde kurallara uygun tutulabilir)

Kontrol Listesi		
Adı Soyadı:		
Sınıf		
No		
Hedef: Tokla olma	Var	Yok
1. Isınma		
2. Hazırlık		
3. Tokluğu olma		
4. Selam verme		

5. GÖZLEM FORMLARI

Gözetimler, öğrenciler hakkında doğru ve çabuk bilgiler sağlar.

Öğretmen, öğrencilerin ;

Soru ve önerilere verilen cevaplarını,

Sınıf içi tartışmalara katılmalarını

Grupla çalışma ve tartışmalara katılmalarını

Öğrencinin yaptığı görevler ve materyallere gösterdiği tepkileri

} gözlemler.

6. DERECELENDİRME ÖLÇEKLERİ

Derecelendirme ölçekleri, ölçülen özelliğe ilişkin performans çeşitli düzeyleriyle tanımlar ve ölçütlerin ne dereceye kadar karşılandığını görmeye olanak sağlar.

Kontrol listeleri belli davranışların yerine getirilip getirilmediğini belirlemeden öte gidemez. Derecelendirme ölçekleri ise ölçülen özelliğe ilişkin performans çeşitli düzeyleriyle tanımlayabilir ve detaylı ölçütlerin ne dereceye kadar karşılandığını gösterir.

Öğrencinin Adı Soyadı :		Tarih :			
ÖLÇÜTLER	Çok İyi (4)	İyi (3)	Orta (2)	Zayıf (1)	
1. Dinleyicileri göz teması kuruyor					
2. Beden dilini kullanıyor					
3. Anlatılır tonla konuşuyor					
4. Vurguları yerinde yapıyor					
5. Akıcı konuşuyor					
6. Gereksiz cümle kurmuyor					
7. Düzensiz ifadeler kuruyor					
8. Gereksiz tekrar yapmıyor					

(Sözlü Sunum Becerisine Yönelik Derecelendirme Ölçeği)

7. KELİME İLİŞKİLENDİRME TESTLERİ (KIT)

Öğrenci bu teknikte, belli bir süre içinde (30 sn) herhangi bir konu ile ilgili verilen bir anahtar kavramın çağrıştırdığı kelimeleri cevap olarak verir.

Öğrencinin hafızasından, bir kavrama verdiği sıralı cevapların bilişsel yapıdaki kavramlar arasındaki bağlantıyı gösterdiği farz edilir.

8. OTANTİK DEĞERLENDİRME

Yapılan öğrenmelerin gerçek hayat durumları ile bağlantılandırılması ve becerilerinin değerlendirilmesini kapsar.

1. Uygulamaları vurgular : Bir öğrencinin ne bildiğinin yanı sıra bilgisini kullanıp kullanmadığını da değerlendirir.

2. Doğrudan değerlendirmeye odaklanır : Belirlenmiş hedefi direkt ölçer.

3. Gerçekçi problemleri kullanır : Ödevleri, günlük yaşamın bir parçası ve gerçekçi yolla oluştururlar.

4. Açık uçlu düşünmeye teşvik eder : Birden fazla doğru yanıtı, yanıtı birden fazla yolla ifade etmeye teşvik eder.

9. YAPILANDIRILMIŞ GRID

En önemli amacı; öğrencilerin bilgi seviyesini, eksiklerini ve kavram yanlışlarını tespit etmektir.

Bu teknikte öğrencinin seviyesine uygun olarak 9, 12, 16 kutucuktan oluşan bir tablo hazırlanır ve tabloya her kutucuk numaralandırılır. Öğretmen, konuyla ilgili bir soru hazırlar ve sorunun yanıtını veya yanıtlarını, rastgele kutucuklara yerleştirir.

1. Balina	4. Kelebek	7. Yengeç
2. Soğan	5. Yılan	8. Örümcek
3. Solucan	6. Kartal	9. Kurbağa

S.1. Kutucuklardan hangisi meneli hayvanlar grubunda yer alır?

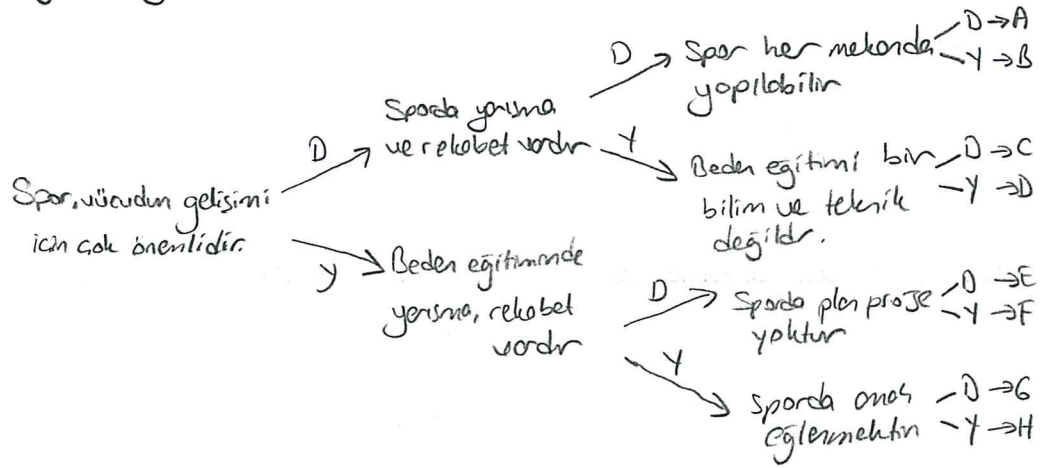
Cevap : (1)

S.2. Hangileri hem karada hem suda hareket edebilir.

Cevap : (5, 7, 9)

10. TANILAYICI DALLANMIŞ AĞAÇ (TDA)

Birbiriyle bağlantılı DİY tipindeki soruları içeren TDA tekniğinde; her bir DİY kararı bir sonraki DİY kararı etkileyen veya belirleyen sonuçlar içerir.



11. PROJE YÖNTEMİ

- Öğrencilerin bireysel veya grup olarak istedikleri bir konu veya alanda inceleme, araştırma ve yorum yapma; yeni bilgilere ulaşma,
- Özgün düşünce üretme ve çalışmalarında bulunmaları amacıyla ders öğretmeni rehberliğinde yapacakları çalışmalar.
- Geliştirme süreci; uzun, karmaşık ve zorludur.
- Geniş içerikli ve uzun süreli performans ödevleridir.
- Ünitelerdeki kazanımları kapsayan ayrıntılı ödevlerdir.
- Proje konusu; öğrenci tarafından veya öğretmenin belirlediği.
- listeden seçme yoluyla belirlenebilir.
- Öğrenci, projenin amacı, izlenecek yolları, kullanılacak materyaller önceden planlar

12. KAVRAM HARİTALARI

Kavram Haritaları; bilgiyi organize etmek ve sunmak amacıyla yapılmış grafiksel araçlardır.

Bu araçlar daire ya da bir çeşit kutucuklar içine yazılmış olan kavramları içerir.

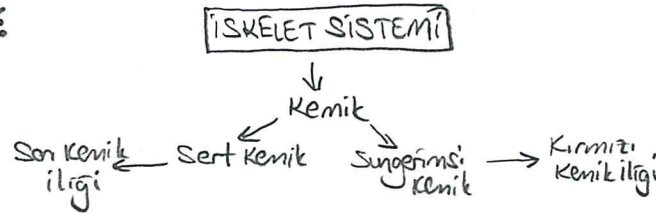
Kavram haritalarında; iki kavram arasındaki ilişki belirtilmektedir.

Kavram Haritaları :

Bir konunun öğretiminde / Öğrenmeyi kolaylaştırmada
 Öğrenme sürecini kontrol etmede / Değerlendirme yapmada
 Kavram yapılarını ortaya çıkarmada

} kullanılır

Örnek:



13. GÖRÜŞME TEKNİĞİ

Görüşme; Öğretmenin öğrenciye belli bir öğrenme konusu hakkında bir takım sorular sorması ve sorulara karşılık sözlü cevaplar almasıdır.

Amaçta ulaşabilmesi için öğrenciler kendini rahat hissetmeli ve gerçekçi bilgiler vermelidir.

Görüşme sırasında yazılı, sesli veya görsel kayıtlar alınabilir.

Yazılı kayıtlar yüzeysel bilgiler verirken, sesli ve görsel kayıtlar daha detaylı bilgiler sağlar.

14. ÖZ DEĞERLENDİRME

Bireysel veya kendini değerlendirme olarak da adlandırılır.

Öğrenciler; kendi öğrenme süreçlerini, özellikle başarı düzeylerini ve öğrenme sonuçlarını yargırlarlar.

Temel amacı, öğrencilerin öz değerlendirme becerilerini geliştirmek. Değerlendirme sürecinde öğrenci; öğrenme tarzını, güçlü ve zayıf yönlerini, becerilerini eleştirme - değerlendirme imkanı kazanır.

ÖRNEK:

KAZANIMLAR	😊	😐	😞
1. Kendime güvenerek konuşurum	✓		
2. Kelimeleri anlamlarına uygun kullanırım	✓		
3. Topluluk önünde konuşabilirim	✓		
4. Kitaplarımı düzenli kullanırım	✓		
5. Okuduklarımdan konusunu belirtirim	✓		
6. Önemli gördüğüm yerlerin altını çizirim		✓	
7. Sayfa temizliğine dikkat ederim	✓		

15. AKRAN DEĞERLENDİRME

Bir grup içinde yer alan bireylerin davranışlarını değerlendirmeleridir.

Kişinin öğretmen dışında başka birinden de değer alınmasına yardımcı olur.

Yanlış davranışlar olabilir, arkadaşları olan kişiye daha yalınca puan verilebilir.

Sonuçları öz değerlendirme becerilerini geliştirmek de kullanılabilir.

ÖRNEK:

	1. Arkadaşıma göre ben	2. Arkadaşıma göre ben	3. Arkadaşıma göre ben	Bana göre ben
1. Çalışmalara gönüllü katılır.				
2. Bildiklerini arkadaşlarıyla paylaşıp				
3. Gerekliğinde arkadaşlarına yardım eder.				
4. Arkadaşlarının görüşlerine saygılıdır.				
5. Tartışmalarda kınacı olmadan konuşur.				

"Ölne- Değerlendirme Araç ve Yöntemleri"	Nedir?	Yararları nelerdir, Ne amaçla kullanılır?	Sınırlılıkları nelerdir?	Afyon!
"Ö2 Değerlendirme"	Bireyin; belli bir konuda, belli özellikler çerçevesinde kendini değerlendirmesi	"öğrenciler kendi öğrenme süreç- lerine aktif katılırlar kendilerini sürecin parçası olarak görürler Kendi güçlü ve zayıf yönlerini tanıma imkanı bulurlar	"öğrenciler, kendilerini değeri- lendirdikten sonra davranış- bilgiler	"Ö2, grup ve akron değerlendirme so- nucuları not vermek için kullanılır"
Akron Değerlendirme	"öğrencilerin, belli özellikler çerçevesinde birbirlerini değerlendirmeleri"	Eğitimsel düşünme becerile- rinin gelişmesine katkı sağlar. Eğitimsel kavramı hoşgörülü olma özelliğini geliştirir.	Doğruluk fazla olur. Yakın ortadastra ⊕ notlar, uzak ortadastra ⊖ notlar verilebilir.	Çalışmaya uygun doğru beceriler mi? derece puanları özelliklerinin ve derece ölçeklerinin kullanılması önerilir onca;
Performans "Ödevleri"	"öğrencilerin sahip oldukları bilgi ve becerileri" günlük yaşamları ilişkileri direnerek ortaya koymaları, gerektiren kısır süreçli çalışmalar	öğrencilerin, bilgiyi kullon- malarını ve gerçek yaşam durumlarına yakın ürünler ortaya koymalarını sağlar.	Uzun süreli ve esnek yapıl- dığı için sadece öğrenciye ait olmayabilir "öğrencinin davranışları göre hazırlanmış da gel- dir."	bu evlerde birlikt kavram, intelli, öğren formları öze ve akron değerlendirme formlarını kulland- masında fayda vardır.
Proje "Ödevleri"	Geniş içerikli ve uzun süreli performans ödevleridir.	Belirli bilmelere odak- landığında öğretimin her bir parçası gözlemlenebilir	olmayan sonuçlar çıkar.	

"Öğrenci Ürün Dosyası (Portfolio)"	"Öğrencinin bir öğrenme alanındaki çaba, başarı ve gelişimini gösteren, bir dönemdeki çalışmalarını, organize ettiği dosyaları."	"Öğrencinin zaman içindeki gelişimini bir çizelgeyle izlenebilir. Üst düzey becerileri geliştirir."	"İçerik önceden belirlenmez. Nitelikli veya niteliksiz birer çalışmaya içerebilir. Kolaylık sunu. Herda dosyanın sürdürülmesi zorundur. Evde bulunanlar için öğrenciye bir örneği verilebilir."	İçerik önceden belirlenmez. Nitelikli veya niteliksiz birer çalışmaya içerebilir. Kolaylık sunu. Herda dosyanın sürdürülmesi zorundur. Evde bulunanlar için öğrenciye bir örneği verilebilir.
Dereceli Puanlara Anahtar (Rubric)	"Ölçülen başarıya ilişkin performans, çeşitli düzeylerle tanımlanabilir ve ölçüler detaylı belirleyen ölçektir."	"Öğretmenin beklentisini somut hale getiren. Belleğin performansın beklentilere oranı gösteren bir detaylı gösterim."	Horizantalı çok zonen alır ve zamanlık gerektirir.	Horizantalı çok zonen alır ve zamanlık gerektirir.
Derecelendirme Ölçetleri	Bir çalışmanın, etkinliğin veya cevabın değerlendirilmesinde kullanılan puanlama rehberidir."	Horizantalı kolaydır. Öğretmenin beklentilerini somut hale getiren. Birde davranışları denek için kullanılabilir."	Gösteren, zonen alır. Oran alımları ve değer yargıları, duyguları sonuçları etkileyebilir."	Gösteren, zonen alır. Oran alımları ve değer yargıları, duyguları sonuçları etkileyebilir."
Kontrol Listeleri	"Öğrenci performansının eksik noktalarını belirlemek amacıyla kullanılan araçlardır."	Eksik noktaları belirler. İstenen davranışları net olarak ortaya koyar. Tekrar kullanılır. Karşılaştırma yapılabilir."	Sadece gözlenebilir ve gözlemlenmeyen zamanlarda bir okutma zor verilir."	Sadece gözlenebilir ve gözlemlenmeyen zamanlarda bir okutma zor verilir."
Kavram Haritaları	"Ünile veya kavramlar geçen kavramların birbirleriyle ilişkilerinin yazıldığı, tablo veya grafiklerdir."	Öğrenmeyi kolaylaştırır. Kavram ilişkilerini belirler. Öğrenmeyi etkili kılar. Tekrar kullanılabilir."	Ayrıntılı bilgi vermez. Haritaların zorudur."	Ayrıntılı bilgi vermez. Haritaların zorudur."

Klasik Değerlendirme Yöntemleri

1. Değerlendirme sürecinde yalnız öğretmen aktif.
2. Öğretmen merkezli eğitim değerlendirilir.
3. Ürün (sonuç) önemlidir.
4. Ezbere dayalıdır.
5. Kağıt, kalem, standart testler kullanılır.
6. Geri bildirim ve sonuç not olarak verilir.
7. Ölüt önceden belirlenir.
8. Puanlamada sayısal semboller kullanılır.
9. Özetleyici ve yorguluyucudur.

Çağdaş Değerlendirme Yöntemleri

1. Değ. sürecinde öğretmen ve öğrenci aktiftir.
2. Öğrenci merkezli eğitim değerlendirilir.
3. Süreç odaklıdır.
4. Düşünme ve değerlendirmeye dayalıdır.
5. Çok boyutlu değer, ürün ya da performans örnekleri kullanılır.
6. Geri bildirim süreç içinde verilir.
7. Ölüt öğrencilerle belirlenir.
8. Rubric puanlama kullanılır.
9. Geliştirici ve teşhis edicidir.

Soru Sorarken Dikkat Edilecek Hususlar

1. İstenilen cevap açık ve anlaşılır olmalı : Öğrenci soruyu okurken kendinden hangi cevabın istendiğini net olarak anlamalıdır. Dil ve gramer kurallarına dikkat edilmeli.

2. Test maddeleri hedef davranışlara uygun olmalı : Sınavdan önce belirtke tablosu hazırlanmalı. Bu şekilde hangi konularda, hangi düzeyde soruların sorulması gerektiği daha net ortaya konulur.

3. Objektif testlerde bir maddeyle birden fazla bilgi yoklanmamalıdır. Ve, veya gibi bağlaçlarla iki ya da daha çok bilginin yoklanması öğrencileri yönlendirebilir.

4. Maddeyi oluşturan cümle bir kaynaktan alınmalıdır : Özellikle öğrencilerin elinde bulunan ders kitaplarından soru alınıp sorulmalıdır. (Tıradık sorular sınav geçerliğini düşürür)

5. Soruda ipucu bulunmamalı : Bir test maddesi başka bir test maddesi için cevap özelliği taşımamalı

6. Aynı sorular her yıl üst üste kullanılmamalı : Bilete bilmeceyi net olarak ayırtmayacağı için güvenirlik ve geçerliği düşürecektir.

7. Puanlama objektifliğini artırıcı tedbirler alınmalı : Sınavdan önce detaylı cevap anahtarı hazırlanmalı, cevap anahtarı soru yapılırken hazırlanmalı.

TEST PUANLARI ÜZERİNDE İSTATİSTİKİ İŞLEMLER

1. Bir testten herhangi bir anlam çıkarmak için; öncelikle notlar büyükten küçüğe ya da küçükten büyüğe sıralanır.
2. Bu sıralama bize öğrencinin sınıf içindeki başarı sırasını ve en büyük puanla en küçük puan arasındaki farkı (range) belirlememizi sağlar.
3. Puanlar hakkında daha iyi fikir edinebilmek için frekans dağılımı hazırlanır.
4. Frekans dağılımı, test puanlarının nasıl bir dağılım gösterdiğini ve her puanın kaç kez tekrarlandığını gösterir.
5. Tüm puanlar büyükten küçüğe sıralanır ve her puanı alan öğrenci sayısı karşısına yazılır.
6. Puan türleri arttıkça tablo uzayacağı için puanlar grupeleme yapılır.
7. Grup ortalgı belirlenirken, önce puanların kaç grupta toplanacağına karar verilir. Daha sonra range, belirlenen grup sayısına bölünür.
 Örn: Puanlar 90 ile 20 arasındaysa ve 15 grupta toplanacaksa grup ortalgı ne olmalı?

$$(90-20) \div 15 = 5$$
8. **Grup ortalgılarının gerçek noktası:** Grup ortalgılarının alt sınırının 0,5 altı ve üst sınırının 0,5 üstüdür.
9. **Grup ortalgılarının orta noktası:** Puan ortalgılarının her iki puanı toplanarak 2'ye bölünür.
10. **İlgilenmeli (kümülatif) frekans:** Frekans sıralamasının alttan üste doğru toplanmasıyla bulunur.

ÖLÇMEDE İSTATİSTİKİ İŞLEMLER

1. TEST İSTATİSTİKLERİ

A. Merkezi Eğilim Ölçüleri

1. Aritmetik Ortalama
2. Ağırlıklı Ortalama
3. Ortanca (medyan)
4. Tepe değer (mod)

B. Merkezi Dağılım Ölçütleri

1. Standart Sapma
2. Bağıl Değişkenlik Katsayısı
3. Dizi Genişliği (R_{max})
4. Çeyrek Sapma
5. Varyans

C. Standart Puanlar

1. Z Puanı
2. T Puanı

D. Normal Dağılım Eğrileri

1. Normal Dağılım
2. Gözpik Dağılım

E. Testin Ayırt Ediciliği

1. Cronbach'a göre
2. A.O'ya göre

F. Testin Ortalama Gücü

G. Mutlak Başarı Yüzdesi

2. MADDE İSTATİSTİKLERİ

1. Madde Standart Sapması

2. Güçlük İndeksi

3. Ayırt Edicilik

4. Madde Varyansı

1. TEST İSTATİSTİKLERİ

20 öğrenciye 100 puan üzerinden uygulanan testten elde edilen sonuçları inceleyelim:

1. (76, 41, 76, 80, 74, 47, 85, 95, 71, 54, 85, 65, 85, 83, 55, 66, 66, 50, 83, 69)

Bu sonuçlar sadece öğrenciler hakkında bilgi verir ve bu bilgileri yorumlamak, grup hakkında fikir yürütmek zordur.

Puanlar sıralanmadan Aritmetik Ort, Standart Sapma, RST ve Ortanca gibi bilgilere ulaşılamaz.

Öncelikle elimizdeki verileri sıralamak gerekir.

2. Puanların frekans dağılımı yapılarak sıralanması

Bu adımda elimizdeki verileri, küçükten büyüğe veya büyükten küçüğe sıralamak gereklidir. 20 öğrenciye ait veriler, sıralanmış olarak aşağıdaki gibidir.

Puan(x)	Frekans(f)	Puan(x)	Frekans(f)
95	1	66	2
85	3	65	1
83	2	55	1
76	2	54	1
74	1	50	2
71	1	47	1
69	1	41	1

Frekans dağılımı yapıldıktan sonra sırasıyla şu işlemler yapılır:

3. Grup Frekans Analizi : Puanların daha geniş bir aralığa dağılması halinde tablodaki farklı puan sayıları daha büyük olur. Puan dağılımında daha rahat işlemler yapılabilmek için puanlar gruplanır.

4. Puan Aralığının Orta Noktası : Aritmetik olarak puan aralığının orta noktası tespit edilir.

5. Yığılmalı (Kümülatif) Frekans : Frekanslar dizinin altından itibaren toplanarak bulunur. Toplam frekanstır.

6. Grup Gerçek Analizi : Grubun gerçek aralığı üst sınıp +0,5 eklenir, alt sınırından -0,5 çıkarılarak bulunur.

Puan Aralığı	Frekans	Grup Aralığının Orta Noktası	Yığılmalı Frekans	Gerçek Grup Aralığı
91-95	1	93	20	90,5 - 95,5
86-90	0	88	-	85,5 - 90,5
81-85	5	83	19	80,5 - 85,5
76-80	2	78	14	75,5 - 80,5
71-75	2	73	12	70,5 - 75,5
66-70	3	68	10	65,5 - 70,5
61-65	1	63	7	60,5 - 65,5
56-60	0	58	-	55,5 - 60,5
51-55	2	53	6	50,5 - 55,5
46-50	3	48	4	45,5 - 50,5
41-45	1	43	1	40,5 - 45,5

A. MERKEZİ EĞİLİM ÖLÇÜLERİ

1. Aritmetik Ortalama (\bar{X})

Verilerin ağırlık merkezidir. Dizideki puanların toplamı ölçüm sayısına (frekans) bölünmesiyle elde edilir.

Formül: $\bar{X} = \frac{\sum X}{N}$

\bar{X} : Aritmetik Ortalama
 $\sum X$: Toplam Puan
 N : Öğrenci Sayısı

a) Öğrencilerin Puanları Birer Kez Tekrar Etmiş İse;

Puanlar toplanır, öğrenci sayısına bölünür.

ÖRNEK: Altı öğrencinin aldığı sınavın puan dağılımı aşağıdadır.

P	Cözüm: $\bar{X} = \frac{\sum X}{N}$
65	
50	
82	50+65+70+
70	82+87+90 = (74)
87	6
90	

b) Puanların Frekansı Birten Fazla ise;

Önce puanlar frekanslarla çarpılır, sonuçlar toplanır, eleman sayısına bölünür.

Puan	Frekans	Pxf
90	3	270
77	4	308
70	2	140
66	2	132
	(N=11)	Σx=850
		$\frac{850}{11} = (77,2)$

c) Puanlar Gruplanmış İse;

Önce puan aralığının orta noktası bulunur.

Bulunan sonuç frekansla çarpılır, puanlar toplanır.

Toplanan puanlar frekans sayısına bölünür.

ÖRNEK: Puan Aralığı	Frekans	Puan Aralığının Orta Noktası	Σ f.X	
86-90	1	88	88	
81-85	3	83	249	
76-80	4	78	312	
71-75	2	73	146	
66-70	2	68	136	
61-65	1	63	63	
	(12)		(994)	$\frac{994}{12} = (82,8)$

2. Ağırlıklı Aritmetik Ortalama

Bazen puanların ortalamaya katkıları farklı olabilir.

Farklı ağırlık verilerek hesaplanan ortalamalardır.

Üniv. ve liselerde yıl sonu notlarının hesaplanmasında kullanılır.

Formül: $\bar{X} = \frac{\sum (Kredi \cdot Puan)}{\sum Kredi}$

Dersin kredisi ile o dersten alınan not çarpılarak sonuçlar toplanır ve toplam kredi sayısına bölünerek bulunur.

ÖRNEK:

Dersler	Kredi	Puan	Puan x Kredi
Matematik	6	80	480
Edebiyat	6	70	420
Dil Anlatım	4	65	260
Tarih	2	79	158
Coğrafya	2	65	130
Fizik	2	72	144
Kimya	2	56	112
Biyoloji	2	66	132
	<u>26</u>		<u>1826</u>

Çözüm: $\bar{X} = \frac{\sum (Kredi \cdot Puan)}{\sum Kredi} = \frac{1826}{26} = 70,2$

3. Ortanca (Medyan)

Ortanca, üstten-alta ya da soldan sağa sayı dizisinin tam ortasında bulunan puandır. Sıralanmış veriyi iki eşit yarıya böler.

ÖRNEK: Bir öğrenci girdiği sınavlardan

(80, 87, 74, 68, 90, 67, 55, 76, 67, 75, 92...) puanlarını almıştır.

Ortanca'yı bulmak için puanları sıralanarak gerekir:

92, 90, 87, 80, 76, 75, 74, 68, 67, 67, 55
Medyan

d. Veriler Tek Sayıda Olursa;

Baştan veya sondan orta sıradaki

ölçme sonucu sayılarak bulunur.

ÖRNEK: 50, 53, 60, 60, 65, 70, 77, 84, 90
M

b). Veriler Çift Sayıda Olursa;

Baştan ve sondan ortadaki 2 veri

ortanca olarak ortaya çıkar.

İki sayı toplanıp ikiye bölünür.

ÖRNEK: 2, 3, 4, 4, 5, 5, 7, 7, 8, 8, 9
 $\frac{5+7}{2} = 6$ kabul edilir.

c). Gruplanmış Ölçümlerin Ortancasını Hesaplama

Bu tür hesaplamada, bir orolığa rastlayan ölçümlerin oroluk boyunca eşit olarak dağıldığı kabul edilir.

Not: Bu işlemlerde frekans sütunundan faydalanılır.

ÖRNEK:

Puan Aralığı (x)	Frekans (f)	Yığılmış Frekans (Σf)
81-85	1	44
76-80	3	43
71-75	4	40
66-70	9	36
<u>61-65</u>	<u>10</u>	<u>27</u>
56-60	6	17
51-55	5	11
46-50	3	6
41-45	3	3

medyan burada.

61-70... 65

63 medyan alın

$\frac{27}{2} = 13.5$

4. TEPE DEĞER (MOD)

Bir puan grubunda en çok tekrar eden ölçme sonucudur.

Mod, merkezi yığılma ölçütleri içinde en az bilgi veren ölçüdür.

Örnek:

Puan(x)	Frekans(f)
90	4
80	5
70	8
60	10
50	6
40	5
30	4
20	1

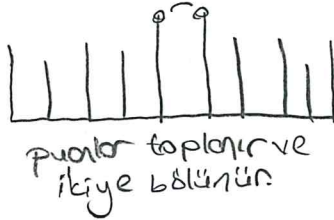
a). Gözlenen frekanslar birbirine eşit ise mod olmaz, dağılım modsuzdur.



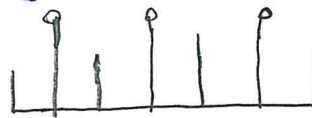
b). Sadece bir değer diğerlerinden fazla tekrar etmişse dağılım tek modludur.



c). Ardışık 2 veya daha çok ölçüm birbirine eşit sayıda ve diğerlerinden çok tekrar etmişse mod ardışık ölçümlerin ortasıdır.



d). Ardışık olmayan 2 ya da daha çok ölçüm eşit sayıda ve diğerlerinden daha çok tekrar etmiş ise dağılım çok modludur.



e). Eğer 2 değer diğerlerinden fazla tekrar etmiş ve eşit sayıda tekrar etmişse dağılım çift modludur.



f). Gruplanmış verilerde frekansı en çok olanın ortası mod olarak alınır.

B. MERKEZİ DAĞILMA ÖLÇÜLERİ

1. Standart Sapma (Kayma)

Standart Sapma (SS), bir puan dağılımında her bir puanın Aritmetik Ortalamadan (\bar{X}) ne kadar uzaklaştığının ortalamasını veren bir değişiklik ölçüsüdür.

Formül $S_x = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$

S_x : Standart Sapma

n : Öğrenci sayısı

$\sum (x - \bar{x})^2$: Bütün puanların A.O'dan farklarının karesi.

Ölçme sonuçlarının aritmetik ortalamadan farklarının kareleri toplanır, kişi sayısının bir eksiğine bölünür.

Puanların A.O değerini hesaplamak için;

- Elde edilen puanların A.O'sı hesaplanır
- Her puanın A.O'dan farkı alınır
- Puanların A.O'dan farklarının karesi alınır.
- Farkların karesi toplanır öğrenci sayısının bir eksiğine bölünür ve altı sayının karekökü alınıp işlem yapılır.

Puanlar(x)	A.O(\bar{x})	Puanın A.O'dan farkı ($x - \bar{x}$)	$(x - \bar{x})^2$ Puanın A.O'dan farkının karesi	$\sum (x - \bar{x})^2$	$\frac{\sum (x - \bar{x})^2}{n - 1}$
21	24	21-24 = -3	9	28	$\frac{28}{7-1} = \frac{28}{6}$ $= 4,66$ $\sqrt{4,66} = 2,15$
22	24	22-24 = -2	4		
23	24	23-24 = -1	1		
24	24	24-24 = 0	0		
25	24	25-24 = 1	1		
26	24	26-24 = 2	4		
27	24	27-24 = 3	9		

Not: Bir dizideki ölçümlerin birbirinden farkı arttıkça standart sapma büyür, ölçümler birbirine yaklaştıkça küçülür.
 Testin Standart Sapması büyüdükçe testin güvenirliği artar.
 İki ya da daha fazla ölçümler dizisinin ortalamaları aynı olsa bile S.S'leri farklı olabilir.

Standart Sapmanın Yorumu

Ss. Küçük Olan Gruplar

Ölçümler arasında değişim azdır
 Puanlar, A.O etrafında toplanmıştır
 Alınan notlar birbirine yakındır
 Grup homojendir (Benzer)
 (Kolay sınavlar)

Ss. Büyük Olan Gruplar

Güçlü değişim (farklılaşma) büyüktür
 Puanlar, A.O'dan uzaklaşmıştır.
 Alınan notlar birbirinden uzaktır
 Grup heterojendir (Farklıdır)
 (Sevici zor sınavlar)

Not: Standart Sapma; tek başına yorum yapmaya için yeterli değildir. Sınıfın başarısı hakkında bakılacak ilk ölçü "A.O"dır.

ÖRNEK

Dersler	Ortalama	Sx
Türkçe	64	5
Matematik	85	6
Geometri	93	3
Fizik	55	5
Kimya	50	7
Tarih	70	8

S dersinde S.S'sı farklıdır:

(Geo) dar bir alana yayıldığı için Ss. en küçük olandır.

(Tar) geniş alana yayıldığı için Ss daha büyüktür.

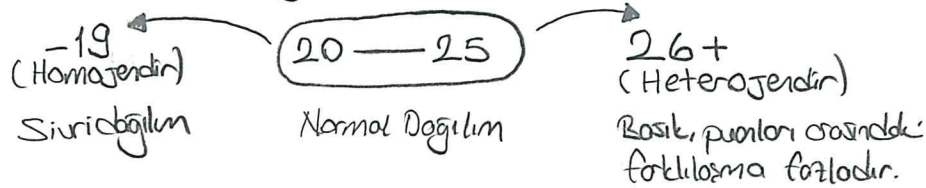
2. Bağıl Değişkenlik Katsayısı (BDK)

BDK, Standart kaymanın A.O'ya bölünerek, yüzle çarpılması sonucunda elde edilen bir yüzdendir.

BDK(V): Hem A.O, hem S.S'yi içeren, daha çok bilgi veren ölçümdür.

$$\text{Formül} = V = \frac{S_x}{\bar{X}} \cdot 100$$

BDK ve Dağılımlar Arasındaki İlişki



3. Dizi Genişliği (RanJ)

Bir puan dağılımında en büyük ölçme sonucu ile en küçük ölçme sonucu arasındaki farka (RanJ) denir.

Sadece dağılımın en büyük ve en küçük puanını ifade eden

$$\text{RanJ} = \text{ERP} - \text{EKP}$$

Öz: (55, 38, 44, 81, 23, 75, 74, 35, 52, 61, 76) $\text{RanJ} = 81 - 23 = 58$

RanJ değeri düşük olursa: Öğrencilerin puanları birbirine benzediği için bilen ile bilmeyeni ayırt edemediği için güvenirlik düşüktür.

RanJ değeri yüksek olursa: Bilenle bilmeyeni iyi ayırdığı için güvenirlik yüksektir.

A.O'ya Bağlı Olarak Verilen Kararlar	Ss'ya Bağlı Olarak Verilen Kararlar
--------------------------------------	-------------------------------------

- | | |
|--|---|
| x Grubun başarı düzeyi nedir? | x Başarılı ve başarısız sınıf hangisidir? |
| x Grubun mutlak başarı düzeyi nedir? | x Öğrenciler arasında farklılaşma var mı? |
| x Öğrencilerin Ort. başarı düzeyi nedir? | x Grup homojen mi, heterojen mi? |
| x Öğrencilerin öğrenme düzeyi nedir? | x Grup A.O'ya ne kadar yakın ya da uzak? |

A.O / MOD / MEDYAN KARŞILAŞTIRMA

★ A.O, ölçümlerin hepsini kullanır, bu yüzden daha kullanışlıdır.
Medyan yorisinden, Mod bir ya da birkaç ölçümler yorisilerek hesaplanır.

★ Ortolamanın değeri, ölçümlerin sayısal değerine bağlıdır, medyan ölçüm sayısına, mod tekrar sayısına bağlıdır.

★ A.O aşırı ölçümlerden etkilenir, diğerleri etkilenmez.

★ Kaymalarda A.O kullanılır.

4. Çeyrek Sapma

Bir puan dizisi 4 eşit parçaya bölünerek, dördlüklerin sınırları bulunur. Küçük puanlardan itibaren birinci dördlüğün sınırına Q_1 , ikinci dördlüge Q_2 üçüncü dördlüge Q_3 adı verilir.

Üçüncü çeyreğin sınırından birinci çeyreğin sınırı çıkarılır ($Q_3 - Q_1$) fark ikiye bölünür.

(%75'inci sayıdan %25'inci sayı çıkarılır 2'ye bölünür)

$$\text{Formül} = Q = \frac{Q_3 - Q_1}{2}$$

Q : Çeyrek sapma
 Q_3 : Üçüncü çeyrek (%75)
 Q_1 : Birinci çeyrek (%25)

ÖRNEK : 59, 63, 64, 65, 65, 65, 66, 66, 66, 67, 68, 68, 70, 71, 73, 74

\downarrow \downarrow \downarrow
 Q_1 Q_2 Q_3

C. STANDART PUANLAR

- *). Bir öğrencinin farklı derslerdeki başarı durumunu kıyaslamak için standart puanlar kullanılır.
- *). Farklı ölçümlerden elde edilen puanları karşılaştırmak, öğrenciyi yanlış yorumlama yapmaya yönlendirir. Çünkü her testteki modellerin güçlük düzeyleri ve gruptaki öğrencilerin özellikleri farklılık gösterir.
- *). Toplanan ham verileri, anlamlı hale getirmek için standart puanlar kullanılır. \Rightarrow "Z Puanı" - "T Puanı"

1. Z PUANI

Aritmetik Ortalaması 0, Standart Sapması 1, kabul edilen puanlardır.

$$Z = \frac{X - \bar{X}}{S}$$

X : Ham puan
 \bar{X} : Aritmetik ortalama
 S : Standart sapma

ÖRNEK: Esma'nın beş farklı dersten aldığı puanlar, sınıfın ortalaması ve standart sapması aşağıdadır. Esma hangi derste daha başarılıdır.

Ders	X	\bar{X}	S	Z Puanı
Matematik	60	50	5	$\rightarrow (60-50)/5 = +2$
Türkçe	80	90	10	$\rightarrow (80-90)/10 = -1$
İngilizce	40	60	10	$\rightarrow (40-60)/10 = -2$
Resim-İş	70	80	2	$\rightarrow (70-80)/2 = -5$
Beden Eğt.	70	60	10	$\rightarrow (70-60)/10 = +1$

Yorum: Z puanı hesaplanmadan yapılacak bir yorumda Esma'nın Türkçe dersindeki başarısı daha yüksek denilebilir. Z puanı hesaplandığında Esma'nın en başarılı olduğu dersin Mat. ($Z = +2$) olduğu görülür. En başarısız olduğu ders ise puanının en düşük olduğu İng. değil, daha yüksek puan olduğu Resim ($Z = -5$) dersidir.

2. T PUANI

Aritmetik ortalaması 50, standart sapması 10, kabul edilen puarlardır.

2 puanının negatif olabilmesi ve 0 değerine yakın sonuçlar vermesi yorumcuları zora sokar. Bu nedenle 2 puanı, T puanına dönüştürülür.

$$\text{Formül} = T = 50 + 10 \left(\frac{x - \bar{x}}{s} \right) = 50 + 10z.$$

ÖRNEK: 2 puanları hesaplanan Esma'nın T puanları nedir?

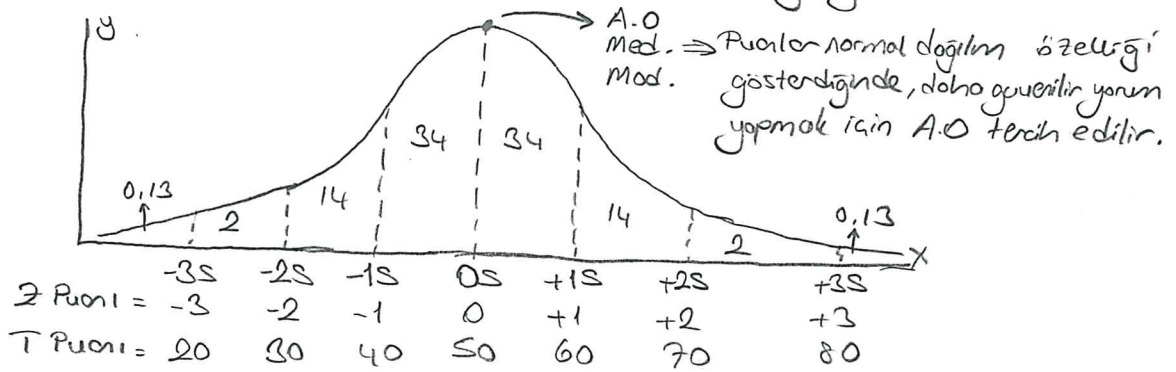
Ders	2 puanı	T puanı
Matematik	+2	$50 + 10(2) = 70$
Türkçe	-1	$50 + 10(-1) = 40$
İngilizce	-2	$50 + 10(-2) = 30$
Resim-İş	-5	$50 + 10(-5) = 0$
Beden Eğt.	+1	$50 + 10(1) = 60$

Yorum = Puanlar incelendiğinde, 2 puanındaki sıralamanın değişmediği görülür.

D. DAĞILIM EĞRİLERİ

1. Normal Dağılım Eğrisi

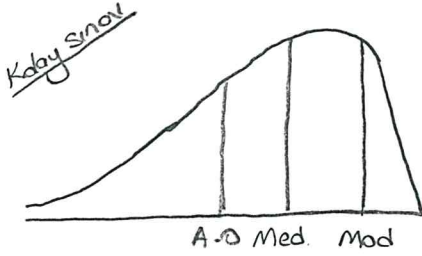
Birçok istatistikî işlemlerde temel alınan A.O ve Ss'ye göre düzenlenir.



- $\bar{x} \pm 1s$ alanı içinde tüm ölçümlerin %68.26'sı bulunur.
- $\bar{x} \pm 2s$ alanı içinde tüm ölçümlerin %95.44'ü bulunur.
- $\bar{x} \pm 3s$ alanı içinde tüm ölçümlerin %99.74'ü bulunur.

2. Çarpıklık Katsayısı

a. Sola Çarpık Dağılım



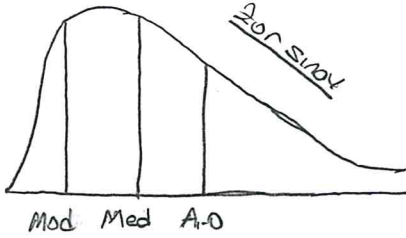
Punların yüksek olduğunu ve öğrencilerin hedef davranışların çoğunu kazandıklarını gösterir.

En yüksek değer = Mod / En düşük değer = A.O.

Ölçüt olarak Medyan kullanılmalıdır.

ÖRNEK: Bir dağılımda A.O 70, medyan 80, Mod 90 ise; ölçüt olarak medyan kullanılır.

b. Sağa Çarpık Dağılım



Öğrencilerin puanları düşüktür, öğrenciler hedef davranışların çoğunu kazanamamıştır.

En düşük değer = Mod / En yüksek değer = A.O.

Ölçüt olarak medyan kullanılmalıdır.

ÖRNEK: Öğrencilerin 3 farklı dersten aldıkları puanlar tabloda verilmiştir. Bu puanlar nasıl bir dağılım gösterir? Öğretmen ölçüt olarak hangi değeri almalıdır?

Tarih

P	f
1	3
2	5
3	7
4	4
5	2

Normal Dağılım
A.O ölçüt
olarak kullanılır.

Matematik

P	f
1	9
2	8
3	6
4	5
5	2

Sağa Çarpık
Ölçüt = Medyan
kullanılır.

Türkçe

P	f
1	1
2	4
3	6
4	7
5	10

Sola Çarpık
Ölçüt = Medyan
kullanılır.

1. Ronj'a Göre

E. TESTİN AYIRT EDİCİLİĞİ

- Testin amacı bilen ile bilmeyen öğrenciyi ayırt etmekse; testin ayırt ediciliği Ronj ile hesaplanabilir.
- Ronj değeri testte yer alan soru sayısının yarısına eşit ya da pozitif yönde büyük olursa; ayırt edicilik sağlanmış olur.
- Ronjin büyük olması o testten en yüksek ve en düşük puan alan öğrenci arasındaki farkın yüksek olduğunu gösterir.
- Testte ayırt ediciliğin olması için Ronjin yüksek olması gerekir.

ÖRNEK:

Dersler	Ronj	K
Matematik	26	50/2
Fizik	20	40/2
Kimya	45	80/2
Biyoloji	45	100/2
Türkçe	75	110/2

⇒ Öğrencilerin hangi dersteki ayırt ediciliği en yüksektir?

$$C: M = 26 - 25 = +1$$

$$F = 20 - 20 = 0$$

$$K = 45 - 40 = +5$$

$$B = 45 - 50 = -5$$

$$T = 75 - 55 = +20 \times \text{Türkçe}$$

2. A.O'ya Göre

- Testin A.O'sından yararlanarak testin ayırt ediciliği hesaplanabilir.
- Bir testin ayırt ediciliğinin yüksek olması için; Testin A.O'sının o testten alınabilecek en yüksek puanın ya da soru sayısının yarısına eşit ya da yakın olması gerekir.

ÖRNEK:

Dersler	A.O	K
Fizik	11	20/2
Kimya	25	50/2
Matematik	63	100/2
Geometri	19	35/2
Biyoloji	37	60/2

⇒ Öğrencilerin hangi dersteki ayırt ediciliği en yüksektir?

$$C: F = 11 - 10 = +1$$

$$K = 25 - 25 = 0 \times \text{Kimya}$$

$$M = 63 - 50 = +13$$

$$G = 19 - 18 = +1$$

$$B = 37 - 30 = +7$$

F. TESTİN ORTALAMA GÜÇLÜĞÜ

Aritmetik ortalama, testten alınabilecek en yüksek puana bölünür.

$$\bar{P} = \frac{\bar{X}}{K}$$

\bar{P} = Ortalama güçlük

\bar{X} = Aritmetik ortalama

K = Testteki madde sayısı,

Testin güçlük derecesi, 0 ile +1 arasında olur. Testin güçlüğü 0'a yakınsa zor, +1'e yakınsa kolay olur.

Testin güçlüğü kullanılış amacına göre değişir.

Aday sayısı çok, alınacak aday az ise, ort. güçlük 0'a yakın olmalı

Testin Ortalama Güçlüğü Yapılacak Sonuç

+0,50↑ → içerik iyi öğrenilmiş veya test öğrencilere kolay gelmiştir.
0,50 → ideal bir test.
-0,50↓ → içerik iyi öğrenilmemiş veya test öğrencilere zor gelmiştir.

ÖRNEK

: Bir testten alınabilecek en yüksek 500 puan alınabilmektedir. Testin A.O'sı 185 ise ortalama güçlüğü kaçtır?

$$\bar{P} = \frac{\bar{X}}{K} = \frac{185}{500} = \boxed{0,37}$$

G. MUTLAK BAŞARI YÜZDESİ

Testin ortalama güçlüğü bulunduysa; Ortalama güçlük düzeyi bulunmadıysa;

Çıkan sonuç 100 ile çarpılır ve mutlak başarı düzeyi bulunur.

Testin A.O'sı testten alınabilecek en yüksek nota veya soru sayısına bölünür ve 100 ile çarpılır.

$$MBY = P_{test} \cdot 100$$

$$MBY = \frac{\bar{X}}{K} \cdot 100$$

ÖRNEK

: Her biri 0 ve 1 şeklinde puanlanan 100 soruluk genel başarı sınavında, eşit sayıda öğrenci olan beş sınıfın puan dağılımları, A.O'ları ve standart sapmaları verilmiştir. En başarılı sınıf hangisidir?

İki sınıfta mutlak başarı eşittir. Sı'ısı küçük olan daha başarılıdır.

Sınıf	\bar{X}	S_x	Cözüm
I. Sınıf	50	7	$\bar{X} = \frac{50}{100} \cdot 100 = 50$
II. Sınıf	70	7	$\bar{X} = \frac{70}{100} \cdot 100 = 70$
III. Sınıf	50	10	$\bar{X} = \frac{50}{100} \cdot 100 = 50$
IV. Sınıf	60	5	$\bar{X} = \frac{60}{100} \cdot 100 = 60$
V. Sınıf	70	5	$\bar{X} = \frac{70}{100} \cdot 100 = 70$

MADDE İSTATİSTİKLERİ

Öğrencilere uygulanan testin modelleri ve seçenekleri üzerinde yapılan çalışmaya "madde analizi" denir.

Madde analiziyle ; madde istatistikleri hesaplanır
teste konulacak maddeler seçilir.
düzeltilecek ve testten çıkarılacak maddeler seçilir.

Madde istatistiği hesaplamalarında iki yöntem kullanılır:

←
%27'lik alt-üst grup yöntemi
"Öğrenci sayısı çok olduğu zaman"

→
Henryson yöntemi
"Öğrenci sayısı az olduğu zaman"

1. %27'lik Üst-Alt Grup Yöntemi

Öğrencilerin doğru cevaplarının sayısı ile test puanı hesaplanır.

Puanlar yüksekten-düşüğe doğru sıralanır.

Test puanında %27 üst grup ile %27 alt grup alınır. Aradaki %46'lık grup dikkate alınmaz.

Madde No	SEÇENEKLER					Toplam
	A	B	C	D	E	
Üst Grup	(A _ü)	(B _ü)	(C _ü)	(D _ü)	(E _ü)	N _{üst}
Alt Grup	(A _a)	(B _a)	(C _a)	(D _a)	(E _a)	N _{alt}
Toplam	n(A)	n(B)	n(C)	n(D)	n(E)	N

2. Henryson Yöntemi

Bu yöntemde öncelikle öğrencilerin modellere verdikleri cevaplar doğruysa 1, yanlış, boş bırakılmış veya birden fazla cevap işaretlenmişse 0, olarak puanlanır.

Her bir öğrencinin kâğıdından; öğrenciler satırda, maddeler sütunda olmak üzere madde puan matrisi oluşturulur.

Öğrenciler	Sorular					Puan(x)
	1	2	3	4	5	
Hakan	1	1	1	0	1	4
Ayşe	1	1	0	1	0	3
Harun	0	0	0	1	0	1
Deniz	1	1	0	1	0	3
Ufuk	0	1	0	1	0	2
Merve	1	1	1	1	1	5

Not: Her sütun 1 veya 0 olursa etkisizdir.
Farklılıkları ortaya çıkarmayan sorudur.

1. MADDE GÜÇLÜK İNDEKSİ (P_j)

Bir maddeye doğru cevap verenlerin gruptaki toplam öğrenci sayısına oranıdır.

$$P_j = \frac{n_{\bar{u}} + n_a}{n_t}$$

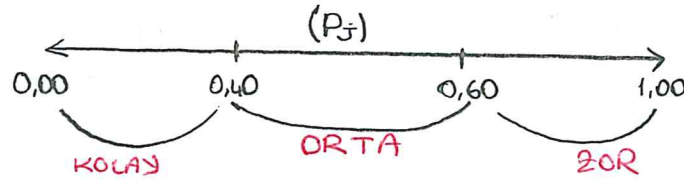
P_j : Madde güçlüğü

$n_{\bar{u}}$: Maddeyi doğru cevaplayanlar (üst Grup)

n_a : Maddeyi doğru cevaplayanlar (alt Grup)

n_t : Gruptaki toplam öğrenci sayısı

Bir maddenin güçlük indeksini hesaplamak için o maddeye doğru cevap verenlerin sayısı belirlenir. Bu sayı tüm cevap verenlere bölünür.



En ideal madde = 0,50

Madde güçlük indeksi "1"e yaklaştıkça maddenin güçlük düzeyi yükselir. Madde güçlük indeksi arttıkça soru kolaylaşır. Madde güçlük indeksi düştükçe de madde zorlaşır.

Not₁: Sınavı başaran kişi sayısı çok fazla ve alınacak kişi sayısı az ise maddelerin zor olması istenir. (0,00-0,40)

Not₂: Sınavı amaç öğrenci başarısını yüksek tutmak ise soruların çoğunluğunun kolay olması istenir. (0,60-1,00)

ÖRNEK:

Madde No=12	A	B	C	D	E*	Baş	Toplam
Üst Grup	5	3	6	4	30	2	50
Alt Grup	9	5	7	5	21	3	50

*Anotatörlü doğru cevap

Verilen sorunun madde güçlük indeksi kaçtır?

Gözüm : $P_j = \frac{n + n}{N} = \frac{30 + 21}{100} = \frac{51}{100} = 0,51$

2. MADDE AYIRT EDİCİLİK İNDEKSİ (r_j)

r_j : Bir maddenin bilen öğrenci ile bilmeyen öğrenciyi birbirinden ayırt etmesidir.

$$r_j = \frac{n_{\bar{u}} - n_a}{n}$$

r_j = Madde ayırt edicilik indeksi

$n_{\bar{u}}$ = Üst grupta doğru cevap verenlerin sayısı

n_a = Alt " " " " "

n = Üst ve Alt gruptaki toplam öğrenci

Bir madde başarılı ve başarısız öğrenciyi ayırt ettiği kadar ayırıcıdır.

Bir maddenin göreceli bilen ve bilmeyen öğrenciyi ayırmaktır.

Bir maddenin ortalamının altında kalan öğrenciler tarafından bilinmemesi, ortalamının üstünde kalan öğrenciler tarafından bilinmesi durumunda o maddenin ayırt ediciliği pozitif olur.

0,40 ve üstü → çok iyi madde

0,30 — 0,39 → oldukça iyi madde

0,20 — 0,29 → düzeltilmeli/geliştirilmeli

0,19 ve altı → değiştirilmeli

Ayırt edicilik indeksi -1,00 ile +1,00 arasında değer alabilir.

Ancak negatif değerler olan maddeye önemli bir kusur vardır.

Negatif ayırt edici madde, testte ölçülmek istenen amac dışında başka değişkenleri ölçer. Teste alınmamalı.

Teste alınacak madde en az 0,30'dan büyük olmalıdır.

Madde ayırt edicilik gücünün yüksek olması, maddenin dolayısıyla testin geçerliğinin yüksek olmasını sağlar.

Madde ayırt edicilik gücü yüksek bir maddenin seçenek dağılımları aşağıdaki gibi ya da bu tabloya yakın değerler olmalıdır.

1. Madde	SEÇENEKLER					Toplam
	A*	B	C	D	E	
Üst Grup	100	-	-	-	-	100
Alt Grup	-	25	25	25	25	100
Toplam	100	25	25	25	25	200

*Annohtorlanmış doğru cevap

Yukarıdaki tabloda: Sınıfın başarılı öğrencilerinden hiç biri gelinmeden doğru seçeneği işaretlemiş.

Sınıfın başarısız, alt grup öğrencileri doğru cevabı işaretlememiş.

Beklenen ve istenen sonuç: iyi bir maddeyi üsttekilerin doğru alttekilerin yanlış cevaplamasıdır.

Madde bilene bilmeyeni, başarılı ve başarısız çok iyi ayırmıştır.

Maddenin güçlüğü 0.50 ve ayırt ediciliği +1.00'dir. (mükemmel madde)

Annohtorlanmış cevabın dışında kalan modeller geldiricidir ve tüm geldiricilerin aynı sayıda öğrenciyi getmeleri beklenir.

Eğer zayıf ya da çok getmiş olan geldiriciler varsa bunlar birbiriyle denge içinde olması için düzeltilmelidir.

Kullanılacak ya da teste alınacak maddelerin, madde güçlük indeksleri 0,50 civarında ve madde ayırtıcılık gücü indeksleri de +0,30'dan büyük ve mümkünse +1,00'e yakın olmalıdır.

ÖRNEK:

Aşağıda beş sorunun madde güçlük indeksi (P_j) ile ayırt edicilik değerleri (r_j) verilmiştir.

Soru	P_j	r_j
1	0,55	0,35
2	0,80	0,45
3	0,40	0,18
4	0,25	0,05
5	0,50	-0,10

① Sorulardan hangisi, testten mutlaka çıkarılmalıdır?

Ç: Madde ayırt ediciliği 0,19 altında olanlar değiştirilmeli fakat burada 3, 4 ve 5. Sorular 0,19 altındadır. Bu durumda en düşük, negatif olan madde değiştirilir. (5. madde)

② Sorulardan hangisi hem kolay hem de ayırt edicidir?

Ç: Kullanılacak sorularda maddelerin; güçlük indeksleri 0,50 civarında ve madde ayırt edicilik gücü +0,30'dan büyük ve mümkün olursa +1,00'e yakın olmalıdır. (2. madde)

ÖRNEK:

	A	B	C	D	E	Bos	N
Üst grup	18	30	10	10	10	22	100
Ait grup	2	26	15	14	10	33	100

① Yukarıda verilen maddede doğru cevap hangi seçeneklerde olsaydı ayırt edicilik en yüksek olur?

$$\begin{aligned} \text{Ç: } A &= \frac{18-2}{100} = 0,16 \\ B &= \frac{30-26}{100} = 0,04 \\ C &= \frac{10-15}{100} = -0,05 \\ D &= \frac{10-14}{100} = -0,04 \\ E &= \frac{10-10}{100} = 0 \end{aligned}$$

② Yukarıda verilen maddenin doğru cevabı D ise; madde ayırt edicilik indeksi kaç olur?

$$\text{Ç: } \frac{10-14}{100} = 0,04$$

Madde Ayırt Edicilik İndeksi Yorumları:

- Madde ayırt ediciliği; bilen ile bilmeyeni birbirinden ayırt etmede işe yarayıp yoramadığı hakkında bilgi verir.
- Madde ayırtıcılık gücü indeksi $-1,00$ ile $+1,00$ arasında değerler alır. Kullanışlımına göre değişir; Başuuran kişi çok seçilecek kişi az ise, sorular ayırt edicilik bakımından $+1$ 'e yakın olmalı.
- Maddelerin kullanılabilirliği hakkında bilgi verir.
- Ayırt edicilik gücü $+1$ 'e yaklaştıkça bilen ile bilmeyen farkı fazladır. (İstenen durum)

Ayırt edicilik gücü 0 'a yaklaştıkça bilen-bilmeyen farkı azdır.

Ayırt edicilik gücü -1 'e yaklaştıkça bilmeyen öğrenciler bu maddeye daha fazla doğru cevap vermiştir. (İstenmeyen durum)

MADDE VARYANSI

Soruyu doğru bilenler ile bilmeyenlerin cevaplarıyla elde edilir. ($P_j \times q_j$)

Test puanlarının varyansı, madde standart sapmasının karesine eşittir.

Madde varyansının $0,25$ olması istenir.

$$S_x^2 = P_j \cdot q_j$$

S_x^2 : Madde varyansı
 P_j : Doğru cevaplayanların oranı
 q_j : Doğru cevaplanmayanların oranı

ÖRNEK

Bir teste 200 öğrenci katılmıştır. 5. soruyu doğru cevaplayan 110 kişidir. Testin madde varyansı kaçtır?

$$P_j = \frac{110}{200} = 0,55$$

$$q_j = 1 - p_j = 1 - 0,55 = 0,45$$

$$S_x^2 = P_j \cdot q_j = 0,2475$$

MADDENİN STANDART SAPMASI

Madde standart sapması, madde varyansının kareköküne eşittir.

$$S_x = \sqrt{P_j \cdot q_j}$$

ÖRNEK: Madde varyansının $0,64$ olduğu bir testin madde standart sapması kaçtır?

$$S_x = \sqrt{P_j \cdot q_j} = \sqrt{0,64} = 0,8$$