

MODÜL 2 (DEĞERLENDİRME NOTLARI)

ÖLÇME VE DEĞERLENDİRME

ÖLÇME VE DEĞERLENDİRMEİNİN TEMEL KAVRAMLARI

Sistem	- En az bir hedefi gerçekleştirmek - Uygun ve değişik öğelerden oluşan dirik bir örüntü
Ögeleri	1) Girdi 2) Süreç 3) Çıktı
Eğitim	Bir sistemdir .
Çünkü	1) Gerçekleştirmek istediği hedefleri var. 2) Hedefleri gerçekleştirmek için farklı öğelerden oluşur.

Hedef Çeşitleri

1) Uzak Hedefler	Devletlerin hedefleridir.
2) Genel Hedefler	Daha işevuruk hedeflerdir. (MEB gibi kurum hedefleri)
3) Özel Hedefler	Dersin hedefleridir. (Kazanımları kapsar.)

EĞİTİMİN ÖGELERİ

1) GİRDİ	Eğitim-öğretim için gerekli her şey <i>Öğrenci, öğretmen, okul, çalışanlar, yönetim, eğitim programı</i>
2) SÜREÇ	-Girdiler, hedefler doğrultusunda bir araya getirilir. -Ders hedeflerini öğrenciye kazandırılan eğitim durumları
3) ÇIKTI	Öğrencilerin sahip olduğu niteliklerdir . Gözlenmeyen süreçler ve ürün değerlendirilemez!

Eğitim Sistemi İçin Çıktılar Nedir?

1) Dar Anlam	Öğrencinin ders sonunda ulaştığı hedef ve kazanımlar
2) Geniş Anlam	- Öğretimin baştan sona kadar sahip olduğu bilgi beceri ve yetkinlik

21. Yüzyıl Değişen Eğitim Paradigması

İnsan Davranışları	a) Bilişsel b) Duyuşsal c) Devinimsel (Psikomotor)
---------------------------	---

Ekonomistler son yıllarda eğitime neden **vurgu** yapıyor?

* Paradigma değişimine uygun insan gücü kaynağı nitelikli eğitimle gerçekleşir.
* Eğitim 1 (bir) birim arttığında => Ekonomik büyüme %0,7 artar. - Bu büyümenin yarısı doğrudan eğitimin katkısı iken - Diğer yarısı eğitim ve ekonomi arasındaki etkileşimden kaynaklanmakta.

Uluslararası

En Geniş Çaplı Eğitim Araştırmalarını Yapan **Örgütler**

1) PISA	2) TIMSS	3) PIRLS
Doğrudan ekonomi veya siyasi-ekonomik politikaların üretirler.		

Dünya Ekonomik Forumuna Göre

2025 Yılında Çalışanların Sahip Olması Gereken **10 Temel Beceri**

1. Analitik düşünme	6. Eleştirel düşünme ve Analiz
2. Yenilik	7. Yaratıcılık
3. Yaratıcılık	8. Orijinallik
4. Etkin öğrenme ve öğrn. stratejileri	9. Liderlik ve sosyal etki
5. Karmaşık problem çözme	10. Teknoloji kullanımı

Dünyanın **Eğitimcilerden Beklediği**

Görev ve Sorumlulukların **Etkileri Nelerdir?**

1) Eğitim ve öğretim ortamlarını yeniden tasarlamamıza
2) Denetleme biçimlerimizi farklılaştırmamıza neden oldu.

DEĞERLENDİRMEİNİN BASAMAKLARI

a. Ölçme	b. Ölçüt	c. Karar
← Köprü görevi görür →		
"Kerim gıda mühendisliği programına yerleştirilmiştir / yerleştirilememiştir." "Zeynep dersten geçmiştir / kalmıştır."		

ÖLÇME, ÖLÇÜT VE DEĞERLENDİRME (KAVRAMSAL TEMELLERİ)

(1) Ölçme	* Öğrencinin belirli bir özelliğini gözleyerek - O özelliğe sayı, sembol ya da sıfat / kategori adı verme ... "Zeynep yapılan test sonucunda 1. oldu."
(2) Ölçüt	Fark: * Ölçme farktan doğmuştur. - Fark, bireyler ya da özellikleri arasında olabilir.
(3) Değerlendirme	* Ölçülen özellik hakkında - Karar alabilmek ve değerlendirme yapabilmek için - Dayanak alınan referans noktası ya da referans aralığıdır . Referans noktası → "60 alan geçer" ... Referans aralığı → "90-109 arası normal zekâdır"
* Bir ölçümü en az bir ölçüte vurarak - Karar verme ve yargıda bulunma işlemidir.	
Eğitim sisteminde ölçmesiz denetleme ve kontrol mekanizması işlemez!	

Ölçüt Değişirse Değerlendirme de Değişir

a. Ölçme	b. Ölçüt	c. Karar
← Köprü görevi görür →		
Ölçme	Ölçüt	Değerlendirme
Irmak Yabancı Dil Bilgisi Seviye Tespit Sınavından (YDS) 50 almıştır.	50	Yüksek lisansa başvuru yapabilir.
Irmak Yabancı Dil Bilgisi Seviye Tespit Sınavından (YDS) 50 almıştır.	55	Doktora başvuru yapamaz.

ÖLÇME TÜRLERİ

(1) Doğrudan Ölçme	* Özellik doğrudan gözlenerek ölçüm elde edilir. * Genellikle beş duyu organı ile algılanan özellikler "Bir sınıftaki öğrenci sayısı" "Öğrencilerin cinsiyetleri- boyları" "Kalem tutma becerisi- Pas verme becerisi"
(2) Dolaylı Ölçme	* Göstergeyle Ölçme * Bazı özellikler doğrudan gözlenemez! (Dolaylı) "Dört işlem becerisi" "Derse yönelik tutumlar" "Genel muhakeme düzeyleri" ...
(3) Türetilmiş Ölçme	* Özellik kendisinden farklı iki ya da daha fazla özelliğin arasındaki Matematiksel bir bağlantı ... * Dört işlem yardımıyla belirlenir. "Hız (yol / zaman)" "yoğunluk (kütle / hacim)"

ÖLÇÜT TÜRLERİ

(1) Mutlak Ölçüt	* Grubun özelliklerinden bağımsız olarak belirlenir. * Genellikle ölçme işlemi öncesinde ilan edilir. * Birinin notu, diğeri etkilemez! "Dersten geçme notu 50'dir", "Doktora için ALES'ten en az 60 almak gerekir"
(2) Bağıl Ölçüt	(Norm Referanslı Ölçüt) * Ölçme grubunun belirli bir özelliğine dayalıdır . * Bağıl ölçüt ancak ölçme işlemi sonrasında belirlenebilir. * Birinin notu, diğeri etkiler! * Rekabetçi bir eğitsel ortam vardır. "Aritmetik ortalamaya denk ve üzerinde puan alan geçer."

NOTLAR

Değerlendirme	* Mutlak ölçüte göre yapılıyorsa → " Mutlak değerlendirme " * Bağıl ölçüte göre yapılıyorsa → " Bağıl değerlendirme " adını alır. "Selim fen lisesine yerleştirilmiştir." → Bağıl değerlendirme "Güzin lisans programları için tercih yapabilecek." → Mutlak d.
* Eğitsel kararların verilmesinde mutlak ve bağıl bir arada kullanılabilir. "Yükseköğretim Kurumları Sınavı" "KPSS Ortaöğretim Memur Atamaları" Ortalama => Bağıl ← 40 Mutlak ya da Bağıl 60 → Mutlak	
* Öğrenci ile ilgili önemli kararlarda → Mutlak değerlendirme gerekir. " Tek ders sınavları " vb....	
- Başvuran kişi sayısının → çok Alınacak kişi → az olduğunda (Arz-talep dengesizliği olan sınavlarda) → Bağıl değerlendirme zorunlu " Yerleştirme / atama " → LGS, YKS, KPSS, TUS vb.	
* Muafiyet sınavlarında → Mutlak değerlendirme yapılması gerekir. "Yabancı dil muafiyet sınavları " vb.	

AMACA GÖRE DEĞERLENDİRME TÜRLERİ		
Öğrenci başarısına odaklı Değerlendirmeler	(1) Tanım ve Yerleştirme (Diyagnostik Değerlendirme)	Şıf içi ölçme ve değerlendirme: * Öğretim sürecinin başında * Hazırbulunuşluluk düzey tespiti... * Amaç not vermek değildir! * " Hazırbulunuşluluk Testi" denir. Okul / kurum geneli ölçme ve değer. * Öğretim sürecinin başında * Tanım - uygun prog. yerleştirme * Niteliklerine uygun gruplar oluşur * Zaman, para - emek tasarrufu sağlar "Muafiyet" ve "seviye tespiti sınav."
	(2) Biçimlendirme ve Yetiştirme (Formatif Değerlendirme)	* Belirli bir konu / ünite vb. sonunda * Davranışların ne kadarını kazandı * Üniteye öğrenme eksiklerini belirle * " İzleme " - " Tarama testi" denir "Notsuz kısa sınavlar (quiz)" "Ünite tarama testleri"
	(3) Değer Biçme (Summatif Değerlendirme)	* Öğretim sürecinin sonunda - Not vermek amacıyla yapılır. * Öğretim sonundaki durum tespiti . * " Alanda erişimi testi" denir. "Bitirme sınavları" "Sertifika sınavları" "Üniversitedeki vize ve finaller",

Yetenek, ilgi, tutum, kaygı, öğrenme kapasiteleri	(4) Rehberlik Amaçlı	1) ÖZEL EĞİTİM: * Belirli alanlarda özel gereksinimi olan çocuk. Özel gereksinim kavramında - Zihinsel gerilik - Öğrenme güçlüğü - Otizm - Diğer spesifik gerilik alanları - İşitme/görme engel alanları - Üstün zekâlılar * " Sandart testler" - " Ölçüt bağımlı testler" * Testlerin sonuçlarıyla BEP hazırlanır. (Bireysel Eğitim Programları) * Tıbbi ve eğitsel / gelişimsel olarak tanı alırlar. * RAM 'lar Eğitsel değerlendirme bağlamında tek yetkili ... 2) MESLEKİ REHBERLİK HİZMETLERİ: * Öğrencilerin alan seçmeleri ve - Bir üst öğretim kurumuna yönlendirilmesinde * Kendi yetenekleri, ilgileri vb. belirlemek... " Standart testler "den yararlanılmaktadır. (5) Program Değerlendirme
		* Eğitim programları bileşenlerinin - Yapısı ve niteliğine ilişkin olarak => Sistematik bilgi toplama ve değerlendirme Amaç Eğitim program ürünlerinin geliştirilmesi

BLOOM
"Bir üniteye öğrenme eksikleri belli bir düzeyin üzerindeyse - O öğrenme açığı kar topu gibi devam eder. - Öğrenemez hâle gelmemesi için ünite sonu kontrol mekanizmaları olmalıdır."

BİR ÖLÇME ARACINDA BULUNMASI GEREKEN PSİKOMETRİK NİTELİKLER	
GÜVENİLİRLİK	* Ölçme işleminden elde edilen puanların - Tesadüfi hatalardan arınık olma derecesidir. - Puanların kararlı, tutarlı ve duyarlı olmasıdır.
KULLANIŞLILIK	* Ölçme aracının → Geliştirilme, uygulanma ve puanlanmasının - Kolay ve ekonomik olmasıdır. (zaman, para, emek, araç gereç vb. açıdan)
Bir ölçmenin Geçerli ve güvenilir olmasının tek koşulu: Puanların hatasız ya da az hatalı olmasıdır.	

Ölçme İşlemlerinde Hataları En Aza İndirebilmek İçin
Öncelikle hataların tanımlanması gerekir: Ölçme işlemi yapan kişi neye müdahale edeceğini

HATALARIN DÖRT KAYNAĞI?	
HATA	1) Ölçme İşlemini Yapan Kişiden Kaynaklanan Hata 2) Ölçme Aracından Kaynaklanan Hata 3) Bireyden / Öğrenciden Kaynaklanan Hata 4) Fiziksel Ortamdan Kaynaklanan Hata

HATA ÇEŞİTLERİ		
Ölçme sonuçlarına karışan Hata miktarı, yönü ve kaynağı bellidir	1) SABİT	* Miktarı ölçmeden ölçmeye değişmez! * Her ölçme işlemine aynı miktarda karışır. "Bir terazinin boş ağırlığı 120 g ise Her tarttığımız 120 g eksik ölçülecektir."
	2) SİSTEMATİK	* Miktarı ölçmeye göre değişen hatalardır. - Ölçülen büyüklüğe - Öğretmene - Ölçme koşullarına göre hata değişir. "Terazi, her bir kilogramda 120 g eksik tartıyorsa Ağırlığı arttıkça hata miktarı da artacaktır."
Ölçme sonuçlarına karışan hata miktarı, yönü ve kaynağı bellidir. Hatayı engellemek -düzeltmek daha kolaydır.		

hata miktarı, yönü ve kaynağı gen. belli değildir!	3) TESADÜFİ	* Tesadüfi (Rastlantısal) hata * Şansla (tesadüfen) ortaya çıkar. * Ne yönde ve ne ölçüde karıştığı => Genellikle bilinemeyen hatalardır. * Ölçme sonuçlarına tek yönlü olarak karışmaz! → Bazen Pozitif bazen de Negatif yönde etki eder
	Ölçme sonuç. karışan hata miktarı, yönü ve kaynağı genellikle belli değildir! Hatayı engellemek -düzeltmek daha zordur!	

* Sabit ve tesadüfi olmayan hatalar → Tesadüfi değişken olma özelliğindedir. (Psikometride hata kuramı tesadüfi hatalar üstüne kurulmuştur.)
--

TESADÜFİ HATALARIN DÖRT KAYNAĞI?
1) Ölçme İşlemini Yapan Kişiden / Öğretmenden Kaynaklanan Hata:
1. Ölçme işlemi sürecinde: - Öğretmen, test katılımcısının dikkatini dağıtabilir - kaygısını artrabilir 2. Ölçme işlemi sonrasında: - Puanlamadaki dikkat ve titizliğin değişmesi, yorgunluk, maddi hata vb.
2) Ölçme Aracından Kaynaklanan Hata:
1. Ölçme araçları hazırlanırken maddelerin iyi ifade edilmemesi 2. Araçta kullanılan dil, dil bilgisi ve imla sorunları 3. Maddelerde anlatım bozukluğu bulunması 4. Muğlak (kapalı) - belirsiz (müphem) ifadeler 5. Ölçme aracının yapısı sorunu - Homojen (benzeşik) maddelerden oluşan bir araç(test) - Heterojen (ayrışık) maddelerden oluşan bir testten daha güvenilir dir. 6. Ölçme aracının uzunluğu sorunu * Madde sayısı ile güvenilirlik arasında doğru orantılı ilişki vardır. * Ölçme aracındaki madde sayısı arttıkça - Yorgunluk, dikkat azalması vb. sebebiyle hata miktarı artar. 7. Seçenekli testlerde Şans başarısı faktörü * Bilişsel özellikleri ölçen testlerde - Eğer seçenek verilmişse şans başarısı karışma olasılığı artar. (Şans başarısı : Salt tahminle doğru yanıt bulma olasılığı)
3) Bireyden / Öğrenciden Kaynaklanan Hata:
* Bireylerin fiziksel, fizyolojik ve psikolojik durumları (Uykusuzluk -açlık) * Grubun homojen ya da heterojen olması Daha heterojen gruplarda güvenilirlik katsayısı → Artar Daha homojen gruplarda güvenilirlik katsayısı → Azalar.
4) Fiziksel Ortamdan Kaynaklanan Hata:
* Ölçmenin yapıldığı fiziksel ortamdan kaynaklanan hatalardır: " Koku, görsel uyarcıların fazlalığı, ergonomi vb.... "



GÜVENİLİRLİK → GEÇERLİLİK için bir ön şarttır

Ancak Yeterli şart değildir.

GÜVENİLİR olması → **GEÇERLİ** olacağı anlamına gelmez!
GEÇERLİ ise → Büyük olasılıkla **GÜVENİLİR**'dir.

Ölçme aracının geçerliliği ve güvenilirliği diye bir şey yoktur!
 Ölçme araçlarından elde edilen "Puan" → - Geçerli ve güvenilir olabilir.

Geçerlilik ve güvenilirlik Varlık-yokluk sorunu değil! (Derece sorunudur)

* Bir testten elde edilen puanlar :
 Yüksek / Orta / Düşük düzeyde geçerli
 Yüksek / Orta / Düşük düzeyde güvenilir
 (biçiminde nitelendirilir.)

Ölçme Araçlarının Psikometrik Nitelikleri Şu Durumlara Göre Değişebilir:

1. Ölçme amacının değişmesi	4. Madde değişiklikleri
2. Uygulama grubunun değişmesi	5. Farklı kültürler
3. Dilin eskimesi / değişmesi	6. Kuramsal bilgi değişiklikleri

KORELASYON

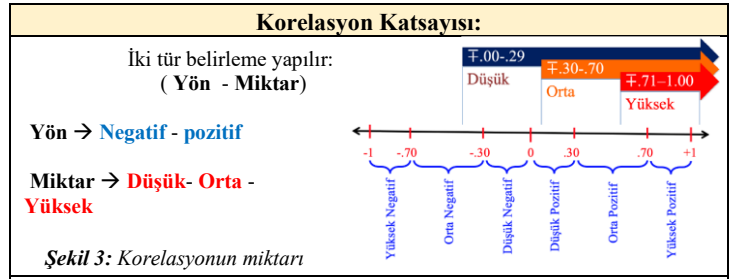
KORELASYON:
 * En az iki değişken arasında karşılıklı bir ilişkinin varlığını inceler. İlişkinin yön ve miktar bilgisini veren **istatistik** bir tekniktir.
 "Öğrencilerin ders tutumları ile başarıları arasında bir ilişki var mıdır?"
 "Saç uzunluğu ile zekâ arasında bir ilişki var mıdır?"

* Geçerlilik ve güvenilirlik genellikle **korelasyon temelli analizlerle** bulunur.
 * **Korelasyon**(co-relation) ("r" ile sembolize edilir.)

* **Korelasyon** "-1" ile "1" arasında değer alır.

a) **Pozitif korelasyon** → İki değişken arasında **doğru orantılı** ilişkidir.
 "Ders çalışma süresi ile sınav notu" → Pozitif bir korelasyon
 b) **Negatif korelasyon** → İki değişken arasında **ters orantılı** ilişkisidir.
 "Ders süresi ile dikkat → Negatif bir korelasyon

* **±1** Korelasyon → **Mükemmel korelasyon**
 1 Korelasyon → **Mükemmel doğru orantılı ilişki**
 -1 Korelasyon → **Mükemmel ters orantılı ilişki**
 0 (sıfır) Korelasyon → **İki değişken arasında sistematik bir ilişki yok!**
 → (Mükemmel ilişkisizlik)
 "Öğretmen boyları ile aylık ücretleri" → Arasındaki korelasyonun sıfırdır.



Not 1:
 * Bir korelasyon katsayısının **negatif** ya da **pozitif** olması
 ⇒ **Büyük-küçük belirtmez!** (Yön bildirir.) (Mutlak Değer)

Not 2:
 * **Korelasyon katsayısı ile neden-sonuç ilişkisi kurulamaz!**

GEÇERLİLİK SORGULAMA YÖNTEMLERİ

1) KAPSAM GEÇERLİLİĞİ:
 * Özellikle başarı testlerinde aranan bir geçerlilik sorgulamasıdır.
 * Bir testin ölçülmek istenen davranışları ne derece kapsadığıyla ilgilidir.

2) ÖLÇÜT DAYANAKLI GEÇERLİLİK:
 * Puanların ölçüt bir puanla karşılaştırılması ile araç geçerlilik ilişkisi

3) YORDAMA GEÇERLİLİĞİ:
 * Yordama, tahmin demektir. (Ancak her tahmin yordama değildir!)
 (Geçerli ve güvenilir bir veri → Bu verinin sınanabilir olması gerekiyor.)
 * Yordama, eldeki bu nitelikteki veriden yola çıkarak,
 → Geleceğe, henüz gerçekleşmemiş bir olguya ilişkin yapılan tahmindir.

	YKS Puanı	Akademik Ortalama
Tarik	438.636	3.54
Esra	348.545	2.45
Begüm	456.344	3.39
Ozan	399.987	2.23
Defne	382.654	2.34
.....	480.334	3.45
.....	287.987	1.67
.....	381.954	3.88
.....	250.897	3.91
Umud	343.019	3.26

* Seçme / yönlendirme amaçlı testlerde Y.G. → **Yüksek** olmalı

Seçme amaçlı testler YKS, KPSS, LGS vb.
Yönlendirme amaçlı Yetenek testleri
İlgi envanterleri

Ölçüt gelecekte belli olacağından beklemek gerekmez!
 Korelasyon (r) 1'e yakın
 Geçerlilik Katsayısı

Şekil 5: Bir yordama geçerliliği çalışma örneği

* **Ölçüt puan** testin tahmin etmeye çalıştığı → **Özelliktir.**
 * **Ölçüt puan** gelecekte belli olacağından → **Beklenir**

4) ZAMANDAŞ GEÇERLİLİK:

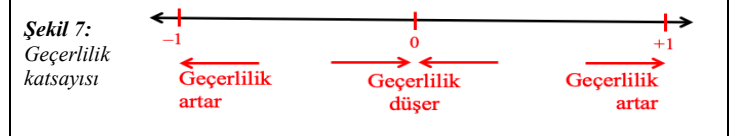
* Diğer adları: **Hâlihazır / Benzer ölçekler / Uygunluk** geçerliliği
 * **Ölçüt puan** eş zamanlı olarak elde edilebilir.
 ⇒ Geçerliliği yüksek bir **aracın puanı** → **Ölçüt puan** olarak ele alınabilir.
 ⇒ Sonra **ölçme aracı** ve **ölçüt araç aynı zamanda** uygulamış
 - İki araçtan elde edilen puanlar arasındaki ilişki incelenir....

Şekil 6: Bir zamandaş geçerlilik çalışma örneği

	Benlik Algısı Ölçeği Puanları	Ölçüt Duan Öz-saygı Ölçeği Puanları
İrem	54	32
Metin	98	54
Erdal	128	98
Dilek	89	65
Çaydaş	101	54
---	87	56
---	98	58
---	67	65
---	122	80
Sedef	79	88

Daha önce geliştirilmiş ve geçerliliğinin yüksek olduğu bilinen araç
 Korelasyon (r) 1'e yakın
 Geçerlilik Katsayısı

* Z.G. te **geçerlilik katsayısının -1 ile 1 arasında değişmesi** beklenir.
 * Ölçüte bağlı olarak **geçerlilik katsayısı**
 → **±1'e** yaklaştıkça **artar**, **0'a** yaklaştıkça **düşer**.



5) YAPI GEÇERLİLİĞİ:

Yapı geçerliliği,
 -Bir testin dayandığı kuramsal temelleri ne derece iyi örneklediğiyle ilgilidir.

* **Yapı geçerliliği**
 - Testin ölçtüğü **niteliklerin neler olduğunu araştırma** ve
 - Kişilerin elde ettikleri **puanların ne anlama geldiğini** açıklama çabasıdır.