

MODÜL 2

ÖLÇME VE DEĞERLENDİRME

ÖLÇME VE DEĞERLENDİRMENİN TEMEL KAVRAMLARI

Dirik: Canlılığını sürdüren, sürekli kendisini yenileyen ve gelişen.

Örüntü: Ögelerden oluşan ve aralarında ilişki bulunan bir bütün.

İşevrük: Doğrudan gözlemediğimiz bir kavramın var olup olmadığını saptayabilmek için ne gibi işler ya da işlemler yapıldığını belirterek tanımlama.

Paradigma: Bir şeyin nasıl üretileceği konusunda örnek, model.

Sistem

- En az bir **hedefi** gerçekleştirmek üzere
- Uygun ve değişik öğelerden oluşan **dirik** bir örüntüdür.
- * Bir bütüne sistem diyebilmek için **girdi, süreç ve çıktı** öğeleri olmalıdır.

* Eğitim de bir **sistem**dir.

Birinci Neden: Gerçekleştirmek istediği hedefler bulunmaktadır.

İkinci nedeni: Bu hedefleri gerçekleştirmek için farklı öğelerden oluşur.

Hedefler (Nedir?)

1. **Uzak hedefler:** Devletlerin hedefleridir.
2. **Genel hedefler:** Daha işevrük hedeflerdir.
 - MEB, MEB genel müdürlükleri, okul yönetimleri gibi kurumların hedefleridir.
3. **Özel hedefler:** Dersin hedefleridir. (Kazanımları kapsar.)

EĞİTİM FARKLI ÖGELERDEN OLUŞUR: (GİRDİ – SÜREÇ – ÇIKTI)**1. Girdi:**

- * Eğitim-öğretim için gerekli her şey eğitim sisteminin girdisini oluşturur.
- Örn. Öğrenci, öğretmen, okul, çalışanlar, yönetim, eğitim programı vb.

2. Süreç:

- **Girdilerin hedefler** doğrultusunda bir araya getirildiği,
- Dersin **hedeflerinin** öğrenciye kazandırıldığı **eğitim durumlarını** kapsar.
- * **Süreç, öğretmenler, okul yönetimi, il yöneticileri kontrol edilir:**
 - Ülke genelinde **belirli aralıklarla kontrol** edilmektedir.
 - Kontroller **olumlu** ise süreç **devam eder** değilse sürece **müdahale edilir**.

3. Çıktı:

- * **Öğrencilerin** sahip olduğu **niteliklerdir**.
- * **Eğitim sistemi için çıktılar nedir?**
 - a) **Dar anlam:** Öğrencinin ders sonunda ulaştığı hedefler ve kazanımlar
 - b) **Geniş anlam:** Öğretimin baştan sona kadar sahip old. bilgi beceri ve yetkinlik...
- * **Yönetimler girdiyi kontrol ederler.**
 - Eğitim sisteminde **hazırbulunuşluk** dar anlamda eğitim sisteminin **girdileridir**.
 - Öğretmenler **hazırbulunuşluk** girdisini ölçerek **kontrol ederler**.
 - **Çıktılar** hedeflerle **tutarlı**ysa eğitim sistemi doğru bir şekilde işliyordur.
 - Eğitimde **kontrol (denetleme) ölçme ve değerlendirme** ile mümkündür.
- * Gözlenmeyen süreçler ve ürün **değerlendirilemez!**

21. YÜZYIL DEĞİŞEN EĞİTİM PARADİGMASI:

- * Bu yüzyılda eğitimde ve ölçme ve değ.de bir **paradigma değişimi** söz konusudur.
- * 20. yüzyıla hâkim olan güç **kültürel iklim ve sanayileşme** idi.
 - **Eğitim** bu iklimle uyumlu biçimde **şekillenmişti**.
 - ABD’de 1930-1950’lerde **okullara “eğitim fabrikaları”** dendi.
- * **Ülkemizde** de o yıllarda **benzer bir anlayışın** vardı.
- * Atatürk’ün ortaya koyduğu **“muasır medeniyetler seviyesine ulaşmak”** fikri
 - ⇒ Türkiye’nin **hedefi** ve bir nevi eğitimin de **uzak hedefi** oldu.
- * **Eğitimin genel hedefi** de pek çok ülkede kısaca **“Toplunun bireylerden beklediği görev ve sorumluluklar doğrultusunda gerekli bilgi ve donanımına sahip olma”** olarak tanımlandı.

- * Geçtiğimiz yüzyılın özellikle son çeyreğinde
 - Bilgi ve iletişim teknolojisinin önemi arttı.
 - Toplumlar, üretimde ve eğitimde bir paradigma değişimi oldu.
 - ⇒ **Sanayi toplumlarından bilgi toplumlarına** hızlı bir **dönüşüm** gerçekleşti.
 - ⇒ Üretimdeki **değişim** eğitim-öğretim ortamlarına da **yansımaya** başladı.

- * **Beşerî bilimlerde** (eğitim bilimlerinde) **insan davranışları**, birbirinden ayrı kümelermiş gibi ele alındı, araştırıldı, hakkında bilgi üretilmeye çalışıldı.
 - [**İnsan Davranışları** : Bilişsel, Duyuşsal ve Devinimsel (Psikomotor)]
 - ⇒ Halbuki insan bu parçaların toplamından daha anlamlı bir varlıktı.

- * 20. yüzyılın **son çeyreğinde** ve özellikle **21. yüzyılda**
 - Ekonomi çevreleri ve ekonomistler **eğitime** yoğun bir **ilgi** göstermeye başladı.

Ekonomistler son yıllarda eğitime neden daha çok vurgu yapıyorlar?

- * Araştırmalar, **paradigma değişimine uygun insan gücü kaynağının**
 - Ancak **nitelikli eğitimle** gerçekleşebileceğini ortaya koymakta.
 - **Niteliğin ölçülmesi çok zor**
 - Ancak** nicel veriler açısından değerlendirdiğimizde bir ülkede
 - ⇒ Ortalama **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.

PISA, TIMSS ve PIRLS

- Dünyanın en geniş çaplı uluslararası eğitim araştırmalarını yapan **örgütler**dir.
- ⇒ Doğrudan **ekonomi** veya **siyasi-ekonomik politikaların** üretilen **örgütler**dir.

Dünya Ekonomik Forumuna Göre**2025 Yılında Çalışanların Sahip Olması Gereken 10 Temel Beceri:**

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

- ⇒ Dünyanın eğitimcilerden **beklediği** görev ve sorumlulukların **etkileri** nelerdir?
 - Eğitim ve öğretim **ortamlarını yeniden tasarlamamıza**
 - **Denetleme biçimlerimizi farklılaştırmamıza** neden oldu.
 - Bu **değişim devam** etmektedir.

* Eğitimde **denetleme (kontrol) ⇒ Ölçme ve değerlendirme** ile olanaklıdır.

*** Değişen hedeflerde**

- Kazandırılacak olan yeni **davranışların** ölçülmesi ve değerlendirilmesinde
- ⇒ Doğal olarak bir **paradigma değişimi** söz konusu.
- ⇒ Geleneksel ölçme anlayışı bu **paradigma değişimine** pek uygun değildir.

* Bireylerden beklenen becerilerin **karmaşıklığı** arttıkça

⇒ Ölçme modellerini **revize etme** gereksinimi ortaya çıkmıştır.

Bir ölçme işleminde yalnızca **“görünce tanıma, sorulunca söyleme”** gibi davranışların ölçülmesi bizi **yeni hedeflere ne kadar ulaştırabilir?**

- 20. yüzyıl eğitim ve ölçme **zihniyeti** ile 21. yüzyıl becerileri **kazandırılmaz!**

ÖLÇME, ÖLÇÜT VE DEĞERLENDİRME (KAVRAMSAL TEMELLERİ)**ÖLÇME:**

- * Öğrencinin belirli bir **özellikliğini** gözleyerek
- O özelliğe **sayı, sembol** ya da **sıfat/kategori adı verme** işlemidir.
- “Zeynep yapılan test sonucunda 1. oldu.”
- “Kerim Temel Yeterlilik Testinden (TYT) 450 puan aldı.”

Ölçmede

- * Gözlenen özelliğin bir **betimlemesi** yapılır.
- ⇒ Bu nedenle ölçme **tek başına bir anlam ifade etmez**.
- ⇒ Ölçmenin **anlam** ifade etmesi **değerlendirme** ile olur.

* Ölçmeye **temel** olan durum **“fark”**tır. (**Ölçme farktan doğmuştur.**)

Fark

1. **Bireyler arasında** olabilir.
2. Bireylerin farklı zamanlarda ortaya koydukları **özellikler arasında** olabilir.

- * Eğitimde **ölçme** dediğimizde aklımıza **ilk gelen** genellikle **sınavlardır**.
- ⇒ Eğitimde **ölçme** yalnızca **sınav** (akademik başarı) dan ibaret **değildir**.
- ⇒ **Yetenek, ilgi, tutum, özel gereksinim gereken alanlar** ölçülebilir.

* Bir eğitim sisteminde **ölçme ve değerlendirme** yoksa

- * Eğitim sisteminin **geçerli ve güvenilir kontrolü sağlanamaz**.
- * Bu **kontrol**, sistemin tüm öğeleri için **geçerlidir**.
 - Ölçülmeyen, kontrol edilmeyen **süreçler denetlenemez**
 - Denetlenemeyen **süreçler** ise **değerlendirilemez**.
 - Ölçme yapılmadan sisteme yönelik **girdi, süreç ve çıktılar kontrol edilemez**.

⇒ Eğitim sisteminde **Ölçmesiz denetleme ve kontrol mekanizmasını işlemez!**

ÖLÇÜT:

- * Ölçülen özellik hakkında
 - **Karar** alabilmek / **yargıya** varabilmek / **değerlendirme** yapabilmek için
 - **Dayanak** alınan **referans noktası** ya da **referans aralığıdır**.
 - Referans noktası** → “60 alan geçer” ...
 - Referans aralığı** → “90-109 arası normal zekâdır”, “DC alan şartlı geçer”

DEĞERLENDİRME:

- * Bir ölçme sonucunu (ölçüm) en az bir **ölçüte** vurarak
 - **Ölçülen nitelik hakkında karar verme / yargıda bulunma** işlemidir.

DEĞERLENDİRMENİN BASAMAKLARI

a. **Ölçme** ⇒ Ölçme sonucu (ölçüm) b. **Ölçüt** c. **Karar**

“Kerim gıda mühendisliği programına yerleştirilmiştir / yerleştirilememiştir.”
“Zeynep dersten geçmiştir / kalmıştır.”

- ⇒ **Ölçüt, ölçme ve değerlendirme** arasında **köprü görevi görür**.
- **Ölçüt** değişirse **değerlendirme** de değişir. Örneğin

Ölçme	Ölçüt	Değerlendirme
İrmak Yabancı Dil Bilgisi Seviye Tespit Sınavından (YDS) 50 almıştır.	50	Yüksek lisansa başvuru yapabilir.
İrmak Yabancı Dil Bilgisi Seviye Tespit Sınavından (YDS) 50 almıştır.	55	Doktora başvuru yapamaz.

1) Doğrudan Ölçme 2) Dolaylı Ölçme 3) Türetilmiş Ölçme

Doğrudan Ölçme (Temel Ölçme):

- * Ölçmeye konu olan **özellik doğrudan** gözlenerek **ölçüm** elde edilir.
- * Genellikle **beş duyu organı** ile algılanan özellikler **doğrudan** gözlenebilirler.
“Bir sınıftaki öğrenci sayısı” → **Sayılabılır**
“Öğrencilerin cinsiyetleri- boyları” → **Görülebilir**
“Kalem tutma becerisi- Pas verme becerisi” → **Beceriler** doğrudan gözlenebilir.
- * Pek çok **psikomotor davranış** doğrudan gözlenebilir.
“Yanlışsız okuyabilme”, “diksiyon”
⇒ Eğitimde ölçmeye konu olan özelliklerin **pek azı doğrudan** ölçülebilir.

Dolaylı Ölçme (Göstergeyle Ölçme):

- * **Bazı özellikler** doğrudan **gözlenemez!**
- * Bu özellik onun **göstergesi** olan **davranışlar** aracılığıyla **gözlenerek ölçülebilir**.
“Dört işlem becerisi” “derse yönelik tutumlar”, “genel muhakeme düzeyleri”,
“sınav kaygıları”, “motivasyon düzeyleri” → Ancak **dolaylı** olarak **gözlenebilir**.

- * “Sınıfta durumuyla karşılaştığınızda nasıl davranırsınız?” sorusuna
- Öğretmenin gelecekte **nasıl davranacağı**nın bir **göstergesi** olduğu kabul edilir.
 - * “İyi yurttaş olma” → Doğrudan gözlenemez!
İyi yurttaş olmanın **göstergesi** “vergi vermesi” → ona “iyi yurttaş” dedirtir.
- Not:** 18 yaşın altında bir öğrencinin bu özelliklere sahip olma durumu
⇒ **Varsayımsal** durumlar üzerinden **gözlenebilir**.

Türetilmiş Ölçme:

- * Ölçülmek istenilen **özellik**
- Kendisinden farklı **iki** ya da **daha fazla** özelliğin arasındaki
- **Matematiksel** bir bağıntı (**dört işlem**) yardımıyla belirlenir.
“Hız (yol / zaman)”, “yoğunluk (kütle / hacim)”

ÖLÇÜT TÜRLERİ

1) Mutlak Ölçüt 2) Bağlı Ölçüt

Mutlak Ölçüt (Kriter Referanslı Ölçüt):

- * **Ölçüt**
- Ölçme işleminin yapıldığı grubun özelliklerinden **bağımsız** olarak belirlenir.
- **Mutlak ölçüt** genellikle **ölçme** işlemi **öncesinde** ilan edilir.
“Dersten geçme notu 50’dir”, “doktora için ALES’ten en az 60 almak gerekir”
- * **Mutlak ölçütün** temel alındığı ölçme işlemlerinde
- **Birinin** notu, **diğerinin** değerlendirilmesini (geçme / kalma) **etkilemez!**
- Başarının ölçüldüğü bir **örnekte** herkes **geçebileceği** gibi herkes **kalabilir** de.

Bağlı Ölçüt (Norm Referanslı Ölçüt):

- * **Ölçüt**
- Ölçme işleminin yapıldığı grubun **belirli bir özelliğine / normuna** dayalıdır.
- **Bağlı ölçüt** ancak **ölçme** işlemi **sonrasında** belirlenebilir.
“Aritmetik ortalamaya / ortancaya / moda denk ve üzerinde puan alan geçer.”
- * **Bağlı ölçütün** temel alındığı ölçme işlemlerinde
- **Birinin** notu, **diğerinin** değerlendirilmesini (geçme / kalma) **etkiler!**
- Bu anlamda öğrenciler **açısından** daha **rekabetçi** bir eğitsel ortam vardır.

Not 1:**Değerlendirme**

- **Mutlak ölçüte** göre yapılıyorsa “**mutlak değerlendirme**”
- **Bağlı ölçüte** göre yapılıyorsa “**bağlı değerlendirme**” adımı alır.
“Selim fen lisesine yerleştirilmiştir.” → **Bağlı değerlendirme**
“Güzin lisans programları için tercih yapabilecektir.” → **Mutlak değerlendirme**

Not 2:

- Eğitsel kararların verilmesinde **mutlak** ve **bağlı** ölçütler **bir arada kullanılabilir**.
“Yükseköğretim Kurumları Sınavı”, “KPSS Ortaöğretim Memur Atamaları”
Ortalama
- 40-60 ise → **Mutlak** ya da **Bağlı** değerlendirme yapılır
- 60’ın üzerindeyse → **Mutlak** değerlendirme yapılır,
- 40’in altındaysa → **Bağlı** değerlendirme **zorunludur**.

Not 3:

- * Öğrenci ile ilgili **önemli / hayati** kararlarda → **Mutlak** değerlendirme gerekir.
“Tek ders sınavları” vb.....

Not 4:

- Başvuran kişi sayısının **çok** / Alınacak kişi sayısının **az** olduğunda
(Arz-talep dengesizliği olan) sınavlarda → **Bağlı** değerlendirme zorunludur.
“Yerleştirme / atama” → LGS, YKS, KPSS, TUS vb.

Not 5:

- * **Muafiyet sınavlarında** → **Mutlak** değerlendirme yapılması gerekir.
“Yabancı dil muafiyet sınavları” vb.

AMACA GÖRE DEĞERLENDİRME TÜRLERİ

1) Tanıma ve Yerleştirme 2) Biçimlendirme ve Yetiştirme 3) Değer Biçme
4) Rehberlik Amaçlı 5) Program Değerlendirme**1) Tanıma-Yerleştirmeye Yönelik Değerlendirme (Diyagnostik Değerlendirme):****Sınıf içi ölçme ve değerlendirme:**

- * Öğretim sürecinin **başında**
- Öğrencilerin **önceki kritik nitelikteki** öğrenmelerini belirlemek amacıyla yapılır.
- * Hedef öğrencinin ders bağlamındaki **hazırbulunuşluluk** düzeyini belirlemektir.
- * Amaç **not vermek değildir!**
- * Bu amaçla uygulanan ölçme araçlarına da “**hazırbulunuşluluk testi**” adı verilir.

Okul / kurum geneli ölçme ve değerlendirme:

- * Yine öğretim sürecinin **başında**
- Öğrenciyi tanımak ve onu uygun olan programa yerleştirmek amacıyla yapılır.
- * Amaç **not vermek değildir!**
- * Hedef öğrencinin niteliklerine uygun grupların oluşturulmasıdır.
- * **Zaman, para** ve **emek** tasarrufu sağlar.
- “Muafiyet sınavları” ve “seviye tespit sınavları” buna örnek gösterilebilir.

2) Biçimlendirme-Yetiştirmeye Yönelik Değerl. (Formatif Değerlendirme):

- * Belirli bir konu / ünite vb. **sonunda**
- Öğrencilerin öğretime konu olan davranışların ne kadarını kazandığını,
- Diğer bir deyişle **ünitedeki** öğrenme **eksiklerini belirlemek amacıyla yapılır**.
- * Amaç **not vermek değildir!**
- * **Öğrenciye**
- O ünite **deki** öğrenme **eksikleri** bildirilir,
- Bir sonraki kapsama geçmeden öğrenme **eksikleri giderilir/ denetlenir**.
- * Bu değerlendirmede kullanılan testlere “**izleme testi**” ya da “**tarama testi**” denir.
“Not verilmeden yapılan kısa sınavlar (quiz)”, “ünite tarama testleri”
- * Bir ünite **deki** öğrenmeler bir **başka ünite** için **ön öğrenme** olabilir
- * Ön öğrenme **eksiklerin giderek artması**
- Öğrencinin **öğrenemez** hâle gelmesine kadar **gidebilir**.

BLOOM’a göre

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

3) Değer Biçmeye / Düzey Belirlemeye Yönelik Değ. (Summatif Değerlendirme):

- * Belirli bir öğretim sürecinin **sonunda**
- **Not vermek** amacıyla yapılır.
- * Öğrencinin belirli bir öğretim süreci **sonundaki durum tespitidir**.
- * Bu amaçla uygulanan testlere “**alandaki erişim testi**” denir.
“Bitirme sınavları”, “sertifika sınavları”, “üniversitedeki vize ve finaller”,
“eğitim-öğretim kurumlarında not vermek amacıyla uygulanan yazılı yoklamalar”

Not: Yukarıdaki ilk üç değerlendirme türü → **Öğrenci başarısına** odaklıdır.

- Öğrencinin akademik **başarısının ölçülüp değerlendirilmesi** söz konusudur.
- * Eğitimde **yalnızca akademik başarının ölçümü - değerlendirmesi olmaz!**
Öğrencilerin yetenekleri, ilgileri, tutumları, kaygıları, öğrenme kapasiteleri ve karşılaşılan güçlükler ile uygulanan eğitim programı etkililik değerlendirilmesi...

4) Rehberlik Amaçlı Değerlendirme:

- * İki bağlamda ele alınabilir: 1) **Özel eğitim** 2) **Mesleki rehberlik hizmetleri**

ÖZEL EĞİTİM:

- Belirli alanlarda **özel gereksinimi** olan çocuklara uygun **eğitim ortamları**
- Programı uygulamak amacıyla **ölçme** ve **değerlendirme süreçleri** vardır.
- * **Özel gereksinim kavramında**
- **Zihinsel gerilik / Öğrenme güçlüğü / Otizm / Diğer spesifik gerilik alanları**
- **Üstün zekâ / işitme/görme engel alanları** → Çok geniş bir yelpaze vardır.
- * **Özel gereksinimi olan çocuklara** uygun eğitim olanaklarının sunulabilmesi için
* Çocuğun **ihtiyacı** olan **alanlar** belirlenmeye çalışılır. (Eğitsel tanılama):
- **Standart testler** veya **ölçüt bağımlı testler** uygulanır.
- * Testlerin sonuçları doğrultusunda **BEP** (Bireysel Eğitim Programları) hazırlanır.
- Bu çocuklar **tıbbi ve eğitsel / gelişimsel** olarak **tanı** alan çocuklardır.
- * Eğitsel **değerlendirme** bağlamında **RAM**’lar bu konudaki **tek yetkilidir**.

MESLEKİ REHBERLİK:

- * Öğrencilerin **alan seçmelerinde** ve bir üst öğretim kurumuna **yönlendirilmesinde**
- Özelliklerine **uygun** alan / meslek seçmeleri
- Öğrenci ve ülkenin insan gücü kaynağının doğru **planlanmasında** çok önemlidir.
- * Öğrencilerin kendi yetenekleri, ilgileri vb. ile özelliklerini belirlemek amacıyla
⇒ **Standart testlerden** yararlanılmaktadır.

5) Program Değerlendirme:

- * **Eğitim programları** bileşenlerinin yapısı ve niteliğine ilişkin olarak
- **Sistematik** bilgi **toplama** ve **değerlendirme** sürecidir.
- * **Program değerlendirme**
- Eğitim programlarının **planlanması**,
- Mevcut program ürünlerinin etkililiğinin **değerlendirilmesi**
- Eğitim program ürünlerinin **geliştirilmesi** amacıyla yapılır.

BİR ÖLÇME ARACINDA BULUNMASI GEREKEN PSİKOMETRİK NİTELİKLER

* **Ölçmenin** her zaman belirli bir **amacı** vardır:
- Kişiler hakkında değerlendirme yapmak
- Elde edilen değerlendirme sonuçlarına göre belirli kararlar vermek için yapılır.

* **Ölçüm**, değerlendirmenin **dayanağıdır**.
- Verilen **kararların doğru ve isabetli** olması
- Ölçümün **doğrudan ilgili** ve olabildiğince **az hatalı** olmasına **bağlıdır**.

* **Psikolojik** özelliklerin ölçülmesinde genellikle **ölçme araçları** kullanılır.
* Bu **araçlardan** elde edilen **puanlar** → **Hatasız** ya da **az hatalı** olmalı
* Bu **araçların** → Belirli **psikometrik** niteliklere sahip olması **gerekir**.
⇒ Ölçülen özelliğin başka **doğru ve tam** olarak ölçülebilmesidir.
⇒ Ölçme aracından elde edilen **puanın** → Amaca **hizmet etme derecesidir**.

Güvenilirlik:

* **Ölçme** işleminden elde edilen **puanların**
- Tesadüfi hatalardan **arınık** olma **derencesidir**.
- **Puanların kararlı, tutarlı ve duyarlı** olmasıdır.

Kullanışlılık:

* **Ölçme** aracının **geliştirilmesinin, uygulanmasının ve puanlanmasının**
- **Kolay ve ekonomik** (zaman, para, emek, araç gereç vb. açıdan) olmasıdır.

* **Geçerli ve güvenilir ölçme** yapmanın tek koşulu:
- Bir ölçme **aracından** elde edilen **puanların** **hatasız** ya da **az hatalı** olmasıdır.

* **Ölçme** işlemlerinde **hataları en aza indirebilmek için**
- Öncelikle **hataların tanımlanması** gerekir:
- Ölçme işlemini yapan kişi **neye müdahale edeceğini**
- Hangi durumlara karşı **önlem alması gerektiğini** bilmelidir.

HATA:

* Sadece eğitimde değil bütün hayatımızdaki ölçmelerimize **hata karışabilir**:

* Bu hataların Kaynakları Nelerdir?

- Ölçme yapan kişilerin dikkati ve titizliği zamandan zamana **değişebilir**
- Ölçme aracına ilişkin bazı **sorunlar** söz konusu olabilir
- Ölçülen özelliğin doğası gereği bazı **sıkıntılar** olabilir
- **Test katılımcısı** yeterince güdülenmemiş olabilir
- Ölçme işleminin yapıldığı **ortamdan** kaynaklı sorunlar vb. **olabilir**.

Ölçmede

* **Gözlenen** bir özelliğin **gerçek değeri** (sayı, sembol ya da sıfat) **bulunmak** istenir.
- Ancak ölçmedeki **hatalar, gerçek değer (puan)** yoluyla doğrudan elde **edilemez**.
- **Bu kavram**, psikometride **gerçek puan** kuramı olarak ele alınır.
- **Bu kavram** basit bir **eşitlik** ile gösterilir.

HATA ÇEŞİTLERİ

Hatalar; **kaynağı, yönü ve miktarı** göz önünde tutularak

a) Sabit b) Sistematik c) Tesadüfi olmak üzere üç türde incelenebilir:

a) Sabit Hata:

- **Miktarı** ölçmeden ölçmeye **değişmeyen**
- Her ölçme işlemine **aynı miktarda** karışan hatalardır.
“ Bir terazinin boş ağırlığı 120 g ise → Her tarttığını 120 g **eksik** ölçülecektir.”
“ Sınavda herkese 10 puan fazla verirse → **Yine karışan hata sabittir.**”

b) Sistematik Hata:

- Ölçülen **büyükliğe, öğretmene** ya da ölçme **koşullarına** göre
⇒ Ölçme **değişen** hatalardır.
“ Terazî, **her bir kilogramda 120 g** **eksik** tartıyorsa
→ Üzerine konulan nesnenin **ağırlığı arttıkça hata miktarı da artacaktır.**”
“ Sınavda yazısı kötü olandan puan kırmada **yine karışan hata sistematiktir.**”

Not: Yukarıdaki iki hatada da (**sistematik - sabit**)

⇒ **Ölçme sonuçlarına karışan hata miktarı, yönü ve kaynağı** **bellidir**.
- Bu tür hataların ölçme sonuçlarına **karışmasını engellemek**
- Bu tür hataları **düzeltilmek** daha kolaydır.

c) Tesadüfi (Rastlantısal) Hata:

- **Şansla** (tesadüfen) ortaya çıkan
- Ne **yönde** ve ne **ölçüde** **karıştığı** genellikle bilinemeyen hatalardır.

* **Ölçme sonuçlarına tek yönlü** olarak **karışmaz!**
- Bazen **Pozitif** bazen de **Negatif** yönde etki eder.

* **Sabit ve tesadüfi olmayan hatalar** → **Tesadüfi değişken olma** özelliğindedir.
- Bu nedenle psikometride hata kuramı **tesadüfi hatalar** üstüne kurulmuştur.

Tesadüfi Hataların Dört Kaynağı Vardır:

- 1) Ölçme İşlemini Yapan Kişiden / Öğretmeninden Kaynaklanan Hata
- 2) Ölçme Aracından Kaynaklanan Hata
- 3) Bireyden / Öğrenciden Kaynaklanan Hata
- 4) Fiziksel Ortamdan Kaynaklanan Hata

TESADÜFİ HATALARIN DÖRT KAYNAĞI

1) Ölçme İşlemini Yapan Kişiden / Öğretmeninden Kaynaklanan Hata:

* Öğretmenin **test etme sürecine** ilişkin **davranışları** iki aşamadır:

1. Ölçme işlemi sürecinde:

- Öğretmen, test katılımcısının dikkatini **dağıtabilir** - kaygısını **artırabilir**

2. Ölçme işlemi sonrasında:

- Puanlamadaki dikkat ve titizliğin **değişmesi, yorgunluk, maddi hata** vb....

2) Ölçme Aracından Kaynaklanan Hata:

* Ölçme araçları hazırlanırken **maddelerin** iyi ifade **edilmemesi**
- Test katılımcılarının yanlış anlamalarına ve **hataya** neden olur.

* Hangi özellik ölçülürse ölçülsün

- Araçta kullanılan **dil, dil bilgisi ve imla** kurallarına uygun **olmalı**
- **Maddelerde anlatım bozukluğu** bulunmaması
- **Muğlak** (kapalı) - **belirsiz** (müphem) ifadeler **bulunmamalıdır**.

Ölçme Aracının Yapısı:

* Ölçme aracı kaynaklı bir başka hata kaynağı → **Aracın yapısı** ile ilgilidir.

* Ölçtüğü özellik ve kapsam bakımından

- **Homojen** (benzeşik) maddelerden oluşan bir **araç** (test)

⇒ **Heterojen** (ayrışık) maddelerden oluşan bir **testten daha güvenilirdir**.

* **Ölçülecek özellikleri** yeterince temsil etmeyen **maddelerle** ölçüm yapılsa
- Bu araç araçtan **elde edilen puanlara** da **hata karışır**.

Ölçme Aracının Uzunluğu:

* **Ölçme aracının yapısıyla** ilgili diğer bir konu ise **aracın uzunluğudur**.

* **Madde sayısı ile güvenilirlik** arasında **doğru orantılı bir ilişki** vardır.

⇒ Bu **sonsuz bir doğru orantı değildir!**

⇒ Ölçme aracındaki **madde sayısı arttıkça**

- **Yorgunluk, dikkat azalması** vb. sebebiyle **hata miktarı** artar.

* **Ölçme araçlarında Bilişsel** özellikleri ölçen testlerde

- Eğer **seçenek verilmişse şans başarısı** **karışma olasılığı** artar.

(**Şans başarısı** : Salt tahminle doğru yanıt bulma olasılığı)

- Bu, ölçülen özelliğin **gerçekte olduğundan yüksek** görünmesine yol açar.

3) Bireyden / Öğrenciden Kaynaklanan Hata:

* Bireylerin ölçme işlemi sürecinde içinde bulundukları

⇒ **Fiziksel, fizyolojik ve psikolojik** durumlar **test puanını** etkiler.

(**Uykusuzluk, açlık, hastalık, ağrı, motivasyon eksikliği, kaygı** vb.)

⇒ Bu faktörler nedeniyle bireyler **gerçek puanlarından uzaklaşabilir...**

* Grubun **homojen** ya da **heterojen** olması

- **Hata kaynağını artırır** ya da **azaltır**.

⇒ Daha **heterojen** gruplarda güvenilirlik katsayısı → **Artar**

⇒ Daha **homojen** gruplarda güvenilirlik katsayısı → **Azalar**.

4) Fiziksel Ortamdan Kaynaklanan Hata:

* Ölçmenin yapıldığı **fiziksel ortama** da ölçme sonuçlarına **hata karıştırabilir**.

“ **Sıcaklık, ışık, ses, koku, görsel uyarıların fazlalığı, ergonomi** vb.... ”

* **Testler** bireylere **fiziksel ortam** açısından

- **Eşit ve standart koşullar** altında **uygulanmalıdır**.

GEÇERLİLİK - GÜVENİRLİLİK VE HATA İLİŞKİSİ

Klasik test kuramına göre:

Geçerlilik → Tüm hata kaynaklarından etkilenir

Güvenilirlik → Yalnızca **Tesadüfi Hatalardan** etkilenir.



* **Güvenilirlik** → Geçerlilik için bir **ön şarttır**.

⇒ Ancak yeterli şart **değildir**.

- Bir testin **güvenilir** olması onun **geçerli olacağı anlamına gelmez!**

- Bir test **geçerli** ise → Büyük olasılıkla **güvenilirdir**.

* Bir ölçme aracının **geçerliliği** ve **güvenilirliği** diye **bir şey yoktur!**

- **Geçerli ve güvenilir olan** → Ölçme araçlarından elde edilen **puanlardır**.

Ölçme araçlarının psikometrik nitelikleri şu durumlara göre **değişebilir**:

- a) Ölçme amacının **değişmesi**
- b) Uygulama grubunun **değişmesi**
- c) Dilin **eskimesi / değişmesi**
- d) Maddelerde - alt ölçeklerde yapılan **değişiklikler**
- e) Farklı **kültürler**,
- f) Kuramsal bilgi birikiminde **değişiklikler**
(**Psikometri: Bireysel** özellik ve **nitelikleri değerlendiren** bir bilim dalıdır.)

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

* Bir testten elde edilen puanlar :

- Yüksek **düzeyde** geçerli
 - Orta **düzeyde** geçerli
 - Düşük **düzeyde** geçerli
 - Yüksek **düzeyde** güvenilir
 - Orta **düzeyde** güvenilir
 - Düşük **düzeyde** güvenilir
- (biçiminde nitelendirilir.)

* Bir testin şöhretli olması onun geçerli olduğu anlamına **gelmez!**

- **Modern test kuramı**na göre geliştirilen **ölçekler** eleştiriye açıktır.
- Böylece testlerin niteliği ve test bilgisi **gelişir**.

KORELASYON

* Geçerlilik ve güvenilirliği belirlemeye yönelik **yöntemler**

- Genellikle **korelasyon temelli analizlere** dayanır.

Korelasyon(co-relation) ("r" ile sembolize edilir.):

- * En az iki **değişken** arasında karşılıklı bir ilişkinin varlığını inceler.
- Var olan ilişkinin → **Yönü** ve **miktar** bilgisini veren **istatistik** bir tekniktir.
- "Öğrencilerin ders tutumları ile ders başarıları arasında bir ilişki var mıdır?"
- "Saç uzunluğu ile zekâ arasında bir ilişki var mıdır?"

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

* Bu iki değer arasında **matematiksel olarak** sonsuz birim vardır.

- Kullanışlılık açısından
- * 100 birim **negatif** korelasyonda,
- * 100 birim **pozitif** korelasyonda,
- * Sıfır ile birlikte toplam 201 birim.

⇒ Değişkenler arasındaki **ilişkinin yönü** ve **miktarı** değerlendirilir.

* **Pozitif korelasyon** → İki değişken arasında **doğru orantılı** ilişkidir.

"Ders çalışma süresi ile sınav notu" → **Pozitif bir korelasyon**

* **Negatif korelasyon** → İki değişken arasında **ters orantılı** ilişkisidir.

"Ders süresi ile dikkat" → **Negatif bir korelasyon**

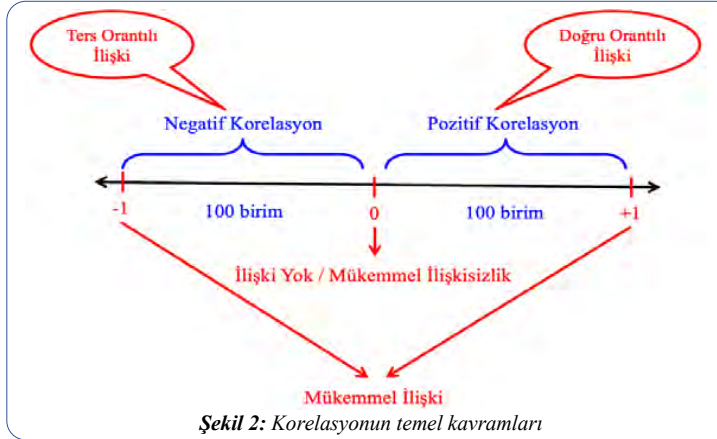
* **±1** Korelasyon → **Mükemmel korelasyon**

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.



Şekil 2: Korelasyonun temel kavramları

Korelasyon Katsayısı:

*İki tür belirleme yapılır:

Birincisi → **Yön** **İkincisi** → **Miktar**dır.

Yön → **Negatif** ya da **pozitif**

Miktar → **Düşük**, **orta** ya da **yüksek**



Şekil 3: Korelasyonun miktarı

Not 1:

* **Korelasyon mutlak değer** olarak değerlendirilmelidir.

* Bir korelasyon katsayısının **negatif** ya da **pozitif** olması

⇒ **Büyük-küçük** **belirtmez!** (Yön bildirir.)

Not 2:

* **Korelasyon katsayısı** ile **neden-sonuç** ilişkisi **kurulamaz!**

* **Değişkenler** arasında **doğru** ya da **ters orantılı** bir **ilişki** olması

- **Değişkenler** arasında bir **neden-sonuç** ilişkisinin varlığı **anlamına gelmez!**

GEÇERLİLİK SORGULAMA YÖNTEMLERİ



Şekil 4: Geçerlilik yöntemleri sınıflaması

1. KAPSAM GEÇERLİLİĞİ:

* Özellikle **başarı testlerinde** aranan bir **geçerlilik** sorgulamasıdır.

* **K.G.** bir testin ölçülmek istenen davranışları **ne derece kapsadığıyla** ilgilidir.

* **Test**, kapsamı açısından davranışları **yeterli** ve **dengeli** biçimde **temsil etmeli**.

Bir Testin Kapsam Geçerliliğinin Yüksek Olması (İki kritere bağlı)

1. Testteki **sorular** ölçülecek **içeriği yeterli** ve **dengeli** biçimde ölçmeli.

2. Her bir **soru** ölçmek istediği özelliği **doğrudan** ölçmeli.

(Kazanımla doğrudan ilgili olması gerekir.)

Kapsam Geçerliliği Sorgulama Yöntemleri:

a) Mantıksal / rasyonel yöntemler:

* Bu yöntemler belirtke tablosu hazırlanması ve uzman görüşüne başvurulmasıdır.

a1) Belirtke tablosunun hazırlanması:

* **B.T.** ölçülen kapsam dâhilinde davranışların belirlenmesi **yöntemidir**.

"Öğretmen satıra davranışları, sütuna hedefleri yazar; belirlediği madde sayısı doğrultusunda davranış ve hedefleri yeterli ve dengeli bir biçimde temsil eden alanları seçer."

a2) Uzman görüşüne başvurulması:

* Uzman, ölçme ve değerlendirme tekniklerini de bilmeli.

* Uzman **belirtke tablosu** ve **maddeler** sunulur:

- Soruların konu kapsamını
- Yeterli ve dengeli bir ölçme durumunu
- Soruların kazanımlarla doğrudan ilgili olma durumunu **değerlendirmesi** istenir.

b) İstatistiksel yöntemler:

Uzmanlardan alınan dönütler betimsel / muhakemeye dayalı bir yolla çözümlenebilir ya da uzmanların "uygundur / uygun değildir" ya da "uygundur / düzeltme gerekir / soru kullanılmamalıdır" vb. biçimde değerlendirme yapması istenebilir. Eğer uzmandan ikinci yolla dönüt istenmişse uzmanlar arasında uyum olup olmadığı, çeşitli istatistiksel yöntemlerle test edilir. Alanda uzmanlar arasında uyumu test eden pek çok indeks bulunmaktadır.

2. ÖLÇÜT DAYANAKLI GEÇERLİLİK:

* Ölçme aracından elde edilen puanların ölçüt bir puanla karşılaştırılarak

- Geliştirilen **ölçme aracının geçerliliğine** ilişkin **nitelendirme yapılır**.

3. YORDAMA GEÇERLİLİĞİ:

* **Yordama**, **tahmin** demektir. (Ancak **her tahmin yordama değildir!**)

* Bir **tahminin yordama** olabilmesi için

- Elde geçerli ve güvenilir bir **veri** olması
- Bu **verinin** **sınanabilir**, **sayısal nitelikte**, **analizlere** ediliyor olması gerekiyor.

* **Yordama**, eldeki bu nitelikteki veriden yola çıkarak,

→ Geleceğe, henüz **gerçekleşmemiş** bir **olguya** ilişkin **yapılan tahmindir**.

* **Ölçme araçlarının** çoğunda gelecekteki davranışların **kestirilmesi** vardır.

* **Yordama geçerliliği**, ölçme aracının bu **kestirim işini**

→ Ne ölçüde **doğru** ve **isabetli** yaptığı ile ilgili bir **sorgulamadır**.

* **Seçme** ya da **yönlendirme amaçlı** testlerde

→ **Yordama geçerliliğinin yüksek olması** istenir.

Seçme amaçlı testler → **YKS, KPSS, LGS** vb. araçlar; diğer taraftan

Yönlendirme amaçlı testler → **Yetenek testleri, ilgi envanterleri...**

* **Yordama geçerliliğinde** **ölçme** aracından elde edilen puanlar

- Ölçme aracının **tahmin ettiği puanla (ölçüt puan)** karşılaştırılır

- **Tahminin** ne ölçüde **doğru** olduğu belirlenmeye çalışılır.

* **Ölçüt puan** testin **tahmin etmeye çalıştığı** → **Özelliktir**.

Yordama

geçerliliğinde

ölçüt puan

gelecekte belli

olacağından

→ **Beklenir**

Y.G. 'de

geçerlilik

katsayısının

0 ile 1

arasında

değişmesi beklenir.

	YKS Puanı	Akademik Ortalama
Tarik	438.636	3.54
Esra	348.545	2.45
Begüm	456.344	3.39
Orzan	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
Umut	343.019	3.26

Ölçüt gelecekte belli olduğundan beklenmektedir.

Korelasyon (r) 1'e yakın

Geçerlilik Katsayısı

Şekil 5: Bir yordama geçerliliği çalışma örneği

* Yordama geçerliliğinde **en zor** ve **önemli nokta**
=> Ölçütün **doğru** bir biçimde **belirlenmesidir**.

Ölçütün Belirlenmesinde Nelere Dikkat etmelidir?

Ölçüt

- 1) Ölçme aracının **yordamaya** çalıştığı **değişkenle doğrudan ilgili** olmalı
- Ölçme aracı **yordamaya** çalıştığı özelliğin **doğrudan bir temsili** olmalı
- 2) **Kararlı** olmalı, günden güne değişme**melidir**.
- 3) Bireylerin **özellikliğini** gerçekten **yansıtan nesnel** ve **güvenilir** olmalı
- 4) **Kolay** elde edilmeli ve **ekonomik** olmalıdır.

4. ZAMANDAŞ GEÇERLİLİK:

- * Diğer adları: **Hâlihazır geçerlilik**
Benzer ölçekler geçerliliği
Uygunluk geçerliliği
- * Zamandaş geçerlilik sorgulamalarında
=> **Ölçüt puan eş zamanlı** olarak elde edilebilir.
=> Ölçme aracı ile ilişkili **nitelikleri ölçen** ve **geçerliliği yüksek** bir aracın **puanı**
- **Ölçüt puan** olarak **ele alınabilir**.
=> Sonra **ölçme aracı** ve **ölçüt araç** aynı zamanda **uygulanıp**
- İki araçtan elde edilen **puanlar arasındaki ilişki** incelenir.

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

Şekil 6: Bir zamandaş geçerlilik çalışma örneği

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



Şekil 7: Geçerlilik katsayısı

- * Z.G. te de **en zor** ve **önemli nokta**
=> Ölçütün **doğru** bir biçimde **belirlenmesidir**.

Ölçütün Belirlenmesinde Nelere Dikkat etmelidir?

Ölçüt

- 1) Ölçme aracının **ölçmeye** yöneldiği **özellikle doğrudan ilgili** olmalı
- Bu ilişki **doğru orantılı** ya da **ters orantılı** olabilir.
- 2) **Geçerliliği yüksek** olmalıdır.
Geçerliliği yüksek olmayan bir ölçüt puanla bakılacak **korelasyonun**
→ **Düşük** olması kaçınılmazdır.

5. YAPI GEÇERLİLİĞİ:

- * Beşerî bilimlerde
* **Atılganlık, güvensizlik** ...öz.nin **ölçülmesi** amacıyla **ölçme aracı** geliştirmek için
- **Önce**, bu özellikleri **belirlemek** (yapıyı tanımlamak)
- **Sonra**, tanımlanmış yapıdan **sinanabilir denenceler** çıkarmak
- **Denenceleri sinamak** için **deneyssel** ve **istatistiksel** çalışmalar yapmak gerekir.

- * **Yapı**, birbirleriyle **ilgili** olduğu düşünülür
- Belli **öğelerin** ya da **öğeler arasındaki ilişkilerin** oluşturduğu bir **örüntüdür**.
- * Bir testin yapısını **geçerleme süreci**,
- Testin maddelerine verilen **yanıtlar** arasındaki ilişkilerin analizine **dayanır**.
- * **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**.

"**Geleneksel aile ile çocuk yetiştirme biçimi** arasındaki **ilişkiyi** incelemek amacıyla
- Bir **ölçek** geliştirip bu ölçeğin **yapı geçerliliğini** ortaya koymak istediğinde
1) **Geleneksel aile** yapısı ve **çocuk yetiştirme** kavramlarının ne anlamlara geldiğini
2) Ölçme araçındaki maddelerin bu anlamlara uygunluğunu araştırarak
=> **Yapı geçerliliği** konusunda **karar verebilir**."

GÜVENİLİRLİK

- * **Güvenilir** bir **ölçme aracı aynı özelliklerle** ilgili olarak
- **Arka arkaya** yapılan **ölçmelerde** yaklaşık olarak **aynı sayısal sonucu** verir.
- * Bir test, aynı gruba **iki ya da üç kez uygulandığında**
- **Her bir kişi bütün uygulamalarda** yaklaşık olarak **aynı puanı** almaktadır.
- * Bir testin **ölçmek** istediği **özellikliği ölçebilmesi** için
- O testin **özellikliği kararlı olarak ölçmesi** gerekir.

Güvenirlilik çalışmalarının odak noktası:

- * "Eğer kişi **iki defa** teste tabi tutulursa **iki testten aldığı puanlar** birbirine **benzer** midir ya da birbirine ne kadar **yakındır**?"

Örneğin

- * **Murat**, dakikada **54 sözcük** yazsın.
- Bu puan Murat'ın becerisini **ne ölçüde yansıtmaktadır**?
- * Dakikada **50** sözcük yazılması yeterli kabul edilse
- Murat bu düzeyin **gerçekten üzerinde midir**?
- * Murat geçen hafta dakikada **45** sözcük yazmıştı.
- Murat'ın bugünkü puanı olan **54**
=> Onun kendisini **geliştirdiğinin** göstergesi midir?
=> Yoksa bir **değişim dalgalanması** mıdır?

İki Ölçüm Arasındaki Uyumsuzluğun Birçok Nedeni Olabilir:

- * Bir hareketten diğerine, "**dikkat** ve **çaba**" **değişebilir**.
* Uzun periyotlarda **puan değişmesi**
- Fiziksel **büyüme**, **öğrenme**, **sağlık** ve **kişilikteki değişim** kaynaklıdır.
* İkinci ölçümde **daha açık soruların** kullanılması

Gerçek puan:

- * Psikometride **hata kavramı istenmeyen değışkene** işaret eder.
* **Ölçme, hatalar giderilene kadar sürdürülmeli**
- **Gerçek puan** elde edilmelidir.
- * Ancak **davranış örneği sınırlı** olduğu için
- **Gözlenen puan gerçek puandan farklılık gösterir**.
- Buradaki farklılık **ölçme hatasıdır**.

- * Örneğin yarışlara hazırlanan bir koşucu
- Bir mesafeyi farklı zamanlarda 23.7, 24.0, ... 25.1, 25.2 saniyelerde koşmuş olsun.
- Bu **ölçümlerin ortalaması 24.7** ise → **Gerçek puanı 24.7**'dir.

- * Ölçme işleminde **hata miktarı arttıkça**
- **Gözlenen puanın gerçek puana olan farkı artar**.

Ölçmenin **standart hatası** $S_h = s\sqrt{1-r_x}$ formülü ile **bulunur**.

S_h = Ölçmenin standart hatası s = Standart sapma r_x = Güvenirlilik katsayısı

- * Bir ölçme işleminde **standart hatanın düşük olması**
- **Güvenirlilik katsayısının yüksek ve**
- **Standart sapmanın düşük olmasına** bağlıdır.

- * Ölçmenin **standart hatası yükseldikçe**
- Ölçme işlemi bireylere **rastgele puan** vermiş gibidir!
- Bu nedenle ölçme işlemlerinde **güvenirliliğin kestirilmesi** çok önemlidir.

- * **Güvenirlilik katsayısı 0 ile 1** arasında **değişir**.
- **Güvenirlilik katsayısı için 0.70** ve **üstü** ölçütü kabul edilir.
- **Genel yetenek** gibi bilişsel özellik ölçen testlerde
=> **Güvenirlilik katsayısının 0.90** ve **üzerinde olması arzu edilir**.

Şekil 8: Güvenirlilik katsayısı



GÜVENİLİRLİK YÖNTEMLERİ

a) Birden çok uygulamaya dayalı yöntemler:

1. **Test-tekrar test** yöntemi
2. **Eşdeğer (paralel) testler** yöntemi

b) Tek uygulamaya dayalı yöntemler:

1. **Eşdeğer yarılar** yöntemi
2. **KR-20** yöntemi
3. **KR-21** yöntemi
4. **Cronbach alfa** yöntemi
5. **Hoyt'un varyans analizi** yöntemi
6. **McDonald omega** yöntemi



Şekil 9: Güvenirlilik kestirim yöntemleri

A) BİRDEN ÇOK UYGULAMAYA DAYALI YÖNTEMLER

1) TEST-TEKRAR TEST YÖNTEMİ:

- * Test güvenilirliğini test etmek için
 - * Bir test, aynı gruba, belli bir zaman aralığıyla iki kez uygulanır.
 - * İlk uygulama puanıyla ikinci uyg. puanı arasındaki korelasyon hesaplanır.
 - => Elde edilen korelasyon katsayısına Kararlılık (devamlılık / istikrarlılık) katsayısı adı verilir.

	Anksiyete Ölçeği Puanları	Anksiyete Ölçeği Puanları
	X ₁	X ₂
Ceren	54	50
Filiz	98	110
Aydın	128	120
Ece	89	80
Deniz	101	101
.....	87	85
.....	98	106
.....	67	59
.....	122	119
Dursun	79	82

Şekil 10: Test-tekrar test yöntemi çalışma örneği

- * Güvenilirliğin kararlılık, tutarlılık ve duyarlılık sorunu vardır.
- * Test-Tekrar Test Yöntemi ile elde edilen güvenilirlik kanıtı
 - Güvenilirliğin yalnızca kararlılık boyutuna karşılık gelir.
 - Güvenilirliğe ilişkin tümel bir sonuç ortaya koymaz!
- * Bu nedenle bir ölçme aracının güvenilirliğine ilişkin kanıt toplamak istersek
 - Kararlılığın yanı sıra tutarlılık ve duyarlılık ilişkini sorgulamaları yapmalıyız.

- * Test-Tekrar Test Yöntemi
 - Ölçülen özelliğin kararlı olduğu durumlarda uygulanması gerekir.
 - Testin ölçtüğü özellik sürekli değişkenlik gösteriyorsa kullanılmamalıdır!
 - İki uygulama arasında kolaylıkla değişmeyen özellikleri ölçen testler içindir.

("Genel zihin yetenekleri, kişilik testleri, ilgi envanterleri, tutum ölçekleri vb.")

Test-T.T. Yönteminde Test Güvenilirliğini Kestirmede Karşılaşılan Önemli Sorun:

- * Testin iki uygulaması arasındaki zaman aralığının ne kadar olması gerektiğidir.
- Bu soruya kesin bir yanıt vermek olanaksızdır!

Aradan geçen zaman,

- Yanıtlayıcıların önemli ölçüde değişmelerine izin vermeyecek bir uzunlukta ve
- İlk yanıtlarının hatırlanmasına izin vermeyecek bir uzunlukta olmalıdır.

İki Uygulama Arasındaki Zamanın Belirlenmesindeki Temel İlke Şudur:

- Birinci uygulamada maddeleri hatırlamayacak kadar uzun olmalı
- Ölçülen özellikte köklü değişimler oluşmayacak kadar kısa olmalıdır."

Aradaki Zamanın Belirlenmesinde Dikkat Edilmesi Gereken Başka Nokta:

- * Özelliğin değişim hızıdır.
- Bazı özellikler daha hızlı değişime açıktır bazıları ise daha zor değişir.
- Bireylerin genel yetenek düzeyleri → Kısa bir zamanda değişmez!
- Bireylerin tutumu → Daha kolay değişebilir.

- => Genel yetenek testi için aradaki zaman → Daha uzun tutulabilir.
- => tutum ölçeği için daha kısa tutulabilir.

- * Özelliklerin değişimi yaşla da ilişkilidir.
- Bazı özellikler belli yaşlarda daha durağanken
- Bazı yaşlarda daha değişkendir.
- "Yaşamın ilk evrelerinde çocukların gelişimi daha hızlıdır."

- * Test-tekrar test yöntemi, başarı testleri için pek tercih edilmemez!
- Kısa vadede başarı kolay değişebilen, kararsızlık gösterebilen bir özelliktir.

- * Test-tekrar test yöntemi
 - * Yetenek testleri, kişilik envanterleri gibi
 - Psikolojik ölçme puanlarının güvenilirlik kanıtlarını üretmede tercih edilir.

- * Test-tekrar test yöntemi
 - Benzer ölçekler geçerliliği, hâlihazır geçerlilik, uygunluk geçerliliği olarak da geçer.

- * Güvenilirliğin bir boyutu da testin tutarlılığıdır.
- "Tek uygulamaya dayalı güvenilirlik sorgulama yöntemleri" ile
- "Test kendi içinde tutarlı bir bütün oluşturur mu?" sorusuna yanıt aranır.
- * Bu nedenle bu yöntemlerin tümünden elde edilen katsayı
 - => İç tutarlılık katsayısı olarak adlandırılır.

B) TEK UYGULAMAYA DAYALI YÖNTEMLER

1) TEST YARILAMA (Eşdeğer Yarılar / İki Yarı Güvenilirliği) YÖNTEMİ:

- * Güvenilirliği tahmin etmede uygulanmış bir test
 - İki eşdeğer yarıya bölünür ve
 - İki yarıdan alınan puanlar arasındaki tutarlılık incelenir.

En temel sorunlardan biri test iki eşdeğer yarıya nasıl bölünür?

* En sık başvurulan yöntemler:

- a) İlk yarı ve son yarı b) Tek ve çift c) Rastlantısal yöntem

=> İlk yarı ve son yarı yöntemi her test için uygun değildir!

- * Uygun olmadığı durumlar:
 - Testteki maddeler basitten zora doğru sıralanmış ise
 - Maddeler konu içeriklerine - faktörlere göre kümelenecek yerleştirilmiş ise
 - Madde sayısı çok fazla ise uygun değildir!

	Türkçe Testi	X ₁ YARI	X ₂ YARI
Deniz	48	20	28
Şebnem	98	50	48
Ufuk	66	32	34
İşıl	88	44	44
Sinem	32	20	12
.....	70	30	40
.....	62	28	34
.....	40	20	20
.....	60	28	32
Sinan	66	36	30

Şekil 11: Test yarılama yöntemi çalışma örneği

- * Test Yarılama ile güvenilirliği tahmin etmede uygulanmış bir test
 - İki eşdeğer yarıya bölünür ve
 - İki yarıdan alınan puanlar arasındaki tutarlılık incelenir. (Şekil 11)
 - İki eşdeğer yarıdan alınan toplam puanlar arasında korelasyon hesaplanır.

- * Ancak elde edilen katsayı
 - Korelasyonun yarı puanlar üzerinden elde edilmiş oldu.
 - Bu nedenle testin tamamına ilişkin bir tutarlılık bilgisi vermez!
 - Elde edilen katsayı eşdeğer yarılarından birinin güvenilirliğidir.

- * Testin tümüne ilişkin bir güvenilirlik katsayısı
 - => Spearman-Brown formülü aracılığıyla hesaplanır.
- * Elde edilen katsayı testin tamamına ilişkin
 - => İç tutarlılık bağlamındaki güvenilirlik katsayısı olarak kabul edilir.

2) KUDER-RICHARDSON 20 VE 21 YÖNTEMLERİ:

- * Kuder-Richardson 20 ve 21 (KR-20 ve KR-21)
 - Testin kendi içinde tutarlı bir bütün oluşturup oluşturmadığı ile ilgili bilgi verir.
 - => KR-20 ve KR-21'de testin iç tutarlılığı değerlendirilir.
 - * Bu yöntemlerden elde edilen katsayıya iç tutarlılık katsayısı denir.

- * KR-20 ve KR-21 ile güvenilirlik kestirimi
 - Puanlamanın kategorik olduğu ölçme araçları için uygundur.
- * Puanlamanın kategorik olması
 - Doğru yanıtlara 1 puan, yanlış ve boş ise 0 puan vererek puanlanır
 - Ya da maddenin yanıtının iki seçenekli verildiği durumlardır.

- * Eğer testteki maddeler
 - Farklı ağırlıklarla puanlanmışsa ya da
 - Test puanları şans başarısı için düzeltilmişse => Bu formüller kullanılmaz!

3) CRONBACH ALFA YÖNTEMİ:

- * Eğer ölçme aracının puanlaması çok kategorili ise
 - KR-20 ile aynı mantık üzerine kurulu Cronbach alfa hesaplanır.
- * Elde edilen katsayının ismi yine iç tutarlılık katsayısıdır.
- Örnek: "Derecelendirme ölçekleri puanlamanın çok kategorili olduğu araçlar"

GÜVENİLİRLİĞİ VE GEÇERLİLİĞİ ARTIRMA YOLLARI

- * Güvenilirlik temelde tesadüfi hatalardan arınık olma durumudur.
- Tesadüfi hata kaynaklarını azaltma → Güvenilirliği artıracaktır.

Güvenilirliği Artıran Temel Kavramlar ve Eylemler:

- 1. Bir testteki madde sayısı arttıkça → Birimler küçülür.
 - Duyarlılık artar
 - => Bu nedenle hata miktarı azalır → Güvenilirlik artar.

- 2. Bir ölçme işleminde
 - Puanlayıcı sayısı arttıkça → Güvenilirlik artar.
 - => Birden çok kişinin yaptığı puanlamanın ortalamasını almak
 - Bir kişinin verdiği puandan → Daha güvenilir.

3. Puanlama nesnelliği → Güvenilirliği büyük ölçüde etkiler.

* Bir testin **farklı kişiler** tarafından **puanlanması** ya da

- Aynı kişinin **farklı zamanlarda** verdiği **puanlar** arasındaki **tutarlılığa**
⇒ **Puanlama güvenilirliği** adı verilir.

* Bir testten elde edilen **puan**

- Puanlayıcıya ya da zamana göre **değişmiyorsa** → Testin **güvenilirliği artar**.

* Ölçme aracının puanlamasının **nesnel** ve **öznel** olması çok önemlidir.

Nesnel Araçlar: “Çoktan seçmeli, derecelendirme ölçeği, kontrol listeleri...”

Öznel Araçlar: “Açık uçlu sorulardan oluşan araçlar, mülakatlar vb...”

* **Öznel araçların** puanlama biçiminin **nesnel hâle getirilmesi**

⇒ **Güvenilirlik** açısından önemlidir.

4. KR-20 ve KR-21 yöntemlerinde bahsedildiği gibi

⇒ **Testteki maddeler** açısından

- **Benzeşiklik (homojenlik) arttıkça** → **Güvenilirlik artar**.

- **Ayrışıklık (heterojenlik) arttıkça** → **Güvenilirlik düşer**.

(Benzer maddeler arasından seçim zorlaşır. Tesadüfi cevap azalır)

5. Puanların güvenilirlik kestirimi için

- Veri elde edilecek **grubun büyüklüğü artarsa** → Grup Heterojenleşir
→ **Güvenilirlik artar**.

(Grup üyelerinin fazlalığı, elde edilecek veriyi çeşitlendirir)

6. Bir üst maddeyle de paralel bir biçimde

* **Maksimum performans ölçen** testlerde

- **Ortalama güçlüğe yaklaştıkça** ($\bar{P} = .50$) → Grup Heterojenleşir
→ **Güvenilirlik artar**.

7. Maddelerin dil bilgisi kurallarına uygun, açık ve anlaşılır yazılması

→ **Güvenilirliği artırır**.

→ Bu koşul sağlanmadığı takdirde diğer koşulların fazla önemi olmayabilir!

→ İyi madde yazarı olmada **ana dilini iyi kullanma** becerisi önemlidir.

⇒ **Temel ilke**

- Bir **maddenin** tüm öğrenciler tarafından **aynı biçimde anlaşılmasıdır**.

8. Öğrencilere yönerge vermek

- Test almayı → **Güdüler** **Hazırbulunuşluluğu** → **Arttırır**.

* **Bireyden** kaynaklanabilecek **hataları asgariye indirir**.

9. Fiziksel ortamın uygun hâle getirilmesi → **Güvenilirliği artırır**.

(Ses, sıcaklık, ıskık, koku, dikkat dağıtıcı uyarılar, ergonomi vb.)

10. Eğer süreli test uygulaması ise **süre** yeterli verilmelidir.

- Bu bir **hız testi değilse** → Herkesin maddelerin tümüne ulaşması beklenir.

11. Uygulamada bireylerin

- **Dikkatini dağıtacak** - **kaygısını artıracak** davranışlardan kaçınılmalıdır.

12. - Maddeler teste düzgün yerleştirilmeli

- Okumayı **güçleştirecek** bir **unsur** bulunmamalı

- Test katılımcısının **yaşına, gelişim düzeyine uygun bir punto**

- Baskı hataları bulunmamalı

13. Tesadüfi - sabit ve sistematik hatalardan arındırılmalıdır.

TEST GELİŞTİRME VE MADDE / SORU TÜRLERİ

Test Nedir?

* Eğitimde **birey** özelliklerini belirlemeye yönelik **ölçme aracıdır**.

* **Test terimi** günlük dilde yalnızca çoktan seçmeli anlaşılrsa da

⇒ Test çok daha geniş bir **anlamda** kullanılmaktadır.

TEST TÜRLERİ

Testi Alan Kişi Sayısına Göre	Testin Uygulanış Süresine Göre	Testin Ölçtüğü Niteliğe Göre	Değerlendirme Yaklaşımına Göre	Hazırlanış Biçimine Göre	Veri Toplama Tekniğine Göre
1. Bireysel	1. Süreli	1. Hız testleri	1. Objektif	1. Standart	1. Performans
2. Grup	2. Süresiz	2. Güç Testleri	2. Subjektif	2. Öğretmen yapımı	2. Kağıt-Kalem

Tablo 4: Test türleri

TEST GELİŞTİRME

* **Önceden belirlenmiş özelliklerin ölçülmesi** ile ilgili bir **test geliştirilmesi**

- **Sistematik işlem basamaklarına** uyulmasını gerektirir.

Test Geliştirmenin İşlem Basamakları:

- Amacın belirlenmesi,
- Kapsamın belirlenmesi ve belirtke tablosunun oluşturulması,
- Denemelik maddelerin / soruların yazılması,
- Maddelerin / soruların gözden geçirilmesi (redaksiyon),
- Denemelik test formunun hazırlanması,
- Testin uygulanması,
- Test ve madde istatistiklerinin hesaplanması,
- Seçilen maddelerden oluşan nihai formun oluşturulması.

* Test geliştirme adımları olarak bunlar verilebilir

* Sınıf içi ölçme ve değerlendirme etkinliklerinde

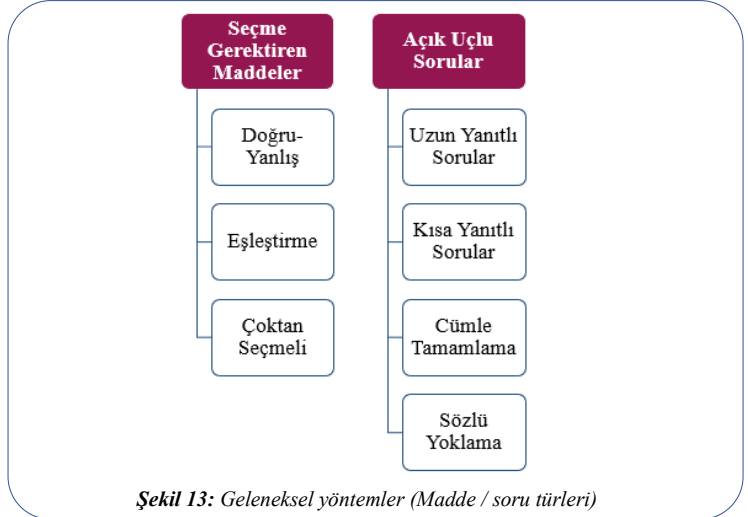
-Bu basamaklara **uygun test geliştirmek pratikte çok olanaklı değildir**.

Öğretmenler açısından daha olanaklı bir test planı:

1. Sınavın amacının belirlenmesi
2. Sınavda yoklanacak davranışların belirlenmesi
3. Sınavın kapsayacağı konuların listelenmesi
4. Okulun takvimine uygun sınav gününün belirlenmesi
5. Bir ders saatine uygulanabilecek uzunlukta bir sınav süresinin belirlenmesi,
6. Belirtke tablosunun hazırlanması
7. Soru/madde türlerinin belirlenmesi
8. Soru/madde sayısının belirlenmesi
9. Sınav süresinin belirlenmesi
10. Soruların “ortalama güçlüğünün” ve “güçlük dağılımının” belirlenmesi
11. Soruların yazımında, izlenecek yolun belirlenmesi
12. Cevap anahtarının ve puanlama yönteminin belirlenmesi
13. Ölçme aracını yazma ve çoğaltma yönteminin belirlenmesi
14. Sınavın uygulanma kurallarının belirlenmesi (yönerge yazılması)
15. Sınavın uygulanması
16. Test ve madde istatistiklerinin hesaplanması

BAŞARININ ÖLÇÜLMESİNDE YÖNTEMLER

Şekil 12: Ölçme araç ve yöntemleri



Şekil 13: Geleneksel yöntemler (Madde / soru türleri)

GELENEKSEL ÖLÇME YÖNTEMLERİ**DOĞRU-YANLIŞ MADDELERİ:**

* Cevaplayıcıdan verilen ifadelerin

- **Doğru** mu, **yanlış** mı olduğunu belirlemesinin **istendiği** madde türüdür.

* **Doğru- Yanlış** maddelerinde verilen bir tür “**önerme**”dir.

“Cumhuriyet 1920 yılında kurulmuştur” (**D**) (**Y**)”

* **Çoktan seçmeli madde türünden sonra** en çok kullanılan **ikinci madde tipidir**.

- Bilimsel gerçekleri, tarihî olayları, kesin yargıları **tamamı** ve **hatırlama**
- Olguları ve gerçekleri kişisel yargı ve görüşlerden **ayırt etme ölçülmek istenir**.

Avantajları:

1. Hazırlaması, uygulaması ve puanlaması kolay(kullanışlıdır.)

2. Soru sayısı artırılabilir. **3.** Sistematik hata karışma olasılığı pek yoktur.

Dezavantajları:

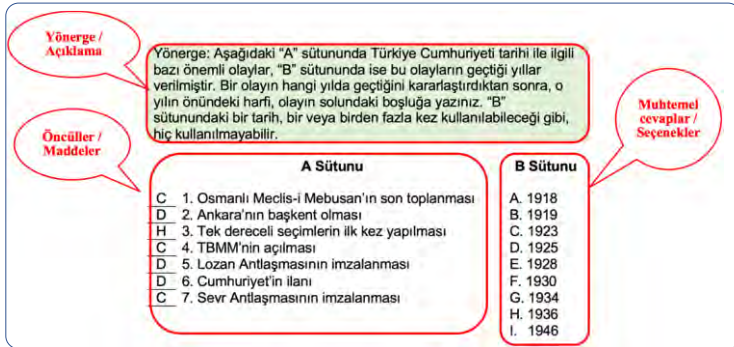
1. Şans başarısı olasılığı yüksektir. **2.** İyi yazılmazsa genellikle üst düzey bilişsel özellikleri ölçmez. **3.** Öğrenme eksiklerini belirleyemez. **4.** Yanlış üzerinden öğretim yapılmaz. **5.** Her derste/konuda kesin yanlış denebilecek durumlar belirlemek zordur.

ÇOKTAN SEÇMELİ MADDELER

- 1) Her madde **tek** ve **belirli** bir **fikri** belirtmelidir.
(Aynı maddede biri **doğru**, öteki **yanlış** **iki fikir olmamalı!**)
- 2) **Doğru yanlış maddesi, kesinlikle doğru** ya da **kesinlikle yanlış** olmalıdır.
(Başka bir **açıklamaya** **gerek kalmadan** belirlenebilmelidir.)
- 3) Maddenin **yanlışlığı** **önemsiiz ayrınıtıda / aldatıcı bir noktada olmamalıdır.**
(Bir yargı temelden **yanlış** olmalıdır.)
- 4) Mümkün olduğunca **olumsuz ifade** kullanılmamalıdır.
(**İki olumsuz ifadedden kesinlikle kaçınılmalıdır.**)
- 5) **Kanı ifadeleri** bir **kaynağa** dayandırılmalıdır.
- 6) **Maddenin ifadesi kısa, açık** ve **yalın** olmalıdır.
- 7) Maddenin okuduğunu **üst düzey anlama süreçlerine yönelik** olabilmesi için metindeki **cümleler aynen** yazılmamalıdır.
- 8) Özellikle **vurgulamak** istenen bir ifadenin **altı çizilebilir.**
- 9) Maddenin özellikle **yanlış olduğu durumlarda**, şans faktörünü engellemek için **ifadeyi düzeltmesi** ya da **yanlışlığın ne olduğunu belirtmesi istenebilir.**
- 10) **Doğru-yanlış maddeleri** ile **üst düzey öğrenmeler** de **yoklanabilir.**
- 11) **Maddeler aynı kapsamda ise ortak soru kökü altında birleştirilebilir.**
- 12) Doğru ve yanlış **maddelerin testteki sıralanışı rastgele olmalıdır.**
- 13) **Maddelerin ifadesi** yaklaşık olarak **aynı uzunlukta** olmalıdır.
- 14) **Doğru ve yanlış maddelerin sayısı**, yaklaşık olarak **birbirine eşit** olmalıdır.
- 15) **Madde işaretleme yöntemi, açık** ve **anlaşılır** bir **biçimde açıklanmalıdır.**
- 16) Mümkünse **kısa, basit cümle yapısı** olmalı / "**ve**" gibi **bağlaçlara** dikkat et.
- 17) **Olumsuz anlamlı bir sözcük** ya da **ifade** varsa **altı çizilmelidir.**
- 18) **İpucu** vermekten **kaçınmalıdır:**
 - * "Asla, daima, hiçbir, tümü" vb. ifadelerin → **Yanlış** olma olasılığı yüksektir.
 - * "Genellikle, bazen, sıklıkla" vb. ifadelerin → **Doğru** olma olasılığı yüksektir.
- 19) **Üst düzey beceriler için** **tablo, grafik, harita** gibi bir **öncül** kullanılmalıdır.

EŞLEŞTİRME MADDELERİ:

- * Eşleştirme maddeleri, **iki grup hâlinde verilen** ve **birbirleriyle ilgili olan** - **Bilgi öğelerinin** belli bir açıklamaya göre **eşleştirilmesini** gerektirir.



Avantajları (Eşleştirme Maddeleri):

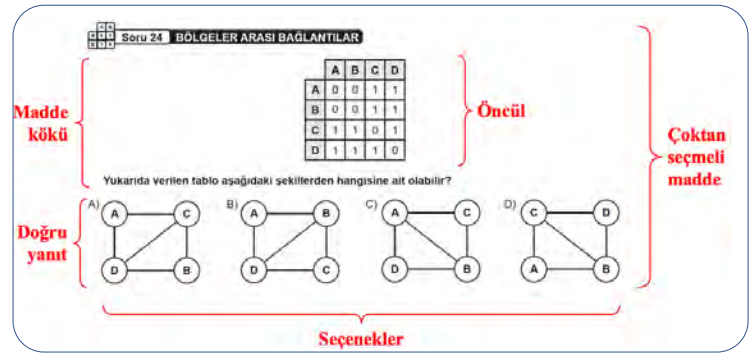
- a. Hazırlaması, uygulaması ve puanlaması **kolaydır**. (Kullanışlılık)
b. Soru sayısı artırılabilir.
c. **Sistematiik hata** karışma olasılığı **pek yoktur!**

Dezavantajları (Eşleştirme Maddeleri):

- a. Şans başarısı olasılığı **yüksektir!**
b. İyi yazılmazsa genellikle üst düzey bilişsel özellikleri **ölçemez!**

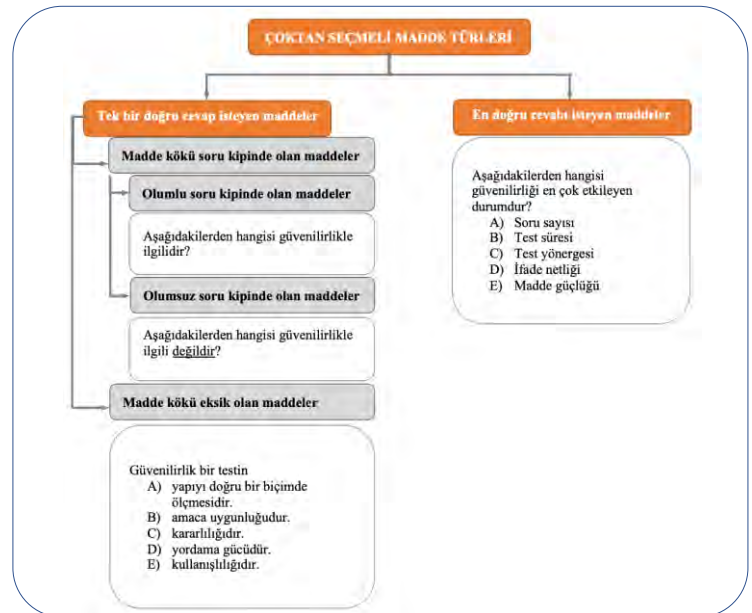
EŞLEŞTİRME MADDELERİ YAZIM KURALLARI:

- 1) Soru grubunda **yönerge/açıklama** yazılmıştır.
- 2) Seçeneklerin kullanılacak mı? / Kaç kez kullanılacağı **belirtilmelidir**.
 - Bir eşleşme takımındaki **öncüller** ile **cevaplar eşit sayıda** olmamalıdır.
- 3) * Eşleştirmede **ortaokul ve üstü öğretim kademelerinde**
 - Öncülün başındaki boşluğa seçeneğin harfini yazmaları istenebilir.
 - * **İlkokul düzeyinde ise**
 - **Öncüllerle seçenekler** arasında oklarla eşleştirme yapmaları istenebilir.
- 4) Okul öncesi, özel eğitimde **eşleştirmede şekil – resim** kullanılabilir.
- 5) Öncüller ile seçeneklerin her biri **benzeşik (homojen)** öğelerden **oluşmalıdır**.
- 6) **Uzun ifadeler** öncül olarak **kullanılmalı**
 - **Öncüller sütunu** sayfanın sol sütununa
 - **Cevaplar sütunu** ise sayfanın sağ tarafına yerleştirilmelidir.
- 7) **Madde sayısı en az 6, en çok 15** olmalıdır.
- 8) Maddelerin tümü **aynı sayfada** bulunmalıdır.
- 9) Cevapların seçileceği sütun, bir **kelime listesi ise**
 - **Alfabetik** sıraya göre (*küçükten büyüğe ya da büyükten küçüğe*) olmalıdır.
- 10) **Hatırlama düzeyi** dışında,
 - **Kavrama düzeyi** öğrenmeler de **yoklanabilir**. (grafik / tablo / metin / harita)



ÇOKTAN SEÇMELİ MADDELERİN ÖZELLİKLERİ:

- 1) Hazırlanması **zor** ancak puanlaması **kolaydır**.
- 2) Puanlama **objektiftir**.
- 3) Üst düzey bilişsel becerilerle ilgili madde **yazmak zordur!**
- Maddeler genellikle **bilgi**, **kavrama** ve **uygulama** basamaklarındadır.
- 4) **Doğru cevabı içindedir:**
=> **Sentez** düzeyinde soru sormak **olanaklı değildir!**
=> **Yaratıcılığın denetlenmesi zordur!**
=> **Puanlara şans başarısı karışma ihtimali** vardır.
- 6) **Çok sayıda soru** sorulabilir.
- 7) **Sistematiik hata karışma ihtimali** çok **düşüktür**.
- 8) **Çeldirme mantığı** temellidir.



ÇOKTAN SEÇMELİ MADDE YAZIM KURALLARI:

- 1) **Bilen bir öğrenci**, seçenekleri görmeden **doğru cevabı verebilmelidir**.
- 2) Öğrencinin henüz öğrenmediği **sözcük, kavram, terim** vb. kullanılmamalıdır.
- 3) **Bir durum, resimle** anlatılıyorsa çok **cümle** yerine **görselden** yararlanılabilir.
- Ancak maddenin **cinsiyet** gibi **yanlılık oluşturmamasına dikkat** edilmelidir.
- 4) **Gereksiz görsel** kullanılmamalıdır.
- 5) Öğrencinin öğretmenin **zihninden geçeni** keşfetmesi beklenmemeli.
- 6) **Karışık bir dil → Soruyu zorlaştırmak değildir!**
- Ortalama bir öğrencinin **anlayabileceği bir dil** kullanılmalıdır.
* Bir çoktan seçmeli maddeyi **zorlaştırmak için:**
 - a) **Çeldiricileri doğru yanıtla yaklaştırmak** (güçlendirmek)
 - b) **Davranışı üst düzey becerileri ölçen bir formda** hazırlamak.
- 7) **Madde kökünde gereksiz sözcük** kullanılmamalıdır.
- Bu hata önemli yerleri **vurgularken** yapılmaktadır.
- 8) **Seçeneklerde → Gereksiz sözcük tekrarı**ndan kaçınılmalıdır.
Tekrar edilen sözcük → Madde köküne alınabilir.
- 9) **Bir maddenin doğru yanıtı**
- **Başka bir maddenin kökünde** ya da **seçeneklerinde** bulunmamalıdır.
- 10) **Seçenekler birbirleriyle tutarlı ve ayrı gramatik yapıya** sahip olmalıdır.
* Eğer **seçenekler** ifade bakımından **birbirinden farklıysa**
- Öğrencilere **ipucu verir**. (Özellikle **eklerin** kullanımına dikkat etmek gerekir.)
- 11) **İpucu vermektan kaçınmak** gerekir.
* “Asla, daima, hiçbir, tümü” vb. ifadelerin → **Yanlış** olma olasılığı yüksektir.
* “Genellikle, bazen, sıklıkla” vb. ifadelerin → **Doğru** olma olasılığı yüksektir.
- 12) Birbirinin **tam zıttı** olan durumların seçeneklerde verilmesi **uygun değildir**.
- Gen. **biri doğrudur. Şans başarısı artar / geçerlilik ve güvenirlilik düşer.** →

- 13) Seçeneklerde birbirini **anlamca içeren** ifadeler kullanılmamalıdır.
- 14) Seçeneklerin **uzunlukları** birbirine **eşit** olmalıdır.
- 15) Doğru cevapların bir **örüntü** göstermemesi gerekir. (Dağıtılırken dikkat!)
- 16) “**Hepsi**” ve “**hiçbiri**” bir seçenek olarak kullanılmamalıdır.
- 17) **Seçenek sayısı** genellikle 4 ya da 5 olsa da seçenek sayısını
- * Öğrencinin içinde olduğu **gelişim dönemi** belirler:
 - Lise ve daha üstü öğretim düzeylerinde → 5
 - Ortaokul düzeyinde → 4
 - İlkokul düzeyinde → 3 seçenek kullanılabilir.
- 18) Seçenekler **bir sıra ile** verilmelidir.
- Alfabetik veya **büyüklik** olarak sıralanmalıdır.
- 19) **Dil bilgisi** ve **imla kuralları** çok önemlidir.
- * **İyi bir madde yazarı**
 - a. Madde yazılan **alanı** çok iyi bilmeli
 - b. Madde yazma **teknik** ve **yöntemlerinden** haberdar olmalı
 - c. Maddelerin yazılacağı **dili kullanmada becerikli** olmalı
 - d. Öğrencilerin gelişim düzeyini çok iyi **bilmelidir**.
- 20) Seçeneklerin **baş harfleri büyük** olmalıdır.
- 21) **Madde kökü** ile **seçenekleri aynı sayfada** olmalıdır.
- 22) **Ortak köke dayalı madde grubu yazılacaksa**
- **Hangi soruların ortak köke göre yanıtlanacağı** belirtilmelidir.
 - (İlkokulda madde sayısını çok artırmamak gerekir.)

ACIK UÇLU SORULAR

- * Açık uçlu sorular ikiye ayrılır [**Pikometri alanyazınında** (literatür)]:
 - a) **Uzun yanıt gerektiren maddeler**
 - b) **Yanıtı sınırlı maddeler**

- * Ancak bu testler (**sahada**):
 - a) **Essay (yazılı yoklama)**
 - c) **Boşluk doldurma (cümle tamamlama)**
 - b) **Kısa yanıt**
 - d) **Sözlü yoklama** olarak geçmekte.

ESSAY (YAZILI YOKLAMA)

- Yazılı yoklamaların temel özellikleri:**
- * Sorulara cevap verme konusunda **sınırsız bir özgürlüğü** vardır.
- * **Sınırsız cevap özgürlüğünün:**
 - Avantajı: Öğrencinin **bilgisi** ile **zihninin nasıl işlediğini** anlama olanağı sağlar.
 - Dezavantajı: Öğrenci **sınırlı bilgiye sahipse** soruyla **ilgisi olmayan, aklına estiği gibi** uzun cevaplar yazabilir.

- * **Sınırsız cevap özgürlüğü**
 - **Anlatım becerisi iyi olan öğrencilere** bir **avantaj** sağlayabilir.

Essay (Yazılı Yoklama)

- * Dünyada **en eski** ve **hâlâ en çok tercih edilen** testlerden biridir.
- * Bunun iki nedeni vardır:
 - a. **Pratik nedenler:** Hazırlaması kolaydır.
 - b. **Eğitsel nedenler:** Üst düzey bilişsel becerileri ölçmek için geleneksel yöntemler arasında **en avantajlı** yazılı test etme yoludur (Sentezde tek yol.).
- * **Yazma, problem çözme, bilgileri organize etme, analiz vb. özgün bir ürün ortaya koyma (yaratıcı düşünme), eleştirel düşünme, yeni durumlara beceriyi transfer etme, analitik düşünme, bilimsel düşünme, hipotez oluşturma ve neden-sonuç ilişkilerini açıklama, veri düzenleme, güçlü ve zayıf yönleri belirleme gibi becerileri ölçmek için çok kullanışlıdır.**

Essay (Yazılı Yoklama)

- * **Çok fazla soru** sorulmıyor → Bir **dezavantajıdır**.
- * **Puanlar**, sadece **bilgiyi değil**, öğrencinin
 - **Anlatım biçimini, yazı güzelliğini, bilgisini örgütleme biçimini** de yansıtır.
 - Bu ise **ölçme sonuçlarındaki sistematik hata miktarını** artırabilir.
- * **Puanlama özneliği** belki de **en önemli dezavantajıdır**.

“Essay”de Nesnelliği Artırma Yolları:

- 1) **Ayrıntılı bir cevap anahtarı** çıkarılmalıdır.
- 2) Öğrenci **ne yaparsa kaç puan verileceğinin** belirlenmesi gerekir.
- 3) **Bütüncül (holistik)** ya da **analitik rubrik** hazırlanabilir.

KISA CEVAPLI SORULAR

- * **Kısa cevaplı testler**
 - Bir **sözcük, rakam, tarih** ya da **en çok bir cümle** ile **cevaplanan** sorulardan oluşur.
 - Öğrencinin aklının estiği cevabı **yazmasını engeller**.
 - Cevaplar çok **kısa**dır.
 - (Testlerin puanlanmasını genellikle **nesnel** kılar. ,
 - Eğer konu **kapsamı** bakımında **homojen** bir yapıda ise
 - **Gruplandırılarak sorulması daha uygun** olacaktır.
 - **Uzun cevap** gerektiren sorularda **olduğu gibi** ayrıntılı cevap anahtarı olmalıdır.

Avantajları:

1. **Hazırlaması, uygulaması ve puanlaması kolay** (kullanışlıdır)
2. **Soru sayısı artırılabilir.**
3. **Sistematik hata karışma olasılığı pek yoktur.**

Dezavantajları:

- * **Üst düzey bilişsel özelliklerin ölçülemediği** yine bir **dezavantajıdır**.

- * **Cümle tamamlama** (boşluk doldurma), **kısa cevaplı / sınırlı cevap** gerektirir.
- * Öğretmen
 - **Önemli** gördüğü bir **cümleyi** alır,
 - **Kritik** gördüğü bir ya da birden fazla **yeri çıkarır**
 - **Yerine bir boşluk** koyar
 - Öğrenciden o cümleyi **anamlı bir biçimde tamamlaması** istenir.
- * **Hazırlanması** oldukça **kolaydır**.
- * **Kim, ne, nerede, ne zaman?** sorularına cevabını **ölçmeye** sorulardır.
- * **Cümleler** ders kitabından **aynen alınmamalı**
 - Öğretmen **kendi** sözcükleriyle **yeniden** yazmalıdır.

Avantajları:

- a. **Hazırlaması, uygulaması ve puanlaması kolaydır** (Kullanışlı)
- b. **Soru sayısı** artırılabilir.
- c. **Sistematik hata karışma olasılığı pek yoktur!**

Dezavantajları:

- **İyi yazılmazsa** genellikle **üst düzey bilişsel özelliklerin** ölçülemediği

CÜMLE TAMAMLAMA SORULARI YAZIM KURALLARI:

- 1) Her **kısa cevap maddesinin** → Yalnızca **tek doğru cevabı olmalı** (açık-net)
- 2) **Maddenin cevabının bulunmasında** → **İpuçları** verilmemeli.
- 3) Cümlelerin **gramer yapısında** → **İpucu** genellikle **eklerle** verilmektedir.
- 4) **Cevapta birden çok ayrıntı** varsa → O ayrıntıların her birine **ayrı puan** vermeli.
- 5) **İlköğretimde cevapların** ayrı bir **kâğıdına** işaretletilmesi **uygun değildir!**
 - Ancak ortaöğretimde olabilir.
- 6) Bir cümlede **çok sayıda boşluk** bırakılmamalı
- 7) **Cümleden sadece anahtar** nitelikli **anamlı** ve **önemli** sözcükler **çıkartılmalıdır**.
- 8) Herkesçe **aynı biçimde** alıngınlacak biçimde **soru yazmak** gerekir.
 - Öğrenciye “*Acaba öğretmenim bununla neyi kastediyor?*” dedirtmemelidir.

SÖZLÜ YOKLAMA

- * **Sözlü sınavlar ciddi bir ölçme hatası barındırma potansiyeline** sahiptir.
- * **Dile dayalı becerilerin ölçülmesi** için **geleneksel** yöntemler içindeki **tek türdür**.
 - “Sözlü anlatım, diksiyon, yabancı dil dersinde konuşma, solfej ve şarkı söyleme...”
- * Sorular **sözlü olarak sorulur** ve **cevap sözlü olarak verilir**.
- * Öğretmen ile öğrenci arasında **karşılıklı ve devamlı** bir **etkileşim** vardır.
- * **Bireysel bir testtir** → Her öğrenciye **ayrı soru sorma zorunluluğu vardır**.
- * **Bireysel bir testtir** → Tüm öğrencileri test etmek için **çok süre gereklidir**.
- * Öğrencinin cevaplarını gözden geçirme **şansı yoktur!**
- * **Cevaplarda** çoğu zaman **düşünme** ve **tasarlamaya** pek **zaman vermez!**
- * **Puanlama hemen yapılır** → Cevap **doğrulama genel izlenimle** yapılır.
- * Öğrencinin **ayağa / tahtaya kaldırılması ve kaygı** ...vb
 - ⇒ **Bireyden kaynaklanan hata miktarını** artırabilir.
- * Öğretmenin **ölçme sonucuna sistematik hata karıştırma olasılığı** yükselir.
- * **Sözlü anlatım becerisi** iyi olan (daha az biliyor olsa da) → **Yüksek puan** alabilir.
- * Öğrencinin **kılık-kıyafeti, diksiyonu, hâl ve hareketleri**
 - ⇒ Puanlamaya **artı** ya da **eksi** yönde **etki** edebilir.
- * **Çok fazla soru sorulamayacağı için** → **Kapsam geçerliliği problemi** olabilir.

SÖZLÜ YOKLAMA UYGULAMA KURALLARI:

- 1) **Sözlü sınavlarda** da (yazılı sınavlar gibi) yapılacak **gün** ve **saat** ilan edilmeli
- 2) Sorular **önceden hazırlanmalı**
- 3) Sorular **amaca uygun** ve **her öğrenciye benzer güçlükte** olmalı
- 4) **Cevap anahtarı** hazırlanmalı
- 5) Sınavda öğrencilerin **kaygısını** artıracak **davranışlardan** kaçınılmalı
- 6) **Hazırbulunuşluluk** artırılmalı
- 7) **Basit olan sorudan** başlanmalı
- 8) **Kaydedilemiyorsa** puanlama **hemen** yapılmalı
- 9) **Puana ilişkin dönüt** **hemen** verilmeli
- 8) **Puanlamada sistematik hata kaynaklarına karşı dikkatli** olunmalı

Not: Yoklanacak olan **davranışlar**, **yazılı sınav türleriyle de yoklanabiliyorsa** ⇒ **Sözlü yoklama** yapmaktan **kaçınılmalıdır**.

GENEL DEĞERLENDİRME

- 1) Her soru türünün **avantajları** olduğu gibi **dezavantajları** da vardır.
- 2) **Dezavantajları asgariye indirmenin yolu**
 - ⇒ **Madde / soru çeşitlenmesi** yapmaktır (İlkokulda çok artırmamak gerekir)
- 2) **Geleneksel test etme yaklaşımları ürün odaklıdır**.
 - ⇒ Eğitim sürecinde **ürün** elbette **önemlidir** ancak
 - * **Ürün kadar**
 - Öğrencinin süreçte **bilgiyi nasıl yapılandırdığının izlenmesi** önemlidir.
 - ⇒ **Geleneksel ölçme yöntemleri destekleyici yöntemlerle desteklenmelidir**.