

GEÇERLİLİK SORGULAMA YÖNTEMLERİ



Şekil 4: Geçerlilik yöntemleri sınıflaması

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 olmalı)

Kapsam Geçerliliği Sorgulama Yöntemleri

1) Mantıksal / rasyonel yöntemler:

a) Belirtke tablosunun hazırlanması:

* **Ölçülen kapsam dâhilinde davranışların belirlenmesi yöntemidir.**

b) Uzman görüşüne başvurulması:

* **Uzmana belirtke tablosu ve maddeler sunulur:**

- Soruların konu kapsamını
- Yeterli ve dengeli bir ölçme durumu
- Soruların kazanımlarla ilgili değerlendirilmesi istenir.

2) İstatistiksel yöntemler:

* **Uzmanlardan alınan dönütler** çözümlenebilir ya da
- Uzmanların değerlendirmesi istenir. (Uzmanlar arasında uyum test edilir.)

GÜVENİLİRLİK

Güvenilirlik **tesadüfi hatalardan arınık olma durumudur.**
Güvenilir bir ölçme aracı

Aynı özelliklerle ilgili arka arkaya ölçmelerde → Aynı sayısal sonucu verir.

İki Ölçüm Arasındaki Uyumsuzluk Nedenleri?

- 1) Bir hareketten diğerine, "dikkat ve çaba" değişebilir.
- 2) Uzun periyotlarda puan değişmesi
- 3) İkinci ölçümde daha açık/kapalı soruların kullanılması

Gerçek Puan – Standart Sapma

* **Ölçme, hatalar giderilene kadar sürdürülürse Gerçek puan** elde edilir.
- Ancak davranış örneği sınırlı → Gözlenen puan gerçek puandan **farklıdır.**
- Buradaki farklılık → **Ölçme hatasıdır.**

"Bir koşucu farklı zamanlarda 23.7, 24.0, 25.1... saniyelerde koşmuş olsun.
- Bu ölçümlerin ortalaması **24.7** ise → **Gerçek puanı 24.7** dir.

* **Ölçmenin standart hatası:** $S_h = s\sqrt{1-r_x}$ formülü ile bulunur.
(S_h = Ölçmenin standart hatası s = Standart sapma r_x = Güvenilirlik katsayısı)

* Bir ölçme işleminde **standart hatanın düşük** olması:

1. **Güvenilirlik katsayısının** → **Yüksek** ve
2. **Standart sapmanın** → **Düşük** olmasına bağlıdır.

* **Ölçmenin standart hatası yükseldikçe** (Bireylere rastgele puan vermiş gibidir)
- **Güvenilirliğin** kestirilmesi çok önemlidir.

* **Güvenilirlik katsayısı 0 ile 1** arasında değişir.

- **Kabul edilen aralık: 0.70** ve **üstü ölçüt**
- **Genel yetenek gibi bilişsel özellikli ölçen testlerde**
⇒ **Güvenilirlik katsayısının 0.90** ve **üzerinde** olması arzu edilir.

Şekil 8: Güvenilirlik katsayısı



GÜVENİLİRLİK KESTİRİM YÖNTEMLERİ



Şekil 9: Güvenilirlik kestirim yöntemleri

GÜVENİLİRLİK (KESTİRİM) 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

A) BİR DEN ÇOK UYGULAMAYA DAYALI GÜVENİLİRLİK YÖNTEMLERİ

1) TEST-TEKRAR TEST YÖNTEMİ:

* Bir test, aynı gruba, belli bir zaman aralığıyla **iki kez** uygulanır.
* İki uygulamanın **puanı** arasındaki **korelasyon** hesaplanır.

⇒ **Korelasyon katsayısına** elde edilir. (**Kararlılık katsayısı**)

* **Test-tekrar test** yöntemi

* **T.T.T. Yöntemi** Benzer ölçmeler geçerliliği, hâlihazır geçerlilik, uygunluk geçerliliği olarak da geçer.

	Anksiyete Ölçeği Puanları X ₁	Anksiyete Ölçeği Puanları X ₂
Çaren	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

Kararlılık Katsayısı

Korelasyon (r)
1'e yakın

Şekil 10: Test-tekrar test yöntemi çalışma örneği

Not: Güvenilirliğin **kararlılık, tutarlılık** ve **duyarlılık** sorunu vardır.

- * **T.T.T. Yöntemi** ile güvenilirlik kanıtı → **Kararlılık** boyutuyla ilgilidir.
- * **T.T.T. Yöntemi** → **Güvenilirliğe** ilişkin **tümel bir sonuç ortaya koymaz!**

* **T.T.T. Yöntemi** → Özelliğin **kararlı** olduğu durumlarda uygulanmalıdır.
(Testin ölçtüğü özellik **sürekli değişkenlik** gösteriyorsa kullanılmamalıdır!)
"Genel zihin yetenekleri, kişilik testleri, ilgi envanterleri, tutum ölçekleri" vb.

* **T.T.T. Yöntemi** → **Yetenek testleri, kişilik envanterlerinde** tercih edilir.

* **T.T.T. Yöntemi** → **Başarı testleri** için pek tercih **edilmez!**

2) EŞDEĞER (PARALEL) TESTLER YÖNTEMİ

Bu bölüm ana PDF'de atlanmış!

B) TEK UYGULAMAYA DAYALI GÜVENİLİRLİK YÖNTEMLERİ

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

* **Test iki eşdeğer yarıya** bölünür.

- İki yarıdan alınan **puanlar** arasındaki **tutarlılık** incelenir.

* Yarılardan alınan **toplam puanlar** arasında **korelasyon** hesaplanır.

* Ancak elde edilen katsayı **korelasyonun yarı puanlardan** elde edildiği için

- Testin **tamamına** ilişkin bir **tutarlılık** bilgisi **vermez!** (Yarı güvenilirlik)

- Testin tümüne ilişkin bir güvenilirlik katsayısı: **Spearman-Brown** formülü

* Bir test iki eşdeğer yarıya **nasıl** bölünür?

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

(Her teste uygun değil!)

* Basitten zora ise

* Kümelene varsa

* Madde sayısı çok ise

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

Testin yarılarına ilişkin tutarlılığı verir. Bu korelasyon testin tamamına nasıl genellenir?

Korelasyon (r)

Spearman-Brown Eşitliği

İç tutarlılık katsayısı

Şekil 11: Test yarılama yöntemi çalışma örneği

B) TEK UYGULAMAYA DAYALI GÜVENİLİRLİK YÖNTEMLERİ (DEVAM)

2-3) KUDER-RICHARDSON 20 VE 21 YÖNTEMLERİ:

- * Kuder-Richarson 20 ve 21 (KR-20 ve KR-21)
 - Test kendi içinde **tutarlı bir bütün** oluşturmuş mu? (**İç tutarlılık**)
 - **İç tutarlılık katsayısı** elde edilir.
- * Puanlamanın **katégorik olduğu** ölçme araçları için uygundur.
 - (D:1 Y:0 Boş:0 puan)
 - (İki seçeneği cevap)
- * Eğer **maddeler farklı ağırlıklarla puanlanmışsa** ya da **Puanlar şans başarısı için düzeltilmişse** => Bu formüller **kullanılmaz!**

4) CRONBACH ALFA YÖNTEMİ:

- * Ölçme aracının puanlaması **çok katégorili** ise kullanılır.
- KR-20 ile **aynı mantık** üzerine kuruludur.
- * Elde edilen katsayının ismi yine **İç tutarlılık katsayısıdır.**

5) HOYT'UN VARYANS ANALİZİ YÖNTEMİ

Bu bölüm ana PDF'de atlanmıştır!

6) MCDONALD OMEGA YÖNTEMİ

Bu bölüm ana PDF'de atlanmıştır!

GÜVENİLİRLİĞİ VE GEÇERLİLİĞİ ARTIRMA YOLLARI

Güvenilirliği Artıran Temel Kavramlar ve Eylemler

(Güvenilirlik **tesadüfi hatalardan** arınık olma durumudur)

1. Testteki **madde sayısı arttıkça** → **Duyarlılık artar / Güvenilirlik artar.**
2. Bir ölçmede **puanlayıcı sayısı arttıkça** → **Güvenilirlik artar.**
3. **Puanlama nesnelligi** → **Güvenilirliği etkiler.**
 - Testi farklı kişiler veya aynı kişi farklı zamanlarda **puanlarsa** => Puanlar arasındaki **tutarlılığa puanlama güvenilirliği** denir.
- * Testin puanı
 - => Puanlayıcıya ya da zamana göre değişmiyorsa → **Güvenilirlik artar.**
- * Nesnel Araçlar: “Çoktan seçmeli, derecelendirme ölçüğü, kontrol listeleri...”
Öznel Araçlar: “Açık uçlu sorulardan oluşan araçlar, mülakatlar vb...”
=> **Öznel araçlar puanlamada nesnel hâle getirilirse** → **Güvenilirlik artar.**
4. * Testteki maddeler açısından
 - **Benzeşiklik (homojenlik) arttıkça** → **Güvenilirlik artar.**
(Benzer maddeler arasından seçim zorlaşır. Tesadüfi cevap azalır)
 - **Ayrışıklık (heterojenlik) arttıkça** → **Güvenilirlik düşer**
5. - Veri elde edilecek **grubun büyüklüğü arttıkça** → **Güvenilirlik artar.**
(Grup üyelerinin fazlalığı, elde edilecek veriyi çeşitlendirir- Grup **Heterojenleşir**)
6. * **Maksimum performansı ölçen** testlerde
 - **Ortalama güçlüğe yaklaştıkça** ($P = .50$) → **Güvenilirlik artar.**
(Grup **Heterojenleşir**)
7. - Bir **madde tüm öğrencilerce aynı anlaşılması** → **Güvenilirliği artırır.**
(Maddelerin dil bilgisi kurallarına uygun, açık ve anlaşılır yazılması)
8. **Fiziksel ortamın uygun hâle getirilmesi** → **Güvenilirliği artırır.**
(Ses, sıcaklık, ışık, koku, dikkat dağıtıcı uyarıcılar, ergonomi vb.)
9. Öğrencilere **yönerge vermek**
 - Test almayı → **Güdüler** **Hazırbulunuşluğu** → **Artırır.**
 - * **Bireyden kaynaklanabilecek hataları asgariye indirir.**
10. Eğer **sürelili 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** (Okuma **güçleştirilmemeli**)
 - Test katılımcısının yaşına, gelişim düzeyine uygun bir **punto**
 - Baskı hataları **bulunmamalı**
13. **Tesadüfi - sabit ve sistematik hatalardan** arındırmalıyız.

TEST GELİŞTİRME VE MADDE / SORU TÜRLERİ

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. Sürelili	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

Test Geliştirme İşlem Basamakları:

1. Amaç belirleme
2. Kapsam belirlenme - belirtke tablosu hazırlama
3. Denemelik maddeleri / soruları yazma
4. Maddeleri gözden geçirme (redaksiyon),
5. Denemelik test formu hazırlama
6. Test uygulama
7. Test ve madde istatistikleri hesaplama
8. Seçilen maddelerden nihai form oluşturma

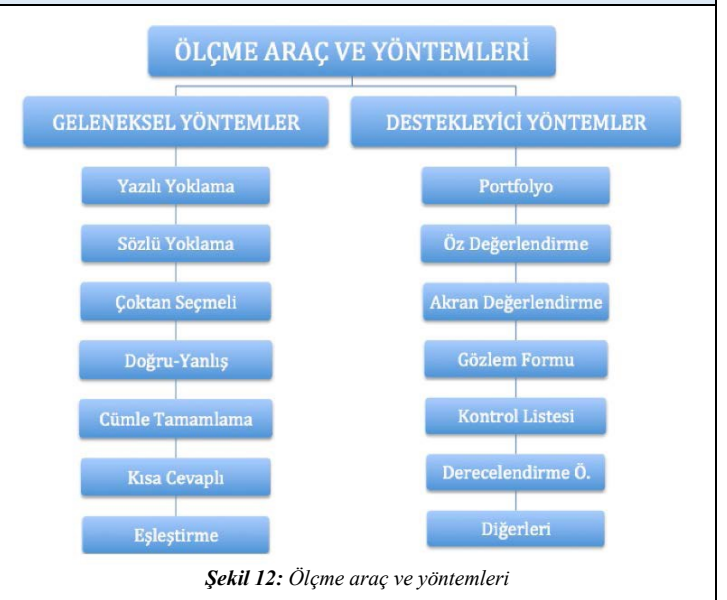
TEST GELİŞTİRME

Test Geliştirme İşlem Basamakları:

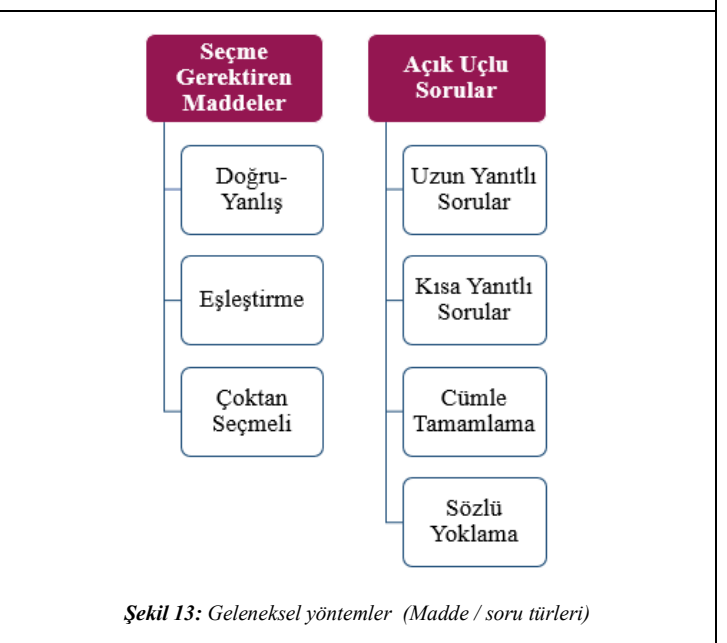
(Öğretmenler açısından daha olanaklı bir test planı)

1. Amaç - davranışları - konuları belirleme
2. Sınav günü - sınav süresi belirleme
3. Belirtke tablosu hazırlama
4. Soru türleri - soru sayısı belirleme
5. Soruların “ortalama güçlüğü” - “güçlük dağılımı” belirleme
6. Soru yazımında izlenecek yolu belirleme
7. Cevap anahtarı - puanlama yöntemi belirleme
8. Ölçme aracını yazma - çoğaltma yöntemi belirleme
9. Sınavın uygulanma kurallarını belirleme
10. Sınavı uygulama - Test ve madde istatistiklerini hesaplama

BAŞARININ ÖLÇÜLMESİNDE YÖNTEMLER



GELENEKSEL YÖNTEMLER (MADDE / SORU TÜRLERİ)



GELENEKSEL ÖLÇME YÖNTEMLERİ

1) DOĞRU-YANLIŞ MADDELERİ:

- * İfadelerin **doğru mu, yanlış mı** olduğunu denetler.
- * **Doğru- Yanlış** maddelerinde verilen bir tür "**önerme**" dir.
" Cumhuriyet 1920 yılında kurulmuştur (D) (Y) "
- * **Çoktan seçmeli madde türünden sonra en çok kullanılan 2. madde** tipidir.

Avantajları:

1. Hazırlanması, uygulaması ve puanlaması kolay(kullanışlı)
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 ölçemez.
3. Öğrenme eksiklerini belirleyemez.
4. Yanlış üzerinden öğretim yapılmaz.
5. Her derste/konuda kesin yanlış denebilecek durumlar belirlemek zordur.

- * **İpucu** vermekten kaçınılmalıdır:
"Asla, daima, hiçbiri, tümü" vb. ifadelerin → **Yanlış** olma olasılığı yüksektir.
"Genellikle, bazen, sıklıkla" vb. ifadelerin → **Doğru** olma olasılığı yüksektir.

2) EŞLEŞTİRME MADDELERİ:

- * Birbirleriyle ilgili **iki grup** bilginin **öğelerinin eşleştirilmesini** gerektirir.

Avantajları:

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

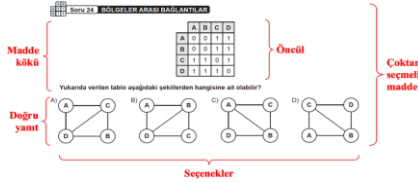
Dezavantajları:

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

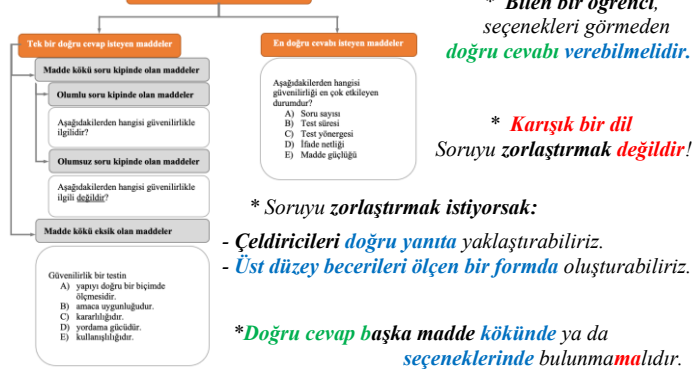
- * **Uzun ifadeler öncül** olarak kullanılmalı
- **Öncüller sütunu sayfanın soluna / Cevaplar sağ tarafına** yerleştirilmelidir.
- * **Madde sayısı en az 6, en çok 15** olmalıdır.

3) ÇOKTAN SEÇMELİ MADDELER

- 1) Hazırlanması zor puanlaması kolaydır.
- 2) Puanlama objektiftir.
- 3) Üst düzey becerilerle ilgili madde **yazmak zordur!**
- 4) **Doğru cevabı içindedir:**
=> **Sentez** düzeyinde soru sormak **olanaklı değildir!**
=> **Yaratıcılığı denetlenmesi zordur!**
=> **Puanlara şans başarısı karışma ihtimali** vardır.
- 5) **Çok sayıda soru** sorulabilir.
- 6) **Sistematik hata karışma ihtimali** çok düşüktür.
- 7) **Çeldirme mantığı** temellidir.



ÇOKTAN SEÇMELİ MADDE TÜRLERİ



- * **Seçenekler birbirleriyle tutarlı** olmalıdır. → **İpucu vermemelidir.**

* Bazı kelimeler **ipucu** vermek demektir!

"Asla, daima, hiçbiri, tümü" vb. ifadelerin → **Yanlış** olma olasılığı yüksektir.
"Genellikle, bazen, sıklıkla" vb. ifadelerin → **Doğru** olma olasılığı yüksektir.

- * Birbirinin **tam zıttı** olan durumların seçeneklerde **verilmemelidir.**
- **Biri doğru** olacağından **şans başarısı artar / geçerlilik ve güvenilirlik düşer.**

- * Doğru cevapların bir **örüntü** göstermemesi gerekir. (Dağıtılırken dikkat!)

- * "**Hepsi**" ve "**hiçbiri**" bir seçenek olarak kullanılmamalıdır.

- * **Seçenekler bir sıra ile** verilmelidir. (Alfabetik veya büyüklük)

GELENEKSEL ÖLÇME YÖNTEMLERİ (DEVAM)

4) AÇIK UÇLU SORULAR

- * **Açık uçlu sorular ikiye ayrılır** [Pikometri alanyazımında(literatürde)]:

- a) Uzun yanıt gerektiren maddeler
- b) Yanıtı sınırlı maddeler

- * **Ancak bu testler (sahada):**

A) ESSAY (YAZILI YO KLAMA):

- * Sorulara **cevap** vermede **sınırsız bir özgürlüğü** vardır.
Avantajı: - Öğrencinin bilgisi ile zihninin nasıl işlediğini anlayabilme
- Anlatım becerisi iyi olan öğrencilere avantaj sağlar.
Dezavantajı: Sınırlı bilgi ile aklına estiği gibi uzun cevaplar yazabilir.

- * **Dünyada en eski ve en çok tercih edilen testlerden biridir.**
- **Pratik nedenler:** Hazırlanması kolaydır.
- **Eğitsel nedenler:** Üst düzey bilişsel becerileri ölçmede geleneksel yöntemler arasında en avantajlı olanıdır.
- * **Yazma, problem çözme vb. becerileri ölçmek için çok kullanışlıdır.**
- * **Çok fazla soru sorulamıyor** → **Bir dezavantajdır.**

- * **Puanlar, sadece bilgiyi ölçmemesi** → **Sistematik hata miktarını artırabilir.**
Puanlama özneliği belki de en önemli **dezavantajdır.**

"Essay"de Nesnellığı 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.

B) KISA CEVAPLI SORULAR:

- * **Bir sözcük, rakam, tarih ya da en çok bir cümle ile cevaplanır.**
- * **Öğrencinin aklının estiği cevabı yazmasını engeller. (Nesnellik artar)**
- * **Uzun cevaplı** sorularda olduğu gibi **ayrıntılı cevap anahtarı** olmalıdır.

Avantajları:

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

Dezavantajları:

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

C) CÜMLE TAMAMLAMA (BOŞLUK DOLDURMA) SORULARI:

- * **Cümle tamamlama** (boşluk doldurma), **kısa cevaplı / sınırlı cevap** gerektirir.
- * **Hazırlanması oldukça kolaydır.**
- * **Kim, ne, nerede, ne zaman?** sorularına cevap olabilecek **ölçme** sorularıdır.
- * **Cümleler ders kitabından aynen alınmamalı (Kendi sözcükleriyle tekrar...)**

Avantajları:

- a. Hazırlanması, 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**

- * **Her maddenin** → Yalnızca **tek doğru cevabı olmalı** (açık-net)
- * **Maddenin cevabının bulunmasında** → **İpuçları** verilmemeli.
- * **İlköğretimde cevapların ayrı bir kâğıdına** işaretlenmesi **uygun değildir!**
- * **Bir cümlede çok sayıda boşluk** bırakılmamalı
- * **Cümleden sadece anahtar nitelikli anlamlı ve önemli sözcükler çıkarılmamalıdır.**

D) SÖZLÜ YO KLAMA:

- * **Sözlü sınavlar ciddi bir ölçme hatası barındırma potansiyeline** sahiptir.
- * **Dile dayalı becerilerin ölçülmesi için geleneksel yöntemler içindeki tek türdür.**
"Sözlü anlatım, diksiyon, y. dil dersinde konuşma, solfej ve şarkı söyleme..."
- * **Sorular sözlü olarak sorulur ve cevap sözlü olarak verilir.**
- * **Bireysel bir testtir** → Her öğrenciyi ayrı soru sorma zorunluluğu vardır.
- * **Bireysel bir testtir** → Tüm öğrencileri test etmek için çok süre gereklidir.
- * **Cevaplarda çoğu zaman düşünme ve tasarlamaya pek zaman vermez!**
- * **Öğrencinin ayağa / tahtaya kaldırılması ve kaygı ...**
=> **Bireyden kaynaklanan hata miktarını artırabilir.**
- * **Öğretmenin ölçme sonucuna sistematik hata karıştırma olasılığı yükselir.**
- * **Sözlü anlatım becerisi iyi olan** (daha az bilse de) → **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 sınavlarda da** (yazılı sınavlar gibi) **gün ve saat** ilan edilmeli

- * **Sorular ve cevaplar önceden hazırlanmalı**

- * **Basit olan sorudan** başlanmalı ve **puana ilişkin dönüt hemen** verilmeli

Not: Davranışlar, yazılı sınavla yoklanabiliyorsa sözlü sınav yapılmamalı!