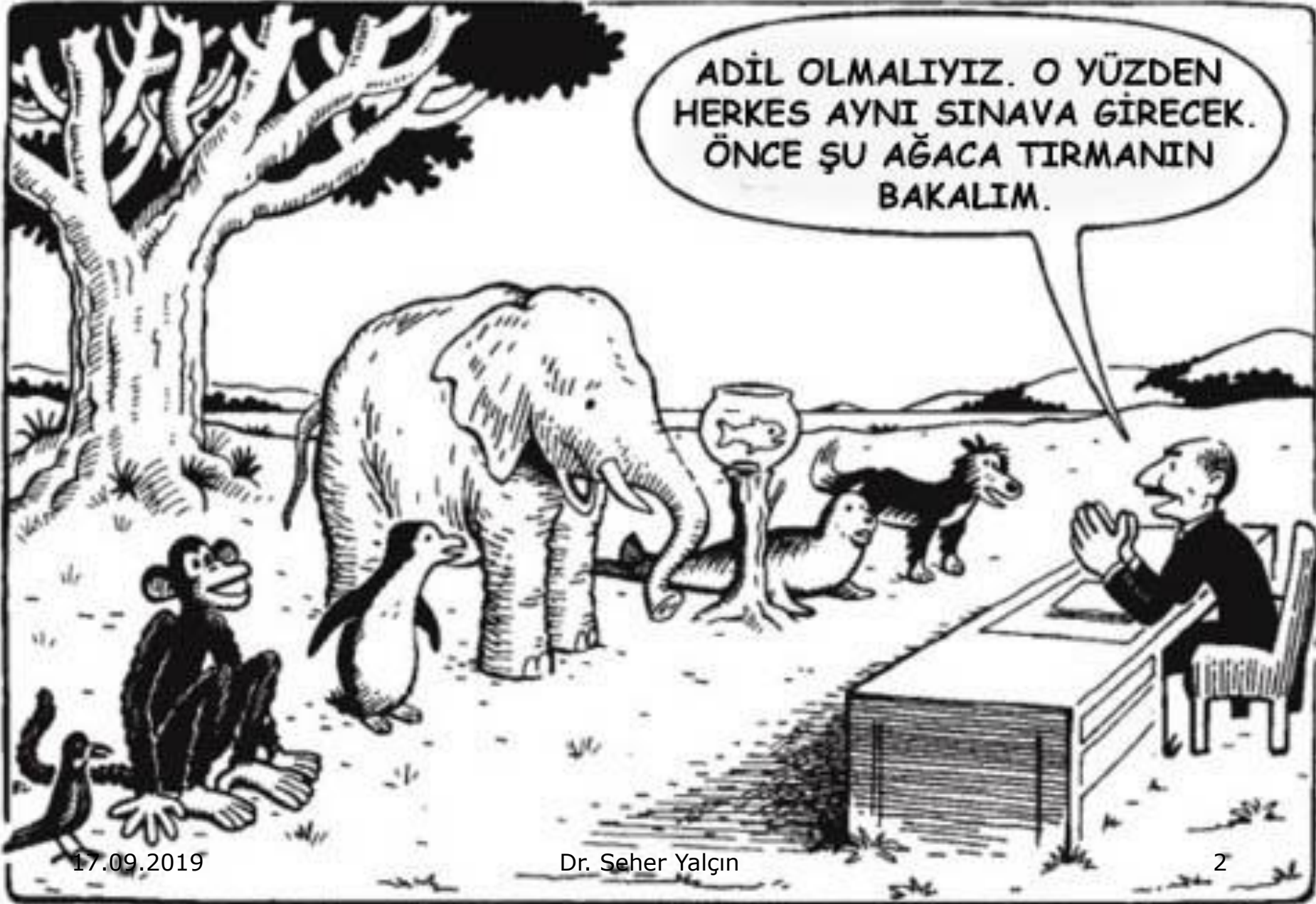


# BİLİMSEL ARAŞTIRMA YÖNTEMLERİ

---

## **Ölçme Araçlarının Taşınması Gerekten Özellikler: Geçerlik ve Güvenirlik**

Doç. Dr. Seher Yalçın



ADİL OLMALIYIZ. O YÜZDEN  
HERKES AYNI SINAVA GİRECEK.  
ÖNCE ŞU AĞACA TIRMANIN  
BAKALIM.

# ÖLÇME HATASI

---

- Hata, ölçülen özelliğın gerçek değeri ile ölçme sonucunda elde edilen (gözlenen) değeri arasındaki farktır (Atılgan, Kan ve Dođan, 2011).
- Dolaylı ölçmeler hatayı arttırır.
- Gerçek Puan = Gözlenen Puan + Hata <sup>3</sup>

# Hata Kaynakları (Atılğan ve diğ., 2011; Baykul, 2000)

---

- **Ölçmeci** (dikkatsizlik, öznel puan verme, yorgunluk)
- **Ölçme aracı** (yönerge, kapsam, maddelerin kapsamı örnekleme durumu, anlaşılabilirliği, uzunluğu)

# Hata Kaynakları (Atılğan ve diğ., 2011; Baykul, 2000)

---

- **Ölçme işleminin yapıldığı ortam**  
(sessiz, ışık)
- **Ölçme işleminin yapıldığı grup**  
(dikkat, stres)
- **Ölçme yöntemi** (psikomotor-  
essey)

# Hata Türleri (Tekin, 1996)

---

- **Sabit Hata** (bir ölçmeden diğerine miktarı değişmeyen)
  - Ölçme aracı, ölçmeciden kaynaklı olabilir.
  - Doğrudan geçerliği düşürür.
- **Sistemik Hata** (ölçülen büyüklüğe, ölçmeciye ve ölçme koşullarına bağlı olarak miktarı değişen)
  - Ölçme aracı, ölçmeciden kaynaklı olabilir.
  - Doğrudan geçerliği düşürür.

# Hata Türleri (Tekin, 1996)

---

- **Tesadüfi Hata** (yönü, kaynağı ve miktarı bilinmeyen)
  - Ölçmeci, şans başarısı, öğrenciler, ölçme aracı, ortam vb kaynaklanabilir.
  - Doğrudan güvenilirliği dolaylı olarak geçerliği düşürür.

# Geçerlik

---

- ❑ Bir ölçme aracının amaca hizmet etme derecesidir.
- ❑ Bir ölçme aracının ölçmek istediği özelliği başka özelliklere karıştırmadan ölçebilmesidir (Tekin, 1996).
- ❑ Geçerlik, ölçme aracının bir özelliğinden ziyade o ölçme aracından elde edilen puanların anlamıyla ilgilidir.



# Geçerlik

---

- Bir testin geçerliğinden çok belirli bir grup ve belirli bir amaç için geçerliği söz konusudur (Köse, 2012).
- Geçerlik katsayısı -1 ile +1 arasında değişir.
- Temel olarak geçerlilik belirleme yöntemlerinin hepsi bireyin testteki performansı ile bireyin o özellikle ilişkili gözlenebilir davranışları arasındaki ilişkiye dayanır (Köse, 2012).

# Geçerlik Kanıtlama Yöntemleri



# Kapsam Geerliđi

---

- Ölme aracının ölçmeyi amaçladığı davranış evrenini temsil etme gücü olarak açıklanabilir (Köse, 2012).
- Bir testin bu testle ölçölmek istenen davranışları ne derece kapsadığı ile ilgilidir (Atılgan ve diđ., 2011).

# Kapsam Geçerliđi (Köse, 2012)

---

- Kapsam geçerliđi yüksek öğretmen yapımı bir test, öğretmenin öğrettiđi konu alanını temsil etmelidir (Köse, 2012).
- Kapsam geçerliđini belirlemede, iki farklı yöntem kullanılmaktadır. Bunlar;
  - mantıksal (yargısal) yaklaşım ve
  - istatistiksel yaklaşımdır.

# Kapsam Geçerliđi (Atılgan ve diđ., 2011)

---

□ Kapsam geçerliđini belirlemede en çok kullanılan yöntem mantıksal yaklaşım içerisindeki uzman kanılarına başvurmadır.

□ Bunun için ölçme aracındaki maddelerin dağılımının ölçme aracının ölçmeye yöneldiđi davranışları kapsayıp kapsamadığına bakılmaktadır.

Belirtke tablosu, bir boyutunda bir derse ya da alana ait konu ve içerik diğer boyutunda öğrencilere kazandırılacak hedef ve davranışları içeren iki boyutlu bir tablodur (Atılgan ve diğ., 2011). Aşağıda ölçme dersi için hazırlanan belirtke tablosunda her bilişsel düzeyden kaç soru yer aldığı tabloda verilmiştir.

	<b>Bilgi</b>	<b>Kavrama</b>	<b>Uygulama</b>	<b>Analiz</b>	<b>Sentez</b>	<b>Değerlendirme</b>
Temel kavramlar	3	1	2	1	-	-
Ölçme araçlarının özellikleri	2	3	2	1	2	2
Madde türleri	4	2	1	2	1	1
<b>Toplam</b>	<b>9</b>	<b>6</b>	<b>5</b>	<b>4</b>	<b>3</b>	<b>3</b>



# Kapsam Geerliđi

---

□ Kapsam geerliđi **istatistiksel** olarak, geliřtirilen testle aynı kapsamı ölçtüđü bilinen bir testin geliřtirilen testle aynı gruba uygulanarak iki testten elde edilen puanlar arasındaki korelasyonla hesaplanmaktadır (Tavřancıl, 2010).

# Kapsam ve Görünüş Geçerliđi

(Köse, 2012)

---

- Kapsam geçerliđi ile görünüş geçerliđi farklı kavramlardır.
- Görünüş geçerliđi teknik anlamda bir geçerlik türü olmayıp, testin ölçmek istediđi özelliđi ölçüyor görünmesi olarak algılanabilir.



# Yapı Geçerliđi

---

- Yapı sözcük anlamı olarak doğrudan gözlenemeyen psikolojik özellik anlamına gelmektedir (Köse, 2012).
- Yapı, birbiri ile ilgili olduđu düşünölen belli öğelerin ya da öğeler arasındaki ilişkilerin oluşturduđu bir örüntüdür. Yetenek, zeka, güvensizlik, içe dönüklük, dışa dönüklük vb. özellikler psikolojide yapı olarak adlandırılırlar (Tekin, 1996).

# Yapı Geçerliđi

---

- ❑ Ölçme aracının yapı geçerliđi ise ölçme aracının ölçmeyi amaçladığı psikolojik özelliđi temsil etme gücü olarak tanımlanabilir (Köse, 2012).
- ❑ Bireylerin psikolojik özellikleri hakkında teşhis yapılmak istendiđinde bu özellikleri ölçmek için geliştirilen ölçme araçlarının yapı geçerliklerinin yüksek olması beklenmelidir (Köse, 2012).

# Yapı Geçerliđi (Köse, 2012)

---

- Yapı geçerliđini ortaya koymada faktör analizi oldukça sık kullanılmaktadır. Faktör analizi, çođunlukla psikolojik ölçme aracı geliřtirmek veya ölçme aracının temel aldıđı varsayılan kuramsal yapıyı test etmek amacıyla kullanılan bir analiz türüdür.

# Yapı Geçerliđi (Köse, 2012)

---

- Yapı geçerliđi, ölçülmek istenen yapının farklı gruplarda incelenmesi ile de ortaya konabilir. Bu yöntem farklı-zıt gruplar çalışmaları (contrasted group studies) olarak adlandırılmaktadır.
- Örneđin, sınav kaygısı yüksek ve normal bireyler kaygısı yüksek olan bireyler testten yüksek puan, normal olanlar da düşük puanlar

# Ölçüt Dayanaklı Geçerlik

---

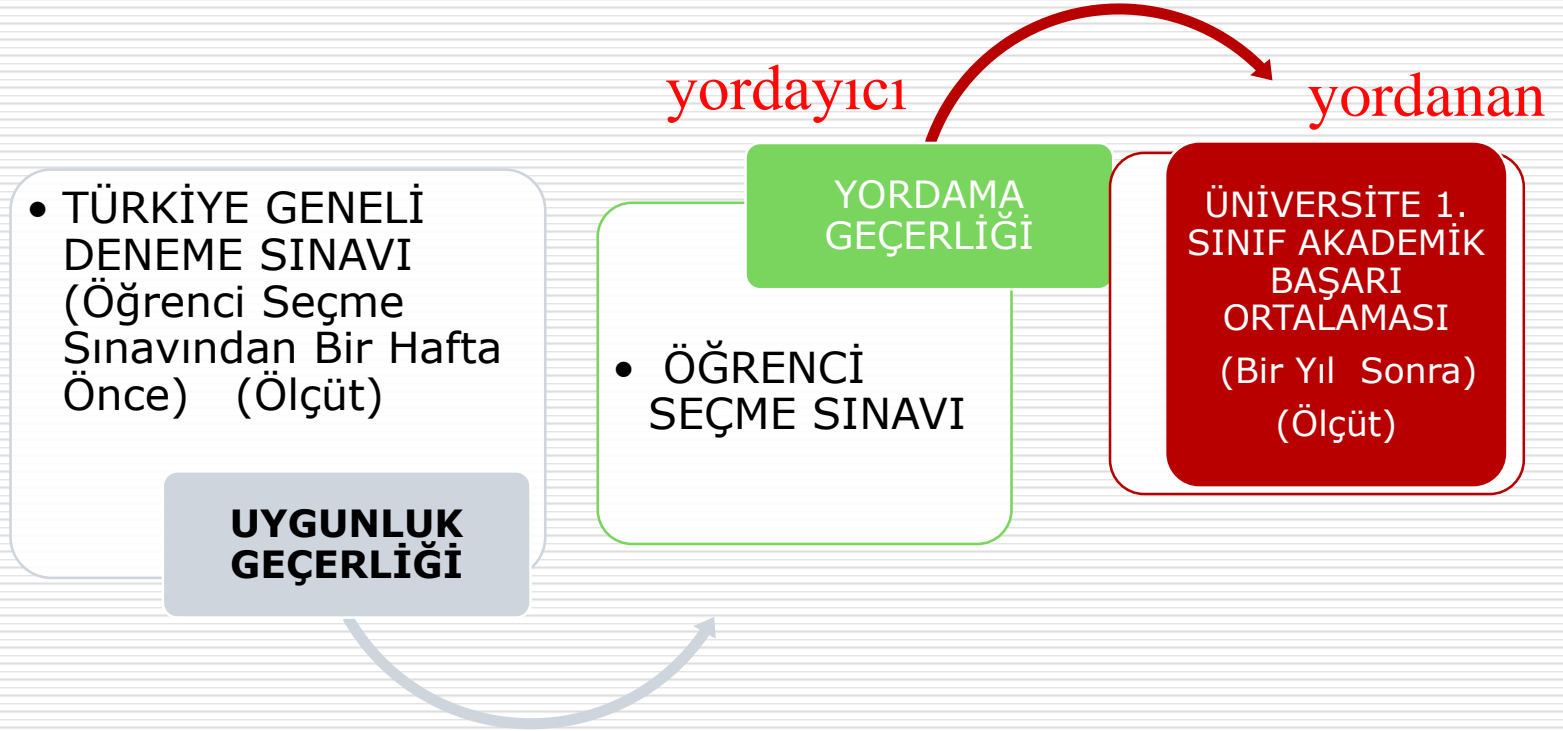
- Bir ölçüte dayalı geçerlik belirlenirken, ölçütten elde edilen puanlar, ölçme aracından elde edilen puanlarla aynı zamanda veya daha önceden elde edilmişse bu puanlar arasındaki korelasyona dayalı geçerliğe **uygunluk/zamandaş geçerlik** adı verilmektedir (Atılgan ve diğ, 2011; Baykul, 2000).

# Ölçüt Dayanaklı Geçerlik

---

- Ölçüt durumundaki puanlar, geçerliği belirlenecek ölçme aracından elde edilen puanlardan daha sonra elde edilmişse, bu puanlar arasındaki korelasyona dayalı geçerliğe **yordama geçerliği** adı verilmektedir (Atılgan ve diğ, 2011; Baykul, 2000).

# Ölçüt Dayanaklı Geçerlik (Köse, 2012)



Ölçüt puanlar, testin uygulanmasından sonra, ileriki bir tarihte elde edilmek zorundaysa, yordama geçerliğine başvurulur.

# GÜVENİRLİK

---

- Aynı bireyler üzerinde yapılan bir niteliğe ait ölçmelerin benzer şartlar altında tekrar edilebilirliğidir (Köse, 2012).
- Ölçme sonuçlarının tesadüfi hatalardan arınlık derecesi
- Ölçme aracının ölçülmek istenen özelliğe olan duyarlılık derecesi (Atılğan ve diğ., 2011).



# Güvenirlilik

---

- Güvenirlilik kavramı, geçerlik kavramında olduğu gibi ölçme araçları için değil, o ölçme araçlarından elde edilen puanlarla veya sonuçlarla ilgili bir kavramdır (Reynolds, Livingston ve Wilson, 2006; Akt. Köse, 2012).

# Güvenirlik Kestirimi

## Korelasyonel Yöntemler

## Puan Cinsinden Güvenirlik

### Tek Uygulamaya Dayanan Yöntemler

### Birden Çok Uygulama Dayanan Yöntemler

### Ölçmenin Standart Hatası (ÖSH)

1. İki Yarım Yöntemi
2. KR 20-21
3. Cronbach Alfa

1. Test tekrar Test
2. Eşdeğer (Paralel) Formlar Yöntemi

# Tek Uygulamaya Dayanan Yöntemler

---

## İki Yarıya Bölme Yöntemi

- Tesadüfi, Tek-çift, konu alanı vb.
- Tutarlılık
- İlişkinin düşük olması,
  - Testin gelişigüzel yanıtlanması
  - İki yarının farklı davranışları ölçmesinden kaynaklanabilir.

# Spearman Brown (Düzeltilme Formülü)

---

$r_{11}$ : İki alt teste ait güvenilirlik katsayısı

$r_x$ : Testin tümüne ait güvenilirlik katsayısı

- Testin içindeki alt test sayısı 2'den fazla ve  $k$  tane ise

$$r_x = \frac{k(r_{11})}{1 + (k - 1)r_{11}}$$

# Tek Uygulamaya Dayanan Yöntemler

## □ KR 20-21

- Testteki her maddenin aynı değişkeni ölçtüğü sayılıtısına dayanır
- Testteki tüm maddeler arasındaki tutarlılığı (Maddelerin birbirleriyle ve testin tümüyle) (İç tutarlılık)

$$KR_{20} = \frac{K}{K-1} \left[ 1 - \frac{\sum pq}{S_x^2} \right]$$

$K$  = Testin soru sayısı

$p$  = Madde güçlüğü

$q = 1 - p$

$S_x^2$  = Testin varyansı

# Tek Uygulamaya Dayanan Yöntemler

---

## Cronbach-Alfa

Genellikle ağırlıklı puanlama veya dereceleme yöntemiyle puanlama uygulandığında

İç tutarlılık

$$\alpha = \frac{K}{K-1} \left[ 1 - \frac{\sum s_j^2}{S_x^2} \right]$$

$K$  = Testin soru sayısı  
 $S^2_j$  =  $j$  maddesinin varyansı  
 $S^2_x$  = Testin varyansı

# Güvenirlik Kestirimi

## Korelasyonel Yöntemler

## Puan Cinsinden Güvenirlik

### Tek Uygulamaya Dayanan Yöntemler

### Birden Çok Uygulama Dayanan Yöntemler

### Ölçmenin Standart Hatası (ÖSH)

1. İki Yarım Yöntemi
2. KR 20-21
3. Cronbach Alfa

1. Test tekrar Test
2. Eşdeğer (Paralel) Formlar Yöntemi

Dr. Seher Yalçın

# Birden fazla Uygulamaya Dayanan Yöntemler

---

- **Test Tekrar Test Yöntemi** (Tekin, 1996)
- Kararlılık (Uzun süreli), grup içindeki sırasında tutarlılık
- Yüksek güvenilirlik katsayısı, testin kararlı olduğunu ve ölçme aracının uygulamadan kaynaklı tesadüfi hatalardan da arınık olduğunu gösterir.



# Birden fazla Uygulamaya Dayanan Yöntemler

---

- **Test Tekrar Test Yöntemi** (Tekin, 1996)
- Ölçülen nitelik sürekli değişkenlik gösteriyorsa, uygun değil
- Süre hatırlanmayacak kadar uzun, öğrenme olmayacağı kadar kısa olmalı

# Birden fazla Uygulamaya Dayanan Yöntemler

---

## Paralel Testler (Eşdeğer Formlar) Yöntemi

- Ölçülen davranışlar, soru sayısı, ortalama vb. eşit
- Tutarlılık (Kısa süreli)
- Yüksek güvenirlik katsayısı,
  - iki eşdeğer testten elde edilen puanların tutarlı olduğunu,
  - iki testin aynı davranışı ölçtüğünü gösterir.

# Güvenirliğin Puan Olarak Kestirimi: Ölçmenin Standart Hatası ve Testlerde Kullanımı (Köse, 2012)

---

- Güvenirlik, hatasızlığın ölçüsü ise hatanın ölçüsü de ölçmenin standart hatasıdır. Ölçmenin standart hatası;

$$\text{Standart Hata} = ss. \sqrt{1 - \text{Güvenirlik}}$$

# **Güvenirlilik Katsayısını Etkileyen Koşullar** (Baykul, 2000; Köse, 2012; Tekin, 1996)

---

- Puanlamanın Nesnelliği**
- Puanlayıcı Güvenirliliği**
- Test Edilen Grubun Değişkenliği**  
(Homejen- heterojen olması)
- Testin uzunluğu**
- Testin Güçlük Düzeyi**

# Karşılaştırma

---

- Geçerlik tüm hata türlerinden etkilenirken, güvenirlik sadece **TESADÜFİ HATA**lardan etkilenir.
- Güvenirlik 0 ile 1 aralığında değer alırken
- Geçerlik -1 ile +1 aralığında değer alır

# Karşılaştırma

---

- Bir ölçme aracının güvenilir olması geçerli olduğu anlamına gelmez.
- Bir test geçerli olmasa da güvenilir olabilir. Yani güvenilirlik geçerlik için gerekli fakat yeterli değildir.

# Kullanışlılık

---

- Ekonomiklik (zaman, emek, para)
- Uygulanabilirlik (formatı, anlaşılabilirliği)
- Puanlanabilirlik (güvenilir, kolay)