



6. EĞİTİMDE KULLANILAN ÖLÇME ARAÇLARININ NİTELİKLERİ

Kapsam
geçerliđi

Ölçüt
geçerliđi

Yapı
geçerliđi

Görünüş
geçerliđi

Türleri



**Ölçme
Aracının
Nitelikleri**



Güvenirliđi tahmin yöntemleri

Test-tekrar test

Testi yarılama

Paralel testler

KR-20 ve 21
formülleri

Cronbach
Alpha

Puanlayıcılar
arası tutarlılık

Güvenirlik; duyarlılık, kararlılık, tutarlılık ve objektiflik kavramlarıyla doğrudan bağlantılıdır.

A. ÖLÇME ARACININ GÜVENİRLİĞİ

Bir ölçme aracının; farklı zamanlarda aynı sonucu vermesi veya tesadüfi hatalardan arınık olma derecesi güvenilirliği ile ilgilidir.

Güvenirlik; bir ölçme aracıyla aynı koşullarda tekrarlanan ölçümlerde elde edilen ölçüm değerlerinin kararlılığının (tutarlılığının) bir göstergesidir.

Örneğin; metre ile bir uzunluğun ölçülmesinde, farklı zamanlarda aynı sonuçlar elde edilebiliyorsa, kullanılan ölçme aracının güvenilirliğinin yüksek olduğu söylenir.

Ölçme araçlarının **güvenirliğini** etkileyen dört temel özellik vardır:

Kararlılık, Tutarlılık, Duyarlılık ve Objektiflik

a) Kararlılık:

Bir nitelik aynı araçla birden fazla ölçüldüğünde, birbirinden farklı olmayan sonuçlar çıkarsa, ölçme aracı kararlıdır.

Ölçmeler arası zaman farkının fazla olduğu durumda ve ölçmenin aynı araçla yapılması durumunda benzer sonuçlar elde ediliyorsa kararlılıktan söz edilir.

b) Tutarlılık:

Tutarlılığı yüksek olan ölme araçları ile yapılan ölçümler ne kadar tekrarlanırsa tekrarlansın elde edilen değerler birbirine yakın olur.

Ölçmeler arası zaman farkının az olduğu ve ölçmenin aynı ya da farklı araçla yapıldığı durumlarda tutarlılıktan söz edilir.

Örneğin; bir kasa elmayı aynı terazide 5'er gün arayla tarttığımızda her seferinde 12 kg değerini elde ediyorsak, bu ölçme aracının tutarlı sonuçlar verdiğini söyleriz.

Hatırlatma: Bir testin kısa zamanda tekrarlı uygulaması benzer sonuçlar veriyorsa **tutarlıdır**.

Uzun zamanda benzer sonuçlar veriyorsa **hem tutarlı hem de kararlı** olur. Fakat zaman artarken testin **kararlılığı** ön plana çıkar.

c) Duyarlılık:

Bir ölçme aracı, ne kadar **hassas** ölçüm yapabiliyorsa duyarlılığı da o ölçüde artar.

Bir ölçme aracının duyarlı ölçüm yapması için ise birimleri arasındaki mesafenin küçük olması gerekir.

Örneğin; bir altın bilezik alırken kuyumcunun kilogramla tartım yapması mı, yoksa miligramla tartım yapması mı, daha doğrudur?
100 soruluk test 10 soruluk testten daha duyarlıdır.



d) Objektiflik

Ölçme işlemini **kim** yaparsa yapsın **aynı sonuca** ulaşmasıdır.

Örnek: Bir matematik sınavını okuyan farklı iki öğretmeninin yakın notlar vermesidir.

Güven Aralığı

- Ölçme araçlarının güvenirlik düzeyi, **güvenirlik analizi** ile belirlenir. Güvenirlik katsayısı (r) **“0” ile “+1”** arasında değişir.
- Yapılan ölçme işlemlerinde tesadüfî hata miktarı ne kadar az olursa güvenirlik düzeyi de o oranda yüksek olur.
- Eğer ölçme işleminde tesadüfî hata yapılmamış ise güvenirlik katsayısı **“1”**e yakın bir değer alacaktır.
- Fakat sosyal bilimlerde yapılan ölçümlerde, kontrol edilemeyen birçok değişken olduğundan, hatasız ölçüm yapmak neredeyse imkânsızdır.

Hatırlatma!

- ➡ Güvenirlik, ölçme aracıyla incelenen özelliğe ilişkin bir derece belirleme işlemi olduğundan “**güvenirlik var**” **veya** “**güvenirlik yok**” ifadeleri kullanılmaz.
- ➡ Ölçme araçlarının güvenilirlikleri için “**güvenirliği yüksek**” **veya** “**güvenirliği düşük**” ifadelerin kullanılması yerinde olacaktır.

Ölçme Araçlarının Güvenirliğini Belirleme Yöntemleri:

Ölçme araçlarının ne kadar güvenilir olduğunu belirlemek için birçok yöntem vardır. Bu yöntemler:

- 1. Test Tekrarı Yöntemi**
- 2. Eşdeğer Formlar Yöntemi**
- 3. Testi Yarılama Yöntemi**
- 4. KR-20 ve KR-21 Yöntemleri**
- 5. Cronbach Alpha Yöntemi**
- 6. Puanlayıcılar Arası Tutarlılık**

1) Test Tekrarı (Test Tekrar Test) Yöntemi

- Bir ölçme aracının, aynı gruba farklı zamanlarda tekrar uygulanması ile elde edilen güvenirlik katsayısıdır.
- Farklı uygulamalarda elde edilen değerlerin tutarlı olması, test tekrarı yöntemine göre güvenirliğin yüksek olduğu anlamına gelir.
- Aradan geçen **zaman** cevaplayıcıların ölçülen bilgilerin unutacakları kadar uzun, ilk uygulamada verdikleri cevabı hatırlayacakları kadar kısa olmamalıdır.
- Bu yöntemle bir ölçme aracının güvenirlik düzeyi belirlenirken, farklı uygulamalardan elde edilen puanlar arasındaki ilişki düzeyine bakılır. Test tekrar test yöntemi sonucunda hesaplanan katsayıya “**kararlılık katsayısı**” da denilmektedir.

2) Eşdeğer Formlar (Paralel Testler) Yöntemi:

- Bütün özellikleri aynı olan ve aynı amaca yönelik olarak hazırlanmış iki farklı ölçme aracının aynı veya denk bir gruba, aynı yada farklı zamanlarda uygulanması sonunda elde edilen puanların tutarlılığının incelendiği yöntemdir.
- Bu özdeş testin sonuçları birbirine ne kadar benzerse, eşdeğer formlar yöntemine göre güvenirliğinin yüksek olduğu söylenir.
- **Eşdeğer formlar;** Kapsam, ölçülen özellik, soru sayısı, soru türü, güçlük düzeyi, ayırt edicilik yönünden denk olmalıdır.
- Paralel testler yöntemiyle ölçme araçlarının güvenirliğini belirlemede elde edilen katsayıya **“eşdeğerlik katsayısı”** denir. Yani iki testin ne ölçüde eşdeğer olduğunu ifade eden bir değerdir. Bu değer de 0 ile 1 arasında değer alır. Bu değer 1’a yaklaştıkça güvenirlik yüksek, 0’a yaklaştıkça güvenirlik düşük olur.

3) Testi Yarılama Yöntemi:

- ➡ Bir ölçme aracının, bir gruba uygulanması sonunda elde edilen verilerin iki eşdeğer parçaya bölünmesi ve bu parçaların puanlarının karşılaştırılması ile yapılan güvenirlik belirleme işlemlerine “**testi yarılama yöntemi**” denir.
- ➡ Bu iki eşit parçadan elde edilen puanlar ne kadar tutarlı olur ise testi yarılama yöntemine göre ölçme aracının güvenirliği de o kadar yüksek olur.
- ➡ Testi yarılama yöntemiyle elde edilen katsayıya “**iç tutarlılık katsayısı**” da denilmektedir.
- ➡ Bu yöntemle hesaplanan güvenirlik katsayısı, genellikle diğer yöntemlere göre daha yüksek çıkmaktadır.
- ➡ Bilimsel nitelik taşımayan çalışmalarda, özellikle bir testi birden fazla uygulama imkânı olmadığında, kullanılması tavsiye edilmektedir.

4) KR-20 ve KR-21 (Kuder Richardson) Yöntemi:

- ➡ Bir ölçme aracı **kesinlikle doğru veya kesinlikle yanlış** cevaplar içeriyorsa yani **akademik başarıyı** ölçen bir test ise güvenirliğini belirlemek için en fazla KR-20 ve KR-21 yöntemleri tercih edilmektedir.
- ➡ **Kuder ve Richardson**'un 1937 yılında yaptıkları çalışma ile ortaya çıkardıkları bu yönteme göre, ölçme aracında bulunan her bir madde teker teker analiz edilmektedir.
- ➡ Özellikle her bir maddenin **güçlük düzeyinin** hesaplanması gereklidir.
- ➡ **Çoktan seçmeli soru** türlerinde kullanılan bir yöntemdir.

KR-20 formülünün kullanılması için,

- 1. Testte yer alan her bir maddenin aynı puan ağırlığına sahip olması.**
Örneğin; testte 20 soru varsa hepsi 5 puan olmalıdır.
- 2. Soruların güçlük düzeylerinin birbirinden farklı olması.**
Örneğin; bazı sorular zor iken, bazıları kolay, çoğunluğu ise orta güçlükte olmalıdır.
- 3. Düzeltme formülünün kullanılmamış olması.**
Örneğin; dört yanlış cevabın bir doğru cevabı götürdüğü testlerde KR-20 yerine KR-21 kullanılır.
- 4. Doğru cevapların 1, yanlış cevapların 0 şeklinde puanlanmış olması.**
Örneğin; öğrencilerin doğru cevaplarından 1, yanlış cevaplarından ise 0 aldığı kabul edilmelidir.

KR-21 formülünün kullanılması için,

1. Testte yer alan her bir maddenin farklı puan ağırlığına sahip olması.

Örneğin; bir testteki bazı sorular 5, bazıları 4 puan değerinde ise KR-20 yerine KR-21 tercih edilmelidir.

2. Soruların güçlük düzeyinin birbirine yakın olması.

Örneğin; bir testteki soruların güçlük düzeyleri 0.45 ile 0.55 veya 0.30 ile 0.35 arasında değişiyorsa KR-21 formülü tercih edilmelidir.

3. Düzeltme formülünün kullanılmış olması.

Örneğin; öğretmen yanlış cevapların doğru cevapları götürdüğünü belirtmiş ise KR-21 formülü kullanılmalıdır.

KR-20 ve KR-21 Hangi Durumlarda Kullanılır?

KR-20 ve KR-21 yöntemleri **çoktan seçmeli, doğru-yanlış ve kesin doğru cevabı olan açık uçlu maddelerden** meydana gelen testlerin güvenirlik düzeyini belirlemede kullanılmaktadır.

Bu iki yöntemle elde edilen güvenirlik katsayıları, ölçme aracının iç tutarlılığı ile ilgili olduğundan **“iç tutarlılık katsayısı”** olarak da bilinmektedir.

5) Cronbach Alpha Yöntemi:

- Bu güvenirlik katsayısı, yanıtları iki kategorili olmayan **derecelendirilmiş** (*Tamamen katılıyorum-Kesinlikle katılmıyorum ve Her zaman-Hiçbir zaman aralıkları gibi*) ölçeklerin iç tutarlılık katsayısını hesaplamada kullanılan bir yöntemdir.
- Bu yöntem, KR-20 ve KR-21'in daha genel halidir. Bu yöntemin temel mantığı, ölçme aracındaki maddelerin bütün olası yarılarının karşılaştırılmasına dayanmaktadır.
- Bu yöntemde, ölçme aracı yalnızca bir defa kullanıldığı ve analizlerde bu maddelerin birbiriyle ilişkisi incelendiği için elde edilen değer **“iç tutarlılık katsayısı”** olarak kabul edilir.
- Eğer Alpha değeri yüksek çıkarsa alt testler birbiriyle uyumlu ve tutarlı olacaktır. Dolayısıyla testin iç tutarlılığı yüksek ve alt testleri homojen olacaktır.

6) Puanlayıcılar Arası Tutarlılık:

- ➡ **Bir ölçme işleminde güvenirliği arttırmanın bir diğer yolu ise elde edilen verilerin iki puanlayıcı tarafından değerlendirilmesi ve uyuşumuna bakılmasıdır.**

Ölçme Aracının Güvenirliğini Etkileyen Faktörler:

- Test türü
- Testte yer alan soru sayısı
- Testin güçlük düzeyi
- Testi cevaplama süresi
- Testin puanlanması
- Testin uygulandığı grubun uygunluğu
- Test anlaşılabilirliği
- Testin uygulandığı ortam

Güvenirliği Artırmanın Yolları

- Soru sayısını arttırmak
- Soruların açık, anlaşılır ve cevaplanabilir olması
- Öğrencilerin güdülenmesi
- Sınav süresinin yeterli olması
- Testin orta güçlükte olması
- Testin objektif ve dikkatli puanlanması güvenilirliği artırır
- Duyarlılığı yüksek ölçme aracının kullanılması
- Ölçme sonuçlarının hatasız kaydedilmesi
- Uygun bir test yönergesinin hazırlanması
- Testin uygulanacağı öğrenci sayısını arttırmak

Ölçmecinin Güvenirliği

- ➡ Ölçme ve değerlendirme sürecinde, kullanılan ölçme aracının güvenilirlik düzeyinin yüksek olmasının yanında ölçmeyi yapan kişinin veya araştırmacının güvenilirlik düzeyi de oldukça önemlidir.
- ➡ Güvenilir ölçme yapan kişiler, olaylara “*nesnel*” veya “*objektif*” bakabilen kişilerdir.
- ➡ Ölçme yapan kişi; ölçme aracından elde ettiği verileri **belli bir zaman aralığında iki kez puanlar** ve bu puanlar arasındaki **tutarlıklara** bakarsa ölçme işleminin **puanlama güvenilirliği artar**.

Bir okulda, fen bilgisi proje ödevi verilmiştir. Ürünlerin puanlanmasının sınıfta derse giren öğretmen yerine, diğer sınıfın öğretmeni tarafından yapılması sağlanmıştır. Projelere kod numaraları verilmiş ve öğrenci isimleri, projeleri değerlendiren öğretmenden gizlenmiştir. Ayrıca değerlendirme için (kapsamlı, detaylı ve açık) bir puanlama rehberi oluşturulmuş, öğretmenin bu rehber göre puan vermesi istenmiştir.

Yukarıda anlatılan işlemler, kullanılan ölçme aracının hangi özelliğini iyileştirmek amacıyla gerçekleştirilmiştir? (KPSS: 2009)

- ☒ A) Objektiflik
- ☐ B) İç tutarlık
- ☐ C) Kullanışlılık
- ☐ D) Görünüş geçerliği
- ☐ E) Yapı geçerliği

Bir öğretmenin cevap kağıtlarını aynı dersi veren diğer öğretmenlere de puanlattırmasının gerekçesi aşağıdakilerden hangisi olabilir?
(KPSS:2003)

- A) Değerlendirme için ölçüt belirleme
- B) İyi bir cevap anahtarı oluşturma
- ☒ C) Puanlama güvenirliğini artırma
- D) Puanların geçerliğini sağlama
- E) Puanlamayı kolaylaştırma

Aşağıda özellikleri verilen ölçme araçlarından hangisinin diğerlerine göre daha güvenilir ölçme sonuçları vermesi beklenir? (KPSS:2005)


- A) Kısa cevaplı 10 soru
- B) Boşluk tamamlama türü 25 soru
- C) Doğru-yanlış turu 50 soru
- D) Üçer seçenekli 25 soru
- ☒ E) Beşer seçenekli 50 soru

Aşağıdakilerden hangisi bir başarı testinden elde edilen puanların güvenirliğinin bir tanımı olabilir?
(KPSS: 2003)

- A) Soruların konulara uygun şekilde dağıtılmış olması
- ☒ B) Test puanlarının tesadüfi hatalardan arınık olma derecesi
- C) Soruları, sadece yoklanan davranışı öğrenmiş olanların cevaplayabilmesi
- D) Test puanlarıyla, öğrencinin gelecekteki başarısının yordanabilmesi
- E) Soruların, testin amacına hizmet etme derecesi


Aşağıdakilerden hangisi, testin güvenilirliğini düşürücü bir etken değildir? (KPSS: 2001)

- A) Cevaplama süresinin çok kısa olması
- B) Sorularda yoruma açık ifadeler kullanılması
- C) Puanların bir puanlayıcıdan diğerine değişebilmesi
- ☒ D) Puanlama şeklinin öğrenciye önceden bildirilmesi
- E) Soruların çok zor olması

- 
- ➔ Bir öğretmen sınav sorularına verilen yanıtları dört beş gün arayla iki kez puanlamış; bu iki puanlamadan elde ettiği sonuçları karşılaştırarak aralarındaki tutarlılığa bakmıştır.


Öğretmen bu çabasıyla aşağıdakilerden hangisini belirlemeye çalışmaktadır? (KPSS: 2001)

- ☒ A) Puanlama güvenirliğini
- ☐ B) Soruların doğruluğunu
- ☐ C) Soruların güçlük derecesini
- ☐ D) Puanlama anahtarının tutarlılığını
- ☐ E) Sınavın geçerliğini

- 
- Bir öğretmen İngilizce kelime bilgisini yoklayan bir test hazırlamıştır. Bu testi birkaç hafta arayla öğrencilere iki kez uygulamış ve bu iki uygulamadan elde edilen puanlar arasındaki korelasyonu 0.90 olarak bulmuştur.

Bulunan bu değer testin en çok hangi özelliği hakkında bilgi verir?
(KPSS-2004)

- A) KR-20 güvenirliği
- ☒ B) Test-tekrar-test güvenirliği
- C) Kullanışlılığı
- D) Yapı geçerliği
- E) Kapsam geçerliği

- 
- Bir öğretmen, bir testin iki paralel formunu 2 hafta arayla öğrencilerine uygulamıştır. Formların uygulanması veya puanlanması esnasında karşılaşılabilecek olası durumlar aşağıda verilmiştir:

I. Öğrencilerin ikinci uygulamada yorgun olmaları

II. Öğretmenin her iki uygulamada da toplam puanları hesaplarken işlem hataları yapması


III. Toplam puanları düşük bulan öğretmenin, puanları her iki uygulamada da ortalaması 65, standart sapması 10 olacak şekilde artırması

Bu durumlardan hangileri, paralel formlar yöntemi ile hesaplanan güvenirlik katsayısını düşürür? (KPSS: 2012)


- A) Yalnız I B) Yalnız II **C) I ve II** D) I ve III E) II ve III

B. ÖLÇME ARACININ GEÇERLİĞİ



- Bir ölçme aracının, ölçülmek istenen özelliği başka özelliklerle karıştırmadan ölçebilme yeteneğine geçerlik denir.
- Başka bir tanıma göre ölçme aracının geçerliği, ölçülmek istenen nitelikleri ölçmeleri ve amaca uygun ölçme yapmalarıdır.
- Geçerlik; bir ölçme aracının, amacı ne kadar gerçekleştirdiği ile ilgili bir derecedir.
- **ÖRNEK:** Uzunluk ölçmek için geliştirilmiş olan metre, kişilerin boylarını ölçme amacına hizmet eder. Bu demektir ki bir ölçme aracı olan metre, uzunluk ölçmede geçerlidir fakat ağırlık ölçmede geçerli değildir.



➡ Bir aracın geçerli olması için, onun sadece ölçmek için düzenlendiği özelliği ölçmesi yani onu başka özelliklerle karıştırmadan ölçmesi de gerekir.



ÖRNEK: Bir öğretmen yaptığı bir Biyoloji sınavında yüksek bir okuma yeteneği ve kelime bilgisi gerektirecek sorular sormuş ve puanlama yaparken de yazı güzelliği, sayfa düzeni vb. özelliklere de fazla puan vermiştir. Böyle bir test ölçmeyi amaçladığı özelliği başka özellikleri de karıştırarak ölçtüğü için geçerliği düşüktür.

- 
- 
- ➔ Bir ölçme aracının “**geçerliği var**” **veya** “**geçerliği yok**” gibi ifadelerle tanımlanması doğru değildir.
 - ➔ Geçerlik bir derece olduğundan, güvenirlikte olduğu gibi, “**geçerliği yüksek**” **veya** “**geçerliği düşük**” ifadelerinin kullanılması daha uygundur.
 - ➔ Geçerli ölçme araçları iki temel özellik taşır.
Bunlar: **Uygunluk** ve **Güvenirlik** tir.

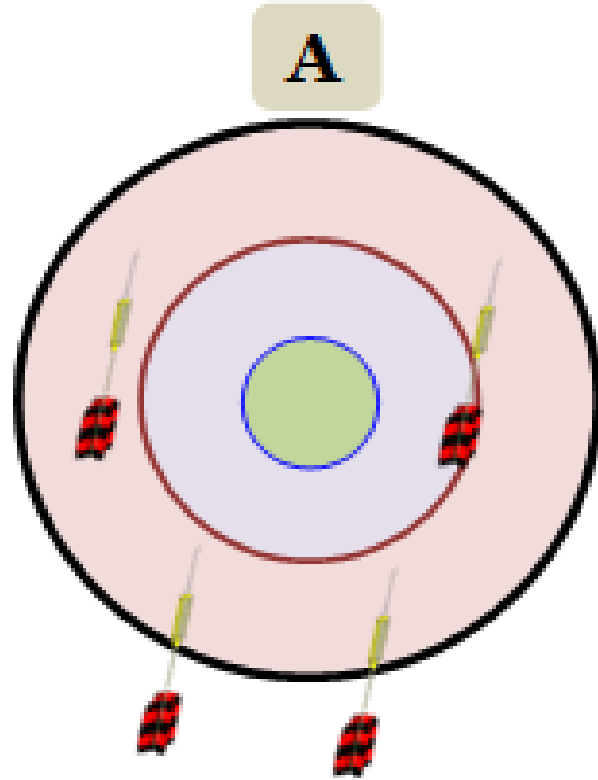
Uygunluk:

- Bir ölçme aracının ölçülecek **özelliği tam olarak ortaya çıkarabilme yeteneği** olarak tanımlanmaktadır.
- Yani hangi özellik ölçülmek isteniyorsa, ölçme işleminde kullanılacak ölçme aracı bu **özelliği ortaya çıkarıcı** nitelikte olmalıdır.
- Ölçme araçlarının uygunluğu, istatistiksel analizlerle belirlenemez. Bir ölçme aracının uygunluk özelliğini taşımasında, hazırlayan kişilerin **alan uzmanı** olmasının büyük rolü vardır.

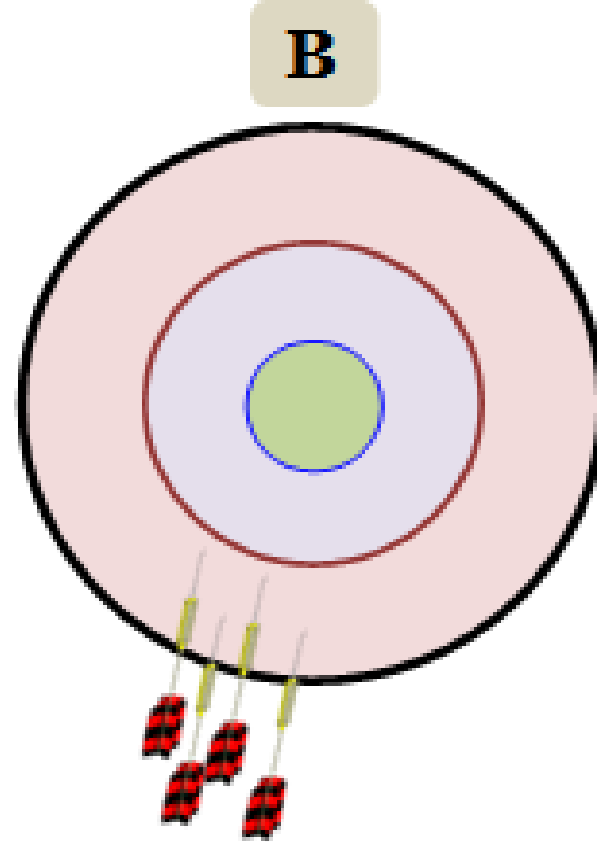
Güvenirlilik:

- Ölçme aracının farklı uygulamalarda tutarlı sonuçlar vermesidir.
- Bir ölçme aracının **geçerli** olabilmesi için **mutlaka güvenilir** olması gerekir.
- Yani güvenirlik, geçerlik için ön şarttır. Güvenirliği düşük olan ölçme araçlarının geçerliği de kesinlikle düşük olur.
- Fakat güvenirliği yüksek olan her ölçme aracının geçerliğinin yüksek olmayacağı da bilinmelidir.

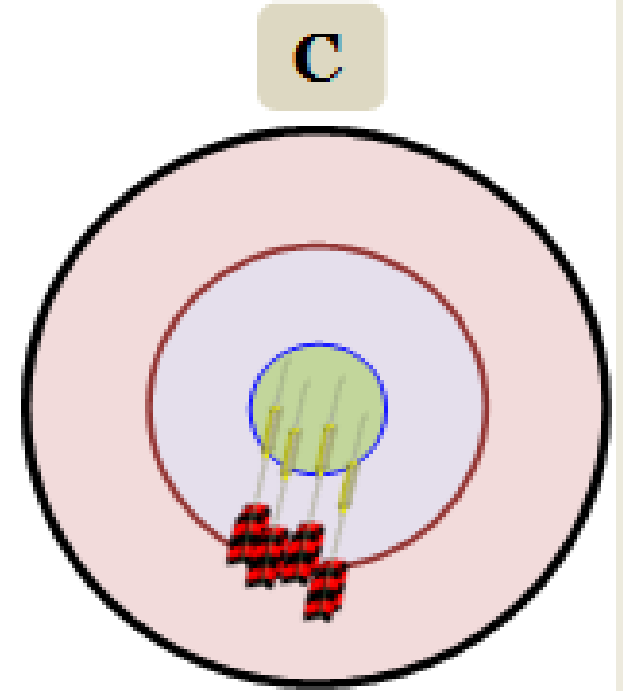
Geçerlik-Güvenirlilik İlişkisi



*A-Geçerliği düşük,
güvenirliliği düşük*



*B-Geçerliği düşük,
güvenirliliği yüksek*



*C-Geçerliği ve
güvenirliliği yüksek*

Ölçme Araçlarında Aranan Geçerlik Türleri

1. **Kapsam Geçerliği**
2. **Ölçüte Dayalı Geçerlik**
(Ölçüt Geçerliği, Yordama Geçerliği)
3. **Yapı Geçerliği**
4. **Görünüş Geçerliği**

1. Kapsam geçerliği:

- ➡ Ölçme aracının, ölçülmek istenen konu, ünite, kazanımları kapsama derecesidir.
- ➡ Diğer bir tanıma göre; ***kapsam geçerliği***, bir bütün olarak ölçeğin ve ölçekteki her bir maddenin amaca ne derece hizmet ettiği ile ilgili özelliğidir.


Kapsam geçerliğini belirleme yöntemleri

İstatistiksel Yol: Ölçme aracı ile elde edilen veriler, aynı konuda daha önceden hazırlanmış ve kapsam geçerliği yüksek olan başka bir ölçme aracının sonuçlarıyla karşılaştırılır.

- Bu testlerin sonuçları birbirine ne kadar yakın ise yeni hazırlanan ölçme aracının kapsam geçerliğinin de o ölçüde yüksek olduğu belirlenir.

Mantıksal Yol: Bir ölçme aracının geçerliğini belirlemede en fazla tercih edilen yöntemdir.

- ➔ Bir testteki soruların kapsamı ne ölçüde örneklendirdiğini belirlemenin en etkili yolu, test maddelerini **belirtke tablosu** ile karşılaştırmaktır.
- ➔ Bunun için testteki her bir maddenin, belirtke tablosundaki dağılımları incelenir ve kapsama uygunluğu hakkında bir karara varılır.
- ➔ Bu yöntemin kullanılması için, ilgili alanda deneyim sahibi olmak gerekir.
- ➔ **Uzman kişilerin görüşünün alınması önemlidir.**



Bir arařtırmacı oklu zekâ alanlarını ölçmek amacıyla geliřtirdiđi ok boyutlu testi, dil bilimi, matematik, resim, heykel, müzik, beden eđitimi, iletiřim, psikoloji bölümlerindeki öđrencilere uygulamıřtır. Uygulama sonucunda farklı bölümlerdeki öđrencilerin kendi alanıyla ilgili sorulardan yüksek, diđer alanlardaki sorulardan ise daha düşük puan aldıklarını saptamıřtır.

Buna göre, ölçme aracı için ařađıdakilerden hangisi söylenebilir? (KPSS: 2009)


- A) Amaca hizmet etme derecesi düşüktür.
- B) Hata düzeyi düşüktür.
- ☒ C) Geçerliđi yüksektir.
- D) Kullanıřlılıđı yüksektir.
- E) Güvenirliđi düşüktür.

Bir başarı testinin “**kapsam geçerliğinin yüksek olduğu**” ifadesinden aşağıdaki sonuçların hangisi çıkarılabilir?
(KPSS: 2003)

- A) Tutarlı ve kararlı sonuçlar veren bir testten yararlanıldığı
- B) Testteki soruların pek çoğunu öğrencilerin doğru olarak yanıtlayabildikleri
- ☒ C) Testteki soruların konu alanını iyi bir şekilde örneklediği
- D) Testteki soruların güçlük düzeylerinin, gerçekleşmesi beklenen öğrenme düzeylerine uygun olduğu
- E) Öğrencilerin, testteki soruları yanıtlamaya çalıştıkları

Aşağıdakilerden hangisi sınavın geçerliğini artırıcı etkenlerden biridir? (KPSS:2001)

- A) Yanıtların puanlanmasında yanlı davranılması
- B) Sınav sorularının güçlük indeksi
- C) Soruların tek bir üniteden hazırlanması
- D) Çeldiricilerin uzun cümlelerden oluşması
- E) Sınavda her hedef-davranışla ilgili sorular sorulması**



Bir matematik öğretmeni derste problem çözme becerilerine yönelik etkinliklere ağırlık verdiği halde, uyguladığı sınavı daha çok dört işlem becerisini ölçen sorulardan oluşturmuştur.

Öğretmen bu yaklaşımla, aşağıdakilerden hangisini tehlikeye atmıştır? (KPSS: 2001)

- A) Test-tekrar test güvenirliğini
- B) Sınavın kullanışlılığını
- ☒ C) Testin kapsam geçerliğini
- D) Soruların ayırt etme gücünü
- E) Puanlayıcı güvenirliğini

2) Yordama Geçerliği:

- Yordama, davranışların önceden tahmin edilebilmesi ile ilgili bir özelliktir.
- Diğer bir ifadeyle **yordama**; istatistiksel teknikler kullanılarak ve bilinen özelliklerden yararlanılarak bilinmeyen özellikler hakkında yapılan geleceğe yönelik tahminlerde bulunma işlemidir.

Örneğin; ÖSS ile yerleşilen bölümdeki akademik başarı arasındaki ilişki.

Bir firmaya eleman almak amacıyla yapılan bir sınav için, **“Oldukça geçerli bir sınav yapıldı”** diyen bir uzman sınavın hangi özelliğini ifade etmektedir?
(KPSS: 2006)

- A) Zor olduğu
- B) Adayların yeteneklerine uygun olduğunu
- C) Objektif olduğunu
- D) Kolay olduğunu
- ☒ E) İşteki başarıyı yansıtacak yapıda olduğunu

Aşağıdakilerden hangisinde, yordama geçerliğinin belirlenmesine çalışılmaktadır? (KPSS: 2003)

- A) Hazırlanan soruların, belirli bir alanda öğrenilenleri temsil edip etmediğinin incelenmesi
- ☒ B) Ölçme aracının öğretiminde bir üst düzeyde başarılı olacak öğrencileri seçme özelliğinin incelenmesi
- C) Bir grup öğrencinin iki derste başarılarının karşılaştırılması
- D) Bir ölçme aracının, birkaç hafta arayla aynı gruba tekrar uygulanması ve elde edilen puanlar arasındaki ilişkinin incelenmesi
- E) Hazırlanan bir testin farklı düzeylerdeki öğrenci gruplarına uygulanması ve bu grupların başarılarının karşılaştırılması

Aşağıdakilerden hangisinde yordama geçerliliği vardır? (KPSS:2004)

- A. Belirli alanda hazırlanan soruların öğrenci seviyesine uygun olduğunun belirlenmesi
- B. İki arada yapılan sınavın birbiriyle ilişkisi
- ☒ C. Yapılan bir sınavın, bir üst eğitime geçebilecek öğrencileri seçebilme düzeyinin incelenmesi
- D. Sınav sonunda bir notun verilmesi
- E. Elde edilen ortalamanın diğer derslerle karşılaştırılması

Pazarlama elemanları seçiminde kullanılmak üzere bir test geliştirilmiş ve işe alımlarda bu test uygulanmıştır. Daha sonra, işe alınan elemanlarının ilk aydaki başarılı satış sayıları ile işe girişte aldıkları test puanları arasındaki uyum incelenmiştir.

Yukarıda belirtilen inceleme, uygulanan testin hangi özelliği hakkında bilgi verir? (KPSS: 2007)

- ☒ A) Yordama geçerliği
- ☐ B) Yapı geçerliği
- ☐ C) Kapsam geçerliği
- ☐ D) İç tutarlığı
- ☐ E) Puanlayıcılar arası güvenirliği

3) Yapı Geçerliği:


- Yapı, birbiriyle ilgili olduğu düşünülen belli öğelerin veya bu öğeler arasındaki ilişkilerin oluşturduğu bir bütündür.
- Birbiriyle ilişkili olan bu öğeler, incelenecek özelliğin yapısını meydana getirir.
- Bir ölçme aracının yapısı ise, testte yer alan maddelerin birleştirilmesi ile oluşur.
- Bu yönüyle, bir testte yer alan maddeler testin amacına ne kadar uygun olursa, testin yapı geçerliği de o ölçüde yüksek olur.

... Yapı Geçerliği:

- ➔ Bir ölçme aracının yapı geçerliğini belirlemek için faktör analizi yapmak gerekir.
- ➔ Bu analiz, bir ölçme aracının kaç tane alt özelliği ölçtüğünü gösteren istatistiksel bir işlemdir.
- ➔ **Örneğin;** zeka düzeyini ölçen bir testimiz olsun. Bu testte yer alan 47. sorunun yapı geçerliği düşük çıkmış olsun. Bunun anlamı yapı bakımından 47. soru zeka ölçemiyor, uygun değil demektir. Bu anlamda testin yapı geçerliği, temelde test maddelerine verilen cevaplar arasındaki ilişkilerin analizine dayanır. Bir testte yer alan her madde, testin bütünü ile ilişkili olmalıdır.

4) Görünüş Geçerliği

- Bu geçerlik türü, ölçme aracının neyi ölçtüğünü değil de neyi ölçer görüldüğünün belirlenmesi işlemidir.
- Görünüş geçerliği olan testler, yalnızca kapsadığı özellikleri ölçer, diğer özellikleri işleme katmaz.
- **Örneğin**; bir yaprak testin ya da deneme sınavının üzerinde ÖSS yaprak testi, tarama sınavı yazıp içindeki sorular tamamen ALES soruları ise görünüş geçerliği yok ya da çok düşük demektir.



“Test puanının yüksek olması kazandırılmak istenilen bilgi ve beceriler açısından bireyin ileri düzeyde olduğunu gösterir.”

Bu varsayımı kanıtlamaya yönelik çalışmalar testin hangi niteliği hakkında bilgi sağlar?
(KPSS: 2003)

- A) Kullanışlılığı
- B) Güvenirliği
- ☒ C) Geçerliği
- D) Objektifliği
- E) Tutarlılığı

**Aşağıdakilerden hangisi yabancı dil muafiyet
sınavında kullanılan test için bir geçerlik kanıtı
olabilir? (KPSS: 2005)**

- A) Testin hatasız puanlanabilmesi
- B) Testin çoktan seçmeli maddelerden oluşması
- C) Testten alınan puanların ranjının geniş olması
- ☒ D) Yabancı dili iyi bilenlerin testten yüksek puan alması
- E) Testte farklı yapılarda soruların bulunması

Ölçme Aracının Geçerliğini Etkileyen Faktörler

1. Ölçme aracının güvenirliği üzerinde etkili olan özellikler geçerlik üzerinde de etkilidir.
2. Ölçme aracında yer alan soruların sayısının arttırılması geçerliği yükseltir.
3. Testte yer alan soruların güçlük düzeyleri normal dağılım eğrisine benzemelidir.
4. Ölçme aracında yer alan maddelerin düzeyi, öğrencilerin seviyesine ne kadar uygun olursa geçerlik düzeyi de o ölçüde artar.
5. Testte yer alan soruların, çalışan ve çalışmayan öğrencileri ayırt edebilir olması, geçerliğin yükselmesini sağlayacaktır.

... Ölçme Aracının Geçerliğini Etkileyen Faktörler

6. Testi cevaplamak için verilen sürenin fazla veya az olmaması geçerliği yükseltir.
7. Testte yer alan maddelerin kolaydan zora, basitten karmaşığa doğru sıralanması geçerliği olumlu yönde etkiler.
8. Testin uygulaması sırasında kopya çekilmesi geçerliği düşürür.
9. Testin puanlanmasında hata yapılmaması geçerliğin yükselmesini sağlayacaktır.
10. Testte yer alan soruların doğru cevaplarının sistematik olarak sıralanmaması geçerliği yükseltir.
11. Ölçme aracında, dersin veya konunun her hedef davranış ile ilgili sorulara yer verilmesi geçerliği yükseltir.

GEÇERLİK VE GÜVENİRLİK ARASINDAKİ İLİŞKİ

1. Bir ölçme işleminde güvenirlilik “nasıl ölçülecek?”, geçerlik ise “ne ölçülecek?” sorusuna odaklanır. Yani ölçme aracının geçerliği, testte yer alan maddelerin neyi ölçtüğü ile ilgiliyken, güvenirliliği bu testin nasıl uygulanacağı ile ilgilidir.
2. Ölçme aracının geçerlik düzeyi; bütün hatalardan (sabit, sistematik ve tesadüfî hatalar) olumsuz etkilenirken, güvenirlilik düzeyi yalnızca tesadüfî hatalardan olumsuz etkilenir.
3. Güvenirlilik düzeyi düşük olan ölçme araçlarının geçerliği de düşük olur. Yani güvenirlilik, geçerlik için en önemli ön şarttır.

... Geçerlik ve Güvenirlik Arasındaki İlişki

4. Geçerliği yüksek olan bir testin güvenirliği de mutlaka yüksek olur. Çünkü geçerliğin yüksek olması için, ölçme sürecine hiçbir hata karışmamış olması gerekir.
5. Bir ölçme aracının güvenirlik düzeyi yüksek olsa bile geçerlik düzeyi düşük olabilir.
6. **Geçerlik ve güvenirlik düzeyi** düşük olan bir ölçme aracının geçerliğini yükseltmek için öncelikle güvenirlik düzeyini yükseltmek gerekir. Çünkü bir ölçme aracının geçerlik düzeyi, ne kadar çaba gösterilirse gösterilsin, güvenirlik katsayısının karesinden (r^2) daha fazla yükseltilemez.

Aşağıdakilerden hangisi “anlatım becerisini” ölçen bir yazılı sınavdan elde edilen puanların güvenilirliğini ve geçerliğini artırmaz? (KPSS:2002)

- A. Objektif tipte olmayan sorular için ayrıntılı puanlama anahtarının hazırlanması
- ☒ B. Öğrencinin sadece cevaplarının değil, aynı zamanda sınıftaki genel durumunun da dikkate alınarak puan verilmesi
- C. Sınav kağıtlarının birden fazla puanlayıcı tarafından puanlanarak puanların ortalamasının alınması
- D. Cevapların aynı puanlayıcı tarafından, sıraları değiştirilerek ikinci kez puanlaması ve puanlar arasında fark varsa bir daha puanlama yapılması
- E. Çabuk cevaplanabilen türden, çok sayıda soru sorulması

Bir “testin gvenirliđi yksek olduđu halde geerliđinin dřk olması durumunu ařađıdakilerden hangisi aıklar? (KPSS:2002)

- A. Testin hazırlanması g ve zaman alıcı olduđu halde puanlaması olay olmaktadır.
- ☒ B. Soru sayısı ve soruların glk dzeyi uygun olduđu halde programdaki tm konuları kapsamamaktadır.
- C. Test programındaki konuların tmn kapsadıđı halde bilen đrenciyle bilmeyeni ayıramamaktadır.
- D. Test puanlarının kararlılıđı yksek olmasına rađmen, sınav sırasında kopya ekilmesi engellenememiřtir.
- E. Testteki maddelerin glk derecesi yksek olmasına rađmen puanlamaya hata karıřmamaktadır.

Başarıyı ölçmek için geliştirilmiş beş testle ilgili olarak tablodaki bilgiler elde edilmiştir.

Test	Güvenirlik katsayısı	Yordama geçerliği	Kapsam geçerliği	Ortalama güçlük
1	0,72	0,45	Orta	0,70
2	0,65	0,29	Yüksek	0,35
3	0,83	0,64	Yüksek	0,58
4	0,92	0,18	Düşük	0,40
5	0,61	-0,55	Düşük	0,50

Sınıfınızdaki öğrenciler arasında bu dersteki başarıları yüksek olanları seçmek isteyen bir öğretmenin hangi testi kullanması en uygun olur? (KPSS: 2007)

- A)1 B) 2 **C)3** D) 4 E) 5



Test	Güvenirlik katsayısı	Yordama geçerliği	Kapsam geçerliği	Ortalama güçlük
1	0,72	0,45	Orta	0,70
2	0,65	0,29	Yüksek	0,35
3	0,83	0,64	Yüksek	0,58
4	0,92	0,18	Düşük	0,40
5	0,61	-0,55	Düşük	0,50

“Bir testin güvenirliğinin yüksek olması, geçerli bir test elde etmek için yeterli değildir” yargısını desteklemek için hangi test örnek gösterilebilir?

- A) 1 B) 2 C) 3 **D) 4** E) 5



Test	Güvenirlik katsayısı	Yordama geçerliği	Kapsam geçerliği	Ortalama güçlük
1	0,72	0,45	Orta	0,70
2	0,65	0,29	Yüksek	0,35
3	0,83	0,64	Yüksek	0,58
4	0,92	0,18	Düşük	0,40
5	0,61	-0,55	Düşük	0,50

Diğerlerinden daha kolay olduğu halde kabul edilebilir düzeyde güvenilir ve geçerli olan test hangisidir?

- A) 1 B) 2 C) 3 D) 4 E) 5**

C. ÖLÇME ARACININ KULLANIŞLILIĞI

- Kullanışlılık, bir ölçme aracında aranan üçüncü temel özelliktir.
- Kullanışlılık; bir ölçme aracının geliştirilmesinin, çoğaltılmasının, uygulanmasının, cevaplanmasının ve puanlanmasının kolay olmasıdır.
- Ayrıca bir testin **maliyeti** ne kadar düşük olursa, kullanışlılığı o ölçüde yüksek olur.

Ölçme Aracının Kullanışlılığı İle İlgili Genel Özellikler:

- a) **Hazırlanması:** Öğretmenin ölçme aracını hazırlarken zorlanmaması ve kısa sürede tamamlaması,
- b) **Uygulanması:** Öğretmenin ölçme aracını uygularken güçlük çekmemesi ve kısa sürede uygulaması,
- c) **Cevaplanması:** Öğrencilerin ölçme aracını cevaplarken zorluk çekmemeleri ve kısa sürede cevaplamaları,
- d) **Puanlanması:** Öğretmenin ölçme aracını puanlarken zorlanmaması ve kısa sürede puanlaması,
- e) **Maliyetin düşüklüğü:** Ölçme aracının okula maddi yönden yük getirmemesidir.