

ÖLÇME VE DEĞERLENDİRME (1)

ÖLÇME VE DEĞERLENDİRMENİN ÖNEMİ

Girdi (Öğretmen, öğrenci, araç-gereç, bina vb.)

Süreç (Öğretim süreci, öğretim yöntemleri)

Çıktı (İstendik davranışlar, kazanımlar)

Kontrol (Değerlendirme)

-Eğitim sürecinin etkililiğini ortaya koymak için öğrenci davranışlarına bakılır. Bu da ölçme ve değerlendirmeye yapılır.

-Ölçme ve değerlendirme etkinlikleri öğrenme amaçlarının ne olduğunun tanımlanmasıyla başlar ve bu amaçlara ne derecede ulaşıldığının ortaya konmasıyla devam eder.

-Ölçme ve değerlendirme öğrenci başarısıyla birlikte okulun başarısı, öğretmenin başarısı, programın başarısı hatta ülkenin eğitim alanındaki başarısı hakkında bilgi verir.

-Ölçülmeyen hiçbir şey geliştirilemez, düzeltilemez.

Ölçme ve Değerlendirmenin İşlevi

-Öğrencilerin neyi, ne kadar öğrendiği hakkında bilgi verir.

-Eksik öğrenmeler hakkında bilgi vererek önlem alınmasına imkân sağlar.

-Öğretmenlerin başarısı hakkında bilgi verir.

-Öğretim yönteminin etkililiği hakkında bilgi verir.

-Okul başarısı hakkında bilgi verir.

-Öğrenim sürecinin daha etkili hale gelmesini sağlar.

-Ülkede uygulanan öğretimin etkililiği hakkında bilgi verir.

-Bireylerin, özelliklerine uygun almalarına yönlendirilmelerini sağlar.

ÖLÇME VE DEĞERLENDİRMEYLE İLGİLİ TEMEL KAVRAMLAR

-**Değişken:** Bir durumdan diğerine farklı değerler alabilen, değişkendir.

-İnsanların boyları, ağırlıkları, zekâ düzeyleri, tutumları, başarıları, cinsiyetleri, başarıları, cinsiyetleri, milliyetleri, dilleri, ekonomik gelirleri değişkendir.

-Bir özelliğin değişken olması o özelliğin ölçülmesini gerekli kılar. Özellikler değişken olmasaydı ölçülmelerine gerek olmazdı.

Nicel Değişken

-Sayılarla ifade edilebilen değişkenlerdir, bunlar bir büyüklüğü anlatır. (Uzunluk, ağırlık, zekâ düzeyi...)

Nitel Değişken

-Sembollerle ya da sıfatlarla ifade edilebilen değişkenlerdir. (Cinsiyet, milliyet, din...)

Sürekli Değişken

-İki değeri arasında başka bir değer bulunabilen değişkenler sürekli değişkendir. Uzunluk; km, m, cm şeklinde küçültülebilir. (Ağırlık, uzunluk, yaş, zekâ düzeyi...)

Süreksiz Değişken

-İki değeri arasına sınırlı sayıda değer alan değişkendir, listelenebilir. (Cinsiyet: erkek, kadın; okul türleri: fen lisesi, meslek lisesi...)

Değişken Türleri

-Değişkenler nicel, nitel, sürekli ve süreksiz değişken olarak sınıflandırılır. Bunun yanında değişkenler, araştırmanın amacına göre de bağımlı ve bağımsız değişken olmak üzere ikiye ayrılır.

Bağımlı ve Bağımsız Değişken

-Bağımlı değişken, başka bir değişkene bağlı olarak değişir. (Sigara içme ile kanser arasındaki ilişkiye ait bir araştırmada sigara bağımsız, kanser ise bağımlı değişkendir.)

-Neden, bağımsız; sonuç bağımlı değişkendir.

S5

Ölçme ve Değerlendirme

-Ölçme, bir özelliğin gözlenip gözlem sonuçlarının sembolle ya da sayıyla ifade edilmesidir. Ölçmenin aşamaları şunlardır:

*Ölçülecek özellikler tanımlanır. (Ali'nin boyu, Veli'nin zekâsı...)

*Ölçülecek özelliğe uygun araç seçilir ya da hazırlanır. (Cetvel, tartı, tutum ölçeği...)

*Ölçülecek özellik ve araç eşleştirilerek sayı ya da başka bir sembolle ifade edilir. (Ali'nin boyu 180 cm, Veli sınavdan 85 puan aldı...)

-Öğrencinin boyunun sayılarla ifade edilmesi de öğrencilerin boy sırasına dizilmesi de ölçmedir. Bu durumda sınıftaki en uzun öğrenci Ali'dir, ifadesi de bir ölçme sonucudur. Sınıftaki öğrencileri esmerler, sarışınlar şeklinde sınıflamak da ölçme sonucudur.

Ölçme Kuralı: Ölçülen özellikle sayıların eşleşme yollarıdır. (Sınavda her sorunun 10 puan olması)

Değerlendirme: Ölçme sonuçlarının bir ölçütle karşılaştırılarak değer yargısına varılma sürecidir. Değerlendirme sürecinde üç temel öge vardır:

*Değerlendirme sürecinin başlaması için öncelikle ölçme sonucuna ihtiyaç vardır. (Sınavdan 80 puan aldı.)

*Ölçme sonucunu değerlendirmek için bir ölçüte ihtiyaç vardır. (Başarılı sayılmak için sınavdan en az 50 puan almak gerekir.)

*Ölçme sonucu ölçütle karşılaştırılarak karar verilir. (80 puan alan Ali başarılıdır.)

-Ölçme daha çok sayı ve sembollerle ifade edilirken değerlendirme daha çok sözcüklerle ifade edilir.

-Ölçme nesnel, değerlendirme öznelidir. (Ali 180 cm'dir, ölçme; Ali uzundur, değerlendirmedir.)

-Ölçme hatalıysa değerlendirme de hatalı olur.

-Ölçüt, ölçme sonuçlarına uygun olmazsa değerlendirme amacına ulaşmaz.

S7,8

Ölçütüne Göre Değerlendirme Türleri

Mutlak Değerlendirme: (Kriter Dayanaklı)

-Ölçüte diğer öğrencilerin etkisi yoktur.

-Ölçüt sınavdan önce belirlenir. (50 alan sınıfı geçer.

Bağıl Değerlendirme: (Norm Dayanaklı)

-Çan eğrisi, bağıl değerlendirmenin tipik örneğidir.

-KPSS bağıl değerlendirmeye güzel bir örnektir.

S10,11,12,13,14

Ölçme Türleri

Doğrudan (Temel) Ölçme:

-Ölçülmek istenen özellik başka bir özelliğe ihtiyaç duyulmadan gözlenerek ölçülüyorsa buna doğrudan ölçme denir.

-Bir masanın uzunluğu, sınıftaki öğrencilerin boy sırasına dizilmesi, bir kitabın sayfa sayısı...

Dolaylı Ölçme:

-Ölçmek istediğimiz özelliği başka özelliklerle ilişkilendirerek ölçme işlemidir.

-İlgi, tutum, kişilik özellikleri; KPSS ile eğitim bilimleri bilgisinin ölçülmesi, resim yapma yeteneğinin ölçülmesi, termometreyle sıcaklığın ölçülmesi, bir çocuğun zekâsının ölçülmesi...

Türetilmiş Ölçme:

-Bazı durumlarda ölçmek istediğimiz özellik, bir araç yardımıyla ölçülemez; başka ölçme işlemlerine ihtiyaç duyulur.

-Nüfus yoğunluğu=Toplam kişi/Alan, Kişi başına düşen gelir=Toplam gelir/Kişi sayısı

NOT: Eğitimde kullanılan ölçmeler genellikle dolaylı ölçmelerdir ve dolaylı ölçmelere doğrudan ölçmelere göre daha çok hata karışır.

S16,17

Ölçmede Birim

-Ankara-Amasya arası 320'dir dersek zihinlerde herhangi bir büyüklük oluşmaz. Ama 320 km dersek birimi ifade etmiş oluruz.

-Birimler tanımlanmış ve doğal birim olarak ikiye ayrılır. (km tanımlanmış bir birimken öğrencileri sayarken her öğrencinin bir birim olması doğal bir birimdir.)

-Birimlerin temel olarak taşınması gereken üç temel özellik eşitlik, genellik ve kullanışlılıktır.

Eşitlik:

-Birimin her zaman aynı büyüklükte olması durumdur. (Metre, kilogram her yerde aynı büyüklüktedir.)

Genellik:

-Aynı birimin insanlar tarafından anlaşılması demektir. (Türkiye’de santigratın Amerika’da fahrenheitin sıcaklık birimi olarak kullanılması genelliği bozmaz. Türkiye’de bazı yerlerde santigrat bazı yerlerde fahrenheit kullanılsaydı bu durum genelliği bozardı.)

Kullanışlılık:

-Birimin kullanışlı olması ölçülecek büyüklüğe uygun olması demektir. (Gezegenler arası ışık yılıyla, Ankara-Amasya arası km ile ölçülür.)

Ölçmede Sıfır

-Ölçmede iki tür sıfır vardır.

-Sınıfta hiçbir öğrencinin olmaması doğal (gerçek) sıfırdır.

-Bir öğrenci Türkçe dersinden sıfır alması onun hiç Türkçe bilmediği anlamına gelmediğinden bu sıfır, bağıl (görelî, tanımlanmış) sıfırdır.

Ölçek ve Türleri

-Ölçek, ölçme sonuçlarını gösteren sayı ya da sembollerin nitelikleridir. Dört temel ölçek vardır.

Sınıflama (Adlandırma) Ölçekleri

-Bu ölçek türünde, ölçülen özelliğin belli bir kategoriye girip girmediği belirlenir. (Bir sınıftaki öğrencileri cinsiyetine göre ayırmak, futbolcuları forma numarasıyla ayırmak, illere plaka numarası vermek)

Sıralama Ölçekleri

-Ölçülecek değişkenlerin belli bir büyüklüğe göre sıraya konulması işlemidir.

-Ali, Ahmet ve Mehmet boy sırasına göre dizildiğinde 1. Ali, 2. Ahmet ve 3. Mehmet’sen en uzun Ali, sonra Ahmet, en kısa da Mehmet’tir. Bu özellik sıralama ölçütlerinin geçişlilik özelliğidir.

-Sıralama ölçeğinde birimler arası eşit değildir. Yani Ali’yle Ahmet arasındaki farkla Ahmet’le Mehmet arasındaki fark eşit değildir.

-Sıralama ölçeğinde sıfırıncı sıra olmadığından sıfır anlamsızdır. Ölçekte başlangıç noktası da değişkendir.

Eşit Aralıklı Ölçekler

-Başlangıç noktası bağıl (görelî) sıfırdır. Ölçekteki birimler birbirine eşittir. (Dağın yüksekliğinde sıfır noktası deniz seviyesidir, deniz seviyesine bağlıdır. Miladi takvimde sıfır Hz. İsa'nın doğumudur.)

-Bu ölçekte objeler arasındaki farkı ölçebiliriz. (İki dağ arasındaki yükseklik farkı gibi)

-Eğitimde ve sosyal işlemlerde kullanılan ölçeklerin büyük bir bölümü eşit ağırlıklı ölçektir. Testlerden elde edilen standart puanlar da eşit ölçek niteliğindedir.

-Başlangıç noktası bağıl sıfır olduğundan 80 alan bir öğrencinin 40 alan bir öğrenciden iki kat fazla bilgiye sahip olduğu söylenemez.

-Bu ölçekte toplama çıkarma yapılabilirken çarpma ve bölme yapılamaz.

Eşit Oranlı Ölçekler

-Eşit oranlı ölçeğin eşit aralıklı ölçekten temel farkı başlangıç noktasının gerçek (mutlak) sıfır olmasıdır.

-Birimler arası eşittir ve bu özelliğinden dolayı toplama, çıkarma, çarpma ve bölme yapılabilir.

-Cisimlerin ağırlıkları, uzunlukları ve hacimleri eşit oranlı ölçekle ölçülebilir.

S21,22

ÖLÇME ARAÇLARINDA BULUNMASI GEREKEN NİTELİKLER

-Ölçme araçlarının etkili olabilmesi için gerekli özellikler "güvenirlilik, geçerlik ve kullanılabilirlik."

-Ölçme araçlarının kalitesi genellikle hatayla ilişkilidir.

HATA VE TÜRLERİ

-Özelliklerin gerçek değerleriyle gözlenen değerleri arasındaki farka "hata" denir.

-En hassas araçlarla yapılan ölçmelere bile bir miktar hata karışabilir.

-Hatalar üçe ayrılır:

Sabit Hata

-Ölçme sonuçlarına aynı yönde ve miktarda karışan hatalardır.

-Geçerliliği etkiler.

-Her öğrenciye 5 puan fazla verilmesi, bir cetvelin 1 cm eksik ölçmesi...

Sistematiik Hata

-Ölçülen büyüklüğe, ölçmeciye ya da ölçme koşullarına bağılı olarak miktarı değışen hatalardır. Bu hatalara yanlışlık adı da verilir. Kontrol edilebiliridir.

-Geçerliliğı etkiler.

-Öğretmenin yazısı güzel öğrencilere 10 puan fazla vermesi, öğretmenin önce okuduğı yazılılara az sonra okuduklarına daha fazla puan vermesi, kız öğrencilere fazla puan verilmesi...

Tesadüfi Hata

-Kaynağı ve yönü belli olmayan hatalardır. Genellikle ölçme anında ortaya çıkar. Puanları artı ya da eksi yönde etkiler.

-Kaynağı belli olmadığı için kontrol edilemez ve engellenemez.

-Güvenirliğı etkiler.

-Öğretmenin yazılı kâğıtlarına rastgele fazla puan vermesi, öğrencinin sınavda hastalanması, kaygı taşınması, öğretmenin dikkatsizlik sonucu puanları yanlış toplaması veya not defterine yanlış yazması...

(Ölçme teorileri daha çok tesadüfi hatalara odaklanmıştır. Tesadüfi hataların az olması ölçme aracını güvenilir kılar.)

S29

GÜVENİRLİK (0-1)

-Bir ölçme aracının tesadüfi hatalardan arınık olma derecesidir.

-100 soruluk bir test 5 soruluk bir testten daha hassas olacağından güvenilirliğı artar. Burada güvenilirlik "duyarlılık" anlamında kullanılmıştır.

-Kilomuz değışmediğı halde bizi bir 80 kg, bir 90 kg tartan bir baskülün doğru tarttığını düşünmeyiz. Ölçülmek istenen özellik değışmedikçe aynı aracın benzer sonuçlar vermesi aracın kararlı ve güvenilir olduğunu gösterir. Burada güvenilirlik "kararlılık" anlamında kullanılmıştır.

-Eğitimde kullanılan güvenilirlik türü de "tutarlılık" anlamında güvenilirliktir.

-Bir test maddesine verilen cevapların, testin tamamıyla tutarlı olması durumdur.

-Bir testin gvenirlik deęeri 0 ile 1 arasında deęerler alır.

lme aracının gvenilir olması iin alınması gereken nlemler:

-Soru sayısını artırmak

-Soru aık ve anlařılır bir dille yazmak

-Testteki soruların dil seviyesini herkesin anlayabilmesi iin testin uygulanacaęı grubun dil seviyesinin altında belirlemek

-ęrencileri sınav sorularını cevaplamaya motive etmek

-Sınav sresini uygun belirlemek

-Soruların zorluk seviyesini orta dzeyde belirlemek

-Sınav puanlamasını objektif olarak yapmak

-Sınav ortamını uygun dzenlemek (ısı, ses, ıřık...)

-Sınavda ka soru bulunduęunu, puanlamanın nasıl yapılacaęını bir ynergeyle ęrencilere bildirmek

-Soruların soruluř sırasına dikkat etmek (ok zor soruları bařa koymamak)

-Testin yapısını, dizgisini, harflerin puntosunu sınava gireceklerin dzeyine uygun olarak belirlemek

-řans bařarısını azaltmak (5 seenekli bir test, 4 seenekli bir testten daha gvenilirdir.)

-Sınav kaygısı gibi ęrencilerin gerek performansını etkileyebilecek tutumları azaltmaya alıřmak

S31

Korelesyon (r)

-İki deęiřken arasındaki iliřkinin miktarını ve ynn tanımlayan istatistiksel bir tekniktir.

-Korelasyon katsayısı hesaplama yolları ve trleri sınav kapsamı dıřındadır.

-Korelasyon katsayısının pozitif olması, iki deęiřkenin birlikte artması demektir. (Zekâ ve bařarı arasındaki korelasyon 0,85)

-Korelasyon katsayısının negatif olması, deęiřkenlerden birinin artarken dięerinin azalması demektir. (Sigara ime ile saęlık arasındaki korelasyon -0,75)

-Korelasyon katsayısının sıfıra yakın olması, iki değişken arasındaki ilişkinin çok zayıf olması demektir. (Boy ile sigara içme arasındaki korelasyon 0'dır ya da 0'a yakındır.)

Neden-Sonuç İlişkisi ve Korelasyon

-İki değişken arasındaki ilişkinin yüksek korelasyon göstermesi, her zaman birinin diğerinin nedeni olduğu anlamına gelmez. (Okuma hızı ile tüketilen yiyecek miktarı arasındaki korelasyonun yüksek çıkması okuma hızıyla yaş arasındaki korelasyondan kaynaklanır. Yaşı daha çok olan bireyler de daha hızlı okur.)

S33,34

Güvenirlilik Kestirme Yolları

Test-Tekrar Test Yöntemi

-Bir gruba bir test uygulanır, puanlar kaydedilir. Bir süre sonra aynı test tekrar uygulanır. İki uygulama arasındaki tutarlılığa bakılır. İki uygulama arasındaki korelasyon katsayısı testin güvenirlik katsayısı olarak yorumlanır. İki uygulama arasındaki süre çok uzun ya da çok kısa olmamalıdır.

Paralel (Eşdeğer) Formlar Yöntemi

-Test-tekrar test yönteminde meydana gelen hatırlama problemini ortadan kaldırmak amacıyla hazırlanmıştır.

-Paralel form, aynı davranışı ölçen, aynı zorluk düzeyine ve aynı soru sayısına sahip iki teste verilen addır.

-İki paralel test aynı gruba uygulanır.

-Farklı zamanda ama test-tekrar test yönteminden daha kısa aralıkla yeniden uygulanır. İki uygulama arasındaki korelasyon katsayısı güvenirlik katsayısı olarak yorumlanır.

Eşdeğer Yarılar Yöntemi

-Bu yöntemde bir gruba bir test uygulanır. Daha sonra test rastgele iki yarıya ayrılır. (10 soruluk bir testin tek sayılı soruları bir gruba çift sayılı soruları diğer gruba, gibi) Elde edilen iki puan arasındaki korelasyon katsayısı güvenirlik katsayısı olarak yorumlanır.

Kuder Richardson Yöntemi (KR-20 ve KR-21 Formülleri)

-Bu yöntem iç tutarlılık katsayılarının ölçümünde kullanılır. Cevapların 1 ya da 0 (doğru ya da yanlış) olarak sınıflanabildiği testler için uygundur. Eğer bu katsayı yüksekse (1'e yakınsa) testin tesadüfi hatalardan arınık olduğu sonucuna varılır.

Cronbach Alpha Katsayısı

-Cevapların 1 ya da 0 olarak (doğru ya da yanlış) sınıflanamadığı araçlar için uygun KR-20 ve KR-21 yöntemine benzer bir güvenirlik katsayısıdır.

-Duyuşsal özelliklerin ölçülmesinde kullanılır. (Tutum ölçekleri gibi)

Puanlayıcılar Arası Güvenirlik

-Özellikle puanlaması sübjektif olan araçlarda (yazılı, sözlü gibi) kullanılan güvenirlik kestirme yöntemidir. Cevap kağıtlarının birden çok puanlayıcı tarafından puanlanarak korelasyon katsayısının bulunması ile bulunur.

Not: Test-tekrar test ve paralel formlar yöntemleri kararlılığa yönelikken eşdeğer yarılar ile KR-20 ve KR-21 yöntemleri iç tutarlılığa yöneliktir.

Tesadüfî Hata ve Güvenirlik

-Aynı özelliğin farklı zamanlardan ölçülmesi sonucunda meydana gelen tesadüfî hatalar arttıkça bu iki olayın kararlı ve tutarlı sonuçlar vermesi mümkün olmaz. Bu yüzden tesadüfî hatalar testlerin güvenirliğiyle doğrudan ilgilidir. Bu hataların önüne geçmek için test-tekrar test ya da paralel formlar yöntemi uygulanarak testlerin tutarlılığı buna bağlı olarak da güvenirliği hakkında bilgi sahibi olunur.

-Sabit ve sistematik hatalar güvenirliği etkilemez. Her öğrencinin puanını 5 puan artırırsak sonuçlar yine tutarlı dolayısıyla da güvenilir olacaktır. Ancak sabit ve sistematik hataların geçerliğe etkisi iki türdür:

*Sabit hata KPSS gibi bağıl değerlendirmelerde geçerliği etkilemez. Her öğrenciye 5 puan eklense de atanan öğrenciler değişmez.

*Mutlak değerlendirme yapılacaksa o zaman sabit hata geçerliği etkiler. 50 puan alanın Başarılı sayılacağı bir testte her öğrenciye 5 puan eklenmesi bazı başarısız öğrencilerin başarılı olması anlamına gelir.

Asgari Güvenirlik Katsayısı

-Güvenirliğin ideal değer 1 ise de buna ulaşmak imkansız olduğundan uygulayıcılar daha düşük güvenirlik katsayıları elde ederler.

-Her test için ortak bir güvenirlik katsayısı söylemek mümkün değildir ancak genel olarak standart çoktan seçmeli testlerin güvenirliği .85-.95; açık uçlu, yazılı yoklamaların .65-.80; portfolyo puanlama .40-.60 arasında değişir.

-Burada önemli noktalardan biri de bireyler hakkında verilecek kararlardır. KPSS gibi bireylerin geleceğini ilgilendiren testlerde geçerliğin olabildiğince yüksek olması gerekir. Bu durumda çok sayıda ölçme aracından elde edilecek kararlar puanların geçerliğini de artırır.

S39,40,41

GEÇERLİK (-1,+1)

-Güvenirlik, bir ölçme aracından elde edilen puanların tesadüfi hatalardan arınık olması ya da özellik pek değişmediği sürece ölçme sonuçlarının da değişmemesi olarak tanımlanır. Ancak araçların güvenilir ölçme yapması yeterli değildir.

-Araçların, ölçülmek istenen nitelikleri ölçmeleri, amaca uygun ölçmeleri "geçerli" olmaları anlamına gelir.

-Geçerlikte ölçme araçlarının neyi ölçtüğü yani içerik önem kazanmaktadır.

-Geçerlik kestirme yolları, genellikle korelasyon tekniğine dayanır. Geçerlik -1 ile +1 arasında değerler alır. +1'e yaklaşması artması, uzaklaşması ise düşmesi anlamına gelir.

Geçerlik Türleri

1)Kapsam Geçerliği

-Bir ölçme aracının, ölçülmek istenen konu, ünite ya da davranışları kapsama derecesidir.

-Bir testin hangi davranışları kapsayacağını belirtke tablosuyla belirleyebiliriz. Belirtke tablosu, davranışların konulara göre dağılımını gösteren bir çizelgedir.

Kapsam Geçerliğini Belirleme Yolları

-Uzman yardımına başvurma

-Aynı kapsamı ölçen başka bir testle korelasyon <http://egitimvaktim.com/>

S44,45

2)Bir Ölçüte Dayalı Geçerlik

a.Yordama Geçerliği

-Geleceğe yönelik tahmin amacı güden testin bu tahmininin ne derecede gerçekleştiği araştırılır.

-Bir öğrencinin LYS’de aldığı puanlarla yerleştiği bölümdeki puanları karşılaştırılırsa bu sınavın yordama geçerliği ortaya konmuş olur.

S46

b.Uygunluk (Uyum) Geçerliği

-Yordama geçerliğiyle uygunluk geçerliği arasındaki temel fark, ölçüt puanların elde edildiği zamandır. Yordama geçerliğinde ölçüt puanlar yordayıcı puandan sonra elde edilirken uygunluk geçerliğinde daha önce elde edilir.

-Deneme sınavları ile LYS arasındaki korelasyon katsayısı uygunluk geçerliği olarak yorumlanır.

Bir Ölçüte Dayalı Geçerlik Çalışmalarında Dikkat Edilmesi Gerekenler

-Yordayıcı ve ölçüt puanların güvenilirliklerinin yüksek olması gerekir. Güvenirlik düşükse geçerlik de düşük olacaktır.

-Ölçüt puanlar ile yordayıcı puanlar birbiriyle ilişkili olmalıdır. Öğrencilerin beden eğitimi puanlarına bakarak sayısal-sözel alan seçmelerine karar vermek geçerli olmayacaktır.

3)Yapı Geçerliği

-Bir insanın müziğe olan ilgisini gözleyemeyiz ama her fırsatta bir müzik aleti çalmak isteyen kişinin müziğe ilgisi olduğunu anlayabiliriz.

-Bir testin yapı geçerliği, testle ölçülmek istenen yapının ortaya konulma derecesidir.

-Nasıl ki psikolojik yapılar doğrudan gözlemlenemeyip bağlı olduğu düşünülen davranışlar aracılığıyla belirlenir, ölçme aracının yapı geçerliği de böyledir.

4)Görünüş Geçerliği

-Testteki soruların başka hiçbir işleme gerek kalmadan istenilen değişkeni ölçüyor görünüp görünmemesi durumudur.

Geçerliği Etkileyen Faktörler

-Güvenirliği düşük bir testin geçerliği de düşük olacaktır.

-Soru sayısının artması kapsam geçerliğini artıracaktır.

-Sabit ve sistematik hatalar geçerliği düşürür.

- Sınavdan önce soruların açıklanması geçerliđi düşürür.
- Puanlamadan kaynaklı hatalar geçerliđi düşürür. (Yazı güzelliđine puan verilmesi gibi)
- Öğrencilerin kopya çekmesi geçerliđi düşürür.
- Her bir maddenin ölçölmek istenen değışkeni ölçmesi, testin tamamının geçerliđini artırır.
- Cevaplama süresinin çok uzun ya da çok kısa olması geçerliđi düşürür.
- Soruların bilenle bilmeyeni ayırması geçerliđi artırır.
- Sınavla ilgili bir yönergenin verilmesi (soru sayısı, süre...) geçerliđi artırır.
- Test, dil seviyesi açısından grubun altındaysa geçerlik yükselir.
- Maddelerde yer alan belirsiz ifadeler geçerliđi düşürür.
- Sorular kolaydan zora doğru sıralanmışsa geçerlik artar.
- Testin çok kısa olması geçerliđi düşürür.
- Ölçüt olarak seçilen grubun güvenilirliđi de testin geçerliđini etkiler.

S48,49

Güvenirlik ile Geçerlik Arasındaki İlişki

Testin en önemli özelliđi "geçerliđidir" çünkü,

- Bir test geçerliyse zaten güvenilirirdir. (Test, istenilen özelliđi ölçüyorsa hatalardan arınık demektir.)
- Güvenilir bir test, geçerli olmayabilir. (Güvenirliđi yüksek bir matematik testi, fen bilgisi ölçmek için geçerli değildir.)
- Testin güvenilirliđi geçerliđini sınırlar. (Güvenirliđi 0,90 olan bir testte 0,10 hata var demektir, bu da geçerliđi sınırlar.)
- Güvenirlik ve geçerlik bir derecedir, varlık ya da yokluk durumu değildir. (Bir testin güvenilirliđi ve geçerliđi vardır/yoktur yerine yüksektir/düşüktür, denir.)

50,51,52

KULLANIŞLILIK

-Ölçme aracı ekonomik olmalıdır.

-Araçların hazırlanması kolay olmalıdır.

-Puanlama hızlı, ekonomik ve kolay olmalıdır.

-Puanlar kolayca yorumlanabilmelidir.

EĞİTİMDE KULLANILAN TESTLERİN SINIFLANDIRILMASI

Ölçülen Davranışın Yapısına Göre

-Tipik Performans Testleri: Amacı, bireylerin ortaya koydukları tipik davranış örüntülerini belirlemektir. (Tutum, ilgi, kişilik özellikleri gibi)

-Maksimum Performans Testleri: Amacı, bireylerin belirli bir zamanda ortaya koyacakları maksimum performansı ölçmektir. (Başarı, yetenek, zekâ, performans testleri gibi)

Hazırlanma Şekline Göre

-Standart Testler: Bir uzman grubu tarafından hazırlanır ve standart şartlarda uygulanır. (Zekâ, yetenek testleri gibi)

-Öğretmen Yapımı Testler: Öğretmen tarafından sınıf düzeyini belirlemek amacıyla hazırlanır. (Yazılı gibi)

Puanlama Türüne Göre

-Objektif Testler: Genellikle bilgisayarlar yardımıyla puanlanan testlerdir. Puanlama hatası taşımazlar. (Çoktan seçmeli testler gibi)

-Sübjektif Testler: Puanlaması kişisel kanaatlerden etkilenen testlerdir. Yanlılık varsa sistematik hata karışabilir; puanlama güvenilirliği ve geçerliği düşer. (Yazılı, sözlü gibi)

Ayrılan Zamana Göre

-Hız Testleri: Bireylerin verilen zaman aralığında cevapladığı soruları ölçer.

-Güç Testleri: Genellikle soruların güçlük düzeyine göre sıralandığı testlerdir.

Cevap Formatına Göre

-Kâğıt-Kalem Testleri: Öğrencilerin, cevapları yazılı olarak verdikleri testlerdir. (Yazılı gibi)

-Seçme Gerektiren Testler: Öğrencilerin, cevapların birini seçtiği testlerdir. (Çoktan seçmeli test gibi)

Uygulanan Kişi Sayısına Göre

-Bireysel Testler: Tek tek bireylere uygulanan testlerdir. (Yazılı gibi)

-Grup Testleri: Aynı anda çok sayıda kişiye uygulanan testlerdir. Ekonomiktir.