

MODÜL 2

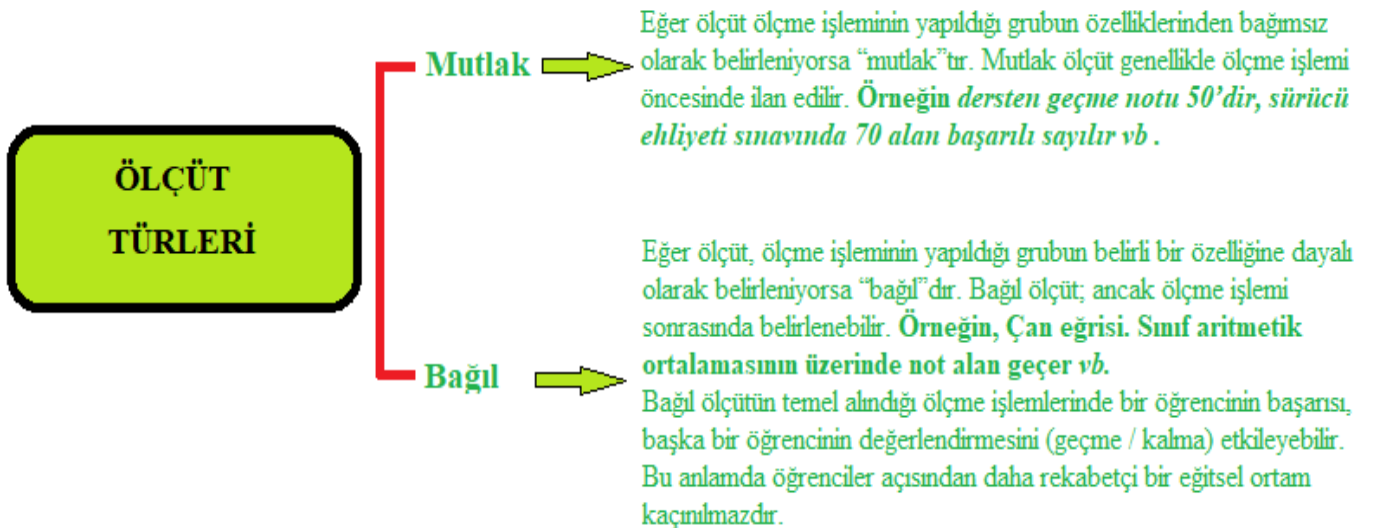
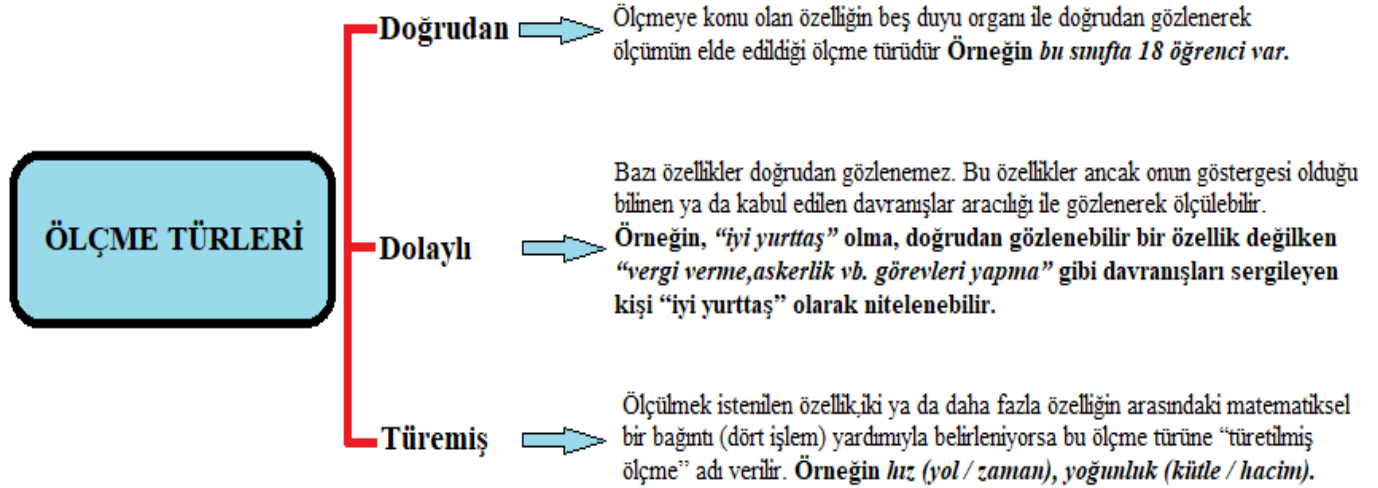
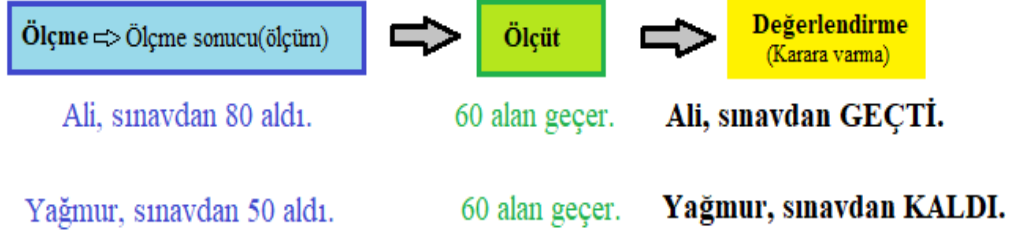
ÖLÇME VE DEĞERLENDİRME

1. ÖLÇME, ÖLÇÜT VE DEĞERLENDİRMENİN KAVRAMSAL TEMELLERİ

Ölçme: Öğrencinin belirli bir özelliğini gözleyerek o özelliğe sayı, sembol ya da sıfat/kategori adı verme işlemidir.

Ölçüt: Ölçülen özellik hakkında karar alabilmek / yargıya varabilmek / değerlendirme yapabilmek için dayanak alınan referans noktası ya da referans aralığıdır.

Değerlendirme: Bir ölçme sonucunu (ölçüm) en az bir ölçüte vurarak ölçülen nitelik hakkında karar verme / yargıda bulunma işlemidir.



DEĞERLENDİRME

Tanım- Yerleştirmeye Yönelik

Sınıf içi ölçme ve değerlendirme:

Öğretim sürecinin başında, öğrencilerin önceki kritik nitelikteki öğrenmelerini ders bağlamındaki hazırbulunuşluk düzeyini belirlemek amacıyla yapılır. "hazırbulunuşluk testi"

Amaç not vermek değildir.

Okul / kurum geneli ölçme ve değerlendirme:

Yine öğretim sürecinin başında, öğrenciyi tanımak ve onu uygun olan programa / kura / sınıfa yerleştirmek amacıyla yapılır. Amaç not vermek değildir. Hedeflenen, öğrencilerin niteliklerine uygun grupların oluşturulmasıdır. Zaman, para ve emek tasarrufu sağlar. Muafiyet sınavları ve seviye tespit sınavları buna örnek gösterilebilir.

Bıçimlendirme-Yetiştirmeye Yönelik

Belirli bir konu / ünite vb. sonunda öğrencilerin öğretime konu olan davranışların ne kadarını kazandığını, ünitedeki öğrenme eksiklerini belirlemek amacıyla yapılır. Amaç yine not vermek değildir.

Bu değerlendirme türünde öğrenciyeye o ünitedeki öğrenme eksikleri bildirilir, bir sonraki kapsama geçmeden öğrenme eksiklerini giderilip giderilmediği denetlenir.

Bıçimlendirme-yetiştirme amacıyla yapılacak değerlendirmede kullanılan testlere alanda "izleme testi" ya da "tarama testi" denir. Not verilmeden yapılan kısa sınavlar (quiz), ünite tarama testleri bu amaçla yapılan ölçme ve değerlendirmeye örnek olarak verilebilir.

Değer Bıçmeye-Düzyel Belirlemeye Yönelik

Belirli bir öğretim sürecinin sonunda, not vermek amacıyla yapılan değerlendirmedir.

Değer bıçmeye yönelik değerlendirme, öğrencinin belirli bir öğretim süreci sonundaki durum tespittir; çünkü artık hedeflenen öğretim sürecinin sonuna gelinmiştir.

Bu amaçla uygulanan testlere "**alanda erişme testi**" denir. Bitirme sınavları, sertifika sınavları, üniversitedeki vize ve finaller, okullarda not vermek amacıyla uygulanan yazılı yoklamalar bu ölçme ve değerlendirme türüne örnek olarak verilebilir.

Rehberlik Amaçlı

Özel eğitim: Belirli alanlarda özel gereksinimi olan çocuklara uygun eğitim ortamları ve/veya programı uygulamak amacıyla ölçme ve değerlendirme süreçleri işe koşulmaktadır. Özel gereksinimi olan çocuklara uygun eğitim olanaklarının sunulabilmesi için çocuğun ihtiyacı olan alanlar, eğitsel tanımlama amacıyla genellikle standart testler veya ölçüt bağımlı testlerle (beceri öğretimde) belirlenmeye çalışılır. Bu tür testlerin sonuçları doğrultusunda, tıbbi ve eğitsel / gelişimsel olarak tanı alan çocuklar için bireysel eğitim programları (BEP) hazırlanır.

Eğitsel değerlendirme bağlamında RAM'lar bu konudaki tek yetkilidir.

Mesleki rehberlik: Öğrencilerin alan seçmelerinde ve özellikle bir üst öğretim kurumuna yönlendirilmesinde, özelliklerine uygun alan / meslek seçmeleri hem öğrenci hem de ülkenin insan gücü kaynağının doğru planlanması açısından çok önemlidir. Bu amaçla öğrencilerin kendi yetenekleri, kişilik özellikleri, ilgileri vb. ile özelliklerini belirlemek amacıyla standart testlerden yararlanılmaktadır.

Program Değerlendirme

Eğitim programları bileşenlerinin yapısı ve niteliğine ilişkin olarak sistematik bilgi toplama ve değerlendirme sürecidir. Program değerlendirme eğitim programlarının planlanması, mevcut programlar ve/veya ürünlerin etkililiğinin değerlendirilmesi ve eğitim programlarının ve/veya ürünlerin geliştirilmesi amacıyla yapılır.

Eğitim programlarını değerlendirme, değerlendirilmekte olan programın boyutlarıyla ilgili sorularla yönlendirilen uygulamalı bir etkinliktir. Program değerlendirmede program taslağının hazırlanması, programın uygulanması ve programın etkisine ilişkin yargılar elde edilir.

Mustafa Uçan



Geçerlilik ölçmek istenilen özelliğin, başka özelliklerle karıştırılmadan, doğru ve tam olarak ölçülebilmesidir. Diğer bir ifadeyle ölçme aracından elde edilen puanın amaca hizmet etme derecesidir. **Örneğin, otobüs bileti uçağa binemeyiz. O bilet uçak için değildir ve dolayısıyla geçersiz olur.**

Güvenilirlik ölçme işleminden elde edilen puanların tesadüfi hatalardan arınık olma derecesidir. Diğer bir ifadeyle puanların kararlı, tutarlı ve duyarlı olmasıdır. Kelime kökü güvenmektir. **Örneğin, bir insan sürekli doğru konuşuyor; saygılı, tutarlı ve kararlı davranışlarda bulunuyorsa ona GÜVENİLİR. Yani hatası, tutarsızlığı, kararsızlığa yol açan durumu yoktur. İşte ölçme araçlarındaki puanların da bunun gibi istikrarlı, kararlı, tutarlı sonuçlar vermesini bekleriz. Bunu sağlarsa güvenilir ölçme aracı olur.**

Kullanışlılık ise ölçme aracının geliştirilmesinin, uygulanmasının ve puanlanmasının kolay ve ekonomik (zaman, para, emek, araç gereç vb. açıdan) olması ile ilgilidir.

Bir ölçme aracından elde edilen puanların hatasız ya da az hatalı olması geçerli ve güvenilir ölçme yapmanın tek koşuludur. Ölçme işlemlerinde hataları en aza indirebilmek için öncelikle hataların tanımlanması gerekir ki ölçme işlemi yapan kişi neye müdahale edeceğini, hangi durumlara karşı önlem alması gerektiğini bilsin.

HATA

Hata, gerçek değer ve ölçülen değer arasındaki farktır. Ölçmede gözlenen bir özelliğin gerçek değeri (sayı, sembol ya da sıfat) bulunmak istenir. Ancak ölçmeye karışan çeşitli hatalar nedeniyle gerçek değer (puan) ölçme yoluyla doğrudan elde edilemez. Psikometride gerçek puan kuramı olarak ele alınan bu kavram basit bir eşitlik ile gösterilir.

$$X = T + E$$

X = Bireylerin ölçme aracından elde ettiği gözlenen puanı

T = Bireylerin gözlenemeyen gerçek puanı

E = Ölçmeye karışan hata miktarı

Sabit ve sistematik hatalar zaman zaman birbirine karıştırılabilir

Örnek	Sabit Hata	Sistematik Hata
Oğretmenin sevdiği öğrenciye puan vermede bonkor davranması		✘
Oğretmenin herkesin puanına %10 kadar eklemesi		✘
Bir sorunun silik olması nedeniyle öğrencilerin soruyu cevaplayamamaları	✘	
Sınav için verilen zamanın yetmemesi nedeniyle bazı öğrencilerin soruların tamamına ulaşamaması		✘
20 sorudan oluşan bir testin arka yüzünün baskısının alınmaması	✘	
Terzimin mezurasının 150 cm yerine aslında 145 cm olması	✘	✘

HATA

Sabit Hata

Her ölçme işlemine aynı miktarda karışan hatalardır.
Örneğin, Bir öğretmen, sınavında herkese 10 puan fazla veriyorsa yine karışan hata sabit olacaktır.

Sistemik Hata

Ölçülen büyüklüğe, öğretmene ya da ölçme koşullarına göre miktarı değişen hatalardır.
Örneğin, bir öğretmen, sınavında yazısı kötü olandan puan kırılırsa yine karışan hata sistemik olacaktır.

Tesadüfi (Rastlantısal) Hata

Şansla ortaya çıkan ne yönde ve ne ölçüde karıştığı genellikle bilinmeyen hatalardır. Ölçmelere tek yönlü olarak karışmaz, ölçme sonuçlarına bazen pozitif bazense negatif yönde etki eder. Psikometride hata kuramı tesadüfi hatalar üstüne kurulmuştur.

Ölçme İşlemini Yapan Kişiden Kaynaklanan Hata

Öğretmenin
Ölçme işlemi sürecinde: Test katılımcısının dikkatini dağıtacak veya kaygısını artıracak davranışlar vb.
Ölçme işlemi sonrasında: Puanlamadaki dikkat ve titizliğin zamandan zamana değişmesi, yorgunluk, öncelik sonralık yanlılığı, maddi hata vb.

Ölçme Aracından Kaynaklanan Hata

Ölçme araçları hazırlanırken maddelerin iyi ifade edilmemesi, test katılımcılarının yanlış anlamalarına ve dolayısıyla hataya neden olacaktır.
Ölçme aracındaki madde sayısı arttıkça bireyde yorgunluk, dikkat azalması vb. etkiler nedeniyle hata miktarı artacaktır.
Bilişsel özellikleri ölçen testlerde eğer seçenek verilmişse (D-Y testi) şans başarısı (doğruyu bulma) karışma olasılığıdır.

Bireyden Kaynaklanan Hata

Bireylerin ölçme işlemi sürecinde içinde buldukları fiziksel, fizyolojik ve psikolojik durumlar test puanına etki eder. Uykusuzluk, açlık, hastalık, ağrı, kaygı vb. faktörler nedeniyle bireyler gerçek puanlarından uzaklaşabilirler. Ayrıca grubun homojen ya da heterojen olması da hata kaynağını artıran ya da azaltan bir etkiye sahiptir. Daha heterojen gruplarda güvenilirlik katsayısı artarken daha homojen gruplarda güvenilirlik katsayısı azalır.

Fiziksel Ortamdan Kaynaklanan Hata

Ölçme işleminin gerçekleştiği fiziksel ortama ilişkin bazı etmenler de ölçme sonuçlarına hata karıştırabilir. Sıcaklık, ışık, ses, koku, görsel uyarıların fazlalığı, ergonomi vb. etmenler hata miktarını artırabilir. Bu anlamda testler bireylere bu etmenler açısından eşit ve standart koşullar altında uygulanmalıdır, aksi takdirde hata miktarının artması kaçınılmazdır.

Gerek sistemik gerekse sabit hatalarda ölçme sonuçlarına karışan hata miktarı, yönü ve kaynağı bellidir. Bu nedenle bu tür hataların ölçme sonuçlarına karışmasını engellemek ya da bu tür hataları düzeltmek görece daha kolaydır. Tesadüfi hatalarda ise kaynak çoğu zaman belli değildir; düzeltmek daha zordur.

GEÇERLİLİK, GÜVENİLİRLİK VE HATA İLİŞKİSİ



Bir ölçme işleminde geçerlilik, bütün hata kaynaklarından olumsuz yönde etkilenir. Çünkü bu hatalar beni ölçmek istediğim özelliği ölçebilme amacımdan dolayısıyla geçerlilikten uzaklaştırabilir.

Güvenilirlik ise sadece tesadüfi hatalardan etkilenmektedir. Çünkü işin içine şans faktöründen kaynaklı hatalar karışacağından ölçmek istediğim özellik ile ilgili tutarsızlıklar, kararsızlıklar meydana gelecek ve bu da ölçme aracımın güvenilirliğini olumsuz etkileyebilir.

Güvenilirlik, geçerlilik için bir ön şart ancak yeterli şart değildir. Diğer bir ifade ile **bir testin güvenilir olması onun geçerli olacağı anlamına gelmez ancak bir test geçerli ise büyük olasılıkla güvenilirdir.** Çünkü bir testin tesadüfi hatalardan arınmış olması geçerli olması için yeterli değildir. Tesadüfi hata ile beraber sabit ve sistematik hatalardan da arınmış olması gerekir.

Ölçme araçlarının **psikometrik nitelikleri (geçerlik ve güvenilirliği)** şu durumlara göre değişebilir:

a. Ölçme amacının değişmesi: Örneğin, üniversiteye öğrenci seçme sınavı bu amaç için geçerliken KPSS için kullanılamaz. Amaca hizmet etmez ve geçerlilik, dolayısıyla güvenilirlik değişir.

b. Uygulama grubunun değişmesi: İlkokul öğrencileri için geliştirilen bir test veya sınav ortaokul öğrencileri için geçerli ve güvenilir olmayabilir.

c. Dilin eskimesi / değişmesi: Uzun yıllar önce geliştirilmiş bir testin dili eskimiş olabilir veya dilde değişimler olabilir. Günümüzde anlaşılması zor kelimeler içerebilir ki bu da geçerlilik ve güvenilirliği etkiler.

d. Maddelerde ve/veya alt ölçeklerde yapılan değişiklikler: Bazen amacımız doğrultusunda maddelerde, sorularda, ifadelerde değişiklikler yaparız. Bu da geçerlilik ve güvenilirliği etkileyebilir.

e. Farklı kültürler: Bir kültürde geliştirilen bir ölçme aracı başka kültürlerde aynı geçerlilik ve güvenilirliğe sahip olmayabilir.

f. Kuramsal bilgi birikiminde değişiklikler: Yeni bilgiler eklendikçe, çıkarıldıkça veya değiştirildikçe puanlar dolayısıyla geçerlilik ve güvenilirlik de değişir.

Geçerlilik ve güvenilirlik bir varlık-yokluk sorunu değil, **derece sorunudur.** Bu nedenle 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.

Geçerlilik ve güvenilirliği belirlemeye yönelik yöntemler genellikle **korelasyon temelli analizlere dayanır.**

Mustafa Uçan

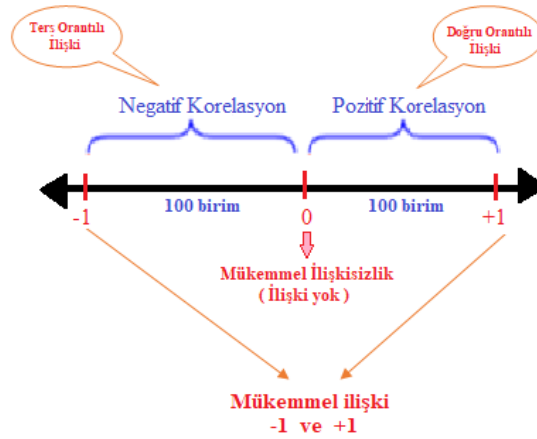
KORELASYON (co-relation): En az iki değişken arasında karşılıklı bir ilişki bulunup bulunmadığı, ilişki varsa bu ilişkinin yönü ve miktarı hakkında bilgi veren istatistik bir tekniktir. **Başarım hakkında bilgi vermez.**

KORELASYON = İlişki var mı, yok mu? İlişki yönü ve miktarı ne?

- ✓ Korelasyon -1 ile 1 arasında değer alır. Bu iki değer arasında matematiksel olarak sonsuz birim vardır ancak kullanılabilirlik açısından 100 birim negatif korelasyonda, 100 birim pozitif korelasyonda, sıfır ile birlikte toplam 201 birimlik bir skala üzerinden değişkenler arasındaki ilişkinin yönü ve miktarı değerlendirilir.
- ✓ Pozitif korelasyon iki değişken arasında doğru orantılı ilişki anlamına gelir. ++ veya -- olur gibi düşün. **Örneğin ders çalışma süresi ile sınav notu ya da gelir ile tüketim arasında pozitif bir korelasyonun elde edilmesi beklenir.**
- ✓ Negatif korelasyon, iki değişken arasında ters orantılı ilişki anlamına gelir. +,- veya -,+ gibi düşün. **Örneğin ders süresi ile dikkat ya da yükseklik ile sıcaklık arasında negatif bir korelasyonun elde edilmesi beklenir.**

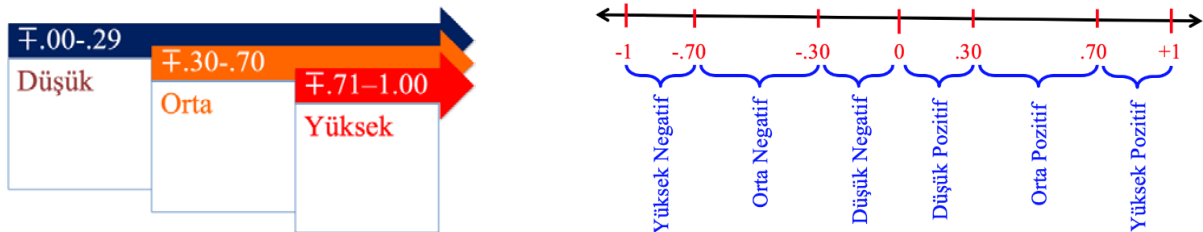
∓ 1 mükemmel korelasyon anlamına gelir; ancak +1 mükemmel doğru orantılı ilişki anlamına gelirken, -1 ise mükemmel ters orantılı ilişki anlamına gelir.

- ✓ Diğer taraftan sıfır korelasyon iki değişken arasında sistematik bir ilişkinin olmadığı anlamına gelir. Bu duruma mükemmel ilişkisizlik de denir. **Örneğin öğretmenlerin boyları ile aylık ücretleri arasındaki korelasyonun sıfır olması beklenen bir durumdur.**



Korelasyon katsayısı hakkında kabaca iki tür belirleme yapmak gerekir. Bunlar:

Yön, negatif ya da pozitif olarak değerlendirilirken; **Miktar**, kabaca düşük, orta ya da yüksek olarak nitelendirilir.



Not 1: Korelasyon mutlak değer olarak değerlendirilmelidir. Bir korelasyon katsayısının negatif ya da pozitif olması büyüklük-küçüklük belirtmez, yön bildirir.

Not 2: Korelasyon katsayısı ile neden-sonuç ilişkisi kurulamaz. Değişkenler arasında doğru ya da ters orantılı bir ilişki olması, söz konusu değişkenler arasında bir neden-sonuç ilişkisinin varlığı anlamına gelmez.



A. KAPSAM GEÇERLİLİĞİ:

Kapsam geçerliliği, özellikle başarı testlerinde aranan bir geçerlilik sorgulamasıdır.

Örneğin, bir öğretmen başarıyı ölçecekse geçerlilikle ilgili zihindeki öncelikli problem kapsam geçerliliği olmalıdır.

Kapsam geçerliliği, bir testin ölçülmek istenen davranışları(kazanımları) ve öğretime konu olan içeriği (konu kapsamını) ne ölçüde temsil ettiğiyle ilgilidir.

Bir testin kapsam geçerliliğinin yüksek olması için; **Testteki soruların / maddelerin**

- 1.İçeriği (konu kapsamını) yeterli ve dengeli bir biçimde ölçmesi,
- 2.Kazanımla doğrudan ilgili olması gerekir.

'O kadar çalıştım öğretmen sadece 1 konudan sormuş.' durumu olmasın.

Kapsam geçerliliğinin sorgulanmasında **belirtke tablosu(Kazanım tablosu)** kullanılabilir veya **uzmanlardan** istatistiksel yöntemle(*Uygundur / uygun değildir* ” ya da “*uygundur / düzeltme gerekir / soru kullanılmamalıdır*” vb.) yardım alınabilir.

B. YORDAMA GEÇERLİLİĞİ:

Yordama, eldeki geçerli ve güvenilir verilerden yola çıkarak geleceğe, henüz gerçekleşmemiş bir olguya ilişkin yapılan tahmindir. **Bilinenden yola çıkarak bilmeyen hakkında tahminde bulunmaz. Şu ana bakarak geleceği tahmin etme.**

Örneğin, benim bugün yağmur yağıyor. Demem tahminken; meteorolojinin bunu demesi yordama olur. Çünkü meteoroloji bunu elindeki ölçümlere göre yapıyor.

Ölçme araçlarının çoğunda bireylerin gelecekteki davranışlarının kestirilmesi söz konusudur. Yordama geçerliliği, ölçme aracının bu kestirim işini ne ölçüde doğru ve isabetli yaptığı ile ilgili bir sorgulamadır.

Özellikle **seçme amaçlı** (KPSS, LGS vb.)ya da **yönlendirme amaçlı** (yetenek testleri) uygulanan ölçme araçlarının yordama geçerliliğinin yüksek olması istenir.

Yordama geçerliliğinde, ölçme aracından elde edilen puanlar, ölçme aracının tahmin ettiği puanla (ölçüt puan) karşılaştırılır ve tahminin ne ölçüde doğru olduğu belirlenmeye çalışılır. **Ölçüt puan**, testin tahmin etmeye çalıştığı özelliktir.

Yordama geçerliliğinde ölçüt puan gelecekte belli olacağından beklemek gerekmektedir. Yordama geçerliliğinde geçerlilik katsayısının +1 yönünde olması beklenir.

	KPSS Puanları	İş Performansı (Ölçüt Puan)
Uğur	87.32	78
Dilara	89.24	90
Hakan	75.09	87
Arzu	93.46	88
Eda	78.87	76
.....	90.29	95
.....	77.76	98
.....	87.33	87
.....	76.77	80
Engin	69.55	78

Ölçüt gelecekte belli olacağından beklemek gerekmektedir.

Korelasyon (r)
1'e yakın

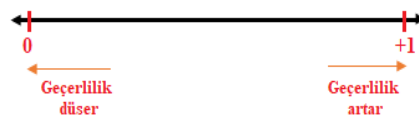
Geçerlilik Katsayısı

YORUM: Öğretmen adayları, KPSS'den belirli puan almışlardır. KPSS yüksek not alanın öğretmen olarak atandığında yetkinliğinin (iş performansı) yüksek, düşük not alanın da yetkinliğinin (iş performansı)düşük olacağı kurgusuna sahiptir.(YORDAMA)

Sınavdan sonra tercih yapıp atandılar ve göreve başladılar.

Atananların iş performansları ölçülüyor.(Ölçüt Puan)
Kişilerden eğer yüksek not alan yüksek ve düşük not lan düşük performans gösterdiyse; yani korelasyon +1'e yakınsa, **KPSS YORDAMAYI İYİ BİÇİMDE YAPIYORDUR.**

KPSS Puanları ile İş Performansı puanları arasındaki korelasyondan (ilişki) yordama geçerliliği katsayısı belirlenir.Buradaki korelasyon;



C. ZAMANDAŞ GEÇERLİLİĞİ:

Zamandaş geçerlilik, sorgulamalarında ölçüt puan eş zamanlı olarak elde edilebilir.

Geliştirilen ölçme aracı ile ilişkili olabilecek nitelikleri ölçen ve daha önceden geçerliliği yüksek olduğu bilinen ölçüt araç aynı zamanda uygulanıp arasındaki ilişki incelenir. İ

	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

Daha önce geliştirilmiş ve geçerliliğinin yüksek olduğu bilinen araç

Korelasyon (r) 1'e yakın

Geçerlilik Katsayısı

Öğrencilere Ben kimim? sorusuna cevap verdikleri Benlik Algısı Testi uygulanmış ve notlar alınmış olsun. Daha önceden geliştirilmiş ve geçerliliğinin yüksek olduğunu bildiğim Öz-Saygı Ölçeği(Ölçüt Puanı)ile aralarındaki korelasyona (ilişki) bakıyorum.

YORUM: Normalde Öz-saygısı yüksek bireylerin de benlik algısının yüksek olmasını yani korelasyonun +1 olmasını beklerim. Eğer böyleyse, Öz-Saygı Ölçeğinin geçerliliğinin YÜKSEK olduğunu biliyorum. O halde uyguladığım Benlik Algısı Ölçeğim de geçerlidir. derim.

Zamandaş geçerliliğinde geçerlilik katsayısının -1 ile 1 arasında değişmesi beklenir. Ölçüte bağlı olarak geçerlilik katsayısı ∓ 1 'e yaklaştıkça artar, 0'a yaklaştıkça düşer.



Not: Sınavlar şuan + şuan olabilir veya şuan+ önce(yapılmış veri elimde)

D. YAPI GEÇERLİLİĞİ:

Ölçme aracı geliştirmek için önce, söz konusu olan özellik ya da özellikleri belirlemek yani yapıyı tanımlamak; daha sonra, tanımlanmış yapıdan sınanabilir denenceler çıkarmak ve çıkarılan denenceleri sınamak için deneysel ve istatistiksel çalışmalar yapmak gerekmektedir. Yapı, birbirleriyle ilgili olduğu düşünülen belli öğelerin ya da öğeler arasındaki ilişkilerin oluşturduğu bir örüntüdür.

Yapı, aslında genellikle gözleyemediğimiz, 5 duyu organımızla algılayamadığımız psikolojik özelliklere (başarı, zeka, kişilik, tutum, ilgi vb.) karşılık gelir.

Yapı geçerliliği, bir testin dayandığı kuramsal temelleri ne derece iyi örneklediğiyle ilgilidir.

Örneğin, bir kişi, geleneksel aile biçimi ile çocuk yetiştirme biçimi arasındaki ilişkiyi incelemek amacıyla bir ölçek geliştirip bu ölçeğin yapı geçerliliğini ortaya koymak istediğinde geleneksel aile yapısı ve çocuk yetiştirme kavramlarının ne anlamlara geldiğini, ölçme aracındaki maddelerin bu anlamlara uygunluğunu araştırarak yapı geçerliliği konusunda karar verebilir.

Mustafa Uçan

GÜVENİLİRLİK

Güvenilir bir ölçme aracı, aynı özelliklerle ilgili olarak arka arkaya yapılan ölçmelerde yaklaşık olarak aynı sayısal sonucu vermelidir. Bir testin ölçmek istediği özelliği ölçebilmesi için o testin söz konusu olan özelliği kararlı, tutarlı, duyarlı olarak ölçmesi gerekir.

Örneğin, Murat, üç dakikalık bir sözcük çalışmasında, dakikada 54 sözcük yazmıştır. Bu puan Murat'ın becerisini ne ölçüde yansıtmaktadır. Varsayalım ki dakikada 50 sözcük yazılması yeterli kabul edilmektedir. Murat bu düzeyin gerçekten üzerinde midir? Murat geçen hafta dakikada 45 sözcük yazmıştır. Murat'ın bugünkü puanı olan 54, onun kendisini geliştirdiğinin göstergesi midir ya da bir değişim dalgalanması mıdır?

İki ölçüm arasındaki uyumsuzluğun birçok nedeni olabilir. Bir hareketten diğerine, "dikkat ve çaba" değişebilir. Özellikle uzun zamanlarda puan değişmesi, fiziksel büyüme, öğrenme ya da sağlık ve kişilikteki değişimlerden; kaynaklanabilir. Yine ikinci ölçümde daha açık soruların kullanılması diğer bir faktör olabilir. İki ölçüm arasındaki puan farklarını yorumlayabilmek için gerçek puan kuramına bakalım.

Örneğin, yarışlara hazırlanan bir koşucu, bir mesafeyi farklı zamanlarda 23.7, 24.0, 24.2, 25.1, 25.2 saniyelerde koşmuş olsun. Bu ölçümlerin ortalaması 24.7 ise gerçek puanı 24.7'dir. Koşucu bu puana daha önce 23.7 saniyede koştuğu ölçümü göstererek itiraz edebilir. Bu durumda koşucuya 23.7 ve 25.2 saniyede koştuğu durumlara birçok faktörün etkisinin olabileceği ve 23.7 saniyede tekrar koşmasının belki de hiç olanaklı olamayacağını, gerçek puanının 24.7 olduğunu söyleyebiliriz. Bu söylem gerçek puan kuramına dayanır.

$$X = T + E$$

X = Bireylerin ölçme aracından elde ettiği gözlenen puanı

T = Bireylerin gözlenemeyen gerçek puanı

E = Ölçmeye karışan hata miktarı

Bir ölçme işleminde hata miktarı sıfır ise eşitlik $X = T + 0$ ve dolayısıyla $X = T$ olmuş olur.

Ölçme işleminde hata miktarı arttıkça gözlenen puanın gerçek puana olan farkı artmaya başlar. Ölçmenin standart hatası formülü ile bulunur.

$$S_h = s\sqrt{1 - r_x}$$

$S_h =$ Ölçmenin standart hatası

$s =$ Standart sapma

$r_x =$ Güvenilirlik katsayısı

Ölçmenin standart hatası $S_h = s\sqrt{1 - r_x}$ Bir ölçme işleminde $s=8$ ve $r_x=75$ ise formülü ile bulunur.

$S_h = 8\sqrt{1 - 0.75}$
 $S_h = 8\sqrt{0.25}$
 $S_h = 8 \times 0.5$
 $S_h = 4$

$S_h =$ Ölçmenin standart hatası **$Sh=E$**
 $s =$ Standart sapma
 $r_x =$ Güvenilirlik katsayısı **0-1 arasında**. Bu sonuç bireylerin puanlarına ± 4 puan ölçmenin standart hatası nasıl düşük çıkar? hata karıştığı anlamına mı gelir?

%68 olasılık için bireyin puanından bir standart hata çıkarılır ve puanına eklenirken %95 olasılık için bireyin puanından iki standart hata çıkarılır ve puanına eklenir, son olarak %99 olasılık için bireyin puanından üç standart hata çıkarılır ve puanına eklenir.

%68 olasılık için $X \mp 1 \times Sh$

%95 olasılık için $X \mp 2 \times Sh$

%99 olasılık için $X \mp 3 \times Sh$

Bir kişi 100 üstünden 50 puan almış ve $Sh=4$ olsun.

Yorum: Ben 50 alan bu öğrenciye sonsuz kere ölçme yaptığımda gerçek notu:

%68 olasılık: $X \mp 1 \times Sh \Rightarrow 50 \mp 1 \times 4 =$ Bireyin gerçek puanı %68 olasılıkla 46-54 arasında değişir.

%95 olasılık için $X \mp 2 \times Sh \Rightarrow 50 \mp 2 \times 4 =$ Bireyin gerçek puanı %95 olasılıkla 42-58 arasında değişir.

%99 olasılık için $X \mp 3 \times Sh \Rightarrow 50 \mp 3 \times 4 =$ Bireyin gerçek puanı %99 olasılıkla 38-62 arasında değişir.

Güvenilirlik katsayısı 0 ile 1 arasında değişir. Güvenilirlik katsayısı için genellikle 0.70 ve üstü ölçütü kabul edilir.



Mustafa Uçan

GÜVENİLİRLİK KESTİRİM YÖNTEMLERİ

Birden Çok Uygulamaya Dayalı Yöntemler

Test-Tekrar Test

Eşdeğer (Paralel) Testler

Bir test, aynı gruba, belli bir zaman aralığıyla iki kez uygulanır. İki uygulama sonundaki puanlar arasındaki korelasyon hesaplanır. Elde edilen korelasyon katsayısına **kararlılık (devamlılık / istikrarlılık)** katsayısı adı verilir.

	Ankışete	
	Ölçü Puanları	Ölçü Puanları
	X ₁	X ₂
Canan	54	50
Filiz	98	10
Aydn	128	120
Erik	89	80
Deniz	101	101
.....	87	85
.....	98	106
.....	67	59
.....	122	19
Özcan	79	82

Kararlılık Katsayısı

Korelasyon (r)
1'e yakın

Güvenilirliğin kararlılık, tutarlılık ve duyarlılık sorunu olduğuna daha önce değinilmişti. Bu yöntemle güvenilirliğin yalnızca kararlılık boyutuna karşılık gelir; tutarlılığa ve duyarlılığa ilişkin değildir.

Bu yöntem daha çok iki uygulama arasında kolaylıkla değişmeyen özellikleri ölçen testler için uygundur.

Örneğin genel zihin yetenekleri hemen değişen bir özellik değildir.

Önemli bir sorun, testin iki uygulaması arasındaki zaman aralığının ne kadar olması gerektiğidir. "Birinci uygulamada maddeleri hatırlamayacak kadar uzun, ölçülen özellikte köklü değişimler oluşmayacak kadar kısa olmalıdır."

Test Yarılama

Bir test iki eşdeğer yarıya bölünür ve bireylerin iki yarıdan aldıkları puanlar arasındaki **tutarlılık** incelenir.

Örneğin; 30 soruluk bir test olsun.

- 1-15 sorular, 16-30 sorular
- tek sorular I.yarı, çift sorular II.yarı
- Rastlantısal olarak kura ile 15-15 seçilebilir.

İlk ve son yarı yöntemi her test için uygun değildir.

Testteki,

- Maddeler basitten zora doğru sıralanmış** ise kolay sorular 1-15, zorlar 16-30 arasında olur.
- Maddeler konu içeriklerine ya da faktörlere göre kümelenecek yerleştirilmiş** ise 4 ünite işlemiş olayım. 1-15 arası 1. ve 2. ünite, 16-30 3. ve 4. ünitelerden olabilir.
- Madde sayısı çok fazla** ise yorgunluk, ağrı gibi tesadüfi hata olasılığı artacak ve güvenilirlik azalacaktır.

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

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

Korelasyon (r)

Spearman-Brown Eşitliği

İç tutarlılık katsayısı

Tek Uygulamaya Dayalı Yöntemler

KR-20

KR-21

Cronbach Alfa

Kuder-Richarson 20-21 veya KR-20 ve KR-21, testin kendi içinde tutarlı bir bütün oluşturup oluşturmadığı ile ilgili bilgi verir. Bu nedenle her iki yöntemle de testin iç tutarlılığı değerlendirildiğinden bu yöntemlerden elde edilen katsayıya **İç tutarlılık katsayısı** adı verilir.

KR-20 ve KR-21 ile güvenilirlik kestirimi, puanlamamın kategorik olduğu ölçme araçları için uygundur.

Doğru yanıtlara 1 puan, yanlış ve boş bırakılan maddelere 0 puan vererek puanlamamın yapıldığı ya da maddenin yanıtının iki seçeneğe verildiği durumlarıdır. (Çoktan seçmeli, Doğru-Yanlış testleri). Soru 5 puansa, doğru yanıt 5 boş ve yanlış 0 puan alır.

Eğer testteki maddeler farklı ağırlıklarla puanlanmışsa ya da test puanları şans başarısı için

Eğer ölçme aracının puanlaması çok kategorili ise KR-20 ile aynı mantık üzerine kurulu Cronbach alfa hesaplanır.

Elde edilen katsayının ismi yine **İç tutarlılık katsayısıdır.**

Çok kategori demek en az 3 ve üzeri derece olması demektir.

Derecelendirme ölçekleri puanlamamın çok kategorili olduğu araçlara örnek verilebilir.



Mustafa Uçan

NOT: Tek bir test, tek bir grup ve tek bir uygulama gerektirmektedir. Dolayısıyla güvenilirlik kanıtı elde etmek daha pratiktir. Bir test bir gruba, bir kez uygulanıyorsa test kendi içinde değerlendirilecek demektir. Bu da **İÇ TUTARLILIK KATSAYISIYLA** ilgilidir.

GÜVENİLİRLİĞİ VE GEÇERLİLİĞİ ARTIRMA YOLLARI

Güvenilirlik temelde tesadüfi hatalardan arınık olma durumu ise şüphesiz ki **tesadüfi hata kaynaklarını azaltmaya** yönelik eylemler **güvenilirliği artıracaktır**.

1. Bir testteki **madde sayısı arttıkça** birimler küçüldüğünden duyarlılık artar, bu nedenle hata miktarı azalacağından güvenilirlik artar. *Örneğin, 1 soru ile benim Türkçe başarıyı ölçemezsiniz. 10 soruda daha iyi, 100 soruda daha da iyi ölçersiniz.*

2. Bir ölçme işleminde genel bir ilke olarak **puanlayıcı sayısı arttıkça** güvenilirlik artar. Birden çok kişinin yaptığı puanlamanın ortalamasını almak, bir kişinin verdiği puandan daha güvenilir olma eğilimindedir.

3. **Puanlama nesneliği**, 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. Elbette kullanılan ölçme aracının puanlamasının nesnel ve öznel olması bu noktada çok önemlidir. Çoktan seçmeli, derecelendirme ölçeği, kontrol listeleri gibi araçlar puanlaması nesnel araçlara örnek iken açık uçlu sorulardan oluşan araçlar, mülakatlar vb. ise puanlaması öznel araçlara örnek olarak verilebilir. Öznel araçların puanlama biçiminin mümkün olduğunca standartlaştırılarak nesnel hâle getirilmesi güvenilirlik açısından önemlidir.

4. KR-20 ve KR-21 yöntemlerinde bahsedildiği gibi testteki maddeler açısından **benzeşiklik (homojenlik) arttıkça** güvenilirliğin artması, diğer taraftan ayrışıklık (heterojenlik) arttıkça güvenilirliğin düşmesi beklenen bir durumdur.

5. Testten elde edilen puanların güvenilirlik kestirimi için veri elde edilecek **grubun (örneklem) büyüklüğü arttıkça grubun heterojenleşmesinden dolayı** güvenilirlik artar. Herkes ayrı olsun ve herkesin puanını ölçeyim. Mesela baskül herkesi 50 kg ölçse sıkıntı var deriz. Ama ne kadar farklı ölçerse o kadar iyi. Bunun için de ne kadar çok öğrenci o kadar ayırtedici olur.

6. Maddelerin **dil bilgisi kurallarına uygun**, açık ve anlaşılır yazılması belki de güvenilirliği en çok artıran etmenlerden biridir.

7. Öğrencilere test uygulaması öncesinde ve gerekliyse süreçte yönerge vermek, test almaya **güdülenme ve hazırbulunuşluğun** artması açısından önemlidir. Buna yönelik alınacak önlemler özellikle bireyden kaynaklanabilecek hataları asgariye indirir.

8. **Fiziksel ortamın** ses, sıcaklık, ışık, koku, dikkat dağıtıcı uyarıcılar, ergonomi vb. etmenler açısından **uygun hâle getirilmesi** yine güvenilirliği artıran bir başka etmendir.

9. Eğer süreli bir test uygulaması ise **sürenin yeterli verilmesi** gerekir. Eğer bir hız testi değilse tüm katılımcıların maddelerin tümüne ulaşması beklenen bir durumdur. Tabii ki hız testleri bu durumdan bağımsız olarak değerlendirilebilir.

10. Uygulamada bireylerin **dikkatini dağıtacak ve/veya kaygısını artıracak davranışlardan kaçınılması**, puanlama ya da veri girişi yapılırken dikkatli ve titiz davranılması ölçme işlemi yapan kişiden kaynaklanabilecek hataları azaltacağından dolayı güvenilirliği artırır.

11. **Maddelerin teste düzgün yerleştirilmesi**, okumayı güçleştirecek bir unsurun bulunmaması; test katılımcısının yaşına, gelişim düzeyine uygun bir punto büyüklüğü kullanılması; baskı hataları bulunmaması vb. etmenler güvenilirliği artırır.

Güvenilirliği artıran faktörler geçerliliği de artırır.

KURALIMIZ
TESADÜFİ HATAYI AZALT GÜVENİLİRLİK ARTSIN.
TESADÜFİ, SİSTEMATİK VE SABİT HATAYI AZALT GEÇERLİLİK ARTSIN.

Mustafa Uçan

TEST GELİŞTİRME VE MADDE / SORU TÜRLERİ

Test, eğitimde bireylerin özelliklerini belirlemeye yönelik ölçme araçlarına verilen genel bir kavramdır.

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

TEST GELİŞTİRME: Test geliştirmenin işlem 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.

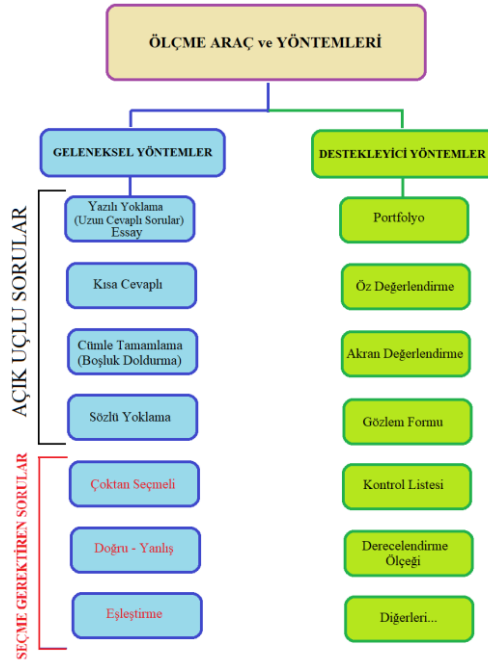
Öğretmenler açısından sınıf içi ölçme ve değerlendirmelerde uygulanması daha olanaklı bir test planı önerisi

- Sınavın amacının belirlenmesi (puanların hangi amaçlarla kullanılacağı)
- Sınavda yoklanacak davranışların belirlenmesi
- Sınavın kapsayacağı konuların listelenmesi
- Okulun takviminin ve sınav sonuçlarının kullanılacağı zamanın dikkate alındığı bir sınav gününün belirlenmesi
- Bir ders saatine uygulanabilecek uzunlukta bir sınav süresinin belirlenmesi,
- Belirtke tablosunun hazırlanması
- Soru/madde türlerinin belirlenmesi [Soru/madde türleri sınavda kullanılacak davranışlara göre belirlenir. Bir sınavda tek bir soru/madde türü kullanılabileceği gibi birden fazla da kullanılabilir (soru/madde çeşitlemesi).]
- Soru/madde sayısının belirlenmesi
- Sınav süresinin belirlenmesi
- Sınavda kullanılacak soruların/maddelerin “ortalama güçlüğü” ve “güçlük dağılımının” belirlenmesi
- Soruların/maddelerin yazımında, redaksiyonunda ve teste alınacak soruların/maddelerin seçiminde izlenecek yolun belirlenmesi
- Cevap anahtarının ve puanlama yönteminin belirlenmesi
- Ölçme aracını yazma ve çoğaltma yönteminin belirlenmesi
- Sınavın uygulanma kurallarının belirlenmesi (yönerge yazılması)
- Sınavın uygulanması
- Test ve madde istatistiklerinin hesaplanması.

SORU: Cevaplayıcının (öğrenci) cevap özgürlüğünün olduğu format varsa soru diyoruz.

MADDE: Cevaplayıcının (öğrencinin) sadece seçeneklerden seçme hakkı varsa madde diyoruz.

Mustafa Uçan



1. DOĞRU-YANLIŞ MADDELERİ: Cevaplayıcının verilen ifadelerin doğru mu, yanlış mı olduğunu belirlemesinin istendiği madde türüdür.

Çoktan seçmeli madde türünden sonra **en çok kullanılan ikinci madde tipidir**. Genellikle öğrencinin bilimsel gerçekleri, tarihi olayları, kesin yargıları tanıması ve hatırlaması; olguları ve gerçekleri kişisel yargı ve görüşlerden ayırt etmesi ölçülmek istendiğinde kullanışlıdır.

<u>Avantajları</u>	<u>Dezavantajları</u>
<p>a. Hazırlaması, uygulaması ve puanlaması kolay kullanışlıdır.</p> <p>b. Soru sayısı artırılabilir.</p> <p>c. Sistemik hata karışma olasılığı pek yoktur.</p>	<p>a. Şans başarısı olasılığı yüksektir.</p> <p>b. İyi yazılmazsa genellikle üst düzey bilişsel özellikleri ölçemez.</p> <p>c. Öğrenme eksiklerini belirleyemez.</p> <p>d. Yanlış üzerinden öğretim yapılmaz.</p> <p>e. Her derste/konuda kesin yanlış denebilecek durumlar belirlemek zordur.</p>

Doğru-Yanlış Maddeleri Yazım Kuralları:

• Her madde tek ve belirli bir fikri belirtmelidir. Özellikle aynı maddede biri doğru, öteki yanlış iki fikir ifade etmekten kaçınılmalıdır.

Yanlış örnek: Atatürk,1938 yılında ölmüş ve aynı yıl Anıtkabir'e defnedilmiştir. (D) (Y)

Doğru Örnek: Atatürk,1938 yılında ölmüştür. (D) (Y)

Atatürk,1938 yılında Anıtkabir'e defnedilmiştir. (D) (Y)

• Doğru yanlış maddesi, kesinlikle doğru ya da kesinlikle yanlış olmalıdır.

Yanlış örnek: Bugüne kadar Türkiye Cumhuriyeti'nin üç beş cumhurbaşkanı olmuştur. (D) (Y)

Doğru Örnek: Bugüne kadar Türkiye Cumhuriyeti'nin on iki cumhurbaşkanı olmuştur. (D) (Y)

• Bir maddenin yanlışlığı önemsiz bir ayrıntıda ya da aldatıcı bir noktada olmamalıdır.

Yanlış örnek: Fransız İhtilali 1798 yılında olmuştur. (D) (Y)

Doğru Örnek: Fransız ihtilali 1789 yılında olmuştur. (D) (Y)

• Mümkün olduğunca olumsuz ifade kullanılmamalıdır. Özellikle iki olumsuz ifadeden kesinlikle kaçınılmalıdır.

Yanlış örnek: Atatürk, ülkenin kurtuluşunu milletin mücadelesinde görmemiş değildir. (D) (Y)

Doğru Örnek: Atatürk, ülkenin kurtuluşunu milletin mücadelesinde görmüştür. (D) (Y)

• Kanı ifadeleri belli bir kaynağa ya da bir otoriteye dayandırılıp verilmelidir.

Yanlış örnek: Bilgi deneyle başlar; ancak deneyle sona ermez. (D) (Y) (Kişiden kişiye değişen kanıdır)

Doğru Örnek: Immanuel Kant'a göre bilgi deneyle başlar; ancak deneyle sona ermez.(D) (Y)

• Bir maddenin ifadesi kısa, açık ve yalın olmalıdır. Madde, ayrıntılarla şişirilerek uzatılmamalıdır.

Yanlış örnek: Millattan önce 2000 yıllarında Nil Vadisi'ne yerleşen halk birçok tanrıya taptı; onların en önemlileri iyilik tanrısı Osiris ile kötülük tanrısı Seth idi. (D) (Y) (Bir sürü bilgi ile şişirilmiş.)

Doğru Örnek: Eski Mısırlılar politeist idi. (D) (Y)

• Bir metne bağlı olarak yanıtlanan doğru-yanlış maddelerinin anlam geliştirme, yorumlama, değerlendirme gibi daha üst düzey okuduğunu anlama süreçlerine yönelik olabilmesi için, metindeki cümleler, seçeneklerde **aynen yazılmamalıdır.**

• Öğrencinin bir maddede ifadenin doğruluğuna ya da yanlışlığına karar vermesi gerektiğinde öğretmen özellikle vurgulamak istediği bir sözcük / terim vs. varsa onun altını çizebilir.

Doğru Örnek: Yutak, sindirim sisteminin yanı sıra solunum sisteminde de yer alır. (D) (Y)

• Doğru-yanlış maddesinin özellikle **yanlış olduğu durumlarda**, öğrencinin şansa puan almasını engellemek amacıyla ifadeyi düzeltmesi ya da yanlışlığın ne olduğunu belirtmesi

Doğru Örnek: Kızılçam, soğuk iklimler için uygun bir çam türüdür. (D) (Y)
Öğrenci: Kızılçam, ılıman iklimde yetişmeye uygun bir çam türüdür.

• İyi kurgulandığı takdirde doğru-yanlış maddeleri ile üst düzey öğrenmeler de yoklanabilir. **Örneğin**, bir grafik verilir. Grafikle ilgili yorum cümleleri (D) (Y) şeklinde sorulabilir.

• Doğru-yanlış maddeleri aynı kapsamdan geliyorsa ortak bir soru kökü altında birleştirilebilir.

Doğru örnek

Çoktan seçmeli testlerle ilgili ifadeleri dikkatle okuyunuz. Eğer ifade çoktan seçmeli testlerin bir özelliği ise D'yi, değilse Y'yi daire içine alınız.

1. Puanlama güvenilirliği yüksektir.	(D) (Y)
2. Geliştirmesi kolaydır.	(D) (Y)
3. büyük bir örnekleme hatasına sahiptir.	(D) (Y)
4. Puanlaması az zaman alır.	(D) (Y)

• Doğru ve yanlış maddeler, teste belirli bir örüntü ile yerleştirilmemeli.

Yanlış örnek: D-D-Y-Y-D-D2 Doğru-2 Yanlış örüntü. Öğrenci keşfederse okumadan hepsini yapar.

Doğru Örnek: D-Y-DD-YYY-D-D-Y-Y

• Doğru ve yanlış maddelerin toplam sayıları birbirine yakın veya dengeli olmalıdır.

Doğru Örnek: D-Y-DD-YYY-D-D-Y-Y 5 Doğru -6 Yanlış cevapla dengeli dağılmış.

• Maddeleri işaretleme yöntemi, öğrenciye açık ve anlaşılır şekilde açıklanmalıdır.

• Mümkün olduğunca kısa yazılmalı, basit cümle yapıları tercih edilmeli, bileşik yapı cümlelerden kaçınılmalı, özellikle **ve** gibi bağlaçlara dikkat edilmelidir. Çünkü ve bağlacının olduğu cümlelerde birden çok yargı olma olasılığı yüksektir.

• Olumsuz anlama sahip bir sözcük ya da ifade varsa altı çizilmelidir.

• İpucu vermekten kaçınılmalı. Özellikle **asla, daima, tümü, hiçbiri** vb. ifadelerin **yanlış olma olasılığı**, öte yandan **genellikle, çoğunlukla, bazen, zaman zaman, sıklıkla** vb. ifadelerin **doğru olma olasılığı** yüksektir.

• Üst düzey becerileri ölçmek için; tablolar, grafikler, haritalar, veya okuma parçaları bir öncül olarak kullanılabilir.

Mustafa Uçan

2.EŞLEŞTİRME MADDELERİ: Eşleştirme maddeleri, iki grup hâlinde verilen ve birbirleriyle ilgili olan bilgi öğelerinin belli bir açıklamaya göre eşleştirilmesini gerektirir.

Yönerge / Açıklama

Yönerge: Aşağıdaki "A" sütununda Türkiye Cumhuriyeti tarihi ile ilgili bazı önemli olaylar, "B" sütununda ise bu olayların geçtiği yıllar verilmiştir. Bir olayın hangi yılda geçtiğini kararlaştırdıktan sonra, o yılın önündeki harfi, olayın solundaki boşluğa yazınız. "B" sütunundaki bir tarih, bir veya birden fazla kez kullanılabileceği gibi, hiç kullanılmayabilir.

Öncüller / Maddeler

A Sütunu

C 1. Osmanlı Meclis-i Mebusan'ın son toplanması
D 2. Ankara'nın başkent olması
H 3. Tek dereceli seçimlerin ilk kez yapılması
C 4. TBMM'nin açılması
D 5. Lozan Antlaşmasının imzalanması
D 6. Cumhuriyet'in ilanı
C 7. Sevr Antlaşmasının imzalanması

B Sütunu

A. 1918
B. 1919
C. 1923
D. 1925
E. 1928
F. 1930
G. 1934
H. 1936
I. 1946

Muhtemel cevaplar / Seçenekler

<u>Avantajları</u>	<u>Dezavantajları</u>
a. Hazırlaması, uygulaması ve puanlaması kolay, kullanışlıdır. b. Soru sayısı artırılabilir. c. Sistematiik hata karışma olasılığı pek yoktur.	a. Şans başarısı olasılığı yüksektir. b. İyi yazılmazsa genellikle üst düzey bilişsel özellikleri ölçemez.

Eşleştirme Maddeleri Yazım Kuralları:

• Her eşleştirme soru grubunda yönerge/açıklama yazılmalıdır. Yönergede eşleştirmeye konu olan kapsam belirtilmeli, öncüllere ve seçeneklere nelerin konulduğuna değinilmelidir. Eşleştirmenin nasıl yapılacağı konusunda bilgi verilmelidir.

Yönerge: Aşağıdaki "A" sütununda Türkiye Cumhuriyeti tarihi ile ilgili bazı önemli olaylar, "B" sütununda ise bu olayların geçtiği yıllar verilmiştir. Bir olayın hangi yılda geçtiğini kararlaştırdıktan sonra, o yılın önündeki harfi, olayın solundaki boşluğa yazınız. "B" sütunundaki bir tarih, bir veya birden fazla kez kullanılabileceği gibi, hiç kullanılmayabilir.

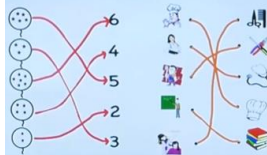
• Seçeneklerin kullanılıp kullanılmayacağı, birden fazla kez kullanma durumu belirtilmelidir. Bir eşleşme takımındaki öncüller ile cevaplar eşit sayıda olmamalıdır. Eğer eşit sayıda olursa öğrenci, hemen ilk bakışta bildiklerini eşleştirecek; geri kalan öncüllerin cevaplarını da tahmin edecektir.

Yönerge: Aşağıdaki "A" sütununda Türkiye Cumhuriyeti tarihi ile ilgili bazı önemli olaylar, "B" sütununda ise bu olayların geçtiği yıllar verilmiştir. Bir olayın hangi yılda geçtiğini kararlaştırdıktan sonra, o yılın önündeki harfi, olayın solundaki boşluğa yazınız. "B" sütunundaki bir tarih, bir veya birden fazla kez kullanılabileceği gibi, hiç kullanılmayabilir.

• Eşleştirmede ortaokul ve üstü öğretim kademelerinde öncülün başında bir boşluk bırakılarak öğrenciden seçeneğin harfini yazmaları istenebilir. İlkokul düzeyinde ise öğrencilerden öncüllerle seçenekler arasında oklarla eşleştirme yapmaları istenebilir.

A Sütunu	B Sütunu
1. $18 \times 9 + 6$	9
2. $44 + 4 - 9$	5
3. $35 + (5 + 2)$	2
.....	168
.....

• Okul öncesinde, özel eğitimde eşleştirme şekillerle ya da resimlerle / fotoğraflarla yapılabilir.



• Bir eşleştirme maddesinde yer alan öncüller ile seçeneklerin her biri benzeşik (homojen) öğelerden oluşmalıdır. Diğer bir ifadeyle aynı kapsamdan gelmelidir.

A Sütunu	B Sütunu
C 1. Çok buğday yetişir.	A Atina
B 2. Çok kahve yetişir.	B Belgrat
E 3. Avusturya'nın başkentidir.	C Çin Halk Cumhuriyeti
A 4. Yunanistan'ın başkentidir.	D Kanada
D 5. Çok pirinç yetişir.	E Viyana
	F Venezuela
	G Brezilya

Başkentler ve yetişen bitkiler ayrı sorularda sorulmalıydı.

Mustafa Uçan

• Uzun ifadeler öncül olarak kullanılmalı ve öncüller sütunu sayfanın sol sütununa, cevaplar sütunu ise sayfanın sağ tarafına yerleştirilmelidir. Böyle olması, cevaplama zaman kaybını önler ve cevabın seçimini kolaylaştırır.

• Bir eşleştirmeli maddeler grubundaki madde sayısı, en az 6, en çok 15 olmalıdır. Madde sayısının 6'dan az olması, salt tahminle doğru cevabın bulunma olasılığını artırır. 15'ten çok madde kullanıldığında ise cevaplayıcıların doğru olarak eşleştirilecek ifadeleri ayıklayıp seçmesi çok zaman alır. Üstelik madde sayısı arttıkça öncüller ile cevaplar takımının benzeşikliğini sağlamak giderek güçleşir.

• Maddelerin tümü aynı sayfada bulunmalıdır. Aksi hâlde cevabın seçimini güçleştirir.

• Cevapların seçileceği sütun, bir kelime listesi ise alfabetik sıraya göre (küçükten büyüğe ya da büyükten küçüğe) düzenlenmelidir.

• Bu eşleştirme maddeleri öğrencilerin yalnızca hatırlama düzeyi dışında, kavrama düzeyindeki öğrenmelerinin de yoklanabileceğine ilişkin örnektir. **Örneğin, grafik / tablo / metin / harita**

3.ÇOKTAN SEÇMELİ MADDELER:

Soru 24 BÖLGELER ARASI BAĞLANTILAR

	A	B	C	D
A	0	0	1	1
B	0	0	1	1
C	1	1	0	1
D	1	1	1	0

Öncül

Yukarıda verilen tablo aşağıdaki şekillerden hangisine ait olabilir?

Doğru yanıt

A) B) C) D)

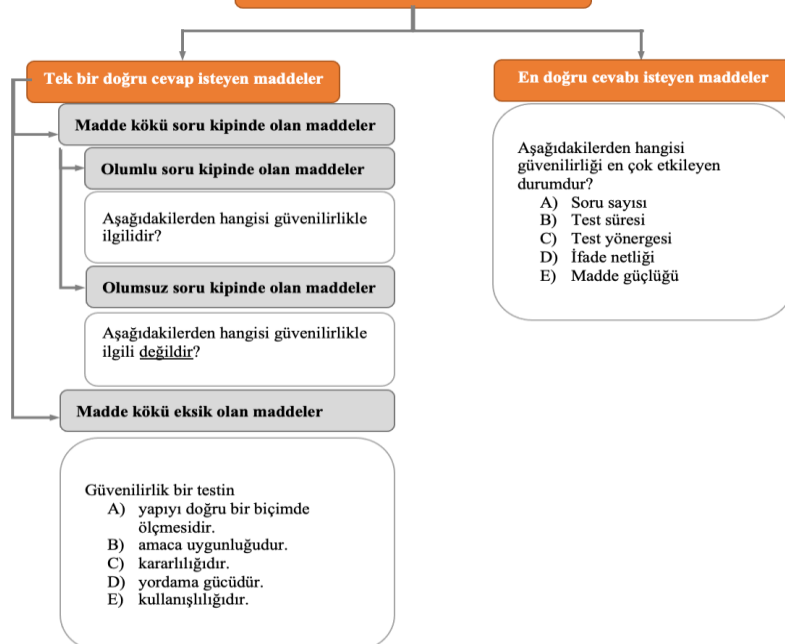
Seçenekler

Çoktan seçmeli madde

Çoktan Seçmeli Maddelerin Özellikleri:

- Hazırlanması zor ancak puanlaması kolaydır.
- Puanlama objektiftir.
- Üst düzey bilişsel becerilerle ilgili madde yazmak zordur. O yüzden yazılan maddeler genellikle bilgi, kavrama ve uygulama basamaklarındadır.
- Doğru cevabı içinde barındırması nedeniyle sentez düzeyinde soru sormak olanaklı değildir. Bu nedenle bir test yalnızca çoktan seçmeli maddelerden oluşuyorsa yaratıcılığın denetlenmediği yönünde haklı bir eleştiri vardır.
- Doğru cevabı içinde barındırması nedeniyle puanlara şans başarısı karşısına ihtimali vardır.
- Çok sayıda soru sorulabilir.
- Sistematik hata karşısına ihtimali çok düşüktür.
- Çeldirme mantığı temellidir.

ÇOKTAN SEÇMELİ MADDE TÜRLERİ



Mustafa Uçan

Çoktan Seçmeli Madde Yazım Kuralları:

- İdeal çoktan seçmeli maddede, bilen bir öğrenci, seçenekleri görmesine gerek kalmaksızın doğru cevabı verebilmelidir.

Yanlış örnek	Doğru örnek
Analizler sonucunda bir testin geçerliliğinin düşük olduğu belirlenmiştir. Bu durumun nedeni aşağıdakilerden hangisi olabilir? Birçok neden olabilir. A) Soru sayısı azdır. B) Testten elde edilen puanların güvenilirliği düşüktür. C) Öğretme konu olan kapsam, sınava yeterince yansıtılmamıştır. D) Test hazırlama aşamalarına uyulmamıştır. E) Testin kullanılabilirliği düşüktür.	Öğretme konu olan ünitelerdeki kazanımları, yeterli ve dengeli bir biçimde ölçmeyen bir testin geçerliliğinin düşük olduğu belirlenmiştir. Yukarıda bahsedilen durum, hangi geçerlilik türü bağlamında değerlendirilir? A) Yordama geçerliliği B) Yapı geçerliliği C) Zamandaş geçerlilik D) Görünüş geçerliliği E) Kapsam geçerliliği


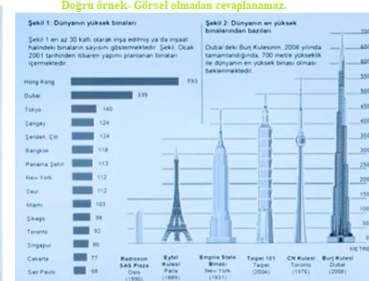
- Öğrencinin sözcük repertuarında bulunmayan, henüz öğrenmediği sözcük, kavram, terim vb. kullanılmamalıdır.

Yanlış örnek
Bir testte madde sayısının az olması nedeniyle güvenilirlik katsayısı düşüktür. Bu nedenle teste yeni maddeler eklenmiştir. Ayrıca yüksek olarak bulunan standart hata da düşürülmek istenmiş, bunun için başarı düzeyi açısından homojen bir grup üzerinde pilot uygulama yapılmıştır. İşlemediğin yeri sorma Pilot uygulama sonucunda aşağıdaki durumlardan hangisinin olması beklenir?

- Bir durumun resim ya da fotoğrafla anlatılabildiği durumda, çokça cümle kurmak yerine görsellerden yararlanılabilir. Ancak maddenin cinsiyet gibi alt gruplarda yanlılık oluşturmamasına dikkat edilmelidir.

Doğru örnek	Yanlış örnek
 <p>Derya arkadaşlarıyla yapacağı basketbol maçı öncesi yukarıda şekli verilen evlerin bulunduğu sitenin çevresinde ısınma koşusu yapmıştır. Derya toplam 3 turda 810 metre koştuğuna göre, sitenin kısa kenarı kaç metredir? A) 30 B) 45 C) 60 D) 90</p>	 <p>Serkan arkadaşlarıyla yapacağı futbol maçı öncesi yukarıda şekli verilen sahanın çevresinde ısınma koşusu yapmıştır. Derya toplam 3 turda 810 metre koştuğuna göre, sahanın kısa kenarı kaç metredir? A) 30 B) 45 C) 60 D) 90</p>

- Gereksiz görsel kullanılmamasına da özen gösterilmelidir. Görsel çıkarıldığında anlamca bir kayıp oluşmuyorsa o görsel kullanılmamalıdır.

Yanlış örnek	Doğru örnek- Görsel olmadan cevaplanamaz																																																																																																		
 <p>Her yazının birinci paragrafı önemlidir. Başlık, giriş, gelişme, sonuç ve kapanış bölümlerine sahiptir. Bu bölümlerin her birinin yazılma amacı farklıdır. Bu bölümlerin yazılma sırası da farklıdır. Bu bölümlerin yazılma sırasını aşağıdaki seçeneklerden hangisi doğru olarak göstermektedir? A) Giriş, gelişme, sonuç, kapanış, başlık B) Başlık, giriş, gelişme, sonuç, kapanış C) Giriş, gelişme, sonuç, kapanış, başlık D) Giriş, gelişme, sonuç, başlık, kapanış</p>	 <p>Şekil 1: Dünyanın yüksek binaları (metre cinsinden)</p> <table border="1"><thead><tr><th>Bina</th><th>Yükseklik (m)</th></tr></thead><tbody><tr><td>Hong Kong</td><td>318</td></tr><tr><td>Doha</td><td>140</td></tr><tr><td>Singapur</td><td>124</td></tr><tr><td>Şangay Çin</td><td>124</td></tr><tr><td>Bangkok</td><td>118</td></tr><tr><td>Parana Şehri</td><td>113</td></tr><tr><td>New York</td><td>112</td></tr><tr><td>Sao Paulo</td><td>112</td></tr><tr><td>Miami</td><td>103</td></tr><tr><td>Shanghai</td><td>94</td></tr><tr><td>Toronto</td><td>92</td></tr><tr><td>Singapur</td><td>86</td></tr><tr><td>Osaka</td><td>77</td></tr><tr><td>Sao Paulo</td><td>64</td></tr></tbody></table> <p>Şekil 2: Dünyanın en yüksek binasından başlıklar</p> <table border="1"><thead><tr><th>Bina</th><th>Yükseklik (m)</th></tr></thead><tbody><tr><td>Worldwide Financial Center</td><td>333</td></tr><tr><td>Keybank Tower</td><td>324</td></tr><tr><td>Bank of China Tower</td><td>300</td></tr><tr><td>Bank of China Tower</td><td>293</td></tr><tr><td>Bank of China Tower</td><td>283</td></tr><tr><td>Bank of China Tower</td><td>273</td></tr><tr><td>Bank of China Tower</td><td>263</td></tr><tr><td>Bank of China Tower</td><td>253</td></tr><tr><td>Bank of China Tower</td><td>243</td></tr><tr><td>Bank of China Tower</td><td>233</td></tr><tr><td>Bank of China Tower</td><td>223</td></tr><tr><td>Bank of China Tower</td><td>213</td></tr><tr><td>Bank of China Tower</td><td>203</td></tr><tr><td>Bank of China Tower</td><td>193</td></tr><tr><td>Bank of China Tower</td><td>183</td></tr><tr><td>Bank of China Tower</td><td>173</td></tr><tr><td>Bank of China Tower</td><td>163</td></tr><tr><td>Bank of China Tower</td><td>153</td></tr><tr><td>Bank of China Tower</td><td>143</td></tr><tr><td>Bank of China Tower</td><td>133</td></tr><tr><td>Bank of China Tower</td><td>123</td></tr><tr><td>Bank of China Tower</td><td>113</td></tr><tr><td>Bank of China Tower</td><td>103</td></tr><tr><td>Bank of China Tower</td><td>93</td></tr><tr><td>Bank of China Tower</td><td>83</td></tr><tr><td>Bank of China Tower</td><td>73</td></tr><tr><td>Bank of China Tower</td><td>63</td></tr><tr><td>Bank of China Tower</td><td>53</td></tr><tr><td>Bank of China Tower</td><td>43</td></tr><tr><td>Bank of China Tower</td><td>33</td></tr><tr><td>Bank of China Tower</td><td>23</td></tr><tr><td>Bank of China Tower</td><td>13</td></tr><tr><td>Bank of China Tower</td><td>3</td></tr></tbody></table>	Bina	Yükseklik (m)	Hong Kong	318	Doha	140	Singapur	124	Şangay Çin	124	Bangkok	118	Parana Şehri	113	New York	112	Sao Paulo	112	Miami	103	Shanghai	94	Toronto	92	Singapur	86	Osaka	77	Sao Paulo	64	Bina	Yükseklik (m)	Worldwide Financial Center	333	Keybank Tower	324	Bank of China Tower	300	Bank of China Tower	293	Bank of China Tower	283	Bank of China Tower	273	Bank of China Tower	263	Bank of China Tower	253	Bank of China Tower	243	Bank of China Tower	233	Bank of China Tower	223	Bank of China Tower	213	Bank of China Tower	203	Bank of China Tower	193	Bank of China Tower	183	Bank of China Tower	173	Bank of China Tower	163	Bank of China Tower	153	Bank of China Tower	143	Bank of China Tower	133	Bank of China Tower	123	Bank of China Tower	113	Bank of China Tower	103	Bank of China Tower	93	Bank of China Tower	83	Bank of China Tower	73	Bank of China Tower	63	Bank of China Tower	53	Bank of China Tower	43	Bank of China Tower	33	Bank of China Tower	23	Bank of China Tower	13	Bank of China Tower	3
Bina	Yükseklik (m)																																																																																																		
Hong Kong	318																																																																																																		
Doha	140																																																																																																		
Singapur	124																																																																																																		
Şangay Çin	124																																																																																																		
Bangkok	118																																																																																																		
Parana Şehri	113																																																																																																		
New York	112																																																																																																		
Sao Paulo	112																																																																																																		
Miami	103																																																																																																		
Shanghai	94																																																																																																		
Toronto	92																																																																																																		
Singapur	86																																																																																																		
Osaka	77																																																																																																		
Sao Paulo	64																																																																																																		
Bina	Yükseklik (m)																																																																																																		
Worldwide Financial Center	333																																																																																																		
Keybank Tower	324																																																																																																		
Bank of China Tower	300																																																																																																		
Bank of China Tower	293																																																																																																		
Bank of China Tower	283																																																																																																		
Bank of China Tower	273																																																																																																		
Bank of China Tower	263																																																																																																		
Bank of China Tower	253																																																																																																		
Bank of China Tower	243																																																																																																		
Bank of China Tower	233																																																																																																		
Bank of China Tower	223																																																																																																		
Bank of China Tower	213																																																																																																		
Bank of China Tower	203																																																																																																		
Bank of China Tower	193																																																																																																		
Bank of China Tower	183																																																																																																		
Bank of China Tower	173																																																																																																		
Bank of China Tower	163																																																																																																		
Bank of China Tower	153																																																																																																		
Bank of China Tower	143																																																																																																		
Bank of China Tower	133																																																																																																		
Bank of China Tower	123																																																																																																		
Bank of China Tower	113																																																																																																		
Bank of China Tower	103																																																																																																		
Bank of China Tower	93																																																																																																		
Bank of China Tower	83																																																																																																		
Bank of China Tower	73																																																																																																		
Bank of China Tower	63																																																																																																		
Bank of China Tower	53																																																																																																		
Bank of China Tower	43																																																																																																		
Bank of China Tower	33																																																																																																		
Bank of China Tower	23																																																																																																		
Bank of China Tower	13																																																																																																		
Bank of China Tower	3																																																																																																		

- Bir maddeyi okuyan bir öğrenci öğretmenin zihninden geçeni keşfetmek zorunda değildir.

Yanlış örnek	Doğru örnek
Bir testte hangisi en önemlidir? A) Madde güçlüğü B) İfade netliği C) Soru sayısı D) Madde ayırtediciliği E) Test düzeni	Güvenlilik söz konusu olduğunda aşağıdakilerin hangisi en önemlidir? A) Madde güçlüğü B) İfade netliği C) Soru sayısı D) Madde ayırtediciliği E) Test düzeni

- Madde kökünde gereksiz yere sözcük kullanılmamalıdır. Zaman zaman öğretmenler kendi doğrularını ya da önemli yerleri vurgularken bu hatayı yapmaktadırlar.

Yanlış örnek	Doğru örnek
Basın özgürlüğü indeksinin son sıralarında yer alan ülkelerde, demokrasinin olmazsa olmazı olarak nitelenen basın özgürlüğünün tesis edilmesi için, hangi yasal düzenlemenin öncelikle yapılması gerekir?	Basın özgürlüğü indeksinin son sıralarında yer alan ülkelerde, basın özgürlüğünün tesis edilmesi için, hangi yasal düzenlemenin öncelikle yapılması gerekir?

• Bazı öğretmenler karışık bir dille madde yazdığında sorunun zorlaşacağını düşünür. Oysa bu oldukça yanlış bir yöntemdir. Karışık bir dil yerine ortalama bir öğrencinin anlayabileceği bir dil kullanılmalıdır. Çoktan seçmeli maddeyi zorlaştırmak isteyen öğretmen temelde iki yöntem kullanabilir:

a. Çeldiricileri doğru yanıtı yaklaştırmak (güçlendirmek) ve b. Davranışı üst düzey becerileri ölçen bir formda hazırlamak.

Yanlış örnek	Doğru örnek
<p>Testin bütününden bağımsız alt boyutların her birisinin, bütünle ve diğer alt boyutlarla en az korele olduğu ve alt test içi maddeler arasındaki ilişkinin maksimuma yükseldiği durumda, testten elde edilen puanlar için aşağıdaki yargılardan hangisine varılabilir?</p> <p>A) Konu kapsamı dengeli bir biçimde temsil edilmektedir. B) Testin görünüş geçerliliği düşüktür. C) Hedeflerde bilişsel alanlara göre dengeli bir dağılım söz konusudur. D) Güvenilirlik katsayısı yüksektir. E) Yordama gücü düşüktür.</p>	<p>İç tutarlılık katsayısı yüksek bir test için aşağıdakilerden hangisi söylenebilir?</p> <p>A) Konu kapsamı dengeli bir biçimde temsil edilmektedir. B) Testin görünüş geçerliliği düşüktür. C) Hedeflerde bilişsel alanlara göre dengeli bir dağılım söz konusudur. D) Güvenilirlik katsayısı yüksektir. E) Yordama gücü düşüktür.</p>

• Seçeneklerde gereksiz sözcük tekrarı kaçınılmalıdır. Tekrar edilen sözcük madde köküne alınabilir.

Yanlış örnek	Doğru örnek
<p>Buna göre, Sivas Kongresi'nin Kurtuluş Savaşı için aşağıdakilerden hangisini sağlamadığı söylenemez?</p> <p>A) Amaç birliği B) Kültür C) Milli birliği D) Güç birliği</p>	<p>Buna göre, Sivas Kongresi'nin Kurtuluş Savaşında hangi konuda birlik sağlamaya etkisi <u>olmamıştır</u>?</p> <p>A) Amaç B) Kültür C) Milliyetçilik D) Güç</p>

• Bir maddenin doğru yanıtı, başka bir maddenin kökünde ya da seçeneklerinde bulunmamalıdır.

Yanlış örnek	Aynı Sınav
<p>1. Karamazov Kardeşler adlı romanın yazarı kimdir?</p> <p>7. Suç ve Ceza, Karamazov Kardeşler, Kumbarbaz gibi kült eserlerin yazarı Dostoyevski hangi ulustandır?</p>	<p>1. sorunun cevabını 7. sorudan verebilir.</p>

• Seçeneklerin hepsi birbirleriyle tutarlı ve ayrı gramatik yapıya sahip olmalıdır. Eğer seçenekler ifade bakımından birbirinden farklıysa çoğu durumda bu öğrencilere ipucu veren bir yapı sergiler. Özellikle eklerin kullanımına dikkat etmek gerekir.

Yanlış örnek	Doğru örnek	Yanlış örnek
<p>Başarıyı ölçmek üzere geliştirilmiş bir testte, geçerlilik aşağıdakilerden hangisiyle öncelikle ilgilidir?</p> <p>A) Soruların konu kapsamını dengeli bir biçimde yansıtmasıyla B) Testte madde/soru çeşitlenmesine gidilmesi C) Testi uzmanların hazırlaması D) Soru sayısının artırılması E) Heterojen bir gruba uygulama yapılması</p>	<p>Başarıyı ölçmek üzere geliştirilmiş bir testte, geçerlilik aşağıdakilerden hangisiyle öncelikle ilgilidir?</p> <p>A) Soruların konu kapsamını dengeli bir biçimde yansıtması B) Testte madde/soru çeşitlenmesine gidilmesi C) Testi uzmanların hazırlaması D) Soru sayısının artırılması E) Heterojen bir gruba uygulama yapılması</p>	<p>Aşağıdakilerden hangisi güvenilirliği arttıran bir işlem <u>değildir</u>?</p> <p>A) Soru sayısını arttırmak B) Anlaşılır madde yazmak C) Teste yönerge eklemek D) Soruların zor olması E) Testi heterojen gruplara uygulamak</p>

• İpucu vermektan kaçınmak gerekir. Özellikle “bazen, çoğunlukla, genellikle, sıklıkla ya da zaman zaman” vb. ifadelerin doğru olma olasılığı yüksektir. Diğer taraftan “hiçbir zaman, asla, daima, hiçbir, her zaman, tümü” vb. ifadelerin de yanlış olma olasılığı yüksektir. Bu nedenle bu tür sözcükler seçeneklerde kullanılmamalıdır.

Yanlış örnek
<p>Testteki soru sayısının artırılması halinde geçerlilik nasıl etkilenir?</p> <p>A) Her zaman artar. B) Daima azalır. C) <u>Bazen artar.</u> D) Etkilenmez.</p>

• Birbirinin tam zıttı olan durumların seçeneklerde verilmesi çoğunlukla uygun değildir. Çünkü genellikle ikisinden biri doğrudur. Şans başarısı artar, geçerlilik ve güvenilirlik düşer.

Yanlış örnek
<p>Testteki soru sayısının artırılması halinde testin psikometrik niteliklerinde nasıl bir değişim olması beklenir?</p> <p>A) <u>Güvenilirlik artar.</u> B) Güvenilirlik azalır. C) Kullanışlılık azalır. D) Geçerlilik artar. E) Test zorluğu artar.</p>

Mustafa Uçan

- Seçenekler yazılırken anlamca birbirini içeren / kapsayan ifadelerin kullanılmamasına dikkat edilmelidir.

Yanlış örnek	Doğru örnek
<p>Parmak emme davranışı yaşamın belirli bir döneminde olağan bir davranış olarak nitelendirilirken, bu davranışın başka gelişim dönemlerinde görülmesi bireyin bir travması olduğunu düşündürülebilir.</p> <p>Parmak emme davranışı öncelikle hangi gelişim döneminde görülürse travmatik bir davranış olarak değerlendirilmelidir?</p> <p>A) Bebeklik B) Çocukluk C) Ergenlik D) Genç yetişkinlik E) Yetişkinlik</p> <p>cocukluk bebekliği kapsar.</p>	<p>Parmak emme davranışı yaşamın belirli bir döneminde olağan bir davranış olarak nitelendirilirken, bu davranışın başka gelişim dönemlerinde görülmesi bireyin bir travması olduğunu düşündürülebilir.</p> <p>Parmak emme davranışı öncelikle hangi yaş döneminde görülürse travmatik bir davranış olarak değerlendirilmelidir?</p> <p>A) 0-2 B) 3-11 C) 12-18 D) 19-25 E) 26 yaş ve üstü</p> <p>Yaş aralığı daha uygun</p>

- Seçeneklerin uzunlukları birbirine eşit olmalıdır. Özellikle doğru cevabın daha uzun ya da daha kısa olmamasına özen gösterilmelidir.

Yanlış örnek

Aşağıdaki cümlelerin hangisinde abartılı anlatım vardır?

A) Sahildeki kızgın kum taneçikler zerreleri ayağımı yakıyordu.
B) Gittiği tatil köyünde pek çok arkadaş edindi.
C) Çarşısı, bütün şehri dolaştım, dünyayı alt üst ettim, gezmediğim yer kalmadı, çarşıda aradığımı bulamadım.
D) Anadolu türkleri kültür birikimimizin örnekleridir.

- Doğru cevaplar teste dağıtılırken dikkat edilmeli, doğru cevapların bir örüntü göstermemesi sağlanmalıdır.



- “Hepsi” ve “hiçbiri” bir seçenek olarak kullanılmamalıdır.

Yanlış örnek

Pamuk aşağıdaki illerden hangisinde yetişir?

A) Antalya
B) Adana
C) Aydın
D) Hepsi

**iki seçeneği bulan direk
hepsi der.**

- Çoktan seçmeli maddelerde seçenek sayısı, Lise ve daha üstü öğretim düzeylerinde 5, ortaokul düzeyinde 4 ve ilkokul düzeyinde 3 seçenek kullanılabilir.

- Seçenekler bir sıra ile verilmelidir. Seçenekler baş harflerine göre alfabetik olarak, eğer sayı iseler büyüklük olarak sıralanmalıdır.

Doğru örnek

Ali, 8 Eylül'de Hatay'a saat 18.00'da ulaştığına göre uçuşuna İstanbul'dan kaçta başlamıştır?

A)06.00 B)14.00 C) 18.00 D)20.00

Seçenekler küçükten büyüğe sıralanmış

- Hangi özellik ölçülürse ölçülsün, çoktan seçmeli madde yazımının bütün aşamalarında hem madde kökünde hem de seçeneklerde dil bilgisi ve imla kuralları konusunda yüksek bir titizlik ve özen göstermek çok önemlidir. **İyi bir madde yazarı;**

- madde yazılan alanı çok iyi bilmeli,
- madde yazma teknik ve yöntemlerinden haberdar olmalı,
- maddelerin yazılacağı dili kullanmada becerikli olmalı ve
- testin uygulanacağı öğrencilerin gelişim düzeyini çok iyi bilmelidir.

- Çoktan seçmeli maddelerde seçeneklerin baş harfleri büyük olmalıdır.

- Bir madde kökü ile seçenekleri aynı sayfada olmalıdır. Yarıya başka sayfada, yarıya diğer sayfada olmamalıdır. Bu durum ortak köke dayalı maddeler için de geçerlidir.

- Eğer ortak köke dayalı madde grubu yazılacaksa hangi soruların ortak köke göre yanıtlanacağı belirtilmelidir. İlkokulda madde sayısını çok artırmamak gerekir.

Örneğin, bir grafik veya tablo verilip onunla ilgili birkaç soru sorulur. '3.,4. Ve 5. Soruları tabloya göre yorumlayınız.' tarzında.

AÇIK UÇLU SORULAR

4. ESSAY (YAZILI YOKLAMA):

- Cevaplayıcıların sorulara cevap verme konusunda sınırsız bir özgürlüğü vardır. Sınırsız cevap özgürlüğünün hem avantajı hem de dezavantajı bulunmaktadır.

Doğru örnek

Kurbağalar 200 milyon yıldır tundralardan tropikal bölgelere kadar uzanan geniş bir alanda çoğalıp yayıldılar. Günümüzde hızla ve belki de geri dönülemez bir biçimde ortadan kayboluyorlar.

National Geographic, Mayıs 2011

Kurbağa neslinin yok olmaya yüz tutmasının nedenlerini açıklayınız.

Avantajı: Öğrencinin yalnızca sahip olduğu bilgiyi değil, aynı zamanda zihninin nasıl işlediğini anlama olanağı sağlar.

Dezavantajı: Öğrenci sınırlı bir bilgiye sahipse soruyla ilgisi olmayan, aklına estiği gibi uzun cevaplar yazabilir.

- Sınırsız cevap özgürlüğü, anlatım becerisi iyi olan öğrencilere bir avantaj sağlayabilir.

• Bu sınav türü, dünyada en eski ve günümüzde de öğretmenlerin hâlâ en çok tercih ettiği 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.

• Çok fazla soru sorulamaması (Okuma ve yazma eylemlerinin süreler arasında ortalama 10 kat fark bulunmakta.) bir dezavantajdır. Kapsam açısından yeterince soru sorulamadığından **Kapsam Geçerliliği riske girer.**

• Puanlar sadece öğrencinin sahip olduğu bilgiyi değil; öğrencinin anlatım biçimini, yazı güzelliğini, bilgisini örgütlenme biçimini de yansıtır. Bu ise ölçme sonuçlarındaki sistematik hata miktarını artırabilir.

• Puanlama öznelliği belki de en önemli dezavantajdır.

"Essay"de Nesnellik Artırma Yolları:

- Bilgi dışında; kağıt düzeni, yazı güzelliği gibi durumlara itibar edilmemeli.
- Kağıttaki isimleri puan vermeden önce kapatmalı.
- Kağıttaki soruları tek tek değil; soru soru tüm sınıfı okumalı.

Ayrıntılı bir cevap anahtarı çıkarılmalıdır. Öğrenci ne yaparsa kaç puan verileceğinin belirlenmesi gerekir.

Anahtar	Puan
Yararlar	
a) Işık enerjisini soğurma yüzeyi artar (genişlik)	2
b) Işık enerjisi içeriye girebilir (yassılık)	2
c) Karbondioksit alıp oksijen vermeyi kolaylaştırır	2
d) Su buharlaşması kolay olur (suyun gereğinden fazla olduğu hal)	2
Güçlükler	
e) Su kaybı çok olur (suyun kıt olduğu hal)	2

Bütüncül (holistik) Rubrik Örneği

Ölçütler	Başlangıç Düzeyinde (1)	Geliştirilmesi Gerekir (2)	Kabul Edilebilir (3)	Tam Olarak Başarılı (4)	Başarı Puanı
İçerik	Yazılanlar (seçilen sözcükler, cümleler, deyimler, örnekler vb.) verilen konuyla ya da olayla ilişkili değildir ve birbirini desteklememektedir. Belirlenen bağlanlarda yazılanların neredeyse tümü hatalı ve iletsini aktarmada yetersizdir.	Yazılanlar (seçilen sözcükler, cümleler, deyimler, örnekler vb.) verilen konuyla ya da olayla çok az ilişkili ve yeteri kadar destekleyici değildir. Belirlenen bağlanlarda yazılanların çok azı, doğru ve iletsini aktarmada yeterlidir.	Yazılanlar (seçilen sözcükler, cümleler, deyimler, örnekler vb.) verilen konuyla ya da olayla oldukça ilişkili ve birbirini destekleyicidir. Belirlenen bağlanlarda yazılanların çoğu doğru ve iletsini aktarmada yeterlidir.	Yazılanlar (seçilen sözcükler, cümleler, deyimler, örnekler vb.) verilen konuyla ya da olayla tam olarak ilişkili ve birbirini destekleyicidir. Belirlenen bağlanlarda yazılanların tümü doğru ve iletsini aktarmada yeterlidir.	

Analistik Rubrik Örneği

5. KISA CEVAPLI SORULAR:

Kısa cevaplı testler; öğrencinin bir sözcük, bir rakam, bir tarih ya da en çok bir cümle ile cevaplandırabileceği sorulardan oluşur. Kısa cevaplı testler, öğrencinin aklının estiği cevabı yazmasını büyük ölçüde engelleyecek bir özelliğe sahiptir. Bundan ötürü maddeleri yanıtlayan bir cevaplayıcının yazacağı cevaplar çok kısadır. İstenen cevapların kısa ve belirgin olması, kısa cevaplı testlerin puanlanmasını genellikle nesnel kılar.

- Kısa cevap, soru cümlesi ya da eksik cümle yapısında olabileceği gibi hatırlama ya da tanımayı gerektiren bir yapıda da olabilir.

Bir çoktan seçmeli maddedeki yanlış cevaplara ne ad verilir?

Çeldirici sözcüğünü tanımlayınız.

- Eğer konu kapsamı bakımında homojen bir yapıda ise gruplandırılarak sorulması daha uygun olacaktır.

Aşağıdaki cümlenin öğelerini (altı çizilmiş kısımların) adlarını, onlar için ayrılan boşluklara yazınız.
"İhtiyar delikanlı, polisten kaçan çocuğu yakalayiverdi."
(A) (B) (C)
A) Özne
B) Nesne
C) Yüklem

Aşağıda Türkiye'de üretilen bazı ürünlerin adları verilmiştir. Bu ürünlerin her birinin Türkiye'de en çok yetiştirildiği bölgenin adını karşısındaki boşluğa yazınız.
A. Fındık _____
B. Tütün _____
C. Buğday _____
D. İncir _____
E. Zeytin _____

- Uzun cevap gerektiren sorularda olduğu gibi bu soru türünde de ayrıntılı bir cevap anahtarı hazırlanmalıdır. Öğrencilerin verecekleri cevap çeşitliliğinin öğretmenin cevap repertuarından her zaman daha fazla olacağı göz önüne alındığında, anahtarın zümrece hazırlanması önerilebilir.

Avantajları:

- a. Hazırlaması, uygulaması ve puanlaması kolay; kullanışlıdır.
- b. Soru sayısı artırılabilir.
- c. Sistemik hata karışma olasılığı pek yoktur.

Dezavantaj: İyi yazılmazsa genellikle üst düzey bilişsel özelliklerin ölçülememesi yine bir dezavantajdır.

6. CÜMLE TAMAMLAMA SORULARI: Cümle tamamlama, kısa cevaplı / sınırlı cevap gerektiren soru grubudur. Bu tür sorularda öğretmen önemli gördüğü bir cümleyi alır, kritik gördüğü bir ya da birden fazla yeri çıkarır ve yerine bir boşluk koyar. Öğrenciden o cümleyi anlamlı bir biçimde tamamlaması istenir. Cümle tamamlama soruları boşluk doldurma olarak da geçmektedir.

Örnek

Dekametre, metrik sistemde metrenin _____ katıdır.

- Hazırlanması oldukça kolaydır. Bu yargı, özellikle *kim, ne, nerede, ne zaman?* sorularına cevap olabilecek olgusal bilgileri ölçmeye yönelik soruların yazılmasında geçerlidir.
- Cümleler ders kitabından aynen alınmamalı, öğretmen kendi sözcükleriyle soruları yeniden yazmalıdır.

Avantajları:

- a. Hazırlaması, uygulaması ve puanlaması kolay; kullanışlıdır.
- b. Soru sayısı artırılabilir.
- c. Sistemik hata karışma olasılığı pek yoktur.

Dezavantaj: İyi yazılmazsa genellikle üst düzey bilişsel özelliklerin ölçülememesi yine bir dezavantajdır.

Mustafa Uçan

Cümle Tamamlama Soruları Yazım Kuralları:

• Her kısa cevap maddesi, yalnızca tek bir doğru cevabı olacak biçimde yapılandırılmalıdır. Tartışma götürecektir, yoruma açık durumların soru yapılmasından kaçınılması ve sorunun ifadesinin açık, net ve anlaşılır olması gerekir.

Yanlış örnek

Toplumda toprak ağalarının egemenliğini işleyen tanınmış roman _____ tarafından yazılmıştır.

Doğru örnek

İnce Memed adlı roman _____ tarafından yazılmıştır.

• Bir maddenin ifadesinde, o maddenin cevabının bulunmasında işe yarayacak ipuçları vermektten kaçınılmalıdır. Doğru cevabın verilmesinde kullanılabilecek ipucu, maddede verilen gereksiz bir bilgi ya da cümlenin gramer yapısı olabilir.

Yanlış örnek

Osmanlı Devleti, 1914'te başlayan Birinci Dünya Savaşına _____ yılında girmiştir.

Doğru örnek

Osmanlı Devleti, Birinci Dünya Savaşına _____ yılında girmiştir.

• Cümlenin gramer yapısı bağlamında ipucu genellikle eklerle verilmektedir.

Yanlış örnek

Huzur romanının yazarı Eke uygun isim uydurur. 'dir.

Doğru örnek

Huzur romanı _____ tarafından yazılmıştır.

• Maddenin cevabında birden çok ayrıntı varsa o ayrıntıların her birine ayrı puan vermek gerekir.

Örnek

___ ve ___ vitaminleri suda çözünür ve pek çoğu vücutta depolanmaz.

• İlköğretimde cevapların ayrı bir cevap kâğıdına işaretlenmesi uygun görülmemektedir; ancak ortaöğretimde bunun pek bir sakıncası bulunmamaktadır.

• Bir cümlede çok sayıda boşluk bırakılmamalı ve cümlelerden anahtar niteliğindeki anlamlı ve önemli sözcükler çıkarılmamalıdır. Eksik cümle istenileni anlatacak biçimde yapılmış olmalıdır.

Yanlış örnek

Dilim, pastayı bıçakla bölmekle elde edilir.
şeklinde yazılsa puan verilmeli.

Doğru örnek

Bölüm bölüneni bölene bölmekle elde edilir.

• Herkesçe aynı biçimde algılanacak, öğrencinin *“acaba öğretmenim bununla neyi kastediyor?”* sorusunu sormayacağı biçimde soru yazmak gerekir. **Öğrenci, öğretmenin zihninden geçeni keşfetmek zorunda değildir.**

Yanlış örnek

Telefon Tarik mi? bulunmuştur.
isim mi?
olay mı?

Doğru örnek

Telefon _____ yılında bulunmuştur.

Telefon _____ tarafından bulunmuştur.

7. SÖZLÜ YOKLAMA: Sözlü sınavlar çok iyi organize edilse / kurgulansa dahi ciddi bir ölçme hatası barındırma potansiyeline sahiptir. Ancak bütün bu **potansiyel sınırlılıklarına rağmen sözlü yoklamalar özellikle dile dayalı becerilerin ölçülmesi için geleneksel yöntemler içindeki tek türdür. Örneğin sözlü anlatım, diksiyon, yabancı dil dersinde konuşma, solfej ve şarkı söyleme vb.**

- Sorular sözlü olarak sorulur ve cevap sözlü olarak verilir.
- Öğretmen ile öğrenci arasında karşılıklı ve devamlı bir etkileşim vardır.
- Bireysel bir test olması nedeniyle her öğrenciye ayrı soru sorma zorunluluğu vardır.
- Bireysel bir test olması nedeniyle tüm öğrencileri test etmek için gereken süre oldukça fazladır.
- Öğrencinin cevaplarını gözden geçirme şansı yoktur.
- Cevaplama çoğu zaman üzerinde düşünmeye ve tasarlamaya olanak olmadan verilir.
- Puanlama hemen yapılır. Bu ise cevabın doğruluğunun genel izlenimle yapılmasına neden olur.
- Öğrencinin yerinde ayağa kaldırılması ya da tahtaya kaldırılması kaygı vb. bireyden kaynaklanan hata miktarını artırabilir.
- Öğretmenin ölçme sonucuna sistematik hata karıştırma olasılığı yükselir.
- Sözlü anlatım becerisi iyi olan bir öğrenci, iyi olmayan bir başka öğrenciden 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 Kuralları:

- Sözlü sınavların yapılacağı gün ve saat, yazılı sınavlarda olduğu gibi ilan edilmelidir.
- Sorular önceden hazırlanmalıdır. Soruların seçiminde amaca uygunluğa ve her öğrenciye benzer güçlükte soru belirlenmesine dikkat edilmelidir.
- Cevap anahtarı hazırlanmalıdır.
- Sınavın yapılması aşamasında öğrencilerin kaygısını artıracak davranışlardan kaçınılmalıdır. Hazırbulunuşluluk artırılmalıdır.
- Basit olan sorudan başlanmalıdır.
- Kayıt altına alın(a)mıyorsa puanlama hemen yapılmalıdır.
- Puana ilişkin dönüt hemen verilmelidir.
- Puanlamada sistematik hata kaynaklarına karşı dikkatli olunmalıdır.

Not: Yoklanacak olan davranışlar, yazılı sınav türleriyle de yoklanabiliyorsa sözlü yoklama yapmaktan kaçınılmalıdır.

GENEL DEĞERLENDİRME:

Her madde / soru türünün avantajları olduğu gibi dezavantajları da vardır. Dezavantajları asgariye indirmenin yolu madde / soru çeşitlemesi yapmaktır ancak özellikle ilkökulda, çocukların gelişim süreçleri de dikkate alındığında, madde / soru çeşidi sayısını çok artırmamak gerekir.

Geleneksel test etme yaklaşımları ürün odaklıdır. Eğitim sürecinde ürün elbette önemlidir ancak ürün kadar öğrencinin süreçte bilgiyi nasıl yapılandırdığının izlenmesi de beklenen ürünün ortaya konulabilmesi açısından çok önemlidir. Bu nedenle geleneksel ölçme yöntemleri, destekleyici yöntemlerle desteklenmelidir.

Mustafa Uçan

DESTEKLEYİCİ DEĞERLENDİRME YAKLAŞIMLARI

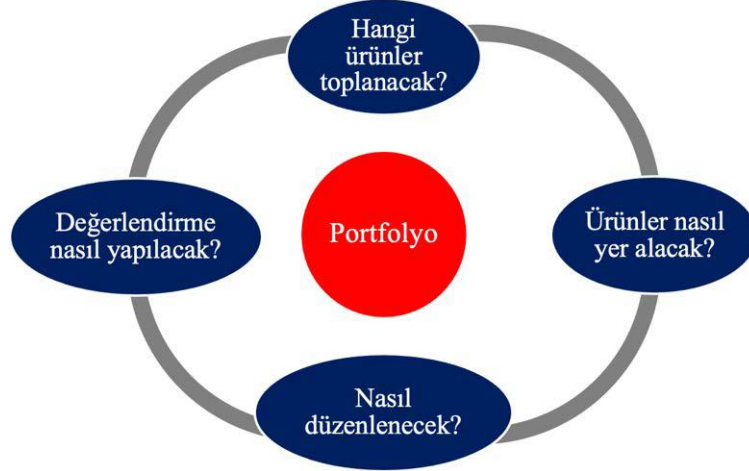
1. PORTFOLYO: Öğrencinin bir bütün olarak gelişim ve öğrenme süreci ile ürünlerini gösteren, aynı zamanda değerlendirilmesini de sağlayan sistemli ve amaçlı olarak oluşturulmuş dosyalardır. Portfolyolar ile amaçlanan, öğrenciyi çok yönlü olarak ve öğrencinin gelişimini bir süreç olarak izlemek ve değerlendirmektir. Portfolyolar ile öğrencinin öz geçmişi, ilgileri, tutumları, becerileri, yetenekleri, öğrenme biçimi, sosyal özellikleri vb. özelliklerinin tanınması amaçlanır. Sürecin paydaşları; öğrenci, öğretmen, arkadaş, aile ve iletişimde olduğu diğer kişilerdir. Öğrencinin yalnızca okulda değil; aile ortamında, oyun ortamında ve dış çevrede tanınması amaçlanır.

- Portfolyolar öğrencinin yaptığı çalışmaların bir araya getirildiği herhangi bir çalışma dosyası değildir.
- Öğrenciyi başarılı / başarısız olarak sınıflandırmayı sağlayan bir araç değildir.
- Öğrencilerin özelliklerini birbirleri ile karşılaştırmak amacıyla oluşturulmuş araçlar değildir.

Öğrenci portfolyosunun kabaca iki türü vardır. Bunlar:

a. Süreci yansıtan portfolyolar: Öğrencinin öğrenme ve gelişim sürecini yansıtır. Başlangıç çalışmalarını, süreçteki çalışmalarını, karşılaşılan güçlükleri ve öğrenme ürünlerini içerir.

b. Ürünü yansıtan portfolyolar: Öğrenme sürecinden çok bitmiş görevleri içerir. Öğrencinin en iyi olduğunu düşündüğü çalışmaları içerir.

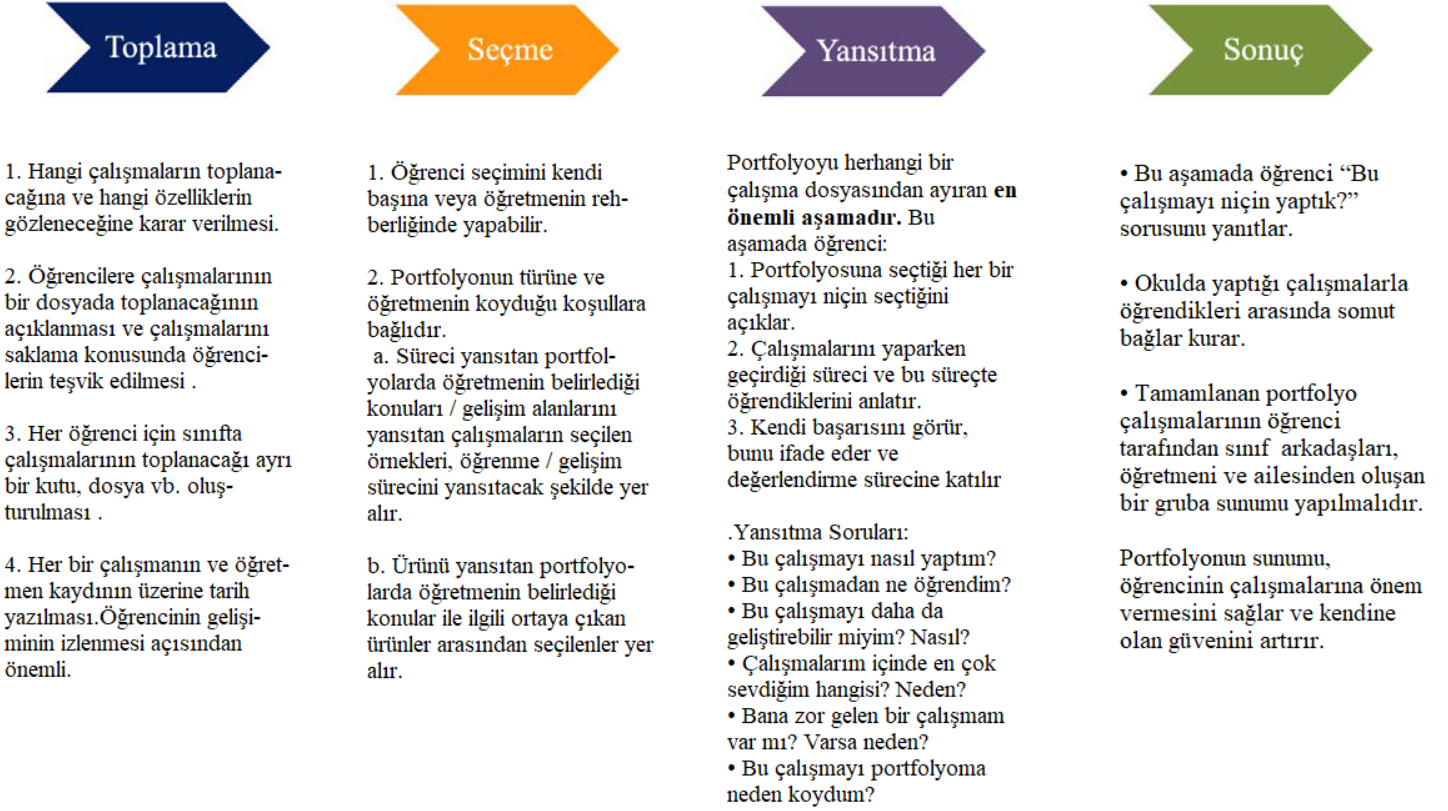


Şekil : Portfolyo değerlendirme süreci

Portfolyonun içeriği:

- Öğretmen kayıtları (gözlemler, anekdot kayıtları)
- Öğrencinin çalışmaları
- Öğrencinin sözel ve psikomotor becerilerini gösteren teyp ve video kayıtları
- Öğrencinin kendi çalışmaları hakkındaki düşünceleri, günlükler
- Öğrenciye yazılan mektuplar
- Öğrencinin yazdığı mektuplar
- Öğretmenin aileye ve diğer öğretmenlere yazdığı mektuplar

Portfolyo Oluşturma Süreci



2.PERFORMANS DEĞERLENDİRME:



✚ **ÖZ DEĞERLENDİRME:** Öğrencinin belirli bir konuda kendi kendisini değerlendirmesidir.

- Öğrencilerin kendi özellikleriyle (yetenek, ilgi, beceri vb.) ilgili farkındalığının artmasını, zayıf ve güçlü yönlerini keşfetmesini sağlar.
- Öz düzenleme becerisi artar. Yani bir işi tanımlama, takip etme, düzenleme ve sonuçlandırma, kendini disipline edebilme becerisi artar.
- Ölçütlü düşünme becerisi artar.
- Öğrenme motivasyonunu artırır.

Öz Değerlendirmenin Olası Dezavantajları:

- a. Öğrencilerin kendi performanslarını değerlendirirken yanlış davranışları söz konusu olabilir.
- b. Başlangıçta deneyimsizlik nedeniyle performansın değerlendirilmesinde yanlışlar olabilir.

Mustafa Uçan

✚ **AKRAN DEĞERLENDİRME:** Öğrencinin ortaya koyduğu performansa ilişkin arkadaşlarının değerlendirmesine denir.

- Akranların değerlendirme sürecine katılması nedeniyle daha katılımcı, aktif bir eğitim ortamı sağlanabilir (Sorumluluk duygusu artar.)
- Öğrencilerin eleştirel düşünme becerileri artar.
- Öğretmen dışındaki bir kaynaktan dönüt almak öğrencinin performansını artırabilir.
- Eleştiri kültürü (olumlu-olumsuz) gelişebilir.

✚ **RUBRİK (DERCELE PUANLAMA ANAHTARI):** Performans ürünlerinin değerlendirilmesinde en sık kullanılan araçlardan biridir. Amaç, öğretmen tarafından ürünün genel izlenimle puanlamasındaki özneliğini azaltmaktır.

Rubrikler ikiye ayrılır: a. Bütüncül (Holistik) rubrik b. Analitik rubrik

Ölçüt	Puan
Maketi yaptığı materyal kalitelidir.	10
Ek yerlerinde pürüz bulunmamaktadır.	15
Özgün bir fikir ortaya koymuştur.	20
Maket çevre düzenlemesi gibi görsel öğelerle zenginleştirilmiştir.	15
Renklendirmeler uyumludur.	10
.....
.....
.....
Toplam	100

Bütüncül (Holistik) Rubrik Örneği

Ölçüt	1 Puan	2 Puan	3 Puan	Puanı
Maketi yaptığı materyal kalitelidir.	Karton	Kontrplak	Masif	
Ek yerlerinde pürüz bulunmamaktadır.	>0.5 cm	0.3-0.5 cm	<0.3 cm	
Özgün bir fikir ortaya koymuştur.	Ders örnekleriyle aynı	Ders örneklerine yeni öğeler katmış	Tamamen özgün bir fikir	
Maket çevre düzenlemesi gibi görsel öğelerle zenginleştirilmiştir.	Zenginleştirme yetersiz	Ağaç, park vb. yapmış ancak yetersiz	Çevreyi tüm öğelerle yansıtmış	
Renklendirmeler uyumludur.	Renklendirmede özensiz	Kısmen uyumlu	Renklendirmeler çok iyi	
.....	
.....	
.....	
Toplam				

Analitik rubrik örneği (Maket yapımı)

- Rubrikler performans görevleriyle birlikte öğrenciye verilmelidir. Öğrenciler kendilerinden ne beklediğini ve nasıl puanlandırılacağını başta bilir.
- Rubriklerin geliştirilmesi uzmanlık gerektirir. Rubrikler için geçerlilik ve güvenilirlik çalışmaları yapılmalıdır.

✚ **KONTROL LİSTELERİ:** Gözlenen performans ürününün ölçütlere uygunluğunu “evet-hayır”, “var-yok”, “gösterdi-göstermedi” vb. bir biçimde kategorik (1-0) olarak puanlama amacıyla kullanılan araçlardır.

- Özellikle sergilenen performans detaylı ve ardışık eylemler gerektirdiği zamanlarda kullanışlıdır. Örneğin bir deneyin eyleme dökülmesi vb. Burada öğrencinin hangi aşamalarda işi yapacağı ayrıntılı olarak listelenmiştir. Meslek Liselerinde çok kullanılır.
- Pek çok işlem adımıyla oluşan performanstaki eksik adımları belirlemek için oldukça uygundur.

✚ **DERECELEME ÖLÇEKLERİ:** Bu araçların kullanımında performans dayalı işlemler ilk baştan sonuna kadar listelenir ve davranışın karşısına davranışın gösterilme derecesi en az üçlü [örneğin tam gösterildi (3), kısmen gösterildi (2) ve gösterilmedi (1)] bir biçimde derecelendirilir.

	Her zaman yapar	Çoğunlukla yapar	Nadiren yapar	Destekle yapar	Yapamaz
Sirasını bekleyebilir.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Duyularını kontrol edebilir.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kendi arkadaş seçebilir.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Oyunun kurallarını açıklayabilir.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kendi başına 20 dk. oynayabilir.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Kendiliğinden özür dileyebilir.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Duyularını ifade edebilir.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Oyuncak vb. eşyalarını paylaşabilir.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

✚ **GÖZLEM FORMLARI:** Öğrenme çıktılarının somut olarak gözlenebildiği bazı alanlarda bu yöntem oldukça kullanışlıdır.

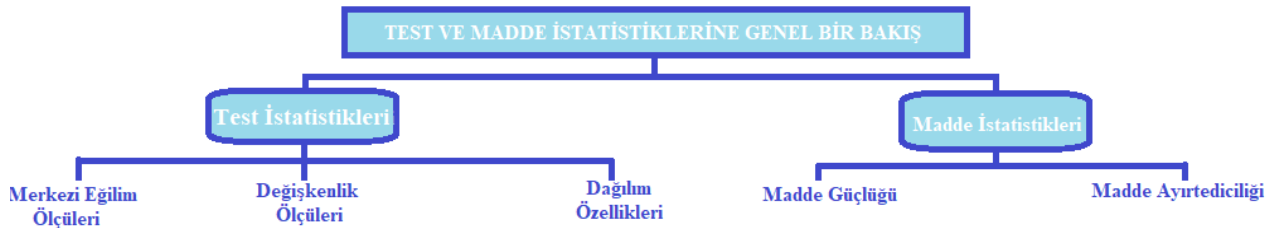
- Özellikle fen derslerinde, meslek liselerinin somut performans ürünlerinin geliştirildiği vb. alanlar için oldukça uygundur.
- Gözlemler öğrenciler hakkında doğru ve hızlı bilgi elde edilmesini sağlar.
- Gözlem formları yarı yapılandırılmış biçimde olabileceği gibi tam yapılandırılmış bir formatta da olabilir.

Mustafa Uçan

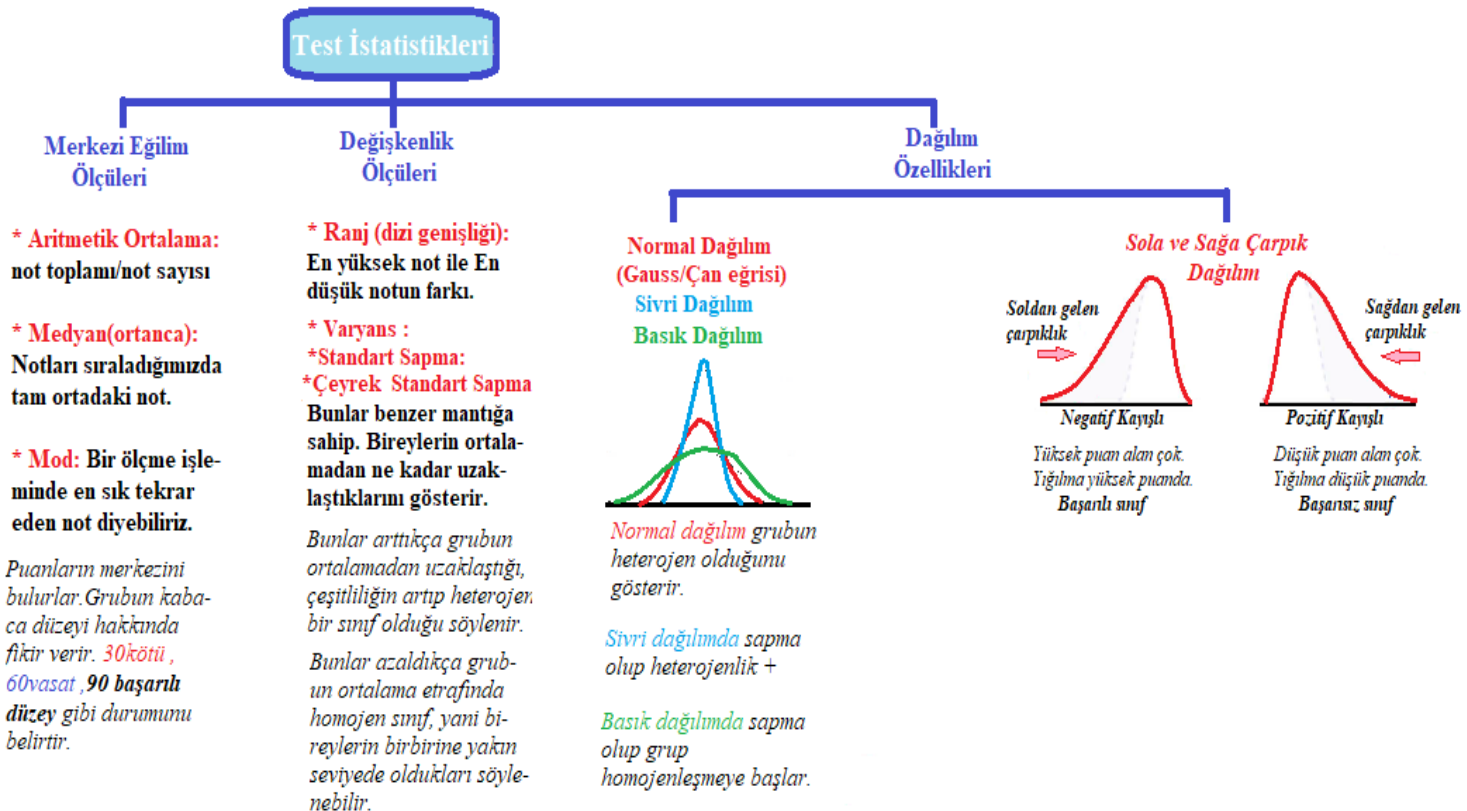
GELENEKSEL VE DESTEKLEYİCİ YÖNTEMLERİN KARŞILAŞTIRILMASI

GELENEKSEL YÖNTEMLER	DESTEKLEYİCİ YÖNTEMLER
Ürün değerlendirilir.	Süreç ve ürün birlikte değerlendirilir.
Öğrencinin ulaştığı noktanın tespiti önemlidir.	Ne öğrendikleri yanında, öğrendiklerini nasıl kullandıklarıyla ilgilenilir.
Essay dışında genellikle üst düzey düşünme becerilerinin değerlendirilmesinde yetersizdir.	Üst düzey bilişsel düşünme becerilerine odaklanır.
Değerlendirme öğrenmeden ayrıdır.	Değerlendirme öğrenmeyle bütünleşmiştir.
Bireyden ziyade gruba odaklıdır.	Odak noktası gruptan ziyade bireydir.
Başarının bireyler arası değerlendirilmesine odaklıdır.	Öğrencinin bireysel olarak gelişimine odaklıdır.
Geçerlilik ve güvenilirlik kontrolü daha kolaydır.	Geçerlilik ve güvenilirlik problemi olabilir.

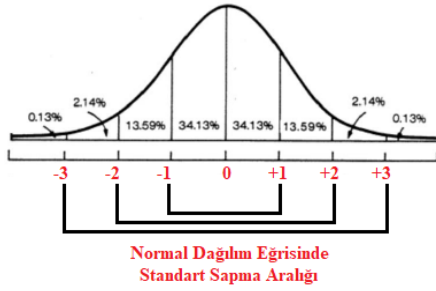
Tablo : Geleneksel ve destekleyici yöntemlerin karşılaştırmalı avantaj ve dezavantajları



Ölçme işlemi sonucunda bireylerin toplam puanları üzerinden hesaplanan istatistiklerdir.



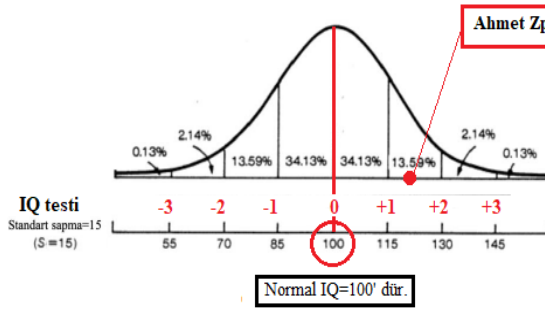
Normal dağılım, aritmetik ortalama ve standart sapma gibi basit iki istatistikle özetlenebilir. Normal dağılım ortalama etrafında üç standart sapma sağa ve sola, toplamda altı alanda özetlenebilir. (Grubun %99,74'lük kısmını açıklar + ve - 3 standart sapma ile özetlenir) Dolayısıyla bir kişinin ortalamadan kaç standart sapma uzaklaştığı, grup içerisindeki yerini gösterir.



$$Z \text{ puanı} = \frac{\text{ham puan} - \text{aritmetik ortalama}}{\text{standart sapma}} = \frac{X - \bar{X}}{S}$$

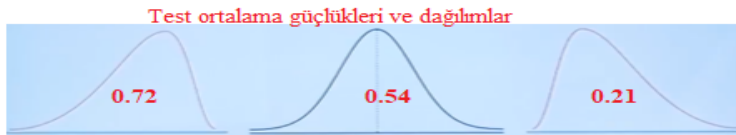
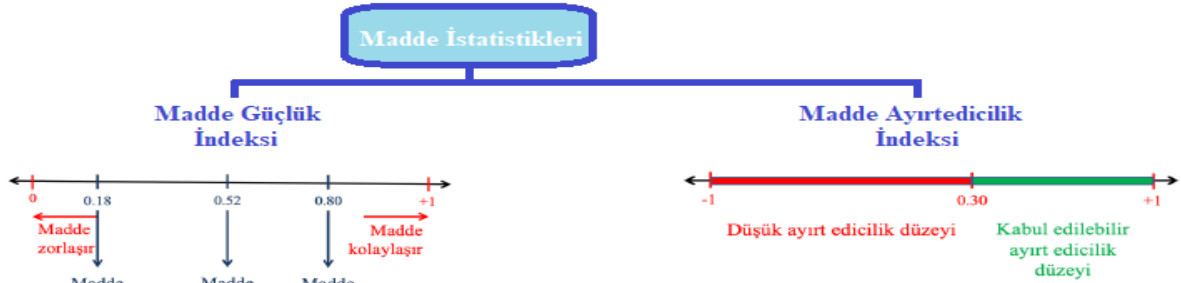
Z puanı: Bireyin puanının ortalamadan ne kadar uzaklaştığını kaç standart sapma yaptığını gösterir. z-1, z+2 vb. Böylece öğrencinin z puanı ile hangi %lik dilimde olduğunu bulabiliriz.

Örnek: Normalde insanların IQ'ları ortalama 100'dür. Bir gruba IQ testi yapılmış olsun. Ahmet isimli birey IQ testi ham puanı 125 almış ve grubun standart sapması da 15 olarak hesaplanmış olsun.



Yorum: Ahmet'in grubun ortalamadan uzaklığı +1 ile +2 olan ve grubun ortalamasının %13.59'lük dilimi yani +1.18 standart sapma üzerindedir. Grup ortalamasının üzerinde bir IQ'ya sahiptir.

Madde İstatistikleri, Ölçme işlemi sonucunda bireylerin madde puanları üzerinden hesaplanan istatistiklerdir.



SOLA ÇARPIK DAĞILIM

*Test çok kolaydır.
*Öğrenciler çalışmış.

NORMAL DAĞILIM EĞRİSİ

* Test orta zorlukta.
* Sınıf heterojen. kötü,orta,iyi ve çok iyi alan var;ancak genelde sınıf orta düzeyde
*Bilen ve bilmeyen öğrenciyi iyi ayırdığından en güvenilir testtir.

SAĞA ÇARPIK DAĞILIM

*Test zordur.
*Öğrenciler çalışmamış.

Mustafa Uçan

EĞİTİM İZLEME ARAŞTIRMALARI: TARİHİ, ÖNEMİ VE TÜRKİYE'DEN SONUÇLAR

İlk kısımda genel olarak kavramların nasıl ele alındığı ve bu kavramlar hakkındaki farklı görüşleri ele alarak başlamak faydalı olacaktır.

Başarı (achievement), Erişi: Belirli bir konu alanında veya uygulama alanında kazandırılan/kazandırılmaya çalışılan kazanımlara bireylerin ulaşma/erişme düzeyleri olarak anlamlandırıyoruz. Yani, sunulan bilgilere, sunulan içeriğe o eğitimi alan bireylerin ne kadar ulaşabildiğine dair bir gösterge, erişim düzeyi olarak ifade edilmektedir.

Beceri: Başarıda, öğrencilerin bilgi ve becerilere sahip olduktan sonra bunları ne kadar kullanabildiğidir. Elde ettiği bilgileri ve kazanımlarını hayatta uygulayabilme yeteneğidir. Geleneksel yöntemde bilgi edinmek önemliyken günümüzde bilgileri ve kazanımları uygun durumlarda kullanmak daha önemlidir. Hatta bunun ötesinde artık okuryazarlık boyutu daha önemli olmuştur.

Okuryazarlık (literacy): Okuryazarlık, özellikle belirli bir alanda ilgili doğru bilgileri bulma, bu bilgileri yanlış bilgilerden ayıklama, bunun için temel okuryazarlık becerilerini kullanma, ardından doğru bilgileri seçerek ve uygun analiz yöntemlerini kullanarak gerçek ya da tasarlanmış durumlarda kullanma becerisi olarak tanımlanmaktadır. Eğitim sistemimiz de dâhil olmak üzere birçok çağdaş eğitim sisteminde, aslında öğrencilere salt bilginin yanında çeşitli beceriler kazandırmak ve kazandırılan bu becerileri doğru ortamlarda, doğru yorumlayacak ve kullanabilecek şekilde gerekli tüm altyapıyı kazandırmak amaçlanmaktadır. Uluslararası Öğrenci Değerlendirme Programında (PISA) ölçülen özellikler “okuryazarlık” olarak tanımlanmaktadır.

Bilimsel ve teknolojik gelişmeler sonucunda bilgiye ulaşmanın kolaylaşmasıyla birçok eğitim sistemi ve iş dünyası artık bilgi ve beceriden çok okuryazarlık aramaktadır. Bu yüzden öğrencilerine bilgilere ulaşım, analiz etme ve böylece doğru bilgi ve beceriyi elde etme, yorumlayarak gerçek hayatta karşılaşılabilecekleri problemleri çözebilecek becerileri kazandırmaya çalışmaktadır. Bu yetkinliğe sahip bireylerin istihdam edilebilirliği artmıştır. Bu nedenle ülkeler eğitim sistemlerine daha fazla daha yatırım yaparak ve gerekli iyileştirmeleri sağlayarak öğrencilerine gerekli becerileri kazandırmaya, onların okuryazarlıklarını geliştirmeye gayret etmektedir.

Uluslararası İzleme Çalışmaları: Aslında bu tarz izleme çalışmalarına ihtiyaç duyulmasının ve bu çalışmaların popüler olmasının sebeplerinden ilki karşılaştırılabilir veri ihtiyacıdır. Ülkelerin eğitime, insan kaynağına ve onları geliştirmek ile ilgili eğitim yatırımları vardır. Bu anlamda ülkeler kendi eğitimlerinin Dünya’da nerede olduklarını bilmeleri ve diğer ülkelerle karşılaştırmaları için veriye ihtiyaç duymakta. Bu yüzden de yıllardır uluslararası izleme çalışmaları yapılmakta.

**Örneğin, PISA-Uluslararası matematik, fen ve okuma becerileri alanlarında;
TIMSS-Uluslararası matematik ve fen,
PIRLS-Uluslararası okuma becerileri değerlendirilmektedir.**

Ayrıca izleme çalışmalarının kapsamı günümüzde sosyal ve duygusal beceriler, vatandaşlık becerileri, yetişkinlik becerileri, dijital okuryazarlık gibi çok farklı alanları da içerecek şekilde genişlemiştir. Ülkelerin eğitim performanslarına dair artan geri bildirim ihtiyacını karşılamak için daha fazla izleme araştırması yapılıyor. Üç-Dört yıllık periyotlarla yapılan bu çalışmalar eğitim sisteminde yapılan işlemler ve bu işlemlerin ayrı ayrı etkilerine dair sonuçlar sağlamamaktadır; eğitim performansının zaman içinde nereden nereye geldiği hakkında bütüncül bir değerlendirmeye imkân sağlamaktadır.

Mustafa Uçan

GÜNÜMÜZDE EN YOĞUN KATILIM GÖSTERİLEN ÇALIŞMALAR

PISA
Programme for International Student Assessment
(Uluslararası Öğrenci Değerlendirme Programı)



*Ekonomik Kalkınma ve İşbirliği Örgütü (OECD) tarafından geliştirilen ve üç yıllık periyotlarla uygulanan izleme çalışmasıdır.
*OECD, Türkiye'nin kurucu üyesi olduğu bir örgüt olup **temel amacı** tüm üye ülkeler arasındaki ekonomik iş birliklerini ve kalkınmayı destekleyecek programlar oluşturarak birlikte kalkınmayı sağlamaktır.
*Bu bağlamda OECD, eğitimi, ekonomik kalkınmaya katkı sunan ve çağdaş vatandaşlık becerilerine sahip insan kaynağının geliştirilmesinde en önemli araç olarak görmektedir.

15 yaş grubu öğrenciler

TIMSS
Trends in International Mathematics and Science Study
(Uluslararası Matematik ve Bilim Çalışmalarında Eğitlimler)



* Uluslararası Eğitim Başarılarını Değerlendirme Kuruluşu (IEA) tarafından geliştirilmekte ve dört yıllık periyotlarla uygulanmaktadır.
*Türkiye uzun süredir TIMSS'e de katılım göstermektedir.
* Dünyadaki en büyük ve en kapsamlı uluslararası öğrenci başarılarını değerlendirme çalışmasıdır.
*Öğrencilerin matematik ve fen alanlarında kazandıkları bilgi ve becerilerin değerlendirilmesine yönelik bir tarama araştırmasıdır.

4.-8.sınıf öğrencileri

PIRLS
Progress in International Reading Literacy Study
(Uluslararası Okuma Becerilerinde Gelişim Araştırması)



*Uluslararası Eğitim Başarılarını Değerlendirme Kuruluşu (IEA) tarafından geliştirilmekte ve beş yıllık periyotlarla gerçekleştirilen bir izleme çalışmasıdır.
*Türkiye uzun bir aradan sonra 2021'de tekrar PIRLS'e katılmış ancak bu uygulamanın sonuçları 2022 yılının sonunda açıklanacaktır.
*Öğrencilerin matematik ve fen alanlarında kazandıkları bilgi ve becerilerin değerlendirilmesine yönelik bir tarama araştırmasıdır.

4.-8.sınıf öğrencileri

PISA KAPSAMI VE SONUÇLARI

OECD, ekonomik kalkınmanın ve sürdürülebilirliğinin sağlanmasındaki en önemli kaynak olan beşerî sermayeyi üreten faktör olarak eğitimi konumlandırmaktadır. Eğitimin ekonomik kalkınma üzerindeki belirleyici rolü dolayısıyla eğitim niteliğini detaylı olarak ele alan ve **üç yıllık periyotlarla uygulanan PISA** çalışmasını gerçekleştirmektedir.

PISA, 15 yaş grubunu hedef grup olarak almaktadır. Bunun nedeni birçok Avrupa ve OECD ülkesinde 15 yaşa tekabül eden dönemin yaklaşık olarak ortaokulun sonuna denk gelmesi ve bu dönemde zorunlu eğitimin tamamlanmasıdır. Dolayısıyla **PISA**, zorunlu eğitim dönemini bitiren öğrencilerin henüz iş gücü piyasasına katılmadan ya da eğitimlerine devam etmeden **matematik, fen ve okuma becerileri alanlarında** hangi yeterliklere sahip olduğuna dair bilgi vermektedir.

TIMSS KAPSAMI VE SONUÇLARI

TIMSS, IEA tarafından dörder yıllık döngüler şeklinde gerçekleştirilmektedir. PISA ile TIMSS arasındaki en temel fark, **PISA**'da zorunlu eğitim dönemini tamamlayan öğrencilerin iş gücü piyasası tarafından ve ekonomik kalkınma açısından istenen, uygulamaya dönük okuryazarlık becerilerinin hangi seviyede olduğu değerlendirilmekte iken **TIMSS**'te eğitim programı (curriculum based) ile ilişkili becerilerin değerlendirilmesidir.

TIMSS'te 4 ve 8. sınıf düzeyindeki öğrencilerin ilgili dönemin eğitim programlarından beklenen özelliklere hangi düzeyde sahip olduklarına dair matematik ve fen alanlarında ayrı ayrı uygulamalar yapılmakta ve ilgili eğitim programlarında öğrencilerin okul başarısına dair daha doğrudan çıktılar elde edilmektedir. **TIMSS**, ilk defa 1995 yılında uygulanmaya başlanmıştır.

PIRLS KAPSAMI VE SONUÇLARI

PIRLS de IEA tarafından uygulanmakta olup okuma boyutunu içermesi sebebiyle **TIMSS**'i tamamlar nitelikte bir çalışmadır. **PIRLS**, öğrencilerimizin okuma becerilerine ve okuma düzeylerine dair önemli çıktılar sağlamaktadır. Türkiye **PIRLS**'e ilk defa 2001 yılında katılmış, 20 yıllık bir aradan sonra alınan kararla 2021 yılında tekrar dâhil olmuştur.

Mustafa Uçan

OECD SOSYAL VE DUYGUSAL BECERİLER ARAŞTIRMASI

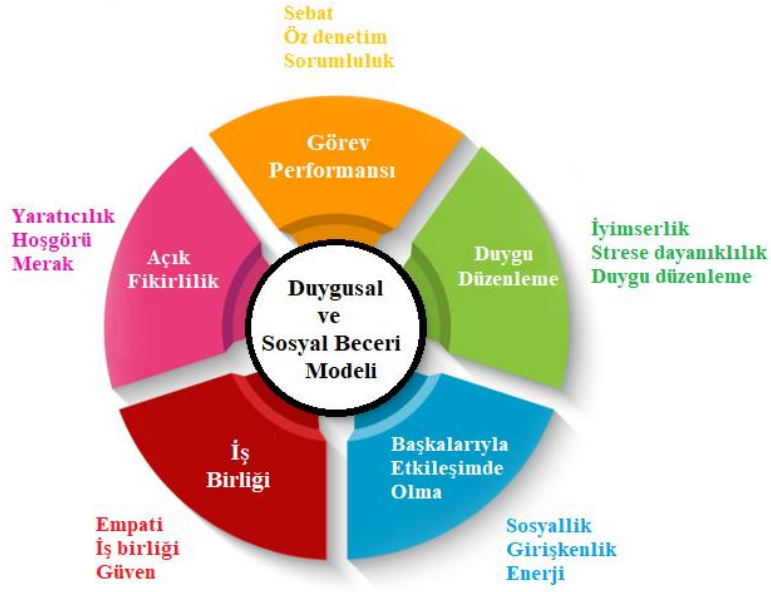
Bu çalışmada;

Birinci olarak, tümüyle sosyal ve duygusal beceriler dikkate alınmıyor ki bu diğer tüm izleme araştırmalarından en önemli farkıdır.

İkincisi, 10 ve 15 yaş grupları ayrı ayrı ele alınmaktadır ki duygusal özellikler, sosyal ve duygusal özellikler yaşa bağlı olarak, özellikle de ergenlik dönemine bağlı olarak önemli değişimler gösterebilmektedir. Ergenlik sürecini de içerecek şekilde iki ayrı yaş düzeyinin dikkate alınması bu becerilerin nasıl değiştiği hakkında önemli çıkarımlar sağlamaktadır.

Üçüncüsü ölçümler, dünyanın farklı bölgelerinde kabul gören beş faktörlü kurama dayalı olarak (Big five modeli) yapılmaktadır. Bu beş faktörlü model:

Büyük Beşli Sosyal ve Duygusal Beceriler Modeli



Beş faktörlü modeli oluşturan ana alanlar ve bu alanları oluşturan alanlar:

- **açık fikirlilik** (yaratıcılık, hoşgörü, merak),
- **iş birliği** (empati, iş birliği, güven),
- **duygu düzenleme** (iyimserlik, strese dayanıklılık, duygu düzenleme),
- **görev performansı** (sebat, öz denetim, sorumluluk)
- **başkalarıyla etkileşimde olma** (sosyallik, girişkenlik, enerji) olarak tanımlanmıştır.

Bu çalışmayı özel kılan diğer bir özelliği de veri çeşitlenmesinin kullanılmasıdır. Veri çeşitlenmesiyle kastedilen aynı sorular üzerinden hem öğrencilerden hem öğretmenlerinden hem de velilerinden bilgi toplanmasıdır ki bu tüm girdilerin/verilerin birleştirilerek sonuçların daha güvenilir hâle getirilmesine katkı sağlamaktadır.

Veri çeşitlenmesi kullanılmasının sebebi, utangaçlık ya da kendine güven gibi duygusal becerilere ilişkin soruları, öğrencilerin oldukları gibi değil olmak istedikleri veya görünmek istedikleri şekilde cevaplandırma eğilimlerinin önüne geçebilmek; daha gerçekçi sonuçlara ulaşabilmektir.

Bu çalışmaya katılan ülkeler ve şehirler şu şekildedir:

Kanada (Ottawa)	ABD (Houston)	Kolombiya (Manizales ve Bogota)
Portekiz (Sintra)	Finlandiya (Helsinki)	Türkiye (İstanbul)
Güney Kore (Daegu)	Çin (Suzhou)	Rusya Federasyonu (Moskova)

Mustafa Uçan