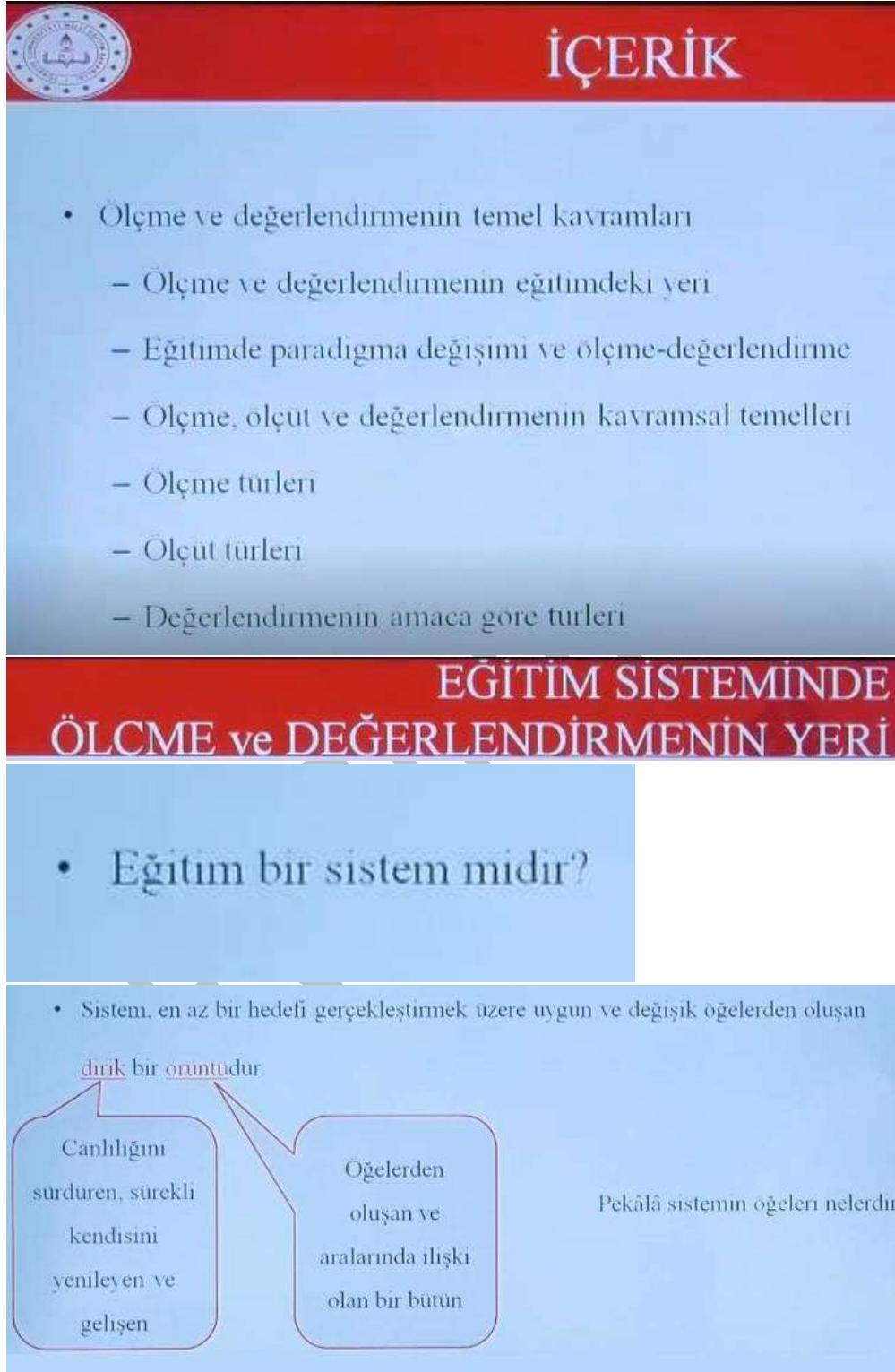


UZMAN ÖĞRETMEN EĞİTİM VİDEOLARINDAKİ GÖRSELLER

2.ÖLÇME VE DEĞERLENDİRME

1- ÖLÇME VE DEĞERLENDİRMENİN TEMEL KAVRAMLARI-1 (Sayfa: 69-71)

EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ



İÇERİK

- Ölçme ve değerlendirmenin temel kavramları
 - Ölçme ve değerlendirmenin eğitimdeki yeri
 - Eğitimde paradigma değişimi ve ölçme-değerlendirme
 - Ölçme, ölçüt ve değerlendirmenin kavramsal temelleri
 - Ölçme türleri
 - Ölçüt türleri
 - Değerlendirmenin amaca göre türleri

EĞİTİM SİSTEMİNDE ÖLÇME ve DEĞERLENDİRMENİN YERİ

- Eğitim bir sistem midir?

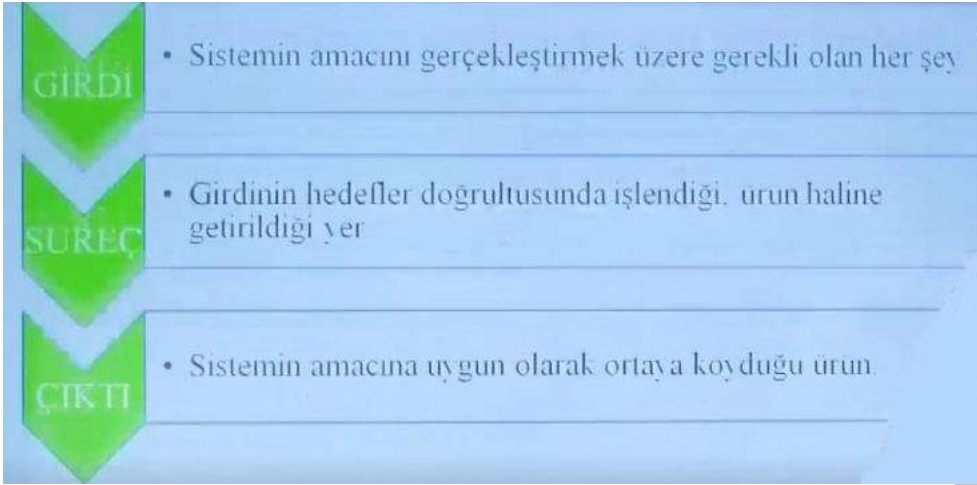
- Sistem, en az bir hedefi gerçekleştirmek üzere uygun ve değişik öğelerden oluşan dirik bir oluntudur.

Canlılığını sürdüren, sürekli kendisini yenileyen ve gelişen

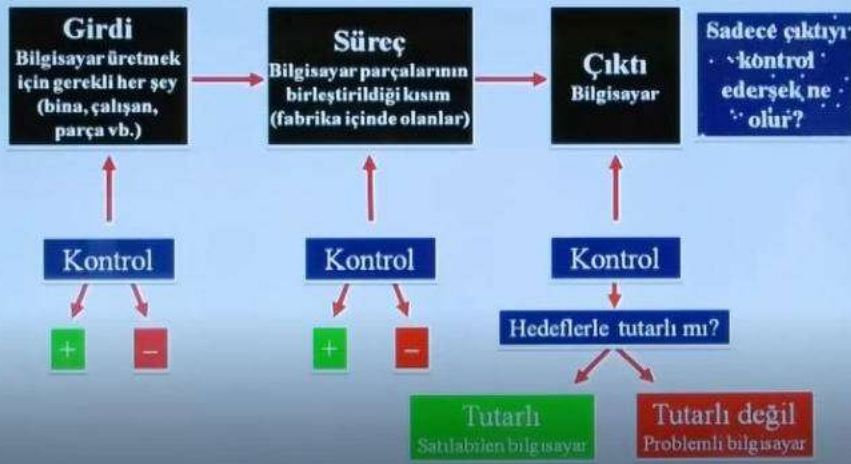
Öğelerden oluşan ve aralarında ilişki olan bir bütün

Pekâlâ sistemin öğeleri nelerdir?

SİSTEMİN ÖĞELERİ



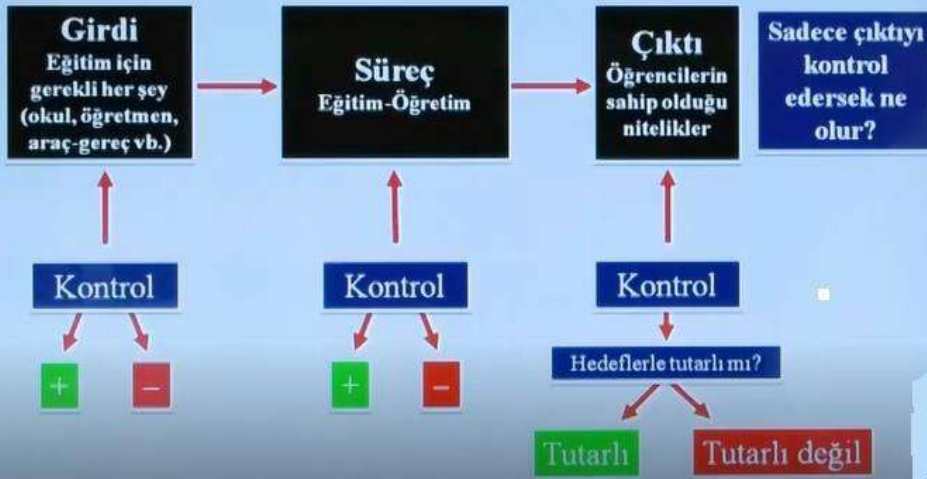
Bir sistem örneği: Bilgisayar imal eden bir işletme



- Eğitim bir sistemdir, çünkü birincisi gerçekleştirmek istediği hedefler vardır.
 - Uzak hedefler: Devletlerin hedefleridir, bazı ülkelerde çerçevesi anayasa ile güvenceye alınmıştır.
 - Genel hedefler: Daha işevuruk hedeflerdir. Kurumların hedefleri vb.
 - Özel hedefler: Dersin hedefleridir, bir dersin kazanımlarını kapsar.
- İkincisi bu hedefleri gerçekleştirme için farklı öğelerden oluşmaktadır.
 - Girdi: Eğitim-öğretim için gerekli her şey eğitim sisteminin girdisini oluşturur. Örneğin öğrenci, öğretmen, okul, çalışanlar, yönetim, eğitim programı vb.

- İkincisi bu hedefleri gerçekleştirme için farklı öğelerden oluşmaktadır.
 - b. Süreç: Girdilerin hedefler doğrultusunda bir araya getirildiği, dersin hedeflerinin öğrenciye kazandırıldığı eğitim durumlarını kapsar. Süreç, dar anlamda, öğretmenin farklı öğretim yöntemlerini kullanarak ders işlediği bölümü oluşturur. Geniş anlamda ise öğrencinin okula başladığı ve okuldan mezun olduğu süreci içerir.
- İkincisi bu hedefleri gerçekleştirme için farklı öğelerden oluşmaktadır.
 - c. Çıktı: Girdilerin işlemler bölümünden hedefler doğrultusunda biçimlendirilip oluşturulmasının sonunda ortaya çıkan ürünlerin tümüdür. Bir okulun mezun ettiği bireyler eğitim sisteminin çıktısını oluşturur. Çıktılar sistemin hedefleriyle nicelik ve nitelik açısından eşleşiyorsa sistem tam işliyor demektir.

EĞİTİM SİSTEMİ



Eğitimde kontrol nasıl sağlanır?

- Eğitimde kontrol (denetleme) ölçme ve değerlendirme ile mümkündür. Gözlenmeyen süreçler ve ürün değerlendirilemez.

Diğer taraftan içinde bulunduğumuz yüzyıl ile birlikte bir eğitimde ve dolayısıyla ölçme ve değerlendirmede bir paradigma değişimi söz konusudur.



21. YÜZYIL DEĞİŞEN EĞİTİM PARADİGMASI

20. Yüzyıla hakim olan kültürel iklim sanayileşme idi ve dolayısıyla eğitim bu iklimle uyumlu biçimde şekillenmişti. Özellikle ABD’de 1930-1950’lerde okullardan «eğitim fabrikaları» olarak söz ediliyor, eğitim de bu anlayışa uygun biçimde kurgulanıyordu. Ülkemizde de o tarihlerde benzer bir anlayışın olduğunu söylemek olanaklı.

Atatürk’ün ortaya koyduğu «muasır medeniyetler seviyesine ulaşmak» Türkiye’nin hedefi ve bir nevi eğitimin de uzak hedefi oldu. Eğitimin genel hedefi de pek çok ülkede kabaca «toplumun bireylerden beklediği görev ve sorumluluklar doğrultusunda gerekli bilgi ve donanımına sahip olma» olarak tanımlandı.

Geçtiğimiz yüzyılın özellikle son çeyreğinde bilgi ve iletişim teknolojisinin öneminin artması toplumlarda, üretimde ve eğitimde bir paradigma değişimine yol açtı. Artık sanayi toplumlarından bilgi toplumlarına hızlı bir dönüşüm baş gösterdi. Tanımlamalar, kavramlar değişti.



Sanayi Paradigması	Bilgi ve İletişim Teknolojisi Paradigması
Enerji yoğun	Bilgi yoğun
Çizim bürolarında yapılan tasarım ve mühendislik	Bilgisayar destekli tasarım
Ardışık tasarım ve üretim	Eş zamanlı mühendislik
Standart tasarım	Sipariş üzerine tasarım
Tahsis edilmiş tesis donanım	Esnek üretim sistemleri
Otomasyon	Sistemasyon
Tek firma	Ağ yapı
Hiyerarşik yapılar	Düz yapılar
Bölünmüş	Butünleşik

Sanayi Paradigması	Bilgi ve İletişim Teknolojisi Paradigması
Servisi olan üretim	Ürünleri olan servis
Merkezi yapı	Dağılmış yapı
Özgün beceri	Çoklu beceri
Hükümet kontrolü	Hükümetin bilgi, eşgüdüm ve üretimi yönlendirmesi
Planlama	Vizyon
Kaynak: DISK-AR, 1994	

- Üretimde yaşanan bu değişim eğitim öğretim ortamlarına da yansımaya başladı.

Sanayi Paradigması	Bilgi ve İletişim Teknolojisi Paradigması
Sınıflarda eğitim	Bireysel araştırma
Pasif öğrenme	Yaparak-yaşayarak öğrenme
Bireysel çalışma	Grupla çalışma
Her şeyi bilen öğretmen	Rehberlik eden öğretmen
Sabit içerik	Esnek içerik
Homojen	Heterojen

Diğer taraftan beşeri bilimler ve özellikle eğitim bilimlerinde bilişsel, duyuşsal ve devinimsel (psiko-motor) olarak uç boyutta tanımlanan insan davranışları, birbirinden ayrı kümelermiş gibi ele alındı, araştırıldı, hakkında bilgi üretilmeye çalışıldı.

Halbuki insan bu parçaların toplamından daha anlamlı bir varlıktı.

20. yüzyılın son çeyreğinde ve özellikle 21. yüzyılda ekonomi çevreleri ve ekonomistler eğitime yoğun bir ilgi göstermeye başladı. Makro ekonomiyi ve gelecek ekonomi perspektiflerini ele alan değerlendirmelerde eğitime vurgu yapmayan ekonomist pek kalmadı.

Ekonomistler son yıllarda eğitime neden daha çok vurgu yapıyorlar?

Araştırmalar paradigma değişimine uygun insan gücü kaynağının, ancak nitelikli eğitimle gerçekleşebileceğini ortaya koymakta. Niteliğin ölçülmesi çok zor ama nicel veriler açısından değerlendirdiğimizde, bir ülkede ortalama eğitim bir birim arttığında, ekonomik büyüme %0.7 artmakta. Bu büyümenin yarısı doğrudan eğitimin katkısı iken diğer yarısı eğitim ve ekonomi arasındaki etkileşimden kaynaklanmakta.

PISA, TIMSS ve PIRLS gibi dünyanın en geniş çaplı uluslararası eğitim araştırmalarını yapan örgütler ya doğrudan ekonomi ya da yine siyasi-ekonomik politikaların üretildiği örgütler.

Top 10 skills of 2025

	Analytical thinking and innovation	• Analitik düşünme ve yenilik/yaratıcılık
	Active learning and learning strategies	• Etkin öğrenme ve öğrenme stratejileri geliştirme
	Complex problem-solving	• Karmaşık problem çözme
	Critical thinking and analysis	• Eleştirel düşünme ve analiz
	Creativity, originality and initiative	• Yaratıcılık, orijinallik ve inisiyatif alabilme
	Leadership and social influence	• Liderlik ve sosyal etki
	Technology use, monitoring and control	• Teknoloji kullanımı, izleme/takip ve kontrol
	Technology design and programming	• Teknoloji desenleme ve programlama
	Resilience, stress tolerance and flexibility	• Strese tolerans ve bilişsel esneklik
	Reasoning, problem-solving and ideation	• Muhakeme becerisi ve problem çözme

Kaynak: Dünya Ekonomik Forumu Twitter Hesabı

Bu değişim doğrultusunda dünyanın biz eğitimcilerden beklediği görev ve sorumluluklar, eğitim ve öğretim ortamlarını yeniden tasarlamamıza ve denetleme biçimlerimizi farklılaştırmamıza neden oldu, bu değişim devam ediyor.

Eğitimde denetleme (kontrol) ölçme ve değerlendirme ile olanaklıdır, ancak değişen hedeflerde bireylere kazandırılacak olan yeni davranışların ölçülmesi ve değerlendirilmesinde de doğal olarak bir paradigma değişimi söz konusu.

Geleneksel ölçme anlayışının ne yazık ki, bu paradigma değişimine pek uygun olmadığı söylenebilir. Bireylerden beklenen becerilerin karmaşıklığı arttıkça ölçme modellerini revize etme gereksinimi ortaya çıkmakta.

Bir ölçme işleminde yalnızca «görünce tanıma, sorulunca söyleme» gibi davranışların ölçülmesi bizi yeni hedeflere ne kadar ulaştırabilir?

20. yüzyıl eğitim ve ölçme zihniyeti ile 21. yüzyıl becerileri kazandırılmaz.

2- ÖLÇME VE DEĞERLENDİRMENİN TEMEL KAVRAMLARI-2 (Sayfa: 71- 73)

EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ



ÖLÇME VE DEĞERLENDİRME

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

Fiziksel,
fizyolojik,
demografik,
psikolojik

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

Yapay uyarıcılara
maruz bırakarak
gerçekleştirilen
yapılandırılmış bir
gözlem

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

Ölçüm=Ölçme

sonucu

Örneğin 60, CC, orta

vb.

Örneğin

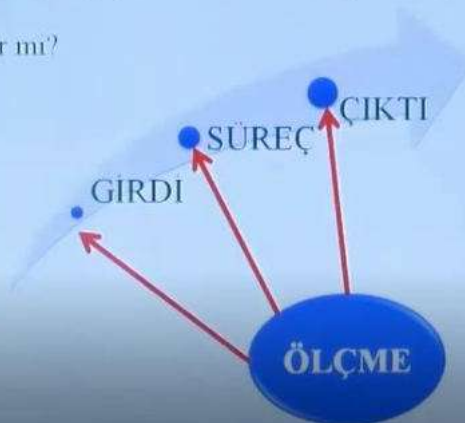
- Kerim Temel Yeterlilik Testinden (TYT) 450 puan aldı.
- Can'ın genel yetenek testi sonucu, zeka bölümünün 132 olduğu belirlendi.
- Zeynep yapılan test sonucunda 1. oldu.

Ölçme bir betimlemedir.

Ölçme tek başına bir anlam ifade etmez.

- Ölçmeye temel olan durum «fark»tır. Ölçme farktan doğmuştur. Fark
 - bireyler arasında olabileceği gibi
 - bireylerin farklı zaman dilimlerinde ortaya koydukları özellikler arasında olabilir.
- Eğitimde ölçme dediğimizde aklımıza ilk gelen genellikle sınavlardır. Eğitimde ölçme yalnızca sınavlardan mı ibarettir?
 - Başarı
 - Yetenek
 - İlgi
 - Tutum
 - Özel gereksimin gereken alanlar

Bir eğitim sisteminde ölçme yapılmısa da olur mu ya da ölçme olmadan bir eğitim sistemi kurgulanabilir mi?



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

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

Örneğin 60 alan geçer.

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

Örneğin 90-109 arası normal zekadır,
DC alan şartlı geçer (aritmetik
ortalama - $\frac{1}{2}$ standart sapma) vb.

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

1. Ölçme $\Rightarrow$ Ölçme sonucu (ölçüm)
2. Ölçüt
3. Karar

Örneğin

- Kerim gıda mühendisliği programına yerleştirilmiştir / yerleştirilememiştir.
- Can üstün yetenekliler sınıfına kabul edilmiştir / edilmemiştir.
- Zeynep dersten geçmiştir / kalmıştır.
- Mehmet öğretmen olarak atanmıştır / atanamamıştır.
- Hakan burs almaya hak kazanmıştır / burs alamamıştır.

- Ölçüt ölçme ve değerlendirme arasında köprü görevi görür. Ölçüt değişirse değerlendirme de değişir.

Örneğin

Ölçme	Ölçüt	Değerlendirme
Irmak Yabancı Dil Bilgisi Seviye Tespit Sınavından (YDS) 50 almıştır.	50	Yüksek lisansa başvuru yapabilir.
Irmak Yabancı Dil Bilgisi Seviye Tespit Sınavından (YDS) 50 almıştır.	55	Doktora başvuru yapamaz.



ÖLÇME TÜRLERİ

- Eğitimde ölçme bağlamında ölçme üçe ayrılır. Bunlar:

1. Doğrudan ölçme (Temel ölçme)
2. Dolaylı ölçme (Göstergeyle ölçme)
3. Türetilmiş ölçme

1. Doğrudan Ölçme (Temel Ölçme): Ölçmeye konu olan özelliğin doğrudan gözlenerek ölçümün elde edildiği ölçme türüdür. Genellikle beş duyu organı ile algılanan özellikler doğrudan gözlenebilirler. Örneğin
– bir sınıftaki öğrenci sayısı, parktaki salıncak sayısı, öğrencilerin cinsiyetleri, boyları, kalem tutma becerisi, pas verme becerisi vb.
özellikler doğrudan gözlenebilir

2. **Dolaylı Ölçme (Göstergeyle Ölçme):** Bazı özellikler doğrudan gözlenemezler. Bu özellikler, ancak onun göstergesi olduğu bilinen ya da kabul edilen davranışlar aracılığı ile gözlenerek ölçülebirlir. Örneğin

- Öğrencilerin dört işlem becerisi, derse yönelik tutumları, genel muhakeme düzeyleri, öz-düzenleme becerileri, sınav kaygıları, motivasyon düzeyleri vb. özellikler ancak dolaylı olarak gözlenebilir.

Eğitimde ölçmeye konu olan özelliklerin pek azı doğrudan ölçülebirlir. Söz gelimi yanlışsız okuyabilme, diksiyon, pek çok psiko-motor davranış doğrudan gözlenebilir. Diğer taraftan bir öğretmen adayına «sınıfta durumıyla karşılaştığınızda nasıl davranırsınız?» diye sorduğunuzda, vereceği sözlü ya da yazılı tepkinin gelecekte nasıl davranacağının bir göstergesi olduğu kabul edilir.

Benzer biçimde «iyi yurttaş» olma doğrudan gözlenebilir bir özellik değilken, iyi yurttaş olmanın göstergesi olarak kabul edilen «vergi verme, seçme ve seçilme hakkını kullanma, askerlik vb. görevleri yapma» gibi davranışları sergileyen kişi «iyi yurttaş» olarak nitelenebilir.

Diğer taraftan 18 yaşının altında bir öğrencinin bu özelliklere sahip olma durumu, varsayımsal durumlar üzerinden gözlenebilir.

3. **Türetilmiş Ölçme:** Ölçülmek istenilen özellik kendisinden farklı iki ya da daha fazla özelliğin arasındaki matematiksel bir bağıntı (dört işlem) yardımıyla belirleniyorsa bu ölçme türüne türetilmiş ölçme adı verilir. Örneğin

- hız (yol / zaman), yoğunluk (kütle / hacim), zeka bölümü (zeka yaşı / takvim yaşı x 100) vb.

özellikler türetilmiş ölçmeye örnek olarak verilebilir.

• Eğitimde ölçüt ikiye ayrılır. Bunlar:

1. Mutlak Ölçüt (Kriter Referanslı Ölçüt)
2. Bağıl Ölçüt (Norm Referanslı Ölçüt)

3- ÖLÇME VE DEĞERLENDİRMENİN TEMEL KAVRAMLARI-3 (Sayfa: 73-75)

EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ



ÖLÇÜT TÜRLERİ

- Eğitimde ölçüt ikiye ayrılır. Bunlar:

1. Mutlak Ölçüt (Kriter Referanslı Ölçüt)
2. Bağıl Ölçüt (Norm Referanslı Ölçüt)

1. Mutlak Ölçüt (Kriter Referanslı Ölçüt): Eğer ölçüt ölçme işleminin yapıldığı grubun özelliklerinden bağımsız olarak belirleniyorsa mutlakdır. Mutlak ölçüt genellikle ölçme işlemi öncesinde ilan edilir. Örneğin

- *dersten geçme notu 50 'dir, doktora başvurusu için ALES'ten en az 60 almak gerekir, sürücü ehliyeti sınavında 70 alan başarılı sayılır vb.*

Ölçüt daha önceden belirlenmiş ve değişmez standart(lar) olarak saptanmıştır, diğer bir deyişle kuramsal bir yetkinlik standardı esas alınır. Örneğin *«tam öğrenme kuramı doğrultusunda 100 üzerinden 70, başarılı olma koşuludur» vb.*

Mutlak ölçütün temel alındığı ölçme işlemlerinde bir öğrencinin notu, diğer bir öğrencinin değerlendirmesini (geçme/kalma) etkilemez. Dolayısıyla örneğin başarının ölçüldüğü bir örnekte, herkes geçebileceği gibi, herkes de kalabilir

1. Bağıl Ölçüt (Norm Referanslı Ölçüt): Eğer ölçüt ölçme işleminin yapıldığı grubun belirli bir özelliğine / normuna dayalı olarak belirleniyorsa bağıldır. Bağıl ölçüt ancak ölçme işlemi sonrasında belirlenebilir. Örneğin

- *aritmetik ortalamaya / ortancaya / moda denk ve üzerinde puan alan geçer;*
aritmetik ortalama ile yarım standart sapma arasında puan alanlar şartlı geçer;
akademik başarıları yüksek ilk on öğrenci proje için seçilecektir vb.

Bağıl ölçütün temel alındığı ölçme işlemlerinde bir öğrencinin başarısı, başka bir öğrencinin değerlendirmesini (geçme / kalma) etkileyebilir. Bu anlamda öğrenciler açısından daha rekabetçi bir eğitsel ortam kaçınılmazdır.

- Eğer değerlendirme mutlak ölçüte göre yapılıyorsa mutlak değerlendirme, bağıl ölçüte göre yapılıyorsa bağıl değerlendirme adını alır. Örneğin
 - Selim fen lisesine yerleştirilmiştir. → Bağıl değerlendirme
 - Güzin lisans programları için tercih yapabilecektir. → Mutlak değerlendirme

- Eğitsel kararların verilmesinde mutlak ve bağıl ölçütler bir arada da kullanılabilir.

Örneğin

- Yükseköğretim Kurumları Sınavı
- KPSS Orta Öğretim Memur Atamaları
- Ortalama 40-60 ise mutlak ya da bağıl değerlendirme yapılır, ortalama 60'ın üzerindeyse mutlak değerlendirme, 40'ın altındaysa bağıl değerlendirme zorunludur.

- Öğrenci ile ilgili önemli / hayati kararlar verileceği zaman mutlak değerlendirme yapılması gerekir. Örneğin

- Tek ders sınavları

- Başvuran kişi sayısının çok alınacak kişi sayısının az olduğu (arz-talep dengesizliği olan) sınavlarda yerleştirme / atama vb. bağıl değerlendirme ile yapılmak zorundadır.

- LGS, YKS, KPSS, TUS vb.

- Muafiyet sınavlarında mutlak değerlendirme yapılması gerekir. Örneğin

- Yabancı dil muafiyet sınavları



AMACA GÖRE DEĞERLENDİRME TÜRLERİ

- Amaca göre değerlendirme beşe ayrılır. Bunlar:
 1. Tanıma-Yerleştirmeye Yönelik Değerlendirme (Diyagnostik Değerlendirme)
 2. Biçimlendirme-Yetiştirmeye Yönelik Değerlendirme (Formatif Değerlendirme)
 3. Değer Biçme / Düzey Belirlemeye Yönelik Değerlendirme (Summatif Değerlendirme)
 4. Rehberlik Amaçlı Değerlendirme
 5. Program Değerlendirme

1. Tanıma-Yerleştirmeye Yönelik Değerlendirme

Sınıf içi ölçme ve değerlendirme. Öğretim sürecinin başında, öğrencilerin önceki kritik nitelikteki öğrenmelerini belirlemek amacıyla yapılır. Diğer bir ifadeyle amaç öğrencinin ders bağlamındaki hazırbulunuşluluk düzeyini belirlemektir.

- Amaç not vermek değildir
- Bu amaçla uygulanan ölçme araçlarına da hazırbulunuşluluk testi adı verilir.

1. Tanıma-Yerleştirmeye Yönelik Değerlendirme

Okul / kurum geneli ölçme ve değerlendirme. Yine öğretim sürecinin başında, öğrenciyi tanımak ve onun için uygun olan programa / kura / sınıfa yerleştirmektir.

- Amaçlanan not vermek değildir.
- Hedellenen öğrencinin niteliklerine uygun grupların oluşturulmasıdır.
- Zaman, para ve emek tasarrufu sağlar.
- Muafiyet sınavları ve seviye tespit sınavları buna örnek gösterilebilir.

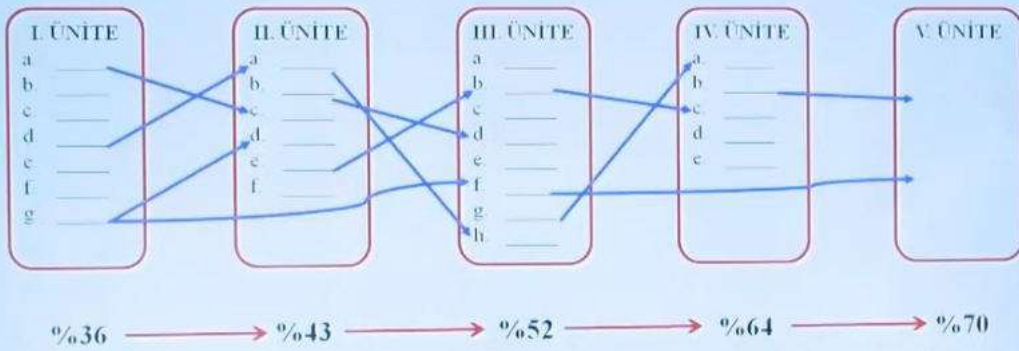
2. Biçimlendirme-Yetiştirmeye Yönelik Değerlendirme

Belirli bir konu / unite vb. sonunda öğrencilerin öğretime konu olan davranışları ne kadarını kazandığını, diğer bir deyişle unitedeki öğrenme eksiklerini belirlemek amacıyla yapılır.

- Amaç yine not vermek değildir.

2. Biçimlendirme-Yetiştirmeye Yönelik Değerlendirme

- Bu değerlendirme türünde öğrenciye o ünitedeki öğrenme eksikleri bildirilir, bir sonraki kapsama geçmeden öğrenme eksikleri giderilir ya da giderilip giderilmediği denetlenir.
- Öğrenme eksikleri giderilmezse ne olur?



Öğrenme eksikliği kar topu gibi büyür.

- Biçimlendirme-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.

3. Değer Biçmeye / Düzey Belirlemeye Yönelik Değerlendirme

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

Öğrenci yapılan her ölçme işleminden sayı değeri ya da sembol alır. Her sayı değeri not anlamına gelmez. Eğer puan öğrencinin geçme/kalma vb. bir kararla kullanılıyorsa o ölçüme not adı verilir.

- Değer biçmeye yönelik değerlendirme öğrencinin belirli bir öğretim süreci sonundaki durum tespittir, çünkü artık hedeflenen öğretim sürecinin sonuna gelinmiştir.
- Bu amaçla uygulanan testlere alanda *eriş test*i denir.
- Bitirme sınavları, sertifika sınavları, üniversitedeki vize ve finaller, Millî Eğitimde not vermek amacıyla uygulanan yazılı yoklamalar bu ölçme ve değerlendirme türüne örnek olarak verilebilir.

İlk üç değerlendirme türü öğrenci başarısına odaklıdır, diğer bir ifadeyle öğrencinin akademik başarısının ölçülüp değerlendirilmesi söz konusudur. Ancak eğitimde yalnızca akademik başarının ölçülüp değerlendirilmesi ile ilgilenilmez. Öğrencinin yetenekleri, ilgileri, tutumları, kaygıları, öğrenme kapasiteleri ve karşılaşılan güçlükler ile uygulanan eğitim programının etkililiğinin değerlendirilmesi de çok önemlidir.

4. Rehberlik Amaçlı Değerlendirme:

- Özel eğitim: Belirli alanlarda özel gereksinimi olan çocuklara uygun eğitim ortamları ve/veya programı uygulamak amacıyla ölçme ve değerlendirme süreçleri işe koşulmaktadır. Özel gereksinim kavramı zihinsel gerilikten öğrenme güçlüğüne, otizmden diğer spesifik gerilik alanlarına, üstün zekadan işitme/görme vb. engel alanlarına kadar çok geniş bir yelpazede ele alınır.

4. Rehberlik Amaçlı Değerlendirme:

- Özel eğitim: Özel gereksinimi olan çocuklara uygun eğitim olanaklarının sunulabilmesi için, çocuğun ihtiyacı olan alanlar, eğitsel tanılama amacıyla genellikle standart testler veya ölçüt bağımlı testler (beceri öğretimde) ile 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.

4. Rehberlik Amaçlı Değerlendirme

- 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.

4. Rehberlik Amaçlı Değerlendirme

- Mesleki rehberlik: Böylece öğrencinin kendi özellikleri hakkındaki farkındalıklarının artırılması ve dolayısıyla buna uygun alan / meslek seçmeleri sağlanmış olur.

5. 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 ve bilimsel araştırma sürecinin adımlarıyla büyük ölçüde benzerlik gösterir.

Program değerlendirmede program taslağının hazırlanması, programın uygulanması ve programın etkisine ilişkin yargılar elde edilir.



GENEL DEĞERLENDİRME

Bu bölümde içinde ölçme ve değerlendirmenin eğitim sistemi içindeki rolü, içinde bulunduğumuz yüzyılda eğitimin ve ölçme-değerlendirmenin değişimi, ölçme, değerlendirme ve ölçütlerle ilgili temel kavramlar ile türler ve son olarak eğitim sistemi içinde değerlendirmenin amaca göre farklı türleri üzerinde durduk.

4- BİR ÖLÇME ARACINDA BULUNMASI GEREKEN PSİKOMETRİK NİTELİKLER-1 (Sayfa: 75-77)

EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ



İÇERİK

- Ölçme araçlarında bulunması gereken psikometrik nitelikler
 - Geçerlilik, güvenilirlik ve kullanılabilirliğin kavramsal temelleri
 - Ölçme ve hata kavramı
 - Geçerlilik türleri
 - Kapsam geçerliliği
 - Yordama geçerliliği
 - Zamandaş geçerlilik
 - Yapı geçerliliği
 - Güvenilirlik kestirim yöntemleri
 - Test-tekrar test
 - Test yarılama
 - KR20 ve KR21
 - Cronbach alfa
 - Güvenilirliği ve geçerliliği artırma yolları



PSİKOMETRİK NİTELİKLER

Ölçme her zaman belirli bir amaç için yapılır. Amaç ölçmeye konu olan psikolojik özellik bakımından kişiler hakkında değerlendirme yapmak ve elde edilen değerlendirme sonuçlarına göre belirli kararlar vermek için yapılır. Verilen kararların doğru ve isabetli olması, değerlendirmenin dayandığı ölçümün özellikle doğrudan ilgili ve olabildiğince az hatalı olmasına bağlıdır.

Psikolojik özelliklerin ölçülmesinde genellikle ölçme araçları kullanılır ve bu araçlardan elde edilen puanların hatasız ya da az hatalı olması ve bu araçların belirli psikometrik niteliklere sahip olması gerekir. Bunlar:

- 1) **Geçerlilik (Validity)**
- 2) **Güvenilirlik (Reliability)**
- 3) **Kullanılabilirlik (Practicability)**

Geçerlilik ölçmek istenilen özelliğin, başka özelliklerle karıştırmadan, doğru ve tam olarak ölçülebilmesidir. Diğer bir ifadeyle ölçme aracından elde edilen puanın amaca hizmet etme derecesidir.

Güvenilirlik ölçme işleminden elde edilen puanların tesadüfi hatalardan arınık olma derecesidir. Diğer bir ifadeyle puanların kararlı, tutarlı ve duyarlı olmasıdır.

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 işlemini yapan kişi neye müdahale edeceğini, hangi durumlara karşı önlem alması gerektiğini bilsin.



HATA

Yalnızca eğitimde değil bütün bilim dallarında, hatta günlük yaşamımızda dahi yapılan ölçme işlemlerinde bile ölçme sonuçlarına hata karışma olasılığı vardır.

Ölçme yapan kişilerin dikkati ve titizliği zamandan zamana değişebilir, ölçme aracına ilişkin bazı sorunlar söz konusu olabilir, ölçülen özelliğin doğası gereği bazı sıkıntılar olabilir, test katılımcısı yeterince güdülenmemiş olabilir, ölçme işleminin yapıldığı ortam kaynaklı sorunlar olabilir vb.

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. Bu eşitlik:

$$X = T + E$$

$$X = T + E$$

X = Bireylerin ölçme aracından elde ettiği gözlenen puanı (?)

T = Bireylerin gözlenemeyen gerçek puanı (?)

E = Ölçmeye karışan hata miktarı (?)



HATA TÜRLERİ

Bu çerçevede hatalar kaynağı, yönü ve miktarı göz önünde tutularak üç türde incelenebilir:



1) Sabit hata (constant error)



2) Sistemantik hata (systematic error)



3) Tesadüfi (rastlantısal) hata (random error)



SABİT HATA

Miktarı ölçmeden ölçmeye değişmeyen, diğer bir deyişle her ölçme işlemine aynı miktarda karışan hatalardır. Örneğin marketteki terazi üzerinde herhangi bir nesne yokken -120 gr. gösteriyorsa ne tartılırsa tartılsın 120 gr. eksik ölçülecektir. Bir öğretmen sınavında herkese 10 puan fazla veriyorsa yine karışan hata sabit olacaktır.



SİSTEMATİK HATA

Ölçülen büyüklüğe, öğretmene ya da ölçme koşullarına göre miktarı değişen hatalardır. Örneğin marketteki terazi her bir kilogramda 120 gr. eksik tartıyorsa üzerine konulan nesnenin ağırlığı arttıkça hata miktarı da artacaktır. Bir öğretmen sınavında yazısı kötü olandan puan kırıyorsa yine karışan hata sistematik olacaktır.



SABİT ve SİSTEMATİK HATA ÖRNEKLERİ

Sabit ve sistematik hatalar zaman zaman birbirine karıştırılabilir.

Örnek	Sabit Hata	Sistematik Hata
Öğretmenin sevdiği öğrenciye puan vermede bonkor davranması		×
Öğretmenin 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	×	
Terzinin mezurasının 150 cm yerine aslında 145 cm olması	×	×

Gerek sistematik gerekse sabit hatalarda ölçme sonuçlarına karışan hata miktarı, yönü ve kaynağı bellidir. Bu nedenle bu tür hataların ölçme sonuçlarına karışmasını engellemek ya da düzeltmek nispeten daha kolaydır.



TESADÜFİ HATA

Şansla ortaya çıkan ne yönde ve ne ölçüde karıştığı genellikle bilinemeyen hatalardır. Ölçmelere tek yönlü olarak karışmazlar, ölçme sonuçlarına bazen pozitif bazense negatif yönde etki ederler. Sabit ve tesadüfi olmayan hatalar tesadüfi değişken olmayan özelliğine sahiptir. Bu nedenle psikometride hata kuramları tesadüfi hatalar üzerine kurulmuştur.

Tesadüfi hataların genellikle dört kaynağı olduğu kabul edilir. Bunlar:

- 1) **Ölçme işlemi yapan kişi (öğretmen)**
- 2) **Ölçme aracı**
- 3) **Birey (öğrenci)**
- 4) **Fiziksel ortam**

1. Ölçme işlemi yapan kişiden / öğretmenden kaynaklanan hata: Öğretmenin test etme sürecine ilişkin davranışlarını iki aşamada ele alabiliriz:

- a. Ölçme işlemi sürecinde: Test katılımcısının dikkatini dağıtacak ve/veya kaygısını arttıracak davranışlar vb.
- b. Ölçme işlemi sonrasında: Puanlamadaki dikkat ve titizliğin zamandan zamana değişmesi, yorgunluk, öncelik-sonralık yanılgısı, maddi hata vb.

2. Ölçme aracından kaynaklanan hata: Ölçme araçları hazırlanırken maddelerin iyi ifade edilmemesi, test katılımcılarının yanlış anlamalarına ve dolayısıyla hataya neden olacaktır. O nedenle hangi özellik ölçülürse ölçülsün, araçta kullanılan dilin dilbilgisi, imlâ kurallarına uygun olması, maddelerde anlatım bozukluğu bulunmaması, muğlak ya da belirsiz (müphem) ifadelerin bulunmaması gerekir.

2. Ölçme aracından kaynaklanan hata: Ölçme aracından kaynaklanabilecek bir başka hata kaynağı ise aracın yapısı ile ilgilidir. Ölçtüğü özellik ve kapsam bakımından homojen (benzeşik) maddelerden oluşan bir araç, heterojen (ayrışık) maddelerden olan bir testten daha güvenilirdir. Diğer taraftan ölçülecek özellikler evrenini yeterince temsil etmeyen maddelerden oluşan bir araçtan elde edilen puanlara da hata karışmaktadır.

2. Ölçme aracından kaynaklanan hata: Ölçme aracının yapısıyla ilgili diğer bir konu ise aracın uzunluğudur. Madde sayısı ile güvenilirlik arasında doğru orantılı bir ilişki vardır, ancak bu sonsuz bir doğru orantı anlamına gelmez. Ölçme aracındaki madde sayısı arttıkça bireyde yorgunluk, dikkat azalması vb. etkenler nedeniyle hata miktarı artacaktır.

2. Ölçme aracından kaynaklanan hata: Ölçme aracıyla ilgili diğer bir faktör ise bilişsel özellikleri ölçen testlerde eğer seçenek verilmişse şans başarısı karışma olasılığıdır. Bu durum ölçülen özelliğin gerçekte olduğundan yüksek görünmesine yol açmaktadır.

Salt tahminle doğru cevabı bulma olasılığı

3. **Bireyden / öğrenciden kaynaklanan hata:** Bireylerin ölçme işlemi sürecinde içinde bulundukları fiziksel, fizyolojik ve psikolojik durumlar test puanına etki eder.

Uykusuzluk, açlık, hastalık, ağrı, motivasyon eksikliği, kaygı vb. faktörler nedeniyle bireyler gerçek puanlarından uzaklaşabilirler.

Ayrıca grubun homojen ya da heterojen olması da hata kaynağını arttıran ya da azaltan bir etkiye sahiptir.

4. **Fiziksel ortamdan kaynaklanan hata:** Ölçme işleminin gerçekleştiği fiziksel ortama ilişkin bazı etmenler de ölçme sonuçlarına hata karıştırabilir. Sıcaklık, ışık, ses, koku, görsel uyarıcılarının fazlalığı, ergonomi vb. etmenler hata miktarını arttırabilir.

Bu anlamda testler bireylere bu etmenler açısından eşit ve standart koşullar altında uygulanmalıdır, aksi takdirde hata miktarının artması kaçınılmazdır.

SEVGİNİN

5- BİR ÖLÇME ARACINDA BULUNMASI GEREKEN PSİKOMETRİK NİTELİKLER-2 (Sayfa:77-79)

EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ

GEÇERLİLİK ve GÜVENİLİRLİK İLİŞKİSİ



Şekil 1: Geçerlilik, güvenilirlik ve hata ilişkisi

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.

Bir ölçme aracının geçerliliği ve güvenilirliği diye bir şey yoktur, ölçme araçlarından elde edilen puanların geçerliliği ve güvenilirliği diye bir şey vardır. Ölçme araçlarının psikometrik nitelikleri şu durumlara göre değişir:

- ☐ Ölçme amacının değişmesi
- ☐ Uygulama grubunun değişmesi
- ☐ Dilin eskimesi / değişmesi
- ☐ Maddelerde ve veya alt ölçeklerde yapılan değişiklikler
- ☐ Farklı kültürler
- ☐ Kuramsal bilgi birikiminde değişiklikler

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.

Bir ölçme aracının puanlarından yapılan değerlendirmenin doğruluğu araştırmayı gerektirir. Bu puan gerçekten ölçmek istediğimiz özelliği gösteriyor mu?

Aracın ölçmek istediğimizi ölçüp ölçmediği bir araştırma işidir. Araştırma ile değerlendirmenin doğruluğu incelenir. Bir testin şöhretli olması onun geçerli olduğu anlamına gelmez. Modern test kuramına göre geliştirilen ölçekler eleştiriye açıktır. Böylece testlerin niteliği ve test bilgisi gelişir.

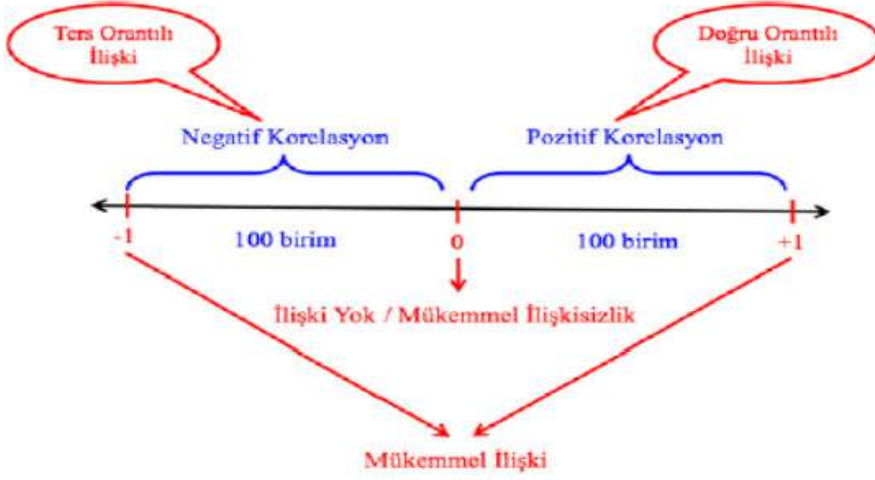


KORELASYON

Geçerlilik ve güvenilirliği belirlemeye yönelik yöntemler genellikle korelasyon temelli analizlere dayanır. Bu nedenle kısaca korelasyon kavramından bahsetmek gerekmektedir.

Korelasyon (co-relation) en az iki değişken arasında karşılıklı bir ilişki bulunup bulunmadığı, eğer ilişki varsa bu ilişkinin yönü ve miktarı hakkında bilgi veren istatistik bir tekniktir. r ile sembolize edilir. Örneğin

- Öğrencilerin derse ilişkin tutumları ile ders başarıları arasında bir ilişki var mıdır?
- Öğretmenlerin ders saati yükleri ile iş doyumları arasında bir ilişki var mıdır?
- Saç uzunluğu ile zeka arasında bir ilişki var mıdır?



X	Y	X	Y	X	Y
2	9	2	27	2	5
4	12	4	24	4	5
5	13	5	21	5	5
8	16	8	17	8	5
9	17	9	16	9	5
11	21	11	13	11	5
12	24	12	12	12	5
15	27	15	9	15	5

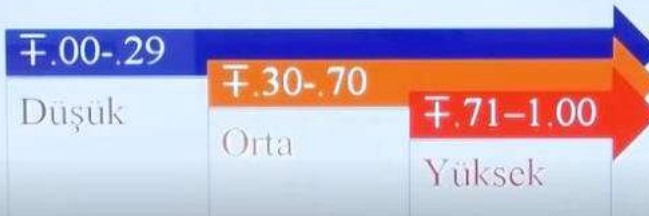
$r = 1.00$

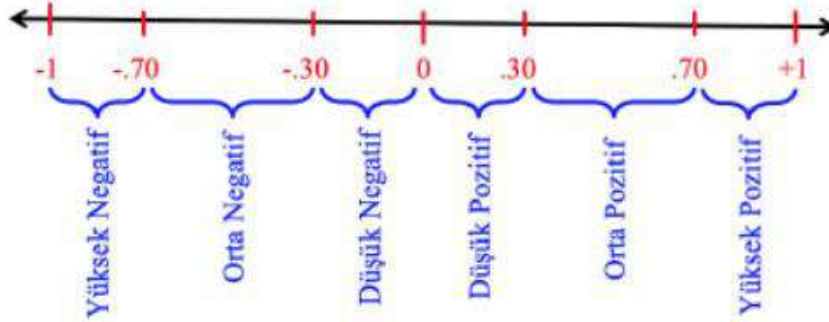
$r = -1.00$

$r = 0.00$

Korelasyon katsayısı hakkında kabaca iki tür belirleme yapmak gerekir.

1. **Yönü** (pozitif mi? negatif mi?)
2. **Miktarı** (düşük mü? orta mı? yüksek mi?)





Şekil 3: Korelasyonun miktarı

DİKKAT: Korelasyon mutlak değer olarak değerlendirilmelidir. Bir korelasyon katsayısının negatif ya da pozitif olması büyüklük-küçüklük belirtmez, yön bildirir.

DİKKAT: Korelasyon katsayısı ile neden-sonuç ilişkisi kurulamaz.



Şekil 4: Geçerlilik yöntemleri sınıflaması



KAPSAM GEÇERLİLİĞİ

Kapsam geçerliliği özellikle başarı testlerinde aranan bir geçerlilik sorgulamasıdır.

Eğitimde öğretmen bir program dahilinde önceden belirlenen davranışları kazandırmayı amaçlar. Belirli bir zaman sonra davranışların ne kadarının kazanıldığını belirlemek isteyen öğretmen, geliştirdiği testte öğretime konu olan içeriğin testte ne ölçüde temsil edildiğini belirlemek durumundadır.

Bu anlamda kapsam geçerliliği bir testin ölçülmek istenen davranışları ne derece kapsadığıyla ilgilidir. Ne amaçla kullanılırsa kullanılsın test kapsamı açısından, ölçmeye konu olan davranışları yeterli ve dengeli bir biçimde temsil etmelidir.

Temsil edilmezse ne olur?

Bir testin kapsam geçerliliğinin yüksek olduğunun söylenebilmesi için

1. Testteki soruların / maddelerin ölçülecek özellikler evrenini (konu kapsamını / içeriği) yeterli ve dengeli bir biçimde ölçüyor olması ve
2. Her bir sorunun / maddenin ölçmek istediği özelliği doğrudan ölçmesi, diğer bir deyişle kazanımla doğrudan ilgili olması gerekir.

Bu iki kriter birden sağlanmalıdır.

6- BİR ÖLÇME ARACINDA BULUNMASI GEREKEN PSİKOMETRİK NİTELİKLER-3 (Sayfa:79-82)

EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ

SEVGİNİN SESİYLE



KAPSAM GEÇERLİLİĞİ

Kapsam geçerliliği sorgulama yöntemleri:

1. Mantıksal / Rasyonel Yöntemler
 - a. Belirtke Tablosu Hazırlanması
 - b. Uzman görüşüne Başvurulması
2. İstatistiksel Yöntemler

1. Mantıksal / rasyonel yöntemler → **Belirtke tablosunun hazırlanması**: Bir kapsam geçerliliği sorgulamasında öncelikle ölçmeye konu olan kapsam dahilinde davranışların belirlenmesi gerekir. Bu noktada en çok kullanılan yöntemlerden biri belirtke tablosu hazırlamaktır. Öğretmen satırda davranışları, sütunda hedefleri yazar; belirlediği madde sayısı doğrultusunda davranış ve hedefleri yeterli ve dengeli bir biçimde temsil eden alanları seçer.

Tablo 3. *Belirtke tablosu örneği*

	Hedefler							
Konular	Bilgi	Kavrama	Uygulama	Analiz	Sentez	Değerlendirme	Toplam	
1. Ölçme ve değerlendirmenin temel kavramları							%43,3	
1.a. Ölçme ve değerlendirmenin eğitim sistemi içindeki yeri		1				1	2	
1.b. Eğitimde paradigma değişimi ve bu değişimin ölçme ve değerlendirme uygulamalarına yansımaları	1			2			3	
1.c. Ölçme, değerlendirme ve ölçütün kavramsal temelleri		2					2	
1.d. Ölçme türleri		1					1	
1.e. Ölçüt türleri	1					1	2	
1.f. Değerlendirmenin amaca göre türleri		1			1	1	3	
2. Bir ölçme aracında bulunması gereken psikometrik nitelikler							%56,7	
2.a. Geçerlilik, güvenilirlik ve kullanılabilirliğin kavramsal temelleri		1					1	
2.b. Ölçmede hata kavramı	1						1	
2.c. Hata türleri		1		1			2	
2.d. Geçerlilik ve güvenilirlik ile hata ilişkisi						1	1	
2.e. Geçerlilik ve güvenilirlikle ilgili temel kavramlar		1	2				3	
2.f. Geçerlilik sorgulama yöntemleri		2		1			3	
2.g. Güvenilirlik kestirim yöntemleri	1	1	1			1	4	
2.h. Geçerliliği ve güvenilirliği artırma yolları		1		1			2	
TOPLAM	4	12	4	5	1	5	30	

1. Mantıksal / rasyonel yöntemler → Uzman görüşüne başvurulması. Uzman ile kastedilen öncelikle ölçme ve değerlendirme tekniklerini de bilen bir alan uzmanıdır. Uzmanla belirtke tablosu ve maddeler sunulur, soruların / maddelerin konu kapsamını yeterli ve dengeli bir biçimde ölçme durumunu ve soruların / maddelerin kazanımlarla doğrudan ilgili olma durumunu değerlendirmesi istenir.

2. İstatistiksel yöntemler: Uzmanlardan alınan dönütler betimsel / muhakemeye dayalı bir yolla çözümlenebilir ya da uzmanların «*uygundur* / *uygun değildir*» ya da «*uygundur* / *düzeltilme gerekir* / *soru kullanılmamalıdır*» vb. biçimde soruları değerlendirmesi istenebilir. Eğer uzmandan ikinci yolla dönüt istenmişse uzmanlar arasında uyum olup olmadığı, çeşitli istatistiksel yöntemlerle test edilir. Alanda uzmanlar arasında uyumu test eden pek çok indeks bulunmaktadır.



ÖLÇÜT DAYANAKLI GEÇERLİLİK

Ölçme aracından elde edilen puanların ölçüt bir puanla (testin tahmin etmeye çalıştığı ve geçerliliği yüksek bir puan) karşılaştırılarak, geliştirilen ölçme aracının geçerliliğine ilişkin nitelendirme yapılır.





YORDAMA GEÇERLİLİĞİ

Yordama nedir?

Her tahmin bir yordama mıdır?

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 iki amaçla uygulanan ölçme araçlarının yordama geçerliliğinin yüksek olması istenir. Bunlar seçme ya da yönlendirme amaçlı kullanılan araçlardır.

Özellikle iki amaçla uygulanan ölçme araçlarının yordama geçerliliğinin yüksek olması istenir. Bunlar **seçme** ya da yönlendirme amaçlı kullanılan araçlardır.

Örneğin

YKS

KPSS

LGS

vb.

Özellikle iki amaçla uygulanan ölçme araçlarının yordama geçerliliğinin yüksek olması istenir. Bunlar seçme ya da **yönlendirme** amaçlı kullanılan araçlardır.

*Örneğin alan seçiminde ve/veya
bir üst öğretim kurumuna
yönlendirmek için uygulanan
yetenek testleri, ilgi envanterleri,
mesleki kişilik envanterleri vb.*

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 GEÇERLİLİĞİ ÇALIŞMA ÖRNEKLERİ

	KPSS Puanları	Ölçüt Puan İş Performansı
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

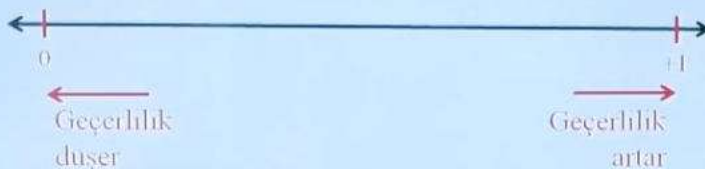
	YKS Puanı	Akademik Ortalama
Tarık	438.636	3.54
Esra	348.545	2.45
Begüm	456.344	3.39
Ozan	399.987	2.23
Defne	382.654	2.34
.....	480.334	3.45
.....	287.987	1.67
.....	381.954	3.88
.....	250.897	3.91
Umut	343.019	3.26

Ölçüt gelecekte belli olacağından beklemek gerekmektedir.

Korelasyon (r) 1'e yakın

Geçerlilik Katsayısı

Yordama geçerliliğinde geçerlilik katsayısının 0 ile 1 arasında değişmesi beklenir



Yordama geçerliliğinde en zor ve önemli nokta ölçütün doğru bir biçimde belirlenmesidir. Ölçütün belirlenmesinde şu noktalar dikkate alınmalıdır: Ölçüt

1. ölçme aracının yordamaya çalıştığı değişkenle doğrudan ilgili olmalı, ölçme aracı hangi özelliği kestirmeyi amaçlıyorsa onun doğrudan bir temsili olmalıdır.
2. kararlı olmalı, günden güne değişmemelidir. Açıktır ki kendisi kararsız olan bir özellik hiçbir araçla yordanamaz.

Ölçüt(ün)

3. bireylerin özelliğini gerçekten yansıtan nesnel ve güvenilir bir ölçü olmalıdır.

Sözelimi okulda alınan notlar bir ölçüt olarak alınmışsa öğrencilere verilen notlara başarının dışındaki etmenler etki etmemelidir.

4. elde edilmesi kolay ve ekonomik olmalıdır.



ZAMANDAŞ GEÇERLİLİK

Bu geçerlilik türü alanyazında halı hazır geçerlilik, benzer ölçekler geçerliliği, uygunluk geçerliliği adı ile de anılmaktadır.

Zamandaş geçerlilik sorgulamalarında ölçüt puan eş zamanlı olarak elde edilebilir.

Geliştirilen ölçme aracı ile ilişkili olabilecek nitelikleri ölçen ve geçerliliği yüksek bir aracın puanı ölçüt puan olarak ele alınabilir. Daha sonra geliştirilen ölçme aracı ve ölçüt araç aynı zamanda uygulanıp, iki araçtan elde edilen puanlar arasındaki ilişki incelenir.

ZAMANDAŞ GEÇERLİLİK ÖRNEKLERİ

	Benlik Algısı	Ölçüt Puan
	Ölçeği Puanları	Öz-saygı
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

	Benlik Algısı	Ölçüt Puan
	Ölçeği Puanları	Depresyon
İrem	54	87
Metin	98	34
Erdal	128	23
Dilek	89	43
Çağdaş	101	32
-----	87	23
-----	98	49
-----	67	78
-----	122	45
Sedef	79	34

Daha önce geliştirilmiş ve geçerliliğinin yüksek olduğu bilinen araç

Korelasyon (r) -1'e yakın

Geçerlilik Katsayısı

Zamandaş geçerliliğinde geçerlilik katsayısının -1 ile 1 arasında değişmesi beklenir.

Ölçüte bağlı olarak geçerlilik katsayısı ∓ 1 'e yaklaştıkça artar, 0'a yaklaştıkça düşer.





YAPI GEÇERLİLİĞİ

Beşeri bilimlerde atılganlık, güvensizlik, içe donukluk vb. özelliklerin ölçülmesi amacıyla ölçme aracı geliştirmek için önce, söz konusu olan özellik ya da özellikleri belirlemek yani yapıyı tanımlamak, daha sonra tanımlanmış yapıdan sinanabilir denenceler çıkarmak ve çıkarılan denenceleri sinamak için deneysel ve istatistiksel çalışmalar yapmak gerekmektedir.



ZAMANDAŞ GEÇERLİLİK

Zamandaş geçerlilikte de en zor ve önemli nokta ölçütün doğru bir biçimde belirlenmesidir. Ölçütun belirlenmesinde şu noktalar dikkate alınmalıdır: Ölçüt(un)

- 1 ölçme aracının ölçmeye yöneldiği özellikle doğrudan ilişkili olmalıdır. Bu ilişki doğru orantılı ya da ters orantılı olabilir.
- 2 geçerliliği yüksek olmalıdır. Geçerliliği yüksek olmayan bir ölçüt puanla bakılacak korelasyonun düşük olması kaçınılmazdır.

7- BİR ÖLÇME ARACINDA BULUNMASI GEREKEN PSİKOMETRİK NİTELİKLER-4 (Sayfa:82-86)**EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ****YAPI GEÇERLİLİĞİ**

Beşeri bilimlerde atılganlık, güvensizlik, içe dönüklük vb. özelliklerin ölçülmesi amacıyla 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 nedir?

Yapı birbirleriyle ilgili olduğu düşünülen belli öğelerin ya da öğeler arasındaki ilişkilerin oluşturduğu bir örüntüdür. Bu anlamda, bir testin yapısını geçerleme süreci, temelde testin maddelerine verilen yanıtlar arasındaki ilişkilerin analizine dayanır.

BAŞARI**ZEKÂ****KİŞİLİK****TUTUM****İLGİ**

Yapı geçerliliği, bir testin dayandığı kuramsal temelleri ne derece iyi örneklediğiyle ilgilidir. İnsan davranışları ve özellikleri çoğunlukla soyut bir yapıya sahiptir. Zeka, sevgi, merak, sosyal uyum ve ruh sağlığı gibi soyut kavramları psikolojik testlerle ölçebilmek için, önce bu kavramların içeriğinin bilinmesi gerekir.

Birey ne tür davranışlar gösterdiği zaman hangi niteliklerin varlığına ya da yokluğuna karar verilebileceğine ilişkin ölçütler geliştirilmesi gerekmektedir. Yapı geçerliliği bir yandan testin ölçtüğü niteliklerin neler olduğunu araştırma, diğer yandan testi alan kişilerin elde ettikleri puanların ne anlama geldiğini açıklama çabalarıyla 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.



GÜVENİLİRLİK

Güvenilir bir ölçme aracı, aynı özelliklerle ilgili olarak arka arkaya yapılan ölçmelerde yaklaşık olarak aynı sayısal sonucu verir. Güvenilir bir test, aynı gruba iki ya da üç kez uygulandığında gruptaki her bir kişi bütün uygulamalarda yaklaşık olarak aynı puanı almalıdır. Bir testin ölçmek istediği özelliği ölçebilmesi için o testin söz konusu olan özelliği kararlı olarak ölçmesi gerekir.

Güvenirlilik çalışmalarının odak noktası şudur:

Eğer kişi iki defa teste tabi tutulursa iki testten aldığı puanlar birbirine benzer midir? ya da birbirine ne kadar yakındır?

Bu noktada gözlenen puan, gerçek puan, ölçmenin standart hatası ve güvenilirlik katsayısı kavramlarının bilinmesi gerekir.

Murat, üç dakikalık bir sözcük çalışmasında 162 sözcük ya da diğer bir ifadeyle 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 bir çok nedeni olabilir. Bir hareketten diğerine, "dikkat ve çaba" değişebilir. Özellikle uzun periyotlarda 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 değinmek gerekmektedir.

Psikometride hata kavramı istenmeyen değişkene işaret eder. Ölçme hataları giderilene kadar sürdürülmeli ve böylece gerçek puan elde edilmelidir. Ancak davranış örneği sınırlı olduğu için gözlenen puan gerçek puandan farklılık gösterir. Buradaki farklılık ölçme hatasıdır. Geleneksel olarak hataların varlığı gözlenen puanın, gerçek puandan yüksek ya da düşük olmasına neden olur.

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 bir çok 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

$$X = T + 0 \Rightarrow X = T$$

Ölçmenin standart hatası $S_h = s\sqrt{1 - r_x}$ formülü ile bulunur.

S_h = Ölçmenin standart hatası

s = Standart sapma

r_x = Güvenilirlik katsayısı

Ölçmenin standart hatası nasıl düşük çıkar?

Bir ölçme işleminde s=8 ve $r_x=.75$ ise

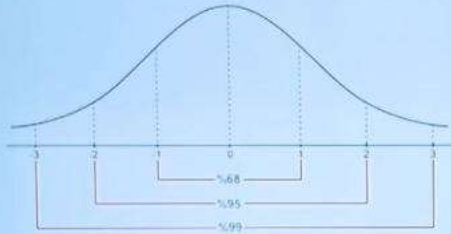
$$S_h = 8\sqrt{1 - .75}$$

$$S_h = 8\sqrt{.25}$$

$$S_h = 8 \times .5$$

$$S_h = 4$$

Bu sonuç bireylerin puanlarına ± 4 puan hata karıştığı anlamına mı gelir?



%68 olasılık için $X \pm 1 \times S_h$

%95 olasılık için $X \pm 2 \times S_h$

%99 olasılık için $X \pm 3 \times S_h$

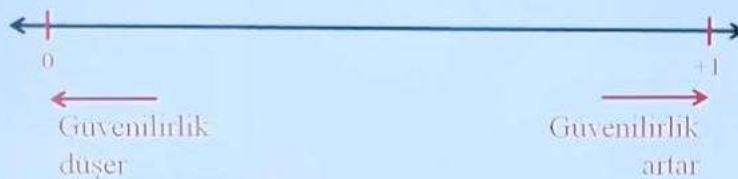
Bir kişi 100 üstünden 50 puan almış ve $S_h=4$ ise

%68 olasılık için $X \pm 1 \times S_h \Rightarrow 50 \pm 1 \times 4 = 46-54$

%95 olasılık için $X \pm 2 \times S_h \Rightarrow 50 \pm 2 \times 4 = 42-58$

%99 olasılık için $X \pm 3 \times S_h \Rightarrow 50 \pm 3 \times 4 = 38-62$

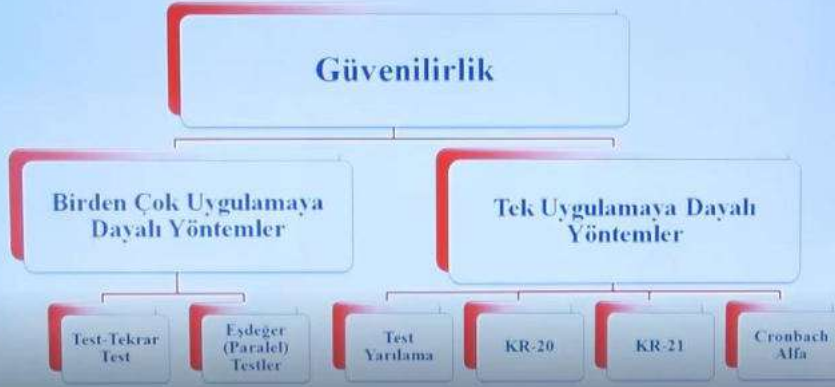
Güvenilirlik katsayısı 0 ile 1 arasında değişir.



Güvenilirlik katsayısı için alanyazında genellikle 0.70 ve üstü ölçütü kabul edilir. Ancak bu değerin de çok yüksek bir değer olmadığı, 0.70-0.80 arası güvenilirliğin ancak ön bilgi elde etmek amacıyla kullanılabileceği ifade edilir. Genel yetenek gibi bilişsel özellikleri ölçen testlerde güvenilirlik katsayısının 0.90 ve üzerinde olması arzu edilir.



GÜVENİLİRLİK KESTİRİM YÖNTEMLERİ



TEST-TEKRAR TEST YÖNTEMİ

Bu yöntem ile test güvenilirliğini test etmek için bir test, aynı gruba, belli bir zaman aralığıyla iki kez uygulanır. Daha sonra bireylerin birinci uygulamadan aldıkları puanlarla ikinci uygulamadan aldıkları puanlar arasındaki korelasyon hesaplanır. Elde edilen korelasyon katsayısına **kararlılık (devamlılık/istikrarlılık) katsayısı** adı verilir.

	Anksiyete Ölçeği Puanları		Anksiyete Ölçeği Puanları
	X ₁		X ₂
Ceren	54		50
Filiz	98		110
Aydın	128	Z	120
Ece	89	A	80
Deniz	101	M	101
.....	87	A	85
.....	98	N	106
.....	67		59
.....	122		119
Dursun	79		82

Kararlılık
Katsayısı

Korelasyon (r)
1'e yakın

Güvenilirliğin kararlılık, tutarlılık ve duyarlılık sorunu olduğuna daha önce değinilmişti.

Bu yöntemle elde edilen güvenilirlik kanıtı güvenilirliğin yalnızca kararlılık boyutuna karşılık gelir; güvenilirliğe ilişkin tümel bir sonuç ortaya koymaz. Bu nedenle bir ölçme aracının güvenilirliğine ilişkin kanıt toplamak isteyen bir kişi kararlılığın yanı sıra tutarlılığa ve duyarlılığa ilişkin sorgulamaları yapmak durumundadır.

Bu yöntem ölçülen özelliğin kararlı olduğu durumlarda uygulanması gereken bir yöntemdir. Testin ölçtüğü özellik sürekli değişkenlik gösteriyorsa bu yöntemle testin güvenilirliği hesaplanmamalıdır. Diğer bir deyişle bu yöntem daha çok iki uygulama arasında kolaylıkla değişmeyen özellikleri ölçen testler için uygundur. Örneğin genel zihin yetenekleri, kişilik testleri, ilgi envanterleri, tutum ölçekleri vb. gibi testlerin kararlılık bağlamında güvenilirlikleri bu yöntemle hesaplanabilir.

Bu yolla test güvenilirliğini kestirmede karşılaşılan önemli bir sorun, testin iki uygulaması arasındaki zaman aralığının ne kadar olması gerektiğidir. Bu soruya kesin bir yanıt vermek olanaksızdır. Aradan geçen zaman, hem ölçme aracının ölçtüğü özellik bakımından yanıtlayıcıların önemli ölçüde değişmelerine hem de birinci uygulamada verilmiş olan yanıtların hatırlanmasına izin vermeyecek bir uzunlukta olmalıdır.

Diğer bir ifadeyle iki uygulama arasındaki zamanın belirlenmesindeki temel ilke şudur:

“Birinci uygulamada maddeleri hatırlamayacak kadar uzun, ölçülen özellikte köklü değişimler oluşmayacak kadar kısa olmalıdır”.

Aradaki zamanın belirlenmesinde dikkat edilmesi gereken bir başka nokta özelliğin değişim hızıdır. Bazı özellikler daha hızlı değişime açıken, bazıları ise daha zor değişir.

Hangi özellikler daha hızlı ve yavaş değişir?

Ayrıca özelliklerin değişimi yaşla da ilişkilidir. Bazı özellikler belli yaşlarda daha durağanken, bazı yaşlarda daha değişkendir.

Hangi özellikler belli yaşlarda daha durağan ve değişkendir?

Test tekrar test yöntemi başarı testleri için çok tercih edilmemektedir. Çünkü kısa vadede başarı kolay değişebilen, kararsızlık gösterebilen bir özelliktir. Bu yöntem daha çok yetenek testleri, kişilik envanterleri vb. psikolojik ölçme araçlarından elde edilen puanların güvenilirlik kanıtlarını üretmek için tercih edilmektedir.

SEVGİ

8- BİR ÖLÇME ARACINDA BULUNMASI GEREKEN PSİKOMETRİK NİTELİKLER-5 (Sayfa:86-89)**EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ****TEK UYGULAMAYA DAYALI YÖNTEMLER**

Birden çok uygulamaya dayalı yöntemler alanyazında genellikle

1)**Test yarılama (iki yarı güvenilirliği / eşdeğer yarılar)****2)****KR-20 ve KR-21****3)****Cronbach alfa****4)****Hoyt'un varyans analizi**

olarak geçer.

Birden çok uygulamaya dayalı yöntemlerle güvenilirlik kanıtı elde etmek biraz zahmetlidir. Test tekrar test yönteminde testi farklı zamanda ikinci kez uygulamak için grup bulmak bazen zor olabilir. Ayrıca katılımcı kaybı yaşamak da olasıdır.

Tek uygulamaya dayalı yöntemler tek bir test, tek bir grup ve tek bir uygulama gerektirmektedir. Dolayısıyla güvenilirlik kanıtı elde etmek daha pratiktir (özellikle kararlılık katsayısının hesaplanmadığı durumlarda güvenilirliğin bir boyutu tabiri ki eksik kalır). Bir test bir gruba, bir kez uygulanıyorsa test kendi içinde değerlendirilecek demektir.

Güvenilirliğin bir boyutu da testin tutarlılığıdır. Tek uygulamaya dayalı güvenilirlik sorgulama yöntemleri ile "test kendi içinde tutarlı bir bütün oluşturur mu?" sorusuna yanıt aranır. Bu nedenle bu yöntemlerin tümünden elde edilen katsayı **iç tutarlılık katsayısı** olarak adlandırılır.



TEST YARILAMA YÖNTEMİ

Bu yöntemle güvenilirliği tahmin etmede uygulanmış bir test iki eşdeğer yarıya bölünür

ve bireylerin iki yarıdan aldıkları puanlar arasındaki tutarlılık incelenir. Buradaki en

temel sorunlardan biri testin iki eşdeğer yarıya nasıl bölünmesi gerektiği ile ilgilidir.

En sık başvuru yöntemler şunlardır:

1. İlk yarı ve son yarı
2. Tek ve çift
3. Rastlantısal

İlk ve son yarı yöntemi her test için uygun değildir. Uygun olmadığı durumlar:



Testteki maddeler basitten zora doğru sıralanmış ise



Testteki maddeler konu içeriklerine ya da faktörlere göre kümelenecek şekilde yerleştirilmiş ise



Testteki madde sayısı çok fazla ise

TEST YARILAMA YÖNTEMİ ÇALIŞMA ÖRNEĞİ

	<u>Türkçe Testi</u>	<u>X₁ YARI</u>	<u>X₂ YARI</u>
Deniz	48	20	28
Şebnem	98	50	48
Ufuk	66	32	34
Işıl	88	44	44
Sinem	32	20	12
.....	70	30	40
.....	62	28	34
.....	40	20	20
.....	60	28	32
Sinan	66	36	30

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

Korelasyon (r)

Spearman-Brown Eşitliği

İç tutarlılık katsayısı



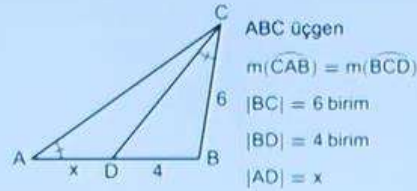
KUDER-RICHARDSON 20 ve 21

Kuder-Richarson 20 ve 21 ya da kısaca 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 **iç tutarlılık katsayısı** adı verilir.

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

Puanlamanın kategorik olması nedir?

KR-20 ve KR-21 doğru yanıtlara 1 puan, yanlış ve boş bırakılan maddelere 0 puan vererek puanlamanın yapıldığı ya da maddenin yanıtının iki seçenekli verildiği durumlarda güvenilirliği belirlemede kullanılır. Eğer testteki maddeler farklı ağırlıklarla puanlanmışsa ya da test puanları şans başarısı için düzeltilmişse bu formüller kullanılmaz.



Yukarıdaki verilere göre, x kaç birimdir?

- A) 3 B) 4 C) 5
D) $\frac{7}{2}$ E) $\frac{9}{2}$

16. Tanışmadığınız yeni birisine de dostça davranmak gerekir mi?
49. Belli etmeseniz bile çevrenizdekilere sık sık darıldığınız olur mu?
95. Aileniz, yeterli düzeyde başarılı olduğunuz kanısında mıdır?
156. Bazen sevdiğiniz kişiyi kırmaktan hoşlandığınız olur mu?

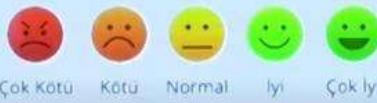
Evet	Hayır
☐	☐
☐	☐
☐	☐
☐	☐



CRONBACH ALFA

Eğer ölçme aracının puanlaması çok kategorili ise KR-20 ile aynı mantık üzerine kurulu Cronbach alfa hesaplanır. Elde edilen katsayının ismi yine **iç tutarlılık katsayısı**dır.

Puanlamanın çok kategorili olması ne demek?



1 Madde sayısı arttıkça güvenilirlik artar. Neden?

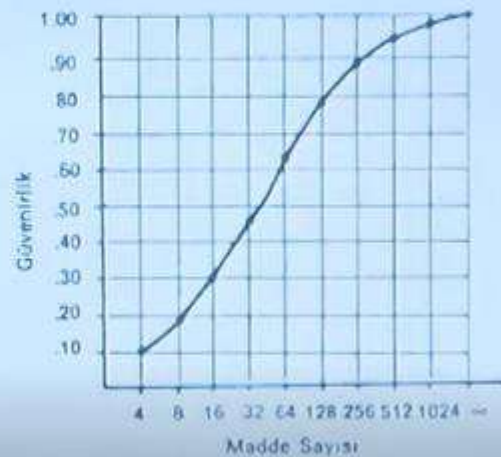
Birim		Duyarlılık		Hata		Güvenilirlik	Örnek
Küçük	⇒	Artar	⇒	Azalı	⇒	Artar	Kuyumcu terazisi 20 maddelik bir test
Büyük	⇒	Azalı	⇒	Artar	⇒	Azalı	Kantar 10 maddelik bir test

Ölçme yapılan en küçük aralığa denir.

1 Madde sayısı arttıkça güvenilirlik artar. Neden?

TABLO 7. Test Uzunluğunun Test Güvenirliğiyle İlişkisi

Madde Sayısı (K)	Güvenirlik (r_{xx})	Güvenirlikteki Artış
4	.10	—
8	.18	.08
16	.30	.12
32	.46	.16
64	.63	.17
128	.77	.14
256	.89	.12
512	.94	.05
1024	.97	.03
∞	1.00	.03



ŞEKİL 7. Test Uzunluğuyla Test Güvenirliğinin İlişkisi.

- 2) Puanlayıcı sayısı arttıkça güvenilirlik artar.
- 3) Maddelerin dilbilgisi kurallarına uygun, açık ve anlaşılır yazılması.
- 4) Puanlama nesnelliği
- 5) Bireylerin test almaya güdülenmesi
- 6) Fiziksel ortamın uygun hale getirilmesi (ses, sıcaklık vb.)
- 7) Süreli bir test uygulaması ise sürenin yeterli verilmesi
- 8) Maddelerin testte düzgün yerleştirilmesi, okumayı güçleştirecek bir unsurun bulunmaması, test katılımcısının yaşına gelişim düzeyine uygun bir punto büyüklüğü kullanılması, baskı hataları bulunmaması vb.
- 9) Puanlama ya da veri girişi yapılırken dikkatli ve titiz davranılması.
- 10) Uygulamada bireylerin dikkatini dağıtacak ve/veya kaygısını arttıracak davranışlardan kaçınılması

Güvenilirliği artıran faktörler geçerliliği de artırır. Ancak geçerlilik yalnızca tesadüfi hatalardan değil, aynı zamanda sabit ve sistematik hatalardan da etkilenir. O nedenle öğretmenin tesadüfi hataların yanı sıra sabit ve sistematik hatalardan arındırmaya yönelik alacağı tüm tedbirler geçerliliği de arttıracaktır.

9- TEST GELİŞTİRME VE MADDE/SORU TÜRLERİ-1 (Sayfa: 89-90)

EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ



İÇERİK

- ☐ Test ve test türleri
- ☐ Test geliştirme aşamaları
- ☐ Soru/madde kavramsal temel
- ☐ Geleneksel soru/madde türleri
 - ☐ Doğru-yanlış
 - ☐ Eşleştirme
 - ☐ Çoktan seçmeli
 - ☐ Uzun yanıtli sorular
 - ☐ Kısa yanıtli sorular
 - ☐ Cümle tamamlama
 - ☐ Sözlü yoklama

- ☐ Destekleyici yöntemlerin kavramsal temelleri
 - ☐ Portfolyo
 - ☐ Performans değerlendirme



TEST NEDİR?

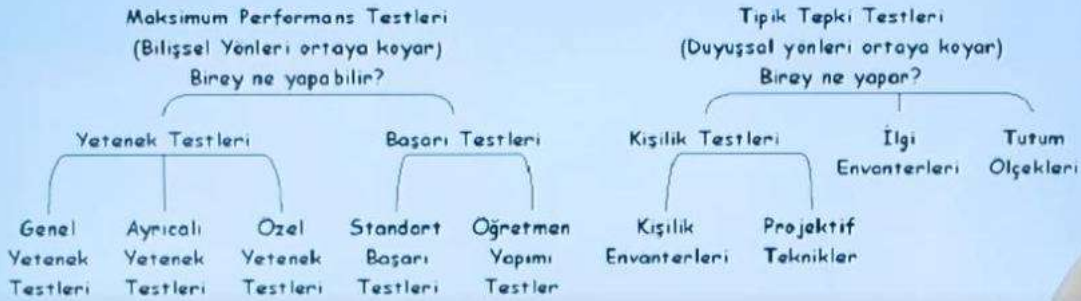
Test eğitimde bireylerin özelliklerini belirlemeye yönelik ölçme araçlarına verilen genel bir kavramdır. Bu terim zaman zaman günlük dilde yalnızca çoktan seçmeli maddelerin bulunduğu ölçme araçları için kullanılsa da, test bu belirlemeden çok daha geniş bir anlamda kullanılmaktadır.



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

TESTLER VE TESTLERİN SINIFLANDIRILMASI



TEST GELİŞTİRMENİN AŞAMALARI

Önceden belirlenmiş özelliklerin ölçülmesi ile ilgili bir test geliştirilmesi, sistematik işlem basamaklarına uyulmasını gerektirir. Test geliştirmede farklı yaklaşımlar olmakla birlikte, burada alanda yaygın kabul gören işlem basamakları üzerinde durulacaktır. Test geliştirme işleminin işlem basamakları:

1. Amacın belirlenmesi
2. Kapsamın belirlenmesi ve belirtke tablosunun oluşturulması
3. Denemelik maddelerin / soruların yazılması
4. Maddelerin / soruların gözden geçirilmesi (redaksiyon)
5. Denemelik test formunun hazırlanması
6. Testin uygulanması
7. Test ve madde istatistiklerinin hesaplanması
8. Seçilen maddelerden oluşan nihai formun oluşturulması

Pek çok kaynakta test geliřtirmenin adımları olarak bunlar olsa da, sınıf ii lme ve deęerlendirme etkinliklerinde bu basamaklara uygun test geliřtirmek pratikte ok mmkn deęildir. Bu durumda ğretmenler aısından sınıf ii lme ve deęerlendirmelerde uygulanması daha olanaklı bir test planı nerelim

1. Sınavın amacının belirlenmesi (elde edilecek puanların hangi amalarla uygulanacağı),
2. Sınavda yoklanacak davranışların belirlenmesi,
3. Sınavın kapsayacağı konuların listelenmesi,
4. Okulun takviminin ve sınav sonuçlarının kullanılacağı zamanın dikkate alındığı bir sınav gntünün belirlenmesi,
5. Bir ders saatine uygulanabilecek uzunlukta (en fazla iki) bir sınav sresinin belirlenmesi,
6. Belirtke tablosunun hazırlanması
7. Soru/madde trlerinin belirlenmesi. Soru/madde trleri sınavda kullanılacak davranışlara gre belirlenir. Bir sınavda tek bir soru/madde tr kullanılabileceęi gibi, birden fazla da kullanılabilir (soru/madde eřitlemesi),

8. Soru/madde sayısının belirlenmesi,
9. Sınav süresinin belirlenmesi,
10. Sınavda kullanılacak soruların/maddelerin «ortalama güçlüğü» ve «güçlük dağılımının» belirlenmesi,
11. Soruların/maddelerin yazımında, redaksiyonunda ve teste alınacak soruların/maddelerin seçiminde izlenecek yolun belirlenmesi,
12. Cevap anahtarının ve puanlama yönteminin belirlenmesi,
13. Ölçme aracını yazma ve çoğaltma yönteminin belirlenmesi,
14. Sınavın uygulanma kurallarının belirlenmesi (yönerge yazılması),
15. Sınavın uygulanması
16. Test ve madde istatistiklerinin hesaplanması



SORU MU? MADDE Mİ?



ÖLÇME ARAÇ VE YÖNTEMLERİ

GELENEKSEL YÖNTEMLER

- Yazılı Yoklama
- Sözlü Yoklama
- Çoktan Seçmeli
- Doğru-Yanlış
- Cümle Tamamlama
- Kısa Cevaplı
- Eşleştirme

DESTEKLEYİCİ YÖNTEMLER

- Portfolyo
- Öz Değerlendirme
- Akran Değerlendirme
- Gözlem Formu
- Kontrol Listesi
- Derecelendirme Ö.
- Diğerleri



GELENEKSEL YÖNTEMLER

Seçme Gerekli Maddeler

- Doğru-Yanlış
- Eşleştirme
- Çoktan Seçmeli

Açık Uçlu Sorular

- Uzun Yanıtlı Sorular
- Kısa Yanıtlı Sorular
- Cümle Tamamlama
- Sözlü Yoklama

10- TEST GELİŞTİRME VE MADDE/SORU TÜRLERİ-2 (Sayfa: 90-92)**EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ**



DOĞRU-YANLIŞ MADDELERİ

Cevaplayıcının verilen ifadelerin doğru mu, yanlış mı olduğunu belirlemesinin istendiği madde türüdür. Doğru yanlış maddelerinde verilen bir tür «önerme»dir.



Cumhuriyet 1920 yılında kurulmuştur. (D) (Y)

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.



DY MADDELERİ AVANTAJLAR

Hazırlaması kolaydır.	Kullanışlılık artar.
Uygulaması kolaydır.	
Puanlaması kolaydır.	
Çok sayıda soru sorulabilir.	1. Güvenilirlik artar. 2. Geçerlilik artar (?)
Sistemik hata karışması ihtimali yoktur.	Geçerlilik artar.



DY MADDELERİ DEZAVANTAJLAR

Şans başarısı olasılığı yüksektir.	Güvenilirlik düşer.
İyi yazılmazsa genellikle üst düzey bilişsel özellikleri ölçemez.	Analiz Değerlendirme Yaratıcılık
Öğrenme eksiklerini belirleyemez.	Formatif Değerlendirme için uygun değildir.
Yanlış üzerinden öğretim yapılmaz.	
Her derste/konuda kesin yanlış denebilecek durumlar belirlemek zordur.	



DY MADDELERİ YAZIM KURALLARI

Her madde tek ve belirli bir fikri belirtmelidir. Özellikle aynı maddede biri doğru, ötek 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)

Yanlış örnek

Balinalar tatlı suda yaşayan memeli hayvanlardır. (D) (Y)

Doğru örnek

Balinalar,
tatlı suda yaşarlar. (D) (Y)
memeli hayvanlardır. (D) (Y)

Doğru yanlış maddesi, kesinlikle doğru ya da kesinlikle yanlış olmalıdır. Maddenin doğruluğu ya da yanlışlığı, başka bir açıklamaya gerek kalmadan belirlenebilmelidir.

Yanlış örnek

Bugüne değin Türkiye Cumhuriyeti'nin üç beş cumhurbaşkanı olmuştur. (D) (Y)

Doğru örnek

2022 yılına kadar Türkiye Cumhuriyeti'nin 12 cumhurbaşkanı olmuştur. (D) (Y)

Yanlış örnek

İki kıta levhası çarpıştığında yerkabuğunu yukarı doğru itilerek dağ sıraları oluşabilir. (D) (Y)

Yanlış örnek

Bir testte soru sayısı arttıkça kapsam geçerliliği artar. (D) (Y)

Yanlış örnek

Kurtuluş savaşı, Yunanlılara karşı olmuştur. (D) (Y)

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

Bir yargı temelden yanlış olmalıdır.

Yanlış örnek

Fransız Devrimi 1798 yılında olmuştur. (D) (Y)

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

Yanlış örnek

Atatürk siyasal bağımsızlığın ekonomik bağımsızlığa bağlı olduğunu görmemiş değildir. (D) (Y)

Doğru örnek

Atatürk siyasal bağımsızlığın ekonomik bağımsızlığa bağlı olduğunu görmüştür. (D) (Y)

Kam ifadeleri kendiliklerinden doğru ya da yanlış olmadığından bir kaynağa dayandırılarak verilmelidir. Belli bir kaynağa ya da bir otoriteye dayandırılmadan verilen kam ifadelerinin yer aldığı maddelerin ayırt etme gücü sıfır ya da negatif olma eğilimindedir.

Yanlış örnek

Kültür ile uygarlık arasında fark yoktur. (D) (Y)

Doğru örnek

Atatürk'e göre kültür ile uygarlık arasında fark yoktur. (D) (Y)

Yanlış örnek

Bilgi deneyle başlar, ancak deneyle sona ermez. (D) (Y)

Doğru örnek

Immanuel Kant'a göre bilgi deneyle başlar, ancak deneyle sona ermez. (D) (Y)

Bir maddenin ifadesi kısa, açık ve yalın olmalıdır. Madde, onda sorulan ana fikrin doğruluk ya da yanlışlığının belirlenmesi için gerekli olmayan ayrıntılarla şişirilerek uzatılmamalıdır.

Yanlış örnek

Milattan önce 2000 yıllarında Nil Vadisine yerleşen halk birçok tanrıya taptı, onların en önemlileri iyilik tanrısı Osiris ile kötülük tanrısı Seth idi. (D) (Y)

Bir maddenin ifadesi kısa, açık ve yalın olmalıdır. Madde, onda sorulan ana fikrin doğruluk ya da yanlışlığının belirlenmesi için gerekli olmayan ayrıntılarla şişirilerek uzatılmamalıdır.

Yanlış örnek

Milattan önce 2000 yıllarında Nil Vadisine yerleşen halk birçok tanrıya taptı, onların en önemlileri iyilik tanrısı Osiris ile kötülük tanrısı Seth idi. (D) (Y)

Doğru örnek

Eski Mısırlılar politeist idi. (D) (Y)

Okuduğunu anlama süreçlerinin ilk basamağı, metin içerisinde açıkça verilmiş bilgiden doğrudan çıkarım yapmaktır. Bir metne bağlı olarak yanıtlanan doğru-yanlış maddelerinin anlam geliştirme, yorumlama, değerlendirme gibi daha üst düzey okuduğunu anlama süreçlerine yönelik olabilmesi için, metindeki cümleler aynen yazılmamalıdır.

Yanlış örnek

"Bir varmış, bir yokmuş. Bir yavru fil varmış. Bir gün canı çok sıkılmış, hortumunu sallayarak kocaman kocaman sesler çıkartmış. Sonra da kendi sesini öyle çok beğenmiş ki, bir türlü susmak bilmemiş. Gece olmuş, herkes yatmış ama o hala hortumunu sağa sola çevirip kocaman sesler çıkartıyor ve çıkardığı seslerin çok güzel olduğunu düşünüyor. Filin sesinden kimse uyuyamaz olmuş. Sonunda maymun filin kapısını çalmış. 'Bak gece oldu, yıldızlar çıktı. Artık hepimiz uyumak istiyoruz. Lütfen gürültü yapma!' diye fili uyarmış. Ancak fil yıldızlara şöyle bir bakıp, 'Hayır, gece olmadı! Onlar yıldız değil, güneş.' diye maymunu terslemiş."

Aytül Akal, Geceyi Unutan Fil

Açıklama: Okuduğunuz metne ilişkin aşağıda belli durumlar verilmiştir. Verilen durumlar metne göre doğruysa D'yi, yanlışsa Y'yi daire içine alınız.

Maymun filin gürültüsünden rahatsız olmuş. (D) (Y)
Fil kocaman sesleri hortumundan çıkartıyordu. (D) (Y)

Öğrencinin bir maddede ifadenin doğruluğuna ya da yanlışlığına karar vermesinde, öğretmenin özellikle vurgulamak istediği bir sözcük/terim vs. varsa altını çizebilir.

Doğru örnek

Yutak, sindirim sisteminin yanı sıra solunum sisteminde de yer alan bir organdır. (D) (Y)

Doğru-yanlış maddesinin özellikle yanlış olduğu durumlarda, öğrencinin şansla puan almasını engellemek amacıyla ifadeyi düzeltmesi ya da yanlışlığın ne olduğunu belirtmesi istenebilir.

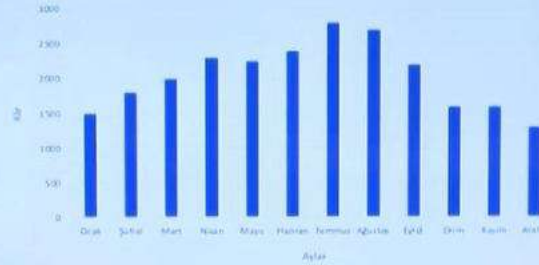
Doğru örnek

Kızılçam, soğuk iklimler için uygun bir çam türüdür. (D) (Y)

(Kızılçam Akdeniz havzası gibi ılıman iklimde yetişmeye uygun bir çam türüdür vb.)

Doğru örnek

Bir içecek firmasının, yeni çıkardığı ürünün satışından elde ettiği bir yıllık kârın aylara göre dağılımı gösteren grafik şöyledir:



Açıklama: Aşağıda verilen her bir ifadeyi dikkatle okuyun. Grafikte verilen verilere göre ifade doğruysa D'yi, yanlışsa Y'yi daire içine alınız.

1. Bu ürünün satışından en fazla yaz aylarında kâr elde edilmiştir. (D) (Y)
2. Aylık kârın en düşük olduğu ay Aralık'tır. (D) (Y)
3. Şubat-Mart aylarındaki toplam kâr, Eylül-Ekim aylarının toplam kârından düşüktür. (D) (Y)

İyi kurgulandığı
takdirde doğru-yanlış
maddeleri ile üst düzey
öğrenmeler de
yoklanabilir.

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, testte belirli bir örüntüye göre yerleştirilmemelidir.

Maddelerin sıralanışı, baştan itibaren iki doğru bir yanlış, bir doğru iki yanlış gibi bir örüntü gösterirse öğrencilerin bunu fark etmeleri olanaklı olabilir. Bu nedenle, doğru ve yanlış maddelerin testteki sıralanışı rastgele olmalıdır.

❑ Doğru ve yanlış maddelerin ifadesi yaklaşık olarak aynı uzunlukta olmalıdır.

❑ Bir doğru-yanlış testindeki doğru ve yanlış maddelerin sayısı, yaklaşık olarak birbirine eşit olmalıdır.

❑ Maddeleri işaretleme yöntemi, öğrenciye açık ve anlaşılır bir biçimde açıklanmalıdır.

❑ Mumkun olduğunca kısa yazılmalı, basit cümle yapısı tercih edilmeli, özellikle «ve» gibi bağlaçlara dikkat edilmelidir.

❑ İpucu vermekten kaçınmak gerekir. Özellikle «asla, daima, hiçbir, tümü» vb.

ifadelerin yanlış olma olasılığı yüksektir. Öte yandan «genellikle, çoğunlukla, bazen, zaman zaman, sıklıkla» vb. ifadelerin de doğru olma olasılığı yine yüksektir.

❑ Üst düzey becerileri yoklamak için tablo, grafik, harita ya da okuma parçası gibi bir öncülün kullanılması önerilir.

11- TEST GELİŞTİRME VE MADDE/SORU TÜRLERİ-3 (Sayfa:92-94)**EGİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ**



EŞLEŞTİRME MADDELERİ

Eşleştirme maddeleri, iki grup halinde verilen ve birbirleriyle ilgili olan bilgi öğelerinin, belli bir açıklamaya göre eşleştirilmesini gerektirir.

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.

A Sütunu	B Sütunu
C. 1. Osmanlı Meclis-i Mebusan'ın son toplanması	A. 1918
D. 2. Ankara'nın başkent olması	B. 1919
H. 3. Tek dereceli seçimlerin ilk kez yapılması	C. 1923
C. 4. TBMM'nin açılması	D. 1925
D. 5. Lozan Antlaşmasının imzalanması	E. 1928
D. 6. Cumhuriyet'in ilanı	F. 1930
C. 7. Sevr Antlaşmasının imzalanması	G. 1934
	H. 1936
	I. 1946

Yönerge / Açıklama

Öncüller / Maddeler

Muhtemel cevaplar / Seçenekler

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.

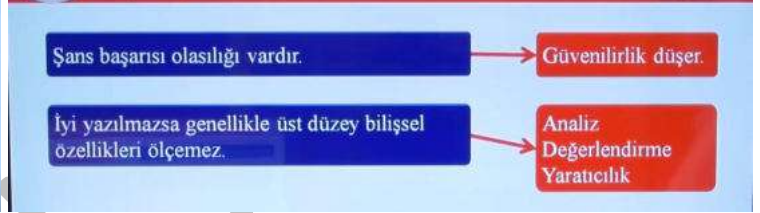
A Sütunu	B Sütunu
C. 1. Osmanlı Meclis-i Mebusan'ın son toplanması	A. 1918
D. 2. Ankara'nın başkent olması	B. 1919
H. 3. Tek dereceli seçimlerin ilk kez yapılması	C. 1923
C. 4. TBMM'nin açılması	D. 1925
D. 5. Lozan Antlaşmasının imzalanması	E. 1928
D. 6. Cumhuriyet'in ilanı	F. 1930
C. 7. Sevr Antlaşmasının imzalanması	G. 1934
	H. 1936
	I. 1946

Olası Durum	İlgili Yanıt Seti
İcat ve keşifler	Kişiler
Önemli olaylar	Tarihler
Tanımlar	Terim ve ifadeler
Örnekler, uygulamalar	Kurallar, ilkeler ve sınıflamalar
Kavramlar	Semboller ve işaretler
Eser başlıkları	Yazar ve oyuncu
Yabancı kelimeler	Türkçe karşılıkları
Kullanım biçimleri ve fonksiyonlar	Parçalar ve makineler
Nesnelerin isimleri	Nesnelerin şekilleri

EŞLEŞTİRME MADDELERİ AVANTAJLAR



EŞLEŞTİRME MADDELERİ DEZAVANTAJLAR



EŞLEŞTİRME MADDELERİ YAZIM KURALLARI

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 kullanılabilir, hiç kullanılmayabilir.

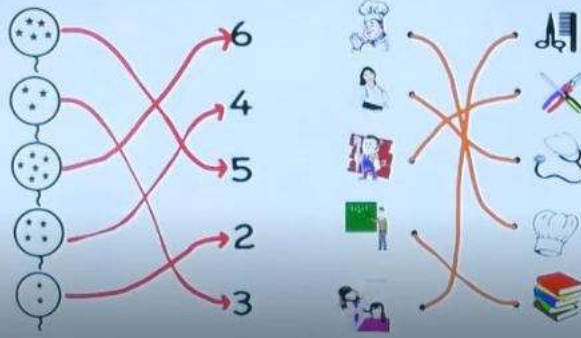
Seçeneklerin kullanılıp kullanılmayacağı, birden fazla kez kullanma durumu belirtilmelidir. Bir eşleşme takımındaki *öncüller* ile *cevaplar* eşit sayıda olmamalıdır. Eğer eşit sayıda öğrenci, hemen ilk bakışta bildiklerini eşleştirecek, geri kalan öncüllerin cevaplarını da tahmin edecektir.

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 kullanılabilir, 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. $44 \div 4 - 9$	5
3. $35 \div (5 + 2)$	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.

Yanlış örnek

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.

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

Uzun ifadeler öncül olarak

kullanılmalı ve öncüller sütunu

sayfanın sol sütununa, cevaplar

sütunu ise sayfanın sağ tarafına

yerleştirilmelidir. Böyle olması,

cevaplamada zaman kaybını önler

ve cevabın seçiminin kolaylaştırır

Doğru örnek

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.

A Sütunu	B Sütunu
C 1. Osmanlı Meclis-i Mebusan'ın son toplanması	A. 1918
D 2. Ankara'nın başkent olması	B. 1919
H 3. Tek dereceli seçimlerin ilk kez yapılması	C. 1923
C 4. TBMM'nin açılması	D. 1925
D 5. Lozan Antlaşması'nın imzalanması	E. 1928
D 6. Cumhuriyet'in ilanı	F. 1930
C 7. Sevr Antlaşması'nın imzalanması	G. 1934
	H. 1936
	I. 1946

- 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 arttırır. 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 halde cevabın seçimini güçleştirir.
- Cevapların seçileceği sütun, bir kelime listesi ise alfabetik sıraya göre (küçükten büyüğe ya da büyükten küçüğe) düzenlenmelidir.

Doğru örnek

Yönerge: Kanal K, bu yıl yılbaşı için beş farklı merkezden yayın yapmaktadır. Bu merkezler Los Angeles, Rio de Janeiro, Ankara, Mumbai (Bombay) ve Şangay'dır. Aşağıdaki "A" sütununda şehirler, "B" sütununda ise yeni yıla girdiği tarih ve saatler verilmiştir. Ankara'daki ekip yeni yıla girdiği anda diğer merkezlerdeki tarih ve saati kararlaştırdıktan sonra, o şehrin önündeki harfi, seçeneğin solundaki boşluğa yazınız. "B" sütunundaki bir seçenek, bir veya birden fazla kez kullanılabilir gibi, hiç kullanılamayabilir.



A Sütunu

- B. 1. Los Angeles
- E. 2. Rio de Janeiro
- C. 3. Mumbai (Bombay)
- G. 4. Şangay

B Sütunu

- A. 1 Ocak - 14.00
- B. 31 Aralık - 14.00
- C. 1 Ocak - 03.20
- D. 1 Ocak - 04.50
- E. 31 Aralık - 19.00
- F. 31 Aralık - 20.30
- G. 1 Ocak - 06.00

Bu eşleştirme maddeleri öğrencilerin yalnızca hatırlama düzeyi dışında, kavrama düzeyindeki öğrenmelerin de yoklanabileceğine ilişkin örnektir. Ör. Grafik / tablo / metin / harita



ÇOKTAN SEÇMELİ MADDELER

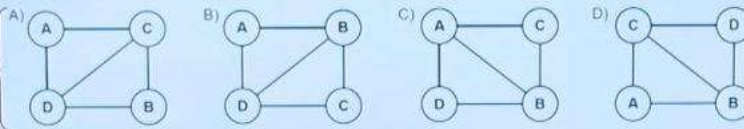


Soru 24 BÖLGELER ARASI BAĞLANTILAR

	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?



Seçenekler

Çoktan seçmeli madde

Doğru yanıt

Madde kökü



ÇOKTAN SEÇMELİ MADDELER

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 karışma ihtimali vardır.

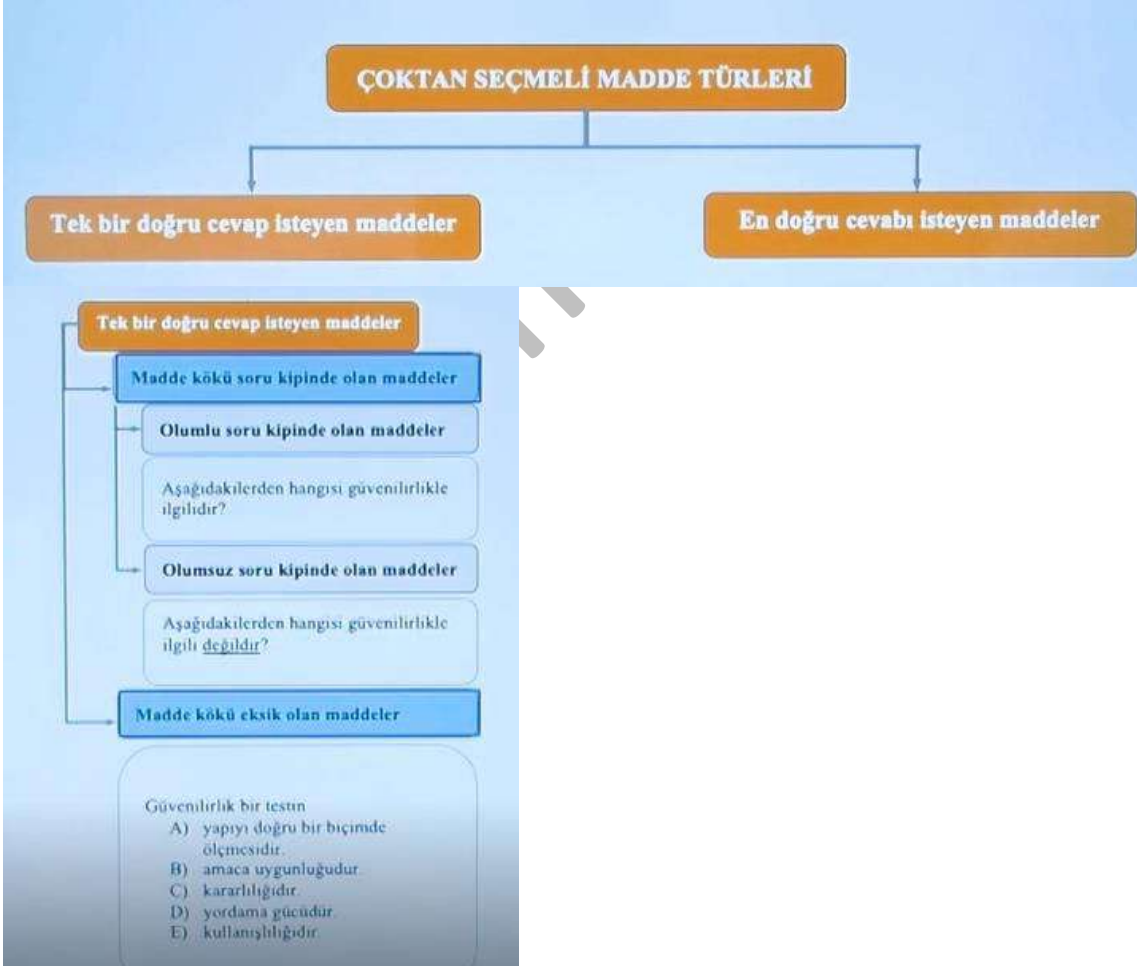
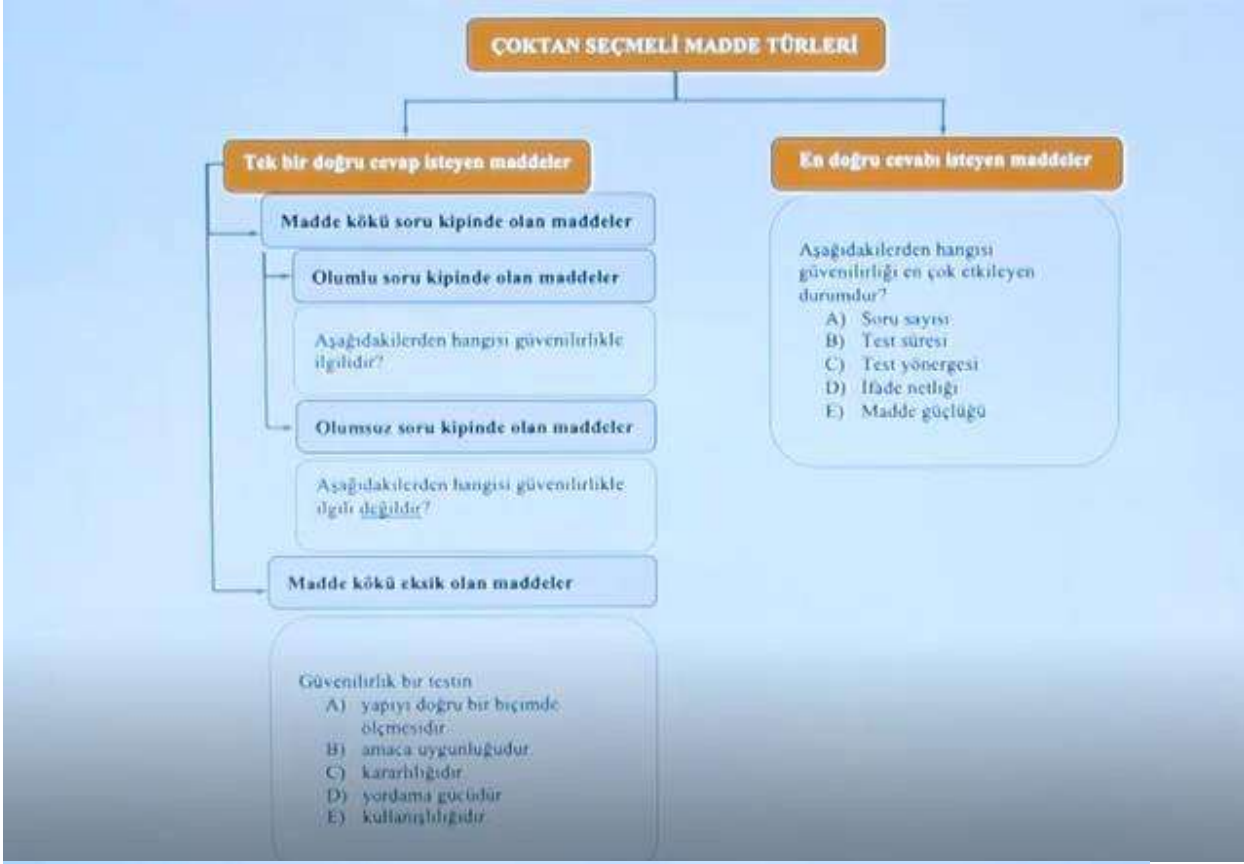
Seçenek Sayısı					Şans Başarısı
A	B				%50
A	B	C			%33.3
A	B	C	D		%25
A	B	C	D	E	%20
20	20	20	20	20	
		↓			
		20			

$$20 \text{ doğru, } 80 \text{ yanlış} = 20/80 = 1/4$$

Çok sayıda soru sorulabilir.

Sistemik hata karışma ihtimali çok düşüktür.

Çeldirme mantığı temellidir.



Madde kökü soru kipinde olan maddeler

Olumlu soru kipinde olan maddeler

Aşağıdakilerden hangisi güvenilirlikle ilgilidir?

Olumsuz soru kipinde olan maddeler

Aşağıdakilerden hangisi güvenilirlikle ilgili değildir?

Madde kökü eksik olan maddeler

Güvenilirlik bir testin

- A) yapıyı doğru bir biçimde ölçmesidir.
- B) amaca uygunluğudur.
- C) kararlılığıdır.
- D) yordama gücüdür.
- E) kullanılışlılığıdır.

En doğru cevabı isteyen maddeler

Aşağıdakilerden hangisi güvenilirliği en çok etkileyen durumdur?

- A) Soru sayısı
- B) Test süresi
- C) Test yönergesi
- D) İfade netliği
- E) Madde güçlüğü

12- TEST GELİŞTİRME VE MADDE/SORU TÜRLERİ-4 (Sayfa: 94-95)

EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ

SEVGİNİN SESİYLE



ÇSM YAZIM KURALLARI

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

Yanlış örnek

Analizler sonucunda bir testin geçerliliğinin düşük olduğu belirlenmiştir.

Bu durumun nedeni aşağıdakilerden hangisi olabilir?

- A) Soru sayısı azdır.
- B) Testten elde edilen puanların güvenilirliği düşüktür.
- C) Öğretime konu olan kapsam, sınava yeterince yansıtılmamıştır.
- D) Test hazırlama aşamalarına uyulmamıştır.
- E) Testin kullanılabilirliği düşüktür.

Doğru örnek

Öğretime 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.

Pilot uygulama sonucunda aşağıdaki durumlardan hangisinin olması beklenir?

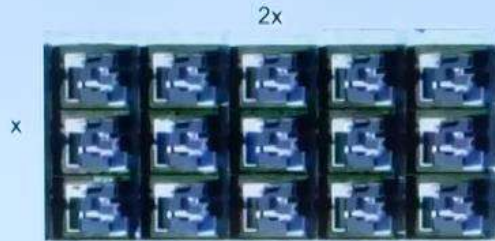
Doğru örnek

Bir testte madde sayısının az olması nedeniyle güvenilirlik katsayısı düşüktür. Bu nedenle teste yeni maddeler eklenmiştir. Revize edilmiş testin güvenilirlik katsayısını belirlemek için, bir grup üzerinde pilot uygulama yapılmıştır.

Pilot uygulama sonucunda aşağıdaki durumlardan hangisinin olması beklenir?

Doğru örnek

Bir durumun resim ya da fotoğrafla anlatılabildiği durumda, çokça cümle kurmak yerine görsellerden yararlanılabilir.



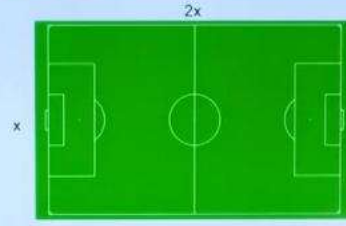
Derya arkadaşlarıyla yapacağı basketbol maçı öncesi yukarıda şekli verilen evlerin bulunduğu sitenin çevresinde ısınma koşusu yapmıştır.

Derya toplam 3 turda 810 metre koştuğuna göre, sitenin kısa kenarı kaç metredir?

- A) 30
- B) 45
- C) 60
- D) 90

Ancak madde böyle
hazırlanırsa cinsiyete göre
bir yanlışlık oluşma
olasılığı söz konusudur.

Yanlış Örnek



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

Yanlış Örnek



Bu reportajda yazar, gazetecinin hangi sorusunu yanıtlamıştır?

- A) Sizi diğer yazarlardan ayıran özellik nedir?
B) Yazarlar neden esrarengiz olmaya çalışır?
C) Okuyucularımız hakkında ne düşünüyorsunuz?
D) Sizce anlaşılır bir yazar olmanın zorlukları nelerdir?

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

YÜKSEK BİNALAR

"Yüksek Binalar" 2008 yılında bir Norveç dergisinde yayınlanmış bir makaledir.

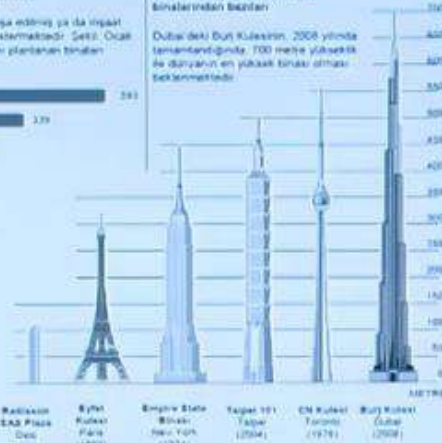
Şekil 1: Dünyanın yüksek binaları

(Şekil 1'te en az 30 katlı olarak inşa edilmiş ya da inşaat halinde olan binaların kayıtları gösterilmektedir. Şekil: Ocak 2001 tarihinden itibaren yapılmı planlanan binalar gösterilmektedir.)

Hong Kong	379
Dubai	379
Taipei	345
Singap	334
Sendai, Çin	324
Shanghai	318
Kuala Lumpur	313
Hacı Yek	312
Sao	312
Kuala Lumpur	308
Osaka	308
Toronto	301
Singapore	300
Colombo	277
Sat. Kule	268

Şekil 2: Dünyanın en yüksek binalarından bazıları

Dünya'daki bu kulelerin, 2008 yılında tamamlanmış ve 700 metre yükseklik ile dünyanın en yüksek binası olmaları beklenmektedir.

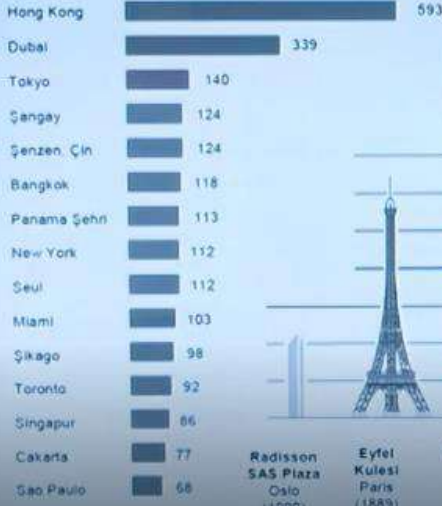


Şekil 1'de ne tür bir bilgi verilmektedir?

- A) Farklı binaların yüksekliklerinin bir karşılaştırması
B) Farklı şehirlerdeki binaların toplam sayısı
C) Çeşitli şehirlerdeki belirli bir yüksekliği aşan binaların sayısı
D) Farklı şehirlerdeki binaların mimari tarzları hakkında bilgi

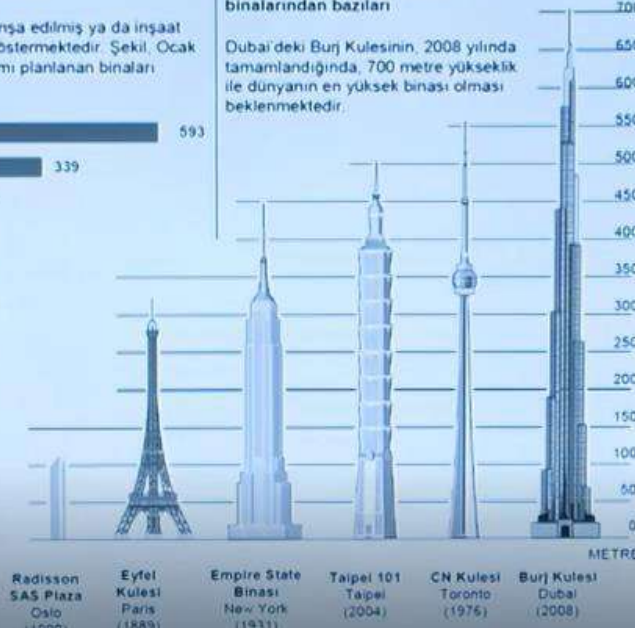
Şekil 1: Dünyanın yüksek binaları

Şekil 1 en az 30 katlı olarak inşa edilmiş ya da inşaat halindeki binaların sayısını göstermektedir. Şekil, Ocak 2001 tarihinden itibaren yapımı planlanan binaları içermektedir.



Şekil 2: Dünyanın en yüksek binalarından bazıları

Dubai'deki Burj Kulesinin, 2008 yılında tamamlandığında, 700 metre yükseklik ile dünyanın en yüksek binası olması beklenmektedir.



Şekil 1'de ne tür bir bilgi verilmektedir?

- A Farklı binaların yüksekliklerinin bir karşılaştırması
- B Farklı şehirlerdeki binaların toplam sayısı
- C Çeşitli şehirlerdeki belirli bir yüksekliği aşan binaların sayısı
- D Farklı şehirlerdeki binaların mimarî tarzları hakkında bilgi



ÇSM YAZIM KURALLARI

Maddeyi okuyan kişinin aklına hemen "hangi bakımdan en önemli" sorusu gelecektir.

Bu soruyu okuyan herkes kendince yorumda bulunabilir. Biri güvenilirlik açısından bakarken, diğeri geçerlilik açısından soruyu yorumlayabilir.

Yanlış örnek

Bir testte hangisi en önemlidir?

- A) Madde güçlüğü
- B) İfade netliği
- C) Soru sayısı
- D) Madde ayırtediciliği
- E) Test düzeni

Doğru örnek

Güvenilirlik söz konusu olduğunda aşağıdakilerin hangisi en önemlidir?

- A) Madde güçlüğü
- B) İfade netliği
- C) Soru sayısı
- D) Madde ayırtediciliği
- E) Test düzeni

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.

Yanlış örnek

Testin bütününden bağımsız alt boyutların her birisinin, bütüne 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?

- A) Konu kapsamı dengeli bir biçimde temsil edilmektedir.
- B) Testin görünüş geçerliliği düşüktür.
- C) Hedeflerde bilişsel alanlara göre dengeli bir dağılım söz konudur.
- D) Güvenilirlik katsayısı yüksektir.
- E) Yordama gücü düşüktür.

Doğru örnek

İç tutarlılık katsayısı yüksek bir test için aşağıdakilerden hangisi söylenebilir?

- A) Konu kapsamı dengeli bir biçimde temsil edilmektedir.
- B) Testin görünüş geçerliliği düşüktür.
- C) Hedeflerde bilişsel alanlara göre dengeli bir dağılım söz konudur.
- D) Güvenilirlik katsayısı yüksektir.
- E) Yordama gücü düşüktür.

Bir çoktan seçmeli maddeyi zorlaştırmak isteyen öğretmen temelde iki yöntem kullanabilir:

1. Çeldiricileri doğru yanıtla yaklaştırmak (güçlendirmek)
2. Davranışı üst düzey becerileri ölçen bir formda hazırlamak

!/? Hazırbulunuşluluğu ölçmeyi ya da öğrenme eksiklerini tespit etmeyi ya da not vermeyi amaçlayan bir öğretmen testinin/sınavının ortalama güçlüğüne ne düzeyde tutmalıdır?

- !/? Bir sınav zor olursa ne olur?
- !/? Bir sınav kolay olursa ne olur?
- !/? Bir sınav ortalama güçlükte olursa ne olur?

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

Yanlış 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?

Doğru örnek

Basın özgürlüğü indeksinin son sıralarında yer alan ülkelerde, basın özgürlüğünün tesis edilmesi için, hangi yasal düzenlemenin öncelikle yapılması gerekir?

Yanlış örnek

Sivas Kongresi, yurdun her yerinden gelen temsilcilerin katılımı ile gerçekleşti. Kongreye katılan temsilciler yaptıkları konuşmalarda mücadelenin millet adına yapıldığını, amaçlarının vatanın kurtuluşu ve milletin bağımsızlığı olduğunu açıkladılar. Kongrede Anadolu ve Rumeli'de faaliyet gösteren bütün millî cemiyetlerin "Anadolu ve Rumeli Müdafaa-i Hukuk Cemiyeti" adı altında birleştirilerek vatanın kurtuluşu için birlikte hareket etmelerinin sağlanması kararlaştırıldı.

Buna göre, Sivas Kongresi'nin Kurtuluş Savaşı için aşağıdakilerden hangisini sağladığı söylenemez?

- A) Amaç birliği
- B) Kültür
- C) Millî birliği
- D) Güç birliği

Seçeneklerde gereksiz sözcük tekrarından

kaçınılmalıdır. Tekrar edilen sözcük madde

köküne alınabilir.

Doğru örnek

Buna göre, Sivas Kongresi'nin Kurtuluş Savaşında hangi konuda birlik sağlamaya etkisi olmamıştır?

- A) Amaç
- B) Kültür
- C) Milliyetçilik
- D) Güç

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

Yanlış örnek

Karamazov Kardeşler adlı romanın yazarı kimdir?

Yanlış örnek

Suç ve Ceza, Karamazov Kardeşler, Kumarbaz gibi kült eserlerin yazarı Dostoyevski hangi ulustandır?

Yanlış örnek

Başarıyı ölçmek üzere geliştirilmiş bir testte, geçerlilik aşağıdakilerden hangisiyle öncelikle ilgilidir?

- 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

Doğru örnek

Başarıyı ölçmek üzere geliştirilmiş bir testte, geçerlilik aşağıdakilerden hangisiyle öncelikle ilgilidir?

- 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

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

Aşağıdakilerden hangisi güvenilirliği arttıran bir işlem değildir?

- A) Soru sayısını arttırmak
- B) Anlaşılır madde yazmak
- C) Teste yönerge eklemek
- D) **Soruların zor olması**
- E) Testi heterojen gruplara uygulamak

Doğru örnek

Aşağıdakilerden hangisi güvenilirliği arttıran bir işlem değildir?

- A) Soru sayısını arttırmak
- B) Anlaşılır madde yazmak
- C) Teste yönerge eklemek
- D) Soruları zorlaştırmak
- E) Testi heterojen gruplara uygulamak

İpucu vermekten 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

Testteki soru sayısının arttırılması halinde geçerlilik nasıl etkilenir?

- A) Her zaman artar.
- B) Daima azalır.
- C) **Bazen artar.**
- D) Etkilenmez.

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

Yanlış örnek

Testteki soru sayısının arttırılması halinde testin psikometrik niteliklerinde nasıl bir değişim olması beklenir?

- A) **Güvenilirlik artar.**
- B) Güvenilirlik azalır.
- C) Kullanışlılık azalır.
- D) Geçerlilik artar.
- E) Test zorluğu artar.

Seenekler yazılırken anlamca birbirini ieren/kapsayan ifadelerin kullanılmamasına dikkat edilmelidir.

Yanlış rnek

Parmak emme davranışı yaşamın belirli bir döneminde olağan bir davranış olarak nitelendirilirken, bu davranışın başka gelişim dönemlerinde görülmesi bireyin bir travması olduğunu düşündürebilir.

Parmak emme davranışı öncelikle hangi gelişim döneminde görülürse travmatik bir davranış olarak değerlendirilmelidir?

- A) Bebeklik
- B) Çocukluk
- C) Ergenlik
- D) Genç yetişkinlik
- E) Yetişkinlik

Doğru rnek

Parmak emme davranışı yaşamın belirli bir döneminde olağan bir davranış olarak nitelendirilirken, bu davranışın başka gelişim dönemlerinde görülmesi bireyin bir travması olduğunu düşündürebilir.

Parmak emme davranışı öncelikle hangi yaş döneminde görülürse travmatik bir davranış olarak değerlendirilmelidir?

- A) 0-2
- B) 3-11
- C) 12-18
- D) 19-25
- E) 26 yaş ve üstü

Seeneklerin uzunlukları birbirine eşit olmalıdır. Özellikle doğru cevabın daha uzun ya da daha kısa olmamasına dikkat gösterilmelidir.

Yanlış rnek

Aşağıdaki cümlelerin hangisinde abartılı anlatım vardır?

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

Doğru cevaplar teste dağıtılırken dikkat edilmeli bir örüntü göstermemesi sağlanmalıdır.

MATEMATİK				
1	A	B	C	D
2	A	B	C	D
3	A	B	C	D
4	A	B	C	D
5	A	B	C	D
6	A	B	C	D
7	A	B	C	D
8	A	B	C	D
9	A	B	C	D
10	A	B	C	D
11	A	B	C	D
12	A	B	C	D
13	A	B	C	D
14	A	B	C	D
15	A	B	C	D
16	A	B	C	D
17	A	B	C	D
18	A	B	C	D
19	A	B	C	D
20	A	B	C	D

Hepsi ve hiçbirini bir seçenek olarak kullanılmamalıdır.

Yanlış örnek

Pamuk aşağıdaki illerden hangisinde yetişir?

- A) Antalya
- B) Adana
- C) Aydın
- D) Hepsi

Doğru örnek

- I. Antalya
- II. Adana
- III. Aydın

Pamuk aşağıdaki illerden hangisinde/hangilerinde yetişir?

- A) Yalnızca II
- B) I ve II
- C) II ve III
- D) I, II ve III

Doğru örnek olarak verilen maddede de bir yanlış vardır. Adana kesinlikle doğrudur.

Çoktan seçmeli maddelerde seçenek sayısı genellikle 4 ya da 5 olsa da, seçenek sayısını belirleyen temel faktör öğrencinin içinde olduğu gelişim dönemidir.

- Lise ve daha üstü öğretim düzeyleri: 5
- Ortaokul düzeyi: 4
- İlkokul düzeyi: 3*-4

ÇSM'lerin seçenekleri

bir sıra ile verilmelidir.

Seçenekler baş harflerine

göre alfabetik olarak eğer

sayı iseler büyüklük

olarak sıralanmalıdır.

Doğru örnek



Soru 16 HAREKET HALİNDEKİ LEYLEKLER

Robert, 8 Eylül'de İskenderun'a saat 18.00'de ulaştığına göre uçuşuna İstanbul'dan kaçta başlamış olabilir?

- A) 06.00
- B) 14.00
- C) 18.00
- D) 23.00

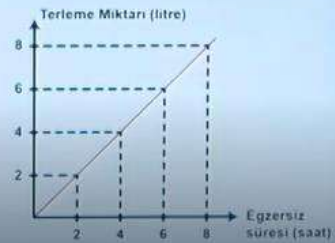
Doğru örnek



Soru 42 BOŞALTIM SİSTEMİ

Yandaki grafikte bir sporcunun yaptığı egzersiz ve bu egzersizden dolayı kaybettiği su miktarı verilmiştir. Bu grafiğe göre günde 2,5 saat egzersiz yapan bir sporcu günde yaklaşık en az kaç litre su içmelidir?

- A) 2-3 litre
- B) 4-5 litre
- C) 6-7 litre
- D) 8-9 litre



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

1. Madde yazılan alanı çok iyi bilmeli,
2. Madde yazma teknik ve yöntemlerinden haberdar olmalı,
3. Maddelerin yazılacağı dili kullanmada becerili olmalı ve
4. Testin uygulanacağı öğrencilerin gelişim düzeyini çok iyi bilmelidir.

- Çoktan seçmeli maddelerde seçeneklerin baş harfleri büyük olmalıdır.
- Bir madde kökü ile seçenekleri aynı sayfada olmalıdır. Yarıyı başka sayfada, yarıyı diğer sayfada olmamalıdır. Bu durum ortak köke dayalı maddeler için de geçerlidir.

Doğru örnek

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

3-5 maddeleri aşağıdaki tabloya göre yanıtlayınız.

Ölçme İşlemi	n	$\bar{X}$	X_{ort}	Mod	Skew ¹	Kur ²	K-S ³	S-W ⁴
1	40	112	110	114	256	-349	200	007
2	150	88	95	117	-1.76	-451	009	000
3	457	34	30	38	675	067	000	000
4	300	67	67	67	0	0	200	060
5	95	68	71	72	-651	-1.72	052	039
6	386	42	29	12	2.01	-988	000	000

¹Skewness ²Kurtosis ³Kolmogorov-Smirnov ⁴Shapiro-Wilk

3. 5 nolu ölçme işlemindeki dağılımla ilgili aşağıdaki yorumlardan hangisi doğrudur?

- A) Normale yakın
- B) Sola çarpık
- C) Sivri
- D) Basık
- E) Sağa çarpık

4. 1 nolu ölçme işlemindeki dağılımla ilgili aşağıdaki yorumlardan hangisi doğrudur?

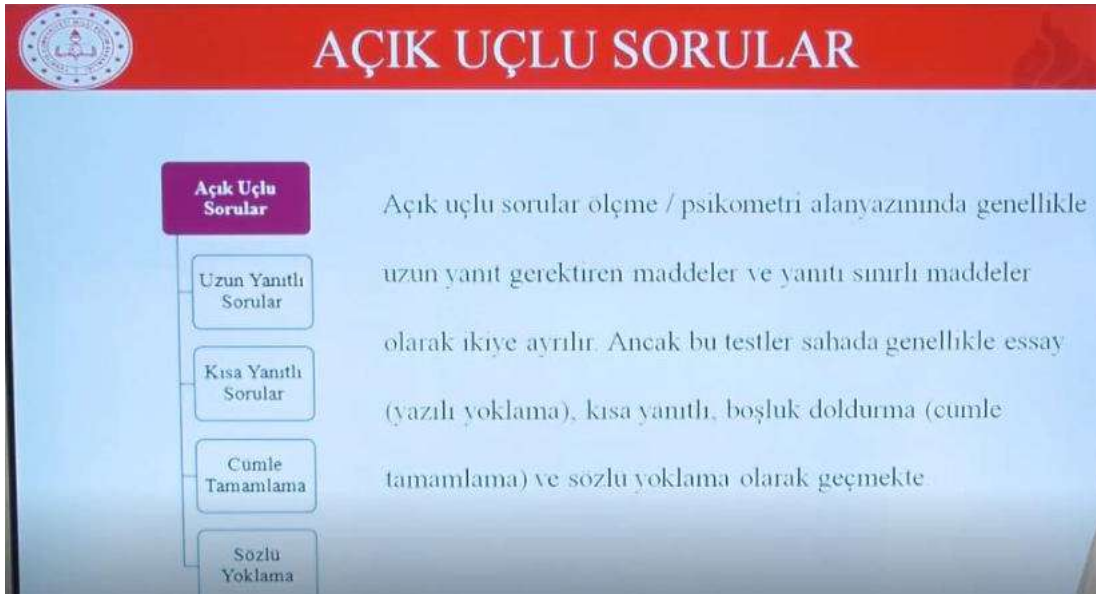
- A) Sola çarpık
- B) Simetriğe yakın
- C) Sağa çarpık
- D) Basık
- E) Normale yakın

5. 2 nolu ölçme işlemindeki dağılımla ilgili aşağıdaki yorumlardan hangisi doğrudur?

- A) Sola çarpık
- B) Normale yakın
- C) Sivri
- D) Sağa çarpık
- E) Simetriğe yakın

13- TEST GELİŞTİRME VE MADDE/SORU TÜRLERİ-5 (Sayfa:95-97)

EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ

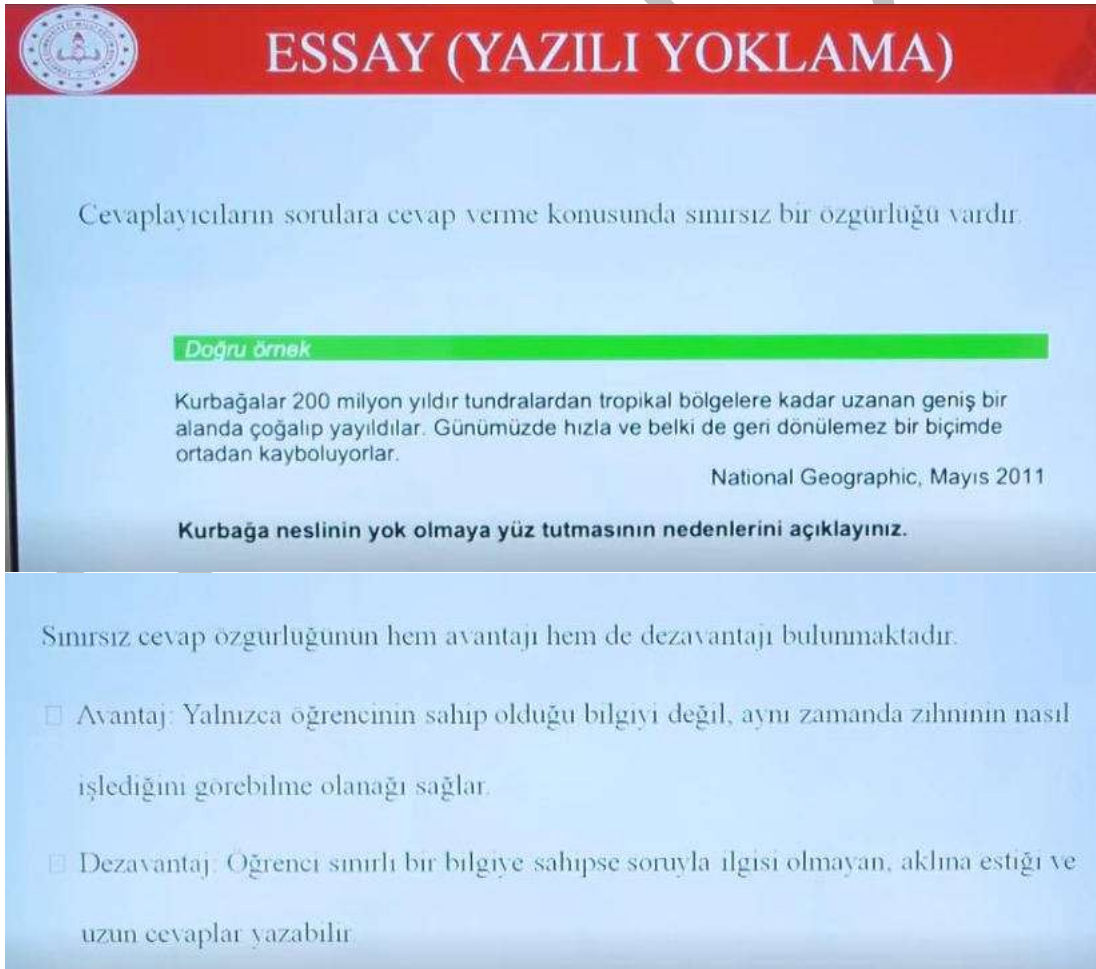


AÇIK UÇLU SORULAR

Açık Uçlu Sorular

- Uzun Yanıtlı Sorular
- Kısa Yanıtlı Sorular
- Cümle Tamamlama
- Sözlü Yoklama

Açık uçlu sorular ölçme / psikometri alanyazınında genellikle uzun yanıt gerektiren maddeler ve yanıtı sınırlı maddeler olarak ikiye ayrılır. Ancak bu testler sahada genellikle essay (yazılı yoklama), kısa yanıtlı, boşluk doldurma (cümle tamamlama) ve sözlü yoklama olarak geçmekte.



ESSAY (YAZILI YOKLAMA)

Cevaplayıcıların sorulara cevap verme konusunda sınırsız bir özgürlüğü vardır.

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.

Sınırsız cevap özgürlüğünün hem avantajı hem de dezavantajı bulunmaktadır.

- ☐ Avantaj: Yalnızca öğrencinin sahip olduğu bilgiyi değil, aynı zamanda zihninin nasıl işlediğini görebilme olanağı sağlar.
- ☐ Dezavantaj: Öğrenci sınırlı bir bilgiye sahipse soruyla ilgisi olmayan, aklına estiği ve uzun cevaplar yazabilir.

Üstelik sınırsız cevap özgürlüğü anlatım becerisi iyi olan öğrencilere bir avantaj sağlayabilir.

Örnek

"Bu roman şimdiye dek okuduğum romanların ne en iyisi ne de en kötüsüdür. Yazar, başlıca karakterleri, gerçekten de çok belirgin olarak çizmiş. Ancak aynı başarıyı, ikinci derecedeki karakterlerin betimlenmesinde gösterdiği söylenemez..."

Bu sınav türü dünyada en eski ve günümüzde de hala öğretmenlerin en çok tercih ettiği testlerden biridir. Bunun iki nedeni vardır:

- ☐ Pratik nedenler: Hazırlaması kolaydır.
- ☐ Eğitsel nedenler: Üst düzey bilişsel becerileri ölçmek için geleneksel yöntemler arasında en avantajlı yazılı test etme yoludur (sentezde tek yol).

Özellikle şu becerileri ölçmek için çok kullanışlıdır:

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

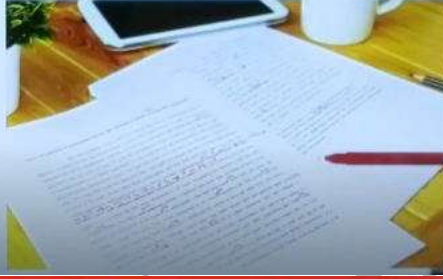
Güçlü ve zayıf yönleri belirleme

Essay'in Dezavantajları:

- ❑ Çok fazla soru sorulamaz (okuma ve yazma eylemlerinin süreler arasında ortalama 10 kat fark bulunmakta)
- ❑ Puanlar sadece öğrencinin sahip olduğu bilgiyi değil, öğrencinin anlatım biçimini, yazı güzelliği, bilgisini örgütleme biçimini de yansıtır. Bu ise ölçme sonuçlarındaki sistematik hata miktarını arttırabilir.

Essay'in Dezavantajları:

- Puanlama subjektifliği belki de en önemli dezavantajdır.



ESSAY'DE NESNELLİĞİ ARTIRMA YOLLARI

Bilgi dışında kağıt düzeni, yazı güzelliği gibi durumlara itibar edilmemelidir.

Kağıtlardaki isimleri kapatmalıdır.

Kağıtların her birini tek tek ayrıca okumak yerine soru soru okunabilir.

Ayrıntılı bir cevap anahtarı çıkarılmalıdır. Cevap anahtarı bütüncül (holistik) rubrik biçiminde olabileceği gibi, analitik rubrik biçiminde de olabilir.

Örnek

Özümlenme organı olarak yaprakların yassı ve geniş yüzeyli oluşları bitkiye ne gibi yararlar sağlar ve ne gibi güçlükler çıkarır?

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

Butüncül (Holistik)

Rubrik Örneği

Örnek

Özümlenme organı olarak yaprakların yassı ve geniş yüzeyli oluşları bitkiye ne gibi yararlar sağlar ve ne gibi güçlükler çıkarır?



Resim 1

Gökçöy'ün balıkçılıkla uğraşırken dönemi hali



Resim 2

Gökçöy'ün fabrikalar açıldıktan sonraki hali

Gökçöy'ün eski ve yeni halinin avantajlı ve dezavantajlı yönlerini yazınız. Siz hangi halinde yaşamak istersiniz? Neden?

Sınıflamalar	Tanımlamalar	Puan
Mükemmel	Köyün eski ve yeni halinin avantajlı ve dezavantajlı yönlerine ilişkin iki unsuru belirtmiştir. Köyün hangi halinde yaşamak istediğine ilişkin mantıklı iki gerekçe açıklamıştır.	3
İyi	Köyün eski ve yeni halinin avantajlı ve dezavantajlı yönlerine ilişkin iki unsuru belirtmiştir. Köyün hangi halinde yaşamak istediğini belirtmeden yazmıştır.	2
Gelişimi için gerek	Köyün eski ve yeni halinin avantajlı ve dezavantajlı yönlerini eksik yazmış. Köyün hangi halinde yaşamak istediğini yazmamıştır.	1
Yetersiz	İlgisiz yanıtlar vermiştir.	0

Yazılı Anlatım

Analytik Rubrik

Örneği

Örnek

Ölçütler	Başlangıç Dönemi (1)	Gelişim Dönemi (2)	Kabul Edilebilir (3)	Tam Olarak Başarılı (4)	Başarı Puanı
İçerik	Yanıtlarda (seçilen örnekler, örnekler, örnekler, örnekler vb.) verilen konuyla ya da olayla ilgili açıklama ya da yorum yeterince detaylı değildir. Belirtilen bağlamlarda yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Yanıtlarda (seçilen örnekler, örnekler, örnekler vb.) verilen konuyla ya da olayla ilgili açıklama ya da yorum yeterince detaylı değildir. Belirtilen bağlamlarda yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Yanıtlarda (seçilen örnekler, örnekler, örnekler vb.) verilen konuyla ya da olayla ilgili açıklama ya da yorum yeterince detaylı değildir. Belirtilen bağlamlarda yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Yanıtlarda (seçilen örnekler, örnekler, örnekler vb.) verilen konuyla ya da olayla ilgili açıklama ya da yorum yeterince detaylı değildir. Belirtilen bağlamlarda yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	
Seçilen Yöntem	Belirtilen bağlamlarda yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Belirtilen bağlamlarda yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Belirtilen bağlamlarda yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Belirtilen bağlamlarda yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	
Anlatım	Yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	
Baglamsal Uygunluk	Belirtilen bağlamlarda yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Belirtilen bağlamlarda yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Belirtilen bağlamlarda yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Belirtilen bağlamlarda yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	
Yazım Kuralları	Yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	
Ölçüt	Yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	Yanıtlanmayan soruların yanı sıra, yanlış ve doğru aktarmada yetersizdir.	

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 bağlamlarda yazılanların neredeyse tümü, hatalı ve iletişimi aktarmada yetersizdir.	Yazılanlar (seçilen sözcükler, cümleler, deyimler, örnekler vb.) verilen konuyla ya da olayla çok az ilişkili ve yeteri kadar destekleyici değildir. Belirlenen bağlamlarda yazılanların çok azı, doğru ve iletişimi aktarmada yeterlidir.	Yazılanlar (seçilen sözcükler, cümleler, deyimler, örnekler vb.) verilen konuyla ya da olayla oldukça ilişkili ve birbirini yeterince destekleyicidir. Belirlenen bağlamlarda yazılanların çoğu doğru ve iletişimi aktarmada yeterlidir.	Yazılanlar (seçilen sözcükler, cümleler, deyimler, örnekler vb.) verilen konuyla ya da olayla tam olarak ilişkili ve birbirini destekleyicidir. Belirlenen bağlamlarda yazılanların tümü doğru ve iletişimi aktarmada yeterlidir.	

KISA CEVAPLI SORULAR

Kısa cevaplı testler, öğrencinin bir sözcük, bir rakam, bir tarih ya da en çok bir cümle ile cevaplandırabileceği sorulardan oluşur.

Soru 3 Pufi

Pufi ile Mehmet Amca'nın listelerini düşünün. Pufi'nin listesinde olabilecek bir etkinlik yazınız?

Pufi'nin listesinde olabilecek bir etkinlik yazınız?

Kısa cevaplı testler, öğrencinin aklının estiği cevabı yazmasını büyük ölçüde engelleyecek bir özelliğe sahiptir. Bundan ötürü maddeleri cevaplandıran 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 yetiştiği bölgenin adını karşısındaki boşluğa yazınız.

- A. Fındık _____
B. Tütün _____
C. Buğday _____
D. İncir _____
E. Zeytin _____



KISA CEVAPLI SORULAR AVANTAJLAR

Hazırlaması kolaydır.

Uygulaması kolaydır.

Puanlaması kolaydır.

Çok sayıda soru sorulabilir.

Sistemik hata karışması ihtimali yoktur.

Kullanışlılık artar.

1. Güvenilirlik artar.
2. Geçerlilik artar (?)

Geçerlilik artar.



KISA CEVAPLI SORULAR AVANTAJLAR

İyi yazılmazsa genellikle üst düzey bilişsel özellikleri ölçemez.

Analiz
Değerlendirme
Yaratıcılık



KISA CEVAPLI SORULARIN PUANLANMASI

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.

Soru 3: YÜKSEK BİNALAR

R419Q05 - 0 1 2 9

Norveç'te, Oslo'da bulunan Radisson SAS Plaza binası sadece 117 metredir. Bu bina neden Şekil 2'de yer almıştır?

YÜKSEK BİNALAR PUANLAMA 3

Tam Puan

Bu makalenin bir Norveç dergisinden alındığı ya da okuyucuların Norveçli olduğu gerçeğine değinir.

- Yazı bir Norveç dergisinden alınmış
- Yazı Norveç'te yaşayan insanlar için yazılmış, bu şekilde onlara bir bakış açısı kazandıracaktır
- Norveç'te yaşayanlara, bu ülkede gerçekten yüksek binalar bulunmadığını göstermek için!

Kısmi Puan

Norveç'ten söz etmeden, SAS Plaza'nın referans ya da karşılaştırma noktası olduğuna değinir.

- Karşılaştırma için
- Referans noktası olması için
- Bu şekilde yapılan bu derecelendirmenin ne anlama geldiğini anlarsınız.

Sıfır Puan

Yetersiz veya belirsiz bir yanıt verir.

- 117 metre halen oldukça yüksek sayılır
- Çünkü o dünyadaki en yüksek binalardan birisidir
- O, 30 ya da daha fazla kat yüksekliğindedir
- O, Norveç'teki en yüksek binadır. [fazla belirsiz - metnin okuyucusu ile bağlantı kurmuyor]
- Çünkü o Norveçli. [fazla belirsiz - metnin okuyucusu ile bağlantı kurmuyor]

Metni doğru anlamadığını ortaya koyar veya akla uygun olmayan ya da ilgisiz bir yanıt verir.

- Otel olan sadece o. [doğru, fakat bu metinde yer alma nedeni değil]
- Sivri bir kulesi olmayan sadece o. [doğru, fakat bu metinde yer alma nedeni değil]

Boş

Tam Puan

Bu makalenin bir Norveç dergisinden alındığı ya da okuyucuların Norveçli olduğu gerçeğine değinir.

- Yazı bir Norveç dergisinden alınmış.
- Yazı Norveç'te yaşayan insanlar için yazılmış, bu şekilde onlara bir bakış açısı kazandıracaktır.
- Norveç'te yaşayanlara, bu ülkede gerçekten yüksek binalar bulunmadığını göstermek için!

Kısmi Puan

Norveç'ten söz etmeden, SAS Plaza'nın referans ya da karşılaştırma noktası olduğuna değinir.

- Karşılaştırma için.
- Referans noktası olması için.
- Bu şekilde yapılan bu derecelendirmenin ne anlama geldiğini anlarsınız.

Sıfır Puan

Yetersiz veya belirsiz bir yanıt verir.

- 117 metre halen oldukça yüksek sayılır.
- Çünkü o dünyadaki en yüksek binalardan birisidir.
- O, 30 ya da daha fazla kat yüksekliğindedir.
- O, Norveç'teki en yüksek binadır. *[fazla belirsiz – metnin okuyucusu ile bağlantı kurmuyor]*
- Çünkü o Norveçli. *[fazla belirsiz – metnin okuyucusu ile bağlantı kurmuyor]*

Metni doğru anlamadığını ortaya koyar veya akla uygun olmayan ya da ilgisiz bir yanıt verir.

- Otel olan sadece o. *[doğru, fakat bu metinde yer alma nedeni değil]*
- Sıvı bir kulesi olmayan sadece o. *[doğru, fakat bu metinde yer alma nedeni değil]*

Boş.





Soru 31 BESİN PİRAMİDİ

Düzenli koşu yapan bir kişinin **mayıs ayı başındaki kilosu 65 kg'dır**. Bu kişi, bir süre iki farklı beslenme programı uygulamıştır. Beslenme programlarını tamamladıktan sonra elde ettiği sonuçlar aşağıda verilmiştir.

Mayıs Ayı Beslenme Programı

- Şeker, katı yağ ve surulebilir ürünler ağırlıklı beslenmiştir.
- Mayıs ayının sonunda tartıldığında **70 kg'a** ulaşmıştır.
- Bu beslenme programı sonucunda **kilo aldığını** fark etmiştir.

Haziran Ayı Beslenme Programı

- Sebze ve meyve ağırlıklı beslenmiştir.
- Haziran ayının sonunda tartıldığında **65 kg'a** düşmüştür.
- Bu beslenme programı sonucunda **kilo verdiğini** fark etmiştir.

İki farklı beslenme programını uygulayan bu kişinin cevaplamaya çalıştığı soru nedir? Yazınız.



31.

Beslenme programları ile kilo alıp verme arasındaki ilişkiyi ifade eden sorular.

Örnekler:

- Beslenme programlarımız kilo alıp vermemizi etkiler mi?
- Hangi beslenme programını uygularsak kilo alırız?
- Hangi beslenme programı kilo vermemizi sağlar?

KD: Mayıs ayı beslenme programının kilo aldırıldığını ya da haziran ayı beslenme programının kilo verdirdiğini ifade eder.

Diğer tüm yanıtlar



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.

Örnek

Dekametre, metrik sistemde metrenin _____ katıdır.

- ☐ Cümle tamamlama soruları boşluk doldurma olarak da geçmektedir.
- ☐ Hazırlanması oldukça kolaydır. Bu yargı, özellikle *kim?*, *ne?*, *nerede?*, *ne zaman?* sorusuna cevap olabilecek olgusal bilgileri ölçmeye yönelik soruların yazılmasında geçerlidir.
- ☐ Cümleler ders kitabından aynen alınmamalı, öğretmen kendi sözcükleriyle soruları yeniden yazmalıdır.



CÜMLE TAMAMLAMA SORULARININ AVANTAJLARI

Hazırlaması kolaydır.

Uygulaması kolaydır.

Puanlaması kolaydır.

Çok sayıda soru sorulabilir.

Sistematiik hata karışması ihtimali yoktur.

Kullanışlılık artar.

1. Güvenilirlik artar.
2. Geçerlilik artar (?)

Geçerlilik artar.



CÜMLE TAMAMLAMA SORULARININ DEZAVANTAJLARI

İyi yazılmazsa genellikle üst düzey bilişsel özellikleri ölçemez.

Analiz
Değerlendirme
Yaratıcılık

14- TEST GELİŞTİRME VE MADDE/SORU TÜRLERİ-6 (Sayfa: 97-99)

EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ



CÜMLE TAMAMLAMA SORULARI YAZIM KURALLARI

Her kısa cevap maddesi, yalnızca tek bir doğru cevabı olacak biçimde yapılandırılmalıdır. Tartışma götürecek, 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 vermekten kaçınılmalıdır. Doğru cevabın verilmesinde kullanılabilecek ipucu, maddede verilen gereksiz bir bilgi ya da cümlelerin gramer yapısı olabilir.

Yanlış örnek

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

Doğru örnek

Osmanlı Devleti, Birinci Dünya Savaşına _____ yılında girmiştir.

Cumlenin gramer yapısı bağlamında ipucu genellikle eklerle verilmektedir.

Yanlış örnek

Huzur romanının yazarı _____'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 kağıdına işaretletilmesi uygun görülmemektedir, ancak ortaöğretimde bunun pek bir sakıncası bulunmamaktadır.

Bir cümlede çok sayıda boşluk bırakılmamalı ve cümlelerden sadece anahtar niteliğindeki anlamlı ve önemli sözcükler çıkarılmalıdır. Eksik cümle istenileni anlatacak biçimde yapılanmış olmalıdır.

Yanlış örnek

_____, _____, _____ bölmekle elde edilir.

Doğru örnek

_____ bölüneni bölene bölmekle elde edilir.

Herkesçe aynı biçimde algılanacak, öğrencinin «acaba öğretmenin bununla neyi kastediyor?» sorusunu sormayacağı biçimde soru yazmak gerekir. Öğrenci, öğretmenin zihninden geçeni keşfetmek zorunda değildir.

Yanlış örnek

Telefon _____ bulunmuştur.

Doğru örnek

Telefon _____ yılında bulunmuştur.

Telefon _____ tarafından bulunmuştur.



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.



SÖZLÜ YOKLAMALARIN ÖZELLİKLERİ

- ☐ 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 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 kıyafeti, diksiyonu, hal ve hareketleri puanlamaya artı ya da eksi yönde etki edebilir.
- ☐ Çok fazla soru sorulamayacağı için kapsam geçerliliği problemi olabilir.



SÖZLÜ YOKLAMA KURALLARI

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



SÖZLÜ YOKLAMA

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 ilköğretimde, çocukların gelişim süreçleri de dikkate alındığında, madde/soru çeşidi sayısını çok arttırmamak gerekir.

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



DESTEKLEYİCİ DEĞERLENDİRME YAKLAŞIMLARI

Destekleyici Yöntemler (Alternative Assessment) nedir?

Desteklenen nedir?

Performans görevi nedir?

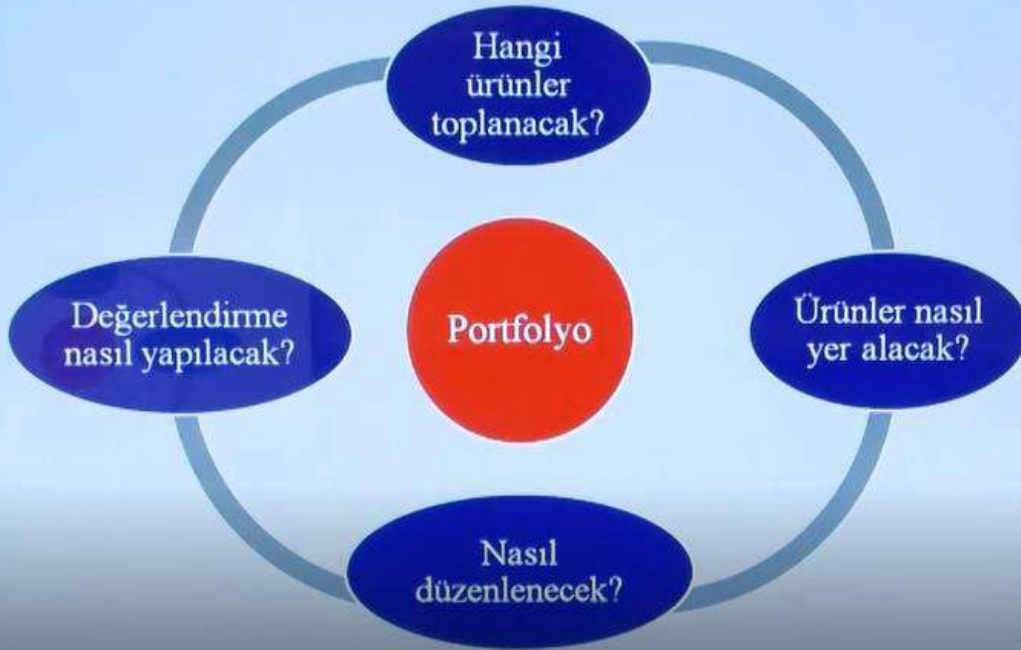
Portfolyo (ürün seçki dosyası) nedir?

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 özgeçmişi, 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:

- Süreci yansıtan portfolyolar: Öğrencinin öğrenme ve gelişim sürecini yansıtır. Başlangıç çalışmalarını, süreçteki çalışmaları, karşılaşılan güçlükleri ve öğrenme ürünlerini içerir.
- Ürünü yansıtan portfolyolar: Öğrenme sürecinden çok bitmiş görevleri içerir. Öğrencinin en iyi olduğunu düşündüğü çalışmaları içerir.

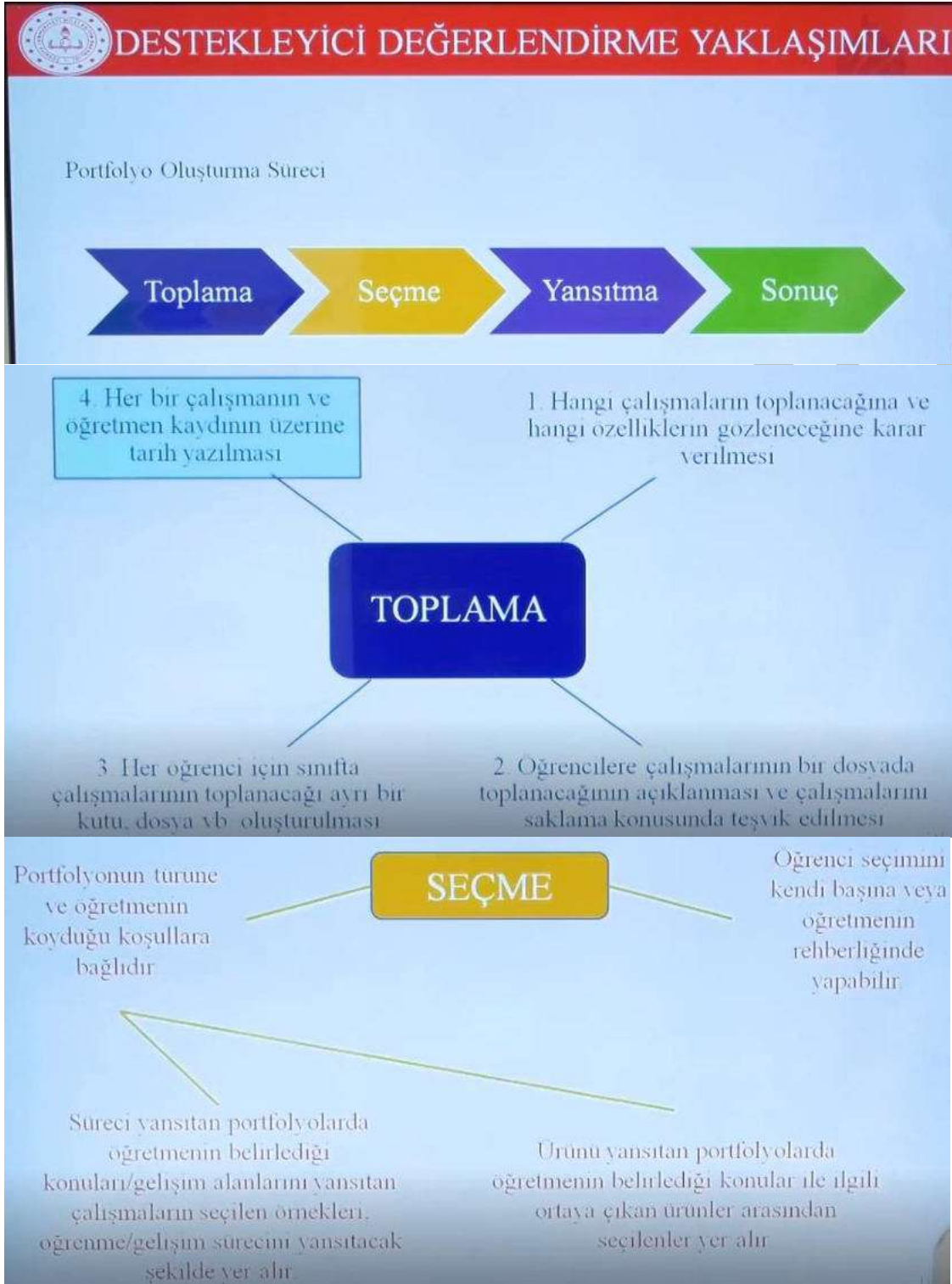


Portfolyonun içeriği:

- Öğretmen kayıtları (gözlemler, anekdot kayıtları)
- Öğrencinin çalışmaları
- Öğrencinin sözel ve psiko-motor 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

15- TEST GELİŞTİRME VE MADDE/SORU TÜRLERİ-7 (Sayfa: 99-102)

EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ



YANSITMA

Portfolyoyu herhangi bir çalışma dosyasından ayıran en önemli aşamadır. Bu aşamada öğrenci

1. portfolyosuna seçtiği her bir çalışmayı niçin seçtiğini açıklar,
2. 2. çalışmalarını yaparken geçirdiği süreci ve bu süreçte öğrendiklerini anlatır,
3. 3. kendi başarısını görür, bunu ifade eder ve değerlendirme sürecine katılır.



SONUÇ

- ❖– Bu aşamada öğrenci, “bu çalışmayı niçin yaptık?” sorusunu yanıtlar
- ❖– Okulda yaptığı çalışmalarla öğrendikleri arasında somut bağlar kurar.
- ❖– Tamamlanan portfolyo çalışmalarının öğrenci tarafından sınıf arkadaşları, öğretmeni ve ailesinden oluşan bir gruba sunumu yapılmalıdır.
- ❖– Portfolyonun sunumu, öğrencinin çalışmalarına önem vermesini sağlar ve kendine olan güvenini artırır.

PERFORMANS DEĞERLENDİRME



ÖZDEĞERLENDİRME: Öğrencinin belirli bir konuda (örneğin bir ürünü ortaya koymada gösterdiği performans vb.) kendi kendisini değerlendirmesine denir.

- ❑ Öğrencilerin kendi özellikleriyle (yetenek, ilgi, beceri vb.) ilgili farkındalığının artmasını, zayıf ve güçlü yönlerini keşfetmesini sağlar.
- ❑ Öz-düzenleme becerisi artar.
- ❑ Ölçütlü düşünme becerisi artar.
- ❑ Öğrenme motivasyonunu artırır.

ÖZDEĞERLENDİRMENİN OLASI DEZAVANTAJLARI:

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

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.

AKRAN DEĞERLENDİRMENİN OLASI DEZAVANTAJLARI:

- ❑ Öğrencilerin yanlış davranışları söz konusu olabilir. Kişisel ilişkiler olumlu ya da olumsuz yönde değerlendirmeye etki edebilir.
- ❑ Genel izlenimle puan verme söz konusu olabilir.

RUBRİK (DERCELİ PUANLAMA ANAHTARI): Performans ürünlerinin değerlendirilmesinde en sık kullanılan araçlardan biridir. Amaç öğretmen tarafından ürünün genel izlenimle puanlamasındaki özneliğini azaltmaktır. Rubrikler ikiye ayrılır:

1. Butuncül (Holistik) Rubrik
2. Analitik Rubrik

Butuncül (Holistik) Rubrik Örneği: Bir Maket Yapımı Örneği

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

Analitik Rubrik Örneği: Bir Maket Yapımı 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 öğeleriyle yansıtmış	
Renklendirmeler uyumludur.	Renklendirmede özensiz	Kısmen uyumlu	Renklendirmeler çok iyi	
*****	*****	*****	*****	
*****	*****	*****	*****	
*****	*****	*****	*****	
Toplam				

- ☐ Rubrikler performans görevleriyle birlikte öğrenciye verilmelidir.
- ☐ Rubriklerin geliştirilmesi uzmanlık gerektirir. Rubrikler için geçerlilik ve güvenilirlik çalışmaları yapılmalıdır.

Kontrol Listeleri Gözlenen performans ürününün ölçütlere uygunluğunu «evet-hayır», «var-yok», «gösterdi-göstermedi» vb. bir biçimde kategorik (1-0) olarak puanlama amacıyla kullanılan araçlardır.

- ☐ Özellikle sergilenecek performans detaylı ve ardışık eylemler gerektirdiği zamanlarda kullanışlıdır. *Örneğin bir deneyin eyleme dökülmesi vb.*
- ☐ Pek çok işlem adımında oluşan performanstaki eksik adımları belirlemek için oldukça uygundur.

ÇOKTAN SEÇMELİ MADDE YAZMA KONTROL LİSTESİ

	Evet	Hayır
1. Öğrencinin seçenekleri görmeden yanıt vermesi olanaklıdır.	☐	☐
2. Öğrencinin sözcük/kavram repertuarında bulunmayan terim vb. <u>kullanılmamıştır.</u>	☐	☐
3. Öncülde verilen resim/tablo vb., maddenin yanıtlanması için gereklidir.	☐	☐
4. Madde içeriği cinsiyet vb. gruplar açısından yanlılık <u>oluşturmamaktadır.</u>	☐	☐
5. Tüm öğrenciler aynı şeyi anlar.	☐	☐
6. Anlamayı güçleştirecek bir dil <u>kullanılmamıştır.</u>	☐	☐
7. Tüm çeldiricilerin bir gerekçesi vardır.	☐	☐
8. Gereksiz sözcük ve/veya cümle <u>bulunmamaktadır.</u>	☐	☐
9. Seçeneklerde sözcük tekrarı <u>yapılmamıştır.</u>	☐	☐
10. Seçenekler gramatik açıdan birbirleriyle uyumludur.	☐	☐
11. Seçeneklerde ipucu niteliğinde ek vb. <u>bulunmamaktadır.</u>	☐	☐
12. "Bazen, çoğunlukla, genellikle, sıklıkla, zaman zaman, hiçbir zaman, asla, daima, hiçbir, her zaman, tümü" vb. sözcükler seçeneklerde <u>bulunmamaktadır.</u>	☐	☐
13. Birbirinin anlamca zıttı olan durumlar seçeneklerde <u>bulunmamaktadır.</u>	☐	☐
14. Seçenekler birbirini kapsamamaktadır.	☐	☐
15. Seçeneklerin uzunlukları eşittir/yakındır.	☐	☐
16. "Hepsi ve hiçbir" seçenek olarak <u>kullanılmamıştır.</u>	☐	☐
17. Madde kökü koyu yapılmıştır.	☐	☐
18. Seçenekler uygun bir sırayla verilmiştir.	☐	☐
19. Seçeneklerin baş harfleri büyüktür.	☐	☐
20. Maddenin tamamı aynı sayfadadır.	☐	☐
21. Ortak öncüllü sorularda madde sayısı üçü geçmemektedir.	☐	☐
22. Öncül numara olarak verilmiş ise numaralar seçeneklerde dengeli bir biçimde dağıtılmıştır.	☐	☐
23. Anlatım bozukluğu, ikili/üçlü anlam, muğlak anlatım vb. <u>bulunmamaktadır.</u>	☐	☐
24. Gereksiz yere uzatılmamıştır.	☐	☐
25. Doğru yanıt başka bir maddede <u>bulunmamaktadır.</u>	☐	☐
26. Doğru yanıtlar seçeneklere örüntü <u>ıçermeyecek</u> bir biçimde dağıtılmıştır.	☐	☐
27. Doğru yanıtlar seçeneklere eşit dağıtılmıştır.	☐	☐

Dereceleme Ölçekleri: Bu araçların kullanımında performansa 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
Sırasını bekleyebilir.	☐	☐	☐	☐	☐
Duygularını kontrol edebilir.	☐	☐	☐	☐	☐
Kendi arkadaş seçebilir.	☐	☐	☐	☐	☐
Oyunun kurallarını açıklayabilir.	☐	☐	☐	☐	☐
Kendi başına 20 dk. oynayabilir.	☐	☐	☐	☐	☐
Kendiliğinden özür dileyebilir.	☐	☐	☐	☐	☐
Duygularını ifade edebilir.	☐	☐	☐	☐	☐
Oyuncak vb. eşyalarını paylaşabilir.	☐	☐	☐	☐	☐

Kesinlikle Katılmıyorum	Katılmıyorum	Kararsızım	Katılıyorum	Kesinlikle Katılıyorum
Kesinlikle Haksız	Büyük Bir Olasılıkla Haksız	Şüpheli	Büyük Bir Olasılıkla Haklı	Kesinlikle Haklı
Hiç	Az	Orta	Çok	Tam
Çok Düşük	Ortanın Biraz Altında	Orta	Ortanın Biraz Üstünde	Çok Yüksek
Hiç Hoşlanmıyorum	Pek Hoşlanmıyorum	Kararsızım	Biraz Hoşlanıyorum	Çok Fazla Hoşlanıyorum
Hiçbir Zaman	Nadiren	Bazen	Genellikle	Her Zaman
Çok Zayıf	Zayıf	Orta	İyi	Çok İyi
Kesinlikle Onaylamıyorum	Onaylamıyorum	Onaylıyorum	Kesinlikle Onaylıyorum	
Kesinlikle Katılmıyorum	Katılmıyorum	Katılıyorum	Kesinlikle Katılıyorum	
Kötü	Orta	İyi	Çok İyi	
Hayır Katılmıyorum	Kısmen Katılıyorum	Evet Katılıyorum		
Hiçbir Zaman	Bazen	Her Zaman		

Gözlem Formları: 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.

GELENEKSEL YÖNTEMLER

- ☐ Ürün değerlendirilir.
- ☐ Öğrencinin ulaştığı noktanın tespiti önemlidir.
- ☐ Essay dışında genellikle üst düzey düşünme becerilerinin değerlendirilmesinde yetersizdir.
- ☐ Değerlendirme öğrenmeden ayırdır.
- ☐ Bireyden ziyade gruba odaklıdır.
- ☐ Başarının bireyler arası değerlendirilmesine odaklıdır.
- ☐ Geçerlilik ve güvenilirlik kontrolü daha kolaydır.

DESTEKLEYİCİ YÖNTEMLER

- ☐ Süreç ve ürün birlikte değerlendirilir.
- ☐ Ne öğrendikleri yanında, öğrendiklerini nasıl kullandıklarıyla ilgilenilir.
- ☐ Üst düzey bilişsel düşünme becerilerine odaklanır.
- ☐ Değerlendirme öğrenmeyle bütünleşmiştir.
- ☐ Odak noktası gruptan ziyade bireydir.
- ☐ Öğrencinin bireysel olarak gelişimine odaklıdır.
- ☐ Geçerlilik ve güvenilirlik problemi olabilir.

16- TEST VE MADDE İSTATİSTİKLERİNE GENEL BİR BAKIŞ (Sayfa: 102-104)

EĞİTİMEN: DOÇ.DR. GÜÇLÜ ŞEKERCİOĞLU-AKDENİZ ÜNİVERSİTESİ



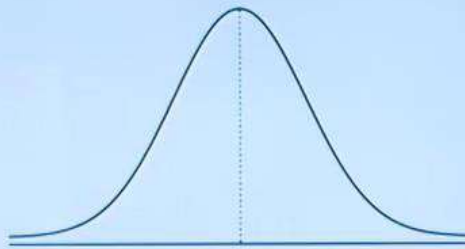
İÇERİK

- Test istatistikleri
 - Merkezi eğilim ölçüleri
 - Değişkenlik ölçüleri
 - Dağılım özellikleri
- Madde istatistikleri
 - Madde Güçlüğü
 - Madde Ayırtediciliği

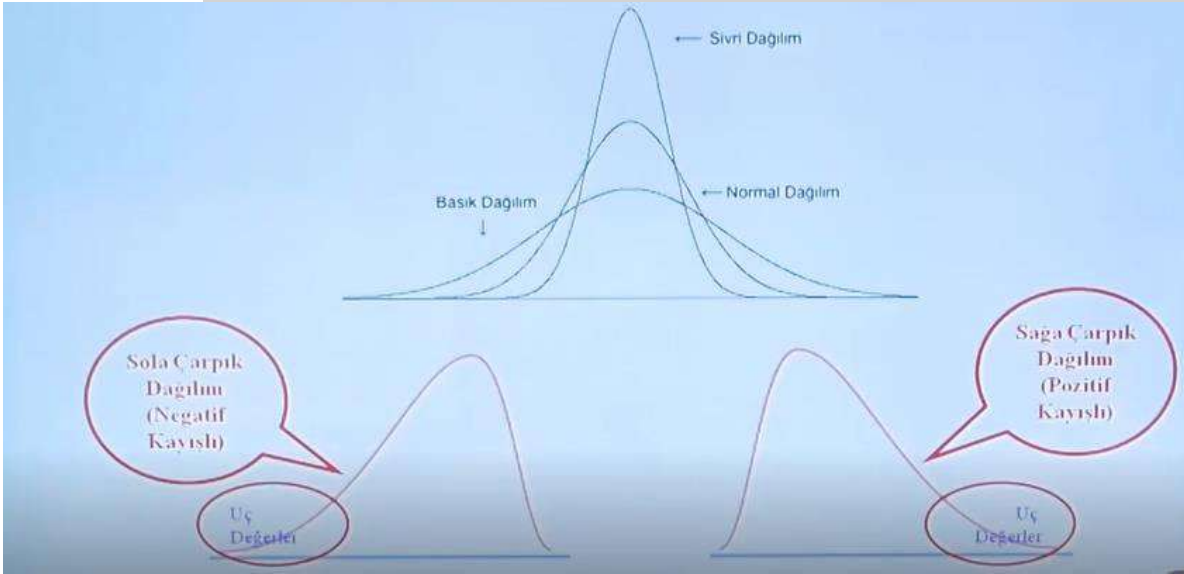


TEST İSTATİSTİKLERİ

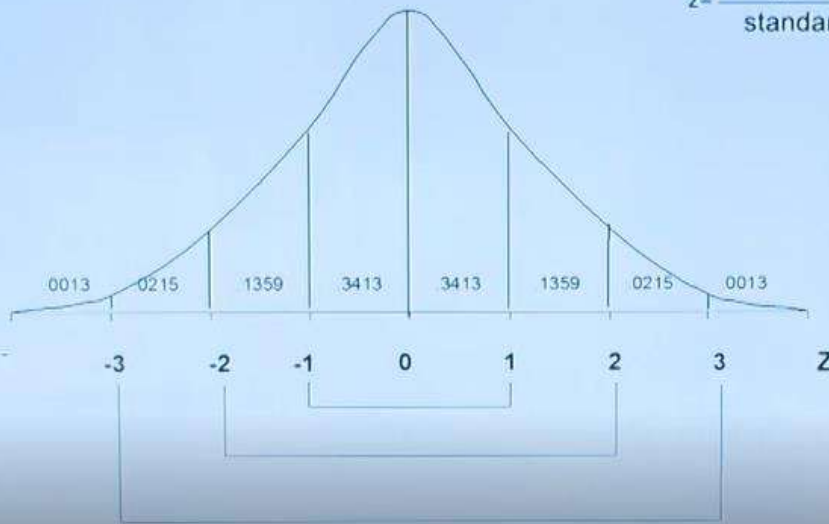
- Merkezi eğilim ölçüleri aritmetik ortalama, medyan (ortanca) ve moddur.
- Merkezi eğilim ölçüsü hesaplamak ne işimize yarar? Ne tür kararlar almamızı sağlar?
- Değişkenlik ölçüleri ranj (dizi genişliği), varyans, standart sapma ve çeyrek sapmadır.
- Değişkenlik ölçülerini hesaplamak ne işimize yarar? Ne tür kararlar almamızı sağlar?
- Normal dağılım (Gauss dağılımı / çan eğrisi) nedir?



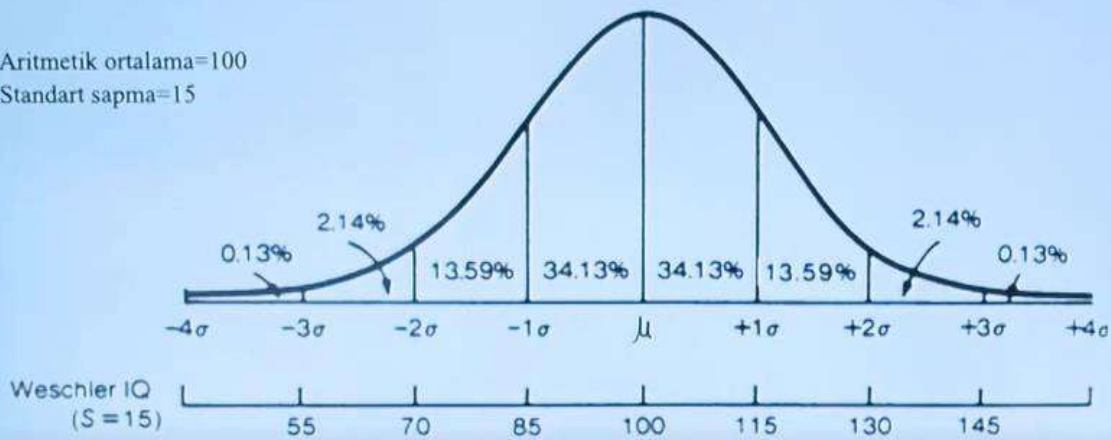
- Bir ölçme işleminde dağılımın normal olması ne anlama gelir?
- Bir ölçme işleminde dağılımın hep normal mi dağılır?



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



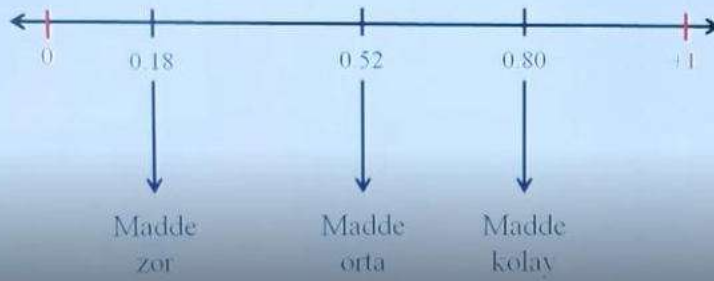
Aritmetik ortalama=100
Standart sapma=15



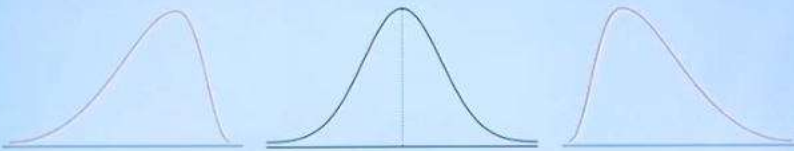


MADDE İSTATİSTİKLERİ

- Madde güçlük indeksi nedir?
- Madde güçlük indeksi nasıl yorumlanır?



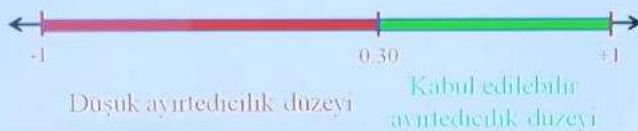
- Testin ortalama güçlüğü nedir?
- Testin ortalama güçlüğü ile dağılımlar arasında nasıl bir ilişki vardır?

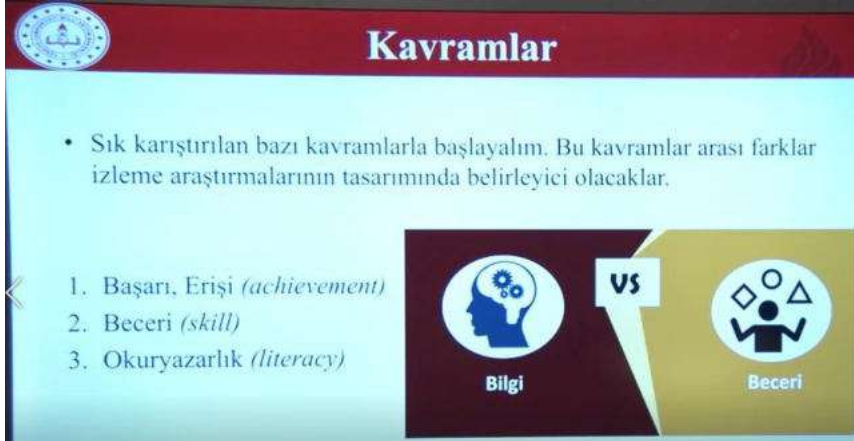


- Testin ortalama güçlüğü ile güvenilirlik arasında nasıl bir ilişki vardır?

- ! Hazırbulunuşluluğu ölçmeyi ya da öğrenme eksiklerini tespit etmeyi ya da not vermeyi amaçlayan bir öğretmen testinin/sınavının ortalama güçlüğüne ne düzeyde tutmalıdır? Soruların güçlük dağılımı nasıl olmalıdır?
- ! Bir sınav zor olursa ne olur?
- ! Bir sınav kolay olursa ne olur?
- ! Bir sınav ortalama güçlükte olursa ne olur?

- Madde ayırtedicilik indeksi nedir?
- Madde ayırtedicilik indeksi nasıl yorumlanır?



17-ULUSLARARASI EĞİTİM İZLEME ARAŞTIRMALARI (Sayfa: 105-107)**EĞİTİM: DR. HAYRİ EREN SUNA-BAKANLIK MÜŞAVİRİ**

Kavramlar

- Sık karıştırılan bazı kavramlarla başlayalım. Bu kavramlar arası farklar izleme araştırmalarının tasarımında belirleyici olacaklar.



Bilgi

VS

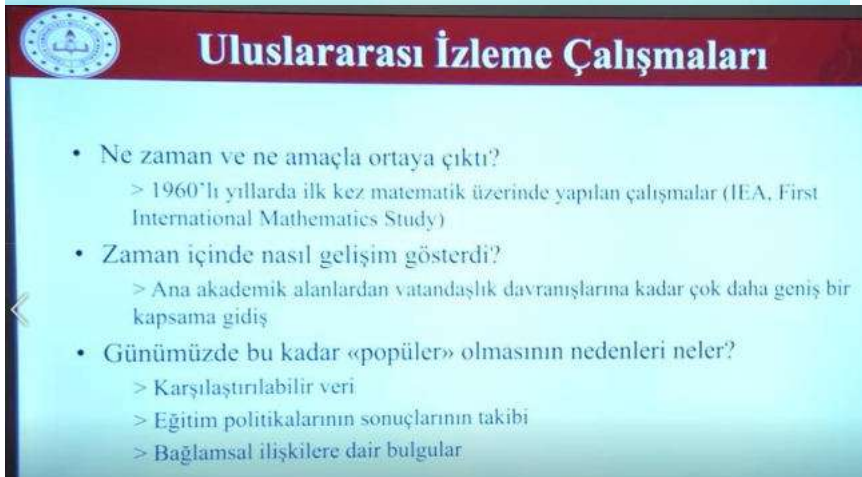


Beceri

1. Başarı, Erişi (*achievement*)
2. Beceri (*skill*)
3. Okuryazarlık (*literacy*)

- Eğitim ve kariyer açısından hangi kavram daha önemli?
- Güncel eğitim politikaları hangi kavram(lar) daha fazla odağına alıyor?

1. Başarı, Erişi (*achievement*)
2. Beceri (*skill*)
3. Okuryazarlık (*literacy*)



Uluslararası İzleme Çalışmaları

- Ne zaman ve ne amaçla ortaya çıktı?
 - > 1960'lı yıllarda ilk kez matematik üzerinde yapılan çalışmalar (IEA, First International Mathematics Study)
- Zaman içinde nasıl gelişim gösterdi?
 - > Ana akademik alanlardan vatandaşlık davranışlarına kadar çok daha geniş bir kapsama gidiş
- Günümüzde bu kadar «popüler» olmasının nedenleri neler?
 - > Karşılaştırılabilir veri
 - > Eğitim politikalarının sonuçlarının takibi
 - > Bağlamsal ilişkilere dair bulgular

Günümüzde en yoğun katılım gösterilen çalışmalar

1. Uluslararası Öğrenci Değerlendirme Programı (PISA)
 - > 3 yıllık döngülerle Ekonomik Kalkınma ve İşbirliği Örgütü (OECD) tarafından gerçekleştirilmektedir.
2. Uluslararası Matematik ve Fen Eğilimleri Araştırması (TIMSS)
 - > 4 yıllık döngülerle Uluslararası Eğitim Başarılarını Değerlendirme Kuruluşu (IEA) tarafından gerçekleştirilmektedir.
3. PIRLS-Uluslararası Okuma Becerilerinde Gelişim Araştırması (PIRLS)
 - > 5 yıllık döngülerle Uluslararası Eğitim Başarılarını Değerlendirme Kuruluşu (IEA) tarafından gerçekleştirilmektedir.

18-PISA: KAPSAMI VE SONUÇLARI (Sayfa: 107-112)

EĞİTMEN: DR. HAYRİ EREN SUNA-BAKANLIK MÜŞAVİRİ



PISA DÖNGÜLERİNDE TEMEL VE AĞIRLIKLI ALANLAR*

PISA 2000	PISA 2003	PISA 2006	PISA 2009	PISA 2012	PISA 2015	PISA 2018
<i>Okuma Becerileri</i>	Okuma Becerileri	Okuma Becerileri	<i>Okuma Becerileri</i>	Okuma Becerileri	Okuma Becerileri	<i>Okuma Becerileri</i>
Matematik Okuryazarlığı	<i>Matematik Okuryazarlığı</i>	Matematik Okuryazarlığı	Matematik Okuryazarlığı	<i>Matematik Okuryazarlığı</i>	Matematik Okuryazarlığı	Matematik Okuryazarlığı
Fen Okuryazarlığı	Fen Okuryazarlığı	<i>Fen Okuryazarlığı</i>	Fen Okuryazarlığı	Fen Okuryazarlığı	<i>Fen Okuryazarlığı</i>	Fen Okuryazarlığı

PISA 2018 TÜRKİYE ÖRNEKLEMİNİN OKUL TÜRLERİNE GÖRE DAĞILIMI

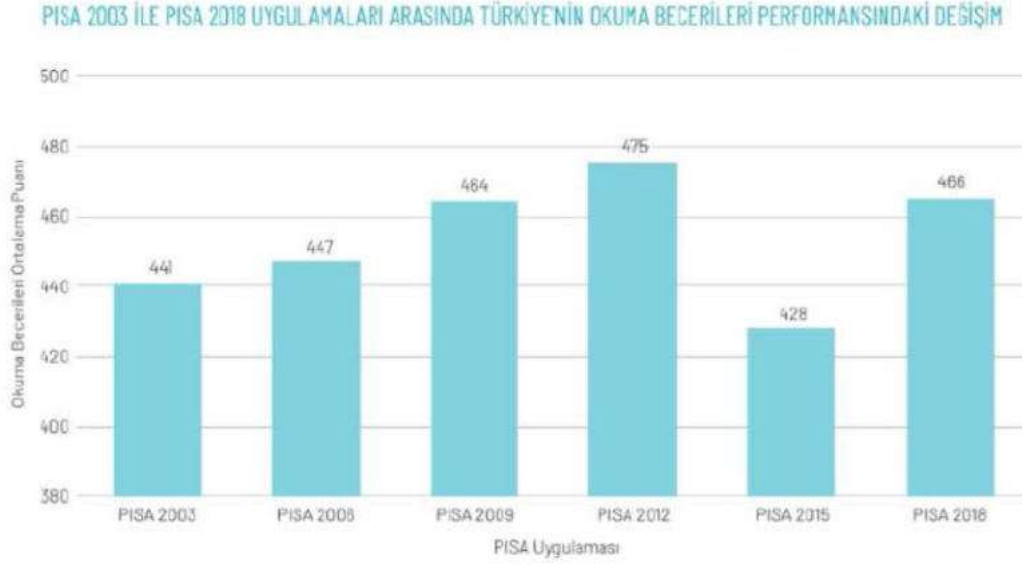


TÜRKİYE ÖRNEKLEMİNİN İBBS DÜZEY 1 BÖLGELERİNE DAĞILIMI

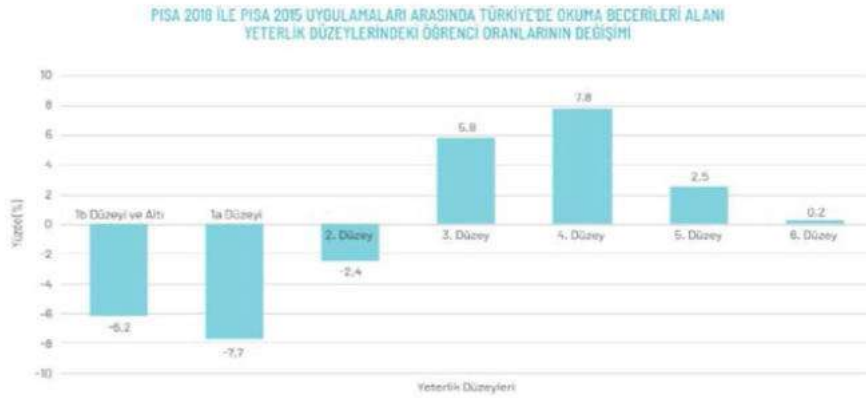


OKUMA BECERİSİ

Ortalama Puanlardaki Değişim



Yeterlik Düzeylerindeki Değişim

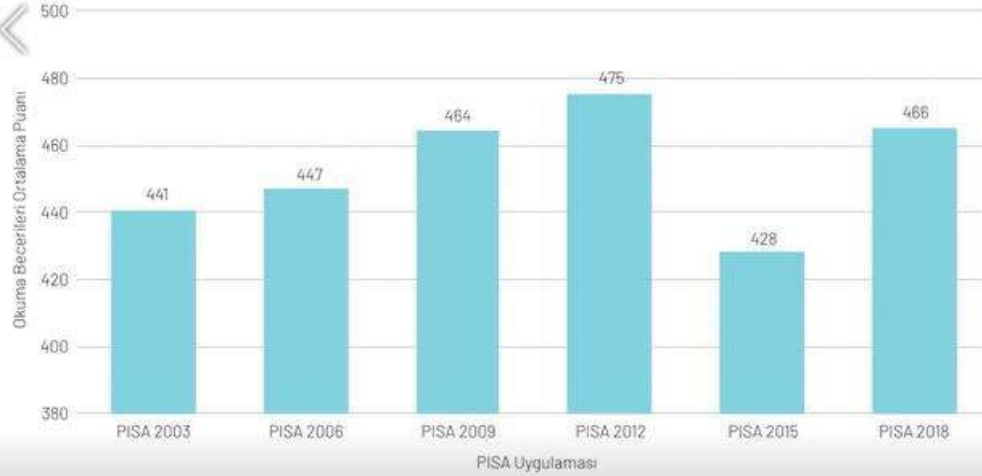


Şekil 1. Okuma becerisi ortalama puanlarındaki ve yeterlik düzeylerindeki değişim

Okuma Becerisi

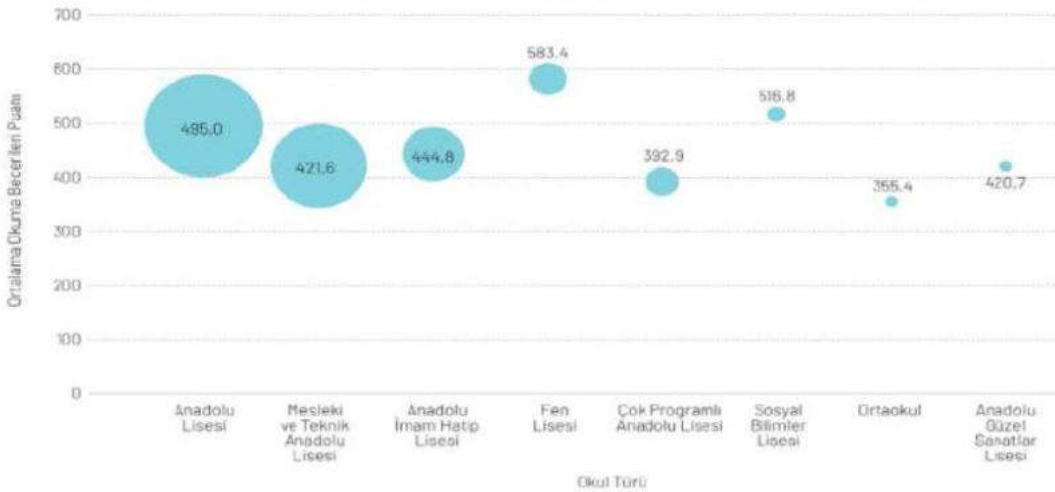
Ortalama Puanlardaki Değişim

PISA 2003 İLE PISA 2018 UYGULAMALARI ARASINDA TÜRKİYE'NİN OKUMA BECERİLERİ PERFORMANSINDAKİ DEĞİŞİM

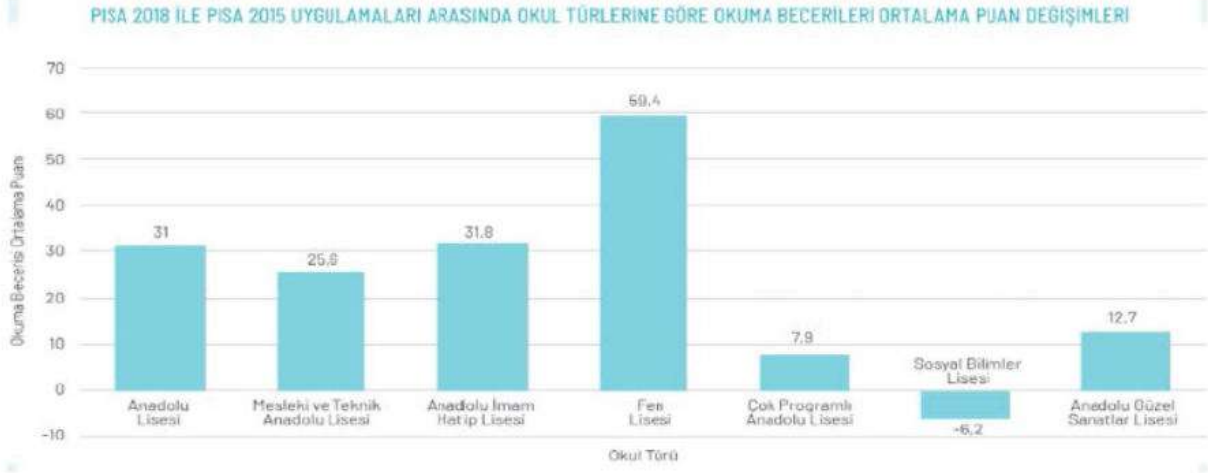


Okul Türlerine Göre Puan Ortalamaları

PISA 2018 OKUMA BECERİLERİ PERFORMANSLARININ OKUL TÜRÜNE GÖRE DEĞİŞİMİ***



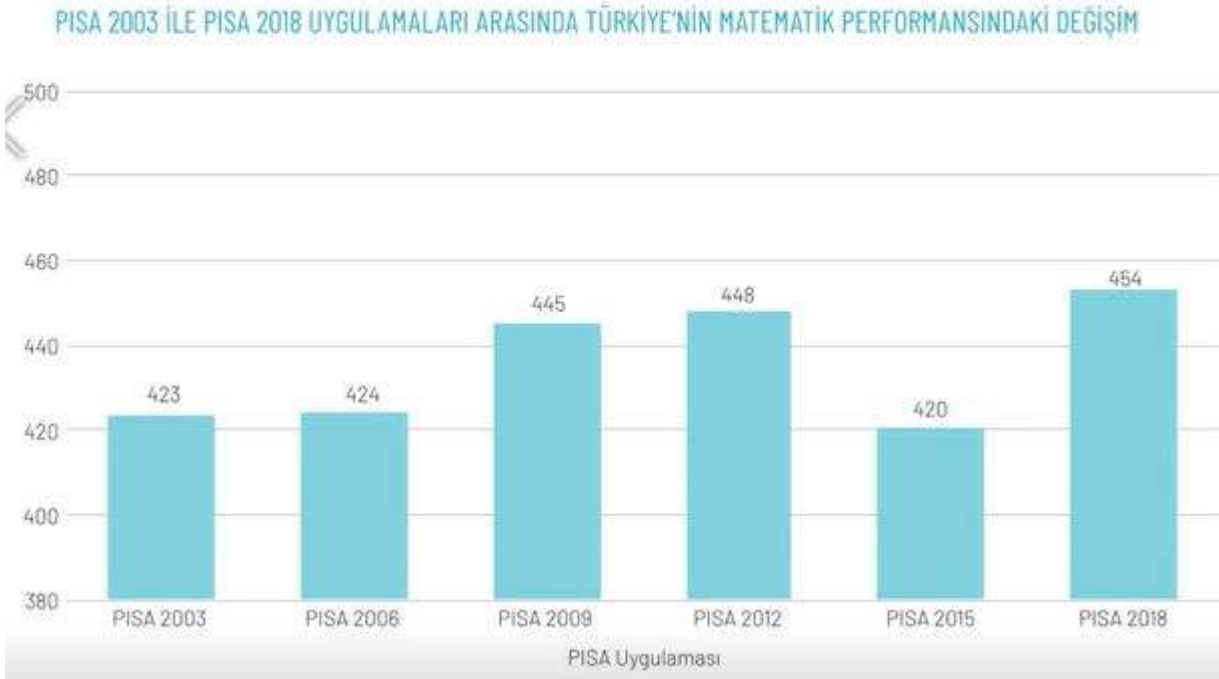
Önceki Döngü ile Performans Farkı



Şekil 2. Okul türlerine göre okuma becerisi puan ortalamaları ve performans farkları^W

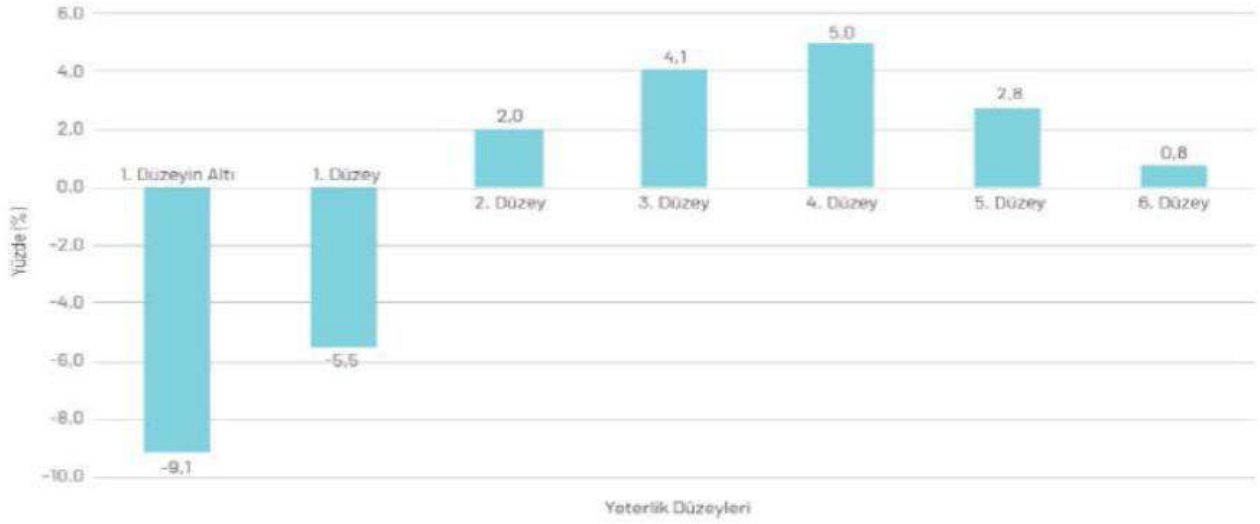
Matematik Okuryazarlığı

Ortalama Puanlardaki Değişim



Yeterlik Düzeylerindeki Değişim

PISA 2018 İLE PISA 2015 UYGULAMALARI ARASINDA TÜRKİYE'DE MATEMATİK ALANI YETERLİK DÜZEYLERİNDEKİ ÖĞRENCİ ORANLARININ DEĞİŞİMİ

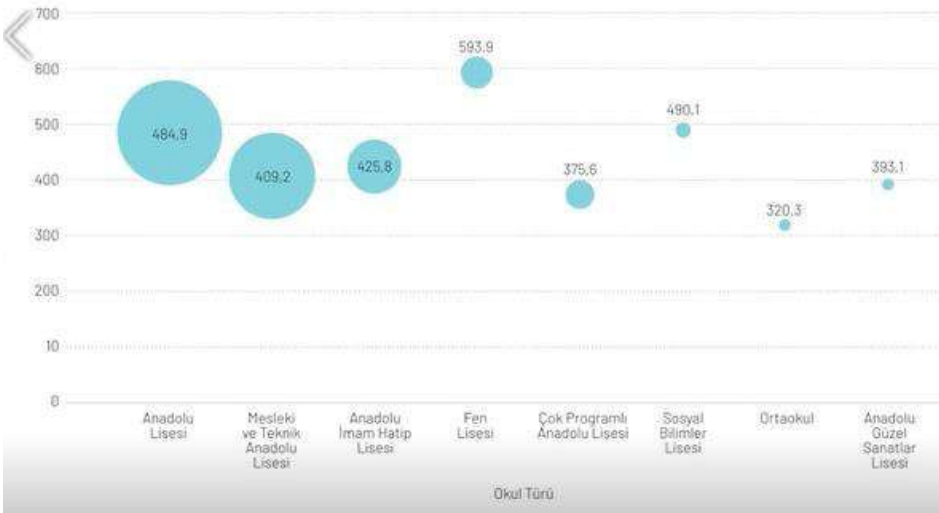


Şekil 3. Matematik okuryazarlığı ortalama puanlarındaki ve yeterlik düzeylerindeki değişim

Matematik Okuryazarlığı

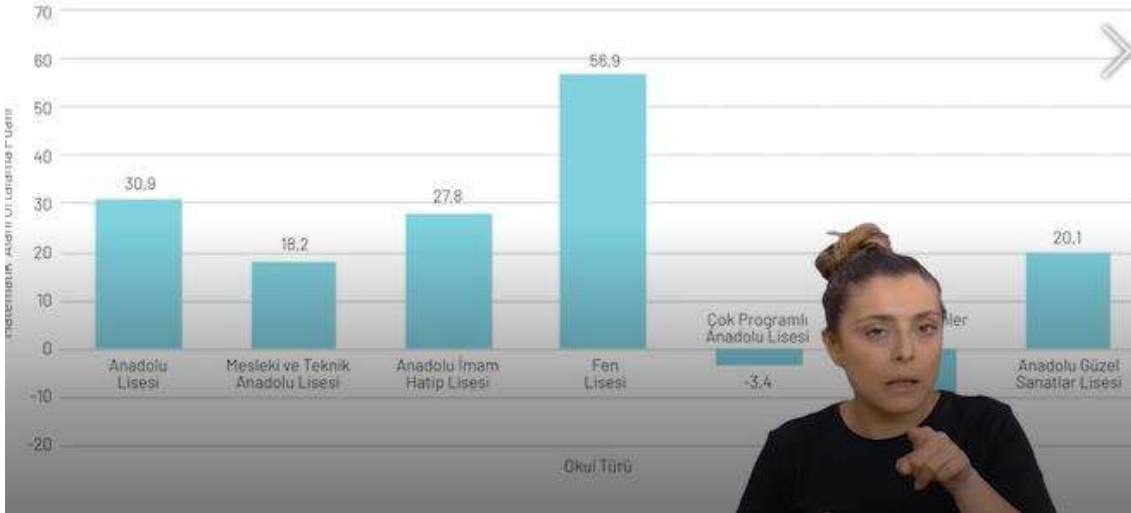
Okul Türlerine Göre Puan Ortalamaları

PISA 2018 MATEMATİK PERFORMANSLARININ OKUL TÜRÜNE GÖRE DEĞİŞİMİ***



Önceki Döngü ile Performans Farkı

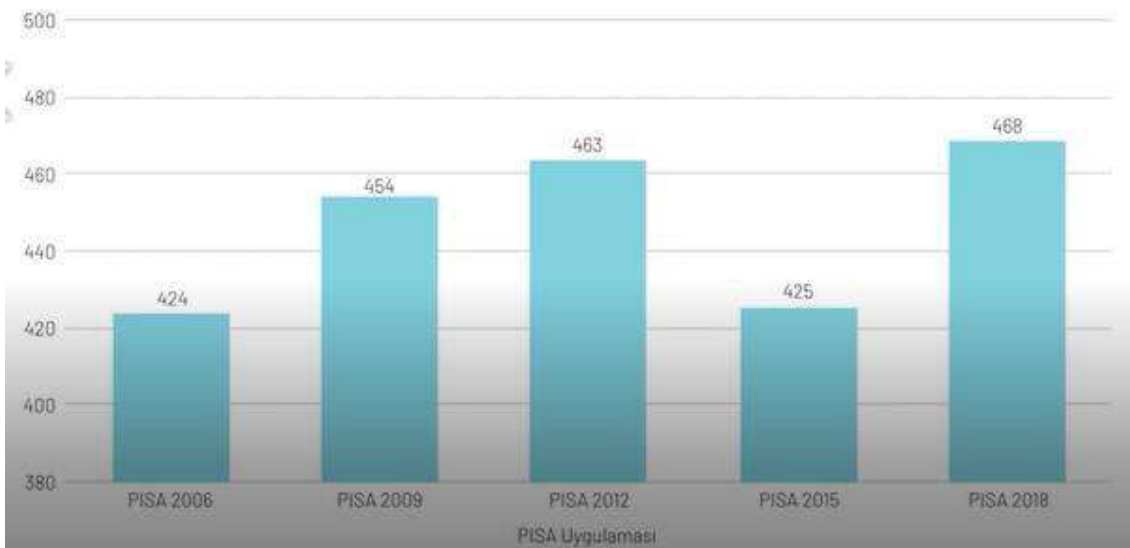
PISA 2018 İLE PISA 2015 UYGULAMALARI ARASINDA OKUL TÜRLERİNE GÖRE MATEMATİK ORTALAMA PUAN DEĞİŞİMLERİ



Fen Okuryazarlığı

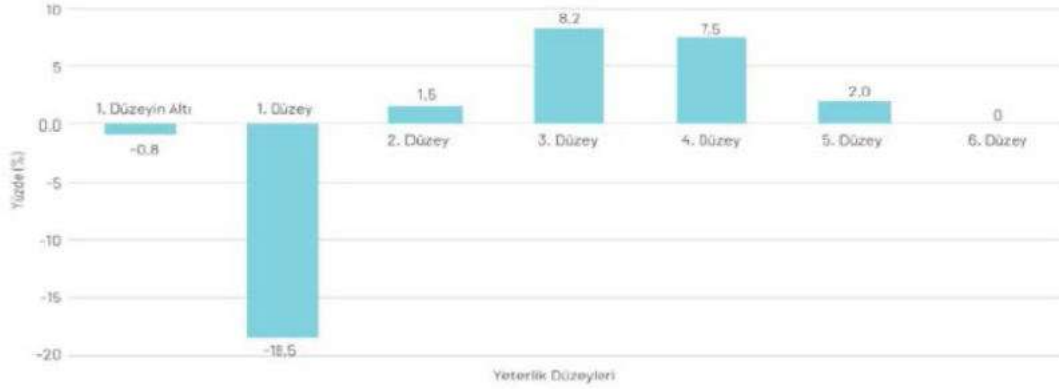
Ortalama Puanlardaki Değişim

PISA 2006 İLE PISA 2018 UYGULAMALARI ARASINDA TÜRKİYE'NİN FEN PERFORMANSINDAKİ DEĞİŞİM



Yeterlik Düzeylerindeki Değişim

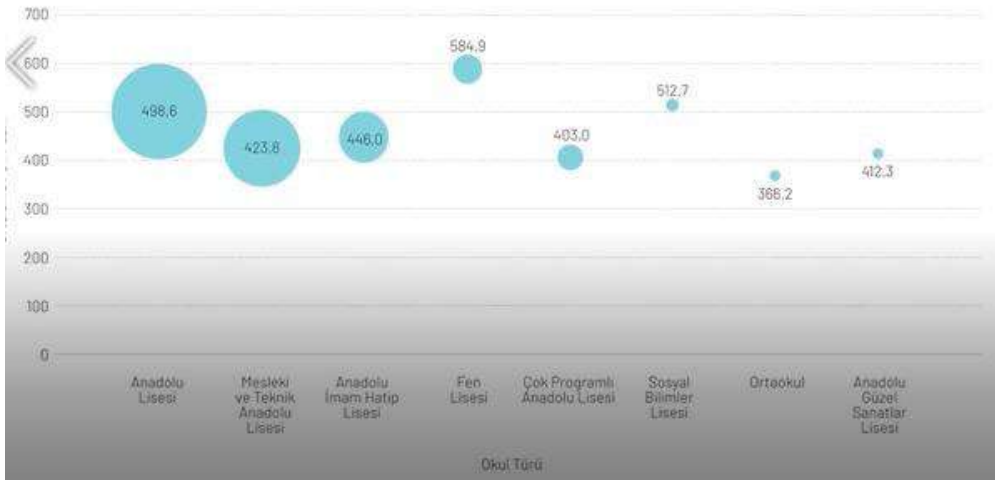
PISA 2018 İLE PISA 2015 UYGULAMALARI ARASINDA TÜRKİYE'DE FEN ALANI YETERLİK DÜZEYLERİNDEKİ ÖĞRENCİ ORANLARININ DEĞİŞİMİ



Fen Okuryazarlığı

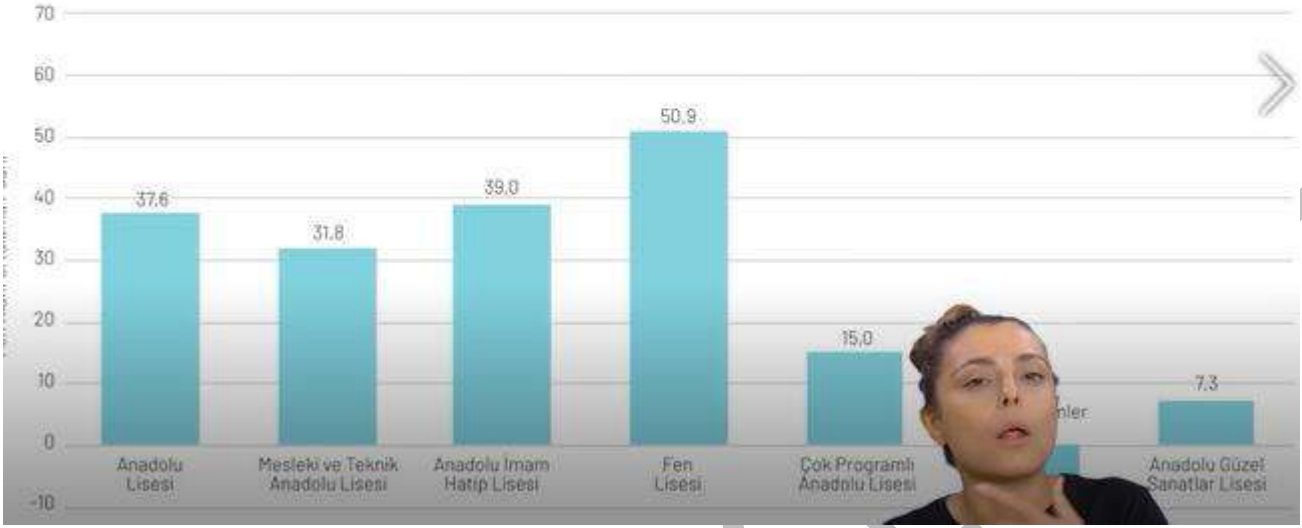
Okul Türlerine Göre Puan Ortalamaları

PISA 2018 FEN PERFORMANSLARININ OKUL TÜRÜNE GÖRE DEĞİŞİMİ***



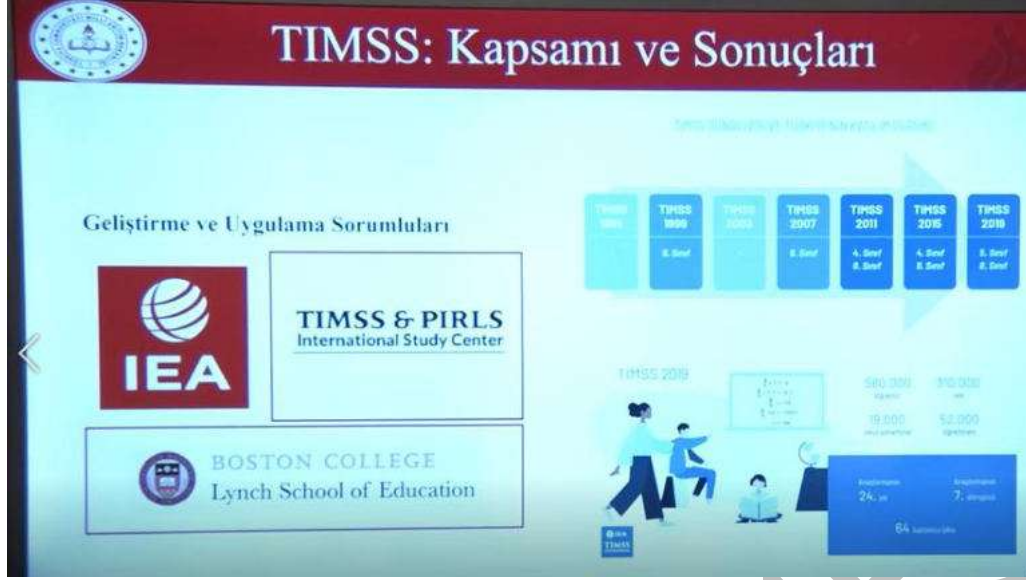
Önceki Döngü ile Performans Farkı

PISA 2018 İLE PISA 2015 UYGULAMALARI ARASINDA OKUL TÜRLERİNE GÖRE FEN ORTALAMA PUAN DEĞİŞİMLERİ



19-TIMSS: KAPSAMI VE SONUÇLARI (Sayfa: 112-115)

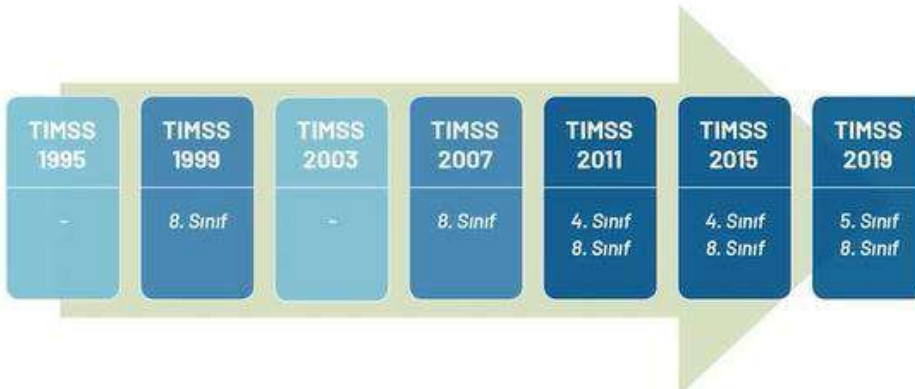
EGİTMEN: DR. HAYRİ EREN SUNA-BAKANLIK MÜŞAVİRİ



Geliştirme ve Uygulama Sorumluları



TIMSS DÖNGÜLERİ VE TÜRKİYE'NİN KATILIM DURUMU

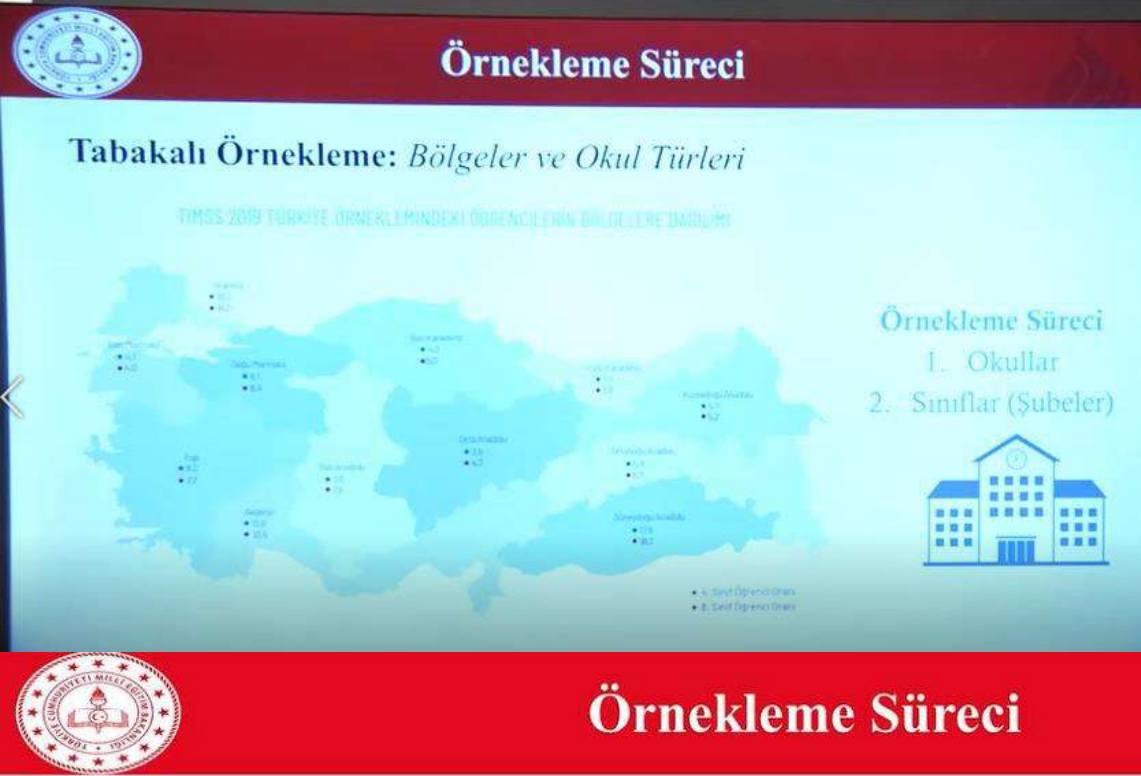




Bir TIMSS Döngüsünün Hikayesi



Şekil 5. Bir TIMSS döngüsünün hikâyesi

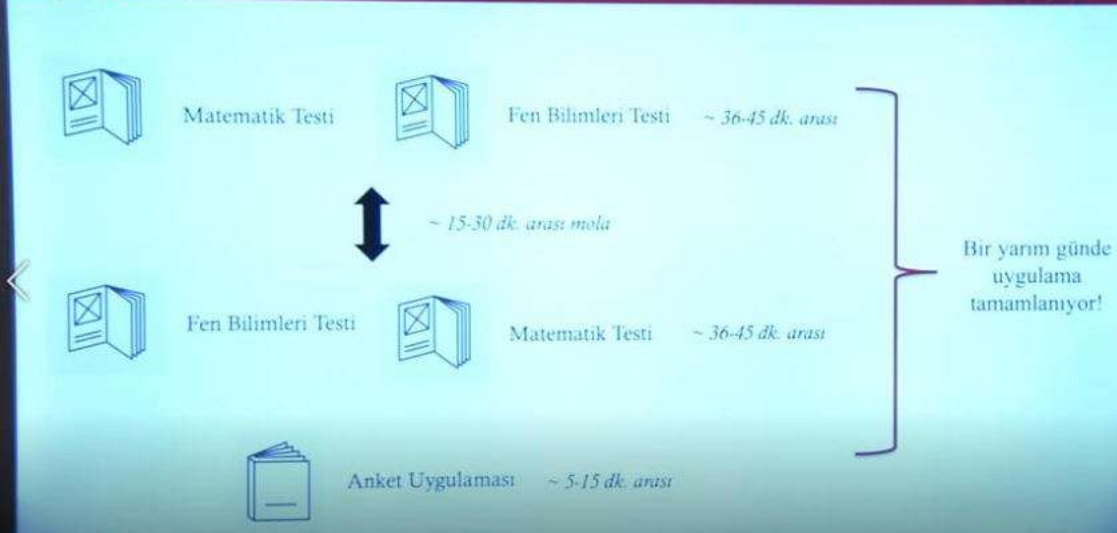


Tabakalı Örnekleme: Bölgeler ve Okul Türleri

TIMSS 2019 TÜRKİYE ÖRNEKLEMİNDEKİ ÖĞRENCİLERİN BÖLGELERE DAĞILIMI



Uygulama Süreci



2019 Döngüsü - Öne Çıkan Sonuçlar

5. sınıf düzeyi örneklemi ile katılım

İlk kez 2019 yılında Türkiye, TIMSS döngüsünün 4. sınıf uygulamasına 5. sınıf örneklemi ile katıldı.

2017 yılında alınan kararla;

1. Yaş
2. Değerlendirme çerçevesi uyumu

ölçütleri dikkate alınmıştır.

2019 Döngüsü - Öne Çıkan Sonuçlar

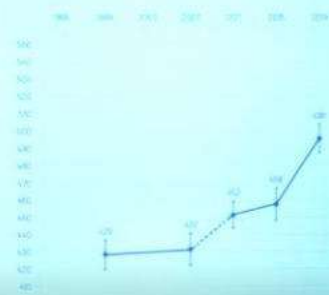
4. Sınıf Düzeyi

TÜRKİYE'NİN TIMSS DÖNGÜLERİNDEKİ
4. SINIF MATEMATİK BAĞARI DEĞİŞİMLERİ*



8. Sınıf Düzeyi

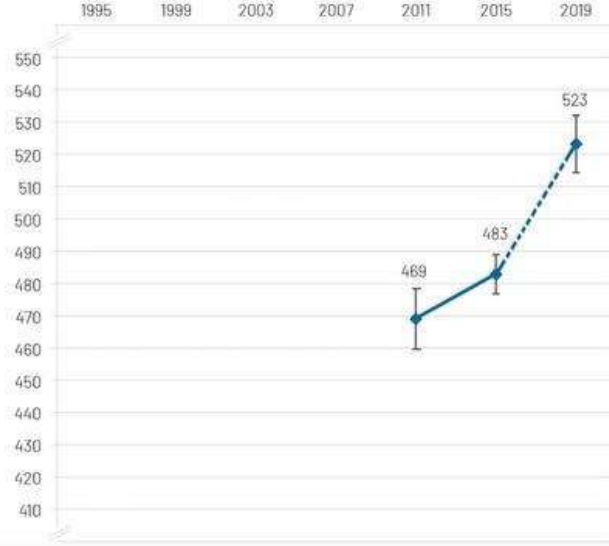
TÜRKİYE'NİN TIMSS DÖNGÜLERİNDEKİ
8. SINIF MATEMATİK BAĞARI DEĞİŞİMLERİ*



Matematik

4. Sınıf Düzeyi

TÜRKİYE'NİN TIMSS DÖNGÜLERİNDEKİ 4. SINIF MATEMATİK BAŞARISI DEĞİŞİMLERİ*

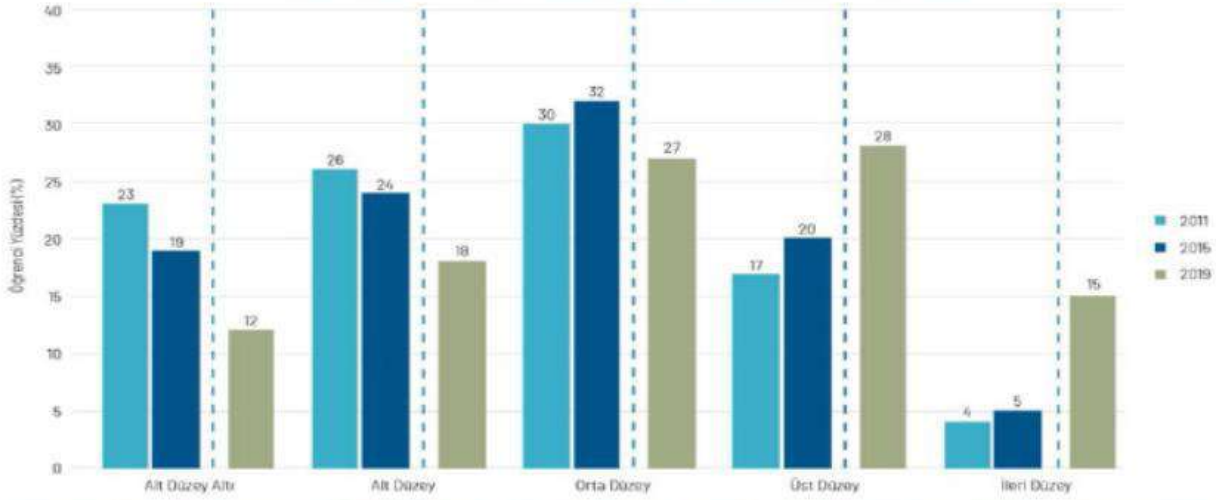


Matematik

*Türkiye'nin katıldığı TIMSS döngülerindeki ortalama matematik ölçek puanları verilmiştir. TIMSS 2019 döngüsüne Türkiye ilk kez beşinci sınıf düzeyinde katıldığı için Türkiye'nin 2019 ve öncesindeki döngülerindeki karşılaştırmalı performansı uluslararası raporda yer almamıştır.

4. Sınıf Düzeyi

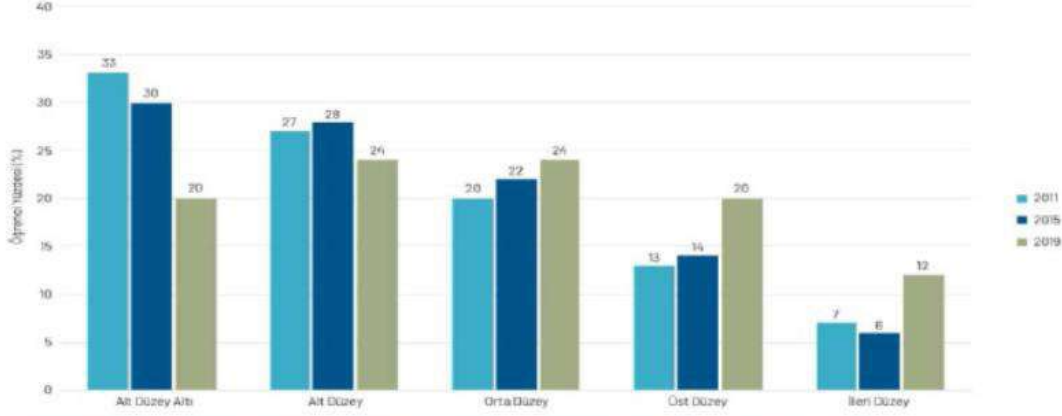
TÜRKİYE'DEKİ ÖĞRENCİLERİN SON TIMSS DÖNGÜLERİNDE 4. SINIF MATEMATİK YETERLİK DÜZEYLERİNDEKİ ORANLARI*



*Türkiye'nin katıldığı TIMSS döngülerindeki matematik yeterlik düzeylerine ulaşan öğrenci oranları. TIMSS 2019 döngüsüne Türkiye ilk kez beşinci sınıf düzeyinde katıldığı için Türkiye'nin 2019 ve öncesindeki döngülerindeki karşılaştırmalı performansı uluslararası raporda yer almamıştır.

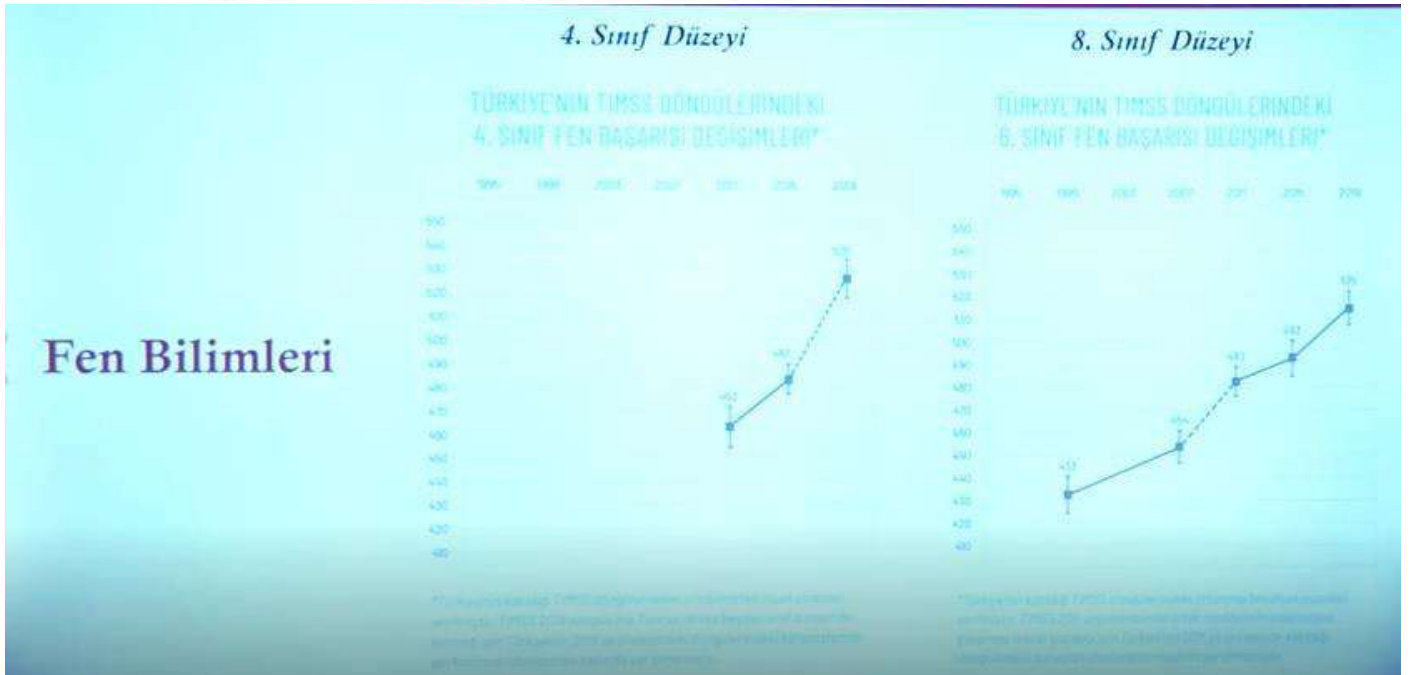
8. Sınıf Düzeyi

TÜRKİYE'DEKİ ÖĞRENCİLERİN SON TIMSS DÖNGÜLERİNDE 8. SINIF MATEMATİK YETERLİK DÜZEYLERİNDEKİ ORANLARI*



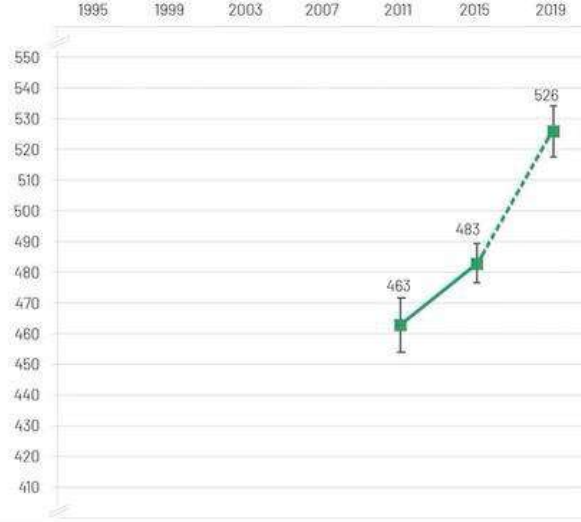
*Türkiye'nin katıldığı TIMSS döngülerindeki matematik yeterlik düzeylerine ulaşan öğrenci oranları, TIMSS 2011 uygulamasında ortak maddelerin adaptasyon çalışması tekrar yapıldığı için Türkiye'nin 2011 yılı öncesinde katıldığı döngülerdeki sonuçları uluslararası raporda yer almamıştır.

Şekil 6. Matematik yeterlik düzeyleri oranları



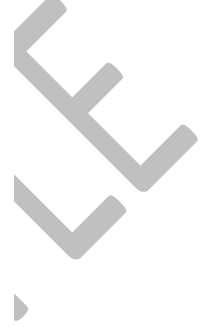
4. Sınıf Düzeyi

TÜRKİYE'NİN TIMSS DÖNGÜLERİNDEKİ 4. SINIF FEN BAŞARISI DEĞİŞİMLERİ*



*Türkiye'nin katıldığı TIMSS döngülerindeki ortalama fen ölçek puanları verilmiştir. TIMSS 2019 döngüsüne Türkiye ilk kez beşinci sınıf düzeyinde katıldığı için Türkiye'nin 2019 ve öncesindeki döngülerindeki karşılaştırmalı performansı uluslararası raporda yer almamıştır.

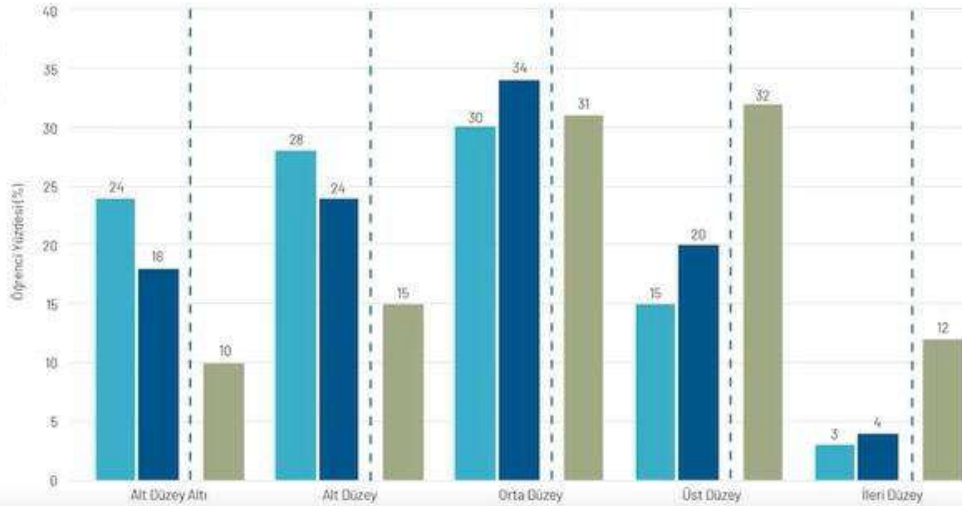
Fen Bilimleri



Fen Bilimleri

4. Sınıf Düzeyi

TÜRKİYE'DEKİ ÖĞRENCİLERİN SON TIMSS DÖNGÜLERİNDE 4. SINIF FEN YETERLİK DÜZEYLERİNDEKİ ORANLARI*



*Türkiye'nin katıldığı TIMSS döngülerindeki fen yeterlik düzeylerine ulaşan öğrenci oranları. TIMSS 2019 döngüsüne Türkiye ilk kez beşinci sınıf düzeyinde katıldığı için Türkiye'nin 2019 ve öncesindeki döngülerindeki karşılaştırmalı performansı uluslararası raporda yer almamıştır.

20- OECD SOSYAL DUYGUSAL BECERİLER İZLEME ARAŞTIRMASI (Sayfa: 115-120)

EĞİTMEN: DR. HAYRİ EREN SUNA-BAKANLIK MÜŞAVİRİ

OECD Sosyal ve Duygusal Beceriler Araştırması

> **Neden önemli?**

OECD tarafından gerçekleştirilen ve sosyal ve duygusal becerileri odağına alan ilk geniş ölçekli çalışma

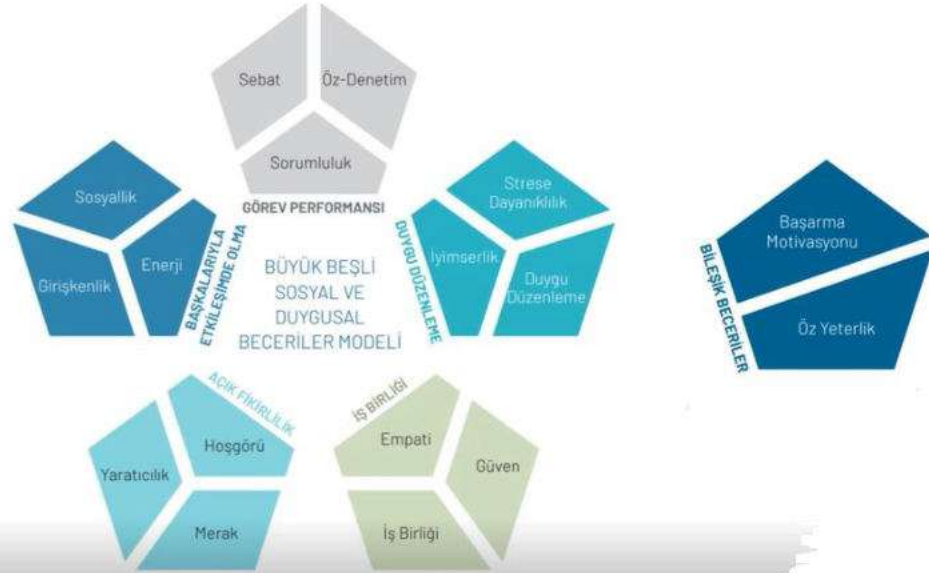
> **Hangi yenilikleri getiriyor?**

10 ve 15 yaş olmak üzere iki ayrı yaş grubunu hedefe alıyor.

5 faktör modelini (Big-five model) kullanıyor.

Veri çeşitlemesini (data triangulation) kullanıyor.

5 Faktörlü Model



Farklı kültürlerden katılımcılar

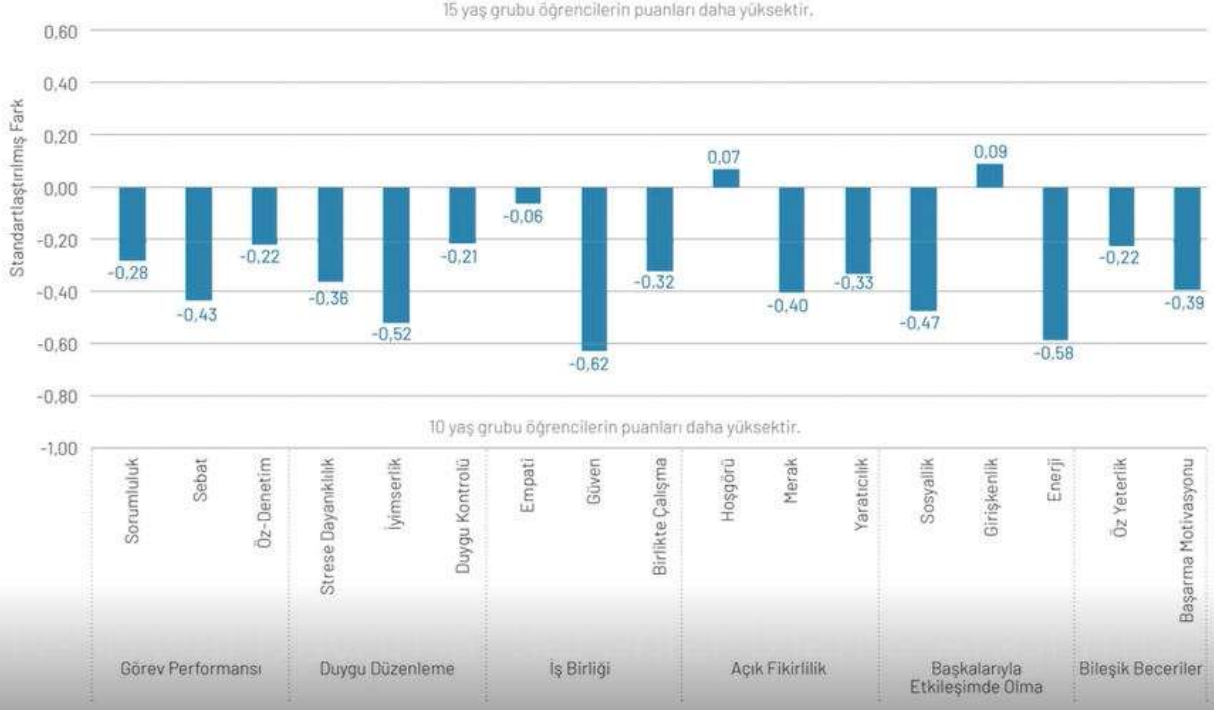


İstanbul Örnekleminin Özellikleri

	10 Yaş Grubu			15 Yaş Grubu		
	Öğrenci	Veli	Öğretmen	Öğrenci	Veli	Öğretmen
Geçerli Katılımcı	2.701	1.974	2.841	3.168	2.033	3.373
Toplam Katılımcı	2.796	2.095	2.903	3.184	2.155	3.416
Geçerli Katılımcı Oranı	96,60	94,22	97,86	99,50	94,34	98,74

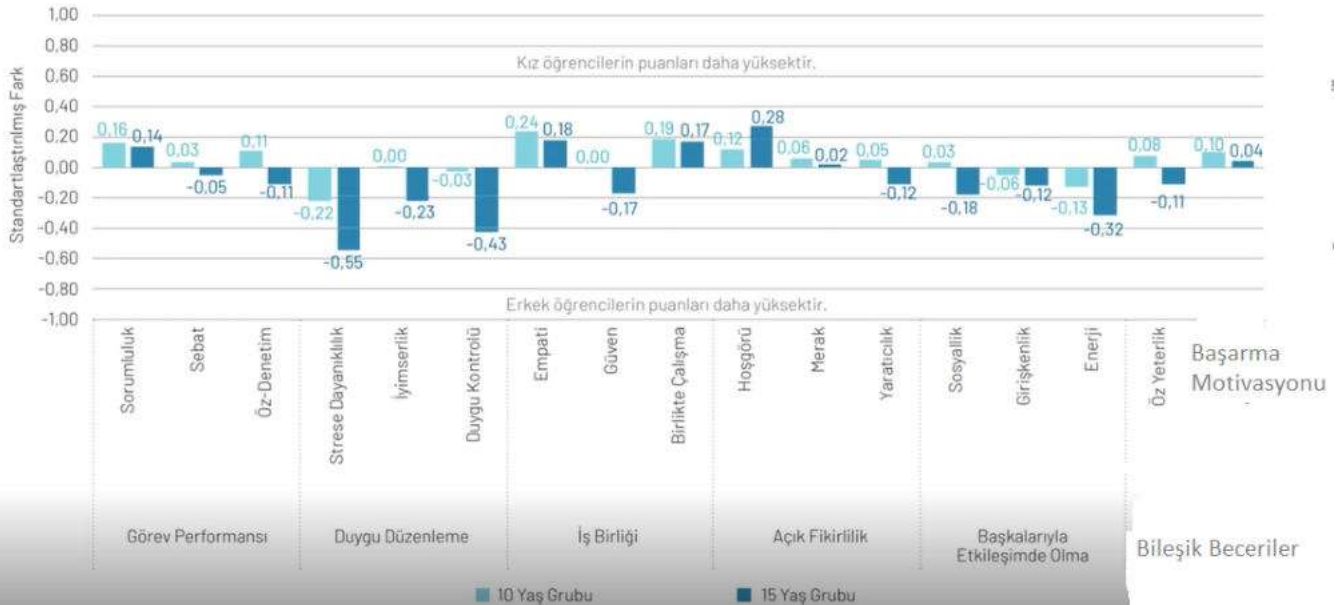
10 Yaş Grubu			15 Yaş Grubu		
Okul Türü	Okul Sayısı	Öğrenci Oranı (%)	Okul Türü	Okul Sayısı	Öğrenci Oranı (%)
Ortaokul	59	80,08	Anadolu Lisesi	38	48,84
İmam Hatip Ortaokulu	9	13,11	Mesleki ve Teknik Anadolu Lisesi	30	36,02
İlkokul	27	6,81	Anadolu İmam Hatip Lisesi	9	11,79
Toplam	95	100,00	Fen Lisesi	2	1,89
			Sosyal Bilimler Lisesi	1	1,48
			Toplam	80	100

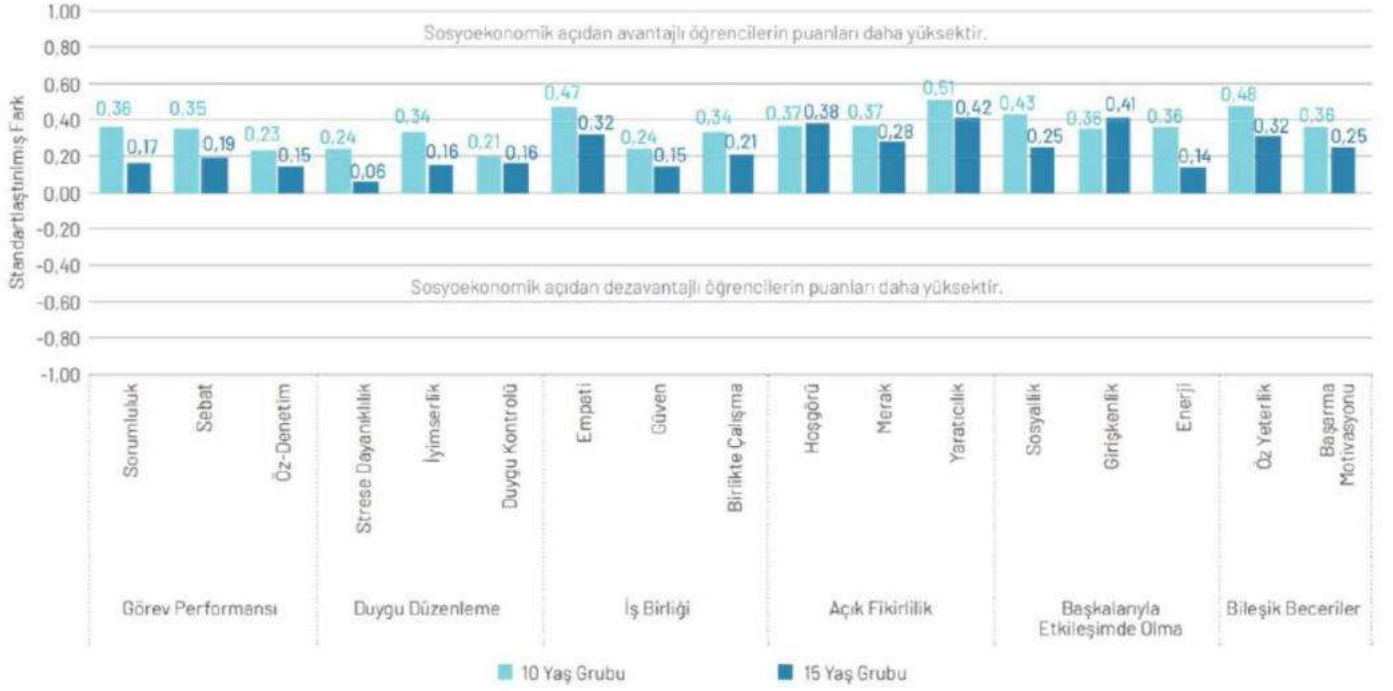
Yaş Düzeylerine Göre Sosyal ve Duygusal Beceriler



Cinsiyete Göre Sosyal ve Duygusal Beceriler

SOSYAL VE DUYGUSAL BECERİLERDE KIZ VE ERKEK ÖĞRENCİLER ARASINDAKİ FARK*

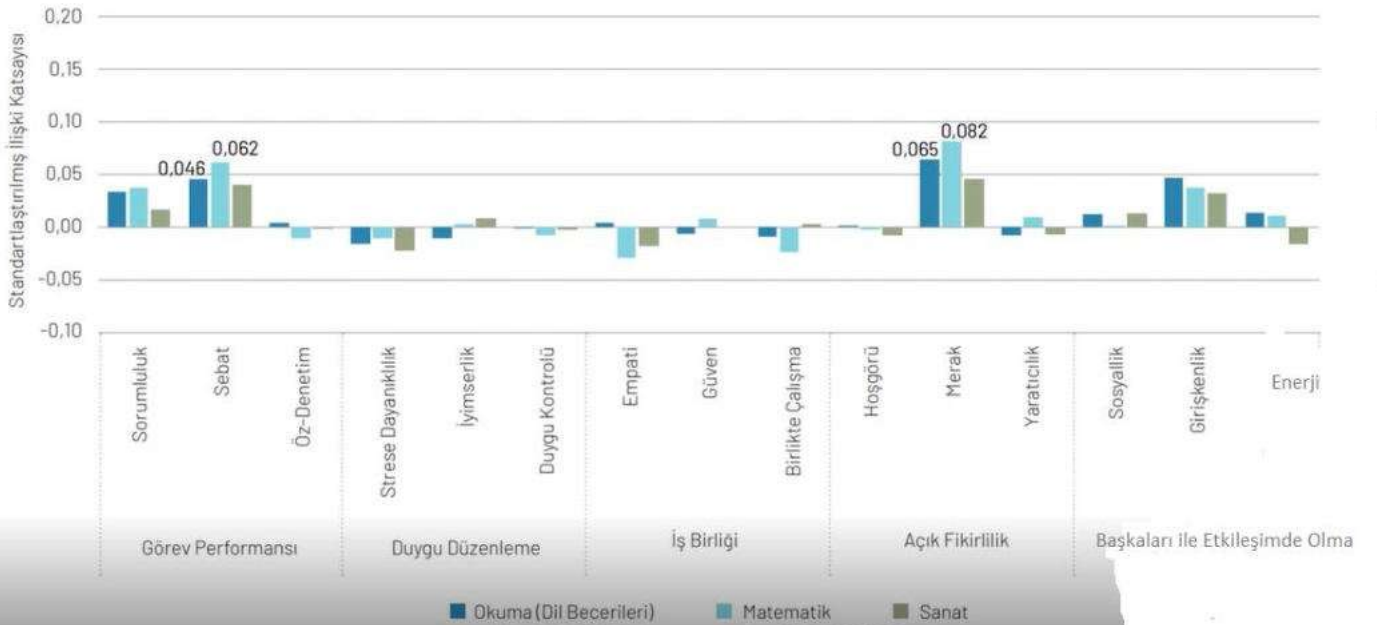




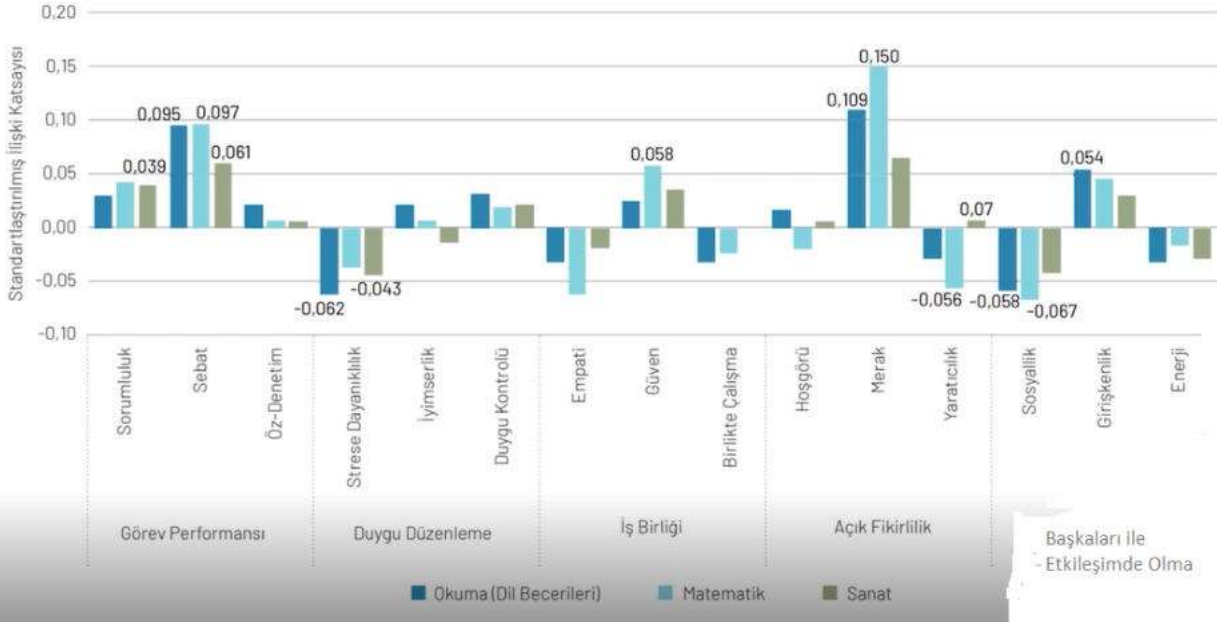
Şekil 10. Sosyoekonomik düzeye göre sosyal ve duygusal beceriler

Windows'u Et

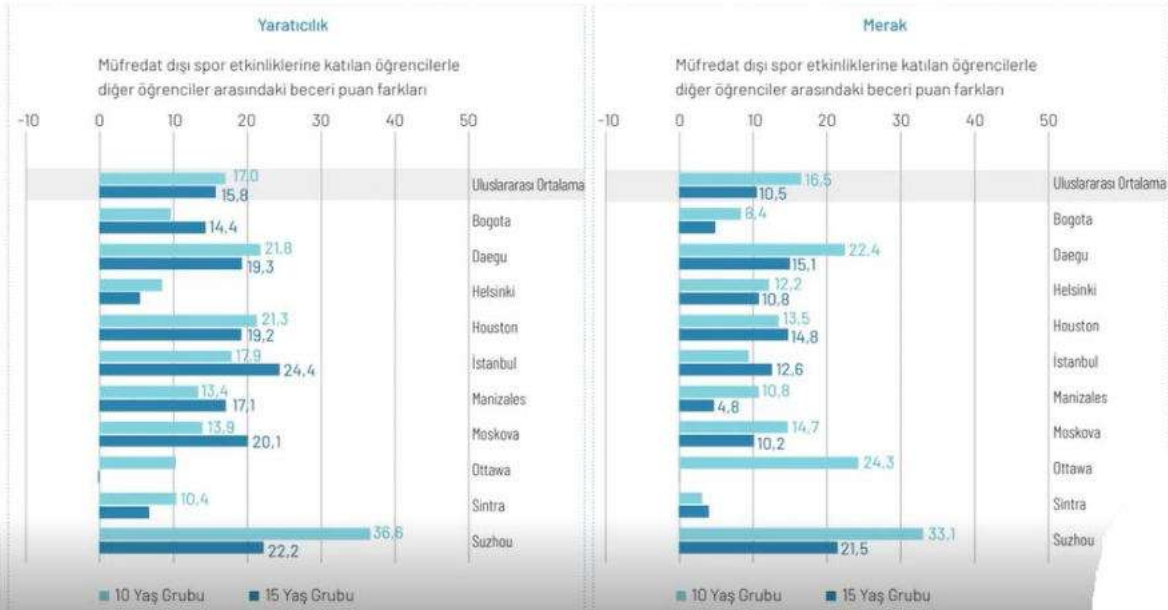
Akademik Başarı ile İlişkisi – 10 Yaş Grubu



Akademik Başarı ile İlişkisi – 15 Yaş Grubu



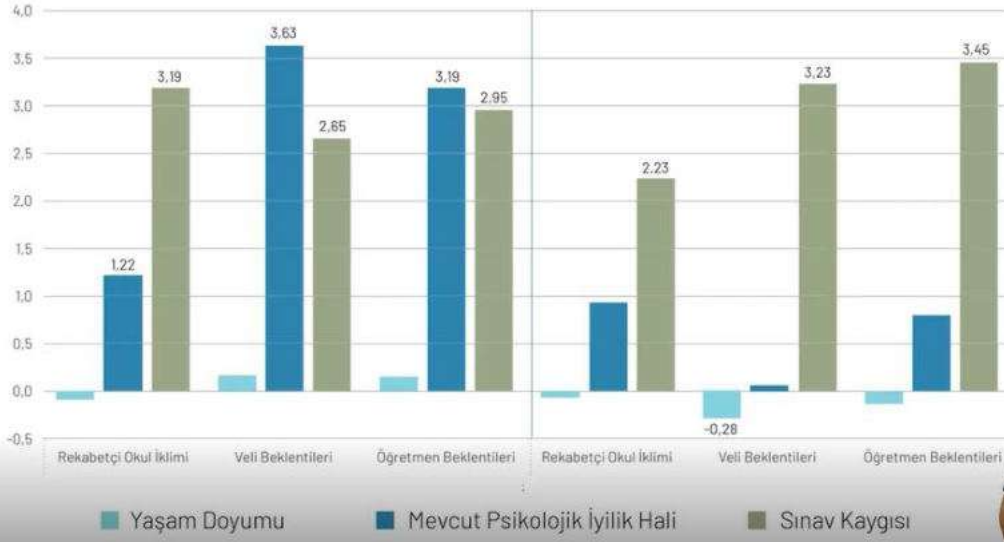
Merak ve Yaratıcılık Becerilerini Geliştirmek için Öneriler – Spor Etkinlikleri



Okul İklimi ile Sosyal ve Duygusal Beceriler Arasındaki İlişki

10 Yaş Grubu

15 Yaş Grubu



Yaşam Doyumu – 10 Yaş Grubu

10 Yaş Grubu

Şehir	Yaşam Doyumu Ortalaması	Öğrenci Oranı			
		Düşük Doyum/ Doyumsuzluk (%)	Orta Doyum (%)	Doyum (%)	Yüksek Doyum (%)
Bogota	8,60	5,31	9,10	15,52	70,07
Daegu	8,10	6,27	13,24	26,48	54,01
Helsinki	8,60	2,81	6,57	25,83	64,79
Houston	7,60	12,67	15,28	22,32	49,73
Istanbul	8,10	7,86	12,88	22,91	56,35
Manizales	8,70	6,25	7,53	12,74	
Moscow	8,30	5,21	11,93	23,23	
Ottawa	7,90	6,90	13,87	31,43	
Sintra	8,50	4,62	10,26	20,59	
Suzhou	8,20	5,76	14,08	25,05	
Uluslararası Ortalama	8,23	6,56	11,61	22,83	

Okul içi İlişkiler

Şehir	10 Yaş Grubu			15 Yaş Grubu		
	Okula Aidiyet - Zorbalığa Maruz Kalma	Okula Aidiyet - Öğrenci ve Öğretmen İlişkileri	Zorbalığa Maruz Kalma - Öğrenci ve Öğretmen İlişkileri	Okula Aidiyet - Zorbalığa Maruz Kalma	Okula Aidiyet - Öğrenci ve Öğretmen İlişkileri	Zorbalığa Maruz Kalma - Öğrenci ve Öğretmen İlişkileri
Bogota	-0,37	0,31	-0,11	-0,23	0,22	-0,06
Daegu	-0,24	0,27	-0,08	-0,21	0,26	-0,07
Helsinki	-0,38	0,32	-0,20	-0,26	0,25	-0,12
Houston	-0,37	0,31	-0,16	-0,27	0,29	-0,10
İstanbul	-0,39	0,35	-0,19	-0,27	0,30	-0,08
Manizales	-0,39	0,34	-0,16	-0,25	0,25	
Moskova	-0,34	0,34	-0,14	-0,29	0,26	
Ottawa	-0,35	0,26	-0,16	-0,22	0,21	
Sintra	-0,33	0,19	-0,10	-0,23	0,13	
Suzhou	-0,40	0,39	-0,18	-0,28	0,33	
Uluslararası Ortalama	-0,36	0,32	-0,15	-0,25	0,26	



Ulusal İzleme Araştırmaları

AKADEMİK BECERİLERİN İZLENMESİ VE DEĞERLENDİRİLMESİ

ABIDE

Akademik Becerilerin İzlenmesi ve Değerlendirilmesi

Üst düzey düşünme becerilerinin ölçülmesi ve beceri temelli yapı kurgusu

T.C. MİLLÎ EĞİTİM BAKANLIĞI

İzleme Araştırması Sonuçları Sayı: 9

Türkçe-Matematik-Fen Bilimleri Öğrenci Başarı İzleme Araştırması (TMF-ÖBA)-I: 2019

4. Sınıf Seviyesi

Kazanım odaklı ve geribildirime dayalı kurgu

T.C. MİLLÎ EĞİTİM BAKANLIĞI

İzleme Araştırması Sonuçları Sayı: 9

Dört Beceride Türkçe Dil Sınavı: İlk Ölçüm Sonuçları

Okuma, dinleme, yazma ve konuşma becerilerini anadilde ölçen ilk izleme araştırması

AKADEMİK BECERİLERİN İZLENMESİ VE DEĞERLENDİRİLMESİ

ABIDE

Akademik Becerilerin İzlenmesi ve Değerlendirilmesi

Üst düzey düşünme becerilerinin ölçülmesi ve beceri temelli yapı kurgusu

T.C. MİLLÎ EĞİTİM BAKANLIĞI

İzleme Araştırması Sonuçları Sayı: 9

Türkçe-Matematik-Fen Bilimleri Öğrenci Başarı İzleme Araştırması (TMF-ÖBA)-I: 2019

4. Sınıf Seviyesi

Kazanım odaklı ve geribildirime dayalı kurgu