

ÖLÇME VE DEĞERLENDİRME MODÜL 2

 **@gunes_ogretmenle_etkinlikler**

ÖLÇME VE DEĞERLENDİRMENİN TEMEL KAVRAMLAR

Sistem, en az bir hedefi gerçekleştirmek üzere uygun ve değişik öğelerden oluşan bir örüntüdür². Bir bütüne sistem diyebilmek için **girdi**, **süreç** ve **çıkış** öğelerinin olması gerekir.

Eğitim de bir sistemdir çünkü gerçekleştirmek istediği hedefler bulunmaktadır. Hedefler şu biçimde tanımlanabilir:

1. Uzak hedefler: Devletlerin hedefleridir, bazı ülkelerde çerçevesi anayasa ile güvenceye alınmıştır.

2. Genel hedefler: Daha işevuruk hedeflerdir. Kurumların (MEB, MEB genel müdürlükleri, il millî eğitim müdürlükleri, okul yönetimleri vb.) hedefleri vb.

3. Özel hedefler: Dersin hedefleridir, bir dersin kazanımlarını kapsar.

Eğitimin bir sistem olmasının ikinci nedeni, bu hedefleri gerçekleştirmek için farklı öğelerden oluşmasıdır:

1. Girdi: Eğitim-öğretim için gerekli her şey eğitim sisteminin girdisini oluşturur. Örneğin öğrenci, öğretmen, okul, çalışanlar, yönetim, eğitim programı vb

2. Süreç: Girdilerin hedefler doğrultusunda bir araya getirildiği, dersin hedeflerinin öğrenciye kazandırıldığı eğitim durumlarını kapsar.

3. Çıkış: Öğrencilerin sahip olduğu niteliklerdir. Eğitim sistemi için çıktılar, dar anlamda öğrencinin bir dersin sonunda ulaştığı hedefler ve kazanımları ifade ederken geniş anlamda öğrencinin öğretim sürecinin başlangıcından sonuna kadar sahip olduğu bilgi beceri ve yetkinlik anlarını ifade etmektedir. Çıktılar hedeflerle tutarlı mı diye kontrol edilmektedir. Çıktılar hedeflerle tutarlı ise eğitim sistemi doğru bir şekilde işliyor demektir.. Eğer çıktılar hedeflerle tutarlı değilse eğitim sisteminde bir şeyler aksıyor demektir. **Eğitimde kontrol (denetleme) ölçme ve değerlendirme ile mümkündür.**

ÖLÇME, ÖLÇÜT VE DEĞERLENDİRMENİN KAVRAMSAL TEMELLERİ

1-Ölçme: Öğrencinin belirli bir özelliğini³ gözleyerek o özelliğe sayı, sembol ya da sıfat/kategori adı verme işlemidir.

Eğitimde ölçme dediğimizde aklımıza ilk gelen genellikle sınavlardır. Eğitimde ölçme yalnızca sınavlardan ibaret değildir. Eğitimde akademik başarı dışında yetenek, ilgi, tutum, özel gereksinim gereken alanlar vb. özellikler de ölçmeye konu olur.

Ölçülmeyen, kontrol edilmeyen süreçler denetlenemez ve denetlenemeyen süreçler ise değerlendirilemez

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

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

Ölçme ⇒ Ölçme sonucu (ölçüm), b. Ölçüt ve c. Karardır. Örneğin• Ayşe burs almaya hak kazanmıştır / burs alamamıştır.

YİRMİ BİRİNCİ YÜZYIL DEĞİŞEN EĞİTİM PARADİGMASI:

20. yüzyıla hâkim olan kültürel iklim, sanayileşme idi ve dolasıyla 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 yıllarda 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 donanıma sahip olma” olarak tanımlandı.

Dünya Ekonomik Forumuna göre 2025 yılında çalışanların sahip olması gereken 10 temel beceri:

- Analitik düşünme ve yenilik/yaratıcılık
- Etkin öğrenme ve öğrenme stratejileri
- Karmaşık problem çözme
- Eleştirel düşünme ve analiz
- Yaratıcılık, orijinallik
- Liderlik ve sosyal etki
- Teknoloji kullanımı



ÖLÇME TÜRLER



1. Doğrudan Ölçme (Temel Ölçme): Ölçmeye konu olan özelliğin doğrudan gözlenerek ölçümün elde edildiği ölçme türüdür.. Örneğin bir sınıftaki öğrenci sayısı, parktaki salıncak sayısı vb. özellikler doğrudan gözlenebilir.

2. Dolaylı Ölçme (Göstergeyle Ölçme): Bazı özellikler doğrudan gözlenemez. Bu özellikler ancak onun göstergesi olduğu bilinen ya da kabul edilen davranışlar aracılığı ile gözlenerek ölçülebilir. Örneğin öğrencilerin dört işlem becerisi, derse yönelik tutumları vb. özellikler ancak dolaylı olarak 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),

ÖLÇÜT TÜRLERİ

1. Mutlak Ölçüt (Kriter Referanslı Ölçüt): Eğer ölçüt ölçme işleminin yapıldığı grubun özelliklerinden bağımsız olarak belirleniyorsa “**mutlak**”tır. Mutlak ölçüt genellikle ölçme işlemi öncesinde ilan edilir. Örneğin dersten geçme notu 50’dir

Mutlak ölçütün temel alındığı ölçme işlemlerinde bir öğrencinin notu, diğer bir öğrencinin değerlendirmesini (geçme / kalma) etkilemez.



2. Bağlı Ölçüt (Norm Referanslı Ölçüt): Eğer ölçüt ölçme işleminin yapıldığı grubun belirli bir özelliğine / normuna dayalı olarak belirleniyorsa “**bağlı**”dır. Bağlı ölçüt ancak ölçme işlemi sonrasında belirlenebilir. Örneğin aritmetik ortalamaya göre puan alan geçer, Bağlı ölçütün temel alındığı ölçme işlemlerinde bir öğrencinin başarısı, başka bir öğrencinin değerlendirmesini (geçme / kalma) etkileyebilir. Bu anlamda öğrenciler açısından daha rekabetçi bir eğitsel ortam sağlar

• Selim fen lisesine yerleştirilmiştir. → Bağlı değerlendirme
• Güzin lisans programları için tercih yapabilecektir. → Mutlak değerlendirme

AMACA GÖRE DEĞERLENDİRME TÜRLERİ:

Amaca göre değerlendirme tanıma ve yerleştirme, biçimlendirme ve yetiştirme, değer biçme, rehberlik amaçlı ve program değerlendirme olarak beşe ayrılır.

1. Tanıma-Yerleştirmeye Yönelik Değerlendirme (Diyagnostik Değerlendirme):

Sınıf içi ölçme ve değerlendirme: Amaç öğrencinin ders bağlamındaki hazırbulunuşluluk düzeyini belirlemektir. not vermek değildir. Bu amaçla uygulanan ölçme araçlarına da “hazırbulunuşluluk testi” adı verilir.

Okul / kurum geneli ölçme ve değerlendirme: Yine öğretim sürecinin başında, öğrenciyi tanımak ve onu uygun olan programa / kura / sınıfa yerleştirmek amacıyla yapılır. Amaçlanan, not vermek değildir. Hedeflenen, öğrencinin niteliklerine uygun grupların oluşturulmasıdır. Zaman, para ve emek tasarrufu sağlar.

2. Biçimlendirme-Yetiştirmeye Yönelik Değerlendirme (Formatif Değerlendirme):

Belirli bir konu / ünite vb. sonunda öğrencilerin öğretime konu olan davranışların ne kadarını kazandığını, diğer bir deyişle üniteadaki öğrenme eksiklerini belirlemek amacıyla yapılır. Bu değerlendirmede kullanılan testlere alanda “izleme testi” ya da “tarama testi” denir. Not verilmeden yapılan kısa sınavlar (quiz) örnek olarak verilebilir

Değer Biçmeye / Düzey Belirlemeye Yönelik Değerlendirme (Summatif Değerlendirme):

Belirli bir öğretim sürecinin sonunda, not vermek amacıyla yapılan değerlendirmedir. Değer biçmeye yönelik değerlendirme, öğrencinin belirli bir öğretim süreci sonundaki durum tespiti Bu amaçla uygulanan testlere “alanda erişim testi” denir. Bitirme sınavları, eğitim-öğretim kurumlarında not vermek amacıyla uygulanan yazılı yoklamalar örnek olarak verilebilir.

İlk üç değerlendirme türü öğrenci başarısına odaklıdır, diğer bir ifadeyle öğrencinin akademik başarısının ölçülüp değerlendirilmesi söz konusudur.

.4. Rehberlik Amaçlı Değerlendirme: Rehberlik amaçlı değerlendirme kabaca iki bağlamda ele alınabilir. Bunlar özel eğitim ve mesleki rehberlik hizmetleridir.

Özel eğitim: Belirli alanlarda özel gereksinimi olan çocuklara uygun eğitim ortamları ve/veya programı uygulamak amacıyla ölçme ve değerlendirme süreçleri işe koşulmaktadır

Mesleki rehberlik: Öğrencilerin alan seçmelerinde ve özellikle bir üst öğretim kurumuna yönlendirilmesinde çalışılır.

5. Program Değerlendirme: Eğitim programları bileşenlerinin yapısı ve niteliğine ilişkin olarak sistematik bilgi toplama ve değerlendirme sürecidir. Program değerlendirmede program taslağının hazırlanması, programın uygulanması ve programın etkisine ilişkin yargılar elde edilir.



BİR ÖLÇME ARACINDA BULUNMASI GEREKEN PSİKOMETRİK NİTELİKLER

Geçerlilik ölçmek istenilen özelliğin, başka özelliklerle karıştırılmadan, 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.

HATA : Kaynağı, yönü ve miktarı göz önünde tutularak sabit, sistematik ve tesadüfi olmak üzere üç türde incelenebilir

Sabit Hata: Miktarı ölçmeden ölçmeye değişmeyen, hatalardır. Bir öğretmen, sınavında herkese 10 puan fazla veriyorsa karışan hata sabit olacaktır.

.2. Sistematik Hata: Ölçülen büyüklüğe, öğretmene ya da ölçme koşullarına göre miktarı değişen hatalardır. Bir öğretmen, sınavında yazısı kötü olandan puan kırıyorsa yine karışan hata sistematik olacaktır. sistematik ve sabit hatalarda ölçme sonuçlarına karışan hata miktarı, yönü ve kaynağı bellidir. ölçme sonuçlarına karışmasını engellemek ya da bu tür hataları düzeltmek görece daha kolaydır.

.3. Tesadüfi (Rastlantısal) Hata: Şansla ortaya çıkan ne yönde ve ne ölçüde karıştığı genellikle bilinemeyen hatalardır.

Tesadüfi hataların genellikle dört kaynağı olduğu kabul edilir. Bunlar:

1. Ölçme işlemi yapan kişiden / öğretmenden kaynaklanan hata:

Ölçme işlemi sürecinde: Test katılımcısının dikkatini dağıtacak ve/veya kaygısını artıracak davranışlar vb.

Ölçme işlemi sonrasında: Puanlamadaki dikkat ve titizliğin zamandan zamana değişmesi, yorgunluk, vb.

2.Ölçme aracından kaynaklanan hata: Ölçme araçları hazırlanırken maddelerin iyi ifade edilmemesi, test katılımcılarının yanlış anlamalarına ve dolayısıyla hataya neden olacaktır.

Bireyden / öğrenciden kaynaklanan hata: Bireylerin ölçme işlemi sürecinde içinde bulundukları fiziksel, fizyolojik ve psikolojik

.4. Fiziksel ortamdan kaynaklanan hata: Ölçme işleminin gerçekleştiği fiziksel ortamda. sıcaklık, ışık, ses, koku, görsel uyarıcıların fazlalığı, ergonomi vb. etmenler hata miktarını artırabilir.durumlar test puanına etki eder.

GEÇERLİLİK, GÜVENİLİRLİK VE HATA İLİŞKİSİ:

Bir testin güvenilir olması onun geçerli olacağı anlamına gelmez ancak bir test geçerli ise büyük olasılıkla güvenilirlerdir.

Güvenilirlik, geçerlilik için bir ön şart ancak yeterli şart değildir

ölçme araçlarından elde edilen puanların geçerliliği ve güvenilirliği diye bir şey vardır..

Ölçme araçlarının psikometrik nitelikleri şu durumlara göre değişebilir:

- a. Ölçme amacının değişmesi,
- b. Uygulama grubunun değişmesi,
- c. Dilin eskimesi / değişmesi,
- d. Maddelerde ve/veya alt ölçeklerde yapılan değişiklikler,
- e. Farklı kültürler,
- f. Kuramsal bilgi birikiminde değişiklikler.

KORELASYON: Geçerlilik ve güvenilirliği belirlemeye yönelik çalışmalardır .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?

Korelasyon -1 ile 1 arasında değer alır

Pozitif korelasyon ,iki değişken arasında doğru orantılı ilişki anlamına gelir. Örneğin ders çalışma süresi ile sınav notu arasında pozitif bir korelasyonun elde edilmesi beklenir. **Negatif korelasyon**, iki değişken arasında ters orantılı ilişki anlamına gelir. Örneğin ders süresi ile dikkat arasında negatif bir korelasyonun elde edilmesi beklenir.

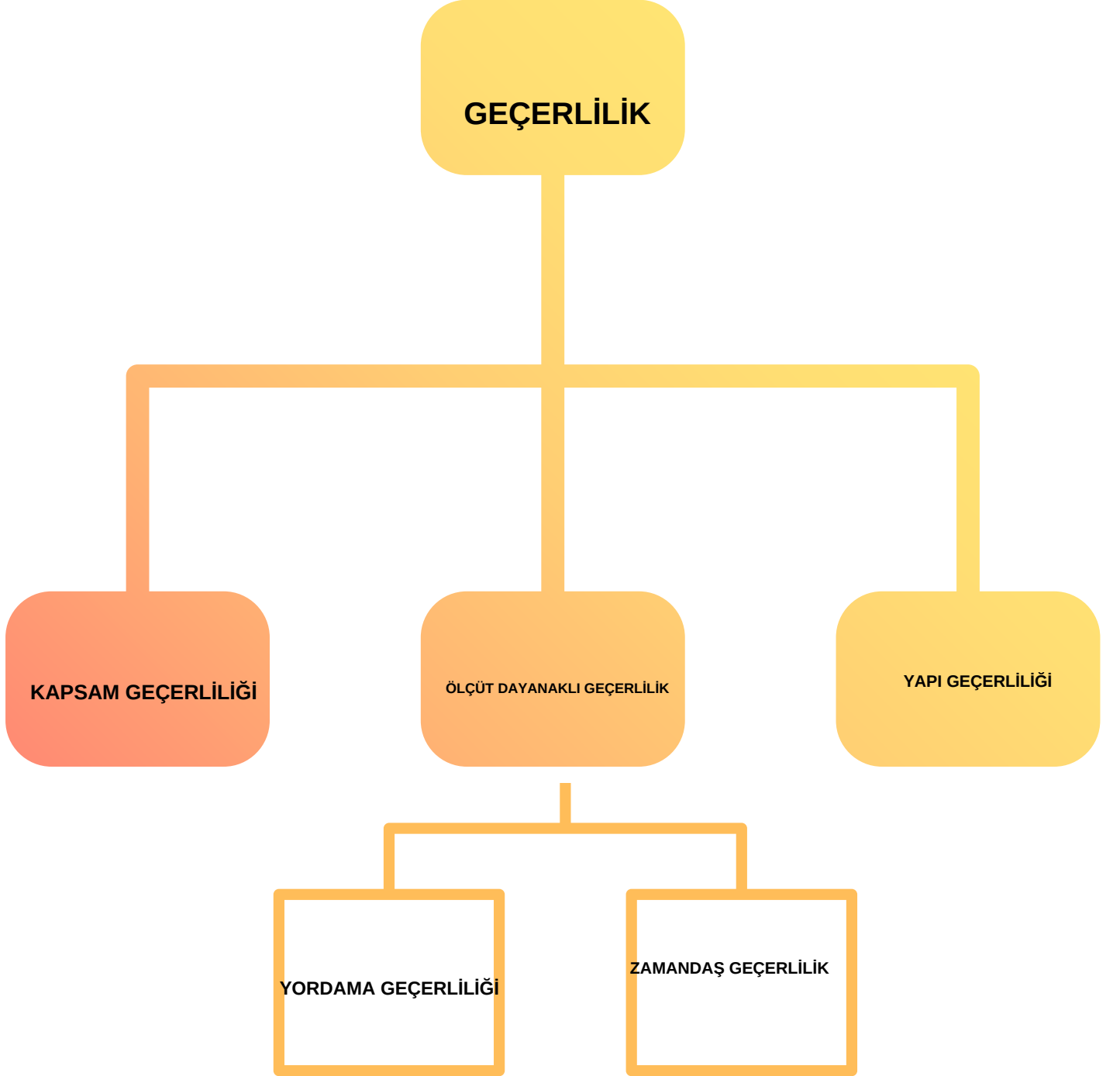
± 1 mükemmel korelasyon anlamına gelir ancak 1 mükemmel doğru orantılı ilişki anlamına gelirken, -1 ise mükemmel ters orantılı ilişki anlamına gelir.

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

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

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

GEÇERLİLİK SORGULAMA YÖNTEMLERİ:



KAPSAM GEÇERLİLİĞİ:

Öğretmen bir program dâhilinde 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. 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

2. Her bir sorunun / maddenin ölçmek istediği özelliği doğrudan ölçmesi, diğer bir deyişle kazanımla doğrudan ilgili olması gerekir.

Testin kapsam geçerliliğinin yüksek olduğunun söylenebilmesi için bu iki kriter birden sağlanmalıdır. Kapsam geçerliliği sorgulama yöntemleri: Mantıksal / rasyonel yöntemler

Bu yöntemler belirtke tablosu hazırlanması ve uzman görüşüne başvurulmasıdır.

BELİRTKE TABLOSUNUN HAZIRLANMASI

Öğretmen satırda davranışları, sütunda hedefleri yazar; belirlediği madde sayısı doğrultusunda davranış ve hedefleri yeterli ve dengeli bir biçimde temsil eden alanları seçer.

UZMAN GÖRÜŞÜNE BAŞVURMA: Uzmanla belirtke tablosu ve maddeler sunulur, uzmandan soruların maddelerin konu kapsamını yeterli ve dengeli bir biçimde ölçme durumunu ve soruların , maddelerin kazanımlarla doğrudan ilgili olma durumunu değerlendirmesi istenir.

İSTATİKSEL YÖNTEMLER: Uzmanlardan alınan dönütler betimsel / muhakemeye dayalı bir yolla çözümlenebilir ya da uzmanların “uygundur / uygun değildir” ya da “uygundur / düzeltme gerekir / soru kullanılmamalıdır” vb. biçimde değerlendirme yapması istenebilir.

GEÇERLİLİK SORGULAMA YÖNTEMLERİ:

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, tahmin demektir. Diğer bir deyişle yordama, eldeki bu nitelikteki veriden yola çıkarak geleceğe, henüz gerçekleşmemiş bir olguya ilişkin yapılan tahmindir. 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 testlerdir. Seçme amaçlı testlere YKS, KPSS, LGS vb. araçlar; diğer taraftan yönlendirme amaçlı testlere ise alan seçiminde ve/veya bir üst öğretim kurumuna yönlendirmek için uygulanan yetenek testleri, ilgi envanterleri, mesleki kişilik envanterleri vb. örnek olarak gösterilebilir. 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 özelliştir. Yordama geçerliliğinde ölçüt puan gelecekte belli olacağından beklemek gerekmektedir. 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(ün);

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.
3. Bireylerin özelliğini gerçekten yansıtan nesnel ve güvenilir bir ölçüt olmalıdır. Söz gelimi okulda alınan notlar bir ölçüt olarak alınmışsa öğrencilere verilen notlara başarının dışındaki etmenler etki etmemelidir.
4. Elde edilmesi kolay ve ekonomik olmalıdır.

ZAMANDAŞ GEÇERLİLİK: 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. 6: Bir zamandaş geçerlilik çalışma örneği Zamandaş geçerliliğinde geçerlilik katsayısının -1 ile 1 arasında değişmesi beklenir. Ölçüte bağlı olarak geçerlilik katsayısı ∓ 1 'e yaklaştıkça artar, 0 'a yaklaştıkça düşer. Geçerlilik katsayısı Zamandaş geçerlilikte de en zor ve önemli nokta ölçütün doğru bir biçimde belirlenmesidir. Ölçütün belirlenmesinde şu noktalar dikkate alınmalıdır: Ölçüt(ün);

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.

YAPI GEÇERLİLİĞİ:

Yapı, birbirleriyle ilgili olduğu düşünülen belli ögelerin ya da ögeler arasındaki ilişkilerin oluşturduğu bir örüntüdür. Bu anlamda, bir testin yapısını geçerleme süreci, temelde testin maddelerine verilen yanıtlar arasındaki ilişkilerin analizine dayanır.. İnsan davranışları çoğunlukla soyut bir yapıya sahiptir. Zekâ, 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; diğer bir ifadeyle 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. Güvenirlik 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 güvenirlik katsayısı kavramlarına değinmek gerekmektedir. Örneğin 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 birçok nedeni olabilir. Bir hareketten diğerine, “dikkat ve çaba” değişebilir. Özellikle uzun periyotlarda puan değişmesi, fiziksel büyüme, öğrenme ya da sağlık ve kişilikteki değişimlerden kaynaklanabilir.

GÜVENİRLİK



Birden Çok Uygulamaya
Dayalı Yöntemler

TEST-TEKRAR TEST

EŞDEĞER(PARALEL) TESTLER



Tek Uygulamaya Dayalı
Yöntemler

TEST YARILAMA

KR-20

KR-21

CRONBACH ALFA



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. 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. 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 (EŞDEĞER YARILAR / İKİ YARI GÜVENİLİRLİĞİ) 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 başvurulanan yöntemler:

- a) ilk yarı ve son yarı
- b) tek ve çift ve
- c) rastlantısalıdır. İlk ve son yarı yöntemi her test için uygun değildir.

Uygun olmadığı durumlar:

Testteki • maddeler basitten zora doğru sıralanmış ise

- maddeler konu içeriklerine ya da faktörlere göre kümelenerek yerleştirilmiş ise
- madde sayısı çok fazla ise

KUDER-RICHARDSON 20 VE 21 YÖNTEMLERİ: 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, doğru yanıtlara 1 puan, yanlış ve boş bırakılan maddelere 0 puan vererek puanlamanın yapıldığı ya da maddenin yanıtının iki seçenekli verildiği durumlardır. Eğer testteki maddeler farklı ağırlıklarla puanlanmışsa ya da test puanları şans başarısı için düzeltilmişse bu formüller kullanılmaz.

ACH ALFA YÖNTEMİ: Eğer ölçme aracının puanlaması çok kategorili ise KR-20 ile aynı mantık üzerine kurulu Cronbach alfa hesaplanır. Elde edilen katsayının ismi yine iç tutarlılık katsayısıdır. Derecelendirme ölçekleri puanlamanın çok kategorili olduğu araçlara örnek verilebilir.

GÜVENİLİRLİĞİ VE GEÇERLİLİĞİ ARTIRMA YOLLARI:

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

Güvenilirliği artıran temel kavramlar

- 1.** Bir testteki madde sayısı arttıkça birimler küçüldüğünden duyarlılık artar, bu nedenle hata miktarı azalacağından güvenilirlik artar.
- 2.** Bir ölçme işleminde genel bir ilke olarak puanlayıcı sayısı arttıkça güvenilirlik artar. Birden çok kişinin yaptığı puanlamanın ortalamasını almak, bir kişinin verdiği puandan daha güvenilir olma eğilimindedir.
- 3.** Puanlama nesnelliği, güvenilirliği büyük ölçüde etkiler. Bir testin farklı kişiler tarafından puanlanması ya da aynı kişinin farklı zamanlarda verdiği puanlar arasındaki tutarlılığa puanlama güvenilirliği adı verilir.
- 4.** KR-20 ve KR-21 yöntemlerinde bahsedildiği gibi testteki maddeler açısından benzeşiklik (homojenlik) arttıkça güvenilirliğin artması, diğer taraftan ayrışıklık (heterojenlik) arttıkça güvenilirliğin düşmesi beklenen bir durumdur.
- 5.** Testten elde edilen puanların güvenilirlik kestirimi için veri elde edilecek grubun (örneklem) büyüklüğü arttıkça grubun heterojenleşmesinden dolayı güvenilirlik artar.
- 6.** Aslında bir üst maddeyle de paralel bir biçimde maksimum performansı ölçen testlerde ortalama güçlüğü yaklaştıkça ($PP\Diamond = .50$) grup heterojenleşir ve dolayısıyla güvenilirlik artar.
- 7.** Maddelerin dil bilgisi kurallarına uygun, açık ve anlaşılır yazılması belki de güvenilirliği en çok artıran etmenlerden biridir.
- 8.** Öğrencilere test uygulaması öncesinde ve gerekliyse süreçte yönerge vermek, test almaya güdülenme ve hazırbulunuşluluğun artması açısından önemlidir.
- 9.** Fiziksel ortamın ses, sıcaklık, ışık, koku, dikkat dağıtıcı uyarıcılar, ergonomi vb. etmenler açısından uygun hâle getirilmesi yine güvenilirliği artıran bir başka etmendir.
- 10.** Eğer süreli bir test uygulaması ise sürenin yeterli verilmesi gerekir. Eğer bir hız testi değilse tüm katılımcıların maddelerin tümüne ulaşması beklenen bir durumdur. Tabii ki hız testleri bu durumdan bağımsız olarak değerlendirilebilir.
- 11.** Uygulamada bireylerin dikkatini dağıtacak ve/veya kaygısını artıracak davranışlardan kaçınılması, puanlama ya da veri girişi yapılırken dikkatli ve titiz davranılması ölçme işlemini yapan kişiden kaynaklanabilecek hataları azaltacağından dolayı güvenilirliği artırır.
- 12.** Maddelerin teste düzgün yerleştirilmesi, okumayı güçleştirecek bir unsurun bulunmaması; test katılımcısının yaşına, gelişim düzeyine uygun bir punto büyüklüğü kullanılması; baskı hataları bulunmaması vb. etmenler güvenilirliği artırır.

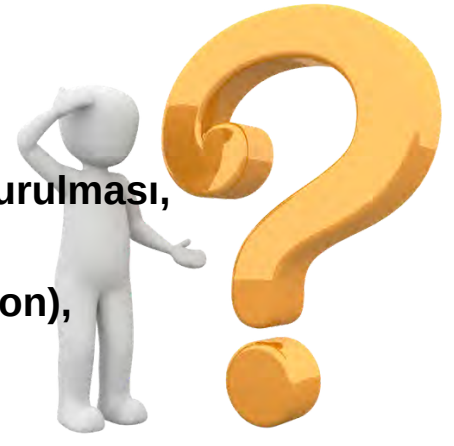
TEST GELİŞTİRME :

Test geliřtirmenin iřlem basamakları:

- a. Amacın belirlenmesi,
- b. Kapsamın belirlenmesi ve belirtke tablosunun oluřturulması,
- c. Denemelik maddelerin / soruların yazılması,
- d. Maddelerin / soruların gözden geçirilmesi (redaksiyon),
- e. Denemelik test formunun hazırlanması,
- f. Testin uygulanması,
- g. Test ve madde istatistiklerinin hesaplanması,
- h. Seçilen maddelerden oluřan nihai formun oluřturulması. **Pek çok**

kaynakta test geliřtirmenin adımları olarak bunlar olsa da sınıf içi ölçme ve deęerlendirme etkinliklerinde bu basamaklara uygun test geliřtirmek pratikte çok olanaklı deęildir. Bu durumda öğretnenler açısından sınıf içi ölçme ve deęerlendirmelerde uygulanması daha olanaklı bir test planı önerelim:

1. Sınavın amacının belirlenmesi (puanların hangi amaçlarla kullanılacağı)
2. Sınavda yoklanacak davranıřların belirlenmesi
3. Sınavın kapsayacağı konuların listelenmesi
4. Okulun takviminin ve sınav sonuçlarının kullanılacağı zamanın dikkate alındığı bir sınav gününün belirlenmesi
5. Bir ders saatine uygulanabilecek uzunlukta bir sınav süresinin belirlenmesi,
6. Belirtke tablosunun hazırlanması
7. Soru/madde türlerinin belirlenmesi [Soru/madde türleri sınavda kullanılacak davranıřlara göre belirlenir. Bir sınavda tek bir soru/madde türü kullanılabileceęi gibi birden fazla da kullanılabilir (soru/madde çeřitilmesi).]
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 çoęaltma yönteminin belirlenmesi
14. Sınavın uygulanma kurallarının belirlenmesi (yönerge yazılması)
15. Sınavın uygulanması
16. Test ve madde istatistiklerinin hesaplanması



. BAŞARININ ÖLÇÜLMESİNDE YÖNTEMLER



Ölçme Araç ve Yöntemleri



GELENEKSEL ÖLÇME YÖNTEMLERİ:

1. DOĞRU-YANLIŞ MADDELERİ:

Cevaplayıcının verilen ifadelerin doğru mu, yanlış mı olduğunu belirlemesinin istendiği madde türüdür. Doğruyanlış maddelerinde verilen bir tür “önerme”dir

.Avantajları: a. Hazırlaması, uygulaması ve puanlaması kolay, diğer bir ifadeyle kullanışlıdır.

b. Soru sayısı artırılabilir. c. Sistemik hata karışma olasılığı pek yoktur.

Dezavantajları: a. Şans başarısı olasılığı yüksektir.

b. İyi yazılmazsa genellikle üst düzey bilişsel özellikleri ölçemez.

c. Öğrenme eksiklerini belirleyemez.

d. Yanlış üzerinden öğretim yapılmaz.

e. Her derste/konuda kesin yanlış denebilecek durumlar belirlemek zordur.

Doğru-Yanlış Maddeleri Yazım Kuralları:

- Her madde tek ve belirli bir fikri belirtmelidir.
 - Doğru yanlış maddesi, kesinlikle doğru ya da kesinlikle yanlış olmalıdır.
 - Bir maddenin yanlışlığı önemsiz bir ayrıntıda ya da aldatıcı bir noktada olmamalıdır..
 - Mümkün olduğunca olumsuz ifade kullanılmamalıdır.
 - Kanı ifadeleri kendiliklerinden doğru ya da yanlış olmadığından bir kaynağa dayandırılarak verilmelidir
- Bir maddenin ifadesi kısa, açık ve yalın olmalıdır.
- Okuduğunu anlama süreçlerinin ilk basamağı, metin içerisinde açıkça verilmiş bilgiden doğrudan çıkarım yapmadır
 - Öğrencinin bir maddede ifadenin doğruluğuna ya da yanlışlığına karar vermesi gerektiğinde öğretmen özellikle vurgulamak istediği bir sözcük / terim vs. varsa onun altını çizebilir.
 - Doğru-yanlış maddesinin özellikle yanlış olduğu durumlarda, öğrencinin şansa puan almasını engellemek amacıyla ifadeyi düzeltmesi ya da yanlışlığın ne olduğunu belirtmesi istenebilir.
 - 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 ve yanlış maddeler, testte belirli bir örüntüye göre yerleştirilmelidir.

- 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.
- Mümkün olduğunca kısa yazılmalı, basit cümle yapısı tercih edilmelidir
- Olumsuz anlama sahip bir sözcük ya da ifade varsa altı çizilmelidir.
- İpucu vermekten kaçınmak gerekir.
- Üst düzey becerileri yoklamak için tablo, grafik, harita ya da okuma parçası gibi bir öncülün kullanılması önerilir.

2.EŞLEŞTİRME MADDELERİ:

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

Avantajları: a. Hazırlaması, uygulaması ve puanlaması kolay, diğer bir ifadeyle kullanışlıdır. b. Soru sayısı artırılabilir. c. Sistematiik hata karışma olasılığı pek yoktur.

Dezavantajları: a. Şans başarısı olasılığı yüksektir. b. İyi yazılmazsa genellikle üst düzey bilişsel özellikleri ölçemez

Eşleştirme Maddeleri Yazım Kuralları

.Her eşleştirme soru grubunda yönerge/açıklama yazılmalıdır.

.Eşleştirmenin nasıl yapılacağı konusunda bilgi verilmelidir.

• Seçeneklerin kullanılıp kullanılmayacağı, birden fazla kez kullanma durumu belirtilmelidir. Bir eşleşme takımındaki öncüller ile cevaplar eşit sayıda olmamalıdır.

• 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.

• 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.

- Uzun ifadeler öncül olarak kullanılmalı ve öncüller sütunu sayfanın sol sütununa, cevaplar sütunu ise sayfanın sağ tarafına yerleştirilmelidir.
- Bir eşleştirmeli maddeler grubundaki madde sayısı, en az 6, en çok 15 olmalıdır. Madde sayısının 6'dan az olması, salt tahminle doğru cevabın bulunma olasılığını artırır. 15'ten çok madde kullanıldığında ise cevaplayıcıların doğru olarak eşleştirilecek ifadeleri ayıklayıp seçmesi çok zaman alır.
- Maddelerin tümü aynı sayfada bulunmalıdır.
- Cevapların seçileceği sütun, bir kelime listesi ise alfabetik sıraya göre
- Bu eşleştirme maddeleri öğrencilerin yalnızca hatırlama düzeyi dışında, kavrama düzeyindeki öğrenmelerinin de yoklanabileceğine ilişkin örnektir. Örneğin grafik

3. ÇOKTAN SEÇMELİ MADDELER:

Çoktan Seçmeli Maddelerin Özellikleri:

- Hazırlanması zor ancak puanlaması kolaydır.
- Puanlama objektiftir.
- Üst düzey bilişsel becerilerle ilgili madde yazmak zordur.
- Doğru cevabı içinde barındırması nedeniyle sentez düzeyinde soru sormak olanaklı değildir. yaratıcılığın denetlenmediği yönünde haklı bir eleştiri vardır.
- Doğru cevabı içinde barındırması nedeniyle puanlara şans başarısı karışma ihtimali vardır.
- Çok sayıda soru sorulabilir.
- Sistemik hata karışma ihtimali çok düşüktür.
- Çeldirme mantığı temellidir.

ÇOKTAN SEÇMELİ MADDE TÜRLERİ

A) Tek bir doğru cevap isteyen maddeler

- Madde kökü soru kipinde(olumlu- olumsuz) olan maddeler
- Madde kökü eksik olan maddeler

B) En doğru cevabı isteyen maddeler



Çoktan Seçmeli Madde Yazım Kuralları:

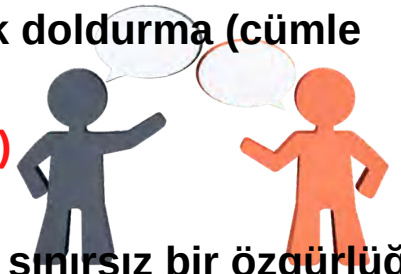
- İdeal bir çoktan seçmeli maddede, bilen bir öğrenci, seçenekleri görmesine gerek kalmaksızın doğru cevabı verebilmelidir.
 - Öğrencinin sözcük repertuvarında bulunmayan, henüz öğrenmediği sözcük, kavram, terim vb. kullanılmamalıdır.
 - Bir durumun resim ya da fotoğrafla anlatılabildiği durumda, çokça cümle kurmak yerine görsellerden yararlanılabilir.
 - Gereksiz görsel kullanılmamasına da özen gösterilmelidir.
 - Bir maddeyi okuyan bir öğrenci öğretmenin zihninden geçeni keşfetmek zorunda değildir.
 - Karışık bir dil yerine ortalama bir öğrencinin anlayabileceği bir dil kullanılmalıdır. Bir çoktan seçmeli maddeyi zorlaştırmak isteyen öğretmen temelde iki yöntem kullanabilir: a. Çeldiricileri doğru yanıtla yaklaştırmak (güçlendirmek) ve b. Davranışı üst düzey becerileri ölçen bir formda hazırlamak.
 - Madde kökünde gereksiz yere sözcük kullanılmamalıdır.
 - Seçeneklerde gereksiz sözcük tekrarı kaçınılmalıdır
- Bir maddenin doğru yanıtı, başka bir maddenin kökünde ya da seçeneklerinde bulunmamalıdır.
- Seçeneklerin hepsi birbirleriyle tutarlı ve ayrı gramatik yapıya sahip olmalıdır.
 - İpucu vermekten kaçınmak gerekir.
 - Birbirinin tam zıttı olan durumların seçeneklerde verilmesi çoğunlukla uygun değildir. Çünkü genellikle ikisinden biri doğrudur. Şans başarısı artar, geçerlilik ve güvenilirlik düşer.
 - Seçenekler yazılırken anlamca birbirini içeren / kapsayan ifadelerin kullanılmamasına dikkat edilmelidir.
 - Seçeneklerin uzunlukları birbirine eşit olmalıdır.
 - Doğru cevaplar teste dağıtılırken dikkat edilmeli, doğru cevapların bir örüntü göstermemesi sağlanmalıdır.
 - “Hepsi” ve “hiçbiri” bir seçenek olarak kullanılmamalıdır.
 - 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 düzeylerinde 5, ortaokul düzeyinde 4 ve ilkokul düzeyinde 3 seçenek kullanılabilir.
 - Seçenekler bir sıra ile verilmelidir. Seçenekler baş harflerine göre alfabetik olarak, eğer sayı iseler büyüklük olarak sıralanmalıdır.
 - Hangi özellik ölçülürse ölçülsün, çoktan seçmeli madde yazımının bütün aşamalarında hem madde kökünde hem de seçeneklerde dil bilgisi ve imla kuralları konusunda yüksek birtitizlik ve özen göstermek çok önemlidir.
- Çoktan seçmeli maddelerde seçeneklerin baş harfleri büyük olmalıdır.
- Bir madde kökü ile seçenekleri aynı sayfada olmalıdır.
- Eğer ortak köke dayalı madde grubu yazılacaksa hangi soruların ortak köke göre yanıtlanacağı belirtilmelidir.

AÇIK UÇLU SORULAR

A) Uzun yanıt gerektiren maddeler

B) Yanıtı sınırlı maddeler olarak ikiye ayrılır. Ancak bu testler sahada genellikle essay (yazılı yoklama), kısa yanıtı, boşluk doldurma (cümle tamamlama) ve sözlü yoklama olarak geçmekte.

1.ESSAY (YAZILI YOKLAMA)



Yazılı yoklamaların temel özellikleri:

- Cevaplayıcıların sorulara cevap verme konusunda sınırsız bir özgürlüğü vardır.

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

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

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

- Bu sınav türü, dünyada en eski ve günümüzde de öğretmenlerin hâlâ en çok tercih ettiği testlerden biridir. Bunun iki nedeni vardır: a. Pratik nedenler: Hazırlaması kolaydır. b. Eğitsel nedenler: Üst düzey bilişsel becerileri ölçmek için geleneksel yöntemler arasında en avantajlı yazılı test etme yoludur (Sentezde tek yol.).

- Yazma, problem çözme, bilgileri organize etme, analiz vb. özgün bir ürün ortaya koyma (yaratıcı düşünme), eleştirel düşünme, yeni durumlara beceriyi transfer etme, analitik düşünme, bilimsel düşünme, hipotez oluşturma ve neden-sonuç ilişkilerini açıklama, veri düzenleme, güçlü ve zayıf yönleri belirleme gibi becerileri ölçmek için çok kullanışlıdır.

- Çok fazla soru sorulamaması (Okuma ve yazma eylemlerinin süreler arasında ortalama 10 kat fark bulunmakta.) bir dezavantajdır.

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

- Puanlama öznelliği belki de en önemli dezavantajdır.

“Essay”de Nesnelliği Artırma Yolları:

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

- Bütüncül (holistik) ya da analitik rubrik hazırlanabilir

2.)KISA CEVAPLI SORULAR

Kısa cevaplı testler; öğrencinin bir sözcük, bir rakam, bir tarih ya da en çok bir cümle ile cevaplandırabileceği sorulardan oluşur. İstenen cevapların kısa ve belirgin olması, kısa cevaplı testlerin puanlanmasını genellikle nesnel kılar.

- Eğer konu kapsamı bakımında homojen bir yapıda ise gruplandırılarak sorulması daha uygun olacaktır.
- Ayrıntılı bir cevap anahtarı hazırlanmalıdır.

Avantajları: a. Hazırlaması, uygulaması ve puanlaması kolay; diğer bir ifadeyle kullanışlıdır. b. Soru sayısı artırılabilir. c. Sistemik hata karışma olasılığı pek yoktur.

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

3) CÜMLE TAMAMLAMA SORULARI

Cümle tamamlama, kısa cevaplı / sınırlı cevap gerektiren soru grubudur.Cümle tamamlama soruları boşluk doldurma olarak da geçmektedir.

Avantajları: a. Hazırlaması, uygulaması ve puanlaması kolay; diğer bir ifadeyle kullanışlıdır. b. Soru sayısı artırılabilir. c. Sistemik hata karışma olasılığı pek yoktur.

Cümle Tamamlama Soruları Yazım Kuralları:

- Her kısa cevap maddesi, yalnızca tek bir doğru cevabı olacak biçimde yapılandırılmalıdır.

.Bir maddenin ifadesinde, o maddenin cevabının bulunmasında işe yarayacak ipuçları vermekten kaçınılmalıdır.

..Cümlelerin gramer yapısı bağlamında ipucu genellikle eklerle verilmektedir.

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

- İlköğretimde cevapların ayrı bir cevap kâğıdına işaretletilmesi uygun görülmemektedir.

.Bir cümlede çok sayıda boşluk bırakılmamalı ve cümlelerden sadece anahtar niteliğindeki anlamlı ve önemli sözcükler çıkarılmalıdır.

.Herkesçe aynı biçimde algılanacak, soru yazmak gerekir.

4) 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 vb

- Sorular sözlü olarak sorulur ve cevap sözlü olarak verilir.
- Öğretmen ile öğrenci arasında karşılıklı ve devamlı bir etkileşim vardır.
- Bireysel bir test olması nedeniyle her öğrenciye ayrı soru sorma zorunluluğu vardır.
- Bireysel bir test olması nedeniyle tüm öğrencileri test etmek için gereken süre oldukça fazladır.
- Öğrencinin cevaplarını gözden geçirme şansı yoktur.
- Cevaplama çoğu zaman üzerinde düşünmeye ve tasarlamaya olanak olmadan verilir.
- Puanlama hemen yapılır. Bu ise cevabın doğruluğunun genel izlenimle yapılmasına neden olur.
- Öğrencinin yerinde ayağa kaldırılması ya da tahtaya kaldırılması kaygı vb. bireyden kaynaklanan hata miktarını artırabilir.
- Öğretmenin ölçme sonucuna sistematik hata karıştırma olasılığı yükselir.
- Sözlü anlatım becerisi iyi olan bir öğrenci, iyi olmayan bir başka öğrenciden daha az biliyor olsa da yüksek puan alabilir.
- Öğrencinin kılık-kıyafeti, diksiyonu, hâl ve hareketleri puanlamaya artı ya da eksi yönde etki edebilir.
- Çok fazla soru sorulamayacağı için kapsam geçerliliği problemi olabilir.

Sözlü Yoklama Uygulama Kuralları:

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

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

DESTEKLEYİCİ DEĞERLENDİRME YAKLAŞIMLARI

1. PORTFOLYO: Öğrencinin bir bütün olarak gelişim ve öğrenme süreci ile ürünlerini gösteren, aynı zamanda değerlendirilmesini de sağlayan sistemli ve amaçlı olarak oluşturulmuş dosyalardır. • Portfolyolar öğrencinin yaptığı çalışmaların bir araya getirildiği herhangi bir çalışma dosyası değildir.

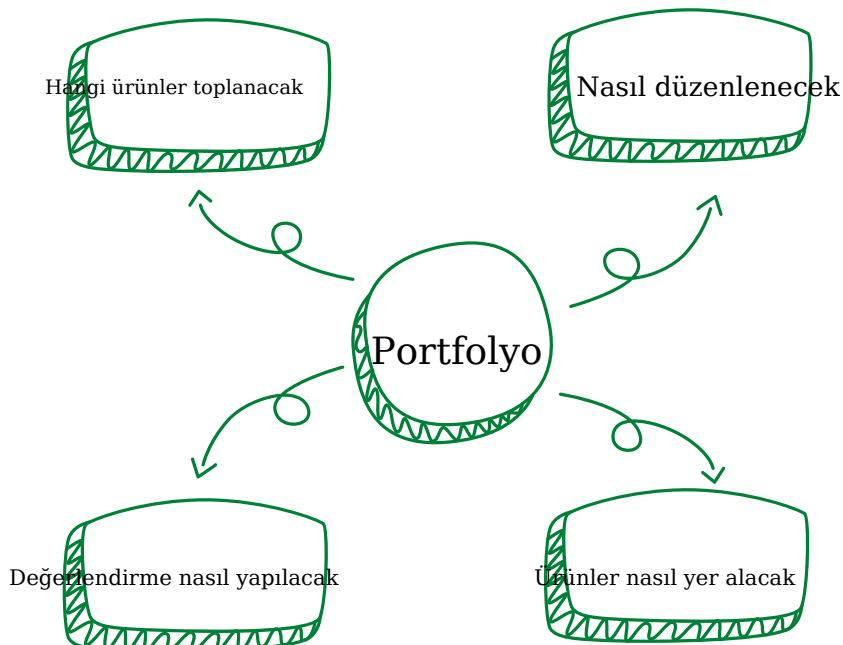
• Öğrenciyi başarılı / başarısız olarak sınıflandırmayı sağlayan bir araç değildir 7

• Öğrencilerin özelliklerini birbirleri ile karşılaştırmak amacıyla oluşturulmuş araçlar değildir. Öğrenci portfolyosunun kabaca iki türü vardır. Bunlar:

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

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

Portfolyo değerlendirme süreci



TOPLAMA:

1. Hangi çalışmaların toplanacağına ve hangi özelliklerin gözleneceğine karar verilmesi
2. Öğrencilere çalışmalarının bir dosyada toplanacağını açıklanması ve çalışmalarını saklama konusunda öğrencilerin teşvik edilmesi
3. Her öğrenci için sınıfta çalışmalarının toplanacağı ayrı bir kutu, dosya vb. oluşturulması
4. Her bir çalışmanın ve öğretmen kaydının üzerine tarih yazılması

SEÇME 1. Öğrenci seçimini kendi başına veya öğretmenin rehberliğinde yapabilir.

2. Portfolyonun türüne ve öğretmenin koyduğu koşullara bağlıdır.
 - a. Süreci yansıtan portfolyolarda öğretmenin belirlediği konuları / gelişim alanlarını yansıtan çalışmaların seçilen örnekleri, öğrenme / gelişim sürecini yansıtacak şekilde yer alır.
 - b. Ürünü yansıtan portfolyolarda öğretmenin belirlediği konular ile ilgili ortaya çıkan ürünler arasından seçilenler yer alır.

YANSITMA: Portfolyoyu herhangi bir çalışma dosyasından ayıran en önemli aşamadır. Bu aşamada öğrenci “Bu çalışmayı niçin yaptık?” sorusunu yanıtlar.

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

Yansıtma Soruları:

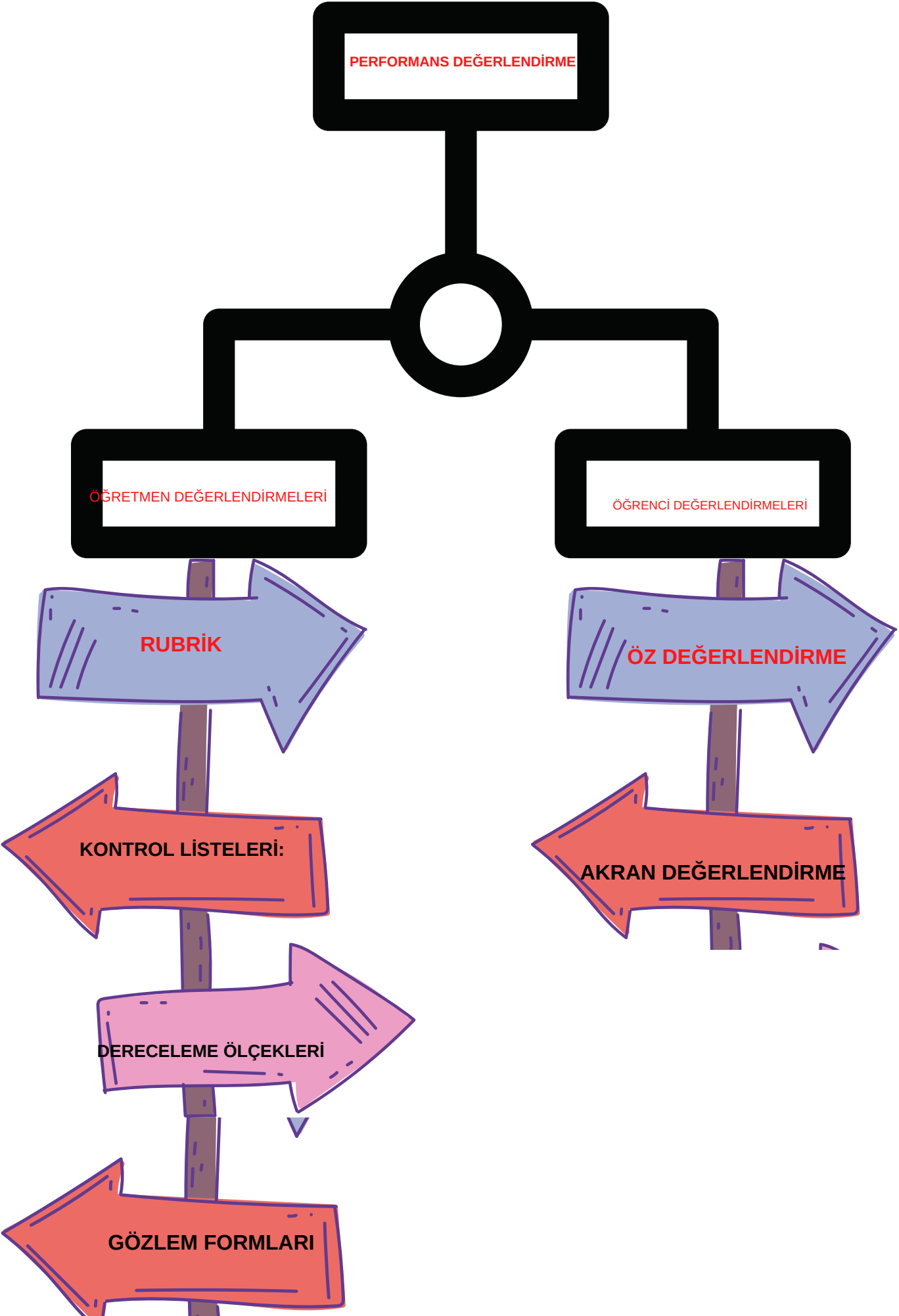
- Bu çalışmayı nasıl yaptım?
- Bu çalışmadan ne öğrendim?
- Bu çalışmayı daha da geliştirebilir miyim? Nasıl?
- Çalışmalarım içinde en çok sevdiğim hangisi? Neden?
- Bana zor gelen bir çalışmam var mı? Varsa neden?
- Bu çalışmayı portfolyoma neden koydum?

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



PERFORMANS DEĞERLENDİRME

ÖZ DEĞERLENDİRME: 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.

Öz Değerlendirmenin Olası Dezavantajları:

a. Öğrencilerin kendi performanslarını değerlendirirken yanlış davranışları söz konusu olabilir.

b. Başlangıçta deneyimsizlik nedeniyle performansın değerlendirilmesinde yanlışlar olabilir.

2. 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 Değerlendirmenin Olası Dezavantajları:

a. Öğrencilerin yanlış davranışları söz konusu olabilir. Kişisel ilişkiler olumlu ya da olumsuz yönde değerlendirmeye etki edebilir.

b. Genel izlenimle puan verme söz konusu olabilir.

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

Rubrikler ikiye ayrılır:

a. Bütüncül (Holistik) rubrik

b. Analitik rubrik

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

4. KONTROL LİSTELERİ: Gözlenen performans ürününün ölçütlere uygunluğunu “evet-hayır”, “var-yok”, “gösterdi-göstermedi” vb. bir biçimde kategorik (1-0) olarak puanlama amacıyla kullanılan araçlardır. • Özellikle sergilenecek performans detaylı ve ardışık eylemler gerektirdiği zamanlarda kullanışlıdır. Örneğin bir deneyin eyleme dökülmesi vb.

5. DERECELEME ÖLÇEKLERİ: 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.

6. GÖZLEM FORMLARI: Öğrenme çıktılarının somut olarak gözlenebildiği bazı alanlarda bu yöntem oldukça kullanışlıdır. • Özellikle fen derslerinde, meslek liselerinin somut performans ürünlerinin geliştirildiği vb. alanlar için oldukça uygundur. • Gözlemler öğrenciler hakkında doğru ve hızlı bilgi elde edilmesini sağlar. • Gözlem formları yarı yapılandırılmış biçimde olabileceği gibi tam yapılandırılmış bir formatta da olabilir.



GELENEKSEL YÖNTEMLER

1. Ürün değerlendirilir.
2. Öğrencinin ulaştığı noktanın tespiti önemlidir.
3. Essay dışında genellikle üst düzey düşünme becerilerinin değerlendirilmesinde yetersizdir.
4. Değerlendirme öğrenmeden ayrıdır.
5. Bireyden ziyade gruba odaklıdır.
6. Başarının bireyler arası değerlendirilmesine odaklıdır.
7. Geçerlilik ve güvenilirlik kontrolü daha kolaydır.

DESTEKLEYİCİ YÖNTEMLER

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



TEST VE MADDE İSTATİSTİKLERİNE GENEL BİR BAKIŞ

1. TEST İSTATİSTİKLERİ: Ölçme işlemi sonucunda bireylerin toplam puanları üzerinden hesaplanan istatistiklerdir. Eğitimde en sık kullanılan istatistikler merkezî eğilim ölçüleri, değişkenlik ölçüleri, dağılım özellikleri, standart puanlar vb.dir. Merkezî eğilim ölçüleri:

Aritmetik ortalama

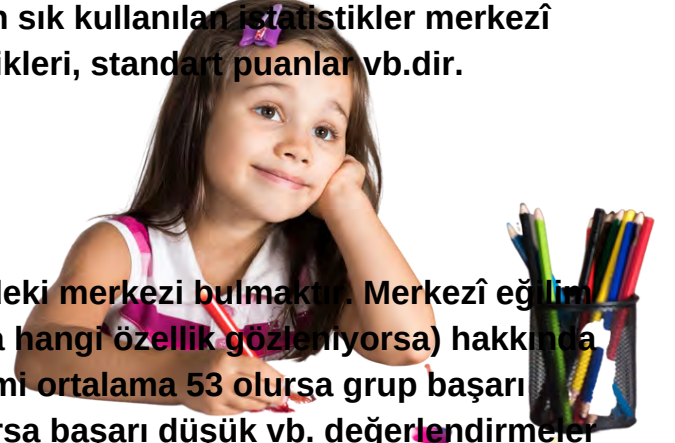
Medyan (ortanca)

Moddur.

Bu üç istatistiğin de temel işlevi bir puan dizisindeki merkezi bulmaktır. Merkezî eğilim ölçüleri, örneğin ortalama, grubun başarısı (ya da hangi özellik gözleniyorsa) hakkında kabaca da olsa fikir yürütmemizi sağlar. Söz gelimi ortalama 53 olursa grup başarı düzeyi açısından vasat, 82 olursa başarılı, 27 olursa başarı düşük vb. değerlendirmeler yapılabilir.

Diğer taraftan değişkenlik ölçüleri ranj (dizi genişliği), varyans, standart sapma ve çeyrek sapmadır. Değişkenlik ölçüleri grubun kabaca homojen ya da heterojen bir özellik gösterip göstermediği, standart sapma örnek olarak verilirse değişkenliğin ortalama etrafında nasıl dağıldığını gösterir. Puanların değişkenliği, grup hakkında biraz daha ayrıntılı bir değerlendirme yapılmasını sağlar. Evrende pek çok fiziksel, fizyolojik ya da psikolojik özelliğin normal dağılım eğrisine göre dağıldığı, diğer bir ifadeyle bu dağılımın “normalin bir ölçüsü” olduğu kabulü yapılır. Sivri, basık, sola çarpık ve sağa çarpık dağılımlar normalden sapma gösteren dağılımlardır. Normalden sapma durumlarının pek çok nedeni olabilir. Örneğin grup sayısının az olması, testin kolay ya da zor olması, grubun ölçülen özellik açısından genel niteliği vb. durumlar. Normal dağılım ortalama ve standart sapma gibi basit iki istatistikle özetlenebilir. Normal dağılım ortalama etrafında üç standart sapma sağa ve sola, toplamda altı alanda özetlenebilir (%99,74'lük kısmı). Dolayısıyla bir kişinin ortalamadan kaç standart sapma uzaklaştığı, grup içerisindeki yerini gösterir.

2. MADDE İSTATİSTİKLERİ: Ölçme işlemi sonucunda bireylerin madde puanları üzerinden hesaplanan istatistiklerdir. Eğitimde en sık kullanılan istatistikler madde güçlük indeksi, madde ayırt edicilik indeksi, madde güvenilirliği vb.dir. Madde güçlük indeksi; maddenin kabaca zor mu, orta mı, kolay mı olduğunu gösteren bir istatistiktir. Madde güçlük indeksi 1'e yaklaştıkça maddenin kolaylaştığı, 0'a yaklaştıkça zorlaştığı anlaşılır. 0.50 civarı ise ortalama güçlükte olduğu biçiminde yorumlanır. Her ne kadar ayrı bir formülü olsa da madde güçlükleri toplamının madde sayısına bölümü testin ortalama güçlüğüne ortaya koyar. Bu noktada öğretmenlerin zihninde şu sorular olabilir: “Hazırlanış şekli 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?” Bu tür amaçlarla hazırlanan sınavlarda kolay ve zor sorular teste konulabilir ancak ortalama güçlükteki sorular her zaman sayıca daha fazla olmalıdır. Nihayetinde testin ortalama güçlüğü hesaplandığında 0.50 olması ideal durumdur. Pekâlâ “test zor ya da kolay olursa ne olur?” Bir test zorlaştıkça dağılımın sağa çarpık, kolaylaştıkça sola çarpık olma olasılığı artar. Diğer bir ifadeyle grup başarı düzeyi açısından homojenleşmeye başlar ve homojen bir dağılımda testin bilenle bilmeyeni birbirinden ayırt etme gücü azalabilir ve güvenilirlik katsayısı düşebilir.





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

Eğitim alanında **başarıyı**, belirli bir konu alanında veya uygulama alanında kazandırılan/kazandırılmaya çalışılan kazanımlara bireylerin ulaşma/erişme düzeyleri olarak anlamlandırıyoruz.

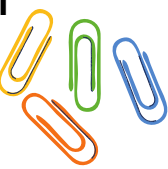
Erişi ya da başarı kavramı, sunulan bilgilere, sunulan içeriğe o eğitimi alan bireylerin ne kadar ulaşabildiğine dair bir gösterge, **erişim düzeyi** olarak ifade edilmektedir.

Beceri ise başarıdan farklı olarak tekil kazanımlardaki erişim düzeylerinden ziyade öğrencilerin bilgi ve becerilere sahip olduktan sonra bunları ne kadar kullanabildiğine dair bir gösterge olarak değerlendirilmektedir. Beceri; bilgilerin bir araya toplanıp, doğru bilgilerin seçilip ardından seçilen bilgilerin doğru şekilde kullanılması, gerçek ya da tasarlanan bir durumda kullanılabilmesini ifade etmektedir.

Eğitimin daha geleneksel olduğu zamanlarda bilgiye sahip olmak, bilgiyi ezberde tutmak ya da yeri gelince hatırlamak ve anlamak kıymetliydi. Bilgiye ulaşmanın gittikçe kolaylaştığı günümüzde artık bu bilgileri uygun durumlarda kullanmak önem kazanmakta ve beceri kavramının önemi öne çıkmaktadır.

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

Bilimsel ve teknolojik gelişmeler sonucunda bilgiye ulaşmanın kolaylaşması, 2000'li yıllardan sonra eğitim sistemlerinin öğrencilere bilgiye ulaşmanın ötesinde bunları uygulama becerisi kazandıracak şekilde revizyona gitmesini gerektirmiştir. Birçok eğitim sistemi artık daha fazla ve daha ağırlıklı şekilde öğrencilerine gerçek hayatta karşılaşacakları problemleri çözebilecek becerileri kazandırmaya çalışmaktadır. Bu becerilere sahip insan kaynağının yetişmesi, ülkelerin en büyük önceliğidir. Eğitim de öğrencilere söz konusu becerilerin kazandırılmasında ülkelerin sahip olduğu en önemli araçtır. Bu nedenle ülkeler eğitim sistemlerine daha fazla daha yatırım yaparak ve gerekli iyileştirmeleri sağlayarak öğrencilerine gerekli becerileri kazandırmaya, onların okuryazarlıklarını geliştirmeye gayret etmektedir.





Uluslararası İzleme Çalışmaları

Eğitim izleme çalışmaları, bugün çok popüler olmakla beraber aslında çok uzun süredir sahada olan uygulamalardır. Bugün birçok ulusal izleme araştırmasına temel teşkil eden ilk çalışmalardan biri 1964 yılında gerçekleştirilen uluslararası matematik izleme çalışmasıdır. TIMSS'in de geliştiricisi olan kurum (IEA) tarafından gerçekleştirilen bu çalışma 1960'larda matematik üzerinde başlamıştır. Aslında bu tarz izleme çalışmalarına ihtiyaç duyulmasının ve bu çalışmaların popüler olmasının sebeplerinden ilki karşılaştırılabilir veri ihtiyacıdır. Yaptıklarımızla eğitimde istediğimiz ivmeyi yakalayabildik mi, yakalayamadık mı? Bu soruları yanıtlamak için ise ülkeler arasında karşılaştırılabilir veriye ihtiyaç duyuluyor. Eğitim izleme araştırmaları da 1960'lı yıllardan itibaren özellikle bu ihtiyacı karşılıyor. 1960'lı yıllarda matematik alanı ile başlayan çalışmaların bugün kapsamı büyük ölçüde genişlemiştir. Örneğin PISA ile beraber değerlendirilen alanlar artmış; matematik, fen ve okuma becerileri üç ana alan olarak belirlenmiştir. Uluslararası Matematik ve Fen Eğilimleri Araştırmasında (TIMSS) matematik ve fen, Uluslararası Okuma Becerilerinde Gelişim Araştırmasında (PIRLS) ise okuma becerileri değerlendirilmektedir. Ayrıca izleme çalışmalarının kapsamı günümüzde sosyal ve duygusal beceriler, vatandaşlık becerileri, yetişkinlik becerileri, dijital okuryazarlık gibi çok farklı alanları da içerecek şekilde genişlemiştir. İzleme araştırmalarının sağladığı ikinci katkı, ihtiyaç duyulan bağlamsal ve bütüncül ilişkilere dair bulgular sunmasıdır. Çalışmalarda sadece eğitim performansının nasıl değiştiğinin ötesinde elde ettiğimiz bilginin kapsamı ve bağlamı bundan çok daha fazlasıdır. Aslında bu çalışmalar öğrenci performansındaki değişimin birçok özellikle ilişkili olup olmadığı konusunda da bilgi vermektedir. Örneğin öğrencilerin akademik başarıları okul iklimi, eğitim yatırımları, öğretmen özellikleri, öğrencilerin sosyoekonomik özelliklerine göre nasıl değişiyor gibi önemli sonuçları yine bu çalışmalar sayesinde görebiliyoruz.

Günümüzde En Yoğun Katılım Gösterilen Çalışmalar

PISA, Ekonomik Kalkınma ve İşbirliği Örgütü (OECD) tarafından geliştirilen ve üç yıllık periyotlarla uygulanan izleme çalışmasıdır. **OECD**, Türkiye'nin kurucu üyesi olduğu bir örgüt olup temel amacı tüm üye ülkeler arasındaki ekonomik iş birliklerini ve kalkınmayı destekleyecek programlar oluşturarak birlikte kalkınmayı sağlamaktır. Bu bağlamda OECD, eğitimi, ekonomik kalkınmaya katkı sunan ve çağdaş vatandaşlık becerilerine sahip insan kaynağının geliştirilmesinde en önemli araç olarak görmektedir. Bir diğer izleme çalışması olan **TIMSS** ise Uluslararası Eğitim Başarılarını Değerlendirme Kuruluşu (IEA) tarafından geliştirilmekte ve dört yıllık periyotlarla uygulanmaktadır. Türkiye uzun süredir hem PISA'ya hem de TIMSS'e katılım göstermektedir. Son olarak PIRLS de yine IEA tarafından beş yıllık periyotlarla gerçekleştirilen bir izleme çalışmasıdır. Türkiye uzun bir aradan sonra 2021'de tekrar PIRLS'e katılmış ancak bu uygulamanın sonuçları 2022 yılının sonunda açıklanacaktır.

PISA KAPSAMI VE SONUÇLARI

OECD, ekonomik kalkınmanın ve sürdürülebilirliğinin sağlanmasındaki en önemli kaynak olan beşerî sermayeyi üreten faktör olarak eğitimi konumlandırmaktadır. Eğitimin ekonomik kalkınma üzerindeki belirleyici rolü dolayısıyla eğitim niteliğini detaylı olarak ele alan ve üç yıllık periyotlarla uygulanan PISA çalışmasını gerçekleştirmektedir. En güncel PISA uygulaması olan 2018 döngüsüne 79 ülke/ekonomi katılım göstermiştir. Dünyanın dört bir yanından farklı coğrafyalardan, farklı kültürlerden, farklı dinamiklere sahip eğitim sistemleri olan ülkeler yaygın bir şekilde PISA'ya katılım göstermektedir. PISA, 15 yaş grubunu hedef grup olarak almaktadır. Bunun nedeni birçok Avrupa ve OECD ülkesinde 15 yaşa tekabül eden dönemin yaklaşık olarak ortaokulun sonuna denk gelmesi ve bu dönemde zorunlu eğitimin tamamlanmasıdır. Dolayısıyla PISA, zorunlu eğitim dönemini bitiren öğrencilerin henüz iş gücü piyasasına katılmadan ya da eğitimlerine devam etmeden matematik, fen ve okuma becerileri alanlarında hangi yeterliklere sahip olduğuna dair bilgi vermektedir. PISA, 2000 yılından bu yana uygulanmakta ve her üç yıllık döngüde bir alana ağırlık verilmektedir. 2000 yılında ağırlıklı alan okuma becerileri iken 2003 yılında matematik 107 okuryazarlığı ve 2006 yılında fen okuryazarlığı, 2018 yılında tekrar okuma becerileri olacak şekilde çapraz bir modelleme kullanılmaktadır. PISA 2018 uygulamasına ait Türkiye örnekleme incelendiğinde, 15 yaş grubunda kısmen 8. sınıfta bulunan ortaokul son sınıf öğrencilerimiz bulunmakta ama yoğun olarak 9 ve 10. sınıf öğrencilerimizin olduğu görülmektedir. Okul türlerine göre dağılım incelendiğinde ise öğrencilerimizin %44'ünün Anadolu lisesi, %31'inin mesleki ve teknik Anadolu lisesi ve %14'ünün Anadolu imam hatip lisesinde eğitim gördüğü, bu dağılımın ülkemizde hâlihazırda lise türlerindeki öğrenci dağılımıyla benzerlik gösterdiği ifade edilebilir. Örneklem seçimi yedi coğrafi bölgeye göre değil, Türkiye'nin sosyoekonomik farklılıklarını da dikkate alan 12 bölgeden oluşan İBBS Düzey-1 istatistiksel örnekleme sistemine göre yapılmaktadır. Bu örnekleme sürecinde bölgesel öğrenci yoğunlukları dikkate alınmakta ve daha büyük illerimizin olduğu bölgeler daha fazla öğrenciyle, daha az olduğu yerler ise daha az öğrenciyle temsil edilmektedir.

YETERLİLİK DÜZEYLERİNDE DEĞİŞİM

@gunes ogretmenle etkinlikler

Türkiye, PISA 2003 yılından bu yana katılım göstermekte olup 2003 senesinde 441 olan okuma puanı genel bir artış eğilimi göstererek 466 puana ulaşmıştır. 2003'ten 2012'ye kadar kesintisiz bir artış, 2015 döngüsünde bir düşüş ve söz konusu 2018 döngüsünde tekrar artış vardır. PISA'nın en önemli çıktılarından bir tanesi öğrencilerin hangi yeterlik düzeyinde olduğuna dair veriler sunmasıdır. Yeterlik düzeyleri son derece önemlidir çünkü her bir yeterlik üzerinde bulunan öğrencinin o alanda ne yapıp ne yapamadığını, davranış bazında hangi davranışları gösterip hangi davranışları gösteremediğine dair somut açıklamalar yapmaktadır. Örneğin en üst düzeye yeterlik düzeyi olan 6. düzeydeki bir öğrencinin çok karmaşık metinleri rahatlıkla okuyabilmekte, anlamlandırabilmekte, metin içerisindeki dinamiklerin farkına varıp bunları ilişkilendirebilmekte, farklı metinlerdeki örtülü anlamları fark edebilmektedir. 2015'ten 2018'e puan seviyesindeki artışın en memnun edici yanı, en düşük yeterlik düzeyleri olan 1-b, 1-a ile kısmen 2. yeterlik düzeylerindeki öğrenci oranları azalırken daha üst yeterlik düzeylerindeki öğrenci oranlarımızın artmasıdır.

OKUL TÜRLERİNE GÖRE OKUMA BECERİSİ PUAN ORTALAMALARI

Okul türlerinin okuma becerisi puan ortalamalarının nasıl farklılıklar gösterdiği incelendiğinde fen lisesi 583 puan ile OECD genel ortalamasının da oldukça üzerindedir. Performans açısından fen liselerini sosyal bilimler liselerinin takip ettiği gözlenmiştir. Okullarımız arasında başarı farkları bulunmakla beraber Bakanlığımız okullar arasındaki başarı farklarını en aza indirmek adına önemli adımlar atmaktadır. Burada bir diğer önemli gösterge de geçen yıllara kıyasla okullar arasındaki başarı farkının kısmen azalmış olmasıdır. Elbette bu farkın daha da azalması hedeflenmektedir. 2015'ten 2018'e performans artışı ve yeterlik düzeylerindeki iyileşme sosyal bilimler liseleri dışında tüm okul türlerimizde gözlenmiştir. Örneğin fen lisesi öğrencilerinin ortalaması 583 puan olup 2015 PISA uygulamasından 59 puan daha yüksektir.

MATEMATİK YETERLİLİK DÜZEYİ

Matematik okuryazarlığına geldiğimizde, 2003 yılında 423 puanla başlayan Türkiye'nin ortalama matematik okuryazarlık puanının on beş yıllık aradan sonra 454 puana çıktığı görülmektedir. Okuma becerisinde olduğu gibi matematik okuryazarlığında da OECD ortalamasının kısmen altında bulunmakla birlikte 1. düzey ve bu düzeyin altındaki öğrenci oranları azalırken daha yüksek performans ifade eden 3, 4 ve 5. düzeydeki öğrenci oranlarında artış gözlenmiştir. Okuma becerisine benzer olarak matematik okuryazarlığı başarısında da fen liselerini; sosyal bilimler, Anadolu, mesleki ve teknik Anadolu ve Anadolu imam hatip liseleri takip etmektedir. Fen liseleri matematikte OECD ortalamasından çok daha yüksek performans göstermektedir. Okullarımız arasında başarı farkını en aza indirmek için diğer projelerle birlikte matematik seferberliği gibi somut adımlar atılmaktadır. Son olarak çok programlı Anadolu Lisesi ve sosyal bilimler liselerinde kısmi puan düşüşü bulunmakla birlikte diğer tüm okul türlerinde önemli puan artışları gözlenmiştir.

FEN OKURYAZARLIĞI

Fen okuryazarlığı alanında 2006'da Türkiye 424 puanla başladığı süreçteki performansını 2012'ye kadar önemli ölçüde artırmış, 2018'de de artırmaya devam ederek en üst düzeye ulaştırmıştır. Birinci düzey ve altında bulunan yaklaşık %20'lik öğrenci artık daha yüksek düzeylerde yer almaktadır ve bu durum 43 puanlık artışın sonucudur. En başarılı iki okul türü olan fen ve sosyal bilimler liselerini; Anadolu, mesleki ve teknik Anadolu ile Anadolu imam hatip liseleri takip etmektedir. Yine 2018 ile beraber sosyal bilimler haricindeki tüm liselerde önemli puan artışları gözlenmiştir. Özetle Türkiye, matematik ve fen okuryazarlığında son 15 yılda, katıldığı tüm PISA uygulamalarından en yüksek ortalama puana ulaşmış; okuma becerilerinde ise performansını önemli ölçüde artırmıştır.

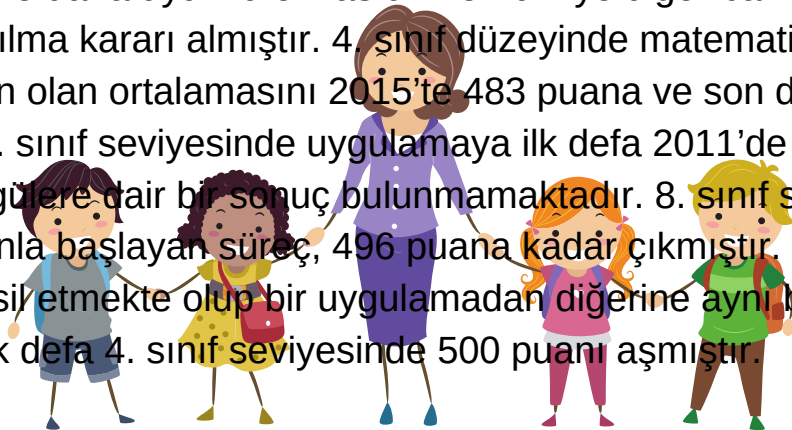




@gunes_ogretmenle_etkinlikler

TIMSS KAPSAMI VE SONUÇLARI(Dr. Hayri Eren SUNA)

TIMSS, IEA tarafından dörder yıllık döngüler şeklinde gerçekleştirilmektedir. PISA ile TIMSS arasındaki en temel fark, PISA'da zorunlu eğitim dönemini tamamlayan öğrencilerin iş gücü piyasası tarafından ve ekonomik kalkınma açısından istenen, uygulamaya dönük okuryazarlık becerilerinin hangi seviyede olduğu değerlendirilmekte iken TIMSS'te eğitim programı (curriculum based) ile ilişkili becerilerin değerlendirilmesidir. TIMSS'te 4 ve 8. sınıf düzeyindeki öğrencilerin ilgili dönemin eğitim programlarından beklenen özelliklere hangi düzeyde sahip olduklarına dair matematik ve fen alanlarında ayrı ayrı uygulamalar yapılmakta ve ilgili eğitim programlarında öğrencilerin okul başarısına dair daha doğrudan çıktılar elde edilmektedir. **TIMSS, ilk defa 1995 yılında uygulanmaya başlanmış olup 2019'da araştırmanın yedinci döngüsü tamamlanmıştır. Türkiye TIMSS'e 1999 yılında ilk defa 8. sınıf düzeyinde katılım gösterirken 2011 yılından bu yana son üç döngüye hem 8. sınıf hem de 4. sınıf düzeyinde katılmıştır.** TIMSS'in son döngüsü olan 2019 döngüsünde tüm dünyadan 580.000 öğrenci, 310.000 veli, 19.000 yönetici ve 52.000 öğretmenden veri toplanmıştır. TIMSS'te tabakalı örnekleme yöntemi kullanılmaktadır: Türkiye'deki tüm okulların listesi uluslararası merkeze gönderilmekte bu merkez iki ayrı sınıf düzeyi için (4 ve 8. sınıf) önce okulları tesadüfen seçmektedir. Burada PISA'da da kullanılan İBB-düzey-1 bölgeleri kullanılmaktadır. Ardından seçilen okullar içerisinde tesadüfi olarak şubeler seçilmekte ve örneklem seçim süreci tamamen bağımsız ve yansız bir şekilde uluslararası merkez tarafından gerçekleştirilmektedir. Uygulama sürecine ise öğrenciler matematik veya fen testi ile başlamaktadır. Sınav yaklaşık 36-45 dakika sürmekte, ardından 15-30 dakika arası mola verilmektedir. Diğer testin uygulaması tamamlandıktan sonra 5-15 dakikalık bir anket uygulaması gerçekleştirilmektedir. Sonuç olarak TIMSS asıl uygulaması yaklaşık yarım gün içerisinde tamamlanmaktadır. Millî Eğitim Bakanlığının 2017 yılında aldığı kararla Türkiye 2019 yılında ilk defa 4. sınıf düzeyindeki uygulamaya 5. sınıf örnekleme ile katılmıştır. Birçok katılımcı ülkede okul öncesi eğitim zorunlu eğitime dâhil olduğu ya da okul öncesi eğitim katılım oranları çok yüksek olduğu için genel ortalamaya göre ülkemizdeki 4. sınıf öğrencilerinin yaş ortalaması biraz düşük kalmaktaydı. Değişikliğin bir ikincisi sebebi ise oluşturulan değerlendirme çerçevesinin Türkiye'nin 5. sınıf eğitim programı ile daha uyumlu olmasıdır ve Türkiye diğer bazı ülkelerle beraber 5. sınıf örnekleme ile katılma kararı almıştır. 4. sınıf düzeyinde matematik alanında Türkiye 2011 yılında 469 puan olan ortalamasını 2015'te 483 puana ve son döngüde 523 puana çıkarmıştır. Türkiye 4. sınıf seviyesinde uygulamaya ilk defa 2011'de katılmaya başladığı için daha önceki döngülere dair bir sonuç bulunmamaktadır. 8. sınıf seviyesinde matematikte 429 puanla başlayan süreç, 496 puana kadar çıkmıştır. TIMSS'te 500 puan önemli bir ölçütü temsil etmekte olup bir uygulamadan diğerine aynı beceriyi ifade etmektedir. Türkiye ilk defa 4. sınıf seviyesinde 500 puanı aşmıştır.





@gunes_ogretmenle_etkinlikler

4. VE 8. SINIF SINAV SONUÇLARI(PİSA, TIMSS,PIRLS)

Matematik yeterlik düzeyleri 4. sınıflar için incelendiğinde 2011 yılında öğrencilerimizin %4'ü, 2015'te %5'i ileri düzey yeterliğe sahipken son uygulamada bu oran %15'e çıkmıştır. 2011'de %7, 2015'te %6 olan ileri düzeydeki öğrenci oranlarımız ise son uygulamada %12'ye çıkmıştır. Alt düzey altı öğrencilerimizin oranı 4. sınıfta %23'ten %12'ye, 8. sınıf düzeyinde %33'ten %20'ye düşmüştür. Fen bilimlerinde de çok benzer bir artışın olduğunu ve bu artışın özellikle 2019'da 4. sınıfta 463 puandan 526 puana ulaştığını ve puan artışının özellikle 2019 uygulamasında ivme kazandığı görülmektedir. 1999'da 433 ile başlayan ortalama puan, bugün 500 ortalama ölçek noktasını da geçip 515 puana ulaşmıştır. Bu anlamda PISA'ya benzer şekilde TIMSS sonuçlarımızda da hem 4 hem de 8. sınıfta matematik ve fen alanlarında önemli iyileşmeler söz konusudur. Fen bilimleri alanında yeterlik dağılımları incelendiğinde ileri düzey seviyesindeki öğrenci oranları 2011'de %3'e, 2015'te %4'e ve son uygulamada %12'ye çıkmıştır. 8. sınıf 114 düzeyinde ise %8'den %13'e çıkmıştır. Alt düzeyin altında kalan öğrenci oranları ise 2014'te 4. sınıf düzeyinde %24'ten %18'e ve en son uygulamada %10'a kadar düşmüştür. 8. sınıfta ise 2011'de %21'den 2015'te %17'ye ve en son uygulamada da %12'ye düşmüştür. Bu sonuçlar, alt düzeylerdeki öğrenci oranlarımızın önemli ölçüde düştüğünü göstermesi açısından son derece önemlidir. PISA'da önemli ölçüde puanlarımız artmış olmakla birlikte OECD ortalamasına henüz ulaşamamak da TIMSS'in eğitim programıyla daha ilişkili olması sebebiyle müfredata dayalı değerlendirmelerde öğrencilerimiz daha yüksek performans göstermektedir. Bununla birlikte öğrencilerimizin daha üst düzey bilişsel becerilerde, iş gücü piyasasının talep ettiği daha karmaşık beceriler diyebileceğimiz becerilerde gelişim göstermeleri beklenen daha fazla alan bulunmaktadır. PIRLS de IEA tarafından uygulanmakta olup okuma boyutunu içermesi sebebiyle TIMSS'i tamamlar nitelikte bir çalışmadır. PIRLS, öğrencilerimizin okuma becerilerine ve okuma düzeylerine dair önemli çıktılar sağlamaktadır. Türkiye PIRLS'e ilk defa 2001 yılında katılmış, 20 yıllık bir aradan sonra alınan kararla 2021 yılında tekrar dâhil olmuştur. Bu uygulamaya ilişkin sonuçlar 2022 yılının sonunda açıklanacaktır.



OECD SOSYAL VE DUYGUSAL BECERİLER ARAŞTIRMASI

Tüm dünyada sosyal ve duygusal becerilerin önemi ile birlikte eğitim sistemlerindeki ağırlığının artmasıyla beraber 2017 yılında OECD tamamen sosyal ve duygusal beceriler odaklı yeni bir araştırma başlatmıştır. Türkiye de İstanbul'la bu çalışmaya katılmıştır. Öncelikle bu çalışma neden önemlidir ve nasıl bir fark yaratmaktadır? Birinci olarak tümüyle sosyal ve duygusal beceriler dikkate alınıyor ki bu diğer tüm izleme araştırmalarından en önemli farkıdır. İkincisi 10 ve 15 yaş grupları ayrı ayrı ele alınmaktadır ki duygusal özellikler, sosyal ve duygusal özellikler yaşa bağlı olarak, özellikle de ergenlik dönemine bağlı olarak önemli değişimler gösterebilmektedir. Ergenlik sürecini de içerecek şekilde iki ayrı yaş düzeyinin dikkate alınması bu becerilerin nasıl değiştiği hakkında önemli çıkarımlar sağlamaktadır. Üçüncüsü ölçümler, dünyanın farklı bölgelerinde kabul gören beş faktörlü kurama dayalı olarak (Big five modeli) yapılmaktadır.

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

- 1) Açık fikirlilik (yaratıcılık, hoşgörü, merak),
- 2) İş birliği (empati, iş birliği, güven),
- 3) Duygu düzenleme (iyimserlik, strese dayanıklılık, duygu düzenleme),
- 4) görev performansı (sebat, öz denetim, sorumluluk)
- 5) Başkalarıyla etkileşimde olma (sosyallik, girişkenlik, enerji) olarak tanımlanmıştır.

Her bir alanı oluşturan alanlar ilişkide ve temelde hepsi kendi içerisinde sosyal dayanağı olan sosyal beceri olarak tanımlanan özelliklerdir. Bir de birleşik beceriler diye tanımlanan ve bu beşliye dâhil olmayan ama onunla ilişkili olup sosyal beceriler açısından önemli görülen iki ayrı beceri ise **başarma motivasyonu** ve **öz yeterlik algısıdır**.



YAŞA VE CİNSİYETE GÖRE SOSYAL VE DUYGUSAL BECERİLER

Çalışmaya Türkiye, 10 yaş grubundan yaklaşık 2.700 ve 15 yaş grubundan ise yaklaşık 3.200 öğrenci ile katılmıştır. Paylaşılan temel bulgular tüm katılımcılardan elde edilen genel bulgular olmakla birlikte Türkiye'nin bulguları genel bulgularla benzerlik göstermektedir. Bu çalışmanın enteresan sonuçlarından biri 10 yaş grubunun sosyal ve duygusal beceri puanlarının 15 yaş grubundan daha yüksek olmasıdır. Şekilde üst tarafa doğru olan çizgiler 15 yaş grubunun puanlarının yüksek olduğunu ifade ederken aşağıya doğru giden çizgiler 10 yaş grubunun sosyal ve duygusal beceri puanlarının daha yüksek olduğunu ifade etmektedir. Buna göre neredeyse tüm becerilerde 10 yaş grubu 15 yaş grubundan daha yüksek puanlara sahiptir. 15 yaş grubunda ergenliğin etkilerinin yoğun görüldüğü düşünüldüğünde sosyal ve duygusal beceriler üzerinde 10 yaştan 15 yaşa geçerken önemli bir düşüş olduğu aşikârdır. Strese dayanıklılık, duygu kontrolü ve enerji gibi bazı özelliklerde erkek öğrencilerin puanları yüksekken sorumluluk, empati, birlikte çalışma, hoşgörü ve merak gibi becerilerde kız öğrencilerin puanları daha yüksektir. 10 yaştan 15 yaşa geçişte yön değiştiren bazı özellikler bulunmaktadır. Örneğin öz denetim puanının 10 yaş grubundaki kız öğrencilerde, 15 yaş grubunda ise erkek öğrencilerde daha yüksek olduğu görülmektedir. Benzer durum öz yeterlik puanı için de geçerlidir. Başarma motivasyonu ise her iki yaş düzeyinde de kız öğrencilerin daha yüksek puana sahip olduğu bir özelliktir. Özetle çıkarmamız gereken en önemli sonuç, sosyal ve duygusal beceriler anlamında kız ve erkek öğrencilerin önemli farklılıklara sahip olduğu ve bu becerileri geliştirmeye dönük her türlü müdahalede cinsiyetler arası farkların dikkate alınması gerekliliğidir.

SOSYO EKONOMİK DURUM

Sosyoekonomik düzey, ailenin sahip olduğu eğitim düzeyi ve gelirden öğrencinin evde sahip olduğu eğitim imkânlarına, okuma imkânlarına hatta fiziksel imkânlara (keni odası olup olmaması, internet bağlantısına sahip olup olmaması) kadar her türlü eğitim ve imkâna dayalı fırsatı kendi içinde barındıran yapı olarak tanımlanmaktadır.

sosyoekonomik açıdan avantajlı öğrencilerin daha yüksek sosyal ve duygusal beceri puanlarına sahip olduğu gözlenmiştir. Ölçülen bütün özelliklerde sosyoekonomik açıdan avantajlı öğrenciler daha yüksek puanlar almıştır. Özellikle yaratıcılık, empati, öz yeterlik gibi becerilerde puanları daha yüksektir. Bu durum sosyoekonomik açıdan dezavantajlı öğrencilerin sosyal ve duygusal açıdan daha fazla desteklenmesi gerektiğini açıkça ortaya koymaktadır. Sosyal ve duygusal beceriler bağlamında bireyin öz yeterliğini geliştirmek, her türlü konuda onu daha yeterli hissettirmek veya kendine yönelik yeterlik inancını artırmak önemlidir ama bunlar akademik başarı ile ne kadar ilişkilidir? Buradaki özellik ön plana çıkıyor: Birinci özellik olan merak, hem okuma hem de matematik becerileriyle anlamlı bir ilişkiye sahiptir. İkincisi de sebat, yani aynı konu üzerinde öğrenene kadar sabırla çalışmaktır. Bu iki özellik aynı grupta akademik başarıyla doğrudan ilişkili görünmektedir. 15 yaş grubu incelenecek olursa yine merak ve sebatı ek olarak kendini güvende hissetme, sorumluluk bilincine sahip olma ve yaratıcılık özelliği de farklı başarı göstergeleri ile ilişki göstermektedir. Bu becerilerin desteklenmesi akademik başarıyı da yükseltme potansiyeline sahiptir. Öğrencileri desteklemek adına bu becerilerini geliştirmek için neler yapılabilir? Çalışmalar, spor etkinliklerine katılımın yaratıcılık ve merak puanlarını önemli ölçüde artırdığını göstermektedir. Örneğin müfredat dışı spor aktivitelerine katılan öğrencilerin katılmayanlara göre yaratıcılık puanları %17 ile %16 arasında daha yüksek, keza merak puanları da %10 ile %15 arasında yaş düzeyine bağlı olarak daha yüksektir. İstanbul'da da bu oranların uluslararası ortalamadan daha yüksek olduğu gözlenmiştir. Sanat etkinlikleri de öğrencilerin merak ve yaratıcılık becerilerine önemli katkı sağlayan etkinliklerdir. Sanat etkinliklerine katılan öğrencilerin yaratıcılık puanları %23 ile %30 arasında, merak puanları ise yaklaşık %15 daha yüksektir. Müfredat dışı sanat ve spor etkinliklerini desteklemek bu öğrencilerin akademik becerileri ile ilişkili olan sosyal ve duygusal becerilerini artırmaktadır.



OKUL İKLİMİ ÇIKTILARI

Okul iklimi ile ilgili çıktılar ise rekabetçi okul iklimi, veli beklentileri ve öğretmen beklentilerini kapsamaktadır. Akademik çalışmalardan, öğrencilerden beklenti arttıkça (beklentinin gerçekçi olması koşuluyla) öğrencilerin akademik performansları ve kendi öz yeterlik algılarının arttığı bilinmektedir. Veli beklentileri ve öğretmen beklentileri arttıkça öğrencinin psikolojik iyilik hâli ve sınav kaygısı da artmakta ancak bir taraftan da psikolojik kaygıları ve yaşam doyumları kısmen artmaktadır. Diğer taraftan rekabetçi okul iklimi ve okul içi rekabeti destekleyen politikalar arttıkça öğrencilerin özellikle psikolojik iyilik hâlinde önemli eksikler olduğu görülmektedir. İstanbul'un yaşam doyumu açısından 10 yaş grubunda uluslararası ortalamaya oldukça yakın ama 15 yaş grubunda uluslararası ortalamanın altında olduğu gözlenmiştir. Tüm katılımcı şehirlerde 10 yaştan 15 yaşa gelindiğinde öğrencilerin yaşam doyumu azalmaktadır. Bu nedenle öğrencilere yaşam doyumu sağlayacak, onları kabul edecek olumlu okul iklimi yaratmak son derece önemlidir. Okullarımızda zorbalığa maruz kalma, büyük oranda öğretmen ve öğrenci ilişkisi bağlamında çözülebilmektedir. Öğretmen ve öğrenci ilişkisi arttıkça, daha etkileşimli hâle geldikçe zorbalığa maruz kalma durumu azalmaktadır. Öğretmen ve öğrenci ilişkisi arttıkça okula aidiyet de artmaktadır. İstanbul, katılımcı şehirler arasında öğrencileri okula aidiyet oranı yüksek şehirlerden biridir. Ulusal izleme çalışmaları kapsamında Akademik Becerilerin İzlenmesi ve Değerlendirilmesi (ABİDE) araştırması, üst düzey bilişsel becerilere öğrencilerimizin ne kadar sahip olduklarını incelenmekte; Türkçe-Matematik-Fen Öğrenci Başarı İzleme Araştırması (TMF-ÖBA) ise öğrencilerimizin eğitim programına bağlı kazanımlara ulaşma düzeylerini belilemek üzere yapılmaktadır. Son olarak 2019 yılında ilk defa tamamen bilgisayar ortamında gerçekleştirilen okuma, dinleme, yazma ve konuşma olmak üzere dört temel becerinin hepsini kapsayan Türkçe dil becerilerinin izlenmesi araştırması yapılmış ve yaygınlaştırılmaktadır.