

UZMAN ÖĞRETMENLİK VE YETİŞTİRME PROGRAMI NOTLARI

MODÜL 2 ÖLÇME VE DEĞERLENDİRME



HÜSEYİN KUVVETLİ
2022



1. ÖLÇME VE DEĞERLENDİRMENİN TEMEL KAVRAMLARI

1.1. EĞİTİM SİSTEMİNDE ÖLÇME VE DEĞERLENDİRMENİN YERİ

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

Bir bütüne sistem diyebilmek için girdi, süreç ve çıktı öğelerinin olması gerekir.

Eğitim de bir sistemdir çünkü birincisi gerçekleştirmek istediği hedefler bulunmaktadır.

Sistemin hedefleri

- 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. Süreç, öğretmenler, okul yönetimi, il yöneticileri ülke genelinde belirli aralıklarla kontrol edilmektedir.
- 3. Çıktı:** Öğrencilerin sahip olduğu niteliklerdir.

1.2. YIRMI BİRİNCİ YÜZYIL DEĞİŞEN EĞİTİM PARADİGMASI:

ABD'de 1930-1950'lerde okullardan "eğitim fabrikaları" olarak söz ediliyordu.

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 "toplumun bireylerden beklediği görev ve sorumluluklar doğrultusunda gerekli bilgi ve donanıma sahip olma" olarak tanımlandı.

Tablo 1. Değişen üretim paradigması

Sanaui Paradigması	Bilgi ve İletişim Teknolojisi Paradigması
Enerji verim	Bilgi verim
Çizim bürolarında yapılan tasarım	Bilgisayar destekli tasarım
Ardışık tasarım ve üretim	Eş zamanlı mühendislik
Standart tasarım	Sipariş üzerine tasarım
Tahsis edilmiş tesis donanım	Esnek üretim sistemleri
Ötomasuon	Sistemmasuon
Tek firma	Ağ yapısı
Hiyerarşik yapılar	Düz yapılar
Bölmüştü	Bütünlük
Servisi olan üretim	Ürünleri olan servis
Merkezi yapı	Dağılım yapı
Özgün beceri	Çoklu beceri
Hükümet kontrolü	Hükümetin bilgi, eş güdümlü ve üretimi yönlendirmesi
Planlama	Vizyon

Tablo 2. Değişen üretim paradigması

Sanaui 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

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

Eğitimde denetleme (kontrol) ölçme ve değerlendirme ile olanaklıdır.

Bireylerden beklenen becerilerin karmaşıklığı arttıkça ölçme modellerini revize etme gereksinimi ortaya çıkmaktadır.

1.3. ÖLÇME, ÖLÇÜT VE DEĞERLENDİRMENİN KAVRAMSAL TEMELLERİ:

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

Örneğin

- Kerim Temel Yeterlilik Testinden (TYT) 450 puan aldı.

Ölçmeye temel olan durum "fark"tır. Bir anlamda ölçmenin farktan doğduğu ifade edilebilir.

Eğitimde akademik başarı dışında yetenek, ilgi, tutum, özel gereksinim gereken alanlar vb. özellikler de ölçmeye konu olur. Eğitim sisteminde ölçme olmadan denetleme ve kontrol mekanizmasını işletmek mümkün değildir.

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

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

Değerlendirmenin basamakları:

- Ölçme ⇒ Ölçme sonucu (ölçüm),
- Ölçüt ve
- Karardır.

Örneğin

- Kerim gıda mühendisliği programına yerleştirilmiştir / yerleştirilememiştir

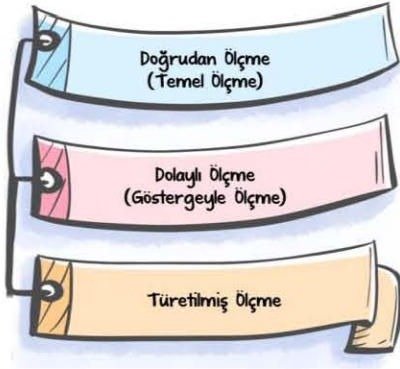
NOT



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

ÖLÇME	ÖLÇÜT	DEĞERLENDİRME
İleriş Yabancı Dil Bilgisi Seviye Tespit Sınavından 50 almıştır.	50	Yüksek Lisansa başvuru yapabilir
İleriş Yabancı Dil Bilgisi Seviye Tespit Sınavından 50 almıştır.	55	Doktora başvuru yapamaz

ÖLÇME TÜRLERİ



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,

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,

Türetilmiş Ölçme: Ölçülme 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),

ÖLÇÜT TÜRLERİ



Mutlak Ölçüt (Kriter Referanslı Ölçüt): Eğer ölçüt ölçme işleminin yapıldığı grubun özelliklerinden bağımsız olarak belirleniyorsa "mutlak"tır. Mutlak ölçüt genellikle ölçme işleminin öncesinde ilan edilir.

Örneğin dersten geçme notu 50'dir, doktora başvurusu için ALES'ten en az 60 almak gerekir...

Bağılı Ölçüt (Norm Referanslı Ölçüt): Eğer ölçüt ölçme işleminin yapıldığı grubun belirli bir özelliğine / normuna dayalı olarak belirleniyorsa "bağılı"dır.

Bağılı ölçüt ancak ölçme işlemi sonrasında belirlenebilir.

Örneğin aritmetik ortalamaya / ortancaya / moda denk ve üzerinde puan alan geçer,

NOT



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

Bağılı ölçütün temel alındığı ölçme işlemlerinde bir öğrencinin başarısı, başka bir öğrencinin değerlendirmesini (geçme / kalma) etkileyebilir.

NOT

Eğer değerlendirme mutlak ölçüte göre yapılıyorsa "mutlak değerlendirme", bağılı ölçüte göre yapılıyorsa "bağılı değerlendirme" adını alır.

Örneğin

- Selim fen lisesine yerleştirilmiştir. → Bağılı değerlendirme
- Güzin lisans programları için tercih yapabilecektir. → Mutlak değerlendirme

NOT

Eğitsel kararların verilmesinde mutlak ve bağılı ölçütler bir arada da kullanılabilir.

Örneğin

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

NOT

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

Örneğin tek ders sınavları vb.

NOT

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

Örneğin LGS, YKS, KPSS, TUS vb.

NOT

Muafiyet sınavlarında mutlak değerlendirme yapılması gerekir.

Örneğin yabancı dil muafiyet sınavları vb.

AMACA GÖRE DEĞERLENDİRME TÜRLERİ

1. Tanıma-Yerleştirmeye Yönelik Değerlendirme (Diyagnostik)
 - a) Sınıf içi ölçme ve değerlendirme
 - b) Okul / kurum geneli ölçme ve değerlendirme
2. Biçimlendirme-Yetiştirmeye Yönelik Değerlendirme (Formatif)
3. Değer Biçmeye / Düzey Belirlemeye Yönelik Değerlendirme (Summatif)
4. Rehberlik Amaçlı Değerlendirme
 - a) Özel eğitim
 - b) Mesleki rehberlik
5. Program Değerlendirme

1. Tanıma-Yerleştirmeye Yönelik Değerlendirme (Diyagnostik)

- a) Sınıf içi ölçme ve değerlendirme
- b) Okul / kurum geneli ölçme ve değerlendirme

a) Sınıf içi ölçme ve değerlendirme:

Amaç öğrencinin ders bağlamındaki hazırbulunuşluluk düzeyini belirlemektir.

Amaç not vermek değildir.

Bu amaçla uygulanan ölçme araçlarına da "hazırbulunuşluluk testi" adı verilir.

b) Okul / kurum geneli ölçme ve değerlendirme:

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.

2. Biçimlendirme-Yetiştirmeye Yönelik Değerlendirme (Formatif)

Belirli bir konu / ünite vb. sonunda öğrencilerin öğretime konu olan davranışların ne kadarını kazandığını, diğer bir deyişle ünitedeki öğrenme eksiklerini belirlemek amacıyla yapılır.

NOT



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 (Summatif Değerlendirme):

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

Bu amaçla uygulanan testlere "alanda erişim testi" denir.

Örneğin; bitirme sınavları, sertifika sınavları, üniversitedeki vize ve finaller...

4. Rehberlik Amaçlı Değerlendirme:

- a) Özel eğitim
- b) Mesleki rehberlik

a) Özel eğitim:

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

Eğitsel değerlendirme bağlamında RAM'lar bu konudaki tek yetkilidir.

b) 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.

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.

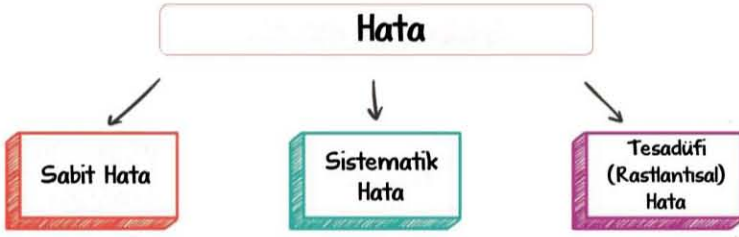
Ölçme aracında bulunması gereken psikometrik nitelikler:

Geçerlilik → Güvenilirlik → Kullanışlılık

Geçerlilik: ölçmek istenilen özelliğin, başka özelliklerle karıştırılmadan, doğru ve tam olarak ölçülebilmesidir.

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

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



1- Sabit Hata:

Miktarı ölçmeden ölçmeye değişmeyen, diğer bir deyişle her ölçme işlemine aynı miktarda karışan hatalardır.

Örneğin marketteki terazi, üzerinde herhangi bir nesne yokken terazi -120 g gösteriyorsa ne tartılırsa tartılsın 120 g eksik ölçülecektir.

2- Sistematik Hata:

Ölçülen büyüklüğe, öğretmene ya da ölçme koşullarına göre miktarı değişen hatalardır.

Örneğin Bir öğretmen, sınavında yazısı kötü olandan puan kırılırsa yine karışan hata sistematik olacaktır.

3- Tesadüfi (Rastlantısal) Hata:

Şansa ortaya çıkan ne yönde ve ne ölçüde karıştığı genellikle bilinmeyen hatalardır.

- Ölçme işlemi yapan kişiden/öğretmenden kaynaklanan hata
Ölçme işlemi sürecinde
Ölçme işlemi sonrasında
- Ölçme aracından kaynaklanan hata
- Bireyden / öğrenciden kaynaklanan hata
- Fiziksel ortamdan kaynaklanan hata

a) 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, öncelik-sonralık yanlılığı, maddi hata vb.

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

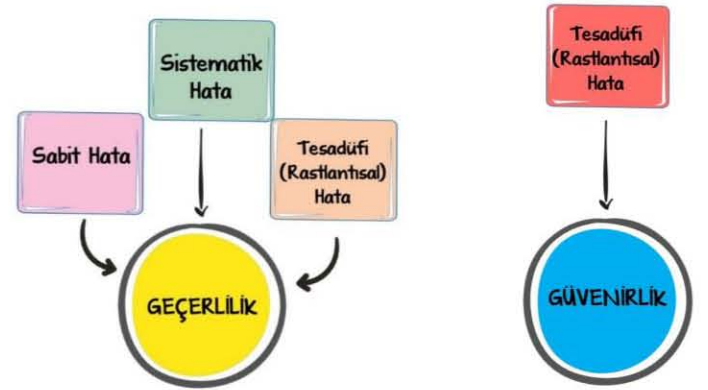
c) Bireyden / öğrenciden kaynaklanan hata: Bireylerin ölçme işlemi sürecinde içinde buldukları fiziksel, fizyolojik ve psikolojik durumlar test puanına etki eder.

d) Fiziksel ortamdan kaynaklanan hata: Ölçme işleminin gerçekleştiği fiziksel ortama ilişkin bazı etmenler de ölçme sonuçlarına hata karıştırabilir.

Sıcaklık, ışık, ses, koku, görsel uyarıcıların fazlalığı, ergonomi vb. etmenler hata miktarını artırabilir.

GEÇERLİLİK, GÜVENİRLİK VE HATA İLİŞKİSİ

Geçerlilik tüm hata kaynaklarından etkilenirken klasik test kuramına göre güvenilirlik yalnızca tesadüfi hatalardan etkilenir.



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

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 değişebilir:

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

KORELASYON

Korelasyon, 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.

- 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 zekâ arasında bir ilişki var mıdır?

NOT

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

NOT

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



1- KAPSAM GEÇERLİLİĞİ:

Kapsam geçerliliği özellikle başarı testlerinde aranan bir geçerlilik sorgulamasıdır.

Kapsam geçerliliği bir testin ölçülmek istenen davranışları ne derece kapsadığıyla ilgilidir.

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.

Kapsam geçerliliği sorgulama yöntemleri:

Mantıksal / rasyonel yöntemler: Bu yöntemler belirtke tablosu hazırlanması ve uzman görüşüne başvurulmasıdır.

Belirtke tablosunun hazırlanması: Bir kapsam geçerliliği sorgulamasında öncelikle ölçmeye konu olan kapsam dâhilinde davranışların belirlenmesi gerekir.

Uzman görüşüne başvurulması: Uygulamada uzman ile kastedilen öncelikle ölçme ve değerlendirme tekniklerini de bilen bir alan uzmanıdır.

İstatistiksel 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.

2-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.

3- YORDAMA GEÇERLİLİĞİ:

Yordama, tahmin demektir ancak her tahmin yordama değildir. Bir tahminin yordama olabilmesi için elde geçerli ve güvenilir bir veri olması ve bu verinin sınılanabilir, sayısal nitelikte, belirli analizlere tabi tutuluyor olması gerekiyor.

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:

Alanyazında **hâlihazır geçerlilik, benzer ölçekler geçerliliği, uygunluk geçerliliği** adı ile de anılmaktadır.

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

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.

Hüseyin KUVVETLİ 2022

GÜVENİLİRLİK

Güvenilir bir ölçme aracı, aynı özelliklere 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 almaktadır.

Meb çalışma kitabı sayfa 83-84 ü inceleyiniz.

**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.

Örneğin genel zihin yetenekleri, kişilik testleri, ilgi envanterleri, tutum ölçekleri vb. gibi testlerin kararlılık bağlamında güvenilirlikleri bu yöntemle hesaplanabilir.

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 başvurulan yöntemler:

- ilk yarı ve son yarı
- tek ve çift
- rastlantsaldır.

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 kümelenecek şekilde yerleştirilmiş ise
- madde sayısı çok fazla ise

Bu yöntemle güvenilirliği tahmin etmede, uygulanmış bir test iki eşdeğer yarıya bölünür ve bireylerin iki yarıdan aldıkları puanlar arasındaki tutarlılık incelenir.

Testin tümüne ilişkin bir güvenilirlik katsayısı **Spearman-Brown formülü** aracılığıyla hesaplanır.

Elde edilen katsayı testin tamamına ilişkin iç tutarlılık bağlamındaki güvenilirlik katsayısı olarak kabul edilir.

KUDER-RICHARDSON 20 VE 21 YÖNTEMLERİ:

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.

CRONBACH ALFA YÖNTEMİ:

Eğer ölçme aracının puanlaması çok kategorili ise KR-2.0 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

- Bir testteki madde sayısı arttıkça birimler küçüldüğünden duyarlılık artar, bu nedenle hata miktarı azalacağından güvenilirlik artar.
- Bir ölçme işleminde genel bir ilke olarak puanlayıcı sayısı arttıkça güvenilirlik artar.
- Bir testin farklı kişiler tarafından puanlanması ya da aynı kişinin farklı zamanlarda verdiği puanlar arasındaki tutarlılığa **puanlama güvenilirliği** adı verilir. Bir testten elde edilen puan, puanlayıcıya ya da zamana göre değişmiyorsa testin güvenilirliği artar.
- KR-2.0 ve KR-2.1 yöntemlerinde bahsedildiği gibi testteki maddeler açısından benzerlik (homojenlik) arttıkça güvenilirliğin artması, diğer taraftan ayrışıklık (heterojenlik) arttıkça güvenilirliğin düşmesi beklenen bir durumdur.
- 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 ($PP = 50$) 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 gerekiyorsa süreçte yönerge vermek, test almaya güdülenme ve hazırbulmuşluğun 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 etmendür.

10. Eğer süreli bir test uygulaması ise sürenin yeterli verilmesi gerekir.

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

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.

Güvenilirliği artıran faktörler geçerliliği de artırır.

3. TEST GELİŞTİRME VE MADDE / SORU TÜRLERİ

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

TEST TÜRLERİ:

Testin 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. Hz Testleri	1. Objektif	1. Standart	1. Performans
2. Grup	2. Süresiz	2. Güç Testleri	2. Subjektif	2. Öğretmen Yapımı	2. Kağıt-Kalem

TEST GELİŞTİRME:

Test geliştirme işlem basamakları:

- Amacın belirlenmesi,
- Kapsamın belirlenmesi ve belirtke tablosunun oluşturulması,
- Denemelik maddelerin / soruların yazılması,
- Maddelerin / soruların gözden geçirilmesi (redaksiyon),
- Denemelik test formunun hazırlanması,
- Testin uygulanması,
- Test ve madde istatistiklerinin hesaplanması,
- Seçilen maddelerden oluşan nihai formun oluşturulması.

Öğretmenlere sınıf içi ölçme ve değerlendirilmelerde uygulanması daha olanaklı bir test planı önerisi:

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

BAŞARININ ÖLÇÜLMESİNDE YÖNTEMLER

ÖLÇME ARAÇ VE YÖNTEMLERİ

GELENEKSEL YÖNTEMLER

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

DESTEKLEYİCİ YÖNTEMLER

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

Seçme Gerektiren Maddeler

- Doğru-Yanlış
- Eşleştirme
- Çoktan Seçmeli

Açık Uçlu Sorular

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

Hüseyin KUVVETLİ 2022

GELENEKSEL ÖLÇME YÖNTEMLERİ:

- Doğru-yanlış maddeleri
- Eşleştirme maddeleri
- Çoktan seçmeli maddeler
- Açık Uçlu Sorular
- Essay (yazılı yoklama)
- Kısa cevaplı sorular
- Cümle tamamlama soruları
- Sözlü yoklama:

1. Doğru-yanlış maddeleri:

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

Genellikle öğrencinin bilimsel gerçekleri, tarihî olayları, kesin yargıları tanıması ve hatırlaması; alguları ve gerçekleri kişisel yargı ve görüşlerden ayırt etmesi ölçülmek istendiğinde kullanışlıdır.

Avantajları:

- Hazırlaması, uygulaması ve puanlaması kolay, diğer bir ifadeyle kullanışlıdır.
- Soru sayısı artırılabilir.
- Sistematik hata karışma olasılığı pek yoktur.

Dezavantajları:

- Şans başarısı olasılığı yüksektir.
- İyi yazılmazsa genellikle üst düzey bilişsel özellikleri ölçemez.
- Öğrenme eksiklerini belirleyemez.
- Yanlış üzerinden öğretim yapılmaz.
- Her derste/konuda kesin yanlış denebilecek durumlar belirlemek zordur.

Doğru-Yanlış Maddeleri Yazım Kuralları:

- Her madde tek ve belirli bir fikri belirtmelidir. Özellikle aynı maddede biri doğru, öteki yanlış iki fikir ifade etmekten kaçınılmalıdır.
- Doğru yanlış maddesi, kesinlikle doğru ya da kesinlikle yanlış olmalıdır. Maddenin doğruluğu ya da yanlışlığı, başka bir açıklamaya gerek kalmadan belirlenebilmelidir.
- Bir maddenin yanlışlığı önemsiz bir ayrıntıda ya da aldatıcı bir noktada olmamalıdır. Bir yargı temelden yanlış olmalıdır.
- Mümkün olduğunca olumsuz ifade kullanılmamalıdır. Özellikle iki olumsuz ifadede kesinlikle kaçınılmalıdır.
- Kanı ifadeleri kendiliklerinden doğru ya da yanlış olmadığından bir kaynağa dayandırılarak verilmelidir. Belli bir kaynağa ya da bir otoriteye dayandırılmadan verilen kanı ifadelerinin yer aldığı maddelerin ayırt etme gücü sıfır ya da negatif olma eğilimindedir.
- Bir maddenin ifadesi kısa, açık ve yalın olmalıdır. Madde, onda sorulan ana fikrin doğruluk ya da yanlışlığının belirlenmesi için gerekli olmayan ayrıntılarla şişirilerek uzatılmamalıdır.
- Okuduğunu anlama süreçlerinin ilk basamağı, metin içerisinde açıkça verilmiş bilgiden doğrudan çıkarım yapmadır. 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.
- Öğrencinin bir maddede ifadenin doğruluğuna ya da yanlışlığına karar vermesi gerektiğinde öğretmen özellikle vurgulamak istediği bir sözcük / terim vs. varsa onun altını çizilebilir.
- 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. 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. 91
- 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 edilmeli, özellikle "ve" gibi bağlaçlara dikkat edilmelidir.
- Olumsuz anlama sahip bir sözcük ya da ifade varsa altı çizilmelidir.
- İpucu vermektan kaçınmak gerekir. Özellikle "asla, daima, hiçbiri, 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.

2. Eşleştirme maddeleri:**Avantajları:**

- Hazırlaması, uygulaması ve puanlaması kolay, diğer bir ifadeyle kullanışlıdır.
- Soru sayısı artırılabilir.
- Sistematik hata karışma olasılığı pek yoktur.

Dezavantajları:

- Şans başarısı olasılığı yüksektir.
- İyi yazılmazsa genellikle üst düzey bilişsel özellikleri ölçemez.

Eşleştirme Maddeleri Yazım Kuralları:

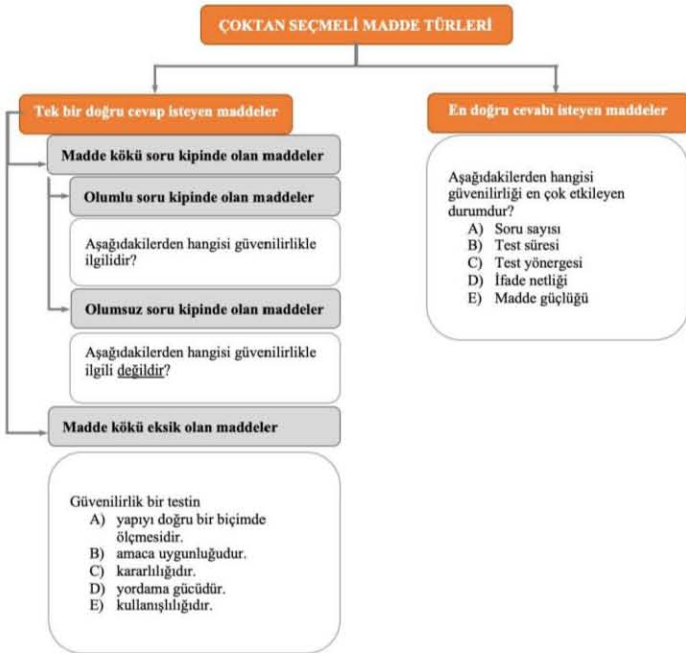
- Her eşleştirme soru grubunda yönerge/açıklama yazılmalıdır. Yönergede eşleştirmeye konu olan kapsam belirtilmeli, öncüllere ve seçeneklere nelerin konulduğuna değinilmelidir. Eşleştirmenin nasıl yapılacağı konusunda bilgi verilmelidir.
- Seçeneklerin kullanılıp kullanılmayacağı, birden fazla kez kullanma durumu belirtilmelidir. Bir eşleştirme takımındaki öncüller ile cevaplar eşit sayıda olmamalıdır. Eğer eşit sayıda olursa öğrenci, hemen ilk bakışta bildiklerini eşleştirecek; geri kalan öncüllerin cevaplarını da tahmin edecektir.
- 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. Diğer bir ifadeyle aynı kapsamdan gelmelidir.
- Uzun ifadeler öncül olarak kullanılmalı ve öncüller sütunu sayfanın sol sütununa, cevaplar sütunu ise sayfanın sağ tarafına yerleştirilmelidir. Böyle olması, cevaplama zaman kaybını önler ve cevabın seçimini kolaylaştırır.
- Bir eşleştirmeli maddeler grubundaki madde sayısı, en az 6, en çok 15 olmalıdır. Madde sayısının 6'dan az olması, salt tahminle doğru cevabın bulunma olasılığını artırır. 15'ten çok madde kullanıldığında ise cevaplayıcıların doğru olarak eşleştirecek ifadeleri ayıklayıp seçmesi çok zaman alır. Üstelik madde sayısı arttıkça öncüller ile cevaplar takımının benzeşikliğini sağlamak giderek güçleşir.

- Maddelerin tümü aynı sayfada bulunmalıdır. Aksi hâlde cevabın seçimini güçleştirir.
- Cevapların seçileceği sütun, bir kelime listesi ise alfabetik sıraya göre (küçükten büyüğe ya da büyükten küçüğe) düzenlenmelidir.
- Bu eşleştirme maddeleri öğrencilerin yalnızca hatırlama düzeyi dışında, kavrama düzeyindeki öğrenmelerinin de yoklanabileceğine ilişkin örnektir. Örneğin grafik / tablo / metin / harita

3. Çoktan seçmeli maddeler:

Çoktan Seçmeli Maddelerin Özellikleri:

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



Çoktan Seçmeli Madde Yazım Kuralları:

- İdeal bir çoktan seçmeli maddede, bilen bir öğrenci, seçenekleri görmesine gerek kalmaksızın doğru cevabı verebilmelidir.
- Öğrencinin sözcük repertuarında bulunmayan, henüz öğrenmediği sözcük, kavram, terim vb. kullanılmamalıdır.
- Bir durumun resim ya da fotoğrafla anlatılabildiği durumda, çokça cümle kurmak yerine görsellerden yararlanılabilir. Ancak maddenin cinsiyet gibi alt gruplarda yanlılık oluşturmamasına dikkat edilmelidir.

- Gereksiz görsel kullanılmamasına da özen gösterilmelidir. Görsel çıkarıldığında anlamca bir kayıp oluşmuyorsa o görsel kullanılmamalıdır.
- Bir maddeyi okuyan bir öğrenci öğretmenin zihninden geçeni keşfetmek zorunda değildir.
- 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. Bir çoktan seçmeli maddeyi zorlaştırmak isteyen öğretmen temelde iki yöntem kullanabilir:
 - Çeldiricileri doğru yanıtla yaklaştırmak (güçlendirmek) ve
 - Davranışı üst düzey becerileri ölçen bir formda hazırlamak.
- Madde kökünde gereksiz yere sözcük kullanılmamalıdır. Zaman zaman öğretmenler kendi doğrularını ya da önemli yerleri vurgularken bu hatayı yapmaktadırlar.
- Seçeneklerde gereksiz sözcük tekrarı kaçınılmalıdır. Tekrar edilen sözcük madde köküne alınabilir.
- Bir maddenin doğru yanıtı, başka bir maddenin kökünde ya da seçeneklerinde bulunmamalıdır.
- 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.
- İpucu vermektan kaçınmak gerekir. Özellikle "bazen, çoğunlukla, genellikle, sıklıkla ya da zaman zaman" vb. ifadelerin doğru olma olasılığı yüksektir. Diğer taraftan "hiçbir zaman, asla, daima, hiçbirini, her zaman, tümü" vb. ifadelerin de yanlış olma olasılığı yüksektir. Bu nedenle bu tür sözcükler seçeneklerde kullanılmamalıdır.
- 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. Özellikle doğru cevabın daha uzun ya da daha kısa olmamasına özen gösterilmelidir.
- Doğru cevaplar teste dağıtılırken dikkat edilmeli, doğru cevapların bir örneği 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 ilkökul 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 bir titizlik ve özen göstermek çok önemlidir. İyi bir madde yazarı; a. madde yazılan alanı çok iyi bilmeli, b. madde yazma teknik ve yöntemlerinden haberdar olmalı, c. maddelerin yazılacağı dili kullanmada becerikli olmalı ve d. testin uygulanacağı öğrencilerin gelişim düzeyini çok iyi bilmelidir.
- Çoktan seçmeli maddelerde seçeneklerin baş harfleri büyük olmalıdır.
- Bir madde kökü ile seçenekleri aynı sayfada olmalıdır. Yarısı başka sayfada, yarısı diğer sayfada olmamalıdır. Bu durum ortak köke dayalı maddeler için de geçerlidir.

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

AÇIK UÇLU SORULAR

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

4. Essay (Yazılı Yoklama):

Yazılı yoklamaların temel özellikleri:

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

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:** Hazırlanması 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 örgütlenme biçimini de yansıtır. Bu ise ölçme sonuçlarındaki sistematik hata miktarını artırabilir.

• Puanlama özneliği belki de en önemli dezavantajdır.

"Essay"de Nesnellik Artırma Yolları:

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

• Bütüncül (holistik) ya da analitik rubrik hazırlanabilir.

5. Kısa cevaplı sorular:

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

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

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

Avantajları:

a. Hazırlanması, uygulaması ve puanlaması kolay; diğer bir ifadeyle kullanışlıdır.

b. Soru sayısı artırılabilir.

c. Sistematik hata karışma olasılığı pek yoktur.

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

6. Cümle tamamlama soruları:

Kısa cevaplı / sınırlı cevap gerektiren soru grubudur.

Bu tür sorularda öğretmen önemli gördüğü bir cümleyi alır, kritik gördüğü bir ya da birden fazla yeri çıkarır ve yerine bir boşluk koyar.

Öğrenciden o cümleyi anlamlı bir biçimde tamamlaması istenir. Cümle tamamlama soruları boşluk doldurma olarak da geçmektedir.

• Hazırlanması oldukça kolaydır. Bu yargı, özellikle kim, ne, nerede, ne zaman? sorularına cevap olabilecek olgusal bilgileri ölçmeye yönelik soruların yazılmasında geçerlidir.

• Cümleler ders kitabından aynen alınmamalı, öğretmen kendi sözcükleriyle soruları yeniden yazmalıdır.

Avantajları:

a. Hazırlanması, uygulaması ve puanlaması kolay; diğer bir ifadeyle kullanışlıdır. b. Soru sayısı artırılabilir. c. Sistematik hata karışma olasılığı pek yoktur.

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

• Cümleler ders kitabından aynen alınmamalı, öğretmen kendi sözcükleriyle soruları yeniden yazmalıdır.

Avantajları:

a. Hazırlanması, uygulaması ve puanlaması kolay; diğer bir ifadeyle kullanışlıdır. b. Soru sayısı artırılabilir. c. Sistematik hata karışma olasılığı pek yoktur.

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

Cümle Tamamlama Soruları Yazım Kuralları:

- Her kısa cevap maddesi, yalnızca tek bir doğru cevabı olacak biçimde yapılandırılmalıdır. Tartışma götürerek, yorumla açık durumların soru yapılmasından kaçınılması ve sorunun ifadesinin açık, net ve anlaşılır olması gerekir.
- Bir maddenin ifadesinde, o maddenin cevabının bulunmasında işe yarayacak ipuçları vermekten kaçınılmalıdır. Doğru cevabın verilmesinde kullanılacak ipucu, maddede verilen gereksiz bir bilgi ya da cümlelerin gramer yapısı olabilir.
- 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 işaretlenmesi 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 yapılmış olmalıdır.
- Herkesçe aynı biçimde algılanacak, öğrencinin "Acaba öğretmenim bununla neyi kastediyor?" sorusunu sormayacağı biçimde soru yazmak gerekir.

7- Sözlü Yoklama:

Dile dayalı becerilerin ölçülmesi için geleneksel yöntemler içindeki tek türdür.

Örneğin sözlü anlatım, diksiyon, yabancı dil dersinde konuşma, solfej ve şarkı söyleme vb.

- Sorular sözlü olarak sorulur ve cevap sözlü olarak verilir.
- Öğretmen ile öğrenci arasında karşılıklı ve devamlı bir etkileşim vardır.
- Bireysel bir test olması nedeniyle her öğrenciye ayrı soru sorma zorunluluğu vardır.
- Bireysel bir test olması nedeniyle tüm öğrencileri test etmek için gereken süre oldukça fazladır.
- Öğrencinin cevaplarını gözden geçirme şansı yoktur.
- Cevaplama çoğu zaman üzerinde düşünmeye ve tasarlama olanak olmadan yapılır.
- 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 yüksektir.
- Sözlü anlatım becerisi iyi olan bir öğrenci, iyi olmayan bir başka öğrenciden daha az biliyor olsa da yüksek puan alabilir.
- Öğrencinin kıyafeti, diksiyonu, hâl ve hareketleri puanlamaya artı ya da eksi yönde etki edebilir.
- Çok fazla soru sorulamayacağı için kapsam geçerliliği problemleri 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 anahtarları hazırlanmalıdır.
- Sınavın yapılması aşamasında öğrencilerin kaygısını artıracak davranışlardan kaçınılmalıdır. Hazırbulunmuşluk artırılmamalıdır.
- Basit olan sorudan başlanmalıdır.
- Kayıt altına alın(a)mıyorsa puanlarına hemen yapılmalıdır.
- Puanla ilişkin dönüt hemen verilmelidir.
- Puanlamada sistematik hata kaynaklarına karşı dikkatli olunmalıdır.



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 çeşitlenmesi yapmaktır ancak özellikle ilköğretimde, çocukların gelişim süreçleri de dikkate alındığında, madde / soru çeşidi sayısını çok artırmamak gerekir.

DESTEKLEYİCİ DEĞERLENDİRME YAKLAŞIMLARI:**1. Portfolyo**

- Süreci yansıtan portfolyolar
- Ürünü yansıtan portfolyolar:

2. Performans değerlendirme

- Öz Değerlendirme
- Akran Değerlendirme
- Rubrik (Dereceli Puanlama Anahtarı)
- Kontrol listeleri
- Dereceleme ölçekleri
- Gözlem Formları

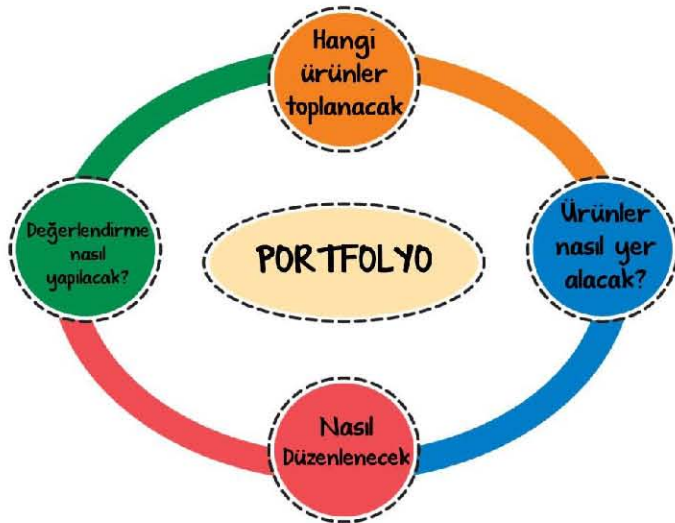
1. Portfolyo

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.
- 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 çalışmalarını, karşılaşılan güçlükleri ve öğrenme ürünlerini içerir.
- Ürünü yansıtan portfolyolar:** Öğrenme sürecinden çok bitmiş görevleri içerir. Öğrencinin en iyi olduğunu düşündüğü çalışmaları içerir.

Portfolyo değerlendirme süreci**Portfolyonun içeriği:**

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

Portfolyo Oluşturma Süreci**TOPLAMA:**

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

SEÇME

- Öğrenci seçimini kendi başına veya öğretmenin rehberliğinde yapabilir.
 - Portfolyonun türüne ve öğretmenin koyduğu koşullara bağlıdır.
- Süreci yansıtan portfolyolarda öğretmenin belirlediği konuları / gelişim alanlarını yansıtan çalışmaların seçilen örnekleri, öğrenme / gelişim sürecini yansıtacak şekilde yer alır.
 - 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

- Portfolyosuna seçtiği her bir çalışmayı niçin seçtiğini açıklar.
- Çalışmalarını yaparken geçirdiği süreci ve bu süreçte öğrendiklerini anlatır.
- 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.

2. PERFORMANS DEĞERLENDİRME



a. Öz Değerlendirme:

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:

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

b. Akran değerlendirme:

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:

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

c. Rubrik (dereceli puanlama anahtarı):

Amaç, öğretmen tarafından ürünün genel izlenimle puanlamasındaki öznelliğini azaltmaktır.

Rubrikler ikiye ayrılır:

- Bütüncül (Holistik) rubrik
- Analitik rubrik

Meb Çalışma kitabı sayfa 101 deki örnekleri inceleyebilirsiniz.

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

d. 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 sergilenen 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.

e. 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.

f. 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 VE DESTEKLEYİCİ YÖNTEMLERİN KARŞILAŞTIRILMASI

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

Hüseyin KUVVETLİ 2022

4. TEST VE MADDE İSTATİSTİKLERİNE GENEL BİR BAKIŞ

1. Test istatistikleri
2. Madde istatistikleri

1. Test istatistikleri

Ölçme işlemi sonucunda bireylerin toplam puanları üzerinden hesaplanan istatistiklerdir.

Eğitimde en sık kullanılan istatistikler

- merkezi eğilim ölçüleri,
- değişkenlik ölçüleri,
- dağılım özellikleri,
- standart puanlar vb.dir.

Merkezi eğilim ölçüleri: aritmetik ortalama, medyan (ortanca) ve moddur. Bu üç istatistiğin de temel işlevi bir puan dizisindeki merkezi bulmaktır.

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.

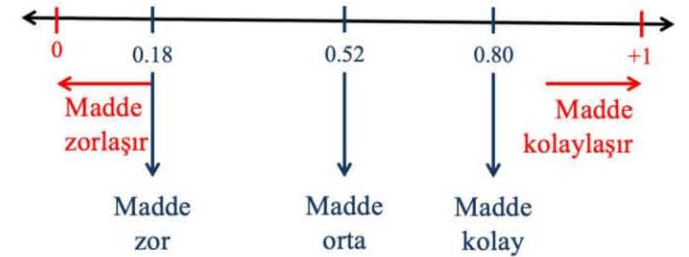
Dağılım özellikleri: Grup hakkında bilgi elde etmemizi sağlayan istatistiklerdir.

2. Madde istatistikleri:

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 ayırt edicilik indeksi ise maddenin bilenle bilmeyeni ayırt edip edemediğine ilişkin bir istatistiktir. -1 ile 1 arasında değer alır ve genellikle alanyazında 0.30 ile 1 arası ayırt edici olarak nitelendirilir.



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

- Kavramlar
- Uluslararası İzleme Çalışmaları
- Günümüzde En Yoğun Katılım Gösterilen Çalışmalar

a. Kavramlar

Başarı (erişi):

Belirli bir konu alanında veya uygulama alanında kazandırılan/kazandırılmaya çalışılan kazanımlara bireylerin ulaşma/erişme düzeyleri.

Konu alanında verilen eğitim sonucunda öğrencilerin beklenen kazanımlara, becerilere ya davranışlara sahip olma düzeyini ifade eden bir kavram.

Sunulan bilgilere, sunulan içeriğe o eğitimi alan bireylerin ne kadar ulaşabildiğine dair bir gösterge, erişim düzeyi.

Beceri:

Öğrencilerin bilgi ve becerilere sahip olduktan sonra bunları ne kadar kullanabildiğine dair bir gösterge olarak değerlendirilmektedir.

Bilgilerin bir araya toplanıp, doğru bilgilerin seçilip ardından seçilen bilgilerin doğru şekilde kullanılması, gerçek ya da tasarlanan bir durumda kullanılabilme yetisi.

Okuryazarlık:

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.

NOT



(PISA)
Uluslararası Öğrenci Değerlendirme Programı

b. Uluslararası İzleme Çalışmaları

Bugün birçok ulusal izleme araştırmasına temel teşkil eden ilk çalışmalardan biri 1964 yılında gerçekleştirilen **uluslararası matematik izleme çalışması**dır.

TIMSS'in de **geliştiricisi** olan kurum (IEA) tarafından gerçekleştirilen bu çalışma 1960'larda **matematik** üzerinde başlamıştır.

Bu çalışmaların popüler olmasının sebeplerinden ilki **karşılaştırılabilir veri ihtiyacıdır**.

Uluslararası Matematik ve Fen Eğilimleri Araştırmasında (TIMSS) matematik ve fen, **Uluslararası Okuma Becerilerinde Gelişim Araştırmasında (PIRLS)** ise okuma becerileri değerlendirilmektedir. İzleme çalışmalarının sağladığı bir katkı da ihtiyaç duyulan bağlamsal ve bütüncül ilişkilere dair bulgular sunmasıdır.

c. Günümüzde En Yoğun Katılım Gösterilen Çalışmalar

NOT



(PISA)
PISA, Ekonomik Kalkınma ve İşbirliği Örgütü (OECD) tarafından geliştirilen ve üç yıllık periyotlarla uygulanan izleme çalışmasıdır.

Ekonomik Kalkınma ve İşbirliği Örgütü (OECD)

Temel amacı tüm üye ülkeler arasındaki ekonomik iş birliklerini ve kalkınmayı destekleyecek programlar oluşturarak birlikte kalkınmayı sağlamak

NOT



(IEA)
Uluslararası Eğitim Başarılarını Değerlendirme Kuruluşu

2. PISA KAPSAMI VE SONUÇLARI

OECD, ekonomik kalkınmanın ve sürdürülebilirliğinin sağlanmasındaki en önemli kaynak olan beşerî sermayeyi üreten faktör olarak eğitimi konumlandırmaktadır.

Eğitimin ekonomik kalkınma üzerindeki belirleyici rolü dolayısıyla eğitim niteliğini detaylı olarak ele alan ve **üç yıllık periyotlarla uygulanan PISA çalışmasını** gerçekleştirmektedir.

PISA, 15 yaş grubunu hedef grup olarak almaktadır.

Bunun nedeni birçok Avrupa ve OECD ülkesinde 15 yaşa tekabül eden dönemin yaklaşık olarak ortaokulun sonuna denk gelmesi ve bu dönemde zorunlu eğitimin tamamlanmasıdır.

Dolayısıyla PISA, zorunlu eğitim dönemini bitiren öğrencilerin henüz iş gücü piyasasına katılmadan ya da eğitimlerine devam etmeden **matematik, fen ve okuma becerileri** alanlarında hangi yeterliklere sahip olduğuna dair bilgi vermektedir.

3. TIMSS KAPSAMI VE SONUÇLARI

TIMSS, IEA tarafından dörder yıllık döngüler şeklinde gerçekleştirilmektedir.

PISA ile TIMSS arasındaki en temel fark,

PISA'da zorunlu eğitim dönemini tamamlayan öğrencilerin iş gücü piyasası tarafından ve ekonomik kalkınma açısından istenen, uygulamaya dönük okuryazarlık becerilerinin hangi seviyede olduğu değerlendirilmekte iken TIMSS'te eğitim programı ile ilişkili becerilerin değerlendirilmesidir.

TIMSS'te 4 ve 8. sınıf düzeyindeki öğrencilerin ilgili dönemin eğitim programlarından beklenen özelliklere hangi düzeyde sahip olduklarına dair matematik ve fen alanlarında ayrı ayrı uygulamalar yapılmakta ve ilgili eğitim programlarında

öğrencilerin okul başarısına dair daha doğrudan çıktılar elde edilmektedir.

TIMSS, ilk defa 1995 yılında uygulanmaya başlanmıştır.

PIRLS 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,

4. OECD SOSYAL VE DUYGUSAL BECERİLER ARAŞTIRMASI

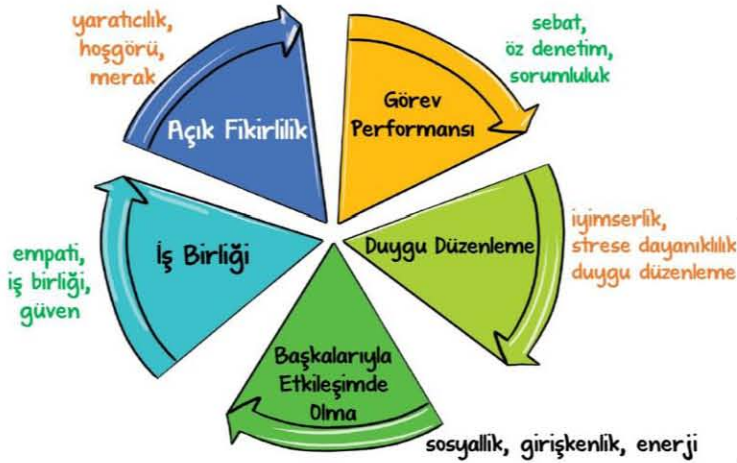
Bu çalışmada;

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 duyuşsal özellikler, sosyal ve duygusal özellikler yaşa bağlı olarak, özellikle de ergenlik dönemine bağlı olarak önemli değişimler gösterebilmektedir.

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

Büyük Beşi Sosyal ve Duygusal Beceriler Modeli



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

açık fikirlilik: (yaratıcılık, hoşgörü, merak),

iş birliği: (empati, iş birliği, güven),

duygu düzenleme: (iyimserlik, strese dayanıklılık, duygu düzenleme),

görev performansı: (sebat, öz denetim, sorumluluk)

başkalarıyla etkileşimde olma: (sosyallik, girişkenlik, enerji)

olarak tanımlanmıştır.

Bu çalışmayı özel kılan diğer bir özelliği de **veri çeşitlenmesinin** (data triangulation) kullanılmasıdır.

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

Bu çalışmaya katılan ülkeler ve şehirler şu şekildedir:

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

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 belirlemek üzere yapılmaktadır.