Click to verify

Ölçüt geçerliliği nedir

Ölçme ve Değerlendirme KPSS ders notlarına Ölçme Araçlarının Özellikleri, Geçerlik ve geçerlik türleri, Güvenirlik ve güvenirlik türleri ve Kullanışlık konuları ile devam ediyoruz.Bir ölçme aracının tesadüfi hatalardan arınık olma derecesidir.Güvenirlik ve tesadüfi hata arasında korelasyonel bir ilişki vardır.Güvenirlik (0 ve +1) aralığında değer alır.+1’e yaklaştıkça güvenirlik artar, 0’a yaklaştıkça güvenirlik azalır.Testteki güvenirlik arttıkça tesadüfi hata miktarı azalır, testteki güvenirlik azalırsa tesadüfi hata miktarı artar.NOT: Eğitimde kullanılan kağıt kalem testleri (yazılı yoklama, çoktan seçmeli, kısa cevaplı test vb.) en az 0.70 olmalıdır.GÜVENİRLİĞİN ANLAMLARI 1. DUYARLILIKBir ölçme aracının birim aralıklarının küçülmesidir.100 soruluk test 50 soruluk testten duyarlıdır.100cm, 1 metreden duyarlıdır.1000 metre, 1 km den duyarlıdır.2. TUTARLILIKBir ölçme aracındaki tüm soruların aynı özelliği ölçmesidir.Tutum ölçeğindeki tüm soruların aynı olması (iç tutarlılık)3. KARARLILIKÖlçme sonuçlarının her seferinde aynı sonucu vermesidir. (PUANLAMA GÜVENİRLİĞİ)Öğretmenin bir sınav kağıdını iki kez okuyup aynı puanı vermesidir.4. OBJEKTİFLİKPuanlama yapan kişinin tarafsızlığıdır.Aynı zamanda ölçme aracının tarafsızlığından da söz edilir. (PUANLAYICI GÜVENİRLİĞİ)İki öğretmenin aynı sınav kağıdına benzer puanlar vermesidir.GÜVENİRLİK BELİRLEME YOLLARI 1. TEST TEKRAR YÖNTEMİBir test, bir grup öğrenciye, iki ayrı zamanda uygulanması sonucunda, her iki uygulamadan elde edilen puanlar arası korelasyon katsayısı test tekrar test güvenirliği anlamına gelir.Bu katsayı aynı zamanda KARARLILIK katsayısıdır.Elde edilen korelasyon 1’e ne kadar yakın olursa ölçme aracının güvenirliği o kadar yüksek olur.Bu yöntemin uygulanması sürecinde hata kaynağı ölçülen özelliğin zaman içinde değişmesidir. Bu değişime neden olan faktör unutma veya yeni öğrenmelerin olmasıdır.İki uygulama arası yeni öğrenmeler veya unutma olursa iki uygulamanın puanları arasındaki korelasyon katsayısı 0’a yakınlaşacağı için güvenirlik azalacaktır. Bu nedenle Test Tekrar Test Yöntemi, daha çok kişilik, yetenek, tutum, ilgi gibi zor değişen özellikleri ölçen testlerde kullanılması tercih edilir.2. PARALEL TESTLER YÖNTEMİ (EŞDEĞER FORMLAR) Her yönüyle aynı özellikte, aynı kapsamda , aynı güçlükte , farklı sorulardan oluşan iki test aynı öğrenci grubuna art arda uygulanır.Her iki testten elde edilen korelasyon katsayısı her bir testin güvenirlik katsayısı anlamına gelir.Bu korelasyon TUTARLILIK ve EŞDEĞERLİK katsayısını verir.Eşdeğerlik katsayısı yüksek çıkarsa; iki formun birbiriyle tutarlı, eşdeğer olduğu, aynı hedefleri ölçtüğü ve soruların birbirine yakın güçlük düzeyinde olduğu söylenebilir.Fakat kademeler ilerledikçe birbirine denk özellikte iki form hazırlamak zordur.KARŞILAŞTIRMAParalel formlar yönteminin, test tekrar test yöntemine göre avantajı iki formun yakın zamanda art arda uygulanması nedeniyle bilgiyi unutma veya yeni öğrenmelerin olması gibi durumlardan etkilenmemektedir.Paralel formlar yönteminin dezavantajı iki formun kapsamı güçlük vb. özellikler bakımından denk olması zordur.Denk iki form hazırlamak emek ve zaman açısından kullanışlı değildir.3. TEST YARILAMA YÖNTEMİ Hazırlanan bir test, bir gruba, bir kez uygulanır.Uygulamanın sonucunda test, iki yarıya bölünür.Daha sonra her iki yarının puanları hesaplanır.Hesaplanan bu korelasyon katsayısı, tek yarı güvenirliğini verir.Daha sonra korelasyon katsayısı SPEARMAN BROWN formülüne yazılarak testin bütününün güvenirliği bulunur.Test yarılama yöntemiyle İÇ TUTARLILIK katsayısı elde edilir.Tek uygulama olduğundan kullanışlığı yüksektir.Bir testin tamamının güvenirliği her zaman yarısından daha fazladır.Testin bütünü için güvenirlik katsayısı formülü;SPEARMAN BROWN: 2*r/1+r şeklindedir. (r: eş yarılar arasında bulunan korelasyondur)ÖRNEK: Bir testin iki yarısının güvenirlik katsayısı 0.50 ise, testin bütününün güvenirliği nedir?2*0,50= 11+0,50= 1,501/1,50= 0,66 olacaktır.4. KR20 – 21 YÖNTEMİ Testteki soruların; birbiriyle uyumunu, homojenliğini, benzerliğini, iç tutarlılığını ölçen,Doğru cevaplara 1, yanlış cevaplara 0 puan verilen,Bir defa uygulanan, güvenirlik belirleme yöntemidir.Testin bütün öğrencilerin %90’ı tarafından cevaplanmış olma şartı vardır.Dereceleme ölçekleri, ağırlıklı puanlama ve düzeltme formülü KR20-21 güvenirlik yönteminde kullanılmaz.NOT: KR20-21 YÖNTEMİ;Maddeler arası uyumu inceler.Eğer maddeler homojense KR20-21 güvenirlik düzeyi yüksek çıkacaktır.Maddelerin güçlük düzeyleri aynı ise KR21 formülü güvenirliğin gerçek değerini hesaplar.Madde güçlük düzeyleri farklı ise KR-21 gerçek değerin altında hesaplar.Ancak KR-20 formülü madde güçlük düzeyleri farklı olsa da testin güvenirliğinin gerçek değerini hesaplar.NOT:Testin Madde Güçlük Düzeyleri AYNIYSA:KR20=KR21Madde standart sapma ve varyans değerleri de eşit olur.Testin Madde Güçlük Düzeyleri FARKLIYSA:KR 20 > KR21Madde standart sapma ve varyans değerleri de farklı olur.KR21, güvenirliğin alt sınırını oluşturur.KR20-21 DEĞERİ YÜKSEK ÇIKARSA YAPILACAK KESİN YORUMLARMaddeler birbiriyle tutarlı, homojen, benzer olduğu,Testin tesadüfi hatalardan arınık olduğu,Testin tek boyutlu olduğu,Testin duyarlı olduğu , yorumları yapılır.KR20-21 DEĞERİ YÜKSEK ÇIKARSA YAPILABİLECEK YORUMLARYapı geçerliği yüksektir, denebilir.Kapsam geçerliği düşüktür, denebilir.Ayırt ediciliği yüksektir, denebilir.5. CRONBACH ALFA YÖNTEMİ Dereceli puanlamaya dayalı olan tek boyutlu likert tipi (tutum ölçeği, ilgi ölçeği vb.) ölçeklerin iç tutarlılığını ölçen güvenirlik yöntemidir.Bu yöntemde maddeler arasındaki uyuma, benzerliğe, iç tutarlılığa bakılır.Düzeltme formülü, ağırlıklı puanlama ve dereceleme ölçeklerin uygulandığı testlerde kullanılır.GÜVENİRLİĞİ ARTIRMA YOLLARI Soruların açık ve anlaşılır olmasıSoru sayısının artırılması (duyarlılık artar)Seçenek sayısının artırılması (şans azalır)Öğrencilerin güdülenmesiSınav süresinin yeterli olmasıSoruların orta güçlükte olmasıTestin dikkatli ve objektif puanlanmasıSoruların kolaydan zora doğru sıralanmasıSınavda yönerge hazırlanıp kullanılmasıTestte katılan kişi sayısının artırılmasıTestin uygulandığı grubun heterojen olmasıKopya çekilmesinin engellenmesiSınav ortamının uygun ölçülerde hazırlanmasıSTANDART HATA Ölçmenin standart hatası, güvenirlik belirleme yönteminde sınava giren her bir öğrencinin sınav puanları tek tek incelenir. Bu inceleme ile öğrencinin gerçek puan aralığı belirlenir.Tesadüfi hatada öğrenci puanına hatanın hangi yönde karıştığı bilinmez. Bu nedenle ölçmenin standart hatası düşünülerek gerçek puan aralığı belirlenir.Ölçmenin standart hatasında öğrencinin gerçek puanı değil gerçek puan aralığı belirlenir.(Se) STANDART HATA = (Sx) STANDART SAPMA * √1-rx şeklindedir. (rx= güvenirlik katsayısı)Hata hesaplandıktan sonra öğrenci puanına,Bir defa eklenip çıkarılırsa %68,İki defa eklenip çıkarılırsa %95,Üç defa eklenip çıkarılırsa %99, olasılıkla gerçek puan aralığı hesaplanmış olur.SORU: Kimya testinin güvenirliği 0.75, standart sapması 10 olduğuna göre 40 puan alan Güler’in %32 hata payı ile gerçek puan aralığı kaçtır ? (2011 KPSS)ÇÖZÜM: (Sx) yani standart sapma = 10Güvenirlik kat sayısı: 0.75Hata payı %32 ise bu %68 olasılığa bakılması gerekir. Bu bilgileri kullanarak cevabımızı artık bulabilirsiniz.√1-0.75 = 0.510*0.5= 5%68 olasılığa baktığımız için 40’a bir kere 5 ekleyip ve çıkaracağız bu da bize 35 ve 45 puan aralıklarını verecektir.SORU: Aşağıda standart sapması ve güvenirlik katsayısı verilen testlerden hangisinin standart hatası en düşüktür ? (2004 KPSS)ÇÖZÜM:Formülümüz belli (Sx) STANDART SAPMA * √1-rxSoruda her şey verilmiş zaten, şimdi bunları formüle yerleştireceğiz.A= 2*√1-0.50= 1.4B= 4*√1-0.50= 2.8C=5*√1-0.30= 4.1D=4*√1-0.70= 2.7E=2*√1-0.90= 0.6cevabımız bu koşullarda E şıkkı olacaktır.SORU: Aşağıda standart sapması ve güvenirlik katsayısı verilen testlerden hangisinin standart hatası en büyüktür ? (2010 KPSS)ÇÖZÜM:Formülümüz bir önceki soruda ki yine belli (Sx) STANDART SAPMA * √1-rxTabii ki sınav esnasında bunları uzun uzun hesaplamayacağız, artık kolay yönteme bakmak lazım. rx’i en düşük olan standart sapmanın en yüksek olduğu değeri belirleyelim. Bu da bize biyoloji dersini işaret ediyor. Kontrolümüzü yapalım;Biyoloji = 15* √1-0.64 = 9 yapacaktır. Yine standart sapması yüksek olan fizik dersini kontrol edelim ardından sonuç zaten bize göz kırpacaktır.Fizik = 15* √1-0.91 = 4.5 evet buradan da gördüğümüz üzere standart hatanın en büyük olduğu ders biyoloji dersidir.NOTSoruda standart hata hangisinde en büyüktür diyorsa; Standard sapma (Sx)’nın en büyük olduğu ve Güvenirlik kat sayısı (rx)’nın en düşük olarak birlikte verildiği şıkka yönelmeliyiz.Eğer soruda standart hata hangisinde en küçüktür diyorsa; Standard sapma (Sx)’nın en küçük olduğu ve Güvenirlik kat sayısı (rx)’nın en büyük olarak birlikte verildiği şıkka yönelmeliyiz.GEÇERLİK Ölçme aracının ölçmek istediği özelliklere başka değişkenler karıştırmadan ölçebilmesine geçerlik denir.Ölçme aracının amacına hizmet etme derecesidir.Amaç , eğitimde bilen öğrenci ile bilmeyen öğrenci ayrımının yapılabilmesidir.Geçerlik (-1) ve (+1) aralığında değerler alır.(+1)’e yaklaştıkça geçerlik artar, 0’ yaklaştıkça düşer.(-) değer ise bilen öğrencilerin yapamayıp, bilmeyen öğrencilerin yapmasıdır.Bir test 0.70 ve üstü değere sahipse geçerlidir.NOT: Her geçerli test, güvenilirdir. Fakat her güvenilir olan test, geçerli olmayabilir. Geçerlik , güvenirlikten daha geniş bir değer aralığına sahiptir. Bir testin geçerlik katsayısı, o testin güvenirliğinin karakökü kadardır.SORU: Bir testin güvenirliği 0.81 ise aynı testin geçerlik katsayısı ne olur?ÇÖZÜM: 0.81’in kareökünü alıyoruz. Bu da 0.9 olacaktır. Bu sonuca göre hem geçerlik katsayısını bulduk hem de testin katsayısı 0.70’den büyük olduğu için geçerli olduğunu görebiliriz.NOTBir testin geçerli olabilmesi için;1. Güvenilir olması ön koşuldur.2. İlgililik (uygunluk) özelliği olmalıdır.Hazırlanan her test, ölçülmek istenen özelliklere uygun hazırlanırsa geçerli olur.GEÇERLİK TÜRLERİ1. KAPSAM GEÇERLİĞİ Bir testte aranan en önemli geçerlik türüdür.Ölçme aracında yer alan soruların, öğrencilere anlatılan konulara uygun olması demektir.Örneklemin (soruların), evreni (anlatılan konulara) temsil etme derecesidir.Kapsam geçerliğini, uzun içerikli testlerde hazırlamak zordur. Bu nedenle bu tür sınavlarda örneklem alanı belirlenip sorular hazırlanır. Bu testler, BAŞARI TESTLERİ’dir.Başarı testlerinde, tüm davranışlar değil, kritik davranışlardan sorular sorulur.Kritik davranış, öğrencinin mutlaka bilmesi gereken önemli konulardır.Ünite testlerinde (İZLEME TESTLERİNDE) ise anlatılan tüm konulardan soru sormak önemlidir.İZLEME testleri, sürecin takibinde kullanılmaktadır.KAPSAM GEÇERLİĞİNİ SAĞLAMA 1. BELİRTKE TABLOSU: Hangi konudan, hangi hedef düzeyinde kaç soru sorulacağının belirtildiği tablodur. Sorular , konulara dengeli paylaştırılırsa testin kapsam geçerliği yüksek olur.2. RASYONEL YAKLAŞIM (UZMAN KANISI): Hazırlanan testlerin aynı branş uzmanlarının görüşlerinin alınmasıdır. Özellikle öğretmen yapımı testlerde kullanılır.3. İSTATİSTİKSEL YAKLAŞIM: Kapsam geçerliği uzmanlarca kanıtlanmış olan test ile yeni hazırlanıp, geçerliği kanıtlanacak testlerin karşılaştırılmasıdır.2. ÖLÇÜT GEÇERLİĞİ (Bir ölçüte göre geçerlik) Ölçüt geçerliği ölçütün elde ediliş zamanına göre ikiye ayrılır.a. YORDAMA GEÇERLİĞİ Bir ölçme aracının gelecekteki başarıyı, performansı tahmin etme derecesidir. Burada ölçüt testin tahmin etmesi gereken başarı veya performanstır.Test puanları önce elde edilir, ölçüt puanları sonra elde edilir.Testten elde edilen puanlarla ölçüt puanları arasındaki ilişki aynı yönde ve kuvvetli ise yordama geçerliği yüksek olur.LYS (mevcut sonuç) ………… ÜNİVERSİTE Başarısı (tahmin edilen)İş sınavı (mevcut sonuç) ………………… işteki performansı (tahmin edilen)KPSS (mevcut sonuç) ………………….. Öğretmenlik performansıb. UYGUNLUK GEÇERLİĞİ (EŞ ZAMAN, ZAMANDAŞ) Ölçüt puanlarının test puanlarıyla yakın zamanda veya daha önce elde edilmesi durumunda, ölçüt puanları ile test puanları arasındaki ilişkiye uygunluk geçerliği denir.Geçmişe tahmin etmektir, tanı koymaktır.Ölçüt puanları önce, test puanları sonra elde edilir.Lise başarı ortalaması (tahmin edilen) ………………… LYS (Test puanı)3. YAPI GEÇERLİĞİ Bir testin içinde varolan soruların birbiriyle olan uyumluğu, benzerliği, iç tutarlılığı bir yapıdır.Özellikle duyuşsal testlerde aranılan bir özelliktir.Eğer bir test içinde birden fazla ders varsa, her ders kendi içinde değerlendirilmelidir, her testin kendi içinde bir yapısı vardır.Bir testin yapı geçerliği artarsa kapsam geçerliği azalır. Çünkü yapı tek boyutluluk, sorularda uyumluluk, homojenliğe bakarken kapsam geçerliği, anlatılan her konudan soru, çok boyutluluk özelliğine bakar.YAPI GEÇERLİĞİ BELİRLEME YOLLARI1. FAKTÖR ANALİZİ: Ölçme aracındaki soruların aynı özelliği ölçüp ölçmediğine bakılır. Farklı özellikteki soru, testten atılır ve testin kalan sorularla iç uyumu, yapısı sağlanmış olur.2. TEST VE MADDE YANLILIĞI: Aynı zeka düzeyine sahip bireylerin sorulara doğru cevap verme olasılıklarının aynı olması gerekir. Aksi durumda test ve madde yanlılığından bahsedilir.ÖRNEK: Bir öğretmen toplama işlemiyle ilgili sorular sorarken yekünü bulunuz demiştir. Ancak toplama işlemini tüm öğrencilerin bilmesine rağmen sadece yekün kelimesini bilenler soruya doğru yanıt vermiştir. Bu durumda maddeler ölçmesi gereken özelliğin dışında bir özellik ölçmüştür. Kelime bilgisi yanlılığa neden olmuştur.3. GRUBA TEST UYGULAMA: Bir test bir gruba uygulanır, testin ölçtüğü özelliğe sahip olanlar yüksek puan , sahip olmayanlar düşük puan alırsa test, özelliğe sahip olanla olmayanı ayırt eder. Ve yapı geçerliği yüksek çıkar.4. UZMAN GÖRÜŞÜ: Testteki maddelerin her birinin ölçülmek istenen yapıyla ilgili olup olmadığını belirlemek için uzman görüşüne başvurmaktır. Yapı geçerliğinde en çok kullanılan yöntemdir.5. CEVAPLAYICILARIN CEVAP VERME SÜRECİNİN İNCELENMESİ: Testin uygulandığı gruptan rastgele bireyler seçilir ve soruları sesli okumaları, yorumlamaları istenir. Aynı zamanda da cevaplayıcının verdiği tepkiler gözlenir. Gerçek hayatta gösterilen tepkilerle soru sorulduğundaki verdiği tepkiler karşılaştırılır. Eğer tutarlılık yüksekse yapı geçerliği sağlanmış olur.NOT: Sabit ve sistematik hatalar yapı geçerliğini düşürür. Özellikle SİSTEMATİK hatalar!4. GÖRÜNÜŞ GEÇERLİĞİ Ölçme aracının ölçmesi gereken özelliği ölçüyor göründüğüyle ilgilidir.Görünüş geçerliği, ölçme aracının kapağının, içeriğiyle uyum halinde olmasına bağlıdır.NOT: Bilişsel alan testlerinde görünüş geçerliğinin yüksek olması istenilen bir durumdur. Ancak duyuşsal alan testlerinde görünüş geçerliğinin yüksek olması istenmez. Çünkü duyuşsal alanda ölçme yapılırken bireyler, kasıtlı olarak kendileri hakkında yanlış bilgi verirler. Bu kasıtlılığı önlemek için görünüş geçerliği düşürülür.GEÇERLİĞİ ARTIRMA YOLLARI Aracın kapsamı ölçülen kapsama ne kadar yakın olursa geçerlik o kadar artar.Güvenirlik artarsa geçerlik artar.Cevaplamaya verilen sürenin yeterli olması geçerliği artırır.Ölçme aracının güçlüğü amaca ne kadar uygun olursa geçerlik artar. (Seçme sınavları zor, izleme testleri kolay, başarı testleri orta güçlüktedir)Soruların tekrar tekrar kullanılmaması, öğrencilerden izole edilmesi, kopyanın engellenmesi geçerliği artırır.Testin uygulandığı grup heterojen olmalıdır.Testteki öğrenci sayısı artarsa geçerlik bir noktaya kadar artar sonra sabitlenir.Aynı seçeneğin doğru cevap olarak alt alta en fazla üç(3) kez gelecek şekilde hazırlanmalıdır.Sabit ve sistematik hatalar yapmamak, objektif puanlamalar yapmak geçerliği artırır.DİKKAT! Ölçme araçları amaca hizmet etme yönüyle ele alınacak olursa; Başarı testleri için kapsam geçerliği,İzleme testleri için kapsam geçerliği,Seçme ve yerleştirme için yordama geçerliği,Duyuşsal özellikleri ölçen (kişilik,tutum,ilgi vb.) testler için yapı geçerliğinin kanıtlanmış olması gerekir.GEÇERLİK ve GÜVENİRLİK ARASINDAKİ İLİŞKİ Güvenirlik,tesadüfi hatalardan etkilenirken; geçerlik sabit ve sistematik hatalardan etkilenir.Güvenirlik ölçme aracının hatalardan arınık olmasıdır, geçerlik ise ölçmenin, amaca hizmet etmesidir.Güvenirlik katsayısı 0 ve +1 aralığında değer alırken, geçerlik -1 ve+1 aralığında değer alır.Güvenirlik, geçerliğin ön koşuludur fakat tek başına yeterli değildir.Güvenirlik, geçerliği sınırlandırır.Güvenirlikle geçerlik arasında orantı yoktur, bağıntı vardır.Bir testin güvenirlik değerinin karakökü o testin geçerliğinden küçük olamaz.3. KULLANIŞLIK Ölçme aracının hazırlanış, uygulanış, ve puanlanış bakımından emek, zaman ve maliyet açısından kazançlı olmasıdır.Testin hazırlanıp uygulanmasının puanların yorumlanmasının kolay olmasıdır.Testlerdeki soru sayılarının azaltılması kullanışlığı artırır. Bir testin geçerliliğini araştırıyorsam ya bunu sonraki bir şeyle karşılaştırırım ya da bunu önceki bir şeyle karşılaştırırım mantığıyla hareket eden, saçma sapan gereksiz bi şey..Eğer bir testi önceki ya da sonraki bir testle karşılaştırıyorsa ölçüt geçerliliğini araştırıyordur.Sen eğer hazırladığın testi önceki bir şeyle karşılaştırıyorsan buradaki baktığın geçerlilik türü Uygunluk Geçerliliğidir.Sen eğer hazırladığın testi sonraki bir şeyle karşılaştırıyorsan buradaki baktığın geçerlilik türü Yordama Geçerliliğidir.Şimdi %100 geçerli olan süper bir pratik bilgi vereceğim sizlere.Bir soruyu okuduğunda baktın ki öğretmen hazırladığı testi ya da aracı herhangi bir şeyle karşılaştırıyor.Uygunluk mu? Yordama mı?Böyle bir soruda sorunun altına aşağıdaki gibi bir doğru çiz. Örnek olarak,Bir uzman öğrencilerin üniversite 1. sınıfda ki genel akademik notlarıyla YGS ’de ki sınav puanlarını karşılaştırmıştır.Bu süreçte YGS sınavının geçerliliğini araştıran uzman falan filan… gibi bir soru üzerinde yukarıdaki pratiği uygulayalım…Burada geçerliliği araştırılan durum YGS sınavı öyle değil mi?Soruda anlaşıldığı üzere YGS’yi üniversite 1. sınıfta ki genel akademik notlarıyla karşılaştırıyorsa demek ki doğrumuzu aşağıdaki gibi çizeceğiz.Soruda YGS sınavındaki puanlarıyla lisedeki notlarını karşılaştırdığını söyleseydi eğer,Ölçüt Geçerliliği seçme sınavlarında önemlidir.Bir soruda Yordama veya Uygunluğu işaretleyebilmen için kesinlikle bu adamın hazırladığı ölçme aracını bir şeyle karşılaştırması gerekiyor.Eğer bir karşılaştırma söz konusu değilse bu soruda Yordama ve Uygunluktan bahsedemezsin. Bir testin belirlenmiş bir karşılaştırma standardıyla (yani bir ölçütle) ne kadar iyi korelasyon gösterdiğine dair bir indekstir. kriter geçerliliği üç türe ayrılır: öngörüsel geçerlik, eş zamanlı geçerlik ve geriye dönük geçerlik. örneğin, bir suç davranışı ölçütü geçerliyse, bir bireyin (a) gelecekte bir suç ihlali nedeniyle tutuklanıp tutuklanmayacağını, (b) şu anda yasaları ihlal edip etmediğini ve (c) daha önce sabıka kaydı olup olmadığını tahmin etmek için kullanmak mümkün olmalıdır. kriter referanslı geçerlik; kriterle ilgili geçerlik olarak da adlandırılır. Araştırmacıların, veri toplamak için kullanmadan önce araştırma araçlarının hem güvenilir hem de geçerli olduğundan emin olmaları gerekir. Veri güvenilirliği ve geçerliliği, araştırmada toplanan verilerin kalitesini değerlendirmek için kullanılan iki önemli kavramdır. Hem veri güvenilirliği hem de geçerliliği, elde edilen sonuçların doğru, tutarlı ve güvenilir olmasını sağladıkları için araştırmalarda çok önemlidir. Araştırmada güvenilirlik ve geçerlilik Güvenilirlik Bir ölçümün zaman içinde veya farklı değerlendiriciler arasında tutarlılığını veya istikrarını ifade eder. Bir ölçüm aracı, aynı şeyi ölçmek için tekrar tekrar kullanıldığında tutarlı bir şekilde aynı sonuçları üretiyorsa güvenilir olarak kabul edilir. Başka bir deyişle, kimin uyguladığına, ne zaman uygulandığına ve hangi koşullar altında uygulandığına bakılmaksızın tutarlı sonuçlar üreten güvenilir bir araçtır. Geçerlilik Bir ölçüm aracının ölçmesi gereken şeyi ne ölçüde ölçtüğünü ifade eder. Bir ölçüm aracı, ölçmeyi amaçladığı kavram veya yapıları doğru bir şekilde ölçüyorsa geçerli kabul edilir. Başka bir deyişle, ölçtüğünü iddia ettiği şeyi ölçen ve araştırma sorusuyla anlamlı ve ilgili sonuçlar üreten geçerli bir araçtır. Özetlemek gerekirse, güvenilirlik bir ölçümün tutarlılığını ifade ederken, geçerlilik bir ölçümün doğruluğunu veya gerçekliğini ifade eder. Güvenilirlik sonuçların tutarlılığını ve istikrarını sağlamak için önemliyken, geçerlilik bir ölçüm aracından elde edilen sonuçların anlamlı ve araştırma sorusuyla ilgili olmasını sağlamak için çok önemlidir. Bu tablo, araştırmada güvenilirlik ve geçerlilik için hızlı bir referans sağlar: KriterlerGüvenilirlikGeçerlilikTanımBir ölçümün zaman içinde ve farklı örneklemler arasında tutarlı bir şekilde aynı sonuçları üretme derecesi.Bir ölçümün ölçmesi gereken şeyi doğru bir şekilde ölçme derecesi.TürleriTest-tekrar test güvenilirliği, değerlendiriciler arası güvenilirlik, iç tutarlılık güvenilirliği.İçerik geçerliliği, ölçüt geçerliliği ve yapı geçerliliği.ÖrnekBir ilacın kan basıncı üzerindeki etkisi, farklı zamanlarda alınan kan basıncı değerlerinin birbiriyle tutarlı olup olmadığı kontrol edilerek ölçülebilir.İş memnuniyetini doğru bir şekilde ölçen bir anket.ÖnemTutarlı sonuçlar üretmek ve ölçüm hatalarını azaltmak için önemlidir.Bir ölçümün geçerliliğini sağlamak için önemlidir.HedefBir ölçüm aracının tutarlı sonuçlar üretmesini sağlamak.Bir ölçüm aracının doğru sonuçlar ürettiğinden emin olmak için. Veri güvenilirliği, bir çalışmanın sonuçlarının ne ölçüde tekrarlanabileceğini ifade eder. Anket veya soru formu gibi bir araştırma aracından elde edilen sonuçların tutarlılığını ve istikrarını ifade eder. Bir araştırma aracı güvenilirse, aynı katılımcı grubuna birden fazla kez uygulandığında tutarlı sonuçlar üretmelidir. Geçerlilik Nedir? Veri geçerliliği, bir araştırma aracından elde edilen sonuçların doğruluğunu ifade eder. Bir araştırma aracının ölçmeyi amaçladığı şeyi ne ölçüde ölçtüğünü ifade eder. Başka bir deyişle, bir araştırma aracı geçerli ise, ölçmek için tasarlandığı kavram veya olguyu doğru bir şekilde ölçmelidir. Görünüş geçerliliği, içerik geçerliliği, yapı geçerliliği ve kriter geçerliliği gibi farklı geçerlilik türleri vardır. Via hoganassessments.com Güvenilirlik ve geçerliliğin değerlendirilmesi, anketler, soru formları veya testler gibi herhangi bir araştırma aracının geliştirilmesi ve değerlendirilmesinde önemli bir süreçtir. Güvenilirlik, ölçümün zaman içinde veya farklı puanlayıcılar arasında tutarlılığı veya istikrarını ifade ederken, geçerlilik aracın ölçmesi gereken şeyi ne ölçüde ölçtüğünü ifade eder. Genel olarak, güvenilirlik ve geçerliliğin değerlendirilmesi, araştırma araçlarının doğru olmasını ve tutarlı sonuçlar üretmesini sağlamak açısından çok önemlidir; bu da araştırmanın kalitesini ve güvenilirliğini artırır. Güvenilirlik Güvenilirliği değerlendirmek için test-tekrar test güvenilirliği, değerlendiriciler arası güvenilirlik ve iç tutarlılık güvenilirliği dahil olmak üzere çeşitli yöntemler vardır. İşte güvenilirliğin araştırmada nasıl çalışabileceğini gösteren örnek bir tablo: Güvenilirlik TürüAçıklamaÖrnekTest-Tekrar Test GüvenilirliğiBir ölçümün aynı gruba iki veya daha fazla kez uygulandığında zaman içinde tutarlı sonuçlar üretme derecesi.Bir grup katılımcıya bir anket uygulamak ve bir hafta sonra aynı anketi aynı katılımcı grubuna uygulamak ve sonuçları karşılaştırmak.Değerlendiriciler Arası GüvenilirlikFarklı puanlayıcıların veya gözlemcilerin aynı şeyi puanlarken veya gözlemlerken tutarlı sonuçlar üretme derecesi.İki veya daha fazla puanlayıcının aynı kompozisyon setini bağımsız olarak puanlaması ve tutarlı olup olmadıklarını belirlemek için puanlarını karşılaştırması.İç Tutarlılık GüvenilirlikBir ölçümdeki farklı maddelerin veya soruların tutarlı sonuçlar üretme derecesi.Özsaygı gibi aynı yapıyı ölçmek için tasarlanmış bir anketteki farklı maddelere verilen yanıtlar arasındaki korelasyonun hesaplanması. Bunun sadece örnek bir tablo olduğunu ve araştırma bağlamına bağlı olarak dahil edilebilecek başka birçok güvenilirlik türü olduğunu unutmayın. Önemli olan, çalışmanız için uygun güvenilirlik türünü (türlerini) seçmek ve araştırmanızı güvenilirliği etkili bir şekilde değerlendirmenize olanak tanıyacak şekilde tasarlamaktır. Geçerlilik Geçerliliği değerlendirmek için içerik geçerliliği, ölçüt geçerliliği ve yapı geçerliliği gibi farklı geçerlilik türleri incelenebilir. Geçerliliğin araştırmada nasıl işleyebileceğini gösteren örnek bir tablo aşağıda verilmiştir: Geçerlilik TürüAçıklamaÖrnekİçerik GeçerliliğiBir ölçümün belirli bir yapının tüm yönlerini ne ölçüde yakaladığı.Depresyonu ölçmek için tasarlanmış bir dizi soru veya maddenin gözden geçirilmesi ve bunların depresyonun ilgili tüm yönlerini kapsayıp kapsamadığının değerlendirilmesi.Ölçüt GeçerliliğiBir ölçütün belirli bir kriter veya sonuçla ne ölçüde ilişkili olduğu.Yeni bir matematik becerisi ölçümü puanlarının, yeni ölçümün mevcut testteki performansı ne kadar iyi tahmin ettiğini değerlendirmek için "altın standart" olarak kabul edilen mevcut bir matematik testindeki puanlarla ilişkilendirilmesi.Yapı GeçerliliğiBir ölçümün amaçlanan yapıyı doğru bir şekilde ölçme derecesini ifade eder.Ölçeğin amaçlanan yapıyı ne ölçüde yakaladığını değerlendirmek için bir dışadönüklük ölçümündeki puanları sosyal aktivite ve dışa dönük davranış gibi ilgili yapıların diğer ölçümleriyle ilişkilendirmek. Mind the Graph bilim insanlarına görsel olarak çekici ve etkili slaytlar oluşturmaları için araçlar sağlayan bir platformdur. Platform, araştırmanız için bilim insanlarının bilgilerini ve illüstrasyonlarını hem estetik hem de etkili bir şekilde düzenlemelerine yardımcı olan bir dizi görsel olarak çekici figür sunar. Sonuç, bilimsel bilgileri daha geniş bir kitleye etkili bir şekilde ileten profesyonel görünümlü bir sunumdur. Mind the Graph ile bilim insanları bilgi ve illüstrasyonları güzel ve etkili sunumlara kolayca entegre edebilirler. Mind the Graph ile Yaratmaya Başlayın Etkili görseller hakkında özel yüksek kaliteli içerik bilimde iletişim. Geçerlilik ve Türleri ölçme işleminin ardından elde edilen puanların kullanış amaçlarına göre farklı türlere ayrılmaktadır.Geçerlilik için bilmemiz gereken iki temel unsur vardır.Geçerlilik Bir ölçme aracının amaca hizmet etme derecesidir.Güvenirlik , geçerliliğin ön koşuludur. Bir testin güvenirli olması geçerli olduğu anlamına gelmez.Örnek olarak bir atış poligonuna bakalım. Bir sınav testinin ölçmek istediği konulara kapsam denir.Başarı testlerinin daima önceliğidir.Hedeflerin sınavda örnek soru olarak karşımıza çıkmasıdır.Konular sınavda ne eksik ne de fazla olmalıdır.Konuların genişliğine göre sorular dengeli bir şekilde sınav kağıdına dağıtılmalıdır.Burada kapsam geçerliliğini kontrol etmek için farklı yöntemler uygulanabilir.Belirtke Tablosu : Kaç ders saati varsa ona oranla soru sayısını tablo ile belirlemek.Uzman Görüşü : En az iki uzmana hazırladığımız testi götürüyor ve onlardan öneri alıyoruz.İstatistiksel Yol :Bizim hazırladığımız test ile daha önceden hazırlanan bir test aynı gruba verilir. Aralarındaki korelasyon yakın ise sınavın kapsamı yüksektir. 2.) Ölçüt Geçerliliği (Yordama ve Uygunluk) a.) Uygunluk GeçerliliğiMevcut bir ölçüm ile geçmişte ya da eş zamanlı performansı karşılaştırmaktır.Bir test hazırladık bunun kapsam geçerliliğini merak ediyoruz ve geçmişte yapılan bir ” kapsam geçerliliği kanıtlanmış ” test ile ilişkisine bakıyoruz.Dikkat edilmesi nokta geçmişe dönüktür. b.) Yordama GeçerliliğiMevcut ölçüm geleceği tahmin etmesidir.Yapılan deneme sınavları gelecekte olan KPSS ‘deki puanımıza benzer ise kapsam geçerliliği başarılı demektir.Dikkat edilmesi gereken nokta gelecekteki sınava dönük olmasıdır. Testin istenilen yapıyı ölçme amacıdır.Davranışın ölçme aracı uyuşmasıdır.Çevre bilinci yüksek birinin , çöpü yere atmamasıdır.Çalışmayan bir öğrenci sınavdan yüksek alırsa yapı geçerliliği düşüktür.Çünkü başarıyı ölçen bir testte başarılı olmayan bir öğrenci yüksek almıştır.KAYNAK: Araştırmacıların, veri toplamak için kullanmadan önce araştırma araçlarının hem güvenilir hem de geçerli olduğundan emin olmaları gerekir. Veri güvenilirliği ve geçerliliği, araştırmada toplanan verilerin kalitesini değerlendirmek için kullanılan iki önemli kavramdır. Hem veri güvenilirliği hem de geçerliliği, elde edilen sonuçların doğru, tutarlı ve güvenilir olmasını sağladıkları için araştırmalarda çok önemlidir. Araştırmada güvenilirlik ve geçerlilik Güvenilirlik Bir ölçümün zaman içinde veya farklı değerlendiriciler arasında tutarlılığını veya istikrarını ifade eder. Bir ölçüm aracı, aynı şeyi ölçmek için tekrar tekrar kullanıldığında tutarlı bir şekilde aynı sonuçları üretiyorsa güvenilir olarak kabul edilir. Başka bir deyişle, kimin uyguladığına, ne zaman uygulandığına ve hangi koşullar altında uygulandığına bakılmaksızın tutarlı sonuçlar üreten güvenilir bir araçtır. Geçerlilik Bir ölçüm aracının ölçmesi gereken şeyi ne ölçüde ölçtüğünü ifade eder. Bir ölçüm aracı, ölçmeyi amaçladığı kavram veya yapıları doğru bir şekilde ölçüyorsa geçerli kabul edilir. Başka bir deyişle, ölçtüğünü iddia ettiği şeyi ölçen ve araştırma sorusuyla anlamlı ve ilgili sonuçlar üreten geçerli bir araçtır. Özetlemek gerekirse, güvenilirlik bir ölçümün tutarlılığını ifade ederken, geçerlilik bir ölçümün doğruluğunu veya gerçekliğini ifade eder. Güvenilirlik sonuçların tutarlılığını ve istikrarını sağlamak için önemliyken, geçerlilik bir ölçüm aracından elde edilen sonuçların anlamlı ve araştırma sorusuyla ilgili olmasını sağlamak için çok önemlidir. Bu tablo, araştırmada güvenilirlik ve geçerlilik için hızlı bir referans sağlar: KriterlerGüvenilirlikGeçerlilikTanımBir ölçümün zaman içinde ve farklı örneklemler arasında tutarlı bir şekilde aynı sonuçları üretme derecesi.Bir ölçümün ölçmesi gereken şeyi doğru bir şekilde ölçme derecesi.TürleriTest-tekrar test güvenilirliği, değerlendiriciler arası güvenilirlik, iç tutarlılık güvenilirliği.İçerik geçerliliği, ölçüt geçerliliği ve yapı geçerliliği.ÖrnekBir ilacın kan basıncı üzerindeki etkisi, farklı zamanlarda alınan kan basıncı değerlerinin birbiriyle tutarlı olup olmadığı kontrol edilerek ölçülebilir.İş memnuniyetini doğru bir şekilde ölçen bir anket.ÖnemTutarlı sonuçlar üretmek ve ölçüm hatalarını azaltmak için önemlidir.Bir ölçümün geçerliliğini sağlamak için önemlidir.HedefBir ölçüm aracının tutarlı sonuçlar üretmesini sağlamak.Bir ölçüm aracının doğru sonuçlar ürettiğinden emin olmak için. Veri güvenilirliği, bir çalışmanın sonuçlarının ne ölçüde tekrarlanabileceğini ifade eder. Anket veya soru formu gibi bir araştırma aracından elde edilen sonuçların tutarlılığını ve istikrarını ifade eder. Bir araştırma aracı güvenilirse, aynı katılımcı grubuna birden fazla kez uygulandığında tutarlı sonuçlar üretmelidir. Geçerlilik Nedir? Veri geçerliliği, bir araştırma aracından elde edilen sonuçların doğruluğunu ifade eder. Bir araştırma aracının ölçmeyi amaçladığı şeyi ne ölçüde ölçtüğünü ifade eder. Başka bir deyişle, bir araştırma aracı geçerli ise, ölçmek için tasarlandığı kavram veya olguyu doğru bir şekilde ölçmelidir. Görünüş geçerliliği, içerik geçerliliği, yapı geçerliliği ve kriter geçerliliği gibi farklı geçerlilik türleri vardır. Via hoganassessments.com Güvenilirlik ve geçerliliğin değerlendirilmesi, anketler, soru formları veya testler gibi herhangi bir araştırma aracının geliştirilmesi ve değerlendirilmesinde önemli bir süreçtir. Güvenilirlik, ölçümün zaman içinde veya farklı puanlayıcılar arasında tutarlılığı veya istikrarını ifade ederken, geçerlilik aracın ölçmesi gereken şeyi ne ölçüde ölçtüğünü ifade eder. Genel olarak, güvenilirlik ve geçerliliğin değerlendirilmesi, araştırma araçlarının doğru olmasını ve tutarlı sonuçlar üretmesini sağlamak açısından çok önemlidir; bu da araştırmanın kalitesini ve güvenilirliğini artırır. Güvenilirlik Güvenilirliği değerlendirmek için test-tekrar test güvenilirliği, değerlendiriciler arası güvenilirlik ve iç tutarlılık güvenilirliği dahil olmak üzere çeşitli yöntemler vardır. İşte güvenilirliğin araştırmada nasıl çalışabileceğini gösteren örnek bir tablo: Güvenilirlik TürüAçıklamaÖrnekTest-Tekrar Test GüvenilirliğiBir ölçümün aynı gruba iki veya daha fazla kez uygulandığında zaman içinde tutarlı sonuçlar üretme derecesi.Bir grup katılımcıya bir anket uygulamak ve bir hafta sonra aynı anketi aynı katılımcı grubuna uygulamak ve sonuçları karşılaştırmak.Değerlendiriciler Arası GüvenilirlikFarklı puanlayıcıların veya gözlemcilerin aynı şeyi puanlarken veya gözlemlerken tutarlı sonuçlar üretme derecesi.İki veya daha fazla puanlayıcının aynı kompozisyon setini bağımsız olarak puanlaması ve tutarlı olup olmadıklarını belirlemek için puanlarını karşılaştırması.İç Tutarlılık GüvenilirlikBir ölçümdeki farklı maddelerin veya soruların tutarlı sonuçlar üretme derecesi.Özsaygı gibi aynı yapıyı ölçmek için tasarlanmış bir anketteki farklı maddelere verilen yanıtlar arasındaki korelasyonun hesaplanması. Bunun sadece örnek bir tablo olduğunu ve araştırma bağlamına bağlı olarak dahil edilebilecek başka birçok güvenilirlik türü olduğunu unutmayın. Önemli olan, çalışmanız için uygun güvenilirlik türünü (türlerini) seçmek ve araştırmanızı güvenilirliği etkili bir şekilde değerlendirmenize olanak tanıyacak şekilde tasarlamaktır. Geçerlilik Geçerliliği değerlendirmek için içerik geçerliliği, ölçüt geçerliliği ve yapı geçerliliği gibi farklı geçerlilik türleri incelenebilir. Geçerliliğin araştırmada nasıl işleyebileceğini gösteren örnek bir tablo aşağıda verilmiştir: Geçerlilik TürüAçıklamaÖrnekİçerik GeçerliliğiBir ölçümün belirli bir yapının tüm yönlerini ne ölçüde yakaladığı.Depresyonu ölçmek için tasarlanmış bir dizi soru veya maddenin gözden geçirilmesi ve bunların depresyonun ilgili tüm yönlerini kapsayıp kapsamadığının değerlendirilmesi.Ölçüt GeçerliliğiBir ölçütün belirli bir kriter veya sonuçla ne ölçüde ilişkili olduğu.Yeni bir matematik becerisi ölçümü puanlarının, yeni ölçümün mevcut testteki performansı ne kadar iyi tahmin ettiğini değerlendirmek için "altın standart" olarak kabul edilen mevcut bir matematik testindeki puanlarla ilişkilendirilmesi.Yapı GeçerliliğiBir ölçümün amaçlanan yapıyı doğru bir şekilde ölçme derecesini ifade eder.Ölçeğin amaçlanan yapıyı ne ölçüde yakaladığını değerlendirmek için bir dışadönüklük ölçümündeki puanları sosyal aktivite ve dışa dönük davranış gibi ilgili yapıların diğer ölçümleriyle ilişkilendirmek. Mind the Graph bilim insanlarına görsel olarak çekici ve etkili slaytlar oluşturmaları için araçlar sağlayan bir platformdur. Platform, araştırmanız için bilim insanlarının bilgilerini ve illüstrasyonlarını hem estetik hem de etkili bir şekilde düzenlemelerine yardımcı olan bir dizi görsel olarak çekici figür sunar. Sonuç, bilimsel bilgileri daha geniş bir kitleye etkili bir şekilde ileten profesyonel görünümlü bir sunumdur. Mind the Graph ile bilim insanları bilgi ve illüstrasyonları güzel ve etkili sunumlara kolayca entegre edebilirler. Mind the Graph ile Yaratmaya Başlayın Etkili görseller hakkında özel yüksek kaliteli içerik bilimde iletişim. Özet Bilimsel çalışmalarda somut ve soyut özelliklerin ölçülmesi için ölçme araçlarına ihtiyaç vardır. Ölçme araçlarının standardize olabilmesi için vazgeçilmez iki temel özelliğin olması gerekmektedir. Bu özellikler geçerlilik ve güvenirliktir. Bu çalışmada geçerlilik ve güvenirliğin tanımı, çeşitleri ve arasındaki ilişki ve farkları bulunmaktadır. Anahtar Kelime Güvenirlilik, Geçerlilik, Ölçek Giriş Ölçme, bir özelliğe sahip oluş derecesini belirtmek için, belirli kurallara uyularak sembolik değerler verme işlemidir. Bir başka anlamda bir değişkenin sembolle ve özellikle sayı ile eşleştirilmesidir. Bilimsel araştırmalarda ölçmenin temel amacı bir konu üzerinde değerlendirme yapmak ve karar vermektir. Değerlendirme sonucunun ve dolayısıyla da kararların doğruluğu ve uygunluğu değerlendirme sonuçlarına dolayısıyla da ölçeğin uygun olmasına bağlıdır. Bunu sağlamak için ölçeğin standardize olması gerekir. Standardize olan ölçekler nesneldir. Ölçme araçlarının standardize olabilmesi için vazgeçilmez iki temel özelliğin olması gerekmektedir. Bu özellikler geçerlilik ve güvenirliliktir (Ercan & Kan, 2004). Her ölçeğin öncelikle ölçülecek özelliği tam ve doğru olarak başka bir özelikle karıştırmadan ölçmelidir. Bu özelliğe geçerlilik denir. Bir ölçeğin kullanış amacına uygunluğunun derecesini geçerlilik belirler. Ayrıca, bir ölçme aracı ölçmek istediği özelliği tutarlı olarak ölçebilmelidir. Ölçeme işlemi aynı şartlar altında aynı ölçekle tekrarlandığında aynı sonuçları verebilmelidir. Bu özelliği güvenirlik olarak adlandırılır (Ergin, 1995). Güvenirlik Ölçeğin taşıması gereken özelliklerden birisi olan güvenirlik, bir ölçme aracıyla aynı koşullarda farklı zamanlarda tekrarlanan ölçümlerde elde edilen ölçüm değerleri arasındaki tutarlılığına o aracın güvenirliliği denir. Güvenirlik sadece ölçme aracına ait bir özellik değildir, aracın sonuçlarına ilişkin bir özelliktir. (Öncü, 1994). Ölçekle toplanan bilgilerin tutarlı özellik taşıdığına bir başa değişle hatasız olduğuna ve aynı hedefle yapılacak başka bir ölçmede aynı sonuçların elde edileceğine güven duyulması gerekir. Güven duyulmayan bir ölçek kullanışsızdır. Örnek olarak, bir elmayı defalarca tarttığımızda hep aynı değeri buluyorsak veya insanlara farklı zamanlarda sorduğumuz aynı sorularda aynı cevapları alıyorsak ölçme aracımızın güvenilir olduğunu söyleyebilir (Ercan & Kan, 2004). Güvenirlik üç öğeden oluşur; 1. Değişmezlik: Değişmezlik bir ölçümün ne derece kararlı olduğunu, zaman içinde ne derece değişmeden kaldığını ifade eder (Bryman, 2004) 2. İç güvenirlik: İç güvenirlik, ölçeği oluşturan değişkenlerin tutarlı olmasını, yani cevaplayıcıların bir değişkene verdikleri cevabın, diğer değişkenlere verdikleri cevaplarla ilişkili olmamasını ifade eder (Bryman, 2004) 3. Gözlemciler arası tutarlılık: Bu kavram da çeşitli gözlemcilerin gözlemlerini kaydetme ya da verileri kategorilere çevirme konusundaki kararlarını ifade eder. Gözlemciler bu konularda önceden alınmış kararlara bağlı kalmalı, özellikle veriler birden fazla kişi tarafından toplanıyor ve kaydediliyorsa farklı gözlemcilerin kayıtlarının tutarlı olması sağlanmalıdır (Bryman, 2004). Bir ölçeğin güvenirliliğinin yüksek olması o ölçeğin ölçmede izlenen süreçler ile kullanılan ölçütlerin detaylı olarak belirlenmesine bağlıdır. Bilimsel çalışmalarda güvenirliği artmak için madde ve soru sayısı artırılarak çok sayıda ölçüt kullanılır. Böylece hataların birbirini dengelemesiyle güvenirliği yüksek sonuçlar alınabilir. Birbiri ardına yapılan ölçmelerde bireyin grup içinde tutarlılığı aranır. Bu yaklaşımla aynı nesneler ile iki ölçüm arası korelasyon hesaplanır. Bulunan bu korelasyon değerine, güvenirlik kat sayısı denir. Güvenirlik katsayısı, gerçek ölçümlerin puanın, gözlenen (ölçülen) puanlara oranıdır (Ergin, 1995) Güvenilirliğin tespiti geçerliliğe göre daha somut işlemlere dayanır ve net sayısal verilerle ifade edilebilir. Güvenirliliği tespit etmek için kullanılan bazı yöntemler aşağıda listelenmektedir. Test-Tekrar Test Yolu: Ölçüm aracının aynı örnekleme benzer koşullar altında aralıklarla iki kere uygulanması ve elde edilen sonuçların karşılaştırılmasıdır. Sonuçlar arasındaki ilişkinin katsayısı, güvenirlik katsayısıdır. Uygulanacak iki ölçüm arasındaki zaman aralığı, cevaplayıcıların önceki uygulamada verdikleri cevapları hatırlamayacakları kadar uzun, ancak ölçülmek istenen özellikte değişiklik meydana gelmesine izin vermeyecek kadar da kısa olmalıdır. (Sencer & Sencer, 1978) Eşdeğer Biçimler Yolu: Ölçüm aracı aynı amaca yönelik olan iki farklı biçimde hazırlanır, örneğin cevap kategorilerinin ya da soruların yerleri değiştirilir ve aynı örnekleme ardı ardına uygulanır. Elde edilen sonuçlar karşılaştırılır ve güvenirlik düzeyi belirlenir (Sencer & Sencer, 1978). Yarıya Bölme Yolu: Bu yolda ölçme tamamlandıktan sonra elde edilen veriler tesadüfî yollarla ikiye ayrılır. Bu iki veri grubu ayrı ayrı ele alınarak birbirleriyle karşılaştırılır ve aralarındaki tutarlılık ölçülerek güvenirlik düzeyi belirlenir. Soruların kötü ifade edilmesi, cevapları kodlayan araştırmacıların bireylerin verdikleri açık uçlu soruları farklı kodlarla kodlamaları ya da soruların ilgisiz ya da konuyla ilgili çok az bilgisi olan kişilere sorulması (örneklemin yanlış seçilmesi) gibi faktörler, güvenirlik düzeyinin düşmesine neden olur (Sencer & Sencer, 1978). İç tutarlılık metodu: Araştırma belli bir kavramı ölçmek için aynı ölçüm düzeyinde birbirinden farklı sorulardan aynı form içerisinde yaralanıyorsa iç tutarlılık metodu uygulanmaktadır. Bu yöntemde araştırmacı testin güvenilirliği açısından cevaplayıcının birbirine benzer sorularda benzer cevaplar vermesini yani cevaplayıcının, cevapları arasında tutarlı olmasını bekler dolayısıyla benzer sorulardan benzer cevapların alınması testin güvenilirliği açısından delil olarak kabul edilir (Ergin, 1995) Gözlemciler arası güvenilirlik: Araştırmanın uygulanmasında denekler değil de gözlemciler ön planda ise araştırmacı aynı kişi, olay veya yeri değerlendirmek üzere farklı gözlemciler ile çalışır ve elde edilen sonuçların birbiri ile uyumu araştırmanın güvenilirlik derecesini verir. Özellikle araştırmanın değişkenlerinin nitel olması veya daha az nicel olması bu yöntemin tercihinde etkili olmaktadır (Ercan & Kan, 2004). Geçerlilik Bir ölçme aracının güvenilir olması doğru bilgi elde edebilmek için yeterli değildir. Her ölçeğin öncelikle ölçülecek özelliği tam ve doğru olarak başka bir özelikle karıştırmadan ölçmelidir. Bu özelliğe geçerlilik denir (Ercan & Kan, 2004). Bir başka deyişle geçerlik, bir ölçme aracının ölçmeyi amaçladığı özelliği, başka herhangi bir özellikle karıştırmadan, doğru ölçebilme derecesidir (Ergin, 1995). Bir ölçeğin geçerliliği; kapsamlılık, kolay uygulanabilirlik, nesnelliği, maddeleri ayırt etme gücü, puanlana bilirlikten etkilenir. Bunlar güvenirlilik değil bir geçerlilik sorunudur. Güvenirlik ve geçerlik birbirinden tamamen ayrı iki kavramlardır. Bulunan istatistik değerleri arasında da hiçbir ilişki yoktur. Bir ölçme aracının mutlaka hem geçerli hem güvenilir olması anlamında birbirlerini tamamlarlar. Bir ölçmenin geçerli sayılabilmesinin ilk koşulu onun güvenilir olmasıdır. Fakat birbirinin ön koşulu değildir. Bir ölçek standardize olabilmesi için her ikisinin de birlikte olması gerekir. Bir ölçeğin güvenilir olması o ölçeğin geçerli olduğu anlamına gelmez (Ercan & Kan, 2004). Bir bilimsel araştırmadan elde edilen sonuçların geçerliği üç açıdan değerlendirilir. Bunlardan ilki ölçüt geçerliği, ikincisi iç geçerlik ve üçüncüsü de dış geçerliktir. Ölçüt Geçerliliği Bir ölçme aracının ölçmeyi hedeflediğini ölçmesidir. Örnek olarak uzunluk ölçme aracı ile ağırlık ölçemeye çalışılırsa bu geçerli bir ölçüm değildir. İç geçerlik, ölçüm aracının ölçmeyi amaçladığı özelliği, bir başka özellikle karıştırmadan doğru bir şekilde ölçebilme derecesidir. Başka bir deyişle bağımlı değişkendeki farklılaşmanın, bağımsız değişkenlerle açıklanma derecesidir (Sencer & Sencer, 1978). İç Geçerlilik Neden sonuç ilişkisinde bulunan sonuca, ölçülmekte olan özellik değil de çalışmanın kendisi, çalışma yolu ya da tekniği neden olursa çalışma iç geçerlik açısından geçersizdir. Örnek olarak, ders çalışma sıklığının final notları üzerindeki etkisinin incelendiği bir araştırmada, ders çalışma sıklığı ve not ortalaması arasındaki ilişkinin iç geçerliliği, notlardaki değişikliklerin gerçekten de ders çalışma sıklığından kaynaklanması olasılığıdır. Eğer notlar araştırmacının kontrolü dışındaki başka etkenler tarafından da etkileniyorsa iç geçerlik düşüktür. Araştırma sürecinde iç geçerliği yükseltmek için dış faktörler kontrol altında tutulmaya çalışılır, bu nedenle iç geçerliğin en yüksek olduğu araştırmalar laboratuvar ortamında yürütülen araştırmalardır (Sencer & Sencer, 1978). İç geçerliği tespit etmek için yüzeysel geçerliğin sınanması ve içerik geçerliğinin sınanması gerekir. Yüzeysel geçerleme, bir aracın ölçülmek istenen konuyu ölçer görünmesidir. Örneğin bir muhafazakârlık tanımı yapılarak buna göre geliştirilen bir ölçek, mantıksal olarak siyasal sistem, mülkiyet, aile gibi konularda gelenekçi tutum ve davranışları ölçmek için geçerlidir. Ancak yüzeysel geçerleme sadece mantıksal düzeyde kaldığı için ölçme aracının gerçekte neyi ölçtüğü konusuna açıklık getirmez. İçerik geçerliği ise, bir ölçme aracının ölçülen kavramla ilgili mantıksal içerimleri yeterince temsil etmesidir. Yani, ölçeği oluşturan soruların ya da maddeler, ölçülen konuya ilişkin çeşitli özellikleri iyi temsil etmelidir. Bir önceki örnek üzerinden düşünmek gerekirse, geliştirilen muhafazakârlık ölçeği içindeki soru ya da maddelerin muhafazakârlığı ölçebiliyor olması ve ölçekte ilgili konuyu ölçebilecek kadar çok sayıda soru ya da madde olması gerekir. İç geçerliği etkileyen faktörlerden başlıcaları örneklemin yanlış seçilmesi ve zaman faktörüdür. Zaman faktörü, bağımsız değişken dışında kalan önemli bazı değişkenlerin bağımlı değişken üzerinde yarattığı etkinin zamanla artması ve bağımlı değişkende meydana gelen değişikliğin asıl nedenini bulmayı zorlaştırmasıdır. (Ergin, 1995). Dış Geçerlilik Bilimsel araştırmanın örneklem hakkında elde edilen bulguların ne dereceye kadar genelleyebileceği, araştırmanın dış geçerliğini oluşturur Örneğin; bir deneyde, grubun iş motivasyonunu artıran bir etkenin, gerçek işyerlerindeki çalışanların da iş motivasyonunu benzer düzeyde artırma olasılığı, bu etkenin dış geçerliğini ifade eder. Buna ek olarak, dış geçerlik, genelleme geçerliği olarak da adlandırılır Başka bir şekilde tanımlayacak olursak bir örneklemden elde ettiğimiz bilgilerin evren hakkında, ya da bir ortamdan elde ettiğimiz bilgilerin başka ortamlar hakkında ne derece bilgi verdiği, dış geçerlik olarak ifade edilir. Örneklemdeki dağılımın ait olduğu evrendeki dağılımı yansıtma derecesi evren geçerliği olarak adlandırılır. Bir ortamda yürütülen bilimsel araştırma sonuçlarının diğer ortamlara genellenebilirlik derecesi ise ekolojik geçerlik olarak adlandırılır. Yapay sosyal ortamlarda ya da laboratuarlarda yürütülen araştırmaların dış geçerliği, özellikle de ekolojik geçerliği genellikle düşüktür. Dış geçerliği etkileyen başlıca faktörler insanların denek olduklarını bildikleri zaman normalden farklı davranma eğilimi göstermeleri, yanlış örneklem seçimi, kavramların yeterince iyi tanımlanmaması ve kavramları yeterince iyi ölçmeyen değişkenler oluşturulmasıdır (Ercan & Kan, 2004) Kaynakça Bryman, A. (2004). Social Research Methods (2nd edition). Oxford: Oxford University Press. Ercan, İ., & Kan, İ. (2004). Ölçeklerde Güvenirlilik ve Geçerlilik. Uludağ Üniversitesi Tıp Fakültesi Dergisi, 30(3), 211-216. Ergin, D. Y. (1995). Ölçeklerde Güvenirlilik ve Geçerlik. M.Ü.Atatürk Eğitim Fakültesi Eğitim Bilimleri Dergisi(7), 125-48. Öncü, H. (1994). Eğitimde Ölçme ve Değerlendirme. Ankara: Matser Basım San. Ve Tic. Ltd. Şti. Sencer, M., & Sencer, Y. (1978). Toplumsal Araştırmalarda Yöntembiiim. Ankara: TODAİ. Eğitimde, psikolojide ve sosyal bilimlerde, ölçüt geçerliliği kavramı büyük önem taşır. Bu kavram, bir testin veya ölçüm aracının belirli bir ölçütle ne derece uyumlu olduğunu ifade eder. Peki, Ölçüt Geçerliliği Nedir? sorusunun yanıtı nedir ve bu kavram pratikte nasıl kullanılır? Ölçüt geçerliliği, bir ölçüm aracının, ölçmeyi amaçladığı kavram veya yapı ile ne kadar doğru bir ilişki kurduğunu belirler. Bu geçerlilik türü, ölçüm sonuçlarının gerçek dünyadaki ilgili bir performans veya davranışla olan ilişkisini inceleyerek değerlendirilir. Ölçüt geçerliliği, eğitimciler, psikologlar ve araştırmacılar için neden bu kadar kritiktir? Çünkü doğru sonuçlar elde etmek, güvenilir veri toplamak ve bunları analiz ederek anlamlı çıkarımlar yapmak, çalışmalarının temelini oluşturur. Ölçüt Geçerliliği Nedir Ölçüt geçerliliği nedir sorusuyla başlayalım. Ölçüt geçerliliği, bir testin veya ölçüm aracının, belirli bir kriter veya standartla ne kadar iyi ilişki kurduğunu gösteren bir geçerlilik türüdür. Bu tür geçerlilik, test sonuçlarının dış bir ölçütle olan korelasyonunu değerlendirir ve testin amacına uygun olup olmadığını belirler. Ölçüt geçerliliği, eğitim, psikoloji, sosyoloji gibi birçok alanda yaygın olarak kullanılır. Ölçüt geçerliliği, iki ana türe ayrılır: eşzamanlı geçerlilik ve yordama geçerliliği. Eşzamanlı geçerlilik, test sonuçlarının aynı anda alınan bir dış ölçütle olan ilişkisidir. Yordama geçerliliği ise test sonuçlarının gelecekteki bir performans veya ölçütle olan ilişkisidir. Bu ayrım, ölçüt geçerliliğinin kullanım alanına ve amacına göre önemlidir. Ölçüt geçerliliği nedir sorusunun cevabı, testlerin doğruluğunu ve güvenilirliğini değerlendirmede büyük bir rol oynar. Doğru bir ölçüm aracı, elde edilen sonuçların güvenilir ve geçerli olmasını sağlar. Bu nedenle, eğitimde ve diğer alanlarda ölçüt geçerliliği büyük bir önem taşır. Test sonuçlarının güvenirliği, bireylerin yeteneklerini, bilgilerini ve performanslarını doğru bir şekilde değerlendirmemizi sağlar. Ölçüt Geçerliliği Türleri Ölçüt geçerliliği türleri, testlerin doğruluğunu ve güvenilirliğini belirlemede kullanılan farklı yöntemlerdir. Bu türler, testin amacına ve kullanılacağı alana göre değişiklik gösterir. Ölçüt geçerliliğinin iki ana türü vardır: eşzamanlı geçerlilik ve yordama geçerliliği. Aynı anda alınan bir dış ölçütle test sonuçları arasındaki ilişkiyi değerlendirir. Mevcut performansın doğru ölçülüp ölçülmediğini belirler. Örneğin: Bir öğrencinin matematik testinden aldığı puan ile dönem sonu notu arasındaki ilişki. Test sonuçlarının gelecekteki bir performans veya ölçütle olan ilişkisini inceler. Gelecekteki başarı veya performansı tahmin etme yeteneğini değerlendirir. Örneğin: Üniversite giriş sınavı sonuçlarının üniversitedeki akademik başarıyı tahmin etmesi Bu iki tür, testlerin geçerliliğini farklı açılardan değerlendirir ve her biri, testlerin belirli kriterlerle ilişkisini ortaya koyar. Bu nedenle, testlerin geçerliliğini ve güvenilirliğini sağlamak için her iki türün de dikkate alınması önemlidir. Ölçüt geçerliliği nasıl ölçülür sorusuna verilecek yanıt, testlerin doğruluğunu ve güvenilirliğini belirlemek için kullanılan yöntemleri kapsar. Ölçüt geçerliliği, test sonuçlarının belirli bir dış ölçütle olan korelasyonunu değerlendirerek ölçülür. Bu süreçte kullanılan yöntemler arasında korelasyon analizi ve regresyon analizi bulunur. Korelasyon analizi, test sonuçları ile dış ölçüt arasındaki ilişkiyi belirlemek için kullanılır. Bu analiz, iki değişken arasındaki ilişkiyi ve bu ilişkinin yönünü ve gücünü belirler. Korelasyon katsayısı, -1 ile +1 arasında bir değer alır ve bu değer, iki değişken arasındaki ilişkiyi açıklar. Pozitif bir korelasyon, test sonuçları ile dış ölçüt arasındaki pozitif ilişkiyi, negatif bir korelasyon ise negatif ilişkiyi gösterir. Regresyon analizi, test sonuçlarının dış ölçütü ne kadar iyi tahmin edebileceğini belirlemek için kullanılır. Bu analiz, bağımsız değişkenin bağımlı değişken üzerindeki etkisini inceler ve bu etkiyi ölçmek için bir model oluşturur. Regresyon analizi, yordama geçerliliğini değerlendirmek için yaygın olarak kullanılır. Kısa bir tablo ile özetleyecek olursak; YöntemAçıklamaKorelasyon AnaliziTest sonuçları ile dış ölçüt arasındaki ilişkiyi belirler.Regresyon AnaliziTest sonuçlarının dış ölçütü ne kadar iyi tahmin edebileceğini değerlendirir.Ölçüt Geçerliliği Nasıl Ölçülür? Bu analizler, testlerin geçerliliğini ve güvenilirliğini sağlamak için kritik öneme sahiptir. Doğru bir geçerlilik analizi, testlerin belirli bir kriterle olan ilişkisini ortaya koyar ve bu ilişkilerin doğruluğunu değerlendirir. Ölçüt geçerliliği ve güvenilirlik arasındaki farklar, testlerin değerlendirilmesinde önemli rol oynar. Her iki kavram da testlerin doğruluğunu ve güvenilirliğini belirlemek için kullanılır, ancak farklı açılardan değerlendirilirler. Geçerlilik, bir testin ölçmek istediği şeyi ne kadar doğru ölçtüğünü belirler. Ölçüt geçerliliği, bu bağlamda test sonuçlarının belirli bir dış ölçütle olan korelasyonunu değerlendirir. Geçerlilik, testlerin amacına uygun olup olmadığını ve doğru sonuçlar verip vermediğini belirler. Güvenilirlik ise bir testin tutarlılığını ve tekrar edilebilirliğini ifade eder. Bir testin güvenilir olması, her uygulandığında benzer sonuçlar vermesi anlamına gelir. Güvenilirlik, testin iç tutarlılığını ve zaman içindeki istikrarını ölçer. Bu, test sonuçlarının tesadüfi hatalardan arındırılmış olduğunu gösterir. Geçerlilik ve güvenilirlik arasındaki temel fark, geçerliliğin testin doğruluğunu, güvenilirliğin ise tutarlılığını ölçmesidir. Bir testin geçerli olması, aynı zamanda güvenilir olmasını gerektirir, ancak bir testin güvenilir olması, mutlaka geçerli olduğu anlamına gelmez. Bu nedenle, testlerin hem geçerliliğini hem de güvenilirliğini değerlendirmek önemlidir. Bu kavramlar, testlerin doğruluğunu ve güvenilirliğini sağlamak için kritik öneme sahiptir. Doğru bir test, hem geçerli hem de güvenilir olmalıdır. Eğitimde ölçüt geçerliliğinin önemi, öğrencilerin performansını doğru bir şekilde değerlendirmek ve eğitim süreçlerini iyileştirmek için kritiktir. Ölçüt geçerliliği, testlerin ve ölçüm araçlarının doğruluğunu ve güvenilirliğini sağlar, bu da eğitimde yüksek kaliteli değerlendirmeler yapılmasına olanak tanır. Ölçüt geçerliliği, öğrencilerin akademik performansını ve yeteneklerini doğru bir şekilde değerlendirmek için kullanılır. Doğru bir değerlendirme aracı, öğrencilerin güçlü ve zayıf yönlerini belirlemeye yardımcı olur ve bu bilgiler doğrultusunda eğitim programları geliştirilebilir. Ölçüt geçerliliği, eğitim programlarının etkinliğini değerlendirmek için de kullanılır. Programların başarısını ölçmek, programların iyileştirilmesi ve öğrenci başarılarının artırılması için önemlidir. Geçerli bir değerlendirme aracı, programların güçlü ve zayıf yönlerini belirlemeye yardımcı olur. Öğretmenlerin performansını değerlendirmek için de ölçüt geçerliliği kullanılır. Öğretmenlerin eğitim süreçlerindeki etkisini belirlemek, öğretim yöntemlerini ve stratejilerini geliştirmek için önemlidir. Bu değerlendirmeler, öğretmenlerin profesyonel gelişimine katkıda bulunur. Eğitimde ölçüt geçerliliği, yüksek kaliteli eğitim süreçlerinin ve doğru değerlendirmelerin temelini oluşturur. Bu nedenle, eğitimde kullanılan testlerin ve değerlendirme araçlarının geçerli ve güvenilir olması büyük bir önem taşır. Üniversite giriş sınavlarının yordama geçerliliği Matematik test sonuçları ile öğretmen değerlendirmeleri arasındaki ilişki İşe alım testlerinin eşzamanlı geçerliliği İş performansı değerlendirmeleri Depresyon ölçeğinin yordama geçerliliği Anksiyete testlerinin eşzamanlı geçerliliği Öğrenci değerlendirme testleri Eğitim programlarının etkinlik ölçümleri Ölçüt geçerliliği örnekleri, bu kavramın farklı alanlarda nasıl uygulandığını ve ne kadar önemli olduğunu göstermektedir. Doğru ve güvenilir değerlendirmeler yapmak, her alanda başarıyı ve kaliteyi artırır. Ölçüt geçerliliği, bir testin belirli bir dış ölçütle olan korelasyonunu ve ilişkiyi değerlendiren bir geçerlilik türüdür. Ölçüt geçerliliği türleri, eşzamanlı geçerlilik ve yordama geçerliliği olarak ikiye ayrılır. Eşzamanlı geçerlilik, test sonuçlarının aynı anda alınan bir dış ölçütle olan ilişkisidir. Yordama geçerliliği, test sonuçlarının gelecekteki bir performans veya ölçütle olan ilişkisidir. Ölçüt geçerliliği, korelasyon analizi ve regresyon analizi gibi yöntemlerle ölçülür. Yazımızı beğendiyseniz, aşağıdan oylayabilir ve yorumlardan fikirlerinizi belirtebilirsiniz

- wixa
- gupiluba
- common core math stupid
- http://minhtiensteel.com/upload/ckfinder/files/79733941234.pdf
- http://inbonehospital.com/ckupload/files/niwudiraj.pdf
- https://gyn-koe70.de/obrazky/file/71373520782.pdf
- wiza