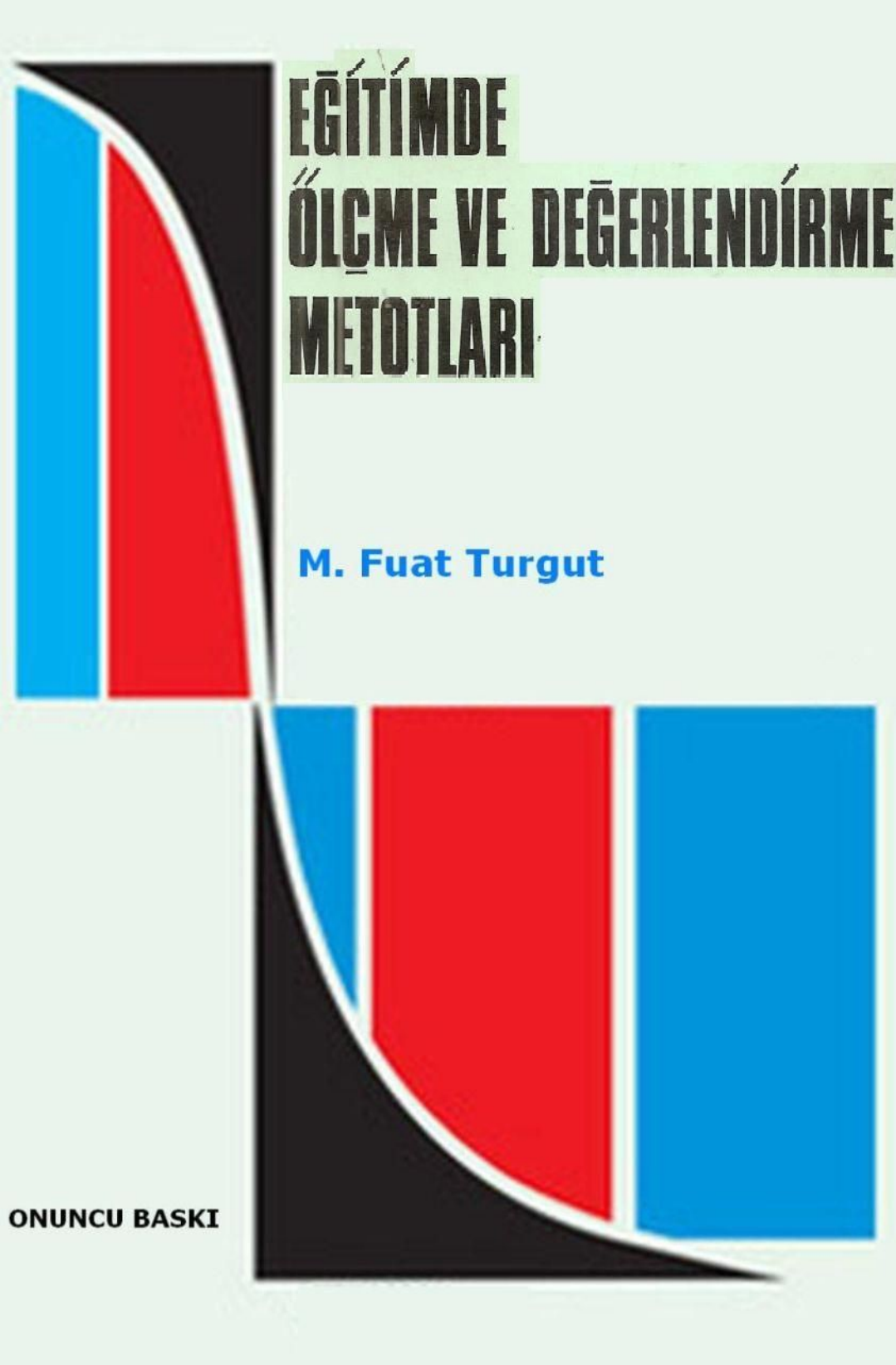




EĞİTİMDE ÖLÇME VE DEĞERLENDİRME METOTLARI

M. Fuat Turgut

ONUNCU BASKI

EGİTİMDE ÖLÇME VE DEĞERLENDİRME METOTLABI

ONUNCU BASKI

Prof.Dr. M. FUAT TURGUT

ANKARA - 1995

Her hakkı yazarına aittir.

Birinci baskı,	1977
İkinci baskı,	1983
Üçüncü baskı,	1984
Dördüncü baskı,	1986
Beşinci baskı,	1987
Altıncı baskı,	1988
Yedinci baskı,	1990
Sekizinci baskı,	1992
Dokuzuncu baskı,	1993
Onuncu baskı,	1995

ISBN - 975 - 95915 - 0 - 2



YARGICI MATBAASI

Tel : 229 68 08 Fax : 229 68 09
Kızılay-Ankara

Ö N S Ö Z

Geçen beş yılda, **Eğitimde Ölçme ve Değerlendirme Metotları**, öğretmen yetiştirme programlarının temel kitaplarından biri haline gelmiş; öğretmenler tarafından da sık sık baş vurulan bir eser olmuştur. Bu ilginin özendirici etkisiyle, kitabın ikinci baskısına karar verilmiştir.

Bu kitap, öğretmenlere ve öğretmen adaylarına, öğrenci başarısını değerlendirmede kullanacakları bilgi ve becerileri kazandırmak amacıyla yazılmıştır. Kitapta, değerlendirmeyi planlama, ölçme araçlarını hazırlama, sınavları uygulama ve puanlama, ölçme sonuçlarını yorumlama ve not verme, notları çeşitli eğitim kararlarında kullanma gibi işlerde öğretmen için gerekli tekniklere yer verilmiştir. Öğretmenler için yararlı olmakla birlikte, zekâ ve kişilik değişkenlerini ölçme yöntemleri, daha çok uzmanlık gerektirdiğinden, bu kitaba alınmamıştır.

İkinci baskı hazırlanırken, bölümlerin hepsi kapsam ve dil yönünden geliştirilmiş, örnekler çeşitlendirilmiş; kitaba, «Değerlendirme Sonuçlarının Kullanılması» adıyla yeni bir bölüm eklenmiştir.

Son yıllarda ülkemizde de bir uzmanlık dalı haline gelmiş olan Ölçme ve Değerlendirme alanında oldukça yeni bir telif olan bu kitabın, eğitim kitaplığımızdaki yerini korumaya devam edeceği; öğretmenlerimizin, öğrenci başarısını değerlendirmedeki davranışlarını etkileyerek mesleğe katkıda bulunacağı umulur.

Kitabın birinci baskısını derslerinde kullanıp ikinci baskının daha mükemmel olması için değerli önerilerde bulunan Dr. Yaşar Baykul ile Dr. Nizamettin Koç'a; ikinci baskıyı hatasızca gerçekleştiren Şükrü Saydam yönetiminde Erol Çevik, Hayrettin Buyraz ve Hasan Mercan'dan oluşan basım ekibine; provaları birkaç kez okuyarak hataları mümkün olduğu kadar azaltan eşim Perihan'a teşekkürü zevkli bir görev sayarım.

Ankara/Mart, 1983

M. Fuat TURGUT

İçindekiler

Bölüm 1 — Giriş	1
Bölüm 2 — Ölçmenin Temelleri	7
Bölüm 3 — Araç ve Yöntemlerin Nitelikleri	27
Bölüm 4 — Yazılı Yoklama	47
Bölüm 5 — Kısa Cevap Gerektiren Testler	63
Bölüm 6 — Sınıflama Gerektiren Testler	75
Bölüm 7 — Seçme Gerektiren Testler	86
Bölüm 8 — Sözlü Sınavlar	109
Bölüm 9 — Hedef Analizi ve Test Maddeleri	116
Bölüm 10 — Bilişsel Davranışların Ölçülmesi	128
Bölüm 11 — Duyuşsal Davranışların Ölçülmesi	149
Bölüm 12 — Psiko - Motor Becerilerin ve Ürünlerinin Ölçülmesi	161
Bölüm 13 — Sınavlar ve Sınav Planları	171
Bölüm 14 — Ölçme Araçlarını Çoğaltma Uygulama ve Puanlama	185
Bölüm 15 — Ölçme Sonuçları Üzerinde İstatistiksel İşlemler	194
Bölüm 16 — Değerlendirme ve Not Verme	220
Bölüm 17 — Değerlendirme Sonuçlarının Kullanılması	249
Bölüm 18 — Madde Analiziyle Test Geliştirme	261
Kaynaklar	276
Ekler	284
Bulduru	289

Şekil ve Tablolar

Ş E K İ L

15.1	Histogram	197
15.2	Frekans Poligonu	198
15.3	Yığılmalı Frekans Grafiği	198
15.4	Simetrik Bir Dağılım	209
15.5	Pozitif Kayışlı veya Sağı Çarpık Bir Dağılım	209
15.6	Negatif Kayışlı veya Solu Çarpık Bir Dağılım	210
15.7	Normal Dağılım	211
15.8	Bir Normal Dağılımda Çeşitli Puanlar	215

T A B L O

4.1	Puanlama Anahtarı	62
4.2	Kompozisyon Puanlama Anahtarı	66
9.1	Bilişsel Davranışlar	122
9.2	Duyuşsal Davranışlar	123
9.3	Psiko - Motor Davranışlar	124
9.4	Belirtke Tablosu	125
12.1	Mikroskop Kullanma Davranışları Çetelesi	166
12.2	Mikroskop Kullanma Davranışlarını Dereceleme Aracı	167
12.3	Kitap Cildi İçin Analitik Anahtar	169
15.1	Sıralanmış Puanlar	194
15.2	Puanların Gruplanması	195
15.3	Ortalama ve Standart Kayma Hesabı İçin Puanların Kodlanması	202
15.4	Korelasyon Katsayısının Sapma Puanlarından Hesabı İçin Tablolama	216
15.5	Sıra Korelasyonunun Hesabı İçin Tablolama	219
16.1	Ölçme Güvenirliğine Bağlı Olarak Değerlendirme Güvenirliği	235
16.2	Mutlak Başarı Yüzdeleriyle Verilen Notlar	236
16.3	Mutlak Başarı Yüzdeleriyle Diğer Bir Değerlendirme	237
16.4	Ortalama Başarı Düzeyi İle Bağlı Değerlendirme	241
16.5	Ham Puanların Yığılmalı Yüzdelerinin Hesaplanması	243
16.6	Yüzdeleri Saptanmış Bir Dağılımla Not Verme	244
E. 1	Çift Serili Korelasyonun Bir Dağılımın Uçlarından Kestirilmiş Değerleri	286

B Ö L Ü M 1

Giriş

Eğitim öğrencide istenilen davranışları geliştirmek, kusurlu davranışları düzeltmek, istenmeyen davranışları silmek gibi amaçlarla yapılır. İlkokul veya lise öğrenimi, meslek öğrenimi, vb. eğitimin genel örnekleridir. Kurslarda daktilo, piyano veya dans öğretilmesi; bir kimsenin Müslümanlaştırılması veya sosyalleştirilmesi ya da bir kimseye Atatürkcü politik görüş kazandırılması da eğitimin çeşitli örnekleridir. Eğitime tabi tutulan kişilerde, eğitim sürecinin sonucu olarak, planlanan davranış değişikliklerinin meydana gelmesi beklenir. Eğitim işleminin tamamlanmasıyla bu davranışlar eğitilenlere kazandırılmış olabilir. İstenilen davranışların hiç oluşmaması, oluşsa bile istenilen derecede bir davranış değişikliği olmaması da mümkündür. İstenilen niteliklerde bir davranış değişikliği uygulanan eğitimin başarısına, beklenen davranış değişikliğinin gerçekleşmesi de başarısızlığına kanıttır.

Uygulanan bir eğitimin başarılı olup olmadığının, başarılı ise ne derecede ve hangi öğrenciler için başarılı olduğunun bilinmesi istenir. Eğitim işlemi devam ederken, başarısızlığın ve başarısız bireylerin erken tanınması, önlem alınmasını kolaylaştırır. Başarı derecesinin bilinmesi ve başarısızlık hallerinin ortaya çıkarılması, ileride girişilecek benzer eğitim etkinliklerinin daha gerçekçi esaslarla planlanmasına yardım eder. Fakat, başarı derecesinin bilinmesi ve başarısız olanların tanınması, ancak, öğrencilerde meydana gelen davranış değişikliklerinin ölçülüp değerlendirilmesiyle mümkün olur.

Herhangi bir eğitim dönemi ardışık basamaklardan oluşur. Bir eğitim sistemi çeşitli düzeylerde okullara, bir okul sınıf veya yıllara, bir yıl ise daha kısa alt dönemlere bölünür. Eğitime kapsam açısından bakılınca da, ardışık bölümlerden oluşmuş ders veya programlar görülür. Öte yandan, öğrenciler okullara alınırlar, başarılarına göre ardışık basamaklarda ilerletilirler. Bazı hallerde bir kısım öğrencilere programın tekrarlatıldığı, bir kısım öğrencilerin sistemden çıkarıldığı olur. Bütün bu işlemler öğrenci hakkında verilen bazı ka-

rarların sonucunda yapılır. Öğrenciler hakkındaki bu çeşit eğitim kararları, çoğu zaman, öğrenci başarısına ilişkin bir takım ölçmele-re ve değer yargılarına dayanır. Ne yazık ki, başarıya dayanan ka-rarlar her zaman isabetli olmaz. İsbetli bir karar verilebilmesi için karara dayanak olacak kadar yeterli bilgi toplanması, toplanan bil-gilerde önemli hatalar bulunmaması ve bu bilgilerin doğru yorum-lanması gerekir.

Herkes şu veya bu konuda karar verme durumunda kalmış ve çeşitli kararlar vermiştir. Buna rağmen, karar işleminin mahiyeti, bir kararda ne gibi öğeler bulunduğu, kararların hangi koşullarla doğru olabileceği pek az kişi tarafından bilinir. Karar işleminin ve öğele-rinin genelde çözümlenmesi, karar verme süreci hakkında önemli bilgiler verir. Bir coğrafya öğretmeninin bütünleme sınavında kullan-mak üzere dört yazılı yoklama sorusu hazırladığını ve bunların doğ ru cevaplarına sırasıyla en çok 10, 30, 40, 50 puan vermeyi kararlaşt ırdığını düşünelim. Öğretmenin buraya kadar yaptığı iş, bir ölçme aracı hazırlamak ve o araç için bir ölçme kuralı saptamak olmuştur. Öğretmen sınavı uyguladıktan sonra, bir öğrencinin cevaplarına sırasıyla 10, 0, 30, 20 puanlarını vermiş olsun. Öğretmenin bu aş a-mada yaptığı iş bir ölçmedir. Bundan sonra öğretmen öğrenciye, $10+0+30+20=60$ ve $60/10=6$ işlemleriyle, «altı» notunu vermiş ol sun. Öğrenciye not verilmesi bir değerlendirme işlemidir. Sonra bu not okul idaresine gider; orada, sınav yönetmeliğinin «Bütünlemede en az beş notu geçer.» hükmüne göre, öğrencinin coğrafya dersin-den geçtiği yargısına varılır. Bu değer yargısına ulaşma da bir de ğerlendirmedir. Bu değer yargısı, varsa öteki bütünlemeli dersler-deki değer yargılarıyla birlikte, öğrenci hakkındaki sınıf geçme ka-rarına dayanak olur.

Her kararın gerisinde bir değer yargısı bulunur. Değer yargıla-rına, gözlem veya ölçme sonuçlarının bir ölçüte vurulmasıyla ulaş ı-lır. Yukarıdaki örnekte öğretmenin $60/10=6$ işlemiyle yaptığı de ğer-lendirmede kullandığı ölçüt, «tam puan 100» veya «tam not 6» esasına göre, o sınavda sorulan soruların tümü olmuştur. Okul idaresinin «coğrafyadan geçer» yargısına da, $6>5$ işlemiyle ve yönetmeliğin koyduğu «en az beş» ölçütüyle ulaş ılmıştır.

Öğrenciler hakkındaki eğitim kararları, başarının yanında baş ka değer yargılarına ve verilere de dayanmak zorundadır. Fakat, öğrenci başarısı hakkındaki değer yargıları, eğitim kararlarının en önemli dayanağıdır. Bu kararların isabetli olması için, öğrenci ba-şarısı hakkında yeterince doğru bir değer yargısına ulaşılması iste-

nir. Bir değer yargısının doğruluğu ise, hem dayandığı gözlem veya ölçmelerin doğruluğuna, hem de seçilen değerlendirme ölçütünün uygunluğuna bağlıdır. Fakat, bugün ülkemizde ölçme ve değerlendirmenin iyi bilinmemesi, ölçme sonuçlarıyla ölçütlerin birbiriyle karıştırılmasına, yetersiz dayanaklarla isabetsiz kararlar verilmesine sebep olmaktadır. Bazen güvenilir olmayan bir ya da birkaç ölçme sonucuna dayanarak karar verildiği; bazen de geçerli olmayan bir ölçüt seçildiği, hatta ölçütün yanlış kullanıldığı görülmektedir.

Ölçme ve değerlendirme, genel kullanımda birbirinden ayrı anlam taşıyan iki kavramdır. Öğretmenlerimizin ve üst düzeydeki bazı eğitimcilerimizin bu kavramları birbirine karıştırarak kullandıkları görülmektedir. Bu yanlışlık, bazı Batı kaynaklarında değerlendirme kavramına, ölçmeyi de içine alacak kadar geniş anlam verilmesinden doğmuştur. Bu kitapta ölçme ve değerlendirme kavramlarının ayrı ayrı kullanılmasına özen gösterilmiştir.

Ölçme, geniş anlamıyla, herhangi bir niteliği gözlemek ve gözlem sonucunu sayılarla ya da başka sembollerle ifade etmektir. Ölçmede eşyanın, olayların veya insanların ölçmeye konu olan niteliklerinin gözlenmesi, sayılması veya bir ölçme aracıyla karşılaştırılması vardır. Ölçme işleminin sonucunda da, ölçmeye konu olan niteliğin ya sayısı, ya derece, ya da sıfatlarla gösterilmesi vardır. Ölçmelere her zaman hata karışır. Bir ölçme sonucunun güvenilir olması, ölçme hatalarının önemsiz dereceye indirilmesiyle sağlanır. Değerlendirme ise, ölçme sonuçlarını bir ölçüte vurarak, ölçülen nitelik hakkında bir değer yargısına varma sürecidir. Değerlendirmenin güvenilir ölçme sonuçlarına dayanması, geçerli bir ölçütle yapılması, değer yargısına ulaşma işlemlerinde yanlışlık bulunmaması doğru bir değer yargısına ulaşmak için gereklidir. Değerlendirmenin dayanacağı ölçme sonuçlarının hatalarını azaltmak, eğitimdeki niteliklerin ölçülmesinde çok zordur. Değerlendirmenin kullanılacağı eğitim kararına uygun bir ölçüt seçimi ise daha zordur. Çünkü, uygun bir ölçüt seçme işleminde, eğitimin amaçlarından öğrencinin özgeçmişine kadar birçok etkenin dikkate alınması gereklidir.

Eğitim kararları öğrenci, öğretmen, programlar, ders araçları, okullar, vb. hakkında olabilir. Kararlara öğrencilerle ilgili işlemler açısından bakılınca, değerlendirmenin en az iki görevi ortaya çıkar. Bu görevlerden biri öğrencileri başarılarına göre sınıflamaktır. «geçti - bütünülemeli - kaldı», «zayıf - orta - iyi» gibi kararlar veya öğretim amaçları için öğrenci başarısına dayanan diğer gruplama işlemleri birer sınıflamadır. İkinci görev öğrenciyi tanımaya yardımcı

olmaktır. Öğrenciyi tanıma daha ziyade rehberlik amaçlarına hizmet ettiğinden, bu göreviyle değerlendirme, öğrencinin eksik yönlerini, kuvvetli olduğu alanları, özel yeteneklerini, öğrenme güçlüklerini ortaya çıkarmaya çalışır.

Sınıftaki öğretimin iyi yürütülebilmesi için, öğrenci başarısının sık sık ölçülüp değerlendirilmesi gerekir. Böyle olunca da öğretmene bu konuda önemli sorumluluklar düşer. Gerçekten, bir sınıftaki öğrencilerin herhangi bir derste başarıları ölçülüp değerlendirilecekse, o dersin öğretmeni, ölçme ve değerlendirme becerilerine sahip olmak kaydıyla, bu işi yapacak en yetkili kimsedir. Onun için, öğrenci başarısının ölçülmesi ve değerlendirilmesi konuları, öğretmenlik meslek programlarının önemli bir parçası haline gelmiştir. Öğretmen, bu görevi yerine getirebilmek için, ölçme ve değerlendirme adıyla bilinen alanın bazı özel bilgilerini rahatlıkla kullanabilecek derecede iyi bilmek, bu alanda bazı beceriler geliştirmiş ve olumlu tutumlar kazanmış olmak zorundadır.

Ölçme ve değerlendirme alanının bilgi, beceri ve tutumlarının bir öğretmene kazandırılması, bu alandaki uzmanların yetişmesi için gerekli eğitim kadar kapsamlı ve zor değildir; fakat, sınav - yanılma yoluyla kazanılabilecek kadar kolay veya bir folklor olmaktan ileri geçmeyen geleneksel uygulamaların öğrenilmesi ve onların gerektirdiği becerilerin kazanılmasıyla yetinilebilecek kadar da dar değildir. Bu bilgi ve beceriler, meslek öncesinde haftada birkaç saatlik bir yıl süreli bir dersle, meslek içinde teksifi bir kursla, hatta meraklı öğretmenlerin kendi gayretleriyle kazanılabilir.

Ölçme ve değerlendirme alanının özel bilgilerini kazanmış bir öğretmenle, bu bilgileri kazanmamış bir öğretmen arasında önemli farklar vardır. Ölçme ve değerlendirme tekniklerine uyarak not veren bir öğretmen, hem daha az hatalı değer yargılarına ulaşmış olur, hem de kendi öğretim yöntemlerini değerlendirip geliştirebilir. Çünkü, iyi bir ölçme, ders programının ve öğrenme sürecinin, ayrıntılarına kadar analiziyle başlar. İşte bu analiz, öğretmenin, kendi öğretim etkinliklerini geliştirmesine yardım eder.

Şüphesiz, eğitim kararları, sadece öğretmenin kendi öğrencileri hakkında verdiği kararlardan ibaret değildir. Öğrenci seçimi, program değerlendirme, araştırma gibi etkinliklerde kullanılacak ölçme araçlarının hazırlanması, daha ileri uzmanlık bilgi ve becerileri gerektirir. Öğretmenin bu alandaki araçlardan ve araştırma sonuçlarından yararlanması beklenir. Bunun için de öğretmen, uzmanlık ürünlerinden yararlanabilecek yetişkinliğe ulaşmalıdır.

Eğitimde birçok değişken ölçülebilir. Ölçülen büyüklükler çoğu halde öğrencilerin bir takım nitelikleridir. Bu sebeple, ölçme işlemi her zaman ölçülecek büyüklüğün tanımlanmasıyla başlar. Ölçme, araç gerektirir. Ölçme araçları, ölçülecek büyüklüğün gözlenip sayılarla ifadesini kolaylaştırır. Her araç, bir bakıma, gözlemin daha duyarlı yapılmasını sağlar. Eğitimde kullanılan ölçme araçları, ölçülecek büyüklüğün tabiatına ve aracın kullanılacağı gruba göre çeşitlilik gösterir. Ölçmeyi yapacak kimse, değerlendirmenin amacına uygun bir araç seçmek, yoksa böyle bir aracı bizzat yapmak ihtiyacını duyar.

Eğitimde kullanılan ölçme araçları, yapımlarının gerektirdiği uzmanlık ve beceriler bakımından çeşitlilik gösterir. Seçme sınavlarında, rehberlikte, kliniklerde kullanılan ölçme araçlarının yapımı yüksek düzeyde uzmanlık ister. Hatta, bu araçlardan bazılarının kullanılması bile özel bilgi ve becerileri gerektirir. Öte yandan, sınıfta öğrenci başarısının ölçülmesinde kullanılan yazılı yoklama ve testler gibi ölçme araçlarının hazırlanmasında ve uzmanlarca hazırlanan standart başarı testleri gibi araçların kullanılmasında daha az özel bilgi ve becerilere gerek vardır.

Ölçme ve değerlendirme sonuçları, çoğu zaman idarecilere, öğrenci velilerine ve öğrencilere bildirilir. Bu kimseler, genellikle, ölçme sonuçlarını kendi başlarına yorumlayıp bir anlam çıkaracak bilgi ve becerilere sahip değildirler. Onun için, ölçme ve değerlendirme sonuçlarının ilgililer için anlamlı bir dille verilmesi zorunlu olur. Bu iş de ölçme ve değerlendirmeyi yapan öğretmene düşer.

Eğitimdeki ölçme araçları hem kullanılır, hem de uygulamadan alınan sonuçlara dayanılarak analiz edilip geliştirilir. Örneğin, bir yazılı yoklamada alınan cevapların önceden hazırlanmış bir anahtarla puanlanması sırasında, sorularda ve anahtarda kusurlar görülebilir. İleride kullanılmak üzere, sorular düzeltilebilir ve anahtar yeniden hazırlanabilir. Böyle analiz ve düzeltme işlemleri testlerde daha çok kullanılır. Ölçme araçlarını analiz etme, kusurlarını bulup düzeltme, bu işlemlerde özel yetiştiklikleri olan öğretmenlerce daha iyi ve daha kolayca yapılır.

Yukarıdaki tartışmadan anlaşılacağı gibi, öğretmenlerden bu alanda uzmanlık istenmese bile, öğrenci başarısının sınıftaki etkinlikler arasında ölçülüp değerlendirilmesiyle ilgili bazı önemli sorumluluklar yüklenmeleri istenir. Bu sorumlulukların gerektirdiği görevler ise, ancak bu alana özgü bazı bilgi ve becerileri kazanmış öğretmenler tarafından hakkıyla yerine getirilebilir. Öğretmenin bu

konudaki sorumlulukları, öğrenci başarısının belirleyicisi olarak ne gibi davranışların ölçüleceğini tayin etmekten, ölçme araçlarının geliştirilip uygulanmasına, ölçme sonuçlarının yorumlanıp değerlendirilmesine ve ilgililere bildirilmesine kadar uzanır.

Bu kitabın amacı, öğretmen adaylarına ve kendilerini yetiştirmek isteyen öğretmenlere, bu alanda, bir kitaptan yararlanılarak ulaşılabilecek bir düzeye erişmelerinde yardımcı olmaktadır.

Bu kitabın iç düzeni, ölçme ve değerlendirme alanının Batıda görülen kitaplarının gelenekselleşmiş düzenlerinden farklıdır. Gelenekselleşmiş kitaplar bu konuya ya İstatistik bölümleriyle, ya da test planlarıyla girerler. İstatistik, ölçme tekniğinin konusu değil, gerekli olduğu zaman kullanılacak bir araçtır. Bir öğretmen ölçme ve değerlendirme işlerine test planıyla başlar. Fakat, bu kitabın okuyucularından çoğu, konuya test planıyla girerlerse ve gerçekçi bir plan yapmaya koyulurlarsa, henüz öğrenmedikleri konuların uygulamalarıyla karşılaşır. Bu düşünceyle, kitabın birinci bölümü, eğitimde ölçme ve değerlendirme problemlerine kuşbakışı bir göz atmaya ayrılmıştır.

İkinci ve üçüncü bölümlerde, bazı kavramların açıklanmasıyla, ölçmeye bir temel hazırlanmaya çalışılmıştır. Bunları izleyip beş bölüm, eğitimde en çok kullanılan ölçme araçlarıyla ve onların getirdiği yöntemlerle uğraşmaktadır. Daha sonraki dört bölümde eğitimde hedefler, davranışlar ve ölçme araçlarının ölçülecek davranışlara göre yapımı ele alınmıştır. On üçüncü ve on dördüncü bölümlerde sınavların hazırlanışı, uygulanışı ve puanlanışıyla ilgili teknikler tartışılmıştır.

On beşinci bölümde ölçme sonuçlarının yorumu ve değerlendirme işlemlerinde gerekli İstatistiksel yöntemler verilmiştir. Bunun amacı, okuyucuya temelli İstatistik bilgileri vermek değil, değerlendirme işlemlerinde kullanılan İstatistik yöntemlerini hatırlatmak ve pratik yollar göstermek olmuştur. On altıncı bölümde değerlendirme ve not verme konuları ayrıntılarıyla tartışılmıştır. On yedinci bölümde değerlendirme sonuçlarının çeşitli eğitim kararlarında kullanılması ve gerektiğinde ilgililere bildirilmesi ele alınmıştır. Kitap, madde analizi ve test geliştirme yöntemlerinin kısaca tanıtıldığı on sekizinci bölümle bitmektedir.

B Ö L Ü M 2

Ölçmenin Temelleri

Öğrenci başarısının değerlendirilmesi için, diğer işlerin yanında, başarının ölçülmesi gerektiğine daha önce işaret edilmişti. Değerlendirmeye hizmet edecek bir ölçmenin değerlendirme amacına uygun olması, değerlendirmenin gerektirdiği kadar duyar ve hatadan arınık olması istenir. Bu niteliklerde bir ölçme yapılabilmesi için ölçmenin temel kavramları ve ilkeleri ana çizgileriyle bilinmelidir. Bu bölümde ölçmenin temel kavramları örneklerle tartışılmaktadır.

Ölçmenin Önemi

Ölçmenin önemi, bilimdeki ve uygulamadaki yeri tartışmasız kabul edilebilecek kadar açıktır. Ölçmek, genellikle bir niteliği nicilemek veya sayısal ifadelerle belirtmek anlamında kullanılır. Bir niteliğin «büyük», «çok», «fena», «pekiyi» gibi sıfatlarla belirlenmesi, sayılarla ifade edilmesinden daha az kesinlik taşır. Sayılarda anlamın kesinliği, çoğu halde, ifadenin prestijini artırır. Ölçme, gözlemlerimize gerçekten kesinlik kazandırıyorsa, o konudaki bilgilerimiz de kesinleşiyor demektir. Ölçmenin bilim ve uygulamada önem kazanması işte bu kesinleştirmede yatar. Ayrıca, sayılarla yapılan işlemler de kolaylaşır. Öte yandan her bilim dalı, özellikle doğa ve toplum bilimleri, teori ile gözlem arasındaki bağıntıların derinleştirilmesiyle gelişir. Ölçme işlemleri hem gözlemleri kolaylaştıran, hem de gözlemlerin duyarlılığını artıran işlemlerdir. Bir bilim dalı daha kesin olmak yolunda ise ve bir uygulamada daha kesin verilere dayanacaksa, ölçme yapılması zorunlu olur.

Eğitimde ölçmenin önemi, eğitimin teorik bir bilim olma gayretleri yanında, pratik kararlarda değer yargılarına dayanak sağlamasından gelir. Eğitimde ölçme gayretleri ilerledikçe, duyar ölçme araç ve metodları geliştirildikçe, bir çok eğitim kararı, subjektif kanılardan ziyade objektif ölçmelere dayanmaya başlamıştır. Nitekim, testlerin ortaya çıkmasından önce, zekâ, yetenek, tutum, kişilik gibi terimlerle ifade edilen niteliklerin kişilerdeki dereceleri gözlemlerle

ancak kabaca kestirilebiliyordu. Bugün, birçok nitelik testlerle ölçülmekte, niteliğin kişideki derecesi standard birimlerle gösterilmekte bu ölçümlere dayanılarak çeşitli kararlar verilmektedir.

Ölçmenin Anlamı

Ölçmenin bir tanımı verilebilir. Ölçmeye hem kavram, hem işlem, hem de ürün olarak matematiksel bir modelle yaklaşılabilir. Böylece ölçme kavramı daha kesin bir anlam kazanır. Fakat, bu kitabın amacı ve kapsamı ölçmeye teorik açıdan bakmaya elvermez. Onun için, gerekli kavramların anlatılmasında örneklere dayanılmıştır. Aşağıdaki cümlelerin herbiri bir ölçme sonucunu ifade etmektedir. Bu örneklerin incelenmesi bazı kavramların anlaşılmasına yardım eder.

1. 6/F sınıfındaki öğrenciler sayılmış, 38 kişi bulunmuştur.
2. Sami'nin boyu ölçülmüş, 152 cm. bulunmuştur.
3. Bu sabah okul bahçesindeki termometre 16°C gösteriyordu.
4. Nermin, giriş sınavı sonucunda, tıp fakültesine on sekizinci olarak girdi.
5. Dünkü konferanslarda Ali Atar çok iyi konuştu, fakat Ömer Tekler sık sık kekeleydi.
6. Yoklama cetvelinde yatılı öğrenciler 2, gündüzlüler 1 kodu ile gösterilmiştir.

Yukarıdaki örneklerin her birinde bir nitelik gözlenmiştir. Sınıftaki öğrencilerin sayısı, Sami'nin boyu, okul bahçesinin sıcaklığı, Nermin'in tıp fakültesine giriş sırası, konuşmacıların konuşma yetenekleri, öğrencilerin yatılı veya gündüzlü oluşları gözlenen niteliklerdir. Bu niteliklerin hepsinde ortak olan özellik, söz konusu niteliğin bir kişiden diğerine veya bir zamandan diğerine sabit olmayışıdır. Sınıftaki öğrencilerin sayısı, havanın sıcaklığı, Sami'nin boyu ve diğer özellikler başka değerlerde olabilirdi. O halde, yukarıdaki örneklerde gözlenen niteliklerin her birinin bir büyüklüğü vardır. Bu büyüklük kişiden kişiye veya yerden yere farklı değerde olabilir. Çeşitli değerler alabilen bir niteliğe, daha doğrusu onun ölçülerine, matematik dilinde **değişken** denir. Bu anlamda yatılılık - gündüzlülük de bir değişkendir, fakat yalnız iki değer alabilen bir değişkendir.

Yukarıdaki örneklerin her birinde bir değişken gözlenmiş, gözlem sonucu çoğu halde bir sayı ile ifade edilmiştir. Konferansçıların konuşma yetenekleri de sayılarla ifade edilebilirdi. Konuşma yeteneğini nicellemek için «çok fena = 1, fena = 2, orta = 3, iyi = 4, çok iyi = 5»- gibi sayısal semboller kullanılabilirdi.

Ölçme sırasında gözlenen değişkenler, eşyanın, olayların veya insanların bazı nitelikleridir. Bu değişkenler dışımızdaki evrenin olgusal varlıklarında bulunur. Ölçme yapan kimse bu varlıklarda ölçmek istediği niteliği gözler ve değerini kestirir; sonra, gözlem sonucunu bir sayı veya başka bir sembolle ifade eder. Çoğu zaman, gözlenen bir değişkenin değeri sayısal bir sembolle ifade edilmiş ise ölçmeden söz edilir. Oysa biz, bazı ölçme sonuçlarını sayılarla değil, diğer sembollerle gösteririz. Sonuçların, «büyük - küçük», «iyi - orta - zayıf» gibi sıfatlarla ifade edildiği ölçme işlemleri de vardır. Ölçme sonuçlarının her zaman sayılarla ifade edilmesi zorunluğu yoktur.

Yukarıdaki ölçme örneklerine tekrar dönülürse, ölçülen değişkenin tabiatı ve kullanılan ölçme işleminin mahiyeti bakımından önemli farklılıklar görülür. Birinci örnekte, sınıftaki öğrencilerin sayısı pozitif bir tam sayı ile ifade edilmiştir. Böyle hallerde, ölçülen değişken ayrık bir değişkendir. Değişkenin bütün değerleri tam elemanlar halindedir. Ölçme işlemi, kümenin elemanlarını saymaktan ibarettir. Saymak için, sıfır ve pozitif tam sayılardan oluşan sayılar kümesi yeter. İkinci örnekte, ölçülen büyüklük bir uzunluktur. Uzunluk sürekli bir değişken olduğu için ölçme sonucunun ifade edileceği sayı sisteminin de sürekli olması gerekir. Nitekim, uzunluk ölçmede kullanılan birimlerin kesirleri de vardır. Metreden küçük uzunlukları, metrenin ve santimetrenin kesirleriyle ifade ederiz. Uzunluk gibi sürekli değişkenleri ifade etmek için, sıfır ve tam sayılardan başka, kesirleri de içine alan bir sayılar kümesi gerekir. Matematikçiler bu kümeye gerçel sayılar kümesi derler. Uzunluk ve kütle gibi değişkenlerin pozitif ve sıfır değerleri olacağı şüphesizdir. Fakat, keyfi olarak biz tanımlamamışsak, böyle değişkenlerin negatif değerleri olmaz.

Üçüncü örnekte ölçülen değişken sıcaklıktır. Sıcaklık ısının bir ölçüsüdür. Havadaki ısı ise, gaz moleküllerinin kinetik enerjilerinin bir ölçüsüdür. Bu bakımdan ısı sürekli bir değişken olarak kabul edilebilir. Öte yandan, bu örnekte ilginç başka bir özellik görülür. Havanın sıcaklığı «santigrat» derecesi denilen bir birimle ölçülmüştür. Bu birimle ölçülen sıcaklığın negatif değerler de aldığı bilinir. Sıcaklık molekül kinetik enerjisinin bir ölçüsü olduğuna göre, negatif derecelerin anlamı nedir? Bir insanın malî durumu «lira» birimiyle ölçülebilir. Bu birimde negatif sayı, o insanın «borçlu» olduğunu ifade eder. Sıcaklık halinde durum böyle değildir. Sıcaklığın «gerçek» ölçüsünde negatif sayı anlamsızdır. Nitekim, Kelvin ölçüğü adıyla bilinen sıcaklık ölçeğinin negatif dereceleri yoktur. Santigrat ve Kelvin ölçeklerindeki ölçmeler şöyle bir eşitlikle birbirine çevrilebilir :

$Y^{\circ} K = X^{\circ} C + 273$. O halde, bu iki sıcaklık ölçeği arasındaki fark sıfır noktasının seçiminden doğmuştur. Keivin ölçeğinde $0^{\circ} K$, gerçekten ısının sıfır değerine karşılık olduğu halde, santigrat ölçeğinde $0^{\circ} C$, gerçek sıfırın, ısı yokluğunun karşılığı değildir. Santigrat ölçeğinde sıfır noktası ısının gerçek sıfırında değil, yukarıdaki denklemden hesaplanabileceği gibi, $+ 273^{\circ} K$ noktasındadır.

Dördüncü örnekte, ölçülen değişken bir büyüklük değil, bir kimsenin belli bir grupta belli bir büyüklüğe göre sırasıdır. Bu örnekte ölçme sonucunu ifade eden sayı ancak bir sıra göstermektedir. Sıralamada kullanılan sayıların anlamlı bir sıfırı olmadığı gibi, her sayı bir birim de göstermez. Ölçme işlemi kişileri veya eşyayı belli bir niteliğe göre bir boyutta sıralamak suretiyle yapılmışsa, ölçme sonuçları da «sıra sayıları» ile ifade edilir.

Beşinci örnekte, ölçme sonuçları sıfatlarla ifade edilmiştir. Yukarıda belirtildiği gibi, bu durumda ölçme sonuçları 1 ile 5 arasında değişen tam sayılarla da ifade edilebilirdi. Sayıların böyle kullanılması, bu ölçmelerde bir birim bulunduğu veya sıfırın «hiç yok» gibi bir karşılığı olduğu anlamına gelmez. Çünkü «fena = 2» ile nitelenen bir konuşmanın, «iyi = 4» ile nitelenen bir konuşmanın ancak yarısı kadar mükemmel olduğu söylenemez. Böyle ölçmelerde ölçme sonuçlarını ifade etmekte kullanılan sayılar ancak sıfatların yerini tutumakla birlikte, ölçülen kişiler konuşma yeteneklerine göre derecelenmiş beş sınıfa ayrıldığı için pekiyi'nin iyi'den, iyi'nin orta'dan, orta'nın fena'dan, onun da çok fena'dan daha iyi olduğu bilinir. O halde, «pekiyi > iyi > orta > fena > çok fena» bağıntısı korunur. Böyle bir ölçme işlemi sıralamaya benzer. Ölçme sonuçları sayılarla ifade edilmişse, bu sayıların mutlak büyüklükleri anlamsız olmakla birlikte, büyüklük -küçüklük bağıntıları anlamlıdır.

Son örnekte, bütün öğrenciler «yatılı - gündüzlü» gibi iki gruba ayrılmıştır. Bu sınıflama işleminde bir sınıfın diğerinden büyük olduğu söylenemez. Böyle bir sınıflamadaki gruplar bir sınıflama boyutunda sıralanamaz.

Formal Anlamda Ölçme

Ölçmeyi daha formal bir yaklaşımla incelemek de mümkündür. Formal yaklaşımlar ölçmenin temel kavramlarını ve işlemlerini aksiyomatik modeller içinde işlerler.^{12/3} Ölçmeyi formal modellerle ince-

¹P. Suppes and J.L. Zinnes, «Basic Measurement Theory» in, R.D. Luce, R.R. Bush and E. Galanter (Eds.), *Handbook of Mathematical Psychology*, Vol. I, s. 1 - 76

²J. Pfanagl, *Theory of Measurement*.

³C.H. Coombs, R.M. Dawes and A. Tversky, *Mathematical Psychology : An Elementary Introduction*, s. 7 - 30.

leyen düşünce tarzı, aşağıdaki üç paragrafta çok kaba olarak ifade edilmiştir.

Herhangi bir nitelik, bir canlılar eşyalar veya olaylar kümesinde ölçülür. Böyle bir küme, elemanlarıyla ve elemanlar arası ilişkileriyle **ampirik ilişkiler sistemi** denilen bir sistem meydana getirir. Öte yandan, bu sistemde yapılan ölçmelerin sonuçlarını ifade etmekte kullanılan sayı veya semboller kümesi de, elemanlarıyla ve elemanlar arası ilişkileriyle, **formal ilişkiler sistemi** denilen bir küme meydana getirir. Ölçme, ampirik ilişkiler sisteminin formal ilişkiler sistemiyle gösterilmesi olarak düşünülür. Bu göstermenin esası, canlılara, olaylara veya eşyaya, onlar arasında gözlenen ilişkileri yansıtacak şekilde sayı veya semboller vermektir.

Herhangi bir alanda ölçme yapacak bir kimse, önce hangi özelliklerin nicelendirileceğini, yani ilgili ampirik ilişkiler sisteminin tanımlamalıdır. Sonra, nicelendirilecek özellikler yeterince yansıtabilecek bir sayı veya semboller sistemi seçmelidir. Daha sonra da gözlenen özelliklere bu sayı veya sembollerin hangi kuralla verileceğini belirtmelidir.

Ölçmeye böyle formal bir modelle yaklaşırsa üç temel problemle karşılaşılır. Bunların birincisi **temsîl** problemi olup, ampirik sistemdeki ilişkilerin formal sistemin işlem ve bağıtlarıyla temsîl edileceğinin kanıtlanmasını gerektirir. İkincisi **teklik** problemi olup, ampirik sistemin elemanlarıyla formal sistemin elemanları arasında yapılan bire - bir eşleme işleminin biricikliğinin kanıtlanmasını gerektirir. **Anlamlılık** problemi olan üçüncüsü ise, ölçmede kullanılan sayıların hangi nitelikte bir ölçek vereceği, ölçme sonuçlarının nasıl yorumlanacağı, ölçme sonuçlarının diğer bir sayı veya sembol sistemine çevrilmesiyle anlamlarının korunup korunmayacağı gibi soruların cevaplandırılmasını gerektirir. Bu üçüncü problem aşağıda, ölçekler kesiminde kısaca ele alınmaktadır.

Ölçmeyi aksiyomatik yoldan formalleştirmenin başlıca yararı herhangi bir ölçme işleminin önemli özelliklerini ve ölçmede kullanılan sayı veya sembol sistemin yapısal özelliklerini teorik olarak ortaya çıkarabilmektedir. Bu yolla, deneysel incelemelere teorik bir temel bulunmuş olur.

Ölçmenin Tanımı

Matematiksel tanımlardan ölçme işlemlerini de içine alan sözel tanımlara kadar, çeşitli ölçme tanımları yapılmıştır. Şu tanım, mate-

matiksel tanımlar kadar formal ve bazı sözel tanımlar kadar geniş olmamakla birlikte, bu kitapta kullanılan ölçme konuları için yeterlidir. **Ölçme, bir niteliğin gözlenip gözlem sonucunun sayılarla veya başka sembollerle gösterilmesidir.**

Bu tanım, yukarıda örneklerle açıklanan ölçme türlerinin hepsini kapsar. Sıralama ve sınıflama işlemleriyle toplanan verileri ölçme dışında bırakan tanımlar da vardır. Mesalâ, Campbell⁴ bugün klasikleşmiş olan eserinde, Jones⁵ da diğer bir eserde ölçülebilirliği tartışılırken sıralama ve sınıflamayı ölçme olarak düşünmemişlerdir. Dar anlamdaki ölçme kavramı, eğitim gibi sosyal bilim dalları için yeterli değildir. Çünkü eğitim ve psikolojideki ölçmelerin bir çoğu sınıflama ve sıralamadan ileri bilgiler vermez. Eğitimciler arasında da ölçme teorisine geniş tanımlarla yaklaşanlar vardır.⁶

Burada verilen tanımdan hareketle, ölçmenin en az üç basamağı olduğu söylenebilir. Canlıların, olayların veya eşyanın ölçülecek niteliğini diğer niteliklerinden ayırt edebilmek ölçmenin birinci basamağıdır. Ölçme işlemi açısından bu basamakta yapılacak iş, ölçülecek ampirik sistemi ölçme değişkeniyle ve ölçme boyutuyla tanımlamaktır. Ölçmenin ikinci basamağı, ölçme sonuçlarının sayı veya sembollerle gösterilmesidir. Ölçme işlemi açısından bu basamakta yapılacak iş, ölçülecek değişkene ve ölçmenin amacına uygun bir sembol veya sayılar kümesi seçmektir. Ölçmenin üçüncü basamağı, gözlenen değişkenin belli değerlerine belli sayılar verilmesidir. Bu basamakta yapılacak iş, gözlenen değişkenin hangi değerine hangi sayının verileceğini gösteren kuralı saptamaktır. Uzunluk ölçmek için arşın veya metre seçme, bir yazılı yoklamayı puanlamak için doğru cevap anahtarı hazırlama ve hangi nitelikteki bir konuşmaya «iyi» denileceğini kararlaştırma gibi işlemler, aslında, ölçme kuralının saptanmasından başka bir şey değildir.

Ölçmede Birimler

Ölçme sonuçlarını sayılarla ifade ederken, çoğu halde bir birim söz konusudur. «152 cm», «16°C», «36 öğrenci» gibi ölçme sonuçlarında cm, °C ve «öğrenci» birer birim ifade etmektedirler. Ölçmenin

⁴N.R. Campbell, *An Account of the Principles of Measurement and Calculation*.

⁵L.V. Jones, «The Nature of Measurements» in R.L. Thorndike (Ed.), *Educational Measurement*, s. 335 - 355.

⁶E.E. Cureton, «Measurement Theory», in R.L. Ebel (Ed.), *Encyclopedia of Education Research*, 4 th ed., s. 785 - 804.

önemli problemlerinden biri, ölçme işleminin yapılmasını kolaylaştıracak ve ölçme sonuçlarının kullanılacağı amaca uygun düşecek bir birim seçmektir.

Ölçülecek değişkenlerden bazılarının doğal birimleri vardır. Bir sınıftaki öğrencileri saymakta kullanılabilecek en uygun birim «bir tek öğrenci» olur. Bir sekreterin daktilo hızını ölçmek için uygun bir birim dakikada yazılan kelime sayısı olabilir. Genellikle, ayrık değişkenlerin doğal birimleri vardır. Zaman ve uzunluk gibi fiziksel büyüklükler, doğal birimlerle değil, tanımlanmış birimlerle ölçülür. Bir öğrencinin matematik dersine karşı tutumu, diğer birinin futbol ilgisi, yapılmış bir masanın mükemmelliği, el yazısının güzelliği, bir çocuğun zekâsı gibi değişkenlerin de doğal birimleri yoktur. Doğal birimleri olmayan değişkenlerin ölçülmesinde birimler, tanımlama ve anlaşma yoluyla seçilir. Metre, saat, kilovat tanımlanmış ve kullanıla gelmiş standart fiziksel birimlerdir. Birimin doğal olmaması önemli bir sakınca doğurmaz. Elmanın tanesine, insanın sayısına alıştığımız gibi, ağırlığın kilogramına, elektriğin kilovatına alışmışızdır. Hatta, doğal birimlerin kullanışsız olduğu hallerde daha kullanışlı birimler tanımlanmıştır.

Ölçmede kullanılacak birimlerde bulunması istenen üç özellik, birimlerin eşitliği, genelliği ve kullanış amacına uygunluğudur. Metre kulaçtan daha iyi bir birimdir. Çünkü metrenin yapımı ve kullanışındaki hataların sınırları içinde, bütün metreler birbirine eşittir; oysa ki, bütün kulaçlar birbirine eşit değildir. Eğitimde öğrencilerin başarılarını 0-10 gibi 11 bölmeli bir ölçekle ölçüyorsak, bu ölçekte birimlerin eşit olduğu söylenemez. Bazı öğrencilerin başarılarının ölçüsü olarak gösterilen bütün 5'ler aritmetik anlamda birbirine eşittir. Fakat bu öğrencilerin başarıları birbirine eşit midir? Hatta, bir öğretmenin bir sınıfta ve bir sınavda takdir ettiği bütün 5'lerin eşit bilgileri gösterdiği bile kesinlikle söylenemez. Bu eşitsizlik yalnızca öğretmenin ölçmede yaptığı hatadan gelmez; daha ziyade hangi derecedeki başarıya 5 takdir edileceği, yani ölçme kuralı, kesinlikle saptanamadığı için ortaya çıkar.

Birimlerin herkes tarafından aynı anlamda kullanılması da çok önemli bir zorunluktur. Metre ve yarda kesinlikle tanımlanmıştır; metre ve yarda birbirine çevrilebilir. Fakat, insanların bir kısmının yarda kullanması iletişim gücünü yaratır. Onun için birimlerde birlik, birimlerin eşitliği kadar önemlidir. Üstelik, farklı birimler kullanıldığı hallerde, biri diğerine hatasızca çevrilemeyen birimlerdeki iki ölçmeyi birbiriyle karşılaştırma olanağı da bulunamaz.

Birimin ölçülecek değişkene uygun büyüklükte seçilmesi, ölçmeyi ve iletişimi kolaylaştırır. Onun için yolların uzunluğu kilometre ile taksit ise metre ile ölçülür.

Ölçmede Sıfır Noktası

Ölçme açısından sıfır önemli bir kavramdır. Ölçmede kullanılan sayı veya sembollerin sıfırı varsa, bu sıfırın ölçülen büyüklüğün sıfır değerine karşılık olması beklenir. Böyle olmazsa ölçme sonuçlarının birbirine oranı anlamını kaybeder.

Ölçülecek değişkenin sıfır değerinin tayin edilebilmesi ve ölçmede bu değerın sıfır sayısıyla gösterilmesi çok istenen bir özelliktir. Bazı değişkenlerin doğal sıfırı vardır. Bir kursa devam eden öğrencileri sayıyorsak, hiçbir öğrencinin gelmediği bir günde devam sıfırdır. Doğal sıfır değerine «mutlak sıfır» demek gelenekselleşmiştir; fakat «gerçek sıfır» deyimini daha anlamlıdır.

Bazı değişkenlerin doğal sıfır değeri bulunmaz. Meselâ zamanı saat, gün, ay gibi birimlerle ifade ediyoruz. Zamanın gerçek sıfırı hangi zamandır? Gerçek sıfırı bulunmayan veya bilinmeyen değişkenlerin ölçülmesinde, bazan seçilen birimden türediği bazan da itibarı olarak tanımlandığı için, bir sıfır değeri bulunur. Sıcaklık ölçmekte kullanılan santigrat ve fahrenheit derecelerinin sıfırları itibarıdır. Bir uzunluk olduğu için yüksekliğin doğal sıfırı anlamlı olduğu halde, yer yüzündeki coğrafi yüksekliklerin ölçülmesinde kullanılan sıfır (deniz düzeyi) da itibarıdır.

Eğitimde kullanılan bazı birimlerin sıfır değerleri vardır. Başarı ölçmede takdir edilen notun sıfırı tanımlanmıştır. Fakat, ne derecede bir başarısızlığa «sıfır» takdir edileceği bilinmez. Bir yazılı yoklamada kopya yapan bir öğrenciye «sıfır» verildiği olur. Bu sıfır, şüphesiz, o öğrencinin gerçek bilgisinin sıfır olduğunu göstermez.

Sıfırın hiç tanımlanmadığı ölçme türleri de vardır. Öğrencilerin yaptıkları resimler en çirkinden en güzele doğru sıralanarak, herbirine bir güzellik sırası verilebilir. Böyle bir ölçmede sıfır yoktur. Sıfır sayısı verilse bile, bir değişkenin hiç bulunmaması gibi bir anlam taşımaz.

Ölçmede İzomorfluk ve Homomorfluk

Bir ölçme işlemi çoğu halde, bir olgusal gerçeği nicilemek amacıyla yapılır ve bu nicilemenin sonucu bazı sayı veya sembollerle

gösterilir. Ölçme sonuçlarının kullanılabilmesi için bu gösterme zorunludur:

Ölçmeye formal yoldan yaklaşıldığında, ölçmenin, bir ampirik ilişkiler kümesini bir formal ilişkiler kümesiyle gösterme olarak düşünüldüğü daha önce belirtilmişti. Ampirik kümenin, ölçülmek istenen bütün özellikleri formal kümede bulunmuyorsa, yani ölçmede kullanılan sayı veya sembol sistemi yetersiz ise, ölçme hatalı olur. Öte yandan, ampirik kümede bulunmayan bir ilişki formal kümede varsa, bu özeliğin kullanılması ölçme sonuçları üzerinde yanlış işlemlere yol açar.

Bir yazılı yoklama, tam doğru cevaplandırılan sorulara tam numara (söz gelşi 1), eksik veya yanlış cevaplara sıfır, verilerek puanlanmış olsun. Bir öğrencinin bir soruya verdiği cevap kısmen doğru ise, bu puanlama işleminde yarı doğru cevaplara değer verilmediği için, öğrencinin yarım bilgisi gözlenemediği halde ölçülmemiş olur. Çünkü bilgiyi göstermekte kullanılan formal küme (0 ve 1) yarım bilgileri göstermemektedir. Bazen bu haliin tersi olur. Ölçme sonuçlarını göstermede kullanılan sayıların özelliklerinden biri, ölçülen değışkende yoktur. Öğrencilerin resimlerine, güzellik derecelerine göre, 1 ile 10 arasında puanlar verildiğini düşünelim. $4 + 4 = 8$ eşitliği bu sayı kümesinde aritmetik bakımından doğru olduğu halde, iki tane «dörtlük» resim bir araya getirilse, güzellik bakımından bir «sekizlik» resme denk olmaz. Bu örnekteki ölçmede «toplama» işleminin anlamı yoktur.

Ampirik kümenin her elemanını formal kümede ancak bir sayı (eleman) karşılıyorsa (bire - birlik), böylece ampirik kümenin hiç bir elemanı gösterilmedik kalmıyorsa (örtme) ve ayrıca elemanlar arası ilişkiler formal kümenin işlemleriyle gösterilebiliyorsa, ölçme **izomorf** olur. Bu özellikleri taşıyan ölçme sonuçlarından ampirik ilişkiler ve ampirik ilişkilerden de ölçme sonuçları elde edilebilir. Bire - bir olma ve örtme şartları yerine getirilmeden yapılan, fakat ilişkileri işlemlerde koruyan bir gösterme de **homomorf** olur. Ölçme işlemlerinde izomorfluk ve homomorfluğun incelenmesi, ölçme sonuçlarının formal özelliklerinin ortaya çıkarılması açısından önem kazanır. Ölçme sonuçlarını ifade eden sembol veya sayıların bu tür formal özellikleri ölçek niteliklerini tayin eder.

Ölçmede Ölçek

Dilimizde ölçek kelimesi bugün çeşitli anlamlarda kullanılmaktadır. Örneğin «1/50 ölçeğinde bir plan» ifadesi gerçek büyüklük-

lerin bu planda 50 defa küçültülerek gösterildiğini ifade eder. Bir haritanın ölçeği de bu anlamı taşır. Bazı harita veya krokilerin altında haritanın ölçeğine göre bölmelenmiş bir ölçek çizgisi bulunur. Harita 1/100.000 ölçeğinde ise, bu çizginin her santimetresi 1 kilometre olarak bölmelenmiştir. O halde ölçek terimi, «belli birimde bölmelenmiş bir ölçme aracı» anlamında da kullanılır. Santigrat veya fahrenheit ölçeğinden söz edilirken, o terim bu anlamda kullanılmaktadır. Ölçek teriminin bazen «birim» yerine de kullanıldığı olur. Bu kitapta özellikle aşağıdaki tartışmalarda ölçek terimi «ölçme sonuçlarını gösteren sembol veya sayıların formal nitelikleri» anlamında kullanılmıştır.

Ölçme işleminin yapılışına ve ölçme sonuçlarını göstermede kullanılan sayı kümesinin özelliklerine bağlı olarak çeşitli ölçekler elde edilir. Ölçmenin formal incelenmesi sonucu, ölçeklerin şu dört ana türde olabileceği ortaya çıkarılmıştır.⁷¹⁸

1. Eşit oranlı ölçekler
2. Eşit aralıklı ölçekler
3. Sıralama ölçekleri
4. Sınıflama ve adlandırma ölçekleri

Bunlardan bazı noktalarda ayrılan, böylece bu türlerin arasında düşen ölçekler de vardır. Aşağıda bu ölçeklerin önemli nitelikleri, dayandıkları aksiyomların tartışılmasına girilmeden özetlenmiş ve bir ölçme işleminin belli bir ölçek niteliğinde sonuç vermesi istendiğinde, işlemin minimum gerekleri belirtilmiştir.

1. Eşit Oranlı Ölçekler : Bir değişkenin iki gerçek değerinin birbirine oranının, bu değerlere karşılık olan iki ölçme sonucunun birbirine oranına eşit olmasına, eşit oranlılık özeliği denir. Yapılan ölçmede ampirik ilişkiler kümesi elemanları arasındaki oranlar korunursa, şüphesiz, eşit oranlı ölçme sonuçları elde edilir. Eşit oranlılık niteliği, ölçmedeki iki özellikten doğar. Bunlardan birincisi, ölçülen değişkenin sıfır değerinin ölçme sonuçlarını gösteren sayılarda sıfıra karşılık olmasıdır. İkincisi ise ölçme sonuçlarını gösteren sayılarda, değişkenin eşit büyüklüklerine, bu büyüklüklerle oranlı sayıların karşılık olmasıdır. Terazilerde kütleye, metre ile uzunluğa ölçmek; bir çokluğu doğal birimleri ile saymak, eşit oranlı ölçek niteliklerinde sonuçlar verir. Bu niteliklerdeki ölçme sonuçları şu bilgileri taşır :

⁷¹⁸S.S. Stevens, «On the Theory of Scales of Measurement», Science, 103 (1948) : 667 - 680.

⁷¹⁹Cureton, op. cit., s. 785 - 804.

a) Herhangi iki ölçme sonucu birbirinden farklı ise, ölçülen değişkenin bu ölçmelere karşılık olan değerleri de farklıdır. Ölçülen değişkenin çeşitli değerleri ölçme sonuçlarına göre sınıflanabilir. (b) Birçok ölçme sonucu, ölçme boyutunda sıralanmışsa, bu sıralama ölçülen değişkenin ölçmelere karşılık olan değerinin sırası ile aynı olur. Ölçülen değişkenin çeşitli değerlerini sıralamak için ölçme sonuçlarını sıralamak yeter. (c) Herhangi iki ölçme arasındaki fark değişkenin bu ölçmelere karşılık olan değerleri arasındaki farkla oranlıdır. Bu nedenle, ölçme sonuçları arasındaki toplama ve çıkarma işlemleri, değişkenin karşılık olan değerleri arasındaki toplama ve çıkarma işlemlerine denktir. (d) Herhangi iki ölçmenin birbirine oranı, değişkenin karşılık olan değerlerinin birbirine oranına eşittir. Bu nedenle, ölçme sonuçları arasındaki çarpma ve bölme işlemleri, değişkenin karşılık olan değerleri arasındaki çarpma ve bölme işlemlerine denktir. Eşit oranlı ölçek niteliklerindeki ölçme sonuçları üzerinde sınıflama, sıralama, farkları veya toplamaları bulma, çarpma veya bölme gibi işlemler, ölçülen değişkende de anlamlarını korurlar. Onun için, bu niteliklerdeki ölçme sonuçlarına her türlü istatistiksel işlemin uygulanabileceği söylenir.^{9,10}

Bir ölçme aracı veya ölçme işlemi ancak şu iki şart altında eşit oranlı ölçek niteliklerinde ölçme sonuçları verir : (i) Ölçekteki sıfır noktası, ölçülen değişkenin sıfır değerine karşılık olmalıdır. (ii) Ölçekteki birimler, ölçeğin her bölgesinde eşit ve ölçülen değişkenin gerçek değerleriyle oranlı olmalıdır. Bir ölçme işleminde bu şartların sağlandığından emin olunursa, ölçme sonuçlarına eşit oranlı ölçek niteliklerinde sayılar olarak bakılabilir. Bu şartlara uyularak hazırlanmış bir ölçme aracıyla yapılan ölçme işlemleri ve doğal birimlerle sayma işlemi, hemen her zaman eşit oranlı ölçek niteliklerinde sonuçlar verir.

2. Eşit Aralıklı Ölçekler : Bir ölçeğin sıfır noktası ile değişkenin sıfır değerinin birbirine karşılık olmadığı halter de vardır. Meselâ, termometre 0°C sıcaklığını gösterirken havanın ısı gerçek anlamda sıfır değildir. Bu hal, ölçekte sıfır noktasının itibarı olarak seçilmesinden doğar. Bununla birlikte, ölçekteki eşit aralıklar değişkenin değerlerinde eşit farklara karşılık oluyorsa, eşit aralıklı bir ölçek meydana gelir. Santigrat ve fahrenheit gibi sıcaklık ölçekleri, zaman ölçekleri ve benzerleri eşit aralıklıdır. Birimlerin eşitliğini sağlayabilen her ölçek eşit aralıklı bir ölçek niteliği kazanır. Bu niteliğe

⁹Ibid, s. 785.

¹⁰H. Arıcı, İstatistik : Yöntemler ve Uygulama, s. 8 - 21.

sahip ölçme sonuçlarında herhangi iki ölçmenin birbirine oranı, değişkenin karşılık olan değerlerinin birbirine oranına eşit değildir. Sıfır noktasının itibarı seçilmiş olması nedeniyle, eşit oranlı ölçeğin yukarıda (d) maddesinde gösterilen özeliği kaybolur. Fakat, (a), (b) ve (c) maddelerinde gösterilen diğer özellikleri korunur. Onun için, eşit aralıklı ölçek niteliğindeki ölçme sonuçları üzerinde sınıflama, sıralama, farkları veya toplamaları bulma gibi işlemler ölçülen değişkende de anlamlarını korurlar; fakat çarpma ve bölme gibi işlemler anlamlarını kaybederler.

Eşit aralıklı ölçekte bir ölçme aracı meydana getirilecekse, bir birim tanımlanır ve pratik amaçlara uygun bir sıfır noktası seçilir. Sıfır noktasının kullanışlı olması yeter; gerçek sıfıra karşılık olması gerekmez. Aracın bölmeleri, tanımlanan birime göre eşit aralıklı ve ölçülecek değişkenin değeriyle oranlı olmalıdır. Eğitimde kullanılan bazı standart testlerin eşit aralıklı ölçek niteliğinde puanlar verecek şartlarda ölçeklenmesine gayret edilir. Genellikle, standartlaştırılmış puanların eşit aralıklı ölçek niteliğinde olduğuna inanılır.

3. Sıralama Ölçekleri : Ölçülecek değişkenin iyice bilinmediği, sıfır değerinin anlamsız olduğu ve bir birimin tanımlanamadığı haller de vardır. Bu hallerde o değişkenin değerleri bir büyüklük sırasına konulabiliyorsa, sıralama işlemiyle ölçme yapılabilir. Öğrencileri boy sırasına dizen bir beden eğitimi öğretmeni böyle bir ölçme yapmış olur. Bir Türkçe öğretmeni de, öğrencilerin kompozisyonlarını «pekiyi - iyi - orta - zayıf» gibi dört gruba ayırdığında, buna benzer bir sıralama yapmış olur.

Sıralamada sıfırın anlamı yoktur. Bir birim söz konusu olmadığından sıraları farklı iki ölçme arasındaki farkın ne büyüklükte olduğu da bilinmez. Sıralama ile elde edilen ölçme sonuçları ancak iki çeşit bilgi taşır : (a) Sırada ayrı yer tutan iki birey, sıralama boyutunda birbirinden farklıdır. (b) Büyüklük sırasında önce gelen bir birey, sonra gelen bireylerin hepsinden daha büyüktür. Bu iki bilgi, bir sıralama işleminden alınan sonuçlarla ancak sınıflama ve sıralama gibi matematiksel işlemler yapılabileceğini; toplama, çıkarma, çarpma ve bölme işlemlerinin anlamlı olmayacağını açıkça gösterir.

Sıralama sonuçları sıra sayıları denilen «birinci, ikinci, üçüncü...» gibi sayılarla veya sembollerle gösterilebilir. Sayılarla gösterme halinde sayılara büyüklük - küçüklük sırası göstermekten ötede bir anlam verilmemesi gerekir.

4. Sınıflama Ölçekleri : Ölçülecek değişkenin belli bir boyutu olmadığını, değişkenin değerlerinin herhangi bir boyutta sıralanmadığını kabul edelim. Bu halde ölçülecek bir nitelik yok mudur? Ölçülecek bir nitelik varsa, ölçme nasıl yapılabilir? Bir sınıfta, bir kısmı erkek bir kısmı kız olmak üzere, N öğrenci bulunsun. Bu sınıfın öğretmenini «Kızlar dışarı çıksın» deyip grubu böylece ikiye ayırdığında bir sınıflama yapmış olur. Bir müdür yardımcısı da sınıf geçme defterinin sonuç hanesine her öğrenci için «geçti», «kaldı», «bütünlemeli» kelimelerinden birini yazdığında bir sınıflama işlemi yapmış olur. Ölçülecek değişken sadece kategorik değerler olabiliyorsa, ancak sınıflama ve her kategorideki bireyleri sayma mümkündür. Sınıflama sonuçları, genellikle, ancak iki bireyin aynı kategoride veya farklı kategorilerde olduğunu bildirir.

Fakat, bir sınıflamada kategoriler (sınıflar) bir boyutta sıralanabilirse, bu kategorilere dahil bireyler de aynı şekilde sıralanabilir. Sınıf geçen öğrencilerin bütünlemelilerden, onların da kalanlardan daha başarılı olduğu kabul edilsin. Bu sayıyı doğru ise, başarı seviyeleri bakımından bu üç kategori arasında «geçti > bütünlemeli > kaldı» bağıntıları doğru olur. Mehmet doğrudan doğruya geçmiş, Cemil bütünlemeye kalmış, Ayşe doğrudan doğruya kalmış bir öğrenci ise, yukarıdaki sayıyı o gruptaki bütün öğrenciler için doğru olmak kaydıyla, başarıları bakımından «Mehmet > Cemil > Ayşe» bağıntıları doğru olur. Bu örnekte görüldüğü gibi, önceden belli bir boyutta sıralanmış kategorilere göre sınıflama işlemi, sıralama ölçeği niteliğinde ölçme sonuçları verir.

Bununla birlikte, bir kümenin bütün elemanlarını sıralamak ile aynı kümenin bütün elemanlarını sıralı kategorilere göre sınıflamak arasında fark vardır. Birinci işlemde her eleman diğerinden ayırt edilebilir. İkinci işlemde ise, bir kategoriye düşen elemanlar arasındaki fark ayırt edilebilir olsa bile dikkate alınmamış olur. Sıralı kategorilere göre sınıflamada, ancak kategori sayısı ile sınırlı ayırım yapılabilir. Kümenin bütün elemanları sıralandığı zaman daha fazla sayıda ayırım yapılabilir.

Sınıflama yoluyla elde edilen ölçme sonuçları üzerinde ancak sınıfların frekanslarına dayanan istatistiksel işlemler anlamlı olur.

Ölçeklerin Birbirine Dönüştürülmesi

Ölçme sonuçlarının matematiksel özellikleri bakımından dört temel ölçekten herhangi birinde olabileceği yukarıda tartışıldı. Bu temel ölçeklerin, en çok bilgi taşıyandan en az bilgi taşıyana doğru

şöyle sıralandıkları okuyucunun dikkatinden kaçmamıştır 1. Eşit oranlı ölçek, 2. Eşit aralıklı ölçek, 3. Sıralama ölçeği, 4. Sınıflama ölçeği. Bir ölçme işlemi bu ölçeklerden herhangi birinde ölçme sonuçları vermiş ise, bu sonuçlar diğer bir ölçeğe dönüştürülebilir mi? Ölçekler arası dönüşümler bazı hallerde mümkündür, bazı hallerde ise mümkün değildir.

Ölçme sonuçları eşit oranlı ölçekte iken, ölçmelerin sıfır noktası gerçek sıfırdan farklı bir değere kaydırılırsa (meselâ, her ölçmeye sabit bir sayı katılırsa), bu dönüştürme sonucunda elde edilen sayılar eşit oranlılık özeliğini kaybeder; fakat diğer özelliklerini korurlar. Demek ki, sıfır noktasını gerçek sıfırdan farklı bir noktaya kaydıran herhangi bir dönüştürme, eşit oranlı ölçümleri eşit aralıklı ölçek niteliğine indirir.

Eşit aralıklı ölçekte birimlerin eşitliği korunduğu için ölçümler arasındaki farklar anlamlıdır. Eşit aralıklı ölçmeler birimlerin eşitliğini bozan, fakat büyüklük sırasını bozmayan bir dönüştürme ile diğer bir ölçeğe çevrilirse, böylece elde edilen sayılar eşit aralıklılık özeliğini kaybeder.

Sıralanmış ölçümleri, büyüklük sırasını parçalara bölerek sınıflamaya dönüştürmek mümkündür. Bu dönüştürmede aynı sınıfa düşen bireyler arasındaki sıra farkları kaybolur.

Böylece çeşitli dönüştürmelerde eşit oranlı ölçekten sınıflama ölçeğine doğru inilebileceği ortaya çıkar. Bu dönüştürmelerin her basamağında bilgi kaybına uğranacağı unutulmamalıdır. Aksi yönde dönüştürmelerle sınıflama ölçeğinden eşit aralıklı ölçeğe doğru çıkmak mümkün olur mu? Aşağıya inişte bilgi kaybı olduğu için, genellikle, yukarıya doğru çıkmak olanağı yoktur. Ancak, eldeki ölçme sonuçlarının dışında bazı bilgi veya bazı sayıltılarla, düşük nitelikteki bir ölçekten daha yüksek nitelikteki bir ölçeğe geçme olanağı sağlanabilir. Psikometri ve psikofizik ile uğraşan bazı yazarlar, nitelikleri düşük bir ölçekten nitelikleri yüksek bir ölçeğe geçme yolları bulmuşlardır.^{11/12} Şüphesiz, bu metotlar ancak dayandıkları sayıltıların doğruluğu oranında geçerlidir.

Eğitimde ve psikolojide, ölçülen birçok değişkenin gerçek sıfır noktası bilinmemekte, bu değişkenlerin ölçülmesi için birimler ta-

¹¹L.L. Thurstone, «A Law of Comparative Judgements», *Psychological Review*, 34 (1927) : 273 - 286.

¹²Stevens, *op. cit.*, s. 667 - 680.

nımlanamamakta ve standart ölçüler geliştirilememektedir. Onun için, eğitim ve psikolojideki birçok ölçmenin ancak sıralama ölçeği niteliğinde sonuçlar verdiği söylenebilir. Bununla birlikte, bu bilim dallarında da standartlaştırma ve ölçmeleri yüksek nitelikte ölçeklere dönüştürme gayretleri önemli yer tutar.

Ölçmede Araçların Rolü

Uzunluk ölçmek için metre, kütle ölçmek için terazi, sıcaklık ölçmek için termometre kullanırız. Birçok kimse ölçme araçlarını kullandıkları halde, bir ölçme aracının ölçmede ne gibi yarar sağladığını düşünmemiştir. Bir kütleyi teraziyle ölçmek, ağırlığın kolumuzda yarattığı çekme duyusuyla onun miktarını kestirmekten daha doğru sonuç verir. Araçlar, ölçülecek değişkenin gözlenmesinde duyarlılığı artırır.

Ölçmenin bir gözlem gerektirdiği daha önce tartışıldı. Ölçülecek değişkenin değerleri gözlenemiyorsa ölçme yapılamaz. Bazı ölçme araçları bu gözlemi kolaylaştırır. Eğitimde başarının ölçülmesinde kullanılan yazılı yoklama, test, vb. gibi araçlar, aslında, öğrenci davranışının veya davranış ürünlerinin gözlenmesini kolaylaştıran araçlardır.

Termometre, metre, vo. gibi bölmeli araçlar, önceden ölçeklenmiş olduklarından, ölçme sonuçlarının standart birimlerle ifade edilmesini sağlarlar. Eğitimde kullanılan standart testler de, terazi veya metredeki kadar doğrulukla olmamakla birlikte, standart ölçümler veren araçlar olarak düşünülebilir.

Eğitimde kullanılan ölçme araçları, hatta ölçme metotları, hem hazırlanışlarında hem de kullanılışlarında gerektirdikleri bilgi ve beceriler yönünden çeşitlilik gösterir. Öğretmenler bazı hazır araçların ve metotların kullanılmasında, bazılarının da yapımında başarılı olacak derecede bilgi ve beceri edinme zorundadırlar. Ayrıca, bir derste amaç edinilen başarı çoğu halde çok çeşitli davranışlardan oluştuğu için, başarının ölçülmesinde kullanılacak araçların çeşitli olması zorunluğu vardır. Öğretmene düşen görev, ölçülecek bilgi veya becerinin ölçülmesine en uygun düşen araçları seçmek, öyle araçlar yoksa kendi gayretleriyle yapabilmektir.

BAŞLICA ÖLÇME TÜRLERİ ve ÖLÇME İŞLEMLERİ

Ölçmenin çeşitli işlemlerle yapılabileceği yukarıdaki tartışmalarda belirtildi. Ölçek türlerinden söz edilirken de belli niteliklerde

ölçümler elde edebilmek için hangi ölçme işlemlerinin gerekli olduğu kısaca belirtildi. Bu tartışmalardan, çeşitli ölçme türleri bulunduğu, bazı ölçme türlerinin ancak belli ölçme işlemleriyle yapılabilceği kanısı doğmaktadır. Gerçekten, ölçme teorilerini kurup tartışanlar da çeşitli ölçme türlerinden söz etmişlerdir.

Ölçülecek değişkenin değerleri doğrudan doğruya gözlenebiliyorsa **temel ölçme** türünden söz edilir. Yukarıdaki kavramlarla ifade edilirse, temel ölçme, ampirik kümedeki elemanların, dolaysız olarak gözlenip nicelendirilmesidir. Bir kovadaki suyu bir tasla boşaltarak suyun hacminin kaç tas olduğunu bulmak, bir bahçenin enini adımlamak, terazi ile ağırlık ölçmek, temel ölçme örnekleridir. Minaralleri birbirine sürterek sertlik sırasına koymak, öğrencileri arka arkaya dizerek boy sırasına koymak da temel ölçme işlemleridir.

Eğitimde temel ölçme ancak bazı hallerde mümkündür. Eğitimde ölçülmesi istenen değişkenlerden birçoğunun doğrudan doğruya gözlenemediği, böyle değişkenlerin temel ölçmeye elverişli olmadığı, ileride daha ayrıntılı olarak tartışılacaktır.

Bazı değişkenlerin temel ölçme işlemleriyle ölçülememesi, başka ölçme türlerinin geliştirilmesine yol açmıştır. Hatta, temel ölçmenin mümkün olduğu hallerde bile, diğer bir ölçme türü daha pratik geliyorsa temel ölçmeden vazgeçilebilir. Kütlenin hem terazi, hem de yaylı kantarla ölçüldüğünü biliriz. Terazi, kütlelerin doğrudan doğruya karşılaştırılmasına imkân verdiği için, temel ölçme işleminde kullanılan bir araçtır. Fakat, bir yaylı kantarla bir cismin kütlesi ölçülürken yapılan işlem kütleleri doğrudan doğruya karşılaştırmak değil, kantardaki yayın, ucuna asılı cismin ağırlığıyla uzama miktarını gözlemektir. Bir termometre ile sıcaklık ölçerken gözlenen şey de, ısı miktarındaki değişme değil, araçtaki cıvanın ısı kaybı ile büzülme veya ısınmakla uzama miktarıdır. Yaylı kantar, termometre ve benzeri araçlar standart birimlerle derecelendikten sonra daha kullanışlı olur. Bu örneklerde olduğu gibi, bir değişkenin başka bir değişken yardımıyla ölçülmesi bir **dolaylı ölçme** işlemidir.

Yaylı kantar, termometre ve benzeri araçların derecelenmesi ya temel ölçme ile, ya da yine temel ölçmelerle doğrulanmış bir ampirik bağıntıya dayanılarak yapılabilir. Henüz derecelenmemiş bir yaylı kantar, ucuna teraziyle ölçülmüş 100, 200, 300 g ağırlığında kütleler asılarak derecelenebilir. Kantardaki yayın her 100 gramlık kuvvetin etkisiyle 1 cm uzadığı deneylerle doğrulanmışsa, yayın uzama miktarı santimetre veya milimetre olarak bölmelenerek kantar yine derecelenmiş, böylece ağırlık veya kuvvet ölçmeye hazırlan-

miş olur. Bu örneklerde yayın uzaması, kuvvetin göstergesi, civa-
nın genişmesi de ısının göstergesi olarak kullanılmıştır. Onun için
bu tür dolaylı ölçmelere **göstergeyle ölçme** denir.

Herhangi bir gösterge ile ölçme işleminin doğru olabilmesi için
derecelemenin doğru yapılmış olması gerekir. Derecelemenin doğ-
ruluğu ise, dayandığı temel ölçme işleminin veya hem temel ölçme
işleminin hem de amplirik bağıntısının doğruluğuna bağlıdır.

Cisimlerin kütleleri ve hacimleri temel ölçme veya gösterge ile
ölçme işlemlerinin herhangi biriyle ölçülebilir. Öte yandan, bir cis-
min kütlesi hacmiyle bölünürse yoğunluğu bulunur. O halde, yoğun-
luk, iki değişken ve bunların arasında bir aritmetik işlemiyle tanımlanmış bir büyüklüktür. Yoğunluk ölçme ise, bu tanıma dayanan bir
dolaylı veya türetilmiş ölçme türüdür.

Bu kitapta başlıca ölçme türlerinin kısaca ele alınmasından
amaç, eğitimdeki değişkenlerin ölçülmesinde kullanılan yaklaşımla-
rı daha kolayca tartışmaya taban hazırlamaktır. Onun için, ölçme
türleri üzerinde daha fazla durulmayacaktır.

DAVRANIŞIN ÖLÇÜLMESİNDE BAŞLICA YAKLAŞIMLAR

Eğitimde boy, ağırlık, yaş gibi değişkenlerin ölçülmesine ihti-
yaç duyulabilir. Bu değişkenler, uzunluk, kütle, zaman gibi temel
boyutlara olduğundan, ölçülmelerinde önemli güçlüklerle karşılaşıl-
maz. Çünkü, böyle fiziksel değişkenlerin ölçülmesiyle ilgili birçok
problem başka alanlarda çözülmüş, bu tür büyüklüklerin birimleri
tanımlanmış, ölçme araçları geliştirilmiştir. Fakat, çeşitli eğitim ka-
rarlarında kullanılan birçok başka değişkenin boyutları kesinlikle bi-
linmez; birimleri tanımlanmamış, ölçme araçları geliştirilmemiştir.
Öğrencilerin aritmetik becerileri, tarih dersindeki başarıları, güzel
sanatları takdir dereceleri, ellşlerinin mükemmelliği, karakterlerinin
sağlamlığı veya ahlâklılık dereceleri, vb. gibi değişkenlerden eğitim-
de sık sık söz edilmesine rağmen, bu büyüklüklerin ölçülmeleri stan-
dartlara bağlanamamıştır. Bu değişkenler ya öğrenci davranışlarıyla
tanımlanan, ya da öğrenci davranışlarında ortaya çıkan nitelikler-
dir. O halde öğrenci davranışları gözlenip nicelenebilirse, bu tür de-
ğişkenler de ölçülebilir. Bazı öğrenci davranışları gözlenebilir. Bazı
davranışların da ürünleri ölçmeye elverişli olur. Bir öğrencinin ya-
zılı yoklama sorularına verdiği cevap, diğer birinin çizdiği resim,
bir diğerinin yaptığı masa, vb. bir davranışlar zinciri sonunda orta-
ya çıkan somut ürünlerdir. Ortaya somut bir ürün çıkaran davranış-
larda ürünün özelliklerini nicelemekle, yani ürüne bir puan vermek

ve bunu o davranışların ölçüsü olarak kabul etmekle bir ölçme yapılabilir. Hem davranış ürününün, hem de o ürünün ortaya çıkarılmasındaki davranışların çeşitli basamaklarda gözlenip nicelendirilmesi de üçüncü bir ölçme yolu olarak düşünülebilir. Bir öğrencinin yaptığı masaya puan takdir edilebileceği gibi, öğrencinin bu masayı yaparken kullandığı teknikler, atelyedeki temizlik ve düzeni, işteki devamlılığı, işi tamamlama süresi vb. de gözlenip puanlanabilir.

Özet olarak denilebilir ki, öğrenci davranışlarının ölçülebilmesi için ya davranışların, ya o davranışlar sonunda ortaya çıkan ürünün, ya da her ikisinin gözlenip nicelendirilmesi gerekir.

Eğitimin amacı öğrencilerde bir takım istendik davranışlar geliştirmek olduğuna göre, eğitimde en geçerli ölçme, o davranışların bireylerde doğrudan doğruya ve gerçek hayat şartları altında gözlenmesiyle yapılan ölçme olmalıdır. Davranış ürünlere bakılarak ölçülecekse, nicelendirilecek ürünler de gerçek koşullar altında ortaya çıkarılmış olmalıdır. Ne yazık ki, ölçülecek davranışların her türünü gözlemek, hele gerçek hayat koşullarında gözlemek çoğu halde mümkün olmaz. Onun için, eğitimde ölçme problemine başka yöntemlerle yaklaşma yolları da aranmıştır.

Davranışın gerçek hayatta gözlenmesi, çeşitli güçlüklerden dolayı mümkün olmayabilir. Hedef davranış, öğrencinin ileride yetişkinlik hayatında yapacağı bir şey, veya bugün göstereceği bir davranış olsa bile, gerçek hayatta istenildiği zaman gözlenmesi mümkün olmayacak derecede seyrek yapılan bir şey olabilir. Böyle durumlarda hedef davranışı değil, gerçek duruma benzeyen yapma bir durumda gösterilen davranışı gözleyip ölçmek daha kolay olabilir. Bu halde ölçme durumu gerçek hayat durumuna mümkün olduğu kadar benzetilir. Seçilen ölçme durumları gerçek durumu ne derece sadakatle yansıtabilirse ölçme de o derece geçerli olur.

Eğitimde çeşitli sınav ve ölçme durumlarının yaratılması, davranışın gerçek hayat koşullarında ölçülmesine imkân vermemekle birlikte, gerçeğe yakın bir durumda ölçülmesine imkân verebilir. Üstelik, böyle yapma durumlar öğretmenin kontrolunda olduğu için, istenildiği kadar çeşitli durumlar yaratılabilir; ölçmeye karışması hata yaratacak yan etkiler ödenebilir ve gözlem için uygun bir zaman seçilebilir. Meselâ, ilkokulda aritmetik öğretiminin hedef aldığı davranışlardan biri «alışverişte ödenecek parayı ve paranın üstünü yanlışsız hesaplama» ise, bu davranışın ölçülmesi için öğrencinin büyümesini beklemek yersiz, çarşıda alışveriş yapısını gözlemek gereksiz olur. Okulda yaratılan bir durumda öğrencilerin alış-

veriŖe benzer bir etkinlik iine sokulmaları, hattâ bu oyunda gerek para verip üstünü almaları bile saėlanabilir. Byle durumlar yaratılırken, ğrencinin o hedef davranıŖı icra edeceėi bir etkinilik bulmaya gayret etmek gerekir.

Bir gerek durumdaki davranıŖın gzlenmesi mmkn olmayan hallerde baŖvurulabilecek diėer bir yntem, ğrenciye bir gerek hayat durumunu anlatmak ve byle bir durumda karŖılaŖtıėında ne yapacaėını sormaktır. Bir ğretmen adayına eŖitli meslek dersleri okutmadaki asıl ama, adaya, ğretmen olduėu zaman gstermesi istenen davranıŖları kazandırmaktır. Adaya henz ğretmenlik yaptırılmadıėına gre, ğretmenlikte gstereceėi bir davranıŖı gzlemeye henz imkân yoktur. Onun yerine, ğretmen adayına meslek dersinde gereėe yakın bir problem verilir ve bu durumda ne yapacaėı sorulur. Adaya «Bir ğrenciyi yazılı yoklamada kopya yaparken yakalasanız, ona nasıl muamele edersiniz?» gibi bir soru sorulabilir. Byle bir soruyu cevaplandıran kimse, Ŗphesiz, gerek davranıŖını deėil, onunla baėıntılı olacaėı umulan szel davranıŖını gstermiŖ olur. Yine bir ğretmen adayının ğretmen olarak bir dersi nasıl yrttėn gzlemek mmkn olmazsa, ondan ileride vereceėi bir dersiñ planı istenebilir.

Hedef davranıŖın yapma durumlarda gzlenmesi, hatta ilerideki davranıŖ iin bir tasarı hazırlatılması da mmkn olmayabilir. Byle hallerde, hedef davranıŖ yerine, onunla baėıntılı diėer davranıŖları lme yoluna gidilir. YurttaŖlık bilgisi dersinde amalardan biri «ğrencinin iyi bir vatandaŖ olabilmesi», bununla ilgili olarak da yetiŖkinlikte, «vergi ykmls olduėunda vergisini kaırmadan deme davranıŖı kazanması» olabilir. ğrenci henz vergi ykmls olmadıėı iin bu davranıŖın gzlenmesine imkân yoktur. Fakat, ğrencinin okulda iken verginin gerekliliėini kavramıŖ olması, vergi kaakılıėının ahlâksızlık olduėunu iyice ėrenmiŖ olması, vergiye karŖı olumlu ve ahlâksızlıėa karŖı olumsuz tutumlar geliŖtirmiŖ olması vb. gibi niteliklerinin ileride vergi kaırmasını nleyeceėini kabul edebiliyorsak; vergi deme davranıŖı yerine, ona temel olacak diėer davranıŖları lebiliriz. Byle bir sayılıya dayanılınca, okulda geliŖtirilecek ve llecek davranıŖların oėu bilgiye, tutum ve duygulara dnŖr. O halde, bazı davranıŖlar yerine onların dayandıėı bilgi, beceri ve tutumlardan ıkan davranıŖları gzleyip lebiliriz.

Eėitimde davranıŖı lmek iin eŖitli yaklaŖımlar vardır. Fakat, bu yaklaŖımların hepsi aynı nitelikte lme sonuları vermeyeceėi gibi, hepsi de aynı derecede kullanıŖlı deėildir. Herhangi bir durum-

da en uygun yaklaşımları seçmek, ölçmeyi yapacak kimsenin görevidir. Genel bir ilke olarak, davranışın gerçek koşullar altında ölçülmesinin en doğru bilgileri vereceği söylenebilir.

Nihayet, hangi yaklaşım kullanılırsa kullanılsın, herhangi bir alanda ölçme, Thorndike ve Hagen¹³ tarafından açıkça belirtildiği gibi, şu üç basamakta gerçekleştirilir : (a) Ölçülecek özeliğın tanınması ve tanımlanması, (b) ölçülecek özeliğın ortaya çıkarılmasını ve gözlenmesini sağlayacak işlemlerin belirtilmesi, (c) gözlem sonuçlarını derece veya miktar olarak niceleyecek işlemlerin bulunması. Herhangi bir değişken gözlenemiyorsa, varlığı bilinse bile, ölçülemez. Herhangi bir özellik derece veya miktar olarak gösterilemiyorsa, gözlenirse bile iyice ölçülmüş olmaz. Onun için, bu kitabın temel amacı, eğitimde ölçülmesi istenen öğrenci davranışlarını gözlenebilir ve nicelenebilir hale getirmede kullanılan başlıca metotları tanıtmaktır.

¹³R.L. Thorndike and E. Hagen, *Measurement and Evaluation in Psychology and Education*, 3rd ed., s. 9.

B Ö L Ü M 3

Araç ve Yöntemlerin Nitelikleri

Bir ölçme aracının veya yönteminin kullanılış amacına uygunluğunu tayin eden birçok etken vardır. Bu etkenler aracın veya yönetimin kendi niteliklerinden doğduğu gibi, uygulanış koşullarından, üzerinde ölçme yapılan kişilerin niteliklerinden de doğabilir. Ölçme araç ve yöntemlerinin niteliklerini **güvenilirlik, geçerlik ve kullanılabilirlik** kavramları altında gruplamak gelenekselleşmiştir.

Güvenilirlik kavramıyla herhangi bir ölçme araç veya yönteminin ölçtüğü değişkeni ne derece duyarlıkla ölçebildiği; başka bir ifadeyle, ölçme sonuçlarının tesadüfi hatalardan ne derece arınık olduğu ifade edilir. Geçerlik kavramıyla bir ölçme aracının veya yönteminin, onunla ölçülmek istenen değişkeni ölçüp ölçmediği; ölçebiliyorsa, onu başka değişkenlerden ne derece arınık olarak ölçtüğü ifade edilir. Kullanılabilirlik kavramıyla, araç veya yöntemin uygulanabilirliği eğitim seviyesi, uygulama süresi, ölçmeciden istediği beceriler, maliyet ve ekonomi gibi etkenlerin o araç veya yöntemin kullanılmasına etkileri ifade edilir.

Gerek güvenilirlik, gerekse geçerlik ölçme hatalarıyla yakından ilgili kavramlardır. Özellikle, güvenilirliğin incelenebilmesi için, ölçmede hata probleminin ve hata türlerinin, hiç değilse ana çizgileriyle bilinmesi gerekir. Bu bölümde ölçme hatalarına kısaca değinildikten sonra, güvenilirlik, geçerlik ve kullanılabilirlik kavramları eğitimdeki araç ve yöntemler açısından tartışılmakta; bu nitelikleri etkileyen etkenler incelenmekte; eğitimdeki ölçmelerde geçerliği, güvenilirliği ve kullanılabilirliği gerektiğinde yükseltebilme yolları genel çizgileriyle verilmektedir.

ÖLÇME HATALARI

Ölçme sonuçlarına çeşitli yollardan hata karışır. En duyar araçlarla ve en doğru sanılan yöntemlerle yapılan ölçmelerde bile bir miktar hata payı vardır. Ölçme hataları, ölçmede kullanılan araç-

tan, ölçme yönteminden, ölçmeyi yapan kimseden, ölçmenin yapıldığı ortamdan, üzerinde ölçme yapılan bireylerin bütün bu etkenlerle etkileşiminden doğabilir. Ölçme hatalarının incelenmesi hem kuramsal olarak, hem de uygulama alanlarında ilgi çeken bir konudur. Eğitimde ölçülen değişkenlere de çeşitli kaynaklardan ve çeşitli türlerde hatalar karışır. Bu sebeple, eğitimdeki ölçmelerde de hataların incelenmesi önemli konulardan bir haline gelmiştir.

Eğitimdeki ölçmelerde hatalardan sakınılmak isteniyorsa, hata kaynakları ve özellikle bu kaynakların ölçme sonuçlarını nasıl etkilediği bilinmelidir. Ölçmelerdeki hataların miktarı kestirilecekse, hem hatanın ne olduğu iyice tanımlanmalı, hem de hata miktarını hesaplamaya imkân veren bir yöntem bulunmalıdır. Ölçme sonuçları önemli kararlara dayanak olacaksa, ölçmelerde ne kadar hata bulunduğunu kestirmek zorunlu olur. Çünkü, verilecek kararın isabet derecesi, başka etkenlere de bağlı olmakla birlikte, ölçme sonuçlarındaki hatalarla sınırlıdır.

Diğer deneysel bilimlerde ve uygulamalı bilimlerde, ölçme hataları problemine kuramsal açıdan yaklaşmış, hataları formalleştiren kavram ve indisler geliştirilmiş, hataları niceleme yolları bulunmuştur. Ne yazık ki, eğitimde kullanılan ölçme ve araç ve yöntemlerinin temelinde farklı oluşu, fizik gibi bu alanda daha gelişmiş bir bilim dalının kavram ve yöntemlerinin eğitimde de kullanılmasını engellemektedir. Onun için, eğitim, psikoloji, vb. gibi bilim dallarında hata problemlerine güvenilirlik ve geçerlik kavramlarıyla yaklaşılması gelenekleşmiştir. Bu sebeple, ölçme hataları bu kitapta da güvenilirlik ve geçerlik kavramlarıyla ele alınmıştır.

Hataların istatistiksel özellikleri dikkate alınırsa, **sabit**, **sistematik** ve **tesadüfi** olmak üzere üç tür hata ayırt edilebilir.

Bir cetvelin yanlışlıkla 0 - 30 cm yerine 1 - 30 cm olarak bölme-lendiği düşünölsün. Bu cetvelde bölmeler sıfır yerine birden başladığı için, araç aslında 1 cm kısadır. Cetvelin bu kusurunun farkında olmayan bir ölçmeci, onunla yaptığı her ölçmeyi gerçek değerinden 1 cm daha uzun okuyacaktır. Bunun sonucu olarak, bu cetvelin bir defa uygulandığı her ölçmede, başka hata olmasa bile, 1 santimetrelık hata bulunacaktır. Bir ölçmeden diğerine miktarı değişmeyen hatalara sabit hatalar denir. Toplam puanı 10 olan bir yazılı yoklamada, 2 puanlık bir soru sınıfın tümü tarafından cevapsız bırakılmış olsun. Cevaplar puanlanırken bu soruya sıfır velirse, en yüksek toplam puan 10 değil, 8 olacaktır. Bu soru, sınav süresinin

yetersizliğinden cevapsız bırakılmış olsun. Süre verildiğinde bu soruyu doğru cevaplandırabilecek öğrencilerin hepsi, böylece 2 puan eksik puanlanmış olacaktır. Bu öğrencilerin puanlarında sabit hata bulunacaktır. Puanlama yöntemi değiştirilip, sanki bu soru doğru cevaplandırılmış gibi, her öğrencinin puanına 2 puan eklenecek olursa, aslında bu sorunun doğru cevabını yazamayacak olan öğrenciler haketmeden ikiye puan kazanmış olacaklardır. Bu ikinci grubun puanında da, bu sefer aksi yönde, sabit hata bulunacaktır.

Yukarıda sözü edilen kusurlu cetvel ile 29 santimetreden uzun cisimlerin boylarının ölçüldüğü düşünölsün. Ölçmeci, cetveli bu cisimlere bazen iki defa (30 - 58 cm arasındakiler için), bazan da üç defa (59 - 87 cm. arasındakiler için) uygulamak zorunda kalsın. Böylece, ölçmelerin bir kısmında $2 \times 1 = 2$ cm, diğesinde $3 \times 1 = 3$ cm hata bulunacaktır. Bu hata, cismin uzunluğuna bağılı olarak sistemli bir biçimde (her 29 santimetrede 1 cm) değişmektedir. Ölçölen büyüklüğe, ölçmeciye veya ölçme koşullarına bağılı olarak miktarı değişen hatalara sistematik hatalar denir. Bir öğretmen bir yazılı yoklamayı puanlarken kız öğrencilere erkeklerden daha fazla puan takdir ederse, öğrencinin cinsiyetine bağılı bir sistematik hata yapmış olur. Aynı tarzda, yazı güzelliğinin puana karıştırılmaması gereken bir sınavda el yazıları daha güzel olan öğrencilere daha fazla puan verilmesi, başta okunan kâğıtlarda kılın kırık yaratılmasına karşılık, sona kalan kâğıtlara gelişigüzel bol puan verilmesi, vb. gibi hatalar da sistematik hatalardır. İstatistik dilinde bu türden sistematik hatalara **yanlılık** denir.

Uzunlukların ölçüldüğü cetvelde bölmeleme kusuru bulunmasın. Fakat, ölçmeci dikkatsiz olsun; gerek cetveli uygulamada, gerekse sonucu okumada gelişigüzel hatalar yapmış olsun. Bu halde ölçme sonuçlarının bir kısmı gerçek değerlerinden büyük, diğeri bir kısmı ise gerçek değerlerinden küçük bulunacaktır. Böyle olmakla birlikte, hatanın hangi yönde olacağı önceden kestirilemez. Bir cismin gerçek değerinden büyük gösterilmesi olasılığı, onun gerçek değerinden küçük gösterilmesi olasılığına eşittir. Böyle hatalar, genellikle, kaynakları iyi bilinmeyen hatalar olup ölçme sonuçlarına gelişigüzel karışır. Bu tür hatalara tesadüfi hatalar denir. Yazılı yoklama cevaplarını dikkatsizce okuyup puanlayan bir öğretmenin puanları tesadüfi hatalarla dolu olur. Cevaplara, hiç okumadan puan takdir eden diğeri bir öğretmenin puanları ise tümüyle hata sayılır.

Bir ölçme sonucunda bu hataların herhangi biri bulunabileceği gibi, iki türden, hattâ üç türden hata birlikte bulunabilir. Oysa ki, eği-

timdeki ölçmeler değerlendirmelere veri sağlamak amacıyla yapılır. Değerlendirmeler de kararlara dayanak sağlar. Kararın isabetli olabilmesi için, ölçmelerin belli bir dereceye kadar hatadan arınık olması gerekir. Ölçme hatalarının, özellikle tesadüfi hataların hangi koşullar altında azaltılabileceği, güvenilirlik kavramına dayalı olarak, aşağıda tartışılmaktadır.

ÖLÇMEDE GÜVENİRLİK

Bir öğretmenin üç öğrenciye laboratuvarında bir metal çubuğun boyunu ölçtirdiği düşünölsün. Bu üç öğrenci, Akın, Bilge ve Cengiz aynı çubuğu beşer defa ölçerek şu sonuçları bulmuş olsunlar :

<u>Akın'ın</u> <u>ölçmeleri</u>	<u>Bilge'nin</u> <u>ölçmeleri</u>	<u>Cengiz'in</u> <u>ölçmeleri</u>
12,2 cm	12,5 cm	120,1 mm
12,1 cm	12,0 cm	120,2 mm
12,2 cm	12,3 cm	120,3 mm
12,3 cm	12,1 cm	120,2 mm
12,2 cm	12,6 cm	120,2 mm

Akın ve Bilge ölçmede milimetre bölmeli bir cetvel kullanmışlar; ölçme sonuçlarını bir ondalık basamakla santimetre olarak kaydetmişlerdir. Cengiz ise milimetrenin onda birini ölçebilen bir kompas kullanmış ve sonuçları bir ondalık basamakla milimetre olarak kaydetmiştir. Akın'ın ölçmeleri 12,1 cm ile 12,3 cm arasında, Bilge'ninkiler 12,0 cm ile 12,6 cm arasında değişmiştir. Çubuğun boyunun, ölçme sırasında bu kadar fazla değişmesi beklenemez. Ölçme sırasında çubuğun boyunun sabit kaldığı farzedilirse, bütün değişkenliğin ölçme hatalarından doğduğunu kabul etmek gerekir. Cengiz bile beş ölçmede aynı sonucu bulamamıştır. Onun ölçmelerinde de hata vardır.

Akın ve Bilge aynı sonuçları bulmamışlardır. O halde, her iki ölçmeci de hata yapmıştır; fakat hataları eşit değildir. Ölçmelerde ne kadar hata olduğu bilinse, hangisinin daha fazla hata yaptığı kestirilebilir. Bunun için, önce, metal çubuğun boyunun gerçekte ne kadar olduğu bilinmeli; sonra da ölçmelerdeki hata miktarı kestirilebilmelidir.

Çubuğun gerçek boyunu mutlak bir kesinlikle bilmeğe olanak yoktur. Yukarıdaki ölçmelerden daha duyar bir ölçme, daha az ha-

talı ve bu nedenle gerçeğe daha yakın olacaktır; fakat gerçeğin kendisi olmayacaktır. Onun için «gerçek değer» kuramsal bir kavramdır; somut bir anlamı yoktur. Böyle olmakla birlikte, bir ölçmeci tarafından elde edilen ölçmelerin ortalamasının, gerçek değere, ölçmelerin herhangi birinden daha yakın olması beklenir. Aritmetik ortalama Akın'ın ölçmelerinde 12,2 cm; Bilge'ninkilerde 12,3 cm; Cengiz'inkilerde ise 12,02 cm bulunur.

Ölçmelerin ortalaması o büyüklük için gerçek değer olarak kabul edilirse, ölçmelerin ortalamadan farkları hata sayılabilir. Bu açıdan bakılınca, Bilge'nin ölçmelerinde daha fazla hata olduğu açıkça görülür. Onun ölçmeleri aritmetik ortalamadan $\mp 0,3$ cm ayrılmıştır; oysa ki, Akın'ınkiler $\mp 0,1$ cm ayrılmıştır. Cengiz'in ölçmelerinde ise ortalamadan ayrılma ancak $\mp 0,01$ santimetredir. Bu örnekteki ölçmelerde hangi türden hatalar vardır? Buradaki hataların tesadüfi hata olduğu söylenebilir.

Bir ölçme sonucu, içindeki tesadüfi hataların azlığı oranında güvenilir sayılır. Güvenirlik, bir bakıma, ölçme sonuçlarının tesadüfi hatalardan arınlık derecesini ifade eden bir kavramdır. Diğer bir anlamda güvenilirlik, ölçmelerin duyarlık derecesini ifade eden bir kavramdır. Ahmet ve Bilge ölçmelerini bir cetvelle yapmışlardır; onların aracı ve yöntemi ancak 1 mm büyüklüğündeki uzunlukların ayırt edilip ölçülebilmesine olanak vermiştir. Cengiz'in aracı ve yöntemi ise 0,1 mm büyüklüğündeki uzunlukların ayırt edilip ölçülebilmesine olanak vermiştir. Onun ölçmelerinin duyarlığı ve bu sebeple güvenilirliği daha yüksektir.

Ölçme sonuçlarının güvenilirliğinden söz edilebileceği gibi, sadece ölçme aracının veya sadece ölçmenin güvenilirliği de tanımlanıp kestirilebilir. Eğitimdeki uygulamalar için, ölçme araç ve yöntemlerinden elde edilen puanların güvenilirliğini tanımlamak ve onunla uğraşmak çoğu halde yeterlidir. Çünkü, araç, yöntem, ölçmeci ve ölçme ortamı gibi etkenlerin ortak etkileri zaten puanlara yansır. Bu kitapta puanların güvenilirliği incelenecek, fakat bu etkenlerin ayrı ayrı incelenmesine imkân veren metotlardan söz edilmeyecektir.

Yukarıdaki örneklerde ölçme hatasının kestirilebilmesi, ölçmelerin tekrarı sayesinde mümkün oldu. Her öğrenci çubuğun boyunu yalnız bir defa ölçmüş olsaydı, herbir ölçmenin ortalaması kendisine eşit olacak, bu ortalamalardan ayrılmalara söz konusu olmayacaktı. O halde, eğitimdeki ölçmelerde hata miktarının veya güvenilir-

liğin kestirilebilmesi için, ölçmelerin tekrarına, hiç değilse ölçmelerin tekrarı sayılabilecek bazı yöntemlere ihtiyaç vardır.

Eğitim, psikoloji, vb. gibi alanlarda ölçmeleri birçok defa tekrar etmek hemen hiçbir halde mümkün olmaz. Bazı hallerde iki, nadiren üç defa tekrar edilen ölçmeler yapılabilir. Çoğu halde, ancak bir defa uygulanan bir sınavdan alınan puanların güvenilirliğini kestirmek gerekir. Onun için, eğitim, psikoloji, v.b. gibi alanların ölçmelerinde güvenilirliğe, yukarıdaki misalde görülen basit anlamdan, oldukça farklı anlamlar verilmiş; bu anlamlara bağlı olarak, çeşitli güvenilirlik kestirme yöntemleri bulunmuştur. Güvenirlik kavramının başlıca anlamları aşağıdaki örneklerde görülmektedir.

Bir test aynı öğrenci grubuna birkaç gün arayla iki defa uygulansın. İki uygulama arasında öğrencilerin o testle ölçülen yeteneklerinde hiçbir değişme olmamışsa, iki uygulama sonuçlarının birbirini tutması beklenir. Bu ölçme sonuçlarında hiç tesadüfi hata yoksa, her iki puan kümesi de öğrencileri aynı başarı sırasına koyar. Bu halde testin güvenilirliği tamdır. Bu örnekte güvenirlilik, aynı araçla yapılan iki ölçme arasındaki **tutarlılık** anlamındadır. Gerçekte bir test veya sınav aynı öğrenci grubuna iki defa uygulansa, yüzde yüz tutarlı sonuçlar alınmaz. İki uygulama arasındaki sürede öğrencilerin yetenekleri değişebilir. Bazı öğrenciler, uygulamaların birinde veya diğerinde bazı sorulara gelişigüzel cevaplar vermiş olabilirler. Bunlar ve benzeri diğer etkiler sebebiyle iki puan kümesi birbirini tutmuyorsa, ölçme araç ve işlemlerinde hata bulunduğu kanısına varılır.

Bir test veya sınavı aynı öğrenci grubuna iki defa uygulamanın önemli sakıncaları vardır. Birinci uygulamadan ikincisine hatırlama, iki uygulama arasındaki sürede bazı soruların cevaplarını öğrenme ve benzeri gibi olaylar ölçme sonuçlarını etkiler. Test güvenilirliğini kestirmenin önemli olduğu hallerde, bir test değil, birbiriyle eşdeğer iki test hazırlanır. Bu iki testten biri birinci, diğeri ikinci uygulamada kullanılır. Böyle hesaplanan ölçme güvenilirliği, bir testin iki eşdeğer formu arasındaki tutarlılık anlamına gelir. Bu yöntem, güvenirlilik kestirmede iyi bir yöntem olmamakla birlikte, eşdeğer araçların hazırlanmasındaki ve eşdeğerliğin kanıtlanmasındaki güçlüklerden dolayı pek seyrek uygulanır.

İki uygulamadan elde edilen puanlar arasındaki korelasyon katsayısı, iki uygulamanın ne derece tutarlı olduğunu gösterir. Bu amaçla hesaplanan momentler çarpımı korelasyonuna¹ **güvenirlilik katsayısı**

¹Bu bölümde kullanılan istatistikler için on besinci bölüme bakınız.

denir. Korelasyon katsayısı — 1,00 ile + 1,00 arasında değişmekle birlikte, güvenilirlik katsayıları hemen her zaman 0,00 ile + 1,00 arasında değişir. Ölçme sonuçları tesadüfi katalardan ne derece arınık ise güvenilirlik katsayısı da üst sınıra o derece yaklaşır.

İster bir testle, ister iki eşdeğer form ile güvenilirlik hesaplanacak olsun, iki uygulama arasındaki süre uzatılabilir. İki uygulama arka arkaya yapılmayıp birkaç gün veya bir iki hafta gibi kısa bir süre verilirse, öğrencilerin bir uygulamadan diğerine fazla değişmesi beklenmez. Fakat iki uygulama arasında altı ay, bir yıl gibi uzun bir zaman olursa, iki ölçme sonucunun birbirini tutması sadece ölçme aracının güvenilir olduğunu göstermekle kalmayıp, testle ölçülen yeteneğin altı ay veya bir yıl gibi bir sürede değişmediğini de gösterir. Böyle durumlarda hesaplanan güvenilirlik katsayılarına kararlılık ölçüsü olarak da bakılır.

İki eşdeğer araç hazırlamak veya bir aracı iki defa uygulamak mümkün olmayabilir. Bu gibi hallerde, bir testin sadece bir defa uygulanmasıyla elde edilen puanlardan ölçme güvenilirliğini kestirme yolları bulunmuştur. Bu amaçla en çok kullanılan yöntem, testin maddelerini görünüşte eşdeğer iki yarıya ayırmak; iki yarıdan ayrı puanlar elde ederek, iki yarı puanları arasında tutarlılık aramaktır. Güvenirlik katsayısı, bir testin iki eşit yarısından elde edilen puanlardan yararlanılarak, şu formülle hesaplanır :

$$(3.1) \quad r = \frac{4 r_{12} S_1 S_2}{S_x^2}$$

r_{12} : İki yarı puanları arasındaki korelasyon
 S_1 : Birinci yarı puanlarının standart kayması
 S_2 : İkinci yarı puanlarının standart kayması
 S_x : Toplam puanların standart kayması

Elde sadece r_{12} korelasyonu varsa, iki yarı puanlarının standart kaymaları eşit farzedilerek, şu formül kullanılabilir :

$$(3.2) \quad r_x = \frac{2 r_{12}}{1 + r_{12}}$$

Bu yöntem uygulandığı takdirde test güvenirliliği, testin iki yarısı arasındaki tutarlılık anlamına gelir. Test gelişigüzel cevaplandırılmış ise, iki yarının puanları öğrencileri aynı sıraya koymaz; böylece, test güvenirliliğinin düşük olduğu kanısına varılır. Fakat, testin iki yarısı farklı yetenekleri ölçüyorsa, öğrencilerin bu yeteneklerde «aynı derecede yetenekli» olmaları beklenemeyeceğinden, iki yarının puanları arasındaki tutarsızlık sadece testteki tesadüfi hatalar-

dan değil, iki yarının ölçtüğü yeteneklerin farklı oluşundan da doğar. Onun için, böyle hesaplanmış bir güvenilirlik katsayısı büyük bulunursa, hem testin tesadüfî hatalardan arınık olduğu, hem de testin iki yarısının eşdeğer olduğu kanısına varılır. Fakat bu katsayı küçük çıkarsa, güvenilirliğin mi düşük olduğu, yoksa iki yarının çok farklı değişkenleri mi ölçtüğü kestirilemez. Bu sebeple, bu yöntemle güvenilirlik hesaplanacaksa, önce, iki yarının soru sayısı, soru güçlüğü ve soruların kapsamı bakımından hiç değilse görünüşte eşdeğer olduğu kanısına varılabilmelidir.

Sadece bir defa uygulanan bir testte güvenilirlik kestirmenin diğer bir yolu, iki yarıdan alınan puanların birbiriyle uyumu yerine, bütün soruların birbiriyle uyumuna bakmaktır. Böylece bulunan yüksek bir güvenilirlik katsayısı, hem test puanlarının tesadüfî hatalardan arınık olduğunu, hem de o testle ölçülen değişkenin tek boyutlu bir değişken olduğunu gösterir. Karmaşık, çok boyutlu bir değişkeni ölçen bir testin bütün sorularının birbiriyle tutarlı olması beklenemez. Onun için, bir testin maddeleri arasındaki tutarlılığa dayanan yöntemler, kapsamı bir örnek olmayan testlere uygulanamaz.

Bir testin bütün maddelerinin birbiriyle ne derece tutarlı olduğu Kuder - Richardson formülleriyle hesaplanabilir. KR - 20 simgesiyle anılması gelenekselleşmiş olan formül şudur :

$$(3.3) \quad r_x = \frac{K}{K-1} \left[1 - \frac{\frac{K}{\sum p_i (1-p_i)}}{S_x^2} \right]$$

K : Testteki madde sayısı
 p_i : j maddesinin güçlük indisi
 S_x : Test puanlarının standart kayması.

Test maddelerinin herbirinin güçlük indisi bilinmezse bu formül uygulanamaz. Bu sebeple, KR - 20 formülü ancak madde analizi yapılmış testlere uygulanır.²

Testteki bütün maddelerin aynı güçlük derecesinde olduğu farzedilerek, KR-20 formülünden, onun özel hali olan ve KR-21 simgesiyle bilinen şu formül çıkarılmıştır :

$$(3.4) \quad r_x = \frac{K}{K-1} \left[1 - \frac{K \bar{X} - (\bar{X})^2}{K S_x^2} \right]$$

K : Testteki madde sayısı
 S_x : Test puanlarının standart kayması.
 $\bar{X}$: Test puanlarının ortalaması.

²Madde güçlük indisi ve madde analizi için on sekizinci bölüme bakınız.

Bu formül madde analizi yapılmamış testlere de uygulanabilir. Fakat, testin maddeleri eşit güçlükte değilse, KR - 21 ile hesaplanan güvenilirlik katsayısı KR - 20 ile bulunacak katsayıdan küçük çıkar. Bundan dolayı, KR - 21 ile hesaplanan değer, testin güvenilirlik katsayısının alt sınırı olarak kabul edilir.

Ham puanlar varyansının, tutarlı veya kararlı varyans ile ölçme hataları varyansının toplamından meydana geldiği düşünülebilir. Bu düşünce şöyle formülleştirilir :

$$(3.5) \quad S_x^2 = S_t^2 + S_e^2$$

S_x : Ham puanların standart kayması
 S_t : Hatasız veya tutarlı puanların standart kayması
 S_e : Ölçme hatalarının standart kayması

Bu eşitliğe dayanılarak; güvenilirlik katsayısı şu oranla tanımlanır :

$$(3.6) \quad r_x = S_t^2 / S_x^2$$

O halde, güvenilirlik katsayısı, test puanlarındaki tutarlı veya kararlı değişkenliğin ham puanlar değişkenliğine oranı olarak yorumlanabilir. İşte bu sebeple, yüzde yüz tutarlı bir testte bu oran 1,00 değerini, puanları tamamen tesadüfi hatadan ibaret bir testte ise 0,00 değerini alır.

Son iki denklemden :

$$(3.7) \quad S_e = S_x \sqrt{1 - r_x}$$

bağıntısı bulunur. Buradaki S_e hataların standart kayması olup, ölçme dilinde **ölçmelerin standart hatası** deyimiyle anılır. Güvenirlik katsayısından farklı bir hata ölçüsü olan bu istatistik, hata miktarını bir oran olarak değil, test puanı biriminde gösterir.

Test puanlarının standart kayması ve güvenilirlik katsayısı biliniyorsa, bu istatistik kolayca hesaplanabilir. Güvenirlik katsayısı 0,84 ve standart kayması 12 olan bir testin puanlarında $S_e = 12\sqrt{1-0,84} = 4,8$ puan hata vardır. Bu testte bir kişinin puanı $X_i = 114$ olsun. Ölçme hatalarının bu puan etrafında normal dağıldığı kabul edilirse, o kişinin gerçek puanının, yüzde altmış sekiz ihtimalle, $X_i - S_e = 114 - 4,8 = 109$ ile $X_i + S_e = 114 + 4,8 = 119$ sınırları arasında bulunduğu söylenebilir.

Eğitimde kullanılan ölçme araç ve yöntemlerinin güvenirliliğini artırmak için alınabilecek tedbirler aşağıda sıralanmıştır.

1. Bir sınavda kullanılan soru sayısı arttıkça, çoğu halde, o sınavdan elde edilen toplam puanın güvenilirliği de artar. Yukarıda sözü edilen öğrencilerden biri metal çubuğun boyunu beş defa ölçecek yerde elli defa ölçseydi, her defasında gittikçe artan hatalar yapmamak kaydıyla, hem gerçek değere daha yakın bir ortalama değer bulurdu, hem de elli ölçmedeki hata oranı daha az olurdu. Zaten ölçmeleri tekrar etmedeki amaç da tesadüfi hataların ortalama değer üzerindeki etkisini azaltmaktır. Bir sınıfta yapılan sınavın puanları için güvenilirlik, bir bakıma, puanların sınıftaki öğrencileri başarılarına göre sıralayabilme gücüdür. Bir tek soruyla yapılmış ve «doğru - yanlış» olarak puanlanmış bir sınav, o sınıftaki öğrencileri ancak iki gruba ayırabilir. Dört sorulu bir sınav da, aynı yöntemle puanlandığında, öğrencileri, hiç doğru cevabı olmayandan dört doğru cevaba kadar en çok beş gruba ayırabilir. Soru sayısı arttıkça bu gruplama bir sıralamaya yaklaşır ve puanların güvenilirliği artar. Bir ders saatında cevaplandırılabilen iki soru yerine, aynı sürede cevaplandırılabilen on soru ile yapılan bir yazılı yoklama daha güvenilir ölçme sonuçları verebilir. Aynı sebeple, 60 maddelik bir testin, benzer nitelikte sorulardan oluşan 40 maddelik bir testten daha güvenilir sonuçlar vermesi beklenir.

2. Bir sınavda kullanılacak soruların açıkça anlaşılır ve kesinlikle cevaplanabilir olması, o sınavdan elde edilen puanın güvenilirliğini artırır. Sınav yazılı yoklama şeklinde ise ve sorulardan bazıları cevaplayıcılarca anlaşılmamışsa, cevaplayıcılar bu soruları dedikleri veya anladıkları yönde yorumlar; istenileni değil, bildiklerini cevap olarak yazarlar. Böyle cevapları objektif bir yöntemle puanlamak imkânsızdır. Böyle bir cevaba verilecek puan puanlayıcının takdirine kalır. Objektif olmayan bir ölçmenin güvenilirliği, genellikle düşük olur. Sınav objektif yöntemlerle puanlanabilen bir testle yapılsa bile, test maddelerinin bazıları anlaşılmıyorsa, cevaplayıcılar o maddeleri gelişigüzel cevaplandırır. Gelişigüzel cevapların bir-biriyle tutarlı olması beklenemez.

3. Bir sınavda cevaplayıcılar her soruyu dikkatle ve hızla cevaplandırmaya teşvik edilmelidir. Öğrencilerin yeterince güdülenmediği sınavlarda cevaplar tesadüfe kalacağı, ya da birçok soru cevapsız bırakılacağı için, sınav puanlarının güvenilirliği düşer. Çünkü ister cevaplayıcıların güdülenmemiş olmasından, ister zamanın yetersizliğinden olsun, birçok sorunun cevapsız bırakılmış olması sınavdaki soru sayısını küçültür; bu da sınav puanlarının güvenilirliğini düşürür.

4. Sınav süresi öğrencilerin hemen hepsinin bütün soruları cevaplandırmalarına yetecek uzunlukta olmalıdır. Soruların tümünün cevaplandırılmadığı bir yazılı sınavın güvenilirliği, bundan önceki maddede belirtildiği gibi, cevaplanan soru sayısı azaldığı için düşer. Objektif testlerle yapılan sınavlarda sürenin gereğinden kısa verilmesi, cevaplayıcıların testin sonundaki soruları okumadan ve gelişigüzel cevaplandırmalarına sebep olur. Gelişigüzel verilen cevapların güvenilirliği ise sifara yakındır.

5. Bir sınavda kullanılacak sorular, o sınavı cevaplayacak bireylerin yaklaşık olarak yarısı tarafından doğru cevaplandırılacak güçlükte olmalıdır. Çok kolay sorulardan oluşan bir sınav, yoklanan grubun öğrencilerini birbirinden yeterince ayıramaz. Çok güç sorulardan oluşan bir sınav da öğrencileri birbirinden yeterince ayıramaz. Üstelik, soruların çok güç olması cevapların gelişigüzel veya soru dışı verilmesine sebep olur.

6. Her sınav objektif yollarla puanlanabilmelidir. Bir puanlamanın, daha genel olarak bir ölçmenin, objektif olması, ölçme uzmanlarının aynı sonuçlarda birleşmesi demektir. Sınavlar halinde objektiflik, sınav kapsamında yetkili olmak kaydıyla, bütün puanlayıcıların aynı cevaba aynı puanı vermesi anlamına gelir.³ Ölçme sonuçlarının objektifliği arttıkça güvenilirliği de artar. Çünkü, objektif yöntemlerle puanlanamayan sınavların sonuçlarında puanlayıcı hataları bulunur. Bu hataların bir kısmı tesadüfi hata türünden olup güvenilirliği düşürür.

7. Ölçme işleminin herhangi bir basamağında yapılan dikkatsizlik hataları, tıpkı puanlama hatalarında olduğu gibi, çoğu halde tesadüfi hata türündedir ve ölçme sonuçlarının güvenilirliğini düşürür. Ölçmenin yapılmasında gösterilecek titizlik ve dikkat hemen her zaman güvenilirliği artırır.

8. Duyarlılığı yüksek bir araç veya yöntem daima güvenilirliği yüksek sonuçlar verir. Güvenirliği artırmak için, duyarlılığı yeterince yüksek araç veya yöntemler kullanılmalıdır.

9. Ölçme sonuçlarının, ölçme işleminin elverdiği duyarlılıkla kaydedilmesi güvenilirliği gerçek değerine yaklaştırır. Duyarlılıkla ölçülebildiği halde, sonuçların kabaca verilmesi güvenilirliğin hatalı kestirilmesine sebep olur. Yukarıdaki misalde Ahmet ve Bilge'den sonuçları ondalık basamaksız ve santimetre olarak kaydetmeleri

³R.L. Ebel, *Essentials of Educational Measurement*, s. 372.

istenseydi, bütün sonuçlar 12 cm olarak verileceğinden, bu iki öğrencinin hata yapmadıkları, diğer bir deyişle güvenilirliklerinin mükemmel olduğu kanısı uyanacaktı. Şüphesiz, bu kanı ancak santimetre basamağındaki ölçmeler için doğru, milimetre basamağına inildiğinde yanlış olurdu. Ölçme sonuçlarını ölçme işlemindeki duyarlıktan ötede bir duyarlılıkla kaydetmeye çalışmak da güvenirliliğin hatalı kestirilmesine sebep olur. Ahmet ve Bilge'den, sonuçları, Cengiz'inki gibi, yüzde bir santimetreye kadar duyarlılıkla kaydetmeleri istenseydi, kullandıkları araç ve yöntemle santimetrenin yüzde birini ayırt edemeyeceklerinden, ölçmelerin yüzdelik basamakları tümüyle tesadüfi hata olurdu. Bu ölçmelere dayanan güvenilirlik de gerçek değerinden küçük bulunurdu. Cevabında sadece beş ana fikri ayırt edilebilen bir yazılı yoklama sorusunun cevaplarını 0 - 10 veya 0 - 20 gibi gereğinden fazla ayrıntılı bir ölçekle puanlamak tesadüfi hataları artırır. Aynı cevapları 0 - 1 gibi yalnız iki değeri bulan bir ölçekle puanlamak ise, duyarlık kaybına sebep olur.

ÖLÇMEDE GEÇERLİK

Eğitimde kullanılan araç ve yöntemlerle ölçülmek istenen değişkenler bireylerin bazı nitelikleridir. Bu nitelikler, çoğu halde, öğrenci başarısı, bireyin yetenekleri gibi deyimlerle tanımlanmaya çalışılan değişkenlerdir. Aracın veya yöntemin, ölçülmek istenen değişkenin ölçüsü olabilecek bir puan vermesi, bu ölçüyü başka değişkenlerle karıştırmaması beklenir. Bir araç ve yöntem bu niteliği sağlayabildiği derecede geçerli puanlar verir. Bir lise öğrencisinin fizik dersindeki başarısı ölçülmek istensin. Bunun için de bir sınav yapılacak olsun. Bu sınavın, fizik dersindeki başarıyı tanımlayan bilgi, beceri, tutum, vb. gibi değişkenlerin tümünü ölçmesi; ölçtüğü bu değişkenleri fizikteki başarının öğelerinden olmayan diğer değişkenlerle karıştırmaması istenir. Sınav bu nitelikteyse, onun verdiği puanlar geçerlidir. Sınav bu niteliklerini kaybettiği oranda geçerliliğini de kaybeder. Sınavda sorulacak soruların derste öğretilmiş konuları öğrenmeden de cevaplandırılacak nitelikte olması, sınav sorularının derste işlenen konuların tümünü içine almaması veya dengeli olarak kapsamaması, kopya yapılmış olması, puanlamada yanlış davranılması; puanlamanın güvenilir olmaması, vb. gibi birçok etken bu sınavdan elde edilecek puanların geçerliliğini düşürür.

Ölçme sonuçlarının geçerliği, güvenilirlikte olduğu gibi, ölçme yapılan öğrenen grubunun niteliklerine ve ölçme ortamına bağlı olarak değişir. Üstelik, ölçme sonuçlarının geçerliği, ölçümlerin kullanılacağı amaca, yani değerlendirme ve karar verme işleminin gerek-

terine de bağılıdır. Yukarıda sözü edilen fizik sınavının puanları, o derste «geçti - kaldı» kararları için geçerli olur da, burs verilecek üstün yetenekli öğrencileri seçmek için geçerli olmayabilir.

Eğitimde çok kullanılan yazılı yoklama, sözlü yoklama, kısa cevaplı ve seçmeli testler gibi araç ve yöntemler için geçerlik problemi, hemen her zaman, bu araçlardan elde edilen puanların, ölçülme istenen davranışları ne derece doğrulukla ölçtüğünün kestirilmesine dönüşür. Bu araç ve yöntemler, çoğu halde, bir derste o dersin programına uygun olarak geliştirilmesi hedef alınan davranışları ölçmek için uygulanır. Söz konusu davranışlar ister bazı bilgi ve beceriler, ister tutumlar ve diğer duyuşsal davranışlar olsun, **hedef davranışlar** adını alır. O halde, bir test bir programdaki başarıyı ölçmek için hazırlanmışsa, o test o programın hedef davranışlarını yeterince temsil edebilmeli, onları başka davranışlarla karıştırmamalı ve yeterince duyarlıkla ölçmelidir. İşte bu gereklilikten dolayı, bir programdan türetilen hedef davranışların ölçülmesinde geçerlik problemi, çoğu halde, o test veya sınavın sorularını doğru cevaplandırmakla ortaya konan davranışların, programın hedef aldığı davranışları ne dereceye kadar temsil ettiğinin tayinine dönüşür. Bu işin yapılması bazı hallerde kolaydır. Bir yabancı dil dersinde program, 50 kelimenin Türkçe anlamlarının bilinmesini hedef alıyorsa, bu kelimelerin tümünün anlamlarını yoklayan bir testin geçerliği, başka etkenler bulunmamak kaydıyla, mükemmel olacaktır. Bu misalde hedef davranışlarla testi cevaplandırma davranışları aynıdır; çünkü hedef davranışların tümü sınava dahil edilmiştir. Testi cevaplandırma davranışlarının hedef davranışlara eşit olmadığı hallerde geçerliğin kestirilmesi için, o dersi okutan öğretmen, o programı hazırlayan programcı veya ölçme uzmanı gibi profesyonel kişilerin kanılarına başvurulması gerekir.

Bir ölçme aracının kapsamı, onun geçerliğini yükseltecek biçimde seçilebilir. Hazırlanması bitmiş bir aracın da, kapsamına ve içeriğine bakılarak geçerliği kestirilebilir. Ölçme araç ve yöntemlerinin kapsamlarına ve içeriklerine bakılarak varılan geçerlik yargılarına **kapsam geçerliği** denir.

Yazılı yoklama, test, sözlü yoklama vb. gibi ölçme araç ve yöntemleri hazırlanırken önce, bir «test planı» yapılır. Bu plan, o araçla, dersin konularıyla ilgili hedef davranışlardan hangilerinin kapsama-
cağını ayrıntılarıyla gösterir. Böylece, araç veya yöntemin hangi davranışları ölçeceği önceden saptanmış olur. Test planının hedef

davranışlar yönünden programı temsil edici ve dengeli olması gerekir. Hazırlanan testin de plana uygun olması gerekir. Planın geçerliği yeterli görülürse, testin plana uygunluğu da onun geçerliği için kanıt sayılabilir.

Yapımı tamamlanmış bir testin maddeleri incelenerek de kapsam geçerliği hakkında kanıt elde edilebilir. Bu inceleme, Thorndike ve Hagen tarafından da işaret edildiği gibi, «konular» ve «süreçler» olarak iki boyutta yapılmalıdır.⁴ Bir sınavda konular, öğrencinin cevaplandırmaya çalıştığı sorulardır; süreçler ise öğrencinin o soruları cevaplandırırken yapmaya mecbur olacağı davranışlardır. Kapsam incelenirken testteki soruların hangi konulardan geldikleri tayin edilmeli, bunların, o sınavda ölçülmek istenen konuları tam ve dengeli temsil edip etmediğine bakılmalıdır. Böylece, bir sömestredeki öğrenmeleri yoklamak için hazırlanan bir testte, o sömestr işlenen konuların önemleri oranında teste girdiği görülürse, test kapsamının program kapsamını temsil ettiği ve bu sebeple testin kapsam geçerliğine sahip olduğu kanısına varılır.

Bir öğrencinin sınavdaki bir soruyu cevaplandırırken gösterebileceği davranışlar üç grupta toplanabilir. Öğrencinin önce doğru çözüm yöntemini bulması, sonra bu yöntemi uygulayarak doğru çözüme varması, en sonunda da çözümünü puanlayıcıya iletmesi beklenir. Üçüncü basamak, sınavın türüne göre, sözlü veya yazılı cevap, ya da doğru cevabı işaretleme şeklinde olur. Bir sınavda ölçülmek istenen davranışlar, ancak nadiren bu üçüncü gruptakilerdir. Çoğu halde yoklanmak istenen davranışlar, ilk iki gruptaki doğru cevaba ulaşma davranışlarıdır.

Yukarıda sözü edilen yazarlar, «süreçlerin incelenmesi» deyişiyle, test sorularının doğru cevaplandırılabilmesi için gerekli davranışlar zincirinin incelenmesini kastetmişlerdir. Bir sınavın soruları teker teker incelenmeli, her bir sorunun doğru cevaplandırılabilmesi için ne gibi davranışlar gerektiği meydana çıkarmalı, işte bu davranışların programda hedef alınan davranışlar olup olmadığına bakılmalıdır. Böylece, testin gerektirdiği cevaplama davranışlarının programın hedef davranışlarını temsil ettiği görülürse, o testin geçerliğe sahip olduğu kanısına varılır. Eğitimde, özellikle öğrenci başarısının ölçülmesinde kullanılan araçların geçerliği için en inandırıcı kanıt

⁴R.L. Thorndike and E. Hagen, *Measurement and Evaluation in Psychology and Education*, 3rd ed., s. 164.

bu yolla elde edilebilir. Aşağıda değinileceği gibi, bu araçların hazırlanmasında kapsam geçerliği de bu yolla sağlanabilir.

Bir ölçme aracının geçerliğini düşüren diğer bir etken, o araçtan elde edilen puana, ölçülmek istenen değişkenden başka etkilerin karışmasıdır. Bir örnek olarak, öğrencilerin tarih dersindeki başarılarının bir yazılı yoklama ile ölçüleceği farzedilsin. Tarih dersinin hedefleri arasında fikirleri yazılı olarak ifade etmek, okunaklı yazı yazmak gibi davranışlar bulunmadığı kabul edilsin. Yazılı yoklama puanlanırken, fikirlerin kolay anlaşılır bir sıraya konulmuş olması puana etki eder. İyi ifade edilmiş bir cevaba da yüksek puan takdir edilmesi çok sık görülen bir olaydır. Böylece, tarih dersindeki başarının göstergesi olan puana yazılı ifade ve yazı okunaklılığı gibi değişkenler de girmiş olur. Bu yazılı yoklamanın amacı sadece tarih bilgisini yoklama olduğundan, ifade düzgünlüğü ve yazı okunaklılığı, birer «istenmeyen değişken» olarak, yazılı yoklama puanlarının geçerliğini düşürür.

Sınav araç ve yöntemlerinin geçerliğini düşüren istenmeyen değişkenler çok çeşitlidir. Meselâ, seçmeli testlerde cevaplama zamanının sınırlı olduğu hallerde, testteki bütün soruları okumaya zaman bulamayan bir öğrenci, okuma hızının yavaşlığından dolayı puan kaybeder. Aynı şekilde, sınavın yapıldığı dilde yetersizlik, cevapların kaydında karşılaşılan güçlük veya dikkatsizlik de istenmeyen değişkenler olarak test geçerliğini düşürür. Bu kitabın ilerideki bölümlerinde çeşitli ölçme araç ve yöntemleri incelenirken, her biri için geçerliği düşürücü etkiler de tartışılmaktadır.

Herhangi bir ölçme aracının geçerliğini kanıtlamakta içerik ve kapsama bakmak önemli bir yol olmakla birlikte, biricik yol değildir. Araç veya yöntemin verdiği puanlar üzerinde yapılacak istatistiksel analiz ve araştırmalarla da test geçerliği kanıtlanabilir. Bu yöntemlerden birkaçı aşağıda kısaca özetlenmektedir.

Tarih dersindeki başarıyı ölçmek için uzmanlarca hazırlanmış ve geçerliği kanıtlanmış bir test bulunsun. Öğretmen bu testi ve kendi hazırladığı yazılı yoklamayı, kısa bir zaman aralığıyla arka arkaya, sınıfında uygulamış olsun. Yazılı yoklama puanlarıyla test puanları arasında pozitif ve yeterince yüksek bir korelasyon bulunursa, testin geçerliği önceden kanıtlandığı için, yazılı yoklama puanlarının da geçerliğe sahip olduğu kanısına varılır. Bu misalde geçerliği kanıtlanacak olan puanlar yazılı yoklama puanlarıdır; test puanlarına ise **geçerlik ölçütü** denir. Böyle bir yöntemle geçerlik kanıtlamanın en önemli gücü, geçerliği önceden kanıtlanmış ölçüt

puanları bulabilme güçlüğünden gelir. Yıgmalı başarı ortalaması, sınıf geçme notu, not ortalaması gibi birçok ölçü, eğitimde kullanılan başarı testlerinin geçerliğini kanıtlamada geçerlik ölçütü olarak kullanılır. Diğer bir güçlük, ölçüt puanlarının geçerliğinin sağlam bir yoldan kanıtlanmamış olmasından doğar. Çoğu halde bu kanıtlama, uzman kanısına veya ölçüt ile testin ölçtüğü davranışlar arasında tutarlı ve mantıklı bir ilişki bulunmasına dönuşür.

Geçerliği kanıtlanacak olan puanlarla ölçüt puanları arasındaki ilişki, korelasyon katsayıları ile hesaplanır. Bu haldeki korelasyon **geçerlik katsayısı** adını alır. Geçerlik katsayıları $-1,00$ ile $+1,00$ arasında değerler alabilir. Geçerlik katsayısının negatif olması testin önemli bir kusuru bulunduğunu, sıfır olması geçersizliğini gösterir. Test, geçerlik katsayısı pozitif ise ve katsayının yüksekliği oranında geçerlidir.

Geçerlik katsayıları sadece test puanları ile ölçüt puanları arasındaki ilişkinin zayıflığından değil, her iki puanın güvenilirliklerinin tam olmayışından da düşer. Bu sebeple, geçerlik katsayıları, güvenilirlik katsayılarıyla birlikte yorumlanır. Hangi yükseklikteki bir korelasyonun yeterli bir geçerlik kanıtı sayılacağı, test puanlarının kullanılış amacına bağlıdır. Test puanlarıyla çok kaba bir ayırım yapılacaksa, düşük bir geçerlik katsayısı bile yeterli sayılabilir.

Test geçerliğini kanıtlamada başvuru olan diğer bir yöntem, ölçüt puanı yerine karşılaştırma grupları ve **ölçüt grupları** kullanmaktır. İlkokul üçüncü sınıflar için bir Türkçe testi hazırlanmış olsun. Bu test, üçüncü sınıflarla birlikte dördüncü ve beşinci sınıflara da uygulandığında, test puanları sınıf düzeyi yükseldikçe büyüyen ortalamalar verirse, testin geçerliğe sahip olduğu kanısına vurılır. Bu halde test, ölçüt gruplarını birbirinden ayırabildiği oranda geçerlidir. Benzer bir yöntem, modern fen programlarının değerlendirilmesinde kullanılan testlerin geçerliğini kanıtlamada uygulanmıştır⁵.

Yukarıdaki örnekte Türkçe yeteneklerinin sınıf düzeyi ile oranlı olarak arttığı kabul edilmiştir. Ölçüt gruplarıyla geçerlik kanıtlamada test geçerliği, bu türden sayıtların doğruluğuna da dayanır. Gerçekte Türkçe yetenekleri sınıf düzeyi yükseldikçe artmıyorsa,

⁵M.F. Turgut, «Pilot Liselerde Denenen Modern Matematik ve Fen Programlarının Değerlendirilmesi», bulunduğu eser : Türkiye Bilimsel ve Teknik Araştırma Kurumu, III. Bilim Kongresinde Bilim Adamı Yetistirme ve Fen Öğretimi Sektörüne Sunulan Tebliğler, s. 35 - 59.

test puanlarında sınıftan sınıfa artan ortalamalar gözlemlense bile bu artış geçerliğe kanıt olamaz.

Ölçme sonuçlarının geçerliğini sağlamak için, gerek araçların yapımında gerekse yöntemlerin uygulanmasında, alınabilecek önlemlerin başlıcaları şunlardır :

1. Kapsam geçerliği sadece konular açısından değil, testteki soruları doğru cevaplandırabilmek için gerekli cevaplayıcı davranışları açısından da anlam kazanır. Bir araç, kapsadığı soruların geçerliği oranında geçerli olabilir. Onun için, her soru, o araçla ölçülmek istenen davranışlardan en az bir tanesini yoklamalıdır.

2. Sınav soruları, kopya, ipuçlarından yararlanma, tahminle doğru cevabı bulma gibi geçerli olmayan davranışlarla da doğru cevaplandırılabilir. Gerek uygulamada, gerekse soru hazırlamada bu etkiler azaltılabilir. Onun için, her soru, o soruyla ölçülmek istenen bilgi ve becerilere sahip öğrencilerin tereddütsüz doğru cevaplandırabileceği; bu bilgi, beceri ve yeteneklere sahip olmayanların doğru cevaplandıramayacağı nitelikte olmalıdır.

3. Sınav belli bir sürede okutulan konular ve o konulara dayanılarak geliştirilmesi beklenen bilgi, beceri ve yetenekler bakımından hem kapsayıcı, hem de dengeli olmalıdır. Kapsayıcılık niteliğiyle sınavın bütün konuları, bilgi, beceri ve yetenekleri temsil etmesi istenir. Bu nitelik hem konuların tümünden, hem de hedef davranışların tümünden iyi bir örneklemeye sağlanır. Dengelilik niteliği ise, sınavın, her konuya ve her hedef davranışa, örneklemeye, programdaki önemleriyle oranlı bir yer vermesiyle sağlanır.

4. Herhangi bir puanın geçerliği için önşart, onun güvenilirliğidir. Güvenirlikten yoksun bir puan sadece tesadüfi hataları gösterir; hata da geçerli bir ölçü olamaz. Her ölçme sonucu ancak güvenilirliği oranında geçerlik kazanır. Onun için, güvenilirliği artıran bütün önlemler geçerliği de artırabilir. Fakat güvenirliliğin sağlanmış olması geçerliğin de mutlaka sağlanacağı anlamına gelmez. Bir öğrenci grubunda baş çevresi yeterince duyarlılıkla, meselâ milimetre basamağına kadar hatasız olarak ölçülebilir. Bu ölçme sonuçları her öğrencinin başının şekli hakkında kaba bir fikir verebilir. Şüphesiz, bu ölçülere dayanarak her öğrencinin kaç numara şapka giyeceği kesinlikle tayin edilebilir. Fakat, baş anormal derecede büyük veya küçük değilse, bu baş ölçülerine dayanılarak, öğrencilerin zekâları kestirilemez. Güvenirliliği ne kadar yüksek olursa olsun, baş ölçüsü şapka numarasını tayin için geçerli bir değişken olduğu halde, zekâ kestirmek için geçerli değildir.

Ölçme araçlarında «homojenlik» anlamındaki güvenilirliği artırmak, bazan geçerliği düşürebilir. Bir fen bilgisi testi fizik, kimya, biyoloji, astronomi gibi alanlardan soruları içeriyorsa ve astronomi sorularındaki öğrenci başarısı testin diğer kısımlarıyla uyuşmuyorsa; bu soruların testten çıkarılması homojenliği artırdığı için güvenilirliği de artırır; fakat, testin dengesi bozulduğu için de kapsam geçerliğini düşürür.

5. Sınavın ve ölçme araçlarının güclüğü de geçerliğe etki eder. Güçlük derecesi o sınavdan elde edilecek puanların kullanılacağı amaca göre ayarlanmalıdır. Matematik problemleri çözmede okul birincisi, ikincisi ve üçüncüsünü seçmek amacıyla kullanılacak bir yarışma sınavının sorularının, bütün okulda ancak birkaç öğrencinin çözebileceği kadar güç olması gerekir. Sorular kolay olursa, birçok öğrenci tarafından sınavın tümü doğru cevaplandırılacağı için, okul birincisini ayırt etme imkânı kalmaz. Eğitimde sınavlar, çoğu halde, öğrencilere not takdir etmek için kullanılır. Öğrencilere 0 ile 10 arasında not vermek için kullanılacak bir sınavın soruları, çeşitli güçlüklerde olmalı, fakat tüm soruların ortalama güclüğü yüzde elli civarında bulunmalıdır. Çok kolay veya çok zor sorulardan oluşan sınavlar, not takdirinde genellikle geçerliği düşük sonuçlar verir.

6. Aynı soruların yıldan yıla hiç değiştirilmeksizin kullanılması, soruların sınavdan önce açıklanması, derslerin sınav sorularını çözmeye ayrılması, kopya yapılması, sınav süresinin yetersizliğinden dolayı bir kısım soruların cevapsız bırakılması, öğrencilerin çeşitli nedenlerle sınavı yarıda bırakması, öğrencinin gürültü ve hastalık gibi nedenlerle soruları tüm yetenekleriyle cevaplandıramaması, vb. gibi etkenler geçerliği düşürür.

ÖLÇMEDE KULLANIŞLILIK

Ölçme araç ve yöntemlerinin kullanılışlılığı hakkında bir yargıya varılırken şu noktaların dikkate alınması gerekir :

1. **Ekonomi** : Araç veya yöntem, maliyet yönünden ekonomik olmalı, öğretmene, öğrenciye, yönetime ağır malî külfet yüklememelidir. Şüphesiz maliyeti aşağıya indirmek için, sınavın geçerliğini yetersiz bir dereceye kadar düşürmek de savunulamaz. Onun için, bir okulun kaldıracabileceği malî külfetle en geçerli sınavı yapmak, maliyet ve geçerlik ilişkisi açısından, en doğru yol olur.

2. **Hazırlama süresi** : Her araç ve yöntem, uygulamadan önce bir hazırlık gerektirir. Testler ve kısa cevaplı araçlar halinde bu sü-

re daha uzun, yazılı yoklamalarda daha kısadır. Hazırlama süresi kısa olan bir sınav, diğer etkenler eşit olmak şartıyla, hazırlama süresi uzun olan bir sınava tercih edilebilir. Bununla birlikte, hazırlama süresinin, puanlama süresiyle birlikte dikkate alınarak değerlendirilmesi gerekir. Yazılı yoklamaların hazırlanması az zaman alır; fakat, kalabalık gruplarda yazılı yoklama cevaplarının puanlanması hazırlıktaki zaman tasarrufunu kat kat aşar. Ayrıca hazırlama süresi, sınavı hazırlayacak kimselerin uygulamadan önce ne kadar zamanı olduğuna göre de değerlendirmelidir. İki gün sonra uygulanması zorunlu olan bir sınav için iyi bir test hazırlamaya imkân yoktur.

3. Uygulama süresi : Her araç veya yöntem ders saatlerinin bir kısmının uygulamaya ayrılmasını gerektirir. Bir yöntemin kullanışlılığı, verdiği puanın önemine oranla, uygulamada gerektirdiği zaman azaldıkça artar. Bir yazılı ve bir sözlü sınav aynı geçerlikte puanlar verse bile, sınav, uygulama süresi kısa olduğu için, çok küçük olmayan gruplarda daha kullanışlıdır.

4. Hazırlayıcı ve uygulayıcıların nitelikleri : Ölçme araçlarının bazıları, hazırlayıcıda ve uygulayıcıda bazı özel bilgi ve beceriler bulunmasını gerektirir. Bu bilgi ve becerilere sahip olmayan kimselerin hazırladıkları araçlardan, uyguladıkları sınavlardan çoğu zaman hatalı sonuçlar elde edilir. Bir aracın uygulanması ne kadar özel beceriler gerektirirse, kullanışlılığı da o kadar sınırlanır. Öğretmenin kendi dersi için hazırlayacağı araçlardan bazıları, meselâ testler, uzmanlık seviyesinde olmasa bile bazı bilgiler gerektirir. Bu bilgileri kazanmamış bir öğretmen için testler kullanışlı değildir.

5. Cevaplayıcının nitelikleri : Ölçme araçlarından yeterince hatasız puan alınabilmesi için, cevaplayıcıda bulunması gerekli bazı nitelikler vardır. Öğrenciler cevaplama yönergesini kolayca anlayabilir ve cevaplarını kolayca kaydedebilirlerse, o araç kullanışlıdır.

6. Uygulama kolaylıkları : Testlerde sayfa düzeni, teksir ve baskı mükemmelliği, cevap kâğıdının düzeni; yazılı yoklamalarda soruların yazılı verilmesi, vb. gibi hususlar uygulama kolaylıkları sağlar; bu yollarla da araçların kullanışlılığı artırılabilir.

7. Puanlama kolaylıkları : Testlerde cevap kâğıdının testten ayrılması, cevap kâğıdının iyi düzenlenmesi, cevapların makinayla puanlanabilmesi; diğer ölçme araçlarında ve yöntemlerde puanlama anahtarı gibi yardımcı araçların kullanılması puanlama kolaylıkları sağlar. Böylece, araçlar ve yöntemler daha kullanışlı hale gelir.

8. Puanları yorumlama kolaylıkları : Bir aracın uygulanmasıyla elde edilen puan anlamlı bir biçimde yorumlanamazsa, değerlendirme ve karar vermede kullanılamaz. Bir araç, yorumlamaya yardımcı, yorumlamayı kolaylaştırıcı ve yorumda hatayı azaltıcı araçlarla beslenirse daha kullanışlı olur. Meselâ, öğretmenin hazırladığı bir yazılı yoklamada, puanlama anahtarının yanında, hangi puandan «geçti - kaldı» kesimi yapılacağını belirtir bir şey de bulunursa, o puanlara göre not verme kolaylaşır. Standart testlerde sadece uygulama yöntemlerinin saptanmasıyla kalınmayıp, ham puanların daha anlamlı ölçeklere çevrilmesini ve böylece puanların yorumunu kolaylaştıran norm tabloları da hazırlanır. Böyle araçlar, test puanının yorumlanmasını kolaylaştırır, test sonuçlarının kullanılmasındaki hataları azaltır. Ayrıca, çeşitli bölümlerden oluşan bir testin, her bölüm için ayrı bir puan vermesi de kullanışlılığı artıran önemli etkenlerdendir.

Yukarıda tartışılan sekiz nitelikten de anlaşılacağı gibi, bir aracın kullanışlılığını artıran önlemler, dolaylı olarak onun güvenilirlik ve geçerliğini de artırır.

B Ö L Ü M 4

Yazılı Yoklama

Yazılı yoklama, geleneksel eğitim sistemlerinde çok kullanılan bir sınav türüdür. Bazan «kompozisyon tipi imtahanlar» diye de anılan bu sınav türünün temel özelliklerinin ve sınırlılıklarının bilinmesi, öğretmenin daha iyi ölçme yapmasına yardım edebilir. Bu bölümde yazılı yoklama bir ölçme yöntemi olarak her yönüyle tartışılmakta, okullarımızda çok kullanılan bu yöntemin mümkün olduğu kadar hatasız yapılması için bazı öneriler verilmektedir.

YAZILI YOKLAMA NEDİR?

Öğrencilere birkaç soru yazdırılıp veya yazılı verilip bunlara belli bir sürede yazılı cevap istenmesi, bir sınav türü olarak her öğretmenin baş vurduğu bir yöntemdir. Böyle bir sınavda öğrenciden istenen işler, cevapları düşünmek, tasarlamak, düzenlemek, yazmak ve gerekirse yazdıklarını kontrol etmektir. Öğretmene düşen işler ise, sınavdan önce soruları hazırlamak, sınavı yönetmek, sınavdan sonra cevapları okuyup puanlamaktır. Stalnaker tarafından da belirtildiği gibi, yazılı yoklama yöntemi, kısa cevaplı sınav yöntemlerinden cevapların bir veya birkaç paragraf gibi uzunca bir kompozisyon gerektirmesiyle ayrılır.¹ Bu sınav türünün ölçme ve değerlendirme amacıyla verilen yazılı ev ödevlerinden farkı ise, cevaplamanın oldukça standart şartlarda ve belli bir sürede yapılmasıdır.

Cevapların yazılı verilmesi zorunluğu, cevapların uzunluğu, cevaplayıcı bağımsızlığı, puanlamanın öğretmen veya puanlayıcı kanısına dayanması gibi özellikler yazılı yoklamaların niteliklerini etkileyen önemli etkenlerdir. Bunlar, çoğu halde, yazılı yoklamaların bir ölçme aracı olarak yeterince güvenilir ve geçerli sonuçlar vermesini engelleyen etkenlerdir. Böyle olmakla birlikte, yazılı yoklama yönteminin temel özelliklerinin bilinmesi, yukarıda sayılan etkenlerin

¹J.M. Stalnaker, «The Essay Type of Examination», in E.F. Lindquist (Ed.), *Educational Measurement*, s. 495.

etkilerinin anlaşılması, ölçme güvenirlik ve geçerliğini düşürücü etkilere karşı önlem alınması, yazılı yoklamaları daha kullanışlı hale getirebilir. Öğretmenlerin yazılı yoklamaları devamlı olarak kullanmaları, onların bu tekniğin sakıncalarını sınama ve yanılma yoluyla azaltabilecekleri anlamında yorumlanmamalıdır . Öğretmenin bu sınav türünü kullanmada iyi bir tecrübeye sahip olabilmesi için, başlangıçta bu sınav yönteminin özelliklerini iyice bilmesi gerekir.

YAZILI YOKLAMALARIN ÖZELLİKLERİ

Yazılı yoklamaların başlıca özellikleri şöyle sıralanabilir :

1. Yazılı yoklamada cevaplayıcı, cevabını kendi düşünüp bulmak zorundadır. Oysaki, cevaplayıcı objektif testlerin bazı türlerinde verilen bir cümlemin doğru olup olmadığına karar vermek, bazılarında ise verilenler arasından doğru cevabı seçmek zorundadır. Bu özellik açısından bakılınca, yazılı yoklama ile daha köklü bilgiler ölçüldüğü görünümü uyanır. Fakat, ileride tartışılacağı gibi, bu görünüm çoğu halde doğru değildir. Aslında sınav türlerinin her biri, iyi hazırlanmak kaydıyla, köklü bilgilerin ve ileri basamaklardaki zihin yeteneklerinin yoklanmasında etkili olabilir.

2. Yazılı yoklamada cevaplayıcı, cevabını yazılı olarak ifade etmek zorundadır. Bu özellikten dolayı her yazılı yoklamada, sınav hızı, vb. gibi bazı değişkenler de puana karışır.

3. Yazılı yoklamada cevaplar çoğu halde sınırlı değildir. Her cevaplayıcı, cevaplarını dilediği gibi ifade etme bağımsızlığına sahiptir. Soruların, cevapları sınırladığı hallerde bile konu dışına çıkan, cevapları başka yöne çeken, gereğinden uzun veya çok kısa cevaplar veren cevaplayıcılar görülür. İşte bu özellikten dolayı, yazılı yoklamalarda soruların çeşitli anlamlarda yorumlanması, daha teknik terimiyle soru belirsizliği, sorulmayan şeylerin yazılmasıyla cevapların kasten şişirilmesi gibi olaylar ortaya çıkar.

4. Genellikle, bir yazılı yoklama sorusuna verilen cevapları tamamen yanlış veya tamamen doğru olarak iki sınıfa ayırmak mümkün değildir. Çoğu halde cevaplar tam doğrudan kısmen doğruya ve tamamen yanlışa kadar çeşitli doğruluk derecelerinde olur. İşte bu özellikten dolayı, yazılı yoklamalarda, herhangi bir soruya verilen cevabın doğruluğu ve bu sebeple de takdir edilen puanın doğruluğu tartışmalara yol açar.

5. Yazılı yoklamalarda cevapların doğruluk derecesini tayin etmek puanlayıcıya düşer. Öğretmenlerin notlarının isabetsizliği

üzerindeki gözlemler, bir öğretmenin iki ayrı puanlamada aynı puanları veremeyişi, iki öğretmenin puanları arasında tam uyum bulunmayışı gibi olaylar, yazılı yoklamalarda puanlamanın puanlayıcıya takdir hakkı bırakışından doğar.

6. Cevaplama işlemi, cevaplayıcının çok zamanını alır. Yazılı yoklamalarda sınav süresinin çok büyük bir kısmı cevapların düzenlenip yazılmasına harcanır. Bu sebeple, bir yazılı yoklamada sorulabilecek soru sayısı, hatta sorulabilecek soruların çeşitleri kısıtlıdır. Böylece, yazılı yoklamalarla, objektif testlere kıyasla; aynı sınav süresinde daha dar bir alan yoklanabilir.

7. Yazılı yoklamaların puanlanması güçtür. Bu güçlük hem her cevabın dikkatle okunması zorunluluğundan, hem de cevapların kesinlikle doğru veya kesinlikle yanlış olarak sınıflanamayışından doğar. Ayrıca, puanlayıcı kanısı gerektiği için, puanlama işlemine çeşitli hatalar karışır.

8. Bir yazılı yoklamanın hazırlanması için gerekli zaman, genellikle, diğer ölçme araçlarınınkinden kısa, uygulamadan sonra yapılacak puanlama işlemleri için gerekli zaman ise oldukça uzundur. Bu özellikten dolayı yazılı yoklamalar çok sayıda adayın katıldığı sınavlar, kalabalık sınıflar ve öğretim yükü çok fazla olan öğretmenler için kullanışlı değildir.

9. Yazılı yoklama sorularının güçlük derecesini objektif yöntemlerle tayin etmek zordur. Soru güçlüğü gerekli kesinlikle kestirilemeyince, soruların güçlük derecelerine dayanan test istatistikleri, yazılı yoklamalar için kolayca ve doğru olarak hesaplanamaz.

YAZILI YOKLAMALARDA GÜVENİRLİK

Yazılı yoklamalar bilgi veya başarı ölçme araçları olarak dikkate alınırsa, yazılı yoklama puanlarının ne derece güvenilir olabileceği, bu araçların temel özelliklerinden hareket edilerek tartışılabilir. Ayrıca, yazılı yoklamaların güvenilirlikleri araştırmalarla ölçülebilir. Bu kesimde yazılı yoklamaların güvenilirliğini etkileyen başlıca etkenler tartışılmaktadır. Yazılı yoklamalarda güvenilirlik araştırmalarından bazılarına, aşağıdaki kesimde değinilecektir.

Üçüncü bölümde belirtildiği gibi, bir sınavdan elde edilecek puanların güvenilirliği, genellikle, soru sayısına bağlı olarak artar. Bir yazılı yoklamada sorulabilecek soru sayısının kısıtlı oluşu, güvenilirliği sınırlayan önemli etkenlerden biridir.

Yazılı yoklamalarda güvenilirliđi düşüren etkenlerden biri de soru belirsizliđidir. Soruda ne istendiđini kesinlikle anlayamayan bir öğrenci, istenen cevabı deđil, kendi yorumuna göre istenenden farklı bir cevabı yazar. Bazı hallerde, cevabı bilmeyen bir öğrencinin, soruyu başka yana çekmesi ve onunla ilgili gibi görünen başka bir şeyi cevap olarak yazması da mümkündür. Cevabın soruda istenenden farklı olması halinde puanlama güçlükleri doğar. Puanlayıcı, bu tür cevapları puanlamada tutarlı olamaz. Bu sebeple, açıkça anlaşılmamış veya kasıtlı olarak başka yana çekilmiş soruların cevapları da güvenilir bir yöntemle puanlanamaz.

Yazılı yoklamalarda puanlama işlemi, her cevabın okunmasını ve doğruluk derecesinin kestirilip ona göre puan verilmesini gerektirir. Dikkatsizce okunmuş ve gelişı güzel puanlanmış yazılı yoklamalardan alınan sonuçların güvenilirliđi çok düşük olur. Puanlayıcı, cevapları okumada dikkatli olsa bile, cevapların doğruluk derecesini ve her cevaba verilecek puanı takdir hakkı puanlayıcıya kaldıđı için, bir puanlayıcıdan diğeri ve bir puanlamadan diğeri tutarlı sonuçlar almak oldukça güçtür. Puanların, bir puanlamadan diğeri veya bir puanlayıcıdan diğeri hiç deđişmemesi halinde **puanlama güvenilirliđi** mükemmeldir. Yazılı yoklamalar objektif yöntemlerle puanlanamadıđından, puanlama güvenilirliđi hemen hiç bir zaman mükemmel olmaz.

Yazılı yoklamalarda, güvenilirliđin düşük olmasının diğeri bir sebebi de soru güçlüğünün sınavın amacına göre ayarlanamamasıdır. Yazılı yoklamalarda soru güçlüğü istatistiksel yollarla hesaplanamadıđı için, bu ayarlama ancak kabaca yapılabilir. Bir sınavın güvenilirliđi, bütün diğeri etkenler sabit kalmak şartıyla, ancak her soru, sınıfın yüzde ellisi tarafından doğru cevaplandırılabilirdiği takdirde en büyük deđerini bulur. Çok zor sorulardan oluşan bir sınavın güvenilirliđi düşük olacađı gibi, çok kolay sorulardan oluşan bir sınavın güvenilirliđi de düşük olur.

Üçüncü bölümde tartııldıđı gibi, test ve sınav güvenilirliđine etki eden daha birçok etken vardır. Yukarıda kısaca sözü edilen dört etken, yazılı yoklamalardan elde edilecek puanların güvenilirliđi açısından önemlidir. Bu etkenler puanların güvenilirliđini düşürücü yönde işlemekle birlikte, gerek soruların hazırlanmasında, gerekse cevapların puanlanmasında bazı noktaların dikkate alınmasıyla, yazılı yoklamaların güvenilirliđini bir dereceye kadar yükseltmek mümkündür. Bu noktalar aşıđıdaki kesimlerin birinde açıklanmaktadır.

YAZILI YOKLAMALARDA GEÇERLİK

Sınıfta kullanılan ölçme araç ve yöntemleriyle elde edilen puanların geçerliği hiç bir zaman tam olmaz. Çünkü herhangi bir derste kazandırılması amaç edinilen bilgi, beceri ve benzerlerinin ancak bir kısmı sınavlarda yoklanabilir; üstelik, sınavlar hangi araç veya yöntemlerle yapılırsa yapılsın, puanlara, geçersiz sayılabilecek başka etkiler de karışır. Bu hal yazılı yoklamalar için de böyledir. Hatta, yazılı yoklamalarda geçerliği düşüren etkenler, objektif araç ve yöntemlere kıyasla daha fazladır.

Bir yazılı yoklamada sorulabilecek soru sayısının sınırlılığı, sınavın kapsamını daraltır. Bir derste sadece bir sô mestrin dörtte birinde işlenen konulardan sınav yapılacak olsa bile, okutulan konuların ancak bir kısmı yazılı yoklamaya dahil edilebilir. Yazılı yoklamalarda geçerliği düşüren en önemli etken, derste işlenen birçok konunun yazılı yoklamaya alınamamasından doğar.

Bir sınavda sorulması gereken konular sayıca çok, öte yandan sınav süresi sınırlı olunca konular arasından bir seçme yapmak zorunluğu doğar. İşte bu seçme dengeli olmazsa, bazı konulardan gereğinden fazla ve bazılarından az soru sorulursa, o sınav ders konularını tam temsil etmez. Ayrıca, sınavda sorulan soruların derste işlenen konuları temsil etme oranı küçüldükçe de sınav puanlarının geçerliği düşer. Yazılı yoklamaların hazırlanmasında, konular arası dengeyi sağlamak bir dereceye kadar mümkün olabilir. Fakat, soru sayısı çok küçük olacağı için, belli bir dönemde işlenen konuları yeterince temsil edebilecek kadar çeşitli soru sormak hemen hiç bir zaman mümkün olmaz.

Bir sorunun yalnız konusu değil, soruluş tarzı da geçerliği etkiler. Cevaplayıcılar, iyice anlayamadıkları sorulara kendi yorumlarına göre cevap verirler. Sorular kesinlikle anlaşılamıyorsa, verilecek cevaplar sınırlanmamışsa, öğrenciler diledikleri gibi ve diledikleri yönde cevaplar yazarlar. Şüphesiz, sorulmamış bir konu doğru yazılmış olsa bile geçerli bir cevap sayılamaz. Onun için, yazılı yoklamalarda soruların belirsizleştiği oranda geçerlik düştüğü gibi, cevapların sınırlanmadığı oranda da geçerlik düşer.

Yazılı yoklamalarda cevapların yazılması sırasında, sınavda ölçülmek istenen bilgi ve becerilere başka beceriler de karışır. Okunaklı yazabilmek, iyi kompoze edebilmek, kısa ve kesin yazabilmek, çabuk yazabilmek bu tür becerilerin en önemlilerindendir. Bu beceriler de, eğitimin genellikle geliştirmek istediği becerilerdendir; fa-

kat, yazılı yoklama bu becerileri ölçme amacıyla yapılmıyorsa, bunların puanlara karışması ölçmenin geçerliğini düşürür.

Yazılı yoklamaların puanlanması sırasında da geçerlik zedeleyebilir. Puanlayıcının bazı cevaplayıcılara karşı ve bazılarında yavaş davranması, bazı cevaplara gereğinden fazla ve bazılarında gereğinden az puan takdir etmesi gibi **puanlama yanlışlıkları** hemen her zaman geçerliği düşürücü etkenlerdir. Aşağıda değinilecek olan diğer puanlama hataları da geçerliği düşürür.

Üçüncü bölümde belirtildiği gibi, güvenilirliği düşüren etkenlerin hepsi geçerliği de düşürür. Yazılı yoklamaların yeterince objektif yöntemlerle puanlanamadığı hallerde, güvenirlik düşük olacağı gibi, geçerlik de genellikle çok düşük olur.

YAZILI YOKLAMALARIN KULLANIŞILIĞI

Yazılı yoklamalar sınavdan önceki hazırlık zamanı dar olan hallerde kullanışlı gibi görünür. Fakat, yazılı yoklamanın, bu bölümde önerildiği gibi teknik esaslara uyularak hazırlandığı hallerde ve okunacak cevap kâğıtları çoğaldıkça bu üstünlük kaybolur. Soru hazırlamanın nisbeten kolay oluşu, sınavdan önceki soru tesiri yükünün azlığı da yazılı yoklamaların kullanışlılığını artıran etkenlerdendir. Fakat, bunların sağladığı üstünlükler de diğer sakıncaların arttığı hallerde kaybolur.

YAZILI YOKLAMALARIN NİTELİKLERİ ÜZERİNE ARAŞTIRMALAR

Yazılı yoklamaların, yöntemin tabiatından gelen sınırlılıkları bulunduğu, yukarıdaki kesimlerde ortaya kondu. Objektif ölçme araçlarının eğitimde kullanılmaya başlamasıyla yazılı yoklamalar üzerinde yapılan araştırmalar, bu sınırlılıkları ve sakıncaların genellikle doğru olduğunu göstermiştir. Bu kesimde, yazılı yoklamalar üzerinde yapılan araştırmalardan bazılarını değinilmektedir.

Yazılı yoklamalarda güvenilirliği düşüren etkenlerden biri, puanlamanın objektif yöntemlerle yapılamayıpıdır. Onun için, yazılı yoklamalarda puanlama güvenilirliği, birçok deneysel araştırmaya konu olmuştur. Bu tür araştırmaların 1929 yılına kadar yapılmış olanlarını özetleyen Ruch², bağımsız puanlayıcıların takdir ettikleri puanlar arasındaki korelasyonların 0,60 civarında, aynı puanlayıcının iki ay-

²G.M. Ruch, The Objective or New Type of Examination, s. 91 - 93.

rı zamanda takdir ettiği puanlar arasındaki korelasyonların ise 0,40 civarında değiştiğini bildirmiştir.

Thorndike ve Hagen³, bir yurttaşlık bilgisi yazılı yoklama kâğıtlarından yalnız bir tek sorunun cevabını kopya edip, yurttaşlık bilgisi öğretmenlerine 20 üzerinden puanlatmışlardır. Puanlamayı yapan 70 öğretmenin hepsi ölçme ve değerlendirme konusunda iki sömetr süren bir ders görmüş olmalarına rağmen, o cevaba takdir ettikleri puanlar 2 - 19 arasında değişmiştir. Ömer Mart⁴ Devlet Ortaokul İmtihanında cevap olarak verilen Türkçe, Matematik ve Tıbbat Bilgisi kâğıtlarından üç tanesini seçip bu dersleri ortaokullarda öğreten öğretmenlere puanlatmış; 10 üzerinden verilen puanların matematikte 2 - 7 arasında, diğer iki derste ise daha geniş sınırlar arasında dağıldığını görmüştür.

Tiegs⁵ ise, 10 öğrencinin yazılı yoklama cevaplarını aynı öğretmene iki ay ara ile iki defa puanlatmıştır. Bu öğretmen, ayrıntılı bir anahtarla 100 üzerinden puanlama yaptığı halde, ikinci sefer, birinci puanlamada en yüksek puan takdir ettiği üç öğrenciden ikisinin cevaplarını altıncı, birininkileri ise sekizinci olarak puanlamıştır. Birinci puanlamada «kalır» puan takdir ettiği en düşük puanlı iki öğrencinin cevaplarını ise, ikinci puanlamada «geçer» puanın oldukça üstünde puanlamıştır.

Daha sonraki araştırmalarda, bu derece kötü sonuçlar bulunmamıştır. Yazılı yoklama kâğıtlarının, birbiriyle karşılaştırılmaksızın, puanlayıcıda bıraktığı genel izlenime göre «iyi-orta-zayıf» gibi üç sınıfa ayrıldığı bir araştırmada, güvenilirlik katsayısı olarak hesaplanan korelasyon 0,67 bulunmuştur.⁶ Bu yöntemle puanlamada, puanlama kategorilerinin sayısı arttıkça puanlama güvenilirliğinin de arttığını gösteren kanıtlar vardır.⁷ Üniversite düzeyindeki bir sınavda Tarih, Yabancı Dil, Biyoloji ve Fizik derslerinin her birinde 1,5 - 2,0 saat süren yazılı yoklamaların kâğıtları, ayrıntılı bir anahtarla ve birbirinden bağımsız çalışan iki puanlayıcıya puanlattırılmıştır. Bu araş-

³R.L. Thorndike and E. Hagen, *Measurement and Evaluation in Psychology and Education*, 2 nd. ed., s. 43 - 47.

⁴Ö. Mart «Yazılı Yoklamalara Bel Bağlanabilir mi?», *Tedrisat*, Sayı 10, 1952.

⁵E.W. Tieges, «Educational Diagnosis», *Educational Bulletin*, No. 13, California Test Bureau, 1952.

⁶A.E. Myers, C.C. McConville and W.E. Coffman, «Simplex Structure in the Grading of Essay Tests», *Educational and Psychological Measurement*, 26 (1966): 41-54.

⁷W.E. Coffman, «Essay Examinations», in R.L. Thorndike (Ed.), *Educational Measurement*, 2nd ed., s. 295.

tırma, iki puanlayıcının puanları arasındaki korelasyonlar 0,70-0,80 arasında değişmiş; iki puanın ortalaması alındığında ise, 0,81 - 0,98 arasında değişen güvenilirlik katsayıları bulunmuştur.⁸ Bu araştırmalar gösteriyor ki, bir yazılı yoklama, oldukça objektif bir yöntemle puanlanmak kaydıyla, güvenilirliği yüksek puanlar verebilir.

Yazılı yoklamalarda güvenilirlik ve geçerliğin, soru yapısına da bağlı olduğu yukarıda belirtilmişti. Araştırmalarda da bu kanıyı doğrulayan sonuçlar bulunmuştur. Finlayson,⁹ cevabı sınırlandırılmamış sorularda puanlayıcılar arasındaki puan farklarının çok büyük olduğunu bildirmiştir. Bazı araştırmalar, yazma ve kompozisyon becerilerinde, dil yeteneklerinde üstün olan öğrencilerin cevaplarına, sınavın amacı bu beceri ve yetenekleri ölçmek olmadığı halde, gerçek değerlerinden yüksek puan takdir edildiğini göstermiştir.^{10,11, 12}

Coffman,¹³ yazılı yoklamalarla ölçülen başarı ile diğer başarı ölçüleri arasında ilişkiler arayarak, yazılı yoklamaların geçerliğini incelenmiş; sorular, müfredatı yeterince temsil ettiği ve puanlama objektifliği sağlandığı takdirde, yazılı yoklama geçerliğinin tatmin edici bir dereceye çıkarılabileceğini görmüştür.

YAZILI YOKLAMALAR NASIL YAPILMALI?

Yazılı yoklama, puanlaması güç, geçerliği harcanan zaman ve kulfete oranla düşük bir sınav türüdür. Böyle olmakla birlikte, eğitim ve öğretimde bu sınav türünden tamamen vazgeçmek mümkün olmaz. Diğer sınav türlerini iyi bilmeyen öğretmenler, yazılı yoklamaları hazırlama ve puanlama tekniklerini iyi bilmeseler bile, geleneksel olarak sınavlarında yazılı yoklama yöntemine fazla ağırlık verirler. Ayrıca, kompozisyon becerilerinin ve herhangi bir alanda yeni bilgi sentezleri yapma yeteneğinin ölçüleceği hallerde, yazılı yoklama yerine konulabilecek daha iyi bir teknik yoktur. Yazılı yok-

⁸F. Swineford, *Test Analysis, Advanced Placement Tests, Form EBP*.

⁹DS. Finlayson, «The Reliability of the Marking of Essays», *British Journal of Educational Psychology*, 21 (1951) : 126 - 134.

¹⁰J.A. Carroll, «A Factor Analysis of Verbal Ability», *Psychometrika*, 6 (1941) : 279 - 307.

¹¹J.R. Wittenborn, «Manual Ability, Its Nature and Measurements», *Educational and Psychological Measurement*, 5 (1945) : 241 - 280, 395 - 409.

¹²C.W. Taylor, «A Factor Study of Fluency in Writing», *Psychometrika*, 12 (1974) : 239 - 262.

¹³W.E. Coffman, «On the Validity of Essay Tests of Achievements», *Journal of Educational Measurement*, 3 (1966) : 151 - 156.

lamaları kullanma zorunlu olunca, gerek soru hazırlamada gerekse puanlamada belirli esaslara uyarak, bu yöntemin sakıncalarını mümkün olduğu kadar azaltmak akla yakın bir yoldur.

Yazılı yoklamalar başlıca iki amaçla kullanılır. Bu amaçlardan biri, herhangi bir derste bilişsel davranışları yoklama, diğeri ise kompozisyon becerilerini ölçmedir. Yazılı yoklamalar birinci amaçla kullanıldığında, yazı ve kompozisyon, aslında sınav için bir iletişim (fikirleri başkalarına ulaştırma) aracı olur. Bu halde, yazı güzelliği, kompozisyonda ustalık, dersin hedefleri dışında kalan ifade ustalıkları gibi değişkenler puanlamada dikkate alınmamalıdır. Bunların puana karışması, sınav kapsamıyla yoklanmak istenen bilgilerin ölçülmesinde geçerliği düşürür. Amaç anadilinde, bir yabancı dilde veya herhangi bir bilim ya da sanat alanında yazılı ifade becerilerini ölçmek olunca, yazı ve kompozisyon hem bir iletişim aracı hem de sınavın ölçmek istediği şey olur. Bu halde aslında ölçülmek istenen değişken kompozisyon becerileri olduğundan, ifade güzelliği, kompozisyon mükemmelliği ve orijinallliği, imlâ kurallarına uyma, vb. puanlanmalı; fakat, kapsamın bilgisel değeri puanlanmamalıdır.

Bu kesimde yazılı yoklamaların bilgi yoklama amacıyla nasıl kullanılacağı tartışılmaktadır. Yazılı yoklamaların kompozisyon becerilerini yoklama amacıyla kullanılışı ise daha sonraki kesimde ele alınacaktır.

Yazılı Yoklama Planı

Yazılı yoklamalar da, diğer sınavlarda olduğu gibi, planlama, soru yazma, soru seçimi ve çoğaltılması, uygulama, puanlama ve puanların kullanılması basamaklarından geçmelidir. İyi bir sınav, gelişigüzel veya sınama -yanılma yoluyla değil, iyi bir planlama ve ayrıntılar üzerinde durularak hazırlanabilir.

Bir sınav planında sınavın amacı, yani o sınavda elde edilecek puanların ne amaçla kullanılacağı, ne gibi ölçme araç ve yöntemlerinden yararlanılacağı, yoklanacak konular ve ölçülecek davranışlar bakımından sınavın kapsamı gibi ayrıntılar belirtilir. Bir yazılı yoklama da, iyi planlandığı ve plana uyularak hazırlanıp uygulanabildiği oranda iyi yapılmış olur.

Soruların Hazırlanması

Her sınav aracı, nihayet, bir takım sorulardan meydana gelir. Sınavdan elde edilecek puanların güvenilirliğine ve geçerliğine, ara-

cın kullanışlığına etki eden en önemli etken, soruların niteliğidir. Bir yazılı yoklama da ancak sorularının niteliği oranında iyi olabilir. Bir yazılı yoklamada kullanılacak soruların hazırlanmasında ve seçiminde aşağıdaki esaslara uyulmalıdır.

1. Sorular cevaplayıcıların tümü tarafından doğru olarak anlaşılabilmelidir. Açıkça anlaşılmayan sorular başka başka anlamlarda yorumlanır, bu sebeple farklı yönlerde cevaplandırılır. Bu nitelikteki cevapları puanlamak zor olur.

2. Sorulara verilecek cevaplar sınırlandırılmalıdır. Sınırlanmayan cevaplar hem uzunluk, hem de konunun ayrıntıları bakımından çeşitli olur. Ayrıca, iyi sınırlandırılmayan sorular başka konulara çekilmeye ve gereksizce şişirilmeye elverişli olur. Yazılı yoklamalarda cevapları sınırlandırmanın birkaç yolu olabilir. Bir yol, genel sorular yerine, aynı konuda daha özel ve kısaca cevaplandırılabilen sorular sormaktır. Şu örnekleri inceleyelim :

- I. Işığın kırılmasını anlatınız.
- II. Işık iki saydam ortamı ayıran yüzeyde hem kırılır, hem yansır. Bu olayı açıklayınız.
- III. Bitkilerde özümlemeyi anlatınız.
- IV. Yeşil yaprakların bitki beslenmesindeki rolü nedir?

Birinci soru cevaplayıcıyı başı boş bırakmaktadır. İkinci soru epeyce sınırlıdır, buna rağmen, doğru cevaplandırılabilmesi için «kırılma» olayının iyice bilinmesi gerekir. Hatta, kırılma olayını anlamış olmayı ortaya çıkarabilmesi bakımından ikinci sorunun daha geçerli olduğu söylenebilir. Benzer bir analizle, dördüncü sorunun üçüncü sorudan daha iyi olduğu sonucuna varılır.

3. Uzun ve az sayıda soru yerine kısa ve çok sayıda soru sorulmalıdır. Kısa sorularla, sınırlı bir sınav süresinde daha çeşitli bilgi ve yetenek alanları yoklanabileceği için, sınavın güvenilirliği ve geçerliği yükseltilebilir. Ayrıca, kısaca cevaplandırılabilen sorular, daha özel cevaplar gerektirir; böylece cevapların puanlanması kolaylaşır ve puanlama objektifliği artırılabilir.

4. Sorular ders kitaplarından ve diğer okuma kaynaklarından aynen alınmamalıdır. Hele öğrencinin elindeki standart kaynaklarda çözümü bulunan sorular, cevapları ders kitaplarının standart malzemesi haline gelmiş olan konular sınavda kullanılmamalıdır. Bu çeşit sorular çoğu halde hatırlamayı ölçer. Oysaki, hemen bütün derslerde öğretimin amacı, hatırlamadan ötede, bazı bilgilerin kavranması, yeni durumlara uygulanması, vb. gibi daha ileri basamaktaki

zihin yeteneklerini geliřtirmektir. Bu sebeple, ezberlenebilir cevapları yoklayan sorular yerine, bilgilerin yorumunu, yeni durumlara uygulanmasını, yeni sentezleri gerektiren sorular tercih edilmelidir. Öğretmenlerimizde, ders kitabının bir kesimini veya sınıfta anlatılan bir konuyu tekrar edebilen bir öğrenciyi başarılı sayma eğilimi vardır. Bilgilerin hatırdaki tutulması ve bu sebeple hemen unutulmamacak kadar iyi öğrenilmesi de önemli bir amaçtır. Fakat, bu öğrenme, öğrenmenin ilk basamağıdır. Bilgilerin kullanıldığı daha yüksek düzeydeki zihin yetenekleri daha ileri derecede öğrenme gerektirir.

5. Sorular birbirinden bağımsız olarak cevaplandırılabilir. Birinin cevaplandırılması ondan önceki sorunun doğru olarak çözülmesine bağı olan sorular, yazılı yoklamalar için kullanışlı değildir. Bu çeşit sorularda, önceki sorunun yanlış cevaplandırılması halinde, sonraki sorulardan doğru cevabı ona bağı olanların puanlanmasında güçlük çıkar. Bir soru şıklara ayrılırsa, şıklar arasındaki cevap bağımlılıkları da buna benzer puanlama güçlükleri doğurur.

6. Sınavda sorulacak toplam soru sayısından daha fazla sayıda soru hazırlanmalıdır. Böylece, hazırlanan sorulardan maksada en uygun, teknik özellikleri üstün olanları seçme imkânı doğar. Genellikle, sınavda kullanılacak soru sayısının üç katı kadar soru hazırlanırsa kolayca yeter sayıda soru seçilebilir.

7. İmkân bulunursa soruların hepsi sınavdan önce denenmelidir. Çoğu halde, soruların deneneceğı bir öğrenci grubu bulunmaz. Öğretmen aynı programı birkaç yıl üst üste uygulamışsa, önceki yıllarda verilen sınavları şimdiki yıl için «deneme» olarak değerlendirebilir. Böylece, soruların cevaplama süreleri daha doğru olarak kestirilebileceğı gibi, güçlük dereceleri de kabaca hesaplanabilir. Soruları deneme imkânı yoksa, bir veya iki meslekdaşın eleştirmeleğini almakta yarar vardır.

Sınavda Kullanılacak Soruların Seçimi

Bir sınavda kaç soru sorulacağı, sınav planında saptanır. Genellikle, bu sayıdan fazla soru hazırlanacağı için, eldeki sorulardan bir kısmının, bazı ölçütlere dayanılarak seçilmesi gerekir. Seçmede dikkate alınacak ölçütlerin en önemlileri şöyle sıralanabilir :

1. Soru seçimi hem konular hem de ölçülecek zihin davranışları bakımından, sınav planında gösterilen oranlara uygun olarak yapılmalıdır. Plan, bu açıdan ne kadar dengeli olursa olsun, soru

seiminde konu ve zihin davranışları oranlarına uyuımazsa, sınavın dengesi bozulur ve bu sebeple geçerliđi zedelenir.

2. Seilen soruların glk dereceleri, sınavın amacına uygun olmalıdır. Yazılı yoklamalarda ok sayıda soru sorulamaz. Soruların glk derecelerini nceden kesinlikle bilmek de mmkn deđildir. Bununla birlikte, ğretmenin kendisi veya soruları incelediđi meslekdařlar, btn soruları «ok g - g - orta - kolay - ok kolay» gibi beřli bir sınıflamaya tabi tutabilirler. Sorular, ğretmen kanısıyla bylece sınıflandırılmıř ve derecelendirilmıřse, soru glđ dađılımının simetrik olması ve yazılı yoklamaya dahil edilen btn soruların ortalama glđnn «orta» olması sađlanabilir. Sz geliři, yazılı yoklamaya 10 soru dahil edilecekse, bu sorular řu gk derecelerinde seilebilir : 1 soru ok g, 2 soru g, 4 soru orta, 2 soru kolay, 1 soru ok kolay.

3. Seilen sorularda ifade ve iml hataları bulunmamalıdır. Aynı konuda istenilen glkte birka soru varsa, diđer teknik zellikleri daha iyi olan soru tercih edilmelidir.

Soruların ođaltılması ve Aıklamalar

Yazılı yoklamalarda sınav sresinin nemli bir kısmı cevapları yazmaya harcandıđı iin ok sayıda soru sorulamaz. Bazı hallerde verilen sre btn soruların cevaplandırılmasına yetmeyebilir. Soruların, sınavın bařında đrencilere yazdırılması, memleketimizde ok kullanılan bir yol olmakla birlikte, zaten sınırlı olan cevaplama sresini daha da kısalttıđı iin sakıncalıdır. Bu sebeple, soruların đrencilere yazılı olarak verilmesi gerekir. Bunun iin de soru listesi-nin sınavdan nce ođaltılması gerekir.

Sorular ođaltılacaksa, listenin bařında bir «aıklama» veya «ynerge» bulunmalıdır. Bu aıklamada cevaplama sresi, her sorunun cevaplanması iin tahminen ne kadar zaman harcanacađı, cevapların sınırlılıđı, yazma řekli ve cevaplama ile ilgili diđer bilgiler bulunmalıdır. Gerekirse her soruya verilecek puan da aıklamada belirtilebilir.

Soruların hazırlanmasında gizliliđe titizlikle uyulması gerektiđi gibi, ođaltma sırasında da soruların đrenci eline gemesine karřı sıkı nlemler alınmalıdır. Bunun iin, msveddelerin iyice sayılması, ođaltma malzemelerinin sayılması, iře yaramayacak kopyaların yok edilmesi, kullanılacak kopyaların da numaralanmıř olarak emniyet altında bulundurulması gerekir.

Yazılı Yoklamaların Puanlanması

Yazılı yoklamaları puanlamada kullanılabilecek birkaç yöntem vardır. Burada, genel izlenimle puanlama, **sınıflama**, **sıralama** ve **anahtar** ile puanlama yöntemleri tartışılacaktır. Bu yöntemler, hem verdikleri puanın özellikleri bakımından hem de puanlamayı objektifleştirme dereceleri yönünden birbirinden farklıdır.

Genel İzlenimle Puanlama : Bu yöntemde puanlayıcı, bir cevap kâğıdını baştan sona okur; sonra, edindiği genel izlemine göre, bütün cevaplara bir toplam puan takdir eder. Ülkemizde çok kullanılan bu puanlama yönteminin birçok sakıncası vardır. Bununla birlikte, aşağıda verilen bazı önerilere uyulduğu takdirde, bu yöntemin sakıncalarından bir kısmı azaltılabilir.

Sınıflama : Bu yöntemde, ele alınan bir cevap kâğıdı tamamen okunduktan sonra «iyi - orta - fena» diye adlandırılan üç sınıftan birine konur. Böylece, bütün kâğıtlar teker teker okunup üç sınıfa ayrılır. Bırdan sonra her sınıfa düşen kâğıtlar tekrar okunarak, o sınıf içinde iki veya üç alt sınıfa ayrılır. Bu ikinci sınıflama işlemi de tamamlandıktan sonra, ilk sınıflamanın sınırlarına düşen alt sınıflardaki (meselâ, dokuz alt sınıf halinde üçüncü ile dördüncü alt sınıflardaki) kâğıtlar birbiriyle karşılaştırılır; gerekirse, bazı kâğıtların alt sınıfları değiştirilir. Bu işlem de tamamlandıca bütün kâğıtlar mükemmelliklerine göre sınıflanmış olur; fakat kâğıtlara sayısal puan takdir edilmiş olmaz. Sınıflar kendi aralarında sıralı olduğu için, sübjektif olarak her sınıfa bir sayısal puan verilebilir. (Yukarıdaki örnekte en fena kâğıtların bulunduğu alt sınıfa 1, sonra gelen sınıfa 2,..... ve en iyi kâğıtların bulunduğu alt sınıfa 9 puanı verilebilir.) Böyle yapılırsa, aynı alt sınıfa konmuş kâğıtların sayısal puanları eşit olur.

Sınıflamada her sınıfa düşecek kâğıt sayısının, önceden kabul edilmiş bir dağılıma göre saptanmasını öneren yöntemler de vardır. Örneğin, cevap kâğıtları beş sınıfa ayrılacaksa ve puanların yaklaşık olarak bir normal dağılım vermesi istenirse, sınıflama bittiğinde, sınıflardaki kâğıt yüzdeleri şöyle olmalıdır : En alt %5, alt % 25, orta % 40, üst % 25, en üst % 5. Sınıflara düşecek cevap kâğıtları sayısının önceden saptanması, ölçek nitelikleri oldukça farklı puanlar verir. Söz gelişi, yukarıda verilen sınıflamada «en alt» sınıfına «zayıf» anlamında bir puan takdir edilirse, sınıflanmış kâğıtlar arasında ne kadar zayıf kâğıt bulunursa bulunsun, bu puanlamada kâğıtların en kötü yüzde beşine «zayıf» puanı verilmiş olur. Şüphesiz, sınıflardaki cevap kâğıdı sayısının önceden saptandığı bir sınıflamadan elde edilen puanlar mutlak anlamda yorumlanamaz.

Sınıflama yöntemleriyle elde edilen puanlar, genel izlenimle takdir edilen puanlara oranla biraz daha objektif olabilir; fakat, anahtarla puanlamadaki kadar objektif olmaz. Sınıflama yöntemleri, hatalı olmakla birlikte, iyi bir puanlama anahtarının hazırlanamadığı hallerde kullanılabilir.

Sıralama : Okunacak kâğıt sayısının küçük olduğu hallerde, bütün kâğıtlar birbiriyle karşılaştırılarak en iyiden en kötüye doğru bir mükemmellik sırasına konabilir. Sıralamada da kâğıtlara sayısal ve mutlak anlamda bir puan verilmiş olmaz. Üstelik, sıralamada her cevap kâğıdı mutlak mükemmelliğine göre değil, diğer cevap kâğıtlarının mükemmelliklerine kıyasla ve onlara bağlı olarak yer alır. Bu sebeple, sıralama sonuçları, ancak **bağıl** bir anlam taşır. Sıralama sonuçları, hemen her zaman sıra sayılarıyla gösterilir ve «sıralama ölçeği» niteliklerindedir.

Sıralama ölçeğiyle yetinilmez de, sıralama sonuçlarına mutlak bir anlam kazandırmak istenilirse, sonuçlar sayısal puanlara çevrilebilir. Bunun için, sıranın başından, ortasından ve sonundan ikiye kâğıt seçilir. Bu kâğıtlara mutlak mükemmelliklerine göre sayısal puanlar verilir. Diğer kâğıtların puanları da bu altı kâğıdın sayısal puanları referans alınarak kestirilir. Bu yöntem, anahtarla puanlamada olduğu kadar mutlak sonuçlar vermese bile, sıra sayılarının bağıl anlamından oldukça farklı ve mutlak anlamda yorumlanabilir puanlar verir.

Anahtarla Puanlama : Yazılı yoklamada sorulan soruların cevapları kısa ve önemli fikirler halinde listelenir. Böyle bir liste hazırlandıktan sonra her cevaba ve cevapların alt bölümlerine puan takdir edilir. Böylece meydana getirilen puanlı kısa cevap listesine **puanlama anahtarı** denir. Puanlama anahtarının hazırlanması, aslında, bir soruya verilmesi beklenen cevapların iyice analiz edilmesini gerektiren bir işlemdir. Onun için, bu yöntemle puanlamaya, **analitik** puanlama da denir. Anahtarın hazırlanmasında dikkat edilecek esaslar şöyle sıralanabilir :

1. Cevaplara verilecek toplam puanın yuvarlak bir sayı olması, her soruya ve soruların alt bölümlerine verilecek ağırlıkların kolayca hesaplanmasına yardım eder. Toplam puanın büyüklüğü, en küçük ayrıntıya verilecek puan 1 olmak üzere seçilmelidir. Toplam puanın yeteri kadar büyük alınmaması halinde, ya ayrıntıları kaybetme ya da kesirli sayılarla uğraşma durumunda kalınır. Toplam puanın gereğinden büyük alınması halinde ise, anahtarda bulunmayan küçük farklar puanlamaya karışabilir. Puanlayıcının algılayamadığı

için puanlayamayacağı ayrıntıların anahtara girmesi puanlama güvenilirliğini düşürür.

2. Bir sınavdaki soruların bir kısmı diğerlerine oranla daha fazla veya daha önemli bilgileri yokluyorsa veya daha ileri yeteneklerle cevaplandırılabiliriyorsa, bu soruların daha fazla ağırlık alması puanların geçerliğini artırır. Onun için sorulara, kapsamaları ve programdaki önemleri oranında ağırlık verilmelidir. Fakat, sorulara farklı ağırlıklar verilmesini gerektirir önemli farklar yoksa, eşit ağırlıklarla çalışmak, yani her soruya aynı puanı vermek daha kolay olur.

3. Puanlama anahtarı, programı belli derslerde ve cevabı kesinlikle belli sorularda sınavdan önce hazırlanmalıdır. Hatta, anahtar hazırlamakta güçlük görülürse, sınav sorularının düzeltilmesine bile gidilebilir. Cevapların çeşitli olacağı sanılıyor ve cevaplayıcıların bilgi düzeyleri bilinmiyorsa, puanlama anahtarı, birkaç kâğıt okunduktan sonra hazırlanabilir. Bunun için kâğıtlardan bazıları kabaca gözden geçirilmeli, «iyi», «orta» ve «fena» sayılabilecek üç kâğıt iyice incelenerek anahtar hazırlanmalıdır.

4. Cevap kâğıtları anahtarla puanlanırken, bir öğrencinin cevapları arasında orijinal ve anahtarda bulunmayan önemli bir fikre rastlanabilir. Böyle fikirlerin puana dahil edilmemesi sakıncalıdır. Böyle hallerde anahtar tekrar gözden geçirilmeli, önemli ayrıntılar anahtara eklenmelidir.

5. Bazen bir öğrencinin, istenilen bilgiyi değil, onun tam tersini yazdığı; bazen cevapların kasıtlı olarak diğer konulara çekildiği veya şişirildiği; bazen de öğrencinin bir konuda bilgi göstermeye çalışırken diğer konuda bilgisizliğini ortaya koyduğu görülür. Böyle hallerde puan kırılmalı mıdır? Bu soru puanlama anahtarında **negatif puan** düşüncesini doğurmuştur. Bir sorunun ayrıntılarına alt puanlar verilmiş ise, o soruya verilen toplam puanın bir kısmı negatif puanlarla kırılabilir. Fakat, bir sorunun cevabı şişirildiği veya tamamen yanlış yazıldığı için, ondan bağımsız olarak cevaplandırılabilen diğer bir sorunun puanı kırılmamalıdır. Bu ise, bir soru için düşünülebilecek negatif puanların büyüklüğünü sınırlar. Puanlama anahtarında, bir sorunun ne tür cevaplarına negatif puan verileceği dikkatle saptanmalıdır.

Anahtarla puanlama için, bir cevap kâğıdında bir soruya verilen cevap okunur. Anahtarda gösterilen fikirlerin tümü bu cevapta varsa, o sorunun tam puanı o cevaba verilir. Anahtar hazırlanırken tam olmayan cevaplara belirli bir puan verilmesi kabul edilmemişse, puanlama sırasında kısmen doğru cevaplara puan verilmez.

Anahtarla puanlama metodu yazılı yoklamaların sakıncalarından birçoklarını önler. Çok iyi hazırlanmış bir anahtar, puanlayıcının işini kolaylaştırır, puanlama işlemini çabuklaştırır ve puanları oldukça objektif hale getirir. Tablo 4.1'de bir soru ve puanlama anahtarı örnek olarak verilmiştir.

TABLO 4.1
PUANLAMA ANAHTARI

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

Anahtar	Puan
Yararlar :	
a) Işık enerjisini soğurma yüzeyi artar (genişlik)	2
b) Işık enerjisi içeriye girebilir (yassılık)	2
c) CO ₂ alıp O ₂ vermeyi kolaylaştırır	2
d) Su buharlaşması kolay olur (Suyun gereğinden fazla olduğu hal)	
Güçlükler :	
e) Su kaybı çok olur (Suyun kıt olduğu hal)	2
Toplam	10
Negatif puan :	
(c) deki bilgiyi tersine yazmak	—1
(d) ile (e) yi ayıramamak	—1

Bazen, yazılı yoklamaların birden fazla puanlayıcı tarafından puanlanması gerekebilir. Böyle hallerde, genellikle, farklı puanlayıcılar tarafından verilen puanlar arasındaki tutarsızlık, genel izlenimle puanlamada, sıralama ve sınıflamadakinden büyük olur. Buna karşılık, aynı anahtarla çalışan bağımsız puanlayıcılar arasında tutarsızlıklar daha az olur. Bunun sebebi, anahtar hazırlanırken bir tür ortak ölçüt kabul edilmiş olmasıdır.

Puanlama anahtarlarının diğer bir yararı, sınav sonuçlarına bakılarak tamamlayıcı öğretim faaliyetlerinin planlanacağı hallerde görülür. Puanlama anahtarındaki analiz, hangi konuların veya bir konuda hangi kısımların iyi öğrenilmediğini gösterir.

Puanlama anahtarına ve analitik puanlamaya karşı çıkanlar, ayrıntılara bakıldığı için cevaplarda bütünlüğün değerlendirilemediği ve değersiz ayrıntılara, eksik cevaplara puan verildiği iddialarını ileriye sürerler. Bu iddialar, puanlama anahtarının kullanılmasını engellemez. Çünkü, gerekli hallerde, ancak tam cevaplara puan verilecek bir anahtar hazırlanabileceği gibi, cevabın bütünlüğüne, an-

latım tarzına, cevabın orijinalliğine de ayrı puan verecek bir anahtar hazırlanabilir.

Puanlama Hatalarını Azaltan Diğer Önlemler

Aşağıda sıralanan önlemler yazılı yoklamaların puanlanmasında çok rastlanan bazı hataları azaltır veya tamamen önler.

1. Puanlayıcı, okuduğu kâğıdın hangi öğrenciye ait olduğunu bilmemelidir. Böylelikle, puanlayıcının o öğrenci hakkındaki ön kanılarının, öğrenci yararına ya da zararına, puanı etkilemesi önlenir.

2. Puanlayıcı, yazı güzelliğinin veya okunaksızlığının, ifade jüzgünlüğünün ve benzeri yanıltıcı etkilerin hatalara sebep olduğunu bilmelidir. Bu tür hatalara karşı uyanık olmak ve bunlardan bilinçli olarak sakınmak hataları azaltır.

3. Özellikle, anahtarla puanlama yapılmayan hallerde, «iyi» bir kâğıttan sonra okunan «orta» bir kâğıt «zayıf» görünür. Bu etkinin tersi de olabilir. Çok kötü bir kâğıttan sonra okunan orta kâğıt, gerçek değerinden iyi görünür. Bu hatayı önlemek için, kâğıtların, değişik sıra ile iki defa okunması tavsiye edilir. Şüphesiz, anahtarla puanlama halinde bu hata çok azalır. Puanlayıcının bu etkiyi bilmesi, böyle bir hatadan bilinçli olarak kaçınması da hatayı azaltır.

4. Yukarıdaki hata gibi, aynı kâğıtta iyi yazılmış bir cevaptan sonraki cevabın gereğinden yüksek puanlanmasına ve kötü yazılmış bir cevaptan sonraki cevabın da gereğinden düşük puanlanmasına sebep olan bir hata vardır. Bu hata bir öğrencinin ilk soruya verdiği cevabın puanlayıcı üzerindeki etkisinin, diğer cevaplara geçmesinden doğar. Bu hatanın önlenmesi için, bütün kâğıtlarda birinci soruya verilen cevaplar puanlanmalı, sonra kâğıtların sırası karıştırılarak ikinci sorunun cevapları puanlanmalı ve bu işlem böylece devam ettirilmelidir.

5. Birden fazla puanlayıcı çalıştırmak mümkün olursa, her puanlayıcı diğerlerinden bağımsız olarak bütün kâğıtları puanlamalı, sonra bir kâğıda verilen puanların ortalaması alınmalıdır. En az iki, en çok beş kişi puanlayıcı olarak görevlendirilebilir. Puanlayıcı sayısının daha fazla artırılması puanların güvenilirliğini önemli derecede artırmaz. Puanlanacak cevap kâğıtlarının çok fazla olduğu hallerde bütün puanlayıcılar ortak bir anahtar hazırlamalı, her kâğıt iki puanlayıcı tarafından okunmalı ve puanlanmalıdır. İki'den fazla puanlayıcı çalıştırılırsa, bu yöntemle birçok kâğıt okunabilir. Puanlayıcılar ikişer ikişer de gruplanabilir. Bu yöntem, puanlama grup-

ları arasındaki farkları önleyemezse de, bir tek puanlayıcı tarafından verilecek puanlara göre daha güvenilir sonuçlar almaya yardım eder.

KOMPOZİSYON BECERİLERİNİN YAZILI YOKLAMAYLA ÖLÇÜLMESİ

Herhangi bir alanda yeni bir sentez meydana getirme yeteneğinin veya kompozisyon becerilerinin ölçülmesinde, yazılı yoklamanın en etkili yöntem olduğu yukarıda belirtildi. Yazılı yoklama, ancak tekniğine uyularak hazırlanıp puanlandığı takdirde bu amaçlar için etkili olabilir. Yazılı yoklamanın bu iki amaçla kullanılması, yukarıdaki kesimlerde anlatılan yöntemlerden biraz farklı bir yöntemle uygulanmasını gerektirir.

Yeni bir sentez meydana getirme yeteneği ölçülecekse, yazma konusu, alışılmamış, cevabı standart eserlerde kolayca bulunamayacak bir soru olmalıdır. Örneğin, bir psikoloji programında çeşitli öğrenme kuramlarına yer verilmiş olsun ve ders kitabında bu kuramlar teker teker incelensin. Öğrenciye «Çeşitli öğrenme kuramları, unutmaya olayını nasıl açıklar?» gibi bir soru sorulursa, kuramları iyi öğrenmiş bir öğrenci, ders kitabında bu soruyu doğrudan doğruya cevaplandırarak bir kesim bulunmadığı halde, o kuramlardan cevabı çıkarabilir ve o soruya cevap niteliğinde bir kompozisyon meydana getirebilir. Böyle bir sınavda hem cevabın doğruluğuna, hem de kompozisyonun mükemmelliğine puan takdir etmek gerekir.

Ana dilinde kompozisyon becerileri ölçülecekse, kompozisyon konusu, öğrencinin yazmasına yetecek derecede bildiği bir konu veya anlatmaya değer bir tecrübe olmalıdır. Aksi halde, iyi yazamamış bir öğrencinin, yazma becerilerindeki yetersizliğinden mi, yoksa kompozisyon konusunu bilmediğinden mi iyi yazamadığı kestirilemez. Çok çeşitli materyal, kompozisyon konusu olarak kullanılmaya elverişli olabilir. Yaşanmış olaylar, günlük olaylar, genel kültür sayılabilecek bilgiler ve benzerleri, sık kullanılan kompozisyon temalarıdır. Seçilen konu, öğrencinin yaşına, geçmiş yaşantılarına ve eğitim düzeyine uygun olmalıdır.

Kompozisyonlarda sadece konu veya tema vermekle yetinilmeli; özellikle yukarı eğitim basamaklarında, meydana getirilecek kompozisyon için farazi bir maksat da verilmelidir. Böylece, yazara daha gerçekçi bir durum verilmiş olur. Bu açıdan bakılınca, lise düzeyindeki öğrenciler için, aşağıdaki örneklerden ikincisi daha üstündür.

1. «Ormanların canlı tabiatın devamını sağlamadaki rolü» konusunda yaklaşık 500 kelimelik bir kompozisyon yazınız.

2. Büyük halk kitleleri tarafından okunan bir gazete, sizden «Ormanların canlı tabiatın devamındaki rolü» konusunda 500 kelimelik bir yazı istemektedir. Böyle bir yazıyı, önemli değişiklikler ve düzeltmeler gerektirmeyecek mükemmellikte yazınız.

Kompozisyon yazdırılırken yazı miktarı kelime veya standart sayfa olarak sınırlandırılmalıdır. Yazma süresi de tasarlamaya ve yazmaya yetecek kadar bol verilmelidir.

Kompozisyonlar, yukarıda sözü edilen genel izlenim, sıralama ve sınıflama yöntemleriyle puanlanabilir. Fakat, bu yöntemler kullanıldığında, kompozisyon puanlarının güvenilirliği, genellikle, aynı yöntemlerle puanlanan diğer yazılı yoklamaların güvenilirliğinden daha düşük bulunur. Buna sebep, puanlayıcıların bir kompozisyonda bulunması gerekli niteliklerin neler olduğunda anlaşılamamaları ve puanlamanın sübjektif ölçütlerle yapılmasıdır. Bu sakıncanın azaltılması için, kompozisyonların da analitik anahtarlarla puanlanması düşünülmüştür. Analitik anahtar, çoğu halde, puanlayıcılar arası uyumu ve bu nedenle puanların güvenilirliğini artırır. Türkçe kompozisyon için örnek olarak hazırlanmış bir anahtar Tablo 4.2'de gösterilmiştir. Bu örnek, büyük kısmıyla analitik olmakla birlikte, genel izlenimle puanlamaya da yer bırakmıştır. Üslûba ve yeniliğe verilecek 20 puanın takdiri için, kompozisyonun bir bütün olarak genel izlenimle de puanlanması gerekir.

DİĞER YAZILI YOKLAMA ÇEŞİTLERİ

Yazılı yoklamaların, bu bölümde esasları verilen ana yöntemden, gerek hazırlanışta gerekse uygulanışta ayrılan çeşitleri vardır. Aşağıdaki paragraflarda bu yöntemlerden üçüne kısaca değinilmektedir.

Yazılı yoklamalarda gereğinden fazla sayıda soru verip, cevaplayıcılardan, tercih ettikleri belirli sayıda sorunun cevaplandırılmasını isteyen sınav çeşidi de kullanılır. **Tercihli** veya **seçmeli** yazılı yoklama çeşidi denilebilecek olan bu sınavda, meselâ on soru verilip, cevaplayıcılardan, bunların beş tanesini seçerek cevaplandırmaları istenebilir. Soruları birkaç grupta toplayıp her gruptan birer sorunun seçilmesini isteyen sınav çeşidi ile bir gruptaki soruların tümünü zorunlu tutan ve diğer gruptakilerden belirli sayıda soru seçtiren sınav çeşitleri de tercihli yazılı yoklamalardandır.

Tercihli yazılı yoklamaların kullanılabileceği haller vardır. Cevaplayıcıların sınav konularındaki geçmiş yaşantıları bakımından birbirinden çok farklı olduğu hallerde, konuları bakımından tercihli sorular sorulabilir. Böyle bir durumda, güçlük dereceleri ve geçer-

likleri eşdeğer sorular hazırlama zorunluğu doğar. Bu niteliklerde sorular hazırlamak, hele soruların güçlük ve geçerlik bakımından birbirine eşdeğer olduğunu kanıtlamak kolay olmaz. Okuldaki başarının ölçülmesinde tercihli sorulara, ancak, o dersin programında tercihli veya seçimlik konular varsa gerek duyulabilir. Programın tek tip işlendiği hallerde, tercihli sınavlar kullanılmaması daha doğru olur. Tercihli konuların bulunduğu programlarda bile tercihli sı-

TABLO 4.2
KOMPOZİSYON PUANLAMA ANAHTARI

Anahtar	Puan
I. Başlık	
a) Varlığı	1
b) Konuyla ilişkisi	2
c) Kısıtlılığı, anlatıcılığı, etkileyiciliği	3
II. Anlatım Düzeni	
a) Giriş :	
a 1) Varlığı	1
a 2) Konuyla ilişkisi	3
a 3) Sunuş açıklığı	3
a 4) Etkileyiciliği	3
b) Geliştirme :	
b 1) Ana fikrin varlığı	2
b 2) Ana fikrin belirginliği	3
b 3) Ana fikre ulaşmadaki düşünce zincirinin sağlamlığı, tutarlılığı	5
b 4) Yardımcı fikirlerin varlığı	2
b 5) Yardımcı fikirlerin ana fikri destekleme ve tutarlılık dereceleri	5
b 6) Anlatımın açıklığı ve akıcılığı	5
c) Sonuç :	
c 1) Varlığı	2
c 2) Ana fikirle tutarlılığı	3
c 3) Yan fikirlerle desteklenişi	2
c 4) Kısıtlılığı, açıklığı, kesinliği	5
III. İfade zenginliği	
a) Kullanılan kelimelerin çeşitliliği	3
b) Kelimelerin yerindeliği	6
c) Cümle yapısının maksada ve okuyucuya uygunluğu ...	3
IV. Yazım kurallarına uygunluk	
a) İmlâ	3
b) Söz dizimi (gramer)	5
c) Noktalama	3
V. Bir bütün olarak kompozisyon	
a) Üslûp	10
b) Yenilik, kendine özgülük	10
Toplam puan	100

nav sorularından kaçınıp, seçimlik konulardaki öğrenci farklılıklarını ödev veya projelerle değerlendirmekte yarar vardır.

Tercihli sorular kullanıldığında, sınavdaki soru sayısının ve bu sebeple örneklemin temsil gücünün artacağı sanılırsa da, bu doğru değildir. On sorudan herhangi beş tanesinin cevaplandırıldığı bir sınavda, on soruluk bir örneklem değil, ancak herbiri beşer soruluk çeşitli örneklemler yoklanmış olur. Üstelik, bu halde cevaplayıcılar birbirinden farklı soru gruplarını cevaplandırmış olacaklarından, onları birbiriyle karşılaştırmak güçleşir. Böyle bir sınavdan elde edilecek puan sırası, diğer şartlar eşit olsa bile, tek tip sorularla elde edilecek bir puan sırası kadar güvenilir olmaz.

Öğrencilere belli bir soru vermek yerine, «En çok bildiğiniz bir konuyu yazınız.» gibi bir yönergeyle yazılı yoklama yapan öğretmenler de görülmüştür. Ders dışında yapılacak bir projenin ve benzerlerinin öğrencilerce kendi ilgi duydukları alanlardan seçilmesinde yarar görülebilir. Fakat, çelişkili bir deyimle **sorusuz imtihan** adı verilebilecek olan böyle bir yöntemin, yazılı yoklama olarak savunulabilecek tarafı yoktur.

Yazılı yoklama sorularını **ad çekme** ile ve her öğrenciye ayrı takımlar halinde veren öğretmenler de görülmüştür. Bu yöntem de, birçok sakıncalar doğurabileceği için savunulabilecek bir sınav çeşidi değildir.

Sonuç olarak denilebilir ki, ana yöntemden ayrılan yazılı yoklama çeşitleri, kendilerine özgü sakıncalarından dolayı daha az kullanışlıdır ve geçerliği daha düşük sonuçlar verir.

B Ö L Ü M 5

Kısa Cevap Gerektiren Testler

Cevaplayıcının, bir kelime, bir rakam, bir ibare veya en çok bir cümle ile cevaplayabileceği maddelerden meydana gelmiş test tipine, kısa cevaplı test denir.¹ Bu test çeşidi ayrıntılı cevap veya kompozisyon gerektiren yazılı yoklamalardan cevapların kısalığı ile, seçme veya sınıflama gerektiren testlerden ise, cevapların verilenler arasından seçilmeyip cevaplayıcı tarafından bulunması ve yazılması özelliği ile ayrılır.

Kısa cevaplı test maddeleri yapı bakımından çok çeşitli olabilir. Bu bölümde yalnız iki çeşit madde yapısı üzerinde durulacaktır. Örnekler :

1. Işığın rengini hangi özellik tayin eder? (Dalga boyu)
2. 1961 Anayasasına göre, Cumhurbaşkanı tarafından seçilir. (Büyük Millet Meclisi)

Bu örneklerde görülen maddelerden birinciye soru, ikinciye eksik cümle yapısında madde denir.

KISA CEVAP TESTLERİNİN BAŞLICA ÖZELİKLERİ

Bir sınavın niteliklerini tayin eden önemli özelliklerden biri soruların yapısıdır. Kısa cevap testlerinin ayırtedici özellikleri soruların yapısından doğar. Bu özellikler aşağıda sıralanmış ve kısaca açıklanmıştır.

1. Cevaplayıcı, cevapları düşünüp bulmak ve yazmak zorundadır. Bu yapıdaki test maddeleri cevaplayıcının sorulan bilgiyi hatırlama veya bulma yeteneğini yoklar. Sadece bilgi yoklayan, bilginin yorumunu veya uygulamasını istemeyen seçmeli testler dikkate alınır, kısa cevap yapısındaki maddelerin seçmeli testlerle yoklanan yeteneklerden biraz daha farklı bir yetenek yokladığı söylene-

¹Bir testin kendi başına puanlanabilen en küçük birimine «test maddesi» denir.

bilir. Bir bilgiyi hatırlamak, onu görünce tanımaktan veya yanlışlardan ayırtıtmekten oldukça farklı bir yetenektir.

2. Cevaplayıcı, istediği cevabı verme bağımsızlığına sahiptir. Bu özeliği ile kısa cevap testleri, yazılı yoklamalara yaklaşıp. Kısa olmakla birlikte her cevap, nihayet, cevaplayıcının düşünebildiği ve doğru diye verdiği cevaptır. Kısa cevap maddeleriye de, kompozisyon gerektiren sorularda olduğu gibi, istenilenden ve birbirinden çok farklı cevaplar alınabilir. Özellikle sorular açık değil ve cevaplayıcının bilgi düzeyi düşük ise böyle cevaplar çok olur.

3. Cevaplar kısadır. Bu özellikten dolayı kısa cevap testleri seçmeli testlere yaklaşıp. Cevapların cevaplayıcı tarafından hatırlanıp yazılması gerekmele birlikte, cevabı kompoze etme külfeti hemen hemen önemsiz dereceye iner. Kısa cevap testlerinde, cevaplama işlemi cevabı bilen bir cevaplayıcının çok zamanını almayacağı için, belli bir sınav süresinde çok sayıda soru sorulabilir ve daha geniş bilgi alanını yoklamak kabil olur. Böylece programın kapsamını daha iyi temsil edebilen bir örnekleme yapılabilceğinden, aynı sürede uygulanabilecek bir yazılı yoklamaya kıyasla, daha güvenilir ve belki de geçerliği daha yüksek bir sınav yapılabilir.

4. Kısa cevap gerektiren maddeler çok çeşitli bilgi ve yetenek alanlarına uygulanabilir. İlk bakışta, kısa cevap testleriyle sadece hatırlamaya dayanan basit bilgiler yoklanabilecek gibi görünürse de, madde yazmada tecrübeli ve alanlarında bilgili kimseler, daha derin bilgileri ve daha yukarı basamaktan zihin yeteneklerini yoklayabilecek maddeler yazabilirler.

5. Kısa cevap testleri her eğitim düzeyine uygun düşer. Cevaplayıcıdan basit bir cevaplama ve kolay bir cevap kaydı istendiği için, bu yapıdaki testler aşağı eğitim basamaklarında kolaylıkla uygulanabilir. Şüphesiz, cevaplamanın basitliği bu yapıdaki maddelerin yukarı eğitim basamaklarında uygulanmasına engel değildir.

6. Puanlama işlemi nisbeten kolaydır, fakat tamamen objektif değildir. Puanlamanın kolaylığı cevapların, kompozisyon gerektiren sorulara kıyasla, sınırlı, kısa ve kesin oluşundan ileri gelir. Fakat bu testlerde cevaplayıcıların istedikleri cevabı verme bağımsızlığı vardır. Cevapların kısmen doğru olması, doğruluk derecesinin kestirilememesi de mümkündür. Puanlayıcının, kısa da olsa, her cevabı okuması, anahtara göre doğru olup olmadığını kararlaştırması gerekir. Onun için, puanlama işlemi bir makina ile yapılacak kadar mekanikselleştirilemez.

7. Kısa cevap testlerinin hazırlanması, diğer yapıdaki testlere kıyasla daha kolaydır. Kısa cevap gerektiren test maddelerinin yazılması bilgi ve tecrübe gerektirmekle birlikte, bu tekniğin esasları kolayca öğrenilebilir; gerekli tecrübe kısa sürede kazanılabilir.

KISA CEVAP MADDELERİNİN YAZILMASI

Soru ve eksik cümle yapısındaki kısa cevap maddeleri, yazma tekniği bakımından ayrılık gösterir. Aşağıda önce her iki çeşit maddeye uygulanabilecek esaslar sıralanmıştır.

1. Her madde, kesinlikle cevaplandırılacak bir açıklıkla ifade edilmelidir. Maddenin ifadesi açık olmazsa, her cevaplayıcı onu başka bir anlamda yorumlar; bu sebeple, beklenilenden farklı ve çok çeşitli cevaplar alınır. Eksik bir cümle ile bir soru sormak kolay olmadığı için, tecrübesiz madde yazarları, özeltikle eksik cümle yapısındaki maddelerde belirsiz ifadeler kullanırlar. Örnek :

3. Sıtma mikrobi

Bu maddenin yazarı «plazmodidir» cevabını beklemiştir. Fakat «zararlıdır», «kanda yaşar», «sivrisinekle geçer» gibi cevapların da doğru olabileceği açıkça bellidir. Maddede verilen kelimelerden ne sorulduğu anlaşılmamaktadır. Maddenin kusuru şöyle düzeltilebilir:

3a. Sıtma mikrobuna adı verilir.

Böylece madde kesinleşmiştir. Eksik cümle yapısından vazgeçilerek soru kipi kullanılırsa madde daha kolay anlaşılır hale gelir.

3b. Sıtma mikrobuna ne ad verilir?

Maddelerin belirsizliği bazen istenilen cevabın sınırlandırılmayışından, bazen de sorulan sorunun kesin bir cevabı bulunmamasından ileri gelebilir.

Örnekler :

4. Elektrik enerjisinin sağladığı fayda nedir?

5. Küçük hayvanları niçin severiz?

Bu iki maddeye de birçok doğru cevap verilebilir. Alınan cevaplar cevaplayıcının bu soruları yorumlayışlarına göre farklılık gösterir. Kesin bir doğru cevabı olmayan sorular, kısa cevaplı testler için iyi madde olamaz.

2. Her madde önemli bir bilgiyi yoklamalıdır. Bir sınav hazırlanırken madde yazma kolaylığı ön plana alınacak yerde, yoklanması istenen bilgi ve becerilerin toplamı gözönünde bulundurulmalı, bu

toplamdan en önemlileri seçilmelidir. Çoğu zaman öğretmen önemli bir bilgiyi yoklamak ister, fakat soruyu yazmadaki yetersizliği buna engel olur. Örnekler :

6. Kaldıraçlar kaç sınıfa ayrılır?
7. Kaldıraçların çeşitlerini yazınız.
8. Kaldıraçlar hangi özelliklerine göre sınıflanır?
9. Kaldıraçlar hangi özelliklerden dolayı faydalı olur?

Bu sorulardan birincisine, kaldıraç çeşitlerinin adlarını veya özelliklerini belirtmeden «üçü ayrılır» diye bir cevap verilse, bu cevap doğru olur. Fakat bu soru ile önemli bir bilgi yoklanmış olmaz. Böyle bir cevabı veren cevaplayıcının kaldıraçları ne derecede bildiği anlaşılmaz. Diğer sorular gittikçe derinleşen bilgileri yoklamaktadır. Kaldıraçların kuvvet ve karşı kuvvet kollarının sabit noktaya göre konumlarının sınıflamaya esas olduğunu yazan bir öğrenci, kaldıraç çeşitlerini açıklamaksızın sayan bir öğrenciden daha ileri düzeyde bir kavrayışa ulaşmıştır. Kaldıraçların «kuvvetleri belli oranlarda büyütme veya küçültme» özeliğinden dolayı yararlı olduğunu yazan bir öğrenci de, bu araçların kullanıldığı yerleri sayıp döken bir öğrenciden daha derin bir kavrayışa sahiptir.

Hangi düzeyde bir bilgi veya zihin yeteneği yoklanacaksa, test maddesi de ancak o düzeyde bir yetenekle cevaplandırılabilir nitelikte yazılmalıdır.

3. Her madde yalnız bir bilgiyi yoklamalıdır. Aynı soru içinde birden fazla bilgiyi yoklayan maddelerin puanlanması zordur. Bu halde, istenilen bütün bilgileri yazamayan bir cevaplayıcıya tam puanın bir kısmını vermek, belki de eksik cevapları yanlış saymak gerekecektir. Ayrıca, istenilen cevabın ne derece ayrıntılı olacağını kestiremeyen cevaplayıcılar eksik veya fazla yazabilir. Özellikle soru kipindeki maddelerde birden fazla bilgi sorma hatasına düşülebilir. Tecrübesiz yazarlar, soru kipinde madde yazarken bu hataya karşı dikkatli olmalıdırlar. Bir sorudan birden fazla ayrıntı sorulması gerekli ise, maddede istenilen ayrıntıları saydırmak doğru olur. Örnek :

11. Bir katı maddenin suda çözünürlüğü nelere bağlıdır?
12. Bitkiler hangi olaylarda oksijen harcar?
 - a)
 - b)
13. Nazi Almanyası İkinci Dünya Savaşına hangi sebeplerle girmiştir?

Maddeler böyle ifade edilince ayrıntılarının ayrı ayrı puanlanması gerekir. Madde yazılırken ayrıntı sayısının cevap yeri bırak-

mak veya cevap yerlerini numaralamak, cevaplayıcıya gereksiz ipucu verir. 12 numaralı madde bu bakımdan kusurludur.

4. Kısa cevap maddelerinin yazılmasında şekil, harita ve benzeri gibi grafik materyallerden yararlanılabilir. Bir grafiği kullanabilme, cetvel ve grafiklerden yorum ve yordama yapma yeteneklerinin ölçülmesinde bunlardan faydalanma zorunuğu vardır. Grafik materyaller çoğu halde testin yazılmasını kısaltır. Bir grafik, bir şekil veya cetvel birkaç maddenin ortak kaynağı olarak verilebilir. Bir grafik veya tabloya dayanan birkaç soru çıkarılabilir. Fakat aynı grafikten gereğinden fazla soru sorulması, testin kapsamına girecek konuların dengesini bozabilir. Madde yazarı, grafik materyalleri kullanırken bu sakıncaya karşı uyanık olmak zorundadır.

Eksik cümle yapısındaki maddelerin yazılmasında, yukarıdakilere ek olarak, şu noktalara da dikkat edilmelidir :

1. Maddeyi oluşturan cümle, cevaplayıcının tanıdığı bir kaynaktan aynen alınmamalıdır. Madde daha önce öğrenilmiş veya ezberlenmiş bir cümleden ibaret olursa, bilgi yerine basit bir hatırlama ölçme tehlikesi vardır. Ana fikri yoklanacak olan bir cümle, eksik cümle yapısında bir madde olarak yeniden yazılmalıdır. Madde yeniden yazılırken kalıplaşmış tanımlar, tekerleme haline gelmiş ifadeler kullanılmamalıdır. Bir cevaplayıcının bir bilgiyi değişik bir ifade içinde gördüğü zaman tanıyabilmesi, alışılmış bir ifade halinde hatırlayabilmesinden daha ileri bir öğrenmenin kanıtıdır.

2. Eksik cümlede, geri kalan kelimeler ne sorulduğunu anlatabilecek yeterlikte olmalıdır. İçinden bir kaç kelimesi çıkarılmış bir cümle, soru niteliğinde olmadığı gibi, ne söylendiğini açıkça belirten bir ifade de olamaz. Ne söylendiği ancak geri kalan kelimelerden sözün gelişle çıkarılabilir. Aslında, soru veya emir kipinde olmayan bir cümle ile soru sormak güçtür. Anlaşılması imkânsız veya çok güç bir madde yazılmasını önlemek için, bir cümleden bir-iki kelimeyi fazla çıkarılmaması tavsiye edilir.² Ayrıca bir cümleden çıkarılan kelime, o cümlemin sonuna yakın olursa ne sorulduğu daha kolay anlaşılır.³ Örnek :

14. 1912 savaşına ve katıldı. Osmanlı Devleti bu savaştan çıktı.

²J.C. Stanley, *Measurement In Today's Schools*, 4th ed., s. 211.

³R.L. Thorndike and E. Hagen, *Measurement and Evaluation In Psychology and Education*, 2nd, ed., s.70.

Bu maddede Balkan savaşıyla ilgili bilgiler sorulmaktadır. Ortadaki üç başluğa «Bulgaristan, Yunanistan - Karadağ» yazılması istenmektedir. Bu savaş, aynı tarihlere raslayan Trablusgarp savaşıyla karıştırılabilir. Balkan devletlerinden biri yerine Osmanlı Devleti veya Türkiye de yazılabilir. Son başluğa «yenik», «yorgun», «kayıplı» gibi kelimelerden biri uygun düşebilir. Aslında bu madde ile üç fikir sorulmak istenmektedir. Bunlar şöyle sorulabilir :

- 14a. Yunanistan ve Osmanlı Devleti arasındaki 1912 savaşına ne ad verilir?
- 14b. Yunanistan ve Osmanlı Devleti arasındaki 1912 savaşına başka hangi devletler katılmıştır?
- 14c. Yunanistan ve Osmanlı Devleti arasındaki 1912 savaşında Osmanlı Devletinin kayıpları neler olmuştur?

Bu örneklerde de görüldüğü gibi, soru kipi eksik cümle yapısındaki maddelerden daha kolayca kullanılabilir.

3. Maddede, istenilen yeteneğe sahip olmayan bir cevaplayıcının doğru cevabı bulabilmesine yardımcı ipuçları bulunmamalıdır. Cevap yerlerinin uzunluğu ve düzeni, cevaptaki kelime sayısının belirtilmesi, cevaptaki kelimenin baş harfinin belirtilmesi maddeyi kolaylaştırır ve gereksiz ipuçları verir.

TESTİN DÜZENLENMESİ

Testin başında bir yönerge (açıklama) bulunması zorunludur. Bu yönergede cevaplama süresi, cevaplama sırasında öğrencilerin sorularının cevaplandırılmayacağı, puanlama yöntemi, vb. gibi genel açıklamaların yanında, cevaplamanın nasıl yapılacağı, cevapların nasıl kaydedileceği de açıkça belirtilmelidir.

Her testte, maddelerin, kullanışlı bir sayfa düzeninde verilmesi gerekir. Bu düzenlemede okuma ve algılama kolaylığı, cevapları kaydetme kolaylığı, puanlama kolaylığı, okuma ve çoğaltmada ekonomi gibi ölçütler dikkate alınır.

Kısa cevap testlerinde maddelerin soru - cevap yeri düzeninde sıralanması gelenekselleşmiştir. Bu düzenlemede cevap yerlerinin tek tip olması, hepsinin sağ marjinde bitmesi veya sayfanın sağ yarısında sıralanması tercih edilir.⁴ Testin daktilosunda maddeleri alt alta sıralamaya, bir maddenin diğerinden belirli aralıklarla ayrılmasına dikkat edilmelidir.

⁴Stanley, op. cit.,s. 209.

PUANLAMA

Puanlama işlemine başlamadan önce bir puanlama anahtarı hazırlanmalıdır. Anahtar her maddeye verilmesi beklenen doğru cevapları ayrıntılarıyla listelemelidir. Her doğru cevaba 1 puan verilmesi çoğu halde maksada yeter. Soruların güçlük dereceleri ve cevapların ayrıntıları birbirinden çok farklı olmadıkça, maddelere farklı puan vermek gerekmez. Maddelere eşit ağırlık verilmesi, puanlama, madde analizi ve test geliştirme işlemlerinde birçok kolaylık sağlar.

Beklenen cevaptan çok farklı veya yanlış bir cevaba, negatif puan verilmesini savunanlar da vardır. Bir soruya apaçık yanlış cevap vermenin hiç cevap vermemekten daha kötü olduğu kabul edilebilir. Fakat, yanlış bilginin diğer doğru bilgileri yok etmesi akla yakın değildir. Bilgi alacak - borç gibi iki yönü olan cebirsel bir büyüklük değildir. Ancak, test sonuçlarına bakarak, yanlışları düzeltici bir eğitim planlanacaksa, yanlışların ayrıca listelenmesi ve analizi gerekebilir.

B Ö L Ü M 6

Sınıflama Gerektiren Testler

Bazı testlerde, cevaplayıcılardan, maddeleri belli bir ölçüte göre sınıflamaları istenir. Bu yapıdaki testlerin en çok bilineni, maddelerin cevaplayıcıya «doğru - yanlış» olarak iki sınıfa sınıflatıldığı test türüdür. Doğru - yanlış testleri diye bilinen testlerde çeşitli yapıda maddeler kullanılmıştır. Ayrıca, yapısı itibarıyla doğru yanlış testlerine benzeyen, fakat ikiden fazla kategoriye göre sınıflama gerektiren testler de vardır. Bu bölümde önce doğru - yanlış yapısındaki maddeler, sonra ikiden fazla kategoriye göre sınıflama gerektiren madde tipleri tartışılacaktır.

DOĞRU - YANLIŞ TESTLERİ

Doğru - yanlış testleri bazıları yanlış, bazıları doğru önermeler halinde verilen maddelerden oluşur. Cevaplayıcıdan, her maddeyi okuması, madde kapsamındaki fikre göre onu ya doğru ya da yanlış olarak sınıflaması istenir. Böylece, bu sınıflamaya bakarak, cevaplayıcının test kapsamındaki bilgileri ne derece bildiği, hiç değilse maddelerdeki fikirlerin doğru olup olmadığını ne derecede kestirebildiği ölçülür. Örnek :

Açıklama : Aşağıdaki cümlelerden doğru olanların başına D, yanlış olanların başına Y koyunuz.

1. Bir dik üçgenin alanı, dik kenar uzunlukları çarpımının yarısına eşittir.
2. Bir eşkenar dörtgenin alanı, köşegenlerinin çarpımına eşittir.

Bu yapıdaki testler, eğitimde testlerin kullanılmaya başladığı yıllarda çok tutuluyordu. Fakat, sonraları teknik özellikleri daha üstün maddelerin bulunmasıyla, doğru - yanlış testleri daha seyrek kullanılır oldu. Bu yapıdaki testlerle yoklamaya elverişli bilgiler vardır. Hatta, bir yazar, bazı bilgilerin, doğru - yanlış yapısındaki maddelerde daha etkili bir biçimde ölçülebileceğini savunur.¹ Diğer test-

¹R.L. Ebel, *Measuring Educational Achievement*, s. 124 - 148.

lerde olduđu gibi, bu çeşit testlerin de madde yapısından gelen sınıncaları ve üstünlükleri vardır. Doğru - yanlış testlerinin (ve daha genel olarak sınıflama testlerinin) özelliklerinin ve madde yazma tekniklerinin iyi bilinmesi, öğretmene, bu tipteki testleri yerinde kullanabilme ve gerektiği gibi hazırlayabilme yeterliği kazandırır.

Doğru - Yanlış Testlerinin Genel Özellikleri

1. Test maddeleri bir kısmı yanlış bir kısmı doğru önermeler halinde verilmiştir. Cevaplayıcıdan, her cümlemin doğru olup olmadığını tayin etmesi ve cevabını «doğru» veya «yanlış» sembolleriyle kaydetmesi istenir. Cevaplayıcı, maddede verilen önermenin doğruluğunu ya da yanlışlığını, kendi bilgilerine göre ve kendi yargısını kullanarak kararlaştırır. Doğru - yanlış testlerinin değişik yapıdaki maddelerinde izin verilmekle birlikte, ana yapıdaki maddelerde «bilmiyorum», «bazı koşullarla doğru», v.b. gibi cevaplara izin verilmez. Cevaplayıcı bir maddeye ya «doğru» ya da «yanlış» demek zorundadır. Bununla birlikte, bazı cevaplayıcıların bazı maddeleri okudukları halde cevapsız bırakmaları da mümkündür. Test cevapları iki seçenekli olmakla birlikte, «cevapsız» veya «erişilememiş» maddeler de bulunabileceği için, uygulamada üç - dört çeşit cevap elde edilebilir.

2. Cevaplayıcı, cevaplarını yalnız bir harf yazarak veya bir işaret koyarak kaydeder. Onun için sınav süresinin hemen hepsi maddeleri okumaya, doğru olup olmadıklarını düşünmeye harcanır. Böylelikle, bir sınav süresi içinde, diğer test çeşitlerine kıyasla çok geniş bir bilgi alanı yoklanabilir.

3. Cevapların bütün cevaplayıcılar için ortak olan iki işaretle verilmiş olması, bu yapıdaki maddelerin puanlanmasını hem kolaylaştırır hem de tamamen objektifleştirir. Test ile cevap kâğıdı birbirinden ayrılabilmesi için cevap kâğıdına istenilen düzen verilebilir; böylece cevaplar makina ile puanlanabilir.

4. Madde yapısı basittir, fakat basitliğine rağmen, oldukça ileri basamaktaki bilgileri yoklamaya da elverişlidir. Bu yapıdaki maddelerle belleğe dayanan bilgilerin yanında, açıklama, karşılaştırma, yorumlama, genelleme, uygulama gibi zihin etkinliklerini gerektiren bilgiler de yoklanabilir. Fakat, hatırlamadan ileri zihin yeteneklerini yoklayacak nitelikte maddelerin yazılması ustalık ve yaratıcılık gerektirir.

5. Cevaplayıcıya verilecek test yönergesi kısa ve basittir. Cevaplayıcıdan istenen cevaplama işlemi, ölçülen bilgi veya yeteneğin

dışında bir zorluk göstermez. Bu sebeple, doğru - yanlış testleri eğitimin aşağı basamaklarında bile kolaylıkla uygulanabilir.

6. Sadece iki cevap seçeneği bulunduğu için, önermenin doğru olup olmadığını bilmeden işaretleyen bir cevaplayıcının istenilen cevabı tutturabilme şansı oldukça yüksektir. Bu yapıdaki testlerin puanlarına, şansla doğru cevabın bulunmasından doğan bir hata karışır. Bu hata, cevap seçeneklerinin sayısı azaldıkça büyür; onun için sadece iki seçeneği olan doğru - yanlış testlerinde en büyüktür.

Doğru - Yanlış Maddelerinin Yazılması

Her test çeşidinde olduğu gibi, bu test çeşidinde de maddelerin yapısından gelen bazı teknik zorlukların farkına varmak ve bunlardan doğacak sakıncaları önlemek, ancak madde yazmada yeterliğe sahip olmakla mümkündür. Doğru - yanlış testleri için madde yazmada dikkat edilecek esaslar aşağıdaki sekiz noktada toplanabilir.

1. Her maddede sadece bir tane ana fikir bulunmalıdır. Birden fazla ana fikir taşıyan maddelerde fikirlerden biri yanlış diğeri doğru ise, sadece yanlış fikrin yanlışlığını bilen bir cevaplayıcı, diğer fikir hakkında hiç bilgi sahibi olmasa bile, maddeyi doğru cevaplandırabilir. Örnek :

3. Foklar kutup denizlerinde yaşar ve diğer balıklar gibi yumurtlayarak ürerler. (Y)

Bu örnekte iki ana fikir vardır. Fokların nerede yaşadığı sorulacaksa maddenin birinci kısmı, nasıl üredikleri sorulacaksa ikinci kısmı yeter. Böyle maddeleri düzeltmenin en kestirme yolu, her ana fikri bağımsız bir madde haline getirmektir.

2. Her madde kesinlikle doğru, ya da kesinlikle yanlış olmalıdır. Bir maddenin yanlışlığı veya doğruluğu başka bir açıklamaya hacet kalmaksızın tayin edilebilmelidir. Bunun için, önermenin doğru olup olmadığını ortaya koyacak minimum bilgiler maddede verilmiş olmalıdır. Belirsiz ifadeler, yeteri kadar ayrıntısı bulunmayan cümleler cevaplayıcıyı kararsızlığa düşürür. Örnekler :

4. İkinci Dünya Savaşı Avrupa ve Afrika'da cereyan etmiştir. (D)

Bu maddenin yazarına göre, madde «doğru» olarak anahtarılanmıştır. Gerçekte bu madde ne doğru, ne de yanlıştır. İkinci Dünya Savaşının cereyan ettiği kıt'aların hepsi sayılmadığı için madde eksiktir. İkinci Dünya savaşını iyi bilen bir cevaplayıcı bu maddeye

(Y) cevabı verebilir. Anahtarın (D)'den (Y)'ye çevrilmesi maddeyi düzeltmez.

Pi sayısının sayısal değerini soran şu maddeler de, yeteri kadar ayrıntı vermedikleri için kusurludur.

5. Pi sayısının değeri yaklaşık olarak 3 kabul edilir. (D)
8. Pi sayısı 22/7 dir. (D)

Beşinci maddede pi sayısının değerinin hangi basamakta yaklaşıklıkla sorulduğu belli değildir. Madde şöyle yazılırsa, açıklık kazanır:

- 5a. Pi sayısının değeri, dört ondalık basamağa kadar kesinlikle 3,1416 dir. (D)

Altıncı madde ile de pi sayısının daha kesin bir özeliği yoklanacaksa, bu madde şöyle düzeltilebilir :

- 6a. Pi sayısı rasyonel bir sayıdır. (Y)

Bir otoriteye veya bir sayılıya göre doğru, diğer bir otoriteye göre yanlış ya da eksik olan bilgiler, maddede otoriteden söz edilmeksizin verilirse, cevaplayıcıyı kararsızlığa götürebilir. Aşağıdaki maddeler bu yönden kusurludur.

7. Kuşların göçü bugün iyice açıklanabilmektedir. (D)
8. Işık çok hızlı hareket eden çok küçük madde tanecikleri şeklinden ibarettir. (D)

Kusurları yukarıda olduğu gibi eksiklikten, açıklama yetersizliğinden ve ifade belirsizliğinden gelen maddeler çoğu halde bilgi seviyesi daha yüksek olan cevaplayıcıları yanıltır. Onun için, böyle maddelerin geçerliliği (ayırma gücü) negatif olma eğilimindedir.

Bilgilerimizin çoğu kesinlikle doğru veya yanlış olarak sınıflanabilecek nitelikte değildir. Birçok bilgi ve birçok önerme ancak bazı koşullarda kesinlik kazanır. Bu gerçek, doğru - yanlış testlerinin ancak pek az alanda başarıyla uygulanabileceği düşüncesini doğurur. Fakat, bu güçlük madde yazarının yaratıcılığı ile azaltılabilir. Aşağıdaki madde buna bir örnektir.²

9. Bir dairenin içine bir kare ve bir eşkenar dörtgen çizilse, karenin kenarı üçgeninkinden büyük olur. (Y)

Madde yazarı, önermelerini gerekli koşullarla vererek kesinlik kazandırabilir; bunu yaparken de önermenin yanlışlığını veya doğruluğunu açığa vuracak geçersiz ipuçlarından sakınmalıdır.

²A.G. Wesman, «Writing the Test Item», in R.L. Thorndike (Ed.) Educational Measurement, 2nd ed., s. 94.

3. Bir maddenin yanlışlığı, önemsiz veya atlatıcı bir noktada olmamalıdır. Önemsiz ayrıntıları veya kolaylıkla algılanamayacak kısımları değiştirilerek yanlış hale getirilmiş cümleler iyi test maddesi olamaz. Örnekler :

10. İstanbul'un Türkler tarafından fethi 1543 yılına raslar. (Y)
11. Demir uzun zaman havada kalınca oksitlendiği için kararır. (Y)

Bu maddelerden birincisi yanlıştır; doğrusunda tarih 1453 olmalıdır. Bazı cevaplayıcılar 1543 rakamını doğru algılayamadıkları için bu maddeye «doğru» işareti koyabilir. Bu haliyle madde istenilen tarih bilgisinin yanında, okuduğunu algılama gücünü ve dikkati de ölçer. İkinci maddenin yazarı demir oksidin rengini sormak istemiştir. Ona göre, maddeyi yanlış yapan son kelimedir. Gerçekte demirin havada paslanması demir oksit oluşmasından farklı bir olaydır. Bu haliyle madde, demirin paslanması konusunda önemli bir bilgiyi yoklayamamaktadır. Her madde doğru olup olmadığı bilinmesi gerekli bilgilere dayanmalıdır. Onun için, yanlış hale getirilecek bilgi cümlelerin önemli bir fikri olmalıdır. «Yanlış» madde yazmada, sınıf tartışması, yazılı yoklama, ödev gibi etkinliklerde öğrencilerin düştüğü önemli hatalardan yararlanılabilir.

4. Her madde kısa ve açık yazılmalıdır. Bu madde gereksiz ayrıntılarla şişirilmemeli, açıklığı kaybedilmeden mümkün olduğu kadar kısa yazılmalıdır. Maddedeki ana fikrin doğruluğuna etkisi olmayan ifadeler, uzun ve dolaşık cümleler, ebedi olsun diye açıklığı kaybedilmiş önermeler iyi test maddesi olamaz. Bazı niteleyiciler («çoğu zaman», «büyük ölçüde», v.b. gibi) her cevaplayıcı tarafından aynı anlamda yorumlanmadıkları için maddeye belirsizlik verir. Maddede böyle niteleyiciler kullanılmamalıdır.

5. Olumsuz, özellikle çift olumsuz ifadeler kullanılmamalıdır. Doğru bir önermeden «yanlış» bir önerme elde etmek için cümlelerin fiilini olumsuz hale getirme yoluna gidilmemelidir. Böyle maddelerde «me», «ma» gibi olumsuzluk ekleri bazı cevaplayıcının gözünden kaçabilir. Üstelik, böylece «yanlış» hale getirilmiş maddelerin yanlış olduğu ve zorla yanlış hale getirildiği kolayca anlaşılabilir. Örnekler :

12. Sıtma Insandan Insana sıvrisinekle geçmez. (Y)
13. Sıcak havalarda fazla tuzlu yememelidir. (Y)

Bu maddeler aşağıdaki gibi yazılırsa hem algılama hataları azalır, hem de «yanlış» oldukları açıkça belli olmaz.

- 12a. Sıtma Insandan Insana karasinekle geçer. (Y)
- 13a. Sıcak havalarda az tuzlu yemelidir. (Y)

Maddede çift olumsuz kullanılması ifadenin anlaşılmasını güçleştirir. Bir cümledeki çift olumsuzdan kurtulmanın en kısa yolu, o cümleyi olumlu halde ifade etmektir. Örnek :

14. Timur Anadolu'ya gelmemiş değildir.

(D)

14a. Timur Anadolu'ya gelmiştir.

(D)

6. Maddelerin uzunlukları birbirine yakın olmalıdır. Tecrübesiz yazarlar «doğru» maddeleri gereğinden daha uzun yazma eğilimindedirler. Bunun farkına varan bazı cevaplayıcılar, istenilen bilgilere sahip olmadıkları halde, maddenin uzunluğuna bakarak «doğru» yargısına varabilirler; böylece, gerçek bilgilerinin gerektirdiğinden fazla puna kazanırlar.

7. Bir maddede birden fazla yanlış fikir bulunmamalıdır. Bir madde iki veya daha fazla yanlış fikir taşıyorsa, şüphesiz, «yanlış» olmak gerekir. Fakat, bu fikirlerden yalnız bir tanesinin yanlış olduğunu bilen bir cevaplayıcı, maddeye «yanlış» işareti koyar. Başka bir deyişle, madde gereğinden fazla kolaylaşmış olur. Böyle maddelerin ifadelerini yalnız bir yanlışla indirmek, eğer her iki fikir de yoklamaya değer görüyorsa, o maddeyi bağımsız iki madde halinde yeniden yazmak gerekir.

8. Bir maddenin yanlışlığını veya doğruluğunu apaçık belirtecek ipuçları bulunmamalıdır. Maddeyi oluşturacak öneri öyle seçilmelidir ki, o madde ile ölçülmek istenen bilgiye sahip olmayan bir cevaplayıcı, maddenin doğru olup olmadığını başka yollardan meydana çıkaramasın. Örnek :

15. Yavuz Sultan Selim'in Mısır seferi 515 - 517 yılları arasında iki yıl sürmüştür.

(Y)

Bu maddenin yazarına göre, bu cümlelerin yanlış 515 - 517 tarihlerindedir. Bu madde, algılanması zor bir ayrıntıda yanlışla çevrildiği için kusurludur. Ayrıca, rakamları doğru okuyan bir cevaplayıcı, Yavuz'un Osmanlı Padişahlarından biri olduğunu ve Osmanlılar çağını kabaca bilse bile bu maddeye «yanlış» işareti koyabilir. Onun için, bu madde Mısır seferinin tarihini yoklayamaz. Mısır seferi gerçekten iki yıl sürdüyse, maddedeki «iki yıl sürmüştür» ifadesi gereksizdir.

9. Test maddesiyle önemsiz ayrıntılar, hemen her kaynaktan bulunabilecek ezberlenebilir bilgiler yoklanmamalıdır. Her madde ile yoklanacak bilgi, doğru olup olmadığı bilinmeye değer bir bilgi olmalıdır. Hatırlamaya dayanan basit bilgiler yerine, kavrayabilme, yeni durumlara uygulayabilme, v.b. gibi ileri zihin yeteneklerinin de doğru - yanlış yapısındaki maddelerle yoklanması mümkündür. Test

teknğinde tecrübesiz öğretmenler, genellikle, hatırlamaya dayanan basit bilgilerle cevaplandırılabilcek maddeler yazarlar. Şu iki maddede böyledir.

16. Türkiye Cumhuriyetinin kurulmasıyla, Türk toplumuna devrimler de gelmiştir. (D)

17. Ay batıdan doğar ve doğudan batar. (Y)

Buna karşılık, tecrübeli ve yaratıcı öğretmenler daha ileri zihin yeteneklerini yoklayan maddeler yazarlar. Aşağıdaki iki örnek, aynı konularda daha derin bilgi ve yeteneklerle doğru cevaplandırılabilir

18. Cumhuriyet dönemi devrimleriyle Türkiye'deki batılılaşma hareketlerinin kapsamı genişlemiş, hızı artmıştır. (D)

19. Ayın doğuşu ile batışı arasındaki süre, aynı günde güneşin doğuşu ile batışı arasındaki süreden daha kısadır. (Y)

Doğru - Yanlış Testlerinde Diğer Madde Çeşitleri

Doğru - yanlış testlerinin madde yapısından gelen önemli birkaç sakıncası vardır. Bunlardan biri, doğru - yanlış testlerinde şansla doğru cevabın bulunabilmesinden doğan şans hatasının büyük oluşudur. Şans başarısını azaltma gayretlerinden, doğru - yanlış türündeki test maddelerinde çeşitli yapısal değişiklikler ortaya çıkmıştır. Madde yapısındaki bu değişiklikler, başlıca iki grupta toplanabilir. Birinci gruptaki değişikliklerde, ikili sınıflamadan ayrılmaksızın, şans başarısını önleme yoluna gidilmiştir. İkinci gruptaki değişikliklerde ise, sınıflamadaki seçenek sayısı artırılarak, şans başarısı olasılığını küçültme yoluna gidilmiştir. Doğru - yanlış testlerinde ikili sınıflamadan ayrılmadan yapılan madde değişiklikleri bu kesimde, ikiden fazla seçenek ile sınıflama gerektiren madde çeşitleri bundan sonraki kesimde incelenmektedir.

Alışılmış yapıdaki bir doğru - yanlış testi aşağıdaki gibi bir yönergeye uygulanır.

Açıklama : Aşağıdaki cümlelerde doğru olanların başına D, yanlış olanların başına Y koyunuz.

20. Su oksijen ve karbondan oluşur. (Y)

21. Sönmemiş kireç suda eriyince bir baz meydana gelir. (D)

Bu maddelerin önemli yapısal kusurları yoktur. Maddelerin gerektirdiği bilgilere sahip olmayan bir cevaplayıcı, düşünmeden D ve Y koysa, onahtara göre doğru cevabı rasgele tutturabilme olasılığı:

her madde için 0,50 olur. Fakat, cevaplayıcıdan önce her maddenin doğru olup olmadığını kararlaştırması, sonra da «Y» ile işaretlediği maddelerde cümleyi «yanlış» yapan ifadeyi belirtmesi istense, cevaplama sadece şansa kalmış olmaz. Bu değişiklik yapılırsa, test şu açıklama ile uygulanır :

Açıklama : Aşağıdaki cümlelerden doğru olanların başına D koyunuz, yanlış olanların başına Y koyup cümlelerin yanlış kısmının altını çiziniz.

Test bu açıklama ile uygulanınca, «yanlış» olan 20 numaralı maddeye şöyle cevap verilmiş olabilir :

- (a) Y - 20. Su oksijen ve karbondan oluşur.
- (b) Y - 20. Su oksijen ve karbondan oluşur.

Bu örneklerde cevaplayıcıların altını çizdikleri (siyah basılmış) kelimelerden açıkça anlaşılmaktadır ki, kusursuz gibi görünen bu madde, iki cevaplayıcı tarafından farklı anlamlarda yorumlanmıştır. Yanlış bir önerme, ancak, cümlelerin doğru kabul edilen diğer kısımlarına göre düzeltilebilir. Test maddesinde ise, hangi kısımların doğru kabul edileceği belirtilemez. Böyle olunca, doğru - yanlış türündeki maddelerin yukarıdaki değişiklikle uygulanması, testin tam objektif olarak puanlanabilmesini önler.

Doğru - yanlış testlerinde diğer bir değişiklik, aşağıdaki örnekte görüldüğü gibi, önermenin yanlış kısmını düzelttiren madde yapısına gidilerek meydana getirilmiştir.

Açıklama : Aşağıdaki cümlelerden doğru olanların başına D koyunuz; yanlışların yanlış kısımlarını çizip, altına doğrusunu yazınız.

- (c) Y - 20. Su oksijen ve karbondan oluşur.
Karbondioksit
- (d) Y - 20. Su oksijen ve karbondan oluşur.
hidrojenden

Burada görüldüğü gibi, test maddesinde verilen önerme yanlıştır; fakat, her iki düzeltme de doğrudur. Demek ki, maddenin birden fazla doğru cevabı bulunabilmektedir. Bu haliyle madde, doğru - yanlış yapısında bir madde olmaktan uzaklaşıp, kısa cevap gerektiren bir madde yapısına yaklaşmıştır.

Doğru - yanlış testlerine yöneltilen eleştirilerden biri, her cümlelerin ancak bir metinde tam anlam kazandığı, metinden ayrılmış tek tek cümlelerin doğruluğunu tayin etmenin zor olduğu düşüncesine dayanır. Bu sakıncadan kaçınmak için, metinden ayrılmış tek

bir önerme yerine, bir metinden birbirine bağlı birkaç önermeden oluşan test maddeleri yapılması tavsiye edilmiştir.³ Örnekler :

Açıklama : Aşağıdaki maddelerde eksik cümleyi doğru tamamlayan ifadenin sonundaki parantezin içine (+), yanlış tamamlayan ifadenin sonundaki parantezin içine (—) işareti koyunuz.

22. Besinlerimizdeki mineraller

- | | | |
|---|-----|-----|
| (a) vücudumuzda enerji sağlar. | () | (—) |
| (b) kemiklerin ve kanın yapısında kullanılır. | () | (+) |
| (c) vücuttaki bazı biyo - kimyasal olayları düzenlemeye kullanılır. | () | (+) |

Bu örnekte verilen üç önerme kendi başlarına birer test maddesi haline getirilebilirdi. Maddelerin böyle düzenlenmesi, cümlelerin baş tarafı tekrar edilmediği için, ekonomi sağlar; ayrıca, birbirine yakın konular bir araya getirildiğinden, cevaplayıcının düşünme süresini kısaltabilir. Madde bu yapıya getirilince her önermenin ayrı ayrı mı puanlanacağı, yoksa her üç önermenin anahtara uygun olarak işaretlenmesi halinde mi puan verileceği tartışılabilir. Böyle bir maddenin her önermesi, kapsamı bakımından ilişkili olmakla birlikte, birbirinden bağımsız olarak puanlanabilir. Bu sebeple, maddenin her seçeneğine bir puan verilmesi doğru olur.

Önergelerin birbirinden bağımsız olması gerçekten fikirler arası ilişkileri bozuyorsa, bir paragraf tümüyle yeniden yazılabilir. Paragrafta cümlelerin yaklaşık olarak yarısı «doğru», yarısı «yanlış» önergeler haline getirilir. Her cümle ayrı bir madde olarak numaralanır. Cevaplayıcılardan, her cümlemin doğru olup olmadığını tayin etmeleri, cevaplarını ona göre vermeleri istenebilir.

Doğru - yanlış testlerinde başka yapıda maddeler de kullanılmıştır.⁴ Test maddesini cevaplandırma işlemi maddeleri doğru - yanlış ölçütüne göre sınıflama işleminden ayrıldığında, maddenin yapısal özellikleri değişir. Maddenin ayrı bir yapıya dönüşmesi, cevaplamada güçlük yaratabileceği gibi, onunla ölçülen zihin yeteneklerinde de değişiklik yaratabilir. Bu sebeple, yapıları farklı madde çeşitleri kullanılmadan önce, onların testle ölçülecek olan bilgileri yoklayıp yoklayamayacağı iyice incelenmelidir.

SINIFLAMA GEREKTİREN DİĞER MADDE ÇEŞİTLERİ

Yalnız iki sınıfa göre sınıflanan madde çeşitlerinde şans başarısı olasılığının en büyük oluşu, test yapımcılarını, sınıflama ka-

³J.C. Stanley, *Measurement in Today's Schools*, 4th ed., s. 216 - 219.

⁴C.W. Odell, *How to Improve Classroom Testing*, rev. ed.

tegorilerini artırarak şans başarısı olasılığını küçültme düşüncesine götürmüştür. Böylece, temel yapısı bakımından yine sınıflama gerektiren, fakat sınıf sayısı ikiden fazla olan madde çeşitleri türetilmiştir. Örnek :

Açıklama : Aşağıdaki cümleleri okuyunuz. Bir önerme her zaman doğru ise (+), her zaman yanlış ise (—), bazı koşullarla doğru olabilecek ise (0) ile işaretleyerek cevabınızı belirtiniz.

23. Bir paralelkenarda karşılıklı açılar eşittir. () (+)
24. Bir dairenin çapı, çevresinin yarısına eşittir. () (—)
25. İki kenar ve birer açısı eşit olan üçgenler eşittir. () (0)

Bu örnekte testin maddelerinin üçlü bir sınıflama ile cevaplandırılması istenmektedir. Böylece hem şans başarısı olasılığı küçültülmüş, hem de testin daha geçerli bilgileri yoklayabilmesine olanak hazırlanmıştır. Diğer bir örnek şudur :

Açıklama : Aşağıdaki maddelerde her önerme, gerekçesiyle birlikte verilmiştir. Maddenin ilk yarısı önermeyi, «çünkü» kelimesinden sonrası gerekçeyi ifade eder. Bu maddeleri şu anahtarla cevaplandırınız.

Önerme doğru, gerekçe doğru : A
Önerme doğru, gerekçe yanlış : B
Önerme yanlış, gerekçe doğru : C
Önerme yanlış, gerekçe yanlış : D

26. Sıtma sivrisinekle geçer, çünkü sivrisinek mikroplu sularda yaşar. (B)
27. Dengesiz beslenme hastalıklara sebep olur, çünkü vücudun besin maddelerine ancak belli oranlarda ihtiyacı vardır. (A)
28. Sivrisinekle mücadelede kının kullanılır, çünkü kının sıtma mikrobunu öldürür. (C)
29. Sıtma bir karaciğer hastalığıdır, çünkü B vitamini eksikliğinden ortaya çıkar. (D)

Bu örnekteki maddelerin herbiri, birbirine bağlı iki fikirden oluşmuştur. Cevaplayıcının görevi, bu iki fikir birbirine bağlı olarak değerlendirilmesi istendiği için, daha zorlaşmıştır. Maddelerdeki yapısal değişiklik sebebiyle, daha derin bilgilerin yoklanabilmesi imkân arttırılmış; ayrıca, cevaplar dörtlü bir sınıflama ile verileceğinden, şans başarısı olasılığı küçültülmüştür. İkiden fazla katagorilere göre sınıflama gerektiren test maddeleri, yapıları bakımından seçmeli test maddelerine benzetilmektedir. Bundan sonraki bölümde incelenecek olan ortak seçenekli madde yapısı ile, ikiden fazla katagorilere göre sınıflama gerektiren maddeler arasında büyük bir benzerlik vardır.

SINIFLAMA GEREKTİREN TESTLERİN DÜZENLENMESİ

Sınıflama gerektiren testler düzenlenirken, test düzenlemenin genel ilkelerinden başka, madde yapısının gerektirdiği bazı esaslara da uyulur. Doğru - yanlış testlerinde olduğu gibi ikili sınıflama gerektiren testlerde, maddelerin yarısı «doğru», diğer yarısı «yanlış» anahtarlanmış olmalıdır. Böyle yapılmazsa, testin maddelerini düzenli bir dizi izleyerek cevaplandırarak cevaplayıcıların puanlarına sistematik hata karışabilir. Söz gelişi, 50 maddelik bir testte, 35 madde «doğru», 15 madde «yanlış» anahtarlanmış ise, maddeleri hiç okumadan hepsinde D cevabını veren bir cevaplayıcının 35 cevabı doğru ve 15 cevabı yanlış olur. Aynı testte bütün maddelere Y cevabını veren diğer bir cevaplayıcının ise 35 cevabı yanlış ve 15 cevabı doğru olur. Bu puanlar, yanlışlar toplamı doğrular toplamından çıkarılarak düzeltilse, birincinin puanı +20, ikincinininki -20 olur. Gerçekte, her iki cevaplayıcının düzeltilmiş puanlarının sıfır olması gerekirdi. Bu örnekteki hatalar testte doğru madde sayısının yanlış madde sayısına eşit olmayışından doğmuştur.

Maddelerin test içinde sıralanışlarında hiçbir sistemli dizi izlenmemeli, yanlış ve doğru anahtarlı maddeler gelişigüzel karıştırılmalıdır. Maddeler sistemli bir düzende sıralanırsa, sıralanma düzenini çözebilen cevaplayıcılar, bu ipucundan yararlanarak, gereğinden fazla puan kazanırlar.

Aynı testte, sınıflama ölçütleri ve sınıf sayıları birbirinden farklı maddeler kullanmak zorunluğu doğarsa, aynı yapıdaki maddeler bir arada gruplanmalıdır. Böyle hallerde, testin başındaki genel yönerge ile yetinilmemeli; ayrıca, her grubun başında o maddelerin nasıl cevaplandırılacağını gösterir özel açıklamalar verilmelidir.

B Ö L Ü M 7

Seçme Gerektiren Testler

Sorulan bir sorunun cevabını verilen cevaplar arasından seçti-ren maddelerden oluşmuş testlere seçmeli testler denir. Bundan önceki bölümde incelenen sınıflama testlerinin de bir çeşit seçme gerektirdiği düşünülebilir. Bu iki tür madde, farklı yapılarda görüldüğü için, seçmeli test deyimi, bu kitapta daha dar anlamda kullanılmıştır.

Test tekniğinin uygulamaya girişinden kısa bir süre sonra kullanılmaya başlanan seçmeli test maddeleri, bugüne kadar kullanılmış ve şöhretini gittikçe artırmıştır. Buna sebep, seçmeli maddelerin bazı teknik üstünlüklerinin bulunmasına ek olarak, bu madde yapısı üzerinde pek çok araştırma yapılmış olmasıdır. Seçmeli maddeler, yapıları bakımından çeşitlilik gösterir. Bu bölümde seçmeli yapıdaki maddelerin başlıca çeşitleri incelenecektir.

TEMEL KAVRAM VE TANIMLAR

Seçmeli testlerde çok kullanılan bir madde çeşidi, şu örnekteki yapıda olabilir.

1. Çanakkale savaşında Osmanlı orduları hangi devletlere karşı savaşmıştır?
- | | |
|-------------------------|------------------------|
| a) İngiltere ve Amerika | b) Amerika ve İtalya |
| c) İtalya ve Yunanistan | d) İngiltere ve Fransa |
| e) Yunanistan ve Fransa | (D) |

Seçmeli test maddeleriyle ilgili tartışmaları kolaylaştırmak için bu madde yapısına özgü kavramları kısaca açıklamak gerekir.

Madde kökü, madde ile ne sorulduğunu ifade veya ima eden ve sorunun çözümü için gerekli bilgileri veren kısımdır. Yukarıdaki örnekte soru işaretiyle biten cümle o maddenin köküdür.

Cevap şıkları, madde kökündeki soruya, muhtemel doğru cevap olarak verilen ifadelerdir. Yukarıdaki misalde a'dan e'ye kadar sıralanmış olan ifadeler cevap şıkları veya **seçenekler** denir. Maddede verilen seçeneklerden, genellikle yalnız bir tanesi «doğ-

ru» veya «en doğru» cevap olur. Madde yazarının doğru cevap olarak belirttiği bu seçeneğe **anahtara göre doğru cevap**, veya kısaca **anahtarlanmış cevap** denir. **Çeldiriciler**, bir maddenin anahtara göre yanlış seçeneklerdir. Çeldiriciler, bilgisi yetersiz, bilgisiz veya yanlış bilgili cevaplayıcıları yanıltmak için verilir.

Bir test maddesinin, cevaplama yönergesi de dahil olmak üzere, yazıldığı zaman aldığı şekle onun **yapısı** veya **formu** denir. Maddeyi oluşturan sembol, kelime ve ifadelerin düzeni; maddenin nasıl cevaplanacağını belirten açıklamalar; cevapların kaydediliş şekli madde yapısının temel öğeleridir.

Diğer test tiplerinde olduğu gibi, seçmeli testlerde de, madde yapısında değişiklik yaparak o madde ile yoklanacak yetenek ve bilgileri daha iyi şekilde yoklayabilme imkânı sağlanır. Bunun için, test yazarının başlıca madde formlarını tanıması gerekir. Seçmeli yapıdaki maddelerin başlıcaları aşağıda tanıtılmaktadır.

SEÇMELİ MADDELERİN BAŞLICA ÇEŞİTLERİ

Seçmeli testlerde kullanılan maddelerin, yapı bakımından birbirinden farklı birçok çeşidi vardır. Bu çeşitlilik, bazı yazarları, seçmeli test maddelerini yapılarına göre sınıflamaya götürmüştür.¹ Bu keşimde seçmeli madde formları tanıtılırken, madde yapısının belirgin özelliklerini esas alan bir sınıflamaya gidilmiştir. Böylece, doğru cevaba, madde köküne ve maddelerinin gruplanışına göre meydana getirilen üç sınıf aşağıda tartışılmaktadır.

A. Doğru Cevaba Göre Sınıflama

Seçmeli maddeler doğru cevabın niteliği ve sayısına göre sınıflandırıldığında, başlıca beş çeşit madde formu görülür.

1. **Doğru Cevabı Kesin ve Biricik¹ Olan Maddeler** : Madde kökünde sorulan sorunun kesin ve biricik doğru cevabı vardır; diğer cevaplar kesinlikle yanlıştır. Örnek :

2. Türkiye Cumhuriyeti 1961 Anayasası halk oyuna sunulduğunda, devlet başkanlığı hangi makamda idi?
 - a) Kurucu Meclis Başkanlığında
 - b) Başbakanlıkta
 - c) Genel Kurmay Başkanlığında
 - d) Cumhurbaşkanlığında
 - e) Millî Birlik Komitesi Başkanlığında (e)

¹A.G. Wesman, «Writing the Test Item», in R.L. Thorndike (Ed.), Educational Measurement, 2nd, ed., s. 81 - 129.

Bu örnekte görüldüğü gibi, madde kökünde sorulan sorunun biricik doğru cevabı vardır; bu cevap doğru seçenekte verilmiştir. Diğer seçeneklerin hepsi yanlıştır. Böyle bir maddeyle karşılaşan cevaplayıcı, doğru cevabı tayin edebilmek için ya o biricik cevabın doğruluğunu, ya da çeldiricilerin hepsinin yanlış olduğunu bilmelidir.

2. Anahtarlanmış Cevabı En Doğru Olan Maddeler : Bazı hallerde, madde kökünde sorulan sorunun, kesinlikle boğru bir cevabı bulunmayabilir. Hatta, madde kökündeki soruya verilecek cevaplar, doğruluk dereceleri bakımından birbirinden farklı olabilir. Böyle hallerde, madde kökünde kesin bir soru sormaktansa, verilen seçenekler arasından en doğru cevabı seçtirme yoluna gidilir. Örnekler :

3. $\sqrt{46}$ aşağıdakilerden hangisine en yakındır?
a) 6,5 b) 6,6 c) 6,7 d) 6,8 e) 6,9 (d)
4. Okullarımızda en çok kullanılan sınav türü hangisidir?
a) Sözlü sınav b) Yazılı yoklama
c) Kısa cevaplı test d) Seçmeli test (b)
5. Aşağıdaki gözlemlerden hangisi bir gazın moleküllerinin hareket halinde olduğuna en inandırıcı kanıttır?
a) Havanın basıncı vardır.
b) İki gaz çabucak birbirine karışır.
c) Sıcak hava yükselir.
d) Yeteri kadar sıkıştırılan bir gaz sıvılaşır.
e) Soğutulan bir gazın basıncı düşer. (b)

Dördüncü maddenin kökünde «46'nın kare kökü kaçtır?» Sorusu sorulsaydı, seçeneklerde verilen cevapların hiçbirisi doğru olmazdı; karekökün kesin değeri çok basamaklı bir kesir olduğundan, ona uygun çeldiriciler bulmak kolay olmazdı. Madde kökünde istenilen kareköke en yakın değer sorulduğu için, verilen seçeneklerden biri doğru cevap olur. Dördüncü maddenin de objektif olarak saptanabilecek bir doğru cevabı vardır. Öte yandan, beşinci maddenin doğru cevabı bu kadar objektiflikle tayin edilemez. Cevaplayıcının seçimi, daha ziyade, «en inandırıcı» ibaresini yorumlayışına bağlı olacaktır. Madde kökü, seçeneklerin hangi ölçüte göre sıralanacağını açıkça belli etmemektedir. «İnandırıcılık» ölçütü madde kökünde verilmemiştir. Cevaplayıcı böyle bir ölçütü kendisi bulup bu maddeye uygulayacaktır. Cevaplayıcı, madde kökünde seçeneklere yöneltildiği için, bu yapıdaki maddeler, seçeneklerin hepsi birlikte dikkate alınmadan cevaplandırılmaz.

En doğru cevap formu, maddenin yapısal düzeninden değil, madde ile yoklanmak istenen bilgilerin tabiatından (bir sorunun ce-

şitli doğruluk derecelerinde birkaç cevabı bulunuşundan) ve madde yazarının tercihindən doğmaktadır. Bu yapıdaki maddelerin kullanılabileceği birçok durum vardır. Madde ile, bir sıra, bir dizi, bilinmesi gerekli bazı bilgiler yoklanacaksa, dördüncü örnekte olduğu gibi, objektif olarak ve kesinlikle anahtarlanabilen bir soru yazılmalıdır. Bunun için cevaplayıcının hangi ölçüte göre seçme yapacağı, hiçbir yorumu gerektirmeyecek bir açıklıkla, madde kökünde belirtilmelidir. Fakat, cevaplayıcıdan bilginin yanında bir değerlendirme yapabilmesi, en geçerli ölçütü kendisinin bulabilmesi isteniyorsa, madde kökünde kesin bir ölçüt verilmemeli; hatta, cevaplayıcının kullanacağı ölçüt konusunda ipucu verilmemelidir.

Madde yazarı test tekniğinde usta değilse, en doğru cevap formunu seyrek kullanmalıdır. Tecrübesiz yazarlar bu formda test maddeleri üzerinde çalıştıkları takdirde, ölçme gücü yeterli olmayan, basit hatlamaları yoklayan, bazen de anahtarlanması mümkün olmayan maddeler meydana getirirler. Örnekler :

6. Elektrik enerjisinin endüstriye sağladığı en büyük kolaylık nedir?
 - a) Ucuz elde edilebilmesi
 - b) Her yere nakledilebilmesi
 - c) Bol elde edilebilmesi
 - d) Depo edilebilmesi
 - e) Diğer enerji hallerine hemen çevrilebilmesi (?)
7. Aşağıdakilerden hangisi çocuklarda sözlü dil gelişimini engelleyen en önemli sebeptir?
 - a) Uyumsuz kişilik
 - b) Ağır işitme
 - c) Kusurlu ses anatomisi
 - d) Zekâ geriliği
 - e) Aileden ayrı yaşama (b)

Altıncı maddede bir uzmanlar grubunun üzerinde anlaşabilecekleri bir doğru cevap yoktur. Yedinci maddenin anahtarının doğruluğu da tartışılabilir niteliktedir.

Ninayet, bu formdaki maddeler yazılırken dikkat edilmesi gereken iki özel nokta vardır. Aynı derecede doğru iki seçenek, ister çeldirici ister doğru cevap olsun, bu yapıdaki maddelerde kullanılmamalıdır. «Yukarıdakilerin hiçbirisi» ve «yukarıdakilerin hepsi» seçenekleri de bu formda kullanılamaz.

3. Anahtarlanmış Cevapları Birden Fazla Olan Maddeler :

Madde yazarı birden fazla doğru cevabı bulunan sorularla uğraşırken güçlükle karşılaşır. Madde kökündeki sorunun sadece bir tek doğru cevabı yoktur. Doğru cevapların sadece biri seçenekler arasında verilerek, bir tek doğru cevaplı bir madde meydana getirmek düşünülebilir. Böyle bir madde bilgi düzeyi yüksek cevaplayıcıları

rahatsız eder; verilen cevabın eksik veya yetersiz olduğu düşünce-sine kapılırlar. Madde seçenekleri arasında birden fazla doğru ce-vap bulunmasını zararsız, hatta yararlı gören yazarlar vardır. Böy-lece meydana getirilen maddenin seçenekleri arasında, iki veya da-ha fazla doğru cevap bulunur. Hangi seçeneklerin ve kaç tanesinin doğru olduğunu tayin etmek cevaplayıcıya bırakılır. Örnekler :

8. Kanın pıhtılaşmasında hangi etken veya etkenler rol oynar?
a) Hemoglobinin oksitlenmesi
b) Kanın ezik dokularla teması
c) Değişmemiş protrombin bulunması
d) Kanın yabancı bir yüzle teması (b, c)
- 8a. Aşağıdaki etkenlerden hangisi veya hangileri okul çağındaki çocukların zihin gelişimini etkiler?
a) Beslenme b) Eğitim c) Soyaçekim
d) Beden gelişimi e) Sosyal etkileşim (b, c, e)

Bu örneklerde görüldüğü gibi, bu form madde yazmayı kolay-laştırır; fakat puanlamada ve madde analizinde önemli güçlükler yaratır. Doğru cevaplardan herbirine 1 veren bir puanlama yöntemi yukarıdaki maddelere uygulanamaz; uygulandığı takdirde seçenek-lerin hepsini işaretleyen bir cevaplayıcı sekizinci maddede 2, 8a'da 3 puan kazanır. Doğru - yanlış testlerinin puanlanmasında olduğu gibi, her maddede yanlışlar toplamının doğrular toplamından çıkarılacağı bir puanlama yönetimine gitmek gerekir. Bu ise hem test puanla-ma, hem de test ve madde analizi işlemlerini güçleştirir. Ayrıca tes-tin, doğru ve yanlış anahtarlanmış madde sayıları bakımından den-gesiz bir doğru - yanlış testi haline gelmesi tehlikesi de vardır. Onun için, önemli bir zorunluluk yoksa, bu yapıdaki maddeler kullanılma-malıdır. Aşağıda görüleceği gibi, madde kökündeki sorunun birden fazla doğru cevabı bulunduğu hallerde, bileşik cevaplı madde for-mu, çok cevaplı madde formundan daha kullanışlıdır.

Ancak, iki cevabın doğru olduğu, bir tek veya ikiden fazla ce-vap işaretlendiği takdirde o maddeye verilen cevabın yanlış sayıl-dığı test maddeleri de vardır. Genellikle, «en çok - en az», «en bü-yük - en küçük», v.b. gibi ölçütlerle iki cevabı seçtiren bu yapıdaki maddeler daha kolay puanlanır. Örnekler :

9. Diğer koşullar eşit olmak kaydıyla puanlama güvenilirliği en düşük ve en yüksek olan ölçme araç veya yöntemlerini seçiniz.
a) Sözlü sınav b) Grup ödevi c) Seçmeli test
d) Yazılı yoklama e) Bireysel ödev (a, c)
10. Aşağıdaki iyonlardan elektron kazanabilme gücü en büyük ve en küçük olanlar hangileridir?
a) Hg b) H c) Fe d) Ca e) Na (a, e)

Bu yapıdaki maddelerde, cevaplayıcı, doğru cevabın ancak iki seçimle doğru olacağını bilmektedir. Onun için madde, ancak anah-tara göre doğru iki seçenek işaretlendiğinde 1 puan takdir edilecek bir yöntemle puanlanabilir. Bu yöntemle puanlama, şansla doğru cevabın bulunmasından doğan hatayı önemli derecede azaltır.² Herhangi bir sıra, dizi veya elemanların büyüklükleri, bilinmesi gerekli bilgiler ise, böyle bilgileri seçmeli madde formuyla yoklayabilmenin yollarından biri, yukarıdaki gibi iki cevaplı maddeler yazmaktır. Böylece, bir tek cevabı doğru olan maddelerden daha geçerli bir araç hazırlayabilmek olanağı doğar.

4. Bileşik Cevap Gerektiren Maddeler : Birden fazla cevaplı maddelerin yazılmasındaki kolaylığa rağmen, kullanılması zorluk madde yazarlarını birden fazla cevaplı sorular için, başka bir madde formu aramaya zorlamıştır. Bu gayretlerden bileşik cevaplı madde formu doğmuştur. Örnek :

11. Bir K katı maddesi ısıtıldığında, hepsi bir S sıvısı ile bir G gazına ayrışıyor. Bu maddelerden hangisi bir element olabilir?
- | | | |
|-------------|-------------|------------|
| a) Yalnız G | b) Yalnız S | c) K ile G |
| d) K ile S | e) S ile G | (e) |

Bu maddenin doğru cevabı e'dir. a ve b seçeneklerinden herbiri doğru cevap gibi görünebilir. Onun için, dikkatsiz bir cevaplayıcı a veya b'den birini doğru cevap olarak seçebilir. Fakat, maddenin yapısı sonucu, e seçeneği, a ve b seçeneklerini kapsar. Madde yukarıdaki gibi açıkça yazılamıyorsa, böyle bileşik seçeneklerden sakınılmalıdır. Bileşik cevaplı madde yapısına diğer bir örnek şudur :

12. M, N ve P doğrusal uzaklıklar olduğuna göre,
 $M > N$ ve $P < N$ ise,

- I. $N + P > M$
II. $2P > M + N$
III. $M + N > M + P$

eşitsizliklerinden hangisi veya hangileri doğru olabilir?

- | | | |
|-------------|-----------------|------------|
| a) Yalnız I | b) Yalnız II | c) I ve II |
| d) I ve III | e) I, II ve III | (d) |

Bu örnekte görüldüğü gibi, madde kökünde romen rakamlarıyla verilen ifadelerin seçeneklerde tekrarlanmaması, maddenin daha kısa yazılmasını sağlamıştır.

5. Doğru Cevabı Gizli Maddeler : Seçmeli testlere, doğru cevap seçenekler arasında verildiği için cevaplayıcının doğru cevabı

²M.F. Turgut, Şans Başarısının Test Puanlarına Etkisi, s. 47 - 66.

tanımakta güçlük çekmeyeceği ve bundan dolayı seçmeli testlerle önemli bilgilerin yoklanamayacağı eleştirisi yöneltilir. Bu eleştiride bir gerçek payı vardır. Özellikle, madde ustaca yazılmamışsa, madde kökü ile doğru cevap arasında kurulabilecek bir çağrışım bile maddenin doğru cevaplandırılmasına yardım eder. Seçmeli testlerin bu sakıncasını önlemek isteyen yazarlar, doğru cevabı gizleyebilme yolları aramışlar ve bazı yöntemler bulmuşlardır. Örnekler :

13. Telefonu kim icat etmiştir?
a) Bell b) Hertz c) Edison
d) Marconi e) Erickson (a)
14. Telefonu icat eden kişinin soyadı aşağıdakilerden hangisi arasında bir harfle başlar?
a) A - D b) D - F c) G - I
d) K - S e) P - T (a)

Bu maddelerde sorulan bilgi 12 numaralı madde ile yoklanırsa, çağrışım veya görünce tanıma gibi zihin süreçleriyle de doğru cevap bulunabilir. Fakat, 13 numaralı maddede mucidin ismi gizlenmiştir; bu madde ancak mucidin adının hatırlanması ve harflerin alfabeadaki sırasının bilinmesiyle doğru cevaplandırılabilir. Doğru cevabı gizlenen maddelerin kullanılabileceği birçok durum vardır. Meselâ, matematikte veya sayısal problem çözme gerektiren diğer durumlarda, işlemin doğru sonucu değil, önemli bir basamağı sorulabilir. Bu basamak öyle seçilir ki, işlemi doğru yapamayan bir cevaplayıcı maddenin doğru cevabını bulamaz. Aşağıdaki örneği inceleyelim.

15. 132 sayısının kare kökü aşağıdakilerden hangisine en yakındır?
a) 11,2 b) 11,3 c) 11,4 d) 11,5 e) 11,7 (d)

Bazı cevaplayıcılar bu soruya karekök olarak değil, seçeneklerden bazılarının karelerini alarak cevaplama yoluna gidebilirler. Amaç, cevaplayıcının karekök algorizmini bilip bilmediğini yoklamaksa, madde bu amaca hizmet etmeyebilir. Aynı soru şöyle düzeltilirse, bu sakıncadan kurtarılabilir :

16. 132 sayısının kare kökünü alseniz, virgülden sonraki ikinci basamak ne olur?
a) 2 b) 3 c) 5 d) 7 e) 8 (e)

B. Madde Köküne Göre Sınıflama

Seçmeli test maddelerini, madde kökünün yazılış şekline göre de sınıflamak mümkündür. Madde kökü soru kipinde veya eksik cümle halinde olabilir. Kökün ifadesi de ya olumlu, ya da olumsuz kipte yazılabilir.

1. **Kökü Soru Kipinde Olan Maddeler** : Yukarıda verilen örneklerin hemen hemen hepsinde madde kökü soru kipinde yazılmıştır. Madde kökünün açık seçik bir soru sorması, maddenin anlaşılmasını kolaylaştırır. Diğer formlarda tecrübeli olmayan yazarların soru şeklindeki kökü tercih etmeleri, madde kusurlarını önemli derecede azaltır.

2. **Kökü Eksik Cümle Olan Maddeler** : Madde kökü, seçeneklerden herbiri ile tamamlanan bir cümlenin bir kısmından oluşan maddelere, eksik cümle yapısındaki maddeler denir. Örnek :

17. Met olayının sebebi deniz suyunun
- a) kara kütlelerince çekilmesidir.
 - b) güneş ışınlarıyla ısınmasıdır.
 - c) rüzgarlar tarafından itilmesidir.
 - d) ay tarafından çekilmesidir.
 - e) güneş tarafından itilmesidir.

(d)

Bu örnekte görüldüğü gibi, madde kökü bir cümlenin ilk kısmını, seçeneklerden herhangi biri de ikinci kısmını meydana getirmektedir. Bu form bazı hallerde ekonomi sağlamakla birlikte, yazılması güç bir madde formudur. Testler memleketimize Anglo - Sakson kültüründen girmiştir; hâlen de oradan girmektedir. İngilizcenin cümle yapısı, eksik cümle formunda kök yazmaya elverişlidir. O dilde bu form altında daha kısa ve daha basit maddeler meydana getirilebilir. Fakat, bizim tecrübelerimiz göstermiştir ki, Türkçenin cümle yapısı bu formda madde yazmaya pek elverişli değildir. Bu yapıdaki maddeleri yazarken karşılaşılan önemli güçlüklerden biri, ne sorulduğunun madde köküyle açıkça ifade edilemeyeşidir. Soru kipinde olmayan eksik bir ifade ile ne sorulduğunu ima edebilmek kolay değildir. Türkçede fiilin sona gelişi bu güçlüğü artırır. Üstelik, yine fiilin sona gelişi, bazı kelimelerin bütün seçeneklerde tekrarlanması ve maddenin uzamasına sebep olur. Eksik cümle halinde ifade edilen madde kökünde ne sorduğu açıkça belli olmazsa, seçmeli madde, bir arada gruplanmış dört - beş «doğru - yanlış» maddesi gibi görünür. Ekonomi sağlamadıkça veya maddeye açıklık getirmedikçe, kökü eksik cümle halinde madde yazılmamalıdır.

3. **Kökü olumsuz maddeler** : Madde kökünün olumsuz ifade edildiği madde yapısına olumsuz form denir. Bir soruyu veya bir cümleyi olumlu halde ifade etmek varken, niçin olumsuz hale gidildiği sorulabilir. Aşağıdaki örnekler incelenirse bu soruya bazı cevaplar bulunur

1. **Ortak Köklü Maddeler** : Bir paragraf, bir şekil veya tablo birkaç maddenin köklerinin ortak kısmını teşkil ediyorsa, ortak köklü madde meydana gelir. Bu halde, maddelerin hepsi ortak şekil, tablo veya paragrafa dayanır. Örnekler :

21. - 22. sorular için : Bir ABC dik üçgeninin C tepesinden geçen yüksekliği, hipotenüsü O noktasında keser. $b = 2a$ olduğuna göre, şu soruları cevaplandırınız :

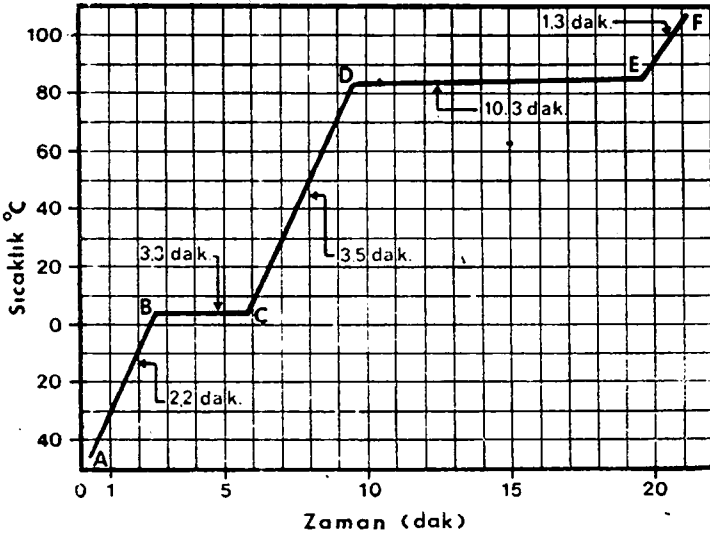
21. AOC üçgeninin alanı OBC üçgeninkinden kaç defa büyüktür?

- a) $\sqrt{2}$ b) $2/\sqrt{2}$ c) 2
d) 3 e) Hiçbiri (e)

22. OA/OB oranı neye eşittir?

- a) $1/\sqrt{2}$ b) $1/2$ c) $3/4$
d) 2 e) 4 (e)

23. - 25. sorular için : Aşağıdaki grafik ısıtılan 78 gr benzenin sıcaklığındaki değişimleri zamanın fonksiyonu olarak gösteriyor. Benzeni ısıtmakta $7,0 \times 10^3$ kal/dak veren sabit bir ısı kaynağı kullanılmıştır. Basınç 1 atmosferde sabit kalmıştır.



23. Grafiğin DE kısmı neyi gösterir?

- a) Katı halde ısınmayı b) Sıvı halde ısınmayı
c) Gaz halde ısınmayı d) Katı halden sıvı hale geçişi
e) Sıvı halden gaz hale geçişi (d)

24. Benzen moleküllerinin ortalama kinetik enerjisinin en yüksek olduğu zaman, grafiğin hangi bölgesinde gösterilmiştir?

- a) AB b) BC c) CD
d) DE e) EF (e)

25. Bu deneyden benzenin erime ısı nasıl hesaplanabilir?

- a) $3,3 \text{ dak} \times 7,0 \times 10^2 \text{ kal/dak}$
- b) $3,3 \text{ dak} \times 78 \text{ gr} \times 700 \text{ kal/dak}$
- c) $(3,3 \text{ dak} \times 7,0 \times 10^2 \text{ kal/dak})/78 \text{ gr}$
- d) $10,3 \text{ dak} \times 700 \text{ kal/dak}$
- e) $(10,3 \text{ dak} \times 700 \text{ kal/dak})/78 \text{ gr}$ (c)

Ortak köklü madde formu pek çok halde uygulanabilir. Bir paragrafı okuyup anlama, bir cetvel veya bir grafiği yorumlayabilme, v.b. gibi bilgi ve beceriler bu yapıdaki maddelerle ekonomik olarak yoklanabilir. Ortak köklü maddelerin hazırlanmasında bir tehlikeden sakınmak gerekir. Bütün maddeler aynı köke dayandığı için, ortak konuya gereğinden fazla ağırlık verilmiş olabilir. Tecrübesiz kişiler ortak kökü hazırladıktan sonra ondan gereğinden fazla madde, hattâ aynı bilgi ve becerileri yoklayan birden fazla madde yazarlar. Bu sebeple toplam testin dengesi bozulabilir. Yukarıdaki örneklerde 21 ve 22 numaralı maddeler aynı bilgi ve aynı işlemlerle cevaplandırılabilir. Gerçek bir testte bu maddelerden biri ile yetinilebilirdi.

2. Ortak Seçenekli Maddeler : Bazı hallerde birkaç maddenin bütün seçenekleri aynı olabilir. Böyle hallerde, ortak seçenekleri her madde kökünden sonra tekrarlamak yerine, maddelerin ortak seçeneklerini başa alıp kökleri daha sonra yazmak suretiyle, farklı bir gruplama meydana getirilebilir. Örnekler :

Aşağıdaki soruları (26. - 30. maddeler) şu seçeneklerle cevaplandırınız :

- a) kare
- b) üçgen
- c) yamuk
- d) dikdörtgen
- e) hiçbir

- 26. Hangi şeklin alanı yalnız bir kenarı bilinmekle hesaplanabilir?
- 27. Hangi şekilde sadece iki kenar birbirine paraleldir?
- 28. Hangi şeklin alanı hesaplanırken π sayısı gereklidir?
- 29. Hangi şeklin iç açıları toplamı iki dik açıya eşittir?
- 30. Hangi şeklin köşegenleri birbirini dik keser?

(26 : a 27 : c 28 : e 29 : b 30 : a)

Aşağıdaki soruları (31. - 33. maddeler) şu seçeneklerle cevaplandırınız :

- a) su
- b) alkol
- c) cıva
- d) zeytinyağı
- e) karbon dioksit

- 31. Hangi madde suyla kolayca karışır?
- 32. Hangi maddenin yoğunluğu en büyüktür?
- 33. Hangi maddenin kaynama noktası en düşüktür?

(31 : b 32 : c 33 : e)

Ortak seçenekli madde formu, şekil, harita, v.b. gibi grafik materyallerle de etkili bir biçimde kullanılabilir. Hatta, ortak seçenek-

ler, madde kökünün bir parçası olan grafik materyal üzerinde bile bulunabilir.

Ortak seçenekli madde formu ekonomi sağlar. Fakat, ortak seçenekli madde formunun en önemli özeliği, diğer yapıdaki maddelerle yoklanması güç olan bazı bilgi ve yetenekleri yoklamaya elverişli oluşudur.

Ortak seçenekli maddelerin yazılmasında en çok dikkat edilecek noktalardan biri, ortak seçeneklerin her madde için gerçekten etkili bir şık olmasını sağlayabilmektir. Yukarıdaki örneklerden ikincisi, bu bakımdan kusurludur. Otuz birinci maddede, hangi maddenin su ile karışabilir olduğu soruluyor. Ortak seçenekler arasındaki «su» şıkkı, bu madde için çeldirici rolü oynamaz. Ortak seçenekli maddelerin cevaplandırılma şekli, cevaplayıcılarca önceden bilinmeyebilir. Cevaplayıcıdan ne istendiğini, cevapların nasıl kaydedileceğini açıkça belirtmek gerekir. Ortak seçenekli maddelerle yalnız hatırlatmaya dayanan bilgilerin ötesinde bazı yetenekler yoklanacaksa, bu nokta daha fazla önem kazanır.

Bazı yazarlar ortak seçenekli madde formunu «eşleştirme» tipindeki testlerle karıştırırlar.³ Eşleştirme tipinde, uyarıcılar ya da sorular sütunundaki elemanların, seçenekler ya da cevaplar sütunundakilerle eşleştirilmesi istenir. Görünüşte ortak şıklı form da eşleştirme tipine benzemektedir. Fakat benzerlik burada biter. Çünkü, eşleştirme tipindeki testlerde sütunlardaki eleman sayısı sınırlı değildir; en kötüsü, kökten köke de seçenek sayısı değişir. Bu özeliğinden dolayı eşleştirme tipindeki testler psikometrik analizlere elverişli değildir. Böyle olduğu için, eşleştirme tipindeki maddeler testlerde kullanılmaz olmuştur.

Bu kesimde seçmeli testlerde kullanılan başlıca madde formları incelendi. Seçmeli testlerde başka yapılarda maddeler de vardır.⁴ Burada tartışılan formlar çok kullanılır. Seçim hemen hemen bütün ihtiyaçları karşılayabilecek derecede geniş, fakat tartışmayı amacından saptırmayacak kadar dar tutulmuştur.

SEÇMELİ TESTLERİN ÖZELİKLERİ

Buraya kadar yapılan tartışmada, seçmeli testler çeşitli madde formlarıyla tanıtıldı. Seçmeli testlerin madde yapısından doğan

³H. Tekin, *Eğitimde Ölçme ve Değerlendirme*, İkinci baskı, s. 157.

⁴C.W. Odel, *How to Improve Classroom Testing*, rev. ed.

özelikleri vardır. Bu özelliklerin bilinmesi, madde yazarlarına bu yapıdaki testleri ustaca kullanma yeteneği kazandırabilir. Onun için, seçmeli testlerin temel özelliklerini burada kısaca tartışmakta fayda vardır.

1. Seçme gerektiren maddelerde, çoğu zaman, sorunun doğru cevabı kendi içinde verilmiştir. Cevaplayıcı, cevabını, verilen seçenekler arasından seçmek zorundadır. Diğer cevaplı testlerde olduğu gibi, seçmeli testlerde de sorulan sorunun cevabı seçenekler arasında verilmiş olduğu için, cevaplayıcının, doğru cevabı tanıması ve çeldiriciler arasından seçebilmesi yeter. Bazı kimseler doğru cevabı tanımanın, cevabı düşünüp bulmaktan daha kolay olacağı, bu sebeple seçmeli testlerin önemli bilgileri yoklayamayacağı düşüncesindedir. Bu düşünce ancak basit hatırlama veya çağrışım ile doğru cevaplandırılabilen maddeler için doğru olabilir. Madde yapısı daha ileri bilgi ve becerileri gerektiriyorsa, doğru cevabı gizli maddelerde olduğu gibi, çağrışım veya hatırlama önlenememişse, bu sakınca ortadan kalkar.

2. Test süresinin büyük bir kısmı, maddeleri okumaya ve doğru cevabı bulmaya harcanır. Seçilen cevabın kaydı pek az zaman alır. Onun için, sınırlı bir test süresinde, yazılı yoklamalara ve hatta kısa cevaplı sınavlara kıyasla, daha çok sayıda soru sorulabilir.

3. Seçmeli testler cevaplayıcıda, testle yoklanmak istenen bilgi ve yeteneklerin yanında, «okuma yeteneği» de gerektirir. Çoğu halde, sınırlı bir sınav süresinde okunacak açıklamalar ve test maddeleri büyük bir bütün teşkil eder. Bir testte okuma yükü ve okuma gücünün kasıtlı olarak aşağı basamaklara indirilememişse, hızlı okuyabilen ve okuduğunu kolayca anlayabilen cevaplayıcılar, okuma yetenekleri sınırlı cevaplayıcılara kıyasla daha yüksek puan kazanırlar.

4. Verilen cevaplar kesinlikle «doğru» veya kesinlikle «yanlış» ya da «en doğru» şeklinde sınıflandırılabilir. Bu sebeple, seçmeli testler objektif olarak puanlanabilir. Puanlama, mekaniksel yollarla da yapılabilir. Anahtar hatalı değilse, seçmeli testlerde önemli puanlama hataları beklenmez.

5. Seçmeli testlerde kullanılabilecek çok çeşitli madde formu vardır. Onun için, seçmeli testler çok çeşitli bilgi, beceri ve yeteneklerin ölçülmesine uygun düşer. Seçmeli testler bu özeliğinden dolayı pek çok alana uygulanmıştır. Seçmeli testlerin öğrencilerin cevabı bulabilme, organize ve kompoze edebilme yeteneklerini ölçmediği

ileri sürülür. Bu yetenekleri ölçmek için bazı seçmeli testler geliştirilmişse de, bu testler pek başarılı olmamıştır. İleri sürülen sınırlılıkta bir gerçek payı vardır.

6. Seçmeli testler, basit madde formları kullanılmak şartıyla, aşağı eğitim basamaklarında da uygulanabilir. Ayrı cevap kâğıdı kullanmadan, ilkokul üçüncü - dördüncü sınıfta bu yapıdaki testler uygulanabilir. Bu bakımdan da seçmeli testlerin geniş bir uygulama alanı vardır.

7. Seçmeli maddelerde doğru cevabın şansa bulunabilmesi ihtimali vardır. Onun için, şans başarısından doğan hata, doğru-yanlış testlerindeki kadar büyük olmamakla birlikte, seçmeli testlerden alınan puanlara da girer. Fakat, maddedeki seçenek sayısı arttıkça şans başarısı ihtimali azalır. Testteki madde sayısı arttıkça da, testin tümünden sadece şans başarısı ile önemli bir puan kazanma ihtimali de hızla küçülür.

8. Seçmeli testlerde maddeler yazılırken test güçlüğünü ayarlamak veya test denendikten sonra istenilen güçlükte madde seçmek mümkündür. Her maddenin objektif olarak puanlanabilmesi, madde güçlüğünün istatistiksel yollardan hesaplanmasına olanak verir. Öte yandan, bir maddenin güçlüğü, hem kapsamındaki problemin güçlüğüne, hem de çeldiricilerin doğru cevaba yakınlık derecesine bağlıdır. Bu değişkenlerle oynayarak madde güçlüğünü ayarlama imkânı vardır.

9. İyi niteliklerde seçmeli testlerin hazırlanması, bu konuda tecrübeli madde yazarlarını gerektirir. Madde yazarlarının test yazma tekniklerini bilmeleri, tecrübeli olmaları ve yaratıcı olmaları gerekir. Bu niteliklere sahip olmayan madde yazarları, hele dikkatli çalışmazlarsa, çoğu halde kusurlu maddeler meydana getirirler.

10. Seçmeli testlerin hazırlanması uzun zaman alır. Buna karşılık, sınavdan sonraki puanlama işlemi az zaman alır. Bu yüzden, seçmeli testlerle yapılacak bir sınav önceden önemli bir hazırlık süresi gerektirir; fakat, sınavın sonuçları oldukça çabuk alınır.

SECMELİ TEST MADDELERİNİN YAZILMASI

Test maddelerinin yazılmasında gözönünde bulundurulması gereken esaslar birçok yazar tarafından tartışılmıştır. Odell,⁵ Ebel,⁶

⁵Ibid.

⁶R.L. Ebel, *Measuring Educational Achievement*, s. 149 - 198.

Thorndike ve Hagen,⁷ Wesman⁸ objektif test maddelerini yazmada uygulanabilecek teknikleri örneklerle açıklamışlardır. Greene ve arkadaşları⁹ orta öğretim seviyesinde, Dréssel ve arkadaşları¹⁰ yüksek öğretimde okutulması gelenekselleşmiş konu alanlarında başarının ölçülmesini tartışırken seçmeli test maddelerinden bol örnekler vermişlerdir. Mosier ve arkadaşları¹¹ seçmeli test maddelerinin yazılmasında kullanılacak bir ölçütler listesi geliştirmişlerdir.

Madde yazarı, madde yazma tekniklerini bilir ve çalışırken göz önünde tutarsa, meydana getireceği maddelerin kalitesi yükselir. Madde yazma teknikleri tecrübeyle öğrenilebilir. Bir yazar, çok çeşitli madde yazdığı, kendi maddelerini bir süre sonra eleştirdiği ve başkalarının maddelerini de eleştirdiği oranda bu tecrübeyi kazanır.

Bu kesimde önce, herhangi bir test maddesinin yazılmasında dikkat edilecek genel esaslar verilmekte; sonra, seçmeli testlerde madde kökünün yazılmasıyla ilgili ilkeler sıralanmakta, daha sonra da madde seçeneklerinin yazımı tartışılmaktadır.

A. Genel İlkeler

1. Test maddesi gayet açık bir dille yazılmalı, ifade kesinlik taşımalı, birbirinden farklı yorumlara imkân verilmemelidir. «Çoğunlukla», «çoğu zaman», «bazen», «nadiren» gibi cümlelerin anlamını değiştirici kelimeler maddeye kesinlik verecek yerde farklı yorumlara yol açar. Çeşitli şekillerde yorumlanabilen bir maddenin, doğru cevabı bilen bazı cevaplayıcılar tarafından bile farklı cevaplandırılması mümkündür. Maddede bu gibi niteleyiciler kullanılmamalıdır.

2. Her maddede ifade açıklığına, kesinliğine ve kısalığına dikkat edilmelidir. Başka bir deyişle, maddenin açıklığı korunmak şartıyla en az kelimeyle ifade edilmesine çalışılmalıdır. Testlerde okuma zamanı çok önemlidir. Okuma yükünü azaltmak testin geçerliğini artırır. Her sorunun gerektiğinden sadece iki - üç kelime uzun yazıldığını kabul edelim. Bu miktar 100 maddelik bir testte 250 - 300

⁷R.L. Thorndike and E. Hagen, *Measurement and Evaluation in Psychology and Education*, 2nd ed., s. 60 - 95.

⁸Wesman, op. cit., s. 81 - 129.

⁹H.A. Greene, et al., *Measurement and Evaluation in the Secondary School*, Chs. 7, 15 - 24.

¹⁰P.L. Dréssel, et al., *Evaluation in Higher Education*, Chs. 4 - 7.

¹¹C.I. Mosier, et al., *Suggestions for the Construction of Multiple-Choice Test Items*, *Educational and Psychological Measurement*, 5 (1945) : 261 - 271.

kelime eder. Bu ise lise düzeyinde 3 - 5 dakikalık okuma zamanının veya 5 - 10 sorunun kaybedilmesi demektir.

3. Test maddelerinde cinaslı, aldatıcı, bilmece gibi ifadeler kullanılmamalıdır. Böyle ifadeler sorunun anlaşılmasını güçleştirir; testle ölçülmek istenen bilgi ve becerilere, yabancı değişkenlerin karıştırılmasına sebep olur.

4. Test maddelerinin dil bakımından güçlük düzeyi, o testi cevaplandıracak grubun ortalama okuma yeteneğinin oldukça altında tutulmalıdır. Teste giren cevaplayıcıların hemen hemen hepsi, bütün maddeleri açıkça anlayabilmelidir. Genellikle kısa cümle, alışılmış terim ve deyimler, düz cümle kipi kullanılırsa okuma güçlüğü azalır. Şüphesiz, test, okuma yeteneklerini yoklamak için hazırlanmışsa, ortak kökün seçiminde ve maddelerin yazılmasında bu tavsiyeye uyulmayabilir.

5. Test maddeleri ve testlerde kullanılacak diğer yazılı malzeme, ders kitabı gibi öğrencinin kullandığı yazılı kaynaklardan kelimeleri kelimesine alınmamalıdır. Bu halde, önceden çözülmüş bir problemin tekrar çözülmesi veya ezberlenmiş bir bilginin hatırlanması yoluyla o madde cevaplandırılabilir. Her madde hiç değilse bir yönüyle bütün cevaplayıcılar için «yeni» olmalıdır. Bir madde, konusu tanınamayacak kadar farklı olmamak kaydıyla, cevaplayıcı için yeni bir durum yaratabildiği oranda geçerlik kazanır.

6. Test maddesi önemsiz ayrıntıları sormamalıdır. Her madde programın kapsamına giren bir konuda öğretimin amacı olan bir zihin davranışını yoklamalıdır. Usta bir yazar, kendine bir konu ve geliştirilmesi amaçlanan bir davranış verildiği zaman, o davranışı yoklayabilecek bir madde yazabilir. Öte yandan, konu ve amaçlarla ilişkisini kesmiş bir yazar, daha ziyade madde formu ile ilgilendiği için formu kusursuz, fakat geçersiz maddeler yaratır.

7. Her test maddesi diğer maddelerden bağımsız olarak cevaplandırılabilirdir. Bir maddenin doğru cevabı, onu izleyen diğer bir madde için veri olarak kullanılmışsa, ikinci maddenin doğru cevaplandırılabilmesi, birincinin doğru cevaplandırılmasına bağlı olur. Bu çeşit maddelerden sakınmak gerekir. Köklerinin bir kısmı ortak olan maddelerle seçenekleri ortak olan maddeler bu anlamda bağımlı sayılmaz.

8. Bir maddede verilen bilgi, diğer bir maddede sorulan sorunun cevabını açıklamamalıdır. Bu husus hem maddeler yazılırken

göz önünde bulundurulmalı, hem de test için madde seçilirken kontrol edilmelidir.

B. Madde Kökünün Yazılması

1. Madde kökü önemli bir soru sormalı veya cevaplayıcının çözeceği problemin ne olduğunu açıkça ifade etmelidir. Aşağıdaki madde bu bakımdan kusurludur.

34. Kan yerine kan plazması

- a) kanda mikrop bulunabileceği için kullanılır.
- b) kan grubunun tayini gerekmediğinden tercih edilir.
- c) plazma sentetik olarak hazırlanabildiği için tercih edilir.
- d) plazmada daha çok akyuvar bulunduğu için kullanılır.

Söz konusu kusur, kökü eksik cümle halindeki maddelerde de ha çok bulunur. Çünkü, eksik bir ifade ile ne sorulduğunu anlatabilmek güçtür. Madde kökünün ne sorduğu, seçenekler okunmasa bile kolayca anlaşılabilir. Yukarıdaki madde şöyle düzeltilir :

34a. Kan vermede tam kan yerine kan plazması niçin tercih edilir?

- a) Kanda mikrop bulunabileceğinden
- b) Plazmada kan grubunun tayini gerekmediğinden
- c) Plazma sentetik olarak hazırlanabildiğinden .
- d) Plazmada daha çok akyuvar bulunduğundan (b)

Ne sorulduğu madde kökünden açıkça anlaşılmayan maddeler, çoğu halde, gelişigüzel bir araya getirilmiş doğru - yanlış maddeleri gibi görünür. Böyle maddeler cevaplayıcıyı, sorulan sorunun doğru cevabını seçecek yerde, seçeneklerin doğru olup olmadığını yoklamaya ve aralarından en doğru görüneni seçmeye götürür.

2. Madde kökü, gereksiz açıklama veya bilgilerle şişirilmemelidir. Kök, probleminin ne olduğunu açıkça ifade etmeli, problemin çözümü için gerekli bilgileri en kısa şekilde vermeli, fakat bunlardan başka hiçbir şey ihtiva etmemelidir.

3. Madde kökü, maddenin mümkün olan en büyük kısmını kapsamalıdır. Özellikle kökün eksik cümle halinde ifade edildiği maddelerde, bazı ifadelerin seçeneklerde tekrar edilmesi maddeyi gereksiz olarak uzatır. Madde kökü soru kipinde yazılmış olsa bile, sorunun yazılışını değiştirerek seçenekleri kısaltmak mümkün olabilir. Örnek :

35. Bir miktar suyun sıcaklığı 0°C nin altına kadar düşürülüp, su dondurulduğunda, meydana gelen buzun

- a) hacmi sabit kalır.
- b) hacmi küçülür.
- c) hacmi büyür.
- d) hacmi küçüldüğü için yoğunluğu büyür.
- e) hacmi büyüdüğü için yoğunluğu küçülür. (e)

Bu maddede kök gereksizce uzatılmıştır; «hacmi» kelimesi köke yazılacak yerde seçeneklerde tekrar edilmiştir. Ayrıca, seçenekler birbirine benzer değildir; maddede söz konusu olan hacimden suyun hacmi mi, yoksa, buzun hacmi mi kastedildiği açıkça belirtilmiş değildir. Bu madde ile, «suyun donarken genişlemesi» yoklamak isteniyorsa, madde şöyle düzeltilebilir :

- 35a. Oda sıcaklığındaki su, donma noktasına kadar soğutulsa, hacmi
- a) değişmediği halde yoğunluğu büyür.
 - b) değişmediği halde yoğunluğu küçülür.
 - c) değişmediği için yoğunluğu sabit kalır.
 - d) küçüldüğü için yoğunluğu küçülür.
 - e) büyüdüğü için yoğunluğu küçülür. (e)

Seçeneklerdeki ifadelerin mümkün olduğu kadar büyük bir kısmının köke alınması, genellikle, daha iyi madde yazmaya imkân vermekle birlikte, bu konuda haddinden fazla gayret göstermek başka bir sakıncaya yol açar. Seçeneklerin birer kelimeye indirilmesi ve bu kelimelerin de ezberlenebilir terimlerden ibaret olması, test maddesini gereğinden fazla kolaylaştırır.

C. Seçeneklerin Yazılması

Seçmeli testlerde soru, madde kökünde sorulur, cevabı da doğru seçenekte verilir. Öyleyse, test maddesi niçin bir bilgi veya yetenek yoklar? Bir seçmeli test maddesinin ölçme gücü sadece sorulan sorunun tabiatından değil, daha ziyade, doğru cevapla birlikte verilen çeldirici seçeneklerin tabiatından gelir. Seçeneklerin yazılmasında, özellikle, çeldiricilerin yazılmasında, uyulması gerekli teknik esaslar vardır. Bu esasların başlıcaları aşağıda verilmektedir.

1. Seçenekler, ifade tarzı, uzunluk ve kapsam bakımından birbirine benzer olmalıdır. Tecrübesiz yazarlar, daha ziyade kökü ek-sik cümle yapısındaki maddelerde, doğru cevabı daha uzun yazma eğilimindedirler. Cümle yapısı bakımından yanlış seçeneklerin köke benzemeyişi veya doğru seçeneğin daha çok benzeyişi, anahtarlanmış cevabın, bilgisiz cevaplayıcılar tarafından da seçilmesini kolaylaştırır. Çeldiricilerle kök arasındaki kapsam ayrılıkları, bazı çeldiricilerin kolayca elenmesine yol açar. Seçeneklerin birbirine benzemeyişi, madde kökünde ne sorulduğu açıkça belli değilse, maddenin cevaplandırılmasını güçleştirir. Aşağıdaki madde, hem kökünde ne sorulduğu belli olmadığı, hem de seçenekleri kapsam bakımından benzer olmadığı için kusurludur.

- a) alternatif akımı doğru akıma çevirir.
- b) doğru akımı alternatif akıma çevirir.
- c) elektrik enerjisini depo eder.
- d) elektrik enerjisini nakletmekte kullanılır. (d)

2. Çeldiriciler bilgisiz veya yanlış bilgili cevaplayıcıları yanıltmalı, fakat bilgileri yeterli olanları yanıltmamalıdır. Örnek :

37. Bir bitkiye fotosentez için gerekli olan ilk maddelerin alınmasını yaprağın hangi dokusu sağlar?
- a) Üst epiderm
 - b) Pallzat dokusu
 - c) Süngerimsi doku
 - d) Alt epiderm (d)

Bu madde dikkatle incelenirse, seçeneklerin kökteki soruya kesin bir cevap vermediği görülür. Bir bitkiye fotosentez için gerekli olan CO₂ yapraktan girer, su ise yaprağa başka yoldan gelir. Bu madde «alt epiderm» dokusunun görevlerinden birini sormak için hazırlanmışsa, madde kökü bu bilgiyi dolaylı olarak sormaktadır ve tam soramamaktadır. Sınıftaki veya ders kitabındaki ifadeleri kabaca öğrenmiş olan cevaplayıcılar, anahtara göre doğru cevabı işaretleyebilirler; fakat, daha bilgili cevaplayıcılar hangi şıkkı seçeceklerine karar veremezler. Böylece, maddenin ayırma gücü negatif olabilir.

3. Kökte ve şıklarda kelime benzerliği, şık uzunluğu gibi bazı cevaplayıcıların doğru cevabı bulmakta kullanabilecekleri ipuçları varsa bunlar çeldiricilerde olmalı, doğru cevapta bulunmamalıdır. Birçok Batı kaynağında böyle ipuçlarının hiç kullanılmaması tavsiye edilir. Cevaplayıcılar böyle ipuçlarının kasten çeldiricilere konulduğunu bilmiyorlarsa, kullanılmasında sakınca yoktur. Aşağıdaki madde, kök ile bir çeldirici arasında kelime benzerliği bulunduğu için kusurludur. Bu benzerliği gören her cevaplayıcı (b) çeldiricisini eler.

38. Aşağıdaki yerlerden hangisinde Akdeniz iklimi görülür?
- a) Kaliforniya kıyıları
 - b) Akdeniz kıyıları
 - c) Güney Afrika kıyıları
 - d) Arjantin'in güney kıyıları
 - e) Avustralyanın güney - batı kıyıları (e)

4. Çeldiricilerin doğru cevaba yakınlık derecesi arttıkça maddenin güçlüğü de artar. Madde güçlüğü ise, testin kullanılış amacına göre ayarlanır. Onun için, çeldiricilerin doğru cevaba yakınlık derecesi testin güçlük derecesine göre ayarlanmalıdır. Konunun uzmanlarının bile doğru sanacağı kadar doğru cevaba yakın çeldiriciler kullanılmamalıdır. Öte yandan, bütün cevaplayıcıların eleyebileceği derecede yanlış çeldiriciler de kullanılmamalıdır. Çeldiricilerden herbiri, o test ile ölçülecek yeteneği bir dereceye kadar yoklu-

yorsa, cevaplayıcılar için, doğru cevaba en uzak çeldiriciyi elemek en kolay, doğru cevaba en yakın çeldiriciyi elemek ise en zor olacaktır. Doğru cevaba en yakın çeldiriciyi eleyebilen bir cevaplayıcının yeteneği, ya elemeyle ya da doğrudan doğruya problemi çözerek doğru cevabı bulmaya yeter. Wesman,¹² çeldiriciler yazılırken, doğru cevaba en yakın çeldiriciyi elemek için gerekli yeteneğin, doğru cevabı seçmek için gerekli yetenekten daha büyük olmamasına dikkat edilmesini tavsiye eder. Bu tavsiyenin pratikteki anlamı şudur : Hiçbir çeldirici, o maddede sorulan problemi çözebilecek yetenekteki bir cevaplayıcının seçimini çeldirecek kadar yanıltıcı olmamalıdır.

5. «Yukarıdakilerin hepsi» seçeneğinin gerek çeldirici, gerekse doğru şık olarak kullanılması özel dikkat ister. Seçeneklerden bir tanesinin yanlışlığı apacık belli ise, «hepsi» şıkkı da yanlış diye elenir. Seçeneklerden iki tanesinin doğru olduğu apaçık biliniyorsa bir tek doğru cevaplı maddelerde, «hepsi» şıkkının anahtarlanmış cevap olduğu anlaşılır. «Yukarıdakilerin hepsi» şıkkı, diğer seçeneklerin gerek ifade gerekse çeldiricilik bakımından birbirine çok benzediği hallerde kullanılmalıdır. «Hepsi» şıkkı, bütün bir testte birkaç defa ve bazen doğru, bazen yanlış seçenek yerinde kullanılmalıdır.

6. «Hiç biri» seçeneğinin kullanılması da özel dikkat ister. Bu seçenek, en doğru cevabı seçtiren maddelerde kullanılmaz. Maddenin kesinlikle doğru bir cevabı varsa ve bu cevap seçenekler arasında verilmişse, «hiç biri» çeldirici olarak verilebilir. Maddede sorulan sorunun doğru cevabı seçenekler arasında verilmemişse, «hiç biri» doğru cevap olur. «Hiç biri» seçeneğini kullanmanın iki sakıncası vardır. Birincisi, cevaplayıcılar bu seçeneği dikkate almayabilirler. Bu sakıncayı önlemek için «hiç biri» bazen çeldirici, bazen doğru cevap olarak kullanılmalıdır. İkinci sakınca «hiç biri» seçeneğinin doğru cevap olarak kullanılmasında görülür. Problemi yanlış çözen bir cevaplayıcı, o yanlış çözüm çeldiriciler arasında yoksa, «hiç biri» seçeneğini seçer ve böylece maddeyi doğru cevaplandırmış sayılır. Bu seçeneğin sayısal çözüm gerektiren maddelerde hiç kullanılmaması, problemi yanlış çözen bir cevaplayıcının, kendi çözümünü seçenekler arasında bulamayınca maddeye geri dönmesine sebep olur. Bundan istenmeyen başka test davranışları da doğabilir. Çeldiricilerin görevlerinden biri de yanlış çözümleri çekmek olduğu için «hiç biri» seçeneğinin çeldirici olarak kullanılması, çözümde doğ-

¹²Wesman, op. cit., s. 117.

ruluk ve hız yoklayan sayısal testlerde geçerliği artırılabilir. Bu tür testlerde «hiç biri» çeldiricilerde bolca kullanılmalı; yukarıda sözü edilen ikinci sakıncanın azaltılması için pek seyrek olarak da doğru cevap yerine konulmalıdır.

7. Birbirinden bağımsız olmayan, biri diğerini içine alan seçenekler yazılmamalıdır. Madde kökünde sorulan sorunun birden fazla doğru cevabı varsa veya doğru cevapta birden fazla fikir bulunması gerekiyorsa, birbirini içine alan seçenekler yazmak yerine bileşik cevap gerektiren madde formuna dönülmesi iyi olur.

8. Biri diğerine zıt düşen seçeneklerin kullanılması da güçtür. Aynı gerekçe ile, «hepsi» ve «hiç biri» seçeneklerinin ikisini birden aynı maddede kullanmaktan sakınılması gerekir. Cevaplayıcılar, birbirine zıt seçeneklerin bulunması halinde, genellikle, seçeneklerin bir kısmını kolayca elerler. Birbirine zıt şıktan biri doğru cevap olarak verilmişse, cevaplayıcılardan bazıları bu ikisinin birden yanlış olmayacağını düşünerek, seçimi, zıt şıklar arasında yaparlar. Böylece madde yalnız iki seçenekli olarak işler.

9. Çeldirici nitelikte cevap yazabilmek dikkat ve ustalık ister. Öğrencilerin diğer sınıflarda yaptığı yanlışlıklar, iyi öğrenilmemiş konular ve benzerleri çeldiriciler için fikir verebilir. Ayrıca, maddede sorulan problemin hatalı çözümleri analiz edilerek iyi çeldiriciler yazmak mümkündür. Madde yazarı, çeldirici bulmakta güçlük çekiyorsa, yanlışlığı apaçık belli olan işlemmez çeldiriciler uydurma yerine, madde formunu değiştirme yoluna gitmelidir.

10. Büyüklük sırası, sayısal sıra veya benzeri gibi anlamlı bir sıraya konulabilecek seçenekler, uygun bir sırada verilmelidir. Meselâ, sayısal seçenekler büyükten küçüğe veya küçükten büyüğe doğru sıralanmalıdır. Böylece, maddenin okunup algılanması kolaylaşır. Doğru cevabın seçenekler üzerine dağılımı rasgele olmalı, doğru cevabın yeri herhangi bir sistematik seçme ile bulunmamalıdır.

11. Seçmeli testlerde dört-beş seçenek kullanılır. Seçenek sayısı dörtten az oluncâ, doğru cevabın şansla bulunabilme ihtimali artar. Bazı maddelerde, maddenin konusu icabı ancak iki - üç anlamlı seçenek olabilir. Böyle hallerde, madde formunu değiştirerek şık sayısını artırma yoluna gidilmelidir. Zorunluk olmadıkça iki - üç şıklı maddeler kullanılmamalıdır. İlkokulun aşağı sınıflarında iki - üç, yukarı sınıflarında üç - dört şıklı maddeler kullanmak daha iyi sonuçlar verebilir.

12. Bir testin bütün maddeleri aynı sayıda seçeneğe sahip olmalıdır. Seçenek sayısı maddeden maddeye değişen testlerin cevap-

landırılmasında cevap kayıdı güçlükleri çıkabileceği gibi, puanlama ve analiz işlemlerinde de çeşitli zorluklar ortaya çıkar.

SEÇMELİ TEST MADDELERİNİN PUANLANMASI

Test maddesinin yapısı, testlerin nasıl puanlanacağını belirler. Seçmeli test maddelerinin puanlanmasında şu üç yöntemden biri kullanılabilir :

1. Madde doğru cevaplandırılmışsa 1, yanlış cevaplandırılmış veya boş bırakılmışsa 0 puan vermek.

2. Her maddeye, ölçtüğü yeteneğin önemiyle oranlı bir puan vermek. (Meselâ, kolay veya kapsamı dar maddelere 1, bu bakımdan orta nitelikte maddelere 2, çok zor veya kapsamı geniş maddelere 3 verilebilir.)

3. Maddenin seçeneklerine, doğruluk derecesine göre farklı ağırlıklar vererek puanlamak.

Bu yöntemlerden birincisi en çok kullanılan puanlama yöntemi olup, kolaylığı, hızla yapılabilmesi, test ve madde analizinin genelleştirilmiş formüllerine uygun olması sebebiyle tercih edilir.

İkinci yöntem, esas itibarıyla, maddelere geçerlikleri oranında ağırlık verilmesini savunan bir yöntemdir. Bir testin maddeleri arasında kapsam ve güçlük bakımından farklar bulunabileceği, bu farkların puan geçerliğini etkileyeceği kabul edilmekle birlikte, hem maddelere farklı ağırlıklar verme güç hem de farklı ağırlıklarla puanlama zor olduğundan, bu yöntem kullanışsızdır.

Bir tek maddenin seçeneklerine farklı ağırlıklar veren puanlama yöntemi de denenmiştir. Davis,¹³ bu yöntemin doğru cevabın en doğru olduğu, çeldiricilerin de gittikçe artan doğruluk dereçelerinde olduğu maddelerde kullanılabileceğini söyler. Böyle bir puanlamada, doğru cevaba en yakın çeldiriciyi seçen cevaplayıcının yeteneğinin, daha zayıf bir çeldiriciyi seçen cevaplayıcının yeteneğinden daha fazla olduğu kabul edilir. Davis, şu maddeyi örnek verir :

39. Bir kalite kontrolcüsü günde iki motoru kontrol edebilmektedir. Üç kontrolcünün çalıştığı bir günde, kontrol edilecek motorlar 72 motor arasından rasgele seçilse, herhangi bir motorun o gün kontrol edilme ihtimali nedir?

a) 1/36 b) 1/24 c) 1/12

d) Yukarıdakilerden başka bir oran

e) Problem, verilen bilgilerle çözülemez.

(c)

¹³F.B. Davis, *Educational Measurements and Their Interpretation*, s. 80 - 81.

Bu madde şu ağırlıklara puanlanmıştır : a : —1, b : 0, c : 1, d : —2, e : —3. Çeldiriciler incelendiğinde, bu ağırlıkların doğru cevaba yakınlık derecesiyle arttığı görülür.

Davis ve Fifer,¹⁴ bu örnekteki gibi hazırladıkları maddelerle yaptıkları bir araştırmada, seçeneklere ağırlık verilerek elde edilen puanların güvenilirliğinin bir miktar yükseldiğini, fakat test geçerliğinin önemli derecede değişmediğini görmüşlerdir. Seçmeli maddeleri seçeneklere farklı ağırlıklar vererek puanlamak, puanlama işlemlerini çok güçleştirir. Buna karşılık elde edilecek yarar, çoğu halde, çekilen külfete değmeyecek kadar küçük olur.

¹⁴F.B. Davis and G. Fifer, «The Effects on the Test Reliability and Validity of Scoring Aptitude and Achievement Test with Weights for Every Choices», Educational and Psychological Measurement, 19 (1959) : 159 - 170.

B Ö L Ü M 8

Sözlü Sınavlar

Soruların genellikle sözlü sorulduğu ve cevapların sözlü verildiği sınav çeşitlerine sözlü sınav denir. Birkaç üyeden oluşan bir komisyonun bir cevaplayıcıya sorular sorması ve sözlü cevaplara hemen not takdir etmesi suretiyle yapılan sınav, sözlü sınavların en çok baş vurulan türlerinden biridir. Öğretmenin sınıftaki öğrencilerden birini yazı tahtası önüne çağırıp birkaç soru sorması ve genellikle sözlü olarak verilen cevaplara not takdir etmesi de okullarımızda çok kullanılan bir sözlü sınav türüdür.

Sözlü sınavlar, çok eskiden beri kullanılmış olmalarına rağmen, eğitimde önemli araştırmalara konu olmamıştır. Diğer sınav çeşitleri, özellikle testler, incelenmeye başlandıktan sonra sözlü sınavların sakıncaları, diğerlerine kıyasla ortaya konmuştur.

Ülkemizde sözlü sınavlar, bütün sakıncalarına rağmen, bu yıla kadar kullanılagelmiştir. İlkokulların son sınıflarındaki bitirme sınavları daha son zamanlara kadar sözlü yapılıyordu. Eskiden «müzakere», bugün «sözlü yoklama» denilen yöntem hâlâ kullanılmaktadır.

Bu bölümde sözlü sınavların özelliklerine ve uygulanma yöntemlerine kısaca değinilmektedir.

SÖZLÜ SINAVLARIN ÖZELİKLERİ

Sözlü sınavı diğer sınav çeşitlerinden ayıran özelliklerin başlıcaları şöyle sıralanabilir :

1. Cevap sözlü olarak verilir. Bazı sözlü sınav çeşitlerinde, cevabın yazı tahtasına yazdırıldığı görülürse de, sözlü sınavların çoğunda cevapların sözlü olarak verilmesi zorunludur. Böyle olunca, sözlü ifade yeteneği, konuşmanın etkililiği, ses tonu ve benzeri değişkenler de puana karışabilir. Bu yeteneklerin ölçülmesini amaç almayan bir sınavda, puanların sözlü ifade yeteneklerinden etkilenmesi, şüphesiz, sınavın gecerliliğini düşürür.

2. Sınavı yapan kişi (ayırtman) ile sınava tabi tutulan kişi (cevaplayıcı) arasında karşılıklı ve devamlı bir etkileşim vardır. Bu etki-kileşim sınavın niteliklerini etkiler.

3. Sözlü sınavlarda bir defada ancak bir tek cevaplayıcı dinlenebilir. Bu sebeple, sözlü sınavların bireysel olarak uygulanması zorunludur. Sınav bireysel olarak uygulandığı takdirde, bir sınıf kalabalığındaki gruplarda bile, uygulama çok zaman alır. Hele, sayıları binleri bulan gruplarda gerekli zamanın uzunluğu, bu sınav türünden vazgeçilmesini zorunlu kılar.

4. Sınav bireysel olarak uygulandığından ve sınavı tamamlanmış bireylerle henüz sınava alınmamış bireyleri birbirinden ayrı tutmak zor olduğundan, her bireye ayrı sorular sorma zorunluğu doğar. Hele sınav, müzakere denilen yöntemde olduğu gibi diğer bireylerin önünde yapılıyorsa, bir bireyin kısmen veya tamamen cevaplandığı bir soruyu diğer bir bireye sormak imkânı kalmaz. Sorular daha önceden dikkatle hazırlanıp güçlük ve geçerlik dereceleri bakımından eşdeğerlilikleri sağlanmadıkça, öğrencileri verdikleri cevaplara göre karşılaştırmak mümkün olmaz.

5. Sorular, genellikle, sözlü sorulur. Böyle olunca, cevaplayıcının soruları tekrar gözden geçirmesi imkânı kalmaz. Gerçekte, bu sınav türünün esasında soruların yazılı verilmesini engelleyecek bir sebep yoktur. Fakat, soruların sözlü sorulması gelenekselleşmiştir.

6. Cevaplama, çoğu halde, üzerinde iyice düşünülüp tasarlanmadan yapılır. Buna sebep, sınavın cevaplayıcı ile ayırtmaların yüz yüze karşılıklı konuşmalarına imkân veren bir havada yapılmasıdır. İstenirse, sözlü sınavlarda da cevaplayıcılara düşünme zamanı, hatta cevaplarını tasarlayabilmeleri için kalem ve kâğıt bile verilebilir.

7. Cevaplar kaydedilmediği için, puanlama ya cevaplar dinlenirken ya da cevaplama tamamlandıktan hemen sonra yapılır. Bu durum, sözlü sınavlarda puanlamanın sadece genel izlenimle yapılmasına yol açmıştır. Cevaplar kayıtlı olmadığı için, cevapları birbiriyle karşılaştırma, bağıl yöntemlerle sıralama veya sınıflama, analitik bir anahtarla puanlama imkânları kalmaz. Puanlama genel izlenimle yapıldığı için, puanların güvenilirliği genellikle çok düşük olur.

SÖZLÜ SINAVLARIN KULLANILACAĞI DURUMLAR

Sözlü sınavların bir ölçme yöntemi olarak yazılı sınav ve benzerleri yerine bütün derslerde sık sık kullanılmasını savunanlar ol-

muştur. Bunlara göre, sözlü sınavlar şu niteliklerinden dolayı savunulur.¹ a) Sözlü sınav, ayırtman ile öğrenci arasında karşılıklı etkileşim içinde cereyan eder. Onun için, öğrenciyi sınav heyecanından ve gereksiz korkulardan kurtarmak kolay olur. b) Sözlü sınavda ayırtman, bir cevabın açıklanmasını, genişletilmesini; cevapta ileri sürülen bir düşüncenin gerekçelerle savunulmasını isteyebilir. Böylece, öğrenci daha tam bir cevap verebilir. Öğretmen de onun bilgilerini istediği derinlikte yoklayabilir. c) Yazılı sınavlarda görülen kopya, cevabı gereksizce şişirme ve başka yana çekme olayları sözlü sınavlarda kolayca önlenabilir. d) Sözlü sınav, öğretmen ve öğrenci arasında karşılıklı etkileşim yarattığı için, sınav tecrübeleriyle öğrenmeye daha fazla imkân verir. e) Sözlü sınav sırasında öğrencinin yazılı sınavlarla ölçülemeyen genel görünüş, sözlü ifade yeteneği, konuşma etkililiği, v.b. gibi kişisel nitelikleri de gözlenip ölçülebilir.

Yukarıda sayılan beş madde, sözlü sınavların yararları veya üstünlükleri olarak değil, sakıncaları olarak da belirtilebilirdi. Birinci maddede sözü edilen ayırtman ve öğrenci etkileşimi sınavın zararına da işleyebilir. Bir öğrenciye yol gösterip konuyu doğru cevaplandırmasına yardım etmek mümkün olduğu gibi, kasıtlı veya kasıtsız müdahalelerle şaşırtmak da mümkündür.

İkinci maddede ileri sürülen nitelik, aslında, soruyu yeni sorularla derinleştirme imkânından doğar. Önceden her soru dikkatle hazırlanır ve her soruda ne kadar derinliğe inileceği açıkça belirtilirse, cevaplama sırasında ek sorular sorulmasına hacet kalmaz. Onun için, ayırtmanın öğrenciye cevaplama sırasında soru sorabilmesi, bu sınav çeşidinin önemli bir üstünlüğü sayılamaz.

Üçüncü maddede sözü edilen kopya, cevabı şişirme ve başka konuya çekme yazılı yoklamalarda da bazı tedbirlerle önlenabilir. Bu gerekçe de, yazılı yoklamalar yerine sözlü sınav konulmasını zorunlu kılmaz.

Dördüncü maddede ileri sürülen gerekçeyle sadece sözlü sınavlar değil, hiçbir sınav yapılmamalıdır. Çünkü sınavlar öğretim amacıyla değil, ölçme ve değerlendirme amacıyla yapılır. Bir öğretmenin amacı bazı konuları yeniden öğretmek veya pekiştirmek ise, bu amaca uygun öğretim faaliyetleri düzenlemesi daha doğru olur.

Sözlü sınav sırasında öğrenciyle ayırtmanın yüz yüze olmasının, ayırtmana, öğrencinin bir takım kişisel özelliklerini gözleme fırsatı

¹R.L. Ebel, *Essentials of Educational Measurement*, s. 266 - 268.

verdiği doğrudur. Fakat, sınav bu niteliklerin ölçülmesi amacıyla değil, öğrencinin belli bir konudaki bilgi ve becerilerini ölçmek amacıyla yapılır. Böyle kişisel niteliklerin, sınavda ölçülecek hedef davranışlarla karıştırılması, hem bilgilere hem de kişisel niteliklere bir puan takdir edilmesi, puanların geçerliğini düşürür.

Bütün sakıncalarına rağmen, sözlü sınavları kullanmanın zorunlu olduğu hallerle karşılaşılabilir. Sözlü ifade becerilerini geliştirmek, genel eğitim veren programların hepsinde önemli bir amaç kabul edilmiştir. Okullarımızda sözlü ifade becerileri, genellikle, Türkçe ve yabancı dil derslerinin hedefleri arasında sayılır. Kelimeleri doğru telâffuz etmek, cümlelerde kelimeleri doğru vurgulamak, düzgün ve etkili bir tarzda konuşmak, bir topluluğa hitap edebilmek ve benzerleri, ana dili dersinin geliştireceği davranışlardır. Fakat, diğer derslerin de kendilerine özgü dilleri, o konulara özgü sözlü ifade biçimleri vardır. Meselâ, matematikte terimlerin özel sembolleri ve bunların söyleniş biçimleri vardır. İşte bu tür özel sözlü ifadeler, o derste ilgili konularla birlikte öğrenilir. Ayrıca, yabancı dil derslerinde de o dilde sözlü ifade becerileri önemli hedeflerdendir.

Sözlü ifade becerileri, ister ana dilinde, ister bir yabancı dilde, isterse bir özel alanda olsun, okulda geliştirilmeye değer davranışlar olduğu için, onların ölçülüp değerlendirilmesi de gereklidir. Sözlü ifade becerilerini gözleyip ölçmenin en uygun yöntemi de, şüphesiz, öğrencinin sözlü ifadeyi kullandığı bir sözlü sınav olur.

Öte yandan, sözlü ifade, diğer alanlardaki bilgi ve becerilerin ölçülmesinde bir iletişim aracı olarak da kullanılabilir. Bu halde sınavın amacı sözlü ifade becerilerini ölçmek değil, sözlü ifade aracılığı ile diğer bir alandaki davranışları ölçmektir. Sözü ifade becerilerinin ölçülmesinde sözlü sınav biricik etkili yöntem olduğu halde, diğer davranışları ölçmek için çeşitli yöntemler vardır. Onun için, önemli bir gereklilik olmadıkça, sözlü sınavlar, diğer alanlardaki bilgi ve becerileri ölçmek için kullanılmamalıdır. Sözlü sınav uygulama zorunluğu, bazen, öğrenci azlığı, sınav süresinin çok kısa oluşu gibi sebeplerden de doğar. Bazen de, öğrencinin sakatlığı gibi haller sözlü sınavları gerekli kılar.

SÖZLÜ SINAVLAR NASIL YAPILMALI?

Memleketimizde sözlü sınavların başlıca iki türü uygulanmaktadır. «Müzakere» veya «sözlü yoklama» adıyla imtihan yönetmeliklerimize girmiş olan yöntemde, öğretmen, diğer öğrenciler karşısın-

da bir tek öğrenciye sözlü sorular sorar. Öğrenci sözlü cevaplar verir. Öğretmen bu cevaplara bir puan takdir eder. Diğer bir yöntemde ise imtihan komisyonu iki - üç öğretmenden oluşur; bir tek öğrenci, komisyon önünde soruları cevaplandırır. Öğrencinin cevapları her üç üyenin anlaşmaları halinde bir tek puanla puanlanır; anlaşma olmazsa üyelerin ayrı ayrı takdir ettikleri puanların ortalaması alınır.

Sözlü yoklama denilen yöntem okullarımızda gittikçe daha seyrek kullanılır hale gelmektedir. Bir komisyon önünde yapılan sözlü sınavlar da hemen hemen hiç kullanılmaz olmuştur. Çünkü, genellikle bitirme ve bütünleme sınavlarında kullanılan bu yöntemin yerine, yine komisyonlarca yapılan yazılı sınavlar uygulanmaktadır. Gerçekte, gerek bir öğretmenin sınıfta uygulayacağı, gerekse bir komisyonca yapılan sözlü sınavlara bazı özel durumlarda ihtiyaç duyulur. Bu sebeple, sözlü sınavların, mümkün olduğu kadar az hata ile yapılması için, aşağıda bazı tavsiyeler verilmektedir.

1. Yazılı yoklamalarda olduğu gibi, sözlü sınavda sorulacak sorular önceden hazırlanmalıdır. Sorular önceden hazırlanmazsa, soru gücünü ayarlamak, sınavın amacına uygun soru seçmek gibi sınavın geçerliğini artırıcı tedbirler alınamaz.

2. Soruların hazırlanmasında, yazılı yoklamalardaki gibi, açıklığa, kolay anlaşılma, kesinliğe ve benzeri niteliklere dikkat edilmelidir. Sözlü sınav bir alandaki bilgileri yoklamak amacıyla yapılıyorsa, istenilen bilgilerin ortaya çıkarılmasını kolaylaştıracak biçimde sorular sorulmalıdır. Sözlü sınav sözlü ifade becerilerinin gözlenip ölçülmesi amacıyla yapılıyorsa, öğrenciye üzerinde kolayca konuşabileceği bir konu verilmeli veya konuşmayı sağlayacak sorular sorulmalıdır.

3. Cevapların puanlanması için önceden puanlama anahtarları hazırlanmalıdır. Sınav sözlü ifade becerilerini gözleyip ölçmek için yapılıyorsa, kompozisyonları puanlamada tavsiye edilen anahtara benzer bir analitik anahtar hazırlanmalıdır. (Bakınız : Bölüm 4, Tablo : 4.2).

4. Sınavı bir komisyon yapacaksa, soruların ve cevap anahtarlarının komisyonca hazırlanması; analitik anahtar başka bir kaynaktan alınmışsa, komisyon üyelerinin anahtarı kullanmada yetiştirilmeleri yararlı olur.

5. Öğrencilere, soruları cevaplandırırken uyacakları genel esaslar, cevapları düşünüp tasarlama yolları, konuşmada dikkat edi-

lecek noktalar ve benzerleri önceden öğretilmelidir. Böylece, öğrenci sözlü sınava hazırlanmış olacağından, sınav heyecanı ile cevaplamının güçleşmesi, kötü konuşma ve yetersiz tasarlama ile cevapların bazulması, vb. ihtimalleri azaltılabilir.

6. Ölçülecek yetenek ve beceriler diğer sınav türleriyle ölçülebiliyorsa sözlü sınavlardan kaçınılmalıdır. Bu tavsiyeye uyulursa, sözlü sınavların sözlü dil becerilerini amaç edinen ana dili ve yabancı dil derslerinde daha sık, diğer derslerde ise pek seyrek yapılması gerekir. Matematik gibi bir derse özgü sözlü ifade becerilerinin ölçülmesi için, her öğrencinin bir ders yılında bir defa beş - on dakikalık bir sözlü tebliğde dinlenmesi yeter.

SÖZLÜ SINAVLARIN DİĞER ÇEŞİTLERİ

Sözlü sınavların zahmetli, zaman alıcı ve pahalı oluşu ile puanlanmasının yeterince güvenilir olmayışı, bu sınav yönteminde değişiklik yapılmasına yol açmıştır. Bu değişikliklerden bazıları aşağıda kısaca tartışılmaktadır.

Sözlü yoklama denilen yöntemde olduğu gibi, bir ders saatında ve sınıf huzurunda yapılan sözlü sınavın en büyük sakıncası, sınav sırasında diğer öğrencilerin pasif veya seyirci kalmalarıdır. Bu sakıncayı önlemek için, okullarımızda kullanılan müzakere yöntemi bir sınav olarak değil, bir tekrar yöntemi olarak yönetmeliklere girmiştir. Öğretmenin sorduğu soruyu bir öğrenci cevaplandırırken, diğer öğrencilerin cevabı dinlemeleri ve böylelikle öğrendiklerini pekiştirmeleri umulur. Gerçekte bu yöntem etkili bir tekrara da imkân vermez. Üstelik, diğer öğrencilerin tekrara katılıp katılmadıkları da kontrol edilemez. Çoğu halde diğer öğrenciler, aktif tekrarlayıcı veya cevapları eleştirici değil, pasif seyirci rolünde kalırlar.

Diğer öğrencileri daha aktif hale getirmek için önerilen bir yöntemde, öğrenciler soru hazırlarlar; öğretmen herhangi bir öğrenciyi soru sormaya, diğerlerini de bu soruyu cevaplandırmaya çağırır. Öğretmen ancak yönetici ve puanlayıcı rolünde kalır. Hatta, puan takdirini de öğrencilere bırakan bir yöntem bile tavsiye edilmiştir.²¹² Böyle bir değişiklikle diğer öğrenciler daha aktif hale getirilebilir; fakat, sınavın geçerliği ve güvenilirliği büyük ölçüde zedelenir. Teklif

²¹²W. Brody and N.J. Powell, «A New Approach to Oral Testings», *Educational and Psychological Measurement*, 7 (1947) : 289 - 298.

²¹³J.R. Hartnett, «Oral Examinations», *Improving College and University Teaching*, 13 : 208 - 209.

edilen yöntem, öğretmen yerine öğrencileri koyduğu için daha kötü bir sınav yönetimi getirmektedir.

Aynı sakıncadan hareket edilerek ileri sürülen bir yöntemde, öğrencilerin küçük gruplar halinde veya bütün sınıf olarak, belli bir konuda «grup tartışması» yapmaları; öğretmenin de bu tartışmalar sırasında öğrencileri gözleyerek bilgilerine, konuşma becerilerine ve diğer kişisel niteliklerine puan vermesi önerilmiştir. Bu değişiklik, grup tartışmalarıyla ilgili bazı davranışların gözlenip ölçülmesi için uygun bir yöntem getirebilir; fakat, bilgilerin veya diğer becerilerin ölçülmesi için daha iyi bir yöntem getirmiş olmaz.

Komisyonlar önünde yapılan sözlü sınavlarda, zaman kazanmak veya her öğrenciye daha fazla soru sorabilmek için komisyon üyelerinin öğrencileri ayrı iki komisyon halinde ve birbirlerinden bağımsız olarak sınava almaları da tavsiye edilmiştir.⁴ Aslında, birden fazla puanlayıcı kullanıldığında puanlama güvenilirliğinin yükselmesi, puanlayıcıların birbirlerinin etkisinde kalmadan puan vermeleriyle mümkündür. Bir öğrenci her iki komisyondan geçmezse, puanlama güvenilirliği artırılmayacağı gibi, başka sakıncalar da ortaya çıkabilir.

İki komisyon yerine, komisyon üyelerinin sınavı ayrı odalarda kendi başlarına sürdürmeleri de düşünülebilir. Söz gelişi, üç üye varsa, sınava dahil konular üçe bölünür; her üye, her öğrenciye kendine ayrılan bölümden sorular sorar. Her öğrenci, sıra ile bu sınav yerlerinden geçer. Üç üyenin takdîr ettiği puanların ortalaması, öğrencinin sınav puanı olur. Böyle bir yöntem sınav süresini kısaltır veya belli bir sürede öğrencilere daha fazla sayıda soru sorulmasına imkân verir. Soru sayısı artırıldığı için sınav puanlarının güvenilirliği ve geçerliği yükselir. Fakat, her üye ancak kendi sorduğu soruların cevaplarını puanladığından, puanlama güvenilirliği, aynı cevaba üç üyenin bağımsız puan takdir ettiği haldeki kadar yüksek olmaz.

⁴Ebel, op. cit., s. 267.

Hedef Analizi ve Test Maddeleri

Öğretmenin öğretimdeki amacı, öğrencide bazı davranışların oluşmasını sağlamaktır; ölçmedeki amacı ise, hedef alınan bu davranışların ne ölçüde oluştuğunu meydana çıkarmaktır. Ölçme açısından önemli bir iş, öğrencide istendik davranışlar oluşmuşsa, bu davranışların hangi durumlarda gözlenip, nasıl ölçülebileceğini belirlemektir. Öğretmen istendik öğrenci davranışlarını, çoğu halde, sınav sorularıyla ölçer. O halde, sınavlarda sorulan sorular, kullanılan testler, v.b. bir bakıma, hedef davranışların ne ölçüde oluştuğunu meydana çıkarmaya yarayan araçlardır. Bir sınavın bu amaca hizmet edebilmesi için, sınavdaki her soru öğretimin hedef aldığı davranışlardan en az birini yoklamalı, onu başka davranışlarla karıştırmamalıdır. Bu nitelikte sorular hazırlamak, hem hedef davranışların hem de soru yazma tekniklerinin iyi bilinmesini gerektirir.

Bu bölümde, eğitim hedeflerinin nasıl analiz edildiği açıklanmakta; hedefler analiz edilmek suretiyle, hedef davranışların nasıl tanımlandığı örneklerle gösterilmektedir.

BİR SINAV SORUSU İLE NE ÖLÇÜLÜR?

Öğretmenler sınavlarda çok çeşitli sorular sorarlar. Soruların konu bakımından çeşitli olması yanında, güçlük bakımından da farklı olduğunu tecrübelerimizle biliriz. Sorular ve test maddeleri, cevaplayıcıdan istedikleri davranışlar bakımından da farklılık gösterir. Bazı sorular sadece bir bilginin hatırlanmasıyla cevaplandırılabilir. Bazı sorular ise cevaplayıcıda, kritik düşünme, problem çözme, çözümünü ve cevabı kompoze edebilme gibi zihin davranışları gerektirir. Bir sorunun ne ölçeceği, hazırlanması sırasında kolaylıkla kestirilemez. Görünüşte kolay bazı soruların da birçok öğrenci tarafından gerektiği gibi cevaplandırılmadığını hayretle görmüş olan öğretmenler çoktur.

Bir Amerikan eğitimcisi, Iowa okullarında fizik okuyan 325 öğrenciye uyguladığı bir fizik başarı testine aşağıdaki dört soruyu da-

hil etmiş ve puanlamadan sonra doğru cevap yüzdelerini hesaplamıştır.¹

1. Buzun erime ısısı kaç kaloridir?
Doğru cevap : 80 kalori. Doğru cevap verenler % 75.
2. 0°C de 1 gr buz eritmek için kaç kalori gerekir?
Doğru cevap : 80 kalori. Doğru cevap verenler % 70.
3. Buzun erime ısını tanımlayınız?
Doğru cevap : 0°C de 1 gr buz 0°C de su haline geçirmek için gerekli ısı miktarı. Doğru cevap verenler : % 50.
4. Bir kaptaki su 0°C ye kadar soğuduğunda 800 kalori veriyor. Bu kaptaki suya, soğumadan önce 0°C de 800 gr buz konsaydı buzun ne kadarı erirdi?
a) Hepsi
b) 10 gramı
c) Hemen hemen hepsi
d) 1 gr ile 2 gr arasında bir miktarı
Doğru cevap : b. Doğru cevap verenler : % 35.

Bu soruların incelenmesinde açıkça görülüyor ki, ilk üç soru sınıfta öğretilen bilgilerin hatırlanmasını ve kavranmış olmasını gerektirmektedir. Dördüncü soru, bu yeteneklerin yanında, problem çözme yeteneğini de gerektirir. Bu sorulara doğru cevap verenlerin yüzdeleri de ilginç sonuçlar vermektedir. Birinci soruya, sınava giren 325 öğrencinin % 75'i doğru cevap vermiştir. Bu soru basit bir hatırlama gerektirir. Aynı bilgi, ikinci soruda biraz değişik bir ifade ile sorulmuştur. Bu küçük değişiklik bile doğru cevap verenler yüzdesini % 75'den % 70'e düşürmüştür. Üçüncü soruda «80 kalori» değil, onu tanımlayan ifade sorulmuştur. Bu bilgi ise öğrencilere ancak yarısı tarafından hatırlanabilmektedir. Dördüncü soruda buzun erime ısısına ilişkin ilkelerin basit bir probleme uygulanması istenmektedir. Bu uygulamada gerekli matematiksel işlem lise düzeyi için çok kolaydır, istenen sonuç çabucak zihnen hesaplanabilir. Buna rağmen, probleme doğru cevap verenler % 35'e düşmüştür. Lindquist'in bu araştırmasından çıkarılabilecek kısa sonuç şudur : Bir sorunun ne ölçebileceği, onun nasıl sorulduğuna da bağlıdır.

Şüphesiz, bu çalışmada sözü edilmeyen diğer bir etken, öğrencilerin öğrenme düzeyleridir. Erime ısısı gibi bir kavramı çok kullanmış veya iyice pekiştirmiş bir öğrenci grubunda, doğru cevap yüzdeleri yukarıdakilerinden büyük olurdu. O halde, bir sorunun ne

¹H.E. Howkes, E.F. Lindquist, C.R. Mann (Eds.), *The Construction and Use of Achievement Examinations*, s. 91 - 93.

gibi yetenekler ölçtüğü, sorusunun yapısı yanında, o soruyu cevaplandıranların yetişkinliklerine de bağlı olur. Bir programa bağlı olarak yapılan öğretimde, öğrencinin başarısı yoklanacaksa, öğrencilerden, neleri ne kadar öğrenmeleri beklendiğinin de bilinmesi gerekir. Bir derste nelerin öğrenileceğini ve ne derinlikte öğrenileceğini, müfredat programlarında gösterilen hedefler tayin eder.

PROGRAMLARDA HEDEFLER

Programlar hazırlanırken hedeflerin nasıl saptanacağı, öğretimi başlamadan önce hedeflerin nasıl analiz edileceği ve benzerleri, program geliştirme alanında incelenen konulardır. Gerek hedeflerin mahiyeti, gerek hazırlanışları, gerekse analiz edilışleri, program geliştirme alanındaki yayınların birçoğunda ayrıntılarıyla açıklanır. (Bakınız : Tyler², Lindwall³, Popham ve ark.⁴, Krathwohl ve Payne⁵, Ertürk⁶). Öte yandan, çoğu halde hedef analizinin hiç yapılmadığı, hedeflerin özel hedefler halinde açık seçik ifade edilmediği görülür. Durum böyle olunca, hedefleri analiz etmek de, çoğu zaman, ölçme araçlarını hazırlayacak kimseye düşer. Öğretmenin kendi hazırlayacağı araçlar söz konusu olunca da, şüphesiz, bu iş o dersin öğretmenine düşer. Onun için, öğretmen, hedeflerin mahiyetini hiç değilse ana çizgileriyle bilmelidir. Bu kesimde, hedeflerin özellikleri kısaca tanıtılmaktadır.

Eğitimde hedefler özellikleri bakımından birkaç boyutta sıralanabilir.^{7,8} Bu boyutlardan birincisi, «genellik - sınırlılık» boyutudur. Bir hedefin genel ifadelerle verilmiş olması hem anlamında bulanıklık yaratabilir, hem de o hedefin davranışlarla tanımlanmasını güçleştirir. Bunun zıddına, bir hedefin gereğinden fazla sınırlandırılmış olması, aksi yönde bir tehlike yaratır. Bu halde hedef çok sınırlı davranışlara iner; hedefe dahil olması gereken diğer davranışlar da dışarıda kalır. Ölçmede işe koşulabilme açısından, çok genel olan hedefler test maddeleri için belirli davranışlar göstermez; çok sınırlı olan hedefler ise ölçme alanını gereğinden fazla daraltır.

²R.W. Tyler, *Basic Principles of Curriculum Development and Instruction*.

³C.M. Lindvall (Ed.), *Defining Educational Objectives*.

⁴W.J. Popham, *Establishing Instructional Goals*.

⁵D.R. Krathwohl and A.D. Payne, «Defining and Assessing Educational Objectives» in, R.L. Thorndike (Ed.), *Educational Measurement*, 2nd ed., s. 17 - 45.

⁶S. Ertürk, *Eğitimde Program Geliştirme*.

⁷*Ibid*, s. 53 - 54.

⁸N.E. Gronlund, *Measurement and Evaluation in Teaching*, s. 20 - 58.

Hedeflerin nitelenebileceği ikinci boyut, ifadesinin açıkça gözlenebilir bir davranışa çevrilebilir olması veya somut davranışa çevrilememesi boyutudur. «Dört aritmetik işlemin bilgileri», genel olmakla birlikte, gözlenebilir davranışlara çevrilebilecek bir hedeftir. Oysa ki, «matematiğin insan hayatındaki rolünü takdir edebilme», o kadar kolayca davranışa çevrilebilir bir hedef değildir. Bir hedefin ölçmede kullanılabilir olması için davranışa çevrilebilecek nitelikte ifade edilmiş olması gerekir. Fakat, bu tavsiyeden, «kolayca ölçülemeyecek nitelikler hedef alınmamalıdır» yargısına varılmamalıdır. Öğrenci «matematiğin hayattaki rolünü takdir eder», inancıyla bir öğretim faaliyetine sokulmuşsa, onun, matematiğin hayattaki rolünü takdir edebildiğini gösterir davranışları bulunmalı; öğretimin sonunda bu davranışların gelişip gelişmediği araştırılmalıdır.

Nihayet, program geliştirme uzmanları hedef ile konu arasındaki ilişkinin ne olabileceğini de tartışırlar. Hedefler, davranışlarla tanımlanırlar. Davranışlar da muhakkak bir konu kapsamında gelişir. Onun için, hedefler dersin konularıyla «kenetlidir.»⁹ Bir programı geliştirirken, hedefsiz konular sıralamak ne derece eksik kalırsa, konusuz hedef saptamak da o kadar eksik kalır. «Pisagor teoremi» gibi konu başlıkları sıralanmakla, bu konuların hangi davranış basamağına kadar öğrenileceği saptanmış olmaz. Aynı şekilde, «ilkelere yeni durumlara uygulayabilme» gibi bir hedef de, hangi ilkenin ne gibi durumlara uygulanacağını açıklamaz. «Pisagor teoremini yeni problemlere uygulayabilme» ise, konusu ve davranış basamağı saptanmış bir hedeftir. Hedefler böyle saptanırsa, her bir hedefi onun kritik davranışlarıyla tanımlamak; sonra kritik davranışların gözlenmesine imkân verecek ölçme durumları yaratmak kolay olur.

Öğretmen hedefler konusunda kararsızlığa düşebilir veya kendisini yetersiz hissedebilir. Böyle hallerde, programın «öğretmen kılavuzu» gibi yardımcı kaynaklarından yararlanılabilir. Eğer varsa, program geliştirme, ölçme - değerlendirme ve konu alanı uzmanlarına baş vurulabilir.

EĞİTİM HEDEFLERİNİN SINIFLANDIRILMASI

Öğretimin öğrencide bazı istendik davranışlar meydana getirmek amacıyla yapıldığı yukarıda belirtilmişti. İstendik davranışlar nelerdir? İşte bu soruya cevap vermek, bir yandan program geliştirme uzmanlarının, bir yandan da ölçme ve değerlendirme ile uğ-

⁹Ertürk, op. cit., s. 54.

raşan kişilerin işi olmuştur. Öğretmen de doğrudan doğruya her iki yönden bu soruyla ilgilenir.

Eğitimin amaçları ve hedef alınan öğrenci davranışları çeşitli okullarda, bir okulun sınıflarında ve bir sınıfta okutulan çeşitli derslerde ayrı ayrı saptanır. Belli bir öğrenim basamağında sadece bir dersin amaçları çeşitli eğitim sistemlerinde gözden geçirilirse, bunların arasında ayrılıklar olduğu; amaçların genel ifadelerle verildiği, hedef davranışlar halinde ifade edilmediği görülür. Oysa ki, hedeflerin anlamda şüphe yaratmayacak kesinlikte ifade edilmiş olması, geliştirilmesi planlanan öğrenci davranışlarını gösterecek derecede ayrıntılı bir analize tabi tutulması istenir. İşte bunun için, müfredat programlarında derslerin hedefleri düzenli bir biçimde ifade edilmeye ve sınıflandırılmaya gayret edilmiştir. İlkokul basamağından orta öğretimin ikinci devresine kadar, çeşitli derslerin hedefleri, Bloom, Hastings ve Madaus¹⁰ tarafından ayrıntılarıyla verilmiştir.

Derslerin hedeflerinin özel olarak saptanması yanında, bütün derslere uygun düşecek genel bir hedefler sınıflaması arama gayretleri de görülmüştür. Bu gayretler sonunda, «aşamalı sınıflama» deyimiyle de bilinen ve çok rağbet gören bir sınıflama ortaya çıkmıştır.^{11,12,13} Bu sınıflamanın genel alanlar itibariyle başlıca özellikleri şu noktalarda özetlenir :

1. Eğitimin sonucu olarak gelişebilecek davranışlar önce üç alana ayrılmıştır. **Bilişsel alan** davranışları bilgilerle ve bilgilerden doğan zihin yetenekleriyle ilgilidir. Bu alandaki davranışlar kendi içerisinde altı basamağa ve her basamakta alt sınıflara ayrılmıştır. **Duygusal alan** davranışları insanın geliştirdiği duygu ve değerlerle ilgilidir. Bu alandaki davranışlar, kendi içerisinde beş basamağa ve her basamak da alt sınıflara ayrılmıştır. **Psiko - motor alan** davranışları insanın kas ve zihin koordinasyonu ile yaptığı becerilerle ve onların temel öğeleriyle ilgilidir. Bu alanın davranışları, kendi içerisinde yedi basamağa ve her basamak da alt sınıflara ayrılmıştır.

2. Davranışlar, gerek bir alanın basamaklarında, gerekse bir basamakta aşamalı olarak düzenlenmiştir. Herhangi bir basamak-

¹⁰B.S. Bloom, J.T. Hastings, G.F. Madaus (Eds), *Handbook on Formative and Summative Evaluation of Student Learning*.

¹¹B.S. Bloom (Ed.), *Taxonomy of Educational Objectives, Handbook I : Cognitive Domain*.

¹²D.R. Krathwohl, B.S. Bloom, B.B. Masia, *Taxonomy of Educational Objectives, Handbook II : Affective Domain*.

¹³Ertürk, op cit., s. 63 - 72.

taki davranışın oluşabilmesi için, ondan önce gelen basamaklardaki davranışların öğrenilmiş ve geliştirilmiş olması gereklidir.

3. Davranışlar, her alanda üç haneli sayısal kodlarla gösterilmiştir. Noktadan önceki sayı, davranış basamağını; noktadan sonrakiler ise, basamak içindeki birinci ve ikinci derecedeki sınıflamaları gösterir.

Bilişsel alana dahil hedefler, kodlarıyla birlikte Tablo 9.1,de gösterilmiştir. Duyuşsal alan sınıflamasıyla, psiko - motor alan sınıflaması, sırasıyla Tablo 9.2 ve Tablo 9.3'de gösterilmiştir (Son iki tablo kısaltılarak Ertürk'ten alınmıştır.¹⁴)

Aşamalı sınıflamaların kullanılmasında şu noktayı göz önünde bulundurmakta yarar vardır. Bu sınıflamanın, bir zamanların «zihin fakülteleri» diye adlandırılan psikolojik sınıflama ile ilgisi yoktur. Bir çok kimse bu sınıflamayı, başlangıçta, «fakülte» psikolojisine, yani İnsan zihninin «bellek», «zekâ», v.b. gibi bağımsız kısımlardan oluştuğu görüşüne dönüş gibi görür. Fakülte psikolojisinin kavramları, çoğu halde gözlenemeyen kuramsal yapılardır. Oysa ki, aşamalı sınıflamada sınıflar insanın gözlenebilir davranışlarıdır. Ölçme tekniği açısından bu sınıflamanın üstünlüklerinden biri, tanımlanan davranışların gözlenebilir olmasından gelir.

Bilişsel alanın bilgi (1.00) basamağında, bir alandaki bilgilerin ve o bilgilerle uğraşma araç ve yöntemlerinin gerektiğinde hatırlanabilecek kadar öğrenilmesi hedef alınır. Kavrama (2.00) basamağında bilgilerin hatırlanmasından ötede, başka ifade biçimlerinde tanınması, bir sembolden diğerine çevirilmesi ve yorumlanması gibi davranışlar vardır. Uygulama (3.00) basamağında bilgilerin yeni durumlara uygulanması amaç edinilmiştir. Analiz (4.00) basamağında bir bilgi bütününde öğelerin ve öğeler arası ilişkilerin ortaya çıkarılması; sentez (5.00) basamağında kapsamı ve iç düzeni bakımından yeni bir bilişsel ürün meydana getirme gibi davranışlar hedef alınmıştır. Değerlendirme (6.00) basamağında bir bilişsel ürün hakkında iç veya dış ölçütlerle değer yargısına varma davranışları yer alır. Burada görüldüğü gibi, 1.00 kodu altında bilgiler sınıflanmıştır; 2.00 ve daha yukarı kodlarda, kazanılan bilgiler vasıtasıyla insanın edineceği zihin yetenekleri sınıflanmıştır.

¹⁴Ibid, s. 67 - 76.

TABLO 9.1
BİLİŞSEL DAVRANISLAR

1.00 Bilgi

1.00 Bir alana özgü bilgiler

1.11 Terimlerin bilgisi

1.12 Olguların bilgisi

1.20 Bir alana özgü bilgilerle uğraşmanın araç ve yolları

1.21 Alışların (teamül) bilgisi

1.22 Yönelim, sıra ve dizilerin bilgisi

1.23 Sınıflar ve kategorilerin bilgisi

1.24 Ölçütlerin bilgisi

1.25 Metotların bilgisi

1.30 Bir alandaki evrensel öğelerin ve soyutlamaların bilgisi

1.31 İlke ve genellemelerin bilgisi

1.32 Kuram ve yapıların bilgisi

2.00 Kavranma

2.10 Çevirme (tercüme)

2.20 Yorumlama(tefsir)

2.30 Öteleme (ekstrapolasyon ve enterpolasyon)

3.00 Uygulama

4.00 Analiz

4.10 Öğelere dönük analiz

4.20 İlişkilere dönük analiz

4.30 Örgütlenme ilkelerine dönük analiz

5.00 Sentez

5.10 Özdeşsiz bir iletişim muhtevası meydana getirme

5.20 Bir plan veya işlemler takımı önerisi meydana getirme

5.30 Bir soyut ilişkiler takımı geliştirme

6.00 Değerlendirme

6.10 İç ölçütlerle değer yargısına varma

6.20 Dış ölçütlerle değer yargısına varma

TABLO 9.2
DUYUŞSAL DAVRANIŞLAR

- 1.0 Alma (algılama)
 - 1.1 Farkındalık
 - 1.2 Almaya açıklık
 - 1.3 Tercihli dönüklük
 - 2.0 Algıya karşılık verme (mukabele)
 - 2.1 Mukabelede uysallık
 - 2.2 Mukabeleye isteklilik
 - 2.3 Mukabelede tatmin
 - 3.0 Değer verme
 - 3.1 Bir değerli kabullenme
 - 3.2 Bir değere düşkünlük
 - 3.3 Bir değere adanma
 - 4.0 Değerde örgütlenme
 - 4.1 Bir değerlin kavramsallaştırılması
 - 4.2 Bir değerler sistemi örgütlenme
 - 5.0 Değerler tümgesince nitelenmişlik
 - 5.1 Genellendik kuruluş
 - 5.2 Karakterlenme
-

Duyuşsal alan davranışlarının sınıflandırılmasında, psikometride incelenmesi gelenekleşmiş olan «ilgiler», «tutumlar», «kişilik ve sosyal uyum» gibi konulardan geniş ölçüde farklı kavramlar getirilmiştir. Burada sözü edilen davranışların daha temelde oluşları bir yana, bu sınıflama yenice kullanılmaya konulduğundan, birçok eğitimci bu davranışları henüz tam anlamıyla kavramış değildir. On ikinci bölümde görüleceği gibi, en modern programlarda bile, duyuşsal alan hedeflerinin birçokları, geleneksel programlardakilere benzer olarak, «tutumlar», «ilgiler», «değerler» halinde ifade edilmiş; ölçme teknikleri de psikometrinin bu alanlarından alınmıştır. Oysa ki, Tablo 9.2'de verilen analiz herhangi bir dersin hedeflerine uygulanabilecek kadar temele inmiş ve genelleştirilmiştir. Psiko-motor davranışlar da geleneksel olarak, burada olduğu gibi bir becerinin temelindeki davranışlar halinde değil, daha karmaşık beceriler halinde programlara girmiştir. Ölçme işlemleri de, bir becerinin temelindeki davranışlara değil, daha ileri basamakta, becerilerin mükemmelliğini ortaya çıkarmaya yönelir.

TABLO 9.3
PSİKO - MOTOR DAVRANIŞLAR

- 1.00 Algılama
 - 1.10 Duyuşsal uyarılma
 - 1.11 İşltme ile uyarılma
 - 1.12 Görme ile uyarılma
 - 1.13 Dokunma ile uyarılma
 - 1.14 Tutma ile uyarılma
 - 1.15 Koklama ile uyarılma
 - 1.16 Devrimsel duyu ile uyarılma
 - 1.20 İşaret seçme
 - 1.30 Çevirme
 - 2.00 Kuruluş
 - 2.10 Zihinsel kuruluş
 - 2.20 Bedensel kuruluş
 - 2.30 Duygusal kuruluş
 - 3.00 Kılavuzlanmış faaliyet
 - 3.10 Taklit
 - 3.20 Deneme - yanılma
 - 4.00 Mekanizma
 - 5.00 Karmaşık dışa - vuruk faaliyet
 - 5.10 Kararsızlığın giderilmesi
 - 5.20 Otomatik icra
 - 6.00 Uyum
 - 7.00 Yaratma
-

BELİRTKE TABLOSU

Müfredat programlarımızda hedeflerden «amaçlar» diye söz edildiği, amaçların daha ziyade genel ifadeler halinde verildiği, okullarımıza konmuş derslerden birkaçının programını incelenmekle, kolayca görülebilir. Programlarımızın diğer bir özeliği, o derste işlenecek konuları sadece başlıklarıyla listelemeleridir. Bir programın ölçme ve değerlendirmede işe yarayabilmesi için, hem hedeflerinin analiz edilmiş olması, hem de konularının ayrıntılarıyla verilmesi gerekir. Öğretmenlerimize hedef analizinde yardımcı olacak program materyalleri henüz yaygın değildir. Fakat, konuların ayrıntıları, hemen bütün derslerin ders kitapları bulunduğu için, öğretmenlerin elinde bulunur.

Aslında, bir dersin özel hedefleri ile konuları arasında sıkı bir bağlantı vardır. Daha önce de belirtildiği gibi, herhangi bir hedef

ancak bir takım konularda davranışa çevrilebilir. O halde, hedef analizi de, herhangi bir ders için yapılacaksa, o dersin konularıyla birlikte yapılacaktır.

Öğretmen dersin özel hedeflerini, yukarıdaki sınıflamalara uyararak, bilişsel, duyuşsal ve psiko - motor alanların kavramlarıyla ifade edebilir. Böylece hedefler öğrencide geliştirilecek yetenekler, beceriler, tutumlar, ilgiler, v.b. olarak tanımlanır. Hatta, daha ayrıntılı olarak, Tablo 9.1, Tablo 9.2 ve Tablo 9.3'te görüldüğü gibi de tanımlanabilir. Öte yandan, her bir hedefin hangi konuya bağlı olarak geliştirileceği de, ayrıntılarına kadar saptanabilir. Böyle bir çalışmayla varılan sonuçlar iki boyutlu bir tabloda özetlenebilir. Bu tablonun bir boyutunda hedefler «davranışsal» yönleriyle; ikinci boyutunda ise «konularıyla» sıralanır. Bu tür tablolara **belirtke tablosu** denir. Tablo 9.4'de, lise fizik dersinden iki konuyu ve altı hedefi kapsayan bir belirtke tablosu örnek olarak gösterilmiştir. Bu tablonun incelenmesinden anlaşılacağı gibi, her iki konunun bütün ayrıntıları «bilgi» ve «kavrama» basamağına kadar öğrenilecektir. Ayrıca I d, II a, II b, II c, ve II f konularında «uygulama» yaptırılacaktır. II'de duyuşsal davranışlar, hem I ve hem de II'de psiko - motor davranışlar geliştirilecektir.

TABLO 9.4
BELİRTKE TABLOSU (FİZİK I)

Konular	Bilişsel				Psiko - Motor	
	1.00	2.00	3.00	3.1	3.00	4.00
I. Işığın davranışı						
a) Işık kaynakları	x	x				
b) Saydamlık	x	x				
c) Yayılma	x	x				
d) Gölgeler	x	x	x		x	
e) Işık demeti ve ışık ışıını	x	x			x	x
II. Yansıma ve aynalarda						
a) Yansıma kanunları	x	x	x		x	
b) Düz aynalar	x	x	x		x	
c) Paraboloid aynalar	x	x	x		x	
d) Teleskoplar	x	x				
e) Gerçek ve zahiri görüntüler	x	x				
f) Dağınık yansıma ve aydınlanma	x	x	x	x		

Öğretmen böyle bir tabloyu ders yılının başında bütün yıl için hazırlamalıdır. Burada küçük bir kısmı örnek olarak gösterilmiş olan tablo, daha geniş ve daha gerçekçi olarak hazırlanmalıdır. Böyle bir ana tablo hazırlandıktan sonra, hedeflerden davranış analizine geçilir. Bu basamakta öğretmenin yapacağı iş, verilen bir hedefe ulaşılabilmesi için, öğrencide ne gibi «kritik» davranışların geliştirilmesi gerektiğini ayrıntılarıyla saptamaktır. Böylece, belirtke tablosu, öğretim metotlarının saptanmasına da yardımcı olur.

Öte yandan, belirtke tablosu, ölçme ve değerlendirmede de çok işe yarar. Hedeflerden kritik davranışlara inildiği takdirde, hangi davranışların ölçüleceği meydana çıkarılmış olur. Böylece, hangi konularda ve hangi hedef basamağında soru yazılacağı da saptanmış olur. Belirtke tablosu, Tablo 9.4'teki ana tablo halindeyken, bir sınavı dahil olacak konu ve hedeflerin saptanmasına; bunlar arasındaki oranların tayin edilmesine ve her konu ve hedeften kaç tane soru sorulması gerektiğinin belirlenmesine de yardım eder.

HEDEFTEN DAVRANIŞLARA DAVRANIŞLARDAN TEST MADDELERİNE

Bir dersin özel hedefleri, tekniğine uygun olarak, açık - seçik bir biçimde ifade edildikten sonra; ikinci adım, hedeflerden hedefi tanımlayan davranışları çıkarmaktır. Her özel hedef, o hedefi oluşturan davranışlar saptanmak suretiyle tanımlanır. Bundan sonra, üçüncü adımda, saptanan her davranış için, o davranışın gözlenebileceği bir ölçme durumu yaratılır. Çoğu halde bu durum, cevaplandırılacak bir soru veya bir test maddesidir.

Ortaokul düzeyindeki matematik derslerinde Pisagor teoremi okutulur ve bu teoremin uygulamaları yaptırılır. Böyle bir öğretimden önce, özel hedefler şöyle saptanmış olabilirdi :

Matematik : Özel hedefler

1. Pisagor teoremini kavrayabilme
2. Pisagor teoremini geometriye ve topogratyaya uygulayabilme

Bu iki özel hedef, daha ayrıntılı kritik davranışlarla tanımlanmalıdır. Bu tanımlama şöyle yapılabilir :

Hedef 1. Pisagor teoremini kavrayabilme

Davranışlar :

- 1 a. Bir dik üçgende Pisagor bağıntısını kavrama.
- 1 b. Pisagor bağıntısını şu üç halde ifade etme :
$$a^2 + b^2 = c^2, \quad a^2 = c^2 - b^2, \quad c = \sqrt{a^2 + b^2}.$$
- 1 c. Pisagor teoreminin ispatını kavrama.

Hedef 2. Pisagor teoremini uygulayabilme

Davranışlar :

- 2 a. Kenar uzunluğu verilen bir karenin köşegen uzunluğunu hesaplama.
- 2 b. Köşegeni ve bir kenarı verilen bir dikdörtgenin diğer kenarını hesaplama.
- 2 c. Bir eşkenar üçgenin alanını kenarı cinsinden formülleştirme
- 2 d. Bir bataklığın verilen bir doğrultudaki uzunluğunu, yapabileceği diğer ölçmelerden hesaplama.

İşte hedefler böylece daha ayrıntılı davranışlarla tanımlandıktan sonra, sınav için soru hazırlamak veya test maddesi yazmak kolaylaşır. Bir öğretmen Pisagor teoremi konusunda yukarıdaki iki bilişsel hedefe ne derece ulaşıldığını meydana çıkaracak bir sınav hazırlayacak olsa, davranış analizi yapmamış ise, o konuda kaç soru soracağını ve sorularının ne gibi bilgileri yoklayacağını açık, seçik olarak ortaya koyamaz. Oysa ki, yukarıdaki davranış analizini yapmış bir öğretmen, Pisagor teoreminin hangi yönlerini, hangi konularda yoklayacağını da saptamış olur.

Okul programlarındaki derslerin hedeflerini çözümleyip öğrenci davranışları halinde ifade etmek, okullarımızda öğretmenlere düşer. Öğretmenin bu konuda yararlanabileceği kaynaklar vardır. Özçelik, bu çözümlemenin yöntemini örnekler üzerinde açıklamaktadır.¹⁵ Sönmez, bilişsel alan, duyuşsal alan ve psiko - motor alan hedeflerinin davranışa dönüştürülmesini çeşitli örnekler üzerinde açıklamaktadır.¹⁶

¹⁵D.A. Özçelik, Okullarda Ölçme ve Değerlendirme, s 66 - 94.

¹⁶V. Sönmez, Program Geliştirmede Öğretmen Ekitabı. 2. Baskı.

Bilişsel Davranışların Ölçülmesi

Davranışların ölçülmesinde genel bir yaklaşım, öğrencinin hedef davranışı göstereceği bir ölçme durumu yaratmak, o durumda ortaya çıkan davranışı gözleyip nicelemektir. Test ve sınav soruları birer uyarıcı olarak işler. Soruları cevaplandıran bir öğrenci, davranışlarının bir kısmını verdiği cevapta açığa vurur. Bilişsel davranışlar bilgi, bilişsel beceri ve yeteneklerden oluşur. Bilişsel davranışları ölçmek için hazırlanan test maddeleri de ölçülmek istenen davranışın ortaya çıkarılıp gözlenebilmesine imkân vermelidir.

Bundan önceki bölümde, eğitimde hedef alınan bilişsel davranışlar sınıflandırıldı. Bu bölümde, bilişsel davranışları ölçmekte kullanılacak test maddelerinin nasıl hazırlanması gerektiği, Tablo 9.1'deki aşamalı sınıflama esas alınarak tartışılmaktadır.

BİLGİ BASAMAĞINDAKİ DAVRANIŞLARIN ÖLÇÜLMESİ

Bilişsel davranışların aşamalı sınıflamasında, 1.00 kodu altında bilgi basamağındaki davranışlar sınıflanmıştır. Burada «bilgi» terimi, herhangi bir konudaki özel bilgilerin, işlem yollarının ve metodların, bir yapının veya olgusal gerçeğin **hatırlanması** anlamında kullanılmıştır.

Hatırlamanın başlıca iki türü vardır. Bunlardan biri **görünce tanıma**, diğeri **sorulunca söyleme** diyebileceğimiz zihin davranışlarıdır. Her iki halde de, daha önce kazanılmış bir yaşantı (öğrenme) akla getirilir ve ona göre hareket edilir. Bir kimseyi görünce tanımamız, onu daha önce gördüğümüzü ve onunla ilgili geçmiş yaşantılarımızı hatırlamakla olur. Bize adı verilen diğeri bir kimseyi, adının dışındaki nitelikleriyle anlatmamız da, yine onunla ilgili geçmiş yaşantılarımızı hatırladığımız için mümkün olur. Aynı şekilde olayları, eşyayı, sembollerini, kavramları, genelleme ve ilkeleri de hem görünce tanıyabilir, hem de sorulunca söyleyebiliriz. Bununla birlikte, görünce tanımak, sorulunca söylemekten farklı bir zihnin etkinliğidir. Birinci halde söz konusu bilgiler ana çizgileriyle verilir, kişi onları daha

önce görmüş veya öğrenmiş olduğunu hatırlar. İkinci halde ise, kişi, sorudaki ipuçlarından hareket ederek, daha önce öğrenmiş olduğu bilgiyi hatırlar, zihindeki diğer bilgiler arasından seçip istenilen bir ifade aracıyla bildirir (söyler veya yazar). Davranışların ölçülmesinde, hatırlamanın bu iki türünün birbirinden ayrılması önemlidir.

Sınav ve testlerde kullanılan soruların bazıları görünce tanımayı, doğruyu yanlıştan ayırmayı, sorulan şeyi sorulmayanların arasından seçmeyi gerektirir; bazıları ise sorulan şeyi hatırlamayı ve istenilen bir yolda ifade etmeyi gerektirir. Bilgi (1.00) kodu altında sınıflanan hedeflerde, bilgilerin, işlem yollarının, yapıların ve durumların bu iki türden herhangi biriyle hatırlanabilecek kadar iyi öğrenilmesi amaç alınır. Onun için, böyle bir eğitim hedefinin gerçekleşip gerçekleşmediğini ölçmenin yolu, söz konusu bilgileri, bu iki tür hatırlamanın herhangi biriyle yoklamaktır.

Tablo 9.1'de görüldüğü gibi, bilgi basamağındaki hedefler önce üç sınıfa, sonra bu sınıflar da kendi aralarında alt sınıflara ayrılmıştır. Aşağıda bu sınıfların anlamları açıklanmakta, söz konusu hedef davranışların nasıl ölçüldüğü gösterilmektedir.

Bir Alana Özgü Bilgiler (1.10)

Her alanın kendine özgü özel bilgileri vardır. Ayrıca, her alanın gene kendine özgü terimleri vardır. Bu alt sınıfta herhangi bir alanın özel bilgilerini ve terimlerini bilmek hedef alınmıştır. Bu bilgiler soyutlamanın en alt basamağını meydana getirir. Onun için bu alt sınıf, daha ziyade somut bilgiler, temel kavramlar ve terminolojiden oluşur.

Eğitimde özel bilgilerin, kavram ve terimlerin bilinmesini hedef almak yeni bir şey değildir. Derslerin amaçları incelendiğinde, hemen her derste, o dersin konusuna özgü bilgileri ve kavramları terimleriyle birlikte öğrenmenin hedef alındığı görülür. Şöyle ifade edilmiş hedeflere raslanabilir :

- Bir terimin bilimsel anlamını, günlük dildeki anlamından ayırabilme
- Sayısal düşünmede kullanılan kavramları ve dili bilebilme
- Çocuk gelişimiyle ilgili başlıca kavramları tanımlayabilme

Yukarıdakilere benzer ifadeler hedef olarak alındığında, geliştirilecek davranışlar «terimlerin anlamını söyleme veya yazma», «kavramları anlama veya tanımlama», v.b. olur. Her bilgi alanının özel terimleri ve kavramları yanında, o alana özgü ifade tarzları da var-

dır. Terimler kavramların sembolleri olarak kullanılır. Bir alana özgü ifade tarzları, anlama kesinlik kazandırmak için kullanılır. Böylece, anlaşma (iletişim) kolaylaşır. Onun için, bu alt sınıfta, ister sözel olsun ister başka sembollerde olsun, sembollerin ve özel ifade tarzlarının anlamını bilebilme hedef alınmıştır.

Her bilgi alanının kendine özgü bilgileri vardır. Bir alanı bilme, terim ve kavramlarıyla birlikte bu özel bilgilerin öğrenilmesiyle başlar. Her eğitim seviyesinin müfredat programlarında, öğrencilere, o seviyeye uygun ölçülerde bilgi kazandırmak en önemli hedeflerden biri olagelmıştır. Bu nokta aşağıdaki hedeflerde görülmektedir.

- Türkiye'nin doğal kaynaklarının bilgisi
- Besin maddelerinin enerji değerlerinin bilgisi
- Ondalık sayıların temel özelliklerinin bilgisi
- İkinci Meşrutiyet devrinin dış politika olaylarının bilgisi
- Meteoroloji tahminlerinin nerden sağlanacağını bilebilme
- Bir kitap hakkındaki eleştirilerin nerede bulunacağını bilebilme

Bu hedefler hem olgusal bilgileri, hem de olgusal bilgilerin elde edilebileceği kaynakları kapsamaktadır.

Tablo 9.1'deki sınıflamada görüleceği gibi, bu alt sınıfta terim bilgisi (1.11) ve özel bilgiler (1.12) birbirinden ayrılmıştır. Ölçme işlemlerinde bu ayrımı yapmak zordur. Çünkü terimler kullanılmaksızın özel bilgiler yoklamaz; özel bilgilere dayanılmaksızın da, terimlerin ve sembollerin anlamları yoklanamaz. Fakat, bu zorluk ölçme ve değerlendirmede önemli bir sakınca doğurmaz; eğitimde kullanılan en analitik ölçmelerde bile, terimleri ve sembolleri özel bilgilerden ayırmak ihtiyacı pek az doğar. Bu sebeple, aşağıda verilen örneklerde 1.11 ve 1.12 ayrımı yapılmamıştır.

Bu alt sınıftaki davranışların gelişip gelişmediği nasıl yoklanabilir? Başka bir ifadeyle, bir kimsenin bir alandaki bilgieri, kavram ve terim bilgileri nasıl ölçülebilir? Bu soruya verilecek cevap bir yönüyle yoklanacak bilgilerin tabiatına, terim ve sembollerin şekline; diğer yönüyle de sınavda kullanılacak ölçme aracının türüne bağlıdır. Aşağıda kısa cevaplı sınavlarda sorulabilecek sorulardan ve test maddelerinden örnekler verilmiştir.

1. Bir şairin şiirlerinde kullandığı takma ada ne denir?

- | | | |
|----------|-----------|-----------|
| A) Lakap | B) Mahlas | C) Nazire |
| D) İmza | E) Nam | |

(1.10)

(B)

2. Aşağıdaki bilgilerden hangisi çok kullanılan bir kişilik kuramı ileri sürmüştür?
(A) Thorndike (B) Freud (C) Watson
(D) Köhler (E) Wundt
(1.10) (B)
3. Adını yeni duyduğunuz bir kimyasal elemanın kaç protonlu olduğunu çabucak bulabilmek için hangi kaynağa başvurursunuz?
(A) Bir ansiklopediye (C) Periyodik cetvele
(B) Ansiklopedik sözlüğe (D) Bir kimya kitabına
(E) Bir kimya dergisine
(1.10) (C)
4. Bir miktar maddayı etkileyen kuvvetin, bu kuvvetin etkisinde o maddenin kazandığı ıvmeye bölünmesiyle bulunan orana ne denir?
(A) Ağırlık (B) İtme (C) Kütle (D) Hız
(1.10) (C)

Birbirine yakın birkaç kavram ve terimi beraberce yoklamanın daha etkili bir yolu, ortak şıklı maddeler kullanmaktır. Örnek :

5. - 8. sorular için : Aşağıdaki soruları şu seçeneklerle cevaplandırınız.
(A) Beceri (B) Zekâ (C) Anıklık
(D) Kişilik (E) Hiçbiri
5. Belirli durumlara karşı gösterilen tepkilerde, yaşantıların etkisiyle meydana gelen, oldukça sürekli değişimler.
6. Bireyin bir davranışlar bütününe ne dereceye kadar öğrenebileceğinin ölçüsü.
7. Karmaşık bir devimsel davranışlar dizisini yeterince hatasız ve hızlı yapabilme yeteneği.
8. Öğrenme, soyutlama ve yeni durumlara uyma yeteneklerinin bileşimi.
- (1.10) (5 : E, 6 : C, 7 : A, 8 : B)

Bir Alana Özgü Bilgilerle Uğraşma Araç ve Metotları (1.20)

Bir bilgi alanında kullanılan alışların (teamüllerin); o alandaki yönelim, sıra ve dizilerin, sınıf ve kategorilerin; o alanın bilgileriyle uğraşma metotlarının; o alanda bilgilerin doğruluk derecesi hakkında hangi kanıtlarla değer yargısına varıldığının bilgisi bu alt sınıfta toplanan hedeflerdir.

Her alanda kabul edilmiş veya öyle kullanılması gelenekselleşmiş bazı alışlar vardır. Bir üçgenin köşelerini belli bir yönde ve belli harflerle gösterme; 3.6.1972 tarihini «Mart 1972» değil, «Haziran 1972» olarak anlama v.b: böyle alışlardanır. Alışların bilinmesi hem iletişimi hem de diğer bilgilerin öğrenilmesini kolaylaştırır.

Aynı şekilde, her alanda sırasıyla bilinmesi gerekli bazı bilgiler vardır. Meselâ, prizmadan geçen ışık belli sırayla sıralanmış bilinen renklere ayrılır. Ayrıca, bazı olayların kendilerine özgü gidiş ve oluşları vardır. Meselâ bir böcek, doğumundan ölümüne kadar belli hayat evrelerinden geçer. Türlerin tarihsel gelişimi belli bir yönelim gösterir. Kimyasal elemanların belli sıra ve kategorileri vardır, kimyasal özellikler bu sıra ve sınıflar üzerinde belli yönelimler gösterir. Müfredat programlarında sıra ve dizilerin, yönelimlerin bilgisi çoğu halde hedef alınır.

Bir alandaki bilgiler geliştikçe, o bilgilerin sınıflaması da gelişir. Sınıflamaların gerek yeni araştırmalarda, gerekse o alanın bilgilerinin öğrenilmesinde önemli yardımları olur. Canlılar sistematığının biyolojiye, periyodik cetvelin kimyaya katkıları bu konuda iyi örneklerdir. Müfredat programlarında, sınıflamaların bilgisi ve onları gerektiğinde kullanabilme hedef alınır. Hatta, bazı derslerin konuları sınıflama sistematığıne göre programa girer.

Her bilgi alanında, işlem yollarının, gözlemlerin, genellemelerin ve görüşlerin doğru olup olmadığı hakkında değer yargısına ulaşmakta kullanılan ölçütler vardır. Ölçütler, bilgi üretmenin çeşitli basamaklarında yararlı olur. Öğrencilerin ölçütleri bilmeleri ve yerinde kullanabilmeleri, o alanın bilgilerini o alanın yöntemleriyle, bir bilim adamı gibi öğrenmelerini kolaylaştırır. Bir öğrenci bir alanın bilgilerini araştırarak, hipotez kurarak, deneyerek ve yoklayarak öğrencekse ölçütleri kullanmak zorundadır. Onun için herhangi bir dersin müfredat programında, o alanın bilgileriyle beraber, ölçütlerinin de öğrenilmesi, hatta kullanılarak öğrenilmesi hedef alınır. Her bilgi alanına özgü bilgi edinme ve araştırma yöntemleri vardır. Hatta, bir alan içinde sadece bazı olayların incelenmesinde kullanılan özel teknikler bile vardır. Her alanın kendine özgü araç ve yöntemlerinin bilgisi ve onların yerinde kullanılabilmesi, o dersin hedeflerindendir.

Aşağıda bu alt sınıf hedef davranışlarının ölçülmesinde kullanılabilecek madde örnekleri verilmiş ve her madde son basamakta kodlanmıştır.

9. Elektrik devrelerinde, akımın hangi yönde aktığı kabul edilir?
(1.21) (Doğru cevap : + uçtan, — uca.)

10. 0,0484 sayısında, kaç tane anlamlı basamak vardır?

A) 1 B) 2 C) 3 D) 4 E) 5
(1.21) (C)

11. Aşamalı sınıflamada değerlendirme (D), uygulama (U), kavrama (K) ve sentez (S) üst düzeye doğru nasıl sıralanmıştır?
A) U - S - K - D B) K - U - D - S C) D - K - S - U
D) K - U - S - D E) D - S - U - K

(1.22) (D)

12. Aşağıdakilerden hangisi Divan Edebiyatının nazım türlerinden değildir?
A) Kaside B) Gazel C) Şarkı
D) Türkü E) Rübâî

(1.23) (D)

Ölçüt bilgilerini ve ölçütleri kullanabilme yeteneğini ölçen maddelerden örnekler :

13. Tarihî bir olayın tarihinin doğruluğunu kanıtlamak için ne gibi ölçütler kullanılır?

(1.24)

14. Aşağıdakilerden hangisinin farklı oluşu, ikizlerin ayrı yumurtalardan türediklerinin kesin kanıtıdır?
A) Kan grubu B) Kromozom sayısı
C) Cinsiyet D) Zekâ düzeyi

(1.24) (C)

Metot bilgilerini ve metotları kullanabilme yeteneğini ölçen maddelerden örnekler :

15. Okulöncesi çağındaki bir çocuğun zekâ düzeyi hangi yöntemlerle kestirilebilir?

(1.25)

16. Bir yıldızda ne gibi kimyasal elementler bulunduğu hangi yolla anlaşılır?

- A) Yoğunluğun tayini B) Parlaklığın ölçülmesi
C) Spektrum analizi D) Uzaklığın ölçülmesi

1.25) (C)

Bir Alandaki Evrensel Öğelerin ve Soyutlamaların Bilgisi (1.30)

Her bilgi alanında özel kavramlardan genel kavramlara, somut yaşantılardan soyut düşüncelere gidilir. Daha ileride, genellemelerden ilkelere varılır; genellemelerden ve diğer soyutlamalardan kuramsal yapılar ve kuramlar (teoriler) kurulur. Atomun yapısını, daha genel olarak maddenin yapısını açıklamaya çalışan kuramlar ve hayatı açıklayan kuramlar bunun örnekleridir. Bir bakıma, bilgi denince birçok kimsenin aklına bu tür soyutlamalar gelir. Bu anlamda bilgi bütün okul programlarında yer alır.

Bu alt sınıftaki bilgiler kendi aralarında ikiye ayrılmıştır : İlke ve genellemeler (1.31), kuramlar ve kuramsal yapılar (1.32). İlke ve ge-

nellemeler bir olaylar grubu üzerindeki gözlemleri özetler; sonra benzer olayları anlatmak, açıklamak veya önceden kestirmekte kullanılır. Burada hedef alınan davranış, öğrencinin ilke ve genellemeleri yeni durumlara uygulayabilmesi değil, tanıyabilmesi veya hatırlayabilmesidir. Okul programlarında bu türden amaçlara örnek olarak aşağıdakiler söylenebilir :

- Enerji dönüşümünün temel ilkelerinin bilgisi
- Freud'ün kişilik kavramının bilgisi

Aşağıda, ilkeler ve kuramsal yapılar bilgilerini yoklayan maddelerden örnekler verilmiştir.

17. Bazı mevsimlerde, günün akşam saatlerine doğru denizden kayera esen rüzgar, hangi fiziksel niteliklerdeki bölgelerde görülür?
(1.31)
18. Aşağıdaki ışık olaylarından hangisi Newton'un tanecik teorisi ile açıklanamamıştır?
A) Yansıma B) Kırılma C) Enerji taşıma
D) Boşlukta yayılma E) Doğrusal yolla yayılma
(1.32) (B)

KAVRAMA BASAMAĞINDAKİ DAVRANIŞLARIN ÖLÇÜLMESİ

Bundan önceki kesimde bilgilerin hatırlanması veya tanınması basamağındaki bilişsel hedefler tartışıldı. Bilginin sadece hatırlanması, bilişsel davranışların en alt basamağını teşkil eder. Bilgiye dayanan zihin davranışlarının, bu aşamalı sınıflamada aşağıdan ikincisi kavrama veya bilgileri anlamadır.

Bir bilginin öğrenci tarafından kavranması, onun taşıdığı anlamın öğrencinin zihninde meydana gelmesidir. Bu bir iletişim sürecidir ve şöyle olur : Bilginin anlamı bir kaynak tarafından kodlanarak (herhangi bir dille veya sembolle ifade edilerek) alıcıya gönderilir. İletişim terimlerinde, kodlanmış bir anlam mesaj adını alır. Alıcı, aldığı mesajın kodunu çözerek (kodu sembolden anlama çevirerek) anlamını kavrar. Aslında bir öğretmen, bir kitap yazarı, bir ressam, bir film yapımcısı birer kaynak olarak bilgi mesajları yayarlar. Bunların hitap ettiği kimseler de, birer alıcı olarak, mesajı alıp sembollere yüklü bilgiyi sembolden anlama çevirebildikleri takdirde kavrayabilirler. Bir öğrencinin bir mesajdaki anlamı tamamen kavrayabilmesi de, ancak, mesajın kayıba uğramadan ona iletilmesi ve onun mesajı sembolden anlama çevirebilmesiyle mümkün olur.

Bu konuda, ölçme ve değerlendirme amaçları için önemli olan soru, bir bilgiyi kavramış bir kimsenin bunu hangi davranışlarıyla

açığa vurduğu sorusudur. Eskiden bazı eğitimciler bu soruya yanlış bir cevap vermişlerdir. Bu eğitimciler bir bilgiyi öğretmenin söylediği veya kitabın yazdığı ifadeyle tekrar edebilen bir öğrencinin, onu kavradığını kabul etmişlerdir. Bu davranışı bir bilgi mesajını anlayan bir kimse gösterebileceği gibi, anlamadan ezberleyen bir kimse de gösterebilir. O halde, bir bilgiyi kavramış olan bir kimse, anlamadan ezberlemiş olan bir kimseden, davranışlarına bakılarak nasıl ayırt edilebilir?

Bloom ve arkadaşları aşamalı sınıflamayı geliştirirken bu soruya daha doğru bir cevap vermişlerdir.¹ Bir bilgi mesajındaki anlamı kavramış bir kimse, onu mesajdakinden farklı bir dilde ifade edebilir; yani, mesajdaki anlamı diğer bir dile veya diğer sembollere **çevrilebilir**. Kavramanın bundan sonraki basamağına ulaşmış bir kimse ise, mesajdaki anlamı **yorumlayabilir**. Kavramanın en ileri basamağındaki bir kimse de, mesajdaki anlama dayanarak, **öteleme** ve **ulama** (ekstrapolasyon ve enterpolasyon) yapabilir.

Müfredat programlarda açıkça ifade edilmemiş olmakla beraber, kavramanın bu üç basamağındaki davranışlar eğitimin önemli hedeflerindendir. Aşağıda bu üç tür davranış açıklanmaktadır.

Çevirme (Tercüme) (2.10)

Bilgiler genellikle bir dille veya başka sembollerle ifade edilmiş halde bulunur. Bir ifadedeki (mesajdaki) anlamı kavrayan bir öğrenci, onu kendi diliyle yeniden ifade edebilir. Bir bilgi öğrencinin onu öğrendiği zamanki dilden farklı bir şekilde ifade edilmişse, öğrencinin onu, farklılığına rağmen tanıması da ancak bilginin kavranmasıyla olur. Bir bilginin anlamı, birbirinden farklı iki dilde ifade edilmişse, yani anlam iki ayrı sembol grubu ile ifade edilmişse, bu ikisinin aynı mesajı taşıdığını anlayabilen bir kimse mesajlardaki anlamı kavramıştır. O halde, burada «çevirme» terimiyle ifade edilen davranışlar, bir anlamın farklı bir dille ifade edilmesi, bir semboller takımından diğer semboller takımına çevrilmesi, iki farklı ifadenin aynı anlamda olduğunun tanınabilmesi gibi davranışlar halinde görülebilir.

Çevirme, bir sözel kalıptan diğer bir sözel kalıba olabileceği gibi, bir sözel kalıptan diğer sembollere, diğer sembollerden sözel

¹B.S. Bloom (Ed.), *Taxonomy of Educational Objectives; Handbook I : Cognitive Domain*.

kalıplara, sembollerden sembollere gibi bir soyutlama türünden diğerine olabilir. Yabancı dilde bir metni okuyan ve anlamını Türkçe olarak yazan bir kimse bir çevirme davranışı göstermiş olur. Bir cebir problemini okuyan ve onun çözümü için bir cebirsel denklem yazan kimse ile bir cebirsel denkleme sözel bir problem yakıştıran bir başka kimse de aynı davranışı göstermektedir. Aynı şekilde, teknik sembollerle ifade edilmiş gerçekleri söze ifade edebilme, bir plan veya tablodaki gerçekleri söze çevirebilme de aynı türden davranışlardır. Burada dikkat edilmesi gereken bir nokta, kavrama olmaksızın da bazı çevirme davranışlarının mümkün olduğudur. Bir telgraftaki sözleri Mors koduyla gönderen bir telgrafçı, bir sembolden diğer bir sembole çevirme yapmaktadır; bunun için telgraf metnindeki anlamı kavramış olması gerekmez. Sabit bir kurala bir sembolden diğerine doğru yapılmış bir çevirme, kuralın uygulandığına kesin kanıttır; fakat sembollerde gizlenmiş olan anlamın kavrandığına kesin kanıt olamaz. Onun için, bir sınavda bir çevirme işlemi yaptırılacaksa, anlamı kavramdan çevirme yapmanın imkânsız olduğu sorular sorulmalıdır.

Kavramayı çevirme basamağında yoklayabilmek için, test ve sınav durumlarında öğrenciden ya bir bilgiyi kendi diliyle ifade etmesi, ya öğrenme durumundan farklı bir dilde ifade edilmiş bir bilgiyi tanıması, ya da bir sembol takımıyla ifade edilmiş bir mesajı diğer bir sembol takımıyla yeniden ifade etmesi istenir. Aşağıda, kavramayı, çevirme davranışlarıyla ölçen sorulardan örnekler verilmiştir.

19. Çözümü, $3x + 4 = 50 - 1$ denkleminin çözülmesiyle bulunabilecek bir problem yazınız.
(2.10)
20. Türkiye'nin 1950 - 1970 arasındaki pamuk üretimini ilgili kaynaklardan bulup, topladığınız bilgileri bir grafikte gösteriniz.
(2.10)
21. Sekiz yaşındaki bir çocuğun, bir testle ölçülen zekâ yaşı 108 ay bulunmuştur. Buna göre, aşağıdaki önermelerden hangisi doğrudur?
A) Çocuk sekiz yaşından büyüktür.
B) Çocuğun zekâ bölümü 108'dir.
C) Çocuğun zekâsı ortalamanın 8 puan üstündedir.
D) Çocuk testte yaşının üstünde başarılı olmuştur.
(2.10) (D)
22. Lityumun atom numarası 3'tür. Bazı lityum atomlarında 3, bazıları 4 nötron bulunur. Lityum kaç izotoplu olabilir?
A) 2 B) 3 C) 4 D) 6 E) 7
(2.10) (A)

Bilgileri grafikle veya başka bir şekilde gösterebilme bir çevirme yeteneği olduğu gibi, bir grafikle veya bir formülle ifade edilmiş bilgileri sözle söyleyebilme, ya da yazı ile yazabilme de çevirme yetenekleridir. Çünkü bu işleri doğru yapabilen bir öğrenci, grafik veya formüldeki mesajı anlamış olduğunu gösterir.

Yorumlama (2.20)

Bir bilgi mesajını yorumlamak, mesajda gizli olan anlamı daha açık ve daha kolay anlaşılır hale getirmektir. Bu yetenek yukarıda incelenen çevirme yeteneğinden biraz daha ileri bir yetenektir. Yorumlamada mesajı diğer bir dille ifade etmekten ötede, verilen bilgilerde bazı bağıntılar bulma ve onları açıklama davranışları da vardır. Yorumlama bazen, verilen bilgilerden bir genellemeye gidilerek, bazen de bir genelleme özel durumlara uygulanarak yapılır. Uzun bir paragrafı anlamını bozmadan özetleme, bir ilkeyi örneklerle açıklama da yorumlama davranışlarıdır.

Bu yeteneğin eğitimdeki önemi açıktır. Günlük hayatta bile, yorumlama gerektiren pek çok durumla karşılaşırız. Onun için, müfredat programlarında yorumlama yeteneğinin geliştirilmesinden bir hedef olarak söz edilir.

Yorumlama yeteneğinin ölçülmesinde bir pragraftaki anlamın açıklanmasını, uzun bir parçanın özetlenmesini, bir bilgi bütünündeki bağıntıların açıklanmasını, özel gözlemlerden genel yargılar çıkarılmasını ve genel ilkelerin özel durumlardaki hallerini soran sorular kullanılır. Aşağıda, yorumlama yeteneğini ölçen bir test maddesi verilmiştir.

23. Dolunay kışın gümüşî parlaklıkta, yazın sarımtırak görünür.

Aşağıdaki ifadelerden hangisi, bu olayı açıklar?

- A) Ay güneşten gelen ışığı yansıtır.
- B) Dolunay yazın ufka daha yakındır.
- C) Atmosfer ışığı yazın daha iyi geçirir.
- D) Dolunaydan yansıyan ışık kışın daha kalın bir atmosfer tabakasından geçer.
- E) Dolunay kışın daha çok ışık yansıtır.

(2.20)

(E)

Cevaplayıcının bu maddeyi doğru cevaplayabilmesi için, madde kökündeki «gümüşî parlaklıkta» ifadesini, «daha çok ışık yansıtır» anlamında yorumlaması gerekir. Sorulan soru, aşağıdaki maddede olduğu gibi farazî bir duruma dayanırsa, öğrencinin doğru cevabı daha önce görmüş olması ihtimali de ortadan kalkar; böylece, yorumlama yeteneği daha geçerli bir madde ile ölçülmüş olur.

24. Bir bitkinin normal büyümesi için belirli miktarda su gereklidir. Bitkiler suyu kökleriyle alıp terlemeyle verirler. Bitkiler terlemeyi tamamen durdurabilecek sistemlere sahip olsalardı, kendi terlemesini durduran bir bitkiye ne olurdu?
- A) Bitki daha hızlı özümleme yapardı.
 - B) Bitki madensel tuzları yeteri kadar alamazdı.
 - C) Özümleme ürünü olan organik maddelerin iletimi dururdu.
 - D) Bitkinin yedek besinleri kullanması dururdu.
 - E) Bitkinin topraktan aldığı su miktarı artardı.

(2.20)

(B)

Bu maddede istenen yorum şudur : «Bitki terlemeyi durdurduğu için kökten yaprağa doğru olan su akımı da durur. Bu sebeple de, suda erimiş olarak bitkiye ulaşan madensel tuzların ulaşımı durur». Şüphesiz, bu yorumu yapabilen bir cevaplayıcı doğru cevap olarak B şıkkını seçer. Başka yorumlar bu kadar doğru olmaz.

Öteleme (Ekstrapolasyon ve Enterpolasyon) (2.30)

Bir bilgiyi, yönelim veya ilkeyi mevcut verilere dayanarak veriler dışına uzatmak; gözlenmiş olaylar alanına bakarak, gözlenmemiş alandaki olaylar hakkında yargıya varmak gibi bilişsel davranışlara bu sınıflamada öteleme denmiştir.

Ancak bazı şartlar altında doğru bir öteleme yapılabilir. Öteleme yapacak kimsenin verilen bilgileri yorumlayabilecek kadar iyi öğrenmiş olması ilk şarttır. Bundan başka, öteleme yapacak kimse verilerdeki yönelimi, olayın gidişini ve sonucunu görebilmelidir. Verilerin sınırları iyi bilinmelidir. Verilerden çıkarılan genellemelerin, bu sınırlar içinde doğru olduğu, ötelemeyle ulaşılan genellemelerin ise, bu sınırlar dışında ancak muhtemelen doğru olabileceği bilinmelidir. Bu anlamda öteleme, sadece mevcut verilere dayanan olayların veri sınırları dışındaki gidişleri hakkında yargıya varmaktır. Fakat, aşağıda sıralanan hedeflerden anlaşılacağı gibi, bu sınıflamada öteleme daha geniş alınmıştır; istatistiksel ekstrapolasyon ve enterpolasyondan başka öteleme davranışları da vardır. Müfredat programlarındaki öteleme hedeflerinden bazı örnekler :

- Yönelimleri, olayların gelişmelerini önceden kestirebilme.
- Verilerdeki boşlukları enterpolasyon ve ekstrapolasyonla doldurabilme.
- Kestirme ve tahminleri yanılta etkenleri tanıyabilme ve bunlara karşı tedbirli olabilme.

Öteleme davranışının yoklanmasında şu esaslara göre soru ve test maddeleri hazırlanmalıdır :

1. Zamanla değişen bir bilgi bütünü verilir. Bu verilerdeki yönelimin, zaman aralığının dışına ötelenmesi istenir.

2. Bir konuda bir genelleme verilir. Bu genellenenin, dayandığı gözlemler dışına ötelenmesi istenir.

3. Bir evrenden alınmış örneklemelere ait bilgiler verilir, bunlara dayanılarak evrene ait genellemeler çıkarılması istenir.

4. Bir bilgi bütününe ait veriler ve onlardan çıkarılan genellemeler verilir. Bunların doğruluk derecesi, hataları, muhtemel hata kaynakları sorulur.

Aşağıda, öteleme davranışlarını ölçen maddelerden oluşan bir test örneği verilmiştir.

25. - 29. sorular için : Bir tuzun suda çözünürlüğünü sıcaklığa bağlı olarak inceleyen bir öğrenci, deneylerden aldığı verileri şu tabloda toplamıştır :

<u>Çözücünün sıcaklığı (°C)</u>	<u>Çözünürlük (gr/100 cm³)</u>
10	22
20	33
40	62
60	111
80	172
100	235

Öğrenci, bu tabloya bakarak, aşağıdaki genellemeleri yapmıştır. Siz onun genellemelerini inceleyip, yukarıdaki verilere göre ne derecede doğru olduklarını tayin edeceksiniz. Cevaplarınızı şu seçeneklerle işaretleyiniz.

A : Kesinlikle doğru, B : Doğru olması daha olası

C : Doğruluğu veya yanlışlığı söylenemez.

D : Yanlış olması daha olası, E : Kesinlikle yanlış

Öğrencinin genellemeleri :

25. Bu tuzun çözünürlüğü, sıcaklığın doğrusal bir fonksiyonudur.

26. 50°C deki çözünürlük 87 gr/100 cm³ olmalıdır.

27. Su 120°C ye kadar ısıtılabilseydi, çözünürlük 300 gr/100 cm³ bulunurdu.

28. Tuzların suda çözünürlüğü sadece sıcaklığa bağlıdır.

29. Tuzlar suda çözünürken ısı verir.

(2.30) (25 : E, 26 : B, 27 : C, 28 : C, 29 : D)

UYGULAMA BASAMAĞINDAKİ DAVRANIŞLARIN ÖLÇÜLMESİ

Burada uygulama deyimiyle ifade edilen şey bilginin veya bilgilerle uğraşma yollarının yeni durumlara uygulanmasıdır. Okullarda bilgiler ve teknikler ileride kullanılacağı, yeni durumlara uygulanacağı inancıyla verilir. Onun için, müfredat programlarında bilgilerin bir işe yaramasından veya gerektiğinde hayatın çeşitli aşamalarında kullanılabilmesinden söz edilir. Gerçekten, öğrenilen bir bilgi ya aynı konuda daha ileri bilgilerin kazanılmasında, ya da karşılaşılan bir pratik güçlüğün çözülmesinde kullanılır.

Bir bilginin yeni bir durumda uygulanabilmesinin ilk şartı, onun, gerektiğinde hatırlanıp kullanılabilecek kadar iyi öğrenilmiş olmasıdır. Onun için, bilgilerin bundan önceki kesimde ele alınan basamaklarda kavranmış olması uygulanabilirliğin ilk şartıdır. Fakat, bilginin kavranmış olması, yeni bir durumda uygulanabilmesi için yeter şart değildir. Uygulama yapması beklenen öğrenci, yeni durumu yorumlayıp problemin ne olduğunu görebilmeli; bu durumdaki problemin çözümünde daha önce öğrendiği ilkelerin, bilgilerin veya çözüm yollarının başarıya götüreceğini kestirebilmeli; hiç değilse, bu yeni problemle daha önce çözdüğü bir problem arasında benzerlikler görebilmelidir. Bu yeteneklerden yoksun bir kimse uygulama için gerekli bilgilere sahip olsa bile, uygulamayı yapamaz. Öğretmenin uygulama gerektiren problemler sorduğu; uygulama veya problemin çözümü yerine, ilgili bilgileri sayıp döken cevaplar aldığı çok görülmüştür.

Her uygulama, basit dahi olsa, bir problem çözme işlemi haline gelir. Problem çözme mekaniksel veya alışılmış bir beceri olabileceği gibi, düşünme ve buluş gerektiren bir işlem de olabilir. Öğrenci için yeniliği olmayan bir fizik veya matematik problemi, çoğu halde, bilinen bir formül üzerinde alışılmış bir işlemden başka bir şey değildir. Böyle bir halde ölçülen davranış, ilk bakışta «uygulama» gibi görünürse de, bundan daha aşağı basamakta bir zihin davranışdır.

Uygulama davranışının ölçülmesinde kullanılacak problem, hiç değilse bazı yönleriyle, öğrenci için yeni olmalıdır. Çoğu zaman, bir problemin bir öğrenci için yeni olup olmadığı kesinlikle bilinemez. Hele bir problemin bir sınıftaki bütün öğrenciler için yeni olduğunu söylemek daha güçtür. Bir problemi öğrenci için «yeni» hale getirmenin bazı yolları şunlardır :

1. Derste veya pratikte karşılaşılmaması mümkün olmayan, ya da üzerinde fiilen çalışılmayan farazî bir durum uydurulur. Bu durumla ilgili problemler düzenlenir.

2. Derste öğrenilmiş veya pratikte karşılaşılmış bir durum, önemli öğeleri değiştirilerek, yeni bir problem haline getirilir.

3. Öğrencinin henüz öğrenmediği yukarı sınıf konularında, öğrenilmiş konuların uygulanmasıyla çözülebilecek problemler düzenlenir.

Bir problem çeşitli derecelerde «yeni» olabilir. Bazı problemlerde yeni öğeler azdır, önemsizdir; böyle problemlerin çözümü için ileri derecede görüş veya buluş gerekmez. Bu tür problemlerle ileri derecede bir uygulama yeteneği ölçülmüş olmaz. Bazı problemler ise, ya çok karmaşık ya da öğrenilen bilgilerle benzerliği çok az problemler olabilir. Öğrenilen bilgilerin bu tür problemlere uygulanması daha zordur. Bu tür problemleri çözebilen öğrencilerde ise daha ileri bir uygulama yeteneği bulunmak gerekir. Problemin yeniliği ve çözümü için gerekli düşünüş yolunun alışılmış yollardan uzaklığı arttıkça, o problemle ölçülen uygulama yeteneği de artar. Findley ve Scates ölçme durumunun haddinden fazla yeni olmasının tehlikesine işaret etmişlerdir.²

Öğrenciden, bilgileri yeni durumlara uygulamakta dört tür davranış beklenebilir.

1. Öğrenciden, probleme uygulanacak ilkeyi bulması, bunu probleme uygulaması, çözümü yazılı veya sözlü olarak vermesi istenir. Bu davranış test tipi dışındaki sözlü ve yazılı sınavlarla yoklanabilir. Çözümde kullanılacak ilkenin soruda açıklanması veya ima edilmesi durumun yeniliğini zedeler. Öte yandan, ilkeyi bulmuş, çözümü başarıyla uygulamış bir öğrencinin, hesaplama, vb. gibi ayrıntılarda hata yapmış olması, ölçülecek davranışı zedelemeyiz.

2. Öğrenciye bir problem verilir ve sadece çözüm için gerekli işlem yolunu göstermesi istenir. Doğru işlem yolunu gösteren veya verilen işlem yolları arasından en uygununu seçebilen bir öğrencinin çözümde kullanılacak ilkeyi ve işlem yollarını bildiği farzedilir. Test sorularının bazıları bu esasa göre hazırlanır. Davranışın bu yolla ölçülmesi, çözüm istemediği için, hem çözümü bulandırabilecek yan değişkenlerin (hesaplama hataları, yazı güzelliği, vb. gibi) ölçmeye girmesini önler, hem de ölçme için gerekli süreyi kısaltır. Çözümün kompoze edilmesi, hesaplama becerilerinin kullanılması

²W.G. Findley and D.E. Scates, «Obtaining Evidence of Understanding», in Nelson B. Henry (Ed.), *The Measurement of Understanding*, s. 44 - 64.

gibi davranışların ölçülmesi gerekmiyorsa, uygulama yeteneğinin ölçülmesinde bu ikinci yol daha etkilidir.

3. Öğrenciden problemi çözmesi ve sadece sonucu vermesi veya verilenler arasından doğru çözümü seçmesi istenir. Bu tür bir ölçme metodu seçmeli test maddelerine daha uygun düşer.

4. Problemin çözümü verilir ve çözümde hangi ilkenin kullanıldığı sorulur veya çözümde kullanılan ilke diğer ilkeler arasından seçtirilir. Bu tür sorular da seçmeli test maddelerine uygun düşer. Fakat test maddesi çoğu halde daha kolay olur.

Aşağıda uygulama yeteneğinin ölçülmesinde kullanılabilecek sorulardan örnekler verilmiştir. Okuyucu bu örnekleri incelerken, her sorunun hangi öğelerinin hangi tip öğrenciler için yeni olabileceğini düşünmeli; ayrıca, problemin nasıl «yeni» hale getirildiğini bulmalıdır.

30. 10 m çapında 60 m yüksekliğinde bir kulenin hacimce 1/1000 oranında bir model yapılacaktır. Modelin yüksekliği ne kadar olacaktır?
A) 0,6 cm B) 6 cm C) 60 cm
D) 6 m E) Hiçbiri
(3.00) (D)
31. İnsan vücudu nişastayı sindiremeseydi, hayatımız hangi yönlerden bugünkünden farklı olurdu?
(3.00)
32. Otomobilinin lastiğini değiştirecek bir adam kriket kullanmakla ne kazanır?
A) Kriket ile daha az iş yapar.
B) Kriket ile daha çok iş yapar.
C) Otomobile uyguladığı kuvveti büyütür.
D) Otomobil daha yükseğe kaldırır.
(3.00) (C)

ANALİZ BASAMAĞINDAKİ DAVRANIŞLARIN ÖLÇÜLMESİ

Olaylar daima karmaşık halde karşımıza çıkar. Bilgi ürünleri, teknik sistemler ve benzerleri de, belli derecelerde karmaşıklığı, bazı öğeleri, öğeler arası bağıntıları ve genel yapısal ilkeleri olan bütünlere sahiptir. Bir olayı, bir bilgi bütünü veya bir sistemi öğrenmenin yollarından biri, onu analiz etmektir.

Analiz işlemi, genellikle, bir sistemin ne gibi öğelerden meydana geldiğini, bu öğeler arasında ne gibi ilişkiler olduğunu, sistemin

hangi ilkelere göre kurulduğunu ve işlediğini meydana çıkarmak amaçlarıyla yapılır. Onun için, aşamalı sınıflamanın analiz basamağındaki davranışlar üç alt sınıfa ayrılmıştır. Bunlar, sınıflamanın ana çizgilerinde görüldüğü gibi, öğelere dönük (4.10), ilişkilere dönük (4.20) ve örgütleme ilkelerine dönük (4.30) analiz türleridir.

Bir bilgi alanının analizinde ne kadar ileriye gidilirse, bilgiler o kadar derinleşir. Bu nedenle, sadece bilimlerde değil, hemen her alanda analiz metotları ve alanın analizi öğretilir. Onun için, hem bilgi alanının analizini tanıyabilmek, hem de bir bütünü analiz edebilmek eğitimin önemli amaçlarından biridir.

Analiz yeteneği gelişmiş bir öğrenciden bilişsel ürünlerin analizinde beklenebilecek davranışlar, Bloom ve arkadaşları tarafından şöyle sıralanmıştır.³

A. Analitik bir ölçüt kullanarak bir bilişsel ürünündeki kelimeleri, ifadeleri ve fikirleri sınıflama.

B. Bir metne özgü nitelikleri, 'açıkça ifade edilmemiş' olsa bile, metindeki ipuçlarından çıkarma.

C. Bir bilişsel ürünündeki bağıntılardan ve ölçütlerden sayıltıları, açıkça ifade edilmemiş nitelikleri ve durumları çıkarma.

D. Bir bilişsel ürünündeki düzen ve sırayı öncelik - sonralık, sebep - sonuç ve geçerlik gibi ölçütleri kullanarak ortaya çıkarma.

E. Bir bilişsel ürünün bütününe örüntüsünü tanıma; bütününe dayandığı ilkeleri meydana çıkarma.

F. Bir bilişsel ürünündeki temel görüşü veya görüş açısını ve amacı meydana çıkarma.

Bu listeden de anlaşılacağı gibi, analiz basamağındaki davranışlar, oldukça ileri bir öğrenme seviyesi ve oldukça yüksek zihin yetenekleri gerektirir.

Analiz basamağındaki hedef davranışların gözlenip ölçülmesi de kolay değildir. Uygulamada olduğu gibi, öğrenciden analizi istenecek bilişsel ürün öğrenci için yeni olmalı, hiç değilse önemli kısımları yeni hale getirilmiş olmalıdır. Test ve sınav materyallerinin öğrenci için nasıl «yeni» hale getirileceğine bundan önceki kesimde değinilmişti. Ayrıca, öğrenci, analiz edeceği bilişsel ürüne ula-

³B.S. Bloom, J.T. Hastings, G.F. Madaus (Eds.), *Handbook on Formative and Summative Evaluation of Student Learning*, s. 150 - 192.

şabilmelidir. Öğrenciden Pasteur'ün deneylerinden birini analiz etmesi istenecekse, onun, o deneyi ya önceden dikkatle incelediğini kabul etmek, ya da deneyi anlatan dokümanı soruyla birlikte vermek zorunluğ u doğ ar.

Bir test sorusu analizin sadece bir alt sınıf davranışını ölçecek derecede basit hazırlanabileceğ i gibi, birden fazla davranış ı ölçecek kadar karmaşık da hazırlanabilir. Bilişsel üründe analizi istenecek niteliklerin yeteri kadar gizli olup olmadığı, öğrencinin analizinin doğru olup olmadığı, o konuyu iyi bilen bir kimse tarafından tayin edilmelidir.

Aşağ ıda analiz davranışlarını ölçen bir örnek verilmiştir.

33. Turşu yapmak için tuzlu su hazırlayan bir kimse, sıya, içindeki bir yumurta yüzünceye kadar tuz koymuştur. Bu uygulamada yumurtanın rolü nedir?
- A) Tuzun sudra erimesini sağ lar.
 - B) Suyun sıcaklığını sabit tutar.
 - C) Suyun belli bir tuzlulukta olmasını sağ lar.
 - D) Suyun bulanmasını önler.
 - E) Tuzun dibe çökmesini önler.

(4.20)

(C)

Bir kompozisyonu oluşturan ifadelerin bir kısmı olgusal gerçekler, bir kısmı yorumlar, bir kısmı tahmin ve hipotezler, bazıları da genelleme ve yargılardır. Bu öğelerin tanınmasıyla kompozisyonun anlaşılması kolaylaşır. Bir bütünün öğeleri tanındıktan sonra, onlar arasındaki ilişkiler meydana çıkarılabilir. Bu ilişkiler öğelerin birbirine göre durumları, birbirine yardımları, bütündeki fonksiyonları, v.b. olabilir. Bir kompozisyonadaki ana fikrin, yardımcı fikirlerin, iddiaların, delillerin, sayıtların birbirine ne şekilde bağılı olduğunu meydana çıkarmak 4.20 basamağ ında bir analizdir.

Her bütün veya her sistem, hem yapısal düzeni hem de işleyiş i bakımından, bir takım temel ilkelere dayanır. Analizin en ileri basamağ ında iş te bu ilkeler meydana çıkarılır.

Analiz davranışlarının cevaplı testlerle ölçülmesi zorunlu değildir. Cevaplı testler bazı üstünlükler taşımakla birlikte, kısa cevaplı sorular veya uzun cevap gerektiren yazılı yoklama soruları da analiz davranışlarının ölçülmesinde başarıyla kullanılabilir. Bu amaçla bir paragraf, bir metin veya bir sistem verilir. Öğrenciden, bunun öğeler arası ilişkilerini, yapısal düzenini ve bütünlük ilkelerini bulması istenir.

SENTEZ BASAMAĞINDAKİ DAVRANIŞLARIN ÖLÇÜLMESİ

Parçalardan ve öğelerden, belli bir yapısal düzene veya örüntüye sahip olan bir bütün meydana getirmek, bir sentez işidir. Sentez çoğu halde, çok yönlü düşünme ve yaratıcılık gerektirir. Bunda önemli olan husus sentezin amacını iyi tayin etmek, amaca uygun öğeleri seçip, etkili bir düzen içinde bir bütün haline getirmektir.

Sentezle ilgili davranışlar müfredat programlarında eskiden beri çeşitli şekil ve derecelerde amaç edinilmiştir. Güzel yazma ve konuşma, bir eliş veya sanat eseri meydana getirme sentez gerektiren etkinliklerdir.

Bu basamaktaki hedefler kendi aralarında üç alt sınıfa ayrılmıştır. Sınıflamada «özdeşsiz bir iletişim muhtevası meydana getirebilme (5.10) olarak verilen hedef, öğrencinin herhangi bir alanda kendi ürünü olan bir kompozisyon yazması veya onun gibi bir şey hazırlaması anlamındadır. O halde, bu hedefte açık ifade etme, fikirleri uygun bir sıraya koyma, bir ana fikri yeterli delillerle savunma, vb. gibi kompozisyon davranışları vardır. Bu alt sınıf sadece yazılı ve sözlü kompozisyonu değil, diğer ifade araçlarıyla yapılabilecek (resim, müzik, vb. gibi) kompozisyonları da kapsar.

İkinci alt sınıf (5.20), «bir plan veya işlemler takımı önerisi meydana getirebilme» diye ifade edilmiştir. Bundan da anlaşılacağı gibi, burada hedef, meydana getirilecek son ürün değil, o ürünün planıdır. Bir ev yapmak, şüphesiz, bir sentez işlemidir. Yapılacak bir evin planını çizmek de sentezdir; fakat bu 5.20 basamağında bir sentezdir. Eğitimde plan veya işlemler dizisi halindeki önerileri hazırlayabilmek de önemli hedeflerdendir. Şu örnekler bu noktaya ışık tutar :

- Bir hipotezin sınanması için bir yöntem teklif edebilme.
- Bir deneyin nasıl yapılacağını gösteren bir açıklama hazırlayabilme.

Sentezin üçüncü alt sınıfı «soyut ilişkiler takımı geliştirebilme» (5.30) şeklinde ifade edilmiştir. Bu basamak daha karmaşık ve daha ileri davranışları hedef alır. Pratikteki bir problemi çözmek için onu soyutlayan, yeni kavram ve ilişkiler kurarak genel bir çözüm bulan bir kimse bu basamakta bir davranış göstermiştir. Fen ve matematik derslerinde yeni soyutlamalarla özdeşsiz ürünler meydana getirebilme bu derslerin nihai hedeflerindendir. Aşağıdaki hedefler bunun örnekleridir.

- Eldeki verilere dayanarak bir hipotez kurabilme.
- Bir durum için matematiksel bir model kurabilme.
- Deney ve gözlemlere dayanarak kavramsal bir sistem kurabilme.

Sentez yeteneğinin ölçülmesinde en sık başvuru olan yol, öğrenciden, belli konuda bir sentez istemek ve öğrencinin ürününü puanlamaktır. Zaten, aşamalı sınıflamada sentez, süreçleriyle değil, davranış ürünleriyle tanımlanmıştır. Bu ölçmede dikkat edilecek şeylerden biri problemin (sentez gerektiren durumun) öğrenci için yeni olmasıdır. Daha önce değinildiği gibi, aksi halde öğrenci kendi sentez ürününü değil, başka bir kompozisyonun hatırladıklarını verir. Böyle olunca, o ürün «özdeşsiz» sayılamaz.

Sentez ürünleri puanlanırken de dikkatli olmak gerekir. Sentez gerektiren problemlerin önceden bilinen kesin ve biricik doğru cevabı bulunmaz. Her öğrenci kendine özgü bir cevap meydana getirir. Bu cevap hem öğrencinin kendi ürünü olması, hem de bir sentezde bulunması gerekli nitelikler dikkate alınarak puanlanmalıdır.

Sentez davranışlarının ürün olarak ölçülmesi, bu basamaktaki hedeflerin kompozisyon gerektiren ödev, proje, yazılı yoklama gibi ölçme yöntemleriyle ölçüleceği kanısını doğurur. Bu kanı doğru olmakla birlikte, sentez davranışlarını ölçen seçmeli testler de vardır.⁴

DEĞERLENDİRME BASAMAĞINDAKİ DAVRANIŞLARIN ÖLÇÜLMESİ

Burada değerlendirme terimi bilgilerin, bilgi edinme yöntemlerinin, fikirlerin, çözümlerin, teori ve yapıların ve benzeri bilişsel ürünlerin değeri hakkında bir amaç için yargıya varabilme anlamında kullanılmıştır.

Değerlendirmede daima bir gözlem veya ölçme, bir de ölçüt (kıstas) kullanılır. Bir değer yargısına ulaşma, belli ölçütlerle objektif olarak yapılabileceği gibi, sübjektif ölçütlerle, güvenilirliği olmayan gözlemlerle kişisel yargılara varılarak da yapılabilir. Öte yandan, kişilerin tercihleri, sevdikleri veya sevmedikleri şeyler de kişisel değer yargılarına dayanır. Eğitimde hedef alınan değerlendirme, bu kişisel tercihlerin dayandığı yargılara varabilme değil, objektif ölçü ve ölçütlerle değerlendirme yapabilme yeteneğidir.

⁴Ibid, s. 193 - 223.

Bir bilişsel ürünün değerlendirilmesi, başlıca iki yolla yapılır. Ölçütler ürünün kendisinde mevcut olabilir. Bu takdirde, bilişsel ürünün kendi içindeki ölçütlere göre doğru olup olmadığı, başka bir ifadeyle, bilgi bütününün kendi içinde çelişkiye düşüp düşmediği aranır. Buna «iç ölçütlerle değerlendirme» (6.10) denir. Bazen de, bir bilişsel ürün, dıştaki bir ölçüte vurularak değerlendirilir (6.20). Bu takdirde değerlendirmeyi yapacak kimse, doğru bir değer yargısına götürecek uygun bir ölçüt seçme durumundadır. Meselâ, «nötron» kavramı belli bir atom teorisinin öğelerinden biridir. Nötronun o teo-rideki diğer kavramlardan herhangi biriyle çelişkili olup olmadığını meydana çıkarmak, teoriyi iç ölçütlerle değerlendirmektir. Nötronun gerçekten maddesel tanecik halinde var olup olmadığını meydana çıkarmak veya nötron kavramı ile fizik biliminin ne kadar ilerlediği hakkında bir yargıya varmak ise, dış ölçütlerle değerlendirmedir.

Değerlendirme oldukça karmaşık bir davranıştır. Bir kimsenin başarılı değerlendirme yapabilmesi için daha aşağı basamaklardaki davranışları (kavrama, uygulama, analiz, sentez gibi) iyice geliştirmiş olması gerekir. Onun için değerlendirme ile ilgili hedeflere programlarda oldukça ileri eğitim basamaklarında yer verilir.

Aşağıda sıralanan davranışlar değerlendirmede gözlenebilecek öğrenci davranışlarıdır.

1. Bir bilişsel ürünün iç tutarlılığı ve doğruluğu, ne derece özenle hazırlandığı hakkında değer yargısına varma (iç ölçütler).

2. Bir bilişsel üründe tartışmaların tutarlılığı, kanıtlar, sayıltılar ve genellemeler arasındaki bağıntıların tutarlılığı ve iç düzenin mantıklılığı hakkında değer yargısına varma (iç ölçütler).

3. Bir bilişsel ürün hakkındaki değer yargısının görüş açısını ve değerini tanıma (iç ölçütler).

4. Bir bilişsel ürünü diğer bir ürünle karşılaştırarak değer yargısına varma (dış ölçütler).

5. Verilen ölçütleri veya standartları kullanarak, ya da kendi kendine ölçüt seçerek bir bilişsel ürün hakkında değer yargısına varma (dış ölçütler).

Öğrencide bu davranışların ölçülmesi için şu noktalara dikkat etmek gerekir : a) Değerlendirilecek bilişsel ürün veya problem durumu öğrenci için yeni olmalıdır. b) Öğrenci değerlendireceği ürüne ulaşabilmelidir; değerlendirme sırasında gerekirse baş vurabil-

melidir. c) Değerlendirme işlemi, yukarıda sayılan davranışlardan en az birini gerektirmelidir. d) Öğrenci tarafından varılan değer yargısı, o konudaki bilirkişilerin değer yargılarıyla karşılaştırılmalıdır.

Değerlendirme davranışları çeşitli tiplerdeki soru ve problemlerle ölçülebilir. Aşağıda bir örnek verilmiştir.

34. - 37. sorular için : Balıkların üremeleriyle ilgili araştırma yapan iki öğrenciden Cengiz bir gölden sazan yumurtaları toplayıp bunları 10°C, 15°C, 20°C, 25°C sıcaklıklardaki çeşme suyunda tutmuş ve her kaptan çıkan yavruları saymıştır. Kaya ise 5 ayrı gölden göl suyuyla birlikte sazan yumurtaları almış, bunları 18°C sıcaklıkta tutarak her kaptan çıkan yavruları saymıştır.

- 34. Cengiz'in deney sonuçlarından hangi soruya cevap bulunabilir?
- 35. Kaya'nın deney sonuçlarından hangi soruya cevap bulunabilir?
- 36. Kaya'nın kontrol ettiği, fakat Cengiz'in dikkate almadığı değişken nedir?
- 37. Her iki deneyde hangi değişkenler deneme dışı kalmıştır?
(34 - 36 : 6.10; 37 : 6.20)

Bu örnek yukarıdaki haliyle kısa cevaplı bir yazılı yoklamada kullanılabilir. İstenirse, soruların her biri madde kökünü oluşturacak biçimde seçmeli test yapısına da dönüştürülebilir.

Bu bölümü kapatırken şu noktayı vurgulayalım : Bir sorunun hangi bilişsel düzeyde işleyeceği, hem sorunun yapısına hem de onu cevaplandıracak kişinin geçmişteki öğrenmelerine bağlıdır. Bir kimse için «yeni» olan bir soru uygulama düzeyinde işlerken, aynı problemle daha önce karşılaşmış bir kimse için kavrama veya hatırlama düzeyinde işleyebilir. Böyle olmakla birlikte, öğretmenin aşamalı sınıflamayı öğrenmesi ve sınavlarında kullanması, ona çeşitli yararlar sağlar.

BÖLÜM 11

Duyuşsal Davranışların Ölçülmesi

Duyuşsal alan davranışlarının aşamalı sınıflanışı Tablo 9.2'de gösterilmiştir. Bu sınıflama yapılmadan önce de, müfredat programlarında çeşitli duyuşsal davranışlar hedef olarak yer almıştır. Öte yandan, sadece okuldaki derslerde hedef alınanlar değil, genel olarak insanın duyuşsal davranışları psikolojide incelenir. Psikolojide bu davranışları ölçme metotlarının aşamalı sınıflamadaki gibi değil, «ilgi», «tutum», «tercih», «değer», «kişilik» gibi başlıklar altında incelenmesi gelenekselleşmiştir.

Eğitimde ölçme ve değerlendirme alanının ders kitaplarına bakılırsa, bazılarında, duyuşsal alanda hedef analizinden söz edilmediği, bu alanın hedef davranışlarının ölçülmesi yerine, genel olarak «ilgilerin ölçülmesi», «tutumların ölçülmesi» gibi konuların psikometriden aktarıldığı görülür.^{1,2,8}

Belki bunun bir sonucu olarak, belli de duyuşsal davranışların bilişsel davranışlardan ayrı düşmesi sebebiyle, öğretmenler, bu alan davranışlarının ölçülmesinde oldukça isteksiz görünmüşlerdir. Okullarda duyuşsal davranışlara ayrı bir değerlendirme notu takdir edilmediği gibi, bir dersin alanına giren duyuşsal davranışlardaki gelişme de, çoğu zaman o dersteki başarı notuna katılmaz.

Öte yandan, öğrencilerin çeşitli alanlardaki ilgilerinin, meslek tercihlerinin, çeşitli derslere veya sosyal kurumlara ilişkin tutumlarının, bazı kişilik boyutlarının ölçülmesi birçok Batı ülkesinde gelenekselleşmiştir. Bu tür ölçmeler okullarda görevli psikolojik danışmanlar tarafından, standart testlerle yapılır; ölçme sonuçları ise daha ziyade bireysel rehberlik ve danışma kararlarında kullanılır.

¹R.M.W. Travers, *Educational Measurement*, Chs. 9 - 12.

²J.S. Ahmann and M.D. Glock, *Evaluating Pupil Growth*, Ch. 14.

⁸H.H. Remmers, N.L. Gage, J.F. Rummel, *A Practical Introduction to Measurement and Evaluation*, Chs 10 - 11.

Bu bölümde, bu kitabın genel amacına uygun olarak, ders programlarında ne gibi duyuşsal hedefler bulunduğu ve bu tür hedef davranışların nasıl ölçüldüğü, sadece birkaç örnekle gösterilecektir.

PROGRAMLARDA DUYUŞSAL DAVRANIŞLAR

Hem okulların genel programları hem de derslerin müfredat programları incelendiğinde, genel hedefler arasında duyuşsal alan davranışlarına raslanabileceği gibi, derslerin özel hedefleri arasında da duyuşsal davranışlara raslanır. Aşağıdaki liste, çeşitli kaynaklardan kasıtlı olarak seçilmiş duyuşsal davranışları gösteriyor.

- Sanat eserlerinin değerlerini takdir etme
- Bir mesleğe karşı gerçek ilgi duyma
- Bilim adamlarının buluşları karşısında heyecanlanma
- Toplumsal sorunların bilincine varma
- Demokratik toplum düzeniyle ilgili olumlu tutumlar geliştirme
- Birlikte çalıştığı gruplarda lidere uyma
- Bulunduğu gruplarda grup adına sorumluluk yüklenme
- Bireysel zorluklar karşısında ruhsal dengesini koruma

Daha geniş bir tarama yapılsaydı, bu liste uzar ve kapsamı çeşitlenirdi. Bu kadarıyla bile, duyuşsal davranışların, bireysel ilgilerden kişilik niteliklerine ve toplumsal tutumlara kadar çok geniş bir alana yayıldıkları görülüyor.

Son onbeş-yirmi yıla gelinceye kadar, duyuşsal alan davranışları programlara düzensiz bir biçimde girmiştir. Bu sebeple de, sınıftaki başarının veya programların etkililiğinin değerlendirilmesi amacıyla duyuşsal alan davranışlarını ölçme çabaları yetersiz kalmıştır. Programlarda duyuşsal alan hedeflerine daha açıklıkla yer verilmesinin ve duyuşsal davranışların da ölçülüp değerlendirilmesinin gerekliliği birçok yazar tarafından dile getirilmiştir.^{4,5} Ancak program geliştirme ve değerlendirmenin gereğince ele alındığı hallerde, duyuşsal alan hedeflerine kesinlik kazandırma ve duyuşsal davranışları ölçme çabaları görülmüştür. Sınıftaki değerlendirme veya program değerlendirme amaçlarıyla bu davranışların kullanılabilecek araç ve yöntemlere kapsamlarında önemli yer ayıran bazı eserlere de raslanmıştır.⁶

⁴P.E. Jacob, *Changing Values in College*.

⁵P.L. Dressel, *Evaluation in the Basic College at Michigan State University*.

⁶E.R. Smith and R.W. Tyler, *Appraising and Reporting Student Progress*.

Son on beş yılda, duyuşsal hedeflerin programlarda daha açıklıkla yer aldığı ve duyuşsal davranışları ölçme gayretlerinde önemli ilerlemeler olduğu görölmektedir.⁷ Bu bölümde, sosyal bilgiler ve matematik programlarında yer alan hedeflerden ve onlarla ilgili davranışları ölçen test maddelerinden örnekler verilecektir.

Orta Öğretim Düzeyinde Sosyal Bilgiler Derslerinde Duyuşsal Hedefler

Aşağıdaki hedefler Orlandi tarafından verilmiş bir belirtke tablosunda alınmıştır.⁸

Duyuşsal Davranışlar

1. İlgili Beceriler

— Demokratik katılma : Gruplara katılma (K).

2. İlgili Tutumlar

— İstendik bilişsel davranışlar : İnsan davranışına bilimsel yaklaşım (L). Diğer kişilerin davranışına insancıl bakış (M).
— İstendik toplumsal davranışlar : Farkında olma ve ilgi (N). Sorumluluk kabul etme (O). Politik hayata katılma (P).

3. İlgili Değerler

— Demokratik davranışlar : Temel demokratik değerler (Q).

Burada hedefler K'dan Q'ye kadar harflerle işaretlenenlerdir. Bu davranışların bazıları, grup çalışmalarına katılma (K) gibi, beceriler altında gösterilmiştir. Aslında burada hedef alınan davranışlar, küçük grup çalışmalarında önemli rol oynayan bazı beceri ve tutumlardır. Diğer davranışların çoğu, tutumlar altında sıralanmıştır. Tutumların, bilişsel yönü bulunduğu da kabul edilmiştir; L ve M davranışları aslında bilişsel ağırlığı olan davranışlardır. Q ise değerlerle ilgili bir hedeftir. Bu analiz, bizim kanımızca, özel hedefler basamağında kalmış; gözlenebilir hedef davranışlara indirilememiştir. Aşağıda, K'dan Q'ye kadar, bu hedeflerle ilgili davranışları ölçmede kullanılan araçlardan örnekler verilmektedir.

Gruplara katılma (K) özel hedefine örnek olarak, aşağıdaki davranışların gözlenip ölçülmesi gerektiği söylenmektedir :

— İkna edici konuşma ve hareketlerde bulunma

⁷B.S. Bloom, J.T. Hastings, G.F. Madaus (Eds.), *Handbook ve Formative and Summative Evaluation of Student Learning*.

⁸R.L. Orlandi, «Evaluation of Learning in Secondary School Social Studies», in B.S. Bloom, J.T. Hastings, G.F. Madaus (Eds.), *Handbook on Formative and Summative Evaluation of Student Learning*, s. 488 - 498.

— Grubu bir karara doğru götürme

Böyle davranışların ölçülmesi için, genellikle iki yoldan biri kullanılabilir. Davranış sözel olarak ölçülecekse, öğrenciye, çekışmeli bir konunun tartışıldığı fakat karara veya uzlaşmaya varılamadığı bir olay verilir. Bu olayın ya hikâyesi yazılıp verilir, veya daha iyisi, bir filmi gösterilir. Sonra öğrencilere «Siz olsaydınız şu tarafı nasıl ikna ederdiniz?» «Grubu şu karara vardiirmek isteseydiniz, lider olarak nasıl bir yol izlerdiniz?» gibi sorular sorup, sözlü veya yazılı cevap istenir. İkinci yöntem, öğrencilere gerçekten tartışacakları bir konu verip, onu küçük bir grupta tartışmalarını sağlamaktır. Tartışma devam ederken, gerek grup liderinin, gerekse diğer üyelerinin davranışları gözlenip nicelendirilir.

Tutum adı altında birçok hedef gösterilmektedir. Bir tutumun açığa vurulması halinde ortaya çıkabilecek davranışların hepsi, tutum ölçen araç ve yöntemlerle ölçülebilir. Daha önce de belirtildiği gibi, tutumların ölçülmesi psikometrinin konularından biridir.⁹ Aşağıdaki örneklerde özel hedefler ve ilgili tutumları ölçecek araçlar gösterilmektedir.

Hedef : İnsan davranışına bilimsel yaklaşım (L)

Test maddesi :

İnsan davranışını açıklayabilecek evrensel ilkeler bulmak mümkündür.

a) Şiddetle katılırim.

c) Katılmam.

b) Katılırim.

d) Şiddetle reddederim.

Bu örnekte görülen yapıdaki araçlara Likert tipi tutum ölçekleri denir. Bu ölçeklerin gerek hazırlanması, gerekse puanlarının yorumu, seçmeli testlerinkinden farklı tekniklerle yapılır.

Hedef : Diğerlerinin davranışlarına insancıl bakış (M)

Test durumu : Bütün sakinlerin «beyaz» insanlar olduğu bir mahallede, bir zenci, gazeteye «kiralık» ilânı verilmiş bir apartmandan daire kiralamak ister. Apartman sahibi ise, dairesini bir zenciye vermek istemez.

1. Siz ev sahibi olsaydınız, dairenizi zenciye kiralar mıydınız?

2. Siz ev sahibi olsaydınız, zenciye nasıl davranırdınız?

3. Siz zenci olsaydınız ve apartman sahibi dairesini bir zenciye veremeyeceğini söyleseydi, ona nasıl davranırdınız?

⁹Tutum ölçme araç ve metotları için bakınız .

R.L. Thorndike and E. Hagen, *Measurement and Evaluation in Psychology and Education*, 2nd ed., s. 345 - 347.

J.S. Stanley, *Measurement in Today's Schools*, 4th ed. s. 281 - 287.

H. H. Remmers, *Introduction to Opinion and Attitude Measurement*.

M. E. Shaw and J. M. Wright, *Scales for the Measurement of Attitudes*.

Bu mısalde, öğrenciden, farazi bir durumda nasıl davranacağını yazılı veya sözlü olarak bildirmesi istenmektedir. Bu tür araçlara «açık uçlu» sorulardan oluşan araçlar denir. Açık uçlu sorunun en önemli özeliği, cevabı ima etmemiş olması, cevaplayıcıya cevap bağımsızlığı tanımış olmasıdır. Açık uçlu maddeler, böyle kompozisyon soruları halinde düzenlenebileceği gibi, eksik cümle halinde de ifade edilebilir. Yukarıdaki soruların son ikisi şöyle verilebilirdi :

- 2 a. Ben ev sahibi olsaydım, o zenciye
3 a. Ben o zenci olsaydım, ev sahibine

Yazılı yoklamalarda olduğu gibi uzunca bir yazılı cevap gerektiren, fakat çok özel olmayan ve uzmanlık bilgileri istemeyen sorular da bu amaçla kullanılabilir.

Hedef : Farkında olma ve ilgi (N)

Test soruları :

4. Türkiye'nin bu aylarda karşılaştığı başlıca sosyal sorunlar nelerdir?
5. Son altı aydaki başlıca iç politika olaylarımız hangileridir?
6. Son bir yılda, memleketimizi ilgilendiren başlıca dış politika olayları nelerdir?

Bu örnekte görülen tipteki sorularla, öğrencinin, kendi toplumdaki toplumsal sorun ve olaylarla ne dereceye kadar ilgilendiği; onların esaslarını ne derinlikte öğrenebildiği meydana çıkarılmak istenir.

Sorumluluk kabul etme (O) hedefi altındaki davranışlar da, yukarıdaki gibi, verilen bir durumda ve bir grup içinde, gerçeğe yakın bir rol oynama, seçmeli sorulara cevap verme ve sözlü veya yazılı ifade ile tutumunu açıklama teknikleriyle yoklanabilir. Politik hayata katılma (P) hedefiyle ilgili davranışların, seçmeli test maddeleriyle nasıl ölçüldüğü aşağıdaki örneklerde görülüyor.

Hedef : Politik hayata katılma (P)

Test maddeleri :

7. Siyasi partileri, toplum yararını düşünen kimseler değil, iktidar tukusuyla hareket eden kimseler yönetiyorlar.
a) Şiddetle katılıyorum. c) Katılmam.
b) Katılıyorum. d) Şiddetle reddederim.
8. Milletvekili seçimlerinde oy kullanmak, en şerefli vatandaşlık görevidir.
a) Şiddetle katılıyorum. c) Katılmam.
b) Katılıyorum. d) Şiddetle reddederim.

Sosyal bilgiler derslerinde hedef alınan değerler, bir bakıma duşşsal hedeflerdir. Değerlerin gelişmesi, esas itibarıyla bilissel dav-

ranışlara bağlıdır. Çünkü, demokratik bir toplumda değerler, rasyo- nel bir düşünce ürünü olarak oluşmalı; demokratik yöntemlerde de kullanılmalıdır. Eğer hedef, akılcı, tutarlı değerlerin oluşması ve de- mokratik toplum düzeninin işleyişinde işe koşulması ise; bir öğren- cinin değerlerinin ölçülmesi yetmez. Değerlerin demokratik toplu- mun kabul ettiği değerler olup olmadığının meydana çıkarılmasından başka, öğrencinin bu değerlere akılcı yargılarla ulaşip ulaşmadığı, değerleri kendi içinde tutarlı bir sistem haline getirip getirmediği, nihayet, değer sistemini demokratik toplum düzeninin işleyişinde kullanıp kullanmadığı da meydana çıkarılmalıdır. Bu amaçla, öğret- menle öğrenci arasında karşılıklı konuşma tekniği kullanılabilir. Ya- zılı yöntemler kullanılacaksa, öğrenciden, bir listeden değerleri seç- mesi istenebilir. Öğrenciden, seçtiği değer maddelerini önem sırası- na koyması ve herbiri için seçim ve önem gerekçeleri yazması da istenebilir.

Orta Öğretim Düzeyinde Matematik Derslerinde Duyuşsal Hedefler

Aşağıdaki hedefler Wilson tarafından geliştirilen bir belirtke tablosundan alınmıştır.¹⁰

Duyuşsal Davranışlar

E.O — İlgi ve Tutumlar : E1. Tutum. E2. İlgi. E3. Güdü. E4. Kaygı.
E5. Benlik tasarımı.

F.O — Takdirler : F1. Dış değerlerle takdir. F2. İç değerlerle takdir.
F3. İşevuruk takdir.

Bir takım duyuşsal davranışlar geliştirmek eğitiminin genel amaçları arasında olmakla birlikte, duyuşsal alanda, yalnız bir tek dersin programına düşen özel hedefler de vardır. Yukarıda verilen ilgiler, tutumlar ve takdirler matematik dersiyle ilgili özel hedefler- dir. Bu hedefler diğer dersler için de örnek alınabilecek niteliktedir. Şimdi bu hedeflerin mahiyetini kısaca görelim ve hedef davranışları ölçen araçlardan örnekler verelim.

Tutum (E.1) : Bir kimsenin herhangi bir olay, eşya veya insan grubuna karşı olumlu ya da olumsuz davranış gösterme eğilimine tutum denir. Bir öğrencinin matematik dersine ilişkin tutumu söz ko- nusu olsun. Bu tutum henüz gelişmemiş ise «tarafsız», matematiği

¹⁰J.W. Wilson, «Evaluation of Learning in Secondary School Mathematics», in B.S. Bloom, J.T. Hastings, G.F. Madaus (Eds.), *Handbook on Formative and Summative Evaluation of Student Learning*, s. 643 - 696.

beğenmeyici veya reddedici yönde ise «olumsuz», aksi yönde ise «olumludur.»

Genellikle, olumlu tutumlar söz konusu olan derste başarılı olmak, dersin konularına ilgi duymak, konularının önemini görebilmek, vb. sonucunda oluşur. O derste başarısız olmak, o dersin öğretmenden ceza görmek veya diğer bir acı verici yaşantı geçirmek, vb. de olumsuz tutumları oluşturur. Olumlu tutumlar öğrenmeyi kolaylaştırır; olumsuz tutumlar ise, öğrenmeye ket vurur. Bu sebeple, öğrenme faaliyetlerine rehberlik edilecek durumlarda, derslere ilişkin tutumların ölçülüp değerlendirilmesi gerekli olur.

Tutumların tutum konusuna göre çeşitleri vardır. Gerçekten de tutumlar, tutum konusuyla birlikte adlandırılırlar. «Azınlıklara ilişkin tutum», «fen derslerine ilişkin tutum» gibi ifadeler kullanılır.

Tutumlar doğrudan doğruya gözlenemez. Fakat tutumlar insanın diğer davranışlarıyla açığa vurulur. Meselâ, matematik dersine ilişkin tutumlar, matematiği sevmek, onu önemli bir ders olarak görme, matematikten kaçma, vb. gibi davranışlarda kendini gösterir. Onun için, tutum ölçülürken, bu tür davranışlar gözlenir. Tutumların ölçülmesinde kullanılan yöntemlerden biri, tutum konusunda öğrencinin kendi ifadesini, kendi düşünüş ve duygularını almaktır. Aşağıdaki sorular bu genel yöntemle hazırlanmıştır.

9. - 13. sorular için : Aşağıdaki cümleleri okuyup, her cümlenin sizin duyuşlarınızı ne derecede yansıttığını şu seçeneklerle belirtiniz :

A) Pek çok B) Çok C) Az D) Pek az

9. Matematik ders kitaplarını okumayı severim.
10. Matematik dergilerini okumayı severim.
11. Büyük matematikçilerin hayat hikâyelerini ve buluşlarını okumayı severim.
12. Yeni ve zor problemleri çözmeyi severim.
13. Matematikte yeni konular öğrenmeyi severim.

(E. 1)

14. - 17. sorular için : Aşağıdaki cümleleri okuyup aynı düşüncede olup olmadığınızı şu seçeneklerle belirtiniz :

- A) Tamamen aynı görüşteyim. B) Aynı görüşteyim,
C) Bir şey söyleyemem. D) Aynı görüşte değilim.
E) Hiç de aynı görüşte değilim.

14. En çok sevdiğim konu matematiktir.
15. Matematik eğlencelidir.
16. Matematik sıkıcı bir konudur.
17. Matematik bana kolay gelir.

(E. 1)

İlgi (E.2) : İlgi de başarıyı etkileyen önemli bir değişkendir. Genellikle öğrenciler ilgi duydukları derslere daha çok zaman ayırır, o konuları daha iyi öğrenirler. Matematik konularına karşı aktif bir ilgi, öğrenmenin ilk adımı olduğu gibi, ilginin sürekliliği bir meslek ilgisi-ne ve sonuçta matematikte veya bir uygulama alanında bir meslek seçimine dönüşebilir.

Okul derslerine karşı ilgiler erken başlar. O derste başarılı olma, zevk duyma, o dersle ilgili davranışların müköfatlandırılması gibi etkenlerle ilgiler olumlu yönde gelişir. Bir öğrencinin bütün derslere ve bütün konulara karşı aynı ilgiyi duyması beklenemez. Çoğu halde, ilgilerin farklılığından doğan tercihlerin eğitimde bilinmesi ve öğrenciye rehberlik edilirken kullanılması gerekir. Matematik dersine karşı duyulan ilginin ölçülmesi de bu açıdan önem kazanır.

Öğrencilerin ilgilerini ölçmenin birçok yöntemi vardır. Bir öğretmen, öğrencilerin sınıftaki faaliyetlerinden bile, hangisinin matematiğe karşı ilgi duyduğunu kestirebilir. Bu yöntemlerden biri, öğrenciye tercihlerini sormaktır. Aşağıda bazı örnekler verilmiştir.

18. - 22. sorular için : Aşağıdaki cümleleri okuyunuz. Her cümleyi, tercihlerinizi ifade edip etmediğine göre, şu seçeneklerle cevaplayınız :

A) Evet

B) Bir şey söyleyemem

C) Hayır

18. Bir matematikçi olmak isterim.

19. Matematik'in devamlı olarak kullanıldığı bir işte çalışmak isterim.

20. Matematikten öğrendiklerimi diğer konularda uygulamak isterim.

21. Büyük matematikçilerin buluşlarını nasıl yaptıklarını okumak isterim.

22. Boş zamanlarımda matematiksel bulmacalar çözmek isterim.

(E. 2)

Güdü (E.3) Güdü, kabaca, bir kimsenin davranışlarının gerisindeki gizli güç veya insanı davranışa götüren etmen olarak anlatılabilir. Güdüler, bir bakıma, hareketlerimizin içimizden gelen psikolojik sebepleridir. Matematik derslerinde geliştirilecek çeşitli davranışlar vardır. Bunların geliştirilebilmesi için, herşeyden önce, öğrencinin matematikte hedef alınan bilgileri edinmeye, becerileri kazanmaya «güdülenmiş» olması, yani istekli olması, kendiliğinden merak duyması gerekir.

Güdüleri ölçmenin de çeşitli yöntemleri vardır. Bir insanın davranışlarını gözleyerek, o hareketlerin gerisinde gizli güdüleri, dolaylı yoldan kestirmek mümkündür. Matematiğe karşı iyice güdülenmiş bir öğrenci, matematikle uğraşmaktan çabuk bıkmayacak, ödevlerini tam ve zamanında yapacak, ders dışı faaliyetlerinde çeşitli mate-

matik konularıyla ilgilenecek ve uğraşacaktır. Güdülerini ölçmek için, aşağıdaki gibi sorulardan oluşan testler de kullanılır.

23. Çözdüğünüz bir problemde, çözümünüzün doğru olup olmadığını yoklar mısınız?

A) Pek seyrek B) Ara sıra C) Sık sık D) Her zaman

(E. 3)

24. Bir ay içinde ödev ve normal ders hazırlığı dışında kaç problem çözdünüz?

A) Hiç B) 1 - 4 C) 5 - 10 D) 10'dan fazla

(E. 3)

Kaygı (E.4) : İnsanda kaygı, gelmesi muhtemel bir tehlikeden korkma halidir. Herkes hayatının bir döneminde, şu veya bu sebeple, kaygılanmıştır. Kaygıların en önemli özelliği, insanın normal davranışlarına etki etmeleridir. İnsan kaygılı olduğu zaman davranışları bozulur; öğrenmesi etkililiğini, becerileri mükemmelliğini kaybeder.

Genellikle okuldaki derslerde başarısızlık ceza ve acı ile sonuçlanır. Ceza, ihtar ve kötü not alma korkusu, o dersle ilgili faaliyetler yaklaşırken öğrencilerin kaygılanmalarına yol açar. Matematik iyi öğrenilemediğinden (şüphesiz, iyi öğretilmediğinden de) en çok korku ve kaygı yaratan dersler arasına girmiştir; belki de başta gelmektedir.

Okullarda derslerden doğan kaygıyı azaltmak, o dersin öğretmenine düşen bir görevdir. (Kaygının hiç bulunmaması da istenmez. Çünkü, diğer davranışları bozmayacak derecede bir kaygı güdüleyici etkiler yaratır.) Onun için, duyuşsal alan değişkenleri arasında kaygının da sözü edilmiş; kaygının ölçülmesi için girişimlerde bulunulmuştur. Örnekler :

25. Matematik sınavlarından önce korku ve heyecan duyarım.

A) Pek seyrek B) Ara sıra C) Sık sık D) Her zaman

(E. 4)

26. Bir problemi çözemeyince heyecanlanır, sonraki problemlere hiç el süremem.

A) Pek seyrek B) Ara sıra C) Sık sık D) Her zaman

(E. 4)

Benlik Tasarımı (E.5) : Benlik tasarımı denilen değişken, bir insanın kendini nasıl gördüğünü ifade eder. Benlik tasarımı, psikolojide çeşitli incelemelere konu olmuştur. Bu araştırmalarda iki tür benlik tasarımı söz edilir. Bunlardan biri bir insanın kendini gerçekte nasıl gördüğü (gerçek benlik), ikincisi bir insanın nasıl ol-

mak istediğidir (ideal benlik). Bir öğrencinin kendini, bir matematik öğrencisi olarak nasıl gördüğü (gerçek benlik), özellikle matematik eğitimi ile uğraşanları ilgilendirmiştir.

Matematik konuları karşısında kendini gerçek halinden küçük, yetersiz veya başarısız gören bir öğrenci, genellikle, güdüsünü kaybeder; sonunda gerçekten başarısız olmaya, matematik derslerinde başı derde girmeye başlar. Bu durumun tersi de olabilir. Kendine gerçek değerinden fazla değer veren, zeki olduğunu sanarak gerektiği gibi hazırlanmayan bir öğrenci de başarısızlığa uğrayabilir. Öğrencilerin bu nitelikleri çeşitli öğretmen - öğrenci ve öğrenci - öğrenci ilişkilerinde gözlenip ölçülebilir. Daha önemli durumlarda, öğrencini cevaplarına dayanan testler veya benzeri araçlar kullanılır.

Takdir Davranışları (F.O) : Gerek bilişsel gerekse duyuşsal davranışlar yeterince geliştikten sonra, insanda o konuyla ilgili takdir davranışları oluşur. Bir şeyi beğenmemiz, onu hem bilmemize hem de onun hakkında olumlu tutumlar geliştirmemize bağlıdır.

Bir öğrencini matematiği takdir etmesi, matematiğin bir işe yaradığını, birçok yerde uygulandığını görmesinden (dış değerle takdir, F.1) veya matematiğin kendi içinde, kendi başına bir değer taşıdığını anlamasından (iç değerle takdir, F.2), ya da matematiği başarıyla kullanabilmesinden, matematik yapabilmesinden (işevuruk takdir, F.3) ileri gelebilir.

Öğrencilerin matematiği takdir davranışları sözel olarak ölçülebileceği gibi, matematikle uğraşmaları, meselâ bir problemin çözümünü sınıfa anlatmaları sırasında da gözlenebilir. Aşağıdaki sorular takdir davranışlarının ölçülmesinde kullanılan sorulardan örneklerdir.

27. Matematik nedir?
28. Niçin matematik öğreniyorsunuz?
(F. 0)
29. Cebir günlük hayatta ne işe yarar?
30. Bu hafta matematik bilgilerinizi nelerde uyguladınız?
(F. 1)
31. Matematiği niçin seversiniz?
32. «Salt güzelliği yalnız Öklid görebilmiştir.»
diyen bir kimse ile aynı düşüncede misiniz? Niçin?
(F. 2)
33. Ondalık kesirlerin kareköklerinin nasıl alınacağını, bir örnek üzerinde açıklayınız.
34. Seçmeli bir testte başarınızı şansla nasıl artırabilirsiniz?
(F. 3)

DUYUŞSAL DAVRANIŞLARIN ÖLÇÜLMESİNDE BAZI GÜÇLÜKLER

Bilişsel davranışlar, çoğu halde, öğrencinin zihin ürünü veya zihin becerileri olarak ortaya çıkar; bu sebeple, gözlenip ölçülmeleri bir dereceye kadar kolaydır. Oysa, duyuşsal davranışlar o derece kolaylıkla gözlenemez.

Duyuşsal davranışların ölçülmesindeki yöntemlerden bir grubu, öğrencinin kendi duyuşlarını rapor etmesi esasına dayanır. Test maddeleri, 7 - 8 ve 9 - 17 numaralı maddelerde olduğu gibi, öğrencinin kendi duyuşlarını alnayabileceği ve testi ona göre cevaplandıracağı sayılıtlarına dayanır. Bir öğrenci kendi duyuşlarını anlayıp değerlendiremezse, teste verdiği cevaplar geçersiz olacaktır. Kendi duyuşlarını anlayabilen bir diğer öğrenci, duyuşlarını açığa vurmamak istemez veya kendini toplumda kabul gören duyuşlarla yüklü göstermek isterse, cevabını, gerçekte olduğu gibi değil, başka bir yönde değiştirerek verecektir. Böylece, gerçek dışı olarak cevaplandırılmış bir test de geçersiz olacaktır.

Gerçek dışı cevaplar, sadece test yapısındaki araçlarda değil, öğrenciyle mülakat, kısa bir kompozisyon yazdırma, vb. gibi diğer yöntemlerde de söz konusudur. Öğrenci gerçek duyuşlarını, isterse, bu yöntemlerde de gizleyebilir. Bunun için duyuşsal davranışları ölçen araçların, bilişsel davranışları ölçen araçlara kıyasla, genellikle daha düşük geçerliğe sahip olması beklenir.

Gerçek dışı cevaplar cevaplayıcının test maddesini anlayamamasından doğabilir. Bu halde test puanları tutarsız, güvenilirliği düşük olur. Test maddelerinin sade bir dille yazılması, cevaplamanın nasıl yapılacağının cevaplayıcılara öğretilmesi, test konusunun kuramsal değişkenler yerine somut davranışlara dayanması gibi önlemler bu hayati azaltır.

Cevaplayıcıyı güdüleyen bazı etmenler, onu, kasıtlı olarak gerçek dışı cevaplar vermeye götürür. Kendini daha iyi, sosyal yönden daha beğenilir niteliklerde gösterme güdüsü bu etmenlerin başında gelir. Örneğin, 8 numaralı maddede seçimlerde oy kullanma davranışı konu edilmektedir. Seçimlere katılmanın görev sayıldığı ve bu davranışın beğenildiği bir toplumda, seçimlere karşı olumsuz tutuma sahip bir kimse kendini beğenilir nitelikte göstermek isterse, kendisi için gerçek olan c veya d seçeneklerinden birini değil, a veya b seçeneklerinden birini işaretler.

Kasıtlı gerçek dışı cevaplarla uğraşma, psikometrenin çetin problemlerinden biri olmuştur. Standart envanterlerde «zorunlu seç-

me» yapısındaki maddeler gerçek dışı cevapları azaltır. Ölçme sonuçlarının, cezalandırıcı veya kınayıcı olarak değil, yapıcı ve düzeltici olarak kullanılması; bu amaçlarla kullanılacağıının cevaplayıcılara bildirilmesi de gerçek dışı cevapları azaltabilir.

Duyuşsal alan davranışlarıyla ilgili diğer bir ölçme zorluğu, bu tür araçların kolayca puanlanamamasından doğar. Bilişsel davranışları ölçen testlerin hepsinde, sorulan sorunun bir doğru cevabı vardır. Onun için testlerde hiç değilse o alandaki bilgili kişilerin doğru kabul edebilecekleri doğru cevaplar vardır. Oysa ki, duyuşsal alan testlerinin çoğunda, 25 ve 26 numaralı test maddelerinde olduğu gibi, bir doğru cevap yoktur; her seçenek ayrı bir anlam taşır. Bu tür araçlar, ya önceden kararlaştırılmış ağırlıklarla puanlanır; ya da «en geçerli anahtar» denemelerle bulunur. Söz gelişi, bu maddeler, $A = 1$, $B = 2$, $C = 3$, $D = 4$ gibi itibari ağırlıklarla puanlanabilir.

Duyuşsal alan davranışlarının ölçülmesindeki bu zorlukların yanında, ölçme sonuçlarının yorumunda ve kullanılmasında da zorluklar vardır. Genellikle olumsuz tutumların, ilgisizliklerin ve takdir etmeyişin başarısızlık gibi yorumlanması; davranışları bu yönde olan öğrencilerin şu veya bu şekilde cezalandırılmalarına yol açar. Bu doğru bir kullanma olmaz. Duyuşsal alan davranışları, bu alandaki eksikliklerin ve yanlışlıkların ortaya çıkarılabilmesi için ölçülür. Eksiklikler ve yanlışlıklar bir defa ortaya çıkarıldıktan sonra yapılacak işlem, öğrenciye, onları tamamlama ve düzeltme yollarını göstermek olmalıdır. Diğer davranış alanlarında tamamlayıcı öğretim mümkün olduğu gibi, duyuşsal alan davranışlarının geliştirilmesi ve düzeltilmesi için de tamamlayıcı öğretim mümkündür.

Psiko-Motor Becerilerin ve Ürünlerinin Ölçülmesi

O kul programlarında bilgiye dayanan hedeflerin payı büyüktür. Onun için, programlardaki hedeflerin çoğunda, zihin becerileri gerektiren ve dille ifade edilebilen davranışlar yer alır. Bununla birlikte, okul programlarında, psiko - motor becerilerin ağır bastığı ders ve etkinlikler de bulunur. Müzik, dans, cimnastik, spor oyunları vb. geniş ölçüde beden hareketlerine dayanır. İş dersleri, çeşitli atelye dersleri, laboratuvar çalışmalarını kapsayan dersler de bazı psiko - motor becerilerin geliştirilmesini hedef alır.

Bu bölümde psiko - motor beceri terimi geniş anlamında kullanılmaktadır. Bu terimle, oldukça karmaşık davranışlar zincirinden oluşan ve psiko - motor yönü ağır basan beceriler ifade edilmektedir. Pişano çalmak, yazı makinasıyla yazı yazmak, bisiklet sürmek, yüzmek, voleybol oynamak, duyar teraziyle kütle ölçmek, mikroskop kullanmak, vb. bu tür becerilerdir. Beceri ürünü deyimiyle de, karmaşık bir becerinin icra edilmesi sonunda ortaya çıkan iş ifade edilmektedir. Resim yapmak, yazı makinası ile yazı yazmak, etek dikmek, masa cilalamak, pasta yapmak karmaşık becerilerdir; bu becerilerin icrasıyla ortaya resim, yazılmış metin, etek, pasta, vb. gibi bitmiş bir iş çıkar. Ürünü beceriden ayırmak, ölçme amaçları için gerekli olabilir.

Programlarda hedef alınan hiçbir becerinin, sadece psiko - motor davranışlardan oluşması düşünülemez. Bütün becerilerin bilişsel dayanakları vardır. Birçok beceri de duyuşsal davranışlar doğurur. Fakat, bu bölümde örnekleri verilen becerilerde, davranışların psiko - motor yönü ağır basar. Becerilerin ölçülmesinde de, daha ziyade psiko - motor davranışlara ağırlık verilir.

Okul programlarında psiko - motor becerilere önemli bir yer verildiği halde, ölçme ve değerlendirmede bu becerilere aynı önem verilmemiştir. Buna sebep, psiko - motor davranışları veya beceri ürünlerini ölçmenin oldukça güç oluşu ve bu tür ölçmede kullanılacak araç ve yöntemlerin geliştirilmesinde yetersizliklerin bulunuşudur.

Psiko - motor hedefler programlara girdiğine göre, güç de olsa, psiko - motor becerilerin ölçülüp değerlendirilmesi zorunludur.

Psiko - motor becerilerin ölçülmesinde şu üç genel yöntemden birine başvurulabilir :

1. Öğrenci, hedef alınan beceriyi icra eder; öğretmen, icra sırasında öğrencinin hareketlerini gözler ve icraya puan takdir eder. Bu yöntem, aslında, davranışın gözlenip ölçülmesidir.

2. Öğrenci, hedef alınan beceriyi icra ederek bir beceri ürünü ortaya çıkarır; öğretmen, bu ürünü inceler ve ona puan takdir eder. Bu yöntemde ölçülen şey, ürünün mükemmelliğidir. Fakat aslında, ürüne bakılarak öğrencinin becerisi hakkında bir değer yargısına varılır.

3. Öğrenci, hedef alınan beceriyi icra ederek bir ürün meydana getirir. Öğretmen, hem öğrencinin davranışlarını hem de ürünün oluşmasını, çeşitli basamaklarda gözler ve her basamakta ara puanları takdir eder, sonunda da bir toplam puana varır.

Bu yöntemlerden hangisinin kullanılacağı, hedefin mahiyetine ve ölçmenin amacına bağlıdır. Hedef, öğrencinin bir millî oyunu oynayabilmesi ise, ortaya çıkacak maddi bir ürün bulunmadığından, birinci yöntem kullanılır. Hedef, bir öğrencinin ağaç işlerindeki becerisi ise, ondan bir masa yapması istenebilir. Bu halde, masaya verilecek puanla yetinilecekse, ikinci yöntem kullanılır. Masaya verilecek puan, hedef davranışları yeterince temsil etmiyorsa, işin ara basamaklarına ve öğrencinin işi yaparken gösterdiği davranışlara da puan vermek gerekeceğinden, üçüncü yöntem kullanılır.

Davranışı gözleyip puanlama ve ürünü inceleyip puanlama, iki ayrı yöntem olarak, aşağıdaki kesimlerde ayrıntılarıyla ele alınmaktadır. Daha sonraki kesimlerde davranışın ve davranış ürününün ölçülmesinde kullanılabilecek ölçme araçlarından örnekler verilmektedir.

PSİKO - MOTOR DAVRANIŞLARIN GÖZLEMLERLE ÖLÇÜLMESİ

Programlarda hedef alınan beceriler, karmaşık davranışlar zincirinden oluşur. Bir becerinin ölçülmesi için, bu zincirin önemli halkaları gözlenebilmeli ve gözlem sonuçları nicelendirilmelidir. Bilişsel bir hedefin ölçülmesinde, önce o hedefi oluşturan kritik davranışlar ortaya çıkarılır; sonra, öğrencinin bu davranışları gösterebileceği ölçme durumları yaratılır. Testler, bir bakıma, öğrencinin hedefle ilgili kritik davranışlarını gösterebileceği durumlardır.

Psik - motor davranışların ölçülmesinde de, genel çizgileriyle aynı yaklaşım kullanılır. O halde, bir beceriyi, onu oluşturan davranışları gözleyerek ölçmede izlenecek yöntemin ana basamakları şöyle olmalıdır :

1. **Beceri**yi oluşturan davranışlar analiz edilmeli, bunlar arasından ölçme için gerekli olanlar (kritik davranışlar) saptanmalıdır.
2. Kritik davranışların gözlenmesini ve puanlanmasını kolaylaştıracak bir araç hazırlanmalıdır.
3. Öğrencinin saptanan kritik davranışları icra edebileceği gözlem ve sınav durumları yaratılmalıdır.
4. Öğrenci bu durumlarda davranışı icra ederken gözlenmeli, gözlenen davranışlar puanlanmalıdır.

Davranış analizinin nasıl yapılacağı, daha önceki bölümlerde tartışılmıştı. Becerileri oluşturan davranışların analizi de, o genel esaslara uyularak yapılır. Bununla birlikte, burada bazı noktalara tekrar değinmekte yarar vardır. Beceri davranışlarının analizi, hedeflerle ilişkili olarak yapılmalıdır. Bu analiz, hedefi, davranışlarıyla tanımlayacak kadar ayrıntılara indirilmelidir. Böylece listelenen davranışlar gözlenebilir ve puanlanabilir nitelikte olmalıdır. Hedefe ne derecede ulaşıldığını yoklamak için o hedefi oluşturan davranışların tümünün ölçülmesi gerekmez. Beceriye oluşturan davranışlar zincirinde, ancak kritik davranış halkaları gözlenip ölçülmelidir. Analiz tamamlandıktan sonra kritik davranışlar belirlenmelidir.

Kritik davranışların gözlenip puanlanmasını kolaylaştıracak bir araç hazırlama işi, analiz iyi yapılırsa kolaylaşır. Analiz sonucunda seçilen kritik davranışlar, kullanışlı bir düzenle araçta yer almalıdır. Bir becerini icra edilmesindeki basamaklara göre sıralama, çoğu halde kullanışlı bir araç meydana getirir. Otomobil kullanma becerisinin ölçülmesinde, imtihan komisyonu üyelerinin kullandığı liste böyle bir araçtır. O araçta sürücü davranışları, «motoru çalıştırma», «dönüşler», «park yapma», vb. gibi alt gruplarda gruplanmıştır. Hazırlanan araç, puanlamanın nasıl yapılacağını da açıkça göstermelidir. Yukarıda sözü edilen şoför davranışlarını ölçme aracında, iyi bir sürücüde bulunması gerekli davranışlar sıralanmış, her davranışa belli bir ağırlık verilerek, tam puan 100 ile ifade edilmiştir. İmtihanda hata yapan bir şoför adayının hatalarına göre puan kırılır. Kırılan puanların toplamı tam puandan çıkarılınca adayın puanı kalır. Hatalardan puan kırılan bir araç yapıldığı gibi, icra edilen kritik davranışların mükemmelliğine göre puan takdir eden bir araç da yapılabilir:

Türkçeye çevrilmiş olan bir kitapta, özellikle iş analizi teknikleri örneklerle açıklanmıştır.¹

Öğrencinin kritik davranışları icra edeceği durum, ya gerçek bir durum olarak seçilir, ya da sınav için özellikle hazırlanır. Öğrencilerin voleybol oynama becerileri gerçek bir maçta gözlenebileceği gibi, sınav için düzenlenmiş bir oyunda da gözlenebilir. Gerçek durumlar daha gerçekçi olduğu için tercih edilebilir. Fakat, çoğu halde, gerçek durumlar sınıfa getirilemez; hizmet veya iş üretme amaç olduğu için, gözlem durumu ölçmenin amacına uydurulamaz. Böyle olunca, ister istemez sınav için özel durumlar yaratmak zorunluğu doğar.

Sınav için yaratılacak durumlar hangi nitelikte olmalıdır? Sınav konusu olarak öğrenciye yaptırılacak beceri, gerçek hayattaki beceriyi temsil edecek derecede ona benzer olmalıdır. Sınav durumu gerçek becerinin ancak bir yönünün icra edilmesine elverişli olursa, sınav kapsamı daraltıldığı için, ölçme geçerliği düşer. Bunun tersine, sınav durumu gereksizce uzatılmış ve ölçülmesi istenmeyen davranışlarla karmaşılaştırılmış olursa, ölçme geçerliği yine düşer. Üstelik, bu halde ölçülecek davranışların gözlenmesi de, diğerleriyle karıştığı için zorlaşır. Sınav durumunda icrası istenen beceri, o becerinin kritik davranışlarının tümünü kapsamalıdır. Ayrıca, sınav durumu, kritik davranışların gözlenip puanlanabilmesine de elverişli olmalıdır.

Gerçek hayat becerilerinin bir bölümünü, o becerinin tümünü temsil edecek benzerlikte sınav durumuna getirmek mümkün olmayabilir. Örneğin, bir pilot adayının sınavı gerçek uçakla uçarken yapılırsa, yanlış hareketler çok tehlikeli olabilir. Bazı gerçek durumların hiçbir kısmı sınıfa getirilemez. Böyle hallerde simülasyon (taklid) denilen bir teknik kullanılır.² Örneğin, pilot adayları gerçek uçakla uçarken değil, yerde, uçağı taklid eden makinalar üzerinde imtihan edilir. Eğitimde birçok gerçek araç ve gerçek durum, küçük değişikliklerle sınıfa getirilebilir.

Sınav durumunda öğrencinin kritik davranışlarının gözlenmesi, oldukça zor bir iştir. Ölçmede kullanılacak araç hazırlandıktan sonra, gözlemcilerin bu aracı kullanmada yetiştirilmesi yararlı olabilir.

¹V. C. Fryklund, *Analysis Techniques for Instructors*. (Çeviren : R. Öncül, Öğretmenler İçin Analiz Tekniği.)

²R. Fitzpatrick and E.J. Morrison, «Performance and Product Evaluations», in R.L. Thorndike (Ed.), *Educational Measurement*, s. 237 - 270.

Öğretmen de, aracın hazırlanışı sırasında neleri gözleyeceğini ve nasıl puanlayacağını kararlaştırmalıdır.

Çok önemli ölçme işlemlerinde, icra filme alınır ve yavaşlatılarak gösterilip davranışın ayrıntıları iyice gözlenir. Bir defa yapılan hareketle, özellikle hareket hızlı ve karmaşık ise, bütün davranışları birlikte gözlemek zordur. Böyle hallerde beceri tekrarlatılır ve her seferinde belli bir kritik davranış gözlenir.

Psiko - Motor Becerileri Puanlama Yöntemleri

Psiko - motor becerileri puanlamada kullanılan ölçeklerin başlıca iki türü var... Beden eğitimi dersinde öğrencilerin kasadan atlatıldığı düşünülün. Bu beceri «atlayamadı = 0, atladi = 1» gibi bir puanlama yöntemiyle puanlanabilir. Eğer becerinin mükemmelliği de ölçülecekse, «atlayamadı = 0, çok kötü atladi = 1, kötü atladi = 2, orta = 3, iyi = 4, çok iyi = 5» gibi bir puanlama yöntemiyle puanlanır. Bu iki tür ölçek, verdikleri sonuçların nitelikleri bakımından farklıdır. Bir beceri çok küçük davranış birimlerine ayrıştırılmış ise, 0 - 1 gibi bir puanlama yöntemi yeterli olabilir. Fakat, beceriyi oluşturan davranışlar daha karmaşık ise ve derecelenebiliyorsa, daha fazla bölmeli bir ölçek kullanmak gerekir. Aksi halde ölçmenin doğruluğu azalır ve puanların güvenilirliği düşer.

Bazı beceriler, bilinen fiziksel büyüklüklerle ölçülebilir. Böyle hallerde o büyüklüklerin ölçekleri kullanılır. Koşularda zaman, dakika - saniye - salise olarak, atlamalarda uzunluk ve yükseklik metre - santimetre olarak ölçülür. Aynı düşünce, başka becerilere de uygulanabilir. Bir daktilonun standart bir metni kaç dakikada yazdığı, bir terzinin bir eteği kaç saatte diktığı, vb. böyle puanlar verir.

Bazı hallerde, davranışın tekrar sayısı puan olarak alınabilir. Voleybol oynarken bir oyuncunun kaç defa servis attığı ve bunlardan kaç tanesini kaçırdığı (dışa attıkları ve ağa taktıkları) sayılabilir. Yukarıdaki kasa atlama örneğinde olduğu gibi, davranışın derecelendiği veya sınıflandığı puanlama yöntemleri de vardır. Standart bir becerinin ne kadar zamanda icra edildiği ölçülebileceği gibi, beceri ürünlerinin sayılabileceği hallerde, belli bir sürede ortaya çıkarılan ürün miktarı ölçü alınabilir.

Herhangi bir becerinin ölçülmesinde kullanılacak puanlama yöntemi, daha ziyade becerinin mahiyeti ve gözlenecek davranışların niteliklerince tayin edilir. Ayrıca, puanın hangi amaçla kullanılacağı da dikkatle alınmalıdır. Puanlama yöntemi, değerlendirmenin amacına hizmet edecek ayrıntıda puanlar vermelidir.

Psiko - Motor Becerileri Ölçmede Kullanılan Araçlar

Psiko - motor becerilerin ölçülmesinde kullanılmak üzere, çok çeşitli araçlar geliştirilebilir. Çok kullanılan araçları iki tipte toplamak mümkündür. Bunlardan biri çetele araçları adıyla bilinenler olup, gözlenecek davranışın varlığını «x» veya «evet» gibi bir işaret veya kelime ile gösterilmesine imkân veren, fakat derecelenmesine imkân vermeyen araç tipidir. Diğer, gözlenen beceri veya davranışın mükemmelliğini de göstermeye imkân verdiğinden, **dereceleme** aracı adıyla anılır. Tablo 12.1'de gösterilen çetele, öğrencilerin mikroskop kullanma becerisini ölçmek için hazırlanmış bir araçtan kısaltılarak alınmıştır.³ Aracın aslı, muhtemel olumsuz ve gereksiz davranışları da listelediği için çok uzundur. Öğretmen, öğrenciden mikroskopla bir kültüre bakmasını ister. Öğrenci kültüre bakarken, öğretmen onun davranışlarını gözleyip her davranışın icra sırasını çeteleye yazar. Söz gelişi, bir öğrenci lamı temizlemeden üstüne kültür damlatırsa, çetelede (b.) karşısına 1, kültürü sulandırmadan ve lamel kullanmadan lamı mikroskobun altına koyarsa (g.) karşısına 2 yazar. Böylece, gözlenen iki davranışın sırası kaydedildiği gibi, çeteledeki (a.), (c.), (d.), (e.) ve (f.) davranışlarının icra edilmediği de kaydedilmiş olur.

TABLO 12.1
MİKROSKOP KULLANMA DAVRANIŞLARI ÇETELESİ

Gözlediğiniz davranışların icra sırasını rakamlarla kaydediniz.		İcra sırası
Gözlenecek davranışlar		
a. Lamı temizleme		—
b. Lamin üstüne kültür damlatma		—
c. Kültürü sulandırma		—
d. Lameli temizleme		—
e. Lameli lamin üstüne yerleştirme		—
f. Fazla sıvıyı temizleme		—
g. Lamı mikroskoba yerleştirme		—
h. Işık kaynağını ayarlama		—
i. Aynaları ayarlama		—
j. Diyaframı ayarlama		—
k. Gerekli objektif çevirme		—
l. Gözünü okülere yaklaştırma		—
m. Kaba ayarı yapma		—
n. İnce ayarı yapma		—
o. Lamı mikroskoptan çıkarma		—
p. Objektifi temizleme		—

³R.W. Tyler, «A Test of Skills in Using a Microscope» Educational Research Bulletin. 9 : 493 - 496.

Çeteleler, genellikle, olumlu veya olumsuz davranışların sayısını kaydetmeye yarayan araçlardır. Olumlu ve olumsuz davranışları birlikte listeleyen araçlara pek seyrek raslanır.

Bir çeteleden toplam puan çıkarmak da oldukça teknik bir iştir. Yukarıda kısaltılmış örneği verilen araçta bir değerlendirme çetelesi de vardır. Burada, onun yerine, basit bir dereceleme aracı adapte edilmiş ve Tablo 12.2'de gösterilmiştir.

TABLO 12.2

MİKROSKOP KULLANMA DAVRANIŞLARINI DERECELEME ARACI

Davranışları şu anahtarla derecelayınız :				
0 : Gözlenmedi		2 : Orta		
1 : Zayıf		3 : İyi		
Davranışlar	Derece			
	0	1	2	3
1. Preparat hazırlama	_____	_____	_____	_____
2. Mikroskoba yerleştirme	_____	_____	_____	_____
3. Obj. oküler ve ayna temiz. yapma	_____	_____	_____	_____
4. Işık ayarı yapma	_____	_____	_____	_____
5. Diyafram ayarı yapma	_____	_____	_____	_____
6. Gerekli objektifi seçme	_____	_____	_____	_____
7. Gözünü okülerlere yanaştırma	_____	_____	_____	_____
8. Görüntü ayarı yapma	_____	_____	_____	_____
9. Mikros. bakımını ve korun. yapma	_____	_____	_____	_____

Bir dereceleme aracından toplam puan elde etmek daha kolaydır. Yukarıdaki araçta işaretlenen davranışların dereceleri toplanırsa, bir toplam puan elde edilir.

PSİKO - MOTOR BECERİ ÜRÜNLERİNİN ÖLÇÜLMESİ

Beceri ürünlerinin ölçülmesiyle ilgili bazı teknikler, bu kitabın dördüncü bölümünde, kompozisyon becerilerinin ölçülmesinden söz edilirken incelenmişti. Bir kompozisyonu puanlamada kullanılan analitik teknik, psiko - motor beceri ürünlerinin ölçülmesinde de kullanılabilir. Daha genel olarak, bir iş veya ürün analitik yolla ölçülecekse, yöntem şu basamaklardan oluşmalıdır.

1. İş analiz edilmeli, ürünü oluşturan öğelerin ayrıntılı bir dö-kümü yapılmalıdır. Bunlardan, ölçülmeye değer olanlar saptanmalıdır.

2. Ölçülecek öğelerin nicelenip puanlamasını kolaylaştıracak bir araç hazırlanmalıdır.

3. Ürünün meydana getirilişindeki işlemler de ölçmeye dahil edilecekse, bu işlemlerin kritik basamakları listelenmelidir.

4. İşlemleri, kritik basamaklarda gözlemeyi ve puanlamayı kolaylaştıracak bir araç hazırlanmalı veya işlemler de ikinci maddedeki araca dahil edilmelidir.

5. İşin yapılması sırasında işlemler gözlenerek ve iş çeşitli basamaklarda incelenerek puanlama yapılmalıdır.

İş analizi, davranış analizi kadar önemlidir; fakat oldukça farklı metodlarla yapılır. Genellikle, o işin teknolojisi ve iş planları işin analizini kolaylaştırır.

İş analizi tamamlandıktan sonra, davranış ölçmede kullanılan araçlar gibi, çetele ve dereceleme araçları hazırlanır. Bu araçların hazırlanmasında izlenecek yöntem, ana çizgileriyle davranış gözleme araçlarınıninkine benzer. İşlemlerin veya iş yapılırken gözlenecek kritik davranışların ölçülmesi de, ana çizgileriyle aynıdır.

İş veya ürünü ölçmede, biraz farklı puanlama yöntemleri kullanılır. Çok ayrıntılı bir analiz yapılmış ise, en küçük ayrıntıya 1 puan veren bir yöntem yeterli olur. Fakat, ürünün çeşitli öğeleri ayrı önemde değilse ve her öğeye farklı bir puan verilecekse, kompozisyon puanlamada kullanılan anahtar gibi, ürünün çeşitli öğelerine önemleriyle oranlı ağırlıklar veren bir puanlama sistemi geliştirilebilir.

Ürünü puanlamada diğer ölçekler de kullanılabilir. Ürün, birbirine çok benzeyen birimlerden oluşuyorsa, belli bir sürede meydana getirilen ürün sayısı anlamlı bir ölçü olur. Bir daktilonun dakikada ortalama kaç kelime yazdığı, bir işçinin saatte kaç düğme diktiği, bir marangozun günde kaç sandalye monte ettiği, vb. böyle ölçüler verir. Bazen, ürün yerine, ürün meydana getirilirken harcanan malzemenin ölçüsü de anlamlı bir puan verebilir. Meydana gelen üründen bütünlük veya genel görünüş de önemli ise, öğelere verilen kısmi puanlardan başka, genel görünüme de ayrı puan takdir edilebilir veya toplam puanın bir kısmı bu genel değerlendirmeye ayrılabilir. İşin temizliğine, kısa zamanda bitirilip bitirilmediğine, malzemenin ekonomik bir biçimde kullanılıp kullanılmadığına, vb. de ayrı puanlar takdir edilebilir.

Tablo 12.3'te bir kitap cildini puanlamada kullanılabilecek bir araç, örnek olarak gösterilmiştir. Bu tablonun incelenmesinden anlaşılacağı gibi, bir işe bu tablodaki analitik yöntemle puan verilecekse, işin, yapılırken de gözlenmesi gerekmektedir. Bu tabloda gösterilen toplam puan (100), malzemenin seçimi de dahil olmak üzere, işin yedi niteliği arasında 10, 20, 20 gibi ağırlıklarla dağıtılmış; sonra bu puanlar da kendi içinde ayrıntılara dağıtılmıştır. Bu dağıtım, işin öğelerinin önemi ve nitelikleriyle oranlı olarak yapılmalıdır.

TABLO 12.3
KİTAP CİLDİ İÇİN ANALİTİK ANAHTAR*

İşin nitelikleri	Puan
A. Kullanılan malzemenin işe uygunluğu	10
A 1. Mukavva	3
A 2. Kapak altı kâğıdı	3
A 3. Kapak kâğıdı ve bez	4
B. Dikiş	20
B 1. Düzgünlük	10
B 2. Sağlamlık	10
C. Tutkallama ve kesim	20
C 1. Tutkallama	5
C 2. Kesim gönyesi	5
C 3. Marjinler	5
C 4. Sayfa başı ve sonu	5
D. Sırt kamburluğu	10
D 1. Kitabın hacmine uygunluk	10
E. Kapak	25
E 1. Mukavva parçalarının kesimi	5
E 2. Kapak kâğıdının kesimi	2
E 3. Kapağın yapımı	12
E 4. Kapağın geçirilmesi	6
F. Sıkıştırma ve kurutma	5
F 1. Sıkıştırmada düzgünlük	3
F 2. Kurutma	2
G. Kapak yazıları	10
G 1. Malzemeye uygunluk	2
G 2. Yazı tekniğine uygunluk	2
G 3. Kapak yüzü kompozisyonu	6
Toplam puan	100

* Bu anahtar Gazi Eğitim Enstitüsü öğretmenlerinden sayın Hidayet Tellî'nin yardımıyla hazırlanmıştır.

İŞ ve BECERİLERİ ÖLÇMEDE DİKKAT EDİLECEK ESASLAR

Bu bölümde önerilen yöntemlerin geçerliği, çoğu zaman, tartışmalara ve şüpheyeye konu olur. Geçerliği sağlayan en önemli şeylerden biri, beceri veya ürün analizini dikkatle yapmaktır. Bu analiz sonunda meydana getirilen araç, gerçek beceriyi ne derece iyi temsil ederse, elde edilecek puan da o derecede geçerli olur. Analiz, bilirkişiler yardımıyla yapılmalı, gözlenmesi gereksiz olan ya da mümkün olmayan ayrıntılar listelenmemelidir.

Araçların kullanılış ve puanlanış yöntemleri de geçerliği etkiler. Onun için, aracı hazırlayan kimse ile kullanacak kimse arasında etkileşim kurulabilmelidir. Birden fazla ölçmecinin aynı aracı kullandığı hallerde, ölçme önemliyse ve özellikle aracın kullanılması tecrübe gerektiriyorsa, puanlayıcıların kısa bir çalışmayla yetiştirilmeleri tavsiye edilir.

Gözlem araçlarının güvenilirliğinin, testlere göre daha düşük olduğu da sık sık söylenir. Bu iddia doğru olabilir. Çünkü, gözlem ve puanlama, esasında sübjektif yöntemlerle yapılmaktadır. Gözlemi kolaylaştıran ve puanlamayı objektifleştiren bütün önlemler, araçtan elde edilecek puanın güvenilirliğini artırır.

Akhun ve Gönen, genel olarak atelye derslerinde hedef alınan becerilerin ölçülmesinde, özellikle «iş yapma testi» denilen araçların hazırlanmasında öğretmenin izleyeceği yöntemi gösterirler.⁴ Öğretmen becerilerin ölçülmesinde kullanılan araçları incelemeli; kendi araçlarını hazırlayıp deneme - düzeltme ile geliştirmelidir. Bu tür çalışmalar ölçmeleri objektifleştirmede, öğrencilerin eksikliklerini tanımada, öğretim program ve yöntemini ıslah etmede öğretmene yardımcı olur.

1. Akhun ve Z. Gönen, *Atelye ve Meslek Dersleri Öğretim Metotları*, s. 113 - 124

Sınavlar ve Sınav Planları

Sınavlar öğrenci davranışlarının gözlenip nicelendirilmesi amacıyla hazırlanmış ölçme durumları olarak tanımlanabilir. Gerçekten, sınavlarda amaç, öğrencini davranışlarını, çoğu halde önceden nazırlanmış uyarıcılara gösterdiği tepkileri gözleyerek ölçmektir. Eğitimde sınav yapma zorunluđu, aslında ölçme zorunluğundan doğar. Daha önce değinildiđi gibi, ölçme zorunluđu da kararlara dayanak olacak değeri yargıları için veri toplama ihtiyacından doğar. Bu açıdan bakılınca, sınavlar, eğitimdeki kararlar için veri toplamak amacıyla kasıtlı olarak hazırlanmış ölçme durumlarıdır.

Sınavlarla eğitim sistemi arasında sıkı bir ilişki vardır. Bu ilişkinin eğitim sisteminin yapısına ve uygulanan programların hedeflerine uygun olması beklenir. Bu sebeple, bir sistemde sistemin yapısına ve programlarına uygun düşen sınav yöntemlerinin seçilmiş olması önemlidir. Fakat, bazı sistemlerde, sistemle sınavlar arasındaki ilişkinin iyi çözümlenmediđi, sisteme ve programlarına uygun olmayan yöntemlerin seçildiđi görülür. Çoğu halde buna sebep, sınavlara yüklenen fonksiyonların gerçekçi olmayışı; seçilen yöntemlerin programlara ve öğrencinin düzeyine uygun düşmeyişidir. Sınavların zamanları, süreleri, kapsamı, uygulama yöntemleri ve sonuçlarının kullanılışı biçimleri, aslında, eğitim sisteminin temelinde yatan eğitim felsefesince ve müfredat programlarının hedeflerince tayin edilir. Sistemin yapısı ve işleyişi ile sınavlar arasındaki ilişkiler de bunlardan doğar.

Sınavlara eğitim kararları için bilgi toplamaya yarar ölçme durumları olarak bakılınca, verilecek kararın mahiyetine uygun düşecek bir sınav yöntemi seçmek kolaylaşır. Bu bölümde, başlıca sınav çeşitleri kısaca ele alınmaktadır. Bölümün ikinci kesiminde ise, sınav planlarının nasıl hazırlanacağı açıklanmaktadır.

BAŞLICA SINAV ÇEŞİTLERİ

Sınavların birçok çeşidi vardır. Sınavlar birkaç bakımdan sınıflandırılabilir. Bu kesimde sınavlar kullanılışı amaçlarına, uygulanışı

yöntemlerine ve uygulanma zamanlarına göre üç ölçütte sınıflandırılarak ele alınmıştır.

Kullanılış Amaçlarına Göre Sınavlar

Aşağıdaki paragraflarda, kullanılış amaçları farklı olan (1) seçme, (2) yarışma, (3) yeterlik, (4) sınıflama, (5) tarama sınavları tartışılmaktadır.

Seçme Sınavı : Farklı niteliklerdeki adaylar arasından istenilen nitelikteki bireyleri seçmek amacıyla uygulanan sınavlara seçme sınavı denir. Okullara öğrenci seçmede, bir işe personel almada uygulanan sınavlar, çoğu halde seçme sınavı türünden sınavlardır. Sınav puanlarının okul veya işte başarılı olabilecekleri ayırt etmesi, başarı ihtimali zayıf olanları elemesi beklenir. Bir sınavın seçme görevini iyi yapabilmesi, o iş veya okulda başarılı olabilmek için gerekli bilgi, beceri ve yetenekleri yeterince duyarlıkla yoklayabilmesine bağlıdır. Bu tür sınavların hazırlanmasında, önce, iş veya okulda başarılı olabilmek için gerekli nitelikler analiz edilir. Sonra bu niteliklerin ne gibi cevaplayıcı davranışlarıyla ölçülebileceği saptanır ve bu davranışları kapsayan bir sınav aracı hazırlanır. Sınav uygulandıktan sonra (testin nitelikleri biliniyorsa uygulamadan önce) **minimum kesme puanı** saptanır. Başarısı bu puanın üzerinde olan bireyler seçilir.

Yarışma Sınavı : Bir okula başvuran adayların sayısının okul kontenjanını aştığı hallerde, sınavın okulda başarılı olabilecekleri ayırt etmesiyle yetinilemez. Sınav puanlarının yetenekleri okulda başarı için yeterli olanları birbirinden ayırt etmesi ve kendi aralarında sıralaması da gerekli olur. Bir yarışma sınavı, seçme sınavında olduğu gibi, o işte veya okulda başarı için gerekli bilgi, beceri ve yetenekleri yoklamalı; ayrıca, bu nitelikleri taşıyan bireyleri, daha ince bir ayırımla kendi aralarında sıralamalıdır. Bu nedenle, yarışma sınavları geçerliği yüksek ve oldukça güç sorulardan oluşur. Güçlük düzeyinin iyi kestirilememesi veya yarışmaya giren aday sayısının çok fazla olması hallerinde, sınavın iki basamakta yapılması zorunluğu doğabilir.

Yeterlik Sınavı : Bir bireyin belli bir alanda minimum yeterlik düzeyine ulaşmış olup olmadığını meydana çıkarmak için yapılan sınav, yeterlik sınavı adı verilir. Okullarımızda uygulanan bitirme sınavları, her dersten ayrı ayrı verilen yeterlik sınavlarıdır. Bütünleme ve tek ders sınavları, üniversitelerde uygulanan yabancı dil muafiyet sınavları birer yeterlik sınavıdır.

Yeterlik sınavlarında sınavın kapsamını saptamak zor olmaz. Okullarda bu kapsamı tayin eden şey ilgili dersin programı ve programdaki hedeflerdir. Fakat, o alandaki bilgi ve becerilerin ne kadarının minimum yeterlik sayılabileceğini saptamak gerçekten güçtür. Programlarda en az yeterlik için hangi davranışların, ne derecede öğrenilmesi gerektiği saptanmış ise minimum yeterlik düzeyinin kesilmesi zor olmaz. Aksi halde, öğretmenin minimum yeterlik için gerekli davranışları saptaması ve sınavını bu davranışları ölçecek kapsamda hazırlaması gerekir.

Bir yeterlik sınavında minimum yeterlik düzeyine ulaşmış bireylerin kendi aralarında daha ince bir ayırma tabi tutulması gerekmediği gibi, bu düzeye ulaşmamış olanların da daha ayrıntılı bir sınıflamaya tabi tutulması gerekmez. Böyle olduğu için, yeterlik sınavlarında, minimum yeterlik düzeyindeki sorular sorulur; onlardan daha zor sorular sorulmaz. Daha zor sorular sorulması minimum yeterlik düzeyini hatalı bir şekilde yükseltir. Daha kolay sorular da yeterlik düzeyini düşürür ve sınavı gereksizce uzatır.

Sınıflama Sınavı : Bu sınav türü, adayları niteliklerine göre gruplara ayırmak amacıyla kullanılır. Öğrencileri yabancı dildeki becerilerine göre sınıflayan hazırlık okulu sınavları, adayları çeşitli meslek gruplarına ayıran sınavlar bu türün örnekleridir. Sınıflama sınavlarıyla adayların çok çeşitli yeteneklerinin yoklanması, bir adayın yeteneklerinin, kendi içinde karşılaştırılması istenir. Bu sebeple bu tür sınavlar çok boyutlu olur ve ayrıntılı alt puanlar verir. Bu tür sınavların hazırlanması ve sonuçlarının kullanılması oldukça ileri psikometrik tekniklerin bilinmesini gerektirir.

Öğretmenler de sınıflama problemleriyle karşılaşır. Sınıflama çeşitli yetenekler üzerinde değil, sadece bir yetenekte ve bir boyutta da yapılabilir. Öğrencileri yabancı dildeki kelime bilgilerine göre sıralamak, sonra bu sıra üzerinden «zayıf - orta - iyi» gibi gruplara ayırmak tek boyutlu bir sınıflamadır. Öğrencilere bir dersteki toplam başarılarına göre 0 - 10 arasında notlar vermek de, bir bakıma, onları 11 kategorili bir sınıflamaya tabi tutmaktır.

Sınıflamada kullanılacak bir sınavın yeterince duyar olması, yani güvenilirliğinin yeterince yüksek olması istenir. Sınıflamada kullanılacak sınıf sayısı arttıkça, gerekli duyarlık düzeyi de artar. Öğrencilere 0 - 10 arasında hatasızca puan verebilen bir sınav, 0 - 100 arasında puan vermek için yeterince duyar olmayabilir.

Tarama Sınavı : Bazı hallerde sınavlar, sistemin tümünde programları ve öğretimi değerlendirmek için kullanılır. Bu halde sınavın

ders programının tümünü kapsaması ve o programı uygulayan sınırların hepsinde verilmesi düşünülür. Pratikte bu mümkün olmadığından, sınav dersin tümünü temsil edecek kadar kapsamlı bir örnekleme ile hazırlanır; okulları temsil edecek büyüklükte bir örnekleme de uygulanır. Bu tür sınavlar konu alanını ve sistemi taradığı için bu adı alır. Tarama sınavlarında sorulacak sorular (çoğu halde test maddeleri) çeşitli güçlük derecelerinde hazırlanır; sınavın ortalama gücü yüzdeleri elli civarında seçilir. Tarama sınavlarının sonuçları, çoğu halde, öğretim ve programları değerlendirme, sistemin işleyişini yoklama gibi araştırma amaçlarıyla kullanılır.

Uygulanış Yöntemlerine Göre Sınavlar

Sınavlara uygulanış yöntemleri açısından bakılınca, alışılmış uygulamaların dışında, iki tür sınav görülür. Bunlardan biri açık kitap sınavı, diğer soruları dışarıda cevaplandırılan sınavdır.

Açık Kitap Sınavı : Sınavların çoğu kavrayış, uygulama, sentez, vb. gibi hatırlamadan yukarı düzeydeki zihin davranışlarını yoklamak için kullanılır. Amaç böyle olunca, kolayca hatırlanabilen, basit bilgilere dayanan, kitaplarda kolayca bulunabilen bilgilerin sınavlarda sorulmaması gerekir. Sınav üst düzeyde zihin yeteneklerini yoklayacağına ve sorulan soruların cevaplarının kitaplarda kolayca bulunamayacağına göre, ders kitabı gibi kaynakların sınav sırasında öğrencinin elinde bulunması, soruları cevaplandırırken öğrencinin onlara başvurması sakıncalı olmayacaktır. Böyle bir düşünceyle açık kitap sınavı denilen sınav uygulaması ortaya atılmıştır.

Açık kitap sınavlarının üstünlüklerinden söz edenler vardır. Tussing şu noktalar üzerinde durur : Açık kitap sınavları a) sınav korkusu ve heyecanını azaltır, b) öğretmenleri cevabı kitaplarda kolayca bulunabilen basit sorular sormaktan caydırır, c) koypa olayını ortadan kaldırır.¹

Açık kitap sınavlarının sakıncalarını ortaya çıkarmış araştırmalar da vardır. Kalish, aynı dersi okuttuğu iki şubeden birine «açık kitap», diğerine «kapalı kitap» yönetimiyle aynı sınavları uygulamış; iki yöntemin sınav soruları aynı olduğu halde farklı bilgileri yokladığı sonucuna varmıştır.² Bu araştırmada, ayrıca, öğrencilerin a) kay-

¹L. Tussing, «A Consideration of Open - book Examination,» *Educational and Psychological Measurement*, 11 : 597 - 602.

²R.A. Kalish, «An Experimental Evaluation of the Open - book Examination», *Journal of Educational Psychology*, 49 : 200 - 204

nakları dikkatle incelemedikleri, b) sınava daha az hazırlıkla geldikleri, c) yeterince derinlikte öğrenmedikleri de saptanmıştır.

Açık kitap sınavları birçok bakımdan sakıncalıdır. Her ne kadar aksi önerilse de, öğretmenler geniş ölçüde kitaplarda kolayca bulunabilecek bilgileri yoklarlar. Üstelik, bir öğrencinin kaynak olarak kullanmasını öğreneceği ve gerektiğinde baş vuracağı yayınlar, ders kitabından ve ders notlarından oldukça farklı şeylerdir. Ayrıca, tipik bir sınav süresi, örneğin bir ders saati, ders kitabı, not, kaynak kitap gibi basılı malzemenin incelenmesine fırsat verecek kadar uzun değildir.

Bütün sakıncalarına rağmen, açık kitap sınavları bazı özel hallerde kullanılır. Ezberlenmesi güç ve hatırdaki tutulmasında fayda görülmeyen formül, rakam, cetvel gibi materyalin çok olduğu bir derste, sınav soruları bunlara baş vurulmasını gerektiriyorsa, açık kitap yöntemi uygulanabilir. Bundan başka, sınav sorularının kitap ve kaynaklardan da yararlanılacak bir uygulama olarak düzenlendiği hallerde de açık kitap yöntemi uygulanabilir.

Açık kitap sınavları için hazırlanacak soruların cevapları, sınavda baş vurulacak kaynaklarda bulunmamalıdır. Sınav süresi ancak gerekli kaynaklara çabucak bakmaya ve cevapları yazmaya yetecek kadar olmalı; sınav sırasında okuyup öğrenmeye teşebbüs eden cevaplayıcılar, soruların büyük bir kısmını cevapsız bırakmak zorunda kalmalıdır.

Dışarıda Cevaplandırılan Sınav : Bazı konular, öğrencinin, çeşitli kaynaklara baş vurarak, kendi düşüncelerini de katmasıyla cevaplandırabileceği sorular hazırlanmasını gerektirir. Böyle konular araştırma, irdeleme ve düzenleme gerektirdiği için ve yazma süresi bir sınavın süresine sığmayacak kadar uzun olduğundan, normal bir yazılı yoklama ile yoklanamaz. Böyle hallerde, bir arada toplu sınav yönteminden vazgeçilerek, öğrencinin, soruları dışarıda ve okul saatleri dışındaki zamanda cevaplandıracağı bir sınav yöntemi düşünülmüştür. Bu yöntemin savunucuları, öğrencinin, sınıftaki bir sınav süresinden daha uzun bir sürede kendi başına daha serbestçe çalışıp ne derece mükemmel bir zihin ürünü meydana getirebileceğini yoklamanın, eğitimin amaçları yönünden daha geçerli bir sınav olacağını söylerler. Dışarıda cevaplanan sınav bu özellikleriyle ev ödevlerine ve projelere benzer; fakat, genellikle bir dönem sonu sınavı olarak kullanılır ve sonuçlarına ödevlerden daha çok ağırlık verilir.

Dışarıda cevaplandırılan sınavların en önemli sakıncası, ödevlerde olduğu gibi, öğrencinin cevapları hazırlarken başkalarından önemli yardımlar alması ve kendi ürünü yerine başkasının ürünü getirmesidir. Sorular güçlük ve geçerlikleri bakımından eşdeğer olarak hazırlanır ve her öğrenciye ayrı sorular verilirse, bu sakınca kısmen önenebilir. Ayrıca, dışarıda cevaplanma süresinin kısıtlanması da, başkalarından yardım alma imkânlarını kısıtlayabilir.

Sakıncalarına rağmen bu sınav türü bazı ülkelerde yükseköğretim düzeyinde kullanılır. Aşağı eğitim basamaklarında, ödev ve projeler özel tekniklerine uyularak yaptırılıp puanlanabiliyorsa, açık kitap sınavlarına gerek kalmaz.

uygulama Zamanlarına Göre Sınavlar

Sınavlara uygulanma süreleri ve zamanları açısından bakılıncu, bazı sınavların kısa süreli ve sık uygulanan türden olduğu, bazılarının dönem içinde, bazılarının ise dönem sonlarında veya bitirmede uygulandığı görülür. Bu dört çeşit sınav aşağıdaki paragraflarda tartışılmaktadır.

Kısa Süreli Sınav : Bir ders süresinin ancak 5 - 10 dakikalık bir kısmında uygulanan çok kısa süreli sınavları kullanan öğretmenler vardır. Bir konuyu işlemeden önce onunla ilgili bazı kaynakları okuma ödevi olarak veren bir öğretmen, bu ödevle öğrencinin hem ilgisini uyandırmak hem de konunun kabaca öğrenilmesini sağlamak ister. Bundan amaç, bir bakıma, öğrencilerin işlenecek konuya hazır oluş düzeylerini artırmaktır. Amaç bu olunca, öğrencilerin konunun işleneceği derse hazırlıklı gelip gelmediklerinin yoklanması gerekir. Bunun için dersin başında çok kısa süreli bir sınav yapılır. Genellikle kısa cevaplı bir yazılı yoklama veya 10 - 15 maddelik bir objektif test ile yapılan bu sınavın kapsamı, okuma ödevi olarak verilen konunun ana çizgilerinden oluşur.

Yukarıdaki paragrafta sözü edilen okuma ödevinin yoklanması bir örnek olarak seçilmiştir. Kısa süreli sınavların başka kullanımları da vardır. Bir derste işlenen konunun hemen o derste yoklanmasını gerektirecek durumlarda da kısa süreli sınavlar kullanılır. Böyle bir durumda öğretmen, o derste öğrettiği konunun ne derecede öğrenildiğini, özellikle hangi kısımlarının iyi öğrenilmediğini meydana çıkarmak ve sonraki dersi ona göre planlamak ister. Bir ders süresi zaten kısa olduğu için, dersin sonunda uygulanacak sınavın süresi ancak 5 - 10 dakika olabilir.

Kısa süreli sınavlardan elde edilecek puanların nihai nota katılıp katılmayacağı tartışma konusudur. Bu tür bir sınav, ister dersin başında ister sonunda verilsin, öğrencinin başarısının bir yönünü ölçer. Önce verilen sınav, öğrencinin kendi kendine öğrenebildiği bilgileri, sonra verilen sınav ise öğrencinin o derste işlenen konulardan hemen hatırlayabildiklerini yoklar. Bu iki tür öğrenmeden birincisi eksik bilgilerden oluşur, ikincisi kolayca unutulur. Üstelik, sınav kapsamı çok dardır. Bu nedenlerle, kısa süreli sınavların puanları ancak küçük ağırlıklarla nota katılabilir. Bazı öğretmenler kısa süreli birkaç sınavın puanlarını birleştirip bir tek ara sınavı gibi kullanırlar. Bu yöntem, başka ara sınavları da yapılmak koşuluyla uygun olur.

Ara Sınavları : Bir ders yılı veya sömestr oldukça uzun bir öğrenim süresidir. Öğrencinin başarısı, genellikle böyle bir dönemin sonunda değerlendirilir. Bu değerlendirme, sadece dönem sonundaki ölçümlere değil, dönem içindeki ölçme ve gözlemlere de dayanmak zorundadır. Ayrıca, sınav yönetmeliklerinde ve benzeri kaynaklarda «değerlendirmenin sürekli yapılması» ilkesi bulunur. Bunun için, bir sömestr uzunluğundaki öğretim dönemlerinde ara sınavları yapılması gerekir. Ara sınavlarının sonuçları dönem sonundaki değerlendirmeye katılabileceği gibi, sınavın teşhis gücü varsa, dönem içinde öğretim etkinliklerinin planlanmasında ve özellikle başarısız öğrencilere rehberlik yapılmasında kullanılabilir.

Ara sınavlarının süreleri dönem sonu sınavlarınınkine göre daha kısa olduğu için, sınavın yorucu etkileri azdır. Son sınavın bir günde yapılmasına karşılık, ara sınavları dönem içinde farklı günlere yayılır. Bu özelliğinden dolayı, ara sınavları öğrencinin başarısındaki gelişmeyi de gösterebilir. Öte yandan, ara sınavları öğrenme zamanına daha yakın günlerde yapıldığı için ancak taze bilgileri yoklar; öğrencinin başarısının uzun sürede unutmaya karşı dayanıklı olup olmadığını göstermez. Oysaki, eğitimde amaç, öğrenciye kalıcı bilgiler kazandırmaktır. Ara sınavları, dönem başına yakınlığı oranında daha dar kapsamlı olur ve daha basit öğrenmeleri yoklar. Bu özelliklerinden dolayı ara sınavlarının geçerlikleri, dönem sonu sınavınıninkine oranla daha düşüktür. Aynı nedenlerle dönem başına yakın bir ara sınavı, dönem sonuna yakın bir ara sınavından daha düşük geçerlikte olur.

Ara sınavlarının değerlendirmede ne kadar ağırlık alacağı, çoğu halde tartışma konusu olur. Ağırlıkların tayininde dikkate alınması gerekli olan husus sınav sonuçlarının geçerliğidir. Ara sınavları son sınava göre, genellikle, daha düşük geçerliğe sahip olduğundan.

nihai nota daha küçük ağırlıkla dahil edilir. Bu konuda dikkate alınması gerekli olan diğer bir nokta şudur : Ara sınavları öğrenci başarısındaki gelişmenin ölçüsüdür; son sınav ise öğrencinin eriştiği son başarı düzeyini gösterir. Ara sınavlarında kazanılan puanlar nihai nota katıldığında, son sınavda ancak sınıf geçecek derecede başarılı olmuş bir öğrencinin notu, ara sınavındaki başarısızlığından dolayı sınıfta kalacak düzeye düşer. Ara sınavlarında yüksek puan almış bir öğrencinin notu ise, son sınavdaki başarısının üstüne çıkar. Böyle iki öğrenci arasındaki fark, ara sınavlarına verilen ağırlıklarla oranlı olarak artar. Ara sınavlarına nisbeten az ağırlık verilmesinin savunulmasındaki gerekçelerden biri de budur.³ Şüphesiz, aşağıdaki paragrafta değinileceği gibi, bir tek son sınav da yeterince geçerli ölçümler vermez. Bu nedenle, ara sınavlarını nota hiç katmamak doğru olmaz. Ayrıca, düzgün çalışmak, bugünkü işini yarına bırakmamak, derslerdeki çeşitli çalışmaları zamanında yapmak gibi davranışlar yalnız öğrencinin nihai başarısını artıracığı için değil, eğitimde hedef alınan genel davranışlar olduğundan geliştirilmek istenir. Ara sınavlarına verilecek ağırlık, bu tür öğrenci davranışlarındaki gelişmeyi son nota yansıtır.

Dönem Sonu Sınavı : Dönem süresince yapılan sınavlarla dersin konularının tümü yoklanmış olsa bile, dönem sonunda bir sınav daha yapılması gereği duyulur. Bunun nedenlerinin başında, ara sınavlarıyla unutulabilir bilgilerin yoklandığı, son sınav ile daha kalıcı bilgilerin yoklanacağı düşüncesi vardır. Ayrıca, son sınav ile, öğrencinin ulaştığı son başarı düzeyi yoklanmak istenir.

Not vermede yalnızca öğrencinin ulaştığı son başarı düzeyinin dikkate alınması gerektiği, notun yalnızca son sınav puanına dayanılarak verilebileceği, ara sınavlarının nota katılmasına gerek olmadığı düşüncesini savunanlar da vardır. Şüphesiz, bir son sınav kapsamlı ve geçerli olarak düzenlenebilir. Bu nedenle, son sınavın fazla ağırlık alması haklı olarak savunulabilir. Fakat sadece bir son sınavı dayanılarak not vermek doğru olmaz. Bir tek sınavın başlıca sakıncaları şu noktalarda özetlenir :

a) Bazı derslerde hedef alınan davranışların bir kısmı, bir tek son sınavda ölçülmeye elverişli değildir. Örneğin, «grup tartışmalarında yapıcı katkıda bulunma», «bir işi planlama, yapma ve değerlendirme» gibi davranışlar, öğrencinin iş başında gözlenmesiyle daha iyi ölçülebilir.

³R.L. Thorndike ve E. Hagen, **Measurement and Evaluation in Psychology and Education**, 2nd ed., s. 506 - 507.

b) Son sınav tek başına yeterince kapsamlı, geçerli ve güvenilir olmayabilir. Ara sınavlarındaki puanların nota katılması, notun güvenilirliğini ve geçerliğini artırır.

c) Öğrenci, son sınavın yapıldığı günde, hastalık, rahatsızlık, vb. gibi nedenlerle gerçek başarı düzeyini gösteremeyebilir.

d) Yalnız son sınav sonuçlarına bakılarak not verildiği takdirde, öğrenci düzgün çalışmayabilir. Normal ve devamlı öğrenme yerine, çalışmaları sona bırakma, sınav öncesi yoğun ve yorucu çalışmaya girme gibi alışkanlıklar kazanabilir.

Bitirme Sınavı : Bitirme sınavları bir derste öğrencilere not vermek için değil, öğrencinin okul süresince kazandığı davranışları genel düzeyde değerlendirmek, ülke çapında standart sağlamak, v.b. gibi gerekçelerle düzenlenir. Gerekçelerin geçerliği ve bu nedenle son sınavın gerekli olup olmadığı tartışılabilir.

Bizim sistemimizde olduğu gibi, öğrencide toplu gelişme aranıyor ve bu gelişmenin düzeyi belirlenmek isteniyorsa, bunun en güvenilir ölçüleri ders yılı süresince yapılan sınavlarda sağlanır. Temel eğitim okulları gibi, öğrencinin toplu gelişimine bakılarak sınıf geçme kararı verilen okullarda, ayrıca bir bitirme sınavına gerek yoktur. Programlarında derslere bağımsızlık tanınan, okulu bitirmek için bütün derslerden belirli bir başarı düzeyine erişerek sınıf geçme koşulu aranan okullarda da ayrıca bir bitirme sınavına gerek yoktur.

Bitirme sınavları da bir tek gün ve saatte yapıldığından, dönem sonu sınavlarının sakıncalarını taşır. Sistemimizde ülke çapında uygulanmış olan «olgunluk» ve «Devlet Ortaokul Bitirme» adlarıyla anılan sınavlar, bir tek sınav olmanın yarattığı sakıncaların yanında, sorularının kapsam ve güçlük derecelerinde yerel farklılıklara imkân vermedikleri için sakıncalı bulunmuşlar ve bugün kullanılmaz olmuşlardır.

ÖDEVLER VE PROJELER

Ödevler ve projeler başlı başına sınav yöntemleri değildir; fakat, ödevlerdeki başarının nota katılması gelenekselleşmiştir. Bu nedenle ödevlere bu kesimde kısaca değinilmektedir.

Ev Ödevleri ve Projeler

Öğrencinin bir problemi ders dışında kendi çabalarıyla çözmesi, bir projeyi tasarlayıp sonuna kadar yürütmesi, bir inceleme sonucu-

nu rapor halinde yazması eğitimin genel amaçları içinde hedef alınan davranışlardandır. Ödev ve projelerdeki başarı, öğrenci başarısının önemli bir yönünü yansıtır. Onun için, ödev ve projelerde hedef alınan davranışların ölçülüp değerlendirilmesi istenir. Ödevlerle ölçülen başarının, dışarıda cevaplandırılan sınavlarda olduğu gibi, sınav heyecanından arınmış ve gerçeğe daha yakın bir davranış ölçüsü olacağı görüşü de savunulur. Bu görüşte bir gerçek payı vardır. Fakat, ödev ve projeler öğretmenin gözetiminden uzakta yapıldığından, ödev puanının geçerliğine şüphe düşüren başka sakıncalar doğar.

Ödev puanlarının geçerliğini gölgeleyen en önemli etken, ödevin öğrenci tarafından değil, başka kişiler tarafından yapılmış olması, ya da öğrencinin ödevi yaparken önemli yardımlar sağlamış olmasıdır. Ödevlerde, öğrencilerin birbirinden kopya etmelerini önlemek için, ödevin tamamlanacağı süreyi kısaltmak, her öğrenciye ayrı bir ödev vermek gibi önlemler düşünülebilir.

Aslında, ödevler ve dışarıda cevaplandırılan sınavlar, iyi bir ölçme yöntemi olmaktan ziyade, iyi bir «kendi kendine öğrenme» yöntemi olarak düşünülebilir. Nitekim, ödevlerin ve projelerin öğrenmeye katkıları en çok bu yönleriyle olur. Öyleyse, ödevler puanlanmasa, yani ölçme yöntemi olarak düşünülmesse olmaz mı? Ne yazık ki, ödev ve projeler, puanlanmadığı takdirde öğrenme yöntemi olarak etkilerini kaybeder. Onun için ödevlerin puanlanması ve puanın değerlendirmeye girmesi zorunludur. Şüphesiz, ödev ve projeler de, herhangi bir başarı ürünü gibi puanlanabilir. Ödev puanları küçük bir ağırlıkla nota katılmalıdır.

Grup Projeleri

Genellikle birkaç öğrenciye verilen ortaklaşa yapılacak bir ödev ve grup ödevi veya grup projesi denir. Bireysel ödevler yanında bu çeşit ödevlerin verilmesinde ana amaç, öğrencilerde birlikte çalışma davranışları geliştirmektir. Oysaki, grup projeleri puanlanırken grup üyelerinin ödevin yapılmasına katkıda bulunan ya da engel olan davranışlarına değil, meydana gelen ürüne puan verilir. Bu ölçme metodunun gerisinde, iyi bir ürünün iyi işbirliği sonucunda meydana getirilebileceği sayılısı yatar. Şüphesiz, bu faraziye doğruluğunu kaybettiği oranda grup projesine verilen notun «birlikte çalışma» ölçüsü olarak geçerliği düşer. Grup projeleri, gerçekten işbirliği gerektirecek nitelikte, gruba dahil her öğrencini önemli bir katkısı olmaksızın bitirilemeyecek kapsamda olursa bu sakınca azalır.

Grup ödevlerinde ortaya çıkan diğer bir sakınca, ürüne verilen puanın grubun bütün üyelerine aynen verilmesi zorunluluğundan doğar. Gerçekte, ürünün meydana getirilmesinde her üyenin katkısı aynı derecede olmamıştır. Bu nedenle, grup projelerinden kazanılan puanın bireysel başarı ölçüsü olarak geçerliği şüpheli olur. Grup üyeleri daha önceki başarıları ve yetenekleri bakımından birbirine benzer seçilir ve birlikte çalışmaları sağlanırsa bu sakınca önemli derecede azaltılabilir.

SINAV VE TEST PLANLARI

Sınavlar ve sınavlarda kullanılacak ölçme araçları önceden dikkatle planlandığı takdirde daha iyi yapılabilir. Planlama, bir bakıma, ardışık basamaklarda seçenekleri dikkate alarak karar verme işlemidir. Öğretmen, çeşitli seçenekler karşısındaki tercihlerini, sınavdan önceki planlama basamaklarında gösterir. Bir sınavın planlamasında izlenecek yol, aşağıdaki basamaklarda kısaca açıklanmaktadır.

1. Sınavın amacının saptanması. Bir sınavın amacını saptamak, o sınavdan elde edilecek puanlarla ne gibi kararlar verileceğini saptamaktır. Her sınavın nitelikleri, amacına göre ayarlanır. Bir sınav bir tek amaca hizmet edecekse daha kolay hazırlanabilir. Çok amaca hizmet edecek sınavların kapsamlı olması ve kapsamının çeşitlendirilmesi gerekir.

Öğretmenin sınıfta uygulayacağı sınavların başlıca iki amacı olabilir. Dönem sonu sınavı, bütünleme sınavı gibi sınavlarda amaç, öğrencinin ulaştığı başarı düzeyini kestirmek ve ona göre sınıf geçme kararı vermektir. Dönem süresince verilen ara sınavları, dönem başında verilen yerleştirme ve gruplama sınavları daha ziyade öğrencileri tanımak ve rehberlik etmek amaçları için hazırlanır.

2. Sınavda yoklanacak davranışların listelenmesi. Sınavın temsil edeceği dönemde öğretim etkinlikleriyle geliştirilmek istenen kritik davranışlar, soru örneğine esas olmak üzere saptanmalıdır. Bunun için, dersi iyi öğrenmiş bir öğrenciden beklenebilecek davranışlar kestirilmelidir. Dersi iyi öğrenmiş bir öğrenciyi, dersi yeterince öğrenmemiş öğrenciden ayırt edebilecek davranışlar da saptanmalıdır. Daha önceki bölümlerde tartışılan hedef ve davranış analizi bu iş için de yararlı olur.

3. Sınavın kapsayacağı konuların listelenmesi. Sınavın temsil edeceği dönemde işlenen konular saptanmalı, her konunun ayrıntı-

ları listelenmelidir. Her kritik davranış bir konuda icra edileceğine göre, konuların ayrıntılarıyla listelenmesi ve konularla davranışlar arasında bağlantı kurulması önem taşır.

4. Sınav zamanının seçimi. Her sınavın sonuçlarının ne zaman kullanılacağı, okulun akademik takviminde bellidir. Bu tarihlerle bağlı olarak, her sınavın ne zaman yapılacağı da hiç değilse kabaca saptanabilir. Öte yandan, bazan yönetmelik hükümleri bazan da dersin programı sınavların ne sıklıkta verilmesi gerektiğini belirler. Okullarımızda öğretmenlerimizden yıllık planlar istenmesi gelenekselleşmiştir. Gerçekçi bir yıllık plan ayrıntılı olur ve takviminde sınavların zamanlarını da gösterir.

Sınavların sık sık yapılmasında yarar vardır. Ne var ki, sınırlı olan ders saatlerinin ancak bir kısmı sınavlara ayrılabilir. Bir programda sınavlara ayrılacak zaman, programın yapısına ve öğrencinin eğitim düzeyine göre değişir. Bir sömestride sınavlara harcanacak zaman, ortaöğretim düzeyinde, sömestr ders saatları toplamının yüzde onundan az veya yüzde yirmisinden çok olmamalıdır. Sınav planında sınavın yapılacağı tarih ve ders saati de gösterilebilir.

5. Sınav süresinin saptanması. Okullardaki ders saatları kırk beş dakika kadardır. Bir sınavın bir ders saatında cevaplandırılacak uzunlukta olması, yönetim açısından birçok kolaylık sağlar. Ders saatinin önemli bir kısmını alan fakat tamamını doldurmayan bir sınav, zaman israfına veya cevaplama süresinin gereksizce uzatılmasına yol açar. Ders saatına sığdırılamayan bir sınav da derslerin birleştirilmesi zorunluğunu doğurur. Süresi bir ders saatiyle sınırlı olmayan sınavlar da düşünülebilir; bu tür sınavlarda sürenin tayininde dikkate alınması gereken en önemli nokta, sınavın aşırı yorgunluğa sebep olacak kadar uzun tutulmamasıdır. Plan, sınav süresini dakika olarak göstermelidir.

6. Sınav aracı için belirtke tablosu hazırlanması. Yukarıdaki maddelerde sözü edilen kritik davranışlar listesi ile konular listesi bir araya getirilerek, her ikisinden dengeli bir örneklem yapılmalıdır. Örneklemde kritik davranışların ağırlıkları yüzde olarak gösterilmeli; her davranışın hangi konuda icra edileceği de açıkça belirtilmelidir. Kritik davranış yüzdeieri belirtke tablosunun bir boyutunda, konuların ağırlıklarını gösteren yüzdeler de diğer boyutunda gösterilmelidir. Böylece, belirtke tablosunun gövdesi, belli bir konuda belli bir kritik davranışın hangi oranda sınava dahil edileceğini gösterir.

7. Sınavda kullanılacak madde tiplerinin saptanması. Belirtke tablosundaki konu ve davranışlar, sınavda kullanılabilen madde türlerini kabaca belirler. Ayrıca, sınavın amacı ile sınav süresi de kullanılabilen soru türlerini sınırlar. Fakat, çoğu halde, aynı davranış birkaç tür madde ile yoklanabilir. Bir davranış istenilen düzeyde yoklayabilen maddelerden daha geçerli ve daha güvenilir olması beklenen tip seçilmelidir.

Bir tek sınavda çok çeşitli yapıda maddeler kullanılmamalıdır. Böyle bir uygulama, cevaplayıcı görevlerinin sık sık değişmesini gerektireceğinden, sınavın geçerliğini düşürebilir; ayrıca, puanlama ve analiz güçlükleri de yaratabilir. Testler ve kompozisyon gerektiren yazılı yoklama soruları aynı sınavda birlikte kullanılmamalıdır. Aynı derste hem bir test hem de bir yazılı yoklama uygulamak zorunluğu doğarsa, bu sınavların herbirine ayrı cevaplama süreleri verilmelidir.

8. Sınavda kullanılacak madde sayısının saptanması. Bir sınavda kullanılabilen soru sayısı, maddelerin yapısına, cevaplayıcıların düzeyine ve sınavın süresine bağlı olarak saptanır. Öğretmenler yazılı yoklamalarda öğrencilerin yazma hızlarını, objektif testlerde okuma ve cevaplama hızlarını kestiremediklerinden hata yaparlar. Çoğu halde, sınav süresine sığmayacak kadar çok soru sorulur. Bir sınavın ilk uygulandığında, soru sayısı ancak kabaca kestirilebilir; fakat, sorular daha önceki yıllarda denenmiş ise daha kesin tahminler yapılabilir. Toplam madde sayısı saptandıktan sonra belirtke tablosuna dönülür; toplam madde sayısı, tablonun yüzdelerinden hareket edilerek, hedef davranışlar ve konular arasında dağıtılır.

9. Sınavın ortalama gücünün ve madde güçlükleri dağılımının saptanması. Bir sınavın ortalama gücü, o sınavdan elde edilebilecek en yüksek puanın, ortalama olarak yüzde kaçına ulaşıldığını gösteren orandır. En çok 50 puan kazanılabilecek bir sınavda sınıfın ortalaması 30 ise, sınavın ortalama gücü 0,60 olur. Bu oran, sınavın amacına göre ayarlanır. Bir derste not vermek için uygulanacak sınavlarda ortalama güçlük 0,50 civarında olmalıdır. Bir tek maddenin gücü de, o maddeye doğru cevap veren cevaplayıcıların yüzdesiyle ölçülür. Bir sınavda kullanılan soruların madde güçlüklerinin ortalaması, o sınavın ortalama gücünü verir. Bu sebeple, ortalama gücü 0,50 olacak bir sınav, ortalamaları 0,50 olmak koşuluyla, çeşitli güçlükteki maddelerden oluşabilir.

Öğrencileri «yeterli - yetersiz» gibi sadece iki gruba ayıracak bir sınavın bütün sorularının «yeterli» sayılabilecek güçlük düzeyinde olması istenir. Öte yandan, not verme gibi çok kategorili bir sınıf-

lama için kullanılacak bir sınavın soruları ise, çeşitli güçlüklerde olmalı; fakat, güçlük derecelerinin ortalaması 0,50 civarında bulunmalıdır.

Sınav planı hazırlanırken madde güçlükleri dağılımı ve testin ortalama güçlüğü için gerekli koşullar saptanmalı; madde yazma ve madde seçmede bu koşullara uyulmaya çalışılmalıdır.

10. Madde yazma ve seçme. Sınav planında soruların kimin tarafından yazılacağı, yazılmış sorulardan sınava dahil edileceklerin hangi yöntemle seçileceği kısaca belirtilmelidir. Öğretmen kendi sınıfında uygulayacağı sınavlar için soru yazma ve seçmede, aynı dersi okutan diğer öğretmenlerden yararlanabilir. Birbirine yakın dersleri okutan öğretmenler arasında, bu konuda işbirliği kurulabilir.

11. Test düzeni ve çoğaltma. Testin nasıl düzenleneceği, sınav ve test yönergelerinin esasları planda kısaca gösterilir. Sınav araçlarının nasıl çoğaltılacağı da kısaca belirtilir. Sınav gizliliğini korumak için, sınav araçlarının yazımı, çoğaltılması ve saklanması safhalarında alınacak önlemler ana çizgileriyle tasarlanır.

12. Sınavı uygulama yöntemi. Sınavın hangi koşullarda, nerede, kimler tarafından ve kimlerin yardımıyla uygulanacağı planda belirtilir. Uygulamanın iyi planlanması, sınavdan önce yapılacak hazırlıkların tasarlanmasına ve gerekli önlemlerin alınmasına, sınavın güvenle uygulanmasına yardımcı olur.

13. Puanlama yöntemi. Sınav cevaplarının nasıl puanlanacağı ana çizgileriyle saptanır. Testlerde cevapsız bırakılmış maddelere nasıl bir işlem uygulanacağı, şans başarısı için düzeltme yapılıp yapılmayacağı planda belirtilir. Diğer sınavlar için puanlama anahtarının esasları da planda yer almalıdır.

14. Sınav sonuçlarını kullanma yöntemi. Sınavın amacına hizmet edebilmesi için, puanların nasıl kullanılacağı planda belirtilmeli; puanlar üzerinde yapılacak istatistiksel işlemler, değerlendirmede kullanılacak ölçüt ve yöntem saptanmalıdır.

Öğretmenin bir yandan kendi konu alanındaki, diğer yandan ölçme ve değerlendirme tekniklerindeki yeterliliği, gerçekçi ve kullanışlı sınav planlarının hazırlanmasında önem taşır. Ayrıca, yapılmış planları incelemek, benzerlerini hazırlamak da öğretmene tecrübe kazandırır. Özçelik, Fen Bilgisi, Biyoloji, İngilizce ve Bıçkı - Dikiş derslerinden ayrıntılı test planı örnekleri vermiştir.⁴

⁴D.A. Özçelik, *Okullarda Ölçme ve Değerlendirme*, s. 104 - 116.

Ölçme Araçlarını Çoğaltma, Uygulama ve Puanlama

B u bölümde testlerin ve benzeri ölçme araçlarının düzenlenmesi, çoğaltılması, uygulanması ve puanlanması gibi işlemler ana çizgileriyle tartışılmakta; bu işlemlerin hatasızca yürütülebilmesi için tavsiyelerde bulunmaktadır.

TEST DÜZENİ ve ÇOĞALTMA

Bir yazılı yoklama için hazırlanmış bir kaç soru veya bir test için yazılmış birçok maddede bir araya getirilip çoğaltıldığında bir ölçme aracı meydana getirilmiş olmaz. Böyle bir sorular grubuna ölçme aracı diyebilmek için soruların başında, hiç değilse cevaplamanın nasıl yapılacağını ve cevaplama süresini belirten bir açıklama bulunmalıdır. Özellikle testlerde, maddelerin test içerisinde sıralanışı ve kâğıt üzerindeki düzeni, harf tipi, cevap kâğıdı ve cevapların kayıt şekli, okuma ve cevaplama işlemleri açısından önem kazanır. Bu kesimde, daha ziyade testler göz önünde tutularak, düzenleme ve çoğaltma işlemleri kısaca tartışılmaktadır.

Test Düzeni

Yazılı yoklamalar gibi birkaç maddeden oluşan ölçme araçlarında düzenleme, gerekli olmakla birlikte, önemli problemler doğurmaz. Fakat, testler gibi çok sayıda maddeden oluşan araçlarda okuma yükü fazladır; iyi bir iç düzenle, algılama kolaylaştırılıp okuma yükü azaltılabilir.

Cevaplayıcılar İçin Yönerge : Cevaplayıcıların testi nasıl cevaplandıracaklarını gösteren talimata **test yönergesi** veya **test açıklamaları** denir. Cevaplayıcının yapacağı cevaplama ve cevap kaydı işlemleri bütün cevaplayıcılarca bilinen şeyler ise, kısa ve basit bir yönergeyle yetinilebilir. Fakat, test, cevaplayıcılar için yeni yapıda maddelerden oluşuyor veya cevaplama karmaşık bir işlemi gerektiriyorsa, ayrıntılı ve açıklamalı bir yönerge verilmelidir. Bu yönergede bulunması gerekli noktalar şöyle sıralanabilir :

1. Testin amacı kısaca belirtilir.
2. Testteki soruların nasıl cevaplanacağı bir paragrafla açıklanır; gerekli hallerde örnek verilir.
3. Cevapların nasıl kaydedileceği, gerektiğinde nasıl değiştirileceği belirtilir.
4. Cevaplayıcının, doğru cevaptan emin olmadığı maddelerde nasıl davranacağı açıklanır.
5. Puanlama yönteminin cevaplayıcı için önemli yönleri belirtilir. Meselâ, seçmeli testlerde birden fazla cevabın işaretlenmesini önlemek için «Birden fazla cevap yeri işaretlenmiş maddeler yanlış sayılır.» gibi bir açıklama verilir.
6. Testteki madde sayısı ve cevaplama süresi bildirilir. Başlama ve bitirme emirleriyle ilgili bir açıklama verilir.
7. Test araç ve gereçlerinin kullanılması ve geri verilmesiyle ilgili açıklamalar verilir.
8. Cevaplama süresince sorulacak soruların cevaplandırılmayacağı belirtilir.

Yönerge çoğu hallerde testin başında yer alır. Farklı yönergeler gerektiren bölümlerden oluşan bir testte ise, başa genel bir yönerge, bölüm başlarına özel yönergeler konulması, çeşitli yararlar sağlar.

Test yönergesi cevaplayıcıya ne yapacağını anlatmak, bazı hallerde öğretmek için verilir. Bu sebeple, test yönergesi hiçbir şüpheye yer kalmayacak açıklıkta yazılmalıdır. Öte yandan testlerde okuma zamanı alıcı olduğundan, test yönergesi, açıklığı kaybedilmeden mümkün olduğu kadar kısa yazılmalıdır.

Testlerde Sayfa Düzeni : Test düzenlemede dikkate alınacak birçok nokta vardır. Bunların hemen hepsi testin cevaplayıcılarca kolay algılanabilmesi ve hızlı okunabilmesi esasları için düşünülmüştür. Testlerde sayfalar pek büyük olmamalıdır. Yazılı sayfa boyutları bu kitabinkinden büyük testler, satırları uzun olacağı için okuma güçlüğü yaratır. Daha büyük sayfalarda ve bazı madde türlerinde çift sütunlu sayfa düzenine gidilir. Maddelerin test içindeki yeri kesinlik kazandıktan sonra, herbirine sıra numarası verilir. Numara sırasının bozulmaması ayrı cevap kâğıtlarına cevap kaydını kolaylaştırır. Maddelerin sayfa içindeki dağılımında, maddeler ara-

sında birbirinden kolayca ayırt edebilmelerini sağlayacak kadar aralık bırakılmasına ve bir maddenin iki sayfaya bölünmemesine dikkat edilmelidir. Seçmeli testlerde, madde seçenekleri, birbirinden ayırt edilebilecek ve hepsi birden algılanabilecek bir düzende sıralanmalıdır. Alt - alta sıralama her zaman kolay algılanabilen bir sıralama değildir. İki - üç satıra yayılmış seçenekler bazı hallerde daha iyi olur. Seçmeli yapıda olmayan test maddelerinin düzenlenmesinde de, madde yapısına uygun bir sıralamaya gidilmelidir. Meselâ, kısa cevaplı maddelerde cevabın yazılmasına yetecek kadar yer bırakmak ve cevap yerlerini cevap kâğıdının bir yanında toplamak gibi kolaylıklara gidilebilir.

Yazı Karakteri : Okullarda hazırlanan testlerin hemen hepsi yazı makinası ile yazılıp basit tekniklerle çoğaltılır. Teksir kâğıdı ile çoğaltma yapılacaksa, mumlu kâğıdı iyi kesen ve harf karakteri okunaklı olan bir yazı makinası kullanılmalıdır. Optik kopya ve küçültülerek kopya imkânları olan hallerde, orijinal testin büyük karakterde yazılması, sonra fotokopi sırasında yüzde yirmi kadar küçültülmesi testin okunaklılığını artırabilir. Elle yazılan materyal, ister mumlu teksirde olsun, ister ispirtolu çoğaltma tekniklerinde olsun, çoğu halde okunaksız kopyalar verir.

Kontrol ve Düzeltme : Testlerde bazen bir tek harfin yanlış yazılmış olması bile, bir sorunun iptaline sebep olabilir. Çoğaltma tamamlandıktan sonra yakalanan bir hata kolayca düzeltilemez. Onun için, orijinal kopya, çoğaltmaya girmeden önce birkaç defa kontrol-dan geçirilmelidir. Orijinal kopyada yapılacak düzeltmeler teknik esaslara uyularak yapılmalı; okunaklılığı kayboymuş düzeltmelerden kaçınılmalıdır.

Düzeni Farklı Formlar : Çok kalabalık sınıflarda ve genel olarak uygulama yeri sıkıntısı çekilen hallerde, cevaplayıcıların birbirlerinden yararlanmalarını önlemek için, bir tek testten iç düzenleri farklı birkaç form türetilir. Böyle hazırlanan formlara **sözde eş formlar** denir. Sözde eş formlar, aynı sorular farklı sıralarda sıralanarak; bazen de, her maddenin seçeneklerinin madde içindeki sırası değiştirilerek hazırlanır. Uygulamada, farklı formların arka arkaya ve yan yana dağılışı düzenlenirse, cevaplayıcıların kopya yapmaları oldukça güçleşir. Fakat, sözde eş form hazırlamak, test düzenleme ve çoğaltma işlemlerini hemen hemen iki katına çıkarır. Üstelik, sorular aynı olsa bile, düzeni farklı test formları, sıra farklarından doğacak hatalar sebebiyle aynı puanı vermeyebilir. Çok zorlayıcı bir sebep olmadıkça, sözde eş form hazırlamaya gidilmemelidir.

Çoğaltma : Şüphesiz, çoğaltma işlemlerinin teknik yönleri vardır. Bu konudaki zorluklar, önemine göre, teknik elemanlardan yardım sağlanarak aşılabılır. Burada üzerinde durulması gereken önemli bir nokta, çoğaltılacak malzeme test olduğu için, çoğaltma işlemleri sırasında gizliliğin korunmasıdır. Bunun için, çoğaltma işlemi yetkili bir kimsenin gözetiminde yapılmalı; bozuk kopyaların tümü okunamayacak bir biçimde yok edilmeli; kusursuz kopyalar dikkatle sayılıp numaralanmalıdır. Ancak böylece, bazı çalınma olayları ortaya çıkarılabilir.

UYGULAMA

Testler, sınavlar ve benzerleri, aslında cevaplayıcıların belirli davranışlarını ortaya çıkarıp nicelendirmek için hazırlanmış ölçme durumlarıdır. Ölçmenin geçerliği için sadece aracın geçerli kapsam ve yapıda olması yetmez. Uygulama da ölçülecek davranışların serbestçe ortaya konmasına imkân verecek ve nicelendirmesini kolaylaştıracak şartlarda yapılmalıdır. Bu ise uygulamadan önce hazırlık, uygulama sırasında disiplin ve uygulamadan sonra kontrol gerektirir.

Uygulamadan Önceki Hazırlık : En basit bir sınav bile uygulamadan önce bazı hazırlıklar gerektirir. Bu hazırlık şu noktalarda özetlenebilir :

1. Uygulama araç ve gereçleri hazırlanır. Yeter sayıda test veya teksir edilmiş sorular, müsvedde kâğıdı, cevap kâğıdı, kalem, silgi ve gerekli diğer malzeme sınav saatinden önce sınav yerinde hazır bulundurulur.

2. Uygulama yeri hazırlanır. Bu yer iyi havalandırılmış ve aydınlatılmış, gürültüden uzak, öğrencilerin kopyaya imkân vermeyecek bir düzenle oturtulmalarına yetecek kadar büyük bir dersane veya salon olmalıdır. Okullarımızdaki sınıflar genellikle öğrenci sayısına göre küçük yerlerdir. Bu gibi hallerde okulun daha büyükçe bir yeri sınav salonu olarak düzenlenir. Dershanenin dar olduğu hallerde, gözcü sağlanarak sınıfın bölünmesine gidilebilir.

3. Öğretmenin tek başına yürütemeyeceği sınavlarda görev alacak uygulayıcı ve gözcüler önceden belirlenir ve görevleri kendilerine anlatılır.

4. Öğrenciler uygulamaya hazırlanır. Bu hazırlık öğrencinin sınavda nasıl davranacağını ve zamanını nasıl kullanacağını öğretmekten, yanında bulunduracağı malzemeye kadar çeşitli şeyleri kap-

sar. Öğrenciler sınavda bilgi ve yeteneklerini en yüksek seviyede ortaya koyabilecek kadar güdülenmeli; fakat, sınavın zorluğu, sınav sonuçlarının öğrencinin eğitim hayatına etkisi gibi şeylere değinilerek aşırı kaygı yaratmaktan sakınılmalıdır.

5. Uygulamadan önce, uygulama sırasında ve uygulamadan sonra yapılacak işleri sırasıyla gösteren bir çetele hazırlanmalıdır. Bu çetelede gösterilen işler tamamlandııkça işaretlenirse, uygulama hazırlıklarının eksiklikleri zamanında görülüp önlem alınabilir ve uygulamanın hatasızca yürütölmesi sağlanabilir.

Bir salondaki uygulamanın sorumlusu uygulayıcıdır. Uygulama sırasında yapılacak işler şu noktalarda toplanabilir :

1. Öğrenciler, sınavın başlamasından kısa bir süre önce sınav salonuna alınıp uygun bir düzende oturtulurlar; gerekirse kimlik kontrolü yapılır.

2. Cevap kâğıtları dağıtılır. Öğrencilerden ad, yaş, sınıf, vb. gibi bilgiler toplanacaksa, bu bilgiler testler dağıtılmadan toplanır.

3. Testler sırasıyla dağıtılır. Dağıtım sırasında, erken başlama, kopya, test çalma gibi olaylara karşı önlemler alınır.

4. Cevaplamanın başlama ve bitiş saatları belirlenir. Cevaplama başlatılır.

5. Uygulama sırasında sorulacak sorular cevaplandırılmaz. Öğrencilerin bilgi alıp verme gibi bahanelerle konuşmalarına izin verilmez. Uygulama disiplini bozucu ve sakın çalışmayı engelleyici her türlü hareket yasaklanır.

6. Gözcülerin öğrencilerle bireysel konuşmalarına izin verilmez.

7. Cevaplama süresi dolmadan sınavı bitirenlerin test malzemeleri kontrol edildikten sonra salondan çıkmalarına izin verilir. Fakat, cevaplama süresinin sonuna doğru erken çıkanların sayısı kabarcacağı için, teker teker çıkmaya izin verilmez. Dışarıya çıkış emri verilmeye kadar yerlerinden kalkmamaları öğrencilere duyurulur.

8. Süre dolduğunda cevaplama durdurulur; önce cevap kâğıtları sonra testler sırayla toplanır. Testlerin ve ilgili malzemenin tamam olup olmadığı kontrol edildikten sonra öğrenciler salondan çıkarılır.

Uygulama ve test malzemelerinin ilk kontrolü tamamlandıktan sonra testler ve cevap kâğıtları üzerinde daha ince bir kontrol ge-

rekebilir. Bunda testlerin numaralanmış olmasının ve uygulama baş larken test numarasının cevap kâğıdının belli bir yerine yazdırılma sının yararı vardır. Böylece testin tamamı veya bir parçasının çalın- mış olması halinde, fail sonradan bulunabilir. Bu ince kontrol de tamamlandıktan sonra testler saklanır; cevap kâğıtları puanlamaya verilir.

PUANLAMA

Ölçme araçlarının çoğu sayısal bir puan verir. Ölçme işleminin sayısal bir sonuca ulaştırılmasına veya gözlenen niteliklerin nice- lendirilmesine **puanlama** denir. Herhangi bir ölçme aracından elde edilen ve henüz üzerinde bir dönüştürme yapılmamış olan sayılara **ham puan** denir. Puanlama işlemi hemen hemen bütün hallerde ham puanlar verir. Ham puanlar daha sonraki işlem ve kararlara taban olduğundan, puanlamanın hatasız yapılması önem taşır.

Ölçme aracının yapısı, onun üzerinde yapılabilecek puanlama işlemlerinin mahiyetini sınırlar. Söz gelişi, bir test seçmeli maddeler- den oluşuyorsa daha kolay, kısa cevaplı maddelerden oluşuyorsa daha zor puanlanır. Seçmeli testlerde de cevaplar ayrı cevap kâğı- dına kaydedilmişse puanlama kolaylaşır.

Ölçme aracının yapısı, puanlamanın hatasızca yapılıp yapılamas- nını da etkiler. Yazılı yoklamalarda ve kısa cevap gerektiren test- lerde olduğu gibi, bir cevabın doğru olup olmadığını tayin etmek pu- anlayıcının takdirine kalıyorsa, puanlama yöntemi sübjektiftir. Bu hallerde puanlama güvenilirliği tam olmaz; puanlara, takdir farklılık- larından doğan puanlama hataları karışır. Bir anahtarla puanlanabi- len seçmeli testlerde olduğu gibi, puanlayıcının görevi, seçilen ceva- bın anahtara uygun olup olmadığını tayin etmekten ibaretse, puan- lama yöntemi objektiftir. Özel cevap kâğıdı kullanılmak şartıyla bu tür testlerin cevapları makinayla puanlanabilir.

Puanlamanın sübjektif olduğu araçlarda, iyi hazırlanmış bir pu- anlama anahtarı, puanlamanın bir dereceye kadar objektifleştirilme- sine yardım eder. Objektif testlerde de puanlamayı kolaylaştıran araçlar kullanılır. Ayrı cevap kâğıdı ile uygulanan testlerde, puan- lama anahtarının bir kopyasında doğru cevapların delinmesiyle ya- pılabilecek bir puanlama aracı, anahtara uygun cevapların sayılma- sında kolaylık sağlar. Bu tür puanlama araçları, gerekli hallerde plastik gibi saydam maddelerden de yapılabilir.

Puanlamada, puanlama araçları, hatta puanlama makinaları kullanılsa bile, puanlama hataları tamamen önlenemeyebilir. Bu

sebeple, her puanlama işleminde bir de kontrol işlemi yapılmalıdır. Kısa cevaplı testlerde, puanlaması tamamlanmış kâğıtların küçük bir kısmının yeniden puanlanması kaba bir kontrol sağlar. Seçmeli test cevapları puanlanırken, «doğru», «yanlış», «birden fazla cevaplı» ve «cevapsız» maddeler sayıları birbirinden bağımsız olarak bulunursa, bu sayıların toplamı testteki toplam madde sayısını verdiği için, bir kontrol imkânı elde edilmiş olur. Toplamın tutmadığı hallerde, puanlama tekrarlanarak hatalar düzeltilir. Makina ile puanlanan çok sayıdaki cevap kâğıdından az sayıda kâğıt örnekleme yoluyla ayrılıp elle puanlanırsa, önemli bir kontrol yapılmış olur.

Testlerde çeşitli puanlama usulleri tavsiye edilmiştir. En basit puanlama kuralı, «doğru» cevaplandırılmış her maddeye «1»; «yanlış», «cevapsız», «birden fazla cevaplı» maddelere «0» vermektir. Daha karmaşık puanlama kuralları uygulanabilir, fakat puanlama işlemleri zorlaşır.

Seçmeli testlerin puanlanmasıyla ilgili olarak tartışılan bir konu, tahminden dolayı kazanılması muhtemel başarı için düzeltme yapılıp yapılmayacağıdır. Aşağıdaki kesimde bu konu ana çizgileriyle özetlenmektedir.

Seçmeli Testlerde Şans Başarısı

Seçmeli test maddelerinde cevabın verilen seçenekler arasından seçilmesi ve sınıflama gerektiren testlerde kategori sayısının sınırlı oluşu, doğru cevabın şansla da bulunabilmesi ihtimalini doğurur. Seçmeli bir test maddesinde doğru cevap şansla işaretlenmişse, o madde puanı, cevaplayıcının gerçek başarısından değil, şanslılığından kazanılmış olur. Böyle hallerde testteki başarı **şans başarısı** adını alır.¹

Seçmeli testlerde doğru cevabı bilen cevaplayıcılar doğru seçeneği, çeldiricilerden birini doğru sanan cevaplayıcılar da o çeldiriciyi işaretlerler. Doğru cevaptan emin olmayan veya madde ve çeldiricileri hakkında hiçbir bilgisi olmayan cevaplayıcıların çoğunluğu tahmine giderler. Bu gibiler şanslı oldukları oranda şans başarısı sağlarlar. Şans başarısı, bir yandan test maddesinin seçmeli olduğundan, bir yandan da cevaplayıcıların tahmine gitmesinden doğar.

Şans başarısı test puanlarına sabit bir hata olarak girmez. Çünkü şans puanı, bir yandan cevaplayıcıların tahmine gidip gitmediklerine, bir yandan tahmine gidenlerin ne derecede şanslı oldukları-

¹M.F. Turgut, Şans Başarısının Test Puanlarına Etkisi, s. 2.

na, bir yandan da bütün testte kaç maddeyi tahminle cevaplandırdıklarına bağlıdır. Ayrıca, şans başarısı, tesadüfî bir değişken gibi görünmekle birlikte, yukarıdaki etkilerin yaratacağı yanlışlıklardan dolayı bir tesadüfî değişken özeliği de göstermez. Şans başarısı, şanslı cevaplayıcıların puanlarını gerçek değerinden büyük gösteren, şanslıya daha çok pay veren bir etkidir. Onun için şans başarısı test puanlarının geçerliğini zedeler.

Şans başarısının test puanlarına katkısının çok büyük olamayacağı, aslında, konunun abartıldığı görüşünde olanlar vardır. Testlerdeki madde sayısı arttıkça da şans başarısının önemsiz dereceye düşeceği görüşünde olanlar da vardır. Her iki görüşte birer gerçek payı bulunmakla birlikte, şans başarısı önemsizlenecek kadar küçük bir hata kaynağı değildir.²

Geçmişte, şans başarısını önlemek için test açıklamalarına «Doğru cevabını kesinlikle bilmediğimiz maddeleri cevapsız bırakınız. Tahmin etmeniz aleyhinize olur.» gibi bir yönerge konulduğu yıllar olmuştur. Aslında bu yönerge işlemez. Çünkü, cevaplayıcılar, tahmin etmenin kendilerine puan kazandıracağını er geç öğrenirler. Üstelik, cevaplayıcıların bir kısmının bu yönergeye uyması bir kısmının ise uymaması halinde, şans başarısının yarattığı hata cevaplayıcılar arası farkların büyümesine sebep olur.

Şans başarısını cevaplama da önlemeye çalışmaktansa, puanlamada düzeltmeye gayret etmenin daha etkili bir yol olacağı düşüncesi doğmuş; böylece **düzeltilme formülü** adıyla bilinen bir formül bulunmuştur. Formül şöyledir :

$$(14.1) \quad R_b = R - \frac{W}{a - 1}$$

R_b : Cevaplayıcının düzeltilmiş puanı
 R : Bir cevap kâğıdındaki doğru işaretilenmiş maddeler toplamı
 W : Bir cevap kâğıdındaki yanlış işaretilenmiş maddeler toplamı
 a : Test maddelerindeki seçenek sayısı

Bu formülde görüldüğü gibi, düzeltilmiş veya **şans başarısından arıtılmış** puan, doğru cevaplar toplamından yanlış cevaplar toplamının bir kısmı çıkarılarak bulunmaktadır. Bu formül bazı sayıtlara dayanılarak kurulmuştur.³

²Ibid, s. 23 - 26.

³Ibid, s. 10 - 17.

Bu sayıtlardan biri, bir cevap kâğıdındaki yanlış cevapların hepsinin, tahmin edip tutturamamaktan doğduğu sayıtlıdır. Şüphesiz, herhangi bir cevaplayıcı için bu sayıtlı doğru değilse, formül hatalı bir düzeltme yapar. Bir cevaplayıcı, tahmin edip tutturamadığı için değil de, çeldiricilerden birini doğru sandığı için maddeyi yanlış işaretlemişse, bu bilgisizliğinden dolayı «cezalandırılmış» olur. Cevaplayıcılarda yanlış bilgiler bulunduğu hallerde formül gereğinden fazla düzeltir.

Formülün çıkarılmasında ikinci sayıtlı, bir cevaplayıcının tahmin ettiği her maddede doğru cevabı bulma ihtimalinin eşit ve $1/a$ değerinde olduğu sayıtlıdır. Böyle bir sayıtlı doğru ise, beş seçenekli maddelerden oluşan bir testte, 40 maddeyi tesadüfen cevaplandırılan bir cevaplayıcının şansla doğru cevaplandırabileceği madde sayısının $(1/5) 40 = 8$ olması beklenir. Şüphesiz şanslı bir kimse bundan daha fazla ve şanssız bir kimse de bundan daha az sayıda maddeyi doğru cevaplandırır. Bu sebeple formül, şanslı bir kimsenin puanını gereğinden az ve şanssız bir kimseninkini gereğinden fazla düzeltir. Ayrıca, bir kimse çeldiricilerden bir veya birkaçının yanlış olduğunu kestirerek eler ve geri kalan seçenekler arasında tahmine giderse, doğru cevabı bulma ihtimali $1/a$ değil, bundan daha büyük olur. Bu sebeple formül, elemeli tahminin işlediği hallerde şans başarısını gereğinden az düzeltir.

Bir cevaplayıcının cevap kâğıdında çekinilmiş ve erişilememiş maddelerin sayısı büyük ise, düzeltme formülü hiç uygulanmamalıdır. Özellikle, çekinilerek cevapsız bırakılmış madde sayısının büyük olması, o cevaplayıcının tahminden kaçındığına işaretler. Böyle bir kişinin ham puanlarına düzeltme formülü uygulanması, yanlış bilgilerden doğan yanlış cevap sayısının bir kısmı doğru cevap sayısından çıkarıldığı için, onun puan kaybına sebep olur.

Cevap kâğıtlarında boş bırakılmış madde sayısı önemsiz denecek kadar az ise, formül uygulanmalı mıdır? Bu sorunun cevabı test puanlarının kullanılacağı maksada göre verilir. Puanların mutlak değeri gerekli değilse, meselâ cevaplayıcıların aldıkları puan sırasına göre karar verilecekse, düzeltme formülünün uygulanması gerekmez. Çünkü, bütün cevaplayıcıların doğru cevabını bilmedikleri her maddeyi gelişigüzel tahmin ettikleri bir testte, düzeltilmemiş ham puan sıralamasıyla düzeltilmiş puan sıralaması aynı sonucu verir. Fakat, bir cevaplayıcının gerçekten kaç maddeyi doğru cevaplandırdığı kestirilecekse, ham puanların düzeltilmesi gerekir. Aynı gerekçeyle, mutlak değerlendirme yöntemleriyle not verilen hallerde de ham puanlar şans başarısından arıtılmalıdır.

B Ö L Ü M 15

Ölçme Sonuçları Üzerinde İstatistiksel İşlemler

Ölçme sonuçları çoğu zaman sayısal veriler halinde elde edilir. Ölçme sonuçlarının betimlenmesi, özetlenmesi, birbiriyle karşılaştırılması ve yorumlanması bu veriler üzerinde istatistiksel işlemler yapılmasını gerektirir. Bu bölümde, ölçme sonuçlarının özetlenip yorumlanmasında kullanılan başlıca istatistiksel yöntemler verilmektedir.

ÖLÇME SONUÇLARININ BETİMLENMESİ

Altmış maddelik bir test 41 kişilik bir sınıfa uygulanmış ve Tablo 15.1'deki ham puanlar elde edilmiş olsun. Puanları betimlemenin bir yolu, bu tabloda olduğu gibi, onları en büyükten en küçüğe doğru sıralamaktır.

TABLO 15.1
SIRALANMIŞ PUANLAR

52	49	48	46	45	44	43	42	41	41	40	40	39	38
38	37	37	36	36	35	35	35	34	33	33	33	32	32
31	31	30	30	29	27	27	26	26	25	24	22	20	

Bu tabloda 52 ile 20 arasındaki puanların bazıları mevcut değildir; bazılarının ise birden fazla vardır. Her puandan kaç tane olduğunu gösteren sayıya **frekans** denir. Tablo 15.1'de 23 puanının frekansı sıfır, 26 puanının ikiidir.

Puanlar sıralandıktan sonra, puan dizisinin üst ve alt sınırları kolayca belirlenir. Tablo 15.1'deki puanlar 52 ile 20 arasında dağılmıştır. Bu dağılımda diziyi niteleyen ölçülerden biri **dizi genişliği** adıyla bilinir ve en yüksek puan ile en düşük puan arasındaki fark ile tanımlanır. Tablo 15.1'de dizi genişliği $52 - 20 = 32$ puandır. Demek ki, bu dizi 32 puanlık bir aralığa dağılmıştır.

Gruplama ve Frekans Dağılımı

Yukarıdaki puanlar hem sıralanır, hem de puandan kaç tane olduğu sayılarak frekanslar bulunabilirdi. Böyle yapılıyorsa, 52'den 20'ye kadar yirmi yedi ayrı puan tablolanırdı. Şüphesiz, puanların daha geniş bir aralığa dağılması halinde tablodaki farklı puan sayısı daha büyük olurdu. Hem puan dağılımının genel gidişini daha iyi görmek, hem de hesaplamaları kolaylaştırmak amacıyla puanlar gruplanır. Tablo 15.2'de yukarıdaki puanların üç puanlık eşit aralıklarla gruplanması görülüyor.

TABLO 15.2
PUANLARIN GRUPLANMASI

Puan aralığı	Frekans (f)	Ağırlığın orta değeri (X')	Yığılmalı frekans (tf)	Ondalık basamaklarla aralık sınırları
50 - 52	1	51	41	49,5 - 52,5
47 - 49	2	48	40	46,5 - 49,5
44 - 46	3	45	38	43,5 - 46,5
41 - 43	4	42	35	40,5 - 43,5
38 - 40	5	39	31	37,5 - 40,5
35 - 37	7	36	26	34,5 - 37,5
32 - 34	6	33	19	31,5 - 34,5
29 - 31	5	30	13	28,5 - 31,5
26 - 28	4	27	8	25,5 - 28,5
23 - 25	2	24	4	22,5 - 25,5
20 - 22	2	21	2	19,5 - 22,5

Gruplamanın esası, birbirine yakın birkaç puanı bir tek puanmış gibi düşünmektir. Tablo 15.1'de 41 - 43 aralığındaki 43, 42, 41, 41 puanlarının hepsi 42 kabul edilmiştir.

Puanların gruplanmasında dikkat edilmesi gerekli bazı noktalar vardır. Puanları gruplayacak kimse, önce grup sayısını seçmek zorundadır. Grup sayısının küçük seçilmesi puan dağılımının gerçek özelliklerinin kaybolmasına sebep olur; büyük seçilmesi ise hesaplama işlemlerinin kolaylaştırılmasına engel olur. Ayrıca, grup sayısının tek olması, puan dağılımında simetri sağlar. Orta için, puanların 9, 11, 13, 15 gibi tek sayıda gruba ayrılması önerilir.

Grup sayısı saptandıktan sonra, dizi genişliği grup sayısına bölünerek **aralık genişliği** bulunur. Tablo 15.2'deki grupta grup sayısı 11 olarak seçilmiş, dizi genişliği $52 - 20 = 32$ bulunmuş ve aralık genişliği de $32/11 \approx 3$ alınmıştır. Gruplamada dikkat edilecek esaslardan biri de aralıkları eşit genişlikte almaktır. Aralık genişliklerinin eşit alınmaması halinde, gruplamadan beklenen hesaplama kolaylıkları azalır.

Aralık sınırlarının saptanmasında başlıca iki yol vardır. Puanlar tam sayılar halinde ifade edildiğinden, Tablo 15.2'deki ilk sütunda olduğu gibi, aralık sınırlarının tam sayılarla saptanması savunulur. Öte yandan, ortancanın ve diğer yüzdelliklerin bulunmasında tam puanın kesirlerine inileceği için, tablonun son sütununda olduğu gibi, aralık sınırlarının kesirli puanlarla ifade edilmesi de istenir. Bir frekans tablosunda aralık sınırları ve frekanslar gösterilir, fakat aralıkların orta değerleri gösterilmez. Zaten bu sayılar, en yüksek aralığın orta değerinden itibaren aralık genişliğindeki basamaklarla azalan sayılardır.

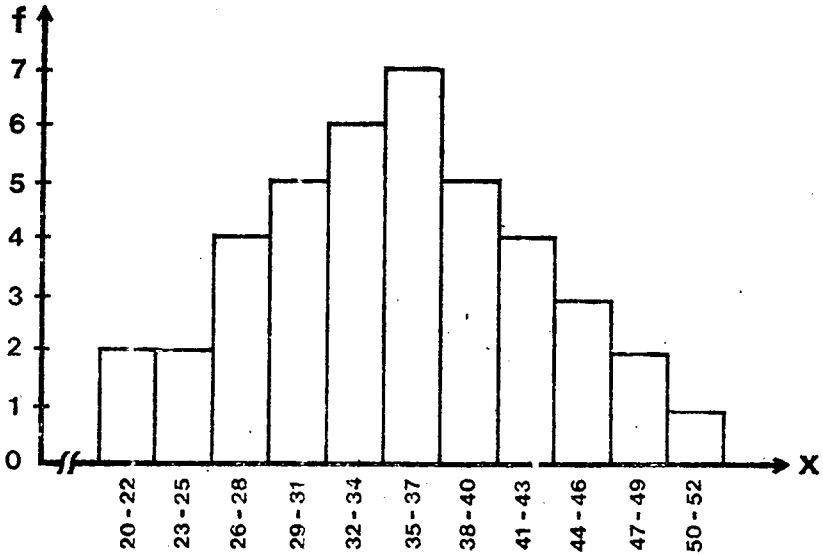
Bazen frekanslarla birlikte **yığmalı frekans** denilen sayılar da tablolanır. Yığmalı frekanslar, dizinin altından itibaren frekanslar toplanarak bulunur. Yukarıdaki örnekte 20 - 22 aralığının frekansı 2, onun üstündeki 23 - 25 aralığının yığmalı frekansı $2 + 2 = 4$, daha sonrakinin yığmalı frekansı $4 + 4 = 8$ ve 29 - 31 aralığının $8 + 5 = 13$ bulunur. Frekanslar toplamı dizideki toplam puan sayısını verdiği için, en yüksek aralığın yığmalı frekansı da bu sayıya eşit olmalıdır. Kullanılacağı yer öyle gerektirirse, yığmalı frekanslar en yüksek puan aralığından en düşük puan aralığına doğru da hesaplanabilir.

Frekans Grafikleri

Puan dağılımını betimlemenin diğer bir yolu, dağılımın grafiğini çizmektir. Frekans dağılımları birkaç yoldan grafiğe dökülebilir. Burada sadece üç grafik tipi örnek olarak verilmektedir.

İstatistik dilinde **histogram** adıyla bilinen grafik, her puanın veya puan aralığının frekansının bir sütun ile gösterildiği grafikdir. Şekil 15.1 yukarıdaki tabloda gruplanmış olan puanların histogramını gösteriyor. Histogramda yatay eksen hem puan eksenidir. Bu ekseninde puan aralıklarının sınırları veya her aralığın orta değeri gösterilir. Düşey eksen ise, frekans eksenidir. Bu grafikte her aralık, frekansı ile orantılı bir sütunla gösterilir. Histogramın şekli yatay ve düşey ekseninde alınan birimlere de bağlı olur. Birimler, grafiğin dü-

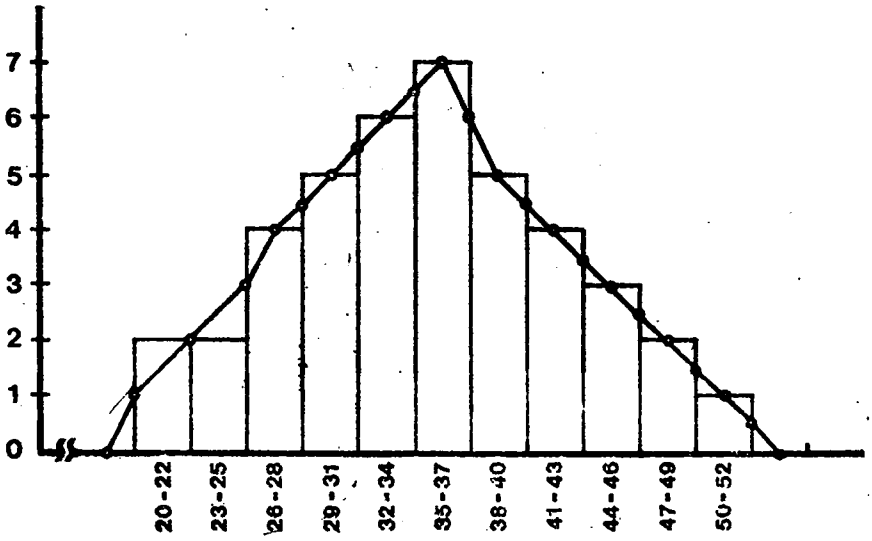
şey eksenindeki boyutunun, yatay eksenindeki boyutunun dörtte üçü olacak ölçekte seçilirse histogram güzel görünür.



Şekil 15.1 - Histogram

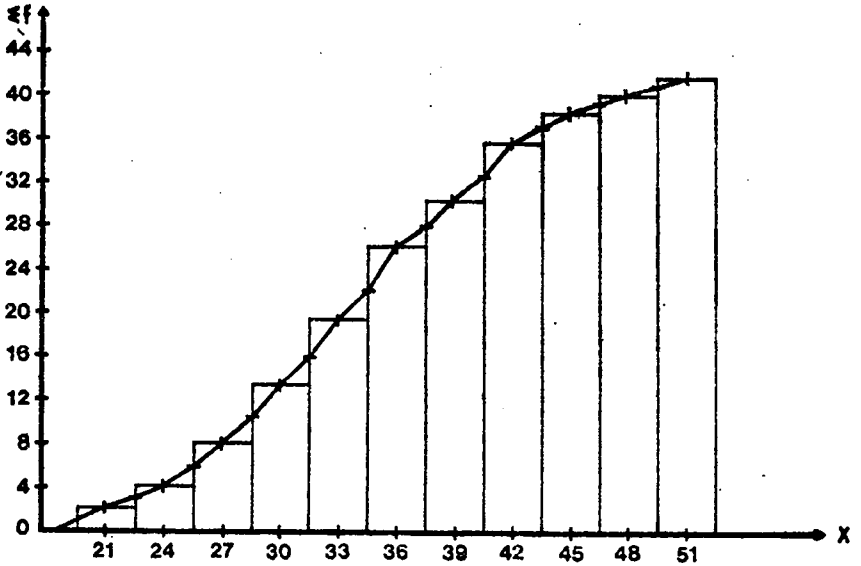
Ham puanların sürekli olduğu farzedilirse, frekans grafiğinin sütunlardan oluşan bir grafik değil, bir eğri olması beklenir. Histogram çizildikten sonra, onun üzerine çizgilerden oluşan bir «poligon» çizilebilir. Bunun için, ham puan ekseninin en düşük aralığının altındaki aralığın orta noktasından başlamak üzere, her ilerleyişin ve her yükselişin orta noktası işaretlenir. Sonra bu noktalar birleştirilir. Böylece elde edilen kırık çizgili grafiğe **frekans poligonu** denir. Frekans poligonu, sürekli frekans eğrisinin şeklini kabaca gösterir. Şekil 15.2'de frekans poligonunun histogram üzerine çizimi görülüyor.

Tablo 15.2'de gösterilen yığılmalı frekanslarla da bir sütun grafiği çizilebilir. Böyle bir sütun grafiği üzerine de, frekans poligonu çizilirken yapıldığı gibi, kırık çizgilerle yaklaşık şekli meydana çıkarılan bir eğri çizilebilir. Böyle bir eğri **yığılmalı frekans grafiği** adını alır. Şekil 15.3'de, Tablo 15.2'deki yığılmalı frekanslara dayanılarak çizilen bir grafik görülüyor. Yığılmalı frekanslar büyüyerek gittiği için, yığılmalı frekans grafiğinde frekans eksenindeki birimi küçültmek gerekir.



Şekil 15.2 Frekans poligonu

Bir dağılımın frekans poligonu kabaca bir çan eğrisine benziyorsa, aynı dağılımın yığılmalı frekans grafiği, yükseliş hızı başta ve sonda az, ortada fazla olan bir eğri verir.



Şekil 15.3 Yığılmalı frekans grafiği

MERKEZE YIĞILMA ÖLÇÜLERİ.

Puanların, dağılımın ortasında yığılma eğilimi gösterdiği, hem Tablo 15.2 gibi frekans dağılımı tablolarında hem de Şekil 15.1 gibi frekans grafiklerinde açıkça görülür. Frekans dağılımlarının bu özelliği epeyce genel bir özellik olduğundan, bir merkezde yığılma ölçüsü olarak bazı istatistikler tanımlanmış ve dağılımları betimlemede kullanılmıştır. Bu kesimde, bu istatistiklerden tepe değer (mod), ortanca (medyan) ve aritmetik ortalama kısaca anlatılmaktadır.

Tepe Değer (Mod)

Bir frekans dağılımının tepe değeri, frekansı en büyük olan puandır. Tablo 15.1'deki puanlar gruplanmadan bir frekans dağılımı yapılırsa, en çok tekrar eden puanların, her biri üçer tane olmak üzere 35 ve 33 oldukları görülür. Böyle bir dağılımın frekanslarının en büyük değerleri 33 ve 35 puanlarında 3 olur. Aynı puanlar Tablo 15.2'deki gibi gruplandığında tepe değer, frekansı 7 olan 35 - 37 aralığına düşer. Bu aralığın orta değeri olan $(34,5 + 37,5)/2 = 36$, bu dağılımın tepe değeri kabul edilir.

Frekans poligonu gibi eğrilerde tepe değer, o eğrinin maksimumuna veya tepe noktasına tekabül eden puandır. Frekans poligonlarında tepe değer bir maksimum olarak görüldüğü halde, yığılma frekans grafiklerinde bir maksimum olarak görünmez.

Bir dağılımın tepe değeri birden fazla olabilir. Ardışık puanların veya gruplanmış puanlarda ardışık aralıkların frekansları eşit ise bunların orta noktası tepe değer olarak alınır. Tablo 15.2'deki puanlarda 32 - 34 ve 35 - 37 aralıklarının frekansları 7 olsaydı, tepe değer 34,5 alınacaktı.

Tepe değer, merkeze yığılma ölçüsü olarak kaba bir istatistik olup ancak en büyük frekansın hangi puanda olduğunu gösterir. Böyle olmakla birlikte, tepe değer, ortanca ve aritmetik ortalama gibi istatistiklerle karşılaştırılarak, frekans dağılımlarının yorumlanmasında kullanılır.

Aritmetik Ortalama

Bir otobüs 600 kilometrelik Ankara - İzmir yolunu 8 saatte almış olsun. Bu otobüsün bir saatte aldığı yol $600/8 = 75$ km bulunur. Otobüs bütün yol boyunca bu hızla gitmemiştir; çoğu zaman daha hızlı gitmiş, bazen yavaşlamış, belki de birkaç defa durmuştur. Bu

sebeple, yolun tümünün yolculukta geçen zamanın tümüne bölünmesiyle bulunan hıza **ortalama hız** denir. Ortalama hız otobüsün herhangi bir andaki hızını göstermemekle birlikte, oldukça kullanışlı bir kavramdır. Meselâ, aynı otobüsün benzer koşullardaki 150 kilometrelik bir uzaklığı 2 saatte alabileceği, ortalama hıza bakılarak kestirilebilir.

Bir dersin üç sınavında 7, 6 ve 4 notlarını almış bir öğrencinin ortalama notu da, $(7 + 6 + 4)/3 = 5,7$ bulunur. Bu ortalama o öğrencini sınavlarının her birinde ne derecede başarı sağladığını göstermez; fakat, üç sınav birlikte düşünüldüğünde, öğrencinin ne derecede başarı sağladığını gösterir.

İstatistikte kullanılan çeşitli ortalamalar vardır. Yukarıdakiler gibi hesaplanan bir ortalamaya, **aritmetik ortalama** denir. Genel olarak $X_1, X_2, \dots, X_n$ gibi N sayının ortalaması, bu sayıların toplamı N ile bölünerek bulunur. Onun için, aritmetik ortalama şu formülle tanımlanır :

$$(15.1) \quad \bar{X} = \frac{\sum X_i}{N}$$

Bu formülde sigma (Σ) simgesi, birinciden sonuncuya kadar bütün puanların toplanacağını gösterir.

Puanlar gruplandıktan sonra ilk puanlar elde bulunmayabileceğinden, her aralığın orta değeri o aralığın frekansı ile çarpılarak, puanlar toplamına yakın bir sayı bulunur; bu sayı N ile bölünerek aritmetik ortalama elde edilir. Aritmetik ortalamanın bu yolla hesaplanması, genel olarak, şu formülle gösterilir :

$$(15.2) \quad \bar{X} = \frac{\sum f X'}{N}$$

f : Aralığın frekansı
X' : Aralığın orta değeri

Puanlar gruplandıktan sonra, bundan daha az hesaplama gerektiren kodlama ile hesaplama yöntemini uygulamak daha doğru olur. Onun için, 15.2 formülü, hesaplamalarda pek seyrek kullanılır.

Aritmetik ortalamanın kodlama ile hesaplanmasında, sırasıyla şu işlemler yapılır :

1. Puanlar eşit aralıklarla gruplanıp her aralığın frekansı bulunur.

2. Frekans dağılımının ortalarındaki bir aralığın kodu (Tablo 15.3'deki D değeri) sıfır kabul edilir; onun üstündeki aralıklara sırasıyla +1, +2,; altındakilere ise -1, -2, kodları verilir.

3. Her aralığın frekansı ile kodu çarpılarak, fD çarpımları bulunur. Bu çarpımlar sıfır kodlu aralıkta sıfır, onun üstündeki aralıklarda pozitif, altındakilerde negatiftir.

4. Pozitif ve negatif çarpımlar ayrı ayrı toplanarak, ΣfD toplamalarının cebirsel değeri bulunur.

5. Aritmetik ortalama şu formülle hesaplanır :

$$(15.3) \quad \bar{X} = X_0' + \frac{a}{N} \Sigma fD$$

X_0' : Sıfır kodlu aralığın orta değeri
 a : Aralık genişliği
 f : Aralığın frekansı
 D : Aralığın kodu

Tablo 15.3'deki gruplanmış puanların aritmetik ortalaması, kodlama yöntemiyle şöyle bulunur :

$$\bar{X} = 36 + \frac{3}{41} (-11) = 35,2$$

Kodlama, aslında, puan aralıklarının orta değerleri olan X' puanlarını D puanlarına çevirme işlemidir. Bu işlem sonucunda elde edilen D puanlarının aritmetik ortalaması daha kolayca hesaplanır. Formül 15.3'ün uygulanmasıyla da X puanlarının aritmetik ortalaması kolayca bulunur. Kodlamanın hangi aralıktan başlatıldığı aritmetik ortalamayı etkilemez. Fakat, dağılımın ortalarındaki bir aralığa sıfır kodu verilirse, üzerinde işlem yapılacak sayılar küçülür ve hesaplamalar kolaylaşır.

Aritmetik ortalama $\bar{X} = (X_1 + X_2 + \dots + X_n)/N$ veya $\bar{X} = (X_1/N) + (X_2/N) + \dots + (X_n/N)$ halinde de yazılabilir. Burada görüldüğü gibi, herhangi bir X_i puanının aritmetik ortalamaya katkısı X_i/n oranındadır. Aritmetik ortalamaya, bu özeliğinden dolayı, **basit aritmetik ortalama** da denir.

Ağırlıklı Ortalama

Puanların ortalamaya katkılarının X/N oranında değil de, başka oranlarda olması istenirse, aritmetik ortalama yukarıdakinden farklı bir yöntemle hesaplanır. Genel olarak, $X_1, X_2, \dots, X_n$ puanlarına

sırasıyla $W_1, W_2, \dots, W_n$ ağırlıkları verilerek hesaplanan aritmetik ortalamaya **ağırlıklı ortalama** denir. Ağırlıklı ortalama şu formülle hesaplanır :

$$(15.4) \quad \bar{X} = \frac{W_1X_1 + W_2X_2 + \dots + W_nX_n}{W_1 + W_2 + \dots + W_n}$$

Bir öğrencinin bir dersteki yazılı, sözlü ve ödev notları sırasıyla 3, 4 ve 8 olsun. Bu notların basit ortalaması $(3 + 4 + 8)/3 = 5$ olur. Bu notlara sırasıyla % 60, % 30 ve % 10 ağırlık verilmek istenirse, ağırlıklı ortalama $(60 \times 3 + 30 \times 4 + 10 \times 8)/(60 + 30 + 10) = 3,8$ olur. Ödev notuna küçük bir ağırlık verildiğinden, onun ortalamaya etkisi azaltılmış; böylece öğrencinin ortalama notu, daha çok ağırlık verilen yazılı ve sözlü notlarına yaklaştırılmıştır.

TABLO 15.3
ORTALAMA VE STANDART KAYMA HESABI İÇİN
PUANLARIN KODLANMASI

Puan ağırlığı	Frekans f	Kod D	fD	fD ²
50 - 52	1	+5	5	25
47 - 49	2	+4	8	32
44 - 46	3	+3	9	27
41 - 43	4	+2	8	16
38 - 40	5	+1	5	5
35 - 37	7	0	+35	0
32 - 34	6	-1	-6	6
29 - 31	5	-2	-10	20
26 - 28	4	-3	-12	36
23 - 25	2	-4	-8	32
20 - 22	2	-5	-10	50
N = 41			-46	$\Sigma fD^2 = 249$
a = 3				
$\Sigma fD = +35 - 46 = -11$			$X_o' = (35 + 37)/2 = 36$	

Ortanca (Medyan)

Diğer bir merkeze yığılma ölçüsü ortancadır. Ortanca, büyüklük sırasına dizilmiş puanlarda, dizinin tam ortasına düşen puandır. Bir öğrenci 5, 8, 3, 6, ve 4 notlarını almış olsun. Bu beş not 8, 6, 5, 4, 3 olarak sıralanır. Bu sıralamada üstten veya alttan üçüncü not dizisinin tam ortasına düştüğünden, öğrencinin notlarının ortancası 5'dir.

Dizideki puan sayısı, N çift ise ortanca iki puan arasına düşer. 9, 8, 6, 5, 4, 3 dizisinin ortancası 6 ile 5 arasına düşer. Böyle hallerde puanların ondalık değerlerine inmek gerekir. 5 sayısının 4,5 - 5,5 aralığında ve 6 sayısının 5,5 - 6,5 aralığında ondalık değerlere sahip olduğu düşünülürse, ortancanın 4,5 - 6,5 aralığının orta noktasında bulunduğu görülür. Bir aralığın orta değeri, aralık sınırları toplamının yarısına eşit olduğundan, yukarıdaki altı sayının ortancası $(4,5 + 6,5)/2 = 5,5$ bulunur. Ortancanın bu yolla hesaplanması, ancak dizideki sayıların frekanslarının eşit olduğu halde mümkündür. Puan veya aralık frekanslarının eşit olmadığı hallerde, ortanca, bir aralığın tam ortasına düşmez.

Puanlar gruplanmış ve yığılmalı frekanslar bulunmuş ise, ortancanın hesaplanmasında şu formülün uygulanması kolaylaşır :

$$(15.5) \quad \text{Ort} = L + \left(\frac{N/2 - tf_a}{f_b} \right) a.$$

L : Ortancanın bulunduğu aralığın alt sınırı
 N : Puan sayısı
 tf_a : L'ye kadar olan frekansların toplamı
 f_b : Ortancanın bulunduğu aralığın frekansı
 a : Ortancanın bulunduğu aralığın genişliği

Bu formül, Tablo 15.2'deki puanlara şu işlemlerle uygulanır :

1. Frekanslar toplamının yarısı bulunur. $N/2 = 41/2 = 20,5$.
2. Bu dizinin ortancası 34,5 - 37,5 aralığına düşer. Çünkü, bu aralığın alt sınırına kadar yığılmalı frekanslar toplamının $tf_a = 19$ ve bu aralığın frekansının $f_b = 7$ olduğu görülmektedir. $N/2 = 20,5$ sayısı 19'dan büyük $19 + 7 = 26$ 'dan küçüktür.

3. Bu aralığın alt sınırı $L = 34,5$ ve aralık genişliği $a = 3$ olduğundan, bu puanların ortancası şöyle hesaplanır.

$$\text{Ort} = 34,5 + \left(\frac{20,5 - 19}{7} \right) 3 = 34,5 + \left(\frac{1,5}{7} \right) 3 = 35,1.$$

Bu hesaplamada görüldüğü gibi, ortanca, 34,5 - 37,5 aralığının alt sınırına, bu aralık genişliğinin yedide bir buçuğu eklenerek bulunmuştur. Ortanca, frekanslar dizinin üst başından itibaren toplanarak da bulunabilir. Bu haldé, 34,5 - 37,5 aralığının üst sınırından, aralık genişliğinin yedide beş buçuğunu çıkarmak gerekir. Hesaplamaların doğruluğunu kontrol etmek için, ortanca, dizinin her iki ucundan da hesaplanıp sonuçların aynı olup olmadığına bakılabilir. Tablo 15.2'deki puanlarda, dizinin üst başından itibaren hesaplanan ortanca;

$$\text{Ort} = 37,5 - \left(\frac{26 - 20,5}{7} \right) 3 = 37,5 - \left(\frac{5,5}{7} \right) 3 = 35,1$$

bulunur.

Ortanca, dizideki puan sıralanışına dayandığı için, aşırı uçlardaki puanların sayısal büyüklüklerinden etkilenmez.

DAĞILMA ÖLÇÜLERİ

Puanlarda ortada toplanma eğilimi yanında, dağılma veya değişkenlik gösterme eğilimi de görülür. Bu sebeple, bir dizi puandaki dağılmayı gösteren bazı istatistikler tanımlanmıştır. Bu kesimde, dağılma ölçülerinden dizi genişliği (ranj), standart kayma ve çeyrek kayma üzerinde durulmaktadır.

Dizi Genişliği (Ranj)

Dizi genişliği, bir kümedeki en yüksek puan ile en düşük puan arasındaki fark olarak tanımlanır. Dizi genişliğinin hesaplanması için puanların sıralanmış olması gerekmez; en yüksek puan ile en düşük puanın bulunması yeter.

Puanlar gruplanmış ise, dizi genişliği, en yüksek aralığın orta değerinden en düşük aralığın orta değeri çıkarılarak bulunur. Tablo 15.2'deki puanlarda dizi genişliği, en yüksek aralığın orta değeri 51 ve en düşük aralığın orta değeri 21 olduğundan, $51 - 21 = 30$ bulunur. Bu değer gruplanmamış puanlarla bulunan değerden biraz farklı olabilir. Bu fark, gruplama işleminden doğar.

Standart Kayma

$X_1, X_2, \dots, X_n$ puanlarının herbirinden bu puanların aritmetik ortalaması çıkarılsa $X_1 - \bar{X} = x_1, X_2 - \bar{X} = x_2, \dots, X_n - \bar{X} = x_n$ farkları

elde edilir. Bu işlemlerle elde edilen fark puanlarına **ortalamadan sapmalar** denir.

Aritmetik ortalamadan sapmaların bir kısmı negatif, bir kısmı pozitif ve cebirsel toplamı sıfırdır. Fakat, aritmetik ortalamadan sapmalar, kareleri alındıktan sonra toplanırsa, negatif sayıların kareleri pozitif olduğundan, kareler toplamı sıfırdan büyük bir sayı çıkar. Bu kareler toplamından

$$(15.6) \quad S^2 = \frac{\sum x_i^2}{N - 1}$$

denklemleriyle hesaplanan istatistiğe **variyan**s denir.

Bir dizi puanın varyansı o dizideki değişkenliğin bir ölçüsüdür. Varyasyonun karekökü alınır**s**a **standart kayma** elde edilir.

$$(15.7) \quad S = \sqrt{\frac{\sum x_i^2}{N - 1}}$$

Standart kayma da bir değişkenlik ölçüsüdür. Varyans ve standart kayma birçok istatistiksel özelliğe sahip olduklarından çok çeşitli işlemlerde kullanılır. Standart kayma, gruplanmamış puanlardan doğrudan doğruya da hesaplanabilir. Hesaplama formülü hemen her istatistik kitabında vardır.¹

Gruplanmış puanlarda standart kayma, X' puan aralıklarının orta değerlerini göstermek üzere, şu formülle hesaplanır :

$$(15.8) \quad S = \sqrt{\frac{\sum f (X')^2 - (\sum f X')^2 / N}{N - 1}}$$

Yukarıdaki formül ile hesaplama oldukça zahmetlidir. Puanlar büyük sayılardan oluşuyorsa, puanların kareleri çok büyük sayılar vereceğinden, makinasız hesaplama çoğu halde yanlış ve zaman alıcı olur. Fakat, puanlar gruplandıktan sonra kodlanırsa, hesaplama işlemleri çok kolaylaşır. Gruplanmış ve kodlanmış puanların standart kayması, aşağıdaki formülle hesaplanır :

$$(15.9) \quad S = a \sqrt{\frac{\sum f D^2 - (\sum f D)^2 / N}{N - 1}}$$

¹H. Arıcı, İstatistik, s. 78 - 79.

Tablo 15.2'deki kodlanmış puanlara bu formül uygulanırsa, standart kayma

$$S = 3 \sqrt{\frac{249 - (-11)^2/41}{41 - 1}} = \sqrt{6,15} = 7,44$$

bulunur.

Variyans ve standart kaymanın formüllerinde N - 1 yerine N alarak kaynaklar vardır. Evren değerleri hesaplanıyorsa formül N ile uygulanır; ayrıca, N çok büyük ise, N - 1 yerine N alınması önemli bir hata yaratmaz. Fakat, küçük örneklerde böyle yapılması, standart kaymanın, bir örneklem istatistiği olarak «yanlı» çıkmasına sebep olur. Örneklerdeki puan sayısı büyük değilse, standart kayma, bu kesimdeki formüllerde gösterildiği gibi N - 1 ile hesaplanmalıdır.

Bir dağılımda standart kaymanın ne büyüklükte olması beklenir? Bu sorunun cevabı, şüphesiz, dağılımın istatistiksel özelliklerine göre değişir. Genellikle, aritmetik ortalamanın üç standart kayma altındaki sınır ile üç standart kayma üstündeki sınır arasına dağılımın hemen hemen hepsi girer. Bu kural Tablo 15.2'deki dağılıma uygulanırsa, ortalama 35,2 ve standart kayma 7,44 olduğundan, alt sınır $35,2 - 3 \times 7,44 = 12,9$ ve üst sınır $35,2 + 3 \times 7,44 = 57,5$ bulunur. Gerçekten, o dağılımda bu sınırların dışına taşan hiçbir puan yoktur.

Çeyrek Kayma

Yukarıdaki kesimlerin birinde ortanca hesabına değinilmişti. Ortanca, puan sırasının «yüzde ellinci» puanıdır; bu sebeple ellinci **yüzdelik** olarak da adlandırılabilir. Ortancanın tanımına ve gruplanmış puanlarda hesabına benzer olarak, herhangi bir Y_p yüzdeliği şu formülle hesaplanır:

$$(15.10) \quad Y_p = L + \frac{(pN/100) - tf_a}{f_b} \times a$$

Tablo 15.2'deki dağılımın alttan itibaren yüzde yirmi beşinci puanı 28,5 - 31,5 aralığına düşer ve bu formülle

$$Y_{25} = 28,5 + \frac{(25 \times 41/100) - 8}{5} \times 3 = 29,85$$

bulunur. Aynı formülle, 37,5 - 40,5 aralığına düşen yüzde yetmiş beşinci puan da şöyle bulunur.

$$Y_{75} = 37,5 + \frac{(75 \times 41/100) - 26}{5} \times 3 = 40,35$$

Yüzdelik puan hesaplama işlemi, aslında, dağılımda puan sıralamasının yüzde olarak ifade edilmiş noktalarına karşılık olan ham puanları bulmaktan ibarettir. Ortancanın bulunmasında puan sırasının altından veya üstünden itibaren hesaplama yapılabilir. Fakat, yüzdelikler daima puan sırasının alt ucundan itibaren hesaplanmalıdır; çünkü tanımları puan sırasının altından itibaren yapılmıştır.

Herhangi bir dağılımda Y_{25} , Y_{50} ve Y_{75} yüzdeliklerine sırasıyla birinci çeyrek, ikinci çeyrek veya ortanca ve üçüncü çeyrek denir; bunlar Q_1 , Q_2 ve Q_3 simgeleriyle de gösterilir. Üçüncü çeyrek ile birinci çeyrek arasındaki farkın yarısına **çeyrek kayma** denir. Bu istatistik şu formülle gösterilir :

$$(15.11) \quad Q = (Q_3 - Q_1)/2.$$

Çeyrek kayma da standart kayma gibi bir dağılım ölçüsüdür. Çeyrek kayma, her zaman standart kaymadan daha küçüktür. Nitekim, $Ort \pm Q$ sınırları arasında, bir dağılımın ortadaki yüzde ellisi girer.

DAĞILIMLARIN YORUMLANMASI

Bir grup hakkında bir değer yargısına varılacağı zaman, o grubun puanlarını bir bütün olarak yorumlamak gerekir. Puanları bir bütün olarak değerlendirmede başvurulan yollardan biri frekans dağılımının yorumudur. Frekans dağılımları değişkenlikleri bakımından birbirinden farklı olabileceği gibi, simetrik veya kayışlı oluşları bakımından da birbirinden farklı olabilirler. Bu kesimde, frekans dağılımlarında bağıl değişkenlik ve kayışlılık kavramları üzerinde kısaca durulacak; özel bir dağılım olan normal dağılıma da değinilecektir.

Bağıl Değişkenlik

Bundan önceki kesimde sözü edilen dizi genişliği, çeyrek kayma ve standart kayma, birimleri farklı iki ayrı ölçmeden elde edildiklerinde birbiriyle doğrudan doğruya karşılaştırılamaz. Herbiri 80 öğrenciden oluşan iki grubun bir testteki ortalamaları eşit fakat standart kaymaları farklı ($S_1 = 5,1$ ve $S_2 = 7,2$) bulunmuş olsun. İki ölçme aynı testten geldiği ve birimler her iki grupta eşit olduğu için, bu misaldeki iki grubun değişkenliği doğrudan doğruya karşılaştırılabilir. İkinci grubun puanları daha değişkendir.

Puanların, birimleri farklı ölçmelerden elde edildiği hallerde, **bağıl değişkenlik katsayısı** denilen ve

$$(15.12) \quad (V = S/\bar{X}) \cdot 100$$

formülüyle tanımlanan istatistik, iki grubun değişkenliğini karşılaştırma imkânı verir.

Bir grubun birimleri farklı iki ölçmedeki değişkenliği de bağıl değişkenlik katsayısı yardımıyla karşılaştırılabilir. Otuz beş kişilik bir sınıfın 32 maddelik bir matematik testindeki ortalaması $\bar{X}_m = 18,0$ ve standart kayması $S_m = 5,4$ bulunmuş; aynı sınıfın 55 maddelik bir tarih testindeki ortalaması $\bar{X}_t = 36,0$ ve standart kayması $S_t = 7,2$ bulunmuş olsun. Bu misalde standart kaymalar doğrudan doğruya karşılaştırılırsa, tarih puanlarının daha değişken olduğu kanısına varılabilir. Şüphesiz, bu kanı yanlış olabilir. Bu iki testin bağıl değişkenlik katsayıları şöyle hesaplanır :

$$V_m = (S_m/\bar{X}_m) \cdot 100 = (5,4/18,0) \cdot 100 = 30.$$

$$V_t = (S_t/\bar{X}_t) \cdot 100 = (7,2/36,0) \cdot 100 = 20.$$

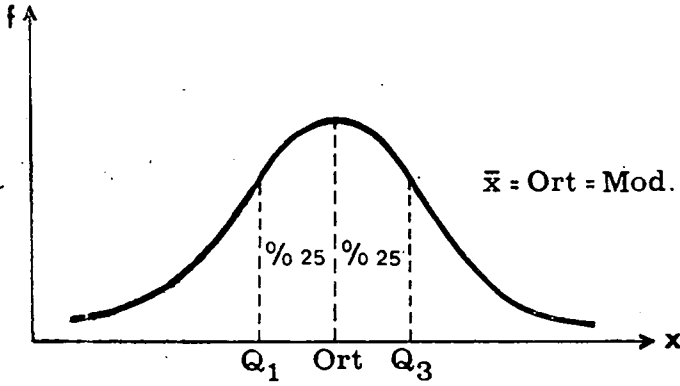
Matematik testi puanlarının bağıl değişkenlik katsayısının daha büyük olduğu görülüyor. Demek ki, o sınıfta matematik puanları tarih puanlarından daha fazla değişkenlik göstermiştir.

Formül 15.12'nin incelenmesinden de anlaşılacağı gibi, bağıl değişkenlik katsayısı yüzdeye benzer bir sayıdır; onun için, bir yüzde gibi yorumlanır.

Kayışlılık (Çarpıklık)

Bir frekans dağılımının şekline bakarak, onu oluşturan puanların dağılımı hakkında bazı yorumlara varılabilir. Şekil 15.4'de görüldüğü gibi, ortalamaya göre simetrik bir dağılımda aritmetik ortalama, ortanca ve mod birbirine eşittir. Ayrıca, simetrik bir dağılımda, sayısal değeri Q_1 ile ortanca arasına düşen puanlar tüm puanların yüzde yirmi beşini, ortanca ile Q_3 arasına düşenler de yüzde yirmi beşini oluşturur. Aynı şekilde, sayısal değeri Q_1 den küçük olan puanlar tüm frekansların dörtte birini, Q_3 den büyük olanlar da diğer dörtte birini oluşturur.

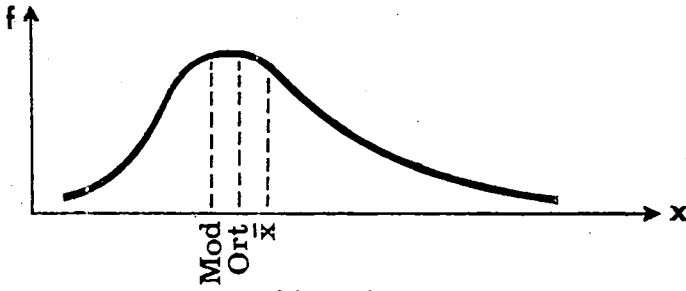
Aritmetik ortalamadan küçük puanların tüm frekansların yarısından fazla olması, puanların, ortalamanın altında yığıldığının belirtisidir. Bu halde frekans dağılımı Şekil 15.5'deki gibi olur. Bu du-



Sekil 15.4 Simetrik bir dağılım

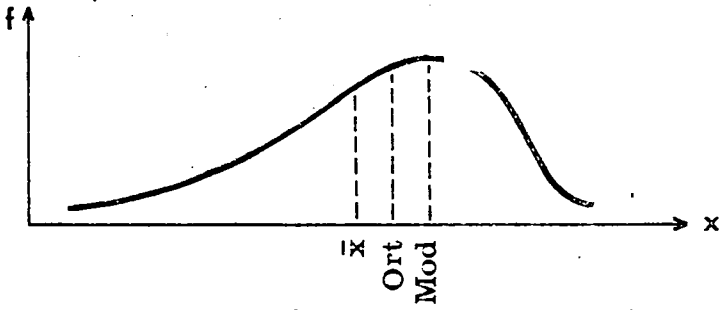
rumda merkeze yığılma ölçüleri arasında $\text{Mod} < \text{Ort} < \bar{X}$ ilişkileri gözlenir. Ortalamadan büyük puanların frekansları gittikçe düşer ve yüksek puanlar sağa doğru dağılır. Böyle bir dağılıma **pozitif kayışlı** veya **sağı çarpık** dağılım denir.

Puanların yarısından fazlasının aritmetik ortalamanın üstünde toplandığı bir dağılımda ise, Şekil 15.6'da görüldüğü gibi, $\bar{X} < \text{Ort} < \text{Mod}$ ilişkileri gözlenir. Böyle bir dağılıma **negatif kayışlı** veya **solu çarpık** dağılım denir.



Şekil 15.5 Pozitif kayışlı veya sağ çarpık bir dağılım

Çarpık dağılımlarda yukarıdaki paragraflarda sözü edilen mod, ortanca ve ortalama ilişkilerinin gözlenmesine sebep, aritmetik ortalamanın, puanların sayısal değerlerinden-etkilenerek aşırı puanlara doğru kaymasıdır. Bir tek modu bulunan düzgün bir dağılımda kayışlılığın yönü ve derecesi yaklaşık olarak;



Şekil 15.6 Negatif kayışlı veya solu çarpık bir dağılım

$$(15.13) \quad Ky = 3 (\bar{X} - \text{Ort}) / S$$

formülüyle hesaplanabilir. Bu istatistiğin pozitif çıkması halinde, aritmetik ortalama ortancanın sağında ve bu sebeple dağılım pozitif kayışlı; negatif çıkması halinde ise dağılım negatif kayışlıdır. Kayışlılığın sıfır çıkması halinde, şüphesiz, dağılım simetriktir. Bu istatistik her ikisi de aynı yönde çarpık iki dağılımın kayışlılık derecelerini karşılaştırmada da kullanılabilir.

Eğitimdeki ölçmelerde dağılımların şekliyle daha ziyade gruplar veya ölçme araçları hakkında yargılara varılacağı zaman ilgilenilir. Meselâ, bir sınıfın bir testteki puanlarının solu çarpık bir dağılım vermesi, bireylerin yarıdan fazlası ortalamanın üstünde puan kazandığından, o testin o sınıfa göre ortanın üstünde kolaylıkta olduğu anlamında yorumlanır.

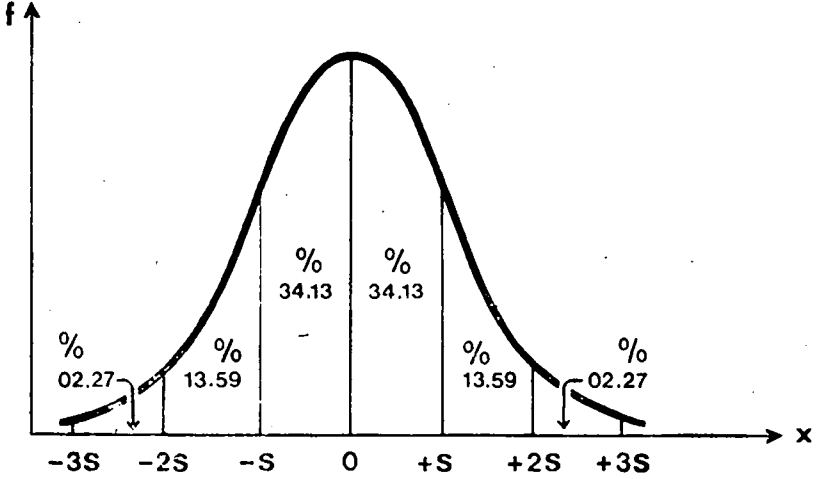
Normal Dağılım

Normal dağılım denen özel bir simetrik dağılım, bazı özelliklerinden dolayı ölçme ve değerlendirmede sık sık kullanılır. Matematiksel olarak ifade edildiğinde normal dağılım veya daha doğru olarak **birim normal eğri**,

$$(15.14) \quad y = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

denkleminin grafiğidir. Bu eğride ortalama $z = 0$, standart kayma $\sigma = 1$ ve eğri ile z eksenleri arasında kalan alanın tümü 1'dir. Ayrıca, eğrinin tepe noktası z 'nin sıfır değerinde olup bu noktada $y = 0,39894$ değerindedir. Şekil 15.7 ortalaması sıfır ve standart kayması σ olan

bir normal eğride, standart kayma birimlerine karşılık olan alanların tüm alana göre yüzdelerini gösteriyor. Bu şekilde f birimi σ biriminden büyük alınmıştır; böylece eğri y eksenini boyunca uzatılmıştır. Bu şekildeki bir eğri normal eğri olmakla birlikte, birim normal eğri değildir.



Şekil 15.7 Normal dağılım

Normal dağılım fonksiyonunun matematiksel özellikleri bilindiğinden, normal olduğu bilinen veya normal olduğu farzedilen bir dağılımın birçok özeliği, bu fonksiyon yardımıyla kestirilebilir. Bir sınıfta uygulanan bir testten alınan puanların ortalaması $\bar{X} = 42,4$ ve standart kayması $S = 6,3$ bulunmuş olsun. Ortalamanın bir standart kayma altı $42,4 - 6,3 = 36,1$ sınırını ve bir standart kayma üstü $42,4 + 6,3 = 48,7$ sınırını verir. O sınıfta puanların normal dağıldığı farzedilirse, puanları 36'dan büyük fakat 49'dan küçük olan öğrencilerin, tüm sınıfın yüzde altmış sekizini oluşturduğu sonucuna varılır. Çünkü bir normal dağılımda bu sınırlar arasına tüm dağılımın yüzde altmış sekizi (daha kesin olarak % 6826) düşer.

HAM PUANLARIN DİĞER PUANLARA ÇEVİRİLMESİ

Ölçme işleminden elde edilen sayılara ham puan denir. Ham puanlar, bazı nitelikleri açısından yetersiz olduklarından, kendi başına yorumlanamaz. Bir testten elde edilen ham puan, hiç değilse o testten elde edilmesi mümkün en yüksek puan bilinmeden yorumlanamaz. Ham puanları farklı iki öğrenci birbiriyle karşılaştırılabilirse

de aralarındaki farkın ne anlamda olduđu, elde başka bir dayanak olmaksızın yorumlanamaz.

Ham puanların kullanışsızlığı ve yorumlama güçlükleri, daha iyi niteliklerde ve daha kolay yorumlanabilir puanlar aranmasına; ham puanların bu tür puanlara çevrilmesine yol açmıştır. Bu kesimde, bu tür puanların hesaplanmalarına ve yorumlanmalarına kısaca yer verilmektedir.

Mutlak Başarı Oranı

En çok 40 puan verebilen bir testte, bir öğrenci 20 puan kazanmış olsun. Bu öğrenci için, hiç değilse «o testten elde edilebilecek maksimum puanın yarısını kazanmıştır» denilebilir. Genel olarak, en çok L değerini alabilen bir ölçmede, bir kişinin puanı X_i olsun. X_i/L oranına **mutlak başarı oranı** denir. Mutlak başarı oranları, $100 X_i/L$ formülüyle, yüzde olarak da ifade edilebilir. Ham puanların böyle yüzdelere dönüştürülmesi, en yüksek puan 100 olacağı için, en yüksek puana göre bir yorum sağlar. Bu yöntemle hesaplanmış bir puanın söz gelişi 84 olması halinde, o kişinin o testte elde edilebilecek en yüksek puanın yüzde seksen dördünü kazandığı anlamı çıkarılır.

Başarı oranlarında mümkün olan en yüksek puan ölçüt alındıđından, bu tür yorumlara bazen «mutlak» bazen de «ölçüt dayanaklı» yorumlar denir.

Standart Puanlar

Bir grup üzerinde yapılmış ölçmelerin sonuçları $X_1, X_2, \dots, X_n$ ham puanlarını vermiş olsun. Bu ham puanların kendi aritmetik ortalamalardan sapmalarını $X_1 - \bar{X} = x_1, X_2 - \bar{X} = x_2, \dots, X_n - \bar{X} = x_n$ ile gösterelim. Her ham puandan aritmetik ortalama çıkarılarak bulunan bu sapma puanları, ham puanlardan biraz daha fazla bilgi taşır. Çünkü, sapma puanlarının bulunmasında, aritmetik ortalama hesaplanıp kullanılmıştır. Bir öğrencinin sapma puanı -4 ise, onun grup ortalamasının 4 ham puan altında bir puana sahip olduđu, başka hiçbir bilgiye dayanılmaksızın söylenebilir.

Sapma puanları ham puanlardan daha fazla bilgi taşımakla birlikte, değerlendirme için yeterli bir yoruma imkân vermez. Sözü edilen öğrencinin puanının ortalamadan 4 puan küçük oluşu, onun, ortadan ne kadar uzak olduğunu göstermez. Bu eksikliği gidermek için ham puanlar üzerinde standart puana çevirme denilen bir dönüştürme daha yapılır.

Ham puanların standart kayması da hesaplanıp, her sapma puanı bu standart kayma ile bölünürse $z_1 = x_1/S$, $z_2 = x_2/S$, $z_n = x_n/S$ gibi puanlar bulunur. Böylece elde edilen z puanlarına **standart** puanlar denir. Çünkü, ham puanların ortalama ve standart kayması ne büyüklükte olursa olsun, z puanlarının ortalaması daima sıfır ve standart kayması daima 1'dir. Gerçekten, ham puanı aritmetik ortalamaya eşit olan bir kimsenin z puanı $z = (\bar{X} - \bar{X})/S = 0/S = 0$ bulunur. Ham puanı aritmetik ortalamanın tam bir standart kayma üstünde olan bir kişinin z puanı ise $z = (\bar{X} + S - \bar{X})/S = S/S = 1$ bulunur.

Aritmetik ortalamanın üç standart kayma altındaki alt sınır ile üç standart kayma üstündeki üst sınır aralığının, bir dağılımın hemen hemen hepsini kapsadığı daha önce belirtilmişti. Ham puanlar z puanlara dönüştürülmekle dağılımın şekli değişmez. Bu sebeple, $-3z$ ile $+3z$ sınırları aralığına da bir dağılımın hemen hemen hepsi girer. Bu özelliklerinden dolayı z puanları, sapma puanlarından daha fazla bilgi taşır ve başka bir bilgiye hacet kalmaksızın yorumlanabilir. Misal olarak, bir öğrencinin puanının $z = +2,00$ bulunduğu farzedilsin. Bu öğrencinin kendi grubunun ortalamasının iki standart kayma üstünde bir puana sahip olduğu, başka bir bilgiye gerek duyulmaksızın söylenebilir. O gruptaki puanların bir normal dağılım gösterdiği de farzedilirse, bir normal dağılımda $+2\sigma$ sınırına kadar tüm grubun yüzde doksan sekizi girdiğinden (Şekil 15.7 de $+2\sigma$ sınırına kadar alanlar toplanırsa $\% 50 + \% 34,13 + \% 13,59 = \% 98$ bulunur.), puanı $+2z$ olan bir kişi grubunun yüzde doksan sekizinden iyidir. Bu yorum, z puanlarının hesaplanmış olması ve dağılımın normal farzedilmesi sayesinde mümkün olmuştur.

Ham puanlar hangi sayılarda olursa olsun, z puanları her zaman hem kesirli, hem de bir kısmı eksi sayılar halinde elde edilir. Kesirli ve negatif sayıların kullanışsızlığı, z puanlarının da, doğrusal bir dönüştürmeyle, istenilen ortalama ve standart kaymada başka standart puanlara çevrilmesi zorunluğunu doğurur.

Çok kullanılan puanlardan biri, ortalaması 50 ve standart kayması 10 kabul edilen T puanlarıdır. Ham puanlardan z puanları hesaplanmışsa, onlardan

$$(15.15) \quad T_i = 10 z_i + 50$$

formülüyle T puanları bulunur. Şüphesiz, ortalamanın 50 ve standart kaymanın 10 olması da elvermeyebilir. Genellikle, ortalaması $\bar{T}$ ve

standart kayması S_x olan bir dönüştürme, ortalaması $\bar{X}$ ve standart kayması S_x olan ham puanlardan

$$(15.16) \quad T_i = \frac{S_x}{S_x} (X_i - \bar{X}) + \bar{T}$$

formülüyle hesaplanabilir. T puanları da z puanları gibi yorumlanır.

Yüzdelik Puanlar

Ham puanlar dağılımından herhangi bir yüzdeliğin nasıl hesaplanacağı, yukarıda formül 15.10 ile gösterilmiştir. Ham puanlar ranjının her puanı bu formülle yüzdeliğe çevrildiğinde, yüzdelik puanlar elde edilir. Formülün incelenmesinden de anlaşılacağı gibi, ham puanlar hangi büyüklükte olursa olsun, yüzdelik puanlar 0 ile 100 (daha kesin olarak 1 ile 99) arasında değişen sayılar verir.

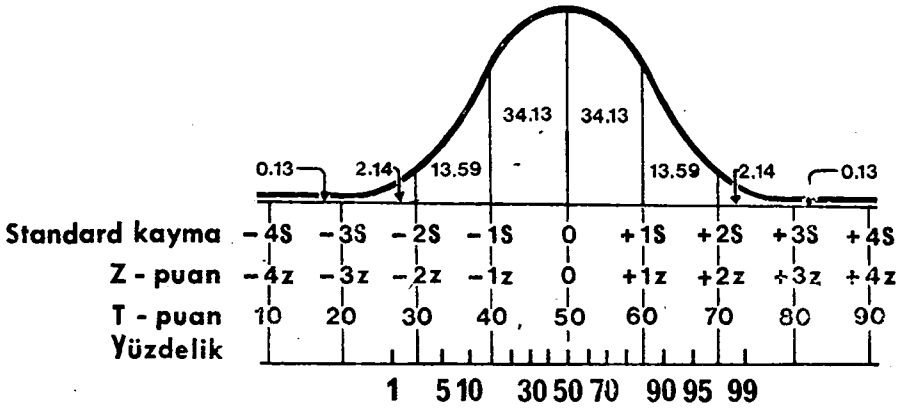
Yüzdelik puan, bir bireyin paunının, en düşük ham puandan itibaren kaçınıcı olduğunu, yüzde olarak ifade eder. Yüzdelik puanı 65 olan bir öğrencinin ham puanının, grubun yüzde altmış dördünün puanlarından yüksek ve yüzde otuz dördünün puanlarından düşük olduğu, başka bilgilere gerek duyulmaksızın söylenebilir. Demek ki, yüzdelik puan da, standart puan gibi kendi başına yorumlanabilen bir puandır.

Yüzdelik puanlar mutlak başarı oranlarından farklıdır. Mutlak başarı oranında söz konusu olan yüzde, herhangi bir ham puanın, mümkün olan en yüksek puana göre yüzdesidir. Oysa ki, yüzdelik puanlarda söz konusu olan «yüzdelik» herhangi bir ham puanın, ham puanlar sırasındaki yeridir. Bu ayırımın yapılabilmesi için, Türkçedeki «yüzde» kelimesi yerine, «yüzdelik» sözcüğü kullanılmaktadır.

Çeşitli Puanların Karşılaştırılması

Standart puanlarda, ister z puan ister herhangi bir T puan olsun, ham puan dağılımının şekli korunur. Ham puanlar çarpık bir frekans dağılımına sahipse, o puanlardan yukardaki yöntemlerin herhangi biriyle elde edilen standart puanlar da aynı yönde çarpık puanlar verir. Onun için, test standartlaştırma gibi büyük gruplar üzerinde yapılan ölçekleme işlemlerinde, ham puanlar hem standartlaştırılır hem de dağılımları normalleştirilir. Bu kitapta normalleştirmenin istatistiksel yöntemleri üzerinde durulmayacaktır. Fakat, bir normal dağılımda çeşitli puanlar, birimleri yönünden karşılaştırılmaktadır.

Şekil 15.8'de standart kayma birimiyle dilimlenmiş bir normal dağılımda, alanlar ve çeşitli puanlar gösterilmiştir.



Şekil 15.8 Bir normal dağılımda çeşitli puanlar

Bu şekilde de görüldüğü gibi, standart puanların eşit artış ve azalışları X eksenini üzerinde eşit uzaklıklara denk gelmektedir. Standart puanlar eşit aralıklı bir ölçek verir. Buna karşılık, yüzdelik puanlarda birimler eşit değildir. Dağılımın ortasında ölçek aralıkları küçük, uçlarda büyüktür. Bu sebeple, yüzdelik puanların eşit aralıklı olmadığı söylenir. Başka bir deyişle, meselâ yüzde beşlik bir fark, ortalama civarında küçük bir ham puan farkına, iki uca ise daha büyük bir ham puan farkına denktir.

İLİŞKİ ÖLÇÜLERİ

Bir sınıfta matematik ve Türkçe sınavları yapılmış olsun. Her öğrenci için, biri matematik puanı diğeri Türkçe puanı olmak üzere, iki puan bulunsun. Türkçe dersindeki başarı ile matematik dersindeki başarı arasında bir ilişki var mıdır? Varsa bu ilişkinin yönü ve derecesi nedir? Bu gibi soruların cevaplandırılabilmesi için puanlar üzerinde, ilişki ölçüleri denilen istatistiklerden birinin hesaplanması gerekir. Bu kesimde ilişki ölçülerinden momentler çarpımı korelasyonu ve sıra korelasyonu kısaca anlatılmaktadır.

Momentler Çarpımı Korelasyon Katsayısı

Tablo 15.4'de on kişilik bir grubun matematik ve Türkçe har puanları ile onlardan hesaplanmış sapma puanları görülmüyor. Bı

tabloda $x \cdot y$ sütunu sapma puanlarının çarpımını gösteriyor. Çarpimlardan herhangi biri sıfır olduğu zaman çarpım sıfır, çarpımların biri pozitif diğeri negatif olduğu zaman da çarpım negatif olacaktır. Bu sebeple, $x \cdot y$ çarpımlarının cebirsel toplamı negatif de olabilir.

Çarpımlar toplamının ortalama değerine **kovaryans** denir. Bu kovaryans daima iki değişken arasında söz konusu olup

$$(15.17) \quad C_{xy} = \frac{\sum x \cdot y}{N}$$

formülüyle hesaplanabilir. Kovaryans iki değişkenin birlikte nasıl değiştiklerini gösterir. Kovaryansın pozitif olması iki değişkenin aynı yönde değiştiğini, negatif olması değişkenlerden biri büyürken diğercinin küçüldüğünü, sıfır olması iki değişken arasında bir ilişki bulunmadığını gösterir.

TABLO 15.4
KORELASYON KATSAYISININ SAPMA PUANLARINDAN HESABI
İÇİN TABLOMALA

Öğr.	Matematik			Türkçe			xy
	X	x	x ²	Y	y	y ²	
A.D.	23	8	64	24	7	49	56
B.C.	19	4	16	17	0	0	0
C.A.	18	3	9	22	5	25	15
D.L.	18	3	9	16	-1	1	-3
E.K.	16	1	1	15	-2	4	-2
F.B.	15	0	0	18	1	1	0
K.S.	13	-2	4	16	-1	1	2
S.M.	11	-4	16	14	-3	9	12
L.N.	10	-5	25	12	-5	25	25
R.E.	7	-8	64	16	-1	1	8
Σ	150		208	170		116	118
							-5
							113

Kovariyansın sayısal değeri iki değişkenin sayısal değerlerine bağlı olduğundan, bir kovariyansın ne derecede bir ilişki gösterdiğini kendi başına yorumlamak mümkün değildir. Fakat, ham puanların z puanlara dönüştürülerek standartlaştırılmasıyla kendi başına yorum sağlandığı gibi, kovariyans ham puanların standart kaymalarıyla bölünürse, kendi başına yorum sağlanabilir. İki değişken arasındaki kovariyansın, o değişkenlerin standart kaymalarıyla bölünmesinden elde edilen istatistiğe **momentler çarpımı korelasyon katsayısı** denir. x ve y değişkenleri arasındaki korelasyon

$$(15.18) \quad r_{xy} = \frac{C_{xy}}{S_x S_y}$$

oranıyla tanımlanır. Bu istatistik bazen «doğrusal» korelasyon, bazen de sadece korelasyon adıyla anılır. Bu oran, kovariyans sıfır olduğu zaman sıfır değerini alır. Kovariyans negatif ise korelasyon negatif, pozitif ise pozitif değer alır. Korelasyon katsayısı en büyük değer olarak +1,00; en küçük çebirsel değer olarak da —1,00 değerini alır. $r_{xy} = +1,00$ halinde iki değişken arasında doğrusal ve mükemmel bir ilişki vardır; iki değişken aynı yönde birbiriyle oranlı olarak değişir. $r_{xy} = -1,00$ halinde ise iki değişken zıt yönde birbiriyle oranlı değişmektedir. Genellikle, korelasyon katsayıları mükemmel olmaz. Demek ki, korelasyon katsayısı iki değişken arasındaki ilişkinin hem yönünü hem de derecesini gösterir.

Tobla 15.4'deki verilerden korelasyon katsayısı şöyle hesaplanır:

$$S_x = \sqrt{\Sigma x^2 / \sqrt{N} - 1} = \sqrt{208 / \sqrt{10} - 1} = 4,81$$

$$S_y = \sqrt{\Sigma y^2 / \sqrt{N} - 1} = \sqrt{116 / \sqrt{10} - 1} = 3,59$$

$$r_{xy} = \frac{\Sigma x \cdot y}{N S_x S_y} = \frac{113}{10 \times 4,81 \times 3,59} = 0,65.$$

Korelasyon, matematiksel olarak formül 15.18'in eşdeğeri olan başka formüllerle de hesaplanabilir. Ham puanlardan z puanları hesaplanmışsa

$$(15.19) \quad r_{xy} = \frac{\Sigma z_x z_y}{N}$$

formülü kolaylıkla uygulanır. Ham puanlar üzerinde hiçbir işlem yapılmamışsa

$$(15.20) \quad r_{xy} = \frac{N\sum XY - \sum X \sum Y}{\sqrt{[N\sum X^2 - (\sum X)^2] [N\sum Y^2 - (\sum Y)^2]}}$$

formülü uygulanır. N büyük ise ve hesap makinası varsa, bu formül kullanışlıdır. Makinasız hesaplamalar zahmetli ve genellikle hatalı olur. Makinasız uygulanabilecek bir yöntem, verileri kodlamaktır. Böyle bir yöntem Arıcı'da ayrıntılarıyla verilmiştir.²

Sıra Korelasyonu

İki değişken arasındaki ilişki derecesinin hesaplanmasında, ham puanların sayısal değerleri yerine, puan sırasını kullanan istatistikler de vardır. Bunlardan biri olan **Spearman sıra farkı korelasyonu** şu formülle tanımlanır :

$$(15.21) \quad r_s = 1 - \frac{6 \sum d^2}{N (N^2 - 1)}$$

Bu istatistiğe kısaca sıra korelasyonu da denir. Bu korelasyonun hesaplanması için, önce, aşağıdaki basamaklarla, Tablo 15.5 gibi bir hesaplama tablosu hazırlanır.

1. Değişkenlerden biri, bireysel kimlikler kaybedilmeden, en yüksek puandan en düşük puana doğru, puan sırasına konur. Tablo 15.4'deki X puanları Tablo 15.5 de sıralanmış, sıra numaraları R (X) sütununda gösterilmiştir. Aynı puandan birkaç tane varsa bunlara önce farklıymış gibi sıra numarası verilir. Sonra sıra numaralarının ortalaması alınır ve herbirine R (X) olarak verilir. Misalde iki 18 vardır. Bunlar farklı olsalardı 3 ve 4 sıra sayılarını alacaklardı. Yukarıdaki kurala göre her ikisine de $(3 + 4)/2 = 3,5$ sıra sayısı verilmiştir.

2. İkinci değişkenin ham puanları, bireysel kimlikler kaybedilmeden başka bir yerde sıralanır. Ham puanlar ve R (Y) sıra sayıları, bireysel kimliklere uyularak tabloya geçirilir.

3. Tablonun her satırında, büyük sıra sayısından küçük sıra sayısı çıkarılarak, sıra sayıları farkının mutlak değeri bulunur.

4. Sıra sayıları farklarının kareleri alınır ve bu kareler sütununun toplamı bulunur.

²Ibid, s. 113 - 118.

TABLO 15.5
SIRA KORELASYONUNUN HESABI İÇİN TABLOLAMA

Öğr.	X	R (X)	Y	R (Y)	d= R(X)-R(Y)	d ²
A.D.	23	1	24	1	0	0
B.C.	19	2	17	4	2	4
C.A.	18	3,5	22	2	1,5	2,25
D.L.	18	3,5	16	6	2,5	6,25
E.K.	16	5	15	8	3	9
F.B.	15	6	18	3	3	9
K.S.	13	7	16	6	1	1
S.M.	11	8	14	9	1	1
L.N.	10	9	12	10	1	1
R.F.	7	10	16	6	4	16
N = 10					$\Sigma = 49,50$	

Hesaplama tablosu böylece tamamlandıktan sonra, Formül 15.21 uygulanarak sıra farkı korelasyonu bulunur. Yukarıdaki tablodan sıra farkı korelasyonu şöyle hesaplanır :

$$r = 1 - \frac{6 \times 49,5}{10 (100 - 1)} = 0,70.$$

Sıra farkı korelasyonu da, momentler çarpımı korelasyon katsayısı gibi yorumlanır.

Değerlendirme ve Not Verme

Öğretmenliğin en güç işlerinden biri not vermedir. Öğretmenin geçerli ölçütlerle ve objektif yöntemlere not verebilecek bilgilere sahip olması, onun güçlüklerini azaltır. Eğitim sistemimizde en çok değişikliğe uğrayan, buna rağmen sağlam ve temelli bir çözüme oturtulamayan kurallar da, sınavlara, not takdirine ve sınıf geçmeye ait kurallardır. Bu da, değerlendirme ve not vermenin oldukça güç bir konu olduğuna kanıttır.

Bu bölümde önce değerlendirme ve notların eğitim kararlarındaki yeri, sonra not vermede uygulanan başlıca yöntemler tartışılmaktadır.

EĞİTİMDE DEĞERLENDİRME VE NOTLAR

Notlar, öğrenci başarısının değerlendirilmesi sonucunda ortaya çıkar ve onun için öğrenci hakkındaki değer yargılarının ölçüleridir. Bir kesimde notların eğitimdeki gerekliliği, değerlendirmenin mahiyeti ve not vermede kullanılan ölçütler üzerinde durulmaktadır.

Eğitimde Notların Gerekliliği

Eğitimde notun gerekliliği, değerlendirmenin gerekliliğinden doğar. Cronbach, değerlendirmenin, bu sebeple de notların, eğitimdeki fonksiyonlarını şöyle ifade eder.¹

1. Değerlendirme, öğrenciye davranışını nasıl değiştireceği, nasıl geliştireceği hakkında bilgi verir.
2. Değerlendirme, yeterince başarılı olan öğrenciyi güdüler.
3. Değerlendirme, öğrenci hakkında verilecek kararlara dayanak olur.

¹L.J. Cronbach, *Educational Psychology*, 2nd ed., s. 539.

4. Değerlendirme, öğretmenin kendi öğretiminin ne derecede etkili olduğunu kestirmesine yardım eder.

5. Değerlendirme, yöneticilere ve diğer ilgililere bilgi verir.

Cronbach, yukarıda sıralanan maddelerde değerlendirmeye iki açıdan bakmıştır. Bunlardan birincisi değerlendirme sonuçlarına bakarak çeşitli kararlar verme, ikincisi öğrenciye ve diğer ilgililere öğrenci başarısı hakkında bilgi vermedir. Thorndike ve Hagen öğrenci başarısı hakkındaki bilgilerin nerelerde kullanılabilceğini tartışmışlardır.² Onlara göre öğrenci başarısının bilinmesi; (1) öğrenciyi okul çalışmalarında güdüler, (2) öğrenim etkinliklerine rehberlik edilmesine yardımcı olur, (3) ilerideki öğrenimin planlanmasına ve meslek seçimine yardımcı olur, (4) kişisel gelişmenin değerlendirilmesine yarar:

Okullarımızda notların en önemli gerekçesi, sınıf geçme kararlarına dayanak sağlamasıdır. Bu açıdan notlara, öğrenim basamaklarında öğrenci akışını düzenleyen ölçüler olarak bakılabilir. Hatta, okullarımızda notların yalnız sınıf geçme amaçları için kullanıldığını söylemek bile bir abartma olmaz. Geleneksel yöntemlerle verilen notların, Thorndike ve Hagen tarafından sayılan dört işgörüğü yerine getirebilecekleri şüphelidir. Sınıf geçme kararlarına dayanak olarak notlar, çoğu halde yetersiz kalmaktadır. Notlardan beklenen işgörülerin yerine getirilebilmesi için, notların hangi niteliklerde olması gerektiği sorulabilir. Bu sorunun cevapları aşağıdaki paragraflarda kısaca tartışılmaktadır.

Güdüleyici Olarak Not : Bir öğrencinin okuldaki bir etkinlik için güdülenmesi çok yönlü psikolojik bir olaydır. Okulda öğrenciyi güdüleyen etkenler çok çeşitlidir. Bir güdüleyici öğrenciyi pozitif yönde güdüleyebileceği gibi, negatif yönde de güdüleyebilir. Her ne suretle verilirse verilsin, notların da güdüleyici etkileri vardır. Notun pozitif bir güdüleyici olarak etkisi, başarılı olmanın insanda başarıyı teşvik edişinden; notu takdir eden öğretmenin takdirini kazanmak isteğinden; notun bir çeşit sicil oluşturacağı ve ileride kullanılacağı bilincine varmaktan; notun okul arkadaşları çevresinde ve ailede yarattığı tavsip ve takdirden doğar. Notun olumsuz bir güdüleyici olarak etkileri ise, başarısızlığın yozdurucu ve utandırıcılığından; öğretmen - öğrenci ilişkilerinde tasvip ve takdiri kaybetme korkusundan; ailede bu korkunun yanında diğer baskıların varıldığından doğar.

²R.L. Thorndike and E. Hagen, *Measurement and Evaluation in Psychology and Education*, 2nd. ed., s. 484 - 491.

Bir güdüleyicinin, günlük öğretim etkinliklerinde etkili olabilmesi için, öğrenmenin hemen peşinden gelmesi gerekir. Bu açıdan bakılınca, seyrek olarak verilen notların pozitif güdüleyici olarak etkilerinin önemli olmayacağı anlaşılr. Böyle olmakla birlikte, ödev ve ara sınavı gibi durumlarda ölçme sonuçlarının öğrencilere kısa sürede bildirilmesi, notun güdüleyicilik etkisini artırabilir.

Notun, olumsuz güdüleyici olarak etkilerinin önlemesi daha zordur. Sadece not vermede alınacak önlemler, olumsuz etkilerin önlenmesi için, notun hakedilmeyen bir dereceye yükseltilmesi veya bireysel ölçütlerle not takdiri gibi yöntemlere götürür ki, bunlar da notlardan beklenen diğer görevlere zıt düşer. Öğrencide olumsuz güdüleri kırmmanın çeşitli yolları vardır. Notun geçerliğini düşürmektense, öğrencilere notları, objektif değeriyle kabul etmelerini öğretmek, üzerinde durulmaya değer bir seçenektir. Notların olumsuz etkilerini abartan, bu yüzden onların ve sınıf geçme yöntemlerinin tümüyle terkedilmesini öneren eğitimciler olmuştur.³ Bu kitabın yazarı, notun niteliklerini korumanın, olumsuz etkilerini de başka yöntemlerle önlemenin daha doğru olacağı konusındadır.

Öğrenme Etkinliklerinin Rehberliği İçin Not : Öğrenme, çoğu halde tam olmaz; hatta yanlış öğrenmeler bile vuku bulabilir. Öğrencinin takıldığı bir nokta daha ileri öğrenmeleri engelleyebilir. Eksiklikler ve yanlışlar zamanında teşhis edilip düzeltilmezse, öğrenmede çeşitli güçlükler doğabilir. Tamamlayıcı ve düzeltici öğretim için rehberlik gereklidir. Engellerin, eksikliğin ve yanlışların ortaya çıkarılmasından hemen sonra durumun öğrenciye bildirilmesi, düzeltici etkinlikleri kolaylaştırır. Şüphesiz, dönemden döneme verilen bir karne notu bu görevi yerine getiremez. Çünkü notlar, eksikliğin veya yanlışlığın nerelerde olduğunu göstermez; not, eksiklikleri gösterici nitelikte olsa bile, öğrenciye zamanında ulaşmaz. Karne döneminden karne dönemine verilen bir başarı notunun tamamlayıcı ve düzeltici öğretim etkinliklerinin planlanmasına yardım edecek teşhis gücü yoktur. Notlar tamamlayıcı ve düzeltici öğretim planları için kullanılacaksa, ölçmeler ayrıntılı yapılmalı; bir derste birkaç «alt not» verilmeli; sonuçlar sık sık öğrenciye bildirilmelidir.

³Bakınız : M.S. Marshall, «The Flotation Technique : Teaching Without Grades», *Improving College and University Teaching*, 8, 23 - 29.

J.I. Goodlad and R.H. Anderson, *The Nongraded Elementary School*, rev, ed R.H. Anderson, «The Importance and Purposes of Reporting», *National Elementary School Principal*, 45, 6 - 11.

Öğrenim ve Meslek Rehberliği İçin Not : Öğrencilerin bir okulda kazandıkları notlarla daha yukarı basamaktaki bir okulda kazandıkları notları arasında, genellikle pozitif korelasyonlar görülür. O halde, aşağı okuldaki başarı, yukarı okuldaki başarıyı bir dereceye kadar yordar. Öte yandan, özellikle orta öğretimin ikinci devresinde, öğrencilerin çeşitli derslerdeki başarılarında sezilebilir bir farklılaşma ortaya çıkar. Kimi öğrenci fen derslerinde, kimisi sosyal derslerde, kimisi el becerilerinde, kimisi de güzel sanatlarda daha başarılı olur. Bütün derslerde ancak bir dereceye kadar başarılı olabilen, hiçbir alanda sıvırlıemeyen öğrenciler de bulunmasına rağmen, başarılarıdaki bu farklılaşma notların ayırıcı gücü bulunabileceğine işarettir.

Sadece notlara bakarak öğrenim dalı ayırmak ve meslek rehberliği yapmak, çoğu zaman isabetsiz kararlar verilmesine yol açar. Buna sebep, notların yeterince güvenilir bir ölçü olmayışı yanında, yordayıcı veya ayırıcı amaçlarla yapılmış ölçmelere dayanmayışıdır. Daha ziyade yordayıcı ve ayırt edici ölçülere dayanan notlar, yeterince güvenilir olmak kaydıyla, öğrenim dalı ve meslek seçiminde diğer yordayıcılarla birlikte kullanılabilir.

Kişisel Gelişmenin Ölçüsü Olarak Not : Okul, nihai analizde, her öğrenciyi kişisel potansiyelinin üst sınırına kadar geliştirmekle, onu yarının toplumuna hazırlamakla görevli olan bir kurumdur. Böyle olunca, her öğrencinin bir bütün olarak ne derecede geliştiği, toplumun istediği insan modeline ne dereceye kadar yaklaştığı ölçülebilmeli; her öğrenci bu kişisel gelişme ölçüsüne bakılarak değerlendirilmelidir. Ne yazık ki, gelenekselleşmiş başarı notları veya onların bir bileşkesi, kişisel gelişmenin ve kişiliğin toplu oluşumunun bir ölçüsü olamaz. Çünkü, dönemden döneme verilen karne notları, hatta üç yıllık okul süresinde kazanılan notların tümü, ancak bazı alanlardaki başarıyı yansıtır; kişilik gelişmesinde önemli olan bazı değişkenleri ölçme dışı bırakır. Okullar bu değişkenleri önemsemediklerinden değil, ölçemediklerinden karneye dahil etmezler. Bununla birlikte, karnelerin, akademik başarının yanında kişilik değişkenlerini de kapsaması; öğrencinin toplu gelişiminin değerlendirilmesi, modern eğitimcilerin sık sık savundukları bir görüştür.

Sınıflama Kararlarının Dayanağı Olarak Not : Öğrencilerin bir okul içinde, derslerdeki başarıları esas tutularak, alt sınıflardan üst sınıflara doğru ilerletilmeleri; bir okulu bitirdikten sonra daha yukarı düzeydeki bir okula kabul edilmeleri, hemen bütün eğitim sistemlerinin uyguladığı bir yoldur. Öğrencilerin eğitim sistemi içinde sınıf-

landırılmalarını ve akışını düzenleyen bu tür kararlarda not, tek ölçü olmasa bile, en önemli dayanaklardan biridir. Sınıflama kararlarından başka, yarışma veya giriş sınavları gibi bazı seçme durumlarında not, ilk eleme değişkeni olarak kullanılabilir. Meselâ, bazı okulların seçme sınavlarına aday olabilmek için, «sınıfta kalmamış olmak», «sınıf geçme notu ortalaması en az iyi olmak» veya «fen derslerindeki sınıf geçme notu 7'den aşağı olmamak» gibi şartlar uygulandığı görülmüştür.

Notlar sınıflama ve seçme kararları için, genellikle, yeterince güvenilir ve geçerli ölçüler değildir. Fakat notlar, diğer ölçülerle birleştirildiği takdirde bu kararların isabet derecesini artırır.

Değerlendirmenin Üç Ögesi

Her değer yargısına bir değerlendirme süreci sonunda ulaşılır. Öğrenciye takdir edilen not, aslında, öğretmenin öğrenci başarısı hakkında ulaştığı değer yargısının bir ifadesidir. Bu nedenle, bir değer yargısına nasıl ulaşıldığını genel çizgileriyle tartışmak, not takdiri probleminin daha açıkça anlaşılmasına yardım eder.

Her değer yargısında üç öge bulunur. Bunlardan biri bir gözlem veya **ölçme sonucu**, ikincisi bir **ölçüt**, üçüncüsü ölçme sonucu ile ölçüt arasında bir **işlemdir**. Değerlendirme ölçme sonuçlarını bir ölçüte vurarak bir değer yargısına ulaşma işlemi olarak tanımlanabilir. Açıkça ifade edilmemiş olsa bile, her değer yargısının gerisinde, bir ölçme veya gözlem sonucu, bir ölçüt ve bir de gözlemi ölçütle karşılaştırma işlemi bulunur. Bir öğrencinin konuşmasını dinleyip not defterinin düşünceler hanesine «kekeliyor» yazan bir öğretmen, bir gözlemi kaydetmiş olur. Ölçmenin geniş anlamıyla, bu bir ölçme işlemidir. Bu öğretmenden o öğrencinin konuşma becerisini değerlendirmesi istense, öğretmenin değer yargısı «fena» olurdu. Bu yargının ölçütü nedir? Burada ölçüt, kekelemenin konuşmayı bozuşunda, çoğunluk kekelemediğinden kekemeli konuşmaya alışkın olmayışımızda bulunur. Kekeleyenlerin çoğunlukta olduğu bir toplumda olsaydık, kekeleme herkese normal gelebilir, kekelemeden konuşan bir öğrenciye «fena» notu verilebilirdi.

Bir dersin bütünleme sınavında, bir öğrencinin sorulan üç sorudan sadece birincisine tam cevap yazdığı, diğer iki soruyu cevapsız bıraktığı farzedilsin. Cevaplar, birinci soruya 20, ikinciye 12 ve üçüncüye 18 puan olmak üzere toplam 50 puanla puanlanmış olsun. Söz konusu öğrencinin bu sınavdaki puanı 20 olur. Bu puan onun

sınavdaki başarısının ölçüsüdür. Bu ölçme sonucuna dayanarak bir değer yargısına ulaşmak, ancak onu bir ölçütle karşılaştırmakla mümkün olur. Bütünleme sınavında başarılı sayılabilmek için, «sınavda elde edilebilecek maksimum puanın en az yarısını kazanmak» şartı konulmuş olsun. Böyle bir şart, değer yargısına ulaşmaya yardım eden bir ölçüttür. Bu misalde maksimum puanın yarısı $50/2 = 25$ eder. Öğrencinin puanı minimum başarı şartından küçük olduğundan «başarısız» yargısına ulaşılır.

Bir değer yargısının doğru olabilmesi, ölçmelerin duyarlılığına ve geçerliğine, seçilen ölçütün amaca uygunluğuna ve değerlendirme işleminin doğru yapılmış olmasına bağlıdır. Bu husus not takdirinde de böyledir. Notlar güvenilir ve geçerliği kanıtlanmış ölçme sonuçlarına dayanmalı; notun kullanılacağı amaca uygun bir ölçüte göre ve hatasız işlemle verilmelidir.

Notur. Dayanacağı Ölçme Sonuçları

Herhangi bir derste, öğrencide gelişmesi beklenen davranışlar bilgi, beceri ve tutum gibi boyutlarda analiz edilirse, bir dersin geliştirmeyi hedef aldığı davranışların çeşitli olduğu; hepsini bir tür yoklama ile ölçmeye imkân olmadığı görülür. Bu nedenle, bir tek derste bile, yoklama ve ölçme yöntemlerinin çeşitlendirilmesi gerekir. Nitekim, imtihan yönetmeliklerimizin yıllardır değişmeyen hükümlerinden biri, kanaat notunun, «sözlü yoklama», «yazılı yoklama» ve «ödev» notu gibi üç ölçmeye dayanacağı hükmüdür.

Son yıllarda müfredat programlarının hedeflerinde önemli değişiklikler olmuştur. Bilgi kazanmanın yanında, bilgiyi anlamanın ve yeni durumlara uygulamanın, bilgi ürünlerini analiz edebilmenin, bilgilerin doğruluğunu yoklayabilmenin ve bilgi üretmenin önemli bilişsel hedefler haline geldiği görülmektedir. Psiko - motor becerilerde ve duyuşsal hedeflerde de çeşitlenme görülmektedir. Bir dersin hedefleri bu derece çeşitlenince, şüphesiz, o derste başarılarının ölçülmesinde geleneksel yöntemler yetersiz kalır.

Ölçme araç ve yöntemlerinin gelişigüzel çeşitlendirilmesi iyi bir ölçme yapılmasını garantilemez. Dersin hedeflerine bağlı olarak ölçülmesi istenen davranışların türleri, o derste kullanılması uygun olan ölçme araç ve yöntemlerini tayin eder. Bu kitabın daha önceki bölümlerinde çeşitli ölçme araç ve yöntemleri ayrıntılarıyla incelenmişti için burada araç ve yöntemler üzerinde durulmayacaktır.

Ölçmelerin çeşitli oluşu not vermede **ağırlık problemi** adıyla bilinen bir problem doğurur. Birkaç ölçme sonucuna dayanılarak bir tek not takdir edilecekse, her ölçme nota hangi oranda girmelidir? Bu soru not takdirinin çetin sorularından biridir. Not takdirinde çeşitli ölçmelere eşit ağırlıklar verilmesinin veya ağırlık kavramının hiç dikkate alınmamasının yarattığı sakıncalar birçok öğretmenin zihnini meşgul etmiştir. Karne notunun, yazılı, sözlü ve ödev notlarının basit aritmetik ortalaması olarak verildiği hallerde en sık duyulan itiraz, ödev notunun geçerliği şüpheli olmasına rağmen karne notunu gereğinden fazla yükseltişidir. Bu ve benzeri itirazlar, aslında, notun dayandığı ölçmelere uygun ağırlıklar verilmek suretiyle karşılanabilir.

Çeşitli ölçmelere farklı ağırlıklar vermenin iki basamağı vardır. Birinci basamakta her ölçmeye verilecek ağırlık yüzde olarak kararlaştırılır; ikincisi basamakta ise, son notun kararlaştırılan ağırlıklarda bir **bileşke** olması için istatistiksel işlemler uygulanır. Bu kesimde ağırlık yüzdelerinin kararlaştırılmasında işe koşulabilecek bazı ilkeler verilmektedir. Ağırlıklı notun hesabı, not vermede istatistiksel yöntemlerin işlendiği, kesime bırakılmıştır.

Ağırlıkların kararlaştırılmasında genel ilke, **her ölçme sonucuna güvenilirliği ve geçerliği ile oranlı bir ağırlık vermektir**. Özel olarak şu ilkeler uygulanmalıdır :

1. Daha kapsamlı bir ölçmenin geçerliği, genellikle daha yüksektir. Kapsamı geniş olan sınavların puanlarına daha fazla ağırlık verilir.
2. Konuları birbiri üzerine yığılan derslerde, sonraki konular öncekileri de kapsadığı için, kapsamları eşit olsa bile, dönem sonu sınavlarının puanlarına dönem içi sınavlarının puanlarından daha fazla ağırlık verilir.
3. Puanlama hataları az veya puanlama güvenilirliği yüksek olan sınavların sonuçlarına daha fazla ağırlık verilir. Bu ilke uygulanırsa, aynı kapsamda olan bir test ve bir yazılı yoklamanın sonuçları söz konusu olduğunda test puanlarına daha fazla ağırlık vermek gerekir.
4. Kopya ve benzeri etkiler, ölçme sonuçlarının geçerliğini düşürür. Öğrencinin kendi ürünü olan sonuçlara, bu yönden şüpheli olan sonuçlardan daha fazla ağırlık verilir. Bu ilke uygulanırsa, okulda öğretmenin gözetiminde yapılan ödevlerin puanlarına, evde yapılan ödevlerin puanlarından daha fazla ağırlık vermek gerekir.

5. Bir sınav hedef davranışlardan yalnız birkaçını ölçüyorsa, sınavın sonuçlarına, o davranışların hedeflerin tümündeki oranına eşit ağırlık verilmelidir.

Değerlendirmede Kullanılan Ölçütler

Değerlendirmede sonuçların kullanılacağı amaca uygun bir ölçüt seçmek de güç problemlerden biridir. Öğretmenlerimiz, ölçme ve değerlendirme işlemlerini birbirinden ayıramadıklarından ve çoğu halde kendi sübjektif kanılarını farkında olmadan ölçüt olarak kullandıklarından, öğrenci başarısı hakkında hatalı değer yargısına ulaşmaktadırlar. Bu sebepten, not vermede ölçüt seçme problemi, ülkemizde daha çetin bir problem olarak ortaya çıkmaktadır. İyi seçilmiş bir ölçüt hem değerlendirme işlemini, hem de değerlendirme sonuçlarının yorumunu kolaylaştırır. Öğretmen ölçütlerin özelliklerini bilirse, değerlendirmenin amacına uygun bir ölçüt seçebilir. Not vermede kullanılan ölçütlerin başlıcaları aşağıdaki paragraflarda tanıtılmaktadır.

Öğretmenin Kanısı : Bir öğretmen, bütünlemde bir yazılı yoklama kâğıdını genel izlenimle puanlayıp «4» notunu vermiş olsun. Bu 4 hem sorulara verilen cevapların bir ölçüsü, hem de «geçti - kaldı» kararına aynen dayanak olacağı için bütünlemedeki değerlendirme- nin notudur. Soruları hazırlayan ve cevapları puanlayan kimse o öğretmen olduğu için, bu ölçmede öğretmenin bilgisi referans kabul edilmiş, öğrencinin bilgisi onunla karşılaştırılmıştır. Üstelik bu ölçme sonucu, «geçti - kaldı» kararında aynen kullanıldığından, kararı etkileyen en büyük etken, soruların hem hazırlayıcısı hem puanlayıcısı rolündeki öğretmen olmuştur. Böyle bir değerlendirmede, öğretmen, soruları güç seçtiği ve puanlamada ölçüyü «kıt» tuttuğu takdirde birçok öğrenci başarısız görünür; aksi halde öğrenciler gerçek düzeylerinden daha başarılı görünürler. Öğretmenin kanısı sabit bir referans çerçevesi olmaz.

Öğretmenin kanısı ölçüt alınarak verilen notlar, birçok bakımdan hatalı olabilir. Sınavda sorulan soruların çok zor veya çok kolay olması, sınavın kapsamlı olmaması, puanlamanın güvenilir olmayışı, vb. bu hataların kaynaklarıdır. Öğretmen kanısıyla verilen notları yorumlamak çok güçtür. Bir öğretmenin verdiği notlardan doğru anlam çıkarmak, ancak sorulan soruları ve cevapların nasıl puanlandığını ayrıntılarıyla bilmekle mümkün olabilir.

Sınıftaki Diğer Öğrencilerin Başarısı : Bir sınıfta sekiz sorulu bir yazılı yoklama yapılmış, sorulardan iki tanesi hiç bir öğrenci ta-

rafından cevaplandırılmamı. Alsun. Bu yazılıya sekiz soru üzerinden (8 tam cevaba 10) not veren bir öğretmen soruların tümünü ölçüt almış olur. Öte yandan, aynı durumda diğer bir öğretmen, hiç bir öğrencinin cevaplandıramadığı iki soruyu puanlamadan çıkararak altı soru üzerinden (6 tam cevaba 10) not takdir ederse, soruların tümünü değil, cevaplandırılan soruların tümünü ölçüt kabul etmiş olur. İkinci öğretmen not verirken sınıftaki başarı durumunu dikkate almış; «genel başarısızlık» durumunun notu etkilemesini istememiştir. Soruların sınıfa göre «zor» olması veya cevaplama süresinin «kısa» gelmesi tehlikesinden, öğretmen kanısının sübjektifliğinden ve bunlara benzer diğer etkilere kaçan eğitimciler, sınıftaki ortalama başarı düzeyini ve başarı dağılımını ölçüt kabul etmeyi düşünmüşlerdir. Bu düşünceden, puanları sınıfın ortalamasına yakın öğrencilere 5 - 6, ortalamanın bir hayli üstündekilere 7 - 8, çok üstündekilere 9 - 10, ortalamanın bir hayli altındakilere 3 - 4, çok altındakilere 1 - 2, gibi not takdir eden yöntemler doğmuştur. Yazılı yoklama kâğıtlarını kendi aralarında en iyiden en kötüye doğru sıralayıp bu sıraya göre not vermede de sınıftaki başarı sırası ölçüt alınmış olur.

Sınıf ortalamasını ve puanların sınıftaki dağılımını ölçüt alan bir değerlendirme, öğretmenin ölçüyü kıt tutması, bazı soruların veya sınavın tümünün sınıfa göre çok zor gelmesi, cevaplama süresinin çok kısa verilmesi vb. gibi etkenlerden etkilenmez. Buna karşılık, üstün başarıların çoğunlukta olduğu bir sınıfa düşen veya öğrenme olanaklarından yoksun bulunan öğrenciler, daima sınıf ortalamasının altında kalacaklarından, orta veya zayıf not alırlar. Üstelik, sınıf ortalamasına göre verilen notlar, bir sınıftaki öğrencinin diğer bir sınıftaki öğrenciyle karşılaştırılmasına imkân vermediği gibi, herhangi bir sınıfın ya da bireyin mutlak başarı düzeyini de göstermez.

Böyle yöntemlerle verilen notlar sınava veya programa göre değil, sınıfın ortalama başarısına ve puanların sınıftaki dağılımına göre anlam kazanır. Onun için bir notun yorumu da, ancak o öğrencinin sınıftaki başarı sırasını belirttiği kadar anlamlı olur.

Ülke Çapında Normlar : Sınıf ortalamasının sınıftan sınıfa büyük değişkenlik göstermesi, bireysel farklardan başka, aynı program eşit koşullarda uygulanmış olsa bile, her öğretmenin farklı güçlükte soru hazırlanmasından ve cevapları farklı ölçülerle puanlamasından doğabilir. Öğretmenler ve okullar arasındaki ölçme farklarını ortadan kaldırmanın bir yolu, ülke çapında uygulanan tek sınava ve bu sınav puanlarının ortalamasının ölçüt alınmasına gitmektir. Geçmişte böyle düşüncelerden standart başarı testleri doğmuştur.

Standartlaştırılmış ve ülke çapında uygulanmış bir testin normlarına göre not verme, ana çizgileri bakımından sınıfın ortalama başarısına göre not vermeye benzer. Referans çerçevesi, daha büyük bir grubun ortalama başarısıdır. Böyle bir yöntemle verilen notlara dayanılarak çeşitli sınıfları veya okulları birbiriyle karşılaştırmak anlamlı olur. Buna rağmen, verilen not başka bir yönden hatalı olabilir. Öğretmeni yeterli olmayan, öğretim araç ve materyalleri bulunmayan, sosyo - ekonomik olanakları sınırlı olan okulların öğrencileri, ülke çapındaki normlar uygulandığında diğerleriyle kıyaslandıklarından, çoğu halde ortanın altında not alırlar. Böyle bir not yerel kararlarda, meselâ sınıf geçme kararında kullanılırsa, bazı öğrenciler ellerinde olmayan etkenlerden dolayı sınıfta kalabilirler.

Öğrencinin Yeteneği : Yetenek ölçülerinin eğitimde geniş çapta kullanılmaya başlamasından sonra, her öğrencinin genellikle yeteneği oranında başarılı olduğu açıkça ortaya çıkmıştır. Sınıfın ortalama başarısını ölçüt alan değerlendirmelerde üstün yetenekli öğrencilerin başta, yetenekleri düşük olanların ise, bütün çabalarına rağmen, ortanın altında not aldıkları görülür. Öğrenciler arasında geniş yetenek farkları bulunduğunu, bir öğrencinin ancak yeteneğinin elverdiği sınıra kadar başarılı olabileceğini gören bazı eğitimciler, her öğrenciyi kendi yeteneğine kıyasla değerlendirmenin daha doğru olacağını savunmuşlardır. Bir misal olarak, bu ölçütle aritmetikte başarının nasıl değerlendirildiğini görelim. «Zekâ yaşı», bazı standart zekâ testlerinden elde edilen bir puandır. Zekâ yaşı 8 yıl olan bir çocuğun testteki başarısı, kendi yaşı (takvim yaşı) ne olursa olsun, takvim yaşı 8 olan çocukların ortalama başarısına denk olarak yorumlanabilir. Aynı şekilde bir «aritmetik yaşı» tanımlanabilir ve standart bir aritmetik testiyle de aritmetik yaşı ölçülebilir. Her iki ölçü elde varsa, her öğrenci için «aritmetik yaşı/zekâ yaşı» oranı hesaplanır; istenirse bu oran 100 ile çarpılır. Aritmetikte zekâsının olanak verdiği başarıyı gösteremeyen bir öğrenci için bu oran 100'den küçük, aksi halde 100'den büyük bulunur. Bir öğrenci, başarı oranının 100'den ayrılma miktarına göre, «zekâsının altında» veya «zekâsının üstünde» başarılı olarak değerlendirilir. Böyle bir değerlendirmede ölçüt, öğrencinin zekâ testinde kazandığı puandır. Şüphesiz, başarı oranlarına bakarak not verebilmek için ikinci bir ölçüt daha gerekir.

Yeteneğe göre değerlendirme veya yetenek ve diğer bir ölçüte göre değerlendirme sonuçları kolayca yorumlanabilir. Fakat, yeteneğe göre değerlendirme, kendi verdiği nota bir mutlak ölçü gözüyle bakan veya öğrencilerin başarılarını birbirleriyle karşılaştırarak not

vermeye alışmış olan öğretmenlere «adaletsiz», hattâ «gülünç» gelebilir. Bir sınavda soruların ancak yarısını doğru cevaplandırmış bir öğrencinin «zeki olmadığı için» «iyi» not alabilmesi, bu gibi öğretmenlerin anlayıp yorumlayabilecekleri bir sonuç değildir.

Öğrencinin Programa Girişteki Başarısı : Bazı derslerde öğrenciler o derste işlenen konuların bir kısmını bilerek gelirler. Okul programlarındaki derslerden ancak bazıları öğrenciler için tümü yeni bilgileri kapsar. Bir öğrencinin bir derste belli bir sürede ne kadar ilerlediğinin en anlamlı ölçüsü, şüphesiz, öğrencinin o süre sonundaki başarısı ile başındaki başarısı arasındaki farktır. Meselâ, bir kursta her kursiyer'in ne kadar öğrendiği meydana çıkarılacaksa, aynı test kursun başında ve sonunda uygulanıp iki ölçme sonucu arasındaki fark bulunur. Böylece referans çerçevesi, her öğrenci için programa girişteki başarı düzeyidir. Onun için girişteki veya başlangıçtaki puan, değerlendirme ölçütüne karışmış olur. Böyle olmakla birlikte, fark puanlarına dayanarak not takdir edebilmek için de ikinci bir değerlendirme ölçütü daha bulmak gerekir. Bu ölçüt, yukarıda tartışılan ölçütlerden birinde olduğu gibi, kursu tamamlayanların fark puanlarının ortalaması olabilir.

Fark puanlarına dayanılarak yapılan değerlendirmede mutlak başarı düzeyinin anlamı kalmaz. Mutlak başarı düzeyi yüksek olan bir öğrenci fark puanı küçük olduğu (derse hazır bilgilerle başladığı) için düşük not, mutlak başarı düzeyi düşük olan bir öğrenci de fark puanı büyük olduğu için yüksek not alabilir.

Programın Hedefleri : Eğitim - öğretim belli hedeflere ulaşmak için belli davranışları geliştirmek amacıyla yapıldığına göre, öğrenci başarısının değerlendirmede kullanılacak en akla yakın ölçüt bu hedeflerle tanımlanan davranışların tümü olmak gerekir. Fakat böyle bir ölçüt iyice analiz edilip ayrıntılarıyla tanımlanmadıkça kullanışlı olamaz. Genel hedeflerin analizinden özel hedeflere ve onları oluşturan davranış birimlerine inebilmek kolay bir iş değildir; kolay olmamakla birlikte, yapılması imkânsız da değildir. (Bakınız : Bölüm 9). Elde iyi hazırlanmış bir belirtke tablosu varsa, ondan dengeli ve sistemli bir örnekleme yoluyla sınav sorularına geçilebilir. Sınavdan elde edilebilecek en yüksek puan da ölçüt olarak kullanılabilir. Böylece, sınav sorularının tümü ölçüt alınmış olacağından, değerlendirme mutlak bir anlam kazanır. Sınav, belirtke tablosundaki davranışları temsil ettiği oranda geçerli, sınav puanlarına ve sınavın tümüne dayanılarak verilecek notlar da hedef davranışlardan ne kadarının gerçekleştiğini gösterdiği oranda anlamlı olur.

Değerlendirmede Ölçüt Seçimi

Değerlendirme sonuçları daima bir amaçla (bazen birden fazla amaçlarla) kullanılır. Seçilen değerlendirme ölçütünün işte bu kullanılış amacına uygun olması istenir. Değerlendirme ölçütü, nota dayanılarak verilecek kararın isabet derecesini artırmalı, karar vermeyi kolaylaştırmalıdır. Nitelikleri farklı olduğu için ölçütler aynı kullanışlılıkta değildir. Not vermede uygun bir ölçüt seçimi ölçütlerin iyi tanınmasıyla ve notun ne amaçla kullanılacağıının açıkça bilinmesiyle mümkündür. Bu sebeple, yukarıda başlıca özellikleri verilen ölçütlerin burada karşılaştırılmaları yararlı olur.

Bir ölçüte, ölçme sonuçlarının yorumlanmasını kolaylaştıran ve böylece değer yargısına ulaşmaya yardım eden bir **referans çerçevesi** olarak bakılabilir. Ölçme sonuçlarından çıkarılacak değer yargısının hangi referans çerçevesinde en anlamlı görüldüğü bu açıdan tartışılabilir.

Yukarıdaki paragraflarda nitelikleri tartışılan ölçütlerden, genellikle iki çeşit değerlendirme yapılabileceği kanısına ulaşılır. Bunlardan biri **mutlak değerlendirme** olup bu işlemde öğrencinin mutlak başarı düzeyi kestirilir ve ona göre not verilir. Başarıyı oluşturan davranışların mutlak sıfır noktası ve standart birimleri bulunmadığından, «mutlak başarı düzeyi» kavramının kesin bir anlamı yoktur. Böyle olmakla birlikte, öğretmenin kanısına ve programın hedeflerine dayanan iki ölçütle yapılan değerlendirmelere mutlak değerlendirme denir. Çünkü bu tür değerlendirmelerde **not**, öğrencinin belli bir bilgi ve beceriler toplamından ne kadarını öğrendiğinin ölçüsüdür. İkinci değerlendirme türü **bağıl değerlendirme** olup bu işlemde öğrenci başarısı başka bir değişkenle kıyaslanarak not takdir edilir. Bir kursun başında ve sonunda başarıyı ölçerek ikisi arasındaki farka not takdir etmek bir bağıl değerlendirmedir. Sınıf ortalamasına veya daha geniş bir norm grubunun ortalanmasına göre, ya da öğrencinin yeteneğine göre not vermek de bağıl değerlendirmelerdir. Yukarıdaki paragraflarda sözü edilen ilk ve son ölçüt mutlak değerlendirmede, aradaki dört tanesi bağıl değerlendirmede daha kullanışlı olur.

Öğretmenin kanısı ve programın hedefleri mutlak ölçütler vermekle birlikte, ikisi arasında önemli farklar vardır. Öğretmenin kanısı daha sübjektiftir; öğrenciden ne istendiği ve ne derecede bir başarının yeterli sayılacağı açıkça belirtilmediğinden, diğer bir ifadeyle referans çerçevesi değiştiğinden, ölçme ve değerlendirme ha-

talı yapılır. Öte yandan, programın hedeflerini konu ve davranışlar düzeyinde iyice analiz eden bir öğretmen, gene kendi kanılarını bu analize katmış olmakla birlikte, hangi davranışların ölçüleceğini ve ne derecede bir davranış değişikliğini yeterli sayacağını iyice saptadığında, daha objektif bir ölçme ve daha az hatalı bir değerlendirme yapabilir.

Bağıl değerlendirmeye olanak veren dört ölçüt de aynı kullanışlılıkta değildir. Bir sömestrin başında ve sonunda yapılan ölçmeler arasındaki farka dayanan bir değerlendirme, o sömestirde ne kazanıldığını ortaya çıkarmak amacıyla en kullanışlıdır. Böyle bir değerlendirme, derse giren öğrencilerin önceki bilgileri arasında büyük farklar bulunduğu durumlarda yararlı bilgiler verir.

Ölçüt olarak sınıf ortalamasını almakla, okul ortalaması, il ortalaması, ülke çapında normlar gibi geniş gruplara dayanan ortalamaları almak arasında önemli farklar vardır. Geniş gruplara dayanan normlar, bireysel puanların değerlendirilmesinde daha genel bir referans çerçevesi sağlar; ayrıca, o gruba dahil bütün alt gruplar arasında çeşitli karşılaştırmalara imkân verir. Fakat, bu tür ölçme ve ölçütlerle mahalli olarak kullanılacak notlar vermek, yukarıda tartışılan sebeplerle hatalı olur. Bölgeler ve okullar arası farkların notları etkilememesi isteniyorsa, referans grubu bir tek okula kadar indirilmelidir. Bir tek okulda öğretmenler arası farkların notlara etki etmemesi isteniyorsa, sınıf ortalaması referans alınmalıdır. Böylece, bağıl değerlendirmede referans çerçevesi bir tek sınıfa kadar indirilince, farklı öğrencilerin notlarını birbiriyle karşılaştırma olanığı kalmaz.

Öğrenciyi yeteneğine göre değerlendirmek, başarının ölçülmesinde zekâ ölçüsü gibi bir ölçüyü ölçüt kabul etmek, ancak bireysel değerlendirmelerde geçerlidir. Böyle değerlendirmeler bireysel rehberlik ve danışma hizmetlerinde çok kullanışlı bilgiler verir. Bir öğrenci, yeteneğinin elverdiği en ileri başarıya erişmişken, onun başarısını mutlak bir ölçütle «yetersiz» olarak değerlendirmek ve onu tamamlayıcı öğretim çalışmalarına sokmak doğru olmaz.

Mutlak değerlendirme için mutlak ölçüler gerekir. Elde mutlak ölçüler varsa, onları bağıl ölçülere çevirerek bağıl değerlendirme yapmak mümkündür. Fakat, bazı sayıltılara dayanmaksızın, bağıl ölçüleri mutlak ölçülere çevirmek mümkün olmaz. Bu sebeple, bağıl ölçülere dayanılarak mutlak nitelikte notlar verilemez. Mutlak nitelikteki ölçüler, her iki tür değerlendirmeye olanak verdiği için daha kullanışlıdır.

Bir deęerlendirmede birden fazla lme sonucu kullanıldıęı gibi, iki ltn bazı niteliklerini taşıyan karma bir lt de kullanılabılır. Aşağıdaki kesimlerde, not takdirindeki istatistiksel yntemler iřlenirken, karma ltle not vermede bařvurulan bir yntem tartıřılmaktadır.

NOT VERMEDE İSTATİSTİKSEL YNTEMLER

lme sonuları toplandıktan ve deęerlendirmede kullanılacak bir lt seildikten sonra, sıra deęerlendirme iřlemine gelir. lme sonularının lte gre yorumlanması, varılan deęer yargılarının sayı, harf veya sıfatlarla gsterilmesi bu iřlemin bařlıca basamaklarıdır. Not verme iřleminin esası, lme sonularını gsteren sayıların, deęerlendirme sonularını ifade eden sembollere dnřtrlmesidir.

Puanlamanın deęerlendirme lęi ile yapıldıęı hallerde, puanlama iřlemi aynı zamanda bir deęerlendirme iřlemi olduęundan, puanları rota evirme gereęi duyulmaz. Fakat, lme sonuları her zaman deęerlendirme lęiyle ifade edilmiř olmaz. lmede duyarlıęı artırmak iin, deęerlendirmede kullanılandan farklı bir lek kabul etmek gerekebilir, hattâ zorunlu olabilir. Ayrıca, lme iřlemiyle deęerlendirme iřleminin birbirinden ayrılmasında, deęerlendirme hatalarını azaltmak bakımından nemli yararlar vardır.

Bu kesimde not verme iřleminin bařlıca yntemleri, bir tek ve birka lme sonucuna dayananlar olmak zere, iki alt kesimde ele alınmıřtır. Her iki alt kesimde de mutlak ve baęıl yntemler ayrı ayrı aıklanmıřtır.

Deęerlendirme lęi

İlkokullarımızda 1 - 5 arasında beř rakamla, orta dereceli okullarımızda 0 - 10 arasında onbir rakamla not verilmesi gelenekselleřmiřtir. Bu sayılarla, **deęerlendirme lęi** adı verilebilecek birer lek tanımlanmıřtır. Kullanıřlı bir deęerlendirme lęinin yeterince duyar (gvenilir) olması, ayrıca derecelerinin btn ilgililerce aynı anlamda yorumlanması istenir.

Notların anlamında birlik saęlamak amacıyla, hem btn okullarda aynı lęin kullanılması emredilir, hem de sayısal leklerdeki notlar sıfatlarla aıklanır. Okullarımızda kullanılan notların anlamları řyle aıklanmıřtır : İlkokullarda :⁴ 5 : Pekiyi, 4 : iyi, 3 : Orta,

⁴Bařbakanlık Resmi Gazete, 14291, 27 - 8 - 1972.

2 : Geçer, 1 : Başarısız. Orta dereceli okullarda :⁵ 10 - 9 : Pekiyi, 8 - 7 : İyi, 6 - 5 : Orta, 4 - 3 : Zayıf, 2 - 1 : Pek zayıf, 0 : Boş.

Bazı okullarımızda sayısal ölçekler yerine harfler kullanılmaktadır. Notlar harflerle gösterildiğinde, her harfin sayısal değerinde bir anlaşmaya varmadan, notlara ağırlık vermek ve not ortalamasını almak gibi istatistiksel işlemler yapılamaz. Not ölçeğindeki her harfin bir sayısal eşdeğeri bulunacağına göre, harfler de sayısal nitelikte bir ölçek verir. Notları harflerle göstermenin, ancak değerlendirme ölçeğinin ölçme ölçeklerinden ayırt edilmesinde yararı olabilir. Sekizinci Milli Eğitim Şûrasına teklif edilip kabul gören bir ölçek şöyle sıfatlandırılmıştır :⁶ A : Pekiyi, B : İyi, C : Orta, D : Geçer, E : Başarısız.

Bu ölçek bir dersten sınıf geçebilecek derecede başarılı öğrencileri dört kategoriye ayırır; başarısızları kendi içinde sınıflamaz. Şûrada harflerin sayısal değerleri kesinlik kazanmamıştır. Fakat, D notu 5 notuna eşdeğer kabul edilirse, D - A notları 5 - 10 arasına; E ise 0 - 4 arasına denk olur. Demek ki, Şûrada kabul edilen ölçek eşit aralıklı değildir. Değerlendirme ölçeğinin eşit aralıklı olması istenen bir özellik olmakla birlikte, E gibi en alt kategoriye sıfır değeri verilmek kaydıyla, «başarısız» öğrenciler arasında hiçbir ayırım yapılmayacağı veya E gibi notlar üzerinde hiçbir sayısal işlem yapılmayacağı hallerde, ölçeğin eşit aralıklı olmayışından veya harflerle gösterilmesinden önemli sakıncalar doğmaz. Bu harfli ölçeğe benzer bir sayısal ölçek, «Ders Geçme ve Kredi Düzeni» nin denendiği okullarda uygulanmıştır.⁷

Notun duyarlılığı, herşeyden önce notun dayandığı ölçmelerin duyarlılığına bağlıdır. Ölçme sonuçlarının not ölçeğine çevrilmesi duyarlılığı artırmaz, fakat azaltabilir. Tümünde dört ana fikrin yoklandığı bir sınavda cevaplar en çok beş kategoriye (hiç bilmiyenler, bir ana fikri yazarlar, iki ana fikri yazarlar,...) ayrılabilir. Bu ölçmeye dayanan notlar ister 10, ister 100 esasına göre verilsin, duyarlık 0 - 4 arasında beş bölmeli bir ölçeğin duyarlığından ileri olamayacağı için, notlar hatalı olabilir. Bu halin tersi de doğrudur. Bir testin 8 ile 47 arasında dağılan puanlarına 0 - 5 ölçeği ile not verildiğinde, ölçmedeki duyarlık, değerlendirmede kaybedilmiş olur. Ölçme

⁵Milli Eğitim Bakanlığı, Milli Eğitim Bakanlığına Bağlı Orta Dereceli Okulların Sınıf Geçme ve İmtihan Yönetmeliği, s. 11 - 12.

⁶Milli Eğitim Bakanlığı, Sekizinci Milli Eğitim Şûrası : Esaslar, Raporlar, Kararlar, s. 26.

⁷Milli Eğitim Bakanlığı, Tebliğler Dergisi, 1837, 28 - 4 - 1975.

sonuçlarında karşılığı olmayan duyarlıkta bir not ölçeği kullanmak doğru değildir. Fakat, notların kullanılacağı amaca uygun ayrıntıda bir değerlendirme ölçeği kabul etmek, ölçmeleri de nottan biraz daha yüksek duyarlıkta yapmak kolayca savunulabilir.

Not ölçeğindeki kategori sayısı arttıkça, notun dayandığı ölçmeler yeterince güvenilir olmak kaydıyla, değerlendirmenin güvenilirliği de artar. Ebel'den alınan Tablo 16.1'de ölçme sonuçlarının güvenilirliği ile değerlendirmenin güvenilirliği arasındaki ilişkinin, not ölçeğindeki kategori sayısı ile arttığı görülür.⁸

TABLO 16.1
ÖLÇME GÜVENİRLİĞİNE BAĞLI OLARAK
DEĞERLENDİRME GÜVENİRLİĞİ

Ölçme güvenirliği	Değerlendirme kategorileri			
	2	5	10	15
0,95	0,63	0,85	0,92	0,94
0,90	0,60	0,80	0,87	0,89
0,80	0,53	0,71	0,78	0,79
0,70	0,47	0,62	0,68	0,69
0,50	0,33	0,45	0,48	0,49

Eğitimde kullanılan standart testlerin güvenilirlik katsayıları 0,90 civarında olabilir. Öğretmenlerin hazırladığı standardlaştırılmamış araçlarda güvenilirlik, genellikle daha düşüktür. Notların dayandığı ölçme sonuçlarının güvenilirliği 0,80 civarında olduğu farzedilirse, 0 - 10 ölçeği ile kesirsiz not vermenin önemli duyarlık kaybına sebep olmayacağı yukarıdaki tabloda görülmektedir.

Bir Tek Ölçmeye Dayanan Değerlendirme

Bir sınavda aşağıdaki puanlar elde edilmiş olsun :

28	26	26	25	25	25	24	24	24	23	23	22	22	22
22	22	21	21	21	21	20	20	20	20	20	19	19	19
19	19	18	18	17	17	17	16	16	15	15	14	12	11

Yalnız bu puanlara dayanılarak ve 0 - 10 ölçeği ile not verilecekse nasıl bir işlem yapılmalıdır? Şüphesiz, puanlara uygulanacak

⁸R.L. Ebel, *Measuring Educational Achievement*, s. 423.

değerlendirme işlemi seçilen değerlendirme ölçütünün gerektirdiği bir işlem olmalıdır. Bu kesimde, biri mutlak ikisi bağıl olmak üzere, üç ölçütle değerlendirme işlemleri verilmektedir.

Mutlak Başarı Oranıyla Not Verme : Yukarıdaki puanlar en çok 40 puan verebilen bir sınavda elde edilmiş olsun. Öğretmen sınavın tümünü ölçüt alırsa, daha önce belirtildiği gibi, bir mutlak değerlendirmeye karar vermiş olur, $40/10 = 4,0$ gibi bir katsayı bulup ham puanları bu katsayıyla bölmek, not vermeye yeter. Böylece $28/4 = 7$; $26/4 = 6,5 \approx 7$; $25/4 = 6,25 \approx 6$, ... gibi notlar verilir.

Herhangi bir öğrencinin ham puanı X_i ve o sınavdan elde edilmesi mümkün en yüksek puan L ise, X_i/L oranına **mutlak başarı oranı**, bu oranın 100 katı olan $100 X_i/L$ oranına da **mutlak başarı yüzdesi** denir. Gerek mutlak başarı oranı gerekse mutlak başarı yüzdesi ham puanlardan türetilen diğer bir puan olup, kendi başına not vermeye yetmez. Hangi yüzdenin «iyi» ve hangisinin «orta» olacağı, daha genel bir ifadeyle notların hangi başarı düzeylerine verileceği de kararlaştırılmalıdır. Yukarıdaki puanlara uygulanan ve ülkemizde gelenekselleşmiş olan yöntemle, başarı yüzdelerine, 0 - 10 ölçeğinde Tablo 16.2'deki gibi not verilir.

TABLO 16.2
MUTLAK BAŞARI YÜZDELERİYLE VERİLEN NOTLAR

G E Ç E R		K A L I R	
%	Not	%	Not
95 - 100	10	35 - 44	4
85 - 94	9	25 - 34	3
75 - 84	8	15 - 24	2
65 - 74	7	5 - 14	1
55 - 64	6	0 - 4	0
45 - 54	5		

Böyle bir yöntemle yapılan mutlak değerlendirmenin en belirgin özelliği, «geçer - kalır» sınırını yüzde kırk beşten çizmiş olmasıdır. Bu sınırı daha yüksek bir başarı düzeyinden çizen yöntemler de vardır. Amerika Birleşik Devletlerinde, bağıl değerlendirme yöntemleri yaygınlaşmadan önce, Tablo 16.3'teki bir yöntemle not vermenin yaygın olduğu belirtilmiştir.⁹

⁹H. Beaumont and F.G. Macomber, *Psychological Factors in Education*, s. 237-262.

TABLO 16.3
MUTLAK BAŞARI YÜZDELERİYLE DİĞER BİR DEĞERLENDİRME

G E Ç E R		K A L I R	
%	Not	%	Not
93 - 100	A	0 - 69	F
86 - 92	B		
79 - 85	C		
70 - 78	D		

Bu tabloda görüldüğü gibi, böyle bir değerlendirmede «geçer» not alabilmek için, sınavda elde edilebilecek en yüksek puanın en az yüzde yetmişine erişmek gerekmektedir.

Sınavın sınıfa göre çok güç olması, bazı soruların sınıfta hiç bir öğrenci tarafından doğru cevaplandırılmaması, cevaplama süresinin yetersizliği gibi öğrencinin sorumlu tutulamayacağı etkiler, testin tümü ölçüt alınarak yapılan değerlendirmede bütün sınıfın notlarının düşmesine sebep olur. Bu sakıncalardan kaçınan eğitimciler, sınıfın başarısını ölçüt alan değerlendirmeyi daha uygun bulmuşlardır. En çok kullanılan yöntem sınıftaki puan dağılımına göre not verme yöntemidir.

Sınıftaki Puan Dağılımına Göre Not Verme : Sınıfın ortalama başarısını ölçüt alarak sınıftaki puan dağılımına göre not verme, başlıca iki sayılıtya dayanır. Bunlardan birinde **puanı sınıf ortalaması civarında olan öğrencilerin gerçekten orta başarıda olduğu**; ikincisinde ise, **test puanları dağılımının sınıftaki başarı dağılımına denk olduğu** kabul edilir. Test puanlarının ortalaması ($\bar{X}$) ve standart kayması (S) hesaplandıktan sonra, 0 - 10 ölçeğinde verilecek notların sınırları şöylece saptanır :

Sınır	$\bar{X}$	$\bar{X} + S$	$\bar{X} + 1,5S$	$\bar{X} + 2 S$	$\bar{X} + 2,5 S$	$\bar{X} + 3 S$
Not	6	7	8	9	10	
Sınır	$\bar{X}$	$\bar{X} - S$	$\bar{X} - 1,5S$	$\bar{X} - 2 S$	$\bar{X} - 2,5 S$	$\bar{X} - 3 S$
Not	5	4	3	2	1	0

Bu kesimin başında verilen puanların ortalaması 20,1 ve standart kayması 3,8 bulunmuştur. Yukarıdaki modele göre şu sınırlar saptanır :

Sınır	10,6	12,5	14,4	16,3	20,1	23,9	25,8	27,7	
Not	1	2	3	4	5	6	7	8	9

Bu dağılıma göre, ham puanlara şu notlar verilir :

28 : 9	24 - 25 : 7	17 - 20 : 5	13 - 14 : 3
26 - 27 : 8	21 - 23 : 6	15 - 16 : 4	11 - 12 : 2

Bu yöntemle not verilirken ortalama ve standart kayma yüzdelik basamağa kadar duyarlıkla hesaplanmalı ve ondalık basamak yuvarlaklaştırmayla saptanmalıdır. Sınırların tam sayılar halinde yuvarlaklaştırılması gerekmez. Fakat, bir sınır tam sayı halinde çıkarsa öğrenci lehine hareket edilmelidir. Öğretmen, gerekli gördüğü hallerde, sınırları öğrenci lehine bir - iki puan değiştirebilir. Yukarıdaki örnekte 20 puanına 5 yerine 6 verilebilir.

Puanların dağılımı simetrik ise, not dağılımı da simetrik olur. Ayrıca, puan dağılımı çok çarpık değilse, 5 ve 6 notlarını alanların sayısı diğer notlara göre fazla olur. Ham puan dağılımının çarpık olduğu hallerde, notların dağılımı da aynı yönde çarpık olur.

Not sınırlarını ortalama dan itibaren eşit $0.5S$ basamaklarıyla saptayan bir yöntem de vardır. Bu yöntemde 5 ve 6 gibi notların frekansları azalır; buna karşılık «7 ve yukarısı» ile «4 ve aşağısı» kategorilerinin frekansları artar. Yönetmeliklerimizde olduğu gibi «4 ve aşağısı» başarısız sayılırsa, bu yöntemde başarısızlık yüzdesi değerine göre daha fazla olur.

Ham puanlar dağılımının simetrik olduğu hallerde bu yöntemle not verilirse, sınıftaki öğrencilerin yaklaşık olarak üçte biri 5, üçte biri 6, diğer üçte biri de gittikçe azalan frekanslarda 4, 3, ... ve 7, 8, 9, ... notlarını alırlar.

Sınıfın ortalamasına ve standart kaymasına göre not verme, diğer not ölçeklerinde de uygulanabilir. Yukarıda sözü edilen 1 - 5 ölçeğinin kullanıldığı hallerde, şu sınırlarla not verilebilir :

Sınır	$\bar{X} - 1,5 S$	$\bar{X} - 0,5 S$	$\bar{X} + 0,5 S$	$\bar{X} + 1,5 S$	
Not	1	2	3	4	5

Bu sınırlarla not verilirse, ham puan dağılımının normal olduğu bir sınıfta, öğrencilerin yaklaşık olarak yüzde yedisi 1, yüzde yirmi

dördü 2, yüzde otuz sekizi 3, yüzde yirmi dördü 4 ve yüzde yedisi 5 notunu alır.

Sınıftaki puan dağılımına göre not verme, sınavın tümünü ölçüt alan değerlendirmenin sakıncalarını gidermekle birlikte, bu yöntemle özgü iki sakıncayı da birlikte getirir.

1. Bu yöntemde ham puanların ortalaması ve standart kayması önemli rol oynar. Puanların bir uca yığılması ortalamayı o uca doğru çeker. Sınıfta çok düşük veya çok yüksek birkaç puan bulunması ortalamayı etkilediği gibi standart kaymanın da büyümesine sebep olur. Bu etkiler notları belli yönlerde değiştirir. Böylece, öğrencinin notu sadece kazandığı ham puanın bir sonucu olarak değil, sınıftaki diğer öğrencilerin puanlarından da etkilenerek ortaya çıkar.

2. Bu yöntemle not vermede, sınavdaki soru sayısı veya elde edilmesi mümkün olan en yüksek puan dikkate alınmaz. Yukarıdaki misalde ham puanlar en çok 40 puan değil de 75 puan verebilen bir sınavdan elde edilmiş olsaydı, bu yöntemle ham puanlara yine aynı notlar verildi. Verilen not mutlak başarı düzeyini göstermez.

Yukarıdaki yöntemlerde aritmetik ortalama ve standart kayma yerine ortanca ve çeyrek kayma kullanılırsa, ham puan dağılımının şekli nasıl olursa olsun, öğrencilerin belirli yüzdeleri daima aynı notları alır. Örneğin, 6 notu verilen Ort. ve Ort. + Q aralığına daima sınıfın yüzde yirmi beşi düşer. Böylece puanların sınıftaki dağılıma özeliği notlara aynen yansıtılmamış olur. Ortanca ve standart kayma ile not veren bir yöntem önerilmiştir.¹⁰ Bu yöntem ortancayı kullanmakla aşırı uçlardaki puanların etkisini notlara yansıtmaz; fakat, standart kaymayı kullanmakla da bu etkiyi notlara yansıtır. Bunun için yöntem kendi içinde çelişkilidir; kullanılması önemli hatalara sebep olabilir.

Sınıftaki puan dağılımına göre not vermede kullanılan başka yöntemler de vardır. Bunlardan biri, her notun sınıftaki öğrencilerin yüzde kaçına verileceğini önceden saptayan yöntemdir. Dressel ve Nelson, bir üniversitede birden fazla şubesi olan derslerde, her şubede şöyle bir not dağılımı kullanıldığını belirtirler:¹¹ O şubedeki öğrencilerin en başarılı % 0 - 15'ine A, sonra gelen % 20 - 30'una

¹⁰H. Tekin, *Eğitimde Ölçme ve Değerlendirme*, 2. Baskı, s. 293 - 298.

¹¹P.L. Dressel and C.H. Nelson, «Testing and Grading Policies», in R.L. Dressel and associates, *Evaluation in Higher Education*, s. 227 - 252.

B, sonra gelen % 40 - 50'sine C, sonra gelen % 10 - 20'sine D ve en başarısız % 0,- 10'una F verilir.

Ham puanları, **standart —9 ölçeği** denilen ölçeğe çevirerek not verme de bu yöntemlerden sayılır. Bu ölçekle verilen notlar 1 ile 9 arasında değişir. Notların aritmetik ortalaması 5, standart kayması yaklaşık olarak 2 ve dağılımı hemen hemen normal olur. Yüzde olarak her notu alanların dağılımı, en başarılı öğrencilerden en başarısız öğrencilere doğru şöyle olur : ¹² % 4 : 9, % 7 : 8, % 12 : 7, % 17 : 6, % 20 : 5, % 17 : 4, % 12 : 3, % 7 : 2, % 4 : 1.

Sınıftaki başarı dağılımının normal olması gerektiğini gösterir sağlam bir kanıt bulunmadıkça, ham puan dağılımını normalleştirerek not verme savunulamaz. Ancak, sınavda sorulan soruların güçlük dereceleri normal bir dağılım gösteriyorsa ve sınav çok kalabalık gruplara uygulanmışsa ham puanlar normalleştirilebilir.

Genellikle, yüzdeleri önceden saptanmış bir dağılımla not vermek, sınıftaki mutlak başarı düzeyi ile başarı dağılımı dikkate alınmadığından sakıncalıdır.

Mutlak Başarı Düzeyi ile Bağlı Değerlendirme : Mutlak değerlendirme sınavın sınıfa göre çok güç geldiği veya cevaplama süresinin yetersizliği gibi etkilerle başarının düştüğü hallerde, öğrencide olmayan kusurlardan dolayı notları düşürdüğü için sakıncalıdır. Bağlı değerlendirme ise, sınıftaki mutlak başarı düzeyinin anormal derecede düşük olduğu hallerde «başarısız» öğrencileri başarılıymış gibi gösterdiğinden; sınıftaki mutlak başarı düzeyinin çok yüksek olduğu hallerde de bazı orta öğrencileri «başarısız» gösterdiğinden sakıncalıdır. O halde, hem sınıftaki başarı dağılımını, hem de sınıfın mutlak başarı düzeyini dikkate alan bir yöntem, bu sakıncaların her ikisini de bir dereceye kadar azaltabilir. Ebel,¹³ sınıfın bütün derslerdeki ortalama başarısına göre not dağılımının yüzdelerini ayarlayan bir yöntem önermiştir. Bu yöntemde, ortalama başarı düzeyi yüksek bir sınıfta iyi ve pekiyi notlarının frekansları, ortalama başarı düzeyi düşük bir sınıfta ise orta ve zayıf notlarının frekansları simetrik bir dağılımdakine göre daha büyük alınır.

Ebel'in düşüncesi bir tek derste bir tek sınava dayanılarak verilen notlara da uygulanabilir. Şüphesiz, bir tek testle, hem test güç-

¹²Bu ölçekte not vermenin ayrıntıları için bakınız : M.F. Turgut, *Değerlendirme ve Not Verme Metotları*, s. 58 - 64.

¹³Ebel, op. cit., s. 428 - 429.

lüğünü hem de ortalama başarı düzeyini hesaplamak mümkün değildir. Sınıfın ortalama başarısı ve test güclüğü, biri diğerine göre tanımlanan iki kavramdır. Bir an için, test güclüğünden söz edilmeksizin, sınıfın o testteki ortalama başarı düzeyinin kestirileceği farzedilsin; her maddesine 1 verilerek puanlanan K maddelik bir testte mümkün olan en yüksek puan $L = K$ olur. Bir sınıfın böyle bir testteki ortalaması $\bar{X}$ ile gösterilirse, $\bar{X}/L$ oranı, **testin tümüne göre sınıfın ortalama başarı düzeyi** denilebilecek bir oran verir. Bu oran 0,50 civarında ise sınıf orta başarıda bir gruptur. Bu sayı 0,50'den büyük olduğu oranda «iyi», küçük olduğu oranda «zayıf» bir grubu gösterir. Ortalama başarısı iyi bir sınıfta 7, 8, 9 gibi notların, ortalama başarısı zayıf bir sınıfta ise 4, 3, 2 gibi notların frekanslarının «orta» bir sınıftakine göre daha fazla olacağı düşünülürse; not dağılımını $\bar{X}/L$ oranına göre ayarlamak akla yakın gelir. Tablo 16.4 böyle düşüncelerle hazırlanmıştır. Bu tablo hazırlanırken $\bar{X}/L$ oranının ortcya yakın (0,45 - 0,54) değerlerinde not dağılımının simetrik olduğu; $\bar{X}/L$ oranı büyüdükçe notların 7, 8, 9, ... gibi yüksek kategorilere, $\bar{X}/L$ oranı küçüldükçe 4, 3, 2, ... gibi düşük kategorilere kayacağı farzedilmiştir.

TABLO 16.4
ORTALAMA BAŞARI DÜZEYİ İLE BAĞIL DEĞERLENDİRME

Notlar		10	9	8	7	6	5	4	3	2	1
$\bar{X}/L$											
0,75 ve yukarı	%	8	12	16	22	20	10	6	3	2	1
(% Toplamları	$\Sigma\%$	8	20	36	58	78	88	94	97	99	100
0,65 - 0,74	%	6	10	14	18	23	12	8	6	2	1
	$\Sigma\%$	6	16	30	48	71	83	91	97	99	100
	%	4	6	12	15	21	18	12	8	3	1
0,55 - 0,64	$\Sigma\%$	4	10	22	37	58	76	88	96	99	100
	%	3	4	10	13	20	20	13	10	4	3
0,45 - 0,54	$\Sigma\%$	3	7	17	30	50	70	83	93	97	100
	%	1	2	6	8	12	23	18	14	10	6
0,35 - 0,44	$\Sigma\%$	1	3	9	17	29	52	70	84	94	100
	%	1	2	3	6	10	20	22	16	12	8
0,34 ve aşağı	$\Sigma\%$	1	3	6	12	22	42	64	80	92	100

Böyle bir tabloyla not vermek, aslında, not yüzdeleri önceden saptanmış bir dağılımla not vermektir. Meselâ, $\bar{X}/L$ oranının 0,55 - 0,64 arasında olduğu bir grupta, grubun yüzde dördüne 10, yüzde altısına 9, ... verilir. Bu yöntemir ortalama ve standart kayma ile not verme yönteminden farkı, ham puanların standart kayması ne olursa olsun, not dağılımının değişmemesidir. Ham puan dağılımının $\bar{X}/L$ oranıyla girilen satırdaki yığılmalı yüzdelerden çok ayrılması halinde, bu bağıl değerlendirme hatalı olur.

Tablo 16.4 kullanılarak not vermek için önce $\bar{X}/L$ oranı hatasızca hesaplanmalıdır. Ham puanlarda şans başarısı varsa, aritmetik ortalama gerçek değerinden büyük çıkar. Her biri a seçenekli K maddeden oluşan bir testte, şans başarısından arıtılmış ortalama

$$(16.1) \quad \bar{X}_c = (\alpha \bar{X}_c - K) / (\alpha - 1)$$

formülüyle hesaplanır.¹⁴ Burada $\bar{X}_c$ şans başarısı etkisindeki puanların aritmetik ortalamasıdır. Bu formül ham puanlarda şans başarısı bulunduğu hallerde uygulanmalıdır. Cevap kâğıtlarının gözden geçirilmesi, ham puanlarda önemli ölçüde şans başarısı bulunup bulunmadığı hakkında ipuçları verebilir. Cevap kâğıtlarında **çekinilmiş** (cevaplayıcılar tarafından okunduğu halde cevapsız bırakılmış) veya **erişilememiş** (cevaplanmış en son maddeden sonda gelen cevapsız maddeler) madde sayısı önemli derecede büyük ise, cevaplayıcıların rasgele cevaplama kaçınmış olmaları daha muhtemeldir. Bu hallerde ortalama düzeltilmemelidir.

Erişilememiş madde sayısının önemli derecede büyük olduğu veya bazı maddelerin baskı ve anahtar hatası gibi sebeplerle puanlanmadan çıkarıldığı hallerde, değerlendirmeye dahil edilmemesi gereken maddelerin sayısı, testten elde edilebilecek maksimum puanın hesabında dikkate alınmalıdır. Grupta hiçbir cevaplayıcının cevaplandırmadığı madde sayısı E'nin önemli derecede büyük olması, hem rasgele cevaplama yapılmadığının hem de testi cevaplamak için verilen sürenin yetersizliğinin belirtisi sayılabilir. Böyle olunca, $\bar{X}/L$ oranı hesaplanırken aritmetik ortalama düzeltilmemeli ve $L = K - E$ alınmalıdır. Tablo 16.4 dağılımlarına göre not verebilmek için ham puanların yığılmalı yüzdeleri de, Tablo 16.5'de olduğu gibi, puanlar gruplanmaksızın hesaplanmalıdır.

¹⁴M.F. Turgut, Şans Başarısının Test Puanlarına Etkisi s. 28.

TABLO 16.5
HAM PUANLARIN YIĞMALI YÜZDELERİNİN HESAPLANMASI

Ham puan	Frekans	Yığmalı frekans	Yığmalı yüzde	Ham puan	Frekans	Yığmalı frekans	Yığmalı yüzde
28	1	1	2	19	5	30	71
26	2	3	7	18	2	32	76
25	3	6	14	17	3	35	83
24	3	9	21	16	2	37	88
23	2	11	26	15	2	39	93
22	5	16	38	14	1	40	95
21	4	20	48	12	1	41	98
20	5	25	60	11	1	42	100

Yüzdeleri önceden saptanmış bir dağılımla not vermek için gerekli istatistiksel işlem, ham puanlar yığmalı dağılımının not dağılımı yüzdelere uydurularak dilimlenmesinden başka bir şey değildir. İşlemin basamakları, bu kesimin başında verilen ham puanlar örnek alınarak aşağıda sıralanmıştır.

1. $\bar{X}/L$ oranı gerekli düzeltmeler yapıp hesaplanır. Gruptaki 42 öğrencinin cevap kâğıtları gözden geçirilip 4 maddenin hiçbir öğrenci tarafından cevaplandırılmadığı, 2 maddenin hatalı olduğu için değerlendirmeye katılmadığı kabul edilsin. Bu durumda ortalamının düzeltilmesi gerekmez. $\bar{X} = 20,1$ ve $L = 40 - (4 + 2) = 34$ olduğu için $\bar{X}/L = 20,1/34 = 0,59$ bulunur.

2. Tablo 16.4'te 0,59 oranına karşı gelen satırlardan, notlar ve yığmalı yüzdeleri diğer bir tablonun ilk iki sütununa kopya edilir. (Bakınız : Tablo 16.6).

3. Ham puanların daha önce hesaplanmış olan yığmalı yüzdeleri, Tablo 16.5'den yararlanılarak ve not yığmalı yüzdesini içine alacak sınırlar belirtilerek, Tablo 16.6'nın üçüncüsü sütununa geçirilir.

4. Notların yığmalı yüzdesine en yakın ham puandan (veya öğrenci lehine hareket edilerek, o yüzdeyi içine alan en küçük ham puandan) kesmek suretiyle, her nota karşılık olan puan bulunur. (Tablo 16.6'da son iki sütuna bakınız). Kullanılan örnekte grubun yüzde dördüne 10 vermek gerekmektedir; 10 notu 28 ham puandan kesilse grubun sadece yüzde ikisine, 26'dan kesilse grubun yüzde yedisine 10 verilmiş olur. Onun için, 10 notunun en düşük puanı 27

olmalıdır. (Bu grupta 27 ile 10 alan hiç kimse yoktur.) Aynı düşüncelerle 25 ve 26 puanlarına 9 verileceği ortaya konur. Bu örnekte kritik puanlardan biri 23'dür. Bu puan «not yığmalı yüzdesine en yakın olma» ölçütüyle 8 notunun dışında kalır; fakat, öğrenci lehine hareket edildiği takdirde, 8 notuna dahil edilebilir. Böyle hallerde öğretmen, öğrenci lehine hareket ederek «tartışmalı» durumdan kurtulmalıdır.

TABLO 16.6
YÜZDELERİ SAPTANMIŞ BİR DAĞILIMLA NOT VERME

Not	Notun yığmalı yüzdesi	Ham puan yığmalı yüzdesi	Ham puan	Not
10	4	2 - 7	28 - 27	10
9	10	7 - 14	26 - 25	9
8	22	21 - 26	24 - 23	8
7	37	26 - 38	22	7
6	58	48 - 60	21 - 20	6
5	76	71 - 76	19 - 18	5
4	88	83 - 88	17 - 16	4
3	96	95 - 98	15 - 12	3
2	99	99 - 100	11	2
1	100	—	—	—

Birkaç Ölçmeye Dayanan Değerlendirme

Ancak seyrek hallerde bir tek ölçmeye dayanılarak not verme zorunluğu doğar. Dönem veya sınıf geçme notu verme gibi çok sık karşılaşılan hallerde, değerlendirme birden fazla ölçme sonucuna dayanır. Birden fazla ölçme sonucuna dayanarak not verme, ölçme sonuçlarının birleştirilmesini gerektirir. Bu da hem ölçme sonuçlarının birleştirilmesi için uygun bir işlem seçilmesi, hem de ölçmelere verilecek ağırlıkların kararlaştırılması gibi iki problem doğurur. Bu kesimde birkaç ölçmeye dayanarak not vermede uygulanabilecek başlıca yöntemler tartışılmaktadır.

Basit Ortalamayla Not Verme : Okullarımızda pek sık kullanılan yöntem, ödev, yazılı sınav, sözlü yoklama gibi ölçme sonuçlarının herbirini not ölçeğinde ifade etmek; sonra bunların basit aritmetik ortalamasını alarak tek nota ulaşmaktır. Basit aritmetik ortalama ile

not vermenin başlıca iki sakıncası vardır. Bunlardan biri, basit aritmetik ortalamaya giren ölçmelerin ne oranlarda ağırlık aldığıının bilinmemesidir. Oysaki, geçerlik ve güvenilirlikleri birbirinden farklı olan ölçmelerin, son nota, geçerlik ve güvenilirlikleri oranında girmesi gerekir. İkincisi, sayısal değeri ve bu sebeple standart kayması daha büyük olan ölçmenin, basit aritmetik ortalama da daha fazla ağırlık almasıdır. Böylece son not, öğretmenin istediği ağırlıklarda bir bileşke değil, ortalamaya giren sayıların niteliklerinde bir bileşke haline gelir.

Not Ölçeğindeki Ölçmelere Ağırlık Verme : Yukarıda sözü edilen sakıncalardan birincisi, ortalama hesaplanırken çeşitli ölçmelere geçerlikleri oranında ağırlıklar vererek önlenabilir. Bir derste son notu belirleyecek iki yazılı, iki sözlü, iki ödev puanı saptanmış olduğu ve puanların 0 - 10 arasında not ölçeğiyle verildiği bir sınıfta, bir öğrencinin puanları $Y_1 = 6$, $Y_2 = 4$; $S_1 = 6$, $S_2 = 5$; $\bar{O}_1 = 9$, $\bar{O}_2 = 10$; olsun. Öğretmen, yazılı sonuçlarının % 50, sözlü sonuçlarının % 40 ve ödevlerin % 10 oranında son nota girmesini kararlaştırmış olsun. Bu durumda nihai not, bir ağırlıklı ortalama olarak hesaplanmalıdır. Son not, önce her tür puanın kendi içindeki basit ortalaması bulunup sonra basit ortalamaların ağırlıklı ortalaması alınarak hesaplanabilir. Örnekteki durumda bu hesaplama şöyle yapılır :

Yazılı, sözlü ve ödev puanlarının kendi ortalamaları, sırasıyla, $Y = (Y_1 + Y_2)/2 = (6 + 4)/2 = 5,00$; $s = (S_1 + S_2)/2 = (6 + 5)/2 = 5,50$; $\bar{O} = (\bar{O}_1 + \bar{O}_2)/2 = (9 + 10)/2 = 9,50$ bulunur. Bu ortalamaların ağırlıklı ortalaması, yukarıdaki ağırlıklarla

$$A.O. = \frac{(5,0 \times 50) + (5,5 \times 40) + (9,5 \times 10)}{50 + 40 + 10} = 5,65 \approx 6$$

olur.

Herbiri not ölçeğinde olan ölçme sonuçlarının ağırlıklı ortalamasını almak, ağırlık problemini çözmekle birlikte, çeşitli ölçmelerin aynı birimle ifade edilmemelerinden doğan hataları önlemez. Nitekim, bir kısmı arşın, bir kısmı yarda olarak ölçülmüş uzunlukların ortalaması anlamlı bir birimde çıkmaz. Onun için, ağırlıklı ortalama hesaplanmadan önce, ölçme sonuçları ortak bir ölçeğe çevirilir.

Ham Puanlara Ağırlık Verme : Ham puanlar genellikle birbirinden farklı birimlerde elde edilir. Bir derste laboratuvar çalışmalarına, ara sınavlarına ve son sınava dayanılarak not verilecek olsun. Bu ölçmelerin dizi genişlikleri ve standart kaymaları hesaplanıp aşağıdaki gibi bulunmuş olsun : (Bu örnekte puanlar kasıtlı olarak büyük alınmıştır.)

	Dizi geniřlięi	Standart kayma
Laboratuvar alıřmaları	63 - 86	4,2
Ara sınavı	32 - 147	21,0
Son sınav	48 - 105	10,5

Öęretmen, laboratuvar ve ara sınavı puanlarına eřit aęırlık, meselâ % 30 ve % 30, vermeyi kararlařtırmıř olsun. Bir öęrencinin laboratuvar alıřmalarında en yüksek, ara sınavında en düşük puanı aldıęı; ikinci bir öęrencinin de laboratuvar alıřmalarında en düşük, ara sınavında en yüksek puanı aldıęı farzedilsin. Bu iki ölçmeye eřit aęırlık verildięinde, iki öęrencinin toplam puanları eřit olmalıdır; ünkü herbiri bir ölçmede kaybettięi puanı ötekinde telâfi etmiř ve sınıfın ortalarında bir toplam puan kazanmıř olur. Ham puanlar aęırlıksız olarak toplanırsa, birinci öęrencinin toplam puanı $86 + 32 = 118$, ikincinininki $63 + 147 = 210$ olur. Görüldüęü gibi, bu iki toplam puan ok farklıdır. İkinci öęrencinin toplam puanı, onun daha başarılı olduęu ara sınavının ham puanları daha büyük sayılarla ifade edildięi, daha fazla daęıldıęı için büyümüřtür. Böylece, ham puanlar olduęu gibi toplanmakla birlikte, standart kayması büyük olan puan daha fazla aęırlık almıřtır. Oysaki, öęretmen bu iki ölçmeye eřit aęırlık vermeyi kararlařtırmıřtı. Ham puanlar kendi standart kaymalarıyla bölünüp toplanırsa bu hata önlenir. Böylece řu toplam puanlar bulunur :

$$\text{Birinci öęrencinin puanı} = (86/4,2) + (32/21,0) = 22,0$$

$$\text{İkinci öęrencinin puanı} = (63/4,2) + (147/21,0) = 22,0$$

Ham puanların aęırlıklı toplamaları alınırken, her puan hem o puanların standart kaymasıyla bölünecek, hem de o puanlara verilmesi kararlařtırılan aęırlıkla arpılacaktır. Hesaplamaları kolaylařtırmak için, ham puanların herbirini standart kayma ile bölmek yerine, kararlařtırılan aęırlık standart kayma ile bölünür; böylece elde edilen düzeltilmiř aęırlık ham puanlarla arpılır. Yukarıdaki örnekte kararlařtırılan aęırlıklar řöyle idi : Laboratuvar alıřmaları = $b_1 = \% 30$; ara sınavı = $b_2 = \% 30$; son sınav = $b_3 = \% 40$. Bu aęırlıklar ham puanların standart kaymalarıyla bölünürse, düzeltilmiř aęırlıklar elde edilir. Örnekteki puanlar için düzeltilmiř aęırlıklar řöyle hesaplanır :

$$W_1 = b_1/S_1 = 30/4,2 = 7,1$$

$$W_2 = b_2/S_2 = 30/21,0 = 1,4$$

$$W_3 = b_3/S_3 = 40/10,5 = 3,8$$

Düzeltilmiř aęırlıklar aralarındaki oranlar ok deęiřtirilmemek řartıyla yuvarlaklařtırılabilir. Yukarıdaki örnekte $W_1 = 10$, $W_2 = 2$,

$W_3 = 5$ alınabilir. Ağırlık yüzdeleri zaten sübjektif olarak kararlaştırıldığından, düzeltilmiş ağırlıkların böyle kabaca alınması önemli hatalar doğurmaz. Hatta, standart kaymaların bile kaba bir yöntemle kestirilmiş olması da önemli hatalara sebep olmaz.

Ham puanlara verilecek ağırlıklar böylece hesaplandıktan sonra, her ham puan kendi ağırlığıyla çarpılarak, ağırlıklı toplam puanlar bulunur. Böylece birçok ölçme sonucu her öğrenci için bir tek ağırlıklı puan halinde birleştirilmiş olur. Ağırlıklı toplam puanlar da, bir tek ölçme sonucuymuş gibi, seçilecek bir değerlendirme işlemiyle nota çevrilebilir. Meselâ, Tablo 16.4'deki yüzdelerle not verilmesi kararlaştırılırsa, önce $\bar{X}/L$ oranının ağırlıklı olarak hesaplanması gerekir. Üç ölçme halinde, $\bar{X}$ ve L 'de gerekli düzeltmeler yapıldıktan sonra ağırlıklı ortalama başarı oranı, şu formülle hesaplanır :

$$A.O.B. = \frac{W_1\bar{X}_1 + W_2\bar{X}_2 + W_3\bar{X}_3}{W_1L_1 + W_2L_2 + W_3L_3}$$

Bundan sonra ağırlıklı toplam puanların yığılmalı yüzdeleri hesaplanır. Daha sonra da, yüzdeleri önceden saptanmış bir dağılımla not verme yönteminin işlemleri sırayla uygulanır.

Ham Puanları Standartlaştırarak Ağırlık Verme : Ham puanların standart puanlara çevrilmesi, birçok halde, z - puanı denilen puanlar hesaplanarak yapılır. Ham puan dağılımı nasıl olursa olsun, z - puanlarının ortalaması sıfır ve standart kayması bir olur. Bu özellik z - puanlarının herhangi bir ağırlıkla çarpılıp toplanmasına veya doğrudan doğruya cebirsel toplamalarının alınmasına olanak verir. Ölçmeler önce z - puanlarına çevrilirse, sonra istenen ağırlıklarla z - puanları çarpılarak, z - puanlarının da ağırlıklı toplamaları bulunabilir. Bu toplam puanlar da, bir tek ölçme sonucu gibi, bağıl değerlendirme yöntemlerinin biriyle nota çevrilebilir.

Ağırlıklı toplam puanların hesaplanmasında en duyar yöntem, her ölçmenin z - puanlarını hesaplayıp onlara ağırlık vermektir. Giriş sınavları gibi önemli psikometrik işlerde bu yöntem kullanılır. Ne yazık ki, z dönüşümü birçok öğretmenin zaman ve araç bulamayacağı kadar zahmetli hesaplamalar gerektirir.

Ağırlıklı Puanlarla Mutlak Değerlendirme : Yukarıda sözü edilen yöntemler, ham puanları bir tek ağırlıklı puana dönüştürme işlemleridir. Bu bir tek puan bağıl değerlendirmede kullanılabileceği gibi, mutlak değerlendirmede de kullanılır. Tablo 16.2'de gösterilen mutlak başarı yüzdeleriyle not verilecek olsa, yukarıda sözü edilen

yazılı, sözlü ve ödev puanlarının ağırlıklı toplamında, 5 ile 6 veya 4 ile 5 gibi kesme sınırları nasıl tayin edilir? Bunun için, Tablo 16.2'deki yüzdeler bir sınavdan elde edilmesi mümkün en yüksek puanın yüzdeleri olduğundan, ağırlıklı toplam puanın mümkün olan en büyük değeri bilinmeli ve bu değer tablodaki yüzdelerle uygun olarak dilimlenmelidir. Yukarıdaki örnekte elde edilebilecek en yüksek ham puanlar, laboratuvar çalışmalarında 90, ara sınavında 150 ve son sınavda 120 olsun. Bu ölçmelere, sırasıyla, $W_1 = 10$, $W_2 = 2$ ve $W_3 = 5$ düzeltilmiş ağırlıkları verilmesi kararlaştırılsın. Bu ağırlıklarla ağırlıklı maksimum puan $90 \times 10 + 150 \times 2 + 120 \times 5 = 1800$ olur. Bu sayının yüzde kırk beşi 5 notunun alt sınırını vereceğinden, 810, beş takdir edilecek en küçük ağırlıklı toplam puandır. En yüksek ağırlıklı toplam puanın diğer yüzdeleri hesaplanarak, diğer notların da alt sınırları bulunur. Bu yöntemde ağırlıklı toplam puanın en yüksek değeri ölçüt alındığından, bir mutlak değerlendirme yapılmış olur.

Ham puanlar arası birim ve ağırlık farklarını göstermek için, gerçek olmayan bu örnekte ham puanlar kasıtlı olarak büyük sayılarla ifade edilmiş; standart kaymalar ve ağırlıklar çok farklı alınmıştır. Bundan dolayı ağırlıklı toplam puan binler basamağına çıkmıştır. Gerçek bir durumda çoğu halde bu kadar büyük sayılar elde edilmez.

Bu bölümü kapatırken şu noktaları tekrar vurgulayalım : Notlar öğrenci başarısının ölçüsü olan çeşitli puanlara dayanır. Bu ölçülere, geçerlikleri ve güvenilirlikleri dikkate alınarak ağırlık verilir. Ağırlıklı not verme ölçmelerden notlara yansıyan hataları azaltılabileceği gibi notun geçerliğini de artırır.

Not vermede kullanılacak ölçütün seçiminde, okulun düzeyine, programın hedeflerine ve notun ne gibi kararlarda kullanılacağına bakılmalıdır. Bağlı değerlendirme yönteminin de, mutlak değerlendirme yönteminin de kendilerine özgü sakıncaları vardır. Not verme yöntemi seçilirken bu sakıncalar göz önünde bulundurulmalı; yöntemde durumun gerektirdiği düzeltmeler yapılmalıdır.

Notların hatasız ve geçerli ölçüler olarak ortaya çıkmasında öğretmenin önemli sorumlulukları vardır. Notun ne gibi ölçmelere dayanacağını kararlaştırmak, her ölçmenin ağırlığını saptamak, uygun bir ölçüt ve yöntem seçmek, puanları hatasızca nota çevirmek bu sorumlulukların başlıcalarıdır. Öğretmenin bu sorumluluklarını başka kurum veya kişilere devretmek sakıncalı olur. Not vermenin ardışık karar noktalarında öğretmen yargılarına dayanılacağından, not vermede son söz yine öğretmendir.

Değerlendirme Sonuçlarının Kullanılması

Eğitim sistemlerinde öğrencilerin okula kabul edilmeleri, sömestr ve yıllara göre iletetilmeleri, program ve kollara ayrılmaları, program değiştirmeleri **öğrenci akışı** terimiyle adlandırılır.¹ Sistemde öğrenci akışı, yönetmeliklerde belirlenen kurallarla düzenlenir. Yönetmeliklerimiz incelendiğinde görülür ki, akış kararlarının dayanağı, öğrenci başarısının ve yeteneklerinin ölçüleridir. Çoğu halde, kararların biricik dayanağı başarı notlarıdır.

Akış kurallarının sistemin yapısına uymaması birçok sorun yaratır. Sınıf geçme ve mezuniyette kuralların program hedeflerinin ötesinde sıkı tutulması; okullara girişte gereksiz sınırlamalar bulunması yetenek kayıplarını ve fireleri çoğaltır. Kurallardaki boşluklar, minimum başarı sınırının program hedeflerinin altında tutulması, gereksiz hoşgörüler sistemin işleyişini verimliymiş gibi gösterir; fakat gerçekte **gizli başarısızlık** yaratır.

Öğrenci akışını düzenleyen kurallarla programların yapısı arasında da önemli ilişkiler vardır. Programların yüklü ve öğrencinin hazıroluşluk düzeyine göre zor olması, programların hedefleri ile akış kuralları arasındaki tutarsızlık, vb. çok sayıda öğrencinin sınıf veya ders tekrara, bazı öğrencilerin sistemden çıkarılmalarına yol açar.

Bir sistemde öğrenci akışı kurallarının incelenmesi, kararlarda notlar ve diğer psikometrik ölçüler kullanıldığından, ölçme ve değerlendirme alanının da konusu olur. Notların sınıf geçme ve ders geçme gibi dikey akış kararlarında kullanılması bu bölümün birinci kesiminde; kollara ve mesleklere ayırma gibi yöneltme tavsiyelerinde kullanılması ikinci kesiminde ele alınmaktadır. Notların program değerlendirme, verimliliği ölçme gibi araştırmalarda kullanılması, bu kitabın kapsamını aşan tekniklerle ilgili olduğundan ele alınma-

¹F. Oğuzkan, M.F. Turgut, S.Ç. Özoğlu, *Ortaöğretim İkinci Devresinde Ders Geçme ve Kredi Düzeni Üzerinde Bir Simülasyon Denemesi*, s. 9.

miştir. Bölüm, değerlendirme sonuçlarının kaydedilmesi ve ilgililere bildirilmesini tartışan kesimle sona erer.

DIKEY AKIŞTA NOTLAR

Öğrencilerin sistemdeki dikey hareketlerinde başlıca ölçüt derslerdeki öğrenci başarısıdır. Bir öğrenci bir sınıfta verilen derslerin hepsinden başarılı olmuş ise, hakkında «sınıf geçti» kararı verilir; o öğrenci ertesi yıl üst sınıfın derslerine devam eder. Buna karşılık, bir öğrenci o yılın derslerinin hepsinden başarısız olmuşsa, hakkında «sınıfta kaldı» kararı verilir; o öğrenci ertesi yıl aynı sınıfın derslerini tekrarlar. Bu iki halde öğrenciyi gruplama kararı tartışmasız kabul edilebilir. Fakat, bir öğrenci bir sınıfın derslerinin bazılarında başarılı, bazılarında başarısız olmuşsa nasıl gruplanacaktır? Uygulanan yöntemlerdeki farklılık, okulların bu soruya verdikleri cevapların farklılığından doğar. Öğrencilerin okulda ilerletilmesini düzenleyen yöntemler, çok çeşitli olmakla birlikte, **sınıf geçme** ve **ders geçme** terimleriyle adlandırılan iki ana türde toplanabilir.

Öğrencileri sınıf esasına göre gruplayan bir okulda, sınıftaki derslerin bir kısmından başarısız olmuş bir öğrenci sınıfta kalır, başarılı olduğu dersler de dahil olmak üzere, kaldığı sınıfın derslerin hepsini ertesi yıl tekrarlar. Bu tür gruplama yöntemlerine «sınıf geçme düzeni» denilmiştir.²

Bir sınıftaki derslerin tümündeki başarı yerine, derslerdeki başarıları ayrı ayrı esas alan gruplama yöntemleri de vardır. Bu yöntemlerin uygulandığı sistemlerde okul programı sınıf programlarından değil, yıl ve sömestrlere dağıtılmış derslerden oluşur. Öğrenci bu dersleri belirli bir sırayla alır. Bir öğrenci bir dersten başarılı olursa ertesi sömestr veya yıl onu izleyen derse devam eder. Bazı derslerden başarısız olmuş bir öğrenci ise, ertesi yıl başarısız olduğu dersleri izler; onlarla birlikte, ilerideki yılların önşartlı olmayan derslerini de alabilir. Belirli bir sürede derslerin tümünü başaran ve belirli sayıda «kredi» kazanan öğrenci mezun olur. Bu tür gruplamalara, sınıf yerine ders esasına dayandığı için, «ders geçme ve kredi düzeni» denilmiştir.³

Aşağıdaki kesimlerde sınıf geçme düzeni ile ders geçme ve kredi düzeni, okullarımızda uygulanmaları da dikkate alınarak tartışılmaktadır.

²Ibid, s. 7.

³Ibid, s. 7.

Sınıf Geçme Düzeni

Okul süresinin yıllara bölünmesinden ve okul programındaki derslerin yıllara dağıtılmasından «sınıf» kavramı doğmuştur. Okul programı açısından sınıf, belli bir öğrenim yılında verilen dersler grubunu ifade eder. Öğrencilerin gruplanması açısından sınıf, belli bir öğrenim yılındaki dersleri izleyen öğrenci grubu anlamına gelir.

Sınıf geçme düzeninin uygulandığı en katı halde, öğrenci bir tek dersten başarısız olsa bile kaldığı sınıfın bütün derslerini tekrarlar. Kuramsal olarak, böyle bir uygulamanın altında (a) sınıf programındaki bütün derslerin birbirine bağlı ve bölünmez olduğu, (b) sınıf geçmek için bu derslerin hepsinin birlikte öğrenilmesi gerektiği, (c) kişinin gelişmesinin toplu olduğu, bir alandaki eksikliğin tüm gelişmenin yetersizliğine yol açacağı sayılıları yatar. Uygulamada bu üç sayıltı ne dereceye kadar doğru olabilir?

Sınıf programının bir bütün olduğu düşüncesi doğrudur. Derslerin birbirinden ayrılamaz olduğu düşüncesi ise ancak ilkökul düzeyindeki programlar için doğrudur. İlkokul programlarındaki dersler birer akademik bütün değil, birbiriyle ilişkili öğrenme ünitelerinden oluşan bütünlere sahiptir. Daha yukarı eğitim düzeylerine çıkıldıkça, dersler akademik nitelik kazanır, birbiriyle daha az ilişkili bağımsız bütünlere haline gelir. Bir öğrencinin sınıf programındaki her dersten yeterli bir başarı düzeyine ulaşması düşüncesi doğrudur; fakat, programdaki derslerin hepsinin aynı zamanda öğrenileceği düşüncesi doğru değildir. Normal koşullarda, sınıf programındaki bütün derslerden başarılı olan öğrenci oranı, ilkökuldan yukarı düzeylere doğru çıktıkça azalır. Buna sebep, yukarı düzeylere doğru derslerin çeşitlenmesi, programlarda yetenek ve ilgiye göre farklılaşmalara izin verilmemesidir. Üçüncü sayıltı ise, ilkökul çağındaki çocuklar için doğrudur. Üst sınıflara çıkıldığında bu ilişki zayıflar; öğrenci bir yönde sıvrılmeye başlar.

Sınıf geçme düzeni ilkökullarda ve ortaokul düzeyinde bazı yumuşatmalarla uygulanabilir. Böyle olmakla birlikte, bu düzeyde sınıfta kalmaya karşı çıkan eğitimciler vardır. Goodlad ve Anderson, ilkökulda öğrencilerin sınıflara ayrılmayıp yaşlarına göre gruplanmalarını; öğretim yöntemlerinin bireyselleştirilmesiyle her öğrenciye kendi yetenekleri ölçüsünde eğitim verilmesini savunurlar ve böyle bir uygulamayı açıklarlar.⁴

⁴J.I. Goodlad ve R.H. Anderson, *Nongraded Elementary School*.

Ortaöğretimin ikinci kademesinde derslerin çeşitliliği, programların yüklü olması, öğrencilerin ilgi ve yeteneklerinde farklılaşma başlaması sınıf geçme düzeninin uygulanmasında önemli sorunlar yaratır. Aşağıda görüleceği gibi, bu düzeydeki program ve öğrencilere ders geçme ve kredi düzeni daha uygun düşer.

Sınıf geçme düzeninin katı haliyle uygulanmasında, sınıfta kalma ve bir yıl önce başarılan dersleri tekrarlama önemli ekonomik kayıplara sebep olur. Ayrıca, sınıfta kalan öğrencinin kişiliğinde ve sosyal ilişkilerinde yaratılan olumsuz etkiler de büyük sorunlara yol açar. Böyle gerekçelerden hareket edilerek, sınıf geçme düzenine yumuşatıcı yöntemler getirilmiştir. Sınıf geçme düzenini yumuşatan mekanizmaların başlıcaları şunlardır :

1. Sınıf programındaki bir veya birkaç dersten başarısız olmuş bir öğrenci, sınıf öğretmenler kurulu kararıyla o derslerden başarılı sayılabilir.

2. Bir veya birkaç dersten başarısız olmuş bir öğrenci, diğer derslerdeki genel başarı ortalaması belli bir sınırın üstünde olmak koşuluyla, o derslerden başarılı sayılabilir.

3. Öğrencinin başarısız olduğu derslerdeki notu başarı sınırına yakınsa (Örneğin, 5 ile geçilen halde 4 ise), o derslerden başarılı sayılabilir.

Bu üç yöntem bir - iki dersten başarısız olmuş bir öğrencinin başarısızlığını hoşgörür; sınıf tekrarını önler. Yöntemlerde başarısız öğrencinin eksikliklerinin tamamlanması yoktur. Bu tür yöntemlere «af ve müsamaha mekanizmaları» denilmiştir.⁶

Sınıf geçme düzenini yumuşatma amacıyla, okullarımızda, «bütünleme», «tamamlama», «borçlu veya sorumlu sınıf geçme», «tek ders sınavı» adlarıyla bilinen yöntemler de uygulanmıştır.⁶ Bunlardan bütünleme, öğrenciye başarısız olduğu derste eksikliklerini gidermesi için süre tanıyan ve o sürenin sonunda uygulanan bütünleme sınavıyla «geçti - kaldı» kararı veren yöntemdir. Tamamlamada ise aynı işlem okulda ek takvim süresince verilen derslerle yapılmıştır. Her iki yöntemin başarısı, tamamlayıcı ve düzeltici öğretim etkinliklerinin iyi düzenlenmesine ve dikkatle uygulanmasına bağlıdır. Geçmişte bu yöntemler iyi kullanılamamıştır.^{7,8}

⁶T. Oğuzkan, M.F. Turgut, S.Ç. Özoğlu, op. cit. 88 - 92.

⁶Millî Eğitim Bakanlığı, Orta Öğretimle İlgili Yönetmelikler, 1976, s. 89 - 122.

⁷M.F. Turgut, «Eğitim Psikolojisi Açısından Bütünleme ve Borçlu Sınıf Geçme», Eğitim Dünyası, 8 : 4 - 9.

⁸T. Oğuzkan (Ed.), Türkiye'de Ortaöğretim ve Sorunları, s. 152 - 156.

Borçlu veya sorumlu olarak sınıf geçme yönteminde öğrencinin üst sınıfa geçmesine izin verilir; sorumlu olduğu derste eksikliklerini bu sınıfta tamamlaması istenir. Sorumluluk üst sınıfta ilgili öğretmenin uygulayacağı bir sınavla kalkar. Bu yöntem, kanımızca, üst sınıfta eksikliği tamamlatacak öğretim imkânları bulunmadığı için iyi işlemediştir.

Okullarda uygulanmakta olan tek ders sınavı, başarısızlığı hoşgörme veya tamamlatma yöntemi değil, sınavın tekrarı yöntemidir. Onun için, ancak, önceki sınıfların güvenilirliğini artırıcı bir etkisi olabilir. Böyle düşünülürse, sınavın yalnız tek dersten başarısız olan öğrencilere değil, başarısız öğrencilerin hepsine verilmesinin daha yararlı olacağı kanısına varılır.

Sınıf geçme düzeninin en katı uygulaması olan «bütün derslerden başarılı olma» şartıyla sınıf geçmede karar, tek tek her dersten ve o dersin öğretmeni tarafından verilir. Sınıf geçme kararı sınıf programının bütünlüğüne ve öğrencinin genel başarı düzeyine bakılmadan verildiği için eleştirilir. Hatalı sınıfta kalmaları önlemek amacıyla, öğrencinin sınıfı geçip geçemeyeceğinin sınıf programının bütünlüğü içinde değerlendirilmesi, bu konudaki kararın «sınıf öğretmenler kurulu» tarafından verilmesi düşünülmüştür. Sınıf öğretmenler kurulu, bir yöneticinin başkanlığında, o sınıfa ders veren öğretmenlerden oluşur. Kurulda öğrencilerin başarı durumları birer birer tartışılır; gerekirse, öğretmenin verdiği not değiştirilir. Kuramsal olarak her iki yönde de not değiştirme mümkün görünmekle birlikte, uygulamada, bir veya birkaç dersten başarısız öğrenciler hakkında «geçer» kararı alınmış; başarılı bir öğrencinin başarısız sayıldığı görülmemiştir.

Sınıf öğretmenler kurulu uygulamaları (1) kurula öğrenci başarısı hakkında yeterli bilgi getirilmediğinden kararların isabetsiz olduğu, (2) objektif ölçütler kullanılmadığından öğrenciden öğrenciye çok farklı kararlar verildiği için eleştirilir. Bu eleştirilerde gerçek payı vardır. Fakat, öğrenci başarısının yeterince ayrıntılı olarak ölçülmesi ve kurul kararlarında kullanılmak üzere ölçütler saptanması halinde, sınıf öğretmenler kurulu daha az sakıncayla işletilebilir. Sınıf öğretmenler kurulu kararlarında şu ölçütler kullanılabilir : Öğrencinin a) başarılı olduğu derslerdeki genel başarı ortalamasına, b) fen dersleri, dil ve edebiyat dersleri gibi ders gruplarındaki başarı ortalamalarına, c) bir üst sınıfta izleyeceği programa, d) başarısız olduğu dersin niteliğine, e) başarısızlık notunun derecesine bakılarak karar verilebilir. Şüphesiz bazı derslerden başarısız olmuş bir

öğrenciyi başarılı sayarak sınıf geçirmede en önemli ölçüt, onun üst sınıf derslerini başarıp başaramayacağıdır.

Bu kesimdeki tartışmalardan şu sonuca varılır : Sınıf geçme düzeni programların yapısına ve öğrencilerin eğitim düzeylerine bakılmaksızın, katı haliyle, hiçbir okulda uygulanmamalıdır. Temeleği-tim okullarında sınıf geçme düzeni, sınıf öğretmenler kurulu gibi bir otorite tarafından ve mümkün olduğu kadar objektif ölçütlerle uygulanabilir. Ortaöğretimin ikinci kademesinde ve yükseköğretimde, programların yapısına ve öğrencilerin yetişkinlik düzeyine daha uygun düşen ders geçme düzeni uygulanmalıdır.

Ders Geçme Düzeni

Ders geçme düzeninin işletilebilmesi okul sisteminin yapısına, okulun düzeyine, programın yapısına, okul yöneticilerinin ve öğretmenlerin bu düzeni anlayışlarına bağlıdır. Programındaki derslere bağımsızlık tanıyan, programlarında tercihler ve seçimlik dersler bulunan, öğrencinin programına bireysellik kazandıran okullar ders geçme düzeni uygularlar.

Okul programları ders geçme düzeninin uygulanışını kolaylaştıracak biçimde düzenlenebilir. Bunun için : (a) Bir ders zorunluluk olmadıkça yıllara bölünmemeli, bir sömestir veya yılda yoğun olarak verilmelidir. Böyle bir düzenleme önşartlı ders sayısını ve önşartlı derslerden dolayı bekleyen öğrenci sayısını azaltır. (b) Seçimlik derslerin sayısı artırılmalıdır. Bu düzeylemeyle, seçimlik bir dersten başarısız olmuş bir öğrencinin, o dersin tekrar verileceği sömestri beklemekten, onun dengi olan diğer bir seçimlik dersi alması sağlanır. (c) Bir sömestirde alınması gerekli ders sayısı ve öğrencinin haftalık ders yükü belirli sınırları aşmamalıdır. Öğrencinin haftalık programında serbest zaman bulunması, bireysel programların düzenlenmesini kolaylaştırır.

Ders geçme ve kredi düzeni yükseköğretim kurumlarımızın bazılarında uygulanmaktadır. Ortaöğretimin ikinci kademesindeki bazı okullarımızda bu düzeni uygulama gayretleri olmuştur. Bu konuda birkaç araştırma da yapılmıştır. Oğuzkan, Turgut ve Özoğlu bu modelin esaslarını analiz etmişler ve uygulanabilirliğini «simülasyon» yöntemiyle yoklamışlardır.⁹ Bu araştırma, ders geçme ve kredi düzeninin, ortaöğretimin ikinci devresindeki okullarda bazı koşullarla

⁹T. Oğuzkan, M.F. Turgut, S.Ç. Özoğlu, *op. cit.*

uygulanabileceğini göstermiştir. Millî Eğitim Bakanlığı bir taslak yönetmelik hazırlayarak, modeli on altı okulda denemeye koymuştur. Yapılan değerlendirmelere göre, deneme beklendiği kadar başarılı olmamıştır.¹⁰ Başarısızlık daha ziyade öğretmen ve yöneticilerin modeli tutmamalarına bağlanmıştır. Programların yapısı da ders geçme ve kredi düzeninin uygulanmasını güçleştirmiş olabilir. Diğer bir araştırmada ders geçme ve kredi düzeninin başarıyla uygulan-
dığı bildirilmiştir.¹¹

Ders geçme ve kredi düzeninin de bazı sakıncaları vardır. Örneğin, öğrencilerin programlarında bütünlüğün kaybedilmesi, derslerin aşamalılığının ve öğrenimin sürekliliğinin bozulmasına sebep olabilir. Modelin en liberal haliyle uygulanması, öğrencinin öğrendiklerini unutmasına ve öğrenim kalitesinin düşmesine yol açabilir. Ders geçme ve kredi düzeninin sakıncalarını azaltmak amacıyla, modelde bir takım sınırlamalar getiren yöntemlere başvurulmuştur. Bunlardan biri, öğrencinin başarısız olduğu derslerin kredileri toplamı belli bir sınırı aştığında, üst düzeyden ders alınmasına izin vermeyip sadece başarısız olunan dersleri tekrarlatan yöntemdir. Diğer bir yöntemde, bir öğrencinin başarılı olduğu derslerdeki not ortalaması belli bir sınırın altına düşünce, üst düzeyden alınabilecek ders sayısını sınırlanır. Başka bir sınırlama, belli bir öğrenim kademesine (örneğin, okulun son yılına), ancak aşağı düzeydeki derslerin hepsini başardıktan sonra girilmesine izin veren yöntemde görülür.

GRUPLAMA VE YÖNELTMEDE NOTLAR

Eğitim sistemlerinde yatay düzenleme işlemlerinden ikisi, öğrencilerin aynı programda gruplanmaları ve farklı programlara yö-neltilmeleridir. Bu kesimde notların gruplama ve yöneltme kararlarında kullanılması tartışılmaktadır.

Aynı Programda Gruplama

Aynı programda aynı dersleri izleyecek olan öğrenci grubu bir sınıftan daha kalabalık olursa, grup hangi ölçütlerle bölünmelidir? Bu soru Batı ülkelerinde çeşitli tartışmalara ve araştırmalara konu

¹⁰Y. Baykul, «Ders Geçme ve Kredi Düzeninin Değerlendirilmesi», *Mesleki ve Teknik Öğretim Dergisi*, 320 : 26 - 37.

¹¹J. Akhun ve H. Doğan, *Okul - Sanayi Ortaklaşa (OSANOR) Eğitim Projesi İkinci Dönem Raporu*.

olmuştur.¹² Batı ülkelerinde, okulun bulunduğu düzeye bağlı olarak, yaşa, cinsiyete, alt sınıflardaki başarıya, yeteneğe, öğrenciler arasındaki sosyal ilişkilere göre gruplamalar savunulmuş ve uygulanmıştır. Ülkemizde öğrencileri şubelere ayırmada kullanılan belli ölçütler yoktur. Uygulama ya okul yönetiminin anlayışına bırakılmış, ya da öğrencilerin okula kaydoluş sırası gibi tesadüflere kalmıştır.

Öğrencilerin başarı ve yetenek bakımından nisbeten homojen gruplara ayrılmaları, daha ziyade, öğretimin grubun genel yetenek düzeyine göre ayarlanmasını isteyenler tarafından savunulur. Heterojen grupları savunanlar; gruplamada öğrenciler arası sosyal etkileşimleri ve yeteneğe dayanmayan başka ölçütleri kullanırlar; sınıftaki yetenek farklılıklarının da öğretimde bireysel yöntemler kullanılarak karşılanmasını isterler. Aslında, gruplamada geçerli olabilecek ölçütler, öğrencilerin birlikte gelişmelerini sağlayacak, toplu öğretim yöntemlerinin verimini artıracak olan ölçütlerdir. Bu nedenle, başarıya ve genel yeteneğe dayanan gruplamalar tercih edilmelidir. İlkokul düzeyinde yaşa, çocuğun zihin gelişimine ve zekâ düzeyine dayanan gruplamalar önerilebilir. Temel eğitimin ikinci kademesinde ve ortaöğretimde yeteneğe ve önceki yıllarda ulaşılan başarı düzeyine göre gruplamalar okulun verimini artırır.

Yöneltme

Okullarda öğrencilerin kol, program veya meslek seçmelerine yardımcı olmak amacıyla yapılan etkinliklere yöneltme denir. Yöneltmede, diğer verilerle birlikte, öğrencinin önceki dönemlerde aldığı başarı notlarından da yararlanılır.

Yöneltmede notların kullanıldığı yerlerden biri, öğrencinin üst düzeydeki bir öğrenimi başarabilmesi olasılığının kestirilmesidir. Öğrencinin alt düzeydeki öğretim basamaklarında aldığı notlara bakılarak, üst düzeyde ne derece başarılı olabileceği kestirilmeye çalışılır. Genel olarak yordama denilen bu kestirme işleminde, önceki başarı ölçülerine «yordayıcı», üst düzeyde elde edilecek olan başarı ölçüsüne «yordanan» değişken denir. Bu iki değişken arasındaki korelasyon katsayısı yordamanın geçerliğini gösterir.

Üst öğrenimdeki başarının yordanmasında mezuniyet derecesi, sınıf geçme derecesi, notların basit ortalaması, ağırlıklı not or-

¹²F. Heathers, «Grouping,» in R.L. Ebel (Ed.), *Encyclopedia of Educational Research*, s. 559 - 570.

talaması, yığılmalı ortalama gibi ölçüler yordayıcı olarak kullanılır. Yordanan değişkenin ölçüsü de üst düzeydeki sınıf geçme notları, bunların ortalaması veya mezuniyet derecesi olur.

Notlarla yapılan yordamalar, hem notlar yeterince güvenilir olmadığı için, hem de alt ve üst düzeylerin ders programları arasındaki ilişki yüksek olmadığından yeterince geçerli olmaz. Yordayıcı olarak yığılmalı başarı ortalaması, derslerin saatlarına veya kredilerine göre ağırlık verilen ağırlıklı not ortalaması kullanılırsa, yordamanın geçerliliği artırılabilir. Genel ortalama yerine, yalnızca üst programın temelini oluşturan derslerdeki ağırlıklı ortalama kullanılması da yordamanın geçerliliğini artırabilir. Okuldaki başarıyı yordamada genel yetenek testleri de geçerlidir. Notların test puanlarıyla birlikte kullanılması da yordamanın geçerliliğini artırır.

Yöneltilmede notların diğer bir kullanıma yeri, öğrencilerin fen - edebiyat gibi program kollarına veya elektrik - muhasebe - ev ekonomisi gibi meslek programlarına ayrılmasıdır. Bu tür ayırma işlemleri de genelde bir yordama olmakla birlikte, genel başarının yordanmasından çok farklıdır. Fen - edebiyat ayırımını örnek alalım. Bu ayırım yapılırken cevaplandırılacak ana soru, öğrencinin hangi programda daha başarılı olacağı sorusudur. Öğrencinin önceki dönemlerde kazandığı notların genel ortalaması bu ayırımı yapamayabilir. Fakat, önceki dönemlerde fen grubu derslerinde daha başarılı olmuş bir öğrencinin, fen kolunda daha başarılı olması beklenir. Bu nedenle, bir tek genel ortalama değil, fen grubu ve sosyal - dil grubu gibi iki ayrı ders grubunda notların grup ortalamaları hesaplanır. Fen grubu ortalaması diğer grup ortalamasından yüksek olan öğrencilerin üst düzeydeki fen kolunda daha başarılı oldukları araştırmalarla saptanırsa, iki ortalama arasındaki farkın ayırıcı gücü kanıtlanmış olur. İsabetli bir ayırım için yordayıcıların güvenilir olması yetmez; iki yordayıcı arasındaki puan farklarının da yeterince büyük olması gerekir. Ne var ki, çeşitli not ortalamaları ancak nadir hallerde böyle ayırıcı farklar gösterir. Aşağı eğitim basamaklarında ders notları arasında gerçekten ayırıcı farklar bulunmaz. Çoğu halde, bir grup dersten yüksek başarı göstermiş öğrenciler diğer derslerden de yüksek başarı gösterirler. Yüksek eğitim basamaklarında da ancak kapsamları çok farklı derslerin notları arasında ayırıcı farklar bulunabilir.

Ülkemizde notlarla bu tür ayırıcılık araştırmaları hiç yapılmamıştır. Ayırıcılıkları kanıtlanmış testlerimiz de yoktur. Böyle olmasına rağmen, notları ayırıcıymış gibi kullanılmaktadır. Örneğin, Fen

Lisesi sınavına girmiş, ortaokul fen grubu derslerindeki başarı ortalamasının belli bir değerden yüksek olması koşulu aranır. Liselerimizdeki kol ayırımında grup dersleri ortalamalarına bakılır. Meslek liselerindeki meslek programlarına yöneltmede de, bazı derslerde kazanılan notlar dikkate alınır.

Batı ülkelerinde, mesleklerdeki ve meslek programlarındaki başarıyı ayırıcı olarak yordayan testler geliştirilmiştir. Notların bu tür testlerle birlikte kullanılması, daha isabetli yöneltme kararları alınmasına yardım eder. Birbirine yakın programlar için ne testler, ne de notlar ayırıcıdır. Bu durumlarda yöneltme kararı, başarı ve yetenek ölçülerine göre değil, kişinin ilgi ve isteğine göre verilir.

DEĞERLENDİRME SONUÇLARININ BİLDİRİLMESİ

Bir dersin öğretmeni tarafından verilen notlar, o öğretmenden başka kişilerce de kullanılır. Bu nedenle, öğrenci başarısının ölçüsü olan notların ve diğer psikometrik ölçümlerin hatasızca elde edilmesini sağlamak, güvenli kayıtlarını tutmak, gerektiğinde kullanmak ve ilgililere bildirmek okulun görevidir. Bu kesimde değerlendirme kayıtlarının tutulması, değerlendirme sonuçlarının öğrenciye ve velilere bildirilmesi üzerinde durulmaktadır.

Değerlendirme Kayıtları

Öğrenci başarısının resmî kayıtları öğretmenin öğrenciye not vermesiyle başlar. Okullarımızda başarı notları sömestr sonlarında öğretmenler tarafından okul yönetimine bildirilir. Bitirme ve bütünleme sınavı gibi sınavların sonuçları da yönetime bildirilir. Bu amaçlarla kullanılan not bildirme formu, öğrencilerin adlarını ve kazandıkları notları listelleyen bir cetveldir. Okulda notların yazıldığı, öğrencinin okul numarası sırasıyla düzenlenmiş bir defter de tutulur. Sınıf geçme kararlarının da geçirildiği bu defter, öğrencinin okuldaki başarısının ana kütüğünü oluşturur. Notlar, ayrıca, «karne» adı verilen belgeye yazılarak öğrenciye ve velilere bildirilir.

Thorndike ve Hagen okulda tutulacak psikometrik kayıtların (a) sürekli ve yığılmalı, (b) geniş kapsamlı, (c) öğrencinin yeteneklerini tanımaya elverişli, (d) gerektiğinde kolayca kullanılabilir niteliklerde olmasının önemini vurgularlar.¹³ Okullardaki akademik kayıt-

¹³R.L. Thorndike ve E. Hagen, *Measurement and Evaluation in Psychology and Education*, 2nd. ed., s. 503.

ların, sistemin tümünde standartlaştırılmasının çeşitli yararları vardır. Akademik kayıtların doğru olması ve güvenle saklanabilmesi de önem taşır.

Öğrencilere Bildirme

Değerlendirme sonuçlarının en yakın ilgilisi öğrencidir. Bu nedenle, öğrencinin başarısı veya başarısızlığı, başarısını nasıl yükseltebileceği kendisine sık sık bildirilir. Öğrenci başarısız ise eksiklerinin ona gösterilmesi ve başarısını nasıl yükseltebileceğinin anlatılması yararlı olur. Öğrenci başarılı ise, başarı durumunu bildirmek başarının devamı için güdüleyici olabileceği gibi başarı düzeyinin yükseltilmesine de yardım edebilir. Başarı durumunun öğrenciyé açıklanması dersin öğretmenine düşen bir görevdir. Başarı durumunu öğrenciyé bildirmenin etkili olabilmesi için sınav, ödev, v.b. gibi ölçme işlemlerinin hemen arkasından gelmesi gerekir.

Başarı durumunu öğrenciyé bildirme aracı olarak karne yetersizdir. Karnenin yetersizlikleri notların ayrıntılı olmayışından, başarısızlığın nasıl telafi edileceğini göstermeyişiinden ve çok seyrek kullanılmasından gelir. Öğretmen, öğrencilere durumlarını bildirme işlemlerine sınıfta toplu açıklamalarla ve öğretmen - öğrenci mülakatlarıyla devamlılık getirirse karnenin yetersizlikleri önemsiz olur.

Öğretmen - öğrenci mülakatlarında öğrencilere başarı durumları hakkında da bilgi verilir. Başarısız öğrenciler, çoğu halde başarısızlıklarını gerçekçi biçimde algılamazlar. Öğrenciler, öğretmenin kıt not verdiği, sınavların zor olduğu, başarılı öğrencilerin kopya yaptığı, öğretmenin gazezli davrandığı gibi savunma mekanizmaları geliştirirler. Notların objektif ölçülerle verilmesi, not verme yöntemlerinin sınıfta tartışılması, öğretmen - öğrenci mülakatlarında öğretmenin «yardımcı» rolünde bulunması, açığa vurulmuş savunma mekanizmaları varsa onların gerçek olmadığını göstermesi bu tür öğrenci davranışlarını önlemede yararlı olur. Başarısızlığın öğrenci üzerindeki olumsuz etkileri, öğrencilerin kendi başarılarını objektif olarak değerlendirmeleriyle azalır. Bu konuda en etkili yardımcı, dersin öğretmenidir. Şüphesiz, rehberlik ve danışma tekniklerinde yetişkin bir öğretmen bu konuda daha başarılı olur.

Velilere Bildirme

Öğrencilerin başarılarıyla yakından ilgilenen diğer bir grup öğrenci velileridir. Onlar da, okul yönetimi ve öğretmenler gibi, öğ-

renciye yardım etme ve onun geleceęi hakkında önemli kararlar alma zorunda kalırlar.

Karne velilere de yeterli bilgi vermez. Strang, karnenin yeter-sizliklerini tamamlamak üzere, velilere řu noktalarda ek bilgi veril-mesini önerir:¹⁴ 1) Öğrencinin notu aynı sınıf düzeyindeki öğren-cilerin notlarıyla karşılaştırılmalı; onun bu öğrenci grubu içindeki sırası belirlenmelidir. 2) Öğrencinin başarısı, akademik yeteneęi ölçüt alınarak değerlendirilmeli; onun, yetenekleri ölçüsünde başa-rılı olup olmadığı belirlenmelidir.

Öğrencinin evi ile okulu arasında etkili bir iletişim sağlamanın en iyi yolu, öğretmenlerle velileri sık sık bir araya getirmektir. Öğ-retmen - veli mülakatlarında her öğrencinin okuldaki başarı duru-mu ayrıntılarıyla açıklanabilir; velilerin çocuklarına ne yolda yar-dım edebilecekleri anlatılabilir. Bu tür bir karşılıklı konuşma, mek-tup gibi tek yönlü iletişim araçlarının eksikliklerini tamamlar ve iyi yürütülebilirse çeşitli yararlar sağlar. Fakat, bu yöntemin işletilme-sinin çeşitli engelleri vardır. İlk engel öğretmenlerin ve özellikle ve-lilerin buluşmalara zaman ayıramamalarından doğar. İkinci engel bazı velilerin okula ve öğrenciye gerekli ilgiyi göstermemelerinden, hatta okula karşı tutumları bulunmasından gelir.

Öğretmen - veli mülakatlarının daha önemli bir sakıncası öğ-retmenin bu konuşmaları amacına ulaştırarak becerilere sahip ol-maması, velilere ve öğrenciye ilişkin tutumlarının olumsuz olması halinde ortaya çıkar. Öğretmen - veli mülakatı da bir danışma işi-dir; danışma ilkelerine uyularak yapılması gerekir. Öğretmenin reh-berlik ve danışma ilkelerini hiç değilse ana çizgileriyle bilmesi, bu konuda başarı sağlamasına yardım eder.

¹⁴ Strang, Reporting to Parents, s. 38.

Madde Analiziyle Test Geliştirme

Bir öğrenci grubunun, bir testin maddelerine verdikleri cevapların analizinden, gerek testin tümü, gerekse tek tek maddeleri hakkında önemli bilgiler elde edilir. Bu bilgilere dayanılarak, testin istatistiksel özellikleri kestirilebilir; kusursuz ve istenilen nitelikteki maddeler bir araya getirilerek, istenilen özelliklerde yeni bir test düzenlenebilir; maddelerin kusurları bulunup düzeltilebilir.

Madde analizi, çeşitli yöntemlerle yapılabilir. Bu yöntemlerin bazıları büyük örneklemelere ve elektronik makinalarla işlemlere göre hazırlanmıştır; bazıları ise fazla zahmet ve makina gerektirmeyecek biçimde basitleştirilmiştir. Bu bölümde, bir öğretmenin kendi imkânlarıyla uygulayabileceği kadar kolay iki yöntem verilmektedir. Ayrıca, madde istatistiklerine dayanılarak, istenilen özelliklerde bir testin nasıl düzenleneceği ve maddelerin nasıl düzeltileceği kısaca belirtilmektedir.

ÜST VE ALT GRUP PUANLARIYLA MADDE ANALİZİ

Bir testin maddelerinin ne derecede iyi işlediğini anlamak için, maddenin «yeteneği yüksek» kişileri «yeteneği düşük» kişilerden ne kadar ayırt edebildiğine bakmak, madde analizinin dayandığı temel düşüncelerden biridir. (Burada «yetenek» sözcüğü, «testle ölçülmek istenilen değişken» anlamında kullanılmıştır.) O testle ölçülmek istenen yetenek bakımından yeterince farklı iki grup bulunabilirse, test bu gruplara uygulanır ve her maddeye iki grubun verdiği cevaplar karşılaştırılır. Yeteneği yüksek olan grubun daha büyük bir oranda doğru cevaplandığı bir maddenin «ayırıcı» ve bu sebeple geçerli olduğu kanısına varılır.

Sınıfta kullanılan başarı testleri için, test dışı ölçütlere göre başarıları farklı gruplar bulmak çok güçtür. Onun yerine, testin toplam puanlarına göre, tüm gruptan bir «üst grup» ve bir «alt grup» tanımlanır. Bir test maddesi gerçekten testin toplam puanıyla ölçü-

len değişkeni ölçebiliyorsa, o maddenin, üst gruptaki kişiler tarafından daha büyük bir oranda doğru cevaplandırılması beklenir. Onun için, üst grupta daha büyük bir oranda doğru cevaplandırılan bir maddenin, ayırıcı ve bu sebeple geçerli olduğu kanısına varılır.

Aşağıda ana çizgileriyle verilen madde analizi yöntemleri, testin toplam puanıyla tanımlanan alt ve üst gruplarla uygulanır. Karşılaştırma grupları testin kendi puanlarıyla seçildiğinden, bu yöntemlere, «iç ölçütle madde analizi» yöntemleri de denir.

Madde analizi şu üç amaçtan birini veya birkaçını gerçekleştirmek için yapılır :

1. Denenen birçok madde arasından, istenilen nitelikteki maddeleri seçmek . Amaç bu olunca, kusursuz ve istenilen, nitelikteki maddeler seçilir; kusurlu maddeler üzerinde hiçbir işlem yapılmaz. Bu tür bir analizde, maddelerin güçlük ve ayırıcılık derecelerini gösteren bazı istatistikler hesaplanır; başka bir işlem yapılmayacağından analiz basitleşir.

2. Madde analizi sonuçlarına bakarak, gerek denenen testin, gerekse yeniden düzenlenecek testin istatistiksel özelliklerini kestirmek Bu tür bir analiz denenmiş olan testin özelliklerini kestirmek amacıyla yapılırsa basitleşir; fakat, çok sayıda denenmiş maddeden belirli niteliklerde ve daha az sayıda maddeden oluşan bir test düzenlenecekse, oldukça teknik madde seçimi işlemleri gerekir.

3. Maddeler üzerinde düzeltme yapmak . Amaç bu olunca, maddelerin bütün seçenekleri analiz edilir; güçlüğü istenilen derecede olmayan, yeterince ayırıcı olmayan, çeldiricileri işlemeyen maddelerin kusurlarının bulunmasına çalışılır.

Aşağıda verilen yöntemlerin ilkinde bu amaçların sadece birincisi, ikincisinde ise her üçü gerçekleştirilebilir.

Basit Bir Analiz

Ana basamakları aşağıda gösterilen yöntem, maddelerin güçlüğü ve ayırıcılığı hakkında kaba bilgiler verecektir. Cevap kâğıtlarının tümü yerine yarısı analiz edildiğinden ve madde seçenekleri analize dahil edilmediğinden, bu yöntemde tablolama ve hesaplama işlemleri oldukça kolaydır. Bu yöntem şu basamaklarla uygulanır :

1. Cevap kâğıtlarının tümü puanlanır. Puanlamada her maddeye verilen cevabın «yanlış» veya «doğru» olduğu cevap kâğıdı üzerine işaretlenir. Bir renkli kalemle «yanlış» cevapları çizmek, «doğ-

ru» cevaplarla «cevapsız» maddeleri işaretsiz bırakmak, bu amaç için yeterli olabilir. Puanlama sonucunda elde edilen ham puan, cevap kâğıdının uygun bir köşesine yazılır.

2. Cevap kâğıtları, en yüksek ham puandan en düşük ham puana doğru, puan sırasına konur. Bu sıralamadan sonra, cevap kâğıtlarının en yüksek puanlı kâğıttan itibaren yüzde yirmi yedisi ayrılır. Bu kâğıtlar «üst grup» kâğıtlarıdır. En düşük kâğıttan itibaren de, aynı sayıda cevap kâğıdı ayrılır. Bunlar da «alt grup» kâğıtlarıdır. Grupta N tane puanlanmış cevap kâğıdı varsa, üst grup ve alt grup kâğıtlarının sayısı, en yakın tam sayı olarak, $N' = 27N/100$ işlemiyle bulunur.

3. Bundan sonra, üst ve alt grup kâğıtlarında ayrı ayrı olmak üzere, her madde için, «doğru», «yanlış» ve «cevapsız - boş» frekansları sayılarak şöyle bir hesaplama tablosuna geçirilir :

Madde No. 1	Doğru	Yanlış	Boş	N'	
Üst	40	10	0	50	p = 0,70
Alt	30	20	0	50	d = 0,20

Bu amaçla, bir tablo yerine, her madde için ayrı bir fiş kullanılabilir.

4. Frekansların tablolanması, her maddede kontrol edilir. Doğru, yanlış ve boş kategorilerindeki frekanslar toplamı, cevap kâğıtları toplam sayısının yüzde yirmi yedisi olan N' sayısına eşit olmalıdır.

5. Tablolama işlemleri testteki bütün maddeler için tamamlandıktan sonra, her madde için şu istatistikler hesaplanır.

a) **Maddenin tahmini güçlüğü** : Cevap kâğıtlarının bir kısmı analize dahil edilmediği için, madde güçlüğü, ancak üst ve alt gruplardaki doğru cevap oranlarının ortalaması alınarak kestirilebilir. Aşağıda genel olarak madde güçlüğü formülü ve 1 numaralı maddenin güçlüğü gösterilmiştir.

$$(18.1) \quad p = \frac{D_u + D_a}{2 N'}$$

p : Maddenin güçlük indisi
 D_u : Üst grupta doğru cevap sayısı
 D_a : Alt grupta doğru cevap sayısı
 N' : Tüm grubun % 27'si

$$p_1 = (40 + 30)/2 \times 50 = 0,70.$$

b) **Maddenin ayırıcılık indisi** : Maddeyi doğru cevaplandıranların sayısının üst grupta daha büyük olması beklenir. Maddeyi üst

grupta doğru cevaplandıranların sayısı ile alt grupta doğru cevaplandıranların sayısı arasındaki fark ne kadar büyük ise, madde o kadar ayırıcıdır. Aşağıda genel olarak madde ayırıcılık indisi formülü ile 1 numaralı maddenin ayırıcılık indisi gösterilmiştir.

$$(18.2) \quad d = (D_a - D_s)/N'. \quad d_1 = (40 - 30)/50 = 0,20$$

Bu basamaklarda görüldüğü gibi, bu analizin gerek tablolama gerekse hesaplama işlemleri çok basittir. Bu analiz, basit olmasına rağmen, oldukça yararlı sonuçlar verir. Böyle bir analizin sonuçlarından nasıl yararlanılacağı ilerideki kesimlerde tartışılacaktır.

Maddenin Seçenekleriyle Birlikte Analiz

Yukarıda ana basamakları verilen analiz, maddenin sadece «doğru», «yanlış» ve «cevapsız» frekansları sayıldığından, seçeneklerin analizini gerektiren durumlarda uygulanmaz. Özellikle çok seçenekli test maddelerine uygulanan bir analiz yöntemi, ana basamaklarıyla aşağıda verilmektedir.

1. Bu yöntem de, testin ham puanlarıyla tanımlanan üst ve alt grupların her maddeye verdikleri cevapların karşılaştırılmasına dayanır. Onun için, cevap kâğıtlarının puanlanması, ham puan sırasına konulması ve yüzde yirmi yedi üst ve alt grupların seçimi, yukarıdaki basit analiz tekniğindeki gibi yapılır.

2. Cevapların tablolanmasında, sadece doğru ve yanlış cevaplar değil, maddenin bütün seçeneklerinin frekansları sayılır. Beş seçenekli maddelerde bu sayımın sonuçları aşağıdaki gibi bir çalışma tablosuna veya fişe geçilir.

Madde No. 2	A	B	C*	D	E	Boş	N'	%	
Üst	4	3	24	6	2	1	40	60	p = 0,50
Alt	6	4	16	8	4	2	40	40	r = 0,20

Bu tablolama için her maddenin doğru cevabının veya yanlışların cevap kâğıtları üzerinde işaretlenmiş olması gerekmez.

3. Her maddenin doğru cevabı testin doğru cevap anahtarından alınıp tablodaki seçenekler üzerine işaretlenir. Örnekteki 2 numaralı maddenin doğru cevabı C olduğundan, tabloda bu seçenek bir yıldız (*) ile işaretlenmiştir. Anahtar hatalarının meydana çıkarılabilmesi ve madde hakkında doğru yorum için, doğru cevapların, puanlamada kullanılan anahtardan aynen alınması gerekir.

4. Üst ve alt grupta maddeyi doğru cevaplandırma yüzdeleri ayrı ayrı hesaplanır. Örnekteki maddede bu yüzdeler, üst grup için $p_a = D_a/N' = 24/40 = 0,60$ ve alt grup için $p_s = D_s/N = 16/40 = 0,40$ bulunur.

5. Maddenin güçlük derecesi, doğru cevap yüzdelerinin ortalaması alınarak veya buna denk olan $p = (D_a + D_s)/2N'$ formülüyle kestirilir. Üst ve alt gruplardaki doğru cevap yüzdeleriyle girilen ve madde ayırıcılık indisiyle birlikte madde güçlüğü de veren madde analizi tabloları vardır.¹

6. Maddenin ayırıcılık indisi olarak, Ek. 1'deki tablolardan, çift serili korelasyon bulunur.² Çift serili korelasyon, esasında, madde puanları ile testin toplam puanları arasındaki momentler çarpımı korelasyon katsayısıdır. Ek. 1'deki çift serili korelasyon tablosuna alt ve üst gruplardaki doğru cevap yüzdeleriyle girilir; tablonun verdiği sayı 100 ile bölünerek korelasyon bulunur. Örnekte üst grupta doğru cevap yüzdesi 60, alt grupta 40 olduğu için Ek. 1'deki tablolarda 40 sütunuyla 60 satırından 20 değeri okunur. O halde, bu maddenin ayırıcılık indisini gösteren çift serili korelasyon $r = 0,20$ değerindedir. Ekteki tablolar üst grupta 20 ile 98 arasındaki, alt grupta 01 ile 88 arasındaki çift değerleri kapsar. Bu sınırlar arasındaki tek değerler için, tablonun açıklanmasında belirtilen yöntemlerle ulama (enterpolasyon) yapılabilir. Alt grupta doğru cevap yüzdesinin üst gruptakinden büyük olduğu hallerde bu korelasyon negatiftir. Negatif değerlerin nasıl bulunacağı da tablonun açıklamalarında gösterilmiştir.

İç ölçütlerle analiz tekniğinden ve yüzde yirmi yedi üst ve alt grupların kullanılmasından doğan bazı güçlüklerle aşağıdaki paragraflarda kısaca değinilmektedir.

Böyle bir analiz için test kaç kişiye uygulanmalıdır? Madde istatistiklerinin yeterince güvenilir olarak elde edilebilmesi için, optimum cevaplayıcı sayısının $N = 370$ olduğu söylenir. Bu halde, alt ve üst gruplardaki kâğıt sayısı tam 100 olduğundan, doğru cevap yüzdelerinin ayrıca hesaplanması gerekmez. Üzerinde tablolama işlemi yapılacak kâğıt sayısı da 200 olur. Çok sayıda kişiye uygulanan testlerde 370 cevap kâğıdı, uygun bir yöntemle tesadüfi olarak seçilir. Öğretmenin sınıfta uygulayacağı testlerde bu kadar büyük sayıda öğrenci hemen hiç bulunmaz. Sınıf mevcutları genellikle 50

¹C - T. Fan, *Item Analysis Table*.

²H. Arıcı, *İstatistik*, s. 127 - 130.

civarında olduğundan, alt ve üst gruplardaki cevap kâğıtları sayısı 15'ten fazla olmaz. Bu da, özellikle seçeneklerin analizi için yetersiz bir sayıdır; çünkü, madde seçeneklerinin frekansları 1 - 3 gibi çok küçük sayılara iner. Bir sınıfın birkaç şubesi birleştirilebilirse, örneklem büyütülebilir. $N = 185$ olduğu takdirde, alt ve üst gruplar ellişer kâğıda çıkar; böylece madde istatistiklerinin güvenilirliği nispeten yükseltilmiş olur.

Madde analizi yapılacak test kaç maddelik olmalıdır? Bu konuda belirlenmiş bir sınır yoktur. Fakat, analizden sonra düzenlenecek son testte kullanılacak madde sayısının iki - üç katı kadar maddenin denenmesinin yeterli olacağı söylenir. Söz gelişi, 40 maddelik bir son test düzenlenecekse, 80 - 120 maddenin denenmesi gerekecek; bu da uygulamanın en az iki ayrı oturumda yapılmasını gerektirecektir.

Madde analizi yapılacak bir test yeterince bol zamanla uygulanmalıdır. Cevaplama süresi kısa gelen testlerde erişilememiş madde sayısı artar; madde analizinde «boş» frekansları artar, diğer seçeneklerin frekansları ve madde istatistikleri önemli derecede hatalı olur. Cevaplama süresi kısa gelen testlerde, cevaplayıcılar, testin sonuna doğru bazı maddeleri okumadan gelişigüzel cevaplandırdıklarından, madde istatistikleri yine önemli derecede hatalı olur. Onun için, madde analizi yapılacak testler, cevaplayıcıların büyük bir kısmının bütün maddeleri okumalarına yetecek kadar bol zamanla uygulanmalıdır.

İç ölçütle analiz tekniği, kapsamı bir örnek olmayan testlere uygulandığında, test kapsamının çoğunluğuna göre ayrı düşen maddeler, ayırıcılık indisleri düşük olduğu için elenebilir. Fizik, kimya, astronomi gibi fen konularını kapsayan bir fen bilimleri testinde astronomi maddeleri, hem kapsamı farklı, hem de sayıları az olduğundan testin toplam puanıyla yüksek korelasyonlar vermeyebilir. Madde seçmede bu maddeler elenebilir. Bu gibi durumlarda testi alt bölümlere ayırmak ve her alt bölümü kendi içinde analiz etmek daha doğru olur.

MADDE İSTATİSTİKLERİNE GÖRE TEST YAPMA

Bir testin istatistiksel özellikleri, onu oluşturan maddelerin istatistiksel özelliklerine bağlıdır. Maddelerin güçlük ve ayırıcılık indisleri, geçerlikleri ve maddeler arası korelasyonlar bir testin istatistiksel özelliklerini tayin eden etkenlerdir. Test istatistikleri, o tes-

ti oluşturan maddelerin istatistiklerinden kestirilebilir. Bu kesimde, madde istatistiklerinden test istatistiklerini kestirmede kullanılan formüllerden birkaçı verilecek; sınıftaki başarıyı ölçmede kullanılan testler için, madde analizi sonuçlarına göre madde seçme yöntemlerinin birkaçına kısaca değinilecektir.

Madde ve Test Güçlüğü

Bir test oluşturan K tane maddenin güçlük indisleri toplamı, o testin ortalamasını verir. Bu teorem şu formülle ifade edilir :

$$(18.3) \quad \overline{X} = \frac{\sum p_j}{K}$$

p_j : j maddesinin güçlük indisi
 K : Testteki madde sayısı
 $\overline{X}$: Test puanlarının ortalaması

Bu denklemin sağ tarafındaki toplamın K ile bölünmesiyle, madde güçlük indislerinin ortalaması elde edilir. Sol tarafın K ile bölünmesiyle bulunan $\overline{X}/K$ oranına da, testin ortalama güçlüğü denir. Aşağıdaki denklemde görüleceği gibi, bir testin ortalama güçlüğü, onu oluşturan maddelerin güçlük indislerinin ortalamasına eşittir.

$$(18.4) \quad \overline{X}/K = \overline{p}$$

$\overline{p}$: Madde güçlük indislerinin ortalaması

Test ortalamasının veya testin ortalama güçlüğü'nün belli bir değerde olması istenirse, seçilen maddelerin istenilen değeri verip vermediği yukarıdaki denklemlerden biri kullanılarak kontrol edilir. Maddeler, istenilen değer elde edilinceye kadar, değiştirilmeye devam edilir.

Sınıftaki başarının değerlendirilmesinde kullanılacak testlerin ortalama güçlüklerinin, genellikle, 0,50 civarında olması istenir. Orta güçlükte bir test, ya hepsi 0.50 civarında indisine sahip maddelerden, ya da çok çeşitli güçlük indislerine sahip fakat ortalama değeri $\overline{p} = 0.50$ civarında olan maddelerden meydana gelebilir. Bu iki test aynı niteliklerde değildir. Madde güçlük indisleri birbirine yakın bir test, dar bir yetenek veya başarı bölgesinde ayırıcı olur. Madde güçlük indisleri birbirinden farklı olan bir test ise, daha geniş yetenek sınırları arasında ayırıcı olur. Onun için, sınıfta kullanılacak testlerin çeşitli güçlüklerde, çoğunluğu orta güçlükte ve ortalama güçlüğü $\overline{p} = 0.50$ civarında olan maddelerden oluşturulması tavsiye edilir.

Madde İstatistikleri ve Test Standart Kayması

K tane maddeden oluşan bir testin ham puanlarının standart kayması, madde istatistiklerinin fonksiyonu olarak, şu eşitlikle verilir.³

$$(18.5) \quad S_x = \sum_{j=1}^k r_j \sqrt{p_j (1 - p_j)}$$

S_x : Test ham puanlarının standart kayması
 r_j : j maddesinin ayıricılık indisi
 p_j : j maddesinin güçlük indisi

Burada görüldüğü gibi, test ham puanlarının standart kayması, hem madde standart kayması (kare köklü ifade) ile hem de madde - test korelasyonu ile oranlı olarak artar; madde güçlük indisinin her madde için 0.50 olduğu halde de en büyük değerini alır. Sınıfta kullanılan başarı testlerinde, genellikle, ham puanlar standart kaymasının mümkün olduğu kadar büyük olması istenir. Bu sebeple, madde - test korelasyonu mümkün olduğu kadar yüksek ve madde güçlük indisi 0.50 civarında olan maddeler seçilir.

Madde İstatistikleri ve Test Güvenirliği

Madde istatistikleriyle testin güvenilirlik katsayısı arasındaki ilişki, Kuder - Richardson formülü (K - R No. 20) adıyla bilinen ve üçüncü bölümde verilen (3.3) numaralı eşitlikle ifade edilir. O formüldeki $\sum_j (1 - p_j)$ terimi madde variansları toplamı olup, ortalama cinsinden $K (s_j^2)$ halinde yazılabilir. Aynı formüldeki S_x^2 terimi de, formül (18.5) göz önünde bulundurularak, $K^2 (r_j s_j)^2$ halinde yazılır. Buradaki $r_j s_j$ ifadesi j maddesinin ayıricılık indisiyle standart kaymasının çarpımı olup, **madde güvenirlilik indisi** adını alır. O halde, bir testin güvenilirlik katsayısı, o testteki madde sayısının, madde varianslarının ortalama değerinin ve madde güvenirlilik indislerinin ortalama değerinin bir fonksiyonu olarak şu formülle ifade edilir:

$$(18.6) \quad r_x = \frac{K}{K - 1} \left[1 - \frac{(s_j^2)}{K (r_j s_j)^2} \right]$$

r_j : Madde - test korelasyonu
 K : Testteki madde sayısı
 s_j : Madde standart kayması

Bu formülün incelenmesinden anlaşılacağı gibi, test güvenirliliğinin yüksek olması için, büyük parantez içindeki kesirin payı mümkün olduğu kadar küçük, paydası ise mümkün olduğu kadar büyük

³H. Gulliksen, *Theory of Mental Tests*, s. 377.

olmalıdır. Sadece standart kayması büyük olan (güçlük indisleri 0,50 civarında olan) maddelerin seçilmesi hem payı hem paydayı büyütür. Fakat, standart kayması büyük olan maddelerden madde - test korelasyonları mümkün olduğu kadar büyük olan maddeler seçilirse, payda daha fazla büyütülebilir. Şüphesiz, standart kayması küçük olan maddeler de seçilebilir, fakat bu seçim test standart kaymasının büyütülmesi şartına zıt düşer. Gulliksen⁴ tarafından da belirtildiği gibi, s_x ve r_x istatistiklerinin test güvenilirliğine etkileri K'ya göre daha azdır. Onun için, madde - test korelasyonları yüksek maddeler seçilmek kaydıyla, güçlük indisleri 0,50 civarında olan maddeler seçilse bile, yeter sayıda madde ile istenilen bir güvenilirlik seviyesine ulaşılabilir.

Test geliştirmede, test güvenilirliği ile ilgili olarak sık karşılaşılan bir problem, belli bir güvenilirlik katsayısına kaç maddelik bir testle ulaşılabilceğinin kestirilmesidir. Bu amaçla, Spearman - Brown formülü adıyla bilinen bir bağıntı kullanılır. Bu bağıntının yeniden düzenlenmiş şekli aşağıdadır :

$$(18.7) \quad n = \frac{r_n - r_1 r_n}{r_1 - r_1 r_n}$$

r_1 : Birim testin güvenilirlik katsayısı
 n : Uzun testteki birim sayısı
 r_n : Uzun testin güvenilirlik katsayısı

Bu formülün kullanılışı, şu örnekte görülebilir : Beş maddelik bir test seçilmiş ve güvenilirlik katsayısı K—R No. 20 formülü ile hesaplanıp 0,30 bulunmuş olsun. Son testin güvenilirlik katsayısının 0,90 olması istenirse, aynı nitelikteki maddelerden kaç tane seçilmelidir? Yukarıdaki formülden

$$n = (0,90 - 0,30 \times 0,90) / (0,30 - 0,30 \times 0,90) = 63/3 = 21$$

bulunur. Demek ki, güvenilirlik katsayısının 0,90 değerine ulaşabilmesi için, testin 21 birim uzunluğunda veya $5 \times 21 = 105$ maddelik olması gerekir.

Kuder - Richardson formülüyle hesaplanan güvenilirlik katsayısı, test maddelerinin birbirine benzerliğinin veya testin bir örnekliğinin (homojenliğinin) ölçüsüdür. Bu katsayıyı haddinden fazla yükseltmek, test kapsamını daraltabileceğinden, test geçerliğini düşürebilir. Madde seçiminde, Nunnally tarafından işaret edildiği gibi, güvenilirlik katsayısı belli bir değere ulaştıktan sonra, onu daha fazla artırma gayretlerinden vazgeçilebilir.⁵

⁴Ibid, s. 379

⁵J.C. Nunnally, *Psychometric Theory*, s. 263 - 269.

Madde İstatistikleri ve Test Geçerliği

Yukarıda verilen basit madde analizinde üst ve alt gruplardaki doğru cevap sayısından hesaplanan d değeri, maddenin ayırıcılığının ve bu sebeple geçerliğinin ölçüsüdür. Bu ölçü, madde istatistiklerinden test istatistiklerini kestiren formüllerde kullanılmaz; fakat, test istatistikleri kestirilmeyecekse, ayırıcılığı yüksek olan maddeleri seçmekte kullanılabilir. Bunun için, basit analize tabi tutulmuş maddeler, d indisi en büyük olandan en küçük olana doğru sıralanır. Ayırıcılık indisleri sıfır ve negatif olan maddeler elenir. Kalanlar arasından, p güçlük indislerine de bakılarak, d indisi yüksek olan maddeler seçilir. Böylece meydana getirilen bir test, iç ölçüte göre geçerliğe sahip olur.

Ayırıcılık indisleri ve p güçlük indisleri eşit maddeler arasında bir seçim yapmak gerektiğinde, test planının ve test kapsamının dikkate alınması, testin kapsam geçerliğini artırabilir.

Madde - test korelasyonunun hesaplanmasına imkân veren bir analiz uygulanmışsa, bu korelasyon da, bir ayırıcılık indisi olarak değerlendirilir. Madde - test korelasyonu sıfır veya negatif olan maddeler teste dahil edilmez. Ebel, madde - test korelasyonlarının değerlendirilmesinde şu tavsiyelerde bulunur.⁶ Çift serili korelasyon katsayısı 0,40 ve daha yukarı değerde ise madde çok iyi, 0,30 - 0,40 arasında ise iyidir. Bu gibi maddeler, genellikle, düzeltme gerektirmez. Bu korelasyon 0,20 - 0,30 arasında ise, madde, zorunlu hallerde aynen kullanılabilir veya ıslah edilebilir. Ayırıcılık indisi 0,20'den küçük maddeler kullanılmamalıdır.

Madde ve test geçerliğiyle ilgili olarak sık sık karşılaşılan bir problem şudur : Madde geçerlik veya ayırıcılık indisi, ister d ile ister r ile hesaplansın, madde güçlüğüne bağlıdır. Çok güç ve çok kolay maddelerde bu indisler sıfıra yaklaşır. Durum böyle olunca, sınıftaki öğrencilerin hemen hemen hepsi tarafından doğru olarak cevaplandırılan bir madde ile aynı şekilde sınıftaki öğrencilerin hemen hemen hepsi tarafından yanlış olarak cevaplandırılan bir madde teste dahil edilmeyecek midir? Bu problem yüzde yirmi yedi alt ve üst gruplar ölçüt alındığı için doğar. Böyle bir ölçüt seçilince, daha başta, ayırıcı olmayan maddelerin eleneceği kabul edilmiş olur. Çok güç veya çok kolay bir maddenin teste bulunması kapsama dayalı gerekçelerle savunulabiliyorsa, şüphesiz bu tür maddeler de,

⁶R.L. Ebel, *Measuring Educational Achievement*, s. 364.

başka kusurları bulunmamak şartıyla teste dahil edilebilir. Fakat daha iyi bir çözüm, bu nitelikteki maddeleri yeniden gözden geçirerek güçlüklerini orta derecelere getirmektir.

Madde Seçme Yöntemleri

Yukarıdaki kesimde madde seçimi, test istatistiklerinin herbiri için tek tek ele alındı. Aslında, bir test geliştirilirken bu istatistiklerin hep bir arada düşünülmesi gerekir. Öyle olunca, test istatistiklerinin hepsini birden istenilen değerlere çıkarmanın mümkün olmadığı; bazı isteklerin birbiriyle çeliştiği açıkça görülür. Nitekim, test standart kaymasını büyüten madde istatistiklerinin, test güvenilirliğini düşürebileceği yukarıda tartışıldı. Bir test için madde seçiminde birbirine karşı işleyen etkiler bulunduğunda, bunları uzlaştıracı bir yöntem veya orta yol aranır. Test uzmanları, madde güçlüğü, madde ayıricılık indisi ve madde geçerliği gibi istatistiklerin test güvenilirliğini ve özellikle test geçerliğini yükseltmede uzlaştırılmalarını sağlayan yöntemler bulmuşlardır. Gulliksen grafik metotlar verir.⁷ Thorndike bu problemi çeşitli yöntemleriyle tartışıp tavsiyeler verir.⁸ Nunnally bazı özel durumlarda uzlaştıracı çözümün nasıl bulunabileceğini gösterir.⁹

Sınıfta kullanılacak başarı testleri için, madde analizi sonuçlarına göre madde seçiminde hepsi birden dikkate alınması gerekli beş şart vardır.

1. Seçilen maddelerin kapsamı testin belirtke tablosuna uygun olmalı; gerek konular gerekse ölçülen davranışlar bakımından testin kapsamı daraltılmamalı ve dengesi bozulmamalıdır.
2. Testin ortalama güçlüğü için başka bir zorunluk yoksa, seçilen maddelerin küçük indislerinin ortalaması 0,50 değerine yakın olmalıdır.
3. Seçilen maddelerin güçlük indislerinin dağılımı simetrik olmalı; orta güçlükteki maddeler çok sayıda, güç ve kolay maddeler ise gittikçe azalan sayılarda olmalıdır.
4. Test standart kayması mümkün olduğu kadar büyütülmelidir.
5. Test güvenilirliği istenilen basamağa kadar yükseltilmelidir.

⁷Gulliksen, *op. cit.*, s. 393 - 384.

⁸R.L. Thorndike, *Personel Selection*, s. 227 - 256

⁹Nunnally, *op. cit.*, s. 236 - 286.

Bu şartların hepsini birden yerine getirmek çok zor bir iştir. Eğer üzerinde madde analizi yapılmış maddelerin sayısı, nihai teste dahil edilecek madde sayısından çok fazla değilse veya maddelerin çoğu kusurlu ise istenilen niteliklerde madde bulma mümkün olmayabilir. Gerekli seçmeyi yapmaya yetecek kadar madde varsa, izlenebilecek pratik bir yol şudur :

Önce, bütün maddelerin analiz sonuçları gözden geçirilerek kusurlu maddeler, güçlük indisleri ve ayırma güçleri bakımından teste dahil edilemeyecek maddeler ayıklanır. Sonra, geri kalan maddeler kapsamları ve ölçtükleri davranışlar açısından incelenerek, belirtke tablosundaki yerleri işaretlenir. Bundan sonra, bu maddeler ayırıcılık indislerine göre sıralanır. Madde güçlükleri ve kapsamları da dikkate alınarak, bu sıranın başından itibaren en ayırıcı olan maddeler seçilir. Seçim tamamlandıktan sonra tekrar belirtke tablosuna dönülür ve nihai testin kapsamı kontrol edilir. Gerekirse eklemeler ve çıkarmalar yapılır.

ANALİZ SONUÇLARINA BAKARAK MADDE DÜZELTME

Madde analizinin önemli amaçlarından biri, cevapların madde seçenekleri üzerindeki dağılımına bakarak, maddenin işleyişini kestirmek; varsa, kusurlarını ortaya çıkarmaktır. Bu tür bir analiz hemen her yapıdaki maddelere uygulanabilmekle birlikte, en ziyade seçmeli maddelere uygulanır. Onun için, bu kesimdeki örnekler çoktan seçmeli maddelerden seçilmiştir.

İdeal bir maddede cevapların seçenekler üzerine dağılımının nasıl olacağı farazi bir misalle açıklanabilir. Bir maddenin doğru cevabını «gerçekten» bilenlerin üst grupta %60, alt grupta % 20 olduğu kabul edilsin; her iki grupta doğru cevabı bilmeyenlerin de maddeyi tesadüfen cevaplandırırları farzedilsin. O halde, doğru cevabı bilmeyenler (üst grupta % 40 ve alt grupta % 80) bu maddenin seçenekleri üzerine eşit ihtimalle dağılacaklardır. Üst ve alt grupta yüzer kişi varsa ve madde beş seçenekliyse dağılımın şöyle olması beklenir :

Madde No. 3	A	B	C*	D	E	Boş	N'
Üst	8	8	68	8	8	0	100
Alt	16	16	36	16	16	0	100

Burada görüldüğü gibi, iki grupta doğru cevabı bilenlerin oranları % 60 ve % 20 değil, % 68 ve % 36 olacaktır. Buna sebep, mad-

deyi tesadüfen cevaplandıranların bir kısmının da doğru cevabı şansa tutturmuş olmalarıdır.

Seçmeli testlerde madde analizi tecrübesi olanlar, bu ideal dağılımın pek seyrek görüldüğünü bilirler. Daha sık görülen dağılım, çeldiricilerin de kendi aralarında «çeldiricilik» sırasına dizilmeleridir. Gerçek bir durumda, yukarıdaki ideal dağılımın değil, aşağıdaki gibi bir dağılımın ortaya çıkması daha muhtemeldir.

Madde No. 4	A	B	C*	D	E	Boş	N'
Üst	6	17	64	3	9	1	100
Alt	15	26	31	8	17	1	100

Bu dağılımda görüldüğü gibi, doğru cevap frekansları her iki grupta da en yüksektir. Çeldiriciler, en kuvvetli çeldiriciden en zayıf çeldiriciye doğru, B - E - A - D sırasıyla sıralanmışlardır. Madde kusursuz ise, doğru cevabın en yüksek frekansları toplamaları, çeldiricilerin de çeldiricilik sırasına göre gittikçe azalan frekanslarda dizilmeleri beklenir. Ayrıca, doğru cevapta üst grup frekansının büyük, alt grup frekansının küçük olması; çeldiricilerde ise bunun tersi beklenir.

Seçmeli test maddelerinde sık görülen kusurlardan bazıları ve bunların madde cevaplarındaki belirtileri aşağıda kısaca tartışılmaktadır.

1. Çeldiricilerinden biri doğru cevabından daha çekici olan madde : Aşağıdaki madde, modern fizik programlarını değerlendirmede kullanılan testlerin birinden alınmıştır.

5. Aşağıdaki olayların hangisinde güneş ışığı renklere ayrılır?

- | | |
|--------------------------|--------------------------|
| A) Prizmada kırılma | B) İnce zarlarda yansıma |
| C) Dar yarıktaki girişim | D) Yukarıdakilerin hepsi |

Bu maddenin doğru cevabı D'dir. Madde analizi aşağıdaki sonuçları vermiştir.

Madde No. 5	A	B	C	D*	Boş	N'	
Üst	45	11	12	32	0	100	$p = 0,23$
Alt	72	8	6	14	0	100	$r = 0,25$

Bu dağılımda görüldüğü gibi, A çeldiricisi doğru cevaptan fazla frekans toplamıştır. Maddenin kusuru, A çeldirisinde aranmalıdır.

Madde incelenince görülür ki, prizmada renklere ayrılma olayı ilk okuldan itibaren okullarda öğretilen bir olaydır. B ve C çıkırları modern fizikte öğrenilen olaylardır; bunları bilmeyenler elbette A'yı işaretlerler. Üstelik, doğru cevabı bulabilecek kadar bilgili olanların bir kısmı da acelecilik sebebiyle A'yı işaretleyebilirler. A yerine daha az çekici ve B ile C'ye benzer bir seçenek bulunabilirse, bu madde ıslah edilebilir.

Cevapların çeldiricilerden birinde yığılması, çok zor maddelerde görüldüğü gibi, çeldiricinin öğrenciler arasında yaygın bir «yanlış bilgi» ifade ettiği hallerde de görülür. Bu tip maddeler, genellikle, düşük ayırıcılık indisi verir.

2. Doğru cevabı bulunmayan veya çok güç olan madde : Böyle maddelerde bütün seçeneklerin frekansları aşağı yukarı eşit çıkar. Böyle olmasa da seçenekler çeldiricilik sırasına dizilseler bile, maddenin ayırıcılık indisi sıfır civarında bulunur.

3. Çok kolay olan veya çeldiricileri çekici olmayan madde : Böyle maddelerde, hem üst hem alt grubun büyük çoğunluğu doğru cevapta toplanır. İki grup arasındaki fark azaldığından, maddenin ayırıcılık indisi sıfıra yakın bulunur. Böyle bir madde, doğru cevaba daha yakın çeldiriciler bulmak suretiyle ıslah edilebilir.

4. Çeldiricileri bir örnek olmayan, çeldiricileri üst gruba alt grubu ayırt etmeyen madde : Aşağıdaki madde, bir kelime bilgisi testinden alınmıştır. Bu testte, cevaplayıcılardan, baştaki kelimeyle eş anlamlı olan kelimeyi bulmaları istenir.

6. İzafi :

A) Palyatif

B) Nispi

C) İndi

D) Sübjektif

Bu madde analizde şu sonuçları vermiştir :

Madde No. 6	A	B*	C	D	Boş	N'	%	
Üst	13	30	4	32	1	70	43	$p = 0.40$
Alt	4	26	12	34	4	70	37	$r = 0.06$

Bu madde pek zor olmadığı halde, ayırıcı değildir. Buna sebep, çeldiricilerin iyi işlemeyişidir. İyi bir çeldiricinin frekansı üst grupta küçük, alt grupta büyük olmalıdır. Frekanslar arasındaki fark ne kadar büyükse, çeldirici o kadar iyidir. Böyle bakılınca, yukarıdaki maddenin çeldiricilerinden sadece C'nin işlediği, A ile D'nin kusurlu

oldukları görülür. A'nın kusuru üst grubu daha çok çeldirmesindedir. D'nin kusuru ise hem çok fazla çekici olması, hem de iki grubu yeterince ayıramamasıdır. Bu madde, analizden sonra şöyle düzeltilmiştir :

6a. İzafî

A) Zahirî

B) Nispi

C) Indî

D) Keyfî

Bu düzeltmeden sonra yapılan uygulamada, madde analizi şu sonuçları vermiştir.

Madde								
No. 6a.	A	B*	C	D	Boş	N'	%	
Üst	8	42	9	11	0	70	60	$p = 0,39$
Alt	19	14	20	16	1	70	20	$r = 0,42$

Bu tabloda görüldüğü gibi, maddenin çeldiricilerinin değiştirilmesi, madde gücünü hemen hiç etkilemediği halde ayırıcılık indisini önemli derecede artırmıştır.

5. Anahtarı hatalı olan madde : Madde analizi, bazı hallerde anahtar hatalarını da gösterir. Aslında madde iyi yazılmış ve doğru anahtarlanmış; fakat, puanlama anahtarı hazırlanırken hata yapılmış olabilir. Bu halde, maddenin ayırıcılık gücü negatif çıkar; anaharda doğru cevap olarak gösterilen seçeneğin frekansları küçük, diğer bir seçeneğin frekansları daha büyük ve pozitif yönde ayırıcı bulunur. Şu analiz sonuçlarını inceleyelim :

Madde								
No. 7	A*	B	C	D	E	Boş	N'	%
Üst	7	2	32	3	2	0	46	15 $p = 0,21$
Alt	13	8	16	4	5	0	46	28 $r = -0,18$

Bu test, bu maddenin doğru cevabını A kabul eden bir anahtarla puanlanmıştır. Madde analizinden sonra anahardan şüphelenilmiş; maddenin aslı incelendiğinde doğru cevabın C olduğu görülmüştür. Nitekim, doğru cevap C olduğuna göre, yukarıdaki frekanslar gerçek kabul edilse, bu madde için $p = 0,53$ ve $r = 0,35$ bulunur.

Madde analizi sonuçlarını yorumlayıp test maddelerini ıslah etmek, tecrübeyle kazanılan bilgi ve beceriler gerektirir. Öğretmen kendi yazdığı maddeleri, burada verilen tekniklerle analiz eder; düzeltir ve tekrar analiz ederse bu tecrübeyi kazanır.

EK 1

ÇİFT SERİLİ KORELASYONUN BİR DAĞILIMIN ALT VE ÜST YÜZDE YİRMİ YEDİ UÇLARINDAN KESTİRİLMESİ

Bir test maddesinin geçerlik indisi olarak, çoğu halde, ayırma gücü denilen madde - test korelasyonu hesaplanır. Bu korelasyon bir tek maddenin puanları ile testin toplam puanları arasındaki korelasyondur.

Madde puanı sadece 0 ve 1 değerlerini alabildiğinden iki değerli bir değişken, test toplam puanı ise çok değerli bir değişkendir. Böyle hallerde çift serili korelasyon hesaplanır. Madde - test korelasyonu, madde puanları dağılımının normal olduğu kabul edilerek, madde ve test istatistiklerinden, bir çift serili korelasyon olarak hesaplanabilir. Fakat, hesaplamalar makina kullanılmaksızın yapılamayacak kadar uzun zaman alır. Çift serili korelasyonun madde ve test istatistiklerinden hesaplanmasındaki bu güçlük, onun zahmetsizce kestirilmesine yarar bazı araçların geliştirilmesine yol açmıştır. Bu ekte verilen tablo, test puanları dağılımının % 27 alt ve üst uçlarında maddeyi doğru cevaplandıranların yüzdelerinden, çift serili korelasyonun bulunması için hazırlanmıştır.

Bu ekte verilen tablo, alt grupta doğru cevap yüzdesi p_a için 0,01'den 0,88'e kadar, üst grupta doğru cevap yüzdesi p_u için de 0,20'den 0,98'e kadar çift değerleri kapsar. Fakat, tek değerlere karşılık olan korelasyonlar, ulama (enterpolasyon) yöntemiyle bulunabilir.

Basım kolaylığı sağlanması için tabloda sıfır ve ondalık virgül gösterilmemiştir. Tablodaki sayıların hepsi virgülden sonraki iki basamak olarak okunmalıdır.

Tablonun kullanılmasında aşağıdaki yöntem izlenir.

1. Üst grup ve alt grup yüzdeleri çift sayılar ise ve üst grupta doğru cevap yüzdesi alt gruptakinden büyükse, çift serili korelasyonun değeri tablodan doğrudan doğruya okunur. Meselâ $p_a = 0,50$ ve $p_u = 2,24$ için tablonun birinci sayfasında 50 satırıyla 24 sütununda 28 okunur. Aranılan korelasyon $r = 0,28$ 'dir.

2. İki grupta doğru cevap yüzdeleri eşitse, çift serili korelasyonun değeri sıfırdır.

3. Alt gruptaki doğru cevap yüzdesi üst gruptaki doğru cevap yüzdesinden büyük ise korelasyon negatif çıkar. Tablonun pozitif ve negatif değerleri simetrik olduğundan, negatif değerler tabloya alınmamıştır. Negatif değerler, p_a ile p_a 'nın değerleri birbiriyle değiştirilerek tablodan bulunabilir. Meselâ, $p_a = 0,20$ ve $p_a = 0,28$ için, tablonun 28 satırında ve 20 sütununda 10 okunur bunun negatif değeri, aranan korelasyondur; $r = -0,10$ bulunur.

4. Doğru cevap yüzdelerinin tek değerleri için tablodan ulama yapılabilir. Meselâ, $p_a = 0,50$ ve $p_a = 0,20$ için $r = 0,33$; $p_a = 0,52$ ve $p_a = 0,20$ için $r = 0,35$ okunur. Tabloda $p_a = 0,51$ satırı bulunmamakla birlikte, $p_a = 0,51$ ve $p_a = 0,20$ için aranan korelasyonun $0,33$ ile $0,35$ değerleri arasında olduğu bilinir. Onun için, $r = (0,33 + 0,35)/2 = 0,34$ alınabilir. Aynı şekilde $p_a = 0,52$ ve $p_a = 0,25$ için $r = (0,30 + 0,28)/2 = 0,29$ bulunur.

5. Tabloda her iki boyutta da ulama yapılabilir. Meselâ, $p_a = 0,41$ ve $p_a = 0,19$ için aranan korelasyonun $0,26$ ile $0,24$ 'ün ortalama değerinden büyük, fakat $0,28$ ile $0,26$ 'nın ortalama değerinden küçük olduğu düşünülürse

$$r = \left(\frac{0,26 + 0,24}{2} + \frac{0,28 + 0,26}{2} \right) / 2 = 0,26$$

bulunur.

TABLO E.1
ÇİFT SERİLİ KORELASYONUN BİR DAĞILIMIN
UÇLARINDAN KESTİRİLMİŞ DEĞERLERİ

	Alt grupta doğru cevap yüzdesi														
	01	02	04	06	08	10	12	14	16	18	20	22	24	26	28
20	53	45	35	28	22	18	14	10	06	03	00	—	—	—	—
22	55	47	38	31	25	20	16	12	09	06	03	00	—	—	—
24	57	49	40	33	28	23	19	15	12	08	05	03	00	—	—
26	58	51	42	35	30	25	21	18	14	11	08	05	03	00	—
28	60	53	44	38	32	28	24	20	17	13	10	08	05	02	00
30	61	54	46	39	34	30	26	22	19	16	13	10	07	05	02
32	63	56	48	41	36	32	28	25	21	18	15	12	10	07	05
34	64	57	49	43	38	34	30	27	23	20	17	15	12	09	07
36	65	59	51	45	40	36	32	29	25	22	20	17	14	12	09
38	66	60	52	47	42	38	34	31	27	24	22	19	16	14	11
40	67	61	54	48	43	39	36	32	29	26	24	21	18	16	13
42	68	63	55	50	45	41	38	34	31	28	26	23	20	18	15
44	69	64	57	51	47	43	39	36	33	30	28	25	22	20	17
46	70	65	58	53	48	44	41	38	35	32	29	27	24	22	19
48	71	66	59	54	50	46	43	39	37	34	31	29	26	24	21
50	72	67	61	56	51	48	44	41	38	36	33	30	28	26	23
52	73	68	62	57	53	49	46	43	40	37	35	32	30	28	25
54	74	69	63	58	54	51	47	45	42	39	36	34	32	29	27
56	75	70	64	59	56	52	49	46	43	41	38	36	33	31	29
58	76	71	65	61	57	54	51	48	45	43	40	38	35	33	31
60	76	72	66	62	58	55	52	49	47	44	42	39	37	35	33
62	77	73	68	63	60	57	54	51	48	46	44	41	39	37	35
64	78	74	69	64	61	58	55	52	50	48	45	43	41	39	36
66	79	75	70	66	62	59	57	54	52	49	47	45	43	40	38
68	80	76	71	67	64	61	58	56	53	51	49	47	44	42	40
70	80	77	72	68	65	62	60	57	55	52	50	48	46	44	42
72	81	78	73	69	66	63	61	59	56	54	52	50	48	46	44
74	82	79	74	70	67	65	62	60	58	56	54	52	50	48	46
76	82	79	75	72	69	66	64	62	60	57	56	54	52	50	48
78	83	80	76	73	70	68	65	63	61	59	57	55	54	52	50
80	84	81	77	74	71	69	67	65	63	61	59	57	56	54	52
82	85	82	78	75	73	71	68	67	65	63	61	59	57	56	54
84	85	83	79	77	74	72	70	68	66	65	63	61	60	58	56
86	86	84	80	78	76	74	72	70	68	67	65	63	62	60	59
88	87	85	82	79	77	75	73	72	70	68	67	65	64	62	61
90	88	86	83	80	79	77	75	74	72	71	69	68	66	65	63
92	89	87	84	82	80	79	77	76	74	73	71	70	69	67	66
94	90	88	86	84	82	80	79	78	77	75	74	73	72	70	69
96	91	89	87	86	84	83	82	80	79	78	77	76	75	74	73
98	92	91	89	88	87	86	85	84	83	82	81	80	79	79	78

TABLO E. 1 (Devamı)

		Alt grupta doğru cevap yüzdesi														
		30	32	34	36	38	40	42	44	46	48	50	52	54	56	58
Üst grupta doğru cevap yüzdesi	30	00	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	32	02	00	—	—	—	—	—	—	—	—	—	—	—	—	—
	34	05	02	00	—	—	—	—	—	—	—	—	—	—	—	—
	36	07	04	02	00	—	—	—	—	—	—	—	—	—	—	—
	38	09	07	04	02	00	—	—	—	—	—	—	—	—	—	—
	40	11	09	07	04	02	00	—	—	—	—	—	—	—	—	—
	42	13	11	09	06	04	02	00	—	—	—	—	—	—	—	—
	44	15	13	11	08	06	04	02	00	—	—	—	—	—	—	—
	46	17	15	13	11	08	06	04	02	00	—	—	—	—	—	—
	48	19	17	15	13	10	08	06	04	02	00	—	—	—	—	—
Üst grupta doğru cevap yüzdesi	50	21	19	17	15	12	10	08	06	04	02	00	—	—	—	—
	52	23	21	19	17	14	12	10	08	06	04	02	00	—	—	—
	54	25	23	21	18	16	14	12	10	08	06	04	02	00	—	—
	56	27	25	23	20	18	16	14	12	10	08	06	04	02	00	—
	58	29	27	25	22	20	18	16	14	12	10	08	06	04	02	00
	60	31	29	26	24	22	20	18	16	14	12	10	08	06	04	02
	62	33	30	28	26	24	22	20	18	16	14	12	10	08	06	04
	64	34	32	30	28	26	24	22	20	18	17	15	13	11	08	06
	66	36	34	32	30	28	26	25	23	21	19	17	15	13	11	09
	68	38	36	34	32	30	29	27	25	23	21	19	17	15	13	11
Üst grupta doğru cevap yüzdesi	70	40	38	36	34	33	31	29	27	25	23	21	19	17	15	13
	72	42	40	38	36	35	33	31	29	27	25	23	21	19	17	15
	74	44	42	40	39	37	35	33	31	29	28	26	24	22	20	18
	76	46	44	43	41	39	37	35	33	32	30	28	26	24	22	20
	78	48	47	45	43	41	39	38	36	34	32	30	29	27	25	23
	80	50	49	47	45	44	42	40	38	36	35	33	31	29	28	26
	82	52	51	49	48	46	44	43	41	39	37	36	34	32	30	28
	84	55	53	52	50	48	47	45	43	42	40	38	37	35	33	31
	86	57	56	54	52	51	49	48	46	45	43	41	39	38	36	34
	88	60	58	57	55	54	52	51	49	47	46	44	43	41	39	38
Üst grupta doğru cevap yüzdesi	90	62	61	59	58	57	55	54	52	51	49	48	46	44	43	41
	92	65	64	62	61	60	58	57	56	54	53	51	50	48	47	45
	94	68	67	66	64	63	62	61	59	58	57	56	54	53	51	50
	96	72	71	70	69	68	66	65	64	63	62	61	59	58	57	55
	98	77	76	75	74	73	72	71	70	69	68	67	66	65	64	63

TABLO E. 1 (Devamı)

		Alt grupta doğru cevap yüzdesi														
		60	62	64	66	68	70	72	74	76	78	80	82	84	86	88
Yüzdesi	60	00	—	—	—	—	—	—	—	—	—	—	—	—	—	—
	62	02	00	—	—	—	—	—	—	—	—	—	—	—	—	—
	64	04	02	00	—	—	—	—	—	—	—	—	—	—	—	—
	66	07	04	02	00	—	—	—	—	—	—	—	—	—	—	—
	68	09	07	04	02	00	—	—	—	—	—	—	—	—	—	—
Doğru cevap	70	11	09	07	05	02	00	—	—	—	—	—	—	—	—	—
	72	13	11	09	07	05	02	00	—	—	—	—	—	—	—	—
	74	16	14	12	09	07	05	02	00	—	—	—	—	—	—	—
	76	18	16	14	12	10	07	05	03	00	—	—	—	—	—	—
	78	21	19	17	15	12	10	08	05	03	00	—	—	—	—	—
Doğru grupta	80	24	22	20	17	15	13	10	08	05	03	00	—	—	—	—
	82	26	24	22	20	18	16	13	11	08	06	03	00	—	—	—
	84	29	27	25	23	21	19	17	14	12	09	06	03	00	—	—
	86	32	31	29	27	25	22	20	18	15	12	10	07	04	00	—
	88	36	34	32	30	28	26	24	21	19	16	14	11	07	04	00
Üst grupta	90	39	38	36	34	32	30	28	25	23	20	18	15	12	08	04
	92	43	42	40	38	36	34	32	30	28	25	22	20	16	13	09
	94	48	47	45	43	41	39	38	35	33	31	28	25	23	19	15
	96	54	52	51	49	48	46	44	42	40	38	35	33	30	26	23
	98	61	60	59	57	56	54	53	51	49	47	45	43	40	37	34

Bu tablodaki sayılar şu kaynaktan kısaltılarak alınmıştır :

C - T. Fan, Item Analysis Table, Princeton, N.J., Educational Testing Service, 1952.

Bulduru

Yazar Soyadları Dizisi

- Akhun, I. 170, 265, 276.
Ahmann, J.S., 149, 277.
American Educational Research Association, 278.
Anastasi, A., 278.
Anderson, R.H., 222, 251, 278, 280.
Arici, H., 17, 205, 218, 265, 276.
Başbakanlık, 233, 276.
Baykul, Y., 255, 276.
Bean, K.L., 278.
Beaumont, H., 236, 278.
Blood, D.F., 278.
Bloom, B.S., 118, 135, 143, 146, 151, 155, 278, 280, 281, 283.
Bormuth, J.R., 278.
Brody, W., 114, 278.
Brown, F.G., 278.
Bush, R.R., 10, 282.
Campbell, N.R., 12, 278.
Carroll, J.A., 54, 278.
Carter, H.D., 278.
Carter, R.S., 278.
Cartwright, C.A., 278.
Chase, C.I., 278.
Chauncey, H., 278.
Coffman, W.E., 53, 54, 278, 280, 281.
Coombs, C.H., 10, 278.
Coppedge, F., 282.
Cox, R.C., 280.
Cronbach, L.J., 220, 279.
Cureton, E.E., 12, 16, 17, 279.
Davis, F.B., 107, 108, 279.
Dawes, R.M., 10, 278.
Dobbin, J.E., 278.
Doğan, H., 255, 276.
Dresel, P.L., 100, 150, 239, 279.
Ebel, R.L., 12, 37, 75, 99, 111, 115, 235, 240, 256, 270, 279, 280.
Ertürk, S., 118, 119, 120, 121, 276.
Fan, C.-T., 265, 279, 288.
Fifer, G., 108, 279.
Findley, W.G., 141, 279.
Finlayson, D.S., 54, 279.
Fitzpatrick, R., 164, 279.
Flynn, J.T., 279.
French, W., 279.
Fryklund, V.C., 164, 279.
Furst, E.J. 280.
Gage, N.L., 149, 282.
Galanter, E., 10, 282.
Garber, H., 279.
Gerberich, R., 280.
Glaser, R., 280.
Glock, M.D., 149, 277.
Godshalk, F.I., 280.
Goodlad, J.I., 222, 251, 280.
Gönen, Z., 170, 276.
Greene, H.A., 280.
Gronlund, N.E., 118, 280.
Gulliksen, H., 268, 269, 271, 280.
Hagen, E., 26, 40, 53, 72, 100, 152, 178, 221, 258, 282.
Hart, F.M., 280.
Hartnett, J.R., 114, 280.
Hastings, J.T., 118, 135, 143, 146, 151, 155, 281, 283.
Heathers, F., 256, 280.
Hedges, W.D., 280.
Henry, N.B., 141, 279.
Howkes, H.E., 117, 280.
Huddleson, E.M., 280.
Jackson, D.N., 280.
Jacob, P.E., 151, 150, 280.
Jones, L.V., 12, 280.
Jorgensen, 280.
Kallish, R.A., 174, 280.

- Karnes, M.R., 281.
 Klein, S.P., 280.
 Koç, N., 276.
 Krathwohl, D.R., 118, 120, 280.
 Lado, R., 280.
 Lange, A., 280.
 Lehmann, I.J., 280, 281.
 Lindeman, R.H., 281.
 Lindquist, E.F., 117, 280, 281.
 Lindvall, C.M., 118, 281.
 Luce, R.D., 10, 282.
 Ludlow, H.C., 278.
 Lyman, H.B., 281.
 Macomber, F.G., 236, 278.
 Madaus, G.F., 118, 143, 146, 151, 155, 281, 283.
 Mager, R.F., 281.
 Mann, C.R., 117, 280.
 Marshall, M.S., 222, 281.
 Mart, Ö., 53, 276.
 Masla, B.B., 120, 280.
 Mayhew, L.B., 279.
 McConville, C.M., 53, 281.
 McCune, G.H., 281.
 Mehrens, W.A., 280, 281.
 Messick, S., 280.
 Michaels, W.J., 281.
 Miller, D.M., 281.
 Milli Eğitim Bakanlığı, 234, 252, 276.
 Millman, J., 281.
 Morrison, E.J., 164, 279.
 Morse, H.T., 281.
 Mosier, C.I., 100, 281.
 Myers, A.E., 53, 281.
 Myers, M.C., 281.
 National Council of Teachers of Mathematics, 281.
 Nelson, C.H., 239, 279.
 Noll, V.H., 281.
 Nunnally, J.C., 269, 271, 281.
 Odell, C.W., 83, 97, 99, 281.
 Oğuzkan, T., 249, 250, 252, 254, 276.
 Orlandi, R.L., 151, 281.
 Osburn, H.G., 281.
 Owens, R.E., 282.
 Öncül, R., 164, 276.
 Özçelik, D.A., 127, 184, 277.
 Özoğlu, S.Ç., 249, 250, 252, 254, 276.
 Pauk, W., 281.
 Payne, D.A., 118, 280.
 PDK National Study Committee on Evaluation, 282.
 Pfanzagl, J., 10, 282.
 Popham, W.J., 118, 282.
 Powell, N.J., 114, 278.
 Price, H.C., 281.
 Remmers, H.H., 149, 152, 282.
 Ruch, G.M., 52, 282.
 Rummel, J.F., 149, 282.
 Scates, D.E., 141, 279.
 Schultz, M., 282.
 Shaw, M.E., 152, 282.
 Smith, E.R., 150, 282.
 Stalnaker, J.M., 47, 282.
 Stanley, J.C., 72, 73, 83, 152, 282.
 Stevens, S.S., 16, 20, 282.
 Strang, R., 260, 282.
 Suppes, P., 10, 282.
 Swineford, F., 52, 282.
 Tan, H., 277.
 Taylor, C.W., 54, 282.
 Tekin, H., 79, 239, 277.
 Thorndike, R.L., 12, 26, 40, 53, 72, 78, 87, 100, 152, 164, 178, 221, 258, 271, 278, 280, 282, 283.
 Thurstone, L.L., 20, 283.
 Tieges, E.W., 53, 283.
 Travers, R.M.W., 149, 283.
 Turgut, M.F., 42, 91, 191, 192, 240, 242, 249, 250, 252, 254, 276, 277, 283.
 Tussing, L., 174, 83.
 Tversky, A., 10, 278.
 Tyler, R.W., 118, 150, 166, 282, 283.
 Valette, R.M., 283.
 Varış, F., 277.
 Wesman, A.F., 78, 87, 100, 105, 283.
 Wilson, J.W., 154, 283.
 Wittenborn, J.R., 54, 283.
 Wolf, R.M., 283.
 Wood, D.A., 283.
 Wood, R., 283.
 Wright, J.M., 152, 282.
 Yurt, I., 277.
 Zinnes, J.L., 10, 282.

Konular Dizisi

- Ağırlıklı puanlar, 201
ile mutlak değerlendirme, 247
- Ağırlık verme, 226, 244
Ham puanlara, 245
Not ölçeğindeki ölçmelere, 245
Standart puanlara, 247
- Alt grup (Bkz : Madde analizi)
Alt test (Bkz : Test düzeni)
Anahtar (Bkz : Puanlama)
Aralık orta değeri, 195
Ayırma gücü (Bkz : Madde analizi)
Bağıl değerlendirme, 237
Bağıl değişkenlik, 207
Katsayısı, 208
Başarı oranları, 236, 240, 247
Belirtke tablosu, 124
Bilgi basamağındaki hedefler (Bkz.
Bilişsel davranışlar)
Bilişsel davranışların ölçülmesi, 128
Analiz davranışları, 142
İkkelere dönük, 143
İlişkilere dönük, 143
Öğelere dönük, 143
Bilgi davranışları, 128
Araç ve metodların bilgisi, 131
Bir alana özgü bilgiler, 129
Evrensel öğelerin bilgisi, 133
Değerlendirme davranışları, 146
Dış ölçütlerle değerlendirme, 147,
İç ölçütlerle değerlendirme, 147
Kavrama davranışları, 134
Çevirme, 135
Öteleme, 138
Yorumlama, 137
Sentez davranışları, 145
Uygulama davranışları, 140
Bilişsel hedefler, 128
Cevaplar
Çekinilmiş, 242
Erişilememiş, 242
Çeyrek kayma, 206
Dağılım, 207
Normal, 210
Sağı çarpık, 209
Simetrik, 210
Solu çarpık, 209
Yorumlanması, 207
Değerlendirme, 3, 146,
Bağıl, 237
Birkaç ölçmeyle, 244
Mutlak, 231, 236
Ölçekleri, 233
Ölçütleri, 227
Sonuçlarının bildirilmesi, 258
Öğrencilere, 259
Velilere, 259
Sonuçlarının kullanılması, 249
Ders geçmede, 254
Gruplamada, 255
Sınıf geçmede, 251
Yönetmede, 256
Tek ölçmeyle, 235
Üç öğesi, 224
Ders geçme düzeni, 254
Dizi genişliği, 194, 204
Doğru cevap (seçmeli testlerde), 87
Bileşik, 91
Bir tek, 87
Birden fazla, 89
En doğru, 88
Gizli, 91
Dolaylı ölçme, 22
Duyuşsal davranışlar, 149
Benlik tasarımı, 157
Güdü, 156
İlgi, 156
Kaygı, 157
Ölçülmesi, 150
Ölçülmesinde güçlükler, 159
Programlarda, 150
Takdir, 158
Tutum, 154
Elemeye cevaplama (seçmeli testler-
de), 193
Eşdeğer formlar, 32
Frekans (istatistik), 195
Dağılımı, 196
Grafiği, 196
Poliyonu, 197
Yığılmalı, 197
Geçerlik, 38
Etkileyen etmenler, 40

- Geçerlik ölçütü, 41
 Kapsam geçerliği, 39
 Katsayısı, 42
 Madde istatistikleri ve test geçerliği, 270
 Ödevlerde, 180
 Ölçmede, 38
 Sözlü sınavlarda, 110
 Uygulamada, 189
 Yazılı yoklamalarda, 51
 Yordamada, 257
- Gerçek puan, 35
 Gözlem, 162
- Gruplama (istatistik), 195
 Grup aralıkları, 195
 Grup orta noktası, 195
- Güçlük
 İndisi, 263
 Madde güçlüğü, 263
 Ortalama, 267
 Test güçlüğü, 267
 Geçerliğe etkisi, 44
 Güvenirliğe etkisi, 37, 58
- Güdü (sınavlarda), 189
- Güvenirlilik, 30
 Eşdeğer formlarla, 32
 Etkileyen etmenler, 36
 Hata, 30
 Kararlılık, 33
 Katsayısı, 32
 K - R formülleri, 34, 268
 Madde istatistikleri ve test güvenirliliği, 268
 Ölçmede, 30
 S - B formülü, 33, 269
 Sözlü sınavlarda, 110
 Standart hata, 35
 Test tekrarı ile, 32
 Tutarlılık, 32
 Yarı puanlarla, 33
 Yazılı yoklamalarda, 49
- Hatırlama ve tanıma, 128
 Hedef analizi, 116
 Hedeflerin sınıflanması, 119
 Programlarda hedefler, 118
 Test maddeleri ve hedefler, 126
 Hedef davranışlar, 116
 Bilişsel, 122, 128
 Duyuşsal, 123, 149
- Psiko - Motor, 124, 161
 İlişki (istatistik), 215
 Momementler çarpımı
 korelasyon katsayısı, 215
 Sıra korelasyonu, 218
 İmtihan (Bkz : Sınav)
 İpuçları (testlerde)
 Geçersiz, 104
 Kısa cevaplı maddeler, 68
 Eksik cümle formu, 68
 Puanlanması, 74
 Soru formu, 68
 Yazılması, 70
- Kompozisyonlar, 64
 Kopya, 44
 Korelasyon katsayısı, 215
 Çift serili, 270, 284
 Momentler çarpımı, 217
 Sıra farkı, 218
 Spearman, 218
- Kovaryans, 216
 Kök (Bkz : Test maddeleri)
 K - R formülleri, 34, 268
 Madde analizi, 261
 Basit bir analiz, 262
 Düzeltme, 272
 İşlemler, 261
 Seçeneklerin analizi, 264
 Tablosu, 286
 Test düzenleme, 256
 Yöntemler, 261
- Madde düzeltme (Bkz : Madde analizi)
- Madde istatistikleri, 261
 Ayırıcılık indisi, 263
 Güçlük indisi, 263
 Test istatistikleri, 266
 Ortalama güçlük, 267
 Standart kayma, 268
 Test geçerliği, 270
 Test güvenirliliği, 268
- Medyan (Bkz : Ortanca)
 Mutlak değerlendirme, 231, 238
 Norm dayanaklı yorum, 213, 228
 Not verme, 220
 Ağırlıklı, 245
 Basit, 245
 Bir tek ölçmeyle, 235
 Birkaç ölçmeyle, 244

- İstatistiksel yöntemleri, 233
Mutlak başarı oranlarıyla, 236
Notların
 Gerekliliği, 220
 Güdüleyiciliği, 221
 Güvenirliliği, 234
 Kullanılması, 249
 Ölçekleri, 234
 Ölçütleri, 227
 Yordayıcılığı, 257
 Rehberlikteki rolü, 223
Sınıftaki puan dağılımıyla, 237
Objektiflik, 98
Ortalama, 201
 Ağırlıklı, 201
 Arıtmelik, 199
 Başarı düzeyi, 240
Ortalamadan sapmalar, 205
Ortanca, 203
 Hesabı, 203
Ödevler, 179
 Ev ödevleri, 179
 Grup projeleri, 180
Öğrenci akışı, 249
Ölçme, 7
 Anlamı, 8
 Araçların rolü, 21
 Birimleri, 12
 Dolaylı, 22
 Formelleştirilmesi, 10
 Geçerliliği, 38
 Göstergöyle, 23
 Güvenirliliği, 30
 Hataları, 27
 Sabit, 28
 Sistematik, 28
 Standart, 35
 Tesadüfi, 28
 Homomorf, 14
 İşlemleri, 21
 İzomorf, 14
 Önemi, 7
 Ölçekleri, 15
 Sıfır noktası, 14
 Tanımı, 11
 Temel, 22
 Türetilmiş, 23
 Türleri, 21
Ölçme araçlarında kullanışlılık, 44
 Cevapların nitelikleri, 45
 Cevaplayıcıların nitelikleri, 45
 Ekonomi, 44
 Hazırlama süresi, 44
 Hazırlayıcı nitelikleri, 45
 Puanlama kolaylıkları, 45
 Uygulama kolaylıkları, 45
 Uygulayıcı nitelikleri, 45
 Yorumlama kolaylıkları, 46
Psiko - Motor
 Becerilerin ölçülmesi, 161
 Çetele araçları, 166
 Dereceleme araçları, 167
 Dikkat edilecek hususlar, 167
 Dolaysız gözlemle ölçme, 162
 Beceri ürünlerinin ölçülmesi, 167
 Davranışlar, 123
Puan
 Ham, 21
 Karşılaştırılması, 214
 Standart, 212
 Yüzdelik, 214
 Z - puanları, 213
Puanlama, 59, 74, 107, 190
 Anahtarı, 60, 61, 66, 169
 Analitik, 60, 169
 Hatalarını önleme, 63
 Genel izlenimle, 59
 Sınıflamayla, 59
 Sıralamayla, 60
Ranj (Bkz : Dizi genişliği)
Seçenekler, 87
 Çeldirici olarak, 87
 Doğru cevap olarak, 87
 Seçmeli maddelerde, 87
 Hepsi, 105
 Hiçbiri, 105
 Yazılması, 103
Seçmeli testler, 86
 Başlıca madde türleri, 87
 Doğru cevabı birden fazla, 89
 Bileşik cevaplı, 91
 Bir tek cevabı doğru, 87
 Doğru cevabı gizli, 91
 Doğru cevabı kesin, 87
 En doğru cevaplı, 88
 Kökü eksik cümle, 93
 Kökü olumsuz, 93
 Kökü soru kipinde, 93
 Ortak köklü, 95

Ortak seçenekli, 96
Sınav, 171
Açık kitap, 174
Ara, 177
Bitirme, 179
Çeşitleri, 171
Dışarıda cevaplanan, 175
Dönem sonu, 178
Kısa süreli, 176
Planları, 181
Seçme, 172
Sınıflama, 173
Sözlü, 109
Tarama, 173
Yarışma, 172
Yeterlik, 172
Sınıf geçme düzeni, 251
Sınıflama gerektiren maddeler, 75
Doğru - Yanlış, 75
Diğer türleri, 83
Sıra farkı korelasyonu, 218
Sıralama, 60
Spearman - Brown formülü, 33, 269
Spearman sıra farkı kor., 218
Sözde eş formlar, 187
Sözlü sınavlar, 109
Diğer çeşitleri, 114
Kullanılacağı durumlar, 110
Nasıl yapılmalı, 112
Özellikleri, 109
Standart - 9 ölçeği, 240
Standart hata, 35
Standart kayma, 204
Şans başarısı (seçme testlerde), 191
Düzeltilmesi, 192
Ortalamanın düzeltilmesi, 242
Şıklar (Bkz : Seçenekler)
Tanıma ve hatırlama, 128
Test
Açıklamaları, 185
Analizi, 286
Çoğaltılması, 188

Düzeni, 73, 85, 185
Sayfa düzeni, 186
Sözde eş formlar, 187
Gizliliği, 188, 190
Madde türleri
Doğru - Yanlış, 75
Kısa cevaplı, 88
Seçmeli, 86
Sınıflama gerektiren, 75
Planı, 181
Puanlanması, 190
Süresi, 189
Uygulanması, 188
Uzunluğu, 183

Tutum

Duyuşsal davranış olarak, 151, 154
Ölçme araçları, 153, 154

Uygulama (Bkz : Test)
Üst grup (Bkz : Madde analizi)

Yazılı: yoklama, 47

Diğer çeşitleri, 65
Geçerliliği, 51
Güvenirliliği, 49
Kompozisyon, 64
Kullanışlılığı, 52
Nasıl yapılmalı, 54
Özellikleri, 48
Puanlanması, 59

Anahtarlar, 60
Genel izlenimle, 59
Sınıflamayla, 59
Sıralamayla, 50

Yordayıcılığı (notların), 257

Yüzde (İstatistik)

Yıgmalı, 243

Yüzdelemlerle mutlak değerlendirme,
236

Yüzdelik puanlar, 214

Varians, 205

puanı, 213

M. FUAT TURGUT 1926'da Aydın'da doğmuş, ilk ve orta öğrenimini orada tamamlamış; Balıkesir Necatibey Öğretmen Okulunu ve Necatibey Eğitim Enstitüsünü bitirdikten sonra, orta öğretim okullarında öğretmenlik ve yöneticilik yapmıştır. Daha sonra öğrenimine Amerika Birleşik Devletlerinde devam eden M. Fuat Turgut Columbia Üniversitesinden psikolojide Lisans, ölçme ve değerlendirmede uzmanlık, psikometride doktora derecelerini almıştır.

Amerika Birleşik Devletlerinde çeşitli araştırma ve öğretim görevlerinde bir süre çalışan Turgut 1962'de yurda dönerek Milli Eğitim Bakanlığı Test ve Araştırma Bürosunda görev almış; daha sonra Orta Doğu Teknik Üniversitesinde ve Hacettepe Üniversitesinde öğretim üyeliği yapmıştır.

Yazarın eğitimde ölçme ve değerlendirme alanında birçok yayını, psikometride ve fen eğitiminde çeşitli araştırmaları vardır.

BU KİTAP öğrenci başarısını değerlendirmede bir öğretmen için gerekli olan teknikleri, her derecedeki öğretmen ve öğretmen adaylarına kazandırmak amacıyla yazılmıştır. Kitap, ölçme ve değerlendirme alanının temel bilgilerini ve sınıftaki değerlendirme etkinliklerinde kullanılan teknikleri kapsar. Sınav ve testlerin planlanmasında, ölçme araçlarının hazırlanmasında, uygulama ve puanlamada, puanlar üzerindeki istatistiksel işlemlerde, ölçme sonuçlarının yorumlanmasında ve not vermede öğretmenin karşılaşabileceği teknik soruların hemen hemen hepsine bu kitapta örnekleriyle ve açıklamalarıyla yer verilmiştir.

İSTEME ADRESİ :

M. Fuat TURGUT
Esat Cad. 55 / 9
KÜÇÜKESAT - ANKARA

ISBN-975-95915-0-2

GENEL DAĞITIM :

GÜL YAYINEVİ
Ş. Adem Yavuz Sk. 11 / B
KIZILAY - ANKARA
Tel :: (0 - 312) 418 75 41