

Tıp Eğitiminde Ölçme ve Değerlendirme

H.Çetin EKERBİÇER¹, Ahmet SALTİK²

ÖZET

Ölçme ve değerlendirme teknikleri eğitim bilimin dinamik tartışma alanlarındandır. En azından eğitim yöntembiliminde olduğu kadar, eğitimde ölçme ve değerlendirme konuları da sürekli olarak gündemde kalmaktadır. Performansı yetenek ve başarı testleri ile ölçmek mümkündür. Başarı testleri arasında "yazılı", "kısa yanıtlı", "tamamlamalı", "esleştirmeli", "doğru-yanlış", "çoktan seçmeli" soru türlerinden söz edilebilir. Testlerin hazırlama ve değerlendirme aşamalarında gözden kaçmaması gereken önemli noktalar vardır. Değerlendirme sırasında seçenek (item) analizi yaparak soruların doğru yanıtlanma oranları ve yanıtlayıcıların seçeneklere dağılımı elde edilebilir. Bu da, soruların kalitesi, öğrencilerin düzeyi ve olası kopya olayları ile ilgili önemli ipuçları verir. Test puanlarının değerlendirilmesinde ham puanları kullanmak yerine standart puanlara dönüşüm uygulanır. Genelde z ve T puanı olmak üzere iki standart puan kullanılır. Ancak, farklı ortalama ve standart sapmalara göre çok çeşitli puanlar elde edilebilir.

Anahtar sözcükler: Ölçme-değerlendirme, tıp eğitimi, sınav, standart puan, item analizi

SUMMARY

MEASURING AND EVALUATING MEDICAL EDUCATION

Measurement technics and evaluation are the major topics of the educational sciences always take place in the agenda. These topics are always actual as well as methodology of education. In the context of measuring medical education "written", "short answered", "filling in the blanks", "matching", "true/false" type questions are some of the question types used for the success tests. There are a number of topics that must be taken into consideration in the process of preparing and evaluating the tests. Answerers' distribution to the items of the question and the proportion of true items can be calculated while evaluating the tests, by making item analysis. This helps in judging the quality of the questions, the global level of the students and supplies some cues on probable cheatings. While evaluating the test points, crude points are transformed into standard points generally using z and T standard points. Medical teachers at all level should have a basic knowledge and skill on measuring and evaluating medical education.

Keywords: Measuring & Evaluation, medical education, exams, standard points, item analysis

Okullarda başarının değerlendirilmesi, 19. yüzyılın ortalarına kadar "sözlü" sınavlarla yapılmaktaydı. Sözlü sınavların yanına yazılı sınavların da eklenmesi çabaları, 1845'lerde H. Mann'ın (1a) eleştirileri ile başlamıştır. Mann, yazılı sınavların öğrencilerin hepsi için standart bir durum yarattığını, daha fazla sayıda soru sorulabileceğini, şans hatasını azalttığını ve sözlü sınavlarda bireyler için soru seçiminde

takınılabilecek öznel tutumları önlemesi gibi etmenlerden dolayı daha avantajlı olacağını belirtiyordu. Bu etmenler, aynı zamanda daha geniş ölçüde "nesnel teknikleri" yeglemenin nedenlerini de o dönemde ortaya koymuş oluyordu.

Nesnel (objektif) testlerin bir başlangıcı olarak, Fisher'in (1a) kalitatif nitelikleri ölçümede, bugünkü derecelendirme ölçüklerine benzer, 5 dereceli örneklerden oluşan ve bireyin yanıtlarını bu

¹ Araş. Gör. Dr., Trakya Üniversitesi Tıp Fakültesi Halk Sağlığı Anabilim Dalı, EDİRNE

² Doç. Dr., Trakya Üniversitesi Tıp Fakültesi Halk Sağlığı Anabilim Dalı, EDİRNE

yakınılığa göre sınıflandırarak değerlendirme esasına dayanan "ölçek" yaklaşımı ile ilk olarak dilbilgisi ile ilgili nesnel bir test hazırlayan Rice'ın (1a) (1894) çabaları sayılabilir.

Başarı testlerinin gelişmesine ise en büyük katkı Thorndike (1a) tarafından yapılmıştır. Thorndike, başarı testlerinin hazırlanması, geliştirilmesi ve bu amaçla kullanılan istatistiksel yöntemler yönünden günümüzde kullanılan başarı testlerinin öncülüğünü yapmış, çeşitli başarı testleri ve ölçükler hazırlamıştır. Testlerin yaygın olarak kullanılması O'nun çabaları ile olmuştur.

Başarı testleri, eğitim hedeflerini ölçmek, okullarda başarıyı değerlendirmek için olduğu kadar, sanayi ve kamu kesiminde personel seçimi ile giriş sınavlarında da yaygın olarak kullanılmaktadır. Test hazırlama yöntemi ve testlerde kullanılan sorular bakımından basit "kavrayış" ve "sorun çözme" yi ölçen sorulara doğru bir gelişme olmuştur (1a).

Başarı Testleri: Maksimum performans testleri; "yetenek" ve "başarı" testleri olmak üzere ikiye ayrılmaktadır. Başarı testleri, belirli bir ders ve programdan bireylerin ne derecede yararlandıklarını göstermesi bakımından yetenek testlerinden ayrılır. Çünkü yetenek testlerinin, belirli bir öğrenimi gerektirmeyen çok çeşitli zihin becerilerine ilişkin sorulardan hazırlanması istenmektedir. Günümüzde başarı testleri, bireylerin akademik özgeçmişlerinden ne derecede yararlandıklarını göstermesi bakımından başarıların değerlendirilmesi amacı ile kullanıldığı gibi, gelecekteki başarılarını yordama amacı ile de kullanılmaktadır. Buradaki temel sayılı, bireylerin geçmişteki yaşıtlarından yararlanmış olmalarının, gelecekteki akademik yaşıtlarından da yararlanacaklarına ilişkin ipuçları vermesine dayanmaktadır (1b).

Başarı testleri ile yetenek testlerini kullanış amaçları yönünden birbirinden kesin olarak ayırmak oldukça güçtür. Her ikisi de bireyin geleceğine yönelik karar alırken "yordayıcı" olarak kullanılmaktadır. Gerçi başarı testleri genellikle bireyin belli bir öğretimden sonra ne öğrendiğini, geçmiş yaşıtlarından ne derecede yararlandığını ölçmek için kullanılmaktansa da, yaygın olarak geleceği yordamak amacıyla kullanılmaktadır (1b).

Başarı testleri, yetenek, kişilik, ilgi ve tutum gibi çeşitli alanlara ilişkin psikolojik testler arasında en yaygın olarak kullanılan testlerdir. Çok çeşitli eğitim programlarında okutulan ders konularına ilişkin olarak eğiticiler tarafından hazırlanan başarı testleri ile uzmanlar tarafından hazırlanan standart başarı testleri birlikte düşünüldüğünde; diğer alanlardaki test sayısını

aşabilecegi kolayca kestirilebilir. Bu konuda Buros (1b) (1978) Sekizinci Zihinsel Ölçüm Yıllığı'nda (Mental Measurement Yearbook) yalnızca standart başarı testleri olarak çeşitli konularda hazırlanmış 350 dolayında test olduğunu belirtmektedir.

Başarı testleri; genel yetenek, özel yetenek ve kişilik testlerinden farklı olarak, ölçülmesi amaçlanan niteliği doğrudan ölçmektedir. Oysa zeka testleri ve özel yetenek testleri, dolaylı olarak bireyin performansına göre yordanan ölçülerdir (1b).

Sınavlarda Sorulacak Soruların Seçimi

Sınavlar, bilinen eğitim süreçlerinin ardından, eğitilenlerin kazanmaları beklenen bilgi, beceri ve uygulamaların ne ölçüde gerçekleştirilebildiğini ölçme amacıyla düzenlenen *tamamlayıcı eğitim süreçler* dir.

Avrupa Tıp Dekanları Birliği

Yani sınavın kendisi de doğrudan bir eğitim süreci olmak durumundadır; bir bütün olarak eğitim sürecinin tamamlayıcı temel ögesidir.

Ölçme ve değerlendirme teknikleri eğitim bilimin dinamik tartışma alanlarındandır. En azından eğitim yöntem-biliminde olduğu kadar, *eğitimde ölçme ve değerlendirme* konuları da sürekli olarak gündemde kalmaktadır (2).

Eğitim; kişi, grup ya da daha büyük toplulukların seçilmiş konularda bilgilendirilmeleri, yeni ve *istendik davranışlar geliştirmeleri*, yeni uygulama becerileri kazanmaları, yanlış ya da istenmedik davranışların doğruları ile yer değiştirmesi amaçları ile yürütülen bir süreçler bütünü olarak tanımlanabilir. Bu süreçte "ürün"ün (süje'nin) değerlendirilmesi doğal olarak kaçınılmazdır. Bunun yolu da değişik tekniklere dayalı olarak yürütülecek ölçme-değerlendirme etkinliklerinden, bir anlamda sınavlardan geçmektedir. Sınav tekniklerinin, ölçülecek eğitim alanına uyumlu olarak belirlenmesi gerekmektedir. Yazılı, sözlü, uygulamalı, kuramsal ya da bunların değişik bileşimleri biçiminde sınavlar yapılabilmektedir. Etkin bir sınavın ya da ölçme-değerlendirme sürecinin, hem bilişsel (cognitive) hem de uygulama (psychomotor) alanlarına ilişkin olarak "süje"nin (eğitilenin) bilgi birikimini, sorun çözme dahil olmak üzere uygulama becerilerini ortaya koyması beklenir. Bu amaçlarla sınavlarda 3 soru türünün dengeli olarak sorulması arzulanır. Bunlar aşağıda verilmektedir (2,3):

1. Belleğe yığılı bilgiyi sorgulayan sorular (recall of facts)

2. Kazanılmış bilginin yorumunu ölçen sorular (*interpretation of data*)
3. Kazanılmış bilgi ile sorun çözüme becerisini ölçen sorular (*problem solving*)

Başarı Testlerinde Kullanılan Soru Türleri: Başarı testlerinde, eğitim hedeflerini ya da bireyin niteliklerini ölçmek amacıyla çok çeşitli soru türlerinden yararlanılmaktadır. Bunlar arasında "yazılı", "kısa yanıtlı", "tamamlamalı", "eşleştirmeli", "doğru-yanlış", "çoktan seçmeli" soru türlerinden söz edilebilir. Bu soru türlerinin ilk üçünde, sorular "açık uçlu" olarak verilmekte, beklenen yanıt yönünden kişi "bağımsız" bulunmaktadır. Soruya verilecek yanıt bir "sözcük olabileceği gibi (tamamlamalı)", "birkaç sayfa" (yazılı) da olabilir. Bu tür sorulara "yazmalı soru" türleri denmektedir. Bu sorularda hazır yanıt seçenekleri verilmemekte, sorunun yanıtını kişinin kendisi bulmaktadır. Bu gruptakilerden en kısa yanılı olan tamamlamalı soruda bile yanıt olarak "anlamlı bir sözcük" bulunması istenmektedir (1b).

Buna karşın son üç soru türünde, diğer deyimle eşleştirmeli, doğru-yanlış ve seçmeli soru türlerinde ise 1 soru sorulmakta ve bu sorunun olası yanıtları da verilmektedir. Testi alan birey verilen "seçenekler" arasında "doğru yanıtı" bulmakta, bir sembol (genellikle a,b,c,d,e) kullanarak işaretlemektedir. Soruya yazılı yanıt bulmak söz konusu olmamakta, verilen belli sayıda (genelde 35 tane) seçenekler içinden bir tanesi seçilmektedir. Bu tür sorulara "işaretlemeli soru" türleri denilmektedir (1b).

Yazmalı Soru Türlerinin Özellikleri: "Yazmalı" soru türlerinin en önemli avantajı, "işaretlemeli" sorularla ölçülemeyen bir düşünceyi organize etmek, ilişkileri belirtmek, açıklamak gibi nitelikleri ölçebilmesidir. Yazmalı soru türlerini hazırlamak oldukça kolaydır. Yanıtlamada raslantısal kestirimlere göre doğru yanıt verme olasılığı düşüktür. Yanıtlama süresi çok zaman aldığı için 50 dakikalık bir sınav süresinde 56 soru sorulabilir. Az sayıda soru sorulduğu için yazmalı sınavlar öğrencilerin bilgi, beceri, yetenek ve öğrenilmesi gereken konuları örneklemede yetersiz kalmaktadır. Yazılı sınavların olumsuz yönlerinden biri de sınav veya testin ölçmeyi amaçladığı nitelikten ayrı olarak; yazının düzeni, yazının okunaklı ve güzel olması v.b. niteliklerin bireyin alacağı not veya puanı etkilemesidir. Ayrıca, yazılı sınavların puana dönüştürülmesi "öznel" olduğu kadar, zaman yönünden de ekonomik değildir. Aynı yazılı yanıtları okuyan kişiler arasında, matematik dersinde bile değerlendirmeler % 27 ile 71 arasında değişkenlik gösterebilmektedir (1b). Yazmalı sınavların puanlaması, puanlayıcının kişisel

ölçütlerine göre değiştiği için bu tür sınavlara "öznel sınavlar" denmektedir.

Bir kural olarak; "Eğer bir nitelik testlerle ölçülebiliyorsa, yazılı sınavlar yeğlenmemelidir"

Eğer yazmalı sınavlar kullanılacaksa yazılı soruları hazırlayan kişi soruları olabildiğince "nesnel" olarak hazırlamalıdır. Bunun için, aşağıdaki kurallara özen gösterilmelidir (1b):

1. Soruları açık-seçik yazmalı, sınırlamalı, "tartışınız" yerine "açıklayınız", "karşılaştırunız" gibi anlatımları yeğlemelidir.
2. Sınav'a girenlerin hepsinin ulaşabileceği ölçüde az sayıda soru sormalı ya da zamanı iyi ayarlamalıdır.
3. Sorulara verilen yanıtlar puanlanırken "biri diğerinden daha iyi olacak biçimde" sınıflandırılmalı, "kağıt okuma" yerine, tüm öğrenciler için "soru okuma" yöntemi uygulanmalı ve puanlama buna göre yapılmalıdır.

İşaretlemeli Soru Türlerinin Özellikleri: "İşaretlemeli" soru grubu içinde en yaygın olarak kullanılan sorular "doğru-yanlış", "çoktan seçmeli" ve "eşleştirmeli" sorulardır. Bu sorulara "nesnel test soruları" da denmektedir.

İşaretlemeli soruları hazırlamak, yazmalı sorulara göre daha çok zaman alır ve bazı beceriler gerektirir. Ancak, avantajlı bir yönü olarak, testler kısa zamanda uygulanabilir ve puanlaması çok kolay ve nesneldir. Yanıtlanması az zaman aldığı için testlerde kısa sürede çok soru yanıtlanabilmektedir. Belirli bir zamanda çok sayıda soru sorulabilmesi, ölçülmesi istenen nitelikleri daha iyi örnekleme olanağı vermekte, az sayıda sorunun getirdiği şans hatası azalmaktadır.

Test hazırlamada sorular açık-seçik, yalın ve dilbilgisi kurallarına uygun olmalı, sözcüklerin seçimi testi alacak kişilerin düzeyine uygun olmalıdır. Soru yapısını oluşturan anlatımlar kitaplardan aynen alınmamalı, yeni bir sentez olmalıdır. Sorularda anlamayı zorlaştıran gereksiz bilgiler ve sözcükler bulunmamalı, her soru bağımsız olmalı ve bir soru diğer bir sorunun yanıtında ipuçları içermemelidir.

Aşağıda "test" (işaretlemeli) soruları hakkında kısa bilgiler verilmektedir.

Doğru-Yanlış Tipi Sorular: Doğru-yanlış türündeki sorular, hazırlanması ve okunması kolay olduğundan, çok sayıda soru sorabileme olanağı verirler. Böylece ölçülmesi istenen nitelikleri daha iyi/kapsamlı olarak örnekleme olanağı artmaktadır. Doğru-yanlış soruları, 2 seçenekten birini seçmeye dayandığı için % 50 şansa yer verir. Bu nedenle,

eğer bir test yalnızca doğru-yanlış sorularından oluşuyorsa, soru sayısının 50'den az olmamalıdır. Soruların yazımında, doğruluğu ve yanlışlığı gösteren ipuçlarının bulunması, şans hatasını daha da artıracaktır.

Araştırmalar, bu tür sorularda “hepsi”, “daima” ve “asla” gibi sözcüklerin anlatumın yanlışlığını; “genellikle”, “bazen”, “sık sık” gibi sözcüklerin ise doğruluğu çağrıştıran ipuçları verdiğiğini göstermiştir. Soru yazımının kitaplardan aynen alınması da yanlış uygulamalardan biridir.

“Doğru-yanlış” soruları hazırlarken dikkat edilmesi gereken birkaç nokta aşağıdadır: (1b)

1. Doğru ve yanlış anlatımların ikisi de kısa ve özenle hazırlanmalıdır.
2. Olumsuz, karmaşık ve çelişkili anlatımlardan kaçınılmalıdır.
3. Görüş ve kanı bildiren anlatımlarda kaynak belirtilmelidir.
4. Doğru ve yanlış anlatımların uzunluğu yaklaşık olarak aynı olmalıdır.
5. Toplam soru sayısı içinde doğru ve yanlış anlatımların sayısı eşit olmalıdır.

Aşağıda “doğru-yanlış” türü soruya bir örnek verilmiştir (4) :

Açıklama: Aşağıdaki anlatımları doğru (D) veya yanlış (Y) olarak işaretleyiniz.

- () 150 Başlıklı A Listesi, Hastalıklar ve Travmalar ile onların sonuçlarını kodlar.
- () Ölüm nedenleri, Devlet İstatistik Enstitüsü'nce 50 Başlıklı B Listesi'ne göre kodlanır.
- () Hastalık tanı kodları, Sağlık Ocaklarında Kişisel Sağlık Fişleri'ne işlenmektedir.
- () Hastanelerde, her ay Form 53/A doldurularak Sağlık Müdürlüğü'ne yollanmaktadır.
- () Hastanelerimizde 999 Başlıklı Liste kullanılmaktadır.

Eşleştirme Türü Sorular: Eşleştirme sorusu bir soru olmaktan çok bir soru grubudur. Eşleştirme soruları yapısal olarak bir “soru” diğeri “yanıt” olmak üzere iki uyarıcılıpki grubundan oluşmaktadır. Soru grubu 5-15 maddeden, yanıt grubu ise soru grubunun sayısından 3-4 madde daha fazla olmalıdır. Eşleştirme sorusunu yanıtlayan kişiden, “soru” grubundaki maddelerle ilgili olanı “yanıt” sütunundan bulup eşlestirmesi istenmektedir.

Eşleştirme sorularının soru ve yanıt maddelerinin “türdeş” (homojen) olması, aynı konudan ya da birbirine yakın konulardan

hazırlanması gereklidir. Eşleştirme sorusu hazırlamak oldukça kolaydır. Ancak bu sorular “tanıma ve belleme” gibi zihin süreçlerini ölçebilmektedir. Eşlestirirken soruları hazırlarken dikkat edilecek bazı noktalar aşağıda verilmiştir (1b):

1. Soru sütunu solda, yanıt sütunu sağda olmalıdır.
2. Soru sütunu 6-15 kadar olmalı, yanıt sütunu ise soru sayısından 3 fazla olmalıdır.
3. Eşlestirmeye esas olan bir açıklama verilmelidir.
4. Eşlestirilecek soru ve yanıtlar bölümmeden aynı sayfada olmalıdır.

Aşağıda, eşleştirme sorusuna bir örnek verilmektedir (4) :

SORU: Aşağıdaki 2 grup kavramı uygun olarak eşleştiriniz.

- | | |
|---------------------------------|-------------------------------|
| A. 0.1 N HCl | P. Hemoglobin ölçüm aygıtı |
| B. % 20'lük triklor asetik asit | Q. Gram boyama maddesi |
| C. Rosin ayıracı | X. Ürobilinojen arama ayıracı |
| D. Sahle seti | W. Ürobilin ayıracı |
| E. Kristal viyole | Y. İdrarda protein ayıracı |
| F. Schlesinger ayıracı | U. Hemoglobin ayıracı |
| G. Ehrlich ayıracı | Z. İdrarda bilirubin ayıracı |

Çoktan Seçmeli Sorular: Çoktan seçmeli sorular, bir “soru” ve soruya yanıt olabilecek olası “seçenek”lerden oluşur. Çoğunlukla seçeneklerden biri “doğru” diğerleri “yanlış” olarak düzenlenir. En yaygın kullanılan soru türlerindendir. Yalın veya oldukça karmaşık nitelikleri ölçüme olanağı verir. Puanlar şans hatasından daha az etkilenmektedir. Soruların seçenekleri testin uygulanacağı bireylerin eğitim düzeylerine göre 3 ile 5 seçenek arasında değişmektedir. 5 seçenekli bir soruda şansa dayalı olarak sorunun doğru yanıtlanması olasılığı % 20 kadardır. Ancak bu şansın ardışık olarak diğer sorularda da yakalanması olasılığı giderek düşmektedir. Örn. on sorunun birden şansa bağlı olarak doğru yanıtlanması olasılığı $(1/5)^{10}$ düzeyindedir.

Çoktan seçmeli sorular hemen her konuya uygun düberler. Standart Başarı Testleri'nde yaygın olarak çoktan seçmeli soru türü kullanılmaktır ve bilgisayar ile puanlama (optik okuyucu) yapılmaktadır. Sonuçlar üzerinde madde (item) analizi yapınak ve seçeneklerin işlerliğini görmek olanaklıdır. Soruya verilen olası seçeneklerin hepsinin, konuyu iyi bilmeyen birisine doğru gibi görünmesi gereklidir. Hazırlanmaları oldukça zordur.

Çoktan seçmeli soruların hazırlanmasında dikkat edilmesi gereken bazı noktalar aşağıda özel olarak verilmiştir (1b):

Soru "kökü"nde yanıtlanacak soru, bir soru tümcesi ile açık-seçik anlatılmış olmalıdır.

Seçeneklerde aynı sözcüklerin yinelenmemesi ve seçeneklerin kısa olması için anlatımın çoğu olabildiğince soru kökünde verilmelidir.

Seçenek sayısı 4-5 olmalıdır ancak, konunun özelliğine göre 2-3 seçenekli sorular da kullanılabilir.

Seçeneklerin hepsi yaklaşık aynı uzunlukta, dil kurallarına ve soru köküne uygun olmalıdır.

Seçeneklerin yaş, sayı, tercih gibi doğal bir özelliği varsa, küçüllererek veya büyütüerek sıralı olarak verilmeli, bunun dışındaki rasgele veya anlatımın ilk harfine göre sıralanmalıdır.

Seçenekler özenle hazırlanmalı, doğru yanıtı bilmeyenlere hepsi doğru gibi görünümlü, fakat yalnızca 1 tanesi doğru olmalıdır.

Aşağıda çoktan seçmeli soru tipine birkaç örnek verilmektedir (4):

SORU: Sağlık Ocağı işlevler piramidinde yer alan işler, hasta sağaltımı ve poliklinik hizmetleri bir yana bırakılırsa, tabandan tepeye doğru hangi seçenekte DOĞRU sıralanmaktadır?

- Anaçocuk sağlığı hizmetleri + çevre sağlığı hizmetleri + bağılıklıma hizmetleri + sağlık eğitimi + taramalar + yönetim işleri
- Anaçocuk sağlığı hizmetleri + sağlık eğitimi + çevre sağlığı hizmetleri + bağılıklıma hizmetleri + taramalar + yönetimsel işler
- Sağlık eğitimi + ana çocuk sağlığı hizmetleri + bağılıklıma hizmetleri + çevre sağlığı hizmetleri + taramalar + yönetim işleri
- Sağlık eğitimi + çevre sağlığı hizmetleri + taramalar anaçocuk sağlığı hizmetleri + bağılıklıma hizmetleri + yönetimsel işler
- Sağlık eğitimi + taramalar + çevre sağlığı hizmetleri + bağılıklıma hizmetleri + anaçocuk sağlığı hizmetleri + yönetim işleri

SORU: Ülkemizin 5 önemli beslenme sorunu hangi seçenektedir?

- PKM/PEM /anemi/iyot eksikliği/quatr/obesite
- PKM/anemi/iyot eksikliği/quatr/vitamin eksiklikleri/ şişmanlık
- PEM/anemi/quatr/vitamin eksiklikleri/malnürisyon
- PKM/anemi/quatr/vitamin eksiklikleri/% 20 malnürisyon
- Anemi/quatr/malnürisyon/PKM/şişmanlık

Seçmeli Test Sorularının Puanlanması: Test sorusunun yapısı, testlerin nasıl puanlanacağını belirler. Seçmeli test sorularının puanlanmasıında şu üç yöntemden biri kullanılabilir:

- Soru doğru yanıtlanmışsa 1, yanlış yanıtlanmışsa veya boş bırakılmışsa sıfır puan vermek.
- Her soruya, ölçüdüğü yeteneğin önemiyle orantılı bir puan vermek. Örn. kolay veya kapsamlı dar

sorulara 1, bu açıdan orta nitelikteki sorulara 2, çok zor veya kapsamlı geniş sorulara üç puan verilebilir.

- Sorunun seçeneklerine, doğruluk derecesine göre farklı ağırlıklar vererek puanlamak.

Bu yöntemlerden ilki en çok kullanılan puanlama yöntemi olup (ağırlıksız puanlama); kolaylığı, hızla yapılabilmesi, test ve soru analizinin gelenekselleşmiş formüllerine uygun olması nedeniyle yeğlenir. İkinci yöntem, 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 ağırlık vermek güç, hem de farklı ağırlıklarla puanlama zor olduğundan, bu yöntem kullanışsızdır (5).

Bir tek sorunun seçeneklerine farklı ağırlıklar veren puanlama yöntemi de denenmiştir. Bazı yazarlar bu yöntemin, doğru yanıtın en doğru olduğu, çeldiricilerin de gittikçe artan doğruluk derecelerinde olduğu sorularda kullanılabileceğini belirtmiştir (5,6). Böyle bir puanlamada, doğru yanıta en yakın çeldiriciyi seçen yanıtlayıcının yeteneğinin, daha zayıf bir çeldiriciyi seçen yanıtlayıcının yeteneğinden daha fazla olduğu kabul edilir (5).

Testin Düzenlenmesi: Yeterince test sorusu yazıldıktan sonra, soruların düzene konularak "testin" hazırlanması gereklidir. Testin hazırlanması sırasında testin "açıklaması", "uzunluğu", "soruların düzeni", "yanıt kağıdı" gibi noktaların da düşünülmesi gereklidir.

Testin Açıklanması: Testin "açıklama"sı ön kapak sayfasında yer alır, ön kapağın yetmemesi durumunda "arka kapak"ta devam eder. Amaç, gerekli açıklamalar yapılp bitirilmeden test alan kişinin sorularla yüz yüze gelmemesidir. Açıklamada testin amacı, sorulara nasıl yanıt verileceği, soru sayısı, testin süresi, yanıtlanmadan dikkat edilecek noktalar, boş bırakılmış bırakılmayacağı, şans hatasını düzeltmek için bir işlem yapılp yapılmayacağı, zamanın iyi kullanılması v.b. gibi noktalar yer alır. Açıklama açık bir dille yapılmalı ve yalın olmalıdır. Gereken durumlarda soruların nasıl yanıtlanacağına ilişkin "örnek sorular"da konabilir. Yanıt kağıdı kullanılabaksa yanıtların yanıt kağıdına nasıl işaretleneceği belirtilir (1b).

Testin Uzunluğu: Testin uzunluğu, testin uygulanacağı kişilerin sınıf ve gelişmişlik düzeyi ile testin uygulanacağı zaman dilimi ile yakından ilgilidir. Testteki soruların uzunluksalığı ve soruların türü de testin uzunluğunu etkiler. Her soru

Tablo I. Seçenek Analizi Yanıt Dağılımı Tablosu.

Test.....	Tarih.....						
Form.....	Çeteleyen.....						
Grup.....	Kontrol Eden.....						
SORU	A	B	C	D	E	BOŞ	TOPLAM
1							
2							
3							
4							
5							
Toplam							

bir birim olarak düşünüldüğünde, ortaokul ve daha yukarı düzeyde, çoktan seçmeli sorular için 1 dakika, doğru-yanlış sorular için 30 saniye süre hesaplanması oldukça yaygındır. Normal bir ders süresinde 50 soruluk çoktan seçmeli veya 100 soruluk bir doğru-yanlış testi uygun olabilir. Bir güç testinde, test uygulanan grubun % 90'ının testi tamamladığı zaman, test için uygun bir süre olarak düşünülebilir. Bu kurallar ilkokul düzeyinde daha toleranslı olarak uygulanmalıdır (1b).

Soruların Düzenlenmesi: Soruların genel düzenlemesi, soru türlerine göre, altölçeklere göre veya soruların güçlük derecelerine göre yapılabilir. "Soru türlerine" göre yapılan düzenlemeye aynı tür sorular bir araya konur. Genellikle doğru-yanlış, çoktan seçmeli ve eşleştirmeli sorular olmak üzere bölümler halinde düzenlenir. Eğer testin altölçekleri varsa ve ayrı puanlar elde edilecekse "altölçeklere" ilişkin sorular ayrı alttestler halinde düzenlenir. Bazen sorular rasgele olarak dağıtılp, her altölçek için ilgili soruları içeren ayrı yanıt anahtarları hazırlanabilir. Düzenleme "soruların güçlük derecelerine" göre yapılacaksa, kolay sorular en başta olmak üzere sıralanır. Sorular hangi biçimde düzenlenirse düzenlensin, testin en başına, test alan kişiyi motive etmek için *birkaç kolay soru* konmalıdır.

Testin düzenlenmesinde soruların "doğru yanıtları" da dikkate alınmalı, her seçenek'e düşen doğru yanıt sayısı yaklaşık aynı olmalı ya da eşit tutulmalıdır. Doğru yanıtın seçeneklere dağıtımasında izlenebilecek bir diğer yöntem rasgele dağıtmadır.

Sorular sayfalara yerleştirilirken bölünmemeli, aynı sayfada olmalıdır. Soru kökü sık satır aralıklarıyla yazılmalı; soru ile seçenekler arasında birbirince satır aralık bırakılmalıdır. Seçeneklerin sık satır aralıklı ve soruların iki satır aralıklı yazılması uygundur (1b).

Yanıt Kağıdı: Testteki soruların yanıtlarının test üzerine mi yoksa ayrıca hazırlanacak yanıt kağıdı üzerine mi işaretleneceği, testin uygulanacağı kişilerin olgunlaşma düzeyi, ekonomik nedenler ve puanlama kolaylığı ile ilgilidir. Okul öncesi ve okulun ilk yıllarda yanıtlama test üzerine yapılmakta aynı yanıt kağıdı kullanılmamaktadır. Bu tür uygulamada yanıtlama ile ilgili işaretlerin belirli bir yere yaptırılması veya doğru yanıtla ilişkin harfin soru numarasının hizasına uygun bir yere yazdırılması puanlamada kolaylık sağlar. İlkokulun ikinci devresinden başlayarak ayrı yanıt kağıdı kullanılabilir.

Eğer puanlama optik okuyucularla yapılacaksa, test uygulanırken buna uygun hazırlanmış yanıt kağıtları kullanmak gereklidir. Elle puanlamada bir yanıt kağıdı hazırlanmalıdır. Yanıt kağıdının üst bölümünde test alan kişiye, kuruma ve teste ilişkin bilgiler bulunur; alt bölüm ise yanıtlama yeri olarak düzenlenir (1b).

Testin uygulanması: Testlerin uygulanması test kullanma sürecinin en önemli aşamalarından birini oluşturur. Testle toplanacak bilgilerin bireyin gerçek durumunun bir ölçüsü olabilmesi için uygulama aşamasında puanı etkileyebilecek durumlar çok iyi denetlenmelidir. Testin uygulandığı yer, zaman, ıslı, aydınlatma, gürültü, nem, koku, sıraların/sandalyelerin ergonomisi, özürlülerin ve solakların durumu, havalandırma gibi çevre koşulları yönünden uygun ortam sağlanmalıdır. Test uygulayacak kişi testin açıklamasına ve içeriğine ilişkin hazırlık yapmalıdır (1b).

Test alan bireyler, kopya yapmamaları için aralıklı bir düzende oturulmalıdır. Aralıklı bir oturma düzeninin olanaklı olmadığı durumlarda aynı test soruları, sırası ve düzeni değiştirilerek A ve B grupları gibi iki takım haline getirilerek kullanılabilir.

Seçenek (item) Analizi

Seçenek analizi (ya da madde analizi), çoktan seçmeli sınavlarda yanıtlayıcıların her soru için seçeneklere dağılımını ortaya çalışma işlemidir. Bu analiz, her sorunun doğru yanıtlanma oranını görmemizi sağlamasının yanı sıra; yanıtlanma oranı düşük sorularda, dersin sunumundan kaynaklanabilecek hatalar ve eksiklikler, sorunun hazırlanmasında yapılmış hatalar veya öğrencilerin sorulan konuya ilgili bilgi düzeyleri hakkında önemli ipuçları verir. Elde ettigimiz sonuçlara göre sorunun iptali, bütün seçeneklerin doğru kabul edilmesi, doğru seçenek dışında en çok yığılma olan seçenekin/seçeneklerin de doğru kabul edilmesi veya öğrencilerin o konuya ilgili bilgi açıkları ve/veya yanlış bilgilenmelerinin düzeltilmesi yoluna gidilebilir. Ayrıca seçenek analizi bize olası kopya olayları hakkında da ipuçları verir (1c,7,8). Örnek tablo aşağıdadır.

Seçenek analizinin nasıl yapılacağına ilişkin aşağıda bir örnek verilmektedir (8):

SORU 1: Yaşamın ilk 2 yılı içindeki beslenmenin zeka gelişimine etkisini incelemek isteyen bir araştırmacı için hangi tür epidemiyolojik çalışma en uygun olur?

- a. Kesitsel, deskriptif.
- b. Kesitsel, analitik.
- c. Longitudinal, prospektif kohort.
- d. Retrospektif, olgukontrol.
- e. Deneysel klinik

Doğru yanıt 'c' seçeneği olarak kabul edilmiştir. Bu sorunun öğrencilerin yaklaşık %10'u tarafından doğru yanıtlanmasının bekendiği varsayılmıştır. Yanıt dağılımı Tablo II'dedir.

Tablo II. Örnek Soru 1'e Verilen Yanıtların Seçeneklere Dağılımı (9).

Seçenekler	Sayı	%
a	42	20.5
b	50	24.5
c	25	12.3
d	49	24.1
e	38	18.6
Toplam	204	100.0

Seçeneklere yanıtların dağılımının homojen olup olmadığı χ^2 testi ile araştırılır (10). Örnekte $\chi^2 = 10.02$ 'dir ($p < .05$). Farkın kaynağı 'c' seçeneği dışında bırakılarak işlem yinelendiğinde, 2.53 gibi küçük bir χ^2 değeri bulunmaktadır ($p > .05$). Bu durumda yaniltıcı olan 4 seçenekin eşit çekicilikte, isabetle düzenlendiği anlaşılmaktadır. Bu soruyu doğru yanıtlayan öğrencilerin testin genelinde de başarılı olmaları beklenir (8). Ayrıca, bu zorluk

düzeyindeki bir soruda doğru veya yanlış seçenekte asını bir yığılma, olası bir kopya olayını düşündürmelidir.

SORU 2: Aşağıdaki önlemlerin Birincil (B), İkincil (I) ya da Üçüncü (Ü) oluşlarına göre nitelendirmeleri hangi seçenekte doğru sıralanmaktadır?

- * Yeterli ve dengeli beslenme
- * Sosyo-ekonomik önlemlerle yoksulluğu giderme
- * Zamansız ölümüleri önlemeye
- * Sosyal rehabilitasyon
- * Vektör denetimi

- a. B/B/B/Ü/B b. B/I/B/Ü/B c. B/B/I/Ü/B
d. I/B/I/Ü/I e. B/B/B/Ü/I

Bu soruda (10) doğru yanıt 'c' seçeneğidir. Bu soruyu da öğrencilerin %10'undan az bir bölümünün doğru yanıtlayamayacağı beklenmektedir. Tablo III'te yanıt dağılımı verilmiştir:

Tablo III. Örnek Soru 2'ye Verilen Yanıtların Seçeneklere Dağılımı.

Seçenekler	Sayı	%
a	7	3.4
b	6	2.9
c	182	89.2
d	4	1.9
e	5	2.4
Toplam	204	100.0

607.97 gibi oldukça büyük bir χ^2 elde edilmekte ($p < .05$), farkın kaynağı olan 'c' seçeneği dışında bırakıldığında $\chi^2 = .90$ 'a düşmektedir ($p > .05$). Bu da bize yanlış seçeneklerin eşit derecede itici olduğunu göstermektedir.

SORU 3: DSÖ hakkında YANLIŞ olan hangi seçeneklerdir?

- a. Birleşmiş Milletler'in teknik bir uzmanlık kuruluşudur.
- b. Anayasası, TBMM tarafından bir yasa ile onanmıştır.
- c. Aldığı kararlar üye ülkeleri bağlayıcıdır.
- d. Bölge toplantıları her yıl Eylül ayında yapılır ve Genel Kurul'a hazırlıkları.
- e. Tanı yöntemleri ve biyolojik ürünler için standartlar koyar

Sorunun (10) doğru yanıtı 'c' seçeneğidir. Yanıtların seçeneklere dağılımı TABLO IV'tedir

Tablo IV. Örnek Soru 3'e Verilen Yanıtların Seçeneklere Dağılımı.

Seçenekler	Sayı	%
a	20	9.8
b	20	9.8
c	40	20.6
d	104	50.0
e	20	9.8
Toplam	204	100.0

Burada doğru yanıt dışında bırakıldığında $X^2 = 123.21$ ($p < .05$). % 50'lik yiğilma olan 'd' seçenekti dışarda bırakıldığında ise $X^2 = 0$ 'dır ($p > .05$). Yanlış 4 seçenekin eşit derecede itici (yeterince yakın) olmadığı ortadadır. Yiğilma olan seçenekte yanlış bir vurgu veya dersin anlatımı sırasında yapılmış bir yanlış söz konusu olabilir. Bu durumda 2 yol düşünülebilir. İlkisi isabetli sorulamadığı için sorunun iptali, ikincisi ise; soruyu doğru yanıtlayan % 20.6'lık öğrenci kesiminin hakkını yememek için yiğilma olan seçenekin de doğru sayılmasıdır.

Test Puanlarının Yorumlanması: Test puanlarının değerlendirilmesinde *ham puanları* kullanmak yerine *standart puanlara dönüşüm* uygulanır. Genelde z ve T puanı olmak üzere iki tür standart puan kullanılır. Ancak, farklı ortalama ve standart sapmalara göre çok çeşitli puanlar elde edilebilir (1d).

z Puanı: z puanı, ortalaması sıfır ve standart sapması 1 olan bir puandır. Aşağıdaki eşitlige göre örnekteki gibi hesaplanır (1d, 8):

$$z = \frac{X - \bar{X}}{S} = \frac{65 - 60}{5} = +1$$

Bu eşitlikte, x puan dağılımının aritmetik ortalaması, X bireyin testten aldığı puan, S ise puan dağılımının standart sapmasıdır. Örn. ortalaması 60, standart sapması 5 olan bir puan dağılımında 65 alan bir bireyin z puanı +1 olur.

T Puanı: T puanı, ortalaması 50 ve standart sapması 10 olan bir puandır. T puanının eşitliği bir örnekle aşağıda verilmiştir (1d):

$$T = 50 + 10\left(\frac{X - \bar{X}}{S}\right) = 50 + 10\left(\frac{65 - 60}{5}\right) = 60$$

Buna göre, örnekte verilen öğrencinin T puanı 60 olacaktır.

Bir testten elde edilen en yüksek ve en düşük puanlar arasında kalan ham puanların karşılığı olan standart puanlar bir liste halinde gösterilebilir. Herhangi bir ham puanın yorumu için bu listeye başvurularak ham puanın temsil ettiği başarı düzeyi standart puan listesinden hazır olarak elde edilir. Böyle bir işlem, özellikle seçme sınavlarında bir zorunluluktur. Sınavla ilişkin genel başarı listesinin hazırlanabilmesi için bireylerin sınavda kullanılan ölçme araçlarının her birinden elde ettikleri puanların, toplam puana dönüştürülebilmesi gereklidir. Bunun için, her puan türünde ham puanların karşılığı olan standart T puanları hesaplanmalı ve bireylerin çeşitli testlerden aldığı

ham puanların karşılığı olan standart puanlar toplanmalıdır.

Daha sonra bireylerin standart puanları toplanarak, toplam puana göre sınava girenlerin tümünü kapsayan bir genel başarı listesi hazırlanmaktadır. Bu liste üzerinden kontenjan sayısı kadar kişi sırlararak/sırlanarak seçilmektedir. Tablo V'te, ortalaması 65 ve standart sapması 15 olan bir puan dağılımına ait ham puanların karşılığı olan standart T puanları bir liste halinde verilmiştir (1d):

Tablo V. Standart Puan Listesi.

Ham Puan	T Puanı	Ham Puan	T Puanı
120	87	60	47
115	83	55	43
110	80	50	40
105	77	45	37
100	73	40	33
95	70	35	30
90	67	30	27
85	63	25	23
80	60	20	20
75	57	15	17
70	53	10	13
65	50	5	10

(Standart T puanı, $\bar{X} = 50$, $S = 10$)

Tablo V'te puanlar 5'er atlayarak verilmiştir. Bir seçme durumunda standart puanların, her ham puan için tek tek hesaplanması gereklidir. Bireysel puanların yorumlanması için de standart puanlardan yararlanılır. Örn. ham puanı 45 olan bir öğrencinin standart puanı 37'dir.

KONUYA İLİŞKİN OLARAK KENDINI DEĞERLENDİRME SORULARI

Aşağıda, okuyucunun kendini değerlendirebilmesi amacıyla 4 adet çoktan seçmeli soru bulunmaktadır. Okuyucu doğru yanıtları bulmada zorlanırsa, ilgili konuları yeniden gözden geçirmeli ve/veya diğer arkadaşlarıyla tartışmalıdır.

SORU 1: Seçenek analizi, aşağıdakilerden hangisi hakkında fikir vermez?

- Konuya ilgili bilgi yanlış olup olmadığı.
- Sürenin yeterli olup olmadığı.
- Kopya yapılp yapılmadığı.
- Doğru yanıtların oranı.
- Yanıtların seçeneklere dağılımı.

SORU 2: Çoktan seçmeli bir test sorusunu çözmek için erişkinlerde genellikle önerilen süre aşağıdakilerden hangisidir?

- 30 sn
- 40 sn
- 50 sn
- 60 sn
- 90 sn

SORU 3: Aşağıdaki önermelerin doğru-yanlış sıralaması hangi seçenektedir?

- Eğer bir nitelik testlerle ölçülebiliyorsa yazılı sınavlar yeğlenmemelidir.
- Doğru-yanlış tipi sorular için soru başına 30 sn çözme süresi önerilir.
- Düzeneğmede kolay soruların sona bırakılması daha doğru olur.
- Test puanlarının yorumlanması genellikle Z ve T standart puanları kullanılır.

- a. D/D/Y/D b. D/Y/D/Y c. D/D/D/D
d. Y/Y/D/D e. D/Y/D/D

SORU 4: Eğer bir test sınavı yalnızca doğru-yanlış tipi sorulardan oluşuyorsa, sınavda en az kaç soru olmalıdır?

- a. 20 b. 30 c. 40 d. 50 e. 60

Yanıt anahtarları kaynakçanın altındadır.

KAYNAKLAR

1. Özgüven İ. E. Psikolojik Testler, PDREM, Ankara, 1994 a) s.2425 b) s.144152 c) s.117 d) s.7174
2. Saltık A. Ülkemizde ve Dünyada Tıp ve Halk Sağlığı Eğitimi, Ders Notları, Edirne, s.1415, 1995
3. Association of Medical Schools in Europe, Newsletter, McGraw Hill Med. Pub. Grp., s.7, July 1995
4. Saltık A. Trakya Univ. Tıp Fak. Halk Sağlığı 3. Dönem 1. Vize Sınavı Soruları, Edirne, 1994
5. Turgut M. F. Eğitimde Ölçme ve Değerlendirme Metodları, Nüve Mtb. Ankara, s.106, 1977
6. Davis F. B. Educational Measurements and Their Interpretation, s. 8081.
7. Hubbard J.P. Measuring Medical Education, Lea and Febirger, Philadelphia, 2nd ed., s.4957, 1978
8. Saltık A. Bir Halk Sağlığı Finali ve Düşündürdükleri. Ege Univ. Tıp Fak. Dergisi. 31: (1), 6166, (1992). [II. Ulusal Halk Sağlığı Kongr. Bild., Mayıs 1990, İstanbul]
9. Saltık A., Yorulmaz F. Bir Epidemiyoloji Sınavının Pedagojik Değerlendirmesi ve Tıp Eğitimi Üzerine Düşünceler. Ege Univ. Tıp Fak. Derg. 31:(1), 4752, (1992). [II.Uluslararası Halk Sağlığı Kong. Bild., Mayıs 1990, İstanbul]
10. Sümbüloğlu K. Biyoistatistik, Özdemir Yay. 4. Baskı, Ankara, s.172173 ve s.280, 1993
11. Saltık A. Trakya Univ. Tıp Fak. Halk Sağlığı 3. Dönem Bütünleme Sınavı Soruları, Edirne, 1995

Yanıt anahtarları: 1b 2d 3a 4d