

**Gaziosmanpaşa Üniversitesi Tıp Fakültesi'nde Uygulanan Çoktan Seçmeli Sınavların
Betimsel Analizi**

**Descriptive Analysis of the Multiple-Choice Exams Applied in the Medicine Faculty of
the Gaziosmanpasa University**

Kürşat Volkan Özcan¹, Yunus Aydoğan¹, İsmail Bulut¹

¹Gaziosmanpaşa
Üniversitesi Tıp
Fakültesi, Tıp Eğitimi
ve Bilişimi Anabilim
Dalı, Tokat.

Yazışma Adresi:

**Öğr. Gör. Kürşat
Volkan Özcan**

Gaziosmanpaşa
Üniversitesi Tıp
Fakültesi, Tıp Eğitimi
ve Bilişimi Anabilim
Dalı

E-mail:

kursatvolkan@hotmail.com

Tel: 0 5532434828

Özet

Amaç: Tıp eğitiminde öğrenme hedeflerine ulaşmadaki durumların belirlenmesi için değerlendirme ve bu değerlendirmenin sağlıklı yapılması önemlidir. Bu çalışmanın amacı Gaziosmanpaşa Üniversitesi Tıp Fakültesi öğrencilerine uygulanan çoktan seçmeli sınavların analizlerinin yapılarak sınavlarda kullanılan maddeler hakkında ilgili geri dönütler vererek daha nitelikli bir ölçme değerlendirme yapılmasına katkı sağlamaktır.

Yöntem: Gaziosmanpaşa Üniversitesi Tıp Fakültesinde 2013 – 2014 Güz öğretim yılında Dönem 1, 2 ve 3'üncü sınıf öğrencilerine uygulanan 11 kurul sonu teorik değerlendirme sınavı ve bu sınavlarda kullanılmış 941 madde bu çalışma kapsamında incelenmiştir. Araştırma modeli olarak "tarama modeli" kullanılmıştır. Araştırma kapsamında sınavların çarpıklık katsayıları, maddelerin güçlük indeksleri, ayırt edicilik indeksleri ve doğru cevap seçeneğe dağılımları hesaplanmıştır.

Bulgular: Madde ve test analizleri yapılan 11 kurul sonu değerlendirme sınavının çarpıklık katsayılarının genel itibariyle 0'ın altında, güçlük düzeylerinin ise 0,64 seviyesinde olduğu görülmüştür. İncelenen testlerde kullanılan maddelerin % 62'lik bir diliminin 0,29 ile 0,79 güçlük indeks aralığında yer aldığı belirlenmiştir. Ayrıca maddelerin doğru cevap seçeneğe dağılımları incelendiğinde ise doğru cevabının seçeneklere eşit dağılmadığı görülmüştür.

Sonuç: Tıp Fakültesinde kurul sonu değerlendirme sınavlarında kullanılan, güçlük düzeyleri çok düşük ve çok yüksek maddeler üzerinde düzeltmeler yapılması daha nitelikli ölçme araçlarının oluşmasına imkân sağlayabilir. Ayrıca öğretim elemanlarının ölçme ve değerlendirme alanında verilecek bir eğitimle daha etkin değerlendirme yapılması sağlanabilir.

Anahtar kelimeler: Tıp eğitiminde değerlendirme, çoktan seçmeli testler, madde güçlük indeksi.

Abstract

Aim: For the determination of the cases to achieve to the learning objectives in medical education is important to make a healthy assessment. The aim of this study is that multiple-choice exams, which are applied to students at Faculty of Medicine of Gaziosmanpaşa University, analyzes and their knowledge of the ingredients used in the exams is to ensure that.

Materials and Methods: In this study we investigated that the final 11 committees of theoretical evaluation exams were applied to the 1st, 2nd and 3rd grade students at Faculty of Medicine of Gaziosmanpaşa University in Fall Semester in the academic year 2013 - 2014 and 941 questions in the exams. "Screening model" has been used as a research model. The distributions of coefficient of skewness of examinations, difficulty index of question and correct answer choice under investigation were calculated.

Findings: It was showed that the skewness coefficient of the investigated final 11

Giriş

Öğrenme-öğretme sürecinde bireylerin kazanmış oldukları bilgi, beceri ve yeterlilikler ile öğretim uygulamalarının etkililiğinin belirlenebilmesi adına ölçme ve değerlendirme faaliyetlerine ihtiyaç duyulmaktadır (1). Eğitim sistemi girdi, süreç, çıktı ve kontrol öğelerinden oluşur (2). Ölçme ve değerlendirme ise eğitim sisteminin kontrol ögesi niteliğindedir (3). Hedeflenen eğitim çıktılarının gerçekleşip gerçekleşmediğini veya ne derece gerçekleştirildiğini ortaya çıkarma, öğrenme güçlüklerini belirleme, eğitim programlarının, uygulanan yöntem ve tekniklerin etkililiğini saptama, öğrencileri yönlendirme ve benzeri amaçlar için ölçme ve değerlendirme çalışmaları yapılmaktadır (4). Eğitim sistemi içinde öğrenci, öğretici

committees of evaluation exams in general were below 0, while the difficulty level was found to be 0.64 level. It also was determined that 62 percent of the questions used in the final committees of evaluation exams were in the range of appropriate difficulty index. The right answer choices in multiple-choice exams analyzing the distribution of the correct answer options were not found to be evenly distributed.

Conclusion: On the ingredients of difficulty levels of very low and very high of questions used in the assessment exam in the Faculty of Medicine to make any corrections would permit the formation of a more qualified assessment tools. Also in the field of measurement and evaluation of the teaching staff will be given an through education can be achieved more effectively assessment.

Key words: Assessment in medical education, multiple choice tests, question difficulty index.

ve programlar hakkında doğru karar verilmesinde ölçme ve değerlendirme büyük bir öneme sahiptir (5).Günümüz tıp fakültelerinde verilen eğitimin usta-çırak ilişkisinden yapılandırılmış bir modele geçmesiyle, tıp eğitiminde de objektif ve uluslararası normlarda ölçme ve değerlendirme yapmak önemli hale gelmiştir (6). Öğretimin kalitesini artırabilmek ve etkin bir değerlendirme yapabilmek için tıp fakültesi öğretim elemanlarının tümünün ölçme ve değerlendirme becerilerine sahip olmaları beklenir (7,8).

Değerlendirme, yapılmasındaki amaca göre tanımaya yönelik değerlendirme, biçimlendirmeye yönelik değerlendirme ve ürüne yönelik değerlendirme olarak 3 gruba ayrılır (2). Tıp fakültelerinde belirli periyodlarla

eđitim programının öngördüđü hedeflere ulařılıp ulařılmadıđına bakılarak öđrenci, öđretim elemanı ve programa iliřkin yargıda bulunulması ürüne yönelik deđerlendirme yapıldıđını göstermektedir. Düzey belirlemeye yönelik deđerlendirme olarak da bilinen bu deđerlendirme türünde çoktan seçmeli sorular sık olarak kullanılmaktadır (9,10).

Deđerlendirmenin sađlıklı yapılabilmesi için ölçme sonuçlarının geçerli ve güvenilir olması kadar kullanılan ölçme aracının da güvenilirlik, geçerlik gibi bazı kriterleri taşıması beklenmektedir (4, 11, 12). Bu beklentilerin karşılanabilmesi için sınavlarda çoktan seçmeli soruların kullanılması bir avantaj olarak karşımıza çıkmaktadır. Çoktan seçmeli soruların deđerlendirilmesi ve objektifliđinin sađlanması diđer soru türlerine nazaran daha kolaydır (13). Öđrencilerin performans iyileřtirme ve öđrenme hedeflerine ulařmadaki durumlarının belirlenmesinde iyi bir testin kullanılması kadar bu testin yorumlanması da ayrı önem taşımaktadır. Testte veya testin deđerlendirilmesinde meydana gelecek olası sorunlar öđrenme çıktılarının yorumlanma sürecini de olumsuz anlamda etkileyecektir (14). Tıp fakóltesi öđretim elemanlarının deđerlendirme süreçlerine dair literatürün az olması ise bu arařtırmanın önemini göstermektedir (15).

Bu çalışmanın amacı Gaziosmanpařa Üniversitesi Tıp Fakóltesi öđrencilerine uygulanan çoktan seçmeli sınavların ve bu sınavlarda kullanılan maddelerin analizlerinin yapılması ve sınavlarda kullanılan maddeler hakkında düzeltmeler önermek, soru ve sınav kalitesini geliřtirmektir. Bu amaç çerçevesinde ařađıdaki 4 arařtırma sorusuna cevap aranmıřtır.

1. Uygulanan çoktan seçmeli sınavların ortalama güçlük düzeyleri nedir?
2. Uygulanan çoktan seçmeli sınavlardaki maddelerin güçlük indeksleri nedir?
3. Uygulanan çoktan seçmeli sınavlardaki maddelerin dođru cevap seçenek dađılımları nasıldır?
4. Uygulanan çoktan seçmeli sınavlardaki maddelerin ayırt edicilik indeksleri nedir?

Gereç ve Yöntem

Gaziosmanpařa Üniversitesi Tıp Fakóltesinde 2013 – 2014 Güz öđretim yılında Dönem 1, 2 ve 3'üncü sınıf öđrencilerine uygulanan 11 kurul sonu teorik deđerlendirme sınavında sorulmuş 941 madde bu çalışma kapsamında incelenmiřtir. Bu süreçte arařtırma modeli olarak "tarama modeli" kullanılmıřtır. Tarama modeli geçmişte ya da günümüzde olan bir durumu, var olduđu řekli ile betimlemeyi amaçlayan bir yaklařımdır (16).

Arařtırma kapsamında her sınav için o sınava giren öđrencilerin puanlarının sınav ortalamasına göre dađılımlarına bakılarak uygulanan sınavların çarpıklık katsayıları hesaplanmıřtır. Öđrencilerin test puanlarına ait çarpıklık katsayısı hesaplanırken řekil 1'deki formülden yararlanılmıřtır.

$$\text{Çarpıklık Katsayısı} = \frac{n \sum_{i=1}^n (x_i - \bar{x})^3}{(n-1)(n-2)S^3}$$

$$\text{Standart Sapma} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{(n-1)}}$$

řekil 1. Çarpıklık katsayısının hesaplanması (17).

Elde edilen çarpıklık katsayısının negatif olması testin sola çarpık ve kolay olduğunu, pozitif olması sağa çarpık ve uygulanan sınavın zor olduğunu ve sifra yakın olması ise dağılımın simetrik yani normal dağılım olduğunu göstermektedir (18,19). Çarpıklık katsayısına göre testin gücüyle ilgili yorumlar Tablo 1'e göre yapılmıştır.

Tablo 1. Sınavların Çarpıklık Katsayıları ve Testin Ortalama Güçlükleri (20).

Çarpıklık Katsayısı	Grup için test gücü
Negatif	Kolay
Pozitif	Zor
0,10'dan küçük	Hafif zor
0,10- 0,25 arası	Orta güçlükte
0,25'ten büyük	Çok zor

Çarpıklık katsayısının yanı sıra testin gücüyle ilgili yorum yapabilmek için testte yer alan maddelerin güçlük indekslerinin (p) ortalamasının da testin ortalama gücünü verdiği belirtilmiştir (18). Bu açıklamadan yola çıkarak madde güçlük indeksleri hesaplanmıştır. Madde güçlük indeksi o maddeyi doğru cevap verenlerin sayısının sınava girenlerin sayısına oranı şeklinde hesaplanmıştır. Madde güçlük indeksini hesaplamak için kullanılan formül Şekil 2'de verilmiştir.

$$\text{Madde Güçlük İndeksi} = \frac{\text{Maddeyi doğru cevaplayan sayısı}}{\text{Sınava giren öğrenci sayısı}}$$

Şekil 2. Madde güçlük indeksin hesaplanması (21).

Literatür incelendiğinde güçlük indeksi sonuçlarının farklı aralıklarda değerlendirildiği bilinmektedir (21,22). Bu hesaplama sonucunda maddelere ait güçlük değerlerine ait değerlendirmeler Tablo 2'ye göre yapılmıştır.

Tablo 2. Maddelerin güçlük indeksi ve karşılıkları (22,23).

Güçlük İndeksi	Anlamı
0,80 – 1,00	Çok kolay
0,65 – 0,79	Oldukça kolay
0,35 – 0,64	Orta düzey
0,20 – 0,34	Oldukça zor
0 – 0,19	Çok zor

Madde güçlük indeksleri belirlendikten sonra ölçülmek istenen özelliğe sahip olan ile sahip olmayana ayırabilen maddeler tespit edilmiştir. Madde ayırt edicilik (r_{jx}) gücü olarak bilinen bu işlemde belirlenen amaca hizmet eden maddeler ortaya çıkarılmıştır (12). Bu indeks korelasyona dayalı olduğu için +1 ile -1 arasında değişen değerler alır. Bu indeks +1'e yaklaştığında maddenin, testle ölçülmek istenen özelliği ölçtüğü, 0'a yaklaştığında testle ölçülmek istenen özelliği ölçmediği, -1'e yaklaştığında da maddenin testin bütününde ölçülmek istenen özellikten başka bir özelliği ölçtüğü anlamına gelir (2). Madde ayırt edicilik gücünü hesaplamak için kullanılan formül Şekil 3'te verilmiştir.

$$r_{jx} = \frac{\bar{X}_{jd} - \bar{X}_x}{S_x} \sqrt{\frac{p}{1-p}}$$

Şekil 3. Madde ayırt edicilik gücünün hesaplanması (12).

Ayrt edicilik gücüne ilişkin yapılmıştır. değerlendirmeler Tablo 3'e göre

Tablo 3. Madde ayrt edicilik gücü ve karşılıkları (12, 24).

Ayrt edicilik indeksi	Anlamı
-1,00 – 0,19	Ayrt ediciliği düşük teste yer almamalı
0,20 – 0,29	Düzeltilerek teste konulabilir
0,30 – 0,39	Düzeltilme yapmadan veya küçük düzeltmelerle teste alınabilir
0,40 – 1,00	Teste aynen konulabilir ve ayrt edici bir maddedir

Son olarak ise çalışma sürecinde değerlendirilen sınavlarda yer alan maddelerin doğru seçenek dağılımları yüzde cinsinden tespit edilmiştir.

Bulgular

Bu kısımda araştırma kapsamında incelenen 4 araştırma sorusu paralelinde bulgular sırasıyla sunulmuştur.

Tıp fakültesi öğrencilerine uygulanan çoktan seçmeli sınavların çarpıklık katsayıları ve testin ortalama güçlükleri Tablo 4'te verilmiştir.

Tablo 4. Sınavların Çarpıklık Katsayıları ve Ortalama Güçlükleri.

Dönem	Sınav	Çarpıklık Katsayısı	Testin Ortalama Güçlüğü	Testin Güçlüğü
Dönem 1	Kurul 1	-0,39	0,69	Oldukça Kolay
	Kurul 2	-1,68	0,71	Oldukça Kolay
	Kurul 3	-0,62	0,64	Orta Düzey
	Kurul 4	-0,25	0,60	Orta Düzey
Dönem 2	Kurul 1	-0,63	0,66	Oldukça Kolay
	Kurul 2	-0,04	0,58	Orta Düzey
	Kurul 3	-0,40	0,66	Oldukça Kolay
	Kurul 4	-0,18	0,55	Orta Düzey
Dönem 3	Kurul 1	0	0,74	Oldukça Kolay
	Kurul 2	-0,52	0,55	Orta Düzey
	Kurul 3	-1,18	0,69	Oldukça Kolay

Tablo 4'te yer alan 11 kurul sonu değerlendirme sınavının çarpıklık katsayıları incelendiğinde 1 sınav haricinde diğer tüm sınavlarda bu katsayının negatif olduğu görülmüştür. Dönem 3'lere uygulanan Kurul 1 sınavında ise çarpıklık katsayısının 0 olduğu görülmüştür. Tablo 3'te yer alan kurul sonu değerlendirme sınavlarının ortalama güçlük değerleri

incelendiğinde çok zor ve oldukça zor olarak nitelendirilebilecek bir sınavın olmadığı, 5'inin orta düzey, 6 tanesinin ise oldukça kolay olduğu görülmüştür.

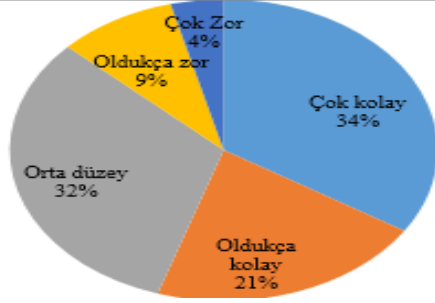
Tıp fakültesi öğrencilerine uygulanan çoktan seçmeli sınavlarda kullanılan maddelerin güçlük indeksleri Tablo 5'de verilmiştir.

Tablo 5. Sınavlardaki Maddelerin Güçlük İndeksi Dağılımları.

Dönem	Sınav	Çok Kolay (0,80 – 1,00)	Oldukça Kolay (0,65 – 0,79)	Orta Düzyey (0,35 – 0,64)	Oldukça Zor (0,20 – 0,34)	Çok Zor (0 – 0,19)
Dönem 1	Kurul 1	0,41	0,21	0,29	0,07	0,03
	Kurul 2	0,46	0,19	0,25	0,10	0,00
	Kurul 3	0,30	0,21	0,36	0,08	0,06
	Kurul 4	0,22	0,24	0,41	0,07	0,05
Dönem 2	Kurul 1	0,35	0,24	0,28	0,08	0,05
	Kurul 2	0,20	0,23	0,41	0,11	0,06
	Kurul 3	0,33	0,27	0,31	0,07	0,02
	Kurul 4	0,17	0,17	0,47	0,12	0,07
Dönem 3	Kurul 1	0,47	0,23	0,23	0,04	0,04
	Kurul 2	0,22	0,16	0,35	0,17	0,10
	Kurul 3	0,46	0,18	0,26	0,07	0,03
Ortalama		0,34	0,21	0,32	0,09	0,04

Tablo 5’te yer alan kurul sonu değerlendirme sınavlarında yer alan maddelerin güçlük indekslerinin ortalamaları incelendiğinde maddelerin %34’ünün “çok kolay”, %21’inin “oldukça kolay”, %32’sinin “orta düzey”, %9’unun “oldukça zor” ve %4’ünün ise “çok zor” olduğu görülmüştür. Uygulanan 11 sınavda yer alan 941 maddenin güçlüklerine göre dağılımı grafik olarak Şekil 4’te belirtilmiştir.

Şekil 4. Maddelerin Güçlük İndekslerine Göre Dağılımı (%).



Uygulanan çoktan seçmeli sınavlarda kullanılan maddelerin ayırt edicilik indeksleri Tablo 6'da verilmiştir.

Tablo 6. Maddelerin Ayırt Edicilik İndekslerinin Dağılımı (%).

Dönem	Sınav	Ayırt ediciliği düşük testte yer almamalı (-1,0 – 0,19)	Düzeltilerek kullanılabilir (0,20 – 0,29)	Küçük düzeltmelerle kullanılabilir (0,30 – 0,39)	Ayırt ediciliği yüksek aynen kullanılabilir (0,40 – 1,0)
Dönem 1	Kurul 1	44,74	32,89	18,42	3,95
	Kurul 2	34,02	25,77	28,87	11,34
	Kurul 3	35,56	28,89	26,67	8,89
	Kurul 4	39,66	31,03	20,69	8,62
Dönem 2	Kurul 1	19,77	17,44	33,72	29,07
	Kurul 2	28,79	30,30	24,24	16,67
	Kurul 3	19,28	22,89	37,35	20,48
	Kurul 4	29,33	16,00	26,67	28,00
Dönem 3	Kurul 1	46,34	20,00	21,82	11,82
	Kurul 2	38,00	20,00	20,00	22,00
	Kurul 3	29,00	21,00	22,00	28,00

Ortalama	33,14	24,20	25,49	17,17
-----------------	--------------	--------------	--------------	--------------

Tablo 6’da yer alan kurul sonu değerlendirme sınavlarında yer alan maddelerin ayırt edicilik indekslerinin ortalamaları incelendiğinde maddelerin % 33, 14’ünün “ayırt ediciliği düşük”, % 24,20’inin “ayırt ediciliği orta düzeyde düzeltilerek kullanılabilir”, %25,49’unun “küçük düzeltmelerle veya düzeltme

yapmadan kullanılabilir”, %17,17’sinin ise “ayırt ediciliği yüksek aynen kullanılabilir” olduğu görülmüştür.

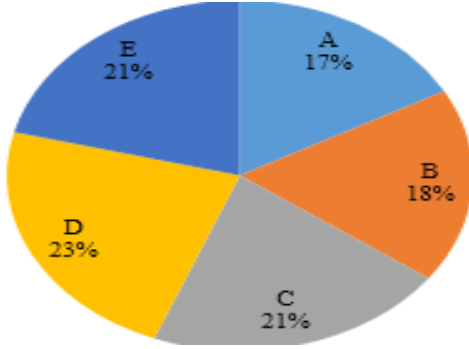
Çoktan seçmeli sınavlarda kullanılan maddelerin doğru cevap seçenek dağılımları Tablo 7’de verilmiştir.

Tablo 7. Maddelerin Doğru Cevap Seçenek Dağılımları (%).

Dönem	Sınav	A	B	C	D	E	Toplam Madde
Dönem 1	Kurul 1	10,53	27,63	17,11	22,37	22,37	76
	Kurul 2	18,56	15,46	22,68	21,65	21,65	97
	Kurul 3	17,78	16,67	23,33	21,11	21,11	90
	Kurul 4	15,52	17,24	20,69	24,14	22,41	58
Dönem 2	Kurul 1	13,95	18,60	20,93	23,26	23,26	86
	Kurul 2	19,70	16,67	21,21	22,73	19,70	66
	Kurul 3	20,48	21,69	21,69	14,46	21,69	83
	Kurul 4	21,33	17,33	16,00	22,67	22,67	75
Dönem 3	Kurul 1	12,73	15,45	20,91	29,09	21,82	110
	Kurul 2	14,00	13,00	25,00	27,00	21,00	100
	Kurul 3	18,00	20,00	21,00	25,00	16,00	100
Ortalama		16,60	18,16	20,96	23,04	21,24	85,55

Tablo 7’de yer alan kurul sonu değerlendirme sınavlarında yer alan maddelerin doğru cevap seçenek dağılımları incelendiğinde öğrencilere yöneltilen maddelerin % 16,60’ının doğru cevabı “A”, % 18,16’sinin doğru cevabı “B”, % 20,96’sinin doğru cevabı “C”, % 23,04’ünün doğru cevabı “D” ve % 21,24’ünün ise doğru cevabı “E” seçeneğinde olduğu görülmüştür. Uygulanan sınavlardaki maddelerin doğru cevap seçenek dağılımının grafiksel gösterimi Şekil 5’te belirtilmiştir.

Şekil 5. Maddelerin Doğru Cevap Seçenek Dağılım Ortalaması (%).



Sonuç ve Tartışma

Tıp eğitiminde öğrenmelerin değerlendirilmesi öğretim elemanları için genellikle ciddi bir endişe konusudur. Sık olarak kullanılan çoktan seçmeli soruların değerlendirmede tek başına kullanılmasının yeterli olmadığı belirtilmektedir. Fakat alan yazın incelendiğinde ise öğrencilerin değerlendirilmesinde kendi kendine eğitim ve klinik akıl yürütme becerilerinin geliştirilmesi gibi alanlarda çoktan seçmeli sorulardan yararlandığı görülmektedir (25). Öğretimin etkinliğinin belirlenmesi için belirli kriterleri taşıyan bir ölçme aracı ile iyi bir değerlendirme yapabilmek her aşamada önemlidir (26). Bu bağlamda Gaziosmanpaşa Üniversitesi Tıp Fakültesi öğrencilerine uygulanan ölçme araçlarının niteliğini belirlemek amacıyla çoktan seçmeli sınavların ve bu sınavlarda kullanılan maddelerin analizleri yapılmıştır.

2013-2014 Eğitim-Öğretim yılı Güz Yarıyılında Dönem 1, 2 ve 3'üncü sınıfta uygulanan 11 kurul sonu değerlendirme sınavının çarpıklık katsayıları

incelendiğinde Dönem 3 Kurul 1 sınavı haricinde diğer tüm sınavların kolay olduğu söylenebilir. Elde edilen çarpıklık katsayılarının negatif olması sınavın kolay olduğuna işaret ettiği gibi sınava giren öğrencilerin başarısını da göstermektedir. Bu yüzden sadece çarpıklık katsayılarına bakarak sınavın güçlüğü hakkında yorum yapmak doğru olmayacaktır. Bu sınavlarda yer alan maddelerin güçlük indeksleri incelendiğinde 5 sınavın güçlük indeksinin 0.35 – 0.64 aralığında yani “orta düzey güçlükte”, 6 sınavın güçlük indeksinin ise 0.65 – 0.79 aralığında yani “oldukça kolay” olduğu görülmüştür. Beş seçenekli çoktan seçmeli testlerde kullanılan maddeler için ideal güçlük indeksinin 0.6 olması gerektiği bilinmektedir (27). İlgili alan yazın incelendiğinde bu seviyenin altında güçlük indeksine sahip yani daha zor olarak nitelendirilebilecek sınavların uygulandığı görülmüş ve bu durum eleştiri konusu olmuştur (27). Bunun yanı sıra uygulanan sınavlarda sorulan maddelerin güçlük indeksinin 0,70’ler seviyesinde olduğunu belirten çalışmalar da mevcuttur

(28). Ayrıca literatürde çarpıklık katsayısına göre kolay olarak nitelendirilebilecek sınavların uygulandığı bilinmektedir (29).

Çalışma kapsamında 941 maddenin güçlük indeksleri incelendiğinde 319 maddenin (%34) “çok kolay”, 197 maddenin (%21) “oldukça kolay”, 300 maddenin (%32) “orta düzey”, 83 maddenin (%9) “oldukça zor” ve 42 maddenin (%4) ise “çok zor” olduğu görülmüştür. Çoktan seçmeli maddeler için ideal güçlük indeksinin 0.20 – 0.79 aralığında olması gerektiği belirtilmiştir (30). Bu açıklamadan hareketle çalışma kapsamında incelenen soruların “çok kolay” olarak nitelendirilebilecek %34’ü ile “çok zor” olarak nitelendirilebilecek %4’ünün belirtilen aralığın dışında olduğu söylenebilir. Alanyazın incelendiğinde bu aralığın farklı sınırlarda belirtildiği ($p < 0.3$ çok zor, $p > 0.70$ çok kolay) çalışmalar mevcuttur. Bu çalışmalarda “çok zor” olarak nitelendirilen maddelerin (%5.97 - %7.6) oranında ve çok kolay olarak nitelendirilen maddelerin ise (%44.88 - %41.20) oranında kullanıldığı bilinmektedir. (5,27). Gaziosmanpaşa Üniversitesi Tıp Fakültesi öğrencilerine uygulanan çoktan seçmeli sınavlardaki maddelerin güçlük değerlendirmeleri $p > 0.7$ ’ye göre yapıldığında %50’ye yakının “çok kolay” olduğu söylenebilir.

Çalışma kapsamında ayırt edicilik indeksleri analiz edilen 941 maddenin %33,14’ünün “ayırt ediciliği düşük” ve %17,17’sinin ise “ayırt ediciliği yüksek aynen kullanılabilir” olduğu görülmüştür. Literatüre incelendiğinde çalışmalarda analiz edilen maddelerin %15,2-%18,29’unun ayırt ediciliklerinin kabul edilebilir alt sınırdan düşük olduğu görülmüştür (10,29). Elde edilen sonuçlar incelenen maddelerin ayırt ediciliklerin

literatüre nazaran daha zayıf olduğu göstermektedir.

Uygulanan çoktan seçmeli sınavlardaki maddelerin doğru cevaplarının seçeneklere dağılımları incelendiğinde doğru cevapların her seçenekte eşit olarak dağılması gerektiği söylenebilir. 941 maddeden 155 maddenin (%17) doğru cevabının “A” seçeneğinde, 169 maddenin (%18) doğru cevabının “B” seçeneğinde, 199 maddenin (%21) doğru cevabının “C” seçeneğinde, 219 maddenin (%23) doğru cevabının “D” seçeneğinde ve 199 maddenin (%21) doğru cevabının “E” seçeneğinde olduğu yani eşit dağılmadığı görülmüştür. Benzer durumlara literatür tarandığında da rastlamak mümkündür (31). Fakat Dönem 1 Kurul 1 sınavında doğru cevabı “A” ve “C” seçeneğinde bulunan madde sayısının toplamının doğru cevabı “B” seçeneğinde bulunan madde sayısına eşit olması veya Dönem 3 Kurul 1 sınavında doğru cevabı “A” ve “B” seçeneğinde bulunan madde sayısının toplamının doğru cevabı “D” seçeneğinde bulunan madde sayısından az olması şans faktörüne bağlı başarı elde edilmesine imkân sağlayacağından eleştirilebilir.

Tıp eğitiminde farklı amaçlar ile web temelli değerlendirme çalışmalarının yapıldığı bilinmektedir (32). Bu bağlamda öğrencilerin kazanımlarını değerlendirmek amacıyla kullanılan değerlendirmelerin objektifliğini ve etkinliğini artırmak için web tabanlı ortamların geliştirilmesi tavsiye edilebilir. Geliştirilecek bu ortamların farklı değerlendirme yöntemlerine imkân sağlaması sağlanarak değerlendirmelerin sadece çoktan seçmeli testlerle yapılması kısıtlılığını ortadan kaldırılabilir. Ayrıca tıp fakültesi öğretim elemanlarına “Eğitim Becerileri” ve “Tıp Eğitiminde Ölçme ve Değerlendirme” gibi hizmet içi programlar düzenlenerek ölçme

ve değerlendirme alanında yetkinlikleri geliştirilebilir ve farklı ölçme yöntemlerinin kullanılması önerilebilir.

Kaynaklar

1. Karip E. (Ed.), Ölçme ve Değerlendirme (5.Baskı). Ankara, Pegem Yayıncılık, 2012.
2. Yaşar M. Ölçme ve Değerlendirme ile İlgili Temel Kavramlar, Tekindal, S. (Ed.), Eğitimde Ölçme ve Değerlendirme, 3. Baskı, Ankara, Pegem Akademi, 2011; 2: 33.
3. Çıkrıkçı-Demirtaşlı N. (Ed.), Eğitimde Ölçme ve Değerlendirme, Ankara, Elhan Yayınları, 2012.
4. Atılğan H, Kan A, Doğan N. Ölçmenin Temel Kavramları, Atılğan, H. (Ed.), Eğitimde Ölçme ve Değerlendirme, Geliştirilmiş 3. Baskı, Ankara, Anı Yayıncılık, 2009;2:24.
5. Bahar M, Nartgün Z, Durmuş S, Bıçak B. Geleneksel-Tamamlayıcı Ölçme ve Değerlendirme Teknikleri (5. Baskı), Ankara, Pegem Yayıncılık, 2012.
6. Sayek İ, Odabaşı O. “Türk Tabipleri Birliği Mezuniyet Öncesi Tıp Eğitimi Raporu 2010”. Ankara, Türk Tabipleri Birliği Yayınları, 2010.
7. Çakan M. Öğretmenlerin ölçme-değerlendirme uygulamaları ve yeterlik düzeyleri: İlk ve ortaöğretim. Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi. 2004;37(2):99-114.
8. Yıldız İ, Uyanık N. Matematik eğitiminde ölçme-değerlendirme üzerine. Gazi Üniversitesi Kastamonu Eğitim Dergisi. 2004;97.
9. Tekin H. Eğitimde Ölçme ve Değerlendirme, Gözden Geçirilmiş 9. Baskı, Ankara, Yargı Yayınları, 1996.
10. Topal K, Aybek H, Kara CO, Büke A, Aybek Z. “PAÜ Tıp Fakültesi Dönem I Öğrencilerine 2006-2007 Eğitim ve Öğretim Yılında Uygulanan Çoktan Seçmeli Sınavların Madde ve Test Analizleri”. Pamukkale Tıp Dergisi. 2008;1(3):100-25.
11. Turgut M. Eğitimde Ölçme ve Değerlendirme Metodları, 10. Baskı, Ankara, Yargıcı Matbaası, 1997.
12. Turgut MF, Baykul Y. Eğitimde Ölçme ve Değerlendirme, 4. Baskı, Ankara, Pegem Akademi. 2012;7:119-20.
13. Çallı İ, Torkul O, Taşbaşı N. İnternet destekli öğretimde kullanılmak üzere web erişimli veri tabanı yönetim sistemiyle ölçme ve değerlendirme sistemi tasarımı. The Turkish Online Journal of Educational Technology-TOJET. 2003;2(3):108.
14. Hamzah, M., & Abdullah, S. Test Item Analysis: An Educator Professionalism Approach. US-China Education Review A. 2011;307-22.
15. Altahawi F, Sisk B, Poloskey S, Hicks C, Dannefer E. Student perspectives on assessment: experience in a competency-based portfolio system. Medical Teacher, 2012;34(3):221-5.
16. Karasar N. Bilimsel Araştırma Yöntemi, 23. Basım. Ankara: Nobel Akademik Yayıncılık. 2012;79-87.
17. Alpar R. Uygulamalı istatistik ve geçerlik-güvenirlilik: spor, sağlık ve eğitim bilimlerinden örneklerle. Detay Yayıncılık. 2010;83-9.
18. Özçelik DA. Ölçme ve değerlendirme, 4. Baskı, Ankara, Pegem Akademi. 2010;3:84.
19. Baykul Y. Eğitimde ve Psikolojide Ölçme: Klasik test teorisi ve

- uygulaması, 2. Baskı, Ankara, Pegem Akademi. 2010;9:290.
20. Tekin H. Eğitimde ölçme ve değerlendirme, 19. Baskı, Ankara, Yargı Yayınevi. 2008.
 21. Güler N. Eğitimde Ölçme ve Değerlendirme, 5. Baskı, Ankara, Pegem Akademi, 2013;130.
 22. Sözmen EY, Çalışkan A. Ege Tıp Fakültesi 1. Sınıf Öğrencilerine Uygulanan Sınavlarda Bir Grup Biyokimya Sorusunun Analiz Sonuçları. Tıp Eğitimi Dünyası. 2002;(7):38-42.
 23. Tuğut N, Gölbaşı Z. Öğrenci Hemşirelere Yönelik Cinselliği Değerlendirme Bilgi Testi'nin Geliştirilmesi, Geçerlik ve Güvenirliği. Anadolu Hemşirelik ve Sağlık Bilimleri Dergisi/Journal of Anatolia Nursing And Health Sciences. 2013;16(3):148-53.
 24. Crocker L, Algina J. Introduction to Classical and Modern Test Theory. New York: CBS College Publishing, 1986.
 25. Gupta M, Sharma G, Gopal Krishan Pal R, Thaman R, Tikoo D. Strategic Use of MCQs In Undergraduate Medical Students To Improve Objectivity of Formative Assessment. National Journal of Integrated Research In Medicine. 2012;3(2):113-8.
 26. Chandratilake MN, Davis MH, Ponnampereuma GG. Evaluating design in assessments for medical education: the utility formula. Internet Journal of Medical Education. 2010;1(1):7.
 27. Theodorsson T, El Shafie K, Al Wardy N, Khan A, Al Mahrezi A, Al Shafae M. Assessment of Family Doctors in Oman: getting the questions right Preliminary findings of a performance analysis of multiple choice questions. Internet Journal of Medical Education. 2010;1(1):2.
 28. Ware J, Viki T. Quality assurance of item writing: during the introduction of multiple choice questions in medicine for high stakes examinations. Medical Teacher. 2009;31(3):238-43.
 29. Eskiocak S, Süer Gökmen S, Çakır E, Gülen Ş. Trakya Üniversitesi Tıp Fakültesinde Son 5 Yılda Yapılan Biyokimya Sınav Sorularının Analizi. Türk Biyokimya Dergisi Turkish Journal of Biochemistry-Turk J Biochem. 2004;29(4):273-6.
 30. Tongchai A, Sharma M, Johnston I, Arayathanitkul K, Soankwan C. Developing, evaluating and demonstrating the use of a conceptual survey in mechanics. International Journal of Science Education. 2009;31(18):2437-57.
 31. Cankur Şimşek, N., Bülent, E. & Sırmalı, Ş. A., Entegre Eğitim Sisteminde Uygulanan Sınavlardaki Doğru Yanıt Seçeneklerinin Dağılımı. Uludağ Üniversitesi Tıp Fakültesi Dergisi. 2002; 28: 1-3.
 32. Kim S, Brock D, Prouty C, Odegard P, Shannon S, Robins L, Gallagher T. A Web-Based Team-Oriented Medical Error Communication Assessment Tool: Development, Preliminary Reliability, Validity and User Ratings. Teaching & Learning in Medicine. 2011;23(1):68-77.

