

Düzce Üniversitesi Tıp Fakültesi Öğrencilerine Uygulanan Sınavlarda Soruların Geçerlik ve Güvenirlik Analizi

The Validity and Reliability Analysis of Questions in Exams Applied to
Duzce University Faculty of Medicine Students

Handan ANKARALI¹, Şengül CANGÜR^{2*}, Mehmet Ali SUNGUR², Safinaz ATAÖĞLU³

¹Medeniyet Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıp Bilişimi AD, İstanbul

²Düzce Üniversitesi Tıp Fakültesi Biyoistatistik ve Tıbbi Bilişim AD, Düzce

³Düzce Üniversitesi Tıp Fakültesi Fizik Tedavi ve Rehabilitasyon AD, Düzce

ÖZ

Amaç: Bu araştırmanın amacı Düzce Üniversitesi Tıp Fakültesi Dönem I, Dönem II ve Dönem III öğrencilerine uygulanan sınav sorularının geçerlik ve güvenilirlik analizlerini yapmaktır.

Gereç ve Yöntemler: Düzce Üniversitesi Tıp Fakültesi Öğrenci İşlerinden 2014-2015 ve 2015-2016 eğitim-öğretim yıllarına ait Dönem I, Dönem II ve Dönem III öğrencilerine uygulanan sınav sonuçları elde edilmiştir. 2014-2015 ve 2015-2016 eğitim-öğretim yıllarında Dönem I, Dönem II ve Dönem III öğrencilerine uygulanan toplam 38 kurul sınavı ve 6 dönem sonu sınavı olmak üzere 44 sınav sonucu incelenmiştir. Her bir sınav için 100 soru seçilmiş ve incelenen toplam soru sayısı 4400 adettir. Sınava giren her birey analize dahil edilmiştir. Sınav sorularının geçerlik ve güvenilirliğini değerlendirmek için her bir sınava ayrı ayrı "Madde Yanıt Analizi" uygulanmıştır. Her bir sınavın güvenilirliğini değerlendirmek için Kuder-Richardson 20 (KR20), KR21 ve Spearman-Brown güvenilirlik katsayıları hesaplanmıştır.

Bulgular: Çalışma sonucunda, 2014-2015 ve 2015-2016 eğitim-öğretim yıllarında Dönem I ve Dönem II ders kurul sınavlarına ait soruların oldukça kolay düzeyde olduğu ve ayırt edici sorulardan oluştuğu, Dönem III ders kurul sınavına ait soruların oldukça kolay düzeyde olduğu ve düzeltilmesi gereken sorular olduğu söylenebilir. Dönem I, Dönem II final sınavlarına ait soruların ise orta düzeyde zor oldukları ve düzeltilmesi gereken sorular olduğu gözlenmiştir. Dönem III final sınavına ait soruların oldukça zor düzeyde olduğu ve çok zayıf nitelikte yani sınavda yer almaması gereken sorular olduğu söylenebilir.

Sonuç: Sınav niteliğini ve öğrenci başarısını değerlendirmek için uygulanan sınavların hem genel hem de soru bazında farklı teknikler kullanılarak değerlendirilmesi önerilmektedir. Ayrıca sınav sorularının kalitesini artırmak amacıyla tıp fakültesi öğretim üyelerine Tıp Eğitiminde Ölçme ve Değerlendirme konularında eğitimlerin verilmesi tavsiye edilmektedir.

Anahtar kelimeler: Tıp fakültesi; sınav; soru zorluk derecesi; ayırt etme indeksi.

ABSTRACT

Aim: The purpose of this research is to analyze the validity and reliability of the exam questions applied to the students of Duzce University Medical Faculty Term I, Term II and Term III.

Material and Methods: Exam results were obtained from students of Term II, Term II and Term III for the period of 2014-2015 and 2015-2016 academic years from Duzce University Faculty of Medicine Registrar's office. Exams of 44 exams including 38 committee exams and 6 final exams applied to the students in Term I, Term II and Term III during 2014-2015 and 2015-2016 academic years are examined. For each exam 100 questions were selected and the total number of questions examined was 4400. Each individual taking the exam has been included in the analysis. In order to evaluate the validity and reliability of the exam questions, "Item Response Analysis" was applied to each exam separately. Kuder-Richardson 20 (KR20), KR21 and Spearman-Brown reliability coefficients were calculated to evaluate the reliability of each exam.

Results: It is seen that the questions regarding the Term I and Term II committee exams in 2014-2015 and 2015-2016 academic years are quite easy and distinctive questions, the questions regarding the Term III committee exam are quite easy and the questions that need to be corrected. It is observed that the questions of the term I and II final exams are moderately difficult and the questions to be corrected. It can be said that the questions of the term III final exam are very difficult and very poor quality, namely they are questions that should not be included in the exam.

Conclusion: We recommended that examinations applied to assess the quality of the exam and student achievement should be assessed using different techniques, both on a general and a question-based basis. In addition, we recommended that medical faculty academic staffs have to give training in Measurement and Evaluation in Medical Education in order to increase the quality of exam questions.

Keywords: Faculty of medicine; exam; question difficulty; discrimination index.

Sorumlu Yazar / Corresponding Author: Şengül CANGÜR, sengulcangur@duzce.edu.tr

Geliş Tarihi / Received: 10.02.2017 Kabul Tarihi / Accepted: 30.04.2018

GİRİŞ

Günümüz tıp fakültelerinde verilen tıp eğitiminde objektif ve uluslararası normlarda ölçme ve değerlendirme yapmak önemli hale gelmiştir (1). Tıp fakültelerinde sınıf yeterliklerin ölçülmesine yönelik ağırlıklı olarak çoktan seçmeli maddelerden oluşan testler kullanılmaktadır. Bu durumun nedenleri arasında sınıf mevcutlarının fazla olması, optik okuyucular kullanılarak puanlamanın oldukça hızlı yapılabilmesi ve puanlama güvenirliliğinin daha kolay sağlanabilmesi vb. hususlar sıralanabilir. Ancak bilindiği üzere çoktan seçmeli maddeleri hazırlamak diğer madde türlerini hazırlamaktan çok daha güç olabilmektedir (2). Buna eğitim vermekte olan akademisyenlerin nitelikli madde yazımı konusunda yaşadıkları sorunlar da eklendiğinde ölçme sürecine karışan hata miktarı radikal biçimde artabilmekte ve bu durum verilen kararların isabetliliğini olumsuz yönde etkileyebilmektedir. Bu açıdan bakıldığında yapılan sınavlar da bir bakıma amacını yitirmektedir. Tıp fakültelerinde yapılan sınavların niteliği ile ilgili yaşanan bu ve benzer sorunlar; sınavlara ve sınavları yapan kuruma olan güveni zedeleyebilmekte, ayrıca eğitimde fırsat eşitliği ve adalet ilkesi açılarından da sorun oluşturmaktadır (3).

Bu sorunları önlemek amacıyla tıp fakülteleri öğrencilere uygulanan sınavlara yönelik olarak; madde (soru) analizi yapmak, soru bankası oluşturmak, belirli zaman aralıklarında test-ölçek geliştirme, çoktan seçmeli madde yazımında dikkat edilecek hususlara yönelik eğitim planlamak ve vermek gibi çeşitli faaliyetler sürdürmektedir (3-6).

Bu araştırmanın amacı Düzce Üniversitesi Tıp Fakültesi Dönem I, Dönem II ve Dönem III öğrencilerine uygulanan sınav sorularının geçerlik ve güvenilirlik analizlerini yapmaktır. Elde edilecek sonuçlar doğrultusunda sınav sorularının niteliklerini artırmak için önerilerde bulunmaktadır.

GEREÇ VE YÖNTEMLER

Düzce Üniversitesi Tıp Fakültesi Öğrenci İşlerinden 2014-2015 ve 2015-2016 eğitim-öğretim yıllarına ait Dönem I, Dönem II ve Dönem III öğrencilerine uygulanan sınav sonuçları elde edilmiştir. Verileri elde etmeden önce Düzce Üniversitesi Tıp Fakültesi Dekanlığından gerekli izin alınmıştır.

2014-2015 ve 2015-2016 eğitim-öğretim yıllarında Dönem I, Dönem II ve Dönem III öğrencilerine uygulanan toplam 38 kurul sınavı ve 6 dönem sonu sınavı olmak üzere 44 sınav sonucu incelenmiştir. Her bir sınav için 100 soru seçilmiş olup incelenen toplam soru sayısı 4400 adettir. Sınava giren her birey analize dahil edilmiştir. Sınavlarda laboratuvar/pratik soruları veya iptal edilen sorular analiz dışı tutulmuştur.

Tıp fakültesinde her bir dönem farklı ders kurullarından oluşmaktadır. Her bir kurul sonunda uygulanan sınav notlarının ortalaması ders kurul sınav ortalama notunu verir. Dönem sonunda uygulanan sınavdan alınan not ise dönem sonu notunu gösterir. Ders kurul sınav ortalama notunun %60'ı ve dönem sonu sınav notunun %40'ı alınarak dönem sonu notu hesaplanmaktadır. Ayrıca ders kurul sınavlarından ortalama 80 puan ve üzeri alan kişi talep ederse final sınavına girmeme hakkını kullanabilir. Bu durumda ortalama ders kurul puanı ayrıca dönem sonu sınav notu olarak geçer. Bu nedenle dönem sonu sınavlarına giren kişi sayısı ara sınavlara göre daha düşüktür.

İstatistiksel Analiz

Sınav sorularının geçerlik ve güvenirliliğini değerlendirmek için her bir sınava ayrı ayrı Madde Yanıt Analizi (Item Response Analysis) uygulanmıştır. Her bir sınavın güvenirliliğini değerlendirmek için Kuder-Richardson 20 (KR20), KR21 ve Spearman-Brown güvenirlilik katsayıları hesaplanmıştır.

BULGULAR

2014-2015 ve 2015-2016 eğitim-öğretim yıllarında Dönem I, Dönem II ve Dönem III öğrencilerine uygulanan toplam 38 kurul sınavı ve 6 dönem sonu sınavı olmak üzere 44 sınav sonucu

incelenmiştir. Her bir sınav için 100 soru seçilmiş olup incelenen toplam soru sayısı 4400 adettir.

2014-2015 Eğitim-Öğretim Yılı

2014-2015 eğitim-öğretim yılında Dönem I, Dönem II ve Dönem III öğrencilerine uygulanan toplam 19 kurul sınavı ve 3 dönem sonu sınavı olmak üzere 22 sınav sonucu (toplam 2200 adet soru) incelenmiştir. Her bir sınav için ayrı ayrı uygulanan Madde Yanıt analiz sonuçlarının genel değerlendirmesi Tablo 1'de verilmiştir. Sorular, "ortalama soru zorluk derecesi" bakımından incelendiğinde; Dönem I, Dönem II ve Dönem III Ders kurul sınavlarının ve Dönem III final sınavı sorularının "oldukça kolay" nitelikte oldukları saptanmıştır. Dönem I ve Dönem II final sınav sorularının ise "orta düzeyde zor" oldukları görülmüştür. Soruların zorluk derecesinin bir diğer göstergesi olan çarpıklık katsayıları incelendiğinde ise Dönem III final sınavı dışındaki diğer sınavların çarpıklık katsayısının negatif çıkması bu sınavlardaki soruların öğrenciler açısından oldukça kolay olduğunu göstermektedir. Bu bilgi ortalama zorluk derecesi sonucu ile uyum içindedir. Buna karşın Dönem III final sınavına ait çarpıklık katsayısı sıfıra yakın ancak pozitif değer almıştır. Bu sonuç ise Dönem III final sorularının orta düzeyde zorluğuna işaret etmektedir. Ortalama zorluk derecesine göre kolay bir sınav olarak yorumlanan Dönem III final sınavının, çarpıklık katsayısına göre orta düzeyde zorluk taşıdığı bilgisine ulaşılmaması söz konusu sınav sorularının zorluk derecesini iyi yansıtmadığı söylenebilir (Tablo 1).

Sınav sorularının, bilen ve bilmeyeni ayırt etme başarısı "soru ayırt etme indeksi" yardımıyla incelendiğinde Dönem I, Dönem II ders kurul sınav sorularının "bilen ve bilmeyeni ayırmada" oldukça iyi düzenlendiği, Dönem III ders kurulu sınavları, Dönem I final ve Dönem II final sınav sorularının "düzeltilmesi ve geliştirilmesi" gerektiği ve Dönem III final sınavı sorularının ise "çok zayıf nitelikte sorular" oldukları saptanmıştır (Tablo 1). Soru güçlük ve soru ayırt edicilik indeksleri birlikte değerlendirildiğinde;

- Dönem I, II ve III ders kurul sınavlarına ait soruların genel olarak "tipik iyi soru" özelliğinde olduğu
- Dönem I ve II final sınav sorularının üzerinde çalışılması ve geliştirilmesi gereken sorular olduğu
- Dönem III final sorularının ise zor fakat ayırt edici sorulardan oluştuğu gözlenmiştir.

Dönem I ve Dönem II ders kurul sınavlarının KR-20 güvenirlilik katsayı değerleri sırasıyla 0,92 ve 0,90 bulunmuştur. Bu değerlere göre ilgili dönemlere ait ders kurulu sınavlarının "mükemmel güvenirlilikte" oldukları görülmüştür. Dönem I ve Dönem II final sınavlarının KR-20 güvenirlilik katsayı değerleri sırasıyla 0,81 ve 0,85'tir. Buna göre ilgili dönemlere ait final sınavlarının "iyi güvenirlilikte" oldukları söylenebilir. Dönem III ders kurul sınavının KR-20 katsayı değeri (0,84) "iyi güvenirliliği" göstermektedir. Ancak Dönem III final sınavına ait KR-20 katsayısı 0,53 bulunmuştur. Bu değer düşük çıkmasının nedenleri arasında çok kolay soruların olması, çok zor soruların olması, ayrıcalık özelliği olmayan kötü yazılmış soruların varlığı sayılabilir (Tablo 1).

Soru ayırt etme indeksi olarak da kullanılan ortalama nokta çift serili korelasyon katsayı (point biserial correlation coefficient) değerleri incelendiğinde; Dönem I, Dönem II ve Dönem III Ders Kurul sınavlarının sorularının "iyi" soru özelliği taşıdıklarını, Dönem I ve Dönem II final sınav sorularının "orta düzeyde" soru oldukları söylenebilir. Dönem III final sınav sorularının "zayıf" nitelikte sorular olduğu ve revize edilmeleri gerektiği açıktır.

Soruların geçerliliğini değerlendirmek için kullanılan ortalama çift serili korelasyon katsayıları incelendiğinde; Dönem III final sınav soruları dışında diğer tüm sınavlara ait soruların genel olarak geçerli sorular oldukları söylenebilir. Dönem III final sınav sorularının geçerliliğinin düşük olduğu ve revize edilmeleri gerektiği açıktır. Bu araştırmadaki geçerlik kavramı ise sınav sonuçlarının başarıyı ölçme derecesi olarak tanımlanabilir (Tablo 1).

Tablo 1. 2014-2015 eğitim-öğretim yılı Dönem I, Dönem II ve Dönem III öğrencilerinin girdikleri sınavlara ait analiz sonuçları

	Dönem I		Dönem II		Dönem III	
	Ders Kurul Sınavları (4 Kurul)	Final Sınavı	Ders Kurul Sınavları (6 Kurul)	Final Sınavı	Ders Kurul Sınavları (9 Kurul)	Final Sınavı
İncelenen toplam sınav kâğıdı	542	97	755	93	850	72
Minimum puan	15	26	11	28	0	48
Maksimum puan	97	79	97	78	99	85
Medyan puan değeri	71,625	62,75	70	58	74,13	66,17
Ortalama puan değeri	69,60	63,43	70	57	73,89	66,22
Standart sapma	13,94	9,34	13	10	10,16	5,64
Çarpıklık katsayısı	-0,75	-0,93	-0,59	-0,52	-0,52	0,49
Basıklık katsayısı	0,92	2,04	1,06	0,31	1,53	0,08
Ortalama soru zorluk derecesi	0,70	0,63	0,70	0,57	0,74	0,66
Ortalama soru ayırt etme indeksi	0,32	0,20	0,30	0,25	0,22	0,14
Ortalama Nokta Çift Serili korelasyon katsayısı	0,35	0,23	0,32	0,23	0,30	0,16
KR20 güvenilirlik katsayısı	0,92	0,81	0,90	0,85	0,84	0,53
KR21 güvenilirlik katsayısı	0,90	0,74	0,87	0,79	0,77	0,24
Yüksek gruba ait minimum puan	77	67	73	63	66	66
Düşük gruba ait maksimum puan	68	61	71	50	92	69
Spearman-Brown katsayısı	0,88	0,75	0,86	0,78	0,79	0,34
Ortalama Çift Serili korelasyon katsayısı	0,50	0,32	0,47	0,39	0,40	0,19

KR: Kuder-Richardson

Tablo 2. 2015-2016 eğitim-öğretim yılı Dönem I, Dönem II ve Dönem III öğrencilerinin girdikleri sınavlara ait analiz sonuçları

	Dönem I		Dönem II		Dönem III	
	Ders Kurul Sınavları (4 Kurul)	Final Sınavı	Ders Kurul Sınavları (6 Kurul)	Final Sınavı	Ders Kurul Sınavları (9 Kurul)	Final Sınavı
İncelenen toplam sınav kâğıdı	568	105	793	98	1052	85
Minimum puan	8	27	19	18	28	47
Maksimum puan	77	63	95	79	99	88
Medyan puan değeri	67,13	60,33	66	58	75,29	64,67
Ortalama puan değeri	65,09	59,95	65	56	74,84	65,36
Standart sapma	13,44	10,77	13	11	9,20	6,36
Çarpıklık katsayısı	-0,91	-0,63	-0,41	-0,67	-0,42	0,14
Basıklık katsayısı	0,98	0,65	-0,04	0,65	0,18	-0,12
Ortalama soru zorluk derecesi	0,68	0,62	0,65	0,56	0,75	0,65
Ortalama soru ayırt etme indeksi	0,32	0,26	0,31	0,25	0,22	0,14
Ortalama Nokta Çift Serili korelasyon katsayısı	0,35	0,27	0,32	0,28	0,26	0,17
KR20 güvenilirlik katsayısı	0,92	0,86	0,90	0,86	0,84	0,62
KR21 güvenilirlik katsayısı	0,89	0,81	0,87	0,81	0,77	0,42
Yüksek gruba ait minimum puan	65	64	65	61	82	69
Düşük gruba ait maksimum puan	68	56	62	54	69	62
Spearman-Brown katsayısı	0,88	0,81	0,82	0,74	0,78	0,58
Ortalama Çift Serili korelasyon katsayısı	0,52	0,39	0,47	0,39	0,37	0,20

KR: Kuder-Richardson

Yüksek not alan bireylerin %27'sinin en düşük not ortalaması 63 puan ve daha yüksek iken, düşük not alan bireylerin %27'sinin en yüksek not ortalaması 50 puan ve üstüdür.

Spearman-Brown güvenilirlik katsayısı bakımından incelendiğinde; Dönem I ve Dönem II ders kurul sınavlarının "iyi güvenilirlikte" oldukları, Dönem III ders kurulu, Dönem I ve Dönem II final sınavlarının "orta düzeyde güvenilir" oldukları, Dönem III final sınav sorularının ortalama "güvenirliğinin oldukça düşük" olduğu söylenebilir.

2015-2016 Eğitim-Öğretim Yılı

2015-2016 eğitim-öğretim yılında Dönem I, Dönem II ve Dönem III öğrencilerine uygulanan toplam 19 kurul sınavı ve 3 dönem sonu sınavı olmak üzere 22 sınav sonucu (toplam 2200 adet soru) incelenmiştir. Her bir sınav için ayrı ayrı uygulanan Madde Yanıt analiz sonuçlarının genel değerlendirmesi Tablo 2'de verilmiştir. Sorular ortalama soru zorluk derecesi bakımından incelendiğinde; Dönem I, Dönem II ve Dönem III Ders kurul sınavlarının ve Dönem III final sınavı sorularının "oldukça

kolay” nitelikte oldukları saptanmıştır. Dönem I ve Dönem II final sınav sorularının ise “orta düzeyde zor” oldukları görülmüştür (Tablo 2).

Sınav soruları soru ayırt etme indeksi bakımından incelendiğinde; Dönem I, Dönem II ders kurul sınav sorularının “oldukça iyi” oldukları, Dönem III ders kurulu sınavları, Dönem I final ve Dönem II final sınav sorularının ise “düzeltilmesi ve geliştirilmesi” gerektiği, Dönem III final sınavı sorularının ise “çok zayıf nitelikte sorular” oldukları saptanmıştır (Tablo 2).

Soru güçlük ve soru ayırt edicilik indeksleri birlikte değerlendirildiğinde;

- Dönem I ve Dönem II ders kurul sınavlarına ve Dönem I final sınavına ait soruların genel olarak “tipik iyi soru” özelliğinde olduğu,
- Dönem III final sınav sorularının “üzerinde çalışılması ve geliştirilmesi” gereken sorular olduğu,
- Dönem II final sorularının ise “zor fakat ayırt edici” sorulardan oluştuğu gözlenmiştir (Tablo 2).

Dönem I ve Dönem II ders kurul sınavlarının KR-20 güvenilirlik katsayı değerleri sırasıyla 0,92 ve 0,90 bulunmuştur. Bu değerlere göre ilgili dönemlere ait ders kurulu sınavlarının “mükemmel güvenilirlikte” oldukları görülmüştür. Dönem I ve Dönem II final sınavlarının KR-20 güvenilirlik katsayı değerleri 0.86’dır. Bu değerlere göre ilgili dönemlere ait final sınavlarının “iyi güvenilirlikte” olduğu söylenebilir. Dönem III ders kurul sınavının KR-20 katsayı değeri (0,84) “iyi güvenirligi” göstermektedir. Ancak Dönem III final sınavına ait KR-20 katsayısı 0,62 bulunmuştur ve bu sınavın “marjinal test (düzeltilebilir)” özelliği taşıdığı söylenebilir. Bu değer düşük çıkmasının nedenleri arasında çok kolay soruların olması, çok zor soruların olması, ayırt edicilik özelliği olmayan kötü yazılmış soruların varlığı aklı gelmektedir (Tablo 2).

Ortalama nokta çift serili korelasyon katsayısı değeri açısından incelendiğinde; Dönem I ve II Ders Kurul sınavlarının sorularının “iyi” soru özelliği taşıdığını, Dönem III ders kurulu sınavlarının ve Dönem I ve II final sınav sorularının “orta düzeyde” soru oldukları söylenebilir. Dönem III final sınav sorularının “zayıf” nitelikte sorular olduğu ve revize edilmeleri gerektiği açıktır (Tablo 2).

Soruların geçerliğini değerlendirmek için kullanılan ortalama çift serili korelasyon katsayıları incelendiğinde; uygulanan tüm sınavlara ait soruların genel olarak geçerli sorular oldukları söylenebilir (Tablo 2).

Yüksek not alan bireylerin %27’sinin en düşük not ortalaması 61 puan ve düşük not alan bireylerin %27’sinin en yüksek not ortalaması 56 puandır.

Spearman-Brown güvenilirlik katsayısı bakımından incelendiğinde; Dönem I, Dönem II ders kurul sınavlarının ve Dönem I final sınavlarının “iyi güvenilirlikte” oldukları, Dönem III ders kurulu ve Dönem II final sınavlarının “orta düzeyde güvenilir” oldukları, Dönem III final sınav sorularının oldukça “düşük güvenilirlikte” olduğu söylenebilir (Tablo 2).

TARTIŞMA VE SONUÇ

Bu çalışmada Düzce Üniversitesi Tıp Fakültesi 2014-2015 ve 2015-2016 eğitim-öğretim yıllarında ilk 3 döneme uygulanan sınav sorularının geçerlik ve güvenilirlikleri ayrıntılı bir şekilde incelenmiştir. Bu araştırmanın sonucunda; 2014-2015 ve 2015-2016 eğitim-öğretim yıllarında Dönem I ve Dönem II ders kurul sınavlarına ait soruların oldukça kolay düzeyde olduğu ve ayırt edici sorulardan oluştuğu, Dönem III ders kurul sınavına ait soruların oldukça kolay düzeyde olduğu ve düzeltilmesi gereken sorular olduğu söylenebilir. Dönem I, Dönem II final sınavlarına ait soruların ise orta düzeyde zor oldukları ve düzeltilmesi gereken sorular olduğu gözlenmiştir. Dönem III final sınavına ait soruların oldukça zor düzeyde olduğu ve çok zayıf nitelikte yani sınavda yer almaması gereken sorular olduğu söylenebilir.

Literatürde tıp fakültesinde uygulanan sınav sorularının (çoktan seçmeli, objektif yapılandırılmış klinik sorular-OYKS vb.)

geçerlik ve güvenilirliklerinin incelendiği çalışmalar bulunmamaktadır (7-10).

Özcan ve ark. (7) Gaziosmanpaşa Üniversitesi tıp fakültesinde yaptıkları çalışmada Dönem I, II ve III Ders kurul sınavlarının genel olarak %34’ünün “çok kolay”, %21’inin “oldukça kolay”, %32’sinin “orta düzey”, %9’unun “oldukça zor” ve %4’ünün ise “çok zor” sorulardan oluştuğunu bulmuşlardır. Ayrıca sınav sorularının %33,14’ünün “ayırt ediciliği düşük”, %24,20’sinin “ayırt ediciliği orta düzeyde ve düzeltilerek kullanılabilir”, %25,49’unun “küçük düzeltmelerle veya düzeltme yapmadan kullanılabilir”, %17,17’sinin ise “ayırt ediciliği yüksek aynen kullanılabilir” olduğunu belirlemişlerdir.

Tomak ve Bek (8) Ondokuz Mayıs Üniversitesi Tıp Fakültesi sınavlarında kullanılan çoktan seçmeli testlerin değerlendirmesini farklı madde analizi modellerini kullanarak yapmışlardır. İnceledikleri 84 soruluk sınavda soruların madde zorluk ve madde ayırt etme indekslerini sırasıyla 0,57 ve 0,277 olarak bulmuşlardır. Soruların çoğunun orta düzeyde zor ve ayırt ediciliği iyi düzeyde olduğunu ifade etmişlerdir.

Fallatah ve ark. (9) King Abdulaziz Üniversitesi tıp Fakültesinde 2012-2013 yıllarında yapılan yazılı ve klinik sınavların dahiliye bölümlerini içeren sınavları için geçerlik ve güvenilirlik çalışması yapmışlardır. Çoktan seçmeli sınav sorularının ve objektif yapılandırılmış klinik sınav sorularının geçerli ve güvenilir olduklarını saptamışlardır. Trejo-Mejia ve ark. (10) Meksika Ulusal Otonom Üniversitesi Tıp Fakültesinde intern öğrencilerine uygulan objektif yapılandırılmış klinik sınavların güvenilirliğini genelleştirilmiş teori ile incelemişlerdir. İlgili sınavların güvenilir olduklarını saptamışlardır.

Sınav niteliğini ve öğrenci başarısını değerlendirmek için uygulanan sınavların hem genel hem de soru bazında farklı teknikler kullanılarak değerlendirilmesi önerilmektedir. Ayrıca sınav sorularının kalitesini artırmak amacıyla tıp fakültesi öğretim üyelerine Tıp Eğitiminde Ölçme ve Değerlendirme konularında eğitimlerin verilmesi tavsiye edilmektedir.

KAYNAKLAR

1. Sayek İ, Odabaşı O. Türk Tabipleri Birliği Mezuniyet Öncesi Tıp Eğitimi Raporu 2010. Ankara: Türk Tabipleri Birliği Yayınları; 2010.
2. Erkuş A. Sınıf Öğretmenleri için Ölçme ve Değerlendirme: Kavram ve Uygulamalar. Ankara: Ekinoks Eğitim Danışmanlık Yayınevi; 2006.
3. Selvi H. Tıp Fakülterinde Uygulanan Sınavların Güvenirlik ve Geçerliğinin Sağlanmasına Yönelik Yapılan Çalışmalar: Mersin Üniversitesi Örneği. Tıp Eğitiminde İyi Uygulamalar Sempozyumu; 23-24 Ekim 2015, İzmir. s.48-52.
4. McGahee TW, Ball J. How to read and really use an item analysis. Nurse Educator. 2009;34(4):166-71.
5. McMorrisson S, Nibert A, Flick J. Critical thinking and test item writing. Houston: Health Education Systems; 2006.
6. Sadaf S, Khan S, Ali SK. Tips for Developing a Valid and Reliable Bank of Multiple Choice Questions (MCQs). Education for Health. 2012;25(3):195-7.
7. Özcan KV, Aydoğan Y, Bulut İ. Gaziosmanpaşa Üniversitesi Tıp Fakültesi’nde Uygulanan Çoktan Seçmeli Sınavların Betimsel Analizi. Gaziosmanpaşa Üniversitesi Tıp Fakültesi Derg. 2014;6(4):281-94.
8. Tomak L, Bek Y. Item analysis and evaluation in the examinations in the faculty of medicine at Ondokuz Mayıs University. Niger J Clin Pract. 2015;18(3):387-94.
9. Fallatah HI, Tekian A, Park YS, Al Shawa L. The validity and reliability of the sixth-year internal medical examination administered at the King Abdulaziz University Medical College. BMC Medical Education. 2015; 15: 10. doi:10.1186/s12909-015-0295-4.
10. Trejo-Mejía JA, Sánchez-Mendiola M, Méndez-Ramírez I, Martínez-González A. Reliability analysis of the objective structured clinical examination using generalizability theory. Med Educ Online. 2016;21:31650.