

Psikiyatri Eğitiminde Klinik Akıl Yürütmenin Değerlendirilmesi: Türkçe Otomatik Soru Üretimi ile Çoktan Seçmeli Soruların Geliştirilmesi



Esra EMEKLİ¹, Emre EMEKLİ², Yavuz Selim KIYAK³, Yasemin HOŞGÖREN ALICI⁴,
Özlem COŞKUN⁵, Işıl İrem BUDAKOĞLU⁶

ÖZET

Amaç: Bu çalışmanın amacı, psikiyatri alanında Türkçe olgu temelli çoktan seçmeli sorular (ÇSS) üretmek için otomatik soru üretimi (OSÜ) yönteminin uygunluğunu değerlendirmektir.

Yöntem: Çalışma tanımlayıcı olarak tasarlanmıştır. İlk aşamada konu başlıkları belirlenmiş ve alanın uzmanları tarafından bir bilişsel model oluşturulmuştur. İkinci aşamada soru şablonu oluşturularak değişkenler belirlenmiş, cevap şıklarının nasıl olması gerektiği düzenlenmiş ve farklı kombinasyonları içeren soru içeriğinin iki eşdeğer şablonu oluşturulmuştur. Son aşamada bu modeller temel alınarak Python tabanlı bir yazılım ile sorular üretilmiştir. Soruların üretilmesinin ardından, rastgele örnekler seçilmiş ve yapılandırılmış bir form aracılığıyla psikiyatri alanında tecrübeli eğitimciler tarafından değerlendirilmiştir.

Bulgular: Toplamda 1189 soru üretilmiş ve her bir tanı için 11 soru örneklendirilmiştir. Uzmanlar tarafından yapılan değerlendirmede soruların altısı her parametre için uygun olarak değerlendirilmiş, beş soruda minör düzeltme gerekliliği belirtilmiştir. Soruların tamamının ezber bilgiyi değil klinik akıl yürütme becerisini ölçtüğü ifade edilmiştir.

Sonuç: OSÜ yöntemi, tıp eğitiminin ihtiyaç duyduğu yüksek kaliteli soruların hızlı ve etkili bir şekilde üretilmesine olanak sağlamaktadır. Çalışma, Türkçe dilinde OSÜ ile klinik akıl yürütmeyi ölçen ÇSS üretiminin psikiyatri alanında uygulanabilir olduğunu göstermiştir. Bu yöntemle kısa sürede çok sayıda soru üretilmiş ve çeşitli kombinasyonlarla zenginleştirilmiştir.

Anahtar Sözcükler: Çoktan seçmeli sorular, klinik akıl yürütme, otomatik soru üretimi, psikiyatri eğitimi, tıp eğitimi

ABSTRACT

Assesment of Clinical Reasoning in Psychiatric Education: Development of Multiple-Choice Questions with Automatic Item Generation in Turkish

Objective: This study aims to evaluate the suitability of the automatic item generation (AIG) for producing Turkish case-based multiple-choice questions (MCQs) in psychiatry.

Method: The study was planned as a descriptive study. In the first stage, topics were determined and a cognitive model was created by subject matter experts. In the second stage, a question template was created, variables were determined, the format of answer options was organized, and two equivalent templates of question content with different combinations were created. In the final stage, questions were generated using Python-based software based on these models. Following the question generation, random samples were selected and evaluated by experienced educators using a structured form.

Results: A total of 1189 questions were generated, with 11 questions sampled for each diagnosis. In the evaluation conducted by experts, six of the questions were deemed appropriate for each parameter, while minor corrections were suggested for five questions. It was stated that all the questions assess clinical reasoning skills rather than factual recall.

Conclusion: The template-based AIG method allows for the rapid and effective production of high-quality questions needed in medical education. The study demonstrated that AIG in the Turkish language for generating MCQs that assess clinical reasoning is applicable in the field of psychiatry. This method enables the production of a large number of questions in a short time, enriched with various combinations.

Keyword: Automated item generation, clinical reasoning, medical education, multiple choice question, psychiatry education

Atf için: Emekli E, Emekli E, Kıyak YS, ve ark. (2025) Psikiyatri Eğitiminde Klinik Akıl Yürütmenin Değerlendirilmesi: Türkçe Otomatik Soru Üretimi ile Çoktan Seçmeli Soruların Geliştirilmesi. *Türk Psikiyatri Derg* 36:336–343. <https://doi.org/10.5080/u27540>

Geliş Tarihi: 12.06.2024, **Kabul Tarihi:** 29.11.2024, **Çevrim İçi Yayın Tarihi:** 07.04.2025

¹Psikiyatrist, ⁴Doç., Başkent Üniv., Tıp Fak., Psikiyatri AD., Ankara; ²Dr. Öğr. Üyesi, Eskişehir Osmangazi Üniv., Tıp Fak., Radyoloji AD., Eskişehir; ³Dr. Öğr. Üyesi, ⁵Doç., ⁶Prof., Gazi Üniv., Tıp Fak., Tıp Eğitimi ve Bilişimi AD., Ankara, Türkiye

Emre Emekli, e-posta: emreemekli90@gmail.com

GİRİŞ

Klinik akıl yürütme; hekimin bilgi edindiği, bu bilgiyi sentezlediği, hipotez ürettiği ve buradan yola çıkarak klinik yöntem, prognoz, tanı, tedavi, hasta bakımı planı oluşturduğu bir süreç olarak tanımlanmaktadır (Durning ve ark. 2013). Bunlara ek olarak klinik akıl yürütme, hekimin hastalık tanısı ve tedavisi süreçlerinde bir sorunla karşılaştığında gerçekleştirdiği zihinsel süreçler olarak da tanımlanabilir (Cate ve ark. 2018). Kritik bir beceri olduğundan, tıp eğitiminde klinik akıl yürütme becerisinin öğrenciler tarafından kazanılması elzemdir (Schmidt ve Mamede 2015). Bu becerinin ölçülmesi amacıyla tıp eğitiminde birçok ölçme-değerlendirme yöntemi geliştirilmiştir (Daniel ve ark. 2019).

Miller'in tıp eğitiminde kabul gören dört seviyeli (bilir - nasıl olduğunu bilir - nasıl olduğunu gösterir - yapar) piramidinin "nasıl olduğunu bilir - nasıl olduğunu gösterir - yapar" seviyelerindeki ölçme-değerlendirme yöntemleri ile klinik akıl yürütme becerisi değerlendirilebilir (Miller 1990). Ancak klinik öncesi öğrenciler için Miller'ın "nasıl olduğunu gösterir" ve "yapar" seviyelerine yönelik ölçme-değerlendirme faaliyetleri klinik ortam gerektirdiğinden, daha az uygulanmaktadır. Bu nedenle klinik akıl yürütme-yi ölçmeyi hedefleyen yöntemlerin birçoğu klinik döneme odaklanmıştır. Klinik dönem öncesi öğrencilerin hastalarla karşılaşmadan önce klinik akıl yürütme becerilerinin değerlendirilmesi için yazılı sınavlar, test sınavları ve probleme dayalı öğrenme (PDÖ) yöntemleri ön plana çıkmaktadır (Hawks ve ark. 2023). PDÖ'de öğrenciler gerçek hayattan alınan veya kurgulanan karmaşık sorunlarla karşılaşır, öğrenci merkezli bu eğitim yönteminde sorunlar çözülmeye çalışılırken klinik akıl yürütme becerilerinin de gelişmesine katkı sağlanır. Bu yöntemle klinik akıl yürütme becerilerinin daha derinlemesine ve uygulamalı olarak değerlendirilmesi mümkündür (Manuaba ve ark. 2022). Ancak bu yöntemin küçük grup dersleri için daha fazla eğitime ihtiyaç duyması, eğitici yetersizliği, öğrenci motivasyonunun yeterli olmaması gibi sorunlar klinik akıl yürütmenin değerlendirilmesi için ülkemizde test sınavlarının daha sık kullanılmasına neden olmaktadır (Rochmawati ve Wiechula 2010). Test sınavları ile öğrenciler tek seferde test edilebilir, standartlar belirlenebilir ve güvenilir puanlar elde edilebilir.

Çoktan seçmeli soruların (ÇSS) ezber yerine klinik akıl yürütme gibi üst düzey becerileri ölçmeye yönelik oluşturulması için en güncel tıbbi bilgilerin klinikle bütünleştirilerek dikkate alınması gerekir. Ancak, tıbbi bilginin yarılanma hızı sürekli ve hızlı bir şekilde artmaktadır (Densen 2011). Bu nedenle, tıp eğitimcilerinin soru üretiminde güncel ayak uydurması zor olabilir. ÇSS'lerin üretilmesi de zaman alıcıdır ve bu süreci otomatikleştirebilen uygulamalar tıp eğitimcileri için oldukça değerlidir. ÇSS üretiminin otomasyonunda iki temel yöntem vardır: Şablon tabanlı otomatik soru

üretimi (OSÜ) ve şablon tabanlı olmayan OSÜ (Gierl ve ark. 2021). Söz konusu iki yöntem de geleneksel ÇSS yazım sürecinden daha verimli bir süreç sağlamaktadır.

Şablon tabanlı olmayan yöntemler daha çok yapay zekâ araçlarına dayalıdır. Yapay zekâ kullanılarak ÇSS üretiminde verimlilik açısından umut verici sonuçlar elde edilmişse de (Cheung ve ark. 2023, Laupichler ve ark. 2023, Zuckerman ve ark. 2023, Emekli ve Karahan 2024, Kıyak ve Emekli 2024) yapay zekânın kalitesiz ve tutarsız bilgi üretebiliyor olması (Deng ve ark. 2023, Walker ve ark. 2023) ve hatta zaman zaman bilimsel kaynak uydurması (Masters 2023) nedeniyle önemli sorunlar ortaya çıkmaktadır. Dahası, yapay zekânın "kapalı kutu" olmasından, yani karmaşık algoritmalar tarafından üretilen çıktıların nasıl oluştuğunun insanlar tarafından henüz tam olarak anlaşılamamasından dolayı çıktıları üzerinde tam bir kontrol sağlamak mümkün değildir. Bu da bilgi yanlışlarının ve tutarsızlıkların kolayca düzeltilmemesine sebep olmaktadır. Yapay zekâ kullanan yöntemlere göre daha fazla emek gerektirse de, şablon tabanlı yöntem kullanılarak ÇSS üretmede bilgi yanlışlığı oluşma problemi daha az olmakta ve şablonlar üzerinden ÇSS oluşturma mekanizması insanlar tarafından direkt olarak gözlenebilir ve tam olarak kontrol edilebilir olduğu için hatalar kolaylıkla düzeltilabilmektedir (Gierl ve ark. 2021). Uygun bilişsel modellerin kullanılması durumunda bu yöntemin verimliliğinin ve üretilen soruların psikometrik özelliklerinin geleneksel yöntemlere benzer olduğu belirtilmektedir (Pugh ve ark. 2020). Ancak, bu yöntem, yapay zekâ tabanlı yöntemlere kıyasla daha fazla emek ve zaman gerektirmektedir.

Gierl ve ark. (2012) şablon tabanlı OSÜ sürecini üç aşamalı olarak tanımlamışlardır. Birinci adımda, soru üretimi için gerekli olan içerik, alanın uzmanları tarafından belirlenir. Bu içerik, hekimlerin belirli bir tıbbi tanıya ulaşmak için ihtiyaç duydukları bilgi, beceri ve problem çözme süreçlerini vurgulayan bir bilişsel model şeklinde ifade edilir. İkinci adımda, yeni sorular oluşturmak için bilişsel model içeriğinin oturtulması gereken bir soru modeli (şablon) geliştirilir. Bu soru modeli her yeni üretilen sorunun değişkenlerini (soru içeriği ve cevap seçenekleri) içeren bir şablon şeklinde oluşturulur. Bu şablonda yer alan değişkenlere atanacak kelimeler ya da kelime grupları da belirlenir. Son adımda bilgisayar tabanlı yazılımlar ile tek bir madde modelinden yüzlerce soru üretilebilir (Gierl ve Lai 2013, Gierl ve Lai 2015). OSÜ kullanımının başarılı sonuçlar verdiği farklı dillerde gösterilmiştir (Gierl ve ark. 2021). Tıp eğitiminde Türkçe olarak şablon tabanlı OSÜ yöntemiyle ÇSS üretmeye odaklanan sadece iki çalışma vardır ve ikisi de hipertansiyon konusunda yapılmıştır (Kıyak ve ark. 2023a, Kıyak ve ark. 2023b).

Psikiyatri, diğer branşlar gibi tanı koyma ve uygun tedaviyi belirlemek için doğru bir semiyoloji bilgisinin gerekli olduğu bir tıbbi uzmanlık alanıdır. Bunun yanı sıra psikiyatrik semiyoloji oldukça karmaşıktır ve güncellenen tanı sistemleri

(Diagnostic and Statistical Manual of Mental Disorders - DSM-5, International Classification of Disease - ICD-11) nedeniyle tanı ve terminolojiler de sürekli değişmektedir (Rejon 2012). Laboratuvar tetkikleri ve görüntüleme gibi biyobelirteçler psikiyatride tanı koyma aracı olarak daha az sıklıkta kullanılabilir. Psikiyatrik görüşmeler ile vaka değerlendirmesi tanı koymada esas belirleyici olmaktadır. Son yıllarda tıp eğitiminde benimsenen güncel yaklaşımlardan biri, semptomdan hastalığa doğru ilerleyen öğretim modelidir (Moghadami ve ark. 2021). Bu modelde, öğrencilerin öncelikle semptomları tanımlamaları ve bu semptomlara sahip olabilecek olası tanıları zihinsel olarak canlandırmaları beklenmektedir. Ülkemizde de tıp fakültesi eğitim programlarının temelini oluşturan Ulusal Çekirdek Eğitim Programı (UÇEP) çerçevesinde, bu yaklaşım desteklenmektedir (Ulusal Çalışma Grubu 2020). Bu açıdan vaka tabanlı eğitim, psikiyatri alanında tanı ve ayırıcı tanı becerisinin öğrencilere kazandırılmasında daha da önem taşımaktadır. Bilindiği kadarıyla psikiyatri alanında soru üretimine olanak veren bu alanda yapılan bir çalışma mevcut değildir.

Bu çalışma, tıp eğitiminde ihtiyaç duyulan yüksek kaliteli soruların üretim sürecini hızlandırmak ve klinik akıl yürütme becerilerini ölçen soruların otomatik olarak üretilmesinin uygulanabilirliğini değerlendirmeyi amaçlamaktadır.

YÖNTEM

Çalışma Tasarımı

Çalışma, tanımlayıcı bir çalışma olarak planlanmıştır. Bu çalışma, insan katılımcıları veya hassas kişisel veriler içermediği için etik kurul onayı gerektirmemektedir. Çalışma, otomatik olarak üretilen soruların geliştirilmesine odaklanmış ve değerlendirme anonim uzman görüşleri ile yapılmıştır.

Soruların Üretilmesi

Birinci aşama: Bilişsel modelin oluşturulması.

Öncelikle tıp eğitiminde doktora derecesine sahip bir hekim (yazar 2) ve bir psikiyatri asistan hekimi (yazar 1) tarafından tıp fakültesi öğrencilerine yönelik olarak psikiyatri soruları hazırlanması kararlaştırılmıştır. Bu aşamada Ulusal Çekirdek Eğitim Programı (Ulusal Çalışma Grubu 2020) ve Başkent Üniversitesi Tıp Fakültesi ders programı temel alınarak hazırlanacak soruların konu başlıkları belirlenmiştir. Konu başlıkları DSM-5'te yer alan toplamda 15 tanıdan oluşturulmuştur (American Psychiatric Association, 2013). Ortak ve ön planda olan klinik özelliklerine göre (duygudurum dönemi barındırma/ bedensel belirtilerin eşlik etmesi/ psikotik belirtiler barındırma) şeklinde üç tanı grubu belirlenmiştir (birinci, ikinci, üçüncü tanı grubu):

Birinci tanı grubu (depresyon, distimi, bipolar bozukluk, postpartum depresyon, şizoafektif bozukluk),

İkinci tanı grubu (konversiyon bozukluğu, bedensel belirti bozukluğu, hastalık kaygısı bozukluğu, yapay bozukluk, bakım verenin yapay bozukluğu),

Üçüncü tanı grubu (kısa psikotik bozukluk, şizofreniform bozukluk, şizofreni, madde kullanımına bağlı psikoz, postpartum psikoz)

Semptom kümeleri de DSM-5'te belirtilen tanı kriterlerindeki temel semptomlara dayanmaktadır. Her tanı grubu için ortak belirtiler listelendikten sonra, bu belirtilerden rastgele atamalar ile temel bir belirti kümesi oluşturulmuştur. Ardından, uygun sosyodemografik veriler, süre ve ek ayırtıcı özellikler eklenerek en uygun tanının bulunmasına yönelik sorular tasarlanmıştır. Bu yaklaşım, tıp eğitiminde semptomdan hastalığa ilerleyen öğretim modeline uygun olarak, öğrencilerin klinik akıl yürütme becerilerini geliştirmeyi hedeflemektedir (Moghadami ve ark. 2021). Tanı ve semptom gruplamaları, soru üretimini kolaylaştırmak amacıyla benzer ve ilişkili semptomları paylaşan tanıları bir araya getirilerek yapılmıştır. Bu nedenle, bazı tanıların gruplandırılmasında zorluklar yaşansa da mutlak bir sınıflamadan ziyade beşinci sınıf tıp fakültesi öğrencilerine yönelik belirtiden tanıya uygun soru üretimi hedeflenecek şekilde bu gruplar pragmatik bir yaklaşımla oluşturulmuştur.

Sonraki aşamada bilgi kaynakları olarak hastanın bilgileri (yaş, cinsiyet, alkol/madde kullanım durumu) ve hastalığın bilgileri (semptom, öykü, süre) belirlenmiştir. Belirlenen bilgi kaynakları temel alınarak her bir tanı için uygun yaş aralıkları, uygun cinsiyet ve alkol/madde kullanım durumu (madde kullanmadığı/sosyal içici olarak alkol tükettiği ve madde kullanmadığı/her gün alkol tükettiği ve sıklıkla madde kullandığı) belirlenmiştir. Her bir tanı için olası semptomlar yazılmıştır. Semptomlar arasından her bir tanı için benzer olanlar gruplandırılarak semptom grupları oluşturulmuştur (Semptom A, B, C, D). Benzer şekilde hastalığın öyküsü için birden fazla tanıyı kapsamayı hedeflenerek tanıları için hastalık öyküleri yazılmıştır. Öyküler gruplandırılmıştır (Öykü A, B, C, D). Son olarak ise her bir hastalığın DSM-5 tanı kriterleri dikkate alınarak hastalığın senaryoda ne kadar süredir olmasının tıbbi açıdan uygun olduğu bilgisi eklenmiştir.

Bu aşama sonunda şizoafektif bozukluk tanısı, bulunduğu gruptaki tanıları gibi duygudurum belirtilerine sahip olmasına karşın psikotik özellikleri ile ön plana çıktığı için ve soru şablonunda ortak olarak verilmesi güç olduğu için soru üretiminden çıkarılmıştır. Yapay bozukluk, hastalık kaygısı bozukluğu ve bakım verenin yapay bozukluğu tanısına götürecek belirleyiciler benzer şekilde soru kökünde otomatik soru üretimine engel oluşturacak şekilde karmaşıklığa yol açtığı için bu tanıları da soru üretiminden çıkarılmıştır. Bu dört tanı yalnızca cevap sıklarında çeldirici şık vazifesi görecek şekilde kullanılmıştır.

Tablo 1. Bilişsel Model Bileşenleri

Hastalık Grubu	Tanı	Yaş	Cinsiyet	Semptom A	Semptom B	Semptom C	Semptom D	Öykü A	Öykü B	Öykü C	Öykü D	Süre
Grup 1	Depresyon	18-65	Kadın/Erkek	Hayır	Hayır	Evet	Hayır	Evet	Hayır	Hayır	Hayır	1-12 ay
	Distimi	18-65	Kadın/Erkek	Hayır	Hayır	Evet	Hayır	Evet	Hayır	Hayır	Hayır	2-4 yıl
	Bipolar bozukluk	18-65	Kadın/Erkek	Hayır	Hayır	Evet	Hayır	Hayır	Hayır	Hayır	Evet	1-24 ay
	Postpartum depresyon	18-35	Kadın	Hayır	Hayır	Evet	Hayır	Hayır	Evet	Hayır	Hayır	2-3 hafta
Grup 2	Somatizasyon	18-65	Kadın/Erkek	Evet	Hayır	Hayır	Hayır	Hayır	Hayır	Evet	Hayır	1-24 ay
	Konversiyon	18-65	Kadın/Erkek	Hayır	Hayır	Hayır	Evet	Hayır	Hayır	Evet	Hayır	1-24 ay
Grup 3	Kısa Psikotik Bozukluk	18-65	Kadın/Erkek	Hayır	Evet	Hayır	Hayır	Evet	Hayır	Hayır	Hayır	1-29 gün
	Şizofreniform Bozukluk	18-65	Kadın/Erkek	Hayır	Evet	Hayır	Hayır	Evet	Hayır	Hayır	Hayır	1-5 ay
	Şizofreni	18-65	Kadın/Erkek	Hayır	Evet	Hayır	Hayır	Evet	Hayır	Hayır	Hayır	6-24 ay
	Madde Kullanımına Bağlı Psikoz	18-65	Kadın/Erkek	Hayır	Evet	Hayır	Hayır	Evet	Hayır	Hayır	Hayır	1-24 ay
	Postpartum Psikoz	18-35	Kadın	Hayır	Evet	Hayır	Hayır	Hayır	Evet	Hayır	Hayır	2-3 hafta
Semptom A	terleme, çarpıntı, ateş basması, nefes darlığı, bayılma hissi, ellerde titreme, bulantı, sıkıntı hissi, karın ağrısı											
Semptom B	takip edildiği, fikirlerinin çalındığı, peygamber olduğunu düşünme, ne yapması gerektiğinin söylendiği, beynine fikir sokulduğu, kulağına bir şeyler söylendiği, bir yerlerden sesler duyma, kendisine zarar verileceğini düşünme, kendi kendine konuşma											
Semptom C	üzüntülü hissetme, sık sık ağlama, çaresiz hissetme, boşlukta hissetme, umutsuzluk, hiçbir şeyden zevk almama, günlük aktiviteleri yapmak istememe, iştahsızlık, gelecekte bir beklentisi olmadığı, karamsar olma, yataktan çıkmak istememe, değersiz olduğunu düşünme, ölmek isteme											
Semptom D	konuşamama, bayılma, ayaklarında güçsüzlük, her iki kolda güçsüzlük, ellerde hissizlik, yutamama, konuşmada zorlanma, yürüyememe, görmeme, duymama, ses çıkaramama, dengesini sağlayamama											
Öykü A	bir süredir evden dışarı çıkmak istemediğini, özbakımının azaldığını											
Öykü B	yakın zamanda doğum yaptığını											
Öykü C	benzer yakınmalarla acil servise başvurduğunu ve kendisine tıbbi bir hastalığının olmadığını söylediğini											
Öykü D	bazı zamanlar az uyuduğunu, yerinde duramayacak kadar çok enerjik hissettiğini ve aşırı para harcadığı dönemlerin olduğunu											

Sonuç olarak geriye kalan 11 tanıdan soru üretilmesine karar verilmiştir. Bilişsel modelin ilk taslağı oluşturulduktan sonra psikiyatri öğretim üyesi hekim (yazar 4) ve tıp eğitiminde doktora seviyesine sahip bir tıp hekimi (yazar 3) tarafından bilişsel model gözden geçirilmiştir. İki yazar tarafından semptomlar, öyküler gözden geçirilerek bazı düzeltmeler yapılmıştır. Bu düzeltmelerle birlikte bilişsel modelin son haline karar verilmiştir (Tablo 1).

İkinci aşama: Soru modelinin (şablonun) oluşturulması.

Oluşturulan modelin soru üretimi için uygun formata getirilmesi amaçlanmıştır. Bu amaçla ilk etapta örnek bir soru (parent item) oluşturulmuştur. Daha sonra bu örnek sorudan bir soru şablonu oluşturulmuştur. Soru şablonunun içeriği beş parçaya bölünmüştür (başvuru nedeni, öykü, süre, alkol-madde kullanım bilgisi, soru cümlesi). Üretilen soruların birbirlerine benzerliğini azaltmak amacıyla, her bir soru içeriği parçası için iki farklı şekilde soru cümlesi aynı anlama gelecek şekilde yazılmıştır. Böylelikle, farklı kelimelerle aynı ifadenin elde edilebilmesi sağlanmıştır. Her üretilen soru

için, bu soru şablonlarından birinin kullanılarak soruların oluşturulması hedeflenmiştir (Tablo 2).

Çalışmada ortak ve ön planda olan klinik özelliklerine göre üç tanı grubu ilk aşamada belirlenmiştir (birinci, ikinci, üçüncü tanı grubu). Oluşturulan sorularda ise bu tanı gruplarından yola çıkarak her bir senaryo için tanının dâhil olduğu gruptaki beş tanının karışık olarak cevap şıkkı olarak atanması hedeflenmiştir.

Üçüncü aşama: Yazılım kullanılarak soruların üretilmesi.

Bu aşamada önceki aşamalarda ortaya konulan bilgilerden yola çıkarak soruları üretmek üzere bir yazılım oluşturulması gerekmektedir. Bu nedenle yazarlar tarafından bir yazılımcı yardımıyla Python 3.1 (Rossum ve Drake, 1995), tabanlı tek kullanımlık bir kod oluşturulmuştur. Bu sayede Tablo 1'de verilen semptom ve öyküyü tarif eden kelimeler Tablo 2'de verilen şablonlarda yer alan ilgili yerlere otomatik olarak atanarak sorular üretilmiştir. Cevap seçenekleri de önceki aşamada ifade edildiği şekilde belirlenmiştir (Tablo 3).

Tablo 2. Soru Modeli

ÜRETİLECEK SORU İÇERİĞİ: [Başvuru Nedeni] [Öykü] [Süre] [Alkol] [Soru Cümlesi]

	Şablon 1	Şablon 2
Başvuru Nedeni	<YAŞ> yaşındaki <CİNSİYET> hasta, <SEMPOTOM1> ve <SEMPOTOM2> şikayetleri ile psikiyatri polikliniğine başvuruyor.	<SEMPOTOM1> ve <SEMPOTOM2> yakınmaları nedeniyle ruh sağlığı ve hastalıkları uzmanına başvuran <YAŞ> yaşındaki <CİNSİYET> hasta.
Öykü	Alınan öyküsünde hasta <ÖYKÜ> ifade ediyor.	Öyküsü, hastanın <ÖYKÜ> ortaya çıkarıyor.
Süre	Bahsettiği şikayetlerin <SÜRE> gibi bir süredir olduğunu söylüyor.	Hasta, yakınmalarının <SÜRE> gibi bir süredir mevcut olduğunu belirtiyor.
Alkol ve Madde	Alkol ve madde kullanımı sorgulandığında <ALKOL> ifade ediyor.	Alkol ve madde kullanıp kullanmadığı sorulduğunda <ALKOL> belirtiyor.
Soru Cümlesi	Bu hastada, aşağıda verilen tanılardan hangisi diğerlerine göre daha muhtemeldir?	Aşağıdakilerden hangisi bu hasta için diğerlerine göre en olası tanıdır?

Tablo 3. Depresyon Tanısı İçin Üretilen Soru Örnekleri**Soru 1**

25 yaşındaki kadın hasta, üzüntülü hissetme ve sık sık ağlama şikayetleri ile psikiyatri polikliniğine başvuruyor. Öyküsü, hastanın bir süredir evden dışarı çıkmak istemediğini, özbakımının azaldığını ortaya çıkarıyor. Hasta, yakınmalarının 6 ay gibi bir süredir mevcut olduğunu belirtiyor. Alkol ve madde kullanımı sorgulandığında kullanmadığını ifade ediyor. Bu hastada, aşağıda verilen tanılardan hangisi diğerlerine göre daha muhtemeldir?

Panik bozukluk

Distimi

Postpartum depresyon

Bipolar bozukluk

Depresyon

Soru 2

İştahsızlık ve karamsar olma yakınmaları nedeniyle 19 yaşındaki kadın hasta ruh sağlığı ve hastalıkları uzmanına başvuruyor. Öyküsü, hastanın bir süredir evden dışarı çıkmak istemediğini, özbakımının azaldığını ortaya çıkarıyor. Hasta, yakınmalarının 10 ay gibi bir süredir mevcut olduğunu belirtiyor. Alkol ve madde kullanıp kullanmadığı sorulduğunda kullanmadığını belirtiyor. Aşağıdakilerden hangisi bu hasta için diğerlerine göre en olası tanıdır?

Panik bozukluk

Postpartum depresyon

Distimi

Depresyon

Bipolar bozukluk

Soru 3

Ölmek isteme ve günlük aktiviteleri yapmak istememe yakınmaları nedeniyle 49 yaşındaki kadın hasta ruh sağlığı ve hastalıkları uzmanına başvuruyor. Alınan öyküsünde hasta bir süredir evden dışarı çıkmak istemediğini, özbakımının azaldığını ifade ediyor. Hasta, yakınmalarının 2 ay gibi bir süredir mevcut olduğunu belirtiyor. Alkol ve madde kullanımı sorgulandığında kullanmadığını ifade ediyor. Aşağıdakilerden hangisi bu hasta için diğerlerine göre en olası tanıdır?

Depresyon

Distimi

Panik bozukluk

Bipolar bozukluk

Postpartum depresyon

Üretilen Soruların Değerlendirilmesi

Üretilen soruların sayısı fazla olduğu için bunlar içerisinde Statistical Package for the Social Sciences (SPSS versiyon 22.0) yardımı ile her bir tanıyı içerecek şekilde örneklem yapılmıştır. Seçilen sorular yazarlar tarafından oluşturulan bir değerlendirme formu ile eğitim tecrübesi olan psikiyatristler ile paylaşılmıştır (ek dosyaya bakınız). Değerlendirme formu literatürde var olan değerlendirme formları temel alınarak araştırmacılar tarafından oluşturulmuştur (Pugh ve ark.

2020). Değerlendirme kriterleri ilk beş soruda evet/hayır sorusu olacak şekilde toplamda altı sorudan oluşmuştur:

1. Soru metni anlaşılabilir.
2. Soru, klinik açıdan uygundur.
3. Sorunun tek bir doğru cevabı vardır.
4. Soruda verilen bilgiler, doğru cevabı bulmak için yeterlidir.
5. Çeldiriciler mantıklıdır.

6. Soru esas olarak şunu ölçmeye yöneliktir (lütfen aşağıdaki iki seçenekten birini işaretleyiniz): Ezber bilgiyi () ; Klinik akıl yürütme becerisini ()

Bunlara ek olarak yazarlardan sorularla ilgili olarak varsa görüşlerini belirtmeleri istenmiştir.

BULGULAR

Toplamda 1189 soru üretilmiştir. Üretilen sorular içerisinde her bir tanı için bir tane olmak üzere 11 soru örneklendirilmiştir (ek dosyaya bakınız). Tıp fakülteleri için soru hazırlama tecrübesi olan üç psikiyatrist soruları değerlendirmiştir. Psikiyatristlerin ortalama klinik tecrübesi 7 yıldır.

2., 3., 5., 9., 10., 11. sorular üç değerlendirici tarafından da her bir parametre için uygun olarak değerlendirilmiştir. İkinci değerlendirici soru 1, 4, 6, 8 için, üçüncü değerlendirici soru 4 ve 7 için önerilerde bulunmuştur. Soru 1 ve 8 içinse ikinci değerlendirici “Sorunun tek bir doğru cevabı vardır.” ve “Soruda verilen bilgiler, doğru cevabı bulmak için yeterlidir.” sorularına “hayır” yanıtını vermişlerdir. Tüm soruların ise klinik akıl yürütmeyi ölçtüğü ifade edilmiştir.

Uzmanların yaptığı açık uçlu yorumlar şu şekildedir:

Soru 1 için “Hastanın yaşı ve cinsiyeti dikkate alındığında postpartum da doğru cevap olarak düşünülebilir. Bu nedenle kadın hastalar için hastanın yaşı daha yüksek belirlenebilir” (Değerlendirici 2).

Soru 4 için iki değerlendirici benzer yorum yapmıştır. “Madde kullanımına bağlı psikoz demek için madde kullanımının zamanı ile psikotik belirtilerin başlangıcı arasındaki ilişkinin daha net belirtilmesi daha uygun olabilir.” (Değerlendirici 3).

Soru 6’da “Sağlık ve bedensel belirtiyeye yönelik uğraş ile ilgili daha net bilgi verilmesi gerekmektedir” (Değerlendirici 2).

Soru 7 için “Konversiyon tanısına ulaşılması için daha net dışlama kriterlerine yer verilmesi gerekmektedir.” (Değerlendirici 3). Soru 8 için “Distimi ve depresyon tanısı sorgulanan hasta için geçmişte manik atak dönemi olup, olmadığı belirtilmesi uygundur (Değerlendirici 2).

TARTIŞMA

Psikiyatri eğitiminde olgu temelli öğrenme gerek pratik gerek ise teorik öğrenmede önemli bir yer tutmaktadır (Morreale 2019). Bu alanda yazılı senaryolar üzerinden, video gösterimleri, hasta simülasyonları ya da gerçek hastaların aynalı oda üzerinden izlenmesi şeklinde farklı yöntemler ile tıp fakültesi öğrencilerinin psikiyatrik vakaları gözlemlemesi; bu sayede tanı ve ayırıcı tanı becerilerini kazanması hedeflenmektedir (Hassoulas 2017). Aynı zamanda olgu örnekleri, teorik ve pratik sınavlarda değerlendirme aracı olarak da

kullanılmaktadır. Ülkemizdeki tıp fakültelerindeki psikiyatri eğitimi (staj sıklığı, soru miktarı vs.) dikkate alındığında büyük bir soru yükü olduğu gözlenmektedir. Bu çalışmada psikiyatri alanında Türkçe olarak olgu temelli klinik akıl yürütmeyi ölçen ÇSS üretilmesi amaçlanmıştır. Bu amaçla öncelikle bilişsel bir model oluşturulduktan sonra bir soru taslağı oluşturularak Python 3.1 (Rossum ve Drake, 1995), temelli bir kod yardımıyla kısa sürede 1189 soru üretilmiştir. Böylece psikiyatri eğitiminde hem sınavlarda hem öğrencilerin sınava yönelik çalışmalarında kullanılmak üzere farklı klinik durumlara yönelik çok sayıda olgu temelli soru üretimi ilk kez mümkün kılınabilmektedir.

Örneklenen 11 sorunun değerlendirmesi ise soru hazırlama tecrübesi olan psikiyatristler tarafından yapılmıştır. Bu değerlendirme sonucunda değerlendirme formu parametrelerinden her biri için altı soru sınavlarda kullanmak için tamamen uygun olarak değerlendirilmiş, dört soru için birer değerlendirici, bir soru için ise iki değerlendirici düzeltilmesi gerektiği yönünde görüş bildirmiştir. Bu sonuçlarla bazı sorularda düzeltmeye gidilmesi gerekse de ilk kez Türkçe olarak OSÜ ile soru üretiminin psikiyatri alanında klinik akıl yürütmeyi ölçmeye yönelik olarak gerçekleştirilebileceği gösterilmiştir. Üstelik verilen öneriler algoritmanın herhangi bir kısmındaki soru kökünün değiştirilmesi ile düzenlenebilir niteliktedir. Bu bulgu soru üretiminin temel algoritmasının doğru işlediğini göstermektedir. Literatürde İngilizce, Fransızca, Çince, İspanyolca ve Korece OSÜ ile sorular üretilmiş ve başarılı kullanımı gösterilmiştir (Gierl ve ark. 2021). Ancak psikiyatri alanında OSÜ ile yapılmış bir çalışmaya literatürde rastlanmamıştır. Literatürde yapılan çalışmalar cerrahi, ilaçlar, yenidoğan sarılığı, üst gastrointestinal kanama, yetişkinlerde karaciğer hastalığı ve acil tıp üzerine yapılmıştır.

Bu çalışmada soru üretme alanında tecrübeli psikiyatri hekimleri tarafından üretilen soruların değerlendirilmesi ve kabul görmesi ile bu zorluğun aşılması hedeflenmiştir. Sonuç olarak OSÜ kullanımının psikiyatri alanında da uygulanabilirliğinin gösterilmiş olması nedeniyle bu çalışma önemlidir.

Tüm sorular için değerlendiriciler soruların klinik akıl yürütmeyi ölçtüğünü ifade etmişlerdir. Yüksek kaliteli ÇSS yazmak tıp fakülteleri için yoğun kaynak gerektiren bir iştir. Literatürde özellikle klinik akıl yürütme becerilerini ölçmek gibi üst düzey değerlendirme yapmayı hedefleyen ÇSS’lerin geliştirilmesinin tıp eğitimcilerinin vaktinin önemli bir kısmını harcamasına neden olduğu ifade edilmektedir (Schuwirth ve van der Vleuten 2004). Tıp fakültelerinde çok fazla sınav ve ders içeriği olması nedeniyle soru bankalarından ihtiyaç duyulan soru sayısı çok fazladır (Wrigley ve ark. 2012). Bunun yanı sıra özellikle Türkiye’de eğitici pozisyonundaki doktorların klinik sorumlulukları da göz önünde bulundurulduğunda soru hazırlamaya harcanan vakit ciddi bir iş yüküne sebep olmaktadır. Bu nedenlerle kısa sürede çok sayıda sınavda

kullanılabilecek soruların üretilmesi önemlidir. Bu çalışma ile kısa sürede çok sayıda kaliteli soru üretebilmenin de mümkün olduğu gösterilmiştir.

Bu çalışmanın literatüre diğer önemli katkısı soru içeriğinde farklı kombinasyonların kullanılması ile soru köklerinin çeşitlendirilmesidir. Literatürde OSÜ kullanılarak Türkçe ÇSS üretmek üzerine yapılan çalışmalara bakıldığında soru kökünün ve cevap seçeneklerinin hep aynı olduğu görülmektedir (Kıyak ve ark. 2023a, Kıyak ve ark. 2023b). Bu çalışmada aynı tanıyı sorgulayan soruların birbirine benzer olma olasılığının azaltılması da hedeflenmiştir. Soru şablonlarında varyasyonlar ve farklı semptom/öykü kombinasyonları kullanılarak bu benzerlik en aza indirilmeye çalışılmıştır. Buna ek olarak ikişer varyasyon içeren beş alt başlığa bölünmüş (başvuru nedeni, öykü, süre, alkol-madde kullanımı, soru cümlesi) eşdeğer iki ayrı şablon oluşturulmuştur. Bu sayede aynı tanıyı sorgulayan soru için bile toplamda 32 farklı formatta soru içeriğinin oluşturulması sağlanmıştır. Ancak, yine de aynı tanıya ait soruların bazı ortak özellikler taşıması kaçınılmazdır. Soruların benzerliğini ölçen özel bir algoritma kullanılmamıştır. Gelecekteki çalışmalarda, ileri istatistiksel teknikler ve doğal dil işleme teknikleri kullanılarak soruların benzerliğini ölçen ve minimize eden algoritmaların geliştirilmesi faydalı olacaktır.

Bu çalışmadaki soru grupları duygudurum bozukluğu belirtilerinin bulunması, bedensel belirtilerin bulunması, psikotik özelliklerin bulunması gibi temel belirtilere dayanmaktaydı. Bu temel mantık ve soru üretme teknolojisi ile psikiyatride vaka tabanlı soru üretiminin mümkün olabildiği sadece üç belirti grubu üzerinden test edilmiştir. Ancak psikiyatrideki zengin tanı grupları ve ortak semptomatoloji düşünüldüğünde bundan sonra yapılacak kapsamlı çalışmalarda bu gruplar genişletilebilir.

Çalışmanın güçlü yönlerinin yanı sıra bazı kısıtlılıkları da bulunmaktadır. İlk olarak oluşturulan sorular içerisinden her bir tanıyı temsil edecek şekilde rastgele örneklem ile sadece 11 soru belirlenmiş ve değerlendirilmiştir. Bu yaklaşım, her tanı için bir örnek soru üzerinden genel bir değerlendirme yapmamıza olanak sağlamıştır. Ancak, bu sınırlı sayıdaki soru, tüm soru evrenini tam olarak temsil etmiyor olabilir. Bunun sebebi, araştırma sürecinde kaynakların sınırlı olması ve değerlendirme sürecinin zaman alıcı olmasıdır. Özellikle, uzman değerlendiricilerin zaman kısıtlamaları nedeniyle daha fazla sorunun incelenmesi pratik olarak mümkün olmamıştır. Ayrıca, çalışma, metodolojik bir ön değerlendirme niteliği taşıdığından, ilk aşamada daha küçük bir örneklemle soruların genel uygunluk ve kalite düzeyini belirlemek amaçlanmıştır. İkinci olarak üretilen soruların benzerliğini en aza indirmek için soru şablonlarında varyasyonlar ve farklı semptom/öykü kombinasyonları kullanılmıştır. Yine de, aynı tanıya ait soruların bazı ortak özellikler taşıması kaçınılmazdır. Gelecekteki çalışmalarda, ileri istatistiksel teknikler ve doğal

dil işleme yöntemleri kullanılarak soruların benzerliğini ölçen ve minimize eden algoritmaların geliştirilmesi bu benzerlikleri en aza indirmek için kullanılabilir. Diğer bir kısıtlılık olarak değerlendirici sayısı azdır ancak değerlendiricilerin hepsi üniversitelerde çalışan eğitici pozisyonunda ve soru hazırlama tecrübesi olan kişilerdir. Ayrıca soruların cevap şıklarının oluşturulmasında çeldiricilerin kalitesini artırmak için özel bir algoritma belirleme yöntemi (Lai ve ark. 2016) kullanılmamıştır. Ancak bu sorun ortak belirti kümesine sahip psikiyatrik hastalıklar için gruplar oluşturularak, cevap şıklarına ortak belirti kümesindeki diğer tanıların atanması şeklinde gerçekleştirilmiştir. Diğer bir kısıtlılık olarak bu çalışmada, anksiyete bozuklukları ve psikotik özellik gösteren duygudurum bozuklukları dahil edilmemiştir. Anksiyete bozuklukları, semptomlarının diğer tanı gruplarıyla örtüşmesi ve çok sayıda alt tipi içermesi nedeniyle bilişsel model ve soru şablonlarının karmaşıklığını artırmaktadır. Benzer şekilde, psikotik özellikli duygudurum bozukluklarının dahil edilmesi, semptom kombinasyonlarını ve olası tanı seçeneklerini önemli ölçüde genişleteceği için çalışma kapsamının dışında bırakılmıştır. Bu durum, çalışmamızın önemli bir kısıtlılığı olup, gelecekteki araştırmalarda bu tanı gruplarının da dahil edilmesi faydalı olacaktır. Ayrıca tanı ve semptom gruplarının pragmatik amaçlarla yapılmış olup, belirli sınıflandırma sistemlerine tam olarak uymayabilir, ancak eğitimsel hedefler doğrultusunda bu yaklaşımın benimsenmiştir. Son olarak ise sorular öğrencilere uygulanmamıştır. Dolayısıyla öğrenciler tarafından sorunun doğru cevaplanma oranını ifade eden zorluk indeksi ve sorunun yüksek performans gösteren öğrenciler ile düşük performans gösteren öğrenciler arasında ne kadar iyi ayırım yapabildiğini gösteren ayırt edicilik indeksi gibi psikometrik özellikleri (Tavakol ve ark. 2024) değerlendirilmemiştir. İlerleyen çalışmalarda, bu yöntemle üretilen soruların öğrencilere uygulanması ve bu psikometrik analizlerin yapılması planlanmaktadır. Bu sayede, soruların gerçek dünyadaki uygulamaları ve etkinlikleri daha net bir şekilde ortaya konulabilecektir.

Bu çalışma, psikiyatride Türkçe tıp eğitimi literatürüne önemli bir katkı sağlayarak, şablon tabanlı OSÜ yönteminin psikiyatri alanında klinik akıl yürütmeyi ölçen olgu temelli soruların hızlı ve etkili bir şekilde üretilmesinde kullanılabileceğini ortaya koymuştur. Ayrıca, soru içeriğinde çeşitlilik sağlanarak soru köklerinin farklı kombinasyonlar ile oluşturulması, soru bankalarının zenginleştirilmesine ve öğrencilerin değerlendirilmesinde daha dinamik ve kapsamlı bir yaklaşımın benimsenmesine olanak tanımaktadır.

EK

https://www.turkpsikiyatri.com/upload/39_27540_TR_EK.pdf

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazarlar herhangi bir çıkar çatışması olmadığını beyan etmişlerdir.

Finansal Destek: Yoktur.

Teşekkür: Soruların oluşturulmasında kullanılan Phyton kodunun yazılmasındaki desteklerinden ötürü Mehmet Ali Akyol'a teşekkür ederiz.

KAYNAKLAR

- American Psychiatric Association. (2013) Diagnostic and statistical manual of mental disorders (5th ed.).
- Cate O, Custers E, Durning SJ (2018) Principles and Practice of Case-based Clinical Reasoning Education: A Method for Preclinical Students. Cham, Springer Copyright, s. 75-83.
- Cheung BHH, Lau GKK, Wong GTC ve ark. (2023) ChatGPT versus human in generating medical graduate exam multiple choice questions—A multinational prospective study (Hong Kong S.A.R., Singapore, Ireland, and the United Kingdom). PLOS ONE 18: e0290691.
- Daniel M, Rencic J, Durning SJ ve ark. (2019) Clinical reasoning assessment methods: a scoping review and practical guidance. Acad Med 94: 902-12.
- Deng J, Zubair A, Park Y-J (2023) Limitations of large language models in medical applications. Postgrad Med J 99: 1298-9.
- Densen P (2011) Challenges and opportunities facing medical education. Trans Am Clin Climatol Assoc 122: 48-58.
- Durning SJ, Artino AR, Jr., Schuwirth L ve ark. (2013) Clarifying assumptions to enhance our understanding and assessment of clinical reasoning. Acad Med 88: 442-8.
- Emekli E, Karahan BN (2024) Comparison of Automatic Item Generation Methods in the Assessment of Clinical Reasoning Skills. Revista Española de Educación Médica, 6: (1).
- Gierl MJ, Lai H, Turner SR (2012) Using automatic item generation to create multiple-choice test items. Med Educ 46: 757-65.
- Gierl MJ, Lai H (2013) Evaluating the quality of medical multiple-choice items created with automated processes. Med Educ 47: 726-33.
- Gierl M, Lai H (2015) Using Automated Processes to Generate Test Items And Their Associated Solutions and Rationales to Support Formative Feedback. Interact Des Archit(s) 25: 9-20.
- Gierl MJ, Lai H, Tanygin V (2021) Advanced Methods in Automatic Item Generation, 1. Baskı. New York, Routledge, s. 42-66.
- Hassoulas A, Forty E, Hoskins M ve ark. (2017) A case-based medical curriculum for the 21st century: The use of innovative approaches in designing and developing a case on mental health. Med Teach 39: 505-11.
- Hawks MK, Maciuba JM, Merkebu ve ark. (2023) Clinical Reasoning Curricula in Preclinical Undergraduate Medical Education: A Scoping Review. Acad Med 98: 958-65.
- IBM Corporation Released 2013. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.
- Kıyak YS, Coşkun Ö, Budakoğlu İİ ve ark. (2023a). Psychometric Analysis of the First Turkish Multiple-Choice Questions Generated Using Automatic Item Generation Method in Medical Education. Tıp Eğitimi Dünyası 22: 154-61.
- Kıyak YS, Budakoğlu İİ, Coşkun Ö ve ark. (2023b) The first automatic item generation in Turkish for assessment of clinical reasoning in medical education. Tıp Eğitimi Dünyası 22: 72-90.
- Kıyak YS, Emekli E (2024) ChatGPT prompts for generating multiple-choice questions in medical education and evidence on their validity: a literature review. Postgrad Med J 6: qgae065.
- Lai H, Gierl MJ, Touchie C ve ark. (2016) Using Automatic Item Generation to Improve the Quality of MCQ Distractors. Teach Learn Med 28: 166-73.
- Laupichler MC, Rother JF, Grunwald Kadow IC ve ark. (2023) Large Language Models in Medical Education: Comparing ChatGPT- to Human-Generated Exam Questions. Acad Med 99: 508-12.
- Manuaba IBAP, No Y, Wu CC (2022) The effectiveness of problem based learning in improving critical thinking, problem-solving and self-directed learning in first-year medical students: A meta-analysis. PLoS One 22: e0277339.
- Masters K (2023) Medical Teacher's first ChatGPT's referencing hallucinations: Lessons for editors, reviewers, and teachers. Med Teach 45: 673-5.
- Miller GE (1990) The assessment of clinical skills/competence/performance. Acad Med 65 (Suppl 9): 63-7.
- Moghadami M, Amini M, Moghadami M ve ark. (2021) Teaching clinical reasoning to undergraduate medical students by illness script method: a randomized controlled trial. BMC Med Educ 21: 87.
- Morreale MK (2019) Teaching and Learning Through the Use of Patient Cases Approach to the Psychiatric Patient: Case-Based Essays, Second Edition. By John W. Barnhill. Acad Psychiatry 43: 311.
- Pugh D, De Champlain A, Gierl M ve ark. (2020) Can automated item generation be used to develop high quality MCQs that assess application of knowledge? Res Pract Technol Enhanc Learn 15: 12.
- Rejón AC (2012) Logic structure of clinical judgment and its relation to medical and psychiatric semiology. Psychopathology 45: 344-51.
- Rochmawati E, Wiechula R (2010) Education strategies to foster health professional students' clinical reasoning skills. Nurs Health Sci 12: 244-50.
- Schmidt HG, Mamede S (2015) How to improve the teaching of clinical reasoning: a narrative review and a proposal. Med Educ 49: 961-73.
- Schuwrth LW, van der Vleuten CP (2004) Different written assessment methods: what can be said about their strengths and weaknesses? Med Educ 38: 974-9.
- Tavakol M, O'Brien DG, Sharpe CC ve ark. (2024) Twelve tips to aid interpretation of post-assessment psychometric reports. Medical Teacher 46: 188-95.
- Ulusal Çalışma Grubu Ulusal Cep-2020, Ulusal Yetkinlik Ve Yeterlilikler Çalışma Grubu Ulusal Cep-2020, Davranış Sosyal Beseri Bilimler Çalışma Grubu Ulusal Cep-2020 (2020) Medical Faculty - National Core Curriculum 2020. Tıp Eğitimi Dünyası 19: 1-146.
- Van Rossum G, Drake FL (1995). Python reference manual. Centrum voor Wiskunde en Informatica Amsterdam.
- Walker HL, Ghani S, Kuemmerli C ve ark. (2023) Reliability of medical information provided by ChatGPT: Assessment against clinical guidelines and patient information quality instrument. J Med Internet Res 25: e47479.
- Wrigley W, van der Vleuten CP, Freeman A ve ark. (2012) A systemic framework for the progress test: strengths, constraints and issues: AMEE Guide No. 71. Med Teach 34: 683-97.
- Zuckerman M, Flood R, Tan RJB ve ark. (2023) ChatGPT for assessment writing. Med Teach 45: 1224-7.