

Joanna Gotlib¹, Aleksander Zarzeka¹, Mariusz Panczyk¹, Piotr Gębski², Lucyna Iwanow³, Marcin Malczyk⁴, Jarosława Belowska¹

¹Zakład Dydaktyki i Efektów Kształcenia, Wydział Nauki o Zdrowiu, Warszawski Uniwersytet Medyczny

²Centralny Ośrodek Egzaminacyjny Wydziału Lekarskiego, Uniwersytet Medyczny w Łodzi

³Studenckie Koło Naukowe Prawa Medycznego, Warszawski Uniwersytet Medyczny

⁴Biuro ds. Egzaminów Uczelnianych, Warszawski Uniwersytet Medyczny

Analiza jakości pytań z końcowego testu zaliczeniowego z przedmiotu „Prawo w medycynie” dla studentów pielęgniarstwa na platformie egzaminów elektronicznych ASK Systems

Słowa kluczowe: egzaminy wspomagane komputerowo, jakość kształcenia, studenci nauk o zdrowiu, pomiar dydaktyczny

Wstęp: Od roku 2014/2015 istnieje w Warszawskim Uniwersytecie Medycznym możliwość przeprowadzania egzaminów testowych w wersji elektronicznej. Pierwszym przedmiotem, którego zaliczenie odbyło się na platformie egzaminów elektronicznych było „Prawo w medycynie” dla studentów Pielęgniarstwa.

Cel pracy: Analiza jakości pytań z testu zaliczeniowego.

Materiał i metody: 200 pytań z 3 możliwościami odpowiedzi, jedna prawidłowa. Analiza rzetelności, łatwości i mocy różnicującej pytań w zależności od tury egzaminu i liczby losowań pytania. Analiza: łatwości bazy pytań, częstości wykorzystania pytań w bazie, ocena zgodności zakładanej względem wyliczonej łatwości pytań, ocena łatwości w zależności od częstości wykorzystania pytania. Statistica 10.0, licencja WUM, ANOVA rang Kruskala-Wallis.

Wyniki: 7 tur egzaminu, łatwość bazy pytań: średnia: 0,715, mediana: 0,750, SD: 0,2346, ocena zgodności łatwości zakładanej względem wyliczonej: korelacja istotna ($p < 0,05$), średni stopień zgodności ($\gamma = 0,433$). Parametry pytań testowych nie zależały od tury egzaminu: ANOVA rang Kruskala-Wallis, łatwość $H = 2,946$; $p = 0,8156$, moc różnicująca $H = 2,842$; $p = 0,8283$. Wraz ze wzrostem liczby losowań danego pytania spada wyliczona dla niego łatwość - na każde jedno dodatkowe wylosowanie pytania spada jego łatwość o wartość 0,002, jednakże zależność ta wymaga potwierdzenia.

Wnioski: 1. Skala pomiarowa prezentowanych pytań była krótka (30 pytań, 3 odpowiedzi), dlatego badania mają charakter pilotażowy i wymagają kontynuacji. 2. Łatwość pytań była wyższa niż zakładana, dlatego też należy rozważyć konieczność modyfikacji pytań. 3. Powtarzalność pytań wpływa na spadek łatwości pytań, dlatego też nieuzasadnione wydają się obawy o ponowne wykorzystywanie wylosowanych już wcześniej pytań. 4. Konieczne jest przeprowadzenie cyklu szkoleń dla nauczycieli obejmujących problematykę układania pytań testowych i ewaluacji egzaminów testowych.

Analysis of quality of questions in the final test for the "Law in Medicine" course for nursing students on ASK Systems e-exam platform

Keywords: computer-aided examinations, quality of education, students of health sciences, didactic measurement

Introduction: It is possible to conduct electronic tests at Warsaw Medical University (WMU) in the academic year 2014-2015. The final test for the "Law in Medicine" course was the first one in the history of WMU that was conducted on the e-exam platform.

Aim of the study: To analyse the quality of questions.

Material and methods : 200 questions with three answer options, one correct answer. Analysis of reliability, easiness and differentiation power of particular questions by the round of exam and number of draws of a particular question. Analysis of the easiness of the question database, frequency of using questions in the database, assessment of interrater reliability assumed for the estimated question easiness, assessment of easiness by the frequency of using a particular question. Statistica 10.0, licensed to WMU, Kruskal-Wallis ANOVA test.

Results: Seven rounds of exams, easiness of question database: medium: 0.715, median: 0.750, SD: 0.2346, assessment of interrater reliability assumed for the estimated question easiness: significant correlation ($p < 0.05$), moderate agreement ($\gamma = 0.433$). Parameters of the test questions were independent of the round of exam: Kruskal-Wallis ANOVA test, easiness of $H = 2.946$; $p = 0.8156$, differentiation power of $H = 2.842$; $p = 0.8283$. Estimated easiness of a particular question dropped with the increase of the number of drawings of this question: question's easiness dropped by 0.002 on each additional drawing.

Conclusions: 1. Owing to a small measuring scale of the questions, the study is only a pilot and needs to be continued. 2. Due to the fact that easiness of questions used in the test was higher than expected, modification of questions needs to be considered. 3. Repeatability of questions in the test decreased the easiness of questions. Concerns about the reuse of previously drawn questions seem to be unjustified. 4. It seems necessary to conduct a series of trainings for teachers covering broad issues concerning phrasing test questions and evaluating test exams.