

УДК 378.16

DOI:

Загребельний С. Л., Костіков О. А., Брус М. В.

ОЦІНКА ВАЛІДНОСТІ ТЕСТОВИХ ЗАВДАНЬ В СИСТЕМІ ДИСТАНЦІЙНОГО НАВЧАННЯ MOODLE

Постановка проблеми у загальному вигляді та її зв'язок з важливими науковими чи практичними завданнями. Проблема оцінки якості тестових завдань для дистанційного контролю знань студентів має велике теоретичне та практичне значення. На співбесіді зі студентом досвідчений викладач може дуже легко виявити дійсний рівень знань студента і виставити відповідну оцінку. Але при дистанційному навчанні виникає питання: наскільки результати тестування відображають дійсний рівень знань опитуваних. Тому велике значення набуває оцінка валідності тестових завдань, яка показує, наскільки оцінки, отримані студентом в результаті тестового випробування відповідають його справжньому рівню знань. У зв'язку з впровадженням дистанційного навчання на кафедрі інформатики і інженерної графіки Донбаської державної машинобудівної академії виникла проблема оцінки валідності розроблених тестових завдань. Вирішенню цієї задачі присвячена ця стаття.

Аналіз основних досліджень і публікацій. Питання теорії діагностики знань за допомогою тестування, форми та структура тестових завдань, алгоритми обробки результатів оцінювання та їх інтерпретації досліджено в роботах Аванесова В. С., Богачкова Ю. М., Бочарникова В., Бродського Я. С., Землякова А. Н., Іванова Ю. І., Інгенкампа К., Равена Дж., Тализіна Н., Шиян Н. та ін. Питанням розробки якісних тестів присвячені праці Кіма В. С. [1], І.Є. Булах І. Є. [2], Анастасі А. [3]. Аналіз валідності тестових завдань у системі Moodle був проведений у роботах Протасової та ін. [4], Коржика та ін. [5], Толстоброва та ін. [6]. Автори цієї статті також приділили увагу аналізу якості тестів засобами системи MOODLE [7].

Формування цілей статті (постановка завдання). Мета даної статті полягає в тому, щоб виділити неінформативні питання та питання з низькою розподільною здатністю серед тестових завдань по дисциплінам, які викладаються на кафедрі інформатики і інженерної графіки Донбаської державної машинобудівної академії за допомогою оцінки валідності тестів засобами системи MOODLE.

Виклад основного матеріалу. При використанні засобів дистанційного навчання перед викладачем виникає проблема створення банку тестових завдань. При цьому неминуче виникає питання про їх якість з точки зору їхньої здатності адекватно оцінювати рівень підготовки студентів. Якісні тестові завдання повинні задовольняти умові валідності. Найбільш вдале визначення валідності педагогічного тесту дав Анастасі [3]. Валідність тесту – це “поняття, яке визначає, що вимірює тест і наскільки якісно це здійснюється” [2]. Невалідність тесту означає, що одержані результати не вимірюють того, що було визначено метою тестування. У нашому випадку це означає, що оцінки, які отримав студент в результаті комп'ютерного тестування не відповідають його справжньому рівню знань. Аналіз валідності результатів базується на спеціально розроблених методиках зіставлення з критеріями аналізу об'єктивності результатів, в основі яких – оцінка міри відхилення від нормального закону результатів тестування та процедура їх нормалізації [8]. Ступінь відхилення від нормального закону визначається за такими параметрами, як асиметрія A_s і ексцес E_x , та встановленням їх статистичної достовірності за t -критерієм Стьюдента. Залежно від значення та знаку асиметрії або ексцесу результати можна поділити на дев'ять загальних груп, що дають змогу визначити критерії прийняття рішення щодо валідності процесу вимірювання та наявності або відсутності статистично достовірної дії факторів, що впливають на результати вимірювання.

Залежно від значення і знаків асиметрії та ексцесу закону розподілу можна виділити дев'ять таких критеріїв[2]:

1) Тест є валідним, якщо асиметрія та ексцес закону розподілу дорівнюють нулю ($A_s = 0, E_x = 0$).

2) Якщо асиметрія є від'ємною, а ексцес додатним ($A_s = \text{“-“}, E_x = \text{“+“}$), то тест є невалідним (легкий і з малою розподільною здатністю). Це результат дії кількох факторів: – тест сконструйовано з невалідних тестових завдань (він мусить валідизуватися шляхом їх ускладнення та збільшення їх розподільної здатності); – процедура тестування є невалідною.

3) Якщо асиметрія та ексцес є від'ємними ($A_s = \text{“-“}, E_x = \text{“-“}$), то тест є легким, але з великою розподільною здатністю. Його валідизація полягає в ускладненні тестових завдань.

4) Якщо асиметрія є від'ємною, а ексцес відсутній ($A_s = \text{“-“}, E_x = 0$), то тест є легким. Його валідизація полягає в ускладненні тестових завдань.

5) Якщо асиметрія є додатною, а ексцес від'ємним ($A_s = \text{“+“}, E_x = \text{“-“}$), то тест є складним з великою розподільною здатністю. Його валідизація полягає у валідизації змісту шляхом вилучення тестових завдань, які не відповідають цій вимозі.

6) При додатних асиметрії та ексцесі ($A_s = \text{“+“}, E_x = \text{“+“}$) тест є складним з малою розподільною здатністю. Це є наслідком дії двох факторів, що впливають один на одного: порушення валідності змісту, невалідність тестових завдань. Валідизація тесту полягає у валідизації тестових завдань.

7) При додатній асиметрії та відсутності ексцесу ($A_s = \text{“+“}, E_x = 0$) тест є складним. Це є наслідком дії двох факторів, що впливають один на одного: порушення валідності змісту, невалідність тестових завдань. Валідизація тесту полягає у валідизації тестових завдань. Особливої уваги потребують два останні критерії (8 і 9), які дають можливість визначити конструктивну валідність тесту, а саме перевірити тест на однорідність. Нагадаємо, що тест вважається однорідним, якщо його сконструйовано з тестових завдань з близькими характеристиками. Неоднорідним вважається такий тест, який складається з окремих субтестів, кожний з яких можна розглядати як тест з певними характеристиками.

8) Якщо крива розподілу симетрична, асиметрія дорівнює нулю, а ексцес додатний ($A_s = 0, E_x = \text{“+“}$), то тест є неоднорідним і сконструйованим із двох субтестів: складного і легкого. Валідизація тесту полягає в його поділі на два окремих тести.

9) Якщо крива розподілу симетрична, асиметрія дорівнює нулю, а ексцес від'ємний ($A_s = 0, E_x = \text{“-“}$), то тест є неоднорідним, сконструйованим із трьох субтестів: легкого, середнього і складного. Валідизація тесту полягає в його поділі на окремі тести.

Застосуємо ці критерії до отриманих результатів тестування. В результаті тестування були отримані значення асиметрії та ексцесу за допомогою внутрішніх засобів Moodle. Ці характеристики наведені в таблиці 1.

Таблиця 1

Значення асиметрії та ексцесу, отримані в результаті тестування по дисциплінам кафедри

Назва дисципліни	Значення асиметрії розподілу (для найвище оціненої спроби)	Значення ексцесу розподілу (для найвище оціненої спроби)
Обчислювальна техніка та програмування	-1,846	3,639
Чисельні методи та моделювання на ЕОМ	-0,6541	0,4091
Економічна інформатика	-0,8872	0,2748

З таблиці 1 видно, що для всіх трьох дисциплін значення асиметрії є від'ємним, а ексцесу - додатним, тому, згідно з другим критерієм тест є невалідним (легкий і з малою розподільною здатністю). Для валідизації тесту треба ускладнити тестові питання та збільшити їх розподільну здатність.

Оскільки тести виявилися невалідними, то виникає необхідність провести більш ретельний аналіз з метою виявлення легких питань та питань з незадовільною розподільною здатністю. Це можливо зробити засобами системи MOODLE.

Для реалізації цієї мети викладачами кафедри інформатики і інженерної графіки Донбаської державної машинобудівної академії за допомогою вбудованого в MOODLE статистичного модуля були отримані наступні характеристики для тестових завдань: успішність, стандартне відхилення, оцінка навмання, призначена вага, ефективна вага, розрізнення, ефективність розрізнення.

Параметр успішність p_j визначає ту частину студентів, яка правильно відповіла на поставлене запитання і обчислюється по формулі

$$p_j = \frac{\sum_{i=1}^n x_{ij}}{n}, (j = 1, 2, \dots, m),$$

де m - кількість тестових завдань, n - кількість студентів, що брали участь у тестуванні, $x_{i,j}$ - оцінка успішності виконання j -го завдання i -м студентом. Чим більше величина коефіцієнта p_j , тим більша кількість студентів виконала j -те запитання [5-6,10].

Складність завдань тесту мусить відповідати рівню підготовки студентів, що проходять тестування. Тест в цілому повинен включати в себе комплекс завдань різної складності - від легких до важких [1, 9]. Занадто прості завдання, на які правильно відповідають всі випробувані ($p_j = 1$) і занадто складні завдання, на які не може відповісти ніхто з опитуваних ($p_j = 0$) не володіють здатністю диференціювати тестуємих за рівнем їх підготовки і повинні з виключатися з переліку тестових завдань.

На рисунках 1, 2 і 3 наведені гістограми значень коефіцієнта успішності та ефективності розрізнення тестових завдань, які були отримані в результаті підсумкового тестування по дисциплінам курсу "Обчислювальна техніка та програмування", "Чисельні методи та моделювання на ЕОМ" і "Економічна інформатика"

Статистика за позиціями питань

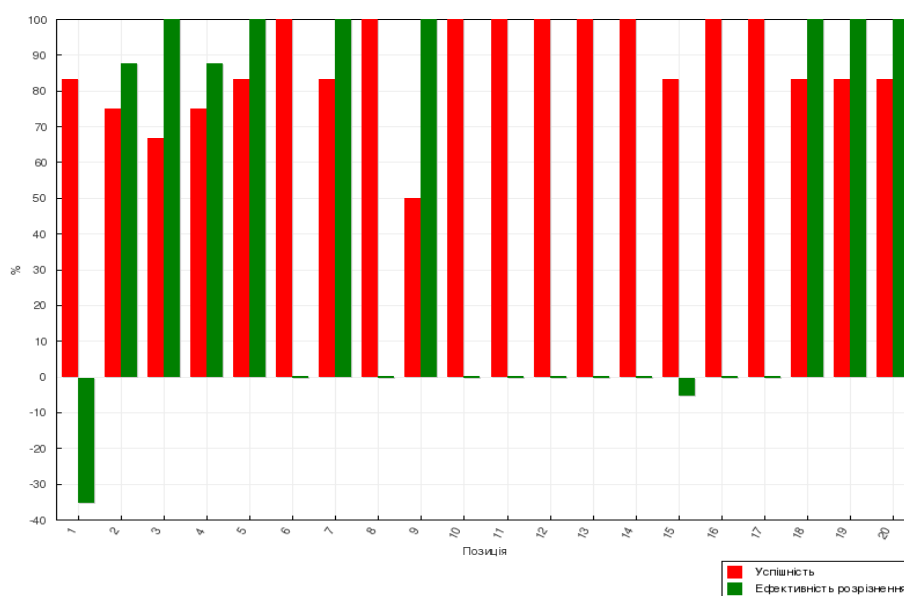


Рис. 1. Значення коефіцієнтів успішності та ефективності розрізнення завдань підсумкового тесту по курсу "Обчислювальна техніка та програмування"

Статистика за позиціями питань

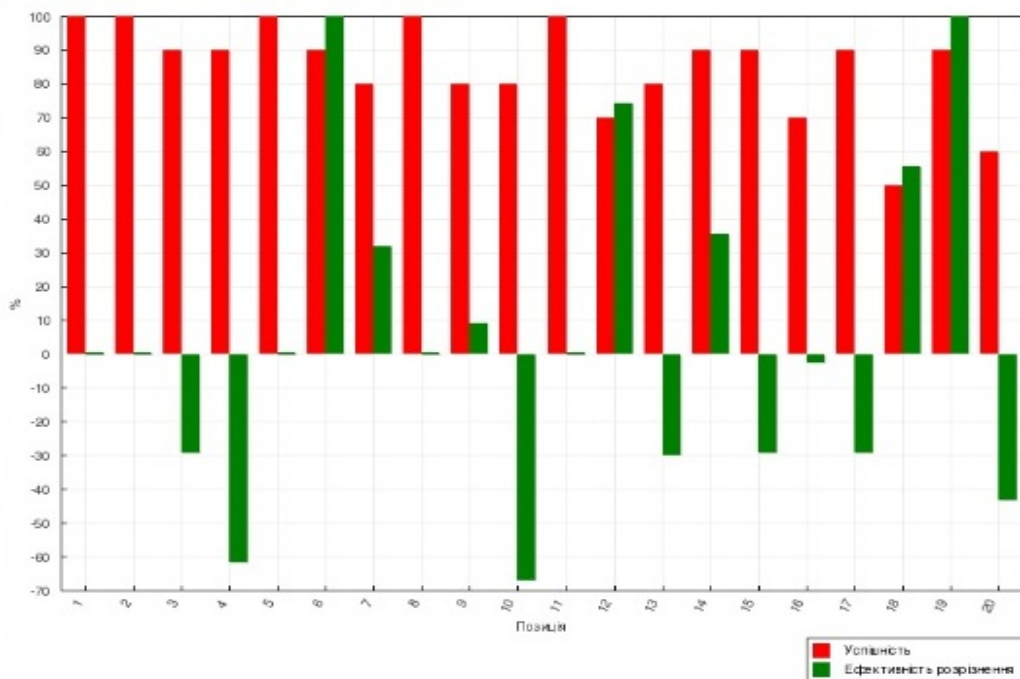


Рис. 2. Значення коефіцієнтів успішності та ефективності розрізнення завдань підсумкового тесту по курсу "Чисельні методи та моделювання на ЕОМ"

Оцінка індексу успішності дозволила виявити в тесті неінформативні питання. Так, у тесті по дисципліні "Обчислювальна техніка та програмування" зустрічаються дуже легкі питання з номерами 6,10,11,12,13,14,16,17 на які відповіли всі опитувані. У тесті по дисципліні "Чисельні методи та моделювання на ЕОМ" дуже легкими виявилися питання з номерами 1, 2, 5, 8, 11. З графіку на рисунку 3 видно, що серед завдань з дисципліни "Економічна інформатика" дуже легкими є запитання 1, 11 і 12.

Статистика по позиціям вопросов

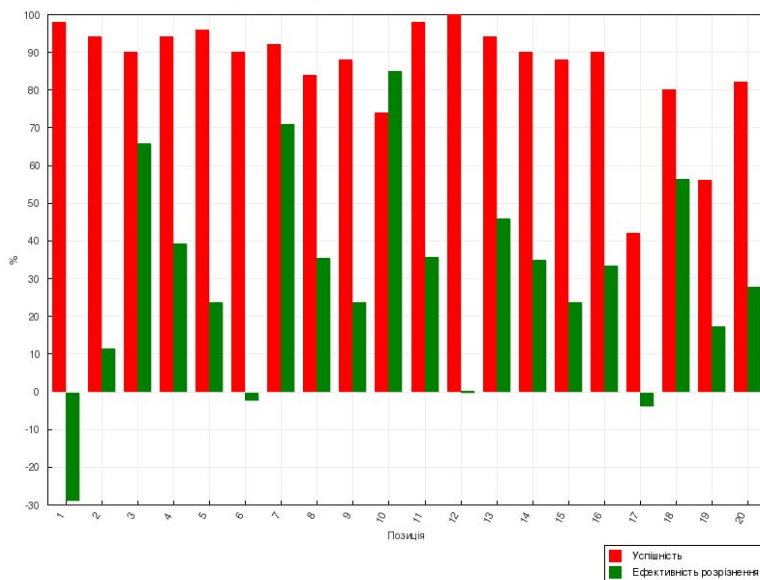


Рис. 3. Значення коефіцієнтів успішності та ефективності розрізнення завдань підсумкового тесту по курсу "Економічна інформатика"

Стандартне відхилення результатів випробовуваних по j -му завданню тесту ($j = 1, 2, \dots, m$) обчислюється за формулою:

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n-1}},$$

де $x_{i,j}$ – оцінка успішності виконання j -го завдання i -м студентом, \bar{x}_j – середня оцінка, отримана i -м студентом за виконання тесту. Аналіз величин стандартного відхилення оцінки для кожного питання дозволяє виявити її внесок в розподільну здатність тесту. Для більшості тестових питань в досліджених тестових завданнях воно має значення більше 0,30, що, відповідно з вимогами педагогічної теорії вимірювань [1, 2, 11, 12, 13,], є задовільним показником їх розподільної здатності (рисунок 3). Завдання, для яких це значення менше 0,30, не мають розподільної здатності і повинні бути перероблені. Наприклад, у випадку тесту по дисципліні "Чисельні методи та моделювання на ЕОМ" завдання з номерами 1,2,5,11 мають стандартне відхилення, яке дорівнює нулю і тому їх треба видалити з тесту (таблиця 2). Аналіз таблиці 3 показує що серед тестових завдань по дисципліні "Обчислювальна техніка та програмування" завдання з номерами 8, 10-14,16,17 мають нульове стандартне відхилення і тому їх треба видалити з переліку тестових завдань. З даних таблиці 4 можна зробити висновок, що стандартне відхилення питань 1,2,4,5,7,12,13 становить менше 30 відсотків і тому мають незадовільну розподільну здатність.

Таблиця 2

Статистичні характеристики, отримані за допомогою статистичного модуля системи Moodle на основі результатів підсумкового тестування по дисципліні "Чисельні методи та моделювання на ЕОМ"

№ питання	Успішність	Станд. відхилення	Оцінка навмання	Призначена вага	Ефективна вага	Розрізнення	Ефективність розрізнення
1	100.00 %	0.00 %	20.00 %	5.00 %	0.00 %		
2	100.00 %	0.00 %	20.00 %	5.00 %	0.00 %		
3	90.00 %	31.62 %	25.00 %	5.00 %		-19.82 %	-29.03 %
4	90.00 %	31.62 %	25.00 %	5.00 %		-40.14 %	-61.29 %
5	100.00 %	0.00 %	25.00 %	5.00 %	0.00 %		
6	90.00 %	31.62 %	25.00 %	5.00 %	12.45 %	53.85 %	100.00 %
7	80.00 %	42.16 %	25.00 %	5.00 %	12.45 %	26.38 %	31.82 %
8	100.00 %	0.00 %	33.33 %	5.00 %	0.00 %		
9	80.00 %	42.16 %	25.00 %	5.00 %	10.17 %	7.14 %	9.09 %
10	80.00 %	42.16 %	25.00 %	5.00 %		-54.17 %	-66.67 %
11	100.00 %	0.00 %	25.00 %	5.00 %	0.00 %		
12	70.00 %	43.30 %	20.00 %	5.00 %	16.07 %	53.29 %	74.36 %
13	80.00 %	42.16 %	20.00 %	5.00 %		-26.04 %	-29.63 %
14	90.00 %	31.62 %	20.00 %	5.00 %	10.17 %	26.67 %	35.48 %
15	90.00 %	31.62 %	20.00 %	5.00 %		-19.82 %	-29.03 %
16	70.00 %	48.30 %	25.00 %	5.00 %	10.17 %	-1.54 %	-2.04 %
17	90.00 %	31.62 %	25.00 %	5.00 %		-19.82 %	-29.03 %
18	50.00 %	52.70 %	25.00 %	5.00 %	16.07 %	41.52 %	55.56 %
19	90.00 %	31.62 %	25.00 %	5.00 %	12.45 %	53.85 %	100.00 %
20	60.00 %	51.64 %	25.00 %	5.00 %	0.00 %	-31.36 %	-42.86 %

Таблиця 3

Статистичні характеристики, отримані за допомогою статистичного модуля системи Moodle на основі результатів підсумкового тестування по дисципліні "Обчислювальна техніка та програмування"

№ питання	Успішність	Станд. відхилення	Оцінка навмання	Призначена вага	Ефективна вага	Розрізнення	Ефективність розрізнення
1	33.33 %	40.32 %		5.0 %		-32.64 %	-35.00 %
2	76.00 %	41.83 %		6.0 %	10.57 %	34.34 %	87.76 %
3	66.67 %	51.64 %		6.0 %	11.19 %	73.29 %	100.00 %
4	76.00 %	41.83 %		6.0 %	10.57 %	34.34 %	87.76 %
5	33.33 %	40.82 %		6.0 %	10.78 %	92.20 %	100.00 %
6	100.00 %	0.00 %		5.0 %	0.00 %		
7	83.33 %	40.82 %		5.0 %	10.78 %	92.20 %	100.00 %
8	100.00 %	0.00 %		5.0 %	0.00 %		
9	50.00 %	54.77 %		5.0 %	10.78 %	60.19 %	100.00 %
10	100.00 %	0.00 %		5.0 %	0.00 %		
11	100.00 %	0.00 %		5.0 %	0.00 %		
12	100.00 %	0.00 %		5.0 %	0.00 %		
13	100.00 %	0.00 %		5.0 %	0.00 %		
14	100.00 %	0.00 %		5.0 %	0.00 %		
15	33.33 %	40.82 %		5.0 %	2.99 %	-4.82 %	-5.00 %
16	100.00 %	0.00 %		5.0 %	0.00 %		
17	100.00 %	0.00 %		5.0 %	0.00 %		
18	83.33 %	40.82 %		5.0 %	10.78 %	92.20 %	100.00 %
19	83.33 %	40.82 %		5.0 %	10.78 %	92.20 %	100.00 %
20	83.33 %	40.82 %		5.0 %	10.78 %	92.20 %	100.00 %

Введемо деякі позначення $s, S, T_s, x_j(s), X_j(s)$ і визначимо їх наступним чином: s - номер спроби здачі тесту, S - множина всіх спроб здачі тесту, T_s - кількість набраних студентом балів за тест, $x_j(s)$ - оцінка студента за визначене завдання, $X_j(s)$ - оцінка студента за всі інші завдання:

$$T_s = \sum_{j=1}^m x_j(s);$$

$$X_j(s) = T_s - x_j(s)$$

Дисперсія показника позначатиметься $V(x_j(s))$, наприклад:

$$V(x_j(s)) = \frac{1}{S-1} \sum_{x \in S} (x_j(s) - \bar{x}_j)^2$$

Коваріація позначатиметься $C(x_j, X_j)$, наприклад:

$$C(x_j, X_j) = \frac{1}{S-1} \sum_{s \in S} (x_j(s) - \bar{x}_j)(X_j(s) - \bar{X}_j)$$

Таблиця 4

Статистичні характеристики, отримані за допомогою статистичного модуля системи Moodle на основі результатів підсумкового тестування по дисципліні "Економічна інформатика"

№ питання.	Успішність	Станд. відхилення	Оцінка навчання	Призначена вага	Ефективна вага	Розрізнення	Ефективність розрізнення
1	98.00 %	14.14 %	33.33 %	5,00 %		-11.54 %	-28.62 %
2	94.00 %	23.99 %	33.33 %	5,00 %	3.41 %	6.93 %	11.28 %
3	90.00 %	30.30 %	16.67 %	5,00 %	6.75 %	44.07 %	65.85 %
4	94.00 %	23.99 %	20.00 %	5,00 %	4.66 %	22.91 %	39.11 %
5	96.00 %	19.79 %	20.00 %	5,00 %	3.34 %	11.72 %	23.66 %
6	90.00 %	30.30 %	14.29 %	5,00 %	3.18 %	-1.53 %	-2.22 %
7	92.00 %	27.40 %	14.29 %	5,00 %	6.32 %	43.38 %	70.79 %
8	84.00 %	37.03 %	20.00 %	5,00 %	6.48 %	26.08 %	35.44 %
9	88.00 %	32.83 %	20.00 %	5,00 %	5.32 %	17.26 %	23.76 %
10	74.00 %	44.31 %	20.00 %	5,00 %	9.75 %	68.31 %	84.84 %
11	98.00 %	14.14 %	20.00 %	5,00 %	2.85 %	14.64 %	35.69 %
12	100.00 %	0.00 %	16.67 %	5,00 %	0.00 %		
13	94.00 %	23.99 %	33.33 %	5,00 %	4.93 %	27.00 %	45.87 %
14	90.00 %	30.30 %	33.33 %	5,00 %	5.51 %	23.81 %	34.88 %
15	88.00 %	32.83 %	25.00 %	5,00 %	5.32 %	17.26 %	23.76 %
16	90.00 %	30.30 %	33.33 %	5,00 %	5.51 %	23.81 %	33.33 %
17	42.00 %	49.86 %	20.00 %	5,00 %	5.23 %	-2.60 %	-3.57 %
18	80.00 %	40.41 %	33.33 %	5,00 %	8.11 %	45.46 %	56.25 %
19	56.00 %	50.14 %	20.00 %	5,00 %	7.01 %	13.80 %	17.23 %
20	82.00 %	38.81 %	20.00 %	5,00 %	6.32 %	21.10 %	27.68 %

Важливою статистичною характеристикою розподільної здатності тестових завдань, яку можна отримати засобами Moodle, є розрізнення D_j [10]:

$$D_j = 100\% \cdot \frac{C(x_j, X_j)}{\sqrt{V(x_j) \cdot V(X_j)}}$$

Цей коефіцієнт показує, наскільки взаємопов'язані правильність відповіді на дане питання і інші питання тесту. Для гарного тестового питання передбачається, що студенти з високими оцінками за нього також будуть мати більш високі оцінки і за тест в цілому. На це вказує позитивне значення коефіцієнта. При цьому наголошується [14,15], що якщо коефіцієнт успішності питання відмінний від 50 %, то D_j не може дорівнювати 100 %. Якщо коефіцієнт успішності близький до нуля або 100 %, то розрізнення завжди буде дуже маленьким, а якщо він буде дорівнювати нулю або 100 %, то D_j буде не визначений.

Ефективність розрізнення De_j – це нормований коефіцієнт розрізнення [16].

Ефективність розрізнення є мірою здатності конкретного завдання розділяти сильних і слабих випробовуваних. Позитивні значення відповідають завданням, які дійсно розділяють «сильних» і «слабих» студентів, в той час як від'ємне значення De_j свідчить про те, що погано підготовлені студенти відповідають на дане завдання в середньому краще, ніж добре підготовлені. Очевидно, що такі завдання не є тестовими, тому що не здатні адекватно розділяти випробовуваних за рівнем їх підготовки, і їх слід вибракувати. Вважається що завдання має достатню розподільну здатність, якщо ефективність розрізнення має значення більше або рівне 0,30 [1, 13, 17]. Аналіз величини De_j для розглянутих тестів дозволив виявити завдання, що не мають достатньої розподільної здатності (0,30). Аналіз даних на рисунку 1 свідчить про те, що серед тестових питань по дисципліні "Обчислювальна техніка та програмування" завдання з номерами 1,15 мають від'ємну ефективність розрізнення і потребують переробки. Для тестових завдань по дисципліні "Чисельні методи та моделювання на ЕОМ" наприклад, питання з номерами 3, 4, 9, 10, 13, 15 не вищенаведеним вимогам, крім того, для деяких з наведених питань значення ефективності розрізнення виявилось від'ємним, що свідчить про необхідність перегляду цих питань (рисунок 3.2). Щодо тестових завдань по дисципліні "Економічна інформатика" питання 1, 6, 17 мають від'ємне значення коефіцієнту ефективності розрізнення, а питання 2, 5, 9, 15, 19 мають значення цього коефіцієнту менше 30 відсотків. Тому ці питання слід переробити.

ВИСНОВКИ

За допомогою статистичних характеристик, обчислених у системі MOODLE були знайдені легкі тестові питання та питання з низькою розподільною здатністю. Переробка цих питань дозволить отримати валідні тестові завдання. Отримання валідних тестів буде темою подальших досліджень.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Ким В. С. Тестирование учебных достижений. Усурийск : Издательство УГПИ, 2007. 214 с.
2. Булах І. Є., Мруга М.Р. Створюємо якісний тест: навч. посіб. К.: Майстер;клас, 2006. 160 с.
3. Анастаси А. Психологическое тестирование: Пер. с англ. в 2 кн. Кн. 1. М. : Педагогика, 1982. 316 с.
4. Протасова И. В, Толстобров А. П., Коржик И. А. Методика анализа и повышения качества тестов в системе электронного обучения Moodle. *Вестник Воронежского гос.ун.-та:Серия системный анализ и информационные технологии*. Воронеж, 2014. №3. С.61-72.
5. Коржик И.А., Протасова И.В., Толстобров А.П. Тестовая система Moodle и качество тестовых заданий. *Современные информационные технологии и ИТ-образование: сб. избранных трудов VII междунар. науч.-практич. конф.* Москва: ИНТУИТ.РУ, 2012. С. 187–196.
6. Толстобров А. П., Коржик И. А. Возможности анализа и повышения качества тестовых заданий при использовании сетевой системы управления обучением MOODLE. *Вестник Воронежского гос. ун- та. системный анализ и информационные технологии*. Воронеж, 2008. № 2. С. 100–106.
7. Брус М., Костіков О., Загребельний С. Оцінка якості тестових завдань в Moodle. *Проблеми та перспективи розвитку сучасної науки в країнах Європи та Азії* : матеріали XVI Міжнародної науково-практичної інтернет-конференції 31 травня 2019р. Переяслав-Хмельницький, 2019. С. 71-74.
8. Булах І. Є. Теорія комп'ютерного тестування. Київ: ЦМК МОЗ України, 1994. 59 с.
9. Hambleton, R. K., & Jones, R. W. Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. *Educational Measurement: Issues and Practice*. 1993. No 12. P. 38–47. DOI:<https://doi.org/10.1111/j.1745-3992.1993.tb00543.x>
10. Коржик И.А., Толстобров А.П. Оценка качества тестов в системе электронного обучения MOODLE URL: <http://www.infoco.ru/mod/data/view.php?id=4&gid=114>
11. Андреев А.А, Солдаткин В.И. Дистанционное обучение: сущность, технология, организация. Москва: Издательство МЭСИ, 1999. 196 с.
12. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Copenhagen, Denmark: Danish Institute for Educational Research, 1960. 199p.

13. Бурлуцький С.В. Методичні рекомендації щодо розроблення електронних навчально- методичних матеріалів в системі Moodle DDMA для студентів всіх форм навчання. Краматорськ: ДДМА, 2017. 48 с.
14. Quiz statistics calculations. URL: http://docs.moodle.org/dev/Quiz_statistics_calculations.
15. Нестеров С.А., Сметанина М.В. Оценка качества тестовых заданий средствами среды дистанционного обучения MOODLE. *Научно-технические ведомости СПбГПУ. Информатика. Телекоммуникации. Управление*. 2013. №5(181). С. 87-92.
16. Аванесов В.С. Понятие и методы математической теории педагогических измерений (Item Response Theory): статья третья. *Педагогические Измерения*. 2009. №4. С. 5.
17. Тавгень, И.А. Дистанционное обучение: опыт, проблемы, перспективы. 2-е изд., исправл. и доп./под редакцией Ю.В.Позняка. Минск: БГУ, 2003. 227 с.

REFERENCES

1. Kim V.S. Learning Achievement Testing. Ussurijsk : Press UGPI, 2007. 214 p. *(in Russian)*.
2. Bulakh I. Ye., Mruga M.R. Quality test creation: textbook. Kyiv: Majster;klas, 2006. 160 p. *(in Ukrainian)*.
3. Anastasi A. Psychological Testing: Translation from English in 2 books. Book. 1. Moscow : Pedagogika, 1982. 316 p.
4. Protasova I.V, Tolstobrov A.P., Korzhik I.A. Methodology for analyzing and improving the quality of tests in the electronic learning system Moodle. *Proceedings of Voronezh State University: Ser. System analysis and information technologies*. Voronezh, 2014, 3, pp.61-72. *(in Russian)*.
5. Korzhik I.A., Protasova I.V., Tolstobrov A.P. Moodle test system and quality of test items.. *Modern Information Technology and IT-education: reports of VII International scientific-practical conference*. Moscow: INTU-IT.RU. 2012, pp. 187–196. *(in Russian)*.
6. Tolstobrov A.P., Korzhik I.A Opportunities for analyzing and improving the quality of test items using the MOODLE network learning management system. *Proceedings of Voronezh State University: Ser. System analysis and information technologies*. Voronezh, 2008,2, pp. 100–106. *(in Russian)*.
7. Brus M., Kostikov O., Zagrebelniy S. Quality assessment of test item in Moodle. Problems and prospects for the development of modern science in Europe and Asia. Proceedings of the XVI International scientific and practical internet conference, May 31, 2019. Pereiaslav-Khmelnytskyi, 2019. pp. 71-74 *(in Ukrainian)*.
8. Bulakh I. Ye. Computer Testing Theory. KiiV: CMK MOZ Ukraine, 1994. 59 p. *(in Ukrainian)*.
9. Hambleton, R. K., & Jones, R. W. Comparison of Classical Test Theory and Item Response Theory and Their Applications to Test Development. *Educational Measurement: Issues and Practice*. 1993, no 12, pp. 38–47. DOI:<https://doi.org/10.1111/j.1745-3992.1993.tb00543.x>
10. Korzhik I.A., Tolstobrov A.P. Assessing the quality of tests in the MOODLE e-learning system URL: <http://www.infoco.ru/mod/data/view.php?d=4&rid=114> . *(in Russian)*.
11. Andreev A.A, Soldatkin V.I. Distance learning: essence, technology, organization. Moscow: Press MESI, 1999. 196 p. . *(in Russian)*.
12. Rasch G. Probabilistic Models for Some Intelligence and Attainment Tests. Copenhagen, Denmark: Danish Institute for Educational Research, 1960. 199p.
13. Burluc'kij S.V. Methodical recommendations for the development of electronic teaching materials in the Moodle DDMA system for students of all forms of education. Kramatorsk : DDMA, 2017. 48 p *(in Ukrainian)*.
14. Quiz statistics calculations. [Online]. http://docs.moodle.org/dev/Quiz_statistics_calculations.
15. Nesterov S.A., Smetanina M.V. Assessment of the quality of test items by MOODLE distance learning environment. *St. Petersburg State Polytechnical University Journal. Computer Science. Telecommunication and Control Systems*. 2013, 5(181), pp 87-92 *(in Russian)*..
16. Аванесов В.С. The concept and methods of the mathematical theory of pedagogical measurements (Item Response Theory): article three. *Pedagogicheskie Izmerenija*. 2009, 4, P. 5 *(in Russian)*.
17. Tavgen I.A. Distance Learning: Experience, Problems, Prospects. 2-nd edition./ editor Ju.V.Poznjak. Минск: BGU, 2003. 227 p *(in Russian)*.

АВТОРИ / АВТОРЫ / AUTORS

Загребельний С. Л. – канд. пед. наук, доц. ДДМА;

Загребельный С. Л. – канд. пед. наук, доц. ДГМА;

Zagrebelnyi S. L. – candidate of pedagogical science, associate professor DSEA.

E-mail: szagrebelniy@gmail.com; <https://orcid.org/0000-0002-6246-4519>

Костіков О. А. – канд. фіз.-мат. наук, доц. ДДМА;
Костиков А. А. – канд. физ.-мат. наук, доц. ДГМА;
Kostikov A. A. – candidate of physico-mathematical sciences, associate professor DSEA.
E-mail: alexkst63@gmail.com; <https://orcid.org/0000-0003-3503-4836>

Брус М. В. – ст. викладач ДДМА;
Брус М. В. – ст. преподаватель ДГМА;
Brus M. V. – senior lecturer DSEA.
E-mail: draw@dgma.donetsk.ua; <https://orcid.org/0000-0002-0168-1712>

Донбаська державна машинобудівна академія (ДДМА), м. Краматорськ.
Донбасская государственная машиностроительная академия (ДГМА), г. Краматорск.
Donbass State Engineering Academy (DSEA), Kramatorsk.

АНОТАЦІЯ / АННОТАЦИЯ / ANNOTATION

Загребельный С. Л., Костіков О. А., Брус М. В. Оцінка валідності тестових завдань в системі дистанційного навчання MOODLE. Вісник Донбаської державної машинобудівної академії. 2020. № 1 (48).

У зв'язку з широким впровадженням дистанційних технологій в навчальний процес великого значення набуває адекватне оцінювання знань студентів за допомогою тестових завдань. Тому оцінка валідності тестів є актуальним завданням. У статті розглядаються можливості системи дистанційного навчання MOODLE для перевірки валідності тестових завдань, розроблених для контролю знань студентів. Проведена перевірка на валідність тестів з трьох дисциплін: «Обчислювальна техніка та програмування», «Чисельні методи і моделювання на ЕОМ», «Економічна інформатика», які були розроблені авторами. Перевірка на валідність проводилася засобами системи MOODLE, які полягають в оцінці ступеня відхилення від нормального закону результатів тестування і в процедурі їх нормалізації. На основі знайдених значень асиметрії і ексцесу закону розподілу результатів контролю, були виявлені валідність і невалідність завдання. Також були отримані наступні характеристики розроблених завдань для контролю знань студентів: індекс легкості, стандартне відхилення, випадково вгадана оцінка, намічений вага, ефективний вага, коефіцієнт диференціації, ефективність диференціації. Детально розглянуто, як за допомогою цих характеристик оцінити результативність тестового випробування, наведені математичні формули, за допомогою яких вони обчислюються і проведено аналіз тестових завдань на їх основі. Користуючись значеннями вищенаведених характеристик, виявлені тестові питання, які є дуже легкими для студентів, питання, які не володіють достатньою диференційною здатністю. Грунтуючись на результатах перевірки, дані рекомендації щодо підвищення валідності і надійності тестових завдань. Таким чином, в нашій статті за допомогою статистичних характеристик, обчислених в системі MOODLE, були знайдені легкі тестові питання і питання з низькою диференційною здатністю. Зміна цих питань дозволить отримати валідніє тестові завдання. Ключові слова: тестові завдання, валідність тестів, система дистанційного навчання MOODLE, диференцирующая здатність тестів.

Ключові слова: тестові завдання, валідність тестів, система дистанційного навчання MOODLE, диференцирующая здатність тестів.

Загребельный С. Л., Костиков О. А., Брус М. В. Оценка валидности тестовых заданий в системе дистанционного обучения MOODLE. Вестник Донбасской государственной машиностроительной академии. 2020. № 1 (48).

В связи с широким внедрением дистанционных технологий в учебный процесс большое значение приобретает адекватное оценивание знаний студентов с помощью тестовых заданий. Поэтому оценка валидности тестов является актуальной задачей.

В статье рассматриваются возможности системы дистанционного обучения MOODLE для проверки валидности тестовых заданий, разработанных для контроля знаний студентов. Проведена проверка на валидность тестов по трем дисциплинам: «Вычислительная техника и программирование», «Численные методы и моделирование на ЭВМ», «Экономическая информатика», которые были разработаны авторами. Проверка на валидность проводилась средствами системы MOODLE, которые заключаются в оценке степени отклонения от нормального закона результатов тестирования и в процедуре их нормализации. На основе найденных значений асимметрии и эксцесса закона распределения результатов контроля, были выявлены валидные и невалидные задания. Также были получены следующие характеристики разработанных заданий для контроля знаний студентов: индекс легкости, стандартное отклонение, случайно угаданная оценка, намеченный вес, эффективный

вес, коэффициент дифференциации, эффективность дифференциации. Подробно рассмотрено, как с помощью этих характеристик оценить результативность тестового испытания, приведены математические формулы, с помощью которых они вычисляются и проведен анализ тестовых заданий на их основе. Пользуясь значениями вышеприведенных характеристик, обнаружены тестовые вопросы, которые являются очень легкими для студентов, вопросы, которые не обладают достаточной дифференцирующей способностью. Основываясь на результатах проверки, даны рекомендации по повышению валидности и надежности тестовых заданий.

Таким образом, в нашей статье с помощью статистических характеристик, вычисленных в системе MOODLE, были найдены легкие тестовые вопросы и вопросы с низкой дифференцирующей способностью. Изменение этих вопросов позволит получить валидные тестовые задания.

Ключевые слова: тестовые задания, валидность тестов, система дистанционного обучения MOODLE, дифференцирующая способность тестов.

Zahrebelny S. L., Kostikov O. A., Brus M. V. Evaluation of the quiz validity in the learning management system MOODLE. *Herald of the DSEA*. 2020. № 1 (48).

In connection with the widespread introduction of remote technologies in the educational process, adequate assessment of students' knowledge with the help of the quizzes is of great importance. Therefore, assessing the validity of quizzes is an urgent task.

The article considers the possibilities of the learning management system MOODLE for checking the validity of the quizzes designed to control students' knowledge. The quizzes validity was evaluated in three disciplines: "Computer Technique and Programming", "Numerical Methods and Computer Modeling", "Economic Informatics", which were developed by the authors. Validation was performed using the MOODLE system, which consists in assessing the degree of deviation from the normal law of quizzes results and in the procedure for their normalization. Based on the founded values of asymmetry and excess of the distribution law of control results, valid and invalid quizzes were identified. Also the following statistics of the developed quizzes for control of student's knowledge were received: facility index, standard deviation, random guess score, the intended question weight, the effective question weight, discrimination index, discriminative efficiency. It is considered in detail how to use these statistics to evaluate the effectiveness of the quiz, mathematical formulas to calculate these statistics are given also. Using the values of the above statistics, the questions with low item discrimination and the very easy quiz questions were found. Based on the quiz results, recommendations are given to increase the validity and reliability of the quiz questions.

Thus, in our article, with the help of statistics calculated in the MOODLE system, easy quiz questions and questions with low item discrimination were found. Changing these questions will allow to get valid quiz.

Keywords: quiz questions, quiz validity, Learning Management System MOODLE, item discrimination.

Стаття надійшла до редакції 11.09.2019 р.