

Висновки

1. Застосування IRT-методів дає змогу об'єктивніше поглянути на тестові завдання, які переважно складають на підставі досвіду та інтуїції викладача, і значно спрощує первинний аналіз результатів тестів.
2. Впровадження методики оцінювання параметрів за моделлю Андерсена підтверджує її зручність та ефективність порівняно з моделлю Раша–Мастерса.
3. Збільшення кількості студентів, охоплених тестуванням, підвищило вірогідність результатів.
4. Автори вбачають подальшу перспективність втілення відповідних методик для аналізу контролю знань студентів різних форм освіти.

1. Алексеева I.B., Гайдей В.О., Диховичний О.О., Коновалова Н.Р., Федорова Л.Б. *Статистичний аналіз тестових завдань із застосуванням сучасних математичних моделей. – Інноваційні технології у вищій школі / Матеріали III Наук.-практ. конф.– Львів, 18–20 жовтня 2011. – С. 119–123.* 2. Rasch G. *Probabilistic Models for Some Intelligence and Attainment Tests.* – Chicago: Univ. of Chicago Press, 1980. 3. Linden W., Hambleton R. *Handbook of Modern Item Response Theory.* – NY: Springer-Verlag, 1997. – 510 p. 4. Masters, G. N. *A Rasch model for partial credit scoring // Psychometrika.* – 1982. – Vol. 47, №2, June 1982. – p. 149-174. 5. Andersen E.B. *A general latent structure model for contingency table data / Wainer H., Messick S. Principles of modern psychological measurement.* – New Jersey: Lawrence Erlbaum Associates, 1983.

УДК 709.4, 710.5

О.Б. Біленька

Національний університет “Львівська політехніка”

МОЖЛИВОСТІ АНАЛІЗУ ТА ПІДВИЩЕННЯ ЯКОСТІ ТЕСТОВИХ ЗАВДАНЬ З ВИКОРИСТАННЯМ СИСТЕМИ УПРАВЛІННЯ НАВЧАННЯМ MOODLE

© Біленька О.Б., 2012

Описано, як за допомогою вбудованих засобів для автоматизації обчислень статистичних показників за результатами виконання тестових завдань можливо контролювати та об'єктивно оцінювати якість тестів.

Ключові слова: тест, тестове завдання, курс, студент, віртуальне навчальне середовище.

Control and objective estimation of a test' quality are possibly carry out with help of build-in means for statistic indeces calculation by executive results of the test tasks.

Keywords: test, test task, e-course, student, virtual learning system.

Постановка проблеми

Широке використання електронних систем управління навчанням не в останню чергу обумовлене наявністю в них засобів, які дають змогу зробити технологічнішою найважливішу складову процесу навчання, яка пов'язана з оцінюванням рівня засвоєння студентами навчального матеріалу, що вивчається. Йдеться про використання автоматизованих тестових систем як засобу вимірювання рівня підготовки студентів.

Для цього в системі Moodle, як і в інших сучасних системах аналогічного призначення, є підсистема, що дає змогу формувати банки тестових завдань всіх основних форм: “у закритій формі”, “обчислювані”, “коротка відповідь”, “числові”, “на відповідність”, “вкладені відповіді” – та вибудовувати з них тести, спрямовані на самоконтроль та тренінг, тематичні тести, тести для поточного та підсумкового оцінювання рівня знань студентів та ін.

Основний матеріал

Для викладача, який використовує такого роду засоби для викладання своїх навчальних курсів, найбільша складова його праці полягає в створенні банку тестових завдань для перевірки знань студентів та оцінювання їхньої успішності. При цьому завжди виникає питання про якість розроблених та використаних тестових завдань з погляду їх здатності адекватно розв’язати задачу вимірювання рівня підготовки студентів, правильного їх диференціювання відповідно до рівня їх підготовки. Річ у тім, що тестові завдання, що задовольняють формальні вимоги, наприклад, правильні за формою, логічно чіткі та несуперечливі тощо, на практиці можуть не володіти якостями, які давали б змогу розглядати їх як повноцінний засіб вимірювання рівня підготовки студентів.

Питання про спроможність конкретних контрольно-вимірювальних матеріалів насправді слугувати засобом вимірювання рівня підготовки студентів виникає під час використання будь-якої системи оцінювання, зокрема й під час традиційних екзаменаційних випробувань в усній чи письмовій формі. Але саме з використанням автоматизованих тестових систем виникає можливість реально застосовувати формалізовані підходи для оцінювання якості використовуваних контрольно-вимірювальних матеріалів, що ґрунтуються на методології теорії педагогічних вимірювань. Стає можливим накопичити та статистично обробити результати досліджень та отримати кількісні характеристики для оцінювання якості тестових контрольно-вимірювальних матеріалів.

Практичне використання підходів, що пропонуються цією теорією, полегшується тим, що в системі керування навчанням Moodle є вбудовані засоби для автоматизації обчислень за результатами виконання тестових завдань статистичних показників, які дають змогу об’єктивно оцінювати якість тестових завдань з погляду їх здатності слугувати засобом вимірювання рівня підготовки студентів.

Обробка результатів випробувань, необхідна для отримання характеристик якості тестових завдань, полягає в такому. Відповідно до педагогічної теорії вимірювань отримані студентами результати випробувань подають у вигляді матриці (таблиця).

Матриця результатів тестових завдань

	Тестовані студенти (i=1,2,...N)					$P_j = \sum_i x_{ij}$
Завдання (j=1,2,...K)	x_{11}	x_{12}	...	$x_{1,N-1}$	x_{1N}	P_1
	x_{21}	x_{22}	...	$x_{2,N-1}$	x_{2N}	P_2

	$x_{K-1,1}$	$x_{K-1,2}$...	$x_{K-1,N-1}$	$x_{K-1,N}$	P_{K-1}
	x_{K1}	$x_{K,2}$...	$x_{K,N-1}$	x_{KN}	P_K
$s_j = \sum_i x_{ij}$	s_1	s_2	...	s_{N-1}	s_N	

У стовпцях цієї матриці знаходяться результати, отримані конкретними студентами за кожним тестовим завданням, а горизонтально розміщуються результати кожного студента за відповідним тестовим завданням. У нижньому рядку наведено сумарні значення балів, які набрав кожен студент за тест ($S_i = \sum_j x_{ij}$). Ці значення використовують для формування оцінки рівня підготовки студентів за навчальним матеріалом, який перевіряє даний тест. У правому стовпці наведено суми балів, отриманих всіма студентами для кожного тестового завдання (для j-го

завдання $P_j = \sum x_{ij}$). За величиною P_j вже можна судити про ступінь реальної легкості (або складності) конкретних тестових завдань для студентів цієї групи – що більше студентів правильно відповіло на конкретне завдання, то це завдання в середньому є легшим для виконання. Для практичного використання тим не менше як міри легкості (складності) завдань придатніші є відносні значення такого роду показників. Це – індекс легкості завдання (ІЛ) та індекс складності завдання (ІС):

$$IL_j = \frac{x_{серj}}{x_{максj}} = \frac{\sum_i x_{ij}}{N \cdot x_{максj}} = \frac{P_j}{N \cdot x_{максj}}, \quad IC_j = 1 - IL_j,$$

де $x_{серj}$ – середня кількість балів, що набрали всі студенти за виконання j -го завдання, $x_{максj}$ – максимально можлива кількість балів за виконання j -го завдання. Конкретні значення ІЛ та ІС тестових завдань залежать від рівня підготовки студентів та не залежать від їх кількості. Важливість отримання кількісних характеристик легкості (складності) завдань полягає в тому, що для того, щоб ці завдання володіли здатністю диференціювати студентів за рівнем їх підготовки, тобто служити вимірювання цього рівня, їх складність повинна відповідати середньому рівню підготовки студентів. Тест загалом має містити комплекс завдань різної складності, але занадто легкі та занадто складні завдання не можуть бути мірилом диференціювання студентів за рівнем підготовки і тому не можуть бути справжніми тестовими завданнями. Отже, за впорядкованими значеннями індексу легкості/складності тестових завдань (як на рис. 1), можна оцінити ступінь відповідності завдань конкретного тесту середньому рівню підготовки студентської групи.

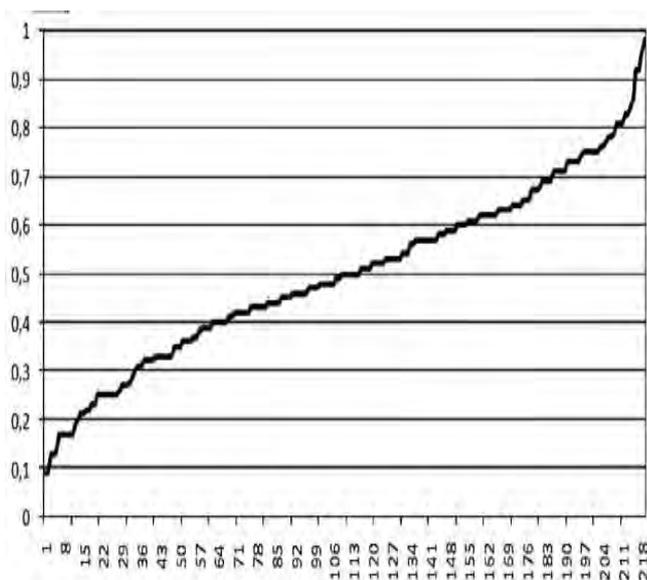


Рис. 1. Значення індексу легкості завдань для тестованої групи

Іншою характеристикою, яку дають змогу обчислити вбудовані в систему Moodle засоби аналізу тестових завдань, є дисперсія цих результатів (їх середньоквадратичне відхилення):

$$s_j^2 = \frac{1}{N} \sum (x_{ij} - x_{серj})^2,$$

яка характеризує розкид балів, отриманих всіма студентами, які

відповідали на конкретне завдання тесту. Якщо відповідь всіх студентів однакова, то розкид дорівнює нулеві. Нульове або низьке значення розкиду результатів свідчить про низьку диференціальну здатність цих завдань, тобто їх слабку здатність розділяти студентів в групі за ступенем їх підготовки. Такі завдання виключаються з тесту. Що більша диференціальна здатність тестових завдань, то вищою є якість тесту.

Як видно з рис. 2, для більшості тестових завдань значення середньоквадратичного відхилення є більшим за 0,3, що є добрим показником диференціовальної здатності тестових

завдань. Ще однією важливою характеристикою якості тестових завдань, яку можна обчислити засобами Moodle, є коефіцієнт кореляції множини значень відповідей, отриманих студентами під час виконання конкретного завдання, з результатами виконання тесту загалом:

$$KD_j = \frac{1}{N \cdot s_j \cdot s_s} \sum_i (x_{ij} - x_{сепj})(s_i - s_{сеп}); \text{ де } s_s^2 = \frac{1}{N} \sum_i (s_i - s_{сеп})^2, \quad s_{сеп} = \frac{1}{N} \sum_i s_i$$

де $s_{сеп}$ – середнє значення балів, отриманих всіма N студентами за тест загалом, s_i – сума балів i -го студента за виконання всіх завдань тесту. Цей показник (коефіцієнт диференціації КД, рис. 3) може набувати значення між -1 та $+1$, тому що в якісному тесті в середньому вищі бали у разі відповіді на кожне конкретне питання повинні отримувати студенти, які набирають вищу суму балів за виконання тесту загалом, тобто між цими значеннями повинна бути кореляція – від’ємні значення коефіцієнта свідчать, що слабші студенти в середньому отримують вищі бали, ніж сильніші. Ясна річ, що такі завдання (внаслідок допущених помилок в їх формулюванні або за інших причин) не є справжніми тестовими завданнями, отже, їх слід також видаляти з тесту. Вважається допустимим, якщо значення цього коефіцієнта перевищує $+0,3$. З рис. 4 видно, що близько 30-ти тестових завдань не задовольняють ці вимоги, що свідчить про їх явну дефектність.

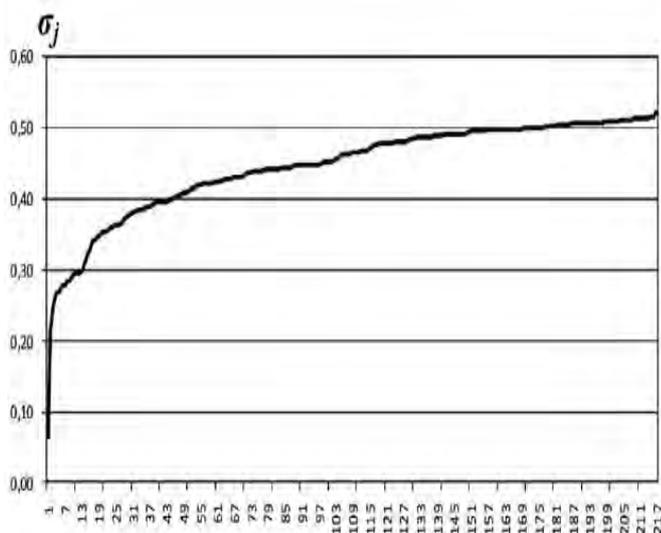


Рис. 2. Середньоквадратичне відхилення результатів виконання завдань

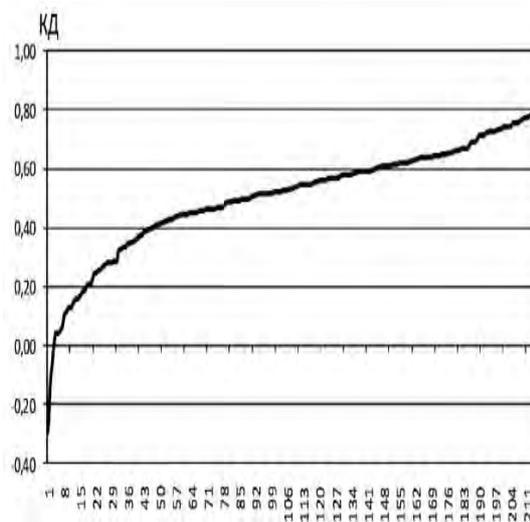


Рис. 3. Коефіцієнт диференціації тестових завдань

Висновки

Отже, засоби Moodle не лише гарантують можливість практичної реалізації автоматизованого контролю рівня підготовки студентів, але й надають викладачеві ефективний механізм покращання якості тестових завдань та підвищення точності та об’єктивності оцінювання рівня студентів. Для протидії можливій компрометації процедури тестування (доступ студентів до бази даних питань тесту тощо) використовується цілий ряд відомих засобів. По-перше, використання в тестових завданнях множинного вибору опції випадкового перемішування відповідей при кожному завданні, що виключає пов’язування правильної відповіді з порядковим номером варіантів відповідей. Крім того, фіксований набір завдань не може бути невідомим та несподіваним для студентів протягом тривалого часу, тому найефективнішим є формування сценарію тесту шляхом випадкового вибору кожного конкретного тестового завдання з множини однотипних завдань, однорідних за рівнем складності. Так сформувати тести можливо лише за наявності достатньо великого банку тестових завдань за кожним розділом навчальної дисципліни. Але слід враховувати, що вищенаведені показники тестових завдань є статистичними величинами, які отримуються шляхом усереднення, отже, забезпечення точності їх експериментального визначення принципово вимагає їх апробації на

більшій кількості студентів. Але ця вимога явно суперечить необхідності практичного формування сценарію тестування шляхом випадкового вибору тестових завдань з більш-менш великого банку завдань, метою якого є зниження імовірності введення одних і тих самих тестових завдань у різних сеансах тестування за одним і тим самим тестом. Зрозуміло, що в цьому випадку, навіть за більшої загальної кількості студентів, що реально беруть участь у тесті, кількість тих, хто отримує однакові тестові завдання, буде значно нижчою, що знижує надійність статистичних показників якості тестових завдань під час їх аналізу. Це протиріччя обов'язково слід враховувати як під час формування сценаріїв тестування, так і під час інтерпретації статистичних показників якості тестових завдань, що розраховуються системою, для пошуку необхідного компромісу.

1. Створення навчальної дисципліни у віртуальному навчальному середовищі Львівської політехніки: Посібник / Укл. Д.В. Федасюк, Л.Д. Озірковський, В.М. Якубенко. – Львів: Вид-во Нац. ун-ту “Львівська політехніка”, 2009. – 60 с. 2. Мясникова Т.С., Мясников С.А. Система дистанційного обучения MOODLE. – Харьков, 2008. – 232 с. 3. Алексеев А.Н. Дистанционное обучение инженерным специальностям. – М.: Университетская книга, 2005. – 333 с. 4. www.moodle.org. 5. <http://moodle.udec.ntu-kpi.kiev.ua>

УДК 37.013: 303.448

Г.А. Буцак

Національний університет “Львівська політехніка”

РОЗРОБЛЕННЯ ТЕСТОВИХ ЗАВДАНЬ ТА ЇХ ЗАСТОСУВАННЯ ДЛЯ ПІДСУМКОВОГО КОНТРОЛЮ ЗНАТЬ СТУДЕНТІВ З ГУМАНІТАРНИХ ДИСЦИПЛІН: АНАЛІЗ ДОСВІДУ

© Буцак Г.А., 2012

Тестування має низку важливих переваг. Ці переваги існують лише у разі дотримання технології розроблення тестів. Підсумкове тестування сприяє повнішій і гармонійній реалізації дидактичних принципів. Тестове оцінювання навчальних досягнень студентів виконує функцію відслідковування структури знань студентів, що дає змогу удосконалювати зміст окремих тем та методи їх викладання.

Ключові слова: тест, дидактичні принципи, репрезентативність тесту, структура знань студента.

Testing has several important advantages. These benefits exist only if the technology development test. Final testing contributes to a more complete and harmonious implementation of didactic principles. Test evaluation of educational achievements of students serves as tracking students' knowledge structure that allows to improve the content of individual topics and methods of teaching.

Key words: test, didactic principle, representative test, student knowledge structure.

Вступ

Нова освітня філософія спрямована на розкриття потенційних можливостей та здібностей студентів. Такі пріоритети спонукають до розвитку педагогічних технологій. В умовах інформаційного суспільства реалізацію дидактичних принципів забезпечують педагогічні технології, що