ПЕДАГОГІЧНІ АСПЕКТИ НАВЧАЛЬНОГО ПРОЦЕСУ

УДК37.011

МЕТОДИКА ОЦІНЮВАННЯ РІВНІВ СКЛАДНОСТІ НАВЧАЛЬНИХ ТЕСТІВ

Зіньковський Ю.Ф., Мірських Г.О.

Сучасний освітянський процес неможливо уявити собі без використання тих чи інших форм тестування. Навчальні тести (тестові завдання) все ширше використовуються на різних етапах очної та заочної форм навчання, становлять основу контролю знань за умови використання у навчальному процесі сучасних інформаційних технологій, утворюють методологічну основу дистанційного навчання. Розроблення та використання навчальних тестів вимагає розв'язання ряду задач, які, за звичай, не виникають під час використання традиційних форм контролю знань, пов'язаних з безпосереднім спілкуванням студента та викладача. Серед таких задач чи не найголовнішим є визначення рівня складності тестових завдань, що мають бути запропоновані студенту під час тестування, аби адекватно визначити рівень його знань.

Поняття рівня складності тестового завдання має системний характер. З одного боку, це поняття невідділиме від студента, якому слід пройти тестування, тобто може визначатися (бути оцінено) з точки зору студента. При цьому, за звичай, суттєву роль відіграє самооцінка студента, відносно його знань, яка в значній мірі пов'язана зі сприйняттям цим студентом комплексу факторів, серед яких і якість викладання відповідної дисципліни, і авторитет викладача, і соціальні умови оточуючого його середовища, і умови проведення тестування, і навички розв'язання подібних завдань тощо. З другого боку, це поняття, звичайно, тісно пов'язане з викладачем відповідної дисципліни, з його самооцінкою, сприйняттям своєї ролі та ролі студентів, як учасників навчального процесу тощо. Звичайно фактори, що впливають на рівень складності, як і сам рівень складності не можуть бути описані кількісними категоріями. Мова може йти лише про їх якісну оцінку, яка базується на ранжирування тестових завдань за рівнем складності. При цьому слід враховувати, що навчальний тест та компоненти, якими визначається його складність, відносяться до об'єктів нечислової природи і, звичайно, їх аналіз слід здійснювати з використанням саме нечислової статистики [1,2].

Оцінювання рівня складності навчальних тестів

З наведеного вище випливає, що оцінювання складності тестових завдань має відбуватися з урахуванням як з точки зору студента, якому належить пройти тестування, так і точки зору викладача відповідної дисципліни, який одночасно (найчастіше) і є розробником цих тестових завдань.

При проведенні експертизи рівня складності навчальних тестів задається шкала порядку (порядкова шкала) з N ступенями відмінності, кожен з яких відповідає визначеному рівню складності. Група експертів (за звичай, в їх ролі виступають викладачі, що проводять заняття з відповідної дисципліни) визначає складність, ставлячи у відповідність i-му завданню відповідний ранг $x_{ij} \in [1,...,N]$, де j-номер експерта. Результатом такої роботи експертів є відповідний масив даних, який слугує основою розподілення тестових завдань за рівнем складності.

В якості прикладу в таблиці представлено експертне оцінювання рівня складності для 10 тестових завдань [1], яке виражене у порядковій шкалі з трьома ступенями відмінності:

1 – завдання зі складністю, що визначається як базова; надання правильних відповідей на завдання цього рівня складності гарантує отримання оцінки "задовільно";

2 – завдання з підвищеною складністю; надання правильних відповідей на завдання цього рівня складності гарантують отримання оцінки "добре";

3 – складні завдання; надання правильних відповідей на завдання цього рівня складності гарантує отримання оцінки "відмінно".

В таблиці інформація щодо рівня складності надана у порядковій шкалі, і тому для отримання середньої величини використовується не математичне сподівання, а медіана [2]. Видно, що одне завдання з номерами 1,5,7,10 є складними, завдання з номерами 2,4,6,8,9 мають підвищену складність, завдання з номером 3 відповідає базовому рівню складності.

Таблиця. Результати експертизи рівнів складності завдань

		Номер завдання									
		1	2	3	4	5	6	7	8	9	10
експерт	1	2	2	1	2	3	1	3	2	2	3
	2	3	1	1	2	3	3	3	3	2	3
	3	3	2	2	2	3	2	3	2	3	3
	4	3	2	1	2	2	2	3	2	2	1
	5	3	2	1	2	3	2	3	2	2	3
медіана (p)		3	2	1	2	3	2	3	2	2	3

Зважаючи на те, що в ролі експертів, за звичай, виступають викладачі, за результатами наведеної експертизи можна оцінити рівень складності завдання з точки зору його укладачів.

Для оцінки рівня складності завдань з точки зору студентів, яким належить пройти тестування, необхідно здійснити аналіз результатів тестування контрольної групи. При цьому в якості міри складності, за звичай, використовують відношення кількості студентів W, що відповіли на дане завдання невірно, до загальної кількості студентів L у контрольній групі, а

саме: $p^* = \frac{W}{L}$. Значна різниця між величинами p та p^* (після переведення останньої у порядкову шкалу відповідно до N ступенів складності) може слугувати ознакою некоректності формулювання тестового завдання або невірного сприймання його з боку студентів (що і призводить до невірних відповідей).

Встановлення рівнів складності навчальних тестів

Встановлення рівнів складності навчальних тестів (що за своєю суттю означає розроблення моделі тестування з урахуванням рівня складності навчальних тестів) базується на аксіоматичному положенні: рівень складності одного складного завдання не дорівнює рівню складності двох, трьох, або більшої кількості легких завдань, при цьому можна (і слід) говорити про факт відмінності рівнів складності окремих тестових завдань. Щодо величини цієї відмінності, то для її визначення використовуються положення статистики нечислових даних.

Вважатимемо, що сукупність всіх підготовлених тестів утворює множину X, а сукупність тестів, які подаються для тестування в конкретному випадку, - множину X^*. При цьому, звичайно $X^* \subset X$ та, відповідно

$$\forall\left\{i \in [1...N(X^*)]\right\} \exists\left\{j \in [1...N(X)]\right\} : x_i^* \equiv x_j,$$

де $N(X^*)$ - кількість тестів, що подаються для тестування в конкретному випадку (тобто кількість завдань в базі X^* – кількість елементів множини X^*); $N(X)$ - загальна кількість тестових завдань (тобто кількість завдань в базі X – кількість елементів множини X); x_i^* - елемент множини X^* - окреме завдання, з тих що подані для тестування; x_j - елемент множини X - окреме завдання, з тих що входять до бази завдань.

Звичайно, кількість тестів, що подані для тестування не може перевершувати загальну кількість тестів у базі, тобто $N(X^*) \le N(X)$.

Крім того, має виконуватись вимога, згідно якої у множині завдань, що надаються студенту в ході тестування, однакові завдання відсутні:

$$\forall\left\{i, j \in [1...N(X^*)]\right\} : i \ne j \to x_i \ne x_j.$$

З метою коректного врахування рівня складності завдань необхідно, щоб виконувались умови:

- будь-яке завдання з числа тих, що подані для тестування, має відповідати одному з визначених рівнів складності, тобто

$$\forall\left\{i \in [1...N(X^*)]\right\} \exists\left\{j \in [1...N(S)]\right\} : S(x_i) = s_j,$$

де $N(S)$ – кількість рівнів складності; - $S()$ – упорядкована за зростанням множина рівнів складності завдань; $S(x_i)$ – складність завдання x_i; s_j – j-

й рівень складності завдання, тобто елемент множини $S()$.

Кожне тестове завдання з множини X або X^*. має бути віднесене до відповідного рівня складності. При цьому, звичайно $\bigcup_{i=1}^{N(S)} X_{s,i} = X$ та $\bigcap_{i=1}^{N(S)} X^*_{s,i} = X^*$, де $X_{s,i}$ - підмножина множини X, що складається з тестів рівня складності s_i; $X^*_{s,i}$ - підмножина множини X^*, що складається з тестів рівня складності s_i. Природно виставити вимогу, згідно якої ситуація, що передбачає віднесення одного й того ж тесту до різних рівнів складності не припустима, тобто $\forall\{i, j \in [1...N(S)]\} : i \neq j \rightarrow X_{s,i} \cap X_{s,j} \equiv \varnothing$, та $\forall\{i, j \in [1...N(S^*)]\} : i \neq j \rightarrow X^*_{s,i} \cap X^*_{s,j} \equiv \varnothing$.

При виконанні всіх наведених умов оцінка результатів тестування може проводитися, наприклад, наступним чином.

1. Для кожного рівня складності s_i задається пара натуральних чисел (N_i, P_i), якими визначається кількість завдань даного рівня складності та поріг виконання рівня (кількість завдань даного рівня, що має виконати особа, яка проходить тестування, щоб отримати залік відповідно до визначеного рівня складності розв'язаних завдань). Ці числа мають відповідати наступним вимогам:

- сума завдань кожного з рівнів складності має дорівнювати загальній кількості завдань, що надаються для тестування: $\sum_{i-1}^{N(S)} N_i = N(X^*)$;

- в тесті мають бути присутні завдання всіх рівнів складності:

$$\forall\{i \in [1...N(S)]\} : N_i > 0;$$

- кількість завдань, які мають бути виконані для того, щоб відповідний рівень складності був пройдений: $\forall\{i \in [1...N(S)]\} : N_i \geq P$;

- при зростанні рівня складності кількість завдань не повинна збільшуватися: $\forall(s_i, s_j) : s_i > s_j \rightarrow N(X_{s,i}) \geq N(X_{s,j})$.

2. Вибираємо та забезпечуємо порядок подання завдань в порядку зростання їх складності, тобто $\forall\{i, j \in [1...N(X^*)]\} : i > j \rightarrow S(x^*_i) \geq S(x^*_j)$.

3. Після виконання завдань, що відносяться до рівня складності S_i, підраховується кількість R_i правильно виконаних завдань даного рівня складності.

4. За умови, що $R_i \geq P_i$, відповідний рівень складності вважається пройденим і тому, хто тестується, пропонуються завдання більш високого рівня складності. Якщо ж ця умова не виконується, то тестування припиняється, а результатом Y тестування вважатиметься останній рівень склад-

ності завдань, що пройдений тим, хто тестується, тобто $Y = s_{i-1}$.

Відмітимо характерну особливість наведеного алгоритму, яка полягає в тому, що студенту (тому, хто тестується) послідовно пропонуються завдання (точніше, сукупності завдань), рівень складності яких збільшується, при цьому кінець тестування визначається можливостями студента щодо надання правильних відповідей на чергову сукупність завдань. Цей алгоритм доречно визначити як тестування за супремумом складності. За таких умов, на наш погляд, доцільно мати не чотири рівня складності завдань, а значно більше (наприклад, десять). При цьому викладач і студент в результаті тестування можуть отримати для себе значно більше інформації, ніж просте визначення за п'ятибальної (на практиці за чотирибальною) шкалою рівня засвоєння відповідного курсу.

З одного боку, система тестування за супремумом складності дозволяє викладачеві

- отримати інформацію про рівень здібності студентів поза рамками традиційного оцінювання;

- виявити тих студентів, що відрізняються підвищеним рівнем здібності, тобто дозволяє здійснити додаткове ранжирування студентів за рівнем здібності щодо вирішення завдань (в тому числі, нестандартних) з даної дисципліни;

- отримати додаткову інформацію для визначеності з рівнем подання навчального матеріалу в рамках відповідної (а при визначеному підборі завдань і з суміжної) дисципліни.

З другого боку, система тестування за супремумом складності показує студенту його особисті можливості (в тому числі у порівнянні з можливостями інших студентів) з даної (та суміжних) дисциплін, що сприятиме у подальшому з визначенням його особистих схильностей щодо виду професійної діяльності та вибору відповідної галузі знань (промисловості, народного господарства) для реалізації своїх професійних інтересів.

Для реалізації всіх позитивних рис та переваг алгоритму тестування за супремумом складності необхідно визначитися з характером зростання складності тестових завдань поза межами області традиційних рівнів оцінювання ("відмінно", "добре", "задовільно", "незадовільно"). На нашу думку, поза вказаної області складність тестових завдань має зростати нелінійно, тобто "відстань" за шкалою складності має зростати при переході від одного рівня до наступного (таку схему, взагалі кажучи, доцільно використовувати і в межах традиційної області оцінювання, наприклад, на переході від оцінки "добре" до оцінки "відмінно"). Для визначення ступеня вказаної нелінійності можна, наприклад, використати методику експертного оцінювання, подібну до тої, що використана вище при визначенні рівнів складності тестових завдань.

Визначення дискримінативності навчальних тестів

Важливим етапом моделювання процесу тестування є визначення правильності підбору тестів з точки зору їх розподілення за рівнями складності. На цьому етапі слід визначити дискримінативність (коефіцієнт розподільної здатності) тестового завдання, яка характеризує його здатність відділяти учасників тестування з різним рівнем навчальних досягнень. Відповідно до рівнів складності вся множина тестів має розділятися на визначену кількість підмножин, причому складність тестів, що належать різним множинам не може бути однаковою (вказані підмножини не можуть пересікатися). Ця задача може бути розв'язана на підставі визначення зв'язку між статистичними даними щодо розв'язання студентами тестів, які належать до різних рівнів складності [1]. Звичайно, тести, що належать до різних рівнів складності але проявляють тісний вказаний статистичний зв'язок мають бути видалені з бази тестів як недискримінативні. Для отримання необхідної інформації щодо такого видалення можна використати результати експертного оцінювання рівня складності тестових завдань, доповненого проведенням тестування контрольної групи. При цьому на підставі попереднього експертного оцінювання задається величина $S(x_i)$, що визначає відповідний рівень складності тестового завдання. Цю величину можна вважати оцінкою складності та дискримінативності тестів з боку викладачів, які, за звичай, виступають у ролі експертів. Надалі проводиться серія тестів T (тобто здійснюється перевірка знань контрольної групи студентів, використовуючи сформовані набори тестових завдань), причому $\bigcup_{i=1}^{N(T)} X_i^* = X$, де $N(T)$ – загальна кількість тестових завдань, яка буде мінімальною за умови: $\forall\{i, j \in [1...N(t)]\} : i \neq j \rightarrow X_i^* \cap X_j^* \approx \varnothing$.

На підставі отриманих результатів тестування згідно кожної групи тестових завдань слід розрахувати медіанні значення кількості студентів, які правильно відповіли на відповідні тестові завдання, а отже за результатами тестування можуть бути віднесені до однієї з чотирьох головних підгруп (відповідно отриманим оцінкам "відмінно", "добре", "задовільно", "незадовільно"). За отриманими даними можна визначити коректність експертної оцінки та відсіяти недискримінативні тестові завдання, якщо реалізувати наступний алгоритм

- для будь-якої пари медіанних значень (M_i, M_j), для яких $s_i > s_j$, має виконуватися умова $M_{ik} \geq M_{jk}$, де K – кількість підгруп за рівнями знань в контрольній групі;

- за допомогою коефіцієнту конкордації [3] визначається відстань між медіанними значеннями кожної пари (M_i, M_j), а саме

$$C = \sum_{k=1}^{K} \frac{\left| M_{ik} - M_{jk} \right|}{K(N_k - 1)}.$$

Тестові завдання вважаються дискримінативними, за умови, що визначений коефіцієнт конкордації не перевершує заданого порогового значення.

Представлена методика формування бази (множини) тестових навчальних завдань з урахуванням їх рівня складності та методика оцінювання результатів тестування дозволяють обґрунтовано підійти до визначення рівня знань студентів за результатами тестового контролю, зменшивши при цьому вплив суб'єктивного чинника, в найбільшій мірі врахувати погляди як викладачів, так і студентів щодо відповідності складності тестових завдань рівню викладання відповідної дисципліни та рівню її засвоєння студентами.

Література

1. Kintsel D.A., Kuznetsov A.V. Non-numerical Approach to the Test Model and Estimation of the Test Parameters//Educational Technology & Society. №10.-2007.-pp. 276-281.

2. Орлов А.И. Нечисловая статистика. М.: МЗ-Пресс, 2004, - 513с.

3. Челышкова М.Б. Теория и практика конструирования педагогических тестов. М.: Логос, 2002, – 432 с.

Зіньковський Ю.Ф., Мірських Г.О. ***Методика оцінювання рівнів складності навчальних тестів.*** *Наведена модель навчальних тестів з оцінюванням рівня їх складності. Запропоновано алгоритм тестування, названий тестуванням за супремумом складності.*

Ключові слова: *навчальний тест, рівень складності навчального тесту*

Зиньковский Ю.Ф., Мирских Г.А. ***Методика оценивания уровней сложности учебных тестов.*** *Приведена модель учебных тестов с оценкой уровня их сложности. Предложен алгоритм тестирования, названый тестированием по супремуму сложности.*

Ключевые слова: учебный тест, уровень сложности учебного теста

Zinkovsky J.F., Mirskikh G.O. ***Method of estimate level complication of the educate test****. A model of educational tests assessing their level of complexity. An algorithm for testing, named testing for the supremum of complexity.*

Key words*: educational test, the difficulty level of educational test*