

РОЗДІЛ IV. ОСВІТНІЙ ПРОЦЕС У ВИЩІЙ ШКОЛІ

УДК 373.512.63

О. М. Алексєєв

Сумський державний університет

Г. В. Алексєєва

Сумський державний педагогічний
університет ім. А. С. Макаренка

ПОРІВНЯЛЬНИЙ АНАЛІЗ ПРОЕКТНОГО ТА КОРЕКТУВАЛЬНОГО МЕТОДІВ ВИЗНАЧЕННЯ СКЛАДНОСТІ ТЕСТОВИХ ЗАВДАНЬ

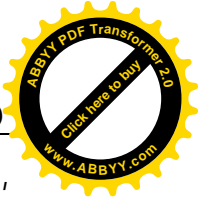
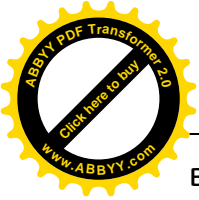
У статті за результатами порівняльного аналізу експериментальних методів визначення складності тестових завдань розроблені рекомендації щодо раціонального використання проектного та коректувального розрахунків індексів складності завдань під час проектування тестів для контролю успішності навчання студентів.

Ключові слова: тестовий контроль, імітаційна модель, тестове завдання, складність завдань, індекс складності завдання, проектний розрахунок, коректувальних розрахунок, експертне оцінювання.

Постановка проблеми. При розробленні методів оцінювання навчальних досягнень за допомогою системи тестових завдань призначення індексів складності (вагових коефіцієнтів) є однією з основних процедур, що визначають достовірність результатів контролю знань. За допомогою індексів складності зазначається ступінь труднощів тестового завдання стосовно решти завдань тесту, задається максимальна кількість балів, яку студент може отримати при правильному виконанні завдання і можлива загальна кількість балів за всю відповідь, тобто в кінцевому підсумку встановлюється кількісна та якісна (для встановленої шкали оцінок) характеристика знань.

При цьому практика, що склалася, коли всім питанням апріорно призначаються однакові показники складності, є неприпустимою. Аналогічно до того, коли при усному контролі викладач задає питання студенту, виходячи зі сформованих у нього уявлень про їх складність, при комп'ютеризованому контролі необхідно враховувати, що завдання можуть істотно відрізнитися за складністю одне від одного.

Аналіз актуальних досліджень. У педагогічній науці неодноразово підкреслювалася важливість достовірного виявлення дійсних знань як з позиції діагностики процесу навчання, так і з метою розвитку, виховання студентів та стимулювання їх до отримання знань [1–6]. Актуальність цієї проблеми не знижується і сьогодні, особливо у світлі тенденції до все більшого впровадження засобів комп'ютерної техніки в освітній процес і



використання технічних пристроїв під час комп'ютеризованого тестування, що спостерігається останнім часом.

У публікаціях, присвячених тестовому контролю знань, проблемі якості тесту, і зокрема визначення складності тестових завдань, приділяється значна увага. При цьому існують два основні підходи до кількісного виміру складності завдань. Перший полягає у різного роду абстрактних висновках, що здійснюються на основі порівняння передбачуваного числа і характеру розумових операцій, необхідних для успішного виконання завдання [1, 5]. Такий підхід більшою мірою має теоретичне значення, тому що його практичне застосування обмежене порівняно вузькою галуззю знань, при контролі якої можна з високою ймовірністю встановити кількість виконуваних розумових операцій і пов'язати їх специфіку з мірою складності контрольованого навчального матеріалу. Другий підхід базується на визначенні складності тестових завдань за результатами експериментальних вимірювань [1, 4, 6]. Однак зазначеному підходу властиві обмеження, пов'язані з експериментальним характером вимірювань, і тому для успішного використання конкретних експериментальних методик необхідні додаткові дослідження для уточнення їх меж застосування.

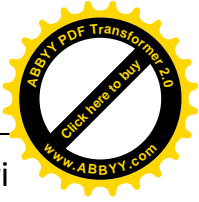
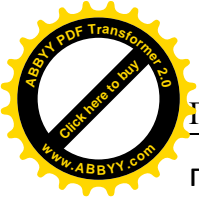
Мета статті – уточнення сфери застосування експериментальних методів визначення складності тестових завдань, що базуються на експертному оцінюванні та обробки результатів тестового контролю знань.

Виклад основного матеріалу. У межах імітаційної моделі розроблено дві методики визначення ступеня складності тестових занять – під час розроблення тестів за допомогою проектного розрахунку і після тестування на основі коректувального розрахунку [2, 3].

Для знаходження сфер раціонального застосування розрахункових методик був виконаний порівняльний аналіз результатів розрахунку складності завдань, отриманих за новими розробленими методиками, і оцінок складності, які встановлювалися традиційно за підсумками експертного оцінювання.

Порівняння полягало в тому, що під час розроблення тестів використовувалися одразу всі три порівнювані методики визначення складності тестових завдань і потім зіставлялись одержані за їх допомогою результати.

З цієї метою під час проектування прототипів тестових завдань



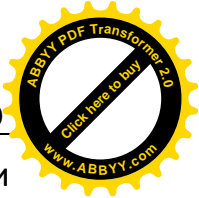
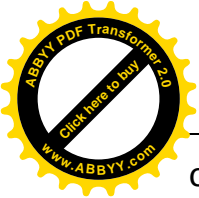
проводилися двоступеневі експертизи. Особливістю було те, що одні й ті самі експерти, незалежно від того, який вид навчальної діяльності студентів проходив під їх керівництвом, давали оцінки складності і теоретичному матеріалу, і практичним умінням, які студенти мали опанувати за результатами вивчення дисципліни.

Перша частина експертизи виконувалася традиційним методом. Експерти давали незалежні оцінки складності тестових завдань, які згодом усереднювались, і для врахування розкиду оцінок розраховувалися дисперсія і довірчий інтервал.

У другій частині експертиза проводилася як компонент проектного розрахунку на базі методики попарних порівнянь [3]. При цьому експерти не виносили прямо судження про складність тестових завдань, а давали оцінку «складніше», «легше» або «однакової складності», порівнюючи завдання. Для визначення меж застосування методу порівняння виконувалося за різної кількості повних циклів зіставлення пар прототипів тестових завдань. Остаточний результат експертизи встановлювався математичною обробкою результатів попарних порівнянь. Додатково експертні оцінки складності тестових завдань усереднювались і розраховувалися дисперсія та довірчий інтервал.

Для забезпечення більшої достовірності висновків експерти з отриманими результатами не ознайомлювались і черговість проведення першої та другої частин експертизи обиралась випадково. Якщо перший експерт спочатку давав оцінки традиційним методом, а потім методом попарного порівняння, то другий і наступні експерти могли проводити експертизу або в тій самій послідовності, або у зворотній – спочатку попарно порівнюючи тестові завдання, а потім незалежно оцінюючи складність кожного завдання. Аналогічно і перший експерт міг починати оцінювати складність завдань спочатку попарним порівнянням, а потім традиційним методом.

Дані експертиз зіставлялися з результатами коректувального розрахунку, що виконувався за результатами тестового контролю [2, 3]. При цьому вважалося, що коректувальний розрахунок найбільш точно відображає складність тестових завдань, оскільки розраховані тут чисельні значення індексів складності безпосередньо залежали від того, наскільки успішно студенти справлялися із завданням. І якщо для одного або



обмеженого числа студентів незадовільне виконання завдання могло мати різні причини, то значення індексу складності, розраховане за результатами тестового контролю великої групи студентів, здебільшого визначалося складністю завдання. Обґрунтованості такого припущення сприяло і те, що під час виконання коректування розрахунку з вибірок попередньо були виключені всі прототипи завдань, які не задовольняли критерії якості [2].

За основу оцінювання прийнята 100-бальна система, як така, використання якої дозволяє точніше диференціювати тестові завдання за складністю, і при цьому доступна для сприйняття експертами під час проведення експертиз традиційним способом. При проектному розрахунку експертиза спрощена, тому що для отримання 100-бальних оцінок експерти працюють за трибальною шкалою – завдання складніше за порівнюване, завдання легше за порівнюване, порівнювані завдання однаково складні.

На рисунку 1 наведена діаграма, що відображає відсоткове відношення довірчого інтервалу Δ до середнього арифметичного індексу складності $Q_{ср}$, встановлене для трьох методів оцінювання складності тестових завдань за різних умов: 1) коректувального розрахунку (КР), при розмірах вибірок, за якими розраховувалися статистичні оцінки, $n = 16$, $n = 156$, $n = 189$ та $n = 252$; 2) експертного оцінювання (ЕКС) при шести експертах та 3) проектного розрахунку (ПР) із числом повних циклів порівняння 1, 5 і 14 (максимальне при 29 тестових завданнях, порівнюваних попарно).

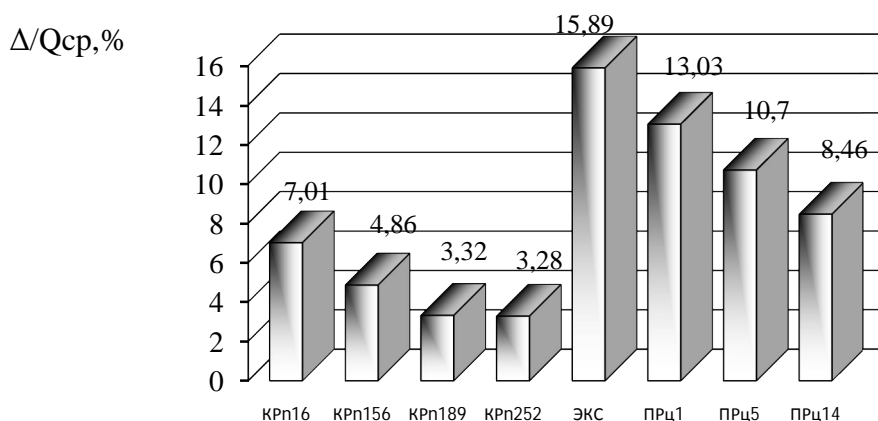
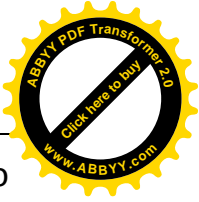
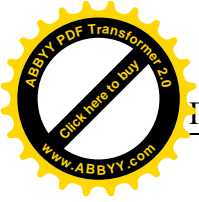


Рисунок 1 – Точність прогнозування складності тестових завдань

$\Delta/Q_{ср}$ – відношення довірчого інтервалу до середнього арифметичного індексу складності; КР – коректувальний розрахунок (n – розмір вибірки); ЕКС – експертне оцінювання; ПР – проектний розрахунок (ζ – кількість повних циклів порівняння)

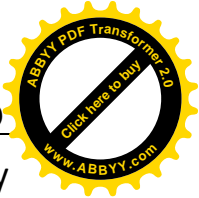
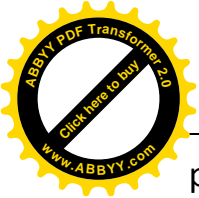


Інтерпретуючи наведені дані, необхідно констатувати, що коректувальний розрахунок дозволяє встановити шукані значення з найменшим довірчим інтервалом, у межах якого математичне сподівання дійсних характеристик складності тестових завдань може відрізнятись від розрахункових. Точність проектного розрахунку при обмеженому обсязі вихідних даних за критерієм $\Delta / Q_{\text{ср}}$ сумірна з експертним оцінюванням. При цьому його прогностичні можливості можна поліпшити, оскільки за збільшенням кількості повних циклів попарних порівнянь зростає точність прогнозу, і при 14 циклах наближається до точності коректувального розрахунку ($\Delta / Q_{\text{ср}} = 8,46\%$ у проектного розрахунку з чотирнадцятьма повними циклами порівнянь і $7,01\%$ у коректувального розрахунку з розміром вибірки $n = 16$).

Однак, незважаючи на більш високу точність, коректувальний розрахунок виконується за результатами вже проведеного контролю і тому не дозволяє забезпечити однакові умови тестування для всіх студентів, у тому числі й тих, навчальні досягнення яких контролюються на самому початку, коли ще немає статистичних даних, необхідних для проведення коректувального розрахунку.

Виходячи з можливості реалізації, коректувальний розрахунок потрібно застосовувати для визначення складності тестових завдань після того, як тестування пройшла хоча б одна студентська група (навіть і для групи з 16 студентів співвідношення $\Delta / Q_{\text{ср}}$ більш ніж удвічі менше, ніж при експертному оцінюванні, і тому його використання доцільніше). Але щоб забезпечити однотипні умови контролю навчальних досягнень для всіх, у тому числі й першої групи студентів, необхідно спочатку для прогнозування складності тестових завдань використовувати розроблений у межах імітаційної моделі проектний розрахунок. Якщо ж навчальну дисципліну вивчає обмежена кількість студентів, наприклад, як це часто буває на випускових кафедрах, де спеціальні дисципліни проходить одна-дві групи студентів на рік, то єдиним способом диференціювання тестових завдань за складністю залишається тільки проектний розрахунок.

Незважаючи на те, що зі збільшенням кількості повних циклів попарне порівняння відношення довірчого інтервалу до математичного сподівання індексу складності зменшується і граничні, найбільша і найменша, прогностичні оцінки складності тестових завдань наближаються до єдиного детермінованого значення, проведення проектного



розрахунку в повному обсязі не завжди може бути рекомендоване у зв'язку з можливими недопустимо великими витратами часу на його проведення. У таблиці 1 наведені приблизні тривалості виконання проектного розрахунку при різних вихідних даних. Таблиця відображає особистий досвід авторів із вибору найбільш прийнятних умов, у яких необхідно використовувати той або інший варіант розрахунку (комірки з кращими умовами виділені затемненим фоном, при нерекондованих або неприпустимих умовах – стоїть знак прочерку).

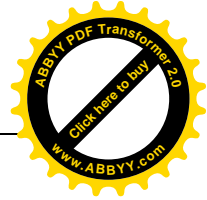
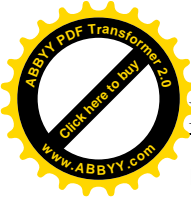
Таблиця 1

Тривалість проектного розрахунку, хв*

Повних циклів	Кількість тестових завдань													
	10	15	20	25	30	35	40	45	50	60	70	80	90	100
1	7	11	14	18	22	26	30	34	38	47	55	65	74	83
2	8	13	17	22	27	33	39	47	52	66	82	99	120	140
3	9	14	20	26	33	40	50	56	68	93	109	132	153	175
5	11	20	28	35	45	58	73	90	113	150	181	213	248	283
7	-	26	35	47	56	69	84	105	123	154	188	233	273	315
10	-	-	44	63	80	99	120	143	175	220	268	307	345	-
15	-	-	-	-	123	149	180	214	250	315	-	-	-	-
20	-	-	-	-	-	-	221	255	283	-	-	-	-	-

* При використанні комп'ютеризованої експертної системи

Отже, залежно від конкретних умов, у яких визначаються індекси складності тестових завдань, може бути застосований проектний або коректувальний метод розрахунку. Проектний метод доцільно використовувати на початку тестування і виконувати з рекомендованою кількістю повних циклів попарних порівнянь. Після того як накопичені статистичні дані, потрібно переходити до коректувального методу розрахунку. При цьому перехід рекомендується здійснювати у кілька етапів. Спочатку – після того як тестування буде проведене в одній студентській групі, а потім у міру накопичення статистичних даних уточнювати прогнозні значення індексів складності відповідно до одержаних нових відомостей про результати контролю успішності навчання досить значних груп студентів. Під час розроблення тестів для дисциплін, які вивчаються невеликим контингентом студентів (як правило, професійно орієнтовані дисципліни), такий перехід може бути здійснено не в поточному, а в подальшому навчальному семестрі за тієї обставини,



що не змінилися статистично значущі умови тестування.

Висновки. Виконаний порівняльний аналіз методів визначення складності тестових завдань дає можливість виявляти найбільш раціональні сфери їх застосування під час розроблення тестів контролю успішності навчання студентів. Урахування рекомендацій що містяться в статті дозволяє для визначення індексів складності завдань використовувати саме ті методики, які найбільшою мірою підходять для конкретних умов тестування. Водночас необхідно вказати на деяку неповноту вироблених рекомендацій. Так, подальшої деталізації потребують рекомендації щодо визначення граничної чисельності студентів, при якій допустимо використовувати коректувальний метод встановлення складності тестових завдань. Необхідно провести додаткове дослідження в частині уточнення величини похибки, що вноситься до результатів тестування при необґрунтованому встановленні складності тестових завдань. Продовження робіт у даному напрямку, уточнення згаданих і, можливо, розроблення нових встановлених рекомендацій дозволить підвищити якість проєктованих тестів і забезпечити необхідну достовірність результатів педагогічних вимірів.

ЛІТЕРАТУРА

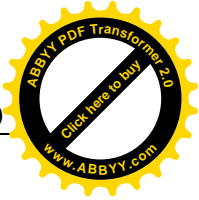
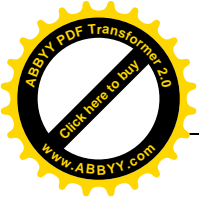
1. Аванесов В. С. Научные проблемы тестового контроля знаний / В. С. Аванесов. – М. : Изд-во ИЦПКПС, 1994. – 136 с.
2. Алексеев А. Н. Дистанционное обучение инженерным специальностям / А. Н. Алексеев. – Сумы : ИТД «Университетская книга», 2005. – 333 с.
3. Алексеев А. Н. Имитационная модель тестового контроля знаний / А. Н. Алексеев, Г. В. Алексеева // Открытое образование. – М., 2010. – №1. – С. 4–11.
4. Ефремова Н. Ф. Тестовый контроль в образовании : учеб. пособ. для студентов, получающих образование по педагогическим направлениям и специальностям / Н. Ф. Ефремова. – М. : Логос, 2007. – 368 с.
5. Майоров А. Н. Теория и практика создания тестов для системы образования: Как выбирать, создавать и использовать тесты для целей образования / А. Н. Майоров. – М. : Народное образование, 2000. – 352 с.
6. Челышкова М. Б. Теория и практика конструирования педагогических тестов : учеб. пособ. / М. Б. Челышкова. – М. : Логос, 2002. – 432 с.

РЕЗЮМЕ

А. Н. Алексеев, Г. В. Алексеева. Сравнительный анализ проектного и корректировочного методов установления сложности тестовых заданий.

В статье по результатам сравнительного анализа экспериментальных методов определения сложности тестовых заданий разработаны рекомендации по рациональному использованию проектного и корректировочного расчетов индексов сложности заданий при проектировании тестов для контроля успешности обучения студентов.

Ключевые слова: тестовый контроль, имитационная модель, тестовое задание, сложность заданий, индекс сложности задания, проектный расчет, корректировочный расчет, экспертное оценивание.



SUMMARY

A. Alekseev, G. Alekseeva. Comparative analysis of design and corrective methods for determining complexity of test questions.

Using the results of comparative analysis of experimental methods for determining the complexity of test questions, the article offers recommendations for rational usage of designed and corrective estimates for indices of test complexity when test questions are being designed.

Key words: test control, imitational model, test question, test complexity, index of test complexity, designed estimate, corrective design, expert evaluation.

УДК 37.011:008/009

І. П. Анненкова

Одеський національний університет
імені І. І. Мечникова

АНАЛІЗ ОСНОВНИХ ПІДХОДІВ ДО ВИЗНАЧЕННЯ ЯКОСТІ ОСВІТИ

У статті проаналізовано сутність поняття «якість освіти». Показано, що якість освіти має два аспекти: внутрішній (відповідність нормі, стандарту) і зовнішній (відповідність потребам споживача). Виділено п'ять основних підходів до визначення поняття «якість освіти»: соціальний, соціокультурний, інституціонально системний, управлінський, комплексний.

Ключові слова: *освіта, якість, якість освіти.*

Постановка проблеми. Соціальні й економічні зміни, що відбулися у нашій країні за останні десятиріччя, призвели до жорсткої конкуренції на ринку праці, яка пред'являє до молодих фахівців все більш високі вимоги. Формується конкурентне середовище, в якому здатний розвиватися і працювати з найбільшою ефективністю тільки фахівець високого рівня кваліфікації. В таких умовах підвищується відповідальність ВНЗ за якість підготовки випускників, причому не тільки перед державою, але й перед молодими фахівцями, які мають пройти через жорсткий відбір на ринку праці. Закон України «Про вищу освіту» надає вищим навчальним закладам великі повноваження у виборі стратегії свого розвитку, педагогічних концепцій і технологій, організаційних методів роботи і формуванні змісту освіти. Це також накладає на навчальні заклади додаткову відповідальність за результати своєї освітньої діяльності. В означених умовах вищі навчальні заклади стають перед необхідністю розробки таких важелів впливу на освітній процес, які б гарантували досягнення поставлених цілей. У зв'язку з цим при розгляді питань сучасного функціонування і розвитку вищої освіти на перший план виходять проблеми якості освіти, управління якістю освіти у вищих навчальних закладах.

Аналіз актуальних досліджень. Тим часом дослідженню цієї