

УДК 373.512.63:004

Алексєєв Олександр Миколайович

доцент, доктор педагогічних наук, професор кафедри «Технології машинобудування, верстатів та інструментів»

Сумський державний університет, м. Суми, Україна

alekseev@tmvi.sumdu.edu.ua

Король Олена Миколаївна

аспірантка

Сумський державний педагогічний університет імені А. С. Макаренка, м. Суми, Україна

Korol.9@mail.ru

**АДАПТИВНІ МОЖЛИВОСТІ ІМІТАЦІЙНОЇ МОДЕЛІ
ТЕСТОВОГО КОНТРОЛЮ ЗНАНЬ**

Анотація. Розглянуто особливості тестового контролю знань з використанням імітаційної моделі. Імітаційна модель базується на сучасних інформаційно-комунікаційних технологіях і наразі максимально спирається на традиції активної участі викладача в діагностиці успішності навчання студентів. Поєднання переваг комп'ютеризованого тестування з математичним обґрунтуванням процедур прийняття рішень екзаменатором розширює область ефективного застосування тестових форм контролю. Відзначено, що застосування імітаційної моделі дозволяє знизити вплив умов, у яких проходило навчання студентів, на достовірність результатів контролю. Описано процедури імітаційної моделі, що мають адаптивні можливості.

Ключові слова: тестування; адаптивний тестовий контроль; формування тесту; імітаційна модель; багаторівневий контроль; оцінювання знань.

1. ВСТУП

Постановка проблеми. У педагогічній літературі неодноразово підкреслювалася важливість педагогічних вимірювань як з позиції діагностики процесу навчання, так і з метою розвитку і виховання студентів, їх стимулювання до здобуття нових знань. Контроль знань набуває особливої значущості у зв'язку з широким упровадженням комп'ютерної техніки в освітній процес і передачею технічним пристроям навчальних і контролюючих функцій викладача. На відміну від традиційного усного іспиту, з більшістю процедур комп'ютеризованого тестового контролю виключена особистість викладача. Такий контроль може не мати повного набору властивостей, необхідних для врахування конкретних умов, у яких відбуваються контрольні заходи. Тому, для підвищення достовірності результатів тестового контролю, що здійснюється з широким використанням комп'ютерних технологій, необхідно проведення додаткових досліджень, спрямованих на розширення його адаптивних можливостей.

Аналіз останніх досліджень і публікацій. Прогресу в розробці нових методів тестування значною мірою сприяли прикладні й теоретичні дослідження таких учених, як І. Є. Булах, М. Р. Мруга, М. М. Олійник, Ю. А. Романенко, S. M. Downing, T. M. Haladyna, K. S. Shultz, D. J. Whitney, M. J. Zickar, D. D. Williams, M. Hricko, S. L. Howell [1–5] та інших. Одним із напрямків, що найдинамічніше розвивається зараз, є той, що пов'язаний із проектуванням адаптивних тестів, у яких запитання, що знову задаються, обираються з урахуванням результатів виконання попередніх завдань. У міру вдосконалення інформаційних і телекомунікаційних технологій, у теорії і практиці педагогічних вимірювань усе ширше знаходить відображення комп'ютеризований тестовий контроль.

Водночас більшість дослідників (О. Г. Колгатін, П. А. Ротаєнко, S. M. Legg, D. C. Buhr, S. J. Osterlind, T. E. Howard, D. V. Smoline [6–10] та інші) відзначають суттєві недоліки тестового контролю, які не вдалося усунути до теперішнього часу. Наукові дані, що отримали вчені, а також результати наших досліджень свідчать, що тестування, засноване на більшості сучасних методик, як і раніше, залишається недостатньо точним і, якщо не вживати заходів, то підміна усного іспиту тестуванням не сприятиме підвищенню достовірності педагогічних вимірювань.

Виходячи з цього, **мета статті** полягає в тому, щоб уточнити адаптивні властивості імітаційної моделі тестового контролю знань, які можуть дозволити повніше враховувати умови, у яких здійснюються контрольні заходи, проаналізувати можливості її використання для підвищення достовірності результатів контролю.

2. МЕТОДИ ДОСЛІДЖЕННЯ

Для досягнення поставленої мети використовувалися методи дослідження: теоретичні: аналіз філософської, психолого-педагогічної, науково-педагогічної та методичної літератури з проблеми дослідження, аналіз освітніх стандартів і навчальних програм — для узагальнення теоретичних питань сутності, структури, функцій, особливостей застосування тестового контролю знань у навчальному процесі; емпіричні: аналіз результатів діяльності викладачів і студентів, спостереження, співбесіда, анкетування, експертне оцінювання — для збору даних про практичний стан досліджуваної проблеми й експериментальної перевірки достовірності окремих теоретичних положень імітаційної моделі тестування; математичні: статистичні та ймовірнісні методи обробки даних — для аналізу експериментальних даних і перевірки ефективності розробленої методики тестування.

3. РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

Для проведення тестового контролю за участю авторів розроблено імітаційну модель [11], що заснована на моделюванні діагностуючих функцій викладача під час усного іспиту. У табл. 1 перераховані процедури імітаційної моделі, які, як встановлено, у більшій мірі мають адаптивні можливості.

Таблиця 1

Процедури імітаційної моделі

Розробка тестових завдань і формування тестів	Контроль результатів навчання	Підведення підсумків контролю
<ul style="list-style-type: none"> ✓ Типізація тестових завдань ✓ Розрахунок кількості завдань і формування тесту 	<ul style="list-style-type: none"> ✓ Використання нечіткої логіки ✓ Багаторівневий контроль 	<ul style="list-style-type: none"> ✓ Аналіз достовірності статистичних відмінностей ✓ Оцінювання знань

3.1. Розробка тестових завдань і формування тестів

3.1.1. Типізація тестових завдань

Комп'ютеризоване тестування знань має на меті постановку завдання у формалізованому вигляді так, щоб відповідь можна було б віднести до одного з типізованих видів і потім проконтролювати відповідними методами.

В імітаційній моделі передбачена можливість використання тринадцяти типів тестових завдань (встановлення відповідності, класифікація термінів, позиціонування малюнка, введення символів, підтвердження правильності альтернативного твердження, множинний вибір, підстановка пропущених слів, відтворення правильної послідовності, наявність або відсутність ключових слів, виправлення помилки в тексті, послідовний вибір, послідовність дій, управління об'єктом), що описані в роботах [12, 13]. Адаптивність тестової системи в даному випадку припускає можливість вибирати ті з її варіантів, які найбільше відповідають виду контролюваного навчального матеріалу. Наприклад, якщо потрібно проконтролювати знання однієї з хронологічних дат, то завдання можна сформулювати за типом множинного вибору — запропонувати у відповіді кілька дат, одна з яких буде правильною. Однак доцільніше буде використовувати тип завдання на введення символів. У цьому випадку студент, відповідаючи, має сам написати в полі введення правильну дату. Перевага такого формулювання завдання в тому, що студент не бачить і, як наслідок, не запам'ятовує неправильні відповіді. Попри це, йому важче відгадати відповідь. Аналогічно, якщо потрібно проконтролювати знання деякої хронологічної черговості подій, то слід використовувати тип завдання на відтворення правильної послідовності і т. д. Імітаційна модель надає в розпорядження викладача різні типи завдань й, отже, дає можливість точніше сформулювати завдання, яке більшою мірою буде відповідати критерію адаптивності щодо змісту навчального матеріалу.

3.1.2. Розрахунок кількості завдань і формування тесту

Щоб мати можливість порівнювати результати перевірки знань різного за змістом навчального матеріалу, наприклад, окремих тем з фізики і хімії, необхідно щоб складність тестів, що використовуються для контролю, була б зіставлена між собою.

За чисельну міру складності одиничного тестового завдання може бути прийнятий рівень складності q_i [14]. У такому випадку сума рівнів складності всіх завдань, включених у тест, буде відображати загальну складність тесту Q_Σ

$$Q_\Sigma = \sum_{i=1}^n q_i, \quad (1)$$

де n — кількість завдань у тесті.

Отже, щоб тест був адаптований щодо складності, необхідно, щоб виконувалася умова

$$Q_{\Sigma_1} = Q_{\Sigma_2} \Rightarrow A \quad (2)$$

або

$$\sum_{i=1}^n q_i = \sum_{j=1}^k q_j \Rightarrow A, \quad (3)$$

де A — критеріальне значення, що обмежує складність тесту, $A = \text{const}$.

Утім, якщо врахувати, що тестові завдання для включення в тест відбираються випадково, то визначити їх кількість простим підбором стає проблематичним. Тому в імітаційній моделі склад тестових завдань обирається відповідно до генетичного алгоритму [15, 16].

Під час формування тестів утворення вихідної популяції на початку еволюційного періоду передбачає створення безлічі з N векторів, кожен з яких є геном окремої особини, формується випадково і є одним із можливих варіантів тесту.

Перехід від вихідної популяції X_0 до X_1 і потім послідовно від X_k до X_{k+1} проводиться шляхом поетапного застосування механізмів репродукції і порівняння якості новостворюваних популяцій з критеріями оптимальності:

$$\begin{array}{ccccccc} & \text{selection} & \text{crossing} & \text{mutation} & \text{survival} & \Psi \in \Psi_0 & \\ \mapsto X_k & \mapsto X_k^s & \mapsto X_k^c & \mapsto X_k^m & \mapsto X_{k+1} & \mapsto Y & \\ & & & & \text{generation } k = k+1 & \downarrow & \Psi \notin \Psi_0 \end{array} \quad (4)$$

де selection, crossing, mutation і survival — відповідно оператори відбору, схрещування, мутації і виживання.

У ролі оператора відбору використовується механізм випадкового відбору, заснований на тому, що більш життєздатна особина має найбільшу ймовірність брати участь у формуванні нащадків. При формуванні тестів це означає, що чим ближче сумарна складність завдань, складена геном особини, до критерію оптимальності A , тим вище ймовірність участі такої особини у створенні потомства.

Сутність дії оператора схрещування полягає в тому, що він повинен забезпечити обмін генами (тестовими завданнями) між двома особинами, що утворюють батьківську пару. В імітаційній моделі для схрещування використовується двоточковий оператор схрещування.

Оператор мутації служить для моделювання природного процесу мутації, покликаний розширити генофонд і, отже, перешкоджати багаторазовому повторенню поколінь з геномоподібними особинами (заціклення задачі).

Послідовне застосування операторів відбору, схрещування і мутації призводить до створення нових особин (варіантів формування тесту), які разом з особинами вихідної популяції складають розширену популяцію, що складається з удвічі більшого числа особин. Загальна чисельність такої популяції доводиться до початкового рівня відбором найбільш сильних особин. Наразі працює оператор виживання, в основі якого лежить механізм визначення якості особини за показниками критерію оптимальності. Особини найгіршої якості виключаються з популяції.

Роботою оператора виживання закінчується репродуктивний цикл й у випадку не отриманого рішення він повторюється відносно нової популяції відповідно до встановленого генетичним алгоритмом. Таке циклічне виконання операторів повторюється до тих пір, поки не буде досягнутий оптимум цільової функції ($F = A - Q_{\Sigma} \cong 0$), тобто поки не сформується тест, який відповідає вимогам критерію адаптивності щодо складності.

3.2. Контроль результатів навчання

3.2.1. Використання нечіткої логіки

Під час тестового контролю висновок про рівень знань студента робиться на основі зіставлення його відповіді з еталоном. Студентові треба дати відповідь на поставлене запитання, вибравши один із запропонованих варіантів або сформулювавши з обмеженого набору слів, букв, цифр або графічних символів свій варіант відповіді. Коли відповідь-еталон точно збігається з відповіддю студента, робиться висновок про те, що тестований знає контрольований матеріал. І, навпаки, якщо відповіді не збігаються — студент відповів неправильно. Однак такий висновок не цілком правомірний, якщо знання студента недостатні, щоб дати однозначну відповідь, він не впевнений, але змушений відповідати в термінах суворої логіки, виходячи за межі власних знань. У цьому випадку в результат тестування вноситься додаткова похибка, тому що враховуються не тільки знання студента, але й можливе вгадування ним правильності відповіді.

Для усунення помилки й адаптації тесту до оцінювання істинних знань студентів в імітаційній моделі пропонується використовувати експертну систему, засновану на апараті нечіткої логіки [17, 18] (рис. 1).

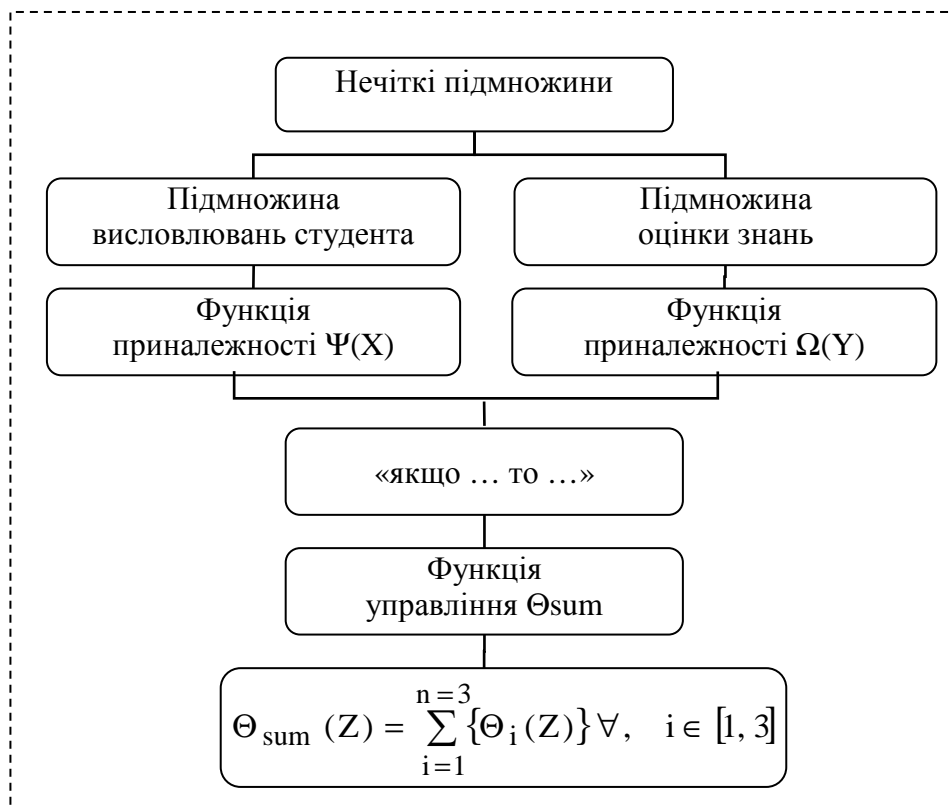


Рис. 1. Експертна система на основі нечіткої логіки

У розробленій експертній системі приналежність оцінки знань студента і висловлюваного ним судження до нечітких логічних підмножин описується кусково-неперервними функціями $\Psi(X)$ і $\Omega(Y)$, що мають прямолінійні початкові і кінцеві ділянки з нульовими рівнями достовірності, і проміжні ділянки, що відповідають переходу до одиничного (максимального) рівня достовірності. Для побудови функції управління Θ_{sum} використовується метод correlation — product encoding. Узагальнення результату за нечіткими множинами у постулюючій $\Psi(X)$ і в констатуючій $\Omega(Y)$ частинах логічного висновку виконується методом Sum combination. Наразі суперпозиція функцій приналежності нечітких множин визначається як

$$\Theta_{sum}(Z) = \sum_{i=1}^{n=3} \{\Theta_i(Z)\} \nabla, \quad i \in [1, 3], \quad (5)$$

де $\Theta(Z)$ — відповідно $\Psi(X)$ або $\Omega(Y)$.

Наявність експертної системи робить можливим оперування не тільки класичними значеннями логічних змінних «хиба» й «істина», а й вживанням їхніх проміжних значень, що плавно переходять від одного крайнього значення («хиба») до іншого крайнього значення («істина»). Наразі для інтерпретації однозначного рішення, висловленого студентом на основі нечітких логічних тверджень, використовується «центр ваги» функції приналежності нечіткої множини.

На рис. 2 зображений елемент управління, включення якого до інтерфейсу програми, що тестує, дозволяє студенту переміщати бігунок в інтервалі «хиба» — «істина» і таким способом висловлювати сумніви в правильності відповіді. Координата

бігунка перераховується із залученням апарату нечіткої логіки і відповідь оцінюється так, що з неї виключається складова, пов'язана з неточною інтерпретацією ступеня впевненості студента.



Рис. 2. Елемент управління

3.2.2. Багаторівневий контроль

Для прийняття рішення щодо успішності навчання сума балів, набрана під час тестування, зазвичай, порівнюється з деяким граничним значенням шкали оцінок. Якщо сумарна кількість балів вище критеріальної величини, то вважається, що студент достатньо засвоїв навчальний матеріал і справляється із завданнями навчання. У протилежному випадку робиться висновок, що студент матеріал не знає, і йому має бути поставлена більш низька оцінка успішності навчання.

Такий висновок правомірний, якщо набрана сума балів істотно більша або менша встановленого критерію. Однак він не цілком виправданий за їх відносної близькості. Дійсно, якщо припустити, що пороговим числом для оцінки рівня знань прийнято 100 балів, то відповідь, оцінена 200 балами, буде однозначно свідчити про достатній рівень знань студента. Аналогічно, відповідь, оцінена 10 балами, буде свідчити про те, що знання студента — незадовільні. Якщо ж за критерієм 100 балів, студент набирає 101 або 99 балів, то вихідні висновки про якість знань уже не є настільки очевидними.

Більшість схем тестового контролю не дозволяють усунути неоднозначність в оцінці знань, що виникає. Водночас послідовність дій, яка має бути виконана для її усунення, добре відома і широко застосовується під час контролю без тестів. У разі виникнення таких труднощів під час очного контролю викладач задає студенту додаткові запитання. З урахуванням відповіді на вже задані і додаткові запитання робиться остаточне судження про знання студента.

Щоб адаптувати тестовий контроль і звести його до аналогічної схеми в імітаційній моделі запропоновано встановити критеріальні значення, за недотримання яких необхідно приймати рішення про проведення додаткової сесії контролю. Для підвищення об'єктивності контролю критеріальні значення запропоновано розраховувати із залученням апарату математичної статистики [19].

У термінах теорії ймовірності дана задача формулюється так. За результатами визначення рівня знань, вираженого оцінками, вимірними в балах X_1, X_2, \dots, X_N , з нормальної сукупності оцінок з генеральною середньою \bar{X}_i дисперсії σ^2 необхідно визначити ймовірність того, що випадкова величина X , буде більше або менше деяких заданих граничних величин. Тоді нижня q_L і верхня q_U межі контрольованого параметра можуть бути виражені залежностями

$$q_L = \frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^L e^{-\frac{(x-\bar{X})^2}{2\sigma^2}} dx, \quad q_U = \frac{1}{\sqrt{2\pi\sigma}} \int_U^{\infty} e^{-\frac{(x-\bar{X})^2}{2\sigma^2}} dx. \quad (6)$$

Проводячи представницькі експерименти, достатні для достовірного визначення генеральних оцінок \bar{X} і σ^2 , можна розрахувати теоретично точні значення критеріальних величин q_L і q_U .

Висновок про досягнутий рівень засвоєння знань робиться на підставі такого. За відомих q_U і q_L приймальне число C прирівнюється до верхньої критеріальної межі q_U , а бракувальне число D — до нижньої критеріальної межі q_L . Якщо набране студентом число балів більше приймального числа C , то знання студента оцінюються як достатні для відповідного рівня успішності.

Якщо кількість балів менша ніж бракувальне число D , то робиться висновок, що студент знає контрольований матеріал нижче, ніж передбачено оцінкою даного рівня успішності.

При числі балів в інтервалі між приймальним і бракувальним числами приймається рішення про неможливість винести висновок про істинні знання студента, і проводиться додаткова сесія контролю. Якщо сума балів, набрана за підсумками основної й додаткової сесій контролю, більше приймального числа, то робиться висновок, що студент знає контрольований навчальний матеріал на відповідному рівні. В іншому випадку вказується на те, що знання студента недостатні. Тестовий контроль припиняється, і вважається, що знання студента оцінені. Якщо і в цьому випадку оцінювання неможливе, то проводяться додаткові сесії контролю.

Щоб забезпечити зіставленість результатів багаторівневого контролю в імітаційній моделі запропоновано кількість тестових завдань для першої сесії розраховувати, виходячи зі значущості контрольованого матеріалу [13]. Для другої і наступних сесій — з урахуванням допустимої складності. У першому випадку міра значущості приймається пропорційно до кількості годин, відведених на вивчення контрольованого навчального матеріалу. У другому випадку як міра складності встановлюється величина, зворотна нормованій сумі балів, набраних студентами на попередніх сесіях контролю. Можливий і варіант, коли обидві міри визначаються за результатами експертного оцінювання.

Попри це, оператор виживання в генетичному алгоритмі доповнено умовою, що ймовірність відбору тестового завдання для включення в додаткову сесію тим вища, чим менш успішно на попередніх сесіях контролю студент впорався із завданнями з тієї ж теми, що й нові завдання для додаткової сесії. Наразі виключаються варіанти, коли одним студентам будуть запропоновані завдання переважно за матеріалом тем, які вони добре знають, і за рахунок цього вони отримують оцінку, вищу дійсного рівня знань. А для інших студентів, навпаки, завдання будуть в, основному, з «важких» для них тем, і їх знання будуть недооцінені. Отже, під час проведення додаткових сесій усі студенти знаходяться в однаковому становищі — кожному з них пропонуються завдання тільки з тих тем, матеріал яких вони ймовірно знають гірше.

3.3. Підведення підсумків контролю

3.3.1. Аналіз достовірності статистичних відмінностей

Оцінка успішності є інтегруючим проявом великої кількості факторів, що діють в процесі навчання. Те, наскільки добре студенти складуть іспит, залежить від самих студентів, від викладачів, від методичного, організаційного забезпечення навчального процесу і т. д. Будь-які зміни в навчальному процесі, у тому числі і в процедурах контролю, можуть призвести до спотворення статистичної картини тестування.

За використання імітаційної моделі більшість рішень приймається на основі статистичного аналізу оцінок, отриманих студентами в тестуванні. Тому щоб адаптувати тестову систему до умов, у яких проводилися контрольні заходи, необхідно перед підведенням підсумків контролю оцінити статистичну істотність змін, внесених протягом часу, відведеного на вивчення навчального матеріалу. Зробити це можна за величиною коефіцієнта достовірності статистичних відмінностей [20], який стосовно достовірності відмінності в результатах тестування можна виразити залежністю

$$K_0 = \frac{|\bar{X} - \bar{X}^*|}{\sqrt{\frac{\Delta\text{sum} + \Delta\text{sum}^*}{n + n^*} \cdot \frac{n + n^*}{n \cdot n^*}}} \quad (7)$$

де \bar{X} , \bar{X}^* , Δsum , Δsum^* — статистичні характеристики, що розраховуються за результатами оцінювання в балах X_i , X_i^* , які були отримані всіма студентами під час виконання всіх завдань тесту до і після (*) внесення змін до навчального процесу. Достовірність відмінностей перевіряється за критерієм Стьюдента C_k . Якщо $K_0 < C_k$, то робиться висновок про відсутність статистично значущих змін у результатах тестування. Якщо $K_0 \geq C_k$, то вважається, що такий вплив встановлено і при прийнятті рішень слід враховувати дані тільки поточного контролю.

3.3.2. Оцінювання знань

Основним способом переведення набраних у тестовому контролі балів в оцінку успішності є порівняння суми балів з лінійною, рідше, нелінійною, шкалою оцінок. Однак подібний підхід тільки перетворить набрані бали в новий масштаб відповідно до поділок шкали оцінок, але не відобразить зусилля студентів, витрачені на досягнення зафіксованого рівня знань. Навіть той факт, що прийшов новий викладач або виданий новий підручник, може істотно вплинути на характер зусиль, які повинен докласти студент, щоб отримати хорошу оцінку. Тому застосування однієї і тієї ж усередненої шкали оцінок для різних вузів, факультетів або кафедр НЕ БУДЕ виправдано, якщо студенти навчалися в неоднакових умовах.

Щоб під час тестового контролю врахувати умови, у яких відбувалося формування знань студентів, в імітаційній моделі додатково вводиться ще одна оцінка. Вона використовується, щоб охарактеризувати здатність студентів досягти найкращих результатів у навчанні в однакових для всіх умовах.

У роботі [21] на основі висновку про те, що результати тестового контролю знань підкоряються нормальному закону розподілу, пропонується пов'язати шкалу оцінок з генеральними середнім і середньоквадратичним відхиленням (σ) оцінок, отриманих студентами у тестуванні. За аналогією з цим в імітаційній моделі приймається, що студенти, які набрали кількість балів більше $\bar{X} + 2\sigma$, тобто досягли кращих результатів, ніж більшість інших, заслуговують найвищої оцінки. І, дзеркально, успішність навчання студентів, що набрали в тестуванні суму балів, меншу $\bar{X} - 2\sigma$, тобто впоралися з тестом гірше інших, має бути визнана незадовільною. Інтервал від $\bar{X} + 2\sigma$ до $\bar{X} - 2\sigma$ (відповідно з властивостями нормального розподілу включає 95,45 % усіх можливих оцінок) ділиться на додаткові інтервали так, щоб шкала оцінок, що розраховується, відповідала національному стандарту навчання.

Можливі й інші варіанти побудови шкали переходу від набраної суми балів до показників успішності навчання. Однак загальний підхід має бути єдиним і ґрунтуватися на порівнянні результатів навчання, отриманих студентами в приблизно

однакових умовах. Студенти, які здатні досягти кращого рівня знань в тих же умовах, що й інші студенти, заслуговують більш високої оцінки.

3.4. Створення загальних і зрозумілих процедур підведення підсумків тестування

Алгоритмізація тестового контролю, математичне обґрунтування прийнятих рішень є в імітаційній моделі основою для адаптації процедур контролю до конкретних умов тестування і створення на цій основі загальних для всіх студентів умов тестування. При цьому важливо не тільки реалізувати закладені в імітаційній моделі процедури, а й довести до студентів сутність підходів, що використовуються, пояснити призначення і принципи прийняття рішень.

Тестовий контроль організовується викладачем, але він є не єдиним суб'єктом цього процесу, а знаходиться в постійній взаємодії зі студентом, кінцевим користувачем сформованої системи тестування. Від того, наскільки студент буде готовий об'єктивно сприйняти і правильно інтерпретувати свою участь у процесі контролю, залежать висновки, до яких він прийде, і чи будуть ці висновки стимулювати його до подальшого розвитку пізнавальної діяльності.

Студент проходить обмежене число процедур контролю, і багато в чому оцінює ефективність тестування тільки з особистого досвіду. Тому для формування повного уявлення про контроль рекомендується надати йому максимально докладні інструкції до виконання операцій контролю, у яких він бере участь. Крім цього, слід викласти для самостійного ознайомлення найбільш повний комплект матеріалів про сутність використовуваних процедур контролю.

Проілюструємо необхідність цього такими прикладами.

Імітаційною моделлю на завершальному етапі тестового контролю, у якому студент не бере участь, передбачений аналіз результатів, спрямований на бракування тестових завдань з низькими вимірювальними характеристиками [11]. Його проведення дозволяє виключити з бази тестових завдань ті, що не відповідають критеріям якості, у тому числі й завдання з неприпустимо малою статистичною ймовірністю правильних відповідей.

У результаті, студентам під час тестування пропонуються лише ті завдання, які відповідають критеріям якості, що в цілому сприяє підвищенню точності педагогічних вимірювань. Крім цього, студент, ознайомлений з умовами відбору тестових завдань, у разі невдалої відповіді буде усвідомлювати, що завдання можливо складне, але може бути розв'язане. Тому низька оцінка не викличе переконання, що в ній винна недосконалість тестової системи. Студент буде знати, що частина його колег успішно впоралися із завданням. Навпаки, замість негативного ставлення до тестування, як засобу об'єктивного контролю, буде сформоване переконання, що причина невдачі — у власній недостатній підготовці, і потрібні додаткові зусилля для подолання отриманого результату.

Найбільшу зацікавленість і критику з боку студентів, зазвичай, викликає те, як формується шкала оцінок. Неоднозначність особистісної участі викладачів у її побудові в більшості систем для тестування призводить до того, що навчальні досягнення студентів не отримують об'єктивного відображення, що знижує стимулюючий вплив оцінки на творчу освітню діяльність студента, на якість навчального процесу в цілому.

Навпаки, як показують спостереження, прийнятий в імітаційній моделі принцип формування шкали оцінок легко сприймається студентами, так як добре аргументований і підтверджений математичними викладками. Студенти позитивно ставляться до того, що критерії оцінювання формуються без чисі-небудь участі, об'єктивні, і відображають лише диференціацію зусиль кожного з них на тлі однакової

для всіх можливості досягти успіху. У імітаційній моделі побудова шкали оцінювання не є простим масштабуванням набраних за підсумками тестування балів, що відображують абсолютні досягнення. Тут абсолютні результати фіксуються, але в якості основного критерію успішності береться до уваги відносна характеристика, що відображає рівень, досягнутий студентом у зіставленні з іншими студентами. Позитивною мотивацією є те, що здібності кожного студента досягти кращих результатів оцінюються, виходячи із загальних для всіх умов (конкретний ВНЗ, форма навчання, викладачі, методичне забезпечення тощо).

3.5. Вимірювальні властивості

Для оцінювання впливу адаптивних можливостей імітаційної моделі на вимірювальні властивості тестового контролю було проведено експериментальне дослідження.

Усього в експерименті брало участь 228 студентів, з яких були сформовані експериментальна і контрольна групи. Формування груп виконувалося на основі поточної успішності студентів. Перевірка за критерієм Манна – Уїтні | Mann – Whitney показала, що групи однорідні.

Студентам експериментальної групи під час тестування надавалася тестова програма SSUquestionnaire, у якій були реалізовані всі запропоновані імітаційною моделлю нововведення [22]. У тестовій програмі, що використовувалася в контролі знань студентів контрольної групи, були відключені модулі, які реалізують процедури, що характеризуються адаптивними можливостями (табл. 1). Виняток склало те, що для формування тестів в експериментальній і контрольній групах використовувалася загальна база тестових завдань. Наразі для контрольної групи константи тестової програми, що визначають умови контролю на всіх етапах від формування тестів до підбиття підсумків контролю, призначалися екзаменаторами і не змінювалися в процесі контролю.

Результати тестування в експериментальній і контрольній групах математично оброблялися. На їх основі оцінювалася тіснота зв'язку між успішністю виконання окремих тестових завдань і підсумковими оцінками студентів, отриманими за всі завдання тесту. Для цього за результатами тестування в кожній із груп, незалежно одна від одної, розраховувався коефіцієнт кореляції Пірсона. Вважалося, що чим ближче абсолютна величина коефіцієнта кореляції Пірсона до одиниці, тим тісніше зв'язок і тим краще вимірювальні властивості відповідного тесту.

Порівняння отриманих даних показало, що в експериментальній групі спостерігається більш тісна лінійна залежність між результатами виконання окремих завдань і загальними підсумками тесту, ніж у контрольній. Коефіцієнт кореляції Пірсона в експериментальній групі щодо контрольної зріс з 0,46 до 0,72, що свідчить про кращі вимірювальні властивості тесту.

4. ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

Розроблене в рамках імітаційної моделі математичне обґрунтування прийнятих екзаменатором рішень значною мірою пом'якшує недоліки тестового контролю як інструменту педагогічних вимірювань. Використання у тестуванні процедур імітаційної моделі, що має адаптивні можливості, дозволяє повніше врахувати особливості процесу навчання і проведених контрольних заходів, знизити їх вплив на точність результатів контролю.

Формування тестових завдань з використанням пропонованих процедур, проведення багаторівневого комп'ютеризованого тестового контролю із застосуванням нечіткої логіки, нові підходи до вдосконалення адаптивних властивостей на етапі підведення підсумків контролю, дають можливість удосконалити систему контролю та підвищити за рахунок цього достовірність його результатів. Перспективним завданням дослідження є аналіз інших підходів до проектування адаптивних методів контролю і доведення їх до рівня практичного використання. Попри це, потребує подальшого дослідження питання про перерозподіл функцій контролю між викладачем і комп'ютером, уточнення місця і ролі комп'ютеризованого тестового контролю в загальній системі контролю знань.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Булах І. Є. Створюємо якісний тест: навчальний посібник / І. Є. Булах, М. Р. Мруга. — К.: Майстер-клас, 2006. — 160 с.
2. Олійник М. М. Тест як інструмент кількісної діагностики рівня знань в сучасних технологіях навчання / М. М. Олійник, Ю. А. Романенко. — Донецьк: Донецький національний університет, 2001. — 83 с.
3. Downing S. M. Handbook of test development / S. M. Downing, T. M. Haladyna. — Mahwah, N.Y.: L. Erlbaum, 2006. — 778 p.
4. Shultz K. S. Measurement theory in action: case studies and exercises / K. S. Shultz, D. J. Whitney, M. J. Zickar. — New York: Routledge, Taylor & Francis Group, 2014. — 396 p.
5. Williams D. D. Online assessment, measurement, and evaluation: emerging practices / D. D. Williams, M. Hricko, S. L. Howell. — Hershey, PA: Information Science Pub, 2006. — 343 p.
6. Ротаєнко П. А. Достовірність комп'ютерного тестування: результати дослідження / П. А. Ротаєнко // Комп'ютер у школі та сім'ї. — 2006. — № 3. — С. 7–11.
7. Колгатін О. Г. Педагогічна діагностика та інформаційно-комунікаційні технології / О. Г. Колгатін. — Х.: ХНПУ ім. Г. С. Сковороди, 2009. — 324 с.
8. Legg S. M. Computerized Adaptive Testing with Different Groups / S. M. Legg, D. C. Buhr // Educational Measurement: Issues and Practice. — 1992. — № 11(2). — P. 23–27.
9. Osterlind S. J. Differential item functioning / S. J. Osterlind, T. E. Howard. — Thousand Oaks, Calif.: SAGE, 2009. — 87 p.
10. Smoline D. V. Some problems of computer-aided testing and “interview-like tests” / D. V. Smoline // Computers & Education. — 2008. — Vol. 51. — No. 1. — P. 743–756.
11. Алексеев О. М. Імітаційна модель тестового контролю знань і умінь / О. М. Алексеев, Г. В. Алексеева // Комп'ютерно-орієнтовані системи навчання: збірник наукових праць. — К.: НПУ ім. М. П. Драгоманова, 2009. — Вип. 7(14). — С. 65–71.
12. IMS Global Learning Consortium [Електронний ресурс]. — Режим доступу: <http://www.imsglobal.org/question>. — Загл. с екрана.
13. Алексеев А. Н. Дистанционное обучение инженерным специальностям / А. Н. Алексеев. — Сумы: Университетская книга, 2005. — 333 с.
14. Hontangas P. The choice of item difficulty in self adapted testing / [S. Hontangas, V. Ponsoda, J. Olea, S. L. Wise] // European Journal of Psychological Assessment. — 2000. — № 16(1). — С. 3–12.
15. Goldberg D. E. Genetic algorithms in search, optimization, and machine learning / D. E. Goldberg. — Reading, Mass.: Addison-Wesley, 1989. — 412 p.
16. Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы / Д. Рутковская, М. Пилиньский, Л. Рутковский. — М.: Горячая линия-Телеком, 2006. — 452 с.
17. Gerla G. Fuzzy logic: mathematical tools for approximate reasoning / G. Gerla. — Dordrecht, Boston, London: Kluwer Academic Publishers, 2001. — 269 с.
18. Volkov N. I. Creation of a library of E-manuals for students of mechanical engineering with tests based on fuzzy logic / N. I. Volkov, A. N. Alexeyev, A. N. Kochevsky // Bulletin of the Petroleum and Gas University of Ploiesti. — Ploiesti, 2004. — № 1. — Volume LVI. — P. 42–46.
19. Alexeyev A. N. Multi-level test control of quality of knowledge by quantitative parameters / A. N. Alexeyev // Computer modeling and new technologies. — Riga, 2007. — № 3. — Volume 11. — P. 43–45.
20. Урбах В. Ю. Статистический анализ в биологических и медицинских исследованиях / В. Ю. Урбах. — М.: Медицина, 1975. — 295 с.

21. TenBrink T. D. An educator's guide to classroom assessment / T. D. TenBrink. — Boston : Houghton Mifflin, 2003. — 55 p.
22. Новые возможности тестового контроля знаний с использованием программы SSUquestionnaire версии 4.10 [Электронный ресурс]. — Режим доступа : <http://test.sumdu.edu.ua>. — Загл. с экрана.

Материал поступил в редакцию 16.04.2014 г.

АДАПТИВНЫЕ ВОЗМОЖНОСТИ ИМИТАЦИОННОЙ МОДЕЛИ ТЕСТОВОГО КОНТРОЛЯ ЗНАНИЙ

Алексеев Александр Николаевич

доцент, доктор педагогических наук, профессор кафедры «Технологии машиностроения, станков и инструментов»

Сумский государственный университет, г. Сумы, Украина

alekseev@tmvi.sumdu.edu.ua

Король Елена Николаевна

аспирантка

Сумский государственный педагогический университет им. А. С. Макаренко, г. Сумы, Украина

korol.9@mail.ru

Аннотация. Рассмотрены особенности тестового контроля знаний с использованием имитационной модели. Имитационная модель базируется на современных информационно-коммуникационных технологиях и при этом максимально опирается на традиции активного участия преподавателя в диагностировании успешности обучения студентов. Сочетание преимуществ компьютеризированного тестирования с математическим обоснованием процедур принятия решений экзаменатором расширяет область эффективного применения тестовых форм контроля. Отмечено, что применение имитационной модели позволяет снизить влияние условий, в которых проходило обучение студентов, на достоверность результатов контроля. Описаны процедуры имитационной модели, обладающие адаптивными возможностями.

Ключевые слова: тестирование; адаптивный тестовый контроль; формирование теста; имитационная модель; многоуровневый контроль; оценивание знаний.

ADAPTIVE CAPABILITIES OF KNOWLEDGE TESTING SIMULATION MODEL

Oleksandr M. Alekseev

Associate Professor, Ed.D., Professor of Manufacturing Engineering, Machines and Tools Chair

Sumy State University, Sumy, Ukraine

alekseev@tmvi.sumdu.edu.ua

Olena M. Korol

Ph.D. student

A.S. Makarenko Sumy State Pedagogical University, Sumy, Ukraine

korol.9@mail.ru

Abstract. This article describes peculiarities of knowledge test control using a simulation model. The simulation model is based on modern information and communication technologies. At the same time, it relies as much as possible on the traditions of active participation of an examiner in the assessment of students' learning success. Combination of advantages of computerized testing with mathematical foundations of examiner's decision-making procedures broadens the field in which test control can be used effectively. The article describes procedures of the simulation model that have adaptive capabilities. It is noted that use of the simulation model reduced the impact of students' learning environment on the accuracy of testing results.

Key words: testing; adaptive test control; test formation; simulation model; multilevel control; knowledge assessment.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1. Bulakh I. E. Creating quality test : textbook / I. E. Bulakh, M. P. Mruha. — K. : Master-class, 2006. — 160 p. (in Ukrainian)
2. Oliinyk M. M. Test as an instrument for qualitative diagnostics of the level of knowledge in today's education technologies / M. M. Oliinyk, Y. A. Romanenko. — Donetsk : Donetsk National University, 2001. — 83 p. (in Ukrainian)
3. Downing S. M. Handbook of test development / S. M. Downing, T. M. Haladyna. — Mahwah, N.Y. : L. Erlbaum, 2006. — 778 p. (in English)
4. Shultz K. S. Measurement theory in action: case studies and exercises / K. S. Shultz, D. J. Whitney, M. J. Zickar. — New York : Routledge, Taylor & Francis Group, 2014. — 396 p. (in English)
5. Williams D. D. Online assessment, measurement, and evaluation : emerging practices / D. D. Williams, M. Hricko, S. L. Howell. — Hershey, PA : Information Science Pub, 2006. — 343 p. (in English)
6. Rotaienko P. A. Reliability of computerized testing: study results / P. A. Rotaienko // Computer in school and in family. — 2006. — №3. — P. 7–11. (in Ukrainian)
7. Kolhatin O. H. Pedagogical Diagnostics and information-communication technologies / O. H. Kolhatin. — X. : HNPУ named after H.S. Skovoroda, 2009. — 324 p. (in Ukrainian)
8. Legg S. M. Computerized Adaptive Testing with Different Groups / S. M. Legg, D. C. Buhr // Educational Measurement : Issues and Practice. — 1992. — № 11(2) . — P. 23-27. (in English)
9. Osterlind S. J. Differential item functioning / S. J. Osterlind, T. E. Howard. — Thousand Oaks, Calif. : SAGE, 2009. — 87 p. (in English)
10. Smoline D. V. Some problems of computer-aided testing and “interview-like tests” / D. V. Smoline // Computers & Education. — 2008. — Vol. 51, No. 1. — P. 743-756. (in English)
11. Alexeyev A. N. A simulation model of the test control of knowledge and skills / A. N. Alexeyev, G. V. Alexeyeva // Computer-oriented educational system . — K. : NPU. M. Dragomanov, 2009. — № 7(14) . — P. 65-71. (in Ukrainian)
12. IMS Global Learning Consortium [online]. — Available from : <http://www.imsglobal.org/question>. (in English)
13. Alexeyev A. N. Distance Engineering Learning / A. N. Alexeyev . — Sumy : University Book, 2005. — 333 p. (in Russian)
14. Hontangas P. The choice of item difficulty in self adapted testing / S. Hontangas, V. Ponsoda, J. Olea, S. L. Wise // European Journal of Psychological Assessment. — 2000. — № 16(1) . — p. 3-12. (in English)
15. Goldberg D. E. Genetic algorithms in search, optimization, and machine learning / D. E. Goldberg. — Reading, Mass. : Addison-Wesley, 1989 . — 412 p. (in English)
16. Rutkovskaya D. Neural Networks, Genetic Algorithms and Fuzzy Systems / D. Rutkovskaya, M. Pilinskiy, L. Rutkovskiy. — M. : Hotline Telecom, 2006. — 452 p. (in Russian)
17. Gerla G. Fuzzy logic: mathematical tools for approximate reasoning / G. Gerla. — Dordrecht, Boston, London : Kluwer Academic Publishers, 2001. — 269 p. (in English)
18. Volkov N. I. Creation of a library of E-manuals for students of mechanical engineering with tests based on fuzzy logic / N. I. Volkov, A. N. Alexeyev, A. N. Kochevsky // Bulletin of the Petroleum and Gas University of Ploiesti. — Ploiesti, 2004. — № 1. — Volume LVI. — P. 42-46. (in English)
19. Alexeyev A. N. Multi-level test control of quality of knowledge by quantitative parameters / A. N. Alexeyev // Computer modeling and new technologies. — Riga, 2007. — № 3. — Volume 11. — P. 43-45. (in English)
20. Urbach V. Y. Statistical Analysis in Biological and Medical Research / V. Y. Urbach. — M. : Medicine, 1975. — 295 p. (in Russian)
21. TenBrink T. D. An educator's guide to classroom assessment / T. D. TenBrink. — Boston : Houghton Mifflin, 2003. — 55 p. (in English)
22. New capabilities of knowledge test control utilizing SSUquestionnaire version 4.10 [online]. — Available from : <http://test.sumdu.edu.ua>. (in Russian)