

ВИКОРИСТАННЯ КОМП'ЮТЕРНИХ ТЕХНОЛОГІЙ ДЛЯ ЗАБЕЗПЕЧЕННЯ ДОСТОВІРНИХ РЕЗУЛЬТАТІВ ПЕДАГОГІЧНИХ ВИМІРЮВАНЬ

О. С. Ільків, В. І. Матвійв, О. О. Пришляк

Львівський державний університет фізичної культури

Аногація. В статті розглянуто проблеми оперативного виявлення результатів навчання через проведення різнопланових педагогічних вимірювань та методики отримання достовірних результатів педагогічних вимірювань, способи контролю за динамікою навчального процесу, побудову індивідуальних стратегій його продовження та прогнозування результатів. На емпіричних даних, отриманих при впровадженні тестових технологій у навчальний процес університету, обґрунтовано методику оцінювання валідності та надійності результатів педагогічних вимірювань.

Ключові слова: педагогічні вимірювання, тести, комп'ютерні технології.

Постановка проблеми. Формування сучасного фахівця – процес складний та багатогранний. Забезпечення високої якості навчання потребує постійної корекції навчальних досягнень. Оперативне виявлення результатів навчання отримують через проведення різнопланових педагогічних вимірювань. Поряд із традиційними формами контролю знань вищі навчальні заклади запроваджують використання тестових методик визначення рівня досягнутих результатів навчання. Використання тестів є науково обґрунтованим інструментом оцінювання навчальних досягнень. Широко висвітлена у педагогічній літературі і методика отримання достовірних результатів педагогічних вимірювань. Впровадження комп'ютерних технологій у навчально-виховний процес та адаптація їх до проведення педагогічних вимірювань створює проблему відповідності отриманих результатів.

Дослідження виконано в рамках загальноуніверситетської комплексної наукової теми.

Мета дослідження полягає в тому, щоб, посилаючись на сучасний досвід педагогічної діяльності у сфері проведення педагогічних вимірювань із застосуванням комп'ютерних технологій, сформулювати методику підтвердження відповідності отриманих результатів до рівня досягнутих навчальних результатів.

Для досягнення визначеної мети передбачалося вирішення таких основних завдань:

- 1) провести теоретичний аналіз та узагальнити досвід світової та української практики перевірки якості засвоєння вивченого матеріалу;
- 2) вивчити специфіку комп'ютерно-орієнтованих методик педагогічних вимірювань;
- 3) охарактеризувати рівень розвитку сучасної педагогічної тестології.

Організація та методи дослідження. Програма дослідження становить науково обґрунтовану систему теоретико-методичних положень про особливості проведення педагогічних вимірювань із застосуванням комп'ютерних технологій та забезпечення достовірних результатів педагогічних вимірювань. Для вирішення поставлених завдань були використані такі методи:

- 1) теоретичний аналіз розробки цієї проблеми в науково-методичній літературі;
- 2) метод системно-структурного аналізу і синтезу;
- 3) методи математичної статистики.

Аналіз останніх досліджень і публікацій. Процес навчання нерозривно пов'язаний із постійною перевіркою якості засвоєння вивченого матеріалу. Орієнтація сучасних навчальних досягнень на використання комп'ютерних технологій створює умови для розвитку нового напрямку педагогічної діяльності – індивідуалізоване навчання. Ідеологічною основою цього напрямку є ідеї особисто центрованої освіти американського філософа Дж. Дьюї, в центрі якої

він поставив особистість (1). Педагогічне обґрунтування індивідуалізованого навчання є в працях Песталоцці, Дж. Керрола, Б.Блумба та інших. Завдання підготовки до індивідуалізованого навчання полягає у створенні психолого-педагогічних умов для повного засвоєння навчального матеріалу. Н. Ф. Талізін окреслює п'ять рівнів засвоєння навчального матеріалу (2). Згідно із такою організацією навчального процесу, перехід на наступний рівень відбувається за результатами педагогічного вимірювання результатів навчання на попередньому. За основу педагогічних вимірювань прийнято класичні роботи Лорда зі статистичної теорії (3). Вчені Росії та Західної Європи приймають за теоретичну основу педагогічних вимірювань математичну теорію вимірювання латентних якостей особистості. Математичний апарат цієї теорії застосовують для проведення вимірювань у політології, психології, соціології. Проте використання на сучасному етапі широкомасштабних вимірювань із застосуванням комп'ютерних технологій вимагає змін і доповнень у концепції моделі педагогічних вимірювань. За теоретичну основу побудови моделі сучасних педагогічних вимірювань можна, на нашу думку, прийняти дослідження В. С. Аванесова(4).

Виклад основного матеріалу. Однією із комп'ютерно-орієнтованих методик педагогічних вимірювань є застосування тестів. Результати комп'ютерного тестування надзвичайно інформативні і дозволяють контролювати динаміку навчального процесу, будувати індивідуальні стратегії його продовження, прогнозувати результати. Вони можуть використовуватися як статистична основа для проведення аналітичних досліджень і вивчення сучасних методик формування спеціаліста в умовах вищого навчального закладу.

Під час підготовки і проведення тесту ми виходимо з того, що його результати повинні відображати рівень умінь застосовувати знання, давати об'єктивну оцінку ступеня оволодіння навичками, вміннями та видами діяльності практичного характеру. Проте „головним аргументом проти того, щоб вважати результати тестів надійним показником якості освіти, є те, що при акцентуванні уваги на підсумкових результатах, не беруться до уваги інші фактори, які впливають на успішність або неуспішність...”(5) Множина неврахованих чинників педагогічного вимірювання має соціально-економічний, мотиваційний та кадровий характер, що значно нівелюється в рамках одного навчального закладу. Використання комп'ютерних технологій під час проведення тестування виключає суб'єктивну складову оцінки, яка завжди присутня у проведенні оцінювання знань викладачем. Комп'ютерне тестування передбачає проведення об'єктивного кількісного зіставлення певної властивості студента із деяким сформованим еталоном. У такій постановці проблеми „педагогічне вимірювання – це процес виявлення відповідності між характеристиками студентів, які оцінюються, і точками емпіричної шкали, в якій відношення між різними оцінками характеристик виражені властивостями числового ряду”(6). Найбільшу надійність, як показує педагогічна практика, мають адаптивні тестові технології. Адаптивні тести, як правило, розпочинають із завдань середнього рівня і коригують відповідно до результатів відповідей на кожному окремому етапі. Використання таких методик можливе тільки при використанні спеціалізованих середовищ із досконало організованим аналітичним блоком оперативного відслідковування результатів. Системний модульний контроль за якістю підготовки студентів із дисципліни „Інформатика і комп'ютерна техніка” у нашому університеті ведеться із використанням програмного комплексу Module. Він призначений для організації дистанційного навчання. Результати тестування, проведеного із використанням згаданого комплексу можна отримати у вигляді узагальненого документа у форматі Excel. На жаль, цих даних недостатньо для проведення поглиблених педагогічних досліджень запропонованого тесту. Провівши дослідження програмного комплексу Module, ми прийшли до висновку, що первинні дані, які використовуються програмою для оцінювання результатів можна самостійно досліджувати, попередньо імпортувавши їх із бази даних у програму Excel. Нескладний запит до бази SQL, на основі якої працює програма, дозволить отримати детальнішу інформацію про відповіді кожного студента.

Недоліком сучасної педагогічної тестології є відсутність єдиної класифікації основних категорій вимірювання знань. Низька дослідженість основних характеристик педагогічного

тестування як однієї із категорії вимірювання знань призупиняє їхнє впровадження у освітню практику. Серед основних категорій вимірювання знань педагогічна література виділяє:

- метод вимірювання;
- інструментарій вимірювання;
- процедура вимірювання;
- процедура оцінювання.

Достовірні результати педагогічного вимірювання можна отримати тільки при високій валідності цих категорій. Серед багатьох визначень валідності у педагогічній літературі зупинимося на одному, яке дає А. Анастасі: "поняття, яке вказує, що вимірює тест і наскільки добре він це робить"(7).

Поняття валідності за своїм змістом складне та багатоаспектне. Окремі автори виділяють валідність – методу, інструментарію, процедури тестування та процедури оцінювання.

І. Є. Булах (8) пропонує аналізувати валідність тестових завдань, посилаючись на результати статистичної обробки даних тестування. Комплексний аналіз результатів обробки та їх інтерпретація дає можливість оцінити якість запропонованих тестів.

Для проведення статистичного аналізу даних, які отримані в результаті використання програми Module, формуємо матриці тестових результатів. До матриці вводимо тільки результати правильних відповідей, які позначаємо 1. Допускаємо також, що результати вимірювання, які проводилися в однакових умовах, збігаються. Для підготовки статистичних даних часто використовують метод повторного тестування, проведення паралельного тестування та зведення результатів тестування навпіл. Готуючи статистичну вибірку ми використали технологію поділу результатів. Згідно з таким підходом,

до першої вибірки X_n^1 зараховуємо відповіді студентів на питання з парним номером, до другої X_n^2 - відповіді на непарні питання.

Емпіричні дані вимірювань для N студентів, що пройшли тестування, заносимо у двовимірну таблицю програми Excel. Використовуючи команду *Сервіс / Аналіз даних / Кореляції*, отримуємо кореляційну матрицю для чинників X_n^1 та X_n^2 .

	Парні	Непарні
1	1	1
2	1	1
3	1	1
4	1	1
5	1	1
6	1	1
7	1	1
8	1	1
9	1	1
10	1	1
11	1	1
12	1	1
13	1	1
14	1	1
15	1	1
16	1	1
17	1	1
18	1	1
19	1	1
20	1	1
21	1	1
22	1	1

Коефіцієнт кореляції = 0,9607

Рис. 1. Результати обчислення коефіцієнта кореляції для парних та непарних завдань

За результатами аналізу матриці визначаємо характер зв'язку між незалежними чинниками. При значенні коефіцієнта кореляції близькому до 1 результати можна вважати сприятливими для подальшого опрацювання.

Графічна інтерпретація емпіричних результатів у вигляді гістограми або згладженої кривої дає додаткову інформацію про нормативно-орієнтований тест. Як показує практика, такі тести забезпечують групування до 80 % вибірки у центральній частині розподілу. Якщо ж тест виконали майже всі студенти або майже ніхто не виконав, то такий тест

відбракується і підлягає переробці, оскільки він визнається недосконалим, або тестування проведено із грубими порушеннями процедури вимірювання.

Педагогічна література дає декілька методик для проведення аналізу результату тестування. Запроваджена в університеті методика передбачає, що на першому етапі оцінюються міри центральної тенденції сукупності. Для цього виявляємо моду та середнє арифметичне розподілу

$$\left(\bar{X} = \frac{\sum_{i=1}^N X_i}{N}\right).$$

Міри центральної тенденції допомагають оцінити якість тесту та репрезентативність вибірки. Високо валідний тест забезпечує нормальний розподіл отриманих студентами балів, при якому до 70 % емпіричних результатів знаходяться у центральній частині. Якщо мода і середнє арифметичне суттєво відрізняються, тест вважають неправильно сконструйованим, він не може бути визнаним для використання в педагогічних вимірюваннях. Несприйнятливим для використання є тест із бімодальною конфігурацією розподілу результатів. Такий результат свідчить про те, що в тесті використано дві групи питань – легкі та складні, що і привело до розподілу студентів на дві підгрупи.

Останнім етапом кількісного аналізу є визначення асиметрії та ексцесу кривої розподілу емпіричних даних.

Формула для обчислення асиметрії у нашому випадку має такий вигляд:

$$A_s = \frac{\sum_{i=1}^N (X_i - \bar{X})^3}{S_x^3 * N},$$

де A_s – асиметрія, X_i – елементи вибірки, \bar{X} – середнє арифметичне вибірки,

S_x – стандартне відхилення, що дорівнює кореню квадратному з дисперсії,

N – розмір вибірки.

Із використанням вказаних позначень ексцес обчислюємо за формулою:

$$E_s = \frac{\sum_{i=1}^N (X_i - \bar{X})^4}{S_x^4 * N}.$$

Обчисливши значення асиметрії та ексцесу, можна зробити висновок про валідність тесту та його недоліки.

Тест вважається валідним при $A_s=0$ та $E_s=0$. Практика експериментальних досліджень великої кількості тестів, які використовуються в університеті показує, що такі параметри має невелика кількість із них.

Інтерпретацію інших комбінацій A_s та E_s виконуємо за методикою, яку запропонували Л. І. Паращенко, В. Д. Леонський, Г. І. Леонська (9) Так? при $E_s < 0$ тест характеризується високою розподільчою здатністю, а при $E_s > 0$ – низькою. Валідний тест, за нашим переконанням, повинен мати високу розподільну здатність, що забезпечує достатню вибіркочість час проведення педагогічних вимірювань.

Показник асиметрії вибірки свідчить про складність тесту, при $A_s > 0$ тест вважається надійним, при значеннях $A_s < 0$ тест легкий. Наведені на рис. 2 результати аналізу тесту дають можливість зробити висновок, що запропонований викладачем тест виявився легким для студентів ($A_s = -0.90$) і має малу розподільну здатність ($E_s = 4,038$).

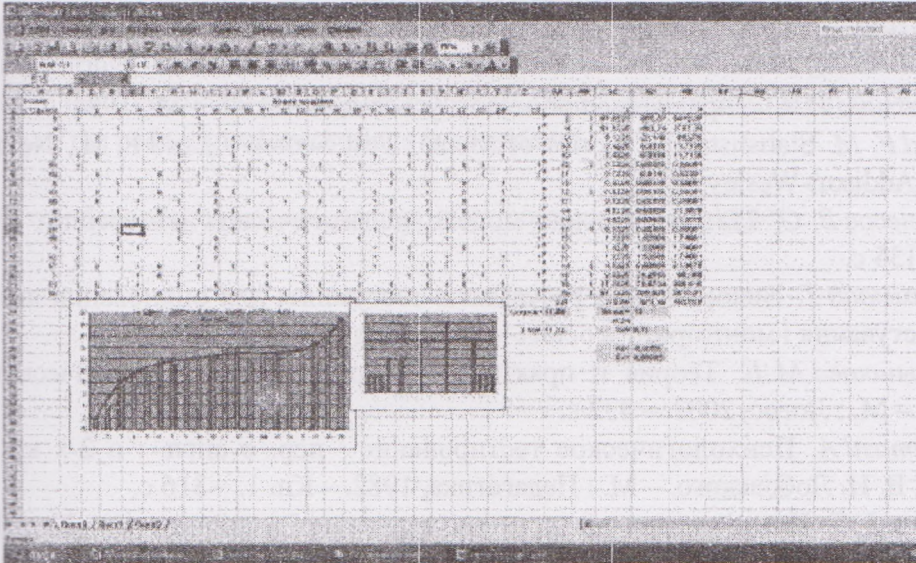


Рис. 2. Результати дослідження тесту на валідність

Близькість середнього значення і моди дають підставу рекомендувати викладачеві додати цей тест, зберігши його конфігурацію та спрямованість. Для визначення причин неважкості валідності необхідно також проаналізувати педагогічні умови проведення тестів та спрямованість тестування.

Завершуємо аналіз результатів тестування визначенням показника надійності тесту за допомогою прогностичної формули Спірмена-Брауна. Параметром функції для обчислення надійності є коефіцієнт кореляції.

Формула Спірмена-Брауна для цього випадку має такий вигляд:

$$C = \frac{2 * r}{1 + r},$$

де C – коефіцієнт надійності, r – коефіцієнт кореляції, розрахований для даних, отриманих за допомогою поділу навпіл. За нашими обчисленнями $r = 0,98$, отже,

$$C = \frac{2 * 0,98}{1 + 0,98} = 0,98.$$

Цей коефіцієнт характеризує послідовність вимірювань і дає відповідь на низку важливих питань. Використаний тест виявився достатньо надійним, це означає, що і при іншій серії вимірювань, ми отримали б такі ж результати.

Надійні тести із різними показниками асиметрії можуть використовуватися в моделі адаптивних вимірювань. При отриманні низьких результатів повторне тестування проводимо використанням легших тестів.

Висновок

Проблема забезпечення достовірних результатів педагогічних вимірювань при використанні тестових завдань може бути реалізована через використання надійних методик статистичного аналізу отриманих результатів на етапі апробації та впровадження, що і буде предметом подальших досліджень. На підставі достовірних результатів аналітичного дослідження результатів тестування можна буде зробити висновок про шляхи удосконалення педагогічних вимірювань у навчальному закладі.

Список літератури

1. Якиманская И. С. Личностно-ориентованное обучение в современной школе. – М. : 1996.
2. Талызина И. С. Управление процессом усвоения знаний. – М., 1984.
3. Lord F. M. Statistical Theories of Mental Test Scores. / Lord F. M., Novick M. R. – Reading, MA : Addison-Wesley, 1968.
4. Аванесов В. С. Математические модели педагогических измерений. – М. : Исс. Центр, 1994. – 135 с.
5. Драммонд Т. Використання результатів тестування для оцінювання якості освіти за і проти // Тестування і моніторинг в освіті. – 2006. – № 6. – С. 31-34.
6. Чельщикова М. Б. Теория и практика конструирования педагогических тестов : учеб. пособие. – М. : Логос, 2002. – 432 с.
7. Анатази А. Психологическое тестирование / пер. с англ. : в 2-х кн.; под ред. К. М. Гуревич, В. И. Лобовского. – М. : Педагогика, 1982. – Кн. 1. – 316 с.
8. Булах І. Є. Комп'ютерна діагностика навчальної успішності. – К. : ЦМК МОЗ України, УДМУ. – 1995. – 221 с.
9. Паращенко Л. І. Тестові технології у навчальному закладі : метод. посіб. / Паращенко Л. І, Леонський В. Д., Леонська Г. І.; наук. ред. О. І. Ляшенко – К. : Майстерня книги, 2006. – 217 с.

ИСПОЛЬЗОВАНИЕ КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ ДЛЯ ОБЕСПЕЧЕНИЯ ДОСТОВЕРНЫХ РЕЗУЛЬТАТОВ ПЕДАГОГИЧЕСКИХ ИЗМЕРЕНИЙ

Ільків О. С., Матвиїв В. І., Пришляк О. О.

Львовский государственный университет физической культуры

Аннотация. Использование компьютерных технологий для обеспечения достоверных результатов педагогических измерений. В статье рассмотрены проблемы оперативного выявления результатов обучения путём проведения разноплановых педагогических измерений и методики получения достоверных результатов педагогических измерений, способы контроля за динамикой учебного процесса, построение индивидуальных стратегий его продолжения и прогнозирования результатов. На эмпирических данных, полученных при внедрении тестовых технологий в учебный процесс университета, рассмотрено методику оценки валидности и надежности результатов педагогических измерений.

Ключевые слова: педагогические измерения, тесты, компьютерные технологии.

THE USE OF COMPUTER TECHNOLOGIES FOR PROVIDING OF RELIABLE RESULTS OF PEDAGOGICAL MEASURINGS

Il'kiv O. S., Matviiv V. I., Pryshliak O. O.

Lviv State University of Physical Culture

Annotation. To the problem of providing the meaningful results of the pedagogical measurings with the use of computer technologies. The problems of receipt of meaningful results of the pedagogical measurings are considered in the article. On empiric information, got at introduction of test technologies in the educational process of university, the method of estimation of validiti and reliabi-ty of results of the pedagogical measurings is considered.

Key words: pedagogical measurings, tests, computer technologies.