УДК 378.091.26

Лузан П. Г.
доктор педагогічних наук, професор,
головний науковий співробітник лабораторії науково-методичного
супроводу підготовки фахівців у коледжах і технікумах
Інституту професійної освіти НАПН України,
Київ, Україна

КРИТЕРІЇ ЯКОСТІ ТЕСТОВОГО ОЦІНЮВАННЯ РЕЗУЛЬТАТІВ НАВЧАННЯ СТУДЕНТІВ

Суттєве підвищення якості підготовки фахівців у коледжах наразі є основним, пріоритетним напрямом подолання фаховою передвищою освітою тих викликів, які виникли внаслідок збройної агресії РФ проти України. Це зумовлено посиленими суспільства до кваліфікації випускників коледжів, які без тривалої адаптації, вже з перших днів роботи за фахом мають компетентно розв'язувати виробничо-технологічні ситуації, демонструвати уміння самостійно опановувати новітні виробничі технології, успішно працювати в команді.

Доведено, що об'єктивне оцінювання навчальних досягнень студентів є домінантною умовою забезпечення якості підготовки майбутніх техніків, технологів, менеджерів тощо. Аналіз досвіду розвинених країн світу засвідчує, що вкрай важливу роль в оцінюванні якості освіти відіграють об'єктивні методики, зокрема педагогічне тестування. Це й не дивно, тому що у порівнянні з узвичаєними емпіричними способами вимірювання результатів навчання тестування має такі переваги: при дидактично правильній організації тестування можна перевірити значно більший обсяг засвоєного навчального матеріалу студентами у порівнянні з традиційними методиками контролю; тестування дозволяє перевірити результати навчання на певному рівні засвоєння: у структурі тесту для цього застосовують завдання різного ступеня складності, що дозволяє не інтуїтивно, а на науковій основі визначити рівень опанування студентом запрограмованим змістом; тестування не має усталених недоліків (суб'єктивність оцінювання; неможливість відновлення результатів перевірки; порівняно малий обсяг перевірки змісту навчального матеріалу; різна кількість і складність додаткових запитань), притаманним традиційним методам контролю знань; тестовий контроль позитивно відрізняється від усних і письмових випробувань тим, що він є спеціально підготовленим контрольним набором завдань, які дозволяють кількісно оцінити результати навчання студентів за допомогою статистичних методів; тестування є більш точнішим, але й більш «м'якшим» інструментом вимірювання результатів навчання, так як воно ставить усіх студентів у рівні умови, використовуючи єдину процедуру і єдині критерії оцінювання та ін.

У вітчизняній освітній практиці активізацію широкого запровадження тестування як методу об'єктивного оцінювання навчальних досягнень студентів пов'язують із наказом Міністерства освіти України від 31.07.1998 р. № 285 «Про порядок розробки складових нормативного та навчально-методичного забезпечення підготовки фахівців з вищою освітою», у якому наголошувалося: *«Серед засобів об'єктивного контролю, найбільш науково обґрунтованим є метод тестування із залученням технічних засобів».*

Натомість помилкові рекомендації огульного застосування тестів «*при всіх видах педагогічного контролю»* без визначення валідності та надійності цього інструментарію, без опанування викладачами елементарними основами тестування, принципів створення

якісних тестів фактично стали причиною «*дискредитації самого метода вимірювання*» (І. Булах). Як наслідок, освітня практика напрацювала за ці роки масу попередньо не перевірених на валідність тестів, які, переважно, складаються із альтернативних і одновибіркових завдань. Природно, такий інструментарій здатний виміряти найнижчі рівні володіння навчальним матеріалом (наприклад, на рівні «впізнавання» за В. Беспалько, чи на рівні «знання» за Б. Блумом), проте не може надійно і точно диференціювати студентів за їх досягненнями.

Відтак, вимірювання рівнів навчальних досягнень тих, хто навчається має ґрунтуватися на критеріях, які визначають відповідність інструментів і процедур оцінювання встановленим вимогам. При аналізі якості тесту будемо послуговуватися більш широким поняттям «педагогічне тестування», що включає і тест як інструмент вимірювання, і саму процедуру вимірювання та оцінювання навчальних досягнень студентів.

До критеріїв оцінки якості педагогічного тестування відносять такі найважливіші характеристики: *об'єктивність; валідність; надійність; точність*. Схарактеризуємо зміст кожного з цих критеріїв.

Об'єктивність як критерій оцінювання якості тестування показує, на скільки мінімізовано вплив суб'єктивних факторів. Педагогічне тестування об'єктивне тоді, коли його результати максимально незалежні від педагогічних працівників. Це означає, що різні дослідники, застосовуючи один і той самий тест, контролюючи одну й ту ж вибірку студентів, мають одержувати однаковий результат. Така якість тесту досягається шляхом стандартизації умов проведення вимірювання: регламентованими процедурами складання тесту і відбору валідних тестових завдань; уніфікованими процедурами проведення вимірювання та оцінювання тестових результатів, статистичними методами обробки та інтерпретації отриманих даних.

Найбільш важливою, і водночас, складною характеристикою педагогічного тестування є *валідність*. Дослідники наводять декілька визначень цієї характеристики: «Валідність тесту – це поняття, яке вказує нам, що тест вимірює і наскільки добре він це робить» (А. Анастазі); «Валідність означає придатність тестових результатів для тієї мети, заради чого проводилося тестування» (В. Аванесов); «Валідність – це комплексна характеристика, що визначається як параметрами засобу та процедури вимірювання, так і властивостями досліджуваної ознаки» (І. Булах).

Отже, *валідність педагогічного тестування* – це комплексна характеристика, що визначається як параметрами педагогічного тесту та процедури оцінювання, так і властивостями досліджуваних освітніх результатів студентів. Як комплексна, інтегративна характеристика валідність педагогічного тестування визначається такими функціональними складниками (І. Булах, Л. Кухар, В. Сергієнко): валідність методу (валідність змісту; валідність відповідності; валідність прогнозу); валідність тесту (валідність тестових завдань; валідність процедури тестування; валідність процедури оцінювання).

*Валідність метод*у – це відповідність того, що вимірюється засобами педагогічного тестування, тому, що він має вимірювати. Іншими словами, цей критерій валідності встановлює, чи цей спосіб вимірювання дає статистично значущі результати при його застосуванні для оцінювання рівнів навчальних досягнень тих, хто навчається. Наприклад, якщо за мету тестування визначено вимірювання освітніх результатів рівня «*застосування*» (за таксономією Б. Блума), то і тестові завдання мають відповідати такому рівню навчальних досягнень студентів. Мовиться про те, що такі тестові завдання не можна правильно виконати засобами пізнавальної діяльності більш низького рівня («*знання (відтворення*)» чи «*розуміння*»). При цьому говорять про *функціональну валідність педагогічного тесту*.

При вимірюванні навчальних досягнень студентів засобами педагогічного тестування валідність методу можна розподілити на якісно відмінні критерії: валідність змісту, валідність відповідності і валідність прогнозу. *Валідність змісту* – це критерій, який визначає паритет між еталонними результатами навчання (освітні результати, задекларовані в освітніх стандартах, освітньо-професійних програмах тощо) і тим змістом, який перевіряється засобами тестових завдань. В ідеальному варіанті зміст тестових завдань відображає вказані нормативні вимоги і, разом з тим, узгоджується з тим змістом навчання, який був запропонований студентам для опанування. *Валідність відповідності* – це характеристика, яка вказує на відповідність, зіставленість результатів вимірювання, здобутих тестуванням та іншими методами. Валідність відповідності можна визначити через коефіцієнт кореляції між результатами тестування та результатами інших вимірювань, проведених у тій самій групі студентів щодо опанування ними того ж самого матеріалу. *Прогностична валідність* – це ступінь того, наскільки отримані результати тестових випробувань прогнозують успішність подальшого навчання. Прогностичність тесту – це інформація, що отримана при аналізі результатів виконання тесту, яка має містити достовірні дані задля визначення змісту та прогнозування результатів подальших педагогічних дій.

Валідність тестового завдання – це критерій, що характеризується складністю та розподільчою здатністю тестового завдання. Складність тестового завдання є його характеристикою, яка визначає, наскільки конкретне тестове завдання було посильним для осіб, які тестувалися, наскільки воно було трудним для виконання. Розподільча здатність тестового завдання показує, наскільки добре результати його виконання диференціюють тестованих за рівнями навчальних досягнень.

І процедура тестування, і процедура оцінювання мають відповідати вимогам *валідності*. Зокрема, при проведенні тестування необхідно дотримуватися певних правил стосовно процедури контролю, вимог, які ґрунтуються на положеннях психології тестування. Разом з тим, не менш важливим етапом педагогічного тестування є оцінювання – один із завершальних етапів вимірювання рівня знань. Процедура та методика оцінювання суттєво впливають на кінцеві результати, на валідність тестового контролю в цілому. Процедура оцінювання полягає в конвертації одержаного при тестуванні результату у певну нормовану шкалу балів – оцінку.

Для унаочнення зазначених позицій на рис. 1. показано взаємозв’язок щойно схарактеризованих валідностей як цілісної структури, що має знаходитися у стані рівноваги. Мова про те, що вилучення зі схеми будь-якого складника («валідність змісту», «валідність тестових завдань» тощо) призводить до «руйнування» структури, та, власне, не дозволяє говорити про валідність тестового оцінювання.

Надійність є показником якості й ефективності педагогічного тестування і визначається ступенем стійкості результатів при повторних вимірюваннях за допомогою того ж тесту або його рівноцінної заміни. Його суть полягає у тому, що повторне виконання одного й того ж тесту одним і тим же студентом має давати практично однакові результати. Надійність вимірювання залежить від об’єктивності процедур тестування, параметрів інструменту вимірювання та стабільності характеристики, що вимірюється. Ступінь надійності тестування визначається коефіцієнтом кореляції між двома рядами величин: результатами (оцінками групи) при першому тестуванні і результатами повторного тестування. Відтак, коефіцієнт надійності змінюється від 0 до 1, і має бути більшим 0,9 при вимірюваннях, за результатами яких приймаються серйозні адміністративні висновки.

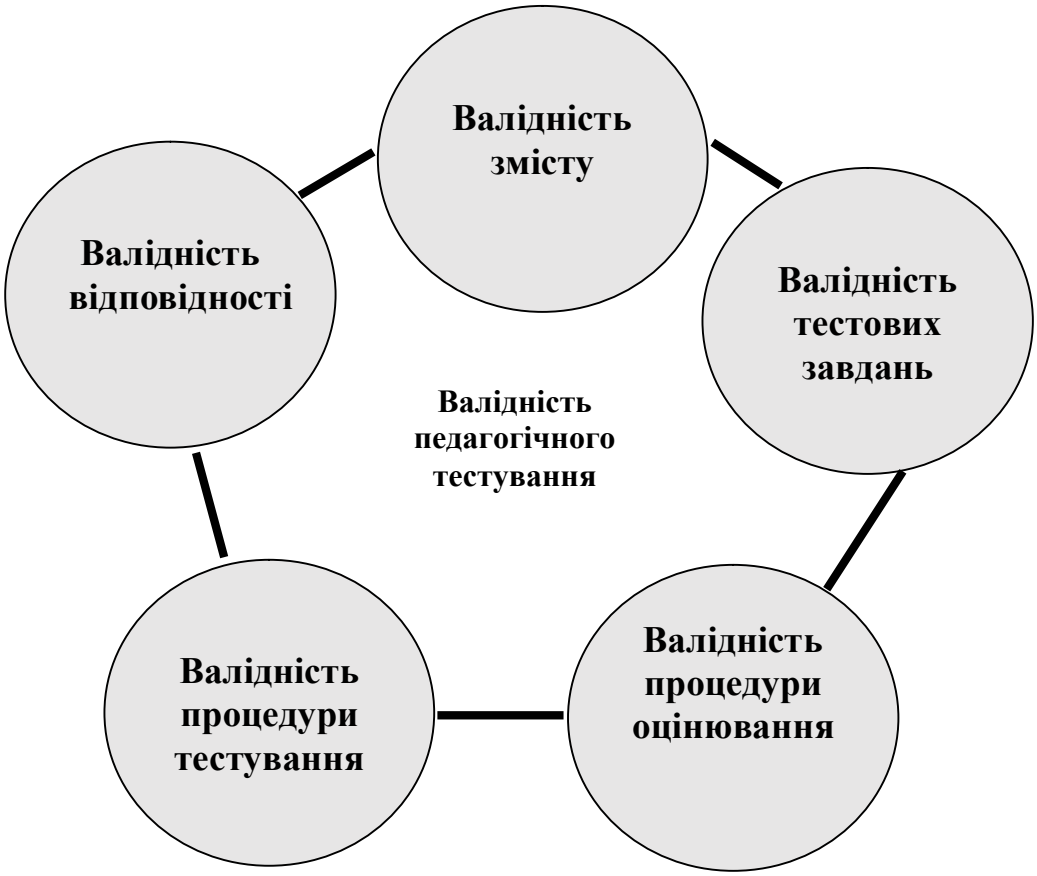

Рис. 1. Схема зв'язку різновидів валідності педагогічного тестування

Коефіцієнт надійності прямо залежить від помилки при тестуванні: будь-який метод вимірювання дозволяє отримати результати з певною точністю, яка характеризує ступінь коливання ознаки, що вимірюється. *Точність* методу вимірювання (тестування) характеризується мінімальною або систематичною помилкою, якої можна припуститися навіть при дотриманні необхідних процедур. З теорії помилок відомо, що коливання результатів вимірювання підпорядковується певним статистичним закономірностям. Це надає можливість розрахувати коефіцієнт точності тестування та мінімізувати помилки вимірювання освітні результати студентів.

Перспективи подальших наукових розвідок будуть пов'язані з обґрунтуванням технології тестового оцінювання якості підготовки фахівців техніко-технологічних спеціальностей у коледжах.