

УДК 371.27:378(73)

ГОЛОВНІ УМОВИ ЕФЕКТИВНОГО ВИКОРИСТАННЯ ТЕСТІВ УСПІШНОСТІ У ВИЩІЙ ШКОЛІ США

Наталія Чорна

*Вінницький державний педагогічний університет
імені Михайла Коцюбинського
вул. Острозького, 32, 21100 Вінниця, Україна*

Розглянуто головні вимоги до ефективного тестування успішності у вищій школі США, в основі яких лежать погляди американських дослідників щодо прийняття попередніх рішень про використання тестів, зокрема, рішення про частотність тестування, форму тестових завдань, їх кількість, рівень складності, рішення про поправку на здогадку.

Ключові слова: тест успішності, тестове завдання, об'єктивний тест, тест-твір, тест-задача, поправка на здогадку, валідність, надійність.

У концепціях української освітньої реформи помітне місце посідає проблема тестів успішності (досягнень). Зокрема, передбачається відмінити вступні іспити у вищі навчальні заклади і проводити так зване зовнішнє тестування рівня знань абітурієнтів.

Зростає інтерес до використання тестів успішності й у педагогічному процесі вищої школи. На думку багатьох викладачів, тестування є ефективним засобом контролю. Воно може забезпечити успішну реалізацію мети і всіх функцій контролю і дає можливість у досить короткий термін сформулювати об'єктивне уявлення про рівень навчальних досягнень студентів.

З огляду на актуальність цієї проблеми для системи освіти України важливо, на наш погляд, ознайомитися із досвідом тестової методики у США, де вона стала універсальним засобом перевірки знань, умінь, професійної підготовки в усіх галузях людської діяльності.

Незважаючи на свою важливість, проблема тестів успішності у системі освіти США ще не стала предметом спеціального дослідження у нашій педагогічній науці (ні в цілісному, ні в аспектному вигляді) і потребує розв'язання.

Ми ставили за мету висвітлити погляди американських дослідників щодо умов ефективного тестування у школах і коледжах США, що може мати певне теоретичне і прикладне значення.

Тестування успішності потребує попередньої підготовки (планування), що передбачає прийняття певних рішень, необхідних для ефективного використання тестів.

Однією з умов ефективного тестування, на думку Р. Ібела, є рішення про частотність тестування. Частотність проведення тестів успішності регулюється певною мірою навчальним закладом. Більшість американських викладачів вважають за необхідне проводити тестування принаймні двічі на семестр. Деякі викладачі тестують навчальні досягнення студентів кожні три – чотири тижні. У цьому є свої переваги. Тести, що подаються через короткі проміжки часу, охоплюють меншу частину навчального матеріалу, але є детальнішими. Не можна не погодитись з Р. Ібелом, що детальний тест, який охоплює меншу частину матеріалу, є ліпшим від того, що охоплює більшу частину матеріалу, але є менш детальним [2, с. 58]. Крім того, важливо зауважити, що часте тестування забезпечує надійнішу основу для оцінювання і для усвідомлення викладачами й студентами результатів їхніх досягнень. Проте, якщо викладач не має великої кількості заздалегідь підготовлених тестових завдань і досвіду їх швидкої перевірки, часте тестування забирає багато часу. У деяких випадках підготовку до тестування і тестування можна проводити у час, відведений на пояснення нового навчального матеріалу. Отже, надто часте тестування може призвести до явної деформації навчальних пріоритетів і формування у студентів хибного уявлення, що успішне складання тестів є чи не головною метою навчання. Уже доведено, що вдало сконструйовані й раціонально використані тести мають певну дидактичну цінність. Виконуючи тестові завдання, студент ніби ще раз проходить навчальну програму. Однак важливо пам'ятати, що головна функція тестів не навчальна, а перевірка.

Важливою умовою ефективного тестування є прийняття рішення про використання певної форми (типу) тестових завдань. Форми тестових завдань, які використовують найчастіше, такі: тести-твори (есе), або дискусійний тип, об'єктивний тип, або тип короткої відповіді, й тип математичної задачі (тест-задача). Тест-задача широко використовується у процесі тестування навчальних досягнень з таких предметів, як математика, деякі природничі науки, інженерія тощо. Для багатьох інших предметів більш придатні або тести-твори, або об'єктивні тести.

Подаємо коротку порівняльну характеристику зазначених типів тестових завдань, яку наводять американські дослідники у контексті виявлення умов ефективного тестування.

На думку американських авторів, насамперед, потрібно позбутися деяких хибних уявлень. Некоректним, зокрема, є уявлення про те, що один тип тестів забезпечує цілковиту перевірку знань, тоді як інші типи тестів перевіряють лише поверхові знання. Як стверджують М. Річардсон і Дж. Столнейкер, “сама форма тесту ще не є запорукою виявлення тієї здатності, рівень наявності якої перевіряється” [5, с. 427]. Помилковою також є думка, нібито везіння відіграє велику роль в одному типі тестування й

цілком відсутнє в іншому. Навпаки, усі три типи можуть виявити один і той же рівень можливостей і у разі ретельного проведення дати результати майже однакової надійності. Досконалий тест-твір або досконалий об'єктивний тест можуть визначити рейтинг групи студентів майже у тій самій послідовності, що і тест-задача. Але це аж ніяк не означає, що усі три типи тестів можуть з однаковою легкістю та ефективністю взаємозамінюватись. Ф. Вернон звернув увагу на очевидність того, що "тести однакових цілей, але різних форм, здатні давати суперечливі результати. Тести однакової форми, розроблені для різних предметів та розумових функцій, добре інтеркорелюють. Для багатьох цілей простіші тести засвідчують більшу валідність, і є сумнівним те, наскільки складніші тести можуть адекватно перевірити вищі розумові функції, на виявлення яких вони спрямовані" [6, с. 27].

Як вважає Р. Ібел, у порівнянні з об'єктивним тестом тест-твір та тест-задачу легше підготувати, але оцінювання й облік успішності об'єктивний тест забезпечує швидше, ніж інші типи тестів, особливо тест-твір. Коли йдеться про тестування великої кількості студентів, використання об'єктивного тесту дає змогу досягнути ефекту з малою або нульовою втратою валідності. Коли ж тестування проводиться у малих групах чи класах, навпаки, перевага повинна надаватись тестам-творам та тестам-задачам, оскільки у цьому випадку вони забезпечують більшу валідність [2, с. 59].

Не можна не погодитись з Р. Ібелом, що серед усіх типів тестів тест-задача має перевагу у внутрішній релевантності. Цей тип тесту більше, ніж інші, дає можливість виявити знання у діяльності, ступінь оволодіння знаннями для розв'язання конкретної проблеми. Тести-твори та об'єктивні тести містять багато поверхових і суто академічних завдань, чого, можливо, і потрібно уникати.

Водночас об'єктивні тести мають свої переваги порівняно з тестами-творами та тестами-задачами. Можливості використання тестів-задач певною мірою обмежуються обсягом і складністю відповідей, які на них треба дати (відповіді повинні бути написані від руки). Правда, наявні твердження, що здатність вибирати відповідь є менш важливою, ніж здатність її продукувати. Проте "більшість тверджень засвідчують, що ці можливості тісно пов'язані" [2, с. 60].

Розглядаючи порівняльні переваги тесту-твору, тесту-задачі й об'єктивного тесту, важливо, на наш погляд, пам'ятати, що єдиним корисним компонентом будь-якого тестового бала (оцінки) є компонент, який би його об'єктивно підтверджував, незважаючи на тип тесту, на підставі якого цей бал виставлений.

З огляду на це дуже важливо уникати суб'єктивних вражень і непереверених суджень з боку тих, хто тестує й оцінює студентів.

Об'єктивним тестам (тестам короткої відповіді) та тестам-задачам об'єктивність притаманна більшою мірою, ніж тестам-творам. У них чіткіше визначені завдання, і тому підстави для вищої чи нижчої оцінки тут очевидні. Хоч відомо, що навіть об'єктивні тести ґрунтуються на великій кількості суб'єктивних рішень стосовно того, що тестувати і як тестувати. А це не може не впливати на їхню об'єктивність. Відповідно, тести-задачі мають додатковий елемент суб'єктивності оцінювання, якого немає в об'єктивному тесті. У них питання, зокрема, як оцінювати недосконалу відповідь і які моменти у тому чи іншому випадку брати до уваги, часто вирішуються спонтанно й суб'єктивно.

З метою якісного оцінювання екзаменатор повинен намагатися, незважаючи на тип тесту, оцінювати якомога об'єктивніше. Оцінювання може бути об'єктивним настільки, наскільки воно незалежно підтверджене іншими компетентними екзаменаторами. Зрозуміло, що оцінки (бали), одержані за досконалий тест-твір, є набагато об'єктивнішими, ніж оцінки (бали), одержані за недосконалий тест множинного вибору. Водночас, не можна не погодитись з твердженням Р. Ібела про те, що «викладачі, які використовують тести-твори, мають тенденцію менше дбати про об'єктивність своїх оцінювань та оцінок, ніж ті, хто використовує тести множинного вибору» [2, с. 64].

Крім того, зазначає Р. Ібел, у більшості випадків викладачі вибирають ті типи тестових завдань, які їм видаються найкориснішими, або ті, у використанні яких вони почуваються особливо компетентними. Хоча, можливо, що зумовлено звичкою чи якимись помилковими уявленнями, вони не використовують інші типи тестів, коефіцієнт корисної дії яких міг би бути вищим [2, с. 64].

Деякі науковці США вважають, що під час тестування повинен використовуватись якомога ширший спектр різнотипних тестових завдань. На їхню думку, такий підхід може поліпшити валідність тесту та зробити його цікавішим. Інші дослідники зазначають, що розробник тестів повинен насамперед вибирати той тип тестового завдання, який найбільше відповідає змісту теми та навчальним цілям, досягнення яких заплановано перевірити. На думку Б. Такмана, вибір типів тестових завдань ґрунтується на дієслові, що використовується для формулювання кожної мети. Такі дієслова, як *називати*, *перелічувати* й *констатувати*, вимагають короткої відповіді (зокрема, у завданнях, де потрібно закінчити речення), а дієслова *з'ясувати*, *виявити відмінності* чи *розпізнати* спонукають студента до вибору правильної відповіді з багатьох можливих (зокрема, у типах завдань, де можна вибрати відповідь із багатьох чи підібрати відповідну). Дієслово *класифікувати* вимагає віднесення предметів до певних категорій, що, зазвичай, передбачає наявність двох варіантів вибору. Ці типи дієслів

вичерпують формат короткої відповіді на запитання у завданні, оскільки завдання, що потребують короткої відповіді, виявляють спроможність студента вказати на правильну відповідь або придумати дуже коротку відповідь.

Довших відповідей (описових, аналітичних чи пояснювальних) вимагають завдання тесту-твору (есе). Хоч тут, як і в завданнях із наявністю варіантів множинного вибору, розробник тестів пропонує варіанти вибору правильної відповіді, а студент повинен вибрати лише яку-небудь конкретну відповідь. У завданні типу есе розробник вказує на те, що слід описати, а студент описує.

Деякі дієслова, наприклад, *продемонструвати* й *побудувати*, передбачають фактичне виконання команди. Багато цілей у математиці, природничонаукових предметах або у мистецтві вимагають, щоб студенти продемонстрували методику виконання.

У цілях тесту міститься вся інформація, яка необхідна, щоб вирішити, якого типу завдання повинні бути використані у тесті. Якщо обраний тип тестового завдання відображає дієслово дії у меті тесту, але не відображає типу дії, яка передбачається викладачем, то таку мету тесту слід переписати так, щоб вона містила дієслово. Як цілі тесту, так і завдання, призначені для їх виявлення, повинні якомога більше відповідати одне одному за характеристиками дії, умовами й критеріями оцінювання результатів [1, с. 461–462].

Таблиця дій і типів тестових завдань (за Б. Такманом) [1, с. 462]:

Дія	Тип завдання
Назвати Констатувати Перелічити	Завершення
Виявити Вирізнити Розпізнати	Множинний вибір, підбір відповіді
Класифікувати	Альтернативний вибір
Описати Проаналізувати Пояснити	Твір (есе)
Продемонструвати Побудувати	Виконання дії

На наш погляд, запропонований Б. Такманом підхід до розробки тестових завдань має багато переваг, однак його не слід розглядати як абсолютний імператив під час конструювання тесту. Ми поділяємо погляд Р. Ібела, який вважає, що “якість тесту успішності більше залежить від того, скільки уваги у ньому приділяється вимірюванню різних видів навчальних

досягнень студентів, а також від якості тестових завдань, а не від їхніх типів” [2, с. 64].

Навчальні досягнення у багатьох курсах, зазвичай, охоплюють оволодіння певною сумою корисних знань та умінь виконувати практичні завдання. Знання досить зручно можна поділити на знання вокабуляра та на знання фактів. Уміння, здебільшого, охоплюють уміння пояснювати та застосовувати знання для виконання відповідних дій у практичній ситуації. Деякі навчальні предмети спрямовані на розвиток інших умінь, таких як уміння вести розрахунки, прогнозувати тощо.

Відповідно більшість тестових завдань, як вважає Р. Ібел, можна порівняно легко й точно віднести до таких семи категорій:

1. Розуміння термінології (вокабуляра).
2. Розуміння фактів і принципів (узагальнення).
3. Здатність пояснювати чи ілюструвати (розуміння взаємовідносин, зв'язків).
4. Здатність обчислювати (числові проблеми).
5. Здатність прогнозувати (що може трапитись за певних умов).
6. Здатність рекомендувати певні дії (у певних практичних проблемних ситуаціях).
7. Здатність продукувати оцінювальні судження.

Завдання, що належать до категорії розуміння термінології, завжди визначають термін, який потрібно з'ясувати чи перефразувати. Завдання, що стосуються фактів і принципів, базуються на описових твердженнях. Завдання, що перевіряють уміння пояснювати, зазвичай, містять слова “чому” чи “тому, що”. Завдання, що належать до категорії здатності обчислювати, вимагають від студентів використання математичних дій для того, щоб вибрати правильний результат з усіх поданих. Завдання, що належать до категорій здатності прогнозування та рекомендування певних дій, ґрунтуються на описі конкретних ситуацій. Завдання на здатність передбачати (прогнозувати) визначають специфіку всіх необхідних для цього умов і вимагають подальших дій, а завдання на уміння рекомендувати певні дії у якій-небудь ситуації конкретизують деякі умови і вимагають відповіді на те, які інші умови чи дії приведуть до певного результату. У завданнях на здатність продукувати оцінювальні судження містяться твердження, цінність чи якість яких повинна бути оцінена на підставі критеріїв, зазначених у тесті.

Корисність таких категорій класифікації тестових завдань із багатьох аспектів навчальних досягнень полягає у тому, що вони виражені переважно мовою змісту навчання, а не мовою якихось гіпотетичних розумових процесів чи логічних операцій, необхідних для відшукування правильної відповіді. Відповідне співвідношення (пропорція) тестових завдань кожної із семи категорій залежить від специфіки навчального предмета. Однак у будь-

якому випадку найкращі тести – це ті, які більше акцентують на застосуванні знань, а не на простій репродукції їхнього вербального оформлення. Написати завдання на перевірку здатності оцінення суджень значно важче, ніж розробити тестові завдання на перевірку здатності відтворювати знання. І доти, справедливо зазначає Р. Ібел, доки розробник тесту попередньо й чітко не вирішить, у якому співвідношенні тестові завдання повинні відображати специфічні аспекти навчальних досягнень, його тест ніколи не стане досконалим [2, с. 72].

Третьою важливою умовою ефективного планування тестів, на думку американських авторів, є рішення про кількість тестових завдань. “Перед написанням тесту необхідно вирішити, які завдання і скільки їх необхідно ввести у тест”, – зазначає Б. Такман [1, с. 459].

На питання про кількість завдань у тесті американські дослідники не дають однозначної відповіді. На їхню думку, з якою неможливо не погодитись, це залежить від кількості й типу навчальних цілей, результати реалізації яких потрібно виявити, типу тестових завдань, часу, що відводиться на тестування, і від віку студентів (Р. Ібел, Б. Такман, Т. Крол, С. Камінські та ін.).

Розв’язуючи питання про кількість завдань у тесті, корисно, на думку Р. Ібела, розглядати їх як певну модель (зразок) усєї гіпотетичної кількості завдань, що могли би бути введені у тест. Викладач іноземної мови, наприклад, міг би ввести у тест на правопис 100 слів із 500 тих, які вивчали протягом семестру (500 – це і є та кількість, з якої сформована модель із 100 слів). Така модель є реально визначеною. Однак для більшості тестів ситуація може бути іншою. Наприклад, майже необмеженою є кількість задач, які можна використовувати у тесті з алгебри, як і необмеженою є кількість різних запитань, які можна сформулювати для використання у тесті з історії. Розробники тестів із цих предметів, як і з багатьох інших, не мають заздалегідь визначеного, обмеженого списку, з якого можна було б вибирати завдання для введення у тест. Однак їхні тести є моделями, оскільки вони охоплюють лише певну частину завдань, які можуть бути наведені у кожному випадку.

Чим більший обсяг змісту навчального предмета, чим ширше коло умінь і навичок, якими повинен оволодіти студент, тим більшою має бути кількість потенційно можливих завдань у тесті, які планує розробник із певного предмета. Уся можлива кількість цих завдань є верхньою межею тієї моделі, яка може бути сформована з цієї кількості. Це означає, що модель не може бути більшою від усєї кількості можливих завдань. Але вся можлива кількість завдань не визначає нижньої межі моделі. Наприклад, 1000 можливих тестових завдань може бути подана тестовою моделлю з десятьма,

п'ятдесятьма чи сотнею завдань. Це ж стосується і 100 000 потенційних завдань та ін.

“Дуже важливо, – зазначає Р. Ібел, – розробити таку тестову модель, яка б адекватно відображала всю потенційно можливу кількість завдань, властивих певному тесту” [2, с. 66].

Простежується тенденція створювати тести з такою кількістю завдань, щоб студенти мали можливість виконувати їх, працюючи з нормальною швидкістю [2]. Однією з причин є те, що швидкість виконання завдань не вважається важливою навчальною метою щодо більшості предметів, які вивчаються у середніх школах і коледжах, і, отже, не впливає на показники навчальних досягнень. Іншою причиною є хвилювання й переживання студентів у процесі тестування без часових обмежень, які значно зростають з вимогою працювати швидко й точно у визначених часових межах. Викладач, який багато зусиль затратив на розробку тесту, цілком природно зацікавлений у тому, щоби більшість студентів устигли виконати усі тестові завдання. У деяких ситуаціях тести на швидкість можуть бути цінними й доцільними, але такі ситуації є більшою мірою винятком, ніж правилом.

“Дуже важливо, – вважають автори відомого американського посібника “Психологія навчання”, – щоби тестові завдання охоплювали всі шість когнітивних рівнів таксономії навчальних цілей Блума (1956)” [3, с. 324]. Це, на їхній погляд, є запорукою розробки досконалих та високоефективних тестів.

Як зазначає Б. У. Такман, на кожен мету потрібно давати щонайменше два тестових завдання. Чим більше тестових завдань є для такого контролю, тим вищою є його точність (надійність) [1, с. 459]. Цю думку поділяє Р. Ібел, який пише: “Чим більшою є кількість завдань у тесті, тим точнішими будуть результати тестування навчальних досягнень студентів у певній галузі знань” [2, с. 66].

Отже, необхідно написати достатню кількість завдань, щоби заповнити час, відведений на тестування, і мати якомога більше завдань на перевірку досягнення кожної мети навчання.

Однак не всі цілі є однаково важливі, більше навчального часу витрачається на важливіші цілі. З огляду на це тести повинні містити більше завдань для визначення важливіших цілей. “Важливість кожної мети, – слушно зазначає Р. Ібел, – потрібно зважити, порівнюючи її з важливістю інших цілей, і кількість завдань для кожного визначення слід підбирати з урахуванням такого зважування” [2, с. 461].

Одним із важливих рішень у процесі підготовки до тестування є, як вважають американські дослідники, рішення про рівень складності тестових завдань. На їхню думку, можливі два шляхи розв'язання цієї проблеми. Перший – це введення до тесту лише тих проблем і завдань, на які студент,

котрий навчався успішно, спроможний дати відповідь. За такої умови більшість студентів теоретично повинні відповісти на більшу кількість завдань правильно. Отож, правильних відповідей може бути так багато, що такі завдання втратять можливість ефективно виявляти різні рівні навчальних досягнень – найкращий, добрий, середній, слабкий, поганий. Інший підхід передбачає вибір таких тестових завдань, які би за своєю складністю відображали різні рівні навчальних досягнень тих, кого тестують. Це переважно важчі завдання, які приблизно половина студентів може виконати, а інші – ні. Цей підхід, загалом, забезпечує надійніші результати за той же час тестування. Однак він більш обтяжливий для тих студентів, навчальні досягнення яких незначні.

Деякі викладачі вважають, що довершений тест повинен містити частину важких завдань для тестування сильних студентів і частину легких завдань для тестування слабших. Це можливо, якщо кожна нова тема навчального курсу чи кожне нове його положення вимагають оволодіння усіма його попередніми темами чи положеннями. У межах такого курсу студенти відрізнятимуться тим, наскільки успішно вони оволоділи навчальним матеріалом, а не тим, скільки окремих положень курсу вони засвоїли. На жаль, лише незначна кількість навчальних курсів чи предметів демонструють таку довершену послідовність тем та ідей.

Студент, який пропустив деякі попередні положення чи не опрацював окремі попередні теми курсу, зазвичай, надалі відставатиме у навчанні. Однак послідовність змісту курсу рідко настільки чітко визначена, що будь-яка прогалина гальмуватиме подальше навчання. Курси іноземних мов і курси з певних галузей математики та інженерії демонструють більшу послідовність й узгодженість своїх тем, аніж курси з інших галузей знань, але навіть у них ця послідовність і узгодженість далека від абсолюту. У більшості навчальних предметів відмінність між сильними й слабкими студентами виявляється не стільки глибиною й повнотою оволодіння ними усім навчальним матеріалом, скільки кількістю правил, що вони запам'ятали, чи кількістю речей, що вони навчилися робити. Стосовно таких навчальних предметів мета урізноманітнення складності тестових завдань утрачає свій сенс, принаймні, доти, поки склад студентів і надійність тесту не зазнають суттєвих змін. Теоретичний аналіз та експериментальні дослідження американських науковців переконливо засвідчують, що у більшості випадків тести з не дуже складними і не дуже легкими завданнями є найліпшими. Зокрема, М. Річардсон зазначає, що “тест, який містить завдання середньої складності, має загальну валідність, що є вищою, ніж валідність тих тестів, які містять завдання будь-якого іншого рівня складності” [5, с. 37]. А. Г. Гуліксен на підставі теоретичного аналізу робить такий висновок: “З метою підвищення надійності й варіативності будь-якого тесту його завдання

повинні тісно співвідноситись, мати один рівень складності, і цей рівень повинен якомога більше наблизитись до середнього” [4, с. 91]. Аналогічну думку висловлює і Б. У. Такман, який, зокрема, пише: “Щоб розробити надійніші тести, викладачі повинні складати завдання середньої складності, уникаючи написання завдань, які є занадто важкими або занадто легкими” [1, с. 511].

Як вважає Р. Ібел, однією з умов ефективного тестування є прийняття рішення про поправку на здогадку.

Показники об’єктивних тестів інколи “корегують на вгадування” шляхом віднімання якоїсь частини неправильних відповідей від загальної кількості правильних відповідей. Метою таких поправок є заохочення студентів до не зазначення окремих відповідей замість сліпого вгадування у випадку, коли деякі тестові завдання є занадто складними для них. Незважаючи на те, що поправка на здогадку використовується у процесі виставлення балів і значною мірою не впливає на підготовку самого тесту, рішення щодо неї повинно прийматись до подання тесту. По-перше, запитань щодо можливої поправки і всього, що з нею пов’язано, від студентів можна очікувати, коли тест уже поданий, і викладач повинен бути готовим відповісти на ці запитання однозначно й чесно. По-друге, відчутний ефект від поправки на здогадку (якщо він узагалі є) слід вважати більш психологічним, ніж математичним. Якщо оголошення поправки на здогадку заохочує студентів уникати відповідей, які ґрунтуються на сліпому чи майже сліпому вгадуванні, воно може певною мірою підвищити надійність і валідність показників тесту [2, с. 75–76].

Ми поділяємо думку Р. Ібела, що всі плюси і мінуси від уведення поправки на здогадку можна звести до таких восьми пунктів [2, с. 76]:

1. Бали, виставлені з урахуванням поправки на здогадку, здебільшого, визначають той самий рейтинг студентів, як і у випадку без неї.
2. Імовірність одержання високого бала у досконалому об’єктивному тесті лише через сліпу здогадку є дуже малою.
3. Добре мотивовані студенти, котрі мають достатньо часу на розв’язання усіх тестових завдань, майже не вгадують відповіді, а якщо і вгадують, то лише окремі з них.
4. Зазвичай, немає ніякої ні моральної, ні навчальної шкоди у заохоченні студентів до раціональних здогадок.
5. Раціональна здогадка студента може надати корисну інформацію про його загальний рівень розвитку.
6. Якщо тест обмежений у часі і виконується на швидкість, поправка на здогадку спонукає повільних студентів до сліпого вгадування.

7. Результати з урахуванням поправки на здогадку можуть охоплювати недоречні моменти, пов'язані з надмірною обачністю студента та його небажанням ризикувати.

8. Поправка на здогадку ускладнює систему оцінювання і веде до зниження точності в оцінюванні.

На нашу думку, з огляду на кількість і складність аргументів за і проти поправки на здогадку і зважаючи на вагомість заперечень прибічників будь-якої з цих альтернатив рішення про прийняття чи неприйняття поправки повинно прийматись одразу і без будь-яких подальших вагань. Оскільки використання поправки призводить до ускладнення процесу оцінювання, незначні переваги, пов'язані з її застосуванням, цілком і повністю залежать від совісті її прибічників. Більшість об'єктивних тестів досягнень, зазвичай, може бути оцінена цілком задовільно шляхом лише врахування правильних відповідей. Водночас ми допускаємо, що за певних обставин прийняття поправки на здогадку перед тестуванням може бути не тільки виправданим, але й необхідним.

Отже, підсумовуючи усе вищезазначене, можна зробити висновок, що з метою створення сприятливих умов для ефективного тестування американські дослідники висловлюються за прийняття відповідних рішень, в основі яких лежать такі міркування:

1. Частіше тестування забезпечує надійнішу основу для оцінювання.
2. Форма тесту суттєво не впливає на вимірювання навчальних досягнень.
3. Кількість і типи тестових завдань повинні адекватно відображати мету та зміст навчання.
4. Тестові завдання не повинні бути ні надто простими, ні надто складними.
5. Незважаючи на те, яку форму тесту чи тип тестового завдання використовує екзаменатор, він повинен намагатись оцінити студентів якомога об'єктивніше.
6. Більшість тестових завдань не повинні бути тривалими за часом їх виконання з тією метою, щоби майже усі студенти встигли їх виконати.
7. Використання поправки на здогадку суттєво не впливає на підвищення валідності більшості тестів успішності.

Ми розглянули лише окремі найважливіші умови ефективного використання тестів успішності у системі освіти США, що не вичерпує усіх аспектів цієї проблеми.

Перспективи подальших розвідок у цьому напрямку вбачаємо у висвітленні поглядів американських дослідників на інші умови ефективного тестування, що дасть можливість сформулювати повніше і глибше уявлення про стан цієї проблеми у педагогіці США.

1. *Такман Б. У.* Педагогическая психология: от теории к практике / Пер. с англ. М., 2002.
2. *Ebel R.* Measuring Educational Achievement. N.J., 1965.
3. *Crowl Th. K., Kaminsky S., Podell D. M.* Educational Psychology: Windows on teaching. Chicago, 1997.
4. *Gulliksen H.* The Relation of Item Difficulty and Inter-item Correlation to Text Variance and Reliability // *Psychometrika*. 1945. Vol. 10. P. 79–91.
5. *Richardson M. W. and Stalnaker J. M.* Comments on Achievement Examinations // *Journal of Educational Research*. 1935. Vol. 28. P. 425–32.
6. *Vernon Ph. E.* Educational Testing and Test Form Factors // *Research Bulletin*. 1958. Vol. 3. P. 23–30.

BASIC REQUIREMENTS FOR EFFECTIVE TESTING OF EDUCATIONAL ACHIEVEMENT IN HIGHER SCHOOL OF THE USA

Nataliya Chorna

*Mikhailo Kotsiubynsky State Pedagogical University of Vinnytsia
Ostroz'ky Str., 32, UA–21100 Vinnytsia, Ukraine*

The article deals with general considerations of American educators on principle requirements for effective testing of educational achievement in American higher school. They are – frequency of testing, types of test questions, their number, level of difficulty, correction by guessing.

Key words: achievement test, test question, objective test, essay-test, problem-test, correction by guessing, validity, reliability.

Стаття надійшла до редколегії 08.04.2004
Прийнята до друку 19.11.2004