УДК 37.012-048.24:373.3

DOI: 10.12958/2227-2844-2021-8(346)-1-207-223

Хорошковська-Носач Тетяна Петрівна,
науковий співробітник відділу навчання мов
національних меншин та зарубіжної літератури
Інституту педагогіки Національної академії
педагогічних наук України, м. Київ, Україна.
khoroshkovska@ukr.net
<https://orcid.org/0000-0001-9505-0817>

ТЕСТУВАННЯ ЯК ОДИН ІЗ МЕТОДІВ ПЕДАГОГІЧНОГО ВИМІРЮВАННЯ

Проблема вимірювання результатів навчально-пізнавальної діяльності учнів є однією з важливих і актуальних у сьогоденні. У сучасній дидактиці та методиці навчання виокремлюють різні групи методів, зокрема й контрольно-перевірні. Контроль результатів навчально-пізнавальної діяльності учнів здійснюється за допомогою методів усної та письмової, а також практичної перевірок. Одним із методів контролю є систематичне спостереження вчителем за навчально-пізнавальною діяльністю учнів. Однак, найбільш об'єктивним методом вимірювання в педагогіці вважається тестування, а у якості вимірників використовуються педагогічні тести.

Проблемою розробки тестів у різні часи займався багато вчених. Серед них В. Аванесов, А. Анастасі, Т. Балихіна, М. Бернштейн, П. Блонський, Ф. Гальтон, Н. Гронланд, А. Кетле, П. Клайн, О. Майоров, В. Маккол, І. Рапопорт, Р. Сельг, І. Соттер, М. Челишкова та інші. В Україні цю проблему досліджують І. Булах, Л. Ващенко, А. Гривко, Ю. Жук, Ю. Ковальчук, Г. Леонська, В. Леонський, О. Ляшенко, С. Науменко, Л. Парашенко та інші.

Тестування, за Ф. Гальтоном, – це метод, який передбачає застосування однакових дослідів щодо великої кількості індивідів із подальшою статистичною обробкою одержаних результатів і визначенням еталонів оцінки. Він, скориставшись роботами А. Кетле, визначив 3 основні умови тестування: 1) застосування серії однакових випробувань щодо великої кількості людей, яких тестують; 2) статистична обробка одержаних результатів; 3) створення еталонів оцінки як основи стандартизованого тестування.

Американець В. Маккол, досліджуючи проблему тестування, розділив тести на педагогічні та психологічні. Основними завданнями педагогічних тестів стало вимірювання успішності учнів зі шкільних дисциплін за визначений період навчання, а також ефективність методів

і технологій викладання (Паращенко & Леонський & Леонська, 2006, с. 13). Однак засновником педагогічних вимірювань вважають американського психолога Едуарда Лі Торндайка, якому належить розробка першого педагогічного тесту.

Мета статті – висвітлити основні питання, пов'язані з проблемою тестування як найбільш об'єктивного методу вимірювання результатів навчально-пізнавальної діяльності учнів.

Дослідження проблеми тестування як одного з методів педагогічного вимірювання потребує уточнення поняття «тест». У літературі немає єдиного погляду на визначення цієї дефініції.

Так, В. Аванесов (Аванесов, 1989, с. 12) поняття «тест» трактує як метод педагогічного вимірювання і як результат застосування тесту як методу вимірювання, що складається з обмеженої кількості завдань. О. Майоров розглядає тест як «інструмент, що складається з кваліметрично вивіреної системи тестових завдань, стандартизованої процедури проведення й заздалегідь спроектованої технології обробки й аналізу результатів, призначений для вимірювання якостей і властивостей особистості, вимірювання яких можливе в процесі систематичного навчання» (Майоров, 2000, с. 33). Таким чином, О. Майоров розглядає тест і як метод, і як засіб вимірювання навчально-пізнавальної діяльності учнів. М. Челишкова (Челышкова, 2002, с. 25–26) визначає тест як сукупність завдань, які дозволяють об'єктивно оцінити якість підготовки учня в певній освітній галузі.

На сучасному етапі в Україні тест розглядають з однієї сторони як «стандартизоване завдання, виконання якого дає можливість вимірювати психологічні та інтелектуальні характеристики учня, рівень його знань», а з іншої – як «метод випробування, що застосовується у різних галузях науки для одержання кількісної характеристики певних явищ» (Бондар, 2005, с. 253). О. Ляшенко визначає тест як «спеціально дібраний набір тестових завдань, стандартизованих процедур проведення і заздалегідь визначених методик оброблення та аналізу одержаних даних, що в сукупності призначені для отримання кількісної характеристики певної властивості особи, яка підлягає оцінюванню» (Ляшенко & Лукіна & Жук & Ващенко & Науменко & Гривко, 2014, с. 15). Поняття «тест» також тлумачиться як сукупність запитань, які переважно вимагають однозначної відповіді, укладені за певними правилами та процедурами, передбачають попередню експериментальну перевірку й відповідають таким характеристикам ефективності, як валідність і надійність (Короткий текстологічний словник-довідник, 2008, с. 122).

Головною відмінністю тесту від контрольної роботи є те, що він передбачає вимірювання (Короткий текстологічний словник-довідник, 2008, с. 122).

Тест, як правило, складається з двох частин: 1) інформаційної, яка містить інструкцію та приклади виконання завдань; 2) операційної, що складається з певної кількості завдань і запитань (Короткий текстологічний словник-довідник, 2008, с. 123).

Науковці, які вивчають питання тестування в освіті (Аванесов В., Булах І., Майоров О.), висувають певні вимоги до структурування тесту. Некоректними вважаються тести з т. зв. «ланцюговими завданнями», тобто такими, коли правильна відповідь на наступне завдання залежить від точності виконання попереднього (Лаврова, 1999, с. 46). Також вважається, що дистрактори – неправильні, але найбільш правдоподібні варіанти відповідей – у тестових завданнях із множинним вибором будуть ефективними, якщо відволікатимуть увагу погано підготовлених учнів (Лаврова, 1999, с. 54). Коли ж один із відволікаючих варіантів не обрав жоден з учнів, такий дистрактор не виконує своєї функції, тому не є ефективним, його потрібно доопрацювати. Удало дібрані дистрактори зменшують імовірність угадування правильної відповіді. У тестології визначено два прийоми формулювання дистракторів: «1) використання правильних тверджень, які, однак, правильно не відповідають на запитання в основі; 2) використання знайомих виразів і фраз. Оскільки їх часто використовують, такі відповіді можуть бути привабливими для тестованих, які мають поверхові знання з предмета» (Короткий текстологічний словник-довідник, 2008, с. 22).

Запитання й завдання тесту, які ще не пройшли попередню експериментальну перевірку, називаються завданнями в тестовій формі, а завдання, що пройшли обов'язкову перевірку статистичних властивостей, називаються тестовими (Короткий текстологічний словник-довідник, 2008, с. 51).

Завдання в тестовій формі мають відповідати низці вимог, а саме: відповідати меті тестування; бути лаконічними; логічно сформульованими; послідовно розміщеними; містити місце для відповіді; мати однакову інструкцію, адекватну формі та змісту завдання; оцінювання повинне здійснюватись за однаковими вимогами (Короткий текстологічний словник-довідник, 2008, с. 32).

За характером відповіді на запитання тестові завдання поділяють на кілька типів, видів, форм (у літературі відсутня єдина класифікація їх). Так, І. Булах (Булах, 1995) розрізняє: 1) відкриті (завдання з відкритою відповіддю, мікротвір-есе); 2) напіввідкриті (з пропусками, на доповнення); 3) закриті тестові завдання (альтернативні та вибіркові). Р. Бобовський (Бобовський, 2010, с. 49) виокремлює чотири основні форми тестових завдань: 1) закриті; 2) відкриті завдання; 3) на встановлення відповідності; 4) на встановлення правильної послідовності. У міжнародному дослідженні якості освіти PISA (Programme for International Student Assessment) визначено такі типи тестових завдань: 1) з вибором відповіді; 2) із закритою вільно

конструйованою відповіддю; 3) із відкритою вільно конструйованою відповіддю; 4) з короткою відповіддю.

На нашу думку, усі завдання в тестовій формі по суті становлять два типи: відкриті й закриті, які у свою чергу можуть мати певні модифікації.

Завдання закритого типу – їх також називають «закриті завдання», «завдання з вибірковими відповідями», «завдання на вибір відповіді» «завдання з множинним вибором» (Паращенко & Леонський & Леонська, 2006; Балыхина, 2006) – містять передбачені відповіді, учневі необхідно обрати одну або кілька правильних серед запропонованих.

До завдань закритого типу дослідники ставлять низку вимог. Серед них:

- відсутність двозначності чи неточності у формулюваннях завдань;
- основа завдання (операційна частина тесту) має формулюватися у стверджувальній формі, яка після підстановки відповіді перетворюється на істинне або хибне твердження;
- основу завдання треба формулювати дуже коротко, одним реченням (до 8 слів) у формі простої синтаксичної конструкції;
- якщо варіанти відповідей починаються з одного слова (кількох однакових слів), то їх потрібно перенести в основу завдання, формулюючи варіанти відповідей двома–трьома словами;
- усі варіанти відповідей мають бути приблизно однакової довжини або в деяких завданнях правильний варіант відповіді може бути дещо коротший за дистрактори;
- частота вибору варіанта правильної відповіді в усіх завданнях тесту має бути приблизно однаковою;
- у відповідях не рекомендується використовувати слова «все», «жоден», «ніколи», «завжди», «усі названі», «жоден із названих» тощо;
- не можна використовувати завдання, у яких одна відповідь впливає з іншою (т. зв. «ланцюжкові завдання»);
- усі варіанти відповідей мають бути граматично узгодженими з основою завдання (Короткий тестологічний словник-довідник, 2008, с. 41–42).

До закритих завдань дослідники відносять такі: 1) завдання з вибором однієї або кількох правильних відповідей із запропонованих варіантів; 2) завдання на встановлення відповідності (утворення логічних пар); 3) завдання на встановлення правильної послідовності.

І. Завдання з однією правильною відповіддю містять кілька варіантів відповідей (як правило, 4 або 5), серед яких учень повинен обрати одну. У практиці вітчизняного й зарубіжного тестування ця форма має кілька модифікацій: 1) завдання з альтернативними

відповідями; 2) завдання на встановлення причиново-наслідкових зв'язків; 3) завдання з пропуском.

Завдання з альтернативними відповідями «використовують, щоб перевірити правильність констатації фактів. Вони складаються з основи – твердження або запитання, яке тестований має визначити як правильне чи неправильне – та двох альтернатив. Учень має відповісти «так» або «ні», «правильно» чи «неправильно», «істинне твердження» чи «хибне» тощо» (Короткий тестологічний словник-довідник, 2008, с. 33). Як і будь-які завдання, завдання з альтернативними відповідями мають свої переваги та недоліки. Серед переваг визначають такі: оцінювання результатів є швидким, простим, об'єктивним; вимагають менше часу, щоб прочитати завдання; тестований може відповісти на значно більшу кількість запитань протягом відведеного проміжку часу, ніж коли він виконує завдання інших форм. До недоліків відносять те, що результати тестування на 50 % залежать від угадування.

Завдання на встановлення причиново-наслідкових зв'язків складаються з певного твердження й пояснення до нього (формулювання причини). Учень має встановити, чи є твердження істинним, а потім визначити, чи правильно пояснено причину.

Завдання з пропуском, або клоуз-тести, – різновиди завдань із вибором однієї правильної відповіді, у яких на місці ключового слова або словосполучення в основі завдання стоїть пропуск, позначений рисою («__») або крапками («...»). Учень має обрати одну правильну відповідь серед кількох запропонованих.

Серед переваг завдань з однією правильною відповіддю тестологи визначають такі: оцінювання результатів швидке, легке, об'єктивне, надійне; завдання добре структуровані й чіткі; дають можливість вимірювати результати навчання від простих до складних; неправильні варіанти відповідей дають діагностичну інформацію про рівень сформованості знань і вмінь; результати менше залежать від угадування порівняно із завданнями з альтернативними відповідями. Недоліки: складання запитань потребує багато часу; часто складно дібрати правдоподібні дистрактори.

Завдання з кількома правильними відповідями передбачають вибір кількох правильних відповідей серед низки запропонованих. Розрізняють два різновиди завдань такого типу: із заздалегідь указаною кількістю правильних відповідей і з незазначеною кількістю правильних відповідей. При цьому «в інструкції треба обов'язково написати, скільки правильних відповідей має завдання, або зазначити, що відповідей може бути кілька» (Короткий тестологічний словник-довідник, 2008, с. 40).

II. *Завдання на встановлення відповідності* (утворення логічних пар) складаються з основи та низки альтернатив – варіантів відповідей – розміщених у дві колонки. «Одна колонка (ліворуч) – це перелік

вихідних умов (слів, словосполучень, речень, дат, формул, термінів тощо), до яких тестований має відшукати відповідну відповідь у другій колонці (праворуч), яку називають списком відповідей. Тестований має порівняти матеріал лівої й правої колонок і утворити правильні логічні пари» (Короткий тестологічний словник-довідник, 2008, с. 43). Бажано, щоб кількість варіантів відповідей правої колонки була на 1–2 більше, ніж умов лівої колонки.

Серед переваг таких завдань визначають: оцінювання результатів легке, об'єктивне та надійне; запитання мають компактную форму, оскільки набір відповідей стосується тієї самої основи; щоб прочитати завдання та обрати варіанти відповідей, учень витрачає небагато часу. До недоліків відносять той факт, що складання завдань, які містять достатню кількість однорідних відповідей, є досить складним процесом (Короткий тестологічний словник-довідник, 2008, с. 44).

III. *Завдання на встановлення правильної послідовності* передбачають, що учень повинен встановити правильну послідовність елементів, дій, подій, операцій тощо.

Переваги: можна перевірити сформованість алгоритмічного мислення; читання варіантів відповідей потребує небагато часу; завдання добре структуровані й чіткі; результати менше залежать від угадування, ніж у завданнях з однією правильною відповіддю.

Серед недоліків завдань закритої форми переважно визначають ефект угадування (Паращенко & Леонський & Леонська, 2006, с. 90), а також обмеження творчої діяльності учнів (Короткий тестологічний словник-довідник, 2008, с. 45).

Незважаючи на певні недоліки закритих завдань, саме вони набули найбільшого поширення в практиці тестування. Такі завдання найчастіше використовуються при проведенні підсумкового контролю, оскільки його можна провести швидко, підсумкові бали учнів просто обчислюються, з допомогою їх можна повніше охопити зміст навчального матеріалу, який перевіряється, підвищується обґрунтованість підсумкових оцінок учнів (Паращенко & Леонський & Леонська, 2006, с. 90).

Завдання відкритого типу – їх також називають «завдання на надання відповіді», «завдання з конструйованою відповіддю», «завдання з вільною відповіддю» тощо (Паращенко & Леонський & Леонська, 2006; Балыхина, 2006) – не містять указаних можливих варіантів відповідей і «вимагають від тестованого самостійно сформулювати правильну відповідь на поставлене запитання та відповідно її оформити (записати одним словом, цифрою, буквою, словосполученням або підготувати розгорнуту відповідь)» (Короткий тестологічний словник-довідник, 2008, с. 48).

До завдань відкритого типу дослідники відносять завдання з короткою і розгорнутою відповіддю.

Завдання з короткою відповіддю передбачають стислу відповідь, сформульовану словом, словосполученням, цифрою чи символом. Така форма завдань вимагає, щоб учень самостійно дав відповідь, а не обрав її із запропонованих варіантів.

Перевагами цих завдань визначають такі: відповідь на запитання формулюється коротко й однозначно; завдання такої форми легко складати; угадування відповіді менш імовірне, ніж у завданнях на вибір відповіді; простота перевірки. Недоліки: перевірка завдань певною мірою є суб'єктивною; на оцінювання результатів може впливати рівень грамотності учня, якщо відповідь треба написати словами.

Завдання з розгорнутою відповіддю передбачають побудову висловлювання. Відповіді в таких завданнях можуть варіюватися від стислої й точної, сформульованої кількома реченнями, до розгорнутої, довільної форми та обсягу.

Переваги: оцінюють результати навчання найвищого рівня (уміння аналізувати, синтезувати, оцінювати); підготовка завдань потребує менше часу порівняно із завданнями закритої форми. Недоліки: оцінювання результатів є трудомістким і потребує багато часу; оцінювання таких завдань є найбільш суб'єктивним порівняно з іншими формами завдань і часто знижує надійність тесту; складно розрахувати час, необхідний для виконання завдання (Короткий тестологічний словник-довідник, 2008, с. 37–38).

Завдання відкритого типу є складнішими, оскільки в них виключається можливість угадування. Загалом, як зазначає Л. Парашенко (Парашенко & Леонський & Леонська, 2006, с. 89), вибір форми завдання залежить від змісту навчального матеріалу та від цілей контролю.

Для того, щоб за допомогою тесту правильно та об'єктивно оцінювати навчальні досягнення учнів, він як інструмент вимірювання повинен бути якісно створеним. Якість, або ефективність тесту, розуміється як «ступінь вимірної «сили» тесту й визначається, наскільки ефективно й точно, незалежно від обставин, тест діагностує учня» (Рапопорт & Сельг & Соттер, 1987, с. 139). Аналіз літератури з тестування свідчить, що основними критеріями якості тесту є його валідність, надійність, репрезентативність, практичність, або технологічність, та економічність.

Валідність (з англ. *valid* – *придатний*) у тестології визначається як спрямованість тесту на те, що він повинен виміряти. Як зазначають учені (Рапопорт І., Сельг Р. та Соттер І.), «валідність показує, що вимірює тест і наскільки добре він це робить. Іншими словами, ця характеристика тесту показує більший чи менший ступінь придатності тесту для його використання з певною метою» (Рапопорт & Сельг & Соттер, 1987, с. 141). Валідність співвідноситься з адекватністю тесту меті вимірювання. Вона перевіряється шляхом порівняння результатів

використання розробленого тесту з результатами перевірки рівня навченості учнів іншими методами (усне опитування, традиційна контрольна робота), зіставлення одержаних результатів з поточною успішністю учнів. Таким чином, валідність – один із основних критеріїв якості тесту, спрямований на об'єктивне вимірювання результатів навчально-пізнавальної діяльності учнів.

Надійність, як і валідність, – важливий критерій якісно створеного тесту. Надійність тесту розглядається як точність і усталеність результатів вимірювання при повторних проведеннях (Паращенко & Леонський & Леонська, 2006, с. 41; Балыхина, 2006, с. 50). В. Аванесов визначає надійність тесту як «узгодженість показників, одержаних у тих самих учнів за умови повторного тестування, тим самим тестом, або еквівалентною його формою» (Аванесов, 1989, с. 35). «Надійність визначається стійкістю, з якою тест виконує свою функцію вимірного інструменту» (Рапопорт & Сельг & Соттер, 1987, с. 157). Надійність тесту є одним із критеріїв його якості й показує, наскільки точно даний тест вимірює вивчене явище. Таким чином, поняття «надійність» має два значення: 1) якщо тест надійний, то при повторному виконанні він дає подібні результати; 2) забезпечує високу точність вимірювання.

Виокремлюють дві основні групи факторів, які впливають на точність педагогічних оцінок: 1) випадкові фактори – психологічні та фізіологічні, вплив яких на результати передбачити неможливо; 2) не випадкові фактори – відображають особливості методів і засобів вимірювання, обраних для контролю (Паращенко & Леонський & Леонська, 2006, с. 41).

Критерію надійності тестів важливе значення надавали такі тестологи, як Л. Гутман, Г. Кюдер, Ч. Спірмен, М. Ричардсон, Ф. Рулон, Дж. Стенлі та ін. Ця якість тесту визначається на основі математико-статистичного аналізу результатів його виконання. Водночас, слід зважати на те, що дія зазначених вище факторів може призвести до похибки вимірювання. Тому деякі дослідники зазначають, що «результат експерименту завжди містить помилку як би якісно не проводилося вимірювання» (Яноши, 1965, с. 15). Так само, на думку В. Свириденко (Свириденко, 1972, с. 24), неможливо знати абсолютне значення величини, що вимірюється. Як зазначає Л. Паращенко, «завдяки вдосконаленню тесту ця похибка в оцінюванні учнів буде зменшуватись, а надійність обраного засобу вимірювання – збільшуватись. Тобто, чим надійнішим є тест, тим об'єктивніше він відображатиме дійсний рівень підготовки учнів» (Паращенко & Леонський & Леонська, 2006, с. 41).

У тестології (Короткий тестологічний словник-довідник, 2008, с. 77; Лаврова, 1999, с. 47; Балыхина, 2006, с. 37–38) існує кілька методик визначення надійності тесту:

1) метод «пара – непара» для визначення коефіцієнта внутрішньої стійкості тесту як показника кореляції між результатами двох послідовних тестувань одного й того самого контингенту тестованих одним і тим самим тестом;

2) метод визначення внутрішньої узгодженості тесту за формулою Кюдера – Ричардсона. Розраховується як коефіцієнт кореляції між результатами відповідей тестованих на парні й непарні завдання одного й того самого тесту;

3) метод встановлення ретестової кореляції, або метод повторного тестування. Він полягає в тому, що один і той самий тест використовують кілька разів в одній групі тестованих.

Репрезентативність перевірних завдань передбачає повноту охоплення завданнями всього навчального матеріалу, або основних тем, чи найбільш значущого матеріалу.

Практичність, або технологічність, тесту полягає в доступності інструкцій і змісту завдань і передбачає простоту як організації тестування, так і перевірки робіт, підрахунку результатів, підведення підсумків (Балыхина, 2006, с. 65).

Економічність полягає в можливості за порівняно короткий час здійснити перевірку досить значного обсягу знань, умінь і навичок учнів. Основними параметрами економічності визначено: 1) компактність матеріалу, що перевіряється; 2) раціональне розміщення тестових завдань у тесті; 3) час, необхідний для виконання тесту, і час, що витрачається на перевірку результатів (Лаврова, 1999, с. 52).

Під час створення тесту важливо враховувати ступінь складності завдань. *Складність завдання* – це його здатність «оцінити належний рівень знань і навичок при допустимому рівні (частці) помилок. Складність завдання залежить від відсотка помилкових відповідей, допущених тестованими під час його виконання» (Короткий тестологічний словник-довідник, 2008, с. 114). Її визначають у відсотках ділення кількості завдань з помилкою на загальну кількість запропонованих завдань, і вона може варіюватися від 1 до 100 %. Деякі вчені (Короткий тестологічний словник-довідник, 2008, с. 114) стверджують, що для нормативно зорієнтованого тесту оптимальними є показники від 40 % до 70 %, оскільки це дає можливість підвищити надійність тесту. У табл. 1 наведено характеристику тестових завдань за коефіцієнтом складності.

Таблиця 1

Характеристика тестових завдань за коефіцієнтом складності

Ступінь складності тесту	Інтервали індексу складності
Дуже легкі	Понад 80
Легкі	60 – 79
Оптимальні	40 – 59
Складні	20 – 39
Дуже складні	19 і нижче

Існує й інша думка щодо складності завдань. Так, окремі вчені (Harris, 1970, р. 212–213) вважають, що тест повинен складатися із завдань, ступінь складності яких перебуває в діапазоні 0,3–0,9. У сучасній тестовій теорії (Item Response Theory) уже розроблені моделі, які визначають складність завдання незалежно від специфічної вибірки тестованих. У такій моделі коефіцієнт складності завдання позначається числом p (інша назва – індекс складності P_{diff} або P -величина) і обчислюється як частка завдань, що виконані правильно від загальної кількості завдань. Якщо складне завдання правильно виконали лише 10 % тестованих, то матимемо число p 0,1 (або 10 %); легке завдання, яке виконали 80 % учнів, матиме число p 0,80 (Короткий тестологічний словник-довідник, 2008, с. 140). Таким чином, «складність тестового завдання визначає, наскільки конкретне тестове завдання було складним для осіб, які тестувалися» (Булах & Мруга, 2006, с. 24). Уважається, що чим вищим є цей показник, тим легшим є тестове завдання: якщо 100 % учнів відповіли правильно, то отримуємо показник 1. Залежно від значення складності тестове завдання кваліфікується таким чином:

- $1,20 < P_{diff} \leq 0,36$ – надто складне;
- $0,36 < P_{diff} \leq 0,84$ – середньої складності;
- $0,84 < P_{diff}$ – надто легке.

За цим показником не складно диференціювати завдання – легке, середньої складності, складне або відгадане (Булах & Мруга, 2006, с. 24). Складні й надто легкі завдання виключаються з тесту, оскільки вони не дають об'єктивної інформації про рівень сформованості знань, умінь і навичок учнів.

Як наголошують дослідники (Белоус & Куцевич & Титаренко & Стужко, 2011, с. 10), для визначення якості тестів необхідно за результатами попереднього тестування виокремити завдання, які не працюють, і завдання низької якості. До завдань, які не працюють, належать надто легкі завдання, на які дали правильну відповідь усі учні, або надто складні, на які ніхто не дав відповіді, а також завдання, що не відповідають поставленій меті. До завдань низької якості належать завдання, що потребують корегування з причин неточності формулювання.

Деякі дослідники (Гуртовий & Пандюк, 2011, с. 24) вважають, що дані порівняльного аналізу методів вимірювання рівня знань, які використовуються в сучасній педагогіці: усного опитування, письмової роботи, інтерв'ю та тестування, свідчать про те, що останній метод найбільше відповідає критеріям якості при визначенні рівня теоретичних знань, умінь і навичок учнів.

Серед *переваг тестування* перед іншими методами контролю можна визначити такі:

- тестування – більш якісний та об'єктивний метод вимірювання, оскільки він виключає суб'єктивність в оцінці, дозволяє зробити процедуру вимірювання однаковою для всіх учнів;

- вища, на відміну від традиційних методів контролю, ефективність, оскільки його можна одночасно проводити серед великої кількості учнів, витрачаючи при цьому мінімум навчального часу;

- швидша обробка результатів тестування, ніж, наприклад, перевірка письмової контрольної роботи;

- можливість використання тестів для самоконтролю.

Водночас можна виокремити й низку *недоліків* використання тестового контролю рівня сформованості знань, умінь і навичок учнів. Серед них висока трудомісткість розробки якісних тестів, можливість угадування відповіді в тестах із закритими тестовими завданнями, що не дає повної гарантії наявності в учня певних знань, умінь і навичок. Окрім того, дослідники (Бобовський, 2010, с. 49) відзначають ще й такий недолік тестових завдань, як неможливість перевірки за допомогою тестів умінь наводити власні приклади, логічно й обґрунтовано висловлювати думки, аналізувати та узагальнювати тощо. Однак, наявність низки недоліків не заперечує доцільності використання тестування як методу вимірювання для об'єктивного контролю знань, умінь і навичок учнів.

Особливе місце в тестуванні належить оцінюванню одержаних результатів. На думку дослідників, оцінювання на основі враження від відповіді учня, тобто без використання спеціальних шкал, породжує суб'єктивізм.

Під *шкалою оцінювання* розуміють числову систему, яка визначена залежно від прийнятої норми й дає можливість перетворити якісні показники (наприклад, «правильно», «неправильно») у кількісні (Лаврова, 1999, с. 35). «Шкала – засіб фіксації результатів вимірювання властивостей об'єктів шляхом впорядкування їх у певну систему чисел, у якій відношення між окремими результатами виражене у відповідних числах» (Паращенко & Леонський & Леонська, 2006, с. 77). У літературі розрізняють такі шкали вимірювання: нормативна, порядкова та рейтингова.

Нормативна (або номінальна) шкала побудована на дихотомічному принципі «правильно-неправильно». Цією шкалою вимірюються результати виконання тестових завдань альтернативного типу, множинного вибору, завдань на встановлення відповідностей, послідовності, виключення зайвого елемента, завершення (Балыхина, 2006, с. 111–112).

Механізм використання нормативної шкали В. Аванесов пояснює таким чином: «Якщо тестований обирає в завданні правильну відповідь, цей результат можна позначити одним балом, а неправильну відповідь – нулем. Чим більше буде в тесті балів, тим впевненіше можна судити про

наявність <...> знань» (Аванесов, 1996, с. 63–64). Таким чином, оцінка за весь тест буде складатися з суми балів, одержаних за кожну правильну відповідь. Водночас у завданнях із множинним вибором або альтернативного типу, коли учням пропонуються варіанти відповідей, не виключається вгадування правильної відповіді. Щоб правильно оцінити результати тесту, В. Аванесов (Аванесов, 1996, с. 42) радить урахувати т. зв. скорегований бал, який визначається за допомогою використання коефіцієнта вгадування, або формули корекції:

$$X = R - W / (K - 1), \text{ де}$$

R – тестовий бал до корекції; W – бали, одержані за неправильні відповіді (тобто кількість неправильних відповідей); K – кількість варіантів, які були запропоновані для вибору відповіді в завданні.

З метою попередження вгадування тестологи рекомендують збільшувати кількість дистракторів до 4–5. Н. Лаврова (Лаврова, 1999, с. 36) зазначає, що із збільшенням варіантів відповіді вірогідність угадування зменшується (більше 5 варіантів достатньо для того, щоб зняти коефіцієнт угадування). Крім того, ще одним дієвим методом попередження вгадування є зниження оцінки за неправильну відповідь (Лаврова, 1999, с. 37). Так, за неправильну відповідь учень одержує не нуль балів, а мінус один, за відсутність відповіді – нуль, за правильну – один бал. Відповідно, у разі сумніву в правильності відповіді, тестованому краще пропустити дане завдання, ніж позначати відповідь навмання. Крім зазначених методик попередження вгадування, існує т. зв. розрахунок за формулою Раша (Rasch, 1980). Однак, слід зазначити, що ця процедура достатньо складна й використовується при автоматичній системі обробки результатів.

Порядкова шкала – це шкала, яка має більше ніж 2 поділки. Останніх може бути стільки, скільки важливих параметрів визначено тестологом для оцінки результату виконання завдань. Кількість поділок збільшується за аналогією з порядковими числами, звідси й назва шкали (Лаврова, 1999, с. 37). «За менш правильну відповідь можна дати один бал, за правильну відповідь – два бали, за найбільш правильну відповідь – три бали» (Аванесов, 1996, с. 64). На нашу думку, терміни «більш чи менш правильна відповідь» потребують уточнення. Так, доцільніше вживати поняття «неправильна відповідь», «правильна, але неповна відповідь», «правильна повна відповідь».

Порядкова шкала може застосовуватись при оцінюванні результатів виконання завдань типу клоуз (коли необхідно встановити відсутні елементи тесту), завдань на розкриття понять, узагальнення, завершення (Лаврова, 1999, с. 38).

Крім нормативної та порядкової шкал, для оцінювання тестових завдань використовується *рейтингова шкала* оцінювання. Вона будується як «шкала для вимірювання відношень між об'єктами, що вивчаються, на основі експертних оцінок» (Психологія, 1990, с. 338).

Рейтингова шкала відрізняється від порядкової тим, що всередині кожної поділки можливий певний інтервал (т. зв. «люфт») балів. Оцінювання за допомогою такого типу шкали здійснюється, як правило, кількома експертами (Лаврова, 1999, с. 38–39).

Таким чином, вважається, що тестування є одним із об'єктивних методів вимірювання результатів навчально-пізнавальної діяльності учнів. У літературі достатньо висвітлено його переваги й недоліки порівняно з іншими методами контролю. Надійним засобом вимірювання визнано тест. У тестології визначено вимоги до завдань у тестовій формі, їх класифікацію, переваги й недоліки. Тестологами досить повно розглянуто такі критерії якості тесту, як валідність, надійність, репрезентативність, практичність, економічність. З метою запобігання суб'єктивізму в оцінюванні результатів навчально-пізнавальної діяльності учнів дослідниками пропонується використання спеціальних шкал.

Список використаної літератури

- 1. Паращенко Л. І.,** Леонський В. Д., Леонська Г. І. Тестові технології у навчальному закладі: метод. посібник; наук. ред. О. І. Ляшенко. Київ: ТОВ «Майстерня книги», 2006. 217 с.
- 2. Аванесов В. С.** Основы научной организации педагогического контроля в высшей школе: учебное пособие. Москва: Исследовательский центр, 1989. 167 с.
- 3. Майоров А. Н.** Теория и практика создания тестов для системы образования. (Как выбирать, создавать и использовать тесты для целей образования). Москва, 2000. 352 с.
- 4. Чельшкова М. Б.** Теория и практика конструирования педагогических тестов: учебное пособие. Москва: Логос, 2002. 432 с.
- 5. Бондар В. І.** Дидактика. Київ: Либідь, 2005. 264 с.
- 6. Ляшенко О. І.,** Лукіна Т. О., Жук Ю. О., Ващенко Л. С., Науменко С. О., Гривко А. В. Тестові технології оцінювання ключових і предметних компетентностей учнів основної і старшої школи: монографія / за ред. Ляшенко О. І., Жука Ю. О. Київ: Педагогічна думка, 2014. 200 с.
- 7. Короткий** тестологічний словник-довідник. Київ: Грамота, 2008. 160 с.
- 8. Лаврова Н. В.** Тестирование как средство измерения сформированности коммуникативной компетенции в русском языке как иностранном в сфере профессионального общения: дисс. ... канд. пед. наук: 13.00.02. Москва, 1999. 206 с.
- 9. Булах І. Є.** Комп'ютерна діагностика навчальної успішності. Київ: ЦМК МОЗ України, УДМУ. 1995. 221 с.
- 10. Бобовський Р. П.** Контроль та оцінювання якості самостійної роботи студентів вищих навчальних закладів. *Науковий часопис Національного педагогічного університету імені М. П. Драгоманова. Серія № 5. Педагогічні науки: реалії та перспективи* / за ред. В. П. Сергієнка. Вип. 22. Київ: Вид-во НПУ імені М. П. Драгоманова, 2010. С. 47–50.
- 11. Балыхина Т. М.** Словарь

терминов и понятий тестологии. 2-е изд. стер. Москва : Рус. яз., 2006. 160 с. **12. Рапопорт И.,** Сельг Р., Соттер И. Тесты в обучении иностранным языкам в средней школе : пособие для учителей. Таллин : Валгус, 1987. **13. Яноши Л.** Теория и практика обработки результатов измерений. Москва : Мир, 1965. **14. Свириденко В. М.** О гносеологической природе постулата неизбежности погрешности измерения. *Вопросы философии*. 1972. № 6. С. 20–26. **15. Harris D-P.** Report on an Experimental Group-Administred Memory. Spantest. *TESOL quarterly*. 1970. Vol. 4, № 3. pp. 210–213. **16. Булах І. Є.,** Мруга М. Р. Створюємо якісний тест : навч. посіб. Київ : Майстер-клас, 2006. 160 с. **17. Белоус И. А.,** Куцевич И. В., Титаренко А., Стужко С. Н. Структурная схема подсистемы определения качества тестового материала и распределения заданий по уровням сложности. *Підготовка молодих викладачів та аспірантів в галузі освітніх вимірювань : тези Міжнародного семінару молодих науковців*. Кіровоград : РВВ КДПУ ім. В. Винниченка, 2011. С. 10–13. **18. Гуртовий Ю. В.,** Пандюк С. С. Тестування під час вивчення математичного аналізу. *Підготовка молодих викладачів та аспірантів в галузі освітніх вимірювань : тези Міжнародного семінару молодих науковців*. Кіровоград : РВВ КДПУ ім. В. Винниченка, 2011. С. 23–25. **19. Аванесов В. С.** Композиция тестовых заданий. Москва : Ассоциация инженеров-педагогов г. Москвы, 1996. 191 с. **20. Rasch G.** Probabilistic Models for Some Intelligence and Attainment Test. With a foreword and Afterword by B. D. Wright. *The University of Chicago Presse*. Chicago & London, 1980. 199 p. **21. Психология :** словарь / под общ. ред. Петровского А. В., Ярошевского М. Г. 2-е изд. Москва : Политиздат, 1990. 494 с.

References

1. Parashchenko, L. I., Leonskyi, V. D., & Leonska, H. I. (2006). Testovi tekhnolohii u navchalnomu zakladi [Test technologies in an educational institution]. Kyiv: TOV «Maisternia knyhy» [in Ukrainian]. **2. Avanesov, V. S.** (1989). Osnovy nauchnoj organizacii pedagogicheskogo kontrolya v vysshej shkole [Fundamentals of scientific organization of pedagogical control in higher education]. Moscow: Issledovatel'skij centr [in Russian]. **3. Majorov, A. N.** (2000). Teoriya i praktika sozdaniya testov dlya sistemy obrazovaniya. (Kak vybirat', sozdavat' i ispol'zovat' testy dlya celej obrazovaniya) [Theory and practice of creating tests for the education system. (How to select, create and use tests for educational purposes)]. Moscow [in Russian]. **4. Chelyshkova, M. B.** (2002). Teoriya i praktika konstruirovaniya pedagogicheskikh testov [Theory and practice of designing pedagogical tests]. Moscow: Logos [in Russian]. **5. Bondar, V. I.** (2005). Dydaktyka [Didactic]. Kyiv: Lybid [in Ukrainian]. **6. Liashenko, O. I.,** Lukina, T. O., Zhuk, Yu. O., Vashchenko, L. S., Naumenko, S. O., & Hryvko, A. V. (2014). Testovi tekhnolohii otsiniuvannia kliuchovykh i

predmetnykh kompetentnostei uchniv osnovnoi i starshoi shkoly [Test technologies for assessing the key and subject competencies of primary and secondary school students]. Kyiv: Pedahohichna dumka [in Ukrainian].

7. Korotkyi testolohichnyi slovnyk-dovidnyk [Short testological dictionary-reference book]. (2008). Kyiv: Hramota [in Ukrainian].

8. Lavrova, N. V. (1999). Testirovanie kak sredstvo izmereniya sformirovannosti kommunikativnoj kompetencii v russkom yazyke kak inostrannom v sfere professional'nogo obshcheniya [Testing as a means of measuring the formation of communicative competence in Russian as a foreign language in the field of professional communication]. *Candidate's thesis*. Moscow [in Russian].

9. Bulakh, I. Ye. (1995). Kompiuterna diahnostyka navchalnoi uspishnosti [Computer diagnostics of academic performance]. Kyiv: TsMK MOZ Ukrainy, UDMU [in Ukrainian].

10. Bobovskyi, R. P. (2010). Kontrol ta otsiniuvannia yakosti samostiinoi roboty studentiv vyshchych navchalnykh zakladiv [Control and evaluation of the quality of independent work of students of higher educational institutions]. *Naukovyi chasopys Natsionalnoho pedahohichnoho universytetu imeni M. P. Drahomanova. Seriya № 5. Pedahohichni nauky: realii ta perspektyvy – Scientific journal of the National Pedagogical University named by M. P. Drahomanov, Series 5. Pedahohichni nauky: realii ta perspektyvy*, 22, 47-50. Kyiv: Vyd-vo NPU imeni M. P. Drahomanova [in Ukrainian].

11. Balyhina, T. M. (2006). Slovar' terminov i ponyatij testologii [Glossary of terms and concepts of testology]. Moscow: Rus. yaz. [in Russian].

12. Rapoport, I., Sel'g, R., & Sotter, I. (1987). Testy v obuchenii inostrannym yazykam v srednej shkole [Tests in teaching foreign languages in high school]. Tallin: Valgus [in Russian].

13. Yanoshi, L. (1965). Teoriya i praktika obrabotki rezul'tatov izmerenij [Theory and practice of processing measurement results]. Moscow: Mir [in Russian].

14. Sviridenko, V. M. (1972). O gnoseologicheskoy prirode postulata neizbezhnosti pogreshnosti izmereniya [On the epistemological nature of the postulate of the inevitability of the measurement error]. *Voprosy filosofii – Questions of philosophy*, 6, 20-26 [in Russian].

15. Harris, D-P. (1970). Report on an Experimental Group-Administred Memory. *Spantest. TESOL quarterly*, 4, 3, 210-213 [in English].

16. Bulakh, I. Ye., & Mruha, M. R. (2006). Stvoriuiemo yakisnyi test [We create a quality test]. Kyiv: Maister-klas [in Ukrainian].

17. Belous, I. A., Kucevich, I. V., Titarenko, A., & Stuzhko, S. N. (2011). Strukturnaya skhema podsistemy opredeleniya kachestva testovogo materiala i raspredeleniya zadanij po urovnyam slozhnosti [Block diagram of the subsystem for determining the quality of the test material and the distribution of tasks by difficulty levels]. *Pidhotovka molodykh vykladachiv ta aspirantiv v haluzi osvitnikh vymyruvan – Training of yo ung teachers and graduate students in the field of educational measurements*. (pp. 10-13). Kirovohrad: RVV KDPU im. V. Vynnychenka [in Russian].

18. Hurtovyi, Yu. V., & Pandiuk S. S. (2011). Testuvannia pid chas vyvchennia matematychnoho analizu [Testing

while studying mathematical analysis]. *Pidhotovka molodykh vykladachiv ta aspirantiv v haluzi osvitykh vymiriuvan – Training of young teachers and graduate students in the field of educational measurements*. (pp. 23-25). Kirovohrad: RVV KDPU im. V. Vynnychenka [in Ukrainian].

19. Avanesov, V. S. (1996). *Kompoziciya testovyh zadaniy* [Composition of test items]. Moscow: Associaciya inzhenerov-pedagogov g. Moskvy [in Russian].

20. Rasch, G. (1980). *Probabilistic Models for Some Intelligence and Attainment Test*. With a foreword and Afterword by B. D. Wright. *The University of Chicago Presse*. Chicago & London [in English].

21. Petrovskij, A. V., & Yaroshevskij, M. G. (Eds.). (1990). *Psihologiya: slovar'* [Psychology: vocabulary]. 2-e izd. Moscow: Politizdat [in Russian].

Хорошковська-Носач Т. П. Тестування як один із методів педагогічного вимірювання

У статті актуалізовано відомості про тестування як один із об'єктивних методів педагогічного вимірювання. Автором проаналізовано різні погляди зарубіжних і вітчизняних дослідників на поняття «тест». Розкрито основні вимоги до структурування тесту, а також до завдань у тестовій формі. Детально розглянуто класифікацію тестових завдань, зазначено переваги та недоліки кожного їх виду. У роботі проаналізовано критерії якості тесту, а також висвітлено переваги та недоліки тестування порівняно з іншими методами контролю результатів навчально-пізнавальної діяльності учнів. Зроблено висновок про доцільність використання тестування як об'єктивного методу вимірювання знань, умінь і навичок. Зазначено, що з метою запобігання суб'єктивізму в оцінюванні результатів навчально-пізнавальної діяльності учнів використовуються спеціальні шкали вимірювання, а саме: нормативна, порядкова та рейтингова.

Ключові слова: тестування, тест, завдання, метод, вимірювання, шкала.

Хорошковская-Носач Т. П. Тестирование как один из методов педагогического измерения

В статье актуализировано сведения о тестировании как одному из объективных методов педагогического измерения. Автором проанализированы различные взгляды зарубежных и отечественных исследователей на понятие «тест». Раскрыты основные требования к структурированию теста, а также к заданиям в тестовой форме. Подробно рассмотрена классификация тестовых заданий, указаны преимущества и недостатки каждого их вида. В работе проанализированы критерии качества теста, а также освещены преимущества и недостатки тестирования по сравнению с другими методами контроля результатов учебно-познавательной деятельности учащихся. Сделан вывод о целесообразности использования

тестирования как объективного метода измерения знаний, умений и навыков. Отмечено, что с целью предотвращения субъективизма в оценке результатов учебно-познавательной деятельности учащихся используются специальные шкалы измерения, а именно: нормативная, порядковая и рейтинговая.

Ключевые слова: тестирование, тест, задания, метод, измерения, шкала.

Khoroshkovska-Nosach T. Testing as one of the methods of pedagogical measurement

The article updates information about testing as one of the objective methods of pedagogical measurement. The author analyzes different views of foreign and domestic researchers on the concept of "testing". It was found that it is considered as a method, tool, measuring instrument, task, set of questions, set of test tasks, concluded under certain rules, and which differs from the examination work in terms of involving measurement. The article reveals basic requirements for structuring the test, as well as for tasks in test form. The author concludes that all tasks in test form may be divided in two types: open and closed, which in turn have certain modifications; the features of each type of test tasks are described in detail, their advantages and disadvantages are indicated. The article analyzes the quality criteria of the test, namely: its validity, reliability, representativeness, efficiency, or processability, cost-effectiveness, and complexity of tasks. In addition, the article highlights the advantages and disadvantages of testing comparing with other methods of monitoring the results of educational and cognitive activities of students. It is concluded that the use of testing as an objective method of measuring knowledge, skills and abilities is expedient. It is noted that in order to prevent subjectivism in evaluation of the results of educational and cognitive activities of students, special measurement scales should be used, namely: normative, ordinal and rating; features of application of each of these scales are being considered.

Key words: testing, test, task, method, measurement, scale.

Стаття надійшла до редакції 09.11.2021 р.

Прийнято до друку 26.11.2021 р.

Рецензент – д. п. н., проф. Караман О. Л.