

ФОРМАТ ТЕСТУ З МАТЕМАТИКИ ЯК ФАКТОР ВПЛИВУ НА ЯКІСТЬ ОЦІНЮВАННЯ УЧНІВ

MATHEMATICS TEST FORMAT AS A FACTOR OF INFLUENCE ON STUDENT ASSESSMENT QUALITY

Підготовка учнів до обов'язкової державної підсумкової атестації з математики для учнів 11-х класів у формі зовнішнього незалежного оцінювання та державної підсумкової атестації з математики для учнів 9-х класів за «Орієнтовними вимогами до змісту атестаційних робіт» спонукає вчителів до використання в навчальному процесі тестів. Зазвичай це авторські, нестандартизовані тести для тематичного й підсумкового оцінювання. Аналіз змісту чинних підручників із математики для базової середньої школи свідчить про те, що для самостійної роботи учнів із підготовки до різних видів контролю знань автори підручників віддають перевагу коротким тематичним тестам, укладеним лише з тестових завдань із вибором однієї правильної відповіді із запропонованих чотирьох її варіантів чи з домінуючою часткою тестових завдань цієї форми.

З метою виявлення впливу формату тесту для підсумкового оцінювання з математики на результати учасників тестування проведено експериментальне дослідження, у якому взяли участь 198 учнів.

Автор розкриває зміст поняття «формат тесту». Виокремлює три формати тесту: відкритий, змішаний (комбінований) та закритий. Вказує на те, що формат будь-якого тесту детально описує його специфікація.

Об'єктивність отриманих результатів тестованих різного рівня навчальних досягнень із математики забезпечена якістю розробленого тесту та процедурами тестування.

Встановлено за рівних умов таке:

1. Учні, які виконували завдання тесту закритого формату, не продемонстрували статистично значимо вищих результатів на кожному з рівнів навчальних досягнень, ніж учні, які виконували завдання тесту змішаного формату.

2. Учні, які виконували завдання тесту закритого формату, продемонстрували статистично значимо вищий результат, ніж учні, які розв'язували відповідні завдання тесту відкритого формату і на середньому рівні навчальних досягнень ($p \leq 0,01$), і на достатньому рівні навчальних досягнень ($p \leq 0,05$), і на високому рівні навчальних досягнень ($p \leq 0,05$).

Автор вказує, що обмеження (можливі обмеження) змістового та когнітивного доменів підсумкових тестів із математики закритого й змішаного форматів також впливають на якість оцінювання.

Ключові слова: тест, формат тесту, оцінювання, підсумкове оцінювання, якість

оцінювання, базова середня освіта, математика.

Preparation of students for the obligatory state final attestation in mathematics for 11th grade students in the form of external independent assessment and state final attestation in mathematics for 9th grade students according to the "Approximate requirements for the content of attestation works" encourages teachers to use tests in the educational process. Typically, those are author's, non-standardized tests for thematic and final assessments. Analysis of the content of existing textbooks in mathematics for basic secondary school shows that for student's independent preparatory work for different types of knowledge control, the authors of the textbooks prefer short thematic tests, concluded only from test tasks with the choice of one correct answer from the four proposed variants or with the dominant share of test tasks of this form.

To identify the influence of the test format for the final assessment in mathematics on the results of test participants, an experimental study was conducted in which 198 students participated. The author reveals the meaning of the term "test format". There are three test formats: open, mixed (combined) and closed. It indicates that the format of any test is described in detail by its specification.

Objectivity of the obtained test results of different levels academic achievement in mathematics is ensured by the quality of the developed test and testing procedures.

Under equal conditions, the following was established:

1. Students who performed the tasks of the closed format test did not demonstrate statistically significantly higher results at each level of academic achievement than students who performed the tasks of the mixed format test.

2. Students who performed the tasks of the closed format test, demonstrated a statistically significantly higher result than students who solved the relevant tasks of the open format test and at the average level of academic achievement ($p \leq 0.01$), and at a sufficient level of educational achievements ($p \leq 0.05$), and at a high level of educational achievements ($p \leq 0.05$).

The author points out that the limitations (possible limitations) of the content and cognitive domains of the final tests in mathematics of closed and mixed formats also affect the quality of assessment.

Key words: test, test format, evaluation, final assessment, quality of assessment, basic secondary education, mathematics.

УДК 37.013

DOI <https://doi.org/10.32843/2663-6085/2020/30-2.2>

Дворецька Л.П.,

науковий співробітник відділу моніторингу та оцінювання якості загальної середньої освіти Інституту педагогіки Національної академії педагогічних наук України

Постановка проблеми в загальному вигляді.

Запровадження зовнішнього незалежного оцінювання (далі – ЗНО) з використанням тестових технологій дало поштовх науковим дослідженням та поширенню практичного досвіду у галузі освітніх вимірювань в Україні. Успішна реформа

вступних іспитів до закладів вищої освіти підкреслила виняткову роль стандартизованих тестів, як надійного інструменту для швидкого й якісного оцінювання навчальних досягнень учасників тестування. Державна підсумкова атестація (далі – ДПА) у формі ЗНО дала старт поширенню автор-

ських, нестандартизованих, тестів у навчальний процес закладів загальної середньої освіти. Вчителі взяли за самостійне укладання тестів для підсумкового й тематичного оцінювання учнів, стали використовувати готові, надруковані, авторські тести. Завдання у тестовій формі й тести з'явилися на сторінках шкільних підручників. Висновок науковців про те, що «для поточного контролю не обов'язково мати професійно високоякісний тест» [1, с. 19], певною мірою легітимізував використання авторських тестів для контролю знань учнів. Водночас щорічне інформування суспільства про якість використаних стандартизованих предметних тестів для системи ЗНО та якість тестових завдань [2] лише підкреслювало факт впливу якості тесту, як інструменту вимірювання, на якість оцінювання. Практики (переважно вчителі математики, експерти з розроблення тестових завдань для ЗНО з математики) раз у раз стали звертати увагу на формат тесту, як фактор впливу на якість оцінювання учнів. Перспектива створення в Україні системи зовнішнього незалежного оцінювання результатів навчання учнів за курс базової середньої освіти (2023 р.), підготовка до апробації нової моделі атестаційної роботи з математики для учнів 9-х класів, актуалізували роботу з дослідження впливу форматів тесту з математики на результати учасників тестування.

Аналіз останніх досліджень і публікацій. Двадцять років тому Є.А. Михайличев констатував, що на формування педагогічної тестології негативно впливає термінологічна анархія, та назвав одну з проблем типологізації тестів і тестових завдань – ототожнення авторами типологізацій понять «тест» і «тестове завдання» [3, с. 37–38]. За цей час відбулося певне усталення термінології, однак семантичної ідентифікації системи термінів науки у різних мовах немає.

Означимо поняття «формат тесту» як ключове поняття дослідження. *Форматом тесту* вважаємо спосіб упорядкування, спосіб внутрішньої організації тесту за використаними формами тестових завдань. Дійсно, тест, як інструмент вимірювання, має певні характеристики, детальний опис яких міститься у специфікації. Специфікація – один із основних документів *технічної* документації. За логікою, однією з таких технічних характеристик і є *формат* тесту. Нагадаємо, що формою тестових завдань (за В. Аванесовим) є спосіб організації, впорядкування й існування змісту тесту.

Виокремимо три формати тесту: закритий (далі – ЗФ), змішаний чи комбінований (далі – КФ) та відкритий (далі – ВФ). Оскільки існує різноманіття форм тестових завдань та способів їхнього впорядкування в тест, то у специфікації будь-якого тесту поряд із загальною характеристикою його формату (на кшталт, ЗФ) є і детальний опис формату.

Зазначимо, що саме тести закритого формату, які містять тестові завдання з вибором однієї правильної відповіді із запропонованих її варіантів, найбільше піддають критиці за можливість вгадування правильних відповідей. Різні формули для корегування тестових балів з урахуванням угадування на практиці не підтвердили присутній приріст надійності тестової оцінки і валідності результатів [4, с. 542].

Проведені експериментальні дослідження впливу форматів тестових завдань на результати педагогічного тестування з біології та географії здобувачів базової середньої освіти науковцями Інституту педагогіки НАПН України дозволили виявити ефект підвищеної оцінки (30 % і 24 % відповідно) тим тестованим, які виконували завдання лише з вибором однієї правильної відповіді із запропонованих її варіантів [5; 6].

Водночас Р. Лізітц та С. Гоу довели відсутність істотного впливу на результати учасників тестування, якщо замінити в тесті завдання відкритого формату на завдання закритого формату зі збереженням складності завдань [7]. Р. Трауб, досліджуючи вплив на результати тестованих різних форматів тестів (усі завдання на вибір однієї відповіді й усі завдання на продукування відповіді), зробив висновок про еквівалентність, у цілому, отриманих результатів, якщо це тест на розуміння прочитаного чи на перевірку математичних знань. Прямі порівняння завдань на вибір відповіді й продукування відповіді (традиційні завдання письмового екзамену, есе) за такими критеріями, як економічність, оперативність та прогностична валідність, проведені Г. Анастасі, Р. Беннетт, Р. Лукхеле, переконливо свідчать на користь завдань закритого формату [8; 9].

Незаперечним є той факт, що дослідження впливу форм тестових завдань і форматів тестів на результати тестованих у фокусі уваги науковців світу ще й тому, що запит на такі дослідження пов'язаний із розвитком систем стандартизованих оцінювань навчальних досягнень учнів.

Зазначимо, що у практиці підсумкового (загальнонаціонального) оцінювання з математики різних країн (державна підсумкова атестація, MATURA тощо) домінують тестові технології, різноманіття моделей, перманентне реформування і модернізація. Грузія – комп'ютерне адаптивне тестування (2013-2019 рр.). Польща і Болгарія – стандартизовані тести (змішаного формату). MALK MATURA з математики – для учнів, які є випускниками базової середньої школи. MATURA – для випускників старшої школи. Тести MALK MATURA і MATURA мають подібні формати, що обумовлено технологічними й технічними можливостями організації зовнішнього незалежного оцінювання, затвердженою нормативно-правовою базою та тяглістю однієї і

тієї самої методики підготовки учнів різного віку до екзаменів *high stakes* тощо. Естонія – традиційний екзамен із математики. Екзаменаційна робота містить кілька структурованих/неструктурованих задач з вільно конструйованими письмовими відповідями учнів, які перевіряють вчителі (комісія вчителів).

В Україні за останні двадцять років відбулися суттєві зміни моделі державної підсумкової атестації з математики для учнів 9-х класів. А саме: від унормованого Міністерством освіти і науки України одного для всіх вчителів і учнів «Збірника завдань для державної підсумкової атестації з алгебри» [10] (з певною кількістю традиційних контрольних робіт на два варіанти з 9 завдань для оцінювання учнів загальноосвітніх класів) до «Орієнтовних вимог до змісту атестаційних робіт з математики для учнів 9 класу» (2019 р.), що дають право кожному вчителю самостійно укласти чи використати готові авторські, не менше 10 варіантів, атестаційні роботи із тестових завдань різних форм, відповідно до затверджених лише загальних характеристик тестів. Тож щороку маємо безліч нестандартизованих тестів (атестаційних робіт) для проведення ДПА з математики у 9-х класах. Стандартизовані тести для ДПА з математики у 11-х класах адмініструє Український центр оцінювання якості освіти.

У вітчизняній практиці ЗНО з математики, що слугує й державною підсумковою атестацією учнів 11-х класів, використовують, згідно зі специфікацією тесту ЗНО з математики 2019 р., ТЗ з вибором однієї правильної відповіді (з п'яти її варіантів), на встановлення відповідності, відкритої форми з короткою відповіддю (структуровані й неструктуровані завдання) й відкритої форми з розгорнутою відповіддю.

Підготовка учнів 9-х та 11-х класів до ДПА у формі ЗНО з математики заохочує вчителів опрацювати з учнями тести різних форматів для тематичного й підсумкового оцінювання. Аналіз змісту підручників свідчить про те, що популярними для домашньої самостійної роботи учнів базової середньої школи є короткі тематичні тести закритого формату (до 12 завдань), укладені лише з тестових завдань з вибором однієї правильної відповіді із запропонованих чотирьох її варіантів, чи змішаного формату з домінуючою часткою тестових завдань з вибором однієї правильної відповіді [11, с. 138]. Тестові бали за виконання 12 таких завдань «легко» конвертуються у шкільні оцінки (1–12 балів) учнів.

Суспільний запит на підвищення якості оцінювання, на тлі використання тестових технологій, у проекції на навчальний процес у школі, спонукає до пошуку відповіді на запитання: «Тест якого формату найточніше оцінює рівень навчальних досягнень учня?».

Відомо, що свобода автора у виборі формату тесту почасти обмежується метою тестування, специфікою знань, умінь, навичок, компетентностей, що підлягають оцінюванню. Однак дискусійним залишається питання впливу формату тесту на результати учасників тестування.

Потребує дослідження проблема вияву впливу формату підсумкового (авторського) тесту з математики на результати учасників тестування та визначення характеру впливу на учнів із різним рівнем навчальних досягнень.

Метою статті є ознайомлення з розробленою методикою експериментального дослідження впливу формату підсумкового тесту з математики на результати учасників тестування та оприлюднення отриманих результатів дослідження. Метою цього дослідження є виявлення впливу форматів підсумкового (авторського) тесту з математики (ЗФ, КФ і ВФ) на результати учнів із різним рівнем навчальних досягнень.

Виклад основного матеріалу. *Організація дослідження.* З метою визначення впливу форматів тесту з математики на результати його виконання учнями було організовано дослідно-експериментальну роботу серед учнів сьомих класів середньої загальноосвітньої школи I–III ступенів № 12 м. Сєвєродонецька та гімназії «Міленіум» № 318 м. Києва. В експерименті брали участь 198 учнів, які виконували завдання очно, протягом одного уроку, з хорошою мотивацією (підсумкове оцінювання). Експеримент відбувався у два етапи (2018 р. та 2019 р.).

Інструментарій дослідження. Відповідно до чинної програми з математики було укладено тест з алгебри для підсумкового оцінювання навчальних досягнень учнів 7 класу. Після схвальних експертних висновків щодо якості ТЗ та апробації тесту, його було запропоновано у трьох форматах: ЗФ (закритий формат, усі ТЗ на вибір однієї правильної відповіді з чотирьох її варіантів); ВФ (відкритий формат, усі завдання відкритої форми з розгорнутою відповіддю); КФ (комбінований формат, перша половина завдань тесту – ТЗ на вибір однієї правильної відповіді з чотирьох її варіантів, друга половина – завдання відкритої форми з розгорнутою відповіддю).

Тест містить 15 завдань, які оцінюються дихотомічно. Кожне завдання оцінювалося 1 балом. Загальний бал учня визначався сумою балів за правильно виконані завдання. Кількість завдань співставна з можливостями учнів різного рівня навчальних досягнень розв'язати завдання і записати відповіді (у випадку роботи із ВФ тесту, де всі завдання потребували розв'язування, яке оцінював екзаменатор). Усім учням вистачило часу (40 хв.) на виконання завдань тесту (у різних форматах). Фактор часу досліджувався окремо.

У процесі розв'язування завдань учні повинні були продемонструвати набуття спеціальних предметних математичних компетентностей: обчислювальної, процедурної, практико-орієнтованої, графічної та логічної. Хорошу розподільну здатність тесту забезпечили, орієнтовно, 20 % легких (і дуже легких) завдань, 60 % – оптимальної складності, 20 % – складних (і дуже складних). Завдання 1, 5, 6 – легкі (і дуже легкі); 8, 12, 14 – складні (і дуже складні); усі інші завдання – оптимальної складності. Змістовий та когнітивний домен описані у специфікації тесту. Тест критеріально-орієнтований. Інструктаж перед тестуванням та однакові умови проведення тестування забезпечили отримання об'єктивних результатів тестованих різного рівня навчальних досягнень з математики.

Про порядок укладання одного й того самого змісту у тести різного формату зазначимо таке. Спочатку – ЗФ, далі – КФ і ВФ (ВФ і КФ). Рух у зворотному напрямі (від відкритого формату тесту (класичної контрольної роботи) до закритого) може негативно вплинути на точність одержаних результатів дослідження, збільшить похибку вимірювання чи призведе до хибних висновків. Перетворення будь-якої задачі з підручника на якісне тестове завдання (на вибір однієї правильної відповіді) іноді потребує лише відповідного досвіду й майстерності виконавця, а іноді є неможливим. Наведемо приклади: «Серед розв'язків рівняння $x + 3y = 20$ знайдіть пару рівних між собою чисел», «Доведіть, що при будь-якому цілому значенні a значення виразу $(a^3 - a) : 6$ є числом цілим» тощо.

Контингенти класів фрагментовані за ознакою «рівень навчальних досягнень учня», а саме: низький, середній, достатній та високий. Рівень

навчальних досягнень учня визначено вчителем математики, про що свідчить оцінка учня в класному журналі – офіційному джерелі інформації про оцінювання учнів класу. На кінець навчального року така оцінка вчителя є комплексною й базується на результатах оцінювання учня протягом тривалого часу різними способами.

Кожна вибірка учнів одного рівня навчальних досягнень розподілена на три групи. Перша група – учні, які виконували тест ЗФ, друга – учні, які виконували тест КФ, третя група – ВФ. Розподіл учнів за групами (ЗФ, КФ, ВФ) і рівнями навчальних досягнень подано у таблиці 1.

Опрацювання результатів. Порівняння середнього тестового балу результатів учнів на кожному з чотирьох рівнів навчальних досягнень за виконання тестів трьох форматів (Табл. 2) свідчить про те, що найвищі результати показали учні, які виконували завдання тесту ЗФ, а найнижчі результати в учнів, які виконували завдання ВФ (окрім учнів з низьким рівнем навчальних досягнень).

Порівнюємо результати учнів кожної з трьох груп (парами) на кожному рівні навчальних досягнень та встановимо статистичну значимість отриманих результатів за критерієм U Манна – Вітні. Цей критерій використовується для оцінки відмінностей між двома вибірками за рівнем певної ознаки, що вимірюється кількісно.

Відмінності між двома вибірками можна вважати значимими ($p < 0,05$), якщо $u_{\text{емп.}} \leq u_{0,05}$, і ще більше достовірними ($p < 0,01$), якщо $u_{\text{емп.}} \leq u_{0,01}$. Критичні значення критерію U Манна – Вітні для рівнів статистичної значимості $p \leq 0,05$ і $p \leq 0,01$ визначаємо за Таблицею II (за Є.В. Гублером і А.А. Генкіним, 1973) [12, с. 316].

Таблиця 1

Розподіл учнів за рівнями навчальних досягнень і форматами тесту, завдання якого виконували

Рівень навчальних досягнень	Кількість учнів, які виконували завдання тесту певного формату		
	ЗФ	КФ	ВФ
Низький	6	7	3
Середній	23	19	17
Достатній	28	31	28
Високий	6	13	17
<i>Разом</i>	63	70	65

Таблиця 2

Розподіл учнів за значенням середнього тестового балу

Рівень навч. досягнень	Середній тестовий бал учнів		
	ЗФ	КФ	ВФ
Низький	4,8333	3,2857	3,6666
Середній	6	4,8947	2,8235
Достатній	7,9642	7,2258	6,5357
Високий	11,1666	10,6923	9,1764
<i>Разом</i>	7,4903	6,5246	5,5506

Приклад обчислення критерію U Манна – Вітні

ЗФ	($n_1 = 6$)	КФ	($n_2 = 7$)
Результат учня, у бал.	Ранг	Результат учня, у бал.	Ранг
10	13	2	3
2	3	2	3
4	7,5	2	3
2	3	3	6
5	10	4	7,5
6	12	5	10
	$\Sigma = 48,5$	5	10
			$\Sigma = 42,5$

Наведемо *приклад* обчислення критерію U Манна – Вітні (рангова непараметрична перевірка гіпотези) для порівняння результатів учнів з *низьким рівнем* навчальних досягнень, що виконували тести ЗФ і КФ (Табл. 3).

$$\Sigma = (13 \cdot 14) : 2 = 91 = 48,5 + 42,5$$

Більша рангова сума відповідає вибірці, що виконувала тест ЗФ. Сформулюємо *гіпотезу* H_0 : учні, які виконували завдання тесту ЗФ продемонстрували не вищі результати, ніж учні, які виконували завдання тесту КФ.

Гіпотеза H_1 : учні, які виконували завдання тесту ЗФ продемонстрували вищі результати, ніж учні, які виконували завдання тесту КФ.

$$u_{емп} = 7 \cdot 6 + (7 \cdot 6) : 2 - 48,5 = 14,5$$

$$u_{емп 2} = 7 \cdot 6 + (7 \cdot 8) : 2 - 42,5 = 27,5$$

Вибираємо менше значення $u_{емп}$. Для порівняння з критичним значенням u .

$$U_{кр.} = 4 \quad (\rho \leq 0,01); \quad U_{кр.} = 8 \quad (\rho \leq 0,05).$$

Оскільки $u_{емп} > U_{кр.}$ ($14,5 > 8$; $14,5 > 4$), то констатуємо відсутність достовірних відмінностей.

Відповідь: гіпотезу H_0 приймаємо.

Отже, учні, які виконували завдання тесту ЗФ, продемонстрували не вищі результати, ніж учні, які виконували завдання тесту КФ.

Порівняння результатів учнів за виконання завдань тесту ЗФ і КФ, ЗФ і ВФ на кожному рівні навчальних досягнень учнів за U-критерієм Манна – Вітні дало змогу встановити статистичну значимість отриманих результатів.

Висновки. У результаті дослідження з'ясовано:

1. Учні, які виконували завдання тесту ЗФ не продемонстрували статистично значимо вищих результатів, на кожному з рівнів навчальних досягнень, ніж учні, які виконували завдання тесту КФ.

2. Учні, які виконували завдання тесту ЗФ, продемонстрували статистично значимо вищий результат, ніж учні, які розв'язували відповідні завдання тесту ВФ і на середньому рівні навчальних досягнень ($\rho \leq 0,01$), і на достатньому рівні навчальних досягнень ($\rho \leq 0,05$), і на високому рівні навчальних досягнень ($u_{емп} = 76,5 > 18 = u_{кр.}$ для $\rho = 0,01$ і $u_{емп} = 76,5 > 26 = u_{кр.}$ для $\rho = 0,05$,

однак $u_{емп.2} = 25,5 \leq 26 = u_{кр.}$, що вказує на статистично значимі відмінності на рівні $\rho \leq 0,05$).

Отримані результати свідчать про те, що закритий формат підсумкового тесту з математики впливає на результати учнів, отже і на якість оцінювання. Найбільший приріст середнього тестового балу виявлено у парі «тест ВФ; тест ЗФ» на вибірці учнів з середнім рівнем навчальних досягнень: від 2,8 до 6. Саме учні з середнім рівнем навчальних досягнень з математики, у випадку підсумкового оцінювання у тестовій формі з використанням тесту лише із завдань з вибором однієї правильної відповіді з кількох її варіантів, отримують найбільше переваг. Їх оцінка, з імовірністю не менше 0,99, підвищиться. Цікаво, що навіть учні з високим рівнем навчальних досягнень з імовірністю не менше 0,95 також підвищують оцінку.

Варто зауважити, що тести ЗФ мають обмеження щодо перевірки уміння учнів доводити тотожності, нерівності, твердження, звужуючи як змістовий, так і когнітивний домени підсумкового тесту.

Як позитивне зазначимо, що підсумкові тести ЗФ і КФ створюють для учнів більш комфортну психологічну атмосферу під час випробування, ніж тести ВФ. Учні отримують завдання і відповіді до завдань. Нехай і з дистракторами.

Запропонована методика (виявлення впливу формату тесту на результати учасників тестування) може стати у нагоді розробникам тестових матеріалів, як доказова база для обґрунтування вибору автором тесту найефективнішого формату для здійснення певного виду контролю навчальних досягнень учнів.

Подальший розвиток дослідження вбачаємо у вивченні впливу різних моделей підсумкового авторського тесту з математики (алгебра і геометрія) комбінованого формату на результати учасників тестування.

БІБЛІОГРАФІЧНИЙ СПИСОК:

1. Тестові технології оцінювання ключових і предметних компетентностей учнів основної і старшої школи : монографія / за ред. О.І. Ляшенко, Ю.О. Жука. Київ : Педагогічна думка, 2014. 200 с.

2. УЦОЯО. URL: <http://testportal.gov.ua/>
3. Михайлычев Е.А. Дидактическая тестология. Москва : Народное образование, 2001. 432 с.
4. Введение в классическую и современную теорию тестов : учебник/ Л. Крокер, Дж. Алгина; пер. с англ. Н.Н. Найденовой, В.Н. Симкина, М.Б. Челышковой; под общ. ред. В.И. Звонникова, М.Б. Челышковой. Москва : Логос, 2010. 668 с.
5. Жук Ю.О., Ващенко Л.С. Вплив форми тестових завдань і профільної спеціалізації класів на результати виконання здобувачами середньої освіти тесту з біології. *Український педагогічний журнал*. 2019. № 2. С. 90–99.
6. Науменко С.О. Вплив форми тестових завдань на рівень освітніх результатів здобувачів базової середньої освіти. *Анот. результати наук.-дослід. роботи Ін-ту педагогіки НАПН України за 2018 рік / Нац. акад. пед. наук України, Ін-т педагогіки*. Київ : Інститут педагогіки, 2018. С. 242–243.
7. Lissitz R.W., Hou Xiaodong, Slater Sh. The Contribution of Constructed Response Items to Large Scale Assessment: Measuring and Understanding their Impact / University of Maryland; Educational Testing Service. *Journal of Applied Testing Technology*. 2012. Vol. 13. Is. #3. URL: <http://www.jattjournal.com/index.php/atp/article/view/48366>.
8. Анастаси А., Урбина С. Психологическое тестирование. 7-е изд. Санкт-Петербург : Питер, 2007. 688 с.
9. Lukhele R., Thissen D., Wainer H. On the relative value of multiple-choice, constructed response, and examinee-selected items on two achievement tests. *Journal of Educational Measurement*. 1994. № 31. P. 234–250.
10. Збірник завдань для державної підсумкової атестації. 9 кл. / за ред. З.І. Слєпкань. Харків : Гімназія, 2003. 144 с.
11. Дворецька Л.П. До проблеми оцінювання якості тестових матеріалів, що розміщені в підручниках з алгебри та геометрії для учнів 7 класу. *Наукові записки / ред. кол. : В.Ф. Черкасов, В.В. Радул та ін. Вип. 150. Серія: Педагогічні науки*. Кропивницький : РВВ КДПУ ім. В. Винниченка, 2017. С. 135–141.
12. Сидоренко Е.В. Методы математической обработки в психологии. Санкт-Петербург : ООО «Речь», 2000. 350 с.