

## **ВИЯВЛЕННЯ Й АНАЛІЗ ЧИННИКІВ ВПЛИВУ НА РЕЗУЛЬТАТИ ПІДСУМКОВОГО ТЕСТУВАННЯ ВИПУСКНИКІВ ГІМНАЗІЇ**

**канд. пед. наук Гривко Антоніна Вікторівна**  
**д-р. пед. наук Жук Юрій Олексійович**  
*Інститут педагогіки НАПН України*  
*Україна*

Реформування освітньої галузі в Україні є різноспрямованим процесом, що передбачає зміни як структурно-організаційних і змістових компонентів освіти, так і контрольньо-оцінювальних процедур. Передбачене трансформування форми державного підсумкового оцінювання випускників гімназії актуалізує дослідження, пов'язані з вивченням особливостей застосування тестових технологій для незалежного об'єктивного контролю навчальних досягнень учнів. Зокрема, такі дослідження спрямовані на визначення й обґрунтування оптимальних форматів тестових завдань і структури тесту.

Незважаючи на численні публікації, у яких висвітлено результати наукового компаративного аналізу діагностичного потенціалу та складності тих чи інших форм закритих і відкритих тестових завдань (Аванесов В. [1], Вірц М. [9], Дятлова К. [2], Карпова І. [3], Мозафарі Ф. [7], Лізітц Р. [6], Хасані Л. [5] та ін.), недостатньо розкритим залишається питання щодо визначення впливу їх на результати тестування.

Пропоноване дослідження спрямовано на комплексне виявлення чинників, які впливають на результативні показники тестування (РП), та вивчення особливостей такого впливу.

Розроблення методики дослідження ґрунтувалось на виснуваному за результатами теоретичного аналізу припущенні, що на результати тестування впливають такі чинники: 1) неконтрольовані – характеристики респондентів: особистісні (когнітивний – знання й уміння, афективний – перцептивно-емоційне оцінювання предмета діяльності – тестових завдань) й групові (зокрема, профільність класу); 2) контрольовані: змістові (зміст тестових завдань (ТЗ) і зміст тесту загалом) і зовнішньоорганізовані конструкти тесту (зокрема, формат тестових завдань і форма передбачуваних відповідей). З метою вивчення впливу перелічених чинників розроблено критеріально-орієнтовані тести з української мови в двох варіантах (як передбачалося, паралельних за змістом і складністю). За змістом тести містили здебільшого завдання за матеріалом 8 класу (тестування у 8 класах проводилось наприкінці навчального року, у 9 класах – на початку навчального року). Кожен варіант тестів містив 4 блоки завдань різних форматів, спрямованих на перевірку засвоєння учнями одних і тих самих навчальних конструктів – знань і вмінь (закриті завдання, що передбачали вибір однієї правильної відповіді за текстом (I блок), закриті з вибором однієї і декількох правильних відповідей (II блок), завдання на

встановлення відповідностей (III блок), відкриті завдання, що передбачали конструйовану реченням відповідь (IV блок)). Для вивчення афективного чинника було розроблено опитувальник ставлення учнів до різних форматів розв'язуваних ними завдань за методом семантичного диференціала [8].

У дослідженні взяли участь 203 учні 8 і 9 класів (99 і 104 ос. відповідно) хіміко-біологічного (ХБ), фізико-математичного (ФМ) та економічного (Екон) профілів, які в ситуації реального освітнього процесу спершу проходили тестування з української мови, а потім оцінювали виконані ними завдання за запропонованим опитувальником (семантико-емотивне оцінювання): із застосуванням процедури шкалювання за визначеними біполярними ознаками учні оцінювали виокремлені блоки тестових завдань.

Результати тестування було опрацьовано за допомогою статистичних методів у MS Excel та програмному пакеті STATISTICA. Визначено відповідність закону нормального розподілу як загальних результатів, так і часткових (за кожним окремим блоком завдань і в кожному класі). Здійснено багатовимірний дисперсійний аналіз впливу виявлених чинників на результати. Точність розрахунків, що в кожному окремому разі перевищувала 99% ( $p < 0,01$ ), сила впливу чинників визначалися за методом М. Плохинського [4]. Оскільки в пропонованому дослідженні РП розглядаються як вияв сукупності впливів виокремлених чинників ( $РП = f\{C_1; C_2; C_3; C_n\}$ ), у процесі аналізу результатів перевірялися такі статистичні гіпотези про вплив на характеристики РП: 1) чинника «формат тестового завдання» ( $C_1$ ); 2) когнітивного чинника ( $C_2$ ); 3) чинника «профільна спеціалізація класа» ( $C_3$ ); 4) афективного чинника ( $C_4$ ). За результатами опрацювання бланків опитувальника було визначено домінантні ознаки (маркери), що виражають сформоване в процесі виконання тестів узагальнене перцептивно-емотивне ставлення учнів до запропонованих форматів завдань. Результати опитування (вербально-числові оцінки) зіставлено з РП тестування (тестовими балами) в якісному вимірі з огляду на різні системи та інструменти вимірювання.

Узагальненими результатами висвітлюваного дослідження є такі:

1. Сила впливу чинника «формат тестового завдання» виявляється на рівні 33%. Порівняльний аналіз результатів тестування учнів 8 і 9 класів за I і II варіантами дав можливість виявити, що сила впливу розглядуваного чинника коливається залежно від текстури тесту – для I варіанта вона становить 33,1%, для II – 17,3%, притому тенденція слабшого вияву сили впливу форми тестових завдань на РП у II варіанті спостерігається як у 8, так і в 9 класі. Виявлені відмінності зумовили окремий подальший розгляд результатів за варіантами. Водночас у процесі аналізу з'ясовано, що сила впливу розглядуваного чинника змінюється відповідно до варіацій поєднання різних форм завдань в одному тесті.

2. Сила впливу когнітивного чинника у 8 і 9 класах коливається від 67 до 83% (залежно від варіанта).

3. Сила впливу чинника «профільна спеціалізація класу» по-різному виявляється в I і II варіантах тестів (різниця досягає 14%), спостерігається також залежність її від освітнього рівня учнів (8 чи 9 клас) – різниця досягає

10%. Порівняння результатів учнів за профільністю класів, дало змогу з'ясувати, що освітній рівень впливає на результати в межах 25%. Порівняння показників сили впливу чинника «профільна спеціалізація» в класах із різними профілями (ХБ, ФМ, Екон) показало близькість їхніх значень.

4. Найбільш стійкий показник сили впливу, який майже не залежить від інших чинників (варіанта, освітнього рівня, профільності класу) має «невиявлений чинник» і становить 42%.

5. Аналіз афективного чинника підтвердив обумовленість загального ставлення учнів до діяльності з розв'язування завдань тих чи інших форматів їх перцептивно-емотивним образом. Так, вербально-числові оцінки запропонованих форматів завдань співвідносяться з тестовими балами за їх виконання: загалом позитивно респонденти оцінили завдання з вибором однієї правильної відповіді («приємний», «швидкий»), які за результатами тестування виявилися для них найлегшими; завдання, спрямовані на довільне конструювання відповідей, здебільшого викликають у респондентів почуття невпевненості у своїх силах і сприймаються ними як найбільш складні, а тому непривабливі («потворний», «холодний», «повільний»). Натомість сприйняття учнями завдань на встановлення відповідностей як оптимально складних зумовило формування позитивного інтересу до їх розв'язування («яскравий»).

Узагальнення результатів дослідження дає змогу зробити висновки щодо впливу розглянутих чинників на результати тестування.

Наявність впливу чинника «формат завдань» та змінюваність показників його сили свідчить про вірогідність зміни результатів тестування відповідно до обраного формату завдань та їх співвідношення в одному тесті (що зумовлює необхідність додаткового дослідження щодо вивчення складності поєднання та кількісного співвідношення різних форматів завдань в одному тесті).

Найбільшою силою впливу на РП характеризується когнітивний (знаннєвий) чинник. Варіювання показників сили його впливу (як і чинника «формат ТЗ») у I і II варіантах тестів з огляду на збереження загальної тенденції (на різних освітніх рівнях і за результатами різних вибірок) можна пояснити чинником «зміст завдань», оскільки паралельні завдання в двох варіантах використовуюваного тесту були спрямовані на перевірку засвоєння одного й того ж елемента навчального матеріалу, але засобами різного операційного матеріалу (використання різних лексем, словосполучень і речень на застосування теоретичних знань).

Близькість значень показників сили впливу фактора «профільна спеціалізація» у класах різних профілів можна пояснити зі специфікою предметного поля тестування – усі групи тестованих (учні ХБ, ФМ та Екон класів) характеризуються наявністю діяльній опори в навчанні української мови на одному рівні.

Зіставлення результатів аналізу афективного чинника з результатами тестування дає змогу зробити висновки про те, що формат ТЗ є сенсорно-аферентним стимулом оперативного образу діяльності, що, зокрема, виявляється у формуванні інтересу й мотивації до розв'язування завдань.

Відповідно, доцільним є поєднання різних форматів ТЗ в одному тесті як для формування позитивних настанов і мотиваційного оптимізму в процесі їх виконання учнями, так і для перевірки й уточнення результатів тестування.

Визначена в процесі пропонованого дослідження залежність РП та сили впливу аналізованих чинників від освітнього рівню респондентів, змісту ТЗ, структури тесту, а також достатньо високий і стійкий показник сили впливу на РП «невиявленого» чинника зумовлюють урахування їх у подальших наукових пошуках, спрямованих на поглиблення та пояснення представлених результатів.

#### Список використаних джерел:

1. Аванесов В. С. Композиция тестовых заданий : учеб. пособие. М.: Центр тестирования, 2002. 239 с.
2. Дятлова К. Д. Исследование влияния разнообразия форм тестовых заданий на статистические характеристики теста. *Вопросы тестирования в образовании*. 2006. № 4. С. 65 – 75.
3. Карпова И. П. Сравнение открытых и выборочных тестов. *Открытое образование*. 2010. № 3. С. 32 – 38.
4. Плохинский Н. А. Биометрия : учеб. пособие. Новосибирск : Изд-во АН СССР, 1961. 362 с.
5. Hassani L. A Study of Students' Reading Performance in Two Test Formats of Summary Writing and Open-ended Questions / Leila Hassani, Tengku Nor Rizan, Mohamad Maasum. *Procedia – Social and Behavioral Sciences*. 2012. № 69. P. 915 – 923. DOI: 10.1016/j.sbspro.2012.12.016
6. Lissitz R. W., Hou X. Multiple Choice Items and Constructed Response Items: Does It Matter? / University of Maryland. 2012. 37 p. URL : <https://www.marces.org/multiple%20choice%20items%20and%20constructed%20response%20items.pdf>
7. Mozaffari F. Investigating the Impact of Response Format on the Performance of Grammar Tests: Selected and Constructed / Farzaneh Mozaffari, S. Mohammad Alavi, Abbasali Rezaee. *Journal of Teaching Language Skills (JTLS)*. 36 (2), Summer 2017. P. 103 – 128.
8. Osgood Ch. E., Suci G. J., Tannenbaum P. The Measurement of Meaning. USA, University of Illinois Press, 1957. 348 p.
9. Wirtz M. Effects of different response formats in measuring educational research literacy / Schladitz, Sandra; Groß Ophoff, Jana; Wirtz, Markus. *Journal for educational research online*. 2017. № 9 (2). P. 137 – 155.