



**Юрій Жук** – доктор педагогічних наук, доцент, завідувач відділу моніторингу та оцінювання якості загальної середньої освіти Інституту педагогіки НАПН України, м. Київ, Україна.

**Коло наукових інтересів:** педагогічна кваліметрія; середовища навчання.

✉ [zhukyrij@gmail.com](mailto:zhukyrij@gmail.com)

id <https://orcid.org/0000-0002-6932-2484>

**Світлана Науменко** – кандидат педагогічних наук, старший науковий співробітник, провідний науковий співробітник відділу моніторингу та оцінювання якості загальної середньої освіти Інституту педагогіки НАПН України, м. Київ, Україна.

**Коло наукових інтересів:** якість загальної середньої освіти; моніторинг якості освіти; тестові технології.

✉ [sveta\\_naum@ukr.net](mailto:sveta_naum@ukr.net)

id <https://orcid.org/0000-0002-8279-4427>



УДК 37.091.26:[316.422:37(045)]

<https://doi.org/10.32405/2411-1317-2024-2-104-117>

## ОСВІТНІ РЕФОРМИ У ДЗЕРКАЛІ ЗОВНІШНЬОГО НЕЗАЛЕЖНОГО ОЦІНЮВАННЯ

**Анотація.** У статті розглянуто проблему аналізу ефективності управлінських, методичних та економічних втручань у систему освіти, які спрямовані на покращення академічних результатів учнів на популяційному рівні, як одного з важливих завдань педагогічної науки. У дослідженні автори виходили з того, що зовнішнє незалежне оцінювання (ЗНО), яке можна розглядати як об'єктивний показник рівня академічної успішності випускників закладів загальної середньої освіти (ЗЗСО), дає можливість охарактеризувати стан системи освіти в Україні на момент сесії ЗНО.

На основі аналізу офіційних звітів зовнішнього незалежного оцінювання, які опубліковані у відкритому доступі, зроблено спробу з'ясувати вплив загальної результативності заходів, реалізованих у системі освіти України, на рівень навчальних досягнень випускників ЗЗСО за період 2009–2021 рр. Для аналізу обрано навчальні предмети, іспити з яких випускники склали на всіх сесіях ЗНО у цей проміжок часу, і результати оцінювання з яких містяться в офіційних звітах Українського центру оцінювання якості освіти (УЦОЯО). Низку сесій ЗНО розглянуто як ряд рандомізованих неконтрольованих квазіекспериментів національного масштабу, які спрямовані на вимірювання певного параметра – рівня академічної успішності випускників ЗЗСО. Результатом зазначеного квазіексперименту є дані, що отримані в результаті проведення певної кількості

сесій ЗНО. Методами математичної статистики на основі отриманих даних проведено аналіз результативності впливу заходів, які здійснювалися в системі освіти (освітніх втручань) упродовж тривалого часу, на рівень академічної успішності випускників ЗЗСО.

У статті зроблено висновок, що у визначений період в цілому підвищилася успішність виконання випускниками ЗЗСО як окремих предметних тестів, так і середня успішність за результатами ЗНО. Висунуто гіпотезу, що одним з основних чинників такого підвищення стала позитивна поведінкова реакція соціуму на впровадження ЗНО, яке є складовою множини освітніх реформ в Україні.

**Ключові слова:** загальна середня освіта; зовнішнє незалежне оцінювання; освітні реформи; величина (розмір) ефекту.

**Постановка проблеми.** Предметом більшості досліджень із соціальних наук є виявлення можливих причин, детермінант, чинників тощо певного набору спостережень. Зокрема, в освітній системі під час планування та/або вироблення управлінських рішень важливо знати, до яких наслідків призводять ті чи інші втручання у цю систему (Russo et al., 2006).

У соціальних науках під час аналізу причинно-наслідкових зав'язків особлива увага звертається на існування коваріат, тобто на зміни, які впливають на відгук (результат, що спостерігається), але не є цікавими для дослідження, та на чинники, що можуть спотворювати результат. У всіх випадках найважливішим є накопичення об'єктивних даних (спостережень), які, у свою чергу, забезпечуються відповідними інструментами, процедурами та заходами, які було застосовано для їх отримання.

У Законі України «Про освіту» (Про освіту, 2017) до інструментів, процедур та заходів забезпечення і підвищення якості освіти належить зовнішнє незалежне оцінювання результатів навчання (стаття 41, п. 1). У Законі також визначено, що ЗНО є «оцінюванням результатів навчання, здобутих особою на певному рівні освіти» (Про освіту, 2017 (стаття 47, п. 1)). Отже, ґрунтуючись на положеннях Закону, можна стверджувати, що в результатах ЗНО відображається якість освіти на основі оцінювання результатів навчання.

Низку сесій ЗНО можна розглядати як ряд рандомізованих неконтрольованих квазіекспериментів національного масштабу, що спрямовані на вимірювання певного параметра – рівня академічної успішності випускників ЗЗСО. За такого підходу вся множина чинників впливу втручань у систему освіти, включаючи коваріати та чинники, що можуть потенційно спотворювати результат (об'єктивні, суб'єктивні, ситуативні тощо), є деякою інтегрованою причиною змін, яка спостерігається у певний період реалізації системи ЗНО.

Найважливішим результатом зазначеного квазіексперименту є набори даних, які отримані в результаті проведення певної кількості сесій ЗНО. Адже сесія ЗНО, дає унікальну можливість оцінити стан системи освіти в даний момент часу шляхом вимірювання параметра, який було обрано для характеристики суб'єктів вимірювання – учнів як кінцевих користувачів системи загальної середньої освіти. Відтак, актуальним є дослідження з використанням методу порівняння змін у результатах ЗНО як наслідку зовнішніх впливів на систему освіти у контексті результативності освітніх реформ.

**Формулювання цілей статті.** Мета цієї статті полягає у розгляді одного із методів, які доступні для вирішення проблем, що пов'язані з можливістю оцінювання заходів, які здійснюються в системі освіти (освітніх втручань). Ми сподіваємося, що вітчизняні дослідники серйозно сприйматимуть ці проблеми, адже результати впливу освітніх втручань можуть відрізнятися із багатьох причин, включаючи розбіжності як у реалізації методів та способів втручання, так і в рівні використання результатів оцінювання втручання з урахуванням соціальних ситуацій. У цьому контексті доцільним є вивчення можливості оцінювання ефекту від втручань у систему освіти на основі результатів ЗНО.

**Аналіз останніх досліджень і публікацій.** Найпопулярнішим напрямом досліджень у галузі освіти є вивчення впливу зовнішніх втручань на оцінки учнів ЗЗСО. В ідеалі результати цих

оцінок мають використовуватися для прийняття обґрунтованих політичних, управлінських та методичних рішень на основі виявлення результативності втручань, наприклад, щодо того, чи треба прийняти навчальну програму або скасувати освітню програму (Stuart, 2010; Weiss et al., 2017). У статті Е. Тіптон (Tipton E.) та Р.Б. Олсена (Olsen R.B.) виявлено, що за останні 30 років значно збільшилися знання, інструменти та методи оцінювання втручань в освіту (Tipton & Olsen, 2018). Дж. Ледвіг (Ladwig J.), А. Люк (Luke A.), Д. Хейс (Hayes D.) та інші дослідники у своїй праці (Ladwig et al., 1999) зауважують, що зміст поняття «ефективність освіти» відрізняється від спрощеного уявлення про те, що вважається досягненням, оскільки передбачається, що результати шкільного навчання можна виміряти за допомогою звичайних показників сформованості навичок, моделей поведінки, знань та компетентностей. При цьому вчені розглядають ці вузькі концепції як історичні артефакти індустріальної епохи. У звіті (Schochet et al., 2014) здійснено узагальнення досліджень, що присвячені кількісним методам оцінювання того, як вплив освітніх втручань на практику викладання і навчання учнів відрізняється на рівні учнів, викладачів та закладу загальної середньої освіти.

Зарубіжні дослідники в галузі освіти особливу увагу звертають на вимірювання й інтерпретацію розміру ефекту впливу на результати залежно від типів втручань у систему освіти. Розумінню величини ефекту присвячена робота Б. Вольф (Wolf B.) та Е. Гарбаткін (Harbatkin E.) (Wolf & Harbatkin, 2023). У ній зазначено, що величина ефекту безпосередньо пов'язана з типом показника результату. В інших дослідженнях виявлено, що величина ефекту описує оцінку ефективності професійної діяльності вчителів (Soland, 2016), ефективності ЗСО (Ladd & Walsh, 2002), розриви в успішності учнів (Quinn et al., 2016), сезонні моделі навчання (von Hippel et al., 2018) та освітні заходи (Bloom, 1988). Натомість, як показує огляд вітчизняних і зарубіжних публікацій, нині є мало даних про чинники, які визначають ефект від більшості різноманітних втручань у систему освіти.

Основні фундаментальні та експериментальні дослідження в галузі оцінювання рівня навчальних досягнень учнів ЗСО України останніми роками були зосереджені переважно на особливостях підсумкового оцінювання «з високими ставками», що пояснюється впровадженням ЗНО (Жук та ін., 2011; Ляшенко та ін., 2014; Ляшенко та ін., 2017; Ляшенко та ін., 2018; Ляшенко та ін., 2019). У наведених роботах досліджено особливості внутрішнього та зовнішнього контролю знань, умінь, навичок та компетентностей, критерії оцінювання навчальних досягнень учнів, методи, принципи та види оцінювання тощо. Однак у цих роботах не розглядалася проблема оцінювання результативності освітніх реформ з використанням тестових технологій.

Треба констатувати той факт, що відстеження результативності освітніх реформ в Україні не притаманне вітчизняній системі освіти, на відміну від більшості країн світу, де саме аналізу результативності освітніх реформ приділяється велика увага. Наприклад, у роботах (Alexander et al., 1998; Bracey, 2000; Berends et al., 2002; Heertum & Torres, 2011; Klein & Rice, 2012; Let's Go Learn, n.d.; Young, 2018) на основі масивів даних, переконливих статистичних доказів та результатів спеціальних досліджень аналізуються освітні досягнення у США, розглядаються позитивні та негативні зміни в системі освіти США, які викликані освітніми реформами, заснованими на стандартах, та висвітлюються наслідки цих змін і потенційні загрози, які вони становлять для ролі освіти у соціальному розвитку держави. З величезного обсягу публікацій, які присвячені проблемі, що тут розглядається, наведено саме ці праці, оскільки в них висвітлюються підходи до аналізу результативності освітніх реформ, які актуальні для вітчизняних дослідників у галузі педагогіки. Адже система реформ в Україні також ґрунтується на державних стандартах. Згідно із Законом України «Про повну загальну середню освіту» (Про повну загальну середню освіту, 2020) «якість повної загальної середньої освіти – це відповідність результатів навчання, здобутих учнем на відповідних рівнях повної загальної середньої освіти, державним стандартам» (Стаття 1, п. 20).

#### **Виклад основного матеріалу дослідження.**

#### **Організація дослідження й опрацювання його результатів.**

Перегляд результатів ЗНО за різні роки показує, що середні бали, набрані здобувачами, та максимально можливий бал, який може набрати учень, з кожного конкретного предмету,

відрізняються. Виходячи з того, що зміст і структура тестів, які використовуються в процедурах ЗНО, відповідають вимогам, визначених Державними стандартами для кожного предмету, можна припустити, що у випадку, коли середній бал дорівнює максимально можливому, спостерігається досягнення найвищої можливої якості освіти для даної популяції випускників ЗЗСО. Інакше кажучи, на момент тестування стан системи освіти повністю забезпечує вимоги до запланованої якості (принаймні за вимогами до рівня якості декларативних та процедурних знань випускників ЗЗСО). Розбіжність між середнім та максимальним балом може характеризувати рівень відповідності результатів тестування вимогам Державного стандарту, тобто характеризує стан системи освіти, що сформувалася на момент виміру (сесії ЗНО).

У нашому дослідженні одиницею виміру стану системи освіти обрано «академічну успішність» учнів, яка визначається як відсоток середнього набраного балу до максимально можливого балу, який може отримати здобувач за виконання конкретного тесту з конкретного навчального предмета у конкретній сесії ЗНО. Запропоноване визначення основної одиниці вимірювання (академічна успішність) не є прямим синонімом поняття «академічні досягнення», адже у системі ЗНО застосовуються нормативно-орієнтовані тести, основне завдання яких – ранжування учнів за результатами тестування (у системі абсолютних балів за виконання тесту). Однак, висока валідність тестів, які використовуються у ЗНО, дозволяє допустити пряму залежність академічної успішності від академічних досягнень випускників ЗЗСО.

Висновки, що ґрунтуються на статистичних даних, які отримані в результаті вивчення сформульованої проблеми шляхом спостереження за результатами ЗНО, є індуктивними, оскільки вони припускають одержання висновку про те, що не спостерігалось (причини) на основі того, що спостерігалось (результати). З іншого боку, вивчення деяких особливостей проблеми через спостереження пов'язане з проблемою накопичення інформації в процесі накопичення спостережень. У нашому випадку накопичення спостережень здійснювалося у процесі реалізації системи ЗНО в досліджуваній період, що є важливим чинником для причинного аналізу (причини повинні передувати своїм наслідкам у часі). У нашому випадку незаперечним фактом є те, що випускники кожного року піддавалися прямим або опосередкованим впливам освітніх втручань упродовж значного періоду часу (11 років), що передують здійсненню процесу обраного способу вимірювання.

Важливо також звернути увагу, що у соціальних, психологічних і педагогічних дослідженнях треба враховувати наявність двох рівнів причинності: щодо популяції і щодо індивідуума. У контексті нашого дослідження ці відмінності полягають у тому, що глобальні освітні втручання стосуються поведінки популяції (як багатовекторної реакції соціуму на втручання в систему освіти), а на поведінку індивіда впливають особистісні мотиви, які ініціюються втручанням, і визначаються неоднорідністю індивідуальних характеристик. Вивчення співвідношення цих рівнів не входить у завдання нашого дослідження, адже є не педагогічною, а епістемологічною проблемою (Courgeau, 2003).

У представленому дослідженні використано офіційні звіти про результати проведення зовнішнього незалежного оцінювання, які розміщені у відкритому доступі на сайті Українського центру оцінювання якості освіти (УЦОЯО) (<https://testportal.gov.ua/ofzvrit/>). Дослідження обмежене періодом 2009–2021 рр. у зв'язку з тим, що з 2009 року несуттєво змінювався перелік навчальних предметів, що підлягали оцінюванню на сесіях ЗНО, а з 2022 року у зв'язку із воєнним станом було введено новий формат оцінювання майбутніх абітурієнтів – національний мультипредметний тест (НМТ).

У дослідженні використано дані про результати виконання тестів ЗНО з таких навчальних предметів: українська мова та література, історія України, математика, фізика, хімія, біологія, географія, англійська мова, німецька мова, французька мова та іспанська мова. Ці предмети склалися в усіх сесіях ЗНО в окреслений період часу. Навчальні предмети, які епізодично включалися до сесій ЗНО у цей проміжок часу, не розглядалися у нашому дослідженні.

Такий підхід дозволив організувати оброблення статистичних даних на основі подій, що безперервно спостерігалися, з періодом (лагом) в один рік для кожного навчального предмета,

обраного для дослідження. Отже, це дослідження можна віднести до ретроспективних лонгітудних досліджень на незалежних вибірках різного розміру. Його особливістю є те, що для аналізу використовувалися дані про результати тестування учнів, які були отримані за допомогою вимірального інструментарію (предметних нормативно-орієнтованих тестів) і відрізнялися для кожного навчального предмета в різних сесіях ЗНО.

Як зазначалося вище, для порівняння оцінок різних досліджень, показників та розмірів вибірки в освіті, якими є результати сесій ЗНО, часто використовується величина ефекту (*d* Коена). Наприклад, ця величина використовується для порівняння приросту успішності учнів з часом або у дослідженнях до і після якогось втручання, або у зв'язку із відсутністю втручання, наприклад, для оцінювання розривів в успішності (Жук та Науменко, 2023; Soland & Thum, 2019).

Значення величини ефекту (відмінності) (так званої effect size, яку позначають *d*, або Cohen's *d*) обчислюється для кожного дослідження за формулою (1):

$$d = \frac{|M_1 - M_2|}{\sqrt{\frac{(N_1 - 1)SD_1^2 + (N_2 - 1)SD_2^2}{N_1 + N_2 - 2}}} \quad (1),$$

де  $M_1, M_2$  – середні значення груп вибірок,  $N_1, N_2$  – обсяги вибірок,  $SD_1, SD_2$  – стандартні відхилення (standard deviation).

За Дж. Коеном (Cohen, 1988) якщо *d* набуває значення 0,2, то величину відмінності можна вважати малою (невеликою), 0,5 – середньою, 0,8 – великою. Відповідно до досліджень (Harris, 2009; Kraft, 2020) два останні значення величини *d* (0,5 та 0,8) приймаються як практично значущі.

За даними офіційних звітів про проведення ЗНО, які розміщені на сайті УЦОЯО, нами було створено таблицю результатів сесій ЗНО за 2009–2021 рр. (табл. 1). У табл. 1 також подано величини ефекту (*d*, ES), які обчислено за допомогою формули (1), та відповідні довірчі інтервали (CI, confidence interval), які обраховано для рівня достовірності 95%.

Таблиця 1

Результати сесії ЗНО з різних предметів за 2009–2021 р.

Предмет	2009 р.			2019 р.			2021 р.			2009-2021 рр.		2019-2021 рр.	
	N	M	SD	N	M	SD	N	M	SD	d	CI	d	CI
Українська мова та література	434210	48,4	18,4	339385	44,8	22,45	222611	57,43	23,72	0,444	0,438; 0,449	0,550	0,545; 0,555
Історія України	178945	35	13	222537	38,8	13,54	203138	39,09	15,34	0,286	0,20; 0,293	0,020	0,01; 0,025
Математика (базовий рівень)	235305	13,89	11,05	155319	22,94	13,04	244542	19,59	14,46	0,442	0,436; 0,447	0,241	0,234; 0,248
Фізика	32142	13,49	7,43	21407	22,18	10,53	23407	24,27	12,34	1,100	1,082; 1,118	0,182	0,163; 0,2
Хімія	28650	40,03	20,07	13700	33,77	20,27	9889	31,64	16,06	0,439	0,416; 0,462	0,114	0,089; 0,14
Біологія	102072	31,18	11,39	76020	38,03	13,21	82854	42,34	13,13	0,915	0,905; 0,924	0,327	0,31; 0,33
Географія	50550	41,99	14,79	75032	44,11	16,07	115916	44,5	14,89	0,169	0,158; 0,179	0,025	0,016; 0,035
Англійська мова	56701	27,3	11,8	91177	38,69	15,78	128 926	39,58	16,78	0,796	0,786; 0,806	0,055	0,046; 0,063
Німецька мова	2256	31,9	15,2	1951	39,25	18,79	1861	41,16	17,05	0,576	0,514; 0,640	0,106	0,04; 0,167
Французька мова	865	27	12,3	553	41,39	17,11	419	42,19	16,69	1,094	0,97; 1,218	0,047	-0,174; 0,08
Іспанська мова	165	34	15	114	43,83	20,09	151	44,03	20,08	0,566	0,344; 0,795	0,010	-0,231; 0,256

На рис. 1–11 зображено динаміку успішності виконання предметних тестів ЗНО (у відсотках середнього набраного балу до максимально набраного балу для даного предмету) за період з 2009 по 2021 рр. На рис. 12 зображено середню успішність виконання предметних тестів ЗНО за цей же період.



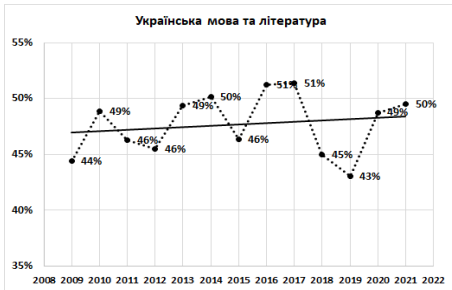


Рис. 1. Динаміка успішності виконання тестів ЗНО з української мови та літератури (2009–2021 рр.)

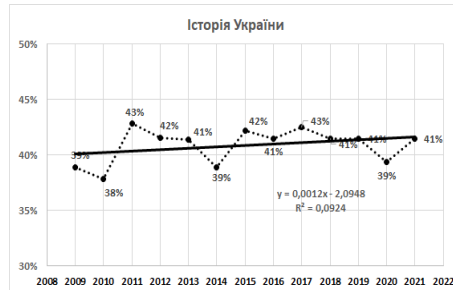


Рис. 2. Динаміка успішності виконання тестів ЗНО з історії України (2009–2021 рр.)

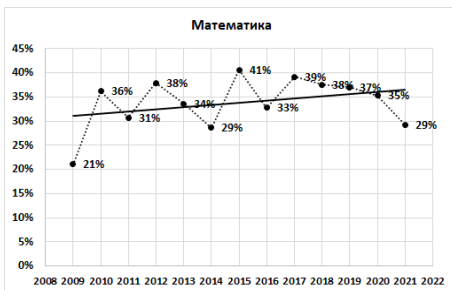


Рис. 3. Динаміка успішності виконання тестів ЗНО з математики (2009–2021 рр.)

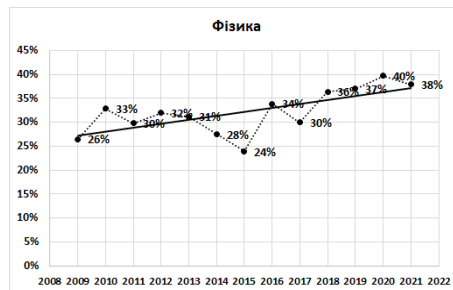


Рис. 4. Динаміка успішності виконання тестів ЗНО з фізики (2009–2021 рр.)

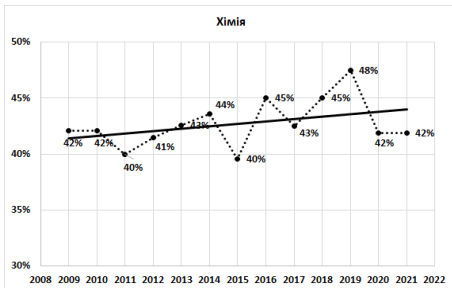


Рис. 5. Динаміка успішності виконання тестів ЗНО з хімії (2009–2021 рр.)

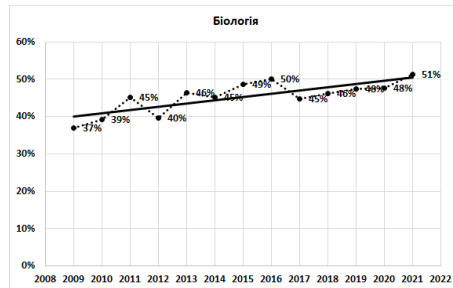


Рис. 6. Динаміка успішності виконання тестів ЗНО з біології (2009–2021 рр.)

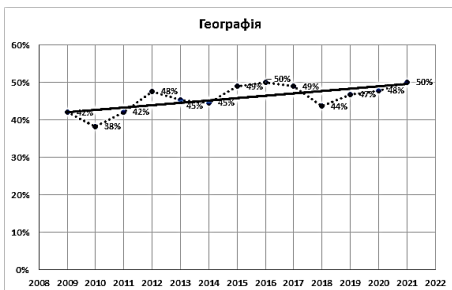


Рис. 7. Динаміка успішності виконання тестів ЗНО з географії (2009–2021 рр.)

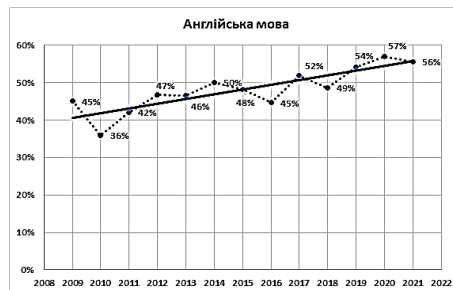


Рис. 8. Динаміка успішності виконання тестів ЗНО з англійської мови (2009–2021 рр.)

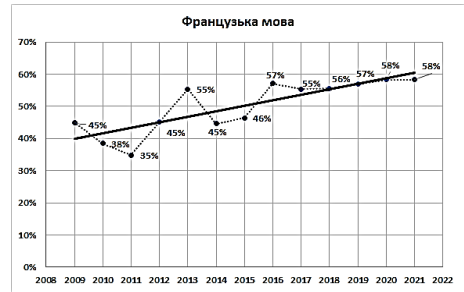
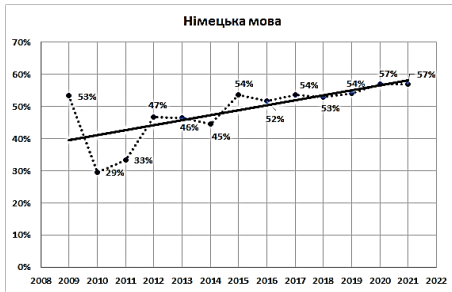


Рис. 9. Динаміка успішності виконання тестів ЗНО з німецької мови (2009–2021 рр.)

Рис. 10. Динаміка успішності виконання тестів ЗНО з французької мови (2009–2021 рр.)

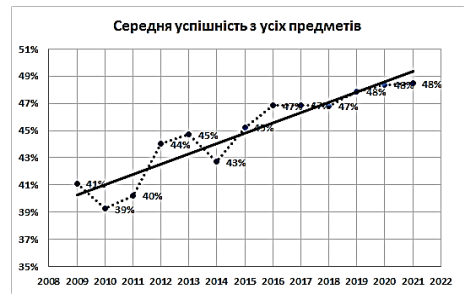
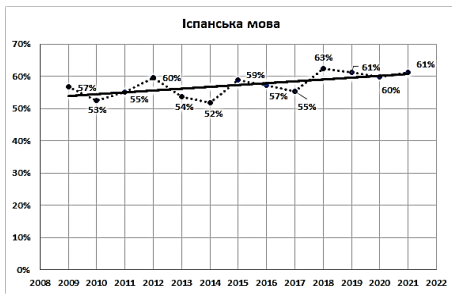


Рис. 11. Динаміка успішності виконання тестів ЗНО з іспанської мови (2009–2021 рр.)

Рис. 12. Середня успішність виконання предметних тестів ЗНО за 2009–2021 рр.

Як видно з рис. 1–12 у 2021 р. у порівнянні з 2009 р. зросла успішність виконання усіх предметних тестів ЗНО та середня успішність виконання предметних тестів ЗНО.

На рис. 13 та 14 подано результати успішності виконання предметних тестів ЗНО сесій 2009 р. та 2021 р. (у відсотках середнього набраного балу до максимально можливого балу для даного предмету).

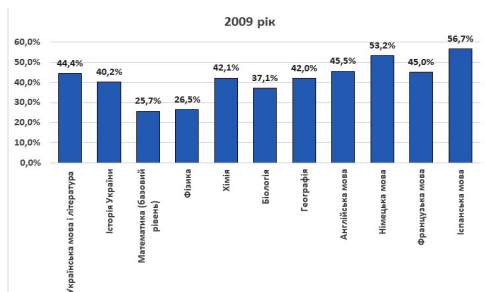


Рис. 13. Успішність виконання предметних тестів ЗНО сесії 2009 р.

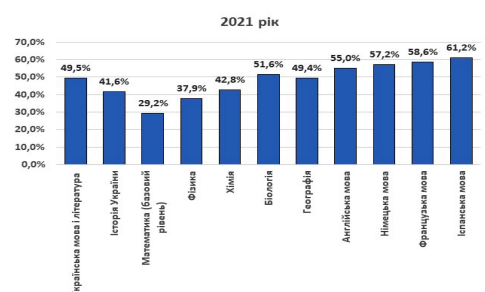


Рис. 14. Успішність виконання предметних тестів ЗНО сесії 2021 р.

Як видно з рис. 13–14 у 2021 р. як і у 2009 р. найвища успішність виконання тестів ЗНО була з іспанської мови, а найнижча – з математики. У розрізі щодо приросту успішності виконання

предметних тестів ЗНО за 2021 р. у порівнянні з 2009 р., який зображено на рис. 15, найвищий приріст успішності виконання тестів ЗНО був з біології (14,5%), а найнижчий – з хімії (0,6%).



Рис. 15. Загальний приріст успішності виконання предметних тестів ЗНО (2009–2021 рр.)



Рис. 16. Загальний розмір впливу (d) освітніх заходів на успішність виконання предметних тестів ЗНО (2009–2021 рр.)

Щодо розміру ефекту впливу (d) освітніх заходів на успішність виконання предметних тестів ЗНО упродовж 2009–2021 років, який зображено на рис. 16, то у цей період освітні заходи найбільше вплинули на успішність виконання тесту ЗНО з фізики (1,1), а найменше – з географії (0,169).

На рис. 17 зображено зміну успішності виконання предметних тестів ЗНО за 2021 р. у порівнянні з 2019 р. (У цей проміжок часу входить період (з 2020 р.) коли через Covid-19 заклади освіти України перейшли на змішану та дистанційну форми навчання). Як видно з рис. 17 у 2021 р. у порівнянні з 2019 р. на 6,4% (найбільший показник) збільшилася успішність виконання тестів ЗНО з української мови та літератури та на 7,8% (найнижчий показник) знизилася успішність виконання тестів ЗНО з математики. Тобто, у порівнянні з усіма предметами, які ми розглядаємо у даному дослідженні, лише на успішність виконання тестів ЗНО з математики негативно вплинули освітні заходи 2019–2021 рр., в тому числі й навчання у змішаній та дистанційній формах.



Рис. 17. Зміна успішності виконання предметних тестів ЗНО за 2021 рік відносно 2019 року



Рис. 18. Загальний розмір ефекту (d) впливу освітніх заходів у 2019–2021 рр. на успішність виконання предметних тестів ЗНО

Щодо показника розміру ефекту (d) впливу освітніх заходів, в тому числі й карантину через Covid-19, на успішність виконання предметних тестів ЗНО упродовж 2019–2021 рр., то найбільший цей показник спостерігається у тестах ЗНО з української мови та літератури (0,550), а найнижчий – з іспанської мови (0,010) (рис. 18).

У 2020 р. у зв'язку із пандемією COVID-19 з'явилися форс-мажорні обставини, які не відносяться до освітніх реформ (25 березня 2020 р. Кабінет Міністрів України запровадив на всій території України режим надзвичайної ситуації), однак загальний позитивний тренд академічної



успішності випускників в цілому не змінився (рис. 12). Це свідчить про те, що освітні реформи мають пролангований характер.

Тобто, у 2021 р. у порівнянні з 2019 р. тест ЗНО з української мови та літератури мав найвищий показник розміру ефекту ( $d$ ) впливу освітніх заходів (0,550) та характеризувався найвищим показником зростання успішності (на 6,4%) виконання тесту ЗНО (див. рис. 16, рис. 17). Однією із причин цього, на нашу думку, є те, з 2020 р. ЗНО з цього предмету, який завжди був обов'язковим, складала лише випускники, які планували вступати до закладів вищої освіти (ЗВО) України. До 2020 р. (з 2015 р.) тести ЗНО з української мови та літератури складала усі випускники ЗЗСО, адже оцінки за цей тест зараховувалися як оцінки за державну підсумкову атестацію (ДПА) усім учням 11-х класів.

На рис. 17–18 ми спостерігаємо різні результати впливу втручань на успішність виконання тестів ЗНО з різних предметів. Вочевидь це може бути предметом подальших досліджень. Як видно з рис. 17–18, запропонований нами спосіб оцінювання академічної успішності показав зміни у результативності успішності виконання тестів ЗНО, які були викликані також форс-мажорними обставинами карантину.

**Висновки дослідження та перспективи подальших розвідок.** Ми не інтерпретуємо отримані нами результати як певну закономірність (соціальну або педагогічну), але оцінюємо як деяку тенденцію, яка залежить від багатьох чинників. Зокрема, на нашу думку, зростання успішності виконання предметних тестів ЗНО пов'язано також із тим, що більшість випускників ЗЗСО під час останнього року або двох років навчання цілеспрямовано готуються (навчаються) до складання тестів ЗНО з репетитором, на курсах, самостійно тощо.

У нашому дослідженні ми не розглядаємо результати ЗНО як показники якості освіти у тому загальному сенсі, як воно трактується в нормативних документах. Ці результати свідчать про рівень підготовленості учнів до предметного тестування. Запропонований метод аналізу не може пояснити, як і чому виникає той чи інший конкретний ефект від конкретних освітніх втручань (або від системи втручань). Для цього необхідно провести докладніші дослідження, які дадуть змогу ідентифікувати причинні механізми та сформулювати теоретичні пояснення причинних ефектів.

Перспективними є дослідження щодо контрольних показників, перелік яких має передбачати, разом із академічною успішністю, також аналіз економічної ефективності (як стандартну частину аналізу політики) різноманітного роду і рівнів втручань у систему освіти, для отримання більшої кількості фактичних даних щодо підвищення достовірності контрольованих показників і, зрештою, щодо поліпшення політичних рішень у сфері освіти.

Зовнішнє незалежне оцінювання є єдиним постійним фактором, який реалізовано у системі освіти України, і який має постійне цільове фінансування. Його впровадження як інструменту боротьби з корупцією під час вступу до закладів вищої освіти України стало позитивною дією в централізованій системі прийому до ЗВО. Держава перехопила у ЗВО ініціативу щодо прийому студентів, що значно обмежило автономність ЗВО. Якщо розглядати ЗНО як постійний вимірник, який уведено до системи освіти, то можна припустити, що результати вимірювань (тестів ЗНО) фіксують вплив великої кількості прихованих від прямого спостереження чинників, які впливають на результати вимірювань.

Можна також припустити, що позитивна поведінкова реакція соціуму на впровадження ЗНО стала тим зворотним зв'язком, який значною мірою забезпечив позитивну динаміку успішності результатів виконання іспитів ЗНО (сесії 2009–2021 рр.). Незважаючи на критику національної освітньої політики України, яка часто з'являється у ЗМІ, це дослідження показує, що у визначений період зростає академічна успішність випускників ЗЗСО з навчальних предметів, які виносилися на ЗНО.

За час існування ЗНО в Україні безпосередньо тестуванням були охоплені сотні тисяч випускників ЗЗСО. Опосередковано у цьому процесі брали участь батьки випускників, вчителі, шкільні адміністратори та органи управління освітою. Так можна припустити, що саме існуван-

ня ЗНО мало певний вплив на соціум. Тому можна зробити висновок, що спостерігається прямий причинно-наслідковий зв'язок між існуванням ЗНО (незалежна змінна) та рівнем навчальних досягнень випускників ЗЗСО (залежна змінна), яка заснована на тому, що:

- 1) результати ЗНО відображають результати багаторічних рішень та зусиль, які застосовуються упродовж одинадцяти років навчання учня у ЗЗСО;
- 2) ЗНО є прикладом універсального втручання у систему освіти, яке полягає в об'єктивізації результатів оцінювання на основі рівності вимог до учасників тестування;
- 3) ЗНО – яскравий зразок системи, в якій використовуються тести, що набувають для учнів характеру тестів з «високими ставками».
- 4) результати ЗНО дозволяють аналізувати досліджувану проблему з урахуванням фактичних даних, які отримані у процесі об'єктивних вимірів.

Вплив зміни політики ЗНО на результати тестування потребує додаткових досліджень. Аналіз досвіду, який отримано у процесі реалізації ЗНО, що базувався на використанні предметних тестів, дозволить уникнути помилок у процесі реалізації будь-яких інших форм тестування подібного масштабу та з подібними цілями. Також практичний інтерес представляє дослідження того, як вплинули перелічені вище зміни в політиці ЗНО на результати зарахування учнів до ЗВО.

Для визначення значущості впливу освітніх реформ в Україні на рівень академічних досягнень учнів ЗЗСО необхідно провести додаткові дослідження з огляду на той факт, що випускники, наприклад 2009 року, розпочинали своє навчання у ЗЗСО у 1999 році. У майбутніх дослідженнях необхідно враховувати демографічну ситуацію в конкретний момент часу (в даній сесії ЗНО) як один із параметрів вимірювання. Залишається відкритим й питання про прогностичну валідність тестів, тобто як співвідносяться результати виконання учнями тестів ЗНО з їхніми академічними здобутками у ЗВО.

Одним із напрямів подальших досліджень може бути з'ясування причинно-наслідкових зв'язків вибору майбутньої спеціальності, яка відображається у виборі випускником ЗЗСО навчальних предметів крім обов'язкових, які запропоновані у конкретній сесії ЗНО. Щоправда, ситуація ускладнюється у зв'язку з переходом на 12-річну систему освіти та нову структуру системи закладів загальної середньої освіти.

Ми сподіваємося, що дослідники, які будуть займатися питанням оцінювання впливу освітніх втручань, особливо на рівень навчальних досягнень учнів ЗЗСО, будуть використовувати дані цієї статті як основу для важливих дискусій щодо методологічних та методичних проблем у дослідженнях з цих питань.

### Використані джерела

- Жук, Ю. О., Ляшенко, О. І., Лукіна, Т. О., Ващенко, Л. С., та Полянський, П. Б. (2011). *Організаційно-методичне забезпечення моніторингових досліджень якості загальної середньої освіти: монографія*. Київ: Педагогічна думка. <https://lib.iitta.gov.ua/2967>.
- Жук, Ю. О., та Науменко, С. О. (2023). *Засоби та інструменти діагностики освітніх втрат*. У О. М. Топузов (Ред.), *Діагностика та компенсація освітніх втрат у загальній середній освіті України: методичні рекомендації* (с. 26–46). Київ: Педагогічна думка. <https://doi.org/10.32405/978-966-644-736-7-2023-190>. <https://undip.org.ua/library/diahnostyka-ta-kompensatsiia-osvitnikh-vtrat-u-zahalniy-seredniy-osviti-ukrainy-methodychni-rekomendatsii/>.
- Ляшенко, О. І., Жук, Ю. О., Ващенко, Л. С., Гривко, А. В., та Науменко, С. О. (2017). *Тестові технології оцінювання компетентностей учнів: посібник*. Київ: Видавничий дім «Сам». <https://lib.iitta.gov.ua/711199/>.
- Ляшенко, О. І., Лукіна, Т. О., Жук, Ю. О., Ващенко, Л. С., Гривко, А. В., та Науменко, С. О. (2018). *Теоретико-методичні засади побудови моніторингових систем оцінювання якості загальної середньої освіти: монографія*. Київ: ТОВ «КОНВІ ПРИНТ». <https://lib.iitta.gov.ua/713252/>.
- Ляшенко, О. І., Лукіна, Т. О., Жук, Ю. О., Ващенко, Л. С., Гривко, А. В., Науменко, С. О., та Топузова, А. В. (2019). *Запровадження моніторингових систем оцінювання якості освіти на основі тестових технологій: методичні рекомендації*. Київ: Педагогічна думка. <https://lib.iitta.gov.ua/719870/>.

- Ляшенко, О. І., Лукіна, Т. О., Жук, Ю. О., Ващенко, Л. С., Науменко, С. О., та Гривко, А. В. (2014). *Тестові технології оцінювання ключових і предметних компетентностей учнів основної і старшої школи: монографія*. Київ: Педагогічна думка. <https://lib.iitta.gov.ua/9410/>.
- Про освіту. Закон України. № 2145-VIII. (2017, Вересень 5). (зі змінами 2018–2023 рр.). *Верховна Рада України. Законодавство України*. <https://zakon.rada.gov.ua/laws/show/2145-19#Text>.
- Про повну загальну середню освіту. Закон України. № 463-IX. (2020, Січень 16). (зі змінами 2020–2023 рр.). *Верховна Рада України. Законодавство України*. <https://zakon.rada.gov.ua/laws/show/463-20#Text>.
- Alexander, D., Heavyside, Sh., Farris E., & Westat, Inc. (1998). *Status of Education Reform in Public Elementary and Secondary Schools: Teachers' Perspectives*. National Center for Education Statistics. <https://nces.ed.gov/pub99/1999045.pdf>.
- Berends, M., Chun, JoAn, Ikemoto, G. S., Stockly, S., & Briggs, R. J. (2002). *Challenges of Conflicting School Reforms: Effects of New American Schools in a High-Poverty District*. Santa Monica, CA: RAND Corporation. [https://www.rand.org/pubs/monograph\\_reports/MR1483.html](https://www.rand.org/pubs/monograph_reports/MR1483.html).
- Bloom, B. S. (1988). Helping all children learn in elementary school and beyond. *Principal*, 67(4), 12–17.
- Bracey, G. W. (2000). The 10th Bracey Report on the Condition of Public Education. *Phi Delta Kappan*, 82(2), 133–144. <https://doi.org/10.1177/003172170008200210>.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, New Jersey: Erlbaum. 567 p.
- Courgeau, D. (Ed.) (2003). *Methodology and Epistemology of Multilevel Analysis*. The Netherlands: Kluwer Academic Publishers. <https://doi.org/10.1007/978-1-4020-4675-9>.
- Harris, D. N. (2009). Toward Policy-Relevant Benchmarks for Interpreting Effect Sizes: Combining Effects With Costs. *Educational Evaluation and Policy Analysis*, 31(1), 3–29. <https://doi.org/10.3102/0162373708327524>.
- Heertum, R. V., & Torres, C. A. (2011). Educational Reform in the U.S. in the Past 30 Years: Great Expectations and the Fading American Dream. *Educating the Global Citizen*, 3–27. <https://doi.org/10.2174/978160805268411101010003>.
- Klein, Jo. I., & Rice, C. (2012). U. S. Education Reform and National Security. *Independent Task Force Report*, 68.
- Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, 49(4), 241–253. <https://doi.org/10.3102/0013189X20912798>.
- Ladd, H. F., & Walsh, R. P. (2002). Implementing value-added measures of school effectiveness: getting the incentives right. *Economics of Education Review*, 21(1), 1–17. [https://doi.org/10.1016/S0272-7757\(00\)00039-X](https://doi.org/10.1016/S0272-7757(00)00039-X).
- Ladwig, J., Luke, A., Hayes, D., Mills, M. D., & Lingard, R. L. (1999). The search for productive schooling: Beyond 'school effectiveness' and 'authentic. Australian Association for Research in Education Conference, Melbourne, 27 Nov – 2 Dec 1999. Coldstream, Vic.: AARE.
- Let's Go Learn*. (n.d.). K-12 Education Reform: A New Paradigm. <https://www.letsgolearn.com/resources/education-reform/>.
- Quinn, D. M., Cooc, N., McIntyre, J., & Gomez, C. J. (2016). Seasonal dynamics of academic achievement inequality by socioeconomic status and race/ethnicity: Updating and extending past research with new national data. *Educational Researcher*, 45(8), 443–453. <https://doi.org/10.3102/0013189X16677965>.
- Russo, F., Mouchart, M., Ghins, M & Wunsch, G. (2006). Statistical Modelling and Causality 1 STATISTICAL MODELLING AND CAUSALITY. <https://www.researchgate.net/publication/266878390>.
- Schochet, P. Z., Puma, M., & Deke, J. (2014). *Understanding Variation in Treatment Effects in Education Impact Evaluations: An Overview of Quantitative Methods*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Analytic Technical Assistance and Development. <http://ies.ed.gov/ncee/pubs/20144017/pdf/20144017.pdf>.
- Soland, J. (2016). Is Teacher Value Added a Matter of Scale? The Practical Consequences of Treating an Ordinal Scale as Interval for Estimation of Teacher Effects. *Applied Measurement in Education*, 30(1), 52–70. <https://doi.org/10.1080/08957347.2016.1247844>.
- Soland, J. & Thum, Y. M. (2019). Effect Sizes for Measuring Student and School Growth in Achievement: In Search of Practical Significance (EdWorkingPaper: 19–60). Retrieved from Annenberg Institute at Brown University. <https://doi.org/10.26300/b5as-wr12>.

- Stuart, E. A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Statistical Science*, 25(1), 1–21. <https://doi.org/10.1214/09-STS 313>.
- Tipton, E., & Olsen, R. B. (2018). A Review of Statistical Methods for Generalizing From Evaluations of Educational Interventions. *Educational Researcher*, 47(8), 516–524. <https://doi.org/10.3102/0013189X18781522>.
- von Hippel, P. T., Workman, J., & Downey, D. B. (2018). Inequality in Reading and Math Skills Forms Mainly before Kindergarten: A Replication, and Partial Correction, of “Are Schools the Great Equalizer?”. *Sociology of Education*, 91(4), 323–357. <https://doi.org/10.1177/0038040718801760>.
- Weiss, M. J., Bloom, H. S., Verbitsky-Sawitz, N., Gupta, H., Vigil, A. E., & Cullinan, D. N. (2017). How Much Do the Effects of Education and Training Programs Vary Across Sites? Evidence from Past Multisite Randomized Trials. *Journal of Research on Educational Effectiveness*, 10(4), 843–876. <https://doi.org/10.1080/19345747.2017.1300719>.
- Wolf, B., & Harbatkin, E. (2023). Making sense of effect sizes: Systematic differences in intervention effect sizes by outcome measure type. *Journal of Research on Educational Effectiveness*, 16(1), 134–161. <https://doi.org/10.1080/19345747.2022.2071364>.
- Young, V. M. (2018). Assessing the Cornerstone of U.S. Education Reform. *The Journal of Educational Foundations*, 31(3–4), 74–99. <https://eric.ed.gov/?id=EJ1212039>.

### References

- Alexander, D., Heavside, Sh., Farris E., & Westat, Inc. (1998). *Status of Education Reform in Public Elementary and Secondary Schools: Teachers' Perspectives*. National Center for Education Statistics. <https://nces.ed.gov/pub99/1999045.pdf>. (in English).
- Berends, M., Chun, JoAn, Ikemoto, G. S., Stockly, S., & Briggs, R. J. (2002). *Challenges of Conflicting School Reforms: Effects of New American Schools in a High-Poverty District*. Santa Monica, CA: RAND Corporation. [https://www.rand.org/pubs/monograph\\_reports/MR 1483.html](https://www.rand.org/pubs/monograph_reports/MR 1483.html). (in English).
- Bloom, B. S. (1988). Helping all children learn in elementary school and beyond. *Principal*, 67(4), 12–17. (in English).
- Bracey, G. W. (2000). The 10th Bracey Report on the Condition of Public Education. *Phi Delta Kappan*, 82(2), 133–144. <https://doi.org/10.1177/003172170008200210>. (in English).
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, New Jersey: Erlbaum. 567 p. (in English).
- Courgeau, D. (Ed.) (2003). *Methodology and Epistemology of Multilevel Analysis*. The Netherlands: Kluwer Academic Publishers. <https://doi.org/10.1007/978-1-4020-4675-9>. (in English).
- Harris, D. N. (2009). Toward Policy-Relevant Benchmarks for Interpreting Effect Sizes: Combining Effects With Costs. *Educational Evaluation and Policy Analysis*, 31(1), 3–29. <https://doi.org/10.3102/0162373708327524>. (in English).
- Heertum, R. V., & Torres, C. A. (2011). Educational Reform in the U.S. in the Past 30 Years: Great Expectations and the Fading American Dream. *Educating the Global Citizen*, 3–27. <https://doi.org/10.2174/978160805268411101010003>. (in English).
- Klein, Jo. I., & Rice, C. (2012). U.S. Education Reform and National Security. *Independent Task Force Report*, 68. (in English).
- Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, 49(4), 241–253. <https://doi.org/10.3102/0013189X20912798>. (in English).
- Ladd, H. F., & Walsh, R. P. (2002). Implementing value-added measures of school effectiveness: getting the incentives right. *Economics of Education Review*, 21(1), 1–17. [https://doi.org/10.1016/S 0272-7757\(00\)00039-X](https://doi.org/10.1016/S 0272-7757(00)00039-X). (in English).
- Ladwig, J., Luke, A., Hayes, D., Mills, M. D., & Lingard, R. L. (1999). The search for productive schooling: Beyond ‘school effectiveness’ and ‘authentic. Australian Association for Research in Education Conference, Melbourne, 27 Nov – 2 Dec 1999. Coldstream, Vic.: AARE. (in English).
- Let's Go Learn*. (n.d.). K-12 Education Reform: A New Paradigm. <https://www.letsgolearn.com/resources/education-reform/>. (in English).

- Liashenko, O. I., Lukina, T. O., Zhuk, Yu. O., Vashchenko, L. S., Hryvko, A. V., & Naumenko, S. O. (2018). *Teoretyko-metodychni zasady pobudovy monitorynhovykh system otsiniuvannia yakosti zahalnoi serednoi osvity: monohrafiia*. Kyiv: TOV «KONVI PRINT». <https://lib.iitta.gov.ua/713252/>. (in Ukrainian).
- Liashenko, O. I., Lukina, T. O., Zhuk, Yu. O., Vashchenko, L. S., Hryvko, A. V., Naumenko, S. O., & Topuzova, A. V. (2019). *Zaprovadzhenia monitorynhovykh system otsiniuvannia yakosti osvity na osnovi testovykh tekhnolohii: metodychni rekomendatsii*. Kyiv: Pedahohichna dumka. <https://lib.iitta.gov.ua/719870/>. (in Ukrainian).
- 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 kompetentnosti uchniv osnovnoi i starshoi shkoly: monohrafiia*. Kyiv: Pedahohichna dumka. <https://lib.iitta.gov.ua/9410/>. (in Ukrainian).
- Liashenko, O. I., Zhuk, Yu. O., Vashchenko, L. S., Hryvko, A. V., & Naumenko, S. O. (2017). *Testovi tekhnolohii otsiniuvannia kompetentnosti uchniv: posibnyk*. Kyiv: Vydavnychi dim «Sam». <https://lib.iitta.gov.ua/711199/>. (in Ukrainian).
- Pro osvitu. Zakon Ukrainy. № 2145-VIII. (2017, Veresen 5). (zi zminamy 2018–2023 rr.). *Verkhovna Rada Ukrainy*.:// zakon.rada.gov.ua/laws/show/2145–19#Text. <https://zakon.rada.gov.ua/laws/show/2145-19#Text>. (in Ukrainian).
- Pro povnu zahalnu seredniu osvitu. Zakon Ukrainy. № 463-IX. (2020, Sichen 16). (zi zminamy 2020–2023 rr.). *Verkhovna Rada Ukrainy. Zakonodavstvo Ukrainy*. <https://zakon.rada.gov.ua/laws/show/463-20#Text>. (in Ukrainian).
- Quinn, D. M., Cooc, N., McIntyre, J., & Gomez, C. J. (2016). Seasonal dynamics of academic achievement inequality by socioeconomic status and race/ethnicity: Updating and extending past research with new national data. *Educational Researcher*, 45(8), 443–453. <https://doi.org/10.3102/0013189X16677965>. (in English).
- Russo, F., Mouchart, M., Ghins, M & Wunsch, G. (2006). Statistical Modelling and Causality 1 STATISTICAL MODELLING AND CAUSALITY. <https://www.researchgate.net/publication/266878390>. (in English).
- Schochet, P. Z., Puma, M., & Deke, J. (2014). *Understanding Variation in Treatment Effects in Education Impact Evaluations: An Overview of Quantitative Methods*. Washington, DC: U. S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Analytic Technical Assistance and Development. <http://ies.ed.gov/ncee/pubs/20144017/pdf/20144017.pdf>. (in English).
- Soland, J. & Thum, Y. M. (2019). Effect Sizes for Measuring Student and School Growth in Achievement: In Search of Practical Significance (EdWorkingPaper: 19–60). Retrieved from Annenberg Institute at Brown University. <https://doi.org/10.26300/b5as-wr12>. (in English).
- Soland, J. (2016). Is Teacher Value Added a Matter of Scale? The Practical Consequences of Treating an Ordinal Scale as Interval for Estimation of Teacher Effects. *Applied Measurement in Education*, 30(1), 52–70. <https://doi.org/10.1080/08957347.2016.1247844>. (in English).
- Stuart, E. A. (2010). Matching Methods for Causal Inference: A Review and a Look Forward. *Statistical Science*, 25(1), 1–21. <https://doi.org/10.1214/09-STS 313>. (in English).
- Tipton, E., & Olsen, R. B. (2018). A Review of Statistical Methods for Generalizing From Evaluations of Educational Interventions. *Educational Researcher*, 47(8), 516–524. <https://doi.org/10.3102/0013189X18781522>. (in English).
- von Hippel, P. T., Workman, J., & Downey, D. B. (2018). Inequality in Reading and Math Skills Forms Mainly before Kindergarten: A Replication, and Partial Correction, of “Are Schools the Great Equalizer?”. *Sociology of Education*, 91(4), 323–357. <https://doi.org/10.1177/0038040718801760>. (in English).
- Weiss, M. J., Bloom, H. S., Verbitsky-Sawitz, N., Gupta, H., Vigil, A. E., & Cullinan, D. N. (2017). How Much Do the Effects of Education and Training Programs Vary Across Sites? Evidence from Past Multisite Randomized Trials. *Journal of Research on Educational Effectiveness*, 10(4), 843–876. <https://doi.org/10.1080/19345747.2017.1300719>. (in English).
- Wolf, B., & Harbatkin, E. (2023). Making sense of effect sizes: Systematic differences in intervention effect sizes by outcome measure type. *Journal of Research on Educational Effectiveness*, 16(1), 134–161. <https://doi.org/10.1080/19345747.2022.2071364>. (in English).
- Young, V. M. (2018). Assessing the Cornerstone of U. S. Education Reform. *The Journal of Educational Foundations*, 31(3–4), 74–99. <https://eric.ed.gov/?id=EJ1212039>. (in English).
- Zhuk, Yu. O., & Naumenko, S. O. (2023). Zasoby ta instrumenty diahnostryky osvity vtrah. U O. M. Topuzov (Red.), *Diahnostryka ta kompensatsiia osvity vtrah u zahalnyi seredni osvity Ukrainy: metodychni rekomendatsii*



(s. 26–46). Kyiv: Pedahohichna dumka. <https://doi.org/10.32405/978-966-644-736-7-2023-190>. <https://undip.org.ua/library/diahnostyka-ta-kompensatsiia-osvitnikh-vtrat-u-zahalniy-seredniy-osviti-ukrainy-metodychni-rekomendatsii/>. (in Ukrainian).

Zhuk, Yu. O., Liashenko, O. I., Lukina, T. O., Vashchenko, L. S., & Polianskyi, P. B. (2011). *Orhanizatsiino-metodychne zabezpechennia monitorynhovykh doslidzhen yakosti zahalnoi serednoi osvity: monohrafiia*. Kyiv: Pedahohichna dumka. <https://lib.iitta.gov.ua/2967>. (in Ukrainian).

**Yurii Zhuk**, D. Sc. (*Pedagogy*), Associate Professor, Head of the Department of Monitoring and Assessment of the Quality of General Secondary Education, Institute of Pedagogy of the National Academy of Educational Sciences of Ukraine, Kyiv, Ukraine.

**Research interests:** pedagogical quality metrics; learning environment.

**Svitlana Naumenko**, PhD (*Pedagogy*), Senior Researcher, Leading Researcher of the Department of Monitoring and Assessment of the Quality of General Secondary Education, Institute of Pedagogy of the National Academy of Educational Sciences of Ukraine, Kyiv, Ukraine.

**Research interests:** quality of general secondary education; monitoring the quality of education; test technologies.

## EDUCATIONAL REFORMS IN THE MIRROR OF EXTERNAL INDEPENDENT EVALUATION

**Abstract.** The article examines the problem relating the analysis of the effectiveness of managerial, methodical and economic interventions in the education system that are aimed at improving the academic results of students at the population level, as one of the important tasks of pedagogical science. In the study, the authors assumed that the external independent evaluation that can be considered an objective measure of the level of academic success of graduates of general secondary education institutions makes it possible to characterize the state of the education system in Ukraine at the time of the session of the external independent evaluation.

An attempt was made to find out the impact of the overall effectiveness of the measures implemented in the education system of Ukraine on the level of educational achievements of the graduates of general secondary education institutions for the period 2009–2021, based on the analysis of official external independent evaluation reports that are publicly available. An educational subject was chosen for such analysis, as well as the exams that graduates passed at all sessions of the external independent evaluation during this period of time, and the above-mentioned evaluation results could be found in the official reports of the Ukrainian Center for the Evaluation of the Quality of Education. A series of external independent evaluation sessions is considered as a series of randomized uncontrolled quasi-experiments of a national scale, which are aimed at measuring a certain parameter – the level of academic success of general secondary education institutions graduates. The result of the specified quasi-experiment is the sets of data obtained as a result of conducting a certain number of external independent evaluation sessions. Based on the obtained data, an analysis of the effectiveness of the measures implemented in the education system (educational interventions) over a long period of time as to the level of academic success of graduates of general secondary education institutions was carried out using the methods of mathematical statistics.

The article concludes that during the specified period, the success rate of the graduates of the general secondary education institutions both in individual subject tests and the average success rate according to the results of the external examinations increased. It is hypothesized that one of the main factors of such an increase was the positive behavioral reaction of society to the introduction of extracurricular activities, which is a component of many educational reforms in Ukraine.

**Keywords:** general secondary education; external independent evaluation; educational reforms; magnitude (size) of the effect.