

## Математико-статистичний апарат у сучасних психологічних дослідженнях: методологічні виклики, Mixed Methods та стандарти Open Science у вищій освіті

Державна академія інтелектуальних технологій та зв'язку (м.Одеса)

### Анотація

Статтю присвячено модернізації системи підсумкового контролю магістрів-психологів в умовах цифровізації та переходу до стандартів Відкритої науки (Open Science). Розглянуто методологічні виклики вищої школи, пов'язані з кризою відтворюваності результатів, практиками p-hacking та HARKing. Обґрунтовано перехід від репродуктивного відтворення статистичних формул до розвитку методологічної інтуїції студентів. На прикладі розроблених трикомпонентних екзаменаційних комплексів проілюстровано інтеграцію методів квантифікації якісних ознак у парадигмі Mixed Methods. Визначено етичні межі та легітимні інструменти залучення генеративного штучного інтелекту як асистента дослідника на етапі експлораторного аналізу даних (AI Disclosure, модель Human-in-the-loop). Запропоновано банк прикладних кейс-завдань, який позиціонується як гнучкий методичний конструктор для оптимізації та технологізації викладацької практики без додаткового бюрократичного навантаження.

**Ключові слова:** вища освіта, магістри-психологи, криза відтворюваності, Open Science, Mixed Methods, квантифікація, генеративний штучний інтелект, екзаменаційні кейси.

### Abstract

The article is devoted to the modernization of the final assessment system for Master's students in psychology under digitalization and the transition to Open Science standards. The study addresses methodological challenges in higher education linked to the replication crisis, p-hacking, and HARKing. A transition from reproductive memorization of statistical formulas to the development of students' methodological intuition is substantiated. Using the developed three-component exam complexes as an example, the integration of qualitative data quantification methods within the Mixed Methods paradigm is illustrated. Ethical boundaries and legitimate tools for incorporating generative artificial intelligence as a research assistant at the stage of exploratory data analysis are defined (AI Disclosure, Human-in-the-loop model). A bank of applied case tasks is proposed, serving as a flexible methodological framework to optimize and technologize teaching practices without introducing additional bureaucratic burdens.

**Keywords:** higher education, Master's students in psychology, replication crisis, Open Science, Mixed Methods, quantification, generative artificial intelligence, examination cases.

### Вступ

Сучасна вища школа функціонує в умовах кардинальної зміни наукових парадигм, викликаній цифровізацією науки, масовим розгортанням генеративного штучного інтелекту та переходом до жорстких стандартів відкритих даних. У контексті підготовки магістрів за напрямом 053 «Психологія» ці зміни актуалізують гостру академічну проблему: необхідність подолання традиційного розриву між гуманітарною складовою психологічного знання та вимогами Data Science.

## Криза відтворюваності та виклики для вищої школи

Психологічна наука XXI століття продовжує переживати глибоку методологічну кризу — кризу відтворюваності (*Replication Crisis*), за якої результати класичних емпіричних експериментів не підтверджуються при їхньому незалежному повторенні верифікованими дослідницькими групами. Однією з фундаментальних причин цього феномену визнано недостатню математико-статистичну грамотність дослідників, що породжує латентні маніпуляції з даними:

- **p-hacking** — селективний відбір значущих коефіцієнтів ( $p < .05$ ) при ігноруванні загального масиву нульових гіпотез;
- **HARKing** (*Hypothesizing After the Results are Known*) — формулювання гіпотез після того, як результати вже отримані, що спотворює індуктивно-дедуктивну логіку науки;
- **ігнорування статистичної потужності** — проведення досліджень на мікро-вибірках, де ймовірність виявити реальний психологічний ефект прагне до випадкової;

У зв'язку з цим реформація математичної освіти для студентів гуманітарних спеціальностей стає пріоритетним завданням вищої школи. Студент-магістр повинен сприймати математичний апарат не як формальний бар'єр перед захистом, а як наскрізний інструмент забезпечення внутрішньої та зовнішньої валідності свого дослідження.

### Парадигма Open Science та архітектоніка концептуального апарату

Перехід до постнекласичної наукової парадигми вимагає від вищої освіти впровадження принципів руху «Відкрита наука» (*Open Science*). Сучасний стандарт психологічного дослідження, зокрема за вимогами American Psychological Association (APA 7), базується на наступних умовах:

- пререєстрація дослідницьких протоколів (*Pre-registration*): фіксація гіпотез, дизайну та плану статистичного аналізу в незалежних репозиторіях (наприклад, OSF) до початку збору даних, що повністю виключає феномен HARKing;
- відкриті бази даних (*Open Data*): публікація деідентифікованих «сирих» матриць даних разом із текстом роботи для можливості проведення незалежного аудиту та реплікації;
- строгий синтаксис таблиць та індексів: оформлення статистичних висновків у курсивних шрифтах із точним зазначенням ступенів свободи ( $df$ ), рівнів значущості ( $p$ ) та величини ефекту ( $d$  Коена,  $\eta^2$ ).

Математична підготовка магістрантів має змістити фокус із репродуктивного відтворення формул на розвиток методологічної інтуїції. Дослідник зобов'язаний чітко співвідносити об'єкт, предмет та ієрархію гіпотез. науковому психологічному дослідженню виділяють чотири рівні гіпотез, кожен з яких потребує математичного обґрунтування (Таблиця 1).

Таблиця 1. Ієрархічна матриця наукових гіпотез у психологічному дослідженні

Рівень гіпотези	Сутність та психологічний зміст	Математичне вираження / Інструмент	Приклад формулювання
Теоретична	Припущення про наявність зв'язку між абстрактними конструктами	Якісні концептуальні моделі	Генеративний III дестабілізує професійну ідентичність співробітників
Емпірична	Переклад теорії на рівень конкретних методик та спостережуваних феноменів	Опис операціоналізованих змінних, показники шкал	Показники шкал Опитувальника професійної ідентичності (VIQ за Дж. Марсія) пов'язані з характером візуальних аватарів геймерів
Операціональна	Фіксація конкретних очікувань у вимірюваних балах та кодах якісного аналізу	Оцифровані індекси та метричні шкали	Бал вигорання за MBI буде статистично вищим у співробітників на віддаленій роботі
Статистична	Пари взаємовиключних гіпотез: нульова ( $H_0$ ) та альтернативна $H_1$	Математичні критерії порівняння ( $t$ , $F$ , $\chi^2$ , $r$ )	$H_0: \mu_1 = \mu_2$ (відмінностей немає); $H_1: \mu_1 \neq \mu_2$ (існують значущі відмінності)

## Планування вибірки та апріорний розрахунок потужності (G\*Power)

Наслідком слабкої математичної підготовки магістрів є проєктування обсягу вибірки «зі стелі» або на основі суб'єктивного принципу «чим більше, тим краще». Недобір вибірки веде до помилки другого роду ( $\beta$ -помилка): дослідник приймає хибну нульову гіпотезу і втрачає реальний психологічний ефект через слабкість «математичного ліхтарика». Перебір вибірки є економічно недоцільним і може зробити статистично значущими мікроскопічні, життєво безглузді відмінності. У навчальний процес вищої школи має бути інтегроване обов'язкове освоєння програмного середовища G\*Power. Розглянемо покроковий алгоритм розрахунку обсягу вибірки для двофакторного експерименту зі змішаним планом (дизайн: *Repeated measures, within-between interaction*, що включає порівняння експериментальної та контрольної груп у двох часових точках — до і після психологічного впливу).

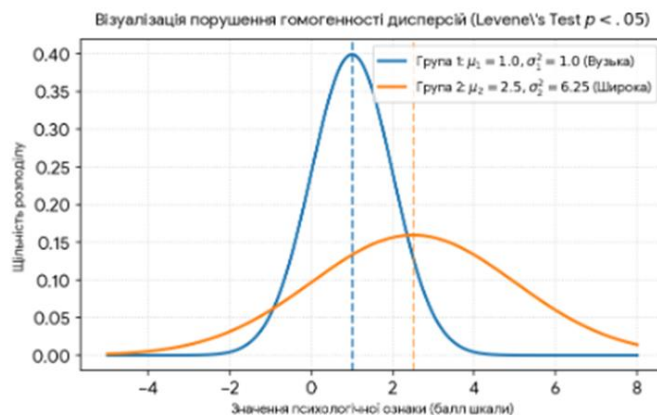
1. Вибір сімейства тестів (*Test family*): *F*-тести
2. Статистичний критерій: *ANOVA: Repeated measures, within-between interaction*
3. Тип аналізу: *A priori: Compute required sample size* — розрахунок обсягу вибірки на етапі планування
4. Фіксація вхідних параметрів (*Input parameters*):
  - Розмір ефекту (*Effect size f*) = 0.25 (консервативний середній рівень за стандартами Коена);
  - Альфа-рівень ( $\alpha$ -помилка) = 0.05 (загальноприйнятий науковий ценз)
  - Статистична потужність ( $1-\beta$ ) = 0.80 (ймовірність зафіксувати ефект, якщо він існує);
  - Кількість груп (*Number of groups*) = 2 (експериментальна та контрольна);
  - Кількість замірів (*Number of measurements*) = 2 (пре-тест і пост-тест).

Програма видає показник *Total Sample Size* (загальний обсяг вибірки), який дослідник зобов'язаний пропорційно розділити між групами. Навчання магістрів цієї процедури повністю страшує наукові проєкти від методологічної сліпоти.

*Верифікація статистичних допущень та подолання помилок.* Грубою помилкою багатьох кваліфікаційних магістерських робіт є автоматичне застосування параметричних критеріїв (таких як, *t*-критерій Стьюдента або ANOVA) до масивів даних без попередньої перевірки математичних допущень. Вища освіта має виробити у студентів жорсткий алгоритм експрес-аудиту даних.

*Характер розподілу ознаки.* Перевірка на нормальність здійснюється не лише візуально, а й з використанням строгих критеріїв: Шапіро-Вілка (для вибірок ( $N < 50$ ) та Колмогорова-Смирнова (для великих вибірок). Якщо розподіл значущо відхиляється від нормального ( $p < .05$ ), параметрична статистика є неприйнятною, і магістр зобов'язаний переходити до непараметричних аналогів (критерії Манна-Витні, Краскела-Валліса, Вілкоксона).

*Гомогенність дисперсій* (критерій Лівена (*Levene's test*)). Перед розрахунком класичного *t*-критерію Стьюдента для незалежних вибірок критично важливо перевірити однорідність розкиду даних у групах за допомогою критерію Лівена. Якщо допущення порушено ( $P_{Levene} < .05$ ), дисперсії значущо відрізняються, дослідник зобов'язаний застосувати критерій Велча (*Welch's t-test*), який автоматично апроксимує ступені свободи (*df*) у менший бік, роблячи критерій більш суворим до помилок першого роду. Нижче представлена візуалізація поведінки щільності розподілу двох вибірок із порушенням допущенням про гомогенність дисперсій (різний розкид навколо середніх значень), що ілюструє суть критерію Лівена.



## Проблема мультиколінеарності в регресійних моделях

При побудові моделей множинної лінійної регресії магістранти часто включають незалежні предиктори змінні, які тісно корелюють між собою (наприклад, «Рівень доходу» та «Суб'єктивне економічне благополуччя»). Це породжує *мультиколінеарність*, що спотворює ваги регресії ( $\beta$ -коефіцієнти) і робить модель математично нестабільною. Для діагностики цього спотворення використовується показник **VIF** (Variance Inflation Factor). Якщо показник VIF перевищує критичний поріг 5 ( $VIF > 5$ ), це свідчить про те, що предиктори дублюють один одного і один з них має бути вилучений або інтегрований через факторний аналіз (EFA) (Таблиця 2)

Таблиця 2. Діагностика мультиколінеарності та регуляризація моделі

Предиктор моделі	Значення VIF	Статус предиктора	Методологічне рішення дослідника
Рівень доходу	1.20	Стабільний (колінеарність відсутня)	Залишається у фінальному рівнянні регресії
Економічне благополуччя	8.45	Критичний (виражена мультиколінеарність)	Дія. Видалити предиктор з моделі або об'єднати його з доходом через факторний аналіз (EFA)

## Парадигма Mixed Methods та квантифікація якісних даних

Сучасне психологічне дослідження вийшло за межі виключно кількісних тестів. Міжнародним золотим стандартом став дизайн змішаних методів (*Mixed Methods Designs*). Він вимагає інтеграції стандартизованої психометрики та якісних джерел (дискурсивно – графічні тести, нарративні інтерв'ю, пострисункові бесіди). Якісні дані не повинні бути просто ілюстрацією («Студент А намалював сумний малюнок»); вони переводяться в жорсткі матриці кодів. Виклик для вищої школи — навчити магістрів проводити *квантифікацію (оцифровку)* якісного контенту.

*Контроль суб'єктивної помилки (Каппа Коена).* Для верифікації даних дослідження залучаються незалежні експерти. Рівень їхньої узгодженості вимірюється коефіцієнтом Каппа Коена ( $k$ ). Магістр повинен розуміти градацію цього індексу:

$k < 0.60$  — незадовільний рівень, висока частка суб'єктивної помилки; дані використовувати в статистиці заборонено;

$0.60 \leq k < 0.70$  — помірний рівень, що вимагає проведення дебрифінгу експертів та перекодування спірних випадків;

$k \geq 0.70-0.75$  — високий науковий ценз, що дозволяє включати оцифровані коди до матриць багатовимірного аналізу;

Після успішної квантифікації якісна ознака (наприклад, дихотомічний код: 0 — бар'єра на малюнку немає, 1 — бар'єр є) з'являється з метричним інтервальним балом тесту за допомогою точково-бісеріального коефіцієнта кореляції  $rpb$ , забезпечуючи справжню триангуляцію даних.

## Інтеграція штучного інтелекту та етичний протокол

Стрімкий вибух генеративного ШІ (LLM-моделей: ChatGPT, Claude, Gemini) у 2022–2026 роках поставив вищу школу перед вибором: забороняти технології чи очолити їхню інтеграцію. Загальні заборони неефективні та ведуть до академічного регресу. Стратегічний крок сучасної освіти — навчання магістрів легітимним методам *AI-assisted research* та методології *Prompt Engineering* (зокрема технології *Few-Shot Prompting*). При аналізі великих масивів текстів ШІ виступає як асистент для розвідувального (*exploratory*) контент-аналізу за умови чіткого кодування та виведення результатів у форматі CSV-таблиць. Зважаючи на те, що психологічні дослідження магістрів вивчають як взаємодію людини з технологіями, так і розробляють корекційно-реабілітаційні програми, спрямовані на відновлення психічного здоров'я (наприклад, після бойових травм, втрат або ізоляції), використання ШІ тут вимагає найвищого рівня обережності. Тут ШІ виступає і як об'єкт дослідження, і як інструмент реабілітації, що подвоює етичні ризики для магістра. Заклади вищої освіти зобов'язані впроваджувати жорсткий етичний протокол використання ШІ, який розмежовує легітимну роботу та академічне шахрайство (Таблиця 3).

Таблиця 3. Етичний комплаєнс та дихотомія застосування інструментів генеративного ШІ в академічному середовищі

Легітимні методи (дозволено)	Грубі порушення (заборонено)
Первинний аналіз методик: ШІ допомагає підібрати валідні опитувальники під об'єкт дослідження.	Генерація відповідей респондентів: симуляція ШІ заповнення тестів замість реальної вибірки людей.
Обробка відкритих відповідей: класифікація великої кількості якісних текстових відповідей респондентів за категоріями.	Порушення анонімності: завантаження в ШІ реальних протоколів консультацій чи ПІБ досліджуваних.
Написання коду для SPSS/R: генерація скриптів для розрахунку кореляцій, Т-критерію чи регресії.	ШІ-діагностика: використання ШІ для встановлення психологічних діагнозів чи висновків без перевірки психолога.
Огляд теорій та брейнштормінг: структурування підходів різних психологічних шкіл, пошук патернів та сіток кодів.	Підгонка результатів та обхід систем: прохання до ШІ "змінити цифри", прихована фабрикація даних, обхід систем «Антиплагіат».
Адаптація кейсів та VR-сценаріїв: створення гіпотетичних описів клієнтських випадків, генерація описів віртуальних середовищ для терапії.	Заміна психолога алгоритмом та ретригер: проведення реабілітаційних сесій через ШІ з реальними особами без нагляду лікаря, неконтрольований ретригер травми.
Аналіз поведінки та маркерів: обробка знеособленої статистики геймерів, пошук депресивних маркерів у закритих базах даних.	Ігнорування "цифрового розриву" та завантаження медичних карток: внесення до ШІ історій хвороб (МКХ-10), видача онлайн-рекомендацій без урахування відсутності доступу до технологій у пацієнтів.

Ключовими зонами особливої відповідальності магістра-психолога при роботі з ШІ є: обов'язкове декларування за правилами *AI Disclosure* (із зазначенням назви моделі, версії та характеру запитів), дотримання принципу *Human-in-the-loop* (верифікація людиною), суворя заборона на завантаження терапевтичних сесій у «хмари», збереження критичного мислення щодо «ідеальної» статистики та адаптація висновків під реальний контекст вибірки вручну.

### Метод кейс-завдань як основа контролю якості освіти

Перехід від репродуктивної моделі освіти («зазубрювання формул») до компетентнісної вимагає радикального оновлення екзаменаційних матеріалів. Традиційні білети мають бути замінені банком наскрізних методологічних кейс-завдань. Екзаменаційний білет сучасного магістра-психолога бажано розробляти як трикомпонентну структуру, що містить питання теоретико-методологічних основ (Блок 1), інструментально-статистичний апарат (Блок 2) та практичну кейс-задачу з реальним контекстом (Блок 3). З метою комплексного контролю сформованості математичних компетенцій у магістрів до екзаменаційного комплексу інтегровано завдання з психометричного аудиту та аналізу якісних ознак. Концептуальна інтеграція банку екзаменаційних кейс-завдань до системи підготовки реалізується через наскрізне поєднання теоретичних знань магістрів-психологів із прикладними задачами цифрової психометрії та кіберпсихології. У межах науково-методичного обґрунтування цієї моделі банк екзаменаційних кейсів виступає не лише інструментом фінального контролю, а й об'єктивним маркером здатності випускника розв'язувати реальні дослідницькі задачі. Практична реалізація зазначеного підходу у структурі статті ілюструється двома базовими типами комплексних завдань:

*Кейси з психометричного аудиту цифрового інструментарію.* У межах цього блоку магістрам пропонується завдання на верифікацію надійності та валідності нових онлайн-тестів або діагностичних ШІ-додатків. Наприклад, у кейсі «Психометричний аудит надійності (*a* Кронбаха)» магістрант проводить пілотну валідизацію нового опитувальника цифрового стресу з 10 пунктів ( $N = 30$ ). Первинний розрахунок показує незадовільну ( $a = .52$ ), але в таблиці виводу SPSS «Item-Total Statistics» у стовпці «Cronbach's Alpha if Item Deleted» навпроти пункту №7 стоїть значення .79, а в стовпці «Corrected Item-Total Correlation» — значення -0.24. Студент має надати аргументовану відповідь щодо математичної причини низької надійності (наприклад, неперекодована зворотна шкала) та описати практичні дії для покращення властивостей методики.

*Кейси з математичного аналізу якісних (нечислових) ознак.* Цей напрям орієнтований на обробку текстової та категоріальної інформації. Типовим прикладом є завдання на статистичний аналіз контенту соціальних мереж: магістр має категоризувати якісні маркери психологічного вигорання або

агресії у текстових повідомленнях користувачів, перевести ці дані у номінативну шкалу та математично довести наявність зв'язку між ознаками за допомогою критерію  $\chi^2$  Пірсона.

Комплексна архітектура підсумкового контролю, що включає інноваційні екзаменаційні кейси, дозволяє перевести атестацію магістрів-психологів із площини формальної перевірки знань у площину реальної оцінки їхньої професійної спроможності. Узагальнення описаного нами науково-методичного досвіду, а також аналіз результатів його практичного застосування на кафедрі, дають підстави для формулювання низки узагальнюючих висновків. Окрім суто академічного ефекту, запропонований підхід має чітко окреслену практичну значущість для модернізації вищої психологічної освіти в умовах цифровізації науки. Підсумовуючи вищесказане, детально окреслимо висновки щодо ефективності моделі та прикладні рекомендації для колег-освітян, які прагнуть оптимізувати процес оцінювання дослідницьких компетенцій магістрів.

### Висновки та практична значущість

1. Орієнтація на подолання кризи відтворюваності результатів: головним вектором модернізації оцінювання у вищій освіті визначено відмову від пасивного репродуктивного відтворення статистичних формул на користь розвитку методологічної інтуїції магістрів. Запропонована архітектура комплексних кейс-завдань навчає студентів суворого дотримання математичних допущень критеріїв і повністю відповідає міжнародним стандартам руху Відкритої науки (*Open Science*).
2. Валідність методології змішаних методів (*Mixed Methods*): на прикладі розроблених екзаменаційних комплексів обґрунтовано, що якісні, дискурсивні та графічні методи кіберпсихологічного аналізу є не просто ілюстративним матеріалом у наукових роботах, а мають проходити обов'язкову процедуру квантифікації (переведення у номінативні чи порядкові шкали) для подальшого багатовимірного статистичного аналізу.
3. Генеративний ШІ як інноваційний інструмент дидактики: штучний інтелект визнано легітимним асистентом магістра-дослідника на етапі експлораторного (розвідувального) аналізу даних та первинної класифікації якісних ознак. Проте його залучення є припустимим виключно за умови обов'язкового прозорого декларування правил використання (*AI Disclosure* відповідно до вимог АРА 7) та фінальної верифікації всіх результатів людиною (модель *Human-in-the-loop*).
4. Технологізація та оптимізація викладацької практики: розроблений трикомпонентний банк екзаменаційних кейсів дозволяє мінімізувати суб'єктивізм та перевести підсумкову атестацію у площину прозорого аудиту дослідницьких компетенцій. Описаний досвід репрезентує гнучкий методичний конструктор для викладачів інших кафедр психологічного профілю, оскільки дозволяє адаптувати та масштабувати завдання під будь-яку прикладну тематику без створення додаткового бюрократичного навантаження на професорсько-викладацький склад.

### СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Дзьобань О. П. Методологія, організація та технологія наукових досліджень : навч. посіб. для аспірантів /; ДНУ «Ін-т інформації, безпеки і права Нац. академії прав. наук України». Київ; Одеса: Фенікс, 2025. 284 с.
2. Методика та організація наукових досліджень з психології: навчальний посібник для студентів спеціальності 053 Психологія /Укладачі: В.Й. Бочелюк, В.В. Нечипоренко, О.Л. Позднякова, Н.Є.Завацька, Ю.А. Завацький, І.Є. Жигаренко, В.Ю. Завацький. [3-ге вид., перероб. і доп.]. Київ: КНТ, 2024. 384 с.
3. Освітні вимірювання та статистичні методи обробки даних [Електронний ресурс] : електрон. метод. рекомендації до практичних занять з курсу для здобувачів спеціальності 011 "Освітні, педагогічні науки" / уклад.: В. В. Павлова, Н. В. Нагорна. Одеса, 2023. 33 с. URL: <https://nupp.edu.ua/page/osnovi-matematichnoi-statistiki-dlya-psikhologiv-053.html>
4. Стиль АРА: вимоги до оформлення списку використаних джерел кирилицею та латиницею в наукових виданнях : метод. рек. / уклад.: М. П. Гребенюк, Г. З. Шевчук, Л. І. Гаврилюк, Ю. А. Кліванська, Р. А. Ткачук. Луцьк : ВІППО, 2025. 68 с.
5. Crayne, M. P., & Medeiros, K. E. (2024). Making sense of the machine: Professional identity transformation in the age of AI. *Journal of Vocational Behavior*, 148, Article 103952. [doi.org](https://doi.org/10.1016/j.jvb.2024.103952)

6. URL: <https://osf.io> (Міжнародний репозиторій Відкритої науки).G\*Power Official Software & Manual (HHU Düsseldorf)
7. URL: <https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower> G\*Power. Statistical Power Analyses for Mac and Windows

***Костюніна Олена Володимирівна*** – кандидат психологічних наук, доцент кафедри кіберпсихології та реабілітації, Державна академія інтелектуальних технологій та зв'язку, м.Одеса, email : [favoritelena09@gmail.com](mailto:favoritelena09@gmail.com)

**Olena V. Kostyunina** – PhD in Psychological Sciences, Associate Professor of the Department of Cyberpsychology and Rehabilitation, State Academy of Intellectual Technologies and Communications, Odesa, Ukraine, email : [favoritelena09@gmail.com](mailto:favoritelena09@gmail.com)