

ВАСЮКОВА Наталія – кандидат філологічних наук, начальник навчально-наукового центру іноземних мов, Національний університет оборони України, просп. Повітряних сил, 28, м. Київ, 03049, Україна

ORCID: <https://orcid.org/0000-0002-6184-3314>

КРИКУН Вікторія – кандидат педагогічних наук, старший дослідник, начальник науково-дослідного відділу мовного тестування наукового центру мовного тестування навчально-наукового центру іноземних мов, Національний університет оборони України, просп. Повітряних сил, 28, м. Київ, 03049, Україна

ORCID: <https://orcid.org/0000-0001-7409-3383>

ResearcherID: <https://www.webofscience.com/wos/author/record/2398821>

ГРИЩУК Юлія – доктор філософії, старший науковий співробітник наукового центру мовного тестування навчально-наукового центру іноземних мов, Національний університет оборони України, просп. Повітряних сил, 28, м. Київ, 03049, Україна

ORCID: <https://orcid.org/0000-0002-3999-3420>

КРАВЧУК Андрій – молодший науковий співробітник наукового центру мовного тестування навчально-наукового центру іноземних мов, Національний університет оборони України, просп. Повітряних сил, 28, м. Київ, 03049, Україна

ORCID: <https://orcid.org/0009-0007-6897-0790>

DOI: <https://doi.org/10.24919/2413-2039.19/51.1>

Бібліографічний опис статті: Васюкова, Н., Крикун, В., Грищук, Ю., Кравчук, А. (2024). Експериментальна перевірка результативності методики комп'ютерного адаптивного мовного тестування відповідно до вимог НАТО STANAG 6001. *Людинознавчі студії. Серія «Педагогіка»*. № 19 (51), 9–17, doi: <https://doi.org/10.24919/2413-2039.19/51.1>

ЕКСПЕРИМЕНТАЛЬНА ПЕРЕВІРКА РЕЗУЛЬТАТИВНОСТІ МЕТОДИКИ КОМП'ЮТЕРНОГО АДАПТИВНОГО МОВНОГО ТЕСТУВАННЯ ВІДПОВІДНО ДО ВИМОГ НАТО STANAG 6001

Анотація. У статті розглядається експериментальна перевірка результативності методики комп'ютерного адаптивного мовного тестування (КАМТ) відповідно до стандарту НАТО STANAG 6001. Актуальність дослідження обумовлена необхідністю оптимізації процесу мовного тестування для забезпечення об'єктивного визначення рівня володіння іноземною мовою військовослужбовців Збройних Сил України з рецептивних видів мовленнєвої діяльності (читання та аудіювання) відповідно до вимог стандарту НАТО STANAG 6001 та швидкого оброблення результатів мовного тестування. Методика КАМТ, що передбачає адаптацію рівня складності тестових завдань залежно від відповідей тих, хто тестується, була апробована на вибірці слухачів Національного університету оборони України протягом 2023–2024 років. Результати експерименту засвідчили високий рівень валідності, достовірності та об'єктивності отриманих результатів за методикою КАМТ, що було підтверджено методами математичної статистики. З метою підтвердження гіпотези дослідження було також проведено кореляційний аналіз щодо наявності високих прямих кореляційних зв'язків між тестуванням за двома методиками – КАМТ та паперового тестування. Аналіз даних показав сильний позитивний кореляційний зв'язок між результатами читання, проведеного із використанням паперового тестування та результатами читання за допомогою методики КАМТ. У ході дослідження було виявлено низку переваг методики КАМТ, а саме: зменшення середнього часу тестування, прискорення процесу оцінювання результату тесту, контроль точності вимірювання результатів всіх учасників тестування, створення унікального тесту завдяки алгоритму підбору тестових завдань, забезпечення адаптивності алгоритму,

який сприяє гнучкості і пристосуванню мовного тесту до індивідуального рівня підготовки кожного учасника тестування. Таким чином методика КАМТ є результативним інструментом оцінювання рівня мовленнєвої компетентності військовослужбовців.

Ключові слова: комп'ютерне адаптивне мовне тестування, педагогічний експеримент, методика, адаптивний підхід, військова освіта, офіцер, іншомовна компетентність, лінгвістичні принципи.

VASIUKOVA Nataliia – Phd (Philology), Chief of Foreign Languages Education and Research Centre, National Defence University of Ukraine, 28 Povitrianyh Syl Ave., Kyiv, 03049, Ukraine

ORCID: <https://orcid.org/0000-0002-6184-3314>

KRYKUN Viktoriia – Phd (Education), Chief of the Scientific Language Testing Department of the Language Testing and Research Centre of the Foreign Languages Education and Research Centre, National Defence University of Ukraine, 28 Povitrianyh Syl Ave., Kyiv, 03049, Ukraine

ORCID: <https://orcid.org/0000-0001-7409-3383>

ResearcherID: <https://www.webofscience.com/wos/author/record/2398821>

HRYSCHUK Yuliia – Phd (Education) Senior Research Fellow at the Language Testing and Research Centre of the Foreign Languages and Education and Research Centre, National Defence University of Ukraine, 28 Povitrianyh Syl Ave., Kyiv, 03049, Ukraine

ORCID: <https://orcid.org/0000-0002-3999-3420>

KRAVCHUK Andrii – Junior Research Fellow at the Language Testing and Research Centre of the Foreign Languages Education and Research Centre, National Defence University of Ukraine, 28 Povitrianyh Syl Ave., Kyiv, 03049, Ukraine

ORCID: <https://orcid.org/0009-0007-6897-0790>

DOI: <https://doi.org/10.24919/2413-2039.19/51.1>

To cite this article: Vasiukova, N., Krykun, V., Hryshchuk, Yu., Kravchuk, A. (2024). Eksperymentalna perevirka rezultatyvnosti metodyky kompiuternoho adaptivnoho movnoho testuvannia vidpovidno do vymoh NATO STANAG 6001 [Experimental verification of the effectiveness of the computer adaptive language testing methodology in accordance with the requirements of NATO STANAG 6001]. *Human Studies. Series of Pedagogy*, № 19 (51), 9–17, doi: <https://doi.org/10.24919/2413-2039.19/51.1>

EXPERIMENTAL VERIFICATION OF THE EFFECTIVENESS OF THE COMPUTER ADAPTIVE LANGUAGE TESTING METHODOLOGY IN ACCORDANCE WITH THE REQUIREMENTS OF NATO STANAG 6001

Summary. The article deals with the experimental verification of the effectiveness of the Computer Adaptive Language Testing (CALT) methodology in accordance with NATO standard STANAG 6001. The relevance of the study is stipulated by the need to optimize the language testing process to ensure an objective determination of the level of foreign language proficiency of the military personnel of the Armed Forces of Ukraine in receptive skills (reading and listening) in accordance with the requirements of NATO STANAG 6001 and to quickly process the results of language testing. The CALT methodology, which provides for the adaptation of the level of difficulty of test tasks depending on the answers of the test takers, was tested on students of the National Defence University of Ukraine in 2023–2024. The results of the experiment showed a high level of validity, reliability, and objectivity of the results obtained using the CALT methodology, which was confirmed by the methods of mathematical statistics. In order to confirm the research hypothesis, a correlation analysis was conducted to determine whether there were high direct correlations between the two methods – CALT and paper-based testing. The data analysis showed a strong positive correlation between the results of paper-based reading and the results of CALT. The study revealed a number of advantages of the CALT methodology, namely: reducing the average testing time, speeding up the process of evaluating the test result, controlling the accuracy of measurement of all test participants, creating a unique test due to the algorithm for selecting test tasks, ensuring the adaptability of the algorithm, which provides flexibility and adaptation of the language test to the individual level of training of each test participant. Thus, the CALT methodology is an effective tool for assessing the level of language competence of military personnel.

Key words: computer adaptive language testing, pedagogical experiment, methodology, adaptive approach, military education, officer, foreign language competence, linguistic principles.

Вступ. Питання оптимізації процесу мовного тестування під час вивчення іноземних мов та під час складання підсумкового тесту є дуже важливим з огляду на зростаючу кількість тестованих у Збройних Силах України (далі – ЗС України) та необхідності швидкої обробки результатів тестування. Наразі постає питання не тільки забезпечення високого рівня валідності мовних тестів та надійності результатів мовного тестування, а й швидкої обробки результатів тестування, доступність тесту в будь-якому місці з підключенням до мережі Інтернет, багаторівневість тесту та скорочення часу на його проходження. З огляду на зазначене вище, робимо припущення, що методика комп'ютерного адаптивного мовного тестування відповідно до вимог НАТО STANAG 6001 (далі – КАМТ) може розглядатися, як рішення окреслених вище завдань.

Аналіз останніх досліджень. Особливості комп'ютерного адаптивного тестування у своїх наукових розвідках висвітлили David J. Weiss, Alper Sahin (2024) (теоретико-практичні аспекти комп'ютерного адаптивного мовного тестування), M. Izadi, F. Heidari (2024) (використання тьютора на основі адаптивного принципу з метою діагностування та оцінювання компетентності з аудіювання), N. Sherkuziyeva (2023) (порівняльний аналіз комп'ютерного адаптивного тестування та традиційного). Українські вчені також досліджували питання адаптивного тестування. Відтак, О. Радкевич (2023) у своїх працях розглядав актуальні аспекти адаптивного тестування в контексті використання електронних засобів навчання. Своєю чергою, І. Постоленко (2024) зосередилася на дослідженні використання комп'ютерного тестування для контролю і оцінювання знань з англійської мови; Я. Сікора (2023) надала інструментарій створення адаптивного тесту. Незважаючи на низку наукових досліджень, які висвітлюють особливості комп'ютерного адаптивного мовного тестування, залишилися поза увагою тести з іноземної мови професійного спрямування, зокрема тести відповідно до вимог НАТО STANAG 6001.

Мета статті – експериментальна перевірка результативності методики комп'ютерного адаптивного мовного тестування відповідно до стандарту НАТО STANAG 6001. Завдання статті полягають у 1) висвітленні методики комп'ютерного адаптивного мовного тестування відповідно до стандарту НАТО STANAG 6001 та 2) висвітлення результатів

експерименту, метою якого є перевірка її результативності.

Виклад основного матеріалу. Методика КАМТ є різновидом комп'ютерного тестування, під час якого послідовність подання тестових завдань (їхня складність), а також кількість завдань, залежать від відповідей того, хто тестується на попередні тестові завдання. Тобто КАМТ враховує рівень розвитку іншомовних компетентностей того, хто тестується під час процедури тестування. Зокрема, під час КАМТ той, хто тестується, отримує спочатку тестове завдання середньої складності, яке відповідає стандартизованому мовленнєвому рівню 2 відповідно до вимог НАТО STANAG 6001 (далі – СМР 2). Якщо він/вона обирає правильну відповідь на визначену кількість завдань, тоді його/її наступне завдання буде складнішим, якщо відповіді неправильні – простішим, доки програма КАМТ не визначить остаточний рівень того, хто тестується. КАМТ, на думку зарубіжних вчених, має багато переваг щодо оцінювання рівня володіння іноземною мовою, зокрема: наявність інструменту високої точності тестування, економія часу, різні варіанти оцінювання в залежності від поставлених цілей.

Розроблена методика КАМТ ґрунтується на таких принципах:

- науковості, що полягає в організації проведення КАМТ на засадах використання досягнень сучасної науки, сучасних інформаційно-комунікаційних технологій навчання, якісного науково-методичного забезпечення та науково обґрунтованих методик і технологій тестування з іноземної мови;

- системності, що передбачає здійснення процесу КАМТ на основі єдиної нормативної бази, чітко визначених потреб, плановості заходів, стандартизації методики викладання;

- послідовності, який передбачає, що кожний наступний рівень діагностування володіння іноземною мовою базується на досягненні попереднього рівня;

- службової необхідності, що полягає у відповідності змісту та обсягу мовної підготовки конкретним потребам ЗС у підготовці особового складу, який володіє іноземною мовою;

- суб'єкт-суб'єктної взаємодії, який полягає в тому, що всі учасники процесу мовного тестування є рівноправними суб'єктами діяльності. Під час мовного тестування відбувається їх міжособистісна взаємодія на основі активного діалогу, співпраці та взаємопорозуміння. Цей принцип передбачає орієнтування

на того, хто тестується, як на суб'єкта іншомовної підготовки з урахуванням його можливостей, інтересів, потреб, індивідуально-психологічних особливостей.

Також методика КАМТ передбачає наявність лінгвістичних принципів, а саме:

- фахової спрямованості, який передбачає, що комп'ютерне адаптивне мовне тестування повинне спрямовуватися на перевірку у тих, хто тестується, здатності й готовності виконувати чітко визначені завдання, які постають під час їхньої військово-професійної діяльності;

- комунікативності, відповідно до якого необхідно використовувати тестові матеріали, які стимулюють тих, хто тестується до активного міжособистісного суб'єкт-суб'єктного усного та письмового спілкування іноземною мовою;

- автентичності, який полягає в тому, що зразки усного та письмового мовлення та мовний інвентар, які є оригінальними, тобто створеними носіями англійської мови для інших носіїв із повним збереженням їх лінгвістичних особливостей;

- комплексності (інтегративності), який забезпечує цілісність КАМТ з іноземною мовою, що проявляється у оцінюванні рівня розвитку компетентностей тестованих в окремих видах мовлення, а не в опануванні окремих мовних аспектів (лексики, граматики, вимови тощо).

Алгоритм КАМТ відповідно до вимог НАТО STANAG 6001 складається з блоків (Gawliczek, Krykun, Tarasenko, Tyshchenko & Shapran, 2021):

- 1) вихідної точки (the Starting Point block), що відповідає за створення набору тестових завдань окремо для СМР 1, СМР 2, СМР 3 відповідно до вимог НАТО STANAG 6001;

- 2) блоку представлення тестових завдань (Item Selection Algorithm), який безпосередньо відповідальний за адаптацію тестових завдань залежно від відповіді тестованого;

- 3) блоку оцінювання та завершення тестування (the Scoring and Termination Criterion), функція якого полягає у визначенні остаточного рівня мовленнєвої компетентності того, хто тестується, та завершенні мовного тестування.

КАМТ проводиться у межах стандартизованих мовленнєвих рівнів, які відповідають рівням мовленнєвої компетентності за мовним стандартом НАТО STANAG 6001, а саме: СМР 0 (Відсутність практичного вміння); СМР 0+ (Завчений мінімум); СМР 1

(Виживання); СМР 1+ (Виживання +); СМР 2 (Функціональний), СМР 3 (Професійний). КАМТ складається з двох субтестів за основними видами мовленнєвої діяльності і проводиться у такій послідовності: аудіювання та читання. Перед початком КАМТ і перед виконанням кожного субтесту проводиться інструктаж, виконання вимог якого для всіх є обов'язковим.

Алгоритм формулювання завдань за КАМТ будується за принципом зворотного зв'язку, коли при правильній відповіді того, хто тестується, наступне завдання пропонується вищого рівня складності, а неправильна відповідь вказує на необхідність пред'явлення більш легкого тестового завдання, у порівнянні з тим, на яке було дано неправильна відповідь. Якщо той, хто тестується, відповідає на запитання правильно, тоді наступне запитання пропонується складнішого рівня. І навпаки, якщо той, хто тестується, відповідає неправильно, тоді наступне запитання буде легшим.

Отже, варіанти розгортання алгоритму можна зобразити у такий спосіб:

- *варіант 1 алгоритму* – той, хто тестується, отримує СМР3 за умови, що він починає тестування із СМР 2, надає 11 правильних відповідей цього рівня, переходить до виконання завдань СМР3 і також надає 11 правильних відповідей. Можлива кількість помилок – 4 на кожному рівні;

- *варіант 2 алгоритму* – той, хто тестується, отримує СМР2+ за умови, що він надає 11 правильних відповідей рівня СМР2, на СМР3 – 8–10 правильних відповідей. На рівні СМР2 кількість помилок може сягати 4, на СМР3 – від 5–7;

- *варіант 3 алгоритму* – той, хто тестується, отримує СМР2 за умови, що він надає 11 правильних відповідей цього рівня та від 0 до 7 правильних відповідей СМР3;

- *варіант 4 алгоритму* – той, хто тестується, отримує СМР1+ за умови, що на СМР2 він допускає 5 помилок, потім виконує завдання на СМР1, де надає 11 правильних відповідей, що дозволяє йому повернутися на рівень СМР2, де він надає від 8 до 7 правильних відповідей. Кількість помилок на рівні СМР2 – 7, на СМР1 – 0–4;

- *варіант 5 алгоритму* – той, хто тестується, отримує СМР1+ за умови, що на СМР2 надає 8–10 правильних відповідей та 11 правильних – на СМР 1. На СМР2 він може допустити 5 помилок, на СМР 1, на СМР2 від 0 до 4;

- *варіант 6 алгоритму* – той, хто тестується, отримує СМР1 за умови, що на рівень

СМР2 він робить 5 помилок, потім надає 11 правильних відповідей на рівень СМР1, після чого в нього є можливість повернутися на рівень СМР2, де він надає тільки 0–7 правильних відповідей. На СМР2 той, хто тестується, робить щонайменше 8 помилок. На СМР1 – 0–4 помилки;

– *варіант 7 алгоритму* – той, хто тестується, отримує СМР0+ за умови, що він робить 5 помилок на рівні СМР2, потім надає тільки 8–10 правильних відповідей на СМР1;

– *варіант 8 алгоритму* – той, хто тестується, отримує СМР0+ за умови, що він робить 5 помилок на рівні СМР2, потім надає тільки 0–7 правильних відповідей на СМР1. Кількість помилок на СМР1 – 8.

З метою визначення результативності методики КАМТ у навчально-науковому центрі іноземних мов Національного університету оборони України проводився педагогічний експеримент протягом 2023–2024 рр. і передбачало три етапи його проведення:

1 *етап* – *підготовчий*, під час якого було здійснено теоретичний аналіз наукової літератури з метою удосконалення методики КАМТ (читання та аудіювання).

2 *етап* – *основний* – апробація методики КАМТ, в основу якого було покладено послідовний план тестування слухачів за двома методиками: паперового тестування (із використанням стандартизованого мовного тесту відповідно до стандарту НАТО STANAG 6001) та методики КАМТ. Тестування за методикою КАМТ відбулось 24 червня, а паперове – 26 червня 2024 року. На даному етапі з метою обґрунтування результативності методики КАМТ було здійснено порівняльний аналіз результатів тестування, проведених з використанням методики КАМТ та паперового тестування за рецептивними видами мовленнєвої діяльності (читання та аудіювання) відповідно до вимог стандарту НАТО STANAG 6001 (рис. 1).

3 *етап* – *впроваджувальний*, сутність якого передбачає підтвердження H_0 – нульової гіпотези того, що статистично значущої різниці між отриманими результатами тестувань за двома методиками – КАМТ та паперового тестування – немає.

Вибіркою для проведення педагогічного експерименту стали 30 слухачів, які навчались протягом чотирьох місяців з січня по квітень 2024 року на курсах іноземних мов у навчально-науковому центрі іноземних мов Національного університету оборони України. Під час дослідження слухачі пройшли обидва

види тестування, що дало змогу всебічно оцінити ефективність і надійність обох методик.

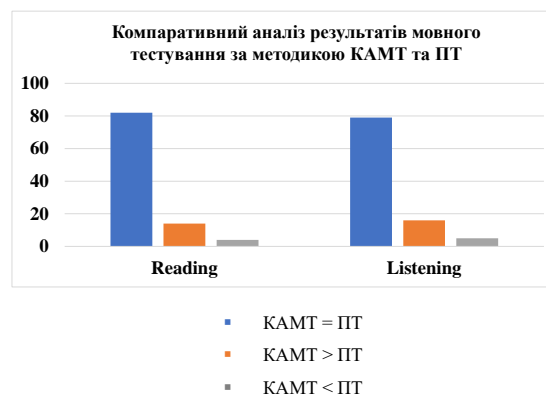


Рис. 1. Компаративний аналіз результатів мовного тестування за методикою КАМТ та ПТ

Результати КАМТ показали, що 70% учасників тестування підтвердили свій рівень володіння англійською мовою з аудіювання (21 особа) та 76,6% – з читання (23 особи). Це свідчить про валідність та надійність тестування за методикою КАМТ. Порівняльний аналіз виявив лише незначну розбіжність результатів між паперовим тестуванням та КАМТ: 16,6% учасників тестування продемонстрували вищі результати з аудіювання, а 13,3% – з читання під час комп'ютерного тестування. Така розбіжність пояснюється тим, що під час КАМТ учасники тестування мали можливість отримати вищі результати, ніж СМР 2. Натомість 10% (3 особи) з читання, та 13,3% (4 осіб) з аудіювання продемонстрували нижчий СМР, які тестувалися КАМТ. Такі дані пояснюються різним рівнем підготовленості учасників тестування до КАМТ [10] (рис. 2).

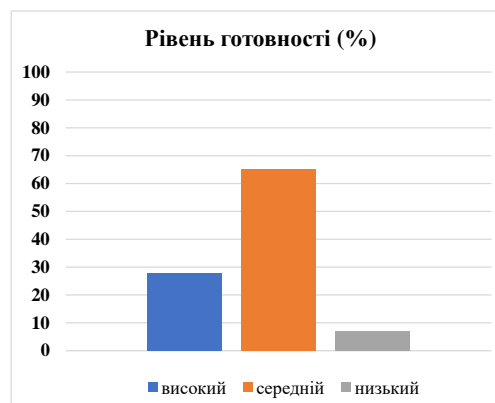


Рис. 2. Результати опитування щодо готовності слухачів курсів іноземних мов щодо проходження мовного тестування за методикою КАМТ

Аналіз відповідей респондентів та їх порівняння з результатами КАМТ і паперового тестування показує, що 28% учасників тестування мали досвід проходження комп'ютерного тестування, що позитивно вплинуло на їх результати під час КАМТ. Водночас, 7% учасників продемонстрували розгубленість під час сесії КАМТ через низький рівень комп'ютерної грамотності. Вони не звикли до роботи з комп'ютерами, що могло збільшити час на навігацію та взаємодію з тестовою системою і вплинути на їхню продуктивність. Окрім того, алгоритм КАМТ передбачає зміну складності питань залежно від відповідей кандидата, що може створити додатковий тиск та відчуття невпевненості.

Порівнюючи результати експерименту, можна стверджувати, що методика КАМТ забезпечує не менш результативне визначення рівня володіння іноземною мовою у рецептивних видах мовленнєвої діяльності (читання та аудіювання) відповідно до вимог стандарту НАТО STANAG 6001, як і методика паперового тестування. Більш того, результати експерименту свідчать про дотримання валідності, достовірності та об'єктивності отриманих результатів. Одночасно можна стверджувати, що застосування методики КАМТ уможливило скорочення тривалості проведення мовного тестування на 50 або більше відсотків. Це можна пояснити низкою переваг методики КАМТ:

- завдяки адаптивній функції алгоритму КАМТ, середній час тестування зменшився на 22%, тоді як метод паперового тестування мав часові обмеження та фіксований набір завдань;

- процес оцінювання результатів тесту значно прискорився завдяки здатності алгоритму автоматично визначати рівень володіння іноземною мовою відповідно до вимог стандарту НАТО STANAG 6001;

- однією з переваг КАМТ є контроль точності вимірювання. КАМТ може вимірювати всіх учасників тестування з однаковим ступенем точності, що забезпечує об'єктивну оцінку результатів. Усі учасники тестування отримали однаковий рівень точності результатів, але з індивідуальним набором завдань, що дозволяє краще адаптувати тест до їхнього рівня підготовленості;

- створення унікального тесту є ще однією суттєвою перевагою комп'ютерного адаптивного мовного тестування (КАМТ). Вважаємо, що «використання банку завдань, відкаліброваних на основі теорії відповідей на завдання, забезпечує алгоритму КАМТ оптимальний варіант тесту.

Банк завдань і алгоритм КАМТ гарантують, що кожен учасник тестування отримає оптимальний тест. Алгоритм вибирає завдання з банку завдань відповідно до здібностей учасника тестування, визначених на основі його відповідей під час тестування» (Hryshchuk & Krykun, 2022);

- адаптивна здатність алгоритму КАМТ, який є дуже гнучким і пристосовується до компетентності учасників тестування під час сесії. Було виявлено вищий рівень мотивації у кандидатів тестування, які мали вищий рівень володіння іноземною мовою, оскільки вони не витрачали час на виконання легких завдань. Водночас учасники тестування з нижчим рівнем володіння іноземною мовою не втрачали мотивацію через занадто складні завдання. Іншими важливими перевагами методики КАМТ є негайне повідомлення про результати та ефективне управління тестуванням (реєстрація, збір даних, створення банку завдань і бази даних, звітність тощо).

З метою підтвердження результативності методу КАМТ для визначення СМР відповідно до стандарту НАТО STANAG 6001 та для підтвердження того, що статистично значущої різниці між отриманими результатами тестувань за двома методиками – КАМТ та паперового тестування – немає, на завершальному етапі педагогічного експерименту було застосовано статистичні методи аналізу.

У процесі вибору статистичних критеріїв оцінювання результативності критерію спиралися на класифікацію завдань дослідження і визначення потужності критерію (його здатності виявляти розбіжності, якщо вони є). Крім того, підставами для вибору критерію були його простота, широта діапазону застосування (на малих вибірках), можливість застосування на нерівних за обсягом вибірках. Алгоритм вибору статистичних критеріїв відповідав загальній класифікації завдань дослідження і методів їх вирішення.

З урахуванням того, що вибірка дослідження є малою відносно генеральної сукупності, а також того, що за результатами оцінювання відповідності емпіричного розподілу нормальному закономірному методом перевірки статистичних гіпотез із використанням λ -критерію Колмогорова-Смірнова показники майже третини шкал не відповідають нормальному розподілу (Гауса-Пуассона), було обрано для статистичної обробки даних непараметричний критерій Вілкоксона (Wilcoxon test), який застосовується для зіставлення показників, виміряних в двох різних умовах на одній і тій же вибірці випробуваних.

Було обрано традиційний рівень значущості розбіжностей 0,05%, що означає, що ми готові прийняти 5% відхилення нульової гіпотези, коли вона насправді вірна. Тобто, застосувавши Т-критерій Вілкоксона з рівнем значущості 0,05, ми порівняли дві вибірки (тестування за методикою КАМТ та паперовим тестуванням) і перевірили, чи є ймовірність менш ніж 5%, що спостережувані відмінності виникли випадково. Якщо p -значення тесту менше 0,05, то відхиляється нульова гіпотеза, що означає, що між вибірками є значуща відмінність.

Одержані результати було проаналізовано за допомогою статистичного пакету для соціальних досліджень статистичний пакет IBM SPSS Statistics 23.0 for Windows та Microsoft Excel 2019 для розрахунку, який дає змогу виконувати будь-які операції із зазначеними у дослідженні даними, обирати для оцінювання відповідний статистичний критерій.

У ході аналізу результатів мовного тестування за двома методиками з читання ми виявили, що значення $p = 0,705$, тобто більше, ніж 0,05, що підтвердило нульову гіпотезу про відсутні статистичні зміни (рис. 3).

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between samples Reading_na_papere and Reading_KAMT equals 0.	Related-Samples Wilcoxon Signed Rank Test	,705	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Рис. 3 Статистичні дані аналізу результатів мовного тестування з читання за методикою КАМТ та паперового тестування

Аналіз результатів мовного тестування за двома методиками з аудіювання ми продемонстрував, що значення $p = 0,739$, що більше, ніж 0,05, що також підтверджує нульову гіпотезу про відсутні статистичні зміни (рис. 4).

	Null Hypothesis	Test	Sig.	Decision
1	The median of differences between samples Listening_na_papere and Listening_KAMT equals 0.	Related-Samples Wilcoxon Signed Rank Test	,739	Retain the null hypothesis.

Asymptotic significances are displayed. The significance level is ,05.

Рис. 4. Статистичні дані аналізу результатів мовного тестування з читання за методикою КАМТ та паперового тестування

Отже, відсутність статистично значущих змін під час вимірювання результатів тестування за двома методиками – КАМТ

та паперового тестування свідчить про об'єктивність та валідність методики КАМТ, а отже, про можливість його використання під час проведення мовного тестування.

Додатково для підтвердження гіпотези дослідження було проведено кореляційний аналіз щодо наявності високих прямих кореляційних зв'язків між тестуванням за двома методиками – КАМТ та паперового тестування. Аналіз даних показав сильний позитивний кореляційний зв'язок ($\rho = 0.807$) між результатами читання на папері та результатами читання за допомогою КАМТ. Цей зв'язок є статистично значущим на рівні $p < 0.01$ ($p = 0.000$), що підтверджує високу ймовірність того, що такий зв'язок не є випадковим, а отже, свідчить про об'єктивність і валідність КАМТ (табл. 1).

Висновки. Обґрунтовано методику КАМТ відповідно до вимог стандарту НАТО STANAG 6001, в основі якої є алгоритм КАМТ, що складається з блоків: 1) вихідної точки (the Starting Point block), що відповідає за створення набору тесових завдань окремо для CMP 1, CMP 2, CMP 3 відповідно до вимог НАТО STANAG 6001; 2) блоку представлення тестових завдань (the Starting Point block), який безпосередньо відповідальний за адаптацію тестових завдань залежно від відповіді тих, хто тестується; 3) блоку оцінювання та завершення тестування (the Scoring and Termination Criterion), функція якого полягає у визначенні остаточного рівня ПК того, хто тестується, та завершенні мовного тестування. КАМТ проводиться у межах стандартизованих мовленнєвих рівнів, які відповідають рівням мовленнєвої компетентності відповідно до вимог стандарту НАТО STANAG 6001, а саме: CMP 0 (Відсутність практичного вміння); CMP 0+ (Завчений мінімум); CMP 1 (Виживання); CMP 1+ (Виживання +); CMP 2 (Функціональний), CMP 3 (Професійний).

Для експериментальної перевірки результативності методу КАМТ було проведено педагогічний експеримент. Педагогічний експеримент складався з підготовчого, основного та впроваджувального етапів. Результатом підготовчого етапу педагогічного експерименту стало аналіз теоретичних джерел для подальшого обґрунтування результативності методики КАМТ.

Для визначення результативності методики КАМТ на основному етапі педагогічного експерименту було проведено тестування

Таблиця 1

Кореляційні зв'язки між тестуванням за КАМТ та паперового тестування

	Паперове тестування	КАМТ	Паперове тестування	КАМТ
Correlation Coefficient	1,000	,807**	,327	,490**
Sig. (2-tailed)		,000	,078	,006
N	30	30	30	30
	,807**	1,000	,115	,476**
Sig. (2-tailed)	,000		,546	,008
N	30	30	30	30
Correlation Coefficient	,327	,115	1,000	,730**
Sig. (2-tailed)	,078	,546		,000
N	30	30	30	30
Correlation Coefficient	,490**	,476**	,730**	1,000
Sig. (2-tailed)	,006	,008	,000	
N	30	30	30	30

тридцяти слухачів за двома методиками, КАМТ та паперового тестування, та здійснено порівняння їхніх результатів. На впровадженню етапі експерименту було здійснено статистичний аналіз результатів оцінювання за допомогою критерію визначення статистичної значущості зсувів оцінок – *T*-критерію Вілкоксона. Було з'ясовано, що статистично значущої різниці між отриманими результатами тестувань за двома методиками – КАМТ та паперового тестування – немає, а отже, гіпотеза дослідження підтвердилась.

Отже, відсутність статистично значущих змін під час вимірювання результатів тестування за двома методиками (КАМТ та паперове тестування) підтверджує об'єктивність

та валідність методики КАМТ, що робить її придатною для використання у мовному тестуванні відповідно до стандарту НАТО СТАНАГ 6001. Крім того, кореляційний аналіз показав сильний позитивний зв'язок ($\rho = 0.807$) між результатами читання та результатами аудіювання за двома методиками. Зазначений кореляційний зв'язок є статистично значущим ($p < 0.01$), що додатково підтверджує об'єктивність і валідність КАМТ.

Перспективи подальших досліджень у цьому напрямі вбачаємо в стандартизації методики КАМТ, що сприятиме підвищенні точності, об'єктивності, валідності та надійності результатів тестування відповідно до стандарту НАТО СТАНАГ 6001.

ЛІТЕРАТУРА

1. Постоленко І. С. Використання комп'ютерного тестування для контролю і оцінювання знань з англійської мови. *Актуальні питання у сучасній науці*. 2024. № 6 (24). С. 968–978. DOI: 10.52058/2786-6300-2024-6(24)-968-979.
2. Радкевич О. П. Адаптивне тестування в контексті використання електронних засобів навчання: суть, розроблення та оцінювання. *Professional Pedagogics*. 2023. № 1. 58–73. DOI: 10.32835/2707-3092.2023.26.58-73.
3. Сікора Я. Б. Створення адаптивного тесту: інструменти та технології. *Наукова конференція викладачів та молодих науковців Житомирського державного університету імені Івана Франка з нагоди Днів науки: тези доп.*, 16–17 трав. / Житомир, 2024. С. 424–427.
4. Gawliczek P., Krykun V., Tarasenko N., Tyshchenko M., Shapran O. Computer Adaptive Language Testing According to NATO STANAG 6001 Requirements. *Advanced Education*. 2021 № 8 (17). 19–26. DOI: 10.20535/2410-8286.225018.
5. Hryshchuk Yu., V. Krykun. Current trends in adaptive training at higher military educational institutions. *Проблеми впровадження дистанційного навчання в освітньому процесі ВВНЗ та можливі шляхи їх вирішення* : матеріали V міжнародної наук.-практ. конф., м. Київ: НУОУ, 21 листоп. 2022. С. 43–46. DOI: 10.36059/978-966-397-261-9/16.4.
6. Izadi M., Izadi M., Heidari F. The potential of an adaptive computerized dynamic assessment tutor in diagnosing and assessing learners' listening comprehension. *Educ. Inf. Technol.* 2024 № 29. 3637–3661. <https://doi.org/10.1007/s10639-023-11871>
7. Sherkuziyeva N., Imamutdinovna Gabidullina F., Ahmed Abdel-Al Ibrahim K. et al. The comparative effect of computerized dynamic assessment and rater mediated assessment on EFL learners' oral proficiency, writing performance, and test anxiety. *Lang Test Asia*. 2023 № 13, 15. <https://doi.org/10.1186/s40468-023-00227-3>
8. Weiss J. D., Sahin A. Computer Adaptive Testing: From Concept to Implementation (2024). *Methodology in the Social Sciences Series*. Guilford Publications. P. 360.

REFERENCES

1. Postolenko, I. (2024). Vykorystannia kompiuternoho testuvannia dlia kontroliu i otsiniuvannia znan z anhliiskoi movy [Using computer-based testing to monitor and assess English language skills]. *Aktualni pytannia u suchasni nauksi*, 6 (24). pp. 968–978 [in Ukrainian].
2. Radkevych, O. (2023). Adaptivne testuvannia v konteksti vykorystannia elektronnykh zasobiv navchannia: sut, rozroblennia ta otsiniuvannia [Adaptive testing in the context of using e-learning tools: essence, development and evaluation]. *Professional Pedagogics*, 1. pp. 58–73 [in Ukrainian].
3. Sikora, Ya. (2024). Stvorennia adaptivnoho testu: instrumenty ta tekhnolohii. Naukova konferentsiia vykladachiv ta molodykh naukovtsiv Zhytomyrskoho derzhavnoho universytetu imeni Ivana Franka z nahody Dniv nauky [Creating an adaptive test: tools and technologies. Scientific Conference of Teachers and Young Scientists of Ivan Franko Zhytomyr State University on the occasion of the Days of Science]. Zhytomyr, 424–427 [in Ukrainian].
4. Gawliczek, P., Krykun, V., Tarasenko, N., Tyshchenko, M., & Shapran, O. (2021). Computer Adaptive Language Testing According To NATO STANAG 6001 Requirements. *Advanced Education*. 8 (17). 19–26.
5. Hryshchuk Yu., V. Krykun (2022). Current trends in adaptive training at higher military educational institutions. Kyiv: NDUU, 43–46.
6. Izadi, M., Izadi, M., & Heidari, F. (2024). The potential of an adaptive computerized dynamic assessment tutor in diagnosing and assessing learners' listening comprehension. *Educ. Inf. Technol.* 29. 3637–3661
7. Sherkuziyeva, N., Imamutdinovna Gabidullina, F., Ahmed Abdel-Al Ibrahim, K. et al. (2023). The comparative effect of computerized dynamic assessment and rater mediated assessment on EFL learners' oral proficiency, writing performance, and test anxiety. *Lang Test Asia*. 13, 15.
8. Weiss J., D, Sahin, A. (2024). Computer Adaptive Testing: From Concept to Implementation. *Methodology in the Social Sciences Series*. Guilford Publications. P. 360.