

# Рекомендовані вимоги до валідизації та застосування в роботі шкал та методик в області діагностики та надання медичної та психологічної допомоги

Олег Чабан

Навчально-науковий інститут психічного здоров'я  
Національного медичного університету ім.  
О.О.Богомольця

Андрій Бурдейний

Навчально-науковий інститут психічного здоров'я  
Національного медичного університету ім.  
О.О.Богомольця

**Актуальність.** Сучасні медичні й психологічні виклики, такі як посттравматичні розлади, депресія, тривожні стани та пост-COVID синдром, вимагають точного та систематизованого підходу до діагностики та оцінки стану пацієнтів. Використання валідованих шкал дозволяє фахівцям ухвалювати обґрунтовані рішення на основі даних, що сприяє ефективному наданню допомоги.

Окрім цього, науковий прогрес і постійні зміни в стандартах лікування вимагають оновлення та вдосконалення методик для їх застосування в різних клінічних і психологічних контекстах. Рекомендації допоможуть уніфікувати підходи до оцінки інструментів, що дозволить забезпечити високу якість діагностики та лікування, незалежно від місця або специфіки проблем пацієнта.

**Мета.** Забезпечити науково обґрунтовані стандарти для валідизації та ефективного використання діагностичних шкал і методик у медичній та психологічній допомозі, що сприятиме підвищенню якості діагностики та оптимізації лікувальних підходів для покращення стану пацієнтів.

## Вступ

Ідея валідизації в наукових дослідженнях і розробці діагностичних інструментів у психології виникла як відповідь на потребу в забезпеченні точності, надійності та адекватності використовуваних методик.[1] Це важливо тому, що психологічні конструкти, такі як тривожність, депресія або стрес, є абстрактними поняттями, які неможливо виміряти безпосередньо. Тому розробка і валідиція інструментів для вимірювання цих конструктів спрямована на те, щоб забезпечити їхню точну операціоналізацію – перетворення на чітко визначені, вимірювані параметри.[2-4] У цьому контексті валідність виступає гарантією того, що інструмент дійсно виконує свою функцію належним чином.

Процес валідизації включає низку етапів, таких як перевірка конструктивної, змістової та критеріальної валідності, які допомагають визначити, чи відповідає інструмент своєму завданню. [5-6] Конструктивна валідність означає, що методика корелює з іншими інструментами, які вимірюють аналогічні конструкти, в той час як змістова валідність свідчить про відповідність питань або завдань інструменту загальному змісту конструкту. Критеріальна валідність перевіряє, наскільки результати інструменту узгоджуються з іншими, зовнішніми

критеріями, які можуть вказувати на достовірність результатів.

Валідизація є особливо важливою у медичній та психологічній допомозі, оскільки помилкові або некоректні діагностичні висновки можуть призвести до неефективних або навіть шкідливих втручань.[7-9] Для пацієнтів з посттравматичними розладами, депресією чи тривожними станами, правильний діагноз є критично важливим для визначення подальшого лікування і реабілітації.[10-12] Використання інструментів, які пройшли ретельний процес валідизації, дозволяє лікарям і психологам ухвалювати більш обґрунтовані рішення та забезпечувати точність діагностики.

Загалом, важливість валідизації інструментів полягає не лише в їхній науковій точності, але й у практичній корисності. В медичній та психологічній сфері валідовані інструменти дозволяють стандартизувати підходи до діагностики та лікування пацієнтів, зменшуючи варіативність результатів залежно від людського фактору чи конкретних умов.[13-16] Це особливо важливо в умовах, коли медична допомога надається дистанційно або в екстрених ситуаціях, де точність і швидкість ухвалення рішень є вирішальними.[17-19]

Таким чином, розробка та валідизація шкал і методик у галузі діагностики медичних і психологічних розладів є необхідною умовою для забезпечення якісної допомоги пацієнтам. Забезпечення відповідності інструментів сучасним науковим стандартам підвищує ефективність лікування та діагностики, а також дозволяє розробляти нові, адаптовані до сучасних викликів підходи до надання допомоги.

Історія цього питання знає певні концептуальні етапи розвитку ідеї валідизації:

1. Витоки в класичній теорії тестів - як пізнавального інструменту:
  - В класичній тестовій теорії, що була розроблена на початку ХХ століття, важливою була проблема точності вимірювань. Валідність розглядалася як ключовий критерій якості тестів, що вимагало відповідності між теоретичним конструктом та емпіричними даними.
2. Зростаюча увага до психометрії, як уніфікованого механізму досліджень:
  - У середині ХХ століття розвиток психометрії та застосування статистичних методів для аналізу тестів привели до більш систематичного підходу до валідизації. Це включало концептуалізацію різних видів валідності: конструктної валідності, критеріальної валідності та вмістової валідності.
3. Стандартизація, варіфікація та професійні вимоги:
  - Поява стандартів, таких як "Стандарти для освітніх та психологічних тестів" (Standards for Educational and Psychological Testing), розроблені Американською психологічною асоціацією (APA), підкреслили необхідність забезпечення валідності як обов'язкової частини розробки будь-якого тесту.
4. Актуальність у сучасних багатопрофільних дослідженнях:
  - В сучасних умовах зростання різноманітності та складності психометричних інструментів, включаючи цифрові і мобільні платформи, питання валідизації стає ще більш важливим. Це зумовлено потребою забезпечення надійності результатів в нових контекстах і для різних груп населення.

Таким чином, ідея валідизації виросла з потреби гарантувати, що інструменти вимірювання в психології та психіатрії забезпечують точні, надійні та адекватні результати, що є ключовими для наукових досліджень та клінічної практики.

Дані рекомендації допоможуть дотримуватися принципу доказовості, ефективності та надійності в формуванні протоколів та клінічних настанов, галузевих регламентуючих документів, методик та протоколів медико - психологічних служб та інших робочих документів в області психіатрії, медичної психології та психотерапії, а також клінічної психології.



Рекомендації побудовані за простим принципом який прийнятий у світовій науці і є загальноприйнятим для подібних запитів і ґрунтується на принципах від найбільш доказових до найменш доказових, або емпіричних спостережень та описаннях одиночних випадків. З вибором звичайно найбільш доказових та таких, які відповідають нижчеописаним критеріям.

## **Концепція створення тесту**

Оцінка та розуміння поведінки людини зазвичай потребують детального спостереження або взаємодії з нею протягом певного часу. Крім того, ми можемо отримати інформацію про людину від тих, хто регулярно з нею спілкується, або ж через наявні записи, що її стосуються. Важливо також враховувати контекст, у якому відбувається взаємодія, оскільки це може впливати на поведінку та сприйняття людини.

Швидший спосіб зібрати дані — це використовувати самозвіт, де людина відповідає на запитання на основі свого минулого досвіду або уявлених ситуацій. Цей метод дозволяє швидко зібрати інформацію про думки та почуття особи. Зібрані таким чином відповіді можуть створити чітке уявлення про її поведінку, що в подальшому допоможе в діагностиці чи розробці індивідуальних стратегій допомоги.

Цей прямий метод самозвіту має кілька переваг: він є швидким, дозволяє одночасно отримувати інформацію від кількох людей, забезпечує можливість порівняння та базується на безпосередніх відповідях осіб, а не на висновках спостерігачів. Проте, у нього є й певні недоліки. Наприклад, існує ризик отримання ненадійної інформації про почуття та поведінку, якщо людина має спотворене самосприйняття або неправильно розуміє питання; можливість фальсифікації відповідей з метою справити певне враження; забування; поверхневі відповіді через брак участі; а також, що найважливіше, проблеми, пов'язані з формулюванням тестових завдань, такими як їхня релевантність, варіанти відповідей та їхня формулювання.[20] Одне й те саме питання, сформульоване по-різному та адресоване різним особам, може призводити до різних відповідей. Наприклад, розглянемо три запитання, які мають на меті визначити, чи страждає людина від депресії: «Чи у Вас відсутній настрій?», «Враховуючи скарги Вашої сім'ї на Ваш постійний поганий настрій, чи можете Ви визнати, що у вас депресія?» та «Як би Ви оцінили свій рівень депресії за шкалою від 1 до 10, де 10 означає дуже високий рівень?». Перше запитання може призвести до простої відповіді «Так» або «Ні», але не відображає ступінь тяжкості депресії. Друге запитання може спровокувати у респондента захисну реакцію, в результаті якої він може помилково відповісти «Ні». Третє запитання, за умови, що респондент правильно його зрозуміє, може дати уявлення не лише про наявність депресії, але й про її ступінь.

Добре розроблений тест, який науково досліджує всі аспекти концепції чи ситуації та демонструє свою стабільність, може вирішити багато з цих проблем, стаючи швидким і точним інструментом.[21-23] Психологічні тести все частіше використовуються в дослідженнях, клінічній практиці та освітніх установах завдяки своїй надійності, а також легкості та швидкості збору, порівняння й аналізу даних. Нова шкала зазвичай розробляється, коли наявні інструменти чи тести для вимірювання цікавого конструкту відсутні або коли існуючі тести не повністю відповідають вимогам або не доступні мовою, що потрібна.

Психологічний тест — це систематична процедура для отримання зразків поведінки, що мають відношення до когнітивного чи афективного функціонування, а також для підрахунку й оцінки цих зразків відповідно до стандартів. [24] Психологічні тести відіграють важливу роль у різних аспектах життя, включаючи клінічні та кар'єрні ситуації, а також у саморозумінні особи. Вони слугують корисними інструментами для прийняття рішень, дозволяючи краще зрозуміти як себе, так і інших. Окрім цього, тести можуть використовуватися як індикатори особистісного зростання, що робить їх важливими для саморозвитку.

Крім практичного застосування в житті, психологічні тести також є важливими у наукових

дослідженнях, допомагаючи зрозуміти психологічні феномени та індивідуальні відмінності. Вони можуть охоплювати різні категорії, такі як тести здібностей, тести особистості та нейропсихологічні тести. Кожен з цих інструментів має свої специфічні цілі та методи оцінювання, що сприяє глибшому аналізу особистісних характеристик.

Однак важливо зазначити, що термін «тест» слід використовувати лише для інструментів, у яких відповіді оцінюються на основі їх якості або правильності. Інші інструменти, такі як анкети, опитування та контрольні списки, які не передбачають оцінювання відповідей, належать до категорії особистісних тестів. Це розмежування підкреслює важливість точності у визначенні та використанні різних типів психологічних інструментів для досягнення ефективності в дослідженнях та практичній діяльності.

## **Побудова тесту**

### **Створення чернетки**

Початковий етап створення чорнових пунктів вимагає глибокого розуміння концепції, що досягається через читання спеціалізованої літератури, консультації з експертами, розвиток власного критичного мислення та активне спостереження. Важливо також з'ясувати, чи є концепція єдиною або складається з кількох підаспектів.[25] Це дозволяє сформулювати чітку та структуровану основу для подальшої розробки і реалізації ідей, забезпечуючи їхню адекватність і відповідність досліджуваній темі. Розглянемо приклад тривоги, яка охоплює не лише емоційні аспекти, такі як страх, розчарування та тривога, але й когнітивні, включаючи постійні думки про можливі невдачі, проблеми та негативні наслідки. Фізіологічні прояви, що супроводжують тривогу, можуть включати тремор, прискорене серцебиття, відчуття непритомності, холодні липкі руки, сухість у роті та розлади шлунка. Для адекватного вимірювання цих складових важливо, щоб конструктор тесту підготував детальний набір елементів, бажано за участю команди експертів. Будь-які недоліки на етапі формулювання тестових завдань можуть призвести до помилок, які відображаються як дисперсія помилок — випадкові та нерелевантні джерела помилок, на відміну від справжньої дисперсії, яка свідчить про реальні відмінності.[26]

### **Формат тесту**

Необхідно також визначити формат тесту: чи буде він відкритим (відомим також як тест з вільною або сконструйованою відповіддю) чи закритим (об'єктивним або з примусовим вибором). Відкриті відповіді можуть бути представлені у формі письмових зразків (як у тестах здібностей), вільних усних відповідей (як у проєктивних тестах) або виконання завдань. Цей формат є корисним для дослідження понять, ситуацій і явищ, оскільки він надає ширший спектр можливостей, дозволяє отримати багатші зразки поведінки індивіда та виявити його унікальні риси. Проте підрахунок балів у таких тестах є більш складним і трудомістким, а їх надійність і валідність зазвичай нижчі в порівнянні з тестами закритого формату (з примусовим вибором).[24]

Формат примусового вибору передбачає обмежену кількість альтернатив, з яких респондент має обрати одну. Вибір може бути здійснений через множинний вибір, істинні/неправдиві твердження, рейтинги, оцінювання та відповідності, а також шляхом перегрупування наданих варіантів. Однією з найбільш відомих шкал у цьому контексті є шкала Лайкерта, де респонденти можуть висловити свою згоду, прийняття або частоту повторення, обираючи з непарної кількості (3, 5, 7 або навіть 9) варіантів відповіді. Зазвичай середня точка на цій шкалі означає нейтральну або середню позицію.[27] Сума відповідей на всі пункти використовується для оцінки результатів. Найбільш популярною є 5-бальна шкала Лайкерта, яка містить такі варіанти відповідей: «повністю згоден», «згоден», «ні згоден, ні не згоден», «не згоден» та «зовсім не згоден».[28]

Формат примусового вибору є найпоширенішим варіантом для тестових завдань завдяки простоті та швидкості оцінювання. Об'єктивність відповідей робить його більш надійним, і його можна використовувати як в індивідуальному, так і в груповому тестуванні. Проте можливі помилки, якщо респондент просто вгадує у тестах здібностей або відповідає нещиро чи безвідповідально у особистісних тестах, що може негативно вплинути на надійність і валідність тесту. Крім того, створення тесту з примусовим вибором потребує глибокого розуміння концепції та навичок у розробці спеціалізованих тестів і формулюванні завдань.

## **Якісний аналіз предметів**

Організація, тип і формулювання елементів вимірювання, а також вступ, інструкції, переходи та завершення тесту повинні, в ідеалі, заохочувати респондентів надавати точну та адекватну інформацію. Вони також мають запобігати можливості респондентів уникати чи відмовлятися відповідати на певні питання. Формулювання тестових завдань повинно враховувати специфіку цільової групи. Нижче наведені рекомендації, які допоможуть досягти цих цілей.

- Використовуйте просту мову, короткі речення та терміни, які учасник розуміє (наприклад, у пункті «Я боюся панічних атак», респондент може не знати, що таке «панічні атаки»).
- Переконайтеся, що всі можливі варіанти відповідей надано для запитань із вимушеним вибором (закритою відповіддю) (наприклад, у запитах про «Частоту нападів паніки», якщо варіанти відповідей коливаються від «один раз на рік» до «кілька разів на день», той, хто ніколи не відчував нападів паніки, змушений вибрати варіант, який його не стосується).
- Переконайтеся, що формулювання питання відповідає варіантам відповіді (наприклад, якщо питання звучить так: «Чи відчуваєте ви страх без причини?», варіанти відповіді «повністю згоден», «згоден», «не згоден», «зовсім не згоден» не будуть бути придатними, оскільки вони призначені для твердження, а не запитання).
- Переконайтеся, що всі предмети актуальні та потрібні.
- Еникайте двостволок (наприклад, «У вас проблеми з травленням і задишка?» – якщо пацієнт має лише одне, а не інше, як він відповість?).
- Еникайте навідних, упереджених формулювань (наприклад, «Тривога часто супроводжується дратівливістю; оцініть свій рівень дратівливості за шкалою від 1 до 10»).
- Уникайте двозначних слів із широким або абстрактним значенням. Наприклад, такі слова, як «загалом», «у середньому» та «регулярно» можуть означати різні речі для різних людей.
- Уникайте предметів, які потребують зусиль і віддаленого пригадування (наприклад, «Чи відчували ви напади тривоги 5 років тому?»)
- Уникайте запитувати інформацію, що загрожує самому, викликає збентеження або приватну інформацію, якщо вона не є важливою (наприклад, дохід, моральна позиція, особисті звички).

Проведення пілотного тестування з участю невеликої групи респондентів (15–30 осіб), які є репрезентативними для цільової аудиторії, може допомогти перевірити зрозумілість та відповідність завдань тесту. На цьому етапі важливо провести якісний аналіз елементів, що включає оцінку охоплення змісту, формулювання та структуру речень, щоб налаштувати їх відповідно до вимог. Після цього результати передаються рецензентам для отримання коментарів і подальшого перегляду, якщо це необхідно.

Після внесення коригувань на цих трьох рівнях, розробник тесту може підготувати його до наступного етапу — кількісного аналізу, який проводиться на значно більшій вибірці, що складається з 200 і більше респондентів, використовуючи дані з пілотного дослідження та відгуки рецензентів.

## Кількісний аналіз предметів

Кількісний аналіз елементів тесту включає низку статистичних процедур, які застосовуються для фінального відбору завдань на основі відповідей, отриманих від респондентів під час розробки тесту. Два ключові аспекти, які враховуються при виборі елементів з пулу, — це їхній рівень складності та дискримінаційна здатність. Рівень складності – це показник частки або відсотка респондентів, які правильно відповіли на питання; з цієї причини його часто називають *P*-value. [24,29]

Індекс складності може варіюватися від 0,0 до 1,0, де 1,0 вказує на те, що 100% респондентів правильно відповіли на запитання, що свідчить про легкість завдання. У тестах, що базуються на критеріях, де результати респондентів порівнюються безпосередньо зі стандартом, складність завдань зазвичай знаходиться в межах 0,7–0,8, оскільки вони повинні бути зрозумілими для більшості учасників. У протилежність цьому, тести, орієнтовані на норми, що мають на меті ранжування та порівняння респондентів, розробляються з урахуванням більшої складності, і їхній індекс складності коливається від 0,4 до 0,6.[24]

Індекс дискримінації використовується для оцінки здатності тестового елемента (запитання або завдання) відрізнити учасників з різними рівнями знань або навичок. Він показує, наскільки ефективно дане запитання визначає, хто відповідає правильно (високий рівень навичок) і хто відповідає неправильно (низький рівень навичок). Можливий діапазон індексу дискримінації становить від -1,0 до 1,0. Індекс негативної дискримінації може вказувати на те, що елемент вимірює щось інше, ніж те, що вимірює решта тесту. Частіше це ознака того, що оцінка предмета була введена неправильно. Це також може статися через необережність або коли предмети написані в зворотному напрямку.

## Достовірність та валідність тесту

Подібно до того, як ми очікуємо, що ваги точно відображають вимірювану масу (валідність) і показують стабільні результати при кожному використанні (надійність), психологічні тести також повинні відповідати цим критеріям, навіть якщо вони вимірюють абстрактні поняття. Надійність вказує на те, наскільки послідовними та стабільними є результати тесту протягом певного часу, тоді як валідність свідчить про те, що тест дійсно вимірює те, що він має виміряти.

Цікаво, що тест може демонструвати високу надійність, але не бути валідним, тобто він може стабільно давати однакові результати, не вимірюючи при цьому бажане явище. Однак, щоб тест вважався валідним, він повинен бути надійним: якщо результати тесту коливаються, важко стверджувати, що він вимірює що-небудь з точністю. Таким чином, надійність і валідність є взаємопов'язаними характеристиками, які важливі для забезпечення достовірності психологічного тестування.

Валідизація — це процес оцінки того, наскільки точно і надійно інструмент, методика або тест вимірює те, що він/вона призначені вимірювати.

Це критично важливий етап у розробці психодіагностичних методик, оскільки валідність визначає, чи можна довіряти результатам тесту і наскільки вони відповідають реальним характеристикам або станам, які досліджуються.

Існує кілька видів валідності:

1. Контентна валідність — оцінює, наскільки тест охоплює всі аспекти досліджуваної характеристики. Наприклад, чи включає тест усі релевантні питання для вимірювання конкретного психічного стану.

2. Конструктна валідність — визначає, наскільки тест дійсно вимірює ту теоретичну конструкцію або концепт, який він має на меті виміряти. Це один з найскладніших і найважливіших типів валідності.
3. Критеріальна валідність — визначає, наскільки результати тесту корелюють з іншими зовнішніми критеріями або стандартами, які вважаються показниками досліджуваного явища.
4. Екологічна валідність — стосується того, наскільки результати тесту можуть бути узагальнені на реальні ситуації або умови.

Валідизація включає збір даних, аналіз результатів та оцінку, щоб переконатися, що інструмент є надійним і точним у вимірюванні того, для чого він був створений. **Таблиця 1** нижче може слугувати основою для розробки, або оцінки психодіагностичних методик, які будуть використовуватись для оцінки психічного стану у військовослужбовців. [30-31]

Таблиця 1. Критерії та рекомендовані вимоги для розробки та оцінки психодіагностичних методик.

Критерій	Опис	Рекомендовані вимоги
Надійність	Ступінь стабільності та консистентності результатів вимірювання.	Коефіцієнт надійності (наприклад, альфа Кронбаха) повинен бути не нижчим за 0,70 для внутрішньої консистентності.
Валідність	Ступінь, до якої методика вимірює те, що вона призначена вимірювати.	Необхідна доказова база контентної, конструктної та критеріальної валідності. Специфічні індикатори повинні відображати військовий контекст.
Чутливість	Здатність методики виявляти різницю між групами або зміни стану обстежуваних.	Методика повинна бути здатна відрізнити різні рівні тяжкості симптомів, зокрема ПТСР або інші стресові розлади.
Специфічність	Здатність методики чітко виявляти цільову характеристику без перехреснення з іншими змінними.	Методика повинна мати високу специфічність для виявлення саме військових стресових розладів, мінімізуючи помилкові позитивні або негативні результати.
Екологічна валідність	Наскільки результати методики можуть бути узагальнені на реальні умови військової служби.	Психодіагностична методика повинна враховувати специфічні умови та контекст служби, включаючи бойові дії, стресові ситуації та військову культуру.
Етика	Дотримання етичних стандартів у проведенні діагностики.	Необхідно забезпечити конфіденційність, добровільність участі та інформовану згоду військовослужбовців.
Крос-культурна валідність	Відповідність методики культурним та мовним особливостям військових різних національностей.	Методика повинна бути адаптована до культурних і мовних особливостей військовослужбовців, щоб забезпечити точність та зрозумілість питань та результатів.
Практичність	Легкість та швидкість проведення методики у військових умовах.	Методика повинна бути зручною у використанні, швидкою в адмініструванні та відповідати ресурсним можливостям польових умов або військових медичних установ.
Часовий аспект валідизації	Актуальність та часова стабільність валідизації методики.	Валідизація повинна бути нещодавною та враховувати зміни у військовій структурі та умовах служби, що відбулися за останні роки.

Ця ж таблиця може бути доповненням до **Таблиці 2**, яка буде основою для вибору тієї чи іншої методики для робочих програм, протоколів та інших регламентуючих документів роботи військових психологів, психіатрів та інших спеціалістів в області ментального здоров'я.

Формування протоколів з надання психологічної допомоги військовослужбовцям повинно ґрунтуватися на доказових підходах та принципах валідизації. Це означає, що всі методи, інтервенції та практики, які включені в протоколи, повинні бути підкріплені науковими доказами і пройшли відповідну перевірку ефективності та безпечності.

## **Ключові етапи та принципи, яких слід дотримуватися при розробці таких протоколів:**

1. Доказова база (Evidence-Based Practice)
2. Валідизація інтервенцій (Validation of Interventions)
3. Принципи безпеки та етики (Principles of Safety and Ethics)
4. Моніторинг і оцінка ефективності (Monitoring and Evaluation of Effectiveness)
5. Практичність та адаптивність (Practicality and Adaptability)
  - *Систематичний огляд літератури:* Протоколи повинні базуватися на надійних і актуальних даних. Це передбачає проведення систематичного огляду літератури для визначення найкращих практик та методів, які мають найбільшу доказову підтримку.
  - *Клінічні випробування:* Включення в протоколи лише тих методів, які пройшли клінічні випробування і були доведені як ефективні у вибраній популяції (військовослужбовці).
  - *Методологічно якісні дослідження:* Пріоритет надається дослідженням з високим рівнем методологічної якості, таким як рандомізовані контрольовані випробування (RCT), метааналізи та систематичні огляди.
  - *Оцінка ефективності:* Усі інтервенції повинні пройти процес валідизації, тобто доведення їхньої ефективності саме у контексті військових умов. Це включає аналіз даних щодо їх впливу на психічне здоров'я військовослужбовців, зокрема в умовах стресу, травми або бойових дій.
  - *Культурна адаптація:* Інтервенції повинні бути валідизовані з урахуванням культурних, соціальних та психологічних особливостей військових різних національностей і середовищ.
  - *Адаптація до військового контексту:* Протоколи повинні бути розроблені з урахуванням специфіки військових умов, включаючи бойові операції, стресові фактори, пов'язані зі службою, та інші особливості.
  - *Етичні стандарти:* Всі процедури та інтервенції повинні відповідати міжнародним етичним стандартам, зокрема щодо інформованої згоди, конфіденційності та добровільності участі.
  - *Безпечність інтервенцій:* Протоколи повинні включати лише ті методи, які пройшли перевірку на безпечність для військовослужбовців, особливо в умовах бойових дій або після травматичних подій.
  - *Регулярний перегляд протоколів:* Протоколи повинні регулярно оновлюватися на основі нових досліджень і даних, що з'являються. Це гарантує, що методики залишаються актуальними і відповідають сучасним стандартам.
  - *Зворотний зв'язок:* Збір та аналіз даних щодо ефективності та безпечності застосовуваних методик і протоколів в реальних умовах, зокрема за допомогою опитувань військовослужбовців, моніторингу результатів та фідбеку від спеціалістів.
  - *Зручність у використанні:* Протоколи повинні бути розроблені таким чином, щоб їх



можна було легко застосовувати в різних військових умовах, включаючи польові умови, госпіталі, реабілітаційні центри тощо.

- *Гнучкість*: Протоколи повинні передбачати можливість адаптації до конкретних ситуацій або потреб військовослужбовців.

6. Протоколи та методика повинні бути доступними, як при безпосередній роботі з військовослужбовцем так і для проведення супервізій роботи фахівців («розмовляти однією мовою»).

Ці принципи допоможуть забезпечити високу якість, ефективність і безпечність надання психологічної допомоги військовослужбовцям, сприяючи їхньому психічному здоров'ю та здатності адаптуватися до складних умов служби.

Останні напрямки розробки як самих тестів так і досягнення їх більшої ефективності направлені на можливість працювати з ними дистанційно (пілот виконаний на PHQ-9) з метою забезпечення їхньої точності та надійності порівняно зі "золотим стандартом" діагностики, таким як діагностичні інтерв'ю на основі DSM-5 та в останні роки за даними публікацій і нашого спілкування з колегами за кордоном іде робота в напрямку максимально переводу паперових носіїв в варіант дистанційного опрацювання.

Другий напрямок - долучення штучного інтелекту до вибору та опрацювання даних тестування з певними етичними та конфіденційними обмеженнями особливого того, що торкається безпеки. Для практичного вибору доказових методик, шкал та методів в області надання медико-психологічної допомоги рекомендуємо заповнити нижче наведену таблицю з послідоючою фахово - експертною оцінкою різними спеціалістами цього напрямку, або як найбільш достовірний варіант оцінки - засліпленим методом.

Таблиця 2. Рекомендовані критеріїв для визначення валідності психодіагностичних методик

Рекомендовані критеріїв для визначення валідності психодіагностичних методик, які можуть використовуватися для проведення психологічної діагностики та оцінки якості психологічної допомоги		
№	Критерій	Визначення
1	Назва психодіагностичної методики українською та англійською мовою	
2	Короткий опис психодіагностичної методик	
3	Вікова група	
4	Цільовий споживач	
5	Кількість цитувань цієї психодіагностичної методики у наукових публікаціях за останні п'ять років	
6	Чи входить методика до офіційних клінічних рекомендацій національного або міжнародного рівня? Якщо так, зазначте назву документу та опишіть роль методики в цих документах.	
7	Чи згадується ця психодіагностична методика у базах даних Scopus? Якщо так, будь ласка, надайте приклади цитувань, оформлених у стилі APA (за можливості додайте DOI)	
8	Чи згадується ця психодіагностична методика у базах даних Web of Science? Якщо так, будь ласка, надайте приклади цитувань, оформлених у стилі APA (за можливості додайте DOI)	
9	Чи була методика згадана у наукових фахових виданнях України категорії	

	"А"? Якщо так, будь ласка, надайте приклади цитувань, оформлених у стилі APA (за можливості додайте DOI)	
10	Чи була методика згадана у наукових фахових виданнях України категорії "Б"? Якщо так, будь ласка, надайте приклади цитувань, оформлених у стилі APA (за можливості додайте DOI)	
11	Чи була ця методика використана у рандомізованих клінічних дослідженнях, що стосуються психічних розладів? Якщо так, будь ласка, надайте приклади цитувань, оформлених у стилі APA (за можливості додайте DOI)	
12	Чи включалася психодіагностична методика до метааналітичних оглядів? Якщо так, надайте приклади цитувань, оформлених у стилі APA (за можливості додайте DOI)	
13	Чи була психодіагностична методика офіційно перекладена та адаптована для українського контексту? Якщо так, будь ласка, надайте приклади цитувань, оформлених у стилі APA (за можливості додайте DOI)	
14	Чи проводилися дослідження, що перевіряли валідність цієї методики саме в Україні? Якщо так, будь ласка, надайте приклади цитувань, оформлених у стилі APA (за можливості додайте DOI)	
15	Чи використовується ця методика в навчальних програмах університетів або в курсах підвищення кваліфікації для спеціалістів сфери психічного здоров'я? Якщо так, вкажіть назву програми та заклад вищої освіти, де прописана дана методика.	

Звичайно можна в ситуації цейтноту спростити оцінку вибору шкал, чи методик за принципом: максимальна кількість позначок в рядках 6,7,8,11,12 це і буде критерій вибору, а далі виникають інші критерії вибору, а також:

1. Доступність: складність/простота шкали/ методики як для пацієнта так і для фахівця;
2. Наявність фахового перекладу на мову працюючого та опитуваного. Бажано і якщо дотримуватися наукових вимог – потрібна лінгвістична експертиза;
3. Можливість аналізу, публікацій в науковій періодиці та експертної оцінки цієї наукової роботи як в нашій країні так і за кордоном (значить ці методики також повині там використовуватися);
4. Досягнення принципу наступності. Перехід військовослужбовця в статус ветерана і надання йому фахової допомоги за межами армії повині дотримуватися принципу передачі не просто «з теплих рук в теплі руки» (американський психолого-реабілітаційний сленг роботи з ветеранами) а також з дотриманням наукових даних відстеження його ментального здоров'я.

При аналізі та зборі інформації про методики які опубліковані журналах наукометричної бази Scopus рекомендовано дотримуватися пошуку методик згаданих перш за все у журналах першого квартилю (за SJR) та щільністю за імпакт фактором. Нижче приводимо зведену

табличку значень подібних цитувань:

Журнали, що входять до бази даних Scopus, оцінюються за кількома показниками, які дозволяють класифікувати їх за рівнем впливу та наукового авторитету. Основні градації такі:

1. Quartiles (Квартилі):

- Q1 (Перший квартиль): Журнали з найвищим показником цитованості у своїй галузі. Це журнали, які входять у топ 25% за показниками SJR (SCImago Journal Rank). Приклад: Lancet, American Journal of Psychiatry, Journal of the American Academy of Child & Adolescent Psychiatry (JAACAP), Psychological Bulletin, Clinical Psychology Review...
- Q2 (Другий квартиль): Журнали, що займають позиції від 25% до 50% у своїй галузі. Приклад: Journal of Clinical Psychology, International Journal of Psychiatry in Clinical Practice, Psychiatry Research, Journal of Anxiety, Disorders Journal of Affective Disorders...
- Q3 (Третій квартиль): Журнали, що знаходяться у діапазоні від 50% до 75%.
- Q4 (Четвертий квартиль): Журнали з найнижчим показником цитованості, які знаходяться в останніх 25%.

1. Impact Factor (Імпакт-фактор):

- Визначає середню кількість цитувань статей, опублікованих у журналі протягом певного періоду часу (зазвичай двох років). Журнали з високим імпаکت-фактором зазвичай мають високий рівень авторитету.

1. CiteScore:

- Альтернативний показник, який вимірює середню кількість цитувань на одну статтю, опубліковану в журналі протягом чотирьох років.

1. SJR (SCImago Journal Rank):

- Враховує не тільки кількість цитувань, але й значущість та престиж джерел цих цитувань. Журнали з високим SJR мають більший вплив у науковій спільноті.

1. SNIP (Source Normalized Impact per Paper):

- Оцінює вплив цитувань, нормалізуючи їх для різних галузей знань, що дозволяє порівнювати журнали з різних дисциплін.

Ці показники використовуються для оцінки якості та впливу журналів, що дозволяє науковцям обирати найбільш авторитетні платформи для публікації своїх досліджень.

## Посилання

1. Cohen, R. J., Swerdlik, M. E. (2018). Psychological testing and assessment: An introduction to tests and measurement, 9th Edition. Mayfield Publishing Co.
2. Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, 52(4), 281. <https://doi.org/10.1037/h0040957>
3. DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage publications.
4. Clayson, P. E. (2024). The psychometric upgrade psychophysiology needs. *Psychophysiology*, 61(3), e14522. <https://doi.org/10.1111/psyp.14522>
5. Meaning, S. I. I. S. (1995). *Validity of Psychological Assessment*. <https://doi.org/10.1037/0003-066X.50.9.741>
6. Domenech Rodríguez, M. M., Erickson Cornish, J. A., Thomas, J. T., Forrest, L., Anderson,

- A., & Bow, J. N. (2014). Ethics education in professional psychology: A survey of American Psychological Association accredited programs. *Training and Education in Professional Psychology*, 8(4), 241–247. <https://doi.org/10.1037/tep0000043>
7. Groth-Marnat, G. (2009). *Handbook of psychological assessment*. John Wiley & Sons.
  8. Kazdin, A. E. (2004). Evidence-based treatments: Challenges and priorities for practice and research. *Child and Adolescent Psychiatric Clinics*, 13(4), 923-940. <https://doi.org/10.1016/j.chc.2004.04.002>
  9. ASSESSMENT, A. T. F. O. P. (2020). for *Psychological Assessment and Evaluation*.
  10. Cooper, R. (2018). *Diagnosing the diagnostic and statistical manual of mental disorders*. Routledge. <https://doi.org/10.4324/9780429473678>
  11. Keane, T. M., Wolfe, J., & Taylor, K. L. (1987). Post-traumatic stress disorder: Evidence for diagnostic validity and methods of psychological assessment. *Journal of clinical psychology*, 43(1), 32-43. [https://doi.org/10.1002/1097-4679\(198701\)43:13.0.co;2-x](https://doi.org/10.1002/1097-4679(198701)43:13.0.co;2-x)
  12. Castro-Rodríguez, J. I., Olariu, E., Garnier-Lacueva, C., Martín-López, L. M., Pérez-Solà, V., Alonso, J., ... & INSAyD Investigators. (2015). Diagnostic accuracy and adequacy of treatment of depressive and anxiety disorders: A comparison of primary care and specialized care patients. *Journal of affective disorders*, 172, 462-471. <https://doi.org/10.1016/j.jad.2014.10.020>
  13. Furr, R. M. (2021). *Psychometrics: an introduction*. SAGE publications.
  14. Coaley, K. (2014). *An introduction to psychological assessment and psychometrics*.
  15. Kazdin, A. E. (2004). Evidence-based treatments: Challenges and priorities for practice and research. *Child and Adolescent Psychiatric Clinics*, 13(4), 923-940. <https://doi.org/10.1016/j.chc.2004.04.002>
  16. Kazdin, A. E. (2011). Evidence-based treatment research: Advances, limitations, and next steps. *American Psychologist*, 66(8), 685. <https://doi.org/10.1037/a0024975>
  17. El-Miedany, Y. (2017). Telehealth and telemedicine: how the digital era is changing standard health care. *Smart Homecare Technology and Telehealth*, 43-51. <https://doi.org/10.2147/SHTT.S116009>
  18. Adeghe, E. P., Okolo, C. A., & Ojeyinka, O. T. (2024). A review of emerging trends in telemedicine: Healthcare delivery transformations. *International Journal of Life Science Research Archive*, 6(1), 137-147. <https://doi.org/10.53771/ijlsra.2024.6.1.0040>
  19. Bitar, H., & Alismail, S. (2021). The role of eHealth, telehealth, and telemedicine for chronic disease patients during COVID-19 pandemic: A rapid systematic review. *Digital health*, 7, 20552076211009396. <https://doi.org/10.1177/20552076211009396>
  20. Ackroyd, S., & Hughes, J. A. (1992). *Data collection in context*.
  21. Reynolds, C. R., Altmann, R. A., & Allen, D. N. (2021). The problem of bias in psychological assessment. In *Mastering modern psychological testing: Theory and methods* (pp. 573-613). Cham: Springer International Publishing. [https://doi.org/10.1007/978-3-030-59455-8\\_15](https://doi.org/10.1007/978-3-030-59455-8_15)
  22. Loewenthal, K. M., & Lewis, C. A. (2020). *An introduction to psychological tests and scales*. Routledge.
  23. Burger, J., Isvoranu, A. M., Lunansky, G., Haslbeck, J., Epskamp, S., Hoekstra, R. H., ... & Blanken, T. F. (2023). Reporting standards for psychological network analyses in cross-sectional data. *Psychological methods*, 28(4), 806. <https://doi.org/10.1037/met0000471>
  24. Urbina, S. (2014). *Essentials of psychological testing*. John Wiley & Sons.
  25. Irvine, S. H., & Kyllonen, P. C. (Eds.). (2013). *Item generation for test development*. Routledge.
  26. Cohen, R. J., Swerdlik, M. E., & Phillips, S. M. (1996). *Psychological testing and assessment: An introduction to tests and measurement*. Mayfield Publishing Co.
  27. Dawes, J. (2008). Do data characteristics change according to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. *International journal of market research*, 50(1), 61-104. <https://doi.org/10.1177/147078530805000106>
  28. Derrick, B., & White, P. (2017). Comparing two samples from an individual Likert question. *International Journal of Mathematics and Statistics*, 18(3), 1-13.
  29. Anastasi, A. (1976). *Psychological testing*.
  30. Чабан О.С., Хаустова О.О., Омелянович В.Ю. Психічні проблеми воєнного часу /



Монографія, Медкнига, Бібліотека практикуючого лікраря, Київ, 2023 232 ст.

31. Чабан О.С., Хаустова О.О., Асанова А.Е., Практична психосоматика: діагностичні шкали, 3 видання, виправлене і доповнене / Навчальний посібник, за ред Чабана О.С., Хаустової О.О., К., Видавничий дім Медкнига, 2021, 200 с.