

РОЗДІЛ 1

ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ПРОБЛЕМИ СОЦІОЛОГІЇ

DOI 10.26565/2227-6521-2025-54-01

УДК 316.422:303.42(075)

СТАНДАРТИЗАЦІЯ СОЦІОЛОГІЧНОГО ТЕСТУ В КОНТЕКСТІ ПРОБЛЕМИ ВИМІРЮВАННЯ В СОЦІОЛОГІЇ

Кислова Ольга Миколаївна – кандидат соціологічних наук, доцент, доцент кафедри методів соціологічних досліджень соціологічного факультету Харківського національного університету імені В.Н. Каразіна, майдан Свободи, 4, Харків, 61022, Україна, e-mail: olga.kyslova@karazin.ua, ORCID iD: <https://orcid.org/0000-0003-2241-1313>

Ніколаєвська Алла Михайлівна – кандидат соціологічних наук, доцент, доцент кафедри соціології соціологічного факультету Харківського національного університету імені В.Н. Каразіна, майдан Свободи, 4, Харків, 61022, Україна, e-mail: alla.nikolaievska@karazin.ua, ORCID iD: <https://orcid.org/0000-0002-7705-9671>

Стаття присвячена аналізу низки проблем, пов'язаних із застосуванням специфічного методу отримання соціологічної інформації – соціологічного тесту. Представлено огляд наукових підходів до розуміння самого поняття «соціологічний тест», уточнено деякі споріднені з ним поняття (зокрема «тест», «норма», «нормативна шкала»), розглянуто процедурні моменти, пов'язані із розробкою соціологічних тестів, прийомами забезпечення їхньої діагностичної спроможності. Особливу увагу приділено питанням стандартизації соціологічних тестів, оскільки ця процедура, по-перше, уніфікує процес тестування та забезпечує об'єктивність і відтворюваність, а, по-друге, є важливою передумовою перевірки соціологічного тесту на надійність, валідність та інваріантність, що визначають загалом якість цього вимірювального інструменту. Зазначено, що, на відміну від психологічної науки, де стандартизація стала невід'ємною частиною дослідницької практики, соціологічні опитувальники часто створюються без комплексної перевірки внутрішньої структури, що призводить до значної варіабельності отриманих результатів і ускладнює порівняння даних різних дослідницьких центрів. За відсутності інструментів соціологічні дослідження ризикують стати фрагментарними та непридатними для процедури порівняння, що, у свою чергу, обмежує можливість формування кумулятивного соціологічного знання, застосування отриманих даних у міждисциплінарних і міжнародних порівняльних дослідженнях. Особливо гостро ці проблеми знаходять прояв у вимірюванні багатовимірних і контекстно залежних соціальних конструктів (соціальної довіри, громадянської активності, соціального статусу тощо). У статті також окреслено низку методологічних орієнтирів для практичної реалізації процедури стандартизації соціологічних тестів, зокрема схарактеризовані два аспекти стандартизації та загальна її процедура, процес та процедура нормалізації шкали тощо.

Ключові слова: соціологічні методи, психологічні тести, соціологічний тест, вимірювання соціальних процесів, стандартизація, шкалювання.

Для цитування: Кислова О. М., Ніколаєвська А. М. Стандартизація соціологічного тесту в контексті проблеми вимірювання в соціології. *Вісник Харківського національного університету імені В. Н. Каразіна. Серія Соціологічні дослідження сучасного суспільства: методологія, теорія, методи*. 2025. Випуск 54. С. 7-16. doi: <https://doi.org/10.26565/2227-6521-2025-54-01>

How to Cite: Kyslova, O., & Nikolaievskaya, A. (2025) 'Standardization of sociological tests in the context of measurement problem in sociology'. *Visnyk of V. N. Karazin Kharkiv National University. Series Sociological studies of contemporary society: methodology, theory, methods*. 54, pp. 7-16. doi: <https://doi.org/10.26565/2227-6521-2025-54-01> [in Ukrainian]

Постановка проблеми. Арсенал методів сучасної соціології перебуває в стані динамічного розвитку: удосконалюються традиційні методи збору інформації, з'являються нові інструменти завдяки досягненням у сфері інформаційних технологій, модернізується процес аналізу та обробки зібраних даних тощо. Постійне оновлення інструментів збору, аналізу та інтерпретації емпіричних даних зумовлюється як внутрішніми потребами соціологічної науки, так і впливом суміжних дисциплін, передусім психології.



Одним із напрямів міждисциплінарного запозичення стало впровадження тестових методик — інструментів, які в психології здобули статус провідних засобів психодіагностики, а в соціології дедалі частіше розглядаються як ефективний засіб вивчення соціальних феноменів (див. [1; 2; 3]).

Як зазначала Н. Паніна [2], проблема соціологічного тестування є частиною ширшого кола питань, пов'язаних із вимірюванням у соціології. Водночас, на її думку, емпірична соціологія значно поступається психології за рівнем методичної розробленості та технічного забезпечення вимірювальних процедур. Психометрія може слугувати орієнтиром у розв'язанні теоретико-методологічних і практичних труднощів соціологічного вимірювання. Серед таких труднощів дослідниця називає:

- неможливість порівнювати дані про один і той самий об'єкт, отримані в різних дослідженнях;
- низьку точність вимірювальних шкал і, як наслідок, їхню слабку чутливість до змін у соціальному об'єкті;
- відсутність чітко зафіксованих характеристик якості методик (таких, як надійність і валідність);
- відсутність нормативних показників, що унеможлиблює стандартизовану інтерпретацію результатів [2, с. 59].

Виправити ситуацію, на її думку, могли б зусилля, спрямовані на створення стандартизованих соціологічних методик, які мають бути водночас компактними, надійними й валідними. Саме на цьому перетині – між теоретичними викликами та потребою в практичних інструментах – й актуалізується проблема стандартизації соціологічного тесту.

На відміну від психології, де стандартизація тестів давно стала невід'ємною частиною дослідницької практики і передбачає ретельне апробування, нормування, перевірку надійності, валідності та уніфікацію процедур проведення, у соціології подібна увага часто відсутня. Соціологічні опитувальники та шкали часто створюються *ad hoc*, без комплексної перевірки їхньої внутрішньої структури, що призводить до значної варіабельності результатів і ускладнює порівняння даних.

Особливо гостро це проявляється у вимірюванні багатовимірних і контекстно залежних соціальних конструктів, наприклад соціальної довіри, громадянської активності, соціального статусу тощо. Без стандартизованих інструментів дослідження ризикують стати фрагментарними, їхні дані неможливо порівнювати з результатами інших досліджень, що суттєво обмежує можливість формування кумулятивного соціологічного знання, а також застосування отриманих даних у міждисциплінарних і міжнародних порівняльних дослідженнях.

Стандартизація є основою шкалювання (побудови шкали вимірювання), тобто визначення, які числові значення відповідають різним рівням вираженості досліджуваної ознаки. У контексті розробки тестів (психодіагностичних, педагогічних чи соціологічних) шкалювання полягає у перетворенні сирих тестових балів на шкальні значення, які дозволяють порівнювати результати між респондентами, групами чи вибірками незалежно від конкретного тесту, кількості запитань чи особливостей розподілу результатів. Отже, стандартизація відіграє вирішальну роль у розробці тесту, адже саме вона забезпечує можливість інтерпретації та порівняння результатів тестування.

Аналіз останніх досліджень і публікацій. Обговорення сутності поняття «соціологічний тест» та його відмінностей від «психологічного тесту» представлено у публікаціях Н. Паніної, Є. Головахи, С. Дембіцького [2; 3-7]. Соціологічні тести розглядаються в контексті проблем вимірювання психологічних, соціально-психологічних та соціальних феноменів «...як комплексні вимірювальні інструменти, рівень обґрунтованості котрих (як засобів отримання достовірного соціологічного знання) дещо вищий за той, якого можливо досягти в соціологічній практиці» [7, с. 7-8].

Водночас варто зауважити, що в англomовній соціологічній літературі термін «sociological test» у значенні спеціальної вимірювальної методики, адаптованої для соціологічного дослідження, майже не використовується. Натомість частіше застосовується ширше поняття «tests», яке охоплює різноманітні психометричні методики, зокрема ті, що використовуються для вивчення соціальних феноменів. Яскравим прикладом є методика вивчення цінностей Ш. Шварца, що має придатний для масових опитувань формат, орієнтована на аналіз соціальних явищ, а результати її застосування широко використовуються для оцінки стану різних суспільств [3, с. 37]. Крім того, в науковому дискурсі набув поширення термін «sociology of testing» [8; 9; 10], що позначає напрям досліджень, зосереджений на вивченні розповсюджені тестової практики в умовах цифрових інновацій, а також на пов'язаних з цим соціальних процесах, зокрема трансформації повсякденного життя під впливом перманентного тестування. У цьому контексті тест розглядається не лише як інструмент дослідження (вимірювання) певних соціальних характеристик, а як форма соціального впливу, що бере участь у регуляції доступу до ресурсів, соціальної мобільності тощо.

Проблематика стандартизації тестів, зокрема соціологічних, зазвичай розглядається у зв'язку з розробкою тестових методик, призначених для вимірювання складних абстрактних конструктів і забезпеченням їхньої валідності та надійності [11-13]. Витоки сучасних підходів сягають робіт таких дослідників, як Льюїс Терман [14], Френк Н. Фріман [15], Чарльз Спірмен [16], Лі Кронбах [17], чії концепції стали основою для сучасної психометрії.

Незважаючи на значні досягнення, проблема стандартизації тестів залишається актуальною і сьогодні: тривають дискусії щодо вдосконалення традиційних підходів, пропонуються нові рішення, активно осмислюються різні аспекти цієї процедури у світлі сучасних методологічних викликів [18-22]. У межах сучасних академічних обговорень можна виокремити низку концептуальних напрямів, що відображають трансформацію уявлень про стандартизацію та її роль у забезпеченні надійності соціологічних вимірювань:

- по-перше, тривають дискусії щодо статистичних аспектів стандартизації. [23-25], включаючи питання нормалізації, обробки сирих балів, побудови шкал та валідації результатів. Водночас стандартизація дедалі частіше розглядається не лише як технічна процедура уніфікації, а й як засіб забезпечення об'єктивності, етичності та репрезентативності тестових вимірювань [26; 27];

- по-друге, значна увага приділяється проблемі інваріантності вимірювання (Measurement Invariance), обговорюються різні формалізовані моделі [28-29], а також дискутуються проблеми та обмеження формалізації, особливо в контексті міжкультурних досліджень [30; 31];

- по-третє, у сфері соціологічного тестування зростає увага до проблем адаптації тестів, що виходять за межі лінгвістичних питань та потребують урахування соціального й культурного контексту цільової вибірки для забезпечення обґрунтованості порівняльних результатів [32; 33].

Попри активізацію академічного інтересу, проблеми стандартизації соціологічних тестів потребують подальшого системного вивчення як у теоретичному, так і в прикладному вимірах, оскільки ця процедура не лише уніфікує процес тестування та забезпечує об'єктивність і відтворюваність, а також є передумовою перевірки надійності, валідності, інваріантності, що визначають якість тесту як вимірювального інструменту.

Метою статті є визначення можливих підходів до стандартизації соціологічного тесту, уточнення пов'язаних з цим понять та окреслення методологічних орієнтирів для практичної реалізації процедури стандартизації в умовах соціологічного дослідження.

Виклад основного матеріалу. Насамперед варто звернутися до уточнення поняття соціологічного тесту та його специфічних рис. Як відомо, у психології метод тестування спочатку виник як основний інструмент дослідження в межах диференційної психології. Згодом тести набули статусу одного з ключових засобів психодіагностики. У 1890-х роках психологічні та антропометричні тести почали активно застосовуватися у США, а вже на початку XX століття вони стали в цій країні основним джерелом наукової інформації про дитячу психіку та ефективним інструментом професійного відбору. Починаючи з кінця 1950-х років, соціологи почали виявляти інтерес до тестових методик, досліджуючи їхні евристичні можливості в контексті соціологічних досліджень [1, с. 303]. Зокрема, відомі дослідження американських авторів, в яких використовувалися ігрові методики, тест незавершених речень (SCT), тематичний апперцептивний тест (ТАТ) та інші інструменти [34]. Безумовно, не всі психодіагностичні методики придатні для використання в соціологічних дослідженнях, проте окремі з них вже продемонстрували свою ефективність і практичну цінність. Насамперед йдеться про модифікації згаданих вище проєктивних методик, тестів, побудованих за зразком методики С. Розенцвейга (PF-Study), а також малюнкових методик.

В українській соціології тестові методики почали використовувати лише наприкінці 1990-х років. Насамперед йдеться про дослідження Н. Паніної, Є. Головахи, а згодом С. Дембіцького. Особливої уваги заслуговують напрацювання Н. Паніної, яка у 1998 році в статті «Вимірювання в соціології та принципи застосування тестів у масових опитуваннях» наголосила не лише на необхідності адаптації класичних психологічних тестів до потреб соціологічного дослідження, а й підкреслила важливість створення повноцінних соціологічних тестів як самостійного інструменту вимірювання.

Дослідниця визначила низку критеріїв, яким має відповідати соціологічний тест:

- універсальність: методика повинна охоплювати найголовніші, ключові індикатори досліджуваного феномену, що забезпечує її придатність для використання в різних типах соціологічних досліджень;

- інтегральність: тест має дозволяти узагальнення різних вимірів в єдиний інтегрований показник;

- якісність: тестова методика повинна мати підтверджені статистичні характеристики надійності та валідності;

- стандартизованість: необхідно встановити нормативи для досліджуваного явища, що уможливило б інтерпретацію результатів з позицій інтенсивності прояву відповідного феномену;

- компактність: тест має бути достатньо стислим, щоб його можна було легко включати до масових опитувань;

- «чутливість»: незважаючи на стислість, тест повинен забезпечувати достатній рівень точності вимірювання;

- об'єктивність: цілі тестування мають бути неочевидними для респондента, що дозволяє зменшити ймовірність свідомого викривлення відповідей [2, с. 67].

Соціологічний тест суттєво відрізняється від психологічного як за своїм призначенням, так і за підходами до побудови й застосування. Насамперед ці два типи тестів орієнтовані на різні об'єкти

вимірювання. Психологічний тест розробляється для діагностики таких індивідуальних психологічних особливостей, як риси особистості, інтелектуальні здібності, емоційні стани чи мотиваційні настанови. Його головна мета – отримати інформацію про психічний стан окремої людини або спрогнозувати її поведінку в певних умовах. Натомість соціологічний тест фокусується на вивченні соціальних характеристик – цінностей, соціальних установок, рівня довіри, політичних орієнтацій тощо. Його завдання полягає не в індивідуальній діагностиці, а у виявленні типових закономірностей і тенденцій у масштабі великих соціальних груп або всього суспільства.

Ці відмінності зумовлюють різний масштаб застосування. Психологічні тести найчастіше використовуються в індивідуальному форматі або в невеликих групах, зокрема в клінічній практиці, системі освіти або під час відбору персоналу. Соціологічні тести, навпаки, розраховані на масове використання, зокрема в репрезентативних опитуваннях, де важлива швидкість, ефективність і компактність інструменту.

Соціологічне тестування можна розглядати у трьох взаємопов'язаних вимірах (див.[3; 4]). Передусім йдеться про методичний вимір, який передбачає дотримання вимог до валідності й надійності методики, що дозволяє застосовувати її в масових опитуваннях. У цьому сенсі соціологічне тестування охоплює як спеціально розроблені соціологічні інструменти, так і адаптовані психологічні методики. Другий вимір пов'язаний зі змістовною спрямованістю тесту: він має бути орієнтований на вивчення соціальних явищ і процесів. Це зумовлює необхідність побудови концептуальної основи тесту саме на засадах соціологічного знання. Третій вимір стосується аналітичного потенціалу соціологічного тесту, тобто можливості використовувати отримані результати не лише для опису чи вимірювання, але й для ширших теоретичних висновків щодо стану та функціонування соціальної системи. Як зазначає С. Дембіцький, навіть за відсутності попередньої теоретичної моделі такі висновки можуть формуватися в процесі емпіричного застосування тесту [3, с. 36-37; 4, с. 144].

Таким чином, дослідники не лише актуалізували необхідність адаптації психологічних методик для соціологічних цілей, а й окреслили підходи до створення власне соціологічних тестів як самостійних вимірювальних інструментів, що розширюють можливості збору соціологічної інформації. Так, Н. Паніна підкреслювала важливість дотримання процедур вимірювання, характерних для психометрії, акцентуючи на потребі розробки тестів, здатних адекватно відображати соціальний контекст. С. Дембіцький представив концепцію соціологічного тесту як індикаторної структури, що дозволяє інтегрувати результати емпіричних досліджень і узагальнювати їх у межах складних соціальних конструктів.

Попри міждисциплінарний характер соціологічного тестування, саме класичні підходи до стандартизації, напрацьовані в рамках психометрії, залишаються ключовим методологічним орієнтиром. Їхнє застосування дає змогу уніфікувати процедури розробки, апробації та інтерпретації тестів, підвищити об'єктивність вимірювання та забезпечити порівнюваність результатів у різних дослідницьких контекстах. У зв'язку з цим ми й пропонуємо звернутися до проблематики стандартизації тестів, зокрема до основних концепцій і процедур, що лежать в її основі.

Тест визначається як опитувальник (анкета), що містить низку пунктів (завдань, відкритих чи закритих запитань), відповіді на які оцінюються певною кількістю балів та об'єднуються відповідним чином для отримання загального (складеного) балу за весь тест. Всі пункти тесту повинні стосуватися вимірюваного конструкту, у тест не включають запитань, що не мають до нього відношення, щоб загальна (складена) оцінка також була індикатором цього конструкту (докладніше див. [20]). Такі додаткові запитання, як соціально-демографічний блок, не входять до тесту і аналізуються окремо. Якщо порівнювати тести зі звичними для соціологів соціологічними анкетами, то тести як засоби складних конструктів можуть бути їхньою частиною. Отже, основними характеристиками тесту є такі: (а) серія питань, на які відповідає респондент (піддослідний), та (б) складений бал, який виникає в результаті оцінювання відповідей на ці питання. Отриманий набір питань разом називається «тестом», «шкалою» або «вимірюванням» [35]. Популярність тестів зумовлюється тим, що вони дозволяють сконструювати кількісну шкалу для вимірювання латентного конструкту, ґрунтуючись на множинній інформації, пов'язаній з ним.

Стандартизація тесту необхідна для забезпечення можливості порівняння тестових показників, отриманих різними респондентами, а також їхнього зіставлення з результатами генеральної сукупності або окремих груп.

Стандартизація тесту має два аспекти:

1. *Стандартизація процесу тестування.* Вона передбачає створення однакових умов проходження тесту, тобто уніфікованість та регламентованість процедури тестування, зокрема розробку точних інструкцій щодо тестування, визначення умов, за яких не можна використовувати методику, встановлення критеріїв придатності методики для певних груп піддослідних (респондентів) тощо.

2. *Стандартизація інтерпретації результатів тестування.* Йдеться про створення системи норм – стандартизованих статистичних показників, що відображають положення індивіда щодо вибірки стандартизації (нормативної групи). На цьому етапі «сирі» оцінки (первинні бали) трансформуються у зручні для аналізу шкальні показники, що дають можливість здійснювати прозору та обґрунтовану

інтерпретацію результатів, а також порівнювати їх між різними групами чи дослідженнями. На цьому етапі залучається арсенал методів математичної статистики.

Результати тестування кожного з учасників тестування представляються у вигляді певного числа, тобто кількості набраних балів. Це число називають первинною тестовою оцінкою або «сирим» балом. Проте для інтерпретації цих чисел, для розуміння отриманого результату у порівнянні з іншими здійснюється штучне перетворення, засноване на аналізі емпіричного розподілу тестових оцінок у репрезентативній вибірці піддослідних. Процедура такого перетворення часто зветься нормалізацією, оскільки вона пов'язана зі встановленням норм та побудовою нормативної шкали.

Норма – це стандартизований статистичний показник, що відображає типовий рівень виконання тесту в межах певної референтної групи. Вона слугує основою для інтерпретації індивідуального результату: дозволяє оцінити, наскільки він відрізняється від середнього (нормативного) рівня, чи перебуває в межах «нормального» діапазону, або ж істотно вищий чи нижчий за нього. Зазвичай норми розробляються окремо для різних соціальних груп з урахуванням віку, статі, рівня освіти, культурного контексту тощо. Вони виступають статистично обґрунтованим орієнтиром, який фіксує те, що вважається типовим або очікуваним для певної категорії респондентів.

Основою визначення норм є статистичний підхід, який передбачає емпіричне тестування репрезентативної вибірки (вибірки стандартизації). Після збору результатів проводиться їхній кількісний аналіз: будують розподіл вимірюваної ознаки (конструкту), обчислюють середнє значення та стандартне відхилення. На основі цих статистичних параметрів створюється нормативна шкала або таблиця, яка й використовується для інтерпретації індивідуальних результатів.

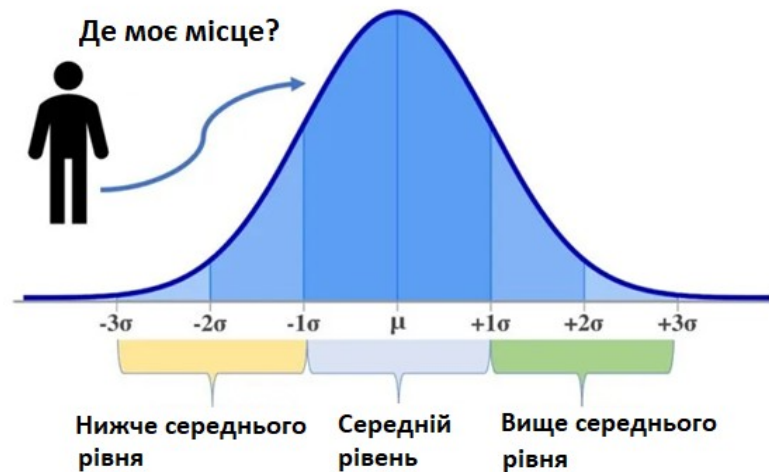


Рисунок 1. Нормальний розподіл як основа для визначення норм (μ – середнє значення, σ – стандартне відхилення).

Figure 1. Normal distribution as a basis for determining norms (μ – mean, σ – standard deviation).

На рисунку зображено класичну модель нормального (гаусівського) розподілу, яка широко використовується в тестуванні. Вона демонструє, як індивідуальний результат співвідноситься із середнім значенням і стандартними відхиленнями, що дозволяє віднести його до певного рівня: середнього, нижчого або вищого від середнього (див. рис. 1).

На основі цих даних створюється нормативна таблиця або шкала, яка й використовується для інтерпретації індивідуальних результатів. Саме вона дає змогу співвіднести первинну оцінку («сирий» бал) конкретної особи з розподілом оцінок, отриманих на вибірці стандартизації, та з'ясувати, яке місце посідає ця оцінка у розподілі: чи відповідає її показник середньому результату групи, де проводилася стандартизація тесту, чи він вищий або нижчий за середній.

Нормативна шкала – це інструмент для інтерпретації «сирих» результатів тесту шляхом їхнього зіставлення зі статистичними показниками, отриманими на вибірці стандартизації. Вона дає змогу визначити, наскільки індивідуальний бал відрізняється від типових (нормативних) значень у популяції. Залежно від цілей дослідження, нормативна шкала може подаватися у формі процентильних рангів, категоріальних рівнів («низький», «середній», «високий») або стандартизованих оцінок.

Слід зазначити, що нормативні шкали завжди ґрунтуються на таких стандартизованих показниках, як процентилі (відсоткові ранги), Z-оцінки, T-бали тощо. Вони дають змогу перетворити первинний «сирий» бал (тобто кількість балів, набраних респондентом під час тестування) на уніфіковану й інтерпретовану оцінку. Ці статистичні стандарти є універсальними засобами порівняння результатів у межах референтної групи. Втім, така інтерпретація може бути недостатньо зручною або зрозумілою для

практичного використання, особливо у прикладних дослідженнях. Саме тому розробляються спеціалізовані нормативні шкали, адаптовані під конкретні конструкти та цільові групи. Прикладами таких шкал у психології можна, наприклад, назвати відповідні інструменти в тестах інтелекту (Wechsler Adult Intelligence Scale (WAIS) чи прогресивних матрицях Равена) або у відомому особистісному опитувальнику MMPI (Minnesota Multiphasic Personality Inventory), призначеному для оцінювання особистісних характеристик і психічних відхилень. Ці шкали створюються на основі стандартизованих даних, але мають власну систему балів, інтерпретаційні межі та рекомендації для практичного тлумачення результатів.

Підкреслимо, що всі ці шкали пов'язані з нормальним розподілом та стандартизованими показниками, наприклад на рис. 2 показано такий зв'язок шкали IQ.



Рисунок 2. Нормальна крива та стандартизовані оцінки

Figure 2. Normal curve and standardized scores

Водночас у контексті психологічних досліджень також широко застосовуються інтерпретаційні шкали типу стенонів (stanine), які ще називають рівневими шкалами. Вони поділяють континуум результатів на обмежену кількість інтервалів (зазвичай дев'ять), кожен з яких має фіксовану інтерпретацію від «дуже низького» до «дуже високого» рівня (зв'язок з нормальним розподілом та стандартними статистичними показниками див. на рис. 2).

У соціологічній практиці також широко застосовуються п'ятибальні шкали — простіші рівневі системи, які класифікують результати за п'ятьма категоріями: дуже низький, низький, середній, високий і дуже високий рівень. Такі шкали часто використовуються як в опитувальниках (на рівні респондентських відповідей), так і на етапі інтерпретації узагальнених результатів після статистичної обробки «сирих» балів.

Переваги таких інтерпретаційних шкал полягають у тому, що вони:

- полегшують інтерпретацію даних для широкої аудиторії, включно з нефаківами;
- підвищують комунікативну ефективність, зводячи складні статистичні оцінки до зрозумілих рівнів;
- сприяють узагальненню результатів при побудові типологій, порівнянь між групами, виявленні закономірностей;
- адаптовані до масових опитувань, в умовах яких важливо уникати перевантаження деталями.

У соціології, де результат повинен бути не лише статистично обґрунтованим, але й зрозумілим у публічному дискурсі, такі шкали забезпечують баланс між точністю та доступністю. У соціологічних дослідженнях перетворення шкал відіграє важливу роль, оскільки дозволяє зробити результати вимірювань осмисленими, порівнюваними та інтерпретованими в ширшому контексті. Сирі бали, які отримує дослідник після застосування тесту чи опитувальника, не дають повної уяви про значення отриманого результату. Самі по собі вони не показують, чи є результат «високим», «низьким» або «середнім» — для цього потрібне нормування, тобто співвіднесення індивідуальних результатів із розподілом результатів у генеральній сукупності чи вибірці.

Висновки та перспективи подальших досліджень. Проблема вимірювання в соціології як «внутрішня» методологічна проблема наукової соціологічної спільноти на рівні суспільної свідомості й громадської думки перетворюється на питання довіри до вимірювання складних соціальних явищ, а врешті-решт – й до самої соціології. Такі явища вимагають певних стандартизованих вимірювальних інструментів (насамперед може йтися про так званий соціологічний тест), розробка яких наразі, на жаль, поки не входить до переліку пріоритетних напрямів діяльності фахівців з методології та методів соціологічних досліджень. Створення соціологічного тесту вимагає, у свою чергу, дотримання низки вимог, які напрацьовані психологічною наукою в рамках психометрії й залишаються ключовим методологічним орієнтиром, незважаючи на часом принципові відмінності між психологічним і соціологічним тестом. Особливої уваги вартують питання стандартизації соціологічного тесту, оскільки вона є передумовою його перевірки на надійність, валідність та інваріантність, що визначають загалом не тільки якість цього вимірювального інструменту, а й врешті-решт і якість соціологічної інформації.

Перспективи подальших досліджень у галузі стандартизації соціологічних тестів пов'язані перш за все з необхідністю подолання низки методологічних викликів, що ускладнюють забезпечення валідності та порівнюваності результатів у сучасних соціокультурних умовах. Насамперед актуальною є потреба у розробці адаптивних моделей стандартизації, здатних враховувати специфіку локальних контекстів, динаміку соціальних змін і нові форми соціальної реальності, що потребують вимірювання. Подальші дослідження можуть зосередитися на комбінуванні нормативного та критеріального підходів до побудови шкал, вивченні інваріантності соціологічних тестів у міжгрупових і міжкультурних порівняннях, а також на етичних аспектах стандартизації. Перспективним напрямом є також інтеграція цифрових технологій у процес тестування, що відкриває нові можливості для стандартизації процедур і збору емпіричних даних. Загалом розвиток теоретико-методологічного підґрунтя стандартизації сприятиме підвищенню наукової обґрунтованості соціологічного вимірювання та його практичної релевантності.

Список використаної літератури:

1. Ніколаєвська А. М. Використання проєктивних методик в соціологічних дослідженнях. *Вісник Харківського державного університету*. 1999. № 439. *Матеріали III Харківських Міжнародних психологічних читань*. Част. 4, 5. С. 303-305.
2. Паніна Н. Вимірювання в соціології та принципи застосування тестів у масових опитуваннях. *Соціологія: теорія, методи, маркетинг*. 1998. № 1-2. С. 53-68.
3. Дембіцький С. *Розробка соціологічних тестів: методологія і практики її застосування*. Київ: Інститут соціології НАН України, 2019. 304 с.
4. Дембіцький С. С. Соціологічні тести: сутність і валідизація. *Соціологія: теорія, методи, маркетинг*. 2016. № 3. С. 140-155.
5. Дембіцький С. Принципи визначення критеріальних рівнів соціологічного тесту. *Український соціум*. 2017. № 4. С. 22-34.
6. Дембіцький С. С. Вимірювання геополітичних орієнтацій в Україні: конструювання соціологічних тестів "RF-geopol" і "EU-geopol". *Соціологія: теорія, методи, маркетинг*. 2018. № 4. С. 105-134.
7. *Комплексні вимірювальні інструменти в соціологічних дослідженнях: розроблення, адаптація, обґрунтування достовірності* / Є. Головаха, С. Дембіцький, Н. Паніна та ін.; за наук. ред. Є. Головахи та С. Дембіцького. Київ: Інститут соціології НАН України, 2022.
8. Pinch T. "Testing-One, Two, Three... Testing!": Toward a Sociology of Testing. *Science, Technology, & Human Values*. 1993. Vol. 18(1). P. 25-41. DOI: <https://doi.org/10.1177/016224399301800103>
9. Marres N., Stark D. Put to the test: For a new sociology of testing. *British Journal of Sociology*. 2020. Vol. 71(3). P. 423–443. DOI: <https://onlinelibrary.wiley.com/doi/full/10.1111/1468-4446.12746>
10. Potthast J. The sociology of conventions and testing. *Social theory now*, edited by Claudio E. Benzecry, Monika Krause and Isaac Ariail Reed, Chicago: University of Chicago Press, 2017, 337-360. DOI: <https://doi.org/10.7208/9780226475318-013>
11. Edwards N. T. The historical and social foundations of standardized testing: In search of a balance between learning and evaluation. *Shiken and Shiken Research Journal: JALT Testing and Evaluations SIG Newsletter*. 2006. Vol. 10. P. 8-16.
12. Friedrich W. N., Olafson E., Faller K. C. Standardized tests and measures. *Interviewing children about sexual abuse: Controversies and best practice*. Oxford: Oxford University Press, 2007. P. 207-225.
13. Fischer R., Milfont T. L. Standardization in psychological research. *International Journal of Psychological Research*. 2010. Vol. 3(1). P. 88-96. DOI: <https://doi.org/10.4236/psych.2018.911145>
14. Terman L. M. *The measurement of intelligence: an explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon intelligence scale*. Boston: Houghton Mifflin Company, 1916. URL: <https://www.gutenberg.org/files/20662/20662-h/20662-h.htm>
15. Freeman F. N. *Mental tests: Their history, principles and applications*. Houghton, Mifflin and Company: 1926. <https://doi.org/10.1037/10738-000>
16. Spearman C. The Proof and Measurement of Association Between Two Things. *Studies in individual differences: The search for intelligence* / J. J. Jenkins & D. G. Paterson (Eds.). New York: Appleton-Century-Crofts, 1961. P. 45–58. <https://doi.org/10.1037/11491-005>
17. Cronbach L.J. *Essentials of Psychological Test*. New York: Harper Routledge, 1970.

18. Havârneanu G. Validity Criteria of a Standardized Test as an Opportunity for Efficient Assessment Created from the Teleological Perspective of Incremental Learning. *Science Journal of Education*. 2023. Vol. 11 (4). P. 142-149. DOI: <https://doi.org/10.11648/j.sjedu.20231104.14>
19. Bas A., et al. Understanding the development, standardization, and validation process of alternative in vitro test methods for regulatory approval from a researcher perspective. *Small*. 2021. Vol. 17(15). DOI: <https://doi.org/10.1002/sml.202006027>
20. Kyriazos T. A., Stalikas A. Applied psychometrics: The steps of scale development and standardization process. *Psychology*. 2018. Vol 9(11).P. 2531-2560. DOI: <https://doi.org/10.4236/psych.2018.911145>
21. Eteng-Uket S. The Development, Validation, and Standardization of a New Tool: The Dyscalculia Test. *Numeracy*. 2023. Vol.16(2). P. 1. DOI: <https://doi.org/10.5038/1936-4660.16.2.1417>
22. Himelfarb I. A primer on standardized testing: History, measurement, classical test theory, item response theory, and equating. *Journal of Chiropractic Education*. 2019. Vol. 33(2). P. 151-163. DOI: <https://doi.org/10.7899/JCE-18-22>
23. Lee S. Advanced Statistical Techniques for Test Standardization: Enhancing Reliability and Validity in Neuropsychological Assessments. *NUMBERANALYTICS*, 2025. URL: https://www.numberanalytics.com/blog/advanced-statistical-techniques-for-test-standardization#google_vignette
24. Flake J. K., Fried E. I. Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in methods and practices in psychological science*. 2020. Vol. 3(4). P. 456-465. DOI: <https://doi.org/10.1177/2515245920952393>
25. Fabian M. Scale norming undermines the use of life satisfaction scale data for welfare analysis. *Journal of Happiness Studies*. 2022. Vol. 23(4). P. 1509-1541. DOI: <https://doi.org/10.1007/s10902-021-00460-8>
26. Leitgöb H., et al. Measurement invariance in the social sciences: Historical development, methodological challenges, state of the art, and future perspectives. *Social science research*. (2023). Vol. 110. DOI: <https://doi.org/10.1016/j.ssresearch.2022.102805>
27. Meuleman B., et al. Why measurement invariance is important in comparative research. A response to Welzel et al. *Sociological Methods & Research*. 2023. Vol. 52(3). P. 1401-1419. DOI: <https://doi.org/10.1177/00491241221091755>
28. Welzel, Christian, et al. Non-invariance? An overstated problem with misconceived causes. *Sociological Methods & Research*. 2023. Vol.52 (3). P. 1368-1400. DOI: <https://doi.org/10.1177/004912412195521>
29. Kusano K., Napier J., Jost J. The mismeasure of culture: When measurement invariance requirements hinder cross-cultural research in psychology. *Preprint*. (2024). DOI: <https://doi.org/10.31234/osf.io/9qe2k>
30. Bommasani R., Liang P. Trustworthy social bias measurement. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*. 2024. Vol. 7. P. 210-224). DOI: <https://doi.org/10.1609/aies.v7i1.31630>
31. Jacobs A. Z., & Wallach H. Measurement and fairness. *FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. 2021. P. 375 – 385. DOI: <https://doi.org/10.1145/3442188.3445901>
32. Ponce Gea A. I., Serrano Pastor F. J. The Cultural Element in the Adaptation of a Test: Proposals and Reflections on Internal and External Influences. *Education Sciences*. 2022. Vol.12(5). P. 291. DOI: <https://doi.org/10.3390/educsci12050291>
33. Council of the International Test Commission. ITC guidelines for translating and adapting tests. *International Journal of Testing*. 2018. Vol. 18(2). P. 101-134. <https://doi.org/10.1080/15305058.2017.1398166>
34. Oppenheim A. *Questionnaire Design, Interviewing and Attitude Measurement*. London, New York: Pinter Publishers, 1992. 303 p.
35. Zumbo, B. D., Gelin, M. N., & Hubley, A. M. (2002). The construction and use of psychological tests and measures. In the Psychology theme of the *Encyclopedia of Life Support Systems (EOLSS)*, Eolss Publishers, Oxford, UK.

Внесок авторів: всі автори зробили рівний внесок у цю роботу

Конфлікт інтересів : автори повідомляють про відсутність конфлікту інтересів

Отримано 15.02.2025

Прийнято до друку 26.03.2025

STANDARDIZATION OF SOCIOLOGICAL TESTS IN THE CONTEXT OF MEASUREMENT PROBLEM IN SOCIOLOGY

Kyslova Olga – PhD (Sociology), Associate Professor, Department of Methods of Sociological Research, V. N. Karazin Kharkiv National University, 4, Svobody Sq., Kharkiv, 61022, Ukraine, e-mail: olga.kyslova@karazin.ua, ORCID iD: <https://orcid.org/0000-0003-2241-1313>

Nikolaievska Alla – PhD.(Sociology), Associate Professor, Department of Sociology, V. N. Karazin Kharkiv National University, 4, Svobody Sq., Kharkiv, 61022, Ukraine, e-mail: alla.nikolaievska@karazin.ua, ORCID iD: <https://orcid.org/0000-0002-7705-9671>

The article is devoted to the analysis of a number of problems associated with the use of a specific method of obtaining sociological information - a sociological test. An overview of scientific approaches to understanding the concept of "sociological test" is presented, some related concepts are clarified (in particular, "test", "norm", "normative scale"), procedural aspects related to the development of sociological tests, methods for ensuring their diagnostic capacity are considered. Particular attention is paid to the issues of standardization of sociological tests, since this procedure, firstly, unifies the testing process and ensures objectivity and reproducibility, and, secondly, is an important prerequisite for

checking a sociological test for reliability, validity and invariance, which generally determine the quality of this measuring tool. It is noted that, unlike psychological science, where standardization has become an integral part of research practice, sociological questionnaires are often developed without comprehensive verification of the internal structure, which leads to significant variability in the results obtained and complicates the comparison of data from different research centers. In the absence of tools, sociological research risks becoming fragmentary and unsuitable for the comparison procedure, which, in turn, limits the possibility of forming cumulative sociological knowledge, applying the obtained data in interdisciplinary and international comparative studies. These problems are particularly acute in the measurement of multidimensional and context-dependent social constructs (social trust, civic engagement, social status, etc.). The article also outlines a number of methodological guidelines for the practical implementation of the procedure for standardization of sociological tests, in particular, two aspects of standardization and its general procedure, the process and procedure for scale normalization, etc. are characterized. The article also outlines a number of methodological guidelines for the practical implementation of the procedure for standardizing sociological tests; in particular, it describes two aspects of standardization and its general procedure, the process and procedure for normalizing the scale, etc.

Keywords: sociological methods, psychological tests, sociological test, measurement of social processes, standardization, scaling.

References:

1. Nikolaievskaya A. M. (1999). 'Using projective tests in sociological research'. *Bulletin of Kharkiv State University*. No. 439. *Proceedings of the 3rd Kharkiv International Psychological Readings*. Part. 4, 5. P. 303-305 [In Ukrainian]
2. Panina, N.V. (1998). 'Measurement in sociology and the principles of using tests in mass surveys' *Sociology: theory, methods, marketing*, 1-2, 53-68 [In Ukrainian]
3. Dembitsky, S. (2019). Development of sociological tests: methodology and practices of its application. Kyiv: Institute of Sociology of the NAS of Ukraine. [In Ukrainian]
4. Dembitsky, S. (2016). 'Sociological tests: Essence and validation'. *Sociology: Theory, Methods, Marketing*, (3), 140-155 [In Ukrainian]
5. Dembitsky, S. S. (2017). 'Principles for determining criterion levels of a sociological test'. *Ukrainian Society*, (4), 22-34 [In Ukrainian]
6. Dembitsky, S. (2018). 'Measuring geopolitical orientations in Ukraine: Construction of sociological tests «RF-geopol» and «EU-geopol»'. *Sociology: Theory, Methods, Marketing*, (4), 105-134 [In Ukrainian]
7. Holovakha, Ye., Dembitsky, S., Panina, N., et al. (2022). Complex measurement instruments in sociological research: Development, adaptation, and validation (Ye. Holovakha & S. Dembitsky, Eds.). Kyiv: Institute of Sociology of the National Academy of Sciences of Ukraine [In Ukrainian]
8. Pinch, T. (1993). '«Testing-One, Two, Three... Testing!»: Toward a Sociology of Testing'. *Science, Technology, & Human Values*, 18(1), 25-41. DOI: <https://doi.org/10.1177/016224399301800103>
9. Marres, N., & Stark, D. (2020). 'Put to the test: For a new sociology of testing'. *British Journal of Sociology*, 71(3), 423-443. DOI: <https://onlinelibrary.wiley.com/doi/full/10.1111/1468-4446.12746>
10. Potthast, J. (2017). The sociology of conventions and testing. Social theory now, edited by Claudio E. Benzecry, Monika Krause and Isaac Ariail Reed, Chicago: University of Chicago Press, 337-360. DOI: <https://doi.org/10.7208/9780226475318-013>
11. Edwards, N. T. (2006). 'The historical and social foundations of standardized testing: In search of a balance between learning and evaluation'. *Shiken and Shiken Research Journal: JALT Testing and Evaluations SIG Newsletter*, 10, 8-16.
12. Friedrich, W. N., Olafson, E., & Faller, K. C. (2007). 'Standardized tests and measures'. *Interviewing children about sexual abuse: Controversies and best practice*, Oxford University Press, 207-225
13. Fischer, R., & Milfont, T. L. (2010). 'Standardization in psychological research'. *International Journal of Psychological Research*, 3(1), 88-96. DOI: <https://doi.org/10.4236/psych.2018.911145>
14. Terman, L. M. (1916). The measurement of intelligence: an explanation of and a complete guide for the use of the Stanford revision and extension of the Binet-Simon intelligence scale. Boston: Houghton Mifflin Company. Available at: <https://www.gutenberg.org/files/20662/20662-h/20662-h.htm>
15. Freeman, F. N. (1926). Mental tests: Their history, principles and applications. Houghton, Mifflin and Company. <https://doi.org/10.1037/10738-000>
16. Spearman, C. (1961). 'The Proof and Measurement of Association Between Two Things'. In J. J. Jenkins & D. G. Paterson (Eds.), *Studies in individual differences: The search for intelligence*, New York: Appleton-Century-Crofts, 45-58. <https://doi.org/10.1037/11491-005>
17. Cronbach, L.J. (1970) Essentials of Psychological Test. New York: Harper Routledge.
18. Himelfarb, I. (2019). 'A primer on standardized testing: History, measurement, classical test theory, item response theory, and equating'. *Journal of Chiropractic Education*, 33(2), 151-163. DOI: <https://doi.org/10.7899/JCE-18-22>
19. Bas, A., Burns, N., Gulotta, A., Junker, J., Drasler, B., Lehner, R., ... & Rothen - Rutishauser, B. (2021). 'Understanding the development, standardization, and validation process of alternative in vitro test methods for regulatory approval from a researcher perspective'. *Small*, 17(15), DOI: <https://doi.org/10.1002/sml.202006027>
20. Kyriazos, T. A., & Stalikas, A. (2018). 'Applied psychometrics: The steps of scale development and standardization process'. *Psychology*, 9(11), 2531-2560. DOI: <https://doi.org/10.4236/psych.2018.911145>
21. Eteng-Uket, S. (2023). 'The Development, Validation, and Standardization of a New Tool: The Dyscalculia Test'. *Numeracy*, 16(2), 1. DOI: <https://doi.org/10.5038/1936-4660.16.2.1417>

22. Havârneanu, G. (2023). 'Validity Criteria of a Standardized Test as an Opportunity for Efficient Assessment Created from the Teleological Perspective of Incremental Learning'. *Science Journal of Education*, 11(4), 142-149. DOI: <https://doi.org/10.11648/j.sjedu.20231104.14>
23. Lee, S. (2025) 'Advanced Statistical Techniques for Test Standardization: Enhancing Reliability and Validity in Neuropsychological Assessments'. *NUMBERANALYTICS*. Available at: https://www.numberanalytics.com/blog/advanced-statistical-techniques-for-test-standardization#google_vignette
24. Flake, J. K., & Fried, E. I. (2020). 'Measurement schmeasurement: Questionable measurement practices and how to avoid them'. *Advances in methods and practices in psychological science*, 3(4), 456-465. DOI: <https://doi.org/10.1177/2515245920952393>
25. Fabian, M. (2022). 'Scale norming undermines the use of life satisfaction scale data for welfare analysis'. *Journal of Happiness Studies*, 23(4), 1509-1541. DOI: <https://doi.org/10.1007/s10902-021-00460-8>
26. Bommasani, R., & Liang, P. (2024). 'Trustworthy social bias measurement'. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, Vol. 7, 210-224. DOI: <https://doi.org/10.1609/aies.v7i1.31630>
27. Jacobs, A. Z., & Wallach, H. (2021). 'Measurement and fairness'. *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. P. 375 – 385. DOI: <https://doi.org/10.1145/3442188.3445901>
28. Leitgöb, H., Seddig, D., Asparouhov, T., Behr, D., Davidov, E., De Roover, K., ... & van de Schoot, R. (2023). 'Measurement invariance in the social sciences: Historical development, methodological challenges, state of the art, and future perspectives'. *Social science research*, 110. DOI: <https://doi.org/10.1016/j.ssresearch.2022.102805>
29. Meuleman, B., Žóltak, T., Pokropek, A., Davidov, E., Muthén, B., Oberski, D. L., ... & Schmidt, P. (2023). 'Why measurement invariance is important in comparative research. A response to Welzel et al.(2021)'. *Sociological Methods & Research*, 52(3), 1401-1419. DOI: <https://doi.org/10.1177/00491241221091755>
30. Welzel, C., Brunkert, L., Kruse, S., & Inglehart, R. F. (2021). 'Non-invariance? An Overstated Problem With Misconceived Causes'. *Sociological Methods & Research*, 52(3), 1368-1400. DOI: <https://doi.org/10.1177/0049124121995521>. (Original work published 2023)
31. Kusano, K., Napier, J., & Jost, J. (2024). 'The mismeasure of culture: When measurement invariance requirements hinder cross-cultural research in psychology'. *Preprint*. DOI: <https://doi.org/10.31234/osf.io/9qe2k>
32. Ponce Gea, A. I., & Serrano Pastor, F. J. (2022). 'The Cultural Element in the Adaptation of a Test: Proposals and Reflections on Internal and External Influences'. *Education Sciences*, 12(5), 291. DOI: <https://doi.org/10.3390/educsci12050291>
33. Council of the International Test Commission. (2018). 'ITC guidelines for translating and adapting tests'. *International Journal of Testing*, 18(2), 101-134. DOI: <https://doi.org/10.1080/15305058.2017.1398166>
34. Oppenheim, A. (1992). *Questionnaire Design, Interviewing and Attitude Measurement*. London and New York: Pinter Publishers.
35. Zumbo, B. D., Gelin, M. N., & Hubley, A. M. (2002). 'The construction and use of psychological tests and measures'. In the Psychology theme of the *Encyclopedia of Life Support Systems* (EOLSS), Eolss Publishers, Oxford, UK.

Authors Contribution : All authors have contributed equally to this work

Conflict of Interest: The authors declare no conflict of interest

Received 15.02.2025

Accepted 26.03.2025