

УДК 303.214, 316.023.6

Дембіцький С.С., канд. соціол. наук, старший науковий співробітник відділу методології та методів соціології Інституту соціології НАН України, вул. Шовковична, 12, Київ, 01021, Україна, e-mail: e_forge@me.com

ПРИНЦИПИ ВИЗНАЧЕННЯ КРИТЕРІАЛЬНИХ РІВНІВ СОЦІОЛОГІЧНОГО ТЕСТУ

Досліджено поняття "норма" у прив'язці до соціологічного тестування. На базі методологічних міркувань запропоновано його заміну на більш прийнятний термін – "критеріальний рівень", що розуміється як числове значення соціологічного тесту, яке позначає якісно відмінний рівень досліджуваної характеристики. Властивостями критеріальних рівнів є партикулярність, якісність і дискримінативність. Для демонстрації теоретичних положень статті наведено результати визначення критеріальних рівнів соціологічного тесту SCL-9-NR.

Ключові слова: соціологічне тестування, норми, критеріальні рівні, психологічний дистрес, SCL-9-NR.

Під соціологічним тестом автор розуміє систему відібраних соціологічними методами висловлювань, які пред'являються респондентам з метою отримання надійної та валідної інформації про ознаки, що досліджуються [1, с. 41]. При цьому соціологічні тести розглядаються в трьох вимірах, які частково перетинаються: методичному (стислість при збереженні показників надійності та валідності), концептуальному (звернення до вивчення соціальних феноменів) і прогностичному (діагностичний потенціал тесту при оцінці стану соціальної системи) [2, с. 144]. На нашу думку, забезпечення принципів принаймні методичного виміру дозволяє вважати шкалу або інтегральний індекс¹ соціологічним тестом.

Крім того, існує думка, що ключовою ознакою тесту (психологічного чи соціологічного) є наявність норм, які дозволяють здійснювати інтерпретацію індивідуальних результатів. Такої думки дотримується ряд наукових співробітників відділу методології та методів соціології Інституту соціології НАН України. Не вважаючи таку постановку питання досить обґрунтованою, тому що поняття "норма" в тестології є дуже неоднозначним, ми визнаємо його важливість у контексті розробки та валідації соціологічних тестів.

Оскільки поняття "норма" в соціологічному тестуванні не досліджувалося раніше, **метою статті** є його критичний розгляд і соціологічно обґрунтована концептуалізація. Для досягнення зазначеної мети стаття передбачає вирішення таких завдань:

- розглянути застосування норм у психологічному тестуванні;
- проаналізувати поняття "норма" та "критеріальний рівень" у контексті соціологічного тестування;
- продемонструвати положення статті на прикладі визначення критеріальних рівнів соціологічного тесту SCL-9-NR.

¹ Про відмінність між шкалами та індексами див.: [3].

Застосування норм у психологічному тестуванні. Поняття “норма” стосовно психологічного тестування належить до результатів, які показують представники певної групи за заданим тестом [4, р. 53]. Такі норми обчислюються за допомогою досліджень стандартизації та надалі використовуються для оцінки індивідуальних результатів психологічного тесту.

Дослідження стандартизації полягають у вивченні репрезентативних груп з метою використання отриманих результатів як орієнтирів порівняння індивідуальних показників за тестом. Це дозволяє з'ясувати відносно положення учасника дослідження щодо вибірки стандартизації, а також забезпечує порівнюваність результатів різних тестів [5, с. 64–65].

Незважаючи на те, що цей підхід є загальноприйнятим, деякі його аспекти вимагають більш детального розгляду. Насамперед, це стосується поняття “репрезентативність”. Так, П. Клайн говорить про необхідність використання стратифікованої вибірки, що дозволяє врахувати у дослідженні стандартизації найважливіші параметри, які впливають на результати використання тесту. Як приклад, він наводить своє дослідження, де учні шостих класів були стратифіковані за статтю, специфікою фінансування школи, її географічним розміщенням і статусом. Для кожної категорії було відібрано по дві школи. При цьому всі школи знаходилися на півночі Англії. У фінальну вибірку увійшло по 1000 хлопців і дівчат [6, с. 219]. Отже, йдеться не про статистичне узагальнення (застосовується випадкова вибірка необхідного обсягу), а про аналітичне (застосовується вивчення релевантних соціальних контекстів)².

Для тестів, які мають універсальну природу, дослідження стандартизації може бути здійснено і на підставі випадкової репрезентативної вибірки, хоча це не є принциповою умовою. Останнє пов'язано з тим, що головна мета такого дослідження – отримання опису властивості, що вивчається, в різних контекстах, а не опис генеральної сукупності в цілому. Тут доречно згадати зауваження Д. Кемпбелла про те, що “в науках, які успішно розвиваються, узагальнення ніколи ще не були “індуктивними” в сенсі підсумовування фактів, які спостерігалися в тих межах, в яких проводяться узагальнення, навпаки, вони завжди були гіпотетичними, навіть якщо і керувалися раніше, відкритими законами” [8, с. 207–208].

Наступне питання стосується перетворення або статистичної нормалізації сирих балів тесту, яке дозволяє ефективно вирішувати вже згадані завдання – порівняння з результатами вибірки стандартизації, а також результатами інших тестів. При цьому актуальним завданням в цьому випадку є саме порівняння результатів різних тестів. Що стосується порівняння з результатами вибірки стандартизації, то статистичні перетворення пов'язані з феноменом нормального розподілу. Отже, чим ближчим є емпіричний розподіл результатів використання тесту до нормального, тим менше сенсу в його статистичній нормалізації. Як зазначає А. Анастасі, при наближенні розподілу первинних показників до нормального стандартні показники, отримані за допомогою лінійного перетворення, практично не відрізняються. Відповідно, статистична нормалізація в такому випадку мало що або нічого не змінить. Необхідно домагатися нор-

² Докладніше про відмінність між статистичною та аналітичною генералізацією див.: [7, с. 69–70].

мального розподілу первинних результатів використання тесту, а не нормалізувати статистичними засобами ненормальний розподіл [5, с. 80].

Крім того, слід взяти до уваги визначення рівнів виразності досліджуваної властивості на підставі результатів використання певного тесту. Дослідження стандартизації саме по собі нічого не говорить про такі рівні. По суті, воно проводиться з іншими цілями. Єдине рішення цієї проблеми, яке зустрічається в методичній літературі (воно напрошується природним чином), полягає у використанні критеріальних норм замість статистичних. У цьому випадку шкала сирих тестових балів калібрується рівнями, які визначають імовірність досягнення деякого критерію або рівня, що дозволяють припустити якісно інший стан учасників дослідження за досліджуваною характеристикою [9, с. 153–154]. Водночас про те, як визначати критеріальні норми, не йдеться.

Соціологічне тестування: норми або критеріальні рівні? Методологічні напрацювання щодо норм у соціологічному тестуванні відсутні. Отже, єдиним орієнтиром тут є положення, що використовуються при розробці психологічних тестів.

З огляду на це, критично оцінимо три зазначених поняття, з точки зору специфіки соціологічного дослідження. Повертаючися до розмежування статистичної й аналітичної репрезентативності, ми вважаємо обґрунтованим використання останньої таким самим чином, як це рекомендується у психологічних дослідженнях. Це пов'язано з тим, що сучасні суспільства характеризуються досить високою гетерогенністю. Відповідно, важливо розуміти, які результати соціологічний тест покаже при використанні в різних соціальних контекстах (різних аудиторіях). При цьому апробація тесту в статистично репрезентативному опитуванні також є необхідним кроком його нормалізації. Отримані в такому опитуванні результати окреслюють загальні рамки порівняння та інтерпретації результатів інших опитувань.

Говорячи про нормалізацію результатів, слід визнати її обмеженість стосовно соціологічних досліджень. По-перше, в соціологічних опитуваннях акцент робиться на узагальненні отриманих результатів, а не на оцінці індивідуальних показників, щоб визначити їх місце щодо показників генеральної сукупності (таке завдання вирішується, насамперед, у тестах оцінки здібностей). По-друге, цілий спектр соціальних якостей на противагу психологічним не підпорядкований нормальному закону. І справа тут не в низькій якості вимірювального інструменту, а історичних особливостях певних спільнот. По-третє, на думку П. Клайна, коли йдеться про наукове дослідження, статистичні норми є не настільки важливими [6, с. 214]. Крім того, є і методичний аргумент на користь низької релевантності нормалізації результатів. З огляду на вимогу компактності соціологічних тестів, інтервали їх можливих значень часто будуть досить “вузькими” порівняно з багатьма психологічними тестами. А це вносить свої обмеження при побудові досить точних інтервалів з використанням статистичної нормалізації.

Що стосується критеріальних рівнів, то дане питання є одним з найважливіших при розробці соціологічного тесту. Це пов'язано з рішенням двох взаємопов'язаних завдань соціологічного тестування. Перше – продемонструвати, що тест володіє диференціальним потенціалом. Наприклад, якщо показники використання певного тесту авторитаризму в групі звичайних респондентів мало відрізняються від відповідних

характеристик крайніх анархістів і крайніх націоналістів, то критеріальних рівнів тест не має (як і не характеризується він критеріальною валідністю). Отже, для визначення критеріальних рівнів соціологічного тесту його автор повинен бути обізнаний про існування в суспільстві таких груп, які характеризуються низьким і (або) високим рівнем досліджуваної властивості, а також мати доступ до них. Це дозволить здійснити критеріальну нормалізацію тесту, що надасть можливість вирішення другого завдання – виокремлення рівнів у результатах тестування, що характеризують респондентів, з точки зору не тільки кількісних, але й якісних відмінностей.

Таким чином, поняття “норма” в соціологічному тестуванні недостатньо точно відображає суть відповідної методичної діяльності. Більш вдалим терміном у цьому випадку, на нашу думку, є поняття “критеріальний рівень”, розуміючи під ним числове значення в підсумкових результатах тесту, яке маркує якісно відмінний стан респондента за досліджуваною характеристикою. Відмітними властивостями критеріального рівня є: 1) партикулярність у сенсі прив’язки до специфічних, з точки зору досліджуваної властивості, соціальних груп; 2) якісність, яка зазначає рівні виразності вимірюваної властивості; 3) дискримінативність як розділовий потенціал тесту.

Кейс-стаді: критеріальні рівні SCL-9-NR. Для презентації викладених положень звернемося до визначення критеріальних рівнів методики SCL-9-NR, що являє собою соціологічний тест для вимірювання психологічного дистресу в масових опитуваннях. Його конструювання та валідизація описані в нещодавніх публікаціях [10; 11]. В останній з них зазначаються два критеріальних рівні: 14 балів (підвищений рівень) і 18 балів (високий рівень). На момент написання цієї публікації зібрано значно більше емпіричних даних, придатних для уточнення критеріальних рівнів SCL-9-NR. Отже, розглянемо відповідні результати.

Масиви даних. Уточнення критеріальних рівнів здійснювалося за результатами п’яти опитувань (табл. 1). Вибірка онлайн-досліджень будувалася на підставі учасників комерційних опитувань компанії Factum Group. Початкова мета онлайн-досліджень полягала не в оцінці рівня дистресу в Києві або Львові, а у побудові досить гетерогенної вибірки, що дозволяє проаналізувати специфіку відповідей на тест SCL-9-NR³. У випадку дослідження пацієнтів з психіатричним діагнозом у Чернігові та Харкові були опитані всі згодні взяти участь у дослідженні. Єдиним критерієм відбору був діагноз (розлади настрою або афективні розлади; невротичні, пов’язані зі стресом і соматоформні розлади; розлади особистості та поведінки в зрілому віці)⁴. Що стосується соціологічного моніторингу “Українське суспільство”, то його результати є репрезентативними щодо дорослого населення України (від 18 років і старше).

За даними наведених досліджень виокремлено шість груп респондентів, релевантних, з точки зору уточнення критеріальних рівнів SCL-9-NR. До першої увійшли респонденти онлайн-опитувань 2015 р., до другої – респонденти онлайн-опитування 2016 р., до третьої – респонденти з психіатричним діагнозом, до четвертої – група респондентів, які оцінюють стан свого здоров’я, як поганий або дуже поганий (соціоло-

³ Докладніше див.: [10].

⁴ Про причини фокусування саме на цих групах пацієнтів див.: [11].

гічний моніторинг), до п'ятої – група респондентів, які мешкали у зоні бойових дій на Донбасі (соціологічний моніторинг), до шостої – респонденти соціологічного моніторингу “Українське суспільство” крім тих, які проживали у зоні бойових дій. Останні були виключені з основного масиву, оскільки ця група з'явилась у нашому суспільстві нещодавно і за дуже короткий строк, а отже, є нехарактерною для його рівноважного стану, що спостерігався протягом років незалежності.

Рішення про поділ респондентів онлайн-досліджень на дві групи було прийнято тому, що рівень виразності психологічного дистресу в дослідженнях 2015 р. є дещо вищим, ніж у 2016 р. Про можливі причини такого розходження буде сказано нижче. Що стосується уваги до двох додаткових груп респондентів соціологічного моніторингу “Українське суспільство”, то вона зумовлена нетипово високим рівнем психологічного дистресу в них.

У зазначених шести групах респонденти додатково ділилися на дві підгрупи за статтю, оскільки у всіх дослідженнях чоловіки демонстрували більш низькі показники тесту. Розподіл відповідно до вікових груп не проводився, тому що при контролі самооцінки стану здоров'я взаємозв'язок між віком і індексом психологічного дистресу зникав.

Таблиця 1

**Дослідження, результати яких використані
для визначення критеріальних рівнів SCL-9-NR**

Дослідження	Збір даних	Метод відбору	Характеристики вибірки
Методичне дослідження у Львові, 2015 р. (Factum Group)	Онлайн-анкета для самозаповнення	Випадковий зі стратифікацією за віком і статтю	Чоловіки та жінки – по 50%, N = 200
Методичне дослідження у Києві, 2015 р. (Factum Group)	Онлайн-анкета для самозаповнення	Випадковий зі стратифікацією за віком і статтю	Чоловіки та жінки – по 50%, N = 200
Методичне дослідження у Києві, 2016 р. (Factum Group)	Онлайн-анкета для самозаповнення	Випадковий зі стратифікацією за віком і статтю	Чоловіки та жінки – по 50%, N = 205
Методичне опитування серед пацієнтів з психіатричним діагнозом, 2016–2017 рр. (ЧОПНЛ, ХОКПЛ)	Самозаповнення у присутності лікаря	Цілеспрямований зі стратифікацією за статтю	Чоловіки – 47%, Жінки – 53%, N = 217
Соціологічний моніторинг “Українське суспільство”, 2016 р. (Інститут соціології НАН України)	Самозаповнення	Випадковий зі стратифікацією за віком, статтю й освітою	Чоловіки – 45%, Жінки – 55%, N = 1802

Примітка: ЧОПНЛ – Чернігівська обласна психоневрологічна лікарня; ХОКПЛ – Харківська обласна клінічна психіатрична лікарня.

Джерело: складено автором.

Характер розподілу індексу SCL-9-NR. Оскільки психологічний дистрес розглядається як природна властивість особистості, його розподіл повинен відповідати нормальному закону. Відповідно, аналіз рівнів психологічного дистресу було розпочато з оцінки характеру розподілу індексу SCL-9-NR. Для цього використовувався критерій Шапіро-Уїлка, розрахований для підвбірок розміром 25 спостережень, відібраних по 1000 разів для кожної з шести зазначених груп респондентів ⁵ (табл. 2). За відповідними результатами, визначено кількість випадків, в яких приймається нульова гіпотеза ($n_p > 0,05$).

Отже, у всіх випадках, за винятком жінок, які проживали в зоні бойових дій, можна припустити нормальний розподіл індексу. Щодо єдиної групи, де цей висновок неприйнятний, то суттєве відхилення від нормальності пов'язане не з нечутливістю методики, а з нетипово важкою соціальною ситуацією, яка призводить до підвищеного психологічного дистресу.

Надалі з причини більшої або меншої міри асиметрії розподілів значень тесту як порогових значень критеріальних рівнів використовуватимуться не середні значення, а медіани.

Рівні виразності психологічного дистресу в різних групах респондентів. Перш ніж перейти до аналізу рівнів психологічного дистресу в різних групах, необхідно позначити ті спостереження, що були виключені з аналізу. До таких належать респонденти, які показали мінімально або максимально можливі значення SCL-9-NR: 0 або 27 балів. В обох випадках є припущення про нещирість відповідних спостережень, яке базується на двох аргументах.

Таблиця 2

Результати використання критерію Шапіро-Уїлка для 1000 підвбірок розміром 25 спостережень, $n_p > 0,05$, %

Група респондентів	Чоловіки	Жінки
Онлайн-дослідження у 2015 р. (далі – ОД2015)	82,7	84,0
Онлайн-дослідження у 2016 р. (далі – ОД2016)	83,1	81,2
ЧОПНБ, ХОКПБ	93,8	95,7
Соціологічний моніторинг: здоров'я (далі – СМ: 3)	93,3	90,4
Соціологічний моніторинг: зона бойових дій (далі – СМ: ЗБД)	94,0	35,8
Соціологічний моніторинг: без зони бойових дій (далі – СМ: без ЗБД)	86,6	93,1

Примітка: ЧОПНЛ – Чернігівська обласна психоневрологічна лікарня; ХОКПЛ – Харківська обласна клінічна психіатрична лікарня.

Джерело: складено автором.

⁵ Це зроблено через чутливість критерію до розміру вибірки. Гістограми самих розподілів не наводяться у статті, але доступні за: [12].

По-перше, контекст сучасного суспільства характеризується постійним впливом стресорів, що унеможлиблює відсутність відповідних ефектів у житті індивідів. По-друге, максимальна виразність за всіма симптоматичними вимірами не була виявлена серед респондентів з психіатричним діагнозом. З огляду на це, робиться припущення, що результат з максимально можливим значенням SCL-9-NR є настільки малоймовірною подією, що нею можна знехтувати.

У цілому показники психологічного дистресу для різних груп респондентів (табл. 3) відповідають теоретичним очікуванням. Черговість груп впорядкована відповідно до рівня виразності психологічного дистресу в них.

Єдиною групою, яка не потрапила в рамки початкових теоретичних очікувань, є респонденти онлайн-дослідження у 2016 р. Їх результати є значно ближчими до результатів соціологічного моніторингу “Українське суспільство”, ніж до результатів онлайн-дослідження у 2015 р. Спочатку передбачалося, що при онлайн-опитуванні респонденти будуть більш щирими і це спричинить отримання більш високих значень SCL-9-NR. Якщо розглядати частотні розподіли (рис. 1, 2), то це припущення частково підтвердилося. Проте відповідно до медіанного значення, спосіб самозаповнення (друкованої або онлайн-анкети) не пояснює відмінності між результатами онлайн-опитувань у 2015 р. і в моніторинговому дослідженні.

Найбільш логічним поясненням такої ситуації автор вбачає у впливі календарного часу опитування. У 2015 р. респонденти проходили опитування в кінці листопада і на початку грудня, а в 2016 р. – в травні (онлайн-дослідження) і червні (соціологічний моніторинг). Можливо, саме загальний емоційний стан в кінці року і в кінці весни – на початку літа пояснює ці відмінності.

В умовах відсутності інших релевантних даних результати цих трьох опитувань можна використовувати як орієнтири для визначення середнього рівня психологічного дистресу серед загального населення в Україні, виміряного за допомогою SCL-9-NR. Так, він коливається в діапазоні від 9 до 11 балів для чоловіків і від 10 до 12 балів для жінок. Водночас ця інформація сама по собі не говорить про межі підвищеного та високого рівнів психологічного дистресу.

Таблиця 3

**Показники SCL-9-NR для чоловіків
і жінок у різних дослідженнях**

Показник	СМ: без ЗБД		ОД2016		ОД2015		СМ: З		СМ: ЗБД		ЧОПНБ, ХОКПБ	
	Ч	Ж	Ч	Ж	Ч	Ж	Ч	Ж	Ч	Ж	Ч	Ж
1-й квар- тиль	5	7	7	8	9	9	9	9	9	10	9	13
Медіана	9	10	9	10	11	12	13	13	14	15	14	16
3-й квар- тиль	12	14	11	13	14	14	17	16	18	21	18	19

Примітка: ЧОПНЛ – Чернігівська обласна психоневрологічна лікарня; ХОКПЛ – Харківська обласна клінічна психіатрична лікарня.

Джерело: складено автором.

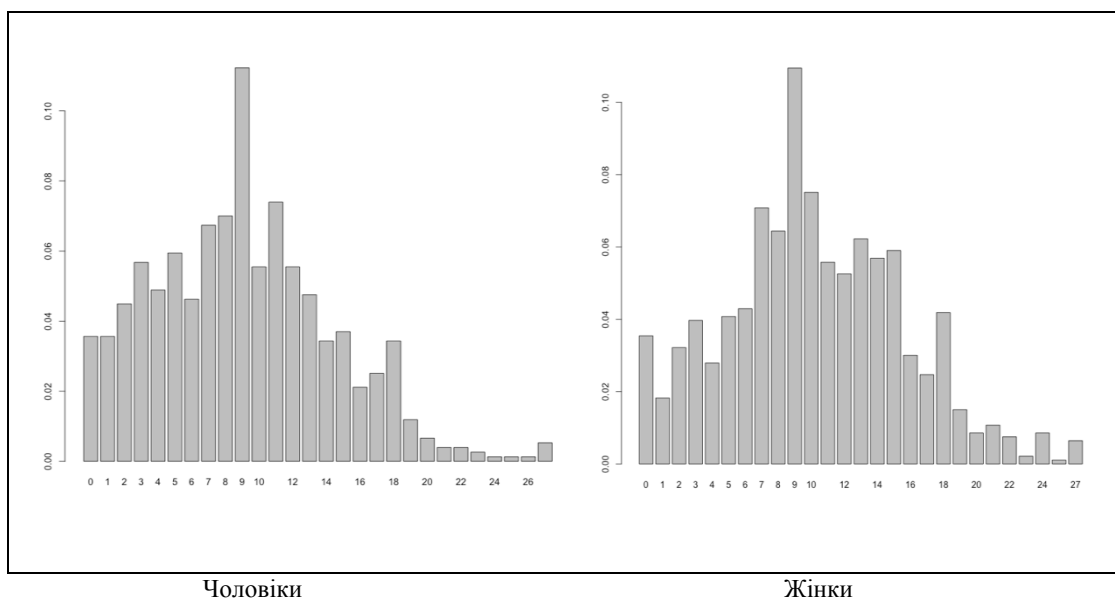


Рис. 1. Розподіл сумарного значення SCL-9-NR, за результатами соціологічного моніторингу “Українське суспільство” для чоловіків і жінок
Джерело: складено автором.

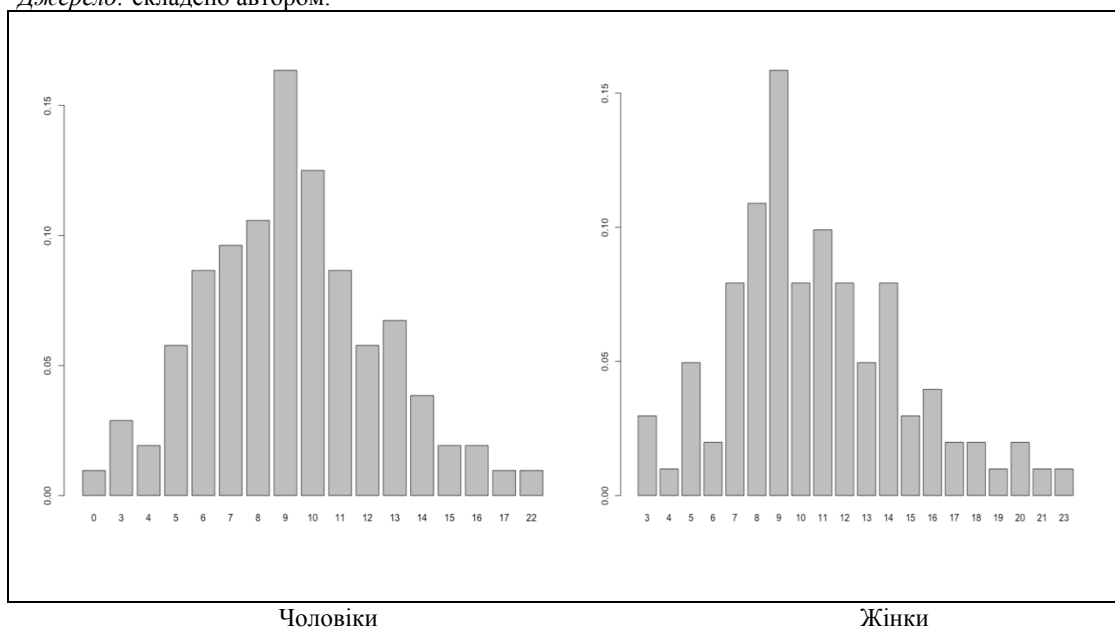


Рис. 2. Розподіл сумарного значення SCL-9-NR, за результатами онлайн-опитування 2016 р. для чоловіків і жінок
Джерело: складено автором.

Для їх визначення слід звернутися до інших трьох груп, що характеризуються наявністю факторів, які обтяжують психологічний стан індивіда. Відповідні результати можна використати шляхом аналізу значень медіан і третіх кватилів розподілів.

Медіанне значення для респондентів (як чоловіків, так і жінок), які оцінюють стан свого здоров'я як поганий або дуже поганий, дорівнює 13 балам. Для респондентів, які жили на території бойових дій або мають психіатричний діагноз, медіанне значення дорівнює, відповідно, 14 балів для чоловіків, а також 15 і 16 балів для жінок. Отже, можна припустити, що рівень підвищеного психологічного дистресу починається з позначки в 13 балів для респондентів обох статей.

Щодо межі високого рівня, то для чоловіків вона починається з 17 балів (третій кватиль "по здоров'ю"). У випадку жінок спостерігається досить великий розкид значень (16, 21 і 19 балів). При цьому медіана в групі з психіатричним діагнозом перетинається з третім кватилем у групі респондентів, які негативно оцінюють стан свого здоров'я (в обох випадках – по 16 балів). На цій підставі я пропоную використовувати 17 балів як межі високого рівня психологічного дистресу, вимірюваного за допомогою SCL-9-NR і серед жінок.

Дискримінативний потенціал SCL-9-NR. Для демонстрації розділового потенціалу методики використано два варіанти класифікації із застосуванням дерев рішень⁶. Перший варіант полягає у використанні дерев класифікації на базі значень підсумкового (сумарного) індексу SCL-9-NR, другий – на основі окремих індикаторів методики. В останньому випадку відповіді респондентів були перекодовані, відповідно до дихотомічної шкали: 1) "ніколи" та "зрідка"; 2) "періодично" та "майже постійно".

Як спостереження використовувалися дані, зібрані в ЧОПНБ і ХОКПБ, а також результати онлайн-опитування у 2016 р. Відповідно, респонденти були розділені на дві групи: з психіатричним діагнозом і без нього. Результати чоловіків і жінок аналізувалися окремо. Кожен з чотирьох підмасивів був випадковим чином розбитий на дві підгрупи, що дозволило сформувати навчальні та контрольні дані для використання в деревах класифікації:

- чисельність жінок з діагнозом становила 56 респондентів у контрольній групі та 57 у навчальній;
- чисельність жінок без діагнозу становила 50 респондентів у контрольній групі та 51 у навчальній;
- чисельність чоловіків з діагнозом становила по 51 респонденту в контрольній і навчальній групах;
- чисельність чоловіків без діагнозу становила по 52 респонденти в контрольній і навчальній групах.

У цілому якість класифікації є вищою для жінок (табл. 4). Сумарний індекс відчутно поступається за якістю класифікації окремим індикаторам, особливо якщо йдеться про контрольні дані. З огляду на те, що методика SCL-9-NR розроблялася не для виявлення психічних захворювань, а для вимірювання величини психологічного

⁶ Аналіз здійснено за допомогою бібліотеки "tree" мови програмування R. Докладніше про дерева класифікації та їх реалізацію в R див.: [13, с. 328–363].

дистресу, наведені результати свідчать про досить високий дискримінативний потенціал методики.

Таблиця 4

Кількість класифікаційних помилок, %

Група	Сумарний індекс SCL-9-NR	Окремі індикатори
Чоловіки: навчальні дані	23,5	23,5
Чоловіки: контрольні дані	33,0	24,3
Жінки: навчальні дані	20,1	15,0
Жінки: контрольні дані	31,1	22,6

Джерело: складено автором.

Слід також зазначити про те, що у випадках чоловіків і жінок як набір предикторів, так і структура дерев відрізняються (рис. 3). Так, у випадку чоловіків спрацьовують симптоматичні виміри обсессивно-компульсивного розладу, фобічної тривожності, соматизації та депресії, а у випадку жінок – психотизму, депресії, фобічної тривожності та соматизації. У рамках даного дослідження це не стало несподіваним. Проте, з точки зору загальноприйнятого підходу, який оцінює психологічний дистрес через депресію та тривожність, такі класифікаційні моделі мають, скоріше, спростовуючий характер.

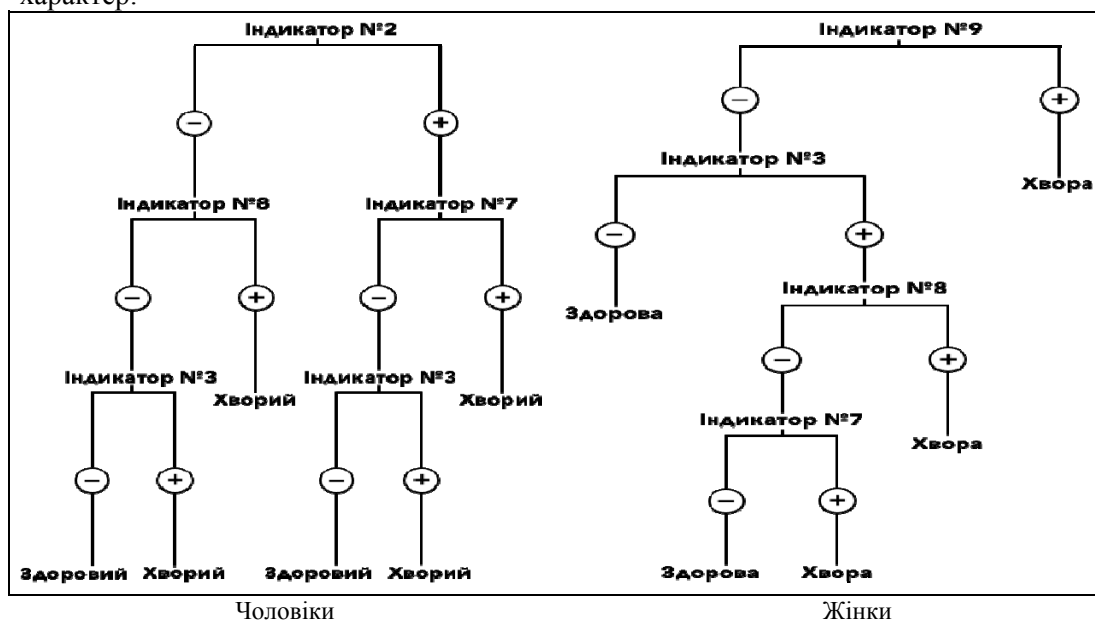


Рис. 3. Класифікаційні дерева на підставі окремих індикаторів SCL-9-NR для чоловіків і жінок

Джерело: складено автором.

Висновки. З огляду на особливості соціологічного тестування, слід говорити не про норми тесту, а про його критеріальні рівні, визначення яких пов'язане з необхідністю проведення спеціальних методичних досліджень, які відповідають таким вимогам:

– партикулярність: націленість на дослідження груп, які є нетиповими, з точки зору виразності властивості, вимірюваної соціологічним тестом (ступінь виразності може бути як більшим, так і меншим);

– якість: спрямованість не на позначення формально відмінних рівнів у результатах використання тесту, а на визначення тих меж, які свідчать про якісно відмінні стани індивідів;

– дискримінативність: забезпечення прийнятної класифікаційного потенціалу вимірювальної методики, з точки зору завдань, на вирішення яких вона спрямована.

У методичному сенсі визначення критеріальних норм передбачає вирішення таких завдань: ретельний відбір спостережень у рамках методичних досліджень; оцінка характеру розподілу підсумкових значень тесту, отриманих у різних методичних дослідженнях; зіставлення підсумкових показників різних методичних досліджень з метою визначення критеріальних рівнів тесту; оцінка дискримінативних особливостей тесту в різних групах.

Список використаних джерел

1. Аванесов В. Тесты в социологическом исследовании. Москва: Наука, 1982. 200 с.
2. Дембицкий С. Социологические тесты: сущность и валидизация. *Социология: теория, методы, маркетинг*. 2016. № 3. С. 140–155.
3. Дембицкий С. Измерительные шкалы VS интегральные индексы. URL: http://soc-research.info/blog/index_files/ms-and-ii.html
4. Kaplan R., Saccuzzo D. *Psychological Testing: Principles, Applications, and Issues*. Belmont: Wadsworth, 2009. 715 p.
5. Анастаси А., Урбина С. *Психологическое тестирование*. 7-е изд. Санкт-Петербург: Питер, 2007. 688 с.
6. Клайн П. *Справочное руководство по конструированию тестов: Введение в психометрическое проектирование*. Киев: ПАН ЛТД, 1994. 283 с.
7. Дембицкий С. Теоретическая валидизация в социологическом исследовании: Методология и методы. Москва: УРСС, 2016. 200 с.
8. Кэмпбелл Д. *Модели экспериментов в социальной психологии и прикладных исследованиях*. Москва: Прогресс, 1980. 391 с.
9. Шмелев А. *Основы психодиагностики*. Ростов-на-Дону: Феникс, 1996. 544 с.
10. Дембицкий С. Экспресс-тест SCL-9-NR: методика оценки выраженности психологического дистресса для массовых опросов. *Социология: теория, методы, маркетинг*. 2016. № 1. С. 52–64.
11. Дембицкий С., Яценко В., Сидоров Н., Александрова С. Критериальная валидизация экспресс-теста SCL-9-NR. *Социология: теория, методы, маркетинг*. 2016. № 2. С. 102–116.
12. Дембицкий С. Гистограммы распределений индекса SCL-9-NR. URL: http://soc-research.info/blog/index_files/scl-plots.html
13. Джеймс Г., Уиттон Д., Хасты Т., Тибширани Р. *Введение в статистическое обучение с примерами на языке R*. Москва: ДМК Пресс, 2016. 460 с.

Отримано 24.10.17

References

1. Avanesov, V. (1982). Tests in sociological research. Moscow: Nauka [in Russian]
2. Dembitskiy, S. (2016). Sociological tests: the essence and validation. *Sotsiologiya: teoriya, metody, marketing – Sociology: theory, methods, marketing*, 3, 140–155 [in Russian]
3. Dembitskiy, S. (2017). Measurement scales VS integral indices. URL: http://soc-research.info/blog/index_files/ms-and-ii.html [in Russian]
4. Kaplan, R., Saccuzzo, D. (2009). *Psychological Testing: Principles, Applications, and Issues*. Belmont: Wadsworth [in English]
5. Anastazi, A., Urbina, S. (2007). *Psychological testing*. Saint Petersburg: Piter [in Russian]
6. Klain, P. (1994). *A Handbook of Test Construction: Introduction to Psychometric Design*. Kiev: PAN LTD [in Russian]
7. Dembitskiy, S. (2016). Theoretical validation in sociological research: methodology and methods. Moscow: URSS [in Russian]
8. Campbell, D. (1980). *Models of experiments in social psychology and applied research*. Moscow: Progress [in Russian]
9. Shmelev, A. (1996). *Fundamentals of Psychodiagnostics*. Rostov-on-Don: Feniks [in Russian]
10. Dembitskiy, S. (2016). SCL-9-NR: a brief test for measurement of psychological distress in survey research. *Sotsiologiya: teoriya, metody, marketing – Sociology: theory, methods, marketing*, 1, 52–64 [in Russian]
11. Dembitskiy, S., Yashchenko, V., Sidorov, N., Aleksandrova, S. (2016). Validation criterion of brief test SCL-9-NR. *Sotsiologiya: teoriya, metody, marketing – Sociology: theory, methods, marketing*, 2, 102–116 [in Russian]
12. Dembitskiy, S. (2017). Histograms of SCL-9-NR index distributions. URL: http://soc-research.info/blog/index_files/scl-plots.html [in Russian]
13. James, G., Witten, D., Hastie, T., Tibshirani, R. (2016). *An Introduction to Statistical Learning with Applications in R*. Moscow: DMK Press [in Russian]

Received on 24.10.17

Дембицкий С.С., канд. социол. наук, старший научный сотрудник отдела методологии и методов социологии Института социологии НАН Украины, ул. Шелковичная, 12, Киев, 01021, Украина, e-mail: e_forge@me.com

ПРИНЦИПЫ ОПРЕДЕЛЕНИЯ КРИТЕРИАЛЬНЫХ УРОВНЕЙ СОЦИОЛОГИЧЕСКОГО ТЕСТА

Исследовано понятие “норма” в привязке к социологическому тестированию. На базе методологических размышлений предложена его замена на более подходящий термин – “критериальный уровень”, понимаемого в качестве числового значения социологического теста, которое обозначает качественно отличный уровень исследуемой характеристики. Свойствами критериальных уровней являются партикулярность, качественность и дискриминативность. Для демонстрации теоретических положений статьи приведены результаты определения критериальных уровней экспресс-теста SCL-9-NR.

Ключевые слова: социологическое тестирование, нормы, критериальные уровни, психологический дистресс, SCL-9-NR.

Dembitskiy S.S., PhD (Sociology), senior scientific fellow, department of methodology and methods of sociology, Institute of Sociology National Academy of Sciences of Ukraine, 12, Shovkovychna Str., Kyiv, 01021, Ukraine, e-mail: e_forge@me.com

DETERMINATION PRINCIPLES OF THE CRITERIAL LEVELS OF THE SOCIOLOGICAL TEST

The article evaluates the concept of the “norm” in its relation to sociological testing. Since there are no methodological works regarding norms in sociological testing, a starting point of the analysis is a study of standardization in psychological testing. For this matter the author considered the following components of standardization studies: achievement of representativeness, statistical normalization and finding of investigated property’s levels. Based on this analysis, paper presents general model of standardization studies of sociological tests, and proposes a substitution of the term “norm” for a more appropriate one – a “criterial level”. It is understood as a numerical value of the sociological test, which denotes a qualitatively different level of a studied characteristic. The properties of the criterial levels are specificity, quantitateness and discriminativeness. To demonstrate the theoretical principals of the article, the work presents results of criterial levels exploring of the brief-test SCL-9-NR. The corresponding analysis is based on data from five studies: 1) online survey in Lviv (2015, N = 200); 2) online survey in Kyiv (2015, N = 200); 3) online survey in Kyiv (2016, N = 205); 4) survey of patients with psychiatric disorders in Chernihiv and Kharkiv (2016-2017, N = 217); 5) sociological monitoring “Ukrainian society” (2016, N = 1802). The exploring of the criterial levels of SCL-9-NR involved three stages: analysis of the SCL-9-NR distribution properties (Shapiro-Wilk test); analysis of psychological distress levels in various groups of respondents and determination of the criterial levels on this basis (median analysis); assessment of the discriminatory potential of SCL-9-NR (classification using decision trees). According to the results, it is determined that an increased level of distress with the use of SCL-9-NR starts at 13 points, and the high level - at 17 points.

Keywords: sociological testing, norms, criterial levels, psychological distress, SCL-9-NR.