



НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
ІНСТИТУТ СОЦІОЛОГІЇ

ДЕМБІЦЬКИЙ СЕРГІЙ

**РОЗРОБКА
СОЦІОЛОГІЧНИХ ТЕСТІВ:
МЕТОДОЛОГІЯ І ПРАКТИКИ
ЇЇ ЗАСТОСУВАННЯ**

Київ 2019

УДК 303.2
Д 30

Рекомендовано до друку Вченою радою Інституту соціології НАН України
(протокол № 5 від 23 жовтня 2018 року)

Рецензенти:

Афонін Е., доктор соціологічних наук, Національна академія державного управління при Президентіві України

Бевзенко Л., доктор соціологічних наук, Інститут соціології НАН України

Паніотто В., доктор філософських наук, Національний університет «Києво-Могилянська академія»

Дембіцький С.

Д 30 Розробка соціологічних тестів: методологія і практики її застосування : монографія. — Київ : Інститут соціології НАН України, 2019. — 304 с.

ISBN 978-966-02-8792-1

У монографії здійснено концептуальне й методологічне обґрунтування особливостей розробки соціологічних тестів. Для цього автор розглянув соціологічні тести як інструменти соціологічного вимірювання, критично оцінив теорію вимірювання в соціальних науках, запропонував альтернативний підхід до вимірювання в соціології, узагальнив та описав вітчизняний досвід розробки соціологічних тестів, проаналізував досвід розробки комплексних вимірювальних інструментів у дослідженнях змішаного типу, оцінив кількісні дослідження в контексті розробки соціологічних тестів і на цій основі виокремив релевантні процедури й методи, визначив практичні засоби розробки соціологічних тестів.

Книга адресована спеціалістам з емпіричної соціології, політологічних, маркетингових, епідеміологічних і психологічних досліджень, а також усім тим, кого цікавлять проблеми вимірювання в соціальних науках.

УДК 303.2

ISBN 978-966-02-8792-1

© Дембіцький С., 2019

© Інститут соціології НАН України, 2019

ЗМІСТ

Вступ	7
Розділ I. СОЦІОЛОГІЧНІ ТЕСТИ ЯК ЗАСІБ РОЗВ'ЯЗАННЯ ПРОБЛЕМ СОЦІОЛОГІЧНОГО ВИМІРЮВАННЯ	13
<i>Глава 1.</i> Різновиди емпіричної соціології	14
<i>Глава 2.</i> Теоретична валідизація вимірювальних шкал	18
<i>Глава 3.</i> Соціологічний тест як вимірювальний інструмент у соціальних науках	31
<i>Глава 4.</i> Теорія соціального вимірювання: припущення, латентні змінні та соціальні феномени	38
<i>Глава 5.</i> Процес розробки соціологічних тестів	43
<i>Глава 6.</i> Приклади соціологічних тестів: соціальне самопочуття, політична культура, психологічний дистрес	53
Висновки до першого розділу	58
Розділ II. ІНДИВІДУАЛЬНЕ БЛАГОПОЛУЧЧЯ ЯК ОБ'ЄКТ СОЦІОЛОГІЧНОГО ТЕСТУВАННЯ	59
<i>Глава 7.</i> До розуміння феномену індивідуального благополуччя	60
<i>Глава 8.</i> Конструювання і валідизація соціологічного тесту SCL-9-NR.	63
<i>Глава 9.</i> Застосування L-шкали за опитування населення України ...	92
<i>Глава 10.</i> Інтегральний індекс соціального самопочуття: альтернативний підхід	108
<i>Глава 11.</i> Вимірювання індивідуального благополуччя в Європейському соціальному дослідженні: концептуальні та методологічні аспекти.	127
Висновки до другого розділу	142

Розділ III. ПОЛІТИЧНА КУЛЬТУРА У ФОКУСІ СОЦІОЛОГІЧНОГО ТЕСТУВАННЯ	143
<i>Глава 12.</i> Соціологічний тест «Типи політичної культури»: вплив способу кластеризації на результати типології.	145
<i>Глава 13.</i> Валідизація шкали громадської активності	159
<i>Глава 14.</i> Конструювання, валідизація та інтерпретація соціологічного тесту «Типи політичної культури – II».	174
<i>Глава 15.</i> Незалежна Україна в геополітичному дискурсі: віхи і акценти.	190
<i>Глава 16.</i> Вимірювання геополітичних орієнтацій в Україні: формулювання індикаторів і апробація в методичному опитуванні	209
Висновки до третього розділу	229
Розділ IV. ПРИНЦИПИ І ПРОЦЕДУРИ РОЗРОБКИ СОЦІОЛОГІЧНИХ ТЕСТІВ	231
<i>Глава 17.</i> Розробка комплексних вимірювальних інструментів у дослідженнях змішаного типу	232
<i>Глава 18.</i> Розробка соціологічних тестів у контексті кількісних досліджень	249
<i>Глава 19.</i> Рекомендації до розробки соціологічних тестів.	266
Висновки до четвертого розділу	274
Висновки	275
Додатки	279
Список використаних джерел	290

ВСТУП

Вимірювання особистісних властивостей пов'язане з появою психометрики, яка є втіленням кількісної психології. На сучасному етапі психометрика включає набір статистичних моделей і методів, які були розроблені для підсумування, опису і побудови висновків на базі емпіричних даних, зібраних у психологічних дослідженнях. Відповідні статистичні моделі та методи можуть бути поділені на три класи. До першого входить психологічне шкалювання. Воно являє собою набір технік присвоювання кількісних значень об'єктам або подіям із використанням даних, отриманих за допомоги суджень людей. Другий клас містить великий набір методів і процедур, виведених із базової ідеї факторного аналізу — пояснення коваріації, що спостерігається між набором змінних, шляхом звернення до варіації латентної змінної, яка може пояснювати таку коваріацію. Третій клас моделей має стосунок до теорії тестів, що найповніше представлена математичною теорією вимірювання (*item response theory*). Остання являє собою клас статистичних моделей, які поєднують базові компоненти психологічного шкалювання і використання пояснювальних латентних змінних з традиції факторного аналізу. Частиною соціології, що зазнала найпотужнішого впливу психометрики, є кількісна соціологія, в межах якої суттєвий внесок був зроблений у моделі структурних рівнянь, і латентний класовий аналіз (Jones & Thissen, 2007). До найвідоміших представників психометрики належать К. Гаусс, В. Вундт, Ф. Гальтон, Ч. Спірмен, А. Біне, Р. Кеттел, Л. Терстоун, К. Кумбс та ін. Перелік науковців, що мають стосунок до розробки шкал вимірювання, є вельми широким, тому згадаємо лише ті персоналії, напрацювання яких ми використали у пропонованому дослідженні: Л. Дерогатіс (SCL-90-R), Р. Кесслер (К10 та К6), С. Хатгей, Д. Маккінлі (ММРІ), Л. Редлофф (CES-D).

Отже, найвпливовіша вимірювальна традиція в науках про людину як представника соціального світу аналітично може бути поділена на дві великі частини. Перша пов'язана зі статистичним аналізом, друга — з розробкою вимірювальних інструментів. У межах соціології доробок було здійснено в тій частині, яка стосується статистичного аналізу. Щодо розробки вимірювальних інструментів соціологічна традиція базується на відповідних положеннях психометрики (якщо йдеться про комплексні вимірювальні інструменти¹), а також на принципах формулювання анкетних питань для масових опитувань². Стосовно останнього випадку говорити слід про міждисциплінарну сферу, в якій важливе місце посідає і соціологія.

¹ Див., напр.: (DeVellis, 2012).

² Див., напр.: (Садмен, Брэдберн, 2002).

Такі історичні особливості емпіричної соціології потягли за собою методологічні суперечності, які характерні для неї і зараз. Сутність цих суперечностей полягає в тому, що соціологічні шкали та індекси розробляють з акцентом на вимірюванні латентної змінної (яка, по суті, є тією чи тією особистісною властивістю), натомість у фокусі соціологічної уваги в більшості випадків перебуває не індивідуальний рівень, а соціальний — не характеристики окремої особистості, а соціальні феномени.

Конвенційно прийнятим підходом до розв'язання цієї суперечності є агрегація даних масових опитувань. Вона може набувати форми елементарної описової статистики або ж, навпаки, здійснюватися з використанням найновіших її методів. Утім, цей загальний принцип лише відволікає увагу від зазначеної суперечності — використання психологічних інструментів для опису суспільних явищ. Кількісна психологія також ґрунтується на узагальненні даних із використанням тих самих статистичних процедур, що й у соціології. Але при цьому ніхто не говорить про дослідження соціальних феноменів. Таким чином, одним із проблемних орієнтирів є необхідність формулювання власне соціологічного підходу до розробки комплексних вимірювальних інструментів, які в тому числі матимуть потенціал і для дослідження соціальних феноменів.

Проблеми розробки соціологічних інструментів вимірювання вітчизняні соціологи вивчають уже понад 20 років. Наталія Паніна, наприклад, виокремила три головні причини скептичного ставлення до вимірювання в соціології (Паніна, 1998):

- 1) соціально-політичні — «пульсуючий» характер інтересу громадсько-політичного істеблїшменту до соціології;
- 2) професійно-соціальні — недостатня кількість соціальних показників у соціології, які мають безпосередній інтерес за будь-яких соціальних умов;
- 3) професійно-методичні — легковажне ставлення соціологів до методичного інструментарію збору даних і подальшого їх аналізу.

Головним методологічним засобом розв'язання цих проблем вона визначила соціологічні тести, які трактувала як спеціальні соціологічні методики вимірювання, призначені для використання в масових опитуваннях. Зробивши порівняльний аналіз застосування тестів (як вимірювальних методик) у психологічних і соціологічних дослідженнях, Н. Паніна запропонувала основні принципи конструювання соціологічних тестів, або адаптації психологічних тестів до особливостей масових опитувань: універсальність, інтегральність, якісність, стандартизованість, «чутливість», компактність і об'єктивність.

Дослідниця не обмежилася суто теоретико-методологічними напрацюваннями щодо соціологічного тестування. Разом з Євгеном Головахою вони розробили декілька соціологічних тестів, із яких найвідомішими є Інтегральний індекс соціального самопочуття, Індекс дестабілізаційного потенціалу та Соціологічний тест «Типи політичної культури», а також адаптували низку відомих психологічних тестів до використання в масових опитуваннях (Шкала аномії, Шкала цинізму, Шкала національної дистанції тощо). Варто відзначити й інших науковців, академічні напрацювання яких є релевантними з точки зору теми соціо-

логічного вимірювання, — С. Садмена, Н. Бредберна, Н. Шварца, Д. Рогозіна (застосування когнітивного аналізу в масових опитуваннях), В. Ядова, В. Паніотто (оцінка якості соціологічного дослідження), Р. ДеВелліса (основи вимірювання в соціальних науках), В. Аванесова (соціологічне тестування), А. Анастасі, П. Клайна, А. Шмельова (психологічне тестування).

Утім, незважаючи на великий обсяг проведеної роботи, багато методологічних і методичних проблем соціологічного вимірювання так і не отримали наукового розв'язання. З огляду на це ми фіксуємо **проблемну ситуацію**, яка полягає у відсутності в соціології власної концепції вимірювання, що тягне за собою низку особливостей, серед яких — психологічна мімікрія методології соціологічного вимірювання, недостатність відповідного термінологічного арсеналу, а також обмеженість шляхів розвитку. Розглянемо їх докладніше.

Психологічна мімікрія полягає в некритичному прийнятті всіх принципів розробки вимірювальних інструментів і відповідної термінології, що запропоновані у психометрії. Це призвело до появи відповідного поля методологічної активності під назвою «соціальне вимірювання», що включає багато дисциплінарних напрямків, серед яких — соціологія, соціальна психологія, психологія особистості, політологія, освітні дослідження тощо. Але така міждисциплінарність має для соціального вимірювання уявний характер, оскільки домінують становище в ньому мають вимірювальні стандарти психометрії.

Це цілком очікувано веде до звуження термінологічного арсеналу емпіричної соціології, оскільки орієнтація йде не на власні методологічні проблеми, а на методологічні проблеми суміжної дисципліни. Те, що не потрапило до поля інтересів психологічного вимірювання, але має безпосередній стосунок до соціологічного (особливості розробки методик вимірювання, об'єктом яких є соціальні феномени, адекватна класифікація таких методик, перехід від рівня індивідуальних відповідей до рівня аналізу інституційних особливостей тощо), не отримує необхідної концептуалізації, а розв'язання відповідних проблем відбувається інтуїтивно або взагалі не здійснюється.

Така ситуація створює контекст, у якому повноцінний розвиток соціологічного вимірювання взагалі є неможливим — складно аргументовано сформулювати релевантну методологічну проблему. Але навіть за наявності певної проблеми, у дослідника відсутні концептуальні засоби її розв'язання.

Досвід розробки соціологічних тестів у вітчизняній соціології частково згладжує цю проблему. Утім відповідні наукові результати є початковим фундаментом для подальшої дослідницької діяльності, а не відповіддю на перелічені особливості проблемної ситуації. Через це виникає **проблема** невідповідності практики розробки комплексних вимірювальних інструментів у соціології та тих методологічних основ, які для цього використовують на сучасному етапі. При цьому під комплексними вимірювальними інструментами ми розуміємо шкали, індекси і класифікатори, що задовольняють вимоги використання в масових опитуваннях. Також у подальшому з метою універсалізації пропонування методологічних положень і врахування найважливіших напрацювань із цієї проблеми (Н. Паніна, Є. Головаха) розглядаємо соціологічні шкали, індекси і класифікатори як різновиди соціологічних тестів.

Отже, **метою** пропонованого дослідження є концептуальне і методологічне обґрунтування особливостей розробки соціологічних тестів.

Досягнення цієї мети передбачає розв'язання таких дослідницьких **завдань**:

- а) описати соціологічний тест як інструмент вимірювання в соціології;
- б) критично розглянути теорію вимірювання, яку використовують у соціальних науках;
- в) запропонувати альтернативний підхід до вимірювання в соціології;
- г) накопичити й описати вітчизняний досвід розробки соціологічних тестів;
- д) узагальнити досвід розробки комплексних вимірювальних інструментів у дослідженнях змішаного типу;
- е) оцінити кількісні дослідження в контексті розробки соціологічних тестів і на цій основі виокремити релевантні процедури і техніки;
- ж) визначити практичні засоби розробки соціологічних тестів.

Об'єктом дослідження є методологія вимірювання в соціології, а **предметом** дослідження — розробка соціологічних тестів.

Окреслена проблема, мета і завдання, об'єкт і предмет дослідження зумовлюють відповідну структуру дослідження.

Розділ I присвячений побудові альтернативного підходу до вимірювання в соціології. Його зміст полягає в необхідності прояснення: 1) термінології, яка враховує соціологічну специфіку вимірювання, різноманіття відповідних вимірювальних інструментів і об'єктів, що потрапляють у фокус соціологічного тестування; 2) аргументів, які обґрунтовують перехід від індивідуального рівня на рівень вимірювання соціальних феноменів; 3) процесу розробки соціологічних тестів. Останні є ключовою категорією як першого розділу, так і всієї роботи. З огляду на це окрему увагу буде приділено трьом вимірам, у яких соціологічне тестування розкривається з усією повнотою — відповідність вимогам використання в масових опитуваннях, спрямованість на дослідження соціальних феноменів і діагностичний потенціал з точки зору дослідження соціальної системи. Такий тривимірний погляд на соціологічне тестування буде застосовано для оцінки існуючих соціологічних тестів (Інтегрального індексу соціального самопочуття, соціологічного тесту «Типи політичної культури», Індекс психологічного дистресу SCL-9-NR).

У наступних двох розділах викладено накопичений досвід розробки соціологічних тестів, що є важливим для переходу від загальних методологічних положень до конкретних рекомендацій. У **розділі II** увагу зосереджено на розробці й аналізі соціологічних тестів, що вимірюють окремі аспекти індивідуального благополуччя. Першим із них є психологічний дистрес, вимірювання якого має усталену традицію і, відповідно, велику кількість запропонованих методик. При цьому варіанти шкал, придатних для застосування в масових опитуваннях, мають психологічний характер. З огляду на це спробуємо розробити саме соціологічний тест психологічного дистресу. Вихідною базою цього буде відомий психологічний тест SCL-90-R Л. Дерогатіса. Розглянемо ще один аспект індивідуального благополуччя, а саме аспект соціального самопочуття. Враховуючи відповідний соціологічний тест (Інтегральний індекс соціального

самопочуття), акцентуємо увагу на конструюванні альтернативного підходу до роботи з ним. Це зумовлене великим методичним потенціалом указаної методики, який був закладений у процесі її розробки. Приділимо також увагу оцінці методик вимірювання особистісного і соціального благополуччя, які було використано в Європейському соціальному дослідженні.

Розділ III присвячений особливостям вимірювання такого соціального феномену, як політична культура. Аналіз розпочнемо з оцінки результатів використання соціологічного тесту «Типи політичної культури», призначеного для вимірювання установок на демократію і громадську активність в українському суспільстві. Крім цього, буде зроблена спроба розвитку методологічних ідей, використаних у побудові цього тесту. У разі успіху відповідних пошуків буде запропоновано новий соціологічний тест — «Типи політичної культури — II», що змістовно виходить за межі установок виключно на демократію і громадську активність. Із метою комплексного охоплення феномену політичної культури увагу зосереджено на вимірюванні геополітичних орієнтацій населення України. Це потребує розгляду незалежної України в геополітичному дискурсі з подальшим переходом до формулювання індикаторів і розробки на їх основі соціологічних тестів.

Розділі IV є завершальною частиною монографії, в якій підсумовано досвід розробки соціологічних тестів, подано принципи і процедури їх розробки. У цьому контексті акцентовано на розробці комплексних вимірювальних інструментів у дослідженнях змішаного типу (порівняльний аналіз існуючих підходів), а також розробці соціологічних тестів у межах кількісних досліджень (опис процедур і методів, що пройшли концептуальну і методичну перевірку). В останній частині розділу запропоновано рекомендації до розробки соціологічних тестів.

Теоретико-методологічним підґрунтям дослідження є традиція розвідок змішаного типу (mixed methods research), а також елементи кількісної парадигми, що відповідають критеріям підходу до соціологічного вимірювання, запропонованого в цій монографії. При цьому дослідження змішаного типу розглянуто з позицій перспективи практики (Creswell & Tashakkori, 2007). Відповідно до неї необхідність звернення до змішаного підходу може проявитися на будь-якому етапі дослідження як відповідь на розв'язання завдань. У такій ситуації дослідник удається до використання кількісних і якісних елементів у межах його звичного дослідницького дизайну (неважливо — кількісного чи якісного), якщо є необхідність і можливість їх інтеграції в поточний дослідницький проект. Це, по суті, прагматична позиція. Не менш важливим підґрунтям дослідження є власні напрацювання автора, що стосуються теоретичної валідації вимірювальних шкал, які можуть бути використані як вихідна точка методологічних пошуків (Дембицкий, 2012).

Методи дослідження зумовлені характером поставлених завдань. Опис соціологічних тестів як інструментів вимірювання в соціології, розгляд існуючої теорії соціального вимірювання, а також формулювання альтернативного підходу до нього передбачають використання таких загальнонаукових методів, як аналіз, порівняння і синтез. Саме вони використані для узагальнення досвіду

розробки вимірювальних інструментів у дослідженнях змішаного типу, оцінки кількісних досліджень у контексті розробки соціологічних тестів і формулювання рекомендацій їх розробки.

Завдання, пов'язані з накопиченням і узагальненням практичного досвіду розробки соціологічних тестів, потребують використання соціологічних і математичних методів. До перших, зокрема, належать вибіркового метод, анкетування, кейс-наративний аналіз, експертне опитування, до других — методи описової статистики, графічні методи, різноманітні критерії перевірки статистичних гіпотез, багатовимірні методи як пошукового, так і верифікувального характеру. Для розширення можливостей оперування кількісними даними їх обробка здійснюватиметься за допомоги мови програмування R.

Емпірична частина роботи полягає в аналізі кейсів, що подані результатами використання комплексних вимірювальних методик у масових опитуваннях. Кейси виокремлено на базі результатів таких загальнонаціональних опитувань в Україні, як соціологічний моніторинг «Українське суспільство» (Інститут соціології НАН України) і Європейське соціальне дослідження (результати і для України, і для інших країн-учасниць цього дослідження). Крім цього, буде проведено окремі методичні дослідження, якщо соціологічні тести розроблені «з нуля».

Пропонована монографія узагальнює дослідження автора за останні роки, що були викладені в понад 20 публікаціях у фахових виданнях в Україні та за кордоном. Вона також містить результати, отримані під час участі автора у виконанні планових тем відділу методології та методів соціологічних досліджень Інституту соціології НАН України: «Розробка й адаптація інструментів вимірювання для масових соціологічних опитувань в умовах суспільства, що трансформується» (2014–2017), а також «Розробка й апробація соціологічного інструментарію дослідження процесів адаптації, інклюзії та зміни ідентичностей в умовах масових міграцій» (2017–2018).

Оскільки монографія має методологічну спрямованість, вона адресована перш за все спеціалістам з емпіричної соціології, політологічних, маркетингових, епідеміологічних і психологічних досліджень, а також усім тим, кого цікавлять проблеми вимірювання в соціальних науках. Звернення до особливостей вимірювання конкретних соціальних феноменів (соціальне самопочуття, політична культура) і психологічних особливостей індивідів (психологічний дистрес, особистісне благополуччя) розширює її потенційну аудиторію за рахунок широкого кола фахівців із соціології, політології та психології. Положення, викладені в монографії, будуть цікавими також для викладачів ВНЗ, аспірантів і студентів.

Автор не претендує на те, щоб ця монографія поставила фінальну крапку у проблемах соціологічного вимірювання. Правильніше сказати, вона є продовженням роботи з відповідної проблематики, яку розпочали відомі вітчизняні соціологи — Н. Паніна і Є. Головаха. Ми намагатимемося не лише ретранслювати їхні ідеї, але і творчо розвинути на двох рівнях. Перший стосується концептуальних і методологічних особливостей соціологічного вимірювання, другий — конкретних соціологічних тестів.

Розділ I

СОЦІОЛОГІЧНІ ТЕСТИ ЯК ЗАСІБ РОЗВ'ЯЗАННЯ ПРОБЛЕМ СОЦІОЛОГІЧНОГО ВИМІРЮВАННЯ

Глава 1

Різновиди емпіричної соціології

Історію зародження емпіричної соціології пов'язують із соціальною фізикою Адольфа Кетле, монографічним методом Фредеріка Ле Пле, Емілем Дюркгеймом і його емпіричним дослідженням феномену самогубства, Марселем Моссом і його вивченням феномену дару шляхом аналізу етнографічних описів (Зборовский, 2007; Лапин, 2004). Як видно, французи, якщо не брати до уваги одного бельгійця, не тільки плідно попрацювали над створенням соціологічної науки, а й задали основні напрямки соціологічних досліджень. Не варто, звичайно, забувати і про психометричну традицію, що веде свій початок від Френсіса Гальтона (Jones & Thissen, 2007).

Ця і низка інших дослідницьких традицій сформували те, що зараз прийнято називати емпіричною соціологією. І оскільки ця сфера наукової та практичної діяльності добре відома більшості соціологів, які мають досвід емпіричних досліджень на тому або тому рівні відповідної практики, ми пропонуємо додати до її назви слово «традиційна», тобто традиційна емпірична соціологія. Логічно, що може бути й «інша» емпірична соціологія, а можливо, і не одна. Тому спробуємо розглянути різновиди емпіричної соціології з тим, щоб правильно визначити місце отриманих нами результатів в історико-методологічному просторі емпіричної соціології.

§ 1.1. «Інша» емпірична соціологія

Щоб відшукати і правильно іменувати «іншу» емпіричну соціологію, слід звернутися до емпіричного матеріалу, на підставі якого відбувалося становлення соціології як науки. Із праць соціологічних першовідкривачів — Анрі Сен-Симона, Огюста Конта та інших можна довідатися, що таким матеріалом була людська історія. За їхньої епохи головна мета могла бути лише амбітною — науково витлумачити людську історію і сформулювати прогноз розвитку суспільств. Своєю чергою, розрив із метафізикою і звернення до позитивної філософії був гарантією того, що соціологія від самого початку — емпірична наука. Якщо проаналізувати праці інших представників класичної епохи соціології, можна виявити, що більшість із них тією чи тією мірою апелювали до історичного матеріалу (Барт, 1902). На сучасному етапі результати роботи перших соціологів «застигли» у вигляді соціологічної класики, яка сама по собі стала предметом історичного вивчення, але вже не в широкому сенсі.

Отже, початкова емпірична соціологія спиралася саме на вивчення історичного матеріалу. Виходячи з цього, пропонуємо назвати її класичною емпіричною соціологією.

Звичайно ж, у класичній соціології поряд з історичним компонентом був присутній і спекулятивний. При цьому їх співвідношення у працях різних класиків істотно варіюється. Наприклад, праці, які використовували метафору організму, є вищою мірою спекулятивними, натомість у прихильників економічного детермінізму історична складова посідає принципово важливу позицію.

Така історико-спекулятивна дуальність є генетичною для сучасної соціології. І йдеться тут не тільки про такі очевидно спекулятивні побудови, як, наприклад, теорії соціальних систем Толкота Парсонса або Нікласа Лумана, а й про ті, що належать перу загальноновизнаних історичних макросоціологів. Наприклад, Еммануїл Валлерстайн знайшов натхнення в марксистській соціології, а Рендалл Коллінз — у соціології Макса Вебера. І лише Чарльз Тіллі шукав музу в архівах.

Праці Валлерстайна, Коллінза і Тіллі, що далеко не вичерпують увесь перелік відомих імен сучасної історичної макросоціології, дозволяють зробити висновок, що класична емпірична соціологія отримала свій розвиток і оформилася у відповідну неокласику. І якщо вірити її головному літописцю, то саме зараз маємо спостерігати її золотий вік (Коллінз, 2015). Докладніше з періодизацією розвитку історичної соціології можна ознайомитися у праці Ірини Савельєвої (Савельєва, 2014).

У цьому контексті цікавим є питання про те, коли ж класична емпірична соціологія йде з арени соціологічних пошуків. Думаю, цей рубіж пов'язаний із творчістю Питирима Сорокіна, яка не тільки стала кінцем класики та її кульмінацією, а й останньою спробою захисту (Sorokin, 1961).

Тепер, відшукавши і найменувавши «іншу» емпіричну соціологію», треба повернутися до її звичного, «традиційного» різновиду і зрозуміти, як саме вони пов'язані між собою.

§ 1.2. Традиція vs Неокласика

Але чи можна говорити, що традиційний і неокласичний різновиди емпіричної соціології мають бути взаємопов'язаними? Вважаємо, так, і ось чому. Традиційна емпірична соціологія експансивна за своїм характером і в основі такої її якості лежить ринкова в широкому сенсі орієнтація. Якісний і кількісний підходи з 1930-х років, а також змішаний з початку 1990-х розвивалися переважно у вигляді відповідей на економічні, політичні та наукові запити різних соціальних груп. Завдяки цьому з'явилася безліч дослідницьких традицій, а також методів і технік. Але навіть незважаючи на очікуване в цьому контексті зростання дослідницького потенціалу соціальних наук, їх ключовими проблемами, як і раніше, залишаються ставлення релевантних наукових питань і наявність надійних засад інтерпретації отриманих результатів. Заковика тут у тому, що розв'язання останніх лежить за межами традиційної емпіричної со-

ціології. Якщо у маркетологів і політтехнологів, які хоча і цікавляться такими основними соціальними явищами, як влада і багатство, але на дуже конкретному рівні й у короткостроковій перспективі, розв'язання згаданих проблем не викликає особливих труднощів, то більшість тих, хто претендує на виробництво академічного знання, хронічно страждає від різного роду ідеологічних недуг, що позначаються як на сприйнятті повсякденності, так і на їхній історичній пам'яті. Відповідно, дані, зібрані в рамках традиційної емпіричної соціології, дуже часто стають заручниками позанаукових і навколонукових дискурсів, замість того, щоб використовуватися за призначенням.

Поставити по-справжньому актуальні дослідницькі питання, а також наділити релевантним значенням емпіричні факти про суспільство можна лише в рамках валідних уявлень про те, що відбувається навколо. І, на наше глибоке переконання, саме історична, або неокласична емпірична соціологія може запропонувати науково обґрунтований нейтральний базис для розв'язання цих теоретико-методологічних проблем.

Очевидним орієнтиром тут є те, що ми живемо у світі національних держав, об'єднаних у єдину глобальну систему. Це той контекст, який забарвлює наше і будь-яке інше сучасне суспільство. І щоб його зрозуміти, треба вникнути в історію становлення цієї світової системи, а також визначити наше місце в ній. Звичайно ж, звернення до історичних досліджень є першим і принципово важливим кроком на цьому шляху.

Отже, розуміти суспільство й отримані про нього дослідницькі дані можна і потрібно в контексті історичної дійсності, під якою ми розуміємо людські дії на різних рівнях соціальної взаємодії, спрямовані на створення і підтримку соціальних інститутів. Разом з тим не слід забувати і про те, що «паливом» історичного процесу як сьогодні, так і завжди, є незмінно специфічна в рамках конкретної соціальної спільності комунікація між індивідами, що здійснюється в контексті осягнення ними навколишнього світу. Це підводить нас до останнього розділу сучасної емпіричної соціології — аналізу людського досвіду, або аналізу дискурсів, або аналізу внутрішніх світів індивідів.

§ 1.3. Суспільство і текст

Використовуючи ідею Карла Поппера про третій світ (Поппер, 2002), наважимося припустити, що будь-яке суспільство в «сухому залишку» — це текст. А якщо мова заходить про людський досвід/дискурси/внутрішні світи індивідів, то текст — найбільш оптимальна форма їх фіксації з метою подальшого вивчення. Водночас з усіх об'єктів соціологічного дослідження саме текст є найбільш проблематичним. По-перше, через неймовірну динаміку індивідуального досвіду в сучасному світі тексти можна порівняти зі світлом далеких і таких, що давно змінилися, а подеколи й узагалі згаслих зірок. По-друге, всесвіт текстів розширюється так само швидко, як і наша світобудова. Соціальні науки готові відповісти на це великою різноманітністю дослідницьких стратегій аналізу тексту і дискурсів, а також постійно зростаючими можливостями технічних засобів і мов програмування.

Якщо говорити про дослідницькі стратегії, то ще 2000 року Стефан Тічер із своїми колегами здійснив порівняльний аналіз 12 таких підходів (Тічер, Мейер, Водак, & Веттер, 2009). Окрім того, стрімкий розвиток комп'ютерних технологій за минулі з тих часів роки відкрив дорогу у світ «Big Data» для багатьох представників наукової спільноти. Об'єднавши два ці аспекти, можна розраховувати на доступ до уявних спільнот сьогодення. І за цей доступ має відповідати третій різновид емпіричної соціології — емпірична соціологія дискурсів.

§ 1.4. Комплексні дослідницькі стратегії. Висновки

З усього сказаного вище логічно випливає можливість використання комплексних дослідницьких стратегій, які є комбінаціями кількох різновидів емпіричної соціології (Дембицкий, 2015с, с. 99). Трохи раніше ми не випадково згадали уявні спільноти, адже одне зі зразкових комплексних досліджень у соціології було здійснено Бенедиктом Андерсоном (Андерсон, 2001), який об'єднав історичні дослідження з вивченням людського досвіду, а в нашій термінології — неокласичну емпіричну соціологію й емпіричну соціологію дискурсів. Отримані результати назавжди увійшли до «золотого фонду» соціологічної думки.

Українська соціологія також має все необхідне для реалізації комплексних дослідницьких стратегій. Наприклад, дані соціологічного моніторингу «Українське суспільство» можна проаналізувати в контексті основних історичних подій і процесів у сучасній Україні, що передбачає проведення спеціального історико-соціологічного дослідження.

Підбиваючи підсумки короткого екскурсу в історію емпіричної соціології, слід наголосити на необхідності розширеного розуміння останньої, оскільки її можливості значно перевищують потенціал традиційних соціологічних досліджень. Надалі використаємо це розширене розуміння для специфікації ролі та місця соціологічних тестів в емпіричній соціології.

Глава 2

Теоретична валідизація вимірювальних шкал

Відправною точкою наших наукових пошуків є розробки підходу до теоретичної валідизації вимірювальних шкал, подальше опрацювання якого привело до методологічної роботи над основами соціологічного вимірювання за допомоги соціологічних тестів. Тому вважаємо за необхідне приділити особливу увагу теоретичній валідизації вимірювальних шкал. Засадничими принципами теоретичної валідизації є такі³.

1. Теоретична валідність являє собою міру відповідності теоретичного конструкту досліджуваному феномену.
2. Теоретична валідність є комплексною характеристикою, що включає очевидний, змістовий і композиційний компоненти.
3. Теоретична валідизація більшою мірою пов'язана з досягненням необхідного рівня відповідності теоретичного конструкту досліджуваному феномену, ніж із перевіркою.
4. Завершений процес теоретичної валідизації пов'язаний із забезпеченням інших видів валідності — конструктної, зовнішньої, внутрішньої та прогностичної.
5. Теоретична валідизація має здійснюватися з застосуванням комплексних дослідницьких стратегій.
6. Теоретична валідизація передусім пов'язана з парадигмою досліджень змішаного типу⁴.

Відштовхуючись від цих положень, спробуємо продемонструвати роль теоретичної валідизації у здійсненні соціологічного вимірювання — починаючи від критеріїв побудови шкали і закінчуючи специфікою інтерпретації отриманих даних. Як приклад використовуємо методику, призначену для вимірювання зусиль, докладених студентами у процесі навчання.

Вона базується на теоретичному конструкті, сформульованому в рамках дослідження, присвяченого феномену навчання за умов обмеженого часу (Дембицкий, 2016с, с. 109–124). Конструкт описує три типи студентів залежно від їхнього ставлення до навчання: «байдужий», «декларативно зацікавлений» і «відповідальний». Попри те, що теорія є справедливою стосовно студентів стаціонару, котрі поєднують роботу і навчання, поширення окремих її елементів

³ Докл. див.: (Дембицкий, 2016с).

⁴ Докладніше про дослідження змішаного типу див.: (Дембицкий, 2015b).

на всіх студентів стаціонару також є справедливим. Відповідні пояснення наведено далі.

§ 2.1. Перший етап: підготовка теоретичної бази

На відміну від узвичаєного підходу валідизації вимірювальних шкал у рамках теоретичної валідизації дослідник повинен розпочинати підготовку вимірювального інструменту, маючи у своєму розпорядженні валідний теоретичний конструкт, а ще краще — валідну теорію. У подальшому всі отримані результати вимірювання треба зіставляти, а в разі потреби й модифікувати відповідно до цих теоретичних положень.

У процесі побудови вимірювальної методики в якості теоретичної бази використано результати двох досліджень. Перше присвячене поведінковому репертуару студентів, який оформлюється в їхніх поведінкових стратегіях (Дембицкий, 2016с, с. 90–96), друге — згаданим вище типам студентів (Дембицкий, 2016с, с. 109–124).

Зупинімося докладніше на змістових характеристиках цих типів. Кожен із них можна описати двома поведінковими стратегіями — основною і додатковою (рідше трапляються ситуації, коли стратегії мають рівноцінне значення). Відповідні стратегії включають різний поведінковий репертуар (табл. 2.1). Опис поведінкових стратегій отримано завдяки дослідницькій стратегії узгодження концептів⁵, що дає змогу забезпечити змістову валідність⁶ теоретичного конструкту.

Типи студентів визначено в рамках використання дослідницької стратегії «обґрунтованої теорії»⁷ і можуть бути схарактеризовані так. Якщо студент належить до «байдужого» типу, він може використовувати: а) стратегію нероби як основну і стратегію спритника як додаткову; б) стратегію спритника як основну і стратегію нероби як додаткову (або однаковою мірою). Коли студент належить до «декларативно зацікавленого» типу, як основну використовує стратегію спритника, а як додаткову — стратегію трудяги. Зрештою, якщо студент належить до «відповідального» типу, він може використовувати: а) стратегію спритника і трудяги однаковою мірою; б) стратегію трудяги як основну, а стратегію спритника як додаткову.

Незважаючи на те, що у фокусі дослідження перебували студенти стаціонару, які поєднують роботу і навчання, отримані дані можна розподілити на дві частини — ту, що стосується тільки студентів, котрі працюють, і ту, що є спільною як для тих, хто працює, так і для тих, хто не працює. Зокрема, співвідношення типів студентів і відповідних поведінкових стратегій є справедливим для всіх студентів, що забезпечує високу зовнішню валідність⁸. Фактично оби-

⁵ Докладніше про побудову концептуальних карт див.: (Kane & Trochim, 2007).

⁶ Концептуальне представництво і значимість елементів теоретичного конструкту.

⁷ Докладно про використання «обґрунтованої теорії» див.: (Страус & Корбин, 2001).

⁸ Можливість вибірки виступати основою для узагальнення отриманих результатів щодо інших людей, контекстів або періодів.

дві групи з точки зору поведінкового репертуару різняться лише одним — наявністю або відсутністю поведінкової групи «Компроміси на роботі».

Таблиця 2.1

Змістові характеристики поведінкових стратегій студентів

<i>Назва групи поведінкових практик</i>	<i>Зміст групи</i>
1. Стратегія нероби	
1.1. «Неробство», сподівання на те, що поталанило	Давати обіцянки, що в подальшому обов'язково навчатимешся; використовувати моменти, коли викладач поспішає, і списувати борги; бути нахабним, йти на ризик; нічого не відпрацьовувати, йти з лекцій (семінарів) без дозволу; не працювати впродовж семестру і надолужувати заборгованість лише на сесії; складати іспити багаторазово (доки у викладача терпець не урветься) тощо.
1.2. «Мета виправдоує засоби»	Використовувати зв'язки; сказати, що роботу забув удома і попросити дозволу відповідати без потрібних матеріалів; якщо здаєш невчасно, списувати готові роботи і здавати їх; умовляти викладача і при цьому нічого не робити тощо.
1.3. Посилання на зовнішні причини	Пояснювати прогалини і невідвідування зовнішніми причинами і «тиснути» на те, що був змушений; посилатися на побутові проблеми, які заважають бути присутнім; використовувати тему здоров'я для виправдання тощо.
2. Стратегія спритника	
2.1. Робота з викладачем	Вивчити викладача і знайти до нього індивідуальний підхід; якщо раніше навчався добре, використовувати колишній авторитет; ранжувати викладачів за вимогливістю і залежно від цього мати час на підготовку.
2.2. Використання слухних ситуацій і моментів	У разі присутності на парах демонструвати активність; коли здаєш борги, перевести обговорення на матеріал, який знаєш; готуватися у транспорті, тощо.
2.3. Мінімізація зусиль	Показувати виконані завдання і відпрошуватися з пар; якщо є певний мінімум необхідних завдань, виконувати тільки його; якщо робиш завдання рідко, то, принаймні, якісно, бо це підкуповує; відвідувати найважливіші види занять (наприклад, практичні, а не лекції) тощо.
3. Стратегія тудяги	
3.1. Самоорганізація	Планувати, розкладати за пунктами те, що необхідно здійснити; всляко використовувати самоорганізацію тощо.
3.2. Опертя на дружні стосунки в навчальній групі	Допомагати студентам, які допомагають тобі, розподіляти підготовку завдань між різними людьми, брати в однокласників необхідні матеріали тощо.
3.3. Увічливість	Добре поводитися на парах; демонструвати позитивне ставлення до дисципліни; коли йдеш із пар, завжди погоджувати це з викладачем.
3.4. Опертя на власні сили	Завжди намагатися схопити суть матеріалу; якщо необхідно, залишатися після пар, щоби перескласти борги; за потреби витратити весь вільний час на підготовку тощо.

§ 2.2. Другий етап: розробка і коригування емпіричних індикаторів відповідно до використовуваної теорії

Спершу на підставі цієї інформації для кожної з указаних стратегій було сформульовано набір індикаторів (табл. 2.2). Усі індикатори оцінювалися за таким запитанням: «Студенти використовують різноманітні способи поведінки в різних ситуаціях, що стосуються навчання. Далі описано деякі з цих способів. Скажіть, як часто Ви особисто використовуєте кожен із них за такою шкалою: (1) майже ніколи не використовую (2), використовую доволі рідко (3), іноді використовую (4), використовую дуже часто (5), майже завжди використовую».

Таблиця 2.2

Набір індикаторів для вимірювання виразності поведінкових стратегій

Назва стратегії	Зміст індикаторів
Стратегія нероби	Велика кількість боргів, накопичених упродовж семестру, не є для мене серйозною проблемою, адже з ними можна впоратися і перед самою сесією
	Я здаю домашні завдання з відставанням від переважної частини групи. Це дає мені можливість звірити правильність виконання роботи і виправити помилки
	Свої пропуски я пояснюю особистими обставинами, що завадили бути присутнім на парі
	Для складання багатьох заліків/іспитів мені зазвичай достатньо ознайомитися з матеріалами конспекту за день, ба навіть у день складання
	Щоб отримати потрібну оцінку, я іноді просто чемно прошу про це викладача
	Коли я пояснюю свою відсутність на заняттях викладачеві, то посилаюся на здоров'я або на сімейні обставини
Стратегія спритника	Коли мені випадає нагода, я демонструю навчальну активність на парі
	Я намагаюся знайти правильний підхід до викладача, зважати на його індивідуальні особливості
	У разі боргів, складаючи іспит/залік, я намагаюся перевести обговорення на матеріал, який знаю
	Я витрачаю більше часу для підготовки до предметів вимогливіших викладачів
	Я переважно відвідую ті заняття (лекції або практичні), які важливіші для складання іспиту з предмета на сесії. Решту занять відвідую не завжди
	Якщо викладач задає певний мінімум необхідних завдань на семестр, для мене достатньо виконати тільки їх
Стратегія трудяги	Підготовку домашніх завдань ми (я та інші студенти моєї групи) розподіляємо між собою
	Я намагаюся планувати виконання моїх навчальних обов'язків
	У навчальних цілях я вдаюся до співпраці з моїми одногрупниками
	Моя поведінка на парах є правильною
	Я намагаюся вхопити суть навчального матеріалу і добре зрозуміти його
	За необхідності я використовую весь вільний час для підготовки до навчання

Під час пілотного опитування ($N = 16$) були також запитання щодо відвідування занять, частоти підготовки до практичних занять і успішності складання іспитів останньої сесії. Пілотаж показав, що такий спосіб визначення типу студента є незадовільним.

По-перше, отримані результати погано співвідносилися з контрольними запитаннями щодо частоти відвідувань занять, інтенсивності підготовки й успішності сесії. Головною причиною цього виявилось те, що індикатори, від початку призначені для вимірювання конкретної стратегії, учасники дослідження могли трактувати зовсім інакше й, отже, не відповідати закладеному в них змісту.

По-друге, результати, отримані на підставі індикаторів, що описують одну спільну стратегію, мали низьку внутрішню узгодженість. Так, найбільше значення α Кронбаха, яке дорівнює 0,44, було зафіксовано для індикаторів, що описують стратегію нероби. Для стратегій спритника і трудяги це значення дорівнювало 0,11 і 0,30 відповідно.

По-третє, специфіка такого вимірювання не відповідала особливостям наявної теорії. Річ у тім, що згідно з теорією, розробленою на основі вивчення феномену навчання за умов обмеженого часу, поведінкові стратегії є свого роду континуумом, що розпочинається стратегією нероби, продовжується стратегією спритника і завершується стратегією трудяги. Відтак поведінковий репертуар студента щонайліпше можна схарактеризувати віднесенням його до двох суміжних стратегій (*нероба і спритник*, *спритник і трудяга*), як це описано вище. У процесі дослідження за стратегією «обґрунтованої теорії» не трапилося жодного випадку, коли б студент використовував усі три стратегії однаковою мірою або ж коли поряд з основною стратегією мав дві додаткові як рівноцінні. За вимірювання з використанням набору індикаторів такі випадки траплялися. Гіпотетично можливі навіть такі ситуації, коли результати вимірювання набувають максимального значення для кожної зі стратегій, незважаючи на те, що такі випадки не фіксуються в емпіричній дійсності у процесі спостереження.

На наш погляд, усе це є свідченням того, що такий спосіб вимірювання багато в чому є артефактним і, відповідно, незастосовним для вивчення поведінкових стратегій студентів. Тому було розроблено інший спосіб вимірювання — на основі одного комплексного запитання, зміст якого також визначався описаними вище поведінковими стратегіями, тобто студентів безпосередньо запитували, яку поведінкову стратегію вони використовують, хоча, власне, йшлося саме про тип студента:

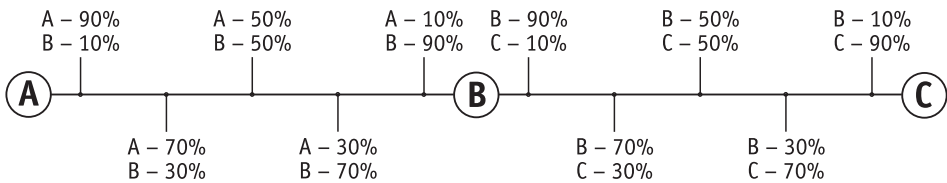
Припустімо, що назагал студентів можна розподілити на три групи, залежно від зусиль, що їх докладають у процесі навчання:

- a) *найменш старанні студенти, які здебільшого не докладають зусиль для навчання, сподіваючись на вдалий збіг обставин і на те, що всі борги можна буде закрити наприкінці семестру, а іспити і заліки скласти завдяки одногрупникам, навіть практично нічого не роблячи;*
- b) *студенти, які балансують між підготовкою найважливіших завдань та ігноруванням завдань, що не мають принципового значення, які використовують*

ранжування викладачів за суворістю і вимогливістю з метою відповідної концентрації зусиль і зазвичай відвідують ті пари, що мають більшу вагу в рамках дисципліни;

- с) найстаранніші, які ставляться до навчання максимально відповідально, намагаються готувати всі завдання, передбачені навчальними вимогами, намагаються здобути максимум знань і постійно відвідують пари.

А тепер, використовуючи наведену далі шкалу, вкажіть, на яке місце Вас можна було б поставити, якщо порівняти з представниками цих трьох груп. При цьому не можна обирати «чисту» групу (А, В або С), необхідно обрати одне зі значень між двома «чистими» групами (між А і В або між В і С)⁹.



§ 2.3. Третій етап: перевірка конструктивної валідності та ретестової надійності

Мінімальна перевірка конструктивної валідності пов'язана з забезпеченням двох її складових — конвергентної та дискримінантної валідності (Campbell & Fiske, 1959). Перша ґрунтується на тому, що результати шкали демонструють зв'язок із даними, отриманими за допомоги методик, призначених для вимірювання концептуально близьких соціальних феноменів. Друга, навпаки, ґрунтується на тому, що результати не пов'язані з даними методик, призначених для вивчення концептуально відмінних феноменів.

Головне побоювання, пов'язане з використанням поданої вище 10-бальної шкали, стосувалося можливості отримання соціально бажаних відповідей, тобто запитання могло виявитися сензитивним і зсунути відповіді в бік завищення оцінок. Для перевірки цього припущення до анкети було включено шкалу брехні з опитувальника ММРІ.

Після вдалого¹⁰ пілотного опитування ($N = 24$) шкалу перевірено на двох вибірках різних ВНЗ ($N_1 = 61$, $N_2 = 62$). Здогадно в разі адекватного відбиття шкалою теоретичного конструкту результати її застосування мали корелювати з відповідями на запитання про кількість виступів на практичних заняттях і про успішність останньої сесії, негативно корелювати з відповідями на запитання про кількість невідвідувань і не мати кореляції зі шкалою брехні. Для відповідної перевірки було використано ранговий коефіцієнт кореляції Спірмена (табл. 2.3).

⁹ В оригіналі для 10-пунктової шкали наведено таблицю, а не малюнок. Тут складно її навести через недостатню ширину сторінки.

¹⁰ Мається на увазі, що не спостерігалось зміщення відповідей до вищих градацій шкали.

Таблиця 2.3

Сила і напрямок зв'язку шкали з контрольними змінними

Контрольна змінна	Сила зв'язку		
	Перший ВНЗ ($N = 61$)	Другий ВНЗ ($N = 62$)	Разом ($N = 123$)
Кількість виступів на практичних заняттях	0,74**	0,69**	0,71**
Успішність складання останньої сесії	0,67**	0,67**	0,67**
Кількість пропусків	-0,56**	-0,64**	-0,59**
Шкала брехні	-0,16	0,20 ¹¹	-0,09 ¹²

** $p < 0,01$

Отримані результати засвідчили, що ця шкала відображає теоретичний конструкт, який описує поведінкові стратегії студентів, а також слабко піддається зсувам, пов'язаним із соціально бажаними відповідями.

Разом із тим наведені дані стосуються перевірки шкали, першою чергою на рівні самозвіту, який багато в чому є суб'єктивним. Тому було здійснено додаткову її перевірку з залученням експертів.

До складу останніх увійшли по два викладачі та одному студенту на кожній із кафедр. У кожному ВНЗ оцінювали всіх студентів однієї навчальної групи ($N_1 = 20$, $N_2 = 24$). Викладачі, відібрані як експерти, були добре знайомі з навчальними групами, а студенти-експерти навчалися у складі цих груп і працювали на цих кафедрах. Оцінювання проводили за 10-бальною шкалою комплексного запитання. Крім цього, для кожного зі студентів розраховано середнє значення оцінок на сесіях, отриманих ним за всі роки навчання¹³. Для визначення сили зв'язку використано ранговий коефіцієнт кореляції Спірмена. Оскільки оцінки експертів високою мірою узгоджені (для першого ВНЗ значення Альфа Кронбаха дорівнює 0,89, для другого — 0,96), додатково розраховано середнє значення оцінок експертів для кожної кафедри та обчислені відповідні коефіцієнти кореляції (табл. 2.4).

Наведені дані ґрунтуються на точнішій інформації, ніж самозвіти. По-перше, це пов'язане з узгодженістю результатів, отриманих у перебігу оцінювання різними експертами, по-друге, забезпечується відсутністю втрати інформації стосовно успішності навчання, оскільки враховано всі роки. Отже, результати експертної перевірки цілковито підтверджують сформульовані раніше висновки.

¹¹ Кількість осіб у другому ВНЗ, опитаних за допомоги шкали брехні, становила так само 17. Надалі цей взаємозв'язок було перевірено на більшій вибірці студентів цього ВНЗ, але статистично значимого зв'язку так і не було виявлено ($N = 41$, $r_s = -0,03$).

¹² $N = 76$.

¹³ Використано дані про успішність складання сесії перших чотирьох років навчання.

Таблиця 2.4

**Сила і напрямок зв'язку результатів експертного оцінювання
з середніми значеннями оцінок, отриманих на всіх сесіях**

<i>Експерти</i>	<i>Сила зв'язку</i>	
	<i>Перший ВНЗ</i>	<i>Другий ВНЗ</i>
Перший викладач	0,83**	0,93**
Другий викладач	0,72**	0,91**
Студент	0,80**	0,93**
Середнє значення трьох оцінок	0,85**	0,95**

** $p < 0,01$

Після перевірки конструктивної валідності у трьох навчальних групах проведено опитування, щоб перевірити ретестову надійність. Узгодженість двох опитувань, оцінена з застосуванням коефіцієнта кореляції Спірмена, дорівнює 0,75 ($N = 37, p < 0,01$).

§ 2.4. Четвертий етап: специфіка інтерпретації отриманих результатів

На особливу увагу заслуговує специфіка інтерпретації результатів, отриманих із застосуванням цієї шкали. Річ у тім, що за змістом шкала ґрунтувалася на інформації про поведінкові стратегії студентів, а кінцевою метою її застосування було отримання інформації про належність кожного студента до одного з трьох типів — «байдужого», «декларативно зацікавленого» і «відповідального». З огляду на це інтерпретація шкали здійснювалася на підставі суто теоретичних положень, що пов'язують поведінкові стратегії з типами студентів (табл. 2.5).

Таблиця 2.5

Співвідношення значень шкали з типами студентів

<i>Тип студента</i>	<i>Поведінкові стратегії</i>	<i>Значення шкали</i>
1. «Байдужий»	1.1. Стратегія нероби — основна, стратегія спритника — додаткова	1, 2
	1.2. Стратегія спритника — основна, стратегія нероби — додаткова або рівноцінна	3, 4, 5
2. «Декларативно зацікавлений»	2.1. Стратегія спритника — основна, стратегія трудяги — додаткова	6, 7
3. «Відповідальний»	3.1. Стратегії спритника і трудяги використовуються рівною мірою	8
	3.2. Стратегія трудяги — основна, стратегія спритника — додаткова	9, 10

Отже, студентів, котрі обрали п'ять перших градацій шкали, віднесено до «байдужого» типу, тих, що обрали шосту і сьому градації, — до «декларативно зацікавленого», тих, які обрали останні три, — до «відповідального».

Для перевірки правильності такої інтерпретації її результати були співвіднесені з результатами відповідей щодо успішності складання останньої сесії (табл. 2.6, 2.7), попередньо перетворених у три градації: 1) під час складання сесії були трійки; 2) сесію складено на чотири і п'ять; 3) сесію складено на відмінно.

Для кожної групи студентів обчислено ранговий коефіцієнт кореляції Гамма. Для обох ВНЗ він дорівнює 0,83 ($p < 0,01$).

Таблиця 2.6

Зв'язок типів студентів з успішністю складання сесії (перший ВНЗ)

Результати складання останньої сесії	Тип студента			Разом
	«Байдужий»	«Декларативно зацікавлений»	«Відповідальний»	
Були трійки	15	6	1	22
Чотири і п'ять	4	10	2	16
На відмінно	1	6	14	21
Разом	20	22	17	59

Таблиця 2.7

Зв'язок типів студентів з успішністю складання сесії (другий ВНЗ)

Результати складання останньої сесії	Тип студента			Разом
	«Байдужий»	«Декларативно зацікавлений»	«Відповідальний»	
Були трійки	21	4	3	28
Чотири і п'ять	5	11	8	24
На відмінно	0	1	9	10
Разом	26	16	20	62

Попри те, що сила і напрямок зв'язку однакові для обох ВНЗ, характер зв'язку істотно різниться. Проте в рамках пропонованої роботи це питання не розглянуто, оскільки воно пов'язане не зі специфікою шкали, а з особливостями навчального закладу¹⁴. Назагал ці результати підтверджують правомірність використання запропонованої інтерпретації, хоч і вказують на необхідність її специфікації стосовно вищих навчальних закладів різних типів, що, своєю чергою, є важливою умовою досягнення композиційної валідності шкали.

Разом із тим у контексті групування пунктів шкали цілком логічним може виглядати сумнів стосовно пріоритетності саме такого варіанта групування.

¹⁴ Відповідне питання розглянуте в: (Дембицкий, 2016с, с. 136–149).

Найочевиднішими альтернативами є такі подані далі варіанти розподілу шкали, які б поділяли її на приблизно рівні групи.

Перший варіант (3/4/3):

1 група / «Байдужий» тип	A — 90% B — 10%
	A — 70% B — 30%
	A — 50% B — 50%
2 група / «Декларативно зацікавлений» тип	A — 30% B — 70%
	A — 10% B — 90%
	B — 90% C — 10%
	B — 70% C — 30%
3 група / «Відповідальний» тип	B — 50% C — 50%
	B — 30% C — 70%
	B — 10% C — 90%

Другий варіант (3/3/4):

1 група / «Байдужий» тип	A — 90% B — 10%
	A — 70% B — 30%
	A — 50% B — 50%
2 група / «Декларативно зацікавлений» тип	A — 30% B — 70%
	A — 10% B — 90%
	B — 90% C — 10%
3 група / «Відповідальний» тип	B — 70% C — 30%
	B — 50% C — 50%
	B — 30% C — 70%
	B — 10% C — 90%

Третій варіант (4/3/3):

1 група / «Байдужий» тип	A — 90% B — 10%
	A — 70% B — 30%
	A — 50% B — 50%
	A — 30% B — 70%
2 група / «Декларативно зацікавлений» тип	A — 10% B — 90%
	B — 90% C — 10%
	B — 70% C — 30%
3 група / «Відповідальний» тип	B — 50% C — 50%
	B — 30% C — 70%
	B — 10% C — 90%

Якщо ці варіанти групування є вдалішими у плані відображення типу студента, то вони мають демонструвати вищі порівняно з ініціальним групуванням (5/2/3) коефіцієнти кореляції (Гамма) з успішністю навчання студентів (табл. 2.8).

Таблиця 2.8

Величина коефіцієнта Гамма для різних варіантів групування пунктів шкали
($p < 0,01$)

Варіант групування шкали	Сила зв'язку для першого ВНЗ	Сила зв'язку для другого ВНЗ	Сила зв'язку для обох ВНЗ*
5/2/3	0,83	0,83	0,81
3/4/3	0,84	0,80	0,79
3/3/4	0,82	0,89	0,83
4/3/3	0,81	0,78	0,77

* Масиви об'єднано.

Як можна бачити з наведених даних, найефективнішим варіантом групування є варіант 3/3/4, а вже потім 5/2/3. Проте якщо порівняти розподіли таблиць зіставності для двох варіантів групування, то матимемо протилежну картину (табл. 2.9, 2.10). В обох таблицях наведено інформацію одразу про два ВНЗ.

Таблиця 2.9

Зв'язок типів студентів з успішністю складання сесії (групування 3/3/4)

Результати складання останньої сесії	Тип студента			Разом
	«Байдужий»	«Декларативно зацікавлений»	«Відповідальний»	
Були трійки	16	29	5	50
Чотири і п'ять	1	20	19	40
На відмінно	1	3	27	31
Разом	18	52	51	121

Таблиця 2.10

Зв'язок типів студентів з успішністю складання сесії (групування 5/2/3)

Результати складання останньої сесії	Тип студента			Разом
	«Байдужий»	«Декларативно зацікавлений»	«Відповідальний»	
Були трійки	36	10	4	50
Чотири і п'ять	9	21	10	40
На відмінно	1	7	23	31
Разом	46	38	37	121

Якщо уявити функціональний зв'язок цієї залежності, логічним буде припустити, що всі байдужі студенти мали би трійки, декларативно зацікавлені навчалися би на чотири і п'ять, а відповідальні — на відмінно. Виходячи з такого критерію, можна оцінити, наскільки кожен варіант групування відхиляється від цієї ідеальної моделі. За варіанта 3/3/4 відхилення спостерігаємо у 48% випадків (58 із 121), а за варіанта 5/2/3 — у 34% випадків (41 із 121)¹⁵. У таблиці, побудованій для варіанта 3/3/4, видно, що значна частина студентів, які мають трійки, потрапила в декларативно зацікавлену групу, що не відповідає поданим вище теоретичним положенням. Власне, саме останній аспект, а не сила встановлених взаємозв'язків, має вирішальне значення для оцінювання різних варіантів групування.

Загалом більшу силу коефіцієнтів кореляції для варіанта 3/3/4 можна пояснити насамперед особливостями обчислення самого показника Гамма і, відповідно, вони несуть радше технічне, ніж смислове навантаження¹⁶.

§ 2.5. Деякі аспекти використання методики

За формулювання висновків необхідно сказати як про методологічні аспекти теоретичної валідизації, так і про методичні аспекти використання комплексної 10-бальної шкали, розглянутої вище.

У методологічному плані можна дати дві головні рекомендації стосовно теоретичної валідизації вимірювальних шкал.

1. Перш ніж переходити до соціологічного вимірювання, треба отримати теоретичні положення, що описують конструкти соціальних феноменів, які будуть об'єктом вимірювання. При цьому розроблені теоретичні положення мають бути валідними (тобто відповідати емпіричній дійсності, що лежить у фокусі дослідження) уже до вимірювання. У цьому сенсі теорія є первинною, а вимірювання — вторинним, похідним від неї. Вимірювальну шкалу розробляють, модифікують і оцінюють у контексті наявної теорії. Завдяки цьому можна досягти змістової та логічної валідності вимірювальної шкали, яка слугує базою для досягнення конструктної валідності.

2. Після того як шкала пройшла перевірку на конструктну валідність, дослідник має здійснити вимірювання щодо тих соціальних груп, для вивчення яких ця шкала придатна. Це необхідно переважно для того, щоб мати уявлення про засади інтерпретації результатів, отримуваних за її використання. Тут знову ж таки неможливо обійтися без валідних теоретичних положень, що вписує конструкт, покладений в основу вимірювання, в ширше концептуальне поле. У цьому разі вже вимірювальну шкалу (а точніше, її результати) використовують для збагачення тієї теорії, на підставі якої вона була побудована.

¹⁵ Для варіантів групування 3/4/3 і 4/3/3 величина відхилення становить 43% і 42% відповідно.

¹⁶ Для докладного ознайомлення з особливостями розрахунку коефіцієнта Гамма див., напр.,: (Хили, 2005, с. 416–420).

У методичному плані можна рекомендувати таке¹⁷.

1. Не включати методику в анкети, призначені для самозаповнення, або пересвідчуватися, що респондент правильно її зрозумів. Слова Сеймура Садмена про те, що респонденти є когнітивними скупцями, тут цілком вочевиднюються. Часто студенти навіть не дочитують інструкцію до кінця і позначають один з описаних у запитанні типів, навіть після прохання уважно прочитати інструкції запитання повністю. Тому бажано застосовувати методику в інтерв'ю типу face-to-face. Якщо ж проводити групове інтерв'ю (наприклад, із навчальною групою), необхідно докладно пояснювати специфіку запитання, а також перевіряти результати відповідей, щоб мати можливість унести поправки в разі неправильного заповнення.

2. У комплексних освітніх дослідженнях, котрі передбачають формування бази даних щодо кожного студента, слід опитувати викладачів, а не студентів — отримувана інформація про студентів матиме точніший характер.

3. Слід дуже акуратно ставитися до продемонстрованих у цій роботі взаємозв'язків. Річ у тім, що специфіка освітньої установи може вносити в них суттєві корективи. Останнє має конче важливе значення з точки зору інтерпретації отриманих результатів. Найліпший спосіб уникнути хибних висновків — триангуляція методів (наприклад, із використанням включеного спостереження). З цих міркувань описана вище методика є найпридатнішою в дослідженнях змішаного типу, що передбачають отримання даних із різних джерел.

§ 2.6. Загальна оцінка застосовності теоретичної валідизації вимірювальних шкал у соціології

На наш погляд, незважаючи на загальне правильне фокусування на специфіці розробки вимірювальних інструментів у соціології (досягнення теоретичної валідності => розробка методики => її оцінка в контексті теоретичних положень), у межах цієї роботи наведені результати висувують більше методологічних запитань, аніж дають відповідей. До найважливіших належать такі.

1. З огляду на поданий підхід, що можна сказати про більш комплексні вимірювальні інструменти (шкали та індекси) в соціології?

2. Чи виправдане впровадження категорії «соціологічний тест» за наявності загальноприйнятої категорії «шкала»?

3. Чи є придатними положення класичної теорії вимірювання з точки зору соціологічного вимірювання?

4. Якщо ні, то які є релевантні припущення альтернативного підходу?

5. Яким є повний процес розробки вимірювальних методик у соціології?

У подальшому матеріалі розділу розглянуто відповіді на ці запитання.

¹⁷ Наведені далі рекомендації ґрунтуються на досвіді, отриманому у процесі збору інформації, який не відображено в монографії. Ця інформація досить звична з точки зору проведення прикладних досліджень, але вельми важлива для тих, хто відчує необхідність використання шкали у своєму дослідженні.

Глава 3

Соціологічний тест як вимірювальний інструмент у соціальних науках

§ 3.1. Вступ

У соціальних дослідженнях вельми поширені комплексні вимірювальні інструменти, звернення до яких є першим кроком у розумінні соціологічного тестування. До них автор відносить шкали та індекси, розуміючи під їх комплексністю те, що вони складаються з набору спеціально підібраних індикаторів, які розв'язують єдине наукове завдання — вимір вираженості певного явища. Їх комплексність зумовлена і тим, що у процесі розробки необхідні спеціальні методичні дослідження, покликані продемонструвати наявність у таких інструментів конвенційно затверджених показників якості — надійності та валідності¹⁸.

Якщо говорити загально про відмінності між шкалами та індексами (DeVelis, 2012, р. 12), то в першому випадку йдеться про вимірювання певної властивості, яка безпосередньо не спостерігається (латентної змінної). Саме вона зумовлює ті фактори свого прояву (непрямі індикатори), які закладені в методику. Предметом вимірювання інтегральних індексів може бути як латентна змінна, так і соціальний феномен. Ще одна відмінність між шкалами та індексами полягає в тому, що за використання індексу об'єкт його вимірювання обумовлений тими факторами, на підставі яких сформульовані індикатори¹⁹.

Один із ключових факторів, що визначає вибір між шкалою та індексом, обумовлений природою досліджуваного явища. У разі соціологічного дослідження вона може бути як індивідуальною, так і суспільною. Щоб прояснити цю думку, розглянемо як приклад твердження, які можуть бути покладені в основу шкали або індексу авторитаризму:

¹⁸ У цьому розділі йтиметься тільки про валідність. Це обумовлено тим, що з точки зору методології соціальних наук вона має дискусійний характер, натомість надійність вимірювання — поняття досить уніфіковане і таке, що не потребує докладного розгляду, враховуючи його детальну розробленість. Надійність соціологічного вимірювання буде розглянуто в четвертому розділі, присвяченому методичним рекомендаціям.

¹⁹ Принциповим методичним результатом такого поділу є нерелевантність внутрішньої надійності в рамках обґрунтування якості інтегральних індексів. У цьому разі наявність взаємозв'язку між різними причинними факторами не є обов'язковою.

- єдина політична воля забезпечує стабільність у суспільстві краще, ніж набір різних точок зору;
- для нормального розвитку країні потрібна «сильна рука», а не розмови про демократію;
- тільки політика «сильної руки» може зберегти порядок у суспільстві;
- сильна держава можлива лише тоді, коли на чолі країни стоїть один лідер;
- наявність єдиної партії та сильного політичного лідера дає людям упевненість у завтрашньому дні.

З одного боку, подані твердження можуть бути оцінені як непрямі індикатори латентної змінної, з іншого — кожен із них говорить про одну з цінностей, які характеризують суспільство, — стабільність, розвиток, порядок, сила, соціальний оптимізм. Якщо підходити до аналізу тверджень з точки зору суспільних цінностей, то не буде концептуальної та методологічної необхідності апелювати до поняття латентної змінної. Наприклад, згода з тим, що в авторитарному суспільстві порядку більше, не має викликати очікувань щодо згоди з тим, що авторитарний лідер забезпечить розвиток або соціальний оптимізм. Тут респондент постає вже не як «пасивне спостереження», яке повідомляє нам інформацію про свій внутрішній стан, а як експерт, який демонструє ціннісні характеристики певного суспільства. Відповідно, з методологічної точки зору цілком може йтися не про шкалу, а про індекс, спрямований на вимірювання соціального феномену — ціннісної структури.

Щодо використання комплексних вимірювальних інструментів слід зазначити, що емпірична соціологія, як відомо, принципів підстави вимірювання запозичує з психології (класична теорія вимірювання) або навіть педагогіки (математична теорія вимірювання²⁰) так, ніби відповідні методологічні положення априорі наділені універсальною природою. Це значно обмежує сферу використання шкал та індексів у рамках масових опитувань.

§ 3.2. Поняття «вимірювання». Вимірювання в соціальних науках

У загальнонауковому значенні поняття «вимірювання» позначає процес подання властивостей реальних об'єктів у вигляді числової величини. Величину розуміють як усе те, що може бути більшим або меншим, притаманним об'єкту більшою або меншою мірою. Відповідно, числова величина — така, що може бути позначена числом. Отже, вимірювання є встановленням числових відношень між властивостями об'єктів (Никифоров, 2006, с. 138). Відомий український соціолог В. Паніотто визначає вимірювання як особливу процедуру, внаслідок якої виникає числова модель досліджуваних властивостей об'єкта (Паніотто, Максименко, & Харченко, 2004, с. 10).

Говорячи про вимірювання в соціальних науках, ми намагаємося акцентувати увагу на його «сіру зону». Перш за все визначимо, що не входить до цієї зони. Вимірювання на основі змінних, що належать до номінальних, порядко-

²⁰ Широко відома як «item response theory». Використовуваний переклад запропонований В. Аванесовим (Аванесов, 2007).

вих і метричних шкал²¹, є несуперечливим і зрозумілим. Нема ніяких невизначеностей щодо використання цих шкал у соціальних науках — воно може бути реалізовано так само, як і у природничих науках (хоча соціологи часто порушують відповідні принципи). З огляду на це до «сірої зони» можна віднести інтервальне вимірювання в соціальних науках. Як зазначає В. Паніотто, в соціології існують шкали, які близькі до інтервальних і мають орієнтовно, але не точно, рівні інтервали між пунктами шкали (Паніотто, Максименко, & Харченко, 2004, с. 19). Він пропонує називати їх псевдоінтервальними. Кращою назвою, на нашу думку, є «квазіінтервальні шкали». Квазіінтервальне вимірювання посідає проміжну позицію між порядковим та інтервальним. Фактично на даний момент це найточніше вимірювання соціальних установок.

Проблематичність квазіінтервального вимірювання полягає в тих засобах статистичної обробки, що є прийнятними для відповідних даних. Це ж саме є справедливим і для побудови квазіінтервальних показників. З погляду природничих наук такого роду дані слід розглядати як порядкові. У соціальних науках, зокрема й у соціології, дуже часто їх розглядають і обробляють як інтервальні.

З огляду на це вимірювання за допомоги комплексних вимірвальних інструментів у соціології ми розглядаємо як квазіінтервальне. При цьому вважаємо, що воно наближене саме до інтервального, а не порядкового рівня.

§ 3.3. Концептуальні кластери вимірювання в соціальних науках

У відомій «Енциклопедії вимірювання і статистики» за редакцією Н. Салкаінда (Salkind, 2007) на термін «тест» натрапляємо в назвах статей, що переважно належать до трьох тематичних розділів — поняття і проблеми у вимірюванні, статистичні техніки, статистичні тести. У пропонованій роботі соціологічні тести розглянуто першою чергою в контексті понять і проблем вимірювання психологічних, соціально-психологічних і соціальних феноменів.

Цікаво, що попри часте вживання терміна «тест» у назвах статей згаданої енциклопедії його визначення не наведено. При цьому є визначення шкалювання, під яким розуміють діяльність із метою вимірювання чи квантифікації психологічних установок (атитюдів) або властивостей (атрибутивів) (Parker, 2007, р. 862). Саме вимірювання розглядають як із класичної (С. Стівенс), так і з сучасної точок зору. У першому випадку під ним розуміють «присвоєння числових значень об'єктам або подіям згідно з правилами», при цьому вимірювання розглядають як процедуру підготовки даних до статистичного аналізу. У другому випадку вимірювання і статистику вважають фундаментально різними

²¹ У цьому разі шкалу розуміють не як комплексний вимірвальний інструмент, а як набір властивостей об'єкта і чисел, що відповідають цим властивостям (Паніотто, Максименко, & Харченко, 2004, с. 11). Такий зміст цього поняття, що відповідає відомій класифікації рівнів вимірювання С. Стівенса, ми застосовуємо лише тут, тобто далі його не використовуємо. Загалом, щоб відійти від такої термінологічної двозначності, замість поняття «шкала» в разі класифікації С. Стівенса можна застосовувати поняття «рівень».

поняттями із власними допущеннями і результатами використання (Stemler, 2007, р. 582–584). У рамках цієї монографії соціологічні тести розглянуто як вимірювальні інструменти.

Слід коротко розглянути чотири ключові теорії, що стосуються розробки вимірювальних інструментів у соціальних науках, — класичну теорію тестів, математичну теорію вимірювання (Item Response Theory), теорію надійності та теорію валідності, щоб оцінити їх релевантність у ракурсі соціологічного тестування.

Класична теорія тестів. З точки зору цієї теорії, значення, отримане в результаті використання тесту, включає два компоненти — істинне значення і помилку вимірювання. Істинне значення розглядають як середнє ідентичних значень, отриманих у результаті повторюваних без обмеження вимірювань. Отже, ключове допущення цієї теорії полягає в тому, що отримане в рамках тестування значення є сумою істинного значення і помилки. При цьому істинне значення завжди невідоме. Вимірювальні помилки — це будь-які аспекти вимірювання, крім істинного значення. Їх джерелами можуть бути добір пунктів методики, інструкція тесту, підрахунок результатів і будь-які систематичні помилки вимірювання. Перші три джерела ведуть до несподіваних і неузгоджуваних ефектів, тобто справляють несистематичний вплив на результати вимірювання і, відповідно, є випадковими помилками. Ці помилки не пов'язані ні з істинним значенням, ні з помилками інших вимірювань такого ґатунку. Систематичні помилки виникають у тих випадках, коли тест вимірює щось відмінне від того, що він має вимірювати (проблема валідності). Друге допущення стосується паралельних тестів. Воно полягає в тому, що кожен індивідуальний пункт методики сам по собі можна розглядати як тест. Це пов'язане з тим, що кожен із них є похідним від латентної змінної. При цьому вплив латентної змінної та величина випадкових помилок є однаковими в усіх випадках (Alexopoulos, 2007, р. 140–142).

Не відкидаючи загального фокусування теорії (істинне значення + помилка вимірювання), важко погодитися з тими акцентами, що їх зроблено в запитаннях про випадкові помилки. Систематичний зсув може бути внесений будь-яким елементом тесту, а не тільки його невалідністю²². Сумнівним виглядає і допущення про однаковий вплив латентної змінної на всі індикатори тесту, особливо коли йдеться не про вимірювальну шкалу, а про інтегральний індекс (це питання розглянемо далі).

Математична теорія вимірювання. У рамках цієї теорії стверджується ймовірнісний характер взаємозв'язку між відповіддю на той або той пункт тесту і латентною змінною. Математична теорія вимірювання включає статистичні моделі оцінювання якості пунктів тесту. Найпоширенішими і найпопуляр-

²² Систематична помилка є такою, що 1) не виникає випадково; 2) має тенденцію виникати в усій вибірці; 3) може призводити до зміщення в результатах опитування. Для кращого розуміння природи випадкових і систематичних помилок радимо звернутися до парадигми сукупної помилки опитування (Total Survey Error Paradigm). Див. напр.: (McNabb, 2014, р. 60–79).

нішими є моделі для оцінювання тестів, що призначені для вивчення одновимірних характеристик і застосовують пункти з дихотомічною шкалою. Більш комплексні моделі існують і для випадків, коли пункти можуть включати шкали відповідей із більш як двома категоріями і/або коли досліджувані характеристики є багатовимірними (Keller, 2007, р. 493–494).

Ця теорія виглядає прийнятнішою аналітичною схемою порівняно з класичною теорією тестів. Разом із тим, беручи до уваги акцент математичної теорії вимірювання на таких властивостях індикаторів тесту, як рівень підготовленості опитуваного, показник рівня складності завдань, здатність тесту відрізнити підготовлених респондентів від непідготовлених тощо (Аванесов, 2007), вона є слушною насамперед для розробки педагогічних, а не соціологічних тестів.

Теорія надійності. Надійність вимірювання пов'язана з його узгодженістю в різноманітних контекстах. Виокремлюють три аспекти надійності: 1) результати надійного інструменту не залежать від того, хто проводить дослідження; 2) використання того самого надійного інструменту в різний час дає однакові результати; 3) усі частини надійного інструменту є взаємозалежними. Відповідно, можна вирізнити три види надійності — надійність оцінювача або кодувальника (результати оцінювання різними експертами того самого об'єкта за посередництва того самого інструменту ведуть до однакових результатів), тест-ретестова надійність (зміни в результатах повторних вимірювань відповідають теоретичним очікуванням), внутрішня узгодженість (взаємопов'язаність пунктів тесту є високою) (Juni, 2007, р. 834).

Звісно, перевірка надійності є важливою умовою розробки соціологічного тесту. При цьому актуальність того чи того різновиду надійності може істотно варіювати відповідно до тих завдань, для виконання яких розроблено конкретний тест.

Теорія валідності. Розглядаючи цю концептуальну царину з точки зору кількісних досліджень, можна (але з певною мірою умовності) говорити про єдину теорію. Якщо ж брати до уваги імплікації якісного і змішаного підходів, то має йтися про цілий набір теорій валідності. І оскільки, на наш погляд, теорія валідності в рамках кількісного підходу є суттєво лімітованою, далі приділено увагу власне проблемі валідації.

§ 3.4. Поняття «соціологічний тест»

Кажучи коротко, поняття «соціологічний тест» описує ті шкали та індекси, які відповідають низці соціологічних вимог.

Поняття «соціологічний тест» упродовжено в науковий обіг радянським, а нині російським соціологом і педагогом В. Аванесовим, який дав йому два визначення. Відповідно до першого соціологічний тест — це «система відібраних соціологічними методами висловлювань, які пред'являють респондентам із метою отримання надійної та валідної інформації про ознаки, що цікавлять дослідника»; відповідно до другого — «система висловлювань, що дозволяє отримати обґрунтоване відображення емпіричної системи з відносинами, яка

цікавить дослідника, в числову систему з відносинами» (Аванесов, 1982, с. 41). У рамках цієї роботи ми дотримуємося першого визначення.

Двома аргументами В. Аванесова на користь виокремлення соціологічного тесту як самостійного методу дослідження є такі (Аванесов, 1982, с. 38–39):

- 1) необхідність вивчення в соціологічних дослідженнях не тільки психологічних (для цього використовують психологічні тести), а й соціологічних феноменів, таких як ціннісні орієнтації, світоглядні установки, соціальні потреби тощо;
- 2) актуальність цього напрямку досліджень на сучасному етапі розвитку методології соціальних наук.

На наш погляд, через більш ніж 30 років обидва аргументи не втратили своєї ваги, проте методологія розробки і використання соціологічних тестів за ці роки потрапила в поле інтересів лише двох українських соціологів — Н. Паніної та Є. Головахи. Ними розроблено низку соціологічних тестів, які сьогодні використовують у масових опитуваннях.

Н. Паніна та Є. Головаха поширили розуміння соціологічного тесту і на психологічні методики, спеціально адаптовані для використання в масових опитуваннях. І це повністю виправдано, оскільки більшість психологічних тестів є дуже об'ємними, що накладає суттєві обмеження на їх використання в соціологічних дослідженнях. Окрім того, Н. Паніна і Є. Головаха сформулювали низку вимог, які має задовольняти соціологічний тест (Головаха & Паніна, 1997, с. 5):

- універсальність — методика має включати найзагальніші сутнісні індикатори досліджуваного явища, що відкриває можливість її використання практично в будь-якому соціологічному дослідженні;
- інтегральність — можливість зведення різних вимірів соціологічного тесту до одного інтегрального показника;
- якісність — соціологічний тест має бути «паспортизований» статистичними показниками надійності та валідності;
- стандартизованість — установлення «норм» досліджуваного явища, що дозволить інтерпретувати отримані результати з точки зору вираженості досліджуваного феномену;
- компактність — соціологічний тест має бути досить компактним, щоб його можна було використовувати в масових опитуваннях;
- «чутливість» — з огляду на вимогу компактності соціологічний тест має вирізнятися прийнятною точністю вимірювання;
- об'єктивність — мета вимірювання має бути прихованою від респондента, що дозволяє уникнути фальсифікації результатів тестування.

З огляду на викладене вище можна виокремити два виміри соціологічного тестування. Перший, і принциповий, стосується необхідності забезпечити методичні вимоги, які дозволяють повноцінно використовувати вимірювальну методику (неважливо — психологічну чи соціологічну) в рамках масових опитувань. Уже в цьому разі можна говорити про соціологічне тестування. Другий вимір пов'язаний з тими аспектами соціальної дійсності, для вивчення яких

застосовують соціологічний тест. Методики, що відповідають цьому виміру, спрямовані на вивчення соціальних феноменів. Відповідно, підготовка концептуальної бази соціологічного тесту здійснюється переважно соціологічними засобами.

Крім зазначених, вважаємо, слід виокремити і третій вимір соціологічного тестування, який пов'язаний із потенціалом використання отриманих результатів. В ідеалі соціологічний тест має бути не просто розроблений як валідна і надійна методика, спрямована на вивчення соціального або психологічного феномену, а й бути пов'язаним із ширшими теоретичними положеннями про будову і функціонування соціальної системи (Дембицкий, 2016b, с. 144). Завдяки цьому відкривається можливість діагностичних висновків про стан її різних частин, можливі перспективи її функціонування і розвитку. Однак навіть якщо відповідні теоретичні положення недоступні, але при цьому методика задовольняє вимоги першого виміру соціологічного тестування, відповідні положення можуть бути сформульовані у процесі використання соціологічного тесту в емпіричних дослідженнях.

Щодо закордонної практики використання соціологічних тестів, то перш за все слід зазначити, що такого терміна в закордонній соціології ми не знайшли. Але використовуючи окреслені вище виміри соціологічного тестування, легко виявити, що низка застосовуваних у західній соціології вимірвальних методик задовольняє вимоги соціологічного тестування. Наведемо два загальновідомі приклади. Першим вимірвальним інструментом, який можна розглядати як соціологічний тест, є методика вивчення цінностей Ш. Шварца²³. Вона має прийнятний для масових опитувань розмір, спрямована на вивчення соціального феномену, а результати її застосування повсюдно використовують для оцінки стану різних суспільств. Другим прикладом може бути відомий тест для оцінки психологічного дистресу (К6), розроблений Р. Кесслером²⁴. Цей тест широко використовують в епідеміологічних опитуваннях. Цікавим є той факт, що незважаючи на психологічну спрямованість, ця методика завдяки своїй стислості придатна саме для масових опитувань, а не для індивідуальної психологічної діагностики.

Тут у читача може виникнути питання, чи треба в такому разі взагалі говорити про соціологічне тестування. Мається на увазі ситуація, коли соціологічні тести розробляють, незважаючи на відсутність відповідного поняття і спеціальної концептуальної галузі. Наша ствердна відповідь на нього пов'язана з тим, що специфіка соціологічного вимірювання повністю не охоплюється методологічними засобами сучасних вимірвальних підходів у соціальних науках, які спираються переважно на психологію і педагогіку. Отже, специфічні для соціології проблеми розв'язують більшою мірою інтуїтивним шляхом, а це для методології емпіричних досліджень неприйнятно.

²³ Докладніше про методику Ш. Шварца див.: (Магун & Руднев, 2007).

²⁴ Докладніше про методику Р. Кесслера див.: (National Comorbidity Survey, 2018).

Глава 4

Теорія соціального вимірювання: припущення, латентні змінні та соціальні феномени

В основі теорії вимірювання, яку повсюдно використовують сьогодні в соціальних науках (неважливо, роблять це свідомо чи на рівні стереотипних методичних практик), лежить система понять, що описують вимірювальний процес, а також система відповідних наукових припущень. Дисциплінарним полем, яке їх породило, є психологія. Тому важливо оцінити, наскільки відповідні методологічні положення є доречними з соціологічної точки зору. Оскільки в центрі відповідної парадигми стоїть поняття «латентна змінна», то почнемо розгляд із неї (DeVellis, 2012, р. 17–30).

§ 4.1. Принципові основи класичної теорії вимірювання

Латентна змінна — досліджуваний феномен, або конструкт, який відображає вимірювальна шкала. Підкреслимо деякі особливості таких змінних. По-перше, латентні змінні неможливо спостерігати безпосередньо. По-друге, вони є не постійними, а змінними величинами. По-третє, це характеристики індивідів, а не об'єктів, яких оцінюють ці індивіди.

Оскільки латентну змінну неможливо виміряти безпосередньо, для визначення її справжньої величини на підставі значень, отриманих за допомоги індикаторів вимірювальної методики, застосовують низку методологічних і математичних припущень. Перш за все йдеться про припущення класичної теорії вимірювання, до яких належать такі:

- величина помилок²⁵, пов'язана з окремими пунктами методики, варіюється випадковим чином, що за використання великих вибірок веде до отримання незміщених середніх значень цих пунктів;
- помилки різних пунктів не корелюють одна з одною; єдиним джерелом, що пов'язує значення індикаторів методики, є латентна змінна;
- величина помилок не корелює з істинним значенням латентної змінної.

Крім припущень, наведених вище, класична теорія вимірювання базується на припущеннях паралельних тестів:

²⁵ Тут під помилками мається на увазі вплив другорядних факторів, які роблять свій внесок у значення того чи того індикатора нарівні з латентною змінною.

- сила впливу латентної змінної на кожен із пунктів методики є однаковою;
- кожен пункт методики має однакову величину помилок — у тому сенсі, що вплив факторів, відмінних від латентної змінної, є константою.

На підставі цих припущень сформульовано низку положень, що їх доповнюють, і необхідних для обґрунтування переходу від значень, отриманих за допомогою індикаторів вимірювальної шкали, до істинного значення латентної змінної.

За використання цього підходу стає зрозумілим, чому у психологічному тестуванні таку велику увагу приділяють надійності вимірювання — як ретестовій, так і внутрішній. За наявності однієї головної причини мінливості індикаторів, яка до того ж діє на них з однаковою силою, показники різних видів надійності мають бути високими.

Незважаючи на свою, на перший погляд, концептуальну і математичну стрункість, вимірювальний підхід, який базується на понятті латентної змінної, має обмежений потенціал для емпіричної соціології. Це обумовлено кількома суттєвими причинами.

§ 4.2. Класична теорія вимірювання vs прагматичний підхід

По-перше, поняття латентної змінної має стосунок виключно до феноменів психологічної природи. Тільки в цьому разі можна робити акцент на внутрішній властивості індивіда, записуючи решту факторів у «залишки», що взаємно «погашають» одно одне. Взагалі кажучи, положення класичної теорії вимірювання звучать так, ніби їх формували з есенціалістських позицій у трактуванні ідентичності. Не заперечуючи важливості внутрішніх усталених у часі якостей індивіда, сучасна соціологія, яка цікавиться динамічними соціальними процесами і відповідними ефектами серед населення, не може відштовхуватися у своїй дослідницькій діяльності виключно від усталених якостей індивіда, оскільки не залишиться жодних виходів на вимірювання соціальних феноменів, які розглядають не як зручні для соціологів абстракції, а як вираження тих явищ і процесів, які справді існують у досліджуваному суспільстві.

По-друге, самі припущення не виглядають переконливими з точки зору сучасної соціології, яка виходить із полідетермінованості досліджуваних явищ, а також необхідності вивчення таких феноменів, які можуть істотно змінюватися впродовж коротких проміжків часу (наприклад, настрої та думки в суспільстві). Тому в соціології не слід постулювати наявність єдиної головної причини, коли йдеться про той або той індикатор вимірювальної методики. Як мінімум необхідно визнати внесок зовнішніх соціальних факторів не як випадкових, а як системних. Інакше у соціологів просто не буде жодних шляхів для вивчення соціальних феноменів.

Тут саме час заторкнути питання про те, що автор розуміє під соціальними феноменами в контексті використання соціологічних тестів, чому слід удаватися до вивчення перших за допомогою других і яких припущень при цьому необхідно дотримуватися.

Відповідаючи на питання про сутність соціальних феноменів, вивчення яких можливе за допомоги соціологічних тестів, необхідно спиратися на особливості сучасної емпіричної соціології з тим, щоб у подальшому це поняття було придатним для застосування в рамках розгляду методів вимірювання. Що традиційно є предметом масових опитувань? Якщо відштовхуватися від положень відомої праці С. Садмена і Р. Бредберна (Садмен & Бредберн, 2002), до нього належать установки²⁶, поведінка і знання респондентів. Вважаємо, в цій тріаді найбільш релевантними, з точки зору виходу на правильне розуміння соціальних феноменів, є саме установки, що спрямовують поведінку індивідів²⁷ і забарвлюють їхні знання. При цьому установки стосовно повсюдно поширених суспільних явищ і процесів, що склалися історично, є нічим іншим, як інтеріоризованою частиною соціальних інститутів²⁸. В іншому разі говорити про соціальні інститути було б вельми проблематично. Це пов'язане з тим, що хоча респонденти і не утворюють механічну суму людей, які репрезентують ті або ті інститути, вони відіграють активну або пасивну роль в їх функціонуванні, що передбачає засвоєння відповідних соціальних норм і практик. Отже, не обмежуючи зміст поняття «соціальний феномен»²⁹ виключно соціальними інститутами, зазначимо, що в контексті використання соціологічних тестів предметом вимірювання останніх є соціальні феномени інституційної природи. Отримуючи інформацію про внутрішній бік соціальних інститутів, можемо робити наукові припущення і про ті зовнішні соціальні процеси і явища, що їх породжують і надалі супроводжують. Щоб зв'язок між використанням соціологічних тестів і вивченням соціальних інститутів виглядав обґрунтованішим, розглянемо питання, яке стосується методичних переваг саме такого соціологічного вимірювання.

Вивчення окремих установок ще не дає виходу на вимірювання соціального інституту, що цікавить дослідника, а лише його окремих аспектів. Тому таке вимірювання передбачає використання більш комплексних інструментів — соціологічних тестів. Окрім того, що вони дозволяють отримати надійну, ва-

²⁶ Під установками розуміємо оцінні реакції респондентів на будь-які об'єкти, що виявляються в їхніх думках, почуттях і цілеспрямованій поведінці (Майерс, 2003, с. 170).

²⁷ Автор усвідомлює всю складність взаємозв'язків між установками і поведінкою, а також проблематичність визначення того, що саме є первинним у цій діаді. Це фокусування використано як робочу гіпотезу.

²⁸ Під соціальними інститутами ми розуміємо комплексні соціальні форми, що склалися історично і репродукують себе (Miller, 2018). Згідно з докладнішим визначенням, соціальні інститути є комплексом позицій, ролей, норм і цінностей, що встановлені в тих чи тих типах соціальних структур і організують відносно стабільні моделі людської діяльності для розв'язання фундаментальних проблем спільного життя (Turner, 1997, р. 6).

²⁹ Говорячи загалом, наше розуміння соціальних феноменів базується на ідеях, висловлених Є. Головахою: соціальний феномен — це «...безпосередньо спостережуваний аспект соціальної дійсності. Вибір того чи того феномену як предмета дослідження пов'язаний із реальною можливістю зіткнення в соціальній дійсності з чимось, що викликає потребу в систематичному спостереженні» (Головаха, 2004, с. 12–13).

лідну і достовірну³⁰ інформацію про інституційну специфіку сучасного суспільства, завдяки своїй змістовній і структурній комплексності соціологічні тести створюють передумови для наукового вивчення складних соціальних явищ і процесів. У поєднанні з лонгітудним дизайном (наприклад, у рамках моніторингових соціологічних досліджень) їх використання дозволяє вивчати не тільки стан соціальної системи, але й її динамічні особливості.

Для розгляду методологічних припущень, що лежать в основі розробки і використання соціологічних тестів, перш за все необхідно визначитися з тими філософськими підставами, від яких нам слід відштовхуватися. Якщо звернутися до класичної теорії вимірювання, а також пізніших і менш суворих підходів, заснованих на понятті латентної змінної (наприклад, tau-equivalent tests, congeneric model, general factor model), перед нами постає суто позитивістська система ідей. У своєму розвитку сучасна емпірична соціологія прийшла до використання низки філософських концепцій, серед яких разом із позитивізмом можна назвати феноменологію, прагматизм, критичну теорію та інші. На наш погляд, за розробки соціологічних тестів необхідно відштовхуватися від прагматичних міркувань, згідно з якими нас перш за все цікавить методика, що дозволяє отримати актуальну інформацію щодо суспільних явищ, які підлягають вивченню. При цьому правильний шлях отримання цінної соціологічної інформації пов'язаний не зі слідуванням формальним і в дійсності недосяжним принципам класичної теорії вимірювання, а з демонстрацією методичних, концептуальних і статистичних особливостей використання вимірювального інструменту, що дозволяють або що не дозволяють аргументовано стверджувати його цінність з точки зору вивчення соціального феномену, який перебуває у фокусі наукового інтересу.

Отже, до припущень прагматичного підходу до соціологічного вимірювання (Дембицкий, 2017с, с.111–112) належать такі.

1. *У разі вимірювання, предметом якого є соціальні феномени, міркування про якість істинне значення, є досить проблематичними.* Це пов'язано з багатоаспектністю соціальних явищ — немає підстав вважати, що існує єдине і стабільне джерело установок, про які ми запитуємо респондента. Кількість таких факторів, а також сила їх впливу залежить від соціального контексту досліджуваного суспільства і тих актуальних процесів, які в ньому відбуваються.

2. *Величина отриманого значення за вимірювання соціального феномену обумовлена здебільшого двома основними джерелами — домінуючим дискурсом індивіда і навколишньою соціальною дійсністю.* Домінуючий дискурс ми бачимо як особливий спосіб спілкування і розуміння навколишнього світу (Філліпс & Йоргенсен, 2008, с. 18), який використовує респондент для оцінки соціального феномену; навколишню соціальну дійсність — як об'єктивовану частину соціальних інститутів, а також ті суспільні процеси, які мають до неї

³⁰ У рамках цієї роботи достовірність ми розуміємо як можливість, а також бажання респондента повідомити інформацію, яка цікавить дослідника (Шмелёв, 1996, с. 154–155).

стосунок, а по-іншому — відповідну поведінкову активність індивідів, що підлягає спостереженню.

3. *Як залишки або випадкову величину слід розглядати саме психологічні особливості індивіда, а по-іншому — набуті у процесі соціалізації усталені властивості особистості.* Власне, їх впливи на оцінку досліджуваного феномену (за розгляду агрегованого рівня) взаємно компенсуватимуть один одного за використання вибірок великого розміру.

Розглядаючи ці припущення, слід брати до уваги головну мету соціологічного вимірювання, яка полягає в побудові узагальненої картини соціальних явищ і процесів. Орієнтація на отримання істинних значень на рівні окремого респондента, як це роблять у класичній теорії вимірювання, в разі масових опитувань є недосяжним завданням, коли йдеться про вивчення соціальних феноменів за допомогою компактного соціологічного інструментарію. Тому слід докласти зусиль, щоб вимірювальний інструмент давав прийнятно точну інформацію, яка піддається осмисленій соціологічній інтерпретації. Це, своєю чергою, вимагає введення такого додаткового припущення.

4. *Задовільно точна і така, що піддається осмисленій соціологічній інтерпретації, інформація може бути отримана тільки за допомоги соціологічних тестів, які пройшли достатню теоретичну й емпіричну валідацію.*

Це положення має суто прагматичний характер, оскільки стверджує, що за використання методологічно обґрунтованих процедур і принципів забезпечення якості соціологічної інформації дослідник отримує можливість розробки вимірювального інструмента, необхідного для досягнення цілей соціологічного дослідження.

Отже, підсумкове значення методики як інструмента вимірювання — це не оцінка істинності, а формування загального уявлення про стан соціальної системи, а також уточнення наявних валідних теоретичних положень про різні аспекти соціального. При цьому окремі індикатори методики або їх підгрупи можуть нести самостійне евристичне значення в контексті проведеного аналізу.

Щоб чіткіше уявити методологічну специфіку соціологічних тестів, звернімося до процесу їх розробки.

Глава 5

Процес розробки соціологічних тестів

Тут ми торкнемося саме специфіки розробки соціологічних тестів, а не технік, які можуть при цьому застосовуватися. Це пов'язано з тим, що до таких технік належать численні процедури формулювання і відбору індикаторів, формування вибірки (починаючи від пілотних опитувань і закінчуючи репрезентативними дослідженнями з метою виокремлення норм), статистичної обробки результатів використання тесту на різних етапах його розробки. Опис усього вказаного передбачає написання окремої методичної роботи, що виходить за рамки цієї монографії. Крім того, багато з таких технік є специфічними, і аналіз ситуацій їх застосування не видається можливим³¹. Тобто далі ми розглянемо саме методологічний рівень розробки соціологічних тестів.

Комплексний підхід до процесу їх розробки включає чотири етапи: теоретичну валідацію, вибір різновиду соціологічного тесту, емпіричну валідацію, а також розробку норм або критеріальних рівнів тесту.

§ 5.1. Теоретична валідація соціологічного тесту

Щоб уникнути проблеми «дурного собаки, що ганяється за своїм хвостом» (де «дурний собака» — це теорія, яку мають намір верифікувати за допомоги вимірювальних процедур, а «хвіст» — вимірювальна методика, побудована на базі ще неперевіреної теорії), в основу соціологічного тесту мають бути покладені вже валідні теоретичні ідеї. Така постановка питання пов'язана з тим, що і «ламані» теоретичні конструкти можуть приводити до прийнятних результатів статистичного аналізу. З метою підтвердження цього положення нами було проведене методичне дослідження, в рамках якого порівнювалися результати використання методики, побудованої на підставі вже валідних теоретичних положень, з результатами тієї самої методики, яка була навмисно викривлена, щоб порушити її змістовну валідність (Дембицкий, 2014а). За допомоги двох цих варіантів вимірювальної методики викладачі та старости одного з київських ВНЗ оцінювали поведінкові стратегії студентів. Отримані результати були зіставлені з середніми балами студентів, які вони отримали на всіх сесіях бакалаврату

³¹ Докладніше про відповідні техніки див., напр.: (Анастаси & Урбина, 2007; Клайн, 1994; Kaplan & Saccuzzo, 2009).

(табл. 5.1). Як видно, що стосується статистики, обидва варіанти є рівнозначними, чого не можна сказати про теоретичні конструкти, які за ними стоять.

Таблиця 5.1

Сила і напрямок зв'язку результатів експертного оцінювання з середніми значеннями оцінок, отриманих на всіх сесіях бакалаврату ($N = 23$)

Експерти	Сила зв'язку	
	Використання «ламаной» шкали	Використання валідної шкали
Викладач № 1	0,52*	0,63**
Викладач № 2	0,80**	0,74**
Старости	0,75**	0,74**
Середнє значення трьох оцінок	0,80**	0,81**

* $p < 0,05$;

** $p < 0,01$.

Звичайно ж, з точки зору статистичного аналізу (наприклад, знаходження взаємозв'язків з іншими конструктами або передбачення значень інших змінних) обидва варіанти шкали можна оцінити як рівноцінні. Але з точки зору соціологічної інтерпретації, «ламана» шкала втрачає наукову цінність, оскільки відмова від змістовних елементів веде до втрати інформації про один тип студентів із трьох³².

Наступне питання стосується джерел тих теоретичних положень, які можна розглядати як валідні. Оскільки феномени, які потрапляють у фокус соціологічного вимірювання, можуть бути соціальними (наприклад, ставлення до владних інститутів), соціально-філософськими (наприклад, справедливість), соціально-психологічними (наприклад, психологічний клімат у колективі) і психологічними (наприклад, тривожність), остільки до джерел валідних теоретичних положень можуть бути віднесені окремі концепції в рамках соціальних наук, попередні спеціальні дослідження, професійний практичний досвід, а в певних випадках і універсальний досвід, тобто знання, властиві всім представникам певної соціальної спільності. Розглянемо їх докладніше.

Розглядаючи наукові концепції, слід говорити про ті з них, побудова яких спиралася на дослідницьку стратегію кейс-стаді³³, що передбачає неманіпулятивне і неконтролююче вивчення соціальних об'єктів і процесів такими, якими вони виявляються в соціальній дійсності³⁴. Завдяки цьому отримані теоре-

³² За допомоги решти шкали всі студенти можуть бути віднесені або до «Декларативно зацікавленого», або до «Відповідального» типу, натомість «Байдужий» тип випадає з фокуса дослідження.

³³ Докл. див., напр.: (Yip, 2009; Дембицкий, 2016с, с. 125–149).

³⁴ Під соціальною дійсністю ми розуміємо поведінкову активність індивідів на різних рівнях соціальної взаємодії.

тичні положення будуть не гіпотетичними конструкціями, а теоретично валідними моделями тих соціальних контекстів, які було охоплено дослідженням. У цьому сенсі використання кейс-стаді є прямим розв'язанням проблеми теоретичної валідності.

Ми відносимо до таких і деякі концепції історичної макросоціології, що використовують кейс-нарративний аналіз. До них, наприклад, належать результати, подані у працях Ч. Тіллі (Тіллі, 2009) і Б. Андерсона (Андерсон, 2001). Ці теорії хоча і виглядають вельми далекими від рівня соціологічного вимірювання, втім задають правильне фокусування щодо найбільш релевантних процесів у сьогоdnішньому світі національних держав, збільшуючи тим самим теоретичну чутливість дослідника.

Що стосується попередніх комплексних досліджень, то знову-таки йдеться про використання кейс-стаді з тих самих причин, які описані вище. Відмінність лише в тому, що тут дослідження «заточене» під побудову необхідної для подальшого соціологічного вимірювання теоретичної бази. Відповідним прикладом є наше дослідження поведінкових практик студентів, які суміщають роботу з навчанням (Дембіцький, 2016с, с. 99–124).

У разі професійного досвіду все значно простіше. Виконання щоденних функціональних обов'язків дозволяє фахівцям отримати розуміння різних аспектів соціальних феноменів природним шляхом. Тому такий досвід є потенційним джерелом валідних теоретичних положень. Прикладом використання професійного досвіду за побудови вимірювальної методики є шкала психологічного дистресу SCL-90-R Л. Дерогатиса (Derogatis, 2004), який відштовхувався від тих симптомів різних психологічних розладів, які були зафіксовані відповідним професійним співтовариством.

Нарешті, як відправну точку можна використовувати й аксіоматичні наукові положення, валідність яких не викликає сумнівів. Так, це можуть бути твердження, які розглядають як базисні у ключових соціологічних теоріях (функціоналізм, символічний інтеракціонізм, теорія конфлікту, феноменологія, теорія обміну тощо) або ж універсальні соціологічні аксіоми. Прикладом останніх є положення, використане Є. Головахою і Н. Паніною для конструювання інтегрального індексу соціального самопочуття: «...соціальне самопочуття людини визначають ступенем задоволення її соціальних потреб, які, своєю чергою, є похідними від існуючої в суспільстві системи соціальних благ, їх виробництва і розподілу» (Головаха & Панина, 1997, с. 23–24).

Розглянуті джерела можуть робити різний унесок у досягнення теоретичної валідності. Тому природно використовувати кілька з них, якщо це можливо, в рамках одного дослідницького проекту.

Теоретична база дослідження має забезпечувати три компоненти теоретичної валідності: очевидний, змістовний і композиційний. Очевидна валідність пов'язана з аргументацією, в рамках якої феномен, який вивчають, вписується в соціальну дійсність і, відповідно, не є плодом чиїхось домислів; змістовна валідність забезпечує значимість і показність елементів теоретичного конструкту; композиційна — правильність його внутрішньої струк-

тури³⁵. У такому вигляді теоретична база надає все необхідне для подальшої операціоналізації теоретичних положень, а також для оцінки отриманих у рамках емпіричної валідизації результатів. Але перш ніж перейти до останньої, необхідно визначитися зі специфікою вимірювального інструменту.

§ 5.2. Вибір різновиду соціологічного тесту

Після підготовки валідної теоретичної бази дослідник конструює на її основі шкалу, індекс або класифікатор. Оскільки в соціології широко використовують саме шкали та індекси, ми почнемо розгляд з них, а до класифікаторів звернемося наприкінці цієї частини роботи.

Приклад щодо оцінки індикаторів, спрямованих на вимірювання авторитаризму, показав, що визначення типу методики (шкали або індексу) може бути вельми нетривіальним завданням. Правильність такого вибору визначається перш за все соціологічними вподобаннями, а також ерудицією автора. У цьому сенсі визначальний фактор тут — особливості академічної соціалізації дослідника. Чим більше він замкнений в уявленнях традиційної емпіричної соціології (особливо кількісної), тим складніше йому буде вийти за обмеження, які накладають класична теорія вимірювання і суміжні з нею підходи. Навпаки, методологічне мислення з позиції неокласичної емпіричної соціології має відкривати розуміння відмінностей між латентними змінними і соціальними феноменами, що природним чином дозволить правильно зупиняти свій вибір на використанні або шкал, або індексів.

Вихідна точка міркувань полягає в тому, що слід розмежовувати індивідуальні психологічні особливості, а також той специфічний досвід, який формується у процесі взаємодії з соціальними інститутами. Вивчення перших можливе з застосуванням шкал (рух від загальної причини до відповідних проявів) та індексів (рух від різноманітних незалежних причин до відповідного результату). Що стосується другого, то на момент написання цієї частини роботи ми були ознайомлені лише з досвідом розробки відповідних індексів. Тому щоб не переривати логіку викладу, не зупинятимемося тут на шкалах, призначених для вимірювання соціальних феноменів. До них повернемося в четвертому розділі роботи вже після розгляду досвіду практичного використання прагматичного підходу до розробки соціологічних тестів у низці наших досліджень.

Натомість тут пропонуємо читачу розглянути як приклад чотири вимірювальні інструменти, які придатні для використання в соціальних дослідженнях, — особистісний опитувальник Айзенка, індекс психологічного дистресу SCL-9-NR, Інтегральний індекс соціального самопочуття, а також шкалу ситуативної тривожності Спілбергера.

Особистісний опитувальник Айзенка є класичною методикою визначення типу темпераменту, в основі якої лежить дві незалежні шкали, спрямовані на вимірювання екстраверсії-інтроверсії, а також нейротизму. Кожен з індикаторів, що входять в ту або ту шкалу, є непрямим індикатором відповідної осо-

³⁵ Докл. див.: (Дембицкий, 2016с, с. 22–31; Дембицкий, 2014б).

бистісної властивості (латентної змінної). Тобто як екстраверсія-інтроверсія, так і нейротизм є основними причинами, що визначають відповіді на індикатори методики. Тому не виникає жодних сумнівів, що йдеться саме про вимірювальні шкали.

У разі індексу психологічного дистресу SCL-9-NR використовують дев'ять індикаторів, кожен із яких подає свій симптоматичний вимір: ворожість, обсе-сивно-компульсивний розлад, депресію, міжособистісну сензитивність, параноїдальні ідеї, тривожність, соматизацію і психотизм. Саме схильність до таких станів призводить до психологічного дистресу, а не навпаки. Їх можна розглядати як специфічну реакцію на різні стресори, з якими стикається індивід. У цьому сенсі вони є незалежними один від одного джерелами психологічного дистресу. Відповідно, SCL-9-NR — це індекс, а не шкала. При цьому треба розуміти, що психологічний дистрес — це не властивість особистості (латентна змінна), а стан психічної системи індивіда (внутрішній стан)³⁶.

Інтегральний індекс соціального самопочуття являє собою методику, спрямовану на вимірювання соціального феномену. Це твердження ґрунтується на тому, що особливості відповідей на її індикатори визначаються насамперед інституційною специфікою досліджуваного суспільства. Пункти індексу розбиті на 11 відносно автономних сфер — соціальних відносин, соціальної безпеки, національних відносин і т. д. Загальна логіка оцінки соціального самопочуття така, що чим краще задоволені потреби в кожній із сфер, тим вище соціальне самопочуття. А отже, як і в попередньому прикладі, йтися має про індекс, а не шкалу.

Добрим прикладом, що демонструє відмінність між шкалою та індексом, є такий вимірювальний інструмент, як шкала ситуативної тривожності Спілбергера. Традиційно тривожність розглядають із двох точок зору — як особистісну властивість (латентна змінна) і як стан психіки (внутрішній стан). З цієї причини шкала Спілбергера має дві форми, кожна з яких складається з 20 тверджень. Перша спрямована на вимірювання особистісної тривожності, друга — ситуативної. Очевидно, що особистісна тривожність має збільшувати чутливість до факторів, що спричиняють тривожність ситуативну. Разом з тим шкалу ситуативної тривожності слід розглядати як самостійний інструмент (у будь-якому разі в його результатах буде побічно відбиватися й особистісна тривожність). Понад те, цю частину методики Спілбергера можна розглядати не як шкалу, а як індекс.

Річ у тім, що смисловий зміст низки її індикаторів прямо вказує на це — «Вам ніщо не загрожує», «Ви відчуваєте жаль», «Ви відчуваєте себе вільно», «Ви відчуваєте себе людиною, яка відпочила», «Ви стривожені» і т. д. Індикатори сформульовані так, що згода з ними може ґрунтуватися як на внутрішніх особливостях респондента (щось постійне), так і на об'єктивно існуючих при-

³⁶ Незважаючи на їх очевидну відмінність, у класичній теорії вимірювання різниця між латентними змінними і внутрішніми станами не концептуалізується. Замість цього все належить до категорії латентних змінних.

чинах. Можна відчувати загрозу за її наявності, відчувати жаль після неправильного вчинку, відчувати себе невільно в певних обмеженнях, бути втомленим через поганий робочий графік і стривоженим через очікувані життєві складнощі. Тобто причини цих станів різні і можуть не мати жодного стосунку до латентної змінної. А якщо немає вагомих причин вважати джерелом підсумкового вимірювання виключно особистісну чутливість, методику слід віднести до індексів.

Отже, «гонитва» за вимірювальними шкалами — не той методологічний імператив, якого слід неухильно дотримуватися. Навпаки, треба проявляти методологічну гнучкість і в разі необхідності вдаватися до використання інтегральних індексів.

Разом з тим зміст комплексних вимірювальних інструментів у соціології не вичерпується шкалами та індексами. Важливе місце в цій низці посідають класифікатори. Це твердження ґрунтується на сучасній практиці класового аналізу в соціології, в рамках якої дослідники вдаються до побудови різних класових схем³⁷. Розташування респондентів у них, серед іншого, визначають низкою міжнародних класифікаторів, особливе місце серед яких посідає класифікатор професій ISCO-88³⁸. На відміну від шкал та індексів, одержувані в рамках масових опитувань результати тут використовують не для вимірювання виразності певної властивості, а для розподілу респондентів у рамках якісних категорій, що не пересікаються. Такий підхід легко застосовний у рамках соціологічного тестування, де класифікатори можна використовувати як окремі вимірювальні інструменти, так і спільно зі шкалами або індексами. В останньому випадку дослідник отримує два види інформації про досліджуване явище — якісну (класифікатор) і кількісну (шкала або індекс).

Перевірку правильності переходу від теоретично валідної бази до інструмента вимірювання здійснюють за допомоги емпіричної валідації, що спирається на результати методичних опитувань.

§ 5.3. Емпірична валідація соціологічного тесту

Емпірична валідність має три незалежні й один залежний компонент. До незалежних належать такі види валідності, як конструктна, критеріальна і прогностична, до залежного — порівняльна валідність. Ці чотири види слід розглядати як основні в контексті емпіричної валідації соціологічних тестів. Окрім них, у конкретних випадках може знадобитися перевірка додаткових (або неосновних) видів емпіричної валідності. Наприклад, популярна сьогодні факторна валідність (factorial validity) належить до цієї категорії. Тут же ми розглянемо тільки основні види емпіричної валідності.

Конструктну валідність ми розглядаємо в термінах, запропонованих у праці Д. Кемпбелла і Д. Фіске (Campbell & Fiske, 1959). Так, для її підтвердження необхідно продемонструвати конвергентну і дискримінантну складові. Перша

³⁷ Докл. див.: (Симончук, 2007).

³⁸ Докл. див.: (Симончук, 2008).

ґрунтується на тому, що результати методики пов'язані з результатами тих шкал або індексів, які спрямовані на вимірювання явищ, теоретично пов'язаних із феноменом, що перебуває у фокусі нашого інтересу. Говорячи простіше, вимірювання конструктів, які пов'язані в теорії, мають бути пов'язані і фактично. У разі дискримінантної валідності логіка та сама, тільки до уваги беруть аспекти, які теоретично не пов'язані. Відповідні вимірювання мають це продемонструвати. Якщо говорити про статистичні процедури, то тут найприйнятнішим є використання кореляційного (в широкому сенсі) аналізу.

Критеріальна валідність належить до визначення ступеня, в якому результати вимірювання пов'язані з якимось об'єктивним показником. Зазвичай її поділяють на конкурентну і прогностичну (останню часто називають предективною) валідності. Однак із міркувань, які буде наведено далі, під критеріальною валідністю вимірювальної методики ми розуміємо виключно конкурентну, а прогностичну розглядаємо як окремий вид. У цьому сенсі критеріальна валідність перевіряється шляхом порівняння результатів вимірювання та інших спостережуваних результатів, які є об'єктивними показниками досліджуваного феномену. Одним із простих, але переконливих способів демонстрації критеріальної валідності соціологічного тесту є отримання очікуваних відмінностей за опитування «контрастних» з точки зору досліджуваного явища груп, що серед іншого вказує і на загальну «чутливість» методики.

Прогностична валідність пов'язана з можливістю інструмента надавати інформацію, необхідну для прогнозування майбутніх станів об'єкта дослідження. З погляду соціологічного вимірювання це, можливо, найбільш важкодосяжний вид валідності — як через необхідність стислості інструментарію, так і через комплексність досліджуваних у соціології процесів і явищ. Саме тому ми проводимо сутнісне розмежування між критеріальною і прогностичною валідністю. Загалом результати соціологічного вимірювання (якщо ми говоримо про конкретну методику) — лишень одне з необхідних джерел емпіричної інформації, на підставі якої будується соціологічний прогноз. Тому автор методики повинен розуміти, як саме результати його соціологічного тесту вбудовуються в ту чи ту галузь соціології з тим, щоб поліпшити висновки про можливі сценарії розвитку соціальної дійсності.

Порівняльну валідність розуміють як порівняльну перевагу вимірювального інструменту в низці схожих методик. Отже, її обґрунтування пов'язане з необхідністю порівняння показників конструктної, критеріальної та прогностичної валідності різних соціологічних методик, спрямованих на вивчення одного і того самого явища. Тому цей різновид валідності є залежним компонентом емпіричної валідизації — результати її перевірки пов'язані як із попередніми результатами емпіричної валідизації, так і з тією методикою (і результатами її валідизації), яка виступає в ролі головного «конкурента».

§ 5.4. Розробка норм або критеріальних рівнів соціологічного тесту

Однією з принципів умов, яку часто висувають у процесі обґрунтування релевантності тестової методики, є виокремлення її норм. З огляду на це розроб-

ка критеріальних рівнів є принципово важливим етапом розробки соціологічного тесту.

Методологічні напрацювання щодо норм у соціологічному тестуванні відсутні. Отже, як початковий орієнтир тут можна використати положення, які використовують за визначення норм психологічних тестів.

Поняття «норма» стосовно психологічного тестування має стосунок до результатів, які демонструють представники певної групи щодо того або того тесту (Kaplan & Saccuzzo 2009, р. 53). Такі норми обчислюють за допомоги досліджень стандартизації й далі використовують для оцінки індивідуальних результатів психологічного тесту.

Дослідження стандартизації полягають у вивченні репрезентативних груп із 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 (Анастасі & Урбіна, 2007, с. 64–65).

Незважаючи на те, що цей підхід є загальноприйнятим, деякі його аспекти вимагають докладнішого розгляду. Насамперед це стосується поняття «репрезентативність». Так, П. Клайн говорить про необхідність використання стратифікованої вибірки, що дозволяє врахувати в дослідженні стандартизації найважливіші параметри, які впливають на результати застосування тесту. Як приклад він наводить своє дослідження, де учні шостих класів були стратифіковані за статтю, специфікою фінансування школи, її географічним положенням і статусом. Для кожної категорії було відібрано по дві школи. При цьому всі школи були розташовані на півночі Англії. У фінальну вибірку увійшло по 1000 хлопчиків і дівчат (Клайн, 1994, с. 219). Отже, йдеться не про статистичне узагальнення (застосовано випадкову вибірку необхідного розміру), а про аналітичне (застосовано вивчення релевантних соціальних контекстів)³⁹.

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'язане з тим, що головна мета такого дослідження — отримання опису властивості, яку вивчають у різних контекстах, а не опис генеральної сукупності. 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, у яких проводяться узагальнення, навпаки, вони завжди були гіпотетичними, навіть якщо і керувалися раніше відкритими законами» (Кемпбелл, 1980, с. 207–208).

Наступне питання стосується перетворення, або статистичної нормалізації «сирих» балів тесту, що дозволяє ефективно розв'язувати вже згадані завдання — порівняння з результатами вибірки стандартизації, а також із результатами інших тестів. При цьому актуальним завданням у цьому разі є саме порів-

³⁹ Докладніше про відмінність між статистичною й аналітичною генералізацією див.: (Дембицкий, 2016с, с. 69–70).

няння 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, за наближення розподілу первинних показників до нормального стандартні показники, отримані за допомоги лінійного перетворення, практично не відрізняються. Відповідно, статистична нормалізація тут мало що або взагалі нічого не змінить. Необхідно домагатися нормального розподілу первинних результатів використання тесту, а не нормалізувати статистичними засобами ненормальний розподіл (Анастасі & Урбина, 2007, с. 80).

Останнє, що необхідно взяти до уваги, — це визначення рівнів виразності досліджуваної властивості на підставі результатів використання певного тесту. Дослідження стандартизації саме по собі нічого не говорить про такі рівні. По суті, його проводять з іншими цілями. Єдине рішення цієї проблеми, на яке можна потрапити в методичній літературі (воно напрошується природним чином), полягає у використанні критеріальних норм замість статистичних. У цьому разі шкала «сирих» тестових балів калібрується рівнями, які визначають імовірність досягнення певного критерію або ступеня, які дозволяють припустити якісно інший стан учасників дослідження за досліджуваною характеристикою (Шмелёв, 1996, с. 153–154). Утім, про те, як визначати критеріальні норми, не йдеться.

Тепер критично оцінимо три зазначені поняття з точки зору специфіки соціологічного дослідження. Повертаючись до розмежування статистичної й аналітичної репрезентативності, вважаємо обґрунтованим використання останньої так само, як це рекомендують психологічні дослідження. Це пов'язано з тим, що для сучасних суспільств є характерною досить висока гетерогенність. Відповідно, важливо розуміти, які результати соціологічний тест продемонструє за використання в різних соціальних контекстах (різних аудиторіях). При цьому апробація тесту у статистично репрезентативному опитуванні також є бажаним 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 роблять на узагальненні отриманих результатів, а не на оцінці індивідуальних показників, щоб визначити їх місце щодо показників генеральної сукупності (таке завдання розв'язують, 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 в історичних особливостях певних спільнот. По-третє, на думку П. Клайна, коли йдеться про наукове дослідження, статистичні норми є не настільки важливими (Клайн, 1994, с. 214). Крім того, є методичний аргумент на користь низької релевантності нормалізації результатів. З огляду на вимогу компактності соціологічних тестів інтервали їх можливих значень часто будуть доволі «вузькими» порівняно з багатьма психологічними тестами. А це вно-

силь свої обмеження за побудови досить точних інтервалів із використанням статистичної нормалізації.

Що стосується критеріальних рівнів, то це питання є одним із найважливіших за розробки соціологічного тесту. Це пов'язано з розв'язанням двох взаємопов'язаних завдань соціологічного тестування. Перше — продемонструвати, що тест має диференціальний потенціал. Наприклад, якщо показники використання певного тесту авторитаризму у групі звичайних респондентів мало відрізняються від відповідних характеристик крайніх анархістів і крайніх націоналістів, то критеріальних рівнів тест не має (як і не характеризує його критеріальна валідність). Отже, для визначення критеріальних рівнів соціологічного тесту його автор повинен бути обізнаний про існування в суспільстві таких груп, для яких характерний низький і/або високий рівень досліджуваної властивості, а також мати доступ до цих груп. Це дозволить здійснити критеріальну нормалізацію тесту, що надасть можливість розв'язання другого завдання — виокремлення рівнів у результатах тестування, які характеризують респондентів з точки зору не тільки кількісних, але й якісних відмінностей.

Отже, поняття «норма» в соціологічному тестуванні недостатньо точно відбиває суть відповідної методичної діяльності. Вдалішим терміном тут, на нашу думку, є поняття «критеріальний рівень», якщо розуміти 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 означає рівні виразності вимірюваної властивості; 3) дискримінативність як поділовий потенціал тесту.

Глава 6

Приклади соціологічних тестів: соціальне самопочуття, політична культура, психологічний дистрес

Далі коротко розглянуто три соціологічні тести, конструювання яких так чи так пов'язане з іменами відомих київських соціологів Є. Головаха і Н. Паніної. Для кожного тесту стисло подано його призначення, особливості конструювання, статистичної перевірки і валідазації, а також соціологічна специфіка в розрізі трьох описаних раніше вимірів соціологічного тестування (застосовність у масових опитуваннях, спрямованість на вивчення соціальних феноменів, діагностичний потенціал з точки зору аналізу соціальної системи).

§ 6.1. Соціологічний тест «Інтегральний індекс соціального самопочуття»

Інтегральний індекс соціального самопочуття (далі — ІСС; автори — Є. Головаха, Н. Паніна⁴⁰) являє собою стандартизовану тестову процедуру, що була розроблена з огляду на необхідність оцінювання масових емоційно-оцінних станів у постійно змінюваних умовах нестабільного перехідного суспільства (Головаха & Панина, 1997, с. 11). Соціальне самопочуття автори методики розуміють як «емоційно-оцінне ставлення індивіда до системи соціальних відносин і до свого місця в цій системі» (Головаха & Панина, 1997, с. 20).

За конструювання тесту було виокремлено 11 сфер соціальної життєдіяльності індивіда — соціальних відносин, соціальної безпеки, національних відносин, соціально-політична, професійно-трудова, інформаційно-культурна, рекреаційно-культурна, матеріально-побутова (першого і другого рівня), міжособистісних відносин, особистісна. Усередині кожної сфери вирізнено по чотири індикатори, що дало в результаті 44 пункти методики, які описують соціальні блага. Така деталізація забезпечила змістову валідність, обґрунтування якої було здійснене на логічному рівні.

⁴⁰ Статистичний аналіз у процесі розробки тесту здійснено А. Горбачиком.

В якості варіантів відповіді була запропонована «шкала достатності»: *не вистачає; важко сказати, вистачає чи ні; вистачає; не цікавить*. З метою статистичного опрацювання було ухвалено прирівняти четвертий пункт шкали (не цікавить) до другого (важко сказати, вистачає чи ні) (Головаха & Панина, 1997, с. 21–24).

На підставі результатів опитування ($N = 1810$), репрезентативного для дорослого населення України (віком від 18 років), проведено оцінювання валідності та надійності методики загалом, проаналізовано використання кожного з запропонованих у початковому варіанті індикаторів соціального самопочуття, побудовано скорочений варіант тесту, здійснено стандартизацію і нормалізацію шкали (Головаха & Панина, 1997, с. 27–28).

На основі результатів кореляційного аналізу кожного пункту тесту з сумарним індексом оцінювання впливу виключення окремих пунктів на внутрішню узгодженість і перевірки нормальності розподілу сумарного індексу всі 44 пункти методики були включені до фінального варіанта тесту. Успішну перевірку пройшли такі види надійності, як розщеплення тесту навпіл, ретестова і внутрішня узгодженість.

Також перевірено конструктну валідність на підставі зіставлення сумарного індексу для різних категорій респондентів, виокремлених залежно від відповідей на запитання про різні аспекти життєвої задоволеності (Головаха & Панина, 1997, с. 40–43).

Скорочений варіант (20 із 44 пунктів) побудовано на базі формальних (збереження внутрішньої узгодженості) і змістових засад (зберігання в основних рисах початкової багатоаспектності вимірювань). У результаті отримано вимірювальний інструмент, що відповідає стандартам надійності та валідності (Головаха & Панина, 1997, с. 46–47).

Проаналізуємо соціологічну специфіку тесту. ПСС розробляли й оцінювали передусім як інструмент соціологічного вимірювання для масових опитувань. Тому він задовольняє принципіву вимогу соціологічного тестування — застосовності в масових опитуваннях. Сам феномен соціального самопочуття включає як психологічні, так і соціальні компоненти, зі значним переважанням останніх. Отже, показники ПСС залежать насамперед від специфіки суспільних процесів, пов'язаних із задоволенням потреб індивідів. З одного боку, це є аргументом на користь діагностичного потенціалу тесту за аналізу соціальної системи нашого суспільства. Водночас залишається низка нез'ясованих питань, серед яких, наприклад, урахування зміни актуальності різноманітних потреб у контексті масштабніших соціальних процесів (так, війна на Сході України мала позначитися на актуальності сфер соціальної безпеки і національних відносин) і розробки критеріїв оцінювання успішності різних соціальних інститутів, відповідальних за надання основних соціальних благ у нашому суспільстві (хоча потенціал тесту повною мірою дає підстави сподіватися на реалізацію цих завдань).

Узагальнюючи подану інформацію, доходимо висновку, що ПСС є повноцінним соціологічним тестом, який відповідає високим стандартам щодо розробки вимірювальних інструментів у соціальних науках.

§ 6.2. Соціологічний тест «Типи політичної культури»

Соціологічний тест «Типи політичної культури» (далі — СТ ТПК; автор — Є. Головаха⁴¹) є інтегральним тестом, призначеним для визначення типу політичної культури респондента. При цьому під політичною культурою розуміємо «сукупність усталених форм політичної свідомості і поведінки людей, що визначають особливості функціонування різних суб'єктів політичного процесу в рамках певної політичної системи» (Шербак, 2006, с. 155).

За деталізації теоретичного конструкта було виокремлено два виміри політичної культури: демократія/тоталітаризм, активність/пасивність (у сенсі громадської участі). Для кожного з вимірів сформульовано по 50 тверджень, при цьому для збалансованості розроблено по 25 тверджень для демократії та тоталітаризму, а також для активності та пасивності.

Для реєстрації відповідей використано шкалу Лайкерта: «цілком згоден», «скоріше згоден», «важко сказати, згоден чи ні», «скоріше не згоден», «абсолютно не згоден».

З метою скорочення методики до прийнятної для соціологічних опитувань розміру, перевірки конструктивної валідності та ретестової надійності було проведено пілотне опитування респондентів віком від 18 років ($N = 130$). Для оцінювання ретестової надійності кожного пункту використано як кореляційний аналіз (надійним вважався пункт, що показував силу зв'язку не менше 0,3), так і розрахунок шкальних оцінок за формулою середньої арифметичної похибки (не більш як 1).

Далі на підставі факторного аналізу було визначено судження, що дають однофакторне рішення на шкалах демократія/тоталітаризм і активність/пасивність. Також було оцінено внутрішню надійність (Альфа Кронбаха) і силу зв'язку між окремими індикаторами й адитивним індексом. З огляду на всі отримані результати з вихідного набору суджень було відібрано по вісім для кожної зі шкал.

Цей проміжний варіант був включений в опитування в рамках Європейського соціального дослідження. Подальшу апробацію методики також здійснювали з використанням перевірки внутрішньої узгодженості та факторної структури. Для обох шкал отримано двофакторні рішення з поділом на шість і два індикатори. Після видалення чотирьох тверджень, що не входили до числа основних факторів, отримано фінальний варіант тесту, що включав по шість суджень для шкал демократії/авторитаризму й активності/пасивності.

Фінальний варіант СТ ТПК використано для вирізнення типів політичної культури населення України з застосуванням кластерного аналізу методом *k-середніх*. Виходячи з теоретичних очікувань наявності чотирьох типів політичної культури (демократична активна, демократична пасивна, авторитарна активна, авторитарна пасивна), було побудовано чотири кластери, що підтвердили попередні теоретичні припущення (Стегній, 2015, с. 29–32).

⁴¹ Статистичний аналіз у процесі розробки тесту здійснено А. Горбачиком.

Розглянемо тест з точки зору соціології. Як і в попередньому випадку, СТ ТПК був спроектований і реалізований як соціологічний тест. У плані стислості він навіть перевершує ПСС, що забезпечує його відповідність першому виміру соціологічного тестування. Дві концептуальні дихотомії, що лежать в основі цього інструмента, є цілком соціальними, що підтверджує його спрямованість на вивчення саме соціальних феноменів. Немає також жодних сумнівів, що СТ ТПК вимірює результативні ефекти інституційної динаміки і демонструє потенціал демократичних процесів у суспільстві. Відповідно, його результати є найвищою мірою застосовними для оцінювання стану соціальної системи.

При цьому наголосимо один аспект, що викликає питання. Так, усі шість суджень для шкали демократія/авторитаризм сформульовані таким чином, що стосуються саме демократичних установок. Отже, питання, якою мірою за ними можна судити про авторитарні установки, залишається відкритим. Особливо актуальним це питання є з огляду на кризові процеси в сучасній Україні, що збільшують імовірність розвитку авторитарних поглядів серед населення країни.

Утім, тест СТ ТПК є високоякісним соціологічним інструментом, що базується на теоретичних положеннях, які не підлягають сумніву, й емпіричних індикаторах, які пройшли багаторівневу процедуру добору, що робить його постійне використання в соціологічних опитуваннях по-справжньому актуальним.

§ 6.3. Соціологічний тест SCL-9-NR

SCL-9-NR (автор — С. Дембіцький) є модифікацією відомого психологічного тесту SCL-90-R⁴², розробленого Л. Дерогатісом (Derogatis & Fitzpatrick, 2004). Відмітними рисами SCL-9-NR порівняно з SCL-90-R є його стислість (для кожного з дев'яти симптоматичних вимірів залишено тільки один індикатор), а також принципові зміни в завданні методики і шкалі відповідей (але не у формулюваннях самих індикаторів). Головне призначення тесту — оцінювання психологічного дистресу респондентів у рамках масових опитувань.

Оригінальна методика, запропонована Л. Дерогатісом, базується на таких симптоматичних вимірах, як ворожість, obsесивно-компульсивний розлад, депресія, міжособистісна сензитивність, параноїдальні ідеї, тривожність, соматизація, фобійна тривожність, психотизм, кожен із яких містить набір індикаторів (від шести до тринадцяти пунктів). За добору індикаторів для кожного симптоматичного виміру було враховано силу його зв'язку з індексом своєї підшкали, а також з інтегральним індексом SCL-90-N (Дембицкий, 2016е, с. 54–55).

⁴² З метою адаптації методики в Україні SCL-90-R був включений Н. Паніною в моніторингові опитування Інституту соціології НАН України в 1997 і 1999 роках, а також О. Балакіревою до всеукраїнського опитування спільного моніторингу Українського інституту соціальних досліджень імені О. Яременка, Центру «Соціальний моніторинг» і відділу моніторингових досліджень соціально-економічних трансформацій Інституту економіки і прогнозування НАН України в 2014 році.

Зміну формулювання тестового завдання і, відповідно, самої шкали відповідей⁴³ здійснено на підставі рекомендацій, наведених у відомій праці С. Садмена і Н. Бредберна з розробки соціологічного інструментарію (Садмен & Бредберн, 2002, с. 65–96). Головною причиною такого коригування був передбачуваний зсув у відповідях респондентів, що виникає за використання оригінального варіанта формулювання, — респонденти були схильні обирати менші значення на шкалі відповідей, що зменшувало і фінальний рівень індексу психологічного дистресу. Як показала подальша перевірка, таке припущення виявилось слушним і використання нових формулювань суттєво поліпшило фінальний розподіл індексу (Дембицкий, 2016е, с. 59–61).

На підставі онлайн-опитувань у Києві ($N = 200$) і Львові ($N = 200$) була перевірена внутрішня надійність і факторна валідність методики (Дембицкий, 2016е). У подальшому, оскільки тест спрямований на вивчення психологічного дистресу, було проведено опитування в Чернігівській обласній психоневрологічній лікарні ($N = 181$) із метою перевірки критеріальної валідності та впливу ефекту соціальної бажаності. Для оцінювання останнього на звичайній вибірці спільно з кафедрою методології та методів соціологічних досліджень КНУ ім. Тараса Шевченка було проведено онлайн-опитування серед користувачів соціальних мереж ($N = 367$). Отримані результати підтвердили як критеріальну валідність тесту, так і прийнятний рівень впливу ефекту соціальної бажаності на його результати. Крім цього, на підставі опитування у психоневрологічній лікарні виокремлено значення тесту, які характеризують підвищений і високий дистрес (Дембицкий, Ященко, Сидоров, & Александрова, 2016).

З метою перевірки конструктивної та критеріальної валідності, а також отримання додаткової інформації щодо критеріальних рівнів тесту на підставі результатів репрезентативного для дорослого населення України опитування 2016 року SCL-9-NR було включено в моніторингове дослідження «Українське суспільство»⁴⁴.

Оцінюючи методику як соціологічний тест, слід визнавати, що як таку її можна розглядати суто в контексті першого виміру соціологічного тестування — його застосовності в соціологічних опитуваннях.

Сам психологічний дистрес не можна назвати соціальним феноменом, проте він пов'язаний із низкою соціальних характеристик (Pearlin, Menaghan, Lieberman, & Mullan, 1981), що потенційно робить SCL-9-NR інструментом аналізу соціальної системи. Для перевірки цього припущення здійснюватиметься додаткова дослідницька робота⁴⁵.

⁴³ Використано шкалу: «ніколи», «зрідка», «періодично», «майже постійно».

⁴⁴ Організатором дослідження є Інститут соціології НАН України. Результати опитування репрезентативні для населення України старшого 18 років (вибірка — не менше 1800 осіб). Анкета опитування містить понад 300 питань і охоплює різні аспекти життя українського суспільства.

⁴⁵ На момент завершення цієї монографії її автор був керівником науково-дослідного проєкту «Психологічний дистрес в українському суспільстві: фактори і поширеність». При цьому психологічний дистрес вимірювався за допомоги SCL-9-NR.

Висновки до першого розділу

У розділі продемонстровано, що предмет комплексного соціологічного вимірювання — це не тільки латентні змінні, а і внутрішні стани і соціальні феномени. Це, своєю чергою, висуває завдання переходу від індивідуальних результатів, отриманих у рамках масових опитувань, до аналізу специфіки досліджуваного суспільства. Ані психологія, ані педагогіка не оперують методологічними засобами, необхідними для її розв'язання. Водночас інтерес соціологічної науки, а отже, і її емпіричних досліджень багато в чому зосереджений на масштабних соціальних явищах і процесах.

Таким чином, загальноприйняті методологічні принципи вимірювання в соціальних науках насправді відірвані від широкої соціологічної практики. У цьому контексті наше звернення до соціологічних тестів є частковим⁴⁶ розв'язанням зазначеної проблеми. Воно засноване на таких теоретико-методологічних передумовах — проведення відмінності між шкалами, індексами і класифікаторами як різновидами соціологічного тестування, прийняття прагматичних методологічних припущень щодо соціологічного вимірювання, а також правильне розуміння статусу теоретичного й емпіричного етапів валідації⁴⁷.

На нашу думку, подальша проблематизація вимірювальних практик у соціології, які засновані на зовнішніх дисциплінарних традиціях, сприятиме продуктивній методологічній рефлексії, поліпшенню стандартів розробки методик і отриманню точніших результатів за проведення масових опитувань.

Говорячи окремо про соціологічні тести, необхідно категоризувати їх як вимірювальні інструменти традиційної емпіричної соціології. Незважаючи на це, їх правильне розуміння, а також вдала розробка багато в чому залежать від ступеня спрямованості, а також прихильності дослідника до неокласичної емпіричної соціології, що є історичним ядром соціологічного підходу до вивчення соціальної дійсності.

Що стосується переходу від загальних методологічних положень, яким присвячений цей розділ, до конкретніших рекомендацій, то він потребує накопичення методичного досвіду розробки соціологічних тестів. Два наступні розділи присвячені результатам відповідних досліджень, проведених автором цієї роботи. Перший (Розділ II) фокусується на феномені індивідуального благополуччя, другий (Розділ III) — на феномені політичної культури. Отриманий методичний досвід підсумовано в останньому розділі (Розділ IV).

⁴⁶ Частковим з тієї причини, що йдеться не про всі соціальні феномени, а лише про ті з них, що мають інституційну специфіку.

⁴⁷ Слід згадати про типологію валідності в соціологічному дослідженні, що була запропонована набагато раніше у вітчизняній соціології й також оперує категоріями теоретичної та емпіричної валідності. Вона належить відомому вітчизняному соціологу Володимирі Паниотто. Так, до теоретичної валідності він відносить змістовну й очевидну, але залишає проблему класифікації способів теоретичного обґрунтування методик як нерозв'язану (Паниотто, 1986, с. 110). Водночас В. Паниотто пропонує деталізоване бачення емпіричної валідації, що спирається на валідність за критерієм: (Паниотто, 1986, с. 111–118).

Розділ II

ІНДИВІДУАЛЬНЕ БЛАГОПОЛУЧЧЯ ЯК ОБ'ЄКТ СОЦІОЛОГІЧНОГО ТЕСТУВАННЯ

Глава 7

До розуміння феномену індивідуального благополуччя

Ми вважаємо термін «індивідуальне благополуччя» найудалішим з точки зору узагальнення характеристик, що описують низку аспектів якості життя індивіда. При цьому індивіда розуміємо як окремого представника тієї чи тієї соціальної спільності (Бабосов, 2009, с. 145). У цьому сенсі термін «індивідуальне благополуччя» наділений соціологічною специфікою, оскільки нерозривно пов'язує індивідуальне життя з ширшим соціальним контекстом.

Перш ніж перейти до опису найважливіших, на наш погляд, складових індивідуального благополуччя, звернемося до вживання цього терміна в сучасній соціології. З цією метою ми відібрали і проаналізували англomовні праці, опубліковані в наукових журналах видавництва «Sage» з початку 2000-х років, назви яких містили словосполучення «індивідуальне благополуччя»⁴⁸. Як виявилось, таких публікацій лише 12⁴⁹. Сім із них присвячені індивідуальному благополуччю в рамках партнерських відносин сімейного типу (може йтися як про офіційно зареєстровані, так і не зареєстровані стосунки), одна стаття є оглядовою — в ній розкриваються різні складові індивідуального благополуччя, інша стаття розглядає індивідуальне благополуччя в контексті територіальних суперечок і ще одна — благополуччя в контексті розвитку і якості життя в Туреччині. Проаналізуємо їх докладніше з точки зору того, як автори статей оцінювали індивідуальне благополуччя у своїх дослідженнях.

У статті, присвяченій взаємозв'язку розв'язання конфліктів, сімейних взаємин підлітків та індивідуального благополуччя (Tucker, McHale, & Crouter, 2003), останнє фіксовано за допомоги оцінки ризикованої поведінки (Risky Behavior Scale), самооцінки (Competence Scale for Adolescents) і депресивних сипптомів (CES-D). Лонгїтюдне дослідження взаємозв'язку сімейних взаємин та індивідуального благополуччя (Stafford, Kline, & Rankin, 2004) концентрується на вимірюванні трьох показників психологічного благополуччя: гло-

⁴⁸ В оригіналі — «individual well-being» (не «well-being», «personal well-being», «social well-being», «subjective well-being» тощо).

⁴⁹ Пошук здійснено в базах даних відомих наукових видань, що містяться в мережі інтернет.

бального щастя (на підставі одного індикатора), депресії (CES-D) і самооцінки (Rosenberg's self-esteem scale). У дослідженні взаємозв'язку сімейних відносин та індивідуального благополуччя в середині життя (Hansen, Moum, & Shapiro, 2007) оцінюється задоволеність життям (Satisfaction With Life Scale). Стаття, присвячена аналізу взаємозв'язку між причинами спільного життя, індивідуальним благополуччям і якістю відносин (Rhoades, Stanley, & Markman, 2008), містить результати вимірювання депресивної та тривожної симптоматики (Brief Symptom Inventory). У дослідженні впливу структурного розширення можливостей персоналу на його індивідуальне благополуччя і продуктивність (Biron & Vamberger, 2010) вимірюється емоційне вигорання (Maslach's and Jackson's scale). За вивчення впливу формальної реєстрації сімейних відносин між парами однієї статі (Fingerhut & Maisel, 2010) автори оцінювали задоволеність життям (Satisfaction With Life Scale). Дослідження взаємозв'язку поведінки, спрямованої на підтримання відносин, та індивідуального благополуччя (Baker, McNulty, Overall, Lambert, & Fincham, 2013), орієнтується на оцінку депресивної симптоматики (CES-D). Стаття, присвячена впливу взаємин з агресивним партнером на індивідуальне благополуччя (Arriaga, Caperza, Goodfriend, Rayl, & Sands, 2013), спирається на результати використання єдиного питання щодо загального почуття щастя респондента. У межах вивчення впливу міждержавних територіальних суперечок на індивідуальне благополуччя (Miller, 2013) останнє оцінено за допомоги одного запитання — наскільки щасливим відчуває себе респондент. За аналізу взаємозв'язку флуктуацій у взаєминах з індивідуальним благополуччям (Whitton, Rhoades, & Whisman, 2014) автори дослідження звертаються до психологічного дистресу (окремі індикатори з Mood and Anxiety Symptom Questionnaire) і задоволеності життям (Satisfaction With Life Scale). У дослідженні впливу глобалізації, релігії та економічного зростання в Туреччині на індивідуальне благополуччя (Sandıkcı, Peterson, Ekici, & Simkins, 2015) як показник останнього використано задоволеність життям (Satisfaction With Life Scale).

Отже, індивідуальне благополуччя оцінюють переважно в контексті трьох складових — психологічного дистресу, задоволеності життям і загального почуття щастя. Незважаючи на те, що інструментарій наведених досліджень є досить різноманітним, відзначити можна такі шкали, як CES-D і Satisfaction With Life Scale. При цьому в усіх розглянутих статтях відсутня концептуалізація поняття «індивідуальне благополуччя», яке використане як само собою зрозуміле і розкрито в тих або тих індикаторах.

З цієї причини особливий інтерес являє праця Р. Сінгха й А. Арора (Singh & Arora, 2010), у якій деталізовані складові індивідуального благополуччя. Автори визначають останнє як задоволеність життям, що віддзеркалюється у відповідних судженнях індивіда. З метою максимально точного вимірювання індивідуального благополуччя автори використовують підхід, згідно з яким життєва задоволеність диференціюється відповідно до різних життєвих сфер. Оцінки, отримані за кожну з них, узагальнюють, що дає вихід на вимір індивідуального благополуччя загалом. Ці сфери Р. Сінгх і А. Арора розподілили на фізичну (здоров'я, зовнішність, активність), соціальну (сім'я/домашнє господар-

ство, друзі, підтримка, сусіди), економічну (дохід, ресурси), зайнятості (професія та її характер), освітню, енвайроментальну (культура, уряд), досвіду (загальна поінформованість, подорожі всередині своєї країни і за її межами), релігійну/духовну, рекреаційну і сферу свободи. Відповідна методика містить 84 індикатори. Цей підхід дуже близький до вітчизняної методики Є. Головахи і Н. Паніної (Головаха & Панина, 1997), які виокремили 11 сфер життєдіяльності, оцінка яких дозволяє виміряти соціальне самопочуття індивіда.

Проведений аналіз дозволяє зробити низку висновків щодо використання терміна «індивідуальне благополуччя» в сучасних соціальних дослідженнях. По-перше, цю категорію не можна назвати теоретично усталеною. Найчастіше її використовують як синонім суб'єктивного благополуччя, задоволеності життям або психологічного дистресу. По-друге, існує безліч аспектів індивідуального життя, які так чи так визначають індивідуальне благополуччя. По-третє, досить часто індивідуальне благополуччя оцінюють із використанням кількох комплексних вимірювальних інструментів.

З огляду на це в контексті соціологічного вимірювання під індивідуальним благополуччям ми розуміємо визначальні показники якості життя індивіда. Під визначальними показниками ми розуміємо такі, в яких знаходять вираження різні аспекти індивідуальної життєдіяльності, навіть якщо вони не порушені вимірювальною методикою явно. Так, психологічний дистрес є таким показником індивідуального благополуччя, якщо йдеться про мережі соціальної взаємодії індивіда, різноманітні стресори, з якими він стикається, а також про певні типи його установок, знань і поведінки (Дембіцький, 2018b). Крім психологічного дистресу, до визначальних показників індивідуального благополуччя ми відносимо соціальне самопочуття (хоча точніше було б говорити про соціальне благополуччя) як показник якості інституційного контексту, а також соціальний статус індивіда як зовнішній показник успіху його життєдіяльності.

Далі ми зосередили увагу на особливостях вимірювання психологічного дистресу і соціального самопочуття. Що стосується методик оцінки соціального статусу, то відповідна наукова активність залишається для нас у статусі перспективної.

Глава 8

Конструювання і валідизація соціологічного тесту SCL-9-NR

§ 8.1. Вступ

Вивчення психологічного дистресу в соціологічних дослідженнях. Сучасне розуміння психологічного дистресу деталізовано у праці Ш. Ріднер (Ridner, 2004). Вона визначає психологічний дистрес як унікальний емоційно дискомфортний стан, який відчуває індивід у відповідь на специфічний стресор або сильну потребу в чому-небудь і веде до тимчасових або постійних негативних наслідків. Для психологічного дистресу характерне:

- відчуття неспроможності ефективно впоратися з проблемою;
- зміна в емоційному стані;
- дискомфорт;
- наявність негативних наслідків, спричинених дискомфортом.

Становлення цього напрямку соціологічних досліджень пов'язане з ім'ям відомого американського соціолога Леонарда Перліна, головні праці якого присвячені ресурсам і практикам протидії (coping) стресу (Pearlin & Schooler, 1978), процесу стресу (Pearlin, Menaghan, Lieberman, & Mullan, 1981; Pearlin, 1989), взаємозв'язку стресу і життєвого курсу (life course) індивіда (Pearlin & Skaff, 1996; Pearlin, Schieman, Fazio, & Meersman, 2005; Pearlin, 2009).

Щоб проаналізувати сучасний стан вивчення психологічного дистресу в соціальних науках, ми звернулися до двох періодичних видань — «Society and Mental Health» і «Journal of Health and Social Behavior», що виходять під егідою Американської соціологічної асоціації та спеціалізуються на вивченні різних аспектів психологічного здоров'я. Розглянемо всі статті цих часописів починаючи з 2011 року, в яких предметом наукового інтересу є психологічний дистрес. Можна виокремити як мінімум 17 таких публікацій, аналіз яких демонструє специфіку вивчення проблеми психологічного дистресу, властиву і для соціології.

По-перше, більшість досліджень, сфокусованих на психологічному дистресі (посилання на них наведено далі), мають кількісний характер із використанням крос-секційного або лонгітюдного дизайну. Також опублікована праця теоретико-методологічної спрямованості, що подає критику модального кількісного підходу з позицій символічного інтеракціонізму (McLeod, 2012), і дві статті, присвячені класику соціології здоров'я, який зробив неоціненний

внесок у вивчення різних аспектів психологічного дистресу, — Л. Перліну (Aneshensel, 2015; Aneshensel & Avison, 2015).

По-друге, модель усіх емпіричних досліджень, по суті, одна і та сама — психологічний дистрес поданий як залежна змінна, а різні соціальні чинники — як незалежні. Для розв'язання відповідних дослідницьких завдань зазвичай використовують різні види регресійного аналізу або лінійних структурних рівнянь. При цьому різноманіття пояснювальних факторів є доволі широким — історичні скрути і складності в рамках освітньої системи для окремо взятої етнічної групи (Walls & Whitbeck, 2011), соціальний капітал (Song, 2011), рольовий конфлікт (Glavin, Schieman, & Reid, 2011), фізичні стресори з урахуванням статусу сімейних відносин (Bierman, 2012), події стресори з урахуванням подальшої рефлексії (Young & Foy, 2013), релігійність у різних етнічних групах як «буферний» ресурс (Oates & Goode, 2013), доступність страхування для різних груп (Alang, McAlpine, & Henning-Smith, 2014; Jacobs, Hill, & Burdette, 2015), залученість у діяльність релігійних і громадських організацій (Acevedo, Ellison, & Xu, 2014), ув'язнення близьких родичів (Brown, Bell, & Patterson, 2016), різного роду труднощі, з якими індивід стикався в дитинстві (Fothergill, Ensminger, Doherty, Juon, & Green, 2016), трансграничні зв'язки мігрантів із членами їхніх сімей (Torres, Alcántara, Rudolph, & Viruell-Fuentes, 2016), макроекономічні показники країн виїзду іммігрантів (Montazer & Wheaton, 2017), оцінка ризику втратити роботу (Burgard & Seelye, 2017), вплив ідентичності, самооцінки і почуття контролю (mastery) (Gallagher, 2017).

По-третє, статті вказують на вимірювальні шкали, які користуються довірою на даному етапі. Перш за все це Шкала депресії Центру епідеміологічних досліджень (CES-D) або ж її скорочені варіанти (використано в семи зі згаданих досліджень). Другими за популярністю можна назвати методики Р. Кесслера (використані в п'яти зі згаданих досліджень: К6 — чотири рази, К10 — один раз), третя — BSI-18 Л. Дерогатіса (використана у двох із згаданих досліджень). Разом з тим усе розмаїття інструментів, що вимірюють психологічний дистрес, а також придатні для використання в масових опитуваннях, не обмежується трьома зазначеними методиками.

Методики вивчення психологічного дистресу. З метою вимірювання вираженості психологічного дистресу було розроблено безліч методик, але якщо говорити про ті з них, що придатні для широкого використання, можна виокремити такі:

- Beck Depression Inventory (21 індикатор) (Beck, Steer, & Carbin, 1988);
- Distress Disclosure Index (12 індикаторів) (Kahn, Hucke, Bradley, Glinski, & Malak, 2012);
- General Health Questionnaire (4 версії — 60, 30, 28, 12 індикаторів) (Goldberg & Hillier, 1979);
- Center for Epidemiologic Studies Depression Scale (3 версії — 20, 12, 8 індикаторів) (Radloff, 1977);
- Kessler Psychological Distress Scale (2 версії — 10, 6 індикаторів) (Kessler et al., 2002);
- Mental Health Inventory (38 індикаторів) (Walls & Whitbeck, 2011);

- Perceived stress scale (2 версії — 14, 10 індикаторів) (Cohen, 2010);
- Symptom Checklist Revised (6 версій — 90, 53, 27, 18, 14, 9 індикаторів) (Derogatis & Fitzpatrick, 2004).

Російськомовні версії цих методик доступні для Beck Depression Inventory (Тарабрина и др., 2007, с. 169–172), General Health Questionnaire (Ignatyev et al., 2012), CES-D (Беялов, 2018), Kessler Psychological Distress Scale (Transcultural Mental Health Center, 2018), Perceived stress scale (Nelson, 2015) і Symptom Checklist Revised (Тарабрина и др., 2007, с. 155–168).

Зазначимо, що жодна з цих шкал не пройшла повноцінних валідаційних перевірок серед російсько- й україномовних респондентів. Маючи доступ до результатів використання повної версії Symptom Checklist Revised (далі — SCL-90-R) у трьох українських загальнонаціональних опитуваннях (1997, 1999, 2014), що включають як російсько-, так і україномовних респондентів, нами було проведено роботу з розробки соціологічного тесту психологічного дистресу (SCL-9-NR⁵⁰), який відповідає вимогам використання в масових опитуваннях.

§ 8.2. Теоретична валідність SCL-9-NR

На нашу думку, основою теоретичної валідності SCL-9-NR є зміст вихідного тесту, на підставі якого він був розроблений, а саме — SCL-90-R. Теоретична валідність останнього базується на описі симптомів різних психіатричних розладів, отриманих унаслідок професійної практики фахівців психіатричного профілю. З цієї причини є сенс окремо звернутися до розробки SCL-90-R.

Історія створення цієї методики бере свій початок від симптоматичного опитувальника Хопкінса (Hopkins Symptom Checklist) (Derogatis, Lipman, Rickels, Uhlenhuth, & Covi, 1974), який охопив п'ять симптоматичних вимірів (соматизація, obsесивно-компульсивний розлад, міжособистісна сензитивність, депресія і тривожність), але при цьому мав низку обмежень: 1) його використання передбачало роботу тільки зі здоровими індивідами; 2) деякі важливі симптоматичні виміри не були охоплені; 3) значна кількість індикаторів методики була просто «шумом», тобто не мала особливої діагностичної цінності. Виходячи з цих міркувань, Л. Дерогатіс, Р. Ліпман і Л. Кові включили в SCL-90 55 індикаторів симптоматичного опитувальника Хопкінса і доповнили їх 45 новими, які відбивали чотири додаткові симптоматичні виміри. Подальша оцінка отриманого інструмента показала, що окремі показники тривожності й obsесивно-компульсивного розладу мали певні вади. Відповідні пункти були замінені, після чого була проведена чергова, цього разу успішна, перевірка. Цей фінальний варіант методики було названо «The Symptom Checklist-90-Revised», або SCL-90-R⁵¹ (Vaurio, 2011, p. 2448; Тарабрина и др., 2007, с. 146–147).

⁵⁰ Повна назва: Symptom Checklist 9 New Revision.

⁵¹ Для докладного ознайомлення з розробкою методики див.: (Derogatis, & Fitzpatrick, 2004).

SCL-90-R являє собою опитувальний лист, призначений для самозаповнення. Він дозволяє зафіксувати низку психологічних симптомів і дистрес загалом. Метою методики є оцінка дев'яти симптоматичних вимірів — соматизації (somatization), obsesивно-компульсивного розладу (obsessive-compulsive), міжособистісної сензитивності (interpersonal sensitivity), депресії (depression), тривожності (anxiety), ворожості (hostility), фобічної тривожності (phobic anxiety), параноїдальних ідей (paranoid ideation), психотизму (psychoticism). Також вона забезпечує можливість розрахунку трьох індексів психологічного дистресу — глобального індексу тяжкості дистресу (Global Severity Index, GSI), індексу наявного симптоматичного дистресу (Positive Symptom Distress Index), загальної кількості виявлених симптомів (Positive Symptoms Total). Отже, інтерпретувати результати можна на підставі комбінації трьох рівнів агрегованих даних — узагальненого бала за всіма пунктами методики, узагальненого бала за однією і більше підшкал, що відбивають відповідні симптоматичні виміри, а також на підставі зазначених вище індексів (Vaurio, 2011, p. 2447; Vufka & Camp 2010, p. 73–74).

За більш ніж 40-річну історію використання SCL-90-R було проведено безліч валідаційних досліджень у різних країнах і стосовно різних клінічних і звичайних груп. В Україні методику перекладено на російську й українську мови, а також адаптовано Наталією Паніною. З метою валідації SCL-90-R була включена нею в моніторингові опитування Інституту соціології НАН України в 1997 і 1999 роках. 2014 року методика була використана у всеукраїнському опитуванні спільного моніторингу Українського інституту соціальних досліджень імені О. Яременка, Центру «Соціальний моніторинг» і відділу моніторингових досліджень соціально-економічних трансформацій Інституту економіки і прогнозування НАН України (за ініціативи наукового керівника моніторингу О. Балакіревої).

Відповідні дані продемонстрували досить цікавий момент — розподіл підсумкових значень методики (йдеться про GSI⁵² та індекси окремих підшкал) у результатах як повного варіанта методики, так і скорочених. Нижче як приклад наведено розподіл результатів індексу SCL-9-UA⁵³ 1999 року (рис. 8.1). Такий самий патерн виявлено і в дослідженні Р. Клагхофера за конструювання SCL-K-9 (Klaghofer & Brähler, 2001, p. 121). Останнє дозволяє припустити універсальний характер такого розподілу результатів різних варіантів методики Л. Дерогатіса. Разом з тим подібний розподіл не виглядає природним. На наш погляд, розподіл значень психологічного дистресу має бути близьким до нормального, якщо ми справді вважаємо, що відповідна властивість має бутеву, а не артефактну природу.

Допустивши правильність цієї гіпотези, необхідно припустити можливу ключову причину зсуву результатів використання SCL-90-R та її скорочених варіантів. У цій публікації як таку причину розглядають соціальну чутливість

⁵² Розраховується як середнє значення за всіма 90 показниками методики.

⁵³ Докл. див.: (Дембицкий & Середя, 2015, с. 49–50).

теми психічного здоров'я в масових опитуваннях, з одного боку, і не відповідні в цьому контексті формулювання методики — з іншого⁵⁴. Щоб перевірити це припущення, ми змінили завдання і шкалу відповідей скороченого варіанта методики й апробували їх у рамках методичних опитувань.

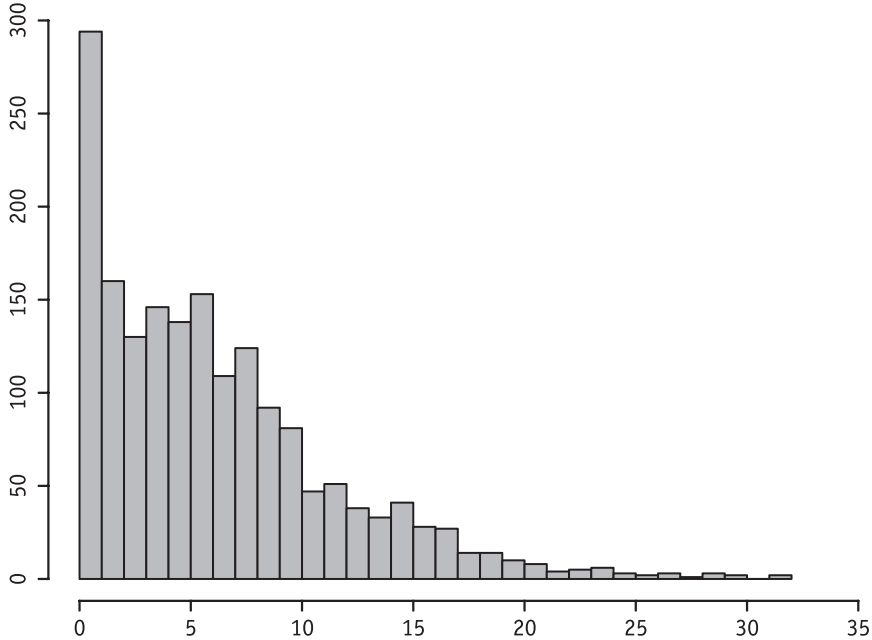


Рисунок 8.1. Розподіл результатів індексу SCL-9-UA 1999 року (Соціологічний моніторинг ІС НАН України «Українське суспільство»)

§ 8.3. Особливості конструювання SCL-9-NR

Відбір індикаторів. Індикатори SCL-9-NR відібрано на підставі SCL-90-R⁵⁵. Для цього було використано підхід Р. Клагхофера й Е. Брелера, які відштовхувалися від сили взаємозв'язку окремих пунктів шкали з глобальним індексом дистресу (GSI), по одному індикатору для кожної підшкали (Klaghofer & Brähler, 2001).

За визначення показників для SCL-9-NR ми взяли до уваги силу зв'язку з GSI (r_{GSI}), а також з індексом своєї підшкали (r_{st}). Основою аналізу був об'єднаний масив ($N = 5689$) на підставі даних соціологічного моніторингу Інституту соціології НАНУ «Українське суспільство» (хвилі 1997 і 1999 років), а також

⁵⁴ Оригінальне формулювання завдання в SCL-90-R: «Нижче наведено перелік проблем і скарг, що іноді виникають у людей. Будь ласка, читайте кожен пункт уважно. Обведіть кружком номер тієї відповіді, яка найточніше описує, наскільки сильно Вас турбувала та або та проблема впродовж останніх двох тижнів, у тому числі сьогодні».

⁵⁵ Перед початком роботи над SCL-9-NR була проведена ретельна методична робота з результатами використання SCL-90-R (Дембицкий & Серед, 2015; Sereda & Dembitskyi, 2016).

всеукраїнського опитування спільного моніторингу Українського інституту соціальних досліджень ім. О. Яременка, Центру «Соціальний моніторинг» і відділу моніторингових досліджень соціально-економічних трансформацій Інституту економіки і прогнозування НАН України 2014 року. Для оцінювання сили зв'язку було використано коефіцієнт кореляції Спірмена (табл. 8.1).

Таблиця 8.1

Сила зв'язку індикаторів SCL-9-NR з GSI та індексами своїх підшкал
($p < 0,001$)

<i>Симптоматичний вимір</i>	<i>Індикатор</i>	r_{GSI}	r_{si}
Ворожість	Легке виникнення досади або роздратування	0,61	0,75
Обсесивно-компульсивний розлад	Те, що Вам важко зосередитися	0,64	0,71
Депресія	Пригнічений настрій, “хандра”	0,61	0,72
Міжособистісна сензитивність	Те, що Ваші почуття легко зачепити	0,63	0,72
Параноїдальні ідеї	Почуття, що більшості людей не можна довіряти	0,53	0,75
Тривожність	Відчуття напруженості або збудженості	0,65	0,70
Соматизація	Відчуття слабкості в різних частинах тіла	0,69	0,76
Фобійна тривожність	Знервованість, коли Ви залишилися наодинці	0,56	0,68
Психотизм	Думки про те, що з Вашим тілом якийсь негаразд	0,57	0,67
Інтегральний індекс на основі наведених індикаторів (сума)		0,91	—

Зміна формулювань методики. У процесі зміни формулювання завдання методики ми виходили з необхідності її «пом'якшення», щоб вона не спричиняла побоювань респондентів, а відтак не зсувала результати опитування в бік менших градацій шкали. Наше фінальне рішення може виглядати неправильним з точки зору рекомендацій щодо впорядкування інструментарію у психодіагностиці⁵⁶. Разом із тим ми орієнтувалися на поради, наведені у відомій книзі С. Садмена і Н. Бредберна, присвяченій темі конструювання опитувального інструментарію для масових опитувань (Садмен & Бредберн, 2002). Вибір на користь соціології в цьому разі зумовлений тим, що мотивація людей, котрі прийшли до психолога по допомогу і заповнюють методику з проблем свого психологічного здоров'я, може бути зовсім іншою, ніж у тих, до кого прийшов інтерв'юєр і запропонував поділитися своїми проблемами. Перші зацікавлені

⁵⁶ Див., напр.: (Клайн, 1994, с. 96–101).

у визначенні своїх проблем та їх розв'язанні, другі незрідка — у приховуванні їх.

Як орієнтири формулювання нового завдання методики ми використовували такі (Садмен & Бредберн, 2002, с. 65–96):

- запитання, що заторкують теми, які спричиняють побоювання, мають бути достатньо довгими;
- для формулювань необхідно користуватися простою мовою;
- під час обговорення таких тем краще не ставити запитань про стан справ у теперішній момент;
- необхідно обґрунтовувати поставлені запитання, принаймні в загальних рисах.

Для відповідності останній вимозі ми скористалися технікою «так учиняє кожен» (Садмен & Бредберн, 2002, с. 86). Зрештою, завдання в SCL-9-NR було сформульовано так: *«У зв'язку з певними життєвими ситуаціями багато людей стикаються з різноманітними переживаннями, відчуттями і думками. Далі наведено перелік найчастіших станів такого роду. Оцініть, як часто кожен із них був характерним для Вас останнім часом: ніколи, зрідка, періодично, майже постійно»*.

Як можна бачити, спершу згадують поважну причину виникнення станів, пов'язаних із психологічним дистресом. Далі зазначають, що такі стани є найчастішими. Наприкінці акцент роблять не на актуальному стані, а на стані останнього часу, що потребує додаткової аргументації. Ми виходимо з того, що психологічний дистрес належить до того класу явищ, наслідки яких неможливо подолати в найкоротші терміни (якщо, звісно, не застосовувати спеціальних медичних препаратів). Звідси, «останній час» — це, власне, той проміжок часу, котрий безпосередньо пов'язаний з актуальним психологічним станом індивіда.

Таким чином, фінальний варіант SCL-9-NR, застосований для валідизації, містив змінене формулювання завдання — 9 пунктів із 90 початкових і 4-бальну шкалу замість 5-бальної (див. Додаток 1).

§ 8.4. Емпірична валідизація і перевірка надійності методики

Дизайн. Для розв'язання завдань валідизації та перевірки надійності методики здійснено низку опитувань.

З метою перевірки специфіки розподілу індексу SCL-9-NR проведено два онлайн-опитування (збір даних здійснила фірма «Factum Group», Київ). Перше — 19 листопада 2015 року серед мешканців Києва, друге — з 25 по 27 листопада 2015 року серед мешканців Львова. У кожному дослідженні опитано по 200 респондентів. Обидві вибірки побудовані на основі практично рівних за розміром квот, сформованих на підставі статі (2 категорії) і віку (3 категорії). Середній вік респондентів в обох випадках дорівнює 39 рокам. Аналогічне опитування у Львові проведене з 17 по 19 травня 2016 року з метою аналізу прогностичного потенціалу методики та її порівняння в цьому аспекті з K10 Р. Кесслера. Середній вік респондентів цього масиву дорівнює 40 років.

У рамках критеріальної валідазації здійснено опитування серед пацієнтів Чернігівської обласної психоневрологічної лікарні (далі — ЧОПНЛ; $N = 181$). Метод опитування — самозаповнення. Це опитування також дозволило перевірити взаємозв'язок значень методики з низкою важливих параметрів, серед яких соціальна бажаність. Збір даних тривав з 5 лютого по 4 березня 2016 року. Кількість чоловіків і жінок у вибірці однакова, середній вік — 44 роки. Надалі (з 25 травня по 31 жовтня 2016 року) в ЧОПНЛ додатково проведено 101 інтерв'ю з метою уточнення прогностичного потенціалу тесту і його порівняння з К10 Р. Кесслера. Кількість чоловіків у вибірці дорівнює 52%, жінок — 48%, середній вік — 42 роки.

Після отримання перших результатів із ЧОПНЛ, що серед іншого показали зв'язок індексу SCL-9-NR зі шкалою брехні MMPI, було ухвалене рішення про проведення додаткового онлайн-опитування серед користувачів соціальних мереж «facebook.com» і «vk.com» із метою перевірки аналогічного взаємозв'язку у групі звичайних респондентів. Відповідне дослідження здійснене спільно з кафедрою методології та методів соціологічних досліджень Київського національного університету імені Тараса Шевченка в період з 24 по 27 березня 2016 року ($N = 380$). Кількість чоловіків у вибірці — 28%, кількість жінок — 72%, середній вік — 25 років.

Методику було включено в загальнонаціональне моніторингове опитування Інституту соціології НАНУ «Українське суспільство» в червні 2016 року ($N = 1800$). Метод опитування — самозаповнення. Відповідні дані дозволили здійснити конструктну валідацію SCL-9-NR, а також розширити критеріальну. Кількість чоловіків у вибірці — 45%, жінок — 55%, середній вік — 45 років.

Спеціальне опитування у два етапи (з 1 по 5 вересня і з 26 по 30 вересня 2016 року) здійснене з метою перевірки ретестової надійності методики ($N = 61$). Метод опитування — самозаповнення. Кількість чоловіків у вибірці — 48%, кількість жінок — 52%, середній вік — 40 років.

Характер розподілу значень індексу. Розподіл значень SCL-9-NR в обох дослідженнях візуально близький до нормального. Графіки квантилів показують відхилення від нормальності для високих значень індексу (рис. 8.2). При цьому результати тесту Шапіро—Вілка вказують на те, що розподіл статистично значимо відмінний від нормального як для дослідження в Києві ($W = 0,98$; $p < 0,01$), так і для дослідження у Львові ($W = 0,97$; $p < 0,01$). Разом із тим цей критерій дуже чутливий до розміру вибірки — зі збільшенням вибірки ймовірність відхилення нульової гіпотези зростає (Field, 2012, р. 182–185).

Щоб перевірити вплив розміру вибірки на результати тесту в нашому випадку, ми сформували випадкові вибірки розміром 100, 50 і 25 спостережень по 1000 разів для кожного з двох досліджень на підставі їх масивів. Для кожної вибірки використано критерій Шапіро—Вілка, що дало змогу визначити кількість випадків, у яких приймається нульова гіпотеза ($n_p > 0,05$). На підставі цих показників можна оцінити відповідність розподілів нормальному (табл. 8.2).

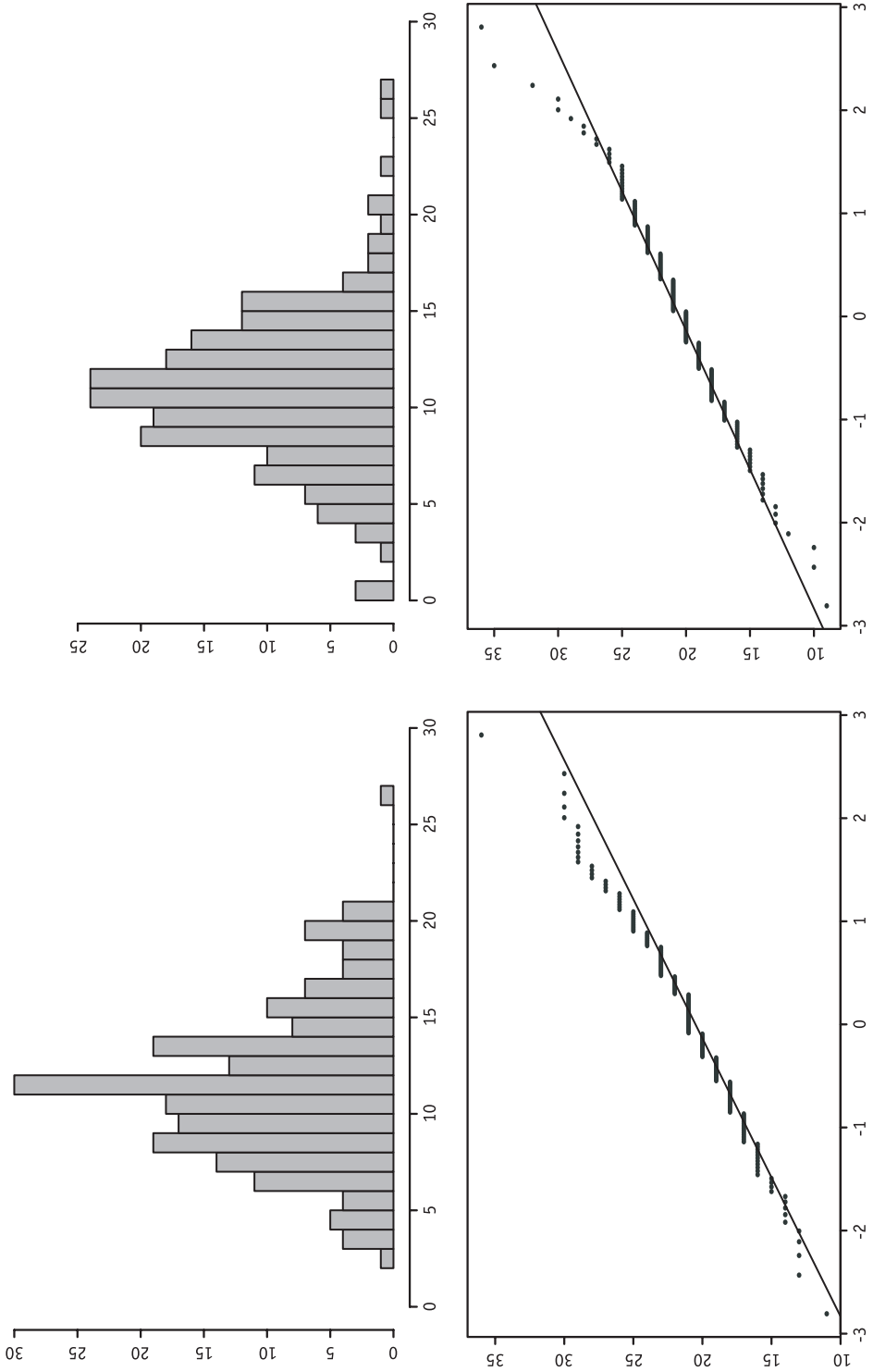


Рисунок 8.2. Гістограми та графіки квантилів для київського (ліворуч) і львівського (праворуч) онлайн-опитувань

Таблиця 8.2

**Результати використання критерію Шапіро—Вілکا для вибірок різного розміру,
 $n_p > 0,05$ (онлайн-опитування)**

Розмір вибірки	Київ	Львів
100	467 із 1000	349 із 1000
50	764 з 1000	659 із 1000
25	875 із 1000	782 зі 1000

Як видно на вибірці у 25 спостережень, нульова гіпотеза підтвердилася у 87,5% випадків для київського дослідження і 78,2% — для львівського. Аналогічну перевірку здійснено щодо результатів використання SCL-9-NR у загальнонаціональному моніторинговому опитуванні Інституту соціології НАН України 2016 року. Для порівняння також було взято результати використання дев'яти відповідних індикаторів з оригінальної методики Л. Дерогатіса (SCL-90-R), яку використовували у хвилі 1999 року того самого моніторингового дослідження (табл. 8.3).

Таблиця 8.3

**Результати використання критерію Шапіро—Вілکا для вибірок різного розміру,
 $n_p > 0,05$ (загальнонаціональні опитування)**

Розмір вибірки	1999	2016
100	0 із 1000	457 із 1000
50	39 із 1000	809 із 1000
25	369 із 1000	901 із 1000

Очевидним є факт того, що зміна формулювань методики істотно вплинула на характер його розподілу, який став значно ближчим до нормального.

Критеріальна валідизація SCL-9-NR: вступні положення. В основі критеріальної валідизації тесту лежить ідея, згідно з якою окремі групи пацієнтів психоневрологічної лікарні повинні показувати вищий рівень психологічного дистресу порівняно з іншими групами пацієнтів, а також простими респондентами⁵⁷. При цьому основний спосіб групування пацієнтів ґрунтується на підставі лікувальних центрів. Розглянемо докладніше ці центри.

У Центрі гострих психічних і поведінкових розладів на лікуванні перебувають пацієнти з гострою психотичною симптоматикою, вираженими агресивними тенденціями стосовно себе й оточення, а також із схильністю до імпульсивних учинків. Такі пацієнти потребують посиленого спостереження й активної фармакотерапії. Зі зрозумілих причин опитування цієї категорії пацієнтів проводили на завершальній стадії їхнього лікування. З урахуванням

⁵⁷ Докладніше про ідею цього критеріального порівняння див.: (Дембицкий, Ященко, Сидоров, & Александрова, 2016b: с. 102–106).

використання фармакотерапії очікувалося, що такі пацієнти не продемонструють високих значень за використання SCL-9-NR.

Центр стаціонарної лікувально-реабілітаційної допомоги призначений для лікування хронічно хворих пацієнтів із вираженими дефіцитарними змінами особистості, сформованим дефектом особистості за основним діагнозом, котрі потребують терапії загострення психічного розладу й активної соціально-трудова реабілітації. У плані психологічних переживань ця група демонструє емоційну обмеженість, апатію й емоційну тупість, що також указувало на можливість отримання помірних значень SCL-9-NR.

У центрі лікування психосоматичних розладів, неврозів, пограничних і кризових станів перебувають пацієнти без виражених порушень нормального функціонування психіки, тобто без вираженої декомпенсації. При цьому пацієнти потрапляють до центру за власним бажанням, і далі використовується режим «відкритих дверей», тобто пацієнти самі визначають графік свого перебування в центрі. Загалом для цієї групи є характерною підвищена увага до своїх психологічних переживань, що давало підстави припускати отримання високих значень SCL-9-NR.

Крім центру перебування пацієнтів за формування вибірки враховано їхню стать. У відповідних лікувальних відділеннях використано суцільний добір, що завершувався в разі достатнього наповнення квоти (табл. 8.4).

Таблиця 8.4

Наповненість квот вибірки, побудованої на базі ЧОПНЛ, $N = 181$

Центр	Стать респондентів	
	Чоловіча	Жіноча
гострих психічних і поведінкових розладів	31	31
стаціонарної лікувально-реабілітаційної допомоги	30	26
лікування психосоматичних розладів, неврозів, пограничних і кризових станів	29	34

Опитували лише тих пацієнтів, котрі добровільно погодилися брати участь у дослідженні, а також могли самостійно ознайомитися з анкетною, зрозуміти і заповнити її.

Спершу ми вдалися до спроби використання SCL-90-R з формулюваннями завдання і варіантів відповідей із SCL-9-NR, але пілотне опитування показало, що методика з 90 індикаторів надто втомлює пацієнтів. Фактично її другу половину багато хто просто ігнорував, і переважна більшість симптомів отримувала оцінку «Ніколи». З огляду на це для основного етапу дослідження було затверджено анкету, що містила два невеликі блоки запитань — SCL-9-NR і шкалу брехні з опитувальника MMPI⁵⁸. Додатково щодо кожного респондента

⁵⁸ Докл. див.: (Мягков, 2002).

психолог ЧОПНЛ фіксував діагноз, відділення, дату потрапляння до лікарні та дату опитування.

Для категоризації розладів психіки використано Міжнародну класифікацію хвороб 10-го перегляду. Серед респондентів зафіксовано такі категорії психічних розладів і розладів поведінки: органічні, включно з симптоматичними, психічними розладами; шизофренія, шизотипові та маячні (маренневі) розлади; розлади настрою (афективні розлади); невротичні, пов'язані зі стресом і соматоформні розлади; розлади особистості та поведінки у зрілому віці (табл. 8.5). З точки зору багатоманіття діагнозів найбільш гетерогенною є підвибірка із Центру лікування психосоматичних розладів, неврозів, пограничних і кризових станів ($IQV = 0,84$). Підвибірки з центрів стаціонарної лікувально-реабілітаційної допомоги ($IQV = 0,61$) і гострих психічних і поведінкових розладів ($IQV = 0,57$) менш гетерогенні через те, що модальною категорією діагнозів для них є шизофренія, шизотипові та маренневі розлади (66% і 71% відповідно).

Таблиця 8.5

Розподіл категорій психічних розладів і розладів поведінки за центрами ЧОПНЛ, $N = 181$

Категорія	Центр		
	ГППР ^a	СЛРД ^b	ЛПРНПКС ^c
Органічні, включно із симптоматичними, психічними розладами	10	15	10
Шизофренія, шизотипові та маренневі розлади	44	37	4
Розлади настрою (афективні розлади)	7	3	10
Невротичні, пов'язані зі стресом і соматоформні розлади	1	0	32
Розлади особистості та поведінки у зрілому віці	0	1	7

a — Центр гострих психічних і поведінкових розладів.

b — Центр стаціонарної лікувально-реабілітаційної допомоги.

c — Центр лікування психосоматичних розладів, неврозів, пограничних і кризових станів.

Критеріальна і конструктна валідація SCL-9-NR — фокусування. Оскільки основою психологічного дистресу може бути багато незалежних факторів, далі йтиметься про конструювання індексу, а не шкали. У разі SCL-9-NR як такі передбачаються ворожість, obsесивно-компульсивний розлад, депресія, міжособистісна сензитивність, параноїдальні ідеї, тривожність, соматизація, фобійна тривожність і психотизм. Наявність в індивіда будь-якого з зазначених розладів посилює спільний психологічний дистрес, при цьому абсолютно не обов'язково, що вони якимось чином взаємозалежні на рівні універсального патерну. Відштовхуючись від такого допущення, засобом емпіричної валідації індексу бачимо критеріальну і конструктну перевірку.

Критеріальна валідація SCL-9-NR: перевірка впливу віку, часу перебування у стаціонарі й ефекту соціальної бажаності. Перш за все слід перевірити вплив

таких факторів, як вік, час перебування у стаціонарі та соціальна бажаність, які можуть викривити результати критеріального порівняння. Для того щоб у подальшому «винести за дужки» ці пояснювальні фактори, приділимо їм окрему увагу.

Вік. Дослідження різняться з точки зору вікового складу респондентів (табл. 8.6). Тому задля їх коректної подальшої перевірки варто оцінити вплив віку на рівень психологічного дистресу.

Таблиця 8.6

Вікові параметри різних досліджень

Параметри розподілу	ООКЛ ^a	ЧОПНЛ ^b	ООКСМ ^c
Мінімум	18,0	17,0	16,0
Перший квартиль	28,0	34,0	19,0
Медіана	36,0	44,0	21,0
Середнє значення	38,9	43,9	25,2
Третій квартиль	52,0	53,0	29,0
Максимум	68,0	75,0	78,0

a – Онлайн-опитування в Києві та Львові.

b – Опитування серед пацієнтів ЧОПНЛ.

c – Онлайн-опитування серед користувачів соціальних мереж.

Для зазначених досліджень сила зв'язку відповідно дорівнює: $r_{ООКЛ} = -0,14$ ($p < 0,001$; $n = 400$), $r_{ЧОПНЛ} = 0,11$ ($p = 0,15$; $n = 175$), $r_{ООКСМ} = -0,09$ ($p = 0,09$; $n = 367$). Як можна бачити, у двох випадках зв'язок не є статистично значимим, а в одному він доволі слабкий, отже, ним можна знехтувати.

Час перебування у стаціонарі. Оскільки в опитуванні на базі ЧОПНЛ ідеться про різноманітні групи респондентів із психіатричним діагнозом, для коректного порівняння їх необхідно оцінити вплив часу, проведеного у стаціонарі, на рівень психологічного дистресу. За аналізу вибірки загалом фіксується дуже слабкий, але статистично значимий взаємозв'язок: $r = -0,15$ ($p = 0,04$; $n = 175$). При цьому, якщо поділити вибірку на підгрупи відповідно до лікувальних центрів, то величина ефектів зменшується і перестає бути статистично значимою, а напрямок зв'язку у двох випадках змінюється з негативного на позитивний. Отже, цей параметр ми не розглядаємо як релевантний з точки зору оцінювання психологічного дистресу.

Соціальна бажаність. На відміну від попередніх двох факторів, ефект соціальної бажаності, оцінений у другому (ЧОПНЛ) і третьому (онлайн-опитування користувачів соціальних мереж) дослідженнях за шкалою брехні, справляє вираженіший вплив на величину індексу SCL-9-NR. І оскільки для всіх досліджень спостерігаються відмінності у вираженості психологічного дистресу між чоловіками і жінками, має сенс оцінювати ці групи окремо.

У вибірці звичайних респондентів тенденція «прикрашати» свої відповіді спостерігається як серед чоловіків, так і серед жінок. При цьому останні щодо психологічних переживань є щирішими: $r_c = -0,27$ ($p = 0,01$; $n = 96$), $r_{ж} = -0,21$ ($p < 0,001$; $n = 266$). Схожа ситуація характерна і для респондентів із психіатричним діагнозом — $r_c = -0,31$ ($p < 0,001$; $n = 86$), $r_{ж} = -0,21$ ($p = 0,05$; $n = 90$). Але в цій групі спостерігаємо істотні варіації в різних лікувальних центрах (табл. 8.7). Як бачимо, проблема соціальної бажаності відповідей особливо виражена у чоловіків, які перебувають у центрі стаціонарної лікувально-реабілітаційної допомоги, а також помірно характерна для всіх пацієнтів центру лікування психосоматичних розладів, неврозів, пограничних і кризових станів.

Таблиця 8.7

**Особливості взаємозв'язку між SCL-9-NR і шкалою брехні
для різних груп респондентів ЧОПНЛ**

Центр	Стать респондента					
	Чоловіча			Жіноча		
	<i>r</i>	<i>p</i>	<i>n</i>	<i>r</i>	<i>p</i>	<i>n</i>
гострих психічних і поведінкових розладів	-0,01	0,97	30	-0,15	0,41	31
стаціонарної лікувально-реабілітаційної допомоги	-0,62	< 0,001	29	-0,12	0,58	25
лікування психосоматичних розладів, неврозів, пограничних і кризових станів	-0,36	0,07	27	-0,30	0,08	34

Двома можливими поясненнями такого сильного взаємозв'язку в підгрупі чоловіків, які перебувають у Центрі стаціонарно-реабілітаційної допомоги, є ситуація опитування і соціально-психологічна специфіка цієї групи. Так, на відміну від інших респондентів, вони проходили не індивідуальне, а групове опитування. При цьому респонденти сиділи за різними столами і не втручалися в заповнення анкет інших опитуваних. Що стосується соціально-психологічної специфіки, то йдеться про чоловіків із хронічними психічними захворюваннями та середнім віком 44 роки (перший квартиль дорівнює 32,5 роки, третій — 55,3). Цілком імовірно, що психіатричний статус поряд із соціальними очікуваннями стосовно чоловіків працездатного віку роблять для них тему психічного здоров'я найвищою мірою сензитивною.

Що стосується пацієнтів Центру лікування психосоматичних розладів, неврозів, пограничних і кризових станів, то для них тема опитування «забарвлена» не так соціально, як психологічно, адже для них характерні гострі психологічні переживання щодо свого внутрішнього стану за високої міри збереженості соціальних і трудових навичок. У цьому контексті наявність такого взаємозв'язку виглядає цілком логічною.

Отже, ефект соціальної бажаності назагал характерний для результатів тестування за використання SCL-9-NR, що й не дивно з огляду на традиційну сензитивність теми здоров'я в масових опитуваннях. При цьому, як показало іні-

ціальне дослідження з валідизації SCL-9-NR (Дембицкий, 2016е), методика за цим показником виглядає вочевидь адекватнішою за оригінальну SCL-90-R.

Критеріальна валідиція SCL-9-NR: загальний рівень вираженості психологічного дистресу в різних групах. Насамперед розглянемо рівень психологічного дистресу серед різних груп пацієнтів ЧОПНЛ. Як ми і передбачали, серед лікувальних центрів найвищий дистрес демонструють пацієнти Центру лікування психосоматичних розладів, неврозів, пограничних і кризових станів, натомість у двох інших групах цей показник значно нижчий. При цьому, якщо розглядати центри окремо, показники жінок дещо вищі, ніж у чоловіків (табл. 8.8). Результати відповідного двофакторного коваріаційного аналізу (як коваріата використано показники шкали брехні; всі змінні розглядаються як незалежні) підтверджують вплив соціальної бажаності ($F = 16,9$; $df = 1$; $p < 0,001$), лікувального центру ($F = 11,4$; $df = 2$; $p < 0,001$) та статі ($F = 6,4$; $df = 1$; $p < 0,05$). Уточнений R^2 для цієї моделі становить 19,4%.

Таблиця 8.8

Середні значення SCL-9-NR в різних групах пацієнтів ЧОПНЛ

Центр	Стать респондента			
	Чоловіча		Жіноча	
	X	sd	X	sd
гострих психічних і поведінкових розладів	10,1	4,7	12,8	5,1
стаціонарної лікувально-реабілітаційної допомоги	9,6	4,9	12,8	4,6
лікування психосоматичних розладів, неврозів, пограничних і кризових станів	14,8	7,0	15,2	4,2

Виражені відмінності в рівні психологічного дистресу спостерігаються у пацієнтів із різними категоріями психічних розладів. Оскільки вибірка не збалансована, результати наведено без поділу за статтю (табл. 8.9). У цьому разі результати двофакторного коваріаційного аналізу (як коваріата використано показники шкали брехні; всі змінні розглядаються як незалежні) знову вказують на статистично значимий вплив усіх пояснювальних змінних — соціальної бажаності ($F = 18,6$; $df = 1$; $p < 0,001$), категорії психічного розладу ($F = 10,9$; $df = 4$; $p < 0,001$) і статі ($F = 8,2$; $df = 1$; $p < 0,01$). Уточнений R^2 для цієї моделі дорівнює 26,9%.

За аналізу значень SCL-9-NR серед звичайних респондентів було ухвалено не об'єднувати спостереження двох досліджень в один масив, а працювати з ними окремо. Це пов'язано з тим, що рівень психологічного дистресу серед чоловіків — учасників київського і львівського онлайн-опитувань — статистично значимо вищий порівняно з відповідним показником серед чоловіків — учасників онлайн-опитування користувачів соціальних мереж ($X_{u1} = 11,3$; $X_{u3} = 10,1$; $t = 2,5$; $df = 242,1$; $p < 0,05$). Виходячи з цього, київський і львівський масиви використано для порівняння індексу SCL-9-NR з відповідним показником масиву, отриманим за опитування пацієнтів ЧОПНЛ, а дані онлайн-

опитування користувачів соціальних мереж — для побудови коваріаційної моделі, що пояснює значення SCL-9-NR на підставі статі та шкали брехні.

Таблиця 8.9

Середні значення SCL-9-NR для різних категорій психічних розладів і розладів поведінки

<i>Категорія</i>	<i>SCL-9-NR</i>	<i>sd</i>	<i>n</i>
Органічні, включно з симптоматичними, психічні розлади	13,2	6,2	33
Шизофренія, шизотипові та маренневі розлади	10,6	4,9	83
Розлади настрою (афективні розлади)	13,5	4,5	20
Невротичні розлади, пов'язані зі стресом і соматоформні розлади	14,5	4,9	33
Розлади особистості та поведінки у зрілому віці	20,5	2,3	8

Звісно, значно цікавіше було б порівняти звичайних респондентів із пацієнтами, що мають різні категорії психічних розладів, але через слабку наповненість відповідних категорій респондентів як групувальний чинник обрано лікувальні центри. Як для чоловіків ($X_{q1} = 11,3$; $sd_{q1} = 4,2$), так і для жінок ($X_{жс1} = 12,0$; $sd_{жс1} = 3,9$) у звичайній вибірці зафіксовано статистично значимі відмінності відповідно до чоловіків ($X_{q2} = 14,8$; $t_q = 2,6$; $df_q = 28,7$; $p_q < 0,05$) і жінок ($X_{жс2} = 15,2$; $t_{жс} = 4,2$; $df_{жс} = 43,6$; $p_{жс} < 0,001$), опитаних в Центрі лікування психосоматичних розладів, неврозів, пограничних і кризових станів.

Однофакторний коваріаційний аналіз на підставі даних онлайн-опитування користувачів соціальних мереж продемонстрував статистично значимий вплив шкали брехні ($F = 21,3$; $df = 1$; $p < 0,001$) і статі ($F = 5,8$; $df = 1$; $p < 0,05$). Уточнений R^2 дорівнює 11,1%.

Подані результати порівняння різних груп респондентів відповідають теоретичним очікуванням і підтверджують критеріальну валідність SCL-9-NR. Окрім цього, її додаткове підтвердження отримано на підставі порівняння рівня психологічного дистресу у звичайних респондентів і тих, хто проживав у зоні бойових дій на Донбасі (інформація отримана в рамках загальнонаціонального моніторингового дослідження ІС НАНУ «Українське суспільство», 2016). І чоловіки ($X_{q4} = 13,3$; $sd_{q4} = 5,7$), і жінки ($X_{жс4} = 14,8$; $sd_{жс4} = 6,1$), які проживали на території бойових дій, демонструють вищий психологічний дистрес порівняно з чоловіками ($X_{q5} = 9,0$; $sd_{q5} = 5,3$; $t_q = 5,0$; $df_q = 52,3$; $p_q < 0,001$) і жінками ($X_{жс5} = 10,2$; $sd_{жс4} = 5,4$; $t_{жс} = 4,8$; $df_{жс} = 46, 3$; $p_{жс} < 0,001$) у звичайній підвибірці.

В останньому випадку основою відмінностей є, першою чергою, не психологічна, а соціальна специфіка, характерна для критеріальної групи. Докладніше про це говоритимемо далі.

Конструктна валідизація SCL-9-NR. У рамках перевірки для аналізу були доступні такі змінні, як самооцінка здоров'я респондента, соціальне самопо-

чуття⁵⁹ і оцінка екологічної ситуації за місцем проживання. Щоб зробити результати конструктивної валідації виразнішими, вибірку поділено на звичайних респондентів і тих, хто проживав у зоні бойових дій (табл. 8.10). Для оцінки поясненої дисперсії використана лінійна покрокова регресія.

Таблиця 8.10

Відсоток поясненої дисперсії SCL-9-NR для різних груп респондентів

Незалежні змінні	Звичайні респонденти		РПЗБД*	
	%	<i>p</i>	%	<i>p</i>
Самооцінка здоров'я	19,1	< 0,001	8,3	< 0,001
Соціальне самопочуття	6,3	< 0,001	26,0	< 0,001
Оцінка екологічної ситуації за місцем проживання	0,3	< 0,01	5,3	< 0,01
Разом	25,7	–	39,7	–

* Тут і далі — респонденти, які проживали в зоні бойових дій.

Як видно, специфіка взаємозв'язків SCL-9-NR з незалежними змінними в наведених групах істотно відрізняється. Для групи звичайних респондентів основним фактором є самооцінка здоров'я, натомість на соціальні фактори сукупно припадає значно менше пояснювального потенціалу. Для групи респондентів, що мешкали в зоні бойових дій, характерне переважання саме соціальних факторів. Цей факт чітко показує, що основою високого психологічного дистресу можуть бути не тільки психічні проблеми, а і складне соціальне становище індивіда (Dembitskyi, 2016).

Тепер порівнюємо відсоток поясненої дисперсії для SCL-9-NR з результатами аналогічних індикаторів з оригінальної методики Л. Дерогатіса (табл. 8.11), яка була використана у хвилях моніторингового дослідження Інституту соціології НАН України 1997 і 1999 років. При цьому для порівняння є доступними тільки дві з трьох незалежних змінних.

Таблиця 8.11

Відсоток поясненої дисперсії індексу на підставі дев'яти індикаторів SCL-90-R

Незалежні змінні	1997 рік		1999 рік	
	%	<i>p</i>	%	<i>p</i>
Самооцінка здоров'я	6,3	< 0,001	11,5	< 0,001
Соціальне самопочуття	2,4	< 0,001	4,5	< 0,001
Разом	8,7	–	16,0	–

⁵⁹ Виміряно за методикою «Інтегральний індекс соціального самопочуття» (Головаха & Панина, 1997).

Незважаючи на те, що самооцінка здоров'я і соціальне самопочуття статистично значимо пов'язані з результатами використання відібраних індикаторів SCL-90-R, величина поясненої дисперсії є на порядок нижчою. Це, на наш погляд, свідчить про кращу конструктну валідність SCL-9-NR порівняно з оригінальною методикою.

Оцінка внутрішньої узгодженості та ретестової надійності. Для оцінки внутрішньої надійності⁶⁰ використано коефіцієнт Альфа Кронбаха. Майже в усіх опитуваннях цей показник для SCL-9-NR перевищує необхідну межу 0,70 (табл. 8.12). Виняток становить онлайн-опитування користувачів соціальних мереж.

Таблиця 8.12

Показники внутрішньої узгодженості SCL-9-NR (коефіцієнт Альфа Кронбаха) для різних досліджень

Дослідження (рік)	N	Альфа Кронбаха
Київське онлайн-опитування (2015)	200	0,75
Львівське онлайн-опитування (2015)	200	0,79
Опитування в ЧОПНЛ (2016)	176	0,83
Онлайн-опитування користувачів соціальних мереж (2016)	380	0,68
Львівське онлайн-опитування (2016)	205	0,79
Загальнонаціональне опитування — звичайні респонденти (2016)	1689	0,89
Загальнонаціональне опитування — РПЗБД (2016)	92	0,91
Опитування для перевірки ретестової надійності — перша хвиля (2016)	61	0,74
Опитування для перевірки ретестової надійності: друга хвиля (2016)	61	0,82

У процесі дослідження з перевірки ретестової надійності між результатами першої та другої хвиль отримано коефіцієнт кореляції Пірсона, що дорівнює 0,76 ($p < 0,001$; $n = 61$). Разом з тим окремі індикатори з точки зору ретестової надійності відрізняються один від одного (табл. 8.13). Єдиним симптоматичним виміром, індикатор якого не демонструє статистично значимого взаємозв'язку між результатами першої та другої хвиль, є фобічна тривожність.

Якщо ж розглянути взаємозв'язок щодо індикатора фобічної тривожності за допомоги таблиці спряженості, то можна зробити висновок, що загалом результати двох опитувань є досить узгодженими (табл. 8.14).

⁶⁰ Узагалі інформація щодо показників внутрішньої узгодженості не є релевантною з точки зору розробки соціологічного тесту у формі індексу. Але ми наводимо її для читачів, які можуть нею зацікавитися.

Таблиця 8.13

Показники ретестової надійності (коефіцієнт Гамма) для окремих індикаторів SCL-9-NR

<i>Індикатор</i>	<i>Гамма</i>	<i>p</i>
Легке виникнення досади або роздратування	0,72	< 0,001
Вам важко зосередитися	0,74	< 0,001
Пригнічений настрій, нудьга	0,69	< 0,001
Ваші почуття легко зачепити	0,52	< 0,01
Почуття, що більшості людей не можна довіряти	0,79	< 0,001
Відчуття напруженості або збудженості	0,55	< 0,01
Відчуття слабкості в різних частинах тіла	0,77	< 0,001
Знервованість, коли Ви залишалися наодинці	0,36	0,13
Думки про те, що з Вашим тілом якийсь негаразд	0,73	< 0,001

Таблиця 8.14

Узгодженість розподілів відповідей за індикатором «Знервованість, коли ви залишаєтеся наодинці»

<i>Результати першої хвили</i>	<i>Результати другої хвили</i>			
	<i>Ніколи</i>	<i>Зрідка</i>	<i>Періодично</i>	<i>Майже постійно</i>
Ніколи	25	13	2	0
Зрідка	8	9	0	0
Періодично	1	2	0	0
Майже постійно	0	0	0	1

Порівняльна валідність: SCL-9-NR і K10. Двома найвідомішими методиками вивчення психологічного дистресу в масових опитуваннях є шкали Л. Редлоффа (скорочений варіант CES-D, 8 індикаторів), і Р. Кесслера (K10 і K6). CES-D використовували в Україні в рамках Європейського соціального дослідження⁶¹, K10 — в рамках Міжнародного дослідження психічного здоров'я ВООЗ⁶². Для проведення нашого порівняння ми зупинилися на K10 Р. Кесслера як новішій і змістовно багатшій методиці⁶³ (включити в опитування і K10, і CES-D не дозволяли організаційні обмеження дослідження).

⁶¹ Скорочену версію CES-D, що складається з восьми індикаторів, застосовували для вивчення особистісного благополуччя (European Social Survey, 2015).

⁶² Докладніше з результатами дослідження можна ознайомитися в: (Kessler & Üstün, 2008).

⁶³ K10 ґрунтується на індикаторах тривожності та депресії, натомість CES-D — виключно депресії.

Методика Р. Кесслера базується на двох основних симптоматичних вимірах — депресія (5 індикаторів) і тривожність (4 індикатори). Ще один індикатор надає симптоматичний вимір соматизації. Так, у змістовному сенсі К10 значно відрізняється від SCL-9-NR. І оскільки К10 є загально визнаним вимірювальним інструментом, виникає питання про доцільність ще однієї шкали, що розв'язує аналогічні завдання і при цьому використовує відмінний концептуальний підхід.

З цієї причини обидві методики включені в онлайн-опитування звичайних респондентів, а також в опитування серед пацієнтів ЧОПНЛ. Через відсутність можливості провести для К10 таку саму багатобічну валідацію, як і для SCL-9-NR, було використано підхід, відомий під назвою ROC-curve⁶⁴. Таке рішення обумовлене тим, що сам Р. Кесслер у рамках обґрунтування валідності своїх шкал використовував саме його (Kessler et al., 2002). Так, ним було отримано значення AUC (площа під кривою), рівне 0,87 — 0,88. Для К10 у нашому дослідженні показники на порядок нижчі — лишень 0,73. Натомість AUC для SCL-9-NR становив 0,80 (рис. 8.3).

За використання бінарної логістичної регресії для прогнозування наявності у респондента психіатричного діагнозу як SCL-9-NR ($\log odds ratio = 0,27$; $z = 7,4$; $p < 0,001$), так і К10 ($\log odds ratio = -0,11$; $z = -5,7$; $p < 0,001$) показують статистично значимі результати. Разом з тим значення AIC⁶⁵ свідчить про вищу якість моделі, побудованої на базі SCL-9-NR ($AIC_{SCL-9-NR} = 280,9$; $AIC_{K10} = 318,0$). Аналогічний висновок можна зробити на підставі аналізу чутливості, специфічності та загальної ефективності моделей⁶⁶ (табл. 8.15).

Таблиця 8.15

Чутливість, специфічність і ефективність моделі для SCL-9-NR і К10, %

Ймовірність наявності діагнозу на основі моделі	Чутливість		Специфічність		Ефективність	
	SCL-9-NR	К10	SCL-9-NR	К10	SCL-9-NR	К10
0,9	7	0	100	100	73	71
0,8	14	5	99	99	74	71
0,7	27	5	97	98	76	71
0,6	37	14	96	97	79	72
0,5	50	23	92	93	79	72
0,4	54	50	89	87	79	76
0,3	71	64	78	73	76	70
0,2	80	80	61	42	67	54
0,1	93	99	24	7	45	34

⁶⁴ Він полягає в побудові кривої, яка демонструє результати бінарної класифікації. У цьому разі класифікували респондентів із психіатричним діагнозом і без нього.

⁶⁵ Інформаційний критерій Акікі.

⁶⁶ У цьому разі чутливість моделі демонструє ефективність передбачення спостережень із психіатричним діагнозом, специфічність — звичайних спостережень, а під загальною ефективністю мається на увазі загальна кількість правильно передбачених випадків.

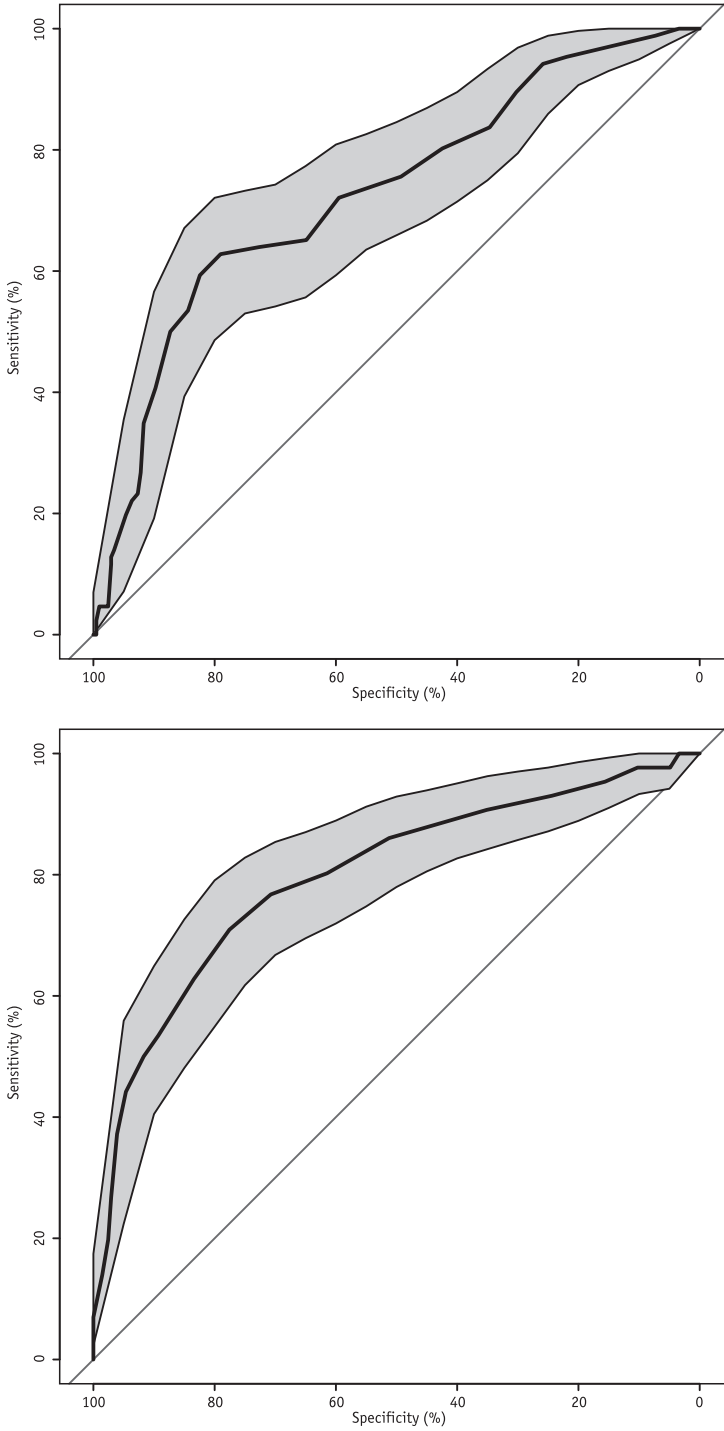


Рисунок 8.3. ROC-криві для K10 (зверху) і SCL-9-NR (знизу)

Отже, незважаючи на те, що SCL-9-NR показує себе краще за K10, з прогностичної точки зору відповідні результати можна назвати добрими, але не відмінними. Втім, основним завданням методики є вимірювання психологічного дистресу, а не визначення наявності психічного розладу у респондента. Аргументом на користь того, що це два різні завдання, є результати конструктивної валідації серед респондентів, які проживали на території бойових дій. Так, рівень психологічного дистресу в цій групі багато в чому обумовлений соціальними, а не психологічними труднощами.

§ 8.5. Критеріальні рівні SCL-9-NR

Дизайн. Уточнення критеріальних рівнів здійснювалося на підставі результатів п'яти досліджень: онлайн-опитувань у Києві та Львові (2015), онлайн-опитування у Львові (2016), опитування серед пацієнтів із психіатричним діагнозом у Чернігові та Харкові (2016–2017)⁶⁷, соціологічного моніторингу «Українське суспільство». Усі вони розглянуті вище.

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 року, до другої — респонденти онлайн-опитування 2016 року, до третьої — респонденти з психіатричним діагнозом, до четвертої — респонденти, які оцінюють стан свого здоров'я, як поганий або дуже поганий (соціологічний моніторинг), до п'ятої — респонденти, які мешкали в зоні бойових дій на Донбасі (соціологічний моніторинг), до шостої — респонденти соціологічного моніторингу «Українське суспільство», крім тих, які проживали в зоні бойових дій. Останніх виключено з основного масиву, оскільки ця група з'явилася в нашому суспільстві нещодавно і за дуже короткий термін, а отже не є характерною для його рівноважного стану, що спостерігався впродовж років незалежності.

Рішення про поділ респондентів онлайн-опитувань на дві групи ухвалено тому, що рівень вираженості психологічного дистресу в дослідженнях 2015 року є дещо вищим, ніж 2016-го. Про можливі причини такого розходження говоритимемо далі. 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 респондентів додатково поділено на дві підгрупи за статтю, оскільки в усіх дослідженнях чоловіки демонструють нижчі показники тесту. Розподілу відповідно вікових груп не проведено, тому що за контролю самооцінки стану здоров'я взаємозв'язок між віком та індексом психологічного дистресу зникав.

⁶⁷ На момент аналізу критеріальних рівнів SCL-9-NR було додатково зібрано 156 інтерв'ю (93 в Чернігові та 63 в Харкові) серед пацієнтів із психіатричним діагнозом. При цьому йдеться тільки про тих із них, діагнози яких належать до трьох категорій: 1) розлади настрою (афективні розлади); 2) невротичні, пов'язані зі стресом, і соматоформні розлади; 3) розлади особистості та поведінки у старшому віці.

Характер розподілу індексу SCL-9-NR. Оскільки психологічний дистрес розглядають як природний стан індивіда, його розподіл має відповідати нормальному закону. Відповідно, аналіз рівнів психологічного дистресу розпочато з оцінки характеру розподілу індексу SCL-9-NR. Для цього використано критерій Шапіро—Вілка, розрахований для підвбірок розміром 25 спостережень, відібраних по 1000 разів для кожної з шести зазначених груп респондентів⁶⁸ (табл. 8.16). За відповідними результатами, визначено кількість випадків, у яких прийнято нульову гіпотезу ($n_p > 0,05$).

Таблиця 8.16

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

<i>Група респондентів</i>	<i>Чоловіки</i>	<i>Жінки</i>
Онлайн-опитування 2015 року (далі — ОО2015)	82,7	84,0
Онлайн-опитування 2016 року (далі — ОО2016)	83,1	81,2
ЧОПНЛ, ХОКПЛ ⁶⁹	93,8	95,7
Соціологічний моніторинг — здоров'я (далі — СМ: З)	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

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

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

Рівні вираженості психологічного дистресу в різних групах респондентів. Перш ніж перейти до аналізу рівнів психологічного дистресу в різних групах, необхідно позначити 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 таких відповідей, яке базується на двох аргументах. По-перше, для контексту сучасного суспільства характерний постійний вплив стресорів, що унеможлиблює відсутність відповідних ефектів у житті індивідів. По-друге,

⁶⁸ Це зроблено через чутливість критерію до розміру вибірки. Гістограми самих розподілів тут не наведені, але доступні за: (Дембицкий, 2017а).

⁶⁹ Харківська обласна клінічна психіатрична лікарня.

максимальна вираженість за всіма симптоматичними вимірами не була виявлена серед респондентів із психіатричним діагнозом. З огляду на це зроблено припущення, що результат із максимально можливим значенням SCL-9-NR є такою малою ймовірною подією, що ним можна знехтувати.

Загалом показники психологічного дистресу для різних груп респондентів (табл. 8.17) відповідають теоретичним очікуванням. Черговість груп впорядкована відповідно до рівня вираженості психологічного дистресу в них.

Таблиця 8.17

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

Показник	СМ: без ЗБД		ОО2016		ОО2015		СМ: З		СМ: ЗБД		ЧОПНЛ, ХОКПЛ	
	Ч	Ж	Ч	Ж	Ч	Ж	Ч	Ж	Ч	Ж	Ч	Ж
перший квартиль	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
третій квартиль	12	14	11	13	14	14	17	16	18	21	18	19

Єдиною групою, яка не потрапила в рамки початкових теоретичних очікувань, є респонденти онлайн-опитування 2016 року. Їхні результати є значно ближчими до результатів соціологічного моніторингу «Українське суспільство», ніж до результатів онлайн-опитування 2015 року. Спочатку ми передбачали, що за онлайн-опитування респонденти будуть щирішими і це спричинить отримання вищих значень SCL-9-NR. Частотні розподіли (рис. 8.4, 8.5) частково підтверджують це припущення. Але відповідно до медіанного значення спосіб самозаповнення (друкованої або онлайн-анкети) не пояснює відмінність між результатами онлайн-опитувань 2015 року і моніторингового дослідження.

Найбільш логічним поясненням такої ситуації ми вбачаємо у впливі календарного часу опитування. 2015 року респонденти проходили опитування наприкінці листопада і на початку грудня, а 2016-го — у травні (онлайн-опитування) і червні (соціологічний моніторинг). Можливо, саме загальний емоційний стан наприкінці року і наприкінці 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 допомоги 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 свідчить про межі підвищеного і високого рівнів психологічного дистресу.

Для їх визначення слід звернутися до інших трьох груп, для яких характерна наявність факторів, що обтяжують психологічний стан індивіда. Відповідні результати можна використати за аналізу значень медіан і третіх квартилів розподілів.

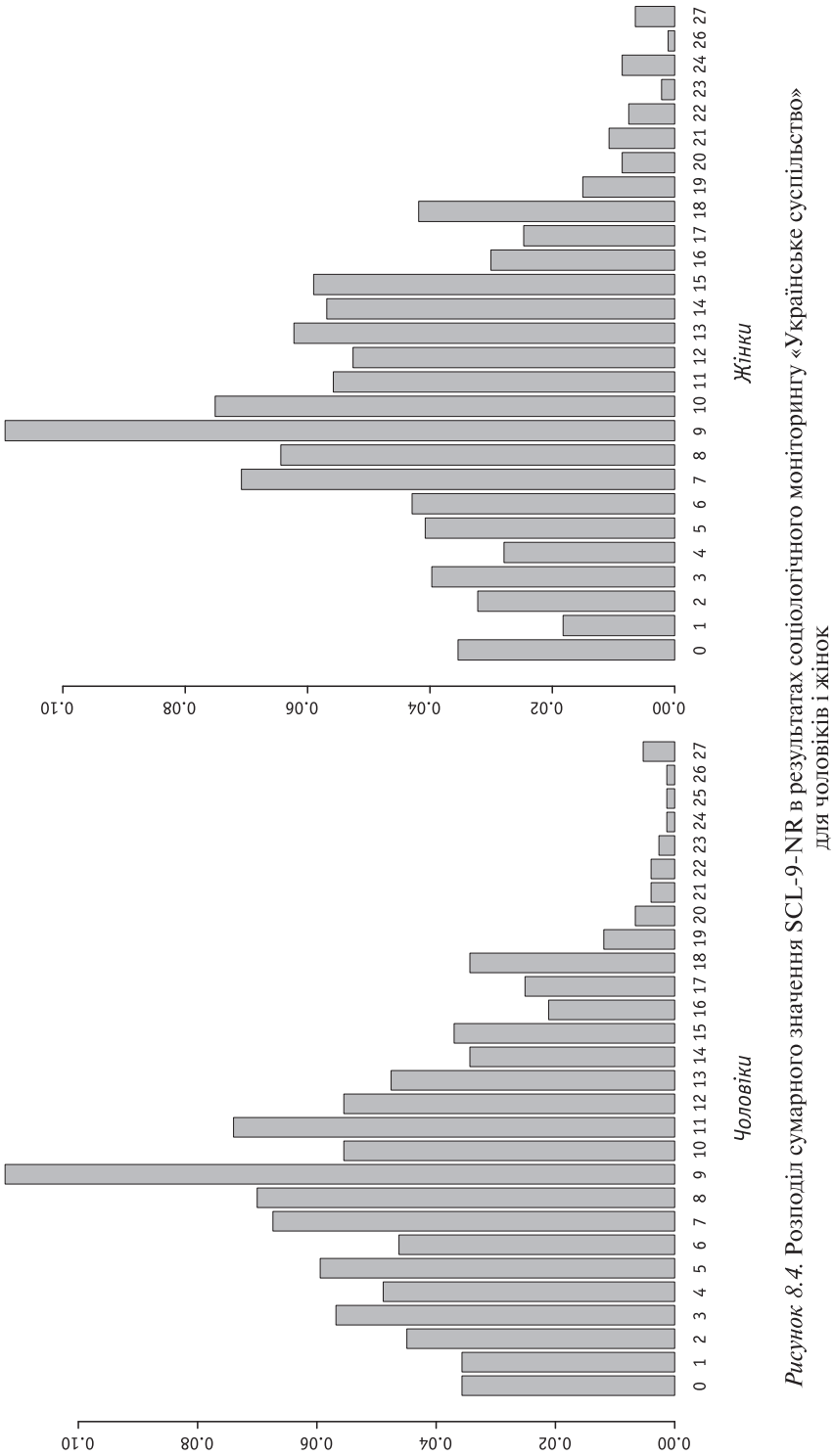


Рисунок 8.4. Розподіл сумарного значення SCL-9-NR в результатах соціологічного моніторингу «Українське суспільство» для чоловіків і жінок

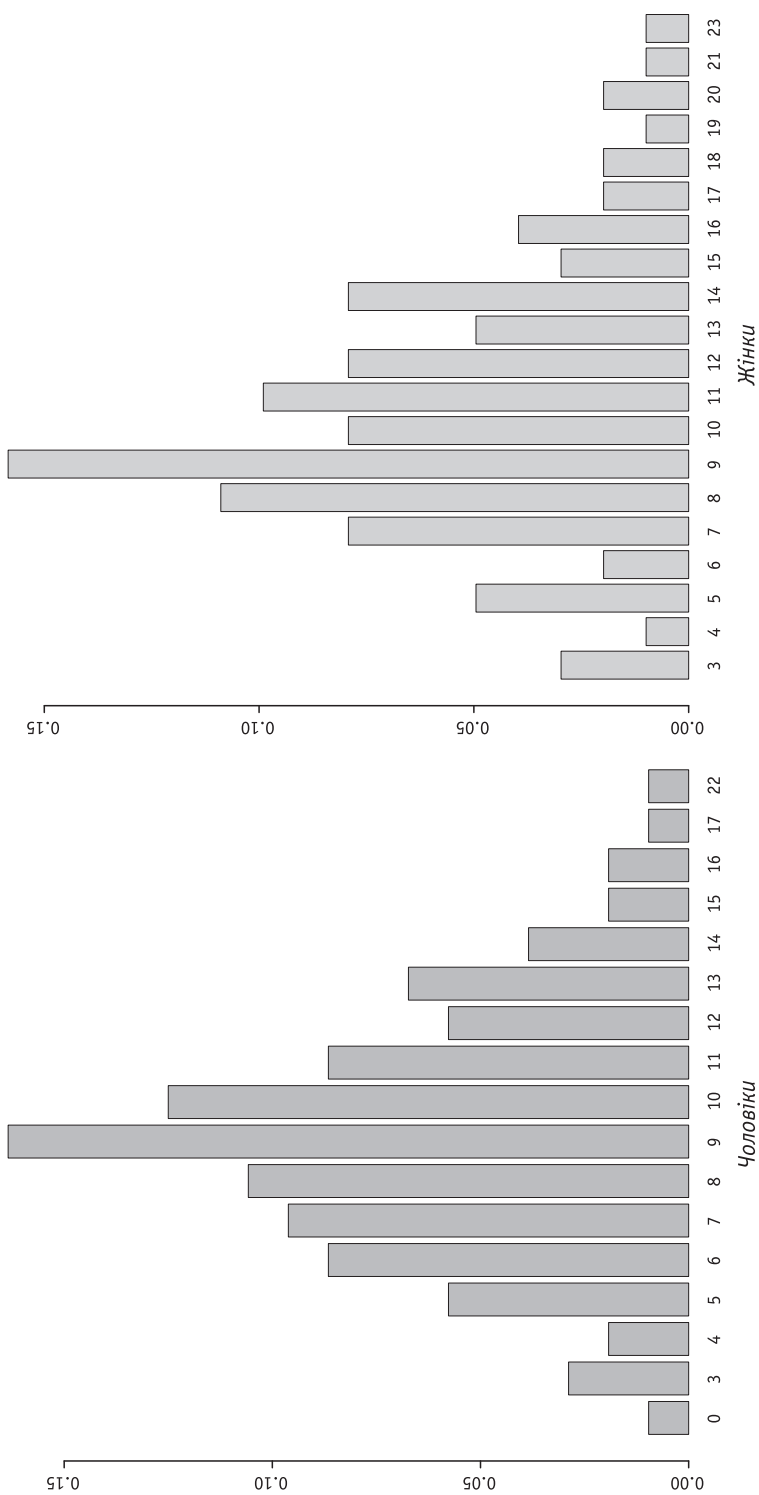


Рисунок 8.5. Розподіл сумарного значення SCL-9-NR в результатах онлайн-опитування 2016 року для чоловіків і жінок

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

Щодо межі високого рівня, то для чоловіків вона починається з 17 балів (третій квартиль «за здоров'ям»). Щодо жінок спостерігаємо досить великий розкид значень (16, 21 і 19 балів). При цьому медіана у групі з психіатричним діагнозом перетинається з третім квартилем у групі респондентів, які негативно оцінюють стан свого здоров'я (по 16 балів). На цій підставі пропонуємо використовувати 17 балів як межу високого рівня психологічного дистресу, вимірюваного за допомогою SCL-9-NR і серед жінок.

Дискримінативний потенціал SCL-9-NR. Для демонстрації поділового потенціалу методики використано два варіанти класифікації з застосуванням дерев рішень⁷⁰. Перший варіант — це використання дерев класифікації на базі значень підсумкового (сумарного) індексу SCL-9-NR, другий — окремих індикаторів методики. За реалізації другого варіанта відповіді респондентів було перекодовано відповідно до дихотомічної шкали: 1) «ніколи» і «зрідка»; 2) «періодично» і «майже постійно».

Як спостереження використовували дані, зібрані в ЧОПНЛ і ХОКПЛ, а також результати онлайн-опитування 2016 року. Відповідно, респондентів було поділено на дві групи: з психіатричним діагнозом і без нього. Результати чоловіків і жінок аналізували окремо. Кожен із чотирьох підмасивів був випадковим чином розбитий на дві підгрупи, що дозволило сформулювати навчальні та контрольні дані для використання в деревах класифікації:

- чисельність жінок із 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 із діагнозом становила по 51 респонденту в контрольній і навчальній групах;
- чисельність чоловіків без діагнозу становила по 52 респонденти в контрольній і навчальній групах.

Загалом якість класифікації є вищою для жінок (табл. 8.18). Сумарний індекс відчутно поступається за якістю класифікації окремим індикаторам, особливо якщо йдеться про контрольні дані. З огляду на те, що методику SCL-9-NR розробляли не для виявлення психічних захворювань, а для вимірювання величини

⁷⁰ Аналіз здійснено за допомоги бібліотеки «tree» мови програмування R. Докладніше про дерева класифікації та їх реалізацію в R див.: (Джеймс, Уиттон, Хасті, & Тибширани, 2016, с. 328–363).

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

Таблиця 8.18

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

Тут слід зазначити, що стосовно чоловіків і жінок як набір предикторів, так і структура дерев відрізняються (рис. 8.6). Так, щодо чоловіків спрацьовують симптоматичні виміри obsесивно-компульсивного розладу, фобічної тривожності, соматизації та депресії, а щодо жінок — психотизму, депресії, фобічної тривожності та соматизації. У рамках цього дослідження це не стало несподіваним. Проте з точки зору загальноприйнятого підходу, який оцінює психологічний дистрес через депресію і тривожність, такі класифікаційні моделі мають, радше, спростовувальний характер.

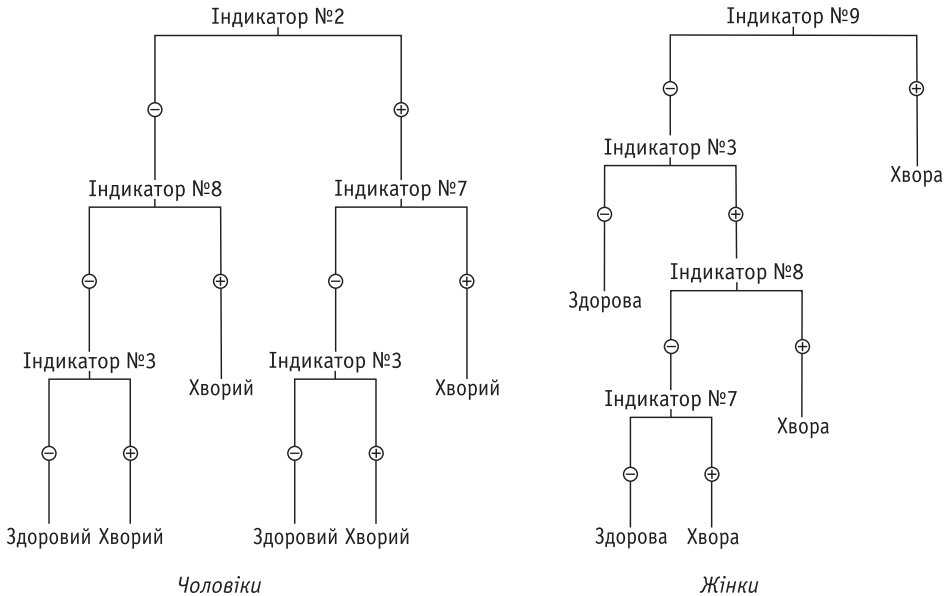


Рисунок 8.6. Класифікаційні дерева на підставі окремих індикаторів SCL-9-NR для чоловіків і жінок⁷¹

⁷¹ Плюси на рисунку позначають варіант відповіді «періодично» або «майже постійно» за відповідними індикаторами, мінуси — «ніколи» або «зрідка».

§ 8.6. Висновки

Методика SCL-9-NR є прийнятною для тих соціальних досліджень, у яких необхідно оцінити загальний рівень психологічного дистресу респондентів без звернення до більш специфічної індивідуальної інформації. При цьому рекомендовано використовувати межу 13 балів як показник підвищеного психологічного дистресу і 17 балів — як показник високого.

Щодо впливу методу збору даних на результати використання методики, то немає підстав вважати, що самозаповнення відрізняється від онлайн-опитування. Інформації про вплив на результати методики використання формату особистого інтерв'ю (face-to-face) на даний момент немає.

SCL-9-NR характерний високими показниками конструктивної та критеріальної валідності, а також внутрішньої узгодженості та ретестової надійності. На підставі даних, отриманих серед респондентів в Україні, можна зробити висновок, що SCL-9-NR — чутливіший інструмент вимірювання психологічного дистресу, ніж K10 P. Кесслера.

Глава 9

Застосування L-шкали за опитування населення України

§ 9.1. Вступ

Згідно з сучасною термінологією, помилки в опитувальних дослідженнях поділяють на помилки вибірки (sampling error) і помилки, не пов'язані з вибірковими процедурами (nonsampling error). До другої групи, приміром, належать вимірювальні помилки, помилки внаслідок невідповідей і дій інтерв'юєрів. Усі вони провокують ті або ті зсуви, що погіршують фінальні результати, покликані відображати генеральну сукупність (McNabb, 2014, р. 42–46).

Одним із важливих джерел виникнення помилок в опитувальних дослідженнях є запитання анкети, що викликають побоювання респондентів. Для оцінювання відповідних зсувів використовують спеціальні методики. Назагал їх можна поділити на тести, що оцінюють щирість респондентів (Crowne & Marlowe, 1960; Groth-Marnat, 2003, р. 244–245), а також більш загальні вимірювальні інструменти для оцінювання респондента як надійного джерела соціальної інформації (Головаха и др., 2008).

Ця частина роботи має на меті розгляд результатів використання L-шкали (шкали брехні) Міннесотського багатофазного особистісного переліку (MMPI), що належить до інструментів оцінки щирості респондентів. Цінність аналізу полягає в тому, що відповідні результати отримано в рамках репрезентативного опитування в Україні (Інститут соціології НАН України, 2006).

Особливості опитувальника дослідження, а також наявність результатів трьох додаткових методичних опитувань із застосуванням L-шкали дали змогу розглянути такі аспекти:

- показники надійності;
- демографічна специфіка відповідей на індикатори шкали;
- застосування вагових коефіцієнтів;
- рівні (не)щирості респондентів.

До згаданих опитувань належать такі: 1) омнібус «Українське суспільство» (Інститут соціології НАН України, 2006 рік, $N = 1800$, середній вік — 45 років, співвідношення чоловіків і жінок у вибірці — 45% і 55%); 2) онлайн-опитування серед користувачів соціальних мереж (2016 рік, $N = 380$, середній вік — 25

років, співвідношення чоловіків і жінок у вибірці — 28% і 72%)⁷²; 3) опитування серед пацієнтів Чернігівської обласної психоневрологічної лікарні (2016 рік, $N = 181$, середній вік — 44 роки, співвідношення чоловіків і жінок у вибірці — 50% і 50%); 4) Методичне опитування А. Мягкова (2002 рік, $N = 387$, дані про демографічні параметри відсутні)⁷³.

§ 9.2. Історична довідка

Історія створення шкали брехні безпосередньо пов'язана з розробкою ММРІ, незважаючи на те, що спочатку вона не була передбачена тестом. Конструювання ММРІ було розпочато Старком Хатусеєм і Джоном Маккінлі 1939 року в Міннесотському університеті. Їхньою головною метою було отримання інструмента, придатного для оцінки тяжкості розладів у дорослих пацієнтів у процесі психіатричної роботи з ними.

Використовуваний для розробки підхід можна назвати емпіричним або критеріальним. Спочатку було відібрано понад 1000 тверджень із різних джерел, що включали раніше розроблені шкали, клінічні звіти, історичні випадки, поради психіатричних інтерв'ю й особистий професійний досвід. Після перевірки цих тверджень на зрозумілість, читабельність, дублі та баланс між позитивними і негативними формулюваннями залишилося 504 твердження. Далі проведено опитування у вибірках звичайних респондентів і респондентів із психіатричним діагнозом. Це дозволило відібрати ті твердження, які показують відмінні розподіли відповідей у зазначених групах.

Надалі стало зрозуміло, що результати відповідей можуть зумовлюватися не тільки відмінностями у психічному стані респондентів, а й різними установками щодо заповнення тесту. З цієї причини автори ММРІ додатково розробили низку шкал, призначених для визначення типу і сили таких установок: «не можу сказати» (?-шкала), брехні (L-шкала), достовірності (F-шкала) і корекції (K-шкала). Головне призначення шкали брехні — визначити наївні зусилля випробуваного, спрямовані на створення надмірно сприятливого враження про себе.

Згодом ММРІ був підданий досить серйозній критиці, яка звертала увагу на такі проблеми тесту, як «старіння» змісту, вади ініціальної процедури побудови, недостатній рівень стандартизації тощо. З цієї причини 1989 року було здійснено ревізію й удосконалення методики, що привело до появи ММРІ-2. У шкалу брехні зміни не вносили (Groth-Marnat, 2003, p. 215–220).

§ 9.3. Показники надійності L-шкали

Передусім розглянемо розподіл відповідей за всіма показниками шкали з метою виявлення тих із них, що частіше спрацьовують як індикатори соціальної

⁷² Дослідження проведено разом із кафедрою методології та методів соціологічних досліджень КНУ ім. Т. Шевченка.

⁷³ Докл. див.: (Мягков, 2002).

бажаності (табл. 9.1). Як можна бачити, пункти шкали істотно різняться під кутом зору виявлення установки на соціально схвалювані відповіді. Звісно, якщо ми виходимо з того, що всі вони працюють саме на розв'язання цієї проблеми. Так, оскільки ММРІ розроблений понад 70 років тому, можна припустити, що окремі показники з L-шкали змінили свій смисловий зміст. Ще однією ймовірною причиною, що впливає на розподіл відповідей, може бути особливість формулювань різних індикаторів.

Таблиця 9.1

Кількість відповідей «Згоден» за індикаторами L-шкали, %⁷⁴

№	Індикатор	ІС НАНУ	Онлайн	ЧОПНЛ
1	Часом мені на думку спадають такі погані думки, що про них краще не розповідати	57,5	66,0	61,9
2	Іноді мені хочеться вилаятися	78,3	80,1	61,1
3	Я не завжди кажу правду	61,5	81,5	59,1
4	Я не щодня читаю передовиці в газетах	81,6	78,2	84,7
5	Іноді я буваю сердитим(-ою)	89,1	92,5	77,5
6	Іноді я відкладаю на завтра те, що можна зробити сьогодні	77,6	95,0	70,2
7	Іноді, коли я погано почуваюся, буваю дратівливим(-ою)	83,6	93,6	76,0
8	Мої манери за столом у себе вдома не такі добрі, як у гостях	46,8	52,8	42,0
9	Я переходжу вулицю в недозволеному місці, коли впевнений(-а), що мене не помітить міліціонер	50,5	38,7	43,8
10	У грі мені приємніше вигравати, ніж програвати	80,4	81,2	83,8
11	Мені подобається мати поважних людей серед моїх знайомих, бо це підвищує мій престиж	55,9	44,5	64,4
12	Не всі, кого я знаю, мені подобаються	85,7	93,6	82,1
13	Іноді я можу трохи попліткувати	66,8	80,4	57,1
14	Іноді я даю добру оцінку людям, про яких знаю дуже мало	55,1	77,3	66,7
15	Трапляється, що я сміюся з непристойного жарту	65,4	89,5	54,4
Середнє значення сумарного індексу		4,64	3,55	5,02

Разом із тим на цьому етапі приймемо робочу гіпотезу, відповідно до якої всі твердження спрямовані на виявлення установки на соціально схвалювані відповіді в перебігу опитування, тобто мають за основу спільну латентну змінну. Лише в такому разі є сенс у перевірці надійності L-шкали в частині її внутрішньої узгодженості.

⁷⁴ Тут і далі в таблицях використано такі скорочення: ІС НАНУ — Омнібус «Українське суспільство», Онлайн — Онлайн-опитування серед користувачів соціальних мереж, ЧОПНЛ — Опитування серед пацієнтів Чернігівської обласної психоневрологічної лікарні, Мягков — методичне опитування А. Мягкова.

Використовуючи підхід, запропонований у статті А. Мягкова (Мягков, 2002, с. 121–123), проаналізуємо взаємозв'язок пунктів шкали один з одним із застосуванням коефіцієнта контингенції, взаємозв'язок кожного пункту з сумарним індексом L-шкали на підставі Ета-коефіцієнта, внутрішню узгодженість на підставі коефіцієнта Альфа Кронбаха й узгодженість двох частин шкали на підставі коефіцієнта Спірмена—Брауна.

Взаємозв'язок пунктів L-шкали. У своїй праці А. Мягков навів усю кореляційну матрицю, а також підрахував середню силу зв'язку з іншими пунктами для кожного індикатора і зазначив щодо кожного з них кількість статистично значимих взаємозв'язків. Тут ми обмежуємося лише середньою силою зв'язку і кількістю статистично значимих відношень (табл. 9.2).

Таблиця 9.2

Взаємозв'язок пунктів L-шкали з сумарним індексом

№ індикатора	Середнє значення коефіцієнта контингенції				Кількість значимих зв'язків			
	ІС НАНУ	Онлайн	ЧОПНЛ	Мягков	ІС НАНУ	Онлайн	ЧОПНЛ	Мягков
1	0,154	0,093	0,151	0,131	12	7	10	7
2	0,207	0,113	0,173	0,110	14	9	9	7
3	0,214	0,132	0,114	0,122	14	10	6	8
4	0,193	0,046	0,089	0,055	14	3	3	4
5	0,184	0,120	0,121	0,094	11	9	4	4
6	0,186	0,042	0,130	0,131	14	3	5	8
7	0,201	0,112	0,157	0,075	14	7	8	4
8	0,188	0,098	0,115	0,090	13	7	5	5
9	0,179	0,100	0,169	0,102	12	8	10	7
10	0,170	0,096	0,118	0,054	13	9	6	2
11	0,182	0,077	0,071	0,063	14	7	3	3
12	0,162	0,068	0,121	0,053	12	4	8	1
13	0,175	0,110	0,175	0,132	14	9	9	3
14	0,160	0,069	0,115	0,101	12	4	5	7
15	0,212	0,097	0,176	0,097	14	5	12	4
Середнє	0,185	0,092	0,133	0,094				

Щоб забезпечити коректність порівняння кількості значимих зв'язків, розмір вибірки соціологічного моніторингу «Українське суспільство» було узгоджено⁷⁵ з відповідним параметром дослідження А. Мягкова ($N = 387$). Для цього з масиву моніторингового дослідження виокремлено вибірку необхідного розміру, на підставі якої здійснювався розрахунок коефіцієнтів контингенції.

⁷⁵ Для цього використано випадковий відбір.

Отже, результати соціологічного моніторингу загалом показують тісніший взаємозв'язок між пунктами шкали, і в переважній більшості випадків вони є статистично значимими. Далі, якщо говорити тільки про силу зв'язку, йде опитування серед пацієнтів психоневрологічної лікарні. Порівнювати тут кількість статистично значимих відношень важко з огляду на меншу вибірку цього опитування. Але, навіть попри це, як мінімум три пункти (№ 1, 9, 15) показують доволі велику кількість статистично значимих взаємозв'язків. Решта два опитування демонструють достатньо посередні результати взаємозв'язаності різних пунктів шкали.

Взаємозв'язок пунктів L-шкали з сумарним індексом. З урахуванням попередніх результатів очікуваним результатом розрахунку Ета-коефіцієнтів були їх вищі значення для моніторингового масиву (табл. 9.3). Оскільки в дослідженні А. Мягкова всі зв'язки були статистично значимими, для розрахунків за моніторинговими даними ми використовуємо усі спостереження.

Таблиця 9.3

Взаємозв'язок пунктів L-шкали з сумарним індексом

№ індикатора	ІС НАНУ	Онлайн	ЧОПНЛ	Мягков
1	0,479	0,454	0,478	0,482
2	0,529	0,420	0,533	0,226
3	0,553	0,485	0,389	0,380
4	0,422	0,217	0,297	0,214
5	0,456	0,383	0,356	0,155
6	0,507	0,185	0,445	0,401
7	0,471	0,377	0,469	0,282
8	0,521	0,465	0,407	0,417
9	0,488	0,474	0,531	0,401
10	0,434	0,420	0,370	0,180
11	0,432	0,367	0,277	0,300
12	0,388	0,269	0,374	0,224
13	0,486	0,449	0,554	0,433
14	0,454	0,357	0,412	0,487
15	0,558	0,340	0,546	0,374
Середнє значення	0,479	0,377	0,429	0,330

Дані моніторингового опитування, а також опитування серед пацієнтів психоневрологічної лікарні показують доволі непогані результати. Результати онлайн-опитування можна назвати прийнятними, натомість опитування А.Мягкова дає вельми слабкий результат.

Внутрішня узгодженість (Альфа Кронбаха). Якщо в опитуванні А. Мягкова величина коефіцієнта Альфа Кронбаха була встановлена на рівні лишень 0,44, то в онлайн-опитуванні вона становила 0,57, в опитуванні пацієнтів психоневрологічної лікарні — 0,70, а в моніторинговому дослідженні «Українське суспільство» — 0,76. Результати знову ж таки вельми неоднозначні.

Узгодженість двох частин шкали (коефіцієнт Спірмена—Брауна). Схожу з внутрішньою узгодженістю картину спостерігаємо і за обчислення показників надійності в рамках моделі *split-half*. Найвищі результати зафіксовано для моніторингового опитування (0,79), дещо нижчі — для опитування пацієнтів психоневрологічної лікарні (0,74), посередні — для результатів онлайн-опитування (0,56), найнижчі — для опитування А. Мягкова (0,33).

Висновки і припущення. Як показують результати аналізу, надійність L-шкали істотно варіює від одного дослідження до іншого. Це, своєю чергою, не дає підстав для однозначних висновків про її якість за даною характеристикою.

Водночас слід замислитися щодо необхідності перевірки внутрішньої надійності L-шкали. Згідно з принципами соціологічного тестування, котрі ми запропонували в першому розділі, такий аналіз має сенс тільки в тому разі, коли йдеться про шкалу. Якщо ж використано індекс, то оцінювати внутрішню надійність немає сенсу. Особисто ми схилиємося до того, що L-шкала є інтегральним індексом. Наша аргументація така.

Результати відповідей за індикаторами методики визначають не латентною змінною, а набором установок щодо соціальної самопрезентації в сучасному суспільстві. Джерелом таких установок є цілий комплекс факторів, що змінюється як від однієї соціальної групи до іншої, так і для конкретного індивіда у процесі його статусних змін і набуття нового соціального досвіду. Отже, говорити про якусь латентну змінну тут проблематично. Якщо це припущення є правильним, тоді розподіл відповідей за різними індикаторами L-шкали має показати варіативність з урахуванням різних демографічних груп.

§ 9.4. Демографічна специфіка відповідей за індикаторами L-шкали

Відповідно до інструкції ММРІ, згоду з твердженнями L-шкали слід розглядати як ширі відповіді. Виходячи з цього, можна оцінити міру ширості за індикаторами шкали брехні в основних демографічних групах (табл. 9.4). Загалом результати доволі варіативні. Найменшу ширість зафіксовано для жінок похилого віку за індикатором «Я переходжу вулицю в недозволеному місці, коли впевнений(-а), що мене не помітить міліціонер» (39,3%), найбільшу — для жінок середнього віку за індикатором «Іноді я буваю сердитим(-ою)» (91,4%). Середній рівень згоди становить 69,3% за стандартного відхилення 14,8%.

Таблиця 9.4

Кількість відповідей «згоден» на твердження L-шкали, %

Вікова група	Стать	
	Чоловіча	Жіноча
1) Часом мені на думку спадають такі погані думки, що про них краще не розповідати		
Молодь (до 29 років)	53,5	56,1
Середній вік (від 30 до 54)	61,8	57,7
Похилий вік (55 і старші)	54,5	57,4

Вікова група	Стать	
	Чоловіча	Жіноча
<i>2) Іноді мені хочеться вилаятися</i>		
Молодь (до 29)	80,3	80,1
Середній вік (від 30 до 54)	85,4	77,6
Похилий вік (55 і старші)	80,9	67,6
<i>3) Я не завжди кажу правду</i>		
Молодь (до 29)	71,7	69,4
Середній вік (від 30 до 54)	66,8	59,3
Похилий вік (55 і старші)	60,0	49,3
<i>4) Я не щодня читаю передовиці в газетах</i>		
Молодь (до 29)	81,8	83,2
Середній вік (від 30 до 54)	81,7	82,8
Похилий вік (55 і старші)	81,4	79,1
<i>5) Іноді я буваю сердитим(-ою)</i>		
Молодь (до 29)	86,9	90,3
Середній вік (від 30 до 54)	90,8	91,4
Похилий вік (55 і старші)	88,2	85,7
<i>6) Іноді я відкладаю на завтра те, що можна зробити сьогодні</i>		
Молодь (до 29)	80,8	82,7
Середній вік (від 30 до 54)	81,0	75,8
Похилий вік (55 і старші)	78,2	70,9
<i>7) Іноді, коли я погано почуваюся, я буваю дратівливим(-ою)</i>		
Молодь (до 29)	81,2	86,8
Середній вік (від 30 до 54)	82,3	86,5
Похилий вік (55 і старші)	82,7	81,6
<i>8) Мої манери за столом у себе вдома не такі добрі, як у гостях</i>		
Молодь (до 29)	47,2	42,3
Середній вік (від 30 до 54)	50,4	45,2
Похилий вік (55 і старші)	48,2	45,9
<i>9) Я переходжу вулицю в недозволеному місці, коли впевнений(-а), що мене не помітить міліціонер</i>		
Молодь (до 29)	56,6	52,6
Середній вік (від 30 до 54)	57,1	51,5
Похилий вік (55 і старші)	48,2	39,3

Закінчення таблиці 9.4

Вікова група	Стать	
	Чоловіча	Жіноча
<i>10) Угрі мені приємніше вигравати, ніж програвати</i>		
Молодь (до 29)	81,3	81,6
Середній вік (від 30 до 54)	86,2	76,7
Похилий вік (55 і старші)	84,5	74,8
<i>11) Мені подобається мати значимих людей серед моїх знайомих, бо це підвищує мій престиж</i>		
Молодь (до 29)	66,7	54,6
Середній вік (від 30 до 54)	61,3	55,7
Похилий вік (55 і старші)	47,3	50,4
<i>12) Не всі, кого я знаю, мені подобаються</i>		
Молодь (до 29)	86,4	86,3
Середній вік (від 30 до 54)	86,2	87,4
Похилий вік (55 і старші)	85,0	83,0
<i>13) Іноді я можу трохи попліткувати</i>		
Молодь (до 29)	53,5	82,7
Середній вік (від 30 до 54)	55,5	80,2
Похилий вік (55 і старші)	49,8	72,0
<i>14) Іноді я даю добру оцінку людям, про яких знаю дуже мало</i>		
Молодь (до 29)	57,6	65,5
Середній вік (від 30 до 54)	50,8	56,1
Похилий вік (55 і старші)	53,2	52,5
<i>15) Буває, що я сміюся з непристойного жарту</i>		
Молодь (до 29)	79,3	70,6
Середній вік (від 30 до 54)	75,1	61,5
Похилий вік (55 і старші)	63,6	50,0

На підставі наведених даних індикатори L-шкали можна згрупувати за схожістю. Для виконання цього завдання застосовано ієрархічний кластерний аналіз. Як вхідні дані було використано шість змінних — рівень згоди з твердженнями методики для кожної з шести демографічних груп. За результатами кластерного аналізу було отримано 6 груп із такими центроїдами (табл. 9.5).

Таблиця 9.5

Центроїди для груп на підставі тверджень L-шкали, %

Гру- па	Демографічна група						Твер- дження
	Ч., до 29	Ч., 30–54	Ч., від 55	Ж., до 29	Ж., 30–54	Ж., від 55	
1	58,6	57,8	50,8	57,2	55,3	49,9	1,9,11,14
2	80,8	84,2	81,2	81,5	76,7	71,1	2,6,10
3	75,5	71,0	61,8	70,0	60,4	49,7	3,15
4	84,1	85,3	84,3	86,7	87,0	82,4	4,5,7,12
5	47,2	50,4	48,2	42,3	45,2	45,9	8
6	53,5	55,5	49,8	82,7	80,2	72,0	13

Перш ніж інтерпретувати отримані результати, слід наголосити важливий аспект використання методики. Фактично високий рівень згоди за тим чи тим індикатором серед загального населення вказує на те, що не погоджуватимуться з ним респонденти з високим рівнем психологічного захисту або сильним бажанням презентувати себе з найкращого боку. Зворотне буде справедливим для тверджень із нижчим рівнем згоди — вони виявляють слабші й дуже поширені механізми психологічного захисту.

Повертаючись до характеристик отриманих груп, треба сказати, що узагальнено їх можна поділити на дві підмножини. Перша складається з груп, що показують однакові патерни згоди з твердженнями серед чоловіків і жінок (групи № 1, 4, 5). Друга підмножина містить групи, в яких чоловіки і жінки по-різному реагують на індикатори методики (групи № 2, 3, 6).

Розгляньмо кожну групу докладніше. Насамперед звернімо увагу на групу № 5. Відповідне твердження отримує найменший рівень згоди респондентів, а отже, виявляє найпоширеніші механізми психологічного захисту. Рівень згоди з ним серед чоловіків різних вікових груп приблизно однаковий, те саме є справедливим для жіночої частини відповідей. При цьому рівень згоди для жінок дещо нижчий порівняно з чоловіками. Загалом достатньо висока гомогенність рівнів згоди в різних демографічних групах свідчить про універсальність восьмого твердження як індикатора схильності презентувати себе з найкращого боку.

Група № 1 також показує дуже низький рівень згоди як серед чоловіків, так і серед жінок. Специфікою цієї групи є те, що рівень згоди трохи падає для літніх респондентів (і чоловіків, і жінок). Схожість відповідей чоловіків і жінок також дає підстави говорити про універсальність тверджень цієї групи.

На протигагу групам № 1 і 5, твердження групи № 4 показують високий рівень згоди, що є свідченням придатності їх до виявлення найсильніших механізмів психологічного захисту. На користь універсальності відповідних тверджень свідчать мінімальні розбіжності в різних демографічних групах.

Решта груп є специфічними щодо рівня згоди з індикаторами L-шкали в різних групах. Група № 2 показує високу згоду для всіх груп чоловіків. Для жінок цей показник також доволі виражений, але падає зі збільшенням віку.

Група № 3 показує назагал трохи вищу згоду серед чоловіків, а також зменшення рівня згоди зі збільшенням віку. Група № 6 узагалі показує низьку згоду серед чоловіків і високу серед жінок. Як бачимо, образ «пліткаря» є значно неприємнішим, аніж «пліткарки».

Якщо погодитися з поділом тверджень методики на універсальні та специфічні, то поряд із використанням повного варіанта шкали брехні вимальовується її скорочена версія на підставі менш чутливих до особливостей різних демографічних груп індикаторів, що включають твердження № 1, 4, 5, 7, 8, 9, 11, 12, 14.

Крім того, дотримуючись логіки поділу механізмів психологічного захисту на сильні та слабкі, а також спираючись на демографічні відмінності, можна розрахувати для кожного індикатора диференційовані вагові коефіцієнти. Це, своєю чергою, дасть змогу побудувати альтернативну версію шкали брехні з урахуванням відповідей за всіма індикаторами методики.

§ 9.5. Вагові коефіцієнти індикаторів L-шкали

Вагові коефіцієнти слід присвоювати так, щоб сильні механізми психологічного захисту робили більший порівняно зі слабкими внесок у підсумковий індекс. Для виконання цього завдання можна використовувати результати, наведені в таблиці 9.4. У цьому разі кожен індикатор має шість можливих вагових коефіцієнтів, застосовуваних відповідно до демографічної категорії того чи того респондента (на підставі статі та вікової групи). Як розміри вагових коефіцієнтів пропонуємо використовувати частку відповідей «згоден» у тій демографічній групі, до якої належить респондент.

Отже, якщо в оригінальній методиці за кожну незгоду респондент отримує один бал за L-шкалою, то за використання вагових коефіцієнтів він отримує кількість балів, що дорівнює самому ваговому коефіцієнту.

Звісно, використання зваженого індексу L-шкали має сенс тільки тоді, коли він дає змогу оцінювати щирість респондентів ефективніше, ніж оригінальний. Тому далі здійснено спробу порівняння чотирьох індексів — оригінального без зважування, скороченого без зважування, оригінального після зважування, скороченого після зважування. Розрахунок зважених коефіцієнтів здійснено з застосуванням мови програмування R (див. Додаток 2).

Оскільки ми були обмежені змінними використовуваного масиву, нам не вдалося продемонструвати тісних взаємозв'язків між відповідями на запитання, що можуть викликати побоювання респондентів, та індексами на підставі індикаторів L-шкали. Як відповідні змінні використано оцінку важливості базових цінностей (матеріальний добробут, міцне здоров'я і міцна сім'я), купівля спиртних напоїв до свят та інших знаменних дат, а також ворожість (вимірювана за підшкалою ворожості з психологічного тесту SCL-90-R)⁷⁶ (див.

⁷⁶ Ідеться про тест для вимірювання психологічного дистресу, автором якого є Л. Дерогатіс (Derogatis & Fitzpatrick, 2004).

Додаток 3). Але перш ніж ставати до аналізу відповідних результатів, розглянемо як різноманітні варіанти індексів, залежні один від одного.

Узаємозв'язок між різноманітними індексами очікувано є дуже високим (табл. 9.6). Особливо це справедливо щодо сили зв'язку між скороченими і повними версіями.

Таблиця 9.6

Взаємозв'язок між значеннями різноманітних варіантів L-шкали, r

Варіант індексу	CI	3OI	3CI
Оригінальний індекс (OI)	0,93	0,99	0,91
Скорочений індекс (CI)		0,91	0,98
Зважений оригінальний індекс (3OI)	0,91		0,92
Зважений скорочений індекс (3CI)	0,98	0,92	

З огляду на такий тісний взаємозв'язок важливо розуміти, чи виправдане скорочення і/або зважування під кутом зору завдань, що їх має виконувати L-шкала.

*L-шкала і базові цінності*⁷⁷. Заперечення важливості матеріального добробуту, міцного здоров'я і міцної сім'ї навряд чи може вказувати на бажання презентувати себе з найкращого боку. Найімовірніше головною причиною є приховані проблеми і небажання говорити правду самому собі, що, своєю чергою, має накладати відбиток на сприйняття навколишньої дійсності й на те, як на неї реагують (табл. 9.7).

Таблиця 9.7

Взаємозв'язок значень різноманітних варіантів L-шкали з оцінкою важливості базових цінностей, r

Цінність	OI	CI	3OI	3CI
Матеріальний добробут	-0,130	-0,129	-0,150	-0,153
Міцне здоров'я	-0,148	-0,140	-0,173	-0,174
Міцна сім'я	-0,131	-0,132	-0,148	-0,156

Цікавим результатом є те, що зважені індекси показують тісніший взаємозв'язок, ніж незважені. Пояснити це можна тим, що більшої ваги набувають твердження, пов'язані з сильнішими механізмами психологічного захисту. Саме останні мають спрацьовувати, коли респондент низько оцінює важливість базових цінностей.

⁷⁷ Для фіксації відповідей про важливість різноманітних цінностей використано шкалу з такими варіантами відповідей: «зовсім не важливо», «скоріше не важливо», «важко відповісти, важливо чи ні», «скоріше важливо», «дуже важливо».

*L-шкала і купівля спиртних напоїв*⁷⁸. На відміну від попереднього випадку, купівля спиртних напоїв є типовим прикладом поведінки, що може викликати побоювання за заповнення анкет (табл. 9.8).

Таблиця 9.8

Взаємозв'язок значень різноманітних варіантів L-шкали з наявністю факту купівлі спиртних напоїв, r

Змінна	OI	CI	ZOI	ZCI
Купівля спиртних напоїв за останні 12 місяців	-0,189	-0,136	-0,184	-0,133

Тут більша вага для окремих індикаторів не спрацьовує, позаяк розмова про спиртні напої активізує не сильні, а звичайні або вельми поширені механізми психологічного захисту. Тому найрізноманітніші групи респондентів однаковою мірою сприймають запитання як незручне. Логічним наслідком є те, що велика кількість індикаторів веде до більшої ефективності індексу.

L-шкала і ворожість. Говорити про свою ворожість не прийнято. Тому тут мають спрацьовувати поширені механізми психологічного захисту. Ще одним важливим моментом, на який слід зважати в цьому разі, є специфіка формулювань, використовуваних у SCL-90-R. Як ми показали раніше, вони майже не придатні для соціологічних опитувань, бо провокують зсуви відповідей до менших градацій шкали, знижуючи тим самим підсумкове значення за методикою загалом та її підшкалами зокрема. У цьому сенсі підшкала ворожості не є винятком (рис. 9.1).

Тому поряд із використанням усіх спостережень розглянемо результати кореляційного аналізу для підвибірки, з якої вилучено респондентів із нульовою ворожістю (табл. 9.9). Як можна бачити, тут вагові коефіцієнти «грають» не на користь якості індексу. Не рятує ситуацію навіть кількість індикаторів. Якщо ж говорити про результати аналізу підвибірки, з якої вилучено респондентів з нульовою ворожістю, то сила зв'язку відчутно слабшає. При цьому зважені індекси втрачають у силі взаємозв'язку більше, ніж прості.

Таблиця 9.9

Взаємозв'язок значень різноманітних варіантів L-шкали з підшкалою ворожості SCL-90-R, r

Змінна	OI	CI	ZOI	ZCI
Ворожість (уся вибірка)	-0,168	-0,170	-0,152	-0,147
Ворожість (підвибірка без респондентів із нульовою ворожістю)	-0,081	-0,102	-0,056	-0,068

⁷⁸ Для фіксації відповідей про купівлю спиртних напоїв використано дихотомічну шкалу з такими варіантами відповідей: «так», «ні».

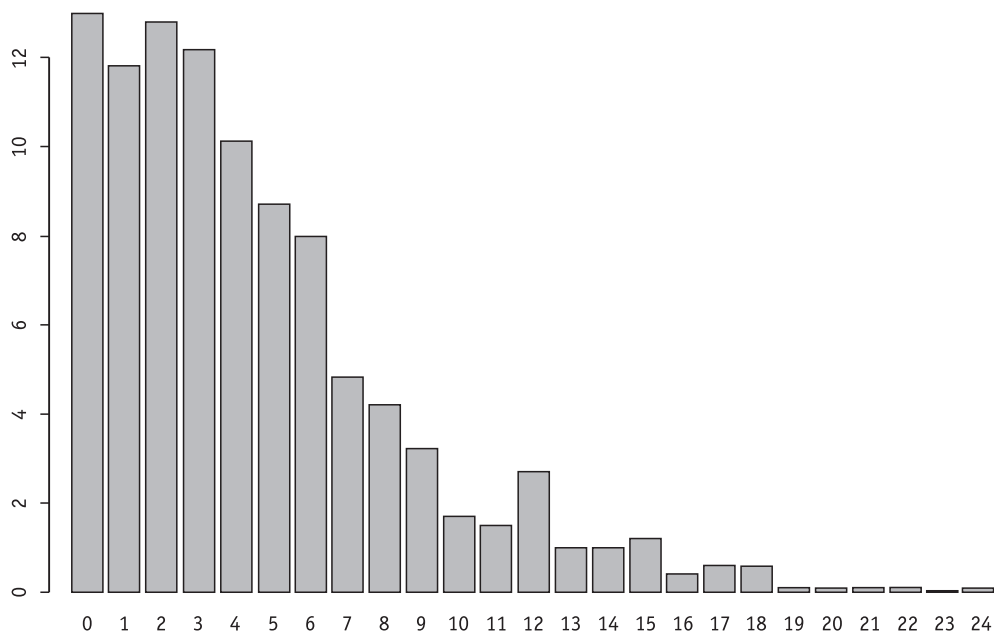


Рисунок 9.1. Розподіл сумарних значень для підшкали ворожості (SCL-90-R), %

Які висновки можна тут зробити? По-перше, респонденти з нульовою ворожістю багато чого про себе приховують. Принаймні, коли йдеться про використання SCL-90-R. По-друге, виходячи з характеру зменшення сили взаємозв'язків, у групі респондентів із нульовою ворожістю переважають потужні механізми психологічного захисту.

Наступним логічним кроком в аналізі відповідей на твердження L-шкали є визначення рівня, що відокремлює достатньо ширих респондентів від тих, котрі такими не є.

§ 9.6. Рівень (не)щирості респондентів

Відповідно до загальних рекомендацій, високі значення за L-шкалою відповідають межі 66 балів, розрахованих на підставі T-показника⁷⁹. Слід зважати й на те, що для різних соціальних груп межа нещирості варіює. Так, для низькокваліфікованих груп межа буде вищою, ніж для висококваліфікованих. Те саме справедливе для вищого статусу зайнятості порівняно з нижчим. Загалом високі значення за L-шкалою свідчать про те, що респондент описує себе в дуже позитивному світлі. Це може бути результатом як свідомого обману, так і нереалістичної самооцінки (Groth-Marnat, 2003, p. 244–245).

⁷⁹ T-показник (*T-score*) — стандартизований бал, що має розподіл із середнім значенням 50 і стандартним відхиленням 10.

У використуваному масиві⁸⁰ середнє значення для шкали брехні дорівнює 4,65, а стандартне відхилення — 3,16. Отже, Т-показник на рівні 66 балів приблизно дорівнює 10 «сирим» балам. Кількість респондентів, що перебувають на цьому рівні або перевищують його, дорівнює 7,2%. Щодо зваженої шкали (середнє — 2,88; стандартне відхилення — 2,10), то рівень високої нещирості починається з 6,24 «сирих» балів, а відповідна частка респондентів дорівнює 7,3%. Загальна величина розбіжності в результатах двох класифікацій респондентів із використанням зазначеного Т-показника становить 1,6%.

Щоб проаналізувати кваліфікаційно-статусні розбіжності, можна звернутися до даних про основний вид занять респондента (табл. 9.10). У відповідних результатах не простежується залежність рівня ширості респондентів від їхньої статусної належності. Наприклад, тимчасово безробітні й індивіди, які працюють за наймом, показують практично однакові середні значення за L-шкалою, а студенти технікумів, ПТУ і коледжів більш ширі за студентів ВНЗ⁸¹.

Крім того, результати зваженого й оригінального індексів важко назвати достатньо узгодженими. Першою чергою це стосується ширості респондентів, які отримують пенсію за віком чи інвалідністю. За використання оригінального індексу вони потрапляють до числа найменш ширих груп, натомість за застосування вагових коефіцієнтів опиняються в середині ранжованого списку.

Таблиця 9.10

**Середнє значення індексу для повних варіантів L-шкали
залежно від виду зайнятості респондента**

<i>Основний вид діяльності</i>	<i>ОІ</i>	<i>ЗОІ</i>	<i>n</i>
Навчання на денному відділенні технікуму, ПТУ, коледжі	3,27	2,11 (1)	15
Домашнє господарство або відпустка з нагляду за дитиною	4,16	2,57 (2)	92
Тимчасова відсутність роботи	4,22	2,71 (4)	149
Робота за наймом на постійних засадах	4,29	2,70 (3)	776
Навчання у ВНЗ	4,46	2,89 (5)	72
Періодична робота за контрактом, трудовою угодою, неповний робочий день	4,86	3,15 (7)	59
Індивідуальна трудова діяльність (дрібне підприємництво, «човникарство» тощо)	4,90	3,18 (8)	38
Приватне підприємництво, бізнес (включно з фермерським)	5,19	3,38 (9)	31
Пенсія за віком або інвалідністю	5,27	3,10 (6)	486
Важко відповісти	6,22	4,11 (10)	32
Силкові структури (армія, міліція (нині — поліція), СБУ)	7,71	5,24 (11)	7

⁸⁰ Мається на увазі масив соціологічного моніторингу «Українське суспільство» за 2006 рік.

⁸¹ Тут слід зазначити, що ймовірність помилковості такого висновку становить 10% для двостороннього критерію і 5% — для одностороннього ($t = -1,7$, $df = 28,4$).

Ці дані не так спростовують залежність результатів використання шкали брехні від кваліфікаційно-статусних характеристик індивідів, як показують неприйнятність її спрощеного використання, а також необхідність докладнішого вивчення.

З огляду на достатню наповненість ($n > 50$) шести з наведених вище груп (див. табл. 9.10) можна визначити для них межу високих показників L-шкали з використанням описаного вище підходу⁸² (табл. 9.11).

Таблиця 9.11

Межа високих значень за L-шкалою для деяких соціальних груп

<i>Основний вид діяльності</i>	<i>OI</i>	<i>ZOI</i>
Домашнє господарство або відпустка з нагляду за дитиною	9	5,60
Тимчасова відсутність роботи	9	5,94
Робота за наймом на постійних засадах	9	5,89
Навчання у ВНЗ	10	6,66
Періодична робота за контрактом, трудовою угодою, неповний робочий день	10	6,35
Пенсія за віком або інвалідністю	11	6,61

Використовуючи оригінальну шкалу, оцінімо вплив цих диференційованих меж на показник кількості нещирих респондентів (стосовно груп, для котрих межі не розраховані, використано загальну межу 10 балів). Для цього скористаймося змінними щодо купівлі алкоголю (табл. 9.12) і ворожості (табл. 9.13). У першому випадку до уваги взято кількість респондентів, які зазначили купівлю алкоголю, у другому — тих, хто артикулює відсутність будь-якої ворожості.

Таблиця 9.12

Вплив способу встановлення високих значень за L-шкалою на розподіл відповідей про купівлю алкоголю, %

<i>Межа високих значень за L-шкалою</i>	<i>Купували алкоголь</i>	
	<i>Прийнятний рівень за L-шкалою</i>	<i>Високий рівень за L-шкалою</i>
Загальна (≥ 10)	47,2	25,5
Диференційована ($\geq 9 / 10 / 11$)	47,1	32,4

Як бачимо, використання диференційованих меж робить достатньо і недостатньо ширі групи більше схожими одна на одну⁸³. Та хоч там як, дати відповідь на запитання про рівень нещирості можна однозначно тільки шляхом спеціальних методичних досліджень.

⁸² T-значення ≥ 66 .

⁸³ Усі розбіжності статистично значимі на рівні $p < 0,001$.

Таблиця 9.13

**Вплив способу встановлення високих значень за L-шкалою
на розподіл відповідей про ворожість, %**

<i>Межа високих значень за L-шкалою</i>	<i>Заперечення будь-якої ворожості</i>	
	<i>Прийнятний рівень за L-шкалою</i>	<i>Високий рівень за L-шкалою</i>
Загальна (≥ 10)	10,6	44,1
Диференційована ($\geq 9 / 10 / 11$)	10,5	36,6

§ 9.7. Висновки

Оцінка використання L-шкали в різноманітних опитуваннях істотно варіює залежно від того, до якого типу вимірювальних інструментів ми її відносимо — вимірювальних шкал чи індексів. Особисто ми схилиємося до другого варіанта. По-перше, це знімає проблему внутрішньої надійності інструмента. По-друге, головна увага з інструменту загалом переходить на його окремі індикатори і те, як їх сприймають різні групи респондентів. Це, своєю чергою, відкриває можливість побудови зваженого варіанта індексу.

Стосовно до L-шкали зважування передбачає різноманітні градації нещирості респондентів — починаючи з «узвичаєної нещирості», коли респондент дотримується мінімальних стандартів позитивної самопрезентації, і до її патологічних форм, коли індивід бажає виглядати як годиться навіть там, де це видається неприродним. Назагал проблема виявлення сильних і слабких механізмів психологічного захисту в масових опитуваннях потребує подальшого дослідження й осмислення. При цьому найактуальнішим напрямком тут бачимо поділ індикаторів щирості на ті, що виявляють схильність добре подавати себе перед іншими, і ті, що виявляють схильність добре виглядати перед самим собою. Мабуть, саме друга група працюватиме ефективно, незалежно від методу збирання даних.

Ще одним нетривіальним питанням щодо використання L-шкали є визначення норм або критеріїв, які дають змогу успішно класифікувати респондентів, схильних давати нещирі відповіді. Загальні рекомендації стосовно використання T-значення на рівні 66 балів виглядають доволі формальними, а також відірваними від дослідницького і соціального контексту. Тут знову ж таки потрібні спеціальні методичні дослідження⁸⁴.

⁸⁴ Слід зазначити, що методичне дослідження, проведене Мягковим (Мягков, 2002, с. 123–126), яке, крім іншого, заторкнуло питання ефективності класифікації респондентів за L-шкалою, показало невтішні результати. Мається на увазі, що ширі та неширі респонденти не відрізнялися за показниками L-шкали.

Глава 10

Інтегральний індекс соціального самопочуття: альтернативний підхід

§ 10.1. Вступ

Інтегральний індекс соціального самопочуття (далі — ПСС) був представлений соціологічній спільноті 20 років тому (Головаха & Панина, 1997). Зараз можна з упевненістю сказати, що Є. Головаха і Н. Панина випередили свій час. Так, з-поміж 83 посилань на статтю «Вимірювання соціального самопочуття: тест ПСС» (Головаха, Панина, & Горбачик, 1998), зафіксованих Google Scholar, лише 15 належать періоду до 2010 року, а найбільша кількість їх (усього 17) узагалі припала на 2016 рік.

З чим пов'язаний цей відтермінований методологічний інтерес? Насамперед збільшився попит на дослідження соціального самопочуття населення, а також його окремих груп, а отже й на відповідні методики. Крім цього, емпірична соціологія у країнах СНД, нарешті, вийшла на рівень, коли валідні та надійні комплексні вимірювальні інструменти в соціології стають постійним, а не епізодичним предметом методологічної рефлексії. Можна виокремити дві групи праць, у яких згадується ПСС — практичної та методологічної спрямованості. У перших описано результати вивчення соціального самопочуття різних соціальних груп, у других — концептуальне осмислення відповідного феномену і/або особливостей його вимірювання.

Праці, в яких фіксується практичний інтерес до ПСС, описують результати дослідження соціального самопочуття населення на Північному Кавказі (взятого як індикатор напруженості) (Гриценко, 2010), молоді етнічних груп (Фатхова, 2010), населення монопрофільних міст (у тому числі й як показник адаптованості до соціально-економічних трансформацій) (Гушина, Кондратович, & Положенцева, 2011; Кашкина, 2012), студентів з особливими потребами (Козлова, 2012), сільського населення північного регіону Росії (Белоножко & Барбаков, 2013), сім'ї (у контексті взаємозв'язку соціології та соціальної роботи) (Колесникова & Шепелева, 2015), працівників кримінально-виконавчої системи (як фактор соціально-фахової діяльності та підвищення професійної мотивації) (Волков, 2015; Волков, 2016), молоді в умовах суспільства, що трансформується (Хамзяєва, 2016), населення Республіки Башкортостан (Хилажева, Леготин, Ягафарова, & Исянгулова, 2017), людей похилого

віку (Бурко & Волкова, 2017), мешканців Луцька (Сальнікова, 2017) та інших соціальних груп.

Що стосується концептуального і методологічного розгляду соціального самопочуття в соціології, то відповідні праці, м'яко кажучи, не є дуже численними. Це дві найбільш цитовані статті відповідної тематики останніх семи років (Чугуєнко & Бобкова, 2013, с. 17–18; Зінурова & Фатыхова, 2011, с. 248). В. Чугуєнко і Є. Бобков вважають, що в ПСС подано «найповніший і методично валідний перелік об'єктивних і суб'єктивних індикаторів соціального самопочуття». Високо оцінюють вони і загальну будову методики, що не спричиняє ускладнень за відповідей і дає змогу отримувати незміщені результати опитування. Р. Зінурова і Ф. Фатыхова спираються у своїй концептуалізації соціального самопочуття серед іншого і на підхід, запропонований Є. Головахою і Н. Паніною, а також вважають, що на сьогодні найдокладніший варіант вимірювального інструмента цього феномену пропонує саме ПСС. Окрім цього, ПСС ставав предметом методологічного інтересу в рамках обґрунтування методики вимірювання соціального самопочуття студентів (Мягков, 2015), типологічного аналізу найманих працівників (Татарова & Бессокирная, 2016), а також у контексті ширших методологічних пошуків (Герасимова, 2017).

З огляду на викладене вище логічним і своєчасним убачається звернення до ПСС як до об'єкта повноцінної методологічної рефлексії. Таке ставлення питання ґрунтується на вимірювальному потенціалі ПСС, котрий, якщо й усвідомлений його авторами, не був ними висвітлений.

Аналіз вимірювальних властивостей ПСС, поданий далі, базується на альтернативному підході опрацювання результатів тесту, що дав змогу по-новому подивитися на особливості оцінювання соціального самопочуття і вможливив перехід від «сирих» значень до інтегрального індексу або класифікаційних категорій, а також побудову скорочених версій методики.

§ 10.2. Оцінювання соціального самопочуття на агрегованому рівні

ПСС дає змогу оцінювати соціальне благополуччя населення України загалом на агрегованому рівні без звернення до значень індексу на рівні індивідуальних спостережень. Для цього розглянемо результати використання скороченої версії методики (ПСС-20), що доступні в масиві соціологічного моніторингу «Українське суспільство» за 2016 рік (Інститут соціології НАН України).

Передусім, кожен аспект соціального благополуччя, що входить до змісту ПСС-20, можна оцінити з точки зору його актуальності для респондентів. Це можна здійснити, використовуючи кількість відповідей «не цікавить» за кожним індикатором. На підставі цього показника різноманітні пункти методики впорядковано від найбільш до найменш актуальних (табл. 10.1). У зазначеній таблиці наведено кількість респондентів, які оцінюють доступність відповідних соціальних благ як достатню (за підрахунків із загальної кількості спостережень щодо кожного пункту було виключено респондентів, які обрали варіант відповіді «не цікавить»).

Таблиця 10.1

Характеристики аспектів соціального самопочуття на агрегованому рівні

№	Потреби	Не цікавить	Достатньо
1	Здоров'я	1,1	44,3
2	Можливість придбати найнеобхідніші продукти	2,2	49,2
3	Можливість харчуватися відповідно до своїх смаків	2,6	27,7
4	Уміння жити в нових суспільних умовах	3,3	36,4
5	Упевненість у своїх силах	3,4	45,1
6	Добре житло	3,7	46,1
7	Ініціатива і самостійність у розв'язанні життєвих проблем	4,0	47,3
8	Необхідна медична допомога	4,8	25,2
9	Необхідний одяг	5,1	57,8
10	Повноцінне дозвілля	5,2	28,3
11	Рішучість у досягненні своїх цілей	7,0	41,2
12	Можливість повноцінно проводити відпустку	8,8	19,0
13	Необхідні меблі	11,2	54,3
14	Можливість працювати з повною віддачею	11,8	40,1
15	Прийнятна робота	13,6	32,8
16	Юридична допомога в захисті своїх прав та інтересів	14,1	23,2
17	Можливість підробляти	15,9	23,7
18	Сучасні економічні знання	17,4	27,4
19	Модний і гарний одяг	24,2	33,8
20	Сучасні політичні знання	26,3	35,3

Як можна бачити, до найактуальніших (маються на увазі ті пункти, неактуальність яких не перевищує 7%, або перші 10) належать потреби, пов'язані з фізичними, матеріальними і психологічними аспектами щоденного існування. Повне задоволення їх коливається від 25,2% (необхідна медична допомога) до 57,8% (необхідний одяг). Середній рівень їх задоволення становить близько 41%. Загалом умовну позначку 50%, окрім уже вказаної долає лише одна потреба — в необхідних меблях.

Друга, менш актуальна половина списку передусім стосується самореалізації та вищих стандартів життя. Очевидно, що за низького рівня задоволеності потреб повсякденного існування говорити про актуальність самореалізації й отримання додаткових благ складніше. І хоча для багатьох респондентів задоволення потреб вищого порядку також є важливим, рівень забезпечення їх очікувано нижчий (у середньому — близько 33%).

Така, дещо похмура картина пов'язана насамперед з обраним способом оцінювання. Якщо взяти до уваги не тільки кількість відповідей «вистачає», а й «важко сказати, вистачає чи ні», то ситуація відчутно зміниться на краще. Інакше кажучи, ефект методу очевидний. Водночас з огляду на такі макроекономічні явища в житті нашої країни, як скорочення реального сектору економіки,

збільшення державних боргів за кредитами, істотне подорожчання енергоресурсів, суттєва депопуляція населення за роки незалежності, а також військовий конфлікт на Сході країни, аналізовані цифри вже не виглядають надто похмурими.

У змістовій площині такий аналіз висуває на перший план низку методологічних завдань — урахування релевантності різноманітних аспектів соціального благополуччя, оцінку смислового групування індикаторів скороченого варіанта індексу, аналіз впливу його змісту на підсумкові результати аналізу.

§ 10.3. Оцінювання індикаторів ІСС-20 стосовно змісту ІСС-44

Спершу розглянемо ІСС-20 за змістом. Для цього проаналізуємо, які з 11 сфер соціальної життєдіяльності, покладені в основу повної версії індексу, враховані у скороченій версії (табл. 10.2). Як бачимо, чотири сфери зовсім не враховані у скороченому варіанті індексу — соціальних відносин, національних відносин, соціально-політична, міжособистісних відносин. Перші три можна віднести до інституційних факторів соціального самопочуття, останню — до соціально-психологічних.

Таблиця 10.2

Змістові характеристики ІСС-20

<i>Сфери соціальної життєдіяльності</i>	<i>Індикатори ІСС-20</i>	<i>Тип(-и) факторів соціального самопочуття</i>
Соціальних відносин	–	Інституційний
Соціальної безпеки	Юридична допомога в захисті своїх прав та інтересів	Інституційний
Національних відносин	–	Інституційний
Соціально-політична	–	Інституційний
Професійно-трудова	1) Можливості працювати з повною віддачею 2) Прийнятна робота 3) Можливість мати додатковий заробіток	1) Соціально-психологічний 2) Інституційний 3) Матеріальний
Інформаційно-культурна	1) Сучасні економічні знання 2) Сучасні політичні знання	Інформаційний
Рекреаційно-культурна	1) Здоров'я 2) Необхідна медична допомога 3) Повноцінне дозвілля 4) Можливість повноцінно проводити відпустку	1) Фізичний 2) Інституційний 3) Соціально-психологічний 4) Матеріальний
Матеріально-побутова сфера (1-го рівня)	1) Можливість купувати найнеобхідніші продукти 2) Добре житло 3) Необхідний одяг 4) Необхідні меблі	Матеріальний

<i>Сфери соціальної життєдіяльності</i>	<i>Індикатори ПСС-20</i>	<i>Тип(-и) факторів соціального самопочуття</i>
Матеріально-побутова (2-го рівня)	1) Можливість харчуватися відповідно до своїх смаків 2) Модний і гарний одяг	Матеріальний
Міжособистісних стосунків	—	Соціально-психологічний
Особистісна сфера	1) Уміння жити в нових суспільних умовах 2) Упевненість у своїх силах 3) Ініціатива і самостійність під час розв'язання життєвих проблем 4) Рішучість у досягненні своїх цілей	Психологічний

Якщо в повній версії більш як третина всіх індикаторів мають стосунок до інституційного рівня, то у скороченій — лише 15%. Менша репрезентованість у скороченій версії характерна і для соціально-психологічних факторів (10% проти 14%). При цьому обидві зазначені групи факторів мають прямий стосунок до соціального благополуччя — інституційні фактори визначають загальний соціально-політичний і соціально-економічний контекст суспільства, а соціально-психологічні зумовлюють якість безпосередньої соціальної взаємодії.

Своєю чергою, на матеріальні фактори, що становлять майже чверть повної версії методики, у скороченій версії припадає 40%. Надмірна репрезентованість характерна і для психологічних факторів — 20% замість 10%.

Така картина схиляє до питання про невідповідність структур повної та скороченої версій індексу. Головною причиною такої невідповідності є спосіб скорочення повної версії. За словами авторів методики (Головаха & Панина, 1997, с. 46), вони базувалися на формальних і змістових засадах. Формальні полягали у збереженні внутрішньої узгодженості індексу, змістові — в його багатоаспектності. Якщо критично поставитися до такого рішення, слід відзначити таке. По-перше, немає потреби у збереженні внутрішньої узгодженості індексу, оскільки його підсумкове значення є результатом впливу не латентної змінної, а певного набору незалежних факторів. По-друге, з огляду на підсумковий результат структура скороченої версії вельми відрізняється від повного варіанта.

Відкидаючи важливість внутрішньої узгодженості індексу, все одно необхідно оцінити актуальність різних сфер соціальної життєдіяльності. Оскільки повну версію ПСС було включено до соціологічного моніторингу «Українське суспільство» 2006 року, для реалізації цього завдання використано саме його результати (табл. 10.3). У технічному плані оцінювання включає обчислення кореляції між індексом ПСС-44 та індексами тих індикаторів, що змістовно описують різноманітні сфери соціальної життєдіяльності, а також розгляд розподілів варіантів відповіді «не цікавить», що варіює від 0 (для кожного з чо-

тирьох індикаторів указано достатність/недостатність благ) до 4 (усі індикатори відповідної підгрупи не цікавлять респондента). На користь більшої актуальності свідчать як вищі показники сили зв'язку, так і менша кількість відповідей «не цікавить».

Таблиця 10.3

Актуальність різних сфер соціальної життєдіяльності

Сфери соціальної життєдіяльності	Сила зв'язку	Показники відсутності інтересу, %				
		0	1	2	3	4
Соціальних відносин	0,68	84,4	10,8	3,4	1,0	0,4
Соціальної безпеки	0,65	83,6	12,9	2,3	0,8	0,3
Національних відносин	0,62	77,0	16,2	4,5	1,7	0,7
Соціально-політична	0,62	78,4	14,2	4,8	2,0	0,7
Професійно-трудова	0,69	72,5	13,8	5,9	5,4	2,4
Інформаційно-культурна	0,64	58,2	19,0	11,9	8,9	2,0
Рекреаційно-культурна	0,70	80,8	13,2	5,5	0,5	0,1
Матеріально-побутова сфера (1-го рівня)	0,69	83,4	11,9	2,7	1,4	0,6
Матеріально-побутова (2-го рівня)	0,68	53,4	27,9	14,1	4,3	0,3
Міжособистісних відносин	0,57	87,5	10,1	2,0	0,3	0,2
Особистісна сфера	0,68	86,5	9,4	3,1	0,7	0,3

Майже всі показники сили зв'язку потрапляють в інтервал від 0,6 до 0,7, за винятком сфери міжособистісних відносин. З іншого боку, саме для неї зафіксовано найбільшу кількість респондентів, які жодного разу не відповіли «не цікавить». Натомість для інформаційно-культурної та матеріально-побутової (другого рівня) сфер за вищих показників сили зв'язку (особливо у другому випадку) кількість незацікавлених респондентів є максимальною. Примітно, що найвищий зв'язок зафіксований для рекреаційно-культурної сфери, для якої характерне найбільше розмаїття індикаторів (кожен із них можна віднести до свого типу).

Загалом, зважаючи на ці дані, немає підстав вважати, що за побудови скороченого варіанта індексу можна знехтувати окремими сферами настільки, щоб вилучити всі їх індикатори. Логічним продовженням цього аналізу є конструювання альтернативної версії скороченого варіанта ІСС.

§ 10.4. Обчислення індексу ІСС: альтернативний варіант

Згідно з оригінальним способом розрахунку підсумкових значень для ІСС-44 і ІСС-20, чотири можливі альтернативні варіанти відповідей прирівнюють до трьох числових значень: 1 — «не вистачає», 2 — «важко сказати, вистачає чи ні»

та «не цікавить», 3 — «вистачає». Отже, автори методики вважають близькими за змістом ситуації, коли респондент вагається сказати, чи вистачає йому певного соціального блага, і коли йому це благо не цікаве.

У рамках нашої роботи використано альтернативний варіант обчислення підсумкового індексу. До уваги взято лише варіанти відповідей «не вистачає» (у числовому виразі дорівнює одиниці), «важко сказати, вистачає чи ні» (у числовому виразі дорівнює двом), «вистачає» (у числовому виразі дорівнює трьом). Відповідні результати щодо кожного спостереження підсумовують, а потім ділять на кількість індикаторів цього спостереження, щодо яких не запропоновано варіант відповіді «не цікавить». Отже, середнє значення індексу для кожного респондента потрапляє в інтервал від 1 до 3, як і за оригінальної процедури розрахунку індексу. Як і в разі оригінального підходу авторів методики, індекс визначають тільки для респондентів, котрі відповіли за всіма індикаторами тесту.

Описані зміни ґрунтуються на двох аргументах. По-перше, ми вважаємо некоректним прирівнювати за змістом варіанти відповіді «важко сказати, вистачає чи ні» та «не цікавить». Перший випадок стосується деякої невизначеності під час оцінювання певного блага, другий — чіткої позиції стосовно того блага, яке не впливає на його соціальне самопочуття на думку самого респондента. По-друге, оригінальний спосіб обчислення зсуває підсумкові результати до їх ідеального середнього — 2, збільшуючи тим самим кількість спостережень за вибіркою, для яких характерне середнє соціальне благополуччя.

Крім того, кількість відповідей «не цікавить» для кожного респондента можна розглядати як додатковий показник соціального благополуччя, що характеризує міру насиченості соціального життя індивіда (далі — соціальна апатія). На користь такої інтерпретації свідчить зв'язок цього показника з віком респондента (кількість повних років) і рівнем його освіти (шість градацій — від неповної середньої до повної вищої). Для ПСС-44 сила зв'язку є слабшою ($r_{вік} = 0,34$; $r_{осв} = -0,20$), ніж для ПСС-20 ($r_{вік} = 0,45$; $r_{осв} = -0,22$)⁸⁵. У першому випадку величина сукупно поясненої дисперсії дорівнює 13,4%, у другому — 22,4%. Причини таких розбіжностей розглянемо далі. Проте загалом можна говорити про звуження сфери актуальних соціальних благ із віком, а також про розширення їх (нехай навіть не дуже суттєве) зі здобуттям додаткової освіти.

§ 10.5. Конструювання ПСС-13

За розробки альтернативного скороченого варіанта ПСС ми спиралися на такі методологічні орієнтири:

- збереження структури ПСС-44 — принаймні один індикатор із кожної сфери соціальної життєдіяльності був відібраний;
- оптимальна стислість — кількість індикаторів (загалом 13) зумовлена значенням, що забезпечує прийнятний діапазон можливих значень індексу,

⁸⁵ Для всіх коефіцієнтів кореляції $p < 0,001$.

- що, своєю чергою, вможливує використання різноманітних статистичних методів аналізу даних⁸⁶;
- внесок у загальний індекс соціального самопочуття та індекс на підставі індикаторів своєї сфери — за добору індикаторів ураховано силу їх зв'язку з підсумковим значенням ПСС-44 (r_g), а також індексами на підставі тих чотирьох індикаторів, що описують відповідну сферу соціальної життєдіяльності (r_p)⁸⁷;
 - загальна релевантність — у разі близькості різноманітних індикаторів за попереднім показником до уваги взято кількість респондентів, які обрали варіант відповіді «не цікавить» (f_{di}).

Розглянемо добір індикаторів для кожної сфери докладніше.

Сфера соціальних відносин. Одразу три індикатори показали достатньо великий внесок у загальний і частковий індекси: 1) стабільність у державі та суспільстві ($r_g = 0,41$; $r_p = 0,63$; $f_{di} = 3,5\%$); 2) впевненість у тому, що ситуація у країні поліпшуватиметься ($r_g = 0,40$; $r_p = 0,65$; $f_{di} = 2,3\%$); 3) справедлива оцінка заслуг людини перед суспільством ($r_g = 0,45$; $r_p = 0,62$; $f_{di} = 6,4\%$).

Третій було відкинуто через порівняно високе значення f_{di} . Коли йшлося про вибір між першим і другим індикаторами, рішення було ухвалене на користь першого, оскільки він більшою мірою стосується соціальної дійсності, а не внутрішніх переживань, які можна індукувати штучним шляхом (особливо коли йдеться про соціальний оптимізм).

Сфера соціальної безпеки. Близькі показники зафіксовано для двох індикаторів: 1) захист від злочинності ($r_g = 0,42$; $r_p = 0,66$; $f_{di} = 4,2\%$); 2) юридична допомога для захисту своїх прав та інтересів ($r_g = 0,46$; $r_p = 0,57$; $f_{di} = 11,8\%$).

Враховуючи величину f_{di} , вибір зроблено на користь першого з них.

Сфера національних відносин. У цій сфері вибір також стосується двох індикаторів: 1) впевненість, що не виникатимуть міжнаціональні конфлікти ($r_g = 0,39$; $r_p = 0,61$; $f_{di} = 5,2\%$); 2) взаєморозуміння між людьми різних національностей ($r_g = 0,35$; $r_p = 0,61$; $f_{di} = 9,0\%$).

Знову ж таки на підставі показника f_{di} вибір варто зробити на користь першого з них.

Соціально-політична сфера. Тут усі показники свідчать на користь єдиного індикатора — дотримання у країні прав людини ($r_g = 0,43$; $r_p = 0,63$; $f_{di} = 2,6\%$).

Професійно-трудова сфера. Доволі «рівна» з точки зору показників сфера: 1) освіта ($r_g = 0,35$; $r_p = 0,55$; $f_{di} = 7,8\%$); 2) прийнятна робота ($r_g = 0,36$; $r_p = 0,59$; $f_{di} = 14,2\%$); 3) можливість мати додатковий заробіток ($r_g = 0,39$; $r_p = 0,59$; $f_{di} = 17,6\%$); 4) можливість працювати з повною віддачею ($r_g = 0,36$; $r_p = 0,62$; $f_{di} = 11,6\%$).

Перший індикатор за найнижчого внеску в індекси має найкраще значення f_{di} . За своїм змістом він стоїть осторонь. Тому є сенс використовувати два інди-

⁸⁶ Як показав досвід валідазації соціологічного тесту SCL-9-NR, діапазон можливих значень від 0 до 27 є достатнім. У разі ПСС-13 цей діапазон потрапляє в інтервал від 13 до 39.

⁸⁷ Для цього використовують коефіцієнт кореляції Спірмена.

катори з цієї групи — стосовно освіти і стосовно самої роботи. В останньому випадку, виходячи з величини f_{di} , обрано четвертий індикатор.

Інформаційно-культурна сфера. У цій групі можна виокремити два індикатори: 1) сучасні економічні знання ($r_g = 0,36$; $r_p = 0,62$; $f_{di} = 18,4\%$); 2) інформація про зміни, що відбуваються в суспільстві ($r_g = 0,29$; $r_p = 0,48$; $f_{di} = 7,7\%$).

Перший із них вирізняється високим унеском в індекси, а також порівняно низьким для цієї групи значенням f_{di} (у випадку індикаторів цієї групи щодо сучасних політичних і науково-технічних знань значення f_{di} перевищують 25%). Що стосується другого, то за низького внеску в індекси (як і в разі сучасних політичних і науково-технічних знань) величина f_{di} є вельми низькою для цієї групи. Вибір між ними зроблено на користь першого індикатора через вищу пріоритетність сили взаємозв'язків з індексами порівняно зі значенням f_{di} .

Рекреаційно-культурна сфера. Оскільки кожен з індикаторів цієї групи належить до власного типу, наведемо кожен із них: 1) здоров'я ($r_g = 0,39$; $r_p = 0,68$; $f_{di} = 0,4\%$); 2) необхідна медична допомога ($r_g = 0,48$; $r_p = 0,70$; $f_{di} = 3,8\%$); 3) повноцінне дозвілля ($r_g = 0,43$; $r_p = 0,45$; $f_{di} = 9,7\%$); 4) можливість повноцінно проводити відпустку ($r_g = 0,40$; $r_p = 0,37$; $f_{di} = 12,0\%$).

Унаслідок суттєвої гетерогенності та важливості цієї сфери з огляду на соціальне самопочуття було відібрано два індикатори. Оскільки перший і другий об'єднані спільною темою здоров'я, а третій і четвертий — вільного часу, то з кожної підгрупи відібрано по одному індикатору.

Другий індикатор виглядає адекватнішим за перший як за своїми показниками, так і з огляду на соціальне самопочуття. Останнє пов'язане з тим, що необхідна медична допомога значно краще вписується у сферу суспільних благ, ніж здоров'я. Третій індикатор, своєю чергою, адекватніший за четвертий. По-перше, це пов'язане зі статистичними показниками; по-друге, дозвілля — ширше поняття, ніж відпустка; по-третє, нормальне дозвілля — постійна умова підтримання нормальної життєдіяльності, натомість відпустка має сезонний характер.

Матеріально-побутова сфера (перший рівень). Два індикатори показують добрі статистичні показники: 1) можливість купувати найнеобхідніші продукти ($r_g = 0,54$; $r_p = 0,66$; $f_{di} = 2,0\%$); 2) необхідний одяг ($r_g = 0,48$; $r_p = 0,66$; $f_{di} = 6,9\%$).

Як можна бачити, перевагу слід віддати першому.

Матеріально-побутова сфера (другий рівень). У цьому випадку найліпше працює такий індикатор: можливість харчуватися відповідно до своїх смаків ($r_g = 0,55$; $r_p = 0,69$; $f_{di} = 1,8\%$).

Сфера міжособистісних стосунків. Тут вибір слід зробити між двома індикаторами: 1) доброта і співчуття оточуючим ($r_g = 0,46$; $r_p = 0,59$; $f_{di} = 3,8\%$); 2) щастя в сімейному житті ($r_g = 0,33$; $r_p = 0,68$; $f_{di} = 3,3\%$).

Оскільки успішність сімейного життя не так сильно пов'язана з загальним соціальним благополуччям (це підтверджує відповідний коефіцієнт кореляції), обрано характеристики соціального оточення.

Особистісна сфера. Один з індикаторів демонструє найкращі результати — впевненість у своїх силах ($r_g = 0,50$; $r_p = 0,75$; $f_{di} = 2,3\%$).

У результаті співвідношення сфер соціальної життєдіяльності, індикаторів і типів факторів соціального самопочуття матиме такий вигляд (табл. 10.4):

Таблиця 10.4

Змістові характеристики ПСС-13

<i>Сфери соціальної життєдіяльності</i>	<i>Індикатори ПСС-13</i>	<i>Тип(-и) факторів соціального самопочуття</i>
Соціальних відносин	Стабільність у державі та суспільстві	Інституційний
Соціальної безпеки	Захист від злочинності	Інституційний
Національних відносин	Упевненість, що не виникатимуть міжнаціональні конфлікти	Інституційний
Соціально-політична	Дотримання у країні прав людини	Інституційний
Професійно-трудова	1) Освіта 2) Можливість працювати з повною віддачею	1) Інституційний 2) Соціально-психологічний
Інформаційно-культурна	Сучасні економічні знання	Інформаційний
Рекреаційно-культурна	1) Необхідна медична допомога 2) Повноцінне дозвілля	1) Інституційний 2) Соціально-психологічний
Матеріально-побутова сфера (1-го рівня)	Можливість купувати найнеобхідніші продукти	Матеріальний
Матеріально-побутова сфера (2-го рівня)	Можливість харчуватися відповідно до своїх смаків	Матеріальний
Міжособистісних стосунків	Доброта і співчуття оточуючих	Соціально-психологічний
Особистісна сфера	Упевненість у своїх силах	Психологічний

§ 10.6. Порівняння ПСС-13 та ПСС-20

У структурному плані ПСС-20 значно відрізняється від ПСС-13 — якщо в першому акцент зроблено на матеріальні (40%) і психологічні (20%) показники, то у другому — на інституційні (46%) і соціально-психологічні (23%).

По-перше, це зумовлює вищий рівень зв'язку ПСС-13 із повною версією індексу ($r = 0,92$) порівняно з ПСС-20 ($r = 0,91$). Оцінюючи ці мінімальні відмінності, варто брати до уваги кількість індикаторів, використаних у скорочених версіях. При цьому сила зв'язку між індексами скорочених версій дорівнює 0,86.

По-друге, середні значення індексів ПСС-20 і ПСС-13 статистично значимо відрізняються ($X_{ПСС-20} = 1,90$; $X_{ПСС-13} = 1,78$; $t = 7,66$; $df = 3492$; $p < 0,001$). Таки доволі суттєві відмінності показників різних варіантів ПСС указують на

необхідність незалежного опрацювання їх критеріальних рівнів і правил інтерпретації. Це ґрунтується на тому, що ПСС-20 можна схарактеризувати передусім як скорочений варіант матеріально-психологічної спрямованості, а ПСС-13 — інституційно-соціально-психологічної. При цьому середнє ПСС-44 ($X_{ПСС-44} = 1,87$) є ближчим до ПСС-20 ($t = 2,08$; $df = 3301$; $p = 0,038$), ніж до ПСС-13 ($t = -6,42$; $df = 3423$; $p < 0,001$).

По-третє, структура взаємозв'язків між соціальною апатією, з одного боку, і віком і освітою — з іншого, краще відтворюється за допомоги значень ПСС-20, ніж ПСС-13 (табл. 10.5). Отже, головна причина цього полягає в тому, що релевантність благ, охоплених ПСС-20, сукупно більшою мірою залежить від вікових змін респондентів.

Такі відмінності мають відбитися на результатах конструктивної валідації скорочених варіантів. Для підтвердження цієї гіпотези було відібрано змінні, що оцінюють матеріальне становище, психологічні та соціально-психологічні характеристики респондентів, а також їхні думки з приводу інституційних особливостей нашого суспільства.

Таблиця 10.5

Взаємозв'язок соціальної апатії з віком і освітою респондентів для повного і скорочених варіантів ПСС

Показник	ПСС-13	ПСС-20	ПСС-44
Вік (r)	0,31	0,45	0,34
Освіта (r)	-0,18	-0,22	-0,20
Пояснена дисперсія	10,8 %	22,4 %	13,4 %

Для оцінювання *матеріального становища* відібрано такі змінні: 1) самооцінка матеріального рівня життя сім'ї (11-бальна числова шкала); 2) розмір заробітної плати (стипендії, пенсії) респондента за останній місяць; 3) оцінка матеріального становища своєї сім'ї за останні 2–3 місяці (7-бальна порядкова шкала: від «часто не маємо грошей та їжі» до «живемо в повному достатку»); 4) оцінка респондентом матеріального становища своєї сім'ї (5-бальна порядкова шкала: від «злиденне» до «заможне»).

У всіх випадках ПСС-20 показує тісніший взаємозв'язок порівняно з ПСС-44 і ПСС-13 (табл. 10.6). Найбільш виражені відмінності зафіксовано за змінною «розмір заробітної плати (стипендії, пенсії)». Ці результати добре узгоджуються зі змістовими особливостями різних варіантів індексу. По-перше, велика відносна представленість матеріальних аспектів соціального самопочуття в ПСС-20 веде до тіснішого взаємозв'язку його значень із відповідними змінними. По-друге, з тієї самої причини для змінної «величина доходу» відмінності в силі взаємозв'язку є найвідчутнішими. По-третє, вищий взаємозв'язок індексів із матеріальним рівнем і матеріальним становищем можна пояснити тим, що ці аспекти матеріального добробуту зумовлюються не тільки особистим фінансовим успіхом індивіда, а й інституційними факторами.

Таблиця 10.6

**Взаємозв'язок різних варіантів індексу
з показниками матеріального добробуту респондентів**

Змінні	Варіанти ПСС		
	ПСС-44	ПСС-20	ПСС-13
Матеріальний рівень	0,40	0,44	0,40
Розмір доходу	0,22	0,28	0,21
Матеріальне становище за останні 2–3 місяці	0,34	0,39	0,36
Матеріальне становище загалом	0,30	0,33	0,32

Отже, незважаючи на переваги ПСС-20 у цьому аспекті аналізу, як ПСС-44, так і ПСС-13 достатньо чутливі щодо матеріальної складової соціального добробуту, попри відмову від певного набору відповідних індикаторів (ПСС-13) або меншу відносну репрезентованість їх (ПСС-44).

Психологічні характеристики оцінено за такими змінними: 1) кількість стресових подій (важка хвороба, смерть близької людини, втрата роботи тощо), що сталися з респондентом останніми 12 місяцями (діапазон можливих значень від 0 до 14); цю змінну застосовано виходячи з припущення, згідно з яким вона є непрямим показником психологічного дистресу; 2) шкала тривожності Спілбергера, форма *B* (діапазон можливих значень — від 20 до 80 балів); 3) локус контролю⁸⁸ (5-бальна порядкова шкала — від «переважно від зовнішніх обставин» до «цілком від мене»).

Виходячи з результатів кореляційного аналізу (табл. 10.7), подані варіанти індексу відрізняються мінімально. Мабуть, трохи вищий взаємозв'язок значень ПСС-20 зі змінною локуса контролю продиктований відносно вищою презентованістю в ньому індикаторів, що оцінюють особистісні характеристики респондента. Проте загалом психологічні характеристики не можна назвати релевантним аспектом порівняння конструктивних особливостей ПСС-44, ПСС-20 і ПСС-13.

Таблиця 10.7

**Взаємозв'язок різноманітних варіантів індексу
з психологічними характеристиками респондентів**

Змінні	Варіанти ПСС		
	ПСС-44	ПСС-20	ПСС-13
Кількість стресових подій	-0,28	-0,27	-0,25
Рівень тривожності	-0,44	-0,43	-0,42
Локус контролю	0,20	0,24	0,22

⁸⁸ Поняття у психології, що характеризує властивість особистості приписувати свої успіхи або невдачі факторам зовнішнього, внутрішнього чи змішаного характеру.

Щодо *соціально-психологічних характеристик*, то було використано такі: 1) шкала аномії⁸⁹ (діапазон можливих значень — від 0 до 18 балів); 2) шкала цинізму⁹⁰ (діапазон можливих значень — від 0 до 14 балів); 3) оцінка свого становища в суспільстві (7-бальна числова шкала); 4) задоволеність своїм становищем у суспільстві на даний момент (3-бальна порядкова шкала: скоріше не задоволений; важко сказати, задоволений чи ні; скоріше задоволений).

Отримані результати (табл. 10.8) важко назвати узгодженими. Показники аномії краще узгоджуються з результатами ПСС-13, а показники цинізму — з ПСС-44, притому, що ПСС-20 поступається іншим варіантам в обох випадках. Водночас ПСС-20 є оптимальним варіантом за аналізу становища респондента в суспільстві. Отже, однозначно стверджувати про перевагу того чи того варіанта індексу немає сенсу, що знову ж таки свідчить про нерелевантність цього набору змінних у плані оцінювання конструктивних особливостей різних варіантів ПСС.

Таблиця 10.8

**Взаємозв'язок різних варіантів індексу
з соціально-психологічними характеристиками респондентів**

Змінні	Варіанти ПСС		
	ПСС-44	ПСС-20	ПСС-13
Аномія	-0,37	-0,35	-0,39
Цинізм	-0,22	-0,16	-0,20
Самооцінка свого становища в суспільстві	0,34	0,38	0,36
Задоволеність становищем у суспільстві	0,43	0,44	0,41

Зрештою, за аналізу зв'язку індексів з *інституційними показниками* використано такі змінні: 1) оцінка актуальної економічної ситуації в Україні (11-бальна числова шкала); 2) оцінка діяльності Президента України (10-бальна числова шкала); 3–5) рівень довіри до президента, парламенту й уряду (в усіх випадках використано 5-бальну порядкову шкалу — від «зовсім не довіряю» до «цілком довіряю»).

Тут отримано результати, що вкладаються в однозначний патерн — ПСС-44 і ПСС-13 показали щільніші взаємозв'язки порівняно з ПСС-20 (табл. 10.4). Така ситуація свідчить, що інституційну специфіку українського суспільства краще «схоплюють» ПСС-44 і ПСС-13. При цьому ПСС-20 найбільше «просідає» за тими змінними, котрі стосуються Президента України.

⁸⁹ Шкала аномії (МакКлоскі, Шаар) має 9 пунктів, котрі вказують на аномійну деморалізованість (психологічна реакція людей на соціальну ситуацію, коли одна система норм і цінностей, що об'єднує людей у спільноту, зруйнована, а інша ще не сформувалася) (Панина, 2012, с. 313).

⁹⁰ Шкала цинізму є додатковою шкалою Міннесотського багатопрофільного особистісного опитувальника (ММРІ). Відтворює міру зневажливого ставлення до узвичаєних цінностей (Панина, 2012, с. 314).

Таблиця 10.9

**Взаємозв'язок різних варіантів індексу
з оцінкою інституційних особливостей України**

Змінні	Варіанти ІСС		
	ІСС-44	ІСС-20	ІСС-13
Оцінка економічної ситуації в Україні	0,28	0,25	0,29
Оцінка президента	0,18	0,10	0,16
Довіра до президента	0,23	0,14	0,21
Довіра до парламенту	0,21	0,15	0,18
Довіра до уряду	0,26	0,19	0,23

Підсумовуючи результати конструктивного аналізу різних варіантів ІСС, необхідно відзначити зв'язок ІСС-20 із заможністю, а ІСС-13 — з інституційною специфікою, за більшої наближеності останнього до повного варіанта методики. В решті аспектів усі варіанти індексу достатньо близькі між собою. Отже, ІСС-20 доцільніше використовувати в ситуаціях, коли дослідникові необхідно зосередитися на тій складовій соціального самопочуття, яка пов'язана з побутовим рівнем життєдіяльності індивіда, а ІСС-13 — у ситуаціях інтересу до ширшого соціального контексту.

З огляду на відмінності між різними версіями індексу необхідно запропонувати можливі варіанти інтерпретації відповідних результатів, щоб від «сирих» значень індексу або профілів відповідей перейти до категорій респондентів, соціальне самопочуття яких піддається осмисленій інтерпретації.

§ 10.7. Категоризація респондентів на підставі результатів використання різних версій ІСС

Авторами ІСС для визначення норм тесту запропоновано використовувати загальноприйнятій у психології підхід, ґрунтований на дослідженнях стандартизації (Головаха & Панина, 1997, с. 48–53). Останнє передбачає вивчення репрезентативних груп із метою використання отриманих результатів як орієнтирів порівняння індивідуальних показників за тестом. Це дає змогу з'ясувати відносну позицію учасника дослідження щодо вибірки стандартизації, а також забезпечує порівнянність результатів різноманітних тестів (Анастаси & Урбина, 2007, с. 64–65).

З низки причин, розглянутих у першому розділі, ми вважаємо такий підхід непридатним у рамках розробки соціологічних тестів і пропонуємо звернутися до визначення критеріальних рівнів тесту, що маркують якісно відмінні стани респондентів. У нашому випадку це може бути низьке, середнє і високе соціальне самопочуття.

Ще одна альтернатива категоризації пов'язана з розробкою класифікатора відповідей — завдяки використанню спеціальних «ключів», із якими порівню-

ють профіль відповідей кожного спостереження, респондента відносять до тієї чи тієї категорії. Далі наведено результати застосування обох указаних альтернатив.

Критеріальні рівні. Для визначення критеріальних рівнів потрібен доступ до результатів відповідей специфічних, з огляду на досліджувану властивість, соціальних груп. Наприклад, за вимірювання психологічного дистресу до таких груп можна віднести людей із психіатричним діагнозом або індивідів, які переживають гострі соціальні негаразди. Зважаючи на статистичні показники відповідної властивості, дослідник отримує орієнтовну інформацію щодо ймовірних меж критеріальних рівнів. На останньому етапі цю інформацію використовують для побудови передбачуваних моделей, що дають змогу оцінити поділовий потенціал тесту.

Проте застосувати цей підхід до ПСС у чистому вигляді не вдалося. Останнє пов'язане з тим, що визначення специфічних або референтних груп виявилось вельми проблематичним. Так, оскільки для індексів характерна висока тематична гетерогенність своїх показників, немає жодних підстав думати, що за низького/високого рівня соціального самопочуття в одній зі сфер соціальної життєдіяльності в решті також спостерігатиметься схожа ситуація. А оскільки референтні групи не було визначено, перевірити поділовий потенціал тесту було неможливо.

Щоб розв'язати проблему референтних груп, принаймні частково, замість показників реальних груп було змодельовано показники, що характеризують ідеально-типових респондентів. Для цього на першому етапі було побудовано лінійні регресійні моделі на підставі п'яти змінних, що добре показали себе⁹¹ під час конструктної перевірки різних варіантів ПСС, а саме: 1) самооцінка матеріального рівня життя сім'ї респондента (11-бальна числова шкала); 2) шкала тривожності Спілбергера, форма *B* (діапазон можливих значень — від 20 до 80 балів); 3) індекс аномії (діапазон можливих значень — від 0 до 18 балів); 4) задоволеність своїм становищем у суспільстві на даний момент (3-бальна порядкова шкала: скоріше не задоволений; важко сказати, задоволений чи ні; скоріше задоволений); 5) оцінка актуальної економічної ситуації в Україні (11-бальна числова шкала).

У кожній із трьох отриманих моделей предиктори статистично значимо впливають на значення ПСС ($p < 0,01$), окрім моделі для ПСС-20, із якої «випадає» оцінка економічної ситуації в Україні. Модель для ПСС-13 пояснює 37% сукупної дисперсії, для ПСС-20 — 38%, для ПСС-44 — 37,5%. За мірками сучасної соціології такі результати можна назвати доволі добрими.

На другому етапі побудовано два ідеально-типові профілі респондентів, для першого з яких характерними є показники «вищий за низький, але нижчий за середній рівень», а для другого — «вищий за середній, але нижчий за високий рівень» для кожного з предикторів (табл. 10.10). Якщо говорити предметніше, то для першого профілю характерна досить низька оцінка матеріального рівня

⁹¹ Відібрано ті змінні, що продемонстрували порівняно сильний взаємозв'язок як із повною версією ПСС, так і з її скороченими варіантами.

життя своєї сім'ї, підвищена тривожність і аномія, скоріше низька задоволеність становищем у суспільстві, а також низька оцінка економічної ситуації в Україні. У другому випадку оцінка матеріального рівня сім'ї, як і економічної ситуації у країні, — вища за середню, рівень тривожності на нормальному рівні, аномія — нижча за середню, задоволеність становищем у суспільстві — скоріше висока.

Таблиця 10.10

Профілі ідеально-типових респондентів

Змінна	Ідеальний тип респондента	
	«Вищий за низький, але нижчий за середній рівень»	«Вищий за середній, але нижчий за високий рівень»
Матеріальний рівень життя сім'ї	3 бали	7 балів
Шкала тривожності	50 балів	40 балів
Індекс аномії	12 балів	6 балів
Задоволеність становищем у суспільстві	1,5 бала	2,5 бала
Оцінка економічної ситуації у країні	3 бали	7 балів

На третьому етапі отримані профілі відповідей було використано як вхідні дані для побудованих раніше моделей. Результати щодо першого профілю вказують на можливу нижню межу середнього рівня соціального самопочуття, результати щодо другого — на можливу нижню межу високого рівня. Для ПСС-13 ці межі дорівнюють відповідно 1,7 і 2,4 бала, для ПСС-20 — 1,8 і 2,5 бала, для ПСС-44 — 1,8 і 2,35 бала. Для груп, що утворюються з огляду на ці межі, є характерним переважання респондентів із низьким і середнім соціальним самопочуттям (табл. 10.11). Результати класифікації ПСС-44 і ПСС-20 збігаються на 80,7%, ПСС-44 і ПСС-13 — на 80,9%, ПСС-20 і ПСС-13 — на 74,7%. Загалом розбіжності достатньо вагомі, але при цьому підлягають осмисленій інтерпретації, ґрунтованій на змістових відмінностях різних версій ПСС.

Таблиця 10.11

Результати групування респондентів на підставі значень різних варіантів ПСС

Рівень соціального самопочуття	Варіанти ПСС		
	ПСС-44	ПСС-20	ПСС-13
Низький	44,6	44,0	46,9
Середній	43,9	41,9	43,7
Високий	11,5	14,0	9,4

Класифікатор відповідей. Далі розглянуто результати опрацювання даних тільки для результатів використання ПСС-44 і ПСС-13, оскільки вони зіставні за структурою, а це відіграватиме суттєву роль у плані коректності порівняння. Але сам підхід може бути застосований і до результатів ПСС-20 без принципових змін.

Оскільки індикатори ПСС побудовані на підставі категоріальних (якісних) шкал, остільки розробка класифікатора полягає у використанні змістових принципів аналізу відповідей респондентів, що дає змогу розподілити спостереження в рамках теоретично обґрунтованих груп. Такий підхід, зокрема, було використано для кластеризації респондентів на підставі результатів соціологічного тесту «Типи політичної культури» (Дембійський, 2017а). У рамках зазначеної праці він дістав назви концептуальної типології, відмітними характеристиками якої є змістовий спосіб визначення схожості між спостереженнями, а також апріорність правил кластеризації.

У випадках ПСС-44 і ПСС-13 було використано два принципи. По-перше, індикатори методик було поділено на дві змістові групи. До першої віднесено індикатори інституційного⁹², а також інформаційного типу (усього 23 і 7 індикаторів відповідно), до другого — соціально-психологічного, матеріального, психологічного і фізичного типів (усього 21 і 6 індикаторів відповідно). Такий поділ ґрунтується на ідеї, згідно з якою соціальне самопочуття назагал визначають двома групами факторів. До першої належать фактори повсякденної життєдіяльності, до другої — суспільно-інституційні. Відповідно аналізувати соціальне самопочуття можна окремо за кожною з цих груп. По-друге, індикатори кожної групи дають респондентові змогу виявити один із чотирьох станів соціального самопочуття:

- соціальна апатія — переважають варіанти відповіді «не цікавить»;
- низьке соціальне самопочуття — варіантів відповіді «не вистачає» більше, ніж варіантів відповіді «Вистачає»;
- високе соціальне самопочуття — варіантів відповіді «вистачає» більше або стільки, як і варіантів відповіді «важко сказати, вистачає чи ні» (за умови відсутності варіантів відповіді «не вистачає»);
- середнє соціальне самопочуття — решта профілів відповідей.

Послідовне застосування цих двох принципів дає очікувані результати — соціальне самопочуття на рівні повсякденної життєдіяльності виглядає краще, ніж під кутом зору ефективності суспільних інститутів (табл. 10.12). Це схиляє до думки, що життєвих успіхів громадяни України досягають переважно всупереч, аніж завдяки діяльності держави.

Порівнюючи розподіл повного і скороченого варіантів індексу, слід наголосити доволі високу загальну схожість результатів у другій групі, натомість на рівні повсякденної життєдіяльності скорочена версія показує оптимістичнішу картину. Якщо дивитися на точний перетин побудованих груп за двома індексами, то збіг для групи індикаторів, що характеризують повсякденну життєдіяльність, становить 70,6%, а для групи, що характеризує суспільні інститути, — 87,9%.

⁹² Виняток становив індикатор «прийнятна робота», що був віднесений до другої групи.

Таблиця 10.12

**Результати застосування класифікатора відповідей
до результатів ПСС-13, %**

Стан соціального самопочуття	Змістова група			
	Повсякденна життєдіяльність		Суспільні інститути	
	ПСС-44	ПСС-13	ПСС-44	ПСС-13
Апатія	1,7	0,6	2,0	1,1
Низьке	46,4	48,1	73,7	75,6
Середнє	46,5	35,9	23,1	20,1
Високе	5,3	15,4	1,2	3,2

Використовуючи дані двох вирізнених підшкал, можна далі диференціювати респондентів на дев'ять груп (табл. 10.13), починаючи з категорії «низьке-низьке» (низька оцінка соціального самопочуття як на рівні повсякденної життєдіяльності, так і щодо суспільних інститутів) і до категорії «високе-високе» (висока оцінка соціального самопочуття в обох випадках). Розгляньмо таку диференціацію на прикладі результатів ПСС-13. При цьому респонденти з апатичними установками виключені з аналізу через мінімальну наповненість відповідних груп.

Таблиця 10.13

**Категоризація рівнів соціального самопочуття
на підставі двох підшкал, %***

Соціальне самопочуття: суспільні інститути	Соціальне самопочуття: повсякденна життєдіяльність		
	Низьке	Середнє	Високе
Низьке	43,8 / 91,1	25,8 / 72,0	5,6 / 36,6
Середнє	3,6 / 7,4	9,3 / 26,0	7,2 / 46,5
Високе	0,3 / 0,6	0,5 / 1,4	2,4 / 15,4

* Перше значення в кожній клітинці показує абсолютну наповненість категорії, друге розраховане як відношення до суми по стовпчику (з урахуванням унеску соціально апатичних респондентів).

Наведені результати дають змогу назагал оцінити результати соціальної адаптації залежно від успішності повсякденної життєдіяльності. За відповідний показник можна взяти суму категорій «середнє» і «високе» для тієї частини соціального самопочуття, що стосується суспільних інститутів. Так, серед респондентів із низьким соціальним самопочуттям на рівні повсякденної життєдіяльності адаптованими можна назвати 8,0% спостережень, із середнім — 27,4%, з високим — 61,9%. Загалом за вибіркою цей показник дорівнює 23,3%.

Якщо зіставити результати категоризації респондентів на підставі ПСС із використанням двох розглянутих підходів (критеріальні рівні vs класифікатор відповідей), то у випадку повсякденної життєдіяльності збіг результатів дорівнює 76,6% (ПСС-44) і 77,0% (ПСС-13), у випадку суспільних інститутів — 58,0% (ПСС-44) і 62,2% (ПСС-13). Такі суттєві розбіжності вказують на необхідність проведення додаткових методичних досліджень.

§ 10.8. Висновки

Здійснена робота свідчить насамперед про те, що можливості використання ПСС виходять далеко за межі, що були встановлені у процесі його створення. Потенціал методики можна цілком розкрити на підставі того, що сьогодні заведено називати *data mining* — аналітичного пошукового процесу опрацювання наявних масивів даних, структура яких згодом приховує релевантні взаємозв'язки. Як продемонстровано вище, звернення до різних варіантів ПСС забезпечує надійне підґрунтя для такої діяльності.

У більш предметному контексті використання ПСС варто наголосити необхідність концептуального фокусування дослідника. Воно передбачає розуміння можливостей і обмежень використовуваної версії ПСС, особливостей її включеності в номологічну мережу, а також специфіки підсумкових результатів. Лише в цьому разі аналіз можна назвати осмисленим із соціологічної точки зору.

Зрештою, подальший розвиток індексу потребує не тільки реактивного (*data mining*), а й активного методологічного пошуку — проведення спеціальних методичних досліджень, спрямованих на адаптацію ПСС до нових умов або специфічних соціальних груп. Так, у скороченій версії ПСС-13 бракує індикатора «впевненість, що не виникатимуть конфлікти на ґрунті політичних розбіжностей». Адже нинішній збройний конфлікт на Сході України зумовлений не міжнаціональними розбіжностями в самій Україні, а політичними і геополітичними конфліктами, що тривалий час індукувалися в нашій країні окремими політичними силами.

Виходячи з цього, з методологічної точки зору пріоритетним є проведення спеціальних методичних досліджень, спрямованих на опрацювання і вивчення вимірювальних властивостей скорочених версій ПСС (зокрема ПСС-13).

Глава 11

Вимірювання індивідуального благополуччя в Європейському соціальному дослідженні: концептуальні та методологічні аспекти

У вітчизняній соціології стандарти вивчення індивідуального благополуччя задані Є. Головахою і Н. Паніною (Головаха & Панина, 1997), які фокусують нашу увагу на необхідності вивчення доступності для індивідів ключових соціальних благ, а по-іншому — задоволеності їхніх соціальних потреб. Загалом цей підхід можна назвати об'єктивним, оскільки доступність/недоступність багатьох соціальних благ обумовлена успішністю функціонування ключових соціальних інститутів.

Натомість у західних соціальних дослідженнях, як було показано раніше, вивчення індивідуального благополуччя спирається переважно на оцінку різних аспектів суб'єктивного благополуччя — задоволеності своїм життям, почуття щастя і психологічного дистресу, що дозволяє назвати відповідний підхід суб'єктивним. У цьому сенсі не є винятком і розробка модуля особистісного і соціального благополуччя в Європейському соціальному дослідженні (далі — ЄСД) (Huppert, Andrew, Frey, Marks & Siegrist, 2005). Наше звернення саме до цього дослідження продиктовано участю в ньому України, що дозволяє розглянути суб'єктивне благополуччя наших громадян на підставі багатого емпіричного матеріалу⁹³.

§ 11.1. Специфіка вимірювання особистісного і соціального благополуччя в ЄСД

У ЄСД до уваги взято два базисні виміри особистісного благополуччя, які можна охарактеризувати як гедоністичне (акцент зроблено на позитивні почуття) і еудаймічне (акцент зроблено на успішну активність). Для їх цілісної інтеграції використовують модель С. Томпсона і Н. Маркса (Thompson & Marks, 2008), у якій благополуччя подане у вигляді динамічного процесу (рис. 11.1).

⁹³ В статті аналізуються дані шостої хвили Європейського соціального дослідження (Norwegian Centre for Research Data, 2012).

У цій моделі успішна активність, або функціонування (еудаймічне благополуччя) є результатом двох факторів — зовнішніх умов (наприклад, робота, побут, сім'я, фізичне здоров'я) і психологічних ресурсів (оптимізм, самооцінка, «гнучкість» тощо). Своєю чергою, функціонування робить зворотний вплив на зовнішні умови життєдіяльності, а також визначає досвід індивіда і його судження про своє життя (щастя, радість, інтерес, нудьга і дистрес). Зрештою, цей досвід і судження чинять зворотний вплив на психологічні ресурси.



Рисунок 11.1. Модель С. Томпсона і Н. Маркса

Разом з тим ця модель стосується саме особистісного, а не соціального благополуччя. Останнє вважають у ЄСД окремим типом і поділяють на міжособистісне і соціетальне. Міжособистісне благополуччя ґрунтується як на отриманні соціальної підтримки, так і на її наданні іншим. Соціетальне — на соціальному капіталі (Jeffrey, Abdallah, & Quick, 2015).

Така концептуально струнка картина помітно втрачає в чіткості, коли йдеться про підбір емпіричних індикаторів. Автори відповідного блоку ЄСД говорять про різні методики, концепти і підконцепти, а також виміри особистісного і соціального благополуччя. І тут давньогрецька філософія «зустрічається» з сучасною соціальною епідеміологією й уявленнями авторів про те, що ж таке особистісне і соціальне благополуччя. Натомість у підсумкових оглядах дослідники приходять до шести вимірів особистісного і соціального благополуччя (табл. 11.1): оцінне благополуччя, емоційне благополуччя, функціональність (functioning), енергійність (vitality), соціетальне благополуччя (community well-being) і підтримувальні взаємовідносини (supportive relationships). Надалі орієнтуватимемося саме на ці групи.

Таблиця 11.1

Індикатори різних вимірів соціального благополуччя

№	Вимір соціального благополуччя / індикатори	Шкала відповідей
<i>Оцінне благополуччя</i>		
1	Враховуючи всі обставини, скажіть, будь ласка, наскільки Ви в цілому задоволені своїм життям в даний час?	11-бальна
2	Говорячи в цілому, наскільки Ви щасливі?	11-бальна
<i>Емоційне благополуччя</i>		
Скільки часу протягом останнього тижня Ви...		
1	... відчували депресію?	4-бальна
2	... були щасливі?	4-бальна
3	... отримували задоволення від життя?	4-бальна
4	... сумували?	4-бальна
5	... відчували тривогу та стурбованість?	4-бальна
6	... відчували себе спокійно?	4-бальна
<i>Функціональність</i>		
1	Я вірю, що можу сам(а) вирішувати, як мені жити	5-бальна
2	У повсякденному житті у мене дуже мало можливостей показати, на що я здатен(а)	5-бальна
3	Майже завжди я відчуваю задоволення від того, що мені вдається	5-бальна
4	Загалом кажучи, як часто Вам цікаво те, що Ви робите?	11-бальна
5	Загалом кажучи, як часто Ви поглинуті тим, що робите?	11-бальна
6	Загалом кажучи, як часто Ви сповнені ентузіазму з приводу того, що робите?	11-бальна
7	В цілому, я гадаю, що те, що я роблю в своєму житті, є цінним та потрібним для людей	5-бальна
8	Наскільки чітко Ви уявляєте, куди рухається Ваше життя?	11-бальна
9	Я завжди дивлюся на своє майбутнє із оптимізмом	5-бальна
10	Є багато речей, в яких я добре розбираюся	5-бальна
11	В цілому я ставлюся до себе дуже добре	5-бальна
12	Часом я відчуваю себе невдахою	5-бальна
13	Коли в моєму житті що-небудь йде не так, звичайно мені потрібно багато часу, щоб прийти до норми	5-бальна
14	Наскільки Вам важко чи легко долати серйозні проблеми?	11-бальна

№	Вимір соціального благополуччя / індикатори	Шкала відповідей
<i>Енергійність</i>		
Скільки часу протягом останнього тижня Ви...		
1	... відчували, що все, що Ви робите, Вам важко дається?	4-бальна
2	... неспокійно спали?	4-бальна
3	... не могли зібратися, мобілізувати себе?	4-бальна
4	... відчували приплив сил та енергії?	4-бальна
<i>Соцістальне благополуччя</i>		
1	Як Ви вважаєте, в цілому більшості людей можна довіряти чи варто бути обережним, маючи справу з людьми?	11-бальна
2	Як Ви думаєте, більшість людей намагатимуться використати або ошукати Вас, якщо будуть мати нагоду, чи вони будуть намагатись поводити себе чесно та порядно?	11-бальна
3	На Вашу думку, люди в основному намагаються допомагати іншим чи дбають про власний інтерес?	11-бальна
4	Ви відчуваєте, що люди, які живуть з Вами по сусідству, допомагають один одному?	7-бальна
5	Я вважаю, що є досить близьким(ою) до людей, що живуть в моєму районі, у тій же місцевості	5-бальна
<i>Підтримуючі взаємовідносини</i>		
1	Скільки часу протягом останнього тижня Ви відчували себе самотньо?	4-бальна
2	Чи є у Вас хтось, з ким би Ви могли поговорити відверто, про особисте, і якщо так, то скільки таких людей у Вашому житті?	7-бальна
3	Якою мірою Ви відчуваєте, що близькі Вам люди Вас цінують?	11-бальна
4	Чи отримуете Ви підтримку та допомогу від близьких Вам людей, коли Ви цього потребуєте?	7-бальна

Як видно, не всі виміри індивідуального і соціального благополуччя мають уніфікований характер з точки зору форми їх індикаторів. Істотно варіюється і кількість останніх у рамках окремих груп. Це не є перешкодою для побудови вимірювальної моделі, покликаної підтвердити описану вище концептуальну схему. А ось з точки зору валідизації вимірювальних методик це накладає певні обмеження⁹⁴. Тому для виокремлення комплексних вимірювальних інструментів у ЄСД ми використали два принципи — належність до однієї групи або виміру, а також ідентичність шкали відповідей. Грунтуючись на них, можна виокремити такі вимірювальні інструменти:

– методика вимірювання емоційного благополуччя (усі шість індикаторів);

⁹⁴ У ЄСД ця суперечність нівельована простою стандартизацією, що не виглядає для нас переконливим.

- методика вимірювання енергійності (усі чотири індикатори);
- методика вимірювання функціональності (індикатори № 1–3, 7, 9–13).

У всіх трьох випадках відповідні індикатори належать до загального конструктору і використовують однакові шкали відповідей. Можна припустити, що ці методики є вимірювальними шкалами. Для перевірки цього припущення необхідно оцінити їх теоретичну валідність, а також продемонструвати коректне відбиття індикаторами відповідних латентних змінних на підставі перевірки факторної валідності та внутрішньої узгодженості.

Не можна також залишити без уваги класичну шкалу вимірювання депресії — Center for Epidemiologic Studies Depression Scale (CES-D), що була використана в ЄСД. По-перше, на відміну від інших шкал, розглянутих на етапі підготовки модуля, CES-D використовує один із її варіантів (що складається з восьми пунктів), без виключення будь-яких індикаторів. По-друге, пункти CES-D автори модуля використовують для побудови двох окремих груп — емоційного благополуччя й енергійності, а повну методику — в рамках фінального звіту щодо оцінки рівня депресії респондентів (European Social Survey, 2015, р. 14). По-третє, безвідносно до особливостей ЄСД, CES-D становить інтерес, оскільки є одним із найпопулярніших інструментів для вивчення психічного здоров'я в соціальних науках за кордоном. Змістовно CES-D складається з чотирьох перших індикаторів емоційного благополуччя, трьох перших індикаторів енергійності, а також одного додаткового пункту, який не потрапив у жоден із наведених вище вимірів: «...відчували себе самотньо?».

§ 11.2. Оцінка теоретичної валідності, перевірка факторної валідності та внутрішньої надійності шкал, запропонованих у ЄСД

В ідеалі *теоретична валідність* має бути забезпечена у трьох її складових: очевидній, змістовній і композиційній⁹⁵. Але оскільки йдеться про вже наявні індикатори, то слід говорити не про досягнення або забезпечення компонентів теоретичної валідності, а про оцінку тих її складових, які надають таку можливість. Тому тут ми зупинимось на очевидній і змістовній валідності.

Оцінюючи очевидну валідність, необхідно розуміти, чи співвідносяться використовувані концепти (емоційне благополуччя, енергійність і функціональність) із феноменами, що мають буттєву природу. Повертаючись до описаної вище концептуальної моделі С. Томпсона і Н. Маркса, зазначимо, що можемо оперувати лишень універсальним досвідом, оскільки можливість проведення спеціального якісного дослідження виходить за рамки нашого аналізу. На наш погляд, використовувана теоретична модель виглядає переконливо з наукової точки зору і логічно з точки зору повсякденного досвіду. Як гедоністичне, так і еудаймічне благополуччя відіграють важливу роль у житті індивідів.

⁹⁵ Докл. див.: (Дембицкий, 2016с, с. 22–31).

Труднощі використання цих концептів починаються там, де вони оформляються в теоретичні конструкти. Так, аналіз змістовної валідності вказує на низку проблемних моментів. По-перше, індикатори емоційного благополуччя і використовувана для їх оцінки шкала відповідей є нічим іншим, як методикою вимірювання психологічного дистресу, в основі якої лежать симптоматичні вимірювання депресії та тривожності. І оскільки аналогічні інструменти вже давно розроблені, наприклад, К6 Р. Кесслера (Kessler et al., 2002), рішення використовувати новий виглядає спірним. Окрім того, одна справа говорити про емоційне благополуччя і зовсім інша — про психологічний дистрес. Кожного разу концептуальні акценти істотно відрізнятимуться. По-друге, шкала енергійності запозичує три чверті своїх індикаторів, а також шкалу відповідей у згаданій раніше CES-D, що робить першу дуже скороченою версією методики для вимірювання депресії. Це тягне за собою тільки що згадані методологічні труднощі. По-третє, індикатори функціональності взагалі виглядають спірно, якщо йдеться про еудаймічне благополуччя, — надто великий акцент зроблено на установки і надто малий — на активність.

Тому з теоретичного боку всі три шкали виглядають непереконливо. Незважаючи на таке оригінальне ставлення питання про різні аспекти благополуччя індивідів, реальний методологічний фокус звернений до сфери психічного здоров'я, що ставить під сумнів не тільки змістовну, а й очевидну валідність вимірювальних інструментів.

Разом з тим усі три шкали можна розглядати у відриві від концепції С. Томпсона і Н. Маркса. У цьому разі аргументом на користь їх якості будуть прийнятні результати факторної валідації. З цією метою нами був проведений відповідний аналіз.

Факторна валідність перевірена на підставі результатів конфірмаційного факторного аналізу (*Confirmatory Factor Analysis* — CFA) з використанням методу діагонально зважених найменших квадратів (*Diagonally Weighted Least Squares* — DWLS). Якість факторних моделей оцінено на підставі таких показників: значення χ^2 -квадрат, величина середньоквадратичної похибки апроксимації (RMSEA), значення порівняльного індексу відповідності (CFI) та індексу Такера — Левіса (TLI). Як прийнятні граничні значення було прийнято: відношення значення χ^2 -квадрат до кількості ступенів свободи (df) менше 5, RMSEA < 0,07 й (CFI, TLI) > 0,95.

Оцінку внутрішньої надійності методик здійснено за допомогою коефіцієнта Альфа Кронбаха. Значення, що перевищують 0,7, розглянуто як прийнятні. Стосовно цього показника слід відзначити те, що індикатори в рамках кожної шкали були уніфіковані з точки зору їх «полярності». Мається на увазі необхідність «перевороту» шкали відповідей, що належать до всіх індикаторів із позитивними формулюваннями або до всіх індикаторів із негативними формулюваннями. Якщо в разі перевірки факторної валідності такі перетворення ніяк не впливають на результати, то в разі підрахунку коефіцієнта Альфа Кронбаха — це принципова вимога, недотримання якої веде до отримання артефактних показників. Відповідні результати наведено далі (табл. 11.2).

Для аналізу використана мова програмування R, версія 3.4.1 (бібліотеки “lavaan” для CFA, “lrm” для розрахунку коефіцієнтів Альфа Кронбаха).

Таблиця 11.2

Результати оцінки внутрішньої надійності та факторної валідності різних шкал благополуччя (Україна)

Показник	Емоційне благополуччя (N = 1821)	Енергійність (N = 1909)	Функціональність (N = 1831)
Альфа Кронбаха (> 0,70)	0,786	0,673	0,736
Хі-квадрат (DWLS)	862,545	2,064	663,457
df (ступені свободи)	9	2	27
Хі-квадрат/df	95,8	1,0	24,6
p-value	0,000	0,356	0,000
RMSEA (< 0,07)	0,228	0,004	0,113
CFI (> 0,95)	0,924	1,000	0,939
TLI (> 0,95)	0,873	1,000	0,918

Як видно, найгіршу якість показала шкала емоційного благополуччя — жоден параметр, за винятком коефіцієнта Альфа Кронбаха, не перевищує встановленої межі. Шкала енергійності, навпаки, має відмінні параметри факторної валідності та внутрішню узгодженість дещо нижчу за норму. Останнє, втім, пов’язане з малою кількістю індикаторів методики. Шкала функціональності показує слабкі результати за всіма параметрами, крім коефіцієнта Альфа Кронбаха. Перш ніж зробити остаточні висновки про якість методик, перевіримо їх факторну валідність і внутрішню узгодженість на підставі масивів двох інших країн — Російської Федерації та Великої Британії. Вибір Росії пов’язаний із тим, що у відповідному дослідженні використано практично ідентичний переклад анкети на російську. Що стосується Великої Британії, то там головним аргументом є оцінка вихідної методики, на підставі якої зроблено переклади в усіх країнах-учасницях ЄСД.

Отримані результати (табл. 11.3, 11.4) відрізняються від наведених раніше. Підсумок перевірки факторної валідності та внутрішньої узгодженості шкали емоційного благополуччя є дещо кращим для даних опитування в Російській Федерації та найкращим у Великій Британії. При цьому показник RMSEA і там, і там залишається незадовільним. Показники факторної валідності для шкали енергійності в обох країнах, як і в попередньому випадку, перебувають на дуже високому рівні, а внутрішньої узгодженості — трохи нижчі за норму. Що стосується шкали функціональності, то результати додаткових перевірок є дещо гіршими для даних опитування в Росії, — величина TLI є нижчою за необхідний рівень, а в іншому показники схожі.

Таблиця 11.3

**Результати оцінки внутрішньої надійності та факторної валідності
різних шкал благополуччя (Росія)**

<i>Показник</i>	<i>Емоційне благополуччя (N = 1821)</i>	<i>Енергійність (N = 1909)</i>	<i>Функціональність (N = 1917)</i>
Альфа Кронбаха (> 0,70)	0,800	0,674	0,740
Хі-квадрат (DWLS)	641,828	4,528	535,742
<i>df</i> (ступені свободи)	9	2	27
Хі-квадрат/ <i>df</i>	71,3	2,3	19,9
<i>p</i> -value	0,000	0,104	0,000
RMSEA (< 0,07)	0,178	0,024	0,093
CFI (> 0,95)	0,957	0,999	0,951
TLI (> 0,95)	0,928	0,998	0,935

Таблиця 11.4

**Результати оцінки внутрішньої надійності та факторної валідності
різних шкал благополуччя (Велика Британія)**

<i>Показник</i>	<i>Емоційне благополуччя (N = 2268)</i>	<i>Енергійність (N = 2267)</i>	<i>Функціональність (N = 2237)</i>
Альфа Кронбаха (> 0,70)	0,820	0,688	0,759
Хі-квадрат (DWLS)	257,413	9,732	561,894
<i>df</i> (ступені свободи)	9	2	27
Хі-квадрат/ <i>df</i>	28,6	4,866	20,8
<i>p</i> -value	0,000	0,008	0,000
RMSEA (< 0,07)	0,110	0,041	0,094
CFI (> 0,95)	0,987	0,998	0,964
TLI (> 0,95)	0,978	0,993	0,952

Отже, успішною можна назвати перевірку внутрішньої узгодженості та факторної валідності тільки для шкали енергійності. Для двох інших шкал результати слід назвати в найкращому разі суперечливими. Особливо це стосується шкали емоційного благополуччя, якщо йдеться про російсько- або україномовну версію методики.

Говорячи про шкалу емоційного благополуччя, вихідна гіпотеза, що пояснює непереконливі результати її факторної валідації, стосувалася наявності в її змісті двох симптоматичних вимірів — депресії (перші чотири пункти) і тривожності (п'ятий і шостий пункти). І хоча таку постановку питання з методологічної точки зору ми вбачаємо правдоподібною і зараз, виявлена в ем-

піричних даних двофакторність має іншу підставу — спосіб формулювання індикаторів. Так, у даних опитувань, проведених в Україні та Російській Федерації, індикатори з позитивними і негативними формулюваннями формують дві кореляційні групи. У разі Великої Британії це виражено значно слабше. Беручи до уваги таку особливість емпіричних даних, перевіримо факторну валідність шкали емоційного благополуччя на підставі двофакторної структури (табл. 11.5) — до першої включено всі індикатори з позитивними формулюваннями (№ 2, 3, 6), до другої — з негативними (№ 1, 4, 5)⁹⁶.

Таблиця 11.5

Результати оцінки факторної валідності шкали емоційного благополуччя в різних країнах (двофакторне рішення)

Показник	Україна	Росія	Велика Британія
Хі-квадрат (DWLS)	58,677	151,502	88,994
<i>df</i> (ступені свободи)	8	8	8
Хі-квадрат/ <i>df</i>	7,3	18,9	11,2
<i>p</i> -value	0,000	0,000	0,000
RMSEA (< 0,07)	0,059	0,090	0,067
CFI (> 0,95)	0,995	0,990	0,996
TLI (> 0,95)	0,992	0,982	0,992

У всіх трьох випадках якість моделі значно покращилася. Особливо показовими є зміни для України. Говорячи про досягнення необхідних меж, лише показник RMSEA для Росії залишився вищим за встановлене значення. Такі результати відкривають простір для нових інтерпретацій, а також свідчать про необхідність подальшого аналізу на підставі даних інших країн-учасниць ЄСД. Однак перш ніж висунути можливі причини отримання таких результатів, оцінимо факторну валідність шкали функціональності, також використавши двофакторну структуру на підставі позитивних і негативних формулювань індикаторів (табл. 11.6).

Як і у випадку шкали емоційного благополуччя, використання двофакторного рішення для шкали функціональності приводить до поліпшення результатів перевірки факторної валідності. Разом з тим воно не має такого вираженого характеру (хоча в разі Росії всі показники потрапляють у задані межі). Тому ми спробуємо визначити можливі причини двофакторності тільки для шкали емоційного благополуччя.

⁹⁶ Слід зазначити, що Ф. Хуперт, яка відповідала за розробку модуля особистісного і соціального благополуччя, була добре обізнана про сутнісну відмінність між позитивними і негативними компонентами особистісного благополуччя (Huppert & Whittington, 2003). Однак у підсумкових презентаційних документах ЄСД це не знайшло відображення.

Таблиця 11.6

**Результати оцінки факторної валідності шкали функціональності
в різних країнах (двофакторне рішення)**

<i>Показник</i>	<i>Україна</i>	<i>Росія</i>	<i>Велика Британія</i>
Хі-квадрат (DWLS)	408,467	280,377	477,986
<i>df</i> (ступені свободи)	26	26	26
Хі-квадрат/ <i>df</i>	15,7	10,8	18,4
<i>p</i> -value	0,000	0,000	0,000
RMSEA (< 0,07)	0,090	0,067	0,088
CFI (> 0,95)	0,963	0,976	0,969
TLI (> 0,95)	0,949	0,966	0,958

Оцінюючи факторну структуру такого конструкту, як емоційне благополуччя, можна висунути кілька пояснювальних гіпотез. Перша полягає в тому, що негативні та позитивні відчуття справді формують два важливі аспекти емоційного благополуччя. Це може проявлятися як у тому, що відсутність негативних емоцій ще не гарантує наявності позитивних переживань, а саме таку гіпотезу висловили автори модуля (Huppert et al., 2005, p. 16), так і в наявності позитивних емоцій, попри наявність негативних переживань. Відповідно, оцінювати емоційне благополуччя лише на підставі негативного емоційного досвіду (психологічного дистресу) не правильно. Друга гіпотеза пов'язана з тим, як респонденти реагують на пункти анкети з негативною і позитивною тональністю. Якщо оцінювати питання про емоційне благополуччя як такі, що викликають побоювання респондентів (така ймовірність існує для обох груп відповідних індикаторів), то наявність двофакторної структури може виявитися результатом дії захисних механізмів, а не відображенням досліджуваного феномену. Третє можливе пояснення полягає в тому, що індикатори з позитивними і негативними формулюваннями надають два окремі феномени, що входять у царину особистісного благополуччя і, відповідно, їх розгляд як єдиного явища в корені є неправильним. Оскільки перевірка всіх цих гіпотез є неможливою в рамках цього дослідження, зупинимось на частковій перевірці першої з них.

З метою проведення відповідного аналізу всі змінні, що характеризують позитивні та негативні переживання, перекодовані у дихотомічні. До першої категорії віднесено респондентів, які дали відповіді «ніколи або майже ніколи» або «меншу частину часу», у другу — «більшу частину часу» або «весь або майже весь час». Далі для кожного респондента окремо у групі позитивних і негативних переживань було розраховано суму відповідей, які потрапили у другу категорію. Так, було отримано дві змінні з діапазоном можливих значень від 0 (переживання відповідної групи не спостерігаються) до 3 (наявні всі переживання відповідної групи). Їх спільний аналіз (табл. 11.7) дозволяє зробити кілька висновків про зв'язок негативних переживань або психологічного дистресу і позитивних переживань, що характеризують, на думку авторів модуля, справжнє емоційне благополуччя.

Таблиця 11.7

**Розподіл кількості позитивних афектів у групах респондентів
щодо кількості негативних афектів, %**

Кількість позитивних афектів	Кількість негативних афектів											
	Нуль			Один			Два			Три		
	УК	РФ	ВБ	УК	РФ	ВБ	УК	РФ	ВБ	УК	РФ	ВБ
Нуль	5,8	7,3	10,8	17,0	25,8	38,5	33,2	43,7	50,0	55,8	77,1	79,5
Один	9,3	9,6	9,8	26,2	27,0	21,1	19,3	22,1	29,7	14,1	10,8	15,9
Два	20,5	18,4	27,4	25,2	30,4	23,4	23,4	25,1	14,1	13,5	6,4	4,5
Три	64,4	64,6	51,9	31,5	16,9	17,0	24,2	9,0	6,2	16,7	5,7	0,0
<i>Середнє значення</i>	2,44	2,40	2,20	1,71	1,38	1,17	1,39	0,99	0,77	0,91	0,41	0,25

По-перше, для найбільшої групи респондентів у вибірці (в Україні, Росії та Великій Британії їхня частка становить відповідно 59,7%, 69,0% і 85,6%), яким властива відсутність негативних переживань, є характерною наявність позитивних емоцій. Більшою мірою це властиво респондентам українського і російського опитувань — лише 5,8% цієї групи в першому випадку і 7,3% — у другому не були щасливі, не отримували задоволення від життя і не відчували себе спокійно і мирно досить часто. Водночас абсолютна більшість респондентів цих груп відчувала все вказане. Що стосується результатів, отриманих у Великій Британії, то кількість спостережень із відсутністю позитивних переживань у відповідній групі є трохи вищою, а кількість спостережень, що відзначають усі три позитивні переживання, — трохи нижчою.

По-друге, у трьох групах респондентів, для яких характерна наявність одного і більше негативного переживання, ситуацію в різних країнах не можна назвати схожою. Щоб описати відмінності між ними, можна звернутися до принципу «чим більше негативних переживань, тим менше позитивних». Цей принцип найменше стосується України і найбільше — Великої Британії, а Росія посідає проміжне місце. Простір можливих інтерпретацій тут досить великий і, як нам видається, для вибору правильної необхідні додаткові дослідження. Тому тут обмежимося вказівкою на той факт, що зазначений принцип зв'язку переживань різної природи працює тим краще, чим більше спостережень усієї вибірки потрапляє в категорію з відсутністю негативних переживань.

Підбиваючи підсумки перевірки гіпотези, треба зазначити, що відсутність негативних переживань усе ж таки щільно пов'язана з наявністю позитивних. При цьому за збільшення негативних афектів кількість позитивних знижується. Однак таке зниження є різним у різних країнах. Ця ситуація задає перспективи подальших досліджень. Перша перспектива пов'язана з дослідженням специфіки взаємозв'язку між позитивними і негативними емоційними афектами в тому вигляді, в якому її подано в ЄСД. Друга, можливо, є ціка-

вішою в науковому значенні, вона полягає в пошуку справді самостійного фактора позитивних переживань, що слабо корелює з індикаторами психологічного дистресу.

§ 11.3. Оцінка теоретичної валідності, перевірка факторної валідності та внутрішньої надійності CES-D

CES-D являє собою коротку шкалу для самозаповнення, розроблену Л. Редлофф для вимірювання депресивної симптоматики серед населення. Пункти методики відібрано на підставі існуючих шкал, які пройшли валідацію. Отже, головне призначення шкали — не діагностика психічного розладу або визначення його тяжкості, а оцінка поточного рівня симптоматики з акцентом на афективному компоненті та депресивних аспектах настрою (Radloff, 1977).

Незважаючи на те, що стаття, присвячена розробці та валідації CES-D, вийшла 1977 року, вона та її коротші варіанти залишаються найбільш затребуваним інструментом для оцінки психологічного дистресу серед загального населення і його окремих груп. Аналіз літератури показує, що частота його використання є навіть вищою, ніж у сучасніших шкал психологічного дистресу Р. Кесслера (K10 і K6) (Дембіцький, 2018b, с. 38–39). Це підтверджує і досвід проведення ЄСД, автори якого зупинилися на варіанті CES-D, що складається з восьми індикаторів.

Оскільки підбір індикаторів автором методики ґрунтувався на загальноприйнятому розумінні симптомів депресії, питання теоретичної валідності в її очевидному і змістовному компонентах слід вважати розв'язаним. Окрім цього в низці досліджень вивчено і внутрішню структуру повного варіанта CES-D, що дозволяє розглянути і композиційний компонент. Л. Редлофф виокремила чотири фактори: 1) депресивні афекти; 2) позитивні афекти; 3) соматичну активність і активність, для якої характерні ускладнення; 4) міжперсональний фактор (Radloff, 1977, р. 397–398). Надалі факторна структура CES-D неодноразово ставала предметом різних досліджень, результати яких узагальнено у статті Л. Рос (Ros et al., 2011, р. 85–87). У більшості досліджень було підтверджено 4-факторну структуру, виявлену Л. Редлофф. Були й альтернативні варіанти. Проте незалежно від дослідження в усіх результатах подано фактор позитивних афектів за наявності інших. Тому у скороченій версії CES-D, яку використовували в ЄСД, логічно орієнтуватися як мінімум на два фактори — позитивних і негативних афектів (про більшу кількість факторів недоцільно говорити через малу кількість індикаторів). На користь цього свідчить і проведений вище аналіз інших шкал ЄСД. Відповідно, перевірку факторної валідності буде здійснено як для однофакторного⁹⁷, так і для двофакторного рішення.

⁹⁷ Європейські науковці, що вивчали особливості використання CES-D у Європейському соціальному дослідженні, схилиються до застосування саме однофакторного рішення (Bracke, Levecque, & Van de Velde, 2008; Van de Velde, Levecque, & Bracke, 2009; Missinne, Vandeviver, Van de Velde, & Bracke, 2014; Karim, Weisz, Bibi, & Rehman, 2015).

Результати перевірки внутрішньої надійності та факторної валідності CES-D (табл. 11.8) дозволяють зробити як мінімум три висновки: 1) шкала показує високу внутрішню узгодженість у всіх випадках (за умови приведення формату відповідей до єдиного стандарту); 2) однофакторна структура є очевидно не адекватною емпіричним даним; 3) двофакторна структура, навпаки, показує високу відповідність; при цьому ступінь поліпшення якості моделей у всіх випадках різний.

Таблиця 11.8

Результати оцінки внутрішньої надійності та факторної валідності CES-D у різних країнах із використанням одно- і двофакторного рішень

Показник	Україна	Росія	Велика Британія
Альфа Кронбаха (> 0,70)	0,831	0,830	0,820
Хі-квадрат (DWLS)	1264,729 / 122,907	1170,506 / 81,495	560,045 / 144,238
df (ступені свободи)	20 / 19	20 / 19	20 / 19
Хі-квадрат/df	63,2 / 4,3	58,5 / 6,5	28,0 / 7,6
p-value	0,000 / 0,000	0,000 / 0,000	0,000 / 0,000
RMSEA (< 0,07)	0,170 / 0,039	0,180 / 0,055	0,109 / 0,054
CFI (> 0,95)	0,944 / 0,997	0,941 / 0,995	0,976 / 0,994
TLI (> 0,95)	0,921 / 0,996	0,917 / 0,992	0,967 / 0,992

Спробуємо внести ясність щодо факторної структури CES-D, поділивши її на дві відповідні підшкали. Оскільки у підшкалу з індикаторами позитивного характеру входить лише два пункти (щастя, задоволення від життя), її легко категоризувати. Для цього, наприклад, шкалу відповідей «ніколи або майже ніколи», «меншу частину часу», «більшу частину часу», «весь або майже весь час» можна перетворити на дихотомічну — «ніколи або майже ніколи» і «меншу частину часу» / «більшу частину часу» і «весь або майже весь час». Маючи два індикатори з дихотомічною шкалою відповідей, усі респонденти у загальному вигляді поділяються на три групи:

- з відсутністю позитивних афектів, які вони відчували більшу частину часу, весь або майже весь час;
- з одним таким афектом;
- з двома такими афектами.

На підставі решти шести індикаторів, що характеризують негативні переживання, можна підрахувати адитивний індекс, великі значення якого свідчать про вищу виразність депресивної симптоматики. Далі проаналізуємо середні значення цього адитивного індексу для трьох категорій респондентів, виокремлених на підставі першої підшкали (табл. 11.9).

Таблиця 11.9

**Середній рівень вираженості депресивної симптоматики
для респондентів із різною кількістю позитивних афектів**

Кількість позитивних афектів	Рівень депресивної симптоматики		
	Україна	Росія	Велика Британія
Нуль	14,6	13,4	12,9
Один	12,2	11,5	10,9
Два	10,5	9,5	8,4

Загалом такі результати можна назвати інтуїтивно зрозумілими. Але погляньмо на ті самі дані під іншим кутом. Для цього категоризуємо також під шкалу на підставі негативних формулювань. Найпростіший спосіб зробити це — розбити розподіл її значень у кожному з трьох масивів на підставі відповідного середнього значення, що дасть нам дві групи респондентів — з вираженістю депресивної симптоматики вище і нижче середнього. Отриману дихотомічну змінну використаємо для побудови таблиці спряженості (табл. 11.10). Патерни, отримані по кожному з трьох масивів, є схожими. При цьому тільки результати для підгруп респондентів із вираженістю негативної симптоматики, нижчою за середню, можна назвати пропорційними результатами, які подано в таблиці 11.9. Натомість для тих респондентів, рівень негативних афектів яких вищий за середній, спостерігаємо зовсім іншу модель. Фактично для них більшою мірою характерна наявність позитивних відчуттів, аніж їх відсутність.

Таблиця 11.10

Зв'язок негативних і позитивних афектів (CES-D), %

Кількість позитивних афектів	Виразність негативних афектів					
	Україна		Росія		Велика Британія	
	Менше за середн. зн.	Більше за середн. зн.	Менше за середн. зн.	Більше за середн. зн.	Менше за середн. зн.	Більше за середн. зн.
Нуль	8,6	33,8	10,3	39,2	4,9	36,0
Один	18,6	25,9	14,8	21,1	8,3	21,7
Два	72,8	40,3	74,9	39,7	86,8	42,3

Отже, для CES-D є характерною висока теоретична і факторна валідність, а також висока внутрішня узгодженість. Питання, однак, викликає специфіка факторної структури методики. Можливі відповіді на нього ті самі, що й у разі шкали емоційного благополуччя, — двоаспектність емоційного досвіду; зміщення, викликані специфікою запитань; зведення в рамках однієї методики двох феноменів.

§ 11.4. Висновки

Узагальнюючи подані результати, вкажемо на низку аспектів вимірювання особистісного і соціального благополуччя в ЄСД — концептуальний, методологічний і методичний. Розглянемо кожен із них.

Незважаючи на оригінальність концептуального фокусування, в теоретичних положеннях С. Томпсона і Н. Маркса не враховано інституційну специфіку досліджуваного суспільства (специфіка соціальних і національних відносин, стан соціальної безпеки, особливості соціально-політичних процесів і т. д.), тобто упущений пласт соціологічної інформації, що має безпосередній стосунок до формування соціального благополуччя респондентів. Ані матеріальне, ані соціально-психологічне або психологічне благополуччя індивіда саме по собі не дозволяє аналізувати успішність розвитку і функціонування суспільства, декларованими цілями якого є дотримання демократичних прав громадян.

Говорячи про методологію, слід зазначити істотну еkleктичність у розгляді феномену особистісного благополуччя через те, що ключові інструменти модуля (шкала емоційного благополуччя, шкала енергійності) сконструйовані на базі CES-D, яка є шкалою, широко використовуваною для вимірювання психологічного дистресу в соціальних науках. Останнє, своєю чергою, базується на зовсім інших концептуальних і методологічних позиціях порівняно з тими, які використовують у ЄСД. Можна сказати, що соціологічне вивчення особистісного і соціального благополуччя (як воно подане в ЄСД) і соціологічне вивчення психологічного дистресу (як воно подане в західній соціології⁹⁸) є альтернативними підходами до вивчення успішності індивідуальної життєдіяльності в сучасному суспільстві. Тому використовувані в ЄСД інструменти і ті феномени, які вони покликані відображати, в методологічному сенсі відірвані одне від одного.

За розгляду методичного рівня і взагалі складається враження, що автори модуля прагнули ускладнити роботу тим, кого зацікавлять результати вивчення особистісного і соціального благополуччя в ЄСД. По-перше, індикатори різних вимірів благополуччя уніфіковані лише частково, що накладає обмеження на перевірку їх якості. По-друге, там, де вони уніфіковані, відповідна перевірка залишає більше питань, аніж дає відповідей. Так, незважаючи на обізнаність про двофакторну структуру використовуваних конструктів (на важливість цього моменту вказано в підготовчих документах), це питання не знайшло ніякого освітлення в підсумкових звітах, присвячених модулю. По-третє, спосіб підсумкової обробки результатів кожного виміру благополуччя (Z-перетворення або стандартизація) взагалі відірваний від попередніх етапів роботи.

Отже, незважаючи на велику кількість матеріалу, зібраного про різні аспекти індивідуального благополуччя, відповідне концептуальне, методологічне і методичне забезпечення виглядає як мінімум спірним і навіть суперечливим. Можливим розв'язанням зазначених проблем є звернення до теоретико-мето-

⁹⁸ Докл. див.: (Дембіцький, 2018b).

дологічної рамки соціології психологічного дистресу. У цьому контексті перспективним виглядає оцінка шкали емоційного благополуччя і CES-D в межах аналізу різних компонентів процесу стресу, а також оцінка якості цих шкал у різних європейських країнах.

Висновки до другого розділу

У репрезентованому розділі розглянуто і проаналізовано різні методики оцінки індивідуального благополуччя, які використовують як у вітчизняній, так і зарубіжній соціології. Головну увагу приділено таким аспектам індивідуального благополуччя, як психологічний дистрес і соціальне самопочуття.

У рамках вимірювання психологічного дистресу обрано методику Л. Дерогатиса SCL-90-R, оскільки її застосовували в низці широкомасштабних опитувань в Україні. На підставі відповідних результатів відібрано дев'ять індикаторів, які лягли в основу розробки соціологічного тесту SCL-9-NR, що задовольняє вимоги соціологічного тестування. Для цього було проведено методичні дослідження, які підтвердили критеріальну і конструктну валідність тесту, його внутрішню і ретестову надійність, а також дозволили визначити критеріальні рівні SCL-9-NR.

У закордонних дослідженнях вимірювання психологічного дистресу найпопулярнішою методикою є шкала CES-D (Л. Редлофф), спрямована на вимірювання депресивної симптоматики. Використання цієї шкали в Європейському соціальному дослідженні дозволило оцінити її вимірювальні властивості серед українського населення. Отримані результати підтвердили її факторну валідність і внутрішню узгодженість, проте залишили питання з приводу специфіки факторної структури. Крім того, порівняно з SCL-9-NR для цієї методики є характерною вужча змістовна валідність.

За вимірювання соціального самопочуття увагу приділено методиці Є. Головахи і Н. Паніної «Інтегральний індекс соціального самопочуття» (ІСС). З огляду на багатий і успішний досвід її використання у вітчизняній соціології методична робота була зосереджена на конструюванні ще одного скороченого варіанта ІСС, а також на альтернативній обробці його результатів. Це дозволило запропонувати ІСС-13, сфокусований на оцінці соціального самопочуття в контексті ефективності соціальних інститутів, а також визначити критерії класифікації індивідуальних результатів. Якщо ж говорити про зарубіжну практику вивчення соціального благополуччя, то аналогів ІСС у ній не спостерігається.

Окрему увагу в розділі приділено використанню L-шкали (ММРІ) серед населення України. Незважаючи на суперечливі результати оцінки внутрішньої надійності методики, ми дійшли висновку про можливість її використання в масових опитуваннях, а також запропонували вагові коефіцієнти для різних індикаторів залежно від соціально-демографічної характеристики респондентів. Окрім того, сконструйовано скорочений варіант цієї шкали.

На завершення проведеної роботи можна констатувати багатий методичний досвід вивчення різних аспектів індивідуального благополуччя, накопичений у вітчизняній соціології, а також істотні можливості його подальшого розширення в контексті міжнародної практики соціологічних опитувань.

Розділ III

ПОЛІТИЧНА КУЛЬТУРА У ФОКУСІ СОЦІОЛОГІЧНОГО ТЕСТУВАННЯ

Щоб зробити подальшу методичну роботу максимально сфокусованою, як відповідну точку ми використали наукові результати, отримані фахівцями відділу соціально-політичних процесів Інституту соціології НАН України під керівництвом Є. Головахи і Н. Паніної з 2002 по 2005 роки.

В межах цього дослідницького проекту політичну культуру визначили як «...сукупність стійких форм політичної свідомості і поведінки, яка визначає особливості функціонування різних суб'єктів політичного процесу в рамках конкретної політичної системи. Під стійкими формами політичної свідомості розуміються політичні орієнтації, установки, переконання (на відміну від мінливих думок, настроїв, оціночних суджень, які стихійно формуються). Стійкі форми політичної поведінки — це вироблені в процесі політичної діяльності типові поведінкові реакції (стійкі форми участі в певних політичних організаціях, кампаніях, акціях)» (Головаха, 2005, с. 5).

Для типології політичної культури використано функціональний підхід, що дозволив виокремити дві визначальні функції політичної культури — орієнтаційну (ставлення до демократії) й активаційну (ставлення до громадської активності). Їх комбінація веде до утворення чотирьох типів політичної культури — активної демократичної, пасивної демократичної, пасивної антидемократичної, активної антидемократичної (Головаха, 2005, с. 39–40).

Для вивчення зазначених типів політичної культури було розроблено дві шкали — для оцінки виразності демократичних установок і установок на громадську активність, які є основою соціологічного тесту «Типи політичної культури».

Перша частина цього розділу присвячена аналізу зазначеного тесту, а також розвитку ідей, покладених у його основу, що привело до розробки соціологічного тесту «Типи політичної культури — II» (далі СТ ТПК-II).

Значну увагу також приділено соціологічному вимірюванню геополітичних орієнтацій, що, на наш погляд, вписуються в категорію політичної культури. Так, якщо демократичні установки й установки на громадську активність описують внутрішньополітичну ситуацію, то геополітичні орієнтації — зовнішньополітичну в тій її частині, яка стосується широких верств населення.

Актуальність вивчення геополітичних орієнтацій продиктовано не тільки академічною доцільністю, а й поточним історичним становищем України, яке потребує як ефективної зовнішньополітичної діяльності українських еліт, так і конструктивної позиції громадян країни з відповідних питань.

З огляду на слабку розробленість проблематики вимірювання геополітичних орієнтацій в Україні з використанням соціологічних тестів ми зробили спробу конструювання авторського підходу до розв'язання цієї проблеми, основу якого закладено в недавніх публікаціях (Дембіцький & Любива, 2017b; Dembitskyi, 2015; Дембіцький & Любива, 2014).

Глава 12

Соціологічний тест «Типи політичної культури»: вплив способу кластеризації на результати типології

§ 12.1. Вступ

Соціологічний тест «Типи політичної культури» (далі — СТ ТПК) спрямований на вивчення сталих форм політичної свідомості та поведінки людей, поданих такими концептуальними вимірами, як громадська активність і підтримка демократії (автор — Є. Головаха). Фінальний варіант методики ґрунтований на 12 твердженнях (див. Додаток 4) і використанні шкали Лайкерта: «цілком згоден», «скоріше згоден», «важко сказати, згоден чи ні», «скоріше не згоден», «абсолютно не згоден».

Процес створення і валідизації, а також специфіку застосування тесту докладно розглянуто у статті О. Стегнія (Стегній, 2015), а також коротко описано в першому розділі цієї монографії. Тому тут ми спинимося лише на одному процедурному аспекті використання СТ ТПК, що видається проблемним, а саме на оригінальній техніці виокремлення типів політичної культури.

Відповідно, метою глави є порівняльний аналіз різних підходів до кластеризації респондентів на підставі результатів застосування СТ ТПК.

Оригінальний підхід кластеризації, запропонований А. Горбачиком, полягає у використанні швидкого кластерного аналізу з застосуванням рухливих (нефіксованих) центрів (*K-means*). З приводу такого вибору А. Горбачик зазначає: «...краще використовувати результати кластерного аналізу, а не класифікацію через точку поділу. Класифікація кластерним аналізом більшою мірою зорієнтована на властивості емпіричних даних, вона “відносна” (тобто це “демократи стосовно інших людей в Україні”, “більш демократично налаштовані, ніж інші”) й не потребує теоретичного обґрунтування норм (що таке “демократ”, як треба відповісти на наш тест, щоб бути демократом)» (Горбачик, 2006, с. 5).

Нам така думка видається вельми спірною. Передусім це пов’язане з тим, що метод швидкого кластерного аналізу не враховує того концептуального змісту, що закладений у його індикаторах і шкалі відповідей. У підсумку фінальні результати визначають не так концептуальним соціологічним аналі-

зом, покладеним в основу конструювання тесту, як особливостями абстрактних математичних процедур. А це надто спірно.

§ 12.2. Способи кластеризації

Вище визначено не тільки основний спосіб кластеризації — метод швидкого кластерного аналізу, а й одну з його альтернатив — ідеальну точку поділу. Останній спосіб полягає в тому, що обирають певну координату, що задає точку перетину двох шкал тесту — активності та демократії, внаслідок чого утворюються чотири сектори, що позначають різні типи політичної культури. І хоча в цьому разі можна говорити про норми, вони більшою мірою є математичними, ніж концептуальними. Розгляньмо цей момент докладніше.

У випадку СТ ТПК маємо дві адитивні шкали з діапазоном можливих значень від 6 до 30. Відповідно, значення, що ділить цей діапазон на дві рівні частини, дорівнює 18. Але числове значення «18», як і більшість інших, може мати найрізноманітніший змістовий смисл, на що тут не зважено. Наприклад: «18» — це шість разів по три, а по-іншому — шість разів по «важко сказати, згоден чи ні». Респондент узагалі не має конкретної думки з приводу того чи того виміру політичної культури. Також «18» — це три рази по одному («цілком згоден») і три рази по п'ять («цілком не згоден»). Тобто йдеться радше про деяку амбівалентність в установках респондента. Ці приклади демонструють, що математична схожість респондентів не має підмінити концептуальну.

Отже, способи кластеризації можуть різнитися за способом визначення схожості — він може бути математичним або змістовим. Окрім того, кластеризація може здійснюватися за спеціальними правилами, встановленими заздалегідь (апостеріорна кластеризація), або реалізовуватися через універсальні алгоритми, що потенційно ведуть до набору можливих рішень (апостеріорна кластеризація). Використовуючи ці параметри, можна виокремити принаймні чотири способи кластеризації результатів СТ ТПК (табл. 12.1).

Таблиця 12.1

Способи кластеризації результатів СТ ТПК

Спосіб визначення схожості	Тип кластеризації	
	Апостеріорний	Апостеріорний
Математичний	Ідеальна точка поділу	Кластерний аналіз
Змістовий	Концептуальна типологія	Латентний класовий аналіз

Як ми вже зазначали, *ідеальна точка поділу* задає координату, що розбиває всю сукупність респондентів на чотири групи. Можливим поліпшенням цього способу є впровадження деякої зони навколо точки поділу (далі це буде $\pm 2,5$ бала на обох осях), потрапляння в яку трактують як невизначену позицію стосовно демократії та громадської активності.

Концептуальна типологія полягає у використанні змістових принципів аналізу відповідей респондентів, що дають змогу згрупувати спостереження в рамках теоретично обґрунтованих груп. Для СТ ТПК ми використовували два такі принципи: 1) «віяло» відповідей «цілком згоден / скоріше згоден / важко сказати, згоден чи ні / скоріше не згоден / абсолютно не згоден» задає три основні модальності — згоди, невизначеності та незгоди; 2) шість індикаторів дають можливість респондентові висловити одну з чотирьох базових установок — підтримки (у відповідях переважає згода), амбівалентності (у відповідях рівною мірою присутня і згода, і незгода), спротиву (у відповідях переважає незгода), невизначеності (у відповідях переважає невизначеність).

Згідно з цими принципами, технічна реалізація концептуальної типології в рамках СТ ТПК полягатиме в перелічуванні всіх можливих комбінацій для шести індикаторів (табл. 12.2). При цьому ми не робили різниці між «цілком згоден» і «скоріше згоден», а також між «цілком не згоден» і «скоріше не згоден».

Таблиця 12.2

Концептуальна типологія СТ ТПК: технічне розв'язання

Підтримка	(1) Згода за чотирма і більше індикаторами (2) Згода за трьома індикаторами і невизначеність, принаймні, за одним із решти
Амбівалентність	(1) Згода за трьома індикаторами і незгода за трьома індикаторами (2) Однакова репрезентованість згоди, незгоди і невизначеності
Спротив	(1) Незгода за чотирма і більше індикаторами (2) Незгода за трьома індикаторами і невизначеність принаймні за одним із решти
Невизначеність	(1) Невизначеність за чотирма і більше індикаторами (2) Невизначеність за трьома індикаторами і наявність як згоди, так і незгоди серед решти

Оскільки використано дві шкали, кількість усіх типів дорівнює 16. Число груп є доволі великим, але воно цілковито покриває змістові характеристики СТ ТПК.

Кластерний аналіз за методом *K-середніх* являє собою підхід для розбивки деякого набору даних на K окремих кластерів, що не перетинаються. Ідея, що лежить в основі цього методу, базується на мінімізації внутрішньокластерної мінливості n -розмірного набору даних (Джеймс и др., 2016, с. 415–416). Для розбивки спостережень на k кластерів використано такий алгоритм: 1) перші k спостережень приймають як k кластерів з однією точкою в кожному кластері; 2) решта $m - k$ спостережень по чергово приписують до одного із k кластерів на підставі найкоротшої дистанції між спостереженням і центроїдом кластера; після кожного приписування центроїд відповідного кластера переобчислюють; 3) після приписування кожного спостереження до одного із k кластерів кластерні центроїди приймають як початкові точки і кожне спостереження знову приписують одному з k кластерів на підставі найближчої до центроїду

дистанції. Цей етап повторюють необхідну кількість разів (Lee, 2007, р. 504). Крім того, у кластерному аналізі методом *K-середніх* можна задати ідеальні центри, з яких починається описаний вище алгоритм. За опрацювання результатів СТ ТПК використовують саме ідеальні центри: (30, 30), (30, 6), (18, 18), (6, 6), (6, 30).

Латентний класовий аналіз є методом, або технікою ідентифікації неозначених груп індивідів, або спостережень у даних на підставі багатовимірних категоріальних змінних. Для цього використовують імовірнісну модель, що описує поділ відповідей за кількома запитаннями у різних групах. Оцінка параметрів моделі веде до прояснення груп, що лежать в основі такого розподілу, у термінах імовірнісних патернів відповідей на запитання (Larsen, 2007, р. 526).

Незважаючи на те, що в оригіналі СТ ТПК передбачає виокремлення чотирьох типів політичної культури, в рамках цієї роботи кластеризація передбачає п'ять (для ідеальної точки поділу, кластерного аналізу і латентного класового аналізу), а також дев'ять (для концептуальної типології) груп⁹⁹. Це зроблено з метою якомога повнішого врахування змістових характеристик СТ ТПК.

§ 12.3. Описові результати різних способів кластеризації¹⁰⁰

Характеризуючи загальний поділ відповідей у двовимірному просторі «ставлення до демократії / громадська активність», зручно вдатися до графічної презентації у вигляді теплової карти¹⁰¹. У нашому випадку вона показує наповненість усіх можливих категорій зазначеного двовимірного простору, обчислену як квадратний корінь відповідних частот¹⁰² (рис. 12.1). Загалом рисунок демонструє переважаючі демократичні установки над недемократичними, а також громадської пасивності над активністю. Звідси найбільш наповненою частиною діаграми є її ліва верхня чверть. При цьому модальна категорія репрезентована респондентами з максимальною установкою на демократію (30 балів) і мінімальною — на громадську активність (6 балів).

Тепер розглянемо групові характеристики різних рішень. По-перше, проаналізуємо наповненість різних груп (табл. 12.3). За розгляду цих результатів слід пам'ятати, що лише в разі концептуальної типології можна говорити про демократизм/антидемократизм і активність/пасивність в абсолютному смислі. В інших випадках такі судження відносні: для ідеальної точки поділу — меншою мірою (завдяки чіткості критерію), для кластерного аналізу і латентного класового аналізу — більшою.

⁹⁹ З аналізу виключено слабо наповнені групи. Розмір шести з них становить менше 1,0%, однієї — 1,2%.

¹⁰⁰ Тут і далі всі обчислення і побудова діаграм здійснено за допомоги мови програмування R.

¹⁰¹ Докладно про побудову теплової карти див.: (Мастицкий, 2016, с. 105–106).

¹⁰² Рішення знайти квадратний корінь ухвалено з тієї причини, що в даних є одна сильно наповнена порівняно з іншими категорія ($X = 1$, $Y = 30$). Це, своєю чергою, погіршує графічну презентацію даних.

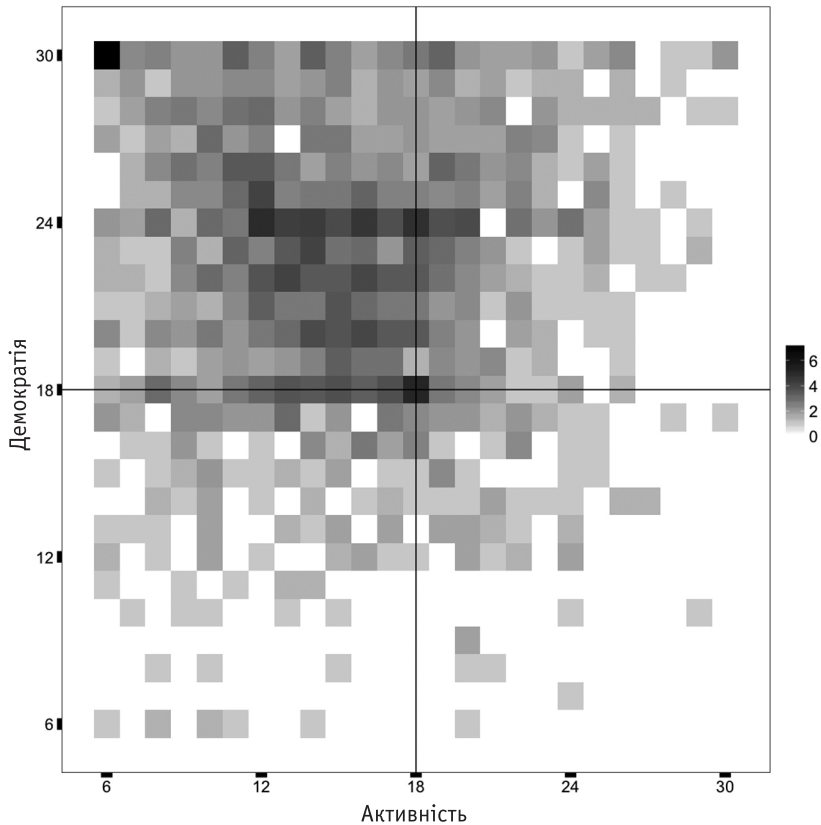


Рисунок 12.1. Теплова діаграма частотної наповненості категорій установок на демократію і громадську активність

Таблиця 12.3

Наповненість груп залежно від різних рішень кластеризації

Група	ІТД	КТ	КА	ЛКА
Активні демократи	22,4	14,7	17,2	14,3
Пасивні демократи	56,1	43,7	21,1	14,5
Не визначилися в обох смислах	10,0	5,2	30,3	—
Пасивні антидемократи	7,4	4,5	14,3	22,9
Активні антидемократи	4,1	2,4	17,1	23,5
Демократи, амбівалентні щодо громадської активності	—	5,4	—	—
Демократи, які не визначилися щодо громадської активності	—	6,2	—	—
Політично невизначені та громадсько активні	—	2,1	—	—
Політично невизначені та громадсько пасивні	—	11,5	—	24,8

Як можна бачити, найбільш схожу наповненість груп демонструють способи поділу за ідеальною точкою і за концептуальною типологією. Цікаво і те, що в рамках латентного класового аналізу п'ята група репрезентована не тими, хто перебуває «посередині», а невизначеними політично і громадсько пасивними респондентами. Цей результат інтуїтивно зрозумілий у контексті того, що на підставі концептуальної типології 60,9% усіх респондентів можна характеризувати як громадсько пасивних, а до групи політично невизначених входить 19,5% вибірки (див. Додаток 5).

По-друге, порівняємо різні способи кластеризації з точки зору загальної «віддаленості» активних демократів і пасивних антидемократів, а також пасивних демократів і активних антидемократів (табл. 12.4).

Таблиця 12.4

Відмінності між полярними групами в різних варіантах кластеризації

Спосіб кластеризації	Активні демократи	Пасивні антидемократи	Різниця
<i>Демократія/антидемократія</i>			
Ідеальна точка поділу	24,9	14,0	10,9
Концептуальна типологія	25,4	12,7	12,7
Кластерний аналіз	26,4	16,5	9,9
Латентний класовий аналіз	27,0	17,2	9,8
<i>Активність/пасивність</i>			
Ідеальна точка поділу	21,1	11,4	9,7
Концептуальна типологія	22,1	11,5	10,6
Кластерний аналіз	21,3	10,7	10,6
Латентний класовий аналіз	20,8	13,8	7,0
Спосіб кластеризації	Пасивні демократи	Активні антидемократи	Різниця
<i>Демократія/антидемократія</i>			
Ідеальна точка поділу	24,1	13,7	10,4
Концептуальна типологія	25,1	13,3	11,8
Кластерний аналіз	26,9	16,6	10,3
Латентний класовий аналіз	28,0	20,1	7,9
<i>Активність/пасивність</i>			
Ідеальна точка поділу	12,2	22,0	9,8
Концептуальна типологія	11,9	22,4	10,5
Кластерний аналіз	10,2	20,0	9,8
Латентний класовий аналіз	9,6	18,5	8,9

У трьох випадках із чотирьох суттєві відмінності демонструє концептуальна типологія, в одному — концептуальна типологія і кластерний аналіз. Найменші відмінності отримано за допомоги латентного кластерного аналізу. Разом із тим концептуальна типологія та ідеальна точка поділу демонструють назагал схожі результати.

§ 12.4. Конструктна перевірка різних способів кластеризації

Ще один важливий спосіб порівняння різних варіантів кластеризації полягає в зіставленні різних груп респондентів за тими змінними, що виявляють взаємозв'язок з установками щодо громадської активності та демократії.

Стосовно громадської активності до таких змінних ми віднесли кількість громадських чи політичних організацій, у яких перебуває респондент, а також аномію (шкала від 0 до 18, де вищі значення свідчать про вищу аномію). Вибір другої змінної пов'язаний із тим, що індивідові, який перебуває у стані аномії, важко, якщо взагалі можливо, знаходити конструктивні шляхи суспільних перетворень і брати участь у них (табл. 12.5).

Таблиця 12.5

Конструктна перевірка шкали для вимірювання громадської активності

Спосіб кластеризації	Орієнтація на громадську активність	Орієнтація на громадську пасивність	Результати порівняння
<i>Участь у громадських організаціях</i>			
Ідеальна точка поділу	0,35	0,17	$t = 5,4; df = 774,8; p < 0,01$
Концептуальна типологія	0,39	0,17	$t = 5,4; df = 499,9; p < 0,01$
Кластерний аналіз	0,34	0,17	$t = 4,8; df = 1137; p < 0,01$
Латентний класовий аналіз	0,30	0,18	$t = 4,1; df = 1283,3; p < 0,01$
<i>Аномія</i>			
Ідеальна точка поділу	10,7	13,7	$t = -12,3; df = 712,8; p < 0,01$
Концептуальна типологія	10,5	13,7	$t = -11,5; df = 488,4; p < 0,01$
Кластерний аналіз	11,0	14,0	$t = -12,5; df = 964,6; p < 0,01$
Латентний класовий аналіз	11,2	13,7	$t = -11,7; df = 1209,8; p < 0,01$

У всіх випадках як для участі у громадських організаціях, так і для аномії зафіксовано статистично значимі відмінності між респондентами з різними установками щодо громадської активності. Найбільші відмінності характерні

за використання концептуальної типології, найменші — для латентного класового аналізу.

Для конструктивної перевірки шкали, що вимірює ставлення до демократії, було використано змінні, що стосуються: а) актуальності для нашої країни сильних лідерів на противагу законодавству і дискусіям (шкала від 1 до 3, де максимальне значення свідчить про повну підтримку необхідності сильних лідерів); б) ставлення до багатопартійної системи (шкала від 1 до 3, де максимальне значення свідчить про повну підтримку багатопартійності); в) оцінки державного управління економікою (шкала від 1 до 3, де 1 свідчить про повну підтримку планової економіки, а 3 — ринкової економіки без жодного державного втручання) (табл. 12.6).

Таблиця 12.6

**Конструктивна перевірка шкали для вимірювання ставлення до демократії,
середні значення**

Спосіб кластеризації	Демократичні орієнтації	Антидемократичні орієнтації	Результати порівняння
<i>Ставлення до необхідності сильного лідера</i>			
Ідеальна точка поділу	2,40	2,39	$t = 0,2; df = 268,4;$ $p = 0,84$
Концептуальна типологія	2,38	2,30	$t = 1,2; df = 186,6;$ $p = 0,25$
Кластерний аналіз	2,33	2,38	$t = -1,0; df = 1055,7;$ $p = 0,30$
Латентний класовий аналіз	2,31	2,36	$t = -1,1; df = 1018,5;$ $p = 0,29$
<i>Ставлення до багатопартійності</i>			
Ідеальна точка поділу	1,84	1,75	$t = 1,6; df = 280,0;$ $p = 0,12$
Концептуальна типологія	1,85	1,74	$t = 1,6; df = 192,2;$ $p = 0,12$
Кластерний аналіз	1,85	1,77	$t = 1,7; df = 1116,1;$ $p = 0,08$
Латентний класовий аналіз	1,84	1,86	$t = -0,4; df = 974,8;$ $p = 0,66$
<i>Ставлення до управління економікою</i>			
Ідеальна точка поділу	1,85	1,77	$t = 1,4; df = 227,1;$ $p = 0,16$
Концептуальна типологія	1,86	1,82	$t = 0,8; df = 162,2;$ $p = 0,42$
Кластерний аналіз	1,88	1,77	$t = 2,9; df = 868,4;$ $p < 0,01$
Латентний класовий аналіз	1,89	1,83	$t = 1,6; df = 929,2;$ $p = 0,10$

У разі прийняття припустимої похибки на рівні 5% статистично значимі результати зафіксовано тільки для одного порівняння — ставлення до управління економікою за виокремлення груп респондентів шляхом кластерного аналізу. Слід також зазначити, що за першого розгляду кластерний аналіз дає найкращі результати конструктивної перевірки для шкали, що вимірює ставлення до демократії.

Понад те, за використання концептуальної типології для порівняння установок щодо необхідності сильних лідерів у країні виникає парадоксальна ситуація, за якої демократично налаштовані респонденти більшою мірою схильні до такої думки, ніж респонденти з протилежного ідеологічного табору. Це могло би свідчити про помилковість цього способу кластеризації, але докладніший аналіз показує інше. Спершу розглянемо категоризовану діаграму розсіювання для змінних «громадська активність» і «ставлення до демократії» (рис. 12.2), побудовану на базі концептуальної типології, а також взаємозв'язок між відповідними групами і ставленням до необхідності сильних лідерів (табл. 12.7).

Таблиця 12.7

**Конструктивна перевірка шкали для вимірювання ставлення до демократії
(концептуальна типологія), %**

Установка щодо демократії	«Кілька сильних керівників можуть зробити для нашої країни більше, ніж усі закони і дискусії»		
	Не згоден	Не знаю	Згоден
Демократична	25,1	11,5	63,4
Амбівалентна	23,1	15,4	61,5
Невизначена	20,7	19,0	60,2
Антидемократична	28,9	12,5	58,6

На діаграмі видно, що респонденти з демократичними установками суттєво відрізняються від тих, хто дотримується протилежних поглядів. Звідси альтернативним поясненням таких несподіваних відмінностей у питанні підтримки сильних лідерів може бути те, що Є. Головаха і Н. Панина назвали феноменом подвійної інституціоналізації, коли, зокрема, вирішальні інститути радянського суспільства, втративши свою легальність, не втратили легітимності (Панина, 2008, с. 76–84). Демократична установка є певним спільним тлом життєдіяльності для основної частини нашого суспільства (відповідно до концептуальної типології, близько 70%), що зумовлено як офіційною орієнтацією нашої держави на демократичний розвиток, так і реальною практикою протистояння громадян авторитаризму у 2004 і 2013–2014 роках. Ця модальна група включає і тих, чий демократизм має суто декларативний характер і хто в інструментальному сенсі готовий запозичувати засоби розв'язання суспільних проблем із радянського «спадку».

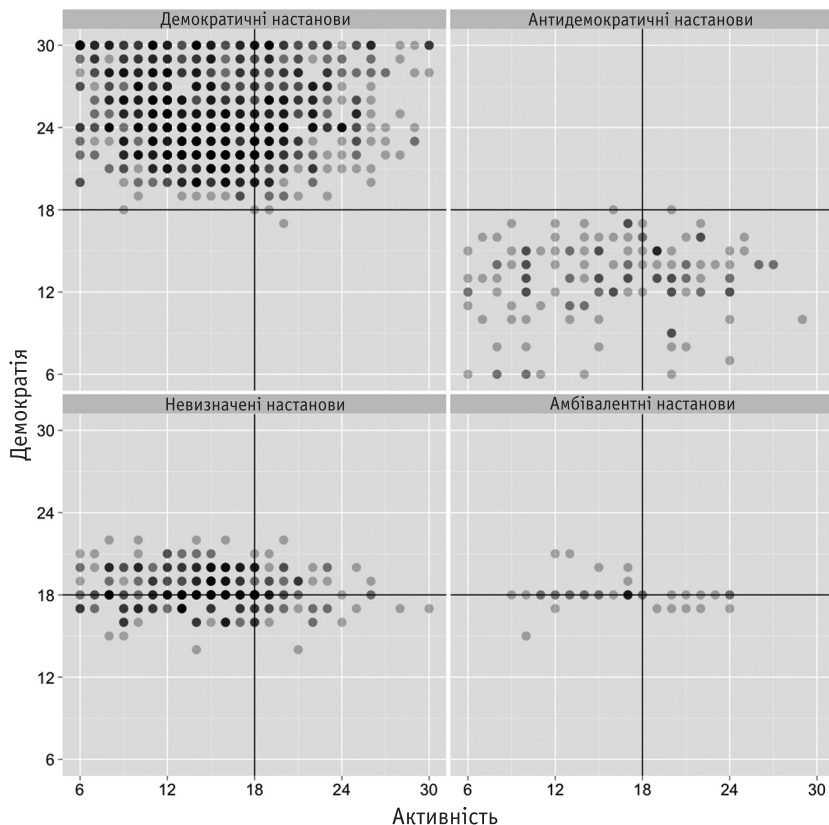


Рисунок 12.2. Категоризована діаграма розсіювання для змінних «ставлення до громадської активності» та «ставлення до демократії»

Таблиця 12.8

Конструктна перевірка шкали для вимірювання ставлення до демократії (кластерний аналіз), %

Установка щодо демократії	«Кілька сильних керівників можуть зробити для нашої країни більше, ніж усі закони і дискусії»		
	Не згоден	Не знаю	Згоден
Демократична	27,9	11,0	61,1
Невизначена	21,2	14,0	64,8
Антидемократична	22,9	15,9	61,2

Тепер розгляньмо аналогічні дані для типології на підставі кластерного аналізу (табл. 12.8, рис. 12.3). Незважаючи на те, що респонденти з демократичними установками меншою мірою схильні підтримувати необхідність сильних лідерів, ніж їхні ідейні опоненти, найбільшу потребу у відповідній категорії ке-

рівників висловлюють респонденти з невизначеними установками щодо демократії.

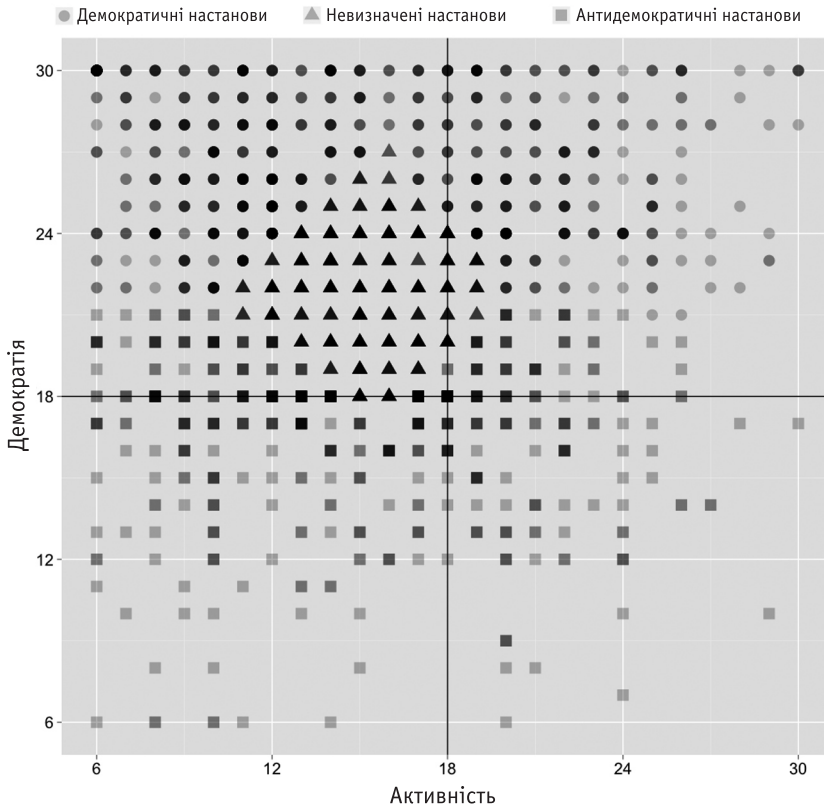


Рисунок 12.3. Діаграма розсіювання за змінними «ставлення до громадської активності» та «ставлення до демократії»

Разом із тим, відповідно до графіка, останню групу можна назвати «невизначеними» лише щодо двох інших. Якщо розглянути поділ відповідей на пункти СТ ТПК, що стосуються демократичних установок цієї групи респондентів, то модальною відповіддю (від 56,5% до 72,5%) буде згода (табл. 12.9). Відповідно, за чотиригрупового рішення кластерного аналізу, в якому група респондентів, що не визначилися, зникає, відмінності між «демократами» й «антидемократами» також перестають фіксуватися ($X_{дем} = 2,39$; $X_{антидем} = 2,36$; $t = 0,6$; $df = 1471,5$; $p = 0,56$).

Отже, позірна на перший погляд невідповідність є насправді показником суперечливості нашого суспільства, в якому багато індивідів, котрі декларують демократичні установки, більшою мірою схильні до «сильної руки» в управлінні державою, ніж їхні ідейні опоненти. Нехай результати порівняння і не є статистично значимими (передусім через слабку наповненість однієї з груп зіставлення), але назагал вони показові.

Таблиця 12.9

Поділ відповідей на пункти СТ ТПК, що стосуються демократичних установок респондентів, котрі не визначилися (кластерний аналіз), %

<i>Твердження щодо демократії</i>	<i>Цілком або скоріше згоден</i>	<i>Важко сказати</i>	<i>Зовсім або скоріше не згоден</i>
Демократія — найкращий політичний устрій для будь-якої сучасної держави	72,5	21,6	5,9
Без демократії неможливий сталий економічний розвиток країни	56,5	36,6	7,2
Я впевнений, що тільки в демократичному суспільстві по-справжньому поважають права людини	68,1	27,7	4,3
Демократія забезпечує людині найбільші можливості для індивідуального політичного вибору порівняно з іншими режимами	61,7	32,5	5,9
Тільки за демократії законодавча влада відображає інтереси всіх основних соціальних груп	60,0	34,1	5,9
Демократія дає людям можливість об'єднуватися в партії та спілки	64,4	31,7	3,9

§ 12.5. Оцінювання змін у різні способи кластеризації

Також важливо порівняти різні способи кластеризації під кутом зору фіксації змін, що відбулися в політичній культурі українського суспільства (табл. 12.10). При цьому з аналізу виключено кластеризацію за допомоги латентного класового аналізу, позаяк вона призводить до формування в різні роки груп, які важко порівнювати, — 2006 року групи виглядають значно гомогеннішими, ніж 2016-го (див. Додаток 6). Окрім того, для кластерного аналізу проаналізовано два рішення — для п'яти і чотирьох груп.

Динаміка змін, зафіксована через ідеальну точку поділу і концептуальну типологію, вказує на схожі патерни. Якщо говорити про п'ять основних груп (активні та пасивні демократи, невизначені, активні та пасивні антидемократи), то відбулося значне збільшення групи пасивних демократів, незначне збільшення груп активних демократів і активних антидемократів, незначне зменшення групи пасивних антидемократів, а також значне зменшення групи респондентів, які не визначилися. Загалом такі зміни виглядають природними і піддаються несуперечливій інтерпретації. Так, збільшення груп активних і пасивних демократів могло відбутися переважно за рахунок респондентів із невизначеними установками, а невелике збільшення групи активних антидемократів — першою чергою за рахунок невеликого зменшення групи пасивних антидемократів. Розгляд концептуальної типології з урахуванням усіх її груп дає додаткову цінну інформацію про динаміку змін політичної культури в Україні.

Таблиця 12.10

Динаміка представленості різних типів політичної культури у 2006 і 2016 роках

Група	2006	Зміна	2016
<i>Ідеальна точка поділу</i>			
Активні демократи	20,6	+ 1,8	22,4
Пасивні демократи	48,1	+ 8,0	56,1
Невизначені	19,0	-9,0	10,0
Пасивні антидемократи	9,0	-1,6	7,4
Активні антидемократи	3,4	+ 0,5	4,1
<i>Концептуальна типологія</i>			
Активні демократи	12,2	+ 2,5	14,7
Пасивні демократи	30,0	+ 13,7	43,7
Невизначені в обох смислах	13,9	-8,7	5,2
Пасивні антидемократи	5,4	-0,9	4,5
Активні антидемократи	1,9	+ 0,5	2,4
Демократи, амбівалентні щодо громадської активності	4,8	+ 0,6	5,4
Демократи, які не визначилися щодо громадської активності	6,3	- 0,1	6,2
Політично невизначені та громадсько активні	3,6	- 1,5	2,1
Політично невизначені та громадсько пасивні	16,2	- 4,7	11,5
<i>Кластерний аналіз (5 груп)</i>			
Активні демократи	18,7	- 1,0	17,7
Пасивні демократи	19,8	+ 5,1	24,9
Невизначені	33,9	- 3,4	30,5
Пасивні антидемократи	17,2	- 3,4	13,8
Активні антидемократи	10,4	+ 2,7	13,1
<i>Кластерний аналіз (4 групи)</i>			
Активні демократи	18,2	+ 8,7	26,9
Пасивні демократи	26,4	+ 8,4	34,8
Пасивні антидемократи	21,2	- 1,1	20,1
Активні антидемократи	34,2	- 15,9	18,3

Якщо ж говорити про кластеризацію з застосуванням двох варіантів кластерного аналізу, то результати не можна назвати узгодженими. За рішення для п'яти груп усі зміни є незначними. П'ятивідсотковий приріст зафіксовано для групи пасивних демократів. Його можна пояснити за рахунок зменшення кількості активних демократів і респондентів, які не визначилися. Невеликий

приріст активних антидемократів можна пов'язати зі зменшенням кількості пасивних антидемократів. У разі чотиригрупової кластеризації динаміка змін міняє не тільки свій характер, а і масштаби. Так, значний приріст спостерігаємо як для групи активних демократів, так і пасивних. При цьому такі зміни відбуваються переважно за рахунок зменшення групи активних антидемократів. У цьому разі інтерпретацію ускладнюють два аспекти — модальність для українського суспільства активної антидемократії 2006 року (понад третина всіх респондентів) і ті причини, з яких активні антидемократи перейшли у протилежний ідеологічний табір.

Звісно, якщо виходити зі змістових характеристик центроїдів різних кластерів, то антидемократів можна назвати такими лише умовно. І з погляду аналізу динаміки змін це початкове рішення авторів тесту (найменування груп безвідносно до їх змістових характеристик) не виглядає переконливим.

§ 12.6. Висновки

Використання ідеальної точки поділу і концептуальної типології приводить до отримання такої структури типів політичної культури, які можна схарактеризувати як демократичну і громадсько пасивну. Така картина доволі добре узгоджується з нинішньою українською дійсністю. Своєю чергою, як кластерний, так і латентний класовий аналіз призводять до структури, для якої є характерною присутність надмірної кількості антидемократичних елементів.

За поділу полярних груп за осями «демократія» і «громадська активність» найкращу ефективність показує кластеризація за концептуальною типологією, найгіршу — кластеризація з застосуванням латентного класового аналізу.

Результати конструктної перевірки різних способів кластеризації у площині демократичних установок демонструють цікаві з наукового погляду результати, що їх можна витлумачувати як проти використання концептуальної типології, так і на її користь (якщо спиратися на припущення про подвійну інституціалізацію українського суспільства).

Щодо громадської активності, то найкращі результати перевірки пов'язані з використанням концептуальної типології. При цьому як ідеальна точка поділу, так і кластерний аналіз достатньо близькі до її показників.

За вивчення динаміки змін найлогічнішу для інтерпретації картину дають концептуальна типологія та ідеальна точка поділу. Кластерний аналіз у цьому аспекті виглядає суперечливим, а латентний класовий аналіз і поготів недоречним для розв'язання такого роду завдань.

Підсумовуючи проведений аналіз, вважаємо логічним і обґрунтованим використання концептуальної типології для детального аналізу типів політичної культури на базі результатів СТ ТПК, а також ідеальної точки поділу як її (концептуальної типології) простішого і менш точного аналога. Що стосується кластерного і латентного класового аналізу, то ці способи кластеризації виглядають за опрацювання СТ ТПК більшою мірою механічними процедурами, що не враховують смислових характеристик тесту і призводять до результатів, які погано вписуються в реалії нашого суспільства.

Глава 13

Валідизація шкали громадської активності

СТ ТПК є єдиним стандартизованим соціологічним інструментом, що був спеціально розроблений для вивчення демократичних установок, а також установок на громадську активність в українському суспільстві. У процесі розробки цієї методики основну увагу приділено її внутрішній структурі. Акцент при цьому зроблено на досягненні внутрішньої узгодженості індикаторів шкал демократії та громадської активності (Головаха, 2005, с. 24–38; Стегній, 2015), а також на особливостях обробки їх результатів за допомоги кластерного аналізу та його альтернатив (Дембіцький, 2017а).

Що стосується різних видів валідизації, то на етапі розробки тесту відповідний аналіз був надто лаконічним і фрагментарним (Головаха, 2005, с. 33–34). З огляду на це головним завданням глави є опис результатів факторної, критеріальної та конструктивної валідизації однієї зі шкал цього тесту, а саме шкали громадської активності. Додатково перевірено інтерпретаційний потенціал результатів відповідей за окремими індикаторами шкали. Нарешті, шкалу громадської активності розглянуто з позицій соціологічного тестування, три рівні якого передбачають стислість, спрямованість на вивчення соціальних феноменів, а також діагностичний потенціал за аналізу соціальних спільнот різного масштабу.

Усі розрахунки здійснено на підставі масиву соціологічного моніторингу «Українське суспільство» за 2016 рік (Інститут соціології НАН України).

§ 13.1. Інтерпретаційний потенціал результатів відповідей за окремими індикаторами шкали

За розробки шкали громадської активності її індикатори (табл. 13.1) відібрано на підставі математичних критеріїв (в основу відповідного рішення покладено результати використання коефіцієнта Альфа Кронбаха). Тому важливо розуміти, наскільки результати, отримані за їх допомоги, піддаються осмисленій соціологічній інтерпретації в сьогоdnішніх умовах політичного життя України.

На підставі індексу якісної варіації (IQV) індикатори шкали можна розбити на дві групи, які мають різний ступінь гетерогенності. До групи з більшою гетерогенністю відповідей належать перший, третій і четвертий індикатори, до групи з меншою — другий, п'ятий і шостий. Перша група показує менше переважання установок на громадську пасивність над активністю, друга — більше

(рис. 13.1)¹⁰³. На цій підставі можна здійснити інтерпретацію отриманих результатів.

Таблиця 13.1

Індикатори шкали громадської активності

№	Формулювання	IQV
1	Мені все одно, яка буде влада, якщо не стане гірше	0,99
2	Я не сподіваюся на вибори, бо не вірю, що від їх результатів зміниться моє життя	0,93
3	Я свій вибір давно зробив, тому не хочу брати участь у сьогоdnішньому політичному житті	0,97
4	Немає сенсу боротися за свої права, якщо влада їх відверто ігнорує	0,98
5	Шляхом голосування ми обираємо владу, ну а далі від нас уже нічого не залежить	0,94
6	Будь-яка спроба що-небудь змінити в політичному житті країни вимагає від людини надто великих жертв, які часто виявляються марними	0,91

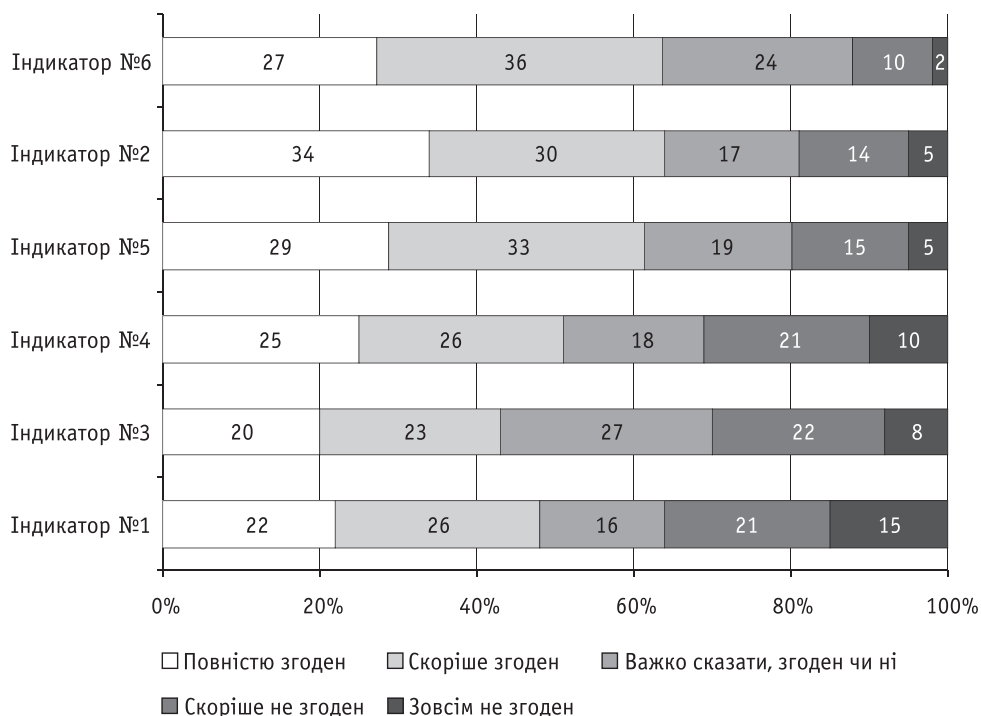


Рисунок 13.1. Розподіл відповідей за індикаторами шкали громадської активності

¹⁰³ На цьому рисунку послідовність індикаторів упорядковано відповідно до зменшення різниці між відсотком відповідей «повністю згоден» і «швидше згоден», з одного боку, і «скоріше не згоден» і «зовсім не згоден» — з іншого.

Чи логічно те, що респонденти 2016 року були більше згодні з тим, що надзвичайно складно змінити політичне життя країни, неможливо змінити своє життя шляхом участі у виборах і дій у невиборний період — через відсутність інструментів впливу на владу? Вони також менше згодні з тим, що немає сенсу боротися за свої права, якщо влада їх явно ігнорує, і не хочуть брати участі у сьогоdnішньому політичному житті, їм байдуже, яка буде влада.

Ми вважаємо, що це цілком уписується в сучасну українську дійсність. Вибори мало що змінюють у політичному житті країни, а в період між ними складно знайти легальні і по-справжньому дієві інструменти впливу на владу. Тому вся політична система України виглядає інертною і контрпродуктивною. При цьому населення має досвід протидії владі, яка порушувала його права, що передбачає як участь у політичному житті країни, так і небайдужість до особливостей українських можновладців.

Отже, відповіді на рівні окремих індикаторів піддаються осмисленій інтерпретації, якщо враховувати сьогоdnішні реалії політичного життя в Україні.

§ 13.2. Факторна валідність і внутрішня узгодженість

Таблиця 13.2

Результати оцінки надійності та факторної валідності шкали громадської активності ($N = 1786$)

Показник	Оригінальний варіант	Після перекодування
Альфа Кронбаха ($> 0,70$)	0,765	0,731
Хі-квадрат (DWLS)	20,124	11,202
df (ступені свободи)	9	9
DWLS / df	2,2	1,3
p -value	0,017	0,262
RMSEA ($< 0,07$)	0,026	0,012
CFI ($> 0,95$)	0,999	0,999
TLI ($> 0,95$)	0,998	0,999

Факторну валідність і внутрішню узгодженість перевірено за допомоги конфірматорного факторного аналізу і коефіцієнта Альфа Кронбаха. Результати отримано для шкали з використанням оригінального віяла відповідей (5 пунктів), а також для його модифікації після перегрупування у тризначний варіант: 1) «цілком згоден» і «скоріше згоден»; 2) «складно сказати»; 3) «скоріше не згоден» і «цілком не згоден». Обидва варіанти показують прийнятні результати за всіма індексами. Показники DWLS і RMSEA є дещо кращими для результатів після перекодування. Що стосується Альфа Кронбаха, то результати цього показника також перебувають на прийнятному рівні для обох варіантів (табл. 13.2). На підставі цих результатів можна зробити висновок, що факторна

структура тесту є виразнішою після перекодування. Відповідно, за опитування краще використовувати оригінальну шкалу (для більшого комфорту респондентів), а аналіз здійснювати на базі перекодованих результатів. Цей факт свідчить і на користь застосування концептуальної типології на підставі модальностей згоди, невизначеності та незгоди за категоризації респондентів на основі результатів СТ ТПК.

Відповідно, далі в рамках критеріальної та конструктивної валідації буде використано результати концептуальної типології, а якщо точніше, — дві основні категорії респондентів, виокремлені за її допомоги — з установками на громадську активність і з установками на громадську пасивність.

§ 13.3. Критеріальна валідність

З метою перевірки критеріальної валідності використано запитання анкети, які очікувано пов'язані з громадською активністю і які стосуються різних аспектів поведінки або статусу респондента, тобто демонструють його об'єктивні характеристики. Виходячи з можливих відповідей на ці запитання, респондентів поділено на відповідні групи. Далі кількість респондентів з установками на громадську активність або пасивність у кожній групі порівняно з еталонними значеннями — як таке використано дві пропорції в цілому по вибірці: 1) респондентів з установками на громадську активність; 2) респондентів з установками на громадську пасивність. Для такого порівняння використано перевірку статистичних гіпотез для пропорцій однієї вибірки (нульова гіпотеза стверджує, що значення у вибірці не відрізняється від значення в генеральній сукупності).

Вибір такого статистичного критерію зроблено тому, що кореляційний аналіз призводить в цьому разі до вельми невиразних результатів, що погано піддаються змістовній інтерпретації.

Пропорція респондентів з установками на громадську активність загалом у вибірці дорівнює 0,197, з установками на громадську пасивність — 0,610.

До соціальних груп, що демонструють більшу кількість респондентів з установками на громадську активність і меншу кількість респондентів з установками на громадську пасивність, належать (табл. 13.3):

- окремі категорії користувачів мережі інтернет;
- індивіди з досвідом керівної роботи;
- респонденти, які приділяють вільний час громадській діяльності;
- ті, що зверталися до місцевої державної адміністрації й отримали допомогу;
- ті, що ніколи не перебували в сімейних відносинах;
- індивіди з неповною і повною вищою освітою;
- респонденти, які оцінюють свій матеріальний стан як прийнятний або добрий;
- ті, що були членами однієї або двох громадських або політичних організацій.

Таблиця 13.3

**Групи, що демонструють більшу громадську активність
і меншу громадську пасивність**

<i>Індикатор</i>	<i>Варіант відповіді</i>	<i>КАР, %</i>	<i>КПР, %</i>
З якою метою Ви використовуєте інтернет?	Відвіую сайти державних, громадських структур і т. п.	39,2***	40,8***
	Шукаю однодумців, партнерів для реалізації нових ідей, проектів, ініціатив	27,1	44,8**
Чи доводиться (доводилося) Вам на Вашій основній роботі керувати співробітниками і відповідати за їхню роботу?	Так	24,8**	—
Нижче наведено перелік занять людей у вільний від основної та домашньої роботи час ... за останні 7 днів (разом із вихідними).	Громадська робота	37,1***	33,9***
Чи зверталися Ви за останні 12 місяців у зазначені установи ¹⁰⁴ або організації для розв'язання Ваших особистих проблем і якою мірою Ви задоволені їх рішеннями?	Зверталися й отримали допомогу	32,1***	54,9
Ваш сімейний стан	Ніколи не перебував(ла) у шлюбі	24,1	52,7**
Ваша освіта	Перший ступінь вищої освіти (бакалавр); повна вища освіта (спеціаліст, магістр, аспірантура, вечний ступінь)	26,2***	54,7***
Як би Ви оцінили матеріальний стан Вашої сім'ї?	Вистачає загалом на проживання; вистачає на все необхідне, але нам не до заощаджень	24,0**	55,9**
	Вистачає на все необхідне, робимо заощадження; живемо в повному достатку	31,9*	48,9
Кількість громадських або політичних організацій, у яких Ви маєте членство	Одна	27,6**	50,2***
	Дві	64,1***	23,1***

Скорочення тут і далі: КАР — кількість респондентів з установками на громадську активність; КПР — кількість респондентів з установками на громадську пасивність.

Позначення в таблиці: * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$. Для решти випадків $p < 0,1$. Позначка «—» говорить, що відповідне значення дорівнює еталонному.

¹⁰⁴ Мається на увазі місцева державна адміністрація.

Виходячи з цих даних можна припустити, що факторами, які тією чи тією мірою позитивно впливають на установки, пов'язані з громадською активністю, є добре матеріальне становище, висока освіта, «необтяженість» сімейними відносинами, а також досвід роботи на керівних посадах і успішної взаємодії з державними структурами. Своєю чергою, до очікуваних наслідків прийняття відповідних установок можна віднести інтерес до державних організацій, а також участь у громадській роботі та діяльності громадських і політичних організацій.

Якщо ж говорити про категорії, які демонструють меншу кількість респондентів з установками на громадську активність і більшу кількість респондентів з установками на громадську пасивність, то до них належать (табл. 13.4):

- індивіди з досвідом проживання в зоні бойових дій (без участі в конфлікті);
- респонденти без оплачуваної роботи;
- ті, хто втратив дружину або чоловіка;
- індивіди з середньою спеціальною освітою і нижче;
- респонденти, які оцінюють матеріальне забезпечення своєї родини як низьке;
- ті, хто не є членами громадських або політичних організацій.

Таблиця 13.4

**Групи, що демонструють меншу громадську активність
і більшу громадську пасивність**

<i>Індикатор</i>	<i>Варіант відповіді</i>	<i>КАР, %</i>	<i>КПР, %</i>
Чи проживали Ви в зоні бойових дій у будь-якій точці України?	Так, проживав і не брав участі в конфлікті	6,8**	85,5***
Чи є Ви зараз вимушеним безробітним і перебуваєте в пошуку оплачуваної роботи?	Так	14,4	67,2
Ваш сімейний стан	Вдівець (вдова)	13,4*	69,5***
Ваша освіта	Початкова, неповна середня	13,6**	—
	Середня повна	16,5	65,7***
Як би Ви оцінили матеріальний стан Вашої сім'ї?	Часто не маємо грошей та їжі, іноді жебракуємо	5,0	—
	Бракує продуктів харчування, іноді голодуємо; вистачає тільки на продукти харчування	15,8**	66,1**
Кількість громадських або політичних організацій, у яких маєте членство	Жодної	17,0*	64,0*

Позначення в таблиці: * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$. Для решти випадків $p < 0,1$. Позначка «—» говорить, що відповідне значення дорівнює еталонному.

Тут можна припустити, що факторами, які негативно пов'язані з установками на громадську активність і позитивно — з установками на громадську пасивність, є різні життєві труднощі, а також низький рівень освіти. Головним наслідком таких установок є зменшення залученості в діяльність громадських організацій.

Загалом наведені результати демонструють взаємозв'язок факторів, що зумовлюють громадську активність/пасивність або мають пряму залежність від них, з результатами використання відповідної шкали й, отже, свідчать на користь її критеріальної валідності.

§ 13.4. Конструктна валідність

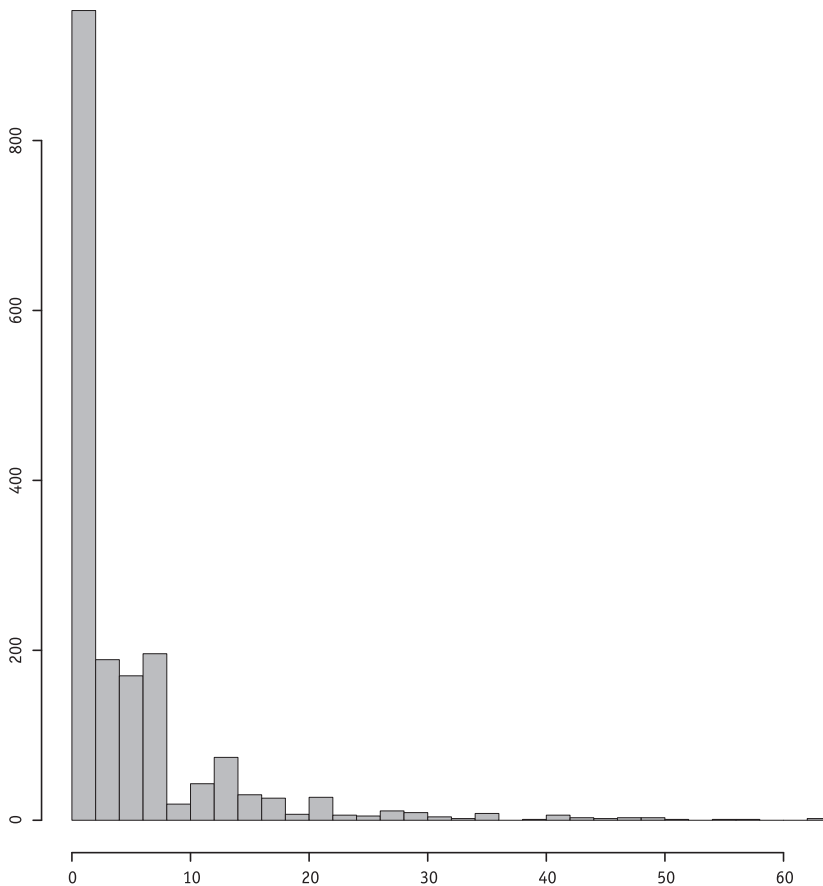


Рисунок 13.2. Розподіл значень індексу соціальної напруги

Для перевірки конструктної валідності використано такий самий підхід до аналізу даних, як і під час критеріальної валідації. При цьому як контрольні змінні відібрано запитання щодо установок, тобто суб'єктивні характеристики респондентів, пов'язані з їхнім ставленням до громадської активності. Крім

простих анкетних запитань також використано результати вимірювальних шкал аномії, цинізму, а також індексів дестабілізаційного потенціалу, соціального самопочуття (ПСС). Оскільки для цих вимірювальних інструментів немає чітких критеріїв виокремлення рівнів виразності відповідних властивостей¹⁰⁵, результати опитування щодо них поділено на три приблизно рівні інтервали, за винятком індексу дестабілізаційного потенціалу. В останньому випадку результати добре описані законом Ципфа (рис. 13.2). Тому в його перший інтервал включено тільки одне значення — 0, у другий потрапляє діапазон значень від 1 до 10, у третій — від 11 і вище (максимальне значення дорівнює 63).

До факторів, що продемонстрували зв'язок з установками на громадську активність/пасивність, належать такі (табл. 13.5, 13.6):

- бажання/небажання розпочати свою справу;
- готовність узяти участь у мітингах і демонстраціях протесту проти падіння рівня життя;
- інтернальність/екстернальність;
- аномія;
- дестабілізаційний потенціал;
- соціальний цинізм;
- соціальне самопочуття;
- оцінка власної пристосованості до нинішньої життєвої ситуації;
- ставлення до українського громадянства;
- готовність/неготовність терпіти матеріальні нестатки задля порядку і спокою у країні;
- оцінка можливості відстояти свої права та інтереси, порушені на рівні місцевої або державної влади;
- ступінь зацікавленості політикою.

Індивіда з установками на громадську активність з вищою ймовірністю можна зустріти серед тих, хто хоче розпочати свою справу, готовий узяти участь у мітингах і демонстраціях протесту, має внутрішній локус контролю, низький або середній рівень аномії та соціального цинізму, середній або високий рівень дестабілізаційного потенціалу та соціального самопочуття, активно включений у нове життя і ринкові відносини, пишається своїм українським громадянством, бажає активно протестувати проти постійного погіршення умов життя, вірить у можливість змінити несправедливі рішення уряду України або місцевої влади, дуже цікавиться політикою.

Отже, установка на громадську активність є характерною насамперед для осіб, орієнтованих на творчу діяльність, захист своїх законних інтересів і конструктивне ставлення до навколишньої соціальної дійсності. Якщо ж говорити про соціальні групи з мінімальним відсотком громадсько активних індивідів, то для них властива соціальна апатія, відсутність віри в можливість захисту своїх інтересів і неконструктивне ставлення до соціального оточення.

¹⁰⁵ Ця частина роботи підготовлена до розробки критеріальних рівнів для ПСС.

Таблиця 13.5

**Групи, що демонструють більшу громадську активність
і меншу громадську пасивність**

<i>Індикатор</i>	<i>Варіант відповіді</i>	<i>КАР, %</i>	<i>КПП, %</i>
Чи хотіли б Ви розпочати свою справу (підприємство, фермерське господарство тощо)?	Так	25,1**	—
Якщо мітинги ¹⁰⁶ і демонстрації протесту відбуватимуться, Ви особисто братимете в них участь чи ні?	Скоріше за все так	26,2***	56,7*
Як Ви вважаєте, від чого здебільшого залежить те, як складається Ваше життя?	Більшою мірою від мене, ніж від зовнішніх обставин	27,3**	52,2**
Шкала аномії	Низький рівень (0–6)	43,8***	33,7***
	Середній рівень (7–12)	24,8**	52,7**
Індекс дестабілізаційного потенціалу	Середній рівень (1–10)	25,2***	57,1*
	Високий рівень (11–63)	32,5***	49,6***
Шкала цинізму	Низький рівень (0–4)	25,8	44,0***
	Середній рівень (5–9)	26,1***	52,5***
Інтегральний індекс соціального самопочуття	Високий рівень (50–60)	26,0**	52,8**
Якою мірою Ви пристосувалися до нинішньої життєвої ситуації?	Активно включився б у нове життя, ринкові відносини уявляю як природний спосіб життєдіяльності	30,7***	49,9***
Якою мірою Ви пишаєтеся чи не пишаєтеся тим, що є громадянином України?	Дуже пишаюся	30,7***	47,5***
Що, на Вашу думку, краще — терпіти будь-які матеріальні труднощі задля збереження у країні порядку, миру і спокою чи в разі значного погіршення умов життя виходити на вулицю з протестом?	Слід активно протестувати проти постійного погіршення умов життя	23,1*	—
Якби уряд України ухвалив рішення, яке утискає Ваші законні права та інтереси, могли б Ви що-небудь зробити проти такого рішення?	Так, зміг би щось зробити	32,1***	44,5***
Якби аналогічне рішення ухвалила місцева влада, чи змогли б Ви щось зробити проти такого рішення?	Так, зміг би щось зробити	32,0***	50,8***
Якою мірою Вас цікавить політика?	Дуже цікавить	34,8***	50,6**

Позначення в таблиці: * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$. Для решти випадків $p < 0,1$. Позначка «—» говорить, що відповідне значення дорівнює еталонному.

¹⁰⁶ Ідеться про масові виступи населення проти падіння рівня життя в місці проживання респондента.

Таблиця 13.6

**Групи, що демонструють меншу громадську активність
і більшу громадську пасивність**

<i>Індикатор</i>	<i>Варіант відповіді</i>	<i>КАР, %</i>	<i>КПР, %</i>
Чи хотіли б Ви розпочати свою справу (підприємство, фермерське господарство тощо)?	Ні	14,0***	66,6**
Якщо мітинги ¹⁰⁷ і демонстрації протесту відбудуться, Ви особисто братимете в них участь чи ні?	Скоріше за все ні	15,8**	66,5**
Як Ви вважаєте, від чого здебільшого залежить те, як складається Ваше життя?	Здебільшого від зовнішніх обставин	12,8**	68,8**
Шкала аномії	Високий рівень (13–18)	12,3***	70,9***
Індекс дестабілізаційного потенціалу	Відсутній (0)	12,9***	66,5***
Шкала цинізму	Високий рівень (10–14)	15,7***	68,4***
Інтегральний індекс соціального самопочуття	Низький рівень (20–30)	12,2***	70,8***
Якою мірою Ви пристосувалися до нинішньої життєвої ситуації?	Не маю бажання пристосуватися до нинішньої ситуації, живу як доведеться, очікую змін на краще	12,2***	70,7***
Якою мірою Ви пишаєтеся чи не пишаєтеся тим, що є громадянином України?	Зовсім не пишаюся, скоріше не пишаюся	12,9**	71,1***
	Важко сказати	13,1***	67,4**
Якби уряд України ухвалив рішення, яке утискає Ваші законні права та інтереси, могли б Ви що-небудь зробити проти такого рішення?	Ні, нічого не зміг би зробити	15,6***	68,6***
Якби аналогічне рішення ухвалила місцева влада, чи змогли б Ви щось зробити проти такого рішення?	Ні, нічого не зміг би зробити	13,8***	69,8***
Якою мірою Вас цікавить політика?	Зовсім не цікавить	10,3***	72,7***

Позначення в таблиці: * $p < 0,05$; ** $p < 0,01$; *** $p < 0,001$. Для решти випадків $p < 0,1$. Позначка «—» говорить, що відповідне значення дорівнює еталонному.

Окреслена картина несуперечливо вписується в теоретичне розуміння феномену громадської активності й, отже, свідчить на користь конструктивної валідності відповідної шкали.

¹⁰⁷ Ідеться про масові виступи населення проти падіння рівня життя в місці проживання респондента.

§ 13.5. Шкала громадської активності як соціологічний тест

Незважаючи на те, що шкала громадської активності є складовим елементом соціологічного тесту «Типи політичної культури», вона може виступати і як незалежний вимірювальний інструмент, більше за те, як повноцінний соціологічний тест, тобто як такий вимірювальний інструмент, який можна розглядати як соціологічний у всіх трьох вимірах. Зупинимось на цьому докладніше.

По-перше, шкала громадської активності за всієї своєї стислості зберігає необхідні показники надійності (Головаха, 2005, с. 24–38) і валідності (останнє продемонстровано вище). Отже, її застосування повністю виправдане в масових опитуваннях.

По-друге, вона спрямована на вивчення соціального феномену. Так, громадську активність у тому чи тому суспільстві обумовлює його історія, а також актуальні соціальні процеси. Натомість на рівні окремого індивіда її виразність пов'язана з набором характеристик, серед яких важливе місце посідають його соціальний досвід, соціальна активність і соціальні статуси.

По-третє, шкала громадської активності надає цінну інформацію для соціальної діагностики на різних рівнях соціальної взаємодії. Оскільки соціальна діагностика є комплексною науковою діяльністю, для її реалізації необхідно залучення інформації з інших релевантних з точки зору соціологічного аналізу показників. Для ілюстрації цього моменту звернемося до запитання про політичні сили, які можуть претендувати на домінуючий статус у сучасній Україні. Щоб відповісти на нього, крім інформації про громадську активність респондентів необхідно залучити дані про їхні політичні погляди (табл. 13.7). Перш за все оцінимо, чи є репрезентовані результати несуперечливими в контексті сьогоденного політичного життя України.

Згідно з отриманими даними, національно-демократичні та націоналістичні сили є найбільш активними і включеними в політичне життя країни. За ними слідує ліберали, що цілком логічно в контексті нещодавньої Революції гідності, одним з ідеалів якої була особиста свобода громадян.

Для решти політичних груп є характерною стандартна репрезентованість політично активних респондентів. Більшість із них — це індивіди з лівими поглядами, які наразі не можна назвати популярними в рамках домінуючого політичного дискурсу.

Своєю чергою, респонденти без політичних уподобань, а також ті, хто не розуміється на відповідних питаннях, показують найменший громадський потенціал. І це цілком очікувано через відсутність інтересу до політики у перших і відсутність необхідних знань у других.

На наш погляд, така картина добре узгоджується з сучасним станом українського суспільства.

Тепер розглянемо ці дані як основу для соціологічного діагнозу розвитку нашого суспільства. Кількість правих політичних сил (християнські демократи, націонал-демократи і націоналісти) дещо менша, ніж лівих (соціал-демократи, соціалісти і комуністи) — 20,2% проти 22,7%. При цьому кількість громадсько активних респондентів є вищою серед перших — 7,3% усієї вибірки

проти 4,5% усієї вибірки. Останній момент створює переваги для прихильників правих політичних ідей — їхнім активістам легше знаходити однодумців у своєму середовищі й, отже, ефективно обстоювати свої інтереси. Найбільшою мірою це стосується націонал-демократів. Своєю чергою їхня потенційно активніша діяльність є доброю підставою для залучення нових послідовників із числа тих, хто ще остаточно не визначив своєї позиції.

Таблиця 13.7

Установка на громадську активність у групах із різними політичними поглядами, %

<i>Політичні погляди</i>	<i>Представленість у вибірці</i>	<i>Установка на громадську активність</i>	<i>СЗВ¹⁰⁸ від загального рівня в суспільстві (19,7)</i>
Національно-демократичні	11,1	45,5	Так
Націоналістичні	3,3	34,5	Так
Ліберальні	2,8	32,7	Так
Соціально-демократичні	9,4	23,8	Ні
Зелені	2,2	22,5	Ні
Соціалістичні	9,1	18,4	Ні
Ще остаточно не визначили своєї позиції	14,3	18,0	Ні
Комуністичні	4,2	14,9	Ні
Християнсько-демократичні	5,8	14,4	Ні
Відсутні	13,4	12,1	Так
Респонденти, що не розуміються на цих течіях	23,2	10,1	Так

Головне питання в тому, якій моделі національної держави віддадуть перевагу прихильники правих політичних ідей як найпопулярніший — громадянській чи етнічній. Обидві мають усі передумови, щоб бути реалізованими. Перша з огляду на відповідність популярним нині ідеалам євроінтеграції, друга з огляду на наявність титульної нації, яка перебуває у процесі протистояння зовнішній агресії. Зазначені політичні дискурси створюють ідеологічні передумови для відповідних ініціатив політичних еліт в Україні, які мають усі шанси бути підтриманими «знизу».

Щоб оцінити ймовірність реалізації кожної з них, подивимося, як респонденти оцінюють мотивацію до об'єднання в українському суспільстві, а також звернемо увагу на кількість респондентів, які демонструють установки на громадську активність, у кожній із відповідних груп (табл. 13.8).

¹⁰⁸ Тут і далі СЗВ означає статистично значимі відмінності.

Таблиця 13.8

Мотивація для об'єднання людей в українському суспільстві, %

Мотивація для об'єднання	Представленість у вибірці	Установка на громадську активність	СЗВ від загального рівня в суспільстві (19,7)
Національна належність	11,2	33,2	Так
Патріотичні почуття громадянина України	23,0	32,0	Так
Віра у краще майбутнє	29,3	24,7	Так
Прагнення спільно долати труднощі життя	21,3	23,6	Так
Релігія (віресповідання)	8,4	23,3	Ні
Родинні та товариські почуття	18,2	21,5	Ні
Спільні труднощі життя	26,9	21,1	Ні
Незадоволеність владою	28,7	20,1	Ні
Політичні погляди	8,1	20,0	Ні
Мова спілкування	6,7	17,5	Ні
Відчуття втрати нормального життя	24,4	16,7	Ні
Ніщо не об'єднує	8,6	15,7	Ні
Важко відповісти	7,6	13,9	Так
Страх перед майбутнім	19,4	13,3	Так

Щоб дістати більше інформації з цих даних, необхідно визначити, які з поданих мотивів респонденти обирають разом. Для цього було використано багатовимірне шкалювання на підставі евклідової відстані (пункти «ніщо не об'єднує» і «важко відповісти» було виключено з аналізу). Двовимірна діаграма на підставі отриманих результатів дозволяє виокремити кілька груп (рис. 13.3). Перша містить мову спілкування, релігію і політичні погляди, друга — страх перед майбутнім, відчуття втрати нормального життя, спільні проблеми і невдоволення владою, третя — прагнення спільно долати труднощі, віру у краще майбутнє і патріотичні почуття громадянина України. Мотиви родинних і товариських почуттів, а також національної належності посідають самотійні позиції.

Подана візуалізація дає можливість здійснювати інтерпретацію за низкою напрямків:

- смислова схожість мотивів, що входять в одну і ту саму групу, а також концептуальні дихотомії, що поділяють різні точки і/або групи у двовимірному просторі;
- кількісна наповненість різних груп;

– кількість респондентів з установками на громадську активність у кожній із груп.

Спочатку розглянемо кожну позицію окремо. Якщо говорити про концептуальні дихотомії, то на осі *X* перша група протистоїть другій. На наш погляд, перша представлена цінностями, друга — інтересами. Відповідно, вісь *X* можна назвати «Ідеалізм — Прагматизм». У разі осі *Y* першій і другій групам протистоїть третя. Ми вважаємо, що перші дві пов'язані з зовнішніми орієнтирами й обставинами, а третя — з внутрішніми прагненнями. Відтак цю вісь можна назвати «Реактивність — Активність».

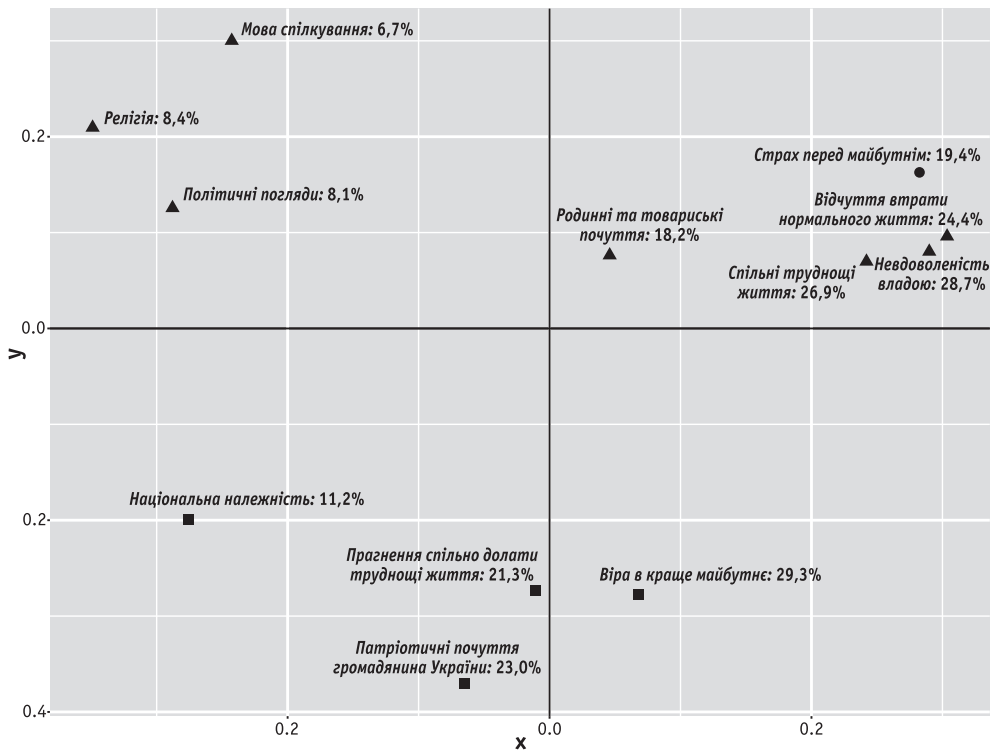


Рисунок 13.3. Результати двовимірного багатовимірного шкалювання на основі індикаторів мотивації об'єднання людей в українському суспільстві

У разі кількісної наповненості показники першої групи перебувають у діапазоні від 6,7 до 8,4%, другої — від 19,4 до 28,7%, третьої — від 21,3 до 29,3%. Це свідчить про домінантність другої та третьої груп мотивів респондентів.

Нарешті, установки на громадську активність помічено за допомогою позначок різної форми. Квадратики показують мотиви, обрані респондентами, серед яких підвищена представленість установок на громадську активність, трикутники — звичайну представленість, круглі позначки — знижену представленість.

Тепер повернемося до питання про ймовірність реалізації різних моделей національної держави в Україні з урахуванням нових даних. Модель етнічної

національної держави могла б опертися на мотивацію, представлену першою групою. У разі ж реалізації моделі громадянської національної держави найкращий «грунт» дає мотивація третьої групи. Як бачимо, і кількісна представленість, і ступінь громадської активності є вищим для мотивів третьої групи. Якщо звернутися до другої групи, то йтиметься не так про побудову чогось, як про розв'язання поточних проблем.

Імовірно, без розв'язання поточних труднощів говорити про побудову успішної національної держави того або того типу не доводиться. Однак характер осмислення проблем сьогоденної дійсності правлячою елітою, а також запропоновані нею шляхи їх розв'язання матимуть безпосередній вплив на ступінь суспільної підтримки політичних рішень серед найактивніших громадян України.

У цьому сенсі суспільний потенціал побудови громадянської національної держави виглядає вагомішим як з точки зору кількості її прихильників, так і з точки зору рівня активності індивідів, що підтримують ідеї громадянської ідентичності.

§ 13.6. Висновки

Репрезентовані в цій главі результати свідчать на користь факторної, критеріальної та конструктивної валідності шкали громадської активності. При цьому важливо розуміти, що в строгому сенсі критеріальна валідизація вимірювальної шкали передбачає спеціально проведені дослідження, що охоплюють такі специфічні групи, як, наприклад, громадські активісти або волонтери. Відповідно, такі дослідження можуть бути подальшим етапом методичного аналізу результатів використання обговорюваної шкали. Зокрема, це допоможе точніше визначити рівні виразності громадської активності, яку вимірюють за її допомогою.

Оцінюючи шкалу громадської активності з позицій соціологічного тестування, можна констатувати таке:

- шкала має оптимальний для масових опитувань розмір;
- вона спрямована на вивчення важливої соціальної характеристики;
- результати її використання мають високий діагностичний потенціал, що дає вихід на — формулювання прогностичних сценаріїв розвитку українського суспільства.

Отже, шкала громадської активності є соціологічним тестом у всіх можливих сенсах.

Глава 14

Конструювання, валіди́зація та інтерпретація соціологічного тесту «Типи політичної культури — II»

Результати використання СТ ТПК у двох хвилях соціологічного моніторингу «Українське суспільство» (2006 і 2016 років) дали підставу припустити деякі обмеження цього тесту. У зв'язку з цим було проведено дослідницьку роботу, спрямовану на створення альтернативного соціологічного тесту «Типи політичної культури — II».

Метою розглядуваної глави є презентація результатів цієї дослідницької роботи, здійсненої у три етапи: 1) постановка проблеми і відбір індикаторів; 2) валіди́зація СТ ТПК-II; 3) визначення способу кластеризації респондентів відповідно до різних типів політичної культури.

§ 14.1. Постановка проблеми і відбір індикаторів

Використання СТ ТПК продемонструвало доволі неоднозначні результати. Так, попри те, що в абсолютній більшості респондентів (70%) виявилися демократичні установки, згода з необхідністю сильного лідера у відповідній групі була виражена сильніше, ніж серед тих, хто показав антидемократичні орієнтації. Можна було припустити, що така ситуація зумовлена різним співвідношенням громадсько активних і громадсько пасивних респондентів у відповідних групах. Так, у групі респондентів із демократичними установками на одного громадсько активного індивіда припадає три пасивні, а у групі з антидемократичними установками це співвідношення становить один до двох. Але ознайомлення навіть із простими описовими показниками (табл. 14.1) спростовує це припущення.

Таблиця 14.1

Рівень згоди з необхідністю сильного лідера, середні значення

Група	Рівень підтримки	n
Громадсько активні антидемократи	2,2	42
Громадсько активні демократи	2,3	262
Громадсько пасивні антидемократи	2,3	81
Громадсько пасивні демократи	2,4	779

Як продемонстровано вище, такі парадоксальні результати можуть бути пов'язані з феноменом подвійної інституціоналізації, в межах якої визначальні інститути радянського суспільства, втративши свою легальність, не втратили легітимності. Проте в рамках цієї роботи буде перевірено іншу гіпотезу, що має не теоретичний, а методичний характер. Але перш ніж її озвучити, звернімося до визначень індикаторів, сформульованих для виявлення демократичних установок: 1) «демократія — найкращий політичний устрій для будь-якої сучасної держави»; 2) «без демократії неможливий сталий економічний розвиток країни»; 3) «я певен, що тільки в демократичному суспільстві посправжньому поважають права людини»; 4) «демократія забезпечує людині найбільші можливості для індивідуального політичного вибору порівняно з іншими режимами»; 5) «тільки за демократії законодавча влада відображає інтереси всіх основних соціальних груп»; 6) «демократія дає людям можливість об'єднуватися в партії та спілки для реального захисту своїх прав та інтересів».

І тут виникає питання: про яку демократію йдеться? Одна річ — говорити про ідеальну демократію, інша — про демократію розвинених країн. Що ж стосується пострадянських країн, то тут поняття «демократія» може бути суто номінальним. Тому стартова точка подальших пошуків — це припущення, що наведені шість тверджень більшість респондентів сприймають саме в контексті демократичних режимів, які можна назвати успішними. Своєю чергою, питання щодо необхідності сильного лідера у своїй країні, ясна річ, сприймають застосовно до національних реалій. І якщо вже вітчизняна демократія часто показує невтішні результати, надії на інституційні поліпшення переносять на образ певного сильного лідера.

З огляду на зазначені суперечності у відповідях звернемося до вихідного набору тверджень, на базі яких сконструйовано СТ ТПК, аби визначити альтернативний список фінальних індикаторів, прийнятніших для оцінювання демократичних установок в Україні. Додатковим аргументом на користь цього стало й те, що в рамках конструювання СТ ТПК як головний орієнтир відбору тверджень використовували їх внутрішню узгодженість. Зрештою, незважаючи на наявність індикаторів, що фіксують як сприйняття, так і несприйняття демократії, до СТ ТПК було включено твердження лише першої групи.

Оскільки дійсність української демократії є далекою від ситуації в розвинених країнах і тим більше від деякої ідеальної демократії, було проведено роботу з тією частиною тверджень, що спрямовані на оцінку несприйняття демократії (антидемократичні установки)¹⁰⁹. Перше, що привертає увагу, — смислова гетерогенність відповідних тверджень. Тому з метою їх категоризації було здійснено експертне опитування¹¹⁰. Як дослідницьку стратегію використано

¹⁰⁹ Цікаво, що всі фінальні твердження СТ ТПК, що оцінюють установки на громадську активність, сфокусовані на відсутності, а не присутності громадської активності.

¹¹⁰ Експертами виступили наукові співробітники відділу методології та методів соціології Інституту соціології НАН України, частина з яких раніше була залучена до розробки СТ ТПК.

стратегію узгодження концептів¹¹¹, результати застосування якої вможливили виокремлення чотирьох груп тверджень (табл. 14.2).

Таблиця 14.2

Категорії індикаторів щодо несприйняття демократії

<i>Колективістські установки</i>	
1	Соціалістична система є справедливішою, ніж капіталізм, оскільки за соціалізму суспільство не поділене на багатих і бідних
2	Надмірні статки треба вилучати і розподіляти серед малозабезпечених
3	У суспільстві не має бути багатих і бідних
4	Справжня рівність може бути забезпечена тільки в такому суспільстві, де колективні інтереси переважають над особистими
5	Не можна допускати, щоб людина власні інтереси ставила вище за державні
<i>Антидемократичні установки</i>	
1	Демократія з усіма її принципами не має сенсу, коли більшість людей живе погано
2	Демократія надає забагато свободи багатим людям
3	Демократичні вибори — це фарс, що аж ніяк не захищає інтереси простих людей
4	Демократія — це тільки слова, якими прикриваються ті, хто має доступ до влади для забезпечення власних інтересів
5	Демократичні вибори, як правило, приводять до влади найкорисливіших людей
6	Демократія надає людині оманливу свободу вибору, якою вона не може скористатися
7	Демократія несе простій людині невпевненість у завтрашньому дні
8	Усі розмови про свободу — лише для того, щоб у країні не було порядку
9	Свобода слова — гасло влади, яким вона заспокоює громадську думку
<i>Тоталітарні установки</i>	
1	Коли немає жорсткого державного контролю, засоби масової інформації сіють у суспільстві зерно аморальності й уседозволеності
2	Держава має здійснювати жорсткий контроль над цінами
3	Держава має жорстко контролювати економіку
4	Державний контроль над усім, що відбувається в суспільстві, забезпечує спокій і добробут більшості людей
5	Держава має бути «всьому головою» і гарантувати гідне життя своїм громадянам
<i>Авторитарні установки</i>	
1	Єдина політична воля забезпечує стабільність у суспільстві краще, ніж набір різних точок зору
2	Для нормального розвитку країні потрібна «сильна рука», а не розмови про демократію
3	Відсутність сильного політичного лідера, здатного підкорити всіх своїй волі, призводить до безладу у країні
4	Сильна держава можлива лише тоді, коли на чолі країни стоїть один лідер
5	Тільки політика «сильної руки» може зберегти порядок у суспільстві
6	Наявність єдиної партії та сильного політичного лідера дає людям упевненість у завтрашньому дні

¹¹¹ Докладніше про використання стратегії узгодження концептів див.: (Дембицкий, 2016с, с. 81–98).

У виборі між цими групами пріоритет віддано твердженням, що відображають антидемократичні й авторитарні установки. Антидемократичні твердження сформульовані в доволі провокаційному стилі, що виводить їх за межі офіційного політичного дискурсу, в рамках якого демократія — це завжди добре. При цьому вони оголюють проблеми, характерні для суспільств, у яких той самий демократичний дискурс стає інструментом досягнення вельми партикулярних цілей окремих груп інтересів за рахунок більшості населення. Що стосується авторитарних тверджень, то вибір на їх користь продиктований як мінімум трьома факторами. По-перше, неухильно високою оцінкою респондентами соціологічного моніторингу «Українське суспільство» такого авторитарного лідера, як О. Лукашенко, впродовж усіх років його президентства в Білорусі. Так, 2016 року його середня оцінка становила 6,3 бала з 10 можливих, що вище за відповідні оцінки П. Порошенка, Б. Обама і В. Путіна. Крім того, до моменту анексії Автономної Республіки Крим, а також збройного конфлікту на Донбасі стабільно високу оцінку отримував і В. Путін. По-друге, 62,5% усіх респондентів згодні з твердженням, що «кілька сильних керівників можуть зробити для нашої країни більше, ніж усі закони і дискусії». По-третє, політичне життя України розгортається переважно не навколо детально розроблених модернізаційних проектів, а навколо популістських обіцянок і осіб, котрі їх дають.

Відмова від тверджень, що фіксують колективістські установки, продиктована тим, що, згідно з думкою респондентів, запитання «індивідуалізм чи колективізм?» не стоїть на політичному порядку денному. Так, прихильників соціалізму підтримують 14,4%, а прихильників капіталізму — 13,4%. Решта або нікого не підтримує (32,0%), або підтримує обидві сторони, щоб обійшлося без конфліктів (23,4%), або вагається з відповіддю (15,8%). Повернення до колективістського світогляду на національному рівні виглядає нині малоімовірним. Тоталітарні установки, знову ж таки, дані не підтверджують. Скажімо, за повернення до планової економіки виступають 25,9% опитаних, натомість за змішану модель — 51,6%, а за суто ринкову — 11,5%. Ми вважаємо, що оцінювати потенційну можливість виникнення тоталітарної держави або її окремих елементів за умов політичного життя України слід саме в контексті підтримки авторитарного лідера, і не як очікуваний результат його діяльності, а радше як непередбачений і небажаний наслідок (латентну функцію).

Отже, ключова зв'язка тут — рівень скепсису щодо дії демократичних інститутів як загалом, так і в Україні, а також прихильне ставлення до постаті авторитарного лідера, здатного увідповіднити таку діяльність декларованим цінностям і цілям.

Фінальний відбір індикаторів. Як показала практика використання СТ ТПК, шести пунктів методики цілком достатньо як для здійснення статистичного аналізу, так і для поділу респондентів на різні типи. Тому для остаточного відбору індикаторів із груп, що описують антидемократичні й авторитарні установки, було необхідно оцінити їх з точки зору вимог до змістової валідності. Останні стосуються двох її принципів складових — концептуального пред-

ставництва (достатня широта охоплення змістового поля) і значимості (важливість елементів).

Оцінюючи пункти антидемократичної спрямованості, можна виокремити три аспекти в їх змісті. Перший із них стосується того, що демократія захищає передусім інтереси тих груп, що перебувають при владі або прагнуть досягти її. Другий — того, що демократія відіграє в житті простих людей негативну, а не позитивну роль. Третій — свободи загалом і свободи слова зокрема. Оскільки останній аспект стоїть дещо осторонь (це показало експертне опитування), відбір проведено на підставі перших двох параметрів (еліти vs маси) — по три твердження для кожного (№ 2–7). Перше твердження було вилучене, адже з ним важко не погодитися.

Щодо тверджень, що оцінюють авторитарні установки, змістовий аналіз показав таке: шість тверджень охоплюють п'ять соціетальних цінностей — стабільність (№ 1), розвиток (№ 2), порядок (№ 3, 5), силу (№ 4), перспективи (№ 6). Щодо цінності порядку обрано п'яте твердження. Також було сформульовано додаткове твердження, пов'язане з такою цінністю, як безпека: «Тільки по-справжньому сильний лідер може захистити те, що дорого всім».

Зрештою, 12 фінальних індикаторів мають уявнити два аспекти в політичній свідомості респондентів: 1) чи вважають вони, що демократія має суто інструментальний характер у руках високопоставлених груп інтересів і при цьому жодною мірою не розв'язує нагальних проблем населення; 2) чи вважають вони, що найефективнішим способом забезпечення базисних соціетальних цінностей є діяльність авторитарного лідера.

§ 14.2. Результати методичного дослідження: валідація СТ ТПК-II

Дизайн. Для тестування запропонованої методики було проведено методичне опитування серед жителів Києва (жовтень — листопад 2017 року), в рамках якого опитано 100 респондентів. Вибірка побудована на базі шести квот, сформованих на підставі статі (2 категорії) та віку (3 категорії)¹¹². Додатковим елементом квотування був район проживання респондентів (по 10 респондентів у кожному з районів Києва). Пропущені значення відсутні. Середній вік респондентів 44,6 року.

Інструмент. Основну частину анкети складено за чотирма методиками. 1. Скорочена версія Інтегрального індексу соціального самопочуття на 14 індикаторів¹¹³ (далі ПСС-14), сконструйована в рамках нашого недавнього до-

¹¹² Групи респондентів серед молоді включали по 10 респондентів, натомість серед представників середнього і похилого віку — по 20. Це пов'язане з тим, що методичне опитування проведено спільно зі спеціалістами кафедри методології та методів соціологічних досліджень КНУ імені Тараса Шевченка (відповідальний виконавець — аспірант кафедри Альона Ковальська), зацікавленими у вивченні соціального статусу жителів Києва.

¹¹³ Розширений за рахунок показника щодо конфліктів, які виникають на політичному ґрунті, ПСС-13.

слідження (Дембіцький, 2017с). Особливістю цієї версії є її фокусування на інституційних особливостях досліджуваного суспільства. 2. Два блоки запитань, що вимірюють антидемократичні установки респондентів, ідентичних з точки зору змісту індикаторів. Перший блок спрямований на вимірювання антидемократичних установок щодо розвинених країн, другий — стосовно України (див. Додаток 7). 3. Блок запитань, що вимірює авторитарні установки. У цьому разі йдеться тільки про Україну (див. Додаток 7). 4. Індекс психологічного дистресу SCL-9-NR (див. Додаток 1).

У цій роботі блоки запитань щодо антидемократичних установок, а також авторитаризму розглянуто як вимірювальні шкали¹¹⁴. Поряд із тим як альтернативні шкали розглянуто два набори індикаторів, що дають змогу вимірювати: 1) авторитарні, а також антидемократичні установки щодо розвинених країн (два виміри)¹¹⁵; 2) як авторитарні, так і всі антидемократичні установки — і стосовно розвинених країн, і щодо України (три виміри).

Крім цього поставлено додаткові запитання (всі з використанням категоріальних шкал відповідей), що стосуються рівня освіти (5 градацій), матеріальної забезпеченості (5 градацій) і суб'єктивного соціального статусу (7 градацій).

Після отримання емпіричних даних першочерговим завданням була спільна перевірка індикаторів, розглядуваних нами як вимірювальні шкали. Дизайн проведеного дослідження дав змогу оцінити факторну валідність і внутрішню узгодженість, а також назагал проаналізувати конструктну валідність.

Факторна валідність перевірена на підставі результатів конфірмаційного факторного аналізу (*Confirmatory Factor Analysis* — CFA) з використанням методу діагонально зважених найменших квадратів (*Diagonally Weighted Least Squares* — DWLS). Якість факторних моделей оцінювали на підставі значень таких показників: χ^2 -квадрат; середньоквадратична похибка апроксимації (RMSEA); порівняльний індекс відповідності (CFI); індекс Такера—Левіна (TLI). За прийнятні граничні значення було прийнято: відношення значення χ^2 -квадрат до ступенів свободи (*df*) менш ніж 5; RMSEA < 0,07; (CFI, TLI) > 0,95.

Оцінювання внутрішньої узгодженості методики здійснено за коефіцієнтом Альфа Кронбаха. Значення, що перевищують 0,7, розглядали як прийнятні.

Конструктну валідність оцінювали шляхом кореляційного аналізу результатів використання шкал антидемократичних установок і авторитаризму (підсумкові значення знаходили шляхом підсумовування результатів за окремими індикаторами), з одного боку, а також значень ПСС-14 і SCL-9-NR (використано коефіцієнт кореляції Пірсона), рівня освіти, матеріальної забезпеченості та суб'єктивного соціального статусу (застосовано коефіцієнт кореляції Спірмена) — з іншого.

¹¹⁴ Альтернативний підхід до індикаторів авторитаризму полягає в розгляді їх як індексу, а не шкали. У цьому разі вважають, що прийнятність авторитарної влади можна оцінити з точки зору забезпечення різних соціетальних цінностей, які не обов'язково пов'язані між собою. Утім, для перевірки цієї гіпотези необхідне здійснення масштабнішого дослідження.

¹¹⁵ Чому вибір зроблено на користь антидемократичних суджень щодо розвинених країн, а не України, буде пояснено далі.

Для статистичного аналізу СТ ТПК використано мову програмування R (версія 3.4.1, бібліотеки «lavaan» для проведення CFA, «lrm» для розрахунку значень Альфа Кронбаха, «rsuch» для кореляційного аналізу).

Внутрішня узгодженість і факторна валідність. Показники внутрішньої узгодженості та факторної валідності наведено далі (табл. 14.3). Для всіх шкал характерні високі показники. Виняток становить показник RMSEA для шкали авторитарних установок.

Таблиця 14.3

**Результати оцінювання надійності та факторної валідності
для різних вимірювальних шкал**

Показник	АДР	АДУ	АВТ	АДР + АВТ	АДР + АДУ + АВТ
Альфа Кронбаха (> 0,70)	0,902	0,886	0,888	0,899	0,913
Хі-квадрат (DWLS)	4,279	12,105	17,328	57,278	139,572
df (ступені свободи)	9	9	9	53	132
Хі-квадрат / df	0,5	1,4	1,9	1,1	1,1
p-value	0,892	0,207	0,044	0,319	0,309
RMSEA (< 0,07)	0,000	0,059	0,096	0,028	0,024
CFI (> 0,95)	1,000	0,999	0,996	0,999	0,999
TLI (> 0,95)	1,003	0,998	0,993	0,999	0,999

Скорочення в таблиці: АДР — шкала антидемократичних установок щодо розвинених країн; АДУ — шкала антидемократичних установок щодо України; АВТ — шкала авторитарних установок; АДР + АВТ — спільне використання шкал антидемократичних установок щодо розвинених країн і авторитарних установок; АДР + АДУ + АВТ — спільне використання всіх трьох шкал.

Підтвердивши внутрішню узгодженість і факторну валідність шкал, можна звернутися до аналізу їх конструктивної валідності (зовнішньої структури), а також оцінити особливості взаємозв'язків шкал (внутрішню структуру).

Внутрішня і зовнішня структура. Для простих шкал (ідеться про окремо взяті шкали авторитарних і антидемократичних установок) оцінити внутрішню структуру можна на підставі сили зв'язку окремих індикаторів із загальним індексом, для комплексних (що включають індикатори двох і більше простих шкал) цю інформацію можна доповнити даними про взаємозв'язки між різними вимірами шкали.

Як можна бачити, індикатори антидемократичних установок демонструють практично однакові результати в разі оцінювання політичної дійсності як у розвинених країнах, так і в Україні загалом (табл. 14.4), усі взаємозв'язки статистично значимі, й серед них нема таких, які б вочевидь «просідали» з точки зору сили зв'язку.

Таблиця 14.4

Сила зв'язку окремих індикаторів із сумарними значеннями різних шкал
($p < 0,001$), r_s

Індикатор	АДР	АДУ	АВТ	АДР + АВТ	АДР + АДУ + АВТ
Демократія надає забагато свободи багатим людям	0,69	0,74	—	0,58	0,53/0,56*
Демократичні вибори — це фарс, що не захищає інтересів простих людей	0,86	0,82	—	0,68	0,69/0,54*
Демократія — це тільки слова, якими прикриваються ті, хто має доступ до влади для забезпечення власних інтересів	0,82	0,83	—	0,76	0,75/0,67*
Демократичні вибори, як правило, приводять до влади найкорисливіших людей	0,82	0,80	—	0,60	0,63/0,62*
Демократія надає людині оманливу свободу вибору, якою вона не може скористатися	0,86	0,83	—	0,77	0,78/0,66*
Демократія несе простій людині невпевненість у завтрашньому дні	0,76	0,74	—	0,60	0,58/0,61*
Єдина політична воля забезпечує стабільність у суспільстві краще, ніж набір різних точок зору	—	—	0,75	0,66	0,53
Для нормального розвитку країні потрібна «сильна рука», а не розмови про демократію	—	—	0,81	0,71	0,64
Тільки політика «сильної руки» може зберегти порядок у суспільстві	—	—	0,83	0,74	0,58
Сильна держава можлива лише тоді, коли на чолі країни стоїть один лідер	—	—	0,81	0,71	0,60
Наявність сильного політичного лідера дає людям упевненість у завтрашньому дні	—	—	0,75	0,62	0,55
Тільки по-справжньому сильний лідер може захистити те, що дорого всім	—	—	0,73	0,53	0,51

* Перше значення в комірці належить до тих індикаторів, що стосуються розвинених країн, друге — до індикаторів, що стосуються України.

Тепер розглянемо особливості взаємозв'язку між підсумковими значеннями простих шкал (табл. 14.5). Цікаво, що авторитарні установки показують виразніший взаємозв'язок з антидемократичними установками щодо розвинених країн, а не України. На наш погляд, це інтуїтивно зрозумілий результат, що піддається соціологічній інтерпретації. Так, феномен демократії, розглядуваний у контексті існування розвинених країн, має бути значно ближчим до ціннісної структури, постульованої з позицій демократії як ідеального типу

людської активності. З відомих причин в Україні феномен демократії є значно віддаленішим від відповідних ідеалів. Тому слід очікувати виразнішого взаємозв'язку між авторитарними установками і скепсисом щодо усталених демократій, а не того, що лише має таку назву.

Таблиця 14.5

Сила зв'язку між сумарними значеннями різних шкал ($p < 0,001$)

Змінні	Сила зв'язку
Антидемократичні установки щодо розвинених країн — антидемократичні установки щодо України	0,52
Антидемократичні установки щодо розвинених країн — авторитарні установки	0,44
Антидемократичні установки щодо України — авторитарні установки	0,30

Далі звернімося до перевірки конструктивної валідності. Як змінні, що відтворюють номологічну мережу, використано показники соціального самопочуття і психологічного дистресу респондентів, а також рівень їхньої освіти, матеріальної забезпеченості та суб'єктивного соціального статусу (табл. 14.6). Для наведених результатів характерні достатньо низька сила зв'язку і доволі різноплановий відгук трьох аспектів політичної культури на елементи номологічної мережі.

Таблиця 14.6

Показники конструктивної валідності для різних шкал

Показники номологічної мережі	Антидемократія, розвинені країни	Антидемократія, Україна	Авторитаризм
Соціальне самопочуття	-0,18 [^]	-0,12	-0,01
Психологічний дистрес	0,22*	0,14	0,22*
Рівень освіти	-0,13	-0,08	-0,20*
Рівень матеріальної забезпеченості	-0,25**	-0,10	-0,07
Суб'єктивний соціальний статус	-0,22*	-0,18 [^]	-0,05

[^] $p < 0,10$; * $p < 0,05$; ** $p < 0,01$.

Низькі значення сили зв'язку є показником того, що політична культура респондентів посідає окреме місце серед характеристик індивіда, а отже, є важливим аспектом, що впливає на поведінку. Показово і те, що шкали антидемократичних установок і авторитаризму вирізняються сильнішими взаємними зв'язками, ніж зв'язками зі змінними для конструктивної перевірки. Це дає змогу розглядати їх укупі як індикатори різних аспектів єдиного феномену.

Що стосується гетерогенності значень у таблиці 14.6, то вони також підлягають обґрунтованій соціологічній інтерпретації. Насамперед треба звернути

увагу на кількість статистично значимих взаємозв'язків у випадку кожної зі шкал. Найбільше їх спостерігаємо для шкали антидемократичних установок щодо розвинених країн (усього чотири), найменше — для шкали антидемократичних установок щодо України (лишень одна). У випадку шкали авторитаризму таких зв'язків два.

Ми вважаємо, що повніша включеність антидемократичних установок щодо розвинених країн у номологічну мережу полягає в тому, що вони більшою мірою ґрунтуються на культурних стереотипах, а тому зазнають виразніших змін залежно від соціальних характеристик індивіда. Так, вище соціальне самопочуття, матеріальна забезпеченість і суб'єктивний соціальний статус пов'язані з вищою оцінкою феномену демократії в тій його формі, що є найближчою до відповідних принципів і цінностей.

Що стосується антидемократичних установок щодо української дійсності, то тут ідеться не так про культурні стереотипи, як про певне загальне знання про політичну ситуацію в Україні. У цьому сенсі такого роду установки є феноменом, що має більшу кількість незалежних об'єктивних підстав, а отже, меншою мірою чутливий до соціального самопочуття і матеріальної забезпеченості.

Натомість у разі авторитарних установок спрацьовує рівень освіти респондента — нижча освіта призводить до більшої ймовірності підтримки авторитарних ідей.

Читач міг помітити, що «за дужками» залишився такий елемент номологічної мережі, як психологічний дистрес. Як бачимо, цей показник у середньому демонструє найтісніші взаємозв'язки зі шкалами, націленими на політичну культуру респондента. На наш погляд, це пов'язане з тим, що загальний емоційний стан певною (нехай навіть незначною) мірою впливає на результати опитування. Альтернативна гіпотеза, висловлена Є. Головахою, стверджує, що рівень психологічного дистресу все ж таки є одним із факторів, які позначаються на політичних установках респондентів.

Після специфікації внутрішніх і зовнішніх взаємозв'язків шкал звернімося до сукупного аналізу їх результатів з метою виокремлення груп респондентів, для яких характерні відмінні патерни політичної культури.

§ 14.3. Результати методичного дослідження: типологія респондентів, інтерпретація результатів

Перш ніж перейти безпосередньо до типології респондентів, розглянемо описові характеристики шкал (табл. 14.7). Як можна бачити, найбільшою мірою виражені антидемократичні установки стосовно України, що доволі очікувано. Останнє можна пояснити загальною низькою ефективністю політичної системи в Україні, де демократія насправді значно ближча до гарних гасел, аніж до відповідної активності як еліт, так і мас. Наступними за виразністю йдуть авторитарні установки, котрі, водночас, доволі є близькими до теоретичної середньої інтенсивності (18 балів). Зрештою, найменш вираженими є антидемократичні установки щодо розвинених країн, що інтуїтивно зрозуміло і цілком очікувано.

Таблиця 14.7

Описові характеристики різних шкал

Установки	Мінімальне значення	1-й квартиль	Медіана	Середнє значення	3-й квартиль	Максимальне значення
Антидемократичні установки (розвинені країни)	6	13	17,0	17,0	21	30
Антидемократичні установки (Україна)	6	17	21,5	21,2	26	30
Авторитарні установки	6	14	19,0	18,7	23	30

Додаткову описову інформацію надають розподіли відповідей за окремими індикаторами (табл. 14.8). При цьому відповіді згруповано за трьома модальностями: незгоди — «-» (варіанти відповідей «скоріше не згоден» і «абсолютно не згоден»), невизначеності — «?» (варіант відповіді «важко сказати, згоден чи ні»), згоди — «+» (варіанти відповідей «цілком згоден» і «скоріше згоден»). Для кожного з індикаторів, що характеризують антидемократичні установки, наведено два розподіли відповідей — верхній рядок відповідає установкам щодо демократії в розвинених країнах, нижній — в Україні.

Отже, відносна більшість респондентів (в одному з випадків абсолютна більшість) не згодна практично з усіма антидемократичними індикаторами, якщо йдеться про розвинені країни. Виняток становить індикатор, згідно з яким демократичні вибори приводять до влади найкорисливіших людей. У цьому разі думки розділилися практично порівну.

Що стосується антидемократичних установок щодо України, то практично з усіма твердженнями погоджується абсолютна більшість респондентів. Винятком є пункт про те, що демократія несе простій людині невпевненість у завтрашньому дні. Утім, і тут маємо справу з більшістю, але відносною.

Натомість індикатори авторитаризму не показують такої однорідності. У двох випадках (цінності стабільності та розвитку) переважання згоди над незгодою мінімальне. В одному випадку переважає незгода (цінність порядку). І в трьох випадках кількість згодних респондентів доволі відчутно переважає над кількістю незгодних (цінності сили, перспективи і безпеки).

Інтерпретація за антидемократичними групами індикаторів є тут зайвою (тим паче, що це запитання ми вже розглядали раніше на рівні шкал), проте для оцінювання смислової гетерогенності індикаторів шкали авторитаризму потрібні подальші дослідження.

З огляду на таку різноплановість установок, фіксованих із застосуванням різних шкал, необхідно сформулювати обґрунтовану систему типології респондентів. Для цього слід співвіднести шкали з точки зору їх важливості сто-

совно політичної культури, а також визначити оптимальний спосіб розподілу або кластеризації респондентів за типами.

Таблиця 14.8

Розподіл відповідей за індикаторами різних шкал, %

Індикатор	–	?	+
Демократія надає надто багато свободи багатим людям	42	29	29
	23	13	64
Демократичні вибори — це фарс, що не захищає інтереси простих людей	43	24	33
	23	20	57
Демократія — це тільки слова, якими прикриваються ті, хто має доступ до влади для забезпечення власних інтересів	42	22	36
	22	19	59
Демократичні вибори, як правило, приводять до влади найкорисливіших людей	35	29	36
	21	23	56
Демократія надає людині оманливу свободу вибору, якою вона не може скористатися	43	23	34
	22	27	51
Демократія несе простій людині невпевненість у завтрашньому дні	52	21	27
	27	33	40
Єдина політична воля забезпечує стабільність у суспільстві краще, ніж набір різних точок зору	32	32	36
Для нормального розвитку країні потрібна «сильна рука», а не розмови про демократію	34	27	39
Тільки політика «сильної руки» може зберегти порядок у суспільстві	40	24	36
Сильна держава можлива лише тоді, коли на чолі країни стоїть один лідер	34	22	44
Наявність сильного політичного лідера дає людям упевненість у завтрашньому дні	26	25	49
Тільки по-справжньому сильний лідер може захистити те, що дорого всім	27	27	46

Що стосується розв'язання першого завдання, то зміст СТ ТПК-II вказує на два загальні виміри політичної культури (демократизм/антидемократизм, авторитаризм/антиавторитаризм), кожен із яких має принципове значення для розуміння відповідного феномену. При цьому авторитарні установки є показовішими за антидемократичні. Це пов'язано з тим, що респондентів, котрі демонструють одночасно авторитарні та демократичні установки, варто віднести до авторитарної групи (далі це питання розглянемо додатково). Крім того, треба правильно оцінити і зіставити важливість двох шкал антидемокра-

тичних установок. На наш погляд, шкала, що відбиває оцінку демократії в розвинених країнах, має більше концептуальне значення як інструмент вивчення ставлення до демократії в Україні. Ми пов'язуємо це з тим, що коли йдеться про розвинені країни, об'єктом установок виступає соціальна дійсність, значно ближча до ідеалів демократії, ніж це має місце у випадку України. Тому перша шкала антидемократичних установок (розвинені країни), а також шкала авторитаризму мають, під кутом зору типології респондентів, першорядне, або вирішальне значення, а друга шкала антидемократичних установок (Україна) — другорядне, або додаткове.

Розв'язок завдання вибору техніки кластеризації респондентів у рамках вивчення політичної культури вже розглядався вище. Так, було здійснено порівняння чотирьох технік кластеризації: ідеальна точка поділу, концептуальна типологія, кластерний аналіз і латентний класовий аналіз. Серед іншого з'ясувалося, що прийнятно точною, найпростішою у використанні, а також такою, що дає інтуїтивно зрозумілі соціологічні результати, є кластеризація через ідеальну точку поділу, тобто визначення точок, або координат, на осях шкал, що розбивають відповіді респондентів на якісно відмінні групи.

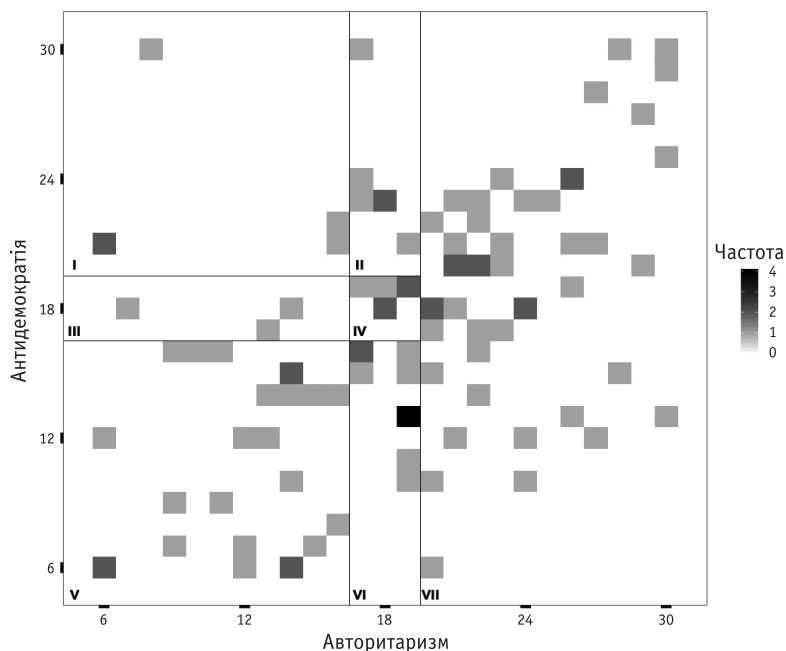


Рисунок 14.1. Теплова діаграма частотної наповненості категорій установок на антидемократію й авторитаризм

Для полегшення кластеризації респондентів на підставі результатів СТ ТПК-II розроблено діаграму, що співвідносить антидемократичні й авторитарні установки з використанням кластеризації на підставі ідеальної точки поділу (рис. 14.1).

Розгляньмо, як використовують цю діаграму. Передусім треба визначити координати на осях, що відповідають шкалам, за якими поділені респонденти. Оскільки застосовано адитивні шкали з діапазоном можливих значень від 6 до 30, ми пропонуємо скористатися центральним значенням, що дорівнює 18 балам. Позаяк одним із варіантів відповіді за кожним з індикаторів є «важко сказати, згоден чи ні», то навколо центральної точки необхідно окреслити зону, потрапляння в яку сигналізуватиме про невизначену позицію респондента. Як таку ми взяли інтервал від 17 до 19 балів за кожною зі шкал¹¹⁶. Виходячи з цього, на кожній шкалі було отримано три інтервали: від 6 до 16 балів (демократичні та антиавторитарні установки), від 17 до 19 балів (невизначені установки) і від 20 до 30 балів (антидемократичні й авторитарні установки). Відповідні результати наведено далі (табл. 14.9).

Таблиця 14.9

Розподіл установок для різних шкал, %

Шкала	Характер установок		
	Негативні (-)	Невизначені (?)	Позитивні (+)
АДР	35	18	47
АДУ	62	15	23
АВТ	32	22	46

Скорочення в таблиці: *АДР* — шкала антидемократичних установок щодо розвинених країн; *АДУ* — шкала антидемократичних установок щодо України; *АВТ* — шкала авторитарних установок.

Оскільки визначальними є результати за шкалою авторитаризму, сукупне використання їх із результатами шкали, що фіксує ставлення до демократії в розвинених країнах, дає нам сім можливих типів респондентів щодо їхньої політичної культури (табл. 14.10). Що стосується результатів шкали демократичних установок щодо української дійсності, то вони лише повідомляють додаткові нюанси про ці типи.

Додатково треба сказати про суперечливі результати, а також про результати, що можуть такими видаватися. Насамперед це стосується спостережень, що демонструють позитивні установки як на демократію, так і на авторитаризм. На наш погляд, у цьому разі йдеться про установки на так званий обізнаний авторитаризм (альтернативна інтерпретація — інструментальний автори-

¹¹⁶ За аналізу результатів СТ ТПК для таких цілей ми використовували інтервал від 16 до 20 балів. Але в цьому разі він видається нам надто широким (що призводить до надмірної категоризації респондентів як таких, котрі вагаються). Це пов'язане як із великою кількістю використовуваних шкал, так і з характером розподілу відповідей. Так, розкид значень СТ ТПК-II в нашому методичному дослідженні значно ближче до 18 балів, ніж ми це спостерігали в разі застосування шкал СТ ТПК. В останньому випадку спостерігалася значна асиметрія з ухилом у «демократичність» і «громадську пасивність» респондентів.

таризм). Утім, що справді видається нам суперечливими результатами, так це наявність позитивних установок щодо української демократії за фіксації негативних установок стосовно демократії розвинених країн. Пояснити такі випадки концептуально (їх зовсім мало) нам не вдалося. Напевно, наявність їх характеризує якість заповнення анкет респондентами.

Таблиця 14.10

Типологія респондентів на підставі СТ ТПК-II

Установки на демократію	Установки на авторитаризм		
	Негативні (-)	Невизначені (?)	Позитивні (+)
Негативні (-)	I. Політичні скептики	II. Скептики щодо демократії	VII. Прихильники авторитаризму
Невизначені (?)	III. Скептики щодо авторитаризму	IV. Ті, що не визначилися	
Позитивні (+)	V. Прихильники демократії	VI. Демократи, що вагаються	

Тепер розгляньмо розподіл респондентів за типами (табл. 14.11). У тих клітинках, де наведено друге значення в дужках, додаткова інформація стосується кількості суперечливих результатів. Що стосується переважання респондентів з авторитарними установками, то ані стать, ані вік не пов'язані з імовірністю потрапляння до цієї групи.

Таблиця 14.11

Результати використання СТ ТПК-II, %

Установки на демократію	Установки на авторитаризм		
	Негативні (-)	Невизначені (?)	Позитивні (+)
Негативні (-)	I. Політичні скептики — 5 (0)	II. Скептики щодо демократії — 5 (0)	VII. Прихильники авторитаризму — 46 (2)
Невизначені (?)	III. Скептики авторитаризму — 3	IV. Невизначені — 6	
Позитивні (+)	V. Прихильники демократії — 24	VI. Демократи, що вагаються — 11	

Як бачимо, двома найбільшими групами є респонденти з демократичними поглядами (прихильники демократії та демократи, що вагаються), а також прихильники авторитаризму. Назагал їх 81%. Подальше вивчення результатів типології на підставі СТ ТПК-II передбачає включення тесту до масштабніших дослідницьких проектів, що мають кращу репрезентативність і охоплюють більше розмаїття анкетних запитань і методик.

§ 14.4. Замість висновків: фінальна конфігурація тесту

Наостанок звернімо увагу на таке важливе питання, як фінальна конфігурація СТ ТПК-II. Для цього передусім слід повернутися до змісту СТ ТПК. Так, перша версія методики мала 12 індикаторів, шість із них були спрямовані на вимірювання демократичних установок, шість — на громадську активність. І якщо прийнятність індикаторів, спрямованих на вимірювання ставлення до демократії, було розглянуто на початку роботи, то решту складових методики ми ще не розглядали. Як показала валідація шкали громадської активності, шість індикаторів, що вимірюють громадську активність, можна розглядати як окрему шкалу чи підшкалу в рамках більш комплексного вимірювального інструменту. Отже, немає жодних підстав відмовлятися від її застосування за вивчення політичної культури респондентів і з використанням СТ ТПК-II. Що стосується аналізованого у цій роботі дослідження, то (під)шкалу громадської активності не було включено в опитування лише через фінансові обмеження¹¹⁷.

Щодо всіх трьох шкал, розглянутих тут, то використання їх є доцільним, оскільки кожна з них надає свою частину релевантної соціологічної інформації. При цьому під кутом зору типології респондентів, достатньо двох шкал — антидемократичних установок щодо розвинених країн, а також авторитарних установок.

Отже, найповніший у плані змісту методики варіант СТ ТПК-II містить 24 індикатори, поділені на чотири (під)шкали — антидемократичних установок щодо розвинених країн, антидемократичних установок щодо українських реалій, авторитарних установок і установок на громадську активність (див. Додаток 7). Це не суперечить можливості вилучення з тесту окремих шкал (які саме шкали мають бути вилучені, залежить від концептуальних рамок конкретного дослідження), а також застосування окремих (під)шкал самих по собі.

¹¹⁷ Дослідження проведене за власні кошти С. Дембіцького й А. Ковальської.

Глава 15

Незалежна Україна в геополітичному дискурсі: віхи й акценти

Актуальність проблеми зовнішньополітичних орієнтацій населення дуже складно переоцінити, тим більше, коли йдеться про таку країну, як Україна, особливо за нинішніх умов. Однак перш ніж говорити про зовнішньополітичні орієнтації (нас першою чергою цікавить специфіка їх фіксації або вимірювання соціологічними засобами), необхідно звернутися до їх базису — до геополітичної активності політичних еліт. Така постановка питання не є випадковою — як буде показано далі, обговорення України в геополітичному контексті відбуватиметься переважно з фокусуванням на рішеннях і діях політичних еліт України, США, Євросоюзу і Російської Федерації.

В аналізі ми також дотримуємося цієї логіки — політичні рішення і дії всіх релевантних учасників геополітичних процесів, у які включена та або та держава, є ключовим фактором, який оформлює зовнішньополітичні орієнтації її населення. Отже, для визначення ключових тем, завдяки яким розкриваються зовнішньополітичні орієнтації населення України, вважаємо за необхідне звернутися до теми незалежної України в геополітичному дискурсі. Останній умовно можна поділити на англomовний і україно-/російськомовний. Тут зупинимося на концептуальних положеннях, поданих в англomовних публікаціях.

Публікації розглядатимемо у хронологічній послідовності, щоб зрозуміти динаміку уявлень про геополітичний статус і особливості незалежної України впродовж її новітньої історії. В окремих випадках робитимемо відступи від цього правила (наприклад, коли публікації зачіпають одну і ту саму тематику).

Ми виходимо з того, що публікації з геополітичної тематики є спробами кейс-наративного аналізу, а отже, їх аналіз потенційно веде до розв'язання завдань теоретичної валідазації, що є принциповим кроком до розробки соціологічних тестів геополітичних орієнтацій населення України.

§ 15.1. 1991–2004: ера перших президентів

Для людини, яка започаткувала обговорення теми геополітичної ролі незалежної України на пострадянському просторі, події, що почалися в Україні 2014 року, ймовірно, не були несподіванкою. Ідеться про Збігнева Бжезінського, який почав академічний розгляд цієї теми далекого 1996 року (Brzezinski, 1996). Не буде перебільшенням те, що він точно визначив основні точки

напруги між Україною і Російською Федерацією, які були і залишаються ключовими в розвитку відносин між двома країнами.

У своїй праці З. Бжезінський говорить про критичну роль України в геополітичному просторі колишнього СРСР. Перш за все автор згадує про роль Л. Кравчука, який брав участь у Біловезькій угоді й наполіг на «цивілізованому розлученні» країн, що входили до складу Радянського Союзу. Загалом це привело до того формату СНД, який існував далі. Це, своєю чергою, спричинило низку глобальних змін. По-перше, істотно трансформувалося європейське геополітичне «рівняння», пов'язане зі зменшенням потужності Росії. На думку З. Бжезінського, незалежний статус України забезпечив безпеку Польщі, Румунії та Туреччині. По-друге, концепція СНД була заснована на кооперації задля захисту безпеки незалежних держав, а не на побудові інтегрованої спільності, що знову-таки могло посилити імперські амбіції Росії. Зрештою, незалежність України була важливою для нашого «великого брата» самого по собі, оскільки ставила його перед фундаментальним питанням самовизначення: «Що таке Росія?». Фактично реакція еліт Російської Федерації на геополітичне самовизначення України і є показником того, чи можна вважати Росію європейською державою.

Далі З. Бжезінський указує на тривожний момент, що полягає в позиції Росії з приводу Чечні, яка стала ареною кровопролитної війни. Для України цей момент був символічним, оскільки сигналізував принципову позицію частини російського політикуму щодо низки територій, важливих для нього. Серед останніх був і Крим (особливо Севастополь). Кілька впливових російських політиків відкрито розглядали його як точку зіткнення інтересів Росії та НАТО. А оскільки З. Бжезінський бачив Україну невід'ємною частиною Центральної Європи, що передбачало подальший рух першої в бік Північноатлантичного альянсу, ситуація видавалася йому потенційно вибухонебезпечною.

Що говорив З. Бжезінський про геополітичне майбутнє України? Перш за все він бачив Україну частиною Західного світу. Тому вважав правильним зміцнення зв'язків між Україною і Польщею, а через Польщу і з Німеччиною. Також він уважав за необхідне налагодження співпраці з Румунією і Туреччиною, оптимізацію економічної взаємодії з Росією, а також початок конструктивного діалогу з Азербайджаном і Узбекистаном для забезпечення геополітичного плюралізму.

Ситуація у Криму як важлива геополітична для України проблема була докладно проаналізована у статті Джейн Доусон (Dawson, 1997). Оскільки перша половина 1990-х років пройшла на півострові під знаком напружених політичних конфліктів ненасильницького характеру, авторка зацікавилася особливістю трьох політичних розколів у Криму — етнічного, ідеологічного і геополітичного з тим, щоб зрозуміти, як удалося зберегти мирний характер політичної боротьби в автономії.

Незважаючи на активні дії проросійських сил, спрямовані на посилення значених розколів, їх політична активність, що спочатку отримала широку підтримку серед виборців, у результаті завершилася колапсом, який розчистив шлях до успіху центристським проукраїнським силам.

Д. Доусон указує, що початковий успіх пов'язаний, по суті, з трьома факторами — економічними успіхами Російської Федерації на тлі економічної кризи в Україні, успіхом ідеологічної та геополітичної пропаганди проросійських сил на півострові, а також успіхом їх звернення до мовного питання. Фактично першочергова проблема полягала в тяжкому економічному становищі населення півострова, а ідеологічне і геополітичне обґрунтування необхідності возз'єднання з Росією було лише гарним приводом для приєднання до успішнішого сусіда. Крім того, з огляду на переважання на півострові російської мови, мовне питання завжди було тут чутливою темою, особливо в контексті вельми поширених на материковій частині України націоналістичних гасел у той період.

Колапс стався внаслідок як помилкової стратегії, так і внутрішніх і зовнішніх суперечностей у проросійському русі. Хибність стратегії полягала в акценті на етнічному розколі між росіянами й українцями у Криму. Фактично асиміляція цих двох груп на півострові була такою сильною, що поділу між росіянами й українцями насправді не існувало. Люди чекали економічних покращень, а не актуалізації етнічного питання. Погіршив ситуацію конфлікт усередині проросійських сил — між президентом АРК та її парламентом. Довершила фіаско позиція політичних еліт Російської Федерації, яка була зайнята Чеченської війною, що робило практично неможливою участь іще в одному серйозному конфлікті.

У результаті політичним елітам півострова (зокрема і «з благословення» міжнародної громадськості) довелося розв'язувати цей конфлікт сам на сам із політичними елітами материкової України, що завдяки низці компромісів привело до ухвалення нової конституції АРК 1996 року, а також до скасування президентської посади в автономії.

Заторкує Д. Доусон і кримськотатарське питання, незважаючи на його вторинність у рамках політичної боротьби на півострові в першій половині 1990-х років. Так, вона говорить про напруженість між двома політичними силами серед кримських татар — Організацією кримськотатарського національного руху (ОКТНР), або Меджлісом, і Національним рухом кримських татар. Для перших пріоритетним було етнічне питання (зокрема відновлення на цій основі кримськотатарської державності), для других — інтеграція в рамках існуючої політичної системи (цей напрямок часто звинувачували у проросійських симпатіях). Крім того, авторка вказує на високий мобілізаційний потенціал протестного руху в цій етнічній групі, що логічно вело до припущення про можливе етнічне протистояння в подальшому. Однак напругу в цьому напрямку знімали орієнтації ОКТНР/Меджлісу на співпрацю з центральними органами влади України.

Перші підсумки політичного і геополітичного курсу України в 1990-х роках підбив Домінік Агель (Agel, 1998). Розуміючи складність геополітичного положення України між Сходом і Заходом, він акцентує увагу на аспектах державного управління, які ще ускладнюють це становище (центральною постаттю аналізу став другий президент України Леонід Кучма). До цих аспектів він відносить:

- 1) відсутність економічних реформ, які б відповідали мінімальним критеріям ефективності, що виявилось у провальній монетарній політиці, неефективній приватизації, невиконанні кредитних зобов'язань перед західними партнерами;
- 2) очевидну (при тому зі схвалення президента) слабкість і неефективність уряду, що відкрила можливість «паразитичного» використання цінних ресурсів владними групами, а також інших проявів корупції;
- 3) «загравання» з мовними і геополітичними вподобаннями електорату виключно в цілях отримання влади; подальша зміна передвиборних гасел на діаметрально протилежні політичні рішення;
- 4) відсутність розв'язання проблеми енергетичної залежності від Російської Федерації, обтяженої корупційним використанням газотранспортної системи України завдяки її безальтернативності для Росії на той момент;
- 5) небажання брати участь у створенні взаємовигідних економічних відносин у рамках СНД;
- 6) суперечності між декларованими європейськими устремліннями еліт України та їх небажання/неспроможність забезпечити мінімально необхідні для цього економічні показники;
- 7) відсутність ефективної мовної політики;
- 8) клановість політичних еліт, що ведуть боротьбу, в якій головний принцип полягає в тому, що «переможець отримує все».

У підсумку автор робить висновок, що в геополітичному сенсі Україні доведеться балансувати між Сходом і Заходом, «російським» і «українським». А для успішного розв'язання цього завдання необхідне стабільне економічне середовище. На завершення Д. Арель зазначає, що незважаючи на поширену думку про обумовленість нестабільності України діями Росії, її (нестабільності) головна причина радше полягає у внутрішньополітичних рішеннях українських еліт.

Михайло Молчанов (Molchanov, 2000) звертається до проблеми націоналізму в посткомуністичних країнах як владного ресурсу (з фокусуванням на Росії та Україні). На його думку, посткомуністичний націоналізм відрізняється від його попередніх типів. Це пов'язане з тим, що його використовують не для консолідації суспільства або створення нової ідентичності, а для суперництва за ресурси, яке перетворило аморфні демократичні рухи на боротьбу за незалежність.

Незважаючи на те, що посткомуністичний націоналізм заведено пов'язувати з етнічною або етнополітичною прихильністю, автор вважає, що його коріння слід шукати у складніших нашаруваннях політики, економіки і міжнародних відносин (а отже, і в геополітичному статусі та амбіціях). У цьому сенсі націоналізм перетворюється на інструмент політичних еліт тих чи тих країн. Конфігурація внутрішніх і міжнародних факторів може вести, на думку М. Молчанова, до однієї з трьох стратегій мобілізації націоналізму — оборонної, селективної та емулятивної.

Оборонний націоналізм виникає на тлі реальних або уявних загроз існуванню нації чи можливостей для національного піднесення. Очевидно, що такого

роду націоналізм на пострадянському просторі відтворюють у Російській Федерації.

Своєю чергою, селективний націоналізм є характерним для групи держав, що складають так званий захисний пояс (наприклад, країни Балтії) між Заходом і його геополітичними «іншими». Селективний націоналізм може реалізуватися за допомоги найрізноманітніших стратегій, проте його головним принципом є прозахідна орієнтація, яка базується на всебічній підтримці з боку західних держав.

Зрештою, емулятивний націоналізм використовують у державах, які виникли насамперед завдяки розпаду СРСР. Їх природний, або історичний націоналізм є слабким або малорозвиненим, що висуває на перший план необхідність конструювання релевантних кодів і практик, покликаних визначити геополітичну позицію молодої незалежної держави (до таких автор відносить країни Центральної Азії).

В Україні, вважає М. Молчанов, спостерігаємо селективний націоналізм, що дозволяє отримувати вигоду з важливого геополітичного положення країни між Росією і Заходом. Також українські еліти стикаються з необхідністю використання емулятивного націоналізму, оскільки кордони сучасної України були оформлені в радянський період, що позначилося на посиленні культурної та політичної гетерогенності нашого суспільства.

Те, як різні країни СНД вибудовують свої зовнішньополітичні стратегії у відповідь на геополітичні амбіції Російської Федерації, докладно проаналізовано у працях Тараса Кузьо (Kuzio, 2000a; Kuzio, 2000b). Він спирається на типологію зовнішньої політики в різних країнах СНД. Використовуючи дихотомії «прагматизм-радикалізм» і «західництво-слов'янофільство», Т. Кузьо виокремлює чотири зовнішньополітичні орієнтації — радикальне західництво (членство в НАТО і ЄС), прозахідну нейтральність (двосторонні відносини в рамках СНД з акцентом на економічній кооперації), просхідну нейтральність (економіко-політична інтеграція в рамках СНД), Російсько-Білоруський союз (економічний, політичний, військовий). До першої групи належать країни Балтії, до другої — ГУАМ¹¹⁸ (Грузія, Україна, Узбекистан, Азербайджан, Молдова), до третьої — Росія, Казахстан, Киргизстан, Таджикистан, Вірменія, до четвертої — Білорусь¹¹⁹.

Головна мета зазначених праць — продемонструвати роль ГУАМ у створенні того, що З. Бжезінський назвав геополітичним плюралізмом. Так, завдяки створенню ГУАМ країни, що входили до нього, мали отримати більше можливостей для того, щоб говорити з Росією на рівних. Це, своєю чергою, мало сприяти подальшій демократизації в регіоні.

¹¹⁸ Тут і далі використано абревіатуру ГУАМ, а не ГУУАМ, оскільки Узбекистан покинув організацію 2005 року.

¹¹⁹ Як показав подальший розвиток подій, така зовнішньополітична орієнтація Білорусі була значною мірою лише піаром для її президента.

Далі Т. Кузьо звертається до теми ідентичності як фактора побудови національної держави в Україні (Kuzio, 2001). При цьому він акцентує увагу на активності українських політичних еліт і тому, як вони визначають «інших» у зовнішньополітичному просторі. Автор підкреслює той факт, що історична і культурна близькість Росії значно актуалізує необхідність осмислення нашої несхожості (тобто правильного розуміння «інших»). У протилежному разі Україна ризикує своєю національною ідентичністю.

В українському політикумі Т. Кузьо виокремлює чотири впливові групи — праві націоналісти, правоцентристи, центристи і ліві. Праві націоналісти бачать усіх росіян (і всередині, і за межами країни) як «інших», натомість ліві бачать «іншими» країни Заходу. Праві центристи підтримують громадянську модель національної держави і, відповідно, включають в неї тих росіян, які є громадянами України. При цьому правоцентристи солідарні з правими націоналістами щодо негативної оцінки імперського і радянського спадку, а також віднесення до «інших» росіян за межами України. Центристи, навпаки, знаходять у радянській історії як позитивні, так і негативні моменти, але поділяють спільну з правоцентристами думку про росіян як «своїх» або «чужих» залежно від їхнього громадянського статусу.

Також Кузьо вказує, що культурна і мовна близькість слов'янських етносів (українців, білорусів і росіян) робить питання національних ідентичностей більш, а не менш конфліктогенним, оскільки впливовіші держави можуть робити спроби нав'язати менш упливовим своє розуміння цієї проблеми.

Тему національної ідентичності автор доповнює у своїй пізнішій праці, присвяченій школам історії України. До таких він відносить радянську, русофільську, українофільську і східнослов'янську (Kuzio, 2006). Значущість цієї теми пов'язана з тим, що спосіб конструювання національної історії має суттєвий вплив на специфіку оформлення образу «інших», а отже, робить її важливим інструментом не тільки оформлення ідентичності, а й державної політики (внутрішньої та зовнішньої). Т. Кузьо доходить висновку про домінування українофільською школи, що забезпечує добру історичну базу для розуміння себе як європейської нації, на протизагу Росії, що має географічне положення поза межами Європи.

Мовний аспект формування національної ідентичності в Україні проаналізовано Тетяною Журженко (Zhurzhenko, 2002). Оцінюючи позицію впливових геополітичних сусідів (Європейський Союз і Російська Федерація) з мовних питань, вона робить висновки про перспективність орієнтації на побудову громадянської національної держави і, відповідно, розв'язання мовних проблем у контексті мультикультуралізму. Авторка вважає, що не дивлячись на фактичний білінгвізм в Україні (за винятком західного регіону) для більшості російськомовних громадян (як росіян, так і українців) питання мови не є пріоритетною політичною проблемою. Це, своєю чергою, суттєво послаблює можливості Росії щодо впровадження мовного «розколу» в політичне життя української держави.

Схожі думки висловлює Віктор Степаненко (Stepanenko, 2003). Зокрема, він говорить про те, що висока схожість двох основних мов в Україні потенційно

веде до можливості загальнонаціонального діалогу і взаємної толерантності, що дозволяють уникнути міжетнічних конфліктів. При цьому він вважає, що вибір між моделлю етнічної та громадянської національної держави є відкритим завданням, яке чекає свого розв'язання з боку політичних еліт України.

Напередодні президентських виборів 2004 року і подальшої Помаранчевої революції Т. Кузьо звернувся до теми періодизації відносин України з Заходом (Kuzio, 2003). Він виокремив три етапи — відсутність інтересу з боку Заходу¹²⁰, співпраця (1995—1999) і розчарування (з 2000). Ключова частина статті присвячена стурбованості Заходу з приводу президентства Л. Кучми, а також результатів майбутніх виборів, у яких, як вважає автор, чесну перемогу міг здобути тільки В. Ющенко (при цьому автор посилається на стабільно високий рейтинг зазначеного кандидата). Слід також указати на важливе зауваження автора статті, згідно з яким характер відносин Заходу з Україною вибудовується в контексті характеру його відносин із Російською Федерацією.

Тему внутрішніх зусиль українського політикуму на регіональному рівні, спрямованих на зближення з Європейським Союзом, розглядає Т. Журженко (Zhurzhenko, 2004a; Zhurzhenko, 2004b). Вона вказує на особливий геополітичний статус регіонів, що межують із Росією (як приклад використовує Слобожанщину). Головний привід звернення саме до них — наявність транскордонної кооперації, що має історичне коріння. Якщо до статусу кордонів України еліти Російської Федерації ставляться нейтрально, то прикордонні території вони оцінюють як такі, що перебувають у сфері їхнього інтересу. Отже, загальна геополітична свідомість у таких регіонах має подвійну важливість — для оформлення європейського майбутнього України необхідним є зсув акцентів від транскордонної кооперації з Росією до усвідомлення факту належності до європейського простору як його прикордонної зони.

§ 15.2. 2004–2013: між Помаранчевою революцією і Революцією гідності

Помаранчева революція стала новою точкою відліку для України щодо її геополітичних перспектив. Докладні результати аналізу відповідних вихідних умов подано у звіті Дослідницької служби конгресу США під авторством Стівена Вохрела (Woehrel, 2005). На 15 сторінках компактно викладено основні геополітичні аспекти внутрішнього і зовнішнього життя України. Розглянуто перспективи тристоронніх відносин Євросоюзу, України і Російської Федерації з урахуванням політичних цілей нового Президента — Віктора Ющенка. Увагу приділено як новим взаємовигідним можливостям, так і точкам напруги у відносинах між зазначеними сторонами. Окремо сказано про внутрішню складність України — мовну дилему, різні зовнішньополітичні устремління

¹²⁰ Такий стан справ був продиктований двома факторами — по-перше, першочерговою увагою до побудови відносин із ліберальною на той момент Російською Федерацією, по-друге, незрозумілим геополітичним статусом України, яка мала величезну армію і була членом клубу ядерних держав.

громадян, усе ще низький рівень підтримки ідеї вступу до НАТО. Плани про співпрацю з Північноатлантичним Альянсом проаналізовано і на рівні діяльності державних інститутів. Розглянуто нові економічні можливості України завдяки вступу до СОТ, а також інструменти розширення енергетичної незалежності в рамках діяльності ГУАМ. Зазначено про можливі дії щодо підтримки з боку США. Уже наступного року С. Вохрел у співавторстві з Джимом Ніколом (Nichol & Woehrel, 2006) подали звіт, присвячений газовому конфлікту між Україною і Росією, що почався після підвищення цін на енергоносії з боку Газпрому.

Геополітичні орієнтації українського політикуму проаналізовані В'ячеславом Нікуліним і Наталією Селютіною (Nikulin & Selyutina, 2007). У фокус їхнього інтересу потрапила активність політичних еліт у рамках перших парламентських виборів після Помаранчевої революції (передвиборчі твердження, виступи і публікації). Результати аналізу показали, що незважаючи на загальну прагматичну орієнтацію, яка полягає в розумінні важливості багатовекторності зовнішньополітичної співпраці, загальна картина є досить неоднорідною. Так, представники Комуністичної партії України демонстрували найвиразнішу прохідну позицію, Партії регіонів і Соціалістичної партії України — помірно прохідну, Блоку Юлії Тимошенко — прозахідну, Народного союзу «Наша Україна» — найбільш прозахідну.

Перехід від позиції, в якій Україна балансує між сходом і заходом, до проголошення нового цивілізаційного вибору України на користь Європи за президента В. Ющенка зафіксований у праці Рілке Драгневої та Антонети Дімітрової (Dragneva & Dimitrova, 2007). На думку авторів, це було вдалим рішенням, необхідним для початку модернізації всередині країни, спрямованої на подальшу європейську інтеграцію. Крім цього, Р. Драгнева й А. Дімітрова фіксують конфліктні та неконфліктні аспекти зовнішньоекономічної діяльності України, коли йдеться про одночасну співпрацю в рамках Європейської економічної зони, з одного боку, і СНД — з іншого.

Проте в пізніших працях (Malek 2009; Natton, 2010) тему цивілізаційного вибору змінює аналіз внутрішньополітичного фіаско в Україні (мається на увазі протистояння В. Ющенка і Ю. Тимошенко), яке перекреслило надії як на реальне зближення з Європейським Союзом, так і на вступ до НАТО. Це залишило Україну в «підвішеному» стані між Заходом і Сходом, а також додало можливостей Росії впливати на ситуацію в Україні.

Уже після президентських виборів 2010 року повернення геополітики в Центральній Європі у прив'язці до Росії та України проголосив експерт корпорації RAND Стефан Лерабі (Larrabee, 2010). За логікою, заданою З. Бжезинським, С. Лерабі називає Україну ключовою з точки зору геополітики країною, яка відкриває Росії шлях до статусу Євразійської імперії.

У цьому контексті Помаранчеву революцію розглядають як важливий поворотний момент, що негативно позначився на ймовірності реалізації імперських амбіцій Російської Федерації. Разом з тим автор добре обізнаний із внутрішнім політичним конфліктом між прем'єром Ю. Тимошенко і третім президентом України — В. Ющенком, який потяг за собою перемогу на прези-

дентських виборах 2010 року В. Януковича. Останнього С. Лерабі оцінює як помірно проросійського політика, стримуваного внутрішньою політичною неоднорідністю Партії регіонів (йдеться перш за все про олігархів). Далі увагу приділено чотирьом ключовим аспектам геополітичних взаємин між Україною і Росією — енергетичним потокам, Чорноморському флоту, статусу Севастополя і кримському питанню, НАТО і Європейському Союзу.

Енергетичне питання, що вперше голосно заявило про себе 2006 року, повторилося 2009-го. Факторами його виникнення можна назвати економічний тиск Росії на Україну, безальтернативність маршруту постачання газу в Європу на той період, а також репутаційні втрати Газпрому.

Дислокацію Чорноморського флоту Російської Федерації автор оцінює як важливий геополітичний інструмент Росії. Зокрема, він звертається до російсько-грузинського конфлікту, у процесі якого Чорноморський флот попрямував до берегів Грузії для ведення військових дій. Подальші протести керівництва України щодо його повернення в Севастополь не дали жодних результатів.

Щодо Севастополя, то автором підкреслено його історико-символічний статус російського міста, до якого звертаються окремі російські політики. Загалом С. Лерабі оцінює Севастополь як фактор можливих гострих конфліктів між Україною і Росією. Схожий потенціал має і вся Автономна Республіка Крим, яка вже ставала плацдармом розвитку сепаратистських настроїв в Україні. Здебільшого вони були інспіровані, в тому числі ззовні, і є додатковим інструментом тиску Росії на Україну.

Нарешті, зближення України з Європейським Союзом і НАТО останніми роками також створює потенціал напруженості між двома країнами. Автор указує, що Україна перебуває в ситуації відсутності чітких перспектив як щодо одного, так і щодо іншого напрямку співпраці. Також висловлено думку, що військова інтервенція з боку Росії в тому контексті, який склався, є малоімовірною, а ось спроби продемонструвати економічну і політичну неспроможність України перед західними партнерами є ймовірнішим варіантом розвитку подій. Важливо відзначити, що наприкінці геополітичного аналізу стану України С. Лерабі говорить про головну внутрішню слабкість України, яка може бути використана Росією, — розділеність українського суспільства.

Докладний аналіз тристоронніх відносин у сфері енергетики між Росією, Україною і Європейським Союзом у цей період подано у статті Олени Кропатчевої (Kropatcheva, 2011). Вона називає їх геополітичними енергетичними іграми, в які включено безліч учасників як на рівні державних інститутів, так і на рівні бізнес-структур. О. Кропатчева говорить, що залежно від обставин можна спостерігати тимчасовий союз будь-яких двох сторін цього політико-економічного трикутника на шкоду інтересам третьої. І незважаючи на високі витрати для всіх учасників цієї гри, остання має перманентний характер. Для Росії та Європейського Союзу пріоритетним напрямком є диверсифікація відповідно експорту й імпорту енергоресурсів, а для України — збереження статусу транзитної держави. При цьому в умовах відомої інтенсивності партнерських відносин між Росією і Європейським Союзом в енергетичній сфері по-

літико-економічна інтеграція з Україною (що відкриває непрямий або прямий контроль над її газотранспортною системою) є як для Російської Федерації, так і для Європейського Союзу важливим результатом у грі з нульовою сумою.

Цікавим є погляд автора на внутрішні особливості сторін цієї геоенергетичної гри. Так, у Росії державні інститути і бізнес-структури надзвичайно взаємопов'язані, Україна постає у вигляді «чорного ящика» через високу конфліктність її внутрішньополітичного життя (йдеться про ключові сили, які ведуть боротьбу за владу), Європейський Союз (і на рівні інститутів, і на рівні країн-учасниць) також можна назвати неоднорідним усередині, але вже в контексті визначення пріоритетності тих чи тих диверсифікаційних проектів.

У рамки такої картини добре вкладається оцінка автором президентських виборів 2010 року. Вона свідчить про те, що на відміну від 2004 року ні Росія, ні Європейський Союз не мали «свого» кандидата, оскільки кожен із претендентів на президентське крісло посідав приблизно однакову позицію щодо розв'язання проблем у сфері зовнішньої політики. У результаті це призвело до послаблення уваги Європи до України, що серед іншого вилалося в ухвалення суперечливих Харківських угод, згідно з якими Україна отримала 30% знижки на російський газ, а натомість зобов'язалася пролонгувати можливість перебування Чорноморського флоту у Криму на період від 25 до 30 років (ані Європейський Союз, ані США не зробили критичних зауважень із цього приводу). З іншого боку, уряд нового президента взяв курс на євроінтеграцію, в тому числі щодо планів із модернізації газотранспортної системи. Отже, головним принципом зовнішньої політики України за нового президента знову стала багатовекторність.

Напередодні найбільшої кризи в історії незалежної України Аттіла Віраг звертається до культурного і геополітичного вимірів будівництва української нації (Virbg, 2012). Він указує на те, що Україна належить до тих молодих країн, у яких демократизація і побудова нації почалися приблизно одночасно. Як і багато інших аналітиків, він вважає, що цей процес обтяжує наявність різних ідентичностей усередині країни, які суперечать одна одній. Більше за те, такий ідейній суперечності сприяють політичні еліти, які розглядають конкретні ідентичності як інструмент досягнення своїх цілей.

Незважаючи на можливість цивілізованого розв'язання проблеми ідейного поділу всередині країни лінією «схід-захід» у принципі, А. Віраг вважає, що в найближчій перспективі внутрішньополітична поляризація лише наростатиме. Це, своєю чергою, негативно позначиться на формулюванні національних цілей та інтересів, оскільки для цього необхідна несуперечлива національна ідентичність, натомість в Україні переважають інтереси різних кланових груп.

У результаті відмітними особливостями національного будівництва в Україні А. Віраг бачить такі: 1) існування двох різновидів історичної пам'яті (східна і західна), які негативно впливають на національне будівництво; 2) зацікавленість владних еліт у закріпленні такої розділеності; 3) відсутність національної концепції, спрямованої на досягнення національного примирення; 4) слабка оформленість міфів, що лежать в основі різних ідентичностей (що загалом пом'якшує конфлікти на ґрунті історичних забобонів).

Вельми цікавими є прогнози автора. Труднощі економічного характеру, на його погляд, витіснятимуть конфлікти на культурному ґрунті. При цьому, незважаючи на нерозв'язаність проблем, пов'язаних із національною ідентичністю, для переходу відповідних конфліктів у гарячу фазу (на кшталт громадянської війни) лише загроз усередині країни недостатньо.

Подальший сплеск інтересу до українського питання в контексті геополітики пов'язаний із конфліктом між Росією і Україною, що почався після перемоги Революції гідності 2014 року.

§ 15.3. 2013 — сьогоднішні дні: європейський вибір і військовий конфлікт

Девід Кадьє звертається до української кризи в контексті протиборства Європейського Союзу і Росії за геополітичний і економічний вплив у країнах, які входять у зону так званого спільного сусідства (Cadier, 2014). У числі цих країн Україна має ключове значення.

З боку Європейського Союзу впроваджено ініціативу «Політика європейського сусідства» (European Neighbourhood Policy), головним мотивом запуску якої були геополітичні міркування — стабілізація європейської периферії, а також конфігурування зовнішньої політики ЄС після розширення. У центрі такої політики — прагнення до стабільності, якої, втім, можна досягти і без членства в ЄС. Окрему позицію посідають країни Центральної Європи, які розглядають політику європейського сусідства, першою чергою, як інструмент ослаблення впливу Росії в регіоні.

Оскільки Росія бачила в такій зовнішньополітичній активності ЄС загрозу своїм інтересам (передусім політико-військовим, що стало очевидним після раптової анексії Криму), вона запропонувала альтернативну геополітичну ініціативу — «Євразійський митний союз», яку запущено 2010 року, анонсовано також «Євразійський економічний союз» (старт було намічено на 2015 рік). Спочатку акцент робили не на політичній чи історичній доцільності, а на перспективі модернізації та економічних вигодах для регіону загалом.

Д. Кадьє указує, що традиційний розгляд цих альтернативних схем регіоналізації як ключових факторів геополітичної орієнтації України є надто спрощеним. Зрозуміти їх протиборство можна лише з огляду на такий опосередковувачий комплекс змінних, як економічні та політичні інтереси еліт усередині самої України. У цьому сенсі рішення В. Януковича про вихід із процесу укладення асоціації з Європейським Союзом було продиктоване насамперед його оцінкою економічної ситуації та перспектив переобрання, а не ступенем лояльності до Росії чи Європейського Союзу.

Підбиваючи підсумки, автор доходить висновку про те, що заходи жорсткої політики з боку Росії вказують на її програв у геополітичному протистоянні альтернативних схем регіоналізації.

Радикальніше про ситуацію в Україні висловлюється Реін Мюллерсон, наживаючи нашу країну жертвою геополітики (Müllerson, 2014). Він зазначає, що операція НАТО в Косово, вторгнення в Ірак, військовий конфлікт між Росією

і Грузією, а тепер і між Росією й Україною демонструють міжнародне право і мораль як геополітичні інструменти. Автор вважає, що якщо під час холодної війни йшлося про протистояння капіталізму і комунізму, то тепер — про протистояння однополярного і багатополярного бачення світу. Причому це суперництво не ідеологічне, а геополітичне, в якому ідеологію, мораль і закон використовують як засоби боротьби за конструювання майбутньої дійсності. Мюллерсон доходить висновку, що внаслідок високої складності сьогоденного світу жодна з моделей (іудео-християнська, англосаксонська, конфуціанська, мусульманська або навіть світська ліберально-демократична) не може стати домінуючою. Більше за те, їх боротьба веде лише до нескінченних конфліктів.

До оцінки конкуруючого пояснення у вигляді концепції зіткнення цивілізацій С. Хантингтона звертається Борис Барканов (Barkanov, 2015). Незважаючи на її початкову привабливість, автор гадає, що релігійна картина в Україні не підтверджує цю концепцію, оскільки розкол між уніатами і православними значно менш важливий (особливо в контексті подій Євромайдану), ніж кластери таких пояснювальних факторів, як мовні практики, національність і геополітичні вподобання. А це, своєю чергою, веде до необхідності визнання зіткнення не цивілізацій, а геополітичних гравців (у цьому разі Росії та Заходу), що звертаються в Україні до різних геополітичних ідентичностей.

На відміну від попередніх прецедентів протистояння, для описуваного тут характерні більша гострота і небезпека (анексія Криму, провал на міжнародних зустрічах навесні 2014 року деескалації конфлікту, що посилювався між Росією й Україною, подальший збройний конфлікт на сході України, що, в тому числі, спричинив трагедію з малайзійським Боїнгом, безпрецедентні економічні санкції щодо Росії). Отже, те, що було конкуренцією геополітичних проектів регіоналізації, перетворилося на дуже напружену дилему регіональної безпеки.

Оцінку ролі дій Російської Федерації в демократизації України і Грузії надано у праці Лаури Делкур і Катаріни Волчук (Delcour & Wolczuk, 2015). Стаття розгортається навколо питання, як слід оцінювати цю роль — як бар'єр чи, навпаки, як підсилювача демократизації (проявами якої виступають три демократичні революції у Грузії та Україні).

Авторки схиляються до думки, що незважаючи на переважання внутрішніх причин демократизації (тиск мас і суперництво між національними елітами) над зовнішніми (зусилля Європейського Союзу і США), Росія проінтерпретувала відповідні процеси як результат утручання Заходу, на який вона має дати симетричну відповідь.

З погляду Л. Делкур і К. Волчук, анексія Криму і військова агресія на сході України, що відбулася слідом за нею, а також наслідки, які вони спричинили як на міжнародному, так і на рівні окремих держав, дозволяють зробити чотири висновки.

По-перше, геополітичні інтереси авторитарних сил, а не страх перед поширенням демократизації, є головною причиною їх боротьби за вплив у країнах, які вони включили у сферу своїх інтересів. Так, саме посилення зв'язків між

Україною і Грузією, а також зближення цих країн із західними структурами призвели до такої відповіді з боку Росії.

По-друге, навіть коли просування і підтримка демократії з боку західних партнерів є слабкою та ілюзорною у країнах, що перебувають у полі інтересів Російської Федерації, остання оцінює такі дії як неприйнятні. Тому події, які свідчать про демократичний прогрес (хоч би якими незначними вони були насправді), можуть бути інтерпретовані як утручання Заходу і призводити до протидії з боку Росії.

Третій висновок стосується механізмів, використовуваних авторитарними гравцями для протидії просуванню демократії. Авторитарні сили протидіють демократизації посиленням авторитарних груп у сусідніх країнах. Підтримка відповідних еліт за допомоги політичних, економічних, а також засобів безпеки, є найпростішим і найефективнішим способом просування своїх інтересів. Утім, за зміни авторитарних еліт у цих країнах на прозахідні, а отже, подальшої неможливості застосування відповідної стратегії, в хід іде широкий спектр інструментів — від економічного і політичного тиску до примусу і військової сили, які використовують на власний розсуд. Готовність діяти таким чином стала несподіванкою для груп, що просувають демократію.

Нарешті, ефекти від дії авторитарних сил щодо успіхів демократизації в інших країнах заслуговують на пильну увагу. Так, підриваючи чужу державність і порушуючи територіальну цілісність, авторитарні сили насправді домагаються зворотного — штовхають відповідні країни в бік Заходу, тобто сприяють подальшій демократизації.

Майкл Джентайл є одним із небагатьох авторів, які звертаються не до макрорівня геополітичних процесів, а до геополітичних орієнтацій населення, в нашому випадку — на сході України (Gentile, 2015). Як емпіричну базу він використовує результати опитування в Луганську ($N = 4000$), отримані наприкінці 2013 року. Дані збирали вже у процесі політичної кризи, пов'язаної з відмовою В. Януковича підписати асоціацію з Європейським Союзом.

Для фіксації геополітичних орієнтацій поставили запитання: «На Ваш погляд, чи було б добре Україні стати членом НАТО і ЄС?». Відповіді: 1) Так, і НАТО, і ЄС; 2) Тільки НАТО; 3) Тільки ЄС; 4) Ні; 5) Не знаю.

Далі результати перекодували. Перші три альтернативи розглядали як західні геополітичні орієнтації, четверту — як східну (проросійську/радянську), п'яту — як проміжну категорію. Автор вважає, що в основі варіанта відповіді «не знаю» може лежати цілий набір детермінант — погана поінформованість, складність у вираженні своєї позиції за доброї поінформованості, бажання піти від відповіді на сензитивне питання.

Як незалежні змінні використано набір демографічних (вік і стать), соціо-економічних (рівень освіти респондента, наявність вищої освіти у батьків респондента, статус зайнятості), соціокультурних (самооцінка етнічної належності, навички володіння англійською мовою, установки щодо гомосексуалізму, загальна задоволеність життям) і ситуаційних індикаторів (орієнтовне місце проживання в місті). Оскільки залежна змінна є номінальною, для аналізу застосовано множинну логістичну регресію.

Результати аналізу показали, що прозахідні геополітичні орієнтації характерні більшою мірою для молодших, не росіян (але інших етнічних меншин), добре освічених, толерантніших щодо секс-меншин, задоволениших життям і таких, що краще володіють англійською, респондентів. Відповідно, респонденти з просхідними установками мають протилежні показники за вказаними параметрами. М. Джентайл зазначає, що Луганськ — це ще не весь схід України, а схід України — не Україна загалом. Відповідно, до отриманих результатів треба підходити з певною мірою обережності.

Ще одне дослідження — на підставі крос-секційного дизайну — подане працею Джона Олокліна, Джерарда Тоала і Володимира Колосова, які звертаються до такого «інформаційно-ідеологічного феномену», як «Новоросія» (O'Loughlin, Toal, & Kolosov, 2017), у контексті його активного використання російськими ЗМІ 2014 року. Головна мета дослідження — визначити, чи є він тією підставою, на якій можна вибудувати геополітичний сепаратизм в Україні. Для її досягнення було проведене репрезентативне опитування (N = 2033) в шести областях України: Харківській, Дніпропетровській, Одеській, Миколаївській, Херсонській і Запорізькій (опитування здійснювалося у грудні 2014 року). Анкета містила близько 140 запитань, які можна поділити на три широкі групи: демографічні питання, питання з геополітичної тематики, а також питання про установки щодо специфічних тем локального характеру.

Щодо такого інформаційно-ідеологічного явища, як «Новоросія», поставлено три запитання: 1) «“Новоросія” — це міф чи історичний факт?»; 2) «Чи може це утворення бути використане як підстава для відокремлення його від України?» (запитання поставлено тільки тим респондентам, які не обрали варіант відповіді «міф»); 3) «Чи є використання терміна «Новоросія» (а) російською політичною технологією щодо знищення України або (б) виразом боротьби жителів південно-східної України за незалежність, або (в) Ви не знаєте, або г) Ви відмовляєтеся відповідати».

На підставі проаналізованих даних автори доходять висновку, що в розглянутих областях «політичний проект “Новоросія”» має мінімальну підтримку. Майже половина опитаних вважала, що цей термін є «історичним міфом». Відповідно, його актуалізація — результат російської політтехнології. Як підставу для відокремлення це явище розглядають старші і менш забезпечені респонденти. Суттєвим фактором підтримки відповідних ідей є також почуття ностальгії за Радянським Союзом. При цьому автори вважають, що недооцінювати загальну поширеність підтримки цього інформаційно-ідеологічного явища на сході України все ж таки не варто. Як альтернативне пояснення його низької підтримки в результатах опитування вони називають сензитивність відповідних запитань, а також високий рівень відповідей із використанням варіанта «не знаю».

§ 15.4. Незалежна Україна в геополітичному дискурсі: ключові теми

Наведений вище матеріал дає узагальнену картину щодо того, яку роль відведено Україні в геополітичних уявленнях західних експертів. Перш за все вона є

важливим елементом, що визначає співвідношення сил у Східноєвропейському регіоні, що пов'язане з транзитом енергетичних ресурсів, а також її роллю у сфері економічних і військово-стратегічних переваг. Усе це неодноразово стало предметом геополітичних конфліктів у регіоні.

По-друге, Україна раз по раз опинялася перед дихотомічним вибором між західним і східним шляхом розвитку. Навіть коли озвучують альтернативні варіанти, які можуть мати кращий потенціал з точки зору інтересів України (наприклад, ГУАМ), їх оцінюють у контексті руху за вказаним вектором. Таке аналітичне фокусування свідчить, перш за все, про першочергову увагу до інтересів політичних гравців на заході і сході, а вже потім — самої України.

По-третє, загальна внутрішня ситуація в Україні, релевантна з точки зору геополітичних можливостей, може бути поділена на два рівні. Базисом тут виступає так званий чорний ящик української політики, в якому політичні еліти країни розв'язують свої перманентні конфлікти. Надбудовою є весь спектр економічних та інституційних проблем, породжуваних результатами конфліктної взаємодії як політичних еліт України, так і нестабільністю зовнішньополітичного середовища (щодо останнього буде справедливим використання ще однієї метафори — «туман війни»).

Нарешті, недостатньо уваги приділено геополітичним уподобанням населення України. А в тих працях, де про це йдеться, її знову ж таки зосереджено на дихотомії «схід-захід». Останнє є поширенням логіки вимушеного (реактивного) геополітичного вибору з рівня міжнародних відносин на рівень повсякденного життя українських громадян. Велику увагу в цьому контексті приділено проблемі їх ідентичності як фундаменту побудови тих чи тих геополітичних проєктів.

Але незважаючи на епізодичну увагу до самих геополітичних орієнтацій, у працях указують важливі з точки зору геополітичних установок особистісні змінні. Крім уже згаданих особливостей національної ідентичності, до них також належать оцінка національної історії та особливостей історичної пам'яті, довіра до тих або тих міжнародних ЗМІ, оцінка мовного питання в Україні, терпимість/нетерпимість до ідейних опонентів усередині країни, оцінка шляхів розвитку національної держави, оцінка російсько-українських енергетичних конфліктів, ціннісні мотиви, ставлення до проявів сепаратизму, оцінка моделей регіонального співробітництва, особливості співвіднесення діяльності великих геополітичних гравців з інтересами України, а також заохочення/осуд їхнього втручання в її внутрішні справи, розуміння різних наслідків того або того геополітичного вибору (зокрема й у військовій сфері). Аналіз релевантності вказаних аспектів є важливим етапом з точки зору підготовки дослідницької стратегії вивчення геополітичних орієнтацій населення країни.

На окрему увагу з погляду вивчення геополітичних орієнтацій заслуговує праця Т. Кузьо (Kuzio, 2000a). Незважаючи на те, що автор зупиняється на дихотомії «схід-захід», він упроваджує другу, принципово важливу, на нашу думку, дихотомію — «радикалізм-прагматизм». Саме геополітичний прагматизм рухає тими гравцями міжнародної взаємодії, які відіграють активну роль у досягненні своїх національних інтересів. Своєю чергою, східний і західний век-

тори розвитку можуть (і повинні) бути розширені за рахунок альтернативних шляхів геополітичного співробітництва. Коротко розглянемо лише деякі з них (ГУАМ, Польща, Туреччина, багатовекторність).

ГУАМ насправді є частиною західної геополітичної стратегії розвитку для Грузії, України, Азербайджану і Молдови. Більшість країн, що входять до нього, були об'єднані спільною проблемою — територіальними конфліктами в межах своїх кордонів (Придністровська Молдавська Республіка, Республіка Абхазія, Республіка Південна Осетія, Нагірно-Карабахська Республіка). Хоч як це парадоксально, але з 2014 року цей «клуб» поповнила й Україна (Крим, Донбас). На думку Т. Кузьо, ГУАМ — це прояв прагматичного інтересу країн, що його утворили. Потенційні вигоди такої співпраці, звичайно ж, не обмежуються політичними і територіальними питаннями, вони охоплюють і економічний вимір.

Роль польсько-українських відносин важко переоцінити у світлі трудової еміграції українців до Польщі, що суттєво посилилася останнім часом. Для української сторони це важливе джерело доходів, для польської — істотний стимул економічного зростання. Логічним продовженням цього стихійно сформованого процесу є економічна співпраця на рівні країн, спрямована на захист прав та інтересів наших співвітчизників у Польщі. З огляду на специфіку описаної ситуації тут ідеться також про прагматичну спрямованість геополітичної активності. Це саме той тип західної інтеграції, який є природним і найпродуктивнішим. Натомість відсутність конструктивного реагування на ситуацію, що склалася, загрожує посиленням економічних проблем в Україні.

Говорити про турецький напрямок дозволяє хоча б те, що відповідні геополітичні орієнтації поширені серед кримськотатарського корінного народу, що проживає в Україні (як відомо, кількість кримських татар у Туреччині перевершує їхню кількість у Криму і на материковій Україні). Але навіть незалежно від цього турецький вектор у контексті геополітичного співробітництва можна розглядати як перспективний. По-перше, це продиктоване можливостями врівноваження сил у геополітичному трикутнику Україна — Туреччина — Росія. По-друге, цінною для України може бути співпраця в економічній сфері, зокрема, ознайомлення з турецьким досвідом розвитку національної економіки, яка характеризується великою кількістю галузей, вигідним географічним положенням і дешевою робочою силою (що, на наш погляд, характерно і для України).

Горезвісна багатовекторність, що в підсумку не найкраще прислужилася Україні, не може бути відкинута лише тому, що вищі керівники нашої держави не змогли вміло нею розпорядитися. З огляду на те, що в найближчому майбутньому Україна не стане членом Європейського Союзу, а також на відносини з Російською Федерацією, які суттєво погіршилися, рух вектором «схід-захід» є істотно обмеженим. При цьому відносини з обома з зазначених партнерів не припиняються і вимагають тих чи тих рішень у контексті зовнішньополітичної дійсності. Не варто скидати з рахунків і всю сукупність можливостей, пов'язаних з альтернативними напрямками (ГУАМ, Польща, Туреччина тощо). У цьому сенсі стратегію багатовекторності можна розглядати як ефек-

тивний інструмент розв'язання геополітичних завдань, що постають перед Україною (питання полягає лише в якості державного управління).

Отже, у стратегічній площині Україна може бути орієнтована на геополітичний радикалізм (постійний рух в одному напрямку незалежно від актуального контексту) або прагматизм (просування своїх національних інтересів за тими напрямками, які є актуальними з точки зору поточної геополітичної ситуації). Що стосується *найважливіших* напрямків зовнішньополітичної діяльності, то немає жодної причини обмежувати їх дихотомією «схід-захід».

Беручи до уваги багатоаспектність геополітичних установок, можливість їх емоційного і раціонального прояву, а також високу складність (векторність) зовнішньополітичного середовища України, спробуємо сформулювати деякі висновки про те, якою насправді має бути фіксація геополітичних уподобань учасників масових опитувань, які проводять в Україні.

§ 15.5. Фіксація геополітичних орієнтацій соціологічними засобами (замість висновку)

Перш за все треба сказати, що спілкування дослідника з респондентами (нехай і у форматі анкетного опитування) на таку тему, як геополітичні вподобання, передбачає вибір певного фокусу обговорення, який визначає результати цієї комунікації. Тут повною мірою виявляється те, що С. Садмен і Н. Бредберн називали контекстуальною зумовленістю відповідей (Садмен & Бредберн, 2002, с. 209–228). І цей контекст не тільки створюється послідовністю попередніх запитань (говорячи загалом, цей фактор може себе навіть не проявити), але і конструюється форматом і змістом самого запитання. Геополітична проблематика є досить складною темою, що має стосунок серед іншого до пам'яті та широти кругозору респондентів. Тому формат і зміст відповідних запитань задають жорсткі рамки когнітивної активності респондента. І зовсім не обов'язково, щоб вони були для нього природними і він з ними ознайомлений. Фактично дослідник змушує респондента мислити про геополітику так, як мислить він сам. А як показав наш аналіз, геополітична проблематика є тією сферою смислів, для якої характерна досить висока гетерогенність.

Отже, соціологічний інструментарій, націлений на геополітичні орієнтації респондентів, має, мірою можливості, торкатися всіх релевантних аспектів геополітичної дійсності України. Це дозволить якнайкраще активізувати пам'ять учасників опитування, а також надасть їм достатню кількість когнітивних інструментів для зваженої інтерпретації цієї комплексної проблеми. Не виконавши цього, дослідник надто багато чого «вносить за дужки», роблячи тим самим свою інтерпретацію отриманих результатів, по суті, волюнтаристською (Дембицкий, 2017d, с. 117).

Спираючись на розглянуті наукові праці, до таких релевантних аспектів ми віднесли:

- 1) ідентичність — різні аспекти національно-громадської ідентичності можуть справляти прямий вплив на геополітичні вподобання респонден-

- тів — це, зокрема, було продемонстровано в одному з наших досліджень (Дембіцький & Любива, 2017b);
- 2) особливості історичної пам'яті — інтерпретація історичних подій і персоналій може бути пов'язана з геополітичними установками, особливо коли ці події/персоналії є співзвучними з актуальними геополітичними процесами;
 - 3) оцінка мовного питання в Україні — ставлення до мовних проблем є показовим не тільки з точки зору спрямованості геополітичних орієнтацій, а і з погляду їх загальної тональності — радикальної або прагматичної;
 - 4) довіра до міжнародних ЗМІ — сприйняття інформації з різних зарубіжних інформаційних джерел як такої, що заслуговує на довіру, або, навпаки, є недостовірною, багато в чому пов'язане з геополітичними вподобаннями;
 - 5) ставлення до територіального устрою України — за уявленнями щодо бажаного територіального устрою країни можуть стояти ті чи ті геополітичні уподобання;
 - 6) оцінка російсько-українських енергетичних конфліктів — у контексті відмови від дихотомії «схід-захід» цей фактор має обмежене значення, водночас його значення важко переоцінити за аналізу ставлення респондентів до Російської Федерації; при цьому питання енергетичної безпеки України може (і повинно) бути винесено за рамки трикутника «Україна — Європейський Союз — Російська Федерація»;
 - 7) ціннісна структура — визнання прихильності до тих або тих наборів цінностей (наприклад, європейських, загальнослов'янських, етнонаціональних або прагматичних) саме по собі мало або практично нічого не свідчить про реальну ціннісну структуру індивіда, натомість така артикуляція може бути важливим виміром за аналізу специфіки зовнішньополітичних орієнтацій респондента;
 - 8) оцінка пріоритетних напрямків співробітництва в регіоні — в цьому разі важлива оцінка перспективності економічної співпраці з можливими геополітичними союзниками, а також думка про те, наскільки тісним воно має бути;
 - 9) особливості розуміння інтересів України в контексті діяльності провідних геополітичних сил, а також оцінка поведінки останніх, тобто оцінка поведінки різних країн як ворожої, дружньої або нейтральної стосовно України, є прямим показником геополітичних орієнтацій респондентів (зокрема оцінка прийнятності втручання в політичне життя України); додатковим параметром тут слід визнати те, якими респонденти бачать зовнішньополітичні рішення українських еліт, спрямовані на досягнення інтересів України;
 - 10) готовність/неготовність прийняти різні наслідки геополітичного вибору політичних еліт України — різні варіанти геополітичного розвитку припускають не тільки певні вигоди, але і витрати, пов'язані з ними; в цьому контексті важливо розуміти, які з можливих негативних наслідків респонденти готові прийняти в рамках того або того вибору.

Наведені аспекти можуть бути використані як вихідна точка для розробки соціологічного тесту, спрямованого на комплексне вимірювання геополітичних орієнтацій населення в Україні. При цьому їх структурування і змістовне оформлення має бути здійснене з опертям на два принципи — врахування радикалізму/прагматизму в судженнях і відмова від геополітичної дихотомії «схід-захід».

Глава 16

Вимірювання геополітичних орієнтацій в Україні: формулювання індикаторів і апробація в методичному опитуванні

Спираючись на зроблені раніше висновки, розглянемо проектні індикатори, спрямовані на вимір геополітичних орієнтацій і покликані забезпечити змістовну валідність соціологічних тестів, що конструюються. Далі щодо кожного аспекту геополітичних орієнтацій подано проектний варіант запитання (або кількох запитань, де це необхідно), після чого дано короткі коментарі.

§ 16.1. Змістовна валідність: індикатори геополітичних уподобань

1) Ідентичність.

Говорючи про різних сусідів України та їх культури, оцініть, наскільки близька кожна з наведених нижче культур особисто Вам (дайте відповідь за кожним пунктом).

- 1 — Не близька зовсім
- 2 — Певною мірою близька
- 3 — Дуже близька
- 4 — Важко сказати

1	Європейська культура	1	2	3	4
2	Мусульманська культура	1	2	3	4
3	Слов'янська культура	1	2	3	4

За вибору показників ідентичності ми звернулися до такого підходу, як «Я-концепція» (Майерс, 2003, с. 73–76) і, зокрема, до тієї його частини, яка стосується усвідомлення своєї належності до ширшої культурної спільноти. На наш погляд, це та частина ідентичності індивіда, яка впливає на оцінку ним геополітичних процесів.

2) Особливості історичної пам'яті.

Як Ви вважаєте, наскільки сильний вплив на українську культуру з точки зору історії вчинили наведені нижче держави? (дайте відповідь за кожним пунктом)

- 1 — Слабкий вплив
- 2 — Сильний вплив не деяких територіях
- 3 — Сильний вплив на більшості територій
- 4 — Важко відповісти

4	Річ Посполита (історична Польща)	1	2	3	4
5	Османська імперія (історична Туреччина)	1	2	3	4
6	Російська імперія	1	2	3	4

Яким, на Ваш погляд, був цей вплив: негативним чи позитивним? (дайте відповідь за кожним пунктом)

- 1 — Переважно негативним
- 2 — Частково негативним, частково позитивним
- 3 — Переважно позитивним
- 4 — Важко відповісти

7	Річ Посполита (історична Польща)	1	2	3	4
8	Османська імперія (історична Туреччина)	1	2	3	4
9	Російська імперія	1	2	3	4

10 Як Ви оцінюєте культурну спадщину радянського періоду в історії України? (оберіть один варіант відповіді, який найкраще підходить)

1. Переважно негативно.
2. Частково негативно, частково позитивно.
3. Переважно позитивно.
4. Важко сказати.

Визнання потужного культурного впливу, а також його позитивна оцінка свідчать, на наш погляд, про відповідні геополітичні вподобання. Тут дещо переважає увага до східного напрямку (Російська імперія, СРСР), що продиктовано виключно історичними особливостями України.

3) Оцінка мовного питання в Україні.

Як Ви вважаєте, чи варто державі підтримувати в Україні наведені далі мови? (дайте відповідь за кожним пунктом)

- 1 — Не варто підтримувати взагалі
- 2 — Варто підтримувати в окремих регіонах
- 3 — Варто підтримувати на всій території України

11	Російська	1	2	3
12	Польська, угорська, румунська (одна чи кілька)	1	2	3
13	Англійська	1	2	3

Ця оцінка спрямована більшою мірою на виявлення геополітичного радикалізму/прагматизму в установках респондентів. Варіант відповіді «не варто підтримувати взагалі» свідчить про радикалізм, інші два — про прагматизм.

4) Довіра до міжнародних ЗМІ.

Як Ви вважаєте, чи можна довіряти офіційним ЗМІ наведених далі країн? (дайте відповідь за кожним пунктом)

1 — Не можна довіряти більшості ЗМІ

2 — Можна довіряти більшості ЗМІ

3 — Важко сказати

14	Країн Європейського Союзу	1	2	3
15	Туреччини	1	2	3
16	Російської Федерації	1	2	3
17	США	1	2	3

Загалом це запитання продовжує логіку вивчення ставлення до наших найближчих геополітичних сусідів. Додатково впроваджено пункт, що стосується США, — як через впливовість їх ЗМІ в міжнародному масштабі, так і через актуальність зовнішньополітичних відносин між Україною і США на сучасному етапі.

5) Ставлення до територіального устрою України.

18 Як Ви вважаєте, який адміністративно-територіальний устрій найкращим чином відповідає успішній реалізації зовнішньої політики України? (оберіть один варіант відповіді, який найкраще підходить)

1. Унітарний (як зараз): усі рішення повинні прийматися з політичного центру.
2. З можливістю створення територіальних автономій, які можуть самостійно визначати свою зовнішню політику в межах конституційних обмежень.
3. Важко сказати.

Це запитання, імовірно, має вказувати, з одного боку, на геополітичний радикалізм (за вибору другого варіанта відповіді), а з іншого — на геополітичні симпатії високої інтенсивності (цей момент слід перевірити у зв'язку з відповідями на інші запитання).

6) Оцінка російсько-українських енергетичних конфліктів.

19 Як Ви вважаєте, хто є головним винуватцем конфліктів між Україною та Росією в енергетичній сфері (маються на увазі транзит газу через територію України та ціни на газ)? (оберіть один варіант відповіді, який найкраще підходить)

1. Як Росія, так і Україна.
2. Переважно Росія.
3. Переважно Україна.
4. Переважно Європейський Союз.
5. Важко сказати.

20 Як Ви вважаєте, що було б найкращим варіантом для України в сфері енергетичної безпеки? (оберіть один варіант відповіді, який найкраще підходить)

1. Покращити співробітництво з Росією.
2. Розвивати перспективні проекти з Європейським Союзом.
3. Розвивати співробітництво з Азербайджаном.
4. Важко сказати.

Енергетичне питання за своєю суттю є геополітичною дилемою для України. У цьому сенсі увага до нього відіграє важливу роль у процесі вивчення геополітичних орієнтацій українського населення. З варіантів відповідей для запитання 19 виключено пункт «Забезпечити Україну газом власного видобутку». По-перше, він дещо розмиває геополітичну тематику; по-друге, є очевидно привабливим; по-третє, складно сказати, наскільки такий сценарій насправді можна реалізувати.

7) Ціннісна структура.

21 Яких цінностей слід дотримуватися керівництву нашої країни в першу чергу, щоб досягти успіху в сфері міжнародної політики? (оберіть один варіант відповіді, який найкраще підходить)

1. Європейських (євроцентризм, мультикультуралізм, толерантність і т.д.).
2. Загальнослов'янських (слов'янська єдність, релігійність, шанування роду і т.д.).
3. Прагматичних (слідування поточним інтересам, нейтральність у політиці та культурі і т.д.).
4. Етнонаціональних (підтримка української культури, націоналізм, громадянство і т.д.).
5. Близького Сходу (слідування ісламу, дотримання соціальної ієрархії та сімейних традицій і т.д.).
6. Важко сказати.

Наскільки близька кожна з цих ціннісних груп особисто Вам? (дайте відповідь за кожним пунктом)

- 1 — Зовсім не близька
- 2 — Певною мірою близька
- 3 — Дуже близька
- 4 — Важко сказати

22	Європейські цінності	1	2	3	4
23	Загальнослов'янські цінності	1	2	3	4
24	Прагматичні цінності	1	2	3	4
25	Етнонаціональні цінності	1	2	3	4
26	Близькосхідні цінності	1	2	3	4

Тут, звичайно, не йдеться про вимірювання ціннісної структури як такої (для цього необхідний більш комплексний інструментарій), а лише про фіксацію додаткових параметрів зовнішньополітичних симпатій.

Крім того, ця група запитань (особливо 21–25) перегукуються з запитаннями про ідентичність (1–3). Рішення про те, яку з цих груп слід залишити (можливо, обидві), може бути ухвалене тільки на підставі результатів методичного дослідження.

8) Оцінка пріоритетних напрямків співробітництва в регіоні.

Будь ласка, оцініть пріоритетність кожного з наведених далі напрямків міжнародного співробітництва для України (дайте відповідь за кожним пунктом).

- 1 — Низький пріоритет, співробітництво необхідно скорочувати
- 2 — Ситуаційний пріоритет, інтенсивність співробітництва залежить від завдань, які вирішуються
- 3 — Високий пріоритет, співробітництво має стратегічно важливий характер
- 4 — Важко сказати

27	Європейський Союз загалом	1	2	3	4
28	Російська Федерація	1	2	3	4
29	Польща (як окремий стратегічний партнер)	1	2	3	4
30	Туреччина	1	2	3	4
31	Група таких країн, як Грузія, Азербайджан і Молдова (ГУАМ)	1	2	3	4
32	Група таких країн, як Польща, Чехія, Словаччина та Угорщина (Вішеградська група)	1	2	3	4
33	Група таких країн, як Естонія, Латвія і Литва (Балтійська асамблея)	1	2	3	4
34	США	1	2	3	4

Як Ви ставитесь до ідеї військового співробітництва за наступними напрямками? (дайте відповідь за кожним пунктом)

- 1 — Таке співробітництво не можна допускати
- 2 — Таке співробітництво є необхідним
- 3 — Важко сказати

35	НАТО	1	2	3
36	Російська Федерація	1	2	3
37	Група таких країн, як Грузія, Азербайджан і Молдова (ГУАМ)	1	2	3

Цей блок є одним із ключових за вимірювання геополітичних орієнтацій, оскільки спрямований на оцінку актуальності різних векторів зовнішньополітичної активності України. Методологічно він ґрунтується на відмові від дихотомії «схід–захід». Особливу увагу приділено військовій співпраці, важливість якої істотно загострилася для України останніми роками.

9) Особливості розуміння інтересів України в контексті діяльності провідних геополітичних сил, оцінка їх поведінки.

Оцініть політику наведених нижче країн або їх об'єднань щодо України на даному етапі? (дайте відповідь за кожним пунктом)

- 1 — Повністю корислива: інтереси України не враховуються
 2 — Частково корислива: інтереси України враховуються частково
 3 — Партнерська: враховується більшість інтересів України
 4 — Важко сказати

38	Європейського Союзу загалом	1	2	3	4
39	Російської Федерації	1	2	3	4
40	Польщі (як окремого стратегічного партнера)	1	2	3	4
41	Туреччини	1	2	3	4
42	Групи таких країн, як Грузія, Азербайджан і Молдова (ГУАМ)	1	2	3	4
43	Групи таких країн, як Польща, Чехія, Словаччина та Угорщина (Вишеградська група)	1	2	3	4
44	Групи таких країн, як Естонія, Латвія і Литва (Балтійська асамблея)	1	2	3	4
45	США	1	2	3	4

Як Ви ставитесь до того, щоб ці країни або їх об'єднання впливали на внутрішньополітичне життя України? (дайте відповідь за кожним пунктом)

- 1 — Негативно: це є руйнівним для України
 2 — Нейтрально: наслідки можуть бути як позитивними, так і негативними
 3 — Позитивно: це піде Україні на користь
 4 — Важко сказати

46	Європейський Союз загалом	1	2	3	4
47	Російська Федерація	1	2	3	4
48	Польща (як окремий стратегічний партнер)	1	2	3	4
49	Туреччина	1	2	3	4
50	Група таких країн, як Грузія, Азербайджан і Молдова (ГУАМ)	1	2	3	4
51	Група таких країн, як Польща, Чехія, Словаччина та Угорщина (Вишеградська група)	1	2	3	4
52	Група таких країн, як Естонія, Латвія і Литва (Балтійська асамблея)	1	2	3	4
53	США	1	2	3	4

54 Як Ви оцінюєте сьогоднішній курс України на євроінтеграцію? (оберіть один варіант відповіді, який найкраще підходить)

1. Дуже негативно, це суперечить інтересам України.
2. Негативно, оскільки це робиться неграмотно.
3. Нейтрально, плюси й мінуси врівноважують один одного.
4. Позитивно, це відповідає актуальним інтересам України.
5. Дуже позитивно, це важливе стратегічне завдання для України.

Цей блок запитань продовжує логіку попереднього. Показники 37–44 спрямовані на виявлення типових проявів геополітичної симпатії/антипатії, натомість показники 45–52 — посилені прояви симпатії (мається на увазі варіант відповіді «Позитивно — це піде Україні на користь»). Запитання щодо оцінки євроінтеграції України продиктоване поточною геополітичною дійсністю.

10) Готовність/неготовність прийняти різні наслідки геополітичного вибору політичних еліт в Україні.

Той чи інший зовнішньополітичний курс держави, що обумовлений як діями її керівництва, так і настроями населення, може призводити до різних позитивних і негативних результатів. Беручи до уваги сьогоднішній зовнішньополітичний курс України, з якими з наведених проблем і як довго Ви готові мати справу? (дайте відповідь за кожним пунктом)

- 0 — Не думаю, що така проблема зараз існує
- 1 — Не готов(-а), ні за яких обставин
- 2 — Готов(-а), якщо ця проблема триватиме не більше року
- 3 — Готов(-а), якщо ця проблема триватиме не більше трьох років
- 4 — Готов(-а), якщо ця проблема триватиме не більше п'яти років
- 5 — Готов(-а) стільки, скільки знадобиться

55	Падіння рівня життя населення (у тому числі й особисто Вашого), пов'язане із виконанням вимог Європейського Союзу про асоціацію з Україною	0	1	2	3	4	5
56	Обмеження політичної самостійності українських еліт в обмін на допомогу впливових закордонних партнерів	0	1	2	3	4	5
57	Військовий конфлікт на Сході України як наслідок напружених відносин між Україною і Росією	0	1	2	3	4	5

Ці питання спрямовані на оцінку загальної зваженості суджень респондентів щодо геополітичної проблематики, тобто вони оцінюють геополітичний радикалізм/прагматизм респондентів. Як заперечення наявних проблем, так і готовність миритися з ними стільки, скільки знадобиться, на нашу думку, може вказувати на геополітичний радикалізм. Своєю чергою, готовність мати

справу з проблемами певний термін має вказувати на геополітичний прагматизм. Статус варіанта відповіді «Не готовий(-а), за жодних обставин» необхідно перевіряти в контексті взаємозв'язків із відповідями на інші запитання.

Запропоновані запитання стосуються широкого спектра тем, теоретично пов'язаних із геополітичними орієнтаціями населення. Проте до проведення спеціального методичного дослідження — це лише припущення. Крім того, в контексті соціологічного тестування 57 індикаторів є надто об'ємним інструментарієм з точки зору масових опитувань, що висуває на перший план завдання його оптимального скорочення.

§ 16.2. Композиційна валідність — гіпотези

Перш ніж розпочати аналіз даних, зробимо спробу структурувати наявні індикатори за орієнтирами, покладеними в основу дослідження, — багатовекторністю геополітичних уподобань і дихотомією «радикалізм/прагматизм», з погляду якої зовнішньополітичним орієнтаціям можна дати додаткову оцінку. Після отримання емпіричного матеріалу така інтерпретація дозволить аргументовано відсіяти запитання, які не розв'язують покладених на них завдань, обрати найінформативніші індикатори, а також загалом оцінити якість запропонованого інструментарію. Крім того, спроба адекватного структурування індикаторів методики є необхідною умовою досягнення її композиційної валідності.

Спочатку визначимо запитання, які вказують на різні вектори геополітичних уподобань (табл. 16.1), а потім — ті, що вказують на їх радикалізм/прагматизм.

Таблиця 16.1

Вектори геополітичних уподобань у поданих питаннях

<i>Напрямок геополітичного співробітництва</i>	<i>№№ індикаторів</i>	<i>Загальна кількість індикаторів</i>
Європейський Союз загалом	1, 14, 19, 20, 21, 22, 27, 35, 38, 46, 54	11
Російська Федерація	3, 6+9, 10, 16, 19, 20, 21, 23, 28, 36, 39, 47	12
Польща	4+7, 12, 29, 40, 48	5
Туреччина	2, 5+8, 15, 21, 26, 30, 41, 49	8
Група таких країн, як Грузія, Азербайджан і Молдова	20, 31, 37, 42, 50	5
Група таких країн, як Польща, Чехія, Словаччина та Угорщина	32, 43, 51	3
Група таких країн, як Естонія, Латвія та Литва	33, 44, 52	3
США	17, 34, 45, 53	4

Незбалансованість кількості індикаторів для різних напрямків обумовлена історичними особливостями й актуальними геополітичними проблемами Ук-

раїни. Для розробки збалансованішого варіанта нам було необхідно зібрати і проаналізувати емпіричні дані. Також слід згадати таку геополітичну орієнтацію, як багатовекторність. Прямих запитань для її фіксації в наведеному вище інструментарії не подано. Проте її визначення не складає труднощів на підставі наявних запитань. Додатково до цього можна поставити сфокусоване на цій темі питання:

58 Як би Ви оцінили співробітництво України з декількома зовнішньополітичними партнерами, що передбачає багатовекторність? (оберіть один варіант відповіді, який найкраще підходить)

1. *Негативно, оскільки така стратегія веде до провалів у зовнішній політиці.*
2. *Нейтрально, оскільки така стратегія не має принципових переваг або недоліків для України в її сьогоденному положенні.*
3. *Позитивно, оскільки така стратегія дозволяє найбільш повно реалізувати переваги України у зовнішньополітичній діяльності.*
4. *Важко сказати.*

На наш погляд, це запитання може також бути індикатором геополітичного прагматизму. Що стосується інших запитань для оцінки радикалізму/прагматизму геополітичних установок, то до них ми віднесли індикатори № 11–13, 18, 24, 55–57.

Пропонований підхід потенційно веде до геополітичної n -векторності респондента, де n може набувати значення від 0 до 8. Нульове значення побічно свідчить про геополітичну закритість респондентів. Що стосується спеціального запитання щодо такої установки, воно може бути таким:

59 Як би Ви оцінили обмеження зовнішньополітичної діяльності керівництва України, тобто ситуацію, в якій максимальний акцент буде робитися на використанні внутрішніх ресурсів країни та мінімальний — на використанні ресурсів закордонних партнерів? (оберіть один варіант відповіді, який найкраще підходить)

1. *Негативно, в сучасному світі — це шлях до деградації.*
2. *Нейтрально, від цього положення України не погіршиться.*
3. *Позитивно, це дозволить Україні повноцінно розвиватися.*
4. *Важко сказати.*

Останній індикатор не тільки спрямований на виявлення тих, хто вірить, що в сучасному світі досягти успіху на державному рівні можна, спираючись лише на внутрішні ресурси і можливості, а й є додатковим індикатором геополітичного радикалізму у формі ізоляціонізму.

§ 16.3. Специфіка реалізації методичного дослідження

Дизайн. Для апробації поданого інструментарію у квітні 2018 року було проведено методичне опитування серед жителів Києва ($n_k = 34$), Одеси ($n_o = 33$) і Львова ($n_l = 33$). У кожному з міст вибірка побудована на підставі шести близьких за розміром квот, сформованих на базі статі (2 категорії) і віку (3 категорії). Середній вік респондентів — 42,9 року. Кількість жінок у вибірці становить 52%, чоловіків — 48%.

Інструмент. Окрім наведених вище запитань, в анкету включено чотири додаткові запитання з використанням категоріальних шкал для вимірювання таких змінних: 1) освіта респондентів (5 категорій); 2) ступінь їхньої зацікавленості політикою (3 категорії); 3) політична течія, якій вони надають переваги (12 категорій); 4) ступінь, на рівні якого респонденти пишаються громадянством України (5 категорій).

З огляду на побудову соціологічних тестів у формі вимірювальних шкал першочерговим завданням була факторна валідизація. Факторну валідність перевірено на підставі результатів конфірмаційного факторного аналізу (*Confirmatory Factor Analysis* — CFA) з використанням методу діагонально зважених найменших квадратів (*Diagonally Weighted Least Squares* — DWLS). Якість факторних моделей оцінено на підставі значень таких показників: χ^2 -квадрат; середньоквадратична похибка апроксимації (RMSEA); порівняльний індекс відповідності (CFI); індекс Такера — Левіна (TLI). Як прийнятні граничні значення було прийнято відношення значення χ^2 -квадрат до ступенів свободи (df) менше 5; $RMSEA < 0,07$; $(CFI, TLI) > 0,95$.

Оцінку внутрішньої узгодженості методики здійснено за допомоги коефіцієнта Альфа Кронбаха. Значення, що перевищують 0,7, розглядалися як прийнятні.

У рамках описового аналізу вибірки використано ієрархічний кластерний аналіз на підставі евклідової відстані з застосуванням методу Варда.

Для статистичного аналізу використано мову програмування R (версія 3.4.1, бібліотеки «lavaan» — для проведення CFA, «lrm» — для розрахунку значень Альфа Кронбаха, «psych» — для кореляційного аналізу).

§ 16.4. Складність відповідей на запитання анкети

Першочерговим завданням після отримання результатів опитування була оцінка змістовної складності індикаторів методики. Для цього в більшості індикаторів анкети було передбачено варіант «Важко відповісти». Тут ми розглянемо лише ті пункти, в яких варіант «Важко відповісти» був обраний респондентами велику кількість разів.

Перш за все треба сказати, що вираз «велика кількість разів» не має однозначної числової інтерпретації. В ідеалі чим менша кількість таких відповідей, тим краще, — це свідчить про близькість смислового наповнення запитань до життєвого досвіду респондентів. З іншого боку, геополітична тематика апріорі належить до категорії складних для пересічного індивіда — одна справа дозвілні роздуми про події, які в усіх на слуху, інша — аналіз поточної ситуації з урахуванням основних зовнішньополітичних напрямків.

Щоб знайти рішення цієї проблеми, ми орієнтувалися на поняття рівня відповідей (*response rate*) у масових опитуваннях¹²¹. Розвиваючи цю логіку, ми

¹²¹ Звичайно, ми не отожднюємо вибір респондентом варіанта відповіді «Важко відповісти» з його недосяжністю на етапі реалізації вибірки.

звернулися до Європейського соціального дослідження (European Social Survey), у якому мінімальна бажана межа рівня відповідей встановлена на позначці 70%. Виходячи з цього справді складними пунктами методики було визначено ті з них, щодо яких понад 30% респондентів висловилися «Важко відповісти». Таких пунктів 11. Ще в 7 запитаннях такий варіант не передбачений. Отже, з 52 пунктів, які надають респонденту можливість артикулювати складність формулювання думки, 21% є справді складними для респондентів. Розглянемо їх докладніше (табл. 16.2).

Таблиця 16.2

Питання, що викликали найбільші труднощі

Тематика запитання	Рівень складності
Статус військового співробітництва України в рамках ГУАМ (пункт 37)	67%
Довіра до офіційних ЗМІ Туреччини (пункт 15)	53%
Ставлення до можливого впливу Туреччини на внутрішньополітичне життя України (пункт 49)	38%
Довіра до офіційних ЗМІ США (пункт 17)	37%
Оцінка політики Туреччини щодо України (пункт 41)	36%
Ставлення до можливого впливу ГУАМ на внутрішньополітичне життя України (пункт 50)	35%
Оцінка історичного впливу Османської імперії на українську культуру (пункт 8)	34%
Оцінка політики Балтійської асамблеї щодо України (пункт 44)	33%
Оцінка шляхів підвищення енергетичної безпеки України (пункт 20)	32%
Міра пріоритетності міжнародного співробітництва з Туреччиною (пункт 30)	32%
Оцінка багатовекторності (пункт 58)	31%

Отже, з 11 розглянутих пунктів 5 мають стосунок до Туреччини або її історичної попередниці. Таку ситуацію можна пояснити специфікою вибірки, до якої не було включено представників кримськотатарського народу. Крім того, стосовно як Туреччини, так і інших геополітичних напрямків, варіант «важко відповісти» можна розглядати і як важливий змістовний показник. Це пов'язане з тим, що він указує на низьку поінформованість респондентів про відповідних зовнішньополітичних партнерів, важливість яких визначають, звичайно ж, не тільки специфікою масової свідомості в якийсь конкретний період часу.

Щоб продемонструвати змістовну цінність варіанта «важко відповісти», звернемося до найскладнішого питання, яке стосується військового співробітництва України в рамках ГУАМ. ГУАМ був задуманий як геополітичне утворення, серед цілей якого було і військове співробітництво. Актуальність ос-

танного для України на даному етапі обумовлена тим, що всі країни-учасниці ГУАМ мають територіальні суперечки або конфлікти, в які включена Росія як їх супротивник. У цьому сенсі 67% виборів альтернативи «важко відповісти» свідчить не про низьку актуальність цього запитання, а радше про актуальність здійснення ефективної інформаційної політики в Україні щодо ГУАМ.

Отже, в контексті вивчення геополітичних орієнтацій категорію «важко відповісти» слід розглядати не стільки як джерело невідповідей, скільки як важливий параметр досліджуваного явища.

§ 16.5. Внутрішня надійність і факторна валідність шкал

Геополітичний прагматизм. Пункти анкети, які, ймовірно, мали вимірювати геополітичний прагматизм, насправді виявилися не пов'язаними одне з одним, що позбавляло будь-якого сенсу подальшу їх перевірку з точки зору внутрішньої надійності та факторної валідності. З огляду на непридатність цих запитань для розв'язання свого головного завдання не розв'язано також часткові завдання, сформульовані щодо окремих індикаторів або їх груп.

Отже, побудова відповідної шкали залишається завданням подальших досліджень.

Геополітичні орієнтації. Достатня кількість індикаторів, спрямованих на вимірювання геополітичних орієнтацій щодо Європейського Союзу і Російської Федерації (11 і 12 відповідно), дозволила оцінити їх як елементи відповідних вимірювальних інструментів.

Незважаючи на досить велику кількість індикаторів (вісім), сформульованих для вивчення зовнішньополітичного ставлення до Туреччини, оцінити їх як єдину методичку не вдалося. Як було сказано раніше, для відповідних індикаторів характерна велика кількість виборів альтернативи «важко відповісти». У процесі попереднього кореляційного аналізу щодо обробки було виключено спостереження, якщо хоча б в одній зі змінних парного зв'язку був присутній варіант відповіді «важко відповісти». Це значно знижувало розмір вибірки. Процедура факторної валідації та перевірки внутрішньої надійності будувалася за схожим принципом — якщо хоча б одна змінна мала варіант відповіді «важко відповісти», спостереження повністю виключали з оцінки. Це, зокрема, призвело до того, що в рамках конфірматорного факторного аналізу індикаторів, що стосуються Туреччини, могло бути використано лише 20 спостережень.

Ця проблема є характерною і для шкал геополітичних орієнтацій щодо Російської Федерації (далі — «RF-geopol») і Європейського Союзу (далі — «EU-geopol»), але значно меншою мірою.

На першому етапі для побудови шкал використано всі індикатори, зазначені в таблиці 30. Спосіб побудови підсумкового значення ґрунтувався на знаходженні різниці між позитивними і негативними оцінками з відповідних запитань. Після цього були побудовані кореляційні матриці між сумарним значенням шкали і значеннями окремих змінних із метою виключення тих пунктів, які показують слабкі зв'язки. В результаті як для «RF-geopol», так і

для «EU-geopol» було залишено по 8 індикаторів, які в подальшому були використані для факторної валідазації.

Для «RF-geopol» (див. Додаток 8) це індикатори щодо специфіки історичного впливу Російської імперії (пункт 9), культурної спадщини СРСР (пункт 10), довіри офіційним ЗМІ Російської Федерації (пункт 16), близькості загальнослов'янських цінностей (пункт 23), пріоритетності міжнародного співробітництва з Російською Федерацією (пункт 28), статусу військового співробітництва з Російською Федерацією (пункт 36), оцінки політики Російської Федерації щодо України (пункт 39), а також впливу Російської Федерації на внутрішньополітичне життя в Україні (пункт 47).

До складу «EU-geopol» (див. Додаток 9) включено пункти про близькість європейської культури (пункт 1), довіру офіційним ЗМІ країн Європейського Союзу (пункт 14), близькість європейських цінностей (пункт 22), пріоритетність міжнародного співробітництва з Європейським Союзом (пункт 27), статус військового співробітництва з НАТО (пункт 35), оцінку політики Європейського Союзу щодо України (пункт 38), вплив Європейського Союзу на внутрішньополітичне життя в Україні (пункт 46), а також оцінку нинішнього курсу України на євроінтеграцію (пункт 54).

Для оцінки факторної валідності та внутрішньої надійності зазначених шкал використано масиви двох типів. Перший ґрунтувався тільки на тих спостереженнях, за якими не було зафіксовано жодної відповіді «Важко відповісти». Другий масив побудовано на підставі всіх спостережень, при цьому відповіді «важко відповісти» були перекодовані в нейтральні варіанти відповідей, тобто ті альтернативи, які не розглядали як позитивні або негативні оцінки. Відповідні результати подано в таблиці 16.3.

Таблиця 16.3

**Результати оцінки надійності та факторної валідності
для шкал «RF-geopol» і «EU-geopol»**

Показник	«RF-geopol» (N = 46)	«RF-geopol» (N = 100)	«EU-geopol» (N = 48)	«EU-geopol» (N = 100)
Альфа Кронбаха (> 0,70)	0,889	0,855	0,836	0,763
Хі-квадрат (DWLS)	11,876	8,304	42,724	28,629
df (ступені свободи)	20	20	20	20
Хі-квадрат / df	0,6	0,4	2,1	1,4
p-value	0,920	0,873	0,002	0,095
RMSEA (< 0,07)	0,000	0,000	0,155	0,066
CFI (> 0,95)	1,000	1,000	0,963	0,984
TLI (> 0,95)	1,004	1,008	0,948	0,977

Результати оцінки шкали «RF-geopol» є високими в обох випадках, що свідчить на користь її використання в подальших дослідженнях. Щодо шкали

«EU-geopol», то тут ситуація є не такою однозначною. Для вибірки, з якої виключені неповні спостереження ($N = 48$), два показники не задовольняють установлені норми. При цьому якщо величина CFI практично збігається з необхідним значенням, то RMSEA значно перевищує його. На базі повної вибірки фіксують результати, що підтверджують високу якість шкали.

Слід зазначити, що результати перевірки є проміжним етапом на шляху до побудови класифікаторів геополітичних уподобань. Оцінки факторної валідності та внутрішньої надійності важливі з точки зору того, чи можемо ми говорити про єдиний конструкт у рамках використання певного набору індикаторів. Своєю чергою, класифікатор дозволяє перейти від оцінки сили виразності установок до категорій респондентів, які мають якісні відмінності.

§ 16.6. Класифікатори геополітичних уподобань

З огляду на специфіку використовуваних запитань в основу класифікатора покладено співвідношення позитивних, нейтральних і негативних відповідей, а також кількість виборів «важко відповісти» (розрахунки наведено в Додатку 10). У результаті виокремлено п'ять категорій респондентів: а) ті, хто мають негативні установки щодо того або того вектора геополітичного співробітництва; б) ті, хто нейтрально ставляться; в) ті, хто посідають невизначену позицію; г) ті, хто демонструють амбівалентне ставлення; д) ті, хто мають позитивні установки.

Для визначення цих категорій використано правила класифікації (табл. 16.4), що поміщають кожного респондента тільки в одну з зазначених груп. З огляду на однакову кількість індикаторів в обох методиках ці правила використано як для «RF-geopol», так і для «EU-geopol».

Таблиця 16.4

Правила класифікації результатів використання шкал «RF-geopol» і «EU-geopol»

Позиція	Правила визначення
Негативна	Кількість позитивних і негативних виборів більше чотирьох (усі вибори можуть бути негативними), при цьому переважання негативних виборів над позитивними становить два і більше.
Невизначена	Кількість виборів «важко відповісти» перевищує загальну кількість позитивних і негативних виборів.
Нейтральна	Загальна кількість позитивних і негативних виборів менше чотирьох, при цьому з аналізу виключають респондентів із невизначеною позицією.
Амбівалентна	Загальна кількість позитивних і негативних виборів більше чотирьох, при цьому їх кількість однакова або переважання є мінімальним (не більше, ніж на один).
Позитивна	Кількість позитивних і негативних виборів більше чотирьох (усі вибори можуть бути позитивними), при цьому переважання позитивних виборів над негативними становить два і більше.

Схожі правила використано для класифікації респондентів у разі геополітичних орієнтацій щодо Польщі (пункти 7, 29, 40, 48), Туреччини (пункти 8, 30, 41, 49), ГУАМ (пункти 31, 37, 42, 50), Вишеградської групи (пункти 32, 43, 51), Балтійської асамблеї (пункти 33, 44, 52) і США (пункти 34, 35, 45, 53)¹²². З огляду на вкрай малу кількість індикаторів відповідні результати класифікації слід розглядати як орієнтовні. Результати подано в таблиці 16.5. Як видно, найбільш «кристалізовані» установки спостерігаємо щодо Європейського Союзу, Російської Федерації та США — загальна кількість негативних і позитивних оцінок для першого становлять 53%, для другої — 70%, для третіх — 55%. Найменшу ясність спостерігаємо щодо ГУАМ і Вишеградської четвірки: 20% і 24% відповідно.

Таблиця 16.5

**Результати використання класифікатора за різними напрямками
геополітичного співробітництва, %**

Напрямок міжнародного співробітництва	Геополітична позиція*				
	НЕГ	НТР	НВЗ	АМБ	ПОЗ
Європейський Союз	23,0	24,0	8,0	15,0	30,0
Російська Федерація	51,0	14,0	5,0	11,0	19,0
Польща (як окремий стратегічний партнер)	19,0	37,0	13,0	10,0	21,0
Туреччина	31,0	45,0	13,0	8,0	3,0
Група країн — Грузія, Азербайджан, Молдова	17,0	44,0	13,0	8,0	3,0
Група країн — Польща, Чехія, Словаччина, Угорщина	13,0	50,0	22,0	4,0	11,0
Група країн — Естонія, Латвія, Литва	14,0	37,0	21,0	5,0	23,0
США	32,0	24,0	10,0	11,0	23,0

* Тут і далі: НЕГ — негативна, НТР — нейтральна, НВЗ — невизначена, АМБ — амбівалентна, ПОЗ — позитивна.

Тепер оцінимо конструктну валідність класифікаторів для Європейського Союзу і Російської Федерації. Для зручності три центральні групи (нейтральна, невизначена і амбівалентна позиції) об'єднані в загальну групу, яку далі називатимемо проміжною. Як змінні номологічної мережі використано місто проживання респондента (для обох шкал), оцінка конфліктів між Україною і Росією в енергетичній сфері (для обох шкал), визначення найкращого варіан-

¹²² Розрахунки на прикладі Польщі наведено в Додатку 10.

та для України у сфері енергетичної безпеки (для «RF-geopol»), цінності, яких слід дотримуватися керівництву країни (для обох шкал).

Місто проживання. Незважаючи на малий обсяг вибірок, описові результати демонструють відмінності в кожному з населених пунктів (табл. 16.6). Вирізняється регіональна специфіка є характерною для відповідей на пункти шкали «RF-geopol».

Таблиця 16.6

Регіональна специфіка результатів класифікації, %

Геополітична позиція	«RF-geopol»			«EU-geopol»		
	Львів	Київ	Одеса	Львів	Київ	Одеса
Негативна	75,8	52,9	24,2	12,1	35,3	21,2
Проміжна	24,2	26,5	39,4	51,5	35,3	54,5
Позитивна	0,0	20,6	36,4	36,4	29,4	24,2

Конфлікти між Україною і Росією в енергетичній сфері. Щоб краще зрозуміти вплив геополітичних уподобань на оцінку енергетичних конфліктів між Україною і Росією, краще брати до уваги одночасно результати двох класифікацій. Так, на підставі результатів аналізу можна виокремити кілька груп респондентів (табл. 16.7), три з яких є основними — з антиросійськими геополітичними орієнтаціями, з проросійськими, а також із нейтральними (респонденти з проміжними поглядами як стосовно Європейського Союзу, так і стосовно Російської Федерації).

Таблиця 16.7

Вплив геополітичної позиції на визначення винуватця енергетичних конфліктів, кількість спостережень

Геополітична позиція	Головний винуватець конфлікту*				
	P+U	P	U	ЄС	ВС
Проєвропейська й антиросійська, антиросійська, антиєвропейська й антиросійська	14	30	0	1	6
Цілком нейтральна	12	3	1	1	5
Проросійська й антиєвропейська, проросійська	9	0	3	5	2
Проєвропейська	3	0	1	0	1
Антиєвропейська	2	0	0	0	1
Разом	40	33	5	7	15

* P+U — як Росія, так і Україна; P — Росія; U — Україна; ЄС — Європейський Союз; ВС — важко сказати

З 51 респондента з антиросійськими поглядами 44 так чи так покладають провину у виникненні конфліктів на Російську Федерацію, а 30 — виключно на неї. Водночас серед респондентів із проросійськими орієнтаціями присутні будь-які відповіді за винятком вибору Росії як єдиного винуватця. Що стосується повністю нейтральних респондентів, то переважна кількість із тих, хто змістовно визначився з відповіддю (тобто не вибрав варіант «важко сказати»), вважають винуватцем конфлікту як Росію, так і Україну.

Забезпечення енергетичної безпеки. Ще виразніший вплив геополітичної позиції щодо Росії спостерігаємо за визначенням найкращого варіанта співпраці у сфері енергетичної безпеки України (табл. 16.8). Переважна більшість респондентів з антиросійськими геополітичними орієнтаціями вважають, що найкращий варіант для України — це розвиток перспективних проектів з Європейським Союзом і лише 2 з 51 сказали, що необхідно покращити співпрацю з Росією. У той же час 16 з 19 респондентів, які позитивно оцінюють східну геополітичну перспективу України, вважають співпрацю з Росією головним пріоритетом.

Таблиця 16.8

Вплив геополітичної позиції на визначення пріоритетного вектора співпраці у сфері енергетичної безпеки, кількість спостережень

Геополітична позиція	Вектор співробітництва*			
	Р	ЄС	А	ВС
Антиросійська	2	27	6	15
Проросійська	16	0	0	3
Проміжна	9	7	0	14

* Р — Росія; ЄС — Європейський Союз; А — Азербайджан; ВС — важко сказати.

Цінності, яких слід дотримуватися керівництву країни. За аналізу ціннісної структури керівних еліт, передписаної їм респондентами, цінними є результати обох класифікаторів (табл. 16.9). Більша частина респондентів (11 із 17), що зупинилися на європейських цінностях, негативно ставиться до східної геополітичної перспективи і позитивно — до західної. Вибір загальнослов'янських цінностей характерний для тих, хто посідає проміжну позицію за результатами обох класифікаторів або ж позитивно оцінює східне геополітичне співробітництво (18 з 26). Цікава картина вимальовується з вибором прагматичних цінностей. Так, чим негативнішою є оцінка проросійського геополітичного вектора, тим більше таких респондентів — 16 у категорії з негативними поглядами, 8 — із проміжними, 4 — з позитивними. Що стосується вибору етнонаціональних цінностей, то всі 13 респондентів відповідної групи потрапляють до категорії з негативною оцінкою східного вектора геополітичного співробітництва. Зрештою, велика частина тих, хто обрав варіант «важко сказати»

(9 із 16), потрапляють до проміжної категорії за результатами використання обох класифікаторів.

Таблиця 16.9

Зв'язок геополітичної позиції з бажаною ціннісною структурою керівних еліт України

Результати класифікації		Бажані цінності*				
«RF-geopol»	«EU-geopol»	Є	ЗС	П	ЕН	ВС
НЕГ	НЕГ	0	3	2	3	1
	НТР, НВЗ, АМБ	2	0	6	7	2
	ПОЗ	11	3	8	3	0
НТР, НВЗ, АМБ	НЕГ	0	1	1	0	1
	НТР, НВЗ, АМБ	1	6	6	0	9
	ПОЗ	2	1	1	0	1
ПОЗ	НЕГ	1	8	1	0	1
	НТР, НВЗ, АМБ	0	4	3	0	1
	ПОЗ	0	0	0	0	0

* Є — європейські; ЗС — загальнослов'янські; П — прагматичні; ЕН — етнонаціональні; ВС — важко сказати.

Загалом наведені результати конструктивної перевірки свідчать на користь застосування в подальших валідизаційних дослідженнях класифікаторів на підставі шкал «RF-geopol» і «EU-geopol».

§ 16.7. Загальні геополітичні орієнтації та соціально-демографічні характеристики

З огляду на те, що в основу цього дослідження закладено 8-векторну модель геополітичних орієнтацій, необхідно, принаймні, на описовому рівні, узагальнити відповіді респондентів, що її характеризують. З цією метою ми застосували ієрархічний кластерний аналіз, як вхідні параметри якого використано різницю між позитивними і негативними відповідями щодо кожного з восьми зазначених геополітичних напрямків. У результаті було виокремлено п'ять кластерів респондентів (табл. 16.10). Для першої групи характерне слабо виражене позитивне ставлення до Європейського Союзу і нейтральне ставлення за іншими напрямками. Друга демонструє негативні оцінки щодо всіх альтернатив геополітичного співробітництва. Третю можна охарактеризувати як помірну, за винятком Російської Федерації та Туреччини. При цьому сила негативних установок щодо Росії тут є найбільш вираженою. Четверту групу можна назвати просхідною. Не дивно, що її геополітичні орієнтації щодо Європейського Союзу і США є негативними. Нарешті, п'ята група є прозахідною й антисхідною.

Таблиця 16.10

Результати кластеризації за вісьма геополітичними напрямками

Показник*	Група				
	1	2	3	4	5
Наповненість	31	9	20	20	20
Центроїд для ЄС (-8, 8)	1,0	-2,2	0,4	-3,9	5,0
Центроїд для РФ (-8, 8)	0,2	-4,1	-5,8	3,7	-4,3
Центроїд для Польщі (-4, 4)	0,2	-1,9	0,4	-1,4	1,6
Центроїд для Туреччини (-4, 4)	-0,3	-2,6	-1,0	-1,2	-0,8
Центроїд для ГУАМ (-4, 4)	0,1	-2,6	-0,3	0,0	0,3
Центроїд для Балтійської асамблеї (-3, 3)	0,3	-1,8	1,1	-0,8	1,3
Центроїд для Вишеградської четвірки (-3, 3)	0,2	-1,9	0,1	-0,9	0,9
Центроїд для США (-4, 4)	0,3	-2,4	0,8	-2,6	1,8

* У дужках указано діапазон можливих значень для відповідних центроїдів.

Для зручності подальшого опису найменуємо ці п'ять груп відповідно помірно проєвропейською, негативістською, антиросійською, антизахідно-проєвропейською і прозахідно-антиросійською. Щоб описати спостережувану гетерогенність за допомогою єдиного числового показника, можна застосувати індекс Лааксо—Таагепера, який традиційно використовують у політологічних порівняльних дослідженнях для вимірювання рівня фрагментації партійної системи:

$$N = \frac{1}{\sum_i^n p_i^2},$$

де N — ефективна кількість партій (тут ефективна кількість соціальних груп), n — номінальна кількість партій (соціальних груп), а p_i — частка i -ої партії на виборах або в легіслатурі (частка i -ої соціальної групи).

У цьому разі його використання також є виправданим, оскільки ми оцінюємо соціально-політичну фрагментацію. З огляду на те, що в результаті кластерного аналізу вибір зупинено на п'яти групах, діапазон можливих значень індексу лежить у діапазоні від одного (мінімальна соціальна гетерогенність; усі спостереження потрапляють лише в одну з п'яти груп) до п'яти (максимальна соціальна гетерогенність; усі п'ять груп мають однаково наповненість).

Індекс Лааксо—Таагепера для нашої вибірки дорівнює 4,46, що свідчить про її високу соціальну неоднорідність. Для Львова цей показник дорівнює 3,74, для Києва — 4,82, для Одеси — 3,10.

Тепер оцінимо зв'язок цих груп із різними соціальними і демографічними характеристиками. До тих із них, які показали в рамках нашого аналізу цікаві результати, належать такі — місто проживання, політичні погляди і вік.

Місто проживання. Кожне з трьох міст показує специфічний патерн розподілу відповідей (табл. 16.11). Респонденти з Києва практично рівномірно розподілені між п'ятьма групами. В Одесі можна виокремити дві основні групи — помірно проєвропейську, а також антизахідно-проросійську. Респонденти Львова практично повністю розподілені між трьома групами — помірно проєвропейською, антиросійською, а також прозахідно-антиросійською.

Таблиця 16.11

Регіональна специфіка результатів кластеризації, %

Група	Місто проживання		
	Одеса	Київ	Львів
Помірно проєвропейська	15	6	10
Негативістська	1	5	3
Антиросійська	3	7	10
Антизахідно-проросійська	10	9	1
Прозахідно-антиросійська	4	7	9

Політичні вподобання. Розглянемо найбільш примітні результати. З дев'яти респондентів, що дотримуються комуністичних або соціалістичних поглядів, шість демонструють антизахідно-проросійські погляди. З восьми респондентів із ліберальними поглядами три потрапляють у помірно проєвропейську групу і ще три — у прозахідно-антиросійську. Схожа ситуація, але з великим антиросійським акцентом є характерною і для націонал-демократів. Із восьми відповідних респондентів два потрапляють у помірно проєвропейську групу, а п'ять — у прозахідно-антиросійську. Що стосується респондентів із націоналістичними поглядами, то з восьми відповідних спостережень шість увійшли в антиросійську групу. Загалом такі результати відповідають логіці співвідношення між внутрішньо- і зовнішньополітичними вподобаннями.

Таблиця 16.12

Вікові показники різних кластерів

Група	Вікова група			Середнє значення (роки)
	Молодь	Середній вік	Похилий вік	
Помірно проєвропейська	7	15	9	43,1
Негативістська	4	2	3	39,2
Антиросійська	11	3	6	37,7
Антизахідно-проросійська	0	11	9	52,6
Прозахідно-антиросійська	9	5	6	39,9

Вік респондентів. Якщо розглядати середню кількість років, то майже всі групи є досить близькими одна до одної (табл. 16.12). Виняток становить анти-західно-проросійська група, яка є найстаршою. Так, у цю групу не потрапив жоден респондент із категорії молоді. Показово і те, що майже три чверті респондентів до 30 років так або так демонструють антиросійські установки.

Оцінюючи наведені вище соціально-демографічні характеристики груп із різними геополітичними вподобаннями, можна відзначити істотну соціально-політичну гетерогенність вибірки, що, взагалі-то, не є дивним, якщо взяти до уваги різноманітність українського суспільства, а також особливості дизайну дослідження.

§ 16.8. Висновки

Незважаючи на успішне конструювання соціологічних тестів «RF-geopol» і «EU-geopol», які відповідають базовим вимогам щодо подальших валідаційних перевірок, низку ідей, важливих з точки зору вивчення геополітичних уподобань, реалізувати не вдалося.

Перш за все недостатньо уваги було приділено концептуалізації геополітичного континууму «радикалізм-прагматизм», що в підсумку вилилося в досить довільні формулювання відповідних індикаторів. Не дивно, що останні виявилися непридатними для розв'язання покладених на них завдань.

Крім цього, якщо стосовно Російської Федерації та Європейського Союзу вдалося сконструювати повноцінні соціологічні тести, то щодо інших напрямків запропоновано тільки обмежений набір релевантних індикаторів, що природним чином звужує їх вимірювальну точність і веде до отримання лишень орієнтовних результатів.

З огляду на згадані проблеми як можливі напрямки подальшої методичної роботи бачимо такі:

- уточнення вимірювальних особливостей соціологічних тестів «RF-geopol» і «EU-geopol», а також їх конструктивна і критеріальна валідація;
- конструювання соціологічного тесту, спрямованого на оцінку геополітичного радикалізму-прагматизму респондентів;
- деталізація індикаторів, покликаних оцінити геополітичні орієнтації за такими напрямками, як Польща, Туреччина, ГУАМ, Вишеградська четвірка, Балтійська асамблея і США.

Висновки до третього розділу

У цьому розділі автор розвиває ідеї вивчення політичної культури соціологічними засобами, сформульованими Є. Головахою і Н. Паніною в середині 2000-х років. Починаючи з емпіричної оцінки вже наявного вимірювального інструмента — соціологічного тесту «Типи політичної культури» в тій його частині, яка стосується особливостей кластеризації респондентів, у подальшому автор приходять до необхідності перегляду шкали демократичних установок. Останнє пов'язане з необхідністю їх фокусування на конкретному політичному об'єкті. В іншому разі демократія, яка сама по собі є абстрактною

ідеєю, в переважній більшості випадків оцінюється позитивно. Що стосується шкали громадської активності, то вивчення її вимірювальних властивостей не виявило жодних складнощів або суперечностей.

Отже, подальшу роботу, на нашу думку, необхідно сфокусувати на дослідженні альтернатив кластеризації респондентів, а також конструюванні додаткових вимірювальних шкал, що оцінюють політичну культуру українського суспільства.

У рамках розв'язання першого завдання проведено порівняння чотирьох способів кластеризації респондентів — кластерного аналізу, розподілу на основі ідеальної точки, концептуальної типології та латентного класового аналізу. Зроблено висновок про пріоритет використання концептуальної типології або ідеальної точки поділу, де друга, з огляду на отримані результати, є менш точним, але простішим у використанні варіантом першої.

У процесі розв'язання другого завдання було апробовано три нові вимірювальні інструменти — антидемократичних установок щодо розвинених країн, антидемократичних установок щодо України, а також авторитарних установок. Усі вони показали добрі результати валідаційних перевірок і є індикаторами важливих аспектів політичної культури українського суспільства. На цій підставі було запропоновано соціологічний тест «Типи політичної культури — II», що містить відповідні шкали, а також шкалу громадської активності.

Разом з тим опис політичної культури був би неповним за звернення лише до внутрішньополітичних особливостей свідомості респондентів. Тому додаткову увагу було спрямовано на соціологічне вимірювання зовнішньополітичних уподобань або геополітичних орієнтацій серед загального населення України. Аналіз публікацій, що розглядають Україну в геополітичному дискурсі, дозволив виокремити релевантні аспекти геополітичної дійсності як основу для побудови опитувального інструментарію. При цьому ми відійшли від дихотомії «Російська Федерація-Європейський Союз», включивши в аналіз такі напрямки геополітичного співробітництва, як Польща (як окремий геополітичний партнер), Туреччина, ГУАМ, Вишеградська четвірка, Балтійська асамблея і США.

На підставі спеціального методичного дослідження було розроблено два соціологічні тести — «RF-geopol» і «EU-geopol», а також запропоновано правила їх використання як класифікаторів геополітичних уподобань щодо Російської Федерації та Європейського Союзу відповідно. Стосовно інших напрямків намічено основні індикатори, які виступають як підстава для побудови класифікаторів, але ще недостатньо деталізовані з точки зору повноцінних соціологічних тестів.

Слід зазначити, що нерозв'язаним залишилося завдання фіксації геополітичного прагматизму/радикалізму. Ця обставина задає можливий напрямок для подальших методичних пошуків.

Розділ IV

ПРИНЦИПИ І ПРОЦЕДУРИ РОЗРОБКИ СОЦІОЛОГІЧНИХ ТЕСТІВ

Глава 17

Розробка комплексних вимірювальних інструментів у дослідженнях змішаного типу

Враховуючи той факт, що ми пропонуємо до використання прагматичний підхід до соціологічного вимірювання, логічним початком цієї частини роботи є звернення до розробки вимірювального інструментарію в дослідженнях змішаного типу (mixed methods reseach)¹²³, що також ґрунтуються на прагматичних засадах.

Незважаючи на те, що проблема розробки інструментів вимірювання в соціології має тривалу і плідну історію, яка базується на таких усталених дослідницьких традиціях, як опитування (survey), соціометрія, тестові дослідження тощо, досвід досліджень змішаного типу є перспективним напрямком удосконалення існуючої практики у сфері розробки і застосування комплексних вимірювальних інструментів у соціології.

Нами було ідентифіковано три підходи до розробки опитувального інструментарію в межах досліджень змішаного типу, а також один кейс, що детально ілюструє розробку вимірювальної методики з використанням досліджень змішаного типу і може бути репрезентований як окремий підхід. Отже, далі йтиметься саме про чотири підходи. Три з них подано в англomовних джерелах (там їх іменують «frameworks») і один у вітчизняних — IDCV¹²⁴ (Е. Онво-б'юзі, Р. Бустаманте, Д. Нельсон), MDVR¹²⁵ (Р. Луйт), PEMI¹²⁶ (П.-М. Дейно, С. Джейкоб) і TVMS¹²⁷ (С. Дембіцький).

§ 17.1. IDCV-підхід

У праці Е. Онво-б'юзі, Р. Бустаманте і Д. Нельсон (IDCV-підхід) (Onwuegbuzie, Bustamante, & Nelson, 2010) подано підхід, що, на думку його авторів, оптимізує розробку кількісного інструментарію і процес конструктної валідазації.

¹²³ Для ознайомлення з дослідженнями змішаного типу див.: (Дембицкий, 2015b).

¹²⁴ Від англ. «Instrument Development and Construct Validation».

¹²⁵ Від англ. «Measurement Development, Validation, and Revision».

¹²⁶ Від англ. «Participatory Evaluation Measurement Instrument».

¹²⁷ Аббревіатура — «теоретична валідація вимірювальних шкал» (у перекладі на англійську — «Theoretical Validation of Measurement Scales»).

Етап 1. Концептуалізація досліджуваного конструкту. Мета — визначення конструкту, що стане основою розробки культурно-релевантного інструментарію. Процедури досягнення: робота з інсайдерами (тобто особами, які мають безпосередній стосунок до досліджуваного феномену і володіють експертними знаннями в цьому контексті), а також робота з власною рефлексивністю (Creswell, 2007, p. 180).

Етап 2. Ідентифікація й опис поведінкових практик, що входять до досліджуваного конструкту. Мета — деталізація обраного на попередньому етапі теоретичного конструкту. Процедури досягнення: відкрите й осьове кодування (стратегія «обґрунтованої теорії») (Glaser, 1967), методи етнографічного аналізу (наприклад, аналіз предметної галузі, таксономічний і компонентний аналіз) (Spradley, 1979), техніка Дельфі.

Етап 3. Розробка інструментарію і його первинна оцінка. Мета — оцінка змістовної валідності методики інсайдерами. Головною процедурою реалізації є побудова таблиці характеристик (Newman, Lim, & Pineda, 2013).

Етап 4. Пілотна апробація інструментарію. Мета — первинна оцінка ефективності методики «в полі». Процедури досягнення: опитування, експертні та фокус-групові обговорення з інсайдерами.

Етап 5. Польова апробація інструментарію. Мета — отримання кількісних і якісних даних, необхідних для подальшої валідації інструменту. Процедурою реалізації є опитування, що ґрунтується на вибірці достатнього розміру. Важливим аспектом цього етапу є включення до опитувальника відкритих питань.

Етап 6. Валідація інструментарію: кількісний аналіз. Мета — перевірка різних видів валідності (змістовної, критеріальної, конструктної). Можна досягти завдяки різноманітним методам статистичного аналізу.

Етап 7. Валідація інструментарію — якісний аналіз. Мета — збагачення наявної інформації завдяки «класичним перевагам» досліджень змішаного типу (Greene, Caracelli, & Graham, 1989) — триангуляції (комбінування різних джерел даних із метою кращого обґрунтування отриманих результатів), додатковості (визначення спільного і відмінного в отриманих результатах), ініціюванню (відкриття нових перспектив розгляду проблеми), розвитку (використання послідовного дизайну, де отримані результати визначають подальші методи), розширенню (забезпечення широти знань про досліджуваний феномен). Процедурами досягнення є різноманітні методи аналізу якісних даних (Leech & Onwuegbuzie, 2007; Leech & Onwuegbuzie, 2008).

Етап 8. Валідація інструментарію — змішано-якісний перехресний аналіз. Мета та сама, що і на попередньому етапі. Метю можна досягти завдяки якісному аналізу кількісних даних (для відповідної трансформації даних використовують методи формування наративного профілю) (Tashakkori & Teddlie, 1998) або кількісному аналізу якісних даних (у цьому разі можна використовувати, наприклад, бінаризацію тем, присутніх у якісних наративах) (Onwuegbuzie, 2003).

Етап 9. Валідизація інструментарію — змішано-кількісний перехресний аналіз. Автори не дають інформації, достатньої для чіткого формулювання мети цього етапу. З процедурної точки зору він являє собою інтеграцію всіх кількісних знахідок, отриманих на попередніх етапах (як на основі кількісних, так і на основі трансформованих якісних даних).

Етап 10. Оцінка отриманого інструменту і процесу його розробки. Мета — ухвалення рішення щодо припинення або продовження валідизації інструменту. Головною процедурою реалізації є проведення дебрифінгу дослідника на предмет його відчуттів щодо процесу роботи й отриманих результатів (Onwuegbuzie, Leech, & Collins, 2008).

Особливістю цього підходу є його ітеративність — у разі необхідності дослідник має повернутися до попередніх етапів для покращення валідизації. До головних переваг цієї методики ми відносимо її акцент на необхідності забезпечення культурної чутливості завдяки активному спілкуванню з інсайдерами на початкових етапах. Щодо вад, то підхід виглядає переобтяженим різного роду процедурами і перевітками, що призводить до втрати фокуса на розв'язанні конкретних завдань валідизації.

§ 17.2. MDVR-підхід

MDVR-підхід Р. Луйта (Luut, 2012) ґрунтується на праці Р. Едкока і Д. Коляе (Adcock & Collier, 2001), яка звертається до поняття вимірювальної валідності, й пояснює останню шляхом інтеграції трьох рівнів вимірювання¹²⁸ — отримані на основі індикаторів (третій рівень) числові значення (четвертий рівень) слід осмислено інтерпретувати з точки зору концепту (другий рівень), покладеного в основу формулювання індикаторів (осмисленими мають бути і результати якісної класифікації числових значень). При цьому Р. Луйт робить акцент на розробці крос-культурних вимірювальних інструментів.

Етап 1. Якісний аналіз інструменту — крос-культурна змістовна валідизація. Мета — перевірка змістовної еквівалентності різних мовних варіантів інструменту. Процедурою досягнення є проведення фокус-групових обговорень смислового навантаження індикаторів із представниками різних цільових груп.

Етап 2. Кількісний аналіз інструменту — крос-культурна конструктивна валідизація. Мета — перевірка схожості факторних структур інструменту в різних цільових групах. Для цього рекомендовано використовувати експлораторний факторний аналіз, хоч інші методи також можна згадати (кластерний аналіз, багатовимірне шкалювання, моделювання структурними рівняннями). Незважаючи на те, що не згадується аналіз індикаторів за допомогою звичайної описової статистики (як, наприклад, розрахунок середніх значень і стандартних відхилень), із контексту роботи випливає необхідність її використання.

¹²⁸ Ідеться про другий, третій і четвертий рівні. Перший рівень перебуває поза межами вимірювальної валідизації й належить до концептуалізації досліджуваного феномену, тобто теоретичної роботи.

Етап 3. Поєднання результатів якісного і кількісного аналізу — ревізія індикаторів. Мета — розгляд і корекція проблемних індикаторів інструментарію. Проблемні індикатори визначають на базі результатів статистичного аналізу (наприклад, хибна позиція у факторній структурі або нетипові значення показників описової статистики). Основою для відповіді на те, чому виникли відповідні проблеми, є тексти фокус-групових обговорень.

У підході Р. Луйта (а точніше Р. Едкока і Д. Кольє), як і в попередньому разі, наголошено на важливості ітеративності під час розробки інструментарію, а також наведено рекомендації щодо ухвалення відповідних рішень. Важливо зазначити, що опис етапів наведено на прикладі розробки конкретної методики — MANI-II, призначеної для вимірювання маскулінної ідеології (masculinity ideology). Цей факт дещо знижує універсальність розглянутих у межах кожного етапу методів. Перевагою підходу ми вважаємо цілісність його концептуальних основ, у яких поєднані ідеї концептуалізації, операціоналізації та статистичної перевірки. При цьому Р. Луйт збагачує базові ідеї цієї концепції за рахунок досліджень змішаного типу. До вад слід віднести певну партикулярність використовуваних методів, яка залишає низку «білих плям» у методологічних побудовах підходу.

§ 17.3. РЕМІ-підхід

РЕМІ-підхід (П.-М. Дейно і С. Джейкоб) (Daigneault & Jacob, 2014) репрезентує досвід розробки методики вимірювання часті зацікавлених сторін в оцінюванні програм. У цьому разі валідацію планували відповідно до типово кількісного підходу, але впровадження відкритих запитань в опитувальник, що окремі науковці трактують, зокрема Дейно і Джейкоб, як елемент досліджень змішаного типу (Johnson & Turner, 2003), дозволило виявити серйозні вади їхньої методики. Це, своєю чергою, призвело до перегляду деяких аспектів останньої. Цікаво те, що автори не мали особливих очікувань щодо якісного елемента опитувальника.

Етап 1. Концептуалізація й операціоналізація досліджуваного явища. Мета — розробка змістовно валідного теоретичного конструкту, формулювання індикаторів і подання їх у вигляді анкетних запитань. Процедури: аналіз літератури, а також традиційні прийоми формулювання анкетних запитань (Садмен & Бредберн, 2002; Феннето, 2004).

Етап 2. Попередня оцінка інструменту. Мета — попередня оцінка опитовального інструменту. Щодо використовуваних процедур зазначено лише те, що це пілотний тест, який має неформальний якісний характер (можливо, йдеться про оцінку методики ключовими інформантами).

Етап 3. Кількісна валідація інструменту. Мета — перевірка конвергентної валідності інструменту. Процедури: кореляційний аналіз.

Етап 4. Змішана валідація інструменту. Мета — оцінка того, чи згодні/незгодні учасники попереднього етапу з індексами, отриманими на підставі отриманих від них відповідей. Процедури: описова статистика (порядкова шкала для

визначення ступеня згоди/незгоди з індексами: 1 — «узагалі не згоден», 2 — «певною мірою згоден», 3 — «повністю згоден»), а також тематичний аналіз (Guest, MacQueen, & Namey, 2012) (використано для аналізу відповідей відкритого запитання з приводу того, чому було надано індексам саме такі оцінки).

Етап 5. Доопрацювання інструменту. Мета — корекція методики з урахуванням результатів попереднього етапу. Процедури визначають відповідно до з'ясованих проблем. Для Дейно і Джейкоб було достатньо дещо змінити шкалу відповідей на кілька питань методики.

Етап 6. Кількісна валідація фінального інструменту. Мета — перевірка того, як зміни в інструменті позначаються на змінах в отриманих відповідях, а також як учасники дослідження оцінюють такі зміни. Процедури: використання додаткових запитань щодо зроблених змін, статистичний аналіз.

Як і у випадку підходу Р. Луйта, автори РЕМІ фокусуються на конкретному прикладі, що зменшує універсальність їхніх побудов. З іншого боку, його доопрацювання не потребує значних зусиль — додаткові можливі етапи знову ж таки базуються на ідеї ітеративності (в розумінні повернення до попередніх етапів із метою відповідного доопрацювання). Цікавим елементом, запропонованим у праці Дейно і Джейкоб, є використання спеціальних питань в опитувальному інструментарії для виявлення його слабких місць. Схожий, але більш комплексний і методологічно обґрунтований підхід, репрезентовано у статті, присвяченій методиці оцінки суб'єктивної надійності (Головаха и др., 2008).

§ 17.4. TVMS-підхід

Підхід С. Дембіцького (*TVMS-підхід*) розкрито у двох працях. У першій розглянуто його процедурні особливості (Дембицкий, 2012), у другій — проблематичність суто кількісної валідації методики (Дембицкий, 2014а). Він ґрунтується на ідеї, що перехід до розробки і валідації методики позбавлений будь-якого сенсу, допоки дослідник не оперує валідною теорією, в контексті якої така розробка має здійснюватися.

Етап 1. Підготовка теоретичної бази. Мета — розробка теоретичного конструкту (а краще — теорії), що буде покладена в основу розробки опитувального інструментарію. Процедури: використання якісних (наприклад, «обґрунтована теорія» (Страусс & Корбин, 2001) або кейс-стаді (Yin, 2009)) і змішаних (наприклад, узгодження концептів (Kane & Trochim, 2007)) стратегій дослідження¹²⁹, придатних для пророблення теоретичної бази валідації. Такі стратегії можна використовувати в комплексі.

Етап 2. Розробка і коригування емпіричних індикаторів відповідно до використовуваної теорії. Мета — приведення конфігурації опитувальної методики до такого формату, який дозволяє отримувати результати, що узгоджуються з

¹²⁹ У цьому разі під стратегією дослідження (або дослідницькою стратегією) розумітимемо набір інтегрованих методів збору й аналізу даних, придатних для розв'язання широкого кола завдань (Дембицкий, 2016с, с. 11).

цією теорією. Процедури: пілотні опитування, триангуляція кількісних і якісних даних на базі теоретичних положень.

Етап 3. Перевірка конструктивної валідності та ретестової надійності. Мета зрозуміла з назви етапу. Процедури: опитування на основі більших вибірок, статистичний аналіз, експертні опитування, аналіз документальних джерел.

Етап 4. Виявлення специфіки інтерпретації отриманих результатів. Мета — з'ясування принципів типології кількісної інформації, отриманої за допомогою розробленої методики (фактично, прояснення основ переходу від кількісних даних до якісних). Процедури: статистичний аналіз різних варіантів типології, отриманих на основі кількісних даних.

Очевидно, що особливістю цього підходу є його теоретична підпорядкованість — перший етап присвячений розробці теоретичних положень, решта здійснюється в її контексті. Перевагою в цьому разі є досягнення науковцем необхідного рівня теоретичної чутливості. Останню можна розуміти як «власність проникливості, здатність розуміти й уміння відокремлювати прийнятне від того, що таким не є» (Страусс & Корбин, 2001, с. 35). Хоча цього не видно з поданих вище етапів, цей підхід також є ітеративним (особливо це стосується другого етапу). Якщо ж говорити про вади, то наведені положення знову ж таки звертаються до досвіду розробки конкретної вимірювальної шкали, що звужує змістовну і процедурну універсальність підходу.

Ця монографія є розвитком ідей, закладених у цей підхід. Так, етапи розробки соціологічних тестів узгоджуються з етапами TVMS-підходу: 1) в обох випадках теоретична валідація; 2) розробка і коригування емпіричних індикаторів, з одного боку, і вибір різновиду соціологічного тесту — з іншого (тут вибір різновиду є концептуальною частиною етапу, а розробка індикаторів — методичною); 3) перевірка конструктивної валідності та ретестової надійності, з одного боку, й емпірична валідація — з іншого (фактично емпірична валідація охоплює перевірку конструктивної валідності як свою складову); 4) виявлення специфіки інтерпретації отриманих результатів, з одного боку, і розробка критеріальних рівнів соціологічного тесту — з іншого (один і той самий різновид методологічної діяльності, але стосовно різних типів вимірювальних інструментів). Завдяки накопиченому практичному досвіду розробки соціологічних тестів можна говорити про покращення змістовної та процедурної універсальності TVMS-підходу.

§ 17.5. Порівняльний аналіз підходів

Для виявлення спільного і відмінного між цими підходами слід звернутися до дизайнів, використовуваних у дослідженнях змішаного типу. Згідно з відомою типологією (Johnson & Onwuegbuzie, 2004, р. 21–22), у найзагальнішому вигляді їх слід поділити на конкурентні та послідовні, при цьому кількісний і якісний компоненти дизайну можуть мати еквівалентний або нееквівалентний статус. На нашу думку, всі чотири підходи використовують послідовний дизайн (хоча послідовність якісних і кількісних компонентів може відрізнятись).

Що ж стосується статусу компонентів, то в цьому значенні підходи різняться. Перші три (IDCV, MDVR і PEMI) використовують еквівалентний статус кількісного і якісного компонентів, натомість TVMS-підхід виходить із домінуючого статусу якісного (це пов'язане з принциповим значенням першого етапу, результати якого застосовують як орієнтири щодо подальшої валідації).

У системі скорочень, яку застосовують у межах розглядуваної типології¹³⁰, проаналізовані вище підходи можна подати у вигляді таких «ланцюгів»:

IDCV: 1) QUAL => 2) QUAL => 3) QUAL => 4) QUAL + quan => 5) QUAN => 6) QUAN => 7) QUAL => 8) QUAL + quan => 9) QUAN + qual => 10) QUAL.

MDVR: 1) QUAL => 2) QUAN => 3) QUAN + QUAL.

PEMI: 1) QUAN => 2) QUAL => 3) QUAN => 4) QUAL + quan => 5) QUAN => 6) QUAN + qual.

TVMS: 1) QUAL + quan => 2) QUAN + QUAL => 3) QUAN => 4) QUAN + QUAL.

У такому вигляді всі підходи значно відрізняються — кожен з авторів розв'язує завдання валідації відповідно до своїх методологічних уподобань. Е. Он-воб'юзі, Р. Бустаманте і Д. Нельсон (IDCV-підхід) пропонують спеціально створений підхід. Р. Луйт (MDVR-підхід) як рекомендації використовує концепцію вимірювальної валідності Р. Едкока і Д. Кольє. П.-М. Дейно і С. Джейкоб (PEMI-підхід) приходять до розуміння необхідності особливого підходу в межах своєї дослідної роботи після отримання несподіваних результатів. С. Дембіцький (TVMS-підхід) звертається до концепції теоретичної валідності¹³¹. Отже, за співставлення цих підходів слід зважати на суттєві відмінності теоретико-методологічних засад, які використовують різні автори (Дембіцький & Хляпатура, 2015, с. 42).

Важливою відмінністю між підходами є специфіка їх ітеративності. В IDCV, MDVR і PEMI ітеративність, умовно кажучи, має індуктивний характер — майже на кожному з етапів можуть бути отримані знахідки, які змусять дослідника або коригувати методику, або повернутися до одного з попередніх етапів для повторного проведення. При цьому визначальною є специфіка саме того етапу, в межах якого отримано відповідні знахідки. Що стосується TVMS, то тут ітеративність має радше дедуктивний характер — усі результати інспектують не лише з урахуванням специфіки етапу, в межах якого вони отримані, а й тих теоретичних положень, які покладені в основу валідації вимірювальної методики і які вважають теоретично валідними.

¹³⁰ QUAL — якісний компонент із домінуючим статусом, QUAN — кількісний компонент із домінуючим статусом, qual — якісний компонент із підпорядкованим статусом, quan — кількісний компонент із підпорядкованим статусом.

¹³¹ Цю концепцію розкрито в монографії «Теоретическая валидизация в социологическом исследовании: методология и методы» (Дембицкий, 2016с).

Оцінивши ці підходи в контексті розробки соціологічних тестів, їх цілком слід віднести до сфери традиційної емпіричної соціології. Відповідно, немає жодних методологічних бар'єрів для їх використання в межах розробки соціологічних тестів. Більше за те, окремі процедури, які використовують у цих підходах, становлять особливий інтерес, оскільки є маловідомими у вітчизняній емпіричній соціології. До таких віднесемо побудову таблиць характеристик, формування нарративного профілю, бінаризацію тем, проведення дебрифінгу.

§ 17.6. Окремі процедури змішаного типу, придатні в межах розробки соціологічних тестів

*Таблиця характеристик*¹³² являє собою процедуру, що призначена для експертної оцінки змістовної валідності методики шляхом зіставлення набору індикаторів із набором складових конструкту, який є основою для конструювання вимірювального інструментарію. Використання цієї техніки передбачає збір як кількісних, так і якісних даних. Кількісні дані дозволяють розрахувати відповідність індикаторів, а також успішність охоплення досліджуваного конструкту за допомоги відсоткових шкал (хоча можна використовувати й інші їх типи). Якісні дані містять пропозиції (у разі необхідності) з приводу додаткових індикаторів, відсутніх у початковому переліку, а також зворотний зв'язок щодо адекватності змісту самого конструкту.

Змістовно таблиця характеристик є таблицею, в якій стовпці являють собою складові конструкту, що підлягають вимірюванню, а рядки — індикатори, сформульовані для виконання цього завдання. Утім, вважаємо, що формат таблиці не є оптимальним через подання назв конструктів, а також формулювання індикаторів. Наочнішим і зручнішим є використання формату стандартизованої анкети. При цьому, на нашу думку, особливу увагу слід приділити формулюванню завдань для експертів з урахуванням природи об'єкта вимірювання. Так, у разі вимірювання латентних змінних формулювання повинні мати інший характер, ніж у разі вимірювання внутрішніх станів або соціальних феноменів.

Як приклади розглянемо таблиці характеристик для оцінки змістовної валідності соціологічних тестів «RF-geopol» (геополітичні орієнтації стосовно Російської Федерації) та «Інтегральний індекс соціального самопочуття» (соціальне благополуччя). Перша частина оцінки стосується складових відповідних конструктів, друга — індикаторів методик.

1) Таблиця характеристик для соціологічного тесту «RF-geopol».

Як Ви вважаєте, чи обумовлені наведені нижче характеристики індивіда його геополітичними орієнтаціями?

1. Так, вони чинять прямий вплив
2. Так, але лише частково
3. Ні, не обумовлені

¹³² Опис цієї техніки ґрунтується на праці І. Ньюмана у співавт.: (Newman, Lim, & Pineda, 2013).

Особливості історичної пам'яті стосовно зовнішньополітичних партнерів	1	2	3
Довіра до ЗМІ різних країн	1	2	3
Ціннісні орієнтації щодо наднаціональних утворень	1	2	3
Ставлення до різних напрямків зовнішньополітичного співробітництва	1	2	3
Оцінка інтересів своєї країни з точки зору діяльності різних зовнішньополітичних сил	1	2	3

Яких характеристик, на Вашу думку, тут не вистачає?

Тепер скажіть, чи охоплюють ці характеристики зміст поданих нижче індикаторів (індикатори модифіковані для оцінки геополітичних орієнтацій стосовно Російської Федерації¹³³).

1. Повністю охоплюють
2. Охоплюють частково
3. Охоплюють незначною мірою або не охоплюють узагалі

Особливості історичної пам'яті стосовно Російської Федерації			
Яким, на Ваш погляд, був вплив Російської імперії на українську культуру?	1	2	3
Як Ви оцінюєте культурну спадщину радянського періоду в історії України?	1	2	3
Довіра до ЗМІ різних країн			
Як Ви вважаєте, чи можна довіряти офіційним ЗМІ Російської Федерації?	1	2	3
Ціннісні особливості наднаціонального характеру			
Наскільки близькі особисто Вам загальнослов'янські цінності (слов'янська єдність, релігійність, шанування роду і т.д.)?	1	2	3
Ставлення до різних напрямків зовнішньополітичного співробітництва			
Будь ласка, оцініть пріоритет міжнародного співробітництва України з Російською Федерацією.	1	2	3
Визначте статус військового співробітництва з Російською Федерацією.	1	2	3
Оцінка інтересів своєї країни з точки зору діяльності різних зовнішньополітичних сил			
Дайте оцінку політиці Російської Федерації стосовно України на даному етапі.	1	2	3
Як Ви ставитеся до того, щоб Російська Федерація впливала на внутрішньополітичне життя України?	1	2	3

Яких індикаторів, на Вашу думку, тут не вистачає?

2) Таблиця характеристик для соціологічного тесту «Інтегральний індекс соціального самопочуття» (далі подано лише першу частину таблиці специфіка-

¹³³ Оскільки ця таблиця специфікації є лише ілюстрацією, варіанти відповідей для індикаторів опущено.

ції, оскільки тільки вона має принципову відмінність порівняно з попереднім прикладом).

Як Ви вважаєте, чи обумовлюють наведені нижче сфери життєдіяльності індивіда його соціальне самопочуття?

1. Так, вони чинять прямий вплив
2. Так, але лише частково
3. Ні, не обумовлюють

Сфера соціальних відносин	1	2	3
Сфера соціальної безпеки	1	2	3
Сфера національних відносин	1	2	3
Соціально-політична сфера	1	2	3
Професійно-трудова сфера	1	2	3
Інформаційно-культурна сфера	1	2	3
Рекреаційно-культурна сфера	1	2	3
Матеріально-побутова сфера	1	2	3
Сфера міжособистісних відносин	1	2	3
Особистісна сфера	1	2	3

Яких сфер життєдіяльності, на Вашу думку, тут не вистачає?

Спершу звернемо увагу на особливості формулювання завдань для перших частин обох таблиць характеристик. У першому прикладі наголошено, що характеристики обумовлені досліджуваною властивістю, у другому — що характеристики, навпаки, обумовлюють її. Така особливість формулювань є принциповою з урахуванням природи об'єктів, що підлягають вимірюванню.

Наступний аспект, на якому треба наголосити, — шкала оцінювання, яку пропонують експертам. Ми відійшли від стандарту, репрезентованого І. Ньюманом і його колегами як основного (відсоткова шкала), оскільки в ньому, на нашу думку, бракує критеріїв для інтерпретації отриманих результатів.

Нарешті, незважаючи на те, що приклад наведено як для соціологічного тесту у формі шкали, так і для соціологічного тесту у формі індексу, релевантність застосування таблиць характеристик для них не є однаковою. Ми вважаємо, що експертне оцінювання є адекватнішим у разі розробки соціологічних індексів, де воно справді може стати у пригоді. Що стосується шкал, то правильність охоплення латентної змінної краще оцінювати за допомоги методичних опитувань цільових груп. Фактично за незбігу результатів факторної валідації та результатів використання таблиці характеристик перевагу слід надати першим. У цьому сенсі результати, отримані за допомоги таблиці характеристик, слід використовувати лише для розширення переліку можливих індикаторів розроблюваної шкали.

З огляду на викладене вище, таблиця характеристик є технікою оцінки і покращення змістовної валідності, яка є особливо актуальною в разі неможливості проведення спеціальних емпіричних досліджень, спрямованих на теоретичну валідизацію.

Формування нарративного профілю являє собою процес, завдяки якому кількісні дані трансформуються в якісні категорії. Прикладом такого процесу є групування розподілу кількісної змінної на інтервали, що репрезентують відмінні якісні категорії (Teddlie & Tashakkori, 2009, p. 271). У цьому контексті виокремлення критеріальних рівнів соціологічного тесту є різновидом формування нарративного профілю.

Вирізняють п'ять типів нарративного профілю — профіль середніх значень, порівняльний профіль, цілісний профіль, модальний профіль і нормативний профіль. Профіль середніх значень базується на аналізі середніх значень із подальшим описом груп на цій основі. З огляду на наш досвід розробки критеріальних рівнів соціологічного тесту SCL-9-NR слід наголосити необхідність використання в окремих випадках медіан замість середніх значень (зокрема це актуально за асиметричності розподілів змінних, які є основою формування нарративних профілів). Використання порівняльного профілю передбачає порівняльний аналіз з акцентом на спільних і відмінних характеристиках відповідних одиниць. Цілісний профіль є описом загального враження, що складається під дією всіх аспектів досліджуваних даних. Модальний профіль акцентує увагу на тих атрибутах досліджуваної сукупності, на які натрапляємо найчастіше. Нормативний профіль є близьким до порівняльного профілю, але базується на порівнянні отриманих даних із певним стандартом. Таким стандартом, наприклад, може бути вибірка стандартизації або певна популяція (Teddlie & Tashakkori, 2009: p. 272).

Скептично налаштований читач може зауважити, що вказані типи нарративного профілю є лише загальними установками щодо операцій із кількісними даними. Зважаючи на те, що Дж. Кресвелл і В. Плано оцінюють цю сферу наукової діяльності як практично не розроблену, такий читач не дуже і помилиться. Утім, це не знецінює саму проблему переходу від кількісних даних до якісних категорій. Можна сказати, що формування нарративного профілю — це спроба розв'язання проблеми інтерпретації кількісних даних. Остання є особливо актуальною в контексті розробки соціологічних тестів.

Ураховуючи отриманий нами досвід, можемо окреслити деякі шляхи побудови нарративного профілю в межах соціологічного тестування. Розглянемо їх із залученням відповідних прикладів. За розробки комплексного запитання для визначення типу студента на базі його поведінкових практик розбиття шкали відповідей на три якісні категорії було здійснено на базі результатів теоретичної валідизації (Дембицкий, 2016с, с. 159–164). Фактично нарративний профіль був сформований ще до проведення кількісного дослідження. Його можна назвати теоретично обумовленим нарративним профілем (QUAL => quan => QUAL). В іншому дослідженні за розробки правил обробки результатів соціологічного тесту «Типи політичної культури» ми дійшли висновку, що сама шкала відповідей на індикатори методики задає якісні стандарти інтерпретації отриманих резуль-

татів, якими не можна нехтувати. У цьому разі відповіді слід розглядати в цілком якісному розрізі — сукупність відповідей респондента дозволяє категоризувати його за допомоги спеціально сформульованих логічних правил. З урахуванням цього можна говорити про нарративні профілі-класифікатори (QUAN => qual). Зрештою, як ми вже згадували, було використано аналіз медіан у різних групах із підвищеним або високим рівнем психологічного дистресу для виокремлення критеріальних рівнів соціологічного тесту SCL-9-NR (профіль середніх значень або медіан; QUAN => qual). Розглянемо специфіку їх застосування.

Теоретично зумовлений профіль передбачає фокусування на якісно відмінних групах, що визначаються ще на етапі теоретичної валідації, й у подальшому фіксуються за допомоги соціологічного тесту або будь-якого іншого вимірювального інструментарію. Ураховуючи концепт теоретичної чутливості, дослідник може натрапити на такі групи і зафіксувати їх у «полі». Подальша розробка вимірювального інструменту здійснюється з орієнтацією на отриману інформацію.

Профілі-класифікатори є способом категоризації респондентів, якщо немає попередньої валідної інформації про якісно відмінні групи, а шкала відповідей на індикатори тесту є категоріальною. При цьому категоріальність передбачає не ступінь виразності (наприклад, «слабко — помірно — сильно») або частоту (наприклад, «ніколи — зрідка — періодично — майже постійно»), а різні якісні стани (наприклад, «згоден — не згоден» або «вистачає — не вистачає»). Такі категорії самі по собі задають логічні кордони можливих інтерпретацій отриманих результатів, на які слід зважати.

Застосування профілів середніх (медіанних) значень, як свідчить наш досвід, передбачає збір великої кількості даних щодо, принаймні, кількох релевантних груп. Так, ми отримали інформацію про три групи з виразним психологічним дистресом — осіб із певними психіатричними діагнозами, осіб, що мають незадовільний стан фізичного здоров'я, а також осіб, що перебувають у соціальній скруті, — респонденти, які проживали в зоні бойових дій. Такі профілі також набувають порівняльного характеру. Так, процедура виокремлення кількісних значень, що характеризують якісні рівні психологічного дистресу, включала зіставлення результатів зазначених груп для знаходження патернів, що їх об'єднують. Очевидно, що в цьому разі шкали відповідей мають схоплювати саме ступінь виразності або частоту досліджуваної ознаки.

Вважаємо, що наведені приклади цілком вкладаються в логіку досліджень змішаного типу, збагачують практику використання нарративних профілів, а також задають напрямок для подальших наукових пошуків.

*Бінаризація тем*¹³⁴ є способом розрахунку величини ефектів¹³⁵ для якісних даних. Отже, бінаризація тем є переходом від якісних даних до кількісних, що відкриває додаткові можливості для аналізу досліджуваного феномену. Е. Он-

¹³⁴ Опис цієї техніки ґрунтується на праці Е. Онвоб'юзі: (Onwuegbuzie, 2003).

¹³⁵ Під величиною ефекту мається на увазі кількісний вимір величини певного статистичного показника.

воб'юзі поділяє розрахунок величини ефектів за допомоги бінаризації на експлаторний і конфірматорний.

Експлаторна бінаризація є продовженням тематичного якісного аналізу. Вона полягає в переведенні ідентифікованих у тексті або під час спостереження тем у цифровий формат, що дозволяє як розрахувати частоту появи тем, так і оцінити їх інтенсивність. Для цього тема, що з'являється в певній одиниці розгляду, отримує числове значення «1», якщо ж тема відсутня, вона отримує числове значення «0». У результаті дослідник може отримати два види матриць — інтерреспондентну та інтрареспондентну. Інтерреспондентна матриця демонструє, які індивіди роблять внесок до кожної теми, що з'являється в дослідженні, натомість інтрареспондентна матриця містить аналогічну інформацію на рівні сукупності тверджень окремого респондента або певного набору досліджуваних подій. Для врахування контексту, в якому розкривається досліджуваний феномен, можна застосувати числове кодування відповідних аспектів.

Експлаторна бінаризація дозволяє розраховувати явні (manifest) і латентні величини ефектів. Явні величини ефектів розраховують на базі змісту, який досліджується безпосередньо, — частотні величини ефектів, а також скориговані величини ефектів. Частотні величини ефектів отримуються завдяки розрахунку частот кожної теми на базі інтерреспондентної матриці. До цієї самої категорії величин ефектів належить розрахунок інтенсивності, що передбачає використання інтрареспондентної матриці, на основі якої визначають кількість значимих тверджень у межах кожної теми. В обох випадках величини ефектів можуть бути конвертовані у відсотковий формат. Скориговані величини ефектів є частотними величинами ефектів, що враховують не тільки частоту згадування, а і довжину аналізованих одиниць.

Для визначення ієрархічної структури аналізованих тем на базі обох видів матриць можна застосовувати експлаторний факторний аналіз. Отримані в результаті мета-теми дозволяють сформувати відповідні матриці.

Латентні величини ефектів можна отримати завдяки статистичному аналізу матриць метатем. Окрім того, латентні величини ефектів розраховують за використання інтерреспондентних матриць для формування наративних профілів (середніх значень, порівняльного, цілісного, модального і нормативного). В цьому разі відбувається повернення до якісного формату даних на базі їх кількісного подання.

Конфірматорну бінаризацію застосовують для статистичної генералізації слів або спостережень, що проявили себе у процесі дослідження, стосовно певного контексту або навіть генеральної сукупності індивідів за умови використання ймовірнісних вибіркового схем достатнього розміру. У цьому сенсі конфірматорна бінаризація є доповненням до аналітичної генералізації, головною метою якої є розповсюдження теоретичних понять на певні контексти (Дембицкий, 2016с, с. 69). Що стосується статистичних методів, то їх прийнятність цілком визначає природа кількісних даних, що підлягають аналізу.

Оцінюючи бінаризацію тем у контексті розробки соціологічних тестів, слід зазначити прийнятність її використання у процесі досягнення теоретичної валидності (аналіз якісних або текстових даних), а точніше, її композиційного компоненту.

Незважаючи на те, що ми не використовували бінаризацію тем у тому вигляді, як її описує Е. Онвоб'юзі, маємо досвід застосування бінаризації тверджень у межах дослідницької стратегії узгодження концептів. Остання дуже добре вписується в парадигму досліджень змішаного типу і є ефективним інструментом структурування концептуального поля досліджуваного феномену. Слід зауважити, що застосовувана в узгодженні концептів бінаризація має експлораторний характер.

Якщо ж говорити про конфірматорну бінаризацію, то близьким до неї є створення дихотомічних категорій на основі порядкових шкал відповідей на індикатори соціологічного тесту SCL-9-NR для подальшого використання цих категорій з метою класифікації респондентів із застосуванням дерев рішень (дискримінативний потенціал тесту).

Загалом можна стверджувати, що бінаризація є універсальною технікою, про яку слід пам'ятати на всіх етапах розробки соціологічного тесту. Що стосується конкретніших рекомендацій, то їх формулювання потребує подальшого накопичення методичного досвіду.

Термін «*дебрифінг*»¹³⁶ почали застосовувати в армії Сполучених Штатів Америки для позначення отримання інформації від пілота або солдата після виконання місії, а також його інструктування щодо того, яку інформацію він може оприлюднити (Bartone, 1995). Цей термін запозичений дослідниками якісного профілю для позначення детального аналізу проведеного дослідження з метою уникнення двох видів зміщень, які можуть виникнути у процесі якісного дослідження. Перший пов'язаний з впливом дослідника на учасників дослідження або ключових інформантів, другий — з впливом учасників дослідження або ключових інформантів на дослідника. Перший різновид має місце, коли дослідник створює загрозу або не адекватно себе позиціонує в певних соціальних або інституційних взаємостосунках, що в подальшому веде до спротиву з боку учасників дослідження. Що стосується другого різновиду, то його спостерігаємо в ситуаціях, коли дослідник надмірно зближується з досліджуваною групою, наприклад, стає повноцінним учасником досліджуваних процесів замість виконання ролі периферійного інсайдера, який користується прийнятною аналітичною перспективою і не бере активної участі в діяльності відповідної групи.

З огляду на це за проведення якісного дослідження особливо гостро постає необхідність забезпечення систематичної рефлексивності дослідника, яка є засобом зменшення зазначених загроз. Такої рефлексивності можна досягти завдяки дебрифінгу, метою якого є визначення і відбиття того, якою мірою зміщення потенційно можуть вплинути на різні аспекти дослідження (форму-

¹³⁶ Опис цієї техніки ґрунтується на праці Е. Онвоб'юзі: (Onwuegbuzie, Leech, & Collins, 2008).

лювання дослідницьких запитань, упровадження процедур збору даних, використання аналітичних процедур), як зміщення змінювалися у процесі дослідження загалом і в окремих інтерв'ю зокрема, а також як вони вплинули на інтерпретацію отриманих результатів. Окрім цього, дебрифінг надає можливість оцінити ініціальні припущення дослідника.

Е. Онвоб'юзі виокремлює п'ять основних типів дебрифінгу якісного дослідника, а також додає до них іще один, ефективність якого він особисто адвокує: 1) дебрифінг колег; 2) дебрифінг з учасниками дослідження після його завершення; 3) дебрифінг «брамника»; 4) дебрифінг дослідників, залучених в один і той самий дослідницький проект; 5) дебрифінг модераторів фокус-груп; 6) дебрифінг інтерпретативного дослідника.

Дебрифінг колег (не залучених до дослідження) допомагає отримати зовнішню оцінку дослідницького процесу. Дебрифінг колег є еквівалентом надійності кодувальників за тієї відмінності, що він має логічну, а не емпіричну природу. Така зовнішня оцінка зобов'язує дослідника залишатися правдивим, відповідати на складні питання щодо дослідження, а також вивільнити ті емоції, які він отримав у процесі дослідницької роботи.

Дебрифінг з учасниками дослідження полягає в інтерв'юванні останніх на етапі його завершення з метою пояснення цілей і результатів дослідження, а також відповіді на будь-які запитання або занепокоєність учасників. Цей тип дебрифінгу також може бути проявом поваги і вдячності учасникам за той важливий внесок, який вони зробили під час дослідження. Більше за те, він може допомогти зняти будь-яку напругу, що виникла під час дослідження між учасниками і дослідником.

Дебрифінг «брамника» має місце тоді, коли дослідник взаємодіє з особою («брамником»), від якої він має отримати дозвіл на доступ до учасників дослідження або до місця його проведення. У цьому разі дослідник підсумовує дослідницькі знахідки, особливо ті з них, що є найцікавішими для «брамника» і/або передбачають важливі наслідки для учасників або місця проведення дослідження. Головна мета цього дебрифінгу — забезпечення подальшого доступу до учасників або місця здійснення дослідження.

Дебрифінг дослідників, що працюють над спільним науковим завданням, може бути присвячений різним аспектам дослідницького процесу, наприклад, непередбаченим знахідкам, етичним дилемам, що виникли у процесі дослідження, модифікації, яких потребує дослідницький план тощо. Можливість обговорювати думки, сприйняття, почуття і досвід може допомогти дослідникам більше заглибитися у проблему, яку вони вивчають.

Дебрифінг модераторів фокус-груп (проводять, коли до дослідження залучено кілька модераторів) проводять одразу після фокус-групової сесії. Він передбачає обговорення початкового сприйняття учасників дослідження, цінних тверджень, які були ними зроблені, неочікуваних знахідок, а також того, чи досягнуто теоретичного насичення (нові фокус-групи не додають нової інформації). Найважливішим при цьому є те, що модератори матимуть можливість змінити фокус групових обговорень у чергових фокус-групах.

Дебрифінг інтерпретативного дослідника полягає в застосуванні спеціальних рекомендацій за участі висококваліфікованого спеціаліста (перш за все у сфері якісних досліджень із широким досвідом інтерв'ювання), який не залучений до дослідження і не є зацікавленою стороною. Останній проводить інтерв'ю з дослідником, яке бажано зафіксувати в аудіо- або навіть відеоформаті. Крім цього, інтерв'юєр має ознайомитися з транскриптами інтерв'ю, які дослідник отримав у процесі свого дослідження. Це дозволить інтерв'юєру сформулювати додаткові релевантні запитання.

Що стосується рекомендованих Е. Онвоб'юзі питань для обговорення, то він акцентує увагу на таких темах, як досвід дослідника, особливості сприйняття ним учасників дослідження, їхня невербальна комунікація, особливості інтерпретації дослідником отриманих знахідок, вплив дослідження на дослідника, вплив дослідження на його учасників, етичні та політичні складнощі, непередбачені ускладнення або дилеми¹³⁷.

За переведення дебрифінгу у площину соціологічного тестування стає очевидною його актуальність на етапі теоретичної валідації, якщо для цього застосовують дослідження на кшталт кейс-стаді. При цьому дебрифінг має комплексний характер — він може проявити спірні аспекти щодо будь-якого виду теоретичної валідності. Також використання дебрифінгу, на нашу думку, можна вивести за межі якісних і змішаних досліджень. До того ж, в дослідженнях кількісного профілю його потенціал ми вбачаємо не меншим, ніж у якісних.

Це пов'язане з тим, що когнітивні зміщення (феномен самовпевненості, упередженість щодо доказів, евристика репрезентативності, помилка базової оцінки, доступність евристики, ілюзорний взаємозв'язок)¹³⁸ можуть виникати в будь-якому соціологічному дослідженні. У цьому контексті розвиток рефлексивності є необхідною умовою професійного зростання соціолога (особливо академічного). Рефлексивність тут слід розуміти як чутливість ученого до тих інтерпретацій, які він привносить у текст і джерела яких не дуже добре усвідомлює.

Орієнтирами дебрифінгу, який покращить рефлексивність дослідника (в тому числі й у межах кількісної парадигми), можуть бути такі запитання (Creswell, 2007, р. 180): 1) Чи є в дослідника будь-яка політична зацікавленість (у широкому значенні), яка впливає на оцінку, узагальнення й опис наукових фактів? 2) Чи не формує дослідник фінальні теоретичні положення в такому стилі, який знімає з нього відповідальність за його інтерпретацію? 3) Якою мірою аналіз дослідника є альтернативним здоровому глузду або домінуювальному дискурсу?

Ці запитання можна використовувати як для проведення дебрифінгу, коли дослідник інспектує себе самостійно, так і у груповому обговоренні пропону-

¹³⁷ Для ознайомлення з відповідними питаннями див.: (Onwuegbuzie, Leech, & Collins, 2008, р. 7)

¹³⁸ Докл. див.: (Майерс, 2003, с. 138–157).

ваних теоретичних положень. При цьому сфера їх застосування включає не тільки результати кейс-стаді, а й інші джерела теоретичної валідності.

§ 17.7. Висновки

Методологічні та методичні напрацювання досліджень змішаного типу містять багато цінних принципів і процедур, які суттєво збагачують і деталізують процес розробки комплексних вимірювальних інструментів у соціальних науках. Багато з них є релевантними в контексті розробки соціологічних тестів.

Перш за все методологічний інтерес становлять підходи (IDCV, MDVR, PEMI, TVMS), що визначають послідовність якісних і кількісних етапів розробки вимірювального інструментарію. Сьогодні через брак досвіду їх використання важко надати перевагу якомусь із них. Хоча правильніше було б говорити про типологізацію ситуацій, у яких виправданим є використання того або того підходу. На наш погляд, це є перспективним напрямком подальших методологічних пошуків.

Цінними також є окремі процедури, які використовують для розробки комплексних вимірювальних інструментів у дослідженнях змішаного типу. Так, їх можна використовувати, щонайменше, для очевидної валідації теоретичних напрацювань (дебрифінг), оцінки і коригування змістовної валідності (таблиця характеристик), прояснення композиційної валідності (нарративний профіль, бінаризація тем), критеріальної валідації (бінаризація тем) і визначення критеріальних рівнів або правил категоризації відповідей (нарративний профіль).

Особливої уваги, на нашу думку, заслуговує процедура дебрифінгу, яка може і повинна бути розширена за межі якісних досліджень.

Глава 18

Розробка соціологічних тестів у контексті кількісних досліджень

§ 18.1. Вступ

У практиці кількісних досліджень накопичено дуже великий досвід розробки комплексних вимірювальних інструментів, який охоплює такі аспекти, як формування і відбір індикаторів, проведення пілотажних досліджень, формування вибірок стандартизації, використання методів статистичного аналізу тощо. Однак з точки зору припущень прагматичного підходу до соціологічного вимірювання будь-яка з частин цього досвіду має пройти як концептуальну, так і методичну перевірку, перш ніж отримати схвальну оцінку. Ті практики, що вже мають статус рутинних у межах кількісної парадигми, не обов'язково ведуть до розробки валідних соціологічних тестів. Це пов'язане передусім із тим, що часто їх сприймають некритично і наділяють технологічним статусом — ніби їх правильне застосування саме по собі веде до необхідних результатів.

Ми ж вважаємо, що розробка соціологічних тестів не може бути зведена до технологічного рівня і, відповідно, в кожному конкретному випадку потребує концептуальної та методологічної рефлексії соціологічного (а не математичного) характеру, яка лише в разі її успіху веде до релевантності виконуваних дослідником дій.

У цій главі з урахуванням набутого досвіду розробки соціологічних тестів, ми акцентуватимемо увагу на низці процедур і технік, які допоможуть спрямувати наукову активність у межах кількісного етапу дослідження у правильне методологічне річище. Їх адекватне застосування передбачає постійну методологічну ретельність науковця. З огляду на це далі говоритимемо про (само)дебрифінг методологічної ретельності, спрямований на низку принципів аспектів розробки соціологічних тестів.

Використовуючи термін «(само)дебрифінг», підкреслимо той факт, що дослідник може здійснювати самоконтроль своєї методологічної та методичної активності шляхом постійної уваги до найпринциповіших питань, які стосуються виконуваної роботи, а також звертатися до колег із метою спільного їх обговорення. Очевидно, що другий варіант є прийнятнішим. Утім, він може бути реалізований лише в тому разі, якщо для цього існують об'єктивні умови.

(Само)дебрифінг методологічної ретельності щодо соціологічного тестування полягає в цілісному баченні процесу вимірювання, а також у постійній

увазі в цьому контексті до застосування дослідницьких технік/методів/стратегій. Останнє передбачає увагу дослідника до конкуруючих пояснень, розбудови номологічної мережі, збільшення інформаційного потенціалу методичних досліджень, а також орієнтацію не лише на перевірку статистичних гіпотез, а й на пошуковий аналіз даних.

§ 18.2. Цілісне бачення процесу соціологічного вимірювання

Цілісність бачення процесу соціологічного вимірювання полягає в урахуванні дослідником чотирьох складових цього процесу, а також взаємозв'язків між ними. Такими складовими є соціологічний тест, респонденти, релевантний соціальний контекст, а також сам дослідник, який розробляє методику і/або інтерпретує її результати.

Базовий взаємозв'язок устанавлюється між соціологічним тестом і респондентами під час опитування. Для його успішності респонденти повинні мати можливість і/або бажання повідомити інформацію¹³⁹, яка цікавить дослідника. У цьому разі йдеться про достовірність отриманої інформації (Шмел'єв, 1996, с. 154–155). По суті навіть використання надійної та валідної вимірювальної методики саме по собі не забезпечує отримання необхідних досліднику даних, якщо респонденти не хочуть повідомляти інформацію або з якихось причин не можуть цього зробити (Дембіцький & Любива, 2017а, с. 52)¹⁴⁰. Відповідні змішення належать до категорії помилок, не пов'язаних із вибіркою (McNabb, 2014, р. 42–46).

Якщо дослідник має обґрунтовані очікування щодо достовірності інформації, отриманої за допомоги соціологічного тесту, наступним кроком буде врахування широкого соціального контексту, в якому формується зв'язок «методика-респондент». Це передбачає обізнаність стосовно сприйняття методики респондентами загалом, а також поінформованість про особливості реагування на її індикатори в різних соціальних групах (принаймні, в релевантних з точки зору вимірювальної властивості). Для більшої ясності розглянемо два приклади.

¹³⁹ Про окремі причини неможливості респондентами повідомити інформацію, яка цікавить дослідника, див.: (Паниотто, 1986, с. 40–43).

¹⁴⁰ Для перевірки достовірності вимірювальних інструментів дослідницьким колективом на чолі з Є. Головахою було розроблено й адаптовано індекс суб'єктивної надійності респондента (Головаха и др., 2008), який слід розуміти як інтегральну оцінку респондентом самого себе як достовірного джерела соціальної інформації. Концепт суб'єктивної надійності містить такі компоненти, що підлягають самооцінюванню: 1) зрозумілість інструменту опитування або його частини; 2) достатність знань для відповіді на запитання чи блоку запитань; 3) інтерес до теми опитування; 4) упевненість у виборі найліпшого варіанта відповіді на поставлені запитання; 5) сталість власної думки; 6) незалежність думки; 7) можливість широко відповідати на поставлені запитання. При цьому вимірювання концепту суб'єктивної надійності може будуватися на самооцінках, які стосуються окремих запитань, блоків запитань або опитувальника загалом.

На прикладі застосування соціологічних тестів «Типи політичної культури» і «Типи політичної культури — II» розглянемо врахування широкого соціального контексту, на прикладі соціологічного тесту SCL-9-NR — сприйняття методики в різних релевантних групах.

У соціологічному тесті «Типи політичної культури» використовують набір тверджень щодо позитивного ставлення до демократії. При цьому не наведено ніякої конкретизації, тобто незрозуміло, про яку демократію йдеться. Така ситуація підштовхує респондента здійснити власну інтерпретацію спрямованості відповідних установок, а це, своєю чергою, значно ускладнює інтерпретаційні можливості дослідника на етапі опису отриманих результатів. Для перевірки коректності такої постановки питання в соціологічному тесті «Типи політичної культури — II» ми двічі використали набір тверджень щодо несприйняття демократії, явно вказавши при цьому два різні об'єкти установок — сучасні розвинені суспільства, з одного боку, й Україна — з іншого. Результати у цих двох групах тверджень виявилися суттєво відмінними.

Соціологічний тест SCL-9-NR спрямований на визначення рівня психологічного дистресу респондентів. Утім, як показали валідаційні дослідження, він не є чутливим для всіх категорій респондентів. Наприклад, показники тесту для респондентів, які за своїм діагнозом належать до категорії людей із шизофренією, шизотиповими і маячними розладами, а також звичайних респондентів, не відрізняються. Водночас тест є чутливим для широкого кола респондентів з іншими психічними розладами (розлади настрою, невротичні розлади, розлади особистості тощо). Також було з'ясовано особливості реагування на індикатори тесту серед респондентів із незадовільним станом здоров'я і серед тих, хто опинився в серйозній життєвій скруті.

У такому вигляді (взаємозв'язок між респондентами і соціологічним тестом у певному контексті) процес соціологічного вимірювання на початку використання конкретної методики для дослідника є «чорним ящиком». У подальшому головним завданням останнього є перетворення «чорного ящика» на зрозумілу йому систему, якої можна досягти завдяки проведенню відповідних методичних досліджень.

Отже, першим завданням (само)дебрифінгу є розуміння дослідником цілісного процесу соціологічного вимірювання і виконання ним усіх необхідних дій для прояснення використання соціологічного тесту в певному контексті в різноманітних групах. Це дозволить застосувати коректні формулювання методики і покращити її інтерпретаційний потенціал, а також визначити групи, які мають бути охоплені опитуванням.

§ 18.3. Формулювання конкуруючих пояснень

Конкуруючі пояснення являють собою дієвий механізм утілення наукового скептицизму з метою попередження когнітивних зміщень на тому етапі теоретичної роботи, коли дослідник визначається з теоретичною перспективою розгляду досліджуваного феномену, тобто здійснює очевидну валідацію соціологічного тесту.

Р. Їін описує дев'ять типів конкуруючих пояснень, три з яких можна використати в межах розробки соціологічних тестів: конкуруюча теорія (rival theory), конкуруюче теоретичне поле (super rival), «соціетальний конкурент» (societal rival) (Yin, 2009, p. 133–135). Для кращого пояснення їх суті звернемося до прикладів, пов'язаних із розглядом феномену психологічного дистресу в соціології.

Конкуруючі теоретичні підходи до феномену психологічного дистресу в соціології розглянуто в нашій нещодавній публікації (Дембіцький, 2018b). Л. Перлін досліджує процес стресу як набір послідовних елементів, що ведуть від стресорів до результатів стресу, одним із яких є психологічний дистрес. Натомість ми розглядаємо процес стресу як послідовність дискретних станів, які можна виокремити аналітично в кожний конкретний момент часу. Цим дискретним станам характерний набір узаємопов'язаних характеристик, одним із яких є психологічний (ди)стрес. Конкуруючими теоретичними підходами в цьому разі будуть і особливості змістовного наповнення психологічного дистресу. Хтось обмежується депресивною симптоматикою, хтось — і депресивною, і симптоматикою тривожності, а хтось інкорпорує більшу кількість компонентів.

Говорячи про конкуруюче теоретичне поле, маємо на увазі ширшу теоретичну перспективу, яка включає і пояснення досліджуваного феномену як своєї складової. Відповідним прикладом є наше теоретичне розуміння індивідуального благополуччя, в якому ми звертаємося до кількох принципових складових — психологічного дистресу, соціального благополуччя і соціального статусу. Перевагою звернення до ширшої теоретичної перспективи є, зокрема, деталізація ідей щодо конструктної та критеріальної валідазації.

Термін «соціетальний конкурент» описує значно глобальніші сили, які насправді визначають сутність досліджуваного явища і які не враховують у дослідженні. Зазначимо, що наш наступний приклад має суто гіпотетичний характер, який, утім, не позбавлений евристичності. Так, можна припустити, що особливості прояву психологічного дистресу в соціумі є результатом особливостей сучасного інформаційного суспільства. Це, зокрема, виявляється в постійному інформаційному навантаженні на індивіда, а також у динамічній мінливості мереж соціальної взаємодії, в які він включений. Приймаючи таку перспективу, необхідно повністю переосмислити підхід до вимірювання психологічного дистресу.

§ 18.4. Використання номологічної мережі

Наступним аспектом прояву методологічної ретельності є увага до номологічної мережі, частиною якої є досліджуваний феномен. Поняття номологічної мережі запропоновано у відомій статті Лі Кронбаха і Пола Міла (Cronbach & Meehl, 1955, p. 290–291) із метою опису коректної конструктної валідазації психологічних тестів. Для його правильного розуміння вони визначили шість фундаментальних наукових принципів.

1. Пояснення того, чим є певний об'єкт, потребує звернення до законів (laws), у межах яких цей об'єкт проявляється. Необхідно звертатися до взаємопов'язаної системи законів, які визначають теорію як номологічну мережу.

2. Закони в номологічній мережі можуть пов'язувати: а) спостережувані властивості або величини між собою; б) теоретичні конструкти зі спостережуваними властивостями або величинами; в) теоретичні конструкти між собою. Такі закони можуть бути як статистичними, так і детерміністичними.

3. Необхідною умовою того, щоб теоретичний конструкт був прийнятним з наукової точки зору, є його включення до такої номологічної мережі, в якій принаймні деякі складові є спостережуваними властивостями або величинами. При цьому певний теоретичний конструкт не визначає їх безпосередньо, а лише разом з іншими теоретичними конструктами. Це, своєю чергою, приводить до того, що між певним теоретичним конструктом і релевантним спостереженням може бути кілька рівнів номологічної мережі.

4. Деталізація знань про теоретичний конструкт — питання розробки номологічної мережі, в якій він проявляється, або, інакше, — покращення визначеності її компонентів, а також збільшення їх кількості.

5. Збагачення номологічної мережі є виправданим тоді, коли зв'язок між теоретичними конструктами й емпіричними спостереженнями стає очевиднішим і щільнішим. Якщо ж цей зв'язок в емпіричному вимірі відрізняється від теоретично очікуваного, дослідник має відчувати певну свободу щодо модифікації мережі.

6. Можемо сказати, що «операції», які дуже відрізняються якісно, «перекриваються» або «вимірюють один і той самий об'єкт», якщо їх розташування в номологічній мережі прив'язує їх до однакового теоретичного конструкту.

Беручи до уваги ці шість загальних принципів в контексті соціологічного тестування, пропонуємо розглядати будь-яку номологічну мережу за допомогою чотирирівневої моделі, де перший (найвищий) рівень відповідає системі взаємопов'язаних теоретичних конструктів, другий — процедурам операціоналізації, третій — релевантним вимірювальним методикам, четвертий (найнижчий) — процесу їх застосування в емпіричних дослідженнях і обробки відповідних результатів (рис. 18.1).

При цьому (і ми знову це наголошуємо) визначальний характер тут мають саме валідні теоретичні положення (перший рівень), що не спростовуються неконсистентними результатами вимірювання (четвертий рівень). Неконсистентні (суперечливі з точки зору теоретичних очікувань) результати або вказують на неправильність процедури операціоналізації, або свідчать про неприйнятність додаткового (контрольного) інструментарію, або ж збагачують теоретичні твердження, на яких базується автор соціологічного тесту. Розглянемо ці ситуації.

Неправильність операціоналізації може бути викликана різними причинами, серед яких, приміром, підбір індикаторів, які змістовно не охоплюють теоретичний конструкт, використання невдалих формулювань або «віяла» від-

повідей. Прикладом невдалої операціоналізації є формулювання як завдання методики, так і варіантів відповідей на неї, що використовують в SCL-90-R. Відповідні корективи, внесені нами, привели, зокрема, до покращення показників конструктивної валідності.

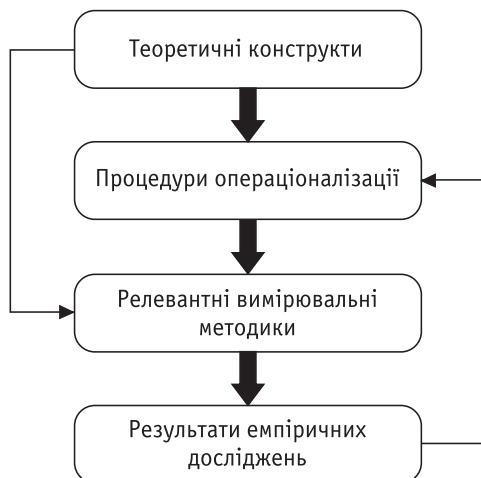


Рисунок 18.1. Чотирирівнева модель номологічної мережі

Неприйнятність додаткового інструментарію полягає в тому, що допоміжні теоретичні конструкти (які не є безпосередньою основою розробки соціологічного тесту) погано «схоплюються» або взагалі «не схоплюються» відповідними методиками, включеними дослідником до анкети. Про це можуть свідчити не тільки слабкі статистичні зв'язки, а й такі статистичні зв'язки, що є зависокими. В першому випадку слід звернути увагу на змістовне наповнення контрольної методики, а також на особливості її формулювань — популярність того чи того інструменту навіть у міжнародній практиці ще не свідчить про її прийнятність у вашому дослідженні (та й узагалі це не гарантує її наукової адекватності). Що стосується виявлення зв'язків високої щільності, то їх слід очікувати лише в разі використання методики-конкурента, тобто методики, заявленої для вимірювання того самого теоретичного конструкту. Якщо ж такий зв'язок демонструє не методика-конкурент, а допоміжна методика, слід звернутися до термінології, яку використовує автор допоміжної методики, і з'ясувати справжнє місце відповідних конструктів у контексті ваших теоретичних положень.

Збагачення теоретичних положень унаслідок неконсистентних результатів має місце тоді, коли результати дослідження вказують на більш тонкі особливості досліджуваного феномену, про які дослідник не здогадувався. Відповідним прикладом є конструктивна валідизація соціологічного тесту «Типи політичної культури», в межах якої ми, зокрема, виявили, що респонденти з демократичними й антидемократичними установками не відрізняються щодо підтримки авторитаризму. Це, як з'ясовано, не спростовує теоретичні акцен-

ти, покладені в основу вимірювання, а лише вказує на різні об'єкти установок. Так, респонденти схвально говорять про абстрактну ідеальну демократію, але потребують «сильної руки» в контексті національних реалій.

Згідно з поданою схемою, етапи розробки номологічної мережі відповідають етапам розробки соціологічних тестів. На теоретичному етапі дослідник концентрується на деталізації теоретичного конструкту, покладеного в основу розробки соціологічного тесту. Більшої уваги потребує внутрішня структура конструкту. Що стосується його зовнішніх зв'язків, то їх можна позначити лише поверхово для демонстрації включеності досліджуваного феномену в дискурсивний і/або інституційний контекст.

На етапі вибору інструмента вимірювання (шкали, індексу, класифікатора) дослідник зіставляє валідні теоретичні знання, реалізовані в теоретичному конструкті або їх сукупності, з релевантними техніками отримання емпіричної інформації від респондентів¹⁴¹.

Велику увагу в кількісній емпіричній соціології приділяють особливостям формулювання анкетних запитань, а також тому, як їх послідовність в анкеті впливає на результати опитування (так звана контекстуальна обумовленість відповідей). Оскільки це питання докладно досліджено й описано, ми не звертатимемося до нього в нашій роботі, а лише зазначимо авторів, які, на нашу думку, досягли статусу експертів у цій сфері. Так, С. Садмен, Н. Бредберн і Н. Шварц не тільки дали практичні рекомендації щодо формулювання різних типів питань (Садмен & Бредберн, 2002), а й виклали когнітивний підхід до аналізу й оцінки якості соціологічних анкет (Садмен, Бредберн, & Шварц, 2003). До цього підходу також звертався відомий російський дослідник Д. Рогозін (Рогозін, 2002).

За отримання першого варіанта соціологічного тесту дослідник може переходити до емпіричної валідазації. Її початок потребує звернення до теоретичної частини номологічної мережі для уточнення зовнішніх зв'язків досліджуваного феномену, що має на меті визначити особливості допоміжних теоретичних конструктів, а відтак і релевантні контрольні методики конструктної валідазації, а також цільові групи респондентів, результати опитування яких будуть використані для критеріальної валідазації. На цьому етапі, якщо змога, визначають і методики, що будуть використані для порівняльної валідазації.

Зрештою, на етапі збору й обробки даних дослідник здійснює оцінку якості переходу від теоретичного до емпіричного рівня розробки соціологічного тесту. У разі її успішності номологічна мережа збагачується за рахунок досліджень із визначення критеріальних рівнів або розробки класифікатора відповідей. Також вона може бути деталізована за рахунок інформації щодо особливостей взаємозв'язку між релевантними конструктами в різних соціальних групах.

¹⁴¹ Цей момент докладніше буде розглянутий у наступному параграфі.

§ 18.5. Збільшення інформаційного потенціалу методичних досліджень

Протилежністю інформаційного збагачення дослідження є те, що ми називаємо мінімальною активністю дослідника. Остання виявляється у проведенні методичних опитувань серед легко доступних груп, нехтуванні соціально-демографічним наповненням вибірки, некритичному підході до вибору методів статистичної обробки даних тощо.

Сумнозвісна «валідизація на студентах» є класичним прикладом мінімальної активності. При цьому ми не заперечуємо прийнятність відповідних опитувань загалом. Очевидно, що таке заперечення не стосується випадків, коли розробляють вимірювальний інструментарій, призначений саме для опитування студентів. Окрім того, відповідне методичне дослідження може мати статус допоміжного. В останньому випадку його результати слід розглядати в контексті результатів інших опитувань¹⁴².

Отже, завдяки чому можна досягти збільшення інформаційного потенціалу методичних досліджень? Перш за все воно передбачає забезпечення соціально-демографічної гетерогенності вибірки. Цього досягають завдяки формуванню таких вибірок методичного опитування, які є максимально гетерогенними за базовими і специфічними з точки зору вимірюваної властивості характеристиками. Ураховуючи досвід масових опитувань в Україні, до базових характеристик ми відносимо стать, вік, освіту, тип поселення й область проживання респондента¹⁴³. Звичайно, в межах методичного опитування включення в дослідження всіх цих характеристик є неможливим. Тому пропонуємо орієнтуватися принаймні на стать і вік, а інші характеристики додавати наскільки можливо. Що стосується специфічних характеристик, то їх важливість визначає специфіка самого дослідження. Фактично такі характеристики мають статус ключових. Наприклад, у межах розробки соціологічного тесту психологічного дистресу до таких, зокрема, належав статус психічного здоров'я респондентів (хворий/здоровий). У певних випадках базові характеристики можуть перетворюватися на специфічні. Приміром, за розробки соціологічного тесту геополітичних орієнтацій регіон проживання є саме специфічною ха-

¹⁴² Як до прикладу звернемося до онлайн-опитування, яке ми провели разом із кафедрою методології та методів соціологічних досліджень КНУ ім. Тараса Шевченка. Особливістю цього опитування було те, що в ньому брала участь переважно студентська молодь, здебільшого респонденти жіночої статі (приблизне співвідношення — 3 до 1). Головним призначенням дослідження була перевірка впливу соціальної бажаності на результати SCL-9-NR серед респондентів, які не мають психіатричного діагнозу. Щоб висновки були коректними, ми перш за все перевірили вплив віку на результати тесту (для порівняння використано результати інших двох опитувань). Що стосується взаємозв'язку між соціальною бажаністю (L-шкала) і результатами SCL-9-NR, для порівняння використано відповідні результати, отримані серед різних груп респондентів із психіатричним діагнозом. Отже, будь-які результати цього та інших опитувань розглянуто в порівняльній перспективі.

¹⁴³ Див., напр.: (Панина, 2012, с. 315).

рактистикою, оскільки регіональна специфіка в Україні, як відомо, тісно пов'язана із зовнішньополітичними вподобаннями населення.

З точки зору врахування базових і специфічних характеристик оптимальною є випадкова вибірка з квотним скринінгом на останньому етапі¹⁴⁴. Дослідник визначає необхідний розмір вибірки (про це йтиметься далі) і відповідно до нього — розмір квот, який дорівнює частці від поділу загального розміру вибірки на кількість категорій тієї чи тієї соціально-демографічної характеристики. Наприклад, якщо стать є важливою з точки зору дослідження характеристикою, то за вибірки 100 спостережень потрібно опитати 50 чоловіків і 50 жінок.

Якщо отримувані квоти виявляються замалими, досліднику необхідно збільшити загальний обсяг вибірки. У разі обмеженості коштів на проведення методичного опитування перевагу знову ж таки необхідно надати гетерогенності вибірки, навіть в умовах отримання квот недостатньо великого розміру. Звернімося, наприклад, до нашого дослідження з розробки соціологічних тестів геополітичних орієнтацій серед населення України («RF-георол» та «EU-георол»).

Проводячи дослідження за власні кошти і маючи обмежений бюджет, ми зупинилися на вибірці 100 спостережень. Базовими параметрами гетерогенності ми визначили стать і вік (три вікові категорії — молодь, середній і старший вік). Як специфічну характеристику визначено регіон проживання. Опитування було замовлено у трьох містах — Києві (34 інтерв'ю), Одесі (33 інтерв'ю) і Львові (33 інтерв'ю). У кожному місті інтерв'ю було поділено на шість майже однакових за кількістю груп (відповідно до статі та віку). У результаті отримано максимально гетерогенну вибірку в контексті наших фінансових можливостей.

Перевага такого підходу до формування вибіркової сукупності полягає в тому, що дослідник отримує можливість аналізувати вплив релевантних факторів на вимірювальну ознаку, контролюючи інші важливі характеристики. Саме за таких умов можна з'ясувати вимірювальні особливості соціологічного тесту. Логіка узагальнення результатів з такої вибірки є аналітичною, а не статистичною. Дослідник не намагається створити модель генеральної сукупності, він намагається зафіксувати певні релевантні контексти. У наведеному прикладі такими контекстами були різні геополітичні вподобання. Незважаючи на те, що в генеральній сукупності різні геополітичні вподобання можуть мати різний підсумковий вплив унаслідок їх більшої або меншої репрезентованості серед населення, в методичному дослідженні вони «урівнюються» в силі свого впливу задля фокусування на потенційних результатах використання методики.

Для рекомендацій щодо розміру вибірки методичного дослідження звернімося до різних видів валідності та надійності. З власного досвіду можемо сказати, що для перевірки конструктивної валідності достатнім є використання

¹⁴⁴ Така вибірка, наприклад, застосовується у соціологічному моніторингу «Українське суспільство» (Інститут соціології НАН України) (Панина, 2012, с. 315).

100 спостережень. Окрім того, оскільки один із пріоритетів емпіричної валідації — перевірка критеріальної валідності з метою оцінки вимірювальної чутливості методики, бажаний розмір кожної критеріальної групи дорівнює 50 спостереженням. При цьому, як ми вже зазначали, дотримання розміру таких груп має нижчий пріоритет порівняно з досягненням загальної гетерогенності. Тому ми не проводили по 50 спостережень у двох містах, а зупинилися на дрібніших кластерах у трьох.

За визначення критеріальних груп головну увагу мало бути приділено як визначенню тих, хто демонструє теоретично очікувану специфіку у відповідях, так і тих, хто, незважаючи на свою специфіку в контексті вимірювальної властивості, демонструє її загальний рівень за застосування соціологічного тесту. Такий підхід дозволить визначити як сильні боки методики, так і обмеження в її використанні. Наприклад, за розробки індексу психологічного дистресу SCL-9-NR ми з'ясували, що він дозволяє зафіксувати підвищений або високий рівень психологічного дистресу лише в окремих групах респондентів із психіатричним діагнозом.

Своєю чергою, прогностична валідність потребує тривалого використання вимірювального інструменту в широкомасштабних репрезентативних опитуваннях і врахування результатів його використання під час розробки сценаріїв розвитку подій на різних рівнях соціальної взаємодії. Розмір унеску відповідних результатів до таких сценаріїв, а також зіставлення останніх із справжніми подіями, дозволить з'ясувати евристичний та інтерпретаційний потенціал методики.

Перевірку надійності¹⁴⁵ зазвичай зводять до демонстрації внутрішньої узгодженості за допомоги Альфа Кронбаха. Це зумовлено як загальною методологічною модою в сучасних соціальних науках, так і прагматичними міркуваннями: організаційною доступністю — необхідністю проведення лише одного методичного опитування. Проте перевірка внутрішньої узгодженості має сенс лише в тому разі, коли розробляють або адаптують вимірювальну шкалу. Відповідні результати можуть бути отримані під час конструктивної валідації без необхідності змінювати обсяг вибірки.

Що стосується інтегральних індексів, то перевірка внутрішньої узгодженості не є доцільною. Так, оскільки окремі частини інтегрального індексу є важливими, але не обов'язково пов'язаними між собою причинними факторами вимірювальної властивості, оскільки демонстрація взаємозв'язку між індикаторами індексу не має методологічної цінності.

За перевірки ретестової надійності треба розуміти доцільність цієї процедури. Це пов'язане з тим, що у фокус соціології часто потрапляють доволі мінливі характеристики, що накладає суттєві обмеження на можливість перевірки стабільності відповідних показників. Коли ж досліджується стабільна властивість, демонстрація її ретестової надійності є дуже цінною як для шкали, так

¹⁴⁵ Докладний огляд статистичних інструментів перевірки надійності соціологічного вимірювання подано в: (Паниотто, 1986, с. 55–76).

і для індексу. При цьому прийнятний розмір вибірки становить 50 спостережень.

Звичайно, аргументація щодо розмірів вибірок із позиції партикулярного досвіду вдовольнить не багатьох. З огляду на це розглянемо додаткові аргументи, що ґрунтуються на методах аналізу потужності. Останні є статистичним інструментарієм, який використовують для планування емпіричних досліджень завдяки співвідношенню таких аспектів статистичного аналізу, як потужність (імовірність відхилити хибну нульову гіпотезу), розмір вибірки, рівень достовірності (імовірність відхилити істинну нульову гіпотезу) і розмір ефекту (наприклад, сила зв'язку або величина відмінностей у стандартизованому вигляді)¹⁴⁶.

Отже, як ми вже зазначали, в межах конструктної валідазації мінімальний обсяг вибірки становить 100 спостережень. За рівня достовірності 0,05 і потужності 0,9 сила зв'язку (фактично розмір ефекту), виявлена в методичному дослідженні, має дорівнювати близько 0,29. Тобто ми виходимо з того, що базовим методом конструктної валідазації буде кореляційний аналіз (односторонній критерій). Що стосується критеріальної валідазації, то як базовий метод запропоновано порівняння середніх значень (також односторонній критерій) або дисперсійний аналіз. У цьому разі за використання груп розміром 50 спостережень і аналогічних достовірності та потужності розміри ефектів мають становити відповідно 0,59 (*d*) і 0,29 (*f*). Застосовуючи правила інтерпретації розмірів ефектів, запропоновані у відомій праці Джакоба Коена (Cohen, 1988, р. 88), можемо сказати, що в методичному дослідженні необхідно продемонструвати хоча б середні розміри ефектів. Вважаємо, що це є правильним з точки зору обґрунтування якості соціологічного тесту. Звісно, такі параметри, як потужність і достовірність, є конвенційними. У своїй праці Коен для прикладів найчастіше використовує потужність, що дорівнює 0,8. Якщо орієнтуватися на цей показник, то бажані розміри ефектів дещо зменшаться: $r = 0,25$; $d = 0,5$; $f = 0,26$.

У межах перевірки ретестової надійності рекомендований розмір вибірки 50 спостережень (за потужності 0,95 і достовірності 0,99) дає бажану силу зв'язку лише 0,52. Тобто такий розмір вибірки може бути навіть дещо зavelиким. Якщо орієнтуватися на розмір ефекту, що дорівнює 0,7 (це значно ближче до бажаних показників ретестової надійності), буде достатньо вибірки 24 спостереження. Однак це тільки статистичні міркування. Пам'ятаючи про необхідність гетерогенності вибірки, не варто надто зменшувати її обсяг.

§ 18.6. Пошуковий аналіз даних

Невід'ємною складовою емпіричної валідазації соціологічного тесту є ретельне проведення статистичного аналізу з особливим акцентом на його пошуковій частині. Це пов'язане з тим, що дуже часто отримані дані містять значно

¹⁴⁶ Докладніше про аналіз потужності й особливості його використання в середовищі R див.: (Кабаков, 2014, с. 335–355).

більше релевантної інформації, ніж може здатися. Звичайно перевірка статистичних гіпотез також є важливою частиною аналітичної роботи, втім, з нею все доволі зрозуміло. А от у частині пошукового аналізу дослідник стикається радше з мистецтвом, аніж із технологією. У цьому сенсі кожний новий масив може бути несхожим на попередній не тільки з точки зору використовуваних змінних, а і з точки зору того, на що необхідно звернути особливу увагу і як це зробити.

На сучасному етапі, коли йдеться про глибокі знання у сфері обробки даних, заведено говорити про так звану науку про дані (data science). У сучасному розумінні вперше цей термін концептуалізував Дрю Конвей (Conway, 2010). Він уклав у його зміст три складові — змістовний досвід, навички управління даними, а також знання з математики і статистики. Цікава і найудаліша, на наш погляд, альтернативна інтерпретація цього поняття належить Гаррету Гролемунду, який виходить із таких складових, як програмування, наукове мислення й осмислення даних (Grolemund, 2014, р. 185). Програмування виконує такі завдання, як зберігання даних, отримання доступу до них, трансформація даних і автоматизація завдань. Наукове мислення передбачає володіння статистичними висновками, науковим методом і оперування формальною логікою. Осмислення даних можна досягти завдяки традиційним та інноваційним методам візуалізації, а також за допомоги описової статистики.

Якщо «винести за дужки» предметний досвід, то інструментарій, необхідний для оперування складовими, зазначеними в обох указаних підходах до розуміння «науки про дані», найповніше реалізований у мові програмування R. Власне, її ми і рекомендуємо як основне середовище обробки даних у процесі розробки соціологічних тестів. Проте викладення основ мови програмування виходить за межі нашої монографії¹⁴⁷.

Нам необхідно сфокусуватися на предметному досвіді обробки результатів методичних досліджень, що є невід’ємною частиною розробки соціологічних тестів. Зауважимо, що далі ми не претендуємо на вичерпний перелік відповідних орієнтирів обробки даних, оскільки орієнтуємося перш за все на власний досвід, отриманий під час розробки таких соціологічних тестів, як SCL-9-NR, СТ ТПК-II, «RF-geopol» і «EU-geopol», а також роботи з результатами інших методик, які були названі соціологічними тестами їх авторами (Інтегральний індекс соціального самопочуття, Типи політичної культури) або які можна такими вважати (CES-D-8, L-шкала, Шкала громадської активності). Загалом накопичення предметного досвіду обробки даних у рамках розробки соціологічних даних перебуває на початковому етапі, а отже, потребує подальших наукових пошуків.

Підсумувавши наявний досвід, ми виокремили такі аналітичні орієнтири.

1. *Увага до розподілу результатів застосування соціологічного тесту: чи відповідає емпіричний розподіл теоретичним очікуванням щодо його форми?*

¹⁴⁷ Для вивчення мови програмування R можемо порекомендувати такі джерела: (Кабаков, 2014; Мاستицкий, 2015; Уикем, 2018).

Приклад: виявлення факту значного відхилення розподілу результатів тесту стало вихідною точкою для розробки тесту, зміна формулювань якого дозволила отримати розподіл, значно ближчий до нормального (Дембицкий, 2016e); проте нормальний розподіл не є еталонним; так, значна правостороння асиметрія результатів тестів аномії та соціального цинізму, включених в інструментарій соціологічного моніторингу «Українське суспільство», не є теоретично суперечливою з огляду на стан загальної недовіри і кризи норм в українському суспільстві.

Необхідні передумови: вибіркова сукупність необхідного розміру (мінімально 100 спостережень, бажано — 200 і більше).

2. Оцінка розподілів окремих індикаторів соціологічного тесту: чи були проінтерпретовані індикатори вимірального інструменту і наскільки така інтерпретація вписується у відповідний соціальний контекст?

Приклади: 1) аналіз розподілів відповідей на індикатори шкали громадської активності, який указав на два важливі аспекти — інертність і контрпродуктивність політичної системи України, а також досвід населення у протидії злочинній владі (Дембицкий, 2016d, с. 58–59); 2) аналіз розподілів відповідей на індикатори соціологічного тесту «Типи політичної культури — II», який, зокрема, продемонстрував значно меншу виразність антидемократичних суджень, якщо йдеться про розвинені країни, а також протилежну ситуацію, коли йдеться про Україну (Дембицкий, 2018a, с. 38–39).

Необхідні передумови: достатня обізнаність дослідника з тією суспільною сферою або інституційним контекстом, що визначає особливості відповідей респондентів на конкретні індикатори методики.

3. Коригування наявної вибіркової сукупності, коли її обсяг чинить безпосередній вплив на статистичну значимість результатів або є релевантним з точки зору порівняння: чи бере дослідник до уваги «фактор вибірки» у процесі статистичного аналізу?

Приклади: 1) зменшення обсягу вибірок за використання критерію Шапіро–Вілка для перевірки розподілу на нормальність, оскільки цей критерій є дуже чутливим до розміру вибірки (чим більша вибірка, тим вища ймовірність відхилення нульової гіпотези) (Дембицкий, 2018, с. 14–15); 2) корекція розміру вибірки за порівняння показників надійності L-шкали в нашому дослідженні (дані соціологічного моніторингу «Українське суспільство») з відповідними показниками іншого дослідження, яке мало менший обсяг вибірки (в іншому разі порівняння не було б коректним).

Необхідні передумови: достатньо великий початковий розмір вибірки; чи є вибірка достатньо великою, визначає дослідник у контексті конкретного емпіричного дослідження.

4. Інтеграція соціально-демографічної специфіки: чи проведено оцінку соціально-демографічного впливу, а також чи вироблено механізми її врахування в межах використання соціологічного тесту?

Приклад: визначення особливостей відповіді на індикатори L-шкали в шести демографічних групах, визначених на основі статі та віку, з подальшим об-

рахунком відповідних вагових коефіцієнтів (кожний індикатор отримав шість можливих вагових коефіцієнтів залежно від тієї демографічної групи, до якої належить респондент) (Дембіцький, 2017b, с. 107–110).

Необхідні передумови: масив даних, отриманий із використанням загальнонаціональної вибірки, який містить релевантну соціально-демографічну інформацію.

5. Поглиблена увага до суперечливих результатів: чи дослідник доклав зусиль, щоб з'ясувати суть суперечностей, виявлених у процесі статистичного аналізу?

Приклад: поглиблений аналіз даних у межах дослідження категоризації респондентів на основі результатів соціологічного тесту «Типи політичної культури»; його відправною точкою став той факт, що респонденти, які декларували демократичні установки, більшою мірою були схильні до підтримки авторитаризму, ніж респонденти, які не підтримують демократію; проведений аналіз дозволив сформулювати низку альтернативних пояснень цього феномену, а також став відправною точкою розробки соціологічного тесту «Типи політичної культури — II».

Необхідні передумови: достатня обізнаність дослідника з суспільною сферою або інституційним контекстом, що визначає особливості відповідей респондентів на конкретні індикатори методики; деталізація соціологічного інструментарію (анкети дослідження), яка дозволяє перевірити припущення дослідника.

6. Урахування попереднього досвіду (за перевірки валідності раніше розроблених вимірювальних інструментів): чи ознайомився дослідник, а також чи використовує він у межах статистичного аналізу важливі змістовні та структурні аспекти методики, яка лежить у фокусі його інтересу?

Приклади: 1) аналіз праць європейських науковців, які оцінювали якість шкали CES-D після її застосування в Європейському соціальному дослідженні; як виявилось, європейські колеги повністю проігнорували структуру досліджуваного явища (депресивної симптоматики), яку було покладено в основу методики її авторкою і яка була докладно вивчена в американських соціальних дослідженнях до проведення Європейського соціального дослідження; 2) дослідження надійності L-шкали, проведене А. Мягковим, яке було використано як вихідна точка нашого дослідження.

Необхідні передумови: доступ до матеріалів провідних періодичних видань із соціальних наук.

7. Застосування статистичних методів «головного конкурента» (за обґрунтування порівняльної валідності): чи було використано ті самі методи статистичного аналізу, що і під час обґрунтування якості конкуруючої методики?

Приклад: порівняльна валідизація соціологічного тесту SCL-9-NR, побудована на порівнянні його результатів із результатами методики вимірювання психологічного дистресу К10; оскільки для демонстрації вимірювальної якості останньої використано спеціальний підхід до оцінки ефективності бінарної класифікації (ROC-curve), цей самий підхід використаний для порівняння вказаних вимірювальних інструментів.

Необхідні передумови: внесення змін до дизайну та інструментарію методичного дослідження, які уможливають застосування статистичних методів «головного конкурента».

8. *Використання еталонних груп: чи є в розпорядженні дослідника дані стосовно еталонної групи і якщо є, чи використано їх у процесі аналізу?*

Приклад: аналіз вимірювальних особливостей шкали громадської активності за допомоги звернення до змінних, що пов'язані з громадською активністю та пасивністю серед населення України (оцінені на основі репрезентативної вибірки); зіставлення показників різних соціально-демографічних груп із цими показниками дозволило провести конструктну і критеріальну валідацію.

Необхідні передумови: дані репрезентативних досліджень, що дозволяють обрахувати еталонне значення; попередні ґрунтовні знання щодо показників еталонних груп.

9. *Аналіз, спрямований на виявлення критеріальних рівнів тесту (за відсутності або неможливості провести спеціальні методичні дослідження): чи існують необхідні дані в масиві, а якщо існують, чи проаналізовані вони дослідником?*

Приклади: 1) використання даних про сферу зайнятості респондента для перевірки оригінального правила щодо визначення нещирості респондентів за допомоги L-шкали; 2) визначення критеріальних рівнів Інтегрального індексу соціального самопочуття за допомоги моделювання показників, які характеризують ідеально-типових респондентів.

Необхідні передумови: достатня деталізація соціологічного інструментарію (анкети дослідження).

10. *Використання різних альтернатив підсумкових значень соціологічного тесту: чи використав дослідник усі релевантні варіанти підсумкових значень методики?*

Приклади: 1) звернення як до індексів, так і до класифікаційних категорій у роботі з результатами соціологічних тестів «RF-geopol» і «EU-geopol»; 2) аналіз використання індексів і класифікаційних категорій Інтегрального індексу соціального самопочуття.

Необхідні передумови: придатність методики для розрахунку альтернативних варіантів підсумкових значень.

11. *Акцент на порівняльному аналізі: чи використав дослідник усі доступні йому релевантні джерела (різні типи підсумкових значень, різні соціально-демографічні групи, різні статистичні методи) для проведення статистичного аналізу в порівняльному розрізі?*

Приклади: 1) порівняльний аналіз різних соціальних груп, для яких характерний підвищений або високий рівень психологічного дистресу, з метою розробки критеріальних рівнів соціологічного тесту; 2) порівняльний аналіз різних типів кластеризації респондентів (ідеальна точка поділу, кластерний аналіз, концептуальна типологія, латентний класовий аналіз) для визначення серед них оптимального в межах застосування соціологічного тесту «Типи політичної культури».

Необхідні передумови (залежно від типу порівняльного аналізу): придатність методики для розрахунку альтернативних варіантів підсумкових значень; наявність релевантних даних щодо різних соціально-демографічних груп; альтернативні погляди на можливий статистичний інструментарій обробки результатів соціологічного тесту.

12. *Гнучкість в інтерпретації отриманих результатів: чи зроблено спроби формулювання альтернативних пояснень щодо найважливіших результатів статистичного аналізу?*

Приклад: оцінка факторної структури шкали емоційного благополуччя на підставі даних Європейського соціального дослідження з точки зору трьох альтернативних пояснень (двокомпонентність емоційного досвіду, різне реагування на різні теми індикаторів методики, необхідність розглядати не один феномен, а два); у подальшому альтернативні пояснення можуть бути покладені в основу додаткового статистичного аналізу.

Необхідні передумови: достатня обізнаність дослідника з феноменом, що вивчається.

13. *Пріоритет концептуальних міркувань над статистичними: чи керується вчений змістовно-соціологічними припущеннями у процесі статистичного аналізу?*

Приклади: 1) розробка альтернативного скороченого варіанта Інтегрального індексу соціального самопочуття, в якому акцентовано на змістовній збалансованості, а не внутрішній узгодженості, як у попередній скороченій версії; 2) відмова від використання статистичних міркувань як основних за відбору індикаторів соціологічного тесту «Типи політичної культури — II»; на противагу цьому проведено експертне опитування з метою оцінки змістовної гетерогенності сукупності антидемократичних суджень; відбір індикаторів здійснено з урахуванням результатів експертного опитування і концептуальних міркувань щодо феномену політичної культури.

Необхідні передумови: визнання математичної статистики підпорядкованим соціології інструментом, а не методологічним орієнтиром у соціальних науках.

14. *Вписування результатів тестування в релевантну соціальну дійсність: чи інтегровані результати використання соціологічного тесту в ширші теоретичні положення, які характеризують поточну ситуацію?*

Приклад: використання результатів шкали громадської активності для формулювання можливих сценаріїв розвитку національної держави в Україні, а також оцінки їх імовірності.

Необхідні передумови: масив даних, отриманий із використанням загальнонаціональної вибірки, який містить релевантну демографічну інформацію; достатня деталізація опитувального інструментарію; достатня обізнаність дослідника з феноменом, який він вивчає.

Отже, орієнтири щодо аналізу даних, отриманих у процесі емпіричної валідації соціологічного тесту, є доволі неоднорідними. Перш за все їх можна диференціювати відповідно до двох вимірів: специфічність/загальність і про-

стота/комплексність у застосуванні. Під специфічними слід розуміти орієнтири, які належать лише до певних ситуацій, що не виникають в обов'язковому порядку. Відповідно, загальні орієнтири є актуальними в будь-якому разі. Простота свідчить про певну рутинність (початкову зрозумілість) необхідної статистичної обробки, натомість комплексні орієнтири передбачають розв'язання творчих завдань, які не мають чітко визначених рішень. Деякі орієнтири можуть вести як до простих, так і до комплексних варіантів обробки даних. Результати цієї типології подано далі (табл. 18.1).

Таблиця 18.1

**Типологія орієнтирів аналізу даних у межах емпіричної валідації
соціологічних тестів**

Ступінь універсальності орієнтирів	Ступінь складності реалізації орієнтирів		
	Прості	Прості/комплексні	Комплексні
Специфічні	№ 2, 3, 8	№ 7	№ 4, 5, 9, 10, 14
Загальні	№ 1	№ 6, 12, 13	№ 11

§ 18.7. Короткі підсумки

Наш огляд застосування кількісної парадигми для розв'язання завдань конструювання соціологічних тестів висуває на передній план дві теми — прагматизм і методологічну ретельність. Прагматизм указує на першочергову увагу до напрацювання методологічного досвіду, який є критерієм оцінювання ступеня прийнятності різноманітних кількісних технік і методичних принципів. Своєю чергою, методологічна ретельність є тим когнітивним інструментом, який дозволяє ефективно структурувати такий досвід, а також підтримувати постійний ефективний зв'язок між перевіреними методологічними установками і конкретною методичною практикою.

Глава 19

Рекомендації до розробки соціологічних тестів

На наш погляд, розробка соціологічних тестів завжди передбачає методологічний пошук. З огляду на це дуже важко давати рекомендації, які б остаточно структурували цей процес. Тому наприкінці нашої роботи ми маємо надати поради, в яких реалізовано баланс між нашим партикулярним досвідом і тими загальними завданнями, які має виконати соціолог, що звертається до розробки соціологічних тестів.

Як ми зазначали в першому розділі, розробка соціологічного тесту має чотири етапи: теоретична валідація, вибір різновиду вимірювального інструменту, емпірична валідація, розробка критеріальних рівнів. Нижче ми розглянемо кожний із них.

§ 19.1. Теоретична валідація: завдання, ситуації та методи

У контексті соціологічного тестування метою теоретичної валідації є уточнення і деталізація теоретичного конструкту, який буде покладено в основу вимірювальної методики. Завдання цього етапу є такими: 1) аргументувати, що теоретичний конструкт репрезентує об'єкт вимірювання, а цей об'єкт, своєю чергою, належить або до оточуючої соціальної дійсності (соціальний феномен), або до психічного стану індивіда (внутрішній стан), або до його особистісної якості (латентної змінної); 2) визначити принципово важливі складові теоретичного конструкту; 3) мірою можливості відтворити внутрішню структуру теоретичного конструкту; 4) визначити ключові зовнішні зв'язки теоретичного конструкту.

Специфіка розв'язання цих завдань пов'язана перш за все з тим, яку дослідницьку стратегію використано для теоретичної валідації. У загальнішому вигляді тут є дві можливі ситуації — використання кейс-стаді або будь-яких інших джерел. Докладно проведене кейс-стаді має розв'язувати всі чотири вказані завдання.

Якщо ж дослідник звертається до інших джерел (професійний досвід, соціологічні концепції, універсальний досвід або соціологічні аксіоми), то відповідна інформація може потребувати додаткових методичних засобів її перевірки й уточнення. Як мінімум ці джерела мають надавати інформацію, що дозволяє виконати принаймні перше завдання, тобто обґрунтувати очевидну

валідність теоретичного конструкту. У цій ситуації також може допомогти проведення дебрифінгу. Що стосується інших завдань, то їх виконання залежить від ступеня деталізованості інформації з відповідного джерела.

За необхідності деталізувати, структурувати, а також визначити зовнішні зв'язки теоретичного конструкту дослідник має провести роботу з ключовими інформантами (фахівці з релевантним професійним досвідом, експерти з відповідних питань).

Друге завдання (змістовна валідація) у цьому разі може бути полегшена завдяки використанню таблиць характеристик. Загальною стратегією роботи з ключовими інформантами є узгодження концептів¹⁴⁸. Вона є особливо придатною для розв'язання другого і третього завдань теоретичної валідації. При цьому вхідні дані (твердження) можна формулювати як за участі ключових інформантів, так і запозичувати з інших джерел.

Прикладом формулювання тверджень самими учасниками ситуації є розробка інструментарію для (само)оцінювання типу студентів з точки зору зусиль, яких вони прикладають під час навчання. Так, у відповідному дослідженні в межах використання узгодження концептів саме студенти визначали поведінкові практики і в подальшому структурували їх (Дембицкий, 2016с, с. 90–96).

Проте залучати ключових інформантів на всіх етапах узгодження концептів доволі складно з організаційної точки зору. Тому в багатьох випадках вважаємо за оптимальне самостійне визначення змістових елементів теоретичного конструкту, а експертів-учасників залучати до узгодження концептів на етапі структурування концептуальних елементів. За приклади слугують дослідження культурних кодів у сфері споживання, базових психологічних потреб (концепція Г. Мюррея) і антидемократичних установок.

За дослідження культурних кодів реклами¹⁴⁹ як стимульовального матеріалу використано зразки комерційної реклами глянцевого журналу. У подальшому завдяки участі ключових інформантів було виокремлено п'ять культурних кодів (краса і сексуальність, сімейне щастя і моральні почуття, багатство і розкішне життя, здоровий спосіб життя, комфорт), а також визначено взаємозв'язки між ними.

Узгодження концептів із застосуванням психологічних потреб, визначених Г. Мюрреем¹⁵⁰ (Дембицкий, 2016а), продемонструвало два виміри детермінант людської поведінки. Перший стосується визначальної мотивації — досягнення успіху або уникнення невдач, другий — базових стратегій поведінки — співробітництва або суперництва.

¹⁴⁸ В оригіналі — концептуальне картографування (concept mapping). Докладно про використання цієї дослідницької стратегії див.: (Дембицкий, 2016с, с. 125–149).

¹⁴⁹ Аналіз проведений в межах дисертаційного дослідження А. Кліменкової. Відповідні результати застосування стратегії узгодження концептів підсумовано в: (Дембицкий, 2015а).

¹⁵⁰ Докладніше про перелік цих потреб див.: (Клейман, 2015: с. 138–139).

Аналіз антидемократичних установок¹⁵¹ (Дембицкий, 2017b) дозволив уточнити їх смислову гетерогенність. Так, було виявлено чотири групи тверджень — колективістські, антидемократичні, тоталітарні й авторитарні. У подальшому це значно полегшило розробку соціологічного інструментарію.

Отже, блок-схема наукової роботи на цьому етапі матиме такий вигляд (рис. 19.1):

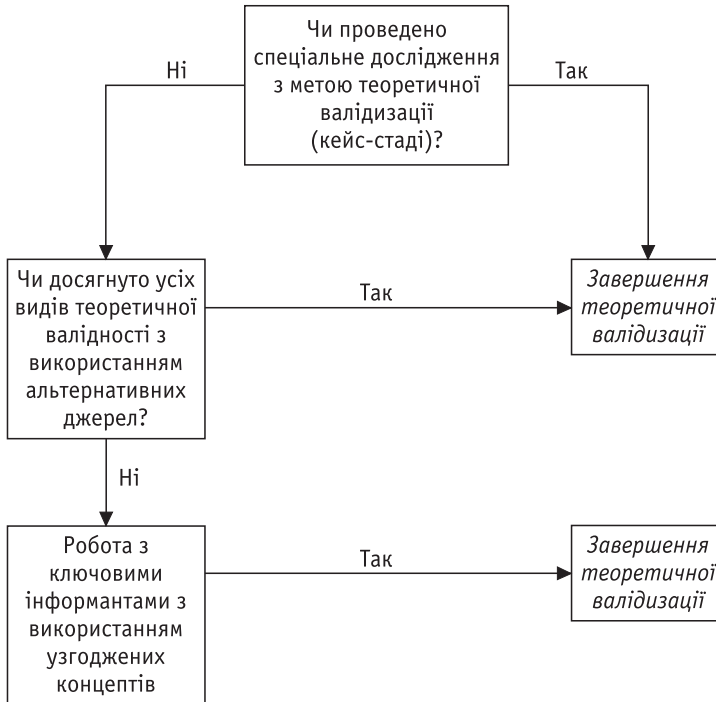


Рисунок. 19.1. Блок-схема теоретичної валідації

§ 19.2. Вибір різновиду вимірювального інструменту: принципи

На цьому етапі розробки соціологічного тесту необхідно обрати між побудованими шкали, індексу і класифікатора. При цьому між шкалою та індексом, з одного боку, а також класифікатором — з іншого, немає принципової відмінності (в тому сенсі, що дуже часто вони можуть бути побудовані на основі одних і тих самих індикаторів). При цьому шкала з індексом є несумісними. Отже, розглянемо принципи вибору між зазначеними інструментами.

Якщо об'єктом вимірювання є особистісна риса респондентів (латентна змінна), то дослідник конструє шкалу. В цьому разі інструмент вимірювання є соціологічним лише у значенні його придатності для використання в масових опитуваннях (сам об'єкт вимірювання є психологічним). Тому виникає

¹⁵¹ Відповідні твердження сформульовано понад 10 років тому за дослідження політичної культури українського суспільства під керівництвом Є. Головахи.

питання, в яких випадках шкали можуть бути спрямовані на вимірювання саме соціальних феноменів?

Для відповіді на це питання необхідно звернутися до основних джерел, що зумовлюють виразність соціального феномену на індивідуальному рівні, а саме до домінувального дискурсу, якого дотримується індивід у цьому питанні, а також до особливостей навколишньої соціальної дійсності. Так, ми вважаємо, що соціологічні шкали конструюють тоді, коли визначальним джерелом є саме домінувальний дискурс. Відповідним прикладом є вимірювання геополітичних орієнтацій, які переважно мають сильне ідеологічне забарвлення. Саме воно робить вплив домінувального дискурсу на геополітичні орієнтації вирішальним, оскільки ідеологія чітко впорядковує відповіді на ключові політичні питання. Звичайно, окремі індивіди можуть керуватися в цьому разі не ідеологічними переконаннями, а прагматичними міркуваннями. Утім, в умовах національної держави, що є домінантною формою організації життя великих сукупностей людей на даному етапі історії, саме ідеологічний вплив є визначальним фактором геополітичних орієнтацій.

Відповідно, коли досліджуване соціальне явище визначають кілька дискурсів, дослідник конструює індекс. Прикладом є індекс авторитаризму, заснований на соціетальних інтересах (стабільність, розвиток, порядок, сила, соціальний оптимізм, захищеність), оскільки кожного з них респондент може розглядати незалежно від інших. Так само індекс конструюють за визначального значення навколишньої соціальної дійсності. Прикладом такого вимірювального інструменту є Інтегральний індекс соціального самопочуття — досвід індивіда, що базується на його сприйнятті ефективності основних соціальних інститутів, є визначальним для результатів вимірювання.

Звичайно, вказані принципи вибору між шкалою та індексом, коли йдеться про соціальні феномени, є певним спрощенням, оскільки не існує ідеальних ситуацій, у яких результати вимірювання повністю визначає домінувальний дискурс або навколишня соціальна дійсність. Тому їх (принципи) слід розглядати як аналітичні орієнтири:

- 1) чим більшою мірою, на думку дослідника, результати вимірювання зумовлені певним дискурсом, тим вагомішою є необхідність конструювання шкали;
- 2) чим більшою мірою, на думку дослідника, результати вимірювання зумовлені різними дискурсами і/або навколишньою соціальною дійсністю респондента, тим вагомішою є необхідність конструювання індексу.

Останнім об'єктом соціологічного вимірювання є стани психічної системи індивіда. Вважаємо, що в цьому разі, залежно від ситуації, також можна використовувати шкали або індекси. Якщо індикатори є причинними факторами певного внутрішнього стану, то вибір слід робити на користь індексу. Прикладом відповідної методики є індекс психологічного дистресу SCL-9-NR. Якщо ж вимірювання базується на тих наслідках, які чинить певний психічний стан, то йдеться про шкалу. Тут треба зазначити, що останнє твердження є гіпотетичним, оскільки ми не маємо досвіду конструювання інструментів вимірю-

вання психічних станів у формі шкал. У цьому напрямку необхідні додаткові наукові пошуки. Як приклад (знову ж таки гіпотетичного характеру) можна назвати вимірювальний інструмент на основі різних соматичних наслідків психологічного дистресу (слабкість, порушення сну, зниження продуктивності тощо).

Залишилося з'ясувати, в яких випадках дослідник може побудувати класифікатор. Ключовим аспектом, що вможливорює цю дію, є інформація про різні якісні стани особистості/психічної системи/установок індивіда, яку повідомляють окремі індикатори методики. У цьому разі дослідник має можливість сформулювати правила класифікації респондентів, відштовхуючись від сукупності окремих відповідей на пункти шкали або тесту, і в подальшому перевірити їх адекватність. З огляду на це головним принципом побудови класифікаторів соціологічних тестів є використання категоріальних шкал відповідей, тобто шкал відповідей із конкретним смисловим наповненням.

Прикладами класифікаторів є правила обробки відповідей для соціологічних тестів ПСС, СТ ТПК-II, «RF-geopol», «EU-geopol» і SCL-9-NR. За вимірювання соціального самопочуття фіксовано такі якісно відмінні стани, як достатність/недостатність соціальних благ, за оцінки політичної культури — такі, як згода/незгода з певними твердженнями, за дослідження геополітичних орієнтацій — такі, як різнопланові оцінки тих чи тих аспектів геополітичної свідомості. Дещо інший підхід використано для побудови класифікатора SCL-9-NR. У цьому разі шкала відповідей являє собою набір різних ступенів частоти певних станів індивіда. З першого погляду видається, що тут слід говорити не про якісно, а про кількісно відмінні стани. Утім, як продемонструвала побудова класифікації з застосуванням дерев рішень, навіть у таких випадках є можливість побудови класифікаторів.

§ 19.3. Емпірична валідація: завдання, ситуації та методи

Метою емпіричної валідації є оцінка правильності переходу від теоретичної бази дослідження до конкретного інструменту вимірювання. Відповідними завданнями є такі: 1) аналіз розподілу підсумкового значення методики; 2) аргументування включеності інструменту до номологічної мережі; 3) демонстрування чутливості інструменту за використання в теоретично визначених групах; 4) оцінка прогностичного потенціалу соціологічного тесту; 5) перевірка факторної валідності методики (лише для вимірювальних шкал); 6) порівняння вимірювальних властивостей соціологічного тесту з відповідними властивостями методики, яка розв'язує аналогічні завдання.

Перше завдання передбачає оцінку розподілу підсумкових значень соціологічного тесту з урахуванням як специфіки вимірювальної властивості, так і ширшого культурного контексту, в який включений респондент, а також особливостей ситуації опитування. Неможливість пояснити розподіл з урахуванням зазначених факторів є важливою пересторогою для застосування методики в подальших дослідженнях. Окрім того, форма зміщеного розподілу може надати інформацію про відповідні фактори викривлення.

За перевірки розподілу, що характеризує особистісну якість, базовим очікуванням (якщо іншого не стверджує використовувана теорія) є нормальність розподілу. Таке саме припущення слід використовувати й у разі психічних станів, якщо йдеться про усталену соціальну ситуацію. Якщо ж на індивідів чинять тиск нетипові для них умови життєдіяльності, розподіли психічних станів можуть істотно відхилитися від нормального.

Коли у фокусі вимірювання перебуває соціальний феномен, то характер розподілу повністю визначають особливостями досліджуваного суспільства або певної його частини. Тут цілком очікуваними є розподіли, зміщені до більших або менших значень шкали. З огляду на це важливо проводити роздільний аналіз різних соціальних груп, коли досліджуваній сукупності є характерною суттєва соціальна гетерогенність.

Для розв'язання цього завдання можна використовувати широкий спектр методів — описову статистику, статистичні тести відповідності окремим розподілам, різні типи діаграм.

Друге завдання полягає в перевірці конструктивної валідності. Воно виокремлюється з кола інших завдань тим, що є мінімально необхідною умовою для перевірки емпіричної валідності. Треба зазначити, що класифікатори теж необхідно перевіряти на конструктивну валідність.

Базовий метод перевірки цього виду емпіричної валідності — кореляційний. Утім, залежно від типу використовуваних даних, можна застосовувати й інші методи математичної статистики. У межах цієї монографії ми не заглиблюватимемося в те, які статистичні методи і коли необхідно застосовувати у процесі емпіричної валідації. Тут ми орієнтуємося на загальновідомі правила застосування тих або тих методів залежно від вхідних даних, а також на актуальну практику статистичного аналізу. За бажання читач знайде достатньо високоякісної літератури з відповідних питань¹⁵².

Розв'язання третього завдання пов'язане з забезпеченням критеріальної валідності. Особливості її перевірки варіюються залежно від типу об'єкта вимірювання. Якщо об'єктом вимірювання є особистісна риса (латентна змінна), то як критерій слід застосовувати або іншу методику, що вимірює ту саму характеристику, або прояви індивідуальної поведінки, що підлягають фіксації й обумовлені відповідною рисою особистості. Прикладом критеріальної валідації в цьому разі може слугувати зіставлення результатів тесту на особистісну ворожість і кількості конфліктів, які мала особа впродовж останнього місяця.

В інших випадках, тобто за вимірювання психічних станів або соціальних феноменів, доцільніше, на нашу думку, використовувати опитування у критеріальних групах. Останні визначають на базі теоретичних очікувань щодо їх низької, звичайної та високої загальної виразності вимірювальної властивості. Також як критерій можна використовувати результати аналогічних вимірювальних інструментів. Приміром, для психічних станів таким є критеріальна валідація індексу психологічного дистресу SCL-9-NR шляхом опитування

¹⁵² Див. напр.: (Хили, 2005; Field et al., 2012; Бослаф, 2015).

пацієнтів психоневрологічної лікарні, а для соціальних феноменів — критеріальна валідація соціологічних тестів «RF-geopol» і «EU-geopol» завдяки дослідженню думок респондентів, які проживають у регіонах із відмінними геополітичними переконаннями.

Основним методом перевірки тут є порівняння середніх значень, хоча можливе і застосування інших методів статистичного аналізу (наприклад, кореляційного аналізу). Знову ж таки, складність статистичних процедур обумовлюють особливості дослідницького дизайну.

За наявності результатів загальнонаціонального опитування з застосуванням анкети, що охоплює багато релевантних із точки зору критеріальної валідації характеристик, можливим є використання еталонного значення (наприклад загальне значення тесту у вибірці), з яким порівнюють результати використання соціологічного тесту в різних групах респондентів. При цьому ознаки, за якими визначають відповідні групи, мають стосуватися самооцінки об'єктивних характеристик респондентів — поведінки або статусу. Звичайно, використання еталонного значення є менш надійним, аніж звернення до іншої валідної методики, дослідження безпосередньої поведінки або опитування у критеріальних групах, але за відсутності інших можливостей його використання є цілком виправданим.

На протилежність конструктивній валідації, критеріальна є не обов'язковою, а лише бажаною. Це пов'язане з більшою організаційною складністю її перевірки, що обумовлено необхідністю збирання додаткових соціологічних даних, доступ до яких може бути об'єктивно обмеженим. Утім слід мати на увазі, що успішна критеріальна валідація є важливим показником якості соціологічного тесту.

Ще більшу складність становить розв'язання четвертого завдання, спрямованого на демонстрацію прогностичної валідності. Воно передбачає побудову прогнозів розвитку досліджуваного явища з використанням результатів вимірювання. Передумовою прогностичної валідації є консистентність результатів використання методики з широкими теоретичними положеннями про функціонування і розвиток соціальної системи. Необхідним також є постійне застосування соціологічного тесту в широкомасштабних дослідженнях, що дозволить перевірити прогностичні можливості методики. З огляду на це можна сказати, що опис прогностичного використання методики є максимальним рівнем її емпіричної валідності. Враховуючи, що наш досвід прогностичної валідації є обмеженим (він наявний лише для шкали громадської активності), докладніші рекомендації потребують подальших емпіричних досліджень і відповідної концептуалізації.

У методичному значенні, як ми бачимо на даному етапі, прогностична валідація передбачає використання широкого спектра засобів роботи з даними.

Нарешті, коли дослідник зупиняється на розробці вимірювальної шкали, він має перевірити факторну валідність методики, якої можна досягти за допомоги здійснення конфірматорного факторного аналізу. У цій роботі містяться приклади відповідної перевірки для шкал громадської активності, антидемократичних установок, авторитаризму, геополітичних орієнтацій.

Ураховуючи те, що дуже часто новий соціологічний інструмент вимірювання є альтернативою вже існуючій методикі (або навіть кількох), бажано здійснити його порівняння з конкурентами. Для цього можна використати як методи, вже застосовані у процесі розробки конкурентних методик, так і ті, що були використані за валідазації власного тесту. Критерії такого відбору визначаються методичними стандартами, встановленими нами для тих чи тих різновидів соціологічних тестів.

§ 19.4. Визначення критеріальних рівнів: існуючі підходи

Незважаючи на те, що соціологічний тест, який пройшов успішну теоретичну й емпіричну валідацію, є цінним вимірювальним інструментом, його повна методологічна завершеність передбачає визначення критеріальних рівнів тесту. Це дозволяє правильно переходити від значень окремих пунктів методики до якісно відмінних станів респондентів за тією властивістю, яка лежить у фокусі соціологічного вимірювання.

З огляду на наші дослідження можемо виокремити два загальні підходи до визначення критеріальних рівнів. Перший використовує дедуктивну логіку, другий — індуктивну. Дедуктивна логіка передбачає використання внутрішнього змістового наповнення варіантів відповідей на пункти тесту. Якщо ми вважаємо, що варіанти відповідей більшою мірою позначають різні якісні стани, то можемо запропонувати їх інтерпретацію з використанням вичерпного набору логічних концептуально несуперечливих правил категоризації респондентів. Своєю чергою, індуктивна логіка полягає у зверненні до результатів емпіричних досліджень із метою з'ясування критеріальних рівнів.

Очевидно, що за використання класифікаторів бажано застосовувати дедуктивну логіку, а в разі індексів і шкал — індуктивну. Однак ситуація ускладнюється тим, що в багатьох випадках шкали та індекси можуть легко набувати форми класифікатора. У цьому разі дослідник має з'ясувати, який спосіб визначення критеріальних рівнів є ефективнішим. Основою для цього є змінні, використані для побудови номологічної мережі. Відповідні зв'язки можна перевірити за допомоги таблиць спряженості або дисперсійного аналізу.

Окремо треба зазначити, що індуктивна логіка передбачає два можливі шляхи визначення критеріальних рівнів — груповий і аналітичний. Групове визначення критеріальних рівнів передбачає проведення додаткових досліджень серед соціальних груп, для яких характерний низький і/або високий рівень вимірювальної властивості. При цьому бажано дослідити якомога більшу кількість таких груп. Наприклад, за визначення критеріальних рівнів індексу психологічного дистресу SCL-9-NR (Дембіцький, 2017d) ми звернулися до таких груп, як пацієнти психоневрологічних лікарень, особи, які оцінюють своє здоров'я як «погане» або «дуже погане», а також особи, що проживали в зоні бойових дій. В подальшому на основі вивчення центральних тенденцій відповідних розподілів значень методики (середніх або медіан) дослідник робить висновок про критеріальні рівні соціологічного тесту.

Аналітичне визначення критеріальних рівнів використовують, коли немає можливості визначити критеріальні групи (наприклад, до них немає доступу). У цьому разі за допомоги кореляційного і регресійного аналізу дослідник може побудувати ідеально-типові профілі респондентів, для яких характерні низькі та високі значення досліджуваної ознаки. Очевидно, що для аналітичного визначення критеріальних рівнів інструментарій дослідження має включати широкий спектр питань, з яких у подальшому буде здійснено відбір окремих індикаторів, що дозволяють побудувати ідеально-типові профілі респондентів. Аналітичне визначення критеріальних рівнів було застосовано нами у процесі роботи з Інтегральним індексом соціального самопочуття.

§ 19.5. Підсумкові положення

Зміст пропонованої роботи вказує на те, що розробка соціологічних тестів, якщо і піддається певній технологізації, то лише в межах розв'язання деяких завдань. Це, першою чергою, обумовлено тематичною гетерогенністю тих об'єктів, які можуть бути об'єктом соціологічного вимірювання. З огляду на це рекомендації до розробки соціологічних тестів можуть лише задавати загальні напрямки дослідницької активності та можливі техніки їх реалізації.

На нашу думку, найкращими методологічними рекомендаціями щодо розробки соціологічних шкал, індексів і класифікаторів є методичні кейси, які містять інформацію про розробку конкретних соціологічних тестів. Ознайомлення з ними, а також критична оцінка використаних підходів і технік дозволять соціологу сформулювати необхідну методологічну компетентність, яка слугуватиме йому як ефективний орієнтир.

Висновки до четвертого розділу

Як ми зазначали у висновках до першого розділу, соціологічні тести є частиною традиційної емпіричної соціології. У цьому розділі продемонстровано, що їх розробка спирається на кількісну парадигму, а також парадигму досліджень змішаного типу. Змішаний підхід домінує у процесі теоретичної валідації, кількісний — емпіричної.

Однак визначення прийняттого різновиду соціологічного тесту виходить за межі традиційної емпіричної соціології. Воно потребує широких теоретико-методологічних міркувань із приводу природи досліджуваного об'єкта. Це вказує на умовність розмежування емпіричної та теоретичної соціології, а також на міждисциплінарну природу соціологічного вимірювання.

В основу розробки критеріальних рівнів може бути покладений як змішаний або кількісний підхід, так і валідні концептуальні положення. Фактично це найбільш комплексний аспект наукових пошуків у процесі конструювання соціологічних тестів.

На завершення ще раз наголосимо на важливості рефлексивності дослідника, який розпочинає розробку соціологічного тесту. Лише науковий скептицизм щодо домінуючого методологічного дискурсу, а також відмова від установки на технологічність дозволять утримувати процес розробки соціологічного тесту у продуктивному річищі.

ВИСНОВКИ

У висновках зосередимо увагу як на головних знахідках нашого дослідження, так і на проблемних моментах, які були виявлені, але не отримали відповіді. Усвідомлення й артикуляція останніх є важливими, оскільки задають напрямки подальших пошуків.

Перший висновок стосується міждисциплінарного поля, що визначає підходи до вимірювання в соціальних науках. Так, розпочинаючи розробку вимірювальної методики, дослідник має усвідомлювати значну гетерогенність відповідних дисциплінарних традицій. Їх доцільність визначають перш за все природою вимірювання, а також тим форматом вимірювальної методики, який є оптимальним з точки зору дизайну конкретного дослідження. Враховуючи той факт, що одним із загальноновживаних у соціології дизайнів є кроссекційні дослідження, розробка соціологічних тестів або адаптація вимірювальних методик, запозичених з інших дисциплін, має відповідати не тільки дисциплінарним стандартам соціології, а і принципам масових опитувань. Це робить використання стандартів психологічного і педагогічного вимірювання в їх чистому вигляді мало прийнятним для соціології.

З огляду на це логічним є впровадження спеціальної методологічної категорії для позначення комплексних вимірювальних інструментів, які відповідають соціологічним вимогам. Як таку ми (вслід за Н. Паніною і Є. Головахою) пропонуємо використовувати соціологічний тест. До відповідних вимог належать універсальність, інтегральність, якісність, стандартизованість, компактність, «чутливість» і об'єктивність. Окрім цього, соціологічні тести можна розглядати у трьох вимірах: методичних вимог, природи властивості, що підлягає вимірюванню, і діагностичного потенціалу.

Наявність серед об'єктів соціологічного вимірювання соціальних феноменів приводить нас до другого висновку, що стосується припущень, які є адекватними характеру не тільки соціологічного вимірювання, а й соціологічних досліджень у широкому значенні. Такими ми вважаємо припущення прагматичного підходу, які акцентують увагу на: 1) проблематичності пошуку істинного значення, як це відбувається у класичній теорії вимірювання і підходах, генетично з нею пов'язаних; 2) релевантних причинах отримуваних у процесі соціологічного вимірювання результатів, до яких належать домінуючий дискурс індивіда, а також навколишня соціальна дійсність; 3) джерелах випадкової величини, якими за вимірювання соціальних феноменів є психологічні особливості індивідів; 4) необхідності теоретичної та емпіричної валідазації соціологічних тестів.

Останнє припущення веде до необхідності конкретизації процесу розробки соціологічних тестів, який включає теоретичну валідазацію, вибір різновиду соціологічного тесту, його емпіричну валідазацію і розробку норм або критеріальних рівнів. До джерел досягнення теоретичної валідності належать окре-

мі концепції в межах соціальних наук (ми надаємо пріоритет тим із них, що побудовані в результаті проведення кейс-стаді), попередні спеціальні дослідження, професійний, а іноді й універсальний досвід. Використання цих джерел сприяє досягненню очевидного змістовного і композиційного компонента теоретичної валідності.

Різновидами соціологічного тесту є шкали, індекси і класифікатори. Шкали використовують, коли в основі відповідей на індикатори методики лежить одна спільна причина. Такою може бути особистісна риса респондентів, певний стан їхньої психічної системи (якщо індикатори розкриваються як наслідки такого стану) або той чи той домінуючий дискурс (якщо йдеться про вимірювання соціального феномену). Розробку індексу здійснюють, коли досліджуване явище зумовлене кількома незалежними причинами. При цьому об'єктом вимірювання може бути стан психічної системи індивіда або соціальний феномен. Що стосується класифікаторів, то їх побудова можлива для будь-якого з указаних об'єктів, але лише тоді, коли дослідник має інформацію про різні якісні стани особистості / психічної системи / установок індивіда, а також знає, як перейти до них на підставі значень шкали або індексу.

Використання шкали, індексу або класифікатора в емпіричному дослідженні відкриває можливість емпіричної валідації. Остання передбачає перевірку конструктивної, критеріальної, прогностичної та порівняльної валідності. В окремих випадках може знадобитися перевірка інших видів емпіричної валідності, наприклад, факторної. Указані види валідності є загальновідомими і спираються на кількісну парадигму в соціальних дослідженнях.

Розробка критеріальних рівнів може здійснюватися з застосуванням одного з двох підходів: 1) дедуктивного (використання внутрішнього змістового наповнення варіантів відповідей на пункти тесту); 2) індуктивного (полягає у зверненні до результатів емпіричних досліджень). При цьому індуктивний підхід передбачає два можливі шляхи визначення критеріальних рівнів — груповий (використовують результати спеціально проведених досліджень серед релевантних груп) і аналітичний (вибудовують ідеально-типові профілі респондентів на підставі наявних даних).

У результаті використання методологічних положень, зазначених у поданих висновках, а також низки емпіричних досліджень ми дійшли висновків щодо релевантності окремих процедур і технік розробки соціологічних тестів. Якщо говорити про дослідження змішаного типу, то такими є: побудова таблиці характеристик, формування наративного профілю і проведення дебрифінгу. У разі кількісних досліджень розробка соціологічних тестів передбачає (само)дебрифінг методологічної ретельності, який спрямований на досягнення цілісного бачення процесу вимірювання, а також постійну увагу в цьому контексті до застосовування дослідницьких технік/методів/стратегій. Останнє полягає у використанні конкуруючих пояснень, розбудові номологічної мережі, збільшенні інформаційного потенціалу методичних досліджень і орієнтації не лише на перевірку статистичних гіпотез, а й на пошуковий аналіз даних.

Тепер звернемо увагу на проблемні моменти, виявлені під час нашого дослідження. У найзагальнішому вигляді їх можна поділити на методологічні, концептуальні та методичні. Почнемо з методологічних.

Перш за все ми приділили недостатньо уваги різновидам емпіричної соціології. Так, оцінка потенціалу неокласичної емпіричної соціології та емпіричної соціології дискурсів з точки зору традиційної емпіричної соціології може суттєво збагатити арсенал методологічних положень, що регулюють розробку соціологічних тестів. На нашу думку, такі пошуки дозволять деталізувати положення, які стосуються визначення характеру властивостей, що підлягають вимірюванню. Останнє безпосередньо стосується правильного вибору між різновидами соціологічних тестів, особливо між шкалами та індексами.

Ще один висновок, пов'язаний із потенціалом різних видів емпіричної соціології, а також міждисциплінарним аналізом, стосується аналізу різних типів об'єктів вимірювання. Як уже було сказано, до таких ми відносимо особистісні риси, стани психічної системи індивідів, а також соціальні феномени. Така диференціація виглядає для нас цілком слушною, але вона є лише вихідним пунктом для розвитку відповідної типології. Особливо актуальним це є для деталізації типів соціальних феноменів, які потрапляють у фокус соціологічного тестування.

Наступний аспект, який потребує додаткової уваги, поданий джерелами теоретичної валідності. Їх перелік і аргументація щодо використання, а також обмежена кількість прикладів, які ми подали в цій монографії, мають бути доповнені чіткішими вказівками щодо релевантності тих або тих джерел у контексті розробки соціологічних тестів різної предметної спрямованості. Одним із перспективних напрямків пошуку при цьому ми вважаємо різноманітні підходи до розробки комплексних вимірювальних інструментів у дослідженнях змішаного типу. Такі підходи потенційно є фреймом, що дозволить об'єднати якщо не всі, то принаймні окремі джерела теоретичної валідності в контексті розробки соціологічних тестів. Загалом потенціал вимірювальних підходів у дослідженнях змішаного типу не обмежується розглядом джерел теоретичної валідності й має бути розповсюджений на інші елементи розробки соціологічних тестів.

Те саме стосується (само)дебрифінгу методологічної ретельності. Усі елементи кількісного підходу, що були оцінені нами як придатні на базі отриманого емпіричного досвіду, необхідно деталізувати шляхом їх цілеспрямованого використання у процесі розробки соціологічних тестів. Це є принциповою вимогою для отримання ними статусу технік, методів або стратегій дослідження.

Маємо також зазначити, що в рамках цієї монографії ми присвятили замало уваги праці відомого вітчизняного соціолога В. Паніотто, погляди якого щодо надійності вимірювання, а також емпіричної валідації (Паніотто, 1986) не втрачають актуальності. Відповідні положення є цінним джерелом збагачення принципів і процедур розробки соціологічних тестів.

Нарешті, останній методологічний висновок стосується деталізації рекомендацій до розробки соціологічних тестів. У тексті ми зазначили складність формулювання відповідних положень. Можливим розв'язанням цієї проблеми є використання диференційованого підходу до формулювання таких рекомендацій, який полягає в розробці кількох їх версій, що орієнтовані на той або той клас соціологічних тестів. Наприклад, як такі класифікувальні ознаки можуть виступити різні об'єкти, що потрапляють у фокус соціологічного вимірювання. Ураховуючи наявний досвід, оптимальним варіантом для форму-

лювання сфокусованих рекомендацій на даному етапі ми бачимо соціологічні тести, спрямовані на вимірювання станів психічної системи індивідів.

Концептуальні проблеми, виявлені під час нашого дослідження, стосуються вивчення індивідуального благополуччя і політичної культури. Щодо вимірювання індивідуального благополуччя, то перспективним напрямком убачаємо підготовку теоретичної бази, що описує статусні характеристики індивіда, з подальшим переходом до конструювання соціологічного тесту. Також у цьому контексті є актуальним дослідження взаємозв'язків між трьома ключовими складовими індивідуального благополуччя — психологічним дистресом, соціальним благополуччям і соціальним статусом. Окрім цього, додатковим висновком, що стосується сфери досліджень індивідуального благополуччя, є необхідність подальшого вивчення особливостей загального прояву в житті індивіда позитивних і негативних афектів, а також того, як він їх сприймає.

Щодо вимірювання політичної культури, нерозв'язаним є завдання фіксації геополітичного прагматизму/радикалізму. Головну проблему тут ми бачимо в недостатній деталізації відповідних теоретичних положень. Більше за те, маємо припущення про можливість використання зазначеного концепту не лише стосовно геополітичних орієнтацій, а і стосовно інших аспектів політичної культури.

Висновки методичного характеру зроблено щодо використання таких соціологічних тестів, як індекс психологічного дистресу SCL-9-NR, L-шкала з Мінесотського багатоаспектного особистісного опитувальника (MMPI) та Інтегральний індекс соціального самопочуття.

Незважаючи на докладне дослідження вимірювальних властивостей SCL-9-NR, ми відчуваємо необхідність подальших методичних розвідок. Вони стосуються того, як метод збору інформації впливає на результати використання тесту. Так, ми не маємо відомостей щодо того, який вплив на них справить використання особистого інтерв'ю (face-to-face). Важливість цієї інформації зумовлена межами, які визначають рівні психологічного дистресу. Цілком можливо, що за використання особистого інтерв'ю вони можуть підлягати коригуванню.

За використання L-шкали напрямків для методичних розвідок значно більше. Перш за все необхідні спеціальні методичні дослідження, спрямовані на адаптацію цієї методики до особливостей масових опитувань. Такі дослідження дозволять відібрати найефективніші індикатори, а також з'ясувати особливості адекватної інтерпретації результатів використання шкали. Важливим напрямком цих досліджень є виявлення різних механізмів психологічного захисту, які застосовують респонденти під час відповіді на соціологічну анкету.

Додаткових методичних досліджень потребує також Інтегральний індекс соціального самопочуття. Зокрема, ми вважаємо за необхідне сфокусуватися на з'ясуванні особливостей правильної інтерпретації його результатів — як, використовуючи результати тесту, слід визначати категорії респондентів у контексті їхнього соціального самопочуття. Тобто існує проблема переходу від кількісних результатів до якісних.

На завершення висловлюємо надію, що такий напрямок методологічних досліджень вітчизняної соціології, як соціологічне тестування, залишатиметься предметом наукової уваги.

ДОДАТКИ

Додаток 1

Експрес-тест психологічного дистресу SCL-9-NR

У зв'язку з певними життєвими ситуаціями багато людей стикаються із різними переживаннями, відчуттями і думками. Далі наведено перелік найчастіших станів такого роду. Оцініть, як часто кожен із них був характерним для Вас останнім часом: ніколи, зрідка, періодично, майже постійно. За кожним пунктом оберіть лише один варіант відповіді.

- 1 — Ніколи
2 — Зрідка
3 — Періодично
4 — Майже постійно

1	Легке виникнення досади або роздратування	1	2	3	4
2	Те, що Вам важко зосередитися	1	2	3	4
3	Пригнічений настрій, «хандра»	1	2	3	4
4	Те, що Ваші почуття легко зачепити	1	2	3	4
5	Почуття, що більшості людей не можна довіряти	1	2	3	4
6	Відчуття напруженості або збудженості	1	2	3	4
7	Відчуття слабкості в різних частинах тіла	1	2	3	4
8	Знервованість, коли Ви залишилися наодинці	1	2	3	4
9	Думки про те, що з Вашим тілом якийсь негаразд	1	2	3	4

Синтаксис розрахунку зваженого індексу для L-шкали на підставі статі та вікової групи респондента

```
for (a in LN) {  
  for (i in 1:3) {  
    test<- prop.table(table(lieplus[[a]][lieplus$voзраст == i],  
                           lieplus$V266[lieplus$voзраст == i]),2)[1,]  
    ind<- lieplus$V266 == "Чоловіча" & lieplus$voзраст == i &  
          lieplus[[a]] == "Не згоден"  
    ind[is.na(ind)] <- F  
    lieplus$newlie[ind] <- lieplus$newlie[ind] + test[[1]]  
    ind<- lieplus$V266 == "Жіноча" & lieplus$voзраст == i &  
          lieplus[[a]] == "Не згоден"  
    ind[is.na(ind)] <- F  
    lieplus$newlie[ind] <- lieplus$newlie[ind] + test[[2]]  
  }  
}
```

, де LN — вектор з іменами індикаторів шкали брехні в масиві,
V266 — змінна з інформацією про стать респондента,
voзраст — змінна з інформацією про вікову групу респондента,
newlie — зважені значення індексу для шкали брехні.

Індикатори підшкали ворожості (SCL-90-R)

1. Легке виникнення відчуття прикрості або роздратування.
2. Вибухи гніву, які їх Ви не могли стримати.
3. Імпульси завдавати тілесні ушкодження або шкоду кому-небудь.
4. Імпульси псувати або трощити щось.
5. Те, що Ви часто вступаєте в суперечки.
6. Те, що Ви кричите або шпурляєте речі.

Додаток 4

Твердження для вимірювання ставлення до демократії

1. Демократія — найкращий політичний устрій для будь-якої сучасної держави.
2. Без демократії неможливий сталий економічний розвиток країни.
3. Я впевнений, що лише в демократичному суспільстві насправді поважають права людини.
4. Демократія забезпечує людині найбільші можливості для індивідуального політичного вибору порівняно з іншими режимами.
5. Тільки за демократії законодавча влада відображає інтереси всіх основних соціальних груп.
6. Демократія надає людям можливість об'єднуватися в партії та спілки для реального захисту своїх прав та інтересів.

Твердження для вимірювання громадянської активності

1. Мені все одно, яка буде влада, тільки б не стало гірше.
2. Я не покладаюся на вибори, бо не вірю, що від їхніх результатів зміниться моє життя.
3. Я свій вибір давно вже зробив, тому не хочу брати участі в теперішньому політичному житті.
4. Який сенс боротися за свої права, якщо влада своїми діями відверто їх ігнорує.
5. Шляхом голосування ми обираємо владу, ну а далі від нас уже нічого не залежить.
6. Будь-яка спроба щось змінити в політичному житті країни потребує від людей великих жертв, котрі, як правило, виявляються марними.

Додаток 5

Повний варіант кластеризації з застосуванням концептуальної типології, %

Ставлення до активності	Ставлення до демократії				
	ДЕМ	АМБ	НЕВИЗН	АНТИДЕМ	Загалом
АКТ	14,7	0,6	2,1	2,4	19,8
АМБ	5,4	0,2	0,7	0,7	7,0
НЕВИЗН	6,2	0,3	5,2	0,9	12,6
ПАСИВ	43,7	1,2	11,5	4,5	60,9
Загалом	70,0	2,3	19,5	8,5	100,3

Скорочення в таблиці: ДЕМ — демократичні установки, АМБ — амбівалентні установки, НЕВИЗН — невизначені установки, АНТИДЕМ — антидемократичні установки, АКТ — установки на громадську активність, ПАСИВ — установки на громадську пасивність.

Додаток 6

Центроїди груп, виокремлених шляхом латентного класового аналізу, 2006

Група	Центроїд групи	
	Шкала демократії	Шкала активності
Перша група	18,2	17,3
Друга група	19,4	11,7
Третя група	13,8	16,1
Четверта група	27,3	15,4
П'ята група	22,0	16,8

Центроїди груп, виокремлених шляхом латентного класового аналізу, 2016

Група	Центроїд групи	
	Шкала демократії	Шкала активності
Активні демократи	27,0	20,8
Пасивні демократи	28,0	9,6
Пасивно невизначені	24,0	12,9
Пасивні антидемократи	17,2	13,8
Активні антидемократи	20,2	18,5

Додаток 7

Соціологічний тест «Типи політичної культури — II»

Люди висловлюють різні судження про демократію. Відзначте, будь ласка, якою мірою Ви особисто згодні чи не згодні з кожним із таких висловлень, якщо йдеться про розвинені країни. (Дайте одну найбільш слушну відповідь по кожному рядку)

- 1 — Цілком згоден
 2 — Скоріше згоден
 3 — Важко сказати, згоден чи ні
 4 — Скоріше згоден
 5 — Повністю згоден

1	Демократія надає надто багато свободи багатим людям	1	2	3	4	5
2	Демократичні вибори — це фарс, що не захищає інтереси простих людей	1	2	3	4	5
3	Демократія — це тільки слова, якими прикриваються ті, хто має доступ до влади для забезпечення власних інтересів	1	2	3	4	5
4	Демократичні вибори, як правило, приводять до влади найбільш корисливих людей	1	2	3	4	5
5	Демократія надає людині оманливу свободу вибору, якою вона не може скористатися	1	2	3	4	5
6	Демократія несе простій людині непевність у завтрашньому дні	1	2	3	4	5

Тепер знову оцініть ці твердження під кутом зору української дійсності.

7	Демократія надає надто багато свободи багатим людям	1	2	3	4	5
8	Демократичні вибори — це фарс, що не захищає інтереси простих людей	1	2	3	4	5
9	Демократія — це тільки слова, якими прикриваються ті, хто має доступ до влади для забезпечення власних інтересів	1	2	3	4	5
10	Демократичні вибори, як правило, приводять до влади найбільш корисливих людей	1	2	3	4	5
11	Демократія надає людині оманливу свободу вибору, якою вона не може скористатися	1	2	3	4	5
12	Демократія несе простій людині непевність у завтрашньому дні	1	2	3	4	5

А що Ви думаєте про постать сильного лідера в політичному житті України?

13	Єдина політична воля забезпечує стабільність у суспільстві краще, ніж набір різних точок зору	1	2	3	4	5
14	Для нормального розвитку країні потрібна "сильна рука", а не розмови про демократію	1	2	3	4	5
15	Тільки політика "сильної руки" може зберегти порядок у суспільстві	1	2	3	4	5
16	Сильна держава можлива лише тоді, коли на чолі країни стоїть один лідер	1	2	3	4	5
17	Наявність сильного політичного лідера дає людям упевненість у завтрашньому дні	1	2	3	4	5
18	Тільки по-справжньому сильний лідер може захистити те, що дорого всім	1	2	3	4	5

Зрештою, чи згодні Ви з такими твердженнями?

19	Мені все одно, яка буде влада, тільки б не стало гірше	1	2	3	4	5
20	Я не покладаюся на вибори, оскільки не вірю, що від їхніх результатів зміниться моє життя	1	2	3	4	5
21	Я свій вибір давно вже зробив, тому не хочу брати участі в теперішньому політичному житті	1	2	3	4	5
22	Який сенс боротися за свої права, якщо влада своїми діями відверто їх ігнорує	1	2	3	4	5
23	Шляхом голосування ми обираємо владу, ну а далі від нас уже нічого не залежить	1	2	3	4	5
24	Будь-яка спроба щось змінити в політичному житті країни вимагає від людей великих жертв, котрі, як правило, виявляються марними	1	2	3	4	5

Додаток 8

Соціологічний тест «RF-geopol»

- 1) Яким, на Ваш погляд, був вплив Російської імперії на українську культуру — негативним чи позитивним?
 1. Переважно негативним.
 2. Частково негативним, частково позитивним.
 3. Переважно позитивним.
 4. Важко сказати.
- 2) Як Ви оцінюєте культурну спадщину радянського періоду в історії України?
 1. Переважно негативно.
 2. Частково негативно, частково позитивно.
 3. Переважно позитивно.
 4. Важко сказати.
- 3) Як Ви вважаєте, чи можна довіряти офіційним засобам масової інформації Російської Федерації?
 1. Не можна довіряти більшості ЗМІ.
 2. Можна довіряти більшості ЗМІ.
 3. Важко сказати.
- 4) Наскільки близькими особисто Вам є загальнослов'янські цінності (слов'янська єдність, релігійність, шанування роду і т.д.)?
 1. Зовсім не близькі.
 2. Певною мірою близькі.
 3. Дуже близькі.
 4. Важко сказати.
- 5) Будь ласка, оцініть пріоритет міжнародного співробітництва України з Російською Федерацією.
 1. Низький пріоритет — співробітництво необхідно скорочувати.
 2. Ситуативний пріоритет — інтенсивність співробітництва залежить від розв'язуваних завдань.
 3. Високий пріоритет — співробітництво має стратегічний характер.
 4. Важко сказати.
- 6) Як Ви ставитеся до ідеї військового співробітництва з Російською Федерацією?
 1. Такого співробітництва не можна допускати.
 2. Таке співробітництво є необхідним.
 3. Важко сказати.

- 7) Оцініть політику Російської Федерації стосовно України на даному етапі.
1. Повністю корислива — інтереси України не враховують.
 2. Частково корислива — інтереси України враховують частково.
 3. Партнерська — враховують більшість інтересів України.
 4. Важко сказати.
- 8) Як Ви ставитеся до того, щоб Російська Федерація впливала на внутрішньо-політичне життя України?
1. Негативно — це є руйнівним для України.
 2. Нейтрально — наслідки можуть бути як позитивними, так і негативними.
 3. Позитивно — це буде корисним для України.
 4. Важко сказати.

Додаток 9

Соціологічний тест «EU-георол»

- 1) Оцініть, наскільки є близькою європейська культура особисто Вам.
 1. Не близька зовсім.
 2. Певною мірою близька.
 3. Дуже близька.
 4. Важко сказати.
- 2) Як Ви вважаєте, чи можна довіряти офіційним засобам масової інформації країн Європейського Союзу?
 1. Не можна довіряти більшості ЗМІ.
 2. Можна довіряти більшості ЗМІ.
 3. Важко сказати.
- 3) Наскільки близькими особисто Вам є європейські цінності (євроцентризм, мультикультуралізм, толерантність і т. д.)?
 1. Зовсім не близькі.
 2. Певною мірою близькі.
 3. Дуже близькі.
 4. Важко сказати.
- 4) Будь ласка, оцініть пріоритет міжнародного співробітництва України з Європейським Союзом.
 1. Низький пріоритет — співробітництво необхідно скорочувати.
 2. Ситуативний пріоритет — інтенсивність співробітництва залежить від розв'язуваних завдань.
 3. Високий пріоритет — співробітництво має стратегічний характер.
 4. Важко сказати.
- 5) Як Ви ставитеся до ідеї військового співробітництва з НАТО?
 1. Такого співробітництва не можна допускати.
 2. Таке співробітництво є необхідним.
 3. Важко сказати.
- 6) Оцініть політику Європейського Союзу стосовно України на даному етапі.
 1. Повністю корислива — інтереси України не враховують.
 2. Частково корислива — інтереси України враховують частково.
 3. Партнерська — враховують більшість інтересів України.
 4. Важко сказати.

- 7) Як Ви ставитеся до того, щоб Європейський Союз впливав на внутрішньо-політичне життя України?
1. Негативно — це є руйнівним для України.
 2. Нейтрально — наслідки можуть бути як позитивними, так і негативними.
 3. Позитивно — це буде корисним для України.
 4. Важко сказати.
- 8) Як Ви оцінюєте сьогоdnішній курс України на євроінтеграцію?
1. Дуже негативно, це суперечить інтересам України.
 2. Негативно, оскільки це роблять неграмотно.
 3. Нейтрально, плюси і мінуси врівноважують одно одне.
 4. Позитивно, це відповідає актуальним інтересам України.
 5. Дуже позитивно, це важливе стратегічне завдання для України.

Додаток 10

Синтаксис розрахунку класифікаторів геополітичних орієнтацій (на прикладі Європейського Союзу і Польщі)

I. Класифікатор для Європейського Союзу

```

geoNA$euroPl <- (geo$V1 == 3) + (geo$V14 == 2) + (geo$V22 == 3) + (geo$V27 == 3) +
(geo$V35 == 2) + (geo$V38 == 3) + (geo$V46 == 3) + (geo$V54 > 3)
geoNA$euroMi <- (geo$V1 == 1) + (geo$V14 == 1) + (geo$V22 == 1) + (geo$V27 == 1) +
(geo$V35 == 1) + (geo$V38 == 1) + (geo$V46 == 1) + (geo$V54 < 3)
geoNA$euro2 <- geoNA$euroPl - geoNA$euroMi
geo$euroDN <- (geo$V1 == 4) + (geo$V14 == 3) + (geo$V22 == 4) + (geo$V27 == 4) +
(geo$V35 == 3) + (geo$V38 == 4) + (geo$V46 == 4)
geoNA$euroFin <- NA
geoNA$euroFin[geoNA$euroDN > (geoNA$euroPl + geoNA$euroMi) & (geoNA$euroPl +
geoNA$euroMi + geoNA$euroDN) > 4] <- 3
geoNA$euroFin[(geoNA$euroPl + geoNA$euroMi) < 4 & is.na(geoNA$euroFin)] <- 2
geoNA$euroFin[(geoNA$euroPl + geoNA$euroMi) >= 4 & (geoNA$euroPl -
geoNA$euroMi) %in% c(-1,0,1)] <- 4
geoNA$euroFin[(geoNA$euroPl + geoNA$euroMi) >= 4 & (geoNA$euroPl -
geoNA$euroMi) <= -2] <- 1
geoNA$euroFin[(geoNA$euroPl + geoNA$euroMi) >= 4 & (geoNA$euroPl -
geoNA$euroMi) >= 2] <- 5
geoNA$euroFin <- factor(geoNA$euroFin, labels = c("Her", "Hтp", "Hвз", "Амб", "Поз"))

```

II. Класифікатор для Польщі

```

geoNA$polPl <- (geo$V7 == 3) + (geo$V29 == 3) + (geo$V40 == 3) + (geo$V48 == 3)
geoNA$polMi <- (geo$V7 == 1) + (geo$V29 == 1) + (geo$V40 == 1) + (geo$V48 == 1)
geoNA$pol2 = geoNA$polPl - geoNA$polMi
geoNA$polDN <- (geo$V7 == 4) + (geo$V29 == 4) + (geo$V40 == 4) + (geo$V48 == 4)
geoNA$polFin <- NA
geoNA$polFin[geoNA$polDN > (geoNA$polPl + geoNA$polMi) & (geoNA$polPl +
geoNA$polMi + geoNA$polDN) > 2] <- 3
geoNA$polFin[(geoNA$polPl + geoNA$polMi) < 2 & is.na(geoNA$polFin)] <- 2
geoNA$polFin[(geoNA$polPl + geoNA$polMi) >= 2 & (geoNA$polPl -
geoNA$polMi) %in% c(0)] <- 4
geoNA$polFin[(geoNA$polPl + geoNA$polMi) >= 2 & (geoNA$polPl -
geoNA$polMi) <= -1] <- 1
geoNA$polFin[(geoNA$polPl + geoNA$polMi) >= 2 & (geoNA$polPl -
geoNA$polMi) >= 1] <- 5
geoNA$polFin <- factor(geoNA$polFin, labels = c("Her", "Hтp", "Hпp", "Амб", "Поз"))

```

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- Аванесов, В. С. (1982). *Тесты в социологическом исследовании*. Москва: Наука.
- Аванесов, В. С. (2007). Истоки и основные понятия математической теории измерений (Item Response Theory). Статья вторая. *Педагогические измерения*, (3), 3–36.
- Анастаси, А., & Урбина, С. (2007). *Психологическое тестирование*. Санкт-Петербург: Питер.
- Андресон, Б. (2001). *Воображаемые сообщества. Размышления об истоках и распространении национализма*. Москва: Канон-Пресс-Ц, Кучково поле.
- Бабосов, Е. М. (2009). *Социология: энциклопедический словарь*. Москва: Книжный дом «Либроком».
- Барт, П. (1902). *Философия истории как социология*. Санкт-Петербург: С.-Петербургская Электротпечатня.
- Белоножко, М. Л., & Барбаков, О. М. (2013). Социальное самочувствие сельского населения северного региона России. *Фундаментальные исследования*, 11(10), 2534–2538.
- Белялов, Ф. И. (2018). *Шкала депрессии CES-D*. Получено из <http://therapy.irkutsk.ru/doc/cesd.pdf>
- Бослаф, С. (2015). *Статистика для всех*. Москва: ДМК Пресс.
- Бурко, В. А., & Волкова, Ж. В. (2017). Социальное самочувствие пожилых людей в современном российском обществе (сравнительный анализ). *Вестник Пермского национального исследовательского политехнического университета. Социально-экономические науки*, (1), 43–62.
- Волков, Е. П. (2015). Социальное самочувствие как фактор повышения профессиональной мотивации офицеров в условиях реформирования УИС. *Правовое и духовно-нравственное воспитание российского офицерства* (с. 64–69). Самара: Самарский юридический институт Федеральной службы исполнения наказаний.
- Волков, Е. П. (2016). Социальное самочувствие как фактор социально-профессиональной деятельности сотрудников уголовно-исполнительной системы. *Современные тенденции развития науки и технологий*, (5), 96–102.
- Герасимова, К. Г. (2017). Конструирование социологических индексов: опыт методической рефлексии. *Вестник Российского университета дружбы народов. Серия: Социология*, 17(1), 106–115.
- Головаха, Е. И. (2004). Социологическое знание: специфика, критерии научности и перспектива развития. *Социология: теория, методы, маркетинг*, (1), 5–14.
- Головаха, Е. И., Горбачик, А. П., Любивая, Т., Панина, Н. В., Серeda, В., & Урсуленко, К. (2008). Субъективная надежность: теория и метод измерения (ИСН). *Социология: теория, методы, маркетинг*, (1), 166–188.

- Головаха, Е. И., Панина, Н. В., & Горбачик, А. П. (1998). Измерение социального самочувствия: тест ИИСС. *Социология: методология, методы, математическое моделирование*, (10), 45–71.
- Головаха, Е., & Панина, Н. (1997). *Интегральный индекс социального самочувствия (ИИСС): конструирование и применение социологического теста в массовых опросах*. Киев: Ин-т социологии НАН Украины.
- Головаха, Є. І. (2005). *Звіт про науково-дослідну роботу «Політична культура українського суспільства: тенденції та закономірності розвитку»*. Київ: Інститут соціології НАН України.
- Горбачик, А. П. (2006). *Пояснення вибору методу*. Київ: Інститут соціології НАН України.
- Гриценко, Г. Д. (2010). Социальное самочувствие населения как индикатор напряженности на Северном Кавказе. В С. Я. Суший (Ред.), *Юг России: проблемы, прогнозы, решения* (с. 186–195). Ростов-на-Дону: Южный научный центр РАН.
- Гущина, И. А., Кондратович, Д. Л., & Положенцева, О. А. (2011). Социальное самочувствие населения моногородов как показатель уровня адаптированности к социально-экономическим трансформациям. *Управление экономическими системами: электронный научный журнал*, (34). Получено из <https://cyberleninka.ru/article/n/sotsialnoe-samochuvstvie-naseleniya-monogorodov-kak-pokazatel-urovnya-adaptirovannosti-k-sotsialno-ekonomicheskim-transformatsiyam>
- Дембицкий, С. (2012). Теоретическая валидизация измерительных шкал. *Социология: теория, методы, маркетинг*, (2), 53–65.
- Дембицкий, С. (2014а). Относительно проблематичности статистической проверки теоретических конструктов. *Социология: теория, методы, маркетинг*, (2), 190–194.
- Дембицкий, С. (2014b). Принципы теоретической валидизации в социологическом исследовании. *Вісник НТУУ «КПІ» Політологія. Соціологія. Право*, 22(2), 53–58.
- Дембицкий, С. (2015а). *Культурные коды потребления*. Получено из http://soc-research.info/blog/index_files/Klimenkova.html
- Дембицкий, С. (2015b). Введение в проблематику исследований смешанного типа. *Социология: теория, методы, маркетинг*, (2), 143–163.
- Дембицкий, С. (2015с). Разновидности эмпирической социологии: реконструкция исторического опыта. *Вісник НТУУ «КПІ» Політологія. Соціологія. Право*, 27/28(3/4), 58–68.
- Дембицкий, С. (2016а). *Генри Мюррей и концептуальные карты*. Получено из http://soc-research.info/blog/index_files/needs.html
- Дембицкий, С. (2016b). Социологические тесты: сущность и валидизация. *Социология: теория, методы, маркетинг*, (3), 140–155.
- Дембицкий, С. (2016с). *Теоретическая валидизация в социологическом исследовании: Методология и методы*. Москва: ЛЕНАНД.
- Дембицкий, С. (2016d). Валидизация шкалы гражданской активности. *Вісник НТУУ «КПІ» Політологія. Соціологія. Право*, 31/32(3/4), 58–68.

- Дембицкий, С. (2016е). Экспресс-тест SCL-9-NR: методика оценки выраженности психологического дистресса для массовых опросов. *Социология: теория, методы, маркетинг*, (1), 52–64.
- Дембицкий, С. (2017а). *Гистограммы распределений индекса SCL-9-NR*. Получено из http://soc-research.info/blog/index_files/scl-plots.html
- Дембицкий, С. (2017б). *Разработка шкалы антидемократии / тоталитаризма: результаты экспертного опроса*. Получено из http://soc-research.info/blog/index_files/totalit.html
- Дембицкий, С. (2017с). Комплексные измерительные инструменты в социологии. *Социологический журнал*, 23(3), 102–124.
- Дембицкий, С. (2017d). Независимая Украина в геополитическом дискурсе: веги и акценты. *Вісник НТУУ «КПІ» Політологія. Соціологія. Право*, 33/34(1/2), 93–105.
- Дембицкий, С. (2018). Конструирование и валидизация социологического теста SCL-9-NR. *Социологический журнал*, 24(3), 8–31.
- Дембицкий, С., & Серeda, Ю. (2015). Симптоматический опросник Леонарда Дерогатиса (SCL-90-R): валидизация в Украине. *Социология: теория, методы, маркетинг*, (4), 40–71.
- Дембицкий, С., Яценко, В., Сидоров, Н., & Александрова, С. (2016). Критериальная валидизация экспресс-теста SCL-9-NR. *Социология: теория, методы, маркетинг*, (2), 102–116.
- Дембіцький, С. (2017а). Априорна та апостеріорна кластеризація результатів соціологічного тесту “Типи політичної культури”: порівняльний аналіз. *Соціологія: теорія, методи, маркетинг*, (1), 128–144.
- Дембіцький, С. (2017б). Застосування L-шкали (ММРІ) щодо загальної популяції в Україні: результати і гіпотези. *Соціологія: теорія, методи, маркетинг*, (3), 102–118.
- Дембіцький, С. (2017с). Інтегральний індекс соціального самопочуття: альтернативний підхід. *Соціологія: теорія, методи, маркетинг*, (4), 109–129.
- Дембіцький, С. (2018а). Конструювання, валідизація та інтерпретація соціологічного тесту “Типи політичної культури — II”. *Соціологія: теорія, методи, маркетинг*, (1), 25–45.
- Дембіцький, С. С. (2017d). Принципи визначення критеріальних рівнів соціологічного тесту. *Український соціум*, 63(4), 22–34.
- Дембіцький, С. С. (2018б). Вивчення психологічного дистресу в соціологічних дослідженнях: історія та методологія. *Український соціум*, 64(1), 36–48.
- Дембіцький, С. С., & Любива Т. Я. (2017а). Вимоги до розробки та адаптації комплексних вимірювальних інструментів в соціології: надійність, валідність і достовірність. *Український соціум*, 62(3), 45–57.
- Дембіцький, С. С., & Любива, Т. Я. (2014). Геополітичні орієнтації населення України: до і після конфлікту з Росією. *Українське суспільство: моніторинг соціальних змін*, 15(1), 141–146.

- Дембіцький, С. С., & Любива, Т. Я. (2017b). Зв'язок територіально-громадянської ідентичності, геополітичних орієнтацій та міграційних настроїв в Україні. *Українське суспільство: моніторинг соціальних змін*, 18(1), 289–298.
- Дембіцький, С. С., & Хляпатура, Б. М. (2015). Підходи до розробки опитувального інструментарію в дослідженнях змішаного типу. *Український соціум*, 55(4), 37–45.
- Джеймс, Г., Уитгон, Д., Хасті, Т., & Тибширани, Р. (2016). *Введение в статистическое обучение с примерами на языке R*. Москва: ДМК-Пресс.
- Зборовский, Г. Е. (2007). *История социологии*. Москва: Гардарики.
- Зинурова, Р. И., & Фатыхова, Ф. Ф. (2011). Индикаторы, репрезентирующие содержание социального самочувствия. *Вестник Казанского технологического университета*, (20), 245–251.
- Кабаков, Р. (2014). *R в действии. Анализ и визуализация данных в программе R*. Москва: ДМК Пресс.
- Кашкина, Л. (2012). Социальное самочувствие населения монопрофильного города. *Арктика и Север*, (8). Получено из <https://cyberleninka.ru/article/v/sotsialnoe-samochuvstvie-naseleniya-monoprofilnogo-goroda>
- Клайн, П. (1994). *Справочное руководство по конструированию тестов*. Киев: ПАН лтд.
- Клейнман, П. (2015). *Психология: люди, концепции, эксперименты*. Москва: Манн, Иванов и Фербер.
- Козлова, Е. (2012). Исследование социального самочувствия студентов с особыми потребностями. *Концепт*, (6). Получено из <https://e-koncept.ru/2012/12080.htm>
- Колесникова, О. Н., & Шепелева, С. В. (2015). Социальное самочувствие семьи: взаимосвязь социологии и социальной работы. *Социология в современном мире: наука, образование, творчество*, (7), 232–236.
- Коллинз, Р. (2017). *Макроистория: очерки социологии большой длительности*. Москва: Эдиториал УРСС.
- Кэмпбелл, Д. (1980). *Модели экспериментов в социальной психологии и прикладных исследованиях*. Москва: Прогресс.
- Лапин, Н. И. (2004). *Эмпирическая социология в Западной Европе*. Москва: Изд. дом ГУ-ВШЭ.
- Магун, В., & Руднев, М. (2007). Жизненные ценности населения: сравнение Украины с другими европейскими странами. В Е. Головаха, С. Макеев (Ред.), *Украинское общество в европейском пространстве* (с. 226–273). Киев: Ин-т социологии НАН Украины; Харьков: Нац. ун-т им. В. Н. Каразина.
- Майерс, Д. (2003). *Социальная психология*. Санкт-Петербург: Питер.
- Мастичкий, С. Э. (2016). *Визуализация данных с помощью ggplot2*. Москва: ДМК-Пресс.
- Мастичкий, С. Э., & Шитиков, В. К. (2015). *Статистический анализ и визуализация данных с помощью R*. Москва: ДМК-Пресс.
- Мягков, А. Ю. (2002). Шкалы лжи из опросника ММРІ: опыт экспериментальной валидации. *Социологические исследования*, (7), 117–130.

- Мягков, А. Ю. (2015). Диагностика и измерение социального самочувствия студентов: к обоснованию методики. *Социология образования*, (7), 102–113.
- Никифоров, А. (2006). *Философия науки: История и теория*. Москва: Идея-Пресс.
- Панина, Н. В. (2008). *Избранные труды по социологии: В трех томах. — Том I. Вопросы теории, методологии, технологии социологического исследования и профессиональной этики*. Киев: Институт социологии НАН Украины.
- Панина, Н. В. (2012). *Избранные труды по социологии: В трех томах. — Том III. Социология политики, национальных отношений, общественного мнения и социальная психиатрия*. Киев: Институт социологии НАН Украины.
- Паниотто, В. И. (1986). *Качество социологической информации*. Киев: Наукова думка.
- Паніна, Н. (1998). Вимірювання в соціології та принципи застосування тестів у масових опитуваннях. *Соціологія: теорія, методи, маркетинг*, (1-2), 53–68.
- Паніотто, В. І., Максименко В. С., & Харченко, Н. М. (2004). *Статистичний аналіз соціологічних даних*. Київ: Вид. дім «КМ Академія».
- Поппер, К. Р. (2002). *Объективное знание. Эволюционный подход*. Москва: Эдиториал УРСС.
- Рогозин, Д. (2002). *Когнитивный анализ опросного инструмента*. Москва: Ин-т Фонда «Общественное мнение».
- Розов, Н. С. (2009). *Историческая макросоциология: методология и методы*. Новосибирск: НГУ.
- Савельева, И. (2014). *Американская историческая социология об изменчивости и специфичности прошлого*. Получено из <https://postnauka.ru/longreads/37740>
- Садмен, С., & Брэдберн, Н. (2002). *Как правильно задавать вопросы: введение в проектирование опросного инструмента*. Москва: Ин-т Фонда «Общественное мнение».
- Садмен, С., Брэдберн, Н., & Шварц, Н. (2003). *Как люди отвечают на вопросы: применение когнитивного анализа в массовых обследованиях*. Москва: Ин-т Фонда «Общественное мнение».
- Сальнікова, С. (2017). Соціальне самопочуття жителів міста Луцька: про що говорить моніторинг. *Социология: теория, методы, маркетинг*, (2), 105–118.
- Симончук, Е. (2007). Классовые структуры в сравнительной перспективе. В Е. Головаха, С. Макеев (Ред.), *Украинское общество в европейском пространстве* (с. 33–104). Киев: Ин-т социологии НАН Украины; Харьков: Нац. ун-т им. В. Н. Каразина.
- Симончук, Е. (2008). Классификатор профессий ISCO-88: история разработки, концептуальные основания, модель операционализации, применение в социологических исследованиях. *Социология: теория, методы, маркетинг*, (3), 24–41.
- Стегній, О. Г. (2015). Соціологічні підходи до вивчення типів політичної культури. *Український соціум*, 53(2), 26–35.
- Страус, А., & Корбин, Дж. (2001). *Основы качественного исследования. Обоснованная теория. Процедуры и техники*. Москва: Эдиториал УРСС.

- Тарабрина, Н. В., Агарков, В. А., Быховец, Ю. В., Калмыкова, Е. С., Макаrchук, А. В., Падун, М. А., ... & Щепина, А. И. (2007). *Практическое руководство по психологии посттравматического стресса. Часть 1: Теория и методы*. Москва: Изд-во «Издательство Когито-Центр».
- Татарова, Г. Г., & Бессокирная, Г. П. (2016). Типологический анализ наемных работников в контексте их социальной адаптации. *Россия реформирующаяся*, (14), 21–49.
- Тилли, Ч. (2009). *Принуждение, капитал и европейские государства. 990–1992 гг.* Москва: Территория будущего.
- Тичер, С., Мейер, М., Водак, Р., & Веттер, Е. (2009). *Методы анализа текста и дискурса*. Харьков: Гуманитарный центр.
- Уикем, Х., & Гроулмунд, Г. (2018). *Язык R в задачах науки о данных: импорт, подготовка, обработка, визуализация и моделирование данных*. Санкт-Петербург: ООО «Альфа-книга».
- Фатыхова, Ф. Ф. (2010). Социальное самочувствие молодежи этнических групп: региональный аспект. *Вестник Университета Российской Академии Образования*, (5), 124–127.
- Феннето, Э. (2004). *Интервью и опросник: формы, процедуры, результаты*. Санкт-Петербург: Питер.
- Филлипс, Л., & Йоргенсен, М. В. (2008). *Дискурс-анализ. Теория и метод*. Харьков: Гуманитарный центр.
- Хамзяева, О. С. (2016). Социальное самочувствие молодежи в условиях трансформации общества. В *Материалы научно-практической конференции Дыльновские чтения «Социологическая диагностика современного общества»* (с. 200–204). Саратов: Изд-во «Саратовский источник».
- Хилажева, Г. Ф., Леготин, Э. В., Ягафарова, Д. Г., & Исянгулова, Э. И. (2017). Социологический анализ социального самочувствия населения Республики Башкортостан. *Вестник ВЭГУ*, (3), 101–109.
- Хили, Дж. (2005). *Статистика. Социологические и маркетинговые исследования*. Киев: ООО «ДиаСофтЮП»; СПб.: Питер.
- Чугуенко, В. М., & Бобкова, Е. М. (2013). Новые тенденции в исследовании социального самочувствия населения. *Социологические исследования*, (1), 15–23.
- Шмелёв, А. Г. (1996). *Основы психодиагностики*. Ростов-на-Дону: Феникс.
- Щербак, М. (2006). Социально-классовые характеристики политической культуры в Украине: опыт применения социологического теста «Типы политической культуры». *Социология: теория, методы, маркетинг*, (3), 153–160.
- Ядов, В. А. (2011). *Стратегия социологического исследования: описание, объяснение, понимание социальной реальности*. Москва: Издательство «Омега-Л»
- Acevedo, G. A., Ellison, C. G., & Xu, X. (2014). Is it really religion? Comparing the main and stress-buffering effects of religious and secular civic engagement on psychological distress. *Society and Mental Health*, 4(2), 111–128.
- Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American political science review*, 95(3), 529–546.

- Alang, S. M., McAlpine, D. D., & Henning-Smith, C. E. (2014). Disability, health insurance, and psychological distress among US adults: an application of the stress process. *Society and mental health, 4*(3), 164–178.
- Alexopoulos, D. S. (2007). Classical test theory. In *Encyclopedia of measurement and statistics* (pp. 140–143). Thousand Oaks, London, New Delhi: Sage.
- Aneshensel, C. S. (2015). Sociological inquiry into mental health: The legacy of Leonard I. Pearlin. *Journal of health and social behavior, 56*(2), 166–178.
- Aneshensel, C. S., & Avison, W. R. (2015). The stress process: An appreciation of Leonard I. Pearlin. *Society and Mental Health, 5*(2), 67–85.
- Arel, D. (1998). Ukraine: the muddle way. *Current History, 97*, 342–346.
- Arriaga, X. B., Capezza, N. M., Goodfriend, W., Rayl, E. S., & Sands, K. J. (2013). Individual well-being and relationship maintenance at odds: The unexpected perils of maintaining a relationship with an aggressive partner. *Social Psychological and Personality Science, 4*(6), 676–684.
- Baker, L. R., McNulty, J. K., Overall, N. C., Lambert, N. M., & Fincham, F. D. (2013). How do relationship maintenance behaviors affect individual well-being? A contextual perspective. *Social Psychological and Personality Science, 4*(3), 282–289.
- Barkanov, B. (2015). Crisis in Ukraine: Clash of Civilizations or Geopolitics? In *Power, Politics and Confrontation in Eurasia* (pp. 210–239). London: Palgrave Macmillan.
- Bartone, P. T., & Adler, A. B. (1995). *Event-oriented debriefing following military operations: what every leader should know*. Army medical research unit Europe.
- Beck, A. T., Steer, R. A., & Carbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical psychology review, 8*(1), 77–100.
- Bierman, A. (2012). Functional limitations and psychological distress: Marital status as moderator. *Society and mental health, 2*(1), 35–52.
- Biron, M., & Bamberger, P. (2010). The impact of structural empowerment on individual well-being and performance: Taking agent preferences, self-efficacy and operational constraints into account. *Human Relations, 63*(2), 163–191.
- Brown, T. N., Bell, M. L., & Patterson, E. J. (2016). Imprisoned by empathy: familial incarceration and psychological distress among African American men in the National Survey of American Life. *Journal of health and social behavior, 57*(2), 240–256.
- Brzezinski, Z. (1996). Ukraine's Critical Role in the Post-Soviet Space. *Harvard Ukrainian Studies, 20*, 3–8.
- Bufka, L. F., & Camp, N. (2010). Brief measures for screening and measuring mental health outcomes. In M. Antony, D. Barlow (Eds.), *Handbook of Assessment and Treatment Planning for Psychological Disorders* (pp. 62–94). New York, London: The Guilford Press.
- Burgard, S. A., & Seelye, S. (2017). Histories of Perceived Job Insecurity and Psychological Distress among Older US Adults. *Society and mental health, 7*(1), 21–35.
- Cadier, D. (2014). Eastern partnership vs Eurasian Union? The EU–Russia competition in the shared neighbourhood and the Ukraine crisis. *Global Policy, 5*, 76–85.

- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin*, 56(2), 81–105.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New Jersey: Lawrence Erlbaum Associates, Publishers.
- Cohen, S. (2010). Perceived stress in a probability sample of the United States. In S. Spacapan, S. Oskamp (Eds.), *The social psychology of health* (pp. 31–67). Newbury Park, California: Sage.
- Conway, D. (2010). *The Data Science Venn Diagram*. Retrieved from <http://drewconway.com/zia/2013/3/26/the-data-science-venn-diagram>
- Creswell, J. W. (2007). *Qualitative inquiry and research design: Choosing among five approaches*. Thousand Oaks; London; New Delhi: Sage.
- Creswell, J. W., & Tashakkori, A. (2007). Differing perspectives on mixed methods research. *Journal of Mixed Methods Research*, 1(4), 303–308.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin*, 52(4), 281–302.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of consulting psychology*, 24(4), 349–354.
- Daigneault, P. M., & Jacob, S. (2014). Unexpected but most welcome: Mixed methods for the validation and revision of the participatory evaluation measurement instrument. *Journal of Mixed Methods Research*, 8(1), 6–24.
- Dawson, J. I. (1997). Ethnicity, ideology and geopolitics in Crimea. *Communist and Post-Communist Studies*, 30(4), 427–444.
- Delcour, L., & Wolczuk, K. (2015). Spoiler or facilitator of democratization?: Russia's role in Georgia and Ukraine. *Democratization*, 22(3), 459–478.
- Dembitskyi, S. S. (2015). Geopolitical orientation of Ukraine's population: before and during the military conflict with Russia. *Українське суспільство: моніторинг соціальних змін*, 16(1), 98–103.
- Dembitskyi, S. S. (2016). Level of psychological distress in different groups of Ukraine's population. *Українське суспільство: моніторинг соціальних змін*, 17(1), 98–103.
- Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioral science*, 19(1), 1–15.
- Derogatis, L., & Fitzpatrick, M. (2004). SCL-90-R, the Brief Symptom Inventory (BSI), and the BSI-18. In M.E. Maruish (Ed.), *The use of psychological testing for treatment planning and outcomes assessment. Vol. 3: Instruments for adults* (pp.1–42). New Jersey, London: Lawrence Erlbaum Associated Publishers.
- DeVellis, R. (2012). *Scale Development: Theory and applications*. Los Angeles, London, New Delhi, Singapore, Washington DC: Sage.
- Dragneva, R., & Dimitrova, A. (2007). Patterns of Integration and Regime Compatibility: Ukraine between the CIS and the EU. In *The CIS, the EU and Russia* (pp. 171–201). London: Palgrave Macmillan.
- European Social Survey. (2015). *Measuring and Reporting on Europeans' Wellbeing: Findings from the European Social Survey*. London: ESS ERIC.

- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Los Angeles, London, New Delhi, Singapore, Washington DC: Sage.
- Fingerhut, A. W., & Maisel, N. C. (2010). Relationship formalization and individual and relationship well-being among same-sex couples. *Journal of Social and Personal Relationships*, 27(7), 956–969.
- Fothergill, K., Ensminger, M. E., Doherty, E. E., Juon, H. S., & Green, K. M. (2016). Pathways from early childhood adversity to later adult drug use and psychological distress: a prospective study of a cohort of African Americans. *Journal of health and social behavior*, 57(2), 223–239.
- Gallagher, M. (2017). Obligatory and Voluntary Identity Discrepancies, Self-Evaluation, and Psychological Distress. *Society and Mental Health*, 7(2), 51–68.
- Gentile, M. (2015). West oriented in the East-oriented Donbas: a political stratigraphy of geopolitical identity in Luhansk, Ukraine. *Post-Soviet Affairs*, 31(3), 201–223.
- Glaser, B. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine.
- Glavin, P., Schieman, S., & Reid, S. (2011). Boundary-spanning work demands and their consequences for guilt and psychological distress. *Journal of Health and Social Behavior*, 52(1), 43–57.
- Goldberg, D. P., & Hillier, V. F. (1979). A scaled version of the General Health Questionnaire. *Psychological medicine*, 9(1), 139–145.
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational evaluation and policy analysis*, 11(3), 255–274.
- Grolemund, G. (2014). *Hands-On Programming with R: Write Your Own Functions and Simulations*. Sebastopol: O'Reilly Media, Inc.
- Groth-Marnat, G. (2003). *Handbook of psychological assessment*. New Jersey: Wiley.
- Guest, G., MacQueen, K. M., & Namey, E. E. (2012). *Applied thematic analysis*. Thousand Oaks, California: Sage.
- Hansen, T., Moum, T., & Shapiro, A. (2007). Relational and individual well-being among cohabitators and married individuals in midlife: Recent trends from Norway. *Journal of Family Issues*, 28(7), 910–933.
- Hatton, D. (2010). Did the Orange Revolution change Ukraine's geopolitical position regarding Russia and the west? *Polis Journal*, 3, 1–26. Retrieved from https://www.europeansocialsurvey.org/docs/round3/questionnaire/ESS3_huppert_proposal.pdf
- Huppert, F. A., & Whittington, J. E. (2003). Evidence for the independence of positive and negative well-being: Implications for quality of life assessment. *British journal of health psychology*, 8(1), 107–122.
- Huppert, F. A., Andrew, C., Frey, B., Marks, N., & Siegrist, J. (2005). *Personal and social well-being: Creating indicators for a flourishing Europe*. Retrieved from <http://essedunet.nsd.uib.no/cms/topics/wellbeing/>

- Ignatyev, Y., Assimov, M., Aichberger, M. C., Ivens, S., Mir, J., Dochshanov, D., ... & Mundt, A. P. (2012). Psychometric properties of a Russian version of the General Health Questionnaire-28. *Psychopathology*, 45(4), 252–258.
- Jacobs, A. W., Hill, T. D., & Burdette, A. M. (2015). Health insurance status and symptoms of psychological distress among low-income urban women. *Society and Mental Health*, 5(1), 1–15.
- Jeffrey, K., Abdallah, S., & Quick, A. (2015). *Europeans' personal and social wellbeing: topline results from Round 6 of the European Social Survey*. London: Centre for Comparative Social Surveys.
- Johnson, B., & Turner, L. A. (2003). Data collection strategies in mixed methods research. In A. Tashakkori (Ed.), *Handbook of mixed methods in social and behavioral research* (pp. 297–319). London, Thousand Oaks, New Delhi: Sage.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher*, 33(7), 14–26.
- Jones, L., & Thissen, D. (2007). History and Overview of Psychometrics. In C. Rao, S. Sinharay (Eds.), *Handbook of Statistics, Volume 26: Psychometrics* (pp.1–27). Amsterdam, Oxford: Elsevier.
- Juni, S. (2007). Reliability Theory. In *Encyclopedia of measurement and statistics* (pp. 834–835). Thousand Oaks, London, New Delhi: Sage.
- Kahn, J. H., Hucke, B. E., Bradley, A. M., Glinski, A. J., & Malak, B. L. (2012). The Distress Disclosure Index: A research review and multitrait–multimethod examination. *Journal of Counseling Psychology*, 59(1), 134.
- Kane, M., & Trochim, W. M. (2007). *Concept mapping for planning and evaluation*. Thousand Oaks; London; New Delhi: Sage.
- Kaplan, R. M., & Saccuzzo, D. P. (2009). *Psychological testing: Principles, applications, and issues*. Belmont: Wadsworth.
- Keller, L. (2007). Item Response Theory. In *Encyclopedia of measurement and statistics* (pp. 493–496). Thousand Oaks, London, New Delhi: Sage.
- Kessler, R. C., & Üstün, T. B. (Eds.). (2008). *The WHO World Mental Health Surveys: global perspectives on the epidemiology of mental disorders* (Vol. 6). New York: Cambridge University Press.
- Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S. L., ... & Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological medicine*, 32(6), 959–976.
- Klaghofer, R., & Brähler, E. (2001). Konstruktion und Teststatistische Prüfung einer Kurzform der SCL-90–R. *Zeitschrift für Klinische Psychologie, Psychiatrie und Psychotherapie*, 49(2), 115–124.
- Kropatcheva, E. (2011). Playing both ends against the middle: Russia's geopolitical energy games with the EU and Ukraine. *Geopolitics*, 16(3), 553–573.
- Kuzio, T. (2000a). Geopolitical pluralism in the CIS: The emergence of GUUAM. *European Security*, 9(2), 81–114.
- Kuzio, T. (2000b). Promoting geopolitical pluralism in the CIS: GUUAM and Western foreign policy. *Problems of Post-Communism*, 47(3), 25–35.

- Kuzio, T. (2001). Identity and nation-building in Ukraine: Defining the 'Other'. *Ethnicities*, 1(3), 343–365.
- Kuzio, T. (2003). Ukraine's relations with the west: disinterest, partnership, disillusionment. *European Security*, 12(2), 21–44.
- Kuzio, T. (2006). National identity and history writing in Ukraine. *Nationalities Papers*, 34(4), 407–427.
- Larrabee, F. S. (2010). Russia, Ukraine, and Central Europe: the return of geopolitics. *Journal of international affairs*, 63(2), 33–52.
- Larsen, M. (2007). Latent Class Analysis. In *Encyclopedia of measurement and statistics* (pp. 526–528). Thousand Oaks, London, New Delhi: Sage.
- Lee, H. (2007). k-Means Cluster Analysis. In *Encyclopedia of measurement and statistics* (pp. 503–507). Thousand Oaks, London, New Delhi: Sage.
- Leech, N. L., & Onwuegbuzie, A. J. (2007). An array of qualitative data analysis tools: a call for data analysis triangulation. *School psychology quarterly*, 22(4), 557.
- Leech, N. L., & Onwuegbuzie, A. J. (2008). Qualitative data analysis: A compendium of techniques and a framework for selection for school psychology research and beyond. *School Psychology Quarterly*, 23(4), 587.
- Luyt, R. (2012). A framework for mixing methods in quantitative measurement development, validation, and revision: A case study. *Journal of Mixed Methods Research*, 6(4), 294–316.
- Malek, M. (2009). The «Western Vector» of the Foreign and Security Policy of Ukraine: Continuities and Ruptures under President Viktor Yushchenko (2005–2009). *The Journal of Slavic Military Studies*, 22(4), 515–542.
- McLeod, J. D. (2012). The meanings of stress: Expanding the stress process model. *Society and Mental Health*, 2(3), 172–186.
- McNabb, D. E. (2014). *Nonsampling error in social surveys*. Los Angeles, London, New Delhi, Singapore, Washington DC : Sage.
- Miller, S. (2011). *Social Institutions*. Retrieved from <https://plato.stanford.edu/entries/social-institutions/>
- Miller, S. V. (2013). Territorial disputes and the politics of individual well-being. *Journal of Peace Research*, 50(6), 677–690.
- Molchanov, M. A. (2000). Post-communist nationalism as a power resource: A Russia-Ukraine comparison. *Nationalities Papers*, 28(2), 263–288.
- Montazer, S., & Wheaton, B. (2017). Economic conditions in countries of origin and trajectories in distress after migration to Canada: results from the national population health survey. *Society and Mental Health*, 7(1), 1–20.
- Müllerson, R. (2014). Ukraine: victim of geopolitics. *Chinese Journal of International Law*, 13(1), 133–145.
- National Comorbidity Survey (2018). *K10 and K6 Scales*. Retrieved from http://www.hcp.med.harvard.edu/ncs/k6_scales.php
- Nelson, L. (2015). *Dr. Cohen's Scales*. Retrieved from <http://www.psy.cmu.edu/~scohen/scales.html>

- Newman, I., Lim, J., & Pineda, F. (2013). Content validity using a mixed methods approach: Its application and development through the use of a table of specifications methodology. *Journal of Mixed Methods Research*, 7(3), 243–260.
- Norwegian Centre for Research Data. (2012). *ESS Round 6: European Social Survey Round 6 Data (2.3)* [Data set]. Retrieved from <https://www.europeansocialsurvey.org/download.html?file=ESS6UA&c=UA&y=2012>.
- Nichol, J. P., Woehrel, S. J., & Gelb, B. A. (2006). *Russia's cutoff of natural gas to Ukraine: context and implications*. Congressional Research Service, Library of Congress.
- Nikulin, V., & Selyutina, N. (2007). Geopolitical orientations of Ukrainian political elites and the electoral campaign of 2006. *Perspectives on European Politics and Society*, 8(4), 544–558.
- O'Loughlin, J., Toal, G., & Kolosov, V. (2017). The rise and fall of «Novorossiya»: examining support for a separatist geopolitical imaginary in southeast Ukraine. *Post-Soviet Affairs*, 33(2), 124–144.
- Oates, G. L., & Goode, J. (2013). Racial differences in effects of religiosity and mastery on psychological distress: Evidence from national longitudinal data. *Society and mental health*, 3(1), 40–58.
- Onwuegbuzie, A. J. (2003). Effect sizes in qualitative research: A prolegomenon. *Quality and Quantity*, 37(4), 393–409.
- Onwuegbuzie, A. J., Bustamante, R. M., & Nelson, J. A. (2010). Mixed research as a tool for developing quantitative instruments. *Journal of Mixed Methods Research*, 4(1), 56–78.
- Onwuegbuzie, A. J., Leech, N. L., & Collins, K. M. (2008). Interviewing the interpretive researcher: A method for addressing the crises of representation, legitimation, and praxis. *International Journal of Qualitative Methods*, 7(4), 1–17.
- Parker, S. (2007). Scaling. In *Encyclopedia of measurement and statistics* (pp. 582–584). Thousand Oaks, London, New Delhi: Sage.
- Pearlin, L. I. (1989). The sociological study of stress. *Journal of health and social behavior*, 30(3), 241–256.
- Pearlin, L. I. (2009). The life course and the stress process: Some conceptual comparisons. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 65(2), 207–215.
- Pearlin, L. I., & Schooler, C. (1978). The structure of coping. *Journal of health and social behavior*, 19(1), 2–21.
- Pearlin, L. I., & Skaff, M. M. (1996). Stress and the life course: A paradigmatic alliance. *The Gerontologist*, 36(2), 239–247.
- Pearlin, L. I., Menaghan, E. G., Lieberman, M. A., & Mullan, J. T. (1981). The stress process. *Journal of Health and Social behavior*, 22(4), 337–356.
- Pearlin, L. I., Schieman, S., Fazio, E. M., & Meersman, S. C. (2005). Stress, health, and the life course: Some conceptual perspectives. *Journal of Health and Social Behavior*, 46(2), 205–219.
- Radloff, L. S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied psychological measurement*, 1(3), 385–401.

- Rhoades, G. K., Stanley, S. M., & Markman, H. J. (2009). Couples' reasons for cohabitation: Associations with individual well-being and relationship quality. *Journal of Family Issues, 30*(2), 233–258.
- Ridner, S. H. (2004). Psychological distress: concept analysis. *Journal of advanced nursing, 45*(5), 536–545.
- Ros, L., Latorre, J. M., Aguilar, M. J., Serrano, J. P., Navarro, B., & Ricarte, J. J. (2011). Factor structure and psychometric properties of the center for epidemiologic studies depression scale (CES-D) in older populations with and without cognitive impairment. *The International Journal of Aging and Human Development, 72*(2), 83–110.
- Salkind, N. J. (2007). *Encyclopedia of measurement and statistics*. Thousand Oaks, London, New Delhi: Sage.
- Sandıkcı, Ö., Peterson, M., Ekici, A., & Simkins, T. (2016). Development and quality of life in turkey: how globalization, religion, and economic growth influence individual well-being. *Journal of Macromarketing, 36*(3), 304–320.
- Sereda, Y., & Dembitskyi, S. (2016). Validity assessment of the symptom checklist SCL-90-R and shortened versions for the general population in Ukraine. *BMC Psychiatry, 16*. Retrieved from <https://bmcp psychiatry.biomedcentral.com/articles/10.1186/s12888-016-1014-3>
- Singh, R. P., & Arora, A. P. (2010). Antecedents of Individual Well-Being. *Vision, 14*(3), 191–205.
- Song, L. (2011). Social capital and psychological distress. *Journal of Health and Social Behavior, 52*(4), 478–492.
- Sorokin, P. A. (1961). *Fads and foibles in modern sociology and related sciences*. Chicago: Henry Regnery.
- Spradley, J. P. (1979). *The ethnographic interview*. Fortworth: Holt, Rinehart, and Winston.
- Stafford, L., Kline, S. L., & Rankin, C. T. (2004). Married individuals, cohabiters, and cohabiters who marry: A longitudinal study of relational and individual well-being. *Journal of Social and Personal Relationships, 21*(2), 231–248.
- Stemler, S. (2007). Measurement. In *Encyclopedia of measurement and statistics* (pp. 862–865). Thousand Oaks, London, New Delhi: Sage.
- Stepanenko, V. (2003). Identities and language politics in Ukraine: the challenges of nation-state building. In *Nation-building, ethnicity and language politics in transition countries* (pp. 107–135). Budapest: Open Society Institute.
- Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, California: Sage.
- Teddlie, C., & Tashakkori, A. (2009). *Foundations of mixed methods research: Integrating quantitative and qualitative approaches in the social and behavioral sciences*. Thousand Oaks, California: Sage.
- Thompson, S., & Marks, N. (2008). *Measuring well-being in policy: issues and applications*. Retrieved from https://b.3cdn.net/nefoundation/575659b4f333001669_ohm6iiogp.pdf

- Torres, J. M., Alcántara, C., Rudolph, K. E., & Viruell-Fuentes, E. A. (2016). Cross-border ties as sources of risk and resilience: do cross-border ties moderate the relationship between migration-related stress and psychological distress for latino migrants in the United States? *Journal of health and social behavior*, 57(4), 436–452.
- Transcultural Mental Health Center. (2018). *Kessler 10*. Retrieved from <http://www.dhi.health.nsw.gov.au/transcultural-mental-health-centre/resources/multilingual-resources-by-title/kessler-10?PagingModule=883&retain=true&pg=2>
- Tucker, C. J., McHale, S. M., & Crouter, A. C. (2003). Conflict resolution: Links with adolescents' family relationships and individual well-being. *Journal of Family Issues*, 24(6), 715–736.
- Turner, J. (1997). *The Institutional Order*. New York: Longman.
- Vaurio, R. (2011). Symptom Checklist-90-Revised. In *Encyclopedia of Clinical Neuropsychology* (pp. 2447–2449). New York, Dordrecht, Heidelberg, London: Springer.
- Veit, C. T., & Ware, J. E. (1983). The structure of psychological distress and well-being in general populations. *Journal of consulting and clinical psychology*, 51(5), 730.
- Virág, A. (2012). The cultural and geopolitical dimensions of nation-building in the Ukraine. *Society and Economy*, 34(4), 619–641.
- Walls, M. L., & Whitbeck, L. B. (2011). Distress among Indigenous North Americans: Generalized and culturally relevant stressors. *Society and Mental Health*, 1(2), 124–136.
- Whitton, S. W., Rhoades, G. K., & Whisman, M. A. (2014). Fluctuation in relationship quality over time and individual well-being: Main, mediated, and moderated effects. *Personality and Social Psychology Bulletin*, 40(7), 858–871.
- Woehrel, S. (2005). *Ukraine's Orange Revolution and US Policy*. Congressional Research Service, Library of Congress.
- Yin, R. K. (2009). *Case study research: Design and methods*. London; Thousand Oaks; New Delhi: Sage.
- Young, W. B., & Foy, S. L. (2013). The influence of meaning-making after spousal loss on trajectories of psychological distress. *Society and Mental Health*, 3(3), 187–202.
- Zhurzhenko, T. (2002). Language Politics in Contemporary Ukraine: Nationalism and Identity Formation. *Questionable Returns*, 12, 12.
- Zhurzhenko, T. (2004). Cross-border cooperation and transformation of regional identities in the Ukrainian–Russian borderlands: towards a Euroregion «Slobozhanshchyna»? Part 1. *Nationalities Papers*, 32(1), 207–232.
- Zhurzhenko, T. (2004). Cross-border cooperation and transformation of regional identities in the Ukrainian–Russian borderlands: towards a Euroregion «Slobozhanshchyna»? Part 2. *Nationalities Papers*, 32(2), 497–514.

Наукове видання

ДЕМБІЦЬКИЙ Сергій Сергійович

**РОЗРОБКА СОЦІОЛОГІЧНИХ ТЕСТІВ:
МЕТОДОЛОГІЯ І ПРАКТИКИ ЇЇ ЗАСТОСУВАННЯ**

Монографія

Наукове редагування — *Любов Бакланова*

Коректура — *Любов Тютюнник, Наталія Мельник*

Дизайн обкладинки — *Михайло Дембіцький*

Макетування — *Михайло Роїк*

Підписано до друку 16.01.2019. Формат 70х100/16. Папір офсетний.
Друк офсетний. Гарнітура Newton. Ум. друк. арк. 23,6. Замов. №1601 Наклад 300.

Надруковано ТОВ "РПФ"Азбука".
04080, Київ-80, вул. Кирилівська 19-21.